Analysis III

Walter Bergweiler
Mathematisches Seminar, CAU Kiel

Wintersemester 2018/19
Fassung vom 13. Februar 2019

Inhaltsverzeichnis

Literatur

Der Teil über Maß- und Integrationstheorie folgt in weiten Teilen einem Vorlesungskript von Markus Nieß (Wintersemester 2014/15) und Michael Gnewuch (Wintersemester 2016/17 und 2017/18) an der CAU.

Darüberhinaus seien zu diesem Thema folgende Bücher genannt (die verlinkten e-Books sollten aus dem Netz der CAU kostenfrei zugänglich sein):

- H. Amann, J. Escher, Analysis III, Birkhäuser, 2001.
  `https://link.springer.com/book/10.1007/978-3-7643-8884-3`
- H. Bauer, Maß- und Integrationstheorie, Walter de Gruyter, 1990.
- J. Elstrodt, Maß- und Integrationstheorie, Springer, 1996.
  `https://link.springer.com/book/10.1007/978-3-662-57939-8`
- O. Forster, Analysis 3, Vieweg, 1981.
  `https://link.springer.com/book/10.1007/978-3-8348-2374-8`
- K. Königsberger, Analysis 2, Springer, 1993.
- W. Rudin, Real and Complex Analysis, McGraw-Hill, 1987

Für Differentialgleichungen sei insbesondere auf folgende Bücher verwiesen:

- H. Amann, Gewöhnliche Differentialgleichungen, de Gruyter, 1983.
- E. A. Coddington & N. Levinson, Theory of Ordinary Differential Equations, McGraw Hill, 1955.
- L. Grüne, O. Junge, Gewöhnliche Differentialgleichungen, Springer, 2009.
  `https://link.springer.com/book/10.1007/978-3-658-10241-8`
- H. Heuser, Gewöhnliche Differentialgleichungen, Teubner, 1989.
- W. Walter, Gewöhnliche Differentialgleichungen, Springer, 1972.
- G. J. Wirsching, Gewöhnliche Differentialgleichungen, Teubner, 2006.
  `https://link.springer.com/book/10.1007/978-3-8351-9044-3`

Beide Themen werden behandelt in

- D. Werner, Einführung in die höhere Analysis, Springer, 2009.
  `https://link.springer.com/book/10.1007/978-3-540-79696-1`

Generell sei gesagt, dass die Literatur zu beiden Themen sehr umfangreich ist und die obige Liste nur eine kleine Auswahl darstellt.

I Maß und Integral

Vorbemerkungen

In Analysis I und II wurde Differentialrechnung in einer und mehreren Veränderlichen sowie Integralrechnung in einer Veränderlichen behandelt. Ein Aspekt dieser Vorlesung ist die Integralrechnung in mehreren Veränderlichen. Dabei werden wir statt des Riemann-Integrals das Lebesgue-Integral behandeln. Wie wir sehen werden, hat das auch im Eindimensionalen gewisse Vorteile. So gelten etwa "schönere" Konvergenzsätze. (Wann folgt aus $f_n \to f$, dass $\int f_n \to \int f$?)

Die Idee beim Riemann-Integral ist, für eine Funktion $f\colon [a,b] \to \mathbb{R}$ das Intervall $[a,b]$ in Teilintervalle I_j zu zerlegen und dann Summen der Form $\sum_j \alpha_j l(I_j)$ zu betrachten. Hier ist $l(I_j)$ die Länge des Intervalls I_j und α_j ist das Infimum bzw. Supremum von f in I_j bei der Unter- bzw. Obersumme, und $\alpha_j = f(x_j)$ mit $x_j \in I_j$ bei der Riemannschen Summe. Der Grundgedanke beim Lebesgue-Integral ist ähnlich. Man nimmt aber allgemeinere Mengen I_j. Statt der Länge benötigen wir jetzt ein allgemeineres Maß die "Größe" einer Teilmenge des $\mathbb{R}^n$. Dies ist die sogenannte Maßtheorie, die wir zunächst entwickeln.

Die hiermit verbundenen Schwierigkeiten werden gut durch das *Banach-Tarski-Paradoxon* veranschaulicht: Man kann eine Kugel im $\mathbb{R}^3$ in endlich viele disjunkte Teilmengen zerlegen und diese dann zu einer Kugel doppelter Größe zusammensetzen. Formal: Es existieren paarweise disjunkte Mengen $M_1, \dots, M_n \subset \mathbb{R}^3$ und starre Bewegungen $T_1, \dots, T_n$ des $\mathbb{R}^3$ mit $\bigcup_{k=1}^n M_k = \{x \in \mathbb{R}^3 \colon \|x\|_2 \le 1\}$ und $\bigcup_{k=1}^n T_k(M_k) = \{x \in \mathbb{R}^3 \colon \|x\|_2 \le 2\}$.

Der Ausweg aus dieser Schwierigkeit ist, nicht allen Mengen ein "Maß" zuzuordnen, sondern dieses nur für gewisse Mengen zu definieren.

1 σ-Algebren

Wir bezeichnen mit $P(X)$ die Potenzmenge einer Menge X, also die Menge aller Teilmengen von X. Für $A \subseteq X$ sei $A^c := X \setminus A$ das Komplement von A (in X).

Definition 1.1. Sei X Menge und $S \subseteq P(X)$. Dann heißt S eine *σ-Algebra* (in X) falls gilt:

(i) $\emptyset \in S$.

(ii) Ist $A \in S$, so ist auch $A^c \in S$.

(iii) Ist $(A_n)_{n\in\mathbb{N}}$ eine Folge in S, so ist auch $\bigcup_{n\in\mathbb{N}} A_n \in S$.

Beispiele. 1. $S = P(X)$.

2. $S = \{\emptyset, X\}$.

3. $S = \{\emptyset, A, A^c, X\}$, mit $A \subseteq X$.

4. $S = \{A \subseteq X \colon A \text{ oder } A^c \text{ ist höchstens abzählbar}\}$.

Aus (iii) folgt (mit $A_n = \emptyset$ für $n > N$), dass

(iii′) $\bigcup_{n=1}^N A_n \in S$ falls $A_1, \dots, A_N \in S$.

Fordert man nur dies statt (iii), so nennt man S eine *Algebra*. Offensichtlich ist jede σ-Algebra eine Algebra, aber die Umkehrung gilt nicht, wie für unendliches X das Beispiel $S = \{A \subseteq X \colon A \text{ oder } A^C \text{ ist endlich}\}$ zeigt. Es genügt, (iii′) für $N = 2$ zu fordern, also $A \cup B \in S$ falls $A, B \in S$. Der allgemeine Fall ergibt sich dann per Induktion.

Wir geben einige einfache Folgerungen aus den Eigenschaften (i)–(iii) aus der Definition der σ-Algebra an:

- Aus (i) und (ii) folgt $X = \emptyset^c \in S$.
- Aus (ii) und (iii) folgt für ein Folge (A_n) in S, dass

$$\bigcap_{n\in\mathbb{N}} A_n = \left(\bigcup_{n\in\mathbb{N}} A_n^c\right)^c \in S.$$

Letzteres gilt auch wieder für endliche Schnitte – und das auch bereits in einer Algebra.

Der folgende Satz kann sehr einfach durch Überprüfung der Bedingungen aus Definition 1.1 bewiesen werden.

Satz 1.2. *Sei M eine Menge von σ-Algebren in einer Menge X. Dann ist auch*

$$\bigcap_{S\in M} S$$

eine σ-Algebra.

Dies führt zu folgender Definition. Hier und im Folgenden sei X immer eine (nichtleere) Menge.

Definition 1.3. Sei $E \subseteq P(X)$. Dann heißt

$$\sigma(E) := \bigcap_{\substack{S\supseteq E \\ S\ \sigma\text{-Algebra}}} S$$

die von E erzeugte σ-Algebra.

Offensichtlich ist $\sigma(E)$ die kleinste σ-Algebra, die E enthält.

Die oben betrachtete σ-Algebra

$$S = \{A \subseteq X \colon A \text{ oder } A^c \text{ ist höchstens abzählbar}\}$$

wird beispielsweise von den endlichen Teilmengen von X erzeugt. Ein anderer (kleinerer) Erzeuger ist die Menge der einelementigen Teilmengen von X.

Die für uns wichtigste σ-Algebra ist die Folgende.

Definition 1.4. Sei (X, d) metrischer Raum. Die von den offenen Teilmengen von X erzeugte σ-Algebra heißt *Borelsche σ-Algebra* und wird mit $\mathcal{B}(X)$ bezeichnet.

Insbesondere interessiert uns der Fall $X = \mathbb{R}^d$ mit $d \in \mathbb{N}$. Hier werden wir üblicherweise die von der Norm $\|\cdot\|_2$ erzeugte euklidische Metrik zu Grunde legen, aber dies ist irrelevant, da alle Normen auf $\mathbb{R}^d$ äquivalent sind. Wir schreiben kurz $\mathcal{B}$ oder $\mathcal{B}^d$ für $\mathcal{B}(\mathbb{R}^d)$.

Für $a = (a_1, \ldots, a_d) \in \mathbb{R}^d$ und $b = (b_1, \ldots, b_d) \in \mathbb{R}^d$ nennt man

$$(a, b] := \{(x_1, \ldots, x_d) \in \mathbb{R}^d \colon a_j < x_j \leq b_j \text{ für alle } j \in \{1, \ldots, d\}\}$$

einen *halboffenen Quader* (oder ein *halboffenes Intervall*). Analog werden $[a, b)$, der offene Quader (a, b) und der abgeschlossene Quader $[a, b]$ definiert. Auch $a = \infty = (\infty, \ldots, \infty)$ bzw. $b = \infty$ ist zugelassen.

Satz 1.5. *Für $d \in \mathbb{N}$ wird $\mathcal{B}^d$ von der Menge der halboffenen Quader $(a, b]$ mit $a, b \in \mathbb{Q}^d$ erzeugt.*

Beweis. Sei $I_{\mathbb{Q}}$ die Menge aller halboffenen Quader $(a, b]$ mit $a, b \in \mathbb{Q}^d$. Dann ist $I_{\mathbb{Q}}$ abzählbar, da $\mathbb{Q}$ abzählbar ist.

Sei nun $U \subseteq \mathbb{R}^d$ offen. Dann ist

$$U = \bigcup_{\substack{A \subseteq U \\ A \in I_{\mathbb{Q}}}} A$$

und damit $U \in \sigma(I_{\mathbb{Q}})$. Es folgt $\mathcal{B}^d \subseteq \sigma(I_{\mathbb{Q}})$.

Umgekehrt ist mit $b^n = (b_1 + 1/n, \ldots, b_d + 1/n)$

$$(a, b] = \bigcap_{n \in \mathbb{N}} (a, b^n) \in \mathcal{B}^d$$

und damit $\sigma(I_{\mathbb{Q}}) \subseteq \mathcal{B}^d$. □

Ebenso wird $\mathcal{B}^d$ von der Menge der halboffenen Quader $(a, b]$ mit $a, b \in \mathbb{R}^d$ erzeugt. Ganz ähnlich sieht man, dass $\mathcal{B}^d$ auch von den Quadern (a, b), $[a, b)$, $[a, b]$, (a, ∞), $[a, \infty)$, $(-\infty, a)$ und $(-\infty, a]$ erzeugt wird, sowohl für $a, b \in \mathbb{R}^d$ wie $a, b \in \mathbb{Q}^d$.

Um nachzuweisen, dass ein Mengensystem eine σ-Algebra ist, werden wir manchmal die folgende Definition benutzen.

Definition 1.6. Sei $D \subseteq P(X)$. Dann heißt D ein *Dynkin-System* (in X) falls gilt:

(i) $\emptyset \in D$.

(ii) Ist $A \in D$, so ist auch $A^c \in D$.

(iii) Ist $(A_n)_{n \in \mathbb{N}}$ eine Folge paarweiser disjunkter Mengen in D, so ist auch $\bigcup_{n \in \mathbb{N}} A_n \in D$.

Der Unterschied zur σ-Algebra ist, dass in (iii) nur Folgen paarweiser disjunkter Mengen betrachtet werden.

Offensichtlich ist also jede σ-Algebra ein Dynkin-System. Die Umkehrung gilt nicht. Man wähle etwa X als eine endliche Menge mit einer geraden Anzahl von Elementen und D als Menge aller Teilmengen mit einer geraden Anzahl von Elementen.

Dieses Beispiel zeigt auch, dass im Gegensatz zur σ-Algebra für $A, B \in D$ nicht $A \cap B \in D$ gelten muss. Ein Mengensystem, welches mit zwei Mengen auch jeweils ihre Schnittmenge enthält, heißt *schnittstabil.* Induktion zeigt, dass für ein schnittstabiles Mengensystem D auch endliche Schnitte von Mengen in D in D enthalten sind.

Lemma 1.7. *Sei D Dynkin-System, $A, B \in D$ und $B \subseteq A$. Dann gilt $A \backslash B \in D$.*

Beweis. Die Mengen A^c und B sind disjunkt und es gilt $A \setminus B = (A^c \cup B)^c$. □

Satz 1.8. *Ein schnittstabiles Dynkin-System ist eine σ-Algebra.*

Beweis. Sei D schnittstabiles Dynkin-System und $(A_n)_{n \in \mathbb{N}}$ Folge in D. Zu zeigen ist, dass $\bigcup_{j=1}^{\infty} A_j \in D$. Durch

$$B_n := A_n \setminus \bigcup_{j=1}^{n-1} A_j = A_n \cap \bigcap_{j=1}^{n-1} A_j^c$$

wird dann eine Folge disjunkter Teilmengen in S definiert. Es folgt

$$\bigcup_{j=1}^{\infty} A_j = \bigcup_{j=1}^{\infty} B_j \in D.$$

□

Analog zu Definition 1.3 betrachtet man für $E \subseteq P(X)$

$$\delta(E) := \bigcap_{\substack{D \supseteq E \\ D \text{ Dynkin-System}}} D.$$

Analog zu Satz 1.2 sieht man, dass $\delta(E)$ tatsächlich ein Dynkin-System ist. Man nennt es das von E erzeugte Dynkin-System.

Satz 1.9. *Sei $E \subseteq P(X)$ schnittstabil. Dann gilt $\delta(E) = \sigma(E)$.*

Beweis. Da jede σ-Algebra ein Dynkin-System ist, gilt $\delta(E) \subseteq \sigma(E)$.

Umgekehrt folgt $\sigma(E) \subseteq \delta(E)$, wenn wir zeigen, dass $\delta(E)$ eine σ-Algebra ist. Nach Satz 1.8 ist dies der Fall, wenn wir zeigen, dass $\delta(E)$ schnittstabil ist. Sei dazu $A \in \delta(E)$. Zu zeigen ist, dass für $B \in \delta(E)$ auch $A \cap B \in \delta(E)$ gilt. Setzt man

$$D_A := \{B \subseteq X \colon A \cap B \in \delta(E)\},$$

so ist also $\delta(E) \subseteq D_A$ zu zeigen.

Nun rechnet man aber leicht nach, dass D_A Dynkin-System ist. Damit reicht es, $E \subseteq D_A$ zu zeigen. Sei dazu $C \in E$. Dann gilt $C \cap B \in E$ für alle $B \in E$, also $E \subseteq D_C$. Weil auch D_C Dynkin-System ist, folgt $A \in \delta(E) \subseteq D_C$. Dies besagt, dass $C \cap A \in \delta(E)$, also $C \in D_A$. □

2 Maße und äußere Maße

Ist S eine σ-Algebra in einer Menge X, so nennt man das Paar (X, S) auch *Messraum* und die Elemente von S *messbar*.

Definition 2.1. Sei (X, S) Messraum. Eine Abbildung $\mu \colon S \to [0, \infty]$ heißt *Maß* (auf S), falls gilt:

(i) $\mu(\emptyset) = 0$.

(ii) Ist $(A_n)_{n \in \mathbb{N}}$ eine Folge paarweise disjunkter Mengen in S, so ist

$$\mu\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n \in \mathbb{N}} \mu(A_n).$$

Man nennt (X, S, μ) dann *Maßraum*.

Man beachte, dass der Fall $\mu(A) = \infty$ zugelassen ist. Man rechnet mit dem Wert ∞ mit den offensichtlichen Regeln, etwa $\infty + x = \infty$ für $x \in [0, \infty]$.

Ein einfaches Beispiel erhält man durch das sogenannte Zählmaß $\mathrm{card} \colon X \to [0, \infty]$: Man definiert $\mathrm{card}(A)$ als Anzahl der Elemente von A.

Die Aussage (ii) gilt analog auch wieder für endliche Vereinigungen, d.h., es gilt

$$\mu\left(\bigcup_{n=1}^{N} A_n\right) = \sum_{n=1}^{N} \mu(A_n).$$

falls $A_1, \ldots, A_N \in S$ paarweise disjunkt. (Man wähle $A_n = \emptyset$ für $n > N$.) Fordert man nur dieses, so erhält man den Begriff des *Inhalts*; siehe Definition 3.4.

Wir stellen einige Eigenschaften von Maßen zusammen.

Lemma 2.2. *Sei (X, S, μ) Maßraum. Dann gilt*

(i) *Sind $A, B \in S$ mit $A \subseteq B$, so ist $\mu(A) \leq \mu(B)$. Genauer gilt $\mu(B) = \mu(A) + \mu(B \setminus A)$.*

(ii) *Sind $A, B \in S$, so gilt $\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B)$.*

(iii) *Ist $(A_n)_{n \in \mathbb{N}}$ eine Folge in S, so ist*

$$\mu\left(\bigcup_{n \in \mathbb{N}} A_n\right) \leq \sum_{n \in \mathbb{N}} \mu(A_n).$$

Beweis. (i) Dies folgt mit $B = A \cup (B \setminus A)$.

(ii) Nach (i) ist $\mu(A \cup B) = \mu(A) + \mu((A \cup B) \setminus A) = \mu(A) + \mu(B \setminus (A \cap B))$ und damit, wiederum mit (i),

$$\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B \setminus (A \cap B)) + \mu(A \cap B) = \mu(A) + \mu(B).$$

(iii) Durch

$$B_n := A_n \setminus \bigcup_{j=1}^{n-1} A_j$$

wird wie im Beweis von Satz 1.8 eine Folge disjunkter Teilmengen in S definiert, und es folgt

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \mu\left(\bigcup_{n\in\mathbb{N}} B_n\right) = \sum_{n\in\mathbb{N}} \mu(B_n) \leq \sum_{n\in\mathbb{N}} \mu(A_n)$$

nach (i). □

Den ersten Teil von (i), also die Implikation "$A \subseteq B \Rightarrow \mu(A) \leq \mu(B)$", nennt man auch Monotonie.

Eine Gleichung wie $\mu(B) = \mu(A) + \mu(B \setminus A)$ in (i) ist auch dann sinnvoll, wenn einige der Werte ∞ sind. Die Umformung $\mu(B \setminus A) = \mu(B) - \mu(A)$ ist aber nur dann möglich, wenn $\mu(A) < \infty$ gilt, da andernfalls der undefinierte Ausdruck $\infty - \infty$ entsteht.

Für Mengen $A, A_1, A_2, \ldots$ schreiben wir $A_n \uparrow A$ falls $A_1 \subseteq A_2 \subseteq A_3 \subseteq \ldots$ und $\bigcup_{n\in\mathbb{N}} A_n = A$. Analog schreiben wir $A_n \downarrow A$ falls $A_1 \supseteq A_2 \supseteq A_3 \supseteq \ldots$ und $\bigcap_{n\in\mathbb{N}} A_n = A$.

Lemma 2.3. *Sei (X, S, μ) Maßraum und $A, A_1, A_2, \dots \in S$.*

(i) *Aus $A_n \uparrow A$ folgt $\mu(A_n) \to \mu(A)$.*

(ii) *Gilt $\mu(A_N) < \infty$ für ein $N \in \mathbb{N}$, so folgt aus $A_n \downarrow A$, dass $\mu(A_n) \to \mu(A)$.*

Beweis. (i) Wie im vorigen Beweis wähle man $B_n := A_n \setminus \bigcup_{j=1}^{n-1} A_j = A_n \setminus A_{n-1}$, mit $A_0 = \emptyset$. Es folgt

$$\mu(A_n) = \mu\left(\bigcup_{j=1}^{n} B_j\right) = \sum_{j=1}^{n} \mu(B_j) \to \sum_{j=1}^{\infty} \mu(B_j) = \mu\left(\bigcup_{j=1}^{\infty} B_j\right) = \mu(A).$$

(ii) Mit $B_n = A_N \setminus A_n$ gilt $B_n \uparrow B := A_N \setminus A$ und damit $\mu(B_n) \to \mu(B)$. Wegen $\mu(A_N) < \infty$ gilt $\mu(B_n) = \mu(A_N \setminus A_n) = \mu(A_N) - \mu(A_n)$ und analog $\mu(B) = \mu(A_N) - \mu(A)$. Die Behauptung folgt. □

Gilt $\mu(A_N) = \infty$ für alle $N \in \mathbb{N}$, so muss (ii) nicht gelten. Ein Beispiel ist durch das Zählmaß card auf der Menge $\mathbb{N}$ mit σ-Algebra $P(\mathbb{N})$ und die Mengen $A_n = \{k \in \mathbb{N} \colon k \geq n\}$ gegeben. Hier gilt $A_n \downarrow \emptyset$, aber $\mathrm{card}(A_n) = \infty \not\to 0 = \mathrm{card}(\emptyset)$.

Definition 2.4. Sei (X, S, μ) Maßraum. Eine Teilmenge N von X heißt *Nullmenge* (oder *μ-Nullmenge*), falls $\mu(N) = 0$ gilt.

Das Maß μ bzw. der Maßraum (X, S, μ) heißt *vollständig*, falls jede Teilmenge einer Nullmenge messbar (also in S) ist.

Aus der Definition des Maßes folgt sofort, dass eine abzählbare Vereinigung von Nullmengen wieder Nullmenge ist.

Ist N Nullmenge in einem vollständigen Maßraum (X, S, μ) und $M \subseteq N$, so folgt also $M \in S$ und nach Monotonie dann $\mu(M) = 0$.

Sind A und C messbar mit $A \subseteq C$ und $\mu(A) = \mu(C) < \infty$ und ist $B \subseteq X$ mit $A \subseteq B \subseteq C$, so folgt, dass auch B messbar ist, mit $\mu(B) = \mu(A) = \mu(C)$. Denn nach Lemma 2.2, (i), ist $C \setminus A$ Nullmenge, also wegen der Vollständigkeit auch $B \setminus A$, und damit ist $B = A \cup (B \setminus A)$ messbar, mit $\mu(B) = \mu(A) + \mu(B \setminus A) = \mu(A)$.

Man kann einen Maßraum (X, S, μ) *vervollständigen*, indem man zu S alle Mengen B hinzufügt, für die $A, C \in S$ mit $A \subseteq B \subseteq C$ und $\mu(C \setminus A) = 0$ existieren, und für diese dann $\mu(B) = \mu(A)$ setzt.

Wir werden Maße über sogenannte äußere Maße konstruieren.

Definition 2.5. Eine Abbildung $\mu^* \colon P(X) \to [0, \infty]$ heißt *äußeres Maß*, falls gilt:

(i) $\mu^*(\emptyset) = 0$.

(ii) Sind $A, B \subseteq X$ mit $A \subseteq B$, so ist $\mu^*(A) \leq \mu^*(B)$.

(iii) Ist $(A_n)_{n \in \mathbb{N}}$ eine Folge in $P(X)$, so ist

$$\mu^*\left(\bigcup_{n \in \mathbb{N}} A_n\right) \leq \sum_{n \in \mathbb{N}} \mu^*(A_n).$$

Nach Definition 2.1 und Lemma 2.2 gelten die Bedingungen (i)–(iii) insbesondere für Maße. Allerdings ist ein Maß im Allgemeinen nicht auf der gesamten Potenzmenge definiert, sondern nur auf einer σ-Algebra.

Das wichtigste Beispiel ist das folgende.

Definition 2.6. Für einen Quader $(a, b]$ mit $a = (a_1, \ldots, a_d), b = (b_1, \ldots, b_d) \in \mathbb{R}^d$ heißt

$$\operatorname{vol}(a, b] = \operatorname{vol}_d(a, b] = \prod_{j=1}^{d} \max\{b_j - a_j, 0\}$$

das (d-dimensionale) *Volumen* des Quaders. Sei E die Menge dieser Quader. Dann heißt die durch

$$\lambda^*(A) = \inf\left\{\sum_{n=1}^{\infty} \operatorname{vol}(Q_n) \colon (Q_n) \text{ Folge in } E \text{ mit } A \subseteq \bigcup_{n=1}^{\infty} Q_n\right\}$$

definierte Abbildung $\lambda^* \colon P(\mathbb{R}^d) \to [0, \infty]$ *äußeres Lebesgue-Maß.*

Dies ist tatsächlich ein äußeres Maß. Denn die Eigenschaften (i) und (ii) aus Definition 2.5 folgen unmittelbar. Eigenschaft (iii) ist trivial, falls $\lambda^*(A_n) = \infty$ für ein $n \in \mathbb{N}$. Andernfalls sei $\varepsilon > 0$. Für jedes $n \in \mathbb{N}$ existiert dann eine Folge $(Q_{n,k})_{k \in \mathbb{N}}$ mit

$$A_n \subseteq \bigcup_{k=1}^{\infty} Q_{n,k} \quad \text{und} \quad \sum_{k=1}^{\infty} \operatorname{vol}(Q_{n,k}) < \lambda^*(A_n) + \frac{\varepsilon}{2^n}.$$

Wegen

$$A \subseteq \bigcup_{n=1}^{\infty} A_n \subseteq \bigcup_{n=1}^{\infty} \bigcup_{k=1}^{\infty} Q_{n,k}$$

und

$$\sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \operatorname{vol}(Q_{n,k}) \leq \sum_{n=1}^{\infty} \left(\lambda^*(A_n) + \frac{\varepsilon}{2^n}\right) = \varepsilon + \sum_{n=1}^{\infty} \lambda^*(A_n)$$

folgt $\lambda^*(A) \leq \varepsilon + \sum_{n=1}^{\infty} \lambda^*(A_n)$ und damit (iii).

Bemerkung 2.7. Analog kann man das Volumen auch für offene oder abgeschlossene Quader definieren. Ähnliche Argumente wie oben zeigen, dass dies das gleiche äußere Maß liefert. Ist beispielsweise $A \subseteq \bigcup_{n=1}^{\infty} Q_n$ mit halboffenen Quadern Q_n, so existieren zu gegebenem $\varepsilon > 0$ offene Quader Q_n' mit $Q_n \subset Q_n'$ und $\operatorname{vol}(Q_n') < \operatorname{vol}(Q_n) + \varepsilon/2^n$, also

$$\sum_{n=1}^{\infty} \operatorname{vol}(Q_n') < \sum_{n=1}^{\infty} \left(\operatorname{vol}(Q_n) + \frac{\varepsilon}{2^n} \right) = \sum_{n=1}^{\infty} \operatorname{vol}(Q_n) + \varepsilon.$$

Hieraus folgt, dass man das gleiche äußere Maß erhält, wenn man nur offene Quader benutzt.

Der Vorteil der halboffenen Quader ist, dass sie eine disjunkte Überdeckung von $\mathbb{R}^d$ (oder gewissen Teilmengen davon) liefern.

Das folgende Lemma ist sehr naheliegend.

Lemma 2.8. *Für jeden Quader $Q = (a, b]$ gilt $\lambda^*(Q) = \operatorname{vol} Q$.*

Beweisskizze. Die Ungleichung folgt $\lambda^*(Q) \leq \operatorname{vol} Q$ sofort aus der Definition mit $Q_1 = Q$ und $Q_n = \emptyset$ für $n \geq 2$. Die umgekehrte Richtung ist etwas aufwändiger. Zunächst zeigt man ein entsprechendes Ergebnis für endliche Überdeckungen mit Quadern, d.h.,

$$\operatorname{vol}(Q) \leq \sum_{n=1}^{N} \operatorname{vol}(Q_n) \text{ für Quader } Q_1, \ldots, Q_N \text{ mit } Q \subseteq \bigcup_{n=1}^{N} Q_n.$$

Sei nun $Q \subseteq \bigcup_{n=1}^{\infty} Q_n$ mit Quadern Q_n. Zu gegebenem $\varepsilon > 0$ findet man wie in Bemerkung 2.7 offene Quader Q_n' mit $Q_n \subset Q_n'$ und $\operatorname{vol}(Q_n') < \operatorname{vol}(Q_n) + \varepsilon/2^n$, so dass $[a, b] \subseteq \bigcup_{n=1}^{\infty} Q_n'$. Da $[a, b]$ kompakt ist, existiert $N \in \mathbb{N}$ mit $[a, b] \subseteq \bigcup_{n=1}^{N} Q_n'$. Damit kann der Fall abzählbarer Überdeckungen auf den Fall endlicher Überdeckungen zurückgeführt werden. □

Die obige Konstruktion lässt sich wesentlich verallgemeinern. In der Tat ergeben die Argumente folgendes Resultat.

Lemma 2.9. *Sei $E \subseteq P(X)$ mit $\emptyset \in E$ und $\nu \colon E \to [0, \infty]$ mit $\nu(\emptyset) = 0$. Dann ist durch*

$$\mu^*(A) = \inf \left\{ \sum_{n=1}^{\infty} \nu(Q_n) \colon (Q_n) \textit{ Folge in } E \textit{ mit } A \subseteq \bigcup_{n=1}^{\infty} Q_n \right\}$$

ein äußeres Maß $\mu^* \colon P(X) \to [0, \infty]$ definiert.

3 Konstruktion von Maßen aus äußeren Maßen

Die Idee ist nun, durch Einschränkung eines äußeren Maßes auf eine geeignete σ-Algebra ein Maß zu gewinnen.

Definition 3.1. Sei μ^* äußeres Maß auf $P(X)$ und $A \subseteq X$. Dann heißt A μ^*-*messbar*, falls

$$\mu^*(T) = \mu^*(T \cap A) + \mu^*(T \cap A^c) \quad \text{für alle } T \subseteq X.$$

Die Menge der μ^*-messbaren Teilmengen von X wird mit $S(\mu^*)$ bezeichnet.

Bemerkung 3.2. Nach Definition des äußeren Maßes gilt immer

$$\mu^*(T) = \mu^*((T \cap A) \cup (T \cap A^c)) \leq \mu^*(T \cap A) + \mu^*(T \cap A^c).$$

Was in Definition 3.1 gefordert wird, ist also

$$\mu^*(T) \geq \mu^*(T \cap A) + \mu^*(T \cap A^c).$$

Satz 3.3. (Maßerweiterungssatz von Carathéodory) *Sei μ^* äußeres Maß auf $P(X)$. Dann ist $S(\mu^*)$ eine σ-Algebra und die Einschränkung von μ^* auf $S(\mu^*)$ ist ein vollständiges Maß.*

Beweis. Zunächst gilt offensichtlich $\emptyset \in S(\mu^*)$ und mit A ist auch A^c in $S(\mu^*)$, da die Bedingung aus Definition 3.1 symmetrisch in A und A^c ist.

Wir zeigen zuerst, dass $S(\mu^*)$ Algebra ist, dann dass $S(\mu^*)$ Dynkin-System ist, und benutzen schließlich Satz 1.8 um zu zeigen, dass $S(\mu^*)$ σ-Algebra ist.

Seien $A, B \in S(\mu^*)$ und $T \subseteq X$. Dann gilt

$$\begin{aligned}
&\mu^*(T \cap (A \cup B)) + \mu^*(T \cap (A \cup B)^c) \\
&= \mu^*((T \cap A) \cup (T \cap B)) + \mu^*((T \cap A^c \cap B^c) \\
&= \mu^*((T \cap A) \cup (T \cap A^c \cap B)) + \mu^*((T \cap A^c \cap B^c) \\
&\leq \mu^*(T \cap A) + \mu^*(T \cap A^c \cap B) + \mu^*(T \cap A^c \cap B^c).
\end{aligned}$$

Nun ist

$$\mu^*(T \cap A^c \cap B) + \mu^*(T \cap A^c \cap B^c) = \mu^*(T \cap A^c),$$

da B μ^*-messbar, und

$$\mu^*(T \cap A) + \mu^*(T \cap A^c) = \mu^*(T)$$

da A μ^*-messbar. Es folgt

$$\mu^*(T \cap (A \cup B)) + \mu^*(T \cap (A \cup B)^c) \leq \mu^*(T),$$

also nach Bemerkung 3.2, dass $A \cup B \in S(\mu^*)$. Damit ist $S(\mu^*)$ Algebra.

Wegen $A \cap B = (A^c \cup B^c)^c$ ist $S(\mu^*)$ auch schnittstabil. Um zu zeigen, dass $S(\mu^*)$ Dynkin-System ist, sei (A_n) Folge paarweise disjunkter Mengen in $S(\mu^*)$. Wir setzen

$$B_n = \bigcup_{j=1}^{n} A_j \quad \text{und} \quad B = \bigcup_{j=1}^{\infty} A_j = \bigcup_{j=1}^{\infty} B_j.$$

Da $S(\mu^*)$ Algebra ist, gilt $B_n \in S(\mu^*)$ für alle $n \in \mathbb{N}$. Zu zeigen ist, dass $B \in S(\mu^*)$.

Wir zeigen zunächst durch Induktion, dass für $T \subseteq X$

$$\mu^*(T \cap B_n) = \sum_{j=1}^{n} \mu^*(T \cap A_j)$$

gilt. Dies ist klar für $n = 1$. Gilt dies aber für ein $n \in \mathbb{N}$, so folgt

$$\begin{aligned}\mu^*(T \cap B_{n+1}) &= \mu^*(T \cap B_{n+1} \cap B_n) + \mu^*(T \cap B_{n+1} \cap B_n^c) \\ &= \mu^*(T \cap B_n) + \mu^*(T \cap A_{n+1}) \\ &= \sum_{j=1}^{n+1} \mu^*(T \cap A_j)\,,\end{aligned}$$

womit die Aussage auch mit $n + 1$ statt n gilt. Für alle $n \in \mathbb{N}$ folgt damit

$$\begin{aligned}\mu^*(T) &= \mu^*(T \cap B_n) + \mu^*(T \cap B_n^c) \\ &\geq \mu^*(T \cap B_n) + \mu^*(T \cap B^c) \\ &= \sum_{j=1}^{n} \mu^*(T \cap A_j) + \mu^*(T \cap B^c).\end{aligned}$$

Es folgt, dass

$$\mu^*(T) \geq \sum_{j=1}^{\infty} \mu^*(T \cap A_j) + \mu^*(T \cap B^c) \geq \mu^*\left(\bigcup_{j=1}^{\infty} (T \cap A_j)\right) + \mu^*(T \cap B^c),$$

also

$$\mu^*(T) \geq \mu^*(T \cap B) + \mu^*(T \cap B^c).$$

Damit ist $B \in S(\mu^*)$, also $S(\mu^*)$ Dynkin-System und nach Satz 1.8 auch σ-Algebra.

Für $T = B = \bigcup_{j=1}^{\infty} A_j$ folgt aus obiger Rechnung insbesondere, dass

$$\mu^*\left(\bigcup_{j=1}^{\infty} A_j\right) \geq \sum_{j=1}^{\infty} \mu^*(A_j)\,.$$

Da auch

$$\mu^*\left(\bigcup_{j=1}^{\infty} A_j\right) \leq \sum_{j=1}^{\infty} \mu^*(A_j)$$

nach Definition des äußeren Maßes, gilt hier Gleichheit. Damit ist die Einschränkung von μ^* auf $S(\mu^*)$ ein Maß.

Es verbleibt zu zeigen, dass $\mu^*|_{S(\mu^*)}$ vollständig ist. Sei dazu N Nullmenge und $M \subseteq N$. Für $T \subseteq X$ ist dann $T \cap M \subseteq N$, also $\mu^*(T \cap M) \leq \mu^*(N) = 0$ und damit

$$\mu^*(T) \geq \mu^*(T \cap M^c) = \mu^*(T \cap M) + \mu^*(T \cap M^c).$$

Nach Bemerkung 3.2 ist M damit messbar. □

Als nächstes wollen wir zeigen, dass für das in Definition 2.6 konstruierte äußere Lebesgue-Maß λ^* jeder Quader und damit nach Satz 1.5 auch jede Borelmenge in $S(\lambda^*)$ liegt. Allgemeiner untersuchen wir die in Lemma 2.9 betrachtete Situation.

Definition 3.4. Sei $H \subseteq P(X)$.

(i) H heißt *Halbring*, falls H schnittstabil ist, $\emptyset \in H$ gilt, und für alle $A, B \in H$ paarweise disjunkte Mengen $C_1, \ldots, C_m \in H$ existieren, so dass

$$A \setminus B = \bigcup_{j=1}^{m} C_j.$$

(ii) Ist H Halbring und $\nu\colon H \to [0, \infty]$, so heißt H *Inhalt*, falls $\nu(\emptyset) = 0$ und und für alle paarweise disjunkten $A_1, \ldots, A_m \in H$ mit $\bigcup_{j=1}^{m} A_j \in H$ gilt, dass

$$\nu\left(\bigcup_{j=1}^{m} A_j\right) = \sum_{j=1}^{m} \nu(A_j).$$

Es ist leicht zu sehen, dass die Menge der halboffenen Quader $(a, b]$ in $\mathbb{R}^d$ ein Halbring ist und das Volumen aus Definition 2.6 ein Inhalt ist. Beispielsweise ist für $(a, b], (c, d] \subset \mathbb{R}^2$ mit $a_1 < c_1 < b_1 < d_1$ und $a_2 < c_2 < b_2 < d_2$

$$(a, b] \setminus (c, d] = (a, c] \cup ((c_1, a_2), (b_1, c_2)] \cup ((a_1, c_2), (c_1, b_2)];$$

vgl. Abbildung 1.

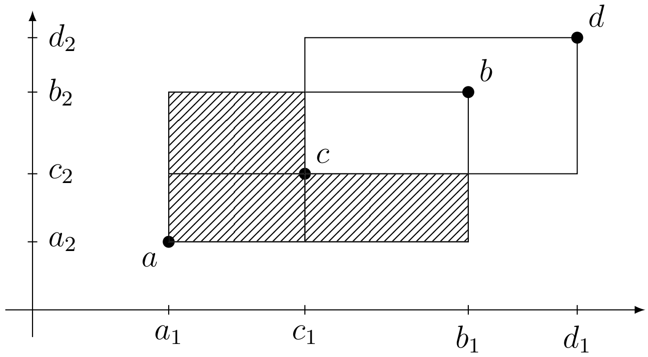

Abbildung 1: $(a, b] \setminus (c, d]$ besteht aus 3 Rechtecken.

Ein Inhalt ist auch monoton, denn für $B \subseteq A$ folgt mit $C_1, \ldots, C_m$ wie in obiger Definition, dass

$$\nu(A) = \nu(B \cup (A \setminus B)) = \nu\left(B \cup \bigcup_{j=1}^{m} C_j\right) = \nu(B) + \sum_{j=1}^{m} \nu(C_j) \geq \nu(B).$$

Satz 3.5. (Maßfortsetzungssatz) *Sei H Halbring, $\nu\colon H \to [0, \infty]$ Inhalt und μ^* das mit $E = H$ gemäß Lemma 2.9 gebildete äußere Maß.*

(i) *Es gilt* $H \subseteq S(\mu^*)$.

(ii) *Gilt*

$$\nu\left(\bigcup_{j=1}^{\infty} A_j\right) \leq \sum_{j=1}^{\infty} \nu(A_j)$$

für jede Folge (A_j) *in* H *mit* $\bigcup_{j=1}^{\infty} A_j \in H$, *so ist* $\mu^*|_H = \nu$.

Aus (i) folgt unmittelbar, dass auch $\sigma(H) \subseteq S(\mu^*)$.

Für das äußere Lebesgue-Maß wählen wir, wie bereits gesagt, H als Menge der halboffenen Quader $(a, b]$ und ν als Volumen. Lemma 2.8 besagt, dass dann auch die Voraussetzung von (ii) gilt, denn mit $A = \bigcup_{j=1}^{\infty} A_j$ ist die linke Seite der Ungleichung $\mathrm{vol}(A)$ und und die rechte ist nach Definition des äußeren Maßes größer als oder gleich $\lambda^*(A)$.

Beweis von Satz 3.5. (i) Sei $A \in H$ und $T \subseteq X$. Zu zeigen ist die Ungleichung

$$\mu^*(T) \geq \mu^*(T \cap A) + \mu^*(T \cap A^c)$$

aus Bemerkung 3.2. Dies ist trivial, falls $\mu^*(T) = \infty$. Sei also $\mu^*(T) < \infty$. Dann existiert zu $\varepsilon > 0$ eine Folge (B_n) in H mit

$$T \subseteq \bigcup_{n=1}^{\infty} B_n \quad \text{und} \quad \sum_{n=1}^{\infty} \nu(B_n) < \mu^*(T) + \varepsilon.$$

Gemäß der Definition des Halbrings ist

$$B_n \setminus A = \bigcup_{k=1}^{m_n} C_{k,n}$$

mit paarweise disjunkten Mengen $C_{k,n} \in H$. Damit können wir B_n als disjunkte Vereinigung

$$B_n = (B_n \cap A) \cup (B_n \setminus A) = (B_n \cap A) \cup \bigcup_{k=1}^{m_n} C_{k,n}$$

schreiben. Es folgt

$$T \cap A \subseteq \bigcup_{n=1}^{\infty} B_n \cap A$$

und

$$T \cap A^c \subseteq \bigcup_{n=1}^{\infty} B_n \cap A^c = \bigcup_{n=1}^{\infty} B_n \setminus A = \bigcup_{n=1}^{\infty} \bigcup_{k=1}^{m_n} C_{k,n}.$$

Es folgt, dass

$$\begin{aligned}
\mu^*(T \cap A) + \mu^*(T \cap A^c) &\leq \sum_{n=1}^{\infty} \nu(B_n \cap A) + \sum_{n=1}^{\infty} \sum_{k=1}^{m_n} \nu(C_{k,n}) \\
&= \sum_{n=1}^{\infty} \left(\nu(B_n \cap A) + \sum_{k=1}^{m_n} \nu(C_{k,n}) \right) \\
&= \sum_{n=1}^{\infty} \nu(B_n) \\
&< \mu^*(T) + \varepsilon
\end{aligned}$$

und damit die Behauptung.

(ii) Sei $A \in H$. Mit $A_1 = A$ und $A_n = \emptyset$ für $n \geq 2$ in Lemma 2.3 folgt $\mu^*(A) \leq \nu(A)$. Umgekehrt existiert für $\mu^*(A) < \infty$ zu $\varepsilon > 0$ eine Folge (B_n) in H mit

$$A \subseteq \bigcup_{n=1}^{\infty} B_n \quad \text{und} \quad \sum_{n=1}^{\infty} \nu(B_n) < \mu^*(A) + \varepsilon.$$

Wegen

$$A = \bigcup_{n=1}^{\infty} (B_n \cap A)$$

und $B_n \cap A \in H$ für alle n folgt

$$\nu(A) \leq \sum_{n=1}^{\infty} \nu(B_n \cap A) < \mu^*(A) + \varepsilon$$

und damit die Behauptung. □

Das folgende Beispiel zeigt, dass die Fortsetzung eines Inhalts zu einem Maß nicht eindeutig sein muss.

Beispiel 3.6. Sei E die Menge der halboffenen Quader $(a, b]$ in $\mathbb{R}^d$ und sei $\nu \colon E \to [0, \infty)$ definiert durch $\nu(Q) = \infty$ falls $Q \neq \emptyset$ und $\nu(\emptyset) = 0$.

Dann sind sowohl durch das Zählmaß $\operatorname{card} \colon P(\mathbb{R}^d) \to [0, \infty]$ wie auch durch $\mu(A) = \infty$ falls $A \neq \emptyset$ Fortsetzungen von ν auf $P(\mathbb{R}^d)$ definiert.

Die Eindeutigkeit der Fortsetzung kann aber durch eine Zusatzvoraussetzung erreicht werden.

Definition 3.7. Sei $E \subseteq P(X)$ und $\nu \colon E \to [0, \infty]$. Dann heißt ν σ-endlich, falls eine Folge (A_n) in E existiert, so dass $\nu(A_n) < \infty$ für alle $n \in \mathbb{N}$ und $X = \bigcup_{n=1}^{\infty} A_n$.

Für das Volumen ist diese Bedingung offensichtlich erfüllt, etwa mit $A_n = (-n, n]$. Die Eindeutigkeit der Fortsetzung von des Volumens auf die Borelmengen erhält man aus folgendem Satz.

Satz 3.8. (Maßeindeutigkeitssatz) *Seien (X, S) Messraum und $\mu_1, \mu_2 \colon S \to [0, \infty]$ Maße. Sei E schnittstabiler Erzeuger von S und es gelte $\mu_1|_E = \mu_2|_E$. Ist $\mu_1|_E$ (und damit auch $\mu_2|_E$) σ-endlich, so gilt $\mu_1 = \mu_2$.*

Wir verzichten hier auf den Beweis; siehe z.B. die Bücher von Elstrodt, Satz II.5.6, oder Werner, Satz IV.3.10.

4 Das Lebesgue-Maß

Wir fassen die Ergebnisse für das in Definition 2.6 definierte äußere Lebesgue-Maß $\lambda^* \colon P(\mathbb{R}^d) \to [0, \infty]$ zusammen. Nach Lemma 2.8 gilt $\lambda^*(Q) = \operatorname{vol}(Q)$ für jeden (halboffenen) Quader. Nach dem Maßerweiterungssatz 3.3 ist $S(\lambda^*)$ eine σ-Algebra und $\lambda^*|_{S(\lambda^*)}$ ein vollständiges Maß. Nach dem Maßfortsetzungssatz 3.5 enthält $S(\lambda^*)$ alle (halboffenen) Quader. Nach Satz 1.5 enthält damit $S(\lambda^*)$ auch alle Borelmengen, d.h., es gilt $\mathcal{B} \subseteq S(\lambda^*)$.

Definition 4.1. $\mathcal{L} := S(\lambda^*)$ heißt σ-Algebra der *Lebesgue-messbaren* Mengen, $\lambda := \lambda^*|_{\mathcal{L}}$ heißt *Lebesgue-Maß* und $\lambda|_{\mathcal{B}}$ heißt *Lebesgue-Borel-Maß*.

Statt $\mathcal{L}$ schreiben wir auch $\mathcal{L}^d$ und statt λ auch λ^d.

Nach dem Maßeindeutigkeitssatz 3.8 ist das Lebesgue-Borel-Maß das einzige Maß auf $\mathcal{B}$, welches auf (halboffenen) Quadern mit dem Volumen übereinstimmt. Tatsächlich ist das Lebesgue-Maß auch das einzige Maß auf $\mathcal{L}$, das diese Eigenschaft hat.

Satz 4.2. *Sei $M \in \mathcal{L}^d$ und $\varepsilon > 0$. Dann existieren $U \subseteq \mathbb{R}^d$ offen und $A \subseteq \mathbb{R}^d$ abgeschlossen mit $A \subseteq M \subseteq U$, so dass $\lambda(U \setminus M) < \varepsilon$ und $\lambda(M \setminus A) < \varepsilon$.*

Beweis. Sei zunächst $\lambda(M) < \infty$. Nach Bemerkung 2.7 kann man in der Definition des äußeren Maßes auch offene Quader nehmen. Es folgt, dass zu gegebenem $\varepsilon > 0$ eine Folge (Q_n) offener Quader mit

$$M \subseteq U := \bigcup_{n=1}^{\infty} Q_n \quad \text{und} \quad \sum_{n=1}^{\infty} \operatorname{vol}(Q_n) < \lambda^*(M) + \varepsilon = \lambda(M) + \varepsilon.$$

Es ist U offen und

$$\lambda(U \setminus M) = \lambda(U) - \lambda(M) \leq \sum_{n=1}^{\infty} \operatorname{vol}(Q_n) - \lambda(M) < \varepsilon.$$

Ist aber $\lambda(M) = \infty$, so wendet man für $n \in \mathbb{N}$ das bereits Bewiesene auf $M_n = M \cap [-n, n]$ an und erhält eine offene Menge U_n mit $M_n \subseteq U_n$ und $\lambda(U_n \setminus M_n) < \varepsilon/2^n$. Jetzt leistet $U = \bigcup_{n=1}^{\infty} U_n$ das Verlangte.

Die abgeschlossene Menge A erhält man, indem man das Bewiesene auf M^c anwendet. Dies liefert eine offene Menge V mit $M^c \subseteq V$ und $\lambda(V \setminus M^c) < \varepsilon$. Mit $A = V^c$ folgt $M \setminus A = M \setminus V^c = M \cap V = V \setminus M^c$ und damit $\lambda(M \setminus A) < \varepsilon$. □

Eine als abzählbarer Durchschnitt offener Mengen darstellbare Menge heißt G_δ-Menge und eine als abzählbare Vereinigung abgeschlossener Mengen darstellbare Menge heißt F_σ-Menge. Offensichtlich sind G_δ-Mengen und F_σ-Mengen Borelmengen.

Satz 4.3. *Sei $M \in \mathcal{L}^d$. Dann existieren eine G_δ-Menge $U \subseteq \mathbb{R}^d$ und eine F_σ-Menge $A \subseteq \mathbb{R}^d$ mit $A \subseteq M \subseteq U$ und $\lambda(U \setminus M) = \lambda(M \setminus A) = 0$.*

Beweis. Nach Satz 4.2 existieren zu $n \in \mathbb{N}$ eine offene Menge U_n und eine abgeschlossene Menge A_n mit $A_n \subseteq M \subseteq U_n$, so dass $\lambda(U_n \setminus M) < 1/n$ und $\lambda(M \setminus A_n) < 1/n$. Die Behauptung folgt mit $U = \bigcap_{n=1}^{\infty} U_n$ und $A = \bigcup_{n=1}^{\infty} A_n$. □

Satz 4.3 impliziert, dass das $(\mathbb{R}^d, \mathcal{L}, \lambda)$ die Vervollständigung von $(\mathbb{R}^d, \mathcal{B}, \lambda|_{\mathcal{B}})$ in dem nach Definition 2.4 beschriebenen Sinne ist.

Gemäß Definition 2.4 heißt eine Teilmenge N von $\mathbb{R}^d$ *Lebesgue-Nullmenge* (oder kurz *Nullmenge*), falls $\lambda(N) = 0$. Satz 4.3 besagt also, dass eine Lebesgue-messbare Menge M von der Form $M = A \cup N$ mit einer F_σ-Menge A und einer Nullmenge N ist.

5 Messbare Abbildungen und Bildmaße

Eine Abbildung zwischen metrischen Räumen ist stetig, wenn Urbilder offener Mengen offen sind, d.h., $f\colon X \to Y$ mit metrischen Räumen X und Y ist genau dann stetig, wenn für jede offene Teilmenge U von Y auch das Urbild $f^{-1}(U)$ offen ist.

Die folgende Definition ist ganz analog.

Definition 5.1. Seien (X, S) und (X', S') Messräume und $f\colon X \to X'$. Dann heißt f *messbar*, wenn für jede messbare Teilmenge B' von X' auch das Urbild $f^{-1}(B')$ messbar ist, d.h., aus $B' \in S'$ folgt $f^{-1}(B') \in S$.

Genauer heißt f auch (S, S')-*messbar*. Wir schreiben auch, dass $f\colon (X, S) \to (X', S')$ messbar ist.

Als Beispiel sei für $A \subseteq X$ die *charakteristische Funktion* $\chi_A\colon X \to \mathbb{R}$,

$$\chi_A(x) = \begin{cases} 1, & x \in A, \\ 0, & x \notin A, \end{cases}$$

betrachtet. Es gilt $\chi_A^{-1}(B) \in \{\emptyset, A, A^c, X\}$ für alle $B \subseteq \mathbb{R}$. Für messbares A ist also χ_A messbar, und legt man auf $\mathbb{R}$ die Borelsche σ-Algebra zu Grunde (oder irgendeine andere σ-Algebra, in der 0 und 1 nicht immer in der gleichen Menge liegen), so gilt auch die Umkehrung.

Satz 5.2. *Seien (X, S) und (X', S') Messräume und $f\colon X \to X'$. Sei E' Erzeuger von S'. Dann ist f genau dann messbar, falls $f^{-1}(B') \in S$ für alle $B' \in E'$.*

Beweis. Wegen $E' \subseteq S'$ folgt, dass für messbares f die im Satz genannte Bedingung gilt.

Für die umgekehrte Richtung sei $f^{-1}(B') \in S$ für alle $B' \in E'$. Es ist also $f^{-1}(E') := \{f^{-1}(B')\colon B' \in E'\} \subseteq S$. Wir betrachten

$$T' := \{B' \subseteq X'\colon f^{-1}(B') \in \sigma(f^{-1}(E'))\}.$$

Nach Übung ist T' σ-Algebra in X'. Offensichtlich ist $E' \subseteq T'$ und damit $S' \subseteq T'$. Für $B' \in S'$ gilt damit $B' \in T'$, also $f^{-1}(B') \in \sigma(f^{-1}(E')) \subseteq \sigma(S) = S$. $\square$

In metrischen Räumen legen wir stets die Borelsche σ-Algebra zu Grunde. Sind also X und Y metrische Räume, so heißt $f\colon X \to Y$ *messbar* (genauer *Borel-messbar*), falls dies bezüglich der σ-Algebren $\mathcal{B}(X)$ und $\mathcal{B}(Y)$ gilt.

Satz 5.3. *Seien X und Y metrische Räume und sei $f\colon X \to Y$ stetig. Dann ist f (Borel)-messbar.*

Beweis. Nach Satz 5.2 reicht es zu zeigen, dass $f^{-1}(M) \in \mathcal{B}(X)$ für alle M aus einem Erzeuger von $\mathcal{B}(Y)$ gilt. Da $\mathcal{B}(Y)$ von den offenen Mengen erzeugt wird, reicht es also, $f^{-1}(M) \in \mathcal{B}(X)$ für alle offenen Teilmengen M von Y zu zeigen. Für diese ist aber wegen der Stetigkeit $f^{-1}(M)$ offen, also auch in $\mathcal{B}(X)$. $\square$

Definition und Satz 5.4. Sei $f\colon (X,S) \to (X',S')$ messbar und sei μ Maß auf S, also (X,S,μ) Maßraum. Dann ist durch

$$f(\mu)(B') := \mu(f^{-1}(B')), \quad B' \in S',$$

ein Maß $f(\mu)$ auf S' gegeben, das sogenannte *Bildmaß von μ unter f.*

Beweis. Zunächst notieren wir, dass wegen der Messbarkeit von f für $B' \in S'$ auch $f^{-1}(B') \in S$ gilt, also $f(\mu)$ definiert ist.

Außerdem ist $f(\mu)(\emptyset) = \mu(f^{-1}(\emptyset)) = \mu(\emptyset) = 0$ und für eine Folge (B'_n) paarweise disjunkter Mengen in S' besteht auch die Folge $(f^{-1}(B'_n))$ in S aus paarweise disjunkten Mengen und es gilt

$$\begin{aligned} f(\mu)\left(\bigcup_{n=1}^{\infty} B'_n\right) &= \mu\left(f^{-1}\left(\bigcup_{n=1}^{\infty} B'_n\right)\right) \\ &= \mu\left(\bigcup_{n=1}^{\infty} f^{-1}(B'_n)\right) \\ &= \sum_{n=1}^{\infty} \mu(f^{-1}(B'_n)) \\ &= \sum_{n=1}^{\infty} f(\mu)(B'_n) \end{aligned}$$

□

Satz 5.5. *Seien $f\colon (X,S) \to (X',S')$ und $g\colon (X',S') \to (X'',S'')$ messbar. Dann ist auch $g \circ f\colon (X,S) \to (X'',S'')$ messbar.*

Der einfache Beweis sei als Übung überlassen. Sind f und g wie in Satz 5.5 und ist μ Maß auf S, so ist also $(g \circ f)(\mu)$ Maß auf S''. Man zeigt leicht, dass $(g \circ f)(\mu) = g(f(\mu))$.

Als Beispiel betrachten wir für $c \in \mathbb{R}^d$ die *Translation* $T_c\colon \mathbb{R}^d \to \mathbb{R}^d$, $T_c(x) = x + c$. Da T_c stetig ist, ist T_c auch messbar. Für einen Quader $Q = (a,b] \in \mathbb{R}^d$ mit $a_j \leq b_j$ für $1 \leq j \leq d$ gilt dann

$$(T_c(\lambda))(Q) = \lambda(T_{-c}(Q)) = \lambda((a-c, b-c]) = \prod_{j=1}^{d}(a_j - b_j) = \lambda((a,b]) = \lambda(Q).$$

Nach dem Maßeindeutigkeitssatz 3.8 folgt hieraus, dass $T_c(\lambda)$ und λ auf $\mathcal{B}$ übereinstimmen. Wir sagen auch, dass das Lebesgue-Borel-Maß (also λ eingeschränkt auf $\mathcal{B}$) *translationsinvariant* ist.

Satz 5.6. *Ist μ translationsinvariantes Maß auf $\mathcal{B}^d$ mit $\mu((0,1]^d) = 1$, so gilt $\mu = \lambda$.*

Beweisskizze. Aus der Translationsinvarianz folgt, dass ein Würfel der Seitenlänge $1/n$ mit $n \in \mathbb{N}$ das Maß $1/n^d$ hat. Hieraus erhält man dann, dass $\mu((a,b]) = \lambda((a,b]]$ für $a, b \in \mathbb{Q}^d$ gilt. (Man wähle n als kleinstes gemeinsames Vielfaches der Nenner der a_j und b_j und zerlege $(a,b]$ in Würfel der Kantenlänge $1/n$.) Da $\mathcal{B}$ nach Satz 1.5 von diesen Quadern $(a,b]$ erzeugt wird, folgt die Behauptung.□

Falls in Satz 5.6 nur $\alpha := \mu((0,1]) < \infty$ vorausgesetzt wird, folgt $\mu = \alpha \cdot \lambda$.

Als weiteres Beispiel für Bildmaße betrachten wir für $c = (c_1, \ldots, c_d) \in \mathbb{R}^d$ mit $c_j > 0$ für alle j die durch $D_c(x_1, \ldots, x_d) = (c_1 x_1, \ldots, c_d x_d)$ definierte Abbildung $D_c\colon \mathbb{R}^d \to \mathbb{R}^d$. Setzt man $1/c = (1/c_1, \ldots, 1/c_d)$, so gilt für einen Quader $Q = (a, b] \in \mathbb{R}^d$ mit $a_j \leq b_j$ für $1 \leq j \leq d$, dass

$$(D_c(\lambda))(Q) = \lambda(D_{1/c}(Q)) = \lambda\biggl(\Bigl(\frac{a}{c}, \frac{b}{c}\Bigr]\biggr) = \frac{1}{\prod_{j=1}^d c_j} \prod_{j=1}^d (a_j - b_j) = \frac{1}{\prod_{j=1}^d c_j} \lambda(Q)$$

und damit

$$(D_c(\lambda)(M) = \frac{1}{\prod_{j=1}^d c_j} \lambda(M)$$

für alle $M \in \mathcal{B}$.

Eine Abbildung $\phi\colon \mathbb{R}^d \to \mathbb{R}^d$ der Form $\phi(x) = Ax + b$ mit einer $(d \times d)$-Matrix A und $b \in \mathbb{R}^d$ heißt *affin* (oder *affin-linear*). Man nennt $\det \phi := \det A$ die Determinante von ϕ. Affine Abbildungen sind stetig und daher auch messbar. Für das Bildmaß $\phi(\lambda)$ auf $\mathcal{B}$ gilt der folgende Satz.

Satz 5.7. *Sei $\phi\colon \mathbb{R}^d \to \mathbb{R}^d$ affin und bijektiv. Dann gilt*

$$\phi(\lambda) = \frac{1}{|\det \phi|} \lambda.$$

Für $M \in \mathcal{B}$ gilt also

$$\lambda(\phi^{-1}(M)) = \phi(\lambda)(M) = \frac{\lambda(M)}{|\det \phi|} = \lambda(M) |\det \phi^{-1}|.$$

Ersetzt man ϕ durch ϕ^{-1} so erhält man

$$\lambda(\phi(M)) = |\det \phi| \cdot \lambda(M).$$

Beweisskizze von Satz 5.7. Sei $\phi(x) = Ax + b$. Wegen der Translationsinvarianz kann man $b = 0$ annehmen. Das obige Beispiel D_c liefert die Aussage für Diagonalmatrizen. Nun hat A eine Zerlegung der Form $A = S_1 D S_2$ mit orthogonalen Matrizen S_1 und S_2 und einer Diagonalmatrix D. (Den Beweis dafür skizzieren wir unten.) Es reicht also, die Behauptung für orthogonale Abbildungen zu beweisen.

Sei also ϕ orthogonal, etwa $\phi(x) = Sx$ mit S orthogonal (also $S^t = S^{-1}$). Für eine Translation T_c gilt nun $T_c \circ \phi = \phi \circ T_a$ mit $a = \phi^{-1}(c)$. Hieraus folgt, dass $\phi(\lambda)$ translationsinvariant ist, also $\phi(\lambda) = \alpha \cdot \lambda$ nach Satz 5.6 und der anschließenden Bemerkung. Da ϕ die Einheitskugel $\{x \in \mathbb{R}^d\colon \|x\|_2 \leq 1\}$ invariant lässt, folgt $\alpha = 1$ und damit $\phi(\lambda) = \lambda$.

Um die obige Zerlegung $A = S_1 D S_2$ mit orthogonalen Matrizen S_1 und S_2 und einer Diagonalmatrix D zu beweisen, notieren wir zunächst, dass AA^t eine positiv definite symmetrische Matrix ist. Nach Hauptachsentransformation existiert damit eine orthogonale Matrix S_1, für die $S_1^t A A^t S_1$ eine Diagonalmatrix mit positiven Einträgen ist. Wir können diese Matrix in der Form D^2 mit einer Diagonalmatrix D mit positiven Einträgen schreiben, d.h., $S_1^t A A^t S_1 = D^2$. Nun ist auch $S_2 := D^{-1} S_1^t A$ orthogonal, denn $S_2 S_2^t = D^{-1} S_1^t A A^t S_1 D^{-1}$ ist die Einheitsmatrix, und $S_1 D S_2 = S_1 D D^{-1} S_1^t A = A$. $\square$

Der folgende Satz zeigt, warum wir das Lebesgue-Maß bzw. Lebesgue-Borel-Maß nicht auf ganz $P(\mathbb{R}^d)$, sondern nur auf geeigneten σ-Algebra definieren konnten.

Satz 5.8. *Es gibt kein translationsinvariantes Maß auf $\mathbb{R}^d$, für welches $(0,1]^d$ endliches und positives Maß hat.*

Beweis. Wir nehmen an, dass so ein Maß μ existiert. Sei $\sim$ die durch

$$x \sim y \Leftrightarrow x - y \in \mathbb{Q}^d$$

definierte Relation auf $(0,1]^d$. Dann ist $\sim$ Äquivalenzrelation. Sei R ein Repräsentantensystem dieser Äquivalenzrelation, d.h., es ist $R \subseteq (0,1]^d$ mit $\bigcup_{r \in R} [r] = (0,1]^d/\!\sim$ und es gilt $r \not\sim s$ für $r, s \in (0,1]^d$ mit $r \neq s$.

Wir zeigen, dass R nicht (μ-)messbar ist. Dazu nehmen wir an, R wäre messbar. Dann ist wegen der Translationsinvarianz für $q \in \mathbb{Q}^d$ auch $q + R$ messbar, mit $\mu(q+R) = \mu(R)$. Sei (q_k) eine Abzählung von $\mathbb{Q}^d \cap (-1,1]^d$. Dann gilt

$$(0,1]^d \subseteq A := \bigcup_{k=1}^{\infty} (q_k + R) \subseteq (-1,2]$$

und wir erhalten, dass A messbar ist mit

$$\mu(A) = \sum_{k=1}^{\infty} \mu(q_k + R) = \sum_{k=1}^{\infty} \mu(R).$$

Es folgt, dass $\mu(A) \in \{0, \infty\}$. Andererseits gilt

$$\mu((0,1]^d) \leq \mu(A) \leq \mu((-1,2]^d) = 3^d \mu((0,1]^d).$$

Das ist ein Widerspruch. $\square$

6 Messbarkeit reellwertiger Funktionen

Wir hatten bereits bei der Definition des Maßes das Intervall $[0,\infty]$ betrachtet, und für den Wert ∞ die Regel $a + \infty = \infty$ für $a \in [0,\infty]$ benutzt. Allgemeiner betrachten wir jetzt $\overline{\mathbb{R}} := [-\infty,\infty] = \mathbb{R} \cup \{\pm\infty\}$, mit $-\infty < a < \infty$ für $a \in \mathbb{R}$. Der Vorteil ist, dass jede Teilmenge von $\overline{\mathbb{R}}$ ein Supremum und ein Infimum hat.

Wir benutzen die naheliegenden Regeln

- $a + \infty = \infty$ für $a \in (-\infty,\infty]$;
- $a - \infty = -\infty$ für $a \in [-\infty,\infty)$;
- $a \cdot (\pm\infty) = \pm\infty$ für $a \in (0,\infty]$;

- $a \cdot (\pm\infty) = \mp\infty$ für $a \in [-\infty, 0)$.

Weiter stellt es sich als sinnvoll heraus, $0 \cdot (\pm\infty) = 0$ zu setzen.

Es ist leicht zu sehen, dass auf $\overline{\mathbb{R}}$ durch

$$\mathcal{B}(\overline{\mathbb{R}}) = \{M \subseteq \overline{\mathbb{R}} \colon M \cap \mathbb{R} \in \mathcal{B}(\mathbb{R})\}.$$

eine σ-Algebra definiert ist. Wir werden bei Abbildungen nach $\overline{\mathbb{R}}$ (oder $\mathbb{R}$) im Folgenden immer diese σ-Algebra zu Grunde legen.

Analog zu Satz 1.5 und der nachfolgenden Bemerkung sieht man, dass $\mathcal{B}(\overline{\mathbb{R}})$ von den Intervallen $[-\infty, a]$ mit $a \in \mathbb{Q}$ (oder auch $a \in \mathbb{R}$) erzeugt wird, ebenso auch von den Intervallen $[a, \infty]$, $[-\infty, a)$, usw., wiederum sowohl für $a \in \mathbb{Q}$ und $a \in \mathbb{R}$.

Nach Satz 5.2 ist damit für einen Messraum (X, S) eine Funktion $f \colon X \to \overline{\mathbb{R}}$ genau dann messbar, wenn $f^{-1}([-\infty, a]) = \{x \in X \colon f(x) \leq a\}$ für alle $a \in \mathbb{Q}$ messbar ist, d.h., $\{x \in X \colon f(x) \leq a\} \in S$ gilt.

Statt $\{x \in X \colon f(x) \leq a\}$ schreiben wir im Folgenden $\{f \leq a\}$. Analog sind Schreibweisen wie $\{f < a\}$ oder $\{f \leq g\}$ mit Funktionen f und g definiert.

Obige Bemerkung nimmt dann folgende Form an.

Satz 6.1. *Sei (X, S) Messraum. Eine Funktion $f \colon X \to \overline{\mathbb{R}}$ ist genau dann messbar, wenn $\{f \leq a\} \in S$ für alle $a \in \mathbb{Q}$.*

Satz 6.2. *Sei (X, S) Messraum und sei (f_n) eine Folge messbarer Funktionen von X nach $\overline{\mathbb{R}}$. Dann sind auch die (punktweise definierten) Funktionen*

$$\sup_{n \in \mathbb{N}} f_n, \quad \inf_{n \in \mathbb{N}} f_n, \quad \limsup_{n \to \infty} f_n \quad \textit{und} \quad \liminf_{n \to \infty} f_n$$

messbar. Insbesondere ist $\lim_{n \to \infty} f_n$ messbar, wenn der Grenzwert existiert.

Beweis. Die Messbarkeit des Supremums folgt aus

$$\left\{\sup_{n \in \mathbb{N}} f_n \leq a\right\} = \bigcap_{n \in \mathbb{N}} \{f_n \leq a\} \in S.$$

Die Messbarkeit des Infimums folgt analog. Die Messbarkeit von Limes superior und inferior folgt aus

$$\limsup_{n \to \infty} f_n = \inf_{n \in \mathbb{N}} \left(\sup_{k \geq n} f_k\right) \quad \text{und} \quad \liminf_{n \to \infty} f_n = \sup_{n \in \mathbb{N}} \left(\inf_{k \geq n} f_k\right). \qquad \square$$

Satz 6.3. *Sei (X, S) Messraum. Eine Funktion $f = (f_1, \ldots, f_d) \colon X \to \mathbb{R}^d$ ist genau dann messbar, wenn jede Koordinatenfunktion f_j messbar ist.*

Beweis. Für $j \in \{1, \ldots, d\}$ ist die durch $(x_1, \ldots, x_d) \to x_j$ definierte Projektionsabbildung $\pi_j \colon \mathbb{R}^d \to \mathbb{R}$ stetig und damit nach Satz 5.3 messbar. Nach Satz 5.5 ist damit auch $f_j = \pi_j \circ f$ messbar.

Seien jetzt umgekehrt alle f_j messbar. Für einen Quader $(a,b]$, mit $a = (a_1,\ldots,a_d]$ und $b = (b_1,\ldots,b_d]$, ist dann

$$f^{-1}((a,b]) = \bigcap_{j=1}^{d} f_j^{-1}((a_j,b_j]) \in S.$$

Da $\mathcal{B}$ nach Satz 1.5 von diesen Quadern erzeugt wird, folgt die Behauptung nach Satz 5.2. $\square$

Satz 6.4. *Seien (X,S) Messraum, $f,g\colon X\to\mathbb{R}$ messbar und $c\in\mathbb{R}$. Dann sind auch $c\cdot f$, $f^+ = \max\{f,0\}$, $f^- := \max\{-f,0\}$, $|f|$, $f\pm g$ und $f\cdot g$ messbar.*

Beweis. Nach Satz 6.3 ist $h\colon X\to\mathbb{R}^2$, $x\mapsto(f(x),g(x))$ messbar. Weiter ist $s\colon\mathbb{R}^2\to\mathbb{R}$, $s(x_1,x_2)=x_1+x_2$, stetig und damit messbar. Wegen

$$(f+g)(x) = f(x)+g(x) = (s\circ h)(x)$$

folgt damit die Messbarkeit von $f+g$ nach Satz 5.5.

Analog sieht man, dass $f\cdot g$ messbar ist. Damit ist auch $c\cdot f$ messbar, insbesondere auch $-f$. Die Messbarkeit von f^+ und f^- folgt dann sofort aus Satz 6.2, und die von $|f|$ wegen $|f| = f^+ + f^-$. $\square$

Satz 6.4 impliziert, dass auch $\{f\le g\}$ messbar ist, denn $\{f\le g\}=\{f-g\le 0\}$. Analog sind auch Mengen wie $\{f=g\}$ oder $\{f\ne g\}$ messbar.

7 Integration nichtnegativer messbarer Funktionen

Definition 7.1. Sei (X,S) Messraum. Eine messbare Funktion $f\colon X\to\overline{\mathbb{R}}$ heißt *Treppenfunktion*, wenn sie nur endlich viele Werte annimmt, d.h., wenn $f(X)$ endliche Teilmenge von $\mathbb{R}$ ist.

Ist f Treppenfunktion und $f(X)=\{\alpha_1,\ldots,\alpha_n\}$, wobei wir die α_j als verschieden annehmen, so gilt mit $A_j = f^{-1}(\alpha_j)$ also

$$f = \sum_{j=1}^{m} \alpha_j \chi_{A_j}.$$

Wir nennen dies die *kanonische Darstellung* von f. Eine Treppenfunktion kann auch auf viele andere Arten dargestellt werden. Beispielsweise ist

$$\chi_{(1,3]} + \chi_{(2,4]} = \chi_{(1,2]} + 2\cdot\chi_{(2,3]} + \chi_{(3,4]} = 1\cdot\chi_{(1,2]\cup(3,4]} + 2\cdot\chi_{(2,3]} + 0\cdot\chi_{(-\infty,1]\cup(4,\infty)}.$$

Die Darstellung auf der rechten Seite ist (für $X=\mathbb{R}$) die kanonische.

Leicht sieht man, dass eine Summe von Treppenfunktionen wieder eine Treppenfunktion ist.

Definition 7.2. Sei (X, S, μ) Maßraum, $M \in S$ und $f\colon X \to [0, \infty]$ Treppenfunktion mit kanonischer Darstellung

$$f = \sum_{j=1}^{m} \alpha_j \chi_{A_j}.$$

Dann heißt

$$\int_M f\, d\mu = \sum_{j=1}^{m} \alpha_j \mu(A_j \cap M)$$

das *(μ-)Integral von f (über M).*

Lemma 7.3. *Seien (X, S, μ) Maßraum und $f\colon X \to [0, \infty]$ Treppenfunktion. Dann ist durch $\nu\colon S \to [0, \infty]$, $\nu(A) = \int_A f\, d\mu$, ein Maß definiert.*

Beweis. Sei

$$f = \sum_{j=1}^{m} \alpha_j \chi_{A_j}$$

die kanonische Darstellung von f. Offensichtlich ist $\nu(\emptyset) = 0$ und für eine Folge (B_l) paarweise disjunkter Mengen in S gilt mit $B = \bigcup_{l=1}^{\infty} B_l$, dass

$$\begin{aligned}
\nu(B) &= \int_B f\, d\mu \\
&= \sum_{j=1}^{m} \alpha_j \mu(A_j \cap B) \\
&= \sum_{j=1}^{m} \alpha_j \mu\left(\bigcup_{l=1}^{\infty} A_j \cap B_l\right) \\
&= \sum_{j=1}^{m} \alpha_j \sum_{l=1}^{\infty} \mu(A_j \cap B_l) \\
&= \sum_{l=1}^{\infty} \sum_{j=1}^{m} \alpha_j \mu(A_j \cap B_l) \\
&= \sum_{l=1}^{\infty} \nu(B_l).
\end{aligned}$$

□

Wir stellen einige Rechenregeln zusammen.

Lemma 7.4. *Seien (X, S, μ) Maßraum, $M, N \in S$ und $f, g\colon X \to [0, \infty]$ Treppenfunktionen. Dann gilt:*

(i) $\int_M (f + g) d\mu = \int_M f\, d\mu + \int_M g\, d\mu$.

(ii) *Für $\alpha \geq 0$ gilt $\int_M \alpha f\, d\mu = \alpha \int_M f\, d\mu$.*

(iii) *Ist $f \leq g$, so ist $\int_M f\, d\mu \leq \int_M g\, d\mu$.*

(iv) *Ist $\mu(M) = 0$, so ist $\int_M f\, d\mu = 0$.*

(v) *Ist $M \subseteq N$, so ist $\int_M f\, d\mu \leq \int_N f\, d\mu$.*

Beweis. (i) Seien

$$f = \sum_{j=1}^{m} \alpha_j \chi_{A_j}. \quad \text{und} \quad g = \sum_{k=1}^{n} \beta_k \chi_{B_k}$$

die kanonischen Darstellungen von f und g. Die Mengen $E_{jk} := A_j \cap B_k \cap M$ sind dann paarweise disjunkt und es gilt, da $\bigcup_{j=1}^m A_j = \bigcup_{k=1}^n B_k = X$, dass

$$A_j \cap M = \bigcup_{k=1}^{n} E_{jk}, \quad B_k \cap M = \bigcup_{j=1}^{m} E_{jk} \quad \text{und} \quad M = \bigcup_{j=1}^{m} \bigcup_{k=1}^{n} E_{jk}.$$

Daher folgt mit Lemma7.3, dass

$$\begin{aligned}
\int_M (f+g) d\mu &= \sum_{j=1}^{m} \sum_{k=1}^{n} \int_{E_{j,k}} (f+g) d\mu \\
&= \sum_{j=1}^{m} \sum_{k=1}^{n} (\alpha_j + \beta_k) \mu(E_{j,k}) \\
&= \sum_{j=1}^{m} \alpha_j \sum_{k=1}^{n} \mu(E_{j,k}) + \sum_{k=1}^{n} \beta_k \sum_{j=1}^{m} \mu(E_{j,k}) \\
&= \sum_{j=1}^{m} \alpha_j \mu(A_j \cap M) + \sum_{k=1}^{n} \beta_k \mu(B_k \cap M) \\
&= \int_M f \, d\mu + \int_M g \, d\mu.
\end{aligned}$$

Aussage (ii) ist klar, (iii) folgt leicht aus (i), da $f - g \geq 0$, und (iv) und (v) folgen direkt aus der Definition. □

Aussage (i) gilt allgemeiner auch für endliche Summen. Hieraus folgt, dass die Gleichung

$$\int_M \sum_{j=1}^{m} \alpha_j \chi_{A_j} \, d\mu = \sum_{j=1}^{m} \alpha_j \int_M \chi_{A_j} \, d\mu = \sum_{j=1}^{m} \alpha_j \mu(A_j \cap M)$$

auch dann gilt, wenn die Treppenfunktion im Integranden nicht in kanonischer Darstellung gegeben ist.

Die Idee ist nun, dass Integral von messbaren Funktionen zu definieren, indem man sie durch Treppenfunktionen approximiert.

Satz 7.5. *Seien (X, S) Messraum und $f\colon X \to [0, \infty]$ messbar. Dann existiert eine monoton steigende Folge (f_n) von nichtnegativen Treppenfunktionen mit $f = \lim_{n\to\infty} f_n$.*

Beweis. Für $n \in \mathbb{N}$ und $0 \leq j \leq n2^n$ sei

$$A_{n,j} = f^{-1}\left(\left[\frac{j}{2^n}, \frac{j+1}{2^n}\right)\right)$$

sei

$$f_n = \sum_{j=1}^{n2^n} \frac{j}{2^n} \chi_{A_{n,j}}.$$

Man überprüft leicht, dass die f_n das Verlangte leisten; vgl. auch Abbildung 2.□

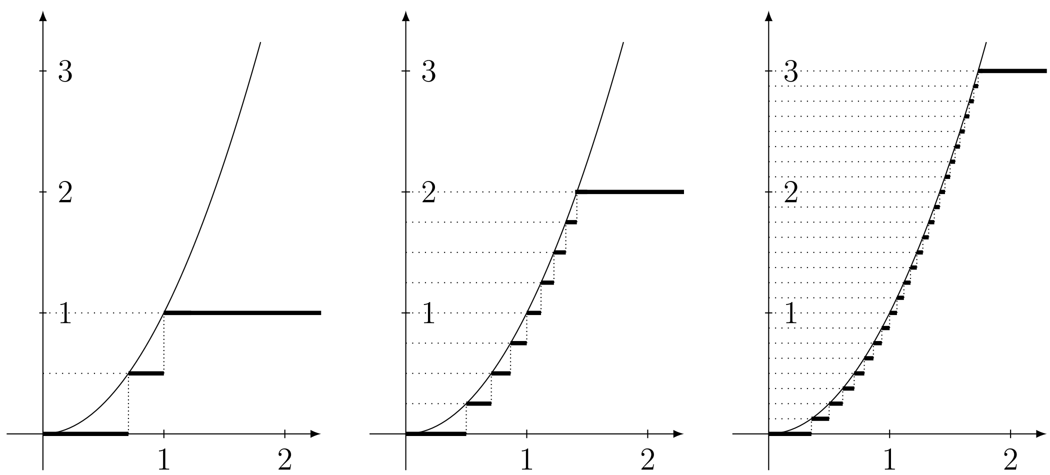

Abbildung 2: Die Treppenfunktionen f_1, f_2 und f_3 aus dem Beweis von Satz 7.5 für $f\colon [0,\infty] \to [0,\infty]$, $f(x) = x^2$.

Mit Satz 6.2 sind die messbaren Funktionen sogar genau die, die punktweiser Grenzwert einer monoton steigenden Folge von Treppenfunktionen sind.

Bemerkung. Auch bei der Definition des Riemannintegrals wird eine Funktion $f\colon [a,b] \to \mathbb{R}$ durch Treppenfunktionen approximiert. Dort wird aber mit einer Zerlegung des Definitionsbereichs $[a,b]$ gearbeitet, während in obigem Beweis der Wertebereich $[0,\infty]$ zerlegt wird und die daraus resultierende Zerlegung des Definitionsbereichs betrachtet wird.

Korollar 7.6. *Ist (X,S) Messraum und $f\colon X \to \mathbb{R}$ messbar, so existiert eine Folge (f_n) von Treppenfunktionen mit $f = \lim_{n\to\infty} f_n$.*

Zum Beweis schreibe man $f = f^+ - f^-$ und wende Satz 7.5 auf f^+ und f^- an.

Definition 7.7. Sei (X,S,μ) Maßraum, $M \in S$ und $f\colon X \to [0,\infty]$ messbar. Dann heißt

$$\int_M f\,d\mu = \sup\left\{\int_M t\,d\mu \colon t \text{ Treppenfunktion mit } 0 \le t \le f\right\}$$

das *(μ-)Integral von f (über M).*

Man beachte, dass für eine Treppenfunktion f diese Definition mit der aus Definition 7.2 übereinstimmt. Dies folgt aus Lemma 7.4, (iii), und weil man in obigem Supremum auch $t = f$ wählen kann.

Die Regeln aus Lemma 7.4 gelten analog.

Lemma 7.8. *Seien (X, S, μ) Maßraum, $M, N \in S$ und $f, g\colon X \to [0, \infty]$ messbar. Dann gilt:*

(i) $\int_M (f+g) d\mu = \int_M f\, d\mu + \int_M g\, d\mu$.

(ii) *Für $\alpha \geq 0$ gilt $\int_M \alpha f\, d\mu = \alpha \int_M f\, d\mu$.*

(iii) *Ist $f \leq g$, so ist $\int_M f\, d\mu \leq \int_M g\, d\mu$.*

(iv) *Ist $\mu(M) = 0$, so ist $\int_M f\, d\mu = 0$.*

(v) *Ist $M \subseteq N$, so ist $\int_M f\, d\mu \leq \int_N f\, d\mu$.*

Für den Beweis des folgenden Satzes wird nur die Monotonieaussage (iii) benötigt. Diese folgt direkt aus der entsprechenden Aussage von Lemma 7.4 und Definition 7.7.

Satz 7.9. (Satz von Beppo Levi; Satz von der monotonen Konvergenz) *Seien (X, S, μ) Maßraum, $M \in S$, (f_n) eine monoton steigende Folge messbarer Funktionen von X nach $[0, \infty]$ und $f = \lim_{n\to\infty} f_n$. Dann ist f messbar und es gilt*

$$\int_M f_n\, d\mu \to \int_M f\, d\mu.$$

Beweis. Die Messbarkeit von f folgt aus Satz 6.2. Da (f_n) monoton steigt, gilt dies auch für die Folge $\left(\int_M f_n\, d\mu\right)$. Damit existiert

$$\gamma := \lim_{n\to\infty} \int_M f_n\, d\mu \in [0, \infty].$$

Wegen der Monotonie ist $f_n \leq f$ und damit $\int_M f_n\, d\mu \leq \int_M f\, d\mu$, also

$$\gamma \leq \int_M f\, d\mu.$$

Für die umgekehrte Richtung ist zu zeigen, dass

$$\int_M g\, d\mu \leq \gamma$$

für jede Treppenfunktion g mit $g \leq f$. Sei dazu $0 \leq c < 1$. Wir setzen $M_n = \{x \in M\colon f_n(x) \geq c\, g(x)\}$. Wegen der Monotonie gilt $M_1 \subseteq M_2 \subseteq \dots$ und es ist $M = \bigcup_{n=1}^\infty M_n$. Mit Lemma 7.4 folgt

$$\gamma \geq \int_M f_n\, d\mu \geq \int_{M_n} f_n\, d\mu \geq \int_{M_n} c\, g\, d\mu = c \int_{M_n} g\, d\mu.$$

Nach Lemma 7.3 ist durch $A \to \int_A g\, d\mu$, ein Maß gegeben. Nach Lemma 2.3, (i), gilt damit

$$\int_{M_n} g\, d\mu \to \int_M g\, d\mu.$$

Da c mit $0 \leq c < 1$ beliebig war, folgt $\int_M g\, d\mu \leq \gamma$. □

Mit dem Satz über monotone Konvergenz können Resultate für Treppenfunktionen leicht auf messbare Funktion nach $[0,\infty]$ übertragen werden, denn nach Satz 7.5 ist jede solche Funktion Grenzwert einer monoton steigenden Folge von Treppenfunktionen. Insbesondere erhält man so die noch nicht bewiesenen Aussagen von Lemma 7.8.

Auch das folgende Lemma folgt mit denselben Argumenten, aber auch direkt aus der Definition.

Lemma 7.10. *Seien (X,S,μ) Maßraum, $M\in S$ und $f\colon X\to[0,\infty]$ messbar. Dann gilt*

$$\int_M f\,d\mu=\int_X \chi_M f\,d\mu.$$

Die Aussage dieses Lemmas wird oft auch zur Definition von $\int_M f\,d\mu$ genommen. Für disjunkte Mengen $M,N\in S$ folgt wegen $\chi_{M\cup N}=\chi_M+\chi_N$, dass

$$\int_{M\cup N} f\,d\mu=\int_M f\,d\mu+\int_N f\,d\mu.$$

Da eine Reihe definitionsgemäß nichts anderes als eine Folge ist, und umgekehrt jede Folge auch als Reihe geschrieben werden kann, hat auch der Satz über monotone Konvergenz ein Analogon für Reihen.

Korollar 7.11. *Seien (X,S,μ) Maßraum, $M\in S$ und (g_n) eine Folge von messbaren Funktionen von X nach $[0,\infty]$. Dann gilt*

$$\sum_{n=1}^\infty\int_M g_n\,d\mu=\int_M\sum_{n=1}^\infty g_n\,d\mu.$$

Zum Beweis muss man nur $f_n=g_1+\cdots+g_n$ setzen.

Der folgende Satz ist eine Konsequenz des Satzes über monotone Konvergenz, gilt aber auch für Folgen, die nicht monoton sind.

Satz 7.12. (Lemma von Fatou) *Seien (X,S,μ) Maßraum, $M\in S$ und (f_n) eine Folge von messbaren Funktionen von X nach $[0,\infty]$. Dann gilt*

$$\int_M\liminf_{n\to\infty} f_n\,d\mu\le\liminf_{n\to\infty}\int_M f_n\,d\mu.$$

Beweis. Die durch $g_n:=\inf_{k\ge n} f_k$ definierte Folge (g_n) ist monoton steigend und nach Definition des Limes inferior ist

$$\liminf_{n\to\infty} f_n=\sup_{n\in\mathbb{N}}\left(\inf_{k\ge n} f_k\right)=\sup_{n\in\mathbb{N}} g_n=\lim_{n\to\infty} g_n.$$

Mit dem Satz von der monotonen Konvergenz und da $g_n\le f_n$ folgt

$$\begin{aligned}\int_M\liminf_{n\to\infty} f_n\,d\mu&=\int_M\lim_{n\to\infty} g_n\,d\mu\\&=\lim_{n\to\infty}\int_M g_n\,d\mu\\&=\liminf_{n\to\infty}\int_M g_n\,d\mu\\&\le\liminf_{n\to\infty}\int_M f_n\,d\mu.\end{aligned}$$ ☐

8 Integrierbare Funktionen

Definition 8.1. Sei (X, S, μ) Maßraum, $M \in S$ und $f\colon X \to \overline{\mathbb{R}}$ messbar. Dann heißt f *integrierbar* (genauer *μ-integrierbar*) über M, wenn

$$\int_M f^+ \, d\mu < \infty \quad \text{und} \quad \int_M f^- \, d\mu < \infty.$$

In diesem Fall heißt

$$\int_M f \, d\mu = \int_M f^+ \, d\mu - \int_M f^- \, d\mu$$

das *Lebesgue-Integral bzgl.* μ (von f über M) oder auch *μ-Integral* (von f über M).

Ist μ das Lebesgue-Maß, also $\mu = \lambda$, so sprechen wir auch kurz vom *Lebesgue-Integral.*

Für $f\colon X \to [0, \infty]$ ist $f^+ = f$ und $f^- = 0$. Daher stimmt das hier definierte Integral für solche Funktionen mit dem aus Definition 7.7 überein.

Da $|f| = f^+ + f^-$ folgt sofort, dass mit f auch $|f|$ integrierbar ist. Ist umgekehrt f messbar und $|f|$ integrierbar, so ist auch f integrierbar, da $\int_M f^+ \, d\mu \leq \int_M |f| \, d\mu$ und da $\int_M f^- \, d\mu \leq \int_M |f| \, d\mu$ nach Lemma 7.8.

Eine Funktion $f\colon X \to \mathbb{C}$ heißt integrierbar (über M), wenn Real- und Imaginärteil von f integrierbar (über M) sind. Man setzt dann

$$\int_M f \, d\mu = \int_M \operatorname{Re} f \, d\mu + i \int_M \operatorname{Im} f \, d\mu.$$

Wir werden uns im Folgenden aber auf reellwertige Funktionen beschränken, auch wenn sich viele Resultate auf komplexwertige Funktionen übertragen lassen.

Es reicht, den Fall $M = X$ zu betrachten; vgl. Lemma 7.10. Statt $\int_X f \, d\mu$ schreiben wir auch $\int f \, d\mu$.

Das folgende Resultat ist analog zu Lemma 7.8.

Satz 8.2. *Seien (X, S, μ) Maßraum, $\alpha \in \mathbb{R}$ und $f, g\colon X \to \mathbb{R}$ integrierbar. Dann gilt:*

(i) *$f + g$ ist integrierbar und $\int (f + g) d\mu = \int f \, d\mu + \int g \, d\mu$.*

(ii) *αf ist integrierbar und $\int \alpha f \, d\mu = \alpha \int f \, d\mu$.*

(iii) *Ist $f \leq g$, so ist $\int f \, d\mu \leq \int g \, d\mu$.*

(iv) *$|f|$ ist integrierbar und $\left| \int f \, d\mu \right| \leq \int |f| \, d\mu$.*

Beweis. Wir beweisen nur die Aussage (i). Dazu notieren wir, dass

$$(f + g)^+ - (f + g)^- = f + g = f^+ - f^- + g^+ - g^-$$

und folglich

$$(f + g)^+ + f^- + g^- = (f + g)^- + f^+ + g^+.$$

Dies liefert mit Lemma 7.8, dass

$$\int (f+g)^+ \, d\mu + \int f^- \, d\mu + \int g^- \, d\mu = \int (f+g)^- \, d\mu + \int f^+ \, d\mu + \int g^+ \, d\mu.$$

Wegen $(f+g)^+ \le f^+ + g^+$ und $(f+g)^- \le f^- + g^-$ sind alle Integrale endlich. Hieraus folgt (i).

Die Beweise von (ii), (iii) und (iv) lassen wir aus. □

Satz 8.3. (Satz von Lebesgue; Satz über majorisierte Konvergenz) *Seien (X, S, μ) Maßraum und (f_n) eine Folge messbarer Funktionen von X nach $\mathbb{R}$, die (punktweise) gegen eine Funktion $f\colon X \to \mathbb{R}$ konvergieren. Es existiere eine integrierbare Funktion $F\colon X \to [0, \infty]$ mit $|f_n| \le F$ für alle n. Dann ist f integrierbar und es gilt*

$$\lim_{n\to\infty} \int |f_n - f| \, d\mu = 0 \quad \textit{und} \quad \int f \, d\mu = \lim_{n\to\infty} \int f_n \, d\mu.$$

Bemerkung 8.4. Ohne die Voraussetzung, dass eine Funktion F wie angegeben existiert, gilt der Satz nicht. Man betrachte etwa die Funktionen $f_n\colon [0,1] \to \mathbb{R}$, $f_n(x) = (n+1)(n+2)x^n(1-x)$. Benutzen wir bereits das später (Satz 9.1) bewiesene Ergebnis, dass das Lebesgue-Integral bzgl. des Lebesgue-Maßes λ über ein kompaktes Intervall mit dem Riemann-Integral übereinstimmt, falls Letzteres existiert, so erhalten wir $\int f_n \, d\lambda = \int_0^1 f_n(x) dx = 1$ für alle $n \in \mathbb{N}$, aber $f_n(x) \to 0$ für $n \to \infty$ und alle $x \in [0,1]$; vgl. Abbildung 3.

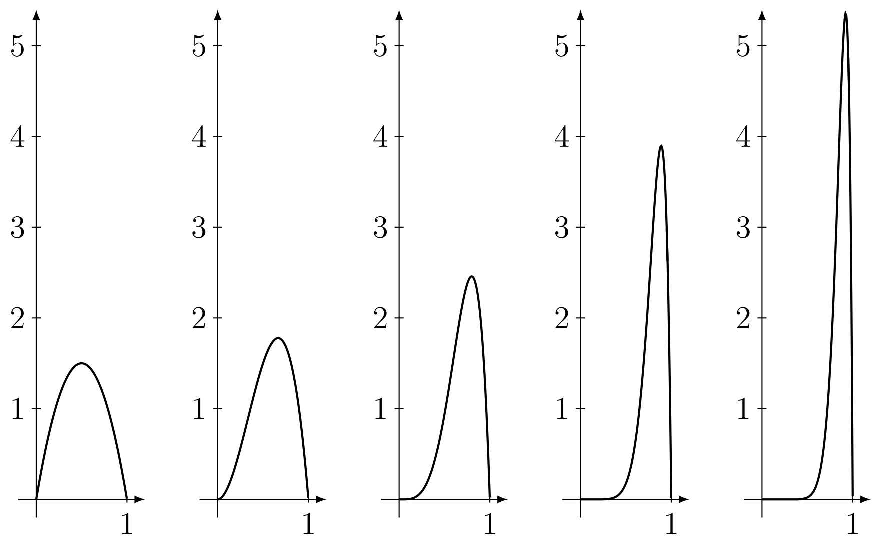

Abbildung 3: Die Funktionen f_1, f_2, f_4, f_8 und f_{12}.

Beweis von Satz 8.3. Die Messbarkeit von f folgt aus Satz 6.2 und die Integrierbarkeit der f_k und die von f folgen aus der Voraussetzung über F.

Wegen $0 \leq |f_n - f| \leq |f_n| + |f| \leq 2F$ folgt mit dem Lemma von Fatou, dass

$$\begin{aligned}\int 2F\,d\mu &= \int \lim_{n\to\infty}(2F - |f_n - f|)\,d\mu \\ &= \int \liminf_{n\to\infty}(2F - |f_n - f|)\,d\mu \\ &\leq \liminf_{n\to\infty} \int (2F - |f_n - f|)\,d\mu \\ &= \int 2F\,d\mu - \limsup_{n\to\infty} \int |f_n - f|\,d\mu\end{aligned}$$

und damit die erste Behauptung.

Die zweite folgt aus der ersten mit

$$\left|\int f_n\,d\mu - \int f\,d\mu\right| = \left|\int (f_n - f) d\mu\right| \leq \int |f_n - f|\ d\mu. \qquad \square$$

Nach Definition 2.4 heißt eine Teilmenge N von X Nullmenge, falls $\mu(N) = 0$ gilt. Wir sagen, dass eine Eigenschaft *fast überall* (genauer *μ-fast überall*) oder *für fast alle x* gilt, falls die Menge der x, für diese nicht gilt, eine Nullmenge ist. Beispielsweise sagen wir, dass $f = g$ fast überall gilt, falls die Menge $\{f \neq g\} = \{x \in X : f(x) \neq g(x)\}$ eine Nullmenge ist, also $\mu(\{f \neq g\}) = 0$.

Satz 8.5. *Seien (X, S, μ) Maßraum und $f : X \to \overline{\mathbb{R}}$ integrierbar. Dann gilt:*

(i) *$|f| < \infty$ fast überall.*

(ii) *Ist $\int |f|\,d\mu = 0$, so ist $f = 0$ fast überall.*

(iii) *Ist $g : X \to \overline{\mathbb{R}}$ messbar und $g = f$ fast überall, so ist g integrierbar und $\int g\,d\mu = \int f\,d\mu$.*

Beweis. (i) Für alle $n \in \mathbb{N}$ ist $n\chi_{\{|f|=\infty\}} \leq |f|$, also

$$n\mu(\{|f| = \infty\}) = \int n\chi_{\{|f|=\infty\}}\,d\mu \leq \int |f|\,d\mu < \infty.$$

Hieraus folgt $\mu(\{|f| = \infty\}) = 0$.

(ii) Sei $M_n = \{|f| \geq 1/n\}$. Dann ist

$$\frac{1}{n}\mu(M_n) = \int_{M_n} \frac{1}{n} d\mu \leq \int_{M_n} |f|\,d\mu \leq \int |f|\,d\mu = 0,$$

also $\mu(M_n) = 0$ und damit $\mu(\{f \neq 0\}) = \mu(\bigcup_{n=1}^{\infty} M_n) = 0$.

(iii) Es ist $N = \{f \neq g\}$ Nullmenge und damit

$$\begin{aligned}\int g^\pm \, d\mu &= \int_{X\setminus N} g^\pm \, d\mu + \int_N g^\pm \, d\mu \\ &= \int_{X\setminus N} g^\pm \, d\mu \\ &= \int_{X\setminus N} f^\pm \, d\mu \\ &= \int_{X\setminus N} f^\pm \, d\mu + \int_N f^\pm \, d\mu \\ &= \int f^\pm \, d\mu < \infty.\end{aligned}$$

Damit ist g integrierbar und $\int g \, d\mu = \int f \, d\mu$. □

Satz 8.5 implizert, dass es für die Integrationstheorie meist irrelevant ist, ob wir Funktionen $f\colon X \to \mathbb{R}$ oder $f\colon X \to \overline{\mathbb{R}}$ betrachten. Im integrierbaren Falle ist $\{f = \infty\}$ eine Nullmenge und Abänderung der Funktion auf dieser Menge ändert das Integral nicht. Es ist also keine Einschränkung, nur Funktionen nach $\mathbb{R}$ zu betrachten.

So gilt der Satz über majorisierte Konvergenz auch dann, wenn man dort Funktionen $f_n\colon X \to \overline{\mathbb{R}}$ und $f\colon X \to \overline{\mathbb{R}}$ zulässt und nur $f_n \to f$ und $|f_n| \leq F$ fast überall fordert. Denn durch Abänderung der f_n und f auf einer Nullmenge kann man erreichen, dass die Voraussetzungen in der oben angegebenen Fassung gelten. Eine analoge Bemerkung gilt für viele andere Sätze.

9 Vergleich von Riemann- und Lebesgue-Integral

Wir wollen das hier eingeführte Lebesgue-Integral (bzgl. des Lebesgue-Maßes λ) und das in Analysis II betrachtete Riemann-Integral miteinander vergleichen.

Satz 9.1. *Seien $a, b \in \mathbb{R}$ mit $a < b$ und $f\colon [a,b] \to \mathbb{R}$. Dann gilt:*

(i) *Ist f Riemann-integrierbar, so ist f auch Lebesgue-integrierbar, und Riemann- und Lebesgue-Integral von f über $[a,b]$ stimmen überein.*

(ii) *f ist genau dann Riemann-integrierbar, wenn f beschränkt ist und die Menge der Unstetigkeitsstellen von f eine Lebesgue-Nullmenge ist.*

Beweis. (i) Sei f Riemann-integrierbar. Dann ist f beschränkt. Für $n \in \mathbb{N}$ und $0 \leq j \leq 2^n$ zerlegen die Punkte $x_{n,j} = a + j(b-a)/2^n$ das Intervall $[a,b]$ in 2^n disjunkte Intervalle $I_{n,1} = [x_{n,0}, x_{n,1}]$ und $I_{n,1} = (x_{n,j-1}, x_{n,j}]$ für $2 \leq j \leq 2^n$. Es ist $\lambda(I_{n,j}) = x_{n,j} - x_{n,j-1} = (b-a)/2^n$ die Länge von $I_{n,j}$. Mit

$$\underline{m}_{n,j} = \inf_{x_{n,j-1} \leq x \leq x_{n,j}} f(x) \quad \text{und} \quad \overline{m}_{n,j} = \inf_{x_{n,j-1} \leq x \leq x_{n,j}} f(x)$$

sind dann durch

$$\underline{f}_n = \sum_{j=1}^{2^n} \underline{m}_{n,j} \chi_{I_{n,j}} \quad \text{und} \quad \overline{f}_n = \sum_{j=1}^{2^n} \overline{m}_{n,j} \chi_{I_{n,j}}$$

Treppenfunktionen mit $\underline{f}_n \leq f \leq \overline{f}_n$ gegeben und

$$\underline{S}_n(f) = \sum_{j=1}^{2^n} \underline{m}_{n,j}\lambda(I_{n,j}) = \int \underline{f}_n \, d\lambda \quad \text{und} \quad \overline{S}_n(f) = \sum_{j=1}^{2^n} \overline{m}_{n,j}\lambda(I_{n,j}) = \int \overline{f}_n \, d\lambda$$

sind die Riemannsche Unter- und Obersumme. Nun ist $\left(\underline{f}_n\right)$ monoton steigend und $\left(\overline{f}_n\right)$ monoton fallend. Damit sind

$$\underline{f} = \lim_{n\to\infty} \underline{f}_n \quad \text{und} \quad \overline{f} = \lim_{n\to\infty} \overline{f}_n$$

messbar und, da f beschränkt ist, auch Lebesgue-integrierbar. Da f Riemann-integrierbar ist, folgt

$$\int \underline{f} \, d\lambda = \lim_{n\to\infty} \int \underline{f}_n \, d\lambda = \lim_{n\to\infty} \underline{S}_n(f) = \int_a^b f(x)dx,$$

wobei die rechte Seite das Riemann-Integral bezeichnet, und ebenso

$$\int \overline{f} \, d\lambda = \lim_{n\to\infty} \int \overline{f}_n \, d\lambda = \lim_{n\to\infty} \overline{S}_n(f) = \int_a^b f(x)dx.$$

Es folgt

$$\int \underline{f} \, d\lambda = \int \overline{f} \, d\lambda, \quad \text{also} \int \left(\overline{f} - \underline{f}\right) \, d\lambda = 0,$$

woraus wegen $\underline{f} \leq f \leq \overline{f}$ mit Satz 8.5, (ii) folgt, dass $\underline{f} = \overline{f}$ fast überall und damit auch $f = \underline{f} = \overline{f}$ fast überall. Mit Satz 8.5, (iii) folgt jetzt die Behauptung.

(ii) Sei zunächst f Riemann-integrierbar. Wie in (i) gezeigt wurde, ist – mit den dortigen Bezeichnungen – die Menge $N := \{\underline{f} \neq \overline{f}\}$ eine Nullmenge. Die Menge $E = \{x_{n,j} \colon n \in \mathbb{N}, 0 \leq j \leq 2^n\}$ der Endpunkte der Teilintervalle ist abzählbar und daher ebenfalls Nullmenge. Damit ist auch $N \cup E$ Nullmenge.

Wir zeigen, dass f in $[a, b] \setminus (N \cup E)$ stetig ist. Sei dazu $x \in [a, b] \setminus (N \cup E)$ und $\varepsilon > 0$. Dann existiert $n \in \mathbb{N}$ mit

$$f(x) - \varepsilon < \underline{f}_n(x) \leq \overline{f}_n(x) < f(x) + \varepsilon.$$

Nun existiert $j \in \{1, \ldots, 2^n\}$ mit $x_{n,j-1} < x < x_{n,j}$. Es folgt, dass

$$f(x) - \varepsilon < \underline{f}_n(x) = \underline{f}_n(y) \leq f(y) \leq \overline{f}_n(y) = \overline{f}_n(x) < f(x) + \varepsilon$$

für $x_{n,j-1} < y < x_{n,j}$, also $|f(y) - f(x)| < \varepsilon$ für y in einem Intervall um x. Damit ist f stetig in x.

Sei umgekehrt f beschränkt. Ist f stetig in $x \in [a, b]$, so folgt $\underline{f}(x) = \overline{f}(x)$. Ist die Menge der Unstetigkeitsstellen eine Nullmenge, gilt also $f = \underline{f} = \overline{f}$ fast überall. Damit folgt die Riemann-Integrierbarkeit wie in (i). □

Wir benutzen im Folgenden auch für das Lebesgue-Integral die Schreibweise $\int_a^b f(x)dx$ statt $\int_{[a,b]} f \, d\lambda$.

Nach Übung gilt Satz 9.1, (i), auch für uneigentliche Riemann-Integrale von Funktionen nach $[0, \infty]$. Satz 9.1, (i), gilt aber *nicht* für uneigentliche Riemann-Integrale von Funktionen nach $\mathbb{R}$. Denn wie nach Definition 8.1 bemerkt ist eine messbare Funktion f genau dann Lebesgue-integrierbar, wenn $|f|$ Lebesgue-integrierbar ist. Die entsprechende Aussage für uneigentliche Riemann-Integrierbarkeit gilt aber nicht; vgl. Bemerkung 6.44 der Analysis-Vorlesung von Herrn Nieß. Ein konkretes Beispiel ist das konvergente uneigentliche Riemann-Integral

$$\int_1^\infty \frac{\sin x}{x} dx.$$

Im Lebesgueschen Sinne existiert dieses Integral nicht, da sonst auch $|(\sin x)/x|$ Lebesgue-integrierbar über $[1, \infty)$ wäre. Es gilt aber

$$\int_1^\infty \left|\frac{\sin x}{x}\right| dx = \infty,$$

sowohl im Riemannschen Sinne wie auch in dem von Definition 7.7.

Satz 9.2. *Sei $f\colon [a, b] \to \mathbb{R}$ differenzierbar. Ist f' beschränkt, so ist f' Lebesgue-integrierbar und es gilt*

$$f(x) - f(a) = \int_a^x f'(t)dt.$$

Bemerkung 9.3. Man vergleiche die Aussage mit dem Hauptsatz der Differential- und Integralrechnung. Dieser liefert die Aussage, falls f' stetig ist.

Beweis. Es gelte $|f'(x)| \le M$ für alle $x \in [a, b]$. Für $k \in \mathbb{N}$ sei $f_k\colon [a, b] \to \mathbb{R}$,

$$f_k(x) = \begin{cases} k\left(f\left(x + \frac{1}{k}\right) - f(x)\right), & a \le x \le b - \frac{1}{k}, \\ 0, & b - \frac{1}{k} < x \le b. \end{cases}$$

Dann ist f_k integrierbar.

Nach Mittelwertsatz gilt

$$|f_k(x)| \le M$$

für alle $x \in [a, b]$. Außerdem gilt

$$f_k(x) \to f'(x)$$

für alle $x \in [a, b)$. Nach dem Satz über majorisierte Konvergenz folgt, dass f' integrierbar ist und dass

$$\begin{aligned} \int_a^b f'(x)dx &= \lim_{k\to\infty} \int_a^b f_k(x)dx \\ &= \lim_{k\to\infty} \int_a^{b-1/k} k\left(f\left(x + \tfrac{1}{k}\right) - f(x)\right) dx \\ &= \lim_{k\to\infty} \left(k \int_{a+1/k}^b f(x)dx - k \int_a^{b-1/k} f(x)dx \right) \\ &= \lim_{k\to\infty} \left(k \int_{b-1/k}^b f(x)dx - k \int_a^{a+1/k} f(x)dx \right). \end{aligned}$$

Nun ist f stetig und damit gilt

$$\left| k \int_a^{a+1/k} f(x)dx - f(a) \right| = \left| k \int_a^{a+1/k} (f(x) - f(a))dx \right| \le \max_{a \le x \le a+1/k} |f(x) - f(a)|.$$

Es folgt, dass

$$k \int_a^{a+1/k} f(x)dx \to f(a).$$

Analog erhält man

$$k \int_{b-1/k}^{b} f(x)dx \to f(b).$$

Insgesamt folgt

$$\int_a^b f'(x)dx = f(b) - f(a).$$

Dies gilt auch, wenn man b durch $x \in [a, b]$ ersetzt. □

10 Parameterabhängige Integrale

Sei X metrischer Raum, (Y, S, μ) Maßraum und $f\colon X \times Y \to \mathbb{R}$. Wir nehmen an, dass für jedes $x \in X$ die durch $y \mapsto f(x, y)$ definierte Funktion integrierbar ist. Dann kann eine Funktion $F\colon X \mapsto \mathbb{R}$,

$$F(x) = \int_Y f(x, y)d\mu(y)$$

definiert werden. (Wir schreiben $d\mu(y)$, um zu verdeutlichen, dass die Integration bezüglich des Maßes μ sich auf die Variable y bezieht.) Wir wollen untersuchen, inwieweit sich die Stetigkeit oder Differenzierbarkeit von f auf F überträgt.

Im Allgemeinen folgt aus der Stetigkeit von f nicht die von F. Etwa für $X = Y = \mathbb{R}$ und $f(x, y) = x \exp\left(-(xy)^2\right)$ folgt für $x > 0$, dass

$$F(x) = \int_{-\infty}^{\infty} x \exp\left(-(xy)^2\right) dy = \int_{-\infty}^{\infty} e^{-t^2} dt = \sqrt{\pi}.$$

Es folgt, dass

$$F(x) = \begin{cases} \sqrt{\pi}, & x > 0, \\ 0, & x = 0, \\ -\sqrt{\pi}, & x < 0. \end{cases}$$

Damit ist F unstetig.

Es gilt aber folgendes Resultat.

Satz 10.1. *Seien $f\colon X \times Y \to \mathbb{R}$ und $F\colon X \to \mathbb{R}$ wie oben. Es gelte:*

(i) *Für alle $y \in Y$ ist $x \mapsto f(x, y)$ stetig.*

(ii) *Es existiert eine integrierbare Funktion $\Phi\colon Y \to [0, \infty]$ so dass $|f(x, y)| \le \Phi(y)$ für alle $x \in X$ und $y \in Y$.*

Dann ist F stetig.

Beweis. Wir müssen zeigen, dass aus $x_k \to x_0$ folgt, dass $F(x_k) \to F(x_0)$ gilt.

Sei also (x_k) Folge in X mit $x_k \to x_0 \in X$. Sei $f_k \colon Y \to \mathbb{R}$, $f_k(y) = f(x_k, y)$. Da die f_k integrierbar sind und $|f_k| \leq \Phi$ gilt, folgt mit dem Satz über majorisierte Konvergenz und $f_0 \colon Y \to \mathbb{R}$, $f_0(y) = f(x_0, y)$, dass

$$F(x_k) = \int_Y f(x_k, y) d\mu(y) = \int_Y f_k d\mu \to \int_Y f_0 d\mu = \int_Y f(x_0, y) d\mu = F(x_0). \qquad \square$$

Satz 10.2. *Seien $X \subseteq \mathbb{R}^d$ offen und $f \colon X \times Y \to \mathbb{R}$ sowie $F \colon X \to \mathbb{R}$ wie vorher. Es gelte:*

(i) *Für alle $y \in Y$ ist $x \mapsto f(x, y)$ stetig differenzierbar (oder äquivalent dazu: alle partiellen Ableitungen $\partial f / \partial x_k$ existieren und sind stetig in X).*

(ii) *Es existiert eine integrierbare Funktion $\Phi \colon Y \to [0, \infty)$, so dass*

$$\left| \frac{\partial f}{\partial x_k}(x, y) \right| \leq \Phi(y)$$

für alle $x \in X$, $y \in Y$ und $k \in \{1, \ldots, d\}$.

Dann ist F stetig differenzierbar.

Außerdem ist die durch

$$y \mapsto \frac{\partial f}{\partial x_k}(x, y)$$

gegebene Funktion von Y nach $\mathbb{R}$ für alle $x \in X$ und alle $k \in \{1, \ldots, d\}$ integrierbar und es gilt

$$\frac{\partial F}{\partial x_k}(x, y) = \int_Y \frac{\partial f}{\partial x_k}(x, y) d\mu(y).$$

Beweis. Sei $\xi \in X$ und e_k der k-te Einheitsvektor. Sei weiter (h_j) Nullfolge in $\mathbb{R} \setminus \{0\}$ und $g_j \colon Y \to \mathbb{R}$,

$$g_j(y) = \frac{f(\xi + h_j e_k, y) - f(\xi, y)}{h_j}.$$

Dann gilt

$$g_j(y) \to \frac{\partial f}{\partial x_k}(\xi, y)$$

für $j \to \infty$ und alle $y \in Y$. Aus dem Mittelwertsatz folgt $|g_j| \leq \Phi$, vgl. den Beweis von Satz 9.2. Damit können wir den Satz über majorisierte Konvergenz auf die Folge (g_j) anwenden. Es folgt, dass $y \mapsto \partial f / \partial x_k(\xi, y)$ integrierbar ist und

$$\int_Y g_j(y) d\mu(y) \to \int_Y \frac{\partial f}{\partial x_k}(\xi, y) d\mu(y).$$

Andererseits gilt

$$\int_Y g_j(y) d\mu(y) = \frac{F(\xi + h_j e_k) - F(\xi)}{h_j}.$$

Es folgt die partielle Differenzierbarkeit von F und die Formel für $\partial F / \partial x_k$. Die Stetigkeit der partiellen Ableitungen folgt aus Satz 10.1. $\square$

Beispiel 10.3. Sei $F\colon (0,\infty) \to \mathbb{R}$,

$$F(x) = \int_0^1 \frac{y^x - 1}{\log y} dy.$$

Das Integral existiert, da $e^t \geq 1 + t$ für $t \in \mathbb{R}$ und damit

$$0 \leq \frac{y^x - 1}{\log y} = \frac{1 - y^x}{-\log y} = \frac{1 - e^{x \log y}}{-\log y} \leq \frac{-x \log y}{-\log y} = x$$

für $0 < y < 1$. Wegen

$$\frac{\partial}{\partial x} \frac{y^x - 1}{\log y} = y^x \leq 1$$

folgt aus Satz 10.2, dass F differenzierbar ist, mit

$$F'(x) = \int_0^1 y^x dy = \frac{1}{1+x}.$$

Obige Abschätzung des Integranden liefert auch $\lim_{x \to 0} F(x) = 0$ und damit

$$F(x) = \int_0^x F'(t) dt = \int_0^x \frac{dt}{1+t} = \log(1+x).$$

11 L^p-Räume

Sei (X, S, μ) Maßraum. Für eine messbare Funktion $f\colon X \to \mathbb{R}$ setzen wir

$$\|f\|_1 = \int |f| \, d\mu.$$

Damit ist f genau dann integrierbar, wenn $\|f\|_1 < \infty$. Sei

$$\mathcal{L}^1(\mu) := \mathcal{L}^1(X, S, \mu) := \{f\colon X \to \mathbb{R}\colon f \text{ integrierbar}\}.$$

Nach Satz 8.2 ist $\mathcal{L}^1(\mu)$ ein Vektorraum und für $f, g \in \mathcal{L}^1(\mu)$ und $\alpha \in \mathbb{R}$ folgt

$$\|\alpha f\|_1 = |\alpha| \cdot \|f\|_1 \quad \text{und} \quad \|f + g\|_1 \leq \|f\|_1 + \|g\|_1.$$

Durch $\|\cdot\|_1\colon \mathcal{L}^1(\mu) \to [0,\infty)$ ist aber keine Norm gegeben, denn aus $\|f\|_1 = 0$ folgt nicht, dass $f = 0$, d.h., dass $f(x) = 0$ für alle $x \in X$. Nach Satz 8.5 gilt dann aber $f(x) = 0$ für fast alle $x \in X$.

Nun ist durch

$$f \sim g \quad :\Leftrightarrow \quad f = g \text{ fast überall}$$

eine Äquivalenzrelation auf $\mathcal{L}^1(\mu)$ gegeben und

$$U = \{f \in \mathcal{L}^1(\mu)\colon f(x) = 0 \text{ fast überall}\} = \{f \in \mathcal{L}^1(\mu)\colon \|f\|_1 = 0\}$$

ist Unterraum von $\mathcal{L}^1(\mu)$. Wir betrachten den Quotientenraum

$$L^1(\mu) := \mathcal{L}^1(\mu)/U.$$

Sei $[f] \in L^1(\mu)$ die Äquivalenzklasse von $f \in \mathcal{L}^1(\mu)$. Aus $f \sim g$, also $[f] = [g]$ folgt dann mit Satz 8.5, (iii), dass $\|f\|_1 = \|g\|_1$. Wir können also eine Abbildung $\|\cdot\|_1\colon L^1(\mu) \to \mathbb{R}$, $\|[f]\|_1 = \|f\|_1$, definieren. Nach Konstruktion gilt $\|[f]\|_1 = 0$ genau dann, wenn $[f] = [0]$. Die anderen Normeigenschaften für die Abbildung $\|\cdot\|_1\colon L^1(\mu) \to \mathbb{R}$ folgen aus denen von $\|\cdot\|_1\colon \mathcal{L}^1(\mu) \to \mathbb{R}$.

Damit erhalten wir folgendes Ergebnis.

Satz 11.1. *$(L^1(\mu), \|\cdot\|_1)$ ist normierter Raum.*

Es ist allgemein üblich, die Äquivalenzklasse von f nicht mit $[f]$, sondern wieder mit f zu bezeichnen.

Statt $\|\cdot\|_1$ betrachtet man allgemeiner für $p > 1$ und messbares $f\colon X \to \mathbb{R}$

$$\|f\|_p = (\| \, |f|^p \, \|_1)^{1/p}.$$

Auch hier folgt aus $\|f\|_p = 0$ nicht $f = 0$, aber dies lässt sich wie im Fall $p = 1$ beheben. Von den weiteren Eigenschaften einer Norm ist $\|\alpha \cdot f\|_p = |\alpha| \cdot \|f\|_p$ für $\alpha \in \mathbb{R}$ klar, die Dreiecksungleichung aber zunächst nicht. Dazu zeigen wir folgendes Lemma.

Lemma 11.2. *Seien $a, b \geq 0$ und $p, q > 1$ mit $\frac{1}{p} + \frac{1}{q} = 1$. Dann gilt*

$$a \cdot b \leq \frac{1}{p} a^p + \frac{1}{q} b^q.$$

Beweis. Wir können $a, b > 0$ annehmen, da sonst die Behauptung trivial ist. Sei $x = \log a$ und $y = \log b$. Mit der Konvexität der Exponentialfunktion folgt

$$a \cdot b = e^x \cdot e^y = e^{x+y} = \exp\left(\frac{1}{p}\,px + \frac{1}{q}\,qy\right) \leq \frac{1}{p}\exp(px) + \frac{1}{q}\exp(qy) = \frac{1}{p}\,a^p + \frac{1}{q}\,b^q$$

wie behauptet. □

Mit Hilfe dieses Lemmas erhalten wir zunächst folgendes Resultat.

Satz 11.3. (Höldersche Ungleichung) *Seien $f, g\colon X \to \mathbb{R}$ messbar sowie $p, q > 1$ mit*

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Dann gilt

$$\|f \cdot g\|_1 \leq \|f\|_p \cdot \|g\|_q$$

Beweis. Gilt $\|f\|_p = 0$, so ist $f = 0$ fast überall, also $f \cdot g = 0$ fast überall und damit $\|f \cdot g\|_1 = 0$. Analoges gilt, falls $\|g\|_q = 0$. Wir können also $\|f\|_p, \|g\|_q \neq 0$ annehmen. Ebenso können wir $\|f\|_p, \|g\|_q \neq \infty$ annehmen.

Sei

$$F = \frac{f}{\|f\|_p} \quad \text{und} \quad G = \frac{g}{\|g\|_q}.$$

Dann gilt $\|F\|_p = \|G\|_q = 1$.

Nach Lemma 11.2 gilt

$$|F(x) \cdot G(x)| \leq \frac{1}{p}|F(x)|^p + \frac{1}{q}|G(x)|^q.$$

Es folgt

$$\|F \cdot G\|_1 \leq \frac{1}{p}\|F\|_p^p + \frac{1}{q}\|G\|_q^q = 1$$

und damit die Behauptung. □

Die gewünschte Dreiecksungleichung liefert nun folgender Satz.

Satz 11.4. **(Minkowskische Ungleichung)** *Seien $f, g\colon X \to \mathbb{R}$ messbar und $p > 1$. Dann gilt*

$$\|f+g\|_p \leq \|f\|_p + \|g\|_p.$$

Beweis. Sei q gemäß der Hölderschen Ungleichung gewählt, also $1/p + 1/q = 1$. Dann gilt

$$\begin{aligned}
\|f+g\|_p^p &= \| \, |f+g|^p \, \|_1 \\
&\leq \left\| \, |f| \cdot |f+g|^{p-1} + |g| \cdot |f+g|^{p-1} \right\|_1 \\
&\leq \left\| \, |f| \cdot |f+g|^{p-1} \right\|_1 + \left\| \, |g| \cdot |f+g|^{p-1} \right\|_1 \\
&\leq \|f\|_p \left\| \, |f+g|^{p-1} \right\|_q + \|g\|_p \left\| \, |f+g|^{p-1} \right\|_q \\
&= (\|f\|_p + \|g\|_p) \left\| \, |f+g|^{p-1} \right\|_q .
\end{aligned}$$

Weiter folgt wegen $(p-1)q = p$ und damit $1/q = (p-1)/p$, dass

$$\left\| \, |f+g|^{p-1} \right\|_q = \left(\left\| \, |f+g|^{(p-1)q} \right\|_1 \right)^{\frac{1}{q}} = \| \, |f+g|^p \|_1^{\frac{p-1}{p}} = \|f+g\|_p^{p-1}.$$

Die Behauptung folgt. □

Sei nun $\mathcal{L}^p(\mu)$ die Menge aller messbaren Funktionen $f\colon X \to \mathbb{R}$, für die $\|f\|_p < \infty$ gilt. Dann ist $\mathcal{L}^p(\mu)$ ein Vektorraum. Wie oben setzt man

$$L^p(\mu) = \mathcal{L}^p(\mu) / \{f \in \mathcal{L}^p(\mu) \colon \|f\|_p = 0\}.$$

De facto identifiziert man $L^p(\mu)$ und $\mathcal{L}^p(\mu)$ wieder miteinander, d.h., die Äquivalenzklasse von $f \in \mathcal{L}^p(\mu)$ in $L^p(\mu)$ wird wieder mit f bezeichnet.

Analog zu Satz 11.1 gilt – mit der offensichtlichen Definition der Norm $\|\cdot\|_p$ auf $L^p(\mu)$ – folgender Satz.

Satz 11.5. *$(L^p(\mu), \|\cdot\|_p)$ ist normierter Raum.*

Für $p = \infty$ geht man wie folgt vor: Eine Funktion $f\colon X \to \mathbb{R}$ heißt *wesentlich beschränkt*, falls $M \in \mathbb{R}$ existiert, so dass $\{x \in X \colon |f(x)| > M\}$ Nullmenge ist. Das Infimum über alle diese M wird mit $\|f\|_\infty$ bezeichnet. Mit $\mathcal{L}^\infty(\mu)$ bezeichnet man den Vektorraum der wesentlich beschränkten messbaren Funktionen $f\colon X \to \mathbb{R}$. Wie oben erhält man hieraus einen normierten Raum $(L^\infty(\mu), \|\cdot\|_\infty)$.

Beispiel. Die durch $x \mapsto 1/x^\alpha$ gegebene Funktion ist genau dann in $L^p((1,\infty))$ bezüglich λ, falls

$$\int_{(1,\infty)} \left| \frac{1}{(x^\alpha)^p} \right| d\lambda(x) = \int_1^\infty \frac{dx}{x^{\alpha p}} < \infty.$$

Dies ist der Fall, wenn $\alpha p > 1$, also $p > 1/\alpha$. Analog ist diese Funktion genau dann in $L^p((0,1))$ ist, wenn $p < 1/\alpha$ gilt.

Satz 11.6. *Für alle $p \in [1, \infty]$ ist $(L^p(\mu), \|\cdot\|_p)$ vollständig.*

Wir zeigen ein allgemeineres Resultat. Dabei heißt eine Folge (f_k) von Funktionen von X nach $\mathbb{R}$ eine L^p-*Cauchyfolge*, falls zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ existiert, so dass $\|f_m - f_n\|_p < \varepsilon$ für $m, n \geq N$.

Satz 11.7. (Satz von Riesz-Fischer) *Sei (f_k) eine L^p-Cauchyfolge in $\mathcal{L}^p(\mu)$. Dann existiert $f \in \mathcal{L}^p(\mu)$ mit*

$$\lim_{k\to\infty} \|f_k - f\|_p = 0.$$

Außerdem existiert eine Teilfolge (f_{k_j}) von (f_k) mit $f_{k_j}(x) \to f(x)$ fast überall.

Bemerkung 11.8. (i) Sei $([f_k])$ Cauchyfolge in $(L^p(\mu), \|\cdot\|_p)$. Dann ist der Satz von Riesz-Fischer auf die Folge (f_k) anwendbar. Ist f wie in diesem Satz gewählt, so ist $[f]$ Grenzwert von $([f_k])$ in $(L^p(\mu), \|\cdot\|_p)$. Dies liefert Satz 11.6, d.h., die Vollständigkeit von $(L^p(\mu), \|\cdot\|_p)$.

(ii) Im Allgemeinen gilt nicht $f_k(x) \to f(x)$ fast überall. Ein Gegenbeispiel erhält man wie folgt: Für $k \in \mathbb{N}$ sei $n \in \mathbb{N}_0$ mit $2^n \leq k \leq 2^{n+1} - 1$. Weiter sei

$$I_k = \left[\frac{k}{2^n} - 1, \frac{k+1}{2^n} - 1\right]$$

und $f_k\colon [0,1] \to \mathbb{R}$, $f_k = \chi_{I_k}$. Es ist also f_k die charakteristische Funktion eines Teilintervalls von $[0,1]$ der Länge $1/2^n$.

Es gilt $\|f_k\|_1 = 1/2^n \to 0$, also $\|f_k - f\|_1 \to 0$ für $f = 0$, aber für $x \in [0,1]$ und $n \in \mathbb{N}_0$ existiert k mit $2^n \leq k \leq 2^{n+1} - 1$ und $f_k(x) = 1$. Also existiert kein $x \in [0,1]$ mit $f_k(x) \to f(x) = 0$.

Beweis von Satz 11.7. Der Einfachheit halber beschränken wir uns auf den Fall $p = 1$. Sei (k_j) wachsende Folge mit

$$\|f_k - f_{k_j}\|_1 < \frac{1}{2^j} \quad \text{für } k \geq k_j.$$

Wir zeigen zunächst, dass (f_{k_j}) fast überall konvergiert.

Sei dazu $g_j = f_{k_{j+1}} - f_{k_j}$. Dann gilt $\|g_j\|_1 < 2^{-j}$ und mit Korollar 7.11 folgt

$$\int \left(\sum_{j=1}^{\infty} |g_j|\right) d\mu = \sum_{j=1}^{\infty} \int |g_j|\, d\mu = \sum_{j=1}^{\infty} \|g_j\|_1 < \sum_{j=1}^{\infty} \frac{1}{2^j} = 1.$$

Sei $h\colon X \to [0,\infty]$,

$$h(x) = \sum_{k=1}^{\infty} |g_j(x)| = \sum_{k=1}^{\infty} |f_{k_{j+1}}(x) - f_{k_j}(x)|.$$

Dann ist h messbar und $\|h\|_1 = \int h\, d\mu \leq 1$. Nach Satz 8.5 existiert damit eine Nullmenge N mit $h(x) \neq \infty$ für $x \notin N$.

Wir definieren nun

$$f(x) = \begin{cases} f_{k_1}(x) + \sum_{j=1}^{\infty} \left(f_{k_{j+1}}(x) - f_{k_j}(x)\right), & x \notin N, \\ 0, & x \in N. \end{cases}$$

Für $x \notin N$ gilt also

$$f(x) = \lim_{n\to\infty} \left(f_{k_1}(x) + \sum_{j=1}^{n-1} \left(f_{k_{j+1}}(x) - f_{k_j}(x) \right) \right) = \lim_{n\to\infty} f_{k_n}(x).$$

Nach dem Satz über majorisierte Konvergenz ist f messbar (denn die Teilsummen der Reihe für f sind durch $|f_{k_1}| + h$ beschränkt). Sei $\varepsilon > 0$. Wir wählen $j \in \mathbb{N}$ mit $1/2^j < \varepsilon$. Für $k \geq k_j$ folgt dann mit dem Lemma von Fatou, dass

$$\begin{aligned} \|f - f_k\|_1 &= \int |f - f_k| \, d\mu \\ &= \int \lim_{j\to\infty} |f_{k_j} - f_k| \, d\mu \\ &\leq \liminf_{j\to\infty} \int |f_{k_j} - f_k| \, d\mu \\ &= \liminf_{j\to\infty} \|f_{k_j} - f_k\|_1 \\ &\leq \varepsilon. \end{aligned}$$

Dies liefert $\|f - f_k\|_1 \to 0$ für $k \to \infty$. □

12 Integration auf Produkträumen

Wir wollen ein Integral bezüglich des Lebesgue-Maßes λ^2 auf $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ durch zweimalige Integration bezüglich des Lebesgue-Maßes λ auf $\mathbb{R}$ berechnen. Zunächst müssen wir dafür die σ-Algebra $\mathcal{B}^2$ mit der Menge $\mathcal{B} \times \mathcal{B}$ in Verbindung setzen. Dies führt zu folgender Definition.

Definition 12.1. Seien (X_1, S_1) und (X_2, S_2) Messräume. Die von den Mengen $A_1 \times A_2$, mit $A_j \in S_j$, erzeugte σ-Algebra heißt *Produkt-σ-Algebra* und wird mit $S_1 \otimes S_2$ bezeichnet.

Sei $\pi_j \colon X_1 \times X_2 \to X_j$, $(x_1, x_2) \to x_j$, die Projektion auf die j-te Komponente.

Lemma 12.2. *Seien (X, S), (X_1, S_1) und (X_2, S_2) Messräume. Eine Funktion $f \colon (X, S) \to (X_1 \times X_2, S_1 \otimes S_2)$, ist genau dann messbar, wenn die Funktionen $\pi_j \circ f \colon (X, S) \to (X_j, S_j)$ messbar sind.*

Beweis. Für $A_1 \in S_1$ ist $\pi_1^{-1}(A_1) = A_1 \times X_2 \in S_1 \otimes S_2$. Damit ist π_1 und analog auch π_2 messbar. Ist f messbar, so folgt die Messbarkeit von $\pi_j \circ f$ aus Satz 5.5.

Sei umgekehrt $\pi_j \circ f$ messbar, für $j = 1, 2$. Um die Messbarkeit von f zu zeigen, reicht es nach Satz 5.2 zu zeigen, dass $f^{-1}(B)$ für alle B aus einem Erzeuger von $S_1 \otimes S_2$. Damit genügt es, $f^{-1}(A_1 \times A_2) \in S$ für alle $A_j \in S_j$ zu zeigen. Dies folgt aber, da

$$\begin{aligned} f^{-1}(A_1 \times A_2) &= f^{-1}((A_1 \times X_2) \cap (X_1 \times A_2)) \\ &= f^{-1}(\pi_1^{-1}(A_1)) \cap \pi_2^{-1}(A_2)) \\ &= (f^{-1} \circ \pi_1^{-1})(A_1) \cap (f^{-1} \circ \pi_2^{-1})(A_2) \\ &= (\pi_1 \circ f)^{-1}(A_1) \cap (\pi_2 \circ f)^{-1}(A_2). \end{aligned}$$

□

Beispiel 12.3. Für $p, q \in \mathbb{N}$ gilt $\mathcal{B}^p \otimes \mathcal{B}^q = \mathcal{B}^{p+q}$. Die Inklusion $\mathcal{B}^p \otimes \mathcal{B}^q \supseteq \mathcal{B}^{p+q}$ folgt hier, da $\mathcal{B}^{p+q}$ von den Quadern in $\mathbb{R}^{p+q}$ erzeugt wird, und diese sind auch in $\mathcal{B}^p \otimes \mathcal{B}^q$.

Für die umgekehrte Inklusion betrachten wir die identische Abbildung

$$\mathrm{id}\colon (\mathbb{R}^{p+q}, \mathcal{B}^{p+q}) \to (\mathbb{R}^p \times \mathbb{R}^q, \mathcal{B}^p \otimes \mathcal{B}^q).$$

Die Inklusion $\mathcal{B}^p \otimes \mathcal{B}^q \supseteq \mathcal{B}^{p+q}$ ist äquivalent zur Messbarkeit dieser Abbildung. Nach Lemma 12.2 ist dies wiederum äquivalent zur Messbarkeit der Projektionen $\pi_1\colon (\mathbb{R}^{p+q}, \mathcal{B}^{p+q}) \to (\mathbb{R}^p, \mathcal{B}^p)$ und $\pi_2\colon (\mathbb{R}^{p+q}, \mathcal{B}^{p+q}) \to (\mathbb{R}^q, \mathcal{B}^q)$. Diese folgt aber wegen Satz 5.3 aus der Stetigkeit der Projektionen.

Ziel ist es nun, zu gegebenen Maßen μ_j auf den Messräumen (X_j, S_j) ein geeignetes Maß $\mu_1 \otimes \mu_2$ auf $(X_1 \times X_2, S_1 \otimes S_2)$ konstruieren, welches

$$(\mu_1 \otimes \mu_2)(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2)$$

für alle $A_j \in S_j$ erfüllt. Insbesondere gilt dann auch $\lambda^p \otimes \lambda^q = \lambda^{p+q}$, da dies für Quader A_j gilt.

Dafür benötigen wir zunächst etwas Terminologie. Sei $A \subseteq X_1 \times X_2$ und $f\colon X_1 \times X_2 \to \overline{\mathbb{R}}$. Für $x_1 \in X_1$ setzen wir

$$A_{x_1} = \{x_2 \in X_2\colon (x_1, x_2) \in A\} \quad \text{und} \quad f_{x_1}\colon X_2 \to \overline{\mathbb{R}},\ f_{x_1}(x_2) = f(x_1, x_2),$$

siehe Abbildung 4. Analog setzt man

$$A^{x_2} = \{x_1 \in X_1\colon (x_1, x_2) \in A\} \quad \text{und} \quad f^{x_2}\colon X_1 \to \overline{\mathbb{R}},\ f^{x_2}(x_1) = f(x_1, x_2).$$

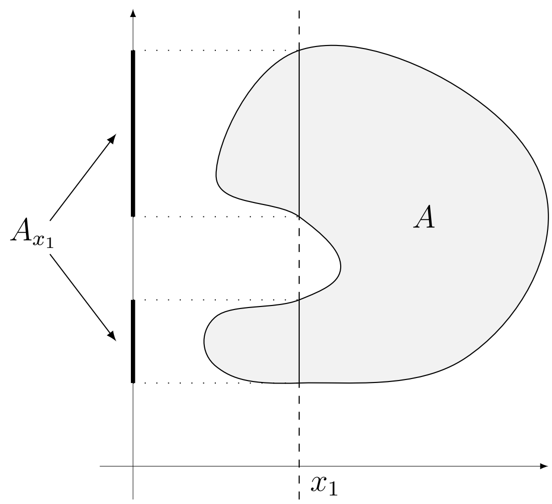

Abbildung 4: A_{x_1} ist die Projektion eines Schnitts von A auf X_2.

Lemma 12.4. *Seien (X_1, S_1) und (X_2, S_2) Messräume und $(x_1, x_2) \in X_1 \times X_2$.*

(i) *Ist $A \in S_1 \otimes S_2$, so gilt $A_{x_1} \in S_2$ und $A^{x_2} \in S_1$.*

(ii) *Ist $f\colon X_1 \times X_2 \to \overline{\mathbb{R}}$ messbar, so sind auch f_{x_1} und f^{x_2} messbar.*

Beweis. Die Funktion $\phi_{x_1}\colon X_2 \to X_1 \times X_2$, $t \to (x_1, t)$ ist nach Lemma 12.4 messbar, da $\pi_1 \circ \phi_{x_1}$ konstant und damit messbar und $\pi_2 \circ \phi_{x_1}$ die Identität und damit ebenfalls messbar ist. Wegen $A_{x_1} = \phi_{x_1}^{-1}(A)$ und $f_{x_1} = f \circ \phi_{x_1}$ folgen die Behauptungen bezüglich x_1. Die für x_2 folgen analog. □

Lemma 12.5. *Seien (X_1, S_1, μ_1) und (X_2, S_2, μ_2) σ-endliche Maßräume und $A \in S_1 \otimes S_2$. Dann sind die Funktionen $\mu_2^A\colon X_1 \to [0, \infty]$, $x_1 \mapsto \mu_2(A_{x_1})$ und $\mu_1^A\colon X_2 \to [0, \infty]$, $x_2 \mapsto \mu_1(A^{x_2})$ messbar.*

Beweis. Aus Symmetriegründen reicht es, μ_2^A zu betrachten. Zunächst nehmen wir an, dass μ_2 sogar endlich ist. Wir betrachten

$$D := \{A \in S_1 \otimes S_2\colon \mu_2^A \text{ ist messbar}\}.$$

Zu zeigen ist, dass $D = S_1 \otimes S_2$.

Dazu zeigen wir, dass D ein Dynkin-System ist, welches den schnittstabilen Erzeuger $S_1 \times S_2$ hat. Aus Satz 1.9 folgt dann, dass

$$D = \delta(S_1 \times S_2) = \sigma(S_1 \times S_2) = S_1 \otimes S_2,$$

also die Behauptung. Die Schnittstabilität von $S_1 \times S_2$ ist klar, da

$$(A_1 \times A_2) \cap (B_1 \times B_2) = (A_1 \cap B_1) \times (A_2 \cap B_2)$$

für $A_j, B_j \in S_j$. Es bleibt zu zeigen, dass D ein Dynkin-System ist. Nun ist $\mu_2^{\emptyset}(x_1) = \mu_2(\emptyset_{x_1}) = \mu_2(\emptyset) = 0$ für alle $x_1 \in X_1$, also $\mu_2^{\emptyset}$ konstant und damit messbar, also $\emptyset \in D$. Weiter gilt für $A \in D$, dass

$$\mu_2^{A^c}(x_1) = \mu_2(A^c_{x_1}) = \mu_2(X_2) - \mu_2(A_{x_1}) = \mu_2(X_2) - \mu_2^A(x_1),$$

also $\mu_2^{A^c} = \mu_2(X_2) - \mu_2^A$. Nach Satz 6.4 ist damit $\mu_2^{A^c}$ messbar, also auch $A^c \in D$.

Weiter gilt für eine Folge (A_j) paarweise disjunkter Mengen mit $A = \bigcup_{j=1}^{\infty} A_j$, dass $\mu_2^A = \sum_{j=1}^{\infty} \mu_2^{A_j}$. Nach Satz 6.2 ist damit μ_2^A als (punktweiser) Grenzwert messbarer Funktionen messbar, also $A \in D$. Damit ist D Dynkin-System.

Ist μ_2 nur σ-endlich, so existiert eine Folge (B_j) in X_2 mit $B_1 \subseteq B_2 \subseteq \dots$, $\mu_2(B_j) < \infty$ für alle $j \in \mathbb{N}$ und $X_2 = \bigcup_{j=1}^{\infty} B_j$. Nach dem bereits Bewiesenen sind die durch $x_1 \mapsto \mu_2(B_j \cap A_{x_1})$ gegebenen Abbildungen ν_j messbar, und wiederum mit Satz 6.2 gilt, dass dann auch für $\mu_2^A = \lim_{j\to\infty} \nu_j$ messbar ist. □

Satz 12.6. *Seien (X_1, S_1, μ_1) und (X_2, S_2, μ_2) σ-endliche Maßräume. Dann gibt es genau ein Maß $\mu_1 \otimes \mu_2$ auf $S_1 \otimes S_2$ mit*

$$(\mu_1 \otimes \mu_2)(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2)$$

für alle $A_j \in Sj$. Für $A \in S_1 \otimes S_2$ gilt

$$(\mu_1 \otimes \mu_2)(A) = \int_{X_1} \mu_2(A_{x_1}) d\mu_1(x_1) = \int_{X_2} \mu_1(A^{x_2}) d\mu_2(x_2).$$

Aus naheliegenden Gründen nennt man das Maß $\mu_1 \otimes \mu_2$ *Produktmaß.*

Beweis von Satz 12.6. Die Eindeutigkeit folgt leicht aus dem Maßeindeutigkeitssatz 3.8. Wir verzichten auf die Details.

Sei τ_1 und τ_2 die durch

$$\tau_1(A) = \int_{X_1} \mu_2(A_{x_1}) d\mu_1(x_1) \quad \text{und} \quad \tau_2(A) = \int_{X_2} \mu_1(A^{x_2}) d\mu_2(x_2)$$

gegebenen Abbildungen von $S_1 \otimes S_2$ nach $[0, \infty]$. Wegen Lemma 12.5 sind diese Integrale definiert. Wir zeigen, dass τ_1 und τ_2 Maße sind. Denn es gilt $\tau_1(\emptyset) = 0$ und für eine Folge (A_j) disjunkte Mengen in $S_1 \otimes S_2$ gilt

$$\begin{aligned}
\tau_1\left(\bigcup_{j=1}^{\infty} A_j\right) &= \int_{X_1} \mu_2\left(\left(\bigcup_{j=1}^{\infty} A_j\right)_{x_1}\right) d\mu_1(x_1) \\
&= \int_{X_1} \mu_2\left(\bigcup_{j=1}^{\infty} (A_j)_{x_1}\right) d\mu_1(x_1) \\
&= \int_{X_1} \sum_{j=1}^{\infty} \mu_2\big((A_j)_{x_1}\big)\, d\mu_1(x_1) \\
&= \sum_{j=1}^{\infty} \int_{X_1} \mu_2\big((A_j)_{x_1}\big)\, d\mu_1(x_1) \\
&= \sum_{j=1}^{\infty} \tau_1(A_j)\,.
\end{aligned}$$

Dabei wurde das Korollar 7.11 zum Satz von Beppo Levi benutzt. Analog sieht man, dass τ_2 ein Maß ist. Weiter gilt

$$\tau_1(A_1 \times A_2) = \int_{X_1} \chi_{A_1}(x_1)\mu_2(A_2) d\mu_1(x_1) = \mu_1(A_1)\mu_2(A_2).$$

Ebenso folgt $\tau_2(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2)$. Wegen der Eindeutigkeit folgt $\tau_1 = \tau_2$. Das Maß $\mu_1 \otimes \mu_2 := \tau_1 = \tau_2$ leistet also das Verlangte. □

13 Der Satz von Fubini

Sei $(X_1, S_1, \mu_1) = (\mathbb{R}^{d-1}, \mathcal{B}^{d-1}, \lambda^{d-1})$ und $(X_2, S_2, \mu_2) = (\mathbb{R}^1, \mathcal{B}^1, \lambda^1)$ in Satz 12.6. Dann ist $(X_1 \times X_2, S_1 \otimes S_2, \mu_1 \otimes \mu_2) = (\mathbb{R}^d, \mathcal{B}^d, \lambda^d)$, denn nach Beispiel 12.3 ist $\mathcal{B}^{d-1} \otimes \mathcal{B}^1 = \mathcal{B}^d$ und es gilt $\lambda^{d-1} \otimes \lambda^1 = \lambda^d$, da diese Maße auf Quadern übereinstimmen, und damit auch auf der von den Quadern erzeugten σ-Algebra $\mathcal{B}^{d-1} \otimes \mathcal{B}^1 = \mathcal{B}^d$. Das Ergebnis ist als Cavalierisches Prinzip bekannt.

Satz 13.1. (Cavalierisches Prinzip) *Sei $M \in \mathcal{B}^d$ und, für $x_d \in \mathbb{R}$,*

$$M^{x_d} = \{(x_1, \dots, x_{d-1}) \in \mathbb{R}^{d-1} \colon (x_1, \dots, x_d) \in M\}.$$

Dann gilt

$$\lambda^d(M) = \int_{\mathbb{R}} \lambda^{d-1}(M^{x_d}) dx_d.$$

Insbesondere folgt für $M, N \in \mathcal{B}^d$ aus $\lambda^{d-1}(M^{x_d}) = \lambda^{d-1}(N^{x_d})$ für alle $x^d \in \mathbb{R}$, dass $\lambda^d(M) = \lambda^d(N)$ gilt. Oft wird auch diese Aussage als Cavalierisches Prinzip bezeichnet.

Beispiel 13.2. Sei $R > 0$ und $K = \{(x, y, z) \in \mathbb{R}^3 \colon x^2 + y^2 + z^2 \leq R^2\}$ die Kugel vom Radius R. Für $|z| \leq R$ ist dann

$$M^z = \{(x, y) \in \mathbb{R}^2 \colon (x, y, z) \in M\} = \{(x, y) \in \mathbb{R}^2 \colon x^2 + y^2 \leq R^2 - z^2\}.$$

eine Kreisscheibe während $M^z = \emptyset$ für $|z| > R$. Mit einer Übungsaufgabe (Serie 4, Aufgabe 1) folgt $\lambda^2(M^z) = \pi(R^2 - z^2)$ für $|z| \leq R$, also

$$\lambda^3(M) = \int_{-R}^{R} \lambda^2(M^z) dz = \pi \int_{-R}^{R} (R^2 - z^2) dz = \pi \left(R^2 z - \frac{1}{3} z^3 \right) \Big|_{-R}^{R} = \frac{4}{3} \pi R^3.$$

Die erste Darstellung des Produktmaßes in Satz 12.6 kann man auch in der Form

$$\begin{aligned} \int_{X_1 \times X_2} \chi_A d(\mu_1 \otimes \mu_2) &= (\mu_1 \otimes \mu_2)(A) \\ &= \int_{X_1} \mu_2(A_{x_1}) d\mu_1(x_1) \\ &= \int_{X_1} \int_{X_2} \chi_A(x_1, x_2) d\mu_2(x_2) d\mu_1(x_1) \end{aligned}$$

schreiben. Für die zweite Darstellung gilt Entsprechendes. Es gilt also

$$\int_{X_1 \times X_2} \chi_A d(\mu_1 \otimes \mu_2) = \int_{X_1} \int_{X_2} \chi_A d\mu_2 d\mu_1 = \int_{X_2} \int_{X_1} \chi_A d\mu_1 d\mu_2.$$

Insbesondere existieren die auftretenden Integrale, d.h., die Integranden sind messbar.

Wegen der Linearität des Integrals kann man in obiger Formel die charakteristische Funktion durch eine Treppenfunktion ersetzen. Da nach Satz 7.5 jede messbare Funktion nach $[0, \infty]$ monotoner Grenzwert einer Folge von Treppenfunktionen ist, erhält man mit dem Satz von Beppo Levi folgenden Satz.

Satz 13.3. (Satz von Tonelli) *Seien (X_1, S_1, μ_1) und (X_2, S_2, μ_2) σ-endliche Maßräume und sei $f \colon X_1 \times X_2 \to [0, \infty]$ messbar (bzgl. $S_1 \otimes S_2$). Dann gilt*

$$\int_{X_1 \times X_2} f \, d(\mu_1 \otimes \mu_2) = \int_{X_1} \int_{X_2} f \, d\mu_2 d\mu_1 = \int_{X_2} \int_{X_1} f \, d\mu_1 d\mu_2.$$

Unser nächstes Ziel ist ein dem Satz von Tonelli entsprechendes Ergebnisse für integrierbare Funktionen. Im Gegensatz zur Messbarkeit folgt aus der Integrierbarkeit von f nicht wie in Lemma 12.4 die der Funktionen f_{x_1} und f^{x_2}. Beispielsweise ist für jede Funktion $g \colon [0, 1] \to \overline{\mathbb{R}}$ die Funktion $f \colon [0, 1]^2 \to \mathbb{R}$,

$$f(x_1, x_2) = \begin{cases} 0 & \text{falls } 0 < x_2 \leq 1, \\ g(x_1) & \text{falls } x_2 = 0, \end{cases}$$

integrierbar bezüglich $\lambda^2 = \lambda^1 \otimes \lambda^1$, da $f(x_1, x_2) = 0$ λ^2-fast überall. Aber $f^0 = g$ muss nicht λ-integrierbar sein.

Satz 13.4. (Satz von Fubini) *Seien* (X_1, S_1, μ_1) *und* (X_2, S_2, μ_2) σ*-endliche Maßräume und sei* $f\colon X_1 \times X_2 \to \overline{\mathbb{R}}$ *integrierbar (bzgl.* $\mu_1 \otimes \mu_2$*). Dann gilt:*

(i) *Für* μ_1*-fast alle* x_1 *ist* f_{x_1} μ_2*-integrierbar und für* μ_2*-fast alle* x_2 *ist* f^{x_2} μ_1*-integrierbar.*

(ii) *Die durch*

$$F_1(x_1) = \begin{cases} \displaystyle\int_{X_2} f_{x_1} d\mu_2 & \textit{falls } f_{x_1} \textit{ integrierbar,} \\ 0 & \textit{sonst,} \end{cases}$$

und

$$F_2(x_2) = \begin{cases} \displaystyle\int_{X_1} f^{x_2} d\mu_1 & \textit{falls } f^{x_2} \textit{ integrierbar,} \\ 0 & \textit{sonst,} \end{cases}$$

definierten Funktionen sind integrierbar und es gilt

$$\int_{X_1 \times X_2} f \; d(\mu_1 \otimes \mu_2) = \int_{X_1} F_1 \; d\mu_1 = \int_{X_2} F_2 \; d\mu_2.$$

Im Grunde ist es egal, wie man F_1 und F_2 an den Stellen definiert, an denen f_{x_1} bzw. f^{x_2} nicht integrierbar sind. Denn dieses sind Nullmengen, und Abänderung des Integranden auf einer Nullmenge ändert den Wert des Integrals nicht (Satz 8.5). In diesem Sinne kann man auch Integrale von Funktionen betrachten, die auf einer Nullmenge überhaupt nicht definiert sind. Interpretiert man die Integrale in diesem Sinne, so stimmt die letzte Formel im Satz von Fubini mit der aus dem Satz von Tonelli überein.

Beweis von Satz 13.4. (i) Die Messbarkeit der Funktionen erhält man aus Lemma 12.4, (i). Der Satz von Tonelli liefert

$$\int_{X_1} \int_{X_2} |f| d\mu_2 d\mu_1 = \int_{X_1 \times X_2} |f| d(\mu_1 \otimes \mu_2) < \infty.$$

Nach Satz 8.5, (i), gilt

$$\int_{X_2} |f_{x_1}| \, d\mu_2 = \int_{X_2} |f(x_1, x_2)| d\mu_2(x_2) < \infty$$

für μ_1-fast alle x_1. Daher ist f_{x_1} μ_2-integrierbar für μ_1-fast alle x_1. Das Ergebnis für f^{x_2} folgt analog.

(ii) Da f_{x_1} für μ_1-fast alle x_1 integrierbar ist, folgt die Messbarkeit von F_1 mit Satz 8.5 und der Zerlegung $f = f^+ - f^-$ wie im Satz von Tonelli. Außerdem liefert der Satz von Tonelli, dass

$$\int_{X_1} |F_1| d\mu_1 \le \int_{X_1} \int_{X_2} |f| d\mu_2 d\mu_1 = \int_{X_1 \times X_2} |f| d(\mu_1 \otimes \mu_2) < \infty,$$

also F_1 integrierbar ist. Analog erhält man die Integrierbarkeit von F_2.

Die letzte Formel gilt nach dem Satz von Tonelli mit $f^\pm$ an Stelle von f. Aber hieraus folgt die Formel auch für f. $\square$

Zusammenfassend erhält man aus den Sätzen von Tonelli und Fubini auch folgendes Resultat.

Satz 13.5. *Seien* (X_1, S_1, μ_1) *und* (X_2, S_2, μ_2) σ*-endliche Maßräume und sei* $f\colon X_1 \times X_2 \to \overline{\mathbb{R}}$ *messbar (bzgl.* $S_1 \otimes S_2$*). Ist*

$$\int_{X_1}\int_{X_2} |f| d\mu_2 d\mu_1 < \infty \quad \textit{oder} \quad \int_{X_2}\int_{X_1} |f| d\mu_1 d\mu_2 < \infty,$$

so ist f *integrierbar und es gilt*

$$\int_{X_1\times X_2} f\, d(\mu_1 \otimes \mu_2) = \int_{X_1}\int_{X_2} f\, d\mu_2 d\mu_1 = \int_{X_2}\int_{X_1} f\, d\mu_1 d\mu_2.$$

14 Die Transformationsformel

Wir wollen ein höherdimensionales Analogon der Substitutionsregel beweisen. Eine Formulierung der (eindimensionalen) Substitutionsregel ist wie folgt: *Ist* $g\colon [a,b] \to \mathbb{R}$ *stetig differenzierbar und* $f\colon g([a,b]) \to \mathbb{R}$ *stetig, so gilt*

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

Falls $g'(x) > 0$ für alle in $x \in [a,b]$, so gilt $g(a) < g(b)$, und falls $g'(x) < 0$ für alle in $x \in [a,b]$, so gilt $g(b) < g(a)$. In beiden Fällen erhalten wir

$$\int_{[a,b]} f(g(x))|g'(x)|dx = \int_{g([a,b])} f(y)dy.$$

Es sei angemerkt, dass unter den obigen Voraussetzungen g bijektiv ist und neben g auch g^{-1} stetig differenzierbar ist.

Definition 14.1. Seien $U, V \subseteq \mathbb{R}^d$ offen. Eine bijektive Abbildung $g\colon U \to V$ heißt *Diffeomorphismus*, wenn g und g^{-1} beide stetig differenzierbar sind.

Für stetig differenzierbares $g\colon U \to V$ bezeichnen wir mit $J_g(x)$ die Jacobi-Matrix von g im Punkte $x \in U$. Ist g bijektiv und $J_g(x)$ für alle $x \in U$ invertierbar, so folgt aus dem Umkehrsatz die Differenzierbarkeit von g^{-1}. Mit $y = g(x)$ ist dann $J_{g^{-1}}(y) = J_g(x)^{-1}$.

Satz 14.2. (Transformationsformel) *Seien* $U, V \subseteq \mathbb{R}^d$ *offen,* $g\colon U \to V$ *ein Diffeomorphismus und* $f\colon V \to \overline{\mathbb{R}}$ *messbar. Dann ist* f *über* V *genau dann integrierbar, wenn* $(f \circ g) \cdot |\det J_g|$ *über* U *integrierbar ist. In diesem Fall gilt*

$$\int_U f(g(x))\, |\det J_g(x)|\, d\lambda(x) = \int_V f(y) d\lambda(y).$$

Vor dem Beweis machen wir eine Plausibilitätsüberlegung: Wir approximieren U durch eine Vereinigung kleiner Quader Q_k und wählen $x_k \in Q_k$. Es scheint naheliegend, dass

$$\int_U f(g(x))|\det J_g(x)|dx \approx \sum_k f(g(x_k))|\det J_g(x_k)|\lambda(Q_k),$$

da die rechte Seite als Riemannsche Summe für das Integral angesehen werden kann. Mit $y_k = g(x_k)$ und $R_k = g(Q_k)$ erhält man ebenso

$$\int_V f(y)dy \approx \sum_k f(y_k)\lambda(R_k) = \sum_k f(g(x_k))\lambda(R_k).$$

In Q_k wird $g(x)$ gut durch $L_k(x) = g(x_k) + J_g(x_k)(x - x_k)$ approximiert. Daher ist $R_k = g(Q_k) \approx L_k(Q_k)$ und damit nach Satz 5.7

$$\lambda(R_k) \approx \lambda(L_k(Q_k)) = |\det L_k|\ \lambda(Q_k) = |\det J_g(x_k)|\ \lambda(Q_k).$$

Beispiel 14.3. (*Polarkoordinaten*). Sei

$$T\colon (0,\infty) \times (-\pi,\pi) \to \mathbb{R}^2 \backslash \{(x,0)\colon x \le 0\},\ \begin{pmatrix} r \\ \varphi \end{pmatrix} \mapsto \begin{pmatrix} r\cos\varphi \\ r\sin\varphi \end{pmatrix}.$$

Es ist T Diffeomorphismus und es gilt

$$J_T(r,\varphi) = \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix}.$$

Es folgt, dass

$$\det J_T(r,\varphi) = r(\cos^2\varphi + \sin^2\varphi) = r.$$

Man beachte, dass $N = \{(x,0)\colon x \le 0\}$ Nullmenge ist, also das Bild von T bis auf eine Nullmenge der ganze $\mathbb{R}^2$ ist.

Sei nun $R > 0$ und $K = \{(x,y) \in \mathbb{R}^2\colon x^2 + y^2 \le R^2, y \ge 0\}$. Mit

$$L = \{(r,\varphi)\colon 0 < r \le R,\ 0 \le \varphi < \pi\}$$

gilt $T(L) = K\backslash N$. Wir erhalten zum Beispiel

$$\begin{aligned}
\int_K y\ d\lambda(x,y) &= \int_{K\backslash N} y\ d\lambda(x,y) \\
&= \int_L r\sin\varphi\ \det J_T(r,\varphi)\ d\lambda(r,\varphi) \\
&= \int_0^R \int_0^\pi r^2 \sin\varphi\ d\varphi\ dr \\
&= \int_0^R r^2(-\cos\varphi)\Big|_{\varphi=0}^{\varphi=\pi} dr \\
&= 2\int_0^R r^2 dr \\
&= \frac{2}{3}R^3.
\end{aligned}$$

In der Übung wurde $\lambda(K) = \frac{1}{2}\pi R^2$ gezeigt. Dies erhält man durch Transformation auf Polarkoordinaten sehr einfach. (Man ersetze den Integranden y durch 1.)

Der Schwerpunkts $x^s = (x_1^s, \dots, x_d^s)$ eines Körpers $K \subseteq \mathbb{R}^d$ ist durch

$$x_j^s = \frac{1}{\lambda(K)} \int_K x_j d\lambda(x_1, \dots, x_d)$$

gegeben. Es folgt, dass der Schwerpunkt von K in $(0, \frac{4}{3\pi}R)$ liegt.

Beispiel 14.4. (*Zylinderkoordinaten*). Wir führen im $\mathbb{R}^3$ Polarkoordinaten in der (x, y)-Ebene ein und lassen die z-Koordinate unverändert.

Wir betrachten also $T\colon (0,\infty)\times(-\pi,\pi)\times\mathbb{R} \to \mathbb{R}^3\setminus\{(x,0,z)\colon x \le 0,\ z \in \mathbb{R}\}$,

$$\begin{pmatrix} r \\ \varphi \\ z \end{pmatrix} \mapsto \begin{pmatrix} r\cos\varphi \\ r\sin\varphi \\ z \end{pmatrix}$$

Es ist T ein Diffeomorphismus und es gilt

$$\det J_T(r,\varphi,z) = r.$$

Außerdem ist $N = \{(x,0,z)\colon x \le 0,\ z \in \mathbb{R}\}$ Nullmenge, also das Bild von T bis auf eine Nullmenge der ganze $\mathbb{R}^3$.

Sei nun $0 < R_1 < R_2$ und

$$K = \left\{(x,y,z) \in \mathbb{R}^3\colon x^2 + y^2 + z^2 \le R_2^2,\ x^2 + y^2 \ge R_1^2\right\}.$$

Es ist K also der Teil der (abgeschlossenen) Kugel vom Radius R_2 um 0, der außerhalb des Zylinders vom Radius R_1 um die z-Achse liegt; vgl. Abbildung 5.

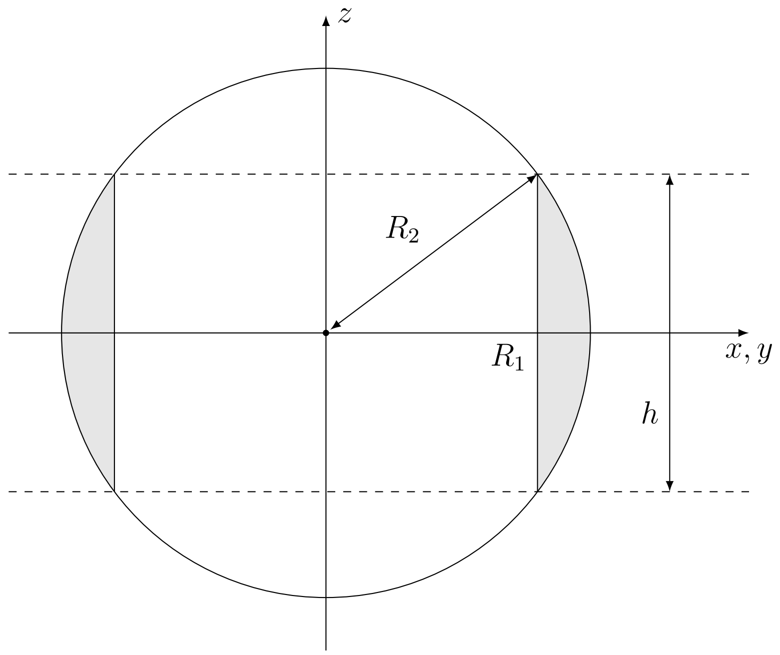

Abbildung 5: Durchbohrte Kugel.

Mit

$$L = \left\{(r,\varphi,z)\colon R_1 \le r \le R_2,\ -\pi < \varphi < \pi,\ -\sqrt{R_2^2 - r^2} \le z \le \sqrt{R_2^2 - r^2}\right\}$$

gilt $T(L) = K\setminus N$.

Es folgt

$$
\begin{aligned}
\lambda^3(K) =& \lambda^3(K\backslash N) \\
=& \int_{K\backslash N} 1 \, d(x,y,z) \\
=& \int_L \det J_T(r,\varphi,z) d(r,\varphi,z) \\
=& \int_{R_1}^{R_2} \int_{-\pi}^{\pi} \int_{-\sqrt{R_2^2-r^2}}^{\sqrt{R_2^2-r^2}} r \, dz \, d\varphi \, dr \\
=& \int_{R_1}^{R_2} \int_{-\pi}^{\pi} 2r\sqrt{R_2^2-r^2} \, d\varphi \, dr \\
=& 2\pi \int_{R_1}^{R_2} 2r\sqrt{R_2^2-r^2} \, dr \\
=& -2\pi \frac{2}{3} \left(R_2^2-r^2\right)^{3/2} \Big|_{r=R_1}^{r=R_2} \\
=& \frac{4\pi}{3} \left(R_2^2-R_1^2\right)^{3/2}.
\end{aligned}
$$

Für $R_1 = 0$ ergibt sich insbesondere – wie bereits in Beispiel 13.2 berechnet – das Volumen einer Kugel von Radius R zu $\frac{4\pi}{3}R^3$.

Ist h die „Höhe“ des Rings K (siehe Abbildung 5), so gilt $\frac{1}{2}h = \sqrt{R_2^2 - R_1^2}$ und damit $\lambda^3(K) = \frac{1}{6}\pi h^3$. Bei gegebenem h hängt $\lambda^3(K)$ also nicht vom Kugelradius R_2 ab. Tatsächlich hängt für $K^z = \{(x,y) \in \mathbb{R}^2 \colon (x,y,z) \in K\}$ bei festem h auch $\lambda^2(K^z)$ nicht von R_2 ab. Damit liefert auch das Cavalierische Prinzip, dass $\lambda^3(K)$ unabängig von R_2 ist.

Beispiel 14.5. (*Kugelkoordinaten*). Sei $T \colon (0,\infty) \times (-\pi,\pi) \times \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \to \mathbb{R}^3$,

$$
\begin{pmatrix} r \\ \varphi \\ \theta \end{pmatrix} \mapsto \begin{pmatrix} r\cos\varphi\cos\theta \\ r\sin\varphi\cos\theta \\ r\sin\theta \end{pmatrix}
$$

Hier ist θ der Winkel, den die Verbindungsstrecke von (x,y,z) und dem Ursprung mit der (x,y)-Ebene bildet; vgl. Abbildung 6. Es ist auch üblich, stattdessen den Winkel zwischen der Verbindungsstrecke und der z-Achse zu nehmen.

Es existiert eine Nullmenge N, so dass $T \colon (0,\infty)\times(-\pi,\pi)\times\left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \to \mathbb{R}^3\backslash N$ ein Diffeomorphismus ist. Eine Rechnung zeigt, dass

$$
\det J_T(r,\varphi,\theta) = r^2\cos\theta.
$$

Für K aus Beispiel 14.4 entspricht die Bedingung $x^2+y^2+z^2 \leq R_2^2$ der Bedingung $r \leq R_2$ und die Bedingung $x^2+y^2 \geq R_1^2$ erhält die Form $r^2\cos^2\theta \geq R_1^2$, also $r \geq R_1/\cos\theta$. Man beachte bei der letzten Umformung, dass $\cos\theta > 0$ für $\theta \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. Aus $R_1/\cos\theta \leq r \leq R_2$ folgt, dass nur Werte von θ betrachtet werden, für die $R_1/\cos\theta \leq R_2$ gilt, also $-\arccos(R_1/R_2) \leq \theta \leq \arccos(R_1/R_2)$.

Mit

$$
L = \left\{(r,\varphi,\theta) \colon -\pi < \varphi < \pi, \ -\arccos\frac{R_1}{R_2} \leq \theta \leq \arccos\frac{R_1}{R_2}, \ \frac{R_1}{\cos\theta} \leq r \leq R_2\right\}
$$

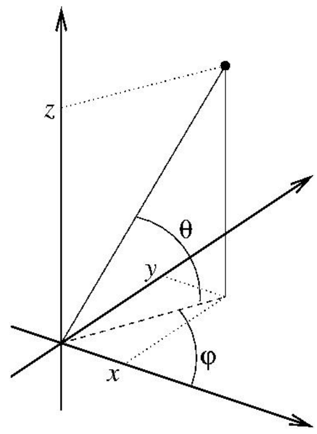

Abbildung 6: Kugelkoordinaten.

ergibt sich mit obiger Nullmenge N, dass $T(L) = K\backslash N$. Hieraus folgt, dass

$$\begin{aligned}
\lambda^3(K) &= \lambda^3(K\backslash N) \\
&= \int_{K\backslash N} 1 \, d(x,y,z) \\
&= \int_L \det J_T(r,\varphi,\theta) d(r,\varphi,\theta) \\
&= \int_{-\pi}^{\pi} \int_{-\arccos(R_1/R_2)}^{\arccos(R_1/R_2)} \int_{R_1/\cos\theta}^{R_2} r^2 \cos\theta \, dr \, d\theta \, d\varphi \\
&= 4\pi \int_0^{\arccos(R_1/R_2)} \frac{1}{3} \left(R_2^3 \cos\theta - \frac{R_1^3}{\cos^2\theta} \right) d\theta \\
&= \frac{4\pi}{3} R_2^3 \sin\theta - R_1^3 \tan\theta \Big|_0^{\arccos(R_1/R_2)} \\
&= \frac{4\pi}{3} \left(R_2^3 - \frac{R_1^3}{\cos\theta} \right) \sin\theta \Big|_0^{\arccos(R_1/R_2)} \\
&= \frac{4\pi}{3} \left(R_2^3 - R_1^2 R_2 \right) \sqrt{1 - \left(\frac{R_1}{R_2} \right)^2} \\
&= \frac{4\pi}{3} \left(R_2^2 - R_1^2 \right)^{3/2}.
\end{aligned}$$

Alternativ kann man auch folgende Integrationsreihenfolge wählen:

$$\lambda^3(K) = \int_{-\pi}^{\pi} \int_{R_1}^{R_2} \int_{-\arccos(R_1/r)}^{\arccos(R_1/r)} r^2 \cos\theta \, d\theta \, dr \, d\varphi = \cdots = \frac{4\pi}{3} \left(R_2^2 - R_1^2 \right)^{3/2}.$$

Beweis der Transformationsformel. Da g, g^{-1} und $\det J_g$ stetig sind, folgt, dass $(f \circ g)|\det J_g|$ genau dann messbar ist, wenn f messbar ist.

Sei zunächst $f\colon V \to [0,\infty]$ messbar. Als Erstes überlegen wir uns, dass die Gleichung aus dem Satz aus der Ungleichung

$$\int_V f(y) d\lambda(y) \le \int_U f(g(x)) |\det J_g(x)| d\lambda(x) \tag{$*$}$$

folgt. (Die Transformationsformel sagt, dass hier Gleichheit gilt.) Denn wenden wir diese Ungleichung auf $(f \circ g)|\det J_g|$ statt f und auf g^{-1} statt g an, so erhalten wir

$$\begin{aligned}\int_U f(g(x))|\det J_g(x)|d\lambda(x) &\leq \int_V f(g(g^{-1}(y)))|\det J_g(g^{-1}(y))|\cdot|\det J_{g^{-1}}(y)|d\lambda(y)\\ &= \int_V f(y)d\lambda(y).\end{aligned}$$

Zum Beweis der Ungleichung $(*)$ benötigen wir mehrere Lemmata.

Lemma 14.6. *Für einen kompakten Würfel $W \subseteq U$ gilt*

$$\lambda(g(W)) \leq \lambda(W) \cdot \max_{x \in W} |\det J_g(x)|.$$

Wir stellen den Beweis noch etwas zurück. Das folgende Lemma verallgemeinert diese Aussage. Es besagt, dass die zu Beginn des Beweises genannte Ungleichung $(*)$ gilt, falls $f = \chi_{g(Q)}$ mit einem kompakten Quader Q ist.

Lemma 14.7. *Für einen Quader $Q = (a, b]$ mit $a, b \in \mathbb{Q}^d$, dessen Abschluss in U liegt, gilt*

$$\lambda(g(Q)) \leq \int_Q \det J_g(x)d\lambda(x).$$

Beweis. Zu gegebenem $\varepsilon > 0$ existiert $\delta > 0$ mit $|\det J_g(x) - \det J_g(y)| < \varepsilon$ für $x, y \in Q$ mit $\|x - y\|_\infty \leq \delta$. Wir zerlegen nun Q in Würfel $W_1, \ldots, W_m$ mit Kantenlänge kleiner als δ; vgl. die Beweisskizze zu Satz 5.6. Es folgt, dass

$$|\det J_g(y)| \geq \max_{x \in W_k} |\det J_g(x)| - \varepsilon \quad \text{für alle } y \in W_k.$$

Mit Lemma 14.6 erhält man nun

$$\begin{aligned}\lambda(g(W_k)) &\leq \lambda(W) \cdot \max_{x \in W_k} |\det J_g(x)|\\ &\leq \int_{W_k} (|\det J_g(y)| + \varepsilon)d\lambda(y)\\ &= \int_{W_k} |\det J_g(y)|d\lambda(y) + \varepsilon\lambda(W_k).\end{aligned}$$

Summation über k liefert zusammen mit $\varepsilon \to 0$ die Behauptung. □

Lemma 14.8. *Für jede Borelmenge M in U gilt*

$$\lambda(g(M)) \leq \int_M |\det J_g(x)|d\lambda(x).$$

Beweisskizze. Durch $\mu := g^{-1}(\lambda)$, also $A \mapsto \mu(A) = g^{-1}(\lambda)(A) = \lambda(g(A))$ und

$$A \mapsto \nu(A) := \int_A |\det J_g(x)|d\lambda(x)$$

sind Maße μ und ν auf $\overline{\mathcal{B}}(U)$ definiert. Nach Lemma 14.7 gilt $\mu(Q) \leq \nu(Q)$ für alle (halboffenen) Quader $Q = (a, b]$ mit $a, b \in \mathbb{Q}^d$, deren Abschluss in U liegt. Hieraus folgt, dass $\mu(A) = \nu(A)$ für alle A aus der σ-Algebra gilt, die von diesen Quadern erzeugt wird; vgl. Satz 3.8. Diese σ-Algebra besteht aber gerade aus den Borelmengen in U; vgl. Satz 1.5. □

Wir fahren nun mit dem Beweis der Transformationsformel fort. Lemma 14.8 impliziert, dass die Ungleichung $(*)$ gilt, falls $f = \chi_A$ für eine Borelmenge A ist. Denn mit $M = g^{-1}(A)$ gilt $\lambda(A) = \lambda(g(M))$, also

$$\int_V \chi_A(y) d\lambda(y) = \lambda(A) \leq \int_M |\det J_g(x)| d\lambda(x) = \int_U \chi_A(g(x)) |\det J_g(x)| d\lambda(x).$$

Damit gilt $(*)$ auch für Treppenfunktionen und wegen Satz 7.5 und des Satzes über monotone Konvergenz auch für messbares $f\colon V \to [0,\infty]$. Damit gilt die Formel aus dem Satz für messbares $f\colon V \to [0,\infty]$.

Insbesondere ist so eine Funktion f über V genau dann integrierbar, wenn $(f \circ g) \cdot |\det J_g|$ über U integrierbar ist. Die Behauptung für $f\colon V \to \overline{\mathbb{R}}$ folgt nun, wenn man Obiges auf f^+ und f^- anwendet. Damit ist die Transformationsformel bis auf den noch fehlenden Beweis von Lemma 14.6 bewiesen.

Beweis von Lemma 14.6. Sei $s_0 > 0$ die Seitenlänge von W und $\alpha \geq 0$ mit

$$\lambda(g(W)) = \alpha \cdot \lambda(W) = \alpha s_0^d.$$

Wir zerlegen W in 2^d kompakte Würfel der halben Seitenlänge $s_0/2$. Für einen dieser Würfel, den wir W_1 nennen, gilt

$$\lambda(g(W_1)) \geq \alpha \cdot \lambda(W_1).$$

Induktiv erhalten wir eine Folge (W_k) kompakter Würfel, wobei W_k die Seitenlänge $s_k := 2^{-k} s_0$ hat und $W_{k+1} \subseteq W_k$ sowie

$$\lambda(g(W_k)) \geq \alpha \cdot \lambda(W_k) = \alpha s_k^d.$$

für alle $k \in \mathbb{N}$ gilt.

Nun existiert $a \in U$ mit

$$\bigcap_{k=1}^{\infty} W_k = \{a\}.$$

Wir können $a = 0 = g(a)$ annehmen. Die Differenzierbarkeit von g liefert

$$g(x) = J_g(0) \cdot x + r(x),$$

wobei $r(x)/\|x\| \to 0$ für $x \to 0$. Sei L die durch $L(x) = J_g(0) \cdot x$ gegebene lineare Abbildung. Für gegebenes $\varepsilon > 0$ erhält man

$$\|g(x) - L(x)\| = \|r(x)\| < \varepsilon s_k$$

für $x \in W_k$, falls k genügend groß ist.

Für $M \subseteq \mathbb{R}^d$, $x \in \mathbb{R}^d$ und $\varepsilon > 0$ sei $\operatorname{dist}(x, M) = \inf\{\|x - y\|\colon y \in M\}$ der Abstand von x zu M und $U_\varepsilon(M) = \{y \in \mathbb{R}^d\colon \operatorname{dist}(y, M) < \varepsilon\}$ die ε-Umgebung von M. Aus der letzten Ungleichung erhält man dann $g(W_k) \subseteq U_{\varepsilon s_k}(L(W_k))$.

Es folgt, dass

$$\lambda(g(W_k)) \leq \lambda(U_{\varepsilon s_k}(L(W_k))) = s_k^d \lambda(U_\varepsilon(L(s_k^{-1} W_k))) = s_k^d \lambda(U_\varepsilon(L([0,1]^d))).$$

Zusammen mit der oben bewiesenen Ungleichung $\lambda(g(W_k)) \geq \alpha s_k^d$ folgt

$$\alpha \leq \lambda(U_\varepsilon(L([0,1]^d))).$$

Mit Satz 5.7 und dem Grenzübergang $\varepsilon \to 0$ erhält man

$$\alpha \leq \lambda(L([0,1]^d)) = |\det J_g(0)| \leq \max_{x \in W} |\det J_g(x)|.$$

□

II Gewöhnliche Differentialgleichungen

1 Beispiele und elementare Lösungsmethoden

Bei vielen Wachstums- und Zerfallprozessen ist die zeitliche Änderung einer Größe proportional zu der Größe selbst. Ist $y(t)$ die Größe zum Zeitpunkt t, so erhält man die Gleichung

$$y'(t) = \alpha y(t)$$

mit einer reellen Konstanten α. Dabei ist $\alpha > 0$ bei Wachstumsprozessen und $\alpha < 0$ bei Zerfallsprozessen. Beispielsweise wird der radioaktive Zerfall durch diese Gleichung beschrieben.

Des Weiteren sei die Größe zu einem Zeitpunkt t_0 bekannt, es gelte etwa $y(t_0) = y_0 > 0$. Ist y eine Lösung der gesuchten Gleichung, so gilt in einem geeigneten Intervall I um t_0 dann $y(t) > 0$ und damit

$$\alpha(t - t_0) = \int_{t_0}^{t} \alpha \, ds = \int_{t_0}^{t} \frac{y'(s)}{y(s)} ds = \log y(t) - \log y(t_0) = \log \frac{y(t)}{y(t_0)},$$

also

$$y(t) = y(t_0) e^{\alpha(t-t_0)} = y_0 e^{-\alpha t_0} e^{\alpha t}$$

für $t \in I$.

Man erkennt leicht, dass durch die letzte Formel eine Lösung der Gleichung $y'(t) = \alpha y(t)$ auf ganz $\mathbb{R}$ gegeben ist und dass diese die eindeutig bestimmte Lösung ist, die der Bedingung $y(t_0) = y_0$ genügt. Denn ist $\tilde{y}$ eine weitere Lösung, so zeigt eine kurze Rechnung, dass $(\tilde{y}/y)' = 0$ und damit $\tilde{y}/y$ konstant ist.

Im Folgenden werden wir Gleichungen des obigen Typs untersuchen, wobei aber auch höhere Ableitungen vorkommen können. Da die Variable nicht notwendigerweise die Zeit sein muss, bezeichnen wir die Variable im Allgemeinen mit x statt t. Allgemein heißt eine Gleichung der Form

$$f\big(x, y(x), y'(x), \dots, y^{(n)}(x)\big) = 0$$

gewöhnliche Differentialgleichung n-ter Ordnung. Dabei ist $n \in \mathbb{N}$ und f eine auf einer geeigneten Teilmenge von $\mathbb{R}^{n+2}$ gegebene Funktion. Gesucht ist eine n-mal (stetig) differenzierbare Funktion $y \colon I \to \mathbb{R}$ mit einem Intervall $I \subseteq \mathbb{R}$, die dieser Gleichung genügt. (Während in Kapitel I auch entartete Intervalle $I = [a, a] = \{a\}$ zugelassen waren, schließen wir dies im Folgenden aus, da für Funktionen auf solchen Intervallen die Ableitung nicht definiert ist.)

Ist die obige Gleichung nach $y^{(n)}$ aufgelöst, also von der Form

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

so spricht man von einer *expliziten* (gewöhnlichen) Differentialgleichung. Andernfalls nennt man die Differentialgleichung *implizit*.

In der Regel sind zusätzlich zu der Differentialgleichung noch sogenannte *Anfangsbedingungen*

$$y(x_0) = y_0, \quad y'(x_0) = y_1, \quad \dots, \quad y^{(n-1)}(x_0) = y_{n-1},$$

gegeben. Man spricht dann auch von einem *Anfangswertproblem.*

Gelegentlich sind statt Anfangsbedingungen auch andere Bedingungen gegeben, etwa sogenannte Randbedingungen, beispielsweise $y(a) = y_a$ und $y(b) = y_b$ für Lösungen y auf dem Intervall $[a, b]$.

Es ergeben sich bei Differentialgleichungen die folgenden Fragestellungen:

(a) Existenz und Eindeutigkeit von Lösungen;

(b) Berechnung von Lösungen (möglich nur in speziellen Fällen);

(c) Eigenschaften von Lösungen (insbesondere falls Berechnung nicht möglich);

(d) Numerische Lösung.

Hier behandeln wir zunächst (b), dann (a). Später wird auch etwas zu (c) gesagt. Punkt (d) wird in der Numerischen Mathematik behandelt.

Wir betrachten zunächst Differentialgleichungen 1. Ordnung, also

$$y'(x) = f\left(x, y(x)\right).$$

Solche Differentialgleichungen kann man geometrisch wie folgt interpretieren. Gesucht wird eine Funktion $x \mapsto y = y(x)$, wobei die Richtung des Tangentenvektors an den Graphen im Punkte (x, y) durch $(1, f(x, y))$ gegeben ist; vgl. Abbildung 7. Man nennt die Menge dieser Vektoren auch *Vektorfeld.* Formal ist ein Vektorfeld nichts anderes als eine Abbildung von einer offenen Teilmenge des $\mathbb{R}^d$ nach $\mathbb{R}^d$.

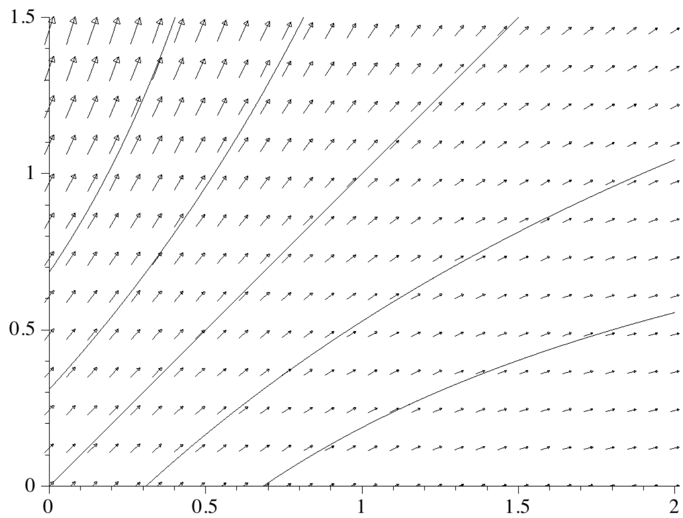

Abbildung 7: Vektorfeld zur Differentialgleichung $y' = (1 + y^2)/(1 + x^2)$. Die Lösungen sind von der Form $y(x) = \tan(\arctan(x) + c)$ mit $c \in \mathbb{R}$.

Wir untersuchen nun Differentialgleichungen gewisser spezieller Typen.

Definition 1.1. Seien I, J Intervalle, $a\colon I \to \mathbb{R}$ und $b\colon J \to \mathbb{R}$ stetig, $x_0 \in I$ und $y_0 \in J$. Die Differentialgleichung

$$y'(x) = a(x)b\left(y(x)\right), \quad y(x_0) = y_0,$$

heißt *Differentialgleichung mit getrennten Veränderlichen.*

Lösungsverfahren. Ist $b(y_0) = 0$, so ist $y(x) \equiv y_0$ eine Lösung. Wir nehmen an, dass $b(y_0) \neq 0$ gilt. Da b stetig ist, kann man J so wählen, dass entweder $b(y) > 0$ für alle $y \in J$ oder $b(y) < 0$ für alle $y \in J$ gilt.

Wegen
$$a(x) = \frac{y'(x)}{b(y(x))}$$
folgt dann
$$\int_{x_0}^{x} a(t)dt = \int_{x_0}^{x} \frac{y'(t)}{b(y(t))}dt = \int_{y_0}^{y(x)} \frac{ds}{b(s)}.$$
Mit
$$A(x) = \int_{x_0}^{x} a(t)dt \quad \text{und} \quad B(y) = \int_{y_0}^{y} \frac{ds}{b(s)}$$
erhalten wir
$$A(x) = B(y(x)).$$
Da $B' = 1/b$ konstantes Vorzeichen hat, ist B streng monoton. Damit existiert die Umkehrfunktion $B^{-1}\colon B(J) \to J$. Durch
$$y(x) = B^{-1}(A(x))$$
ist dann eine Lösung der Differentialgleichung gegeben, falls die rechte Seite definiert ist, das heißt, falls $A(I) \subseteq B(J)$.

Wegen $A(x_0) = 0 = B(y_0) \in B(J)$ existiert ein Intervall I_0 mit $x_0 \in I_0$ und $A(I_0) \subseteq B(J)$. Auf I_0 ist dann durch obige Gleichung eine Lösung y gegeben.

Wir erhalten folgendes Ergebnis.

Satz 1.2. *Seien I, J, a, b, x_0, y_0 wie in Definition 1.1, mit $b(y) \neq 0$ für alle $y \in J$. Seien*
$$A(x) = \int_{x_0}^{x} a(t)dt \quad \textit{und} \quad B(y) = \int_{y_0}^{y} \frac{ds}{b(s)}.$$
Weiter sei I_0 Intervall mit $x_0 \in I_0 \subseteq I$ und $A(I_0) \subseteq B(J)$. Dann ist auf I_0 durch $y(x) = B^{-1}(A(x))$ die eindeutige Löung des Anfangswertproblems
$$y'(x) = a(x)b(y(x)), \quad y(x_0) = y_0,$$
gegeben.

Beispiel 1.3. Wir betrachten das Anfangswertproblem
$$y'(x) = x(1 + y(x)^2), \quad y(0) = 1.$$
Dieses ist vom obigen Typ, mit $I = J = \mathbb{R}$, $a(x) = x$ und $b(y) = 1 + y^2$. Wir erhalten (wobei wir hier und im Folgenden statt $y(x)$ oft nur y schreiben)
$$\int_0^x t\,dt = \int_1^y \frac{ds}{1+s^2},$$
also
$$A(x) := \frac{1}{2}x^2 = \arctan y - \arctan 1 = \arctan y - \frac{\pi}{4} =: B(y).$$

Dies liefert

$$y(x) = y = \tan\left(\frac{x^2}{2} + \frac{\pi}{4}\right).$$

Dies gilt in einem Intervall I_0 um 0.

Zur Bestimmung des Intervalls I_0 beachte man, dass der Arcustangens $\mathbb{R}$ auf $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ abbildet, also $B(J) = B(\mathbb{R}) = \left(-\frac{3\pi}{4}, \frac{\pi}{4}\right)$ gilt. Wir suchen nun ein Intervall I_0 mit $A(I_0) \subseteq \left(-\frac{3\pi}{4}, \frac{\pi}{4}\right)$. Mit $A(x) = \frac{1}{2}x^2$ erhalten wir schließlich $I_0 = \left(-\sqrt{\frac{\pi}{2}}, \sqrt{\frac{\pi}{2}}\right)$.

Man hätte in diesem Beispiel natürlich auch einfach das größte Intervall um 0 suchen können, in der $y(x) = \tan(x^2/2 + \pi/4)$ definiert ist. Aber diese „naive“ Methode kann schiefgehen, wie das folgende Beispiel zeigt.

Beispiel 1.4. Wir betrachten das Auslaufen einer Tonne. Wir bezeichnen mit $h(t)$ die Höhe des Flüssigkeitsstand zur Zeit t und mit $v(t)$ die Auslaufgeschwindigkeit zur Zeit t, siehe Abbildung 8.

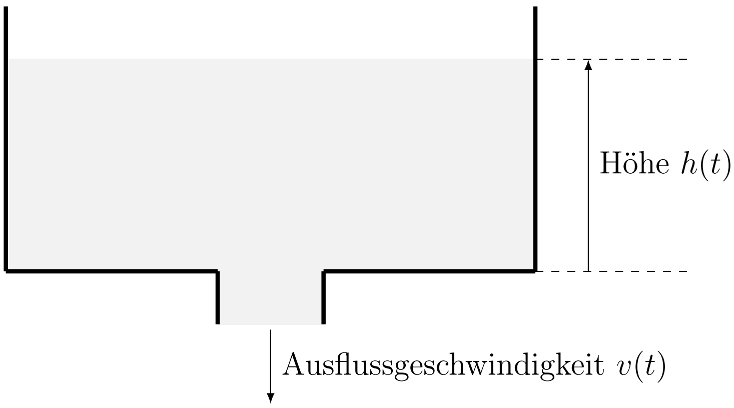

Abbildung 8: Auslaufen einer Tonne.

Die Änderung der potentiellen Energie bei Auslaufen eines Volumens ΔV ist proportional zu $\Delta V \cdot h(t)$. Die Änderung der kinetischen Energie ist proportional zu $\Delta V \cdot v(t)^2$. Wegen der Energieerhaltung ist also $h(t)$ proportional zu $v(t)^2$. Andererseits ist $h'(t)$ offensichtlich proportional zu $v(t)$.

Wir erhalten das Anfangswertproblem

$$h'(t) = -2c\sqrt{h(t)}, \quad h(0) = h_0 > 0,$$

mit einer Konstanten $c > 0$. (Die Konstante ist als $2c$ gewählt, damit die Formeln später schöner werden.)

Mit $a(x) = -2c$ und $b(y) = \sqrt{y}$ erhält man

$$-2ct = \int_0^t a(u)du = \int_{h_0}^{h(t)} \frac{ds}{b(s)} = \int_{h_0}^{h(t)} \frac{ds}{\sqrt{s}} = 2\left(\sqrt{h(t)} - \sqrt{h_0}\right),$$

also

$$h(t) = \left(\sqrt{h_0} - ct\right)^2.$$

Wir untersuchen jetzt, für welche t-Werte dies gilt. Mit $J = (0, \infty)$ und $B(y) = 2\left(\sqrt{y} - \sqrt{h_0}\right)$ erhalten wir $B(J) = \left(-2\sqrt{h_0}, \infty\right)$. Damit folgt, dass $A(t) = -2ct \in B(J)$ für $ct < \sqrt{h_0}$, also für $t < t_1 := \sqrt{h_0}/c$. Also ist durch obige Formel eine Lösung des Anfangswertproblems auf dem Intervall $I_0 = (-\infty, t_1)$ gegeben.

Die durch $h(t) = \left(\sqrt{h_0} - ct\right)^2 = c^2(t - t_1)^2$ gegebene Funktion $h\colon \mathbb{R} \to \mathbb{R}$ ist aber *keine* Lösung der Differentialgleichung auf ganz $\mathbb{R}$, denn für $t > t_1$ ist

$$\frac{d}{dt}\left(c^2(t-t_1)^2\right) = 2c^2(t-t_1) = 2c\sqrt{c^2(t-t_1)^2} \neq -2c\sqrt{c^2(t-t_1)^2}.$$

Physikalisch sinnvoll ist offensichtlich, vgl. Abbildung 9,

$$h(t) = \begin{cases} c^2\,(t-t_1)^2\,, & t < t_1 \\ 0, & t \geq t_1, \end{cases}$$

Tatsächlich ist diese Funktion eine Lösung des Anfangswertproblems.

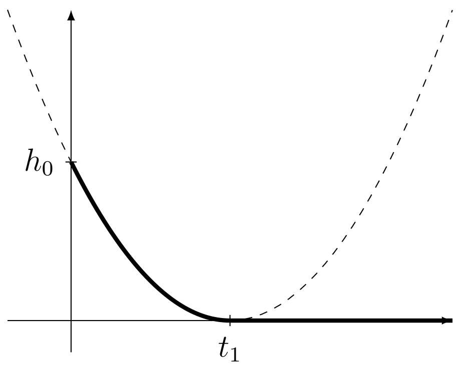

Abbildung 9: Lösung des Anfangswertproblems zum Auslaufen einer Tonne.

Sei $t_0 \geq t_1$. Dann ist obiges h Lösung des Anfangswertproblems

$$h'(t) = -2c\sqrt{h(t)}, \quad h(t_0) = 0,$$

und zwar für jedes $t_1 \in (-\infty, t_0]$. Eine weitere Lösung ist offensichtlich $h(t) \equiv 0$. Die Lösung ist also nicht eindeutig. (Physikalisch bedeutet das, dass wir einer leeren Tonne nicht ansehen, wann sie leergelaufen ist, oder ob sie überhaupt jemals voll war.) Dies ist aber natürlich kein Widerspruch zu Satz 1.2, denn mit den dortigen Bezeichnungen ist $b(y) = \sqrt{y}$ und $y_0 = 0$, also $b(y_0) = 0$.

Definition 1.5. Die Differentialgleichung

$$y'(x) = h\left(\frac{y(x)}{x}\right)$$

heißt *homogene Differentialgleichung*. Dabei ist h auf einem Intervall als stetig vorausgesetzt und es wird eine Lösung y auf einem geeigneten Intervall I mit $0 \notin I$ gesucht.

Bemerkung 1.6. Der Begriff „homogen“ wird auch in anderen Bedeutungen verwandt.

Lösungsverfahren. Man setzt

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

also $y(x) = xz(x)$. Dann gilt $y'(x) = z(x) + xz'(x)$, also

$$z'(x) = \frac{y'(x) - z(x)}{x} = \frac{h(z(x)) - z(x)}{x}.$$

Dies ist eine Differentialgleichung mit getrennten Veränderlichen und alles dort Gesagte gilt sinngemäß auch hier.

Beispiel 1.7. Wir betrachten das Anfangswertproblem

$$y'(x) = \frac{y(x)}{x} - 1 - e^{-y(x)/x}, \quad y(1) = 0.$$

Mit $h(z) = z - 1 - e^{-z}$ und $z = z(x)$ folgt

$$\log x = \int_1^x \frac{dt}{t} = \int_0^z \frac{ds}{h(s) - s} = -\int_0^z \frac{ds}{1 + e^{-s}} = -\left(\log\left(1 + e^z\right) - \log 2\right)$$

also

$$\log(1 + e^z) = \log 2 - \log x = \log \frac{2}{x}.$$

Dies liefert

$$z(x) = z = \log\left(\frac{2}{x} - 1\right)$$

und damit

$$y(x) = x \log\left(\frac{2}{x} - 1\right).$$

Diese Lösung des Anfangswertproblems existiert auf dem Intervall $(0, 2)$.

Definition 1.8. Sei I Intervall, $a, b\colon I \to \mathbb{R}$ stetig. Die Differentialgleichung

$$y'(x) = a(x)y(x) + b(x)$$

heißt *inhomogene lineare Differentialgleichung* 1. *Ordnung*, im Falle $b(x) \equiv 0$ *homogene lineare Differentialgleichung* 1. *Ordnung.*

Lösungsverfahren. Wir betrachten zunächst die homogene Differentialgleichung, also $y'(x) = a(x)y(x)$. Dies ist eine Differentialgleichung mit getrennten Veränderlichen. Mit der Anfangsbedingung $y(x_0) = y_0$ erhalten wir

$$A(x) := \int_{x_0}^x a(t)dt = \int_{x_0}^x \frac{y'(t)}{y(t)}dt = \log \frac{y(x)}{y_0},$$

also $y(x) = y_0 e^{A(x)}$.

Seien nun y_1, y_2 Lösungen der inhomogenen Differentialgleichung. Dann ist $y := y_1 - y_2$ Lösung der homogenen Differentialgleichung. Nun ist die allgemeine Lösung der homogenen Gleichung bekannt, nämlich von der Form $y(x) = ce^{A(x)}$

mit einer Konstanten c. Es genügt also, eine Lösung y_p der inhomogenen Gleichung zu kennen, eine sogenannte *partikuläre* Lösung.

Dazu macht man den Ansatz $y_p(x) = c(x)y_h(x)$, wobei $y_h(x) = e^{A(x)}$ eine Lösung der homogenen Gleichung ist. Man nennt dies auch die Methode der *Variation der Konstanten.* Es gilt dann

$$y_p' = c'y_h + cy_h' = c'y_h + cay_h = c'y_h + ay_p.$$

Aus $y_p' = ay_p + b$ folgt dann $c'y_h = b$. Nimmt man $c(x_0) = 0$ an, so ergibt sich

$$c(x) = \int_{x_0}^{x} \frac{b(t)}{y_h(t)} dt = \int_{x_0}^{x} b(t)e^{-A(t)} dt$$

und damit

$$y_p(x) = c(x)y_h(x) = e^{A(x)} \int_{x_0}^{x} b(t)e^{-A(t)} dt.$$

Insgesamt erhalten wir folgendes Resultat.

Satz 1.9. *Seien I Intervall, $a, b\colon I \to \mathbb{R}$ stetig, $x_0 \in I$ und $A(x) = \int_{x_0}^{x} a(t)dt$. Dann ist für $y_0 \in \mathbb{R}$ durch*

$$y(x) = y_0 e^{A(x)} + e^{A(x)} \int_{x_0}^{x} b(t)e^{-A(t)} dt = e^{A(x)} \left(y_0 + \int_{x_0}^{x} b(t)e^{-A(t)} dt \right)$$

die eindeutige Lösung des Anfangswertproblems $y' = ay + b$, $y(x_0) = y_0$, gegeben.

Beispiel 1.10. Wir betrachten das Anfangswertproblem

$$y'(x) = \frac{1}{x} y(x) + \sqrt{x}, \quad y(1) = 1.$$

Mit

$$A(x) = \int_{1}^{x} \frac{dt}{t} = \log x$$

erhält man auf dem Intervall $(0, \infty)$ die Lösung

$$y(x) = e^{\log x} \left(1 + \int_{1}^{x} \sqrt{t} e^{-\log t} dt \right) = x \left(1 + \int_{1}^{x} \frac{dt}{\sqrt{t}} \right) = x \left(2\sqrt{x} - 1 \right).$$

Definition 1.11. Sei $\Omega \subseteq \mathbb{R}^d$ offen, $f\colon \Omega \to \mathbb{R}^d$ stetig und $U\colon \Omega \to \mathbb{R}$ stetig differenzierbar. Dann heißt U *Potential* (oder *Potentialfunktion*) von f, falls $f = \operatorname{grad} U = (\partial U/\partial x_1, \ldots, \partial U/\partial x_d)$.

Definition 1.12. Sei $\Omega \subseteq \mathbb{R}^2$ offen und $f = (g, h)\colon \Omega \to \mathbb{R}^2$ stetig. Die Differentialgleichung

$$g\left(x, y(x)\right) + h\left(x, y(x)\right) y'(x) = 0$$

heißt *exakt*, wenn $f = (g, h)$ ein Potential hat.

Satz 1.13. *Seien Ω und $f = (g, h)$ wie in Definition* 1.12 *und die dortige Differentialgleichung sei exakt. Sei $(x_0, y_0) \in \Omega$ mit $h(x_0, y_0) \neq 0$. Dann existiert ein offenes Intervall I mit $x_0 \in I$, in dem das Anfangswertproblem*

$$g(x, y(x)) + h(x, y(x))y'(x) = 0, \quad y(x_0) = y_0,$$

eine eindeutige Lösung hat.

Ist U Potential von $f = (g, h)$, so ist $y\colon I \to \mathbb{R}$ genau dann Lösung des Anfangswertproblems, wenn $(x, y(x)) \in \Omega$ und $U(x, y(x)) = U(x_0, y_0)$ für alle $x \in I$.

Beweis. Gilt $U(x, y(x)) = U(x_0, y_0)$, so folgt

$$\begin{aligned} 0 &= \frac{d}{dx} U(x, y(x)) \\ &= \frac{\partial U}{\partial x}(x, y(x)) + \frac{\partial U}{\partial y}(x, y(x))y'(x) \\ &= g(x, y(x)) + h(x, y(x))y'(x). \end{aligned}$$

Gilt außerdem $y(x_0) = y_0$, so löst y das Anfangswertproblem. Umgekehrt gilt auch $U(x, y(x)) = U(x_0, y_0)$ für jede Lösung des Anfangswertproblems.

Die Existenz einer Lösung y unter der Bedingung

$$h(x_0, y_0) = \frac{\partial U}{\partial y}(x_0, y_0) \neq 0$$

folgt aus dem Satz über implizite Funktionen. □

Beispiel 1.14. Gegeben sei das Anfangswertproblem

$$(2xe^y - 1) + (x^2 e^y + 1) \cdot y' = 0, \quad y(0) = 0.$$

Ein Potential ist $U(x, y) = x^2 e^y - x + y$. Wegen $U(0, 0) = 0$ und

$$\frac{\partial U}{\partial y}(0, 0) = x^2 e^y + 1 \Big|_{(x,y)=(1,0)} = 1 \neq 0$$

existiert ein Intervall I mit $0 \in I$ und eine Lösung $y\colon I \to \mathbb{R}$ des Anfangswertproblems. Diese erfüllt

$$x^2 e^{y(x)} - x + y(x) = 0.$$

Explizites Auflösen der Gleichung nach y ist nicht möglich.

Auflösen nach x ist aber möglich und man erhält

$$x = x(y) = \frac{1}{2} e^{-y} \left(1 \pm \sqrt{1 - 4ye^y}\right)$$

für $y < 0{,}20388\ldots$. Wegen $x(0) = 0$ ist für ein Intervall um 0 der negative Wert der Wurzel zu nehmen.

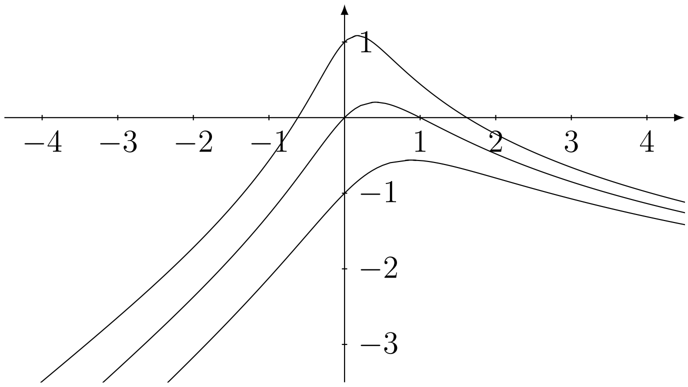

Abbildung 10: Lösungen von $(2xe^y - 1) + (x^2e^y + 1) \cdot y' = 0$ mit den Anfangsbedingungen $y(0) = 0$ und $y(0) = \pm 1$.

Bemerkung 1.15. (i) Sind g und h stetig differenzierbar und hat $f = (g, h)$ ein Potential U, so gilt nach dem Satz von Schwarz

$$\frac{\partial h}{\partial x} = \frac{\partial^2 U}{\partial x \partial y} = \frac{\partial^2 U}{\partial y \partial x} = \frac{\partial g}{\partial y}.$$

Umgekehrt kann man zeigen, dass in konvexen Gebieten im Falle

$$\frac{\partial h}{\partial x} = \frac{\partial g}{\partial y}$$

tatsächlich ein Potential existiert.

(ii) Ist die Differentialgleichung

$$g(x, y) + h(x, y)y' = 0$$

nicht exakt, so kann man versuchen, eine Funktion M zu finden, so dass

$$M(x, y)g(x, y) + M(x, y)h(x, y)y' = 0$$

exakt ist. Es muss dann

$$\frac{\partial}{\partial x}(Mh) = \frac{\partial}{\partial y}(Mg)$$

gelten. Diese Gleichung ist im Allgemeinen nicht lösbar, aber in vielen Fällen findet man durch spezielle Ansätze für M Lösungen, etwa in dem man M nur von x oder nur von y abhängig wählt, oder auch als Funktion von $x^2 + y^2$.

(iii) Man schreibt die Differentialgleichung

$$g(x, y) + h(x, y)\frac{dy}{dx} = 0$$

auch in der symmetrischen Form

$$g(x, y)dx + h(x, y)dy = 0.$$

2 Differentialgleichungssysteme

Bisher haben wir die (explizite) Differentialgleichung

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

betrachtet, wobei $M \subseteq \mathbb{R}^{n+1}$, $f \colon M \to \mathbb{R}$, $I \subseteq \mathbb{R}$ Intervall und $y \colon I \to \mathbb{R}$.

Allgemeiner kann man die Differentialgleichung (1) auch für $y \colon I \to \mathbb{R}^m$, $M \subseteq \mathbb{R}^{mn+1}$ und $f \colon M \to \mathbb{R}^m$ betrachten, mit $m \in \mathbb{N}$. Mit $y = (y_1, \dots, y_m)$ und $f = (f_1, \dots, f_m)$ besteht (1) aus den m Gleichungen

$$y_j^{(n)} = f_j(x, y_1, \dots, y_m, y_1', \dots, y_m', \dots, y_1^{(n-1)}, \dots, y_m^{(n-1)}) \tag{2}$$

für die Funktionen $y_1, \dots, y_m$.

Man nennt (1) bzw. (2) dann (*explizites*) *Differentialgleichungssystem.* Wie im Falle $m = 1$ hat man im Allgemeinen zusätzlich Anfangsbedingungen und spricht dann von einem Anfangswertproblem.

Sei nun die Differentialgleichung (1) mit $m = 1$ gegeben. Sei

$$Y = (Y_1, \dots, Y_n) = (y, y', \dots, y^{(n-1)}).$$

Die Gleichung (1) hat dann die Form

$$Y_n'(x) = f(x, Y_1(x), Y_2(x), \dots, Y_n(x)). \tag{3}$$

Außerdem gilt

$$Y_j'(x) = Y_{j+1}(x), \quad 1 \le j \le n-1. \tag{4}$$

Damit bilden (3) und (4) ein (explizites) Differentialgleichungssystem erster Ordnung. Offensichtlich ist das durch (3) und (4) gegebene System äquivalent zur Gleichung (1).

Somit kann eine explizite Differentialgleichung höherer Ordnung immer auf ein explizites Differentialgleichungssystem erster Ordnung reduziert werden. Es reicht also, letztere zu betrachten.

Beispiel 2.1. Wir betrachten das Differentialgleichungssystem

$$\begin{aligned} x'(t) &= x(t)\left(\alpha - \beta y(t)\right), \\ y'(t) &= y(t)\left(-\gamma + \delta x(t)\right), \end{aligned}$$

wobei $\alpha, \beta, \gamma, \delta > 0$.

Wir interpretieren dies wie folgt: $x(t)$ ist die Größe einer (Beute-)Population zur Zeit t und $y(t)$ die einer (Räuber-)Population. Gilt $y(t) = 0$, so nimmt $x(t)$ mit konstanter Rate zu, aber mit wachsendem $y(t)$ sinkt die Rate, und für $y(t) > \alpha/\beta$ nimmt $x(t)$ ab. Umgekehrt nimmt $y(t)$ mit konstanter Rate ab, wenn keine Beute da ist (das heißt, $x(t) = 0$), aber bei genügend Beute (nämlich für $x(t) > \gamma/\delta$) nimmt $y(t)$ zu. Man spricht daher auch vom Räuber-Beute-Modell. Nach den Mathematikern Lotka und Volterra, die dieses als erste untersuchten, sind die Gleichungen auch als Lotka-Volterra-Gleichungen bekannt.

Eine explizite Lösung des obigen Systems ist nicht möglich. Es lässt sich aber eine Funktion U finden, so dass $t \mapsto U(x(t), y(t))$ für jede Lösung $(x(t), y(t))$ konstant ist.

Dazu schreiben wir mit

$$s(x,y) = -\gamma y + \delta xy \quad \text{und} \quad r(x,y) = \alpha x - \beta xy$$

das obige System zunächst in der Form

$$x' = r(x,y), \quad y' = s(x,y).$$

Dies liefert

$$s(x,y)x' - r(x,y)y' = 0.$$

Setzen wir, rein formal,

$$\frac{dy}{dx} = \frac{\frac{dy}{dt}}{\frac{dx}{dt}} = \frac{y'}{x'},$$

so erhalten wir

$$s(x,y) - r(x,y)\frac{dy}{dx} = 0.$$

Diese Gleichung ist nicht exakt, aber $M(x,y) = 1/xy$ ist ein integrierender Faktor, das heißt,

$$\frac{s(x,y)}{xy} - \frac{r(x,y)}{xy}\frac{dy}{dx} = \left(-\frac{\gamma}{x} + \delta\right) - \left(\frac{\alpha}{y} + \beta\right)\frac{dy}{dx}$$

ist exakt. Ein Potential ist gegeben durch

$$U(x,y) = -\gamma \log x - \alpha \log y + \delta x + \beta y.$$

Nachrechnen zeigt, dass für eine Lösung $(x(t), y(t))$ des Differentialgleichungssystems tatsächlich $U(x(t), y(t))$ konstant ist.

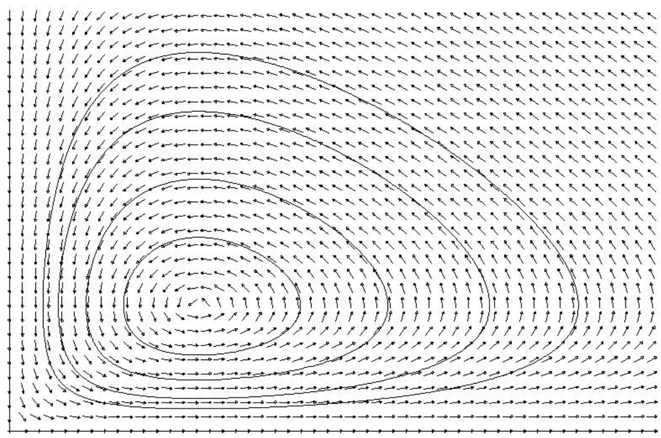

Abbildung 11: Vektorfeld und Lösungskurven des Räuber-Beute-Modells.

Abbildung 11 zeigt für $(\alpha, \beta, \gamma, \delta) = (1, 3, 1, 2)$ einige der Kurven, auf denen $U(x,y)$ konstant ist, sowie das der Abbildung $(x,y) \mapsto (r(x,y), s(x,y))$ entsprechende normierte Vektorfeld. (Normiert bedeutet, dass die Länge der Vektoren normiert wurde.) Die Lösungen $(x(t), y(t))$ sind periodisch in t.

3 Der Satz von Picard-Lindelöf

Wir untersuchen im Folgenden explizite Differentialgleichungen 1. Ordnung, also

$$y'(x) = f(x, y(x)),$$

wobei $f\colon I \times \Omega \to \mathbb{R}^m$ stetig, $I \subseteq \mathbb{R}$ Intervall und $\Omega \subseteq \mathbb{R}^m$. Wir werden I und Ω als offen voraussetzen. Weiter gelte mit $x_0 \in I$ und $y_0 \in \Omega$ die Anfangsbedingung $y(x_0) = y_0$. Gesucht ist eine (stetig differenzierbare) Lösung $y\colon I_0 \to \Omega$, wobei I_0 ein Intervall mit $x_0 \in I_0 \subseteq I$ ist.

Lemma 3.1. *Seien I, Ω, f, x_0 und y_0 wie oben. Sei I_0 Intervall mit $x_0 \in I_0 \subseteq I$ und sei $y\colon I_0 \to \Omega$ stetig differenzierbar. Dann ist y genau dann Lösung des Anfangswertproblems*

$$y'(x) = f(x, y(x)), \quad y(x_0) = y_0,$$

wenn

$$y(x) = y_0 + \int_{x_0}^{x} f(t, y(t))dt$$

für alle $x \in I_0$.

Der *Beweis* folgt unmittelbar aus dem Hauptsatz der Differential- und Integralrechnung.

Auch wenn der Beweis des Lemmas sehr einfach ist, so stellt sich die Aussage als wesentliches Hilfsmittel beim Beweis von Existenzsätzen heraus. Die Idee ist, auf einer geeigneten Funktionenmenge M einen Operator $T\colon M \to M$, $y \mapsto Ty$, zu betrachten, mit

$$(Ty)(x) = y_0 + \int_{x_0}^{x} f(t, y(t))dt$$

Die Lösbarkeit des Anfangswertproblems ist dann äquivalent zur Existenz eines Fixpunktes des Operators T.

Wir erinnern an den Banachschen Fixpunktsatz (Satz 8.44 der Analysis-Vorlesung von Herrn Nieß).

Lemma 3.2. *Sei (M, d) vollständiger metrischer Raum und $T\colon M \to M$. Es existiere $c \in (0, 1)$ mit*

$$d(T(x), T(y)) \leq c \cdot d(x, y)$$

für alle $x, y \in M$. Dann hat T genau einen Fixpunkt, das heißt, es existiert genau ein $\xi \in M$ mit $T(\xi) = \xi$.

Darüberhinaus gilt für alle $x_0 \in X$, dass die rekursiv durch $x_{n+1} = T(x_n)$ definierte Folge (x_n) gegen ξ konvergiert.

Wir wollen einen geeigneten Raum M angeben, auf dem der obige Operator T definiert werden kann. Dazu wählen wir $\delta, r > 0$ mit

$$J_\delta := [x_0 - \delta, x_0 + \delta] \subseteq I \quad \text{und} \quad K_r = \{y \in \mathbb{R}^m \colon \|y - y_0\| \leq r\} \subseteq \Omega.$$

Für eine Teilmenge A von $\mathbb{R}^m$ sei $C(J_\delta, A)$ die Menge der stetigen Funktionen von J_δ nach A. Mit der Supremumsnorm

$$\|u\|_\infty = \max_{x \in J_\delta} \|u(x)\|,$$

ist $C(J_\delta, \mathbb{R}^m)$ ein Banachraum, und $C(J_\delta, K_r)$ ist eine abgeschlossene (und damit vollständige) Teilmenge dieses Banachraums. Durch $u \mapsto Tu$,

$$(Tu)(x) = y_0 + \int_{x_0}^{x} f(t, u(t))dt$$

ist nun eine Abbildung $T\colon C(J_\delta, K_r) \to C(J_\delta, \mathbb{R}^m)$ definiert. Im Allgemeinen gilt aber nicht $T(C(J_\delta, K_r)) \subseteq C(J_\delta, K_r)$.

Lemma 3.3. *Seien J_δ und K_r wie oben. Sei weiter*

$$M_{\delta,r} = \max_{(x,y) \in J_\delta \times K_r} \|f(x, y)\|.$$

Falls $\delta M_{\delta,r} \leq r$, so gilt $T(C(J_\delta, K_r)) \subseteq C(J_\delta, K_r)$.

Bemerkung 3.4. Das Maximum $M_{\delta,r}$ existiert, da f stetig und $J_\delta \times K_r$ kompakt ist. Es ist immer möglich, δ und r mit $\delta M_{\delta,r} \leq r$ zu wählen. Dazu wähle man einfach r beliebig. Wegen $M_{\delta_1,r} \leq M_{\delta_2,r}$ für $\delta_1 < \delta_2$ folgt $\delta M_{\delta,r} \leq r$ für genügend kleine δ.

Beweis von Lemma 3.3. Für $u \in C(J_\delta, K_r)$ und $x \in J_\delta$ gilt

$$\|Tu(x) - y_0\| = \left\| \int_{x_0}^{x} f(t, u(t))dt \right\| \leq \left| \int_{x_0}^{x} \|f(t, u(t))\| \, dt \right| \leq M_{\delta,r}|x - x_0| \leq r,$$

woraus die Behauptung folgt. □

Um den Banachschen Fixpunktsatz anwenden zu können, wird noch eine weitere Voraussetzung benötigt. Wir nehmen an, dass $f(x, y)$ bezüglich y einer Lipschitz-Bedingung genügt, das heißt, dass eine Konstante $L = L_{\delta,r}$ mit

$$\|f(x, y_1) - f(x, y_2)\| \leq L\|y_1 - y_2\|$$

für alle $x \in J_\delta$ und $y_1, y_2 \in K_r$ existiert. Für $u, v \in C(J_\delta, K_r)$ und $x \in J_\delta$ gilt dann

$$\begin{aligned}
\|Tu(x) - Tv(x)\| &= \left\| \int_{x_0}^{x} f(t, u(t))dt - \int_{x_0}^{x} f(t, v(t))dt \right\| \\
&= \left\| \int_{x_0}^{x} \left(f(t, u(t)) - f(t, v(t))\right) dt \right\| \\
&\leq \left| \int_{x_0}^{x} \|f(t, u(t)) - f(t, v(t))\| dt \right| \\
&\leq \left| \int_{x_0}^{x} L\|u(t) - v(t)\| dt \right| \\
&\leq L\|u - v\|_\infty |x - x_0| \\
&\leq \delta L\|u - v\|_\infty.
\end{aligned}$$

Es folgt

$$\|Tu - Tv\|_\infty \leq \delta L \|u - v\|_\infty.$$

Wählt man δ bei festem r klein genug, so gilt $\delta L = \delta L_{\delta,r} < 1$. Dann ist der Banachsche Fixpunktsatz auf T anwendbar. Also hat T einen eindeutig bestimmten Fixpunkt in $C(J_\delta, K_r)$ und damit hat das Anfangswertproblem in J_δ eine eindeutig bestimmte Lösung mit Werten in K_r. (Es folgt auch, dass es keine weitere Lösung gibt, wenn man Werte außerhalb von K_r zulässt.)

Wir fassen die obigen Überlegungen in folgendem Satz zusammen.

Satz 3.5. (Satz von Picard-Lindelöf) *Sei I offenes Intervall, $\Omega \subseteq \mathbb{R}^m$ offen, $f\colon I \times \Omega \to \mathbb{R}^m$ stetig, $x_0 \in I$ und $y_0 \in \Omega$. Es existiere $L \geq 0$ mit*

$$\|f(x, y_1) - f(x, y_2)\| \leq L\|y_1 - y_2\|$$

für alle (x, y) aus einer Umgebung von (x_0, y_0). Dann hat das Anfangswertproblem

$$y'(x) = f(x, y(x)), \quad y(x_0) = y_0,$$

eine eindeutig bestimmte lokale Lösung, d.h., es existiert ein offenes Intervall I_0 mit $x_0 \in I_0 \subseteq I$, in dem das Anfangswertproblem eine eindeutige Lösung hat.

Genauer gilt: Seien $\delta, r > 0$ mit

$$J := [x_0 - \delta, x_0 + \delta] \subseteq I \quad \textit{und} \quad K := \{y \in \mathbb{R}^m\colon \|y - y_0\| \leq r\} \subseteq \Omega$$

und sei

$$M := \max_{(x,y) \in J \times K} \|f(x, y)\|.$$

Weiter gelte

$$\|f(x, y_1) - f(x, y_2)\| \leq L\|y_1 - y_2\|$$

für alle $x \in J$ und $y_1, y_2 \in K$. Gilt dann $\delta M \leq r$ und $\delta L < 1$, so hat das Anfangswertproblem eine eindeutig bestimmt Lösung in J.

Bemerkung 3.6. (i) Die Bedingungen $\delta M \leq r$ und $\delta L < 1$ können durch Verkleinerung von δ immer erreicht werden.

Mit anderen Worten: Sind δ, r, L und M wie im Satz, und ist $\delta' \leq \delta$, $\delta' M \leq r$ und $\delta' L < 1$, so existiert eine eindeutig bestimmte Lösung in $[x_0 - \delta', x_0 + \delta']$. Da man $L, M \neq 0$ annehmen kann, wählt man einfach $\delta' < \min\{r/M, 1/L, \delta\}$.

(ii) Ist f stetig differenzierbar, so ist nach dem Mittelwertsatz die Lipschitzbedingung mit

$$L = \max_{(x,y) \in J \times K} \left\| \frac{\partial f}{\partial y}(x, y) \right\|$$

erfüllt. Dabei ist $\partial f / \partial y$ die Jacobimatrix (oder auch die totale Ableitung) der durch $y \mapsto f(x, y)$ gegebenen Funktion und $\|\partial f / \partial y\|$ die Norm der Matrix (bzw. die Operatornorm).

(iii) Der Beweis zeigt, dass die rekursiv durch $y_0(x) \equiv y_0$ und $y_{n+1} = Ty_n$ definierte Folge (y_n) auf J_δ gegen die Lösung des Anfangswertproblems konvergiert.

(iv) Der Satz gilt analog auch für nichtoffene Intervalle I. Dies ist sofort klar, wenn x_0 im Innern von I liegt, da man einfach I durch sein Inneres ersetzen kann. Entsprechendes gilt aber auch, falls x_0 Randpunkte von I ist. Etwa für $I = [a, b]$ und $x_0 = a$ wählt man dann $J = [x_0, x_0 + \delta]$. Der Beweis geht genauso.

Alternativ kann man den Fall nichtoffener Intervalle auch auf den offener Intervalle zurückführen, indem man etwa für $I = [a, b]$ die Funktion f durch $f(x, y) = f(a, y)$ für $x < a$ und $f(x, y) = f(b, y)$ für $x > b$ zu einer stetigen Funktion $f \colon \mathbb{R} \times \Omega \to \mathbb{R}^m$ fortsetzt und den Satz hierauf anwendet.

(v) Die Voraussetzung, dass Ω offen ist, wurde nur gemacht, damit für alle $y_0 \in \Omega$ ein $r > 0$ mit $K_r \subseteq \Omega$ existiert. Bei gegebenem y_0 reicht es zu fordern, dass f in $I \times K_r$ stetig ist und dort einer Lipschitz-Bedingung genügt.

Beispiel 3.7. Wir hatten in Beispiel 1.3 das Anfangswertproblem

$$y'(x) = x(1 + y(x)^2), \quad y(0) = 1,$$

betrachtet, mit der Lösung

$$y \colon \left(-\sqrt{\frac{\pi}{2}}, \sqrt{\frac{\pi}{2}}\right) \to \mathbb{R}, \quad y(x) = \tan\left(\frac{x^2}{2} + \frac{\pi}{4}\right).$$

Die Differentialgleichung erfüllt die Voraussetzungen des Satzes von Picard-Lindelöf mit $I = \Omega = \mathbb{R}$ und $f \colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $f(x, y) = x(1 + y^2)$. Wir wählen $\delta = r = 1$. Dann gilt $J = [-1, 1]$, $K = [0, 2]$ und

$$M = \max_{(x,y) \in J \times K} |f(x, y)| = 5.$$

Aus dem Mittelwertsatz folgt mit Bemerkung 3.6, (ii), dass wir

$$L = \max_{(x,y) \in J \times K} \left| \frac{\partial f}{\partial y}(x, y) \right| = \max_{(x,y) \in J \times K} |2yx| = 4$$

wählen können. Mit Bemerkung 3.6, (i), liefert der Satz von Picard-Lindelöf eine eindeutig bestimmte Lösung in $[-\delta', \delta']$, falls $\delta' < \min\{\frac{1}{5}, \frac{1}{4}, 1\} = \frac{1}{5}$.

Durch geschickte Wahl von r und δ lässt sich dieses Intervall vergrößern, aber der Satz von Picard-Lindelöf liefert nicht das maximale Intervall.

Beispiel 3.8. Wir betrachten das Anfangswertproblem $y'(x) = \alpha y(x)$, $y(0) = 1$. Wir wissen bereits, dass die eindeutige Lösung durch $y(x) = e^{\alpha x}$ gegeben ist.

Der im Beweis des Satzes von Picard-Lindelöf betrachtet Operator T hat hier die Form

$$(Ty)(x) = 1 + \alpha \int_0^x y(t)dt.$$

Für die in Bemerkung 3.6, (iii), genannte Picard-Lindelöfsche Folge (y_n) erhält man $y_0(x) \equiv 1$,

$$y_1(x) = 1 + \alpha \int_0^x 1 \, dt = 1 + \alpha x,$$
$$y_2(x) = 1 + \alpha \int_0^x (1 + \alpha t)dt = 1 + \alpha x + \frac{(\alpha x)^2}{2}.$$

Durch Induktion sieht man, dass

$$y_n(x) = \sum_{k=0}^{n} \frac{(\alpha x)^k}{k!}.$$

4 Der Existenzsatz von Peano

Der Satz von Picard-Lindelöf besagt, dass die Differentialgleichung

$$y'(x) = f(x, y(x)), \quad y(x_0) = y_0,$$

unter geeigneten Voraussetzungen eine eindeutig bestimmte lokale Lösung besitzt, d.h., es existiert dann ein Intervall I_0 mit $x_0 \in I_0$, in dem die Differentialgleichung eine eindeutig bestimmte Lösung hat. Die entscheidende Voraussetzung ist dabei, dass f bezüglich y einer Lipschitz-Bedingung genügt, d.h.,

$$\|f(x, y) - f(x, \widetilde{y})\| \leq L\|y - \widetilde{y}\|$$

für alle (x, y) und $(x, \widetilde{y})$ aus einer Umgebung von (x_0, y_0).

Ohne diese Voraussetzung ist die Lösung im Allgemeinen nicht eindeutig, wie Beispiel 1.4 (Leerlaufen einer Tonne) zeigt.

Ziel dieses Abschnitts ist der Beweis des Satzes von Peano, der besagt, dass man ohne die Lipschitz-Bedingung noch die Existenz von Lösungen hat.

Satz 4.1. (Satz von Peano) *Sei I offenes Intervall, $\Omega \subseteq \mathbb{R}^m$ offen, $f\colon I \times \Omega \to \mathbb{R}^m$ stetig, $x_0 \in I$ und $y_0 \in \Omega$. Dann hat das Anfangswertproblem*

$$y'(x) = f(x, y(x)), \quad y(x_0) = y_0,$$

eine lokale Lösung, d.h., es existiert ein offenes Intervall I_0 mit $x_0 \in I_0 \subseteq I$, in dem das Anfangswertproblem eine Lösung hat.

Genauer gilt: Sind $\delta, r > 0$ mit

$$J_\delta := [x_0 - \delta, x_0 + \delta] \subseteq I \quad \textit{und} \quad K_r := \{y \in \mathbb{R}^m \colon \|y - y_0\| \leq r\} \subseteq \Omega,$$

und gilt mit

$$M := \max_{(x,y) \in J_\delta \times K_r} \|f(x, y)\|,$$

dass $\delta M \leq r$, so hat das Anfangswertproblem eine Lösung in J_δ.

Bemerkung 4.2. (i) Die Bedingung $\delta M \leq r$ kann durch Verkleinerung von δ immer erreicht werden.

(ii) Im Satz von Picard-Lindelöf wurde zusätzlich noch $\delta L < 1$ verlangt, wobei L die Lipschitz-Konstante war.

Es sei an den Beweis des Satzes von Picard-Lindelöf erinnert. Zunächst wurde gezeigt (Lemma 3.1), dass y genau dann Lösung ist, wenn y Fixpunkt des durch

$$(Ty)(x) = y_0 + \int_{x_0}^{x} f\,(t, y(t))\,dt$$

definierten Operators T ist. Dann wurde gezeigt, dass für geeignete δ, r, M wie angegeben $T\colon C(J_\delta, K_r) \to C(J_\delta, K_r)$ eine Kontraktion ist. Hierauf konnte jetzt der Banachsche Fixpunktsatz (Lemma 3.2) angewandt werden.

Eine Möglichkeit, den Satz von Peano zu beweisen, besteht darin, ähnlich wie dort vorzugehen, aber statt des Banachschen Fixpunktsatzes den (deutlich tiefer liegenden) Fixpunktsatz von Schauder zu benutzen: *Ist X Banachraum, $K \subseteq X$ nicht-leer, kompakt und konvex und ist $T\colon K \to K$ stetig, so hat T einen Fixpunkt.*

Wir werden einen anderen, elementareren Weg beschreiten. Dazu benötigen wir einige Definitionen und Hilfssätze.

Definition 4.3. Seien (X, d_X) und (Y, d_Y) metrische Räume und sei $\mathcal{F}$ eine Menge von Funktionen von X nach Y. Für $x_0 \in X$ heißt $\mathcal{F}$ *gleichgradig stetig in* x_0, falls zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert, so dass $d_Y(f(x), f(x_0)) < \varepsilon$ für $d_X(x, x_0) < \delta$ und alle $f \in \mathcal{F}$.

Für $A \subseteq X$ heißt $\mathcal{F}$ *gleichgradig stetig auf* A, falls zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert, so dass $d_Y(f(x), f(x')) < \varepsilon$ für alle $x, x' \in A$ mit $d_X(x, x') < \delta$ und für alle $f \in \mathcal{F}$.

Bemerkung 4.4. Das Entscheidende ist, dass δ unabhängig von f gewählt werden kann. Bei der gleichgradigen Stetigkeit auf A ist δ zusätzlich noch unabhängig von den Punkten $x, x' \in A$. Dies entspricht der gleichmäßigen Stetigkeit (auf A).

Nun ist eine auf einer kompakten Menge A stetige Funktion gleichmäßig stetig auf A (siehe Satz 6.10 der Vorlesung Analysis I/II von Herrn Nieß). Völlig analog erhält man das folgende Resultat: *Ist A kompakt und $\mathcal{F}$ gleichgradig stetig in jedem Punkt von A, so ist $\mathcal{F}$ gleichgradig stetig auf A.*

Lemma 4.5. *Sei (X, d_X) kompakter metrischer Raum, sei D dichte Teilmenge von X und sei (Y, d_Y) vollständiger metrischer Raum. Sei (f_n) eine (auf X) gleichgradig stetige Folge von Funktionen von X nach Y. Konvergiert $(f_n(x))$ für alle $x \in D$, so ist (f_n) gleichmäßig konvergent.*

Beweis. Sei $\varepsilon > 0$. Sei δ wie in der Definition der gleichgradigen Stetigkeit, also $d_Y(f_n(x), f_n(x')) < \varepsilon$ für alle $x, x' \in X$ mit $d_X(x, x') < \delta$ und alle $n \in \mathbb{N}$. Da D dicht und X kompakt ist, existieren $x_1, \dots, x_N \in D$ mit

$$X \subseteq \bigcup_{j=1}^{N} U_\delta(x_j), \quad \text{wobei } U_\delta(x_j) = \{x \in X\colon d_X(x, x_j) < \delta\}.$$

Für $j \in \{1, \dots, N\}$ existiert nun $n_j \in \mathbb{N}$ mit $d_Y(f_n(x_j), f_m(x_j)) < \varepsilon$ für $n, m \geq n_j$. Wir setzen $n_0 = \max_{j \in \{1,\dots,N\}} n_j$.

Sei nun $x \in X$. Dann existiert x_j mit $x \in U_\delta(x_j)$. Es folgt für $n, m \geq n_0$, dass

$$\begin{aligned} d_Y(f_n(x), f_m(x)) &\leq d_Y(f_n(x), f_n(x_j)) + d_Y(f_n(x_j), f_m(x_j)) + d_Y(f_m(x_j), f_m(x)) \\ &\leq 3\varepsilon. \end{aligned}$$

Damit ist (f_n) Cauchy-Folge im Raum $C(X, Y)$ der stetigen Funktion von X nach Y, versehen mit der durch $d(f, g) = \max_{x \in X} d_Y(f(x), g(x))$ gegebenen Metrik d. Folglich konvergiert (f_n) in $C(X, Y)$, ist also gleichmäßig konvergent. $\square$

Lemma 4.6. *Sei (X, d_X) kompakter metrischer Raum. Dann hat X eine abzählbare dichte Teilmenge von D.*

Beweis. Zu $n \in \mathbb{N}$ existieren $x_1^n, x_2^n, \ldots, x_{k_n}^n \in X$ mit

$$X \subseteq \bigcup_{j=1}^{k_n} U_{1/n}(x_j^n).$$

Damit leistet $D = \{x_j^n : n \in \mathbb{N}, j \in \{1, \ldots, k_n\}\}$ das Verlangte. $\square$

Lemma 4.7. *Seien (X, d_X) und (Y, d_Y) metrische Räume, sei Y kompakt und sei D abzählbare Teilmenge von X. Sei (f_n) Folge von Funktionen von X nach Y. Dann hat (f_n) eine Teilfolge, die auf D (punktweise) konvergiert.*

Beweis. Sei $D = \{d_1, d_2, d_3, \ldots\}$. Da Y kompakt ist, hat die Folge $(f_n(d_1))$ eine konvergente Teilfolge. Wir bezeichnen diese mit $(f_n^1(d_1))$. Die Folge $(f_n^1(d_2))$ hat eine konvergente Teilfolge $(f_n^2(d_2))$. Induktiv erhalten wir so für $k \in \mathbb{N}$ eine Teilfolge (f_n^k) von (f_n) mit der Eigenschaft, dass $(f_n^k(x_j))$ für $1 \leq j \leq k$ konvergiert. Die „Diagonalfolge" (f_n^n) hat dann die Eigenschaft, dass $(f_n^n(x_j))_{n\in\mathbb{N}}$ für alle $j \in \mathbb{N}$ konvergiert. $\square$

Durch Kombination dieser Hilfssätze erhalten wir folgendes Resultat.

Satz 4.8. (Satz von Arzelà-Ascoli) *Seien (X, d_X) und (Y, d_Y) kompakte metrische Räume und sei $\mathcal{F}$ Menge von Funktionen von X nach Y. Ist $\mathcal{F}$ gleichgradig stetig, so hat jede Folge in $\mathcal{F}$ eine (gleichmäßig) konvergente Teilfolge.*

Sind umgekehrt alle Funktionen in $\mathcal{F}$ stetig und besitzt jede Folge in $\mathcal{F}$ eine (gleichmäßig) konvergente Teilfolge, so ist $\mathcal{F}$ gleichgradig stetig.

Beweis. Sei $\mathcal{F}$ gleichgradig stetig. Wir wählen gemäß Lemma 4.6 eine dichte Teilmenge, wenden auf diese Lemma 4.7 an, und benutzen dann Lemma 4.5.

Zur Umkehrung: Ist $\mathcal{F}$ nicht gleichgradig stetig, so existieren $\varepsilon > 0$ und $f_n \in \mathcal{F}$ sowie $x_n, x_n' \in X$ mit $d_Y(f_n(x_n), f_n(x_n')) \geq \varepsilon$ und $d_X(x_n, x_n') \leq 1/k$. Sei nun (f_{n_k}) konvergente Teilfolge von (f_n), etwa $f_{n_k} \to f$. Ohne Einschränkung der Allgemeinheit können wir annehmen, dass auch (x_{n_k}) konvergiert, etwa $x_{n_k} \to x_0$. Dann gilt auch $x_{n_k}' \to x_0$. Es folgt $f_{n_k}(x_{n_k}) \to f(x_0)$ und $f_{n_k}(x_{n_k}') \to f(x_0)$. Andererseits gilt $d_Y(f_{n_k}(x_{n_k}), f_{n_k}(x_{n_k}')) \geq \varepsilon$. Das ist ein Widerspruch. $\square$

Wir kehren nun zum Satz von Peano zurück. Die Idee ist, eine Folge (y_n) von „Näherungslösungen" zu definieren, mit Hilfe des Satzes von Arzelà-Ascoli zu zeigen, dass (y_n) eine konvergente Teilfolge hat, und dann zu zeigen, dass der Grenzwert Lösung ist.

Wie bereits bemerkt, erfüllt eine Löung y die Gleichung

$$y(x) = y_0 + \int_{x_0}^{x} f(t, y(t))\,dt.$$

Für $x_1 < x_2$ gilt dann

$$y(x_2) - y(x_1) = \int_{x_1}^{x_2} f(t, y(t))\,dt.$$

Ist $|x_2 - x_1|$ klein, so weicht $f(t, y(t))$ im Intervall $[x_1, x_2]$ wegen der Stetigkeit von f nur wenig von $f(x_1, y(x_1))$ ab. Es gilt also

$$y(x_2) - y(x_1) \approx (x_2 - x_1) f(x_1, y(x_1)).$$

Alternativ erhält man dies auch aus

$$f(x_1, y(x_1)) = y'(x_1) \approx \frac{y(x_2) - y(x_1)}{x_2 - x_1}.$$

Für eine Zerlegung $(x_0, \dots, x_k)$ des Intervalls $[x_0, x_0 + \delta]$, also $x_0 < x_1 < x_2 < \dots < x_k = x_0 + \delta$ definiert man nun, motiviert durch obige Überlegungen, rekursiv

$$y_{j+1} = y_j + (x_{j+1} - x_j) f(x_j, y_j)$$

für $1 \leq j \leq k$. Sei nun $y \colon [x_0, x_0 + \delta] \to \mathbb{R}^m$ die Funktion, die $y(x_j) = y_j$ für alle $j \in \{0, 1, \dots, n\}$ erfüllt und die in den Intervallen $[x_j, x_{j+1}]$ linear ist.

Es ist zu erwarten, dass für genügend feine Zerlegungen die so erhaltene Funktion y eine gute Annäherung an eine tatsächliche Lösung ist. Wir werden sehen, dass dies wirklich der Fall ist. Man nennt das beschriebene Verfahren zur Gewinnung von Näherungslösungen *Eulersches Polygonzug-Verfahren.* Dieses Verfahren ist das einfachste Verfahren zur numerischen Lösung von Differentialgleichungen. In der Praxis wendet man jedoch in der Regel Verfeinerungen davon an (sogenannte Runge-Kutta-Verfahren). Näheres dazu wird in der Numerischen Mathematik gesagt.

Besonders einfach wird das Verfahren bei äquidistanten Zerlegungen: Für $n \in \mathbb{N}$ und $h = \delta/n$ sei $x_j = x_0 + jh$. Es gilt dann

$$y_{j+1} = y_j + h f(x_j, y_j).$$

Sei y^n die mit dem Eulerschen Polygonzug-Verfahren gewonnene Näherungslösung bei äquidistanter Zerlegung von $[x_0, x_0 + \delta]$ in n Intervalle der Länge $h_n = \delta/n$. Mit $x_j^n = x_0 + jh_n$ und $y_j^n = y^n(x_j)$ ist also

$$y_{j+1}^n = y_j^n + h_n f(x_j^n, y_j^n), \quad y_0^n = y_0.$$

Wir werden zeigen, dass y^n eine auf $[x_0, x_0 + \delta]$ konvergente Teilfolge hat und dass deren Grenzwert eine Löung der Differentialgleichung auf $[x_0, x_0 + \delta]$ ist. Völlig analog erhält man eine Lösung auf $[x_0 - \delta, x_0]$. Hieraus folgt dann die Behauptung.

Zuerst zeigen wir, dass

$$y_j^n \in K_r = \{y \in \mathbb{R}^n \colon \|y - y_0\| \leq r\}$$

für alle $n \in \mathbb{N}$ und $j \in \{0, \dots, n\}$. Wegen

$$\|y_{j+1}^n - y_j^n\| = h_n \|f(x_j^n, y_j^n)\|$$

folgt

$$\|y_{j+1}^n - y_j^n\| \leq h_n M = \frac{\delta}{n} M \leq \frac{r}{n},$$

falls $y_j^n \in K_r$. Wegen $y_0^n = y_0$ erhalten wir aber induktiv

$$\|y_j^n - y_0\| \leq j\frac{r}{n} \leq r$$

und damit schließlich $y_j^n \in K_r$ für alle n und j. Hieraus folgt, dass

$$y^n \colon [x_0, x_0 + \delta] \to K_r.$$

Wir zeigen als Nächstes, dass (y^n) gleichgradig stetig ist. Seien dazu zunächst $x, x' \in [x_j^n, x_{j+1}^n]$. Da y^n im Intervall $[x_j^n, x_{j+1}^n]$ linear ist, folgt

$$\|y^n(x') - y^n(x)\| = \frac{\|y_{j+1}^n - y_j^n\|}{x_{j+1}^n - x_j^n}|x - x'| = \|y_{j+1}^n - y_j^n\|\frac{n}{\delta}|x - x'| \leq \frac{r}{\delta}|x - x'|.$$

Die so erhaltene Abschätzung für $\|y^n(x') - y^n(x)\|$ gilt für alle $x, x' \in [x_0, x_0 + \delta]$, denn ist $x_0 \leq x < x' \leq x_0 + \delta$ und sind j, k mit

$$x_{j-1}^n \leq x < x_j^n \leq x_k^n < x' \leq x_{k+1}^n,$$

so folgt

$$\begin{aligned}
\|y^n(x') - y^n(x)\| \leq \|y^n(x') - y^n(x_k^n)\| + \left\|\sum_{i=j}^{k-1} y^n(x_{i+1}^n) - y^n(x_i^n)\right\| \\
+ \|y^n(x_j^n) - y^n(x)\| \\
&\leq \frac{r}{\delta}\left((x' - x_k^n) + \sum_{i=j}^{k-1}(x_{i+1}^n - x_i^n) + x_j^n - x\right) \\
&= \frac{r}{\delta}(x - x').
\end{aligned}$$

Damit ist (y^n) gleichgradig stetig. Aus dem Satz von Arzelà-Ascoli folgt nun, dass $(y^n)_{n\in\mathbb{N}}$ eine konvergente Teilfolge hat, etwa $y^{n_k} \to y$. Es bleibt zu zeigen, dass y tatsächlich Lösung der Differentialgleichung (auf $[x_0, x_0 + \delta]$) ist.

Äquivalent dazu ist, dass

$$y(x) = y_0 + \int_{x_0}^{x} f(t, y(t))dt$$

für $x \in [x_0, x_0 + \delta]$.

Nun gilt

$$y^n(x_k) = y_k^n = y_0^n + \sum_{j=0}^{k-1}(y_{j+1}^n - y_j^n) = y_0 + h_n \sum_{j=0}^{k-1} f(x_j^n, y_j^n)$$

und

$$y_0 + \int_{x_0}^{x_k} f(t, y^n(t))dt = y_0 + \sum_{j=0}^{k-1} \int_{x_j}^{x_{j+1}} f(t, y^n(t))dt.$$

Zu $\varepsilon > 0$ existiert nun $\eta > 0$ mit

$$\|f(x,y) - f(x',y')\| \leq \varepsilon$$

falls $x, x' \in [x_0, x_0 + \delta]$, $y, y' \in K_r$, $|x - x'| < \eta$ und $\|y - y'\| < \eta$.

Für n groß genug folgt

$$\begin{aligned}
\left\| \int_{x_j}^{x_{j+1}} f(t, y^n(t))dt - h_n f(x_j^n, y_j^n) \right\| &= \left\| \int_{x_j}^{x_{j+1}} \left(f(t, y^n(t)) - f(x_j^n, y_j^n) \right) dt \right\| \\
&\leq \int_{x_j}^{x_{j+1}} \left\| f(t, y^n(t)) - f(x_j^n, y_j^n) \right\| dt \\
&\leq h_n \varepsilon
\end{aligned}$$

und damit

$$\begin{aligned}
\left\| y^n(x_k) - \left(y_0 + \int_{x_0}^{x_k} f(t, y^n(t))dt \right) \right\| &\leq \sum_{j=0}^{k-1} \left\| \int_{x_j}^{x_{j+1}} f(t, y^n(t))dt - h_n f(x_j^n, y_j^n) \right\| \\
&\leq k h_n \varepsilon \\
&= \frac{k}{n} \delta \varepsilon \\
&\leq \delta \varepsilon.
\end{aligned}$$

Für $x \in [x_k, x_{k+1}]$ gilt außerdem

$$\|y^n(x) - y^n(x_k)\| \leq \frac{r}{n}$$

und

$$\left\| \int_{x_k}^{x} f(t, y^n(t))dt \right\| \leq h_n M = \frac{\delta M}{n}.$$

Es folgt

$$\left\| y^n(x) - \left(y_0 + \int_{x_0}^{x} f(t, y^n(t))dt \right) \right\| \leq \delta\varepsilon + \frac{r}{n} + \frac{\delta M}{n}.$$

Mit $n = n_k \to \infty$ erhalten wir

$$\left\| y(x) - \left(y_0 + \int_{x_0}^{x} f(t, y(t))dt \right) \right\| \leq \delta\varepsilon.$$

Da $\varepsilon > 0$ beliebig war, folgt

$$y(x) = y_0 + \int_{x_0}^{x} f(t, y(t))dt.$$

Also löst y die Differentialgleichung auf $[x_0, x_0 + \delta]$. Analog behandelt man das Intervall $[x_0 - \delta, x_0]$ und erhält so eine Lösung auf $[x_0 - \delta, x_0 + \delta]$. Damit ist der Satz von Peano bewiesen.

Bemerkung 4.9. (i) Wir haben den Satz von Peano für offene Intervalle formuliert. Wie beim Satz von Picard-Lindelöf gilt die Aussage auch für nichtoffene Intervalle, auch dann, wenn x_0 Randpunkt ist; vgl. Bemerkung 3.6, (iv).

(ii) Hat die Differentialgleichung eine eindeutige Lösung y auf $[x_0, x_0 + \delta]$, so gilt

$$y(x) = \lim_{n\to\infty} y^n(x) \quad \text{für } x \in [x_0, x_0 + \delta],$$

d.h., das Euler-Polygonzug-Verfahren konvergiert gegen diese Lösung. Denn wäre dies nicht der Fall, so existierte eine Folge (n_k) mit $n_k \to \infty$ und $\varepsilon > 0$ mit

$$\|y^{n_k} - y\|_\infty = \max_{x\in[x_0,x_0+\delta]} |y^{n_k}(x) - y(x)| \geq \varepsilon.$$

Nun hat aber auch (y^{n_k}) nach dem Satz von Arzelà-Ascoli eine konvergente Teilfolge, etwa $y^{n_{k_j}} \to \widetilde{y}$. Es gilt dann $\|\widetilde{y} - y\|_\infty \geq \varepsilon$. Andererseits ist auch $\widetilde{y}$ Lösung. Das ist ein Widerspruch zur Eindeutigkeit.

(iii) Man kann aber auch direkt nachweisen, dass (y^n) konvergiert, falls f bezüglich y einer Lipschitz-Bedingung mit Lipschitz-Konstante L genügt, mit $L\delta < 1$ wie im Satz von Picard-Lindelöf.

5 Abhängigkeit von Parametern

Gegeben seien zwei Anfangswertprobleme

$$y'(x) = f(x, y(x)), \quad y(x_0) = y_0$$

und

$$y'(x) = \widetilde{f}(x, y(x)), \quad y(x_0) = \widetilde{y}_0.$$

Weichen f und $\widetilde{f}$ sowie y_0 und $\widetilde{y}_0$ nur wenig voneinander ab, so wird man erwarten, dass Lösungen $y(x)$ und $\widetilde{y}(x)$ dieser Anfangswertprobleme auch nur wenig voneinander abweichen. Allerdings kann man das nur dann erwarten, wenn die Differentialgleichungen eine eindeutige Lösung haben, denn sonst kann ja auch für $f = \widetilde{f}$ und $y_0 = \widetilde{y}_0$ durchaus $y(x) \neq \widetilde{y}(x)$ gelten.

Wie vorher seien $\delta, r > 0$,

$$J := J_\delta = [x_0 - \delta, x_0 + \delta] \quad \text{und} \quad K := K_r = \{y \in \mathbb{R}^m \colon \|y - y_0\| \leq r\}.$$

Es sei f stetig in $J \times K$ und

$$M := \max_{(x,y)\in J\times K} \|f(x, y)\|.$$

Weiter sei $\widetilde{y}_0 \in \mathbb{R}^m$ mit $\|\widetilde{y}_0 - y_0\| = \varepsilon_0 < r$. Mit $\widetilde{r} = r - \varepsilon_0$ gilt dann

$$\widetilde{K} := \{y \colon \|y - \widetilde{y}_0\| \leq \widetilde{r}\} \subseteq K.$$

Sei $\widetilde{f}$ stetig in $J \times \widetilde{K}$ mit

$$\max_{(x,y)\in J\times\widetilde{K}} \|\widetilde{f}(x, y) - f(x, y)\| \leq \varepsilon_f.$$

Dann gilt

$$\widetilde{M} := \max_{(x,y)\in J\times\widetilde{K}} \|\widetilde{f}(x,y)\| \leq M + \varepsilon_f.$$

Es gelte $\delta M \leq r$ und $\delta\widetilde{M} \leq \widetilde{r}$. Dann hat

$$y'(x) = f(x, y(x)), \quad y(x_0) = y_0$$

eine Lösung y auf J_δ und

$$\widetilde{y}'(x) = f(x, \widetilde{y}(x)), \quad \widetilde{y}(x_0) = \widetilde{y}_0$$

eine Lösung $\widetilde{y}$ auf J_δ.

Satz 5.1. *Seien $x_0, \delta, y_0, \widetilde{y}_0, r, \widetilde{r}, f, \widetilde{f}, M, \widetilde{M}$ wie oben. Es existiere $L \geq 0$ mit*

$$\|f(x,y_1) - f(x,y_2)\| \leq L\|y_1 - y_2\| \quad \text{für } y_1, y_2 \in K_r \text{ und } x \in J_\delta.$$

Dann gilt

$$\|\widetilde{y}(x) - y(x)\| \leq (\varepsilon_0 + \delta\varepsilon_f)e^{L|x-x_0|} \quad \text{für } x \in J_\delta.$$

Zur Vorbereitung des Beweises notieren wir zunächst, dass mit den obigen Bezeichnungen nach Lemma 3.1

$$\begin{aligned}
\widetilde{y}(x) - y(x) &= \widetilde{y}_0 - y_0 + \int_{x_0}^x \Big(\widetilde{f}(x, \widetilde{y}(t)) - f(x, y(t))\Big)\, dt \\
&= \widetilde{y}_0 - y_0 + \int_{x_0}^x \Big(\widetilde{f}(x, \widetilde{y}(t)) - f(x, \widetilde{y}(t))\Big)\, dt \\
&\quad + \int_{x_0}^x \Big(f(x, \widetilde{y}(t)) - f(x, y(t))\Big) dt
\end{aligned}$$

gilt. Es folgt

$$\begin{aligned}
\|\widetilde{y}(x) - y(x)\| &\leq \|\widetilde{y}_0 - y_0\| + \left|\int_{x_0}^x \varepsilon_f dt\right| + \left|\int_{x_0}^x L\|\widetilde{y}(t) - y(t)\| dt\right| \\
&\leq \varepsilon_0 + \delta\varepsilon_f + L\left|\int_{x_0}^x \|\widetilde{y}(t) - y(t)\| dt\right|.
\end{aligned}$$

Hierauf wenden wir jetzt folgenden Lemma an.

Lemma 5.2. (Gronwallsche Ungleichung) *Sei I Intervall, $g\colon I \to [0,\infty)$ stetig uns $x_0 \in I$. Es gebe Konstanten $A, B \geq 0$ mit*

$$g(x) \leq A + B\left|\int_{x_0}^x g(t) dt\right|.$$

Dann gilt

$$g(x) \leq Ae^{B|x-x_0|}.$$

Beweis. Zunächst sei $A > 0$ und $x \geq x_0$. Dann können die Betragsstriche in Voraussetzung und Behauptung entfallen. Sei

$$h(x) = A + B \int_{x_0}^{x} g(t)dt,$$

so dass also $g(x) \leq h(x)$ für $x \geq 0$. Nun gilt $h'(x) = Bg(x)$ nach Hauptsatz der Differential- und Integralrechnung, also $h'(x) \leq Bh(x)$ und damit

$$\frac{h'(x)}{h(x)} \leq B.$$

Es folgt

$$\log h(x) - \log h(x_0) = \int_{x_0}^{x} \frac{h'(t)}{h(t)} dt \leq B \int_{x_0}^{x} dt = B(x - x_0),$$

mit $h(x_0) = A$ also

$$g(x) \leq h(x) = Ae^{B(x-x_0)}.$$

Der Fall $x < x_0$ ist analog.

Betrachten wir eine Folge von A-Werten, die gegen 0 konvergiert, so erhalten wir die Behauptung auch für $A = 0$. □

Um den Beweis von Satz 5.1 abzuschließen, müssen wir die Gronwallsche Ungleichung nun lediglich auf

$$g(x) = \|\widetilde{y}(x) - y(x)\|, \quad A = \varepsilon_0 + \delta\varepsilon_f \quad \text{und} \quad B = L$$

anwenden.

Bemerkung 5.3. Die in Lemma 5.2 formulierte Ungleichung ist Spezialfall einer allgemeineren (ebenfalls auf Gronwall zurückgehenden und nach ihm benannten) Ungleichung. Sind $\alpha, \beta, g \colon I \to [0, \infty)$ stetig und ist $x_0 \in I$ und gilt

$$g(x) \leq \alpha(x) + \left| \int_{x_0}^{x} \beta(t)g(t)dt \right|,$$

so gilt

$$g(x) \leq \alpha(x) + \left| \int_{x_0}^{x} \alpha(t)\beta(t) \exp\left(\left| \int_{t}^{x} \beta(s)ds \right| \right) dt \right|.$$

Der Beweis ist ähnlich.

6 Lineare Differentialgleichungssysteme

Wir betrachten das Differentialgleichungssystem

$$\begin{aligned} y_1'(x) &= a_{11}(x)y_1(x) + \cdots + a_{1n}(x)y_n(x) + b_1(x), \\ &\vdots \\ y_n'(x) &= a_{n1}(x)y_1(x) + \cdots + a_{nn}(x)y_n(x) + b_n(x), \end{aligned}$$

wobei die a_{jk} und die b_j gegebene, auf einem Intervall I stetige Funktionen mit Werten in $\mathbb{R}$ (oder $\mathbb{C}$) sind, und die y_j gesucht sind. Mit der $(n \times n)$-Matrix $A = (a_{jk})$ und den (Spalten-)Vektoren

$$b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \quad \text{und} \quad y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

erhält das Differentialgleichungssystem die Form

$$y'(x) = A(x)y(x) + b(x).$$

Ein Differentialgleichungssystem dieser Art heißt (n-dimensionales) *lineares Differentialgleichungssystem* (auf I). Im Falle $b = 0$ heißt das System *homogen*, anderenfalls *inhomogen*.

Satz 6.1. *Sei I Intervall, $A\colon I \to \mathbb{C}^{n\times n}$ und $b\colon I \to \mathbb{C}^n$ stetig, $x_0 \in I$ und $y_0 \in \mathbb{C}^n$. Dann hat das Anfangswertproblem*

$$y'(x) = A(x)y(x) + b(x), \quad y(x_0) = y_0,$$

eine eindeutige Lösung auf I.

Beweis. Es genügt, den reellwertigen Fall zu betrachten, da ein komplexwertiges System durch Übergang zu Real- und Imaginärteil in ein reellwertiges System der Dimension $2n$ überführt werden kann.

Das Anfangswertproblem hat die Form aus dem Satz von Picard-Lindelöf mit

$$f(x, y) = A(x)y + b(x).$$

Sei I_1 kompaktes Teilintervall von I. Mit

$$L_1 = \max_{x\in I_1} \|A(x)\|,$$

wobei $\|A(x)\|$ die Operatornorm der Matrix $A(x)$ ist, folgt

$$\|f(x, \tilde{y}) - f(x, y)\| = \|A(x)(\tilde{y} - y)\| \le \|A(x)\| \cdot \|\tilde{y} - y\| \le L_1\|\tilde{y} - y\|$$

für $y, \tilde{y} \in \mathbb{R}^n$ und $x \in I$. Nach dem Satz von Picard-Lindelöf existiert zu $x_1 \in I_1$ und $y_1 \in \mathbb{R}^n$ damit $\delta_1 > 0$, so dass das Anfangswertproblem

$$y'(x) = A(x)y(x) + b(x), \quad y(x_1) = y_1,$$

eine eindeutige Lösung auf $[x_1 - \delta_1, x_1 + \delta_1] \cap I_1$ hat.

Wir zeigen, dass δ_1 unabhängig von x_1 und y_1 gewählt werden kann, solange $x_1 \in I_1$. Dazu sei

$$r_1 = \|y_1\|+1, \quad K_1 = \{y \in \mathbb{R}^n\colon \|y-y_1\| \le r_1\} \quad \text{und} \quad M_1 = \max_{(x,y)\in I_1\times K_1} \|f(x, y)\|.$$

Nach dem Satz von Picard-Lindelöf leistet δ_1 das Verlangte, falls $\delta_1 L_1 < 1$ und $\delta_1 M_1 \le r_1$. Mit

$$C = \max_{x\in I_1} \|b(x)\|$$

gilt nun

$$M_1 \leq L_1 \max_{y \in K_1} \|y\| + C \leq L_1(2\|y_1\| + 1) + C = 2L_1\|y_1\| + L_1 + C.$$

Die Ungleichung $\delta_1 M_1 \leq r_1 = \|y_1\| + 1$ gilt also, falls

$$2\delta_1 L_1 \leq 1 \quad \text{und} \quad \delta_1(L_1 + C) \leq 1.$$

gilt. Beispielsweise können wir $\delta_1 = 1/(2L_1{+}C{+}1)$ wählen. (Auch die Bedingung $\delta_1 L_1 < 1$ ist dann erfüllt.)

Sei nun I_1 so gewählt, dass $x_0 \in I_1$. Man erhält eine eindeutige Lösung y^0 in $[x_0 - \delta_1, x_0 + \delta_1] \cap I_1$. Ist etwa $x_1 := x_0 + \delta_1 \in I_1$, liefert das Anfangswertproblem mit der Anfangsbedingung $y(x_1) = y^0(x_1)$ eine Lösung y^1 auf $[x_1{-}\delta_1, x_1{+}\delta_1] \cap I_1$. Auf $[x_0, x_1]$ stimmen y^0 und y^1 überein. Damit erhält man eine Lösung auf $[x_0 - \delta_1, x_1 + \delta_1] \cap I_1 = [x_0 - \delta_1, x_0 + 2\delta_1] \cap I_1$. Induktiv erhält man eine Lösung auf I_1. Da I durch kompakte Intervalle ausgeschöpft werden kann, erhält man schließlich eine Lösung auf I. □

Wie im eindimensionalen Fall betrachten wir zuerst homogene Gleichungen.

Satz 6.2. *Die Lösungsmenge eines n-dimensionalen homogenen Differentialgleichungssystems bildet einen n-dimensionalen Vektorraum.*

Beweis. Man sieht leicht ein, dass die Lösungsmenge L ein Vektorraum ist. Weiter ist $T\colon L \to \mathbb{R}^n$ bzw. $T\colon L \to \mathbb{C}^n$, $y \mapsto y(x_0)$, nach Satz 6.1 eine bijektive Abbildung. Diese ist offensichtlich linear, also ein Isomorphismus. □

Seien $y_1, \ldots, y_n$ Lösungen des homogenen Systems $y' = Ay$. Wir fassen diese zu einer Matrix Y zusammen, also

$$Y = (y_1, \ldots, y_n) = \begin{pmatrix} y_{1,1} & \cdots & y_{n,1} \\ \vdots & & \vdots \\ y_{1,n} & \cdots & y_{n,n} \end{pmatrix}$$

wobei $y_{j,k}$ die k-te Komponente von y_j ist. Man nennt

$$W(x) = W(y_1, \ldots, y_n)(x) = \det Y(x)$$

auch *Wronski-Determinante* von Y.

Sind $y_1, \ldots, y_n$ linear unabhängig sind (und damit nach Satz 6.2 Basis des Lösungsraums), so nennt man $y_1, \ldots, y_n$ *Fundamentalsystem* und Y *Fundamentalmatrix.* Ist Y Fundamentalmatrix und $c = (c_1, \ldots, c_n)^t \in \mathbb{R}^n$ (oder $\mathbb{C}^n$), so ist

$$Yc = \sum_{j=1}^{n} c_j y_j$$

Lösung der Differentialgleichung, und alle Lösungen sind von dieser Form.

Sei nun z_k die Lösung des homogenen Anfangswertproblem mit der Anfangsbedingung $z_k(x_0) = e_k$, wobei e_k der k-te Einheitsvektor ist. Sei $Z = (z_1, \ldots, z_n)$ die zugehörige Fundamentalmatrix. Es ist also $Z(x_0)$ die Einheitsmatrix. Wir nennen Z *Standardfundamentalmatrix* (bezüglich x_0). Aus obigen Überlegungen ergibt sich folgendes Resultat.

Satz 6.3. *Sei Z Standardfundamentalmatrix des Differentialgleichungssystems $y' = Ay$ bezüglich x_0. Sei $y_0 \in \mathbb{R}^n$. Dann hat das Anfangswertproblem*

$$y' = Ay, \quad y(x_0) = y_0,$$

die Lösung $y(x) = Z(x)y_0$.

Sind $y_1, \ldots, y_n$ Lösungen von $y' = Ay$ und ist $Z = (z_1, \ldots, z_n)$ die Standardfundamentalmatrix, so existieren $c_{jk} \in \mathbb{R}$, $j, k \in \{1, \ldots, n\}$ mit $y_k = \sum_{j=1}^n z_j c_{jk}$. Mit $C = (c_{jk})$ und $Y = (y_1, \ldots, y_n)$ folgt also $Y = ZC$. Ist C invertierbar, so ist auch Y Fundamentalmatrix.

Satz 6.4. *Sei I Intervall und $y' = Ay$ homogenes Differentialgleichungssystem auf I. Seien $y_1, \ldots, y_n$ Lösungen, $Y = (y_1, \ldots, y_n)$ und $x, x_0 \in I$. Dann gilt*

$$W(x) = W(x_0) \exp\left(\int_{x_0}^{x} \mathrm{Sp}(A(t))dt \right).$$

Dabei ist $\mathrm{Sp}(A) = \sum_{j=1}^n a_{jj}$ *die Spur der Matrix $A = (a_{ij})$.*

Beweis. Wir zeigen, dass

$$W'(x) = \mathrm{Sp}(A(x))W(x)$$

für alle $x \in I$. Hieraus folgt die Behauptung. Sei dazu $x_1 \in I$ fest und Z die Standardfundamentalmatrix bezüglich x_1. Dann gilt $Y(x) = Z(x)C$ mit einer konstanten Matrix C, und da $Z(x_1)$ die Einheitsmatrix ist, folgt $C = Y(x_1)$.

Mit Hilfe der Darstellung

$$\det Z = \det(z_1, \ldots, z_n) = \sum_{\sigma \in S_n} \mathrm{sign}(\sigma) z_{1,\sigma(1)} \cdot z_{2,\sigma(2)} \cdot \ldots \cdot z_{n,\sigma(n)},$$

wobei S_n die Menge der Permutationen von $\{1, \ldots, n\}$ ist, folgt

$$\begin{aligned}
(\det Z)' &= \sum_{k=1}^n \det(z_1, \ldots, z_{k-1}, z_k', z_{k+1}, \ldots z_n) \\
&= \sum_{k=1}^n \det(z_1, \ldots, z_{k-1}, Az_k, z_{k+1}, \ldots, z_n).
\end{aligned}$$

Es folgt mit $z_k(x_1) = e_k$, dass

$$\begin{aligned}
(\det Z(x))'\big|_{x=x_1} &= \sum_{k=1}^n \det(e_1, \ldots, e_{k-1}, Ae_k, e_{k+1}, \ldots, e_n)\big|_{x=x_1} \\
&= \sum_{k=1}^n a_{kk}(x_1) \\
&= \mathrm{Sp}(A(x_1)).
\end{aligned}$$

Außerdem gilt

$$\begin{aligned}(\det Y(x))' &= (\det(Z(x)Y(x_1)))' \\ &= (\det Z(x) \cdot \det Y(x_1))' \\ &= (\det Z(x))' \cdot \det Y(x_1),\end{aligned}$$

also

$$W'(x) = (\det Z(x))'\big|_{x=x_1} \cdot W(x_1) = \mathrm{Sp}(A(x_1))W(x_1).$$

Die Behauptung folgt. □

Satz 6.5. *Seien $y_1, \ldots, y_n$ Lösungen des homogenen Differentialgleichungssystems $y' = Ay$ auf I. Dann sind äquivalent:*

(i) *$y_1, \ldots, y_n$ bilden ein Fundamentalsystem;*

(ii) *$W(x) \neq 0$ für ein $x \in I$;*

(iii) *$W(x) \neq 0$ für alle $x \in I$.*

Beweis. Die Äquivalenz (ii) $\Leftrightarrow$ (iii) folgt aus Satz 6.4. Offensichtlich gilt auch (iii) $\Rightarrow$ (i), denn sind $y_1, \ldots, y_n$ linear abhängig, so gilt $W(x) \equiv 0$.

Wir zeigen nun, dass (i) $\Rightarrow$ (iii). Sei dazu $y_1, \ldots, y_n$ Fundamentalsystem und es existiere $x_0 \in I$ mit $W(x_0) = 0$. Dann existieren $c_1, \ldots, c_n \in \mathbb{R}$ mit $\sum_{j=1}^n c_j y_j(x_0) = 0$. Nun ist aber $y(x) = \sum_{j=1}^n c_j y_j(x)$ Lösung des Anfangswertproblems $y' = Ay$, $y(x_0) = 0$. Es folgt $y(x) = 0$ für alle $x \in I$, im Widerspruch zur linearen Unabhängigkeit der y_k. □

Die allgemeine Lösung des inhomogenen Systems $y' = Ay + b$ gewinnt man wie im eindimensionalen Fall dadurch, dass man die allgemeine Lösung des homogenen Systems zu einer speziellen (partikulären) Lösung des inhomogenen Systems addiert. Eine solche kann man wieder durch „Variation der Konstanten“ gewinnen, also durch den Ansatz

$$y(x) = \sum_{k=1}^n c_k(x) y_k(x)$$

wobei $y_1, \ldots, y_k$ ein Fundamentalsystem ist.

Wir geben nur das Ergebnis an (vgl. Satz 1.9).

Satz 6.6. *Sei Y Standardfundamentalmatrix des homogenen Systems $y' = Ay$ auf I (zu $x_0 \in I$). Dann ist die Lösung des inhomogenen Anfangswertproblems $y' = Ay + b$, $y(x_0) = y_0$ durch*

$$y(x) = Y(x)y_0 + Y(x)\int_{x_0}^x Y^{-1}(t)b(t)dt = Y(x)\left(y_0 + \int_{x_0}^x Y^{-1}(t)b(t)dt\right)$$

gegeben.

Beweis. Die Anfangsbedingung ist offensichtlich erfüllt. Weiter gilt

$$\begin{aligned}
y'(x) &= Y'(x)\left(y_0 + \int_{x_0}^{x} Y^{-1}(t)b(t)dt\right) + Y(x)Y^{-1}(x)b(x) \\
&= A(x)Y(x)\left(y_0 + \int_{x_0}^{x} Y^{-1}(t)b(t)dt\right) + b(x) \\
&= A(x)y(x) + b(x).
\end{aligned}$$

□

Bemerkung 6.7. Auch wenn Y nur Fundamentalmatrix ist (aber nicht Standardfundamentalmatrix), ist y wie in Satz 6.6 partikuläre Lösung der inhomogenen Differentialgleichung, löst aber nicht das Anfangswertproblem.

Die Berechnung einer partikulären Lösung mit Hilfe von Satz 6.6 setzt voraus, dass man ein Fundamentalsystem des homogenen Systems kennt. Es gibt aber *keine* Methode, mit der man ein solches Fundamentalsystem im Allgemeinen bestimmen kann. Sind einzelne Lösungen bekannt, kann man jedoch die Dimension reduzieren. Diese Methode wird im Folgenden beschrieben.

Reduktionsverfahren von d'Alembert. Sei $v \not\equiv 0$ eine Lösung des n-dimensionalen homogenen Differentialgleichungssystems $y' = Ay$ auf I. Es gelte also $v' = Av$. Wir machen für die Lösung y den Ansatz

$$y(x) = \phi(x)v(x) + z(x) \quad \text{mit } z(x) = \begin{pmatrix} 0 \\ z_2(x) \\ \vdots \\ z_n(x) \end{pmatrix},$$

also $z_1(x) \equiv 0$, wobei $\phi\colon I \to \mathbb{R}$ stetig differenzierbar sei. Es folgt

$$y' = \phi' v + \phi v' + z' = \phi' v + \phi A v + z'.$$

Da wir annehmen, dass auch y Lösung ist, gilt

$$y' = Ay = A(\phi v + z) = \phi A v + Az.$$

Aus diesen beiden Gleichungen erhält man

$$\phi' v + z' = Az.$$

Wegen $z_1 = 0$ folgt

$$\phi' v_k + z_k' = \sum_{j=2}^{n} a_{kj} z_j$$

für $k = 1, \ldots, n$. Für $k = 1$ erhält man im Falle $v_1(x) \neq 0$, dass

$$\phi' = \frac{1}{v_1} \sum_{j=2}^{n} a_{1j} z_j.$$

Es folgt für $k = 2, \ldots, n$, dass

$$z_k' = \sum_{j=2}^{n} a_{kj} z_j - \phi' v_k = \sum_{j=2}^{n} a_{kj} z_j - v_k \frac{1}{v_1} \sum_{j=2}^{n} a_{1j} z_j = \sum_{j=2}^{n} \left(a_{kj} - \frac{v_k}{v_1} a_{1j}\right) z_j.$$

Dieses ist ein $(n-1)$-dimensionales, homogenes System für $z_2, \ldots, z_n$.

Bemerkung 6.8. Wir haben vorausgesetzt, dass $v_1(x) \neq 0$ gilt. Da wir $v \not\equiv 0$ annehmen, existiert zu $x_0 \in I$ aber immer $k \in \{1, \dots, n\}$ und ein Intervall I_0 mit $x_0 \in I_0$, so dass $v_k(x) \neq 0$ für $x \in I_0$. Man kann die entsprechende Reduktion machen, indem man v_1 durch v_k ersetzt.

Beispiel 6.9. Wir betrachten auf $I = (0, \infty)$ das System

$$\begin{aligned} y_1' &= \frac{1}{x} y_1 - y_2, \\ y_2' &= \frac{1}{x^2} y_1 + \frac{2}{x} y_2. \end{aligned}$$

Es gilt also

$$y' = Ay \quad \text{mit } A = \begin{pmatrix} \frac{1}{x} & -1 \\ \frac{1}{x^2} & \frac{2}{x} \end{pmatrix}$$

Eine Lösung ist gegeben durch

$$v(x) = \begin{pmatrix} x^2 \\ -x \end{pmatrix}.$$

Für z_2 erhalten wir die Differentialgleichung

$$z_2'(x) = \left(a_{22}(x) - \frac{v_2(x)}{v_1(x)} a_{12}(x) \right) z_2(x) = \left(\frac{2}{x} - \frac{(-x)}{x^2}(-1) \right) z_2(x) = \frac{1}{x} z_2(x).$$

Eine Lösung ist $z_2(x) = x$.

Mit

$$\phi'(x) = \frac{1}{v_1(x)} a_{12}(x) z_2(x) = \frac{1}{x^2}(-1)x = -\frac{1}{x}$$

können wir also

$$\phi(x) = -\log x$$

wählen und erhalten mit

$$y(x) = \phi(x)v(x) + z(x) = -\log x \begin{pmatrix} x^2 \\ -x \end{pmatrix} + \begin{pmatrix} 0 \\ x \end{pmatrix} = \begin{pmatrix} -x^2 \log x \\ x \log x + x \end{pmatrix}$$

eine weitere Lösung der Differentialgleichung. Eine Fundamentalmatrix ist also

$$\begin{pmatrix} x^2 & -x^2 \log x \\ -x & x \log x + x \end{pmatrix}.$$

Für die Wronski-Determinante $W(x)$ erhalten wir

$$W(x) = x^2(x \log x + x) - x^3 \log x = x^3.$$

Tatsächlich gilt (vgl. Satz 6.4 und seinen Beweis)

$$W'(x) = 3x^2 = \frac{3}{x} W(x) = \mathrm{Sp}(A(x)) W(x).$$

7 Lineare Differentialgleichungssysteme mit konstanten Koeffizienten

Wir betrachten das homogene Differentialgleichungssystem $y' = Ay$, wobei A eine konstante (reelle oder komplexe) $(n \times n)$-Matrix ist. Zur Motivation beginnen wir mit dem Picard-Lindelöf-Verfahren, von dem wir nach Satz 6.1 ja wissen, dass es konvergiert. Seien also $x_0 \in \mathbb{R}, y_0 \in \mathbb{R}^n$ (oder $\mathbb{C}^n$), und mit

$$Ty = y_0 + \int_{x_0}^{x} Ay(t)dt = y_0 + A\int_{x_0}^{x} y(t)dt$$

sei durch

$$y^0(x) \equiv y_0, \quad y^{n+1}(x) = (Ty^n)(x)$$

die Picard-Lindelöfsche Folge gegeben. Es gilt dann

$$y^1(x) = y_0 + A\int_{x_0}^{x} y_0 dt = y_0 + (x - x_0)Ay_0 = (E + (x - x_0)A)y_0.$$

Dabei ist E die Einheitsmatrix. (Wenn wir die Dimension n betonen wollen, schreiben wir auch E_n statt E.) Es folgt

$$\begin{aligned} y^2(x) &= y_0 + A\int_{x_0}^{x} (E + (t - x_0)A)y_0 dt \\ &= y_0 + A\left((x - x_0)E + \frac{(x - x_0)^2}{2}A\right) y_0 \\ &= \left(E + (x - x_0)A + \frac{(x - x_0)^2}{2}A^2\right) y_0. \end{aligned}$$

Analog zu Beispiel 3.8 erhält man induktiv

$$y^k(x) = \left(E + (x - x_0)A + \ldots + \frac{(x - x_0)^k}{k!}A^k\right) y_0.$$

Für eine quadratrische Matrix B definiert man nun

$$\exp(B) := \sum_{k=0}^{\infty} \frac{1}{k!}B^k.$$

Dabei ist $B^0 = E$. Die Reihe konvergiert, denn es gilt

$$\left\|\frac{1}{k!}B^k\right\| = \frac{1}{k!}\left\|B^k\right\| \leq \frac{1}{k!}\|B\|^k.$$

Wir erhalten folgendes Ergebnis.

Satz 7.1. *Das homogene Anfangswertproblem*

$$y' = Ay, \quad y(x_0) = x_0,$$

mit $x_0 \in \mathbb{R}$, $y_0 \in \mathbb{R}^n$ (oder $\mathbb{C}^n$) und einer reellen (oder komplexen) $(n \times n)$-Matrix A hat die (eindeutige) Lösung

$$y(x) = \exp((x - x_0)A)y_0.$$

Es verbleibt die Frage, wie man $\exp(B)$ für eine gegebene Matrix B berechnet. Zunächst überlegt man sich, dass für eine invertierbare Matrix S und $C = S^{-1}BS$ die Gleichung

$$C^k = S^{-1}BSS^{-1}BS \cdots S^{-1}BS = S^{-1}B^kS$$

gilt. Hieraus folgt, dass

$$\exp(C) = S^{-1}\exp(B)S \quad \text{sowie} \quad \exp(B) = S\exp(C)S^{-1}.$$

Am einfachsten ist der Fall einer Diagonalmatrix

$$D = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix}$$

Hier gilt

$$D^k = \begin{pmatrix} d_1^k & & 0 \\ & \ddots & \\ 0 & & d_n^k \end{pmatrix} \quad \text{und} \quad \exp D = \begin{pmatrix} e^{d_1} & & 0 \\ & \ddots & \\ 0 & & e^{d_n} \end{pmatrix}.$$

Im Allgemeinen findet man aber für gegebenes B keine invertierbare Matrix S, so dass $S^{-1}BS$ Diagonalmatrix ist. Es gibt aber immer eine invertierbare Matrix S, so dass $S^{-1}BS$ Jordan-Normalform hat, das heißt, $J := S^{-1}BS$ hat die Form

$$J = \begin{pmatrix} \boxed{J_1} & & & 0 \\ & \boxed{J_2} & & \\ & & \ddots & \\ 0 & & & \boxed{J_q} \end{pmatrix}$$

mit Jordanblöcken $J_1, \ldots, J_q$. Diese sind $(m_l \times m_l)$-Matrizen der Form

$$J_l = \begin{pmatrix} \lambda_l & 1 & & & 0 \\ & \lambda_l & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ 0 & & & & \lambda_l \end{pmatrix}$$

sind, wobei $\lambda_l \in \mathbb{C}$ und $\sum_{l=1}^{q} m_l = n$.

Wir definieren die $(m \times m)$-Matrix $N_m = (\delta_{j+1,k})_{j,k=1,\ldots,m}$, mit dem sogenannten Kronecker-Symbol

$$\delta_{ik} = \begin{cases} 1, & \text{wenn } i = k, \\ 0, & \text{wenn } i \neq k. \end{cases}$$

Dann gilt

$$J_l = \lambda_l E_{m_l} + N_{m_l}.$$

Offensichtlich gilt

$$J^k = \begin{pmatrix} \boxed{J_1^k} & & & 0 \\ & \boxed{J_2^k} & & \\ & & \ddots & \\ 0 & & & \boxed{J_q^k} \end{pmatrix}$$

und damit

$$\exp J = \begin{pmatrix} \boxed{\exp J_1} & & & 0 \\ & \boxed{\exp J_2} & & \\ & & \ddots & \\ 0 & & & \boxed{\exp J_q} \end{pmatrix}.$$

Nun gilt

$$N_m^2 = \begin{pmatrix} 0 & 0 & 1 & 0 & & 0 \\ & 0 & 0 & 1 & \ddots & \\ & & \ddots & \ddots & \ddots & 0 \\ & & & \ddots & \ddots & 1 \\ & & & & \ddots & 0 \\ 0 & & & & & 0 \end{pmatrix} = \left(\delta_{j+2,k}\right)_{j,k=1,\ldots,m}$$

und allgemein

$$N_m^p = \left(\delta_{j+p,k}\right)_{j,k=1,\ldots,m} \quad \text{für } p < m$$

mit

$$N_m^p = 0 \quad \text{für } p \geq m.$$

Allgemein gilt für Matrizen P und Q nicht $\exp(P+Q) = \exp(P)\exp(Q)$. Dies gilt aber, falls $PQ = QP$. Der Beweis sei als Übung überlassen. Für $t \in \mathbb{R}$ erhalten wir damit

$$\begin{aligned} \exp(tJ_l) &= \exp(\lambda_l t E_{m_l} + tN_{m_l}) \\ &= \exp(\lambda_l t E_{m_l}) \exp(tN_{m_l}) \\ &= e^{\lambda_l t} E_{m_l} \sum_{k=0}^{\infty} \frac{t^k}{k!} \left(N_{m_l}\right)^k \\ &= e^{\lambda_l t} \sum_{k=0}^{m_l - 1} \frac{t^k}{k!} \left(N_{m_l}\right)^k, \end{aligned}$$

also

$$\exp(tJ_l) = e^{\lambda_l t} \begin{pmatrix} 1 & t & \frac{t^2}{2} & \cdots & & & \frac{t^{m_l-1}}{(m_l-1)!} \\ & 1 & t & \frac{t^2}{2} & \cdots & & \\ & & 1 & t & \frac{t^2}{2} & & \vdots \\ & & & \ddots & \ddots & \ddots & \vdots \\ & & & & \ddots & \ddots & \frac{t^2}{2} \\ & & & & & \ddots & t \\ 0 & & & & & & 1 \end{pmatrix}.$$

Für das Anfangswertproblem $y' = Ay$, $y(x_0) = y_0$, erhalten wir also zusammenfassend: Sei S invertierbar, so dass

$$J = S^{-1}AS = \begin{pmatrix} \boxed{J_1} & & & 0 \\ & \boxed{J_2} & & \\ & & \ddots & \\ 0 & & & \boxed{J_q} \end{pmatrix}$$

Jordan-Normalform hat. Dann ist

$$\begin{aligned} y(x) &= \exp((x - x_0)A)y_0 \\ &= S\exp((x - x_0)J)S^{-1}y_0 \\ &= S \begin{pmatrix} \boxed{\exp((x-x_0))\,J_1} & & 0 \\ & \ddots & \\ 0 & & \boxed{\exp((x-x_0))\,J_q} \end{pmatrix} S^{-1}y_0 \end{aligned}$$

Lösung des Anfangswertproblems, wobei die Blöcke $\exp((x - x_0)J_l)$ wie oben gegeben sind.

Mit $y = (y_1, \ldots, y_n)^t$ gilt also

$$y_k(x) = \sum_{l=1}^{q} p_l(x)e^{\lambda_l x},$$

wobei $\lambda_1, \ldots, \lambda_q$ die Eigenwerte von A sind und die p_l Polynome sind (deren maximaler Grad um 1 kleiner ist als die Größe des entsprechenden Jordanblocks).

8 Lineare Differentialgleichungen höherer Ordnung

Wir betrachten die Differentialgleichung

$$y^{(n)} + a_{n-1}y^{(n-1)} + \ldots + a_1y' + a_0y = b,$$

wobei $n \in \mathbb{N}$ und $a_0, a_1, \ldots, a_{n-1}, b\colon I \to \mathbb{R}$ (oder $\mathbb{C}$) mit einem Intervall I.

Wie vorher schreiben wir diese Differentialgleichung als System: Mit

$$\overline{y} = \begin{pmatrix} y \\ y' \\ \vdots \\ y^{(n-1)} \end{pmatrix}$$

erhalten wir

$$\overline{y}' = A\overline{y} + \overline{b},$$

wobei

$$A = \begin{pmatrix} 0 & 1 & 0 & & & & 0 \\ & 0 & 1 & 0 & & & \\ & & 0 & 1 & 0 & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & 0 & 1 & 0 \\ 0 & & & & & 0 & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & \cdots & -a_{n-2} & -a_{n-1} \end{pmatrix} \quad \text{und} \quad \overline{b} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ b \end{pmatrix}.$$

Alles über lineare Differentialgleichungssysteme Gesagte überträgt sich unmittelbar auf die lineare Differentialgleichung n-ter Ordnung, etwa Existenz und Eindeutigkeit von Lösungen, wobei die Anfangsbedingung jetzt die Form

$$y(x_0) = y_0, \quad y'(x_0) = y_1, \quad \ldots, \quad y^{(n-1)}(x_0) = y_{n-1}$$

hat.

Ebenso erhält man die allgemeine Lösung der inhomogenen Differentialgleichung aus der allgemeinen Lösung der homogenen Differentialgleichung (d.h., dem Fall $b = 0$) sowie einer partikulären Lösung der inhomogenen Differentialgleichung.

Die Berechnung einer partikulären Lösung vereinfacht sich aber: Seien dazu $y_1, \ldots, y_n$ linear unabhängige Lösungen der homogenen linearen Differentialgleichung. Wir nennen $y_1, \ldots, y_n$ dann wieder Fundamentalsystem, denn

$$Y(x) = (\overline{y}_1, \ldots, \overline{y}_n) = \begin{pmatrix} y_1 & y_2 & \cdots & y_n \\ y_1' & y_2' & \cdots & y_n' \\ \vdots & \vdots & & \vdots \\ y_1^{(n-1)} & y_2^{(n-1)} & \cdots & y_n^{(n-1)} \end{pmatrix}$$

ist Fundamentalmatrix des zugehörigen Systems. Wie zuvor betrachten wir die Wronski-Determinante

$$W(y_1, \ldots, y_n) := W(\overline{y}_1, \ldots, \overline{y}_n) = \det \begin{pmatrix} y_1 & y_2 & \cdots & y_n \\ y_1' & y_2' & \cdots & y_n' \\ \vdots & \vdots & & \vdots \\ y_1^{(n-1)} & y_2^{(n-1)} & \cdots & y_n^{(n-1)} \end{pmatrix}.$$

Oft findet man dies als Definition der Wronski-Determinante – auch dann, wenn die Funktionen $y_1, \ldots, y_n$ nicht als Lösung einer linearen Differentialgleichung erscheinen.

Nach Satz 6.6 und Bemerkung 6.7 ist für $x_0 \in I$ und $c \in \mathbb{R}^n$ durch

$$\overline{y}(x) = Y(x)c + Y(x) \int_{x_0}^{x} Y^{-1}(t)\overline{b}(t)dt$$

eine Lösung des inhomogenen Systems gegeben.

Sei nun $u(t) = Y^{-1}(t)\overline{b}(t)$, kurz $u = Y^{-1}\overline{b}$. Dann gilt $Yu = \overline{b}$ und nach Cramerscher Regel folgt mit $u = (u_1, \ldots, u_n)^t$, dass

$$u_k = \frac{\det V_k}{\det Y} = \frac{\det V_k}{W(y_1, \ldots, y_n)},$$

wobei V_k die Matrix ist, die entsteht, wenn man in Y die k-te Spalte durch $\overline{b}$ ersetzt.

Bis hierher gilt alles auch für allgemeine Systeme. In unserem Fall gilt nun

$$\begin{aligned} \det V_k &= \det \begin{pmatrix} y_1 & \ldots & y_{k-1} & 0 & y_{k+1} & \cdots & y_n \\ y_1' & \ldots & y_{k-1}' & 0 & y_{k+1}' & \cdots & y_n' \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ y_1^{(n-1)} & \ldots & y_{k-1}^{(n-1)} & b & y_{k+1}^{(n-1)} & \ldots & y_n^{(n-1)} \end{pmatrix} \\ &= (-1)^{k+n} b\, W(y_1, \ldots, y_{k-1}, y_{k+1}, \ldots, y_n), \end{aligned}$$

also

$$u_k = (-1)^{k+n} b \frac{W(y_1, \ldots, y_{k-1}, y_{k+1}, \ldots, y_n)}{W(y_1, \ldots, y_n)}.$$

Außerdem interessiert uns nur die erste Komponente von $\overline{y} = (y, y', \ldots, y^{(n-1)})^t$. Wir erhalten folgendes Resultat.

Satz 8.1. *Gegeben sei die Differentialgleichung*

$$y^{(n)} + a_{n-1}y^{(n-1)} + \ldots + a_1 y' + a_0 y = b,$$

mit stetigen Funktionen $a_0, a_1, \ldots, a_{n-1}, b \colon I \to \mathbb{C}$. Seien $y_1, \ldots, y_n$ linear unabhängige Lösungen der homogenen Gleichung. Dann ist die allgemeine Lösung der inhomogenen Gleichung gegeben durch

$$y(x) = \sum_{k=1}^{n} c_k y_k(x) + \sum_{k=1}^{n} (-1)^{k+n} y_k(x) \int_{x_0}^{x} b(t) \frac{W_k(t)}{W(t)} dt,$$

mit $W = W(y_1, \ldots, y_n)$, $W_k = W(y_1, \ldots, y_{k-1}, y_{k+1}, \ldots, y_n)$ und $c_1, \ldots, c_k \in \mathbb{C}$.

Des Weiteren gilt wegen $\mathrm{Sp}(A) = -a_{n-1}$ jetzt

$$W(x) = W(x_0) \exp\left(-\int_{x_0}^{x} a_{n-1}(t)dt\right)$$

für $x, x_0 \in I$.

Wir betrachten nun den Fall konstanter Koeffizienten in der homogenen Differentialgleichung, also die Gleichung

$$y^{(n)} + a_{n-1}y^{(n-1)} + \ldots + a_1 y' + a_0 y = 0$$

mit $a_0, a_1, \ldots, a_{n-1} \in \mathbb{C}$. Auch hier können wir die Ergebnisse für Systeme benutzen, um ein Fundamentalsystem zu bestimmen. Dies ist allerdings recht mühsam. Andererseits zeigen die Ergebnisse für Systeme, dass die Lösungen Linearkombinationen von Funktionen der Form $x \mapsto x^k e^{\lambda x}$ sein müssen.

Wir überprüfen zunächst, wann $y(x) = e^{\lambda x}$ Lösung ist. Mit

$$\frac{d^k}{dx^k}(e^{\lambda x}) = \lambda^k e^{\lambda x}$$

erhält man

$$0 = \lambda^n e^{\lambda x} + a_{n-1}\lambda^{n-1} e^{\lambda x} + \ldots + a_0 e^{\lambda x},$$

wegen $e^{\lambda x} \neq 0$ mit $a_n := 1$ also

$$p(\lambda) := \sum_{k=0}^{n} a_k \lambda^k = 0.$$

Man nennt p das *charakteristische Polynom* der Differentialgleichung. Tatsächlich gilt auch

$$p(\lambda) = (-1)^n \det(A - \lambda E)$$

mit der Matrix A des zugehörigen Systems.

Für eine Nullstelle λ des charakteristischen Polynoms ist $x \mapsto e^{\lambda x}$ also Lösung der homogenen Differentialgleichung. Eine genauere Analyse zeigt: Ist λ eine m-fache Nullstelle, so ist $x \mapsto x^k e^{\lambda x}$ für $0 \leq k \leq m-1$ eine Lösung der Differentialgleichung. Auf diese Weise erhält man n linear unabhängige Lösungen, also ein Fundamentalsystem.

Wir fassen dies in folgendem Satz zusammen.

Satz 8.2. *Seien $n \in \mathbb{N}$, $a_0, a_1, \ldots, a_{n-1} \in \mathbb{C}$, $a_n = 1$ und*

$$p(\lambda) = \sum_{k=0}^{n} a_k \lambda^k.$$

Seien $\lambda_1, \ldots, \lambda_l$ die Nullstellen von p, mit den Vielfachkeiten $m_1, \ldots, m_l$. (Es gilt also $\sum_{k=1}^{l} m_k = n$.) Dann bilden die Funktionen

$$x \mapsto x^k e^{\lambda_j x}, \quad 1 \leq j \leq l,\ 0 \leq k \leq m_j - 1$$

ein Fundamentalsystem der Differentialgleichung

$$y^{(n)} + a_{n-1} y^{(n-1)} + \ldots + a_1 y' + a_0 y = 0.$$

Bemerkung 8.3. Auch wenn alle a_k reell sind, können einige der λ_j komplex sein. Mit λ_j ist dann aber auch $\overline{\lambda}_j$ Nullstelle von p. Sei $\lambda_j = \mu_j + i\nu_j$ mit $\mu_j, \nu_j \in \mathbb{R}$, und damit $\overline{\lambda}_j = \mu_j - i\nu_j$. Dann gilt

$$\frac{1}{2}\left(e^{\lambda_j x} + e^{\overline{\lambda}_j x}\right) = \frac{1}{2} e^{\mu_j x}\left(\cos \nu_j x + i \sin \nu_j x + \cos \nu_j x - i \sin \nu_j x\right) = e^{\mu_j x} \cos \nu_j x$$

und

$$\frac{1}{2i}\left(e^{\lambda_j x} - e^{\overline{\lambda}_j x}\right) = e^{\mu_j x} \sin \nu_j x.$$

Man kann also in obigem Fundamentalsystem die Funktionen

$$x^k e^{\lambda_j x} \quad \text{und} \quad x^k e^{\overline{\lambda}_j x}$$

durch

$$x^k e^{\mu_j x} \cos \nu_j x \quad \text{und} \quad x^k e^{\mu_j x} \sin \nu_j x$$

ersetzen und erhält so ein Fundamentalsystem mit ausschließlich reellwertigen Funktionen.

Beispiel 8.4. Wir betrachten die Differentialgleichung

$$y''(x) + \alpha y'(x) + \beta y(x) = 0,$$

wobei $\alpha, \beta \geq 0$. Physikalisch beschreibt dies eine gedämpfte Schwingung. Dabei ist $y(x)$ die Auslenkung einer an einer Feder hängenden Masse zum Zeitpunkt x. Weiter ist β eine von der Masse und der Feder abhängende Konstante und α ist eine von der Dämpfung abhängige Konstante.

Seien

$$\lambda_{1,2} = \frac{-\alpha \pm \sqrt{\alpha^2 - 4\beta}}{2}$$

die Nullstellen von

$$p(\lambda) = \lambda^2 + \alpha\lambda + \beta.$$

Fall 1. $\alpha^2 < 4\beta$. Dann gilt

$$\lambda_{1,2} = \mu \pm i\nu \quad \text{mit } \mu = -\frac{\alpha}{2}, \ \nu = \frac{\sqrt{4\beta - \alpha^2}}{2}.$$

Damit bilden

$$y_1(x) = e^{\mu x} \cos \nu x, \quad y_2(x) = e^{\mu x} \sin \nu x$$

ein Fundamentalsystem, d.h., die allgemeine Lösung hat die Form

$$y(x) = c_1 y_1(x) + c_2 y_2(x).$$

Im Falle $\alpha = 0$ gilt $\mu = 0$, $\nu = \sqrt{\beta}$, und man erhält

$$y_1(x) = \cos \nu x, \quad y_2(x) = \sin \nu x.$$

Dies ist die ungedämpfte Schwingung.

Fall 2. $\alpha^2 > 4\beta$. Dann gilt

$$\lambda_2 = -\frac{\alpha + \sqrt{\alpha^2 - 4\beta}}{2} < \lambda_1 = -\frac{\alpha - \sqrt{\alpha^2 - 4\beta}}{2} < 0$$

und wir erhalten das Fundamentalsystem

$$y_1(x) = e^{\lambda_1 x}, \quad y_2(x) = e^{\lambda_2 x}.$$

Fall 3. $\alpha^2 = 4\beta$. Dann gilt

$$\lambda_1 = \lambda_2 = \mu := -\frac{\alpha}{2}$$

und wir erhalten das Fundamentalsystem

$$y_1(x) = e^{\mu x}, \quad y_2(x) = xe^{\mu x}.$$

Wir bestimmen in allen drei Fällen die Lösung mit der Anfangsbedingung $y(0) = 1$, $y'(0) = 0$.

Fall 1. Sei

$$y(x) = c_1 e^{\mu x} \cos \nu x + c_2 e^{\mu x} \sin \nu x.$$

Dann gilt

$$y'(x) = (c_1\mu + c_2\nu)e^{\mu x} \cos \nu x + (-c_1\nu + c_2\mu) \sin \nu x.$$

Es folgt $c_1 = 1$ und $c_1\mu + c_2\nu = 0$, also $c_2 = -\mu/\nu$. Wir erhalten

$$y(x) = e^{\mu x} \cos \nu x - \frac{\mu}{\nu} e^{\mu x} \sin \nu x.$$

Der Fall $\alpha = 2$, $\beta = 5$ und damit $\mu = -1$, $\nu = 2$ ist in Abbildung 12 skizziert.

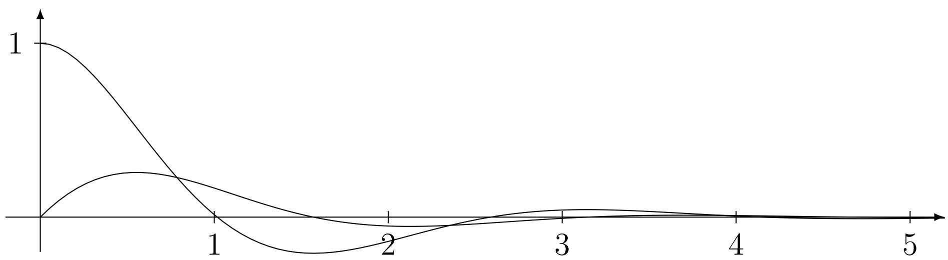

Abbildung 12: Lösungen von $y'' + 2y' + 5y = 0$, mit Anfangswerten $y(0) = 1$, $y'(0) = 0$ sowie $y(0) = 0$, $y'(0) = 1$.

Fall 2. Mit

$$y(x) = c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x}, \quad y'(x) = c_1\lambda_1 e^{\lambda_1 x} + c_2\lambda_2 e^{\lambda_2 x},$$

folgt $c_1 + c_2 = 1$ und $c_1\lambda_1 + c_2\lambda_2 = 0$, also $c_2 = -c_1\lambda_1/\lambda_2$ und damit

$$c_1 \frac{\lambda_2 - \lambda_1}{\lambda_1} = c_1 \left(1 - \frac{\lambda_1}{\lambda_2}\right) = 1.$$

Dies liefert

$$c_1 = -\frac{\lambda_2}{\lambda_1 - \lambda_2}, \quad c_2 = \frac{\lambda_1}{\lambda_1 - \lambda_2},$$

also

$$y(x) = \frac{1}{\lambda_1 - \lambda_2} \left(-\lambda_2 e^{\lambda_1 x} + \lambda_1 e^{\lambda_2 x}\right).$$

Wegen $\lambda_2 < \lambda_1 < 0$ gilt $\lambda_2 e^{\lambda_1 x} < \lambda_1 e^{\lambda_1 x} = \lambda_1 e^{\lambda_2 x} e^{(\lambda_1 - \lambda_2)x} < \lambda_1 e^{\lambda_2 x}$ für $x \geq 0$, also $y(x) > 0$ für $x \geq 0$.

Der Fall $\alpha = 2$, $\beta = \frac{3}{4}$ ist in Abbildung 13 dargestellt.

Fall 3. Mit

$$y(x) = (c_1 + c_2 x)e^{\mu x}$$

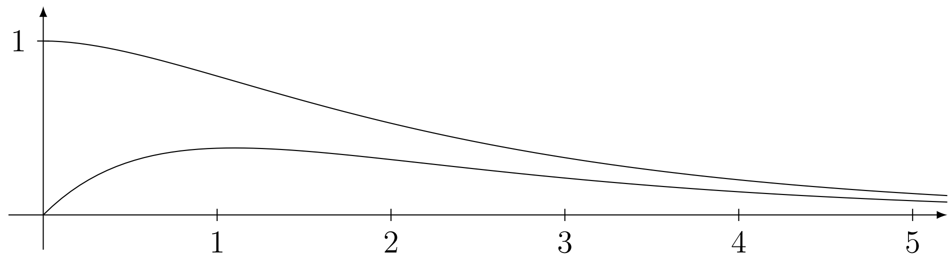

Abbildung 13: Lösungen von $y'' + 2y' + \frac{3}{4}y = 0$, mit Anfangswerten $y(0) = 1$, $y'(0) = 0$ sowie $y(0) = 0$, $y'(0) = 1$.

folgt

$$y'(x) = (c_2 + \mu c_1 + \mu c_2 x)e^{\mu x},$$

also $c_1 = 1$ und $c_2 + \mu c_1 = 0$, also $c_2 = -\mu c_1 = -\mu$. Wir erhalten

$$y(x) = (1 - \mu x)e^{\mu x} = \left(1 + \frac{\alpha}{2}x\right) e^{-\alpha x/2}.$$

Der Fall $\alpha = 2$, $\beta = 1$ ist in Abbildung 14 dargestellt.

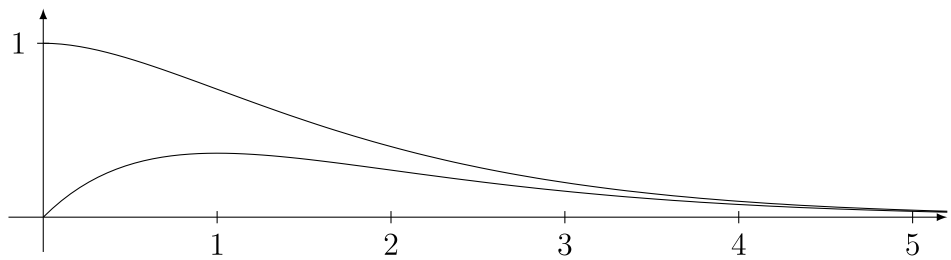

Abbildung 14: Lösungen von $y'' + 2y' + y = 0$, mit Anfangswerten $y(0) = 1$, $y'(0) = 0$ sowie $y(0) = 0$, $y'(0) = 1$.

Bemerkung 8.5. Im Fall 1 kann man natürlich auch wie im Fall 2 vorgehen und erhält mit $\lambda_{1,2} = \mu \pm i\nu$, dass

$$\begin{aligned} y(x) &= \frac{1}{\lambda_1 - \lambda_2}\left(-\lambda_2 e^{\lambda_1 x} + \lambda_1 e^{\lambda_2 x}\right) \\ &= \frac{1}{2i\nu}\left(-(\mu - i\nu)e^{\mu x}(\cos \nu x + i \sin \nu x) + (\mu + i\nu)e^{\mu x}(\cos \nu x - i \sin \nu x)\right) \\ &= \frac{1}{2i\nu}e^{\mu x}\left((-\mu + i\nu + \mu + i\nu)\cos \nu x + (-i\mu - \nu - i\mu + \nu)\sin \nu x\right) \\ &= e^{\mu x}\cos \nu x - \frac{\mu}{\nu}e^{\mu x}\sin \nu x. \end{aligned}$$

Wir betrachten jetzt die inhomogene Differentialgleichung

$$y''(x) + \alpha y'(x) + \beta y(x) = h(x).$$

Physikalisch interpretieren wir das als gedämpfte Schwingung mit zeitabhängiger äußerer Kraft.

Wir betrachten nur den Fall 2, aber mit komplexem $\lambda_{1,2}$ ist damit auch Fall 1 erledigt. Für die Wronski-Determinante erhalten wir

$$W(x) = \det \begin{pmatrix} e^{\lambda_1 x} & e^{\lambda_2 x} \\ \lambda_1 e^{\lambda_1 x} & \lambda_2 e^{\lambda_2 x} \end{pmatrix} = (\lambda_2 - \lambda_1) e^{(\lambda_1 + \lambda_2)x}.$$

Des Weiteren gilt

$$W_1(x) = W(y_2)(x) = y_2(x) = e^{\lambda_2 x}$$

und

$$W_2(x) = e^{\lambda_1 x}.$$

Mit Satz 8.1 erhalten wir

$$\begin{aligned} y(x) &= c_1 y_1(x) + c_2 y_2(x) + \sum_{k=1}^{2} (-1)^k y_k(x) \int_{x_0}^{x} h(t) \frac{W_k(t)}{W(t)} dt \\ &= c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x} - e^{\lambda_1 x} \int_{x_0}^{x} h(t) \frac{e^{\lambda_2 t}}{(\lambda_2 - \lambda_1) e^{(\lambda_1 + \lambda_2)t}} dt \\ &\quad + e^{\lambda_2 x} \int_{x_0}^{x} h(t) \frac{e^{\lambda_1 t}}{(\lambda_2 - \lambda_1) e^{(\lambda_1 + \lambda_2)t}} dt \\ &= c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x} - \frac{e^{\lambda_1 x}}{\lambda_2 - \lambda_1} \int_{x_0}^{x} h(t) e^{-\lambda_1 t} dt + \frac{e^{\lambda_2 x}}{\lambda_2 - \lambda_1} \int_{x_0}^{x} h(t) e^{-\lambda_2 t} dt. \end{aligned}$$

In vielen konkreten Fällen ist es aber einfacher, partikuläre Lösungen durch Ansätze zu gewinnen. Ist etwa h von der Form

$$h(x) = p(x) e^{cx}$$

mit einem Polynom p und $c \in \mathbb{C} \setminus \{\lambda_1, \lambda_2\}$, so existiert eine partikuläre Lösung y von der Form

$$y(t) = q(x) e^{cx},$$

wobei q ein Polynom vom gleichen Grad wie p ist.