Analysis I
Analysis II
Analysis III

von Peter Pfaffelhuber

Version: 20. Januar 2014

Inhaltsverzeichnis

Teil I
Grundlagen

In diesem Abschnitt werden wir das wichtigste Rüstzeug im Umgang mit reellwertigen Funktionen kennen lernen. Nach einer Einführung in Abschnitt 1 über Schreibweisen, Logik und Mengenlehre, werden wir in Abschnitt 2 die reellen Zahlen konstruieren. Schließlich folgt in Abschnitten 3 und 4 der Umgang mit Folgen und Reihen.

1 Grundbegriffe

Aussagenlogik und Mengenlehre kommen überall in der Mathematik vor. Oftmals bemerkt man nicht einmal, dass man gerade Resultate aus diesen Gebieten verwendet. Dies ist Grund genug, am Anfang dieser Vorlesung auf diese beiden Gebiete besonders einzugehen.

1.1 Aussagen und Mengen

Grundlegende Strukturen in der Mathematik sind Mengen und Funktionen. Wir werden hier nur einiges Handwerkszeug an Notation einführen, jedoch nicht auf Feinheiten der modernen Mengenlehre eingehen. Bemerkt sei jedoch, dass es keinesfalls eine einfache Aufgabe ist, genau zu definieren was eine Menge ist.

Definition 1.1 (Aussage und Aussagevariablen). *Eine* Aussage *(oder* Aussageform*) ist ein sprachlicher Ausdruck A, welcher entweder* wahr *oder* falsch *ist. Die verwendete Sprache kann dabei Umgangssprache, oder auch mathematische Formelsprache beinhalten. Eine Aussage A kann* Aussagevariable *(oder auch einfach* Variable*) $x, y, ...$ enthalten. In diesem Fall hängt die Gültigkeit von A (d.h. ob A wahr oder falsch ist) vom Wert der Aussagevariablen ab. Wir schreiben dann $A = A(x, y, ...)$.*

Beispiel 1.2. Beispiele für Aussagen sind etwa:

- $2 \cdot 3 = 6$ (wahr)
- Jede Zahl ist durch 7 teilbar (falsch)
- 3 teilt x (Gültigkeit hängt vom Wert von x ab; ist z.B. wahr wenn $x = 12$)

Definition 1.3 (Aussageoperationen). *Für Aussagen A, B definieren wir folgende weitere Aussagen:*

- *$\neg A$:* Negation *von A. Sie ist genau dann wahr wenn A falsch ist. (In Worten:* Nicht-*A.)*
- *$A \wedge B$:* Konjunktion *von A und B: Sie ist genau dann wahr, wenn sowohl A als auch B wahr sind. (In Worten: A und B.)*
- *$A \vee B$:* Disjunktion *von A und B: Sie ist genau dann wahr, wenn A oder B (oder beide) wahr sind. (In Worten: A oder B)*
- *$A \Rightarrow B$:* Subjunktion *von A und B. Sie ist genau dann wahr, wenn B aus A folgt, d.h. wenn B immer dann wahr ist, wenn A wahr ist. (In Worten: Aus A folgt B. Man beachte, dass hier meistens A und B von Variablen $x, y, ...$ abhängen. $A \Rightarrow B$ bedeutet dann, dass B für alle Werte von $x, y, ...$ gilt, falls dies für A der Fall ist.)*

- $A \iff B$: Bisubjunktion *von A und B. Dies ist gleichbedeutend mit $(A \Rightarrow B) \wedge (B \Rightarrow A)$. (In Worten: A genau dann, wenn B. Manchmal auch kürzer: A gdw B oder A ist äquivalent zu B.)*

Bemerkung 1.4 (Klammerung). Seien A, B, C Aussagen. Die Negation $\neg A$ gilt als einstellige Aussageoperation, die Konjunktion, Subjunktion und Bisubjunktion als zweistellige. Verknüpfungen dieser Operationen erfordern es oftmals, eine Reihenfolge festzulegen, die üblicherweise durch Klammerung erfolgt. Etwa ist $(A \wedge B) \vee C$ nicht dasselbe wie $A \wedge (B \vee C)$. Wir vereinbaren, dass einstellige Operationen immer zuerst ausgeführt werden, und erst anschließend zweistellige, auch wenn dies nicht durch eine Klammer gekennzeichnet wird. Das bedeutet etwa, dass $\neg A \wedge B$ dasselbe ist wie $(\neg A) \wedge B$.

Lemma 1.5[1] **(deMorgan'sche Regeln für Aussagen).** *Seien A, B Aussagen. Dann gilt*

$$\neg(A \wedge B) \iff (\neg A \vee \neg B)$$
$$\neg(A \vee B) \iff (\neg A \wedge \neg B).$$

Beweis. Zunächst zur ersten Aussage. Nach Definition ist $A \wedge B$ genau dann wahr, wenn sowohl A als auch B wahr sind. Damit ist also $\neg(A \wedge B)$ genau dann wahr, wenn entweder A oder B falsch ist. Dies ist genau dann der Fall, wenn $\neg A \vee \neg B$ wahr ist. Bei der zweiten Aussage gilt $A \vee B$ genau dann, wenn entweder A oder B wahr sind. Damit ist $\neg(A \vee B)$ genau dann wahr, wenn sowohl A als auch B falsch sind. Dies ist jedoch gleichbedeutend mit $\neg A \wedge \neg B$. □[2]

Die Rechenregeln des letzten Lemmas lassen sich ergänzen. Es folgen einige Beispiele.

Beispiel 1.6. Seien A, B und C Aussagen (die von Variablen Variablen $x, y, ...$ abhängen).

1. $A \iff \neg\neg A$ ist immer wahr.
2. $A \vee \neg A$ ist immer wahr. (Sie heißt auch die Regel vom ausgeschlossenem Dritten.)
3. $A \wedge \neg A$ ist immer falsch. (Damit ist $\neg(A \wedge \neg A)$ immer wahr, was nach Lemma 1.5 gleichbedeutend ist mit $\neg A \vee A$.)
4. $(A \wedge B) \Rightarrow A$ ist immer wahr.
5. $A \Rightarrow (A \vee B)$ ist immer wahr.
6. Es gilt $(A \wedge B) \wedge C \iff A \wedge (B \wedge C)$ und $(A \vee B) \vee C \iff A \vee (B \vee C)$. Man spricht auch von der *Assoziativität* von $\wedge$ und $\vee$.
7. Es gilt $((A \wedge B) \vee C) \iff (A \vee C) \wedge (B \vee C)$ und $(A \vee B) \wedge C) \iff ((A \wedge C) \vee (B \wedge C)$. Man spricht hier von der *Distributivität* von $\wedge$ und $\vee$.

---

[1] Mit *Lemma* bezeichnen wir einen Hilfssatz. Dieser wird im weiteren Verlauf gebraucht, stellt aber für sich alleine noch keine allzu große Aussage dar. Weiter werden wir mit *Proposition* und *Theorem* wichtigere Aussagen bezeichnen, und mit *Korollar* einfache Folgerungen aus schon Gezeigtem.

[2] Wir bezeichnen das Ende eines Beweises durchgehend mit □.

8. Es gilt $(A \wedge B) \iff (B \wedge A)$ und $(A \vee B) \iff (B \vee A)$. Man spricht von der *Kommutativität* von $\wedge$ und $\vee$.

9. Es gilt $(A \Rightarrow B) \iff (B \vee \neg A)$. Denn $A \Rightarrow B$ bedeutet, dass es nicht sein kann, dass B nicht gilt, A aber schon, also $\neg(\neg B \wedge A)$. Wendet man nun Lemma 1.5 an, erhält man die Behauptung.

10. Es gilt $((A \Rightarrow B) \wedge (B \Rightarrow C)) \Rightarrow (A \Rightarrow C)$. Man spricht von der Transitivität von $\Rightarrow$. Dies lässt sich mit dem letzten Beispiel auch einfach beweisen, nämlich so:

$$\begin{aligned}
(A \Rightarrow B) \wedge (B \Rightarrow C) &\overset{9.}{\iff} (B \vee \neg A) \wedge (C \vee \neg B) \\
&\overset{6.}{\iff} (B \wedge C) \vee (B \wedge \neg B) \vee (\neg A \wedge C) \vee (\neg A \wedge \neg B) \\
&\Longrightarrow C \vee (\neg A \wedge C) \vee \neg A \\
&\Longrightarrow \neg A \vee C \\
&\iff A \Rightarrow C.
\end{aligned}$$

Wir führen nun den Begriff der Menge ein. Allerdings wollen wir nicht sagen, was der – in der Definition wichtige – Begriff der *sinnvollen* Aussage genau ist. Mit anderen Worten führen wir hier nur einen naiven Mengenbegriff ein. Will man Mengen axiomatisch genau definieren, führt das auf das Zermelo-Fraenkel'sche Axiomensystem, auf das wir hier nicht eingehen. In der Praxis sind die allermeisten Aussagen *sinnvoll*, so dass wir keine Schwierigkeiten mit dem naiven Mengenbegriff erwarten.

Definition 1.7 (Menge). *Sei $A = A(x)$ eine "sinnvolle" Aussage, die von einer Variablen x abhängt. Dann bezeichnen wir mit*[3]

$$M := \{x : A(x) \text{ ist wahr}\}$$

die Menge *aller x, für die $A(x)$ gilt. Noch kürzer setzen wir*

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

Wir schreiben $x \in M$ genau dann, wenn $A(x)$ wahr ist. Es bedeutet $x \notin M$ dasselbe wie $\neg(x \in M)$ oder auch, dass $A(x)$ falsch ist.

Beispiel 1.8.
- Die Menge $\emptyset := \{x : x \neq x\}$ ist die leere Menge.
- Die natürlichen Zahlen $\mathbb{N}$ kann man nach Peano axiomatisch etwa so definieren:
  - 1 ist eine natürliche Zahl.
  - Jede natürlich Zahl n hat genau einen Nachfolger, genannt $n + 1$.
  - Die Eins hat keinen Vorgänger.
  - Enthält eine Menge M die Eins und mit jeder natürlichen Zahl n auch deren Nachfolger, so ist M bereits die Menge der natürlichen Zahlen.

[3]Zur Notation: wir unterscheiden in der Mathematik das Gleichheitszeichen "=" vom Zeichen ":=", das das Symbol auf der ":"-Seite (das sogenannte *Definiendum*) durch das Objekt auf der anderen Seite (das *Definiens*) definiert. Sind nämlich x für die $A(x)$ wahr ist bereits bekannt, definieren wir hiermit die Menge M aller dieser Elemente.

- Sei $B(x) : (x = 0) \vee (x \in \mathbb{N}) \vee (-x \in \mathbb{N})$. Dann ist

$$\mathbb{Z} := \{x : B(x)\} = \{0, \pm 1, \pm 2, ...\}$$

die Menge der ganzen Zahlen.

Definition 1.9 (Schnittmenge, Vereinigungsmenge, Komplement, Potenzmenge). *Seien $M = \{x : A(x)\}$ und $N = \{x : B(x)\}$ Mengen. Dann bezeichnet*

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

die Schnittmenge *von M und N,*

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

die Vereinigungsmenge *von M und N sowie*

$$M^c := \{x : \neg A(x)\}$$

das Komplement *von M. Weiter ist $M \setminus N := M \cap N^c$ die* Differenzmenge *von M und N. Die Mengen M und N heißen* disjunkt, *falls $M \cap N = \emptyset$. Wir schreiben $M \subseteq N$ wenn $A(x) \Rightarrow B(x)$ sowie $M = N$ wenn $A(x) \iff B(x)$ gilt. Weiter definieren wir die* Potenzmenge *von M durch*

$$2^M := \{E : E \subseteq M\}.$$

Lemma 1.10 (Distributivgesetze und deMorgan'sche Regeln). *Seien M, N, O Mengen. Dann gilt*

$$\begin{aligned}
&((M \subseteq N) \wedge (N \subseteq O)) \Rightarrow (M \subseteq O), \\
&(M \cap N) \cup O = (M \cup O) \cap (N \cup O), \\
&(M \cup N) \cap O = (M \cap O) \cup (N \cap O), \\
&(M \cap N)^c = M^c \cup N^c, \\
&(M \cup N)^c = M^c \cap N^c.
\end{aligned}$$

Beweis. Alles folgt analog zu den entsprechenden Regeln über Aussagen. Die erste Zeile entsteht so aus Beispiel 1.6.8, die zweite und dritte aus Beispiel 1.6.6 und die letzten beiden aus Lemma 1.5. □

1.2 Quantoren

In der mathematischen Schreibweise von Aussagen sind Quantoren besonders wichtig. Diese legen fest, ob eine Aussage $A(x, ...)$ etwa für alle x oder nur für manche x gelten soll. Sie werden uns im Laufe der Vorlesung sehr oft begegnen.

Definition 1.11 (Quantoren). *Sei $A(x, ...)$ eine Aussage die von Variablen $x, ...$ abhängt.*

- *Dann ist $\forall x \; A(x, ...)$ die Aussage: Für alle möglichen x gilt $A(x, ...)$.*
- *Weiter ist $\exists x \; A(x, ...)$ die Aussage: Es gibt ein x, so dass $A(x, ...)$ gilt*
- *Außerdem ist $\exists! \, x \; A(x, ...)$ die Aussage: Es gibt genau ein x, so dass $A(x, ...)$ gilt.*

Bemerkung 1.12. Typischerweise werden als Variablen $x, ...$ nur Elemente einer Menge zugelassen. Dann schreiben wir etwa $\forall x \in M : A(x, ...)$.

Beispiel 1.13.

- Sei $C(x) : \exists m \in \mathbb{Z}\, \exists n \in \mathbb{N} : x = m/n$. Dann ist
$$\mathbb{Q} := \{x : C(x)\} = \{x = m/n : m \in \mathbb{Z}, n \in \mathbb{N}\}$$
die Menge der rationalen Zahlen.
- Die Aussage $\forall x \in \mathbb{N}\, \exists y \in \mathbb{Z} : x = y$ ist wahr.
- Die Aussage $\forall x \in \mathbb{Q}\, \exists y \in \mathbb{Q} : y = x^2$ ist wahr.
- Die Aussage $\forall x \in \mathbb{Q}\, \exists y \in \mathbb{Q} : y^2 = x$ ist falsch. (Zum Beweis siehe Proposition 1.36)

Beispiel 1.14 (Natürliche Zahlen). Wir greifen noch einmal die Definition der natürlichen Zahlen aus Beispiel 1.8 auf. Mit Quantoren schreibt sich die Definition der natürlichen Zahlen so:

- $1 \in \mathbb{N}$ (Eins ist eine natürliche Zahl.)
- $\forall n \in \mathbb{N}\, \exists!\, m \in \mathbb{N} : n + 1 = m$ (Jede natürlich Zahl hat einen Nachfolger.)
- $\forall m, n \in \mathbb{N} : m + 1 = n + 1 \iff m = n$ (Nachfolger sind eindeutig.)
- $\neg(\exists n \in \mathbb{N} : n + 1 = 1)$ (Die Eins hat keinen Vorgänger.)
- $\forall M : ((M \subseteq \mathbb{N}) \wedge (1 \in M) \wedge (n \in M \Rightarrow n + 1 \in M)) \Rightarrow M = \mathbb{N}$ (Enthält eine Menge M die Eins und mit jeder natürlichen Zahl n auch deren Nachfolger, so ist M bereits die Menge der natürlichen Zahlen.)

Lemma 1.15 (Gesetze für Quantoren). *Seien $A(x, y, ...)$ und $B(x, y, ...)$ Aussagen, die von Variablen $x, y, ...$ abhängen. Dann gilt*

$$\begin{aligned}
\forall x \forall y A(x, y, ...) &\iff \forall y \forall x A(x, y, ...),\\
\exists x \exists y A(x, y, ...) &\iff \exists y \exists x A(x, y, ...),\\
\neg(\forall x A(x, ...)) &\iff \exists x(\neg A(x, ...)),\\
\neg(\exists x A(x, ...)) &\iff \forall x(\neg A(x, ...)).
\end{aligned}$$

Beweis. Die ersten beiden Aussagen sind klar. Für die dritte stellen wir fest, dass $\neg(\forall x A(x, ...))$ genau dann gilt, wenn $A(x, ...)$ nicht für alle x gilt, es also mindestens ein x gibt, für das $A(x, ...)$ falsch ist. Das bedeutet, dass $\exists x(\neg A(x, ...))$ gilt. Die letzte Aussage folgt aus der vorletzten, wenn man dort anstelle von $A(x, ...)$ die Aussage $\neg A(x, ...)$ einsetzt. □

Bemerkung 1.16.

- Wir schreiben $\forall x, y A(x, y, ...)$ anstelle von $\forall x \forall y A(x, y, ...)$. Die erste Zeile des Lemmas zeigt, dass hiermit keine Verwechslungen zu befürchten sind.
- Man kann sicher noch weitere Regeln für das Rechnen mit Quantoren aufstellen. Es ist jedoch etwas Vorsicht geboten. Etwa gilt
$$(\forall x A(x, ...)) \wedge (\forall x B(x, ...)) \iff \forall x(A(x, ...) \wedge B(x, ...)),$$

jedoch nicht

$$(\forall x A(x,...)) \vee (\forall x B(x,...)) \iff \forall x (A(x,...) \vee B(x,...)).$$

(In der unteren Zeile gilt nur $\Rightarrow$.) Weiter gilt

$$\exists x \forall y A(x,y,...) \Rightarrow \forall y \exists x A(x,y,...),$$

jedoch sicher nicht die Umkehrung (d.h. $\forall y \exists x A(x,y,...) \Rightarrow \exists x \forall y A(x,y,...)$).

1.3 Relationen und Funktionen

Funktionen (oft auch mit dem äquivalenten Begriff *Abbildungen* bezeichnet), wurden bereits in der Schule diskutiert. Wir sammeln grundlegende Definitionen und Tatsachen. Wir führen Funktionen als Spezialfälle von Relationen ein, mit denen wir beginnen werden.

Definition 1.17 (Mengenprodukt). *Seien M, N Mengen. Dann ist*

$$M \times N := \{(x,y) : x \in M \text{ und } y \in N\}$$

das Produkt der Mengen *M und N.*

Für Mengen $M_1, ..., M_n$ ist

$$\bigtimes_{i=1}^{n} M_i := M_1 \times \cdots \times M_n := \{(x_1, ..., x_n) : x_1 \in M_1, ..., x_n \in M_n\}$$

das Produkt der Mengen $M_1, ..., M_n$. Ist $M_1 = \cdots = M_n = M$, so schreiben wir auch $M^n := M_1 \times \cdots \times M_n$. Für $n = 1$ ist $M^1 := M$.

Definition 1.18 (Relation). *Seien $M, N \neq \emptyset$ Mengen. Jede Teilmenge $R \subseteq M \times N$ heißt* Relation *(oder auch* Korrespondenz*). Wir bezeichnen mit $\mathcal{D}(R) := \{x \in M : \exists y \in N \ (x,y) \in R\}$ den* Definitionsbereich *von R und mit $\mathcal{W}(R) := \{y \in N : \exists x \in M \ (x,y) \in R\}$ den* Wertebereich *von R. Ist $(x,y) \in R$, so heißt x ein* Urbild *von y und y ein* Wert *von x. Weiter schreiben wir oft xRy falls $(x,y) \in R$. Eine Relation $R \subseteq M \times N$ heißt*

- eindeutig*, falls $(x,y_1),(x,y_2) \in R \Rightarrow y_1 = y_2$. In diesem Fall schreibt man statt $(x,y) \in R$ auch $y = R(x)$,*
- eindeutig umkehrbar*, falls $(x_1,y),(x_2,y) \in R \Rightarrow x_1 = x_2$.*

Ist $M = N$, so heißt R

- reflexiv*, falls $\forall x \in M \ (x,x) \in R$,*
- symmetrisch*, falls $(x,y) \in R \iff (y,x) \in R$,*
- antisymmetrisch*, falls $(x,y),(y,x) \in R \Rightarrow x = y$,*
- transitiv*, falls $(x,y),(y,z) \in R \Rightarrow (x,z) \in R$.*

Für eine Relation $R \subseteq M \times N$ definieren wir die inverse Relation *$R^{-1} \subseteq N \times M$ durch $R^{-1} := \{(y,x) : (x,y) \in R\}$.*

Für Relationen $R \subseteq M \times N$ und $S \subseteq N \times O$ definieren wir die Verknüpfung *(oder* Komposition*) $S \circ R \subseteq M \times O$ durch*

$$S \circ R := \{(x,z) \in M \times O : \exists y \in N \ (x,y) \in R, (y,z) \in S\}.$$

Lemma 1.19 (Rechenregeln für Relationen). *Seien R, S, T Relationen. Dann gilt*

$$\begin{aligned}&\mathcal{D}(R^{-1}) = \mathcal{W}(R), \qquad \mathcal{W}(R^{-1}) = \mathcal{D}(R),\\ &(R^{-1})^{-1} = R,\\ &T \circ (S \circ R) = (T \circ S) \circ R,\\ &(S \circ R)^{-1} = R^{-1} \circ S^{-1}.\end{aligned}$$

Beweis. Die ersten drei Behauptungen sind klar nach Definition. Für die vierte berechnen wir

$$\begin{aligned}(w,z) \in T \circ (S \circ R) &\iff \exists y \ (w,y) \in S \circ R, (y,z) \in T\\ &\iff \exists x, y \ (w,x) \in R, (x,y) \in S, (y,z) \in T\\ &\iff \exists x \ (w,x) \in R, (x,z) \in T \circ S\\ &\iff (w,z) \in (T \circ S) \circ R.\end{aligned}$$

Schließlich folgt die letzte Behauptung aus

$$\begin{aligned}(z,x) \in (S \circ R)^{-1} &\iff (x,z) \in S \circ R\\ &\iff \exists y \ (x,y) \in R, (y,z) \in S\\ &\iff \exists y \ (z,y) \in S^{-1}, (y,x) \in R^{-1}\\ &\iff (z,x) \in R^{-1} \circ S^{-1}.\end{aligned}$$

□

Beispiel 1.20.

- Sei $M = N = \mathbb{Q}$. Wir betrachten die Relation

  $$\leq \ := \{(x,y) : x \text{ kleiner oder gleich } y\}.$$

  (Hier schreibt man meist $x \leq y$ anstelle von $(x,y) \in \leq$.) Diese Relation ist reflexiv, antisymmetrisch und transitiv. Allerdings ist sie weder eindeutig, noch eindeutig umkehrbar noch symmetrisch.

- Sei $M = N = \mathbb{Q}$. Die Relation

  $$= \ := \{(x,y) : x = y\}$$

  ist ebenfalls reflexiv, antisymmetrisch und transitiv, und außerdem noch symmetrisch, eindeutig und eindeutig umkehrbar.

- Sei E eine Menge und 2^E die Potenzmenge von E. Die Relation

  $$\subseteq := \{(M,N) : M \subseteq N\}$$

  auf 2^E ist genau wie $\leq$ reflexiv, antisymmetrisch und transitiv. Allerdings bemerken wir, dass für $x, y \in \mathbb{Q}$ entweder $x \leq y$ oder $y \leq x$ immer gilt, es jedoch Mengen M, N geben kann, so dass weder $M \subseteq N$ noch $N \subseteq M$ gilt.

- Sei $M = N = \mathbb{N}_0$. Wir definieren

  $$\sim := \{(x,y) : |y - x| \text{ ist durch 3 teilbar}\}.$$

  Diese Relation ist reflexiv, symmetrisch und transitiv, allerdings weder eindeutig, eindeutig umkehrbar noch antisymmetrisch.

Definition 1.21 (Funktion). *Seien M, N Mengen. Eine* Funktion *(oder* Abbildung*) F ist eine eindeutige Relation $F \subseteq M \times N$ mit $\mathcal{D}(F) = M$. Sie heißt*

- injektiv*, falls sie eindeutig umkehrbar ist,*
- surjektiv*, falls $\mathcal{W}(F) = N$,*
- bijektiv*, falls sie injektiv und surjektiv ist.*

Eine Funktion $a : \mathbb{N} \to M$ heißt auch (M-wertige) Folge.

Bemerkung 1.22 (Notation).

- Sei $F \subseteq M \times N$ eine Funktion. Wir schreiben hierfür auch
$$F : \begin{cases} M & \to N \\ x & \mapsto F(x), \end{cases}$$
falls $F = \{(x, F(x)) : x \in M\}$.
- Eine Folge $a : \mathbb{N} \to M$ können wir auch schreiben als $a_1, a_2, ...$ oder $(a_i)_{i=1,2,...}$, wobei $a_i := a(i)$ ist. Wir zählen also einfach alle Funktionswerte der Funktion a der Reihe nach auf.
- In Verallgemeinerung sei $F : M \to N, m \mapsto F_m$ eine Funktion. Diese schreiben wir auch als $F = (F_m)_{m \in M}$. Wir sprechen hier auch von der N-wertigen *Familie* $(F_m)_{m \in M}$.
- Sei $(A_m)_{m \in M}$ eine 2^N-wertige Familie, d.h. $A_m \subseteq N$ für alle $m \in M$. Dann ist $\bigcap_{m \in M} A_m := \{n \in N : \forall m \in M \; n \in A_m\}$ der Schnitt der $(A_m)_{m \in M}$ und $\bigcup_{m \in M} A_m := \{n \in N : \exists m \in M \; n \in A_m\}$ die Vereinigung der $(A_m)_{m \in M}$.

Beispiel 1.23.

- Für Mengen M, N und $c \in N$ bezeichnen wir mit
$$c : \begin{cases} M & \to N \\ x & \mapsto c \end{cases}$$
die konstante Abbildung.
- Für eine Menge M bezeichnen wir mit
$$id : \begin{cases} M & \to M \\ x & \mapsto x. \end{cases}$$
die Identitätsabbildung.
- Für Mengen $M_1, ..., M_d$ bezeichnen wir mit
$$\pi_i : \begin{cases} M_1 \times \cdots \times M_d & \to M_i \\ (x_1, ..., x_d) & \mapsto x_i \end{cases}$$
die Projektion auf die i-te Koordinate.

Definition 1.24 (Funktionenmenge). *Seien M, N Mengen. Wir setzen*
$$N^M := \{F : M \to N \text{ Funktion}\}$$
die Menge der Funktionen *von M nach N.*

Bemerkung 1.25.

- Sei $M = \{1, ..., d\}$ und $N = \mathbb{Q}$. Dann wird jede Abbildung $x : M \to N$ durch $x_1 := x(1), ..., x_d := x(d) \in \mathbb{Q}$ definiert. Deshalb schreiben wir auch

$$\mathbb{Q}^d := N^M = \{(x_1, ..., x_d) : x_1, ..., x_d \in \mathbb{Q}\}.$$

- Die Schreibweise N^M erinnert eventuell an die Schreibweise 2^M für die Potenzmenge von M aus Definition 1.9. Dies ist kein Zufall. Wir können nämlich 2^M bijektiv auf $\{0,1\}^M$ abbilden, indem wir

$$\begin{cases} 2^M & \to \{0,1\}^M \\ M' & \mapsto \begin{cases} M & \to \{0,1\} \\ x & \mapsto \begin{cases} 1, & x \in M' \\ 0, & x \notin M' \end{cases} \end{cases} \end{cases}$$

setzen.

Lemma 1.26 (Charakterisierung von Injektivität/Surjektivität). *Sei $F : M \to N$ eine Funktion.*

- *Die Funktion F ist genau dann injektiv, wenn es für alle $y \in \mathcal{W}(F)$ genau ein $x \in M$ gibt mit $F(x) = y$. (Mit Quantoren: $\forall y \in \mathcal{W}(F) \exists! x \in M : F(x) = y$.)*

- *Die Funktion F ist genau dann surjektiv, wenn es für alle $y \in N$ ein $x \in M$ gibt mit $F(x) = y$. (Mit Quantoren: $\forall y \in N \exists x \in M : F(x) = y$.)*

Beweis. Sei F injektiv und $y \in \mathcal{W}(F)$. Nach Definition gilt dann $F(x_1) = y, F(x_2) = y \Rightarrow x_1 = x_2$. Dies ist genau dann der Fall wenn es genau ein $x \in M$ gibt mit $F(x) = y$.

Sei F surjektiv und $y \in N$. Nach Definition ist also $y \in \mathcal{W}(F)$ und damit gibt es ein $x \in M$ mit $F(x) = y$. Die Umkehrung ist klar. □

Lemma 1.27 (Kompositionen von injektiven und surjektiven Funktionen). *Die Komposition von Funktionen, Injektionen, Surjektionen, Bijektionen ist wieder eine Funktion, Injektion, Surjektion, Bijektion.*

Jede injektive Funktion $F : M \to N$ definiert eine Funktion $F^{-1} : \mathcal{W}(F) \to M$ mit $F \circ F^{-1} = F^{-1} \circ F = id$. Ist F bijektiv, so ist $F^{-1} : N \to M$.

Beweis. Seien $F : M \to N$ und $G : N \to O$ Funktionen. Sind F und G injektiv, so gibt es für $z \in \mathcal{W}(G \circ F) \subseteq \mathcal{W}(G)$ wegen der Injektivität von G genau ein $y \in \mathcal{W}(F) \subseteq N$ mit $G(y) = z$. Weiter gibt es wegen der Injektivität von F genau ein x mit $F(x) = y$. Insgesamt gibt es also genau ein x mit $(G \circ F)(x) = G(F(x)) = z$. Sind F und G surjektiv, so gibt es für jedes $z \in O$ ein $y \in N$ mit $G(y) = z$ und für jedes $y \in N$ ein $x \in M$ mit $F(x) = y$. Damit gibt es also für jedes $z \in O$ ein $x \in M$ mit $(G \circ F)(x) = G(F(x)) = z$. Die Behauptung für bijektive Funktionen ist damit auch klar.

Ist F injektiv, so definieren wir $F^{-1}(y)$ für $y \in \mathcal{W}(F)$ durch $F^{-1}(y) = x$ falls $x \in M$ – das wegen der Injektivität einzige $x \in M$ ist – mit $F(x) = y$. In diesem Fall gilt $(F^{-1} \circ F)(x) = F^{-1}(F(x)) = F^{-1}(y) = x$ sowie $(F \circ F^{-1})(y) = F(F^{-1})(y) = F(x) = y$, woraus die restlichen Behauptungen folgen. □

1.4 Mächtigkeit von Mengen

Mengen kann man nach der Anzahl ihrer Elemente unterscheiden. Generell ist eine Menge entweder endlich oder unendlich groß. Letztere Mengen kann man entweder abzählen oder auch nicht.

Definition 1.28 (Mächtigkeit, Gleichmächtigkeit, Abzählbarkeit). *Seien M, N Mengen. Wir bezeichnen mit $|M|$ (oder auch mit $\#M$) die Anzahl der Elemente von M. Diese nennen wir auch die* Mächtigkeit *von M. Die Menge M heißt* endlich*, falls $|M| < \infty$. Die Mengen M und N heißen* gleichmächtig*, wenn es eine Bijektion $M \to N$ gibt. Die Menge M heißt* mächtiger *als N, wenn M nicht gleichmächtig ist wie N, aber ein $M_1 \subseteq M$ gleichmächtig ist wie N.*

Gibt es eine Bijektion $\mathbb{N} \to M$, so heißt M abzählbar*. Ist M entweder endlich oder abzählbar, so heißt M* höchstens abzählbar*. Ist M nicht höchstens abzählbar, so heißt M* überabzählbar*.*

Beispiel 1.29.

- Es ist $|\{1,2,3\}| = |\{2,3,5\}| = 3$.
- Die Menge der geraden Zahlen, $M := \{2,4,6,8,...\}$ ist abzählbar. Es gibt nämlich die Bijektion
$$\begin{cases} \mathbb{N} & \to M \\ n & \mapsto 2n. \end{cases}$$
Diese ist eine Bijektion zwischen $\mathbb{N}$ und M. Man sieht, dass unendliche Mengen zu echten Teilmengen gleichmächtig sein können.
- Wir werden in Bemerkung 1.32 sehen, dass $2^{\mathbb{N}}$, die Potenzmenge von $\mathbb{N}$, überabzählbar ist. Noch weiter unten werden wir sehen, dass die Menge der reellen Zahlen, $\mathbb{R}$, ebenfalls überabzählbar ist.

Lemma 1.30. *Sowohl $\mathbb{Z}$ als auch $\mathbb{Q}$ sind abzählbar.*

Beweis. Die Abzählbarkeit von $\mathbb{Z}$ ist schnell gezeigt. Wir definieren eine Abbildung

$$b_1 : \begin{cases} \mathbb{N} & \to \mathbb{Z} \\ n & \mapsto \begin{cases} n/2, & \text{falls } n \text{ gerade}, \\ -(n-1)/2, & \text{falls } n \text{ ungerade}. \end{cases} \end{cases}$$

und überlegen uns, dass diese bijektiv ist. Alternativ sagen wir einfach, dass $0, 1, -1, 2, -2, 3, -3, ...$ eine Abzählung von $\mathbb{Z}$ liefert.

Um die Abzählbarkeit von $\mathbb{Q}$ zu zeigen, beweisen wir zunächst die Abzählbarkeit von $\mathbb{N}^2$. Hierzu verwenden wir das Cantor'sche Diagonalverfahren an; siehe Abbildung 1.1. Dies führt zu einer Abzählung $(1,1), (1,2), (2,1), (3,1), (2,2), (1,3), (1,4), (2,3), (3,2), (4,1), ...$ von $\mathbb{N}^2$. Mit anderen Worten haben wir hiermit eine Bijektion $b_2 : \mathbb{N} \to \mathbb{N}^2$ definiert. Weiter definiert $b_3 : \mathbb{N}^2 \to \mathbb{Z} \times \mathbb{N}, (n_1, n_2) \mapsto (b_1(n_1), b_2)$ eine Bijektion und $s : \mathbb{Z} \times \mathbb{N} \to \mathbb{Q}, (m,n) \mapsto m/n$ eine Surjektion. Insgesamt ist also $s \circ b_3 \circ b_2 : \mathbb{N} \to \mathbb{Q}$ eine Surjektion. Durch 'Überspringen' von bereits abgezählten Werten aus $\mathbb{Q}$ erhält man die Abzählung

$$0, \cancel{0}, 1, -1, \tfrac{1}{2}, \cancel{0}, \cancel{0}, \tfrac{1}{3}, -\tfrac{1}{2}, 2, \cdots$$

von $\mathbb{Q}$ und die Abzählbarkeit von $\mathbb{Q}$ ist gezeigt. $\square$

$$\begin{tikzcd}
(1,1) \arrow[r] & (1,2) \arrow[dl] & (1,3) \arrow[r] & (1,4) \arrow[dl] & \cdots \\
(2,1) \arrow[d] & (2,2) \arrow[ur] & (2,3) \arrow[dl] & (2,4) & \cdots \\
(3,1) \arrow[ur] & (3,2) \arrow[dl] & (3,3) & (3,4) & \cdots \\
(4,1) & (4,2) & (4,3) & (4,4) & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots
\end{tikzcd}$$

Abbildung 1.1: Das Cantor'sche Diagonalverfahren liefert eine Abzählung von $\mathbb{N}^2$.

Wir betrachten nun noch die Mächtigkeiten einer Menge M und ihrer Potenzmenge 2^M (siehe Beispiel 1.20).

Lemma 1.31. *Sei M eine Menge. Dann ist 2^M mächtiger als M.*

Beweis. Angenommen, es gäbe eine Bijektion $F : M \to 2^M$. Wir definieren die Menge $M' := \{x : x \notin F(x)\} \in 2^M$. Da F bijektiv ist, muss es ein $x \in M$ geben mit $F(x) = M'$. Nun gibt es zwei Möglichkeiten. Entweder ist $x \in M'$ oder $x \notin M'$. Im ersten Fall folgt aus $x \in M'$, dass $x \notin F(x) = M'$, was einen Widerspruch liefert. Im zweiten Fall ($x \notin M'$) ist also $x \notin F(x)$ und damit $x \in M'$ nach Defintion von M'. Wieder ergibt sich ein Widerspruch. Da die Annahme der Existenz der Bijektion F auf jeden Fall auf einen Widerspruch führt, kann sie nicht stimmen. Also gibt es ein solches F nicht. Weiter ist $\{\{x\} : x \in M\} \subseteq 2^M$ und offenbar gibt es eine Bijektion dieser Menge mit M. Damit ist 2^M mächtiger als M. $\square$

Bemerkung 1.32. Aus dem letzten Lemma folgt, dass $2^{\mathbb{N}}$, die Potenzmenge der natürlichen Zahlen, überabzählbar ist.

1.5 Beweise

Beweise sind die Kernstücke der Mathematik. Um sie zu beherrschen müssen ein paar Regeln und Techniken gelernt werden. Insbesondere wollen wir in diesem Abschnitt drei verschiedene Arten mathematischer Beweise kennen lernen: den direkten Beweis, den indirekten Beweis und die vollständige Induktion. Wir beginnen mit dem – konzeptionell einfachsten – direkten Beweis. Bei diesem schließen oder berechnen wir direkt mit Hilfe der zuvor definierten Begriffe. Beispiele für direkte Beweise haben wir bereits in Lemma 1.5, Lemma 1.15, Lemma 1.19, Lemma 1.26, Lemma 1.27 und Lemma 1.30 kennen gelernt. Nachdem wir Summen- und Produktschreibweisen eingeführt haben, bringen wir noch ein Beispiel, in dem eine Formel mittels direkter Berechnung bewiesen wird.

Oftmals werden für die in den Tutorien zu besprechenden Aufgaben Beweise gebraucht. Wir weisen darauf hin, dass die Art und Weise, diese aufzuschreiben, typischerweise von der Art und Weise, wie sie etwa in diesem Vorlesungsskript stehen, abweicht. Auf Lösungen von Übungsblättern wird typischerweise mehr mit Quantoren und Folgepfeilen gearbeitet, wohingegen ein Skript aus ganzen Sätzen besteht. Um diesen Unterschied zu verdeutlichen, bringen wir für ein paar Sätze zwei Beweise, nämlich einen in der Skript- und einen in der Übungsblattversion (siehe die Beweise der Propositionen 1.36, 1.37, 1.38 und 1.40).

Definition 1.33 (Häufige Schreibweisen). *Sei $x_1, x_2, \dots$ eine Folge von Zahlen. Dann definieren wir*

$$\sum_{i=1}^{n} x_i := x_1 + \cdots + x_n,$$
$$\prod_{i=1}^{n} x_i := x_1 \cdots x_n$$

die Summe *und das* Produkt *der ersten n dieser Zahlen. (Als Konvention setzen wir $\sum_{i=1}^{0} x_i = 0$ und $\prod_{i=1}^{0} x_i = 1$.) Weiter setzen wir für $k, n \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$ mit $k \leq n$*

$$n! := \prod_{i=1}^{n} i = 1 \cdots n,$$
$$\binom{n}{k} := \frac{n!}{k!(n-k)!} = \prod_{i=1}^{k} \frac{i+n-k}{i}$$

sowie $\binom{n}{k} = 0$ für $k < 0$ oder $k > n$. (Man beachte, dass $0! := 1$ wegen der Konvention für leere Produkte gilt.) Hier heißt $n!$ die Fakultät *von n und die Zahlen $\binom{n}{k}$ heißen* Binomialkoeffizienten.

Proposition 1.34 (Rechenregel für Binomialkoeffizienten). *Für $k, n \in \mathbb{N}$, $k \leq n$ gilt*

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

Beweis. Wir berechnen direkt

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

□

Wir kommen nun zum indirekten Beweis, der prinzipiell anders als der direkte funktioniert. Formal verwenden wir (siehe Lemma 1.35), dass $A \Rightarrow B$ äquivalent ist zu $\neg B \Rightarrow \neg A$. Um also B aus A zu folgern, genügt es, $\neg A$ aus $\neg B$ zu schließen. Als Beispiel erinnern wir an den Beweis von Lemma 1.31. Hier war A: "M ist eine Menge" und "B: Es gibt keine Bijektion $F : M \to 2^M$". Anstatt direkt zu schließen, haben wir gezeigt dass aus "$\neg B$: Es gibt eine Bijektion $F : M \to 2^M$" ein Widerspruch folgt. Insbesondere kann damit A nicht gelten.

Lemma 1.35 (Grundprinzip von indirekten Beweisen). *Seien A, B Aussagen. Dann gilt*

$$(A \Rightarrow B) \iff (\neg B \Rightarrow \neg A).$$

Beweis. Wir haben bereits gesehen, dass $A \Rightarrow B$ äquivalent ist zu $B \vee \neg A$. (Denn: beide Aussagen bedeuten, dass es nicht sein kann, dass zwar A, aber B nicht gilt.) Genauso ist $\neg B \Rightarrow \neg A$ äquivalent zu $\neg A \vee (\neg\neg B)$, was dasselbe ist wie $\neg A \vee B$. Damit sind beide Seiten äquivalent zu $B \vee \neg A$, haben also insbesondere immer denselben Wahrheitswert. □

Wir bringen ein zweites Beispiel eines indirekten Beweises. Ohne weitere Erläuterung werden wir dabei bekannte Teilbarkeitsregeln für natürliche Zahlen verwenden.

Proposition 1.36 ($\sqrt{2} \notin \mathbb{Q}$). *Es gibt kein $x \in \mathbb{Q}$ mit $x^2 = 2$.*

Beweis in Skript-Version. Angenommen, es gäbe $x \in \mathbb{Q}$ mit $x^2 = 2$. Dann ist $x = m/n$ für $m, n \in \mathbb{N}$. (Es ist auch $m \geq 0$, da mit $x^2 = 2$ auch $(-x)^2 = 2$ gilt.) Durch Kürzen können wir m und n so wählen, dass beide teilerfremd sind. Es gilt also $m^2 = 2n^2$ für teilerfremde $m, n \in \mathbb{N}$. Da 2 somit m^2 teilt, muss 2 auch m teilen. Es gibt also ein $r \in \mathbb{N}$ mit $m = 2r$ und damit $4r^2 = 2n^2$ oder $2r^2 = n^2$. Genauso folgern wir, dass es ein s gibt mit $n = 2s$. Wir sehen, dass 2 sowohl m als auch n teilt. Dies ist ein Widerspruch zur Teilerfremdheit von m und n. Also muss die Annahme $x \in \mathbb{Q}$ falsch gewesen sein. Es gibt also kein $x \in \mathbb{Q}$ mit $x^2 = 2$. □

Beweis in Übungsblatt-Version.
Angenommen, $\exists x \in \mathbb{Q} : x^2 = 2$.
$\Rightarrow \exists m, n \in \mathbb{N} : m^2/n^2 = 2$; oBdA[4] m, n teilerfremd
$\Rightarrow \exists m, n \in \mathbb{N}$ teilerfremd, $m^2 = 2n^2$
$\Rightarrow \exists r, m, n \in \mathbb{N} : m = 2r, 4r^2 = 2n^2$ und m, n teilerfremd, also auch $2r^2 = n^2$
$\Rightarrow \exists r, s, m, n \in \mathbb{N} : m = 2r, n = 2s$ und m, n teilerfremd ↯
Also: $\nexists x \in \mathbb{Q} : x^2 = 2$. □

Eine dritte Beweisart ist die der vollständigen Induktion. Sie funktioniert höchstens dann, wenn wir eine Aussage der Form $\forall n \in \mathbb{N}\, A(n)$ beweisen wollen. Wir erinnern an die Definition der natürlichen Zahlen nach Peano aus Beispiel 1.8.5. Wir haben gesehen, dass jede Menge $M \subseteq \mathbb{N}$, für die $1 \in M$ und $n \in M \Rightarrow n+1 \in M$ gilt bereits die Menge der natürlichen Zahlen sein muss. Daraus lesen wir direkt das *Beweisprinzip der vollständigen Induktion* ab:

Sei $A(n)$ eine Aussage und $M := \{n : A(n) \text{ ist wahr}\}$. Gilt

1. *Induktionsanfang:* $A(1)$,
2. *Induktionsschluss:* $A(n) \Rightarrow A(n+1)$,

dann ist $M = \mathbb{N}$, d.h. $\forall n \in \mathbb{N}\ A(n)$ ist wahr.

Wir verwenden dieses Beweisprinzip um ein paar Rechenregeln für Summen und Produkte kennen zu lernen. Hierbei werden wir wieder jeweils zwei Beweise angeben, einen wie er in einem Vorlesungsskript erscheinen könnte und einen zweiten, wie man ihn eventuell als Lösung eines Übungsblattes verwenden würde.

Proposition 1.37. *Für $n \in \mathbb{N}$ gilt*

$$\sum_{i=1}^{n} i = \binom{n+1}{2}.$$

[4] Diese vier Buchstaben bedeuten *Ohne Beschränkung der Allgemeinheit.* Sie besagen, dass wir keinen Fehler machen, wenn wir eine Annahme treffen.

Beweis in Skript-Version. Wir verwenden vollständige Induktion. Für $n = 1$ ist die Behauptung klar, weil beide Seiten 1 ergeben. Gilt die Behauptung für ein $n-1$, so berechnen wir

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

so dass die Behauptung auch für n gilt. Damit folgt die Aussage. $\square$

Beweis in Übungsblatt-Version. Induktionsanfang: $1 = 1$ ✓
Induktionsschluss $n-1 \to n$:

$$\sum_{i=1}^{n} i = n + \sum_{i=1}^{n-1} i \overset{\text{Induktions-}}{\underset{\text{annahme}}{=}} n + \binom{n}{2} = \frac{2n + n(n-1)}{2} = \frac{(n+1)n}{2} = \binom{n+1}{2}.$$

$\square$

Proposition 1.38 (Geometrische Reihe). *Für $x \in \mathbb{R}$ und $n \in \mathbb{N}$ gilt, falls $x \neq 1$,*

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

Beweis in Skript-Version. Wir verwenden vollständige Induktion. Für $n = 1$ ist die Behauptung klar, weil beide Seiten 1 ergeben. Gilt die Behauptung für ein $n-1$, so berechnen wir

$$\sum_{i=0}^{n} x^i = x^n + \sum_{i=0}^{n-1} x^i = x^n + \frac{1-x^n}{1-x} = \frac{x^n - x^{n+1} + 1 - x^n}{1-x} = \frac{1-x^{n+1}}{1-x},$$

so dass die Behauptung auch für n gilt. Damit folgt die Aussage. $\square$

Beweis in Übungsblatt-Version. Induktionsanfang: $1 = 1$ ✓
Induktionsschluss $n-1 \to n$:

$$\sum_{i=0}^{n} x^i = x^n + \sum_{i=0}^{n-1} x^i \overset{\text{Induktions-}}{\underset{\text{annahme}}{=}} x^n + \frac{1-x^n}{1-x} = \frac{x^n - x^{n+1} + 1 - x^n}{1-x} = \frac{1-x^{n+1}}{1-x}.$$

$\square$

Proposition 1.39 (Bernoulli'sche Ungleichung). *Für $x > -1, x \neq 0$ und $n \in \{2, 3, ...\}$ gilt*

$$(1+x)^n > 1 + nx.$$

Beweis in Skript-Version. Die Behauptung ist für $n = 2$ klar, schließlich ist $(1+x)^2 = 1 + 2x + x^2 > 1 + 2x$. Gilt sie für ein $n-1$, so folgt (wegen $1 + x > 0$)

$$(1+x)^n = (1+x)^{n-1}(1+x) \geq (1 + (n-1)x)(1+x) = 1 + nx + (n-1)x^2 \geq 1 + nx.$$

$\square$

Beweis in Übungsblatt-Version. Induktionsanfang: $(1+x)^2 = 1 + 2x + x^2 > 1 + 2x$ ✓
Induktionsschluss $n-1 \to n$:

$$(1+x)^n = (1+x)^{n-1}(1+x) \underset{\text{annahme, } x>-1}{\overset{\text{Induktions-}}{\geq}} (1+(n-1)x)(1+x)$$
$$= 1 + nx + (n-1)x^2 \geq 1 + nx.$$

□

Proposition 1.40 (Binomische Formel). *Für $x, y \in \mathbb{R}$ und $n \in \mathbb{N}$ gilt*

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

Beweis in Skript-Version. Wir verwenden vollständige Induktion. Für $n = 1$ ist die Behauptung klar, weil beide Seiten $x + y$ ergeben. Gilt die Behauptung für ein $n-1$, so berechnen wir

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

Hier haben wir im vierten Gleichheitszeichen eine Indexverschiebung vorgenommen und im letzten Gleichheitszeichen Proposition 1.34 verwendet. Insgesamt folgt die Behauptung auch für n. □

Die Übungsblatt-Version des Beweises sparen wir uns an dieser Stelle, da sie leicht aus den vorigen Beispielen übertragen werden kann.

Bemerkung 1.41 (Varianten der vollständigen Induktion). Es gibt neben der oben beschriebenen Variante der vollständigen Induktion auch noch weitere Varianten.

- Ist $A(n)$ eine Aussage und $M := \{n : A(n) \text{ ist wahr }\}$. Gilt
  1. *Induktionsanfang*: $A(1), ..., A(k)$ für ein $k = 1, 2, ...$,
  2. *Induktionsschluss*: $A(n) \Rightarrow A(n+1)$ für $n = k, k+1, ...$

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

  Das bedeutet, dass die ersten k Aussagen $A(1), ..., A(k)$ einzeln bewiesen werden, und anschließend erst der Induktionsschluss ausgeführt wird.

- Ist $A(n)$ eine Aussage und $M := \{n : A(n) \text{ ist wahr }\}$. Gilt

1. *Induktionsanfang*: $A(1),...,A(k)$ für ein $k = 1, 2, ...,$
2. *Induktionsschluss*: $A(1),...,A(n) \Rightarrow A(n+1)$ für $n = k, k+1, ...$

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

Das bedeutet, dass beim Beweis von $A(n+1)$ im Induktionsschluss nicht nur auf die Gültigkeit von $A(n)$ zurück gegriffen wird, sondern z.B. auch auf $A(n-1)$.

Bemerkung 1.42 (*Wertigkeit* der Beweisarten). Man möchte meinen, dass es sich bei der Mathematik um eine wertfreie Wissenschaft handelt. Allerdings bemerken wir, dass in der Mathematik manche Beweise höher bewertet werden als andere. Generell gilt (falls möglich): der direkte Beweis ist immer den anderen beiden Beweismethoden vorzuziehen. Wir geben einen – auf Gauß zurückgehenden – alternativen Beweis für Proposition 1.37. Wir beweisen die Formel direkt mittels

$$\sum_{i=1}^{n} i = (1 + \cdots + n) = \tfrac{1}{2}\big((1 + \cdots + n) + (n + \cdots + 1)\big)$$
$$= \frac{1}{2}\Big(\sum_{i=1}^{n} i + \sum_{i=1}^{n}(n - i + 1)\Big) = \frac{1}{2}\sum_{i=1}^{n}(i + n - i + 1) = \frac{(n+1)n}{2} = \binom{n+1}{2}.$$

Mittels eines Rechentricks (der im zweiten Gleichheitszeichen steckt) haben wir also nun auch einen direkten Beweis gefunden.

Auch indirekte Beweise sollte man den direkten nicht vorziehen. Der Grund ist, dass diese oftmals schwieriger zu durchdringen sind, und manchmal auch in einen direkten Beweis übergeführt werden können. (Dies scheint allerdings in den Beweisen von Lemma 1.31 und Proposition 1.36 nicht möglich.)

2 Die reellen Zahlen

Wir haben bereits die – aus der Schule bekannten – Mengen der natürlichen, ganzen und rationalen Zahlen

$$\begin{aligned}\mathbb{N} &:= \{1, 2, 3, ...\},\\ \mathbb{Z} &:= \{0, \pm 1, \pm 2, ...\},\\ \mathbb{Q} &:= \{m/n : m \in \mathbb{Z}, n \in \mathbb{N}\}\end{aligned}$$

kennen gelernt. Wir setzen außerdem

$$\begin{aligned}\mathbb{Z}_+ &:= \mathbb{N}_0 := \{0, 1, 2, ...\},\\ \mathbb{Q}_+ &:= \{m/n : m \in \mathbb{Z}_+, n \in \mathbb{N}\}\end{aligned}$$

Weiter sind bereits die reellen Zahlen bekannt, die wir in Abschnitt 2.1 erstmal formal sauber definieren wollen. Außerdem beschäftigen wir uns in Abschnitt 2.3 mit der Dezimal-Darstellung reeller Zahlen, um dann in 2.4 deren Überabzählbarkeit zu beweisen.

2.1 Konstruktion der reellen Zahlen

Die Konstruktion der reellen Zahlen ist der erste große mathematische Meilenstein der Vorlesung. Zentral bei unserer Konstruktion ist der Begriff der Cauchy- und der Nullfolge. Wir erinnern daran, dass Folgen rationaler Zahlen wie in Bemerkung 1.22 durch $(x_n)_{n=1,2,...}$ gegeben sind.

Definition 2.1 (Cauchy-Folgen und Null-Folgen). *Sei $(x_n)_{n=1,2,...}$ eine Folge rationaler Zahlen.*

- *Die Folge $(x_n)_{n=1,2,...}$ heißt* Cauchy-Folge *(oder* Fundamentalfolge*), falls es für jedes $\varepsilon > 0$ (und $\varepsilon \in \mathbb{Q}$) ein $N \in \mathbb{N}$ gibt, so dass für alle $m, n > N$ gilt, dass*
$$|x_n - x_m| < \varepsilon.$$
In Quantorenschreibweise bedeutet dies
$$\forall \varepsilon > 0 \; \exists N \in \mathbb{N} \; \forall m, n > N : |x_n - x_m| < \varepsilon.$$
Wir definieren die Menge der Cauchy-Folgen durch
$$\mathcal{C} := \{(x_n)_{n=1,2,...} \text{ Cauchy-Folge}\}.$$

- *Die Folge $(x_n)_{n=1,2,...}$ heißt* Null-Folge *falls es für jedes $\varepsilon > 0$ (und $\varepsilon \in \mathbb{Q}$) ein $N \in \mathbb{N}$ gibt mit $|x_n| < \varepsilon$ für alle $n \geq N$. In Quantorenschreibweise bedeutet das*
$$\forall \varepsilon > 0 \; \exists N \in \mathbb{N} \; \forall n > N : |x_n| < \varepsilon.$$
Wir definieren die Menge der Null-Folgen durch
$$\mathcal{N} := \{(x_n)_{n=1,2,...} \text{ Null-Folge}\}.$$
(Es sollte klar sein, dass $\mathcal{N} \subsetneq \mathcal{C}$, d.h. jede Nullfolge ist eine Cauchy-Folge, aber es gibt jedoch Cauchy-Folgen, die keine Nullfolgen sind.)

- *Eine Cauchy-Folge heißt* positiv*, wenn es ein $\varepsilon > 0$ und ein $N \in \mathbb{N}$ gibt mit $x_n > \varepsilon$ für alle $n \geq N$. Wir sagen auch: Für fast alle n, (d.h. für alle bis auf endlich viele n) gilt $x_n > \varepsilon$. Wir definieren die Menge der positiven Cauchy-Folgen durch*
$$\mathcal{P} := \{(x_n)_{n=1,2,...} \text{ positive Cauchy-Folge}\}.$$

Bemerkung 2.2 (Notation: Maximum, Minimum und Mengenschreibweise).

1. Sei $M \subseteq \mathbb{Q}$ mit $|M| < \infty$, d.h. M ist eine endliche Teilmenge von $\mathbb{Q}$. Dann schreiben wir
$$\begin{aligned} \min M := x \qquad &\iff \qquad (x \in M) \wedge \forall y \in M : x \leq y, \\ \max M := x \qquad &\iff \qquad (x \in M) \wedge \forall y \in M : y \leq x. \end{aligned} \tag{2.1}$$
Weiter gilt $\min M = -\max(-M)$ und $\max M = -\min(-M)$, wobei $-M := \{-x : x \in M\}$. Für Zahlen $x, y \in \mathbb{Q}$ schreiben wir
$$a \wedge b := \min\{a, b\}, \qquad a \vee b := \max\{a, b\}.$$
Später werden wir diese Schreibweisen auch für reelle (also nicht notwendigerweise rationale) Zahlen benutzen.

2. Seien $(x_n)_{n=1,2,...}, (y_n)_{n=1,2,...}$ Folgen rationaler Zahlen. Dann schreiben wir $(x_n)_{n=1,2,...} + (y_n)_{n=1,2,...} := (x_n + y_n)_{n=1,2,...}$ und sagen auch, dass die Addition komponentenweise definiert ist.

   Seien $L, M \subseteq \mathbb{Q}$. Dann schreiben wir

$$L + M := \{x + y : x \in L, y \in M\}$$

   und analog für alle anderen Rechenarten.

   Etwas allgemeiner sei $\bullet : K \times K \to K, (x, y) \mapsto x \bullet y$. Für $L, M \subseteq K$ setzen wir dann $L \bullet M := \{x \bullet y : x \in L, y \in M\} \subseteq K$.

   Als Beispiel bemerken wir, dass $\mathbb{Q} + \mathbb{Q} = \mathbb{Q}$ sowie $\mathbb{Q}_+ \cdot \mathbb{Q} = \mathbb{Q}$ gilt.

Lemma 2.3 (Cauchy-Folgen sind beschränkt). *Jedes $(x_n)_{n=1,2,...} \in \mathcal{C}$ ist beschränkt, d.h. es gibt $K \in \mathbb{Q}$ mit $|x_n| \le K$ für alle $n \in \mathbb{N}$.*

Beweis. Da $(x_n)_{n=1,2,...}$ eine Cauchy-Folge ist, gibt es für $\varepsilon = 1$ ein $N \in \mathbb{N}$ mit $|x_n - x_m| \le 1$ für alle $m, n \ge N$. Wir setzen $K := (\max\{|x_1|, ..., |x_{N-1}|\}) \vee (|x_N| + 1)$. Dann gilt nämlich für $n = 1, ..., N-1$, dass $x_n \le \max\{x_1, ..., x_{N-1}\} \le K$ und für $n \ge N$, dass

$$|x_n| \le |x_N| + |x_n - x_N| \le |x_N| + 1 = K.$$

Insgesamt gilt also $|x_n| \le K$ für alle $n = 1, 2, ...$ □

Wir kommen nun zur Definition der reellen Zahlen. Hierfür benötigen wir den Begriff der Äquivalenzrelation.

Definition 2.4 (Äquivalenzrelation und Äquivalenzklasse). *Sei M eine Menge.*

- *Eine* Äquivalenzrelation $\sim$ *auf M ist eine reflexive, symmetrische und transitive Relation.*
- *Sei $\sim$ eine Äquivalenzrelation auf M und $x \in M$. Dann heißt $\bar{x} := \{y : x \sim y\}$ die* Äquivalenzklasse *von x. Weiter bezeichnen wir mit*

$$M/\sim := \{\bar{x} : x \in M\}$$

  die Menge der Äquivalenzklassen.

Lemma 2.5 (Eigenschaften von Äquivalenzrelationen). *Sei $\sim$ eine Äquivalenzrelation auf M und $x, y \in M$. Dann ist $\bar{x} = \bar{y}$ genau dann, wenn $x \sim y$. Falls $x \sim y$ nicht gilt, ist $\bar{x} \cap \bar{y} = \emptyset$. Insbesondere sind die Elemente von $M/\sim$ paarweise disjunkt.*

Beweis. Angenommen, $x \sim y$ und $z \in \bar{x}$. Dann ist mit $x \sim z$ also auch $y \sim z$ und $z \in \bar{y}$ nach der Transitivität von $\sim$. Daraus folgt $\bar{x} \subseteq \bar{y}$ und analog auch $\bar{y} \subseteq \bar{x}$, also $\bar{x} = \bar{y}$. Ist andersherum $\bar{x} = \bar{y}$, so ist per Definition $x \sim y$.

Falls $x \sim y$ nicht gilt und es ein $z \in \bar{x} \cap \bar{y}$ geben würde, wäre $x \sim z$ und $z \sim y$. Dann müsste also auch $x \sim y$ gelten im Widerspruch zur Annahme. Also ist $\bar{x} \cap \bar{y} = \emptyset$. □

Beispiel 2.6. Seien $m, n \in \mathbb{N}_0$. Wir setzen $m \sim n$, falls $|m - n|$ durch 3 teilbar ist. Dann ist $\sim$ eine Äquivalenzrelation auf $\mathbb{N}_0$. Es gibt genau drei Äquivalenzklassen $\bar{0} = \bar{3} = \bar{6} = ..., \bar{1} = \bar{4} = \bar{7} = ...$ und $\bar{2} = \bar{5} = \bar{8} = ...$

Die reellen Zahlen werden als Menge der Äquivalenzklassen von Cauchy-Folgen definiert. Dies sieht zunächst ungewohnt aus, ist aber konsistent mit früher Gelerntem.

Definition 2.7 (Menge der reellen Zahlen). *Seien $(x_n)_{n=1,2,...}, (y_n)_{n=1,2,...} \in \mathcal{C}$. Wir setzen $(x_n)_{n=1,2,...} \sim (y_n)_{n=1,2,...}$ genau dann, wenn $(x_n - y_n)_{n=1,2,...} \in \mathcal{N}$. Dann ist $\sim$ eine Äquivalenzrelation auf $\mathcal{C}$ und wir definieren*

$$\mathbb{R} := \mathcal{C}/\sim = \{\overline{(x_n)_{n=1,2,...}} : (x_n)_{n=1,2,...} \in \mathcal{C}\}.$$

Lemma 2.8 (Eine alternative Beschreibung von $\mathbb{R}$).

1. *Sei $x = \overline{(x_n)_{n=1,2,...}} \in \mathbb{R}$. Dann ist $x = (x_n)_{n=1,2,...} + \mathcal{N}$, wobei*

$$(x_n)_{n=1,2,...} + \mathcal{N} := \{(x_n)_{n=1,2,...} + (y_n)_{n=1,2,...} : (y_n)_{n=1,2,...} \in \mathcal{N}\}.$$

   Insbesondere ist also

$$\mathbb{R} = \{(x_n)_{n=1,2,...} + \mathcal{N} : (x_n)_{n=1,2,...} \in \mathcal{C}\}.$$

2. *Es gilt $\mathcal{N} + \mathcal{N} = \mathcal{N}$ sowie $\mathcal{N} \cdot \mathcal{N} = \mathcal{N}$. Ist weiter $(x_n)_{n=1,2,...}$ eine beschränkte Folge, dann ist $(x_n)_{n=1,2,...} \cdot \mathcal{N} := \{(x_n \cdot z_n)_{n=1,2,...} : (z_n)_{n=1,2,...} \in \mathcal{N}\} = \mathcal{N}$.*

Beweis. 1. Sei $(z_n)_{n=1,2,...} \in \overline{(x_n)_{n=1,2,...}}$. Dann ist $(z_n - x_n)_{n=1,2,...} \in \mathcal{N}$, also $(z_n)_{n=1,2,...} = (x_n)_{n=1,2,...} + (z_n - x_n)_{n=1,2,...} \in (x_n)_{n=1,2,...} + \mathcal{N}$. Ist andersherum $(z_n)_{n=1,2,...} \in (x_n)_{n=1,2,...} + \mathcal{N}$, dann ist $(z_n - x_n)_{n=1,2,...} \in \mathcal{N}$, also $(z_n)_{n=1,2,...} \in \overline{(x_n)_{n=1,2,...}}$.
2. Sind $(x_n)_{n=1,2,...}, (y_n)_{n=1,2,...} \in \mathcal{N}$, so gibt es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$, so dass $|x_n| < \varepsilon/2, |y_n| < \varepsilon/2$ für alle $n \geq N$. Also gilt $|x_n + y_n| \leq |x_n| + |y_n| < \varepsilon$. Daraus folgt, dass $(x_n + y_n)_{n=1,2,...}$ ebenfalls in $\mathcal{N}$ ist. Die übrigen Aussagen folgen analog. □

Es ist bekannt, dass $\mathbb{R}$ auch die rationalen Zahlen umfasst. Da durch unsere Definition $\mathbb{R}$ nicht aus Zahlen, sondern aus (Äquivalenzklassen von) Cauchy-Folgen besteht, formulieren wir nun, in welchem Sinn $\mathbb{Q} \subseteq \mathbb{R}$ zu verstehen ist.

Lemma 2.9 ($\mathbb{Q} \subseteq \mathbb{R}$). *Die Abbildung $\mathbb{Q} \to \mathbb{R}$, $x \mapsto \overline{(x, x, ...)}$ (wobei $(x, x, ...)$ die Folge ist, die konstant x ist) ist injektiv.*

Beweis. Klar. □

Wir schreiben im Folgenden öfter $x \in \mathbb{R}$. Damit ist gemeint, dass x die Äquivalenzklasse einer Cauchy-Folge $(x_n)_{n=1,2,...}$ mit Werten in $\mathbb{Q}$ ist. Zunächst zeigen wir die Körperstruktur[5] von $\mathbb{R}$.

Definition 2.10 (Addition und Multiplikation reeller Zahlen). *Wir definieren auf $\mathbb{R}$ die Operationen $+$ und $\cdot$ mittels*

$$\begin{aligned} \overline{(x_n)_{n=1,2,...}} + \overline{(y_n)_{n=1,2,...}} &:= \overline{(x_n + y_n)_{n=1,2,...}}, \\ \overline{(x_n)_{n=1,2,...}} \cdot \overline{(y_n)_{n=1,2,...}} &:= \overline{(x_n \cdot y_n)_{n=1,2,...}}. \end{aligned} \tag{2.2}$$

[5] Aus der linearen Algebra ist bekannt: Ein Tripel $(\mathbb{K}, +, \cdot)$ ist ein Körper, wenn (i) $(K, +)$ eine kommutative Gruppe mit $0 \in \mathbb{K}$ als neutralem Element ist, (ii) $(\mathbb{K} \setminus \{0\}, \cdot)$ eine kommutative Gruppe mit neutralem Element $1 \in \mathbb{K}$ und (iii) das Distributivgesetz $x \cdot (y + z) = x \cdot y + x \cdot z$ für alle $x, y, z \in \mathbb{K}$ gilt.

Weiter setzen wir

$$-\overline{(x_n)_{n=1,2,\dots}} := \overline{(-x_n)_{n=1,2,\dots}}, \tag{2.3}$$

$$\overline{(x_n)_{n=1,2,\dots}}^{-1} := \overline{(x_{m_n}^{-1})_{n=1,2,\dots}}, \qquad \textit{falls } \overline{(x_n)_{x=1,2,\dots}} \neq 0, \tag{2.4}$$

wobei $m_1, m_2, \dots$ die Folge der Indizes ist, für die $x_n \neq 0$ gilt.

Bemerkung 2.11 (Wohldefiniertheit). Betrachten wir die Definition (2.2). Es fällt auf, dass das Definiendum (die linke Seite) durch eine Äquivalenzklasse definiert wird. Diese Definition ist jedoch nur möglich, wenn das Definiens (das Objekt auf der rechten Seite) nicht von der Wahl von der Äquivalenzklassen-Elemente $(x_n)_{n=1,2,\dots} \in \overline{(x_n)_{n=1,2,\dots}}$ und $(y_n)_{n=1,2,\dots} \in \overline{(y_n)_{n=1,2,\dots}}$ abhängt. (Seien $(x_n)_{n=1,2,\dots}$ und $(x'_n)_{n=1,2,\dots}$ Cauchy-Folgen in $\mathbb{Q}$ mit $\overline{(x_n)_{n=1,2,\dots}} = \overline{(x'_n)_{n=1,2,\dots}}$ und $(y_n)_{n=1,2,\dots}$ wie in (2.2). Dann könnte es ja sein, dass $\overline{(x_n+y_n)_{n=1,2,\dots}} \neq \overline{(x'_n+y_n)_{n=1,2,\dots}}$. In diesem Fall hätten wir $\overline{(x_n)_{n=1,2,\dots}} + \overline{(y_n)_{n=1,2,\dots}} = \overline{(x'_n)_{n=1,2,\dots}} + \overline{(y_n)_{n=1,2,\dots}}$ nicht gut definiert.)

Die *Wohldefiniertheit* der Addition und Multiplikation von Elementen in $\mathbb{R}$ folgt jedoch aus Lemma 2.8.2. Etwa schreiben wir für $(x_n)_{n=1,2,\dots}, (x'_n)_{n=1,2,\dots}$ mit $\overline{(x_n)_{n=1,2,\dots}} = \overline{(x'_n)_{n=1,2,\dots}}$

$$\begin{aligned}\overline{(x_n+y_n)_{n=1,2,\dots}} &= (x_n+y_n)_{n=1,2,\dots} + \mathcal{N} \\ &= ((x'_n+y_n)_{n=1,2,\dots} + (x_n - x'_n)_{n=1,2,\dots} + \mathcal{N} \\ &= ((x'_n+y_n)_{n=1,2,\dots} + \mathcal{N} = \overline{(x'_n+y_n)_{n=1,2,\dots}}.\end{aligned}$$

Somit hängt die rechte Seite von (2.2) nicht von der Wahl der Repräsentanten der Äquivalenzklassen auf der linken Seite ab.

Proposition 2.12 ($\mathbb{R}$ ist ein Körper). *Die Addition und Multiplikation aus Definition 2.10 sind wohldefiniert. Außerdem ist das Tripel* $(\mathbb{R}, +, \cdot)$ *mit den additiv und multiplikativ Inversen aus* (2.3) *und* (2.4) *ein Körper.*

Beweis. Die Wohldefiniertheit haben wir in der letzten Bemerkung eingesehen. Sei $x = \overline{(x_n)_{n=1,2,\dots}}$. Wir müssen noch zeigen, dass $\overline{(x_n)_{n=1,2,\dots}}^{-1}$ für $x \neq 0$ in der Tat eine Cauchy-Folge ist. Zunächst ist die Folge $(m_n)_{n=1,2,\dots}$ nach Lemma 2.3 wohldefiniert. Sei $\varepsilon > 0$. Da $x \neq 0$, gibt es ein $\delta > 0$, so dass $|x_n| > \delta$ für fast alle n gilt. (Andernfalls wäre ja $(x_n)_{n=1,2,\dots}$ eine Nullfolge und damit $x = 0$.) Weiter wählen wir N so groß, dass $|x_n - x_m| < \delta^2\varepsilon$ für alle $n, m \geq N$. Dann gilt für fast alle m, n

$$|x_n^{-1} - x_m^{-1}| = \frac{|x_m - x_n|}{|x_m x_n|} < \frac{\delta^2\varepsilon}{\delta^2} = \varepsilon.$$

Nun übertragen sich die Körper-Eigenschaften von $\mathbb{Q}$. □

2.2 Ordnung und Vollständigkeit der reellen Zahlen

Wir kommen nun zu einer weiteren Eigenschaft der reellen Zahlen. Es ist nämlich möglich, diese anzuordnen.

Definition 2.13 (Totale Ordnung). *Sei M eine Menge. Eine* Ordnung $\leq$ *auf M ist eine reflexive, antisymmetrische und transitive Relation. Gilt außerdem entweder $x \leq y$ oder $y \leq x$ für alle $x, y \in M$, so heißt $\leq$ eine* totale Ordnung.

Beispiel 2.14. Wir haben in Beispiel 1.20 gesehen, dass das übliche '$\leq$' eine totale Ordnung auf $\mathbb{Q}$ definiert. Weiter haben wir bemerkt, dass für eine Menge M die Relation '$\subseteq$' auf 2^M eine Ordnung definiert. Diese ist jedoch nicht total; ist etwa $M = \{1, 2\}$ sowie $A = \{1\}, B = \{2\}$ so ist weder $A \subseteq B$ noch $B \subseteq A$.

Bemerkung 2.15. Für eine Ordnung $\leq$ auf einer Menge M schreiben wir natürlich $m \geq n$ genau dann, wenn $n \leq m$.

Definition 2.16 (Totale Ordnung auf $\mathbb{R}$). *Sei $x := \overline{(x_n)_{n=1,2,...}}, y := \overline{(y_n)_{n=1,2,...}} \in \mathbb{R}$. Wir setzen $x \leq y$ genau dann, wenn $(y_n - x_n)_{n=1,2,...} \in \mathcal{N} \cup \mathcal{P}$ und $x < y$, falls $(y_n - x_n)_{n=1,2,...} \in \mathcal{P}$. Weiter setzen wir*

$$\mathbb{R}_+ := \{x \in \mathbb{R} : x \geq 0\} = \{x = \overline{(x_n)_{n=1,2,...}} : (x_n)_{n=1,2,...} \in \mathcal{N} \cup \mathcal{P}\}.$$

Lemma 2.17 (Eigenschaften der Ordnung auf $\mathbb{R}$). *Die Relation '$\leq$' auf $\mathbb{R}$ ist eine totale Ordnung. Seien $x := \overline{(x_n)_{n=1,2,...}} \in \mathbb{R}$ und $y := \overline{(y_n)_{n=1,2,...}} \in \mathbb{R}$.*

1. *Es gilt eine der Relationen*
$$x \leq 0 \qquad \text{oder} \qquad -x \leq 0.$$
Weiter ist $x \leq 0$ und $-x \leq 0$ genau dann, wenn $x = 0$.

2. *Gilt $x \neq 0$, so ist $x_n \neq 0$ für fast alle $n \in \mathbb{N}$.*

3. *Aus $0 \leq x$ und $0 \leq y$ folgt $0 \leq xy$ und $0 \leq x + y$.*

4. *Für jedes $x \in \mathbb{R}$ gibt es ein $n \in \mathbb{N}$ mit $x \leq n$.*

Beweis. 1. Angenommen, $x := (x_n)_{n=1,2,...}$ ist keine Nullfolge, d.h. $x \neq 0$, und $\varepsilon > 0$. Dann ist $|x_n| > \varepsilon$ für fast alle n. Da $(x_n)_{n=1,2,...}$ eine Cauchy-Folge ist, ist entweder $x_n > \varepsilon$ für fast alle n oder $x_n < -\varepsilon$ für fast alle n. (Andernfalls wäre $x_n > \varepsilon$ für unendlich viele n und $x_m < -\varepsilon$ für unendlich viele m. Andererseits gilt auch $|x_n - x_m| < \varepsilon$ für m, n groß genug im Widerspruch dazu.) Im ersten Fall gilt $x \geq \varepsilon$, im zweiten Fall $x \leq -\varepsilon$. Ist $(x_n)_{n=1,2,...}$ eine Nullfolge, dann gilt sowohl $-x \leq 0, x \leq 0$ als auch $x = 0$.
2. Wie soeben gesehen, ist $x \neq 0$ genau dann, wenn für jedes $\varepsilon > 0$ gilt, dass $|x_n| > \varepsilon$ für fast alle n. Insbesondere ist dann $x_n \neq 0$ für fast alle n.
3. Diese Rechenregeln folgen bereits aus den entsprechenden Rechenregeln für rationale Zahlen, da $xy = \overline{(x_n y_n)_{n=1,2,...}}$ und $x + y = \overline{(x_n + y_n)_{n=1,2,...}}$.
4. Dies folgt bereits aus Lemma 2.3. □

Lemma 2.18 (Rechnen mit Ungleichungen).

1. *Für $x, z \in \mathbb{R}$ mit $x < z$ gibt es $y \in \mathbb{Q}$ mit $x < y < z$.*[6]

2. *Für $x, z \in \mathbb{R}$ mit $x, z > 0$ gibt es $n \in \mathbb{N}$ mit $xn > z$.*

3. *Für $x, y \in \mathbb{R}_+$ mit $x > 1$ gibt es $n \in \mathbb{N}$ mit $x^n > y$.*

4. *Für $x, y \in \mathbb{R}_+$ mit $x < 1$ gibt es $n \in \mathbb{N}$ mit $x^n < y$.*

[6] Wir schreiben $x < y < z$ anstatt $(x < y) \wedge (y < z)$.

Beweis. 1. Sei $\varepsilon = z - x$. Dann gibt es ein $n \in \mathbb{N}$ mit $n > \varepsilon^{-1}$. Wie betrachten $m := \min\{m' : m'/n > x\}$ sowie $y = m/n$. Dann ist $y > x$ nach Definition und außerdem $y = m/n = \frac{m-1}{n} + \frac{1}{n} < x + \varepsilon = z$.
2. Klar, weil es ein $n \in \mathbb{N}$ gibt mit $n > z/x$.
3. Nach 2. gibt es ein $n \in \mathbb{N}$ mit $(x-1)n > y$. Wir schreiben mit Proposition 1.39

$$x^n = (1 + (x-1))^n > 1 + (x-1)n > 1 + y > y.$$

4. Wegen 3. gibt es ein $n \in \mathbb{N}$ mit $(1/x)^n > 1/y$. Daraus folgt aber auch $x^n < y$. □

Definition 2.19 (Intervall). *Sei $a, b \in \mathbb{R}$ mit $a < b$. Dann ist*

$$[a,b] := \{x : a \le x \le b\}$$

das abgeschlossene Intervall von a bis b,

$$(a,b] := \{x : a < x \le b\}$$

das (bei a) halboffene Intervall von a bis b,

$$[a,b) := \{x : a \le x < b\}$$

das (bei b) halboffene Intervall von a bis b,

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

das offene Intervall von a bis b.

Weiter bezeichne ∞ das Symbol unendlich *sowie*

$$\begin{aligned}[a,\infty) &:= \{x : x \ge a\},\\ (a,\infty) &:= \{x : x > a\},\\ (-\infty,b] &:= \{x : x \le b\},\\ (-\infty,b) &:= \{x : x < b\}.\end{aligned}$$

Lemma 2.20 (Intervallschachtelung). *Seien $[a_1,b_1] \supseteq [a_2,b_2] \supseteq [a_3,b_3] \supseteq \cdots$ eine absteigende Folge von Intervallen mit $(b_n - a_n)_{n=1,2,\ldots} \in \mathcal{N}$ und*[7] *$c_n \in [a_n,b_n], n = 1,2,\ldots$. Dann sind $(a_n)_{n=1,2,\ldots}, (b_n)_{n=1,2,\ldots}, (c_n)_{n=1,2,\ldots}$ Cauchy-Folgen und es gilt $\overline{(a_n)_{n=1,2,\ldots}} = \overline{(b_n)_{n=1,2,\ldots}} = \overline{(c_n)_{n=1,2,\ldots}}$.*

Bemerkung 2.21. Sei $[a_1,b_1] \supseteq [a_2,b_2] \supseteq [a_3,b_3] \supseteq \cdots$ eine Intervallschachtelung, d.h. $(b_n - a_n)_{n=1,2,\ldots} \in \mathcal{N}$. Dann existiert nach obiger Proposition genau ein $x \in \mathbb{R}$ mit $a_m \le x \le b_m, m = 1,2,\ldots$.

Denn: Lemma 2.20 impliziert dass $x := \overline{(a_n)_{n=1,2,\ldots}}$ die einzige reelle Zahl mit $a_m \le x \le b_m, m = 1,2,\ldots$ ist.

[7] Hier bedeutet $c_n \in [a_n,b_n], n = 1,2,\ldots$ genauer: $\forall n = 0,1,2,\ldots : c_n \in [a_n,b_n]$.
Ein nachgestelltes “$n = 0,1,2,\ldots$” bedeutet also implizit, dass die vorangegangene Aussage für alle diese n gilt.

Beweis von Lemma 2.20. Für die erste Behauptung genügt es zu zeigen, dass $(c_n)_{n=1,2,\dots}$ eine Cauchy-Folge ist. Sei hierzu $\varepsilon > 0$ und N groß genug, so dass $b_n - a_n < \varepsilon$ für $n > N$. Dann gilt $|c_m - c_n| \leq (b_n - a_m) \vee (b_m - a_n) \leq b_n - a_n < \varepsilon$ für $m > n$. Damit ist $(c_n)_{0,1,2,\dots}$ eine Cauchy-Folge und $b_n - c_n \leq b_n - a_n < \varepsilon$. Damit ist auch $\overline{(b_n)_{n=1,2,\dots}} = \overline{(c_n)_{n=1,2,\dots}}$ gezeigt. Analog folgt $\overline{(a_n)_{n=1,2,\dots}} = \overline{(c_n)_{n=1,2,\dots}}$. □

Proposition 2.22 (Vollständigkeit von $\mathbb{R}$). *Sei $M \subseteq \mathbb{R}$ eine nach oben beschränkte Teilmenge von $\mathbb{R}$ (Dies ist nach Lemma 2.17.4 genau dann der Fall, wenn es ein $N \in \mathbb{N}$ gibt mit $x \leq N$ für alle $x \in M$.) Dann gibt es eine kleinste obere Schranke – genannt* $\sup M$ *oder* Supremum *– von M mit* $\sup M \in \mathbb{R}$.

Bemerkung 2.23 (Supremum, Infimum, Maximum und Minimum).

1. Betrachten wir die Menge $M = (-\infty, 0)$. Diese ist nach oben beschränkt, da etwa $x \leq 1$ für alle $x \in M$ gilt. Weiter ist $\sup M = 0$, da (i) $x \leq 0$ für alle $x \in M$ sowie (ii) $(x \leq d$ für alle $x \in M) \Rightarrow (d \geq 0)$ gilt.

2. Analog zu $\sup M$ kann man auch das *Infimum* einer nach unten beschränkten Menge angeben. Ist M nach unten beschränkt, so ist $-M$ nach oben beschränkt und wir setzen
$$\inf M := -\sup(-M).$$

3. Für $M \subseteq \mathbb{R}$ definieren wir, falls existent, $\min M$ und $\max M$ genau wie in (2.1). Für eine nach oben beschränkte Menge M ist entscheidend, dass immer $\max M \in M$ gelten muss. Insbesondere existiert $\max(-\infty, 0)$ nicht (denn es gibt kein $x < 0$ das $\forall y < 0 : y \leq x$ erfüllt). Jedoch gibt es – siehe Proposition 2.22 – für nach oben beschränkte Mengen immer $\sup M \in \mathbb{R}$, welches die kleinste obere Schranke von M darstellt. Analog verhält es sich mit $\min M$ und $\inf M$ für nach unten beschränkte Mengen M.

Beweis von Proposition 2.22. Sei $a \in \mathbb{Q}$ keine obere Schranke und $b \in \mathbb{Q}$ eine obere Schranke von M. (Um a zu finden, betrachte etwa ein $x \in M$ und eine wachsende Cauchy-Folge in $\mathbb{Q}$, die gegen x wächst. Dann wählt man a als das erste Element dieser Folge. Um b zu finden, erinnern wir daran, dass es ein $n \in \mathbb{N}$ gibt mit $x \leq n$ für alle $x \in M$. Wir können also $b := n$ wählen.) Wir verwenden nun das Prinzip der Intervallschachtelung. Definiere hierzu $[a_1, b_1] := [a, b]$ und $m := (a+b)/2$. Ist m eine obere Schranke, so setzen wir $[a_2, b_2] := [a_1, m]$; ist m keine obere Schranke, so setzen wir $[a_2, b_2] := [m, b_1]$. Dies tun wir, damit die obere Intervallgrenze immer eine obere Schranke ist, die untere jedoch nicht. Wir fahren immer weiter fort und konstruieren so die Intervalle $[a_1, b_1] \supseteq [a_2, b_2] \supseteq [a_3, b_3] \supseteq \cdots$. Es gilt nach Konstruktion $b_n - a_n = (b_{n-1} - a_{n-1})/2, n = 1, 2, \dots$, also auch $b_n - a_n = (b - a)2^{-n}$, d.h. $(b_n - a_n)_{n=1,2,\dots} \in \mathcal{N}$.

Wählen wir $c_n \in [a_n, b_n], n = 1, 2, \dots$, so ist $c := \overline{(c_n)_{n=1,2,\dots}}$ nach Lemma 2.20 eine Cauchy-Folge. (Man beachte, dass ebenfalls $c := \overline{(a_n)_{n=1,2,\dots}} = \overline{(b_n)_{n=1,2,\dots}}$ gilt.) Weiter behaupten wir, dass c kleinste obere Schranke von M ist. Hierzu müssen wir zeigen, dass (i) c obere Schranke ist und (ii) c kleiner oder gleich allen weiteren oberen Schranken von M ist.

Für (i) sei $x \in M$. Wir wählen eine aufsteigende Cauchy-Folge $(x_n)_{n=1,2,\dots}$ mit $x = \overline{(x_n)_{n=1,2,\dots}}$. Dann ist $(b_n - x_n)_{n=1,2,\dots} \in \mathcal{N} \cup \mathcal{P}$ nach Definition 2.16, also $x \leq c$. Damit ist c obere Schranke von M, da $x \in M$ beliebig war. Für (ii) sei d eine weitere obere Schranke von M. Angenommen, es wäre $c > d$. Ohne Einschränkung ist $d \in \mathbb{Q}$ wegen Lemma 2.18.1. Nach Konstruktion ist $a_n \leq d, n = 1, 2, \dots$ (andernfalls wäre d keine obere Schranke) und damit gilt $c = \overline{(a_n)_{n=1,2,...}} \leq d$ im Widerspruch zu $c > d$. Also haben wir $c \leq d$ gezeigt, was in (ii) behauptet war. □

2.3 Darstellung reeller Zahlen

Von früher ist bereits bekannt, dass reelle Zahlen in Dezimalschreibweise dargestellt werden können, also z.B.

$$\pi = 3.14159265358979323846264338327950288419716939937510582097494459230 7... \tag{2.5}$$

Diese Art der Darstellung wollen wir nun für die im letzten Abschnitt definierten rellen Zahlen angeben. Hierzu benötigen wir eine auch im folgenden wichtige Funktion.

Definition 2.24. *Wir definieren die* Abrundungsfunktion

$$[.] : \begin{cases} \mathbb{R} & \to \mathbb{Z} \\ x & \mapsto \max\{n \in \mathbb{Z} : n \leq x\}. \end{cases}$$

Weiter sei $\lfloor . \rfloor := [.]$ *sowie*

$$\lceil . \rceil : \begin{cases} \mathbb{R} & \to \mathbb{Z} \\ x & \mapsto \min\{n \in \mathbb{Z} : n \geq x\}. \end{cases}$$

Damit ist $\lfloor x \rfloor = [x] = \lceil x \rceil$ *für* $x \in \mathbb{Z}$ *und* $\lfloor x \rfloor = [x] = \lceil x \rceil - 1$ *für* $x \in \mathbb{R} \setminus \mathbb{Z}$.

Proposition 2.25 (Dezimaldarstellung von $x \in [0, 10)$). *Für* $x \in [0, 10)$ *definieren wir* $n_0 := [x]$ *sowie* $n_i := [10^i x] - 10[10^{i-1}x], \qquad i = 1, 2, ...$ *Weiter sei* $(x_n)_{n=0,1,2,...}$ *gegeben durch*

$$x_n := \sum_{i=0}^{n} n_i 10^{-i}. \tag{2.6}$$

Dann ist $(x_n)_{n=1,2,...}$ *eine Cauchy-Folge und es gilt* $x = \overline{(x_n)_{n=1,2,...}}$. *Wir schreiben dann*

$$n_0.n_1n_2n_3...$$

für die Dezimaldarstellung von x.

Beispiel 2.26 (Beispiel einer Dezimaldarstellung). Wir geben als Beispiel die Zahlen $n_0, n_1, n_2, ...$ der Dezimaldarstellung von π. Zunächst ist

$$\begin{aligned} n_0 &= [\pi] = 3, \\ n_1 &= [10\pi - 30] = 1, \\ n_2 &= [100\pi - 310)] = 4, \\ &\dots \end{aligned}$$

In diesem Sinne ist also (2.5) die Dezimalzahldarstellung von π.

Beweis von Proposition 2.25. Klar ist, dass $y - [y] \in [0,1)$ für alle $y \in \mathbb{R}_+$. Deshalb ist $[10^i x] - 10[10^{i-1}x] = [10(10^{i-1} - [10^{i-1}x])] \in \{0,...,9\}$ und $(x_n)_{n=0,1,2,...}$ ist eine wachsende Folge rationaler Zahlen. Weiter ist für $m \geq n$

$$\begin{aligned} |x_m - x_n| &= \sum_{i=n+1}^{m} n_i 10^{-i} \leq 9 \sum_{i=n+1}^{m} 10^{-i} \\ &= 9 \cdot 10^{-n+1} \sum_{i=0}^{m-n} 10^{-i} \leq 9 \cdot 10^{-n+1} \frac{1}{1-10^{-1}} \\ &= 10^{-n} \end{aligned}$$

nach Proposition 1.38. Da es zu jedem $\varepsilon > 0$ nach Lemma 2.18.4 ein N gibt mit $10^{-N} < \varepsilon$ folgt, dass $(x_n)_{n=0,1,2,...}$ eine Cauchy-Folge ist.

Um zu zeigen dass $x = \overline{(x_n)_{n=0,1,2,...}}$, zeigen wir $x_n = 10^{-n}[10^n x]$). Diese Aussage ist für $n = 0$ klar. Gilt sie für ein n, so folgern wir

$$\begin{aligned} x_{n+1} &= x_{n+1} - x_n + x_n = ([10^{n+1}x] - 10[10^n x])10^{-(n+1)} + 10^{-n}[10^n x]) \\ &= 10^{-(n+1)}[10^{n+1}x], \end{aligned}$$

woraus die Aussage nach vollständiger Induktion für alle $n = 0,1,2,...$ folgt. Damit ist auch $x - x_n = 10^{-n}(10^n x - [10^n x]) \leq 10^{-n}$, also $(x - x_n)_{n=0,1,2,...} \in \mathcal{N}$ und somit $x = \overline{(x_n)_{n=0,1,2,...}}$. □

Bemerkung 2.27 (Was ist $0.999\cdots$?).

1. Seien $n_0, n_1, ...$ die Koeffizienten der Dezimaldarstellung einer Zahl $x \in [0,10)$. Es stellt sich die Frage, ob alle Möglichkeiten für diese Koeffizienten vorkommen können. Sei also etwa $n_0 = 0, n_1 = n_2 = n_3 = \cdots = 9$. Ist dann $0,999\cdots$ die durch Proposition 2.25 gewonnene Darstellung einer Zahl $x \in [0,10)$?

   Nein! Um dies zu verstehen betrachten wir die Cauchy-Folge $(x_n)_{n=1,2,...}$ mit $x_n = 1 - 10^{-n}$, was genau der Definition von x_n in (2.6) für $n_0 = 0, n_1 = n_2 = n_3 = \cdots = 9$ entspricht. Klar ist, dass $(1 - x_n)_{n=1,2,...} \in \mathcal{N}$ ist, woraus folgt, dass $1 = \overline{(x_n)_{n=1,2,...}}$. Wäre nun $0.999\cdots$ die Dezimaldarstellung einer Zahl $x \in [0,10)$, so haben wir gesehen dass $x = 1$ gelten muss. Die Darstellung der 1 ist jedoch $n_0 = 1, n_1 = n_2 = n_3 = \cdots = 0$. Insbesondere gibt es kein $x \in [0,10)$ mit der Dezimalzahldarstellung $0.999\cdots$.

2. Es stellt sich nach 1. die Frage, welche Folgen $(x_n)_{n=0,1,2,...}$ aus (2.6) mit $n_0, n_1, ... \in \{0,...,\}$ nicht als Dezimalzahldarstellung eines $x \in [0,10)$ auftreten können. Hier übelegt man sich, dass genau die Folgen

   $$M := \{(n_i)_{i=0,1,2,...} : \exists N\ \forall k > N : n_k = 9\}$$

   nicht vorkommen können.

Bemerkung 2.28 (Erweiterungen).

1. Der Einfachheit halber haben wir in Proposition 2.25 die Darstellung im Dezimalsystem nur für $x \in [0,10)$ angegeben. Natürlich ist klar, dass es auch für beliebige $x \in \mathbb{R}$ eine solche Darstellung gibt. Der Unterschied ist nur, dass die Zahl vor dem Komma im Allgemeinen aus mehr als einer Stelle besteht.

2. Man kann jedes $x \in \mathbb{R}$ auch zu jeder beliebigen anderen *Basis* (die nicht unbedingt 10 sein muss) darstellen. Hierzu muss man lediglich eine Basis $a \in \{2,3,...\}$ wählen

und in Proposition 2.25 jedes Vorkommen von 10 durch a ersetzen (und eventuell die Erweiterung auf alle reelle Zahlen aus 1. beachten). Wählt man etwa $a = 2$, so ist die *Binärdarstellung* von π gegeben durch

$$\pi = 11.0010010000111111011010101000100010000101101001100000...$$

2.4 Überarzählbarkeit von $\mathbb{R}$

Nach Definition ist klar, dass $\mathbb{N}$ abzählbar ist. Weiter wissen wir aus Lemma 1.30, dass $\mathbb{Q}$ abzählbar ist und aus Lemma 1.31, dass $2^{\mathbb{N}}$ überabzählbar ist. Wie steht es aber nun mit der Menge der reellen Zahlen, $\mathbb{R}$. Aus Lemma 2.18.1 wissen wir, dass zwischen zwei verschiedenen reellen Zahlen eine rationale liegt. (Genauso liegt natürlich zwischen zwei verschiedenen rationalen Zahlen eine reelle Zahl.) Dies legt nahe, dass die Menge der reellen Zahlen ebenfalls abzählbar ist. Dies ist jedoch nicht der Fall, wie wir nun zeigen werden. Weiter werden wir in Proposition 2.30 zeigen, dass $\mathbb{R}$ und $2^{\mathbb{N}}$ gleichmächtig sind.

Proposition 2.29. *Die Menge der reellen Zahlen $\mathbb{R}$ ist überabzählbar.*

Beweis. Es genügt zu zeigen dass $[0, 1]$ überabzählbar ist. Angenommen, $[0, 1]$ ist abzählbar. Dann ist $[0, 1] = \{x_1, x_2, ...\}$ für eine geeignete Folge $(x_n)_{n=1,2,...}$. Wir konstruieren nun ein $x \in [0, 1]$ mit $x \neq x_n, n = 1, 2, ...$ Dies ist also im Widerspruch zu $[0, 1] = \{x_1, x_2, ...\}$, woraus wir folgern dass die Abzählbarkeit von $[0, 1]$ nicht stimmen kann.

Um $x \in [0, 1]$ mit $x \neq x_n, n = 1, 2, ...$ zu konstruieren, definieren wir eine Intervallschachtelung. Hierzu sei $I_1 = [0, 1/3]$ oder $I_1 = [2/3, 1]$, so dass $x_1 \notin I_1$. Rekursiv definieren wir $I_{n+1} = [a_{n+1}, b_{n+1}] \subseteq I_n = [a_n, b_n]$ mittels $I_{n+1} = [a_n, 2a_n/3 + b_n/3]$ oder $I_{n+1} = [a_n/3 + 2b_n/3, b_n]$, so dass $x_{n+1} \notin I_{n+1}$. (In Worten: wir dritteln das Intervall I_n und wählen entweder das erste oder das dritte Drittel als I_{n+1}, je nachdem wann $x_{n+1} \notin I_{n+1}$ gilt.) Auf diese Art und Weise erreichen wir eine Intervallschachtelung $I_1 \supseteq I_2 \supseteq I_3 \cdots$. Nach Bemerkung 2.21 gibt es genau ein $x \in [0, 1]$ mit $x \in I_n, n = 1, 2, ...$ Allerdings gilt auch $x_n \notin I_n$, also auch $x \neq x_n, n = 1, 2, ...$ Dies ist also der gewünschte Widerspruch zur Abzählbarkeit von $[0, 1]$. $\square$

Proposition 2.30 (Zu $\mathbb{R}$ gleichmächtige Mengen).

1. *Die Mengen $\mathbb{R}$, $(0, 1)$, $[0, 1)$, $(0, 1]$ und $[0, 1]$ sind gleichmächtig.*
2. *Die Mengen $\mathbb{R}$ und $2^{\mathbb{N}}$ sind gleichmächtig.*

Bemerkung 2.31 (Ein falscher Beweis von 2.). Folgende Überlegung stellt *fast* einen Beweis von 2. dar:

Beweisansatz von 2. Nach 1. genügt es zu zeigen, dass $[0, 1)$ und $2^{\mathbb{N}}$ gleichmächtig sind. Hierzu betrachten wir für $x \in (0, 1)$ die Binärdarstellung, also $x_n = \sum_{i=1}^{n} n_i(x)2^{-i}$ für $n_i(x) = [2^i x] - 2[2^{i-1}x]$. Weiter definieren wir die Abbildung

$$h : \begin{cases} [0, 1) & \to 2^{\mathbb{N}}, \\ x & \mapsto \{i : n_i(x) = 1\}. \end{cases}$$

Diese ist sicher injektiv (da zwei verschiedene Zahlen auch unterschiedliche Binärdarstellungen haben) und surjektiv (da jede mögliche Binärdarstellung auch vorkommt). Damit sind $[0, 1)$ und $2^{\mathbb{N}}$ gleichmächtig. $\square$

Leider stimmt es aber nicht, dass jede mögliche Zahlenfolge $(n_i)_{i=1,2,...}$ in $\{0,1\}$ auch die Binärdarstellung einer Zahl $x \in [0,1)$ liefert. In Bemerkung 2.27.2 haben wir nämlich festgestellt, dass genau die Folgen in der Menge

$$M := \{(n_i)_{i=1,2,...} : \exists N \; \forall k > N : n_k = 1\} \tag{2.7}$$

nicht vorkommen können.

Beweis von Proposition 2.30. 1. Wir beginnen mit der Gleichmächtigkeit von $\mathbb{R}$ und $(0,1)$. Hierzu definieren wir die Abbildung[8]

$$f : \begin{cases} \mathbb{R} & \to (0,1) \\ x & \mapsto \begin{cases} \frac{1}{2}e^x, & x \leq 0, \\ 1 - \frac{1}{2}e^{-x}, & x > 0. \end{cases} \end{cases}$$

Diese ist offenbar eine Bijektion, und damit sind $\mathbb{R}$ und $(0,1)$ gleichmächtig.
Um die Gleichmächtigkeit von $(0,1)$ und $[0,1)$ zu zeigen, definieren wir

$$g : \begin{cases} (0,1) & \to [0,1), \\ x & \mapsto \begin{cases} 0, & x = \frac{1}{2}, \\ 2^{-(n+1)}, & x = 2^{-n} \text{ für ein } n = 1,2,..., \\ x, & \text{sonst.} \end{cases} \end{cases}$$

Wir stellen fest, dass g eine Bijektion ist, was die Behauptung zeigt. Die Gleichmächtigkeit von $[0,1)$ und $(0,1]$ ist klar, da $x \mapsto 1-x$ eine Bijektion zwischen den beiden Mengen darstellt. Die Gleichmächtigkeit von $(0,1]$ und $[0,1]$ zeigt man, indem man g mittels $g(1) = 1$ zu einer Funktion $(0,1] \to [0,1]$ erweitert.
2. Wir verwenden die Beweisidee aus Bemerkung 2.31. Insbesondere haben wir dort bereits gesehen, dass es eine Bijektion $h : [0,1) \to 2^{\mathbb{N}} \setminus M$ mit M aus (2.7) gibt. Wichtig ist es zu bemerken dass die Menge M abzählbar ist. Da es unendlich viele Primzahlen gibt, ist die Menge der Primzahlen $\{2,3,5,...\}$ abzählbar. Es gibt also eine Bijektion $\{2,3,5...\} \to M$. Mit Hilfe dieser Bijektion erhalten wir nun auch eine Bijektion $[0,1) \cup \{2,3,...\} \to 2^{\mathbb{N}}$. Es bleibt also zu zeigen, dass es eine Bijektion $[0,1) \cup \{2,3,...\} \to [0,1)$ gibt. Diese lässt sich aber (mit Hilfe desselben Tricks wie in der Definition von g) durch

$$\begin{cases} [0,1) \cup \{2,3,5,...\} & \to [0,1), \\ x & \mapsto \begin{cases} \frac{1}{p}, & \text{falls } x = p \in \{2,3,5,...\}, \\ p^{-(j+1)}, & \text{falls } x = p^{-j} \text{ für ein } p \in \{2,3,5,...\} \text{ und } j = 1,2,..., \\ x, & \text{sonst.} \end{cases} \end{cases}$$

definieren. (Wir bemerken, dass wir in dieser Definition deswegen auf die Menge der Primzahlen zurückgreifen, weil $x = p^{-j}$ nur für höchstens eine Primzahl p für ein j gilt.)

□

[8] Wir setzen hier die Kenntnis der Exponentialfunktion voraus. Wir werden die genaue Definition später nachliefern.

Bemerkung 2.32 (Kontinuumshypothese). Wir haben nun gelernt, dass $(0,1)$ und (damit auch alle Intervalle) dieselbe Mächtigkeit wie $\mathbb{R}$ besitzen. Man kann sich fragen, was denn die kleinsten, überabzählbaren Mengen sind. Insbesondere ist momentan noch nicht klar, ob es überabzählbare Mengen gibt, die weniger mächtig wie $\mathbb{R}$ oder $2^{\mathbb{N}}$ sind. Georg Cantor formulierte die folgende Kontinuumshypothese:

> CH: Es gibt keine überabzählbare Menge, deren Mächtigkeit kleiner als die der reellen Zahlen ist.

Erstaunlicherweise handelt es sich bei dieser Hypothese (im Rahmen der heute verwendeten Mengenlehre) um eine unentscheidbare Aussage. Man hat nämlich zeigen können, dass die Hypothese CH weder bewiesen noch widerlegt werden kann.

3 Folgen reeller Zahlen

Folgen rationaler Zahlen sind uns im letzten Kapitel oft begegnet. Wir erweitern unsere Untersuchungen auf Folgen reeller Zahlen. Dabei richten wir unser Augenmerk vor allem auf die Konvergenz von Folgen. Zentral ist hierbei der Begriff des Grenzwertes, den wir in Abschnitt 3.1 beleuchten werden. Anschließend werden wir in Abschnitt 3.2 wichtige Rechenregeln für Grenzwerte kennen lernen und uns mit Konvergenz von Teilfolgen in Abschnitt 3.3 beschäftigen. Die uneigentliche Konvergenz, d.h. die Konvergenz gegen $\pm\infty$, ist abschließend das Thema von Abschnitt 3.4.

3.1 Konvergenz von Folgen

Wir haben die reellen Zahlen als Äquivalenzklassen von Cauchy-Folgen kennen gelernt. In diesem Sinn ist jede reelle Zahl der Grenzwert einer Cauchy-Folge in $\mathbb{Q}$. Dies verallgemeinern wir nun auf Folgen reeller Zahlen.

Definition 3.1 (Konvergenz, Limes). *Sei $(x_n)_{n=1,2,\ldots}$ eine Folge reeller Zahlen. Sie heißt* konvergent*, wenn es eine Zahl $x \in \mathbb{R}$ gibt, so dass gilt: Für alle $\varepsilon > 0$ gilt $|x_n - x| < \varepsilon$ für fast alle n. Anders ausgedrückt gilt*

$$\exists x \in \mathbb{R}\ \forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n > N : |x_n - x| < \varepsilon.$$

In diesem Fall heißt x der Grenzwert *oder* Limes *der Folge $(x_n)_{n=1,2,\ldots}$. Wir schreiben dann*

$$\lim_{n\to\infty} x_n = x \qquad \text{oder} \qquad x_n \xrightarrow{n\to\infty} x.$$

Eine gegen 0 konvergente Folge heißt auch Nullfolge*. Ist die Folge nicht konvergent, so heißt sie* divergent*.*

Lemma 3.2 (Eindeutigkeit des Limes). *Sei $(x_n)_{n=1,2,\ldots}$ eine Folge reeller Zahlen und $x, y \in \mathbb{R}$, so dass $x_n \xrightarrow{n\to\infty} x$ sowie $x_n \xrightarrow{n\to\infty} y$. Dann ist $x = y$.*

Bemerkung 3.3 (Interpretation). Das Lemma bedeutet, dass der Begriff des Grenzwertes wohldefiniert ist. (Wäre nämlich $\lim_{n\to\infty} x_n = x$ sowie $\lim_{n\to\infty} x_n = y$ mit $x \neq y$, dann wäre unklar, welche reelle Zahl man mit $\lim_{n\to\infty} x_n$ genau bezeichnet.)

Beweis. Angenommen, es wäre $x \neq y$. Wir setzen $\varepsilon = |y - x|/2 > 0$. Dann gibt es nach Definition der Konvergenz ein $N \in \mathbb{N}$, so dass für alle $n > N$ gilt, dass $|x_n - x| < \varepsilon$ sowie $|x_n - y| < \varepsilon$. Daraus folgt $2\varepsilon = |y - x| = |y - x_n + x_n - x| \leq |x_n - y| + |x_n - x| < 2\varepsilon$ und damit ein Widerspruch. □

Bemerkung 3.4 (Notation, $\sup, \inf$). In Proposition 2.22 haben wir bereits für eine nach oben beschränkte Menge $A \subseteq \mathbb{R}$ die Schreibweise $\sup A$ für die kleinste obere Schranke von A kennen gelernt. Dies erweitern wir auf nach oben unbeschränkte Mengen, indem wir $\sup A = \infty$ setzen, falls A nicht nach oben beschränkt ist. Weiter ist $\inf A = -\infty$, falls A nicht nach unten beschränkt ist. Manchmal benötigt man noch $\sup \emptyset = -\infty$ und $\inf \emptyset = \infty$. Ist $A = \{f(m) : m \in I\}$ für eine Menge I und eine Funktion f, so schreiben wir manchmal auch $\sup_{m \in I} f(m) := \sup A$ und $\inf_{m \in I} f(m) := \inf A$. (Insbesondere ist für eine Folge $(x_n)_{n=1,2,...}$ reeller Zahlen $\sup_{n \in \mathbb{N}} x_n := \sup\{x_1, x_2, ...\}$.)

Proposition 3.5 (Konvergenz monotoner Folgen). *Jede beschränkte monotone Folge $(x_n)_{n=1,2,...}$ konvergiert. Ist $(x_n)_{n=1,2,...}$ wachsend, so konvergiert sie gegen $\sup_{n \in \mathbb{N}} x_n$, ist sie fallend, dann gegen $\inf_{n \in \mathbb{N}} x_n$.*

Beweis. Es genügt die Behauptung für wachsendes $(x_n)_{n=1,2,...}$ zu zeigen. Der Fall fallender $(x_n)_{n=1,2,...}$ folgt daraus durch Übergang zu der Folge $(-x_n)_{n=1,2,...}$ Sei also $(x_n)_{n=1,2,...}$ wachsend und $x = \sup_{n \in \mathbb{N}} x_n$ sowie $\varepsilon > 0$. Da x die kleinste obere Schranke von $\{x_1, x_2, ...\}$ ist, gibt es ein $N \in \mathbb{N}$ mit $|x - x_N| < \varepsilon$. Wegen der Monotonie von $(x_n)_{n=1,2,...}$ gilt damit auch $|x - x_n| < \varepsilon$ für alle $n > N$. □

Theorem 3.6 (Wichtige Grenzwerte).

1. *Ist $s > 0$, so ist $\lim_{n \to \infty} n^{-s} = 0$.*
2. *Ist $x > 0$, so gilt $\lim_{n \to \infty} x^{1/n} = 1$.*
3. *Es gilt $\lim_{n \to \infty} n^{1/n} = 1$.*
4. *Für $|x| < 1$ ist $\lim_{n \to \infty} x^n = 0$.*
5. *Sei $k \in \mathbb{N}$ und $|x| > 1$, so gilt $\lim_{n \to \infty} \frac{n^k}{x^n} = 0$.*

Beweis. 1. Sei $\varepsilon > 0$. Dann gilt für $n > N := \lceil \varepsilon^{-1/s} \rceil$, dass $0 < n^{-s} < N^{-s} \leq \varepsilon$.
2. Sei $x_n := x^{1/n} - 1$. Dann gilt mit Proposition 1.39, dass $x = (1 + x_n)^n \geq 1 + nx_n > nx_n$, also auch $x_n < x/n$. Für $\varepsilon > 0$ und $n > N := x/\varepsilon$ folgt daraus, dass $x^{1/n} - 1 = x_n < \varepsilon$.
3. Sei $x_n := n^{1/n} - 1 \geq 0$. Mit Proposition 1.40 gilt

$$n = (1 + x_n)^n \geq 1 + \binom{n}{2} x_n^2,$$

also $n - 1 \geq n(n-1)x_n^2/2$, d.h. $x_n \leq \sqrt{2/n}$. Für $\varepsilon > 0$ und $n > N := 2/\varepsilon^2$ gilt damit $|n^{1/n} - 1| = x_n < \varepsilon$.
4. ist klar nach Lemma 2.18.4.
5. OBdA ist $x > 1$. Sei $z := x - 1 > 0$. Die Binomische Formel, Proposition 1.40, ergibt für $n > 2k$

$$(1 + z)^n > \binom{n}{k+1} z^{k+1} = \frac{n(n-1)\cdots(n-k)}{(k+1)!} z^{k+1} > \frac{n^{k+1} z^{k+1}}{2^{k+1}(k+1)!}$$

und damit

$$\frac{n^k}{x^n} < \frac{2^{k+1}(k+1)!}{z^{k+1}} \cdot \frac{1}{n}.$$

Sei $\varepsilon > 0$. Dann gilt für $n > N := \max\{2^{k+1}(k+1)!/(\varepsilon z^{k+1}), 2k\}$, dass $|n^k/x^n| < \varepsilon$. $\square$

3.2 Rechenregeln

Oftmals geht es in der Mathematik um das Vertauschen von Operationen. In diesem Abschnitt beschäftigt uns beispielsweise die Vertauschung von Grenzwertbildung und den Grundrechenarten.

Proposition 3.7 (Grundrechenarten von Grenzwerten). *Seien $(x_n)_{n=1,2,\dots}$ und $(y_n)_{n=1,2,\dots}$ reellwertige, konvergente Folgen mit $\lim_{n\to\infty} x_n = x, \lim_{n\to\infty} y_n = y$. Dann gilt:*

1. $\lim_{n\to\infty} |x_n| = |x|$,
2. $\lim_{n\to\infty} x_n + y_n = x + y$,
3. $\lim_{n\to\infty} x_n \cdot y_n = x \cdot y$,
4. $\lim_{n\to\infty} x_n/y_n = x/y$, *falls* $y \neq 0$. *In diesem Fall ist* $y_n \neq 0$ *für fast alle* n.

Beweis. 1. Dies folgt sofort aus $||x_n| - |x|| \leq |x_n - x|$.
2. Für $\varepsilon > 0$ gibt es ein $N \in \mathbb{N}$, so dass $|x_n - x| < \varepsilon/2$ und $|y_n - y| < \varepsilon/2$ für $n > N$. Damit gilt

$$|x_n + y_n - (x+y)| \leq |x_n - x| + |y_n - y| < \varepsilon.$$

3. Wie im Beweis von Lemma 2.3 ist $(x_n)_{n=1,2,\dots}$ beschränkt, also etwa $|x_n| \leq c$ für alle $n = 1, 2, \dots$ Sei $\varepsilon > 0$ und N so groß, dass $|x_n - x| < \varepsilon/(2|y|)$ und $|y_n - y| < \varepsilon/(2c)$ für alle $n > N$. Dann gilt

$$|x_n y_n - xy| = |x_n(y_n - y) + y(x_n - x)| \leq c|y_n - y| + |y| \cdot |x_n - x| < \varepsilon.$$

4. Wir zeigen zunächst, dass $1/y_n \xrightarrow{n\to\infty} 1/y$. Aus 3. folgt dann die Behauptung. Hierzu gehen wir wie im Beweis von Proposition 2.12 vor. Zunächst wählen wir $\delta > 0$ und $N' \in \mathbb{N}$, so dass $y_n > \delta$ für alle $n > N'$. Weiter sei $N \in \mathbb{N}$ so, dass $|y_n - y| < \delta^2\varepsilon$ für $n > N$ gilt. Damit gilt für $n > N \vee N'$

$$\Big|\frac{1}{y_n} - \frac{1}{y}\Big| = \frac{|y_n - y|}{|y_n y|} < \frac{\delta^2\varepsilon}{\delta^2} = \varepsilon.$$

$\square$

Beispiel 3.8. Sei $a_0, \dots, a_k, b_0, \dots, b_k \in \mathbb{R}$ und $b_k \neq 0$. Dann ist

$$\lim_{n\to\infty} \frac{a_k n^k + a_{k-1}n^{k-1} + \cdots + a_0}{b_k n^k + b_{k-1}n^{k-1} + \cdots + b_0} = \frac{a_k}{b_k}.$$

Denn: Nach Theorem 3.6.1 ist $a_\ell n^{\ell-k}, b_\ell n^{\ell-k} \xrightarrow{n\to\infty} 0$ für $k < \ell$. Mit Proposition 3.7.2 folgt, dass $a_k + a_{k-1}n^{-1} + \cdots + a_0 n^{-k} \xrightarrow{n\to\infty} a_k$. Wegen Proposition 3.7.4 folgt daraus

$$\lim_{n\to\infty} \frac{a_k n^k + a_{k-1}n^{k-1} + \cdots + a_0}{b_k n^k + b_{k-1}n^{k-1} + \cdots + b_0} = \lim_{n\to\infty} \frac{a_k + a_{k-1}n^{-1} + \cdots + a_0 n^{-k}}{b_k + b_{k-1}n^{-1} + \cdots + b_0 n^{-k}} = \frac{a_k}{b_k}.$$

Proposition 3.9 (Ordnung von Grenzwerten). *Seien* $(x_n)_{n=1,2,...}$, $(y_n)_{n=1,2,...}$, $(z_n)_{n=1,2,...}$ *reellwertige, konvergente Folgen mit* $\lim_{n\to\infty} x_n = x, \lim_{n\to\infty} y_n = y, \lim_{n\to\infty} z_n = z$.

1. *Gilt* $x_n \leq y_n$ *für fast alle* n*, so gilt* $x \leq y$.
2. *Liegen fast alle* x_n *in einem Intervall* $[a,b]$*, so ist* $x \in [a,b]$.
3. *Gilt* $x_n \leq y_n \leq z_n$ *für fast alle* n *und* $x = z$*, so ist auch* $x = y = z$.

Beweis. 1. Für $\varepsilon > 0$ gibt es ein $N \in \mathbb{N}$, so dass $x - \varepsilon < x_n \leq y_n < y + \varepsilon$ für $n > N$. Hieraus folgt $x - y < 2\varepsilon$. Da $\varepsilon > 0$ beliebig war, folgt $x - y \leq 0$.
2. Setzt man $y_n = b, n = 1, 2, ...$, dann folgt $x \leq b$. Analog folgert man $x \geq a$, woraus die Behauptung folgt.
3. Nach 1. ist $y \leq z$ und $y \geq x$. Daraus folgt die Behauptung. □

3.3 Häufungspunkte von Folgen

Geht man von Folgen zu Teilfolgen über, so können sich viele mögliche Grenzwerte ergeben. Diese nennt man dann Häufungspunkte der Folge, die nun definiert werden.

Definition 3.10 (Teilfolge, Häufungspunkt). *Sei* $(x_n)_{n=1,2,...}$ *eine Folge reeller Zahlen.*

1. *Ist* $(k_n)_{n=1,2,...}$ *eine wachsende Folge natürlicher Zahlen, so heißt* $(x_{k_n})_{n=1,2,...}$ *eine* Teilfolge *von* $(x_n)_{n=1,2,...}$.
2. *Ein* $x \in \mathbb{R}$ *heißt* Häufungspunkt *von* $(x_n)_{n=1,2,...}$*, falls es eine Teilfolge* $(x_{k_n})_{n=1,2,...}$ *gibt mit* $\lim_{n\to\infty} x_{k_n} = x$.

Beispiel 3.11.
1. Sei $x_n = (-1)^n$, d.h. $x_1 = -1, x_2 = 1, x_3 = -1, x_4 = 1, ...$ Dann hat $(x_n)_{n=1,2,...}$ genau die Häufungspunkte 1 und -1. Schließlich ist $\lim_{n\to\infty} x_{2n} = 1$ und $\lim_{n\to\infty} x_{2n+1} = -1$.
2. Die Folge $(x_n)_{n=1,2,...}$ mit $x_n = n$ hat keinen Häufungspunkt.

Lemma 3.12 (Charakterisierung von Häufungspunkten). *Sei* $(x_n)_{n=1,2,...}$ *eine Folge reeller Zahlen und* $x \in \mathbb{R}$*. Dann sind äquivalent:*

1. *Es ist* x *ein Häufungspunkt von* $(x_n)_{n=1,2,...}$
2. *Für jedes* $\varepsilon > 0$ *gilt* $|x_n - x| < \varepsilon$ *für unendlich viele* n.

Beweis. 1.⇒2.: Sei $(x_{k_n})_{n=1,2,...}$ eine gegen x konvergente Teilfolge und $\varepsilon > 0$. Nach Definition der Konvergenz gilt dann $|x_{k_n} - x| < \varepsilon$ für fast alle (und damit für unendlich viele) n. Da $(x_{k_n})_{n=1,2,...}$ eine Teilfolge ist, heißt dies insbesondere, dass 2. gilt.
2.⇒1.: Sei $(\varepsilon_n)_{n=1,2,...}$ eine fallende Nullfolge[9]. Wir definieren

$$k_1 := \min\{\ell : |x_\ell - x| < \varepsilon_1\}$$

und rekursiv

$$k_{n+1} := \min\{\ell > k_n : |x_\ell - x| < \varepsilon_n\}.$$

Sei nun $\varepsilon > 0$. Da $(\varepsilon_n)_{n=1,2,...}$ eine Nullfolge ist, gibt es nun ein $N \in \mathbb{N}$ mit $\varepsilon_n < \varepsilon$ für alle $n > N$. Daraus folgt für solche n, dass $|x_{k_n} - x| < \varepsilon_n < \varepsilon$. Daraus folgt, dass $\lim_{n\to\infty} x_{k_n} = x$ und die Behauptung ist gezeigt. □

[9] Hierfür schreibt man oft auch $\varepsilon_n \downarrow 0$.

Theorem 3.13 (Satz von Bolzano-Weierstrass). *Jede beschränkte Folge reeller Zahlen* $(x_n)_{n=1,2,\ldots}$ *besitzt einen Häufungspunkt. Für den größten (kleinsten) Häufungspunkt* x^* (x_*) *und* $\varepsilon > 0$ *gilt, dass* $x_n < x^* + \varepsilon$ $(x_n > x_* - \varepsilon)$ *für fast alle* n. *Wir setzen den* Limes superior *und* Limes inferior

$$\limsup_{n\to\infty} x_n := x^*, \qquad \liminf_{n\to\infty} x_n := x_*.$$

Es gilt

$$\limsup_{n\to\infty} x_n = \inf_{k\in\mathbb{N}} \sup_{n>k} x_n, \qquad \liminf_{n\to\infty} x_n = \sup_{k\in\mathbb{N}} \inf_{n>k} x_n. \tag{3.1}$$

Beweis. Wir zeigen, dass es einen größten Häufungspunkt der Folge $(x_n)_{n=1,2,\ldots}$ gibt. Hierfür benutzen wir das Prinzip der Intervallschachtelung aus Lemma 2.20. Sei hierzu a_1, b_1, so dass $x_n \in [a_1, b_1], n = 1, 2, \ldots$ (Ein solches Intervall gibt es wegen der Beschränktheit der Folge $(x_n)_{n=1,2,\ldots}$.) Rekursiv definieren wir nun Intervalle $[a_1, b_1] \supseteq [a_2, b_2] \supseteq [a_3, b_3]\ldots$ Sei zunächst $m_1 = (a_1 + b_1)/2$. Gilt $x_n \leq m_1$ für fast alle n, so setzen wir $[a_2, b_2] := [a_1, m_1]$. (In diesem Fall findet sich nämlich offensichtlich keine Teilfolge, deren Grenzwert größer als m_1 ist.) Andernfalls ist $x_n \leq m_1$ für unendlich viele n (womit es eine Teilfolge geben kann, deren Grenzwert größer als m_1 ist) und wir setzen $[a_2, b_2] := [m_1, b_1]$. Rekursiv liefert das eine Folge von Intervallen $[a_1, b_1] \supseteq [a_2, b_2] \supseteq [a_3, b_3]\ldots$, so dass für alle k gilt, dass

(i) $x_n \leq b_k$ für fast alle n,

(ii) $x_n \geq a_k$ für unendlich viele n.

Insbesondere gilt also $x_n \in [a_k, b_k]$ für unendlich viele n, es gibt also eine Teilfolge $(x_{k_n})_{n=1,2,\ldots}$, für die sowohl (i) als auch (ii) für alle $n = 1, 2, \ldots$ gilt. Nach dem Prinzip der Intervallschachtelung gibt es genau eine Zahl $x^* \in \mathbb{R}$, die in allen Intervallen liegt, also $x^* \in [a_k, b_k], k = 1, 2, \ldots$

Sei nun $\varepsilon > 0$. Dann ist $(-\infty, b_k] \subseteq (-\infty, x^* + \varepsilon)$ für fast alle k. Da $x_n \leq b_k$ für fast alle n, gilt auch

$$x_n < x^* + \varepsilon \text{ für fast alle } n. \tag{3.2}$$

Damit ist insbesondere x^* der größte Häufungspunkt. Andernfalls gäbe es nämlich ein $x^{**} > x^*$ und eine Teilfolge $(x_{\ell_n})_{n=1,2,\ldots}$ mit $\lim_{n\to\infty} x_{\ell_n} = x^{**}$. Damit müsste mit $\varepsilon = (x^{**} - x^*)/2$ auch $x_{\ell_n} > x^* + \varepsilon$ für unendlich viele n gelten im Widerspruch zu (3.2).

Es bleibt noch, (3.1) zu zeigen. Klar ist, dass $(\sup_{n>k} x_n)_{k=1,2,\ldots}$ eine monoton nichtwachsende Folge ist. Nun gibt es zwei Fälle: (i) es gibt unendlich viele k mit $x_k \sup_{n\geq k} x_n$ und $x_k > \sup_{n>k} x_n$ oder (ii) $k \mapsto \sup_{n>k} x_n$ ist für alle $k > K$ konstant für ein geeignetes K, d.h. für jedes k gibt es ein $\ell > k$ mit $x_\ell \geq x_k$. Im Fall (i) gibt es eine Teilfolge $(\sup_{n>i_k} x_n)_{k=1,2,\ldots}$ mit $\sup_{n>i_k} x_n = x_{i_k}$, $k = 1, 2, \ldots$. Damit ist $(x_{i_k})_{k=1,2,\ldots}$ eine konvergente Teilfolge und $\inf_{k\in\mathbb{N}} \sup_{n>k} x_n$ ein Häufungspunkt der Folge $(x_n)_{n=1,2,\ldots}$. Außerdem ist klar, dass $x_n \leq \sup_{k\geq n} x_k$ und damit für jedes $\varepsilon > 0$ gilt, dass $x_n < \inf_{n\in\mathbb{N}} \sup_{k\geq n} x_k + \varepsilon$ für fast alle n. Genau wie oben ist damit $\inf_{k\in\mathbb{N}} \sup_{n>k} x_n$ der größte Häufungspunkt von $(x_n)_{n=1,2,\ldots}$ womit (3.1) gezeigt ist. Für (ii) ist $\inf_{n\in\mathbb{N}} \sup_{k\geq n} x_k = \sup_{k\geq K} x_k$. Definieren wir rekursiv $i_1 = K$, $i_{k+1} = \min\{i > i_k : x_i > x_{i_k}\}$, so ist klar, dass $\sup_{k\geq K} x_k$ der größte Häufungspunkt von $(x_n)_{n=1,2,\ldots}$ ist, womit (3.1) gezeigt ist.

Die Aussagen über die kleinsten Häufungspunkte folgen, wenn wir zur Folge $(-x_n)_{n=1,2,\ldots}$ übergehen. $\square$

Korollar 3.14. *Sei $(x_n)_{n=1,2,...}$ eine Folge reeller Zahlen und $x \in \mathbb{R}$. Dann sind äquivalent:*

1. *Es gilt $\lim_{n\to\infty} x_n = x$.*
2. *Es gilt $\liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n = x$.*
3. *Die Folge $(x_n)_{n=1,2,...}$ ist beschränkt und hat genau x als Häufungspunkt.*

Beweis. 1.⇒2.: Für $\varepsilon > 0$ gibt es ein $N \in \mathbb{N}$, so dass $|x_n - x| < \varepsilon$ für $n > N$. Damit gibt es ein $k \in \mathbb{N}$ mit $\inf_{n>k} x_n > x - \varepsilon$, also auch $\liminf_{n\to\infty} x_n \geq x - \varepsilon$. Da $\varepsilon > 0$ beliebig war, muss also $\liminf_{n\to\infty} x_n \geq x$ gelten. Analog folgt $\limsup_{n\to\infty} x_n \leq x$. Also gilt

$$x \leq \liminf_{n\to\infty} x_n \leq \limsup_{n\to\infty} x_n \leq x,$$

woraus 2. folgt.
2.⇒3.: Nach dem Satz von Bolzano-Weierstrass ist der kleinste Häufungspunkt der Folge identisch mit dem größten Häufungspunkt. Deswegen kann es nur einen Häufungspunkt geben.
3.⇒1.: Sei $\varepsilon > 0$. Wenn es genau einen Häufungspunkt gibt, folgt nach Lemma 3.12, dass es nur ein $x \in \mathbb{R}$ gibt, so dass $|x_n - x| < \varepsilon$ für unendlich viele $n \in \mathbb{N}$. Dies bedeutet aber wegen der Beschränktheit der Folge, dass $|x_n - x| < \varepsilon$ für fast alle $n \in \mathbb{N}$. Da $\varepsilon > 0$ beliebig war, folgt 1. □

Proposition 3.15 (Jede Cauchy-Folge in $\mathbb{R}$ konvergiert). *Sei $(x_n)_{n=1,2,...}$ eine Folge reeller Zahlen. Dann sind äquivalent:*

1. *Die Folge $(x_n)_{n=1,2,...}$ ist Cauchy, d.h.*

$$\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n, m > N : |x_n - x_m| < \varepsilon.$$

2. *Es gibt ein $x \in \mathbb{R}$ mit $\lim_{n\to\infty} x_n = x$.*

Beweis. 1.⇒2.: Wir haben bereits in Lemma 2.3 gesehen, dass $(x_n)_{n=1,2,...}$ beschränkt ist. Nach Theorem 3.13 gibt es also eine konvergente Teilfolge, d.h. ein $x \in \mathbb{R}$ und $(x_{k_n})_{n=1,2,...}$, so dass $x_{k_n} \xrightarrow{n\to\infty} x$. Sei nun $\varepsilon > 0$. Dann gibt es ein $N \in \mathbb{N}$, so dass für alle $n > N$ gilt, dass $|x_n - x_{k_n}| < \varepsilon/2$ (da $k_n \geq n$) und $|x_{k_n} - x| < \varepsilon/2$. Daraus folgt nun $|x_n - x| = |x_n - x_{k_n} + x_{k_n} - x| \leq |x_n - x_{k_n}| + |x_{k_n} - x| < \varepsilon/2 + \varepsilon/2 = \varepsilon$ und wir haben $x_n \xrightarrow{n\to\infty} x$ gezeigt.

2.⇒1.: Für $\varepsilon > 0$ gibt es ein $N \in \mathbb{N}$, so dass $|x_n - x| < \varepsilon/2$ für alle $n > N$ gilt. Daraus folgt nun für $m, n > N$, dass $|x_n - x_m| = |x_n - x + x - x_m| \leq |x_n - x| + |x_m - x| < \varepsilon/2 + \varepsilon/2 = \varepsilon$. □

Bemerkung 3.16 (Landau-Symbole). Seien $(x_n)_{n=1,2,...}$ und $(y_n)_{n=1,2,...}$ Folgen reeller Zahlen. Wir schreiben

$$\begin{aligned}
x_n &= \mathcal{O}(y_n), \text{ falls } \limsup_{n\to\infty} \left|\frac{x_n}{y_n}\right| < \infty, \\
x_n &= o(y_n), \text{ falls } \limsup_{n\to\infty} \frac{x_n}{y_n} = 0, \\
x_n &= \Omega(y_n), \text{ falls } 0 < \liminf_{n\to\infty} \left|\frac{x_n}{y_n}\right| \leq \limsup_{n\to\infty} \left|\frac{x_n}{y_n}\right| < \infty.
\end{aligned}$$

Beispielsweise ist $n = o(n^2)$, $n = \Omega(2n)$, sowie $n^k = o(x^n)$ für $x > 1$ nach Theorem 3.6.5.

3.4 Uneigentliche Konvergenz

Wir haben bereits in Definition 2.19 das Symbol ∞ kennen gelernt. Wir setzen nun

$$\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$$

und erinnern daran, dass $x < \infty$ und $x > -\infty$ für jedes $x \in \mathbb{R}$ gilt. Wir untersuchen nun Konvergenz in $\overline{\mathbb{R}}$, die auch uneigentliche Konvergenz heißt, weil ∞ und $-\infty$ als Grenzwerte zugelassen sind.

Definition 3.17 (In $\overline{\mathbb{R}}$ konvergente Folgen). *Eine $\overline{\mathbb{R}}$-wertige Folge $(x_n)_{n=1,2,...}$ konvergiert (in $\overline{\mathbb{R}}$) gegen ∞, wenn*

$$\forall K > 0 \; \exists N \in \mathbb{N} \; \forall n > N : x_n \geq K.$$

In diesem Fall schreiben wir $\lim_{n\to\infty} x_n = \infty$. Die Folge $(x_n)_{n=1,2,...}$ konvergiert (in $\overline{\mathbb{R}}$) gegen $-\infty$, wenn $(-x_n)_{n=1,2,...}$ (in $\overline{\mathbb{R}}$) gegen ∞ konvergiert.

Beispiel 3.18.

1. Die Folge $(x^n)_{n=1,2,...}$ konvergiert für $x > 1$ nach Lemma 2.18.3 gegen ∞.
2. Ist $(x_n)_{n=1,2,...}$ eine Nullfolge positiver Zahlen $x_1, x_2, ...$, so konvergiert $(1/x_n)_{n=1,2,...}$ gegen ∞.

   Denn: Sei $K > 0$. Dann gibt es ein $N \in \mathbb{N}$, so dass $0 \leq x_n < 1/K$ für $n > N$ gilt. Dann gilt aber auch $1/x_n > K$.

Bemerkung 3.19 (Rechenregeln für Konvergenz in $\overline{\mathbb{R}}$). Nicht alle Sätze für Konvergenz gegen $x \in \mathbb{R}$ übertragen sich auf Konvergenz in $\overline{\mathbb{R}}$ gegen ∞ (bzw. $-\infty$). Grundlegend ist immer, dass die Rechenregeln für reelle Zahlen $(+, \cdot)$ nicht auf ganz $\bar{\mathbb{R}}$ übertragen werden können. Es gilt

$$\infty + \infty = \infty, \quad \infty \cdot \infty = \infty, \qquad 1/\infty = 0$$

(und analoge Aussagen für negative Zahlen in $\overline{\mathbb{R}}$), jedoch lassen sich

$$\infty - \infty, \quad \frac{\infty}{\infty}, \quad 0 \cdot \infty$$

nicht so definieren, dass die in Proposition 3.7 aufgestellten Rechenregeln weiterhin gelten. 2. gilt nämlich nur, falls $\{x, y\} \neq \{-\infty, \infty\}$. Etwa ist $\lim_{n\to\infty} 2n = \infty, \lim_{n\to\infty} -n = -\infty$, und $\lim_{n\to\infty} 2n - n = \infty$, aber $\lim_{n\to\infty} n - 2n = -\infty$. Gilt bei 3. dass $x = 0$ und $\lim_{n\to\infty} y_n = \infty$, so kann $\lim_{n\to\infty} x_n \cdot y_n$ je nach genauer Form der Folgen $(x_n)_{n=1,2,...}$ und $(y_n)_{n=1,2,...}$ Werte $x \in [0, \infty]$ annehmen. Etwa ist $\lim_{n\to\infty} n^{-1} \cdot n^2 = \infty$, $\lim_{n\to\infty} n^{-1} \cdot cn = c$, $\lim_{n\to\infty} n^{-2} \cdot n = 0$.

Bemerkung 3.20 (Häufungspunkte in $\overline{\mathbb{R}}$). Bisher können nur Werte in $\mathbb{R}$ Häufungspunkte von Folgen sein. Wir sagen aber nun auch, dass ∞ ein Häufungspunkt einer reellwertigen Folge $(x_n)_{n=1,2,...}$ ist, falls

$$\forall K > 0 : x_n > K \text{ für unendlich viele } n$$

(und analog für $-\infty$.)

Wir können in $\overline{\mathbb{R}}$ den Limes superior und Limes inferior genau wie in Theorem 3.13 mittels

$$\liminf_{n\to\infty} x_n := \sup_{k\in\mathbb{N}} \inf_{n>k} x_n, \qquad \limsup_{n\to\infty} x_n := \inf_{k\in\mathbb{N}} \sup_{n>k} x_n$$

definieren. In diesem Fall überträgt sich Korollar 3.14 wiefolgt:

Korollar 3.21. *Sei* $(x_n)_{n=1,2,...}$ *eine Folge reeller Zahlen und* $x \in \overline{\mathbb{R}}$. *Dann sind äquivalent:*

1. *Es gibt ein* $x \in \overline{\mathbb{R}}$ *mit* $\lim_{n\to\infty} x_n = x$.
2. *Es gilt* $\liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n = x$.
3. *Die Folge* $(x_n)_{n=1,2,...}$ *hat genau* x *als Häufungspunkt.*

Beweis. Übung. □

4 Reihen

Als Reihe bezeichnet man eine (unendliche) Summe reeller Zahlen, die wir in Abschnitt 4.1 einführen. Diese sind grundlegend für viele Inhalte der Analysis, etwa die Integralrechnung. Wir besprechen hier vor allem die Konvergenz von Reihen (Abschnitt 4.2). Hier gibt es wesentlich mehr Kriterien als bei den Folgen, die herangezogen werden können, um die Konvergenz zu Reihen zu zeigen. Viele hiervon basieren auf der absoluten Konvergenz (Abschnitt 4.3).

4.1 Grundlegendes

Definition 4.1 (Reihe). *Sei* $(x_n)_{n=1,2,...}$ *eine Folge reeller Zahlen. Dann nennen wir* $(s_n)_{n=1,2,...}$ *mit*

$$s_n := \sum_{k=1}^{n} x_k$$

die Folge der Partialsummen *der Folge* $(x_n)_{n=1,2,...}$ *oder auch* (unendliche) Reihe. *Manchmal bezeichnet man mit* $\sum_{k=1}^{\infty} x_k$ *auch die Folge* $(s_n)_{n=1,2,...}$. *Die* $x_1, x_2, ...$ *heißen Glieder der Reihe* $\sum_{k=1}^{\infty} x_k$. *Ist die Folge der Partialsummen konvergent, so schreiben wir*

$$\sum_{k=1}^{\infty} x_k := \lim_{n\to\infty} s_n.$$

(Man beachte, dass also $\sum_{k=1}^{\infty} x_k$ *sowohl die Reihe als auch deren Grenzwert bezeichnet.)*

Konvergiert die Folge der Partialsummen uneigentlich gegen ∞ *oder* $-\infty$, *so schreiben wir* $\sum_{k=1}^{\infty} x_k = \infty$ *bzw.* $\sum_{k=1}^{\infty} x_k = -\infty$. *Für* $i \in \mathbb{N}$ *definieren wir analog* $\sum_{k=i}^{\infty} x_k$ *für die Folge der Partialsummen* $(\sum_{k=i}^{n} x_k)_{n=i,i+1,...}$

Proposition 4.2 (Geometrische Reihe). *Sei* $x \in (-1, 1)$. *Dann gilt*

$$\sum_{i=0}^{\infty} x^i = \frac{1}{1-x}.$$

Beweis. Mit Proposition 1.38 und Proposition 3.7 ist

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

□

Proposition 4.3 (Divergenz der harmonischen Reihe). *Es gilt*

$$\sum_{n=1}^{\infty} \frac{1}{n} = \infty.$$

Beweis. Sei $N \in \mathbb{N}$. Wir schreiben

$$\begin{aligned}\sum_{n=1}^{2^N} \frac{1}{n} &= 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \cdots + \frac{1}{8}\right) + \cdots + \left(\frac{1}{2^{N-1}} + \frac{1}{2^{N-1}+1} + \cdots + \frac{1}{2^N}\right) \\ &\geq 1 + \frac{1}{2} + 2 \cdot \frac{1}{4} + 4 \cdot \frac{1}{8} + \cdots 2^{N-1} \cdot \frac{1}{2^N} = 1 + \frac{N}{2}.\end{aligned}$$

Hieraus folgt leicht die behauptete Divergenz. □

Lemma 4.4. *Es gilt*

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

Beweis. Wir schreiben

$$\begin{aligned}\sum_{n=1}^{\infty} \frac{1}{n(n+1)} &= \lim_{N\to\infty} \sum_{n=1}^{N} \frac{1}{n(n+1)} = \lim_{N\to\infty} \sum_{n=1}^{N} \frac{1}{n} - \frac{1}{n+1} \\ &= 1 - \frac{1}{2} + \frac{1}{2} - \frac{1}{3} + \frac{1}{3} - \frac{1}{4} + \cdots - \frac{1}{N+1} \\ &= \lim_{N\to\infty} 1 - \frac{1}{N+1} = 1\end{aligned}$$

□

4.2 Konvergenzkriterien

Wir wollen nun hinreichende Bedingungen für die Konvergenz einer Reihe aufstellen. Hierbei werden uns die schon bewiesenen Aussagen über Folgen helfen. Ein oft verwendetes Kriterium ist das Leibniz-Kriterium (Theorem 4.10) für alterniernede Reihen.

Lemma 4.5 (Cauchy-Kriterium). *Sei $(x_n)_{n=1,2,\ldots}$ eine Folge reeller Zahlen. Dann konvergiert $\sum_{n=1}^{\infty} x_n$ genau dann, wenn es für alle $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt, so dass für alle $m > n > N$ gilt, dass*

$$|s_n - s_m| = |x_{m+1} + \cdots x_n| < \varepsilon.$$

Beweis. Klar nach Proposition 3.15. □

Das Cauchy-Kriterium liefert eine notwendige und hinreichende Bedingung für die Konvergenz von Reihen. Nun möchten wir auf eine weitere notwendige (aber nicht hinerichende; siehe Proposition 4.3) Bedingung für die Konvergenz einer Reihe hinweisen.

Lemma 4.6 (Notwendige Bedingung für Konvergenz). *Sei $(x_n)_{n=1,2,\ldots}$ eine Folge reeller Zahlen, so dass $\sum_{n=1}^{\infty} x_n$ konvergiert. Dann gilt $x_n \xrightarrow{n\to\infty} 0$.*

Beweis. Sei $s_N := \sum_{n=1}^N x_n$, also $s_N \xrightarrow{N\to\infty} s$ für ein $s \in \mathbb{R}$. Wir verwenden Proposition 3.7 mittels

$$\lim_{N\to\infty} x_N = \lim_{N\to\infty} s_N - s_{N-1} = s - s = 0.$$

Damit ist die Behauptung gezeigt. □

Lemma 4.7. *Sei $(x_n)_{n=1,2,\ldots}$ eine $\mathbb{R}_+$-wertige Folge. Dann konvergiert $\sum_{n=1}^\infty x_n$ genau dann, wenn die Folge der Partialsummen, $(\sum_{n=1}^N x_n)_{N=1,2,\ldots}$ beschränkt ist, also wenn $\sum_{n=1}^\infty x_n < \infty$.*

Beweis. Klar nach Proposition 3.5. □

Proposition 4.8. *Sei $s > 0$. Dann gilt*

$$\sum_{n=1}^\infty \frac{1}{n^s} \begin{cases} = \infty, & \text{falls } s \leq 1, \\ < \infty, & \text{falls } s > 1. \end{cases}$$

Beweis. Sei zunächst $s \leq 1$. Dann ist $1/n^s \geq 1/n$, also $\sum_{n=1}^N 1/n^s \geq \sum_{n=1}^N 1/n \xrightarrow{N\to\infty} \infty$. Damit ist die Divergenz im Fall $s \leq 1$ gezeigt. Für $s > 1$ sei $N \in \mathbb{N}$. Wir schreiben

$$\begin{aligned} \sum_{n=1}^{2^N-1} \frac{1}{n^s} &= 1 + \Big(\frac{1}{2^s} + \frac{1}{3^s}\Big) + \Big(\frac{1}{4^s} + \cdots + \frac{1}{7^s}\Big) + \cdots \\ &\qquad + \Big(\frac{1}{(2^{N-1})^s} + \frac{1}{(2^{N-1}+1)^s} + \cdots + \frac{1}{(2^N-1)^s}\Big) \\ &\leq 1 + 2 \cdot \frac{1}{2^s} + 4 \cdot \frac{1}{4^s} + \cdots + 2^{N-1} \cdot \frac{1}{2^{s(N-1)}} \\ &= 1 + 2^{-(s-1)} + (2^{-(s-1)})^2 + \cdots + (2^{-(s-1)})^{N-1} = \sum_{n=0}^{N-1} 2^{-(s-1)} \\ &\xrightarrow{N\to\infty} \frac{1}{1 - 2^{-(s-1)n}} \end{aligned}$$

nach Proposition 4.2. Insbesondere ist die linke Seite endlich, woraus die Konvergenz folgt. □

Bemerkung 4.9 (ζ-Funktion). Sie Abbildung

$$\zeta : s \mapsto \sum_{n=1}^\infty \frac{1}{n^s}$$

heißt auch *Riemann'sche ζ-Funktion.* Nach der letzten Proposition nimmt sie genau im Bereich $s \in (1, \infty)$ endliche Werte an. Euler fand als erstes den numerischen Wert

$$\zeta(2) = \sum_{n=1}^\infty \frac{1}{n^2} = \frac{\pi^2}{6}.$$

Weitere Werte für gerade s wurden ebenfalls berechnet. Für ungerade s sind jedoch die genauen numberischen Werte der ζ-Funktion unbekannt.

Wir kommen nun noch zu einem Konvergenzkriterium für alternierende Reihen. Dies sind Reihen, deren Glieder abwechselndes Vorzeichen haben.

Theorem 4.10 (Leibniz-Kriterium für alternierende Reihen). *Sei $(x_n)_{n=0,1,...}$ eine monotone fallende Nullfolge und $s_N := \sum_{n=0}^N (-1)^n x_n$. Dann gilt:*

1. *$(s_N)_{N=1,2,...}$ konvergiert.*
2. *Ist $s_N \xrightarrow{N\to\infty} s$, dann gilt $|s - s_N| \leq x_{N+1}$.*

Beweis. Zunächst bemerken wir, dass $s_{N+2} - s_N = (-1)^N (x_{N+2} - x_{N+1})$. Für N ungerade folgt daraus $s_{N+2} \in [s_N, s_{N+1}]$, für N gerade ist hingegen $s_{N+2} \in [s_{N+1}, s_N]$. Also gilt für alle N

$$s_{N+2} \in [s_N \wedge s_{N+1}, s_N \vee s_{N+1}].$$

Damit ist $A_N := [s_N \wedge s_{N+1}, s_N \vee s_{N+1}], N = 1, 2, ...$ eine Intervallschachtelung mit $s_N \vee s_{N+1} - s_N \wedge s_{N+1} = x_{N+1} \downarrow 0$. Also gibt es genau ein s mit $s \in A_N, N = 1, 2, ...$, woraus 1. folgt. Weiter liegt nach Konstruktion s zwischen s_N und s_{N+1}, woraus 2. folgt, da $|s - s_N| \leq |s_{N+1} - s_N| = x_{N+1}$. □

4.3 Absolute Konvergenz von Reihen

Wir lernen nun einige Kriterien für die Konvergenz von Reihen kennen, insbesondere das Quotientenkriterium (Theorem 4.15) und das Wurzelkriterium (Theorem 4.17). Diese liefern sogar die absolute Konvergenz von Reihen, die wir zunächst definieren wollen.

Definition 4.11 (Abolute Konvergenz). *Sei $(x_n)_{n=1,2,...}$ eine Folge und $s_N := \sum_{n=1}^N x_n$. Dann heißt die Reihe $(s_N)_{N=1,2,...}$* absolut konvergent, *wenn $\sum_{n=1}^\infty |x_n|$ konvergiert.*

Alle der nun folgenden Kriterien für absolute Konvergenz basieren auf folgender wichtigen Aussage.

Theorem 4.12 (Majorantenkriterium). *Seien $(x_n)_{n=1,2,...}$ und $(y_n)_{n=1,2,...}$ Folgen und $y_1, y_2, ... \geq 0$. Außerdem existiere ein N, so dass $|x_n| \leq y_n$ für $n \geq N$ gilt. Dann gilt:*

1. *Konvergiert $\sum_{n=1}^\infty y_n$, so konvergiert $\sum_{n=1}^\infty x_n$ absolut und es gilt*
$$\Big|\sum_{n=N}^\infty x_n\Big| \leq \sum_{n=N}^\infty y_n.$$
2. *Divergiert $\sum_{n=1}^\infty x_n$, so divergiert auch $\sum_{n=1}^\infty y_n$.*

Beweis. Es genügt 1. zu zeigen. Sei $\varepsilon > 0$. Da $\sum_{n=1}^\infty y_n$ konvergiert, gibt es mit dem Cauchy-Kriterium für Reihen (Lemma 4.5) ein $N' \geq N$, so dass $\sum_{n=m}^{m'} y_n < \varepsilon$ für $m' > m \geq N'$ gilt. Daraus folgt auch $\left|\sum_{n=m}^{m'} x_n\right| \leq \varepsilon$. Mit anderen Worten erfüllt auch $\sum_{n=1}^\infty x_n$ das Cauchy-Konvergenzkriterium. Außerdem gilt

$$\Big|\sum_{n=N}^\infty x_n\Big| = \lim_{M\to\infty}\Big|\sum_{n=N}^M x_n\Big| \leq \lim_{M\to\infty}\sum_{n=N}^M y_n = \sum_{n=N}^\infty y_n.$$

□

Beispiel 4.13. 1. Für $0 < x < 1$ konvergiert die Reihe

$$\sum_{n=1}^{\infty} x^{(n^2)}.$$

Es gilt nämlich $n^2 > n$ für alle n und damit $x^{n^2} \leq x^n$. Außerdem haben wir in Proposition 4.2 die Konvergenz von $\sum_{n=1}^{\infty} x^n$ gezeigt.

2. Es gilt

$$\sum_{n=1}^{\infty} \frac{1}{\sqrt{n(n+1)}} = \infty,$$

denn $\frac{1}{\sqrt{n(n+1)}} \geq \frac{1}{2n}$ und $\sum_{n=1}^{\infty} \frac{1}{2n} = \infty$ nach Proposition 4.3.

Korollar 4.14. *Eine absolut konvergente Reihe $\sum_{n=1}^{\infty} x_n$ ist auch konvergent.*

Beweis. Wir wenden das Majorantenkriterium mit $y_n := |x_n|$ an. Da die Reihe absolut konvergiert, folgt aus Theorem 22.26.1 direkt dass auch $\sum_{n=1}^{\infty} x_n$ konvergiert. □

Theorem 4.15 (Quotientenkriterium). *Sei $\sum_{n=1}^{\infty} x_n$ eine Reihe und $x_n \neq 0$ für fast alle n. Außerdem existiere*

$$q := \lim_{n\to\infty} \left|\frac{x_{n+1}}{x_n}\right|.$$

Dann gilt:

1. *Ist $q < 1$, so konvergiert die Reihe absolut.*
2. *Ist $q > 1$, so divergiert die Reihe.*

Beweis. 1. Sei $\tilde{q}$ so, dass $q < \tilde{q} < 1$. Dann gilt $|x_{n+1}/x_n| < \tilde{q}$ für fast alle n. Also gibt es ein N, so dass für $n > N$

$$|x_{n+1}| \leq \tilde{q}|x_n| \leq \tilde{q}^2 \cdots |x_{n-1}| \leq \cdots \leq \tilde{q}^{n-N}|x_N|.$$

Mit

$$\sum_{n=N}^{\infty} |x_N|\tilde{q}^{n-N} = |x_N| \cdot \sum_{n=0}^{\infty} \tilde{q}^n < \infty$$

und dem Majorantenkriterium folgt die Behauptung.
2. Sei $\hat{q}$ so, dass $1 < \hat{q} < q$. Wegen $\lim_{n\to\infty} \left|x_{n+1}/x_n\right| > \hat{q}$ gilt $|x_{n+1}| > \hat{q}|x_n|$ für fast alle n. Insbesondere ist $(x_n)_{n=1,2,\ldots}$ keine Nullfolge und die Behauptung folgt aus Lemma 4.6. □

Beispiel 4.16. 1. Sei $0 < x < 1$. Betrachten wir nochmals die Reihe $\sum_{n=1}^{\infty} x^n$ für $0 < x < 1$. Es gilt $0 < x^{n+1}/x^n = x < 1$, also folgt mit dem Quotienenkriterium wieder die Konvergenz der geometrischen Reihe.

2. Ein weiteres Beispiel liefert die Reihe $\sum_{n=1}^{\infty} nx^n$ (wieder für $0 < x < 1$). Diese konvergiert nach dem Quotientenkriterium, weil nämlich

$$\frac{(n+1)x^{n+1}}{nx^n} = \frac{n+1}{n}x \xrightarrow{n\to\infty} x.$$

3. Bekanntermaßen konvergieren die Summen $\sum_{n=1}^{\infty} 1/n^s$ für $s > 1$. Allerdings gilt
$$\frac{(n+1)^s}{n^s} = \left(\frac{n+1}{n}\right)^s \xrightarrow{n\to\infty} 1$$
womit das Quotientenkriterium nicht anwendbar ist, um diese Konvergenz zu beweisen.

Theorem 4.17 (Wurzelkriterium). *Sei $\sum_{n=1}^{\infty} x_n$ eine Reihe. Außerdem sei*
$$q := \limsup_{n\to\infty} |x_n|^{1/n}.$$
Dann gilt:

1. *Ist $q < 1$, so konvergiert die Reihe absolut.*
2. *Ist $q > 1$, so divergiert die Reihe.*

Beweis. 1. Sei $\tilde{q}$ so, dass $q < \tilde{q} < 1$. Dann gilt $|x_n|^{1/n} < \tilde{q}$, also $|x_n| \leq \tilde{q}^n$, für fast alle n. Mit $\sum_{n=N}^{\infty} \tilde{q}^n < \infty$ für alle N und dem Majorantenkriterium folgt die Behauptung.
2. Sei $\hat{q}$ so, dass $1 < \hat{q} < q$. Wegen $\limsup_{n\to\infty} \big|x_n\big| > \hat{q}$ gilt $|x_n| > \hat{q}^n$ für unendlich viele n. Insbesondere ist $(x_n)_{n=1,2,...}$ keine Nullfolge und die Behauptung folgt aus Lemma 4.6. □

Beispiel 4.18. 1. Sei $|x| < 1$. Betrachtet man die Reihe $\sum_{n=1}^{\infty} x^n$, so stellt man fest, dass $|x^n|^{1/n} = |x| < 1$ gilt. Insbesondere zeigt das Wurzelkriterium, dass die Reihe konvergiert. Ebenfalls liefert eine Anwendung von Theorem 3.6.3, dass $\sum_{n=1}^{\infty} nx^n$ konvergiert.

2. Sei $s > 1$. Wir wissen bereits, dass $\sum_{n=1}^{\infty} \frac{1}{n^s}$ konvergiert. Jedoch ist $(1/n^s)^{1/n} = n^{-s/n} \xrightarrow{n\to\infty} 1$ wie in Theorem 3.6. Also ist hier die Anwendung des Wurzelkriteriums nicht möglich.

3. Es gibt Reihen, deren Konvergenz man mit Hilfe des Wurzelkriteriums, nicht jedoch mit Hilfe des Quotientenkriteriums nachweisen kann. Übung.

Theorem 4.19 (Konvergenzkriterium von Raabe). *Sei $\sum_{n=1}^{\infty} x_n$ eine Reihe und $x_n \neq 0$ für fast alle n.*

1. *Falls ein $d > 1$ existiert, so dass*
$$\left|\frac{x_{n+1}}{x_n}\right| \leq 1 - \frac{d}{n}$$
für fast alle n gilt, so ist die Reihe absolut konvergent.

2. *Falls*
$$\frac{x_{n+1}}{x_n} \geq 1 - \frac{1}{n}$$
für fast alle n gilt, so ist die Reihe divergent.

Beweis. 1. OBdA ist $x_n \geq 0$ und $nx_{n+1} \leq (n-1)x_n - (d-1)x_n$ für alle $n = 1, 2, ...$ Insbesondere ist $(nx_{n+1})_{n=0,1,2,...}$ eine fallende, nach unten durch 0 beschränkte Folge, die also nach Theorem 3.13 gegen ein $y \geq 0$ konvergiert. Weiter können wir schreiben
$$\lim_{N\to\infty} \sum_{n=1}^{N} x_n \leq \frac{1}{d-1} \lim_{N\to\infty} \sum_{n=1}^{N} nx_{n+1} - (n-1)x_n = \lim_{N\to\infty} Nx_{N+1} = y,$$

woraus die absolute Konvergenz von $\sum_{n=1}^{\infty} x_n$ mit dem Majorantenkriterium folgt.
2. OBdA ist $x_n \geq 0$ und $nx_{n+1} \geq (n-1)x_n$ für alle $n = 1, 2, ...$ Also ist die Folge $(nx_{n+1})_{n=0,1,...}$ wachsend, also gibt es ein $y > 0$ mit $x_{n+1} \geq y/n$. Nun folgt mit dem Majorantenkriterium und der Divergenz der harmonischen Reihe die Behauptung. □

Beispiel 4.20. 1. Wir haben bereits gesehen, dass die Reihe $\sum_{n=1}^{\infty} \frac{1}{n^2}$ zwar konvergiert, jedoch weder das Quotienten- noch das Wurzelkriterium diese Konvergenz zeigen. Wir schreiben nun

$$\frac{1/(n+1)^2}{1/n^2} = \frac{n^2}{(n+1)^2} = \frac{1}{1+2/n+1/n^2} \leq \frac{1}{1+2/n} = \sum_{k=0}^{\infty}(-\tfrac{2}{n})^k \leq 1 - \frac{2}{n} + \frac{4}{n^2}.$$

Da $4/n^2 \leq 1/2n$ für fast alle n gilt, gilt für dieselben n auch $\frac{1/(n+1)^2}{1/n^2} \leq 1 - 3/(2n)$. Damit können wir das Konvergenzkriterium von Raabe verwenden, um die Konvergenz der Reihe zu zeigen.

2. Genauso können wir mit Hilfe des Konvergenzkriteriums von Raabe die Divergenz von $\sum_{n=1}^{\infty} \frac{1}{n}$ zeigen. Es ist nämlich

$$\frac{1/(n+1)}{1/n} = \frac{1}{1+\frac{1}{n}} \geq 1 - \frac{1}{n}.$$

4.4 Unbedingte Konvergenz

Oftmals kann man Reihen umordnen. Das bedeutet, dass man die einzelnen Summanden in anderer Reihenfolge durchläuft. Zwar ist man von endlichen Summen gewohnt, dass eine Änderung der Summationsreihenfolge nichts am Ergebnis der Summe ändert, jedoch überträgt sich diese Eigenschaft nicht notwendigerweise auf unendliche Summen. Wir führen nun die Menge der Reihen ein, bei denen Umordnung nichts am Grenzwert der Reihe ändert.

Definition 4.21 (Unbedingte Konvergenz). *Die Reihe $\sum_{n=1}^{\infty} x_n$ konvergiert* unbedingt*, wenn es ein $x \in \mathbb{R}$ gibt, so dass für jede Bijektion $\sigma : \mathbb{N} \to \mathbb{N}$ gilt dass $\sum_{n=1}^{\infty} x_{\sigma(n)} = x$.*

Beispiel 4.22. Betrachten wir die Reihe $s = \sum_{n=1}^{\infty}(-1)^{n+1}/n$ und schreiben s_n als die n-te Partialsumme. Diese Reihe konvergiert nach Theorem 4.10. Allerdings konvergiert sie nicht unbedingt.
Denn: Ordnet man die Reihe um in

$$t := 1 - \frac{1}{2} - \frac{1}{4} + \frac{1}{3} - \frac{1}{5} - \frac{1}{6} + \cdots + \frac{1}{2n-1} - \frac{1}{4n-2} - \frac{1}{4n} + \cdots,$$

d.h. in t folgen einem positiven Glied zwei negative, und bezeichne mit t_n die n-te Partialsumme, so gilt

$$t_{3n} = \tfrac{1}{2}s_{2n}.$$

Es gilt nämlich

$$\frac{1}{2n-1} - \frac{1}{4n-2} - \frac{1}{4n} = \frac{1}{2}\Big(\frac{1}{2n-1} - \frac{1}{2n}\Big).$$

Da $s_n \xrightarrow{n\to\infty} s$ nach Voraussetzung gilt, folgt damit $t_n \xrightarrow{n\to\infty} \frac{1}{2}s$. Damit konvergiert die Reihe nicht unbedingt, da wir eine Umordnung gefunden haben, die gegen einen anderen Grenzwert konvergiert.

Proposition 4.23 (Umordnungssatz). *Sei* $\sum_{n=1}^{\infty} x_n$ *eine Reihe.*

1. *Die Reihe konvergiert genau dann unbedingt, wenn sie absolut konvergiert.*
2. *Ist die Reihe konvergent, aber nicht absolut konvergent, sowie* $x \in \mathbb{R}$*, so gibt es eine Bijektion* $\sigma : \mathbb{N} \to \mathbb{N}$ *mit* $\sum_{n=1}^{\infty} x_{\sigma(n)} = x$.

Beweis. 2. Klar ist, dass für konvergente, aber nicht absolut konvergente Reihen[10] $\sum_{n=1}^{\infty} x_n^+ = \sum_{n=1}^{\infty} x_n^- = \infty$ gelten muss. Wäre nämlich nur eine der beiden Reihen unendlich, so zeigt man leicht, dass die Reihe nicht konvergieren kann. Wären hingegen beide Reihen konvergent, so wäre $\sum_{n=1}^{\infty} |x_n| = \sum_{n=1}^{\infty} x_n^+ + \sum_{n=1}^{\infty} x_n^- < \infty$, woraus die absolute Konvergenz folgen würde. Sei nun $x \in \mathbb{R}$ und $k_1, k_2, ...$ eine Abzählung (in aufsteigender Reihenfolge) der Menge $\{n : x_n \geq 0\}$ und $\ell_1, \ell_2, ...$ eine der Mengen $\{n : x_n < 0\}$. Dann definieren wir eine Bijektion $\sigma : \mathbb{N} \to \mathbb{N}$ wiefolgt:

$$\sigma(1) := \begin{cases} k_1, & \text{falls } x \geq 0, \\ \ell_1, & \text{falls } x < 0. \end{cases}$$

$$\sigma(n+1) := \begin{cases} \inf\{k_m : k_m \neq \sigma(1), ..., \sigma(n)\}, & \text{falls } \sum_{m=1}^{n} x_{\sigma(m)} \leq x, \\ \inf\{\ell_m : \ell_m \neq \sigma(1), ..., \sigma(n)\}, & \text{falls } \sum_{m=1}^{n} x_{\sigma(m)} > x. \end{cases}$$

Mit anderen Worten werwenden wir etwa für $x > 0$ solange positive Folgenglieder von $(x_n)_{n=1,2,...}$, bis $\sum_{m=1}^{n} x_{\sigma(m)} > x$. Anschließend ziehen wir wieder so lange Folgenglieder ab bis $\sum_{m=1}^{n} x_{\sigma(m)} \leq x$ usw. Klar ist, dass σ eine Bijektion ist, da beide Reihen $\sum_{n=1}^{\infty} x_{k_n}$ und $\sum_{n=1}^{\infty} x_{\ell_n}$ divergieren. Dadurch ist nämlich gewährleistet, dass es genügt, endlich viele positive (negative) Folgenglieder aufzusummieren, bevor das nächste negative (positive) Folgenglied in der Umordnung der Reihe verwendet wird. Weiter ist wegen Lemma 4.6 die Folge $(x_n)_{n=1,2,...}$ eine Nullfolge. Sei also $\varepsilon > 0$, so gibt es ein N mit $|x_n| < \varepsilon$ für $n > N$. Sei nun N' groß genug für $\{\sigma(1), ..., \sigma(N')\} \supseteq \{1, ..., N\}$. Dann ist nach Konstruktion

$$\Big|x - \sum_{m=1}^{N'} x_{\sigma(m)}\Big| \leq \sup\{|x - x_n| : n \notin \{\sigma(1), ..., \sigma(N')\} \leq \sup\{|x - x_n| : n \notin \{1, ..., N\} \leq \varepsilon,$$

da nach 'Überspringen' von x von oben (von unten) als nächstes positive (negative) Folgenglieder addiert werden. Hieraus folgt die Konvergenz $x = \sum_{n=1}^{\infty} x_{\sigma(n)}$.
1. Sei zunächst $\sum_{n=1}^{\infty} x_n$ absolut konvergent und $x = \sum_{n=1}^{\infty} x_n$. Weiter sei $\sigma : \mathbb{N} \to \mathbb{N}$ eine Bijektion und $\varepsilon > 0$. Zunächst wählen wir N groß genug, so dass $\sum_{n=N}^{\infty} |x_n| < \varepsilon/3$ sowie $\big|x - \sum_{n=1}^{N} x_n\big| < \varepsilon/3$ und N' groß genug, so dass $\{\sigma(1), ..., \sigma(N')\} \supseteq \{1, ..., N\}$. Dann gilt

$$\begin{aligned} \Big|x - \sum_{n=1}^{\infty} x_{\sigma(n)}\Big| &\leq \Big|x - \sum_{n=1}^{N} x_n\Big| + \Big|\sum_{n=1}^{N'} x_{\sigma(n)} - \sum_{n=1}^{N} x_n\Big| + \sum_{n=N+1}^{\infty} |x_n| \\ &\leq \varepsilon/3 + 2 \sum_{n=N+1}^{\infty} |x_n| < \varepsilon. \end{aligned}$$

Daraus folgt die unbedingte Konvergenz. Andersherum können wir voraussetzen, dass $\sum_{n=1}^{\infty} x_n$ konvergiert. Wir zeigen, dass die Reihe nicht unbedingt konvergiert, wenn die Reihe nicht absolut konvergiert. Dies folgt jedoch direkt aus 2. □

[10] Wir verwenden die Schreibweise $x^+ := x \vee 0$ und $x^- := -(x \wedge 0)$. Man beachte, dass $x^- \geq 0$.

Bemerkung 4.24 (Notation: Summation über Mengen). Bisher haben wir die Summenschreibweise $\sum_{n=1}^{\infty} x_n$ verwendet. Im Falle der unbedingten Konvergenz wissen wir außerdem, dass es für den Grenzwert dieser Reihe keine Rolle spielt, in welcher Reihenfolge die Folgenglieder der Folge $(x_n)_{n=1,2,...}$ aufsummiert werden. Um dies zu verdeutlichen, werden wir nun auch $\sum_{n\in\mathbb{N}} x_n$ schreiben. Allgemeiner sei $I \subseteq \mathbb{N}$ und $k_1, k_2, ...$ eine Abzählung von I. Dann schreiben wir

$$\sum_{n\in I} x_n := \sum_{n=1}^{\infty} x_{k_n},$$

wenn die Summationsreihenfolge keine Rolle spielt, die Reihe $\sum_{n=1}^{\infty} x_{k_n}$ also unbedingt konvergent ist.

Theorem 4.25 (Großer Umordnungssatz). *Sei $\sum_{n=1}^{\infty} x_n$ eine unbedingt konvergente (also nach Proposition 4.23 eine absolut konvergente) Reihe, sowie $I_1, I_2, ... \subseteq \mathbb{N}$ paarweise disjunkt mit $\bigcup_{k=1}^{\infty} I_k = \mathbb{N}$. Dann konvergiert sowohl jedes $s_k := \sum_{n\in I_k} x_n$, $k = 1, 2, ...,$ als auch $\sum_{k=1}^{\infty} s_k$ absolut und es gilt*

$$\sum_{n=1}^{\infty} x_n = \sum_{k=1}^{\infty} \sum_{n\in I_k} x_n. \tag{4.1}$$

Beweis. Da $\sum_{n=1}^{\infty} |x_n| < \infty$ nach Voraussetzung, gilt auch $\sum_{n\in I_k} |x_n| < \infty$ (nach dem Majorantenkriterium) und damit die Existenz jedes s_k, $k = 1, 2, ...$ Weiter wählen wir $I_1' = I_1 \setminus \{N_1, N_1 + 1...\}, I_2' = I_2 \setminus \{N_2, N_2 + 1, ...\}, ...$ endlich, so dass $\sum_{n\in I_k\setminus I_k'} |x_n| < \varepsilon 2^{-k}$. Dann gilt

$$\sum_{k=1}^{\infty} |s_k| = \sum_{k=1}^{\infty} \Big|\sum_{n\in I_k} x_n\Big| = \sum_{k=1}^{\infty} \Big|\sum_{n\in I_k'} x_n + \sum_{n\in I_k\setminus I_k'} x_n\Big| \leq \sum_{n=1}^{\infty} |x_n| + \varepsilon \sum_{k=1}^{\infty} 2^{-k} = \sum_{n=1}^{\infty} |x_n| + \varepsilon < \infty.$$

Hieraus folgt die absolute Konvergenz von $\sum_{k=1}^{\infty} s_k$. Es bleibt noch, (4.1) zu zeigen. Sei $\varepsilon > 0$ und N so groß, dass $\sum_{n=N+1}^{\infty} |x_n| < \varepsilon$. Weiter finden wir ein K, so dass $\bigcup_{k=1}^{K} I_k' \supseteq \{1, ..., N\}$ und $\sum_{k=K+1}^{\infty} |s_k| < \varepsilon$. Dann ist auch

$$\begin{aligned}\Big|\sum_{n=1}^{\infty} x_n - \sum_{k=1}^{\infty} s_k\Big| &\leq \underbrace{\Big|\sum_{n=1}^{N} x_n - \sum_{k=1}^{K} \sum_{n\in I_k'} x_n\Big|}_{\leq \sum_{n=N+1}^{\infty} |x_n| \text{ wegen } \bigcup_{k=1}^{K} I_k' \supseteq \{1,...,N\}} + \sum_{n=N+1}^{\infty} |x_n| + \sum_{k=1}^{K} \sum_{n\in I_k\setminus I_k'} |x_n| + \sum_{k=K+1}^{\infty} |s_k| \\ &< 2\varepsilon + \varepsilon \sum_{k=1}^{\infty} 2^{-k} + \varepsilon = 4\varepsilon.\end{aligned}$$

Daraus folgt die Behauptung. □

Korollar 4.26 (Doppelreihensatz). *Sei $(x_{ij})_{(i,j)\in\mathbb{N}^2}$ eine Familie reeller Zahlen und entweder $\sum_{i=1}^{\infty} \Big(\sum_{j=1}^{\infty} |x_{ij}|\Big) < \infty$ oder $\sum_{j=1}^{\infty} \Big(\sum_{i=1}^{\infty} |x_{ij}|\Big) < \infty$. Dann ist für jede Abzählung $\sigma : \mathbb{N} \to \mathbb{N}^2$*

$$\sum_{n=1}^{\infty} x_{\sigma(n)} = \sum_{i=1}^{\infty} \Big(\sum_{j=1}^{\infty} x_{ij}\Big) = \sum_{j=1}^{\infty} \Big(\sum_{i=1}^{\infty} x_{ij}\Big).$$

Beweis. Die Aussage folgt direkt aus dem großen Umordnungssatz, wenn man ihn auf die Folge $(x_{\sigma(n)})_{n=1,2,\dots}$ und $I_k := \{n : \sigma(n) = (i,k) \text{ für ein } i\}$ anwendet. $\square$

Beispiel 4.27. Es gilt

$$\sum_{n=2}^{\infty}\sum_{k=2}^{\infty}\frac{1}{n^k} = 1.$$

Denn: Wir schreiben

$$\sum_{k=2}^{\infty}\frac{1}{n^k} = \frac{1}{n^2}\sum_{k=0}^{\infty}\frac{1}{n^k} = \frac{1}{n^2}\frac{1}{1-\frac{1}{n}} = \frac{1}{n(n-1)}.$$

Also gilt $\sum_{n=2}^{\infty}\sum_{k=2}^{\infty}\frac{1}{n^k} < \infty$ und daraus folgt

$$\sum_{n=2}^{\infty}\sum_{k=2}^{\infty}\frac{1}{n^k} = \sum_{n=2}^{\infty}\frac{1}{n(n-1)} = 1$$

nach Lemma 4.4.

Definition 4.28 (Cauchy-Produkt). *Seien $(x_n)_{n=0,1,\dots}$ und $(y_n)_{n=0,1,\dots}$ Folgen reeller Zahlen und*

$$d_n := \sum_{i=0}^{n} x_i y_{n-i}, \qquad n=0,1,\dots.$$

Dann heißt $\sum_{n=0}^{\infty} d_n$ das Cauchy-Produkt *der Reihen $\sum_{n=0}^{\infty} x_n$ und $\sum_{n=0}^{\infty} y_n$.*

Proposition 4.29 (Absolute Konvergenz von Cauchy-Produkten). *Seien $(x_n)_{n=0,1,\dots}$, $(y_n)_{n=0,1,\dots}$, $(d_n)_{n=0,1,\dots}$ wie in Definition 4.28. Sind $\sum_{n=0}^{\infty} x_n$ und $\sum_{n=0}^{\infty} y_n$ absolut konvergent, so ist auch $\sum_{n=0}^{\infty} d_n$ absolut konvergent und es gilt*

$$\sum_{n=0}^{\infty} d_n = \Big(\sum_{n=0}^{\infty} x_n\Big)\cdot\Big(\sum_{n=0}^{\infty} y_n\Big).$$

Beweis. Die Aussage ist klar, wenn man den Doppelreihensatz auf die Familie $(x_i y_j)_{i,j=0,1,2,\dots}$ anwendet. $\square$

Beispiel 4.30. 1. Für $|x| < 1$ gilt

$$\sum_{n=0}^{\infty}(n+1)x^n = \frac{1}{(1-x)^2}.$$

Denn: Betrachten wir die Familie $(x^i x^j)_{i,j=0,1,2\dots}$. Dann ist

$$d_n := \sum_{i=0}^{n} x^i x^{n-i} = \sum_{i=0}^{n} x^n = (n+1)x^n.$$

Da $(d_n)_{n=0,1,2,\dots}$ nach Beispiel 4.16.2 summierbar ist, folgt

$$\sum_{n=0}^{\infty}(n+1)x^n = \sum_{n=0}^{\infty} d_n = \sum_{i=0}^{\infty} x^i \sum_{j=0}^{\infty} x^j = \frac{1}{(1-x)^2}.$$

2. Die Voraussetzung der absoluten Konvergenz der Reihe $\sum_{n=1}^{\infty} d_n$ ist wichtig. Sei etwa $x_n = y_n = (-1)^n \frac{1}{\sqrt{n+1}}$. Dann gilt wegen $2\sqrt{xy} \leq x + y$ für $x, y \geq 0$, dass

$$|d_n| = \Big| \sum_{i=0}^{n} x_i y_{n-i} \Big| = \sum_{i=0}^{n} \frac{1}{\sqrt{i+1}\sqrt{n+1-i}} \geq (n+1)\frac{2}{n+2} \geq 1.$$

Damit divergiert $\sum_{n=0}^{\infty} d_n$, jedoch existiert $\left(\sum_{i=0}^{\infty} (-1)^n \frac{1}{\sqrt{n+1}} \right)^2$ nach dem Leibnizkriterium, Theorem 4.10.

Teil II
Stetigkeit, Ableitungen und Integrale

Nachdem wir die grundlegende Hilfsmittel der Folgen und Reihen eingeführt haben, studieren wir Eigenschaften von Funktionen. Es ist hierzu hilfreich, an Funktionen $f : \mathbb{R} \to \mathbb{R}$ zu denken. Zentral werden die Begriffe der Stetigkeit (Abschnitt 5), Differenzierbarkeit (Abschnitt 7) und Integrierbarkeit (Abschnitt 8) von Funktionen sein.

5 Stetigkeit in metrischen Räumen

Um Stetigkeit von Funktionen in allgemeinerem Rahmen untersuchen zu können, führen wir zunächst den Begriff des metrischen Raumes in Abschnitt 5.1 ein. In solchen Räumen ist ein Abstandsbegriff definiert, der auch Metrik genannt wird.

5.1 Metrische Räume

Definition 5.1 (Metrischer Raum). *Sei E eine nicht-leere Menge. Eine Funktion $r : E \times E \to \mathbb{R}_+$ heißt* Metrik *auf E, falls folgendes gilt:*

1. *Definitheit: Für alle $x, y \in E$ gilt $r(x, y) = 0$ genau dann, wenn $x = y$.*
2. *Symmetrie: Für alle $x, y \in E$ gilt $r(x, y) = r(y, x)$.*
3. *Dreiecksungleichung: Es ist $r(x, z) \leq r(x, y) + r(y, z)$ für alle $x, y, z \in E$.*

Das Paar (E, r) heißt metrischer Raum.

Beispiel 5.2. 1. Auf $\mathbb{R}$ definieren wir $r : (x, y) \mapsto |x - y|$. Dann ist r eine Metrik, denn es gilt $|x - z| = |x - y + y - z| \leq |x - y| + |y - z|$.

2. Auf $\mathbb{R}^d$ definieren wir r mittels

$$r_{\mathrm{eukl}}(x, y) := r(x, y) := \sqrt{\Big(\sum_{i=1}^d (x_i - y_i)^2\Big)}.$$

Die Definitheit und Symmetrie von r sieht man sofort. Für die Dreiecksungleichung seien $x, y, z \in \mathbb{R}^d$. Wir schreiben

$$\begin{aligned} r(x, z)^2 &= \sum_{i=1}^d |x_i - z_i|^2 = \sum_{i=1}^d |x_i - y_i + y_i - z_i| \cdot |x_i - z_i| \\ &\leq \sum_{i=1}^d |x_i - y_i| \cdot |x_i - z_i| + \sum_{i=1}^d |y_i - z_i| \cdot |x_i - z_i| \\ &\leq \sqrt{\sum_{i=1}^d (x_i - y_i)^2 \sum_{i=1}^d (x_i - z_i)^2} + \sqrt{\sum_{i=1}^d (y_i - z_i)^2 \sum_{i=1}^d (x_i - z_i)^2} \\ &= (r(x, y) + r(y, z)) \cdot r(x, z) \end{aligned}$$

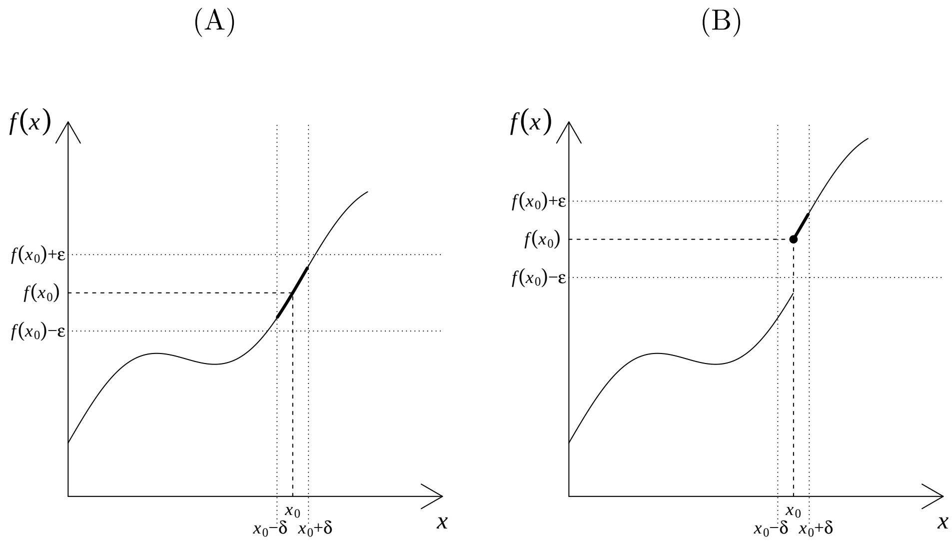

Abbildung 5.1: (A) Die hier abgebildete Funktion f ist stetig in x_0. Schließlich liegen alle Funktionswerte von $x \in (x_0 - \delta, x_0 + \delta)$ in $(f(x_0) - \varepsilon, f(x_0) + \varepsilon)$. (B) Die Funktion ist unstetig in x_0.

wobei wir im vorletzten $\leq$ die aus der linearen Algebra bekannte Cauchy-Schwartz'sche Ungleichung verwendet habe. Dividieren wir durch $r(x, z)$, ist die Dreiecksungleichung gezeigt. Die in 1. und 2. eingeführte Metrik heißt auch *euklidische Metrik*.

3. Auf einer beliebigen Menge E definieren wir die Abbildung $r : E \times E \to \mathbb{R}_+$ mittels

$$r(x, y) = \begin{cases} 0, & \text{falls } x = y, \\ 1, & \text{sonst.} \end{cases}.$$

Diese Abbildung ist eine Metrik, denn die Definitheit und Symmetrie sind offensichtlich, und für die Dreiecksungleichung gilt im Fall $x \neq z$, dass $r(x, z) = 1 \leq r(x, y) + r(y, z)$, da entweder $x \neq y$ oder $y \neq z$. Die Metrik r heißt auch *diskrete Metrik*.

Definition 5.3 (Konvergenz in metrischen Räumen). *Sei (E, r) ein metrischer Raum, $x \in E$ und $(x_n)_{n=1,2,\ldots}$ eine E-wertige Folge. Dann schreiben wir $\lim_{n\to\infty} x_n = x$ oder auch $x_n \xrightarrow{n\to\infty} x$, falls $\lim_{n\to\infty} r(x_n, x) = 0$.*

Beispiel 5.4. Wir verwenden wie üblich die euklidische Metrik auf $\mathbb{R}$. Hier gilt also $x_n \xrightarrow{n\to\infty} x$ genau dann, wenn $x_n - x \xrightarrow{n\to\infty} 0$, was konsistent mit früher gelerntem ist. Durch Änderung der Metrik bekommen wir einen anderen Konvergenzbegriff. Ist nämlich[11] $r(x, y) = 1_{x\neq y}$ wie in Beispiel 5.2.3, so ist $r(x_n, x) \xrightarrow{n\to\infty} 0$ genau dann, wenn $x_n = x$ für fast alle n gilt.

Definition 5.5 (Stetigkeit). *Seien (E, r) und (E', r') metrische Räume. Dann heißt eine Funktion $f : E \to E'$* stetig *in $x_0 \in E$, falls es für jedes $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass $r'(f(x_0), f(y)) < \varepsilon$ für alle $y \in E$ mit $r(x_0, y) < \delta$. In Quantorenschreibweise:*

[11] Hier setzen wir $1_{x\neq y} := \begin{cases} 1, & \text{falls } x \neq y \\ 0, & \text{sonst.} \end{cases}$

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall y \text{ mit } r(x_0, y) < \delta : r'(f(x_0), f(y)) < \varepsilon.$$

Eine Funktion $f : E \to E'$ heißt stetig, *wenn f in allen $x \in E$ stetig ist.*

Beispiel 5.6.

1. Die Darstellung einer stetigen und einer unstetigen Funktion findet man in Abbildung 5.1.

2. Die Funktion $f : x \mapsto x^2$ ist auf ganz $\mathbb{R}$ stetig.

   Denn: Sei $x \in \mathbb{R}$ sowie $0 < \varepsilon < 1$ und $\delta := \varepsilon/(2|x| + 1)$. Dann gilt für y mit $|x - y| \leq \delta$

   $$|x^2 - y^2| = |x - y| \cdot |x + y| \leq \frac{\varepsilon}{2|x| + 1}(2|x| + \delta) \leq \varepsilon.$$

   Damit ist die Stetigkeit in x gezeigt. Da x beliebig war, folgt die Stetigkeit in ganz $\mathbb{R}$.

3. Seien (E, r) und (E', r') metrische Räume. Eine Funktion $f : E \to E'$ heißt *(global) Lipschitz-stetig*, wenn es ein $L \in \mathbb{R}_+$ gibt, so dass

   $$r'(f(x), f(y)) \leq L \cdot r(x, y).$$

   Jede Lipschitz-stetige Funktion ist stetig.

   Denn: Sei $x \in E$ und $\varepsilon > 0$ sowie $\delta := \varepsilon/L \wedge 1$. Dann gilt für $y \in E$ mit $r(x, y) < \delta$

   $$r'(f(x), f(y)) \leq Lr(x, y) \leq L \cdot \delta \leq \varepsilon.$$

4. Die Funktion

   $$f : \begin{cases} \mathbb{R} & \to \{0, 1\} \\ x & \mapsto \begin{cases} 1, & \text{falls } x \in \mathbb{Q} \\ 0, & \text{falls } x \in \mathbb{R} \setminus \mathbb{Q} \end{cases} \end{cases}$$

   ist für kein $x \in \mathbb{R}$ stetig.
   Beweis siehe Übung.

5.2 Äquivalente Formulierungen

Ziel dieses Abschnittes ist es, möglichst viele zur Stetigkeit einer Funktion f im Punkt x äquivalente Formulierungen zu finden. Hierzu werden wir einige aus dem Gebiet der Topologie stammende Begriffe einführen.

Definition 5.7 (Umgebung, offene Menge). *Sei (E, r) ein metrischer Raum und $x \in E$.*

1. *Dann bezeichnen wir für $\varepsilon > 0$ mit*

   $$B_\varepsilon(x) := \{y \in E : r(x, y) < \varepsilon\}$$

   den (offenen) Ball um x mit Radius ε*.*

2. *Weiter heißt $U \subseteq E$ (mit $x \in U$) eine* Umgebung von x, *wenn es ein $\varepsilon > 0$ gibt mit $B_\varepsilon(x) \subseteq U$.*

3. *Eine Menge $O \subseteq E$ heißt* offen*, wenn O für alle $x \in O$ eine Umgebung ist. Eine Menge $A \subseteq E$ heißt* abgeschlossen*, falls A^c offen ist.*

Beispiel 5.8. 1. Betrachten wir den metrischen Raum $(\mathbb{R}, |.|)$. Dann ist $B_\varepsilon(x) = (x - \varepsilon, x + \varepsilon)$. Außerdem sind alle offenen Intervalle (a, b) offen(e Mengen).
Denn: für $x \in (a, b)$ und $\varepsilon := (x - a) \wedge (b - x)$ gilt $(x - \varepsilon, x + \varepsilon) \subseteq (a, b)$. Damit ist (a, b) offen.

2. Sei (E, r) ein metrischer Raum, $x \in E$ und $\varepsilon > 0$. Dann ist $B_\varepsilon(x)$ eine Umgebung von x und offen.
Denn: Die Behauptung, dass $B_\varepsilon(x)$ eine Umgebung von x ist, ist klar. Ist $y \in B_\varepsilon(x))$ und $\delta := \varepsilon - r(x, y) > 0$. Ist $z \in B_\delta(y)$, so folgt $r(x, z) \leq r(x, y) + r(y, z) < r(x, y) + \varepsilon - r(x, y) < \varepsilon$, also $B_\delta(y) \subseteq B_\varepsilon(x)$. Damit ist $B_\varepsilon(x)$ eine Umgebung von y. Da y beliebig war, folgt die Behauptung.

3. Die Menge $\mathbb{Q} \subseteq \mathbb{R}$ ist weder offen noch abgeschlossen.
Denn: Wäre $\mathbb{Q}$ offen und $x \in \mathbb{Q}$, so müsste $(x - \varepsilon, x + \varepsilon) \subseteq \mathbb{Q}$ für ein $\varepsilon > 0$ gelten. Allerdings haben wir in Proposition 2.30 gesehen, dass Intervalle überabzählbar sind, $\mathbb{Q}$ jedoch nach Lemma 1.30 abzählbar.

Lemma 5.9. *1. Sei I eine beliebige Menge und $(O_i)_{i \in I}$ eine Familie offener Mengen. Dann ist auch $\bigcup_{i \in I} O_i$ offen. Ist I endlich, so ist auch $\bigcap_{i \in I} O_i$ offen.*

2. Ist $(A_i)_{i \in I}$ eine Familie abgeschlossener Mengen, so ist auch $\bigcap_{i \in I} A_i$ abgeschlossen. Ist I endlich, so ist auch $\bigcup_{i \in I} A_i$ abgeschlossen.

Bemerkung 5.10. In der Situation obigen Lemmas sagt man auch: Beliebige Vereinigungen und endliche Schnittmengen offener Mengen sind offen, und beliebige Schnitte und endliche Vereinigungen abgeschlossener Mengen sind abgeschlossen.

Beweis von Lemma 5.9. Es genügt, 1. zu zeigen. 2. folgt dann einfach mittels Komplementbildung. Sei $x \in \bigcup_{i \in I} O_i$. Dann ist also $x \in O_i$ für ein $i \in I$. Also gibt es eine Umgebung U von x mit $U \subseteq O_i$ und damit $U \subseteq \bigcup_{i \in I} O_i$. Damit ist $\bigcup_{i \in I} U_i$ offen. Sei $x \in \bigcap_{i \in I} O_i =: O$ mit endlichem I. Dann gibt es eine Familie $(\delta_i)_{i \in I}$ mit $B_{\delta_i}(x) \subseteq O_i$. Sei nun $\delta := \min_{i \in I} \delta_i > 0$. Dann gilt $B_\delta(x) \subseteq B_{\delta_i}(x)$ für alle $i \in I$, also auch $B_\delta(x) \subseteq O$. Damit ist O offen. $\square$

Um uns an die neu erlernten Begriffe zu gewöhnen, folgen nun einfache Eigenschaften.

Lemma 5.11. *Sei (E, r) ein metrischer Raum und $U \subseteq E$.*

1. *Sei $x \in U$. Ist U eine Umgebung von x und $U' \supseteq U$, dann ist auch U' eine Umgebung von x.*

2. *Ist $x \in E$ und sind $U_1, ..., U_n$ Umgebungen von x, dann ist auch $\bigcap_{k=1}^n U_k$ eine Umgebung von x.*

3. *Sei $(x_n)_{n=1,2,...}$ eine E-wertige Folge und $x \in E$. Dann ist $x_n \xrightarrow{n \to \infty} x$ genau dann, wenn jede Umgebung U von x fast alle Folgenglieder von $(x_n)_{n=1,2,...}$ enthält.*

4. *Ist U mit $x \in U$ keine Umgebung von x und $\delta > 0$. Dann gibt es ein $y \notin U$ mit $r(x, y) < \delta$.*

5. *U ist genau dann abgeschlossen wenn gilt: Ist $(x_n)_{n=1,2,...}$ eine U-wertige Folge mit $x_n \xrightarrow{n\to\infty} x$ für ein $x \in E$, dann ist $x \in U$.*

Beweis. 1. Da $B_\varepsilon(x) \subseteq U$ für ein $\varepsilon > 0$, ist auch $B_\varepsilon(x) \subseteq U'$ und damit ist U' eine Umgebung von x.
2. Seien $\varepsilon_1, ..., \varepsilon_n > 0$ so, dass $B_{\varepsilon_k}(x) \subseteq U_k$, $k = 1, ..., n$. Dann gilt für $\varepsilon := \min\{\varepsilon_1, ..., \varepsilon_n\} > 0$, dass $x \in B_\varepsilon(x) \subseteq B_{\varepsilon_k}(x) \subseteq U_k$, also auch $x \in B_\varepsilon(x) \subseteq \bigcap_{k=1}^n U_k$.
3. Sei zunächst $x_n \xrightarrow{n\to\infty} x$ und U eine Umgebung von x. Dann gibt es ein $\varepsilon > 0$, so dass $x \in B_\varepsilon(x) \subseteq U$. Außerdem gilt $r(x_n, x) < \varepsilon$ für fast alle n, also $x_n \in B_\varepsilon(x)$ für fast alle n, und damit auch $x_n \in U$ für fast alle n. Sei andersherum $\varepsilon > 0$. Dann sind fast alle Folgenglieder von $(x_n)_{n=1,2,...}$ in $B_\varepsilon(x)$ (da dieser Ball eine Umgebung von x ist), also gibt es ein N, so dass $r(x_n, x) < \varepsilon$ für $n > N$. Das bedeutet aber, dass $r(x_n, x) \xrightarrow{n\to\infty} 0$, also $x_n \xrightarrow{n\to\infty} x$.
4. Angenommen für alle $y \notin U$ wäre $r(x, y) > \delta$, also $U^c \subseteq (B_\delta(x))^c$. Dann wäre auch $B_\delta(x) \subseteq U$ und damit wäre U eine Umgebung von x.
5. Sei zunächst U abgeschlossen, U^c also offen. Sei $x \in E$ und $(x_n)_{n=1,2,...}$ eine U-wertige Folge mit $x_n \xrightarrow{n\to\infty} x$. Angenommen, $x \in U^c$. Da U^c eine Umgebung von x ist, muss damit U^c fast alle Folgenglieder von $(x_n)_{n=1,2,...}$ enthalten. Dies widerspricht jedoch, dass $x_n \in U$ für alle $n = 1, 2, ...$
Andersherum führen wir ebenfalls einen Widerspruchsbeweis. Nehmen wir an, U sei nicht abgeschlossen, U^c also nicht offen. Damit gibt es ein $x \in U^c$, so dass U^c keine Umgebung von x ist. Sei $\varepsilon_n \downarrow 0$ eine Nullfolge. Nach 4. gibt es $x_n \in (U^c)^c = U$, so dass $r(x_n, x) < \varepsilon_n$. Damit ist $(x_n)_{n=1,2,...}$ eine U-wertige Folge mit $x_n \xrightarrow{n\to\infty} x \in E$, aber $x \notin U$. □

Nun können wir eine Reihe von äquivalanten Formulierungen für Stetigkeit aufstellen.

Proposition 5.12 (Äquivalente Formulierungen der Stetigkeit). *Seien (E, r) und (E', r') metrische Räume, $f : E \to E'$ und $x \in E$. Äquivalent sind:*

1. *f ist in x stetig.*
2. *Für jedes $\varepsilon > 0$ gibt es eine Umgebung U von x, so dass $r'(f(x), f(y)) < \varepsilon$ für alle $y \in U$ gilt.*
3. *Für jede Folge $(x_n)_{n=1,2,...}$ mit $x_n \xrightarrow{n\to\infty} x$ gilt $f(x_n) \xrightarrow{n\to\infty} f(x)$.*
4. *Für jede Umgebung U' von $f(x)$ ist $f^{-1}(U')$ eine Umgebung von x.*

Beweis. 1.⇒2. Nach Voraussetzung gibt es für $\varepsilon > 0$ ein $\delta > 0$, so dass $r'(f(x), f(y)) < \varepsilon$ gilt, falls $r(x, y) < \delta$. Insbesondere können wir $U = B_\delta(x)$ wählen.
2.⇒1. Da U eine Umgebung von x ist, können wir ein $\delta > 0$ wählen mit $B_\delta(x) \subseteq U$. Damit gilt für $y \in B_\delta(x)$, dass $r'(f(x), f(y)) < \varepsilon$ gilt, falls $y \in B_\delta(x)$, d.h. falls $r(x, y) < \delta$.
1.⇒3. Sei $(x_n)_{n=1,2,...}$ eine Folge mit $x_n \xrightarrow{n\to\infty} x$ und $\varepsilon > 0$. Nach Voraussetzung gibt es ein $\delta > 0$, so dass $r(f(x), f(y)) < \varepsilon$ für alle y mit $r'(x, y) < \delta$. Sei N so groß, dass $r(x_n, x) < \delta$ für alle $n > N$. Dann gilt also auch $r'(f(x_n), f(x)) < \varepsilon$ für alle $n > N$ und damit $f(x_n) \xrightarrow{n\to\infty} f(x)$.
3.⇒4. Wir führen einen Widerspruchsbeweis. Angenommen, es gibt eine Umgebung U' von $f(x)$, so dass $f^{-1}(U')$ keine Umgebung von x ist. Dann ist auch $f^{-1}(B_\varepsilon(f(x)))$ keine Umgebung von $f(x)$ für ein geeignetes $\varepsilon > 0$ mit $B_\varepsilon(f(x)) \subseteq U'$. Sei $\delta_n \downarrow 0$ eine fallende Nullfolge. Da $f^{-1}(B_\varepsilon(f(x)))$ keine Umgebung von x ist, aber $x \in f^{-1}(B_\varepsilon(f(x)))$, gibt es

nach Lemma 5.11.4 ein x_n mit $x_n \xrightarrow{n\to\infty} x$, aber $x_n \notin f^{-1}(B_\varepsilon(f(x)))$, $n = 1, 2, ...$ Damit gilt $f(x_n) \notin f(f^{-1}(B_\varepsilon(f(x)))) = B_\varepsilon(f(x))$. Insbesondere gilt $r'(f(x_n), f(x)) > \varepsilon$ für alle n, also kann $f(x_n) \xrightarrow{n\to\infty} f(x)$ nicht gelten.
4.⇒1. Sei $\varepsilon > 0$. Nach Voraussetzung ist $f^{-1}(B_\varepsilon(f(x)))$ eine Umgebung von x. Damit gibt es ein $\delta > 0$, so dass $B_\delta(x) \subseteq f^{-1}(B_\varepsilon(f(x)))$. Daraus folgt $f(B_\delta(x)) \subseteq B_\varepsilon(f(x))$. Das bedeutet, dass für y mit $r(x, y) < \delta$ gilt, dass $r'(f(x), f(y)) < \varepsilon$. Dies war zu zeigen. □

Korollar 5.13. *Seien $(E, r), (E', r')$ metrische Räume und $f : E \to E'$. Dann sind äquivalent:*

1. *f ist stetig für alle $x \in E$.*
2. *$f^{-1}(O')$ ist für alle offenen Mengen $O' \subseteq E'$ offen.*
3. *$f^{-1}(A')$ ist für alle abgeschlossenen Mengen $A' \subseteq E'$ abgeschlossen.*

Beweis. 1.⇒2.: Sei zunächst f stetig und $O' \subseteq E'$ offen. Nun gibt es zwei Möglichkeiten. (i) Angenommen es gibt kein $x \in E$ mit $f(x) \in O'$. Dann ist $f^{-1}(O') = \emptyset$, insbesondere ist $f^{-1}(O')$ offen. (ii) Angenommen $f^{-1}(O') \neq \emptyset$. Dann ist für jedes $x \in f^{-1}(O')$ die Menge O' (da sie offen ist) eine Umgebung von $f(x)$ und damit ist nach Lemma 5.11.4 die Menge $f^{-1}(O')$ eine Umgebung von x. Da x beliebig war, ist also $f^{-1}(O')$ Umgebung aller ihrer Elemente, und damit offen.
2.⇒1.: Andersherum sei $x \in E$ und $\varepsilon > 0$. Dann ist $B_\varepsilon(f(x))$ offen, damit also $f^{-1}(B_\varepsilon(f(x)))$ nach Voraussetzung ebenfalls offen. Da $x \in f^{-1}(B_\varepsilon(f(x)))$, gibt es also ein $\delta > 0$, so dass $B_\delta(x) \subseteq f^{-1}(B_\varepsilon(f(x)))$, also $f(B_\delta(x)) \subseteq B_\varepsilon(f(x))$. Damit gilt $r'(f(x), f(y)) < \varepsilon$ für alle y mit $r(x, y) < \delta$. Damit ist f in x stetig nach Definition. Da x beliebig war, folgt die Stetigkeit von f.
2.⇔3.: Die Menge A' ist genau dann abgeschlossen, wenn $O' := (A')^c$ offen ist. Gilt 2., so ist $f^{-1}((A')^c) = (f^{-1}(A'))^c$ offen und damit ist $f^{-1}(A')$ abgeschlossen. Gilt 3. so ist $f^{-1}((O')^c) = (f^{-1}(O'))^c$ abgeschlossen und damit ist $f^{-1}(O')$ offen. □

5.3 Rechenregeln für stetige Funktionen

Proposition 5.14. *Seien $f, g : E \to I$ für einen metrischen Raum(E, r), $I \subseteq \mathbb{R}$ und $x \in E$. Sind f und g stetig in x, so sind auch $-f$, $|f|$, $f + g$, fg, $f \wedge g$ und $f \vee g$ in x_0 stetig. Ist $g(x) \neq 0$, so ist auch f/g in x stetig.*

Beweis. Wir benutzen Propositionen 5.12 und 3.7. Sei $(x_n)_{n=1,2,...}$ eine E-wertige Folge mit $x_n \xrightarrow{n\to\infty} x$. Dann ist $f(x_n) \xrightarrow{n\to\infty} f(x)$ und $g(x_n) \xrightarrow{n\to\infty} g(x)$ nach Voraussetzung. Damit gilt auch $|-f(x_n) - (-f(x))| \xrightarrow{n\to\infty} 0$, also $-f(x_n) \xrightarrow{n\to\infty} -f(x)$. Nach Proposition 3.7 gilt auch $|f(x_n)| \xrightarrow{n\to\infty} |f(x)|$, $f(x_n) + g(x_n) \xrightarrow{n\to\infty} f(x) + g(x)$, $f(x_n) \cdot g(x_n) \xrightarrow{n\to\infty} f(x) \cdot g(x)$ und im Fall $g(x) \neq 0$ auch $f(x_n)/g(x_n) \xrightarrow{n\to\infty} f(x)/g(x)$. Zuletzt ist $|f(x_n) \vee g(x_n) - f(x) \vee g(x)| \leq |f(x_n) - f(x)| + |g(x_n) - g(x)| \xrightarrow{n\to\infty} 0$. □

Definition 5.15 (Polynome und rationale Funktionen).

1. *Funktionen $p : \mathbb{R} \to \mathbb{R}$ der Form $p(x) = \sum_{k=0}^n a_k x^k$ heißen* Polynome. *Ist $a_n \neq 0$, so sagen wir, das Polynom ist vom Grad n.*
2. *Sind p und q zwei Polynome. Dann heißt p/q eine* rationale *Funktion.*

Korollar 5.16 (Stetigkeit von Polynomen und rationalen Funktionen). *Jedes Polynom $p : \mathbb{R} \to \mathbb{R}$ ist stetig. Jede rationale Funktion $f = p/q$, wobei p und q Polynome sind, ist für $x \in \mathbb{R} \setminus \{x : q(x) = 0\}$ stetig.*

Beweis. Da ein Polynom aus der Hintereinanderausführung von Produktbildung und Addition aus der Identitätsabbildung entsteht, ist die erste Behauptung wegen Proposition 5.14 klar. Die zweite folgt aus derselben Proposition. $\square$

Proposition 5.17 (Kompositionen stetiger Abbildungen sind stetig).
Seien $(E, r), (E', r'), (E'', r'')$ metrische Räume, $f : E \to E', g : E' \to E''$ und $x \in E$. Sind f stetig in x und g stetig in $f(x)$, so ist $g \circ f : E \to E''$ stetig in x.

Beweis. Sei U'' eine Umgebung von $g(f(x))$. Dann ist wegen der Stetigkeit von g die Menge $g^{-1}(U'')$ eine Umgebung von $f(x)$. Damit ist wegen der Stetigkeit von f die Menge $f^{-1}(g^{-1}(U'')) = (g \circ f)^{-1}(U'')$ eine Umgebung von x. Damit folgt die Behauptung aus Proposition 5.12. $\square$

Korollar 5.18. *Die Abbildungen $x \mapsto x^{k/\ell}$ sind stetig auf $[0, \infty)$, $k, \ell = 1, 2, ...$*

Beweis. Wir zeigen zunächst die Stetigkeit von $x \mapsto x^{1/\ell}$. Hierfür benötigen wir, dass $y^\ell - x^\ell > (y - x)^\ell$ für $y > x > 0$ gilt. Wir schreiben

$$y^\ell - x^\ell = y^{\ell-1}(y - x) + x(y^{\ell-1} - x^{\ell-1}) > (y - x)^{\ell-1}(y - x) = (y - x)^\ell.$$

Daraus folgt auch $y^{1/\ell} - x^{1/\ell} < (y - x)^{1/\ell}$, wenn man x durch $x^{1/\ell}$ und y durch $y^{1/\ell}$ ersetzt. Sei nun $\varepsilon > 0$ und $\delta := \varepsilon^\ell$. Ist nun $x, y \geq 0$ mit $|y - x| < \delta$, so folgt $|y^{1/\ell} - x^{1/\ell}| \leq |y - x|^{1/\ell} \leq \varepsilon$. Daraus folgt die Stetigkeit von $x \mapsto x^{1/\ell}$. Nun folgt die Stetigkeit von $x \mapsto x^{k/\ell}$ aus Proposition 5.17, sowie der Stetigkeit von $x \mapsto x^k$ und von $x \mapsto x^{1/\ell}$. $\square$

5.4 Der Zwischenwertsatz

Theorem 5.19 (Der Zwischenwertsatz). *Sei $a < b$ und $f : [a, b] \to \mathbb{R}$ eine stetige Abbildung sowie $f(a) \leq y \leq f(b)$ oder $f(b) \leq y \leq f(a)$. Dann gibt es ein $x \in [a, b]$ mit $f(x) = y$.*

Beweis. OBdA sei $f(a) < f(b)$. Wir konstruieren x mittels einer Intervallschachtelung. Wir beginnen mit dem Intervall $[a_0, b_0] := [a, b]$. Setze $x_1 := (a_0 + b_0)/2$. Ist $f(x_1) < y$, so setze $[a_1, b_1] = [x_1, b_0]$, andernfalls $[a_1, b_1] = [a_0, x_1]$. Setze nun $x_2 := (a_1 + b_1)/2$ und $[a_2, b_2] = [x_2, b_1]$ falls $f(x_2) < y$, andernfalls $[a_1, b_1] = [a_0, x_2]$ usw. So erreichen wir, dass der Funktionswert am linken Intervallrand immer zu klein, der am rechten Intervallrand zu groß ist. Klar ist, dass es ein einziges x gibt, das in allen Intervallen liegt. Da f stetig ist, gilt

$$y \leq \lim_{n\to\infty} f(b_n) = f(x) = \lim_{n\to\infty} f(a_n) \leq y$$

wegen der Stetigkeit von f und da $a_n \xrightarrow{n\to\infty} x$ und $b_n \xrightarrow{n\to\infty} x$. Daraus folgt die Behauptung. $\square$

Korollar 5.20 (Nullstellensatz). *Sei $a < b$ und $f : [a, b] \to \mathbb{R}$ eine stetige Abbildung sowie $f(a) < 0 < f(b)$. Dann gibt es ein $x \in [a, b]$ mit $f(x) = 0$.*

Beweis. Klar nach Theorem 5.19. $\square$

Korollar 5.21. *Sei $a > 0$ und $n \in \mathbb{N}$. Dann gibt es genau ein $x \geq 0$ mit $x^n = a$.*

Beweis. Wegen der Monotonie von $x \mapsto x^n$ ist klar, dass es höchstens ein x mit $x^n = a$ geben kann. Betrachtet man die Funktion $f : x \mapsto x^n - a$ auf $[0, a+1]$, so stellt man fest, dass dies eine stetige Funktion mit $f(0) < 0$ und $f(a+1) = (a+1)^n - a \geq 1 + an - a \geq 1$ gilt. Damit folgt die Aussage aus Theorem 5.19. □

5.5 Stetige Funktionen auf kompakten Mengen

Definition 5.22 ((Folgen-)Kompaktheit). *Sei (E, r) ein metrischer Raum. Eine Menge $K \subseteq E$ heißt* (folgen-)kompakt*, wenn jede K-wertige Folge $(x_n)_{n=1,2,\dots}$ eine Teilfolge $(x_{k_n})_{n=1,2,\dots}$ hat die gegen ein $x \in K$ konvergiert.*

Lemma 5.23. *Sei (E, r) ein metrischer Raum und $K \subseteq E$ kompakt. Dann ist K abgeschlossen.*

Beweis. Klar nach Lemma 5.11.5. □

Theorem 5.24 (Satz von Heine-Borel). *Für $K \subseteq \mathbb{R}^d$ (versehen mit der euklidischen Metrik) sind äquivalent:*

1. *K ist kompakt.*
2. *K ist abgeschlossen und beschränkt.*
3. *K hat die Heine-Borel Überdeckungseigenschaft, das heißt: für jede (unendliche) Überdeckung von K, also $(A_i)_{i\in I}$ mit (unendlicher) Indexmenge I und $K \subseteq \bigcup_{i\in I} A_i$ gibt es eine endliche Menge $J \subseteq I$ mit $K \subseteq \bigcup_{j\in J} A_j$.*

Beweis. 1.⇒2. Angenommen, K ist nicht beschränkt. Dann gibt es eine Folge $(x_n)_{n=1,2,\dots}$ mit $r(x_n, 0) \xrightarrow{n\to\infty} \infty$. Dies gilt auch für jede Teilfolge. Insbesondere ist K nicht kompakt.
2.⇒1. Sei $(x_n)_{n=1,2,\dots}$ eine K-wertige Folge. Da K beschränkt ist, gibt es eine konvergente Teilfolge nach dem Satz von Bolzano-Weierstrass, Theorem 3.13. Da K abgeschlossen ist, ist der Grenzwert nach Lemma 5.11.5 ebenfalls in K. Daraus folgt die Kompaktheit von K.
2.⇒3. Angenommen, K hat nicht die Heine-Borel'sche Überdeckungseigenschaft, ist aber abgeschlossen und beschränkt. Dann gibt es eine Familie $(A_i)_{i\in I}$ mit $K \subseteq \bigcup_{i\in I} A_i$, zu der es keine endliche Teilüberdeckung gibt. Da K beschränkt ist, gibt es ein Intervall $[a_0, b_0]$ mit $K \subseteq [a_0, b_0]$. Durch Halbieren des Intervalls erhält man zwei neue Intervalle. Nun besitzt entweder $K \cap [a_0, (a_0 + b_0)/2]$ oder $K \cap [(a_0 + b_0)/2, b_0]$ keine endliche Teilüberdeckung. (Ansonsten ließen sich diese Teilüberdeckungen ja zu einer endlichen Teilüberdeckung von K bereinigen.) Durch weiteres Halbieren erhalten wir eine absteigende Folge von Intervallen $[a_n, b_n]$ (mit $b_n - a_n \downarrow 0$) mit der Eigenschaft

$$K \cap [a_n, b_n] \text{ kann nicht durch endlich viele der } A_i \text{ überdeckt werden.} \tag{5.1}$$

Der durch diese Intervallschachtelung eindeutig definierte Punkt $x \in K \bigcap_{n\in\mathbb{N}} [a_n, b_n]$ (man erinnere sich dass K abgeschlossen ist, woraus $x \in K$ folgt) wird sicher von einem der A_i überdeckt. Da A_i offen ist, gibt es ein $\varepsilon > 0$ mit $B_\varepsilon(x) \subseteq A_i$. Nun gilt also für N groß genug mit $b_N - a_N < \varepsilon$, dass $K \cap [a_N, b_N] \subseteq K \cap B_\varepsilon(x) \subseteq K \cap A_i$. Insbesondere gibt es eine endliche Teilüberdeckung von $K \cap [a_N, b_N]$ im Widerspruch zu (5.1).

3.⇒2. Zunächst stellen wir fest, dass K beschränkt ist. Schließlich ist $K \subseteq \bigcup_{n\in\mathbb{N}}(-n,n)$, also gibt es nach der Überdeckungseigenschaft ein N mit $K \subseteq (-N,N)$. Wir müssen also zeigen, dass K abgeschlossen ist. Angenommen, K wäre nicht abgeschlossen. Dann gibt es eine Folge $(x_n)_{n=1,2,\dots}$ in A mit $x_n \xrightarrow{n\to\infty} x \notin A$. Wir setzen nun $A_y := B_{|y-x|/2}(y)$ als den offenen Ball um y mit Radius $|y-x|/2$. Dann ist offenbar $\bigcup_{y\in K} A_y$ eine offene Überdeckung von K mit $x \notin \bigcup_{y\in K} A_y$ (denn schließlich liegt x ja gerade außerhalb des Balles A_y nach Konstruktion). Da K die Heine-Borel'sche Überdeckungseigenschaft hat, gibt es $y_1,\dots,y_n \in K$ mit $K \subseteq \bigcup_{j=1}^n B_{|y_j-x|/2}(y_j)$. Sei nun $\varepsilon := \min_{j=1,\dots,n}|y_j-x|/2$. Dann ist $B_\varepsilon(x)\cap\bigcup_{j=1}^n B_{|y_j-x|/2}(y_j) = \bigcup_{j=1}^n B_\varepsilon(x)\cap B_{|y_j-x|/2}(y_j) = \emptyset$. Allerdings ist $x_n \in B_\varepsilon(x)\cap\bigcup_{j=1}^n B_{|y_j-x|/2}(y_j)$ für fast alle n, da $x_n \xrightarrow{n\to\infty} x$ und $x_n \in K \subseteq \bigcup_{j=1}^n B_{|y_j-x|/2}(y_j)$. Dies ist ein Widerspruch. □

Theorem 5.25. *Seien $(E,r),(E',r')$ metrische Räume und $f : E \to E'$ stetig. Ist $K \subseteq E$ kompakt, dann ist auch $f(K) = \{f(x) : x \in K\}$ kompakt.*

Beweis. Sei $(y_n)_{n=1,2,\dots}$ eine $f(K)$-wertige Folge. Dann gibt es eine K-wertige Folge $(x_n)_{n=1,2,\dots}$ mit $f(x_n) = y_n$. Da K kompakt ist, gibt es eine Teilfolge $(x_{k_n})_{n=1,2,\dots}$ und ein $x \in K$ mit $x_{k_n} \xrightarrow{n\to\infty} x$. Also gilt wegen der Stetigkeit von f auch

$$f(K) \ni f(x) = \lim_{n\to\infty} f(x_{k_n}) = \lim_{n\to\infty} y_{k_n}.$$

Damit folgt die Aussage, da $(y_n)_{n=1,2,\dots}$ beliebig war. □

Korollar 5.26. *Sei (E,r) ein metrischer Raum und $f : E \to \mathbb{R}$ eine stetige Funktion. Ist $K \subseteq E$ kompakt, so gibt es $x \in K$ mit $f(x) = \sup_{y\in K} f(y)$. Das bedeutet, dass f sein Supremum auf K annimmt. Analoges gilt für das Infimum.*

Beweis. Nach Theorem 5.25 ist $f(K)$ kompakt, also nach Theorem 5.24 abgeschlossen und beschränkt. Also gibt es wegen der Beschränktheit nach dem Satz von Bolzano-Weierstrass $m := \sup f(K)$. Sei nun $(x_n)_{n=1,2,\dots}$ eine K-wertige Folge mit $f(x_n) \xrightarrow{n\to\infty} m$. Also gibt es wegen Lemma 5.11.5 und der Abgeschlossenheit von $f(K)$ ein $y \in f(K)$ mit $m = f(K)$. Also gibt es $x \in K$ mit $f(x) = m$. □

Definition 5.27 (Gleichmäßige Stetigkeit). *Seien (E,r) und (E',r') metrische Räume. Dann heißt eine Funktion $f : E \to E'$* gleichmäßig stetig, *falls es für jedes $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass $r'(f(x),f(y)) < \varepsilon$ für alle $x,y \in E$ mit $r(x,y) < \delta$. In Quantorenschreibweise:*

$$\forall\varepsilon > 0\ \exists\delta > 0\ \forall x,y \text{ mit } r(x,y) < \delta : r'(f(x),f(y)) < \varepsilon.$$

Proposition 5.28 (Stetige Funktionen auf Kompakta sind gleichmäßig stetig). *Seien (E,r) und (E',r') metrische Räume, $K \subseteq E$ kompakt und $f : K \to E'$ stetig. Dann ist f sogar gleichmäßig stetig.*

Beweis. Angenommen, f wäre nicht gleichmäßig stetig. Dann gibt es ein $\varepsilon > 0$ und Paare (x_n,x_n') mit $r(x_n,x_n') < 1/n$ und $r'(f(x_n),f(x_n')) > \varepsilon$. Da K kompakt ist, gibt es eine Teilfolge $(x_{k_n})_{n=1,2,\dots}$ und $x \in K$ mit $x_{k_n} \xrightarrow{n\to\infty} x$. Da $r(x_n,x_n') < 1/n$, folgt damit auch $r(x_n',x) \le r(x_n',x_n) + r(x_n,x) \xrightarrow{n\to\infty} 0$. Insgesamt können wir also wegen der Stetigkeit von f schreiben

$$\varepsilon \le r'(f(x_n),f(x_n')) \xrightarrow{n\to\infty} r'(f(x),f(x)) = 0.$$

Dies ist ein Widerspruch, also muss f gleichmäßig stetig sein. □

Beispiel 5.29. 1. Die Funktion $x \mapsto 1/x$ ist auf $[t, \infty)$ gleichmäßig stetig für jedes $t > 0$. Denn: Sei $\varepsilon > 0$. Dann wählen wir $\delta > 0$ so, dass

$$\frac{1}{t} - \frac{1}{t+\delta} < \varepsilon, \text{ also } \delta < \varepsilon t(t+\delta).$$

Sei also $\delta := \varepsilon t^2$. Dann gilt für $t \leq x < x'$ mit $x' - x < \delta$, dass

$$\frac{1}{x} - \frac{1}{x'} = \frac{x' - x}{xx'} < \frac{\delta}{t^2} = \varepsilon.$$

Daraus folgt die gleichmäßige Stetigkeit auf $[t, \infty)$.

2. Andererseits gilt: Die Funktion $x \mapsto 1/x$ ist auf $(0, \infty)$ nicht gleichmäßig stetig.

   Denn: Es gilt für $0 < x < x'$ mit $x' - x = \delta$

$$\frac{1}{x} - \frac{1}{x'} = \frac{x' - x}{xx'} = \frac{\delta}{x(x+\delta)}.$$

   Nun ist die Funktion $x \mapsto \frac{\delta}{x(x+\delta)}$ auf $(0, \infty)$ unbeschränkt. Das bedeutet, dass es für jedes $\delta > 0$ und jedes $K > 0$ ein $x \in (0, \infty)$ gibt, so dass $f(x) - f(x + \delta) > K$ gilt. Insbesondere ist f nicht gleichmäßig stetig auf $(0, \infty)$.

5.6 Stetige Fortsetzung von Funktionen

In diesem Abschnitt seien $(E, r), (E', r')$ metrische Räume. Nun beschäftigen wir uns mit folgender Fragestellung: Seien $(E, r), (E', r')$ metrische Räume, $D \subseteq E$ und $f : D \to E'$ stetig sowie $x_0 \in D^c$. Unter welchen Bedingungen gibt es eine stetige Funktion $g : D \cup \{x_0\} \to E'$, so dass $f(x) = g(x)$ für $x \in D$? Wir sagen dann auch, dass g eine stetige Fortsetzung von f ist. Wir beginnen mit der Definition eines Häufungspunktes einer Menge.

Definition 5.30 (Häufungspunkt, isolierter Punkt). *Sei (E, r) ein metrischer Raum und $D \subseteq E$. Dann heißt $x \in E$ ein* Häufungspunkt *von D, wenn $|B_\varepsilon(x) \cap D| = \infty$ für jedes $\varepsilon > 0$ gilt. Weiter heißt $x \in E$* isolierter Punkt *von D, wenn es ein $\varepsilon > 0$ gibt mit $0 < |B_\varepsilon(x) \cap D| < \infty$. Wir schreiben*

$$\overline{D} := \{x \in E : \ x \textit{ Häufungspunkt oder } x \textit{ isolierter Punkt}\}$$

und bezeichnen mit $\overline{D}$ den Abschluss *von D.*

Beispiel 5.31. 1. Es ist 0 ein Häufungspunkt von $(0, \infty)$. Außerdem hat $(0, \infty)$ keine isolierten Punkte.

2. Die Menge $\mathbb{Z} \subseteq \mathbb{R}$ besteht nur aus isolierten Punkten.

3. Sei $(x_n)_{n=1,2,\ldots}$ die reellwertige Folge mit $x_{2n} = 1, x_{2n+1} = 1/n$. Dann sind 0 und 1 Häufungswerte der Folge $(x_n)_{n=1,2,\ldots}$, jedoch ist nur 0 Häufungspunkt der Menge $D := \{x_n : n \in \mathbb{N}\}$.
   Denn: Es gilt $B_\varepsilon(1) \cap D = \{1\}$ falls ε klein genug ist, aber $x_{2n} \xrightarrow{n\to\infty} 1$.

4. Sei $E = \mathbb{R}$ und $D = \{1/n : n \in \mathbb{N}\}$. Dann besteht D nur aus isolierten Punkten und 0 ist der einzige Häufungspunkt von D.
   Denn Für $n \in \mathbb{N}$ ist $\frac{1}{n} - \frac{1}{n+1} = \frac{1}{n(n+1)} < \frac{1}{(n+1)^2}$ und $\frac{1}{n-1} - \frac{1}{n} < \frac{1}{n^2}$. Also ist $B_{1/n^2}(1/n) \cap D = \{1/n\}$.

Lemma 5.32. *Sei (E, r) ein metrischer Raum und $D \subseteq E$.*

1. *Es ist $x \in E$ genau dann ein Häufungspunkt von D, wenn es eine D-wertige Folge $(x_n)_{n=1,2,...}$ mit $x_n \neq x_m, m \neq n, m, n = 1, 2, ...$ gibt mit $x_n \xrightarrow{n\to\infty} x$.*

2. *Sei $x \in E$, $\varepsilon > 0$ und $|B_\varepsilon(x) \cap D| < \infty$. Dann gibt es ein $\delta > 0$ mit $|B_\delta(x) \cap D| \in \{0, 1\}$. Insbesondere ist $x \in D$ für jeden isolierten Punkt von D.*

3. *Die Menge $\overline{D}$ ist abgeschlossen und es gilt*

$$\overline{D} = \bigcap_{\substack{D' \supseteq D \\ D' \text{ abgeschlossen}}} D'.$$

Beweis. 1. Sei x ein Häufungspunkt von D. Dann gibt es nach Definition für jedes n ein $x_n \in D$ mit $x_n \in B_{1/n}(x)$. OBdA ist dabei $x_n \neq x_1, ..., x_{n-1}$. Damit gilt auch $x_n \xrightarrow{n\to\infty} x$. Andersherum nehmen wir an, dass es eine D-wertige Folge $(x_n)_{n=1,2,...}$ unterschiedlicher Punkte gibt mit $x_n \xrightarrow{n\to\infty} x$. Für $\varepsilon > 0$ ist damit $x_n \in B_\varepsilon(x)$ nach Lemma 5.11.3 für fast alle n. Dies impliziert auch $|B_\varepsilon(x) \cap D| = \infty$.
2. Ist $x \in D$ und $|B_\varepsilon(x) \cap D| < \infty$. Dann gibt es ein $n \in \mathbb{N}$ und $x_1, ..., x_n \in D$ mit $B_\varepsilon(x) \cap D = \{x, x_1, ..., x_n\}$. Nun genügt es, $\delta := \min\{r(x, x_1), ..., r(x, x_n)\}/2$ zu setzen. Im Fall $x \in D^c$ setzt man $\delta := \min\{r(x, y) : y \in D\}/2 > 0$. Dann ist $|B_\delta(x) \cap D| = 0$.
3. Zunächst zeigen wir die Abgeschlossenheit von $\overline{D}$. Angenommen, $\overline{D}$ ist nicht abgeschlossen. Dann ist $\overline{D}^c$ nicht offen. Also gibt es ein $x \in \overline{D}^c$, so dass $\overline{D}^c$ keine Umgebung von x ist. Da $D \subseteq \overline{D}$, kann x kein isolierter Punkt von D sein, darf also auch kein Häufungspunkt von D sein. Hierzu werden wir einen Widerspruch konstruieren: Es gibt nämlich nach Lemma 5.11.4 eine $\overline{D}$-wertige Folge $(x_n)_{n=1,2,...}$, so dass $r(x_n, x) < 1/n$ und damit $x_n \xrightarrow{n\to\infty} x$ gilt. Da $x_n \in \overline{D}$, gibt es für jedes n auch $y_n \in D$ mit $r(x_n, y_n) < 1/n$. Damit folgt auch $r(x, y_n) \leq r(x, x_n) + r(x_n, y_n) \xrightarrow{n\to\infty} 0$. Also gilt $y_n \xrightarrow{n\to\infty} x$, insbesondere ist x nach 1. ein Häufungspunkt von D im Widerspruch zur Annahme.

Nun kommen wir zur angegebenen Formel für $\overline{D}$. Nach eben gezeigtem ist '$\supseteq$' klar, da die rechte Seite nach Lemma 5.9.2 eine abgeschlossene Menge ist. Um '$\subseteq$' zu zeigen sei nun $x \in \overline{D}$. Dann ist $0 < |B_\varepsilon(x) \cap D| \leq \infty$ für jedes $\varepsilon > 0$. Angenommen, es gäbe ein abgeschlossenes $D' \supseteq D$, so dass $x \notin D'$, also $x \in (D')^c$. Da $(D')^c$ offen ist, muss es ein $\varepsilon > 0$ geben mit $B_\varepsilon(x) \subseteq (D')^c$, also $B_\varepsilon(x) \cap D' = \emptyset$. Dann wäre aber auch $B_\varepsilon(x) \cap D \subseteq B_\varepsilon(x) \cap D' = \emptyset$ im Widerspruch zu $x \in \overline{D}$. □

Beispiel 5.33. 1. Sei $E = \mathbb{R}$ und $D = (a, b)$ mit $a, b \in \mathbb{R}$. Dann ist $\overline{(a, b)} = [a, b]$.
   Denn: Das abgeschlossene Intervall $[a, b]$ ist die kleinste abgeschlossene Obermenge von (a, b).

2. Es ist $\overline{\mathbb{Q}} = \mathbb{R}$.
   Denn: Für jedes $x \in \mathbb{R}$ gibt es eine $\mathbb{Q}$-wertige Folge $(x_n)_{n=1,2,...}$ mit $x_n \xrightarrow{n\to\infty} x$.

Definition 5.34 (Stetige Fortsetzung). *Seien $(E, r), (E', r')$ metrische Räume, $D \subseteq E$ und $f : D \to E'$ stetig sowie $x_0 \in D^c$. Gibt es eine stetige Funktion $g : D \cup \{x_0\} \to E'$, so dass $f(x) = g(x)$ für $x \in D$, so sagen wir, dass g eine* stetige Fortsetzung *von f auf $D \cup \{x_0\}$ ist. Wir schreiben auch $g|_D = f$, wobei $g|_D : D \to E'$ die Einschränkung von g auf D ist.*

Proposition 5.35 (Eindeutigkeit der stetigen Fortsetzung). *Seien $(E, r), (E', r')$ metrische Räume, $D \subseteq E$ und $f : D \to E'$ stetig sowie $x_0 \in D^c$.*

1. *Ist x_0 kein Häufungspunkt von D, so ist jede Funktion $g : D \cup \{x_0\} \to E'$ mit $g(x) = f(x)$, $x \in D$ eine stetige Fortsetzung von f.*

2. *Ist x_0 ein Häufungspunkt von D, so gibt es höchstens eine stetige Fortsetzung von f auf $D \cup \{x_0\}$.*

Beweis. 1. Zu zeigen ist nur, dass jede Funktion $g : D \cup \{x_0\} \to E'$ mit $g(x) = f(x), x \in D$ stetig in x_0 ist. Nach Lemma 5.32.2 gibt es ein $\delta > 0$ mit $B_\delta(x_0) \cap D = \emptyset$. Ist nun $(x_n)_{n=1,2,...}$ eine $D \cup \{x_0\}$-wertige Folge mit $x_n \xrightarrow{n\to\infty} x_0$. Dann ist $x_n = x_0$ für fast alle n und damit auch $g(x_n) = g(x_0)$ für fast alle n. Insbesondere ist g in x_0 stetig.
2. Nach Lemma 5.32.1 gibt es eine D-wertige Folge $(x_n)_{n=1,2,...}$ mit $x_n \xrightarrow{n\to\infty} x_0$. Seien $g, g' : D \cup \{x_0\} \to E'$ zwei stetige Fortsetzungen von f auf $D \cup \{x_0\}$. Dann gilt

$$g(x_0) = \lim_{n\to\infty} g(x_n) = \lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} g'(x_n) = g'(x_0).$$

□

Bisher haben wir Grenzwerte der Form $x_n \xrightarrow{n\to\infty} x$ innerhalb von metrischen Räumen betrachtet. Grundlegend war hier immer, dass $(x_n)_{n=1,2,...}$ eine Folge war. Nun erweitern wir diesen Begriff, insdem wir sogar Grenzwerte von Funktionen betrachten. Dies werden wir jedoch in Lemma 5.37 wieder auf den Konvergenzbegriff von Folgen zurückführen.

Definition 5.36 (Grenzwert einer Funktion). *Seien $(E, r), (E', r')$ metrische Räume, $D \subseteq E$ und x_0 ein Häufungspunkt von D. Ist die Funktion*

$$F : \begin{cases} f(x), & \text{falls } x \in D \setminus \{x_0\}, \\ y, & \text{falls } x = x_0 \end{cases} \tag{5.2}$$

stetig in x_0, so sagen wir, dass f in x_0 den Grenzwert y hat und schreiben

$$\lim_{x\to x_0} f(x) = y \text{ oder auch } f(x) \xrightarrow{x\to x_0} y.$$

Ist $x_0 \in D$ und ist f stetig in x_0, so gilt also $f(x) \xrightarrow{x\to x_0} f(x_0)$.

Lemma 5.37 (Äquivalente Formulierungen für Grenzwerte von Funktionen). *Seien $(E, r), (E', r')$ metrische Räume, $D \subseteq E$ und x_0 ein Häufungspunkt von D. Ist $f : D \to E'$ und $y \in E'$, so sind äquivalent:*

1. $f(x) \xrightarrow{x\to x_0} y$

2. *Für jedes $\varepsilon > 0$ gibt es ein $\delta > 0$, so dass $r(f(x), y) < \varepsilon$ für alle $x \in D$ mit $x \neq x_0$ und $r(x, x_0) < \delta$.*

3. *Für jede $D \setminus \{x_0\}$-wertige Folge $(x_n)_{n=1,2,\ldots}$ mit $x_n \xrightarrow{n\to\infty} x_0$ ist $f(x_n) \xrightarrow{n\to\infty} y$.*

Beweis. 1.⇒2.: Sei F wie in (5.2). Da F nach Voraussetzung in x_0 stetig ist, und $r(f(x), y) = r(F(x), F(x_0)$ nach Definition gilt, folgt die Behauptung.
2.⇒3. und 3.⇒1.: Klar nach Proposition 5.12. □

Proposition 5.38 (Rechenregeln für Grenzwerte von Funktionen).
Seien $(E, r), (E', r'), (E'', r'')$ metrische Räume, $f, g : E \to E', h : E' \to E''$. Weiter sei $x_0 \in E$ und $f(x) \xrightarrow{x\to x_0} a, g(x) \xrightarrow{x\to x_0} b$.

1. *Ist $E' \subseteq \mathbb{R}$, so gilt*

$$\begin{aligned} f(x) + g(x) &\xrightarrow{x\to x_0} a + b, \\ f(x)g(x) &\xrightarrow{x\to x_0} ab, \\ \frac{f(x)}{g(x)} &\xrightarrow{x\to x_0} \frac{a}{b}, \text{ falls } b \neq 0. \end{aligned}$$

   Gibt es ein $\varepsilon > 0$, so dass $f(x) \leq g(x)$ für alle $x \neq x_0, x \in B_\varepsilon(x_0)$, so gilt $a \leq b$.

2. *Ist h stetig in a, so gilt $h(f(x)) \xrightarrow{x\to x_0} h(a)$.*

Beweis. 1. Alle Aussagen folgen aus Lemma 5.37.3 unter Anwendung der entsprechenden Aussage über Folgen, Proposition 3.7, Proposition 3.9. 2. folgt aus Proposition 5.17. □

Während Grenzwerte der Form $\lim_{x\to x_0} f(x)$ nach Lemma 5.37 äquivalent sind zu $\lim_{n\to\infty} f(x_n)$ für alle Folgen $(x_n)_{n=1,2,\ldots}$ mit $\lim_{n\to\infty} x_n = x_0$, können wir die Auswahl dieser Folgen auch einschränken. Dies führt nun zum Begriff des einseitigen Grenzwertes. Anschließend betrachten wir noch uneigentliche Grenzwerte von Funktionen, wie wir dies im Rahmen der Konvergenz von Folgen schon in Abschnitt 3.4 getan haben.

Definition 5.39 (Einseitige Grenzwerte). *Sei (E', r') ein metrischer Raum, $D \subseteq \mathbb{R}$ und $f : D \to E'$.*

1. *Ist x_0 ein Häufungspunkt von $D^- := D \cap (-\infty, x_0)$, und gilt[12] $f|_{D^-}(x) \xrightarrow{x\to x_0} y$, so sagen wir, dass f in x_0 den linksseitigen Grenzwert y hat und schreiben dafür*

$$f(x_0-) = \lim_{x \uparrow x_0} f(x) = y.$$

2. *Ist x_0 ein Häufungspunkt von $D^+ := D \cap (x_0, \infty)$, und gilt $f|_{D^+}(x) \xrightarrow{x\to x_0} y$, so sagen wir, dass f in x_0 den rechtsseitigen Grenzwert y hat und schreiben dafür*

$$f(x_0+) = \lim_{x \downarrow x_0} f(x) = y.$$

[12] Sei $f : E \to E'$ und $D \subseteq E$. Dann ist

$$f|_D : \begin{cases} D & \to E' \\ x & \mapsto f(x) \end{cases}$$

die Einschränkung von f auf D.

Beispiel 5.40. Wir betrachten die Funktion $f : x \mapsto [x]$ aus Definition 2.24. Es ist f für alle $x \in \mathbb{R} \setminus \mathbb{Z}$ stetig. Außerdem gilt für $x_0 \in \mathbb{Z}$

$$\lim_{x \uparrow x_0} f(x) = x - 1, \qquad \lim_{x \downarrow x_0} f(x) = x.$$

Definition 5.41 (Grenzwerte im Unendlichen).

1. *Sei (E', r') ein metrischer Raum, $D \subseteq \mathbb{R}$ nicht nach oben beschränkt und $f : D \to E'$. Gilt $f(x_n) \xrightarrow{n\to\infty} y$ für jede Folge $(x_n)_{n=1,2,\dots}$ mit $x_n \xrightarrow{n\to\infty} \infty$, so sagen wir, dass f in ∞ den Grenzwert y besitzt und schreiben $\lim_{x\to\infty} f(x) = y$ oder $f(x) \xrightarrow{x\to\infty} y$. Analog definieren wir den Grenzwert für $x \to -\infty$.*

2. *Sei (E, r) ein metrischer Raum, $D \subseteq E$, x_0 ein Häufungspunkt von D und $f : D \to D' \subseteq \mathbb{R}$. Gilt $f(x_n) \xrightarrow{n\to\infty} \infty$ für jede D-wertige Folge $(x_n)_{n=1,2,\dots}$ mit $x_n \xrightarrow{n\to\infty} x_0$, so sagen wir, dass f in x_0 den Grenzwert ∞ besitzt und schreiben $\lim_{x\to x_0} f(x) = \infty$ oder $f(x) \xrightarrow{x\to x_0} \infty$. Analog definieren wir den Grenzwert für $x \to -\infty$.*

Lemma 5.42 (Reduktionslemma). *Sei (E', r') ein metrischer Raum, $D \subseteq \mathbb{R}$ nicht nach oben beschränkt und $f : D \to E'$. Weiter definieren wir $g : \widetilde{D} := \{1/x : x \in D\} \to E'$ mittels $g(x) = f(1/x)$. Dann sind äquivalent:*

1. *f besitzt in ∞ den Grenzwert y.*

2. *Es gilt $g(0+) = y$.*

Analoges gilt für Grenzwerte bei $-\infty$.

Beweis. Ist $(x_n)_{n=1,2,\dots}$ eine D-wertige Folge mit $x_n \xrightarrow{n\to\infty} \infty$, so ist $(1/x_n)_{n=1,2,\dots}$ eine $\widetilde{D}^+$-wertige Folge mit $1/x_n \xrightarrow{n\to\infty} 0$. Deswegen folgt die Behauptung aus Lemma 5.37. □

Bemerkung 5.43. Die Rechenregeln aus Proposition 5.38 gelten für einseitige Grenzwerte und Grenzwerte im Unendlichen, falls sie existieren, entsprechend.

6 Die Exponentialfunktion und von ihr abgeleitete Funktionen

Wir kommen nun dazu, einige spezielle Funktionen einzuführen. Vor allem wird die Euler'sche Zahl e und die Exponentialfunktion eine wichtige Rolle spielen. Von ihr werden wir in Abschnitt 6.2 die Logarithmus-Funktion ableiten. Anschließend erweitern wir die Definition der Exponentialfunktion in Abschnitt 6.3 auf andere Basen, ebenso erweitern wir die Potenzfunktion in Abschnitt 6.4 auf reellwertige Potenzen. Abschließend leiten in Abschnitt 6.5 wir die Winkelfunktionen $\cos, \sin, \tan$ und ihre Umkehrfunktionen her.

6.1 Die Exponentialfunktion

In diesem Abschnitt suchen wir Funktionen f, die die Gleichung

$$f(x + y) = f(x) \cdot f(y) \tag{6.1}$$

erfüllen. Dies kann man in etwa so verstehen: Angenommen, der Bestand einer Population vermehrt sich in einer Zeit x um einen Faktor $f(x)$. Betrachten wir nun den Bestand zur Zeit $x + y$. Einerseits ist dieser um einen Faktor $f(x + y)$ größer ist als zur Zeit 0. Andererseits vermehrt sich zwischen Zeit x und $x+y$ die Population um einen Faktor $f(y)$. Insgesamt ist also der Vermehrungsfaktor zwischen Zeiten 0 und $x+y$ auch gleich $f(x)\cdot f(y)$.

Wir werden sehen, dass die Gleichung (6.1) unendlich viele Lösungen hat. Zunächst untersuchen wir folgende Eigenschaften für eine Funktion $f:\mathbb{R}\to\mathbb{R}$:

$$\begin{aligned} &(1) && f(x+y)=f(x)\cdot f(y),\\ &(2_c) && \lim_{x\to 0}\frac{f(x)-1}{x}=c. \end{aligned}$$

Lemma 6.1 (Folgerungen aus (1) **und** (2_c)**).** *Sei $f:\mathbb{R}\to\mathbb{R}$ eine Funktion, die* (1) *und* (2_c) *erfüllt. Dann gilt $f(0)=1$,*

$$f(x)=\Big(f\Big(\frac{x}{n}\Big)\Big)^n \qquad \textit{und} \qquad x_n:=n\Big(f\Big(\frac{x}{n}\Big)-1\Big)\xrightarrow{n\to\infty} xc.$$

Außerdem ist

$$f(x)=\lim_{n\to\infty}\Big(1+\frac{x_n}{n}\Big)^n.$$

Beweis. Es muss $f(0)=1$ gelten, da andernfalls der Grenzwert in (2_c) nicht existieren würde. Weiter gilt mit (1)

$$f(x)=f\Big(\frac{x}{n}+\cdots+\frac{x}{n}\Big)=\Big(f\Big(\frac{x}{n}\Big)\Big)^n.$$

Außerdem gilt mit $h_n:=x/n$ und wegen (2_c) mit Hilfe von Lemma 5.37

$$n\Big(f\Big(\frac{x}{n}\Big)-1\Big)=x\frac{f(h_n)-1}{h_n}\xrightarrow{n\to\infty} xc.$$

Die letzte Behauptung ist klar wegen $1+x_n=f(x/n)$. □

Lemma 6.2 (Fundamentallemma). *Sei $x\in\mathbb{R}$ und $(x_n)_{n=1,2,\dots}$ eine reellwertige Folge mit $x_n\xrightarrow{n\to\infty} x$. Dann gilt*

$$\lim_{n\to\infty}\Big(1+\frac{x_n}{n}\Big)^n=\sum_{k=0}^\infty\frac{x^k}{k!}.$$

Bemerkung 6.3. Die Existenz des Grenzwertes $\lim_{n\to\infty}\Big(1+\frac{x_n}{n}\Big)^n$ auf der linken Seite ist nicht a priori klar. Jedoch folgt mit dem Quotientenkriterium sofort die Existenz von $\sum_{k=0}^\infty\frac{x^k}{k!}$. Wir werden im Beweis zeigen, dass $\Big(1+\frac{x_n}{n}\Big)^n$ beliebig nahe an diesem Grenzwert liegt, woraus auch die Existenz des linken Grenzwertes folgt.

Beweis von Lemma 6.2. Wir beginnen mit einer Nebenrechnung. Es ist für $0\le k\le n$

$$\binom{n}{k}\frac{1}{n^k}=\frac{n\cdots(n-k+1)}{k!n^k}=\frac{1}{k!}\Big(1-\frac{1}{n}\Big)\cdots\Big(1-\frac{k-1}{n}\Big)\le\frac{1}{k!}.$$

Außerdem sieht man leicht

$$\binom{n}{k}\frac{1}{n^k}\xrightarrow{n\to\infty}\frac{1}{k!}.$$

Damit gilt für beliebiges $K\in\mathbb{N}$, dass

$$\sum_{k=0}^{K-1}\binom{n}{k}\frac{x_n^k}{n^k}-\frac{x^k}{k!}\xrightarrow{n\to\infty}0.$$

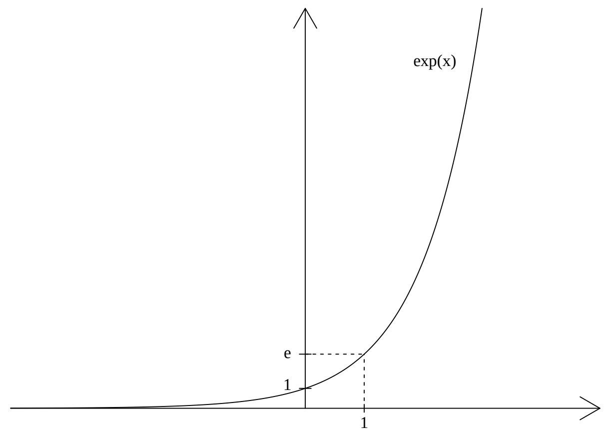

Abbildung 6.1: Die Funktion $x \mapsto \exp(x)$.

Zu $\varepsilon > 0$ wählen wir K so groß, dass

$$\sum_{k=K}^{\infty} \frac{(|x|+1)^k}{k!} < \varepsilon/3, \qquad \text{und} \qquad |x_n| \leq |x| + 1$$

für $n \geq K$. Nun wählen wir $n \geq K$ groß genug, so dass

$$\Big| \sum_{k=0}^{K-1} \binom{n}{k} \frac{x_n^k}{n^k} - \frac{x^k}{k!} \Big| < \varepsilon/3.$$

Dann gilt mit Proposition 1.40

$$\Big| \Big(1 + \frac{x_n}{n}\Big)^n - \sum_{k=0}^{\infty} \frac{x^k}{k!} \Big| \leq \Big| \sum_{k=0}^{K-1} \binom{n}{k} \frac{x_n^k}{n^k} - \frac{x^k}{k!} \Big| + \sum_{k=K}^{n} \binom{n}{k} \frac{|x_n|^k}{n^k} + \sum_{k=K}^{\infty} \frac{|x|^k}{k!} < \varepsilon.$$

Da $\varepsilon > 0$ beliebig war, folgt die Behauptung. □

Definition 6.4 (Exponentialfunktion). *Die Abbildung*

$$\exp : x \mapsto \lim_{n\to\infty} \Big(1 + \frac{x}{n}\Big)^n = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

heißt Exponentialfunktion. *Weiter heißt* $e := \exp(1)$ Euler'sche Zahl.

Bemerkung 6.5. Nach obiger Definition ist also

$$e = \lim_{n\to\infty} \Big(1 + \frac{1}{n}\Big)^n = \sum_{k=0}^{\infty} \frac{1}{k!}.$$

Der numerische Werte er Euler'schen Zahl ist etwa $e \approx 2.71828...$ Die Funktion $x \mapsto \exp(x)$ ist in Abbildung 6.1 dargestellt.

Lemma 6.6 (Grundlegende Eigenschaften). *1. Die Funktion $x \mapsto \exp(cx)$ ist die einzige Funktion, die (1) und (2_c) erfüllt. Sie ist stetig, positiv und monoton. Sie ist streng monoton für $c \neq 0$ (und zwar wachsend für $c > 0$ und fallend für $c < 0$) und es gilt $\exp(-x) = 1/\exp(x)$.*

2. Für $x \in \mathbb{Q}$ gilt $\exp(x) = e^x$. Weiter ist $\exp$ die einzige stetige Fortsetzung von $x \mapsto e^x$ von $\mathbb{Q}$ nach $\mathbb{R}$.

3. Für jedes $y \in \mathbb{R}_+$ gibt es $x \in \mathbb{R}$ mit $y = \exp(x)$.

4. Es gilt $\lim_{x\to\infty} \exp(x) = \infty$, $\lim_{x\to-\infty} \exp(x) = 0$.

5. Es gilt $\lim_{x\to\infty} \frac{\exp(x)}{x^n} = \infty$ sowie $\lim_{x\to\infty} x^n \exp(-x) = 0$.

Beweis. 1. Für (1) schreiben wir

$$\begin{aligned}\exp(x+y) &= \sum_{k=0}^{\infty} \frac{(x+y)^k}{k!} = \sum_{k=0}^{\infty}\sum_{l=0}^{k} \binom{k}{l} x^k y^{l-k} \frac{1}{k!} \\ &= \sum_{l=0}^{\infty}\sum_{k=l}^{\infty} \frac{l!(k-l)!}{x}^k y^{l-k} = \sum_{l=0}^{\infty}\sum_{k=0}^{\infty} \frac{l!k!}{x}^k y^k = \exp(x)\exp(y).\end{aligned}$$

Daraus folgt (1) für alle c. Für (2_c) gilt, falls $|x| < 1$

$$\begin{aligned}\Big|\frac{1}{x}\Big(\sum_{k=0}^{\infty} \frac{(cx)^k}{k!} - 1\Big) - c\Big| &= |c| \cdot \Big|\sum_{k=1}^{\infty} \frac{(cx)^{k-1}}{k!} - 1\Big| \\ &\leq |cx| \cdot \sum_{k=2}^{\infty} \frac{1}{k!}\end{aligned}$$

woraus (2_c) folgt. Angenommen, es gibt eine weitere Funktion f, die (1) und (2_c) erfüllt. Dann muss nach Lemma 6.1 und Lemma 6.2 $(1+x/n)^x \xrightarrow{n\to\infty} \exp(x)$ gelten. Für die Stetigkeit folgt aus (2_c) zunächst $\exp(cx) \xrightarrow{x\to 0} 1$. Damit gilt $\exp(c(x+h)) = \exp(cx)\exp(ch) \xrightarrow{h\to 0} \exp(cx)$. Die Positiviät folgt aus $\exp(x) = (\exp(x/2))^2$, die Monotonie aus $\exp(x+h)/\exp(x) = \exp(h) > 1 + h > 1$. Weiter ist $\exp(0) = 1$ und damit $1 = \exp(x-x) = \exp(x)\exp(-x)$.
2. Für $x \in \mathbb{N}$ ist $\exp(x) = e^x$ klar nach Lemma 6.1. Damit folgt für $x = \frac{m}{n} \in \mathbb{Q}$ und $e^m = \exp(m) = (\exp(m/n))^n$, also $e^x = \exp(x)$. Da $\mathbb{R}$ die Menge der Häufungspunkte von $\mathbb{Q}$ ist, folgt die Behauptung nun aus Proposition 5.35.
3. Sei zunächst $y \geq 1$. Klar ist, dass $\exp(y) \geq 1 + y > y$. Da $\exp(0) = 1$, folgt mit dem Zwischenwertsatz, Theorem 5.19, dass es ein $x \in [0, y]$ gibt mit $\exp(x) = y$. Für $y \leq 1$ folgt die Behauptung daraus, dass es ein x gibt mit $\exp(x) = 1/y$, also $\exp(-x) = y$ nach 1.
4. ist klar wegen 3. und der Monotonie.
5. Es gilt $\exp(x) > x^{n+1}/(n+1)!$, also $\frac{\exp(x)}{n!} > \frac{x}{(n+1)!}$. Daraus folgt bereits die erste Behauptung. Für die zweite Behauptung schreiben wir $\frac{1}{x^n \exp(-x)} = \frac{\exp(x)}{x^n} \xrightarrow{x\to\infty} \infty$. □

Bemerkung 6.7 (Exponentialschreibweise). Wegen Lemma 6.6.2 gilt $\exp(x) = e^x$ für $x \in \mathbb{Q}$. (Für andere Exponenten war bisher der Ausdruck e^x nämlich gar nicht definiert.) Wir werden im Folgenden die Schreibweise $e^x := \exp(x)$ auch für $x \in \mathbb{R} \setminus \mathbb{Q}$ verwenden. Schließlich ist die Exponentialfunktion die einzige stetige Fortsetzung auf $\mathbb{R}$, die diese Beziehung erfüllt.

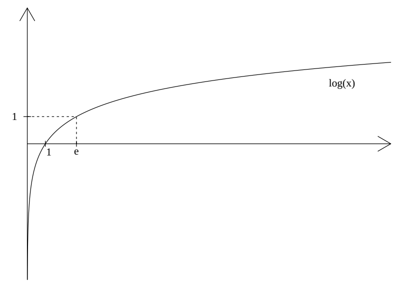

Abbildung 6.2: Die Funktion $x \mapsto \log(x)$.

Lemma 6.8. *Die Zahl* e *ist irrational.*

Beweis. Zunächst zeigen wir für $|x| \leq 1$

$$R_n(x) := \left| \exp(x) - \sum_{k=0}^{n} \frac{x^k}{k!} \right| \leq 2\frac{|x|^{n+1}}{(n+1)!}. \tag{6.2}$$

Hierzu schreiben wir

$$\begin{aligned} R_n(x) &\leq \sum_{k=n+1}^{\infty} \frac{|x|^k}{k!} = \frac{|x|^{n+1}}{(n+1)!}\Big(1 + \frac{|x|}{n+2} + \frac{|x|^2}{(n+2)(n+3)} + \cdots\Big) \\ &\leq \frac{|x|^{n+1}}{(n+1)!}\Big(1 + \frac{1}{2} + \frac{1}{2^2} + \cdots\Big) = 2\frac{|x|^{n+1}}{(n+1)!}. \end{aligned}$$

Nun kommen wir zum eigentlichen Beweis. Angenommen, e wäre rational, d.h. es gibt $m, n \in \mathbb{N}$ mit $e = m/n$. Dann ist $n!e \in \mathbb{N}$ also auch

$$n!R_n(1) = n!\Big(e - \sum_{k=0}^{n} \frac{1}{k!}\Big) \in \mathbb{N}.$$

Allerdings folgt aus (6.2) für $n > 1$

$$0 < n!R_n(1) \leq \frac{2}{n+1} < 1$$

im Widerspruch dazu. □

6.2 Der natürliche Logarithmus

Nach Lemma 6.6.3 ist die Funktion $\exp : \mathbb{R} \to \mathbb{R}_+$ umkehrbar. Diese Umkehrfunktion heißt natürlicher Logarithmus. Siehe Abbildung 6.2 für eine Illustration.

Definition 6.9 (Natürlicher Logarithmus). *Die Abbildung* $\log : \mathbb{R}_+ \to \mathbb{R}$, *für die* $\log(\exp(x)) = \exp(\log(x)) = x$ *gilt, heißt* natürlicher Logarithmus.

Lemma 6.10 (Grundlegende Eigenschaften).

1. *Die Funktion* $\log : \mathbb{R}_+ \to \mathbb{R}$ *ist stetig und streng monoton wachsend.*
2. *Für* $x, y \in \mathbb{R}_+$ *gilt* $\log(xy) = \log x + \log y$, $\log(x^{-1}) = -\log x$ *und* $\log(x^y) = y \log x$.

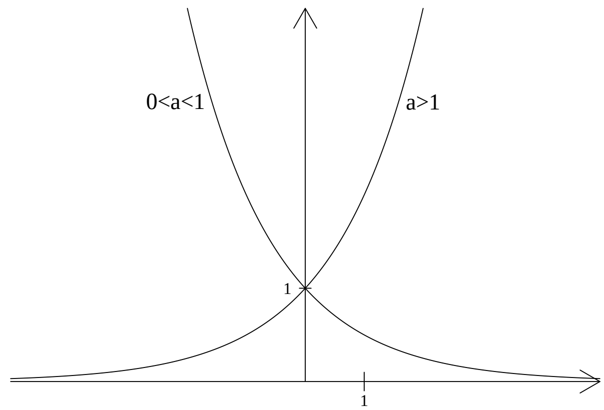

Abbildung 6.3: Die Exponentialfunktionen $x \mapsto a^x$ wachsen für $a > 1$ und fallen für $0 < a < 1$.

3. *Es gilt*
$$\lim_{x \to 0} \frac{\log(1+x)}{x} = 1.$$

4. *Es gilt* $\lim_{x\to\infty} \log(x) = \infty$, $\lim_{x\to 0} \log(x) = -\infty$.

5. *Für* $n \in \mathbb{N}$ *gilt* $\lim_{x\to\infty} \frac{\log(x)}{x^{1/n}} = 0$.

Beweis. 1. und 4. folgen direkt aus den entsprechenden Eigenschaften der Exponentialfunktion.
2. Die erste Behauptung folgt aus $\exp(\log(x) + \log(y)) = \exp(\log(x)) \cdot \exp(\log(y)) = xy = \exp(\log(xy))$, die zweite aus $\exp(-\log(x)) = 1/(\exp(\log(x))) = 1/x = \exp(\log(x^{-1}))$. Die letzte Behauptung zeigen wir im Fall $y \in \mathbb{Q}$ mit $y = m/n$ mit $m, n \in \mathbb{N}$. Der allgemeine Fall folgt dann durch die Stetigkeit. Es gilt

$$n \log(x^{m/n}) = \log((x^{m/n})^n) = \log(x^m) = m \log(x).$$

3. Sei $(x_n)_{n=1,2,\ldots}$ eine Nullfolge und $y_n := \log(1+x_n)$. Da $\exp(\log(1)) = 1$ und $\exp(0) = 1$, ist $\log(1) = 0$ und damit ist $(y_n)_{n=1,2,\ldots}$ wegen der Stetigkeit von log eine Nullfolge. Außerdem gilt
$$\frac{\log(1+x_n)}{x_n} = \frac{y_n}{e^{y_n} - 1} \xrightarrow{n\to\infty} 1$$
nach (2_c) für $x = 1$.
5. Setzt man $y = x^{1/n}$, ist die Aussage äquivalent zu $\frac{\log(y)}{y} \xrightarrow{y\to\infty} 0$. Dies folgt mit $z := \log(y)$ aus $\frac{z}{\exp(z)} \xrightarrow{z\to\infty} 0$ und Lemma 6.6.5. □

6.3 Allgemeine Exponentialfunktionen

Bisher haben wir nur die Exponentialfunktion $x \mapsto e^x$ (für $x \in \mathbb{R}$) kennen gelernt. Dies erweitern wir nun auf beliebige positive Basen ungleich e. Siehe etwa Abbildung 6.3.

Definition 6.11. *Sei $a > 0$. Dann definieren wir die* Exponentialfunktion $x \mapsto a^x := \exp(x \log a)$ zur Basis a.

Lemma 6.12 (Grundlegende Eigenschaften). *Seien $a, b > 0$.*

1. *Die Funktion $x \mapsto a^x$ ist stetig. Sie erfüllt (1) und $(2_{\log(a)})$ aus Abschnitt 6.1. Im Falle $a > 1$ ist sie streng monoton wachsend, im Falle $a < 1$ streng monoton fallend. Es gilt $(a^x)^y = a^{xy}$ und $a^x b^x = (ab)^x$.*
2. *Für $x, y \in \mathbb{R}$ gilt $a^{x+y} = a^x a^y$.*
3. *Es gilt*
$$\lim_{x \to 0} \frac{a^x - 1}{x} = \log a.$$

Beweis. Alle Eigenschaften folgen direkt aus Lemma 6.6. □

6.4 Allgemeine Potenzfunktionen

Wir kommen nun zu Funktionen $x \mapsto x^a$. Bisher wurden diese nur für $a \in \mathbb{Q}$ definiert, was wir hier auf $a \in \mathbb{R}$ erweitern wollen. Siehe auch Abbildung 6.4.

Definition 6.13. *Sei $a \in \mathbb{R}$. Dann definieren wir die Potenzfunktion $x \mapsto x^a := \exp(a \log x)$ für $x > 0$.*

Lemma 6.14 (Grundlegende Eigenschaften). *Sei $a \in \mathbb{R}$.*

1. *Die Funktion $x \mapsto x^a$ ist stetig. Im Falle $a > 0$ ist sie streng monoton wachsend, im Falle $a < 0$ streng monoton fallend.*
2. *Es gelten folgende Grenzwerte für $a > 0$:*
$$\lim_{x \to \infty} x^a = \infty, \qquad \lim_{x \to 0} x^a = 0, \qquad \lim_{x \to \infty} \frac{\log(x)}{x^a} = 0, \qquad \lim_{x \to 0} x^a \log(x) = 0.$$
Für $a < 0$ gilt:
$$\lim_{x \to \infty} x^a = 0, \qquad \lim_{x \to 0} x^a = \infty,$$

Beweis. Alle Behauptungen folgen direkt aus Eigenschaften der Exponentialfunktion. □

6.5 Trigonometrische Funktionen

Aus der linearen Algebra sind die komplexen Zahlen $\mathbb{C}$ und deren Rechenregeln[13] bekannt. Wir erweitern die Definition der Exponentialfunktion
$$e^x := \sum_{k=0}^{\infty} \frac{x^k}{k!} \tag{6.3}$$
auf $x \in \mathbb{C}$. Alle Rechenregeln der Exponentialfunktion übertragen sich auf komplexe Zahlen. Diese Erweiterung ist hilfreich bei der Definition von Cosinus und Sinus in folgendem Sinne.

[13] Insbesondere ist i die imaginäre Einheit, wir schreiben $z = x + iy$ (mit $x, y \in \mathbb{R}$) und $\bar{z}$ ist die komplex Konjugierte von $z \in \mathbb{C}$.

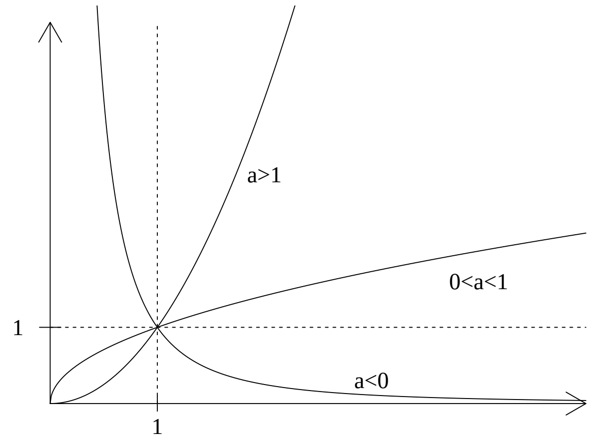

Abbildung 6.4: Die Funktionen $x \mapsto x^a$ wachsen für $a > 0$ und fallen für $a < 0$.

Definition 6.15 (Cosinus und Sinus). *Für $x \in \mathbb{R}$ definieren wir*

$$\cos(x) := \tfrac{1}{2}(e^{ix} + e^{-ix}), \qquad \sin(x) := \tfrac{1}{2i}(e^{ix} - e^{-ix}).$$

Bemerkung 6.16. Aus (6.3) folgt $\overline{e^x} = e^{\overline{x}}$. Daraus berechnen wir für $x \in \mathbb{R}$ sofort $|e^{ix}|^2 = e^{ix}e^{-ix} = e^0 = 1$. Schreibt man $e^{ix} = a + ib$ für $a, b \in \mathbb{R}$, so folgt aus der Definition sofort $a = \cos(x)$ und $b = \sin(x)$, also

$$e^{ix} = \cos(x) + i\sin(x)$$

sowie

$$|e^{ix}|^2 = \cos(x)^2 + \sin(x)^2 = 1.$$

Weiter ist aus der Definition ersichtlich, dass $\cos(-x) = \cos(x)$ und $\sin(-x) = -\sin(x)$ gilt.

Lemma 6.17 (Additionstheoreme). *Es gilt*

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

Beweis. Wir schreiben

$$\begin{aligned}4(\cos(x)\cos(y) - \sin(x)\sin(y)) &= (e^{ix} + e^{-ix})(e^{iy} + e^{-iy}) + (e^{ix} - e^{-ix})(e^{iy} - e^{-iy})\\ &= 2(e^{ix}e^{iy} + e^{-ix}e^{-iy}) = 4\cos(x+y),\\ 4(\sin(x)\cos(y) + \cos(x)\sin(y)) &= -i((e^{ix} - e^{-ix})(e^{iy} + e^{-iy}) + (e^{ix} + e^{-ix})(e^{iy} - e^{-iy}))\\ &= -2i(e^{ix}e^{iy} - e^{-ix}e^{-iy}) = 4\sin(x+y).\end{aligned}$$

□

Proposition 6.18 (Reihendarstellung von Cosinus und Sinus). *Es gilt*

$$\begin{aligned}\cos(x) &:= \sum_{k=0}^{\infty}(-1)^k \frac{x^{2k}}{(2k)!},\\ \sin(x) &:= \sum_{k=0}^{\infty}(-1)^k \frac{x^{2k+1}}{(2k+1)!}.\end{aligned}$$

Beweis. Wir setzen einfach die Definition ein, und verwenden den Umordnungssatz für Reihen

$$\begin{aligned} 2\cos(x) &= \sum_{k=0}^{\infty} \frac{(ix)^k + (-ix)^k}{k!} = \sum_{k=0}^{\infty} \frac{(1+(-1)^k)(ix)^k}{k!} = 2 \sum_{\substack{k=0 \\ k \text{ gerade}}}^{\infty} \frac{(ix)^k}{k!} \\ &= 2\sum_{k=0}^{\infty} \frac{(ix)^{2k}}{(2k)!} = 2\sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!}, \\ 2i\sin(x) &= \sum_{k=0}^{\infty} \frac{(ix)^k - (-ix)^k}{k!} = \sum_{k=0}^{\infty} \frac{(1-(-1)^k)(ix)^k}{k!} = 2 \sum_{\substack{k=0 \\ k \text{ ungerade}}}^{\infty} \frac{(ix)^k}{k!} \\ &= 2\sum_{k=0}^{\infty} \frac{(ix)^{2k+1}}{(2k+1)!} = 2i\sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}. \end{aligned}$$

□

Lemma 6.19 (Abschätzungen für cos **und** sin**).** *Für $x \in (0,2]$ gilt*

$$\begin{aligned} 1 - \frac{x^2}{2} &< \cos(x) < 1 - \frac{x^2}{2} + \frac{x^4}{24}, \\ x - \frac{x^3}{6} &< \sin(x) < x. \end{aligned}$$

Insbesondere ist

$$\lim_{x\to 0} \frac{\sin(x)}{x} = 1.$$

Außerdem ist $x \mapsto \cos(x)$ im Bereich $(0,2]$ streng monoton fallend.

Beweis. Wir bemerken zunächst, dass es sich bei den Reihendarstellungen aus Proposition 6.18 für $x \in (0,2]$ um alternierende Reihen handelt, auf die man Theorem 4.10 anwenden kann. Aus der dort verfügbaren Fehlerabschätzung erhält man sofort die Behauptung.

Um die Monotonie von $x \mapsto \cos(x)$ zu zeigen, folgt zunächst, dass $x \mapsto \sin(x)$ auf $(0,2]$ positiv ist. Sei nun $0 < x < y \le 2$. Wir wenden Lemma 6.17 auf die Paare $((x+y)/2, (y-x)/2)$ und $((x+y)/2, (x-y)/2)$ an und subtrahieren die Ergebnisse, um

$$\begin{aligned} \cos(y) - \cos(x) &= \cos((x+y)/2)\cos((y-x)/2) - \sin((x+y)/2)\sin((y-x)/2) \\ &\quad - \cos((x+y)/2)\cos((x-y)/2) + \sin((x+y)/2)\sin((x-y)/2) \\ &= -2\sin((x+y)/2)\sin((y-x)/2) < 0 \end{aligned}$$

zu erhalten. Daraus folgt die Monotonie. □

Korollar 6.20 (Nullstelle von cos**).** *Die Funktion $x \mapsto \cos(x)$ hat im Bereich $x \in (0,2]$ genau eine Nullstelle.*

Beweis. Klar ist, dass $x \mapsto \cos(x)$ stetig ist. Da $x \mapsto \cos(x)$ auf $(0,2]$ streng monoton fällt, gibt es höchstens eine Nullstelle. Die Existenz einer Nullstelle folgt aus dem Nullstellensatz, Korollar 5.20, und den Abschätzungen aus Lemma 6.19. □

Definition 6.21. *Wir bezeichnen die nach Korollar 6.20 eindeutige Nullstelle von $x \mapsto \cos(x)$ im Bereich $(0,2]$ mit $\pi/2$.*

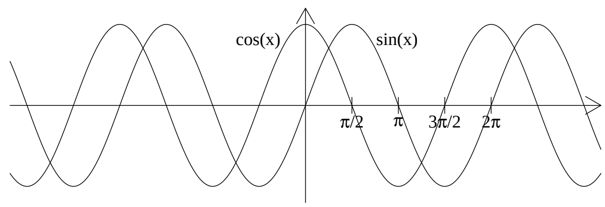

Abbildung 6.5: Die Funktionen $x \mapsto \cos(x)$ und $x \mapsto \sin(x)$ haben eine Periode von 2π.

Bemerkung 6.22. Wegen der Positivität von $x \mapsto \sin(x)$ auf $(0,2]$ und $\cos(x)^2 + \sin(x)^2 = 1$ folgt sofort $\sin(\pi/2) = 1$. Außerdem erkennen wir nun die Formel $e^{i\pi/2} = \cos(\pi/2) + i\sin(\pi/2) = i$. Durch Potenzieren erhalten wir hierdurch $e^{i\pi} = i^2 = -1$ und $e^{2i\pi} = \cos(2\pi) + i\sin(2\pi) = 1$, also $\cos(2\pi) = 1, \sin(2\pi) = 0$. Hier eine kleine Wertetabelle:

x	0	$\frac{\pi}{2}$	π	$\frac{3\pi}{2}$	2π
$\cos(x)$	1	0	-1	0	1
$\sin(x)$	0	1	0	-1	0

Eine Illustration der Funktionen cos und sin findet man in Abbildung 6.5.

Korollar 6.23. *Es gelten die Beziehungen*

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

Die jeweils letzte Beziehung besagt, dass cos *und* sin *eine* Periode *von* 2π *haben.*

Beweis. Aus Lemma 6.17 und den Werten aus der letzten Bemerkung folgert man sofort alle Behauptungen. □

Bemerkung 6.24 (Arcus-Cosinus und Arcus-Sinus). Die Funktionen

$$\cos : \begin{cases} [0,\pi] & \to [-1,1] \\ x & \mapsto \cos(x) \end{cases} \qquad \sin : \begin{cases} [-\pi/2,\pi/2] & \to [-1,1] \\ x & \mapsto \sin(x) \end{cases}$$

sind streng monoton und stetig. Mit dem Zwischenwertsatz folgert man, dass diese Funktionen bijektiv sind. Deswegen können wir ihre Umkehrfunktionen

$$\arccos : \begin{cases} [-1,1] & \to [0,\pi] \\ x & \mapsto \arccos(x) \end{cases} \qquad \arcsin : \begin{cases} [-1,1] & \to [-\pi/2,\pi/2] \\ x & \mapsto \arcsin(x) \end{cases}$$

definieren. Eine Illustration dieser Funktionen findet sich in Abbildung 6.6.

Aus cos und sin lässt sich der Tangens ableiten; siehe auch Abbildung 6.7.

Definition 6.25 (Tangens und Arcus-Tangens). *Für* $x \notin \pi/2 + \{k\pi : k \in \mathbb{Z}\}$ *definieren wir*

$$\tan(x) := \frac{\sin(x)}{\cos(x)}.$$

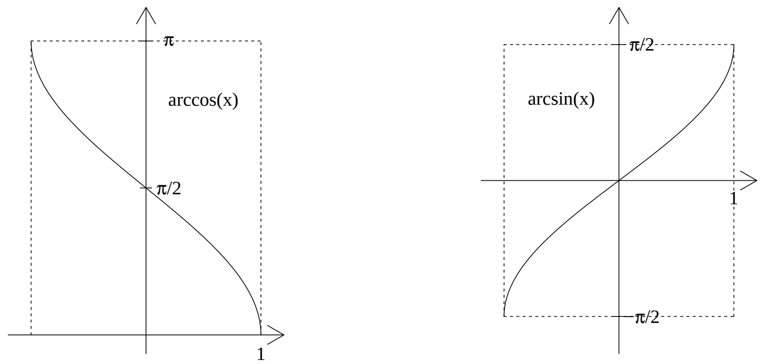

Abbildung 6.6: Die Funktionen $x \mapsto \arccos(x)$ und $x \mapsto \arcsin(x)$ sind die Umkehrfunktionen von cos und sin um Bereich um den Nullpunkt.

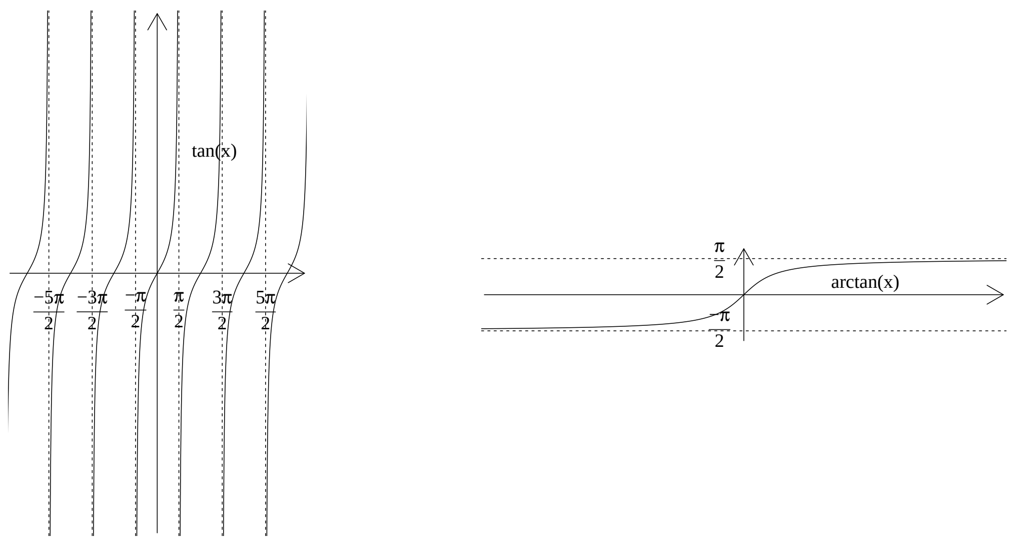

Abbildung 6.7: Die Funktionen $x \mapsto \tan(x)$ und $x \mapsto \arctan(x)$ werden aus cos und sin abgeleitet.

Die Funktion

$$\tan : \begin{cases} (-\pi/2, \pi/2) & \to \mathbb{R} \\ x & \mapsto \tan(x) \end{cases}$$

ist bijektiv. Ihre Umkehrfunktion

$$\arctan : \begin{cases} \mathbb{R} & \to (-\pi/2, \pi/2) \\ x & \mapsto \arctan(x) \end{cases}$$

heißt Arcus-Tangens.

7 Differenzierbarkeit

Nachdem wir nun Stetigkeit von Funktionen kennen gelernt haben, wenden wir uns nun einer stärkeren Eigenschaft, der Differenzierbarkeit, zu. Nachdem wir in Abschnitt 7.1 grundlegen-

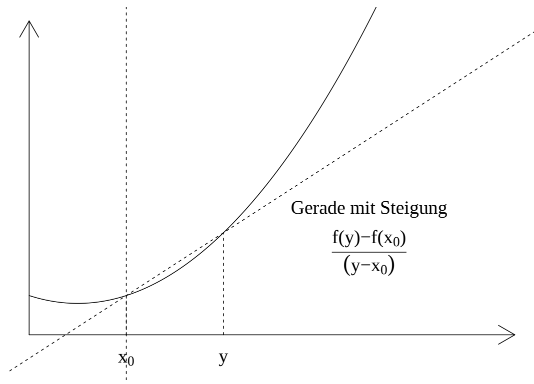

Abbildung 7.1: Die Ableitung einer Funktion f kann man mittels des Differenzenquotienten $\frac{f(y)-f(x_0)}{y-x_0}$ definieren; siehe Definition 7.1. Dieser entspricht der Steigung einer Gerade durch die Punkte $(x_0, f(x_0))$ und $(y, f(y))$.

de Eigenschaften behandeln, beschäftigen wir uns mit Ableitungen der Umkehrfunktion in Abschnitt 7.2. In Anwendungen besonders wichtig sind der Mittelwertsatz – siehe Abschnitt 16.4 – und die Regeln von l'Hospital – siehe Abschnitt 7.4. Zum Abschluss behandeln wir ein paar Sonderfälle in Abschnitt 7.5.

7.1 Definition und grundlegende Eigenschaften

Die Ableitung einer Funktion an einem Punkt x_0 gibt die Steigung der Funktion im Punkt x_0 an. Eine Illustration der Definition der Ableitung findet sich in Abbildung 7.1.

Definition 7.1. *Seien $I \subseteq \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}$. Dann heißt f in $x_0 \in I$ differenzierbar, wenn*

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

existiert. Dieser Grenzwert heißt (erste) Ableitung *von f in x_0 und wird auch mit $\frac{df}{dx}(x_0)$ oder $f'(x_0)$ oder $Df(x_0)$ bezeichnet. Die Funktion f heißt* differenzierbar *(im Intervall I), wenn f für alle $x_0 \in I$ differenzierbar ist. Iterativ definieren wir $f^{(1)}(x_0) := D^1 f(x_0) := f'(x_0)$ und die n-te* Ableitung von *f in x_0, $D^n f(x_0)$ oder $f^{(n)}(x_0)$ mittels $D^n f(x_0) := f^{(n)}(x_0) := \frac{d}{dx} f^{(n-1)}(x_0)$. Wir setzen*

$$\begin{aligned} \mathcal{C}^0(I) &:= \mathcal{C}(I) := \mathcal{C}_{\mathbb{R}}(I) := \{f : I \to \mathbb{R} \text{ stetig}\}, \\ \mathcal{C}^n(I) &:= \mathcal{C}^n_{\mathbb{R}}(I) := \{f : I \to \mathbb{R}\ n\text{-mal differenzierbar},\ f^{(n)} \in \mathcal{C}^0(I)\} \end{aligned}$$

und

$$\mathcal{C}^\infty(I) := \mathcal{C}^\infty_{\mathbb{R}}(I) := \{f : I \to \mathbb{R} \text{ beliebig oft differenzierbar}\}.$$

Proposition 7.2 (Äquivalente Formulierungen). *Seien $I \subseteq \mathbb{R}$ ein Intervall, $x_0 \in I$ und $f : I \to \mathbb{R}$. Dann sind äquivalent:*

1. *f ist in x_0 differenzierbar und $f'(x_0) = y$,*

2. $\lim_{h \to 0} \frac{f(x_0+h) - f(x_0)}{h} = y$,

3. *Es gibt eine in x_0 stetige Funktion φ mit $f(x) - f(x_0) = (x - x_0)\varphi(x)$ und $\varphi(x_0) = y$.*

4. *Es gibt eine lineare Funktion*[14] $L : \mathbb{R} \to \mathbb{R}$, *so dass*

$$\lim_{h \to 0} \frac{f(x_0+h) - f(x_0) - L(h)}{h} = 0$$

und es ist $L(h) = yh$.

Beweis. 1. $\iff$ 2.: Das ist klar, wenn man $x = x_0 + h$ schreibt.
1. $\Longrightarrow$ 3.: Wir definieren

$$\varphi(x) := \frac{f(x) - f(x_0)}{x - x_0}$$

für $x \neq x_0$. Nach 1. wissen wir, dass φ stetig in x_0 fortsetzbar ist.
3. $\Longrightarrow$ 4.: Wir definieren $L(h) := yh$. Damit gilt

$$\frac{f(x_0+h) - f(x_0) - L(h)}{h} = \frac{h\varphi(x_0+h) - yh}{h} = \varphi(x_0+h) - \varphi(x_0) \xrightarrow{h \to 0} 0.$$

4. $\Longrightarrow$ 2.: Da $L(h) = yh$, gilt

$$\lim_{h \to 0} \frac{f(x_0+h) - f(x_0)}{h} - y = \lim_{h \to 0} \frac{f(x_0+h) - f(x_0) - L(h)}{h} = 0,$$

woraus die Behauptung folgt. □

Korollar 7.3 (Differenzierbarkeit und Stetigkeit). *Seien $I \subseteq \mathbb{R}$ ein Intervall, $x_0 \in I$ und $f : I \to \mathbb{R}$. Ist f in x_0 differenzierbar, so ist f in x_0 stetig.*

Beweis. Wir verwenden Proposition 7.2.3. Hieraus folgt nämlich

$$f(x) - f(x_0) = (x - x_0)\varphi(x) \xrightarrow{x \to x_0} 0$$

und damit die Stetigkeit von f in x_0. □

Proposition 7.4 (Einfache Ableitungen). *Es gilt*

1. $\frac{d}{dx} x^n = nx^{n-1}, n = 1, 2, ...$

2. $\frac{d}{dx} e^{cx} = ce^{cx}, c \in \mathbb{R}$,

3. $\frac{d}{dx} \log(x) = \frac{1}{x}$.

[14] Wie in der linearen Algebra gilt also $L(ax + by) = aL(x) + bL(y)$.

Beweis. 1. Es gilt

$$\frac{(x+h)^n - x^n}{h} = \sum_{k=0}^{n} \binom{n}{k} x^k h^{n-k-1} - x^n h^{-1} = \sum_{k=0}^{n-1} \binom{n}{k} x^k h^{n-k-1} \xrightarrow{h\to 0} n x^{n-1}.$$

2. Es gilt

$$\frac{e^{c(x+h)} - e^{cx}}{h} = e^{cx}\frac{e^{ch}-1}{h} \xrightarrow{h\to 0} ce^{cx}$$

nach Lemma 6.6.
3. Es gilt

$$\frac{\log(x+h) - \log(x)}{h} = \frac{1}{x}\frac{\log(1+h/x)}{h/x} \xrightarrow{h\to 0} \frac{1}{x}$$

nach Lemma 6.10. □

Die Rechenregeln 2.–4. aus folgender Proposition heißen *Produktregel*, *Quotientenregel* und *Kettenregel*.

Proposition 7.5 (Rechenregeln). *Seien $f, g : I \to \mathbb{R}$ in $x \in I$ differenzierbar. Dann gilt*

1. $(f+g)'(x) = f'(x) + g'(x)$.
2. $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$.
3. *Ist $g(x) \neq 0$, so ist* $\frac{d}{dx}\frac{f(x)}{g(x)} = \frac{f'(x)g(x)-f(x)g'(x)}{g^2(x)}$.
4. *Ist g differenzierbar in $f(x)$, so gilt* $\frac{d}{dx}g(f(x)) = g'(f(x))f'(x)$.

Beweis. 1. Dies folgt direkt aus

$$\begin{aligned}\frac{f(x+h)+g(x+h)-f(x)-g(x)}{h} &= \frac{f(x+h)-f(x)}{h} + \frac{g(x+h)-g(x)}{h} \\ &\xrightarrow{h\to 0} f'(x) + g'(x).\end{aligned}$$

2. Da g stetig ist, gilt

$$\begin{aligned}\frac{f(x+h)g(x+h)-f(x)g(x)}{h} &= \frac{(f(x+h)-f(x))g(x+h) + f(x)(g(x+h)-g(x))}{h} \\ &\xrightarrow{h\to 0} f'(x)g(x) + f(x)g'(x).\end{aligned}$$

3. Wieder gilt wegen der Stetigkeit von g, da g in einer Umgebung von x unglech 0 ist,

$$\begin{aligned}\frac{\frac{f(x+h)}{g(x+h)} - \frac{f(x)}{g(x)}}{h} &= \frac{1}{g(x+h)g(x)}\Big(\frac{f(x+h)-f(x)}{h}g(x) - \frac{g(x+h)-g(x)}{h}f(x)\Big) \\ &\xrightarrow{h\to 0} \frac{f'(x)g(x)-f(x)g'(x)}{g^2(x)}\end{aligned}$$

4. Nach Proposition 7.2.3 gibt es φ, ψ, die beide stetig in x sind, mit

$$f(x+h) - f(x) = h\varphi(x+h), \qquad g(f(x)+h) - g(f(x)) = h\psi(f(x)+h)$$

sowie $f'(x) = \varphi(x), g'(f(x)) = \psi(f(x))$. Deswegen gilt

$$g(f(x+h)) = g(f(x) + h\varphi(x+h)) = g(f(x)) + h\varphi(x+h)\psi(f(x) + h\varphi(x+h)),$$

also

$$\frac{g(f(x+h)) - g(f(x))}{h} = \varphi(x+h)\psi(f(x) + h\varphi(x+h)) \xrightarrow{h\to 0} g'(f(x))f'(x).$$

□

Beispiel 7.6. 1. Man berechnet leicht mit Hilfe der Produktregel für $n = 1, 2, ...$

$$\frac{d}{dx}x^n e^{-x} = e^{-x}x^{n-1}(n-x).$$

2. Für $a \in \mathbb{R}$ gilt $\frac{d}{dx}x^a = ax^{a-1}$.
   Denn: Wir berechnen

$$\frac{d}{dx}x^a = \frac{d}{dx}e^{a\log x} = e^{a\log x}\frac{d}{dx}a\log x = e^{a\log x}\frac{a}{x} = ax^{a-1}.$$

3. Ein Beispiel für die Quotientenregel sehen wir gleich in der nächsten Proposition bei der Ableitung von $x \mapsto \tan(x)$.

Proposition 7.7 (Ableitung der Winkelfunktionen). *Es gilt*

$$\frac{d}{dx}\cos(x) = -\sin(x), \qquad \frac{d}{dx}\sin(x) = \cos(x), \qquad \frac{d}{dx}\tan(x) = \frac{1}{\cos^2(x)}.$$

Beweis. Wir schreiben

$$\begin{aligned}
\frac{d}{dx}\cos(x) &= \tfrac{1}{2}\frac{d}{dx}(e^{ix} + e^{-ix}) = \tfrac{i}{2}(e^{ix} - e^{-ix}) = -\tfrac{1}{2i}(e^{ix} - e^{-ix}) = -\sin(x),\\
\frac{d}{dx}\sin(x) &= \tfrac{1}{2i}\frac{d}{dx}(e^{ix} - e^{-ix}) = \tfrac{1}{2}(e^{ix} + e^{-ix}) = \cos(x),\\
\frac{d}{dx}\tan(x) &= \frac{d}{dx}\frac{\sin(x)}{\cos(x)} = \frac{\cos^2(x) + \sin^2(x)}{\cos^2(x)} = \frac{1}{\cos^2(x)}.
\end{aligned}$$

□

7.2 Ableitung der Umkehrfunktion

Angenommen, wir wissen dass f' die Ableitung einer umkehrbaren Funktion f ist. Was wir in dieser Situation über die Ableitung f^{-1} aussagen können, formulieren wir nun.

Proposition 7.8 (Ableitung der Umkehrfunktion). *Sei $f : I \to \mathbb{R}$ eine streng monotone Funktion, die in $x_0 \in I$ differenzierbar ist, sowie g die Umkehrfunktion von f. Sei $y_0 := f(x_0)$, also $g(y_0) = x_0$. Dann ist g in y_0 differenzierbar und es gilt, falls $f'(x_0) \neq 0$*

$$g'(y_0) = \frac{1}{f'(x_0)} = \frac{1}{f'(g(y_0))}.$$

Beweis. Da f in x_0 differenzierbar ist, gibt es nach Proposition 7.2.3 eine Funktion φ mit $f(x) - f(x_0) = (x - x_0)\varphi(x)$ und $\varphi(x_0) = f'(x_0)$. Damit gilt

$$y - y_0 = (g(y) - g(y_0))\varphi(g(y)), \qquad g(y) - g(y_0) = (y - y_0)\frac{1}{\varphi(g(y))}.$$

Wieder wegen Proposition 7.2.3 hat g also die Ableitung

$$\frac{1}{\varphi(g(y_0))} = \frac{1}{\varphi(x_0)} = \frac{1}{f'(x_0)} = \frac{1}{f'(g(y_0))}.$$

□

Bemerkung 7.9 (Ableitung des Logarithmus). Wir haben bereits gezeigt, dass $\frac{d}{dx}\log(x) = 1/x$. Dies kann man auch mit Hilfe von Proposition 7.8 zeigen: Es ist log die Umkehrfunktion von exp. Damit ist

$$\frac{d}{dx}\log(x) = \frac{1}{\exp(\log(x))} = \frac{1}{x}.$$

Proposition 7.10 (Ableitung der Arkus-Funktionen). *Es gilt*

$$\begin{aligned}
\frac{d}{dx}\arccos(x) &= -\frac{1}{\sqrt{1-x^2}},\\
\frac{d}{dx}\arcsin(x) &= \frac{1}{\sqrt{1-x^2}},\\
\frac{d}{dx}\arctan(x) &= \frac{1}{1+x^2}.
\end{aligned}$$

Beweis. Wir schreiben

$$\begin{aligned}
\frac{d}{dx}\arccos(x) &= \frac{1}{-\sin(\arccos(x))} = \frac{1}{\sqrt{1-(\cos(\arccos(x))^2)}}\\
&= -\frac{1}{\sqrt{1-x^2}},\\
\frac{d}{dx}\arcsin(x) &= \frac{1}{\cos(\arcsin(x))} = \frac{1}{\sqrt{1-(\sin(\arcsin(x))^2)}}\\
&= \frac{1}{\sqrt{1-x^2}},\\
\frac{d}{dx}\arctan(x) &= \cos^2(\arctan(x)) = \frac{1}{1+\tan^2(\arctan(x))} = \frac{1}{1+x^2}.
\end{aligned}$$

□

7.3 Der Mittelwertsatz

In diesem Abschnitt beschäftigen wir uns mit Extremwerten, einem bereits aus der Schule wohlbekanntem Konzept.

Definition 7.11 (Lokales Maximum/Minimum). *Sei (E, r) ein metrischer Raum und $f : E \to \mathbb{R}$. Gibt es für $x \in E$ ein $\varepsilon > 0$ mit $f(y) \le f(x)$ $(f(y) \ge f(x))$ für $y \in B_\varepsilon(x)$, so heißt x ein* lokales Maximum *(*lokales Minimum*) von f. Gilt die Ungleichung sogar für alle $y \in E$, si heißt x* globales Maximum *(*globales Minimum*). Die Menge der Maxima oder Minima heißt auch die Menge der* Extrema.

Lemma 7.12 (Ableitungen und Extrema). *Sei $f : [a, b] \to \mathbb{R}$ eine stetige und auf (a, b) differenzierbare Funktion. Ist $x_0 \in (a, b)$ ein Maximum oder Minimum von f, so ist $f'(x_0) = 0$.*

Beweis. Wir beweisen das Lemma in dem Fall, wenn x ein Maximum ist. In diesem Fall gilt $(f(x) - f(x_0))/(x - x_0) \leq 0$ für $x \geq x_0$ und $(f(x) - f(x_0))/(x - x_0) \geq 0$ für $x \leq x_0$. Daraus folgt

$$\lim_{x \downarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} \leq 0, \qquad \lim_{x \uparrow x_0} \frac{f(x) - f(x_0)}{x - x_0} \geq 0,$$

also insgesamt $f'(x_0) = 0$. □

Proposition 7.13 (Der Satz von Rolle). *Sei $f : [a, b] \to \mathbb{R}$ eine stetige Funktion, die auf (a, b) differenzierbar ist. Gilt $f(a) = f(b)$, so gibt es ein $x \in (a, b)$ mit $f'(x) = 0$.*

Beweis. Falls f konstant ist, so ist nichts zu zeigen. Andernfalls können wir Korollar 5.26 verwenden, denn f ist eine stetige Funktion auf einem kompakten Intervall. Da f nicht konstant ist, ist entweder das Minimum oder das Maximum von f in (a, b). Damit folgt die Behauptung aus Lemma 7.12. □

Theorem 7.14 (Mittelwertsatz). *Sei $f : [a, b] \to \mathbb{R}$ eine stetige Funktion, die auf (a, b) differenzierbar ist. Dann gibt es ein $x \in (a, b)$ mit*

$$f'(x) = \frac{f(b) - f(a)}{b - a}.$$

Beweis. Wir wenden Proposition 7.13 auf die Funktion $h(x) := f(x) - \frac{f(b)-f(a)}{b-a}(x - a)$ an. Für diese gilt $h(a) = h(b) = f(a)$ und $h'(x) = f'(x) - \frac{f(b)-f(a)}{b-a}$. Damit folgt die Behauptung direkt aus Proposition 7.13. □

Korollar 7.15 (Monotoniekriterien). *Ist $f : (a, b) \to \mathbb{R}$ differenzierbar. Dann gilt*

$$\begin{aligned}
&f' > 0 \text{ in } (a,b) \Rightarrow f \text{ wächst in } (a,b) \text{ streng monoton},\\
&f' \geq 0 \text{ in } (a,b) \Leftrightarrow f \text{ wächst in } (a,b) \text{ monoton},\\
&f' < 0 \text{ in } (a,b) \Rightarrow f \text{ fällt in } (a,b) \text{ streng monoton},\\
&f' \leq 0 \text{ in } (a,b) \Leftrightarrow f \text{ fällt in } (a,b) \text{ monoton},
\end{aligned}$$

Beweis. Die Aussagen "$\Leftarrow$" folgen schon aus der Definition des Differenzenquotienten. Die Aussagen "$\Rightarrow$" folgen aus Theorem 7.14 und $f(x'') - f(x') = (x'' - x')f'(x)$ für ein geeignetes $x \in (x', x'')$. □

Beispiel 7.16. Die Funktion $f : x \mapsto \left(1 + \frac{1}{x}\right)^x$ ist streng monoton wachsend.
Denn: Wir zeigen, dass die Funktion $x \mapsto \log f(x)$ streng monoton wächst. Hierzu berechnen wir

$$\frac{d}{dx} x \log(1 + 1/x) = \log(1 + 1/x) + x\left(\frac{1}{x+1} - \frac{1}{x}\right) = \log(1 + 1/x) - \frac{1}{x+1}.$$

Um zu zeigen, dass diese Funktion positiv ist, betrachten wir[15] $g(x) := \frac{d}{dy} \log f(y)|_{y=1/x} = \log(1+x) - \frac{x}{1+x}$. Es gilt $g(0) = 0$ und $g'(x) = \frac{1}{1+x} - \frac{1}{(1+x)^2} > 0$. Daraus folgt $g(x) > 0$, also auch $\frac{d}{dx} \log f(x) > 0$ und damit die Behauptung.

Korollar 7.17 (Hinreichende Kriterien für Extrema). *Sei $f : (a, b) \to \mathbb{R}$ differenzierbar und $f'(x_0) = 0$ für $x_0 \in (a, b)$.*

1. *Ist $f'(x) \le 0$ für $x \in (a, x_0)$ und $f'(x) \ge 0$ für $x \in (x_0, b)$, dann hat f in x_0 ein lokales Minimum.*
2. *Ist $f'(x) \ge 0$ für $x \in (a, x_0)$ und $f'(x) \le 0$ für $x \in (x_0, b)$, dann hat f in x_0 ein lokales Maximum.*

Beweis. Klar wegen Korollar 7.15. □

Korollar 7.18 (Kriterium für konstante Funktionen). *Sei $f : (a, b) \to \mathbb{R}$ eine differenzierbare Funktion mit $f' = 0$. Dann ist f konstant.*

Beweis. Klar wegen Korollar 7.15. □

Wir kommen nun zu einer Verbindung von Differenzierbarkeit und Lipschitz-Stetigkeit; siehe Beispiel 5.6.3.

Proposition 7.19 (Schrankensatz). *Sei $f : (a, b) \to \mathbb{R}$ eine differenzierbare Funktion mit beschränkter Ableitung, d.h. $|f'(x)| \le L$ für ein $L \in \mathbb{R}_+$. Dann ist f Lipschitz-stetig mit Lipschitz-Konstante L, d.h.*

$$|f(x_2) - f(x_1)| \le L \cdot |x_2 - x_1|$$

für $x_1, x_2 \in (a, b)$.

Beweis. Sei $a < x_1 \le x_2 < b$. Nach dem Mittelwertsatz gibt es ein $x \in (x_1, x_2)$ mit

$$|f(x_2) - f(x_1)| = |f'(x)| \cdot |x_2 - x_1| \le L \cdot |x_2 - x_1|.$$

□

[15] Abundzu muss man bei Ableitungen auf die Notation aufpassen. Hier ein Beispiel: Betrachten wir die Funktion $f : x \mapsto x^2$. Was bedeutet dann der Ausdruck

$$\frac{d}{dx} f(1/x)?$$

Ist $g : x \mapsto 1/x$ und $h : f \circ g$, so bedeutet dies nämlich entweder

$$f'(g(x)) = \frac{2}{x} \qquad \text{oder} \qquad h'(x) = -\frac{2}{x^3}.$$

Um zu verdeutlichen was gemeint ist, schreiben wir manchmal

$$\frac{d}{dx} f(x)\Big|_{x=x_0} := f'(x_0)$$

und sagen zu $\cdots|_{x=x_0}$ auch *ausgewertet bei* $x = x_0$. Nun ist

$$f'(g(x)) = \frac{d}{dx} f(y)\Big|_{y=1/x} = \frac{1}{2x}, \qquad h'(x) = \frac{d}{dy} f \circ g(y)\Big|_{y=x} = -\frac{2}{x^3}.$$

7.4 Die Regeln von l'Hospital

Wir kommen nun zu Grenzwerte der Form $\lim_{x\to x_0} f(x)/g(x)$, falls $f(x), g(x) \xrightarrow{x\to x_0} 0$ oder $f(x), g(x) \xrightarrow{x\to x_0} \infty$.

Proposition 7.20 (Verallgemeinerter Mittelwertsatz). *Seien $f, g : [a,b] \to \mathbb{R}$ stetige Funktionen, die auf (a,b) differenzierbar sind, mit $g'(x) \neq 0$ für $x \in (a,b)$. Dann ist $g(b) \geq g(a)$ und es gibt es ein $x \in (a,b)$ mit*

$$\frac{f'(x)}{g'(x)} = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

Beweis. Nach dem Satz von Rolle gilt zunächst $g(b) \neq g(a)$. (Anderfalls gäbe es ein $x \in (a,b)$ mit $g'(x) = 0$.) Wir definieren nun

$$h(x) := f(x) - \frac{f(b) - f(a)}{g(b) - g(a)}(g(x) - g(a)).$$

Dann gilt $h(a) = h(b) = f(a)$ und

$$h'(x) = f'(x) - \frac{f(b) - f(a)}{g(b) - g(a)} g'(x).$$

Nach dem Satz von Rolle folgt die Behauptung. □

Proposition 7.21 (Die l'Hospitalsche Regel). *Seien $f, g : (a,b) \to \mathbb{R}$ differenzierbare Funktionen mit $g'(x) \neq 0$ für $x \in (a,b)$. Gilt (i) $f(x) \xrightarrow{x\downarrow a} 0$ oder (ii) $f(x) \xrightarrow{x\downarrow a} \infty$, und existiert $\lim_{x\downarrow a} f'(x)/g'(x)$, so gilt*

$$\lim_{x\downarrow a} \frac{f(x)}{g(x)} = \lim_{x\downarrow a} \frac{f'(x)}{g'(x)}.$$

Entsprechendes gilt für $x \uparrow b$, und alle Aussagen gelten auch für $a = -\infty$ und $b = \infty$.

Beweis. Im Fall (i) können wir f und g auf $[a,b)$ stetig mittels $f(a) = g(b) = 0$ fortsetzen. Nach Proposition 7.20 gibt es für jedes $x \in (a,b)$ ein $x' \in (a,x)$ mit

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f'(x')}{g'(x')}.$$

Außerdem geht mit $x \to a$ auch $x' \to a$, woraus die Aussage folgt. Im Fall (ii) sei zunächst $\frac{f'(x)}{g'(x)} \xrightarrow{x\to a} A \in \mathbb{R}$. Sei $\varepsilon > 0$. Mit dem verallgemeinerten Mittelwertsatz folgt, dass

$$\Big|\frac{f(x) - f(y)}{g(x) - g(y)} - A\Big| \leq \varepsilon$$

für alle $x, y \in (a, a+\delta)$ für ein geignetes $\delta > 0$. Sei $y \in (a, a+\delta)$ fest. Nun ist

$$\lim_{x\downarrow a} \frac{f(x) - f(y)}{g(x) - g(y)} = \lim_{x\downarrow a} \frac{f(x)}{g(x)} \frac{1 - f(y)/f(x)}{1 - g(y)/g(x)} = \lim_{x\downarrow a} \frac{f(x)}{g(x)}.$$

Also gilt

$$\lim_{x\downarrow a} \Big|\frac{f(x)}{g(x)} - A\Big| = \lim_{x\downarrow a} \Big|\frac{f(x) - f(y)}{g(x) - g(y)} - A\Big| \leq \varepsilon.$$

Da ε beliebig war, folgt die Behauptung. □

Beispiel 7.22. 1. Es gilt $\sin(x)/x \xrightarrow{x\to 0} 1$.
Da $\cos(x)/1 \xrightarrow{x\to 0} 1$, folgt die Aussage nach Proposition 7.21.

2. Es gilt $\log(x)/x \xrightarrow{x\to\infty} 0$.
Da $\frac{1/x}{1} \xrightarrow{x\to 0} 0$, folgt die Aussage nach Proposition 7.21.

7.5 Sonderfälle

Wir kommen nun zu zwei pathologischen Fällen. Zum einen ist dies eine differenzierbare Funktion mit unstetiger Ableitung, zum anderen eine überall stetige, aber nirgends differenzierbare Funktion.

Beispiel 7.23 (Eine Funktion mit unstetiger Ableitung). Die Funktion $f : x \mapsto x^2\sin(1/x)$ (mit $f(0) = 0$) ist überall (insbesondere in 0) differenzierbar, die Ableitung ist aber in 0 unstetig.
Denn: Wir berechnen für $x \neq 0$

$$\frac{d}{dx}x^2\sin(1/x) = -x^2\cos(1/x)\frac{1}{x^2} + 2x\sin(1/x) = 2x\sin(1/x) - \cos(1/x).$$

Damit kann die Ableitung nicht stetig nach 0 fortgesetzt werden. Für $x = 0$ ist

$$\left|\frac{f(x) - f(0)}{x - 0}\right| \leq \left|\frac{x^2}{x}\right| = |x| \xrightarrow{x\to 0} 0.$$

Damit ist klar, dass f' in 0 unstetig ist.

Beispiel 7.24 (Eine überall stetige aber nirgends differenzierbare Funktion). Wir definieren g_1 mittels

$$g_1(x) = \begin{cases} x - [x], & [x] \text{ gerade}, \\ [x] + 1 - x, & [x] \text{ ungerade}, \end{cases}$$

sowie $g_k(x) := 2^{-k}g_1(2^kx)$; siehe auch Abbildung 7.2. Wir behaupten, dass die Funktion

$$f(x) := \sum_{k=1}^{\infty} g_k(x)$$

gleichmäßig stetig, aber nirgends differenzierbar ist.
Zunächst zur Stetigkeit von f. Klar ist, dass $f_n := \sum_{k=1}^{n} g_k(x)$ für jedes n gleichmäßig stetig ist. Sei also $\varepsilon > 0$, n so groß, dass $2^{-(n-1)} < \varepsilon/2$ und $\delta > 0$ so klein, dass $|f_n(y) - f_n(x)| < \varepsilon/2$ für $|y - x| < \delta$. Nun gilt

$$|f(y) - f(x)| \leq |f_n(y) - f_n(x)| + \left|\sum_{k=n+1}^{\infty} g_k(y) + g_k(x)\right| \leq \varepsilon/2 + 2\sum_{k=n+1}^{\infty} 2^{-k} \leq \varepsilon.$$

Daraus folgt die gleichmäßige Stetigkeit von f.
Wir zeigen nun, dass f in 0 nicht differenzierbar ist; ein analoges Argument liefert dann, dass f nirgends differenzierbar ist. Sei $y_n = 2 \cdot 2^{-n}$, $z_n = -2 \cdot 2^{-n}$. Dann ist

$$g_m(y_n) = \begin{cases} y_n, & \text{falls } m < n, \\ 2^{-m}g_1(2 \cdot 2^{m-n}) = 0, & \text{falls } m \geq n. \end{cases}$$

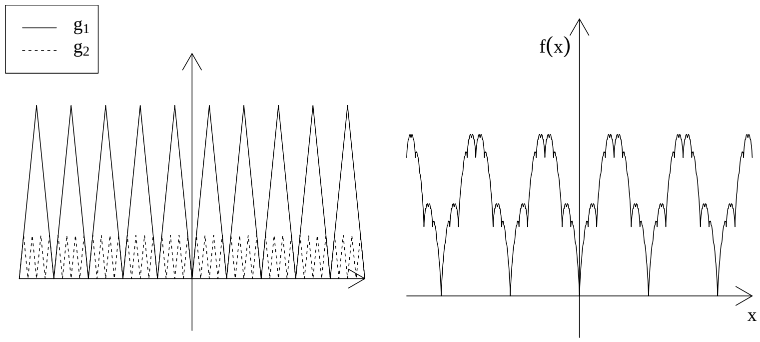

Abbildung 7.2: Die Abbildungen g_k sind stückweise differenzierbar und die Abbildung f ist nirgends differenzierbar.

und analog für z_n. Daraus folgt

$$\frac{f(y_n) - f(y)}{y_n - 0} = \frac{\sum_{k=1}^{n-1} y_n}{y_n} = n - 1,$$
$$\frac{f(z_n) - f(y)}{z_n - 0} = \frac{\sum_{k=1}^{n-1} |z_n|}{y_n} = -(n-1).$$

Insbesondere existiert die Ableitung in 0 nicht.

8 Integrale

Wir kommen nun zum Integralbegriff,

$$\int_a^b f(x)dx. \tag{8.1}$$

Hier wird das Zeichen $\int$ als Grenzwert einer Summe zu verstehen sein.

8.1 Konvergenz von Funktionen

Wir werden bei der Definition des Integrals (8.1) über die Funktion f auf stückweise konstante Funktionen zurückgreifen müssen, d.h. auf eine Folge von Funktionen, die die Funktion f approximieren. Deshalb beschäftigen wir uns zunächst allgemein mit der Konvergenz von Funktionen.

Definition 8.1 (Norm). *Sei E ein $\mathbb{R}$-Vektorraum. Eine Norm ist eine Abbildung $||.|| : E \to \mathbb{R}_+$ mit:*

1. *Positivität: Es gilt $||x|| = 0$ genau dann, wenn $x = 0$ ist.*

2. *Halblinearität: Für alle $\lambda \in \mathbb{R}$ gilt $||\lambda x|| = |\lambda| \cdot ||x||$.*

3. *Dreiecksungleichung: Es gilt $||x + y|| \leq ||x|| + ||y||$.*

Das Paar $(E, ||.||)$ heißt auch normierter Raum.

Beispiel 8.2. 1. Sei $E = \mathbb{R}^d$. Dann ist eine Norm $||.||_2$ mittels

$$||x|| := \Big(\sum_{i=1}^d x_i^2\Big)^{1/2}$$

definiert. Sie heißt auch *euklidische Norm*. Die Dreiecksungleichung folgt genau wie in Beispiel 5.2.2.

2. Sei E eine Menge. Auf dem Raum $\mathbb{R}^E$ (also dem Raum der Funktionen $f : E \to \mathbb{R}$) definieren wir die Supremumsnorm $||.||_\infty$ mittels

$$||f||_\infty := \sup_{x\in E} |f(x)|.$$

3. Jeder normierte Raum $(E, ||.||)$ ist ebenfalls ein metrischer Raum (E, r), wenn man $r(x, y) := ||x - y||$ setzt. Andersherum ist aber nicht jeder metrische Raum normiert. (Insbesondere müssen metrische Räume nicht unbedingt Vektorräume sein.)

Definition 8.3 (Punktweise und gleichmäßige Konvergenz). *1. Seien $f, f_1, f_2, ... : E \to E'$ Funktionen. Gilt $f_n(x) \xrightarrow{n\to\infty} f(x)$ für alle $x \in E$, so sagen wir, die Folge $(f_n)_{n=1,2,...}$* konvergiert punktweise *gegen f und wir schreiben $f_n \xrightarrow{n\to\infty}_{pktw} f$.*

2. Seien $f, f_1, f_2, ... : E \to \mathbb{R}$ Funktionen. Gilt

$$||f_n - f||_\infty \xrightarrow{n\to\infty} 0$$

so sagen wir, die Folge $(f_n)_{n=1,2,...}$ konvergiert gleichmäßig *gegen f und wir schreiben $f_n \xrightarrow{n\to\infty}_{glm} f$.*

3. Seien $f, f_1, f_2, ... : E \to \mathbb{R}$ Funktionen. Für eine kompakte Menge K seien $f|_K, f_1|_K, f_2|_K, ...$ die Einschränkungen der Funktionen auf K. Konvergiert $(f_n|_K)_{n=1,2,...}$ für alle kompakten Mengen $K \subseteq E$ gleichmäßig gegen $f|_K$ so sagen wir, die Folge $(f_n)_{n=1,2,...}$ konvergiert gleichmäßig auf Kompakta *gegen f und wir schreiben $f_n \xrightarrow{n\to\infty}_{glmK} f$.*

Lemma 8.4 (glm ⇒ glmK ⇒ pktw). *Seien $f, f_1, f_2, ... : E \to \mathbb{R}$ Funktionen. Konvergiert die Folge $(f_n)_{n=1,2,...}$ gleichmäßig gegen f, so konvergiert sie auch gleichmäßig auf Kompakta. Konvergiert sie gleichmäßig auf Kompakta, so auch punktweise.*

Beweis. Gilt $f_n \xrightarrow{n\to\infty}_{\text{glm}} f$ gleichmäßig gegen f und ist $K \subseteq E$ kompakt,so gilt

$$||f_n - f||_\infty := \sup_{x\in E} |f_n(x) - f(x)| \geq \sup_{x\in K} |f_n(x) - f(x)| = ||f_n|_K - f|_K||_\infty,$$

woraus die erste Behauptung folgt. Die zweite folgt, da jede Menge $\{x\}$ mit $x \in E$ kompakt ist. □

Beispiel 8.5. Für eine Menge $A \subseteq E$ definieren wir die *Indikatorfunktion*

$$1_A : \begin{cases} E & \to \{0, 1\} \\ x & \mapsto \begin{cases} 1, & x \in A, \\ 0, & x \in E \setminus A. \end{cases} \end{cases}$$

1. Sei $f_n : [0,1] \to [0,1]$ gegeben durch $f(x) = x^n$ und $f := 1_{\{1\}}$. Dann konvergiert $(f_n)_{n=1,2,...}$ punktweise gegen f, aber nicht gleichmäßig (auf Kompakta).
   Denn: Ist $0 \leq x < 1$ und $\varepsilon > 0$, so gibt es ein $N \in \mathbb{N}$ mit $x^n < \varepsilon$ für $n \geq N$ nach Lemma 2.18. Daraus und mit $f_n(1) = 1 = f(1)$ folgt bereits die punktweise Konvergenz. Allerdings gilt für jedes n
   $$||f_n - f|| = \sup_{0 \leq x \leq 1} |f_n(x) - f(x)| = \sup_{0 \leq x < 1} |f_n(x)| = 1,$$
   und da $[0,1]$ kompakt ist, konvergiert $(f_n)_{n=1,2,...}$ nicht gleichmäßig auf Kompakta gegen f.

Lemma 8.6 (Gleichmäßiger Grenzwert stetiger Funktionen ist stetig). *Seien $f, f_1, f_2, ... : E \to \mathbb{R}$ und f_n stetig, $n = 1, 2, ...$. Konvergiert $(f_n)_{n=1,2,...}$ gleichmäßig auf Kompakta gegen f, so ist auch f stetig.*

Beweis. Da f genau dann stetig ist, wenn jede der Funktionen $f|_K$ für $K \subseteq E$ kompakt stetig ist, genügt es, den Fall $||f_n - f||_\infty \xrightarrow{n\to\infty} 0$ zu betrachten.
Sei $x_0 \in E$ und $\varepsilon > 0$. Weiter sei N so groß, dass $||f_N - f||_\infty < \varepsilon/3$ und $\delta > 0$ klein genug, so dass $|f_N(x_0) - f_N(x)| < \varepsilon/3$ für $r(x_0, x) < \delta$. Damit gilt für solche x

$$|f(x) - f(x_0)| \leq |f(x) - f_N(x)| + |f_N(x) - f_N(x_0)| + |f_N(x_0) - f(x_0)| < \varepsilon.$$

Damit ist die Behauptung bewiesen. □

Zur Definition von Integralen benötigen wir folgende Aussage über Treppenfunktionen.

Definition 8.7 (Zerlegung und Treppenfunktion). *Sei $I = [a, b]$.*

1. *Eine Zerlegung von I ist eine Menge $Z := \{x_0, ..., x_n\}$ mit $a \leq x_1 < \cdots < x_n \leq b$. Wir bezeichnen $x_0, ..., x_n$ als Stützstellen von Z, setzen $\Delta x_i := x_i - x_{i-1}$ und definieren die* Feinheit *der Zerlegung durch*
   $$\Delta(Z) := \max_{i=1,...,n} \Delta x_i.$$
2. *Sei $Z = \{x_0, ..., x_n\}$ eine Zerlegung von I. Eine Funktion $f : I \to \mathbb{R}$, die auf (x_{i-1}, x_i) konstant ist, $i = 1, ..., n$, heißt* Treppenfunktion.

Proposition 8.8 (Approximation stetiger Funktionen durch Treppenfunktionen). *Sei $I = [a, b]$ ein Intervall und $f : I \to \mathbb{R}$ stetig. Dann gibt es eine Folge $(\varphi_n)_{n=1,2,...}$ von Treppenfunktionen mit $||f - \varphi_n||_\infty \xrightarrow{n\to\infty} 0$.*

Beweis. Nach Proposition 5.28 ist f auf I gleichmäßig stetig. Sei nun $\varepsilon > 0$. Wir müssen eine Treppenfunktion φ finden, für die $||f - \varphi||_\infty < \varepsilon$ gilt. Hierzu wählen wir $\delta > 0$, so dass $|f(x) - f(y)| < \varepsilon$ für $|x - y| < 2\delta$. Wir wählen $a = x_0 < \cdots < x_n = b$ mit $|x_i - x_{i-1}| < \delta$ sowie $y_i = f(x_i)$, $i = 1, ..., n$ und $\varphi(x) = y_i$ für $x \in (x_{i-1}, x_i]$ und $\varphi(x_0) = y_1$. Dann gilt für $x \in (x_{i-1}, x_i]$
$$|f(x) - \varphi(x)| = |f(x) - f(x_i)| < \varepsilon,$$
da $|x - x_i| < 2\delta$. Daraus folgt schon $||f - \varphi||_\infty < \varepsilon$, also die Behauptung. □

Wir erweitern nun noch Proposition 8.8 auf stückweise stetige Funktionen.

Definition 8.9 (Stückweise stetige Funktionen). *Sei* $I = [a, b]$ *und* $f : [a, b] \to \mathbb{R}$. *Gibt es disjunkte Intervalle* $I_1, ..., I_n$ *mit* $I = \bigcup_{k=1}^n I_k = I$ *und sind die Abbildungen* $f|_{I_k}$ *stetig,* $k = 1, ..., n$, *so heißt* f stückweise stetig.

Korollar 8.10 (Proposition 8.8 für stückweise stetige Funktionen). *Die Aussage von Proposition 8.8 gilt auch für stückweise stetige Funktionen* f.

Beweis. Stückweise stetiges f lässt sich nach Proposition 8.8 stückweise gleichmäßig durch Treppenfunktionen approximieren. Diese stückweise Approximation lässt sich nun leicht zu einer gleichmäßigen Approximation auf I zusammen setzen. □

8.2 Definition und grundlegende Eigenschaften

Ziel dieses Abschnittes ist es zu zeigen, dass alle stückweise stetigen Funktionen (Riemann-)integrierbar sind (Theorem 8.21). Hierzu sind folgende Schritten zentral:

1. Stückweise stetige Funktionen lassen sich gleichmäßig durch Treppenfunktionen approximieren (siehe das eben gezeigte Korollar 8.10).
2. Integration von Treppenfunktionen (Korollar 8.17).
3. Gleichmäßige Konvergenz von Funktionen liefert die Konvergenz der Integrale (Theorem 8.19).

Anschaulich betrachtet soll $\int_a^b f(x)dx$ für eine Funktion $f : I \to \mathbb{R}_+$ der Flächeninhalt von $\{(x, y) : 0 \leq y \leq f(x)\}$ darstellen. (Hier bezeichnet das Symbol $\int$ das Riemann-Integral.) Diesen Flächeninhalt müssen wir jedoch zunächst geeignet approximieren.

Definition 8.11 (Riemann'sche Summe und Integral). *Sei* $f : I = [a, b] \to \mathbb{R}$ *beschränkt.*

1. *Sei* $Z = \{x_0, ..., x_n\}$ *eine Zerlegung von* I *und* $\xi_i \in [x_{i-1}, x_i]$, $i = 1, ..., n$. *Dann heißt mit* $\underline{\xi} = (\xi_1, ..., \xi_n)$
$$s_{Z,\underline{\xi}}(f) := \sum_{i=1}^n f(\xi_i)\Delta_i$$
Riemann'sche Summe *von* f *zur Zerlegung* Z *mit Stützstellen* $\underline{\xi}$.

2. *Die Funktion* f *heißt* Riemann-integrierbar, *wenn es ein* $s \in \mathbb{R}$ *gibt, so dass gilt: Für jedes* $\varepsilon > 0$ *gibt es ein* $\delta > 0$, *so dass* $|s_{Z,\underline{\xi}}(f) - s| < \varepsilon$ *für jede Zerlegung* Z *von* I *mit* $\Delta(Z) < \delta$ *und beliebigen Stützstellen* $\xi_i \in [x_{i-1}, x_i]$. *Dann heißt* s *das* Integral *von* f *über* I *und schreiben auch*
$$s := \int_a^b f(x)dx$$
oder
$$s_{Z,\underline{\xi}}(f) \xrightarrow{\Delta(Z)\to 0} s.$$
Die Menge der Riemann-integrierbaren Funktionen $f : I \to \mathbb{R}$ *bezeichnen wir mit* $\mathcal{R}(I)$. *(Dies ist per Definition eine Teilmenge der Menge der beschränkten Funktionen* $f : I \to \mathbb{R}$.*)*

Lemma 8.12 (Integrierbare Funktionen sind ein Vektorraum). *Sei $I = [a,b]$. Die Menge $\mathcal{R}(I)$ ist ein Vektorraum und $\int_a^b : f \mapsto \int_a^b f(x)dx$ ist ein lineares Funktional, d.h. es gilt für $\lambda, \mu \in \mathbb{R}, f, g \in \mathcal{R}(I)$*

$$\int_a^b \lambda f(x) + \mu g(x)dx = \lambda \int_a^b f(x)dx + \mu \int_a^b g(x)dx.$$

Beweis. Klar ist, dass $f \mapsto s_{Z,\underline{\xi}}(f)$ (für eine feste Zerlegung Z und $\underline{\xi}$) linear ist. Deshalb schreiben wir

$$\begin{aligned}\int_a^b \lambda f(x) + \mu g(x)dx &= \lim_{\Delta(Z)\to 0} s_{Z,\underline{\xi}}(\lambda f + \mu g) = \lim_{\Delta(Z)\to 0} \lambda s_{Z,\underline{\xi}}(f) + \mu s_{Z,\underline{\xi}}(g) \\ &= \lambda \int_a^b f(x)dx + \mu \int_a^b g(x)dx.\end{aligned}$$

□

Beispiel 8.13.
1. Sei $c \in \mathbb{R}$ und $f = c$ konstant. Dann ist $\int_a^b f(x)dx = c(b-a)$.
   Denn: Sei Z eine beliebige Zerlegung von $[a,b]$ und $\underline{\xi}$ eine beliebige Wahl von Stützstellen. Dann gilt
   $$s_{Z,\underline{\xi}}(f) = \sum_{i=1}^n f(\xi_i)\Delta x_i = c \sum_{i=1}^n x_i - x_{i-1} = c(x_n - x_0) = c(b-a).$$
2. Sei $f = 1_{\mathbb{Q}}$ die Indikatorfunktion von $\mathbb{Q}$ und $I = [a,b]$. Dann ist f nicht über I Riemann-integrierbar.
   Denn: Sei $Z = \{x_0, ..., x_n\}$ eine beliebige Zerlegung von I. Dann gibt es $\xi_i \in [x_{i-1}, x_i] \cap \mathbb{Q}$ sowie $\zeta_i \in [x_{i-1}, x_i] \cap (\mathbb{R} \setminus \mathbb{Q})$. Also gilt
   $$\begin{aligned}s_{Z,\underline{\xi}}(f) &= \sum_{i=1}^n f(\xi_i)\Delta x_i = b - a, \\ s_{Z,\underline{\zeta}}(f) &= \sum_{i=1}^n f(\zeta_i)\Delta x_i = 0,\end{aligned}$$
   da $f(\xi_i) = 1$ aber $f(\zeta_i) = 0$ ist. Insbesondere existiert das Riemann-Integral nicht.

Lemma 8.14 (Monotonie des Integrals). *Sei $I = [a,b]$ und $f, g \in \mathcal{R}(I)$ mit $f \leq g$. Dann gilt $\int_a^b f(x)dx \leq \int_a^b g(x)dx$. Insbesondere gilt: ist $|f| \in \mathcal{R}(I)$, so gilt $|\int_a^b f(x)dx| \leq \int_a^b |f(x)|dx$.*

Beweis. Für jede Zerlegung $Z = \{x_0, ..., x_n\}$ und Stützstellen $\underline{\xi}$ gilt

$$s_{Z,\underline{\xi}}(f) = \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}) \leq \sum_{i=1}^n g(\xi_i)(x_i - x_{i-1}) = s_{Z,\underline{\xi}}(g).$$

Nun folgt die Behauptung, indem man auf beiden Seite $\Delta(Z) \to 0$ betrachtet. □

Lemma 8.15. *Sei $I = [a,b]$ und $f, g : I \to \mathbb{R}$, $f \in \mathcal{R}(I)$, sowie $f(x) = g(x)$ für $x \notin \{y_1, ..., y_k\}$. Dann ist $g \in \mathcal{R}(I)$ und $\int_a^b f(x)dx = \int_a^b g(x)dx$.*

Beweis. Sei $Z = \{x_0, ..., x_n\}$ eine Zerlegung von I, sowie $\xi_i \in [x_{i-1}, x_i]$ und $\Delta_i := x_i - x_{i-1}$. Wir schreiben

$$\Big|s_{Z,\underline{\xi}}(g) - \int_a^b f(x)dx\Big| \leq \Big|\sum_{i=1}^n (g(\xi_i) - f(\xi_i))\Delta_i\Big| + \Big|s_{Z,\underline{\xi}}(f) - \int_a^b f(x)dx\Big|$$
$$= \Big|\sum_{i:\exists j:\xi_i = y_j} (g(\xi_i) - f(\xi_i))\Delta_i\Big| + \Big|s_{Z,\underline{\xi}}(f) - \int_a^b f(x)dx\Big| \xrightarrow{\Delta(Z)\to 0} 0,$$

woraus alle Behauptungen folgen. □

Lemma 8.16. *Sei $I = [a, b]$ und $a = a_0 < \cdots < a_n = b$ eine Zerlegung von I. Für $I_k := [a_{k-1}, a_k]$ sei $f|_{I_k} \in \mathcal{R}(I_k)$. Dann ist f in $\mathcal{R}(I)$ und es gilt*

$$\int_a^b f(x)dx = \sum_{k=1}^n \int_{a_{k-1}}^{a_k} f|_{I_k}(x)dx.$$

Beweis. Zunächst beweisen wir, dass $f1_{I_k} \in \mathcal{R}(I)$ und

$$\int_{a_{k-1}}^{a_k} f|_{I_k}(x)dx = \int_a^b f(x)1_{I_k}(x)dx. \tag{8.2}$$

Sei hierzu Z eine Zerlegung von I und $Z' := Z\cup\{x_{k-1}, x_k\}$ die Zerlegung von I, die zusätzlich die Punkte x_{k-1} und x_k beinhaltet sowie $Z|_{I_k} := Z' \cap I_k$ die Zerlegung von I_k, die genau aus den Punkten in Z' besteht, die in I_k enthalten sind. Dann gilt

$$\big|s_{Z,\underline{\xi}}(f1_{I_k}) - s_{Z|_{I_k},\underline{\xi}}(f|_{I_k})\big| \leq \big|s_{Z,\underline{\xi}}(f1_{I_k}) - s_{Z',\underline{\xi}}(f1_{I_k})\big| + \big|s_{Z',\underline{\xi}}(f1_{I_k}) - s_{Z|_{I_k},\underline{\xi}}(f|_{I_k})\big|$$
$$\leq 2||f||_\infty \Delta(Z) \xrightarrow{\Delta(Z)\to 0} 0,$$

woraus (8.2) folgt. Es bleibt nun noch zu bemerken, dass f und $\sum_{k=1}^n f1_{I_k}$ bis auf endlich viele Stellen (nämlich $x_1, ..., x_{n-1}$) übereinstimmen und das Integral nach Lemma 8.12 linear ist. Nun folgt die Aussage aus Lemma 8.15. □

Korollar 8.17 (Integration von Treppenfunktionen). *Sei $I = [a, b]$ und $f : I \to \mathbb{R}$ eine Treppenfunktion, die auf Intervallen $I_k := (a_{k-1}, a_k)$ den wert c_k annimmt, $k = 1, ..., n$. Dann gilt*

$$\int_a^b f(x)dx = \sum_{k=1}^n c_k(a_k - a_{k-1}).$$

Beweis. Da $f|_{I_k}$ bis auf höchstens die beiden Randpunkte mit $c_k|_{I_k}$ übereinstimmt, und wir das Integral konstanter Funktionen bereits in Beispiel 8.13 berechnet haben, folgt die Aussage aus Lemma 8.15 und 8.16. □

Lemma 8.18. *Sei $I = [a, b]$ und $f \in \mathcal{R}(I)$. Dann gilt*

$$\Big|\int_a^b f(x)dx\Big| \leq ||f||_\infty(b - a).$$

Beweis. Sei $Z = \{x_1, ..., x_n\}$ eine Zerlegung von I und $\underline{\xi}$ Stützstellen. Dann gilt

$$\left|s_{Z,\underline{\xi}}(f)\right| = \left|\sum_{k=1}^{n} f(\xi_k)\Delta x_k\right| \leq \sum_{k=1}^{n} ||f||_\infty \Delta x_k \Big| = ||f||_\infty (b-a). \tag{8.3}$$

Mit $\Delta(Z) \to 0$ folgt die Aussage. □

Theorem 8.19 (Gleichmäßige Konvergenz und Integration). *Sei $I = [a,b], f : I \to \mathbb{R}$ und $(f_n)_{n=1,2,...}$ eine $\mathcal{R}(I)$-wertige Folge von Funktionen. Falls $f_n \xrightarrow{n\to\infty}_{glm} f$, so ist $f \in \mathcal{R}(I)$ und es gilt*

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

Beweis. Zunächst ist f wegen $||f||_\infty \leq ||f - f_k||_\infty + ||f_k||_\infty < \infty$ beschränkt. Nun gilt

$$\Big|\int_a^b f_k(x)dx - \int_a^b f_l(x)dx|| \leq ||f_k - f_l||_\infty \cdot (b-a).$$

Insbesondere ist wegen der gleichmäßigen Konvergenz die Folge $\left(\int_a^b f_n(x)dx\right)_{n=1,2,...}$ eine Cauchy-Folge. Wir setzen $s := \lim_{n\to\infty} \int_a^b f_n(x)dx$. Für eine Zerlegung Z von I und Stützstellen $\underline{\xi}$ schreiben wir

$$\left|s_{Z,\xi}(f) - s\right| \leq \underbrace{\left|s_{Z,\xi}(f - f_k)\right|}_{\substack{\leq ||f-f_k||_\infty (b-a) \\ \text{wegen (8.3)}}} + \underbrace{\left|s_{Z,\xi}(f_k) - \int_a^b f_k(x)dx\right|}_{\xrightarrow{\Delta(Z)\to 0} 0} + \underbrace{\left|\int_a^b f_k(x)dx - s\right|}_{\xrightarrow{k\to\infty} 0}$$

Da die linke Seite unabhängig von k ist, gilt also

$$\left|s_{Z,\xi}(f) - s\right| \xrightarrow{\Delta(Z)\to 0} 0.$$

Damit sind alle Behauptungen bewiesen. □

Bemerkung 8.20. Wir haben in Lemma 8.4 gesehen, dass gleichmäßig konvergente Folgen von Funktionen auch punktweise konvergieren. Man kann sich also fragen, ob man in Theorem 8.19 die gleichmäßige Stetigkeit durch die schwächere Bedingung der punktweisen Stetigkeit ersetzen kann. Mit anderen Worten:
Gilt $\int_a^b f_n(x)dx \xrightarrow{n\to\infty} \int_a^b f(x)dx$ schon dann, falls $(f_n)_{n=1,2,...}$ eine $\mathcal{R}(I)$-wertige Folge ist und $f_n \xrightarrow{n\to\infty}_{pktw} f$?
Wir geben zwei Antworten:

1. Ist $(f_n)_{n=1,2,...}$ eine $\mathcal{R}(I)$-wertige Folge und $f_n \xrightarrow{n\to\infty}_{pktw} f$, so gilt nicht notwendigerweise $f \in \mathcal{R}(I)$.
   Denn: Sei $(x_n)_{n=1,2,...}$ eine Abzählung von $\mathbb{Q} \cap [0,1]$ und $Q_n := \{x_1, ..., x_n\}$. Dann ist $1_{Q_n} \in \mathcal{R}(I)$ nach Lemma 8.15 mit $\int_0^1 1_{Q_n}(x)dx = 0$. Außerdem gilt $1_{Q_n} \xrightarrow{n\to\infty}_{pktw} 1_{\mathbb{Q}\cap[0,1]}$, aber $1_{\mathbb{Q}\cap[0,1]} \notin \mathcal{R}(I)$ nach Beispiel 8.13.2.

2. Ist $(f_n)_{n=1,2,...}$ eine $\mathcal{R}(I)$-wertige Folge und $f_n \xrightarrow{n\to\infty}_{pktw} f$ mit $\in \mathcal{R}(I)$, so gilt nicht notwendigerweise $\int_a^b f_n(x)dx \xrightarrow{n\to\infty} \int_a^b f(x)dx$.
   Denn: Sei $I = [0,1]$ und $f_n = n \cdot 1_{(0,1/n)}$. Dann gilt $f_n \xrightarrow{n\to\infty}_{\text{pktw}} 0 =: f$, aber
$$\int_0^1 f_n(x)dx = \int_0^{1/n} ndx = 1 \neq 0 = \int_0^1 f(x)dx.$$

Theorem 8.21. *Sei $I = [a,b]$ und $f : I \to \mathbb{R}$ stückweise stetig, siehe Definition 8.9. Dann ist $f \in \mathcal{R}(I)$.*

Beweis. Der Beweis ist dank unserer Vorarbeit einfach: nach Proposition 8.8 ist f gleichmäßig durch Treppenfunktionen φ_n approximierbar. Also folgt die Aussage aus Theorem 8.19. □

Bemerkung 8.22 (Unter- und Obersummen). Es gibt eine alternative Möglichkeit, das Riemann- Integral zu konstruieren, nämlich über Ober- und Untersummen. Sei $I = [a,b]$ und $f : I \to \mathbb{R}$ beschränkt. Für eine Zerlegung $Z = \{x_0, ..., x_n\}$ von I, $I_k := [x_{k-1}, x_k]$ setzen wir
$$\overline{s}_Z(f) := \sum_{i=1}^n \sup_{x\in I_k} f(x)(x_i - x_{i-1}),$$
$$\underline{s}_Z(f) := \sum_{i=1}^n \inf_{x\in I_k} f(x)(x_i - x_{i-1}).$$

Nun definieren wir das Ober- und Unterintegral durch
$$\int_a^{*,b} f(x)dx := \inf\{\overline{s}_Z(f) : Z \text{ Zerlegung von } [a,b]\},$$
$$\int_{*,a}^b f(x)dx := \sup\{\underline{s}_Z(f) : Z \text{ Zerlegung von } [a,b]\}.$$

(Da f beschränkt ist, sind die Mengen auf der rechten Seite nach unten bzw. oben beschränkte Mengen reeller Zahlen, womit das Infimum bzw. Supremum existiert.)

Nun folgt: Es gilt immer
$$\int_a^{*,b} f(x)dx \leq \int_{*,a}^b f(x)dx.$$
Es ist $f \in \mathcal{R}(I)$ genau dann, wenn Gleichheit des Ober- und Unterintegrals gilt. In diesem Fall ist $\int_a^b f(x)dx$ gerade dem Wert des Ober- oder Unterintegrals.

Denn: Um die Ungleichung zu zeigen seien Z, Z' zwei Zerlegungen von $[a,b]$ und $Z \cup Z'$. Dann gilt
$$\underline{s}_Z(f) \leq \underline{s}_{Z\cup Z'}(f) \leq \overline{s}_{Z\cup Z'}(f) \leq \overline{s}_{Z'}(f).$$

Hieraus folgt auch $\sup_Z \underline{s}_Z(f) \le \inf_{Z'} \overline{s}_Z(f)$. Ist nun $f \in \mathcal{R}(I)$ und $\varepsilon > 0$. Dann gibt es eine Zerlegung Z so dass

$$\int_a^b f(x)dx - \varepsilon \le \inf_{\underline{\xi}} s_{Z,\underline{\xi}} = \underline{s}_Z(f) \le \int_{*,a}^b f(x)dx \le \int_a^{*,b} f(x)dx \le \overline{s}_Z(f) = \sup_{\underline{\xi}} s_{Z,\underline{\xi}}(f)$$
$$\le \int_a^b f(x)dx + \varepsilon.$$

Mit $\varepsilon \to 0$ folgt die Behauptung.
Gilt andersherum die Gleichheit von Ober- und Unterintegral, so gibt es für $\varepsilon > 0$ eine Zerlegung Z mit

$$s_{Z,\underline{\xi}}(f) - \varepsilon \le \overline{s}_Z(f) - \varepsilon \le \int_a^{*,b} f(x)dx = \int_{*,a}^b f(x)dx \le s_{Z,\underline{\xi}}(f) + \varepsilon$$

und die Behauptung folgt mit $\varepsilon \to 0$.

8.3 Der Hauptsatz der Differential- und Integralrechnung

In diesem Kapitel lernen wir Stammfunktionen kennen, mit denen wir Integrale berechnen können. Insbesondere geht es um die Funktion

$$x \mapsto \int_{x_0}^x f(y)dy$$

für eine Riemann-integrierbare Funktion f.

Bemerkung 8.23 (Notation). Sei $I = [a, b]$ und $f \in \mathcal{R}(I)$. Für $x_0, x \in [a, b]$ mit $x_0 \le x$ ist dann $f1_{[x_0,x]} \in \mathcal{R}(I)$ und $f|_{[x_0,x]} \in \mathcal{R}([x_0, x])$. Außerdem gilt (ähnlich wie in Lemma 8.16)

$$\int_{x_0}^x f(y)dy := \int_{x_0}^x f|_{[x_0,x]}(y)dy = \int_a^b (f1_{[x_0,x]})(y)dy.$$

Ist $x_0 > x$, so schreiben wir

$$\int_{x_0}^x f(y)dy := -\int_x^{x_0} f(y)dy.$$

Definition 8.24 (Stammfunktion). *Sei $I = (a, b)$ und $f : I \to \mathbb{R}$. Eine Funktion $F : (a, b) \to \mathbb{R}$ heißt* Stammfunktion *von f, wenn $F' = f$ gilt.*

Bemerkung 8.25 (Aufleitung). Ist F eine Stammfunktion von f, also $F' = f$, so sagt man auch, dass F die *Aufleitung* von f ist. (Im Englischen ist dies ähnlich: *Ableitung* heißt hier *derivative*, und *Aufleitung* heißt *anti-derivative*).

Lemma 8.26. *Sei $I = (a, b)$, $f : I \to \mathbb{R}$ und F, G Stammfunktionen von f. Dann ist $F - G$ konstant.*

Beweis. Es gilt $F' - G' = f - f = 0$. Mit Korollar 7.18 folgt die Behauptung. □

Nun kommen wir bereits zum Hauptsatz.

Theorem 8.27 (Hauptsatz der Differential- und Integralrechnung). *Sei $I = [a, b]$ und $f : I \to \mathbb{R}$ stetig. Dann ist für jedes $x_0 \in I$ die Funktion*

$$F : \begin{cases} I & \to \mathbb{R} \\ x & \mapsto \int_{x_0}^{x} f(y)dy \end{cases}$$

eine Stammfunktion von f.

Bemerkung 8.28. In der Situation des Theorems schreiben wir auch

$$\int f(x)dx = F(x)$$

und sagen, dass $\int f(x)dx$ ein *unbestimmtes Integral* ist, d.h. ein Integral mit unbestimmten Grenzen.

Beweis. Wir müssen $F' = f$ zeigen. Wir schreiben, weil $\mathcal{R}(I)$ ein Vektorraum ist, für $h > 0$,

$$\begin{aligned} \Big|\frac{F(x+h) - F(x)}{h} - f(x)\Big| &= \Big|\frac{1}{h}\int_x^{x-h} f(y)dy - f(x)\Big| \\ &\le \frac{1}{h}\int_x^{x+h} \sup_{x \le z \le x+h} |f(z) - f(x)| dy = \sup_{x \le z \le x+h} |f(z) - f(x)| \xrightarrow{h \downarrow 0} 0 \end{aligned}$$

wegen der Stetigkeit von f. Analog folgt derselbe Grenzwert für $h \uparrow 0$, woraus dann die Behauptung folgt. □

Korollar 8.29 (Berechnung von Integralen). *Sei $I = [a, b]$ und $f \in \mathcal{C}(I)$, sowie F eine Stammfunktion von f. Dann gilt*

$$\int_a^b f(x)dx = F(b) - F(a) =: F(x)|_{x=a}^{x=b}.$$

Beweis. Nach Theorem 8.27 ist die Abbildung $G : x \mapsto \int_a^x f(y)dy$ eine Stammfunktion von f. Außerdem gilt $G(a) = 0$, also gilt nach Lemma 8.26 $F(b) - G(b) = F(a)$ oder auch $G(b) = F(b) - F(a)$, also die Behauptung. □

Theorem 8.30 (Partielle Integration). *Sei $I = [a, b]$ und $f, g \in \mathcal{C}^1(I)$. Dann gilt*

$$\int_a^b f'(x)g(x)dx = f(x)g(x)|_{x=a}^{x=b} - \int_a^b f(x)g'(x)dx.$$

Beweis. Es gilt nach der Produktregel $(fg)' = f'g + fg'$. Mit anderen Worten ist nach Korollar 8.29

$$f(b)g(b) - f(a)g(a) = \int_a^b f'(x)g(x) + f(x)g'(x)dx,$$

woraus die Behauptung folgt. □

Beispiel 8.31. 1. Es gilt

$$\int \log x dx = -x + x \log x.$$

Denn:

$$\int \log(x)dx = \int 1 \cdot \log(x)dx = x \log x - \int x \cdot \frac{1}{x}dx = x \log x - x.$$

2. Es gilt
$$\int \sqrt{1-x^2}dx = \frac{1}{2}\Big(x\sqrt{1-x^2} + \arcsin(x)\Big).$$
Denn:
$$\begin{aligned}\int \sqrt{1-x^2}dx = \int 1 \cdot \sqrt{1-x^2}dx = x\sqrt{1-x^2} + \int \frac{2x^2}{2\sqrt{1-x^2}}dx \\ = x\sqrt{1-x^2} + \int \frac{1}{\sqrt{1-x^2}} - \frac{1-x^2}{\sqrt{1-x^2}}dx \\ = x\sqrt{1-x^2} + \arcsin(x) - \int \frac{1-x^2}{\sqrt{1-x^2}}dx.\end{aligned}$$

3. Es gilt
$$\int \cos^2(x)dx = \tfrac{1}{2}\big(x + \tfrac{1}{2}\sin(2x)\big).$$
Denn:
$$\begin{aligned}\int \cos^2(x)dx = \sin(x)\cos(x) + \int \sin^2(x)dx = \tfrac{1}{2}\sin(2x) + \int 1 - \cos^2(x)dx \\ = \tfrac{1}{2}\sin(2x) + x - \int \cos^2(x)dx.\end{aligned}$$

Theorem 8.32 (Substitutionsregel). *Sei $I = [a,b]$ und $J = [\alpha, \beta]$, $f \in \mathcal{C}(I)$ und $\varphi : I \to J$ mit $\varphi \in \mathcal{C}^1(I)$. Ist F eine Stammfunktion von f, dann ist $F \circ \varphi$ eine Stammfunktion von $(f \circ \varphi) \cdot \varphi'$ und es gilt*
$$\int_a^b f(\varphi(x)) \cdot \varphi'(x)dx = \int_{\varphi(a)}^{\varphi(b)} f(x)dx.$$

Beweis. Zuerst ist $(F \circ \varphi)' = (f \circ \varphi) \cdot \varphi'$, woraus sich die erste Behauptung ergibt. Nach Korollar 8.29 sind nun beide Seiten der zu zeigenden Gleichung gerade $F(\varphi(b)) - F(\varphi(a))$. □

Beispiel 8.33. 1. Es gilt
$$2\int_a^b \frac{\log(x)}{x}dx = ((\log(b))^2 - (\log(a))^2).$$
Denn: Wir verwenden Theorem 8.32 mit $\varphi(x) = \log(x)$ und $f(x) = x$. Dann gilt nämlich
$$\begin{aligned}\int_a^b \frac{\log(x)}{x}dx = \int_a^b f(\varphi(x)) \cdot \varphi'(x)dx = \int_{\varphi(a)}^{\varphi(b)} f(x)dx = \int_{\log(a)}^{\log(b)} xdx \\ = \tfrac{1}{2}((\log(b))^2 - (\log(a))^2).\end{aligned}$$

2. Es gilt
$$2\int_a^b x \cdot \exp(x^2)dx = \exp(b^2) - \exp(a^2).$$
Denn: Wir verwenden Theorem 8.32 mit $\varphi(x) = x^2$ und $f(x) = \exp(x)$. Dann gilt nämlich
$$\begin{aligned}\int_a^b 2x \cdot \exp(x^2)dx = \int_a^b f(\varphi(x)) \cdot \varphi'(x)dx = \int_{\varphi(a)}^{\varphi(b)} f(x)dx = \int_{a^2}^{b^2} \exp(x)dx \\ = \exp(b^2) - \exp(a^2).\end{aligned}$$

3. Sei $I = [a, b]$ und $g \in \mathcal{C}^1(I)$. Dann gilt

$$\int_a^b \frac{g'(x)}{g(x)} dx = \log(g(b)) - \log(g(a)).$$

Denn: Wir verwenden Theorem 8.32 mit $\varphi(x) = g(x)$ und $f(x) = 1/x$. Dann gilt nämlich

$$\int_a^b \frac{g'(x)}{g(x)} dx = \int_a^b f(\varphi(x)) \cdot \varphi'(x) dx = \int_{\varphi(a)}^{\varphi(b)} f(x) dx = \int_{g(a)}^{g(b)} \frac{1}{x} dx$$
$$= \log(g(b)) - \log(g(a)).$$

Bemerkung 8.34 (Anwendung von Theorem 8.32). Um Theorem 8.32 anwenden zu können, ist die dx-Schreibweise nützlich. Betrachten wir hierzu Beispiel 8.33.2. Wir setzen $y = x^2$ und damit $\frac{dy}{dx} = 2x$ oder auch $dx = dy/(2x)$. Schreibt man formal

$$\int_a^b 2x \exp(x^2) dx = \int_{a^2}^{b^2} 2x \exp(y) \tfrac{dy}{2x} = \int_{a^2}^{b^2} \exp(y) dy = e^{b^2} - e^{a^2},$$

so erhält man genau das richtige Resultat.

Wir beenden diesen Abschnitt mit einer mit Theorem 7.14 verwandten Aussage,

Proposition 8.35 (Mittelwertsatz der Integralrechnung). *Sei $I = [a, b]$ und $f, \varphi \in \mathcal{C}(I)$ mit $\varphi \geq 0$. Dann gibt es ein $z \in I$ mit*

$$\int_a^b f(x)\varphi(x) dx = f(z) \int_a^b \varphi(x) dx. \tag{8.4}$$

Beweis. Zunächst ist nach Lemma 8.14

$$\inf_{y \in I} f(y) \int_a^b \varphi(x) dx \leq \int_a^b f(x)\varphi(x) dx \leq \sup_{y \in I} f(y) \int_a^b \varphi(x) dx.$$

Die auf $(\inf_{y\in I} f(y), \sup_{y\in I} f(y))$ definierte Abbildung $m \mapsto m \int_a^b \varphi(x) dx$ ist stetig, also gibt es nach dem Zwischenwertsatz ein $m \in (\inf_{y\in I} f(y), \sup_{y\in I} f(y))$ mit $m \int_a^b \varphi(x) dx = \int_a^b f(x)\varphi(x) dx$. Da f stetig ist, nimmt f jeden Wert in $(\inf_{y\in I} f(y), \sup_{y\in I} f(y))$ an, also gibt es ein $z \in I$ für das (8.4) gilt. □

8.4 Uneigentliche Integrale

Bisher haben wir nur Integrale über kompakte Mengen $I = [a, b]$ betrachtet. Dies werden wir nun erweitern.

Definition 8.36 (Uneigentliches Integral). *Sei $-\infty \leq a < b \leq \infty$. Ist $I = [a, b)$ und $f : I \to \mathbb{R}$, so definieren wir im Falle der Konvergenz*

$$\int_a^b f(x) dx = \lim_{\beta \uparrow b} \int_a^\beta f(x) dx.$$

(Eine analoge Definition verwenden wir im Falle $I = (a, b]$ und auch für $I = (a, b)$.) Wir nennen dieses Integral auch uneigentliches Integral. *Existiert $\int_a^b |f(x)| dx$, so sagen wir, dass f absolut über I integrierbar ist.*

Beispiel 8.37. 1. Es gilt für $c > 0$

$$\int_0^\infty e^{-cx} dx = \frac{1}{c}.$$

Denn: Wir schreiben

$$\int_0^\beta e^{-cx} dx = -\frac{1}{c} e^{-cx}|_{x=0}^{x=\beta} = \frac{1}{c}(1 - e^{-c\beta}) \xrightarrow{\beta\to\infty} \frac{1}{c}.$$

2. Es gilt

$$\int_0^\infty \frac{1}{1+x^2} dx = \frac{\pi}{2}.$$

Denn:

$$\int_0^\beta \frac{1}{1+x^2} dx = \arctan x_{x=0}^{x=\beta} = \arctan \beta \xrightarrow{\beta\to\infty} \frac{\pi}{2}.$$

3. Es gilt

$$\int_1^\infty x^{-s} dx = \begin{cases} \frac{1}{s-1}, & \text{für } s > 1, \\ \infty, & \text{sonst.} \end{cases}$$

Denn: Wir berechnen für $s \neq 1$

$$\int_1^\beta x^{-s} dx = \frac{1}{1-s} x^{-s+1}|_{x=1}^{x=\beta} = \frac{1}{1-s}\big(x^{-\beta+1} - 1\big),$$

woraus die Behauptung mit $\beta \to \infty$ folgt. Für $s = 1$ ist

$$\int_1^\beta x^{-1} dx = \log(x)|_{x=1}^{x=\beta} = \log(\beta) \xrightarrow{\beta\to\infty} \infty.$$

4. Es gilt

$$\int_0^1 x^{-s} dx = \begin{cases} \frac{1}{1-s}, & \text{für } s < 1, \\ \infty, & \text{sonst.} \end{cases}$$

Denn: Wir berechnen für $s \neq 1$

$$\int_\alpha^1 x^{-s} dx = \frac{1}{1-s} x^{-s+1}|_{x=\alpha}^{x=1} = \frac{1}{1-s}\big(1 - x^{-\alpha+1} - 1\big),$$

woraus die Behauptung mit $\alpha \to 0$ folgt. Für $s = 1$ ist

$$\int_\alpha^1 x^{-1} dx = \log(x)|_{x=\alpha}^{x=1} = -\log(\alpha) \xrightarrow{\alpha\to 0} \infty.$$

Beispiel 8.38. Wir zeigen noch, dass die Funktion $f : x \mapsto \frac{\sin(x)}{x}$ über $[0, \infty)$ zwar integrierbar, aber nicht absolut integrierbar ist. Um die Integrierbarkeit zu zeigen, schreiben wir mittels partieller Integration

$$\int_1^\beta \frac{\sin(x)}{x} dx = \frac{\cos(x)}{x}\Big|_{x=1}^{x=\beta} - \int_1^\beta \frac{\cos(x)}{x^2} dx.$$

Das letzte Integral konvergiert, weil

$$\Big|\int_1^\beta \frac{\cos(x)}{x^2}dx\Big| \leq \int_1^\beta \Big|\frac{\cos(x)}{x^2}\Big|dx \leq \int_1^\beta \Big|\frac{1}{x^2}\Big|dx \xrightarrow{\beta\to\infty} 1.$$

Da f auf $[0,\infty)$ stetig ist folgt ihre Integrierbarkeit.

Nun zeigen wir, dass f nicht absolut integrierbar ist. Hierzu bemerken wir, dass

$$\int_{k\pi}^{(k+1)\pi} \Big|\frac{\sin(x)}{x}\Big|dx \geq \frac{1}{(k+1)\pi}\int_{k\pi}^{(k+1)\pi} \sin(x)dx = \frac{2}{(k+1)\pi}.$$

Deswegen gilt

$$\int_0^\infty \Big|\frac{\sin(x)}{x}\Big|dx = \sum_{k=0}^\infty \int_{k\pi}^{(k+1)\pi} \Big|\frac{\sin(x)}{x}\Big|dx \geq \sum_{k=0}^\infty \frac{2}{(k+1)\pi} = \infty.$$

9 Verschiedenes

In diesem Abschnitt sammeln wir noch ein paar Anwendungen von bisher Gelerntem. Insbesondere mag es als Wiederholung einiger Techniken dienen.

9.1 Die Gamma- und Beta-Funktion

Die Gamma- und Beta-Funktion tauchen an vielen Stellen in der Mathematik auf. Beispielsweise werden wir in Lemma 9.2 sehen, dass $\Gamma(n+1) = n!$ für $n = 1, 2, ...$

Definition 9.1 (Gamma- und Beta-Funktion). *Wir definieren die Gamma-Funktion* $\Gamma : (0,\infty) \to \mathbb{R}$ *mittels*

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt,$$

sowie die Beta-Funktion $B : (0,\infty)^2 \to \mathbb{R}$ *mittels*

$$B(x,y) := \int_0^1 t^{x-1}(1-t)^{y-1}dt.$$

Wir zeigen nun ein paar Aussagen über die Gamma- und Beta-Funktion.

Lemma 9.2 (Eigenschaften der Gamma- und Beta-Funktion).

1. *Für* $x > 0$ *ist*

$$\Gamma(x+1) = x\Gamma(x).$$

2. *Für* $x, y > 0$ *gilt*

$$\frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} = B(x,y).$$

3. *Es gilt*

$$\Gamma(1) = 1, \qquad \Gamma(1/2) = \sqrt{\pi}.$$

Aus 1. und 3. folgt insbesondere $\Gamma(n+1) = n!$.

Beweis. Mittels partieller Integration folgt 1. aus

$$\Gamma(x+1) = \int_0^\infty t^x e^{-t} dt = -t^x e^{-t}|_0^\infty + \int_0^\infty x t^{x-1} e^{-t} dt = x\Gamma(x).$$

Für 2. berechnen wir[16] mit Hilfe der Substitutionsregel aus Theorem 8.32

$$\begin{aligned}\Gamma(x)\Gamma(y) &= \int_0^\infty \int_0^\infty s^{x-1} t^{y-1} e^{-(s+t)} ds dt \\ &\overset{r=t+s}{=} \int_0^\infty \int_0^r s^{x-1}(r-s)^{y-1} e^{-r} ds dr \\ &\overset{u=s/r}{=} \int_0^\infty \int_0^1 r u^{x-1} r^{x-1} r^{y-1} (1-u)^{y-1} e^{-r} du dr \\ &= \Gamma(x+y) \cdot \int_0^1 u^{x-1}(1-u)^{y-1} du.\end{aligned}$$

und die Aussage folgt. Für 3. gilt erst einmal durch einfache Integration

$$\Gamma(1) = \int_0^\infty e^{-t} dt = 1.$$

Für die zweite Aussage verwenden wir 2. und schreiben

$$\Gamma(1/2)^2 = B(1/2, 1/2) = \int_0^1 \frac{dt}{\sqrt{t(1-t)}} \overset{s=2t-1}{=} \int_{-1}^1 \frac{ds}{\sqrt{1-s^2}} = \arcsin(s)|_{-1}^1 = \pi$$

und das Lemma ist bewiesen. □

Nun berechnen wir noch das Gauss'sche Integral.

Proposition 9.3 (Gauss'sches Integral). *Es gilt*

$$\int_{-\infty}^\infty e^{-\frac{1}{2}x^2} dx = \sqrt{2\pi}.$$

Beweis. Wir verwenden wieder den Transformationssatz mit $y = \frac{1}{2}x^2$, d.h. $\frac{dy}{dx} = x = \sqrt{2y}$ und schreiben

$$\int_{-\infty}^\infty e^{-\frac{1}{2}x^2} dx = 2\int_0^\infty e^{-\frac{1}{2}x^2} dx = \sqrt{2}\int_0^\infty e^{-y} y^{-\frac{1}{2}} dy = \sqrt{2}\Gamma(1/2) = \sqrt{2\pi}.$$

□

[16] In dieser Rechnung kommt ein Doppelintegral vor. Rechenregeln für solche Mehrfachintegrale werden wir eigentlich erst in Analysis III kennen lernen. In dieser Rechnung benötigen wir jedoch nur, dass

$$\int_a^b f(x)dx \cdot \int_c^d g(y)dy = \int_a^b \Big(\int_c^d f(x)g(y)dx\Big)dy =: \int_a^b \int_c^d f(x)g(y)dxdy$$

für geeignete Funktionen f und g gilt.

9.2 Die Laplace-Methode

Angenommen, man will abschätzen, wie groß in etwa $\int_a^b g_n(x)dx$ für große n ist, dann liefert die Laplace-Methode in manchen Fällen ein sinnvolles Werkzeug, nämlich genau dann, wenn $g_n(x) = e^{nf(x)}$ gilt und f ein eindeutiges Maximum auf $[a,b]$ hat. Dies hat nämlich zur Folge, dass $g_n(x)$ um dieses Maximum herum viel größer ist als an anderen Stellen.

Theorem 9.4 (Laplace-Methode). *Sei $I = [a,b]$, $f \in \mathcal{C}^2(I)$ mit $f(x_0) = \max_{x\in I} f(x)$ für genau ein x_0 mit $f''(x_0) < 0$ (d.h. f hat in x_0 das einzige Maximum auf I). Dann gilt*

$$\lim_{n\to\infty} \frac{\displaystyle\int_a^b e^{nf(x)}dx}{e^{nf(x_0)}\sqrt{\dfrac{2\pi}{-nf''(x_0)}}} = 1. \tag{9.1}$$

Bemerkung 9.5 (Interpretation). Obwohl obiges Theorem als Grenzwertresultat formuliert ist, besteht die Anwendung darin, $\int_a^b e^{nf(x)}dx$ für große n approximativ zu berechnen. Angenommen, $f(x) = -x^2$ und $[a,b] = [-1,1]$. Dann besagt Theorem 9.4, dass

$$\int_{-1}^1 e^{-nx^2}dx \approx \sqrt{\frac{\pi}{n}}$$

gilt.

Beweis. Zunächst ist

$$f(x) = f(x_0) + \tfrac{1}{2}f''(y)(x-x_0)^2 \tag{9.2}$$

für ein y zwischen x_0 und x.
Um dies zu beweisen schreiben wir mittels partieller Integration

$$\int_{x_0}^x (x-z)f''(z)dz = (x-z)f'(z)_{z=x_0}^{z=x} + \int_{x_0}^x f'(z)dz = f(x) - f(x_0).$$

Nach Proposition 8.35 gibt es nun ein y zwischen x_0 und x mit

$$f(x) - f(x_0) = f''(y)\int_{x_0}^x (x-z)dz = \tfrac{1}{2}f''(z)(x-x_0)^2,$$

und damit ist (9.2) gezeigt.

Wir zeigen zunächst, dass für jedes $\delta > 0$ (mit $(x_0-\delta, x_0+\delta) \subseteq (a,b)$)

$$\int_a^b e^{nf(x)}dx \overset{n\to\infty}{\approx} \int_{x_0-\delta}^{x_0+\delta} e^{nf(x)}dx.$$

Hierzu sei $\varepsilon > 0$ und $\delta' > 0$ so klein, dass $f(x) + \varepsilon < \inf_{y\in(x_0-\delta',x_0+\delta')} f(y)$ für $x \notin (x_0 - \delta, x_0+\delta)$. Damit gilt nämlich

$$\begin{aligned}
\frac{\int_a^{x_0-\delta} + \int_{x_0+\delta}^b e^{nf(x)}dx}{\int_{x_0-\delta}^{x_0+\delta} e^{nf(x)}dx} &\le \frac{\int_a^{x_0-\delta} + \int_{x_0+\delta}^b e^{nf(x)}dx}{\int_{x_0-\delta'}^{x_0+\delta'} e^{n\inf_{y\in(x_0-\delta',x_0+\delta')} f(y)}dx} \\
&\le \frac{1}{2\delta'}\int_a^{x_0-\delta} + \int_{x_0+\delta}^b e^{n(f(x)-\inf_{y\in(x_0-\delta',x_0+\delta')} f(y))}dx \\
&\le \frac{b-a}{2\delta'}e^{-n\varepsilon} \xrightarrow{n\to\infty} 0.
\end{aligned}$$

Sei nun $\varepsilon > 0$. Dann gibt es wegen der Stetigkeit von f'' ein $\delta > 0$ so dass $f(x) \leq f(x_0) + \frac{1}{2}(f''(x_0) - \varepsilon)(x - x_0^2)$ für x mit $|x - x_0| < \delta$. Daraus folgt

$$\begin{aligned}\int_a^b e^{nf(x)} dx &\overset{n\to\infty}{\approx} e^{nf(x_0)} \int_{x_0-\delta}^{x_0+\delta} e^{nf(x)} dx \leq \int_{x_0-\delta}^{x_0+\delta} e^{\frac{n}{2}(f''(x_0)-\varepsilon)(x-x_0)^2} dx \\ &= e^{nf(x_0)} \frac{1}{\sqrt{-n(f''(x_0)+\varepsilon)}} \int_{-\delta\sqrt{-n(f''(x_0)+\varepsilon)}}^{\delta\sqrt{-n(f''(x_0)+\varepsilon)}} e^{-\frac{1}{2}y^2} dy \\ &\overset{n\to\infty}{\approx} e^{nf(x_0)} \sqrt{\frac{2\pi}{-n(f''(x_0)+\varepsilon)}}\end{aligned}$$

wobei das letzte $\approx$ aus Proposition 9.3 folgt. Da $\varepsilon > 0$ beliebig war, folgt damit '$\leq$' in (9.1). Ersetzt man in der letzten Rechnung '$-\varepsilon$' und '$\leq$' durch '$+\varepsilon$' und '$\geq$', folgt auch '$\geq$', insgesamt also '$=$'. □

Proposition 9.6 (Stirling's Formel). *Es gilt*

$$\lim_{n\to\infty} \frac{n!}{(n/e)^n\sqrt{2\pi n}} = 1.$$

Beweis. Zunächst betrachten wir $f : x \mapsto \log(x) - x$. Diese Funktion erfüllt die Voraussetzung von Theorem 9.4 und es gilt

$$f'(x) = \frac{1}{x} - 1, \qquad f''(x) = -\frac{1}{x^2}.$$

Damit hat f genau ein Maximum, nämlich bei $x = 1$ mit $f(1) = -1$. Nun verwenden wir die Laplace-Methode und schreiben

$$\begin{aligned}n! &= \int_0^\infty e^{-x} x^n dx \overset{z:=x/n}{=} n^{n+1} \int_0^\infty e^{-nz} z^n dz = n^{n+1} \int_0^\infty e^{-n(z-\log(z))} dz \\ &\overset{n\to\infty}{\approx} n^{n+1} e^{-n} \sqrt{\frac{2\pi}{n}} = (n/e)^n \sqrt{2\pi n}.\end{aligned}$$

□

9.3 Die Partialbruchzerlegung

Rationale Funktionen kamen in der Vorlesung schon häufig vor. Die Partialbruchzerlegung stellt eine vereinheitlichte Darstellung solcher Funktionen das, die vor allem für die Integration von rationalen Funktionen sehr nützlich ist.

Definition 9.7 (Rationale Funktion). *Seien $p, q : \mathbb{C} \to \mathbb{C}$ normierte Polynome, d.h. von der Gestalt*

$$p(z) = \sum_{i=0}^m a_i z^i, \qquad q(z) = \sum_{i=0}^n b_i z^i \tag{9.3}$$

mit Koeffizienten $a_i, b_j \in \mathbb{C}, i = 1, ..., m, j = 1, ..., n$ und $a_m, b_n \neq 0$, also ist m der Grad von p und n der Grad von q. (Gilt $a_m = 1$, so heißt p normiertes Polynom.*) Dann heißt die Funktion $z \mapsto q(z)/p(z)$* (gebrochen) rationale Funktion. *Ist $n < m$, so heißt q/p echt (gebrochen) rationale Funktion.*

Bemerkung 9.8 (Hauptsatz der Algebra). Nach dem Hauptsatz der Algebra ist der Körper $\mathbb{C}$ algebraisch abgeschlossen, d.h. jedes normierte Polynom p wie in (9.3) hat die Darstellung

$$p(z) = \prod_{i=1}^{k} (z - z_i)^{c_i}, \tag{9.4}$$

mit $\sum c_i = m$ und $z_i \neq z_j, i, j = 1, ..., k$. Man sagt auch, das Polynom *zerfällt* in Linearfaktoren. In dieser Darstellung sind $z_1, ..., z_k$ genau die Nullstellen von p und c_i ist die Vielfachheit der Nullstelle z_i. (Dies ist die größte Zahl c, für die $z \mapsto p(x)/(z - z_i)^c$ in z_i stetig ist.)

Angenommen, die Koeffizienten in (9.3) sind alle reell. Gilt dann $p(z) = 0$, so gilt auch $p(\bar{z}) = \overline{p(z)} = \bar{0} = 0$. Insbesondere kann man daraus folgern, dass in der Darstellung 9.4 jede Nullstelle $z = x + iy$ von p der Vielfachheit c eine Nullstelle $\bar{z} = x - iy$ von p derselben Vielfachheit nach sich zieht. Multipliziert man die entsprechenden Terme, so erhält man $(z - z_i)(z - \bar{z}_i) = (z - x_i - iy_i)(z - x_i + iy_i) = (z - x_i)^2 + y_i^2$, also ein Polynom zweiten Grades. Daraus sieht man sofort:
Ist p ein normiertes Polynom mit reellwertigen Koeffizienten, so gibt es $c_i, f_j \in \mathbb{N}$ und $z_i, d_j, e_j \in \mathbb{R}$ mit

$$p(z) = \prod_{i=1}^{k} (z - z_i)^{c_i} \prod_{j=1}^{l} (z^2 + d_j z + e_j)^{f_j}. \tag{9.5}$$

Bemerkung 9.9 (Rationale und echt rationale Funktionen). Sei p/q eine rationale Funktion. Durch Polynomdivision mit Rest kann man diese Funktion als $p/q = r + p'/q$ schreiben, wobei p'/q echt rational ist.

Theorem 9.10 (Partialbruchzerlegung im Komplexen). *Sei p ein Polynom der Form 9.4 (mit komplexen Koeffizienten) und q/p eine echt rationale Funktion. Dann gibt es $\alpha_{ij} \in \mathbb{C}, i = 1, ..., k; j = 1, ..., c_i$ so dass*

$$\frac{q(z)}{p(z)} = \sum_{i=1}^{k} \sum_{j=1}^{c_i} \frac{\alpha_{ij}}{(z - z_i)^j}.$$

Beweis. Der Beweis läuft nach Induktion über $m = c_1 + \cdots + c_k$, dem Grad von p. Für $m = 1$ ist nichts zu zeigen, weil nach Voraussetzung $n < m$, also $n = 0$ oder $n = 1$ gilt. Angenommen, die Behauptung gilt nun schon für alle p mit Grad höchstens m. Ist nun p ein Polynom vom Grad m, so ist $p(z) = (z - z_1)^{c_1} p'(z)$ für ein Polynom p_1 mit $p'(z_1) \neq 0$. Dann gilt für alle $\alpha \in \mathbb{C}$

$$\frac{q(z)}{p(z)} - \frac{\alpha}{(z - z_1)^{c_1}} = \frac{q(z) - \alpha p'(z)}{p(z)}.$$

Ist nun $\alpha := q(z_1)/p'(z_1)$, so ist der Zähler ein Polynom das in z_1 eine Nullstelle besitzt, also gilt $q(z) - \alpha p'(z) = (z - z_1) r(z)$ für ein Polynom z. Nun gilt also

$$\frac{q(z)}{p(z)} = \frac{\alpha}{(z - z_1)^{c_1}} + \frac{(z - z_1) r(z)}{(z - z_1)^{c_1} p'(z)} = \frac{\alpha}{(z - z_1)^{c_1}} + \frac{r(z)}{(z - z_1)^{c_1 - 1} p'(z)}.$$

Der letzte Term ist nun eine rationale Funktion, dessen Nenner ein Polynom vom Grad $m - 1$ ist, auf die man also die Induktionsvoraussetzung anwenden kann. □

Korollar 9.11 (Partialbruchzerlegung im Reellen). *Sei p ein Polynom der Form 9.5 mit reellen Koeffizienten und q/p eine echt rationale Funktion. Dann gibt es $\alpha_{ij} \in \mathbb{R}, i = 1, ..., k; j = 1, ..., c_i$, $\beta_{ij}, \gamma_{ij} \in \mathbb{R}, i = 1, ..., l; j = 1, ..., f_i$ so dass*

$$\frac{q(z)}{p(z)} = \sum_{i=1}^{k} \sum_{j=1}^{c_i} \frac{\alpha_{ij}}{(z - z_i)^j} + \sum_{i=1}^{l} \sum_{j=1}^{f_i} \frac{\beta_{ij} z + \gamma_{ij}}{(z^2 + d_i z + e_i)^j}.$$

Beweis. Die Aussage folgt direkt aus Theorem 9.10 und Bemerkung 9.8. □

Wir haben nun gezeigt, dass sich rationale Funktionen immer als Summe von zwei verschiedenen Typen von Funktionen, nämlich solchen der Bauart $x \mapsto \frac{1}{(z-e)^j}$ und $x \mapsto \frac{\beta z + \gamma}{(z^2+dz+e)^j}$ darstellen lassen. Um rationale Funktionen integrieren zu können, benötigt man also nur noch Stammfunktionen solcher Funktionen. Diese sind nun zusammen gefasst.

Proposition 9.12 (Integration rationaler Funktionen). *Es gilt*

$$\int \frac{1}{z-e} dz = \log |z - e|,$$

$$\int \frac{1}{(z-e)^j} dz = -\frac{1}{j-1} \cdot \frac{1}{z-e)^{j-1}},$$

$$\int \frac{1}{z^2 + cz + e} dz = \frac{2}{\sqrt{4e - c^2}} \cdot \arctan \frac{2z + c}{4e - c^2},$$

$$\int \frac{1}{(z^2 + cz + e)^j} dz = \frac{2z + c}{(j-1)(4e - c^2)(z^2 + cz + e)^{j-1}} + \frac{2(2j-3)}{(j-1)(4e-c^2)} \int \frac{1}{(z^2 + cz + e)^{j-1}} dz,$$

$$\int \frac{\beta z + \gamma}{z^2 + cz + e} = \frac{\beta}{2} \log(z^2 + cz + e) + \left(\gamma - \frac{\beta c}{2}\right) \int \frac{1}{(z^2 + cz + e)^j} dz,$$

$$\int \frac{\beta z + \gamma}{(z^2 + cz + e)^j} dz = -\frac{\beta}{2(j-1)(z^2 + cz + e)^{j-1}} + \left(\gamma - \frac{\beta c}{2}\right) \int \frac{1}{(z^2 + cz + e)^j} dz$$

Beweis. Alle Behauptungen zeigt man am besten durch Ableiten der rechten Seiten. □

9.4 Unendliche Produkte und der Wert von $\zeta(2)$

In Bemerkung 4.9 haben wir bereits bemerkt, dass

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

Hier wollen wir dies nun beweisen. Entscheidend ist hierbei eine verblüffende Darstellung der sinus-Funktion, siehe Lemma 9.17. Hierbei tritt eine unendliches Produkt auf. Dies ist ähnlich wie eine unendliche Reihe. Wir klären zunächst, was die Konvergenz von unendlichen Produkten bedeutet.

Definition 9.13 (Konvergenz unendlicher Produkte). *Sei $(x_n)_{n=1,2,...}$ eine Folge reeller Zahlen. Gilt*

$$\lim_{N \to \infty} \prod_{n=1}^{N} x_n = x$$

mit $x \neq 0$, so sagen wir, dass $\prod_{n=1}^{\infty} x_n$ konvergiert.

Bemerkung 9.14. Sei $\prod_{n=1}^{\infty} x_n$ ein konvergentes unendliches Produkt. Dann ist klar, dass $x_n \xrightarrow{n\to\infty} 1$.

Konvergenzkriterien unendlicher Produkte lassen sich oft auf Konvergenzkriterien für unendliche Summen übersetzen. Etwa gilt folgendes Resultat.

Theorem 9.15 (Konvergenz unendlicher Produkte). *Sei $(x_n)_{n=1,2,...}$ ein Folge nichtneagtiver reeller Zahlen. Dann sind äquivalent:*

1. *$\sum_{n=1}^{\infty} x_n$ konvergiert,*
2. *$\prod_{n=1}^{\infty}(1 + x_n)$ konvergiert,*
3. *$\prod_{n=1}^{\infty}(1 - x_n)$ konvergiert.*

Beweis. Klar ist, dass in jedem der drei Fälle $x_n \downarrow 0$ gelten muss. Weiter ist (wegen der Stetigkeit von log) klar, dass 2. genau dann gilt, wenn $\log \prod_{n=1}^{\infty}(1 + x_n) = \sum_{n=1}^{\infty} \log(1 + x_n)$ existiert (und analog für 3.).
1.⇒2.: Das folgt aus $\log(1 + x) \leq x$.
2.⇒3.: Da $\frac{\log(1+x_n)}{|\log(1-x_n)|} \xrightarrow{n\to\infty} 1$ nach Lemma 6.10.3, gilt $|\log(1 - x_n)| \leq 2 \log(1 + x_n)$ für fast alle n. Damit folgt die Behauptung aus dem Majorantenkriterium, Theorem 22.26.
3.⇒1.: Das folgt aus $x \leq |\log(1 - x)|$. □

Das nächste Resultat wird in der Vorlesung Analysis III wesentlich erweitert werden.

Proposition 9.16 (Majorisierte Konvergenz). *Sei $(y_n)_{n=1,2,...}, (z_n)_{n=1,2,...}$ reellwertige Folgen und $(x_{nk})_{n,k=1,2,...}$ eine unendliche Matrix reeller Zahlen mit $x_{nk} \xrightarrow{k\to\infty} y_n$, sowie $|x_{nk}| \leq z_n$ (und damit $|y_n| \leq z_n$) und $\sum_{n=1}^{\infty} z_n < \infty$. Dann gilt*

$$\lim_{k\to\infty} \sum_{n=1}^{\infty} x_{nk} = \sum_{n=1}^{\infty} y_n,$$
$$\lim_{k\to\infty} \prod_{n=1}^{\infty}(1 + x_{nk}) = \prod_{n=1}^{\infty}(1 + y_n),$$

wobei alle Limiten existieren.

Beweis. 1. Klar ist nach Theorem 22.26, dass $\sum_{n=1}^{\infty} y_n$ existiert. Sei $\varepsilon > 0$ und N groß genug, so dass $\sum_{n=N}^{\infty} z_n < \varepsilon$. Weiter sei K groß genug, so dass $|x_{nk} - y_n| \leq \varepsilon 2^{-n}$ für $k = 1, ..., K, n = 1, ..., N$. Dann gilt für $k > K$

$$\begin{aligned} \Big|\sum_{n=1}^{\infty} x_{nk} - \sum_{n=1}^{\infty} y_n\Big| &\leq \Big|\sum_{n=1}^{N} x_{nk} - \sum_{n=1}^{N} y_n\Big| + \sum_{n=N}^{\infty} |x_{nk}| + \sum_{n=N}^{\infty} |y_n| \\ &\leq \sum_{n=1}^{N} \varepsilon 2^{-n} + 2 \sum_{n=N}^{\infty} z_n < 3\varepsilon. \end{aligned}$$

Da $\varepsilon > 0$ beliebig war, folgt die Behauptung.
2. Sei zunächst N groß genug, so dass $z_n < 1/2$ für $n > N$. Dann gilt $|\log(1 + x_{nk})| \leq 2z_n$ wegen $x_{nk} \leq z_n < 1/2$ für $n > N$. Nun folgt die Behauptung, wenn man 1. auf $(\log(1 + x_{nk}))_{n,k=N,N+1,...}$ anwendet. □

Lemma 9.17 (Produktdarstellung des sinus). *Es gilt*

$$\frac{\sin(\pi x)}{\pi x} = \prod_{n=1}^{\infty}\Big(1 - \frac{x^2}{n^2}\Big).$$

Beweis. Wir setzen

$$p_n(x) := \frac{1}{2i}\Big(\Big(1+\frac{\pi i x}{n}\Big)^n - \Big(1-\frac{\pi i x}{n}\Big)^n\Big) \xrightarrow{n\to\infty} \frac{1}{2i}(e^{\pi i x} - e^{-\pi i x}) = \sin(\pi x).$$

Weiter ist $p_n(x) = 0$ genau dann, wenn für ein $j \in \mathbb{Z}$

$$\Big(\frac{n+\pi i x}{n-\pi i x}\Big)^n = 1 = e^{(2\pi i)j}.$$

Da

$$\tan(x) = \frac{\sin(x)}{\cos(x)} = \frac{1}{i}\frac{e^{ix}-e^{-ix}}{e^{ix}+e^{-ix}} = \frac{1}{i}\frac{e^{2ix}-1}{e^{2ix}+1}$$

ist dies genau dann der Fall, wenn

$$x = \frac{n}{\pi i}\frac{e^{2j\pi i/n}-1}{e^{2j\pi i/n}+1} = \frac{n}{\pi}\tan\frac{j\pi}{n}.$$

Sei nun n gerade. Dann ist p_n ein Polynom vom Grad $n-1$ und, da tan auf $(-\pi/2, \pi/2)$ bijektiv ist, hat p_n genau die Nullstellen

$$x_j = \frac{n}{\pi}\tan\frac{j\pi}{n}, j = -\tfrac{n}{2}+1, ..., \tfrac{n}{2}-1.$$

Damit schreiben wir für gerade n, da der Koeffizient von x in p_n gerade π ist,

$$p_n(x) = \pi x \prod_{j=1}^{\frac{n}{2}-1}\Big(1 - \frac{z}{\frac{n}{\pi}\tan\frac{j\pi}{n}}\Big)\Big(1 + \frac{z}{\frac{n}{\pi}\tan\frac{j\pi}{n}}\Big) = \prod_{j=1}^{\frac{n}{2}-1}\Big(1 - \frac{z^2\pi^2}{n^2\tan^2\frac{j\pi}{n}}\Big)$$

Da $\tan(x) \geq x$ im Bereich $(0, \pi/2)$ monoton ist, gilt für festes $j = 1, 2, ...$

$$\frac{n^2}{\pi^2}\tan^2\frac{j\pi}{n} \geq j^2$$

sowie

$$\frac{n^2}{\pi^2}\tan^2\frac{j\pi}{n} \overset{n\to\infty}{\approx} \frac{n^2}{\pi^2}\sin^2\frac{j\pi}{n} \approx_{n\to\infty} j^2,$$

folgt aus Proposition 9.16, dass

$$\sin(\pi x) = \lim_{n\to\infty} p_n(x) = \lim_{n\to\infty}\prod_{j=1}^{\frac{n}{2}-1}\Big(1 - \frac{z^2\pi^2}{n^2\tan^2\frac{j\pi}{n}}\Big) = \pi x\prod_{j=1}^{\infty}\Big(1 - \frac{x^2}{j^2}\Big)$$

und damit die Behauptung. □

Proposition 9.18 (Der Wert von $\zeta(2)$). *Es gilt*

$$\sum_{n=1}^{\infty}\frac{1}{n^2} = \frac{\pi^2}{6}.$$

Beweis. Wir zeigen nun

$$\frac{1}{x^2}\Big(\frac{\sin(\pi x)}{\pi x} - 1\Big) \xrightarrow{x\to 0} -\frac{\pi^2}{6}$$

sowie

$$\frac{1}{x^2}\Big(\frac{\sin(\pi x)}{\pi x} - 1\Big) \xrightarrow{x\to 0} -\sum_{n=1}^{\infty}\frac{1}{n^2}$$

woraus die Behauptung aus der Eindeutigkeit des Grenzwertes folgt. Der erste Grenzwert ist einfach, es gilt nämlich

$$\frac{1}{x^2}\Big(\frac{\sin(\pi x)}{\pi x} - 1\Big) = \sum_{k=1}^{\infty}(-1)^k\frac{\pi^{2k}x^{2(k-1)}}{(2k+1)!} \xrightarrow{x\to 0} -\frac{\pi^2}{6}.$$

Aus Lemma 9.17 folgtfür $x \to 0$

$$\frac{1}{x^2}\Big(\frac{\sin(\pi x)}{\pi x} - 1\Big) = -\sum_{n=1}^{\infty}\frac{1}{n^2} + \mathcal{O}(x^2) \xrightarrow{x\to 0} -\sum_{n=1}^{\infty}\frac{1}{n^2}.$$

□

Teil III
Approximation von Funktionen

Gegeben sei eine Funktion $f : I \to \mathbb{R}$. Wir sagen, dass die Folge $(f_n)_{n=1,2,...}$ die Funktion f approximiert, falls $f_n \xrightarrow{n\to\infty} f$. Die Konvergenzart die wir betrachten wollen ist die lokal gleichmäßige (was dasselbe ist wie die gleichmäßige Konvergenz auf Kompakta; siehe Lemma 10.2). Wir werden in diesem Teil für die auf einer Menge $K \subseteq I$ gleichmäßige Konvergenz kurz $||f_n - f||_K \xrightarrow{n\to\infty} 0$ schreiben.

10 Lokal gleichmäßige Konvergenz von Funktionen

In Definition 8.3 haben wir bereits von der gleichmäßigen Konvergenz auf Kompakta gehört. Wir werden die hier eingeführte lokal gleichmäßige Konvergenz als dazu äquivalent zeigen. Ziel des Abschnittes ist es einzusehen, durch welche Funktionsklassen (etwa Polynome) beliebige stetige Funktionen approximiert werden können. Beispielsweise haben wir ja bereits gesehen, dass die Exponentialfunktion $x \mapsto \exp(x)$ (zumindest im Bereich $|x| < 1$) durch Polynome approximierbar ist, und zwar gleichmäßig auf Kompakta. Es gilt nämlich nach (6.2), dass

$$\sup_{|x|\leq 1} \Big|\Big| \sum_{k=0}^{n} \frac{x^k}{k!} - \exp(x) \Big|\Big| \xrightarrow{n\to\infty} 0.$$

Wir werden in Abschnitt 10.2 durch den Satz von Weierstrass sehen, dass eine solche Konvergenz auf Kompakta für jede stetige Funktion durch Polynome erreicht werden kann. Eine Verallgemeinerung stellt der Satz von Stone in Abschnitt 10.3 dar, in dem die Menge der Polynome durch eine punktetrennende Algebra ersetzt wird.

10.1 Grundlegendes

In Definition 8.3 haben wir bereits von der gleichmäßigen Konvergenz auf Kompakta gehört. Die lokal gleichmäßige Konvergenz der nächsten Definition sieht zunächst etwas anders aus. Es stellt sich jedoch heraus, dass eine Funktionenfolge genau dann lokal gleichmäßig konvergiert wenn sie gleichmäßig auf Kompakta konvergiert.

Definition 10.1 (Lokal gleichmäßige Konvergenz). *Sei (E, r) ein metrischer Raum und $f, f_1, f_2, ... : E \to \mathbb{R}$. Gibt es für jedes $x \in E$ eine Umgebung U_x so dass $||f_n - f||_{U_x} \xrightarrow{n\to\infty} 0$, so konvergiert $(f_n)_{n=1,2,...}$* lokal gleichmäßig *gegen f.*

Lemma 10.2 (Lokal gleichmäßige und gleichmäßige Konvergenz auf Kompakta). *Sei $I \subseteq \mathbb{R}$ ein Intervall und $f, f_1, f_2, ... : I \to \mathbb{R}$. Dann sind äquivalent:*

1. *$(f_n)_{n=1,2,...}$ konvergiert lokal gleichmäßig gegen f.*
2. *$f_n \xrightarrow{n\to\infty}_{glmK} f$.*

Beweis. 1.⇒2.: Sei $\varepsilon > 0$ und $K \subseteq I$ kompakt. Für $x \in K$ sei U_x eine Umgebung von x mit $||f_n - f||_{U_x} \xrightarrow{n\to\infty} 0$ und $\delta_x > 0$, so dass $A_x := B_{\delta_x}(x) \subseteq U_x$. Damit ist auch $||f_n - f||_{B_{\delta_x}(x)} \xrightarrow{n\to\infty} 0$ und $K \subseteq \bigcup_{x\in K} A_x$. Wegen der Heine-Borel'schen Überdeckungseigenschaft (siehe Theorem 5.24) gibt es $x_1, ..., x_n \in K$ mit $K \subseteq \bigcup_{i=1}^n A_{x_i}$. Setzt man $\delta := \min_{i=1,...,n} \delta_{x_i} > 0$, so folgt, dass

$$\sup_{x \in K} |f_n(x) - f(x)| \leq \sup_{i=1,...,n} \sup_{y \in A_{x_i}} |f_n(y) - f(y)| \xrightarrow{n \to \infty} 0.$$

Das ist aber gerade die gleichmäßige Konvergenz auf Kompakta.
2.⇒1.: Sei $x \in I$. Dann ist der abgeschlossene Ball $B_\varepsilon(x)$ kompakt und eine Umgebung von x. Wegen der gleichmäßigen Konvergenz auf Kompakta folgt also die lokal gleichmäßige Konvergenz. □

Ziel ist es, eine Funktion f (die beliebig, jedoch stetig sein kann) durch einfachere Funktionen zu approximieren. Eine Mindestvoraussetzung an die Menge der Funktionen, die wir zur Approximation heranziehen, ist, dass sie eine Algebra bilden.

Definition 10.3 (Funktionenalgebra). *Sei (E, r) ein metrischer Raum und $\mathcal{M} \subseteq \mathcal{C}(E)$ eine Teilmenge der stetigen Funktionen.*

1. *Dann heißt $\mathcal{M}$* Algebra *(von Funktionen), wenn gilt: (i) $f, g \in \mathcal{M} \Rightarrow f + g \in \mathcal{M}$, (ii) $f \in \mathcal{M}, \lambda \in \mathbb{R} \Rightarrow \lambda f \in \mathcal{M}$, (iii) $f, g \in \mathcal{M} \Rightarrow fg \in \mathcal{M}$. (Mit anderen Worten ist $\mathcal{M}$ ein Vektorraum mit der zusätzlichen Eigenschaft (iii).)*
2. *Eine Algebra $\mathcal{M} \subseteq \mathcal{C}(E)$ heißt* punktetrennend*, wenn es für alle $x, y \in E$ ein $f \in \mathcal{M}$ gibt mit $f(x) \neq f(y)$.*

Beispiel 10.4 (Polynome und trigonometrische Funktionen). Es gibt für unsere weitere Anwendungen zwei wichtige Beispiele für punktetrennende Algebren.

1. Sei
$$\mathcal{M} := \{x \mapsto \sum_{k=0}^n a_k x^k, n \in \mathbb{N}, a_0, ..., a_n \in \mathbb{R}\}$$
die Menge der Polynome. Dies ist eine Algebra, da das Produkt zweier Polynome offenbar wieder ein Polynom ist. Weiter ist $\mathcal{M}$ punktetrennend.
2. Sei
$$\mathcal{M} := \{x \mapsto \sum_{k=0}^n a_k \cos(kx) + b_k \sin(kx), n \in \mathbb{N}, a_0, ..., a_n, b_0, ..., b_n \in \mathbb{R}\}$$
die Menge der Linearkombinationen von trigonometrischen Funktionen. Dies ist ebenfalls eine Algebra, wie man durch Anwendung von Lemma 6.17 sieht. Außerdem ist $\mathcal{M}$ punktetrennend im Raum der 2π-periodischen Funktionen. (Man beachte, dass $\mathcal{M}$ nicht punktetrennend auf dem Raum $\mathcal{C}(\mathbb{R})$ ist, da für jede trigonometrische Funktion Periode 2π besitzt.)

Ziel des später zu zeigenden Satzes von Stone (Theorem 10.12) ist es, beliebige stetige Funktionen durch Funktionen in einer Algebra zu approximieren. Hierzu ist es praktisch, den Abschluss der Algebra zu definieren.

Definition 10.5 (Abschluss von $\mathcal{M} \subseteq \mathcal{C}(E)$). *Sei $\mathcal{M} \subseteq \mathcal{C}(E)$. Dann setzen wir*

$$\overline{\mathcal{M}} := \{f \in \mathcal{C}(E) : \text{ es gibt } p_1, p_2, ... \in \mathcal{M} \text{ mit } p_n \xrightarrow{n \to \infty}_{glmK} f\}$$

und bezeichnen dies mit dem Abschluss von $\mathcal{M}$ (unter der gleichmäßigen Konvergenz auf Kompakta).

Lemma 10.6 (Abschluss einer Algebra ist eine Algebra). *Sei $\mathcal{M} \subseteq \mathcal{C}(E)$ eine Algebra. Dann gilt:*

1. $\overline{\mathcal{M}}$ *ist eine Algebra.*
2. $\overline{\overline{\mathcal{M}}} = \overline{\mathcal{M}}$.

Beweis. 1. Es ist nur zu zeigen, dass $f, g \in \overline{\mathcal{M}} \Rightarrow fg \in \overline{\mathcal{M}}$. Sei $K \subseteq E$ kompakt, $0 < \varepsilon < 1$ und $p, q \in \mathcal{M}$ so, dass $||f - p||_K < \varepsilon/||g||_K$, $||g - q||_K < \varepsilon/(||f||_K + 1)$. Dann gilt $||p||_K \leq ||f||_K + \varepsilon < ||f||_K + 1$ und somit

$$\begin{aligned} ||fg - pq||_K &= ||(f-p)g + p(g-q)||_K \leq ||f-p||_K \cdot ||g||_K + (||f||_K + 1) \cdot ||g-q)||_K \\ &\leq 2\varepsilon. \end{aligned}$$

Hieraus folgt die Behauptung.
2. Klar ist, dass $\overline{\mathcal{M}} \subseteq \overline{\overline{\mathcal{M}}}$. Sei $f \in \overline{\overline{\mathcal{M}}}$ und sei $\varepsilon > 0$. Dann gibt es $g \in \overline{\mathcal{M}}$ mit $||f - g|| \leq \varepsilon$. Außerdem gibt es $p \in \mathcal{M}$ mit $||g - p|| \leq \varepsilon$. Daraus folgt $||f - p|| \leq ||f - g|| + ||g - p|| < 2\varepsilon$ und die Behauptung folgt. □

10.2 Der Satz von Weierstrass

Wir beschäftigen uns nun mit der Approximation von stetigen Funktionen durch Polynome. Dass dies gleichmäßig auf Kompakta immer möglich ist, ist der Inhalt des Satzes von Weierstrass. Die verwendeten Polynome basieren auf den Bernstein-Polynomen, die wir zunächst einführen.

Definition 10.7 (Bernstein-Polynom). *Für $k \leq n$ heißt*

$$r_{k,n}(x) := \binom{n}{k} x^k (1-x)^{n-k}$$

Bernstein-Polynom*.*

Nach der binomischen Formel ist $\sum_{k=0}^n r_{k,n}(x) = 1$. Wir benötigen im Beweis des Satzes von Weierstrass eine weitere Aussage über Bernstein-Polynome, die wir nun formulieren.

Lemma 10.8. *Es gilt*

$$\sum_{k=0}^n (k - nx)^2 r_{k,n}(x) = nx(1-x).$$

Beweis. Wir schreiben

$$\begin{aligned} \sum_{k=0}^n k r_{k,n}(x) &= n \sum_{k=1}^n \binom{n-1}{k-1} x^k (1-x)^{n-k} \\ &= nx \sum_{k=0}^{n-1} \binom{n-1}{k} x^k (1-x)^{n-1-k} = nx, \end{aligned}$$

$$\begin{aligned} \sum_{k=0}^n k(k-1) r_{k,n}(x) &= n(n-1) \sum_{k=2}^n \binom{n-2}{k-2} x^k (1-x)^{n-k} \\ &= n(n-1)x^2 \sum_{k=0}^{n-2} \binom{n-1}{k} x^k (1-x)^{n-2-k} = n(n-1)x^2. \end{aligned}$$

Daraus folgt

$$\begin{aligned}\sum_{k=0}^{n}(k-nx)^2 r_{k,n}(x) &= \sum_{k=0}^{n}\big(k(k-1)-(2nx-1)k+n^2x^2\big)r_{k,n}(x)\\ &= n(n-1)x^2-(2nx-1)nx+n^2x^2 = -nx^2+nx = nx(1-x).\end{aligned}$$

□

Theorem 10.9 (Satz von Weierstrass). *Sei $f \in \mathcal{C}(\mathbb{R})$. Für jede kompakte Menge K und $\varepsilon > 0$ gibt es ein Polynom p mit $||f-p||_K < \varepsilon$.*

Bemerkung 10.10 (Alternative Formulierungen). Folgende Aussagen sind offenbar äquivalent zur obigen Formulierung:

1. Sei $f \in \mathcal{C}(\mathbb{R})$. Dann gibt es Polynome $p_1, p_2, \dots$ mit $p_n \xrightarrow{n\to\infty}_{glmK} f$.
2. Sei $\mathcal{M} \subseteq \mathcal{C}(\mathbb{R})$ die Menge der Polynome. Dann gilt $\overline{\mathcal{M}} = \mathcal{C}(\mathbb{R})$.

Beweis von Theorem 10.9. Es genügt die Behauptung für $K = [0,1]$ und $||f||_K = 1$ zu zeigen. Wir betrachten das auf den Bernstein Polynomen zusammengesetzte Polynom

$$p_n(x) = \sum_{k=0}^{n} f(k/n) r_{k,n} = \sum_{k=0}^{n} f(k/n)\binom{n}{k}x^k(1-x)^{n-k}.$$

Sei $\varepsilon > 0$. Da f auf K gleichmäßig stetig ist, gibt es $\delta > 0$, so dass $|f(x)-f(y)| < \varepsilon$ für $|x-y| < \delta$. Es gilt wegen der binomischen Formel

$$p_n(x) - f(x) = \sum_{k=0}^{n}(f(k/n)-f(x))r_{k,n}(x).$$

Wir setzen $L_x := \{k : |x-k/n| < \delta\}$. Dann ist mit Lemma 10.8

$$\begin{aligned}|p_n(x)-f(x)| &\le \sum_{k\in L_x}|f(k/n)-f(x)|r_{k,n}(x) + \sum_{k\in L_x^c}|f(k/n)-f(x)|r_{k,n}(x)\\ &\le \varepsilon\sum_{k=0}^{n} r_{k,n}(x) + 2||f||_k \sum_{k\in L_x^c} r_{k,n}(x)\\ &\le \varepsilon + 2\sum_{k\in L_x^c}\Big(\frac{x-k/n}{\delta}\Big)^2 r_{k,n}(x)\\ &\le \varepsilon + \frac{2}{n^2\delta^2}\sum_{k=0}^{n}(k-nx)^2 r_{k,n}(x)\\ &= \varepsilon + \frac{2x(1-x)}{n\delta^2} \le \varepsilon + \frac{1}{2n\delta^2}.\end{aligned}$$

Ist also $n > \frac{1}{2\varepsilon\delta^2}$, so gilt

$$|p_n(x) - f(x)| \le 2\varepsilon,$$

woraus die Behauptung folgt. □

10.3 Der Satz von Stone

Der Satz von Stone verallgemeinert den Satz von Weierstrass, indem er an Stelle von Polynomen eine beliebige punktetrennende Algebra $\mathcal{M}$ setzt. Im Beweis des Satzes werden wir verwenden, dass sich die Funktionen $f \vee g$ und $f \wedge g$ für $f, g \in \mathcal{M}$ durch Funktionen in $\mathcal{M}$ approximieren lassen. Der Beweis hierfür verwendet den Satz von Weierstrass.

Lemma 10.11 (Betragsfunktion, Maximum, Minimum in $\overline{\mathcal{M}}$). *Sei $\mathcal{M} \subseteq \mathcal{C}(E)$ eine Algebra. Dann gilt $f, g \in \mathcal{M} \Rightarrow |f|, f \vee g, f \wedge g \in \overline{\mathcal{M}}g$.*

Beweis. Sei $K \subseteq E$ kompakt. Nach Theorem 5.25 ist also auch $f(K)$ kompakt. Offenbar ist $x \mapsto |x|$ eine stetige Funktion. Nach dem Satz von Weierstrass gibt es also für $\varepsilon > 0$ ein Polynom p mit $\sup_{x \in f(K)} |p(x) - |x||_{f(K)} < \varepsilon$. Da $\mathcal{M}$ eine Algebra ist, ist mit f auch $f \circ p \in \mathcal{M}$. Also gilt

$$\sup_{x \in K} ||f(x)| - p(f(x))| \leq \sup_{x \in f(K)} |p(x) - |x|| < \varepsilon.$$

Damit ist $|f| \in \overline{\mathcal{M}}$. Weiter gilt $f \vee g + f \wedge g = f + g$ und $f \vee g - f \wedge g = |f - g|$, also $f \vee g = (f + g + |f - g|)/2$ und $f \wedge g = (f + g - |f - g|)/2$. Da $\overline{\mathcal{M}}$ nach Lemma 10.6 eine Algebra ist, folgt $f \vee g, f \wedge g \in \overline{\mathcal{M}}$. □

Theorem 10.12 (Satz von Stone). *Sei $\mathcal{M} \subseteq \mathcal{C}(\mathbb{R})$ eine punktetrennende Algebra mit $1 \in \mathcal{M}$ und $f \in \mathcal{C}(\mathbb{R})$. Für jede kompakte Menge K und $\varepsilon > 0$ gibt es ein $p \in \mathcal{M}$ mit $||f - p||_K < \varepsilon$.*

Beweis. Sei $y, z \in K$. Da $\mathcal{M}$ Punkte trennt gibt es ein $g \in \mathcal{M}$ mit $g(y) \neq g(z)$. Wir setzen

$$f_{y,z}(x) := f(y) + (f(z) - f(y)) \frac{g(x) - g(y)}{g(z) - g(y)}.$$

Da $1 \in \mathcal{M}$ ist also auch $f_{y,z} \in \mathcal{M}$ mit $f_{y,z}(y) = f(y), f_{y,z}(z) = f(z)$. Wir setzen

$$A_{y,z} := \{x \in K : f_{y,z}(x) < f(x) + \varepsilon\}.$$

Dann ist $A_{y,z}$ als Urbild einer offenen Menge unter einer stetigen Abbildung offen und $y, z \in A_{y,z}$, also auch $\bigcup_{z \in K} A_{y,z} = K$. Da $K \subseteq \mathbb{R}$ kompakt ist gibt es nach Theorem 5.24 Punkte $z_1, ..., z_n \in K$ mit $\bigcup_{i=1}^n A_{y,z_i} = K$. Wir setzen jetzt

$$f_y := \min_{i=1,\ldots,n} f_{y,z_i}.$$

Dann gilt offenbar $f_y \in \overline{\mathcal{M}}$ nach Lemma 10.11, $f_y(y) = f(y)$ und $f_y < f + \varepsilon$ wegen der Überdeckungseigenschaft. Nun setzen wir

$$B_y := \{x \in K : f_y(x) > f(x) - \varepsilon\}.$$

Wieder ist B_y als Urbild einer offenen Menge unter einer stetigen Abbildung offen und $\bigcup_{y \in K} B_y = K$ wegen $y \in B_y$. Da K kompakt ist, gibt es wegen der Überdeckungseigenschaft wieder $y_1, ..., y_m$ mit $\bigcup_{i=1}^m B_{y_i} = K$. Wir setzen jetzt

$$g := \max_{i=1,\ldots,m} B_{y_i}.$$

Dann ist wegen Lemma 10.11 $g \in \overline{\overline{\mathcal{M}}} = \overline{\mathcal{M}}$ und $f - \varepsilon < g < f + \varepsilon$. Mit anderen Worten haben wir ein $g \in \overline{\mathcal{M}}$ gefunden mit $||f - g||_K < \varepsilon$. Nun genügt es noch $h \in \mathcal{M}$ zu wählen mit $||g - h|| < \varepsilon$, dann gilt $||f - h|| \leq ||f - g|| + ||g - h|| < 2\varepsilon$. □

11 Normale Konvergenz

Während es im letzten Abschnitt um Konvergenz von Funktionenfolgen ging, betrachten wir nun Reihen von Funktionen. Als Anwendung wird es im nächsten Abschnitt vor allem um Potenzreihen, also Grenzwerte von Polynomen, gehen. Wir studieren also das Konvergenzverhalten von $\sum_{n=0}^{\infty} f_n$ für Funktionen f_n. Dies spielen wir zunächst zurück auf Reihen von reellen Zahlen. Hierzu sei an die Definition der Supremumsnorm $||.||_I$ aus Beispiel 8.2 erinnert, und die anschließende gleichmäßige Konvergenz.

11.1 Grundlegendes

Definition 11.1 (Normale Konvergenz). *Sei I ein Intervall und $f_0, f_1, ... : I \to \mathbb{R}$ Funktionen. Konvergiert $\sum_{n=0}^{\infty} ||f_n||_I$, so sagen wir, dass $\sum_{n=0}^{\infty} f_n$* normal konvergiert. *Ist f die Grenzfunktion, so schreiben wir hierfür auch $\sum_{n=0}^{\infty} f_n \xrightarrow{n\to\infty}_{normal} f$.*

Lemma 11.2 (Normale und gleichmäßige Konvergenz, Stetigkeit). *Sei I ein Intervall und $f_0, f_1, ... : I \to \mathbb{R}$, so dass die Reihe $f = \sum_{n=0}^{\infty} f_n$ normal konvergiert.*

1. *Es gilt $\sum_{n=0}^{N} f_n \xrightarrow{N\to\infty}_{glm} f$.*
2. *Sind $f_0, f_1, ...$ stetig in $x_0 \in I$, so ist f in x_0 ebenfalls stetig.*

Beweis. 1. Es gilt nach der Dreiecksungleichung

$$\Big|\Big|f - \sum_{n=0}^{N} f_n\Big|\Big|_I \leq \sum_{n=N+1}^{\infty} ||f_n||_I \xrightarrow{N\to\infty} 0.$$

2. Sei $\varepsilon > 0$. Zunächst wählen wir N groß genug, so dass $\sum_{n=N+1}^{\infty} ||f_n||_I < \varepsilon$. Weiter sei $\delta > 0$ klein genug, so dass $|f_n(x) - f_n(x_0)| < \varepsilon/N$ für $|x - x_0| < \delta$ und $n = 0, ..., N$. Dann gilt

$$|f(x) - f(x_0)| \leq \sum_{n=1}^{N} |f_n(x) - f_n(x_0)| + \sum_{n=N+1}^{\infty} |f_n(x)| + |f_n(x_0)| \leq 3\varepsilon.$$

Daraus folgt die Behauptung. □

11.2 Differenzierbarkeit und Integrierbarkeit

Besonders wichtig in der Mathematik sind Vertauschungssätze, von denen wir bereits einige kennen gelernt haben. Nun geht es um die Vertauschbarkeit der Grenzwertbildung sowie von Differenzier- und Integrierbarkeit unter normaler Konvergenz. Zunächst jedoch benötigen wir einen Satz, wann Grenzwertbildung und Ableitung vertauschen.

Proposition 11.3 (Differenzierbarkeit von Grenzwerten). *Sei I ein Intervall und $f_1, f_2, ... : I \to \mathbb{R}$ stetig differenzierbare Funktionen. Gilt $f_n \xrightarrow{n\to\infty}_{pktw} f$ und $f_n' \xrightarrow{n\to\infty}_{glmK} g$ für Funktionen $f, g : I \to \mathbb{R}$, so ist f stetig differenzierbar und $f' = g$.*

Beweis. Nach Lemma 8.6 ist g stetig. Außerdem gilt für $a \in I$

$$f_n(x) = f_n(a) + \int_a^x f_n'(t)dt.$$

Für $n \to \infty$ folgt daraus mit Theorem 8.19

$$f(x) = f(a) + \int_a^x g(t)dt.$$

Nach dem Hauptsatz der Differential- und Integralrechnung folgt hiermit $f' = g$. $\square$

Korollar 11.4 (Differenzierbarkeit und normale Konvergenz). *Sei I ein Intervall und $f_0, f_1, ... : I \to \mathbb{R}$ differenzierbar. Gilt*

1. $\sum_{n=0}^{\infty} f_n \xrightarrow{n\to\infty}_{pktw} f$,
2. $\sum_{n=0}^{\infty} f_n' \xrightarrow{n\to\infty}_{normal} g$,

so ist f stetig differenzierbar und $f' = g$.

Beweis. Wir verwenden Proposition 11.3. Die Reihe $\sum_{n=0}^{\infty} f_n$ erfüllt nämlich die Voraussetzungen dieser Proposition, weil $\sum_{n=0}^{\infty} f_n'$ nach Lemma 11.2.1 gleichmäßig konvergiert. $\square$

Proposition 11.5 (Integrierbarkeit und normale Konvergenz). *Sei I ein Intervall und $f_1, f_2, ... \in \mathcal{R}(I)$, also Riemann-integrierbar. Ist $f = \sum_{n=0}^{\infty} f_n$ normal konvergent, so ist $f \in \mathcal{R}(I)$ und es gilt für $a, b \in I$*

$$\int_a^b f(x)dx = \sum_{n=0}^{\infty} \int_a^b f_n(x)dx.$$

Beweis. Wir verwenden die gleichmäßige Konvergenz $\sum_{n=0}^{N} f_n \xrightarrow{N\to\infty}_{glm} f$ und Theorem 8.19. Daraus folgt nämlich

$$\begin{aligned}\sum_{n=0}^{\infty} \int_a^b f_n(x)dx &= \lim_{N\to\infty} \sum_{n=0}^{N} \int_a^b f_n(x)dx = \lim_{N\to\infty} \int_a^b \sum_{n=0}^{N} f_n(x)dx \\ &= \int_a^b \sum_{n=0}^{\infty} f_n(x)dx = \int_a^b f(x)dx.\end{aligned}$$

$\square$

11.3 Beispiele

Wir beenden den Abschnitt über normal konvergente Reihen mit zwei Beispielen.

Beispiel 11.6. Die Reihe $\sum_{n=1}^{\infty} \frac{1}{n(x-n)}$ für $x \in [a, b] \subseteq \mathbb{R} \setminus \mathbb{N}$ konvergiert normal.
Denn: Sei N so, dass $[a, b] \subseteq (N-1, N)$. Dann gilt

$$\sum_{n=N+1}^{\infty} \left|\left|\frac{1}{n(x-n)}\right|\right| \leq \sum_{n=N+1}^{\infty} \frac{1}{(n-N)^2} = \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty.$$

Also konvergiert auch $\sum_{n=1}^{\infty} \left|\left|\frac{1}{n(x-n)}\right|\right|$, woraus die normale Konvergenz folgt.

Beispiel 11.7 (Potenzreihen). Potenzreihen gehören zu den wichtigsten Reihen in der Analysis und sind Thema des nächsten Abschnittes. Sie sind von der Form

$$p(x) := \sum_{n=0}^{\infty} a_n x^n$$

für Koeffizienten $a_n \in \mathbb{R}$. Konvergiert p für $|x| \leq r$ absolut für ein $r \in [0, \infty]$, dann ist p auf $B_r(0)$ normal konvergent. Schließlich ist dann

$$\sum_{n=0}^{\infty} ||a_n x^n||_\infty = \sum_{n=0}^{\infty} |a_n| \cdot r^n < \infty.$$

12 Potenzreihen

Wir betrachten nun Potenzreihen[17]

$$p(x) := \sum_{n=0}^{\infty} a_n x^n \tag{12.1}$$

für Koeffizienten $a_n \in \mathbb{R}$, $n = 0, 1, ...$ Wir haben bereits einige Konvergenzreihen kennen gelernt, etwa

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

Außerdem ist auch die geometrische Reihe

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

für $|x| < 1$ eine konvergente Potenzreihe.

12.1 Konvergenz von Potenzreihen

Potenzreihen besitzen einfach nachzuprüfende Konvergenzkriterien, mit denen wir uns nun beschäftigen.

Lemma 12.1 (Absolute Konvergenz von Potenzreihen). *Sei p eine Potenzreihe der Form (12.1) und $x \in \mathbb{R}$ so, dass $p(x)$ konvergiert. Dann konvergiert $p(y)$ absolut für jedes $y \in \mathbb{R}$ mit $|y| < |x|$.*

Beweis. Da $p(x)$ konvergiert, gibt es ein $c \in \mathbb{R}$ mit $|a_n x^n| < c$ für alle n. Dann ist $|a_n y^n| = |a_n x^n| \cdot |(y/x)^n| \leq cq^n$ für $q := y/x < 1$. Also besitzt die Reihe $p(y)$ die Majorante $\sum_{n=0}^{\infty} cq^n$ und ist damit absolut konvergent nach dem Majorantenkriterium. □

[17] Analog bezeichnet man auch

$$\sum_{n=0}^{\infty} a_n (x - x_0)^n$$

für ein $x_0 \in \mathbb{R}$ als Potenzreihe. Eine solche Reihe lässt sich jedoch schnell in eine der Form (12.1) mittels $y := x - x_0$ überführen.

Proposition 12.2 (Konvergenzradius). *Sei p eine Potenzreihe der Form* (12.1) *und*

$$R := R(p) := \sup\{|x| : p(x) \text{ konvergiert}\}.$$

Dann ist $p(x)$ für alle x mit $|x| < R$ absolut konvergent, und für alle x mit $|x| > R$ divergent. Die Größe R heißt auch Konvergenzradius *von p.*

Beweis. Sei zunächst $y \in \mathbb{R}$ mit $|y| < R$. Dann gibt es x mit $|y| < |x| < R$, so dass $p(x)$ konvergiert. Damit konvergiert $p(y)$ absolut nach Lemma 12.1. Sei weiter $y \in \mathbb{R}$ mit $|y| > R$. Angenommen, $p(y)$ wäre konvergent. Dann wäre $p(x)$ für $|R| < |x| < |y|$ absolut konvergent im Widerspruch zur Definition des Konvergenzradius. □

Proposition 12.3 (Berechnung des Konvergenzradius). *Sei p eine Potenzreihe der Form* (12.1) *und R ihr Konvergenzradius. Dann gilt die Formel von Cauchy-Hadamard*

$$R^{-1} = \limsup_{n\to\infty} |a_n|^{1/n},$$

und die Formel von Euler

$$R^{-1} = \lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right|,$$

falls der (zumindest uneigentliche) Grenzwert existiert. Wir setzen hierbei generell $1/0 := \infty$ *und* $1/\infty := 0$.

Beweis. Zunächst setzen wir $\ell := 1/\limsup_{n\to\infty} |a_n|^{1/n}$. Für $x \in \mathbb{R}$ mit $|x| < \ell$ ist

$$\limsup_{n\to\infty} |a_n x^n|^{1/n} = |x| \cdot \limsup_{n\to\infty} |a_n|^{1/n} = |x|\ell^{-1} < 1.$$

Nach dem Wurzelkriterium konvergiert also $p(x)$ absolut. Analog folgt die Divergenz im Falle $|x| > \ell$.

Für das zweite Kriterium zur Berechnung des Konvergenzradius setzen wir $\ell := 1/\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right|$, wobei wir annehmen, dass der Grenzwert existiert. Dann gilt für $x \in \mathbb{R}$ mit $|x| < \ell$

$$\lim_{n\to\infty} \left|\frac{a_{n+1}x^{n+1}}{a_n x^n}\right| = |x| \cdot \lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = |x| \cdot \ell^{-1} < 1,$$

und die Behauptung folgt aus dem Quotientenkriterium. □

Beispiel 12.4 (Konvergenzradius bekannter Reihen).

1. Betrachten wir die Potenzreihe $\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$. Ihr Konvergenzradius ist $R = \infty$, es gilt nämlich

$$\frac{1/(n+1)!}{1/n!} = \frac{1}{n+1} \xrightarrow{n\to\infty} 0.$$

   Nach dem zweiten Kriterium in Proposition 12.3 folgt $R = \infty$.

2. Der Konvergenzradius von $\sum_{n=0}^{\infty} x^n$ ist $R = 1$. Es gilt nämlich

$$\limsup_{n\to\infty} 1^{1/n} = 1.$$

Beispiel 12.5 (Konvergenz auf dem Konvergenzradius). Im allgemeinen kann man keine Aussage über die Konvergenz der Potenzreihe p mit Potenzradius R und $x = R$ oder $x = -R$ machen. Betrachte hierzu etwa die Potenzreihen

$$p(x) := \sum_{n=0}^{\infty} x^n, \qquad q(x) := \sum_{n=1}^{\infty} \frac{x^n}{n}, \qquad r(x) := \sum_{n=1}^{\infty} \frac{x^n}{n^2}.$$

Dann ist nach dem Wurzelkriterium und wegen

$$\limsup_{n\to\infty} 1^{1/n} = 1, \qquad \limsup_{n\to\infty} \frac{1}{n^{1/n}} = 1, \qquad \limsup_{n\to\infty} \frac{1}{n^{2/n}} = 1$$

der Konvergenzradius aller drei Potenzreihen gleich 1. Außerdem divergiert $p(1)$ und $p(-1)$ und $q(1)$. Allerdings konvergieren $q(-1)$ (nach dem Leibniz-Kriterium), sowie $r(1)$ und $r(-1)$ (nach dem Majorantenkriterium mit der Majorante $\sum_{n=1}^{\infty} n^{-2}$).

Lemma 12.6 (Restabschätzung). *Die Potenzreihe p der Form* (12.1) *habe Konvergenzradius $R > 0$. Dann gibt es zu jedem $0 < r < R$ und $n \in \mathbb{N}$ ein $c_n > 0$ mit*

$$\Big|\sum_{k=n}^{\infty} a_k x^k\Big| \le c_n \cdot |x|^n$$

für $|x| \le r$.

Beweis. Setze $c_n := \sum_{k=n}^{\infty} a_k r^{k-n} < \infty$. Dann gilt

$$\Big|\sum_{k=n}^{\infty} a_k x^k\Big| = |x|^n \cdot \Big|\sum_{k=n}^{\infty} a_k x^{k-n}\Big| \le c_n \cdot |x|^n.$$

□

Lemma 12.7 (Nullstellen von Potenzreihen). *Sei $p \neq 0$ eine Potenzreihe p der Form* (12.1) *mit Konvergenzradius $R > 0$ und $\mathcal{N} := \{x : p(x) = 0\}$. Dann gib es ein $\varepsilon > 0$ mit $\mathcal{N} \cap B_\varepsilon(0) \subseteq \{0\}$.*

Beweis. Sei $0 < r < R$ und N minimal mit $a_N \neq 0$. Angenommen, es gibt eine Nullfolge $(x_k)_{k=1,2,...}$ mit $x_k \in B_{r/k}(0)$ und $p(x_k) = 0$. Nach Lemma 12.6 gibt es ein c mit $|p(x) - a_N x^N| \le c x^{N+1}$ für $|x| \le r$. Insbesondere gilt also $|p(x_k) - a_N x_k^N| \le c|x_k|^{N+1}$, also $|a_N| \le c|x_k|$. Da $(x_k)_{k=1,2,...}$ eine Nullfolge ist, folgt $a_N = 0$ im Widerspruch zur Annahme. □

Korollar 12.8 (Gleichheit von Potenzreihen). *Seien p, q Potenzreihen mit Konvergenzradien $R(p), R(q) > 0$, so dass*

$$p(x) = \sum_{n=0}^{\infty} a_n x^n, \qquad q(x) = \sum_{n=0}^{\infty} b_n x^n.$$

Gilt $p(x) = q(x)$ in einem Bereich $B_r(0)$ mit $0 < r < R(p) \wedge R(q)$, so gilt $a_n = b_n, n = 0, 1, ...$

Beweis. Man wende Lemma 12.7 auf die Potenzreihe $p - q$ an. □

12.2 Differentiation und Integrierbarkeit von Potenzreihen

Ableitungen und Integration von Polynomen ist einfach: aufgrund der Linearität gilt

$$\frac{d}{dx}\Big(\sum_{n=1}^{N} a_n x^n\Big) = \sum_{n=1}^{N} \frac{d}{dx} a_n x^n, \qquad \int \sum_{n=1}^{N} a_n x^n = \sum_{n=1}^{N} \int a_n x^n.$$

Bei Potenzreihen ist dieses gliedweise differenzieren bzw. integrieren nicht sofort klar, weil es sich bei den Summen ja um unendliche handelt. Wir sehen jedoch in Proposition 12.11 feststellen, dass auch für Potenzreihen gliedweises Differenzieren das richtige Ergebnis liefert. Hierzu müssen wir allerdings zunächst feststellen, dass konvergente Potenzreihen stetige Funktionen darstellen.

Lemma 12.9 (Normale Konvergenz von Potenzreihen). *Sei p eine Potenzreihe der Form* (12.1) *mit Konvergenzradius $R > 0$. Dann ist p normal (und damit nach Lemma 11.2 auch gleichmäßig) konvergent in jedem Ball $B_r(0)$ mit $0 < r < R$, d.h. $\sum_{n=0}^{\infty} |a_n| \cdot r^n$ konvergiert.*

Beweis. Die Aussage ist klar, da wir in Proposition 12.2 bereits gezeigt haben dass Potenzreihen innerhalb ihres Konvergenzradius absolut konvergieren. □

Korollar 12.10 (Stetigkeit von Potenzreihen). *Sei p eine Potenzreihe der Form* (12.1) *mit Konvergenzradius $R > 0$. Dann ist p auf dem offenen Ball $B_R(0)$ stetig.*

Beweis. Sei $x \in B_R(0)$. Dann ist auch $x \in B_r(0)$ für ein $0 < r < R$. Deshalb konvergiert p auf $B_r(0)$ nach Lemma 12.9 normal, also nach Lemma 11.2 auch gleichmäßig. Da der gleichmäßige Grenzwert stetiger Funktionen nach Lemma 8.6 stetig ist, folgt die Aussage. □

Proposition 12.11 (Differenzierbarkeit von Potenzreihen). *Sei p eine Potenzreihe der Form* (12.1) *mit Konvergenzradius $R > 0$. Dann ist p im Punkt x für $|x| < R$ differenzierbar mit*

$$p'(x) := \sum_{n=0}^{\infty} (n+1) a_{n+1} x^n. \tag{12.2}$$

Beweis. Wir verwenden Proposition 11.3. Für den Konvergenzradius R von p gilt $R^{-1} = \limsup_{n\to\infty} |a_n|^{1/n}$. Entscheidend ist es zu bemerken, dass dann auch

$$\limsup_{n\to\infty} (n+1)^{1/n} |a_{n+1}|^{1/n} = \limsup_{n\to\infty} (n+1)^{1/n} \cdot \limsup_{n\to\infty} |a_{n+1}|^{1/(n+1)} = R^{-1}$$

nach Theorem 3.6, der Konvergenzradius der rechten Seite von (12.2) also ebenfalls R ist. □

12.3 Beispiele

Beispiel 12.12 (Die Logarithmus-Reihe). Wir verwenden die Eigenschaften der Differenzierbarkeit und Integrierbarkeit von Potenzreihen, um nun die Potenzreihe der Funktion $x \mapsto \log(1 + x)$ herauszufinden. Die Berechnung basiert auf der geometrischen Reihe

$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$. Dies ist eine Potenzreihe mit Konvergenzradius $R = 1$. Deshalb gilt für $|x| < 1$ nach dem Hauptsatz der Differential- und Integralrechnung

$$\log(1+x) = \int_0^x \frac{d}{dy}\log(1+y)dy = \int_0^x \frac{1}{1+y}dy = \int_0^x \sum_{n=0}^{\infty}(-y)^n dy$$
$$= \sum_{n=0}^{\infty}\int_0^x (-y)^n dy = -\sum_{n=0}^{\infty}\frac{(-x)^{n+1}}{n+1} = -\sum_{n=1}^{\infty}\frac{(-x)^n}{n}.$$

Bemerkung 12.13 (Vorbemerkung zur Binomialreihe). Im nächsten Beispiel behandeln wir die Binomialreihe. Hierzu benötigen wir eine Definition und eine Aussage, die wir hier beweisen wollen. Zunächst verallgemeinern wir die Definition des Binomialkoeffizienten und setzen für $r \in \mathbb{R}$ und $n \in \mathbb{N}_0$

$$\binom{r}{n} := \frac{r(r-1)\cdots(r-n+1)}{n!}. \tag{12.3}$$

Für $r, s \in \mathbb{R}$ gilt dann

$$\sum_{n=0}^{m}\binom{r}{n}\binom{s}{m-n} = \binom{r+s}{m}. \tag{12.4}$$

Denn: Wir zeigen die Aussage erst für $r \in \mathbb{R}$, $s \in \mathbb{N}_0$. Für $s = 0$ ist die Aussage klar. Gilt sie für ein s, so folgt aus Proposition 1.34 (das ganz einfach auf $r \in \mathbb{R}$ erweitert werden kann), dass

$$\binom{r+1+s}{m} = \binom{r+s}{m} + \binom{r+s}{m-1} = \sum_{n=0}^{m}\binom{r}{n}\left(\binom{s}{m-n} + \binom{s}{m-1-n}\right)$$
$$= \sum_{n=0}^{m}\binom{r}{n}\binom{s+1}{m-n}.$$

Damit ist die Aussage für $s \in \mathbb{N}_0$ gezeigt. Um sie auf $s \in \mathbb{R}$ zu erweitern, stellt man fest, dass beide Seiten der Gleichung (12.4) ein Polynom in s vom Grad m darstellen, die auf unendlich vielen Punkten $s \in \mathbb{N}_0$ übereinstimmen. Da ein Polynom ungleich 0 vom Grad m höchstens m Nullstellen hat, muss die Differenz der beiden Seiten gleich 0 sein.

Beispiel 12.14 (Die Binomialreihe). Wir werden nun zeigen, dass für $x \in (-1, 1)$

$$(1+x)^r = \sum_{n=0}^{\infty}\binom{r}{n}x^n. \tag{12.5}$$

Diese Formel erweitert die in Proposition 1.40 formulierte binomische Formel.

Wir bezeichnen mit $\beta_r(x)$ die rechte Seite von (12.5). Der Konvergenzradius R von β_r berechnen wir mittels

$$\left|\frac{\binom{r}{n+1}}{\binom{r}{n}}\right| = \left|\frac{r-n}{n+1}\right| \xrightarrow{n\to\infty} 1$$

als $R = 1$. Weiter verwenden wir (12.4) und die absolute Konvergenz von Potenzreihen für $r, s \in \mathbb{R}$ für

$$\begin{aligned}\beta_r(x)\beta_s(x) &= \sum_{n=0}^{\infty}\binom{r}{n}x^n\sum_{m=0}^{\infty}\binom{s}{n}x^n = \sum_{n=0}^{\infty}\sum_{m=n}^{\infty}\binom{r}{n}\binom{s}{m-n}x^m \\ &= \sum_{m=0}^{\infty}\left(\sum_{n=0}^{m}\binom{r}{n}\binom{s}{m-n}\right)x^m = \sum_{m=0}^{\infty}\binom{r+s}{m}x^m = \beta_{r+s}(x)\end{aligned}$$

wobei wir Proposition (4.23) benutzt haben. Außerdem gilt mit Proposition (9.16)

$$\begin{aligned}\frac{\beta_r(x)-1}{r} &= \sum_{n=1}^{\infty}\frac{r\cdots(r-n+1)}{r\cdot n!}x^n = \sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{n}\underbrace{\frac{(1-r)\cdots((n-1)-r)}{(n-1)!}}_{\xrightarrow{r\to 0}1}x^n \\ &\xrightarrow{r\to 0} -\sum_{n=1}^{\infty}\frac{(-x)^n}{n} = \log(1+x).\end{aligned}$$

Die Funktion $x \mapsto \beta_r(x)$ erfüllt also (1) und $(2_{\log(1+x)})$ aus Abschnitt 6.1. Mit anderen Worten ist nach Lemma 6.6 $\beta_r(x) = \exp(r\log(1+x)) = (1+x)^r$.

12.4 Taylor-Polynome und Taylor-Reihen

Sei f eine n-mal stetig differenzierbare Funktion. Wir suchen ein Polynom $T_n f$, dessen erste n Ableitungen in einem Punkt x_0 mit denen von f übereinstimmen. Die Hoffnung hierbei ist, dass $T_n f$ die Funktion f approximiert. Dies führt zu den Taylor-Polynomen.

Lemma 12.15 (Taylor-Polynom). *Sei $I = [a, b]$ und $f : I \to \mathbb{R}$ eine n-mal in einem Punkt $x_0 \in I$ stetig differenzierbare Funktion. Dann ist*

$$(T_n f)(x; x_0) := (T_n f)(x) := \sum_{k=0}^{n}\frac{f^{(k)}(x_0)}{k!}(x-x_0)^k$$

mit $f^{(0)} := f$ und $f^{(k)}$ der k-ten Ableitung von f das einzige Polynom vom Grad n, dessen ersten n Ableitungen in x_0 mit denen von f übereinstimmt. Wir nennen es Taylor-Polynom *von Grad n im Entwicklungspunkt x_0. Ist $x_0 = 0$, so heißt $T_n f$ auch* Maclaurin-Reihe*.*

Beweis. OBdA ist $x_0 = 0$. Zunächst gilt für $0 \le j \le n$

$$(T_n f)^{(j)}(x_0) = \sum_{k=j}^{n}\frac{f^{(k)}(x_0)}{(k-j)!}x^{k-j}\Big|_{x=x_0} = f^{(j)}(x_0),$$

also stimmen die ersten n Ableitungen von f und $T_n f$ in x_0 überein. Es bleibt also noch die Eindeutigkeit des Polynoms zu zeigen. Hierfür sei $p(x) = a_0 + a_1(x-x_0) + \cdots + a_n(x-x_0)^n$ ein Polynom vom Grad n. Dann ist $p^{(j)}(x_0) = j!a_j$. Ist also $p^{(j)}(x_0) = f^{(j)}(x_0)$, so muss $a_j = \frac{f^{(j)}(x_0)}{j!}$, woraus die Eindeutigkeit folgt. $\square$

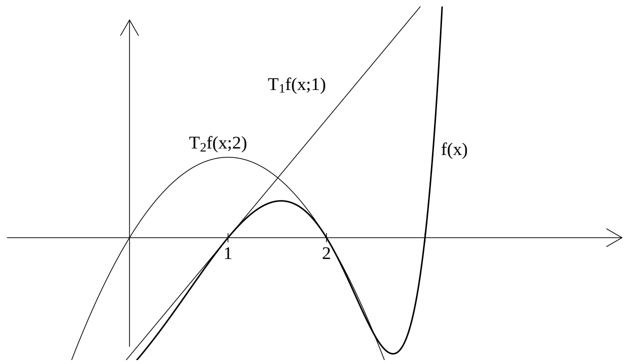

Abbildung 12.1: Eine Funktion f und ihre Taylor-Approximationen $T_1 f$ und $T_2 f$ in den Entwicklungspunkten 2 und 1.

Bemerkung 12.16. Sei f mindestens n-mal in x_0 differenzierbar. Ziel eines Taylor Polynoms $T_n f$ ist es, die Funktion f (zumindest in der Nähe von x_0) möglichst gut zu approximieren. Um einzusehen, dass dies Sinn macht, betrachten wir das Taylor-Polynome $(T_1 f)(x; x_0) = f(x_0) + f'(x_0)(x - x_0)$. Mit anderen Worten stelle (siehe Abbildung 12.1) das Polynom $T_1 f$ eine Tangente an f im Punkt x_0 dar. Weiter ist $T_2 f$ eine Parabel, die sich im Punkt x_0 an f anschmiegt. Für höhere n sollte sich damit $T_n f$ immer näher an f annähern.

Um festzustellen, ob $T_n f \approx f$ für große n gilt, betrachten wir nun die Abweichung $f - T_n f$, die auch Restglied genannt wird.

Theorem 12.17 (Restgliedformeln). *Sei I ein Intervall und $f : I \to \mathbb{R}$ eine $n+1$-mal stetig differenzierbare Funktion. Dann gilt die Integraldarstellung des Restgliedes*

$$R_{n+1}(x; x_0) := R_{n+1}(x) := f(x) - T_n f(x; x_0) = \frac{1}{n!} \int_{x_0}^{x} (x - y)^n f^{(n+1)}(y) dy.$$

Weiter gibt es ein ξ zwischen x_0 und x mit

$$R_{n+1}(x; x_0) = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)^{n+1}.$$

Dies ist die Lagrange-Form des Restgliedes.

Beweis. Die erste Formel für R_{n+1} beweisen wir durch vollständige Induktion. Für $n = 0$ ist die Formel klar, da

$$f(x) - (T_0 f)(x) = f(x) - f(x_0) = \int_{x_0}^{x} f'(y) dy$$

nach dem Hauptsatz der Differential- und Integralrechnung. Gilt die Formel für ein n, gilt also

$$f(x) - (T_n f)(x; x_0) = \frac{1}{n!} \int_{x_0}^{x} (x - y)^n f^{(n+1)}(y) dy,$$

so folgt mittels partieller Integration

$$\begin{aligned} f(x) - (T_n f)(x; x_0) &= \frac{1}{n!} \Big(\frac{1}{n+1} (x - y)^{n+1} f^{(n+1)}(y) \Big|_{x_0}^{x} + \int_{x_0}^{x} \frac{1}{n+1} (x - y)^{n+1} f^{(n+2)}(y) dy \Big) \\ &= \frac{1}{(n+1)!} f^{(n+1)}(x)(x - x_0)^{n+1} + \frac{1}{(n+1)!} \int_{x_0}^{x} (x - y)^{n+1} f^{(n+2)}(y) dy. \end{aligned}$$

Dies ist aber gerade die Behauptung für $n+1$.
Um die zweite Formel zu zeigen, verwenden wir den Mittelwertsatz der Integralrechnung, Proposition 8.35. Es hat nämlich $(x-y)^n$ im Bereich zwischen x_0 und x konstantes Vorzeichen, und damit gibt es ein ξ zwischen x_0 und x mit

$$\int_{x_0}^{x}(x-y)^n f^{(n+1)}(y)dy = f^{(n+1)}(\xi)\int_{x_0}^{x}(x-y)^n dy = \frac{1}{n+1}f^{(n+1)}(\xi)(x-x_0)^{n+1},$$

woraus die zweite Restgliedformel folgt. □

Korollar 12.18 (Hinreichende Bedingung für lokales Extremum). *Sei I ein Intervall, $f \in \mathcal{C}^{(n+1)}(I)$ und $x_0 \in I$ so, dass $f^{(1)}(x_0) = \dots = f^{(n)}(x_0) = 0$ und $f^{(n+1)}(x_0) \neq 0$. Dann gilt:*

- *Ist n ungerade und ist $f^{(n+1)}(x_0) > 0$, so hat f in x_0 ein lokales Minimum.*
- *Ist n ungerade und ist $f^{(n+1)}(x_0) < 0$, so hat f in x_0 ein lokales Maximum.*
- *Ist n gerade, so hat f in x_0 kein lokales Extremum.*

Beweis. Offensichtlich ist $(T_n f)(x;x_0) = f(x_0)$, also $f(x) = f(a) + R_{n+1}(x;x_0)$. Wir verkleinern I so, dass $f^{(n+1)} > 0$ (bzw. $f^{(n+1)} < 0$) auf ganz I gilt. Nun gilt nach der Lagrange-Form von R_{n+1}, dass $f(x) - f(x_0) = f^{(n+1)}(\xi)(x-x_0)^{n+1}$ auf I. Im ersten Fall hat damit f ein Minimum, im zweiten Fall ein Maximum in x_0. Im dritten Fall ist – etwa falls $f^{(n+1)}(x_0) > 0$ – gerade $f(x) - f(x_0) < 0$ für $x < x_0$ und $f(x) - f(x_0) > 0$ für $x > x_0$. Daraus folgt auch die dritte Aussage. □

Bemerkung 12.19 (Ein falsches Resultat). Folgendes Resultat über Taylor-Polynome wäre wünschenswert, ist aber leider falsch, wie das nächste Beispiel zeigt:

> **(Falsche) Proposition** Sei $f \in \mathcal{C}^\infty(I)$ (d.h. unendlich oft auf I differenzierbar). Dann gilt für alle $x_0 \in I$, dass $T_n f(.;x_0) \xrightarrow{n\to\infty}_{glmK} f$.

Diese Proposition gilt auch dann nicht, wenn man die Konvergenz zur punktweisen Konvergenz abschwächt.

Beispiel 12.20 (Taylor-Reihen müssen nicht konvergieren). Wir betrachten die Funktion

$$f : x \mapsto \begin{cases} \exp(-1/x), & x \geq 0, \\ 0, & x < 0. \end{cases}$$

Dann gilt in $x_0 = 0$, dass $f^{(n)} = 0$, $n = 1, 2, ...$ Um dies einzusehen, zeigt man, dass $f^{(n)}(x) = p_n(1/x) \cdot \exp(-1/x)$ für $x \geq 0$ für ein Polynom $2n$-ten Grades p_n. Für $n = 0$ ist diese Behauptung klar. Gilt sie für ein n, so gilt

$$f^{(n+1)}(x) = -\Big(p_n'(1/x)\frac{1}{x^2} + p_n(x)\frac{1}{x^2}\Big)\exp(-1/x).$$

Mit $p_{n+1}(x) = p_n'(x) \cdot x^2 + p_n(x) \cdot x^2$ folgt die Behauptung.

Aus $f^{(n)}(x) = p_n(1/x) \cdot \exp(-1/x)$ folgt, dass $f^{(n)}(x) \xrightarrow{x\downarrow 0} 0$ und damit ist die Taylor-Reihe im Entwicklungspunkt 0 gerade 0. Insbesondere konvergiert also die Taylor-Reihe $T_n f$ nicht gegen f.

13 Approximation periodischer Funktionen

Wie wir in Beispiel 10.4 gesehen haben, ist die Menge der Linearkombinationen von trigonometrischen Funktionen punktetrennend im Raum der stetigen, 2π-periodischen Funktionen. Ähnlich wie bei Potenzreihen kann man also solche Funktionen durch Reihen trigonometrischen Funktionen approximieren. Dies ist Ziel dieses Abschnitts.

13.1 Grundlegendes

Da wir uns mit periodischen Funktionen beschäftigen werden, definieren wir zunächst weitere wichtige Funktionenklassen.

Definition 13.1 (Periodische Funktionen). *Sei $f : \mathbb{R} \to \mathbb{R}$. Gilt $f(x+T) = f(x)$ für alle $x \in \mathbb{R}$, so sagen wir, dass die Funktion f die Periode T hat. Wir erinnern an die Menge $\mathcal{R}(I)$ der Riemann-integrierbaren Funktionen und definieren weiter die Funktionenräume*

$$\begin{aligned} \mathcal{P} &:= \mathcal{P}([-\pi,\pi]) := \{f : \mathbb{R} \to \mathbb{R} : f \text{ hat Periode } 2\pi\}, \\ \mathcal{C}_p(\mathbb{R}) &:= \mathcal{C}(\mathbb{R}) \cap \mathcal{P}, \\ \mathcal{R}_p &:= \mathcal{R}_p([-\pi,\pi]) := \mathcal{R}([-\pi,\pi]) \cap \mathcal{P}. \end{aligned}$$

Da wir durch trigonometrische Polynome approximieren wollen, setzen wir

$$\mathcal{M} := \Big\{ x \mapsto \sum_{k=0}^{n} a_k \cos(kx) + b_k \sin(kx), n \in \mathbb{N}, a_0, ..., a_n, b_0, ..., b_n \in \mathbb{R} \Big\}.$$

Mit dem Satz von Stone und Beispiel 10.4 können wir sofort eine wichtige Eigenschaft schließen.

Theorem 13.2 (Approximationssatz für trigonometrische Polynome). *Sei $f \in \mathcal{C}_p$ und $\varepsilon > 0$. Dann gibt es ein $p \in \mathcal{M}$ mit $||f - p|| < \varepsilon$.*

Beweis. Nach Beispiel 10.4 ist $\mathcal{M}$ eine punktetrennende Algebra in $\mathcal{C}_p$. Damit folgt die Aussage aus dem Satz von Stone, Theorem 10.12. □

Wir erinnern an die wichtige Formel $e^{ix} = \cos(x) + i\sin(x)$. Mit ihrer Hilfe kann man Funktionen $f \in \mathcal{M}$ umschreiben, was im Folgenden sehr nützlich ist.

Lemma 13.3 (Umrechnung trigonometrischen Funktionen in Exponentiale). *Sei*

$$f : x \mapsto \sum_{k=0}^{n} a_k \cos(kx) + b_k \sin(kx) = \sum_{k=-n}^{n} c_k e^{ikx}$$

mit $n \in \mathbb{N}$, $a_0, ..., a_n, b_0, ..., b_n, c_0, ..., c_n \in \mathbb{C}$. Dann gilt

$$\begin{aligned} a_0 &= c_0, & a_k &= c_k + c_{-k}, & b_k &= i(c_k - c_{-k}), \\ c_0 &= a_0, & c_k &= \tfrac{1}{2}(a_k - ib_k), & c_{-k} &= \tfrac{1}{2}(a_k + ib_k). \end{aligned}$$

Insbesondere ist

$$\mathcal{M} = \Big\{ \sum_{k=-n}^{n} c_k e^{ikx}, n \in \mathbb{N}, c_{-n}, ..., c_n \in \mathbb{C}, c_{-k} = \overline{c_k}, k = 0, ..., n \Big\}.$$

Beweis. Wir schreiben einfach, weil cos eine gerade und sin eine ungerade Funktion ist,

$$\sum_{k=-n}^{n} c_k e^{ikx} = \sum_{k=-n}^{n} c_k(\cos(kx) + i\sin(kx)) = \sum_{k=0}^{n}(c_k + c_{-k})\cos(kx) + i(c_k - c_{-k})\sin(kx). \tag{13.1}$$

Hieraus folgen bereits alle Aussagen. □

Wir werden im Folgenden überwiegend mit der Darstellung von Funktionen $f \in \mathcal{M}$ aus (13.1) arbeiten. Wir benötigen hierfür zwei Schreibweisen, die wir gesondert einführen wollen.

Bemerkung 13.4 (Notation).

1. Sei $f : \mathbb{R} \to \mathbb{C}$. Dann definieren wir

$$\int_a^b f(x)dx := \int_a^b \mathrm{Re} f(x)dx + i\int_a^b \mathrm{Im} f(x)dx \tag{13.2}$$

2. Seien $(c_k)_{k\in\mathbb{Z}}$ eine (zweifach unendliche) $\mathbb{C}$-wertige Folge. Dann definieren wir,

$$\sum_{k=-\infty}^{\infty} c_k := \lim_{n\to\infty} \sum_{k=-n}^{n} c_k.$$

   Es muss für die Existenz der Reihe $\sum_{k=-\infty}^{\infty}$ also nur die letzte Summe konvergieren, aber nicht notwendigerweise $\sum_{k=0}^{n} c_k$ und $\sum_{k=0}^{n} c_{-k}$.

13.2 Gleichmäßige Approximation stetiger Funktionen

Ziel dieses Abschnittes ist es, periodische Funktionen durch trigonometrische Polynome zu approximieren. Die trigonometrischen Funktionen haben eine Orthogonalitätseigenschaft, die wir mittels des komplexen Integrals formulieren. Dies werden wir im nächsten Abschnitt für unstetige, periodische Funktionen erweitern.

Lemma 13.5 (Orthogonalität von trigonometrischen Polynomen). *Es gilt*[18]

$$\frac{1}{2\pi}\int_0^{2\pi} e^{ikx}e^{-i\ell x}dx = \delta_{k\ell}.$$

Beweis. Für $k = \ell$ ist die Behauptung klar, da der Integrand 1 ist. Wir berechnen direkt mittels der komplexen Integrale für $k \neq \ell$

$$\int_0^{2\pi} e^{i(k-\ell)x}xdx = \frac{1}{i(k-\ell)}\left(e^{i(k-\ell)2\pi} - 1\right) = 0,$$

weil $e^{2\pi i} = 1$. □

Geht man von einer Darstellung wie in (13.3) aus, kann man sich fragen, wie man denn die Koeffizienten c_k mittels der Funktion f berechnen kann.

[18] Das Kronecker-Symbol δ_{xy} definieren wir durch

$$\delta_{xy} := \begin{cases} 1, & x = y, \\ 0, & \text{sonst.} \end{cases}$$

Korollar 13.6 (Eindeutigkeit von Koeffizienten). *Sei $f \in \mathcal{C}_p$ mit*

$$f(x) = \sum_{k=-\infty}^{\infty} c_k e^{ikx}, \tag{13.3}$$

wobei die rechte Seite gleichmäßig konvergiert. Dann gilt

$$c_k = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-ikx} dx$$

für alle $k \in \mathbb{Z}$.

Beweis. Wir schreiben wegen der gleichmäßigen Konvergenz

$$\begin{aligned} c_k &= \sum_{\ell=-\infty}^{\infty} c_\ell \delta_{k\ell} = \sum_{\ell=-\infty}^{\infty} c_\ell \frac{1}{2\pi} \int_0^{2\pi} e^{-ikx} e^{i\ell x} dx = \frac{1}{2\pi} \int_0^{2\pi} e^{-ikx} \sum_{\ell=-\infty}^{\infty} c_\ell e^{i\ell x} dx \\ &= \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-ikx} dx. \end{aligned}$$

□

Das letzte Korollar führt uns direkt zur Definition von Fourier-Polynomen und der Fourier-Reihe einer periodischen Funktion. Wir erinnern uns, dass die Idee von Taylor-Polynomen die Gleichheit der ersten n Ableitungen einer zu approximierenden Funktion f in einem Punkt x_0 mit dem Taylor-Polynom $T_n f$ war. Bei Fourier-Reihen ist die grundlegende Idee das letzte Korollar. Selbst ohne die gleichmäßige Konvergenz der rechten Seite in (13.3) zu fordern, kann man durch die rechte Seite eine Folge von Funktionen definieren, die wir Fourier-Polynom nennen.

Definition 13.7 (Fourier-Polynom). *Sei $f \in \mathcal{R}_p$, also eine periodische, Riemann-integrierbare Funktion. Dann definieren wir für $n \in \mathbb{N}_0$ das n-te Fourier-Polynom $S_n f$ mittels*

$$(S_n f)(x) := \sum_{k=-n}^{n} \widehat{f}(k) e^{ikx}$$

mit

$$\widehat{f}(k) := \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-ikx} dx.$$

Dabei heißt $\widehat{f}(k)$ der k-te Fourier-Koeffizient und $S_\infty f$ heißt Fourier-Reihe von f.

Bemerkung 13.8 (Alternative Schreibweise). Kombiniert man die letzten beiden Gleichungen, so erhält man

$$(S_n f)(x) = \frac{1}{2\pi} \sum_{k=-n}^{n} \int_0^{2\pi} f(t) e^{ik(x-t)} dt. \tag{13.4}$$

Proposition 13.9 (Parseval'sches Theorem). *Seien $f, g \in \mathcal{R}_p$, so dass für $x \in \mathbb{R}$*

$$f(x) = \sum_{k=-\infty}^{\infty} c_k e^{ikx}, \qquad g(x) = \sum_{k=-\infty}^{\infty} d_k e^{ikx}$$

mit $c_k, d_k \in \mathbb{C}$, $k \in \mathbb{Z}$ und beide Reihen gleichmäßig konvergieren. Dann gilt

$$\frac{1}{2\pi}\int_0^{2\pi} f(x)g(x)dx = \sum_{k=-\infty}^{\infty} c_k \overline{d}_k,$$

insbesondere also

$$\frac{1}{2\pi}\int_0^{2\pi} f(x)^2 dx = \sum_{k=-\infty}^{\infty} |c_k|^2.$$

Beweis. Wir schreiben

$$\frac{1}{2\pi}\int_0^{2\pi} f(x)g(x)dx = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty}\sum_{\ell=-\infty}^{\infty} c_k \overline{d}_\ell \int_0^{2\pi} e^{ikx}e^{-i\ell x}dx = \sum_{k=-\infty}^{\infty} c_k \overline{d}_k.$$

□

Korollar 13.10 (Identitätssatz). *Seien $f, g \in \mathcal{C}_p$ mit $\widehat{f}(k) = \widehat{g}(k), k \in \mathbb{Z}$. Dann ist $f = g$.*

Beweis. Wir betrachten die Funktion $h := f - g$. Dann gilt offenbar $\widehat{h}(k) = 0, k \in \mathbb{Z}$. Insbesondere gilt also

$$\widehat{h}(k) = \frac{1}{2\pi}\int_0^{2\pi} h(x)e^{-ikx}dx = 0. \tag{13.5}$$

Sei nun $(p_n)_{n=1,2,\ldots}$ eine $\mathcal{M}$-wertige Folge (die nach Theorem 13.2 existiert) mit $p_n \xrightarrow{n\to\infty}_{glm} h$. Nun gilt

$$\int_0^{2\pi} |h(x)|^2 = \lim_{n\to\infty}\int_0^{2\pi} h(x)\overline{p}_n(x)dx = 0$$

wegen (13.5). Daraus folgt aber $h = 0$. □

Korollar 13.11 (Darstellungssatz). *Sei $f \in \mathcal{C}_p$, so dass $S_n f$ gleichmäßig konvergiert. Dann gilt $S_n f \xrightarrow{n\to\infty}_{glm} f$.*

Beweis. Nach Voraussetzung, d.h. wegen der gleichmäßigen Konvergenz, definiert $g := S_\infty f = \lim_{n\to\infty} S_n f$ eine stetige Funktion. Für diese gilt

$$\begin{aligned}\widehat{g}(k) &= \frac{1}{2\pi}\int_0^{2\pi} e^{ikx}\sum_{\ell=-\infty}^{\infty}\widehat{f}(\ell)e^{-i\ell x}dx = \frac{1}{2\pi}\sum_{\ell=-\infty}^{\infty}\widehat{f}(\ell)\int_0^{2\pi} e^{ikx}e^{-i\ell x}dx \\ &= \sum_{\ell=-\infty}^{\infty}\widehat{f}(\ell)\delta_{k\ell} = \widehat{f}(k).\end{aligned}$$

Insbesondere sind f und g Funktionen mit denselben Fourier-Koeffizienten. Nach Korollar 13.10 folgt die Behauptung. □

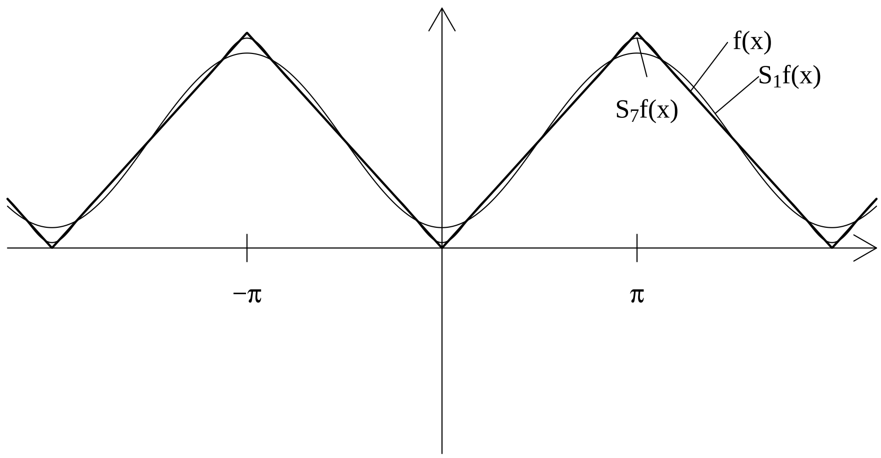

Abbildung 13.1: Die $f : x \mapsto |x|$ aus Beispiel 13.12 und zwei Aprpximationen von ihr. Aus (13.6) folgt $S_n f \xrightarrow{n\to\infty}_{glm} f$.

Beispiel 13.12 (Berechnung von $\zeta(2)$). Wie berechnen nun erneut (siehe Theorem 9.18) den numerischen Wert von $\zeta(2) = \sum_{n=1}^{\infty} \frac{1}{n^2}$. Hierzu betrachten wir die 2π-periodische Funktion $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = |x|$ im Bereich $[-\pi, \pi]$; siehe auch Abbildung 13.1. Ihre Fourier-Reihe ist gegeben durch die Fourier-Koeffizienten

$$\begin{aligned}
\widehat{f}(0) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} |x| dx = \frac{1}{\pi}\int_0^{\pi} x dx = \frac{\pi}{2},\\
\widehat{f}(k) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} |x|\big(\cos(kx) - i\sin(kx)\big)dx = \frac{1}{\pi}\int_0^{\pi} x\cos(kx)dx\\
&= \frac{1}{\pi}\Big(x\frac{\sin(kx)}{k}|_0^{\pi} - \int_0^{\pi}\frac{\sin(kx)}{k}dx\Big) = \frac{1}{k^2\pi}\cos(kx)|_0^{\pi} = \begin{cases} -\frac{2}{k^2\pi}, & k \text{ ungerade},\\ 0, & k \text{ gerade},\end{cases}
\end{aligned}$$

wobei wir partielle Integration verwendet haben. Damit gilt

$$S_\infty f(x) = \frac{\pi}{2} - \frac{4}{\pi}\Big(\cos(x) + \frac{\cos(3x)}{3^2} + \frac{\cos(5x)}{5^2} + \cdots\Big). \tag{13.6}$$

Diese Reihe konvergiert wegen majorisierter Konvergenz absolut und gleichmäßig auf $[0, 2\pi]$. Aus Korollar 13.11 folgt deswegen $S_\infty f(x) = f(x)$. Ausgewertet an der Stelle $x = \pi$ sehen wir

$$\frac{\pi^2}{8} = 1 + \frac{1}{3^2} + \frac{1}{5^2} + \cdots = \sum_{k=0}^{\infty}\frac{1}{(2k+1)^2}.$$

Nun berechnen wir wegen der unbedingten Konvergenz der Reihe

$$\sum_{k=1}^{\infty}\frac{1}{k^2} = \sum_{k=1}^{\infty}\frac{1}{(2k)^2} + \sum_{k=0}^{\infty}\frac{1}{(2k+1)^2} = \frac{1}{4}\sum_{k=1}^{\infty}\frac{1}{k^2} + \frac{\pi^2}{8}.$$

Auflösen ergibt

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}.$$

13.3 Punktweise Approximation unstetiger Funktionen

Solange die Funktion $f \in \mathcal{P}$ Riemann-integrierbar ist, lässt sich ihre Fourier-Reihe $S_\infty f$ bestimmen. Insbesondere kann man sich fragen, inwiefern $S_\infty f$ die Funktion f approximiert, wenn f nicht stetig ist. Hierzu betrachten wir zunächst ein Beispiel.

Beispiel 13.13. Gegeben sei die 2π-periodische Funktion $g : \mathbb{R} \to \mathbb{R}$ mit $g(x) = 1$ für $x \in [0, \pi)$ und $g(x) = -1$ für $x \in [-\pi, 0)$. Da $g(x) = f'(x)$ für jedes $x \notin \pi\mathbb{Z}$ für f aus Beispiel 13.12, liegt folgende Vermutung nahe, die man durch gliedweises Differenzieren aus (13.6) abliest:
Vermutung: Für $x \notin \pi\mathbb{Z}$ gilt

$$g(x) = \frac{4}{\pi}\Big(\sin(x) + \frac{\sin(3x)}{3} + \frac{\sin(5x)}{5} + \cdots \Big).$$

Insbesondere gilt $S_\infty g = g$ an allen Stetigkeitsstellen von g und für $x \in \pi\mathbb{Z}$ ist

$$S_\infty g(k\pi) = 0 = \tfrac{1}{2}(g(x-) + g(x+)).$$

Ein allgemeines Resultat, das diese Vermutung unterstützt, lernen wir in Theorem 13.18 kennen.

Lemma 13.14 (Riemann'sches Lemma). *Sei $f \in \mathcal{R}(I)$ stückweise stetig. Dann gilt für $a, b \in I$*

$$\lim_{k\to\infty} \int_a^b f(x)\sin(kx)dx = 0.$$

Beweis. Sei zunächst $f = c$ konstant. Dann gilt

$$\int_a^b f(x)\sin(kx)dx = -\frac{c}{k}\cos(x)|_a^b \xrightarrow{k\to\infty} 0.$$

Diese Rechnung überträgt sich nun ganz leicht auf Treppenfunktionen f (siehe Definition 8.7). Sei also nun f stückweise stetig und $\varepsilon > 0$. Wegen Korollar 8.10 wissen wir, dass es eine Treppenfunktion φ gibt mit $||f - \varphi|| < \varepsilon$. Deshalb und wegen Theorem 8.19 gilt

$$\begin{aligned}\lim_{k\to\infty}\Big|\int_a^b f(x)\sin(kx)dx\Big| &= \lim_{k\to\infty}\Big|\int_a^b (f(x)-\varphi(x))\sin(kx)dx\Big| \\ &\le \varepsilon \lim_{k\to\infty}\int_a^b |\sin(kx)|dx \le \varepsilon(b-a).\end{aligned}$$

Mit $\varepsilon \to 0$ folgt die Behauptung. □

Um die punktweise Approximation von unstetigen Funktionen durch die Fourier-Transformierte zeigen zu können, benötigen wir nun noch den Dirichlet-Kern. Entscheidend wird hier das Lemma 13.17, das im Fall der Stetigkeit der zu approximiernden Funktion bereits eine punktweise Konvergenz liefert.

Definition 13.15 (Dirichlet-Kern). *Wir definieren den* Dirichlet-Kern *n-ten Grades*

$$D_n(x) := \frac{1}{2\pi} \sum_{k=-n}^{n} e^{ikx}.$$

Lemma 13.16. *1. Es gilt*

$$2\pi \cdot D_n(x) = \begin{cases} \frac{\sin((n+\frac{1}{2})x)}{\sin(\frac{1}{2}x)}, & x \in \mathbb{R} \setminus 2\pi\mathbb{Z}, \\ 2n+1, & x \in 2\pi\mathbb{Z}. \end{cases}$$

2. Weiter ist

$$\int_{-\pi}^{\pi} D_n(x) dx = 1.$$

Beweis. 1. Für $x \in 2\pi\mathbb{Z}$ ist die Behauptung klar. Andernfalls erinnern wir uns an die endliche geometrische Reihe aus Proposition 1.38 und schreiben

$$\begin{aligned} 2\pi D_n(x) &= e^{-inx} \sum_{k=0}^{2n} e^{ikx} = e^{-inx} \frac{e^{i(2n+1)x} - 1}{e^{ix} - 1} = e^{-inx} \frac{e^{i(2n+1/2)x} - e^{-ix/2}}{e^{ix/2} - e^{-ix/2}} \\ &= \frac{e^{i(n+1/2)x} - e^{-i(n+1/2)x}}{e^{ix/2} - e^{-ix/2}} = \frac{\sin((n+\frac{1}{2})x)}{\sin(\frac{1}{2}x)}. \end{aligned}$$

2. Wir erhalten

$$\sum_{k=-n}^{n} \int_{-\pi}^{\pi} e^{ikx} dx = 2\pi + \sum_{k=1}^{n} \int_{-\pi}^{\pi} e^{ikx} + e^{-ikx} dx = 2\pi + \frac{1}{ik}\left(e^{ikx} - e^{-ikx}\right)\Big|_{-\pi}^{\pi} = 2\pi.$$

□

Lemma 13.17 (Dirichlet'sches Lemma). *Sei $f \in \mathcal{P}$ stückweise stetig und in 0 linksseitig und rechtsseitig differenzierbar. Dann gilt*

$$\int_{-\pi}^{\pi} f(x) D_n(x) dx \xrightarrow{n\to\infty} \frac{1}{2}\big(f(0-) + f(0+)\big).$$

Beweis. Wir definieren auf $(0, \pi]$

$$\varphi(x) := \frac{f(x) - f(0+)}{\sin(\frac{x}{2})}$$

und setzen diese Funktion stetig nach 0 fort. Nun ist wegen Lemma 13.16

$$\begin{aligned} \int_0^{\pi} f(x) D_n(x) dx - \frac{1}{2} f(0+) &= \int_0^{\pi} (f(x) - f(0+)) D_n(x) dx \\ &= \int_0^{\pi} \varphi(x) \sin((n+\tfrac{1}{2})x) dx \xrightarrow{n\to\infty} 0 \end{aligned}$$

mit Lemma 13.14. Analog zeigt man $\int_{-\pi}^{0} f(x) D_n(x) dx \xrightarrow{n\to\infty} \frac{1}{2} f(0-)$, woraus die Behauptung folgt. □

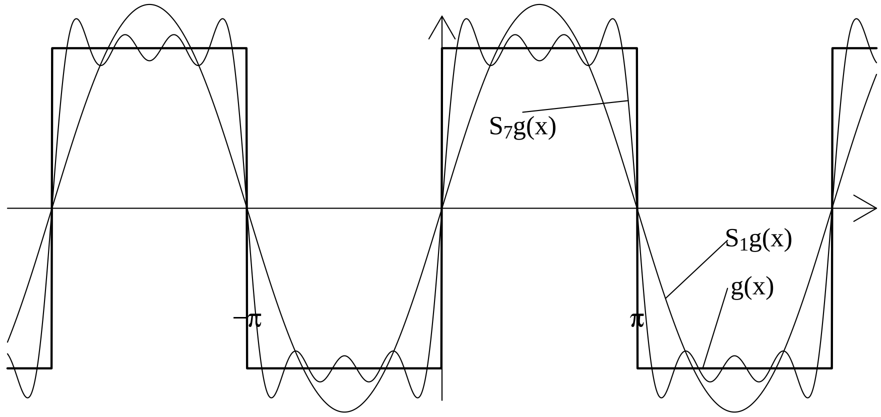

Abbildung 13.2: Die unstetige Funktion g aus Beispiel 13.19 und zwei Approximationen von ihr. Es gilt $S_n g(x) \xrightarrow{n\to\infty} S(x)$ für alle $x \in \mathbb{R} \setminus 2\pi\mathbb{Z}$.

Wir können nun unser Hauptresultat dieses Abschnittes formulieren und beweisen.

Theorem 13.18 (Punktweise Konvergenz von Fourier-Reihen). *Sei $f \in \mathcal{P}$ stückweise stetig und ist in x linksseitig und rechtsseitig differenzierbar. Dann gilt*

1. *Ist f in x stetig, so gilt $S_\infty f(x) = f(x)$.*
2. *Ist f in x unstetig, so gilt $S_\infty^f(x) = \frac{1}{2}\big(f(x+) + f(x-)\big)$.*

Beweis. Wir schreiben unter Verwendung von (13.4), Lemma 13.17 und der Periodizität von f, und da D_n eine gerade Funktion ist,

$$\begin{aligned} S_n f(x) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t) \sum_{k=-n}^{n} e^{ik(x-t)} dt = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t) D_n(t-x) dt \\ &= \frac{1}{2\pi}\int_{-\pi-x}^{\pi-x} f(t+x) D_n(t) dt = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x+t) D_n(t) dt \xrightarrow{n\to\infty} \frac{1}{2}\big(f(x+) + f(x-)\big). \end{aligned}$$

Daraus folgen alle Behauptungen. □

Beispiel 13.19. Wir kommen noch einmal auf die Funktion g aus Beispiel 13.13 zu sprechen. Wir hatten bereits eingesehen, dass

$$(S_\infty g)(x) = \frac{4}{\pi}\Big(\sin(x) + \frac{\sin(3x)}{3} + \frac{\sin(5x)}{5} + \cdots\Big).$$

Mit Theorem 13.18 folgt zumindest die Konvergenz $S_\infty g = g$ punktweise für alle $x \in \mathbb{R} \setminus 2\pi\mathbb{Z}$. Siehe auch Abbildung 13.2.

Teil IV
Mehrdimensionale Differentiation

Wir betrachten im Folgenden Abbildungen $\underline{f} : E \to \mathbb{R}^n$ für $E \subseteq \mathbb{R}^m$. Dabei sind $\mathbb{R}^m, \mathbb{R}^n$ mit der euklidischen Metrik bzw. Norm (siehe Beispiel 5.2 und 8.2) ausgestattet und wir schreiben $|\underline{x}| := ||\underline{x}||_2$. Insbesondere verwenden wir den Konvergenzbegriff aus Definition 5.3. Die Begriffe der offenen Menge, Umgebung oder kompakten Menge übernehmen wir aus Definitionen 5.7 und 19.8. Wir bezeichnen mit $\underline{e}_1, ..., \underline{e}_n$ die kanonische Basis des $\mathbb{R}^n$, also etwa $\underline{e}_j = (0, ..., 0, 1, 0, ..., 0)$. Weiter schreiben wir $\underline{0} = (0, ..., 0)$ für den Nullvektor.

14 Mehrdimensionale Ableitungen

Zentraler Punkt auf den kommenden Seiten sind Funktionen[19] $\underline{f} : E \subseteq \mathbb{R}^m \to \mathbb{R}^n$. Eine Funktion $\underline{f} : E \to \mathbb{R}^n$ ist hierbei eindeutig gegeben durch $f_1, ..., f_n : E \to \mathbb{R}$ mittels $\underline{f}(\underline{x}) := (f_1(\underline{x}), ..., f_n(\underline{x}))$. Andersherum sei

$$\pi_i : \begin{cases} \mathbb{R}^n & \to \mathbb{R} \\ (x_1, ..., x_n) & \mapsto x_i \end{cases} \tag{14.1}$$

die i-te Projektionsabbildung. Dann definiert die Funktion $\underline{f}$ die Funktionen $f_i := \pi_i \circ \underline{f} : E \to \mathbb{R}$.

14.1 Grundbegriffe

Aus der Theorie der Differentiation von Funktionen $f : I \subseteq \mathbb{R} \to \mathbb{R}$ haben wir bereits einige Erkenntnisse gewonnen, etwa:

- Es gibt einige äquivalente Formulierungen für Differenzierbarkeit (siehe Proposition 7.2), etwa ist f in $x \in \mathbb{R}$ genau dann differenzierbar mit $f'(x) = y$, wenn es eine lineare Abbildung $L : \mathbb{R} \to \mathbb{R}$ gibt mit $\frac{f(x+h)-f(x)-L(h)}{h} \xrightarrow{h\to 0} 0$, wobei L durch $L(h) = yh$ gegeben ist.
- Ist $f : I \subseteq \mathbb{R} \to \mathbb{R}$ differenzierbar, dann ist f auch stetig; siehe Korollar 7.3.

Diese Eigenschaften benötigen im mehrdimensionalen zunächst die richtige Definition von Ableitung. Wir stellen nun einige mögliche Begriffe vor.

Definition 14.1 (Ableitungsbegriffe). *Sei $E \subseteq \mathbb{R}^m$ offen und $\underline{f} : E \to \mathbb{R}^n$.*

1. *Die* i-te partielle Ableitung von $\underline{f}$ in $\underline{x} \in E$ *ist definiert als*

$$\frac{\partial \underline{f}(\underline{x})}{\partial x_i} := \frac{\partial}{\partial x_i}\underline{f}(\underline{x}) := \lim_{h\to 0} \frac{\underline{f}(\underline{x} + h\underline{e}_i) - \underline{f}(\underline{x})}{h},$$

falls der Grenzwert existiert, $i = 1, ..., m$. Weiter definieren wir zweite partiellen Ableitungen *durch*

$$\frac{\partial^2 \underline{f}(\underline{x})}{\partial x_i \partial x_j} := \frac{\partial}{\partial x_i}\frac{\partial \underline{f}(\underline{x})}{\partial x_j}, i, j = 1, ..., m.$$

[19] Wir werden konsequent $\underline{x}$ für Vektoren und $\underline{\underline{L}}$ für Matrizen schreiben.

Analog werden höhere partiellen Ableitungen $\frac{\partial^k \underline{f}(\underline{x})}{\partial x_{i_1} \cdots \partial x_{i_k}}$ *definiert.*

2. *Die* Richtungsableitung von $\underline{f}$ in Richtung $\underline{v} \in \mathbb{R}^m$ *ist definiert als*

$$\frac{\partial \underline{f}(\underline{x})}{\partial x_{\underline{v}}} := \frac{\partial}{\partial x_{\underline{v}}} \underline{f}(\underline{x}) := \lim_{h \to 0} \frac{\underline{f}(\underline{x} + h\underline{v})}{h},$$

falls der Grenzwert existiert. Insbesondere ist also

$$\frac{\partial \underline{f}(\underline{x})}{\partial x_i} = \frac{\partial \underline{f}(\underline{x})}{\partial x_{\underline{e}_i}}.$$

3. *Die lineare Abbildung* $\underline{\underline{D}}\,\underline{f}(\underline{x}) : \mathbb{R}^m \to \mathbb{R}^n$ *heißt* Ableitung von $\underline{f}$ in $\underline{x}$, *falls*

$$\lim_{\underline{h} \to 0} \frac{\underline{f}(\underline{x} + \underline{h}) - \underline{f}(\underline{x}) - (\underline{\underline{D}}\,\underline{f}(\underline{x}))(\underline{h})}{|\underline{h}|} = 0 \tag{14.2}$$

gilt. In diesem Fall heißt $\underline{f}$ in $\underline{x}$ differenzierbar.

Bemerkung 14.2 (Berechnung partieller Ableitungen). Sei $E \subseteq \mathbb{R}^m$ offen und $\underline{f} = (f_1, ..., f_n) : E \to \mathbb{R}^n$ sowie $\underline{x} = (x_1, ..., x_m)$. Definiere $g : x \mapsto \underline{f}(x_1, ..., x_{i-1}, x, x_{i+1}, ..., x_m)$. Dann gilt

$$\frac{\partial \underline{f}(\underline{x})}{\partial x_i} = \frac{dg(x_i)}{dx},$$

falls eine der beiden Seiten existiert. Das bedeutet: Die i-te partielle Ableitung von $\underline{f}$ in $\underline{x}$ erhält man durch Festhalten aller Koordinaten bis auf die i-te. Durch diese Erkenntnis übertragen sich alle Eigenschaften von Ableitungen in einer Dimension auf partielle Ableitungen. Als Beispiel betrachte man die einfache Funktion $f : (x, y) \mapsto xy$. Hier gilt offensichtlich

$$\frac{\partial f(x, y)}{\partial x} = y, \qquad \frac{\partial f(x, y)}{\partial y} = x.$$

Beispiel 14.3 (Ableitung der euklidischen Norm, harmonische Funktion). Wir betrachten die euklidische Norm

$$||.|| : \begin{cases} \mathbb{R}^m & \to \mathbb{R} \\ \underline{x} = (x_1, ..., x_m) & \mapsto |\underline{x}| := \sqrt{x_1^2 + \cdots + x_m^2}. \end{cases}$$

Wir berechnen für $\underline{x} \neq \underline{0}$

$$\frac{\partial |\underline{x}|}{\partial x_i} = \frac{2x_i}{2\sqrt{x_1^2 + \cdots + x_m^2}} = \frac{x_i}{|\underline{x}|}.$$

Hieraus ergibt sich für $j \neq i$

$$\frac{\partial^2 |\underline{x}|}{\partial x_i \partial x_j} = -\frac{x_i x_j}{|\underline{x}|^3},$$

sowie

$$\frac{\partial^2 |\underline{x}|}{\partial x_i^2} := \frac{\partial^2 |\underline{x}|}{\partial x_i \partial x_i} = \frac{|\underline{x}| - x_i^2/|\underline{x}|}{|\underline{x}|^2} = \frac{1}{|\underline{x}|}\Big(1 - \frac{x_i^2}{|\underline{x}|^2}\Big),$$

also insgesamt für alle i, j

$$\frac{\partial^2 |\underline{x}|}{\partial x_i \partial x_j} = \frac{1}{|\underline{x}|}\Big(\delta_{ij} - \frac{x_i x_j}{|\underline{x}|^2}\Big).$$

Für die Analysis ist die Lösung der Gleichung

$$\Delta f(\underline{x}) := \sum_{i=1}^{m} \frac{\partial^2 f(\underline{x})}{\partial x_i^2} = 0$$

von zentraler Bedeutung. Der Operator $\Delta := \sum_{i=1}^m \frac{\partial^2}{\partial x_i^2}$ heißt auch Laplace-Operator und Lösungen der Gleichung heißen auch harmonische Funktionen. Mit dem soeben berechneten und einer glatten Funktion $\varphi : \mathbb{R}_+ \to \mathbb{R}$ wollen wir eine harmonische Funktion der Form $f := \varphi \circ |.|$ finden. Wir schreiben für ein solches f

$$\begin{aligned}\Delta f(\underline{x}) &= \sum_{i=1}^m \frac{\partial^2}{\partial x_i^2}\varphi(|\underline{x}|) = \sum_{i=1}^m \frac{\partial}{\partial x_i}\varphi'(|\underline{x}|)\frac{x_i}{|\underline{x}|} = \sum_{i=1}^m \varphi''(|\underline{x}|)\frac{x_i^2}{|\underline{x}|^2} + \varphi'(|\underline{x}|)\frac{1}{|\underline{x}|}\Big(1 - \frac{x_i^2}{|\underline{x}|^2}\Big)\\ &= \varphi''(|\underline{x}|) + \varphi'(|\underline{x}|)\frac{m-1}{|\underline{x}|} = |\underline{x}|^{1-m}\frac{d}{dr}\Big(r^{m-1}\varphi'(r)\Big)\Big|_{r=|\underline{x}|}.\end{aligned}$$

Nun gilt also $\Delta f(\underline{x}) = 0$ genau dann, wenn

$$\frac{d}{dr}\Big(r^{m-1}\varphi'(r)\Big) = 0,$$

also muss für ein $c \in \mathbb{R}$

$$\varphi'(r) = \frac{c}{r^{m-1}},$$

d.h. für ein $d \in \mathbb{R}$

$$\varphi(r) = \int^r \frac{c}{s^{m-1}} ds + d = \begin{cases} c\log(r) + d, & m = 2,\\[2ex] \dfrac{c}{(2-m)r^{m-2}} + d, & m = 3, 4, ... \end{cases}$$

Damit haben wir also alle harmonischen Funktionen f der Form $f(\underline{x}) = \varphi(|\underline{x}|)$ bestimmt.

Beispiel 14.4 (Lineare Ableitungen sind differenzierbar). Sei $\underline{L} : \mathbb{R}^m \to \mathbb{R}^n$ linear. Dann gilt $\underline{\underline{D}}\,\underline{L}(\underline{x}) = \underline{L}$ für alle $\underline{x} \in \mathbb{R}^m$.
Denn: Es ist offensichtlich

$$\frac{\underline{L}(\underline{x}+\underline{h}) - \underline{L}(\underline{x}) - \underline{L}(\underline{h})}{|\underline{h}|} = 0.$$

Also ist $\underline{h} \mapsto \underline{L}(\underline{h})$ eine lineare Abbildung die (14.2) erfüllt, und ist damit die Ableitung von $\underline{L}$.

Proposition 14.5 (Differenzierbare Abbildungen sind stetig). *Sei $E \subseteq \mathbb{R}^m$ und $\underline{f} : E \to \mathbb{R}^n$. Ist $\underline{f}$ in $\underline{x} \in E$ differenzierbar, so ist $\underline{f}$ in $\underline{x}$ auch stetig.*

Beweis. Klar ist, dass lineare Funktionen in mehrdimensionalen Räumen stetig sind[20]. Also genügt es die Stetigkeit der Funktion $\underline{h} \mapsto \varphi(\underline{h}) := \underline{f}(\underline{x}+\underline{h}) - \underline{f}(\underline{x}) - (\underline{\underline{D}}\,\underline{f}(\underline{x}))(\underline{h})$ in $\underline{h} = \underline{0}$ zu zeigen. Es gilt aber $\varphi(\underline{0}) = 0$ und $\varphi(\underline{h}) \xrightarrow{h \to 0} 0$ wegen der Differenzierbarkeit von $\underline{f}$ in $\underline{0}$. □

Klar ist, dass partielle Ableitungen nur eine bestimmte Form der Richtungsableitung sind. Wir werden nun zeigen, dass die Differenzierbarkeit einer Funktion der stärkste Ableitungsbegriff ist, da aus ihm die Richtungsdifferenzierbarkeit (und damit auch die partielle Differenzierbarkeit) folgt. Andersherum werden wir in Theorem 14.13 zeigen, dass aus der Stetigkeit der partiellen Ableitungen die Differenzierbarkeit folgt.

Theorem 14.6 (Differenzierbar⇒Richtungsdifferenzierbar). *Sei $E \subseteq \mathbb{R}^m$ und $\underline{f} = (f_1, ..., f_n) : E \to \mathbb{R}^n$ in $\underline{x} \in E$ differenzierbar mit Ableitung $\underline{\underline{D}}\,\underline{f}(\underline{x})$. Dann existieren alle Richtungsableitungen von $\underline{f}$ in $\underline{x}$ und es gilt für $\underline{v} \in \mathbb{R}^m$*

$$\frac{\partial \underline{f}(\underline{x})}{\partial x_{\underline{v}}} = (\underline{\underline{D}}\,\underline{f}(\underline{x}))(\underline{v}).$$

Bezüglich der Standardbasis von $\mathbb{R}^m$ und $\mathbb{R}^n$ hat $\underline{\underline{D}}\,\underline{f}(\underline{x})$ die darstellende Matrix

$$\begin{pmatrix} \frac{\partial f_1(\underline{x})}{\partial x_1} & \cdots & \cdots & \frac{\partial f_1(\underline{x})}{\partial x_m} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f_n(\underline{x})}{\partial x_1} & \cdots & \cdots & \frac{\partial f_n(\underline{x})}{\partial x_m} \end{pmatrix} \in \mathbb{R}^{n \times m}. \tag{14.3}$$

Diese wird auch Jacobi-Matrix *genannt. Insbesondere ist die Ableitung eindeutig bestimmt und es gilt für $\underline{v} = (v_1, ..., v_m)$ (die Darstellung von $\underline{v}$ in der Standardbasis)*

$$\frac{\partial \underline{f}(\underline{x})}{\partial x_{\underline{v}}} = \begin{pmatrix} \frac{\partial f_1(\underline{x})}{\partial x_1} & \cdots & \cdots & \frac{\partial f_1(\underline{x})}{\partial x_m} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f_n(\underline{x})}{\partial x_1} & \cdots & \cdots & \frac{\partial f_n(\underline{x})}{\partial x_m} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix}$$

Beweis. Sei $\underline{0} \neq \underline{v} \in \mathbb{R}^m$. Dann gilt wegen der Definition der Ableitung

$$\frac{\underline{f}(\underline{h} + h\underline{v}) - \underline{f}(\underline{x})}{h} - (\underline{\underline{D}}\,\underline{f}(\underline{x}))(\underline{v}) = \frac{\underline{f}(\underline{h} + h\underline{v}) - \underline{f}(\underline{x}) - (\underline{\underline{D}}\,\underline{f}(\underline{x}))(h\underline{v})}{h} = 0.$$

Insbesondere ist also mit $h \to 0$

$$\frac{\partial \underline{f}(\underline{x})}{\partial x_{\underline{v}}} = (\underline{\underline{D}}\,\underline{f}(\underline{x}))(\underline{v}).$$

Für $1 \leq i \leq n$ und $1 \leq j \leq m$ folgt daraus mittels der Darstellung in der Standardbasis

$$\frac{\partial f_i(\underline{x})}{\partial x_j} = \big((\underline{\underline{D}}\,\underline{f}(\underline{x}))(\underline{e}_j)\big)_i = (\underline{\underline{D}}\,\underline{f}(\underline{x}))_{ij}.$$

□

[20] Man beachte, dass die hier betrachteten Vektorräume endlichdimensional sind. Im allgemeinen sind lineare Abbildungen in unendlich-dimensionalen Vektorräumen nicht stetig.

Definition 14.7 (Gradient). *Sei $E \subseteq \mathbb{R}^m$ offen und $f : E \to \mathbb{R}$ differenzierbar. Wir schreiben*

$$\mathrm{grad} f(\underline{x}) := \underline{\nabla} f(\underline{x}) := \Big(\frac{\partial f(\underline{x})}{\partial x_1}, ..., \frac{\partial f(\underline{x})}{\partial x_m}\Big).$$

Etwas informell bezeichnen wir (∇ heißt Nabla*)*

$$\underline{\nabla} = \Big(\frac{\partial}{\partial x_1}, ..., \frac{\partial}{\partial x_m}\Big).$$

Bemerkung 14.8 (Eigenschaften des Gradienten). Sei $E \subseteq \mathbb{R}^m$ offen und $f : E \to \mathbb{R}$ differenzierbar sowie $\underline{x} \in E$.

1. Es ist $\underline{\nabla} f(\underline{x})$ der eindeutig bestimmte Vektor mit

$$\langle \underline{\nabla} f(\underline{x}), \underline{v}\rangle = (\underline{D} f(\underline{x}))(\underline{v})$$

   für alle $\underline{v} \in \mathbb{R}^m$. Hierbei bezeichnet $\langle ., . \rangle$ das Standard-Skalarprodukt im $\mathbb{R}^m$.

2. $\underline{\nabla} f(\underline{x}) \in \mathbb{R}^m$ ist die Richtung des des größten Anstieges der Funktion $\underline{x} \mapsto f(\underline{x})$, also der größten Richtungsableitung für eine Richtung $\underline{v}$ mit $|\underline{v}| = 1$.
   Denn: Es gilt wegen der Cauchy-Schwartz-Ungleichung für jedes $\underline{v} \in \mathbb{R}^m$ mit $|\underline{v}| = 1$

$$\frac{\partial f(\underline{x})}{\partial x_{\underline{v}}} = \langle \underline{\nabla} f(\underline{x}), \underline{v}\rangle \leq ||\underline{\nabla} f(\underline{x})||$$

   mit Gleichheit genau dann, wenn $\underline{v}$ und $\underline{\nabla} f(\underline{x})$ linear abhängig, also parallel, sind, d.h. wenn $\underline{v} = \frac{\underline{\nabla} f(\underline{x})}{|\underline{\nabla} f(\underline{x})|}$.

Wie wir in Proposition 14.5 gesehen haben, folgt die Stetigkeit einer differenzierbaren Funktion genau wie im ein-dimensionalen Fall (siehe Korollar 7.3). Da wir im mehrdimensionalen mehrere Ableitungsbegriffe eingeführt haben, können wir uns fragen, ob diese Ableitungsbegriffe nicht zusammenfallen, d.h. ob etwa jede partiell differenzierbare Funktion schon differenzierbar ist. Dass dies nicht der Fall ist zeigt das nächste Beispiel.

Beispiel 14.9 (Eine partiell differenzierbare, aber nicht differenzierbare Funktion). Wir betrachten die Funktion (siehe auch Abbildung 14.1)

$$f : (x, y) \mapsto \begin{cases} \dfrac{xy}{x^2 + y^2}, & (x, y) \neq \underline{0}, \\ 0, & \text{sonst} \end{cases}$$

Dann ist für $x, y \in \mathbb{R}$

$$f(x, 0) = 0, \qquad f(0, y) = 0, \qquad f(x, x) = \frac{x^2}{x^2 + x^2} = \frac{1}{2}.$$

Insbesondere ist f in $\underline{x} = \underline{0}$ nicht stetig, also (nach Proposition 14.5) auch nicht differenzierbar. Allerdings lassen sich die partiellen Ableitungen berechnen, nämlich durch

$$\frac{\partial f(x, y)}{\partial x} = \frac{(x^2 + y^2)y - 2x^2 y}{(x^2 + y^2)^2} = \frac{y(x^2 - y^2)}{(x^2 + y^2)^2}$$

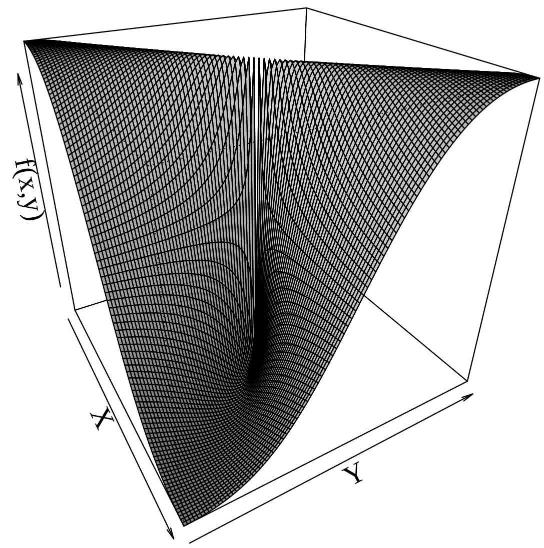

Abbildung 14.1: Die Funktion aus Beispiel 14.9.

und analog

$$\frac{\partial f(x,y)}{\partial y} = \frac{x(y^2 - x^2)}{(x^2 + y^2)^2}$$

sowie

$$\frac{\partial f(0,0)}{\partial x} = \frac{\partial f(0,0)}{\partial y} = 0.$$

Wir sehen damit, dass die Existenz der partiellen Ableitungen nicht einmal die Stetigkeit von f impliziert. Es ist nicht erstaunlich, dass auch die partiellen Ableitungen von f in $\underline{0}$ nicht stetig sind. Es ist nämlich

$$\frac{\partial f(0,y)}{\partial x} = -\frac{1}{y}, \qquad \frac{\partial f(y,y)}{\partial x} = 0.$$

Betrachten wir nun die Richtungsableitungen von f. Sei hierzu $\underline{v} := (v,w) \in \mathbb{R}^2$. Wir berechnen

$$\frac{f(hv,hw) - f(0,0)}{h} = \frac{h^2vw}{h(h^2v^2 + h^2w^2)} = \frac{1}{h}\frac{vw}{v^2 + w^2}$$

und damit existiert die Richtungsableitung in $\underline{x} = \underline{0}$ nur im Fall $v = 0$ oder $w = 0$.

Im letzten Beispiel haben wir gesehen, dass die Begriffe der partiellen Differenzierbarkeit und der Differenzierbarkeit nicht zusammenfallen. Da der Begriff der Richtungsdifferenzierbarkeit (in alle möglichen Richtungen) mehr von einer Funktion fordert, wäre es kein Widerspruch zum letzten Beispiel, wenn jede richtungsdifferenzierbare Funktion auch differenzierbare wäre. Dass dies jedoch nicht gilt, zeigt folgendes Beispiel.

Beispiel 14.10 (Eine richtungsdifferenzierbare, nicht differenzierbare Funktion). Wir betrachten die Funktion

$$f : (x,y) \mapsto \begin{cases} \dfrac{2xy^2}{x^2+y^4}, & (x,y) \neq \underline{0}, \\ 0, & \text{sonst} \end{cases}$$

Dann ist für $x, y \in \mathbb{R}$

$$f(x,0) = 0, \qquad f(0,y) = 0, \qquad f(x^2,x) = \frac{2x^4}{x^4+x^4} = 1,$$

also ist f ebenfalls in $\underline{0}$ unstetig, insbesondere also nach Proposition 14.5 nicht differenzierbar. Diesmal jedoch existieren sogar alle Richtungsableitungen von f in $\underline{0}$. Für Richtungen $\underline{v} = (0,w)$ ist dies klar. Für $(v,w) \in \mathbb{R}^2$ mit $v \neq 0$ gilt außerdem

$$\frac{f(hv,hw) - f(0,0)}{h} = \frac{2vhw^2h^2}{h(v^2h^2 + w^4h^4)} = \frac{2vw^2h^3}{v^2h^3 + w^4h^5} \xrightarrow{h \to 0} \frac{2w^2}{v}.$$

Das bedeutet, dass selbst die Existenz aller Richtungsableitungen die Stetigkeit der Funktion f nicht impliziert.

Wie die letzten Beispiele zeigen, gibt es keine Äquivalenz der drei eingeführten Ableitungebegriffe. Dies ändert sich, wenn man sich auf stetig (partiell) differenzierbare Funktionen einschränkt. Insbesondere sind solche Funktionen differenzierbar, wie Theorem 14.13 zeigen wird.

Definition 14.11 ($\mathcal{C}^k$-Räume). *Sei $E \subseteq \mathbb{R}^m$ offen. Wir definieren für $k = 0, ..., \infty$*

$$\mathcal{C}^k(E, \mathbb{R}^n) := \Big\{ \underline{f} : E \to \mathbb{R}^n : \underline{x} \mapsto \frac{\partial^j \underline{f}(\underline{x})}{\partial x_{i_1} \cdots \partial x_{i_j}} \text{ existiert und ist stetig} \\ \text{für alle } j = 0, ..., k, i_1, ..., i_j = 1, ..., m \Big\}.$$

Ist $n = 1$, so setzen wir $\mathcal{C}^k(E) := \mathcal{C}^k(E; \mathbb{R})$.

Proposition 14.12 (Differenzierbarkeit der Komponenten). *Sei $E \subseteq \mathbb{R}^m$ offen, $\underline{x} \in E$ und $\underline{f} = (f_1, ..., f_n) : E \to \mathbb{R}^n$. Dann ist $\underline{f}$ in $\underline{x}$ genau dann differenzierbar, wenn $f_1, ..., f_n$ in $\underline{x}$ differenzierbar sind.*

Beweis. Zunächst sei $\underline{f}$ in $\underline{x}$ differenzierbar. Dann gilt nach Definition

$$\frac{\underline{f}(\underline{x}+\underline{h}) - \underline{f}(\underline{x}) - (\underline{\underline{D}}\, \underline{f}(\underline{x}))(\underline{h})}{|\underline{h}|} \xrightarrow{\underline{h} \to 0} 0. \tag{14.4}$$

Die linke Seite ist hier also vektorwertig, und die Konvergenz gegen 0 ist gerade die Konvergenz aller Komponenten. Die i-te Komponente ist aber gerade

$$\frac{f_i(\underline{x}+\underline{h}) - f_i(\underline{x}) - \big((\underline{\underline{D}}\, \underline{f}(\underline{x}))(\underline{h})\big)_i}{|\underline{h}|} \xrightarrow{\underline{h} \to 0} 0.$$

Nun ist die i-te Komponente der linearen Abbildung $(\underline{\underline{D}}\,\underline{f}(\underline{x}))$ aber gerade $\underline{D}f_i(\underline{x})$, also

$$\frac{f_i(\underline{x}+\underline{h}) - f_i(\underline{x}) - (\underline{D}f_i(\underline{x}))(\underline{h})}{|\underline{h}|} \xrightarrow{\underline{h}\to 0} 0, \tag{14.5}$$

woraus folgt, dass f_i differenzierbar ist.

Andersherum seien alle $f_1, ..., f_n$ in $\underline{x}$ differenzierbar, es gilt also (14.5) für $i = 1, ..., n$. Dies bedeutet, dass alle Komponenten in (14.4) konvergieren, was dasselbe ist wie die Konvergenz in (14.4). □

Theorem 14.13 (Stetig differenzierbar und differenzierbar). *Sei $E \subseteq \mathbb{R}^m$ und $\underline{f} \in \mathcal{C}^1(E; \mathbb{R}^m)$. Dann ist $\underline{f}$ in jedem $\underline{x} \in E$ differenzierbar.*

Beweis. Wegen Proposition 14.12 genügt es den Fall $n = 1$ zu betrachten. Als Kandidat für die Ableitung $\underline{D}f(\underline{x})$ steht (14.3) zur Verfügung, die wir mit $\underline{\underline{A}}$ bezeichnen.

Für $\underline{x} \in E$ und $\underline{h} = (h_1, ..., h_m) \in \mathbb{R}^m$ setzen wir $\underline{x}_0 := \underline{x}$ und $\underline{x}_k := \underline{x} + h_1\underline{e}_1 + \cdots + h_k\underline{e}_k$, $k = 1, ..., n$.

Nach dem Mittelwertsatz, Theorem 7.14 (angewandt auf die Funktion $h \mapsto f(\underline{x}_{k-1} + h\underline{e}_k)$) gilt für ein $s \in [0, 1]$

$$f(\underline{x}_k) - f(\underline{x}_{k-1}) = f(\underline{x}_{k-1} + h_k\underline{e}_k) - f(\underline{x}_{k-1}) = \frac{\partial f(\underline{x}_k + sh\underline{e}_k)}{\partial x_k} h_k,$$

also gibt es $s_1, ..., s_n \in [0, 1]$ mit

$$\begin{aligned} \frac{|f(\underline{x}+\underline{h}) - f(\underline{x}) - \underline{\underline{A}}\underline{h}}{|\underline{h}|} &= \frac{1}{|\underline{h}|}\Big|\sum_{k=1}^m f(\underline{x}_k) - f(\underline{x}_{k-1}) - \frac{\partial f(\underline{x})}{\partial x_k} h_k\Big| \\ &\leq \sum_{k=1}^m \frac{h_k}{|\underline{h}|}\Big|\frac{\partial f(\underline{x}_{k-1} + s_k h_k \underline{e}_k)}{\partial x_k} - \frac{\partial f(\underline{x})}{\partial x_k}\Big| \end{aligned}$$

Wegen den Stetigkeit der partiellen Ableitungen und $\underline{x}_{k-1} + s_k h_k \underline{e}_k \xrightarrow{\underline{h}\to 0} \underline{x}$ folgt die Behauptung, wenn man $\underline{h} \to 0$ betrachtet. □

14.2 Rechenregeln

Im ein-dimensionalen haben wir in Proposition 7.5 die Produkt- Quotienten- und Kettenregel der Differentiation kennen gelernt. Da partielle Ableitungen genauso funktionieren wir ein-dimensionale Ableitungen mit festgehaltenen Koordinaten, gelten diese Rechenregeln sofort auch für partielle Ableitungen. Jedoch ist die Frage, inwiefern sich diese Regeln auch auf die Ableitung von Funktionen übertragen lassen. Dies werden wir nun beantworten.

Theorem 14.14 (Ableitungsregeln). *Sei $E \subseteq \mathbb{R}^m$ und $E' \subseteq \mathbb{R}^n$ offen.*

1. Linearität*: Für $\underline{f}, \underline{g} : E \to \mathbb{R}^n$ und $\alpha, \beta \in \mathbb{R}$ gilt, falls $\underline{\underline{D}}\,\underline{f}(\underline{x})$ und $\underline{\underline{D}}\,\underline{g}(\underline{x})$ existieren*

$$\underline{\underline{D}}\,(\alpha\underline{f} + \beta\underline{g})(\underline{x}) = \alpha\underline{\underline{D}}\,\underline{f}(\underline{x}) + \beta\underline{\underline{D}}\,\underline{g}(\underline{x}).$$

2. Produktregel*: Für $f, g : E \to \mathbb{R}$ gilt, falls $\underline{D}f(\underline{x}), \underline{D}g(\underline{x})$ existieren*

$$\underline{D}\,(fg)(\underline{x}) = g(\underline{x}) \cdot \big(\underline{D}\,f(\underline{x})\big) + f(\underline{x}) \cdot \big(\underline{D}\,g(\underline{x})\big).$$

3. Quotientenregel*: Für $f, g : E \to \mathbb{R}$ gilt, falls $\underline{D}f(\underline{x}), \underline{D}g(\underline{x})$ existieren und $g(\underline{x}) \neq 0$*

$$\underline{\underline{D}}\frac{f}{g}(\underline{x}) = \frac{\big(\underline{\underline{D}}\,\underline{f}(\underline{x})\big)\underline{D}g(\underline{x}) - \underline{f}(\underline{x})}{g(\underline{x})^2}.$$

4. Kettenregel*: Für $\underline{f} : E \to \mathbb{R}^n, \underline{g} : E' \to \mathbb{R}^p$ und $\underline{f}(E) \subseteq E'$ gilt, falls $\underline{\underline{D}}\,\underline{f}(\underline{x})$ und $\underline{\underline{D}}\,\underline{g}(\underline{f}(\underline{x}))$ existieren*

$$\underline{\underline{D}}\,(\underline{g} \circ \underline{f})(\underline{x}) = \underline{\underline{D}}\,\underline{g}(\underline{f}(\underline{x})) \cdot \underline{\underline{D}}\,\underline{f}(\underline{x}).$$

Bemerkung 14.15 (Partielle Ableitungen bei der Kettenregel). Seien f, g wie bei der Kettenregel. Dann gilt offenbar

$$\frac{\partial(\underline{g} \circ \underline{f})_i(\underline{x})}{\partial x_j} = \sum_{k=1}^{n} \frac{\partial g_i(\underline{f}(\underline{x}))}{\partial x_k} \frac{\partial f_k(\underline{x})}{\partial x_j}.$$

Beweis von Theorem 14.14. Bei der Linearität ist nichts zu zeigen. Wir zeigen zunächst die Kettenregel und schreiben $\underline{\underline{B}} := \underline{\underline{D}}\,\underline{g}(\underline{f}(\underline{x})), \underline{\underline{A}} := \underline{\underline{D}}\,\underline{f}(\underline{x})$. Definiere

$$\varepsilon_{\underline{f}}(\underline{h}) := \frac{\underline{f}(\underline{x} + \underline{h}) - \underline{f}(\underline{x}) - \underline{\underline{A}}\,\underline{h}}{|\underline{h}|}, \qquad \varepsilon_{\underline{g}}(\underline{h}) := \frac{\underline{g}(\underline{f}(\underline{x}) + \underline{h}) - \underline{g}(\underline{f}(\underline{x})) - \underline{\underline{B}}\,\underline{h}}{|\underline{h}|},$$

und stellen fest dass $\varepsilon_{\underline{f}}(\underline{h}) \xrightarrow{h\to 0} 0$ sowie $\varepsilon_{\underline{g}}(\underline{h}) \xrightarrow{h\to 0} 0$ aufgrund der Voraussetzungen. Nun gilt mit $\underline{h}' = \underline{\underline{A}}\,\underline{h} + |\underline{h}|\varepsilon_{\underline{f}}(\underline{h})$

$$\begin{aligned}
&\frac{\underline{g}(\underline{f}(\underline{x} + \underline{h})) - \underline{g}(\underline{f}(\underline{x})) - \underline{\underline{B}}\,\underline{\underline{A}}h}{|\underline{h}|} \\
&\quad= \frac{\underline{g}(\underline{f}(\underline{x}) + \underline{\underline{A}}\,\underline{h} + |\underline{h}|\varepsilon_{\underline{f}}(\underline{h})) - \underline{g}(\underline{f}(\underline{x})) - \underline{\underline{B}}\,\underline{\underline{A}}h}{|\underline{h}|} \\
&\quad= \frac{\underline{g}(\underline{f}(\underline{x}) + \underline{h}') - \underline{g}(\underline{f}(\underline{x})) - \underline{\underline{B}}\,\underline{\underline{A}}h}{|\underline{h}|} \\
&\quad= \frac{\underline{g}(\underline{f}(\underline{x})) + \underline{\underline{B}}\,\underline{h}' + |\underline{h}'|\varepsilon_{\underline{g}}(\underline{h}') - \underline{g}(\underline{f}(\underline{x})) - \underline{\underline{B}}\,\underline{\underline{A}}h}{|\underline{h}|} \\
&\quad= \underline{\underline{B}}\,\varepsilon_{\underline{f}}(\underline{h}) + \frac{|\underline{h}'|}{|\underline{h}|}\varepsilon_{\underline{g}}(\underline{h}').
\end{aligned}$$

Da $\frac{|\underline{\underline{A}}\,\underline{h}|}{|\underline{h}|}$ beschränkt ist, konvergiert die rechte Seite gegen 0, was die Behauptung zeigt.

Wir kommen nun zur Produktregel. Wir definieren $\underline{h} : E \to \mathbb{R}^2$ mittels $\underline{h}(\underline{x}) := (f(\underline{x}), g(\underline{x}))$ und $\varphi : (x, y) \mapsto xy$. Dann ist $f(\underline{x})g(\underline{x}) = \varphi(\underline{h}(\underline{x}))$. Damit gilt also nach der Kettenregel

$$\underline{D}(fg)(\underline{x}) = \underline{D}(\varphi \circ \underline{h})(\underline{x}) = \underline{D}\,\varphi(\underline{h}(\underline{x})) \cdot \underline{\underline{D}}\,\underline{h}(\underline{x}).$$

Nun ist

$$\begin{aligned}
\underline{D}\varphi(x, y) &= \big(y, x\big), \text{d.h. } \underline{D}\varphi(\underline{h}(\underline{x})) = \big(g(\underline{x}), f(\underline{x})\big), \\
\underline{\underline{D}}\,\underline{h}(\underline{x}) &= \begin{pmatrix} \frac{\partial f(\underline{x})}{\partial x_1} & \dots & \frac{\partial f(\underline{x})}{\partial x_m} \\ \frac{\partial g(\underline{x})}{\partial x_1} & \dots & \frac{\partial g(\underline{x})}{\partial x_m} \end{pmatrix}.
\end{aligned}$$

Damit gilt

$$\underline{D}(fg)(\underline{x}) = \big(g(\underline{x})\frac{\partial f(\underline{x})}{\partial x_i} + f(\underline{x})\frac{\partial g(\underline{x})}{\partial x_i}\big)_{i=1,\dots,m} = g(\underline{x}) \cdot \big(\underline{D}\, f(\underline{x})\big) + f(\underline{x}) \cdot \big(\underline{D}\, g(\underline{x})\big).$$

Wegen der eben bewiesenen Produktregel genügt es für die Quotientenregel

$$\underline{D}\frac{1}{g(\underline{x})} = -\frac{\underline{D}g(\underline{x})}{g(\underline{x})^2}$$

zu zeigen. Hierfür setzen wir $\psi(y) = 1/y$ und verwenden die Kettenregel mit

$$\underline{D}\frac{1}{g(\underline{x})} = \underline{D}(\psi \circ g(\underline{x})) = D\psi(g(\underline{x})) \cdot \underline{D}g(\underline{x}) = -\frac{\underline{D}g(\underline{x})}{g(x)^2}.$$

□

Für zweite (partielle) Ableitungen gibt es eine schöne Rechenregel, die zumindest bei stetig partiell differenzierbaren Funktionen besagt, dass die Reihenfolge der Differentiation keine Rolle spielt.

Theorem 14.16 (Satz von Schwartz). *Sei $E \subseteq \mathbb{R}^m$ offen und $f \in \mathcal{C}^2(E)$. Dann gilt*

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}.$$

Beweis. Zunächst bemerken wir

$$\frac{\partial^2 f(\underline{x})}{\partial x_i \partial x_j} = \lim_{h\to 0}\lim_{\ell \to 0} \frac{f(\underline{x} + h\underline{e}_i + \ell \underline{e}_j) - f(\underline{x} + h\underline{e}_i) - f(\underline{x} + \ell\underline{e}_j) + f(\underline{x}))}{h\ell}.$$

Die Behauptung des Theorems ist nun offensichtlich, dass wir auf der rechten Seite die Grenzwerte vertauschen dürfen. Wir schreiben mit dem Mittelwertsatz, Theorem 7.14, für geeignete $\alpha, \beta, \gamma \in [0,1]$, die jeweils von h und ℓ abhängen dürfen,

$$\begin{aligned}
&\frac{f(\underline{x} + h\underline{e}_i + \ell \underline{e}_j) - f(\underline{x} + h\underline{e}_i) - f(\underline{x} + \ell\underline{e}_j) + f(\underline{x}))}{h\ell} \\
&= \frac{1}{h}\Big(\frac{\partial f(\underline{x} + h\underline{e}_i + \alpha\ell\underline{e}_j)}{\partial x_j} - \frac{\partial f(\underline{x} + \beta\ell\underline{e}_j)}{\partial x_j}\Big) \\
&= \frac{\partial}{\partial x_j}\frac{1}{h}\Big(f(\underline{x} + \alpha\ell\underline{e}_j + h\underline{e}_i) - f(\underline{x} + \alpha\ell\underline{e}_j)\Big) + \frac{1}{h}\Big(\frac{\partial f(\underline{x} + \alpha\ell\underline{e}_j)}{\partial x_j} - \frac{\partial f(\underline{x} + \beta\ell\underline{e}_j)}{\partial x_j}\Big) \\
&= \Big(\frac{\partial}{\partial x_j}\frac{\partial}{\partial x_i} f(\underline{x} + \alpha\ell\underline{e}_i + \gamma h\underline{e}_j)\Big) + \frac{1}{h}\Big(\frac{\partial f(\underline{x} + \alpha\ell\underline{e}_j)}{\partial x_j} - \frac{\partial f(\underline{x} + \beta\ell\underline{e}_j)}{\partial x_j}\Big).
\end{aligned}$$

Die Behauptung folgt nun, indem wir zunächst $\ell \to 0$ und dann $h \to 0$ betrachten. □

Korollar 14.17. *Sei $E \subseteq \mathbb{R}^m$ offen und $f \in \mathcal{C}^k(E)$ für ein $k \geq 2$. Dann gilt für $j = 2, ..., k$ und jede Permutation σ von $\{1, ..., j\}$*

$$\frac{\partial^j f}{\partial x_{i_1} \cdot \partial x_{i_j}} = \frac{\partial^j f}{\partial x_{i_{\sigma(1)}} \cdot \partial x_{i_{\sigma(j)}}}.$$

Beweis. Klar durch iteriertes Anwenden von Theorem 14.16. Schließlich wird jede Permutation σ durch hintereinander ausgeführtes Anwenden von paarweisen Vertauschungen erzeugt. □

14.3 Extrema

Im Ein-dimensionalen (d.h. für Funktionen $f : I \to \mathbb{R}$ für ein Intervall I) ist wohlbekannt: Hat eine (glatte) Funktion $f : I \to \mathbb{R}$ für ein $x \in I$ die Eigenschaften $f'(x) = 0, f''(x) > 0$, so hat f in x ein Minimum (und ein Maximum, falls $f''(x) < 0$). Die Aufgabe dieses Abschnittes ist es, die Aussage für glatte Funktionen $f : E \subseteq \mathbb{R}^m \to \mathbb{R}$ zu verallgemeinern. Wir erinnern daran, dass wir bereits in Definition 7.11 definiert hatten, was ein lokales Minimum/Maximum ist. Zunächst zeigen wir die zu Lemma 7.12 äquivalente Aussage.

Lemma 14.18 (Gradient und Extrema). *Sei $E \subseteq \mathbb{R}^m$ offen und $f : E \to \mathbb{R}$ differenzierbar in $\underline{x} \in E$. Falls f in $\underline{x}$ ein lokales Extremum hat, so ist $\underline{D}f(\underline{x}) = 0$.*

Beweis. Hat f in $\underline{x}$ ein Extremum, so hat jede der Abbildungen $g : t \mapsto f(\underline{x} + t\underline{v})$ in $t = 0$ ein Extremum, $\underline{v} \in \mathbb{R}^m$. Weiter gilt mit Lemma 7.12

$$0 = \frac{dg(t)}{dt}\Big|_{t=0} = \frac{\partial f(\underline{x})}{\partial x_{\underline{v}}} = \underline{D}f(\underline{x}) \cdot \underline{v}.$$

Da dies für alle $\underline{v}$ erfüllt sein muss, ist notwendigerweise $\underline{D}f(\underline{x}) = 0$. □

Wir benötigen noch die zur zweiten Ableitung äquivalente Abbildung im mehrdimensionalen.

Definition 14.19 (Hesse-Matrix). *Sei $E \subseteq \mathbb{R}^m$ offen und $f \in \mathcal{C}^2(E)$. Die zweite Ableitung von f in $\underline{x} \in E$ bezeichnen wir mit*

$$\underline{\underline{D}}^2 f(\underline{x}) : \begin{cases} \mathbb{R}^m \times \mathbb{R}^m & \to \mathbb{R} \\ (\underline{v}, \underline{w}) & \mapsto \displaystyle\sum_{i,j=1}^{m} \frac{\partial^2 f(\underline{x})}{\partial x_i \partial x_j} v_i w_j. \end{cases}$$

Die Matrix $\left(\dfrac{\partial^2 f(\underline{x})}{\partial x_i \partial x_j}\right)_{i,j=1,\ldots,m}$ *heißt* Hesse-Matrix *von f in $\underline{x}$ und die quadratische Form $\underline{v} \mapsto \big(\underline{\underline{D}}^2 f(\underline{x})\big)(\underline{v}, \underline{v})$ heißt* Hesse-Form *von f in $\underline{x}$.*

Bemerkung 14.20 (Symmetrie der Hesse-Matrix). Die Hesse-Matrix von f in $\underline{x}$ ist wegen Theorem 14.16 symmetrisch. Es gilt nämlich

$$\underline{\underline{D}}^2 f(\underline{x})(\underline{v}, \underline{w}) = \sum_{i,j=1}^{m} \frac{\partial^2 f(\underline{x})}{\partial x_i \partial x_j} v_i w_j = \sum_{i,j=1}^{m} \frac{\partial^2 f(\underline{x})}{\partial x_j \partial x_i} v_i w_j = \underline{\underline{D}}^2 f(\underline{x})(\underline{w}, \underline{v}).$$

Wir wollen nun formulieren, wann eine Funktion $f : E \to \mathbb{R}$ im mehrdimensionalen ein lokales Minimum (bzw. Maximum) besitzt; hierzu benötigen wir noch die Begriffe der positiv (semi-)definiten quadratischen Form und des Kurvenintegrals.

Definition 14.21 (Positiv (semi-)definite quadratische Form). *Sei*

$$b : \begin{cases} \mathbb{R}^m \times \mathbb{R}^m & \to \mathbb{R} \\ (\underline{v}, \underline{w}) & \mapsto b(\underline{v}, \underline{w}) \end{cases}$$

eine symmetrische Bilinearform. (Das bedeutet $b(\underline{v}, \underline{w}) = b(\underline{w}, \underline{v})$ und $b(\alpha\underline{v} + \beta\underline{v}', \underline{w}) = \alpha b(\underline{v}, \underline{w}) + \beta b(\underline{v}', \underline{w})$.) Dann heißt b

$$\textit{positiv semi-definit, falls } b(\underline{v}, \underline{v}) \geq 0 \textit{ für alle } \underline{0} \neq \underline{v} \in \mathbb{R}^m,$$
$$\textit{positiv definit, falls } b(\underline{v}, \underline{v}) > 0 \textit{ für alle } \underline{0} \neq \underline{v} \in \mathbb{R}^m.$$

Bemerkung 14.22 (Hurwitz-Kriterium). Sei $\underline{\underline{A}}^{m\times m}$ eine symmetrische Matrix. Dann definiert

$$b : (\underline{v}, \underline{w}) \mapsto \underline{v}\underline{\underline{A}}\,\underline{w}^\top$$

eine symmetrische Bilinearform. Folgende Aussagen sind nach dem Hurwitz-Kriterium (das wir hier nicht zeigen werden) äquivalent:

1. b ist positiv definit.
2. Alle Eigenwerte von $\underline{\underline{A}}$ sind positiv.
3. Es gilt $\det((a_{ij})_{i,j=1,...,k}) > 0$ für alle $k = 1, ..., m$.

Lemma 14.23 (Ableitungen von Kurvenintegralen). *Sei $E \subseteq \mathbb{R}^m$ offen, $f \in \mathcal{C}^2(E)$, $I = [a, b]$ ein Intervall und $\underline{\gamma} \in \mathcal{C}^2(I; E)$. Dann gilt*

$$\begin{aligned}(f \circ \underline{\gamma})'(t) &= \underline{D}f(\gamma(t))D\underline{\gamma}(t),\\ (f \circ \underline{\gamma})''(t) &= \underline{\underline{D}}^2 f(\gamma(t))(D\underline{\gamma}(t), D\underline{\gamma}(t)) + \underline{D}f(\underline{\gamma}(t))D^2\underline{\gamma}(t).\end{aligned}$$

Beweis. Nach der Kettenregel ist

$$(f \circ \underline{\gamma})'(t) = \underline{D}f(\gamma(t))D\underline{\gamma}(t) = \sum_{j=1}^m \frac{\partial f(\gamma(t))}{\partial x_j}\gamma_j'(t).$$

Daraus folgt, wieder mit der Kettenregel

$$(f \circ \underline{\gamma})''(t) = \sum_{i,j=1}^m \frac{\partial^2 f(\gamma(t))}{\partial x_i \partial x_j}\gamma_i'(t)\gamma_j'(t) + \sum_{j=1}^m \frac{\partial f(\gamma(t))}{\partial x_j}\gamma_j''(t).$$

□

Lemma 14.24 (Mehr-dimensionales Taylor-Polynom zweiter Ordnung). *Sei $E \subseteq \mathbb{R}^m$ offen und $f \in \mathcal{C}^2(E)$. Dann gilt*

$$\frac{f(\underline{x} + \underline{h}) - f(\underline{x}) - \underline{D}f(\underline{x})\underline{h} - \frac{1}{2}\underline{\underline{D}}^2 f(\underline{x})(\underline{h}, \underline{h})}{|\underline{h}|^2} \xrightarrow{\underline{h}\to 0} 0.$$

Bemerkung 14.25. Für eine Funktion $f \in \mathcal{C}^2(I)$ für ein Intervall I folgt die Behauptung aus Theorem 12.17. Es gilt nämlich für $n = 1$

$$f(x + h) - f(x) - f'(x)h - \tfrac{1}{2}f''(x)h^2 = h^2 \int_0^1 (1 - t)(f''(x + th) - f''(x))dt \xrightarrow{h\to 0} 0,$$

weil $\sup_{t\in[0,1]} |f''(x + th) - f''(x)| \xrightarrow{h\to 0} 0$.

Beweis. Wir verwenden die Funktion $\underline{\gamma} : [0, 1] \to \mathbb{R}^m$, $\underline{\gamma}(t) = \underline{x} + t\underline{h}$ und $\varphi = f \circ \gamma : t \mapsto f(\underline{x} + t\underline{h})$. Dann gilt mit partieller Integration, angewendet mit $u(t) = \varphi'(t) - \varphi'(0)$ und $v(t) = 1$

$$\varphi(1) - \varphi(0) - \varphi'(0) = \int_0^1 (\varphi'(t) - \varphi'(0))dt = \varphi'(1) - \varphi'(0) - \int_0^1 t\varphi''(t)dt = \int_0^1 (1 - t)\varphi''(t)dt.$$

Mit Lemma 14.23 folgt daraus

$$\begin{aligned} f(\underline{x}+\underline{h}) - f(\underline{x}) - \underline{D}f(\underline{x})\underline{h} - \tfrac{1}{2}\underline{\underline{D}}^2 f(\underline{x})(\underline{h},\underline{h}) \\ = \int_0^1 (1-t)\big(\underline{\underline{D}}^2 f(\underline{x}+t\underline{h})(\underline{h},\underline{h}) - \underline{\underline{D}}^2 f(\underline{x})(\underline{h},\underline{h})\big)\,dt. \end{aligned}$$

Wegen der gleichmäßigen Stetigkeit von $t \mapsto \underline{\underline{D}}^2 f(\underline{x}+t\underline{h})(\underline{h},\underline{h})$ gibt es für jedes $\varepsilon > 0$ ein $\delta > 0$, so dass $|\underline{\underline{D}}^2 f(\underline{x}+\underline{y}) - \underline{\underline{D}}^2 f(\underline{x})| < \varepsilon$ für $\underline{y}| < \delta$. Daraus folgt für $|\underline{h}| < \delta$, dass

$$|\underline{\underline{D}}^2 f(\underline{x}+t\underline{h})(\underline{h},\underline{h}) - \underline{\underline{D}}^2 f(\underline{x})(\underline{h},\underline{h})| < \varepsilon|\underline{h}|^2.$$

Da $\varepsilon > 0$ beliebig war folgt die Behauptung. □

Theorem 14.26 (Lokale Extrema und die Hesse-Matrix). *Sei $E \subseteq \mathbb{R}^m$ offen und $f \in \mathcal{C}^2(E)$. Dann gilt:*

1. *Hat f in $\underline{x} \in E$ ein lokales Minimum, so ist $\underline{D}f(\underline{x}) = 0$ und $\underline{\underline{D}}^2 f(\underline{x})$ positiv semidefinit.*
2. *Ist $\underline{D}f(\underline{x}) = 0$ und $\underline{\underline{D}}^2 f(\underline{x})$ positiv definit für ein $\underline{x} \in E$, so hat f in $\underline{x}$ ein lokales Minimum.*

Beweis. 1. Die erste Behauptung hatten wir schon in Lemma 14.18 formuliert. Weiter hat die Abbildung $t \mapsto f(\underline{x}+t\underline{v})$ für jedes $\underline{v} \in \mathbb{R}^m$ in $t = 0$ ein lokales Minimum. Also folgt aus Lemma 14.23

$$0 \le \frac{d^2}{dt^2} f(\underline{x}+t\underline{v}) = \underline{\underline{D}}^2 f(\underline{x}+t\underline{v})(\underline{v},\underline{v}).$$

Insbesondere folgt daraus, dass die Hesse-Matrix in $\underline{x}$ positiv semidefinit ist.
2. Die Menge $A = \{\underline{v} : |\underline{v}| = 1\}$ ist kompakt und die Abbildung $\underline{v} \mapsto \underline{\underline{D}}^2 f(\underline{v},\underline{v})$ ist stetig. Nach Korollar 5.26 existiert also ein $\underline{w} \in A$ mit $\lambda := \inf_{\underline{v}\in A} \underline{\underline{D}}^2 f(\underline{v},\underline{v}) = \underline{\underline{D}}^2 f(\underline{w},\underline{w}) > 0$. Nach Lemma 14.24 gibt es dann wegen $\underline{D}f(\underline{x}) = 0$ für jedes $\varepsilon > 0$ ein $\delta > 0$, so dass für $|\underline{h}| \le \delta$

$$\frac{f(\underline{x}+\underline{h}) - f(\underline{x})}{|\underline{h}|^2} \ge \tfrac{1}{2|\underline{h}|^2}\underline{\underline{D}}^2 f(\underline{x})(\underline{h},\underline{h}) - \varepsilon \ge \tfrac{1}{2}\inf_{\underline{v}\in A} \underline{\underline{D}}^2 f(\underline{x})(\underline{v},\underline{v}) - \varepsilon = \frac{\lambda}{2} - \varepsilon.$$

Wählt man nun etwa $\varepsilon = \lambda/4$ folgt die Behauptung. □

Beispiel 14.27 (Euklidischer Abstand). Wir betrachten die Funktion $f := |.|^2 : \underline{x} = (x_1, ..., x_n) \mapsto \frac{1}{2}(x_1^2 + \cdots + x_n^2)$. Wir wollen nun mittels des Kriteriums aus Theorem 14.26 die wohlbekannte Tatsache nachweisen, dass diese Funktion in $(x, y) = 0$ ein Minimum besitzt. Hierzu berechnen wir wie in Beispiel 14.3

$$\begin{aligned} \underline{D}f(\underline{x}) &= \underline{x}, \\ \underline{\underline{D}}^2 f(\underline{x}) &= \big(\delta_{ij}\big)_{i,j=1,\ldots,m}. \end{aligned}$$

Klar ist, dass $\underline{D}f(\underline{0}) = 0$ und

$$\underline{\underline{D}}^2 f(\underline{0})(\underline{v},\underline{v}) = \sum_{i=1}^m v_i^2.$$

Damit erfüllen f und $\underline{x} = \underline{0}$ die Voraussetzungen von Theorem 14.26.2 und f hat in $\underline{0}$ ein lokales Minimum.

15 Gewöhnliche Differentialgleichungen

Will man eine sich mit der Zeit t verändernde Größe (nennen wir sie $x(t)$) mit ihrer eigenen Änderung (also $x'(t)$) in Beziehung setzen, gelangt man zu einer Differentialgleichung. Solche Gleichungen treten in vielen Anwendungen, etwa in Physik, Biologie, Chemie etc. auf. Wir formulieren ein einfaches Beispiel aus der Chemie. Betrachten wir die chemische Reaktion

$$A + B \stackrel{\kappa}{\to} C,$$

die soviel bedeutet, dass ein Molekül A mit einem Molekül B mit Rate κ (was der Geschwindigkeit der Rekation entspricht) zu einem Molekül C reagiert. Betrachten wir nun etwa die Konzentrationen der Substanz A, B, C und bezeichnen diese zu einem Zeitpunkt t mit $x_A(t), x_B(t)$ und $x_C(t)$. Je größer $x_A(t)$ und $x_B(t)$ sind, desto schneller wächst $x_C(t)$. Nach dem Massenwirkungsgesetz entwickeln sich die Größen x_A, x_B und x_C gemäß

$$\begin{pmatrix} x'_A(t) \\ x'_B(t) \\ x'_C(t) \end{pmatrix} = \begin{pmatrix} -\kappa x_A(t) x_B(t) \\ -\kappa x_A(t) x_B(t) \\ \kappa x_A(t) x_B(t) \end{pmatrix} \qquad \text{oder kürzer} \qquad \frac{d}{dt}\begin{pmatrix} x_A \\ x_B \\ x_C \end{pmatrix} = \begin{pmatrix} x'_A \\ x'_B \\ x'_C \end{pmatrix} = \begin{pmatrix} -\kappa x_A x_B \\ -\kappa x_A x_B \\ \kappa x_A x_B \end{pmatrix}.$$

Die dritte Gleichung interpretiert man in etwa so: Die Änderung der Konzentration der Substanz C steigt proportional zur Rekationsrate, sowie den Konzentrationen der Substanzen A und B.

In diesem Abschnitt beschäftigen wir uns mit solchen Differentialgleichungen, deren Lösungen also Funktionen sind. Dabei sind wichtige Fragen – wie etwa auch beim Lösen von linearen Gleichungssystemen – die Existenz und Eindeutigkeit der Lösung, sowie spezielle Lösungsverfahren. Nach einer Einführung in Abschnitt 15.1 und ein paar Hilfsmitteln in Abschnitt 15.2 werden wir in Abschnitt 15.3 Eindeutigkeits- und Existenzaussagen beweisen. Spezielle Differentialgleichungen sind linear, für die wir in Abschnitt 15.4 auch explizite Lösungsmethoden kennen lernen werden. In Abschnitt 15.5 werden wir dann in speziellen Fällen weitere Lösungsmöglichkeiten kennen lernen.

15.1 Einführung

Ziel des Abschnittes ist es lediglich, einen allgemeinen Rahmen zur Formulierung von (gewöhnlichen) Differentialgleichungen zu geben. Für eine Funktion $\underline{x} : J \subseteq \mathbb{R} \to \mathbb{R}^n$ bezeichnen wir die $n \times 1$-Matrix $\underline{\underline{D}}\,\underline{x}(t)$ im Folgenden auch mit $\underline{x}'(t)$.[21]

Definition 15.1 (Gewöhnliche Differentialgleichung, Anfangswertproblem). *Sei $I \subseteq \mathbb{R}$ ein Intervall, $E \subseteq I \times \mathbb{R}^m$ offen und $f \in \mathcal{C}^0(E, \mathbb{R}^n)$. (Hier und im Folgenden schreiben wir für $(t, \underline{x})$ mit $t \in \mathbb{R}, \underline{x} \in \mathbb{R}^m$ hierfür $f : (t, \underline{x}) \mapsto \underline{f}(t, \underline{x})$.) Dann heißt $x \in \mathcal{C}^1(J, \mathbb{R}^n)$ für ein Intervall $J \subseteq I$ Lösung der (gewöhnlichen) Differentialgleichung (für $\underline{f}$ in J), falls*

$$\underline{x}' = f(., \underline{x}), \;\; d.h. \;\; \underline{x}'(t) = f(t, \underline{x}(t)) \text{ für alle } t \in J.$$

Ist außerdem $(t_0, \underline{x}_0) \in E$ mit $t_0 \in J$, so heißt $\underline{x} \in \mathcal{C}^1(J, \mathbb{R}^n)$ Lösung des $(\underline{f}, J, (t_0, \underline{x}_0))$-Anfangswertproblems, falls $\underline{x}$ Lösung der Differentialgleichung für f in J ist und $\underline{x}(t_0) = \underline{x}_0$ gilt.

[21] In der Physik ist es außerdem üblich, Ableitungen nach der Zeit mit einem Punkt statt einem Strich zu schreiben, also wäre hier $\dot{\underline{x}}(t) := \underline{\underline{D}}\,\underline{x}(t)$.

Bemerkung 15.2 (Notation). Um Schreibarbeit zu sparen, sollen im Folgenden Mengen $J, I \subseteq \mathbb{R}$ immer Intervalle sein.

Beispiel 15.3 (Exponentielles Wachstum). Gegeben sei $f : \mathbb{R} \times \mathbb{R}$ durch $f(t,x) := \alpha x$ (insbesondere hängt also f nicht von t ab). Dann ist bekanntlich $x(t) := e^{\alpha t}$ Lösung der Differentialgleichung $x' = f(t,x) = \alpha x$. Für dieses x gilt nämlich

$$x'(t) = \alpha e^{\alpha t} = \alpha x(t).$$

Sei weiter $t_0 \in \mathbb{R}$ und $x_0 \in \mathbb{R}$. Dann ist $x(t) = x_0 e^{\alpha(t-t_0)}$ Lösung des $(f, \mathbb{R}, (t_0, x_0))$-Anfangswertproblems. Schließlich gilt mit dieser Wahl von x

$$x(0) = x_0 e^{\alpha 0} = x_0 \text{ und } x'(t) = \alpha x_0 e^{\alpha(t-t_0)} = \alpha x(t).$$

Bemerkung 15.4 (Differentialgleichungen höherer Ordnung). Allgemeiner als in Definition 15.1 kann man auch Differentialgleichungen höherer Ordnung betrachten. Sei hierzu $n \in \mathbb{N}$, $E \subseteq I \times \mathbb{R}^n$ offen und $f \in \mathcal{C}^0(E)$. Dann ist $x \in \mathcal{C}^n(J)$ mit $J \subseteq I$ Lösung der Differentialgleichung $x^{(n)} = f(., x, x^{(1)}, ..., x^{(n-1)})$ in J, falls

$$x^{(n)}(t) = f(t, x(t), x'(t), ..., x^{(n-1)}(t)), \qquad t \in J. \tag{15.1}$$

(Ein Beispiel ist die Schwingungs-Differentialgleichung $x'' = -x$ aus der Physik.) Eine solche Differentialgleichung n-ter Ordnung kann man jedoch auch in eine Differentialgleichung erster Ordnung im Sinne von Definition 15.1 umschreiben. Hierzu definieren wir die Funktion $\underline{f} : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ mittels

$$\underline{f}(t, x_0, ..., x_{n-1}) = \begin{pmatrix} x_1 \\ \cdots \\ x_{n-1} \\ f(t, x_0, ..., x_{n-1}) \end{pmatrix}.$$

Dann ist $x \in \mathcal{C}^n(J)$ genau dann Lösung von (15.1) in J, wenn $t \mapsto \underline{x}(t) := (x(t), x'(t), ..., x^{(n-1)}(t))$ Lösung von $\underline{x}' = \underline{f}(., \underline{x})$ in J ist.

Für das entsprechende Anfangswertproblem sei $(t_0, x_0, ..., x_{n-1}) \in E$ mit $t_0 \in J$. Dann heißt $x \in \mathcal{C}^n(J)$ Lösung des $(f, J, (t_0, x_0, ..., x_{n-1}))$-Anfangswertproblems in J, falls zusätzlich $(x_0(t_0) = x_0, x_0^{(1)}(t_0) = x_1, ..., x^{(n-1)}(t) = x_{n-1}$. Man sieht also, dass man bei einem Anfangswertproblem n-ter Ordnung neben dem Startwert in t_0 die ersten $n-1$ Ableitungen in t_0 festlegt.

Bemerkung 15.5 (Partielle Differentialgleichung). Differentialgleichungen im Sinne von Definition 15.1 (oder auch im Sinne von Bemerkung 15.4) heißen *gewöhnliche* Differentialgleichungen: Die Lösungsfunktion einer solchen gewöhnlichen Differentialgleichung hängt nur von einer einzigen Variable (die wir mit t bezeichnen) ab. Anders verhält es sich bei *partiellen* Differentialgleichungen, deren Lösung eine Funktion von zwei oder mehr Variablen ist. Sei hierzu z.B. $f : E \to \mathbb{R}$ mit $E \subseteq \mathbb{R}^8$. Gesucht ist für eine offene Menge $E' \subseteq \mathbb{R}^2$ eine Funktion $u \in \mathcal{C}^2(E')$ mit

$$f\Big(x, y, u(x,y), \frac{\partial u(x,y)}{\partial x}, \frac{\partial u(x,y)}{\partial y}, \frac{\partial^2 u(x,y)}{\partial x^2}, \frac{\partial^2 u(x,y)}{\partial y^2}, \frac{\partial^2 u(x,y)}{\partial x \partial y}\Big) = 0.$$

Entsprechend ist u Lösung des *Randwertproblems* für f mit Randwerten g, falls zusätzlich $u(x,y) = g(x,y)$ für $(x,y) \in \partial E'$ gilt, wobei $\partial E'$ den *Rand* von E' bezeichnet. (Anschaulich sollte klar sein was das ist, auch wenn wir hier keine formale Definition liefern.) Wir werden uns im Folgenden nicht mit partiellen Differentialgleichungen befassen, bemerken jedoch, dass wir in Beispiel 14.3 Lösungen der partiellen Differentialgleichung

$$\frac{\partial u^2(x,y)}{\partial x^2} + \frac{\partial u^2(x,y)}{\partial y^2} = 0$$

kennen gelernt haben. Diese Gleichung heißt auch *Laplace-Gleichung*.

Oftmals ist es hilfreich, eine Differentialgleichung in eine Integralgleichung umzuschreiben. Dies ist mit Hilfe des Hauptsatzes der Differential- und Integralrechnung einfach, wie das nächste Lemma zeigt.

Lemma 15.6 (Differential- und Integralgleichung). *Sei $I \subseteq \mathbb{R}$, $E \subseteq I \times \mathbb{R}^m$ offen, $(t_0, \underline{x}_0) \in E$ mit $t_0 \in J \subseteq I$ und $f \in \mathcal{C}^0(E, \mathbb{R}^m)$. Dann sind äquivalent:*

1. *$\underline{x} \in \mathcal{C}^1(J, \mathbb{R}^n)$ ist Lösung des $(\underline{f}, J, (t_0, \underline{x}_0))$-Anfangswertproblems in J.*
2. *$\underline{x} \in \mathcal{C}^0(J, \mathbb{R}^n)$ löst die Integralgleichung*

$$\underline{x}(t) = \underline{x}_0 + \int_{t_0}^{t} \underline{f}(s, \underline{x}(s))ds \tag{15.2}$$

für $t \in J$, wobei das Integral komponentenweise interpretiert wird.

Beweis. 1.⇒2.: Da $\underline{x}$ Lösung des Anfangswertproblems ist, folgt

$$\underline{x}(t) - \underline{x}_0 = \int_{t_0}^{t} \underline{x}'(s)ds = \int_{t_0}^{t} \underline{f}(s, \underline{x}(s))ds.$$

2.⇒1.: Da die rechte Seite von (15.2) differenzierbar ist, ist auch $\underline{x}$ differenzierbar. Leitet man beiden Seiten von (15.2) ab, folgt die Behauptung. □

15.2 Hilfsmittel

Im nächsten Abschnitt beschäftigen wir uns damit, wann es (eindeutige) Lösungen von Differentialgleichungen gibt. Wir haben bereits gesehen, dass die lineare Differentialgleichung $x' = \alpha x$ eine Lösung hat (und werden später sehen, dass die in Beispiel 15.3 angegebene Lösung des entsprechenden Anfangswertproblems eindeutig ist). Um Existenz und Eindeutigkeit im nächsten Abschnitt zeigen zu können, benötigen wir drei Hilfsaussagen. Den Schrankensatz (was eine Erweiterung von Proposition 7.19 ins Mehrdimensionale ist), das Gronwall-Lemma (Theorem 15.11) und den Banach'schen Fixpunktsatz (Theorem 15.15).

Bemerkung 15.7 (Lokale und globale Lipschitz-Stetigkeit). Wir erinnern an die Definition der Lipschitz-Stetigkeit aus Beispiel 5.6. Wir unterscheiden außerdem zwischen *globaler* und *lokaler* Lipschitz-Stetigkeit in folgendem Sinne:
Sei $E \subseteq \mathbb{R}^m$ und $\underline{f} : E \to \mathbb{R}^n$.

1. Die Funktion $\underline{f}$ heißt *global Lipschitz-stetig* mit Liptschitz-Konstante L, falls

$$|\underline{f}(\underline{y}) - \underline{f}(\underline{x})| \leq L \cdot |\underline{y} - \underline{x}|$$

   für alle $\underline{x}, \underline{y} \in E$ gilt.

2. Die Funktion $\underline{f}$ heißt *lokal Lipschitz-stetig*, falls es für jedes $\underline{x} \in E$ ein $\varepsilon > 0$ und ein $L_x \in \mathbb{R}_+$ gibt mit
$$|\underline{f}(\underline{z}) - \underline{f}(\underline{y})| \leq L_x \cdot |\underline{z} - \underline{y}|$$
für alle $\underline{y}, \underline{z} \in B_\varepsilon(\underline{x})$.

Theorem 15.8 (Schrankensatz). *Sei $E \subseteq \mathbb{R}^m$ offen und $\underline{f} \in \mathcal{C}^1(E, \mathbb{R}^n)$. Dann ist $\underline{f}$ lokal Lipschitz-stetig.*

Bemerkung 15.9 (Eine etwas genauere Aussage). Bezeichnet man mit $|.|$ die Supremumsnorm, also $|\underline{x}| := \max(|x_1|, ..., |x_m|)$ für $\underline{x} = (x_1, ..., x_m) \in \mathbb{R}^m$. Dann sieht man aus dem Beweis des Schrankensatzes, dass
$$|\underline{f}(\underline{x}) - \underline{f}(\underline{y})| \leq |\underline{\underline{D}}\,\underline{f}(\underline{x})| \cdot |\underline{x} - \underline{y}|, \tag{15.3}$$
falls $\underline{x} + t(\underline{y} - \underline{x}) \in E$ für $t \in [0, 1]$. Also sehen wir nicht nur, dass $\underline{f}$ lokal Lipschitz-stetig ist, sondern erhalten auch eine Abschätzung für die Lipschitz-Konstante.

Beweis. Sei $\underline{x} \in E$ und $\varepsilon > 0$ klein genug, so dass $B_\varepsilon(\underline{x}) \subseteq E$. Da stetige Funktionen auf kompakten Mengen ihr Maximum annehmen, gibt es ein L, so dass $\underline{y} \mapsto |\underline{\underline{D}}\,\underline{f}(\underline{y})| \leq L$ für $\underline{y} \in B_\varepsilon(\underline{x})$. Seien nun $\underline{y}, \underline{z} \in B_\varepsilon(\underline{x})$ sowie $\underline{\gamma}(t) := \underline{y} + t(\underline{z} - \underline{y})$. Dann gilt (siehe Lemma 14.23)
$$\begin{aligned} |\underline{f}(\underline{z}) - \underline{f}(\underline{y})| &= |\underline{f}(\underline{\gamma}(1)) - \underline{f}(\underline{\gamma}(0))| = \left| \int_0^1 \underline{\underline{D}}\,\underline{f}(\underline{\gamma}(t)) \underline{\underline{D}}\,\underline{\gamma}(t) dt \right| = \left| \int_0^1 \underline{\underline{D}}\,\underline{f}(\underline{\gamma}(t))(\underline{z} - \underline{y}) dt \right| \\ &\leq \int_0^1 |\underline{\underline{D}}\,\underline{f}(\underline{\gamma}(t))(\underline{z} - \underline{y})| dt \leq L|\underline{z} - \underline{y}|. \end{aligned}$$
□

Wir benötigen noch eine kleine Verallgemeinerung dieses Schrankensatzes.

Korollar 15.10. *Sei $E \subseteq \mathbb{R} \times \mathbb{R}^m$ offen und $\underline{f} \in \mathcal{C}^0(E, \mathbb{R}^n)$. Weiter betrachten wir für festes t die Abbildung $\underline{f}_t : \underline{x} \mapsto f(t, \underline{x})$. Ist $(t, \underline{x}) \mapsto \underline{\underline{D}}\,\underline{f}_t(\underline{x})$ stetig, dann ist $\underline{f}$ lokal Lipschitz-stetig bezüglich $\underline{x}$, das heißt für alle $(s, \underline{x}) \in E$ gibt es ein $\varepsilon > 0$ und $L < \infty$, so dass*
$$|\underline{f}(t, \underline{z}) - \underline{f}(t, \underline{y})| \leq L|\underline{z} - \underline{y}|$$
für alle $(t, \underline{y}), (t, \underline{z}) \in B_\varepsilon((s, \underline{x}))$.

Beweis. Sei $(s, \underline{x}) \in E$ und $\varepsilon > 0$ klein genug, so dass $B_\varepsilon((s, \underline{x})) \subseteq E$. Da $(t, \underline{x}) \mapsto \underline{\underline{D}} f_t(\underline{x})$ stetig ist, gibt es ein L, so dass $y \mapsto |\underline{\underline{D}}_s \underline{f}(\underline{y})| \leq L$ für $(t, \underline{y}) \in B_\varepsilon((s, \underline{x}))$. Der Rest folgt dann genau wie im letzten Beweis. □

Aus Beispiel 15.3 in Verbindung mit Lemma 15.6 sieht man, dass aus $f(t) = a + b \int_0^t f(s) ds$ folgt, dass $f(t) = ae^{bt}$ gelten muss. Das Gronwall-Lemma zeigt nun, dass diese Aussage auch gilt, wenn man jeweils '=' durch '≤' ersetzt.

Theorem 15.11 (Gronwall-Lemma). *Sei $f : [0, \infty) \to \mathbb{R}$ eine stetige Funktion und $a, b \geq 0$. Gilt*
$$f(t) \leq a + b \int_0^t f(s) ds \qquad \text{oder} \qquad f'(t) \leq bf(t) \text{ und } f(0) = a,$$
so folgt
$$f(t) \leq ae^{bt}$$
für alle $t \geq 0$.

Beweis. Aus der Voraussetzung folgt sofort

$$\frac{d}{dt}\Big(e^{-bt}\int_0^t f(s)ds\Big) = \Big(f(t) - b\int_0^t f(s)ds\Big)e^{-bt} \leq ae^{-bt}.$$

Integration und Einsetzen der Voraussetzung liefert

$$\frac{a}{b}\big(\tfrac{1}{a}f(t) - 1\big) \leq \int_0^t f(s)ds \leq e^{bt}a\int_0^t e^{-bs}ds = \frac{a}{b}(e^{bt} - 1).$$

Auflösen nach $f(t)$ liefert die Abschätzung. □

Der Banach'sche Fixpunktsatz ist ein abstraktes Resultat, das in vollständigen metrischen Räumen gilt.

Definition 15.12 (Vollständiger metrischer Raum). *Sei (E, r) ein metrischer Raum. Dieser heißt* vollständig*, wenn jede Cauchy-Folge konvergiert. Das bedeutet: Gilt für eine E-wertige Folge, dass für jedes $\varepsilon > 0$ ein N existiert, so dass für $m, n \geq N$ gilt, dass $r(x_n, x_m) < \varepsilon$. Dann gibt es ein $x \in E$ mit $x_n \xrightarrow{n\to\infty} x$.*

Beispiel 15.13. Sei r die euklidische Metrik in $\mathbb{R}$.

1. $(\mathbb{R}, r)$ ist vollständig. Schließlich besteht $\mathbb{R}$ gerade aus den Äquivalenzklassen aller Cauchy-Folgen.

2. $(\mathbb{Q}, r)$ ist nicht vollständig. Schließlich gibt es eine $\mathbb{Q}$-wertige Cauchy-Folge $x_1, x_2, ...$ mit $x_n \xrightarrow{n\to\infty} \sqrt{2}$.

Lemma 15.14 (Menge stetiger Funktionen ist vollständig). *Sei $I \subseteq \mathbb{R}$. Dann ist $\mathcal{C}(I, \mathbb{R}^m)$, versehen mit dem Supremumsabstand (d.h. für $f, g \in \mathcal{C}(I, \mathbb{R}^m)$ ist $r(f, g) := \sup_{t\in I} |f(t) - g(t)|$), ist vollständig.*

Beweis. Wir wissen bereits, dass $\mathbb{R}^n$, versehen mit der euklidischen Metrik vollständig ist. Sei nun $f_1, f_2, ... \in \mathcal{C}(I, \mathbb{R}^m)$ eine Cauchy-Folge (bzgl. dem Supremumsabstand). Dann ist für jedes $t \in I$ auch $f_1(t), f_2(t), ... \in \mathbb{R}^m$ eine Cauchy-Folge. Deswegen existiert eine Abbildung $f : I \to \mathbb{R}^n$ mit $f_n \xrightarrow{n\to\infty}_{pktw} f$. Sei nun $\varepsilon > 0$. Dann gilt nach Konstruktion für ein $N \in \mathbb{N}$, für das $r(f_m, f_n) < \varepsilon$ für $n > N(\varepsilon)$

$$|f_n(t) - f(t)| = \lim_{m\to\infty} |f_n(t) - f_m(t)| < \varepsilon,$$

und damit auch

$$\sup_{t\in I} |f_n(t) - f(t)| < \varepsilon.$$

Wir haben also gezeigt, dass sogar $f_n \xrightarrow{n\to\infty}_{glm} f$ und damit ist $f \in \mathcal{C}(I, \mathbb{R}^n)$ nach Lemma 8.6. □

Theorem 15.15 (Banach'scher Fixpunktsatz). *Sei (E, r) ein vollständiger metrischer Raum, $E' \subseteq E$ abgeschlossen und $f : E' \to E'$ eine Kontraktion, d.h. es gibt ein $c < 1$ mit*

$$r(f(x), f(y)) \leq c \cdot r(x, y)$$

für alle $x, y \in E'$. Dann gibt es genau einen Fixpunkt von f, d.h. es gibt genau ein $x \in E'$ mit $f(x) = x$. Weiter gilt für alle $x_0 \in E'$ und $x_{n+1} := f(x_n)$, $n = 0, 1, 2, ...$, dass $x_n \xrightarrow{n\to\infty} x$.

Beweis. Zunächst ist die Abbildung f Lipschitz-stetig und damit auch stetig. Wir zeigen die Existenz und Eindeutigkeit des Fixpunktes von f. Zunächst zur Eindeutigkeit: Seien $x, y \in E'$ zwei Fixpunkte. Dann gilt

$$r(x,y) = r(f(x), f(y)) \leq c \cdot r(x,y),$$

also $r(x,y) = 0$ und damit $x = y$.
Nun zur Existenz und zur behaupteten Konvergenz. Sei hierzu $x_0 \in E'$ beliebig und $x_{n+1} = f(x_n), n = 0, 1, 2, ...$ Dann gilt $r(x_{n+1}, x_n) \leq c^n r(x_1, x_0)$ für alle n. Denn: die Behauptung ist für $n = 0$ klar und gilt sie für ein n, so folgt $r(x_{n+2}, x_{n+1}) = r(f(x_{n+1}), f(x_n)) \leq c \cdot r(x_{n+1}, x_n) \leq c^{n+1} r(x_1, x_0)$ nach Induktionsvoraussetzung.

Wir zeigen nun, dass $(x_n)_{n=0,1,2,...}$ eine Cauchy-Folge ist. Sei hierzu $\varepsilon > 0$ beliebig. Dann gilt für $N(\varepsilon)$ groß genug für $c^{N(\varepsilon)} < (1-c)\varepsilon / r(x_1, x_0)$ und $n \geq m \geq N(\varepsilon)$

$$r(x_n, x_m) \leq \sum_{k=m}^{n-1} r(x_{k+1}, x_k) \leq r(x_1, x_0) \sum_{k=m}^{\infty} c^k = r(x_1, x_0) \frac{c^m}{1-c} \leq r(x_1, x_0) \frac{c^{N(\varepsilon)}}{1-c} < \varepsilon.$$

Da nach Voraussetzung (E', r) vollständig ist, gibt es ein $x \in E'$ mit $x_n \xrightarrow{n \to \infty} x$. Dieses x muss ein Fixpunkt von f sein, denn

$$f(x) = \lim_{n \to \infty} f(x_n) = \lim_{n \to \infty} x_{n+1} = x.$$

□

15.3 Existenz und Eindeutigkeit

Wir kommen nun zur Existenz und Eindeutigkeit von Lösungen von Anfangswertproblemen. Während die Eindeutigkeit eine Anwendung des Gronwall-Lemmas aus dem letzten Abschnitt ist (Theorem 15.16), benötigen wir den Banach'schen Fixpunktsatz zum Beweis der lokalen Existenz durch den Satz von Picard-Lindelöf (Theorem 15.19). Außerdem können wir die globale Existenz einer Lösung im Falle von höchstens linear wachsendes Koeffizienten zeigen (Theorem 15.22).

Theorem 15.16 (Eindeutigkeit von Anfangswertproblemen). *Sei $E \subseteq \mathbb{R} \times \mathbb{R}^m$ offen, $\underline{f} \in \mathcal{C}^0(E, \mathbb{R}^n)$, so dass für $\underline{f}_t : \underline{x} \mapsto \underline{f}(t, \underline{x})$ die Abbildung $(t, \underline{x}) \mapsto \underline{D}\,\underline{f}_t(\underline{x})$ stetig ist und $(t_0, \underline{x}_0) \in E$. Dann gibt es für $t_0 \in J \subseteq I$ höchstens eine Lösung des $(\underline{f}, J, (t_0, \underline{x}_0))$-Anfangswertproblems.*

Beweis. Seien $\underline{x}, \underline{y}$ zwei Lösungen des $(\underline{f}, J, (t_0, \underline{x}_0))$-Anfangswertproblems. Wir zeigen zunächst, dass $\underline{x}(s) = \underline{y}(s)$ für $s \in (t_0 - \delta, t_0 + \delta)$ für ein geeignetes $\delta > 0$. Wähle $\varepsilon > 0$ und $L < \infty$ nach Korollar 15.10. Weiter sei $\delta > 0$ so klein, dass $\underline{f}(t, \underline{x}(t)), \underline{f}(t, \underline{y}(t)) \in B_\varepsilon((t_0, \underline{x}_0))$ für $t \in (t_0 - \delta, t_0 + \delta)$. Dann gilt für $u(t) := |\underline{x}(t) - \underline{y}(t)|$ mit der Cauchy-Schwartz'schen Ungleichung

$$\begin{aligned}\frac{d}{dt}u^2(t) &= \frac{d}{dt}\sum_{k=1}^{n}(x_k(t)-y_k(t))^2\\ &= \sum_{k=1}^{n}2(x_k(t)-y_k(t))(x_k'(t)-y_k'(t))\\ &= \sum_{k=1}^{n}2(x_k(t)-y_k(t))(f_k(t,\underline{x}(t))-f_k(t,\underline{y}(t))\\ &\leq 2\sqrt{\sum_{k=1}^{n}(x_k(t)-y_k(t))^2}\sqrt{\sum_{k=1}^{n}(f_k(t,\underline{x}(t))-f_k(t,\underline{y}(t)))^2}\\ &= 2|\underline{x}(t)-\underline{y}(t)|\cdot|\underline{f}(t,\underline{x}(t))-\underline{f}(t,\underline{y}(t))|\\ &\leq 2L\cdot|\underline{x}(t)-\underline{y}(t)|^2 = 2Lu^2(t).\end{aligned}$$

Nun ist $u(0)=0$ und nach dem Gronwall-Lemma folgt mit dem soeben Gezeigten $u^2(t)\leq u^2(0)e^{2Lt}=0$, also $u(t)=0$ für $t\in(t_0-\delta,t_0+\delta)$. Damit ist also $\underline{x}(t)=\underline{y}(t)$ für $t\in(t_0-\delta,t_0+\delta)$.

Nun kommen wir Beweis von $\underline{x}(t)=\underline{y}(t)$ für $t\in J$. Wir setzen $t^*:=\sup\{t\in J:\underline{x}(t)=\underline{y}(t)\}$. Wäre $t^*<\sup J$, so wären sowohl $\underline{x}$ als auch $\underline{y}$ Lösung des $(\underline{f},J,(t^*,\underline{x}(t^*)))$-Anfangswertproblems. Nach oben gezeigtem gibt es ein $\delta>0$ mit $\underline{x}(t)=\underline{y}(t)$ für $t\in(t^*-\delta,t^*+\delta)$ im Widerspruch zur Maximalität von t^*. Analog zeigt man $t_*:=\inf\{t\in J:\underline{x}(t)=\underline{y}(t)\}=\inf J$. □

Das letzte Resultat ist nur dann anwendbar, wenn f' stetig ist. Ist dies nicht der Fall, kann es mehrere Lösungen des entsprechenden Anfangswertproblems geben, wie das nächste Beispiel zeigt.

Beispiel 15.17 (Ein nicht eindeutig lösbares Anfangswertproblem). Wir betrachten $f:(t,x)\mapsto 2\sqrt{|x|}$ und stellen fest, dass $f'(x)=\frac{1}{\sqrt{|x|}}$ in 0 nicht definiert und damit nicht stetig ist; insbesondere ist Theorem 15.16 nicht anwendbar. Weiter sei $(t_0,x_0)=(0,0)$. Dann sind alle Funktionen

$$x:t\mapsto\begin{cases}0, & t\leq t'\\ (t-t')^2, & t>t'\end{cases}$$

Lösungen des $(f,\mathbb{R},(0,0))$-Anfangswertproblems. Es ist nämlich sowohl $\frac{d}{dt}0=0=2\sqrt{|0|}$ als auch $\frac{d}{dt}(t-t')^2=2(t-t')=2\sqrt{(t-t')^2}$ für $t>t'$, also immer $x'(t)=2\sqrt{x(t)}$ und $x(0)=0$.

Die nächste Herausforderung wird nun der Beweis der Existenz einer Lösung des Anfangswertproblems. Wir werden zwei wichtige Aussagen hierzu treffen. Einerseits werden wir mit dem Satz von Picard-Lindelöf die zeitlich lokale Existenz in einem allgemeinen Fall zeigen (Theorem 15.19), andererseits können wir diese Aussage zu einer globalen Existenz erweitern, wenn die Funktion $\underline{f}$ höchstens linear wächst (Theorem 15.22). Wir beginnen mit einem Beispiel, warum bei superlinearem Wachstum von $\underline{f}$ nur eine lokale Lösung (d.h. für eine begrenzte Zeitmenge) zu erwarten ist.

Beispiel 15.18 (Explosion eines Anfangswertproblems). Sei $f : (t,x) \mapsto x^2$ und $(0,x_0) \in \mathbb{R}^2$ (und bemerken, dass f schneller als linear wächst). Wir suchen eine Lösung des $(f,J,(0,x_0))$-Anfangswertproblems für ein $J \subseteq \mathbb{R}$ mit $0 \in J$. Diese ist gegeben durch

$$x(t) = \frac{x_0}{1-tx_0},$$

es ist nämlich

$$x(0) = x_0 \text{ und } x'(t) = \frac{x_0^2}{(1-tx_0)^2} = x^2(t).$$

Allerdings ist diese Lösung nur auf dem Intervall $J := (-\infty, 1/x_0)$ definiert, da x in $t = 1/x_0$ eine Polstelle besitzt. Das bedeutet, dass man die (nach Theorem 15.16 eindeutige) Lösung des Anfangswertproblems nur für $J = (-\infty, 1/x_0)$ bekommt. Man sagt auch, dass die Lösung des $(f,J,(0,x_0))$-Anfangswertproblems bei $t = 1/x_0$ explodiert.

Theorem 15.19 (Kurzzeitexistenzsatz von Picard-Lindelöf). *Sei $E \subseteq \mathbb{R} \times \mathbb{R}^m$ offen, $\underline{f} \in \mathcal{C}^0(E,\mathbb{R}^n)$, so dass für $\underline{f}_t : \underline{x} \mapsto \underline{f}(t,\underline{x})$ die Abbildung $(t,\underline{x}) \mapsto \underline{\underline{D}}\,\underline{f}_t(\underline{x})$ stetig ist und $(t_0,\underline{x}_0) \in E$. Dann gibt es ein $\delta > 0$, so dass das $(\underline{f},(t_0-\delta,t_0+\delta),(t_0,\underline{x}_0))$-Anfangswertporblem eine Lösung besitzt.*

Beweis. Der Trick des Beweises besteht darin, eine Funktion F auf einem Funktionenraum zu definieren und die Aussage des Satzes als Fixpunktgleichung bzgl. F in einem vollständigen metrischen Raum zu schreiben. Zeigt man, dass F eine Kontraktion ist, folgt dann die Aussage unmittelbar aus dem Banach'schen Fixpunktsatz, Theorem 15.15.

Wir setzen zunächst

$$\mathcal{A}_I := \left\{\underline{x} \in \mathcal{C}^0(I,\mathbb{R}^n) : \sup_{t\in I} |\underline{x}(t) - \underline{x}_0| < \varepsilon\right\}$$

und $\underline{F}_I : \mathcal{A}_I \to \mathcal{C}^0(I,\mathbb{R}^n)$ mittels

$$(\underline{F}_I(\underline{x}))(t) := \underline{x}_0 + \int_{t_0}^t \underline{f}(s,\underline{x}(s))ds.$$

Dann ist nach Lemma 15.6 die Funktion $\underline{x}$ genau dann eine Lösung des $(\underline{f},I,(t_0,\underline{x}_0))$-Anfangswertproblems, wenn $\underline{x}(t) = (\underline{F}_I(\underline{x}))(t)$ für $t \in I$ gilt. Dies ist genau dann der Fall, wenn $\underline{x}$ ein Fixpunkt von $\underline{F}_I$ ist. Wir müssen also nun noch I so bestimmen, dass wir den Banach'schen Fixpunktsatz anwenden können. Sei hierzu $\varepsilon > 0$ und $L < \infty$ wie in Korollar 15.10. Außerdem ist $M := \sup |\underline{f}(B_\varepsilon(t_0,\underline{x}_0))|$. Dann gilt für $\delta' := \varepsilon \wedge \frac{1}{M}$ (also ist insbesondere $\delta' \leq \varepsilon$)

$$|(\underline{F}_{(t_0-\delta',t_0+\delta')}(\underline{x}))(t) - \underline{x}_0| = \left|\int_{t_0}^t \underline{f}(s,\underline{x}(s))ds\right| \leq M|t-t_0| \leq \varepsilon.$$

Mit $I = (t_0-\delta,t_0+\delta)$ gilt damit $\underline{F}_I(\mathcal{A}_I) \subseteq \mathcal{A}_I$ falls $\delta \leq \delta'$. Außerdem gilt für $\underline{x},\underline{y} \in \mathcal{A}_I$

$$\begin{aligned}|(\underline{F}_I(\underline{x}))(t) - (\underline{F}_I(\underline{y}))(t)| &= \left|\int_{t_0}^t (\underline{f}(s,\underline{x}(s)) - \underline{f}(s,\underline{y}(s))ds\right| \leq \delta \sup_{s\in I} |\underline{f}(s,\underline{x}(s)) - \underline{f}(s,\underline{y}(s)| \\ &\leq L\delta \sup_{s\in I} |\underline{x}(s) - \underline{y}(s)|.\end{aligned}$$

Wir setzen nun $\delta := \delta' \wedge 1/(2L)$, so gilt $\underline{F}_I(\mathcal{A}_I) \subseteq \mathcal{A}_I$ sowie $\sup_{t\in I} |(\underline{F}_I(\underline{x}))(t) - (\underline{F}_I(\underline{y}))(t)| \leq \frac{1}{2}\sup_{t\in I} |\underline{x}(t) - \underline{y}(t)|$. Mit anderen Worten ist $\underline{F}_I$ eine Kontraktion auf dem vollständigen metrischen Raum $\mathcal{A}_I$ und die Existenz eines Fixpunktes folgt direkt aus dem Banach'schen Fixpunktsatz. $\square$

Beispiel 15.20 (Picard-Iteration). Sei $\underline{f}$ wie in Theorem 15.19. Im gerade durchgeführten Beweis haben wir gesehen, dass die Existenz der Lösung auf den Banach'schen Fixpunktsatz zurückgeführt werden kann. In seiner Formulierung in Theorem 15.15 haben wir außerdem gesehen, dass eine Fixpunktgleichung durch Iteration gelöst werden kann. Dies wollen wir nun an einem Beispiel veranschaulichen. Sei hierzu $f : (t, x) \mapsto \alpha x$. Wir wissen bereits, dass $x(t) = x_0 e^{\alpha t}$ einzige Lösung des $(f, \mathbb{R}, (0, x_0))$-Anfangswertproblems ist.

Wir setzen $x_0(t) := x_0$ und betrachten die Iterationsvorschrift

$$x_{n+1} := F(x_n) : t \mapsto x_0 + \int_0^t \alpha x_n(s) ds.$$

Dann berechnen wir sukzessive

$$\begin{aligned}
x_1(t) &= x_0 + \int_0^t \alpha x_0 ds = x_0(1 + \alpha t), \\
x_2(t) &= x_0 + \int_0^t \alpha x_0 (1 + \alpha s) ds = x_0\Big(1 + \alpha t + \frac{\alpha^2 s^2}{2}\Big), \\
x_3(t) &= x_0 + \int_0^t \alpha x_0\Big(1 + \alpha s + \frac{\alpha^2 s^2}{2}\Big) ds = x_0\Big(1 + \alpha t + \frac{\alpha^2 s^2}{2} + \frac{\alpha^3 t^3}{6}\Big), \\
&\dots
\end{aligned}$$

Hieraus sieht man leicht, dass

$$x_n(t) \xrightarrow{n\to\infty} x_0 \sum_{k=0}^{\infty} \frac{\alpha^k t^k}{k!} = x_0 e^{\alpha t}.$$

Insbesondere konvergiert die Iteration global auf $\mathbb{R}$ gegen die bereits gefundene Lösung.

Theorem 15.19 liefert eine Lösung des Anfangswertproblems, aber nur in einem kleinen Zeitraum. In Beispiel 15.18 haben wir außerdem gesehen, dass es sein kann, dass die Lösung eines Anfangswertproblems *explodiert.* Wir zeigen nun, dass eine solche Explosion im Normalfall die einzige Möglichkeit ist, warum eine Lösung des Anfangswertproblems nur lokal existiert.

Lemma 15.21 (Explosion eines Anfangswertproblems). *Sei $\underline{f} \in \mathcal{C}^0(I \times \mathbb{R}^m, \mathbb{R}^n)$, so dass für $\underline{f}_t : \underline{x} \mapsto \underline{f}(t, \underline{x})$ die Abbildung $(t, \underline{x}) \mapsto \underline{\underline{D}}\,\underline{f}_t(\underline{x})$ stetig ist und $(t_0, \underline{x}_0) \in I \times \mathbb{R}^m$. Weiter sei $J \subseteq I$ maximal, so dass das $(\underline{f}, J, (t_0, \underline{x}_0))$-Anfangswertproblem eine (nach Theorem 15.16 eindeutige) Lösung $\underline{x}$ hat. Dann gilt*

1. *Ist $t^* := \sup J < \sup I$, so gilt $|\underline{x}(t)| \xrightarrow{t\uparrow t^*} \infty$.*

2. *Ist $t_* := \inf J > \inf I$, so gilt $|\underline{x}(t)| \xrightarrow{t\downarrow t_*} \infty$.*

Beweis. Wir zeigen nur die erste Aussage, die zweite folgt analog. Wir definieren für eine kompakte Menge K

$$M_K := \sup\{|\underline{f}(t, \underline{x})| : t \in [t_0, t^*], \underline{x} \in K\} < \infty.$$

Angenommen, die Aussage wäre falsch. Dann gibt es zwei Möglichkeiten:

1. Es gibt eine kompakte Menge $K \subseteq \mathbb{R}^n$, so dass $\underline{x}(t) \in K$ für alle $t \in J$.
   Dann gilt für $t, t' \in [t_0, t^*]$
$$|\underline{x}(t') - \underline{x}(t)| = \left| \int_t^{t'} \underline{f}(s, \underline{x}(s)) ds \right| \leq M_K |t' - t|.$$
   Daraus folgt die Konvergenz $\underline{x}(t) \xrightarrow{t \to t^*} x^*$ für ein $x^* \in K$. Nun hat allerdings das $(\underline{f}, (t^* - \delta, t^* + \delta), (t^*, \underline{x}^*))$-Anfangswertproblem für ein $\delta > 0$ nach Theorem 15.19 eine Lösung im Widerspruch zur Maximalität von t^*.

2. Es gibt eine kompakte Menge K und ein $\varepsilon > 0$, sowie eine Folge $t_n \uparrow t^*$ mit $\underline{x}(t_n) \in K$ und $t'_n := \inf\{t > t_n : r(\underline{x}(t), K) \geq \varepsilon\} \uparrow t^*$. Das bedeutet, dass $\underline{x}(t)$ zwischen K und $B_\varepsilon(K)^c$ oszilliert.
   In diesem Fall setzen wir $K' := B_\varepsilon(K) = \{\underline{x} : \exists \underline{y} \in K : |\underline{x} - \underline{y}| \leq \varepsilon\}$ und stellen fest, dass K' kompakt ist. Damit gilt
$$\varepsilon \leq |\underline{x}(t_n) - \underline{x}(t'_n)| = \left| \int_{t_n}^{t'_n} \underline{f}(s, \underline{x}(s)) ds \right| \leq M_{K'} |t_n - t'_n| \xrightarrow{n \to \infty} 0,$$
   ein Widerspruch.

Da sich in beiden Fällen ein Widerspruch ergibt, ist damit die Aussage gezeigt. □

Theorem 15.22 (Globale Lösung eines Anfangswertproblems). *Sei $\underline{f} \in \mathcal{C}^0(I \times \mathbb{R}^m, \mathbb{R}^n)$, so dass für $\underline{f}_t : \underline{x} \mapsto \underline{f}(t, \underline{x})$ die Abbildung $(t, \underline{x}) \mapsto \underline{\underline{D}}\, \underline{f}_t(\underline{x})$ stetig ist und $(t_0, \underline{x}_0) \in I \times \mathbb{R}^m$. Für $a, b \geq 0$ gelte*
$$|\underline{f}(t, \underline{x})| \leq a + b|\underline{x}|.$$
Dann hat das $(\underline{f}, I, (t_0, \underline{x}_0))$-Anfangswertproblem eine (nach Theorem 15.16 eindeutige) Lösung und es gilt
$$|\underline{x}(t)| \leq (|\underline{x}_0| + at)e^{bt}.$$

Beweis. Sei J das maximale Intervall, für das das $(\underline{f}, J, (t_0, \underline{x}_0))$-Anfangswert eine Lösung hat. Wir werden mit Hilfe von Lemma 15.21 zeigen, dass $J = I$ gilt. Es gilt nach Voraussetzung die Abschätzung
$$|\underline{x}(t)| \leq |\underline{x}_0| + \int_0^t |\underline{f}(s, \underline{x}(s))| ds \leq |\underline{x}_0| + \int_0^t a + b|\underline{x}(s)| ds = (|\underline{x}_0| + at) + b \int_0^t |\underline{x}(s)| ds.$$
Mit dem Gronwall-Lemma folgt daraus
$$|\underline{x}(t)| \leq (|\underline{x}_0| + at)e^{bt},$$
also die gewünschte Abschätzung. Wäre nun $\sup J < \infty$, so müsste nach Lemma 15.21 gelten, dass $|\underline{x}(t)| \xrightarrow{t \uparrow \sup J} \infty$ im Widerspruch zur Abschätzung. □

15.4 Lineare Differentialgleichungen

In diesem Abschnitt behandeln wir Differentialgleichungen der speziellen Form

$$\underline{x}'(t) = \underline{a}(t) + \underline{\underline{B}}(t)\underline{x}(t)$$

für Funktionen $t \mapsto \underline{a}(t)$ und $t \mapsto \underline{\underline{B}}(t)$. Im Falle eines homogenen Systems (d.h. $\underline{a} = 0$) von konstanten Koeffizienten (d.h. $\underline{\underline{B}}$ ist nicht von t abhängig), können wir sogar eine Lösungsmethode für diese Differentialgleichung angeben (siehe Theorem 15.27). Wir beginnen zunächst mit dem einfachen eindimensionalen Fall. Hier kann man Lösungen im allgemeinen Fall explizit angeben.

Proposition 15.23 (Allgemeine Lösung eines linearen, eindimensionalen Anfangswertproblems). *Sei $I \subseteq \mathbb{R}$ sowie $(t_0, x_0) \in I \times \mathbb{R}$ und*

$$f(t,x) = a(t) + b(t)x$$

für b stückweise stetig. Dann wird das $(f, I, (t_0, x_0))$-Anfangswertproblem durch die Funktion

$$x(t) = e^{\int_{t_0}^t b(s)ds}\Big(\int_{t_0}^t a(s)e^{-\int_{t_0}^t b(r)dr}ds + x_0\Big)$$

eindeutig gelöst.

Beweis. Die Eindeutigkeit der Lösung folgt aus Theorem 15.16. Weiter gilt für die angegebene Lösung $x(t_0) = x_0$ sowie

$$\frac{d}{dt}x(t) = b(t)x(t) + e^{\int_{t_0}^t b(s)ds}\frac{d}{dt}\Big(\int_{t_0}^t a(s)e^{-\int_{t_0}^t b(r)dr}ds + x_0\Big) = a(t) + b(t)x(t).$$

$\square$

Im Folgenden soll es um homogene lineare Gleichungen gehen, d.h. Anfangswertprobleme der Form

$$\underline{x}'(t) = \underline{\underline{B}}(t)\underline{x}(t).$$

Im Eindimensionalen ist diese Gleichung mit Hilfe des letzten Resultats schnell gelöst. Nach dem letzten Resultat ist nämlich die Lösung von $x'(t) = b(t)x(t)$ gerade $x(t) := x_0 \exp\Big(\int_{t_0}^t b(s)ds\Big)$. Im Mehrdimensionalen ist es allerdings nicht mehr so einfach, Lösungen anzugeben. Wir behandeln im Folgenden den Spezialfall einer homogenen, linearen Differentialgleichungen. Beginnen werden wir mit einer Beobachtung über den Lösungsraum einer solchen Differentialgleichung.

Lemma 15.24 (Lösungsraum ist linear). *Sei $I \subseteq \mathbb{R}$ und $\underline{\underline{B}} \in \mathcal{C}^0(I, \mathbb{R}^{m\times m})$. Dann hat für jedes $(t_0, \underline{x}_0) \in \mathbb{R} \times \mathbb{R}^m$ und $\underline{f}(t, \underline{x}) := \underline{\underline{B}}(t)\underline{x}(t)$ das $(\underline{f}, I, (t_0, \underline{x}_0))$-Anfangswertproblem genau eine Lösung $\underline{x}$. Weiter ist der Lösungsraum*

$$\mathcal{L} := \big\{\underline{x} \in \mathcal{C}^1(I, \mathbb{R}^m) \text{ Lösung des } (\underline{f}, I, (t_0, \underline{x}_0))\text{-Anfangswertproblems für ein } (t_0, \underline{x}_0)\big\}$$

ein m-dimensionaler Vektorraum und die Abbildung

$$(t_0, \underline{x}_0) \mapsto \underline{x} \in \mathcal{C}^1(I, \mathbb{R}^m) \text{ Lösung des } (\underline{f}, I, (t_0, \underline{x}_0))\text{-Anfangswertproblems für } (t_0, \underline{x}_0)$$

ist für jedes $t_0 \in I$ ein Vektorraum-Isomorphismus.

Beweis. Die Linearlität des Raumes $\mathcal{L}$ sowie die Linearlität der angegebenen Abbildung sind klar. Zu zeigen ist dann nur die zweite Aussage, woraus die erste sofort folgt. Es genügt zu zeigen, dass die Abbildung bijektiv ist. Zur Surjektivität bemerken wir, dass jeder Anfangswert in $\mathcal{L}$ vorkommt, und somit das Bild der linearen Abbildung ganz $\mathcal{L}$ ist. Die Injektivität folgt aus der Eindeutigkeit der Lösung des Anfangswertporblems. □

Ziel ist es nun, eine effiziente Methode anzugeben, wie ein homogenes, lineares Differentialgleichungssystem mit konstanten Koeffizienten gelöst werden kann. Von unserer Kenntnis des eindimensionalen Falls wissen wir, dass $x'(t) = bx(t)$ durch $x(t) = x_0 \exp(b(t - t_0))$ gelöst wird. Analog müssen wir nun im Mehrdimensionalen dem Ausdruck $\exp\big(\underline{\underline{B}}(t-t_0)\big)$ erst einmal einen Sinn verleihen.

Definition 15.25 (Matrix-Exponential). *Für $\underline{\underline{B}} \in \mathbb{R}^{m\times m}$ definieren wir*

$$\exp\big(\underline{\underline{B}}\big) := \sum_{n=0}^{\infty} \frac{\underline{\underline{B}}^n}{n!}.$$

(Es sei bemerkt, dass die Summe auf der rechten Seite immer konvergiert, weil $|\underline{\underline{B}}^n| \le C^n$ für ein $C \in \mathbb{R}$ gilt.) Weiter ist $\underline{\underline{B}}^0 := \underline{\underline{E}}_m$ die Einheitsmatrix.

Lemma 15.26 (Eigenschaften des Matrix-Exponentials). *Seien $\underline{\underline{A}}, \underline{\underline{B}} \in \mathbb{R}^{m\times m}$ und $\underline{\underline{S}} \in GL_m(\mathbb{R})$. Dann gilt:*

1. $\exp\big(\underline{\underline{B}}\big) \in GL_m(\mathbb{R})$,
2. $\exp\big(\underline{\underline{A}} + \underline{\underline{B}}\big) = \exp\big(\underline{\underline{A}}\big)\exp\big(\underline{\underline{B}}\big)$, *falls* $\underline{\underline{A}}\,\underline{\underline{B}} = \underline{\underline{B}}\,\underline{\underline{A}}$.
3. $\exp\big(\underline{\underline{S}}\,\underline{\underline{B}}\,\underline{\underline{S}}^{-1}\big) = \underline{\underline{S}}\exp\big(\underline{\underline{B}}\big)\underline{\underline{S}}^{-1}$.
4. *Ist* $\underline{\underline{D}} = diag(\lambda_1, ..., \lambda_m)$ *eine Diagonalmatrix, so ist* $\exp(\underline{\underline{D}}) = diag(e^{\lambda_1}, ..., e^{\lambda_m})$.
5. *Ist* $\underline{\underline{B}} = (\delta_{i,j+1})_{1\le i,j\le m}$ *(diese Matrizen können in der Jordan'schen Normalform einer Matrix* $\underline{\underline{B}}$ *vorkommen), so ist*

$$\Big(\exp\big(\underline{\underline{B}}t\big)\Big)_{ij} = \begin{cases} \frac{t^{j-i}}{(j-i)!}, & \text{für } j \ge i, \\ 0, & \text{sonst.} \end{cases}.$$

Beweis. Wir zeigen zunächst 2, indem wir wie im Beweis von Lemma 6.6

$$\begin{aligned}\exp\big(\underline{\underline{A}} + \underline{\underline{B}}\big) &= \sum_{n=0}^{\infty} \frac{(\underline{\underline{A}} + \underline{\underline{B}})^n}{n!} = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} \underline{\underline{A}}^k \underline{\underline{B}}^{n-k} \\ &= \sum_{k=0}^{\infty}\sum_{n=k}^{\infty} \frac{1}{k!}\frac{1}{(n-k)!}\underline{\underline{A}}^k\underline{\underline{B}}^{n-k} = \sum_{k=0}^{\infty}\frac{1}{k!}\underline{\underline{A}}^k \sum_{n=0}^{\infty}\frac{1}{(n-k)!}\underline{\underline{B}}^n \\ &= \exp\big(\underline{\underline{A}}\big)\exp\big(\underline{\underline{B}}\big)\end{aligned}$$

schreiben. Hierbei haben wir im zweiten Gleichheitszeichen die Voraussetzung $\underline{\underline{A}}\,\underline{\underline{B}} = \underline{\underline{B}}\,\underline{\underline{A}}$ benötigt. Mit 2. gilt weiter $\exp\big(\underline{\underline{A}}\big)\exp\big(-\underline{\underline{A}}\big) = \exp\big(\underline{\underline{0}}\big) = \underline{\underline{E}}_m$, da $\underline{\underline{A}}(-\underline{\underline{A}}) = (-\underline{\underline{A}})(\underline{\underline{A}})$. Also gilt auch 1.

Für 3. genügt es zu bemerken, dass

$$(\underline{\underline{S}}\,\underline{\underline{A}}\,\underline{\underline{S}}^{-1})^n = \underbrace{\underline{\underline{S}}\,\underline{\underline{A}}\,\underline{\underline{S}}^{-1}\cdots\underline{\underline{S}}\,\underline{\underline{A}}\,\underline{\underline{S}}^{-1}}_{n-\text{mal}} = \underline{\underline{S}}\,\underline{\underline{A}}^n\,\underline{\underline{S}}^{-1}.$$

Damit können wir sofort

$$\exp\left(\underline{\underline{S}}\,\underline{\underline{A}}\,\underline{\underline{S}}^{-1}\right) = \sum_{n=0}^{\infty}\frac{(\underline{\underline{S}}\,\underline{\underline{A}}\,\underline{\underline{S}}^{-1})^n}{n!} = \underline{\underline{S}}\Big(\sum_{n=0}^{\infty}\frac{\underline{\underline{A}}^n}{n!}\Big)\underline{\underline{S}}^{-1} = \underline{\underline{S}}\exp\left(\underline{\underline{A}}\right)\underline{\underline{S}}^{-1}$$

schreiben.

4. ist klar wegen $\underline{\underline{D}}^n = \operatorname{diag}(\lambda_1^n, ..., \lambda_m^n)$. Für 5. bemerken wir, dass $\underline{\underline{B}}$ eine nilpotnte Matrix ist. Für $n \le m$ ist $\underline{\underline{B}}^n = (\delta_{i,j+m})_{1\le i,j\le m}$ und für $n > m$ ist $\underline{\underline{B}}^n = \underline{\underline{0}}$. Hieraus folgt die angegebene Formel. □

Aus der linearen Algebra ist bekannt, dass man jeder Matrix $\underline{\underline{B}}$ eine Jordan'sche Normalform $\underline{\underline{J}}$ zuordnen kann. Außerdem gibt es eine Matrix $\underline{\underline{S}}$ mit $\underline{\underline{J}} = \underline{\underline{S}}^{-1}\underline{\underline{B}}\,\underline{\underline{S}}$. Wir behandeln nun den Fall, dass $\underline{\underline{J}}$ nur aus einem einzigen *Jordan-Kästchen* besteht. Der allgemeine Fall lässt sich hieraus leicht herleiten.

Theorem 15.27 (Allgemeine Lösung der homogenen linearen Differentialgleichung mit konstanten Koeffizienten). *Sei* $\underline{\underline{B}} \in \mathbb{R}^{m\times m}$ *und* $\underline{\underline{S}} \in \mathbb{R}^{m\times m}$*, so dass*

$$\underline{\underline{J}} := \underline{\underline{S}}^{-1}\underline{\underline{B}}\,\underline{\underline{S}} = \begin{pmatrix} \lambda & 1 & 0 & \cdots & \cdots & 0 \\ 0 & \lambda & 1 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & 0 & \lambda & 1 \\ 0 & \cdots & \cdots & \cdots & 0 & \lambda \end{pmatrix}.$$

Für $\underline{f}(t,\underline{x}) = \underline{\underline{B}}\,\underline{x}$ *und* $(t_0,\underline{x}_0) \in \mathbb{R}\times\mathbb{R}^m$ *ist die Lösung des* $(\underline{f},\mathbb{R},(t_0,\underline{x}_0))$*-Anfangswertproblems durch*

$$\underline{x}(t) = \exp\left(\underline{\underline{B}}(t-t_0)\right)\underline{x}_0 = \underline{\underline{S}}\exp\left(\underline{\underline{J}}(t-t_0)\right)\underline{\underline{S}}^{-1}\underline{x}_0 \tag{15.4}$$

gegeben. Außerdem ist

$$\exp\left(\underline{\underline{J}}t\right) = e^{\lambda t}\begin{pmatrix} 1 & t & \frac{t^2}{2} & \cdots & \cdots & \frac{t^{m-1}}{(m-1)!} \\ 0 & 1 & t & \frac{t^2}{2} & \cdots & \frac{t^{m-2}}{(m-2)!} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & 0 & 1 & t \\ 0 & \cdots & \cdots & \cdots & 0 & 1 \end{pmatrix}.$$

Beweis. Der Beweis ist eine Anwendung von Lemma 15.26. Zunächst rechnet man direkt nach, dass $\underline{x}_0\exp\left(\underline{\underline{B}}(t-t_0)\right)$ das Anfangswertproblem löst. Das zweite '=' in (15.4) folgt aus Lemma 15.26.3. Für die angegebene Formel für $\exp(\underline{\underline{J}}t)$ stellen wir fest, dass $\underline{\underline{J}}$ Summe aus einer Diagonalmatrix und einer nilpotenten Matrix ist, die kommutieren. Deshalb ist Lemma 15.26.2 anwendbar. Die Formel folgt dann aus Lemma 15.26.5. □

Beispiel 15.28. Wir betrachten die Funktion $\underline{f} : \underline{x} \mapsto \underline{\underline{B}}\,\underline{x}$ für die Matrix

$$\underline{\underline{B}} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

und das $(\underline{f}, \mathbb{R}, (t_0, \underline{x}_0))$-Anfangswertproblem. Nach Theorem 15.27 wird dies durch

$$\underline{x}(t) = \exp(\underline{\underline{B}}(t - t_0))\underline{x}_0$$

gelöst. Wir berechnen nun das Matrix-Exponential. Da

$$\begin{aligned} \exp(\underline{\underline{B}}(t-t_0)) &= \exp((\underline{\underline{E}}_3 + (\underline{\underline{B}} - \underline{\underline{E}}_3))(t-t_0)) = \exp(\underline{\underline{E}}_3(t-t_0)) \cdot \exp((\underline{\underline{B}} - \underline{\underline{E}}_3))(t-t_0)) \\ &= e^{t-t_0} \begin{pmatrix} 1 & t-t_0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \end{aligned}$$

ist die Lösung des Anfangswertproblems für $\underline{x}_0 = (x, y, z)$

$$\underline{x}(t) = e^{t-t_0} \begin{pmatrix} 1 & t-t_0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = e^{t-t_0} \begin{pmatrix} x + y(t-t_0) \\ y \\ z \end{pmatrix}$$

Zur Kontrolle leiten wir nach t ab und erhalten

$$\underline{x}'(t) = \underline{x}(t) + e^{t-t_0} \begin{pmatrix} y \\ 0 \\ 0 \end{pmatrix} = \underline{\underline{B}}\,\underline{x}(t).$$

15.5 Trennung der Variablen

Wir lernen nun ein weiteres Verfahren kennen, Differentialgleichungen zu lösen. Diese nennt man auch *Trennung der Variablen* und kann bei eindimensionalen Problemen eingesetzt werden.

Bemerkung 15.29 (Getrennte Variablen). Wir betrachten eine ein-dimensionale Differentialgleichung der Form

$$x'(t) = b(t)g(x(t)).$$

Die rechte Seite ist eine spezielle Funktion der beiden Variablen t und x, da man sie als Produkt zweier Funktionen schreiben kann, die jeweils nur von einer Variablen abhängen.

Um ein Anfangswertproblem mit $x(t_0) = x_0$ zu lösen, formen wir zunächst recht informell die Gleichnug $\frac{dx}{dt} = b(t)g(x)$ in

$$\frac{1}{g(x)}dx = b(t)dt$$

um. Integration unter Beachtung des Anfangswertes ergibt nun

$$\int_{x_0}^{x(t)} \frac{1}{g(x)}dx = \int_{t_0}^{t} b(s)ds. \tag{15.5}$$

Kann man beide Seiten berechnen, und anschließend die Gleichung nach $x(t)$ auflösen, erhält man eine Lösung des Anfangswertproblems. Dies ist Inhalt des nächsten Theorems.

Theorem 15.30 (Trennung der Variablen). *Sei $I \subseteq \mathbb{R}$ und $f : (t,x) \mapsto b(t)g(x)$ für $b \in \mathcal{C}^0(I)$ und $g \in \mathcal{C}^1(\mathbb{R})$. Weiter sei $(t_0, x_0) \in I \times \mathbb{R}$.*

1. *Ist $b(t_0) = 0$, so ist $x(t) = x_0$ Lösung des $(f, I, (t_0, x_0))$-Anfangswertproblems.*
2. *Ist $b(t_0) \neq 0$, so hat das $(f, (t_0 - \varepsilon, t_0 + \varepsilon), (t_0, x_0))$-Anfangswertproblems für ein hinreichend kleines $\varepsilon > 0$ eine Lösung. Diese kann durch Auflösen von (15.5) nach $x(t)$ erhalten weden.*

Beweis. In 1. ist die Aussage klar, wie man direkt nachrechnet. Für 2. folgt die Existenz der Lösung aus Theorem 15.19. Weiter wählen wir zunächst $\varepsilon > 0$ so klein, dass g auf $J := (t_0 - \varepsilon, t_0 + \varepsilon)$ keine Nullstelle hat. Insbesondere ist dann das Vorzeichen von g (also auch von $1/g$) auf diesem Intervall einheitlich. Wir definieren die Funktion $G : J \to \mathbb{R}$ durch

$$G(x) := \int_{x_0}^{x} \frac{1}{g(y)} dy$$

und

$$B(t) := \int_{t_0}^{t} b(s) ds.$$

Die Abbildung G hat auf J Ableitung $1/g(x)$, ist damit streng monoton und besitzt eine differenzierbare Umkehrabbildung G^{-1}. Wir setzen $x(t) := G^{-1}(B(t))$ und bemerken erst, dass $x(t_0) = G^{-1}(B(t_0)) = G^{-1}(0) = x_0$. Weiter ist nach Proposition 7.8

$$x'(t) = \frac{d}{dt} G^{-1}(B(t)) = \frac{d}{dx} G^{-1}(B(t)) \frac{d}{dt} B(t) = \frac{1}{G'(G^{-1}(B(t)))} b(t) = g(x(t)) b(t).$$

Damit sind alle Aussagen gezeigt. □

Beispiel 15.31 (Exponentielles Wachstum). Das Anfangswertproblem

$$x'(t) = b(t)x$$

mit $x(0) = x_0$ haben wir zwar bereits in gelöst (siehe Proposition 15.23), können dies jedoch nochmal durch Trennung der Variablen tun. Hierzu schreiben wir

$$\int_{x_0}^{x} \frac{1}{y} dy = \int_{t_0}^{t} b(s) ds,$$

was nach zwei Rechenschritten zur Identität

$$x = x_0 \exp\Big(\int_{t_0}^{t} b(s) ds \Big)$$

führt.

Beispiel 15.32 (Logistisches Wachstum). Wir betrachten das Anfangswertproblem

$$x'(t) = bx(K - x)$$

mit $x(t_0) = x_0$. Wieder können wir die Trennung der Variablen ausnutzen und schreiben

$$\int_{x_0}^{x} \frac{1}{y(K-y)} dy = \int_{t_0}^{t} b ds.$$

Da

$$\int_{x_0}^{x} \frac{1}{y(K-y)}dx = \frac{1}{K}\int_{x_0}^{x} \frac{1}{y} + \frac{1}{K-y}dx = \frac{1}{K}\Big(\log\Big(\frac{x}{x_0}\Big) - \log\Big(\frac{K-x}{K-x_0}\Big)\Big)$$
$$= \frac{1}{K}\Big(\log\Big(\frac{x}{K-x}\Big) - \log\Big(\frac{x_0}{K-x_0}\Big)\Big),$$

gilt

$$\frac{Kx_0}{x} = \Big(\frac{K-x}{x}+1\Big)x_0 = \Big(\frac{K-x_0}{x_0}e^{-Kb(t-t_0)}+1\Big)x_0 = Ke^{-Kb(t-t_0)} + x_0(1-e^{-Kb(t-t_0)}),$$

also

$$x(t) = \frac{Kx_0}{Ke^{-Kb(t-t_0)} + x_0(1-e^{-Kb(t-t_0)})}.$$

16 Anwendungen der mehrdimensionalen Ableitungen

Wir erweitern nun die Anwendungen der mehr-dimensionalen Ableitungen. Einerseits werfen dabei höher-dimensionale Funktionen neue, bisher nicht bekannte Fragestellungen auf (siehe Abschnitt 16.1 und Kapitel 17), andererseits erfordern sie Verallgemeinerungen von bereits aus dem ein-dimensionalen bekannten Resultaten (siehe Abschnitte 16.2, 16.3 und 16.4). In diesem Abschnitt bezeichnet $|.|$ durchgehend die Supremumsnorm.

16.1 Parameterabhängige Integrale

Sei $E \subseteq \mathbb{R}^m$ offen, $I \subseteq \mathbb{R}$ ein Intervall und $f : E \times I \to \mathbb{R}$. Wir beschäftigen uns nun mit Eigentschaften der Funktion

$$\underline{x} \mapsto \int_I f(\underline{x}, t)dt. \tag{16.1}$$

Etwa werden wir sehen, dass diese stetig ist, wenn f stetig ist, sowie eine Aussage über die Möglichkeit, diese Abbildung partiell abzuleiten. Als Mindestvoraussetzung muss gelten, dass für alle $\underline{x} \in E$ die Abbildung $t \mapsto f(\underline{x}, t)$ Riemann-integrierbar ist, sonst macht nämlich die (16.1) angegebene Funktion keinen Sinn. Wir fassen das Ergebnis unserer Untersuchung sofort zusammen.

Theorem 16.1 (Stetigkeit, Differenzierbarkeit parameterabhängiger Integrale). *Sei $E \subseteq \mathbb{R}^m$ offen, I ein kompaktes Intervall und $f \in \mathcal{C}^0(E \times I)$. Dann ist die Funktion*

$$\underline{x} \mapsto \int_I f(\underline{x}, t)dt$$

wohldefiniert und stetig. Existiert außerdem $\frac{\partial f}{\partial x_j}$ für ein $j \in \{1, ..., m\}$ und ist diese partielle Ableitung stetig, so gilt auch

$$\frac{\partial}{\partial x_j}\int_I f(\underline{x}, t)dt = \int_I \frac{\partial f(\underline{x}, t)}{\partial x_j}dt. \tag{16.2}$$

Bemerkung 16.2 (Sprechweise). Zu (16.2) sagt man auch, man habe die Funktion f *unter dem Integral* abgeleitet.

Beweis von Theorem 16.1. Die Wohdefiniertheit ist klar, da für alle $\underline{x} \in E$ der Integrand $t \mapsto f(\underline{x}, t)$ stetig und damit Riemann-integrierbar ist. Sei $I = [a, b]$. Für $\underline{x} \in E$ sei $\delta' > 0$ klein genug, so dass $\overline{B}_{\delta'}(\underline{x}) \subseteq E$. Da $\overline{B}_{\delta'}(\underline{x}) \times I$ kompakt ist, ist f auf dieser Menge gleichmäßig stetig (siehe Proposition 5.28). Also gibt es für $\varepsilon > 0$ ein $\delta > 0$ (mit $\delta \leq \delta'$), so dass $|f(\underline{y}, t) - f(\underline{x}, t)| < \varepsilon/(b-a)$ für $\underline{y} \in B_\delta(\underline{x})$. Damit gilt auch

$$\Big| \int_I f(\underline{y}, t) dt - \int_I f(\underline{x}, t) dt \Big| \leq \int_I |f(\underline{y}, t) - f(\underline{x}, t)| dt \leq \varepsilon.$$

Damit haben wir die geforderte Stetigkeit gezeigt.

Wir kommen nun zur Differenzierbarkeit und der Ableitbarkeit unter dem Integral. Da (genau wie oben) $\frac{\partial f}{\partial x_j}$ auf $\overline{B}_{\delta'}(\underline{x}) \times I$ gleichmäßig stetig ist, gibt es für $\varepsilon > 0$ ein $\delta > 0$ (mit $\delta \leq \delta'$), so dass

$$\Big| \frac{\partial f(\underline{y}, t)}{\partial x_j} - \frac{\partial f(\underline{x}, t)}{\partial x_j} \Big| \leq \varepsilon/(b-a)$$

für $\underline{y} \in B_\delta(\underline{x})$. Nach dem Hauptsatz der Differential- und Integralrechnung ist nun für $h < \delta$

$$\begin{aligned} &\Big| \frac{1}{h} \int_I f(\underline{x} + h\underline{e}_j, t) - f(\underline{x}, t) dt - \int_I \frac{\partial f(\underline{x}, t)}{\partial x_j} dt \Big| \\ &\qquad = \Big| \int_I \int_0^1 \frac{\partial f(\underline{x} + sh\underline{e}_j, t)}{\partial x_j} - \frac{\partial f(\underline{x}, t)}{\partial x_j} ds\, dt \Big| \\ &\qquad \leq \int_I \int_0^1 \Big| \frac{\partial f(\underline{x} + sh\underline{e}_j, t)}{\partial x_j} - \frac{\partial f(\underline{x}, t)}{\partial x_j} ds\, dt \Big| \leq \varepsilon. \end{aligned}$$

Dies aber impliziert (16.2). □

Beispiel 16.3. Wir berechnen nun mit Hilfe von Theorem 16.1 für $x \geq 0$

$$\int_0^1 \frac{t^x - 1}{\log t} dt = \log(1 + x). \tag{16.3}$$

Hierfür schreiben wir

$$\frac{\partial}{\partial x} \int_0^1 \frac{t^x - 1}{\log t} dt = \int_0^1 \frac{\partial}{\partial x} \frac{e^{x \log t} - 1}{\log t} dt = \int_0^1 e^{x \log t} dt = \int_0^1 t^x dt = \frac{1}{x+1}.$$

Das bedeutet, dass auf beiden Seiten von (16.3) Funktionen stehen, die dieselben Ableitungen haben und an der Stelle $x = 0$ übereinstimmen. Also sind beide Funktionen gleich.

Als Anwendung bieten wir nun ein erstes Resultat, das die Vertauschbarkeit von Doppelintegralen begründet.

Proposition 16.4 (Fubini). *Seien $I = [a, b], J = [c, d]$ und $\underline{f} \in \mathcal{C}^0(I \times J)$. Dann gilt*

$$\int_a^b \Big(\int_c^d f(x, y) dy \Big) dx = \int_c^d \Big(\int_a^b f(x, y) dx \Big) dy.$$

Beweis. Wir definieren für $x \in [c, d]$

$$\varphi(x) := \int_a^x \Big(\int_c^d f(t, y) dy \Big) dt, \qquad \psi(x) := \int_c^d \Big(\int_a^x f(t, y) dt \Big) dy.$$

Ziel ist es also, $\varphi(d) = \psi(d)$ zu zeigen. Zunächst ist $\varphi(c) = \psi(c)$ klar. Weiter schreiben wir mit Hilfe von Theorem 16.1 und dem Hauptsatz der Differential- und Integralrechnung

$$\psi'(x) = \frac{d}{dx} \int_a^b \Big(\int_c^x f(t, y) dy \Big) dt = \int_a^b \frac{d}{dx} \Big(\int_c^x f(t, y) dy \Big) dt = \int_a^b f(t, x) dt = \varphi'(x).$$

□

Beispiel 16.5 (Gamma-Funktion). Im Beweis von Lemma 9.2 haben wir bereits die Vertauschung eines Doppelintegrals benutzt.

Oft hat man es mit uneigentlichen parameterabhängigen Integralen zu tun; siehe etwa die Beispiele 16.7 und 16.8. Diese können oft genauso unter dem Integral abgeleitet werden wie in Theorem 16.1.

Korollar 16.6 (Stetigkeit, Differenzierbarkeit parameterabhängiger uneigentlicher Integrale). *Sei $E \subseteq \mathbb{R}^m$ offen, $I = [a, b)$ mit $-\infty < a \leq b \leq \infty$ und $f \in \mathcal{C}^0(E \times I)$. Die partielle Ableitung $\frac{\partial f}{x_j}$ existiere für ein $j \in \{1, ..., m\}$ und sei stetig. Für jedes $\underline{x} \in E$ existiere das uneigentliche Integral $\int_I f(\underline{x}, t) dt$, außerdem konvergiere*

$$\lim_{\beta \uparrow b} \int_a^\beta \frac{\partial f(\underline{x}, t)}{\partial x_i} dt$$

uniform auf Kompakta gegen $\int_a^b \frac{\partial f(\underline{x},t)}{\partial x_i} dt$ (d.h. es gibt eine Funktion g, so dass $\sup_{\underline{x} \in K} \left| \int_a^\beta \frac{\partial f(\underline{x},t)}{\partial x_i} dt - g(\underline{x}) \right| \xrightarrow{\beta \uparrow b} 0$. Wir setzen dann $\int_a^b \frac{\partial f(\underline{x},t)}{\partial x_i} dt := g(\underline{x})$.) Dann gilt

$$\frac{\partial}{\partial x_j} \int_a^b f(\underline{x}, t) dt = \int_a^b \frac{\partial f(\underline{x}, t)}{\partial x_j} dt. \tag{16.4}$$

Beweis. Sei $(\beta_n)_{n=1,2,...}$ mit $\beta_n \uparrow b$ und

$$h_n(\underline{x}) := \int_a^{\beta_n} f(\underline{x}, t) dt.$$

Wir verwenden nun Proposition 11.3, da $h_1, h_2, ...$ stetig differenzierbar sind. Es konvergiert nämlich h_n für $n \to \infty$ punktweise, sowie $\frac{\partial h_n}{\partial x_j}$ gegen g gleichmäßig auf Kompakta. Deshalb folgt die Aussage aus Proposition 11.3 durch eine Anwendung von Theorem 16.1, denn

$$\int_a^b \frac{\partial f(\underline{x}, t)}{\partial x_j} dt = g(\underline{x}) = \frac{\partial}{\partial x_j} \int_a^b f(\underline{x}, t) dt.$$

□

Beispiel 16.7. Als weiteres Beispiel berechnen wir

$$\int_0^\infty \frac{\sin(t)}{t}dt = \frac{\pi}{2}.$$

(Dieses Integral existiert nach Beispiel 8.38.) Hierzu sei für $x > 0$

$$f(x) := \int_0^\infty e^{-xt}\frac{\sin(t)}{t}dt,$$

d.h. wir wollen den Wert $f(0)$ bestimmen. Wir schreiben

$$\int_0^b \frac{\partial}{\partial x}e^{-xt}\frac{\sin(t)}{t}dt = \int_0^b e^{-xt}\sin(t)dt \xrightarrow{b\to\infty}_{glmK} \int_0^\infty \frac{\partial}{\partial x}e^{-xt}\frac{\sin(t)}{t}dt,$$

wobei die gleichmäßige Konvergenz uniform auf Kompakta, also für $x \in K \subseteq (0,\infty)$ kompakt gilt. Damit ist Korollar 16.6 anwendbar. Wieder leiten wir f nach x ab und erhalten mittels partieller Integration für $x > 0$

$$\begin{aligned}\frac{df(x)}{dx} &= -\int_0^\infty e^{-xt}\sin(t)dt = \frac{1}{x}e^{-xt}\sin(t)|_{t=0}^{t=\infty} - \frac{1}{x}\int_0^\infty e^{-xt}\cos(t)dt \\ &= \frac{1}{x^2}e^{-xt}\cos(t)|_{t=0}^{t=\infty} + \frac{1}{x^2}\int_0^\infty e^{-xt}\sin(t)dt = -\frac{1}{x^2}\Big(1 - \frac{df(x)}{dx}\Big),\end{aligned}$$

woraus

$$\frac{df(x)}{dx} = -\frac{1}{1+x^2},$$

also

$$f(x) = -\arctan(x) + c$$

für ein geeignetes $c \in \mathbb{R}$ folgt. Wegen $f(x) \xrightarrow{x\to\infty} 0$ und $\arctan(x) \xrightarrow{x\to\infty} \frac{\pi}{2}$ muss $c = \frac{\pi}{2}$ sein. Insbesondere gilt also

$$\int_0^\infty \frac{\sin(t)}{t}dt = \lim_{x\to 0} f(x) = \frac{\pi}{2}.$$

Beispiel 16.8 (Gauss'sches Integral). Wir blicken zurück auf Proposition 9.3 und berechnen erneut das Integral

$$\int_{-\infty}^\infty e^{-\frac{1}{2}x^2}dx = \sqrt{2\pi}.$$

Mit $t = x/y$ ist

$$\begin{aligned}\frac{d}{dy}\Big(\int_0^y e^{-\frac{1}{2}x^2}dx\Big)^2 &= 2\int_0^y e^{-\frac{1}{2}(x^2+y^2)}dx = \int_0^1 2ye^{-\frac{1}{2}y^2(1+t^2)}dt = -2\int_0^1 \frac{\partial}{\partial y}\frac{e^{-\frac{1}{2}y^2(1+t^2)}}{1+t^2}dt \\ &= \frac{\partial}{\partial y}\Big(-2\int_0^1 \frac{e^{-\frac{1}{2}y^2(1+t^2)}}{1+t^2}dt\Big).\end{aligned}$$

Die Ableitung beider Seiten nach y ist also gleich. Außerdem gilt

$$\int_0^y e^{-\frac{1}{2}x^2}dx - \Big(-2\int_0^1 \frac{e^{-\frac{1}{2}y^2(1+t^2)}}{1+t^2}dt\Big)\Big|_{y=0} = 2\int_0^1 \frac{1}{1+t^2}dt = 2\arctan(1) = \frac{\pi}{2}.$$

Daraus folgern wir

$$\Big(\int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2}\,dx\Big)^2 = \lim_{y\to\infty} 4\Big(\int_0^y e^{-\frac{1}{2}x^2}\,dx\Big)^2 = 2\pi - 8\lim_{y\to\infty}\int_0^1 \frac{e^{-\frac{1}{2}y^2(1+t^2)}}{1+t^2}dt = 2\pi.$$

Daraus folgt der behauptete Wert des Gauss'schen Integrals.

16.2 Lokale Umkehrbarkeit von Funktionen

Von Funktionen $f : I \to \mathbb{R}$ für ein Intervall $I \subseteq \mathbb{R}$ weiß man: f ist in $x_0 \in I$ zumindest dann lokal umkehrbar (d.h. es gibt eine Umgebung $B_\delta(x_0)$ von x_0, so dass $f|_{B_\delta(x_0)}$ umkehrbar ist), wenn $f'(x_0) \neq 0$. Ziel dieses Abschnittes ist es, diese Aussage auf mehrdimensionale Funktionen $\underline{f}$ zu erweitern. Wir werden sehen, dass die Bedingung $f'(x_0) \neq 0$ dadurch ersetzt werden muss, dass die Ableitung $\underline{\underline{D}}\,\underline{f}(\underline{x}_0)$ vollen Rang hat, also invertierbar ist.

Definition 16.9 (Diffeomorphismus). *Seien $E \subseteq \mathbb{R}^m$, $E' \subseteq \mathbb{R}^n$ offene Mengen. Eine Funktion $\underline{f} : E \to E'$ heißt* Diffeomorphismus, *wenn $\underline{f} \in \mathcal{C}^1(E, E')$ und $\underline{f}$ umkehrbar ist mit $\underline{f}^{-1} \in \mathcal{C}^1(E', E)$.*

Beispiel 16.10 ($m = n = 1$).
1. Die Abbildung $f : \mathbb{R} \to (-\frac{\pi}{2}, \frac{\pi}{2})$, gegeben durch $f : t \mapsto \arctan(t)$ ist ein Diffeomorphismus. (Die Umkehrabbildung tan ist genau wie arctan differenzierbar.)
2. Die Abbildung $f : \mathbb{R} \to \mathbb{R}$, gegeben durch $f : t \mapsto t^3$ ist kein Diffeomorphismus. Schließlich ist die Umkehrabbildung $t \mapsto t^{1/3}$ für $t \geq 0$ und $t \mapsto -|t|^{1/3}$ für $t < 0$. Diese Abbildung hat keine stetige Ableitung.

Beispiel 16.11 ($m = n = 2$, Polarkoordinaten). Wir betrachten die Abbildung

$$\underline{f} : \begin{cases} \mathbb{R}_+ \times (0, 2\pi) & \to \mathbb{R}^2 \setminus \{(x,0) : x \geq 0\} \\ (r, \varphi) & \mapsto (r\cos\varphi, r\sin\varphi). \end{cases}$$

Setzt man

$$x = r\cos\varphi, \qquad y = r\sin\varphi,$$

so gilt also $x^2 + y^2 = r^2$, d.h.

$$r = \sqrt{x^2 + y^2}$$

sowie

$$\varphi = \arccos\frac{x}{\sqrt{x^2+y^2}} \text{ oder } \varphi = 2\pi - \arccos\frac{x}{\sqrt{x^2+y^2}},.$$

je nachdem ob $y \geq 0$ oder $y < 0$. Damit ist also

$$\underline{f}^{-1} : \begin{cases} \mathbb{R}^2 \setminus \{(x,0) : x \geq 0\} & \to \mathbb{R}_+ \times (0, 2\pi) \\ (x, y) & \mapsto \begin{cases} (\sqrt{x^2+y^2}, \arccos\frac{x}{\sqrt{x^2+y^2}}), & y \geq 0 \\ (\sqrt{x^2+y^2}, 2\pi - \arccos\frac{x}{\sqrt{x^2+y^2}}), & y < 0 \end{cases} \end{cases}$$

die Umkehrabbildung von $\underline{f}$. Da diese offenbar stetig differenzierbar ist, ist $\underline{f}$ ein Diffeomorphismus. Man spricht bei $(r\sin\varphi, r\cos\varphi)$ für $r \geq 0, \varphi \in (0, 2\pi]$ auch von der Polarkoordinatendarstellung eines Punktes in $\mathbb{R}^2$.

Lemma 16.12 (Grundlegende Eigenschaften von Diffeomorphismen). *Seien $E \subseteq \mathbb{R}^m$, $E' \subseteq \mathbb{R}^n$ offene Mengen und $\underline{f} : E \to E'$ ein Diffeomorphismus. Dann gilt $m = n$ und für $\underline{x} \in E$ ist $\underline{\underline{D}}\,\underline{f}(\underline{x})$ differenzierbar mit $\underline{\underline{D}}\,\underline{f}^{-1}(\underline{f}(\underline{x})) = (\underline{\underline{D}}\,\underline{f}(\underline{x}))^{-1}$.*

Beweis. Da $\underline{f}^{-1} \circ \underline{f} = \mathrm{id}$ und nach Voraussetzung sowohl $\underline{f}$ als auch $\underline{f}^{-1}$ differenzierbar sind, gilt nach der Kettenregel, Theorem 14.14,

$$\underline{\underline{E}}_m = \underline{\underline{D}}\,\mathrm{id}_E(\underline{x}) = \underline{\underline{D}}\,(\underline{f}^{-1} \circ \underline{f})(\underline{x}) = \underline{\underline{D}}\,\underline{f}^{-1}(\underline{f}(\underline{x})) \cdot \underline{\underline{D}}\,\underline{f}(\underline{x}).$$

Daraus folgt schon die angegebene Formel $\underline{\underline{D}}\,\underline{f}^{-1}(\underline{f}(\underline{x})) = (\underline{\underline{D}}\,\underline{f}(\underline{x}))^{-1}$. Analog berechnet man

$$\underline{\underline{E}}_n = \underline{\underline{D}}\,\underline{f}(\underline{f}^{-1}(\underline{y})) \cdot \underline{\underline{D}}\,\underline{f}^{-1}(\underline{y}).$$

Nach Kenntnissen aus der linearen Algebra ist dies nur dann möglich, wenn $m = n$. □

Lemma 16.13. *Seien $E, E' \subseteq \mathbb{R}^m$ offen und $\underline{f} \in \mathcal{C}^1(E, E')$. Gilt*

1. *Für alle $\underline{x} \in E$ ist $\underline{\underline{D}}\,\underline{f}(\underline{x})$ invertierbar,*
2. *$\underline{f}$ ist umkehrbar und $\underline{f}^{-1}$ ist stetig.*

Dann ist $\underline{f}$ ein Diffeomorphismus.

Bemerkung 16.14. Die Aussage des Lemmas ist also, dass man für die Umkehrabbildung eines Diffeomoprhismus nicht fordern muss, dass sie stetig differenzierbar ist, solange sie stetig ist.

Beweis. Nach (14.2) ist $\underline{f}$ in $\underline{x}$ genau dann differenzierbar, wenn

$$\underline{f}(\underline{x} + \underline{h}) - \underline{f}(\underline{x}) = (\underline{\underline{D}}\,\underline{f}(\underline{x}))(\underline{h}) + \underline{R}_{\underline{x}}(\underline{h}) \text{ mit } \lim_{\underline{h} \to 0} \frac{\underline{R}_{\underline{x}}(\underline{h})}{|\underline{h}|} = 0. \tag{16.5}$$

Für $\underline{y} \in E'$ und $\underline{h}'$ so, dass $\underline{y} + \underline{h}' \in E'$ gilt es nun, diese Eigenschaft für $\underline{f}^{-1}(\underline{y} + \underline{h}')$ zu zeigen. Wir setzen

$$\underline{x} = \underline{f}^{-1}(\underline{y}) \qquad \underline{h} = \underline{f}^{-1}(\underline{y} + \underline{h}') - \underline{f}^{-1}(\underline{y}).$$

Wir schreiben mit $\underline{f}(\underline{x} + \underline{h}) = \underline{f}(\underline{f}^{-1}(\underline{y} + \underline{h}')) = \underline{y} + \underline{h}'$

$$\begin{aligned} \underline{f}^{-1}(\underline{y} + \underline{h}') - \underline{f}^{-1}(\underline{y}) = \underline{h} &= (\underline{\underline{D}}\,\underline{f}(\underline{x}))^{-1}\big(\underline{f}(\underline{x} + \underline{h}) - \underline{f}(\underline{x}) - \underline{R}_{\underline{x}}(\underline{h})\big) \\ &= (\underline{\underline{D}}\,\underline{f}(\underline{x}))^{-1}(\underline{h}') + \underline{R}^*_{\underline{y}}(\underline{h}') \end{aligned} \tag{16.6}$$

mit

$$\underline{R}^*_{\underline{y}}(\underline{h}') := -(\underline{\underline{D}}\,\underline{f}(\underline{x}))^{-1}\underline{R}_{\underline{x}}(\underline{f}^{-1}(\underline{y} + \underline{h}') - \underline{f}^{-1}(\underline{y})).$$

Da $\underline{R}_{\underline{x}}(\underline{h})$ die Bedingung (16.5) erfüllt, gibt es ein $\varepsilon > 0$, so dass für $|\underline{h}| < \varepsilon$

$$|\underline{R}_{\underline{x}}(\underline{h})| \leq \frac{1}{2} \frac{|\underline{h}|}{|(\underline{\underline{D}}\underline{f}(\underline{x}))^{-1}|}$$

gilt. Weiter sei $\delta > 0$, so dass (wegen der Stetigkeit von $\underline{f}^{-1}$) für $|\underline{h}'| \leq \delta$

$$|\underline{h}| = |\underline{f}^{-1}(\underline{y} + \underline{h}') - \underline{f}^{-1}(\underline{y})| \leq \varepsilon.$$

Also gilt für $|\underline{h}'| \le \delta$

$$R_{\underline{y}}^*(\underline{h}') \le \left|(\underline{\underline{D}}\underline{f}(\underline{x}))^{-1}\right| \cdot |\underline{R}_{\underline{x}}(\underline{h})| \le \frac{1}{2}|\underline{f}^{-1}(\underline{y}+\underline{h}') - \underline{f}^{-1}(\underline{y})|.$$

Mit (18.2) ist also

$$|\underline{f}^{-1}(\underline{y}+\underline{h}') - \underline{f}^{-1}(\underline{y})| \le 2\left|(\underline{\underline{D}}\underline{f}^{-1}(\underline{x}))\right| \cdot |\underline{h}'|.$$

Damit ergibt sich

$$\begin{aligned}\frac{|\underline{R}_{\underline{y}}^*(\underline{h}')|}{|\underline{h}'|} &\le \left|(\underline{\underline{D}}\underline{f}(\underline{x}))^{-1}\right| \frac{|\underline{f}^{-1}(\underline{y}+\underline{h}') - \underline{f}^{-1}(\underline{y})|}{|\underline{h}'|} \frac{|\underline{R}_{\underline{x}}(\underline{h})|}{|\underline{f}^{-1}(\underline{y}+\underline{h}') - \underline{f}^{-1}(\underline{y})|} \\ &\le 2\left|(\underline{\underline{D}}\underline{f}(\underline{x}))^{-1}\right|^2 \frac{|\underline{R}_{\underline{x}}(\underline{h})|}{|\underline{h}|} \xrightarrow{|\underline{h}|,|\underline{h}'|\to 0} 0,\end{aligned}$$

also die Differenzierbarkeit von $\underline{f}^{-1}$. Aus (18.2) folgt weiter, dass $(\underline{\underline{D}}\underline{f}^{-1}(\underline{x})) = (\underline{\underline{D}}\underline{f}(\underline{x}))^{-1}$. Weiter ist die Abbildung $\underline{x} \mapsto (\underline{\underline{D}}\underline{f}(\underline{x}))^{-1}$ als Komposition stetiger Abbildungen stetig und damit ist $\underline{f}^{-1} \in \mathcal{C}^1(E')$. □

Theorem 16.15 (Lokale Umkehrbarkeit von Funktionen). *Seien $E, E' \subseteq \mathbb{R}^m$ offen und $\underline{f} \in \mathcal{C}^1(E, E')$. Ist für ein $\underline{x}_0 \in E$ das Differential $\underline{\underline{D}}\,\underline{f}(\underline{x}_0)$ invertierbar, dann gibt es eine offene Umgebung $U \subseteq E$ von $\underline{x}_0$ derart, dass $\underline{f}(U)$ eine Umgebung von $\underline{f}(\underline{x}_0)$ ist und die Abbildung $\underline{f}|_U : U \to \underline{f}(U), \underline{x} \mapsto \underline{f}(\underline{x})$ ein Diffeomorphismus ist. Für die Umkehrabbildung $\underline{f}^{-1}$ gilt $\underline{\underline{D}}\,\underline{f}^{-1}(\underline{f}(\underline{x})) = (\underline{\underline{D}}\,\underline{f}(\underline{x}))^{-1}$, $\underline{x} \in U$.*

Beweis. OBdA nehmen wir an, dass $\underline{x}_0 = \underline{f}(\underline{x}_0) = \underline{0}$ mit $\underline{\underline{D}}\,\underline{f}(\underline{x}_0) = \underline{\underline{E}}_m$ gilt. (Andernfalls gehen wir über zur Funktion $\underline{x} \mapsto (\underline{\underline{D}}\,\underline{f}(\underline{x}_0))^{-1}\big(\underline{f}(\underline{x}_0+\underline{x}) - \underline{f}(\underline{x}_0)\big)$. Diese Funktion hat dann die gewünschten Eigenschaften, die sich auf die ursprüngliche Funktion $\underline{f}$ übertragen lassen.)

Der Beweis ist eine Anwendung von Lemma 16.13. Wir gehen in zwei Schritten vor, die die beiden in diesem Lemma geforderten Eigenschaften der Funktion $\underline{f}$ auf einer geeigneten Menge U zeigen.

Schritt 1: $\underline{f}$ ist umkehrbar und $\underline{f}^{-1}$ ist stetig: Wir müssen sicherstellen, dass es für $\varepsilon > 0$ klein genug und $|\underline{y}| < \varepsilon$ ein $\underline{x}$ in einer Umgebung von $\underline{0}$ gibt mit $\underline{y} = \underline{f}(\underline{x})$, oder alternativ

$$\underline{\varphi}_{\underline{y}}(\underline{x}) = \underline{x} \qquad \text{für} \qquad \underline{\varphi}_{\underline{y}}(\underline{x}) := \underline{y} + \underline{x} - \underline{f}(\underline{x}).$$

Dadurch haben wir unsere Aufgabe in eine Fixpunktgleichung umgeschrieben, und der Beweis wird eine Anwendung des Banach'schen Fixpunktsatzes, Theorem 15.15. Um diesen anwenden zu können, müssen wir eine Menge $U \subseteq E$ finden, so dass $\underline{\varphi}_{\underline{y}}|_U$ eine Kontraktion ist. Hierfür sei $\varepsilon > 0$ klein genug, so dass $B_{2\varepsilon}(\underline{0}) \subseteq E$ und $|\underline{\underline{E}}_m - \underline{\underline{D}}\,\underline{f}(\underline{x})| \le \frac{1}{2}$ für $\underline{x} \in B_{2\varepsilon}(\underline{0})$. Dann gilt nämlich für $\underline{x} \in B_{2\varepsilon}(\underline{0})$

$$|\underline{\underline{D}}\,\underline{\varphi}_{\underline{y}}(\underline{x})| = |\underline{\underline{E}}_m - \underline{\underline{D}}\,\underline{f}(\underline{x})| \le \tfrac{1}{2},$$

also nach dem Schrankensatz, (15.3) und Bemerkung 15.9, für $\underline{x}, \underline{x}' \in B_{2\varepsilon}(\underline{0})$

$$|\underline{\varphi}_{\underline{y}}(\underline{x}') - \underline{\varphi}_{\underline{y}}(\underline{x})| \le \tfrac{1}{2}|\underline{x}' - \underline{x}|. \tag{16.7}$$

Weiter ist für $\underline{y} \in B_\varepsilon(\underline{0})$

$$|\underline{\varphi}_{\underline{y}}(\underline{x})| \leq |\underline{\varphi}_{\underline{y}}(\underline{x}) - \underline{\varphi}_{\underline{y}}(\underline{0})| + |\underline{\varphi}_{\underline{y}}(\underline{0})| \leq \tfrac{1}{2}|\underline{x} - \underline{0}| + |\underline{y}| \leq 2\varepsilon,$$

also $\varphi_{\underline{y}}(B_{2\varepsilon}(\underline{0})) \subseteq B_{2\varepsilon}(\underline{0})$. Damit ist gezeigt, dass $\underline{\varphi}_{\underline{y}}$ auf $B_{2\varepsilon}(\underline{0})$ eine Kontraktion ist, und somit gibt es genau ein $\underline{x}$ mit $|\underline{x}| < 2\varepsilon$ und $\underline{\varphi}_{\underline{y}}(\underline{x}) = \underline{x}$ oder äquivalent dazu $\underline{f}(\underline{x}) = \underline{y}$. Wir setzen nun $V := B_\varepsilon(\underline{0})$, $U := \underline{f}^{-1}(V) \cap B_{2\varepsilon}(\underline{0})$ und stellen fest, dass wir die Umkehrfunktion $\underline{f}^{-1}|_V : V \to U$ definiert haben.

Um die Stetigkeit von $\underline{f}^{-1}$ einzusehen, seien $\underline{y}, \underline{y}' \in V$. Dann gilt

$$\underline{f}^{-1}(\underline{y}') - \underline{f}^{-1}(\underline{y}) = \underline{\varphi}_{\underline{0}}(\underline{f}^{-1}(\underline{y}')) - \underline{\varphi}_{\underline{0}}(\underline{f}^{-1}(\underline{y})) + y' - \underline{y},$$

also mit (16.7)

$$|\underline{f}^{-1}(\underline{y}') - \underline{f}^{-1}(\underline{y})| \leq 2|y' - \underline{y}|,$$

woraus die Stetigkeit von $\underline{f}^{-1}$ folgt.
Schritt 2: Invertierbarkeit von $\underline{\underline{D}}\,\underline{f}(\underline{x})$: Mit (16.7) schreiben wir zunächst

$$\left|\left(\underline{\underline{E}}_m - \underline{\underline{D}}\,\underline{f}(\underline{x})\right)\frac{\underline{h}}{|\underline{h}|}\right| \leq |\underline{\underline{E}}_m - \underline{\underline{D}}\,\underline{f}(\underline{x})| \leq \tfrac{1}{2},$$

also

$$\left|\left(\underline{\underline{E}}_m - \underline{\underline{D}}\,\underline{f}(\underline{x})\right)\underline{h}\right| \leq \tfrac{1}{2}|\underline{h}|.$$

Gilt nun $\underline{\underline{D}}\,\underline{f}(\underline{x})\underline{h} = 0$, folgt daraus $|\underline{h}| \leq \frac{1}{2}|\underline{h}|$, was nur für $\underline{h} = 0$ möglich ist. Damit ist $\underline{\underline{D}}\,\underline{f}(\underline{x})$ invertierbar.

Die Formeln für die Ableitung der Umkehrabbildung haben wir bereits in Lemma 16.13 gezeigt. □

Korollar 16.16. *Seien $E, E' \subseteq \mathbb{R}^m$ offen und $\underline{f} \in \mathcal{C}^1(E, E')$ so, dass $\underline{\underline{D}}\,\underline{f}(\underline{x})$ für alle $\underline{x} \in E$ invertierbar ist. Dann ist $\underline{f}(E)$ offen. Ist außerdem $\underline{f}$ injektiv, so ist $\underline{f}$ ein Diffeomorphismus von E auf $\underline{f}(E)$.*

Beweis. Wir zeigen, dass $\underline{f}(E)$ als Vereinigung von offenen Mengen offen ist. Hierzu seien für $\underline{x} \in E$ Umgebungen $U_{\underline{x}}$, so dass $\underline{f}$ auf $U_{\underline{x}}$ ein Diffeomorphismus ist und $\underline{f}(U_{\underline{x}})$ eine Umgebung von $\underline{f}(\underline{x})$ ist. (Solche $U_{\underline{x}}$ gibt es nach Theorem 16.15.) Dann gilt $\underline{f}(E) = \bigcup_{\underline{x} \in E} \underline{f}(U_{\underline{x}})$ und ist somit als Vereinigung offener Mengen offen.

Für die zweite Behauptung genügt es nach Lemma 16.13 zu bemerken, dass $\underline{f}^{-1} : \underline{f}(E) \to E$ stetig ist. Um dies zu sehen sei $U \subseteq E$ offen. Dann gibt es wegen der Injektivität von $\underline{f}$ ein $U' \subseteq E'$ mit $\underline{f}(U) = U'$ und $\underline{f}^{-1}(U') = U$. Nach der ersten Behauptung ist also U' offen und $\underline{f}^{-1}$ ist damit stetig. Aus Lemma 16.13 folgt, dass $\underline{f}$ ein Diffeomorphismus ist. □

Beispiel 16.17 (Polarkoordinaten). Wir greifen nochmals die Polarkoordinaten aus Beispiel 16.11 auf. Insbesondere ist

$$\underline{f} : \begin{cases} \mathbb{R}_+ \times (0, 2\pi) & \to \mathbb{R}^2 \setminus \{(x, 0) : x \geq 0\} \\ (r, \varphi) & \mapsto (r\cos\varphi, r\sin\varphi). \end{cases}$$

Wir berechnen die Jacobi-Determinante von $\underline{f}$

$$\underline{\underline{D}}\,\underline{f}(r,\varphi) = \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix}.$$

Insbesondere ist also

$$\det(\underline{\underline{D}}\,\underline{f}(r,\varphi)) = r.$$

Damit ist $\underline{\underline{D}}\,\underline{f}(r,\varphi)$ für alle $(r,\varphi) \in (0,\infty)\times(0,2\pi)$ invertierbar und aus Korollar 16.16 folgt, dass $\underline{f}$ ein Diffeomorphismus ist. Dies wissen wir zwar schon aus Beispiel 16.11, mussten hierfür jedoch die Umkehrfunktion berechnen.

Beispiel 16.18 (Kugelkoordinaten). Als höher-dimensionales Analogon der Polarkoordinaten gibt es die Kugelkoordinaten. Hierfür definieren wir

$$\underline{f} : \begin{cases} \mathbb{R}_+ \times (0,\pi) \times (0,2\pi) & \to \mathbb{R}^3 \setminus \{(x,0,z) : x \le 0\} \\ (r,\theta,\varphi) & \mapsto (r\sin\theta\cos\varphi, r\sin\theta\sin\varphi, r\cos\theta). \end{cases}$$

Wir berechnen die Jacobi-Matrix

$$\underline{\underline{D}}\,\underline{f}(r,\theta,\varphi) = \begin{pmatrix} \sin\theta\cos\varphi & r\cos\theta\cos\varphi & -r\sin\theta\sin\varphi \\ \sin\theta\sin\varphi & r\cos\theta\sin\varphi & r\sin\theta\cos\varphi \\ \cos\theta & -r\sin\theta & 0 \end{pmatrix}.$$

Damit gilt

$$\det(\underline{\underline{D}}\,\underline{f}(r,\theta,\varphi)) = r^2\cos^2\theta\sin\theta + r^2\sin^3\theta = r^2\sin\theta.$$

Damit ist also $\underline{f}$ ein Diffeomorphismus.

16.3 Implizite Funktionen

Sei $\underline{f} : E \times E' \subseteq \mathbb{R}^m \times \mathbb{R}^n \to E'' \subseteq \mathbb{R}^n$ mit $\underline{f}(\underline{x},\underline{y}) = \underline{0}$ für ein $(\underline{x},\underline{y}) \in E \times E'$. Wir fragen, ob es eine Umgebung U von $\underline{x}$ und eine Funktion $\underline{h} : U \subseteq \mathbb{R}^m \to \mathbb{R}^n$ gibt mit $\underline{f}(\underline{x},\underline{h}(\underline{x})) = \underline{0}, \underline{x} \in U$. Ist dies der Fall so sagen wir, dass $\underline{h}$ durch $\underline{f}(\underline{x},\underline{h}(\underline{x})) = \underline{0}$ implizit definiert ist. Siehe Abbildung 16.1 für eine Illustration.

Bemerkung 16.19 (Notation). Sei $\underline{f} \in \mathcal{C}^1(E \times E', E'')$ für $E \subseteq \mathbb{R}^m, E', E'' \in \mathbb{R}^n$. Für $\underline{x} \in E$ sei weiter $\underline{f}_{\underline{x}} : \underline{y} \mapsto \underline{f}(\underline{x},\underline{y})$ und für $\underline{y} \in E'$ sei $\underline{f}_{\underline{y}} : \underline{x} \mapsto \underline{f}(\underline{x},\underline{y})$. Wir definieren für $\underline{h} \in \mathbb{R}^m, \underline{h}' \in \mathbb{R}^n$

$$\underline{\underline{D}}_{\underline{x}}\underline{f}(\underline{x},\underline{y})(\underline{h}) := \underline{\underline{D}}\,\underline{f}(\underline{x},\underline{y})(\underline{h},\underline{0}),$$
$$\underline{\underline{D}}_{\underline{y}}\underline{f}(\underline{x},\underline{y})(\underline{h}') := \underline{\underline{D}}\,\underline{f}(\underline{x},\underline{y})(\underline{0},\underline{h}').$$

Hierfür schreiben wir auch in der Jacobi-Matrix

$$\underline{\underline{D}}\,\underline{f}(\underline{x},\underline{y}) = \left(\underline{\underline{D}}_{\underline{x}}\underline{f}(\underline{x},\underline{y}) \,\middle|\, \underline{\underline{D}}_{\underline{y}}\underline{f}(\underline{x},\underline{y})\right)$$

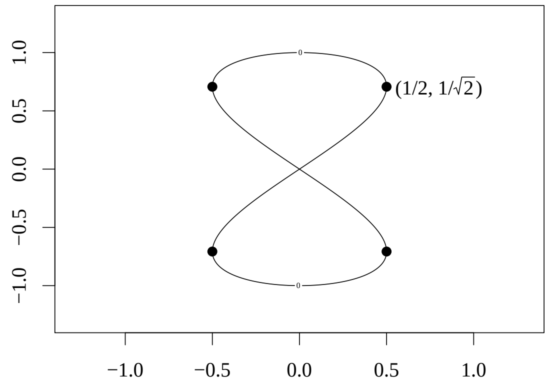

Abbildung 16.1: Um eine Abbildung $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ darzustellen, kann man einen Contour-Plot anfertigen. Hierfür betrachtet man die Punktmenge $\{(x, y, f(x, y)) : x, y \in \mathbb{R}\}$ als *Gebirgslandschaft* und zeichnet jeweils die Höhenlinien, also Linien mit gleichem Wert $f(x, y)$, ein. Anders ausgedrückt stellt man die Funktion durch eine Darstellung von $f(x, y) = c$ für $c \in \mathbb{R}$ dar. Mit anderen Worten sucht man etwa eine Funktion $g_c : \mathbb{R} \to \mathbb{R}$, so dass $f(x, g_c(x)) = c$ für alle (oder viele) x. Im rechten Bild sieht man die Höhenlinie von $f(x, y) = 0$. Hier erkennt man, dass es nicht überall eine lokal definierte Funktion $g_0 : x \mapsto g_0(x)$, so dass $f(x, g_0(x)) = 0$ geben kann.

mit

$$\underline{\underline{D}}_{\underline{x}} \underline{f}(\underline{x}, \underline{y}) = \begin{pmatrix} \frac{\partial f_1(\underline{x},\underline{y})}{\partial x_1} & \cdots & \frac{\partial f_1(\underline{x},\underline{y})}{\partial x_m} \\ \vdots & & \vdots \\ \frac{\partial f_n(\underline{x},\underline{y})}{\partial x_1} & \cdots & \frac{\partial f_n(\underline{x},\underline{y})}{\partial x_m} \end{pmatrix}, \qquad \underline{\underline{D}}_{\underline{y}} \underline{f}(\underline{x}, \underline{y}) = \begin{pmatrix} \frac{\partial f_1(\underline{x},\underline{y})}{\partial y_1} & \cdots & \frac{\partial f_1(\underline{x},\underline{y})}{\partial y_n} \\ \vdots & & \vdots \\ \frac{\partial f_n(\underline{x},\underline{y})}{\partial y_1} & \cdots & \frac{\partial f_n(\underline{x},\underline{y})}{\partial y_n} \end{pmatrix}.$$

Also ist auch

$$\begin{aligned} \underline{\underline{D}}\, \underline{f}(\underline{x}, \underline{y})(\underline{h}|\underline{h}') &= \left(\underline{\underline{D}}_{\underline{x}} \underline{f}(\underline{x}, \underline{y}) \middle| \, \underline{\underline{D}}_{\underline{y}} \underline{f}(\underline{x}, \underline{y})\right) \begin{pmatrix} \underline{h} \\ \hline \underline{h}' \end{pmatrix} \\ &= \underline{\underline{D}}_{\underline{x}} \underline{f}(\underline{x}, \underline{y})(\underline{h}) + \underline{\underline{D}}_{\underline{y}} \underline{f}(\underline{x}, \underline{y})(\underline{h}'). \end{aligned}$$

Theorem 16.20 (Satz über implizite Funktionen). *Seien $E \subseteq \mathbb{R}^m, E', E'' \subseteq \mathbb{R}^n$ offen, $\underline{f} \in \mathcal{C}^1(E \times E', E'')$ und $(\underline{x}_0, \underline{y}_0) \in E \times E'$ so, dass $\underline{f}(\underline{x}_0, \underline{y}_0) = \underline{0}$. Ist $\underline{\underline{D}}_{\underline{y}} \underline{f}(\underline{x}_0, \underline{y}_0)$ invertierbar, so gibt es eine Umgebung $V \times V'$ von $(\underline{x}_0, \underline{y}_0)$ und eine Funktion $\underline{h} \in \mathcal{C}^1(V, V')$, so dass*

$$\underline{f}(\underline{x}, \underline{y}) = \underline{0} \text{ für } (\underline{x}, \underline{y}) \in V \times V' \text{ genau dann, wenn } \underline{y} = \underline{h}(\underline{x}) \text{ für } \underline{x} \in V.$$

Weiter gilt

$$\underline{\underline{D}}\, \underline{h}(\underline{x}) = -\left(\underline{\underline{D}}_{\underline{y}} \underline{f}(\underline{x}_0, \underline{y}_0)\right)^{-1} \cdot \underline{\underline{D}}_{\underline{x}} \underline{f}(\underline{x}_0, \underline{y}_0).$$

Beweis. Wir spielen die Aussage des Satzes zurück auf den Satz über lokale Umkehrbarkeit von Funktionen, Theorem 16.15. Hierzu definieren wir $\underline{\varphi} : E \times E' \to E \times E''$ mittels

$$\underline{\varphi}(\underline{x}, \underline{y}) := (\underline{x}, \underline{f}(\underline{x}, \underline{y})).$$

Dann ist

$$\underline{\underline{D}}\,\underline{\varphi}(\underline{x},\underline{y}) = \left(\begin{array}{c|c} \underline{\underline{E}}_m & \underline{\underline{0}} \\ \hline \underline{\underline{D}}_{\underline{x}}\underline{f}(\underline{x},\underline{y}) & \underline{\underline{D}}_{\underline{y}}\underline{f}(\underline{x},\underline{y}) \end{array}\right).$$

Aus dieser Darstellung sieht man insbesondere, dass $\underline{\underline{D}}\,\underline{\varphi}(\underline{x}_0,\underline{y}_0)$ invertierbar ist. Nach Theorem 16.15 gibt es also offene Mengen $U \subseteq E$, $U' \subseteq E'$, so dass $(\underline{x}_0,\underline{y}_0) \in U \times U'$ und $\underline{\varphi}$ auf $U \times U'$ ein Diffeomorphismus ist. Betrachten wir also die Umkehrabbildung $\underline{\varphi}^{-1}$, es muss nämlich für ein geeignetes $\underline{g} : \underline{f}(U \times U') \to E''$

$$\underline{\varphi}^{-1}(\underline{x},\underline{y}) = (\underline{x},\underline{g}(\underline{x},\underline{y}))$$

gelten. Nun folgern wir

$$\begin{aligned} \underline{f}(\underline{x},\underline{y}) = \underline{0} \quad &\Longleftrightarrow \quad \underline{\varphi}(\underline{x},\underline{y}) = (\underline{x},\underline{0}) \quad \Longleftrightarrow \quad (\underline{x},\underline{y}) = \underline{\varphi}^{-1}(\underline{x},\underline{0}) \\ &\Longleftrightarrow \quad \underline{y} = \underline{g}(\underline{x},\underline{0}). \end{aligned}$$

Nun ist mit $\underline{\varphi}^{-1}$ auch $\underline{g}$ stetig und es gilt $\underline{g}(\underline{x}_0,\underline{0}) = \underline{y}_0$. Also gibt es Umgebungen $V \subseteq U$ von $\underline{x}_0$ und $V' \subseteq U'$ von $\underline{y}_0$, so dass $\underline{g}(V,\underline{0}) \subseteq V'$. Wir setzen also $\underline{h} : V \to V'$ mit $\underline{h}(\underline{x}) := \underline{g}(\underline{x},\underline{0})$ und haben alle Behauptungen erfüllt.

Um das Differential $\underline{\underline{D}}\,\underline{h}(\underline{x})$ zu berechnen, verwenden wir $\underline{f}(\underline{x},\underline{h}(\underline{x})) = \underline{f}(\underline{x},\underline{g}(\underline{x},\underline{0})) = \underline{0}$ und die Kettenregel mit $\underline{\psi}(\underline{x}) = (\underline{x},\underline{h}(\underline{x}))$, um

$$\underline{0} = \underline{\underline{D}}(\underline{f} \circ \underline{\psi})(\underline{x}) = \underline{\underline{D}}\,\underline{f}(\underline{x},\underline{h}(\underline{x})) \cdot \underline{\underline{D}}\,\psi(\underline{x}) = \left(\underline{\underline{D}}_{\underline{x}}\underline{f}(\underline{x},\underline{h}(\underline{x}))\middle|\,\underline{\underline{D}}_{\underline{y}}\underline{f}(\underline{x},\underline{h}(\underline{x}))\right) \cdot \left(\begin{array}{c} \underline{\underline{E}}_m \\ \hline \underline{\underline{D}}\,\underline{h}(\underline{x}) \end{array}\right)$$

zu erhalten. Aus dieser Gleichung folgt sofort die behauptete Form von $\underline{\underline{D}}\,\underline{h}(\underline{x})$. □

Beispiel 16.21 (Abbildung 16.1). Wir betrachten nochmals das Beispiel aus Abbildung 16.1. Hier haben wir

$$f(x,y) = y^2(1-y^2) - x^2$$

gesetzt. Um Theorem 16.20 anwenden zu können, bemerken wir, dass

$$\begin{aligned} D_x f(x,y) &= -2x, \\ D_y f(x,y) &= 2y(1-y^2) - 2y^3 = 2y(1-2y^2). \end{aligned}$$

Insbesondere ist also $D_y f$ genau dann nicht invertierbar, wenn entweder $y = 0$ oder $y = \pm 1/\sqrt{2}$. Also liefert das obige Resultat, dass man außer in den Punkten

$$(0,0), (\tfrac{1}{2}, \tfrac{1}{\sqrt{2}}), (\tfrac{1}{2}, -\tfrac{1}{\sqrt{2}}), (-\tfrac{1}{2}, \tfrac{1}{\sqrt{2}}), (-\tfrac{1}{2}, -\tfrac{1}{\sqrt{2}})$$

eine Funktion $h_0 : x \mapsto h_0(x)$ lokal so definieren kann, dass $f(x, h_0(x)) = 0$.

Beispiel 16.22 (Lineare Abbildung). Sei $\underline{\underline{B}} \in \mathbb{R}^{n\times n}$ und $\underline{g} : \mathbb{R}^m \to \mathbb{R}^n$. Betrachte die Funktion

$$\underline{f}(\underline{x},\underline{y}) := \underline{g}(\underline{x}) + \underline{\underline{B}}\,\underline{y}.$$

Sei $\underline{x} \in \mathbb{R}^m$. Wie aus der linearen Algebra bekannt, hat die Gleichung

$$f(\underline{x}, \underline{y}) = \underline{0}$$

genau dann sicher eine Lösung $\underline{y} \in \mathbb{R}^n$, wenn $\underline{\underline{B}}$ invertierbar ist. (Dann ist sie nämlich $\underline{y} = \underline{\underline{B}}^{-1}\underline{g}(\underline{x})$.) Diese Tatsache kann man auch aus Theorem 16.20 ablesen. Es gilt nämlich

$$\begin{aligned} \underline{\underline{D}}_{\underline{x}}\underline{f}(\underline{x}, \underline{y}) &= \underline{\underline{D}}\,\underline{g}(\underline{x}), \\ \underline{\underline{D}}_{\underline{y}}\underline{f}(\underline{x}, \underline{y}) &= \underline{\underline{B}}. \end{aligned}$$

Damit hat $\underline{f}(\underline{x}, \underline{y}) = \underline{0}$ eine Lösung $\underline{h}(\underline{x})$, wenn $\underline{\underline{B}}$ invertierbar ist. Für deren Ableitung gilt außerdem $\underline{\underline{D}}\,\underline{h}(\underline{x}) = -\underline{\underline{B}}^{-1} \cdot \underline{\underline{D}}\,\underline{g}(\underline{x})$.

16.4 Extrema unter Nebenbedingungen

Die Berechnung von Extremwerten hatten wir bereits in Abschnitt für Funktionen $f : \mathbb{R} \to \mathbb{R}$ und in Abschnitt 14.2 für Funktionen $f : \mathbb{R}^m \to \mathbb{R}$ behandelt. Nun wollen wir uns einem weiteren Fall zuwenden: Gesucht ist $\underline{x} \in \mathbb{R}^m$, der (i) Extremwert einer Funktion $f : \mathbb{R}^m \to \mathbb{R}$ ist, als auch einer *Nebenbedingung* $g(\underline{x}) = 0$ genügt. Wir beginnen mit einem elementar zu berechnenden Beispiel.

Beispiel 16.23 (Minimaler Abstand einer Geraden vom Ursprung). Wir beantworten nun folgende Frage:

> Welcher Punkt (x, y), der auf der Geraden $y = mx + b$ liegt, hat minimalen Abstand zum Ursprung?

(Siehe auch Abbildung 16.2 für eine Illustration.) Hierfür definieren wir $f : \mathbb{R}^2 \to \mathbb{R}$ durch

$$f(x, y) = x^2 + y^2,$$

was gerade dem Quadrat des Abstandes des Punktes (x, y) vom Ursprung entspricht. Außerdem betrachten wir für $b, m \in \mathbb{R}$ die Funktion

$$(x, y) \mapsto g(x, y) = y - mx - b.$$

Somit ist $g(x, y) = 0$ die Gerade durch den Punkt $(0, b)$ mit Steigung m. Obige Frage übersetzt sich also wiefolgt:

> Welcher Punkt (x, y) ist ein Minimum der Funktion $(x, y) \mapsto f(x, y)$ unter der Nebenbedingunge $g(x, y) = 0$?

In diesem Fall lässt sich dieses Minimum unter Nebenbedingung elementar berechnen: Da für das gesuchte (x_0, y_0) nämlich $y_0 = mx_0 + b$ gelten muss, genügt es, die Funktion

$$x \mapsto x^2 + (mx + b)^2 = (m^2 + 1)x^2 + 2bmx + b^2$$

zu minimieren. Leitet man einmal ab und setzt das Ergebnis 0, muss also $2(m^2+1)x_0+2bm = 0$ gelten, also findet sich das Minimum bei

$$x_0 = -\frac{bm}{m^2 + 1}, \text{ und damit } y_0 = mx_0 + b = -\frac{bm^2 - bm^2 - b}{m^2 + 1} = \frac{b}{m^2 + 1}$$

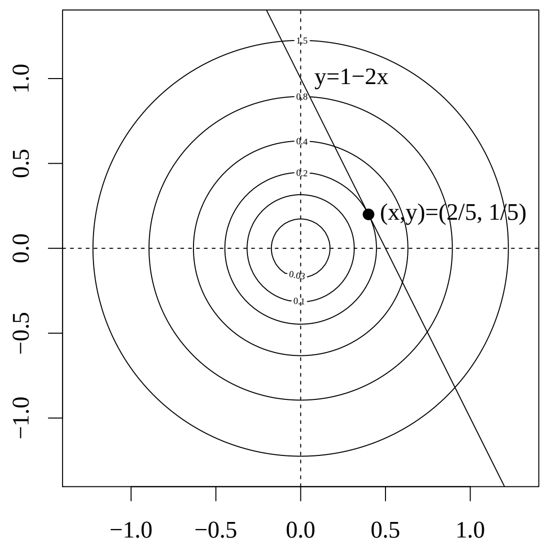

Abbildung 16.2: Die Höhenlinien der Funktion $f : (x, y) \mapsto x^2 + y^2$ sind eingezeichnet. Gesucht ist ein Punkt (x, y) auf der Gerade $y = 1 - 2x$, der den Abstand zu Ursprung, also f, minimiert. Im Beispiel ist dieser durch $(\frac{2}{5}, \frac{1}{5})$ gegeben.

gelten. Das Minimum ist dann außerdem $f(x_0, y_0) = \frac{b^2}{m^2+1}$. Weiterhin bemerken wir, dass

$$\begin{aligned}\frac{\partial f(x_0, y_0)}{\partial x} &= 2x|_{x=x_0} = -\frac{2bm}{m^2+1} = -\lambda m = \lambda \frac{\partial g(x_0, y_0)}{\partial x},\\ \frac{\partial f(x_0, y_0)}{\partial y} &= 2y|_{y=y_0} = \frac{2b}{m^2+1} = \lambda = \lambda \frac{\partial g(x_0, y_0)}{\partial y}\end{aligned}$$

für $\lambda = \frac{2b}{m^2+1}$. Diese lineare Abhängikeit der partiellen Ableitungen von f und g wird nun zentral in der allgemeinen Formulierung von Extrema unter Nebenbedingungen.

Theorem 16.24 (Lagrange-Multiplikatoren). *Sei $E \subseteq \mathbb{R}^m$ offen, $f \in \mathcal{C}^1(E)$, $\underline{g} \in \mathcal{C}^1(E, \mathbb{R}^n)$ und*

$$S := \{\underline{x} \in E : \underline{g}(\underline{x}) = \underline{0}\}.$$

Sei $\underline{x}_0 \in S$ so, dass $f|_S$ in $\underline{x}_0 \in S$ ein lokales Extremum hat. Weiter seien $\underline{\nabla} g_1(\underline{x}_0), ..., \underline{\nabla} g_n(\underline{x}_0)$ linear unabhängig. Dann existieren $\lambda_1, ..., \lambda_n \in \mathbb{R}$ mit

$$\frac{\partial f(\underline{x}_0)}{\partial x_i} = \sum_{k=1}^{n} \lambda_k \frac{\partial g_k(\underline{x}_0)}{\partial x_i}, \qquad i = 1, ..., m. \tag{16.8}$$

Die Zahlen $\lambda_1, ..., \lambda_n$ heißen auch Lagrange-Multiplikatoren.

Bemerkung 16.25 (Anwendung von Theorem 16.24). Wir beschreiben kurz, wie eine typische Anwendung des Theorems aussehen kann. Zunächst fällt auf, dass die Existenz der Lagrange-Multiplikatoren nur eine notwendige, jedoch keine hinreichende für ein Extremum bieten.

Wir definieren zunächst die Funktion

$$F(\underline{x}, \underline{\lambda}) = f(\underline{x}) - \langle \underline{\lambda}, \underline{g}(\underline{x}) \rangle, \tag{16.9}$$

wobei $\langle .,. \rangle$ das Standard-Skalar-Produkt darstellt. Dann ist die Bedingung $\underline{g}(\underline{x}) = \underline{0}$ äquivalent zu $\underline{D}_{\underline{\lambda}} F(\underline{x}, \underline{\lambda}) = \underline{0}$ und die Bedingung (16.8) wird zu $\underline{D}_{\underline{x}} F(\underline{x}, \underline{\lambda}) = \underline{0}$. Insgesamt gilt es also, $(\underline{x}, \underline{\lambda})$ zu finden, die die Gleichung

$$\underline{D}\, F(\underline{x}, \underline{\lambda}) = \underline{0}$$

lösen. Sind $\underline{x}_1, ..., \underline{x}_p$ alle Lösungen in E, so sind neben diesen Punkten noch die Punkte $\underline{z}$ mit $\operatorname{rg}(\underline{\underline{D}}\, \underline{g}(\underline{z})) < n$ mögliche Kandidaten für Extrema von f auf S. Es gilt nun, alle diese Punkte einzeln und direkt auf die Extremaleigenschaft der Funktion f zu prüfen.

Besonders einfach ist der Fall, wenn bei diesem Vorgehen nur ein einziger Punkt $\underline{x}_0$ als möglicher Extremwert übrig bleibt. Kann man dann noch argumentieren, dass f auf S ein Extremum haben muss, ist dieses Extremum bereits gefunden. Beispielsweise zeigt man dafür, dass $E \subseteq K$ mit K kompakt (so dass f auf K sowohl ein Maximum als auch ein Minimum hat) und f auf $K \setminus E$ kein Extremum annimmt.

Beweis von Theorem 16.24. Nach Voraussetzung hat die Matrix $\underline{\underline{D}}\, \underline{g}(\underline{x}_0)$ den Zeilenrang n, und damit auch den Spaltenrang n. Ohne Einschränkung der Allgemeinheit seien dies die n letzten Spalten der Matrix $\underline{\underline{D}}\, \underline{g}(\underline{x}_0)$. Wir schreiben nun $\underline{x} = (\underline{y}, \underline{z})$ mit $\underline{y} \in \mathbb{R}^{m-n}$, $\underline{z} \in \mathbb{R}^n$, und auch $\underline{x}_0 = (\underline{y}_0, \underline{z}_0)$ mit $\underline{z}_0 \in \mathbb{R}^n$. Wie in Bemerkung 16.19 haben wir

$$\underline{\underline{D}}\, \underline{g}(\underline{x}_0) = (\underline{\underline{D}}_{\underline{y}}\, \underline{g}(\underline{y}_0, \underline{z}_0) | \underline{\underline{D}}_{\underline{z}}\, \underline{g}(\underline{y}_0, \underline{z}_0))$$

mit

$$\underline{\underline{D}}_{\underline{y}}\, \underline{g}(\underline{y}, \underline{z}) = \begin{pmatrix} \frac{\partial g_1(\underline{y},\underline{z})}{\partial y_1} & \cdots & \frac{\partial g_1(\underline{y},\underline{z})}{\partial y_{m-n}} \\ \vdots & & \vdots \\ \frac{\partial g_n(\underline{y},\underline{z})}{\partial y_1} & \cdots & \frac{\partial g_n(\underline{y},\underline{z})}{\partial y_{m-n}} \end{pmatrix}, \qquad \underline{\underline{D}}_{\underline{z}}\, \underline{g}(\underline{y}, \underline{z}) = \begin{pmatrix} \frac{\partial g_1(\underline{y},\underline{z})}{\partial z_1} & \cdots & \frac{\partial g_1(\underline{y},\underline{z})}{\partial z_n} \\ \vdots & & \vdots \\ \frac{\partial g_n(\underline{y},\underline{z})}{\partial z_1} & \cdots & \frac{\partial g_n(\underline{y},\underline{z})}{\partial z_n} \end{pmatrix},$$

wobei $\underline{\underline{D}}_{\underline{z}}\, \underline{g}(\underline{y}_0, \underline{z}_0)$ invertierbar ist, weil die letzten n Spalten der Matrix $\underline{\underline{D}}\, \underline{g}(\underline{x}_0)$ linear unabhängig sind.

Da $\underline{g}(\underline{x}_0) = \underline{0}$, können wir also den Satz über implizite Funktionen, Theorem 16.20, auf die Funktion $\underline{g}$ anwenden. Das bedeutet, dass es eine Umgebung $V \times V' \subseteq \mathbb{R}^{m-n} \times \mathbb{R}^n$ von $(\underline{y}_0, \underline{z}_0)$ und eine Funktion $\underline{h} \in \mathcal{C}^1(V, V')$ gibt, so dass $\underline{x} = (\underline{y}, \underline{z}) \in S$ genau dann gilt wenn $\underline{z} = \underline{h}(\underline{y})$. Für diese Funktion $\underline{h}$ gilt dann

$$\underline{\underline{D}}\, \underline{h}(\underline{y}) = -\big(\underline{\underline{D}}_{\underline{z}} \underline{g}(\underline{y}, \underline{h}(\underline{y}))\big)^{-1} \cdot \underline{\underline{D}}_{\underline{y}} \underline{g}(\underline{y}, \underline{h}(\underline{y})).$$

Nach Voraussetzung hat die Funktion $(\underline{y}, \underline{z}) \mapsto f(\underline{y}, \underline{z})$ auf S in $(\underline{y}_0, \underline{z}_0)$ ein lokales Extremum. Damit hat auch die Funktion

$$\varphi : \underline{y} \mapsto f(\underline{y}, \underline{h}(\underline{y}))$$

in $\underline{y}_0$ ein lokales Extremum. (Denn: Für $(\underline{y}, \underline{z})$ in einer Umgebung von $(\underline{y}_0, \underline{z}_0)$m geschnitten mit S ist f in $(\underline{y}_0, \underline{z}_0)$ extremal. Da sich – wegen der Stetigkeit von $\underline{h}$ – Punkte $(\underline{y}, \underline{h}(\underline{y}))$ in einer beliebig kleinen Umgebung von $(\underline{y}_0, \underline{z}_0)$ befinden, gibt es eine Umgebung von $\underline{y}$, in der $\underline{y} \mapsto f(\underline{y}, \underline{h}(\underline{y}))$ in $\underline{y}_0$ extremal ist.) Damit muss $\underline{D}\varphi(\underline{y}_0) = \underline{0}$ gelten. Daraus folgern wir mit der Kettenregel und $\underline{\psi}(\underline{y}) := (\underline{y}, \underline{h}(\underline{y}))$

$$\begin{aligned}\underline{0} &= \underline{D}\varphi(\underline{y}_0) = \underline{D}(f \circ \underline{\psi})(\underline{y}_0) = \underline{D}f(\underline{y}_0, \underline{h}(\underline{y}_0)) \cdot \underline{\underline{D}}\,\underline{\psi}(\underline{y}_0) \\ &= \left(\underline{D}_{\underline{y}} f(\underline{y}_0, \underline{h}(\underline{y}_0)) | \underline{D}_{\underline{z}} f(\underline{y}_0, \underline{h}(\underline{y}_0))\right) \cdot \begin{pmatrix} \underline{\underline{E}}_{m-n} \\ \hline \underline{\underline{D}}\,\underline{h}(\underline{y}_0) \end{pmatrix} \\ &= \underline{D}_{\underline{y}} f(\underline{y}_0, \underline{z}_0) - \underline{D}_{\underline{z}} f(\underline{y}_0, \underline{z}_0) \cdot \left(\underline{\underline{D}}_{\underline{z}} \underline{g}(\underline{y}_0, \underline{z}_0)\right)^{-1} \cdot \underline{\underline{D}}_{\underline{y}} \underline{g}(\underline{y}_0, \underline{z}_0).\end{aligned} \tag{16.10}$$

Wir setzen nun

$$\underline{\lambda} = (\lambda_1, ..., \lambda_n) := \underline{D}_{\underline{z}} f(\underline{y}_0, \underline{z}_0) \cdot \left(\underline{\underline{D}}_{\underline{z}} \underline{g}(\underline{y}_0, \underline{z}_0)\right)^{-1} \in \mathbb{R}^n \tag{16.11}$$

und stellen fest, dass mit (16.10)

$$\underline{D}_{\underline{y}} f(\underline{x}_0) = \underline{\lambda} \cdot \underline{\underline{D}}_{\underline{y}} \underline{g}(\underline{x}_0),$$

also (16.8) für $i = 1, ..., m - n$, und wegen (16.11)

$$\underline{D}_{\underline{z}} f(\underline{x}_0) = \underline{\lambda} \cdot \underline{\underline{D}}_{\underline{z}} \underline{g}(\underline{x}_0),$$

also (16.8) für $i = m - n + 1, ..., m$. □

Beispiel 16.26 (Minimaler Abstand einer Geraden vom Ursprung). Wir betrachten noch einmal das Beispiel 16.23. Hier wollten wir die Funktion $f(x, y) := x^2 + y^2$ unter der Nebenbedingung $g(x, y) := y - mx - b = 0$ minimieren. Geonmetrisch ist klar, dass es genau ein Minimum geben muss. Wie in Bemerkung 16.25 setzen wir

$$F(x, y, \lambda) := x^2 + y^2 - \lambda(y - mx - b)$$

und berechnen die Ableitungen

$$\begin{aligned}\frac{\partial F(x, y, \lambda)}{\partial x} &= 2x + \lambda m, \\ \frac{\partial F(x, y, \lambda)}{\partial y} &= 2y - \lambda, \\ \frac{\partial F(x, y, \lambda)}{\partial \lambda} &= y - mx - b.\end{aligned}$$

Dieses in x, y, λ lineare Gleichungssystem hat genau eine Lösung, nämlich

$$x = -\frac{-bm}{m^2 + 1}, \qquad y = \frac{2b}{m^2 + 1}, \qquad \lambda = \frac{2b}{m^2 + 1}.$$

Mit anderen Worten haben wir nochmals dasselbe Minimum unter $g(x, y) = 0$ berechnet wie in Beispiel 16.23.

Beispiel 16.27 (Geometrisches und arithmes Mittel). Wir zeigen die Ungleichung

$$(z_1 \cdots z_n)^{1/n} \leq \tfrac{1}{n}(z_1 + \cdots + z_n)$$

für $z_1, ..., z_n > 0$, die arithmetisches und geometrisches Mittel miteinander vergleicht.

Hierfür definieren wir $g : \mathbb{R}^n \to \mathbb{R}$ durch

$$g(\underline{x}) = x_1 + \cdots + x_n,$$
$$S := \{(x_1, ..., x_n) \in \mathbb{R}^n_+ : g(\underline{x}) = 1\}.$$

Wir wollen das Maximum der Funktion

$$f(\underline{x}) = x_1 \cdots x_n$$

für $\underline{x} \in S$ bestimmen. Da f auf $\overline{A}$ als stetige Funktion auf einer kompakten Menge ihr Maximum annimmt, und es nicht auf $\overline{A} \setminus A$ liegen kann, nimmt f ihr Maximum auf A an. Es ist $\nabla g(\underline{x}) = (1, ..., 1)$ und damit wird die Bedingung (16.8) zu

$$\prod_{j \neq i} x_j = \tfrac{1}{x_i} \prod_{j=1}^{n} x_j = \lambda$$

für ein $\lambda \in \mathbb{R}$. Dies ist nur dann möglich, wenn

$$\tfrac{1}{x_i} = \lambda, \qquad \text{also} \qquad x_i = \tfrac{1}{\lambda}$$

für alle $i = 1, ..., n$ und ein $\lambda \in \mathbb{R}$. Mit anderen Worten gilt im Maximum von f

$$x_1 = \cdots = x_n = \tfrac{1}{n}$$

und es gilt damit

$$x_1 \cdots x_n \leq \left(\tfrac{1}{n}\right)^n.$$

Sei nun $z_1, ..., z_n > 0$. Dann gilt mit $z = z_1 + \cdots + z_n$

$$\frac{z_1}{z} \cdots \frac{z_n}{z} \leq \left(\frac{1}{n}\right)^n,$$

also

$$(z_1 \cdots z_n)^{1/n} \leq \frac{1}{n}(z_1 + \cdots + z_n).$$

Bemerkung 16.28 (Arithmetisches, geometrisches und harmonisches Mittel). Seien $z_1, ..., z_n > 0$. Neben dem arithmetischen Mittel $\frac{1}{n}(z_1 + \cdots + z_n)$ und dem geometrischen Mittel $(z_1 \cdots z_n)^{1/n}$ gibt es noch das harmonische Mittel

$$\frac{n}{\frac{1}{z_1} + \cdots + \frac{1}{z_n}}.$$

Wir erweitern die Aussage des letzten Beispiels auf

$$\frac{n}{\frac{1}{z_1} + \cdots + \frac{1}{z_n}} \leq (z_1 \cdots z_n)^{1/n} \leq \tfrac{1}{n}(z_1 + \cdots + z_n).$$

Die zweite Ungleichung ist bereits bewiesen. Für die erste verwenden wir die zweite Ungleichung mit den Zahlen $\frac{1}{z_1}, ..., \frac{1}{z_n}$ und erhalten direkt

$$\frac{1}{(z_1 \cdots z_n)^{1/n}} \leq \frac{1}{n}\Big(\frac{1}{z_1} + \cdots + \frac{1}{z_n}\Big).$$

Bildet man auf beien Seiten den Kehrbruch, ergibt sich die erste behauptete Ungleichung.

Beispiel 16.29 (Eigenvektoren symmetrischer Matrizen). Sei $\underline{\underline{A}} = (a_{ij})_{1\leq i,j\leq n} \in \mathbb{R}^{n\times n}$ eine symmetrische Matrix und $f(\underline{x}) := \underline{x}^\top \underline{\underline{A}}\,\underline{x} = \langle \underline{x}, \underline{\underline{A}}\,\underline{x}\rangle$, wobei $\langle .,.\rangle$ wieder das Standard-Skalarprodukt ist. Wir wollen das Maximum der Funktion f unter der Nebenbedingung $\langle \underline{x}, \underline{x}\rangle = x_1^2 + \cdots + x_n^2 = 1$. Wir wollen also herausfinden, welche *Richtung* in der Menge $S^{n-1} := \{\underline{x} : |\underline{x}| = 1\}$ unter der Funktion f maximale Länge erhält. Da S^{n-1} kompakt ist, nimmt f ein solches Maximum an. Wir schreiben

$$F(\underline{x}, \lambda) = \langle \underline{x}, \underline{\underline{A}}\,\underline{x}\rangle - \lambda(\langle \underline{x}, \underline{x}\rangle - 1)$$

und leiten ab, um

$$\begin{aligned} \frac{\partial F(\underline{x},\lambda)}{\partial x_i} &= (\underline{\underline{A}}\,\underline{x})_i + (\underline{x}^\top \underline{\underline{A}})_i - 2\lambda x_i = 2(\underline{\underline{A}}\,\underline{x})_i - 2\lambda x_i, \qquad i = 1,...,n \\ \frac{\partial F(\underline{x},\lambda)}{\partial \lambda} &= \langle \underline{x}, \underline{x}\rangle - 1 \end{aligned}$$

zu erhalten. Die erste Gleichung impliziert, dass $\frac{\partial F(\underline{x}_0,\lambda)}{\partial x_i} = 0$ für $i = 1,...,n$ genau dann erfüllt ist, wenn $\underline{x}_0$ ein Eigenvektor von $\underline{\underline{A}}$ zum Eigenwert λ ist. Ist aber $\underline{x}_0$ ein Eigenvektor zum Eigenwert λ mit $|\underline{x}_0| = 1$, so ist $f(\underline{x}) = \langle \underline{x}, \underline{\underline{A}}\,\underline{x}\rangle = \lambda$. Das bedeutet, dass also λ der größte Eigenwert von $\underline{\underline{A}}$ der maximale Wert der Funktion f ist, und f nimmt im entsprechenden Eigenvektor ihr Maximum an.

17 Kurven

In Lemma 14.23 sind wir bereits Kurven (d.h. Abbildungen $\underline{\gamma} : I \to \mathbb{R}^n$, wobei I ein Intervall ist) begegnet. Genau genommen haben wir dort Ableitungen von $f \circ \underline{\gamma}$ für $\underline{f} \in \mathcal{C}^1(\mathbb{R}^n)$ kennen gelernt. In diesem Abschnitt wollen wir vor allem jedoch Integrale über Kurven betrachten.

17.1 Grundlagen

Zunächst benötigen wir ein paar Begriffe.

Definition 17.1 ($\mathcal{C}^k$-Kurve, wegweise zusammenhängend). *Sei I ein Intervall.*

1. *Ist $\underline{\gamma} \in \mathcal{C}^k(I, \mathbb{R}^n)$, so nennt man $\underline{\gamma}$ eine $\mathcal{C}^k$-*Kurve.
2. *Ist $I = [a,b]$ und $a = t_0 < \cdots < t_n = b$ und ist $\underline{\gamma} \in \mathcal{C}^0(I, \mathbb{R}^n)$ und so, dass $\underline{\gamma}|_{(t_{i-1},t_i)} \in \mathcal{C}^k((t_{i-1}, t_i))$, dann heißt $\underline{\gamma}$ eine stückweise $\mathcal{C}^k$-Kurve.*
3. *Sei $I = [a,b]$. Eine stückweise $\mathcal{C}^k$ Kurve $\underline{\gamma} \in \mathcal{C}^0(I, \mathbb{R}^n)$ heißt* geschlossen, *wenn $\underline{\gamma}(a) = \underline{\gamma}(b)$.*
4. *Eine Menge $E \subseteq \mathbb{R}^n$ heißt* wegweise zusammenhängend, *wenn es zu $\underline{x}, \underline{y} \in E$ ein $\underline{\gamma} \in \mathcal{C}^0([0,1], E)$ gibt mit $\underline{\gamma}(0) = \underline{x}, \underline{\gamma}(1) = \underline{y}$.*

Nun wiederholen wir, was wir bereits über Ableitungen von Kurve wissen.

Bemerkung 17.2 (Ableitungen von Kurven). Aus Lemma 14.23 wissen wir: Ist $\underline{\gamma}$ eine $\mathcal{C}^1$-Kurve und $\underline{f} \in \mathcal{C}^1(E, \mathbb{R}^p)$ mit $E \subseteq \mathbb{R}^n$ und $\underline{\gamma}(I) \subseteq E$, dann gilt

$$(\underline{f} \circ \underline{\gamma})'(t) := \underline{D}\,(\underline{f} \circ \underline{\gamma})'(t) = \underline{\underline{D}}\,\underline{f}(\underline{\gamma}(t)) \cdot \underline{D}\underline{\gamma}(t).$$

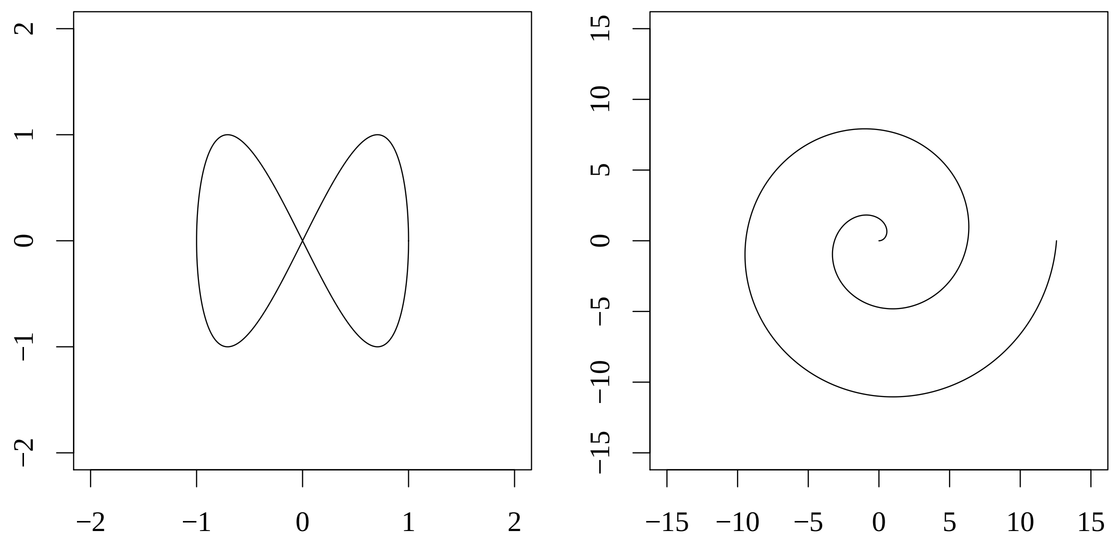

Abbildung 17.1: Links: Die Kurve $\widetilde{\underline{\gamma}}(t) = (\cos(t), \sin(2t))$. Rechts: Die Kurve $\widehat{\underline{\gamma}}(t) = (t\cos(t), t\sin(t))$.

Ist weiter $\underline{\gamma}$ eine $\mathcal{C}^2$-Kurve und $f \in \mathcal{C}^2(E)$, so gilt

$$(f \circ \underline{\gamma})''(t) = \underline{\underline{D}}^2 f(\underline{\gamma}(t))(D\underline{\gamma}(t), D\underline{\gamma}(t)) + \underline{D} f(\underline{\gamma}(t))D^2\underline{\gamma}(t),$$

wobei $\underline{\underline{D}}^2 f(\underline{\gamma}(t))$ die Hesse-Matrix von f an der Stelle $\underline{\gamma}(t)$ ist.

Beispiel 17.3. 1. Sei $\underline{x}, \underline{y} \in \mathbb{R}^n$ und $I = \mathbb{R}$. Dann ist $\underline{\gamma} : t \mapsto \underline{y} + t\underline{x}$ (zumindest im Falle $\underline{x} \neq \underline{0}$) eine Gerade.

2. Die $\mathcal{C}^\infty$-Kurve $\underline{\gamma} : [0, 2\pi] \to \mathbb{R}^2$, definiert durch

$$\underline{\gamma}(t) := (\cos(t), \sin(t))$$

stellt einen Kreis dar. Die Kurve

$$\widetilde{\underline{\gamma}}(t) := (\cos(t), \sin(2t))$$

ist eine *Acht*, die Kurve

$$\widehat{\underline{\gamma}}(t) := (t\cos(t), t\sin(t))$$

ist eine Schraubenlinie (siehe Abbildung 17.1).

3. Die Mengen $A := \{(x, y) : x^2 + y^2 \leq 1\}$ und $B := \{(x, y) : (x-2)^2 + (y-2)^2 \leq 1\}$ sind als Bälle um die Punkte $(0, 0)$ und $(2, 2)$ wegweise zusammenhängend. Allerdings ist $A \cup B$ nicht wegweise zusammenhängend.

In wegweise zusammenhängenden offenen Mengen kann man Wege in der Tat auch als Polygonzüge wählen, was wir nun zeigen werden.

Lemma 17.4 (Wegweise zusammenhängend). *Eine offene Menge $E \subseteq \mathbb{R}^n$ ist genau dann wegweise zusammenhängend, wenn es für $\underline{x}, \underline{y}$ einen Polygonzug (d.h. eine stückweise lineare und damit stückweise $\mathcal{C}^\infty$ Kurve $\underline{\gamma} \in \mathcal{C}^0([0,1], E)$ gibt mit $\underline{\gamma}(0) = \underline{x}, \underline{\gamma}(1) = \underline{y}$.*

Beweis. Die 'wenn'-Richtung ist klar, zu zeigen ist also nur die 'genau dann'-Richtung. Sei also $\underline{x}, \underline{y} \in E$ und $\underline{\gamma} \in \mathcal{C}^0([0,1], E)$ mit $\underline{\gamma}(0) = \underline{x}, \underline{\gamma}(1) = \underline{y}$. Weiter sei für $N \in \mathbb{N}$ die Punkte $t_k^N := k/N, k = 0, ..., N$ und $\underline{\gamma}^N$ so, dass $\underline{\gamma}^N(t_k^N) = \underline{\gamma}(t_k^N)$ und stückweise linear.

Wir zeigen zunächst

(i) $\underline{\gamma}^N \xrightarrow{N\to\infty}_{glm} \underline{\gamma}$.

Hierfür bemerken wir, dass $\underline{\gamma}$ als stetige Funktion einer kompakten Menge gleichmäßig stetig ist. Sei nun $\varepsilon > 0$ und $\delta > 0$ so, dass $|\underline{\gamma}(t) - \underline{\gamma}(s)| < \varepsilon$ für $|t - s| < \delta$. Für $N > 1/\delta$ folgt hieraus

$$\begin{aligned}|\underline{\gamma}^N(t) - \underline{\gamma}(t)| &\leq |\underline{\gamma}^N(t) - \underline{\gamma}^N([tN]/N)| + |\underline{\gamma}^N([tN]/N) - \underline{\gamma}([tN]/N)| + |\underline{\gamma}([tN]/N) - \underline{\gamma}(t)| \\ &\leq |\underline{\gamma}([tN+1]/N) - \underline{\gamma}([tN]/N)| + |\underline{\gamma}([tN]/N) - \underline{\gamma}(t)| < 2\varepsilon,\end{aligned}$$

da $\underline{\gamma}^N([tN]/N) = \underline{\gamma}([tN]/N)$ woraus die gleichmäßige Komvergenz folgt. Weiter behaupten wir, dass es ein $\varepsilon > 0$ gibt mit

(ii) $\inf\{|\underline{x} - \underline{\gamma}(t)| : t \in [0,1], \underline{x} \notin E\} > \varepsilon$.

Andernfalls gäbe es $t_1, t_2, ... \in [0,1]$ mit $t_n \xrightarrow{n\to\infty} t \in [0,1]$ und $\underline{x}_1, \underline{x}_2, ... \notin E$ mit $|\underline{\gamma}(t_n) - \underline{x}_n| \xrightarrow{n\to\infty} 0$. Da $\underline{\gamma}([0,1])$ kompakt ist, wäre dann auch $|\underline{\gamma}(t) - \underline{x}_n| \xrightarrow{n\to\infty} 0$. Da $\mathbb{R}^n \setminus E$ abgeschlossen ist wäre damit $\underline{\gamma}(t) \notin E$ im Widerspruch dazu, dass $\underline{\gamma} \in \mathcal{C}^0(I, E)$.

Sei nun $\varepsilon > 0$ so, dass (ii) gilt. Sei nun mit (i) N groß genug, so dass $\sup_{t\in[0,1]} |\underline{\gamma}^N(t) - \gamma(t)| < \varepsilon$. Dann ist offenbar $\underline{\gamma}^N \in \mathcal{C}^0([0,1], E)$ eine stückweise $\mathcal{C}^\infty$ Kurve mit den geforderten Eigenschaften. □

17.2 Bogenlänge

Kurven sind ein-dimensionale Objekte, die einen Teil des $\mathbb{R}^n$ einnehmen. Als ein-dimensionale Objekte sollten sie auch eine Länge besitzen, die wir nun als Bogenlänge kennen lernen werden.

Definition 17.5 (Bogenlänge). *Sei $I = [a, b]$ und $\underline{\gamma} \in \mathcal{C}^0(I, \mathbb{R}^n)$ eine stückweise $\mathcal{C}^1$-Kurve. Dann ist die Bogenlänge von $\underline{\gamma}$ definiert als*

$$L(\underline{\gamma}) := \sup_{a\leq t_0 < ... < t_n \leq b} \sum_{i=1}^n |\underline{\gamma}(t_i) - \underline{\gamma}(t_{i-1})|$$

Lemma 17.6 (Berechnung der Bogenlänge). *Sei $I = [a, b]$ und $\underline{\gamma} \in \mathcal{C}^0(I, \mathbb{R}^n)$ eine stückweise $\mathcal{C}^1$-Kurve. Dann gilt*

$$L(\underline{\gamma}) = \int_a^b |\underline{\gamma}'(t)| dt.$$

Beweis. Es genügt die Behauptung im Fall dass $\underline{\gamma}$ eine $\mathcal{C}^1$-Kurve ist zu zeigen. Zunächst ist $\underline{\gamma}'$ auf I als stetige Funktion auf einer kompakten Menge gleichmäßig stetig. Sei $\varepsilon > 0$. Wähle dann $\delta > 0$ klein genug, so dass $|\underline{\gamma}(t) - \underline{\gamma}(s)| < \varepsilon/(b - a)$ für $|t - s| < \delta$. Sei nun $a = t_0 \leq t_1 < \cdots < t_n = b$ so, dass $|t_i - t_{i-1}| < \delta, i = 1, ..., n$. Dann gilt

$$\begin{aligned}\Big|\int_a^b |\underline{\gamma}'(t)| dt - \sum_{i=1}^n |\underline{\gamma}(t_i) - \underline{\gamma}(t_{i-1})|\Big| &= \sum_{i=1}^n \int_{t_{i-1}}^{t_i} |\underline{\gamma}'(t)| dt - \sum_{i=1}^n \Big|\int_{t_{i-1}}^{t_i} \gamma'(t) dt\Big| \\ &\leq \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \big(|\underline{\gamma}'(t_{i-1})| + \frac{\varepsilon}{b-a}\big) dt - \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \big(|\gamma'(t_{i-1})| - \frac{\varepsilon}{b-a}\big) dt \leq 2\varepsilon.\end{aligned}$$

Damit ist die Aussage gezeigt. □

Beispiel 17.7. Wir betrachten nochmals die Kurven aus Beispiel 17.3.2. Für $\widehat{\gamma}$ ist

$$\underline{\widetilde{\gamma}}'(t) = (-\sin(t), 2\cos(2t)),$$

und deshalb

$$|\underline{\widetilde{\gamma}}'(t)|^2 = \sin^2(t) + 4\cos^2(2t).$$

Damit gilt für die Bogenlänge

$$L(\underline{\widetilde{\gamma}}) = \int_0^{2\pi} |\underline{\widetilde{\gamma}}'(t)| dt = \int_0^{2\pi} \sqrt{\sin^2(t) + 4\cos^2(2t)} dt.$$

Für $\underline{\widehat{\gamma}}$ berechnen wir

$$\underline{\widehat{\gamma}}'(t) = (\cos(t) - t\sin(t), \sin(t) + t\cos(t)$$

und deshalb

$$\begin{aligned} |\underline{\widehat{\gamma}}'(t)|^2 &= \cos^2(t) + t^2\sin^2(t) - 2t\sin(t)\cos(t) + \sin^2(t) + t^2\cos^2(t) + 2t\sin(t)\cos(t) \\ &= 1 + t^2. \end{aligned}$$

Damit gilt für die Bogenlänge

$$L(\underline{\widehat{\gamma}}) = \int_0^{2\pi} |\underline{\widehat{\gamma}}'(t)| dt = \int_0^{2\pi} \sqrt{(1+t^2)} dt.$$

Definition 17.8 (Umparametrisierung, Parametrisierung nach der Bogenlänge). *Seien I, J Intervalle und $\underline{\beta} : I \to \mathbb{R}^n$ und $\underline{\gamma} : J \to \mathbb{R}^n$ zwei stückweise $\mathcal{C}^1$-Kurven.*

1. *Gibt es eine Bijektion $\varphi \in \mathcal{C}^0(I, J)$, die stückweise $\mathcal{C}^1$ ist und $\varphi' > 0$ oder $\varphi' > 0$ und $\underline{\gamma} = \underline{\beta} \circ \varphi$, so heißt $\underline{\gamma}$ eine Umparametrisierung von $\underline{\beta}$ (und umgekehrt). Im Falle $\varphi' > 0$ heißt φ orientierunserhaltend, im Falle $\varphi' < 0$ orientierungsumkehrend.*

2. *Die Kurve $\underline{\gamma}$ heißt* nach der Bogenlänge parametrisiert, *wenn $|\underline{\gamma}'(t)| = 1$ für alle $t \in I$.*

Lemma 17.9 (Invarianz der Bogenlänge). *Seien I, J Intervalle und $\underline{\beta} : I \to \mathbb{R}^n$ und $\underline{\gamma} : J \to \mathbb{R}^n$ zwei stückweise $\mathcal{C}^1$-Kurven. Dann gilt:*

1. *Ist $\underline{\gamma}$ ein Umparametrisierung von $\underline{\beta}$, so ist $L(\underline{\beta}) = L(\underline{\gamma})$.*

2. *Gilt $|\underline{\gamma}'| > 0$, so gibt eine Umparametrisierung $\underline{\widehat{\gamma}}$ von $\underline{\gamma}$, die nach der Bogenlänge parametrisiert ist.*

Beweis. Es genügt die Aussagen jeweils für $\mathcal{C}^1$-Kurven zu zeigen. 1. Nach Voraussetzung gibt es ein $\varphi \in \mathcal{C}^1(I, J)$, so dass $\underline{\gamma} = \underline{\beta} \circ \varphi$. Deshalb gilt nach Lemma 17.6 wegen $\underline{\gamma}'(t) = \underline{\beta}'(\varphi(t)) \cdot \varphi'(t)$ und wegen dem Transformationssatz

$$L(\underline{\gamma}) = \int_I |\underline{\gamma}'(t)| dt = \int_I |\underline{\beta}'(\varphi(t)) \cdot \varphi'(t)| dt = \int_J |\underline{\beta}'(s)| ds = L(\underline{\beta}).$$

2. Es gilt nun, eine Abbildung $\varphi \in \mathcal{C}^1(I, [0, L(\underline{\gamma})])$ zu finden, so dass $|(\underline{\gamma} \circ \varphi)'| = 1$. Hierzu sei $I = [a, b]$ und wir betrachten wir die Funktion

$$\sigma(t) = L(\underline{\gamma}|_{[a,t]}) = \int_a^t |\underline{\gamma}'(s)| ds.$$

Da $|\underline{\gamma}'| > 0$, ist dies eine streng monotone Funktion mit $\sigma : I \to [0, L(\underline{\gamma})]$. Damit ist σ umkehrbar und wir wissen wegen Proposition 7.8 $(\sigma^{-1})'(t) = 1/\sigma'(\sigma^{-1}(t)) = 1/|\underline{\gamma}'(\sigma^{-1}(t))|$ und damit

$$|(\underline{\gamma} \circ \sigma^{-1})'(t)| = |\underline{\gamma}'(\sigma^{-1}(t))| \cdot |(\sigma^{-1})'(t)| = 1.$$

Damit ist also $\underline{\beta} := \underline{\gamma} \circ \sigma^{-1}$ nach der Bogenlänge parametrisiert. □

17.3 Kurvenintegrale

Die Berechnung der Bogenlänge mittels $L(\underline{\gamma}) = \int_I |\underline{\gamma}(t)'| dt$ ist schon ein Integral über die Kurve $\underline{\gamma}$. Wir werden dies nun verallgemeinern und kommen zu Integralen von Kurven über Vektorfelder.

Definition 17.10 (Vektorfeld, Kurvenintegral). *1. Sei $E \subseteq \mathbb{R}^n$. Eine Abbildung $\underline{f} \in \mathcal{C}^0(E, \mathbb{R}^n)$ heißt* Vektorfeld, *weil jedem $\underline{x} \in E$ ein Vektor $\underline{f}(\underline{x})$ zugeordnet wird.*

2. *Sei $E \subseteq \mathbb{R}^n$, $\underline{f} \in \mathcal{C}^0(E, \mathbb{R}^n)$ ein Vektorfeld sowie $\underline{\gamma} \in \mathcal{C}^0(I, E)$ eine stückweise $\mathcal{C}^1$-Kurve. Dann definieren wir das Kurvenintegral*

$$\int_{\underline{\gamma}} f(\underline{x}) \cdot d\underline{x} := \int_I \langle \underline{f}(\underline{\gamma}(t)), \underline{\gamma}'(t) \rangle dt$$

von $\underline{f}$ über $\underline{\gamma}$. Hierbei bezeichnet $\langle ., . \rangle$ das Standard-Skalarprodukt.

Lemma 17.11 (Eigenschaften des Kurvenintegrals). *Seien I, J Intervalle, $E \subseteq \mathbb{R}^n$, $\underline{f}, \underline{f}_1, \underline{f}_2 \in \mathcal{C}^1(E, \mathbb{R}^n)$ Vektorfelder und $\underline{\beta} \in \mathcal{C}^0(I, E), \underline{\gamma} \in \mathcal{C}^0(J, E)$ stückweise $\mathcal{C}^1$ Kurven.*

1. *Für $\lambda_1, \lambda_2 \in \mathbb{R}$ gilt*

$$\int_{\underline{\gamma}} (\lambda_1 \underline{f}_1 + \lambda_2 \underline{f}_2) \cdot d\underline{x} = \lambda_1 \int_{\underline{\gamma}} f_1(\underline{x}) \cdot d\underline{x} + \lambda_2 \int_{\underline{\gamma}} f_2(\underline{x}) \cdot d\underline{x}.$$

2. *Für $I = [a, b]$ und $a = t_0 < ... < t_n = b$ und $\gamma_i := \gamma_{[t_{i-1}, t_i]}$ gilt*

$$\int_{\underline{\gamma}} f(\underline{x}) \cdot d\underline{x} = \sum_{i=1}^n \int_{\underline{\gamma}_i} f(\underline{x}) \cdot d\underline{x}.$$

3. *Ist $\underline{\beta} = \underline{\gamma} \circ \varphi$ eine Umparametrisierung und φ orientierungserhaltend, so gilt*

$$\int_{\underline{\beta}} \underline{f} \cdot d\underline{x} = \int_{\underline{\gamma}} \underline{f} \cdot d\underline{x}.$$

Ist φ orientierungsumkehrend, so gilt hingegen

$$\int_{\underline{\beta}} \underline{f} \cdot d\underline{x} = -\int_{\underline{\gamma}} \underline{f} \cdot d\underline{x}.$$

Beweis. 1. und 2. folgen direkt aus der Definition des Kurvenintegrals und Eigenschaften des Riemann-Integrals. Für 3. schreiben wir mit Hilfe der Integration durch Substitution

$$\begin{aligned}\int_{\underline{\beta}} \underline{f} \cdot d\underline{x} &= \int_J \langle \underline{f}(\underline{\beta}(t)), \underline{\beta}'(t)\rangle dt = \int_J \langle \underline{f}(\underline{\gamma} \circ \varphi(t)), \underline{\gamma}'(\varphi(t))\rangle \varphi'(t) dt \\ &= \int_I \langle \underline{f}(\underline{\gamma}(t)), \underline{\gamma}'(t)\rangle dt = \int_{\underline{\gamma}} \underline{f} \cdot d\underline{x},\end{aligned}$$

wobei die das dritte Gleichheitszeichen nur gilt, wenn φ orientierungserhaltend ist und sich das Vorzeichen bei diesem Gleichheitszeichen umdreht, wenn φ orientierungsumkehrend ist. □

Lemma 17.12. *Sei I ein Intervall, $E \subseteq \mathbb{R}^n$, $\underline{f} \in \mathcal{C}^1(E, \mathbb{R}^n)$ und $\underline{\gamma} \in \mathcal{C}^0(I, E)$ eine stückweise $\mathcal{C}^1$ Kurve. Dann gilt*

$$\Big| \int_{\underline{\gamma}} \underline{f}(\underline{x}) \cdot d\underline{x} \Big| \leq ||\underline{f} \circ \underline{\gamma}||_I \cdot L(\underline{\gamma}),$$

wobei $||.||_I$ die Supremumsnorm auf I ist.

Beweis. Wir erinnern an die Cauchy-Schwartz-Ungleichung und erhalten

$$\begin{aligned}\Big| \int_I \langle \underline{f}(\underline{\gamma}(t)), \underline{\gamma}'(t)\rangle dt \Big| &\leq \int_I \Big| \langle \underline{f}(\underline{\gamma}(t)), \underline{\gamma}'(t)\rangle \Big| dt \leq \int_I |\underline{f}(\underline{\gamma}(t))| \cdot |\underline{\gamma}'(t)| dt \\ &\leq ||\underline{f} \circ \underline{\gamma}||_I \cdot \int_I |\underline{\gamma}'(t)| dt = ||\underline{f} \circ \underline{\gamma}||_I \cdot L(\underline{\gamma})\end{aligned}$$

□

17.4 Gradientenfelder

Von stetigen Funktionen $f : \mathbb{R} \to \mathbb{R}$ ist man gewohnt, dass sie eine Stammfunktion besitzen. Im Mehrdimensionalen ist dies nicht mehr notwendig der Fall. Vektorfelder $\underline{f}$, die eine solche Stammfunktion F besitzen, heißen Gradientenfelder (und spielen insbesondere in der Physik eine große Rolle).

Definition 17.13 (Gradientenfeld). *Sei $E \subseteq \mathbb{R}^n$ offen. Dann heißt das Vektorfeld $\underline{f} = (f_1, ..., f_n) \in \mathcal{C}^0(E, \mathbb{R}^n)$* Gradientenfeld, *wenn es ein $F \in \mathcal{C}^1(E)$ gibt mit $\underline{\nabla} F = \underline{f}$, d.h. wenn*

$$f_i(\underline{x}) = \frac{\partial F(\underline{x})}{\partial x_i}.$$

Die Funktion F heißt dann Stammfunktion *(oder auch* Potential*) von $\underline{f}$.*

Beispiel 17.14. 1. Sei $|\underline{x}| := \sqrt{x_1^2 + \cdots + x_n^2}$. Das Vektorfeld $\underline{f} : \underline{x} \mapsto \frac{\underline{x}}{|\underline{x}|}$ ist ein Gradientenfeld. In Beispiel 14.3 haben wir nämlich berechnet, dass

$$\frac{\partial |\underline{x}|}{\partial x_i} = \frac{x_i}{|\underline{x}|},$$

d.h. $\underline{x} \mapsto |\underline{x}|$ ist eine Stammfunktion von $\underline{f}$.

2. Wir betrachten das Vektorfeld $\underline{f}(x,y) = (-y, x)$. Wäre dies ein Gradientenfeld, müsste es eine Funktion $(x,y) \mapsto F(x,y)$ geben mit $\frac{\partial F(x,y)}{\partial x} = -y$, d.h. $F(x,y) = -yx + g(y)$ für eine Funktion g. Dann wäre $\frac{\partial F(x,y)}{\partial y} = -x + g'(y)$, was aber für jede Wahl von g ungleich $x \mapsto x$ ist.

Lemma 17.15 (Eindeutigkeit der Stammfunktion). *Sei $E \subseteq \mathbb{R}^n$ offen und wegweise zusammenhängend, sowie $\underline{f} \in \mathcal{C}^0(E, \mathbb{R}^n)$ ein Gradientenfeld und $F_1, F_2 \in \mathcal{C}^1(E)$ mit $\underline{\nabla} F_1 = \underline{\nabla} F_2 = \underline{f}$. Dann ist $F_1 = F_2 + c$ für ein $c \in \mathbb{R}$, d.h. die Stammfunktion von $\underline{f}$ ist bis auf eine konstante eindeutig bestimmt.*

Beweis. Zunächst ist klar, dass $\underline{D}(F_1 - F_2) = \underline{\nabla}(F_1 - \underline{F}_2) = \underline{f} - \underline{f} = 0$. Sei nun $\underline{x}, \underline{y} \in E$. Nach Lemma 17.4 können wir oBdA annehmen, dass es einen Polygonzug von $\underline{x}$ nach $\underline{y}$ in E gibt. Aus (15.3) folgt damit, dass

$$|F_1(\underline{x}) - F_2(\underline{x}) - F_1(\underline{y}) + F_2(\underline{y})| = 0.$$

Mit anderen Worten ist

$$\underline{x} \mapsto F_1(\underline{x}) - F_2(\underline{x})$$

konstant, was ja gerade zu beweisen war. □

Das Integral einer Kurve $\underline{\gamma} \in \mathcal{C}^k([a,b], \mathbb{R}^n)$ über ein Gradientenfeld $\underline{f}$ hat sehr schöne Eigenschaften. Ist F Stammfunktion von $\underline{f}$, so hängt ein solches Kurvenintegral nämlich nur von $F(\underline{\gamma}(a))$ und $F(\underline{\gamma}(b))$ ab.

Theorem 17.16 (Wegunabhängigkeit von Wegintegralen entlang Gradientenfelder). *Sei $E \subseteq \mathbb{R}^n$ offen und wegweise zusammenhängend und $\underline{f} \in \mathcal{C}^0(E, \mathbb{R}^n)$. Dann sind äquivalent:*

1. *$\underline{f}$ ist ein Gradientenfeld.*
2. *Für jede geschlossene stückweise $\mathcal{C}^1$-Kurve $\underline{\gamma}$ ist*
$$\int_{\underline{\gamma}} \underline{f} \cdot d\underline{x} = 0.$$
3. *Für $I = [a,b]$ und $J = [c,d]$ und stückweise $\mathcal{C}^1$-Kurven $\underline{\beta} \in \mathcal{C}^0(I, E)$ und $\underline{\gamma} \in \mathcal{C}^0(J, E)$ mit $\underline{\beta}(a) = \underline{\gamma}(c)$ und $\underline{\beta}(b) = \underline{\gamma}(cd)$ gilt*
$$\int_{\underline{\beta}} \underline{f} \cdot d\underline{x} = \int_{\underline{\gamma}} \underline{f} \cdot d\underline{x}.$$

Beweis. 1. $\Rightarrow$ 2.: Sei $F \in \mathcal{C}^1(E)$ mit $\underline{\nabla} F = \underline{f}$ und $\underline{\gamma} \in \mathcal{C}^0(I = [a,b], E)$. Dann gilt mit Lemma 14.23

$$\begin{aligned}
\int_{\underline{\gamma}} \underline{f} \cdot d\underline{x} &= \int_I \langle (\underline{f} \circ \underline{\gamma}), \underline{\gamma}'(t) \rangle dt \\
&= \int_I \langle ((\underline{\nabla} F) \circ \underline{\gamma}), \underline{\gamma}'(t) \rangle dt \\
&= \int_I (F \circ \underline{\gamma})'(t) dt = F(\underline{\gamma}(b)) - F(\underline{\gamma}(a)) = 0
\end{aligned}$$

2. $\Rightarrow$ 3.: Wir definieren die stückweise $\mathcal{C}^1([0, b-a+d-c], E)$ Kurve

$$\underline{\alpha}: \begin{cases} [b-a+d-c] & \to E \\ t & \mapsto \begin{cases} \underline{\beta}(t), & 0 \leq t \leq b-a, \\ \underline{\gamma}(b-a+d-c-t), & t \geq b-a. \end{cases} \end{cases}.$$

Das bedeutet, dass $\underline{\alpha}$ zunächst $\underline{\beta}$ und danach $\underline{\gamma}$ in umgekehrter Richtung durchläuft. Insbesondere ist wegen den Voraussetzungen in 3. die Kurve $\underline{\alpha}$ geschlossen und es gilt wegen Lemma 17.11.2

$$\begin{aligned} 0 = \int_{\underline{\alpha}} \underline{f}(\underline{x}) \cdot d\underline{x} &= \int_{\underline{\alpha}|_{[0,b-a]}} \underline{f}(\underline{x}) \cdot d\underline{x} + \int_{\underline{\alpha}|_{[b-a,b-a+d-c]}} \underline{f}(\underline{x}) \cdot d\underline{x} \\ &= \int_{\underline{\beta}} \underline{f}(\underline{x}) \cdot d\underline{x} - \int_{\underline{\gamma}} \underline{f}(\underline{x}) \cdot d\underline{x}. \end{aligned}$$

3. $\Rightarrow$ 1.: Ziel ist es, eine Stammfunktion von $\underline{f}$ anzugeben. Hierzu sei für $\underline{z} \in E$ beliebig. Für $\underline{x} \in E$ sei $\underline{\gamma}_{\underline{x}} \in \mathcal{C}^0(I = [a,b], E)$ eine stückweise $\mathcal{C}^1$-Kurve mit $\underline{\gamma}_{\underline{x}}(a) = \underline{z}, \underline{\gamma}_{\underline{x}}(b) = \underline{x}$. Wir definieren

$$F(\underline{x}) := \int_{\underline{\gamma}_{\underline{x}}} \underline{f}(\underline{x}) \cdot d\underline{x}$$

und bemerken, dass F wohldefiniert ist, da $F(\underline{x})$ nach Voraussetzung von $\underline{\gamma}_{\underline{x}}$ nur über den Endpunkt der Kurve, also $\underline{x}$, abhängt.

Es bleibt nun zu zeigen, dass $\underline{\nabla} F = \underline{f}$. Für $j = 1, ..., n$ und $\underline{x} \in E$ betrachten wir nun $\underline{\gamma}_{\underline{x}} \in \mathcal{C}^0([a,b], E)$ und $\underline{\gamma}_{\underline{x}+h\underline{e}_j} \in \mathcal{C}^0([a, b+1], E)$, so dass $\underline{\gamma}_{\underline{x}+h\underline{e}_j}(b+t) = \underline{x} + ht\underline{e}_j$. Das bedeutet, dass $\underline{\gamma}_{\underline{x}+h\underline{e}_j}$ über den Punkt $\underline{x}$ läuft. Nun gilt

$$\begin{aligned} \frac{F(\underline{x}+h\underline{e}_j) - F(\underline{x})}{h} &= \frac{1}{h}\Big(\int_{\gamma_{\underline{x}+h\underline{e}_j}} \underline{f}(\underline{y}) \cdot d\underline{y} - \int_{\gamma_{\underline{x}}} \underline{f}(\underline{y}) \cdot d\underline{y} \Big) \\ &= \frac{1}{h} \int_0^1 \langle \underline{f}(\underline{x} + ht\underline{e}_j), h\underline{e}_j \rangle dt \\ &= \int_0^1 f_j(\underline{x} + ht\underline{e}_j) dt \xrightarrow{h \to 0} f_j(\underline{x}), \end{aligned}$$

woraus die Behauptung folgt. □

Um zu prüfen, ob ein gegebenes Vektorfeld ein Gradientenfeld ist, geben wir nun noch ein Kriterium an.

Korollar 17.17 (Rotationsfreiheit von Gradientenfeldern). *Sei $E \subseteq \mathbb{R}^n$ offen und $\underline{f} \in \mathcal{C}^1(E, \mathbb{R}^n)$ ein Gradientenfeld. Dann gilt*

$$\frac{\partial \underline{f}_i}{\partial x_j} = \frac{\partial \underline{f}_j}{\partial x_i}.$$

Beweis. Es sei $F \in \mathcal{C}^2(E, \mathbb{R}^n)$ eine Stammfunktion von $\underline{f}$. Der Satz von Schwartz, Theorem 14.16, impliziert nun

$$\frac{\partial \underline{f}_i}{\partial x_j} = \frac{\partial}{\partial x_j} \frac{\partial F(\underline{x})}{\partial x_i} = \frac{\partial}{\partial x_i} \frac{\partial F(\underline{x})}{\partial x_j} = \frac{\partial \underline{f}_j}{\partial x_i}.$$

□

Beispiel 17.18. Wir betrachten das Vektorfeld $\underline{f} : (x, y) \mapsto (-y, x)$ aus Beispiel 17.14.2. Wir berechnen

$$\frac{\partial f_1}{\partial y} = -1 \neq 1 = \frac{\partial f_2}{\partial x} = 1.$$

Insbesondere kann $\underline{f}$ kein Gradientenfeld sein (was wir in Beispiel 17.14.2 ja schon gesehen hatten).

Bemerkung 17.19 (Rotation). Sei $E \subseteq \mathbb{R}^3$ offen und $\underline{f} \in \mathcal{C}^1(E, \mathbb{R}^n)$ ein Vektorfeld. Dann heißt das Vektorfeld $\underline{g} \in \mathcal{C}^0(E, \mathbb{R}^n)$, definiert durch

$$\underline{\mathrm{rot}}(\underline{f}) := \begin{pmatrix} \frac{\partial f_3}{\partial x_2} - \frac{\partial f_2}{\partial x_3} \\ \frac{\partial f_1}{\partial x_2} - \frac{\partial f_2}{\partial x_1} \\ \frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2} \end{pmatrix}$$

Rotation von $\underline{f}$. Korollar 17.17 besagt nun, dass für Gradientenfelder die Rotation $\underline{0}$ ist.

18 Ausblicke

Die Vorlesungen Analysis sind Grundpfeiler einer ganzer Mange mathematischer Theorie. Wir geben nun noch einen kleinen Ausblick, was wir mit dem bisher Gelernten anfangen können. Die Funktionentheorie (siehe Abschnitt 18.1) ist eine Fortsetzung der Analysis für Funktionen $f : \mathbb{C} \to \mathbb{C}$. Weiter behandeln wir in Abschnitt 18.2 den Fixpunktsatz von Brouwer. Dieser Satz hat beispielsweise Anwendungen in der Funktionalanalysis.

18.1 Funktionentheorie

Die Funktionentheorie beschäftigt sich mit Funktionen $f : \mathbb{C} \to \mathbb{C}$. Hierzu bemerken wir, dass eine solche Funktion gegeben ist durch eine Funktion $\underline{f} = (f_1, f_2) : \mathbb{R}^2 \to \mathbb{R}^2$ mittels

$$f(x + iy) = f_1(x, y) + if_2(x, y).$$

Wir bezeichnen hier $Ref = f_1$ und $Imf = f_2$. Zentral ist hierbei der Begriff der *komplexen Differenzierbarkeit* oder *Holomorphie.*

Definition 18.1 (Holomorphie). *Sei $E \subseteq \mathbb{C}$ offen und $f : E \to \mathbb{C}$. Dann heißt f in $z = x + iy$* holomorph *mit Ableitung $f'(z)$, wenn*

$$\lim_{h \to 0} \frac{f(z + h) - f(z)}{h} = f'(z)$$

existiert (wobei $h \to 0$ im Komplexen betrachtet wird). Weiter heißt f holomorph auf E, wenn f für alle $z \in E$ holomorph ist.

Bemerkung 18.2 (Ableitungsregelen, $\mathbb{C} \sim \mathbb{R}^2$).

1. Genau wie im Reellen gelten ein paar Ableitungsregeln: Linearität der Differentiation, Produktregel, Quotientenregel, Kettenregel.

2. Wir identifizieren im Folgenden $z = x + iy \in \mathbb{C}$ mit $(x, y) \in \mathbb{R}^2$. Etwa schreiben wir für $c = a + ib$ und $z = x + iy$

$$cz = (ax - by, bx + ay) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Proposition 18.3 (Cauchy-Riemann'sche Differentielgleichungen). *Sei $E \subseteq \mathbb{C}$ offen und $f : E \to \mathbb{C}$ mit $f(x+iy) = f_1(x, y) + if_2(x, y)$. Dann ist f genau dann holomorph, wenn (f_1, f_2) differenzierbar ist mit*

$$\frac{\partial f_1}{\partial x} = \frac{\partial f_2}{\partial y}, \qquad \frac{\partial f_2}{\partial x} = -\frac{\partial f_1}{\partial y}.$$

Beweis. Sei zunächst $f = f_1 + if_2$ komplex differenzierbar mit Ableitung f' und $\underline{f} = (f_1, f_2)$. Dann gilt mit $h = h_1 + ih_2 \in \mathbb{C}$

$$\frac{\underline{f}(x + h_1, y + h_2) - \underline{f}(x, y) - \begin{pmatrix} Ref'(z) & -Imf'(z) \\ Imf'(z) & Ref'(z) \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}}{|h|}$$
$$= \frac{\underline{f}(z + h) - f(z) - f'(z)h}{h} \frac{h}{|h|} \xrightarrow{h \to 0} 0.$$

Damit ist $\underline{f}$ differenzierbar mit

$$\underline{\underline{D}}\,\underline{f}(z) = \begin{pmatrix} Ref'(z) & -Imf'(z) \\ Imf'(z) & Ref'(z) \end{pmatrix}.$$

Daraus folgt bereits die Gültigkeit der Cauchy-Riemann'schen Differentialgleichungen.

Sei umgekehrt $\underline{f}$ differenzierbar und die Cauchy-Riemann'schen Differentialgleichungen gelten, d.h.

$$\underline{\underline{D}}\,\underline{f}(z) = \begin{pmatrix} \frac{\partial f_1(x,y)}{\partial x} & \frac{\partial f_2(x,y)}{\partial x} \\ -\frac{\partial f_2(x,y)}{\partial x} & \frac{\partial f_1(x,y)}{\partial x} \end{pmatrix} = \begin{pmatrix} \frac{\partial f_2(x,y)}{\partial y} & -\frac{\partial f_1(x,y)}{\partial y} \\ \frac{\partial f_1(x,y)}{\partial y} & \frac{\partial f_2(x,y)}{\partial y} \end{pmatrix}.$$

Für

$$f'(z) := \frac{\partial f_1(x,y)}{\partial x} - i\frac{\partial f_2(x,y)}{\partial x} = \frac{\partial f_2(x,y)}{\partial y} + i\frac{\partial f_1(x,y)}{\partial y}$$

ist nun

$$f'(z)h = \underline{\underline{D}}\,\underline{f}(x, y)\underline{h}$$

und damit

$$\frac{f(z + h) - f(z) - f'(z)h}{h} = \frac{\underline{f}(x + h_1, y + h_2) - \underline{f}(x, y) - \underline{\underline{D}}\,\underline{f}(x, y)\underline{h}}{|h|} \frac{|h|}{h}$$
$$\xrightarrow{h \to 0} 0$$

wegen der Differenzierbarkeit. □

Definition 18.4 (Komplexe Integrale und Kurvenintegral). *Sei I ein Intervall und $f : I \to \mathbb{C}$ mit $f = (f_1, f_2)$. Dann setzen wir*

$$\int_I f(t)dt = \int_I f_1(t)dt + i\int_I f_2(t)dt.$$

Ein spezielles Beispiel sind Kurvenintegrale. Ist I ein Intervall und $\gamma \in \mathcal{C}^0(I, \mathbb{C})$ eine stückweise $\mathcal{C}^1$ Kurve, dann ist

$$\int_\gamma f(z)dz := \int_I f(\gamma(t))\gamma'(t)dt.$$

Mit $\gamma = (\gamma_1, \gamma_2)$ gilt dann

$$\begin{aligned}\int_\gamma f(z)dz &= \int_I f_1(\gamma(t))\gamma_1'(t) - f_2(\gamma(t))\gamma_2'(t)dt + i\int_I f_1(\gamma(t))\gamma_2'(t) + f_2(\gamma(t))\gamma_1'(t)dt \\ &= \int_{\underline{\gamma}} (f_1, -f_2)(x, y)\cdot d(x, y) + i\int_{\underline{\gamma}} (f_2, f_1)(x, y)\cdot d(x, y).\end{aligned} \tag{18.1}$$

Bemerkung 18.5 (Merkregel). Um sich die Formel (18.1), die wir im Folgenden noch öfter verwenden werden, merken zu können, schreiben wir $dz = dx + i\,dy$ und damit formal

$$\begin{aligned}\int_\gamma f(z)dz &= \int_\gamma (f_1(z) + if_2(z))(dx + i\,dy) = \int_\gamma f_1(z)dx - f_2(z))dy + i\int_\gamma f_2(z)dx + f_1(z))dy \\ &= \int_{\underline{\gamma}} (f_1, -f_2)(x, y)\cdot d(x, y) + i\int_{\underline{\gamma}} (f_2, f_1)(x, y)\cdot d(x, y).\end{aligned}$$

Beispiel 18.6 (Kurvenintegrale von Polynomen). Wir zeigen nun für $\gamma : [0, 1] \to \mathbb{C}$ mit $\gamma(t) = re^{2\pi it}$

$$\int_\gamma z^n \cdot dz = \begin{cases} 0, & n \neq -1, \\ 2\pi i, & n = -1 \end{cases}.$$

Denn: Wir schreiben direkt

$$\int_\gamma z^n dz = 2\pi i r^{n+1}\int_0^1 e^{2\pi i(n+1)t}dt.$$

Für $n = -1$ folgt daraus direkt die Behauptung. Für $n \neq -1$ hingegen ist damit

$$\int_\gamma z^n dz = \frac{r^{n+1}}{n+1} e^{2\pi i(n+1)t}\Big|_{t=0}^{t=1} = 0.$$

Die folgende Homotopieformel gilt zwar allgemein, wird aber besonders schön, wenn es um Kurvenintegrale von holomorphen Funktionen geht; siehe den Cauchy'schen Integralsatz in Theorem 18.8.

Lemma 18.7 (Eine Homotopieformel). *Sei $E \subseteq \mathbb{R}^n$ sternförmig mit Mittelpunkt $\underline{y} \in E$ und $\underline{f} \in \mathcal{C}^1(E, \mathbb{R}^n)$ ein Vektorfeld. Weiter sei $\underline{\gamma} \in \mathcal{C}^0(I, E)$ eine geschlossene, stückweise $\mathcal{C}^1$ Kurve und $\underline{\gamma}_s(t) := (1-s)\underline{y} + s\underline{\gamma}(t)$. Dann gilt*

$$\int_{\underline{\gamma}} \underline{f}(\underline{x})\cdot d\underline{x} = \int_0^1 s\int_I \langle(\underline{\underline{D}}\,\underline{f}((1-s)\underline{y} + s\underline{\gamma}(t))^\top - \underline{\underline{D}}\,\underline{f}((1-s)\underline{y} + s\underline{\gamma}(t))\underline{\gamma}'(t)), \underline{\gamma}(t)\rangle dtds.$$

Ist insbesondere $\underline{\underline{D}}\,\underline{f}$ symmetrisch, so ist das Kurvenintegral 0.

Beweis. OBdA sei $\underline{y} = \underline{0}$. Es gilt

$$\int_{\underline{\gamma}_0} \underline{f}(\underline{x}) \cdot d\underline{x} = 0$$

sowie, mittels Differentiation unter dem Integral und partieller Integration,

$$\begin{aligned}
\frac{\partial}{\partial s} \int_{\underline{\gamma}_s} \underline{f}(\underline{x}) \cdot d\underline{x} &= \frac{\partial}{\partial s} \int_I \left\langle \underline{f}(\underline{\gamma}_s(t)), \frac{\partial \underline{\gamma}_s(t)}{\partial t} \right\rangle dt \\
&= \frac{\partial}{\partial s} \int_I \langle \underline{f}(s\underline{\gamma}(t))), s\underline{\gamma}'(t) \rangle dt \\
&= \int_I \langle \underline{f}(s\underline{\gamma}(t)), \underline{\gamma}'(t) \rangle dt + \int_I \langle \underline{\underline{D}}\, \underline{f}(s\underline{\gamma}(t))\underline{\gamma}(t), s\underline{\gamma}'(t) \rangle dt \\
&= \langle \underline{f}(s\underline{\gamma}(t)), \underline{\gamma}(t) \rangle |_{t=a}^{t=b} - \int_I \langle \underline{\underline{D}}\, \underline{f}(s\underline{\gamma}(t)) s\underline{\gamma}'(t), \underline{\gamma}(t) \rangle dt \qquad (18.2) \\
&\qquad\qquad + \int_I \langle \underline{\underline{D}}\, \underline{f}(s\underline{\gamma}(t))\underline{\gamma}(t), s\underline{\gamma}'(t) \rangle dt \\
&= s \int_I \langle \underline{\underline{D}}\, \underline{f}(s\underline{\gamma}(t))\underline{\gamma}(t), \underline{\gamma}'(t) \rangle - \langle \underline{\underline{D}}\, \underline{f}(s\underline{\gamma}(t))\underline{\gamma}'(t), \underline{\gamma}(t) \rangle dt \\
&= s \int_I \langle (\underline{\underline{D}}\, \underline{f}(s\underline{\gamma}(t))^\top - \underline{\underline{D}}\, \underline{f}(s\underline{\gamma}(t))\underline{\gamma}'(t)), \underline{\gamma}(t) \rangle dt
\end{aligned}$$

□

Theorem 18.8 (Cauchy'scher Integralsatz). *Sei $E \subseteq \mathbb{C}$ sternförmig mit Mittelpunkt $\underline{y} \in E$ und $f : E \to \mathbb{C}$ holomorph. Weiter sei $\gamma \in \mathcal{C}^0(I, E)$ eine geschlossene, stückweise $\mathcal{C}^1$ Kurve. Dann gilt*

$$\int_\gamma f(z) dz = 0.$$

Beweis. Wir erinnern an (18.1). Wir wenden Lemma 18.7 auf die Funktionen $\underline{g} := (f_1, -f_2)$ und $\underline{h} := (f_2, f_1)$ an. Wir bemerken, dass wegen der Gültigkeit der Cauchy-Riemann'schen Differentialgleichungen

$$\underline{\underline{D}}\, \underline{g}(x,y) = \begin{pmatrix} \frac{\partial f_1(x,y)}{\partial x} & \frac{\partial f_1(x,y)}{\partial y} \\ -\frac{\partial f_2(x,y)}{\partial x} & -\frac{\partial f_2(x,y)}{\partial y} \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1(x,y)}{\partial x} & \frac{\partial f_1(x,y)}{\partial y} \\ \frac{\partial f_1(x,y)}{\partial y} & -\frac{\partial f_1(x,y)}{\partial x} \end{pmatrix},$$

$$\underline{\underline{D}}\, \underline{h}(x,y) = \begin{pmatrix} \frac{\partial f_2(x,y)}{\partial x} & \frac{\partial f_2(x,y)}{\partial y} \\ \frac{\partial f_1(x,y)}{\partial x} & \frac{\partial f_1(x,y)}{\partial y} \end{pmatrix} = \begin{pmatrix} -\frac{\partial f_1(x,y)}{\partial y} & \frac{\partial f_1(x,y)}{\partial x} \\ \frac{\partial f_1(x,y)}{\partial x} & -\frac{\partial f_1(x,y)}{\partial y} \end{pmatrix}$$

impliziert. Insbesondere sind $\underline{\underline{D}}\, \underline{g}$ und $\underline{\underline{D}}\, \underline{h}$ symmetrisch. Nach Lemma 18.7 ist damit die rechte Seite von (18.1) Null. □

Bemerkung 18.9 (Zwei Erweiterungen). Sei $E \subseteq \mathbb{C}$ sternförmig mit Zentrum y. Weiter sei $\gamma_r(t) = e^{2\pi i t}$ für r klein genug, so dass $\gamma_r(I) \subseteq E$.

1. Sei und $f : E \setminus \{y\} \to \mathbb{C}$ holomorph. Dann ist
$$r \mapsto \int_{\gamma_r} f(z)dz$$
konstant.
Denn: Wir gehen wie im Beweis des Cauchy-schen Integralsatzes vor. Wieder betrachten wir die Funktionen $\underline{g} = (f_1, -f_2)$ und $\underline{h} = (f_2, f_1)$, die symmetrische Ableitungen haben. Nach (18.2) ist dann
$$\frac{\partial}{\partial r} \int_{\underline{\gamma}_r} \underline{g}(\underline{x})d\underline{x} = 0$$
und dasselbe gilt für die Funktion $\underline{h}$. Mit (18.1) folgt dann die Behauptung.
2. Ist weiter (etwas allgemeiner als im Cauchy'schen Integralsatz) $f : E \to \mathbb{C}$ holomorph auf $E \setminus \{y\}$ und stetig auf E. Dann ist
$$\int_{\gamma_r} f(z)dz = 0.$$
Denn: Mit Lemma 17.12 gilt
$$\Big|\int_{\gamma_r} f(z)dz\Big| \leq L(\gamma_r) \cdot ||f||_{\gamma_r(I)} \xrightarrow{r\to 0} 0.$$
Mit 1. folgt die Behauptung.

Theorem 18.10 (Cauchy'sche Integralformel). *Sei $E \subseteq \mathbb{C}$ und $f : E \to \mathbb{C}$ holomorph und r so, dass $B_r(z) \subseteq E$. Dann gilt mit $\gamma_r(t) = re^{2\pi i t}$*
$$f(z) = \frac{1}{2\pi i} \int_{\gamma_r} \frac{f(y)}{y - z} dy. \tag{18.3}$$

Bemerkung 18.11 (Interpretation). Das erstaunlichste an der Cauchy'schen Integralformel ist, dass die Funktionswerte einer holomorphen Funktionen auf einem Kreis $\gamma_r(I)$ Funktionswerte innerhalb des Kreises bestimmen. Schließlich hängt die rechte Seite von (18.3) nur von solchen Werten auf dem Kreis γ_r ab.

Beweis von Theorem 18.10. Nach Beispiel 18.6 ist
$$f(z) = \frac{1}{2\pi i} \int_{\gamma_r} \frac{f(z)}{y - z} dy.$$
Wir definieren nun die auf $E \setminus z$ holomorphe und auf E stetige Funktion
$$y \mapsto g(y) = \begin{cases} \frac{f(y)-f(z)}{y-z}, & y \neq z, \\ f'(z), & y = z. \end{cases}$$
Damit folgt die Behauptung mit Bemerkung 18.9.2 aus
$$\int_{\gamma_r} \frac{f(y) - f(z)}{y - z} dy = \int_{\gamma_r} g(y)dy = 0.$$
□

Als Anwendung der Cauchy'schen Integralformel und des Integralsatzes geben wir nun einen Beweis des Fundamentalsatzes der Algebra.

Theorem 18.12 (Fundamentalsatzes der Algebra). *Seien* $a_0, ..., a_n \in \mathbb{C}$ *und* $p : z \mapsto a_0 + a_1 z + a_2 z^2 + \cdots + a_n z^n$*. Dann gibt es mindestens ein* $z \in \mathbb{C}$ *mit* $p(z) = 0$*.*

Beweis. Wir schreiben $p(z) = z \cdot q(z) + a_0$. OBdA ist $z \neq 0$, sonst ist die Aussage trivial. Angenommen, p habe keine Nullstelle. Dann gilt

$$\frac{1}{z} = \frac{p(z)}{zp(z)} = \frac{q(z)}{p(z)} + \frac{a_0}{zp(z)} dz.$$

Nun ist $z \mapsto q(z)/p(z)$ holomorph und damit gilt für $\gamma_r = re^{2\pi it}$ mit Beispiel 18.6

$$2\pi i = \int_{\gamma_r} \frac{1}{z} dz = \int_{\gamma_r} \frac{a_0}{zp(z)} dz.$$

wegen dem Cauchy'schen Integralsatz. Nun ist wegen Lemma 17.12

$$\Big| \int_{\gamma_r} \frac{a_0}{zp(z)} dz \Big| \leq L(\gamma_r) |a_0| \Big|\Big| \frac{1}{zp(z)} \Big|\Big|_{\gamma_r} = 2\pi r |a_0| \cdot \Big|\Big| \frac{1}{p(z)} \Big|\Big|_{\gamma_r} \xrightarrow{r \to \infty} 0$$

da $|p(z)| \xrightarrow{z \to \infty} \infty$. Insgesamt haben wir also gezeigt

$$2\pi = \Big| \int_{\gamma_r} \frac{1}{z} dz = \int_{\gamma_r} \frac{a_0}{zp(z)} dz \Big| \xrightarrow{r \to \infty} 0,$$

ein Widerspruch. Daraus folgt die Behauptung. □

Bemerkung 18.13 (Ausblick in weitere Resultate). Der Cauchy'sche Integralsatz und die Integralformel sind in der Funktionentheorie zentral. Einerseits erfahren sie Verallgemeinerungen, insbesondere bei der Form der Menge E und dem Weg γ, über den integriert wirf. Während wir beim Integralsatz nur sternförmige Gebiete betrachtet hatten, gelingt es in der allgemeinen Theorie, *einfach zusammenhängende* Gebiete zu verwenden. Dies sind Gebiete, in denen sich geschlossene Wege immer auf einzelne Punkte *zusammen ziehen* lassen. Die kreisförmigen Wege im Integralsatz werden außerdem durch beliebige geschlossene stückweise $\mathcal{C}^1$-Kurven ersetzt. Schöne Aussagen über holomorphe Funktionen sind etwa:

- Jede auf einem Gebiet E holomorphe Funktion lässt sich lokal in einer Potenzreihe darstellen. (Wir erinnern daran, dass das für reelle Funktionen nicht gilt, siehe Beispiel 12.20.)

- Ist E einfach zusammenhängend, dann ist f auf E genau dann holomorph, wenn f eine Stammfunktion besitzt. (Auch dies gilt im reellen nicht: es gibt hier sehr wohl Funktionen, die integrierbar sind – und damit eine Stammfunktion besitzen – ohne dass die Funktionen differenzierbar wären.)

- Jede auf ganz $\mathbb{C}$ holomorphe, beschränkte Funktion ist konstant. (Auch dies gilt für Abbildungen im reellen nicht.)

18.2 Der Fixpunktsatz von Brouwer

Im Laufe der Vorlesung haben wir einige Male den Banach'schen Fixpunktsatz (Theorem 15.15) anwenden können, etwa im Kurzzeit-Existenzsatz von Picard-Lindelöf für Differentialgleichungen (Theorem 15.19) und bei der lokalen Umkehrbarkeit von Funktionen (Theorem 16.15). Dies deutet darauf hin, dass Fixpunktsätze in der Mathematik eine große Rolle spielen. Aus diesem Grund geben wir nun noch den wichtigen Fixpunktsatz von Brouwer an. Eine Verallgemeinerung erfährt dieser im Fixpunktsatz von Schauder (in dem $\mathbb{R}^m$ durch einen beliebigen vollständigen normierten Raum ersetzt wird).

Wir schreiben im Folgenden

$$B^n := \{\underline{x} : |\underline{x}| < 1\}, \qquad \overline{B}^n := \{\underline{x} : |\underline{x}| \leq 1\}, \qquad S^{n-1} := \partial B^n := \partial \overline{B}^n := \{\underline{x} : |\underline{x}| = 1\}.$$

Ziel dieses Kapitels ist der Beweis des Brouwer'schen Fixpunktsatzes (Theorem 18.17). Bevor wir mit dem Beweis beginnen können, benötigen wir zwei Lemmas.

Lemma 18.14 (Stetige Abbildungen $\overline{B}^n \to S^{n-1}$). *Es gibt keine Abbildung $\underline{f} \in \mathcal{C}^1(\overline{B}^n, S^{n-1})$ mit $\underline{f}(\underline{x}) = \underline{x}$ für $\underline{x} \in S^{n-1}$.*

Bemerkung 18.15 (Transformationssatz). Im Beweis werden wir einen Satz verwenden, der erst in der Vorlesung Analysis III bewiesen werden wird: Ist $E \subseteq \mathbb{R}^m$ und $\underline{f} : E \to \mathbb{R}^m$ ein Diffeomorphismus, dann ist

$$\lambda(\underline{F}(E)) = \int_E |\det(\underline{\underline{D}}\,\underline{f}(\underline{x}))| d\underline{x}.$$

Hierbei ist $\lambda(\underline{F}(E))$ das Volumen von $\underline{F}(E)$ as Teilmenge des $\mathbb{R}^m$ und die rechte Seite kann als Mehrfachintegral interpretiert werden.

Um die Richtigkeit dieser Formel etwas zu erläutern, betrachten wir $m = 2$ sowie die Abbildung $\underline{f} : \underline{x} \mapsto \underline{\underline{A}}\,\underline{x}$ für eine invertierbare Matrix $\underline{\underline{A}}$. Dann besteht $\underline{f}([0,1]^2)$ genau aus dem von $\underline{\underline{A}}(0,0)$, $\underline{x} := \underline{\underline{A}}\,(1,0)$, $\underline{y} := \underline{\underline{A}}\,(0,1)$ und $\underline{\underline{A}}(1,1)$ aufgespannten Parallelogramm. Um die Fläche von $\underline{f}([0,1]^2)$, also dem von $\underline{x}, \underline{y}$ aufgespannten Parallelogramms zu berechnen, projezieren wir $\underline{y}$ auf $\underline{x}^\perp := (-x_2, x_1)$ und erhalten dabei den Vektor

$$\underline{z} = \frac{\langle \underline{y}, \underline{x}^\perp \rangle}{|\underline{x}^\perp|^2} \underline{x}^\perp.$$

(Dieser Vektor steht offenbar auf $\underline{x}$ senkrecht und es gilt $|\langle \underline{y}, \underline{z} \rangle| = |\underline{z}|^2$.) Die Fläche des von $\underline{x}$ und $\underline{y}$ aufgespannten Parallelogramms ist nun wegen $|\underline{x}| = |\underline{x}^\top|$

$$\lambda(\underline{f}([0,1]^2)| = |\underline{x}| \cdot |\underline{z}| = |\langle \underline{y}, \underline{x}^\top \rangle| = |x_1 y_2 - x_2 y_1| = |\det(\underline{x}, \underline{y})| = |\det(\underline{\underline{A}})| = \int_{[0,1]^2} |\det(\underline{\underline{A}})| d\underline{x}.$$

Beweis von Lemma 18.14. Angenommen es gäbe ein $\underline{f} \in \mathcal{C}^1(\overline{B}^n, S^{n-1})$ mit der geforderten Eigenschaft. Wir definieren dann $\underline{g}(\underline{x}) := \underline{f}(\underline{x}) - \underline{x}$ und

$$\underline{F}(t, \underline{x}) := (1-t)\underline{x} + t\underline{f}(\underline{x}) = \underline{x} + t\underline{g}(\underline{x}).$$

Für $\underline{x} \in \overline{B}^n$ ist damit $|\underline{F}(t,\underline{x})| \leq (1-t)|\underline{x}|+t|\underline{f}(\underline{x})| \leq (1-t)+t = 1$, also $\underline{F}(t,.) \in \mathcal{C}^1(\overline{B}^n, \overline{B}^n)$ und für $\underline{x} \in S^{n-1}$

$$\underline{F}(t,\underline{x}) = (1-t)\underline{x} + t\underline{x} = \underline{x}.$$

Mit $\underline{f}$ ist auch $\underline{g}$ stetig differenzierbar, nach Bemerkung 15.9 insbesondere Lipschitz-stetig, d.h. es gibt ein $L \geq 0$ mit

$$|\underline{g}(\underline{y}) - \underline{g}(\underline{x})| \leq L|\underline{y} - \underline{x}|$$

für $\underline{x}, \underline{y} \in \overline{B}^n$. Wir zeigen nun:

> Behauptung: Es gibt ein $t_0 > 0$, so dass für alle $t \in [0, t_0]$ die Abbildung $\underline{F}(t,.) : B^n \to \underline{F}(t, B^n)$ ein Diffeomorphismus ist.

Wir verwenden Korollar 16.16. Hierzu zeigen wir, dass (i) $\det(\underline{\underline{D}}_{\underline{x}}\underline{F}(t,.)) \neq 0$ für t klein genug und (ii) für $t \in [0, 1/L)$ die Abbildung $\underline{F}(t,.)$ injektiv ist. Die Eigenschaft (i) ist einfach zu zeigen. Schließlich ist $\underline{\underline{D}}_{\underline{x}}\underline{F}(t,.) = (1-t)\underline{\underline{E}}_m + t\underline{\underline{D}}\,\underline{f}$. Damit ist $\underline{\underline{D}}_{\underline{x}}\underline{F}(0,.) = \underline{\underline{E}}_m$ und damit invertierbar. Weiter ist $t \mapsto \det(\underline{\underline{D}}_{\underline{x}}\underline{F}(t,\underline{x}))$ stetig, also muss es ein $t_0 > 0$ geben mit $\det(\underline{\underline{D}}_{\underline{x}}\underline{F}(t,.)) \neq 0$ für $t \leq t_0$. Für (ii) sei $t \in [0, 1/L)$. Angenommen, es wäre $\underline{x}, \underline{y} \in B^n$ mit $\underline{F}(t,\underline{x}) = \underline{F}(t,\underline{y})$, also

$$|\underline{y} - \underline{x}| = t|\underline{g}(\underline{y}) - \underline{g}(\underline{x})| \leq tC|\underline{y} - \underline{x}|.$$

Da $tC < 1$ nach Voraussetzung, ist dies nur für $\underline{y} = \underline{x}$ möglich. Damit ist $\underline{F}(t,.)$ injektiv.

Weiter zeigen wir nun:

> Behauptung: $\underline{F}(t, B^n) = B^n$ für $t \in [0, t_0]$.

Angenommen, es gäbe ein $t \in [0, t_0]$ und $\underline{y} \in B^n \setminus \underline{F}(t, B^n)$. OBdA ist $\underline{x} \in \partial\underline{F}(t, B^n) := \overline{\underline{F}(t, B^n)} \setminus \underline{F}(t, B^n)$, da $\underline{F}(t, B^n)$ offen ist. Sei $\underline{x}_1, \underline{x}_2, ... \in B^n$ so, dass $\underline{F}(t, \underline{x}_n) \xrightarrow{n\to\infty} \underline{y}$. Da $\overline{B}^n$ kompakt ist, können wir zu einer Teilfolge übergehen, die konvergiert. OBdA ist also $\underline{x}_n \xrightarrow{n\to\infty} \underline{x} \in \overline{B}^n$. Wegen der Stetigkeit von $\underline{F}(t,.)$ ist $\underline{F}(t,\underline{x}) = \underline{y}$. Damit muss $\underline{x} = \partial B^n = S^{n-1}$ sein. Nach Voraussetzung folgt daraus also $\underline{y} = \underline{F}(t,\underline{x}) = \underline{x} \in S^{n-1}$, ein Widerspruch zur Annahme, da $B^n \cap S^{n-1} = \emptyset$. Damit ist die Behauptung gezeigt.

Wir konstruieren nun den Widerspruch, der dann die Behauptung des Brouwer'schen Fixpunktsatzes zeigt. Hierfür betrachten wir

$$p : t \mapsto \int_{B^n} |\det(\underline{\underline{D}}_{\underline{x}}\underline{F}(t,\underline{x}))| d\underline{x}.$$

Nach Definition von $\underline{F}(t,\underline{x})$ ist dies ein Polynom in t. Weiter gilt mit dem Transformationssatz (siehe Bemerkung 18.15) für $t \in [0, t_0)$, dass

$$p(t) = \lambda(\underline{F}(t, B^n)) = \lambda(B^n) > 0.$$

Insbesondere ist das Polynom p auf $[0, t_0]$ konstant mit $p(0) > 0$. Dies bedeutet, dass p konstant sein muss, insbesondere also $p(t) > 0$ für alle t gelten muss. Andererseits gilt $|\underline{F}(1,\underline{x})|^2 = 1$ und damit für $j = 1, ..., m$

$$0 = \frac{\partial}{\partial x_j} \sum_{i=1}^{m} (f_i(1,\underline{x}))^2 = 2\sum_{i=1}^{m} f_i(1,\underline{x}) \frac{\partial f_i(1,\underline{x})}{\partial x_j} = 2\underline{F}(1,\underline{x})(\underline{\underline{D}}_{\underline{x}}(1,\underline{x}))_j,$$

also $\underline{F}(1,\underline{x})\underline{\underline{D}}_{\underline{x}}(1,\underline{x}) = \underline{0}$. Insbesondere hat $\underline{\underline{D}}_{\underline{x}}(1,\underline{x})$ den Eigenwert 0 und ist damit nicht invertierbar oder $\det(\underline{\underline{D}}_{\underline{x}}\underline{F}(1,\underline{x})) = 0$. Daraus folgt, dass $p(1) = 0$ gilt im Widerspruch zu $p(t) > 0$ für alle t. □

Lemma 18.16 (Best-Approximation). *Sei $E \subseteq \mathbb{R}^m$ kompakt und konvex. Dann gibt es für jedes $\underline{x} \in \mathbb{R}^m$ genau ein $\underline{p}(\underline{x})$ mit $|\underline{p}(\underline{x}) - \underline{x}| = \inf_{\underline{y}\in E} |\underline{y} - \underline{x}|$. Die Abbildung $\underline{p}$ heißt Best-Approximation, ist stetig und es gilt $\underline{p}(\underline{x}) = \underline{x}$ für $\underline{x} \in E$.*

Beweis. Sei zunächst $\underline{x} \in \mathbb{R}^m$ beliebig. Zunächst ist die Abbildung $\underline{y} \mapsto |\underline{y} - \underline{x}|$ stetig, und nimmt somit auf dem kompakten E ihr Infimum $\underline{p}(\underline{x})$ an. Dies stellt die Existenz der Abbildung $\underline{p}$ sicher. Wir müssen noch zeigen, dass (i) $\underline{p}$ eindeutig ist und (ii) $\underline{p}$ stetig ist. Hierfür setzen wir

$$\gamma_{\underline{x}} := \inf_{\underline{y}\in E} |\underline{y} - \underline{x}|.$$

Für (i) nehmen wir an, dass $\underline{p}'(\underline{x}) \neq \underline{p}(\underline{x})$ mit $\underline{p}'(\underline{x}) \in E$ ein weiterer Minimierer von $\underline{y} \mapsto |\underline{y}-\underline{x}|$ auf E ist. Wir setzen dann $\underline{z} := \frac{1}{2}(\underline{p}(\underline{x}) + \underline{p}'(\underline{x}))$. Da E konvex ist, muss $\underline{y} \in E$ sein. Außerdem gilt

$$|\underline{z} - \underline{x}| = \tfrac{1}{2}\Big|\underline{p}(\underline{x}) - \underline{x} + \underline{p}'(\underline{x}) - \underline{x}\Big| \leq \tfrac{1}{2}\Big|\underline{p}(\underline{x}) - \underline{x}\Big| + \Big|\underline{p}'(\underline{x}) - \underline{x}\Big| = \gamma_{\underline{x}},$$

wobei '=' nur dann gilt, wenn $\underline{p}(\underline{x})-\underline{x}$ und $\underline{p}'(\underline{x})-\underline{x}$ parallel sind, also wenn $\underline{p}(\underline{x})-\underline{x} = \underline{p}'(\underline{x})-\underline{x}$, was aber laut Voraussetzung nicht der Fall ist. Also haben wir gezeigt, dass $|\underline{z} - \underline{x}| < \gamma_{\underline{x}}$, also ein Widerspruch. Damit ist die Eindeutigkeit des Minimierers $\underline{p}(\underline{x})$ gezeigt.
Für (ii) nehmen wir eine Folge $\underline{x}, \underline{x}_1, \underline{x}_2, ... \in \mathbb{R}^m$ mit $\underline{x}_m \xrightarrow{n\to\infty} \underline{x}$. Sei $\varepsilon > 0$ und N groß genug, so dass $|\underline{x}_n - \underline{x}| < \varepsilon$ für $n \geq N$. Nun gilt

$$\gamma_{\underline{x}} \leq |\underline{p}(\underline{x}_n) - \underline{x}| \leq |\underline{p}(\underline{x}_n) - \underline{x}_n| + |\underline{x}_n - \underline{x}| \leq \gamma_{\underline{x}_n} + \varepsilon.$$

Andererseits ist

$$\gamma_{\underline{x}_n} = \inf_{\underline{y}\in E} |\underline{y} - \underline{x}_n| \leq \inf_{\underline{y}\in E} |\underline{y} - \underline{x}| + |\underline{x} - \underline{x}_n| \leq \gamma_{\underline{x}} + \varepsilon.$$

Aus den letzten beiden Gleichungen lesen wir ab, dass

$$\gamma_{\underline{x}} \leq |\underline{p}(\underline{x}_n) - \underline{x}| \leq \gamma_{\underline{x}_n} + 2\varepsilon.$$

Das bedeutet, dass $|\underline{p}(\underline{x}_n) - \underline{x}| \xrightarrow{n\to\infty} \gamma_{\underline{x}} = |\underline{p}(\underline{x}) - \underline{x}|$. Aus der Eindeutigkeit von $\underline{p}(\underline{x})$ folgt die Stetigkeit von $\underline{p}$. □

Nun können wir den Beweis des Brouwer'schen Fixpunktsatzes führen.

Theorem 18.17 (Brouwer'scher Fixpunktsatz). *Sei $E \subseteq \mathbb{R}^m$ konvex und kompakt sowie $\underline{f} \in \mathcal{C}^0(E, E)$. Dann besitzt $\underline{f}$ einen Fixpunkt.*

Bemerkung 18.18 (Spezialfall $m = 1$). Im Spezialfall $m = 1$ ist der Beweis ganz einfach: Dann ist nämlich $E = [a, b]$ ein kompaktes Intervall. Für eine Abbildung $f : [a, b] \to [a, b]$ gilt nun

$$f(a) - a \geq 0, \qquad f(b) - b \leq 0.$$

Damit hat nach Korollar 5.20 die Abbildung $x \mapsto f(x) - x$ mindestens eine Nullstelle, d.h. ein x mit $f(x) = x$, also einen Fixpunkt von f.

Beweis von Theorem 18.17. Der Beweis erfolgt nun in drei Schritten:
Schritt 1: Zunächst nehmen wir an, dass $\underline{f} \in \mathcal{C}^1(B^m, B^m)$. Wir zeigen:

Behauptung: Sei $\underline{f} \in \mathcal{C}^1(B^m, B^m)$. Dann besitzt $\underline{f}$ einen Fixpunkt.

Angenommen, $\underline{f}$ hätte keinen Fixpunkt. Dann gibt es zu jedem $\underline{x} \in B^m$ einen Schnittpunkt der Geraden durch $\underline{x}$ und $\underline{f}(\underline{x})$ durch S^{m-1}. Diese ist gegeben durch die Abbildung $\underline{F}$: $[0,1] \times B^{m-1} \to S^{m-1}$ mittels

$$\underline{F}(t, \underline{x}) = \underline{x} + t\lambda(\underline{x}) \frac{\underline{x} - \underline{f}(\underline{x})}{|\underline{x} - \underline{f}(\underline{x})|},$$

für

$$\lambda(\underline{x}) = \Big(1 - |x|^2 + \Big\langle \underline{x}, \frac{\underline{x} - \underline{f}(\underline{x})}{|\underline{x} - \underline{f}(\underline{x})|} \Big\rangle^2 \Big)^{1/2} - \Big\langle \underline{x}, \frac{\underline{x} - \underline{f}(\underline{x})}{|\underline{x} - \underline{f}(\underline{x})|} \Big\rangle.$$

Denn: Die rechte Seite ist von der Form $\underline{x} + tb_{\underline{x}}(\underline{x} - \underline{f}(x))$ für eine Konstante $b_{\underline{x}}$, was für variables t eine Gerade durch $\underline{x}$ und $\underline{f}(\underline{x})$ darstellt. Weiter ist

$$\begin{aligned} |\underline{F}(1, \underline{x})|^2 &= |x|^2 + \lambda(\underline{x})^2 + 2\lambda(\underline{x}) \Big\langle \underline{x}, \frac{\underline{x} - \underline{f}(\underline{x})}{|\underline{x} - \underline{f}(\underline{x})|} \Big\rangle \\ &= 1 + 2\Big\langle \underline{x}, \frac{\underline{x} - \underline{f}(\underline{x})}{|\underline{x} - \underline{f}(\underline{x})|} \Big\rangle^2 + 2\Big(\lambda(\underline{x}) - \Big(1 - |x|^2 + \Big\langle \underline{x}, \frac{\underline{x} - \underline{f}(\underline{x})}{|\underline{x} - \underline{f}(\underline{x})|} \Big\rangle^2 \Big)^{1/2}\Big) \Big\langle \underline{x}, \frac{\underline{x} - \underline{f}(\underline{x})}{|\underline{x} - \underline{f}(\underline{x})|} \Big\rangle \\ &= 1. \end{aligned}$$

Für $\underline{x} \in S^{n-1}$ ist $\lambda(\underline{x}) = 0$ und damit $\underline{F}(1, \underline{x}) = \underline{x}$. Aus Lemma 18.14 folgt nun, dass es die Abbildung $\underline{x} \mapsto \underline{F}(1, \underline{x})$ nicht geben kann, also einen Widerspruch. Daraus folgt die Behauptung.

Schritt 2: Wir verallgemeinern Schritt 1 nun für Funktionen $\underline{f} \in \mathcal{C}^0(B^m, B^m)$. Sei hierzu $\underline{g}_1, \underline{g}_2, ... \in \mathcal{C}^1(B^m, B^m)$ und so, dass $||\underline{g}_n - \underline{f}|| \to 0$. (Eine solche Folge existiert wegen des Satzes von Stone, Theorem 10.12.) Sei nun $\underline{x}_n$ ein Fixpunkt von $\underline{p}_n$. Durch Übergang zu einer Teilfolge konvergiert $\underline{x}_n \xrightarrow{n \to \infty} \underline{x}$. Damit gilt

$$|\underline{f}(\underline{x}) - \underline{p}_n(\underline{x}_n)| \leq |\underline{f}(\underline{x}) - \underline{f}(\underline{x}_n)| + |\underline{f}(\underline{x}_n) - \underline{p}_n(\underline{x}_n)| \xrightarrow{n \to \infty} 0$$

wegen der Stetigkeit von $\underline{f}$ und $\underline{p}_n \xrightarrow{n \to \infty}_{glm} \underline{f}$. Daraus folgt sofort

$$\underline{f}(\underline{x}) = \lim_{n \to \infty} \underline{p}_n(\underline{x}_n) = \lim_{n \to \infty} \underline{x}_n = \underline{x},$$

also die Aussage.

Schritt 3: Wir verallgemeinern das Resultat von Schritt 2 nun für konvexes, kompaktes E. Sei r so groß, dass $E \subseteq \overline{B}_r(\underline{0})$. Sei $\underline{p} : \overline{B}_r(\underline{0}) \to E$ die stetige Best-Approximation aus Lemma 18.16. Klar ist nun, dass $\underline{f} \circ \underline{p} \in \mathcal{C}^0(\overline{B_r(\underline{0}), B_r(\underline{0})})$ mit $\underline{f} \circ \underline{p}(B_r(\underline{0})) \subseteq E$. Nach Schritt 2 gibt es also ein $\underline{x} \in B_r(\underline{0})$ mit $\underline{x} = \underline{f} \circ \underline{p}(\underline{x}) \in E$. Für $\underline{x} \in E$ ist jedoch $\underline{p}(\underline{x}) = \underline{x}$ und damit ist $\underline{x} = \underline{f}(\underline{x})$ und der Fixpunktsatz ist gezeigt. □

Als Anwendung des Brouwer'schen Fixpunktsatzes zeigen wir nun die Lösbarkeit einer bestimmten Klasse nicht-linearer Gleichungen.

Proposition 18.19. *Sei $E \subseteq \mathbb{R}^m$ und $\underline{g} = (g_1, ..., g_m) \in \mathcal{C}^0(E, E)$, so dass*

$$\langle \underline{g}(\underline{x}), \underline{x} \rangle \geq 0, \text{ für alle } \underline{x} \in B_r(\underline{0}).$$

Dann hat die Gleichung

$$\underline{g}(\underline{x}) = \underline{0}$$

mindestens eine Lösung $\underline{x} \in B_r(\underline{0})$.

Beweis. Angenommen, es gäbe keine Lösung der Gleichung. Dann definieren wir

$$\underline{f} := -r \frac{\underline{g}}{|\underline{g}|}.$$

Da $\underline{g}$ stetig ist und $|\underline{g}(\underline{x})| > 0$ für alle $\underline{x} \in B_r(\underline{0})$, ist $\underline{f}$ stetig. Für $\underline{x} \in B_r(\underline{0})$ gilt außerdem $|\underline{f}(\underline{x})| = r$, und damit $\underline{f}(\overline{B}_r(\underline{0})) \subseteq \overline{B}_r(\underline{0})$. Damit gibt es nach dem Brouwer'schen Fixpunktsatz ein $\underline{x} \in B_r(\underline{0})$ mit $\underline{x} = \underline{f}(\underline{x}) \in \partial B_r(\underline{0})$. Daraus folgt

$$0 \leq \langle \underline{g}(\underline{x}), \underline{x} \rangle = -r|\underline{g}(\underline{x})| \langle \underline{f}(\underline{x}), \underline{x} \rangle = -r|\underline{g}(\underline{x})| \langle \underline{x}, \underline{x} \rangle < 0,$$

ein Widerspruch. □

Beispiel 18.20 (Ein Wirtschaftsmodell). Als weitere Anwendung bringen wir ein Beispiel aus der Wirtschaft. Es soll erforscht werden, ob ein Markt mit n Produzenten ein Gleichgewicht besitzt. Sei x_i das Einkommen des Produzenten i und $f_{ij}(x_i)$ die Ausgaben von Produzenten i an Produzenten j, wenn Produzent i gerade Einkommen x_i hat. Wenn Produzent i all seine Einnahmen wieder ausgeben will, gilt also

$$x_i = \sum_{j=1}^{n} f_{ij}(x_i). \tag{18.4}$$

Andererseits bekommt Produzent j seine Einnahmen von allen anderen Produzenten, so dass

$$x_j = \sum_{i=1}^{n} f_{ij}(x_i) \tag{18.5}$$

gelten muss. Die Frage ist nun, ob es $\underline{x} = (x_1, ..., x_n)$ gibt, so dass (18.4) und (18.5) erfüllt sind. Wir zeigen hierzu:

> Sei $r > 0$ und alle f_{ij} stetig und so, dass (18.4) erfüllt ist. Dann gibt es mindestens eine Lösung $\underline{x}$ in der Menge $E := \{\underline{x} \in \mathbb{R}^n_+ : \langle \underline{x}, \underline{1} \rangle \leq r\}$ von (18.5).

Denn: Die Menge E ist kompakt und konvex. Wir betrachten die stetige Funktion $\underline{g} = (g_1, ..., g_n)$ mit $g_j(\underline{x}) = \sum_{i=1}^n f_{ij}(x_i)$. Dann hat (18.5) genau dann eine Lösung $\underline{x}$, wenn $\underline{g}(\underline{x}) = \underline{x}$. Um die Existenz eines solchen Fixpunktes zu finden, berechnen wir für $\underline{x}$ mit $\langle \underline{x}, \underline{1} \rangle \leq r$

$$\langle \underline{g}(\underline{x}), 1 \rangle = \sum_{j=1}^{n} \sum_{i=1}^{n} f_{ij}(x_i) = \sum_{i=1}^{n} x_i \leq r.$$

Nun folgt die Behauptung nach dem Brouwer'schen Fixpunktsatz.

Bemerkung 18.21 (Don't trust in books). In manchen Skripten findet sich als Anwendung des Brouwer'schen Fixpunktsatzes ein Beweis für den Fundamentalsatz der Algebra:

Beweis von Theorem 18.12. OBdA sei $a_n = 0$ (da das Polynom p und p/a_n dieselben Nullstellen besitzen). Wir setzen

$$r := 2 + |a_0| + \cdots + |a_{n-1}|$$

und

$$f(z) := z - \frac{((|z| \wedge 1)/z)^{n-1}}{r} p(z).$$

Dann ist f stetig und wir zeigen nun, dass $f|_{\overline{B}_r(0)} \subseteq \overline{B}_r(0)$ und wenden dann der Brouwer'schen Fixpunkt satz an. Sei zunächst $z \in \overline{B}_1(0)$ (und damit $|z| \wedge 1 = |z|$). Dann gilt

$$|f(z)| \leq |z| + \left| \frac{(|z|/z)^{n-1}}{r} p(z) \right| \leq 1 + \frac{|a_0| + \cdots + |a_{n-1}| + 1}{2 + |a_0| + \cdots + |a_{n-1}|} \leq 2 \leq r.$$

Für $z \in \overline{B}_r(0) \setminus \overline{B}_1(0)$ (und damit $|z| \wedge 1 = 1$) gilt außerdem

$$\begin{aligned} |f(z)| &= \left| z - \frac{p(z)}{rz^{n-1}} \right| = \left| z - \frac{a_0 + a_1 z + \cdots + a_{n-1} z^{n-1}}{rz^{n-1}} - \frac{z}{r} \right| \\ &\leq \left(1 - \frac{1}{r}\right) r + \frac{|a_0| + |a_1| + \cdots + |a_{n-1}|}{r} \leq r - 1 + 1 = r. \end{aligned}$$

Wir haben also gezeigt, dass $f(B_r(0)) \subseteq B_r(0)$. Also gibt es nach dem Brouwer'schen Fixpunktsatz ein $z \in B_r(0)$ mit $f(z) = z$. Dies ist aber nur dann möglich, wenn $p(z) = 0$. Also hat p eine Nullstelle. $\square$

Allerdings ist der Beweis an einer Stelle fehlerhaft. Können Sie sagen, wo?

Teil V
Maß- und Integrationstheorie

In Kapitel 8 haben wir bereits das Riemann-Integral $\int_a^b f(x)dx$ kennengelernt. Hier war f eine stückweise stetige Funktion. Ziel dieses Abschnittes ist es, diesen Integralbegriff zu erweitern. Zentraler Punkt ist die Entwicklung des Lebensgue-Integrals. (Wir werden hierbei unsere Notation strapazieren und auch für das Lebensgue-Integral $\int_a^b f(x)dx$ schreiben.) Dieses neue Integral ist mindestens aus zwei Gründen eine Erweiterung des alten:

1. Die Funktion f muss nicht notwendigerweise stückweise stetig sein.

2. Konvergenzsätze für das Lebesgue-Integral sind viel stärker als die entsprechenden Aussagen beim Riemann-Integral (siehe etwa Theorem 8.19).

Bevor wir jedoch mit der Konstruktion loslegen können, müssen wir das Lebesgue-Maß (das dem entsprechenden Integral zugrunde liegt) einführen. (Um sich ein Maß vorzustellen, sei $A = [a, b]$ ein Intervall. Ein(e Vorstufe eines) Maß(es) auf der Menge der reellen Zahlen ist dann die Abbildung $A \mapsto \int_a^b 1dx = b - a$.) Da die Konstruktion von Maßen in anderen Bereichen verallgemeinert werden wird (vor allem für die Wahrscheinlichkeitstheorie), werden wir dieses möglichst allgemein halten. Dies bedeutet insbesondere, dass wir auch auf anderen Mengen als den reellen Zahlen Maße definieren wollen. Deswegen bezeichnet $E \neq \emptyset$ in diesem Abschnitt (irgend)eine Menge. Weiter bezeichnet 2^E die Potenzmenge von E und jedes $\mathcal{A} \subseteq 2^E$ ist ein Mengensystem. Alle betrachteten Mengensysteme seien o.E. nicht leer. Bevor wir mit der Konstruktion von Maßen beginnen können, wiederholen wir einige wichtige Grundlagen.

19 Wiederholung Topologie

Topologien werden in der Mathematik immer dann gebraucht, wenn ein Konvergenzbegriff eingeführt wird. Auch wenn Topologien in der bisherigen Teilen nur am Rande behandelt wurden, sind doch einige Konvergenzbegriffe bekannt (siehe etwa die Unterscheidung zwischen punktweiser und gleichmäßiger Konvergenz von Funktionen in Kapitel 8).

Wir lernen Topologien als Mengensysteme kennen, die die Grundlage für weitere Mengensysteme bilden (σ-Algebren), auf denen wir Maße definieren. Wir werden also Maße auf topologischen Räumen betrachten. Dabei werden die zugrunde liegenden σ-Algebren von Topologie erzeugt (siehe Definition 20.7). Deshalb wiederholen wir zunächst Grundbegriffe der Topologie.

19.1 Grundlagen

Unter einer *Topologie* versteht man eine Familie offener Teilmengen eines Grundraumes E[22]. In metrischen Räumen nennt man eine Menge A genau dann offen, wenn für jedes $x \in A$ auch ein offener Ball $B_\varepsilon(x) \subseteq A$ für ein $\varepsilon > 0$. Dieser Fall von metrischen Räumen ist in der Praxis auch am wichtigsten. In der Maßtheorie kommt dem Fall von separablen Topologien, die von

[22] Man beachte hier einen feinen Unterschied zu Definition 5.7: dort hatten wir auf Grundlage einer Metrik definiert, was eine offene Menge ist. Dies bedeutet, dass sich Topologien aus Metriken definieren lassen. Dies ist jedoch nicht notwendigerweise so. Es gibt also auch Topologien (Mengen offener Mengen), die nicht auf Metriken basieren.

vollständigen Metriken erzeugt werden, eine besondere Bedeutung zu. Solche Räume heißen *polnisch*.

Definition 19.1. 1. *Eine Metrik r auf E heißt* vollständig*, wenn jede Cauchy-Folge konvergiert. Ist also $x_1, x_2, \ldots \in E$ mit*

$$\forall \varepsilon > 0 \ \ \exists N \in \mathbb{N} \ \ \forall n, m \geq N : \ \ r(x_n, x_m) < \varepsilon,$$

so gibt es ein $x \in E$ mit $r(x_n, x) \xrightarrow{n\to\infty} 0$.

2. *Ein Mengensystem $\mathcal{O} \subseteq 2^E$ heißt* Topologie*, falls (i) $\emptyset, E \in \mathcal{O}$, (ii) ist $A, B \in \mathcal{O}$, so ist auch $A \cap B \in \mathcal{O}$ (iii) ist I beliebig und ist $A_i \in \mathcal{O}, i \in I$, so ist auch $\bigcup_{i\in I} A_i \in \mathcal{O}$. Das Paar $(E, \mathcal{O})$ heißt* topologischer Raum*. Mengen $A \in \mathcal{O}$ heißen* offen*, Mengen $A \subseteq E$ mit $A^c \in \mathcal{O}$ heißen* abgeschlossen.

3. *Ist $(E, \mathcal{O})$ ein topologischer Raum und $A \subseteq E$. Dann heißt*

$$A^\circ := \bigcup \{O \subseteq A : O \in \mathcal{O}\}$$

das Innere von A *und*

$$\overline{A} := \bigcap \{F \supseteq A : F^c \in \mathcal{O}\}$$

den Abschluss von A.

4. *Ein topologischer Raum $(E, \mathcal{O})$ heißt* separabel*, wenn es eine abzählbare Menge $E' \subseteq E$ gibt mit $\overline{E}' = E$.*

5. *Sei $(E, \mathcal{O})$ ein topologischer Raum und $\mathcal{B} \subseteq \mathcal{O}$. Dann heißt $\mathcal{B}$ eine* Basis *von $\mathcal{O}$, falls*

$$\forall A \in \mathcal{O} \ \ \forall x \in A \ \ \exists B \in \mathcal{B} : x \in B \subseteq A.$$

Dies ist genau dann der Fall, wenn

$$\mathcal{O} = \{A \subseteq E : \ \ \forall x \in A \ \ \exists B \in \mathcal{B} : x \in B \subseteq A\}. \tag{19.1}$$

oder (äquivalent dazu)

$$\mathcal{O} = \Big\{ \bigcup_{i\in I} B_i : B_i \in \mathcal{B}, i \in I, \ I \text{ beliebig} \Big\}. \tag{19.2}$$

6. *Sei $\mathcal{B} \subseteq 2^E$. Dann wird durch* (19.1) *oder* (19.2) *eine Topologie $\mathcal{O}(\mathcal{B})$ definiert, die* von $\mathcal{B}$ erzeugte Topologie.

7. *Sei (E, r) ein metrischer Raum und*

$$\mathcal{B} := \{B_\varepsilon(x) : \varepsilon > 0, x \in E\}. \tag{19.3}$$

Dann ist $\mathcal{O}(\mathcal{B})$ die von r erzeugte Topologie*. Falls speziell $E \subseteq \mathbb{R}^d$ und r der euklidische Abstand ist, heißt die in* (19.1) *oder* (19.2) *definierte Topologie die* euklidische Topologie.

8. *Der Raum $(E, \mathcal{O})$ heißt (vollständig) metrisierbar, wenn es eine (vollständige) Metrik r auf E gibt, so dass* (19.1) *gilt. Der Raum $(E, \mathcal{O})$ heißt* polnisch*, falls er separabel und vollständig metrisierbar ist.*

9. *Seien $(E, \mathcal{O})$ und $(E', \mathcal{O}')$ topologische Räume. Dann heißt eine Abbildung $f : E \to E'$* stetig*, falls $f^{-1}(A') \in \mathcal{O}$ für alle $A' \in \mathcal{O}'$ gilt.*

Beispiel 19.2 (Der Raum $\mathbb{R}^d$). Wir haben bereits den Raum $\mathbb{R}^d$ und die euklidische Metrik r_{eukl} in Definition 5.2 kennen gelernt, und darauf aufbauend den Konvergenzbegriff in Definition 5.3. In Beispiel 15.13 haben wir gesehen, dass $(\mathbb{R}, r)$ vollständig ist (und dasselbe gilt auch für $(\mathbb{R}^d, r_{\text{eukl}})$. Weiter ist

$$\mathcal{B} := \{B_\varepsilon(x) : x, \varepsilon \in \mathbb{Q}, \varepsilon > 0\}.$$

die (abzählbare) Menge aller offenen Intervalle mit rationalen Radien und rationalen Mittelpunkten. Dann gilt nach dem eben gesagten (19.1), d.h. eine Menge $A \subseteq \mathbb{R}$ ist genau dann offen, wenn mit $x \in A$ auch ein rationales $\varepsilon > 0$ existiert mit $B_\varepsilon(x) \subseteq A$. Das bedeutet, dass die obige Definition der Basis $\mathcal{B}$ mit Definition 5.7 im Einklang steht.

Wir zeigen nun: $[a, b]^\circ = (a, b)$ und $\overline{(a, b)} = [a, b]$. (Analog gilt $\left(\prod_{i=1}^d [a_i, b_i]\right)^\circ = \prod_{i=1}^d (a_i, b_i)$ und $\overline{\prod_{i=1}^d (a_i, b_i)} = \prod_{i=1}^d [a_i, b_i]$.)
Denn: Offenbar ist (a, b) offen. (Mit jedem $x \in (a, b)$ ist nämlich $B_{\min(|x-a|,|x-b|)}(x) \subseteq A$.) Also gilt nach der Definition des Inneren, dass $[a, b] \supseteq [a, b]^\circ \supseteq (a, b)$. Wäre nun $a \in [a, b]^\circ$, so müsste auch $B_\varepsilon(a) \subseteq [a, b]^\circ \subseteq [a, b]$ gelten, was ein Widerspruch ist.
Weiter gilt offenbar für jedes $A \subseteq E$

$$\overline{A}^c = \left(\bigcap\{F \supseteq A : F^c \in \mathcal{O}\}\right)^c = \bigcup\{F \subseteq A^c : F \in \mathcal{O}\} = (A^c)^\circ, \tag{19.4}$$

und damit

$$\overline{(a, b)} = \overline{((a, b)^c)^c} = (((a, b)^c)^\circ)^c = ((-\infty, a] \cup [b, \infty))^\circ)^c = ((-\infty, a) \cup (b, \infty))^c = [a, b].$$

Weiter gilt: $\mathbb{R}$ ist separabel und insbesondere gilt $\overline{\mathbb{Q}} = \mathbb{R}$.
Denn: Es genügt, $\overline{\mathbb{Q}} = \mathbb{R}$ zu zeigen. Dies ist nach (19.4) dasselbe wie $(\mathbb{Q}^c)^\circ = \emptyset$. Gäbe es nämlich $x \in (\mathbb{Q}^c)^\circ$, so müsste es (da $(\mathbb{Q}^c)^\circ$ offen ist) ein $\varepsilon > 0$ geben mit $B_\varepsilon(x) \subseteq (\mathbb{Q}^c)^\circ \subseteq \mathbb{Q}^c$, also $B_\varepsilon(x) \cap \mathbb{Q} = \emptyset$. Sei nun $x_1 := x - \frac{\varepsilon}{2} < x_2 := x + \frac{\varepsilon}{2}$, also $x_1, x_2 \in B_\varepsilon(x)$. Nach Lemma 2.18 gibt es $y \in (x_1, x_2) \cap \mathbb{Q}$ im Widerspruch zu $B_\varepsilon(x) \cap \mathbb{Q} = \emptyset$.

Beispiel 19.3 (Der Raum $\overline{\mathbb{R}}$). Wir werden oftmals Funktionen mit Werten in

$$\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\} \quad \text{oder} \quad \overline{\mathbb{R}}_+ := \mathbb{R}_+ \cup \{\infty\}$$

betrachten.[23] Um diese Räume als topologische Räume betrachten zu können, setzen wir

$$\varphi : \begin{cases} \overline{\mathbb{R}} & \to [-1, 1], \\ x & \mapsto \begin{cases} \frac{2}{\pi} \arctan(x), & x \in \mathbb{R}, \\ 1, & x = \infty, \\ -1, & x = -\infty \end{cases} \end{cases}$$

[23] Die Schreibweise $\overline{\mathbb{R}}$ suggeriert, dass hier der Abschluss von $\mathbb{R}$ gemeint ist. Dies stimmt nicht, da die hinzugewonnenen Elemente $-\infty, \infty$ nicht in $\mathbb{R}$ liegen, Abschlüsse von Mengen aber immer höchstens die Elemente des Grundraumes enthalten können. Topologisch gesehen ist $\overline{\mathbb{R}}$ die Zwei-Punkte-Kompaktifizierung von $\mathbb{R}$.

und definieren die Metrik

$$r_{\overline{\mathbb{R}}}(x, y) := |\varphi(x) - \varphi(y)|, \qquad x, y \in \overline{\mathbb{R}}.$$

Der von $r_{\overline{\mathbb{R}}}$ definierte topologische Raum $(\overline{\mathbb{R}}, \overline{\mathcal{O}})$ erweitert die euklidische Topologie $(\mathbb{R}, \mathcal{O})$ auf $\mathbb{R}$ insofern, als dass $\{A \cap \mathbb{R} : A \in \overline{\mathcal{O}}\} = \mathcal{O}$. Dies gilt deshalb, weil φ stetig auf $\mathbb{R}$ ist mit stetiger Umkehrfunktion. Weiter gilt, dass $(\overline{\mathbb{R}}, \overline{\mathcal{O}})$ separabel ist und $r_{\overline{\mathbb{R}}}$ ist eine vollständige Metrik.

Auf $\overline{\mathbb{R}}$ kann man wie in der Analysis gewohnt rechnen. Etwa ist $a \cdot \infty = \infty$ für $a > 0$. Allerdings sind die Ausdrücke wie $\infty - \infty$ und ∞/∞ nicht definiert.

Bemerkung 19.4 (Zusammenhang zwischen Metrik und Topologie). Sei $(E, \mathcal{O})$ ein topologischer Raum und $x, x_1, x_2, \ldots \in E$. Wir definieren

$$x_n \xrightarrow{n\to\infty} x :\Longleftrightarrow \forall O \in \mathcal{O} : x \in O \Rightarrow x_n \in O \text{ für fast alle } n \in \mathbb{N}. \tag{19.5}$$

Insbesondere ergibt so jede Topologie auf E einen Konvergenzbegriff für Folgen in E.

Dieser Konvergenzbegriff stimmt mit dem auf metrischen Räumen bekannten Begriff aus Definition 5.3 überein: ist nämlich r eine Metrik auf E, die $\mathcal{O}$ erzeugt, dann gilt die rechte Seite aus (19.5) genau dann, wenn für alle $\varepsilon > 0$ gilt, dass $r(x_n, x) < \varepsilon$ für fast alle $n \in \mathbb{N}$.

Mit Hilfe des Konvergenzbegriffes aus (19.5) sehen wir nun folgende bekannte Eigenschaft ein:

Lemma 19.5 (Abschluss bei metrischen Räumen). *Sei (E, r) ein metrischer Raum und $\mathcal{O}$ die von r erzeugte Topologie. Für $F \subseteq E$ sind äquivalent:*

1. *F ist abgeschlossen.*
2. *Für alle $x_1, x_2, \ldots \in F$ und $x \in E$ mit $x_n \xrightarrow{n\to\infty} x$ gilt $x \in F$.*

Insbesondere gilt: für jedes $A \subseteq E$ besteht der Abschluss $\overline{A}$ genau aus den Häufungspunkten von A.

Beweis. '1.⇒2.': Angenommen, es gäbe $x_1, x_2, \ldots \in F$, konvergent gegen $x \in F^c$, dann wäre, da $F^c \in \mathcal{O}$ auch $x_n \in F^c$ für fast alle n. Dies ist im Widerspruch zur Voraussetzung.
'2.⇒1.': Angenommen, F wäre nicht abgeschlossen, F^c also nicht offen. Dann gibt es $x \in F^c$, so dass für alle $\varepsilon > 0$ gilt, dass $B_\varepsilon(x) \not\subseteq F^c$. Wähle $\varepsilon_1, \varepsilon_2, \ldots > 0$ mit[24] $\varepsilon_n \downarrow 0$ und $x_n \in B_{\varepsilon_n}(x) \cap F$. Dann gilt $x_1, x_2, \ldots \in F$, $x_n \xrightarrow{n\to\infty} x$, aber $x \in F^c$. □

Lemma 19.6 (Abzählbare Basis und separable Räume). *Sei (E, r) ein separabler, metrischer Raum, $\mathcal{O}$ die von r erzeugte Topologie, E' abzählbar mit $\overline{E'} = E$ und*

$$\widetilde{\mathcal{B}} := \{B_\varepsilon(x) : \varepsilon \in \mathbb{Q}_+, x \in E'\}.$$

Dann ist $\widetilde{\mathcal{B}}$ abzählbar und $\mathcal{O}(\widetilde{\mathcal{B}}) = \mathcal{O}$.

Beweis. Klar ist, dass $\widetilde{\mathcal{B}}$ abzählbar ist und $\mathcal{O}(\widetilde{\mathcal{B}}) \subseteq \mathcal{O}$. Sei $\mathcal{B}$ wie in (19.3). Dann ist für $B_\varepsilon(x) \in \mathcal{B}$

$$B_\varepsilon(x) = \bigcup_{\widetilde{\mathcal{B}} \ni B \subseteq B_\varepsilon(x)} B,$$

also $\mathcal{B} \subseteq \mathcal{O}(\widetilde{\mathcal{B}})$ und damit $\mathcal{O} = \mathcal{O}(\mathcal{B}) \subseteq \mathcal{O}(\widetilde{\mathcal{B}})$. □

---

[24] Wir schreiben $\varepsilon_n \downarrow 0$ falls $\varepsilon_1 \geq \varepsilon_2 \geq \ldots$ und $\varepsilon_n \xrightarrow{n\to\infty} 0$

Beispiel 19.7 (Zwei polnische Räume). 1. Sei $\mathcal{O}$ die euklidische Topologie auf $\mathbb{R}^d$. Diese wird nach Definition 19.1.9 durch die euklidische Metrik definiert. Bekannt ist, dass diese vollständig ist. Weiter ist $\mathbb{Q}^d$ abzählbar und jedes $x \in \mathbb{R}^d$ ist Häufungspunkt einer Folge in $\mathbb{Q}^d$. Insbesondere ist also $\overline{\mathbb{Q}^d} = \mathbb{R}^d$ nach Lemma 19.5, also ist $\mathbb{R}^d$ separabel. Insgesamt ist also $(\mathbb{R}^d, \mathcal{O})$ polnisch.

2. Sei $K \subseteq \mathbb{R}$ kompakt und $E = \mathcal{C}_{\mathbb{R}}(K)$ die Menge der stetigen Funktionen $x : K \to \mathbb{R}$. Auf E sei
$$r(f_1, f_2) := \sup_{x \in K} |f_1(x) - f_2(x)|$$
der Supremumsabstand. Bekannt ist wieder, dass r vollständig ist. Außerdem lässt sich jedes $f \in E$ nach dem Weierstraß'schen Approximationssatz gleichmäßig durch Polynome approximieren. Sei E' die abzählbare Menge der Polynome mit rationalen Koeffizienten. Dann gilt auch, dass $\overline{E'} = E$. Damit ist $(E, \mathcal{O})$ separabel, also polnisch.

19.2 Kompakte Mengen

Topologische Räume können sehr groß sein. Man denke schon an den Raum $\mathbb{R}$, in dem es Folgen gibt, die divergieren. Als Mengen, bei denen Konvergenz zumindest entlang von Teilfolgen gilt, hatte wir schon in Definition 19.8 als folgenkompakte Mengen definiert. Dieses Konzept erweitern wir nun auf den allgemeinen Begriff der *kompakten Menge*. Insbesondere zeigen wir in Proposition 19.11, dass folgenkompakte Mengen meistens genau die kompakten Mengen sind, so dass wir nicht umdenken müssen.

Definition 19.8 (Relativ kompakt, kompakt, relativ folgenkompakt, total beschränkt). *Sei $(E, \mathcal{O})$ ein topologischer Raum und $K \subseteq E$.*

1. *Die Menge K heißt* kompakt*, falls jede offene Überdeckung eine endliche Teilüberdeckung besitzt. Das bedeutet: sind $O_i \in \mathcal{O}, i \in I$ und $A \subseteq \bigcup_{i \in I} O_i$, dann gibt es*[25] *$J \Subset I$ mit $A \subseteq \bigcup_{i \in J} O_i$.*

2. *Die Menge K heißt* relativ kompakt*, falls $\overline{K}$ kompakt ist.*

3. *Die Menge K heißt* relativ folgenkompakt*, falls gilt: für jede Folge $x_1, x_2, \ldots \in K$ gibt es eine konvergente Teilfolge, d.h. $x_{k_1}, x_{k_2}, \ldots \in K$ und $x \in E$ mit $x_{k_n} \xrightarrow{n \to \infty} x$ wie in* (19.5)*.*

4. *Sei r eine Metrik, die $\mathcal{O}$ erzeugt. Dann heißt $K \subseteq E$* total beschränkt*, wenn es für jedes $\varepsilon > 0$ ein $N \in \mathbb{N}$ und $x_1, \ldots, x_N \in K$ gibt, so dass $K \subseteq \bigcup_{n=1}^N B_\varepsilon(x_n)$.*

Beispiel 19.9 (Beschränkte und total beschränkte Mengen $A \subseteq \mathbb{R}^d$). Sei $A \subseteq \mathbb{R}^d$. Dann ist K genau dann total beschränkt, wenn A beschränkt ist (d.h. wenn es ein $K > 0$ gibt mit $A \subseteq B_K(0)$).
Denn: Nach Theorem 5.24 ist jedes $B_K(0)$ kompakt und hat demnach die Heine-Borel'sche Überdeckungseigenschaft. Sei also $\varepsilon > 0$ und $\{B_{\varepsilon/2}(x) : x \in K\}$ eine offene Überdeckung von K. Diese hat eine endliche Teilüberdeckung, also $x_1, ..., x_n \in K$ mit $K \subseteq \bigcup_{i=1}^n B_\varepsilon(x_i)$. Für jedes $x \in K$ mit $r(x, A) < \varepsilon/2$ gibt es ein $y \in A$, so dass $B_{\varepsilon/2}(x) \subseteq B_\varepsilon(y)$ ist. Mit Hilfe dieser y und der endlichen Überdeckung von K erreicht man nun leicht eine endliche Überdeckung

[25] Wir schreiben $J \Subset I$, falls $J \subseteq I$ und J endlich ist.

von A mit ε-Kugeln. Also ist A total beschränkt.
Sei umgekehrt A total beschränkt. Für $\varepsilon = 1$ gibt es also $x_1, ..., x_n \in A$ mit $A \subseteq \bigcup_{i=1}^n B_1(x_i)$. Damit ist der Durchmesser (der größte Abstand zweier Punkte) von A höchstens n, und damit ist A beschränkt.

Lemma 19.10 (Kompakte Mengen sind abgeschlossen). *Sei (E, r) ein metrischer Raum und $\mathcal{O}$ die von r erzeugte Topologie. Ist $K \subseteq E$ kompakt, dann ist K auch abgeschlossen.*

Beweis. Wir zeigen, dass K^c offen ist. Sei hierzu $x \in K^c$. Für alle $x' \in K$ wählen wir $\delta_{x'}$ und $\varepsilon_{x'}$, so dass $B_{\delta_{x'}}(x) \cap B_{\varepsilon_{x'}}(x') = \emptyset$. Dann ist offenbar $\bigcup_{x' \in K} B_{\varepsilon_{x'}}(x') \supseteq K$, also gibt es $J \Subset K$ mit $K \subseteq \bigcup_{x' \in J} B_{\varepsilon_{x'}}(x')$. Setze $\delta := \min_{x' \in J} \delta_{x'} > 0$. Dann ist $B_\delta(x) \cap K \subseteq B_\delta(x) \cap \bigcup_{x' \in J} B_{\varepsilon_{x'}}(x') = \emptyset$, also $B_\delta(x) \subseteq K^c$. Da $x \in K^c$ beliebig war, ist K^c offen, K also abgeschlossen. □

Folgender Satz über kompakte Mengen unterstreicht zum ersten Mal, warum polnischen Räumen eine besondere Bedeutung zukommt.

Proposition 19.11 (Charakterisierung relativ kompakter Mengen). *Sei (E, r) ein metrischer Raum, $\mathcal{O}$ die von r erzeugte Topologie und $K \subseteq E$. Betrachte folgende Aussagen:*

1. *K ist relativ kompakt.*
2. *Es gilt: sind $F_i \subseteq \overline{K}$ abgeschlossen, $i \in I$ und $\bigcap_{i \in I} F_i = \emptyset$, dann gibt es $J \Subset I$ mit $\bigcap_{i \in J} F_i = \emptyset$.*
3. *K ist relativ folgenkompakt.*
4. *K ist total beschränkt.*

Dann gilt

$$4. \Longleftarrow 1. \iff 2. \Longrightarrow 3.$$

Außerdem gilt auch 3. $\Longrightarrow$ 2. falls $(E, \mathcal{O})$ separabel ist und 4. $\Longrightarrow$ 3. falls (E, r) vollständig ist. Insbesondere sind alle vier Aussagen äquivalent, falls $(E, \mathcal{O})$ polnisch ist.

Wir wissen seit Theorem 5.24, dass eine Menge $K \subseteq \mathbb{R}^d$ genau dann (folgen-)kompakt ist, wenn K abgeschlossen und beschränkt ist. Diese Aussage verallgemeinern wir nun auf Teilmengen K eines beliebigen vollständigen und separablen metrischen Raumes.

Korollar 19.12. *Sei (E, r) ein metrischer Raum, $\mathcal{O}$ die von r erzeugte Topologie. Dann sind abgeschlossene Teilmengen kompakter Mengen wieder kompakt.*

Beweis. Sei $K \subseteq E$ kompakt und $A \subseteq K$ abgeschlossen. Eine abgeschlossene Menge ist genau dann kompakt, wenn sie relativ kompakt ist. Aus Proposition 19.11.2 liest man wegen der Relativkompaktheit von K ab, dass für $F_i \in \mathcal{O}^c, i \in I$ mit $F_i \subseteq A \subseteq K$ und $\bigcap_{i \in I} F_i = \emptyset$ ein $J \Subset I$ existiert mit $\bigcap_{i \in J} F_i = \emptyset$. Wieder mit Proposition 19.11.2 folgt daraus, dass A relativ kompakt, also kompakt ist. □

Beweis von Proposition 19.11. '1.⇒4.': Sei $\overline{K}$ kompakt und $\varepsilon > 0$. Offenbar ist $\bigcup_{x\in K} B_\varepsilon(x) \supseteq \overline{K}$ eine offene Überdeckung. Damit gibt es eine endliche Teilüberdeckung, d.h. es gibt $x_1, \ldots, x_N$ mit $\overline{K} \subseteq \bigcup_{n=1}^N B_\varepsilon(x_n)$. Da $\varepsilon > 0$ beliebig war, folgt die Behauptung.

'1.⇒2.': Sei nun $F_i, i \in I$ wie angegeben. Dann ist $\bigcup_{i\in I} F_i^c = \left(\bigcap_{i\in I} F_i\right)^c = E \supseteq \overline{K}$. Da $\overline{K}$ kompakt ist, gibt es $J \Subset I$ mit $\overline{K} \subseteq \bigcup_{i\in J} F_i^c$. Damit ist $\bigcap_{i\in J} F_i = \left(\bigcup_{i\in J} F_i^c\right) \subseteq \overline{K}^c$. Da aber $F_i \subseteq \overline{K}$ vorausgesetzt war, ist $\bigcap_{i\in J} F_i = \emptyset$.

'2.⇒1.': Sei $O_i \in \mathcal{O}, i \in I$ mit $\overline{K} \subseteq \bigcup_{i\in I} O_i$. Setze $F_i = O_i^c \cap \overline{K}$, dann ist $F_i^c \in \mathcal{O}$ und $\bigcap_{i\in I} F_i = \overline{K} \cap \left(\bigcup_{i\in I} O_i\right)^c = \emptyset$. Also gibt es $J \Subset I$ mit $\bigcap_{i\in J} F_i = \emptyset$. Damit ist $\overline{K}^c \cup \bigcup_{i\in J} O_i = \bigcup_{i\in J} F_i^c = E$, also $\bigcup_{i\in J} O_i \supseteq \overline{K}$. Mit anderen Worten ist $\overline{K}$ kompakt.

'2.⇒3.': Sei $x_1, x_2, \ldots \in K$. Wir setzen $F_n = \overline{\{x_n, x_{n+1}, \ldots\}} \subseteq \overline{K}$. Angenommen, es gibt keine konvergente Teilfolge von $x_1, x_2, \ldots$ Dann ist $\bigcap_{n=1}^\infty F_n = \emptyset$. Aus 2. folgt dann, dass es ein $N \in \mathbb{N}$ gibt mit $\emptyset = \bigcap_{n=1}^N F_n = F_N$. Dies ist ein Widerspruch, da F_N nach Konstruktion nicht leer ist; also gibt es eine konvergente Teilfolge.

'3.⇒1.' falls $(E, \mathcal{O})$ separabel ist: Sei E' abzählbar mit $\overline{E'} = E$ und $\mathcal{B} := \{B_{1/n}(x) : x \in E', n \in \mathbb{N}\}$. Damit ist $\mathcal{B}$ eine Basis von $\mathcal{O}$ und auch abzählbar. Wir schreiben $\mathcal{B} = \{B_1, B_2, \ldots\}$.

Angenommen $\overline{K}$ ist nicht kompakt. Das heißt, es gibt $A_i \in \mathcal{O}, i \in I$ mit $\overline{K} \subseteq \bigcup_{i\in I} A_i$ und es gibt keine endliche Teilüberdeckung. Wir setzen für $i \in I$

$$J_i = \{j \in \mathbb{N} : B_j \subseteq A_i\} \subseteq \mathbb{N}$$

sowie $J := \bigcup_{i\in I} J_i \subseteq \mathbb{N}$. Damit ist $A_i = \bigcup_{j\in J_i} B_j$, also

$$K \subseteq \bigcup_{i\in I} A_i = \bigcup_{i\in I} \bigcup_{j\in J_i} B_j = \bigcup_{j\in J} B_j.$$

Klar ist, dass nun $B_j \in \mathcal{O}, j \in J$ eine abzählbare Überdeckung von $\overline{K}$ darstellt. Da es keine endliche Teilüberdeckung für die $A_i, i \in I$ gibt, kann es auch keine endliche Teilüberdeckung für $B_j, j \in J$ geben. (Jedes B_j ist schließlich Teilmenge eines A_i's.) Für $n \in \mathbb{N}$ setzen wir $x_n \in \overline{K} \setminus \bigcup_{j\in J, j\leq n} B_j$. Nach Voraussetzung hat die Folge $x_1, x_2, \ldots \in K$ einen Häufungspunkt $x \in \overline{K}$. Da $\overline{K} \subseteq \bigcup_{j\in J} B_j$, gibt es $k \in J \subseteq \mathbb{N}$ mit $x \in B_k$. Damit liegen einerseits (da B_k offen ist) unendlich viele der x_n in B_k, andererseits ist $x_i \notin B_k$ für alle $i \geq k$ nach Konstruktion. Dies ist ein Widerspruch, also ist $\overline{K}$ kompakt.

'4.⇐'3. falls (E, r) vollständig ist: Sei $x_1, x_2, \ldots \in K$. Wir konstruieren eine Teilfolge, die Cauchy-Folge ist. Diese konvergiert, da (E, r) vollständig ist, und K ist als relativ folgenkompakt erkannt. Wähle eine Folge $\varepsilon_1, \varepsilon_2, \ldots > 0$ mit $\varepsilon_n \downarrow 0$. Da K total beschränkt ist, gibt es endlich viele ε_1-Bälle, die K überdecken. Mindestens einer dieser Bälle muss unendlich viele der x_n enthalten. Diese haben jeweils höchstens Abstand $2\varepsilon_1$. Wähle x_{k_1} als einer dieser unendlich vielen Punkte. Da dieser ε_1-Ball durch endlich viele ε_2-Bälle überdeckt wird, gibt es einen dieser ε_2-Bälle, der unendlich viele der x_n enthält. Diese haben jeweils höchstens Abstand $2\varepsilon_2$. Wähle $x_{k_2} \neq x_{k_1}$ als einen dieser unendlich vielen Punkten. Durch weiteres Vorgehen erhalten wir eine Folge $x_{k_1}, x_{k_2}, \ldots \in K$, so dass $r(x_{k_n}, x_{k_m}) \leq 2\varepsilon_{m\wedge n}$. Mit anderen Worten haben wir wie angekündigt eine Cauchy-Teilfolge in K gefunden. □

Korollar 19.13 (Analogon von Theorem 5.24 für polnische Räume). *Sei (E, r) vollständig und separabel und $K \subseteq E$. Dann ist K genau dann kompakt, wenn K abgeschlossen und total beschränkt ist.*

Beweis. Ist K kompakt, so ist K nach Lemma 19.10 auch abgeschlossen und auch relativ kompakt nach Definition. Nach Proposition 19.11 ist K auch total beschränkt, was die eine Richtung zeigt. Ist andersherum K total beschränkt, so folgt die Relativkompaktheit von K aus derselben Proposition. Klar ist, dass abgeschlossene, relativkompakte Mengen kompakt sind. □

Beispiel 19.14 (Kompakte metrische Räume sind polnisch). Sei (E, r) ein metrischer Raum und $\mathcal{O}$ die von r erzeugte Topologie. Falls E kompakt ist, dann ist $(E, \mathcal{O})$ polnisch.

Wir zeigen, dass r vollständig ist. Sei $x_1, x_2, \ldots \in E$ eine Cauchy-Folge. Da K relativ folgenkompakt nach Proposition 19.11 ist, gibt es eine konvergente Teilfolge $x_{k_1}, x_{k_2}, \ldots$ Sei $x \in E$ der Grenzwert der konvergenten Teilfolge. Sei $\varepsilon > 0$ und $N \in \mathbb{N}$, so dass $r(x_m, x_n) < \varepsilon/2$ für $m, n > N$ und $r(x_{k_n}, x) < \varepsilon/2$ für $k_n > N$. Dann gilt für $m > N$, dass $r(x_m, x) \leq r(x_m, x_{k_n}) + r(x_{k_n}, x) \leq \varepsilon$. Daraus folgt, dass $x_n \xrightarrow{n\to\infty} x$.

Für die Separabilität von $(E, \mathcal{O})$ sei $\varepsilon_1, \varepsilon_2, \ldots > 0$ mit $\varepsilon_n \downarrow 0$. Da K total beschränkt ist, gibt es für alle $n \in \mathbb{N}$ ein k_n und $x_{n1}, \ldots, x_{nk_n}$ mit $K \subseteq \bigcup_{k=1}^{k_n} B_{\varepsilon_n}(x_{nk_n})$. Sei $E' = \{x_{nk} : n \in \mathbb{N}, k = 1, \ldots, k_n\}$. Dann ist E' abzählbar und für jedes $x \in E$ und jedes $n \in \mathbb{N}$ gibt es ein $k(x, n) \in \{1, \ldots, k_n\}$ mit $r(x_{k(x,n)}, x) < \varepsilon_n$. Damit gilt $(x_{k(x,n)}, x) \xrightarrow{n\to\infty} x$. Also ist $\overline{E'} = E$.

20 Mengensysteme

Ein Ziel der Maßtheorie ist es, den *Inhalt* möglichst vieler Mengen zu messen. Deswegen legen wir in diesem Abschnitt fest, was maximale Mengensysteme sind, deren Elemente man einen Inhalt zuordnen kann. Dies führt auf den Begriff der σ-Algebra. Mit anderen Worten werden Elemente von σ-Algebren Ereignisse sein, deren Inhalte wir messen können. (Der wichtigste Fall wird hierbei sein, dass die σ-Algebra von einer Topologie abgeleitet ist, die ja auch schon ein Mengensystem darstellt.) Die anderen in diesem Abschnitt eingeführten Mengensysteme werden dazu dienen, geeignete σ-Algebren zu definieren.

20.1 Halbringe, Ringe und σ-Algebren

Die in diesem Abschnitt eingeführten Begriffe haben einen einfachen Zusammenhang. Ist nämlich $\mathcal{C} \subseteq 2^E$, so gilt

$$\mathcal{C} \text{ ist } \sigma\text{-Algebra} \quad \Longrightarrow \quad \mathcal{C} \text{ ist Ring} \quad \Longrightarrow \quad \mathcal{C} \text{ ist Halbring.}$$

Beziehungen zwischen den Mengensystemen sind in Tabelle 20.1 festgehalten.

Definition 20.1 (Halbring, Ring, σ-Algebra).

1. *Ein Mengensystem $\mathcal{H}$ heißt* schnittstabil, *falls mit $A, B \in \mathcal{H}$ auch $A \cap B \in \mathcal{H}$ gilt. Es heißt* vereinigungsstabil, *falls mit $A, B \in \mathcal{H}$ auch $A \cup B \in \mathcal{H}$ gilt. Es heißt komplementstabil, wenn mit $A \in \mathcal{H}$ auch $A^c \in \mathcal{H}$ gilt. Es heißt* differenzmengenstabil, *falls mit $A, B \in \mathcal{H}$ auch $B \setminus A \in \mathcal{H}$ gilt.*

	$\mathcal{C}$ Halbring	$\mathcal{C}$ Ring	$\mathcal{C}$ σ-Algebra
$\mathcal{C}$ schnittstabil	$\bullet$	$\circ$	$\circ$
$\mathcal{C}$ σ-schnittstabil			$\circ$
$\mathcal{C}$ vereinigungsstabil		$\bullet$	$\circ$
$\mathcal{C}$ σ-vereinigungsstabil			$\bullet$
$\mathcal{C}$ differenzmengenstabil		$\bullet$	$\circ$
$\mathcal{C}$ komplementstabil			$\bullet$
$B \setminus A = \biguplus_{i=1}^n C_i$	$\bullet$	$\circ$	$\circ$
$E \in \mathcal{C}$			$\circ$

Tabelle 20.1: Für $\mathcal{C} \subseteq 2^E$ wird hier die Beziehung zwischen Halbringen, Ringen und σ-Algebren dargestellt. Ein $\bullet$ bedeutet, dass in der Definition des Mengensystems (Spalte) die entsprechende Eigenschaft (Zeile) gefordert ist. Ein $\circ$ bedeutet, dass die entsprechende Eigenschaft aus den definierenden Eigenschaften des Mengensystems folgt.

2. *Ein nicht-leeres Mengensystem $\mathcal{H}$ ist ein* Halbring *(auf E), falls es (i) schnittstabil ist, und (ii) wenn es für $A, B \in \mathcal{H}$ Mengen $C_1, \dots, C_n \in \mathcal{H}$ gibt mit*[26] $B \setminus A = \biguplus_{i=1}^n C_i$.

3. *Ein Mengensystem $\mathcal{R}$ heißt* Ring *(auf E), falls (i) mit $A, B \in \mathcal{R}$ ist auch $A \cup B \in \mathcal{R}$ und (ii) mit $A, B \in \mathcal{R}$ ist auch $B \setminus A \in \mathcal{R}$.*

4. *Ein Mengensystem $\mathcal{F}$ heißt* σ-Algebra *(auf E), falls mit $A, A_1, A_2, \ldots \in \mathcal{F}$ auch $A^c, \bigcup_{n=1}^\infty A_n \in \mathcal{F}$ gilt. Dann heißt $(E, \mathcal{F})$ ein* Messraum.

Bemerkung 20.2 (Beziehungen zwischen den und weitere Mengensystemen).

1. Jeder Ring ist ein Halbring.
   Denn: Um dies einzusehen, ist nur nachzuprüfen, dass jeder Ring $\mathcal{R}$ schnittstabil ist. Dies folgt aus der Differenzmengenstabilität, da mit $A, B \in \mathcal{R}$ auch $A \cap B = A \setminus (A \setminus B) \in \mathcal{R}$.

2. Jede σ-Algebra ist schnittstabil und damit ein Ring.
   Denn: Das ist klar, da $A_1 \cap A_2 = (A_1^c \cup A_2^c)^c$ und $B \setminus A = B \cap A^c$.

3. Eine *Algebra* ist ein vereinigungs- und komplementstabiles Mengensystem. Ein schnitt- und vereinigungsstabiles Mengensystem heißt *Verband*. Unsere Darstellung kommt jedoch ohne diese Begriffe aus.

Beispiel 20.3 (Halbringe, σ-Algebren).

1. Sei $E = \mathbb{R}$. Dann bildet die Menge der halboffenen Intervalle

$$\mathcal{H} := \{(a, b] : a, b \in \mathbb{Q}, a \leq b\}$$

[26] Wir schreiben $A \uplus B$ für $A \cup B$, falls $A \cap B = \emptyset$.

einen Halbring.
Denn: Es gilt für $a_1 \leq b_1, a_1' \leq b_1'$, dass[27] $(a_1, b_1] \cap (a_1', b_1'] = (a_1 \vee a_1', b_1 \wedge b_1']$ und $(a_1, b_1] \setminus (a_1', b_1'] = (a_1, a_1' \wedge b_1] \uplus (b_1', b_1]$, wobei $(a, b] = \emptyset$, falls $a \geq b$.

2. Sei $E = \mathbb{R}^d$. Dann bildet die Menge der halboffenen Quader, also

$$\mathcal{H} := \Big\{ \prod_{i=1}^{d} (a_i, b_i] : a_i \leq b_i, i = 1, ..., d \Big\},$$

einen Halbring.
Denn: Wir verwenden aus Notationsgründen nur $d = 2$. Es ist nur zu zeigen, dass $((a_1, b_1] \times (a_2, b_2]) \setminus (a_1', b_1'] \times (a_2', b_2]$ für $a_1 \leq b_2, a_1' \leq b_1', a_2 \leq b_2, s_1' \leq b_2'$ als disjunkte Vereinigung von Quadern darstellbar ist. Sei hierzu $a_i \leq a_i' \leq b_i' \leq b_i', i = 1, 2$. Dann gilt

$$\begin{aligned} ((a_1, b_1] \times (a_2, b_2]) \setminus (a_1', b_1'] \times (a_2', b_2] =& (a_1, b_1] \times (a_2, a_2'] \\ &\uplus (a_1, a_1'] \times (a_2', b_2'] \\ &\uplus (b_1', b_1] \times (a_2', b_2] \\ &\uplus (a_1, b_1] \times (b_2', b_2]. \end{aligned}$$

Siehe auch Abbildung 20.1.

3. Die einfachsten Beispiele für σ-Algebren sind $\{\emptyset, E\}$ und 2^E. (Beides sind übrigens auch Topologien.) Für überabzählbares E ist ein nicht ganz so triviales Beispiel für eine σ-Algebra das Mengensystem

$$\{A \subseteq E : \ A \text{ abzählbar oder } A^c \text{ abzählbar}\}.$$

(Dieses Mengensystem ist kein Topologie.)

Ein weiteres Beispiel werden wir in Abschnitt 22.1 antreffen: Falls $\mathcal{F}'$ eine σ-Algebra auf E' ist und $f : E \to E'$. Dann ist

$$\sigma(f) := \{f^{-1}(A') : A' \in \mathcal{F}'\} \subseteq 2^E \tag{20.1}$$

eine σ-Algebra auf E. Ist nämlich $A', A_1', A_2', \ldots \in \sigma(f)$, so ist $(f^{-1}(A'))^c = f^{-1}((A')^c) \in \sigma(f)$ und $\bigcup_{n=1}^{\infty} f^{-1}(A_n') = f^{-1}\Big(\bigcup_{n=1}^{\infty} A_n'\Big) \in \sigma(f)$.

In der Praxis am weitaus häufigsten benötigt man Borel'sche σ-Algebren. Dies sind von Topologien abgeleitete σ-Algebren; siehe Definition 20.7.

20.2 Erzeuger und Erweiterungen

Auf der einen Seite ist es Ziel der Maßtheorie, Mengenfunktionen auf σ-Algebren zu definieren. Auf der anderen Seite kann man oftmals nur Halbringe konkret angeben, nicht jedoch σ-Algebren; siehe Beispiele 20.3.1 und 20.3.2. Allerdings gibt es für jeden Halbring $\mathcal{H}$ einen kleinsten Ring, der $\mathcal{H}$ enthält, $\mathcal{R}(\mathcal{H})$. Genauso gibt es eine kleinste σ-Algebra $\mathcal{F}$, die einen Halbring $\mathcal{H}$ (oder einen Ring $\mathcal{R}$) enthält, $\sigma(\mathcal{H})$. Damit erzeugt jeder Halbring $\mathcal{H}$ eine σ-Algebra.

[27] Wie üblich bezeichnet $x \wedge y := \min(x, y)$ und $x \vee y := \max(x, y)$

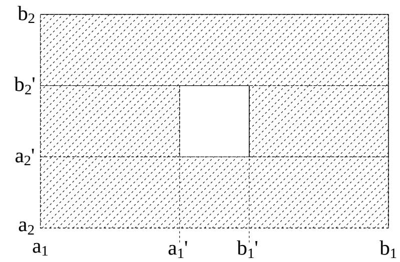

Abbildung 20.1: . Die Differenzmenge von Quadern ist als disjunkte Vereinigung von Quadern darstellbar.

Bemerkung 20.4 (Erzeugte Mengensysteme). Leicht überzeugt man sich, dass der Schnitt von Ringen (σ-Algebren) wieder ein Ring (σ-Algebra) ist. Insbesondere benötigen wir folgende Begriffe:
Sei $\mathcal{C} \subseteq 2^E$, dann ist

$$\mathcal{R}(\mathcal{C}) := \bigcap \Big\{ \mathcal{R} \supseteq \mathcal{C} : \mathcal{R} \text{ Ring} \Big\}$$

der *von $\mathcal{C}$ erzeugte Ring* und

$$\sigma(\mathcal{C}) := \bigcap \Big\{ F \supseteq \mathcal{C} : \mathcal{F}\ \sigma\text{-Algebra} \Big\}$$

die *von $\mathcal{C}$ erzeugte σ-Algebra*. Klar ist, dass $\mathcal{R}(\mathcal{R}(\mathcal{H})) = \mathcal{R}(\mathcal{H})$ und $\sigma(\sigma(\mathcal{H})) = \sigma(\mathcal{H})$.

Lemma 20.5 (Von Halbring erzeugter Ring). *Sei $\mathcal{H}$ ein Halbring. Dann ist*

$$\mathcal{R}(\mathcal{H}) = \Big\{ \biguplus_{k=1}^{n} A_k : A_1, \dots, A_n \in \mathcal{H} \text{ disjunkt}, n \in \mathbb{N} \Big\}$$

der von $\mathcal{H}$ erzeugte Ring.

Beispiel 20.6 (Von Intervallen erzeugter Ring). Sei $\mathcal{H}$ der von halboffenen Intervallen erzeugte Halbring aus Beispiel 20.3. Dann ist also

$$\begin{aligned} \mathcal{R}(\mathcal{H}) = \Big\{ \biguplus_{k=1}^{n} (a_k, b_k] &: a_1, \dots, a_n, b_1, \dots, b_n \in \mathbb{Q}, \\ & a_k < b_k, k = 1, \dots, n \text{ und } a_k < b_{k+1}, k = 1, \dots, n-1 \Big\} \end{aligned}$$

der von $\mathcal{H}$ erzeugte Ring.

Beweis von Lemma 20.5. Klar ist, dass $\mathcal{R}(\mathcal{H})$ schnittstabil ist. Um zu zeigen, dass $\mathcal{R}(\mathcal{H})$ ein Ring ist, zeigen wir zunächst die Differenzmengenstabilität. Sei $A_1, \dots, A_n \in \mathcal{H}$ und $B_1, \dots, B_m \in \mathcal{H}$, jeweils disjunkt. Dann gilt

$$\Big(\biguplus_{i=1}^{n} A_i \Big) \setminus \Big(\biguplus_{j=1}^{m} B_j \Big) = \biguplus_{i=1}^{n} \bigcap_{j=1}^{m} A_i \setminus B_j \in \mathcal{R}(\mathcal{H}).$$

Um die Vereinigungsstabilität von $\mathcal{R}(\mathcal{H})$ zu zeigen sei $A, B \in \mathcal{R}(\mathcal{H})$. Dann ist auch $A \cup B = (A \cap B) \uplus (A \setminus B) \uplus (B \setminus A) \in \mathcal{R}(\mathcal{H})$, da Schnitt- und Differenzmengenstabilität schon gezeigt sind.

Es gibt keinen kleineren Ring, der $\mathcal{H}$ enthält. Schließlich müsste dieser vereinigungsstabil sein, und ist damit eine Obermenge von $\mathcal{R}(\mathcal{H})$. $\square$

Definition 20.7 (Borel'sche σ-Algebra). *Sei $(E, \mathcal{O})$ ein topologischer Raum. Dann bezeichnet man mit $\mathcal{B}(E) := \sigma(\mathcal{O})$ die* Borel'sche σ-Algebra *auf E. Ist $E \subseteq \mathbb{R}^d$, so bezeichnen wir mit $\mathcal{B}(E)$ die Borel'sche σ-Algebra, die von der euklidischen Topologie auf $\mathbb{R}^d$ erzeugt wird. Ist $E \subseteq \overline{\mathbb{R}}$ oder $E = \mathbb{R}^d$, so ist $\mathcal{B}(E)$ die Borel'sche σ-Algebra, die von der Topologie aus Beispiel 19.3 oder aus Beispiel 20.3.3 erzeugt wird. Mengen in $\mathcal{B}(\overline{\mathbb{R}})$ und in $\mathcal{B}(\mathbb{R}^d)$ bezeichnet man auch als* (Borel-)messbare Mengen.

Lemma 20.8 (Abzählbare Basis und Borel'sche σ-Algebra). *Sei $(E, \mathcal{O})$ ein topologischer Raum mit abzählbarer Basis $\mathcal{C} \subseteq \mathcal{O}$. Dann gilt $\sigma(\mathcal{O}) = \sigma(\mathcal{C})$. Insbesondere ist E separabel.*

Beweis. Wir müssen nur zeigen, dass $\mathcal{O} \subseteq \sigma(\mathcal{C})$. Dies ist jedoch klar, da $A \in \mathcal{O}$ als abzählbare Vereinigung von Mengen aus $\mathcal{C}$ dargestellt werden kann. Siehe Lemma 19.6. $\square$

Lemma 20.9 (Borel'sche σ-Algebra wird von Quadern erzeugt). *Sei $E = \mathbb{R}^d$ und*

$$\begin{aligned}
\mathcal{C}_1 &= \Big\{ \prod_{i=1}^{d} (-\infty, b_i] : b_i \in \mathbb{Q} \Big\}, \\
\mathcal{C}_2 &= \Big\{ \prod_{i=1}^{d} (a_i, b_i] : a_i, b_i \in \mathbb{Q}, a_i \le b_i \Big\}, \\
\mathcal{C}_3 &= \Big\{ \prod_{i=1}^{d} (a_i, b_i) : a_i, b_i \in \mathbb{Q}, a_i \le b_i \Big\}, \\
\mathcal{C}_4 &= \Big\{ \prod_{i=1}^{d} [a_i, b_i] : a_i, b_i \in \mathbb{Q}, a_i \le b_i \Big\}.
\end{aligned}$$

Dann ist $\sigma(\mathcal{C}_j) = \mathcal{B}(\mathbb{R}^d)$, $j = 1, \dots, 4$.

Beweis. Wir zeigen das Lemma im Fall $d = 1$. Den Fall $d > 1$ zeigt man analog (mit einer aufwändigeren Notation).
Das Mengensystem $\mathcal{C}_3$ ist eine abzählbare Basis der euklidischen Topologie auf $\mathbb{R}$. Für diesen Fall folgt die Aussage aus Lemma 20.8.

Wir zeigen die Aussage nur für $\mathcal{C}_1$ und $\mathcal{C}_2$, die für $\mathcal{C}_4$ folgt analog. Zunächst ist $\mathcal{C}_2 := \{A \setminus B : A, B \in \mathcal{C}_1\} = \{(a, b] : a, b \in \mathbb{Q}, a \le b\} \subseteq \sigma(\mathcal{C}_2)$ der von $\mathcal{C}_1$ erzeugte Halbring aus Beispiel 20.3. Damit ist $\sigma(\mathcal{C}_1) = \sigma(\mathcal{C}_2)$ und es genügt zu zeigen, dass $\sigma(\mathcal{C}_2) = \mathcal{B}(\mathbb{R})$. Sei hierzu $\mathcal{O}$ wie in Definition 19.1.8 mit $E = \mathbb{R}$. Wir zeigen (i) aus $A \in \mathcal{O}$ folgt, dass $A \in \sigma(\mathcal{C}_2)$, und (ii) aus $A \in \mathcal{C}_2$ folgt, dass $A \in \sigma(\mathcal{O})$. Daraus folgt dann $\mathcal{O} \subseteq \sigma(\mathcal{C}_2) \subseteq \sigma(\mathcal{O})$, also $\sigma(\mathcal{O}) = \sigma(\mathcal{C}_2)$.
Für (i) sei also $A \in \mathcal{O}$. Wir behaupten

$$A = \bigcup \{(a, b] : (a, b] \subseteq A, a, b \in \mathbb{Q}\}, \tag{20.2}$$

und bemerken, dass die rechte Seite ein Element von $\sigma(\mathcal{C}_2)$ ist. In (20.2) ist '$\supseteq$' klar. Um '$\subseteq$' einzusehen, wählen wir $x \in A$. Dann gibt es nach Definition von $\mathcal{O}$ ein $\varepsilon > 0$, so dass $B_\varepsilon(x) \subseteq A$. Allerdings gibt es auch $a, b \in \mathbb{Q}$ mit $a \leq b$ und $x \in (a, b] \subseteq B_\varepsilon(x)$. Damit ist '$\subseteq$' gezeigt und (i) folgt.
Für (ii) gehen wir ähnlich vor; sei $A \in \mathcal{C}_2$. Dann ist offenbar

$$A = \bigcap_{n=1}^{\infty} \big(a, b + \tfrac{1}{n}\big).$$

Da $\big(a, b + \frac{1}{n}\big) \in \mathcal{O}$, ist also $A \in \sigma(\mathcal{O})$. □

Beispiel 20.10 (Borel-messbare Mengen). Natürlich sind alle abzählbaren Durchschnitte und Vereinigungen von Intervallen nach Lemma 20.9 in $\mathcal{B}(\mathbb{R})$. Sei beispielsweise

$$\begin{aligned}
&A_1 = [0, \tfrac{1}{3}] \cup [\tfrac{2}{3}, 1],\\
&A_2 = [0, \tfrac{1}{9}] \cup [\tfrac{2}{9}, \tfrac{3}{9}] \cup [\tfrac{6}{9}, \tfrac{7}{9}] \cup [\tfrac{8}{9}, 1],\\
&A_3 = [0, \tfrac{1}{27}] \cup [\tfrac{2}{27}, \tfrac{3}{27}] \cup [\tfrac{6}{27}, \tfrac{7}{27}] \cup [\tfrac{8}{27}, \tfrac{9}{27}] \cup [\tfrac{18}{27}, \tfrac{19}{27}] \cup [\tfrac{20}{27}, \tfrac{21}{27}] \cup [\tfrac{24}{27}, \tfrac{25}{27}] \cup [\tfrac{26}{27}, 1],\\
&\ldots
\end{aligned}$$

dann bezeichnet $A = \bigcap_{n=1}^{\infty}$ das Cantor'sche Diskontinuum. Diese Menge ist als abzählbarer Schnitt von endlichen Vereinigungen von Intervallen messbar (und übrigens auch abgeschlossen). Wir werden in Beispiel 21.27 ein Beispiel für eine nicht Borel-messbare Menge kennen lernen.

20.3 Dynkin-Systeme

Oftmals muss man in der Maßtheorie zeigen, dass ein bestimmtes Mengensystem $\mathcal{F}$ eine σ-Algebra ist und einen Halbring $\mathcal{H}$ beinhaltet. Die in diesem Abschnitt behandelten Dynkin-Systeme sind dabei sehr hilfreich. Es genügt nämlich wegen Theorem 20.13 zu zeigen, dass $\mathcal{F}$ ein schnittstabiles Dynkin-System mit $\mathcal{H} \subseteq \mathcal{F}$ ist. Dies ist oftmals einfacher, als direkt zu zeigen, dass $\mathcal{F}$ eine σ-Algebra ist.

Definition 20.11 (Dynkin-System). *Ein Mengensystem $\mathcal{D}$ heißt* Dynkin-System *(auf E), falls (i) $E \in \mathcal{D}$, (ii) mit $A, B \in \mathcal{D}$ und $A \subseteq B$ ist auch $B \setminus A \in \mathcal{D}$ und (iii) mit $A_1, A_2, \ldots \in \mathcal{D}$ mit $A_1 \subseteq A_2 \subseteq A_3 \subseteq \ldots$ ist auch $\bigcup_{n=1}^{\infty} A_n \in \mathcal{D}$.*

Beispiel 20.12 (σ-Algebren sind Dynkin-Systeme). Jede σ-Algebra ist ein Dynkin-System.
Denn: Sei $\mathcal{F}$ eine σ-Algebra. Dann ist mit $A, B \in \mathcal{F}$ auch $A^c \in \mathcal{F}$ und damit $E = A \cup A^c \in \mathcal{F}$ sowie $B \setminus A = B \cap A^c \in \mathcal{F}$.

Theorem 20.13 (Schnittstabile Dynkin-Systeme). *Sei $\mathcal{D}$ ein Dynkin-System und $\mathcal{C} \subseteq \mathcal{D}$ schnittstabil. Dann ist $\sigma(\mathcal{C}) \subseteq \mathcal{D}$. Insbesondere ist jedes schnittstabile Dynkin-System eine σ-Algebra.*

Beweis. Genau wie in Lemma 20.4 definiert man das von $\mathcal{C}$ erzeugte Dynkin-System $\lambda(\mathcal{C})$. Da offenbar der Schnitt von Dynkin-Systemen wieder ein Dynkin-System ist, gilt $\lambda(\mathcal{C}) \subseteq \mathcal{D}$. Wir werden zeigen, dass $\lambda(\mathcal{C})$ eine σ-Algebra ist, denn dann ist $\sigma(\mathcal{C}) \subseteq \sigma(\lambda(\mathcal{C})) = \lambda(\mathcal{C}) \subseteq \mathcal{D}$.

Wir zeigen hierzu, dass $\lambda(\mathcal{C})$ schnittstabil ist. Dann ist nämlich für $A_1, A_2, \ldots \in \lambda(\mathcal{C})$

$$\bigcup_{i=1}^n A_i = \Big(\bigcap_{i=1}^n A_i^c\Big)^c \in \lambda(\mathcal{C})$$

und, da $\lambda(\mathcal{C})$ ein Dynkin-System ist $\bigcup_{n=1}^\infty A_n = \bigcup_{n=1}^\infty \bigcup_{i=1}^n A_i \in \lambda(\mathcal{C})$.

Wir müssen also zeigen, dass mit $A, B \in \lambda(\mathcal{C})$ auch $A \cap B \in \lambda(\mathcal{C})$ gilt. Falls $A, B \in \mathcal{C}$, so ist dies wegen der Schnittstabilität von $\mathcal{C}$ klar. Für $B \in \mathcal{C}$ setzen wir

$$\mathcal{D}_B := \{A \subseteq E : A \cap B \in \lambda(\mathcal{C})\}.$$

Dann ist $\mathcal{D}_B$ ein Dynkin-System, da (i) $E \in \mathcal{D}_B$, (ii) für $A, C \subseteq \mathcal{D}_B$ ist $A \cap B, C \cap B \in \lambda(\mathcal{C})$ also mit $A \subseteq C$ auch $A \cap B \subseteq C \cap B$, also $(C \setminus A) \cap B = (C \cap B) \setminus (A \cap B) \in \lambda(\mathcal{C})$ und (iii) für $A_1, A_2, \ldots \in \mathcal{D}_B$ ist $A_1 \cap B, A_2 \cap B, \ldots \in \lambda(\mathcal{C})$ und falls $A_1 \subseteq A_2 \subseteq \cdots$ ist $A_1 \cap B \subseteq A_2 \cap B \subseteq \cdots$, woraus $\Big(\bigcup_{n=1}^\infty A_n\Big) \cap B = \Big(\bigcup_{n=1}^\infty A_n \cap B\Big) \in \lambda(\mathcal{C})$ folgt.

Da $\mathcal{C} \subseteq \mathcal{D}_B$ und $\mathcal{D}_B$ ein Dynkin-System ist, ist auch $\lambda(\mathcal{C}) \subseteq \mathcal{D}_B$. Dies bedeutet, dass für $A \in \lambda(\mathcal{C})$ und $B \in \mathcal{C}$ auch $A \cap B \in \lambda(\mathcal{C})$. Wir setzen nun für ein $A \in \lambda(\mathcal{C})$

$$\mathcal{B}_A := \{B \subseteq E : A \cap B \in \lambda(\mathcal{C})\}.$$

Wie oben zeigt man, dass $\mathcal{B}_A$ ein Dynkin-System mit $\mathcal{C} \subseteq \mathcal{B}_A$ ist. Damit ist wieder $\lambda(\mathcal{C}) \subseteq \mathcal{B}_A$. Insbesondere ist mit $A, B \in \lambda(\mathcal{C})$ auch $A \cap B \in \lambda(\mathcal{C})$ und die erste Behauptung ist gezeigt. Die zweite Behauptung folgt, wenn man $\mathcal{C} := \mathcal{D}$ setzt. □

20.4 Kompakte Systeme

Nachdem wir im Abschnitt 19.2 kompakte Teilmengen von E kennen gelernt haben, stellen wir nun einen wichtigen Zusammenhang zwischen Systemen kompakter Mengen und der Maßtheorie vor. Solche kompakten Systeme spielen eine wichtige Rolle beim Beweis von Theorem 21.9. Hier wird gezeigt, dass aus der Additivität einer Mengenfunktion und einer Approximationseigenschaft bzgl. eines kompakten Systems die σ-Additivität der Mengenfunktion folgt.

Definition 20.14 (Kompaktes System). *Ein schnittstabiles Mengensystem $\mathcal{K}$ heißt* kompaktes System *(auf E), falls aus $\bigcap_{n=1}^\infty K_n = \emptyset$ mit $K_1, K_2, \ldots \in \mathcal{K}$ folgt, dass es ein $N \in \mathbb{N}$ gibt mit $\bigcap_{n=1}^N K_n = \emptyset$.*

Beispiel 20.15 (Kompakte Mengen). Die Menge alles kompakten Teilmengen von E bilden ein kompaktes System.
Etwas allgemeiner gilt auch: Sei (E, r) ein metrischer Raum und $\mathcal{O}$ die von r erzeugte Topologie. Dann ist jedes schnittstabile $\mathcal{K} \subseteq \{K \subseteq E : K \text{ kompakt}\}$ ein kompaktes System.
Denn: sei $\bigcap_{n=1}^\infty K_n = \emptyset$. Dann ist sowohl K_1, als auch $L_n := K_1 \cap K_n \subseteq K_1$ für $n = 1, 2, \ldots$ nach Lemma 19.10 abgeschlossen und wegen der Kompaktheit von K_1 gibt es ein N mit $\bigcap_{n=1}^N K_n = \emptyset$ nach Proposition 19.11.

Lemma 20.16 (Erweiterung kompakter Systeme). *Sei $\mathcal{K}$ ein kompaktes System. Dann ist auch*

$$\mathcal{K}_\cup := \Big\{\bigcup_{i=1}^n K_i : K_1, \ldots, K_n \in \mathcal{K}, n \in \mathbb{N}\Big\}$$

ein kompaktes System.

Beweis. Klar ist, dass $\mathcal{K}_\cup$ schnittstabil ist. Seien $L_1, L_2, \ldots \in \mathcal{K}_\cup$ mit $\bigcap_{n=1}^N L_n \neq \emptyset$ für alle $N \in \mathbb{N}$. Zu zeigen ist, dass dann auch $\bigcap_{n=1}^\infty L_n \neq \emptyset$. Dazu zeigen wir mit Induktion über N folgende Aussage:

> Es gibt für jedes $N \in \mathbb{N}$ Mengen $K_1, \ldots, K_N \in \mathcal{K}$ mit $K_n \subseteq L_n$, $n = 1, \ldots, N$, so dass für alle $k \in \mathbb{N}_0$ gilt dass $K_1 \cap \cdots \cap K_N \cap L_{N+1} \cap \cdots \cap L_{N+k} \neq \emptyset$.

Sei $N = 1$. Da $L_1 = \bigcup_{j=1}^{m_1} K'_j$ für $K'_1, \ldots, k'_{m_1} \in \mathcal{K}$ und $\bigcap_{n=1}^N L_n = \bigcup_{j=1}^{m_1} K'_j \cap \bigcap_{n=1}^N L_n \neq \emptyset$ für jedes $N \in \mathbb{N}$, gibt es ein $j = 1, \ldots, m_1$, so dass $K'_j \cap \bigcap_{n=1}^N L_n \neq \emptyset$ für jedes $N \in \mathbb{N}$. Setze $K_1 := K'_j$, und die Behauptung ist für $N = 1$ gezeigt.

Gelte die Behauptung für $N - 1$. Wieder ist $L_N = \bigcup_{j=1}^{m_N} K''_j$ für $K''_1, \ldots, K''_{m_N} \in \mathcal{K}$. Damit gilt nach der Induktionsvoraussetzung für alle $k \in \mathbb{N}$, dass

$$
\begin{aligned}
&K_1 \cap \cdots \cap K_{N-1} \cap \Big(\bigcup_{j=1}^{m_N} K''_j \Big) \cap L_{N+1} \cap \cdots \cap L_{N+k} \\
&\quad = \bigcup_{j=1}^{m_N} K_1 \cap \cdots \cap K_{N-1} \cap K''_j \cap L_{N+1} \cap \cdots \cap L_{N+k} \neq \emptyset.
\end{aligned}
$$

Damit gibt es ein j, so dass $K_1 \cap \cdots \cap K_{N-1} \cap K''_j \cap L_{N+1} \cap \cdots \cap L_{N+k} \neq \emptyset$ für alle $k \in \mathbb{N}$. Setze $K_N := K''_j$, was den Induktionsbeweis abschließt.

Setzt man $k = 0$ in obiger Behauptung, sieht man, dass es $K_1, K_2, \ldots \in \mathcal{K}$ und $K_n \subseteq L_n$, $n \in \mathbb{N}$ gibt mit $\bigcap_{n=1}^N K_n \neq \emptyset$ für alle $N \in \mathbb{N}$. Da $\mathcal{K}$ ein kompaktes System ist, gilt insbesondere

$$
\emptyset \neq \bigcup_{n=1}^\infty K_n \subseteq \bigcup_{n=1}^\infty L_n
$$

und die Behauptung ist gezeigt. □

21 Maße

Will man Mengen *messen*, also etwa jeder Menge ihren Inhalt zuordnen, scheinen einige Forderungen natürlich. Etwa sollte die leere Menge den Inhalt 0 zugeordnet werden, oder die Massen sollten sich abzählbar additiv verhalten, siehe (21.1). Wie sich herausstellt, sind Maße auf σ-Algebren zu definieren, damit die Forderung der abzählbaren Additivität erfüllt werden kann. In diesem Abschnitt geben wir die wichtigsten Schritte an, solche Maße zu konstruieren. Als wichtigstes Beispiel eines Maßes dient dabei das Lebesgue-Maß.

21.1 Mengenfunktionen

Wir werden nun Funktionen $\mu : \mathcal{C} \to \overline{\mathbb{R}}_+$ betrachten, wenn $\mathcal{C}$ ein Halbring, Ring oder eine σ-Algebra ist.

Definition 21.1 (Maß und äußeres Maß). *Sei $\mathcal{F} \subseteq 2^E$ und $\mu : \mathcal{F} \to \overline{\mathbb{R}}_+$.*

1. *Die Abbildung μ heißt* endlich additiv, *falls für disjunkte $A_1, \ldots, A_n \in \mathcal{F}$ mit $\biguplus_{k=1}^n A_k \in \mathcal{F}$ gilt, dass*

$$
\mu\Big(\biguplus_{k=1}^n A_k \Big) = \sum_{k=1}^n \mu(A_k). \tag{21.1}
$$

Sie heißt subadditiv, *falls für (beliebige)* $A_1, \dots, A_n \in \mathcal{F}$ *mit* $\bigcup_{k=1}^n A_k \in \mathcal{F}$ *gilt, dass*

$$\mu\Big(\bigcup_{k=1}^n A_k\Big) \le \sum_{k=1}^n \mu(A_k). \tag{21.2}$$

2. *Eine Abbildung* $\mu : \mathcal{F} \to \mathbb{R}_+$ *heißt* σ-additiv, *falls* (21.1) *auch für* $n = \infty$ *gilt. Sie heißt* σ-sub-additiv, *wenn* (21.2) *auch für* $n = \infty$ *gilt. Sie heißt* monoton, *wenn aus* $A, B \in \mathcal{F}$ *mit* $A \subseteq B$ *folgt, dass* $\mu(A) \le \mu(B)$.

3. *Falls es eine Folge* $E_1, E_2, \ldots \in \mathcal{F}$ *mit* $\bigcup_{n=1}^\infty E_n = E$ *und* $\mu(E_n) < \infty$ *für alle* $n = 1, 2, \ldots$ *gilt, so heißt* μ σ-endlich.

4. *Sei* $\mathcal{F}$ *eine* σ*-Algebra und* $\mu : \mathcal{F} \to \mathbb{R}_+$. *Ist* μ *auch* σ*-additiv, so heißt* μ *ein* Maß *(auf* $\mathcal{F}$*) und* $(E, \mathcal{F}, \mu)$ *ist ein* Maßraum. *Gilt* $\mu(E) < \infty$, *so heißt* μ endliches Maß *und falls* $\mu(E) = 1$, *so heißt* μ *ein* Wahrscheinlichkeitsmaß *oder eine* Wahrscheinlichkeitsverteilung *oder auch einfach eine* Verteilung. *Außerdem heißt* $(E, \mathcal{F}, \mu)$ *dann ein* Wahrscheinlichkeitsraum.

5. *Sei* $(E, \mathcal{O})$ *ein topologischer Raum und* μ *ein Maß auf* $\mathcal{B}(\mathcal{O})$. *Dann heißt die kleinste abgeschlossene Menge* F *(d.h.* $F^c \in \mathcal{O}$*) mit* $\mu(F^c) = 0$ *der* Träger von μ.

6. *Eine* σ*-subadditiv, monotone Abbildung* $\mu^* : 2^E \to \mathbb{R}_+$ *heißt* äußeres Maß, *falls* $\mu^*(\emptyset) = 0$. *Eine Menge* $A \subseteq E$ *heißt* μ^*-messbar, *falls*

$$\mu(B) = \mu(B \cap A) + \mu(B \cap A^c) \tag{21.3}$$

für alle $B \subseteq E$ *gilt.*

7. *Sei* $\mathcal{F}$ *schnittstabil und* $\mathcal{K} \subseteq \mathcal{F}$ *ein kompaktes System. Dann heißt* μ *von innen* $\mathcal{K}-$*regulär, falls für alle* $A \in \mathcal{K}$

$$\mu(A) = \sup_{\mathcal{K} \ni K \subseteq A} \mu(K).$$

Beispiel 21.2 (Beispiele von Mengenfunktionen). 1. Sehr einfache Beispiele für Mengenfunktionen sind die Dirac-Maße. Ist $x \in E$, so ist

$$\delta_x : \begin{cases} 2^E & \to \{0, 1\} \\ A & \mapsto 1_{\{x \in A\}} \end{cases}$$

ein (Wahrscheinlichkeits-)Maß.

Sei nun $\mu_i = \delta_{x_i}$, $i \in I$. Dann heißt das Maß $\sum_{i \in I} \delta_{x_i}$ ein Zählmaß. Außerdem ist dann für $a_i \in \mathbb{R}_+, i \in I$ auch $\sum_{i \in I} a_i \mu_i$ ein Maß.

2. Sei $E = \mathbb{R}^d$. Wir werden uns oft mit Mengenfunktionen auf

$$\mathcal{H} = \Big\{ \prod_{i=1}^d (a_i, b_i] : a_i, b_i \in \mathbb{Q}, a_i \le b_i \Big\}$$

aus 20.3 beschäftigen. Etwa definiert

$$\mu\Big(\prod_{i=1}^d (a_i, b_i]\Big) = \prod_{i=1}^d (b_i - a_i)$$

offenbar eine additive, σ-endliche Mengenfunktion.
Denn: Wir zeigen die Additivität im Fall $n = 2$. Der allgemeine Fall folgt dann mit Induktion. Seien

$$A_j = \prod_{i=1}^d (a_i^j, b_i^j] \in \mathcal{H}, j = 1, 2$$

disjunkt mit $A = A_i \uplus A_2 \in \mathcal{H}$; wir dürfen also $A = \prod_{i=1}^d (a_i, b_i]$ schreiben. Man überlegt sich, dass $a_i = a_i^1 \wedge a_i^2$ und $b_i = b_i^1 \vee b_i^2$ gelten muss. Außerdem muss $a_i^1 = a_i^2$ und $b_i^1 = b_i^2$ für alle bis auf ein i gelten, das wir j nennen, sowie $a_j^1 \vee a_j^2 = b_j^1 \wedge b_j^2$. Hieraus folgt

$$\begin{aligned}
\mu(A) &= \prod_{i=1}^d (b_i - a_i) = (b_j - a_j) \cdot \prod_{\substack{i=1\\ i\neq j}}^d (b_i - a_i) \\
&= (b_j^1 \vee b_j^2 - a_j^1 \vee a_j^2 + b_j^1 \wedge b_j^2 - a_j^1 \wedge a_j^2) \cdot \prod_{\substack{i=1\\ i\neq j}}^d (b_i - a_i) \\
&= (b_j^1 - a_j^1 + b_j^2 - a_j^2) \cdot \prod_{\substack{i=1\\ i\neq j}}^d (b_i - a_i) \\
&= (b_j^1 - a_j^1) \cdot \prod_{\substack{i=1\\ i\neq j}}^d (b_i^1 - a_i^1) + (b_j^2 - a_j^2) \cdot \prod_{\substack{i=1\\ i\neq j}}^d (b_i^2 - a_i^2) = \mu(A_1) + \mu(A_2).
\end{aligned}$$

Weiter ist $\mu((-n, n]^d) < \infty$ und $\bigcup_{n=1}^\infty (-n, n]^d = \mathbb{R}^d$. Damit ist μ also σ-endlich.

Wir werden die Funktion μ in eindeutiger Weise auf die Borel'sche σ-Algebra $\mathcal{B}(\mathbb{R}^d) = \sigma(\mathcal{H})$ (siehe Lemma 20.9) erweitern und so das Lebesgue-Maß konstruieren, siehe Proposition 21.17.

Bemerkung 21.3 (Inhalte und Prämaße). Endlich additive Mengenfunktionen heißen oft *Inhalt*, σ-additive Mengenfunktionen, die nicht auf σ-Algebren definiert sind, heißen oft Prämaße. Die auf einer Borel'schen σ-Algebra definierten, bezüglich der kompakten Mengen von innen regulären Maße heißen Radon-Maße. Wir werden diese Begriffe nicht verwenden.

Lemma 21.4 (Mengenfunktionen auf Halbringen). *Sei $\mathcal{H}$ ein Halbring und $\mu : \mathcal{H} \to \overline{\mathbb{R}}_+$ endlich additiv. Dann ist μ monoton und subadditiv. Außerdem ist μ genau dann σ-additiv wenn es σ-subadditiv ist.*

Beweis. Wir beginnen mit einer Argumentation, die wir häufiger brauchen werden. Für $A_1, A_2, \ldots \in \mathcal{H}$ und $B_n = A_n \setminus \Big(\bigcup_{i=1}^{n-1} A_i\Big)$ und $n \geq 2$ gibt es $D_1, \ldots, D_k \in \mathcal{H}$ disjunkt mit $B_n = \biguplus_{i=1}^k D_i$. Die Behauptung ist klar für $n = 2$, da $\mathcal{H}$ ein Halbring ist. Gilt die Behauptung für ein n, so ist $B_{n+1} = A_{n+1} \setminus \left(\bigcup_{i=1}^n A_i\right) = \left(A_{n+1} \setminus \left(\bigcup_{i=2}^n A_i\right)\right) \setminus A_1$. Nach Voraussetzung gibt es $D_1, \ldots, D_k$ mit $A_{n+1} \setminus \left(\bigcup_{i=2}^n A_i\right) = \biguplus_{i=1}^k D_i$. Damit ist $B_{n+1} = \left(\biguplus_{i=1}^k D_i\right) \cap A_1^c = \biguplus_{i=1}^k (D_i \setminus A_1)$. Für alle $i = 1, \ldots, k$ gibt es Mengen $E_{i1}, \ldots, E_{k_i} \in \mathcal{H}$ mit $D_i \setminus A_1 = \biguplus_{j=1}^{k_i} E_{ij}$. Damit gilt $B_{n+1} = \biguplus_{i=1}^k \biguplus_{j=1}^{k_i} E_{ij}$ und die Behauptung ist gezeigt.

Wir zeigen nun, dass μ monoton ist. Sei $A, B \in \mathcal{H}$ mit $A \subseteq B$. Dann gibt es disjunkte $D_1, \ldots, D_k \in \mathcal{H}$ mit $B \setminus A = \biguplus_{i=1}^k D_i$, also $B = A \uplus \biguplus_{i=1}^k D_i$ und damit wegen der Additivität von μ auch $\mu(B) = \mu(A) + \sum_{i=1}^k \mu(D_i) \geq \mu(A)$. Genauso zeigt man, dass aus $E_1, \ldots, E_k \in \mathcal{H}$ disjunkt mit $\biguplus_{i=1}^k E_i \subseteq A \in \mathcal{H}$ folgt, dass $\sum_{i=1}^k \mu(E_i) \leq \mu(A)$.

Für die Subadditivität seien $A_1, \ldots, A_n \in \mathcal{H}$ mit $\bigcup_{i=1}^n A_i \in \mathcal{H}$. Wir setzen $B_1 = A_1, B_2 = A_2 \setminus A_1, B_3 = A_3 \setminus (A_1 \cup A_2), \ldots$ Nach oben Gesagtem gibt es für jedes $i = 1, \ldots, k$ ein $k_i \in \mathbb{N}$ und Mengen $D_{ij} \in \mathcal{H}$, $j = 1, \ldots, k_i$ mit $B_i = \biguplus_{j=1}^{k_i} D_{ij}$. Damit gilt $\bigcup_{i=1}^n A_i = \biguplus_{i=1}^n B_i = \biguplus_{i=1}^n \biguplus_{j=1}^{k_i} D_{ij}$. Weiter ist $A_i = B_i \uplus \bigcup_{j=1}^{i-1} A_j$, also

$$\mu\Big(\bigcup_{i=1}^n A_i\Big) = \sum_{i=1}^n \sum_{j=1}^{k_i} \mu(D_{ij}) \leq \sum_{i=1}^n \mu(A_i).$$

Wir zeigen nun μ ist σ-additiv $\iff$ μ ist σ-sub-additiv.
'$\Rightarrow$' folgt genauso wie die Subadditivität von μ. Für Für '$\Leftarrow$' sei $A_1, A_2, \ldots \in \mathcal{H}$ disjunkt mit $A = \bigcup_{n=1}^\infty A_n \in \mathcal{H}$. Da μ monoton ist, gilt

$$\sum_{n=1}^\infty \mu(A_n) = \lim_{N\to\infty} \sum_{n=1}^N \mu(A_n) = \lim_{N\to\infty} \mu\Big(\biguplus_{n=1}^N A_n\Big) \leq \mu(A) \leq \sum_{n=1}^\infty \mu(A_n)$$

wegen der σ-Subadditivität. □

Lemma 21.5 (Fortsetzung von Mengenfunktionen auf Halbringen). *Sei $\mathcal{H}$ ein Halbring, $\mathcal{R}$ der von $\mathcal{H}$ erzeugte Ring aus Lemma 20.5 und μ eine endlich additive Funktion auf $\mathcal{H}$. Definiere die Mengenfunktion $\widetilde{\mu}$ auf $\mathcal{R}$ durch*

$$\widetilde{\mu}\Big(\biguplus_{i=1}^n A_i\Big) := \sum_{i=1}^n \mu(A_i)$$

für $A_1, \ldots, A_n \in \mathcal{H}$ disjunkt. Dann ist $\widetilde{\mu}$ die einzige additive Fortsetzung von μ auf $\mathcal{R}$, die auf $\mathcal{H}$ mit μ übereinstimmt. Außerdem ist $\widetilde{\mu}$ genau dann σ-additiv, wenn μ σ-additiv ist.

Beweis. Es ist nur zu zeigen, dass $\widetilde{\mu}$ wohldefiniert ist. Sei hierzu $A_1, \ldots, A_m, B_1, \ldots, B_n \in \mathcal{H}$ mit $\biguplus_{i=1}^m A_i = \biguplus_{j=1}^n B_j$. Da

$$A_i = \biguplus_{j=1}^n A_i \cap B_j, \qquad B_j = \biguplus_{i=1}^m A_i \cap B_j,$$

gilt wegen der endlichen Additivität von $\widetilde{\mu}$

$$\sum_{i=1}^m \mu(A_i) = \sum_{i=1}^m \sum_{j=1}^n \mu(A_i \cap B_j) = \sum_{j=1}^n \sum_{i=1}^m \mu(A_i \cap B_j) = \sum_{j=1}^n \mu(B_j).$$ □

Proposition 21.6 (Einschluss-/Ausschlussformel). *Sei μ eine additive Mengenfunktion auf einem Ring $\mathcal{R}$ und I endlich. Dann gilt für $A_i \in \mathcal{R}$, $i \in I$, dass*

$$\mu\Big(\bigcup_{i\in I} A_i\Big) = \sum_{J\subseteq I}(-1)^{|J|+1}\mu\Big(\bigcap_{j\in J} A_j\Big)$$

Für $I = \{1,2\}$ gilt also

$$\mu(A_1 \cup A_2) = \mu(A_1) + \mu(A_2) - \mu(A_1 \cap A_2)$$

und für $I = \{1,2,3\}$

$$\begin{aligned}\mu(A_1 \cup A_2 \cup A_3) = \mu(A_1) + \mu(A_2) + \mu(A_3)&\\ - \mu(A_1 \cap A_2) - \mu(A_1 \cap A_3) - \mu(A_2 \cap A_3) + \mu(A_1 \cap A_2 \cap A_3)&.\end{aligned}$$

Beweis. Wir verwenden Induktion über $|I|$. Für $|I| = 2$ ist die Behauptung wegen $A_1 \cup A_2 = A_1 \uplus (A_2 \setminus A_1)$ und $(A_2 \setminus A_1) \uplus (A_1 \cap A_2) = A_2$ klar. Gilt sie für $I = \{1,\dots,n\}$, so ist wegen der Additivität von μ

$$\begin{aligned}\mu\Big(\bigcup_{i=1}^{n+1} A_i\Big) &= \mu\Big(\bigcup_{i=1}^{n}(A_i \cup A_{n+1})\Big)\\ &= \sum_{\emptyset\neq J\subseteq\{1,\dots,n\}}(-1)^{|J|+1}\mu\Big(A_{n+1} \cup \bigcap_{j\in J} A_j)\Big)\\ &= \sum_{\emptyset\neq J\subseteq\{1,\dots,n\}}(-1)^{|J|+1}\Big(\mu(A_{n+1}) + \mu\Big(\bigcap_{j\in J} A_j)\Big) - \mu\Big(\bigcap_{j\in J} A_j \cap A_{n+1})\Big)\Big)\\ &= \mu(A_{n+1}) + \sum_{\emptyset\neq J\subseteq\{1,\dots,n\}}(-1)^{|J|+1}\Big(\mu\Big(\bigcap_{j\in J} A_j)\Big) - \mu\Big(\bigcap_{j\in J} A_j \cap A_{n+1})\Big)\Big)\\ &= \sum_{J\subseteq\{1,\dots,n+1\}}(-1)^{|J|+1}\mu\Big(\bigcap_{j\in J} A_j)\Big).\end{aligned}$$

□

21.2 σ-Additivität

Die endliche Additivität von Mengenfunktionen ist eine Forderung, die man oft nachprüfen kann. Anders sieht es mit der σ-Additivität aus. Wir werden nun alternative Formulierungen für σ-Additivität kennen lernen.

Proposition 21.7 (Stetigkeit von unten und von oben). *Sei $\mathcal{R}$ ein Ring und $\mu : \mathcal{R} \to \overline{\mathbb{R}}_+$ endlich additiv. Betrachte folgende Eigenschaften:*

1. *μ ist σ-additiv*
2. *μ ist σ-subadditiv*
3. *μ ist stetig von unten, d.h. für $A, A_1, A_2, \dots \in \mathcal{H}$ und $A_1 \subseteq A_2 \subseteq \dots$ mit $A = \bigcup_{n=1}^\infty A_n$ ist $\mu(A) = \lim_{n\to\infty}\mu(A_n)$.*
4. *μ ist stetig von oben in $\emptyset$, d.h. für $A_1, A_2, \dots \in \mathcal{H}$, $\mu(A_1) < \infty$ und $A_1 \supseteq A_2 \supseteq \dots$ mit $\bigcap_{n=1}^\infty A_n = \emptyset$ ist $\lim_{n\to\infty}\mu(A_n) = 0$.*

5. *μ ist stetig von oben, d.h. für $A, A_1, A_2, \dots \in \mathcal{H}$, $\mu(A_1) < \infty$ und $A_1 \supseteq A_2 \supseteq \dots$ mit $A = \bigcap_{n=1}^{\infty} A_n$ ist $\mu(A) = \lim_{n\to\infty} \mu(A_n)$.*

Dann gilt

$$1. \iff 2. \iff 3. \implies 4. \iff 5..$$

Außerdem gilt 4. $\implies$ 3., falls $\mu(A) < \infty$ für alle $A \in \mathcal{R}$.

Beweis. 1.$\Leftrightarrow$2. folgt aus Lemma 21.5, da $\mathcal{R}$ ein Halbring ist.

1.$\Rightarrow$3.: Sei μ also σ-additiv und $A, A_1, A_2, \dots \in \mathcal{R}$ wie in 3. Dann gilt mit $A_0 = \emptyset$

$$\mu(A) = \sum_{n=1}^{\infty} \mu(A_n \setminus A_{n-1}) = \lim_{N\to\infty} \sum_{n=1}^{N} \mu(A_n \setminus A_{n-1}) = \lim_{N\to\infty} \mu(A_N).$$

3.$\Rightarrow$1.: Sei $B_1, B_2, \dots \in \mathcal{R}$ paarweise disjunkt und $B = \biguplus_{n=1}^{\infty} B_n \in \mathcal{R}$. Dann gilt für $A_N = \biguplus_{n=1}^{N} B_n$

$$\mu(B) = \lim_{N\to\infty} \mu(A_N) = \sum_{n=1}^{\infty} \mu(B_n).$$

4.$\Rightarrow$5.: Sei $A, A_1, A_2, \dots \in \mathcal{H}$ wie in 5. vorausgesetzt. Weiter sei $B_n := A_n \setminus A$. Dann erfüllt $B_1, B_2, \dots$ die Voraussetzungen von 4., also gilt $\mu(B_n) \xrightarrow{n\to\infty} 0$, also $\mu(A_n) = \mu(B_n) + \mu(A) \xrightarrow{n\to\infty} \mu(A)$.

5.$\Rightarrow$4.: ist klar.

3.$\Rightarrow$4.: Sei $A_1, A_2, \dots \in \mathcal{R}$ wie in 4. vorausgesetzt. Setze $B_n := A_1 \setminus A_n, n \in \mathbb{N}$. Dann erfüllt $B = A_1, B_1, B_2, \dots \in \mathcal{R}$ die Voraussetzungen von 3., und damit gilt $\mu(A_1) = \lim_{n\to\infty} \mu(B_n) = \mu(A_1) - \lim_{n\to\infty} \mu(A_n)$, woraus 4. folgt.

4.$\Rightarrow$3. falls $\mu(A) < \infty$ für alle $A \in \mathcal{R}$. Sei $A, A_1, A_2, \dots \in \mathcal{R}$ wie in 3. vorausgesetzt. Setze $B_n := A \setminus A_n \in \mathcal{R}, n \in \mathbb{N}$. Dann gilt $\bigcap_{n=1}^{\infty} B_n = \emptyset$, also $0 = \lim_{n\to\infty} \mu(B_n) = \mu(A) - \lim_{n\to\infty} \mu(A_n)$, woraus 3. folgt. Hier geht in der letzten Gleichheit die Voraussetzung ein, dass $\mu(A) < \infty$. □

Wir wollen nun von innen bezüglich eines kompakten Systems reguläre Mengenfunktionen näher beleuchten. In vielen Situationen ist die Forderung, dass ein Maß bezüglich eines kompakten Systems von innen regulär ist, erfüllt.

Lemma 21.8. *Ist $(E, \mathcal{O})$ polnisch und μ ein endliches Maß auf $\mathcal{B}(\mathcal{O})$, so existiert zu jedem $\varepsilon > 0$ eine kompakte Menge $K \subseteq E$ mit $\mu(E \setminus K) < \varepsilon$.*

Beweis. Zuerst bemerken wir, dass alle kompakten Mengen nach Lemma 19.10 auch abgeschlossen sind, also sind alle kompakten Mengen in $\mathcal{B}(\mathcal{O})$ und damit ist $\mu(E \setminus K)$ im Lemma wohldefiniert.

Sei $\varepsilon > 0$. Da E separabel ist, gibt es für jedes $n = 1, 2, \dots$ Punkte $x_1^n, x_2^n, \dots \in E$ mit $E = \bigcup_{k=1}^{\infty} B_{1/n}(x_k^n)$. Da μ von oben stetig ist (Proposition 21.7), gilt

$$0 = \mu\Big(E \setminus \bigcup_{k=1}^{\infty} B_{1/n}(x_k^n)\Big) = \lim_{N\to\infty} \mu\Big(E \setminus \bigcup_{k=1}^{N} B_{1/n}(x_k^n)\Big).$$

Damit gibt es ein $N_n \in \mathbb{N}$ mit $\mu\Big(E \setminus \bigcup_{k=1}^{N_n} B_{1/n}(x_k^n)\Big) < \varepsilon/2^n$. Nun ist

$$A := \bigcap_{n=1}^{\infty} \bigcup_{k=1}^{N_n} B_{1/n}(x_k^n)$$

nach Definition total beschränkt, also relativ kompakt nach Lemma 19.11. Außerdem gilt für die kompakte Menge $\overline{A}$

$$\mu(E \setminus \overline{A}) \le \mu(E \setminus A) \le \mu\Big(\bigcup_{n=1}^{\infty} \Big(E \setminus \bigcup_{k=1}^{N_n} B_{1/n}(x_k^n)\Big)\Big) \le \sum_{n=1}^{\infty} \mu\Big(E \setminus \bigcup_{k=1}^{N_n} B_{1/n}(x_k^n)\Big) < \varepsilon.$$

Damit ist die Behauptung gezeigt. □

Theorem 21.9 (Von innen reguläre additive Mengenfunktionen sind σ-additiv). *Sei $\mathcal{H}$ ein Halbring und $\mu : \mathcal{H} \to \mathbb{R}_+$ endlich, endlich additiv und für ein kompaktes System $\mathcal{K} \subseteq \mathcal{H}$ von innen regulär. Dann ist μ auch σ-additiv.*

Beweis. Wie in Lemma 21.5 kann man die Mengenfunktion μ auf eindeutige Weise auf den von $\mathcal{H}$ erzeugten Ring $\mathcal{R}(\mathcal{H})$ (siehe Lemma 20.5) erweitern. Weiter ist nach Lemma 20.16 das System $\mathcal{K}_\cup \subseteq \mathcal{R}(\mathcal{H})$, das durch Vereinigungsbildung aus Mengen von $\mathcal{K}$ besteht, ebenfalls kompakt. Wählt man nun $\varepsilon > 0$ und $A = \bigcup_{i=1}^n A_i \in \mathcal{R}(\mathcal{H})$ mit $A_1, \dots, A_n \in \mathcal{H}$, so gibt es kompakte Mengen $K_1, \dots, K_n \in \mathcal{K} \subseteq \mathcal{H}$ mit $\mu(A_i) \le \mu(K_i) + \frac{\varepsilon}{n}$ für $i = 1, \dots, n$. Damit gilt für die Erweiterung von μ auf den Ring $\mathcal{R}(\mathcal{H})$

$$\mu\Big(\bigcup_{i=1}^n A_i\Big) = \sum_{i=1}^n \mu(A_i) \le \Big(\sum_{i=1}^n \mu(K_i)\Big) + \varepsilon = \mu\Big(\bigcup_{i=1}^n K_i\Big) + \varepsilon.$$

Damit ist μ von innen $\mathcal{K}_\cup$-regulär. Also ist o.E. der Halbring $\mathcal{H}$ ein Ring und $\mathcal{K}$ vereinigungsstabil.

Wir zeigen nun, dass μ stetig von oben in $\emptyset$ ist. Dies genügt nach Proposition 21.7 wegen der Endlichkeit von μ auf $\mathcal{H}$. Seien $A_1, A_2, \dots \in \mathcal{H}$ mit $A_1 \supseteq A_2 \supseteq \cdots$ und $\bigcap_{n=1}^\infty A_n = \emptyset$ und $\varepsilon > 0$. Wähle $K_1, K_2, \dots \in \mathcal{K}$ mit $K_n \subseteq A_n, n \in \mathbb{N}$ und

$$\mu(A_n) \le \mu(K_n) + \varepsilon 2^{-n}.$$

Es gilt $\bigcap_{n=1}^\infty K_n \subseteq \bigcap_{n=1}^\infty A_n = \emptyset$. Deswegen gibt es ein $N \in \mathbb{N}$ mit $\bigcap_{n=1}^N K_n = \emptyset$. Damit gilt

$$A_N = A_N \cap \Big(\bigcup_{n=1}^N K_n^c\Big) = \bigcup_{n=1}^N A_N \setminus K_n \subseteq \bigcup_{n=1}^N A_n \setminus K_n.$$

Daraus folgt wegen der Subadditivität und der Monotonie von μ für alle $m \ge N$

$$\mu(A_m) \le \mu(A_N) \le \sum_{n=1}^N \mu(A_n \setminus K_n) \le \varepsilon \sum_{n=1}^N 2^{-n} \le \varepsilon.$$

Damit ist die Behauptung gezeigt, da $\varepsilon > 0$ beliebig war. □

	Lemma 21.4	Theorem 21.9	Theorem 21.15
μ endlich additiv	$\circ$	$\circ$	
μ endlich		$\circ$	
μ σ-endlich			$\circ$
μ definiert auf Halbring	$\circ$	$\circ$	$\circ$
μ σ-additiv	$\circ/\bullet$	$\bullet$	$\circ$
μ σ-subadditiv	$\bullet/\circ$		
μ von innen $\mathcal{K}$-regulär		$\circ$	
μ eindeutig auf $\sigma(\mathcal{H})$ fortsetzbar			$\bullet$

Tabelle 21.1: Um den Carathéodory'schen Fortsetzungssatz anwenden zu können, spielen Lemma 21.4 und Theorem 21.9 eine Rolle. In der Tabelle bedeuten die $\circ$'s jeweils die Voraussetzungen des Satzes und $\bullet$ die Folgerungen. Wie man leicht ablesen kann, gilt der Carathéodory'sche Fortsetzungssatz z.B. dann, wenn μ endlich und von innen $\mathcal{K}$-regulär ist.

21.3 Eindeutigkeit und Fortsetzung von Maßen

Angenommen, eine additive Mengenfunktion $\mu : \mathcal{H} \to \overline{\mathbb{R}}_+$ ist gegeben. Wir beschäftigen uns damit, wann es möglich ist, diese Mengenfunktion zu einem Maß (d.h. einer σ-additiven Mengenfunktion) auf $\sigma(\mathcal{H})$ auszuweiten. Ziel ist es dabei, Bedingungen aufzustellen, wann das Maß durch μ schon eindeutig gegeben ist. Das Resultat ist in Theorem 21.15 zusammen gefasst. Siehe auch Tabelle 21.1 für eine Übersicht, wie die Resultate vergangener Kapitel damit zusammen hängen.

Proposition 21.10 (Eindeutigkeit von Maßen). *Sei $\mathcal{F}$ eine σ-Algebra und $\mu, \nu : \mathcal{F} \to \mathbb{R}_+$ Maße. Sei $\mathcal{C}$ ein durchschnittstabiles Mengensystem, so dass $\sigma(\mathcal{C}) = \mathcal{F}$ und $\mu|_\mathcal{C}, \nu|_\mathcal{C}$ sind σ-endlich, d.h. $\mu(A) = \sup_{C\in\mathcal{C}} \mu(A\cap C), \nu(A) = \sup_{C\in\mathcal{C}} \nu(A\cap C), A \in \mathcal{F}$. Dann ist $\mu = \nu$ genau dann, wenn $\mu(A) = \nu(A)$ für alle $A \in \mathcal{C}$ gilt.*

Beweis. Die 'genau dann'-Richtung ist klar. Für die 'wenn'-Richtung setzen wir für ein $C \in \mathcal{C}$ mit $\mu(C) = \nu(C) < \infty$

$$\mathcal{D}_C := \{A \in \mathcal{F} : \mu(A \cap C) = \nu(A \cap C)\} \supseteq \mathcal{C}.$$

Wir zeigen, dass $\mathcal{D}_C$ ein Dynkin System ist. Klar ist, dass $E \in \mathcal{D}_C$. Ist weiter $A, B \in \mathcal{D}$ und $A \subseteq B$, so ist $\mu((B\setminus A)\cap C) = \mu(B\cap C) - \mu(A\cap C) = \nu(B\cap C) - \nu(A\cap C) = \nu((B\setminus A)\cap C)$, also $B \setminus A \in \mathcal{D}_C$. Ist $A_1, A_2, \dots \in \mathcal{D}$ mit $A_1 \subseteq A_2 \subseteq A_3 \subseteq \dots \in \mathcal{D}_C$ und $A = \bigcup_{n=1}^\infty A_n \in \mathcal{F}$, so ist wegen Proposition 21.7.

$$\mu(A \cap C) = \lim_{n\to\infty} \mu(A_n \cap C) = \lim_{n\to\infty} \nu(A_n \cap C) = \nu(A \cap C),$$

also $A \in \mathcal{D}_C$. Damit ist $\mathcal{D}_C$ für alle $C \in \mathcal{C}$ mit $\mu(C) < \infty$ ein Dynkin-System und damit gilt wegen Theorem 20.13, dass $\mathcal{F} = \sigma(\mathcal{C}) = \mathcal{D}_C$. Sei $E_1, E_2, \dots \in \mathcal{C}$ mit $E_n \uparrow E$ und

$\mu(E_n), \nu(E_n) < \infty, n = 1, 2, \ldots$ Dann gilt für alle $n = 1, 2, \ldots$, dass $\mu(A \cap E_n) = \nu(A \cap E_n)$ für alle $A \in \mathcal{F}$. Daraus folgt dann für $A \in \mathcal{F}$, da μ und ν stetig von unten sind,

$$\mu(A) = \lim_{n\to\infty} \mu(A \cap E_n) = \lim_{n\to\infty} \nu(A \cap E_n) = \nu(A),$$

d.h. $\mu = \nu$. □

Beispiel 21.11 (Notwendigkeit der σ-Endlichkeit in Proposition 21.10). Proposition 21.10 gilt nicht notwendigerweise, wenn $\mu|_\mathcal{C}$ oder $\nu|_\mathcal{C}$ nicht σ-endlich sind.
Sei hierzu $\mathcal{C} = \{\emptyset\}$, $\mathcal{F} = \sigma(\mathcal{C}) = \{\emptyset, E\}$ und

$$\mu(\emptyset) = \nu(\emptyset) = 0, \qquad \mu(E) = 0, \qquad \nu(E) = 1,$$

also insbesondere $\mu \neq \nu$ aber $\mu|_\mathcal{C} = \nu|_\mathcal{C}$. Allerdings ist $\nu|_\mathcal{C}$ nicht σ-endlich, weil $\nu(E) = 1 \neq 0 = \nu(\emptyset \cap E) = \sup_{C\in\mathcal{C}} \nu(C \cap E)$.

Das folgende Theorem erklärt, warum der Begriff der σ-Algebra zentral ist.

Theorem 21.12 ($\boldsymbol{\mu^*}$-messbare Mengen sind eine $\boldsymbol{\sigma}$-Algebra). *Sei μ^* ein äußeres Maß auf E und $\mathcal{F}^*$ die Menge der μ^*-messbaren Mengen. Dann ist $\mathcal{F}^*$ eine σ-Algebra und $\mu := \mu^*|_{\mathcal{F}^*}$ ein Maß. Außerdem ist $\mathcal{N} := \{N \subseteq E : \mu^*(N) = 0\} \subseteq \mathcal{F}^*$.*

Bemerkung 21.13 (Nullmengen und Eigenschaften fast überall). Mengen N mit $\mu(N) = 0$ heißen auch *(μ-)Nullmengen*. Weiter sagen wir, dass $A \subseteq E$ *(μ)-fast überall* gilt, wenn $A^c \subseteq N$ und N eine μ-Nullmenge ist. Ist μ ein Wahrscheinlichkeitsmaß, sagen wir anstatt 'fast überall' auch *fast sicher*.

Beweis von Theorem 21.12. Wir zeigen zunächst, dass $\mathcal{F}^*$ eine σ-Algebra ist. Klar ist, dass

$$\mu^*(B) = \mu^*(\emptyset) + \mu^*(B) = \mu^*(B \cap \emptyset) + \mu^*(B \cap E),$$

d.h. $\emptyset \in \mathcal{F}^*$. Weiter ist klar, dass aus $A \in \mathcal{F}^*$ auch $A^c \in \mathcal{F}^*$ folgt. Mit $A, A' \in \mathcal{F}^*$ ist wegen $(B \cap A \cap A'^c) \uplus (B \cap A^c) = B \cap (A \cap A')^c$ und der Subadditivität von μ^*

$$\begin{aligned}\mu^*(B) &= \mu^*(B \cap A) + \mu^*(B \cap A^c) = \mu^*((B \cap A) \cap A') + \mu^*((B \cap A) \cap A'^c) + \mu^*(B \cap A^c)\\ &\geq \mu^*(B \cap (A \cap A')) + \mu^*(B \cap (A \cap A')^c) \geq \mu^*(B).\end{aligned}$$

Damit ist $A \cap A' \in \mathcal{F}^*$. Seien nun $A_1, A_2, \dots \in \mathcal{F}^*$ disjunkt und $B_n = \biguplus_{k=1}^n A_k$ und $B = \bigcup_{n=1}^\infty B_n = \biguplus_{k=1}^\infty A_k$. Da $\mathcal{F}^*$ durchschnitt- und komplementstabil ist, ist auch $B_1, B_2, \dots \in \mathcal{F}^*$. Weiter zeigen wir, dass $\mu^*(D \cap B_n) = \sum_{k=1}^n \mu^*(D \cap A_k)$ für alle $D \subseteq E$ gilt. Für $n = 1$ ist dies klar, und gilt es für ein n, so folgt, da $B_n \in \mathcal{F}^*$,

$$\begin{aligned}\mu^*(D \cap B_{n+1}) &= \mu^*(D \cap B_{n+1} \cap B_n) + \mu^*(D \cap B_{n+1} \cap B_n^c)\\ &= \mu^*(D \cap B_n) + \mu^*(D \cap A_{n+1}) = \sum_{k=1}^{n+1} \mu^*(D \cap A_k).\end{aligned}$$

Damit gilt wegen der Monotonie von μ^*

$$\mu^*(D \cap B) \leq \sum_{k=1}^\infty \mu^*(D \cap A_k) = \lim_{n\to\infty} \sum_{k=1}^n \mu^*(D \cap A_k) = \lim_{n\to\infty} \mu^*(D \cap B_n) \leq \mu^*(D \cap B),$$

also

$$\mu^*(D \cap B) = \lim_{n\to\infty} \mu^*(D \cap B_n) = \sum_{k=1}^{\infty} \mu^*(D \cap A_k). \tag{21.4}$$

Als nächstes zeigen wir, dass $B \in \mathcal{F}^*$ ist, woraus folgt, dass $\mathcal{F}^*$ eine σ-Algebra ist. Es gilt nämlich für $D \subseteq E$, wegen (21.4) und der Subadditivität von μ^*

$$\mu^*(D) = \lim_{n\to\infty} \mu^*(D \cap B_n) + \mu^*(D \cap B_n^c) \geq \mu^*(D \cap B) + \mu^*(D \cap B^c) \geq \mu^*(D),$$

woraus $B \in \mathcal{F}^*$ folgt. Weiter folgt aus (21.4), dass μ^* auf $\mathcal{F}^*$ sogar σ-additiv ist, d.h. dass $\mu = \mu^*|_{\mathcal{F}^*}$ ein Maß ist.

Sei nun $N \subseteq E$ so, dass $\mu^*(N) = 0$ und $D \subseteq E$. Dann ist wegen der Monotonie von μ^* auch $\mu^*(D \cap N) = 0$, also

$$\mu^*(D) \geq \mu^*(D \cap N^c) = \mu^*(D \cap N^c) + \mu^*(D \cap N)$$

und damit $N \in \mathcal{F}^*$. □

Proposition 21.14 (Von endlich additiver Mengenfunktion erzeugtes äußeres Maß). *Sei $\mathcal{H}$ ein Halbring und $\mu : \mathcal{H} \to \mathbb{R}_+$ endlich additiv. Für $A \subseteq E$ sei*

$$\mu^*(A) := \inf_{\mathcal{G} \in \mathcal{U}(A)} \sum_{G \in \mathcal{G}} \mu(G)$$

wobei

$$\mathcal{U}(A) := \Big\{\mathcal{G} \subseteq \mathcal{H} \text{ höchstens abzählbar, } A \subseteq \bigcup_{G \in \mathcal{G}} G\Big\}$$

die Menge der höchstens abzählbaren Überdeckungen von A ist und $\mu^*(A) = \infty$ falls $\mathcal{U}(A) = \emptyset$. Dann ist μ^* ein äußeres Maß.

Beweis. Die Abbildung μ^* ist monoton, und da $\emptyset \in \mathcal{H}$ ist $\mu^*(\emptyset) = 0$ wegen der endlichen Additivität von μ; siehe Lemma 21.4. (Es ist nämlich $\mu(\emptyset) = \mu(\emptyset) + \mu(\emptyset)$, woraus $\mu(\emptyset) = 0$ folgt.) Um die σ-Subadditivität von μ^* zu prüfen, wählen wir $A_1, A_2, \dots \subseteq E$. Für $n = 1, 2, \dots$ und $\epsilon > 0$ gibt es Mengen $G_{nk} \in \mathcal{H}$, $k \in \mathcal{K}_n$ höchstens abzählbar mit

$$A_n \subseteq \bigcup_{k \in \mathcal{K}_n} G_{nk},$$
$$\mu^*(A_n) \geq \sum_{k \in \mathcal{K}_n} \mu(G_{nk}) - \varepsilon 2^{-n}.$$

Daraus folgt wegen $\bigcup_{n=1}^{\infty} A_n \subseteq \bigcup_{n=1}^{\infty} \bigcup_{k \in \mathcal{K}_n} G_{nk}$, der Monotonie von μ^* und der Definition von μ^*

$$\mu^*\Big(\bigcup_{n=1}^{\infty} A_n\Big) \leq \sum_{n=1}^{\infty} \sum_{k \in \mathcal{K}_n} \mu(G_{nk}) \leq \varepsilon + \sum_{n=1}^{\infty} \mu^*(A_n).$$

Mit $\varepsilon \to 0$ folgt die σ-Subadditivität von μ^*, d.h. μ^* ist ein äußeres Maß. □

Theorem 21.15 (Fortsetzung einer σ-additiven Mengenfunktion). *Sei $\mathcal{H}$ ein Halbring und $\mu : \mathcal{H} \to \mathbb{R}_+$ σ-endlich und σ-additiv. Weiter sei $\widetilde{\mu} = \mu^*|_{\mathcal{F}^*}$ mit μ^* aus Proposition 21.14 und $\mathcal{F}^*$ aus Theorem 21.12. Dann ist $\sigma(\mathcal{H}) \subseteq \mathcal{F}^*$ und $\widetilde{\mu}|_{\sigma(\mathcal{H})}$ ist das einzige Maß, das auf $\mathcal{H}$ mit μ übereinstimmt.*

Beweis. Zunächst stellen wir fest, dass μ sowohl endlich additiv als auch σ-subadditiv ist nach Lemma 21.4. Nach Proposition 21.14 ist μ^* ein äußeres Maß und nach Theorem 21.12 ist $\mathcal{F}^*$ eine σ-Algebra.

Schritt 1: μ^* stimmt auf $\mathcal{H}$ mit μ überein: Sei $H \in \mathcal{H}$. Wähle $\mathcal{K}$ höchstens abzählbar und $H_k \in \mathcal{H}$, $k \in \mathcal{K}$ mit $H \subseteq \bigcup_{k\in\mathcal{K}} H_k$ und

$$\mu^*(H) \geq \sum_{k\in\mathcal{K}} \mu(H_k) - \varepsilon.$$

Damit gilt wegen $H = \bigcup_{k\in\mathcal{K}} H_k \cap H$ und der σ-Subadditivität von μ

$$\mu^*(H) \leq \mu(H) \leq \sum_{k\in\mathcal{K}} \mu(H_k \cap H) \leq \sum_{k\in\mathcal{K}} \mu(H_k) \leq \mu^*(H) + \varepsilon,$$

wobei wir im zweiten '$\leq$' die endliche Additivität von μ verwendet haben. Mit $\varepsilon \to 0$ folgt, dass $\mu^*(H) = \mu(H)$.

Schritt 2: $\sigma(\mathcal{H}) \subseteq \mathcal{F}^*$: Sei $D \subseteq E, H \in \mathcal{H}$ und $\varepsilon > 0$. Wähle $\mathcal{K}$ höchstens abzählbar und $H_k \in \mathcal{H}, k \in \mathcal{K}$ mit $\mu^*(D) \geq \sum_{k\in\mathcal{K}} \mu(H_k) - \varepsilon$. Dann gilt wegen der endlichen Additivität von μ

$$\mu^*(D) + \varepsilon \geq \sum_{k\in\mathcal{K}} \mu(H_k) = \sum_{k\in\mathcal{K}} \mu(H_k \cap H) + \sum_{k\in\mathcal{K}} \mu(H_k \cap H^c) \geq \mu^*(D \cap H) + \mu^*(D \cap H^c).$$

Mit $\varepsilon \to 0$ und der σ-Subadditivität von μ^* folgt $\mu^*(D) = \mu^*(D \cap H) + \mu^*(D \cap H^c)$, d.h. H ist μ^*-messbar und damit $\mathcal{H} \subseteq \mathcal{F}^*$. Da $\mathcal{F}^*$ nach Theorem 21.12 eine σ-Algebra ist, folgt auch $\sigma(\mathcal{H}) \subseteq \mathcal{F}^*$.

Schritt 3: Eindeutigkeit: Nach Theorem 21.12 ist $\widetilde{\mu}$ ein Maß, also auch $\widetilde{\mu}|_{\sigma(\mathcal{H})} = \mu^*|_{\sigma(\mathcal{H})}$. Sei $\nu : \sigma(\mathcal{H}) \to \mathbb{R}_+$ ein weiteres Maß, das auf $\mathcal{H}$ mit μ übereinstimmt. Da $\mu = \widetilde{\mu}|_{\mathcal{H}}$ als σ-endlich vorausgesetzt war, ist auch $\nu|_{\mathcal{H}}$ σ-endlich. Mit Proposition 21.10 folgt wegen der Durchschnittstabilität von $\mathcal{H}$, dass $\widetilde{\mu} = \nu$ auf $\sigma(\mathcal{H})$ gilt.

Nun sind alle Behauptungen bewiesen. □

In obigem Theorem wird nur klar, dass $\sigma(\mathcal{H}) \subseteq \mathcal{F}^*$. Das nächste Resultat zeigt, wodurch sich Mengen in $\mathcal{F}^*$ von Mengen in $\sigma(\mathcal{H})$ unterscheiden.

Proposition 21.16 (Charakterisierung von $\mathcal{F}^*$ aus Proposition 21.14). *Sei $\mathcal{H}$ ein Halbring, $\mu : \mathcal{H} \to \mathbb{R}_+$ σ-endlich und σ-additiv, μ^* wie in Proposition 21.14 und $\mathcal{F}^*, \mathcal{N}$ wie in Theorem 21.12. Dann ist*

$$\mathcal{F}^* = \{A \setminus N : A \in \sigma(\mathcal{H}), N \in \mathcal{N}\}.$$

Insbesondere ist die rechte Seite eine σ-Algebra.

Beweis. '$\supseteq$': Nach Theorem 21.15 ist $\sigma(\mathcal{H}) \subseteq \mathcal{F}^*$ und nach Theorem 21.12 ist $\mathcal{N} \subseteq \mathcal{F}^*$. Daraus folgt bereits '$\supseteq$', da $\mathcal{F}^*$ komplementstabil ist.

'$\subseteq$': Sei $B \in \mathcal{F}^*$. Weiter seien $E_1, E_2, \dots \in \mathcal{H}$ mit $\mu(E_n) < \infty, n = 1, 2, \dots$ und $E = \bigcup_{n=1}^\infty E_n$. Sei $\varepsilon_1, \varepsilon_2, \dots > 0$ mit $\varepsilon_i \downarrow 0$. Für $B_n := B \cap E_n$ wählen wir $\mathcal{K}_{ni}$ höchstens abzählbar, $A_{nik} \in \mathcal{H}, n \in \mathbb{N}, k \in \mathcal{K}_{ni}$, $B_n \subseteq \bigcup_{k \in \mathcal{K}_{ni}} A_{nik}$ und

$$\mu^*(B_n) \geq \sum_{k \in \mathcal{K}_{ni}} \mu(A_{nik}) - 2^{-n}\varepsilon_i.$$

Klar ist $A_i := \bigcup_{n=1}^\infty \bigcup_{k \in \mathcal{K}_{ni}} A_{nik} \in \sigma(\mathcal{H})$, $B \subseteq A_i, i \in \mathbb{N}$ und $A_i \setminus B = \bigcup_{n=1}^\infty \bigcup_{k \in \mathcal{K}_{ni}} A_{nik} \setminus B_n$. Damit gilt

$$\mu^*(A_i \setminus B) \leq \sum_{n=1}^\infty 2^{-n}\varepsilon_i = \varepsilon_i.$$

Setze $A = \bigcap_{i=1}^\infty A_i \in \sigma(\mathcal{H})$. Dann ist $B \subseteq A$, $N := A \setminus B \subseteq A_n \setminus B, n \in \mathbb{N}$ und

$$\mu^*(N) = \mu^*(A \setminus B) \leq \limsup_{i \to \infty} \mu^*(A_i \setminus B) \leq \limsup_{i \to \infty} \varepsilon_i = 0.$$

Damit folgt die Behauptung, da $B = A \setminus N$. $\square$

21.4 Maße auf $\mathcal{B}(\mathbb{R}^d)$

Wir lernen nun das Lebesgue-Maß auf $\mathcal{B}(\mathbb{R}^n)$ kennen. Dies ist die Fortsetzung der Mengenfunktion aus Beispiel 21.2.2 (diese war definiert auf dem Halbring der halboffenen Quader $\mathcal{H}$ aus Beispiel 20.3) auf $\sigma(\mathcal{H}) = \mathcal{B}(\mathbb{R}^d$. Anschließend charakterisieren wir alle σ-endliche Maße auf $\mathcal{B}(\mathbb{R})$.

Proposition 21.17 (Lebesgue-Maß auf $\mathbb{R}^d$). *Es gibt genau ein Maß λ auf $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, so dass*

$$\lambda\Big(\prod_{i=1}^d (a_i, b_i]\Big) = \prod_{i=1}^d (b_i - a_i) \tag{21.5}$$

für $a_i, b_i \in \mathbb{Q}$ mit $a_i \leq b_i, i = 1, ..., n$.

Beweis. Nach Beispiel 20.3 ist

$$\mathcal{H} = \Big\{\prod_{i=1}^n (a_i, b_i] : a_i, b_i \in \mathbb{Q}, a_i \leq b_i, i = 1, ..., d\Big\}$$

ein Halbring. Genauso ist

$$\widetilde{H} := \Big\{\prod_{i=1}^d \langle a, b\rangle : \langle \in \{[, (\}, \rangle \in \{],)\}, a_i, b_i \in \mathbb{Q}, a_i \leq b_i, i = 1, ..., d\}$$

ein Halbring mit $\sigma(\widetilde{\mathcal{H}}) = \mathcal{B}(\mathbb{R})$. Wir definieren $\widetilde{\lambda}$ auf $\widetilde{\mathcal{H}}$ durch

$$\widetilde{\lambda}\Big(\prod_{i=1}^d \langle a_i, b_i\rangle\Big) = \prod_{i=1}^d (b_i - a_i).$$

Dann ist $\widetilde{\lambda}$ klar σ-endlich. Es ist $\mathcal{K} = \left\{ \prod_{i=1}^{d} [a_i, b_i] : a_i, b_i \in \mathbb{Q}, a_i \leq b_i \right\} \subseteq \widetilde{\mathcal{H}}$ ein kompaktes System nach Beispiel 20.3. Außerdem ist klar $\widetilde{\lambda}$ von innen $\mathcal{K}$-regulär und damit σ-additiv nach Theorem 21.9. Nach Theorem 21.15 gibt es damit genau eine Fortsetzung von $\widetilde{\lambda}$ auf $\sigma(\mathcal{H}) = \mathcal{B}(\mathbb{R}^d)$. □

Beispiel 21.18 (Maß von $\mathbb{Q}$).
Zunächst stellen wir fest, dass $\mathbb{Q} \in \mathcal{B}(\mathbb{R})$.
Denn: Es ist $\{x\} \in \mathcal{B}(\mathbb{R})$, da $\{x\}^c$ offen ist (und damit wegen Abgeschlossenheit unter Komplementbildung mit $\{x\}^c \in \mathcal{B}(\mathbb{R})$ auch $\{x\} \in \mathcal{B}(\mathbb{R})$ gilt), $x \in \mathbb{R}$. Da $\mathcal{B}(\mathbb{R})$ abgeschlossen ist unter abzählbaren Vereiligungen, ist jede abzählbare Teilmenge von $\mathbb{R}$ Element von $\mathcal{B}(\mathbb{R})$. Insbesondere ist also $\mathbb{Q} \in \mathcal{B}(\mathbb{R})$.
Nun zeigen wir noch $\lambda(\mathbb{Q}) = 0$.
Denn: Zunächst gilt für jedes $x \in \mathbb{R}$

$$\lambda(\{x\}) = \lim_{n\to\infty} \lambda([x, x + \tfrac{1}{n})) = \lim_{n\to\infty} \tfrac{1}{n} = 0$$

wegen der Stetigkeit von λ von oben. Deshalb gilt aufgrund der σ-Additivität

$$\mu(\mathbb{Q}) = \sum_{x\in\mathbb{Q}} \mu(\{x\}) = \sum_{x\in\mathbb{Q}} 0 = 0.$$

Man beachte, dass die Abzählbarkeit von $\mathbb{Q}$ entscheidend ist. Ist etwa $A \subseteq \mathbb{R}$ überabzählbar, gilt $\mu(A) = \sum_{x\in A} \mu(\{x\})$ nicht mehr.

Proposition 21.19 (Charakterisierung der $\boldsymbol{\sigma}$-endlichen Maße auf $\mathbb{R}$). *Eine Funktion $\mu = \mathcal{B}(\mathbb{R}) \to \overline{\mathbb{R}}_+$ ist genau dann ein σ-endliches Maß, wenn es eine nicht-fallende und rechtsstetige Funktion $G : \mathbb{R} \to \mathbb{R}$ gibt mit*

$$\mu((a, b]) = G(b) - G(a) \tag{21.6}$$

für $a, b \in \mathbb{Q}$ mit $a \leq b$. Ist $\widetilde{G}$ eine weitere Funktion $\widetilde{G} : \mathbb{R} \to \mathbb{R}$ für die (21.6) gilt, so ist $\widetilde{G} = G + c$ für ein $c \in \mathbb{R}$.

Korollar 21.20 (Charakterisierung der Wahrscheinlichkeitsmaße auf $\mathbb{R}$). *Eine Funktion $\mu : \mathcal{B}(\mathbb{R}) \to [0, 1]$ ist genau dann ein Wahrscheinlichkeitsmaß, wenn es eine nicht-fallende und rechtsstetige Funktion $F : \mathbb{R} \to [0, 1]$ gibt mit $\lim_{b\to\infty} F(b) = 1$ und*

$$\mu((a, b]) = F(b) - F(a) \tag{21.7}$$

für $a, b \in \mathbb{Q}$ mit $a \leq b$. In diesem Fall ist F eindeutig durch μ festgelegt.

Beweis. Die Behauptung folgt direkt aus Proposition 21.19, da der Bildbereich von F nur für eine rechtsstetige, nicht-fallende Funktion genau $[0, 1]$ sein kann. □

Beweis von Proposition 21.19. '⇒': Sei also μ ein σ-endliches Maß auf $\mathcal{B}(\mathbb{R})$. Definiere dann $G(0) = 0$, sowie $G(x) := \mu((0, x])$ für $x > 0$ und $G(x) := \mu((x, 0])$ für $x < 0$. Dann ist G rechtsstetig, nicht fallend, und (etwa für $0 < a < b$) $\mu((a, b]) = \mu((0, b]) - \mu((0, a]) = G(b) - G(a)$.

'⇐': Der Beweis verläuft ähnlich zu dem Beweis von Korollar 21.17. Sei $\mathcal{H} = \{(a, b] : a, b \in \mathbb{Q}, a \leq b\}$ der Halbring der halboffenen Intervalle mit Enden in rationalen Zahlen. Wir zeigen, dass durch (21.6) eine σ-additive Mengenfunktion μ auf $\mathcal{H}$ definiert wird. Dann sieht man mit Theorem 21.15, dass μ auf eindeutige Weise zu einem σ-endlichen Maß auf $\sigma(\mathcal{H}) = \mathcal{B}(\mathbb{R})$ erweitert werden kann. Sei also $a_1, a_2, \dots$ disjunkt so, dass $\bigcup_{n=1}^{\infty}(a_n, a_{n+1}] = (a, b] \in \mathcal{H}$. Ohne Einschränkung ist $a_1 \geq a_2 \geq \dots$ Dann ist $b = a_1$ und $a_n \xrightarrow{n\to\infty} a$. Es gilt wegen der Rechtsstetigkeit von G, dass

$$\mu(a,b] = G(b) - G(a) = G(a_1) - \lim_{N\to\infty} G(a_N) = \sum_{n=1}^{\infty} G(a_n) - G(a_{n+1}) = \sum_{n=1}^{\infty} \mu((a_n, a_{n+1}]),$$

und damit ist die σ-Additivität von μ gezeigt.

Sei nun $\widetilde{G}$ eine weitere Funktion, für die (21.6) gilt. Dann gilt für alle $a \in \mathbb{R}$

$$\widetilde{G}(b) = \widetilde{G}(a) + \mu((a,b]) = G(b) + \widetilde{G}(a) - G(a)$$

und die Behauptung folgt mit $c = \widetilde{G}(a) - G(a)$. □

Definition 21.21 (Lebesgue-Maß und Verteilungsfunktion). *1. Das eindeutig definierte Maß λ aus Proposition 21.17 heißt d-dimensionales Lebesgue-Maß.*

2. Für ein Wahrscheinlichkeitsmaß μ auf $\mathcal{B}(\mathbb{R})$ heißt die Funktion F aus Korollar 21.20 Verteilungsfunktion.

Beispiel 21.22 (Bekannte Verteilungsfunktionen). 1. Sei $f : \mathbb{R} \to \mathbb{R}_+$ stückweise stetig, so dass[28] $\int_{-\infty}^{\infty} f(x)dx < \infty$ gilt. Solche Funktionen definieren vermöge

$$\mu_f((a,b]) := \int_a^b f(x)dx$$

wegen Korollar 21.20 auf eindeutige Art und Weise ein endliches Maß auf $\mathcal{B}(\mathbb{R})$. Die Funktion G aus Proposition 21.19 ist dann etwa gegeben durch

$$G(b) = \int_0^b f(x)dx.$$

In diesem Fall definieren wir

$$f \cdot \lambda := \mu_f.$$

Siehe auch Definition 23.12.

2. Wir geben zwei Beispiele für Verteilungsfunktionen (von Wahrscheinlichkeitsmaßen auf $\mathcal{B}(\mathbb{R})$) an. Es ist für $x \geq 0$

$$F_{\exp(\lambda)}(x) = \int_{-\infty}^{x} 1_{[0,\infty)}(a)\lambda e^{-\lambda a} da = 1 - e^{-\lambda x} \tag{21.8}$$

die Verteilungsfunktion der Exponentialverteilung mit Parameter λ. Außerdem ist

$$F_{N(\mu,\sigma^2)}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{x} \exp\Big(-\frac{(a-\mu)^2}{2\sigma^2}\Big) da =: \Phi(x) \tag{21.9}$$

die Verteilungsfunktion der Normalverteilung $N(\mu, \sigma^2)$ mit dem Erwartungswert μ und der Varianz σ^2.

[28] Wir verwenden hier das Riemann-Integral $\int_a^b f(x)dx$ aus Abschnitt 8. Einen weiteren Integralbegriff, das Lebesgue-Integral, werden wir in Kapitel 22 kennen lernen.

21.5 Bildmaße

Sei μ ein Maß auf $\mathcal{F}$. Wenn man nun den Grundraum vermöge einer Funktion $f : E \to E'$ transformiert, kann man auf E' ein zu der Transformation korrespondierendes Maß definieren, das sogenannte Bildmaß. Sei etwa $E := \mathbb{R}$, $\mathcal{F} = \mathcal{B}(\mathbb{R})$ und $f : u \mapsto u/\lambda$. Wir werden dann sehen, dass das Bildmaß von $\mu_{\exp(1)}$ unter f das Maß $\mu_{\exp(\lambda)}$ ist. Wir erinnern zunächst an die Situation aus Beispiel 20.3.2 und definieren das Bildmaß.

Definition 21.23 (Bildmaß). *Ist $(E, \mathcal{F}, \mu)$ ein Maßraum, $(E', \mathcal{F}')$ ein Messraum und $f : E \to E'$, so dass $\sigma(f) \subseteq \mathcal{F}$ für $\sigma(f)$ aus* (20.1). *Dann definieren wir eine Mengenfunktion $f_*\mu : \mathcal{F}' \to \mathbb{R}_+$ durch*

$$f_*\mu(A') := \mu(f^{-1}(A')) = \mu(f \in A'), \qquad A' \in \mathcal{F}'.$$

Hier heißt $f_*\mu$ auch Bildmaß *von μ unter f. Ist μ ein Wahrscheinlichkeitsmaß, so heißt $f_*\mu$ auch* Verteilung *von f.*

Bemerkung 21.24 (Messbare Funktionen). Ist $\sigma(f) \subseteq \mathcal{F}$ wie in obiger Definition, so sagen wir, dass f (nach $\mathcal{F}/\mathcal{F}'$) messbar ist. Dieses Konzept wird uns im nächsten Abschnitt beschäftigen.

Proposition 21.25 (Bildmaß ist ein Maß). *Sei $(E, \mathcal{F}, \mu)$, $(E', \mathcal{F}')$, $f : E \to E'$ und $f_*\mu$ wie in Definition 21.23. Dann ist $f_*\mu$ ein Maß auf $\sigma(f)$.*

Beweis. Seien $A'_1, A'_2, \dots \in \mathcal{F}'$ disjunkt, so gilt

$$f_*\mu\Big(\biguplus_{n=1}^\infty A'_n\Big) = \mu\Big(f^{-1}\Big(\biguplus_{n=1}^\infty A'_n\Big)\Big) = \mu\Big(\biguplus_{n=1}^\infty (f^{-1}(A'_n)\Big) = \sum_{n=1}^\infty \mu(f^{-1}(A'_n)) = \sum_{n=1}^\infty f_*\mu(A'_n).$$

Damit ist $f_*\mu$ also σ-additiv und die Behauptung ist gezeigt. □

Beispiel 21.26 (Bekannte Transformationen).
1. Sei λ das ein-dimensionale Lebesgue-Maß und $f : x \mapsto x^2$. Dann berechnen wir für $0 \le a \le b$

   $$f_*\lambda([a,b)) = \lambda(f^{-1}([a,b))) = \lambda([\sqrt{a}, \sqrt{b})) = \sqrt{b} - \sqrt{a} = \int_a^b \frac{1}{2\sqrt{x}} dx = g \cdot \lambda$$

   für $g(x) := 1_{x \ge 0} \frac{1}{2\sqrt{x}}$.

2. Für $E = [0, \infty]$ ist $\{[0,b) : g \ge 0\}$ ein schnittstabiles Erzeugendensystem von $\mathcal{B}([0,\infty))$. Sei $\mu_{\exp(1)}$ die Exponentialverteilung auf $[0,\infty)$ mit Verteilungsfunktion $F_{\exp(1)}$ aus (21.8) und $f : u \mapsto u/\lambda$. Dann ist $f_*\mu_{\exp(1)} = \mu_{\exp(\lambda)}$, denn

   $$f_*\mu_{\exp(1)}([0,b)) = \mu_{\exp(1)}(f^{-1}([0,b))) = \mu([0,\lambda b]) = 1 - e^{-\lambda b} = \mu_{\exp(\lambda)}([0,b)).$$

   Damit stimmen $\mu_{\exp(\lambda)}$ und $f_*\mu_{\exp(1)}$ auf einem schnittstabilen Erzeuger überein und die Aussage folgt mit Proposition 21.10.

3. Sei $E = \mathbb{R}^d$, $f_{\underline{y}} : \underline{x} \mapsto x + \underline{y}$ für ein $\underline{y} \in \mathbb{R}$ und λ das d-dimensionale Lebesgue-Maß aus Korollar 21.17. Dann ist $(f_{\underline{y}})_*\lambda = \lambda$, denn

   $$(f_{\underline{y}})_*\lambda\Big(\prod_{i=1}^d [a_i, b_i)\Big) = \lambda\Big(f_{\underline{y}}^{-1}\Big(\prod_{i=1}^d [a_i, b_i)\Big)\Big) = \lambda\Big(\prod_{i=1}^d [a_i - y_i, b_i - y_i)\Big) = \prod_{i=1}^d (b_i - a_i).$$

   Man sagt auch, dass das Lebesgue-Maß translationsinvariant ist.

Beispiel 21.27 (Beispiel einer nicht Borel-messbaren Menge, Vitali-Menge). Bisher gab es noch kein Beispiel für eine Menge, die nicht in $\mathcal{B}(\mathbb{R})$ liegt. Eine solche werden wir nun konstruieren. Sie ist als *Vitali-Menge* bekannt. Hierzu definieren wir auf $\mathbb{R}$ eine Äquivalenzrelation durch $x \sim y \iff y - x \in \mathbb{Q}$. Bezüglich dieser Äquivalenzrelation zerfällt $\mathbb{R}$ in Äquivalenzklassen der Form $\{x + q : q \in \mathbb{Q}\}$. Wir wählen aus jeder Äquivalenzklasse eine Zahl aus $[0, 1]$ aus und bilden damit die Menge V. (Es sei hier angemerkt, dass dieses Auswählen mittels des Auswahlaxioms der Mengenlehre geschieht und damit keinen trivialen Schritt darstellt.) Weiter sei nun für $q \in \mathbb{Q} \cap [-1, 1]$

$$V_q := \{x + q : x \in V\}.$$

Dann ist $[0, 1] \subseteq \biguplus_{q\in\mathbb{Q}\cap[-1,1]} V_q \subseteq [-1, 2]$.

Angenommen, die Menge V wäre messbar. Dann wären auch die Mengen V_q messbar und, wegen der Translationsinvarianz des Lebesgue-Maßes aus Beispiel 21.26.2 wäre $\lambda(V_q)$ nicht von q abhängig. Sei also $\lambda(V_q) =: a \geq 0$. Außerdem wäre dann wegen der Monotonie des Lebesgue-Maßes

$$1 \leq \sum_{q\in\mathbb{Q}\cap[-1,1]} \lambda(V_q) = \sum_{q\in\mathbb{Q}\cap[-1,1]} a \leq 3.$$

Dies ist jedoch nicht möglich, weder für $a = 0$ noch für $a > 0$. Wegen diesem Widerspruch muss also $V \notin \mathcal{B}(\mathbb{R})$ gelten.

22 Messbare Funktionen und das Integral

In diesem Kapitel sei immer $(E, \mathcal{F}, \mu)$ ein Maßraum. Wir können mittlerweile mit dem Maß μ Maßen den Inhalt von Mengen aus $\mathcal{F}$ messen. Ziel der Einführung des Integrals ist es, bei einer solchen Messung die Elemente von E verschieden zu gewichten. Diese Gewichtung wird mit einer Funktion $f : E \to \mathbb{R}$ vorgenommen. Solche Funktionen müssen die Minimalforderung der Messbarkeit genügen. Das Resultat dieser Gewichtung führt auf den Begriff des Integrals.

22.1 Messbare Funktionen

Wir kennen bereits den Begriff der (bezüglich der σ-Algebra $\mathcal{F}$) messbaren Menge, der nun auf den Begriff der messbaren Funktion erweitert wird. Sei $A \in \mathcal{F}$ also eine messbare Menge, so definiert diese mit $x \mapsto 1_A(x)$ eine messbare Funktion in dem nun eingeführte Sinne. Die Linearkombination von solchen Indikatorfunktionen werden wir einfache Funktion nennen. Diesen kommt wegen Theorem 22.8 eine besondere Bedeutung zu, da jede nicht-negative messbare Funktion (im Sinne der punktweisen Konvergenz) von unten durch einfache Funktionen approximiert werden kann.

Bemerkung 22.1 (Notation). Seien E, E' Mengen, $f : E \to E'$ und I beliebig.

1. Wir schreiben $f(A) := \{f(x) : x \in A\}$ für $A \subseteq E$ sowie $f^{-1}(A') := \{f^{-1}(x') : x' \in E'\}$ für $A' \subseteq E'$. Wir bemerken, dass folgende Regeln für $A', A'_i \subseteq E', i \in I$ gelten:

$$f^{-1}((A')^c) = (f^{-1}(A'))^c, \quad f^{-1}\Big(\bigcap_{i\in I} A'_i\Big) = \bigcap_{i\in I} f^{-1}(A'_i), \quad f^{-1}\Big(\bigcup_{i\in I} A'_i\Big) = \bigcup_{i\in I} f^{-1}(A'_i).$$

   Etwas Vorsicht ist jedoch geboten, da für $A, A_i \subseteq E, i \in I$ nur $f(\bigcup_{i\in I} A_n) = \bigcup_{i\in I} f(A_i)$, im allgemeinen jedoch $f(A^c) \neq (f(A))^c$ und $f(\bigcap_{i\in I} A_i) \neq \bigcap_{i\in I} f(A_i)$.

2. Für $\mathcal{C} \subseteq 2^{E'}$ schreiben wir ganz analog

$$f^{-1}(\mathcal{C}) := \{f^{-1}(A') : A' \in \mathcal{C}\}.$$

Lemma 22.2 (Urbilder von σ-Algebren). *Sei E eine Menge und $(E', \mathcal{F}')$ ein Messraum, $f : E \to E'$ und $\mathcal{C}' \subseteq \mathcal{F}'$ mit $\sigma(\mathcal{C}') = \mathcal{F}'$. Dann ist $\sigma(f^{-1}(\mathcal{C}')) = f^{-1}(\sigma(\mathcal{C}'))$. Insbesondere ist $f^{-1}(\mathcal{F}')$ eine σ-Algebra auf E.*

Beweis. '$\subseteq$': Nach Bemerkung 22.1 ist klar, dass $f^{-1}(\sigma(\mathcal{C}'))$ eine σ-Algebra ist. Damit ist $\sigma(f^{-1}(\mathcal{C}')) \subseteq \sigma(f^{-1}(\sigma(\mathcal{C}'))) = f^{-1}(\sigma(\mathcal{C}'))$.
'$\supseteq$': Wir definieren

$$\widetilde{\mathcal{F}}' := \{A' \in \sigma(\mathcal{C}') : f^{-1}(A') \in \sigma(f^{-1}(\mathcal{C}'))\} \subseteq \sigma(\mathcal{C}').$$

Dann ist, wieder wegen Bemerkung 22.1, $\widetilde{\mathcal{F}}'$ eine σ-Algebra und $\mathcal{C}' \subseteq \widetilde{\mathcal{F}}' \subseteq \sigma(\mathcal{C}')$, also $\widetilde{\mathcal{F}}' = \sigma(\mathcal{C}')$. Für $A' \in \sigma(\mathcal{C}')$ ist also $f^{-1}(A') \in \sigma(f^{-1}(\mathcal{C}'))$, was gleichbedeutend ist mit $f^{-1}(\sigma(\mathcal{C}')) \subseteq \sigma(f^{-1}(\mathcal{C}'))$. $\square$

Definition 22.3 (Messbare Funktionen). *Seien $(E, \mathcal{F})$ und $(E, \mathcal{F}')$ Messräume und $f : E \to E'$.*

1. *Die Funktion f heißt $\mathcal{F}/\mathcal{F}'$-*messbar*, falls $f^{-1}(\mathcal{F}') \subseteq \mathcal{F}$. Die σ-Algebra $f^{-1}(\mathcal{F}')$ heißt die* von f erzeugte σ-Algebra *auf E und wird auch mit $\sigma(f)$ bezeichnet.*
2. *Ist $(E, \mathcal{F}, \mathbf{P})$ ein Wahrscheinlichkeitsraum und $X : E \to E'$ messbar, so heißt X eine* E'-wertige Zufallsvariable. *Das Bildmaß $X_*\mathbf{P}$ aus Definition 21.23 heißt dann die* Verteilung von X.
3. *Ist $(E', \mathcal{F}') = (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$, so heißt f eine* reell(wertig)e Funktion. *Ist f nach $\mathcal{F}/\mathcal{B}(\overline{\mathbb{R}})$-messbar, so sagen wir, f sei (Borel-)messbar.*
4. *Ist $E' = \overline{\mathbb{R}}$ und $f = 1_A$ für $A \subseteq E$, so heißt f auch* Indikatorfunktion. *Ist $f = \sum_{k=1}^n c_k 1_{A_k}$ für $c_1, \dots, c_n \in \overline{\mathbb{R}}$ paarweise verschieden und $A_1, \dots, A_n \subseteq E$, so heißt f* einfach.

Beispiel 22.4. Sei $(E, \mathcal{F})$ ein Messraum.

1. Die allermeisten Funktionen $f : E \to \overline{\mathbb{R}}$, die man sich vorstellen kann, sind (Borel-)messbar. Etwa ist die Identitätsabbildung $f : x \mapsto x$ messbar, da $f^{-1}(\mathcal{F}) = \mathcal{F}$. Ein etwas komplizierteres Beispiel einer messbaren Funktion ist $f : x \mapsto x \cdot 1_{\mathbb{Q}}(x)$. Hier gilt für $0 \notin A'$, dass $f^{-1}(A') = A' \cap \mathbb{Q} \in \mathcal{B}(\mathbb{R})$ und für $0 \in A'$, dass $f^{-1}(A') = (A' \cap \mathbb{Q}) \cup \mathbb{Q}^c \in \mathcal{B}(\mathbb{R})$.
2. Es ist wichtig zu sehen, dass für viele messbaren Funktionen f gilt, dass $\sigma(f) \subsetneq \mathcal{F}$, siehe etwas das nächste Beispiel.
3. Eine Funktion $f : E \to \{0, 1\}$ ist genau dann messbar, wenn $f^{-1}(\{1\}) \in \mathcal{F}$. Es ist nämlich in diesem Fall $\sigma(f) = \{\emptyset, f^{-1}(\{1\}), (f^{-1}(\{1\}))^c, E\}$.
4. Sei $\mathcal{F} = \mathcal{B}(\mathbb{R})$. Um eine nicht $\mathcal{F}$-messbare Funktion anzugeben, muss man sich genauso anstrengen, wie ein nicht Borel-messbare Menge zu konstruieren. Es ist etwa für die Vitali-Menge V aus Beispiel 21.27 die Funktion 1_V nicht messbar.

Lemma 22.5 (Eigenschaften von Messbarkeit). *Seien $(E, \mathcal{F})$, $(E', \mathcal{F}')$ und $(E'', \mathcal{F}'')$ Messräume.*

1. *Ist $\mathcal{C}' \subseteq \mathcal{F}'$ mit $\mathcal{F}' = \sigma(\mathcal{C}')$, so ist $f : E \to E'$ genau dann nach $\mathcal{F}/\mathcal{F}'$-messbar, wenn $f^{-1}(\mathcal{C}') \subseteq \mathcal{F}$.*

2. *Ist $f : E \to E'$ nach $\mathcal{F}/\mathcal{F}'$ und $g : E' \to E''$ nach $\mathcal{F}'/\mathcal{F}''$ messbar. Dann ist $g \circ f : E \to E''$ nach $\mathcal{F}/\mathcal{F}''$-messbar.*

3. *Seien $(E, \mathcal{O})$ und $(E', \mathcal{O}')$ topologische Räume, $f : E \to E'$ stetig sowie $\mathcal{F} = \sigma(\mathcal{O})$ und $\mathcal{F}' = \sigma(\mathcal{O}')$ die Borel'schen σ-Algebren. Dann ist f nach $\mathcal{F}/\mathcal{F}'$-messbar.*

4. *Eine reelle Funktion f (d.h. $f : E \to \mathbb{R}$) ist genau dann messbar, wenn $\{x : f(x) \leq y\} \in \mathcal{F}$ für alle $y \in \mathbb{Q}$.*

5. *Eine einfache Funktion $f = \sum_{k=1}^{n} c_k 1_{A_k}$ mit paarweise verschiedenen $c_1, \dots, c_n \in \overline{\mathbb{R}}$ und $A_1, \dots, A_n \subseteq E$ ist genau dann messbar, wenn $A_1, \dots, A_n \in \mathcal{F}$.*

Beweis. 1. Die 'genau dann'-Richtung ist klar. Für die 'wenn'-Richtung verwenden wir Lemma 22.2 und erhalten $f^{-1}(\mathcal{F}') = f^{-1}(\sigma(\mathcal{C}')) = \sigma(f^{-1}(\mathcal{C}')) \subseteq \sigma(\mathcal{F}) = \mathcal{F}$. Damit ist f nach $\mathcal{F}/\mathcal{F}'$-messbar.
2. Wir schreiben direkt $(g \circ f)^{-1}(\mathcal{F}'') = f^{-1}(g^{-1}(\mathcal{F}'')) \subseteq f^{-1}(\mathcal{F}') \subseteq \mathcal{F}$, was die Behauptung bereits zeigt.
3. Nach Definition der Borel'schen σ-Algebra ist $\mathcal{O}'$ ein Erzeuger für $\mathcal{B}(E')$. Da f stetig ist (d.h. $f^{-1}(\mathcal{O}') \subseteq \mathcal{O}$) folgt $f^{-1}(\mathcal{O}') \subseteq \mathcal{O} \subseteq \sigma(\mathcal{O}) = \mathcal{B}(E)$. Nach 1. ist f also $\mathcal{B}(E)/\mathcal{B}(E')$-messbar.
4. Wir verwenden 1. Ist nämlich $E' = \mathbb{R}$, so wird nach Lemma 22.2 $\mathcal{B}(E')$ von $\mathcal{C} = \{(-\infty, y] : y \in \mathbb{Q}\}$ erzeugt. Also ist ein reelles f genau dann messbar, wenn $f^{-1}(\mathcal{C}') = \{\{x : f(x) \leq y\} : y \in \mathbb{Q}\} \subseteq \mathcal{F}$.
5. Sei $A := \Big(\bigcup_{k=1} A_k\Big)^c \in \mathcal{F}$. Dann ist $f^{-1}(\mathcal{B}(\overline{\mathbb{R}})) = \Big\{A \cup \bigcup_{j \in J} A_j, \bigcup_{j \in J} A_j : J \subseteq \{1, \dots, n\}\Big\}$, woraus die Behauptung folgt. □

Lemma 22.6 (Messbarkeit und Rechenregeln). *Sei $(E, \mathcal{F})$ ein Messraum.*

1. *Sind $f, g : E \to \overline{\mathbb{R}}$ messbar. Dann sind auch fg, sowie $af + bg$ für $a, b \in \mathbb{R}$ und f/g messbar, falls $g(x) \neq 0$ für alle $x \in E$.*

2. *Sind $f_1, f_2, \dots : E \to \overline{\mathbb{R}}$ messbar, so sind auch*
$$\sup_{n \in \mathbb{N}} f_n, \quad \inf_{n \in \mathbb{N}} f_n, \quad \limsup_{n \to \infty} f_n, \quad \liminf_{n \to \infty} f_n$$
messbar. Falls existent, ist auch $\lim_{n \to \infty} f_n$ messbar.

Beweis. 1. Betrachte die Abbildung $\psi : E \to \mathbb{R}^2$, definiert durch $\psi(x) = (f(x), g(x))$. Man überlegt sich leicht, dass ψ nach $\mathcal{F}/\mathcal{B}(\mathbb{R}^2)$-messbar ist. Weiter ist $(z, z') \mapsto az + bz'$ und $(z, z') \mapsto zz'$ auf $\mathbb{R}^2$ und $(z, z') \mapsto z/z'$ auf $\mathbb{R} \times (\mathbb{R} \setminus \{0\})$ stetig, also messbar nach Lemma 22.5.3. Damit folgen die Behauptungen nach Lemma 22.5.2.
2. Wir zeigen nur die Messbarkeit von $\sup_{n \in \mathbb{N}} f_n$. Die anderen Aussagen folgen dann mittels $\inf_{n \in \mathbb{N}} f_n = -\sup_{n \in \mathbb{N}}(-f_n)$, sowie $\limsup_{n \to \infty} f_n = \inf_{n \in \mathbb{N}} \sup_{k \geq n} f_k$ und $\liminf_{n \to \infty} f_n = \sup_{n \in \mathbb{N}} \inf_{k \geq n} f_k$. Es gilt für $x \in \mathbb{R}$ nach Lemma 22.5.4
$$\Big\{x : \sup_{n \in \mathbb{N}} f_n(x) \leq y\Big\} = \bigcap_{n=1}^{\infty} \underbrace{\Big\{x : f_n(x) \leq y\Big\}}_{\in \mathcal{F}} \in \mathcal{F}$$

und die Behauptung ist gezeigt. □

Korollar 22.7 (Messbarkeit von Positiv- und Negativteil). *Sei $(E, \mathcal{F})$ ein Messraum und $f : E \to \overline{\mathbb{R}}$. Dann ist f genau dann messbar, wenn $f^+ := f \vee 0$ und $f^- := (-f) \vee 0$ messbar sind. Dann ist auch $|f|$ messbar.*

Beweis. Es gilt $f = f^+ - f^-$ und $|f| = f^+ + f^-$. Damit folgt die Behauptung aus Lemma 22.6.2. □

Theorem 22.8 (Approximation mit messbaren Funktionen). *Sei $(E, \mathcal{F})$ ein Messraum und $f : E \to \overline{\mathbb{R}}$ messbar. Dann gibt es eine Folge $f_1, f_2, \dots : E \to \overline{\mathbb{R}}$ von einfachen Funktionen mit $f_n \uparrow f$ (wobei hier die punktweise Konvergenz gemeint ist, siehe Definition 8.3).*

Beweis. Wir schreiben einfach für $x \in E, n \in \mathbb{N}$

$$f_n(x) = n \wedge 2^{-n}[2^n f(x)],$$

und bemerken, dass $f_n \uparrow f$ per Konstruktion gilt. Außerdem ist $x \mapsto [2^n f(x)]$ nach Lemma 22.5.4 messbar, wenn f messbar ist. □

22.2 Definition

Die Konstruktion des Integrals einer Funktion f nach einem Maß erfolgt nun in mehreren Schritten. Sei $(E, \mathcal{F}, \mu)$ ein Maßraum. Wir bemerken, dass wir für das Integral von $f : E \to \mathbb{R}$ nach μ verschiedene synonyme Schreibweisen verwenden, nämlich

$$\mu[f] = \int f d\mu = \int f(x)\mu(dx). \tag{22.1}$$

Die Definition des Integrals erfolgt zunächst für einfache Funktionen und anschließend (siehe Theorem 22.8) durch Approximation für allgemeine nicht-negative messbare Funktionen. Das Integral für (nicht notwendig nicht-negative) messbare Funktionen wird dann durch das Integral des Positiv- und Negativteils definiert.

Definition 22.9 (Integral von einfachen Funktionen). *Sei $(E, \mathcal{F}, \mu)$ ein Maßraum und $f = \sum_{k=1}^m c_k 1_{A_k}$ eine einfache Funktion mit $c_1, \dots, c_m \geq 0, A_1, \dots, A_m \in \mathcal{F}$. Dann heißt*

$$\mu[f] := \int f d\mu := \sum_{k=1}^n c_k \mu(A_m)$$

das Integral von f bezüglich μ.

Bemerkung 22.10 (Wohldefiniertheit des Integrals). Wir müssen sicherstellen, dass obiges Integral wohldefiniert ist. Sei dazu $f = \sum_{l=1}^n d_l 1_{B_l}$ eine weitere Darstellung von f mit $d_1, \dots, d_n \geq 0$ und $B_1, \dots, B_n \in \mathcal{F}$. Dann gilt

$$\sum_{k=1}^m c_k \mu(A_k) = \sum_{k=1}^m \sum_{l=1}^n c_k \mu(A_k \cap B_l) = \sum_{k=1}^m \sum_{l=1}^n d_l \mu(A_k \cap B_l) = \sum_{l=1}^n d_l \mu(B_l),$$

woraus die Wohldefiniertheit folgt.

Lemma 22.11 (Einfache Eigenschaften). *Seien f, g nicht-negative, einfache Funktionen und $\alpha \geq 0$. Dann gilt*

$$\mu[af + bg] = a\mu[f] + b\mu[g], \qquad f \leq g \Rightarrow \mu[f] \leq \mu[g].$$

Beweis. Klar. □

Beispiel 22.12 (Das Integral von Indikatorfunktionen und Riemann-Integral). Sei $(E, \mathcal{F}, \mu)$ ein Maßraum und $A \in \mathcal{F}$. Dann ist $f = 1_A$ eine einfache Funktion und es gilt

$$\mu[f] = \mu(A)$$

nach Definition 22.9. Es sei hierbei bemerkt, dass die Funktion $f = 1_A$ nicht mehr stückweise stetig sein muss. (Sei etwa A das in Beispiel 20.10 betrachtete Cantor'sche Kontinuum.) Deswegen ist es nicht klar, dass die Funktion 1_A integrierbar bezüglich des Riemann-Integrals ist.

Etwa haben wir in Beispiel 8.13 gesehen, dass $\int 1_{\mathbb{Q}}(x)dx$ nicht existiert, falls $\int$ das Riemann-Integral ist, aber aus Beispiel 21.18 sehen wir, dass $\int 1_{\mathbb{Q}}(x)dx = 0$, falls $\int$ das Lebesgue-Integral ist.

Definition 22.13 (Das Integral von messbaren, nicht-negativen Funktionen). *Sei $(E, \mathcal{F}, \mu)$ ein Maßraum und $f : E \to \overline{\mathbb{R}}_+$ messbar. Das Integral von f bezüglich μ ist gegeben durch*

$$\mu[f] := \int f d\mu := \int f(x)\mu(dx) := \sup\{\mu[g] : g \text{ einfach, nicht-negativ, } g \leq f\}. \tag{22.2}$$

Bemerkung 22.14 (Das Integral als Fortsetzung). Nach Lemma 22.11 ist klar, dass die Definition von $\mu[f]$ für einfache, nicht-negative Funktionen aus Definition 22.9 und Definition 22.13 übereinstimmt. Damit ist obige Definition eine Fortsetzung von $\mu[f]$ auf den Raum der nicht-negativen, messbaren Funktionen.

Weiter ist wichtig zu sehen, dass nach Theorem 22.8 jede der in Definition 22.13 auftretenden Funktionen beliebig genau (punktweise) durch einfache Funktionen approximierbar ist. Insbesondere beinhaltet die Menge, deren Supremum in (22.2) gesucht wird, auch einfache Funktionen g, die beliebig dicht an f liegen.

Proposition 22.15 (Eigenschaften des Integrals). *Sei $(E, \mathcal{F}, \mu)$ ein Maßraum und $f, g, f_1, f_2, \dots : E \to \overline{\mathbb{R}}_+$ messbar. Dann gilt:*

1. *Für $f \leq g$ gilt $\mu[f] \leq \mu[g]$.*
2. *Es gilt monotone Konvergenz, d.h.*

$$f_n \uparrow f \qquad \Rightarrow \qquad \mu[f_n] \uparrow \mu[f].$$

3. *Für $a, b \geq 0$ gilt $\mu[af + bg] = a\mu[f] + b\mu[g]$.*

Beweis. 1. ist klar nach der Definition des Integrals. Für 2. ist zunächst, da $f_n \leq f$ für $n = 1, 2, \dots$

$$\lim_{n\to\infty} \mu[f_n] = \sup_{n\in\mathbb{N}} \mu[f_n] \leq \mu[f].$$

Für '$\geq$' genügt es zu zeigen, dass

$$\mu[g] \leq \sup_{n\in\mathbb{N}} \mu[f_n] \tag{22.3}$$

für alle einfachen Funktionen $g \leq f$. Sei also $g = \sum_{k=1}^m c_k 1_{A_k} \leq f$ für disjunkte Mengen $A_1, \dots, A_m$ und $c_1, \dots, c_m > 0$. Für $\varepsilon > 0$ und $n = 1, 2, \dots$ sei $B_n^\varepsilon := \{f_n \geq (1-\varepsilon)g\}$. Da $f_n \uparrow f$ und $g \leq f$ gilt $\bigcup_{n=1}^\infty B_n^\varepsilon = E$ für alle $\varepsilon > 0$. Also ist

$$\begin{aligned}\mu[f_n] \geq \mu[(1-\varepsilon)g1_{B_n^\varepsilon}] &= \sum_{k=1}^m (1-\varepsilon)c_k\mu(A_k \cap B_n^\varepsilon)\\ &\xrightarrow{n\to\infty} \sum_{k=1}^m (1-\varepsilon)c_k\mu(A_k) = (1-\varepsilon)\mu[g].\end{aligned}$$

Da $\varepsilon > 0$ beliebig war, folgt (22.3).
Für 3. seien $f_1, g_1, f_2, g_2, \dots$ einfache Funktionen mit $f_n \uparrow f$ und $g_n \uparrow g$. Dann ist $af_n + bg_n \uparrow af + bg$ und es folgt

$$\mu[af+bg] = \lim_{n\to\infty} \mu[af_n + bg_n] = \lim_{n\to\infty} a\mu[f_n] + b\mu[g_n] = a\mu[f] + b\mu[g]$$

aus 3. wegen Lemma 22.11. □

Wir können nun das Integral für messbare Funktionen definieren. Hierzu sei zunächst bemerkt, dass für $f: E \to \overline{\mathbb{R}}$ messbar gilt, dass $f^+, f^- \leq |f|$. Ist insbesondere $\mu[|f|] < \infty$, so ist auch $\mu[f^+], \mu[f^-] < \infty$.

Definition 22.16 (Integral messbarer Funktionen). *Sei $(E, \mathcal{F}, \mu)$ ein Maßraum und $f: E \to \overline{\mathbb{R}}$ messbar. Dann heißt f nach μ integrierbar, falls $\mu[|f|] < \infty$ und wir definieren*

$$\mu[f] := \int f(x)\mu(dx) := \int f d\mu := \mu[f^+] - \mu[f^-]. \tag{22.4}$$

Außerdem setzen wir

$$\mathcal{L}^1(\mu) := \Big\{f : E \to \overline{\mathbb{R}} : \mu[|f|] < \infty\Big\}$$

Für $A \in \mathcal{F}$ schreiben wir außerdem

$$\mu[f, A] := \int_A f d\mu := \mu[f 1_A].$$

Bemerkung 22.17 (Erweiterung des Integrals und $\mathcal{L}^p$-Räume). 1. Ist höchstens einer der beiden Terme $\mu[f^+]$ oder $\mu[f^-]$ unendlich, so definieren wir das Integral $\mu[f]$ weiterhin mittels (22.4). In anderen Fällen bleibt das Integral undefiniert.

2. Die Funktionenräume $\mathcal{L}^p(\mu) := \Big\{f : E \to \overline{\mathbb{R}} : \mu[|f|^p] < \infty\Big\}$, $p > 0$, werden in Abschnitt 23 eine besondere Rolle spielen.

22.3 Eigenschaften des Integrals

Wir stellen zunächst einige Eigenschaften des eingeführten Integralbegriffes fest. Diese sind etwa die Monotonie und Linearität. Weiter werden wir sehen, dass sich das Integral einer Funktion nicht ändert, wenn man sie auf einer Nullmenge abändert; siehe Proposition 22.20.

Proposition 22.18 (Einfache Eigenschaften des Integrals). *Sei $(E, \mathcal{F}, \mu)$ ein Maßraum und $f, g \in \mathcal{L}^1(\mu)$. Dann gilt:*

1. *Das Integral ist monoton, d.h.*
$$f \leq g \text{ fast überall} \qquad \Longrightarrow \qquad \mu[f] \leq \mu[g].$$
2. *Es gilt die Dreiecksungleichung*
$$|\mu[f]| \leq \mu[|f|].$$
3. *Das Integral ist linear, seien also $a, b \in \mathbb{R}$, dann ist $af + bg \in \mathcal{L}^1(\mu)$ und*
$$\mu[af + bg] = a\mu[f] + b\mu[g].$$

Beweis. Alle Eigenschaften folgen aus Proposition 22.15.1 und 3., sowie aus der Definition des Integrals für messbare Funktionen. □

Proposition 22.19 (Substitutionssatz). *Sei $(E, \mathcal{F}, \mu)$ ein Maßraum, $(E', \mathcal{F}')$ ein Messraum, $f : E \to E'$ messbar und $f_*\mu$ das Bildmaß von f aus Definition 21.23. Dann gilt für $g \in \mathcal{L}^1(f_*\mu)$, dass $g \circ f \in \mathcal{L}^1(\mu)$ und*
$$\mu[g \circ f] = f_*\mu[g].$$

Beweis. Es genügt, die Behauptung für einfache, nicht-negative Funktionen g zu zeigen. Der allgemeine Fall folgt dann mittels Approximation durch einfache Funktionen. Sei also $g = \sum_{k=1}^m c_k 1_{A'_k}$ mit $A'_k \in \mathcal{F}'$. Dann ist $g \circ f = \sum_{k=1}^m c_k 1_{f \in A'_k}$ und
$$\mu[g \circ f] = \sum_{k=1}^m c_k \mu(f \in A'_k) = \sum_{k=1}^m c_k f_*\mu(A'_k) = f_*\mu[g].$$
□

Proposition 22.20 (Integrale und Eigenschaften fast überall). *Sei $(E, \mathcal{F}, \mu)$ ein Maßraum und $f : E \to \overline{\mathbb{R}}_+$ messbar.*

1. *Es ist $f = 0$ fast überall genau dann, wenn $\mu[f] = 0$.*
2. *Falls $\mu[f] < \infty$, so ist $f < \infty$ fast überall.*

Beweis. 1. Sei $N := \{f > 0\} \in \mathcal{F}$.
'⇒': Es gilt $\mu(N) = 0$. Dann ist $f \leq \infty \cdot 1_N$, also gilt wegen Proposition 22.15.2
$$0 \leq \mu[f] \leq \mu[\infty, N] = \lim_{n \to \infty} \mu[n, N] = 0.$$

Für '⇐' sei $N_n := \{f \geq 1/n\}$ und damit $N_n \uparrow N$ und $nf \geq 1_{N_n}$, also

$$0 = \mu[f] \geq \tfrac{1}{n}\mu(N_n).$$

Damit ist $\mu(N_n) = 0$ und $\mu(N) = \mu(\bigcup_{n=1}^{\infty} N_n) = 0$.
Für 2. sei $A := \{f = \infty\}$. Wegen $f1_{f\geq n} \geq n1_{f\geq n}$ ist

$$\mu(A) = \mu[1_A] \leq \mu[1_{f\geq n}] \leq \tfrac{1}{n}\mu[f, 1_{f\geq n}] \leq \tfrac{1}{n}\mu[f] \xrightarrow{n\to\infty} 0.$$

Damit ist $\mu(f = \infty) = 0$, also $f < \infty$ fast überall; siehe Bemerkung 21.13. □

Zum Abschluss dieses Abschnittes stellen wir noch die Beziehung des (Lebesgue-)Integrals zum bekannten Riemann-Integral dar. Hierzu erinnern wir an Bemerkung 8.22, in der wir die Konstruktion des Riemann-Integrals mittels Ober- und Untersummen gegeben haben. Dies formalisieren wir nochmals mit Hilfe des Lebesgue-Integrals.

Definition 22.21 (Treppenfunktion und Riemann-Integral). *Eine Treppenfunktion ist eine Funktion $f : \mathbb{R} \to \mathbb{R}$, für die es eine Darstellung*

$$f(t) = \sum_{j=-\infty}^{\infty} a_j 1_{[t_{j-1}, t_j)}(t)$$

mit $t_{j-1} \leq t_j, j \in \mathbb{Z}$ gibt, wobei $a_j \in \mathbb{R}, j \in \mathbb{Z}$. Eine messbare Funktion $f : \mathbb{R} \to \mathbb{R}$ heißt (eigentlich) Riemann-integrierbar*, falls $\lambda[|f|] < \infty$ und es Treppen-Funktionen $f_1^+, f_1^-, f_2^+, f_2^-, \ldots$ gibt mit $f_n^- \leq f \leq f_n^+$ und $\lambda[f_n^+ - f_n^-] \xrightarrow{n\to\infty} 0$. Das Riemann-Integral von f ist dann durch $\lambda[f]$ definiert. (Insbesondere stimmen dann Riemann- und Lebesgue-Integral überein.*

Eine Funktion $f : \mathbb{R} \to \mathbb{R}$ heißt uneigentlich Riemann-integrierbar*, wenn $f1_K$ für alle kompakten Intervalle $K \subseteq \mathbb{R}$ eigentlich Riemann-integrierbar ist und $\lambda[f1_{[-n,n]}]$ konvergiert. Dieser Grenzwert ist dann das Riemann-Integral von f bezüglich λ.*

Proposition 22.22 (Riemann-Integrierbarkeit). *Die Funktion $f : \mathbb{R}_+ \to \mathbb{R}$ habe eine diskrete Menge von Sprungstellen. Dann ist f genau dann eigentlich Riemann-integrierbar, wenn f Lebesgue-integrierbar ist und es gilt*

$$\lambda[f] = \lim_{n\to\infty} \sum_{k=1}^{\infty} f(s_{n,k})(t_{n,k} - t_{n,k-1}) \tag{22.5}$$

für $0 = t_{n,0} \leq \ldots \leq t_{n,k_n} = t$ mit $\max_k |t_{n,k} - t_{n,k-1}| \xrightarrow{n\to\infty} 0$ und beliebiges $t_{n,k-1} \leq s_{n,k} \leq t_{n,k}$.

Beweis. Es genügt, die Behauptung für stetiges f zu zeigen. Andernfalls kann man f in die stetigen Teilabschnitte zerlegen. Weiter genügt es, die Behauptung für f mit kompaktem Träger K zu zeigen. Da f auf K gleichmäßig stetig ist, wählt man zunächst $\varepsilon_n \downarrow 0$ und $t_{n,0} \leq \ldots \leq t_{n,k_n}$ so, dass $K \subseteq [t_{n,0}, t_{n,k_n}]$ und $\max_{t_{n,k-1}\leq s<t_{n,k}} |f(t_{n,k-1}) - f(s)| < \varepsilon_n$. Nun ist es leicht, Treppenfuntionen f_n^+ und f_n^- zu finden, so dass $f_n^- \leq f \leq f_n^+$ und $||f_n^+ - f_n^-|| \leq \varepsilon_n$. Daraus folgt die erste Behauptung. Die Formel (22.5) gilt wegen der gleichmäßigen Approximation der Funktion f. □

Beispiel 22.23 (Unterschiede zwischen Riemann- und Lebesgue-Integral). 1. Wie wir bereits in Beispiel 22.12 gesehen haben, ist $f = 1_{\mathbb{Q}\cap[0,1]}$ nicht Riemann-integrierbar. Es ist nämlich $1_{[0,1]} \leq f^+$ für jede reguläre Treppenfunktion $f^+ \geq f$ und $f^- \leq 0$ für jede reguläre Treppenfunktion $f^- \leq f$. Damit können Unter- und Obersumme nicht überein und f ist nicht Riemann-integrierbar.

2. Wie aus der Definition 22.21 des Riemann-Integrals hervorgeht, ist jede eigentlich Riemann-integrierbare Funktion auch Lebesgue-integrierbar. Anders verhält es sich mit uneigentlich Riemann-integrierbaren Funktionen. Sei hierzu f gegeben durch $f(t) = \frac{(-1)^{\lceil t\rceil+1}}{\lceil t\rceil}$. Dann ist

$$\lambda[f1_{[0,2n]}] = \sum_{k=1}^{2n} \frac{(-1)^{k+1}}{k} = \sum_{k=1}^{n} \frac{1}{k} - \frac{1}{k+1} = 1 - \frac{1}{n+1} \xrightarrow{n\to\infty} 1.$$

Damit ist f uneigentlich Riemann-integrierbar. Allerdings gilt

$$\lambda[|f|] = \sum_{k=1}^{\infty} \frac{1}{k} = \infty.$$

Also ist f nicht Lebesgue-integrierbar.

22.4 Konvergenzsätze

Man kann sich fragen, ob es wirklich so wichtig ist, dass man mehr Funktionen bezüglich des Lebesgue-Integrals als bezüglich des Riemann Integrals integrieren kann. Es kommen schließlich in Anwendungen meistens Riemann-integrierbare Funktionen vor. Es gibt allerdings einen weiteren Vorteil des Lebesgue-Integrals, auf den wir nun eingehen werden. Es gelten nämlich Vertauschungsgesetze von Grenzwerten und Integralen (die hier Konvergenzsätze genannt werden), die relativ wenige Voraussetzungen benötigen. Die Wichtigsten sind der Satz von der monotonen und der von der majorisierten Konvergenz.

Theorem 22.24 (Monotone Konvergenz). *Sei $(E, \mathcal{F}, \mu)$ ein Maßraum, $f_1, f_2, \dots \in \mathcal{L}^1(\mu)$ und $f : E \to \overline{\mathbb{R}}$ messbar mit $f_n \uparrow f$ fast überall. Dann gilt*

$$\lim_{n\to\infty} \mu[f_n] = \mu[f],$$

wobei beide Seiten den Wert ∞ annehmen können.

Beweis. Sei $N \in \mathcal{F}$ so, dass $\mu(N) = 0$ und $f_n(x) \uparrow f(x)$ für $x \notin N$. Setze $g_n := (f_n - f_1)1_{N^c} \geq 0$. Damit ist $g_n \uparrow (f - f_1)1_{N^c} =: g$ und mit Proposition 22.18, Proposition 22.20 und Proposition 22.15.2 gilt

$$\mu[f_n] = \mu[f_1] + \mu[g_n] \xrightarrow{n\to\infty} \mu[f_1] + \mu[g] = \mu[f].$$

□

Theorem 22.25 (Lemma von Fatou). *Sei $(E, \mathcal{F}, \mu)$ ein Maßraum und $f_1, f_2, \dots : E \to \overline{\mathbb{R}}_+$ messbar. Dann gilt*

$$\liminf_{n\to\infty} \mu[f_n] \geq \mu[\liminf_{n\to\infty} f_n].$$

Beweis. Für alle $k \geq n$ gilt $f_k \geq \inf_{k \geq n} f_k$ und damit

$$\inf_{k \geq n} \mu[f_k] \geq \mu[\inf_{k \geq n} f_k]$$

wegen Proposition 22.15.1. Deshalb gilt mit $n \to \infty$

$$\liminf_{n \to \infty} \mu[f_n] \geq \sup_{n \in \mathbb{N}} \mu[\inf_{k \geq n} f_k] = \mu[\liminf_{n \to \infty} f_n]$$

wegen der monotonen Konvergenz, da $\inf_{k \geq n} f_k \uparrow \liminf_{n \to \infty} f_n$. □

Theorem 22.26 (Satz von der majorisierten Konvergenz). *Sei $(E, \mathcal{F}, \mu)$ ein Maßraum und $f, g, f_1, f_2, \dots : E \to \overline{\mathbb{R}}$ messbar mit $|f_n| \leq g$ fast überall, $\lim_{n \to \infty} f_n = f$ und $g \in \mathcal{L}^1(\mu)$. Dann gilt*

$$\lim_{n \to \infty} \mu[f_n] = \mu[f].$$

Beweis. Ohne Einschränkung gilt $|f_n| \leq g$ überall. Wir verwenden das Lemma von Fatou, sowie $g - f_n, g + f \geq 0$, also

$$\begin{aligned}
\mu[g + f] &\leq \liminf_{n \to \infty} \mu[g + f_n] = \mu[g] + \liminf_{n \to \infty} \mu[f_n], \\
\mu[g - f] &\leq \liminf_{n \to \infty} \mu[g - f_n] = \mu[g] - \limsup_{n \to \infty} \mu[f_n].
\end{aligned}$$

Nach Subtrahieren von $\mu[g]$ ist also

$$\mu[f] \leq \liminf_{n \to \infty} \mu[f_n] \leq \limsup_{n \to \infty} \mu[f_n] \leq \mu[f].$$

□

Beispiel 22.27 (Beispiele zu den Konvergenzsätzen). 1. Im Lemma von Fatou ist nicht vorausgesetzt, dass eines der f_n integrierbar ist. Wir geben nun ein Beispiel dafür, dass im Lemma von Fatou wirklich '$<$' und nicht '$=$' gilt. Sei hierzu λ das Lebesgue-Maß und $f_n = 1/n$ (also insbesondere f_n konstant), $n = 1, 2, \dots$. Dann gilt $f_n \downarrow 0$, also

$$\liminf_{n \to \infty} \mu[f_n] = \infty > 0 = \mu[0] = \mu[\liminf_{n \to \infty} f_n].$$

2. Im Satz von der majorisierten Konvergenz kann auf die Voraussetzung, dass $|f_n| \leq g$ und $g \in \mathcal{L}^1(\mu)$ nicht verzichtet werden. Sei etwa λ das Lebesgue-Maß auf $[0, 1]$ und $f_n = n \cdot 1_{[0,1/n]}$. Dann ist $\sup_{n \in \mathbb{N}} f_n(x) = \sup\{n : x \leq 1/n\} = \left[\frac{1}{x}\right]$. Also existiert kein $g \in \mathcal{L}^1(\lambda)$ mit $f_n \leq g$. Außerdem ist $\lim_{n \to \infty} f_n = 0$ fast überall (wobei $\{0\}$ die Ausnahmemenge ist) und

$$\lim_{n \to \infty} \mu[f_n] = 1 \neq 0 = \mu[\lim_{n \to \infty} f_n].$$

Anders verhält es sich für $f_n = n \cdot 1_{[0,1/n^2]}$. Hier ist

$$\sup_{n \in \mathbb{N}} f_n(x) = \sup\{n : x \leq 1/n^2\} = \left[\frac{1}{\sqrt{x}}\right] \leq \frac{1}{\sqrt{x}} =: g(x).$$

Hier ist einerseits $g \in \mathcal{L}^1(\lambda)$, der Satz von der majorisierten Konvergenz ist also anwendbar, andererseits ist $\lim_{n \to \infty} f_n = 0$ fast überall und

$$\lim_{n \to \infty} \mu[f_n] = \lim_{n \to \infty} \tfrac{1}{n} = 0 = \mu[0] = \mu[\lim_{n \to \infty} f_n].$$

23 $\mathcal{L}^p$-Räume

Im ganzen folgenden Abschnitt sei $(E, \mathcal{F}, \mu)$ ein Maßraum. Wir werden uns nun mit den messbaren Funktionen $f : E \to \overline{\mathbb{R}}$ beschäftigen, für die $\mu[|f|^p] < \infty$ gilt. Die resultierenden Funktionenräume $\mathcal{L}^p(\mu)$ werden wir als normierte, vollständige Räume (Proposition 23.7) erkennen, was automatisch zu einem neuen Konvergenzbegriff führt. Weiter wird der Raum $\mathcal{L}^2(\mu)$ eine große Rolle spielen. Da man diesen mit einem Skalarprodukt (nämlich $\langle f, g\rangle \mapsto \mu[fg]$) ausstatten kann, stehen hier allgemeine Aussagen zur Verfügung, etwa der Satz von Riesz-Fréchet (Proposition 23.10). Dies werden wir verwenden, um σ-endliche Maße mit Dichte durch den Satz von Radon-Nikodým (Korollar 23.16) zu charakterisieren.

23.1 Grundlagen

Bereits in Bemerkung 22.17 haben wir die Räume $\mathcal{L}^p(\mu)$ erwähnt. Durch die Definition des Integrals im letzten Abschnitt können wir diese nun näher beleuchten. Vor allem zeigen wir die wichtige Hölder- und die Minkowski-Ungleichung, siehe Proposition 23.2.

Definition 23.1 ($\mathcal{L}^p(\mu)$-Räume). *Sei $0 < p \leq \infty$. Wir setzen*

$$\mathcal{L}^p := \mathcal{L}^p(\mu) := \{f : E \to \overline{\mathbb{R}} \text{ messbar mit } ||f||_p < \infty\}$$

für

$$||f||_p := (\mu[|f|^p])^{1/p}, \qquad 0 < p < \infty \tag{23.1}$$

und

$$||f||_\infty := \inf\{K : \mu(|f| > K) = 0\}.$$

Auf den Räumen $\mathcal{L}^p$, $p \geq 1$ zeigen wir nun eine Dreiecksungleichung, die Minkowski-Ungleichung gezeigt. Außerdem sei bemerkt, dass die Hölder-Ungleichung im Spezialfall $p = q = 2$ auch Cauchy-Schwartz-Ungleichung heißt.

Proposition 23.2 (Hölder und Minkowski's Ungleichung). *Seien f, g messbar.*

1. *Sei $0 < p, q, r \leq \infty$ so, dass $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$. Dann gilt*

$$||fg||_r \leq ||f||_p||g||_q \qquad \text{(Hölder-Ungleichung)}$$

2. *Für $1 \leq p \leq \infty$ ist*

$$||f + g||_p \leq ||f||_p + ||g||_p. \qquad \text{(Minkowski-Ungleichung)}$$

Beweis. Wir starten mit dem Beweis der Hölder-Ungleichung. Im Fall $p = \infty$ oder $q = \infty$ ist die Aussage klar, sei also $p, q < \infty$. Ist entweder $||f||_p = 0$, $||f||_p = \infty$, $||g||_q = 0$ oder $||g||_q = \infty$, ist die Aussage ebenso klar. Sei also o.E. $f, g \geq 0$ und $0 < ||f||_p, ||g||_q < \infty$ und

$$\widetilde{f} := \frac{f}{||f||_p}, \qquad \widetilde{g} = \frac{g}{||g||_q}.$$

Dann ist zu zeigen, dass $||\widetilde{f}\widetilde{g}||_r \leq 1$. Wegen der Konvexität der Exponentialfunktion ist

$$(xy)^r = \exp\left(\tfrac{r}{p} p \log x + \tfrac{r}{q} q \log y\right) \leq \tfrac{r}{p} x^p + \tfrac{r}{q} y^q$$

und damit
$$||\widetilde{f}\widetilde{g}||_r^r = \mu[(\widetilde{f}\widetilde{g})^r] \leq \tfrac{r}{p}\mu[\widetilde{f}^p] + \tfrac{r}{q}\mu[\widetilde{g}^q] = 1$$
und die Behauptung folgt.

Zum Beweis der Minkowski-Ungleichung bemerken wir zunächst, dass in den Fällen $p = 1$ und $p = \infty$ die Behauptung klar ist. Im Falle $1 < p < \infty$ gilt mit $q = p/(p-1)$ und $r = 1/p + 1/q = 1$ mit der Hölder-Ungleichung

$$\begin{aligned}
||f+g||_p^p &\leq \mu[|f| \cdot |f+g|^{p-1}] + \mu[|g| \cdot |f+g|^{p-1}] \\
&\leq ||f||_p \cdot ||(f+g)^{p-1}||_q + ||g||_p \cdot ||(f+g)^{p-1}||_q \\
&= (||f||_p + ||g||_p) \cdot ||f+g||_p^{p-1},
\end{aligned}$$

da $||(f+g)^{p-1}||_q = ||(f+g)^{q(p-1)}||_1^{1/q} = ||(f+g)^p||_1^{(p-1)/p} = ||f+g||_p^{p-1}$. Dividieren durch $||f+g||_p^{p-1}$ bringt das Resultat. □

Proposition 23.3 (Zusammenhang zwischen $\mathcal{L}^r$ und$\mathcal{L}^q$). *Sei μ endlich und $1 \leq r < q \leq \infty$. Dann ist $\mathcal{L}^q(\mu) \subseteq \mathcal{L}^r(\mu)$.*

Beweis. Die Behauptung ist klar für $q = \infty$. Sei also $q < \infty$. Wir verwenden die Hölder-Ungleichung. Es gilt für $f \in \mathcal{L}^q$, da $||1||_p < \infty$ wegen der Endlichkeit von μ,

$$||f||_r = ||1 \cdot f||_r \leq ||1||_p \cdot ||f||_q < \infty \tag{23.2}$$

für $\frac{1}{p} = \frac{1}{r} - \frac{1}{q} > 0$, woraus die Behauptung sofort folgt. □

Bemerkung 23.4 (Gegenbeispiel für $\boldsymbol{\sigma}$-endliches $\boldsymbol{\mu}$). Sicherlich gilt Proposition 23.3 nicht, falls μ nicht endlich ist. Sei etwa λ das eindimensionale Lebesgue-Maß und $f : x \mapsto \frac{1}{x} \cdot 1_{x>1}$. Dann ist $f \in \mathcal{L}^2(\lambda)$, aber $f \notin \mathcal{L}^1(\lambda)$.

23.2 $\mathcal{L}^p$-Konvergenz

Wir haben im Satz von der majorisierten Konvergenz (Theorem 22.26) gesehen, dass für eine Folge von Funktionen, die fast überall konvergiert, oft auch deren Integrale konvergieren. Die hier betrachtete $\mathcal{L}^p$-Konvergenz geht jetzt von Konvergenz von Integralen aus. Wir werden sehen, dass der resultierende Konvergenzbegriff zur Folge hat, dass jede Cauchy-Folge konvergiert (Proposition 23.7).

Definition 23.5 (Konvergenz im $\boldsymbol{p}$-ten Mittel). *Eine Folge $f_1, f_2, \dots$ in $\mathcal{L}^p(\mu)$* konvergiert in $\mathcal{L}^p(\mu)$ *(oder* im p-ten Mittel*) gegen $f \in \mathcal{L}^p(\mu)$, falls*

$$||f_n - f||_p \xrightarrow{n\to\infty} 0.$$

Wir schreiben dann auch $f_n \xrightarrow{n\to\infty}_{\mathcal{L}^p} f$.

Proposition 23.6 (Konvergenz im $\boldsymbol{p}$-ten und $\boldsymbol{q}$-ten Mittel). *Sei $\mu(E) < \infty$, $1 \leq r < q \leq \infty$ und $f, f_1, f_2, \dots \in \mathcal{L}^q$. Falls $f_n \xrightarrow{n\to\infty}_{\mathcal{L}^q} f$, so auch $f_n \xrightarrow{n\to\infty}_{\mathcal{L}^r} f$.*

Beweis. Die Behauptung ist klar für $q = \infty$, sei also $q < \infty$. Aus (23.2) folgt sofort dass $||f - g||_r \leq ||f - g||_q$, woraus die Behauptung bereits folgt. □

Proposition 23.7 (Vollständigkeit von $\mathcal{L}^p$). *Sei $p \geq 1$ und $f_1, f_2, \ldots$ eine Cauchy-Folge in $\mathcal{L}^p$. Dann gibt es ein $f \in \mathcal{L}^p$ mit $||f_n - f||_p \xrightarrow{n\to\infty} 0$.*

Beweis. Sei $\varepsilon_1, \varepsilon_2, \ldots$ summierbar, also z.B. $\varepsilon_n := 2^{-n}$. Da $f_1, f_2, \ldots$ eine Cauchy-Folge ist, gibt es für jedes k einen Index n_k mit $||f_m - f_n||_p \leq \varepsilon_k$ für alle $m, n \geq n_k$. Insbesondere gilt

$$\sum_{k=1}^{\infty} ||f_{n_{k+1}} - f_{n_k}||_p \leq \sum_{k=1}^{\infty} \varepsilon_k < \infty.$$

Nach monotoner Konvergenz und der Minkowski'schen Ungleichung gilt damit

$$\Big|\Big| \sum_{k=1}^{\infty} |f_{n_{k+1}} - f_{n_k}| \Big|\Big|_p \leq \sum_{k=1}^{\infty} ||f_{n_{k+1}} - f_{n_k}||_p < \infty.$$

Damit ist insbesondere $\sum_{k=1}^{\infty} |f_{n_{k+1}} - f_{n_k}| < \infty$ fast überall, also $f_{n_1}(x), f_{n_2}(x), \ldots$ für fast alle x eine Cauchy-Folge in $\mathbb{R}$. Damit gibt es eine messbare Abbildung f mit $f_{n_k} \xrightarrow{k\to\infty} f$ fast überall. Nach Fatou's Lemma gilt

$$||f_n - f||_p \leq \liminf_{k\to\infty} ||f_{n_k} - f_n||_p \leq \sup_{m \geq n} ||f_m - f_n||_p \xrightarrow{n\to\infty} 0,$$

d.h. $f_n \xrightarrow{n\to\infty}_{\mathcal{L}^p} f$. □

23.3 Der Raum $\mathcal{L}^2$

Für jedes $1 \leq p \leq \infty$ gilt, dass $||af||_p = |a| \cdot ||f||_p$ für $a \in \mathbb{R}$ gilt. Zusammen mit der Minkowski'schen Ungleichung bedeutet das, dass $\mathcal{L}^p(\mu)$ ein reeller Vektorraum ist. Es ist entscheidend zu bemerken, dass die Abbildung $f \mapsto ||f||_p$ eine Pseudo-Norm ist. Sie kann keine volle Norm sein, weil $||f||_p = 0$ nach Proposition 22.20 nur impliziert, dass $\mu(f \neq 0) = 0$, aber nicht, dass $f = 0$. Wir werden deswegen im Folgenden Funktionen f und g miteinander identifizieren, wenn $f = g$ μ-fast überall gilt. Nach dem eben gesagtem ist dann $(\mathcal{L}^p, ||\cdot||_p)$ ein normierter Raum. Da $||\cdot||_p$ nach Proposition 23.7 vollständig ist, ist $(\mathcal{L}^p, ||\cdot||_p)$ sogar ein Banachraum für jedes $1 \leq p \leq \infty$. Als nächstes betrachten wir den Spezialfall $p = 2$. Wir definieren eine Abbildung $\mathcal{L}^2 \times \mathcal{L}^2 \to \mathbb{R}$ durch

$$\langle f, g \rangle := \mu[fg].$$

Dann ist $\langle \cdot, \cdot \rangle$ offensichtlich linear, symmetrisch und positiv semi-definit, also ein Skalarprodukt Wir schreiben konsequenterweise $f \perp g$ genau dann, wenn $\mu[fg] = 0$. Da $||f|| := ||f||_2 = \langle f, f \rangle^{1/2}$, ist $(\mathcal{L}^2, \langle \cdot, \cdot \rangle)$ also ein Hilbertraum.

Lemma 23.8 (Parallelogrammidentität). *Seien $f, g \in \mathcal{L}^2$. Dann gilt*

$$||f + g||^2 + ||f - g||^2 = 2||f||^2 + 2||g||^2.$$

Beweis. Aus der Definition von $||\cdot||$ und der Symmetrie und Bilinearität von $\langle \cdot, \cdot \rangle$ folgt

$$||f + g||^2 + ||f - g||^2 = \langle f + g, f + g \rangle + \langle f - g, f - g \rangle = 2\langle f, f \rangle + 2\langle g, g \rangle = 2||f||^2 + 2||g||^2.$$

□

Proposition 23.9 (Zerlegung von $f \in \mathcal{L}^2$). *Sei M ein abgeschlossener, linearer Teilraum von $\mathcal{L}^2$. Dann hat jede Funktion $f \in \mathcal{L}^2$ eine fast sicher eindeutige Zerlegung $f = g + h$ mit $g \in M, h \perp M$.*

Beweis. Für $f \in \mathcal{L}^2$ definieren wir

$$d_f := \inf_{g \in M} \{||f - g||\}.$$

Wähle $g_1, g_2, \ldots$ mit $||f - g_n|| \xrightarrow{n\to\infty} d_f$. Nach der Parallelogrammidentität gilt

$$4d_f^2 + ||g_m - g_n||^2 \leq ||2f - g_m - g_n||^2 + ||g_m - g_n||^2 = 2||f - g_m||^2 + 2||f - g_n||^2 \xrightarrow{m,n\to\infty} 4d_f^2.$$

Also ist $||g_m - g_n||^2 \xrightarrow{m,n\to\infty} 0$, d.h. $g_1, g_2, \ldots$ ist eine Cauchy-Folge. Nach Proposition 23.7 gibt es damit ein $g \in \mathcal{L}^2$ mit $||g_n - g||_p \xrightarrow{n\to\infty} 0$, und wegen der Abgeschlossenheit von M auch $g \in M$. Da $||h|| = d_f$ für $h := f - g$, folgt für alle $t > 0, l \in M$, wegen der Definition von d_f,

$$d_f^2 \leq ||h + tl||^2 = d_f^2 + 2t\langle h, l\rangle + t^2||l||^2.$$

Da dies für alle t gilt, folgt $\langle h, l\rangle = 0$, also $h \perp M$.

Zum Beweis der Eindeutigkeit sei $g' + h'$ eine weitere Zerlegung von f. Dann ist aufgrund der Linearität von M einerseits $g - g' \in M$, andererseits ist fast sicher $g - g' = h - h' \perp M$, also $g - g' \perp g - g'$. Dies bedeutet $||g - g'|| = \langle g - g', g - g'\rangle = 0$, also $g = g'$ fast überall. □

Proposition 23.10 (Riesz-Fréchet). *Eine Abbildung $F : \mathcal{L}^2 \to \mathbb{R}$ ist genau dann stetig und linear wenn es ein $h \in \mathcal{L}^2$ gibt, so dass für alle $f \in \mathcal{L}^2$*

$$F(f) = \langle f, h\rangle.$$

Dann ist $h \in \mathcal{L}^2$ fast sicher eindeutig bestimmt.

Beweis. '$\Leftarrow$': Die Linearität von $f \mapsto \langle f, h\rangle$ folgt aus der Bilinearität von $\langle\cdot,\cdot\rangle$. Die Stetigkeit folgt aus der Cauchy-Schwartz Ungleichung mittels

$$|\langle |f - f'|, h\rangle| \leq ||f - f'|| \cdot ||h||.$$

'$\Rightarrow$': Falls $F \equiv 0$ wähle $h = 0$. Falls $F \not\equiv 0$, ist $M = F^{-1}\{0\}$ ein (wegen der Stetigkeit von F) abgeschlossener und (wegen der Linearität von F) linearer Unterraum von $\mathcal{L}^2$. Wähle $f' \in \mathcal{L}^2 \setminus M$ mit der (nach Proposition 23.9 fast sicher eindeutigen) orthogonalen Zerlegung $f' = g' + h'$ mit $g' \in M$ und $h' \perp M$. Da $f' \notin M$, ist $h' \not\equiv 0$ und $F(h') = F(f') - F(g') = F(f') \neq 0$. Wir setzen $h'' = \frac{h'}{F(h')}$, so dass $h'' \perp M$ und $F(h'') = 1$ und es gilt für alle $f \in \mathcal{L}^2$

$$F(f - F(f)h'') = F(f) - F(f)F(h'') = 0.$$

d.h. $f - F(f)h'' \in M$, also insbesondere $\langle F(f)h'', h''\rangle = \langle f, h''\rangle$ und

$$F(f) = \tfrac{1}{||h''||^2} \cdot \langle F(f)h'', h''\rangle = \tfrac{1}{||h''||^2} \cdot \langle f, h''\rangle = \langle f, \tfrac{h''}{||h''||^2}\rangle.$$

Nun folgt die Behauptung mit $h = \frac{h''}{||h''||^2}$.

Zur Eindeutigkeit sei $\langle f, h_1 - h_2\rangle = 0$ für alle $f \in \mathcal{L}^2$; insbesondere ist mit $f = h_1 - h_2$

$$||h_1 - h_2||^2 = \langle h_1 - h_2, h_1 - h_2\rangle = 0,$$

also $h_1 = h_2$ μ-fast sicher. □

Bemerkung 23.11 (Allgemeingültigkeit der letzten Aussagen). Lemma 23.8, sowie die Propositionen 23.9 und 23.10 gelten ebenso, falls man $\mathcal{L}^2$ durch einen anderen Hilbert-Raum ersetzt.

23.4 Satz von Radon-Nikodým

Aus Beispiel 21.22 sind bereits zwei Maße mit Dichten (bezüglich des Lebesgue-Maßes) bekannt. Dieses Konzept wird nun aufgegriffen und in den Kontext von Integralen eingebettet. Sei hierzu ν ein weiteres Maß auf $\mathcal{F}$. Ziel ist es, Bedingungen anzugeben, wann das Maß ν durch eine Dichte darstellbar ist. Die Antwort findet sich im Satz von Radon-Nikodým (Korollar 23.16). Er ist ein Spezialfall des Lebesgue'schen Zerlegungssatzes, Theorem 23.15. Dieser zeigt, dass für je zwei σ-endliche Maße μ, ν das Maß ν (additiv) in zwei Anteile zerlegt werden kann: einen absolutstetigen bzgl. μ und einen zu μ singulären. Der absolutstetige Anteil hat dabei eine Dichte bzgl. μ. Zunächst müssen wir alle Begriffe erklären.

Definition 23.12 (Absolutstetige Maße). 1. *Wir sagen, ν besitzt eine* Dichte f *bzgl. μ, falls für alle $A \in \mathcal{F}$*

$$\nu(A) = \mu[f; A].$$

gilt. Wir schreiben dann $f = \frac{d\nu}{d\mu}$ und $\nu = f \cdot \mu$.

2. *Das Maß ν heißt* absolutstetig *bzgl. μ, falls alle μ-Nullmengen auch ν-Nullmengen sind. Wir schreiben dann $\nu \ll \mu$. Ist sowohl $\nu \ll \mu$ als auch $\mu \ll \nu$, so heißen μ und ν* äquivalent.

3. *Die Maße μ und ν heißen* singulär, *falls es ein $A \in \mathcal{F}$ gibt mit $\mu(A) = 0$ und $\nu(A^c) = 0$. Wir schreiben dann $\mu \perp \nu$.*

Lemma 23.13 (Kettenregel und Eindeutigkeit). *Sei μ ein Maß auf $\mathcal{F}$.*

1. *Sei ν ein σ-endliches Maß. Sind g_1 und g_2 Dichten von ν bzgl. μ, so ist $g_1 = g_2$ μ-fast überall.*

2. *Sei $f : E \to \mathbb{R}_+$ und $g : E \to \mathbb{R}$ messbar. Dann gilt*

$$(f \cdot \mu)[g] = \mu[fg],$$

falls eine der beiden Seiten existiert.

Beweis. 1. Sei $E_1, E_2, \dots \in \mathcal{F}$ so, dass $E_n \uparrow E$ und $\nu(E_n) < \infty$. Setze $A_n := E_n \cap \{g_1 > g_2\}$. Da sowohl g_1 als auch g_2 Dichten von ν bzgl μ sind, folgt

$$\mu[g_1 - g_2; A_n] = 0.$$

Da auf A_n nur $g_1 > g_2$ möglich ist, ist $g_1 = g_2$ $1_{A_n}\mu$-fast überall. Außerdem ist

$$\mu\{g_1 > g_2\} = \mu\Big(\bigcup_{n \in \mathbb{N}} A_n\Big) = 0.$$

Analog folgt $\mu\{g_1 < g_2\} = 0$ und damit $g_1 = g_2$ μ-fast überall.

2. gilt per Definition für $g = 1_A$ mit $A \in \mathcal{F}$. Damit erweitert man die Aussage schrittweise für einfach Funktionen, positive messbare Funktionen und schlussendlich auf den allgemeinen Fall. □

Beispiel 23.14 (Bekannte Dichten). 1. Aus der Vorlesung *Stochastik* sind bereits einige Dichtefunktionen bekannt. Sei etwa für $\mu \in \mathbb{R}, \sigma^2 \in \mathbb{R}_+$

$$f_{N(\mu,\sigma^2)}(x) := \frac{1}{\sqrt{2\pi\sigma^2}} \exp\Big(- \frac{(x-\mu)^2}{2\sigma^2}\Big)$$

und λ das eindimensionale Lebesgue-Maß. Dann heißt das Wahrscheinlichkeitsmaß $f_{N(\mu,\sigma^2)} \cdot \lambda$ auch *Normalverteilung mit Erwartungswert μ und Varianz σ^2*.
Für $\gamma \geq 0$ und

$$f_{\exp(\gamma)}(x) := 1_{x \geq 0} \cdot \gamma e^{-\gamma x}$$

heißt $f_{\exp(\gamma)} \cdot \lambda$ auch *Exponentialverteilung zum Parameter γ*. Man kann nun beispielsweise mit Lemma 23.13 berechnen, dass

$$f_{\exp(\gamma)} \cdot \lambda[\mathrm{id}] = \int_0^\infty \gamma e^{-\gamma x} x dx = -e^{-\gamma x} x\big|_0^\infty + \int_0^\infty e^{-\gamma x} dx = \frac{1}{\gamma}.$$

Den Wert $1/\gamma$ haben wir bereits als Erwartungswert der Exponentialverteilung zum Parameter γ interpretiert.

2. Es gibt natürlich nicht nur Dichten bezüglich des Lebesgue-Maßes. Sei beispielsweise

$$\mu = \sum_{n=0}^\infty \delta_n$$

das Zählmaß auf $\mathbb{N}_0$ (siehe Beispiel 21.2) und $f : \mathbb{N}_0 \to \mathbb{R}_+$, gegeben für ein $\gamma \geq 0$ durch

$$f(k) = e^{-\gamma} \frac{\gamma^k}{k!}.$$

Dann ist $f \cdot \mu$ die Poisson-Verteilung zum Parameter γ auf $2^{\mathbb{N}_0}$ nach Beispiel 21.2.

Theorem 23.15 (Lebesgue'scher Zerlegungssatz). *Seien μ, ν σ-endliche Maße auf $(E, \mathcal{F})$. Dann lässt sich ν eindeutig schreiben als*

$$\nu = \nu_a + \nu_s \qquad \mathit{mit} \qquad \nu_a \ll \mu, \nu_s \perp \mu.$$

Das Maße ν_a hat eine Dichte bzgl. μ, die μ-fast überall endlich ist.

Beweis. Durch eine Ausschöpfung $E_1, E_2, \dots \subseteq E$ mit $E_n \uparrow E$ und $\nu(E_n), \mu(E_n) < \infty$ können wir uns auf den Fall endlicher Maße zurückziehen. Mit Proposition 23.6 ist die lineare Abbildung

$$\begin{cases} \mathcal{L}^2(\mu+\nu) & \to \mathbb{R}) \\ f & \mapsto \nu[f] \end{cases}$$

stetig. Nach Proposition 23.10 gibt es also ein $h \in \mathcal{L}^2(\mu + \nu)$ mit

$$\nu[f] = (\mu+\nu)[fh], \tag{23.3}$$

also

$$\nu[f(1-h)] = \mu[fh] \tag{23.4}$$

für jedes $f \in \mathcal{L}^2(\mu+\nu)$. Wählt man $f = 1_{\{h<0\}}$ in (23.3) folgt

$$0 \le \nu\{h<0\} = (\mu+\nu)[h; h<0] \le 0,$$

also $h \ge 0$ $(\mu+\nu)$-fast überall. Analog kann man mittels $f = 1_{\{h>1\}}$ aus (23.4) folgern, dass

$$0 \le \mu[h; \{h>1\}] = \nu[1-h; \{h>1\} \le 0,$$

also $h \le 1$ $(\mu+\nu)$-fast sicher. Sei nun $f \ge 0$ messbar und $f_1, f_2, \dots \in \mathcal{L}^2(\mu+\nu)$ mit $f_n \uparrow f$. Mit dem Satz von der monotonen Konvergenz folgt

$$\nu[f(1-h)] = \lim_{n\to\infty} \nu[f_n(1-h)] = \lim_{n\to\infty} \mu[f_N h] = \mu[fh],$$

d.h. (23.4) gilt für alle messbaren $f \ge 0$.

Sei nun $E := h^{-1}\{1\}$. Aus (23.4) folgt mit $f = 1_E$, dass

$$\mu(E) = \mu[h; E] = \nu[1-h; E] = 0.$$

Wir definieren für $A \in \mathcal{F}$ zwei Maße ν_a und ν_s durch

$$\nu_a(A) = \nu(A \setminus E), \qquad \nu_s(A) = \nu(A \cap E),$$

so dass $\nu = \nu_a + \nu_s$ und $\nu_s \perp \mu$. Um zu zeigen, dass $\nu_a \ll \mu$ wähle $A \in \mathcal{F}$ mit $\mu(A) = 0$. Damit ist nach (23.4)

$$\nu[1-h; A \setminus E] = \mu[h; A \setminus E] = 0.$$

Da $h < 1$ auf $A \setminus E$ ist, gilt damit $\nu_a(A) = \nu(A \setminus E) = 0$, also $\nu_a \ll \mu$.

Um die Dichte von ν_a bzgl. μ zu bestimmen, setzen wir $g := \frac{h}{1-h} 1_{E \setminus E}$ und verwenden (23.4), so dass

$$\mu[g; A] = \mu\Big[\frac{h}{1-h}; A \setminus E\Big] = \nu(A \setminus E) = \nu_a(A).$$

Also ist $g = \frac{d\nu_a}{d\mu}$.

Um die Eindeutigkeit der Zerlegung zu zeigen, sei $\nu = \nu_a + \nu_s = \widetilde{\nu}_a + \widetilde{\nu}_s$ für $\nu_a, \widetilde{\nu}_a \ll \mu$, $\nu_s, \widetilde{\nu}_s \perp \mu$. Wähle $A, \widetilde{A} \in \mathcal{A}$ mit $\nu_s(A) = \mu(A^c) = \widetilde{\nu}_s(\widetilde{A}) = \mu(\widetilde{A}^c) = 0$. Dann gilt

$$\nu_s(A \cap \widetilde{A}) = \widetilde{\nu}_s(A \cap \widetilde{A}) = \nu_a(A^c \cup \widetilde{A}^c) = \widetilde{\nu}_a(A^c \cup \widetilde{A}^c) = 0$$

und deshalb

$$\begin{aligned}
\nu_a &= 1_{A \cap \widetilde{A}} \cdot \nu_a = 1_{A \cap \widetilde{A}} \cdot \nu = 1_{A \cap \widetilde{A}} \cdot \widetilde{\nu}_a = \widetilde{\nu}_a, \\
\nu_s &= \nu - \nu_a = \nu - \widetilde{\nu}_a = \widetilde{\nu}_s.
\end{aligned}$$

□

Korollar 23.16 (Satz von Radon-Nikodým). *Seien μ und ν σ-endliche Maße. Dann hat ν genau dann eine Dichte bzgl. μ, wenn $\nu \ll \mu$.*

Beweis. '⇒': klar.
'⇐': Nach Theorem 23.15 gibt es eine eindeutige Zerlegung $\nu = \nu_a + \nu_s$ mit $\nu_a \ll \mu, \nu_s \perp \mu$. Da $\nu \ll \mu$, muss $\nu_s = 0$ gelten und damit $\nu = \nu_a$. Insbesondere existiert die Dichte von ν bzgl. μ. □

Beispiel 23.17 (Gegenbeispiel für nicht σ-endliche Maße). Im Zerlegungssatz von Lebesgue 23.15 und im Satz von Radon-Nikodým 23.16 darf man die Voraussetzung, dass μ und ν σ-endlich sind nicht weglassen, wie folgendes Beispiel zeigt:

Sei $(E, \mathcal{F})$ ein Messraum mit überabzählbarem E und

$$\mathcal{F} := \{A : A \text{ oder } A^c \text{ abzählbar}\}.$$

Seien μ und ν unendliche Maße auf $(E, \mathcal{F})$, gegeben durch

$$\nu(A) := \begin{cases} 0, & A \text{ abzählbar,} \\ \infty, & \text{sonst,} \end{cases} \qquad \mu(A) := \begin{cases} |A|, & A \text{ endlich,} \\ \infty, & \text{sonst.} \end{cases}$$

Dann ist offenbar $\nu \ll \mu$. Angenommen, es gäbe eine $\mathcal{F}$-messbare Dichte von ν bzgl. μ, so wäre für alle $x \in E$

$$0 = \nu\{x\} = \mu[f; \{x\}] = f(x)\mu(\{x\}) = f(x).$$

Damit wäre $f = 0$ und $\nu = 0$ im Widerspruch zur Definition von ν.

24 Produkträume

Sei $(E_i)_{i \in I}$ eine Familie von Mengen. Dann heißt

$$E := \bigtimes_{i \in I} E_i := \{(x_i)_{i \in I} : x_i \in E_i\}$$

Produktraum der $(E_i)_{i \in I}$. Weiter definieren wir für $H \subseteq J \subseteq I$ die Projektionen

$$\pi_H^J : \bigtimes_{i \in J} E_i \to \bigtimes_{i \in H} E_i.$$

sowie $\pi_H := \pi_H^I$ und $\pi_i := \pi_{\{i\}}$, $i \in I$. In diesem Kapitel werden wir alle bisher eingeführten Konzepte auf solche Produkträume anwenden. Dabei werden wir meistens nur endliche Produkträume betrachten (d.h. I ist endlich). Unendliche (sogar überabzählbar unendliche Produkträume) werden in der Vorlesung *Stochastische Prozesse* eine wichtige Rolle spielen.

24.1 Topologie

Wir beginnen mit der Definition der Topologie auf Produkträumen. Diese ist gerade so gemacht, dass Projektionen stetige Funktionen sind.

Definition 24.1 (Produktraum und Produkt-Topologie). *Ist $(E_i, \mathcal{O}_i)_{i \in I}$ eine Familie von topologischen Räumen, dann heißt die von*

$$\mathcal{B} := \{\bigtimes_{i \in J} A_i \times \bigtimes_{i \in I \setminus J} E_i : J \Subset I, A_i \in \mathcal{O}_i\}$$

erzeugte Topologie (siehe Definition 19.1.7) $\mathcal{O}$ die Produkttopologie *auf E.*

Bemerkung 24.2 (Stetigkeit der Koordinatenabbildungen). Es sei bemerkt, dass bezüglich der Produkttopologie $\mathcal{O}$ alle Projektionen $\pi_i, i \in I$ stetig sind. Es ist nämlich

$$\pi_i^{-1}(A_i) = A_i \times \bigtimes_{j \neq i} E_j \in \mathcal{O}$$

für $A_i \in \mathcal{O}_i$. Damit ist die Projektion stetig (siehe Definition 19.1.9).

Proposition 24.3 (Abzählbare Produkte polnischer Räume sind polnisch). *Sei $(E_i, \mathcal{O}_i)_{i\in\mathbb{N}}$ eine Familie polnischer Räume. Dann ist der Produktraum $(E, \mathcal{O})$ aus Definition 24.1 polnisch.*

Beweis. Sei $E_i' \subseteq E_i$ abzählbar mit $\overline{E_i'} = E_i$ und r_i eine vollständige Metrik, die $\mathcal{O}_i$ erzeugt, $i \in \mathbb{N}$. Weiter sei $(x_i')_{i\in\mathbb{N}} \in \times_{i\in\mathbb{N}} E_i'$.

Zunächst bemerken wir, dass für $x^1 = (x_i^1)_{i\in\mathbb{N}}, x^2 = (x_i^2)_{i\in\mathbb{N}} \in E$ durch

$$r(x^1, x^2) := \sum_{i=1}^{\infty} 2^{-i}\big(r_i(x_i^1, x_i^2) \wedge 1\big)$$

eine vollständige Metrik auf E definiert wird, die $\mathcal{O}$ erzeugt. Außerdem ist damit die Menge

$$\Big\{(x_i)_{i\in\mathbb{N}} \in \bigtimes_{i\in\mathbb{N}} E_i' : x_i \neq x_i' \text{ für endlich viele } i \in \mathbb{N}\Big\}$$

abzählbar und dicht in E. Also ist $(E, \mathcal{O})$ polnisch. □

24.2 Mengensysteme

Analog zur Topologie ist die Produkt-σ-Algebra gerade so, dass Projektionen messbare Funktionen sind.

Definition 24.4 (Produkt-σ-Algebra). *Ist $(E_i, \mathcal{F}_i)_{i\in I}$ eine Familie von Messräumen, dann heißt die σ-Algebra*

$$\bigotimes_{i\in I} \mathcal{F}_i := \sigma\Big(\{\bigtimes_{i\in J} A_i \times \bigtimes_{i\in I\setminus J} E_i : J \Subset I, A_i \in \mathcal{F}_i\}\Big) \tag{24.1}$$

die Produkt-σ-Algebra *auf E. Ist $(E_i, \mathcal{F}_i) = (E, \mathcal{F}), i \in I$, so setzen wir $\mathcal{F}^I := \bigotimes_{i\in I} \mathcal{F}$.*

Bemerkung 24.5 (Messbarkeit der Koordinatenabbildung). Analog zur Produkttopologie gilt, dass die Projektionen π_i messbar bzgl. $\bigotimes_{i\in I} \mathcal{F}_i$ sind. Es ist nämlich für $A_i \in \mathcal{F}_i$

$$\pi_i^{-1}(A_i) = A_i \times \bigtimes_{j\neq i} E_j \in \bigotimes_{i\in I} \mathcal{F}_i.$$

Außerdem gilt

$$\bigotimes_{i\in I} \mathcal{F}_i = \sigma\Big(\big\{A_i \times \bigtimes_{j\neq i} E_j : A_i \in \mathcal{F}_i\big\}\Big). \tag{24.2}$$

Lemma 24.6 (Produkt-σ-Algebra bei abzählbaren Produkten). *Sei I abzählbar und $(E_i, \mathcal{O}_i)_{i\in I}$ eine Familie polnischer Räume sowie $(E, \mathcal{O})$ der Produktraum, versehen mit der Produkttopologie aus Definition 24.1. Dann ist $\mathcal{B}(E) = \bigotimes_{i\in I} \mathcal{B}(E_i)$. Insbesondere ist $\mathcal{B}(\mathbb{R}^d) = \bigotimes_{i=1}^d \mathcal{B}(\mathbb{R})$.*

Beweis. Da alle $(E_i, \mathcal{O}_i)$, $i \in I$, separabel sind, gibt es nach Lemma 19.6 für jedes $i \in I$ eine abzählbare Basis $\mathcal{C}_i$ von $\mathcal{O}_i$. Damit ist

$$\mathcal{C} := \{\bigtimes_{i\in J} A_i \times \bigtimes_{i\in I\setminus J} E_i : J \Subset I, A_i \in \mathcal{O}_i\}$$

eine abzählbare Basis von $(E, \mathcal{O})$. Nach Lemma 20.8 gilt $\sigma(\mathcal{C}) = \mathcal{B}(E)$. Außerdem ist offenbar $\mathcal{C} \subseteq \bigotimes_{i\in I} \mathcal{B}(E_i)$ und damit $\mathcal{B}(E) \subseteq \bigotimes_{i\in I} \mathcal{B}(E_i)$. Andersherum liefert (24.2), dass für $A_i \in \mathcal{F}_i$

$$A_i \times \bigtimes_{j\neq i} E_j \in \sigma\Big(\big\{A_i \times \bigtimes_{j\neq i} E_j : A_i \in \mathcal{O}_i\big\}\Big) = \pi_i^{-1}(\mathcal{F}_i) \subseteq \mathcal{B}(E),$$

woraus die Behauptung folgt. □

Lemma 24.7 (Produkte von Erzeugern/Halbringen sind Erzeuger/Halbringe). *Seien $(E_i, \mathcal{F}_i)$ Messräume und $E = \bigtimes_{i\in I} E_i$.*

1. *Sei I endlich und $\mathcal{H}_i$ ein Halbring mit $\sigma(\mathcal{H}_i) = \mathcal{F}_i$. Dann ist*
$$\mathcal{H} := \{\bigtimes_{i\in I} A_i : A_i \in \mathcal{H}_i, i \in I\} \tag{24.3}$$
ein Halbring mit $\sigma(\mathcal{H}) = \bigotimes_{i\in I} \mathcal{F}_i$.
2. *Sei I beliebig und $\mathcal{H}_i$ ein schnittstabiler Erzeuger von $\mathcal{F}_i$, $i \in I$. Dann ist*
$$\mathcal{H} := \{\bigtimes_{i\in J} A_i \times \bigtimes_{i\in I\setminus J} E_i : J \Subset I, A_i \in \mathcal{H}_i, i \in J\}$$
ein schnittstabiler Erzeuger von $\bigotimes_{i\in I} \mathcal{F}_i$.

Beweis. Für 1. ist o.E. $I = \{1, \dots, d\}$. Es ist zunächst klar, dass $\mathcal{H}$ schnittstabil ist. Die Eigenschaft (ii) für Halbringe zeigt man durch Induktion über d. Die Behauptung ist klar für $d = 1$, da $\mathcal{H}_1$ ein Halbring ist. Gilt sie für $d-1$, so gilt

$$\begin{aligned}&(A_1 \times \dots \times A_d) \setminus (B_1 \times \dots \times B_d)\\ &\quad = \big(A_1 \times \dots \times A_{d-1} \times (A_d \setminus B_d)\big) \uplus \big((A_1 \times \dots \times A_{d-1}) \setminus (B_1 \times \dots \times B_{d-1})\big) \times (A_d \cap A_d')\end{aligned}$$

Der erste Term der letzten Zeile ist als disjunkte Vereinigung von Mengen aus $\mathcal{H}$ darstellbar, da $\mathcal{H}_d$ ein Halbring ist. Der zweite Term ist als disjunkte Vereinigung darstellbar, da nach Induktionsvoraussetzung $(A_1 \times \dots \times A_{d-1}) \setminus (B_1 \times \dots \times B_{d-1})$ als disjunkte Vereinigung aus Mengen der Form $A_1 \times \dots \times A_{d-1}$ mit $A_i \in \mathcal{H}_i, i = 1, \dots, d-1$ darstellbar ist.

Für 2. ist wieder klar, dass $\mathcal{H}$ schnittstabil ist. Aus (24.1) folgt sofort, dass $\mathcal{H} \subseteq \bigotimes_{i\in I} \mathcal{F}_i$, also $\sigma(\mathcal{H}) \subseteq \bigotimes_{i\in I} \mathcal{F}_i$. Andersherum ist klar, dass für $A_i \in \mathcal{F}_i$

$$A_i \times \bigtimes_{j\neq i} E_j \in \sigma\Big(\big\{A_i \times \bigtimes_{j\neq i} E_j : A_i \in \mathcal{H}_i\big\}\Big) \subseteq \sigma(\mathcal{H}),$$

woraus $\bigotimes_{i\in I} \mathcal{F}_i \subseteq \sigma(\mathcal{H})$ wegen (24.2) und damit die Behauptung folgt. □

In Lemma 20.9 haben wir bereits die Borel'sche σ-Algebra auf $\mathbb{R}^d$ kennen gelernt. Natürlich handelt es sich bei $\mathbb{R}^d$ auch um einen Produktraum. Wir überprüfen nun, ob es sich bei der σ-Algebra $\mathcal{B}(\mathbb{R}^d)$ auch tatsächlich um die hier eingeführte Produkt-σ-Algebra handelt.

Korollar 24.8 (Borel'sche $\boldsymbol{\sigma}$-Algebra auf $\mathbb{R}^d$ wird von Quadern erzeugt). *Sei $E = \mathbb{R}^d$. Für $\underline{a} = (a_1, \dots, a_d), \underline{b} = (b_1, \dots, b_d) \in \mathbb{R}^d$ setzen wir $\underline{a} \le \underline{b}$ genau dann, wenn $a_i \le b_i, i = 1, \dots, d$, sowie mit*

$$(\underline{a}, \underline{b}] := (a_1, b_1] \times \dots \times (a_d, b_d]$$

den halboffenen Quader. Dann definiert

$$\mathcal{H} := \{(\underline{a}, \underline{b}] : \underline{a}, \underline{b} \in \mathbb{Q}, \underline{a} \leq \underline{b}\}$$

einen Halbring mit $\sigma(\mathcal{H}) = \mathcal{B}(\mathbb{R}^d)$.

Beweis. Nach Beispiel 20.3.1 und Lemma 24.7.1 ist $\mathcal{H}$ ein Halbring, der $\bigotimes_{i=1}^d \mathcal{B}(\mathbb{R}) = \mathcal{B}(\mathbb{R}^d)$ erzeugt; siehe Lemma 24.6. □

24.3 Maße und Integrale

Mehrfachintegrale sind schon aus der Analysis bekannt. Wir definieren nun zunächst Maße auf Produkträumen und im gleichen Atemzug auch die dazu gehörigen (Mehrfach-)Integrale. Mit dem Satz von Fubini (Theorem 24.12) können dann Integrale nach Maßen auf Produkträumen als Mehrfachintegrale interpretiert und ausgewertet werden. Hierzu ist es notwendig, dass die in den Mehrfachintegralen auftauchenden Integranden messbar sind. Dies wird in Lemma 24.10 sichergestellt. Um Maße auf Produkträumen in genügender Allgemeinheit definieren zu können, benötigen wir zunächst den Begriff des Übergangskernes.

Definition 24.9 (Übergangskern). *Seien* $(E_i, \mathcal{F}_i), i = 1, 2$ *Messräume. Eine Abbildung* $\kappa : E_1 \times \mathcal{F}_2 \to \mathbb{R}_+$ *heißt ein* Übergangskern *von* $(E_1, \mathcal{F}_1)$ *nach* $(E_2, \mathcal{F}_2)$, *wenn (i) für alle* $x_1 \in E_1$ *ist* $\kappa(x_1, .)$ *ein Maß auf* $\mathcal{F}_2$ *und (ii) für alle* $A_2 \in \mathcal{F}_2$ *ist* $\kappa(., A_2)$ *nach* $\mathcal{F}_1$*-messbar.*

Ein Übergangskern heißt σ*-endlich, wenn es eine Folge* $E_{11}, E_{12}, \dots \in \mathcal{F}_1$ *gibt mit* $E_{1n} \uparrow E_1$ *und* $\sup_{x_1} \kappa(x_1, E_{1n}) < \infty$ *für alle* $n = 1, 2, \dots$ *Er heißt* stochastischer Kern *oder* Markov'scher Kern, *falls für alle* $x_1 \in E_1$ *gilt, dass* $\kappa(x_1, E_2) = 1$.

Lemma 24.10 (Messbarkeit integrierbarer Schnitte). *Seien* $(E_i, \mathcal{F}_i), i = 1, 2$ *Messräume,* κ *ein* σ*-endlicher Übergangskern von* $(E_1, \mathcal{F}_1)$ *nach* $(E_2, \mathcal{F}_2)$ *und* $f : E_1 \times E_2 \to \mathbb{R}_+$ *nach* $\mathcal{F}_1 \otimes \mathcal{F}_2$ *messbar. Dann ist*

$$x_1 \mapsto \kappa(x_1, .)[f] := \int \kappa(x_1, dx_2) f(x_1, x_2)$$

nach $\mathcal{F}_1$*-messbar.*

Beweis. Wir nehmen an, dass $\kappa(x_1, E_2) < \infty$ für alle $x_1 \in E_1$ gilt. (Der allgemeine Fall erfolgt dann mittels einer Folge $E_{11}, E_{12}, \dots \in \mathcal{F}_1$ mit $E_{1n} \uparrow E_1$.) Sei

$$\mathcal{D} := \{A \in \mathcal{F}_1 \otimes \mathcal{F}_2 : x_1 \mapsto \kappa(x_1, .)[1_A] \text{ ist } \mathcal{F}_1\text{-messbar}\}.$$

Dann prüft man leicht nach, dass $\mathcal{D}$ ein schnittstabiles Dynkin-System ist. Weiter ist sicherlich $\mathcal{H} \subseteq \mathcal{D}$, wobei $\mathcal{H}$ wie in (24.3) definiert ist. Damit ist nach Theorem 20.13 $\mathcal{F}_1 \otimes \mathcal{F}_2 = \sigma(\mathcal{H}) \subseteq \mathcal{D} \subseteq \mathcal{F}_1 \otimes \mathcal{F}_2$. Also ist $x_1 \mapsto \kappa(x_1, .)[1_A]$ für alle $A \in \mathcal{F}_1 \otimes \mathcal{F}_2$ nach $\mathcal{F}_1$-messbar. Diese Aussage lässt sich sofort erweitern, indem man anstatt 1_A eine Treppenfunktionen einsetzt. Durch monotone Konvergenz folgt dann auch, dass $x_1 \mapsto \kappa(x_1, .)[f]$ für alle messbaren, nichtnegativen Funktionen nach $\mathcal{F}_1$-messbar ist. □

Theorem 24.11 (Satz von Ionescu-Tulcea). *Seien* $(E_i, \mathcal{F}_i), i = 0, \dots, n$ *Messräume,* μ *ein* σ*-endliches Maß auf* $\mathcal{F}_0$ *und* κ_i *ein* σ*-endlicher Übergangskern von* $\left(\times_{j=0}^{i-1} E_j, \bigotimes_{j=0}^{i-1} \mathcal{F}_j \right)$

nach $(E_i, \mathcal{F}_i)$, $i = 1, \ldots, n$. *Dann gibt es genau ein* σ*-endliches Maß* $\mu \bigotimes_{i=1}^n \kappa_i$ *auf* $\left(\times_{i=0}^n E_i, \bigotimes_{i=0}^n \mathcal{F}_i \right)$ *mit*

$$\Big(\mu \bigotimes_{i=1}^n \kappa_i\Big)(A_0 \times \cdots \times A_n) = \int_{A_0} \mu(dx_0) \Big(\int_{A_1} \kappa_1(x_1, dx_2) \cdots \Big(\int_{A_n} \kappa_n(x_0, \ldots, x_{n-1}, dx_n) \Big) \cdots \Big). \tag{24.4}$$

Beweis. Wir zeigen das Theorem nur für $n = 1$, der allgemeine Fall erfolgt dann durch Induktion.

Der Beweis ist eine Anwendung von Theorem 21.15. Zunächst stellen wir fest, dass nach Lemma 24.7 das in (24.3) definierte Mengensystem $\mathcal{H}$ ein Halbring auf $\times_{i=1}^n E_i$ ist. Wir zeigen zunächst, dass die angegebene Mengenfunktion σ-endlich auf $\mathcal{H}$ ist. Es gibt nämlich $E_{i1}, E_{i2}, \in \mathcal{F}_i$ mit $E_{in} \uparrow E_i, i = 0, 1$ mit $\mu(E_{0n}) < \infty, \kappa_1(x_0, E_{1n}) < \infty, n = 1, 2, \ldots, x_0 \in E_0$ und $\sup_{x_0 \in E_0} \kappa_1(x_0, E_{1n}) =: C_n < \infty$. Damit ist $\mu \otimes \kappa_1(E_{0n} \times E_{1n}) \leq C_n \cdot \mu(E_{0n}) < \infty$ und $E_{0n} \times E_{1n} \uparrow E_0 \times E_1$. Damit ist also $\mu \otimes \kappa_1$ auch σ-endlich. Definiert man $\widetilde{\mu}$ auf $\mathcal{H}$ mittels (24.4), so ist dies also eine σ-endliche Mengenfunktion.

Wir zeigen nun, dass $\widetilde{\mu}$ σ-subadditiv und endlich additiv auf $\mathcal{H}$ ist. Für $A_1, \ldots, A_n \in \mathcal{H}$ und $A = \bigcup_{n=1}^\infty A_n \in \mathcal{H}$ ist wegen der σ-Subadditivität von $\kappa_1(x_0, .)$ für alle $x_0 \in E_0$

$$\begin{aligned} \widetilde{\mu}(A) &= \int \mu(dx_0) \int \kappa_1(x_0, dx_1) 1_A(x_0, x_1) \\ &\leq \sum_{n=1}^\infty \int \mu(dx_0) \int \kappa_1(x_0, dx_1) 1_{A_n}(x_0, x_1) = \sum_{n=1}^\infty \widetilde{\mu}(A_n). \end{aligned}$$

Analog zeigt man die endliche Additivität. Nach Lemma 21.4 ist $\widetilde{\mu}$ damit σ-additiv. Aus Theorem 21.15 folgt nun, dass es genau eine Erweiterung von $\widetilde{\mu}$ auf $\sigma(\mathcal{H}) = \bigotimes_{i=1}^n \sigma(\mathcal{H}_i)$ gibt, welche die im Theorem angegebene ist. □

Wir beschäftigen uns nun mit dem in Theorem 24.11 definierten Maß.

Theorem 24.12 (Satz von Fubini). *Seien* $(E_i, \mathcal{F}_i)$, μ, κ_i *und* $\mu \bigotimes_{i=1}^n \kappa_i$ *wie in Theorem 24.11. Weiter sei* $f : \times_{i=0}^n E_i \to \mathbb{R}_+$ *messbar bezüglich* $\bigotimes_{i=0}^n \mathcal{F}_i$. *Dann gilt*

$$\int f d\Big(\mu \bigotimes_{i=0}^n \kappa_i\Big) = \int \mu(dx_0) \Big(\int \kappa_1(x_1, dx_2) \cdots \Big(\int \kappa_n(x_0, \ldots, x_{n-1}, dx_n) f(x_0, \ldots, x_n) \Big) \cdots \Big). \tag{24.5}$$

Diese Gleichheit gilt auch, falls $f : \times_{i=0}^n E_i \to \mathbb{R}$ *messbar ist mit* $\int |f| d\big(\mu \bigotimes_{i=0}^n \kappa_i\big) < \infty$.

Beweis. Betrachte die Mengenfunktion $\widetilde{\mu}$ auf $\bigotimes_{i=0}^n \mathcal{F}_i$, gegeben durch

$$\widetilde{\mu} : A \mapsto \int \mu(dx_0) \Big(\int \kappa_1(x_1, dx_2) \cdots \Big(\int \kappa_n(x_0, \ldots, x_{n-1}, dx_n) 1_A(x_0, \ldots, x_n) \Big) \cdots \Big).$$

Man sieht, dass $\widetilde{\mu}$ auf $\mathcal{H}$ aus (24.3) mit $\mu \bigotimes_{i=1}^n \kappa_i$ übereinstimmt. Da $\mathcal{H}$ schnittstabil ist, folgt die Gleichheit (24.5) für Indikatorfunktionen wegen Proposition 21.10. Mittels Linearität des Integrals erweitert man (24.5) zunächst auf einfache Funktionen und dann mit Monotonie auf beliebige, nicht-negative, messbare Funktionen. Man beachte hierbei, dass alle vorkommenden Integranden nach Lemma 24.10 messbar sind. □

Korollar 24.13 (Produktmaße). *Sei* $E = \times_{i=1}^n E_i$ *und* $\mathcal{H}_i \subseteq 2^{E_i}$ *ein Halbring,* $i = 1, \dots, n$*, sowie* $\mu_i : \mathcal{H}_i \to \mathbb{R}_+$ σ*-endlich und,* σ*-additiv,* $i = 1, \dots, n$*. Dann gibt es genau ein Maß* $\mu_1 \otimes \dots \otimes \mu_n$ *auf* $\bigotimes_{i=1}^n \sigma(\mathcal{H}_i)$ *mit*

$$\mu_1 \otimes \dots \otimes \mu_n(A_1 \times \dots \times A_n) = \mu_1(A_1) \cdots \mu_n(A_n). \tag{24.6}$$

Für eine Funktion $f : E \to \mathbb{R}_+$ *und jede Permutation* π *auf* $\{1, \dots, n\}$ *gilt dann*

$$\int f d\mu_1 \otimes \dots \otimes \mu_n = \int \Big(\cdots \Big(\int f(x_1, \dots, x_n) \mu_{\pi(1)}(dx_{\pi_1}) \Big) \cdots \Big) \mu_{\pi(n)}(dx_{\pi(n)}).$$

Diese Formel gilt auch für $f : E \to \mathbb{R}$*, falls* $\int |f| d\mu_1 \otimes \dots \otimes \mu_n < \infty$*.*

Beweis. Das Korollar folgt direkt aus Theorem 24.11 und Theorem 24.12, wenn man $\kappa_i(x_0, \dots, x_{i-1}, .) = \mu_i(.)$ für alle $x_0, \dots, x_{i-1}$ setzt. □

Definition 24.14 (Endliches Produktmaß). *Betrachte dieselbe Situation wie in Korollar 24.13. Dann heißt das eindeutige Maß* $\mu_1 \otimes \dots \otimes \mu_n$ *aus Korollar 24.13 das* Produktmaß der $\mu_1, \dots, \mu_n$*. Wir schreiben auch*

$$\bigotimes_{i=1}^n \mu_i := \mu_1 \otimes \dots \otimes \mu_n.$$

Gilt $(E_i, \mathcal{H}_i, \mu_i) = (E_0, \mathcal{H}_i, \mu_0)$*,* $i = 1, \dots, n$*, sind also alle Räume gleich, so bezeichnen wir es auch mit*

$$\mu_0^{\otimes n} := \mu_1 \otimes \dots \otimes \mu_n.$$

Beispiel 24.15 (Mehrdimensionales Lebesgue-Maß).

1. Sei λ das ein-dimensionale Lebesgue-Maß auf $\mathcal{B}(\mathbb{R})$ aus Proposition 21.17. Dann heißt $\lambda^{\otimes d}$ das d-dimensionale Lebesgue-Maß.

2. Sei $f : \mathbb{R}^2 \to \mathbb{R}$ gegeben durch

$$f(x, y) = \frac{xy}{(x^2 + y^2)^2}.$$

   Dann ist für jedes $x \in \mathbb{R}$

$$\int \lambda(dy) f(x, y) = 0,$$

   da $f(x, .) \in \mathcal{L}^1(\lambda)$ und $f(x, y) = -f(x, -y)$. Also gilt insbesondere

$$\int \lambda(dx) \Big(\int \lambda(dy) f(x, y) \Big) = \int \lambda(dy) \Big(\int \lambda(dx) f(x, y) \Big) = 0,$$

   allerdings ist $|f|$ nicht nach $\lambda^{\otimes 2}$ integrierbar, weil f in $(0, 0)$ eine nicht integrierbare Polstelle besitzt. Wie dieses Beispiel zeigt, muss man mit Mehrfachintegralen aufpassen. Insbesondere folgt aus der Gleichheit und Endlichkeit der Mehrfachintegrale nicht, dass der Integrand integrierbar ist.

Wir erweitern nun die partielle Integration aus Theorem 8.30 auf mehrdimensionale Integrale.

Proposition 24.16 (Partielle Integration). *Es seien $f, g \in \mathcal{C}^1(\mathbb{R}^d)$ so, dass $\frac{\partial f}{\partial x_j} g, f\frac{\partial g}{\partial x_j}, fg \in \mathcal{L}^1(\lambda^d)$. Dann gilt*

$$\int \frac{\partial f}{\partial x_j} g d\lambda^d = -\int f\frac{\partial g}{\partial x_j} d\lambda^d.$$

Beweis. Es genügt zu zeigen, dass

$$\int \frac{\partial h}{\partial x_j} d\lambda^d = 0 \tag{24.7}$$

für $h \in \mathcal{C}^1(\lambda^d)$ mit $h, \frac{\partial h}{\partial x_j} \in \mathcal{L}^1(\lambda^d)$.
Ist dies nämlich bewiesen, folgt die Behauptung mit der Wahl $h = fg$.
Sei zunächst $h = 0$ außerhalb $K := \{\underline{x} : |x_j| \le R\}$ für ein $R > 0$. Wir zeigen (24.7) durch Induktion nach d. Für $d = 1$ muss $j = d$ sein und dann folgt die Behauptung nach dem Hauptsatz der Differential- und Integralrechnung. Gilt (24.7) für ein d, zeigen wir die Behauptung nun für $d+1$. Ist $j \le d$, folgt mit Fubini

$$\int \frac{\partial h}{\partial x_j} d\lambda^{d+1} = \int\int \frac{\partial h(\underline{x}, x_{d+1})}{\partial x_j} \lambda^d(\underline{x})\lambda(dx_{d+1}) = \int 0 d\lambda(dx_{d+1}) = 0.$$

Für $j = d+1$ ist wieder nach dem Hauptsatz der Differential- und Integralrechnung

$$\int \frac{\partial h}{\partial x_j} d\lambda^{d+1} = \int\int \frac{\partial h(\underline{x}, x_{d+1})}{\partial x_j} \lambda(dx_{d+1})\lambda^d(\underline{x}) = \int 0 d\lambda^d(d\underline{x}) = 0.$$

Im allgemeinen Fall wählen wir $\varphi \in \mathcal{C}^1(\mathbb{R}, [0,1])$ mit $1_{[-1,1]} \le \varphi \le 1_{[-2,2]}$ und bemerken, dass $\underline{x} \mapsto \varphi(\varepsilon x_j)h(\underline{x})$ die Voraussetzungen von oben erfüllt. Damit gilt mit majorisierter Konvergenz

$$\begin{aligned}\int \frac{\partial h(\underline{x})}{\partial x_j} d\underline{x} &= \lim_{\varepsilon\to 0}\int \varphi(\varepsilon x_j)\frac{\partial h(\underline{x})}{\partial x_j} d\underline{x} \\ &= \lim_{\varepsilon\to 0}\int \frac{\partial}{\partial x_j}\varphi(\varepsilon x_j)h(\underline{x}) - \int f(\underline{x})\varepsilon\varphi'(\varepsilon x_j)d\underline{x} \\ &= \lim_{\varepsilon\to 0}\int \frac{\partial}{\partial x_j}\varphi(\varepsilon x_j)h(\underline{x})d\underline{x} = 0.\end{aligned}$$

□

Definition 24.17 (Normalgebiet). *Eine Menge $E \subseteq \mathbb{R}^d$ heißt ein Normalgebiet basierend auf Funktionen $a_1, ..., a_d, b_1, ..., b_d$, wenn*

$$\begin{aligned}E = \{\underline{x} : a_1 \le x_1 \le b_1, a_2(x_1) \le x_2 \le b_2(x_1), &a_3(x_1, x_2) \le x_3 \le b_3(x_1, x_2), ..., \\ &a_d(x_1, ..., x_{d-1}) \le x_d \le b_d(x_1, ..., x_{d-1})\}.\end{aligned}$$

Proposition 24.18 (Integration über ein Normalengebiet). *Sei E ein Normalengebiet und $f : E \to \mathbb{R}$ messbar. Existiert $\int_E f(\underline{x})d\underline{x}$, so gilt*

$$\int_E f(\underline{x})d\underline{x} = \int_{a_1}^{b_1}\int_{a_2(x_1)}^{b_2(x_1)}\cdots\int_{s_d(x_1,...,x_{d-1})}^{b_d(x_1,...,x_{d-1})} f(x_1, ..., x_d)dx_d\cdots dx_1.$$

Beweis. Dies folgt direkt aus Korollar 24.13. □

Beispiel 24.19 (Inhalt eines Kreises, eines Kegels und einer Kugel). 1. Bekanntermaßen ist der Inhalt von $B_1(\underline{0}) \subseteq \mathbb{R}^2$ gerade π. Dies berechnen wir nun mittels des Satzes von Fubini, Proposition 24.18. Der Vollkreis $B_1(0)$ lässt sich schreiben als

$$B_1(0) = \{(x, y) : -1 \le x \le 1, -\sqrt{1-x^2} \le y \le \sqrt{1-x^2}\}.$$

Es ist

$$\frac{d}{dx} x\sqrt{1-x^2} + \arcsin(x) = \sqrt{1-x^2} - \frac{2x^2}{2\sqrt{1-x^2}} + \frac{1}{\sqrt{1-x^2}} = \frac{1-x^2-x^2+1}{\sqrt{1-x^2}} = 2\sqrt{1-x^2}$$

und damit

$$\begin{aligned}\lambda^2(B_1(0)) &= \int 1_{B_1(\underline{0})}(x,y)d(x,y) = \int_{-1}^1 \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} 1dydx = \int_{-1}^1 2\sqrt{1-x^2}dx \\ &= 2x\sqrt{1-x^2} + \arcsin(x)\Big|_{x=0}^1 = \pi.\end{aligned}$$

Genauso zeigt man

$$\begin{aligned}\lambda^2(B_z(0)) &= \int_{-z}^{z} \int_{-\sqrt{z^2-x^2}}^{\sqrt{z^2-x^2}} 1dydx \overset{x'=x/z}{=} z\int_{-1}^1 \int_{-\sqrt{z^2-x'^2z^2}}^{-\sqrt{z^2-x'^2z^2}} 1dydx' \\ &\overset{y'=y/z}{=} z^2 \int_{-1}^1 \int_{-\sqrt{1-x'^2}}^{-\sqrt{1-x'^2}} 1dy'dx' = z^2\pi.\end{aligned}$$

2. Wir betrachten die Menge

$$E = \{(x, y, z) : 0 \le z \le 1, -z \le x \le z, -\sqrt{z^2-x^2} \le y \le \sqrt{z^2-x^2}\},$$

also eine Teilmenge eines Kegels. Hier ist

$$\lambda^3(E) = \int_0^1 \lambda^2(B_z(\underline{0}))dz = \pi \int_0^1 z^2dz = \frac{\pi}{3}.$$

3. Nun betrachten wir die Kugel im $\mathbb{R}^3$

$$K = \{(x, y, z) : x^2 + y^2 + z^2 \le 1\} = \{(x, y, z) : x^2 + y^2 \le 1 - z^2\}.$$

Dann ist

$$\lambda^3(K) = \int_{-1}^1 \lambda^2(B_{\sqrt{1-z^2}}(\underline{0}))dz = \pi \int_{-1}^1 (1-z^2)dz = \frac{4\pi}{3}.$$

4. Es liegt nahe, den Zusammenhand zwischen $\lambda^2(B_1(\underline{0}))$ und $\lambda^3(B_1(\underline{0}))$ zu verallgemeinern. Sei hierfür $\alpha_d := \lambda^d(B_1(\underline{0}))$. Dann gilt sicher $\lambda^d(B_r(\underline{0})) = r^d\lambda^d(B_1(\underline{0}))$. Hiermit schreiben wir

$$\begin{aligned}\alpha_{d+1} &= \lambda^{d+1}(B_1(\underline{0})) = \int_{-1}^1 \lambda^d(B_{\sqrt{1-z^2}})dz = \alpha_d \int_{-1}^1 (1-z^2)^{d/2}dz \\ &\overset{z=\cos\varphi}{=} \alpha_d \underbrace{\int_0^\pi \sin^{d+1}\varphi d\varphi}_{=:\beta_{d+1}}.\end{aligned}$$

Nun ist mit partieller Integration

$$\beta_{d+1} = \int_0^\pi \sin^{d+1}\varphi d\varphi = d\int_0^\pi \cos^2\varphi \sin^{d-1}\varphi d\varphi = d(\beta_{d-1} - \beta_{d+1}),$$

also $\beta_{d+1} = \frac{d}{d+1}\beta_{d-1}$ mit $\beta_0 = \pi, \beta_1 = 2$. Daraus folgt

$$\beta_d = \begin{cases} \frac{1}{2}\cdot\frac{3}{4}\cdots\frac{d-1}{d}\pi = \pi\frac{\prod_{k<d \text{ ungerade}} k}{\prod_{k<d+1 \text{ gerade}} k}, & d \text{ gerade,} \\ \frac{2}{3}\cdot\frac{4}{5}\cdots\frac{d-1}{d}2 = 2\frac{\prod_{k<d \text{ gerade}} k}{\prod_{k<d+1 \text{ ungerade}} k}, & d \text{ ungerade.} \end{cases}$$

Damit ist

$$\begin{aligned} \alpha_d &= \alpha_1 \cdot \frac{\alpha_2}{\alpha_1}\cdots\frac{\alpha_d}{\alpha_{d-1}} = 2\beta_2\cdots\beta_{d-1}\cdot\beta_d \\ &= \begin{cases} (2\pi)^{d/2}\frac{1}{\prod_{k<d+1 \text{ gerade}} k} = \frac{\pi^{d/2}}{(d/2)!}, & d \text{ gerade,} \\ 2(2\pi)^{(d-1)/2}\frac{1}{\prod_{k<d+1 \text{ ungerade}} k} = \frac{2(2\pi)^{(d-1)/2}}{d\cdot(d-2)\cdots 1} = \frac{\pi^{(d-1)/2}}{\frac{d}{2}\cdot\frac{d-2}{2}\cdots\frac{1}{2}}, & d \text{ ungerade.} \end{cases} \end{aligned}$$

Teil VI
Das Lebesgue-Maß auf $\mathcal{B}(\mathbb{R}^d)$

In dem nun folgenden Teil der Vorlesung wenden wir Gelerntes auf das in Proposition 21.17 konstruierte Lebesgue-Maß λ^d auf $\mathcal{B}(\mathbb{R}^d)$ an. Wir erinnern daran, dass dies das einzige Maß auf $\mathcal{B}(\mathbb{R}^d)$ ist mit

$$\lambda^d\Big(\prod_{i=1}^d (a_i, b_i]\Big) = \prod_{i=1}^d (b_i - a_i) \tag{24.8}$$

für alle $a_1, ..., a_d, b_1, ..., b_d \in \mathbb{Q}$ mit $a_i \leq b_i, i = 1, ...d$. Für Integrale bezüglich des Lebesgue-Maßes gibt es mehrere Schreibweisen. Sei $f : \mathbb{R}^d \to \mathbb{R}$ messbar. Falls $\lambda^d[f]$ existiert, schreiben wir

$$\begin{aligned}
\lambda^d[f] &= \int f d\lambda^d = \int f(\underline{x})\lambda^d(d\underline{x}) = \int f(\underline{x})d\underline{x} \\
&= \int f(x_1, ..., x_d)\lambda^d(dx_1, ..., dx_d) \\
&= \int f(x_1, ..., x_d)d(x_1, ..., x_d) \\
&= \int \cdots \int f(x_1, ..., x_d)dx_1 \cdots dx_d.
\end{aligned}$$

In diesem Abschnitt geht es um Rechenregeln dieses Integrals. Vor allem sind das die Transformationsformel, Theorem 25.13 und der Satz von Gauss, Theorem 27.11. Letzteres Theorem erfordert, dass wir neben λ^d auch noch das Integral auf Untermannigfaltigkeiten des $\mathbb{R}^d$ definieren, wozuAbschnitt 26 vorgesehen ist.

25 Eigenschaften

Wir kommen nun zu wichtigen Eigenschaften des Lebesgue-Maßes bzw. des Lebesgue-Integrals. In Abschnitt 25.1 geben wir Charakterisierungen von Nullmengen, bschäftigen uns dann mit der Translationsinvarianz in Abschnitt 25.2. Schließlich folgt die für konkrete Rechnungen wichtige Transformationsformel in Abschnitt 25.3. Dies ist die Erweiterung der Integration durch Substitution aus Theorem 8.32 auf mehrdimensionale Integrale.

25.1 Nullmengen

Für das Lebesgue-Maß $\lambda^d : \mathcal{B}(\mathbb{R}^d) \to \overline{\mathbb{R}}_+$, ist eine Nullmenge eine Menge $N \in \mathcal{B}(\mathbb{R}^d)$ mit $\lambda^d(N) = 0$. Aus der Konstruktion des äußeren Maßes in Theorem 21.15 wissen wir auch, dass wir für $N' \notin \mathcal{B}(\mathbb{R}^d)$ das Maß $\lambda^d(N') = 0$ definiern können, wenn es $N \in \mathcal{B}(\mathbb{R}^d)$ gibt mit $N' \subseteq N$. In diesem Abschnitt erhalten wir weiter eine geometrische Anschauung von solchen Nullmengen.

Lemma 25.1 (Offene Mengen als Vereinigung von Würfeln). *Sei $A \subseteq \mathbb{R}^d$ offen (in der euklidischen Topologie). Dann gibt es kompakte (oder halboffene) Würfel $W_1, W_2, ...,$ die* *höchstens Randpunkte (keine Punkte) gemeinsam haben mit $A = \bigcup_{n=1}^\infty W_n$. Außerdem gilt*

$$\lambda^d(A) = \sum_{n=1}^\infty \lambda^d(W_n).$$

Beweis. Wir definieren im Falle von halboffenen Würfeln

$$\begin{aligned} \mathcal{V}_k &= \Big\{ \Big(\frac{n}{2^k}, \frac{n+1}{2^k} \Big] : n \in \mathbb{Z} \}, \\ \mathcal{W}_k &= \Big\{ V_1 \times \cdots \times V_d : V_1, ..., V_d \in \mathcal{V}_k \}. \end{aligned} \tag{25.1}$$

Dann haben $W_1 \in \mathcal{W}_k, W_2 \in \mathcal{W}_\ell$ mit $\ell \geq k$ entweder keine Punkte gemeinsam oder es ist $W_2 \subseteq W_1$. (Im Falle von kompakten Würfeln gibt es höchstens Randpunkte gemeinsam.) Wir definieren nun iterativ

$$\begin{aligned} \mathcal{W}_{1,U} &:= \{ W \in \mathcal{W}_k : W \subseteq U \}, \\ \mathcal{W}_{k+1,U} &:= \{ W \in \mathcal{W}_{k+1} : W \subseteq U, W \not\subseteq W' \text{ für alle } W' \in \mathcal{W}_{k,U} \}. \end{aligned}$$

Dann ist nach Konstruktion $A = \bigcup_{k=1}^\infty \bigcup_{W \in \mathcal{W}_{k,U}} W$, woraus die erste Behauptung folgt. Für die zweite Behauptung ist zunächst per Definition klar, dass $\lambda^d(W)$ für einen Würfel W nicht davon abhängt ob der Wüfel halboffen oder kompakt ist. Deshalb schreiben wir im Falle von halboffenen Würfeln mit monotoner Konvergenz

$$\lambda^d(U) = \lambda^d[1_U] = \lambda^d\Big[\sum_{n=1}^\infty 1_{W_n} \Big] = \sum_{n=1}^\infty \lambda^d[1_{W_n}] = \sum_{n=1}^\infty \lambda^d(W_n)$$

und die Behauptung folgt. □

Theorem 25.2 (Charakterisierung von λ^d-Nullmengen). *Sei $N \subseteq \mathbb{R}^d$. Dann sind folgende Aussagen äquivalent:*

1. *N ist eine λ^d-Nullmenge.*
2. *Für jedes $\varepsilon > 0$ gibt es eine offene Menge $A \supseteq N$ mit $\lambda^d(A) < \varepsilon$.*
3. *Für jedes $\varepsilon > 0$ gibt es eine abzählbare Familie $W_1, W_2, ...$ von (kompakten) Würfeln mit*

$$N \subseteq \bigcup_{n=1}^\infty W_n \text{ und } \sum_{n=1}^\infty \lambda^d(W_n) < \varepsilon.$$

Beweis. 1.⇒2.: Wir wissen, dass λ^d als Einschränkung des äußeren Maßes μ^* auf $\mathcal{B}(\mathbb{R}^d)$ entsteht, wobei μ^* die einzige Fortsetzung einer Mengenfunktion auf dem Halbring $\mathcal{H}$ der halboffenen Würfel entsteht; siehe Theorem 21.15 und Lemma 25.1. Es gibt also halboffene Würfel $W_1, W_2, ...$, die N überdecken und $\sum_{n=1}^\infty \lambda^d(W_n) < \varepsilon$. Wir vergrößern Würfel W_n an den abgeschlossenen Seiten zu einem offenen Würfel $\widetilde{W}_n \supseteq W_n$ mit $\lambda^d(\widetilde{W}_n \setminus W_n) < \varepsilon 2^{-n}$. Dann ist $\widetilde{W}_1, \widetilde{W}_2, ...$ eine offene Überdeckung von N mit

$$\sum_{n=1}^\infty \lambda^d(\widetilde{W}_n) \leq \sum_{n=1}^\infty \lambda^d(W_n) + \varepsilon 2^{-n} < 2\varepsilon.$$

Mit $A := \bigcup_{n=1}^{\infty} \widetilde{W}_n$ folgt die Behauptung.
2.⇒3.: Dies folgt direkt aus Lemma 25.1.
3.⇒1.: Wegen der σ-Subadditivität von λ^d gilt

$$\lambda^d(N) \leq \sum_{n=1}^{\infty} \lambda^D(W_n) < \varepsilon.$$

Daraus folgt bereits $\lambda^d(N) = 0$. □

Korollar 25.3 (Bilder von Nullmengen sind Nullmengen). *1. Sei $\underline{\varphi} : \mathbb{R}^d \to \mathbb{R}^d$ Lipschitz-stetig und N eine λ^d-Nullmenge. Dann ist auch $\underline{\varphi}(N)$ eine λ^d-Nullmenge.*

2. Sei N eine λ^d-Nullmenge und $\underline{\varphi} \in \mathcal{C}^1(O, \mathbb{R}^d)$ für eine offene Menge $U \supseteq N$. Dann ist auch $\underline{\varphi}(N)$ eine λ^d-Nullmenge.

Beweis. 1. Wir definieren $r_\infty(\underline{x}, \underline{y}) := \max_{i=1,\dots,d} |x_i - y_i|$. Dann ist r_∞ eine Metrik auf $\mathbb{R}^d$ und es gibt $0 < c < C$, so dass $c|\underline{x} - \underline{y}| \leq r_\infty(\underline{x}, \underline{y}) \leq C|\underline{x} - \underline{y}|$. Insbesondere ist die Funktion φ auch Lipschitz-stetig (mit Lipschitz-Konstante L_∞) bezüglich r_∞. Weiter sind offene Bälle bezüglich r_∞ gerade offene Würfel mit derselben Seitenlänge. Wie wir aus dem Beweis von Lemma 25.1 wissen, können wir die offenen Würfel aus Theorem 25.2.3 so wählen, dass sie Bälle bezüglich r_∞ sind. Es gibt also eine offene Überdeckung $W_1 = \underline{x}_1 + (0, \varepsilon_1)^d, W_2 = \underline{x}_2 + (0, \varepsilon_2)^d, \dots$ mit offenen Bällen von N und $\sum_{n=1}^{\infty} \lambda^d(W_n) < \varepsilon$.

Für den offenen Ball $W = \underline{x} + (0, \varepsilon)^d$ bezüglich r_∞ mit Seitenlänge $\varepsilon > 0$ ist wegen der Lipschitz-Stetigkeit von φ gerade $\varphi(Q) \subseteq \varphi(\underline{x}) + L_\infty(W - \underline{x})$. Nun ist $W_n' = \varphi(\underline{x}_n) + L_\infty(W_n - \underline{x}_n)$, $n = 1, 2, \dots$ eine Überdeckung von $\varphi(N)$ mit

$$\sum_{n=1}^{\infty} \lambda^d(W_n') \leq L_\infty \sum_{n=1}^{\infty} \lambda^d(W_n') \leq L_\infty \varepsilon.$$

Daraus folgt nun, dass $\varphi(N)$ eine Nullmenge ist.
2. Nach Lemma 25.1 und Theorem 25.2 gibt es abzählbar viele kompakte Würfel $W_1, W_2, \dots \subseteq U$, die N überdecken. Nun ist $\varphi|_{W_n} \in \mathcal{C}^1(W_n, \mathbb{R}^d)$ nach dem Schrankensatz, Theorem 15.8, Lipschitz-stetig. Also ist $\varphi(N \cap W_n)$ nach 1. eine λ^d-Nullmenge, also auch

$$\lambda^d(\varphi(N)) = \lambda^d\Big(\varphi\Big(\bigcup_{n=1}^{\infty} N \cap W_n\Big)\Big) \leq \sum_{n=1}^{\infty} \lambda^d(\varphi(N \cap W_n)) = 0.$$

□

25.2 Translationsinvarianz

Wir erinnern an Beispiel 21.26, in dem wir die Translationsinvarianz des Lebesgue-Maßes gezeigt haben. Wir zeigen nun in Proposition 25.8, dass Vielfache des Lebesgue-Maßes die einzigen translationsinvarianten Maße auf $\mathcal{B}(\mathbb{R}^d)$ sind.

Definition 25.4 (Translationsinvariante Maße). *Sei μ ein Maß auf $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. Dann heißt μ translationsinvariant, falls $\mu(\underline{x} + A) = \mu(A)$ für alle $\underline{x} \in \mathbb{R}^d, A \in \mathcal{B}(\mathbb{R}^d)$.*

Beispiel 25.5 (Translationsinvarianz des Lebesgue-Maßes). Das d-dimensionale Lebesgue-Maß λ^d ist translationsinvariant.
Denn: Für $\underline{x} \in \mathbb{R}^d$ sei $\mu_x(A) := \lambda^d(\underline{x} + A)$ für $A \in \mathcal{B}(\mathbb{R}^d)$. Ist $A = \prod_{i=1}^d (a_i, b_i]$ für $a_1, ..., a_d, b_1, ..., b_d \in \mathbb{R}$ mit $a_i \leq b_i, i = 1, ...d$, so gilt $\mu_{\underline{x}}(A) = \lambda^n(A)$ wegen (24.8). Deshalb stimmen λ^d und $\mu_{\underline{x}}$ auf einem schnittstabilen Erzeuger von $\mathcal{B}(\mathbb{R}^d)$ überein und deshalb gilt $\lambda^d = \mu_{\underline{x}}$ wegen Proposition 21.10. Mit anderen Worten ist $\lambda^d(\underline{x} + A) = \mu(A)$.

Lemma 25.6 (Nullmengen translationsinvarianter Maße). *Sei μ ein translationsinvarientes, von innen bezüglich des Systems kompakter Menge reguläres Maß auf $(\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m))$.*

1. *Ist $H := H_{i,c} := \{\underline{x} \in \mathbb{R}^m : x_i = c\}$ für $i = 1, ..., d$ und $c \in \mathbb{R}$. Dann ist $\mu(H) = 0$.*
2. *Ist U offen, $U \supseteq E \in \mathcal{B}(\mathbb{R}^d)$ und $f \in \mathcal{C}^0(U, \mathbb{R}^d)$ Lipschitz-stetig mit Lipschitz-Konstante L. Dann gilt*
$$\mu(\underline{f}(E)) \leq L^d \mu(E).$$
3. *Ist insbesondere $\underline{f} \in \mathcal{C}^0(\mathbb{R}^d, \mathbb{R}^d)$ Lipschitz-stetig und $\mu(N) = 0$, so ist auch $\mu(\underline{f}(N)) = 0$.*

Beweis. 1. Sei $Q = [0,1)^m$ und $G = Q \cap \{\underline{x} : x_i = 0\}$. Wegen der Regularität und der Translationsinvarianz können wir für $s_1, ..., s_N \subseteq [0,1]$ paarweise verschieden

$$N \cdot \mu(G) = \sum_{j=1}^{N} \mu(s_j \underline{e}_i + G) = \mu\Big(\bigcup_{j=1}^{N} s_j \underline{e}_i + G \Big) \leq \mu(Q) < \infty$$

schreiben. Da N beliebig groß war, gilt also $\mu(G) = 0$. Die Behauptung folgt nun aus

$$\mu(H) = \mu\Big(\bigcup_{\underline{x} \in \mathbb{Z}^m, x_i = 0} c\underline{e}_i + \underline{x} + G \Big) = \sum_{\underline{x} \in \mathbb{Z}^m, x_i = 0} \mu(G) = 0.$$

2. OBdA ist $\mu(E) < \infty$ (sonst ist die Aussage klar). Angenommen, es gilt $\mu(\underline{f}(E)) > L^d \mu(E)$. Da μ von innen bezüglich des Systems kompakter Menge regulär ist, gibt es eine kompakte Menge $K \subseteq E$ mit $\mu(\underline{f}(K)) > \mu(K)$. Sei $\varepsilon > 0$. Da K mit endlich vielen Quadern Q mit Kantenlänge $\varepsilon > 0$ überdeckt werden kann, gibt es damit auch einen Quader $Q_\varepsilon(\underline{x}) = \underline{x} + [0, \varepsilon)^d \subseteq U$ mit
$$\mu(\underline{f}(Q_\varepsilon(\underline{x}))) > L^d \mu(Q_\varepsilon(\underline{x})).$$
Wegen der Lipschitz-Stetigkeit ist jedoch $|\underline{f}(\underline{x}) - \underline{f}(\underline{y})| \leq L|\underline{x} - \underline{y}|$, also $f(Q_\varepsilon(\underline{x})) \subseteq Q_{L\varepsilon}(f(\underline{x}))$ und damit
$$\mu(\underline{f}(Q_\varepsilon(\underline{x})) \leq \mu(Q_{L^d\varepsilon}(\underline{f}(\underline{x}))) = L^d \mu(Q_\varepsilon(\underline{x})),$$
wobei die letzte Gleichheit aus der Translationsinvarianz von μ folgt. Dies ist jedoch ein Widerspruch und die Aussage folgt. 3. folgt sofort aus 1. und 2. □

Beispiel 25.7. Sei $\underline{\underline{A}} \in \mathbb{R}^{d \times d}$ und $H := \{\underline{x} \in \mathbb{R}^d : (\underline{\underline{A}}\, \underline{x})_i = c\}$. Dann ist $\lambda^d(H) = 0$.

Proposition 25.8 (Translationsinvariante Maße auf $\mathbb{R}^d$). *Sei μ ein Maß auf $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. Dann sind äquivalent:*

1. *$\mu = c \cdot \lambda^d$ für $c = \mu([0,1]^d)$.*
2. *μ ist translationsinvariant und von innen bezüglich des Systems kompakter Mengen ($\subseteq \mathbb{R}^d$) regulär.*

Beweis. 1.⇒2.: Dies folgt sofort aus Beispiel 25.5.
2.⇒1.: Sei $Q = \prod_{i=1}^{d}(a_i, b_i]$ für $a_1, ..., a_d, b_1, ..., b_d \in \mathbb{Q}$ mit $a_i \leq b_i, i = 1, ...d$. Wir müssen zeigen, dass $\mu(Q) = c \cdot \lambda^d(Q)$ gilt. Wegen der Eindeutigkeit der Maßfortsetzung folgt dann die Behauptung. Sei n groß genug, so dass $na_1, ..., na_d, nb_1, ..., nb_d \in \mathbb{Z}$. Dann gilt

$$Q = \prod_{i=1}^{d} \bigcup_{k_i=na_i}^{nb_i} \frac{1}{n}(k_i, k_i + 1] = \bigcup_{k_1=na_1}^{nb_1} \cdots \bigcup_{k_d=na_d}^{nb_d} \prod_{i=1}^{d} \frac{1}{n}(k_i, k_i + 1].$$

Wegen der Translationsinvarianz von μ ist nun

$$\mu(Q) = n^d \mu(\tfrac{1}{n}[0,1]^d) \prod_{i=1}^{d}(b_i - a_i) = \mu([0,1]^d) \prod_{i=1}^{d}(b_i - a_i) = \mu([0,1]^d)\lambda^d(Q).$$

Daraus folgt die Behauptung. □

Lemma 25.9 (Darstellung invertierbarer Matrizen). *Sei $\underline{\underline{A}} \in Gl(\mathbb{R}, d)$. Dann gibt es $\lambda_1, ..., \lambda_d$ und orthogonale Matrizen $\underline{\underline{S}}_1, \underline{\underline{S}}_2$, so dass $\underline{\underline{A}} = \underline{\underline{S}}_1 \underline{\underline{D}}\, \underline{\underline{S}}_2$ mit $\underline{\underline{D}} = diag(\lambda_1, ..., \lambda_d)$.*

Beweis. Die Matrix $\underline{\underline{A}}\,\underline{\underline{A}}^\top$ ist symmetrisch und hat damit d reelle Eigenwerte. Ist $\underline{x}$ ein Eigenvektor zum Eigenwert λ, so gilt $\lambda\langle \underline{x}, \underline{x}\rangle = \langle \underline{\underline{A}}\,\underline{\underline{A}}^\top \underline{x}, \underline{x}\rangle = \langle \underline{\underline{A}}\underline{x}, \underline{\underline{A}}\,\underline{x}\rangle > 0$, woraus $\lambda > 0$ folgt. Also ist für eine orthogonale Matrix $\underline{\underline{T}}$ und eine Diagonalmatrix $\underline{\underline{D}}$ mit positiven Diagonaleinträgen

$$\underline{\underline{A}}^\top \underline{\underline{A}} = \underline{\underline{T}}\,\underline{\underline{D}}^2 \underline{\underline{T}}^\top.$$

Wir setzen $\underline{\underline{S}}_1 = \underline{\underline{A}}\,\underline{\underline{T}}\,\underline{\underline{D}}^{-1}$, also

$$\underline{\underline{S}}_1^\top \underline{\underline{S}}_1 = \underline{\underline{D}}^{-1} \underline{\underline{T}}^\top \underline{\underline{A}}^\top \underline{\underline{A}}\,\underline{\underline{T}}\,\underline{\underline{D}}^{-1} = \underline{\underline{E}}_d$$

und damit $\underline{\underline{S}}_1$ orthogonal. Weiter ist $\underline{\underline{S}}_2 := \underline{\underline{T}}^\top$ orthogonal und es gilt

$$\underline{\underline{S}}_1 \underline{\underline{D}}\,\underline{\underline{S}}_2 = \underline{\underline{A}}\,\underline{\underline{T}}\,\underline{\underline{D}}^{-1} \underline{\underline{D}}\,\underline{\underline{T}}^\top = \underline{\underline{A}}.$$

Daraus folgt die Behauptung. □

Proposition 25.10 (Bewegungsinvarianz des Lebesgue-Integrals). *Sei $\underline{\underline{A}} \in \mathbb{R}^{d \times d}$ und $\underline{\underline{A}}E = \{\underline{\underline{A}}\,\underline{x} : \underline{x} \in E\}$. Dann gilt*

$$\lambda^d(\underline{\underline{A}}E + \underline{x}) = c \cdot \lambda^d(E)$$

für $\underline{x} \in \mathbb{R}^d$ und $c = \lambda^d(\underline{\underline{A}}[0,1]^d)$. Ist außerdem $\underline{\underline{A}}$ sogar orthogonal, so gilt $\lambda^d(\underline{\underline{A}}E + \underline{x}) = \lambda^d(E)$.

Beweis. Wegen der Translationsinvarianz von λ^d ist oBdA $\underline{x} = 0$. Ist $\underline{\underline{A}}$ nicht invertierbar, so folgt die Behauptung aus Beispiel 25.7. Ist $\underline{\underline{A}}$ invertierbar, so definieren wir

$$\mu(E) := \lambda^d(\underline{\underline{A}}E).$$

Dann gilt

$$\mu(\underline{x} + E) = \lambda^d(\underline{\underline{A}}\,\underline{x} + \underline{\underline{A}}\,E) = \lambda^d(\underline{\underline{A}}\,E) = \mu(E)$$

wegen der Translationsinvarianz des Lebesgue-Integrals. Aus Proposition 25.8 folgt nun

$$\lambda^d(\underline{\underline{A}}E) = c \cdot \lambda^d(E)$$

für $E \in \mathcal{B}(\mathbb{R}^d)$ und $c = \lambda^d(\underline{\underline{A}}[0,1]^d)$ und damit die erste Behauptung. Für die zweite Behauptung benutzen wir $\underline{\underline{A}}B_1(\underline{0}) = B_1(\underline{0})$, falls $\underline{\underline{A}}$ orthogonal ist. (Hierzu berechnet man $||\underline{\underline{A}}\,\underline{x}||_2 = \langle \underline{\underline{A}}\,\underline{x}, \underline{\underline{A}}\,\underline{x}\rangle = \langle \underline{x}, \underline{\underline{A}}^\top \underline{\underline{A}}\,\underline{x}\rangle = ||\underline{x}||_2$.) Also gilt

$$\lambda^d(\underline{\underline{A}}[0,1]^d) \cdot \lambda^d(B_1(\underline{0})) = \lambda^d(\underline{\underline{A}}[0,1]^d) \cdot \lambda^d(\underline{\underline{A}}B_1(\underline{0})) = \lambda^d(B_1(\underline{0})),$$

also $\lambda^d(\underline{\underline{A}}[0,1]^d) = 1$. Daraus folgt nun auch die zweite Behauptung. □

25.3 Transformationsformel

Ziel der Transformationsformel ist die Erweiterung der Integration durch Substitution (siehe Theorem 8.32) auf mehrdimensionale Lebesgue-Integrale. Ist $\varphi \in \mathcal{C}^1(I,J)$ (und I und J Intervalle) und φ eindeutig umkehrbar (insbesondere also $\varphi(I) = J$), so gilt

$$\int_I f(\varphi(x)) \cdot |\varphi'(x)| dx = \int_J f(x) dx.$$

(Man bemerke, dass es hier erlaubt ist, dass φ monoton fällt.) Offenbar erweitert Theorem 25.13 dieses Ergebnis. Wir beginnen mit einer Folgerung aus Proposition 25.10.

Korollar 25.11 (Lineare Transformationsformel). *Sei $\underline{\underline{A}} \in \mathbb{R}^{d\times d}$. Dann gilt für $E \in \mathcal{B}(\mathbb{R}^d)$*

$$\lambda^d(\underline{\underline{A}}E + \underline{x}) = |det(\underline{\underline{A}})| \cdot \lambda^d(E).$$

Insbesondere gilt: Since $\underline{x}_1, ...\underline{x}_d$ linear unabhängige Vektoren. Dann ist das Volumen des von diesen Vektoren aufgespannten Parrallelotops $|det(\underline{\underline{A}})|$, wobei $\underline{\underline{A}})$ die Matrix mit Spaltenvektoren $\underline{x}_1, ..., \underline{x}_d$.

Beweis. Die Behauptung ist für orthogonales $\underline{\underline{A}}$ bereits in Proposition 25.10 gezeigt, da in diesem Fall $|\det(\underline{\underline{A}})| = 1$ gilt. Andernfalls gilt es noch zu zeigen, dass $\lambda^d(\underline{\underline{A}}[0,1]^d) = |\det(\underline{\underline{A}})|$. Ist $\det(\underline{\underline{A}}) = 0$, so folgt die Behauptung aus Beispiel 25.7. Anderfalls verwenden wir Lemma 25.9 und schreiben für orthogonale Matrizen $\underline{\underline{S}}_1, \underline{\underline{S}}_2$, und eine Diagonalmatrix $\underline{\underline{D}}$ mit Diagonaleinträgen $\lambda_1, ..., \lambda_d > 0$

$$\begin{aligned}\lambda^d(\underline{\underline{A}}[0,1]^d) &= \lambda^d(\underline{\underline{S}}_1\,\underline{\underline{D}}\,\underline{\underline{S}}_2[0,1]^d) = \lambda^d(\underline{\underline{D}}\,\underline{\underline{S}}_2[0,1]^d) = \lambda^d(\underline{\underline{D}}[0,1]^d) \cdot \lambda^d(\underline{\underline{S}}_2[0,1]^d) \\ &= \lambda^d([0,\lambda_1] \times \cdots \times [0,\lambda_d]) \cdot \lambda^d([0,1]^d) = \det(\underline{\underline{D}})\lambda^d([0,1]^d) = |\det(\underline{\underline{A}})| \cdot \lambda^d([0,1]^d).\end{aligned}$$

Für die letzte Behauptung genügt es zu sehen, dass das Parallelotop, das von $\underline{x}_1, ..., \underline{x}_d$ aufgespannt wird, gerade $\underline{\underline{A}}[0,1]^d$ ist. □

Lemma 25.12 (Infinitesimale Transformationsformel). *Sei $E \subseteq \mathbb{R}^d$ offen, $\underline{x} \in E$ und $\underline{f} : E \to \mathbb{R}^d$ mit $\underline{\underline{D}}\,\underline{f}(\underline{x})$ invertierbar. Weiter seien W_r Würfel mit Seitenlänge r, so dass $\underline{x} \in W_r$, $r > 0$. Dann gilt*

$$\frac{\lambda^d(\underline{f}(W_r))}{\lambda^d(W_r)} \xrightarrow{r\to 0} |\det \underline{\underline{D}}\,\underline{f}(\underline{x})|.$$

Beweis. OBdE ist $\underline{x} = \underline{f}(\underline{x}) = \underline{0}$. Wir betrachten zunächst den speziellen Fall $\underline{\underline{D}}\,\underline{f}(\underline{x}) = \underline{\underline{E}}_d$. Nach Definition der Differenzierbarkeit gilt dann

$$\frac{||\underline{f}(\underline{x}) - \underline{f}(\underline{0}) - \underline{\underline{E}}_d\underline{x}||}{||\underline{x}||} = \frac{||\underline{f}(\underline{x}) - \underline{x}||}{||\underline{x}||} \xrightarrow{\underline{x}\to 0} 0.$$

Sei $r_j \downarrow 0$ und $\varepsilon > 0$. Für $\underline{x} \in W_{r_j}$ und j groß genug gilt dann

$$||\underline{f}(\underline{x}) - \underline{x}|| \le \varepsilon ||\underline{x}|| \le \varepsilon r_j$$

und damit

$$\lambda^d(W_{r_j(1-2\varepsilon)}) \le \lambda^d(\underline{f}(W_{r_j})) \le \lambda^d(W_{r_j(1+2\varepsilon)}),$$

also

$$(1-2\varepsilon)^d \le \frac{\lambda^d(\underline{f}(W_{r_j}))}{\lambda^d(W_{r_j})} \le (1+2\varepsilon)^d$$

Die Behauptung folgt nun mit $j \to \infty$ und $\varepsilon \to 0$.

Im allgemeinen Fall betrachten wir $\underline{g} = \underline{\underline{D}}\,\underline{f}(\underline{0})^{-1} \cdot \underline{f}$, also $\underline{\underline{D}}\,\underline{g}(\underline{0}) = \underline{\underline{E}}_d$. Mit Korollar 25.11 folgt dann

$$\frac{\lambda^d(\underline{f}(W_r))}{\lambda^d(W_r)} = \frac{\lambda^d(\underline{\underline{D}}\,\underline{f}(\underline{0}) \cdot \underline{g}(W_r))}{\lambda^d(W_r)} = |\det(\underline{\underline{D}}\,\underline{f}(\underline{0}))| \cdot \frac{\lambda^d(\underline{g}(W_r))}{\lambda^d(W_r)} \xrightarrow{r\to 0} |\det(\underline{\underline{D}}\,\underline{f}(\underline{0}))|.$$

□

Theorem 25.13 (Transformationssatz). *Seien $E, E' \subseteq \mathbb{R}^d$ offen und $\underline{\varphi} \in \mathcal{C}^1(E, E')$ ein Diffeomorphismus. Ist $\underline{f}$ nach $\mathcal{B}(\mathbb{R}^d)$ messbar, so gilt*

$$\int_{E'} \underline{f}(\underline{x})d\underline{x} = \int_E \underline{f}(\underline{\varphi}(\underline{x})) \cdot |\det(\underline{\underline{D}}\,\underline{\varphi}(x))|d\underline{x},$$

falls eine der beiden Seiten existiert.

Beweis. Es genügt, die Behauptung für $\underline{f} = 1_A$ für ein $A \in \mathcal{B}(\mathbb{R}^d)$ mit $A \subseteq E'$ zu zeigen. Wie üblich folgt das allgemeine Resultat durch Approximation mit einfachen Funktionen. Wir behaupten nun die Gleichheit der zwei Maße

$$\mu = (\underline{\varphi}^{-1})_*\lambda^d, \qquad \nu = |\det(\underline{\underline{D}}\,\underline{\varphi}(\cdot))| \cdot \lambda^d. \tag{25.2}$$

Ist dies gezeigt, so folgt für $A \in \mathcal{B}(\mathbb{R}^d)$

$$\begin{aligned}\int 1_A d\underline{x} &= \lambda^d(A) = \lambda^d(\underline{\varphi}(\underline{\varphi}^{-1}(A))) = \mu(\underline{\varphi}^{-1}(A)) = \nu(\underline{\varphi}^{-1}(A)) \\ &= \int 1_{\underline{\varphi}^{-1}(A)}(\underline{x}) \cdot |\det(\underline{\underline{D}}\,\underline{\varphi}(x))|d\underline{x} = \int 1_{\underline{\varphi}(\underline{x})\in A} \cdot |\det(\underline{\underline{D}}\,\underline{\varphi}(x))|d\underline{x},\end{aligned}$$

also die Behauptung. Nach Proposition 21.10 genügt es, die Gleichheit der zwei Maße μ und ν aus (25.2) für einen schnittstabilen Erzeuger von $\mathcal{B}(\mathbb{R}^d)$ zu zeigen. Sei hierzu

$$\mathcal{C} = \bigcup_{k=1}^{\infty} \mathcal{W}_k$$

mit $\mathcal{W}_k$ aus (25.1) die Menge aller (halboffenen) Würfel, wobei $\mathcal{W}_k$ die Würfel mit Seitenlängen 2^k und Eckpunkten in $2^{-k}\mathbb{Z}^d$ sind. (Damit ist $\mathcal{C}$ schnittstabil und nach Lemma 25.1 ein Erzeuger von $\mathcal{B}(\mathbb{R}^d)$.) Angenommen, es gibt $W \in \mathcal{C}$ mit Seitenlänge $r > 0$ mit $\mu(W) \neq \nu(W)$. OBdA gibt es also $0 < \theta < 1$ mit $\mu(W) \leq \theta\nu(W)$. (Im Fall von $\nu(W) \leq \theta\mu(W)$ erfolgt die Argumentation analog.) Teilt man W in 2^d Quader mit Kantenlängen $r/2$, so muss es darunter auch einen Würfel W_1 geben mit $\mu(W_1) \leq \theta\nu(W_1)$. Iterativ konstruiert man so eine Folge von Würfeln $W_1, W_2, ...$ mit Kantenlängen $r_1, r_2, , ... \downarrow 0$ und $\mu(W_n) \leq \theta\nu(W_n)$, $n = 1, 2, ...$, insbesondere also

$$\limsup_{n\to\infty} \frac{\mu(W_n)}{\nu(W_n)} \leq \theta < 1. \tag{25.3}$$

Sei nun $\underline{x} \in \bigcap_{n=1}^{\infty} W_n$ und damit

$$\lim_{n\to\infty} \frac{\mu(W_n)}{\nu(W_n)} = \lim_{n\to\infty} \frac{\lambda^d(\underline{\varphi}(W_n))}{|\det(\underline{\underline{D}}\,\underline{\varphi}(x))| \cdot \lambda^d(W_n)} = 1$$

wegen der Stetigkeit von $\underline{x} \mapsto \underline{\underline{D}}\,\underline{\varphi}(\underline{x})$ und Lemma 25.12. Dies widerspricht aber (25.3), woraus nun $\mu = \nu$ folgt. □

25.4 Beispiele

Beispiel 25.14 ($d = 2$, Polarkoordinaten). In Beispiel 16.11 haben wir bereits Polarkoordinaten betrachtet, also den Diffeomorphismus

$$\underline{f} : \begin{cases} \mathbb{R}_+ \times (0, 2\pi) & \to \mathbb{R}^2 \setminus \{(x, 0) : x \geq 0\} \\ (r, \varphi) & \mapsto (r\cos\varphi, r\sin\varphi). \end{cases}$$

Die Jacobi-Matrix ist nach Beispiel 16.17

$$\underline{\underline{D}}\,\underline{f}(r, \varphi) = \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix}$$

mit

$$\det(\underline{\underline{D}}\,\underline{f}(r, \varphi)) = r.$$

Ist also I ein Intervall und $K_I := \{(x, y) \in \mathbb{R}^2 : \sqrt{x^2 + y^2} \in I\}$ eine Kugelschale (etwa ist $I = [0, 1]$ und $K_I = B_1(\underline{0})$) und $\underline{g} \in \mathcal{C}(K_I, \mathbb{R})$ messbar, dann ist

$$\int_{K_I} g(\underline{x})d\underline{x} = \int_I \int_0^{2\pi} g(r\cos\varphi, r\sin\varphi) r d\varphi dr. \tag{25.4}$$

Denn: Polarkoordinaten sind zwar nur für $\mathbb{R}^2 \setminus \{(x, 0) : x \geq 0\}$ definiert, diese Menge ist jedoch nach Lemma 25.6 eine λ^2-Nullmenge.

Wir geben nun ein paar Beispiele für die Berechnung mit Polarkoordinaten an.

1. Als Beispiel berechnen wir erneut den Inhalt des Kreises $B_z(0)$ wie in Beispiel 24.19. Aus (25.4) lesen wir

$$\int_{B_z(\underline{0})} 1\, d\underline{x} = \int_0^z \int_{-\pi}^{\pi} r d\varphi dr = 2\pi \int_0^r r dr = z^2\pi$$

ab.

2. Gegeben sei eine messbare Funktion $\underline{x} \mapsto f(||\underline{x}||_2)$, also eine rotationssymmetrische Funktion. Dann gilt

$$\int_{K_I} f(||\underline{x}||_2)d\underline{x} = \int_I \int_{-\pi}^{\pi} f(||(r\cos\varphi, r\sin\varphi)||_2)rd\varphi dr = 2\pi \int_I f(r)rdr. \tag{25.5}$$

Insbesondere ist f genau dann über K_I integrierbar, wenn $r \mapsto f(r)r$ über I integrierbar ist.

Beispiel 25.15 (Polarkoordinaten in $\mathbb{R}^d$). Allgemeiner als oben kann man auch im $\mathbb{R}^d$ Polarkoordinaten einführen. Dies geben wir hier nur skizzenhaft an. Wir definieren rekursiv (in d)

$$\begin{aligned} \underline{P}_2(r,\varphi_1) &:= (r\cos\varphi_1, r\sin\varphi_1), \\ \underline{P}_{d+1}(r,\varphi_1,...,\varphi_d) &:= (P_d(r,\varphi_1,...,\varphi_{d-1})\cdot\cos\varphi_d, r\sin\varphi_d) \end{aligned}$$

mit $r \in [0,\infty), \varphi_1 \in (-\pi,\pi), \varphi_2,...,\varphi_d \in (-\pi/2,\pi/2)$. Nun berechnen wir die Jacobi-Matrix von P_d rekursiv als

$$\underline{D}P_{d+1} = \begin{pmatrix} \underline{D}P_d\cdot\cos\varphi_d & -P_d\cdot\sin\varphi_d \\ \sin\varphi_d, 0, ..., 0 & r\cos\varphi_d \end{pmatrix}.$$

Daraus ergibt sich rekursiv durch Entwickeln der Determinante nach der letzten Zeile (wegen $r\frac{\partial}{\partial r}P_d = P_d$)

$$\begin{aligned} \det(\underline{D}P_n(r,\varphi_1,...,\varphi_d)) &= -(-1)^d(-1)^{d-1}r\sin^2\varphi_d\cos^{d-1}\varphi_d\det(\underline{D}P_d) \\ &\quad + \det(\underline{D}P_n(r,\varphi_1,...,\varphi_{d-1})\cdot r\cos\varphi_d^{d+1} \\ &= \det(\underline{D}P_n(r,\varphi_1,...,\varphi_{d-1})\cdot r\cos^{d-1}\varphi_d. \end{aligned}$$

Daraus und mit $\det(\underline{D}P_n(r,\varphi) = r$ folgt

$$\det(\underline{D}P_n(r,\varphi_1,...,\varphi_d)) = r^{n-1}\cos^0\varphi_1\cdot\cos^1\varphi_2\cdots\cos^{d-2}\varphi_{d-1}.$$

Ist nun $\underline{x} \mapsto f(||\underline{x}||_2)$ eine messbare Funktion, folgt damit genau wie oben mit $K_I = \{\underline{x} : ||\underline{x}|| \in I\}$ für ein messbares $I \subseteq \mathbb{R}$

$$\begin{aligned} \int_{K_I} f(||\underline{x}||_2)d\underline{x} &= \int_I f(r)r^{d-1}dr\cdot\int|\cos^0\varphi_1\cdots\cos^{d-2}\varphi_{d-1}|d\varphi_{d-1}\cdots d\varphi_d \\ &= \int_I f(r)r^{d-1}dr\cdot\int_0^1 dr^{d-1}dr\cdot\int|\cos^0\varphi_1\cdots\cos^{d-2}\varphi_{d-1}|d\varphi_{d-1}\cdots d\varphi_d \\ &= \int_I f(r)r^{d-1}dr\cdot d\cdot\lambda^d(B_1(\underline{0})) = \int_I f(r)r^{d-1}dr\cdot d\cdot\alpha_d \end{aligned} \tag{25.6}$$

mit α_d aus Beispiel 24.19.4.

Beispiel 25.16 (Das Gauss'sche Integral). In Proposition 9.3 haben wir bereits den Wert

$$\int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2}dx = \sqrt{2\pi} \tag{25.7}$$

mittels der Γ-Funktion berechnet. Weiter haben wir durch parameterabhängige Integrale in Beispiel 16.8 dasselbe Integral erneut berechnet. Nun erfolgt eine dritte Berechnungsweise, die vermutlich die einfachste der drei aufgezeigten Möglichkeiten darstellt. Wir schreiben mit dem Satz von Fubini und (25.5)

$$\Big(\int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2}dx\Big)^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{1}{2}(x^2+y^2)}dydx = 2\pi\int_0^{\infty} e^{-\frac{1}{2}r^2}rd\varphi dr = 2\pi e^{-\frac{1}{2}r^2}\Big|_{r=0}^{r=\infty} = 2\pi,$$

woraus (25.7) folgt.

Beispiel 25.17 (Jacobi-Abbildung). Wir betrachten den Diffeomorphismus

$$\underline{f}:\begin{cases}\mathbb{R}_+\times[0,1] & \to \mathbb{R}_+^2\\ (u,v) & \mapsto (s,t) := (u(1-v),uv),\end{cases}$$

bekannt als Jacobi-Abbildung. Diese Abbildung ist ein Diffeomorphismus und erfüllt $f_1(u,v)+f_2(u,v)=u$. Ihre Jacobi-Matrix ist

$$\underline{\underline{D}}\,\underline{f}(u,v) = \begin{pmatrix}1-v & -u\\ v & u\end{pmatrix}$$

mit

$$\det(\underline{\underline{D}}\,\underline{f}(u,v)) = u.$$

Damit gilt für eine messbare Funktion $g:\mathbb{R}^2\to\mathbb{R}$

$$\int_{\mathbb{R}_+^2} g(s,t)d(s,t) = \int_{\mathbb{R}_+\times[0,1]} g(u(1-v),uv)ud(u,v),$$

falls eine der beiden Seiten existiert.

Als Beispiel für diese Transformation betrachten wir die Beta-Funktion, siehe Abschnitt 9.1. In Definition 9.1 und Lemma 9.2 haben wir gesehen, dass

$$\int_0^1 t^{x-1}(1-t)^{y-1}dt = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} \tag{25.8}$$

mit

$$\Gamma(x) := \int_0^{\infty} t^{x-1}e^{-t}dt$$

gilt. Um dies erneut einzusehen, berechnen wir mit Hilfe des Satzes von Fubini

$$\begin{aligned}\Gamma(x)\Gamma(y) &= \int_0^{\infty}\int_0^{\infty} s^{x-1}t^{y-1}e^{-s-t}dsdt = \int_0^{\infty}\int_0^1 (u(1-v))^{x-1}(uv)^{y-1}e^{-u}udvdu\\ &= \int_0^{\infty} u^{x+y-1}e^{-u}du\int_0^1(1-v)^{x-1}v^{y-1}dv = \Gamma(x+y)\cdot\int_0^1 t^{x-1}(1-t)^{y-1}dt,\end{aligned}$$

woraus (25.8) folgt.

26 Untermannigfaltigkeiten des $\mathbb{R}^d$

Eine (differenzierbare) Untermannigfaltigkeit stellt man sich als eine (glatte) niedrigdimensionale Fläche in einem höher-dimensionalen Raum vor, etwa also die zwei-dimensionale Kugeloberfläche im $\mathbb{R}^3$. Den Spezialfall von ein-dimensionalen Untermannigfaltigkeiten haben wir bereits in Abschnitt 17 als Kurven kennen gelernt.

26.1 Definition und topologische Eigenschaften

Wir werden in diesem Abschnitt drei verschiedene Arten und Weisen kennen lernen, wie man Untermannigfaltigkeiten definieren kann. Siehe auch Tabelle 26.1. Vorher benötigen wir jedoch einen weiteren topologischen Grundbegriff.

Definition 26.1 (Relativ- oder Spurtopologie). *Sei (D, r) ein metrischer Raum und $\mathcal{O}_D$ die von r erzeugte Topologie sowie $E \subseteq D$. Dann heißt die von $r|_{E\times E}$ erzeugte Topologie auf E die* Relativtopologie *oder* Spurtopologie *$\mathcal{O}_E$.*

Lemma 26.2 (Charakterisierung der Spurtopologie). *Sei (D, r) ein metrischer Raum und $\mathcal{O}_D$ die von r erzeugte Topologie, $E \subseteq D$, und $\mathcal{O}_E$ die Relativtopologie auf E.*

1. *Es gilt*
$$\mathcal{O}_E = \{A \cap E : A \in \mathcal{O}_D\}.$$
2. *Weiter ist*
$$\{A \cap E : A^c \in \mathcal{O}_D\}$$
die Menge der in $\mathcal{O}_E$ abgeschlossenen Mengen.
3. *Eine Teilmenge $K \subseteq E$ ist genau dann kompakt bezüglich $\mathcal{O}_E$, wenn K kompakt bezüglich $\mathcal{O}_D$ ist.*

Beweis. 1. Für $x \in E$ und $\varepsilon > 0$ gilt $B_\varepsilon(x) \cap E = B_\varepsilon^E(x)$, wobei $B_\varepsilon^E(x)$ der Ball um x mit Radius ε bezüglich $r|_{E\times E}$ ist, und $B_\varepsilon(x)$ der Ball um x mit Radius ε bezüglich r.
$\subseteq$: Sei $B \in \mathcal{O}_E$. Dann gilt

$$B = \bigcup\{B_\varepsilon^E(x) \subseteq B\} = \bigcup\{B_\varepsilon(x) \cap E : B_\varepsilon(x) \cap E \subseteq B\} = \bigcup\{B_\varepsilon(x) : B_\varepsilon(x) \cap E \subseteq B\} \cap E.$$

Damit haben wir B als Schnitt einer in D offenen Menge und E dargestellt.
$\supseteq$: Sei $x \in A \cap E$ für ein $A \in \mathcal{O}_D$. Dann gibt es $\varepsilon > 0$ mit $B_\varepsilon(x) \subseteq A$ und damit ist $B_\varepsilon^E(x) = B_\varepsilon(x) \cap E \subseteq A \cap E$. Es folgt direkt, dass $A \cap E \in \mathcal{O}_E$.
2. Die in $\mathcal{O}_E$ abgeschlossenen Mengen sind aus Definition

$$\begin{aligned}\{B : B^c \in \mathcal{O}_E\} &= \{B : E \setminus B = A \cap E \text{ für ein } A \in \mathcal{O}_D\}\\ &= \{B : B = E \setminus A \text{ für ein } A \in \mathcal{O}_D\}\\ &= \{A^c \cap E : A \in \mathcal{O}_D\} = \{A \cap E : A^c \in \mathcal{O}_D\}.\end{aligned}$$

3. Sei K kompakt bezüglich $\mathcal{O}_E$ und $(A_i)_{i\in I}$ eine (bezüglich $\mathcal{O}_D$) offene Überdeckung von K. Dann sind nach 1. die Mengen $B_i := A_i \cap E \in \mathcal{O}_E, i \in I$. Dann ist offenbar $(B_i)_{i\in I}$ eine (in $\mathcal{O}_E$) offene Überdeckung von K. Da K bezüglich $\mathcal{O}_E$ kompakt ist, gibt es $J \Subset I$, so

dass $K \subseteq \bigcup_{j\in J} B_j$, also auch $K \subseteq \bigcup_{j\in J} B_j \subseteq \bigcup_{j\in J} A_j$. Insbesondere hat die (in $\mathcal{O}_D$) offene Überdeckung $(A_i)_{i\in I}$ eine endliche Teilüberdeckung. Damit ist K kompakt bezüglich $\mathcal{O}_D$. Die Rückrichtung zeigt man analog.

□

Für $0 \le m \le d$ bezeichnen wir im Folgenden

$$\mathbb{R}_0^m := \{\underline{x} = (x_1, ..., x_d) \in \mathbb{R}^d : x_{m+1} = \cdots = x_d = 0\} \subseteq \mathbb{R}^d. \tag{26.1}$$

Diese Menge ist bereits der Prototyp einer m-dimensionalen Untermannigfaltigkein im $\mathbb{R}^d$.

Definition 26.3 (Differenzierbare Untermannigfaltigkeit). 1. *Eine Menge $E \subseteq \mathbb{R}^d$ heißt m-dimensionale (differenzierbare)* Untermannigfaltigkeit *von $\mathbb{R}^d$, wenn es zu $\underline{x} \in E$ eine Umgebung $U_{\underline{x}}$ von $\underline{x}$ (in $\mathbb{R}^d$) und einen Diffeomorphismus $\varphi_{\underline{x}} : U_{\underline{x}} \to V \subseteq \mathbb{R}^d$ mit V offen gibt, so dass $\varphi_{\underline{x}}(E \cap U_{\underline{x}}) = \mathbb{R}_0^m \cap V$. Die Abbildung $\varphi_{\underline{x}}$ heißt dann eine* Karte *von E und $E \cap U_{\underline{x}}$ ein* Kartengebiet. *Ist $U_{\underline{x}} \supseteq E$, so heißt φ eine* globale Karte *(und E ist ein* Kartengebiet*). Eine Familie $(E \cap U_{\underline{x}})_{\underline{x}\in D}$ von Kartengebieten heißt* Atlas *von E, wenn die Familie E überdeckt. (Existiert eine globale Karte φ von E, so ist E bereits ein Atlas.)*

2. *Seien $(E, \mathcal{O})$ und $(E', \mathcal{O}')$ topologische Räume. Eine stetige, bijektive Abbildung $\underline{\gamma} : E \to E'$ mit stetiger Umkehrabbildung heißt* Homöomorphismus *von E und E'. Gibt es eine solche Abbildung, so heißen E und E' homöomorph.*

Beispiel 26.4 (Kreislinie und Ellipse). 1. Sei

$$E := \{(x, y) : x^2 + y^2 = 1\} \subseteq \mathbb{R}^2$$

die Kreislinie. Dann ist E eine 1-dimensionale Untermannigfaltigkeit.
Denn: Sei $(x, y) \in E$. OBdA ist $y \ge 0$ und wir setzen

$$U_{(x,y)} := \begin{cases} \{(x', y') : -1 < x' < 1, |\sqrt{1-(x')^2} - y'| < \varepsilon\}, & \text{falls } y \neq 0 \\ \{(x', y') : -1 < y' < 1, |\sqrt{1-(y')^2} - x'| < \varepsilon\}, & \text{falls } y = 0 \end{cases}$$

sowie

$$\varphi_{(x,y)}(x', y') = \begin{cases} (x', y' - \sqrt{1-(x')^2}), & \text{falls } y \neq 0 \\ \{(y', x' - \sqrt{1-(y')^2}), & \text{falls } y = 0 \end{cases}.$$

Dann ist $\varphi_{(x,y)}$ für alle $(x, y) \in E$ ein Diffeomorphismus. Außerdem gilt

$$\varphi_{(x,y)}(E \cap U_{(x,y)}) = (-1, 1) \times \{0\}$$

für alle $(x, y) \in E$. (Man beachte, dass dies nun auch für $y = 0$ gilt.) Insbesondere ist $\varphi_{(x,y)}$ eine Karte. Weiter ist etwa (mit einer analogen Konstruktion im Falle $y < 0$) die Familie $(E \cap U_{(1,0)}, E \cap U_{(0,1)}, E \cap U_{(-1,0)}, E \cap U_{(0,-1)})$ ein Atlas von E.

Offenbar kann man E auch anders, nämlich etwa mittels

$$E = \{(\cos(t), \sin(t) : 0 < t < 7\}$$

beschreiben.

2. Wir betrachten für $a, b > 0$ die Abbildung

$$\underline{\gamma} : (x, y) \mapsto (ax, by).$$

Dann ist (mit E wir in 1.)

$$\underline{\gamma}(E) = \left\{(x, y) : \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1\right\}$$

eine Ellipse mit den Halbachsen a und b. Diese ist homöomorph zu E, und ebenfalls eine Untermannigfaltigkeit– wie wir gleich zeigen werden – weil $\underline{\gamma}$ ein Isomorphismus ist.

Lemma 26.5 (Einfache Eigenschaften). *1. Sei $E \subseteq \mathbb{R}^d$ eine m-dimensionale Untermannigfaltigkeit, $U_{\underline{x}}$ eine Umgebung von $\underline{x} \in E$ und $\varphi_{\underline{x}} : U_{\underline{x}} \to V \subseteq \mathbb{R}^d$ mit V offen, so dass $\varphi_{\underline{x}}(E \cap U_{\underline{x}}) = \mathbb{R}_0^m \cap V$. Dann ist $\varphi_{\underline{x}}|_{E \cap U_{\underline{x}}}$ ein Homöomorphismus von $E \cap U_{\underline{x}} \subseteq E$ (versehen mit der Spurtopologie $\mathcal{O}_E$) und $\mathbb{R}_0^m \cap V \subseteq \mathbb{R}^d$ (versehen mit der euklidischen Topologie $\mathcal{O}$ auf $\mathbb{R}^d$).*

2. Sei $E \subseteq \mathbb{R}^d$ und $\iota : \mathbb{R}^d \to \mathbb{R}^d$ ein Isomorphismus. Dann ist E genau dann eine m-dimensionale Untermannigfaltigkeit, wenn $\iota(E)$ eine m-dimensionale Untermannigfaltigkeit ist.

Beweis. 1. ist klar, da der Diffeomorphismus $\varphi_{\underline{x}}$ stetig ist und eindeutig umkehrbar ist. Damit ist auch die Einschränkung $\varphi_{\underline{x}}|_{E \cap U_{\underline{x}}}$ stetig und umkehrbar.

2. Sei E eine m-dimensionale Untermannigfaltigkeit und $\underline{x} \in E$. Weiter sei $U_{\underline{x}}, \varphi_{\underline{x}}$ und V wie in Definition 26.3. Dann gilt mit $\widetilde{\varphi}_{\iota(\underline{x})} := \varphi_{\underline{x}} \circ \iota^{-1}$

$$\widetilde{\varphi}_{\iota(\underline{x})}(\iota(E) \cap \iota(U_{\underline{x}})) = \varphi_{\underline{x}} \circ \iota^{-1}(\iota(E) \cap \iota(U_{\underline{x}})) = \varphi_{\underline{x}}(E \cap U_{\underline{x}}) = \mathbb{R}_0^m \cap V.$$

Da $\iota(U_{\underline{x}})$ eine Umgebung von $\iota(\underline{x})$ ist, ist also $\iota(E)$ eine m-dimensionale Untermannigfaltigkeit. Die Umkehrung zeigt man analog. □

Beispiel 26.6 (Untermannigfaltigkeiten). 1. Laut Definition ist klar, dass jede Menge $\mathbb{R}_0^m \cap V$ für eine offene Menge $V \subseteq \mathbb{R}^d$ eine Untermannigfaltigkeit ist. (Man wähle $\varphi_{\underline{x}} = \mathrm{id}$.) Ist weiter $\iota : \mathbb{R}^d \to \mathbb{R}^d$ ein Isomorphismus, also $\iota(\underline{x}) = \underline{\underline{A}}\,\underline{x} \in \mathrm{Gl}(d, \mathbb{R})$. Dann ist $\{\underline{\underline{A}}\,\underline{x} : \underline{x} \in \mathbb{R}_0^m \cap V\}$ für jede offene Menge V nach Lemma 26.5.2 eine m-dimensionale Untermannigfaltigkeit.

2. Der Graph der Abbildung $\underline{\gamma} : t \mapsto (\sin(t), \sin(2t)), 0 \le t \le 2\pi$, $E = \{\underline{\gamma}(t), 0 \le t \le 2\pi\}$ ist keine (ein-dimensionale) Untermannigfaltigkeit (von $\mathbb{R}^2$). (Siehe auch Abbildung 26.1.) Denn: Angenommen, E wäre eine Untermannigfaltigkeit. Dann müsste es ein $\varepsilon > 0$ klein genug und eine offene Menge $V \subseteq \mathbb{R}^2$ geben, so dass $E \cap B_\varepsilon((0, 0))$ homöomorph zu $\mathbb{R}_0^1 \cap V$ ist. Dies ist jedoch offensichtlich nicht möglich.

Lemma 26.7 (Kompakte Ausschöpfung von Untermannigfaltigkeiten). *Sei $E \subseteq \mathbb{R}^d$ eine (m-dimensionale) Untermannigfaltigkeit und $\mathcal{O}_E$ deren Spurtopologie. Dann gibt es eine Folge von $K_1, K_2, ... \subseteq E$ kompakter (bezüglich $\mathcal{O}_E$) Mengen, so dass $K_i \subseteq K_{i+1}^\circ$ (wobei das Innere von K_{i+1} in der Topologie $\mathcal{O}_E$ gemeint ist) und $\bigcup_{n=1}^\infty K_i = E$.*

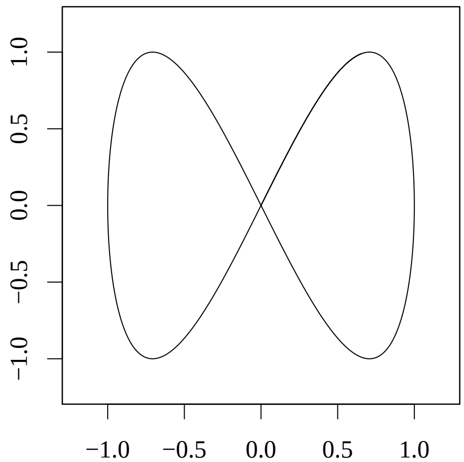

Abbildung 26.1: Der Graph der Abbildung $t \mapsto (\sin t, \sin(2t))$ ist keine ein-dimensionale Untermannigfaltigkeit des $\mathbb{R}^2$.

Beweis. Sei

$$\mathcal{B} = \{E \cap B_\varepsilon(\underline{x}) : \overline{E} \cap \overline{B}_\varepsilon(\underline{x}) \text{ kompakt in } \mathbb{R}^d, \mathbb{Q} \ni \varepsilon > 0, \underline{x} \in \mathbb{Q}^d\}.$$

Dann ist $\mathcal{B}$ offenbar eine abzählbare Basis der Spurtopologie $\mathcal{O}_E$, also können wir $\mathcal{B} = \{B_1, B_2, ...\}$ schreiben. Wir setzen dann $K_1 = \overline{B}_1$ und iterativ $K_n = \bigcup_{i=1}^{k_n} \overline{B}_i$, wobei k_n so gewählt ist, dass $K_n \subseteq K_{n+1}^\circ$. □

Proposition 26.8 (Existenz eines abzählbaren Atlas). *Sei $E \subseteq \mathbb{R}^d$ eine m-dimensionale Untermannigfaltigkeit. Dann gibt es einen abzählbaren Atlas von E, d.h. es gibt $D \subseteq E$ höchstens abzählbar, $U_{\underline{x}} \supseteq \{x\}$ offen in $\mathbb{R}^d$, so dass $U_{\underline{x}} \cap E$ ein Kartengebiet ist, $\underline{x} \in D$ und $E \subseteq \bigcup_{\underline{x} \in D} U_{\underline{x}}$.*

Beweis. Es gibt für jedes $\underline{x} \in E$ eine (in $\mathbb{R}^d$) offene Umgebung $U_{\underline{x}}$, so dass $U_{\underline{x}} \cap E$ eine Untermannigfaltigkeit mit globaler Karte ist. Wir werden nun die abzählbare Ausschöpfung von E aus Lemma 26.7 verwenden, um den abzählbaren Atlas von E zu konstruieren. Seien also $K_1, K_2, ... \subseteq E$ kompakt (bezüglich $\mathcal{O}_E$) mit $K_n \subseteq K_{n+1}^\circ$ und $\bigcup_{n=1}^\infty K_n = E$. Es ist $(U_{\underline{x}} \cap K_{n+1}^\circ)_{\underline{x} \in K_n}$ eine offene Überdeckung von K_n, die also eine endliche Teilüberdeckung $(U_{\underline{x}} \cap K_{n+1}^\circ)_{\underline{x} \in D_n}$ besitzt, $n = 1, 2, ...$ OBdA sei nun $D_n \subseteq D_{n+1}$. Mit der Hilfe dieser endlichen Teilüberdeckungen können wir eine abzählbare Überdeckung $(U_{\underline{x}})_{\underline{x} \in D}$ mit $D = \bigcup_{n=1}^\infty D_n$ konstruieren. Jede der Mengen $U_{\underline{x}} \cap E$ hat eine globale Karte $\varphi_{\underline{x}}, \underline{x} \in D$, ist also ein Kartengebiet. Somit haben wir einen abzählbaren Atlas konstruiert. □

Definition 26.9 (Zerlegung der Eins). 1. *Sei $(E, \mathcal{O})$ ein topologischer Raum und $f : E \to \mathbb{R}$. Dann ist $Tr(f) := \overline{\{f \neq 0\}}$ der* Träger *von f.*

2. *Sei $E \subseteq \mathbb{R}^d$ eine m-dimensionale Untermannigfaltigkeit. Eine* Zerlegung der Eins *von E ist eine (abzählbare) Familie von Funktionen $h_1, h_2, ... \in \mathcal{C}(E, [0, 1])$, so dass (i) es für jedes $\underline{x} \in E$ ein $I_{\underline{x}} \Subset \mathbb{N}$ gibt mit $h_i(\underline{x}) = 0$ für $i \notin I_{\underline{x}} x$, und (ii) $\sum_{i=1}^\infty h_i = 1$.*

   Sei weiter $\mathcal{B} \subseteq \mathcal{O}_E$. Dann heißt $(h_n)_{n=1,2,}$ dem Mengensystem $\mathcal{B}$ untergeordnet, *wenn es für alle n ein $B \in \mathcal{B}$ gibt mit $Tr(h_n) \subseteq B$.*

Beispiel 26.10 (Zerlegung der Eins auf der Kreislinie). Sei

$$E = \{(x,y) : x^2 + y^2 = 1\} = \{(\cos(t), \sin(t) : 0 < t < 7\}$$

die Kreislinie aus Beispiel 26.4. Wir haben bereits gesehen, dass $\mathcal{B} = \{E \cap U_{(1,0)}, E \cap U_{(0,1)}, E \cap U_{(-1,0)}, E \cap U_{(0,-1)}\}$ eine Überdeckung von E ist. Wir geben nun eine $\mathcal{B}$ untergeordnete Zerlegung der Eins auf E an. Hierzu setzen wir (wobei wir t und $t \pm 2\pi$ identifizieren)

$$\begin{aligned} h_k(\cos(t), \sin(t)) &= \cos^2(3t) \cdot 1_{t \in (k\pi/6, (k+2)\pi/6)}, \qquad & k = 1,3,5,7,9,11, \\ h_k(\cos(t), \sin(t)) &= \sin^2(3t) \cdot 1_{t \in (k\pi/6, (k+2)\pi/6)}, \qquad & k = 0,2,4,6,8,10. \end{aligned}$$

Es gilt nämlich etwa $\overline{\{h_0 > 0\}} \subseteq U_{(1,0)}$ sowie $\sum h_i = 1$ nach Konstruktion.

Lemma 26.11 (Existenz der Zerlegung der Eins bei Untermannigfaltigkeiten). *Sei $E \subseteq \mathbb{R}^d$ eine m-dimensionale Untermannigfaltigkeit und $\mathcal{B}$ eine offene Überdeckung von E. Dann gibt es eine $\mathcal{B}$ untergeordnete Zerlegung der Eins von E.*

Beweis. Wir beweisen zunächst folgende Behauptung:

> Sei $O \in \mathcal{O}_E$ und $\underline{x} \in O$. Dann gibt es $\varphi_{\underline{x},O} \in \mathcal{C}(E, [0,\infty))$ mit $\varphi_{\underline{x},O}(\underline{x}) \geq 0$ und $\{\underline{y} : \varphi_{\underline{x},O}(\underline{y}) > 0\} \subseteq O$.

Sei hierzu ε klein genug, so dass $E \cap B_\varepsilon(\underline{x}) \subseteq O$. Weiter definieren wir einfach

$$\varphi_{\underline{x},O} : \begin{cases} E & \to [0,\infty) \\ \underline{y} & \mapsto (\varepsilon - r(\underline{x}, \underline{y}))^+. \end{cases}$$

Nach Konstruktion ist dann klar, dass $\{\underline{y} : \varphi(\underline{y}) > 0\} = E \cap B_\varepsilon(\underline{x}) \subseteq O$.

Nun seien $K_1, K_2, ... \subseteq E$ kompakt, so dass $\bigcup_{n=1}^\infty K_n = E$ (siehe hierzu Lemma 26.7). Wir betrachten zunächst K_1. Für $\underline{x} \in B \in \mathcal{B}$ verwenden wir $\varphi_{\underline{x},B}$ aus obiger Behauptung. Es ist $\{\mathrm{Tr}\varphi^\circ_{\underline{x},B} : \underline{x} \in K_1, B \in \mathcal{B}\}$ eine offene Überdeckung von K_1. Da K_1 kompakt ist gibt es also J_1 endlich und $x_j \in K_1, B_j \in \mathcal{B}$, so dass $K_1 \subseteq \bigcup_{j \in J} \mathrm{Tr}\varphi^\circ_{\underline{x}_j,B_j}$.

Betrachten wir nun K_2. Wie oben ist

$$\{\mathrm{Tr}\varphi^\circ_{\underline{x}_j,B_j} : j \in J_1\} \cup \{\mathrm{Tr}\varphi^\circ_{\underline{x},B} : \underline{x} \in K_2 \setminus K_1, B \in \mathcal{B}\}$$

eine offene Überdeckung, zu er es wegen der Kompaktheit von K_2 eine endliche Teilüberdeckung gibt. Fahren wir so fort, erhalten wir eine abzählbare Menge $J \subseteq E$ und eine abzählbare Überdeckung

$$\{\mathrm{Tr}\varphi^\circ_{\underline{x}_j,B_j} : j \in J\}.$$

Weiter haben wir die Überdeckung so konstruiert, dass für jedes $\underline{y} \in E$ nur endlich viele Werte $\varphi_{\underline{x}_j,B_j}(\underline{y})$ positiv sind. Daraus folgt auch, dass

$$\varphi = \sum_{j \in J} \varphi_{\underline{x}_j,B_j}$$

stetig (insbesondere endlich) ist. Nach Konstruktion ist außerdem $\varphi > 0$. Damit folgt die Behauptung, indem wir $h_j = \frac{\varphi_{\underline{x}_j,B_j}}{\varphi}$ setzen. □

	Voraussetzungen	$E =$	Normalenraum Tangentialraum	Referenzen
Graph	$E' \subseteq \mathbb{R}^m$ offen, $\underline{f} \in \mathcal{C}^1(E', \mathbb{R}^{d-m})$	$\{(\underline{x}, \underline{f}(\underline{x})) : \underline{x} \in E'\}$	$T_{\underline{a}} = \operatorname{Im}(\underline{\underline{A}})$, $N_{\underline{a}} = \operatorname{Ker}(\underline{\underline{A}}^\top)$ für $\underline{\underline{A}} = \begin{pmatrix} \underline{\underline{E}}_m \\ \hline \underline{\underline{D}}\,\underline{f}(\underline{x}) \end{pmatrix}$	Prop. 26.12, Prop. 26.29
Niveaumenge	$E' \subseteq \mathbb{R}^d$ offen, $\underline{f} \in \mathcal{C}^1(E', \mathbb{R}^n)$, $\underline{c} \in \underline{f}(E')$, $\underline{\underline{D}}\,\underline{f}(\underline{a}) \in \mathbb{R}^{n\times d}$ für alle $\underline{a} \in E$ surjektiv, $m + n = d$	$\underline{f}^{-1}(\underline{c})$	$T_{\underline{a}} = \operatorname{Ker}(\underline{\underline{D}}\,\underline{f}(\underline{a}))$ $N_{\underline{a}} = \operatorname{Im}((\underline{\underline{D}}\,\underline{f}(\underline{a}))^\top)$	Prop. 26.14, Prop. 26.31
Einbettung	$E' \subseteq \mathbb{R}^m$ offen, $\underline{\gamma} \in \mathcal{C}^1(E', \mathbb{R}^d)$, $\underline{\underline{D}}\,\underline{\gamma}(\underline{x})$ für alle $\underline{x} \in E'$ injektiv, E' und $\underline{\gamma}(E')$ homöomoprh	$\underline{\gamma}(E')$	$T_{\underline{a}} = \operatorname{Im}(\underline{\underline{D}}\,\underline{\gamma}(\underline{x}))$, $N_{\underline{a}} = \operatorname{Ker}(\underline{\underline{D}}\,\underline{\gamma}(\underline{x})^\top)$ für $\underline{a} = \gamma(\underline{x})$	Prop. 26.25, Prop. 26.33

Tabelle 26.1: Charakterisierungen von Untermannigfaltigkeiten. Siehe auch Theorem 26.26.

26.2 Charakterisierungen

In Definition 26.3 haben wir definiert, was eine Untermannigfaltigkeit ist. Nun lernen wir kennen, in welcher Form Untermannigfaltigkeiten auftreten können, nämlich als Graphen, Niveaumengen und durch Einbettungen. Dies ist in Tabelle 26.1 zusammen gefasst; siehe auch Theorem 26.26.

Proposition 26.12 (Graphen von Funktionen als Untermannigfaltigkeiten). *Sei $m + n = d$, $E' \subseteq \mathbb{R}^m$ offen und $\underline{f} \in \mathcal{C}^1(E', \mathbb{R}^n)$ stetig differenzierbar. Dann ist $E := \{(\underline{x}, \underline{f}(\underline{x})) : \underline{x} \in E'\} \subseteq \mathbb{R}^d$ eine m-dimensionale Untermannigfaltigkeit des $\mathbb{R}^d$ und*

$$\varphi : \begin{cases} E' \times \mathbb{R}^n & \to \mathbb{R}^d \\ (\underline{x}, \underline{y})) & \mapsto (\underline{x}, \underline{f}(x) - \underline{y}) \end{cases}$$

ist eine globale Karte von E. Insbesondere ist E auch ein Kartengebiet.

Beweis. Sei $\underline{x} \in E'$, also $(\underline{x}, \underline{f}(\underline{x})) \in E$ und φ wie angegeben. Dann ist

$$\varphi(E \cap U) = \varphi(E) = \{(\underline{x}, \underline{f}(\underline{x}) - \underline{f}(\underline{x})) : \underline{x} \in E'\} = E' \times \{0\}^n = \mathbb{R}^m_0 \cap V.$$

Daraus folgen alle Behauptungen. $\square$

Beispiel 26.13 (Kreislinie 1). Sei $m = n = 1, d = 2$, $E' = (-1,1)$ und $f : x \mapsto \sqrt{1-x^2}$. Dann ist offenbar $\{(x, \sqrt{1-x^2}) : x \in E'\}$ eine Untermannigfaltigkeit, nämlich der Teil einer Kreislinie im $\mathbb{R}^2$.

Proposition 26.14 (Niveaumengen als Untermannigfaltigkeiten). *Sei $m + n = d$, $E' \subseteq \mathbb{R}^d$ offen und $\underline{f} \in \mathcal{C}^1(E', \mathbb{R}^n)$ stetig differenzierbar. Weiter sei $E := \underline{f}^{-1}(\underline{c}) \subseteq \mathbb{R}^d$ für $\underline{c} \in \mathbb{R}^n$ die Niveaumenge von $\underline{c}$ unter $\underline{f}$ nicht leer, d.h. $E \neq \emptyset$. Ist $\underline{\underline{D}}\,\underline{f}(\underline{a}) \in \mathbb{R}^{n\times d}$ für alle $\underline{a} \in E$ surjektiv (d.h. rg$(\underline{\underline{D}}\,\underline{f}(\underline{a})) = n$), dann ist E eine m-dimensionale Untermannigfaltigkeit.*

Beweis. Sei $\underline{a} \in E$ und $X_0 = \mathrm{Ker}(\underline{\underline{D}}\,\underline{f}(\underline{a}))$. Wegen der Surjektivität von $\underline{\underline{D}}\,\underline{f}(\underline{a}) \in \mathbb{R}^{n\times d}$ ist $\dim(X_0) = d - n$ und es gibt $\underline{y}_1, ..., \underline{y}_n \in \mathbb{R}^d$ mit $Y_0 = \mathrm{span}(\underline{y}_1, ..., \underline{y}_n)$ und $\mathbb{R}^d = \mathrm{span}(X_0, Y_0)$. Dann ist sowohl $\underline{\underline{D}}\,\underline{f}(\underline{a})|_{Y_0}$ als auch $\iota : X_0 \times Y_0 \to \mathbb{R}^d, (\underline{x}, \underline{y}) \mapsto \underline{x} + \underline{y}$ ein Isomorphismus. Wir betrachten $\underline{g} := \underline{f} \circ \iota$. Es gilt mit der Kettenregel

$$\underline{\underline{D}}\,\underline{g}(\underline{x},\underline{y}) = \underline{\underline{D}}\,\underline{f}(\iota(\underline{x},\underline{y})) \cdot \underline{\underline{D}}\,\iota(\underline{x},\underline{y}),$$

also für $(\underline{h}, \underline{k}) \in X_0 \times Y_0$

$$\underline{\underline{D}}\,\underline{g}(\underline{x},\underline{y})(\underline{h},\underline{k}) = \underline{\underline{D}}\,\underline{f}(\underline{x}+\underline{y})(\underline{h}+\underline{k}).$$

Insbesondere gilt damit für $\underline{a} = \underline{x} + \underline{y}$

$$\underline{\underline{D}}_{\underline{Y}_0}\underline{g}(\underline{x},\underline{y})\underline{k} = \underline{\underline{D}}\,\underline{g}(\underline{x},\underline{y})(\underline{0},\underline{k}) = \underline{\underline{D}}\,\underline{f}(\underline{a})\underline{k}.$$

Da $\underline{\underline{D}}\,\underline{f}(\underline{a})|_{Y_0}$ ein Isomorphismus ist, ist $\underline{\underline{D}}_{\underline{Y}_0}\underline{g}(\underline{x},\underline{y}) : Y_0 \to \mathbb{R}^{d-n}$ invertierbar. Damit gibt es nach dem Satz über implizite Funktionen, Theorem 16.20, eine Umgebung $V \times V'$ von $(\underline{x}_0, \underline{y}_0)$ und eine Funktion $\underline{h} \in \mathcal{C}^1(V, V')$, so dass

$$\{(\underline{x},\underline{y}) \in V \times V' : \underline{g}(\underline{x},\underline{y}) = \underline{c}\} = \{(\underline{x}, \underline{h}(\underline{x})) : \underline{x} \in V\}.$$

Insbesondere ist $(E \cap V) \times V'$ der Graph von $\underline{h}|_V$, nach Proposition 26.12 also eine Untermannigfaltigkeit. □

Die Eigenschaft, die $\underline{x}$ bzw. $\underline{c}$ in der letzten Proposition erfüllen muss, fassen wir in folgender Definition zusammen.

Definition 26.15 (Regulärer Wert). *Sei $E \subseteq \mathbb{R}^d$ und $\underline{f} \in \mathcal{C}^1(E, \mathbb{R}^n)$. Dann heißt $\underline{x} \in E$* regulärer Punkt *von $\underline{f}$, wenn $\underline{\underline{D}}\,\underline{f}(\underline{x})$ surjektiv ist. Andernfalls heißt $\underline{x}$ ein* singulärer Punkt. *Sind für ein $\underline{c} \in \underline{f}(E)$ alle $\underline{x} \in \underline{f}^{-1}(\underline{c})$ reguläre Punkte, so heißt $\underline{c}$ ein* regulärer Wert *von $\underline{f}$. Andernfalls heißt $\underline{c}$ ein* singulärer Wert *von $\underline{f}$.*

Nun können wir die Aussage von Proposition 26.14 mit Hilfe von regulären Werten einfacher formulieren.

Korollar 26.16 (Spezialfall regulärer Werte). *Sei $m + n = d$, $E' \subseteq \mathbb{R}^d$ offen und $\underline{f} \in \mathcal{C}^1(E', \mathbb{R}^n)$. Ist $\underline{c} \in \mathbb{R}^n$ ein regulärer Wert von $\underline{f}$, so ist $\underline{f}^{-1}(\underline{c})$ (falls diese Menge nicht leer ist) eine m-dimensionale Untermannigfaltigkeit.*

Beispiel 26.17 (Kreislinie 2). Sei $m = n = 1, d = 2$, $E' = (-1,1)$ und $f : (x,y) \mapsto x^2 + y^2$ sowie $c = 1$ und $E = \{(x,y) : x^2 + y^2 = 1\}$. Dann ist $\underline{\nabla} f(x,y) = (2x, 2y)$, also insbesondere für $(x,y) \in E$ surjektiv. Damit ist E eine Untermannigfaltigkeit, nämlich die Kreislinie in $\mathbb{R}^2$.

Untermannigfaltigkeiten können nach eben gesagtem als Graph einer Funktion oder als Niveaumenge dargestellt werden. Wir kommen nun zu einer dritten Möglichkeit durch eine reguläre Parameterdarstellung. Diese verallgemeinert die Darstellung einer Untermannigfaltigkeit durch einen Graphen.

Definition 26.18 (Reguläre Parameterdarstellung). *Sei $E' \subseteq \mathbb{R}^m$ offen und $\underline{\gamma} \in \mathcal{C}^1(E', \mathbb{R}^d)$. Ist $\underline{\underline{D}}\,\underline{\gamma}(\underline{x}) \in \mathbb{R}^{d\times m}$ für alle $\underline{x} \in E'$ injektiv (d.h. $\mathrm{rg}(\underline{\underline{D}}\,\underline{\gamma}(\underline{x})) = m$), so heißt $\underline{\gamma}$* Immersion *oder* reguläre Parameterdarstellung*. Weiter heißt dann E' der* Parameterbereich *von $\underline{\gamma}$ und $\underline{\gamma}(E')$ die* Spur *von $\underline{\gamma}$.*

Zwei Immersionen $\underline{\gamma}_1 \in \mathcal{C}^1(E_1', \mathbb{R}^d)$ und $\underline{\gamma}_2 \in \mathcal{C}^1(E_2', \mathbb{R}^d)$ heißen äquivalent, wenn es einen Diffeomorphismus $\underline{\psi} : E_1' \to E_2'$ gibt mit $\underline{\gamma}_1 = \underline{\gamma}_2 \circ \underline{\psi}$.

Bemerkung 26.19 (Kurven). In Definition 17.1 haben wir $\mathcal{C}^1$-Kurven als Abbildungen $\underline{\gamma} \in \mathcal{C}^1(I, \mathbb{R}^d)$ für ein Intervall $I \subseteq \mathbb{R}$ definiert. Definition 26.18 verallgemeinert diesen Begriff auf mehr-dimensionale Obkjekte.

Die Bedingung $\mathrm{rg}(\underline{\underline{D}}\,\underline{\gamma}(\underline{x})) = 1$ bedeutet für Kurven, dass $\underline{\gamma}'(t) \neq 0$ gilt. Um zu verstehen, was das bedeutet, betrachten wir die Kurve

$$\underline{\gamma} : \begin{cases} I := (-1,1) & \to \mathbb{R}^2 \\ t & \mapsto (t^3, |t|^3) \end{cases}.$$

Offenbar gilt $\underline{\gamma}'(t) = (3t^2, 3t|t|)$, also ist $\underline{\gamma} \in \mathcal{C}^1(I, \mathbb{R}^2)$. Allerdings ist $\underline{\gamma}'(0) = (0,0)$ und somit ist $\underline{\gamma}$ keine Immersion. Weiter ist

$$\underline{\gamma}(I) = \{(t, |t|) : -1 < t < 1\},$$

also wohl keine 1-dimensionale Untermannigfaltigkeit des $\mathbb{R}^2$. Die Bedingung $\mathrm{rg}(\underline{\underline{D}}\,\underline{\gamma}(\underline{x})) = m$ in Definition 26.18 ist also essentiell dafür, dass $\underline{\gamma}(E')$ eine Untermannigfaltigkeit ist.

Bemerkung 26.20 (Graphen als reguläre Parameterdarstellungen). Wir haben bereits in Proposition 26.12 reguläre Parameterdarstellungen kennen gelernt. Ist nämlich $\underline{f} \in \mathcal{C}^1(E', \mathbb{R}^n)$ für $E' \subseteq \mathbb{R}^m$ offen (und $m+n = d$), so setzen wir $\underline{\gamma}(\underline{x}) = (\underline{x}, \underline{f}(\underline{x})) \in \mathbb{R}^d$. Damit ist $(\underline{\underline{D}}\,(\gamma_i(\underline{x}))_{i=1,\ldots,m} = \underline{\underline{E}}_m$, also insbesondere $\mathrm{rg}(\underline{\underline{D}}\,\underline{\gamma}(\underline{x})) = m$. Damit definiert jeder Graph eine Immersion, und der Graph stimmt mit der Spur der Immersion überein. Insbesondere ist die Spur der Immersion ein Kartengebiet.

Da wir bereits wissen, dass Graphen Untermannigfaltigkeiten sind, liegt es nahe, dass auch Spuren von Immersionen ebenfalls Untermannigfaltigkeiten sind. Dies gilt jedoch nur lokal, wie wir nun sehen werden.

Lemma 26.21 (Immersionen und Untermannigfaltigkeiten). *Sei $E' \subseteq \mathbb{R}^m$ offen und $\underline{\gamma} \in \mathcal{C}^1(E', \mathbb{R}^d)$ eine Immersion. Dann gibt es für jedes $\underline{x} \in E'$ eine Umgebung $U_{\underline{x}}' \subseteq E'$ von $\underline{x}$, eine Permutation $P : \mathbb{R}^d \to \mathbb{R}^d$ der Koordinaten, $V \subseteq \mathbb{R}^m$ offen und einen Diffeomorphismus $\underline{\varphi}_{\underline{x}} : U_{\underline{x}}' \to V$ mit $P \circ \underline{\gamma}|_{U_{\underline{x}}'} = \underline{\gamma}^* \circ \underline{\varphi}_{\underline{x}}$ für ein $\underline{\gamma}^* \in \mathcal{C}^1(V, \mathbb{R}^d)$ und*

$$\underline{\gamma}^*(\underline{y}) = (\underline{y}, \gamma_{n+1}^*(\underline{y}), ..., \gamma_d^*(\underline{y})).$$

Insbesondere sind $\underline{\gamma}$ und $\underline{\gamma}^*$ äquivalent und $\underline{\gamma}(U_{\underline{x}}')$ ist eine m-dimensionale Untermannigfaltigkeit mit globaler Karte.

Beweis. Wir ordnen die Koordinaten so um, dass die ersten m Zeilen von $\underline{\underline{D}}\,\underline{\gamma}(\underline{x}) \in \mathbb{R}^{d\times m}$ linear unabhängig sind. Betrachten wir also die Abbildung $\underline{h} = (\gamma_1, ..., \gamma_m) \in \mathcal{C}^1(E', \mathbb{R}^m)$, so hat deren Jacobi-Matrix vollen Rang, ist also nach Theorem 16.15 lokal umkehrbar. Also gibt es eine Umgebung $U'_{\underline{x}}$ von $\underline{x}$, so dass $V := \underline{h}(U'_{\underline{x}})$ eine Umgebung von $\underline{h}(\underline{x}) \in \mathbb{R}^m$ ist und $\underline{\psi} := \underline{h}|_{U'_{\underline{x}}} \in \mathcal{C}^1(U'_{\underline{x}}, V)$ ein Diffeomorphismus.

Wir setzen nun $\underline{\gamma}^* := \underline{\gamma} \circ \underline{\psi}^{-1} \in \mathcal{C}^1(V, \mathbb{R}^d)$ (also $\underline{\psi} \circ \underline{\gamma}^* = \underline{\gamma}$ wie gefordert) und schreiben

$$\underline{\underline{D}}\,\underline{\gamma}^*(\underline{x}) = \underline{\underline{D}}\,\underline{\gamma}(\underline{\psi}^{-1}(\underline{x})) \cdot \underline{\underline{D}}\psi^{-1}(\underline{x}) = \underline{\underline{D}}\,\underline{\gamma}(\underline{\psi}^{-1}(\underline{x})) \cdot (\underline{\underline{D}}\psi(\underline{\psi}^{-1}(\underline{x})))^{-1},$$

woraus die Injektivität von $\underline{\underline{D}}\,\underline{\gamma}^*(\underline{x})$ folgt. Weiter ist

$$(\underline{\underline{D}}\,\underline{\gamma}^*(\underline{x})_i)_{i=1,...,m} = \underline{\underline{D}}\,\underline{\psi}(\underline{\psi}^{-1}(\underline{x})) \cdot (\underline{\underline{D}}\psi(\underline{\psi}^{-1}(\underline{x})))^{-1} = \underline{\underline{E}}_m,$$

womit $\underline{\gamma}^*$ die gewünschte Form hat.

An der Form von $\underline{\gamma}^*$ sehen wir, dass

$$\underline{\gamma}(U'_{\underline{x}}) = \underline{\gamma}^*(V) = \{(\underline{y}, \underline{f}(\underline{y})) : \underline{y} \in V\}$$

für $\underline{f} = (\gamma_{n+1}, ..., \gamma_d)$. Insbesondere ist $\underline{\gamma}(U'_{\underline{x}})$ der Graph einer Funktion und damit nach Proposition 26.12 eine Untermannigfaltigkeit mit globaler Karte, also auch ein Kartengebiet. □

Beispiel 26.22 (Spur einer Immersion muss keine Untermannigfaltigkeit sein). In Lemma 26.21 haben wir gesehen, dass für eine Immersion $\underline{\underline{\gamma}} : E' \subseteq \mathbb{R}^m \to \mathbb{R}^d$ die Spur $\gamma(E')$ zumindest lokal eine Untermannigkaltigkeit ist. (Schließlich war ja nur behauptet, dass $\underline{\gamma}(U_{\underline{x}})$ für eine Umgebung $U_{\underline{x}}$ von jedem $\underline{x} \in E'$, nicht jedoch, dass $\underline{\gamma}(E')$ eine Untermannigfaltigkeit ist.) Dies ist auch nicht notwendigerweise der Fall wie wir bereits in Beispiel 26.6.2 gesehen haben.

Definition 26.23 (Einbettung). *Sei $E' \subseteq \mathbb{R}^m$ offen und $\underline{\gamma} \in \mathcal{C}^1(E', \mathbb{R}^d)$ eine Immersion. Dann heißt $\underline{\gamma}$ eine* Einbettung*, wenn E' und $\underline{\gamma}(E')$ homöomorph sind.*

Lemma 26.24 (Kartengebiete sind Spuren von Einbettungen). *Sei $E \cap U$ (mit $U \subseteq \mathbb{R}^d$ offen) ein Kartengebiet einer m-dimensionalen Untermannigfaltigkeit $E \subseteq \mathbb{R}^d$. Dann gibt es $E' \subseteq \mathbb{R}^m$ offen und eine Einbettung $\underline{\gamma} \in \mathcal{C}^1(E', \mathbb{R}^d)$ mit $\underline{\gamma}(E') = E \cap U$. Insbesondere sind offene Untermannigfaltigkeiten mit einer globalen Karte Spuren von Einbettungen.*

Beweis. OBdA sei $\varphi : U \to V \subseteq \mathbb{R}^d$ eine Karte (insbesondere ein Diffeomoprphimsus), also $\varphi(E \cap U) = \mathbb{R}^m_0 \cap V$. Wir schreiben nun

$$\mathbb{R}^m_0 \cap V = E' \times \{0\}^{d-m}$$

für ein $E' \subseteq \mathbb{R}^m$ offen. Weiter setzen wir

$$\underline{\gamma} : \begin{cases} E' & \to \mathbb{R}^d \\ \underline{x} & \mapsto \varphi^{-1}(\underline{x}, \underline{0}). \end{cases}.$$

Da φ ein Diffeomorphismus ist, ist $\underline{\gamma}$ ein Homöomorphismus. Weiter ist für $\underline{y} \in E \cap U$ und $\varphi(\underline{y}) = (\underline{x}, \underline{0})$

$$\underline{\underline{D}}\,\underline{\gamma}(\underline{x}) = \underline{\underline{D}}_{\underline{x}}\,\underline{\varphi}^{-1}(\underline{x}, \underline{0}).$$

Letztere Matrix ist die $d\times m$-Matrix, die aus den ersten m Spalten der $d\times d$-Matrix $\underline{\underline{D}}\,\underline{\varphi}^{-1}(\underline{x}, \underline{0})$ entsteht. Da $\underline{\varphi}$ ein Diffeomorphismus ist, müssen alle Spalten von $\underline{\underline{D}}\,\underline{\varphi}^{-1}(\underline{x}, \underline{0})$ linear unabhängig sein, also ist $\underline{\gamma}$ eine Immersion und damit eine Einbettung. □

Proposition 26.25 (Spur einer Einbettung als Untermannigfaltigkeit). *Sei $E' \subseteq \mathbb{R}^m$ offen, $\underline{\gamma} \in \mathcal{C}^1(E', \mathbb{R}^d)$ eine Einbettung und $E := \underline{\gamma}(E')$. Dann gilt:*

1. *E ist eine m-dimensionale Untermannigfaltigkeit.*
2. *Ist $\widehat{E}' \subseteq \mathbb{R}^m$ offen, $\widehat{\underline{\gamma}} \in \mathcal{C}^1(\widehat{E}', \mathbb{R}^d)$ eine weitere Einbettung mit $\underline{\gamma}(E') = \widehat{\underline{\gamma}}(\widehat{E}')$, so sind $\underline{\gamma}$ und $\widehat{\underline{\gamma}}$ äquivalent.*

Beweis. 1. Sei $E = \underline{\gamma}(E')$ und $\underline{a} \in E'$ mit $\underline{\gamma}(\underline{a}) = \underline{x} \in E \subseteq \mathbb{R}^d$. Es genügt zu zeigen, dass es eine Umgebung $U_{\underline{x}}$ von $\underline{x}$ gibt, so dass $U_{\underline{x}} \cap E$ eine Untermannigfaltigkeit ist.

Wie in Lemma 26.21 sei $U'_{\underline{a}}$ eine Umgebung von $\underline{a}$. Da $\underline{\gamma}$ ein Homöomorphismus ist, muss auch $V' := \underline{\gamma}(U'_{\underline{a}}) \supseteq \{\underline{x}\}$ offen in $\mathcal{O}_{\underline{\gamma}(E')}$ sein, also gibt es eine (in $\mathbb{R}^d$) offene Menge $U_{\underline{x}}$, so dass $V' = U_{\underline{x}} \cap E$. Insbesondere ist $U_{\underline{x}}$ eine Umgebung von $\underline{x}$ (in $\mathbb{R}^d$). Weiter ist $\underline{\gamma}(U'_{\underline{a}}) = U_{\underline{x}} \cap E$ nach Lemma 26.21 eine Untermannigfaltigkeit mit globaler Karte.
2. Sei nun $\widehat{\underline{\gamma}}$ eine weitere Einbettung mit $\underline{\gamma}(E') = \widehat{\underline{\gamma}}(\widehat{E}')$. Wir setzen $\underline{\varphi} = \underline{\gamma}^{-1} \circ \widehat{\underline{\gamma}} : \widehat{E}' \to E'$ und bemerken, dass $\underline{\varphi}$ bijektiv ist. Als Verknüpfung von stetigen Abbildungen ist diese Abbildung stetig. Wir müssen zeigen, dass $\underline{\varphi}$ sogar stetig differenzierbar ist. Analog folgt dann, dass auch $\underline{\psi}^{-1}$ stetig differenzierbar ist. Hierzu müssen wir zeigen, dass $\underline{\gamma}^{-1}$ stetig differenzierbar ist.

Nach Lemma 26.21 gilt: Für $\underline{a} \in E'$ gibt es eine Umgebung $U_{\underline{a}}$ und eine offene Menge $V \subseteq \mathbb{R}^m$ und eine Permutation der Koordinaten P (oBdA setzen wir $P = \mathrm{id}$) und einen Diffeomorphismus $\underline{\psi} : U_{\underline{a}} \to V$ mit $\underline{\gamma}|_{U_{\underline{a}}} = \underline{\gamma}^* \circ \underline{\psi}$ und $\underline{\gamma}^* \in \mathcal{C}^1(V, \mathbb{R}^d)$ wie im Lemma, also $\underline{\gamma}^*(V) = \gamma(U_{\underline{a}})$. Für die Projektion $\pi : \mathbb{R}^d \to \mathbb{R}^m$ auf die ersten m Koordinaten gilt für $\underline{y} \in V$, dass $\pi \circ \underline{\gamma}^*(\underline{y}) = \underline{y}$ oder $(\underline{\gamma}^*)^{-1}(\underline{x}) = \pi(\underline{x})$ für $\underline{x} \in \underline{\gamma}(U_{\underline{a}})$. Damit ist $\underline{\gamma}^{-1} = \underline{\psi}^{-1} \circ (\underline{\gamma}^*)^{-1}(\underline{x}) = \underline{\psi}^{-1} \circ \pi(\underline{x})$ für $\underline{x} \in U_{\underline{a}}$. Dies ist offenbar eine stetig differenzierbare Abbildung. □

Wir fassen nun Proposition 26.12, Proposition 26.14, Lemma 26.24 und Proposition 26.25 zusammen.

Theorem 26.26 (Charakterisierungen von Untermannigfaltigkeiten). *Sei $m + n = d$ und $E \subseteq \mathbb{R}^d$. Es sind äquivalent:*

1. *E ist eine m-dimensionale Untermannigfaltigkeit.*
2. *Für jedes $\underline{x} \in E$ gibt es eine Umgebung $U_{\underline{x}} \subseteq \mathbb{R}^d$, sowie $E'_{\underline{x}} \subseteq \mathbb{R}^m$ und eine Einbettung $\underline{\gamma}_{\underline{x}} \in \mathcal{C}^1(E'_{\underline{x}}, \mathbb{R}^d)$ mit $E \cap U_{\underline{x}} = \underline{\gamma}(E'_{\underline{x}})$.*
3. *Für jedes $\underline{x} \in E$ gibt es eine Umgebung $U_{\underline{x}} \subseteq \mathbb{R}^d$ sowie $E'_{\underline{x}} \subseteq \mathbb{R}^d$ offen und $\underline{f}_{\underline{x}} \in \mathcal{C}^1(E'_{\underline{x}}, \mathbb{R}^n)$ mit $\underline{\underline{D}}\, \underline{f}_{\underline{x}}(\underline{a}) \in \mathbb{R}^{n \times d}$ für alle $\underline{a} \in E'_{\underline{x}}$ surjektiv und $E \cap U_{\underline{x}} = \{\underline{f}_{\underline{x}} = \underline{0}\}$.*
4. *Für jedes $\underline{x} \in E$ gibt es eine Umgebung $U_{\underline{x}} \subseteq \mathbb{R}^d$, eine Permutation der Koordinaten P, $E'_{\underline{x}} \subseteq \mathbb{R}^m$ offen und*

$$U_{\underline{x}} \cap E = P(\{(\underline{y}, \underline{h}(\underline{y})) : \underline{y} \in E'_{\underline{x}}\}$$

Beweis. Zunächst stellen wir fest, dass die Eigenschaft der Menge E eine Untermannigfaltigkeit zu sein eine lokale ist, d.h. E ist genau dann eine Untermannigfaltigkeit, wenn $E \cap U$ für jede offene Menge U mit $U \cap E \neq \emptyset$ gilt.
1.⇔2.: Siehe Lemma 26.24 und Proposition 26.25.

1.⇒3.: Sei also $E \subseteq \mathbb{R}^d$ eine m-dimensionale Untermannigfaltigkeit und $\underline{x} \in E$. Nach Definition gibt es eine Umgebung $U_{\underline{x}} \subseteq \mathbb{R}^d$ von $\underline{x}$ und einen Diffeomorphismus $\varphi_{\underline{x}} : U_{\underline{x}} \to V \subseteq \mathbb{R}^d$ mit $\varphi_{\underline{x}}(E \cap U_{\underline{x}}) = \mathbb{R}^m_0 \cap V$. Da $\varphi_{\underline{x}}$ ein Diffeomorphismus ist, gilt $\underline{\underline{D}}\,\varphi_{\underline{x}}(\underline{y}) \in Gl(d, \mathbb{R})$ für alle $\underline{y} \in U_{\underline{x}}$. Bezeichne mit $\pi_n : \mathbb{R}^d \to \mathbb{R}^n$ die Projektion auf die letzten n Koordinaten sowie $\underline{f}_{\underline{x}} := \pi_n \circ \varphi_{\underline{x}} \in \mathcal{C}^1(U_{\underline{x}}, \mathbb{R}^n)$. Dann gilt $\underline{\underline{D}}\,\underline{f}_{\underline{x}}(\underline{y}) = \big(D_i\,(\varphi_{\underline{x}})_j(\underline{y})\big)_{i=1,\ldots,d;j=m+1,\ldots,d} \in \mathbb{R}^{n\times d}$. Da $\underline{\underline{D}}\,\varphi_{\underline{x}}$ vollen Rang hat, ist $\mathrm{rg}(\underline{\underline{D}}\,\underline{f}_{\underline{x}}(\underline{y}))) = n$, d.h. $\underline{\underline{D}}\,\underline{f}_{\underline{x}}(\underline{y})$ ist surjektiv für alle $\underline{y} \in U_{\underline{x}}$. Weiter gilt

$$\{\underline{f}_{\underline{x}} = \underline{0}\} = \{\pi_n \circ \varphi_{\underline{x}} = \underline{0}\} = \{\underline{\varphi}_{\underline{x}} \in \mathbb{R}^m_0\} = \{\underline{\varphi}_{\underline{x}} \in \mathbb{R}^m_0 \cap V\} = E \cap U_{\underline{x}}.$$

3.⇒4.: Das ist klar nach Proposition 26.14. In der Tat haben wir dort im Beweis die Eigenschaft 4. gezeigt, und erst daraus geschlossen dass E eine Untermannigfaltigkeit ist. (Alternativ zeigt man 2.⇒4. mit Hilfe von Lemma 26.21.)
4.⇒1.: Das ist klar nach Proposition 26.12. □

26.3 Tangential- und Normalenräume

Um Untermannigfaltigkeiten beschreiben zu können, führen wir nun die Tangential- und Normalenräume ein. In einem Punkt $\underline{x} \in E$ einer M-dimensionalen Untermannigfaltigkeit ist der Tangentialraum der (m-dimensionale) Vektorraum aller Vektoren, die tangential zur Untermannigfaltigkeit verlaufen. Das orthogonale Komplement heißt Normalenraum.

Definition 26.27 (Tangential- und Normalenraum). *Sei $E \subseteq \mathbb{R}^d$. Ein Vektor $\underline{x} \in E$ heißt* Tangentialvektor *an E im Punkt $\underline{a} \in E$, wenn es ein $\varepsilon > 0$ und ein $\underline{\gamma} \in \mathcal{C}^1((-\varepsilon, \varepsilon), E)$ gibt mit*

$$\underline{\gamma}(0) = \underline{a}, \qquad \underline{\gamma}'(0) = \underline{x}.$$

Weiter bezeichnen wir mit

$$T_{\underline{a}}E := \{\underline{x} \in E \text{ Tangentialvektor an } E \text{ in } \underline{a}\}$$

den Tangentialkegel *an E in $\underline{a}$. Ist dies ein Vektorraum, heißt er auch* Tangentialraum. *In diesem Fall heißt*

$$N_{\underline{a}}E := (T_{\underline{a}}E)^{\perp},$$

das orthogonale Komplement von $T_{\underline{a}}E$, auch Normalenraum *an E in $\underline{a}$.*

Ist E eine Untermannigfaltigkeit, stellt sich heraus, dass $T_{\underline{a}}E$ immer ein Vektorraum ist, der Normalenraum also existiert.

Lemma 26.28 (Tangentialraum ist ein Vektorraum). *Ist $E \subseteq \mathbb{R}^d$ eine m-dimensionale Untermannigfaltigkeit und $\underline{a} \in E$, so ist $T_{\underline{a}}E$ ein m-dimensionaler $\mathbb{R}$-Vektorraum.*

Beweis. Wir zeigen die Behauptung zunächst im Fall $E = \mathbb{R}^m_0 \cap V$ für $V \subseteq \mathbb{R}^d$ offen. In diesem Fall ist für $\underline{a} \in E$

$$T_{\underline{a}}E = \mathbb{R}^m_0.$$

Ist nämlich $\underline{x} \in T_{\underline{a}}E$, so gibt es eine Kurve $\underline{\gamma} : (-\varepsilon, \varepsilon) \to E$, also $\underline{x} = \underline{\gamma}'(0) \in \mathbb{R}^m_0$. Ist andersherum $\underline{x} \in \mathbb{R}^m_0$, so wählen wir $\underline{\gamma}(t) = \underline{a} + t\underline{x}$ und sehen damit $\underline{x} = \underline{\gamma}'(t) \in T_{\underline{a}}E$.

Nun zum allgemeinen Fall. Hier gibt es eine Umgebung U von $\underline{a}$, $V \subseteq \mathbb{R}^d$ offen und einen Diffeomorphismus $\underline{\varphi} : U \to V$ mit $\underline{\varphi}(E \cap U) = \mathbb{R}^m_0 \cap V$. Wir zeigen nun

$$T_{\underline{a}}(E \cap U) = (\underline{\underline{D}}\,\underline{\varphi}(\underline{a}))^{-1}(\mathbb{R}^m_0),$$

woraus die Behauptung folgt. Ist nämlich $\underline{x} \in T_{\underline{a}}(E \cap U)$ und $\underline{\gamma} : (-\varepsilon, \varepsilon) \to E$ mit $\underline{\gamma}(0) = \underline{a}, \underline{\gamma}'(0) = \underline{x}$, so ist $\underline{\underline{D}}\,\underline{\varphi}(\underline{a})\underline{x} = \underline{\underline{D}}\,\underline{\varphi}(\underline{\gamma}(0))\underline{\gamma}'(0) = \underline{\underline{D}}\,(\underline{\varphi} \circ \underline{\gamma})(0) \in \mathbb{R}^m_0$, da $\underline{\varphi} \circ \underline{\gamma}(0) \in \underline{\varphi}(E \cap U) \subseteq \mathbb{R}^m_0$. Ist andererseits $\underline{x} \in (\underline{\underline{D}}\,\underline{\varphi}(\underline{a}))^{-1}(\mathbb{R}^m_0)$, so setzen wir $\underline{\gamma}(t) := \underline{\varphi}^{-1}(\underline{\varphi}(\underline{a}) + t\underline{\underline{D}}\,\underline{\varphi}(\underline{x})) \in \underline{\varphi}^{-1}(\mathbb{R}^m_0 \cap V) = E \cap U$. Weiter ist $\underline{\gamma}(0) = \underline{a}$ und $\underline{\gamma}'(0) = \underline{\underline{D}}\,\underline{\varphi}^{-1}(\varphi(a))\underline{\underline{D}}\,\underline{\varphi}(\underline{x}) = (\underline{\underline{D}}\,\underline{\varphi}(a))^{-1}\underline{\underline{D}}\,\underline{\varphi}(\underline{x}) = \underline{x}$, also $\underline{x} \in T_{\underline{a}}(E \cap U)$. □

Proposition 26.29 (Tangential- und Normalenraum bei Graphen). *Sei $E' \subseteq \mathbb{R}^m$ offen, $\underline{f} \in \mathcal{C}^1(E', \mathbb{R}^{d-m})$ und $E = \{(\underline{x}, \underline{f}(\underline{x})) : \underline{x} \in E'\}$ die m-dimensionale Untermannigfaltigkait aus Proposition 26.12. Dann ist für $\underline{a} = (\underline{x}, \underline{f}(\underline{x})) \in E$ und*

$$\underline{\underline{A}} = \left(\frac{\underline{\underline{E}}_m}{\underline{\underline{D}}\,\underline{f}(\underline{x})} \right)$$

der Tangential- und Normalenraum gegeben durch

$$T_{\underline{a}} = Im(\underline{\underline{A}}), \qquad N_{\underline{a}} = Ker(\underline{\underline{A}}^\top).$$

Beweis. Sei I ein Intervall und $\underline{\gamma} \in \mathcal{C}^1(I, E')$. Dann ist $\widetilde{\underline{\gamma}} : t \mapsto (\underline{\gamma}(t), \underline{f}(\underline{\gamma}(t))$ ein Weg in E. Wir wählen $I = (-\varepsilon, \varepsilon)$ und $\underline{\gamma}^i \in \mathcal{C}^1(I, E')$ mit $\underline{\gamma}^i_j(t) = \underline{x} + \delta_{ij}t$, $i, j = 1, ..., m$. Dann ist

$$(\widetilde{\underline{\gamma}}^i)'(0) = \underline{e}_i + \underline{D}_i\underline{f}(\underline{x}).$$

Da dies für $i = 1, ..., m$ genau m linear unabhängige Vektoren in $\mathbb{R}^d$ ergibt, und $T_{\underline{a}}E$ eine m-dimensionaler Vektorraum ist, spannen diese $T_{\underline{a}}E$ bereits auf. Diese Vektoren sind jedoch gerade die Bilder $\underline{\underline{A}}\,\underline{e}_i$, woraus die erste Behauptung folgt. Sei weiter $\underline{y} \in \mathrm{Im}(\underline{\underline{A}})$, also $\underline{y} = \underline{\underline{A}}\underline{y}'$ für ein $\underline{y}' \in E'$ und $\underline{z} \in \mathrm{Ker}(\underline{\underline{A}}^\top)$. Dann gilt

$$\langle \underline{z}, \underline{y} \rangle = \langle \underline{z}, \underline{\underline{A}}\underline{y}' \rangle = \underline{z}^\top \underline{\underline{A}}\underline{y}' = 0.$$

Da $\dim(\mathrm{Ker}(\underline{\underline{A}}^\top)) = n$ folgt die Behauptung. □

Beispiel 26.30 (Tangential- und Normalraum bei einem Kreis 1). Sei $E = \{(x, y) : x^2 + y^2 = 1\}$ die Kreislinie im $\mathbb{R}^2$. Für $\underline{a} = (x, y)$ mit $y > 0$ ist E lokal durch den Graph $f : (-1, 1) \to (-1, 1)$, $f(x) = \sqrt{1 - x^2}$ darstellen. Dann ist

$$\underline{\underline{A}} = \begin{pmatrix} 1 \\ \frac{-x}{1-x^2} \end{pmatrix}$$

die Matrix aus der Proposition. Damit ist

$$T_{(x,\sqrt{1-x^2})} = \mathrm{span} \begin{pmatrix} 1 \\ \frac{-x}{1-x^2} \end{pmatrix}, \qquad N_{(x,\sqrt{1-x^2})} = \mathrm{span} \begin{pmatrix} \frac{-x}{1-x^2} \\ -1 \end{pmatrix}.$$

Proposition 26.31 (Tangential- und Normalenraum bei Niveaumengen). *Ist $E = \underline{f}^{-1}(\underline{c})$ für ein $\underline{f} = (f_1, ..., f_n) \in \mathcal{C}^1(E' \subseteq \mathbb{R}^d, \mathbb{R}^n)$ und $\underline{c} \in \mathbb{R}^n$ ein regulärer Wert von $\underline{f}$, so gilt*

$$T_{\underline{a}}E = Ker(\underline{\underline{D}}\,\underline{f}(\underline{a})) = \{\underline{x} \in \mathbb{R}^n : \underline{\underline{D}}\,\underline{f}(\underline{a})\underline{x} = \underline{0}\}$$

und

$$N_{\underline{a}}E = span(\underline{\nabla} f_1(\underline{a}), ..., \underline{\nabla} f_n(\underline{a})) = Im((\underline{\underline{D}}\,\underline{f}(\underline{a}))^\top).$$

Beweis. Sei $\underline{x} \in T_{\underline{a}}E$ mit $\underline{\gamma} : (-\varepsilon, \varepsilon) \to E$ mit $\underline{\gamma}(0) = a, \underline{\gamma}'(0) = \underline{x}$. Dann ist $\underline{f} \circ \underline{\gamma} = \underline{c}$ und damit $\underline{0} = \underline{\underline{D}}(\underline{f} \circ \underline{\gamma})(0) = \underline{\underline{D}}\,\underline{f}(\underline{\gamma}(0))\underline{\gamma}'(0) = \underline{\underline{D}}\,\underline{f}(\underline{a})\underline{x}$. Insbesondere ist $\underline{x} \in \mathrm{Ker}(\underline{\underline{D}}\,\underline{f}(\underline{a}))$, also $T_{\underline{a}}E \subseteq \mathrm{Ker}(\underline{\underline{D}}\,\underline{f}(\underline{a}))$. Weiter ist $\dim(T_{\underline{a}}E) = m$ nach 1. Betrachten wir nun die Matrix $\underline{\underline{D}}\,\underline{f}(\underline{a})) \in \mathbb{R}^{n\times d}$. Es ist $\dim(\mathrm{Ker}(\underline{\underline{D}}\,\underline{f}(\underline{a}))) = d - n = m$ wegen der Surjektivität von $\underline{\underline{D}}\,\underline{f}(\underline{a}))$. Daraus folgt nun $T_{\underline{a}}E = \mathrm{Ker}(\underline{\underline{D}}\,\underline{f}(\underline{a}))$. Sei außerdem $\underline{x} \in T_{\underline{a}}E$. Dann gilt

$$\underline{0} = \underline{\underline{D}}\,\underline{f}(\underline{a}))\underline{x} = (\underline{\nabla} f_1(\underline{a})\underline{x}, ..., \underline{\nabla} f_n(\underline{a})\underline{x})$$

und damit stehen $\underline{\nabla} f_1(\underline{a}), ..., \underline{\nabla} f_n(\underline{a})$ auf $T_{\underline{a}}E$ senkrecht und die letzte Behauptung folgt. □

Beispiel 26.32 (Tangential- und Normalraum bei einem Kreis 2). Sei $E = \{(x, y) : x^2 + y^2 = 1\}$ die Kreislinie im $\mathbb{R}^2$. Also ist mit $f(x, y) = 1 - x^2 - y^2$ gerade $E = \{f = 0\}$. Wir berechnen $\underline{D}f(x, y) = (2x, 2y)$ und damit

$$T_{(x,y)} = \mathrm{Ker}(\underline{D}f(x, y)) = \mathrm{span}(-y, x), \qquad N_{(x,y)} = \mathrm{Im}(\underline{D}f(x, y)^\top) = \mathrm{span}(x, y).$$

Proposition 26.33 (Tangential- und Normalenraum bei Einbettungen). *Sei $E' \subseteq \mathbb{R}^m$ offen, $\underline{\gamma} \in \mathcal{C}^1(E', \mathbb{R}^d)$ eine Einbettung und $\underline{\alpha} : I \to \underline{\gamma}(E')$ eine $\mathcal{C}^1$-Kurve auf dem Intervall $I \subseteq \mathbb{R}$. Dann gibt es genau eine $\mathcal{C}^1$-Kurve $\underline{\beta} : I \to E'$ mit $\underline{\alpha} = \underline{\gamma} \circ \underline{\beta}$ und es gilt*

$$\underline{\alpha}'(t) = \underline{\underline{D}}\,\underline{\gamma}(\underline{\beta}(t)) \cdot \underline{\beta}'(t) = \sum_{i=1}^{m} \underline{\underline{D}}_i\underline{\gamma}(\underline{\beta}(t)) \cdot \beta_i'(t). \tag{26.2}$$

Insbesondere ist für $\underline{x} \in \underline{\gamma}(E')$

$$T_{\underline{\gamma}(\underline{x})}\underline{\gamma}(E') = Im(\underline{\underline{D}}\,\underline{\gamma}(\underline{x})), \qquad N_{\underline{\gamma}(\underline{x})}\underline{\gamma}(E') = Ker(\underline{\underline{D}}\,\underline{\gamma}(\underline{x}))^\top.$$

Beweis. Sei $E = \underline{\gamma}(E')$. Wir definieren $\underline{\beta} = \underline{\gamma}^{-1} \circ \underline{\alpha}$. Als Verknüpfung stetiger Abbildungen ist $\underline{\beta}$ stetig. Um zu zeigen, dass $\underline{\beta}$ stetig differenzierbar ist, müssen wir zeigen, dass $\underline{\gamma}^{-1}$ stetig differenzierbar ist. Sei also $\underline{a} \in E'$. Genau wie im Beweis von Proposition 26.25 gilt $\underline{\gamma}^{-1} = \underline{\psi}^{-1} \circ \pi$ auf einer Umgebung $U_{\underline{a}}$ von $\underline{a}$. Dies ist offenbar eine $\mathcal{C}^1$-Kurve. Die Formel für $\underline{\alpha}'$ folgt direkt aus der Kettenregel. Um $T_{\underline{x}}\gamma(E')$ für $\underline{x} \in \gamma(E')$ zu bestimmen, wählen wir einen Weg $\underline{\beta} : (-\varepsilon, \varepsilon) \to E'$ mit $\underline{\beta}(0) = \underline{x}$ und $\underline{\beta}' = \underline{e}_i$. Aus (26.2) folgt nun, dass $(\underline{\gamma} \circ \underline{\beta})' = \underline{\underline{D}}_i\underline{\gamma}(\underline{x})$, also $T_{\underline{\gamma}(\underline{x})}\underline{\gamma}(E') = \mathrm{Im}(\underline{\underline{D}}\,\underline{\gamma}(\underline{x}))$. Sei $\underline{y} \in N_{\underline{\gamma}(\underline{x})}\underline{\gamma}(E')$, also $\underline{y} \cdot \underline{\underline{D}}\,\underline{\gamma}(\underline{x}) \cdot \underline{z} = 0$ für alle $\underline{z} \in \mathbb{R}^m$. Dann muss $\underline{y} \cdot \underline{\underline{D}}\,\underline{\gamma}(\underline{x}) = \underline{0}$ gelten, also $\underline{y} \in \mathrm{Ker}(\underline{\underline{D}}\,\underline{\gamma}(\underline{x}))^\top$. □

Beispiel 26.34 (Tangential- und Normalraum bei einem Kreis 3). Sei $E = \{(x, y) : x^2 + y^2 = 1\}$ die Kreislinie im $\mathbb{R}^2$. Wir schreiben nun $E = \{\underline{\gamma}(t) : 0 < t < 7\}$ mit $\underline{\gamma}(t) = (\cos(t), \sin(t))$. Damit ist $\underline{\gamma}'(t) = (-\sin(t), \cos(t))^\top$ und damit

$$T_{(\cos(t),\sin(t))} = \mathrm{Im}(\underline{\gamma}'(t)) = \mathrm{span}(-\sin(t), \cos(t))^\top,$$
$$N_{(\cos(t),\sin(t))} = \mathrm{Ker}(\underline{\gamma}'(t))^\top = \mathrm{span}(\cos(t), \sin(t))^\top.$$

26.4 Integration über eine Untermannigfaltigkeit

Wollen wir ein Maß (oder Integral) auf einer Umtermannigfaltigkeit E definieren, so ist das im Prinzip ganz einfach, zumindest wenn E durch eine Einbettung gegeben ist. Ist $\underline{\gamma} : E' \subseteq \mathbb{R}^m \to E \subseteq \mathbb{R}^d$ eine reguläre Parameterdarstellung von E, so könnten wir für ein $f \in \mathcal{B}(E, \mathbb{R})$ das Maß λ_E auf $E \cap U$ mittels

$$\int_E f(\underline{x})\lambda_E(d\underline{x}) := \int_{E'} f(\underline{\gamma}(\underline{z}))\lambda^m(d\underline{z}) \tag{26.3}$$

definieren, wobei λ^m das m-dimensionale Lebesgue-Maß ist. Dies geht jedoch schief, wie das folgende Beispiel zeigt:

Beispiel 26.35 (Wohl-Definiertheit in (26.3)**).** Sei

$$E = \{(x, y) : -1 < x < 1, y = \sqrt{1 - x^2}\} = \{(\cos(t\pi), \sin(t\pi)), 0 < t < 1\}.$$

Dann ist sowohl $\gamma_1(x) = (x, \sqrt{1 - x^2})$ als auch $\gamma_2(t) = (\cos(t), \sin(t))$ reguläre Parameterdarstellungen. Für $f = 1$ ist weiter

$$\int_0^1 f(\underline{\gamma}_1(x))\lambda(dx) = 1,$$
$$\int_{-1}^1 f(\underline{\gamma}_2(t))\lambda(dt) = 2.$$

Insbesondere wäre die linke Seite von (26.3) nicht wohl-definiert.

Die Definition in 26.3 ist wohl deshalb gescheitert, weil man Untermannigfaltigkeiten auf verschiedene Arten und Weisen parametrisieren kann. Wir werden in Definition 26.40 für eine reguläre Parameterdarstellung $\underline{\gamma} : E' \subseteq \mathbb{R}^m \to \mathbb{R}^d$ der Untermannigfaltigkeit E

$$\int_E f(\underline{x})\lambda_E(d\underline{x}) := \int_E f d\lambda_E := \lambda_E[f] := \int_{E'} f(\underline{\gamma}(\underline{z}))\sqrt{\det(\underline{\underline{D}}\,\underline{\gamma}(\underline{z})^\top \underline{\underline{D}}\,\underline{\gamma}(\underline{z}))}\lambda^m(d\underline{z}) \tag{26.4}$$

für $f : E \to \mathbb{R}$ setzen. Dabei wird λ_E ein Maß auf $\mathcal{B}(E)$ werden, das lokal dieselben Eigenschaften wie das Lebesgue-Maß besitzt. (Das bedeutet, dass das Lebesgue-Maß einer *kleinen* Menge $A \subseteq E$ den Flächen- bzw Rauminhalt der Menge angeben soll.) Um $\int_E f d\lambda_E$ zu definieren, gehen wir in mehreren Schritten vor.

1. In diesem Abschnitt motivieren und definieren wir das Integral für den Fall, dass es eine globale Karte von E gibt.

2. Ist eine solche Karte nicht verfügbar, benötigen wir eine *Zerlegung der Eins* (durch verschiedene Karten), die im nächsten Abschnitt bereit gestellt wird.

3. Schließlich können wir dann den allgemeinen Fall durch Zusammensetzung des Integrals auf allen Kartengebieten mittels der Schritte 1. und 2. definieren.

Offenbar, siehe (26.4), spielt die Abbildung $\underline{z} \mapsto \sqrt{\det(\underline{\underline{D}}\,\underline{\gamma}(\underline{z})^\top \underline{\underline{D}}\,\underline{\gamma}(\underline{z}))}$ eine wichtige Rolle bei der Definition des Maßes λ_E. Dies liegt daran, wie man lokal Flächen auf einer Untermannigfaltigkeit misst. Dies ist zunächst einfach, wenn wenn E ein Parallelotop ist.

Definition 26.36 (Parallelotop). *Seien $\underline{x}_1, ..., \underline{x}_m \in \mathbb{R}^d$. Dann heißt die Menge*

$$P(\underline{x}_1, ..., \underline{x}_m) := \{a_1\underline{x}_1 + \cdots + a_m\underline{x}_m : a_1, ..., a_m \in [0, 1]\}$$

das von $\underline{x}_1, ..., \underline{x}_m$ aufgespannte Parallelotop.

Proposition 26.37 (Die Gram'sche Determinante als Flächeninhalt). *Es gibt genau eine Funktion $V_m : (\mathbb{R}^d)^m \to \mathbb{R}$, die folgende Eigenschaften besitzt:*

1. *$V_m(\underline{x}_1, ..., \lambda\underline{x}_i, ..., \underline{x}_m) = \lambda V_m(\underline{x}_1, ..., \underline{x}_m)$ für alle $\lambda \in \mathbb{R}$.*
2. *$V_m(\underline{x}_1, ..., \underline{x}_i + \underline{x}_j, ..., \underline{x}_m) = V_m(\underline{x}_1, ..., \underline{x}_m)$ für $i \neq j$.*
3. *$V_m(\underline{x}_1, ..., \underline{x}_m) = 1$, falls $\underline{x}_1, ..., \underline{x}_m$ ein Orthonormalsystem bilden.*

Diese Abbildung ist gegeben durch

$$V_m(\underline{x}_1, ..., \underline{x}_m) = \sqrt{det(\underline{\underline{A}}^\top \underline{\underline{A}})} \text{ mit } \underline{\underline{A}} = (\underline{x}_1^\top, ..., \underline{x}_m^\top) \in \mathbb{R}^{d\times m}.$$

Beweis. Zunächst zeigen wir, dass die angegebene Abbildung die Eigenschaften 1. – 3. erfüllt. Es ist

$$\underline{\underline{A}}^\top \underline{\underline{A}} = \begin{pmatrix} \underline{x}_1 \\ \vdots \\ \underline{x}_m \end{pmatrix} (\underline{x}_1^\top, ..., \underline{x}_m^\top) = \left(\underline{x}_k \underline{x}_l^\top\right)_{1\leq k,l\leq m}.$$

Bekanntlich ist die Determinante die einzige multilineare, alternierende und normierte Abbildung. Damit ist

$$\det \begin{pmatrix} \underline{x}_1 \\ \cdots \\ \lambda\underline{x}_i \\ \vdots \\ \underline{x}_m \end{pmatrix} (\underline{x}_1^\top, ..., \lambda\underline{x}_i^\top, ..., \underline{x}_m^\top) = \lambda^2 \det \begin{pmatrix} \underline{x}_1 \\ \vdots \\ \underline{x}_m \end{pmatrix} (\underline{x}_1^\top, ..., \underline{x}_m^\top)$$

und für $i \neq j$

$$\det \begin{pmatrix} \underline{x}_1 \\ \cdots \\ \underline{x}_i + \underline{x}_j \\ \vdots \\ \underline{x}_m \end{pmatrix} (\underline{x}_1^\top, ..., \underline{x}_i^\top + \underline{x}_j^\top, ..., \underline{x}_m^\top) = \det \begin{pmatrix} \underline{x}_1 \\ \vdots \\ \underline{x}_m \end{pmatrix} (\underline{x}_1^\top, ..., \underline{x}_m^\top).$$

Daraus folgen bereits die Eigenschaften 1. und 2. Weiter ist 3. klar, da in diesem Fall $\underline{\underline{A}}^\top \underline{\underline{A}} = \underline{\underline{E}}_m$.

Nun zeigen wir, dass es höchstens eine Abbildung V_m mit den geforderten Eigenschaften geben kann. Wir unterscheiden zwei Fälle: entweder sind die Vektoren $\underline{x}_1, ..., \underline{x}_m$ linear abhängig oder unabhängig. Im Falle der linearen Abhängigkeit ist etwa $\sum_{j=1}^m a_j\underline{x}_j = 0$ (und oBdA $A_j \neq 0, j = 1, ..., m$). Dann gilt

$$0 = V_m(a_1\underline{x}_1, ..., a_{m-1}\underline{x}_{m-1}, \underline{0}) = V_m(a_1\underline{x}_1, ..., a_m\underline{x}_m) = a_1 \cdots a_m \cdots V_m(\underline{x}_1, ..., \underline{x}_m)$$

und damit ist $V_m(\underline{x}_1, ..., \underline{x}_m) = 0$ eindeutig bestimmt.

Als nächstes zeigen wir, dass

$$V_m(\underline{x}_1, ..., \underline{x}_i + \lambda\underline{x}_j, ..., \underline{x}_m) = V_m(\underline{x}_1, ..., \underline{x}_m)$$

für alle $\underline{x}_1, ..., \underline{x}_m, \lambda \in \mathbb{R}$ schon aus 1. und 2. folgt. Für $\lambda = 0$ ist nichts zu zeigen, und für $\lambda \neq 0$ gilt nämlich.

$$\begin{aligned} V_m(\underline{x}_1, ..., \underline{x}_i + \lambda\underline{x}_j, ..., \underline{x}_m) &= \frac{1}{|\lambda|} V_m(\underline{x}_1, ..., \underline{x}_i + \lambda\underline{x}_j, ..., \lambda\underline{x}_j, ...\underline{x}_m) \\ &= \frac{1}{|\lambda|} V_m(\underline{x}_1, ..., \underline{x}_i, ..., \lambda\underline{x}_j, ..., \underline{x}_m) = V_m(\underline{x}_1, ..., \underline{x}_m) \end{aligned}$$

Sind nun $\underline{x}_1, ..., \underline{x}_m$ linear unabhängig, so wählen wir uns nach dem Gram-Schmidt'schen Orthogonalisierungsverfahren ein Orthonormalsystem von $\mathrm{span}(\underline{x}_1, ..., \underline{x}_m)$ sowie Darstellungen $\underline{x}_i = \sum_{j=1}^{i} a_{ij}\underline{z}_j$. Damit gilt

$$\begin{aligned} V_m(\underline{x}_1, ..., \underline{x}_m) &= V_m(a_{11}\underline{z}_1, a_{21}\underline{z}_1 + a_{22}\underline{z}_2, ..., a_{m1}\underline{z}_1 + \cdots + a_{mm}\underline{z}_m) \\ &= V_m(a_{11}\underline{z}_1, a_{22}\underline{z}_2, ..., a_{mm}\underline{z}_m) = a_{11} \cdots a_{mm} \cdot V_m(\underline{z}_1, ..., \underline{z}_m) = a_{11} \cdots a_{mm}. \end{aligned}$$

Insbesondere ist $V_m(\underline{x}_1, ..., \underline{x}_m)$ bereits durch die Eigenschaften 1. – 3. eindeutig bestimmt. □

Definition 26.38 (Gram'sche Matrix). *Sei $E' \subseteq \mathbb{R}^m$ offen und $\underline{\gamma} \in \mathcal{C}^1(E', \mathbb{R}^d)$ eine Immersion. Die Abbildung*

$$\underline{\underline{g}}_{\underline{\gamma}} : \underline{x} \mapsto \underline{\underline{D}}\,\underline{\gamma}(\underline{x})^\top \cdot \underline{\underline{D}}\,\underline{\gamma}(\underline{x}) \in \mathbb{R}^{m\times m}$$

heißt Gram'sche Matrix *(oder* Maßtensor*) von $\underline{\gamma}$ in $\underline{x} \in E'$. Weiter setzen wir*

$$g_{\underline{\gamma}}(\underline{x}) := det(\underline{\underline{g}}_{\underline{\gamma}}(\underline{x})).$$

Beispiel 26.39 (Lineare Immersion). Sei $E' = \mathbb{R}^m$ und $\underline{\gamma}(\underline{x}) := \underline{\underline{A}}\,\underline{x} + \underline{b}$ für $\underline{\underline{A}} \in \mathbb{R}^{d\times m}$ und $\underline{b} \in \mathbb{R}^m$. Dann ist

$$\underline{\underline{g}}_{\underline{\gamma}} := \underline{\underline{g}}_{\underline{\gamma}}(\underline{x}) = \underline{\underline{A}}^\top \underline{\underline{A}}.$$

Zu bemerken ist hierbei, dass

$$\underline{x} \mapsto \langle \underline{\underline{A}}\,\underline{x}, \underline{\underline{A}}\,\underline{x} \rangle = \langle \underline{x}, \underline{\underline{A}}^\top \underline{\underline{A}}\,\underline{x} \rangle = \underline{x}\,\underline{\underline{g}}_{\underline{\gamma}}\,\underline{x}$$

ein Skalarprodukt auf $\mathbb{R}^m$ definiert, falls $\underline{\underline{g}}_{\underline{\gamma}} \in Gl(m, \mathbb{R})$.

In Beispiel 26.35 haben wir bereits gesehen, dass (26.3) nicht zu einem wohl-definiertem Integralbegriff geführt hat. Mit Hilfe der Gram'schen Determinante können wir diesen Mangel nun beseitigen.

Definition 26.40 (Integration auf einem Kartengebiet). *Sei $E \in \mathbb{R}^d$ eine m-dimensionale Untermannigfaltigkeit mit einer globalen Karte. Sei weiter $\underline{\gamma} \in \mathcal{C}^1(E', \mathbb{R}^d)$ für ein $E' \subseteq \mathbb{R}^m$ offen seine Einbettung mit $\underline{\gamma}(E') = E$ (eine solche existiert immer nach Lemma 26.24). Dann definieren wir für $f \in \mathcal{B}(E, \mathbb{R})$*

$$\int_E f(\underline{x})\lambda_E(d\underline{x}) := \int_E f d\lambda_E := \lambda_E[f] := \int_{E'} f(\underline{\gamma}(\underline{x}))\sqrt{g_{\underline{\gamma}}(\underline{x})}\lambda^m(d\underline{x}) \tag{26.5}$$

wobei $g_{\underline{\gamma}}$ die Gram'sche Determinante aus Definition 26.38 ist.

Bemerkung 26.41 (Wohldefiniertheit in (26.5) **und Erweiterung).** Wir müssen noch zeigen, dass $\int_E f(\underline{x})\lambda_E(d\underline{x})$ unabhängig von der Parametrisierung der Untermannigfaltigkeit E ist. Seien hierzu $\underline{\gamma}_i \in \mathcal{C}(E_i', \mathbb{R}^d)$, $i = 1, 2$ zwei äquivalente Einbettungen mit $\underline{\gamma}_1 = \underline{\gamma}_2 \circ \underline{\psi}$ für einen Diffeomorphismus $\underline{\psi} : E_1' \to E_2'$. Zunächst ist

$$\underline{\underline{D}}\,\underline{\gamma}_1(\underline{x}) = \underline{\underline{D}}\,\underline{\gamma}_2(\underline{\psi}(\underline{x})) \cdot \underline{\underline{D}}\,\underline{\psi}(\underline{x})$$

und damit für die Gram'schen Determinanten $g_{\underline{\gamma}_1}$ und $g_{\underline{\gamma}_2}$ von $\underline{\gamma}_1$ und $\underline{\gamma}_2$

$$\begin{aligned} g_{\underline{\gamma}_1}(\underline{x}) &= \det(\underline{\underline{D}}\,\underline{\gamma}_1(\underline{x})^\top \underline{\underline{D}}\,\underline{\gamma}_1(\underline{x})) = \det(\underline{\underline{D}}\,\underline{\psi}(\underline{x})^\top \underline{\underline{D}}\,\underline{\gamma}_1(\underline{x})^\top \underline{\underline{D}}\,\underline{\gamma}_1(\underline{x}) \underline{\underline{D}}\,\underline{\psi}(\underline{x})) \\ &= (\det(\underline{\underline{D}}\,\underline{\psi}(\underline{x})))^2 \cdot \det(\underline{\underline{D}}\,\underline{\gamma}_2(\underline{x})^\top \underline{\underline{D}}\,\underline{\gamma}_2(\underline{x})) = (\det(\underline{\underline{D}}\,\underline{\psi}(\underline{x})))^2 \cdot g_{\underline{\gamma}_2}(\underline{x}). \end{aligned}$$

Nach dem Transformationssatz, Theorem 25.13, gilt nun

$$\begin{aligned} \int_{E_2'} f(\underline{\gamma}_2(\underline{x}))\sqrt{g_{\underline{\gamma}_2}(\underline{x})}\lambda^m(d\underline{x}) &= \int_{E_1'} f(\underline{\gamma}_2(\underline{\psi}(\underline{x})))\sqrt{g_{\underline{\gamma}_2}(\underline{\psi}(\underline{x}))} \cdot |\det(\underline{\underline{D}}\,\underline{\psi}(\underline{x}))|\lambda^m(d\underline{x}) \\ &= \int_{E_1'} f(\underline{\gamma}_1(\underline{x}))\sqrt{g_{\underline{\gamma}_1}(\underline{x})}\lambda^m(d\underline{x}). \end{aligned}$$

Damit ist die Wohldefiniertheit in (26.5) gezeigt.

Es ist leicht, die Definition (26.5) auf Funktionen $f \in \mathcal{B}(D, \mathbb{R})$ mit $f(\underline{x}) = 0$ außerhalb einer Untermannigfaltigkeit $E = \underline{\gamma}(E') \subseteq D$ für eine Einbettung $\underline{\gamma} \in \mathcal{C}^1(E', \mathbb{R}^d)$ mittels

$$\int f(\underline{x})\lambda_D(d\underline{x}) := \int f|_E(\underline{x})\lambda_E(d\underline{x}) \tag{26.6}$$

für eine Untermannigfaltigkeit $E \subseteq \mathbb{R}^d$ zu definieren.

Bemerkung 26.42 (Ein Maß auf E**).** Sei $E' \subseteq \mathbb{R}^m$ offen und $\underline{\gamma} \in \mathcal{C}^1(E', \mathbb{R}^d)$ eine Einbettung sowie $E = \underline{\gamma}(E')$. Dann können wir mittels

$$\lambda_E : A \mapsto \int_{E'} 1_{\gamma(\underline{x}) \in A} \sqrt{g_{\underline{\gamma}}(\underline{x})}\lambda^m(d\underline{x})$$

ein Maß auf $\mathcal{B}(E)$ definieren.

Proposition 26.43 (Integration über einen Graphen). *Sei $m + n = d$, $E' \subseteq \mathbb{R}^m$ und $\underline{f} \in \mathcal{C}^1(E', \mathbb{R}^{d-m})$. Dann ist $\underline{\gamma} : \underline{x} \mapsto (\underline{x}, \underline{f}(\underline{x}))$ eine Einbettung und es gilt*

$$g_{\underline{\gamma}}(\underline{x}) = det\big(\underline{\underline{E}}_m + \underline{\underline{D}}\,\underline{f}(\underline{x})^\top \cdot \underline{\underline{D}}\,\underline{f}(\underline{x})\big).$$

Ist $m = d - 1$, so gilt

$$g_{\underline{\gamma}}(\underline{x}) = 1 + ||\underline{D}f(\underline{x})||^2.$$

Beweis. Es gilt $\underline{\underline{D}}\,\underline{\gamma}(\underline{x}) = \left(\dfrac{\underline{\underline{E}}_{d-1}}{\underline{D}\,f(\underline{x})}\right)$. Daraus folgt die erste Formel für die Gram'sche Determinante.

Sei nun $m = d - 1$ und $\underline{v}_1, ..., \underline{v}_{d-1}$ eine Orthonormalbasis mit $\underline{v}_1 = \frac{\underline{D}f(\underline{x})}{||\underline{D}f(\underline{x})||}$. Dann gilt

$$\underline{v}_i\big(\underline{\underline{E}}_{d-1} + \underline{D}\,f(\underline{x})^\top \cdot \underline{D}\,f(\underline{x})\big)\underline{v}_j^\top = \delta_{ij} + \delta_{i1}||\underline{D}f(\underline{x})|| \cdot \delta_{j1}||\underline{D}f(\underline{x})|| = \delta_{ij}||\underline{D}f(\underline{x})||^2.$$

Damit gilt für die Matrix O, die $\underline{v}_1, ..., \underline{v}_{d-1}$ als Zeilenvektoren hat

$$\begin{aligned} g_{\underline{\gamma}}(\underline{x}) &= \det\big(O \cdot \big(\underline{\underline{E}}_{d-1} + \underline{D}\, f(\underline{x})^\top \cdot \underline{D}\, f(\underline{x})\big) \cdot O^\top\big) \\ &= \det\big(\underline{\underline{E}}_{d-1}(1 + ||\underline{D}f(\underline{x})||^2)\big) = 1 + ||\underline{D}f(\underline{x})||^2. \end{aligned}$$

□

Beispiel 26.44 (Oberfläche einer d-dimensionalen Kugel). Wir betrachten die Kugelschale (oder Sphäre)

$$S^{d-1} := \partial B_1(\underline{0}) := \{(x_1, ..., x_d) : x_1^2 + \cdots + x_d^2 = 1\}.$$

Offenbar ist für $f(\underline{x}) = \sqrt{1 - x_1^2 - \cdots - x_{d-1}^2}$ gerade

$$S^{d-1} \cap \{\underline{x} : x_d > 0\} = \{(\underline{x}, f(\underline{x})) : \underline{x} \in B_1^{d-1}(\underline{0})\}$$

die halbe Kugelschale. Mit dem letzten Resultat berechnen wir $||\underline{D}f(\underline{x})||^2 = \frac{||\underline{x}||^2}{1-||\underline{x}||^2}$ und damit

$$\begin{aligned} \gamma_d := \lambda_{S^{d-1}}(S^{d-1}) &= 2 \int_{B_1^{d-1}(\underline{0})} \sqrt{1 + ||\underline{D}f(\underline{x})||^2} d\lambda_{B_1^{d-1}(\underline{0})} \\ &= 2 \int_{B_1^{d-1}(\underline{0})} \frac{1}{\sqrt{1 - ||\underline{x}||^2}} \lambda_{B_1^{d-1}(\underline{0})}(d\underline{x}) \\ &= 2(d-1)\alpha_{d-1} \cdot \int_0^1 \frac{r^{d-2}}{\sqrt{1-r^2}} dr \\ &= 2(d-1)d\alpha_{d-1} \cdot \int_0^1 r^{d-2} \cdot \sqrt{1-r^2} dr, \end{aligned}$$

wobei wir im letzten Schritt mit partieller Integration verwendet haben dass

$$\begin{aligned} \int_0^1 \sqrt{1-r^2} r^{d-2} dr &= \frac{1}{d-1} \int_0^1 \frac{(1-(1-r^2))r^{d-1}}{\sqrt{1-r^2}} dr \\ &= \frac{1}{d-1} \Big(\int_0^1 \frac{r^{d-2}}{\sqrt{1-r^2}} dr - \int_0^1 \sqrt{1-r^2} r^{d-2} dr \Big). \end{aligned}$$

Wir zeigen nun

$$\gamma_d = d\alpha_d. \tag{26.7}$$

Hierfür berechnen wir

$$\alpha_d = 2 \int_{B_1^{d-1}(\underline{0})} \sqrt{1 - ||\underline{x}||^2} d\underline{x} = 2(d-1)\alpha_{d-1} \int_0^1 \sqrt{1-r^2} r^{d-2} dr,$$

woraus die Behauptung folgt.

Beispiel 26.45 (Kurven-Integrale). In Verallgemeinerung von Definition 17.5 definieren wir für eine offene Menge $E \subseteq \mathbb{R}^d$, $f \in \mathcal{C}^0(E, \mathbb{R})$, ein Intervall I und $\underline{\gamma} \in \mathcal{C}^1(I, E)$ das Kurveintegral

$$\int_{\underline{\gamma}(I)} f \cdot d\lambda_{\underline{\gamma}(I)} := \int_I f(\underline{\gamma}(t)) |\underline{\gamma}'(t)| dt.$$

Dies ist in völliger Analogie zu (26.5), denn

$$g_{\underline{\gamma}}(t) = \underline{\gamma}'(t)^\top \cdot \underline{\gamma}'(t) = |\underline{\gamma}(t)|^2,$$

also

$$\int_{\underline{\gamma}(I)} f \cdot d\lambda_{\underline{\gamma}(I)} = \int_I f(\gamma(t))\sqrt{g_{\underline{\gamma}}(t)}dt.$$

Wir erweitern nun die Definition 26.5 der Integration über ein Kartengebiet zur Integration auf einer Untermannigfaltigkeit. Ist nämlich $D \supseteq E$ und $f \in \mathcal{B}(D, \mathbb{R})$ mit $f(\underline{x}) = 0$ für $\underline{x} \notin E$, so setzen wir einfach

$$\int_D f d\lambda_D := \int_E f|_E d\lambda_E. \tag{26.8}$$

Zusammen mit der Zerlegung der Eins aus Lemma 26.11 ist es nun leicht, das Integral über eine Untermannigfaltigkeit (die durch mehr als ein Kartengebiet gegeben ist) zu definieren. Zunächst jedoch formulieren wir ein wichtiges Lemma.

Lemma 26.46 (Zerlegungslemma). *Sei $D \subseteq \mathbb{R}^d, E' \subseteq \mathbb{R}^m$, $f \in \mathcal{B}(D, \mathbb{R})$ mit $f(\underline{x}) = 0$ außerhalb einer m-dimensionalen Untermannigfaltigkeit $E = \underline{\gamma}(E') \subseteq D$ für eine Einbettung $\underline{\gamma} \in \mathcal{C}^1(E', \mathbb{R}^d)$. Sei weiter $h_1, h_2, ... \in \mathcal{C}(E, [0, 1])$ eine Zerlegung der Eins von D, so dass gilt:*

1. *$fh_1, fh_2, ...$ sind integrierbar,*
2. *$\sum_{n=1}^\infty \int |f|h_n\lambda_D < \infty$.*

Dann ist f über D integrierbar und es gilt

$$\sum_{n=1}^\infty \int fh_n\lambda_D = \int f\lambda_D.$$

Beweis. OBdA ist $D = E$ und $m = d$. Falls $m < d$, so ist $\int_E f d\lambda_E = \int_{E'} f \circ \underline{\gamma}\sqrt{g} d\lambda^m$ und damit können wir die Aussagen auf $\mathbb{R}^m$ übertragen.

Es gilt immer $f = \sum_{n=1}^\infty fh_n$. Gilt nun 1. und 2., so ist $|f|$ (und damit f) nach dem Satz von der monotonen Konvergenz integrierbar. Weiter folgt die Aussage des Satzes nach dem Satz der majorisierten Konvergenz. □

Folgende Definition erweitert den Integralbegriff auf einem Kartengebiet aus Definition 26.40. Sie benutzt das Zerlegungslemma 26.46.

Definition 26.47 (Integral auf einer Untermannigfaltigkeit). *Sei $E \subseteq \mathbb{R}^d$ eine m-dimensionale Untermannigfalltigkeit und $(E \cap U_{\underline{x}})_{\underline{x} \in D}$ ein Atlas von E mit $D \subseteq E$ abzählbar und $(h_{\underline{x}})_{\underline{x} \in D}$ eine dem Atlas untergeordnete Zerlegung der Eins. (Einen solchen Atlas gibt es nach Proposition 26.8 und die Zerlegung der Eins nach Lemma 26.11.)*

Seien weiter $\underline{\gamma}_{\underline{x}} \in \mathcal{C}^1(E'_{\underline{x}}, \mathbb{R}^d)$ Einbettungen mit $E'_{\underline{x}} \subseteq \mathbb{R}^m$ offen und $\underline{\gamma}_{\underline{x}}(E'_{\underline{x}}) = E \cap U_{\underline{x}}, \underline{x} \in D$ (diese existieren nach Lemma 26.24). Dann definieren wir für $f \in \mathcal{B}(E, \mathbb{R})$

$$\int_E f(\underline{x})\lambda_E(d\underline{x}) := \int_E f d\lambda_E := \lambda_E[f] := \sum_{n=1}^\infty \int_E (f \cdot h_n) d\lambda_E,$$

falls alle Terme auf der rechten Seite existieren. Ist hier auf der rechten Seite $\{h_n > 0\} \subseteq E \cap U_{\underline{x}}$ *für* $\underline{x} \in D$*, so ist (siehe auch* (26.8)*)*

$$\int_E (f \cdot h_n) d\lambda_E := \int_{E \cap U_{\underline{x}}} (f \cdot h_n)|_{E \cap U_{\underline{x}}} d\lambda_E := \int_{E'_{\underline{x}}} (f \cdot h_n)(\underline{\gamma}_{\underline{x}}(\underline{y})) \sqrt{g_{\underline{\gamma}_{\underline{x}}}(\underline{y})} \lambda^m(d\underline{y}).$$

Bemerkung 26.48 (Wohl-Definiertheit des Integrals). Offenbar setzt die obige Definition den Integralbegriff auf einem Kartengebiet aus Definition 26.40 fort. (Dort hatten wir eine globale Karte, damit ist E ein Atlas und 1 ist eine dem Atlas untergeordnete Zerlegung der Eins.) Wir müssen jedoch noch zeigen, dass wir das Integral auf Untermannigfaltigkeiten wohl-definiert haben. Hierzu sei $(E \cap T_{\underline{w}})_{\underline{w} \in C}$ für $C \subseteq E$ abzählbar ein weiterer Atlas und $(g_n)_{n=1,2,...}$ eine dem Atlas untergeordnete Zerlegung der Eins. Weiter seien $\underline{\beta}_{\underline{w}} \in \mathcal{C}_1(E'_{\underline{w}}, \mathbb{R}^d)$ mit $E'_{\underline{w}}$ offen Einbettungen mit $\underline{\beta}_{\underline{w}}(E'_{\underline{w}}) = E \cap T_{\underline{w}}$. Wir müssen nun zeigen, dass

$$\sum_{n=1}^{\infty} \int_E (f \cdot h_n) d\lambda_E = \sum_{m=1}^{\infty} \int_E (f \cdot g_m) d\lambda_E$$

gilt. OBdA sei $f \geq 0$ (sonst zerlegen wir $f = f^+ - f^-$) und stellen fest, dass $(E \cap U_{\underline{x}} \cap T_{\underline{w}})_{\underline{w} \in C, \underline{x} \in D}$ ein weiterer Atlas von E und $(h_n g_m)_{n,m=1,2,...}$ ein diesem Atlas untergeordnete Zerlegung der Eins ist.

Für festes $\underline{x} \in D$ sei $n \in \mathbb{N}$ mit $\{h_n > 0\} \subseteq E \cap U_{\underline{x}} = \underline{\gamma}_{\underline{x}}(E'_{\underline{x}})$. Nun erfüllen $E \cap U_{\underline{x}}$, $f \cdot h_n$ und $(h_n g_m)_{m=1,2,...}$ die Voraussetzungen von Lemma 26.46. Deshalb gilt

$$\int_E (f \cdot h_n) d\lambda_E = \sum_{m=1}^{\infty} \int_E (f \cdot h_n g_m) d\lambda_E.$$

Deshalb gilt (mit einer analogen Argumentation angewandt auf $f \cdot g_m$

$$\begin{aligned} \sum_{n=1}^{\infty} \int_E (f \cdot h_n) d\lambda_E &= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \int_E (f \cdot h_n g_m) d\lambda_E = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \int_E (f \cdot h_n g_m) d\lambda_E \\ &= \sum_{m=1}^{\infty} \int_E (f \cdot g_m) d\lambda_E, \end{aligned}$$

womit die Wohldefiniertheit gezeigt ist.

26.5 Integration über eine $\mathcal{C}^1$-Fläche

Wir erweitern nun den Begriff der λ^d-Nullmenge. Dies benötigen wir, um auf einer m-dimensionalen Untermannigfaltigkeit $E \subseteq \mathbb{R}^d$ eine Teilmenge $N \subseteq E$ zu finden, die bei der Integration über E vernachlässigbar sind. Folgende Definition schließt an Theorem 25.2 an.

Definition 26.49 (m-Nullmenge und $\mathcal{C}^1$-Fläche). *1. Sei* $m \leq d$*. Eine Menge* $N \subseteq \mathbb{R}^d$ *ist eine* m*-Nullmenge, wenn es für jedes* ε *eine abzählbare Familie von Würfeln* $W_1, W_2, ... \subseteq \mathbb{R}^d$ *mit Seitenlängen* $r_1, r_2, ... > 0$ *gibt, so dass*

$$N \subseteq \bigcup_{n=1}^{\infty} W_n \text{ und } \sum_{n=1}^{\infty} r_n^m < \varepsilon.$$

2. *Eine Menge $E \subseteq \mathbb{R}^d$ ist eine m-dimensionale $\mathcal{C}^1$ Fläche, falls es eine m-dimensionale, offene (in $\mathcal{O}_X$) Untermannigfaltigkeit $D \subseteq E$ gibt mit folgenden Eigenschaften:*

   (a) *$E \setminus D$ ist eine m-Nullmenge.*

   (b) *$\overline{D} = E$, wobei $\overline{D}$ der Abschluss von D in $\mathcal{O}_E$ ist. (Das bedeutet, dass E die Menge der Häufungspunkte von Folgen in D ist.)*

Beispiel 26.50 (Kompakte Teilmengen von $\mathbb{R}_0^m$ sind $m+1$-Nullmengen). Sei $m < d$. Dann ist $\mathbb{R}_0^m \cap [0,1]^d$ (siehe (26.1)) eine $m+1$-Nullmenge.
Denn: Sei $\varepsilon > 0$. Wir betrachten Würfel der Kantenlänge $\varepsilon > 0$. Von diesem benötigt man $1/\varepsilon^m$ viele, um $\mathbb{R}_0^m \cap [0,1]^d$ zu überdecken. Außerdem gilt $\varepsilon^{m+1} \cdot \frac{1}{\varepsilon^m} = \varepsilon$, also ist die Summe der Seitenlängen der überdeckenden Würfel kleiner als ε, wie gefordert. Daraus folgt die Behauptung.

Offenbar ist nach Theorem 25.2 jede λ^d-Nullmenge genau eine d-Nullmenge. Wir zeigen nun, dass m-Nullmengen ähnlichen Regeln unterliegen wie λ^d-Nullmengen.

Proposition 26.51 (Eigenschaften von m-Nullmengen). *Sei $m \leq d$.*

1. *Ist N eine m-Nullmenge und $M \subseteq N$. Dann ist auch M eine m-Nullmenge.*

2. *Sei N eine m-Nullmenge und $m < d$. Dann ist N auch eine $m+1$-Nullmenge.*

3. *Seien $N_1, N_2, ... \subseteq \mathbb{R}^d$ alle m-Nullmengen. Dann ist auch $N := \bigcup_{\ell=1}^\infty N_\ell$ eine m-Nullmenge.*

Beweis. 1. ist klar, da jede Überdeckung von N mit Würfeln auch eine Überdeckung von M ist.
2. Sei $0 < \varepsilon < 1$ und $W_1, W_2, ... \subseteq \mathbb{R}^d$ eine Familie von Würfeln mit Seitenlängen $r_1, r_2, ... > 0$ und $\sum_{n=1}^\infty r_n^m < \varepsilon$. Es ist $r_1, r_2, ... < 1$, andernfalls wäre die Summe mindestens 1. Deshalb gilt

$$\sum_{n=1}^\infty r_n^{m+1} \leq \sum_{n=1}^\infty r_n^m < \varepsilon.$$

Insbesondere ist N auch eine $m+1$-Nullmenge.
3. Sei $(W_k^\ell)_{k=1,2,...}$ eine Überdeckung von N_ℓ mit Seitenlängen $(r_k^\ell)_{k=1,2,...}$ und $\sum_{k=1}^\infty (r_k^\ell)^m < \varepsilon 2^{-\ell}$. Dann ist $(W_k^\ell)_{k,\ell=1,2,...}$ eine Überdeckung von N und

$$\sum_{\ell=1}^\infty \sum_{k=1}^\infty (r_k^\ell)^m \leq \sum_{\ell=1}^\infty \varepsilon 2^{-\ell} = \varepsilon.$$

Damit ist auch N eine m-Nullmenge. □

Proposition 26.52 (Untermannigfaltigkeiten sind Nullmengen). *Sei $m < d$ und E eine m-dimensionale Untermannigfaltigkeit von $\mathbb{R}^d$. Dann ist E eine $m+1$-Nullmenge, (insbesondere also auch eine d-Nullmenge).*

Beweis. Da nach Proposition 26.8 eine Untermannigfaltigkeit die Vereinigung abzählbar vieler Kartengebiete ist, genügt es nach Proposition 26.51.3 zu zeigen, dass jedes Kartengebiet eine $m+1$-Nullmenge ist. Eine Karte $\varphi : U \to \mathbb{R}^d$ ist ein Diffeomoprhismus, also ist insbesondere φ^{-1} stetig differenzierbar und bildet die $m+1$-Nullmenge $\mathbb{R}_0^m \cap V$ (für ein $V \subseteq \mathbb{R}^d$ offen;

siehe Beispiel 26.50) auf $U \cap E$ für eine offene Menge $U \subseteq \mathbb{R}^d$ ab. Genau wie in Korollar 25.3 zeigt man, dass Bilder von $m+1$-Nullmengen unter stetig differenzierbaren Abbildungen $m+1$-Nullmengen sind. Insbesondere ist das Kartengebiet $U \cap E$ eine $m+1$-Nullmenge. $\square$

Proposition 26.53 (Teilmengen von Untermannigfaltigkeiten als Nullmengen). *Sei $E \subseteq \mathbb{R}^d$ eine m-dimensionale Untermannigfaltigkeit, $\pi_m : \mathbb{R}^d \to \mathbb{R}^m$ die Projektion auf die ersten m Koordinaten. Dann ist $N \subseteq E$ genau dann eine m-Nullmenge, wenn für jede Karte $\varphi : U \to V$ (d.h. $\varphi(U \cap E) = V \cap \mathbb{R}^m_0$ und damit auch $\pi_m \circ \varphi(U \cap N) = \pi_m(V \cap \varphi(N))$) gilt, dass $\lambda^m(\pi_m \circ \varphi(U \cap N)) = 0$.*

Beweis. Zunächst behaupten wir:

Eine Menge $N \subseteq \mathbb{R}^m_0$ ist genau dann eine m-Nullmenge, wenn $\lambda^m(\pi_m(N)) = 0$.

Denn: wir wissen bereits, dass eine Menge $N \subseteq \mathbb{R}^d$ genau dann eine d-Nullmenge ist, wenn $\lambda^d(N) = 0$. (Siehe hierzu Theorem 25.2 und die Definition 26.49.) Den Beweis der Behauptung führt man entsprechend, konstruiert jedoch alles direkt als Teilmenge von $\mathbb{R}^m_0$.

Nun zum Beweis der Behauptung: '$\Rightarrow$': Sei also N eine m-Nullmenge und $\varphi : U \to V$ eine Karte. Damit ist wie in Korollar 25.3 auch $\varphi(U \cap N) \subseteq \mathbb{R}^m_0$ eine m-Nullmenge, also gilt nach obiger Behauptung $\lambda^m(\pi_m \circ \varphi(U \cap N)) = 0$.

'$\Leftarrow$': Da E nach Proposition 26.8 einen abzählbaren Atlas $(U_{\underline{x}} \cap E)_{\underline{x} \in D}$ mit $D \subseteq E$ abzählbar besitzt und nach Proposition 26.51 die abzählbare Vereinigung von m-Nullmengen wieder eine m-Nullmenge ist, genügt es zu zeigen, dass $U_{\underline{x}} \cap N$ eine m-Nullmenge ist, $\underline{x} \in D$. Nach obiger Behauptung ist $\varphi(U_{\underline{x}} \cap N)$ eine m-Nullmenge und da φ^{-1} stetig differenzierbar ist folgt die Behauptung genau wie im letzten Beweis aus Korollar 25.3. $\square$

Vom Integral einer Funktion $f \in \mathcal{B}(\mathbb{R}^d, \mathbb{R})$ wissen wir, dass sich $\lambda^d[f]$ nicht ändert, wenn man f auf einer λ^d-Nullmenge abändert. Dies ist bei der Integration auf Untermannigfaltigkeiten genauso, wie wir nun zeigen werden.

Theorem 26.54 (Integration auf Untermannigfaltigkeiten und m-Nullmengen). *Sei $E \subseteq \mathbb{R}^d$ eine m-dimensionale Untermannigfaltigkeit, $N \subseteq E$ eine m-Nullmenge und $f, h \in \mathcal{B}(E, \mathbb{R})$. Dann gilt:*

1. *Sei $f(\underline{x}) = h(\underline{x})$ für $\underline{x} \in E \setminus N$. Existiert $\lambda_E[f]$, so existiert auch $\lambda_E[g]$ und es gilt $\lambda_E[f] = \lambda_E[g]$.*

2. *Sei N abgeschlossen. Das Integral $\lambda_E[f]$ existiert genau dann, wenn $\lambda_{E \setminus N}[f|_{E \setminus N}]$ existiert. In diesem Fall sind beide Integrale gleich.*

Beweis. Wir zeigen die Behauptungen nur in dem Fall, wenn E offen ist und es eine globale Karte φ von E gibt. Der allgemeine Fall erschließt sich dann mit Hilfe eines abzählbaren Atlas von E; siehe Proposition 26.8. Nach Lemma 26.24 gibt es $E' \subseteq \mathbb{R}^m$ und eine Einbettung $\underline{\gamma} \in \mathcal{C}^1(E', \mathbb{R}^d)$ mit $\underline{\gamma}(E') = E$. Es ist $N' := \underline{\gamma}^{-1}(N)$ eine m-Nullmenge, also auch eine λ^m-Nullmenge.
1. Wir schreiben direkt

$$\begin{aligned}\lambda_E[f] &= \int_{E'} f(\underline{\gamma}(\underline{x}))\sqrt{g(\underline{x})}\lambda^m(d\underline{x}) = \int_{E'} f(\underline{\gamma}(\underline{x}))1_{\underline{x} \notin N'}\sqrt{g(\underline{x})}\lambda^m(d\underline{x}) \\ &= \int_{E'} h(\underline{\gamma}(\underline{x}))1_{\underline{x} \notin N'}\sqrt{g(\underline{x})}\lambda^m(d\underline{x}) = \lambda_E[h],\end{aligned}$$

falls eines der Integrale existiert.
2. Es ist $\widetilde{\underline{\gamma}} := \underline{\gamma}|_{\underline{\gamma}^{-1}(E\setminus N)}$ eine Einbettung mit Bild $E \setminus N$. (Hierbei haben wir verwendet, dass N abgeschlossen und damit $E \setminus N$ offen ist.) Deshalb schreiben wir, da $E' \setminus \underline{\gamma}^{-1}(E \setminus N)$ eine λ^m-Nullmenge ist

$$\lambda_{E\setminus N}[f|_{E\setminus N}] = \int_{\underline{\gamma}^{-1}(E\setminus N)} f(\underline{\gamma}(\underline{x}))\sqrt{g(\underline{x})}\lambda^m(d\underline{x}) = \int_{E'} f(\underline{\gamma}(\underline{x}))\sqrt{g(\underline{x})}\lambda^m(d\underline{x}) = \lambda_E[f].$$

□

Wir können nun die Integration auf $\mathcal{C}^1$-Flächen definieren. Wieder stellt dies eine Erweiterung von Definition 26.47 dar.

Definition 26.55 (Integration auf einer m-dimensionalen $\mathcal{C}^1$-Fläche). *Sei $E = D \cup (E \setminus D)$ eine m-dimensionale $\mathcal{C}^1$-Fläche, wobei $D \subseteq E$ eine offene m-dimensionale Untermannigfaltigkeit und $E \setminus D$ eine m-Nullmenge und $\overline{D} = E$ ist; siehe Definition 26.49. Dann definieren wir für $f \in \mathcal{B}(E, \mathbb{R})$*

$$\int_E f(\underline{x})\lambda_E(d\underline{x}) := \int_E f d\lambda_E := \lambda_E[f] := \int_D f|_D d\lambda_D,$$

falls die rechte Seite existiert.

Bemerkung 26.56 (Wohl-Definiertheit des Integrals). Wir müssen zeigen, dass die Definition unabhängig von der Untermannigfaltigkeit D ist. Zunächst sei hierzu $\widetilde{D} := \bigcup\{D \subseteq E$ m-dimensionale Untermannigfaltigkeit, $\overline{D} = E\}$. Dann ist $\widetilde{D}$ als Vereinigung von offenen Mengen wieder offen, also auch eine offene Untermannigfaltigkeit. Sind nun $D', D'' \subseteq E$ offene m-dimensionale Untermannigfaltigkeiten, so dass $E \setminus D', E \setminus D''$ jeweils m-Nullmengen (und abgeschlossen, da D', D'' offen) sind und $\overline{D}' = \overline{D}'' = E$. Dann gilt mit Theorem 26.54.2

$$\lambda_{D'}[f] = \lambda_{E\setminus(E\setminus D')}[f] = \lambda_E[f] = \lambda_{E\setminus(E\setminus D'')}[f]\lambda_{D''}[f].$$

27 Der Gauss'sche Integralsatz

In diesem Abschnitt steht der Satz von Gauss (oder *divergence theorem* auf englisch) im Mittelpunkt. Wir formulieren ihn zunächst für ein Vektorfeld, d.h. für eine Funktion $\underline{f} = (f_1, ..., f_d) \in \mathcal{C}^1(\mathbb{R}^d, \mathbb{R}^d)$. Wir setzen

27.1 Die Divergenz eines Vektorfeldes

Wir beweisen hier einen einfachen Fall des Gauss'schen Integralsatzes, der den Zusammenhang zum Hauptsatz der Integral- und Differentialrechnung, veranschaulicht. Hierbei ist $\int_{\partial E}$ das Integral auf einer $\mathcal{C}^1$-Fläche aus dem letzten Abschnitt.

Definition 27.1 (Divergenz eines Vektorfeldes). *Sei $E \subseteq \mathbb{R}^d$ offen und $\underline{f} \in \mathcal{C}^1(E, \mathbb{R}^d)$ ein Vektorfeld. Dann ist die* Divergenz *von $\underline{f}$ gegeben als*

$$div(\underline{f})(\underline{x}) := \frac{\partial f_1}{\partial x_1} + \cdots + \frac{\partial f_d}{\partial x_d}. \tag{27.1}$$

Integration auf	$\int_E f d\lambda_E$	Referenz
Kartengebiet $\underline{\gamma}(E') = E$ mit $E' \subseteq \mathbb{R}^m$ offen, $\underline{\gamma} : E' \to E$ Einbettung	$:= \int_{E'} f(\underline{\gamma}(\underline{x}))\sqrt{g(\underline{x})}\lambda^m(d\underline{x})$, wobei g Gram'sche Determinante aus Definition 26.38	Def. 26.40
Untermannigfaltigkeit E	$:= \int_{E\cap U} f\|_{E\cap U} d\lambda_{E\cap U}$, falls $E\cap U$ Kartengebiet, $\{f \neq 0\} \subseteq E\cap U$	(26.6)
Untermannigfaltigkeit E $(E\cap U_{\underline{x}})_{\underline{x}\in D}$ abzählbarer Atlas von E	$:= \sum_{n=1}^{\infty} \int_E (f \cdot h_n) d\lambda_E$, wobei $(h_n)_{n=1,2,\ldots}$ dem Atlas untergeordnete Zerlegung der Eins	Def. 26.47
$\mathcal{C}^1$-Fläche E	$:= \int_D f\|_D d\lambda_D$, mit $E = D \cup (E \setminus D)$, wobei D eine offene m-dimensionale Untermannigfaltigkeit, $E \setminus D$ eine m-Nullmenge, $\overline{D} = E$	Def. 26.55

Tabelle 26.2: Definition des Integrals $\int_E f d\lambda_E$ für eine Funktion $f \in \mathcal{B}(E, \mathbb{R})$.

Theorem 27.2 (Gauss'scher Integralsatz auf Quadern). *Sei $E = \prod_{i=1}^d [a_i, b_i]$ ein d-dimensionaler Quader und ∂E der Rand von E ($\partial E := E \setminus E^\circ$). Wir schreiben $\partial E = N \uplus D$, wobei D die Vereinigung der (in ∂E) offenen Seiten des Quaders und N die Kanten des Quaders sind. Weiter ist*

$$\underline{\nu} : \begin{cases} \partial E & \to \mathbb{R}^d \\ \underline{x} = (x_1, ..., x_d) & \mapsto \begin{cases} \underline{0}, & \text{falls } \underline{x} \in N, \\ -\underline{e}_i, & \text{falls } \underline{x} \notin N, x_i = a_i, \\ \underline{e}_i, & \text{falls } \underline{x} \notin N, x_i = b_i. \end{cases} \end{cases} \tag{27.2}$$

(Dieses $\underline{\nu}$ werden wir unten als das äußere Einheitsnormalenfeld von E kennen lernen.) Dann gilt

$$\int_E div(\underline{f})(\underline{x}) \lambda^d(d\underline{x}) = \int_{\partial E} \langle \underline{f}(\underline{x}), \underline{\nu}(\underline{x}) \rangle \lambda_{\partial E}(d\underline{x}).$$

Beweis. Wir setzen $E_k := \prod_{j \neq k} [a_j, b_j]$ and $\underline{x}_{-k} := (x_1, ..., x_{k-1}, x_{k+1}, ..., x_d)$. Wir schreiben direkt mit Hilfe des Satzes von Fubini, Theorem 24.12, und des Hauptsatzes der Integral- und Differentialrechnung, Theorem 8.27,

$$\begin{aligned} \int_E \operatorname{div}(\underline{f})(\underline{x}) \lambda^d(d\underline{x}) &= \sum_{k=1}^d \int_E \frac{\partial f_k(\underline{x})}{\partial x_k} \lambda^d(d\underline{x}) \\ &= \sum_{k=1}^d \int_{E_k} \int_{a_k}^{b_k} \frac{\partial f_k(\underline{x})}{\partial x_k} dx_k \lambda^{d-1}(d\underline{x}_{-k}) \\ &= \sum_{k=1}^d \int_{E_k} (f_k(x_1, ..., b_k, ..., x_d) - f_k(x_1, ..., a_k, ..., x_d)) \lambda^{d-1}(d\underline{x}_{-k}) \\ &= \sum_{k=1}^d \int_{E_k \times \{a_k, b_k\}} \langle \underline{f}(\underline{x}), \underline{\nu}(\underline{x}) \rangle \lambda_{E_k \times \{a_k, b_k\}}(d\underline{x}) \\ &= \int_{\partial E} \langle \underline{f}(\underline{x}), \underline{\nu}(\underline{x}) \rangle \lambda_{\partial E}(d\underline{x}). \end{aligned}$$

□

Bemerkung 27.3 (Veranschaulichung der Divergenz). Ist $\underline{f} \in \mathcal{C}^1(E, \mathbb{R}^d)$ für $E \subseteq \mathbb{R}^d$ offen. Für $\underline{x} \in E$ sei Q_ε der Würfel um $\underline{x}$ mit Seitenlänge $\varepsilon > 0$ und äußeredem Einheitsnormalenfeld wie in (27.2). Dann gilt nach Theorem 27.2

$$\operatorname{div}(\underline{f})(\underline{x}) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon^d} \int_{\partial Q_\varepsilon} \langle \underline{f}(\underline{x}), \underline{\nu}(\underline{x}) \rangle \lambda_{\partial E_\varepsilon}(d\underline{x}).$$

Die rechte Seite liefert den Fluss des Vektorfeldes $\underline{f}(\underline{x})$ entlang der Oberfläche des Quaders ∂E_ε. Dies interpretiert man oft auch als Quelldichte des Vektorfeldes $\underline{f}$.

27.2 Das äußere Normalenvektorfeld eine Polyeders

Unser Ziel ist es, Theorem 27.2 auf allgemeinere Gebiete $E \subseteq \mathbb{R}^d$ zu erweitern. Wir werden dies für $\mathcal{C}^1$-Polyeder erreichen, deren Rand eine orientierbare (d.h. es ist klar wo innen und außen ist) Hyperfläche (d.h. eine $d-1$-dimensionale $\mathcal{C}^1$-Fläche) bildet. Siehe Proposition 27.7.

Definition 27.4 (Hypefläche, $\mathcal{C}^1$-Rand, äußere Normale, Polyeder).

1. *Eine $d-1$-dimensionale $\mathcal{C}^1$-Fläche heißt Hyperfläche. Eine $d-1$-dimensionale Untermannigfaltigkeit des $\mathbb{R}^d$ heißt reguläre Hyperfläche.*

2. *Sei $E \subseteq \mathbb{R}^d$ offen. Ein Punkt $\underline{x} \in \partial E$ heißt regulär, wenn es eine Umgebung $U \subseteq \mathbb{R}^d$ und $h \in \mathcal{C}^1(U, \mathbb{R})$ gibt mit $\underline{D}h(\underline{x}) \neq \underline{0}$ für alle $\underline{x} \in U$ und $E \cap U = \{h < 0\}$. Die Menge der regulären Randpunkte heißt $\partial_M G$, weiter heißt $\partial_S E := \partial E \setminus \partial_M E$ singulärer Rand von E. Weiter heißt E ein $\mathcal{C}^1$-Polyeder, wenn $\partial_S E$ eine $d-1$-Nullmenge ist.*

3. *Sei $E \subseteq \mathbb{R}^d$ eine reguläre Hyperfläche und $\underline{x} \in E$. Ein* Einheitsnormalenfeld *ist eine stetige Abbildung $\underline{\nu} : E \to \mathbb{R}^d$, so dass (siehe auch Definition 26.27) $\underline{\nu}(\underline{x}) \in N_{\underline{x}}E$ und $||\underline{\nu}(\underline{x})|| = 1$.*

4. *Eine reguläre Hyperfläche E heißt orientierbar, wenn es ein Einheitsnormalenfeld $\underline{\nu}$ gibt. Dann heißt (E, ν) eine orientierte Hyperfläche. Ist weiter $\underline{f} \in \mathcal{B}(E, \mathbb{R}^d)$, so definieren wir das Integral*
$$\int_{(E,\underline{\nu})} \underline{f} d\underline{\lambda}_E := \int_E \langle \underline{f}, \underline{\nu} \rangle d\lambda_E,$$
falls die rechte Seite existiert.

Beispiel 27.5 (Einheitsnormalenfelder bei Graphen und Niveaumengen).

1. Sei $g \in \mathcal{C}^1(E', \mathbb{R})$ für $U \subseteq \mathbb{R}^{d-1}$. Dann ist nach Proposition 26.29 für $\underline{a} = (\underline{x}, g(\underline{x}))$ der Normalenraum gegeben als $N_{\underline{a}} = \operatorname{Ker}(\underline{\underline{A}}^\top)$ für $\underline{\underline{A}}^\top = (\underline{\underline{E}}_{d-1} | (\underline{D}g(\underline{x}))^\top)$. Für $\underline{y} \in N_{\underline{a}}$ gilt also $y_k + D_k g(\underline{x}) y_n = 0$. Eine Lösung hiervon ist offenbar $\underline{y} = (-\underline{D}g(\underline{x}), 1)$. Damit sind
$$\underline{\nu} : \underline{x} \mapsto \frac{(-\underline{D}g(\underline{x}), 1)}{\sqrt{1 + ||\underline{D}g(\underline{x})||^2}} \text{ und } \underline{\nu} : \underline{x} \mapsto \frac{(\underline{D}g(\underline{x}), -1)}{\sqrt{1 + ||\underline{D}g(\underline{x})||^2}}$$
zwei Einheitsnormalenfelder.

2. Sei $f \in \mathcal{C}^1(E', \mathbb{R})$ für $E' \subseteq \mathbb{R}^d$ offen sowie $E = \underline{f}^{-1}(0)$. Ist $\underline{D}f(\underline{x}) \neq \underline{0}$ für alle $\underline{x} \in E$, so sind
$$\underline{\nu} : \underline{x} \mapsto \frac{\underline{D}f(\underline{x})}{||\underline{D}f(\underline{x})||} \text{ und } \underline{\nu} : \underline{x} \mapsto \frac{-\underline{D}f(\underline{x})}{||\underline{D}f(\underline{x})||}$$
zwei Einheitsnormalenfelder.

Lemma 27.6 (Anzahl von Einheitsnormalenfeldern). *Sei $E \subseteq \mathbb{R}^d$ eine wegzusammenhängende, reguläre Hyperfläche. Dann gibt es entweder kein Einheitsnormalenfeld, oder genau zwei.*

Beweis. Ist $\underline{\nu}$ ein Einheitsnormalenfeld, so ist auch $-\underline{\nu}$ eines. Weiter wissen wir, dass $\dim(N_{\underline{x}}E) = 1$ gilt. Sind also $\underline{\nu}, \widetilde{\underline{\nu}}$ zwei Einheitsnormalenfelder, so ist $s : \underline{x} \mapsto \langle \underline{\nu}(\underline{x}), \widetilde{\underline{\nu}}(\underline{x}) \rangle$ stetig (entlang jeden Weges in E) und $s = \pm 1$. Ist $s = 1$, so muss $\underline{\nu} = \widetilde{\underline{\nu}}$, sonst ist $\underline{\nu} = -\widetilde{\underline{\nu}}$. $\square$

Proposition 27.7 (Rand eines Polyeders als orientierbare Hyperfläche). *Sei $E \subseteq \mathbb{R}^d$ ein $\mathcal{C}^1$-Polyeder mit regulärem Rand $\partial_M E$. Dann gilt:*

1. *Der reguläre Rand $\partial_M E$ ist eine orientierbare Hyperfläche.*
   Insbesondere gilt: Für jedes $\underline{x} \in \partial_M E$ gibt es (nach eventueller Umnummerierung der Koordinaten) einen offenen Quader $Q = Q' \times I$, wobei I ein Intervall ist und eine Abbildung $f \in \mathcal{C}^1(Q', I)$ mit

$$E \cap Q = \{(\underline{x}, y) : f(\underline{x}) < y\} \text{ oder } E \cap Q = \{(\underline{x}, y) : f(\underline{x}) > y\},$$

   sowie $\partial E \cap Q = \{(\underline{x}, f(\underline{x})) : \underline{x} \in Q'\}$.

2. *Auf $\partial_M E$ gibt es genau ein Einheitsnormalenfeld $\underline{\nu}$ derart, dass es für jedes $\underline{x} \in \partial_M E$ ein t_0 gibt, so dass für $0 < t < t_0$ gilt, dass $\underline{x} + t\underline{\nu}(\underline{x}) \notin E, \underline{x} - t\underline{\nu}(\underline{x}) \in E$.*
   Ist $E \cap U = \{h < 0\}$ für eine Umgebung U von $\underline{x}$ für ein $h \in \mathcal{C}^1(U, \mathbb{R})$, so gilt

$$\underline{\nu}(\underline{x}) = \frac{\underline{D}h(\underline{x})}{||\underline{D}h(\underline{x})||}. \tag{27.3}$$

Beweis von Proposition 27.7. 1. OBdA ist die Umgebung U in Definition 27.4.2 ein offener Quader Q. Es gibt also $h \in \mathcal{C}^1(Q, \mathbb{R})$ mit $\underline{D}h(\underline{x}) \neq \underline{0}$ für $\underline{x} \in Q$ und $E \cap Q = \{h < 0\}$, also auch $\partial E \cap Q \subseteq N := \{h = 0\}$.
Wir numerieren die Koorinaten so um, dass die letzte Koordinate von $\underline{D}h(\underline{x})$ nicht 0 ist. Damit haben wir mit $Q = Q' \times I$, dass $h \in \mathcal{C}^1(Q' \times I, \mathbb{R})$ mit $h(\underline{x}, y) = 0$ für $(\underline{x}, y) \in \partial E \cap Q$. Nach dem Satz über implizite Funktionen, Theorem 16.20, gibt es damit ein $f \in \mathcal{C}^1(Q', \mathbb{R})$ mit $\partial E \cap Q \subseteq N = \{(\underline{x}, y) : f(\underline{x}) = y\}$.
Nun hat $\underline{h}$ weder auf $Z_+ := \{(\underline{x}, y) : f(\underline{x}) > y\}$ noch auf $Z_- := \{(\underline{x}, y) : f(\underline{x}) < y\}$ eine Nullstelle (sonst müsste dort ja $f(\underline{x}) = y$ gelten). Weiter hat $\underline{h}$ auf Z_+ und Z_- unterschiedliche Vorzeichen (sonst hätte $\underline{h}$ auf $\partial E \cap Q$ ein Maximum oder Minimum und damit gäbe es $\underline{x} \in Q$ mit $\underline{D}h(\underline{x}) = \underline{0}$). Also muss $\{\underline{h} < 0\} = Z_-$ oder $\{\underline{h} < 0\} = Z_+$ gelten.
Wir müssen noch $\partial E \cap Q = \{(\underline{x}, f(\underline{x})) : \underline{x} \in Q'\}$ zeigen. '$\subseteq$' folgt wegen $\partial E \cap Q \subseteq N$. Weiter ist

$$\{(\underline{x}, f(\underline{x})) : \underline{x} \in Q'\} \subseteq \partial\big(\{(\underline{x}, y) : f(\underline{x}) < y\} \cup \big(\{(\underline{x}, y) : f(\underline{x}) > y\}\big) \subseteq \partial E \cap Q.$$

2. Wir haben bereits in Beispiel 27.5 gesehen, dass $\underline{\nu}$ aus (27.3) ein Einheitsnormalenfeld ist. Es bleibt also zu zeigen, dass $\underline{x} + t\underline{\nu}(\underline{x}) \notin E$ ist. Hierzu betrachten wir $\alpha : t \mapsto h(\underline{x} + t\underline{\nu}(\underline{x}))$ und berechnen $\alpha'(t) = \underline{D}h(\underline{x} + t\underline{\nu}(\underline{x})) \cdot \underline{\nu}(\underline{x})$ und damit $\alpha'(0) = (\underline{D}h(\underline{x}) \cdot \underline{D}h(\underline{x}))/||\underline{D}h(\underline{x})|| = ||\underline{D}h(\underline{x})|| > 0$. Da $E \cap Q = \{h < 0\}$, folgt die Behauptung. □

Definition 27.8 (Äußeres Normalenfeld). *In der Situation wie in Proposition 27.7 heißt $\underline{\nu}$ auch äußeres Einheitsnormalenfeld. In diesem Fall definieren wir für $\underline{f} \in \mathcal{B}(\partial_M E, \mathbb{R}^d)$*

$$\int_{\partial E} \underline{f} d\underline{\lambda}_E := \int_{(\partial_M E, \underline{\nu})} \underline{f} d\underline{\lambda}_E = \int_{\partial_M E} \langle \underline{f}, \underline{\nu} \rangle d\lambda_{\partial_M E},$$

falls die rechte Seite existiert.

Beispiel 27.9 (Äußeres Einheitsnormalenfeld bei Graphen und Niveaumengen). Sei $U \subseteq \mathbb{R}^{d-1}$ offen, $g \in \mathcal{C}^1(U, \mathbb{R})$ und $E = \{(\underline{x}, y), g(\underline{x}) < y\}$. Dann ist

$$\underline{\nu} : \underline{x} \mapsto \frac{(-\underline{D}g(\underline{x}), 1)}{\sqrt{1 + ||\underline{D}g(\underline{x})||^2}}$$

das äußere Einheitsnormalenfeld.
Denn: Wir betrachten $(\underline{x}, g(\underline{x})) + t\underline{\nu}(\underline{x}) = (\underline{x}(t), y(t))$ mit

$$\underline{x}(t) = \underline{x} - t\underline{D}g(\underline{x}), \qquad y(t) = g(\underline{x}) + t.$$

Dann gilt

$$\frac{d}{dt}(y(t) - g(\underline{x}(t)))\Big|_{t=0} = 1 + ||\underline{D}g(\underline{x})||^2 > 0.$$

Daraus folgt nun $(\underline{x}, y) + t\underline{\nu}(\underline{x}) \notin E$ und die Behauptung ist gezeigt.

27.3 Der Satz von Gauss

Definition 27.10. *Sei $E \subseteq \mathbb{R}^d$ offen und $n \geq 1$. Dann bezeichne*

$$\mathcal{C}^1(\overline{E}, \mathbb{R}^n) := \{\underline{f} \in \mathcal{C}^1(E, \mathbb{R}^n) : \underline{f}, \underline{\underline{D}}\,\underline{f} \text{ stetig auf } \overline{E} \text{ fortsetzbar.}\}.$$

Theorem 27.11 (Der Satz von Gauss). *Sei $E \subseteq \mathbb{R}^d$ ein $\mathcal{C}^1$-Polyeder und $\underline{f} \in \mathcal{C}^1(\overline{E}, \mathbb{R}^d)$. Dann gilt*

$$\int_E div(\underline{f}) d\lambda_E = \int_{\partial E} \underline{f} d\underline{\lambda}_{\partial E}. \tag{27.4}$$

Beweis. Wir gehen in drei Schritten vor. Zunächst beweisen wir den Satz für Gebiete mit glattem Rand und Vektorfelder mit kompakten Träger, danach für Gebiete mit glattem Rand. Zuletzt betrachten wir dann die gefordert $\mathcal{C}^1$-Polyeder, müssen uns hier also mit dem singulären Anteil des Randes beschäftigen.
1. Schritt: Sei $U \subseteq \mathbb{R}^{d-1}$ offen, $I = (a, b) \subseteq \mathbb{R}$ und $g \in \mathcal{C}^1(U, I)$ mit Graph $\Gamma = \{(\underline{x}, g(\underline{x})) : \underline{x} \in U\}$. Weiter sei $E = \{(\underline{x}, y) \in U \times I y < g(\underline{x})\}$ und $\underline{f} \in \mathcal{C}^1(\overline{E}, \mathbb{R}^d)$ habe kompakten Träger in $K \subseteq (U \times I) \cup \Gamma$. Dann gilt (27.4).

Die Bedingung, dass der Träger von $\underline{f}$ die kompakte Menge $K \subseteq (U \times I) \cup \Gamma$ ist, bedeutet: im Integral $\int_{\partial E} \underline{f} d\underline{\lambda}_{\partial E}$ spielen nur Werte von $\underline{f}$ auf ∂E eine Rolle. Wegen des kompakten Trägers ist $\overline{\{\underline{f} \neq \underline{0}\}} \subseteq \Gamma$.

Wir behaupten zunächst

$$\int_E \frac{\partial \underline{f}}{\partial x_k} d\lambda_E = \begin{cases} \displaystyle\int_U f_n(\underline{x}, g(\underline{x})) d\underline{x}, & k = n, \\ \displaystyle -\int_U f_k(\underline{x}, g(\underline{x})) \frac{\partial g(\underline{x})}{dx_k} d\underline{x}, & k = 1, ..., n-1. \end{cases} \tag{27.5}$$

Für $k = n$ folgt dies aus dem Hauptsatz der Differential- und Integralrechnung und dem Satz von Fubini und $f(\underline{x}, a) = 0$ mit

$$\int_E \frac{\partial \underline{f}}{\partial x_n} d\lambda_E = \int_U \int_a^{g(\underline{x})} \frac{\partial f_n}{\partial x_n} dx_n d\underline{x} = f_n(\underline{x}, g(\underline{x})).$$

Für $k = 1, ..., n-1$ wählen wir eine Funktion $\eta_\varepsilon \in \mathcal{C}^1(\mathbb{R}, [0,1])$ mit

$$\eta_\varepsilon(x) = \begin{cases} 0, & \text{für } x > -\varepsilon, \\ 1, & \text{für } x < -2\varepsilon. \end{cases}$$

Dann gilt mit majorisierter Konvergenz und partieller Integration (zuerst nach x_k, dann nach x_n) weges des kompakten Trägers von $\underline{f}$

$$\begin{aligned} \int_E \frac{\partial f_k(\underline{x})}{\partial x_k} d\underline{x} &= \lim_{\varepsilon \to 0} \int_E \frac{\partial f_k(\underline{x})}{\partial x_k} \eta_\varepsilon(x_n - g(\underline{x}_{-n})) d\underline{x} \\ &= \lim_{\varepsilon \to 0} \int_E f_k(\underline{x}) \eta_\varepsilon'(x_n - g(\underline{x}_{-n})) \frac{\partial g(\underline{x}_{-n})}{\partial x_k} d\underline{x} \\ &= -\lim_{\varepsilon \to 0} \int_E \frac{\partial f_k(\underline{x})}{\partial x_n} \eta_\varepsilon(x_n - g(\underline{x}_{-n})) \frac{\partial g(\underline{x}_{-n})}{\partial x_k} d\underline{x} \\ &= -\lim_{\varepsilon \to 0} \int_E \frac{\partial f_k(\underline{x})}{\partial x_n} \frac{\partial g(\underline{x}_{-n})}{\partial x_k} d\underline{x} \\ &= -\int_U \Big(\int_a^{g(\underline{x}_{-n})} \frac{\partial f_k(\underline{x})(\underline{x})}{\partial x_n} dx_n \Big) \frac{\partial g(\underline{x}_{-n})}{\partial x_k} d\underline{x} \\ &= -\int_U f_k(\underline{x}, g(\underline{x})) \frac{\partial g(\underline{x})}{\partial x_k} d\underline{x}. \end{aligned}$$

Damit ist (27.5) gezeigt.
Nun folgt die Aussage mit $\underline{\nu}$ aus Proposition 26.43 wegen

$$\begin{aligned} \int_E \operatorname{div}(\underline{f}) d\lambda_E &= \sum_{k=1}^n \int_E \frac{\partial f_k(\underline{x})}{\partial x_k} d\underline{x} \\ &= \int_U f_n(\underline{x}, g(\underline{x})) - \sum_{k=1}^{n-1} f_k(\underline{x}, g(\underline{x})) \frac{\partial g_k(\underline{x})}{d x_k} d\underline{x} \\ &= \int_U \langle \underline{f}(\underline{x}), \underline{\nu}(\underline{x}) \rangle \sqrt{1 + ||\underline{D}g(\underline{x})||^2} d\underline{x} \\ &= \int_{\partial E} \langle \underline{f}, \underline{\nu} \rangle d\lambda_{\partial E} = \int_{\partial E} \underline{f} d\underline{\lambda}_{\partial E}. \end{aligned}$$

2. Schritt: Sei $E \subseteq \mathbb{R}^d$ eine offene, beschränkte Menge, so dass ∂E eine orientierbare $d-1$-dimensionale Untermannigfaltigkeit ist. Weiter sei $\underline{\nu}$ das äußere Einheitsnormalenfeld und $\underline{f} \in \mathcal{C}^1(\overline{E}, \mathbb{R}^d)$. Dann gilt (27.4)*.*
Wir wählen zu $\underline{y} \in \partial E$ eine Quaderumgebung $U_{\underline{y}}$, so dass $\partial E \cap U_{\underline{y}}$ als Graph wie in Schritt 1 dargestellt werden kann. Für $\underline{y} \in E$ sei $U_{\underline{y}} = E$. Dann ist $(U_{\underline{y}})_{\underline{y} \in \bar{E}}$ eine offene Überdeckung der kompakten Menge $\bar{E}$, also gibt es eine endliche Teilüberdeckung $(U_{\underline{y}})_{\underline{y} \in D}$ mit $D \Subset \bar{E}$. Nach Lemma 26.11 gibt es eine $(U_{\underline{y}})_{\underline{y} \in D}$ untergeordnete Zerlegung der Eins $(h_{\underline{y}})_{\underline{y} \in D}$ (und beachten, dass der Träger von $h_{\underline{y}}$ eine kompakte Teilmenge von $U_{\underline{y}}$ ist). Wir betrachten nun die Vektorfelder $\underline{f} \cdot h_{\underline{y}}, \underline{y} \in D$. Dann gilt

$$\int_E \operatorname{div}(\underline{f} h_{\underline{y}}) d\lambda_E = \int_{\partial E} \underline{f} h_{\underline{y}} d\underline{\lambda}_{\partial E}.$$

Für $\underline{y} \in \partial E$ erfüllt nämlich $h_{\underline{y}}\underline{f}$ die Voraussetzungen von Schritt 1 und für $\underline{y} \in E$ sind beide Seiten 0, weil $h_{\underline{y}}\underline{f}$ eine kompakte Teilmenge von E als Träger hat. Damit ist die linke Seite 0 durch partielle Integration und die rechte da $h_{\underline{y}}\underline{f}$ auf ∂E identisch 0 ist. Nun können wir

$$\int_E \operatorname{div}(\underline{f}) d\lambda_E = \sum_{\underline{y}\in D} \int_E h_{\underline{y}} \operatorname{div}(\underline{f}) d\lambda_E = \sum_{\underline{y}\in D} \int_E \operatorname{div}(\underline{f} h_{\underline{y}}) d\lambda_E$$
$$= \sum_{\underline{y}\in D} \int_{\partial E} \underline{f} h_{\underline{y}} d\underline{\lambda}_{\partial E} = \int_{\partial E} \underline{f} d\underline{\lambda}_{\partial E}$$

schreiben.

3. Schritt: Der allgemeine Fall, E ist eine $\mathcal{C}^1$-Polyeder.

Ziel ist es nun noch, den singulären Rand $\partial_S E$ zu behandeln. Nach Definition der rechten Seite in (27.4) spielt bei der Integration nur der reguläre Rand $\partial_M E$ eine Rolle. Zuerst stellen wir fest, dass der Gauss'sche Integralsatz nach Schritt 2 schon gezeigt ist, falls der Träger von $\underline{f}$ eine kompakte Teilmenge von $\bar{E}$ ist, die $\partial_S E$ nicht trifft.

Zunächst ist $\partial_S E$ eine $m-1$-Nullmenge. Wegen der Beschränktheit von E ist $\partial_S E$ kompakt. Für $\varepsilon > 0$ seien $W_1, ..., W_{N_\varepsilon}$ eine Überdeckung von $\partial_S E$ mit Seitenlängen $r_1, ..., r_{N_\varepsilon} > 0$ und $r_1^{d-1} + \cdots + r_{N_\varepsilon}^{d-1} < \varepsilon$. Seien $W_1*, ..., W_{N_\varepsilon}^*$ Würfel mit Seitenlängen $2r_1, ..., 2r_{N_\varepsilon}$ und denselben Mittelpunkten. Wir wählen nun Funktionen $h_1, ..., h_{N_\varepsilon} \in \mathcal{C}^1(\mathbb{R}^d)$, so dass

$$1_{W_k} \leq h_k \leq 1_{W_k^*}, \qquad ||\underline{D}h_k|| \leq \frac{c}{r_k} 1_{W_k^*}$$

für eine (nur von d abhängige) Konstante c. (Um eine solche Funktion zu erhalten, wählen wir zuerst ein $\varphi \in \mathcal{C}^1(\mathbb{R})$ mit $1_{[-1,1]} \leq \varphi \leq 1_{[-2,2]}$ und setzen dann, falls $\underline{0}$ der Mittelpunkt eines Würfels mit Seitenlänge r ist, $h_k(\underline{x}) = \varphi(x_1/r) \cdots \varphi(x_d/r)$.) Wir setzen nun

$$h_\varepsilon = \prod_{k=1}^{N_\varepsilon} (1 - h_k)$$

und bemerken, dass $h \cdot \underline{f}$ eine kompakte Teilmenge von $\bar{E}$ als Träger hat, die $\partial_S E$ nicht trifft. Also folgt nach Schritt 2, dass

$$\int_E \operatorname{div}(h_\varepsilon \underline{f}) d\lambda_E = \int_{\partial E} \langle h_\varepsilon \underline{f}, \underline{\nu} \rangle d\lambda_{\partial E}.$$

Nun konvergiert $h_\varepsilon \underline{f}$ punktweise gegen $1_{(\partial_S E)^c} \underline{f}$ und damit konvergiert die rechte Seite gegen die rechte Seite von (27.4) nach dem Satz von der majorisierten Konvergenz. Für die linke Seite gilt

$$\operatorname{div}(h_\varepsilon \underline{f}) = h_\varepsilon \operatorname{div}(\underline{f}) + \underline{D}h_\varepsilon \cdot \underline{f}.$$

Nun ist

$$\Big| \int_E (\underline{D}h_\varepsilon \cdot \underline{f}) d\lambda_E \Big| \leq C \int_E ||\underline{D}h_\varepsilon|| d\lambda_E \leq C \sum_{k=1}^{N_\varepsilon} \frac{1}{r_k} \int_E 1_{W_k^*} d\lambda_E \leq C \sum_{k=1}^{N_\varepsilon} r_k^{d-1} < C\varepsilon$$

für Konstanten $C > 0$, die von Mal zu Mal verschieden sind. Nun schreiben wir

$$\int_E \operatorname{div}(\underline{f}) d\lambda_E = \lim_{\varepsilon\to 0} \int_E h_\varepsilon \operatorname{div}(\underline{f}) d\lambda_E + \int_E (\underline{D}h_\varepsilon \cdot \underline{f}) d\lambda_E$$
$$= \lim_{\varepsilon\to 0} \int_E \operatorname{div}(h_\varepsilon \underline{f}) d\lambda_E = \lim_{\varepsilon\to 0} \int_{\partial E} \langle h_\varepsilon \underline{f}, \underline{\nu} \rangle d\lambda_{\partial E} = \int_{\partial E} \underline{f} d\underline{\lambda}_{\partial E}.$$

□

Beispiel 27.12 (Oberfläche einer d-dimensionalen Kugel). Wir berechnen nun die ($d-1$-dimensionale) Oberfläche $\lambda^{d-1}(\partial B_1(\underline{0}))$ einer (d-dimensionalen) Kugel $B_1(\underline{0})$. Hierzu definieren wir das Vektorfeld $\underline{f}(\underline{x}) = \underline{x}$. Für das äußere Normalenfeld $\underline{\nu}$ and $B_1(\underline{0})$ gilt $\underline{\nu}(\underline{x}) = \underline{x}$ und damit $\langle \underline{f}(\underline{x}), \underline{\nu}(\underline{x})\rangle = 1$. Deswegen ist mit $\operatorname{div}(\underline{f})(\underline{x}) = d$ gerade

$$\gamma_d := \lambda^{d-1}(\partial B_1(\underline{0})) = \int_{B_1(\underline{0})} \operatorname{div}(\underline{f})(\underline{x})\lambda^d(d\underline{x}) = d\alpha_d.$$

Dieses Resultat kennen wir schon aus Beispiel 26.44, (26.7).

27.4 Die Green'sche Formeln und harmonische Funktionen

Wir bringen nun einige Anwendungen des Gauss'schen Integralsatzes. Diese haben mit harmonischen Funktionen, d.h. Lösungen $u \in \mathcal{C}^2(E,\mathbb{R})$ für eine offene, zusammenhängende Menge $E \subseteq \mathbb{R}^d$ der Gleichung

$$\Delta u := \frac{\partial^2 u}{\partial x_1} + \cdots + \frac{\partial^2 u}{\partial x_d} = 0.$$

Solche Funktionen haben wir bereits in Beispiel 14.3 kennen gelernt. Wir haben dort bereits gesehen, dass (abhängig von der Dimension d)

$$u(\underline{x}) = \begin{cases} \frac{1}{2\pi} \cdot \log|x|, & \text{für } d = 2, \\ -\frac{1}{(d-2)\gamma_d} \cdot \frac{1}{|x|^{d-2}}, & \text{für } d > 2, \end{cases}$$

wobei γ_d wie in (26.7) die Oberfläche der $d-1$-dimensionalen Kugeloberfläche ist, harmonische, rotationssymmetrische Funktionen sind. (Man beachte, dass wir dort solche Funktionen bis auf einen konstanten Faktor bestimmt haben.) Weiter sind mit der Funktion u auch die Funktionen

$$u_{\underline{a}}\underline{x} \mapsto u(\underline{x} - \underline{a}) \tag{27.6}$$

für ein beliebiges $\underline{a} \in \mathbb{R}^d$ harmonisch. Wir beginnen unsere Untersuchung mit den Green'schen Formeln. Diese sind eine Anwendung des Integralsatzes auf ein Vektorfeld der Form $\underline{x} \mapsto \underline{\operatorname{grad}} h(\underline{x})$ für ein $h \in \mathcal{C}^1(E,\mathbb{R})$ mit $E \subseteq \mathbb{R}^d$. (Solche Vektorfelder hatten wir Gradientenfelder genannt.)

Theorem 27.13 (Green'sche Formeln). *Sei $E \subseteq \mathbb{R}^d$ ein $\mathcal{C}^1$-Polyeder, $\underline{\nu}$ das äußere Normalenfeld an E und $\underline{f}, \underline{g} \in \mathcal{C}^2(\overline{E},\mathbb{R})$. Dann gilt*

$$\int_E \langle \underline{grad}\, f, \underline{grad}\, g\rangle = \int_{\partial E} f\langle \underline{grad}\, g, \underline{\nu}\rangle d\underline{\lambda}_{\partial E} - \int_E f\, \Delta g d\lambda_E,$$
$$\int_E (f\,\Delta g - g\,\Delta f) d\lambda_E = \int_{\partial E} (f\langle \underline{grad}\, g, \underline{\nu}\rangle - g\langle \underline{grad}\, f, \underline{\nu}\rangle) d\underline{\lambda}_{\partial E}.$$

Beweis. Wir bemerken zunächst, dass

$$\operatorname{div}(\underline{\operatorname{grad}}\, u) = \operatorname{div}\Big(\frac{\partial u}{\partial x_1}, \cdots, \frac{\partial u}{\partial x_d}\Big) = \Delta u.$$

Wendet man nun den Gauss'schen Integralsatz auf das Gradientenfeld $f\underline{\operatorname{grad}}\, g$ an, so gilt

$$\operatorname{div}(f\underline{\operatorname{grad}}\, g) = \langle \underline{\operatorname{grad}}\, f, \underline{\operatorname{grad}}\, g\rangle + f\,\Delta g$$

und damit durch Anwendung des Gauss'schen Satzes

$$\int_E \langle \underline{\text{grad}}\, f, \underline{\text{grad}}\, g\rangle + \int_E f\,\Delta g d\lambda_E = \int_{\partial E} f \underline{\text{grad}}\, g d\underline{\lambda}_{\partial E} = \int_{\partial E} f\langle \underline{\text{grad}}\, g, \underline{\nu}\rangle d\underline{\lambda}_{\partial E}.$$

Die zweite Formel erhält man, indem man in die Rollen von f und g in der ersten Formel vertauscht und subtrahiert. □

Wir kommen nun zur Anwendung auf harmonische Funktionen. Zunächst benötigen wir ein Lemma.

Lemma 27.14 (Darstellung einer harmonischen Funktion). *Sei $D \subseteq \mathbb{R}^d$ eine offene Menge, $E \subseteq D$ ein beschränktes Polyeder mit äußerem Normalenfeld $\underline{\nu}$ und $h \in \mathcal{C}^2(D, \mathbb{R})$ eine (auf D) harmonische Funktion. Dann gilt (mit der Funktion $u_{\underline{a}}$ aus (27.6))*

$$h(\underline{a}) = \int_{\partial E} (h\langle \underline{\mathit{grad}}\, u_{\underline{a}}, \underline{\nu}\rangle - u_{\underline{a}}\langle \underline{\mathit{grad}}\, h, \underline{\nu}\rangle) d\underline{\lambda}_{\partial E}.$$

Beweis. Wir wählen ein $\varepsilon > 0$ mit $B_\varepsilon(a) \subseteq E^\circ$ und setzen $E_{\underline{a}} := E \setminus \overline{B}_\varepsilon(\underline{a})$. (Damit ist $\partial E_{\underline{a}} = \partial E \uplus \partial B_\varepsilon(\underline{a})$.) Für $\underline{x} \in B_\varepsilon(\underline{a})$ gilt $\underline{\nu}(\underline{x}) = -\frac{x-a}{\varepsilon}$ (das ist das äußere Normalenfeld an $E_{\underline{a}}$) und $\underline{\text{grad}}\, u_{\underline{a}}(\underline{x}) = \frac{1}{\gamma_d}\frac{x-a}{\varepsilon^d}$ mit γ_d aus Beispiel 26.44. Wir berechnen damit direkt für $\underline{x} \in B_\varepsilon(\underline{a})$

$$\langle \underline{\text{grad}}\, u_{\underline{a}}(\underline{x}), \underline{\nu}(\underline{x})\rangle = -\frac{1}{\gamma_d \varepsilon^{d-1}}. \tag{27.7}$$

Die zweite Green'sche Formel (angewandt auf die Funktionen $u_{\underline{a}}$ und h) ergibt somit

$$\int_{\partial E} (h\langle \underline{\text{grad}}\, u_{\underline{a}}, \underline{\nu}\rangle - u_{\underline{a}}\langle \underline{\text{grad}}\, h, \underline{\nu}\rangle) d\underline{\lambda}_{\partial E} = -\int_{\partial B_\varepsilon(\underline{a})} (h\langle \underline{\text{grad}}\, u_{\underline{a}}, \underline{\nu}\rangle - u_{\underline{a}}\langle \underline{\text{grad}}\, h, \underline{\nu}\rangle) d\underline{\lambda}_{\partial E}$$

Wir zeigen nun

$$h(\underline{a}) = -\lim_{\varepsilon \to 0} \int_{\partial B_\varepsilon(\underline{a})} h\langle \underline{\text{grad}}\, u_{\underline{a}}, \underline{\nu}\rangle d\underline{\lambda}_{\partial E}, \qquad 0 = \lim_{\varepsilon \to 0} \int_{\partial B_\varepsilon(\underline{a})} u_{\underline{a}}\langle \underline{\text{grad}}\, h, \underline{\nu}\rangle d\underline{\lambda}_{\partial E},$$

woraus die Behauptung folgt. Für die erste Behauptung genügt es, (27.7) und die Stetigkeit von h zu benutzen, da der Flächeninhalt von $\partial B_\varepsilon(\underline{a})$ gerade $\gamma_d \varepsilon^{d-1}$ ist. Weiter sei $M_\varepsilon := \sup_{\underline{x} \in B_\varepsilon(\underline{0})} |\underline{\text{grad}}\, h(\underline{x})|$. Dann gilt $|\langle \underline{\text{grad}}\, h(\underline{x}), \underline{\nu}(\underline{x})\rangle \leq M_\varepsilon$ und damit auch

$$\Big| \int_{\partial B_\varepsilon(\underline{a})} u_{\underline{a}}\langle \underline{\text{grad}}\, h, \underline{\nu}\rangle) d\underline{\lambda}_{\partial E} \Big| \leq \begin{cases} \frac{M_\varepsilon}{2\pi}(\log \varepsilon) \cdot 2\pi\varepsilon = M_\varepsilon \cdot \varepsilon \log \varepsilon, & \text{für } d = 2 \\ \frac{M_\varepsilon \gamma_d \varepsilon^{d-1}}{(d-2)\gamma_d \varepsilon^{d-2}} = \frac{M_\varepsilon \cdot \varepsilon}{d-2}, & \text{für } d > 2 \end{cases} \xrightarrow{\varepsilon \to 0} 0.$$

Nun folgt also die Behauptung. □

Proposition 27.15 (Mittelwerteigenschaft). *Sei $d \geq 2$, $E \subseteq \mathbb{R}^d$ offen und $h \in \mathcal{C}^2(E, \mathbb{R})$ harmonisch. Dann gilt, falls $B_r(\underline{a}) \subseteq E$,*

$$h(\underline{a}) = \frac{1}{\gamma_d r^{d-1}} \int_{\partial B_r(\underline{a})} h d\underline{\lambda}_{\partial B_r(\underline{a})}.$$

Beweis. Wir verwenden die Darstellung von h aus Lemma 27.14 für $E = B_r(\underline{a})$ und $\langle \underline{\text{grad}}\, u_{\underline{a}}(\underline{x}), \underline{\nu}(\underline{x})\rangle = \frac{1}{\gamma_d r^{d-1}}$ für $\underline{x} \in B_r(\underline{a})$ (siehe den Beweis des letzten Lemmas). Daraus folgt

$$\begin{aligned} h(\underline{a}) - \frac{1}{\gamma_d r^{d-1}} \int_{\partial B_r(\underline{a})} h d\underline{\lambda}_{\partial B_r(\underline{a})} &= -u(r) \int_{\partial B_r(\underline{a})} \langle \underline{\text{grad}}\, h, \underline{\nu}\rangle d\underline{\lambda}_{\partial B_r(\underline{a})} \\ &= -u(r) \int_{B_r(\underline{a})} \text{div}(\underline{\text{grad}}\, h) d\lambda_{B_r(\underline{a})} = 0 \end{aligned}$$

nach dem Gauss'schen Integralsatz wegen $\text{div}(\underline{\text{grad}}\, h) = 0$. □

Wir beenden diesen Abschnitt mit dem Maximumprinzip für harmonische Funktionen, das besagt, dass sie Maxima (und Minima) nur auf dem Rand eines Gebietes annehmen.

Korollar 27.16 (Maximumsprinzip für harmonische Funktionen). *Sei $E \subseteq \mathbb{R}^d$ eine zusamenhängende, offene Menge und $h \in \mathcal{C}^2(E, \mathbb{R})$ harmonisch auf $\bar{E}$. Ist h nicht konstant, dann nimmt h ihr globales Maximum (und globales Minimum) auf ∂E an.*

Beweis. Angenommen, das globale Maximum M von h länge in $\underline{a} \in E$, also $M = h(\underline{a}) = \sup_{\underline{x} \in E} h(\underline{x})$. Dann gilt mit der Mittelwerteigenschaft für ein $\varepsilon > 0$ mit $B_\varepsilon(\underline{a}) \subseteq E$

$$\frac{1}{\gamma_d \varepsilon^{d-1}} \int_{\partial B_\varepsilon(\underline{a})} (M - h(\underline{x})) d\underline{\lambda}_{\partial B_\varepsilon(\underline{a})}(d\underline{x}) = 0.$$

Da $M - h \geq 0$ nach Definition von M, muss also $M = h$ gelten. Damit wäre h konstant, was aber ausgeschlossen war. Dies beweist die Behauptung. □

27.5 Der Satz von Stokes

Bereits in Bemerkung 17.19 haben wir die Rotation eines Vektorfeldes definiert. Diese benötigen wir nun auch im Satz von Stokes. Wir besprechen hier nicht die allgemeinste Version dieses Satzes, sondern nur zwei Versionen, eine in $d = 2$ (Theorem 27.19) und eine in $d = 3$ (Theorem 27.23). Wir wiederholen zunächst die Rotation.

Definition 27.17 (Rotation).
1. *Sei $E \subseteq \mathbb{R}^2$ offen und $\underline{f} = (f_1, f_2) \in \mathcal{C}^1(E, \mathbb{R}^2)$. Dann ist*
$$rot(\underline{f}) := D_1 f_2 - D_2 f_1$$
die Rotation *von $\underline{f}$.*
2. *Sei $E \subseteq \mathbb{R}^3$ offen und $\underline{f} = (f_1, f_2, f_3) \in\in \mathcal{C}^1(E, \mathbb{R}^3)$. Dann ist*
$$\underline{rot}(\underline{f}) := \big(D_2 f_3 - D_3 f_2, \quad D_3 f_1 - D_1 f_3, \quad D_1 f_2 - D_2 f_1\big)$$
die Rotation *von $\underline{f}$.*

Bemerkung 27.18 (Vektorprodukt).
1. Für $\underline{x} = (x_1, x_2, x_3), \underline{y} = (y_1, y_2, y_3) \in \mathbb{R}^3$ ist bekanntlich
$$\underline{x} \times \underline{y} = \big(x_2 y_3 - x_3 y_2 \quad x_3 y_1 - x_1 y_3 \quad x_1 y_2 - x_2 y_1\big)$$
das Vektorprodukt (oder Kreuzprodukt) aus $\underline{x}, \underline{y}$. Dies ist ein Vektor, der sowohl auf $\underline{x}$ als auch auf $\underline{y}$ senkrecht steht (d.h. $\langle x, \underline{x} \times \underline{y}\rangle = \langle y, \underline{x} \times \underline{y}\rangle = 0$) und als Länge den Flächeninhalt des von $\underline{x}$ und $\underline{y}$ aufgespannten Parallelogramms hat. Wegen Proposition 26.37 ist damit $|\underline{x} \times \underline{y}| = \sqrt{\det(\underline{\underline{A}}^\top \underline{\underline{A}})}$ mit $\underline{\underline{A}} = (\underline{x}^\top, \underline{y}^\top)$.

2. Mit Hilfe des Vektorproduktes und $\underline{\nabla} := (D_1, D_2, D_3)$ (im $\mathbb{R}^3$) schreiben wir auch
$$\underline{\text{rot}}(\underline{f}) := \underline{\nabla} \times \underline{f}.$$

Theorem 27.19 (Satz von Stokes in $\mathbb{R}^2$). *Sei $D \subseteq \mathbb{R}^2$ ein beschränktes Gebiet und $E \subseteq D$ sei offen mit stetig differenzierbarem Rand ∂E, d.h. es gibt eine Einbettung $\underline{\gamma} : [a,b] \to \partial E$. Außerdem gelte für das äußere Normalenfeld*
$$\underline{\nu}(\underline{\gamma}(t)) = \frac{1}{|\underline{\gamma}'(t)|} \underline{D}\, \underline{\gamma}'(t)$$
mit $\underline{D}(x,y) = (y, -x)$. (Damit liegt E links *von ∂E.) Ist $\underline{f} = (f_1, f_2) \in \mathcal{C}^1(E, \mathbb{R}^2)$, so gilt*
$$\int_E rot(\underline{f})(\underline{x}) \lambda^2(d\underline{x}) = \int_a^b \langle \underline{f}(\underline{\gamma}(t)), \underline{\gamma}'(t) \rangle dt.$$

Beweis. Wir setzen $\underline{g} = (g_1, g_2)$ mit $g_1 = f_2, g_2 = -f_1$ und verwenden den Gauss'schen Integralsatz, Theorem 27.11. Damit ergibt sich
$$\begin{aligned}
\int_E \underline{\text{rot}}(\underline{f}) d\lambda^2 &= \int_E \big(D_1 f_2(\underline{x}) - D_2 f_1(\underline{x})\big) \lambda^2(d\underline{x}) \\
&= \int_E \text{div}(\underline{g}(\underline{x})) \lambda^2(d\underline{x}) \\
&= \int_{\partial E} \underline{g} d\underline{\lambda}_{\partial E} = \int_{\partial E} \langle \underline{g}, \underline{\nu} \rangle d\underline{\lambda}_{\partial E} \\
&= \int_a^b \frac{1}{|\underline{\gamma}'(t)|} \langle \underline{g}(\underline{\gamma}(t)), \underline{D}\underline{\gamma}'(t) \rangle |\underline{\gamma}'(t)| d\underline{\lambda}_{\partial E} \\
&= \int_a^b \langle (f_2, -f_1)(\underline{\gamma}(t)), (\gamma_2'(t), -\gamma_1'(t)) \rangle dt \\
&= \int_a^b \langle \underline{f}(\underline{\gamma}(t)), \underline{\gamma}'(t) \rangle dt.
\end{aligned}$$
□

Wir verwenden das letzte Resultat zur Berechnung von Flächeninhalten.

Korollar 27.20 (Flächenformel). *Sei $E \subseteq \mathbb{R}^2$ eine Fläche mit einer Einbettung $\underline{\gamma} = (\gamma_1, \gamma_2) \in \mathcal{C}^1(I, \mathbb{R}^2)$ für $I = [a,b]$ des Randes ∂E. Dann gilt*
$$\lambda^2(E) = \frac{1}{2} \int_a^b \big(\gamma_1(t)\gamma_2'(t) - \gamma_2(t)\gamma_1'(t)\big) dt.$$

Beweis. Wir verwenden Theorem 27.19 mit $\underline{f} = (f_1, f_2)$ und $f_1(x,y) = -y, f_2(x,y) = x$. Daraus folgt $\text{rot}(\underline{f}) = 2$ und damit die Aussage. □

Beispiel 27.21 (Kreisfläche). Wieder einmal berechnen wir den Inhalt eines Kreises. Sei also $E = B_r(\underline{0})$ und $\underline{\gamma}(t) = (r\cos(t), r\sin(t))$ für $t \in [0, 2\pi]$, also $\partial E = \gamma([0, 2\pi])$. Damit berechnen wir aus dem letzten Korollar
$$\lambda^2(B_r(\underline{0})) = \frac{r^2}{2} \int_0^{2\pi} \big(\cos^2(t) + \sin^2(t)\big) dt = r^2\pi.$$

Wir kommen nun zum Satz von Stokes in $\mathbb{R}^3$. Zunächst benötigen wir eine Darstellung des Einheitsnormalenfeldes.

Lemma 27.22 (Einheitsnormalenfeld). *Sei $E' \subseteq \mathbb{R}^2$ offen und $\underline{\gamma} \in \mathcal{C}^1(E', \mathbb{R}^3)$ eine Einbettung. Dann ist*

$$\underline{\nu}_E = -\frac{D_1\underline{\gamma} \times D_2\underline{\gamma}}{|D_1\underline{\gamma} \times D_2\underline{\gamma}|} \circ \underline{\gamma}^{-1}$$

mit

$$D_1\underline{\gamma} \times D_2\underline{\gamma} = \big(D_1\gamma_2 D_2\gamma_3 - D_1\gamma_3 D_2\gamma_2, \quad D_1\gamma_3 D_2\gamma_1 - D_1\gamma_1 D_2\gamma_3, \quad D_1\gamma_1 D_2\gamma_2 - D_1\gamma_2 D_2\gamma_1\big)$$

ein stetiges Einheitsnormalenfeld auf $E = \underline{\gamma}(E')$.

Beweis. Nach Proposition 26.33 ist

$$N_{\underline{\gamma}(\underline{x})}E = \operatorname{Ker}(\underline{\underline{D}}\,\underline{\gamma}(\underline{x}))^\top = \operatorname{Ker}\begin{pmatrix} D_1\gamma_1 & D_1\gamma_2 & D_1\gamma_3 \\ D_2\gamma_1 & D_2\gamma_2 & D_2\gamma_3 \end{pmatrix} = \operatorname{Ker}\begin{pmatrix} D_1\underline{\gamma} \\ D_2\underline{\gamma} \end{pmatrix}.$$

Offenbar ist $D_1\underline{\gamma} \times D_2\underline{\gamma}$ ein Vektorprodukt im $\mathbb{R}^3$ und damit sowohl auf $D_1\underline{\gamma}$ als auch auf $D_2\underline{\gamma}$ senkrecht. Also ist $D_1\underline{\gamma} \cdot (D_1\underline{\gamma} \times D_2\underline{\gamma})^\top = D_2\underline{\gamma} \cdot (D_1\underline{\gamma} \times D_2\underline{\gamma})^\top = 0$, also $\underline{\nu}_E(\underline{\gamma}(\underline{x})) \in N_{\underline{\gamma}(\underline{x})}E$. Daraus folgt die Behauptung. □

Theorem 27.23 (Satz von Stokes im $\mathbb{R}^3$). *Sei $D \subseteq \mathbb{R}^3$ eine 2-dimensionale Untermannigfaltigkeit, die die Spur einer Einbettung $\underline{\beta} \in \mathcal{C}^1(D', \mathbb{R}^3)$ für ein $D' \subseteq \mathbb{R}^2$ offen ist. Weiter habe $E' \subseteq D'$ einen stetig differenzierbaren Rand, der durch eine Einbettung $\underline{\gamma}' \in \mathcal{C}^1([a,b], \partial E')$ für $a < b$ parametrisiert ist. Für die Parametrisierung $\underline{\gamma} = \underline{\beta} \circ \underline{\gamma}'$ von ∂E gilt dann für $E = \underline{\gamma}(E)$ und $\underline{f} \in \mathcal{C}^2(D, \mathbb{R}^3)$ und dem Einheitsnormalenfeld $\underline{\nu}_E$ aus Lemma 27.22*

$$\int_E \langle \underline{\operatorname{rot}}(\underline{f}), \underline{\nu}_E \rangle d\lambda_E = \int_{\partial E} \frac{1}{|\dot{\underline{\gamma}}|} \langle \underline{f}, \dot{\underline{\gamma}} \rangle d\lambda_{\partial E}.$$

Bemerkung 27.24 (Rotation). Sei $D \subseteq \mathbb{R}^3$ eine 2-dimensionale Untermannigfaltigkeit und $E_1, E_2, ...$ eine Folge von Teilmengen von D mit $E_k \downarrow \{x\}$, auf die wir den Stokes'schen Satz anwenden können. Dann gilt, falls $\underline{\gamma}_k$ den Rand von ∂E_k nach der Bogenlänge parametrisiert (d.h. $|\dot{\gamma}_k| = 1$)

$$\langle \underline{\operatorname{rot}}(\underline{f})(\underline{x}), \underline{\nu}_E(\underline{x}) \rangle = \lim_{k\to\infty} \frac{1}{\lambda_D(E_k)} \int_{\partial E_k} \langle \underline{f}, \dot{\underline{\gamma}} \rangle d\lambda_{\partial E_k}.$$

Auf der rechten Seite steht die Zirkulation des Feldes $\underline{f}$ entlang des Randes ∂E_k, die man aufgrund der Gleichheit als die Rotation in Richtung $\underline{\nu}_E(\underline{x})$ interpretiert.

Beweis. Zunächst eine Vorüberlegung. Es gilt

$$(\underline{\underline{D}}\,\underline{\beta})^\top \cdot \big(\underline{f} \circ \underline{\beta}\big)^\top = \begin{pmatrix} D_1\beta_1 & D_1\beta_2 & D_1\beta_3 \\ D_2\beta_1 & D_2\beta_2 & D_2\beta_3 \end{pmatrix} \cdot \begin{pmatrix} f_1 \circ \underline{\beta} \\ f_2 \circ \underline{\beta} \\ f_3 \circ \underline{\beta} \end{pmatrix}.$$

Wir schreiben mit Theorem 27.19 (für das vierte Gleichheitszeichen)

$$\begin{aligned}
\int_{\partial E} \frac{1}{|\underline{\dot\gamma}|} \langle \underline{f}, \underline{\dot\gamma}\rangle d\lambda_{\partial E} &= \int_a^b \langle \underline{f}(\underline{\gamma}(t)), \underline{\dot\gamma}(t)\rangle dt \\
&= \int_a^b \langle \underline{f}(\underline{\gamma}(t)), \underline{\underline{D}}\,\underline{\beta}(\underline{\gamma}'(t)) \cdot \underline{\dot\gamma}'(t)\rangle dt \\
&= \int_a^b \langle (\underline{\underline{D}}\,\underline{\beta}(\underline{\gamma}'(t)))^\top \cdot \underline{f} \circ \underline{\beta}(\underline{\gamma}'(t)), \underline{\dot\gamma}'(t)\rangle \\
&= \int_{E'} \big(D_1(D_2\underline{\beta} \cdot (\underline{f} \circ \underline{\beta})^\top) - D_2(D_1\underline{\beta} \cdot (\underline{f} \circ \underline{\beta})^\top)\big) d\lambda_{E'} \\
&= \int_{E'} \big(D_2\underline{\beta} \cdot (D_1(\underline{f} \circ \underline{\beta})^\top) - D_1\underline{\beta} \cdot (D_2(\underline{f} \circ \underline{\beta})^\top)\big) d\lambda_{E'} \\
&= \int_{E'} \big(D_2\underline{\beta} \cdot \underline{\underline{D}}\,\underline{f} \cdot (D_1\underline{\beta})^\top - D_1\underline{\beta} \cdot \underline{\underline{D}}\,\underline{f} \cdot (D_2\underline{\beta})^\top\big) d\lambda_{E'} \\
&= \int_{E'} D_1\underline{\beta} \cdot ((\underline{\underline{D}}\,\underline{f})^\top - \underline{\underline{D}}\,\underline{f}) \cdot D_2\underline{\beta} d\lambda_{E'}.
\end{aligned}$$

Nun ist mit $\underline{\nu}_E$ aus Lemma 27.22 und der Gram'schen Determinante $g_{\underline{\beta}}$

$$\begin{aligned}
&D_1\underline{\beta} \cdot ((\underline{\underline{D}}\,\underline{f})^\top - \underline{\underline{D}}\,\underline{f}) \cdot D_2\underline{\beta} \\
&\quad = \begin{pmatrix} D_1\beta_1 & D_1\beta_2 & D_1\beta_3 \end{pmatrix} \begin{pmatrix} 0 & D_2f_1 - D_1f_2 & D_3f_1 - D_1f_3 \\ D_1f_2 - D_2f_1 & 0 & D_3f_2 - D_2f_3 \\ D_1f_3 - D_3f_1 & D_2f_3 - D_3f_2 & 0 \end{pmatrix} \begin{pmatrix} D_2\beta_1 \\ D_2\beta_2 \\ D_2\beta_3 \end{pmatrix} \\
&\quad = (D_2f_3 - D_3f_2)(D_1\beta_3 D_2\beta_2 - D_1\beta_2 D_2\beta_3) \\
&\qquad\qquad + (D_3f_1 - D_1f_3)(D_1\beta_1 D_2\beta_3 - D_1\beta_3 D_2\beta_1) \\
&\qquad\qquad + (D_1f_2 - D_2f_1)(D_1\beta_2 D_2\beta_1 - D_1\beta_1 D_2\beta_2) \\
&\quad = -\underline{\mathrm{rot}}(\underline{f}) \cdot (D_1\underline{\beta} \times D_2\underline{\beta}) \\
&\quad = \langle \underline{\mathrm{rot}}(\underline{f}), \underline{\nu}_E \circ \underline{\beta}\rangle |D_1\underline{\beta} \times D_2\underline{\beta}| \\
&\quad = \langle \underline{\mathrm{rot}}(\underline{f}), \underline{\nu}_E \circ \underline{\beta}\rangle \sqrt{g_{\underline{\beta}}}
\end{aligned}$$

und damit

$$\int_{\partial E} \frac{1}{|\underline{\dot\gamma}|} \langle \underline{f}, \underline{\dot\gamma}\rangle d\lambda_{\partial E} = \int_{E'} \langle \underline{\mathrm{rot}}(\underline{f}), \underline{\nu}_E \circ \underline{\beta}\rangle \sqrt{g_{\underline{\beta}}} d\lambda_{E'} = \int_E \langle \underline{\mathrm{rot}}(\underline{f}), \underline{\nu}_E\rangle d\lambda_E$$

und die Behauptung ist gezeigt. □