Skript zur Vorlesung

Analysis III

Maßtheorie

Wintersemester 2021/22

Robert Denk

Universität Konstanz

Fachbereich Mathematik und Statistik

14.02.2022

Inhaltsverzeichnis

1 Maßtheoretische Grundlagen 1

a) Inhalte und Maße 1

b) Das Lebesgue-Maß 9

2 Das allgemeine Lebesgue-Integral: erste Eigenschaften 14

a) Messbare Abbildungen und das Lebesgue-Integral 14

b) Konvergenzsätze 17

3 Produktmaße und der Satz von Fubini 22

a) Produkt-σ-Algebren 22

b) Die Sätze von Tonelli und Fubini 24

4 Der Transformationssatz 29

5 Der Satz von Radon-Nikodym 35

6 Die L^p-Räume 40

a) Definition und erste Eigenschaften 40

b) Die Faltung 43

c) Vollständigkeits- und Dichtheitsaussagen 46

Literatur 50

Index 50

1. Maßtheoretische Grundlagen

1.1 Worum geht's? Der Begriff des Maßes ist uns bereits aus der Analysis II bekannt. Während damals vor allem das Lebesgue-Maß und der zugehörige Integralbegriff im Vordergrund standen, sollen jetzt allgemeinere Maße diskutiert werden. Maße werden häufig erzeugt von Inhalten, welche auf einer Algebra (statt auf einer σ-Algebra) definiert sind. Damit stellt sich die Frage nach der Fortsetzung eines Inhaltes zu einem Maß auf die erzeugte σ-Algebra. Die wichtigste Antwort dazu gibt der Fortsetzungssatz von Carathéodory, welcher das erzeugte Maß sogar (in Form des äußeren Maßes) konstruktiv angibt. Die Anwendung auf den elementargeometrischen Inhalt ergibt wiederum das Lebesgue-Maß.

a) Inhalte und Maße

1.2 Definition. Sei X eine Menge, $\mathscr{P}(X) := \{A : A \subset X\}$ die Potenzmenge von X und $\mathscr{A} \subset \mathscr{P}(X)$.

a) $\mathscr{A}$ heißt ein Ring, falls $\emptyset \in \mathscr{A}$ und für alle $A, B \in \mathscr{A}$ gilt: $A \setminus B \in \mathscr{A}$ und $A \cup B \in \mathscr{A}$.

b) $\mathscr{A}$ heißt eine Algebra über X, falls gilt:

(i) $\emptyset \in \mathscr{A}$,

(ii) Für jedes $A \in \mathscr{A}$ gilt $A^c := \{x \in X : x \notin A\} \in \mathscr{A}$.

(iii) Für $A, B \in \mathscr{A}$ gilt $A \cup B \in \mathscr{A}$.

c) Falls in b) statt (iii) gilt

(iii′) Falls $A_n \in \mathscr{A}$ disjunkt sind (d.h. $A_n \cap A_m = \emptyset$ für $n \neq m$), dann ist $\dot{\bigcup}_{n\in\mathbb{N}} A_n \in \mathscr{A}$,

so heißt $\mathscr{A}$ ein Dynkin-System.

d) Falls in c) statt (iii′) sogar gilt:

(iii″) Falls $A_n \in \mathscr{A}$, dann ist $\bigcup_{n\in\mathbb{N}} A_n \in \mathscr{A}$,

so heißt $\mathscr{A}$ eine σ-Algebra, und $(X, \mathscr{A})$ heißt ein Messraum.

1.3 Bemerkung. a) Sei $\mathscr{E} \subset \mathscr{P}(X)$ beliebig. Dann bezeichnen wir mit $\sigma(\mathscr{E})$ die kleinste σ-Algebra, welche $\mathscr{E}$ enthält (die von $\mathscr{E}$ erzeugte σ-Algebra). Diese existiert, da der Durchschnitt beliebig vieler σ-Algebren wieder eine σ-Algebra ist. Analog existieren ein kleinstes Dynkin-System $\mathscr{D}(\mathscr{E})$ und eine kleinste Algebra, die $\mathscr{E}$ enthält.

Die von $\mathscr{E}$ erzeugte Algebra kann man explizit angeben:

$$\Big\{ \bigcup_{i=1}^{n} \bigcap_{j=1}^{n} A_{ij} : A_{ij} \in \mathscr{E} \cup \mathscr{E}^c,\ n \in \mathbb{N} \Big\},$$

wobei $\mathscr{E}^c := \{A^c : A \in \mathscr{E}\}$.

Für σ-Algebren gilt dies keineswegs, auch nicht, wenn man n durch ∞ ersetzt und abzählbar oft iteriert!

b) Sei $X \neq \emptyset$ eine Menge, $(Y, \mathscr{B})$ ein Messraum und $f\colon X \to Y$ eine Funktion. Dann ist $\sigma(f) := f^{-1}(\mathscr{B}) := \{f^{-1}(B) : B \in \mathscr{B}\}$ eine σ-Algebra auf X, die von f erzeugte σ-Algebra.

Analog wird die erzeugte σ-Algebra für eine Familie von Funktionen $\mathcal{F} = \{f_i\colon X \to Y_i \mid i \in I\}$ definiert: Falls $(Y_i, \mathscr{B}_i)$ für $i \in I$ ein Messraum ist, so ist $\sigma(\mathcal{F}) := \sigma\big(\bigcup_{i\in I} f_i^{-1}(\mathscr{B}_i)\big)$ die von $\mathcal{F}$ erzeugte σ-Algebra. Dies ist die kleinste σ-Algebra auf X, für welche alle f_i messbar sind.

c) Man beachte die Zusammenhänge der verschiedenen Mengensysteme. Zum Beispiel ist ein Ring $\mathscr{A}$ genau dann eine Algebra, falls $X \in \mathscr{A}$ gilt. Jede Algebra ist auch ein Ring.

d) Ein Mengensystem $\mathscr{A} \subset \mathscr{P}(X)$ ist genau dann ein Dynkin-System, wenn gilt:

(i) $X \in \mathscr{A}$,
(ii) Für $A, B \in \mathscr{A}$ mit $A \subset B$ gilt $B \setminus A \in \mathscr{A}$.
(iii) Für $A_n \in \mathscr{A}$ $(n \in \mathbb{N})$ mit $A_1 \subset A_2 \subset \ldots$ gilt $\bigcup_{n\in\mathbb{N}} A_n \in \mathscr{A}$.

1.4 Lemma. a) *Ein Dynkin-System $\mathscr{D}$ ist genau dann eine σ-Algebra, falls gilt:*

$$\textit{Für alle } A, B \in \mathscr{D} \textit{ ist } A \cap B \in \mathscr{D}$$

(*d.h. wenn $\mathscr{D}$ $\cap$-stabil ist*).

b) (Dynkin-Lemma). *Sei $\mathscr{E} \subset \mathscr{P}(X)$ $\cap$-stabil. Dann ist $\sigma(\mathscr{E}) = \mathscr{D}(\mathscr{E})$.*

Beweis. a) Sei $\mathscr{D}$ $\cap$-stabil. Dann gilt für $A, B \in \mathscr{D}$:

$$A \setminus B = A \setminus (A \cap B) = A \cap (A \cap B)^c = \Big[A^c \,\dot\cup\, (A \cap B)\Big]^c \in \mathscr{D},$$

also

$$A \cup B = (A \setminus B) \,\dot\cup\, B \in \mathscr{D}.$$

Seien $A_n \in \mathscr{D}$ für alle $n \in \mathbb{N}$. Setze $\widetilde{A}_0 := \emptyset$ und $\widetilde{A}_n := A_1 \cup \ldots \cup A_n \in \mathscr{D}$. Dann ist

$$\bigcup_{n\in\mathbb{N}} A_n = \dot{\bigcup_{n\in\mathbb{N}_0}} \widetilde{A}_{n+1} \setminus \widetilde{A}_n \in \mathscr{D},$$

d.h. $\mathscr{D}$ ist σ-Algebra.

b) Zu zeigen ist nur, dass $\mathscr{D}(\mathscr{E})$ eine σ-Algebra ist. Zu $A \in \mathscr{D}(\mathscr{E})$ definiere

$$\mathscr{D}_A := \{B \in \mathscr{P}(X) : A \cap B \in \mathscr{D}(\mathscr{E})\}.$$

Dann ist $\mathscr{D}_A$ ein Dynkin-System. Da $\mathscr{E}$ $\cap$-stabil ist, gilt

$$\mathscr{E} \subset \mathscr{D}_A \text{ für alle } A \in \mathscr{E}$$

und damit $\mathscr{D}(\mathscr{E}) \subset \mathscr{D}_A$ für alle $A \in \mathscr{E}$, d.h.

$$A \cap B \in \mathscr{D}(\mathscr{E}) \text{ für alle } A \in \mathscr{E},\ B \in \mathscr{D}(\mathscr{E}).$$

Dies heißt $\mathscr{E} \subset \mathscr{D}_B$ für alle $B \in \mathscr{D}(\mathscr{E})$ und damit

$$\mathscr{D}(\mathscr{E}) \subset \mathscr{D}_B \text{ für alle } B \in \mathscr{D}(\mathscr{E}),$$

d.h. $\mathscr{D}(\mathscr{E})$ ist $\cap$-stabil. Mit Teil a) folgt nun die Behauptung. $\square$

1.5 Definition. a) Sei $\mathscr{A}$ ein Ring. Dann heißt $\mu : \mathscr{A} \to [0, \infty]$ ein Inhalt, falls $\mu(\emptyset) = 0$ und

$$\mu(A \cup B) = \mu(A) + \mu(B) \text{ für alle } A, B \in \mathscr{A} \text{ mit } A \cap B = \emptyset$$

(endliche Additivität) gilt. Ein Inhalt μ heißt σ-additiv, falls für alle $(A_n)_{n\in\mathbb{N}} \subset \mathscr{A}$ mit $A_n \cap A_m = \emptyset$ $(n \neq m)$ und $\dot{\bigcup}_{n\in\mathbb{N}} A_n \in \mathscr{A}$ gilt:

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

Ein σ-additiver Inhalt auf einer σ-Algebra heißt ein Maß. In diesem Fall heißt $(X, \mathscr{A}, \mu)$ ein Maßraum.

b) Ein Inhalt μ auf einem Ring $\mathscr{A}$ heißt

- σ-endlich (oder σ-finit oder normal), falls es eine Folge $(A_n)_n \subset \mathscr{A}$ gibt mit $\bigcup_{n\in\mathbb{N}} A_n = X$ und $\mu(A_n) < \infty$ für alle $n \in \mathbb{N}$,
- endlich, falls $\mu(A) < \infty$ für alle $A \in \mathscr{A}$ gilt. (Falls $\mathscr{A}$ eine Algebra oder sogar eine σ-Algebra ist, ist dies äquivalent zur Bedingung $\mu(X) < \infty$.)

c) Ein Maß μ auf einer σ-Algebra heißt Wahrscheinlichkeitsmaß, falls $\mu(X) = 1$ gilt. In diesem Fall heißt $(X, \mathscr{A}, \mu)$ ein Wahrscheinlichkeitsraum.

Wahrscheinlichkeitsräume spielen in der Stochastik eine zentrale Rolle, dort werden sie häufig mit $(\Omega, \mathscr{F}, \mathbb{P})$ bezeichnet.

1.6 Bemerkung. Wir wissen bereits aus Analysis II, dass ein Inhalt μ auf einem Ring $\mathscr{A}$ genau dann σ-additiv ist, falls er stetig von unten ist, d.h. falls

$$A_n \in \mathscr{A},\ A_1 \subset A_2 \subset \dots \ \text{mit}\ A := \bigcup_{n\in\mathbb{N}} A_n \in \mathscr{A} \implies \lim_{n\to\infty} \mu(A_n) = \mu(A).$$

Falls μ endlich ist, ist dies auch äquivalent zur Stetigkeit von oben:

$$A_n \in \mathscr{A},\ A_1 \supset A_2 \supset \dots \ \text{mit}\ \bigcap_{n\in\mathbb{N}} A_n = \emptyset \implies \lim_{n\to\infty} \mu(A_n) = 0.$$

1.7 Beispiele. a) Beispiele für Maße kennen wir schon aus der Analysis II, wie etwa das Dirac-Maß und das Zählmaß. Ein Beispiel für einen Inhalt ist gegeben durch $X = \mathbb{N}$, $\mathscr{A} := \{A \subset \mathbb{N} : A \text{ oder } A^c \text{ endlich }\}$ und

$$\mu(A) := \begin{cases} 0\,, & \text{falls } A \text{ endlich,} \\ 1\,, & \text{falls } A^c \text{ endlich.} \end{cases}$$

Dann ist $\mathscr{A}$ eine Algebra und μ ein Inhalt, aber $\mathscr{A}$ ist keine σ-Algebra und μ ist nicht σ-additiv auf $\mathscr{A}$.

b) Einer der wichtigsten Inhalte ist der elementargeometrische Inhalt, der bereits aus der Analysis II bekannt ist: Zu $a, b \in \mathbb{R}^n$ sei

$$(a, b] := \{x \in \mathbb{R}^n : a_j < x_j \le b_j \quad (j = 1, \dots, n)\}$$

das n-dimensionale Intervall. Sei $\mathbb{I}_n := \{(a, b] : a, b \in \mathbb{R}^n\}$ das System aller Intervalle und

$$\mathbb{A}_n := \Big\{ \dot{\bigcup_{i=1}^k} A_i : A_i \in \mathbb{I}_n, k \in \mathbb{N} \Big\}$$

das System aller endlichen Vereinigungen disjunkter Intervalle.

Definiere den elementargeometrischen Inhalt durch $\lambda((a, b]) := \prod_{j=1}^n (b_j - a_j)$ auf $\mathbb{I}_n$ und durch

$$\lambda\Big(\dot{\bigcup_{i=1}^k} A_i \Big) = \sum_{i=1}^k \lambda(A_i)$$

auf $\mathbb{A}_n$.

Wir werden später (in Lemma 1.19) sehen, dass $\mathbb{A}_n$ ein Ring ist und $\lambda \colon \mathbb{A}_n \to [0, \infty)$ ein σ-additiver und endlicher Inhalt ist.

1.8 Bemerkung. Sei μ ein Inhalt auf einem Ring $\mathscr{A}$. Dann gilt:

(i) μ ist monoton, d.h. für $A, B \in \mathscr{A}$ mit $A \subset B$ gilt $\mu(A) \le \mu(B)$.

(ii) μ ist subtraktiv, d.h. für $A, B \in \mathscr{A}$ mit $A \subset B$ und $\mu(A) < \infty$ gilt $\mu(B \backslash A) = \mu(B) - \mu(A)$.

(iii) μ ist subadditiv, d.h. für $A_1, \dots, A_n \in \mathscr{A}$ gilt

$$\mu\Big(\bigcup_{i=1}^{n} A_i\Big) \le \sum_{i=1}^{n} \mu(A_i)\,.$$

In vielen Fällen ist nicht ein Maß auf einer σ-Algebra gegeben, sondern ein Inhalt auf einer Algebra oder einem Ring. Daher stellt sich die Frage, ob sich dieser Inhalt eindeutig zu einem Maß fortsetzen lässt. Die folgende Konstruktion liefert die Antwort.

1.9 Definition. Ein äußeres Maß auf einer Menge X ist eine Funktion $\nu\colon \mathscr{P}(X) \to [0, \infty]$ mit folgenden Eigenschaften:

(i) $\nu(\emptyset) = 0$,
(ii) $\nu(A) \le \nu(B)$ für alle $A, B \in \mathscr{P}(X)$ mit $A \subset B$,
(iii) $\nu(\bigcup_{n\in\mathbb{N}} A_n) \le \sum_{n\in\mathbb{N}} \nu(A_n)$ für alle $(A_n)_{n\in\mathbb{N}} \subset \mathscr{P}(X)$.

Eine Menge $A \subset X$ heißt ν-messbar, falls

$$\nu(B) = \nu(A \cap B) + \nu(A^c \cap B) \quad \text{für alle } B \in \mathscr{P}(X)\,.$$

1.10 Definition und Satz. *Sei μ ein σ-additiver Inhalt auf einem Ring $\mathscr{A}$. Definiere $\mu^*\colon \mathscr{P}(X) \to [0, \infty]$ durch*

$$\mu^*(A) := \begin{cases} \inf\Big\{ \sum_{n\in\mathbb{N}} \mu(A_n) : A_n \in \mathscr{A},\ A \subset \bigcup_{n\in\mathbb{N}} A_n \Big\}, & \text{falls } \{\cdots\} \ne \emptyset\,, \\ \infty\,, & \text{sonst.} \end{cases}$$

a) μ^* ist ein äußeres Maß und heißt das zu μ gehörige äußere Maß.

b) Jede Menge $A \in \mathscr{A}$ ist μ^*-messbar, und es gilt $\mu^*|_{\mathscr{A}} = \mu$.

Beweis. a) Man rechnet die Eigenschaften eines äußeren Maßes direkt nach (Übung).

b) Sei $A \in \mathscr{A}$. Wir zeigen zunächst

$$\mu^*(B) \ge \mu^*(A \cap B) + \mu^*(A^c \cap B) \quad \text{für alle } B \in \mathscr{P}(X)\,. \tag{1-1}$$

Dazu sei o.E. $\mu^*(B) < \infty$. Da μ ein Inhalt ist, gilt für $(B_n)_{n\in\mathbb{N}} \subset \mathscr{A}$ die Gleichheit

$$\sum_{n\in\mathbb{N}} \mu(B_n) = \sum_{n\in\mathbb{N}} \mu(A \cap B_n) + \sum_{n\in\mathbb{N}} \mu(A^c \cap B_n).$$

Für jede Überdeckung $B \subset \bigcup_{n\in\mathbb{N}} B_n$ ist $A\cap B \subset \bigcup_n (A\cap B_n)$ und $A^c \cap B \subset \bigcup_n (A^c \cap B_n)$. Damit gilt

$$\mu^*(A\cap B) + \mu^*(A^c\cap B) \le \sum_{n\in\mathbb{N}} \mu(B_n).$$

Nimmt man das Infimum über alle Folgen $(B_n)_n$, welche B überdecken, erhält man (1-1).

In (1-1) gilt sogar Gleichheit, denn aus Eigenschaft 1.9 (iii) folgt für $B\in\mathscr{P}(X)$

$$\mu^*(B) \le \mu^*(A\cap B) + \mu^*(A^c\cap B).$$

Es ist noch zu zeigen, dass $\mu(A) = \mu^*(A) \quad (A\in\mathscr{A})$ gilt. Sei $A\in\mathscr{A}$ und $(A_n)_{n\in\mathbb{N}} \subset \mathscr{A}$ mit $A\subset \bigcup_n A_n$. Setze $B_1 := A_1$, $B_2 := A_2\setminus A_1, \dots, B_n := A_n \setminus (A_1\cup\dots\cup A_{n-1})$. Da μ monoton und σ-additiv ist, erhält man

$$\begin{aligned}\mu(A) &= \mu\Big(\bigcup_{n\in\mathbb{N}} (A\cap A_n)\Big) = \mu\Big(\dot{\bigcup_{n\in\mathbb{N}}} (A\cap B_n)\Big) = \sum_{n\in\mathbb{N}} \mu(A\cap B_n)\\ &\le \sum_{n\in\mathbb{N}} \mu(B_n) \le \sum_{n\in\mathbb{N}} \mu(A_n).\end{aligned}$$

Nimmt man das Infimum über alle Folgen $(A_n)_n$, die A überdecken, erhält man $\mu(A)\le\mu^*(A)$. Die andere Ungleichung ist aber klar wegen $A\subset A\cup\emptyset\cup\emptyset\dots$. □

1.11 Satz (Carathéodory)**.** *Sei μ ein σ-additiver Inhalt auf einem Ring $\mathscr{A}$. Dann ist das Mengensystem $\overline{\sigma}(\mathscr{A})$ aller μ^*-messbaren Mengen eine ($\sigma(\mathscr{A})$ enthaltende) σ-Algebra, und $\mu^*|_{\overline{\sigma}(\mathscr{A})}$ ist ein Maß mit $\mu^* = \mu$ auf $\mathscr{A}$. Insbesondere existiert zu μ eine Maßfortsetzung auf $\sigma(\mathscr{A})$.*

Beweis. (i) Nach Satz 1.10 ist $\mathscr{A}\subset\overline{\sigma}(\mathscr{A})$ und $\mu^*|_{\mathscr{A}} = \mu$.

(ii) Man zeigt zunächst, dass $\overline{\sigma}(\mathscr{A})$ eine Algebra und $\mu^*|_{\overline{\sigma}(\mathscr{A})}$ ein Inhalt ist. Dabei ist die $A,B\in\overline{\sigma}(\mathscr{A}) \Rightarrow A\cup B\in\overline{\sigma}(\mathscr{A})$ die einzige Bedingung, die nicht sofort folgt (Übung).

(iii) Wir zeigen, dass $\overline{\sigma}(\mathscr{A})$ ein Dynkin-System ist. Dazu sei $(A_n)_{n\in\mathbb{N}}\subset\overline{\sigma}(\mathscr{A})$ eine Folge paarweiser disjunkter Mengen und $A = \dot{\bigcup_{n\in\mathbb{N}}} A_n$. Aus der μ^*-Messbarkeit von $A_1 \,\dot\cup\, A_2$ folgt durch leichte Rechnung

$$\mu^*((A_1\,\dot\cup\, A_2)\cap B) = \mu^*(A_1\cap B) + \mu^*(A_2\cap B) \quad (B\in\mathscr{P}(X)).$$

Induktiv erhält man

$$\mu^*\Big(\dot{\bigcup_{i=1}^n} A_i\cap B\Big) = \sum_{i=1}^n \mu^*(A_i\cap B) \quad (B\in\mathscr{P}(X)).$$

Nach (ii) ist $\overline{\sigma}(\mathscr{A})$ eine Algebra und damit $\dot{\bigcup\limits_{i=1}^{n}} A_i \in \overline{\sigma}(\mathscr{A})$. Also gilt für alle $B \in \mathscr{P}(X)$

$$\begin{aligned}\mu^*(B) &= \mu^*\Big(\dot{\bigcup_{i=1}^{n}} A_i \cap B\Big) + \mu^*\Big(\Big[\dot{\bigcup_{i=1}^{n}} A_i\Big]^c \cap B\Big)\\ &\geq \sum_{i=1}^{n} \mu^*(A_i \cap B) + \mu^*(A^c \cap B).\end{aligned}$$

Hier wurde $A^c \subset \Big[\bigcup_{i=1}^n A_i\Big]^c$ verwendet. In der letzten Ungleichung nimmt man den Grenzwert $n \to \infty$ und erhält mit 1.9 (iii)

$$\mu^*(B) \geq \sum_{i=1}^{\infty} \mu^*(A_i \cap B) + \mu^*(A^c \cap B) \geq \mu^*(A \cap B) + \mu^*(A^c \cap B).$$

Wiederum mit 1.9 (iii) folgt $\mu^*(B) \leq \mu^*(A \cap B) + \mu^*(A^c \cap B)$ und damit

$$\mu^*(B) = \sum_{i=1}^{\infty} \mu^*(A_i \cap B) + \mu^*(A^c \cap B) = \mu^*(A \cap B) + \mu^*(A^c \cap B) \quad (B \in \mathscr{P}(X)) \tag{1-2}$$

gilt. Also gilt $A \in \overline{\sigma}(\mathscr{A})$, d.h. $\overline{\sigma}(\mathscr{A})$ ist ein Dynkin-System.

(iv) Da $\overline{\sigma}(\mathscr{A})$ $\cap$-stabil ist, zeigt Lemma 1.4 a), dass $\overline{\sigma}(\mathscr{A})$ sogar eine σ-Algebra ist.

(v) Setzt man $B = A$ in (1-2), erhält man

$$\mu^*(A) = \sum_{i=1}^{\infty} \mu^*(A_i),$$

d.h. $\mu^*|_{\overline{\sigma}(\mathscr{A})}$ ist ein Maß. □

1.12 Satz (Eindeutigkeitssatz). *Seien $\mathscr{A}$ eine σ-Algebra und $\mu, \nu \colon \mathscr{A} \to [0, \infty]$ zwei σ-endliche Maße, welche auf einem $\cap$-stabilen Erzeuger $\mathscr{E}$ von $\mathscr{A}$ übereinstimmen. Ferner enthalte $\mathscr{E}$ eine Folge von Mengen $(S_n)_{n\in\mathbb{N}} \subset \mathscr{E}$ mit $S_1 \subset S_2 \subset \dots$, $\bigcup_{n\in\mathbb{N}} S_n = X$ und $\mu(S_n) < \infty \quad (n \in \mathbb{N})$. Dann gilt $\mu = \nu$.*

Beweis. Sei $n \in \mathbb{N}$. Definiere

$$\mathscr{D}_n := \{A \in \mathscr{A} : \mu(S_n \cap A) = \nu(S_n \cap A)\}.$$

Nach Voraussetzung gilt $\mathscr{E} \subset \mathscr{D}_n$. Wir zeigen, dass $\mathscr{D}_n$ ein Dynkin-System ist, indem wir die Kriterien aus Bemerkung 1.3 c) nachrechnen.

Wegen $S_n \in \mathscr{E}$ ist $X \in \mathscr{D}_n$. Für $A, B \in \mathscr{D}_n$ mit $A \subset B$ gilt

$$\mu(S_n \cap (B \setminus A)) = \mu(S_n \cap B) - \mu(S_n \cap A) = \nu(S_n \cap B) - \nu(S_n \cap A) = \nu(S_n \cap (B \setminus A)),$$

also ist $B \setminus A \in \mathscr{D}_n$. Genauso sieht man, dass für eine aufsteigende Folge $(A_k)_{k\in\mathbb{N}} \subset \mathscr{D}_n$ gilt

$$\mu\Big(S_n \cap \bigcup_{k\in\mathbb{N}} A_k\Big) = \nu\Big(S_n \cap \bigcup_{k\in\mathbb{N}} A_k\Big),$$

indem man die Stetigkeit von unten (Bemerkung 1.6) verwendet.

Nach dem Dynkin-Lemma 1.4 folgt nun $\mathscr{D}_n = \mathscr{D}(\mathscr{E}) = \sigma(\mathscr{E}) = \mathscr{A}$ für alle $n \in \mathbb{N}$. Verwendet man nun nochmals die Stetigkeit von unten, so sieht man, dass $\mu(A) = \lim_{n\to\infty} \mu(A \cap S_n) = \nu(A)$ gilt. □

1.13 Definition. Ein Maß μ auf einer σ-Algebra $\mathscr{A}$ heißt vollständig, falls gilt: Aus $A \subset B$, $B \in \mathscr{A}$ und $\mu(B) = 0$ folgt $A \in \mathscr{A}$. Ein vollständiges Maß $\mu : \mathscr{A} \to [0, \infty]$ heißt Vervollständigung des Maßes $\mu_0 : \mathscr{A}_0 \to [0, \infty]$, falls $\mathscr{A}_0 \subset \mathscr{A}$, $\mu|_{\mathscr{A}_0} = \mu_0$ und folgende (universelle) Eigenschaft gilt: Sei $\mu' : \mathscr{A}' \to [0, \infty]$ vollständige Fortsetzung von μ_0. Dann ist $\mathscr{A} \subset \mathscr{A}'$ und $\mu'|_{\mathscr{A}} = \mu$ (d.h. μ ist minimale vollständige Fortsetzung von μ_0).

Insgesamt erhalten wir folgenden Fortsetzungssatz:

1.14 Satz. *Sei $\mathscr{A}$ Ring und $\mu\colon \mathscr{A} \to [0, \infty]$ ein σ-additiver und σ-endlicher Inhalt. Dann existiert genau eine Maßfortsetzung $\widetilde{\mu}$ von μ auf die von $\mathscr{A}$ erzeugte σ-Algebra $\sigma(\mathscr{A})$. Die Fortsetzung ist gegeben durch*

$$\widetilde{\mu} = \mu^*|_{\sigma(\mathscr{A})}.$$

Weiter ist durch $\mu^*|_{\overline{\sigma}(\mathscr{A})}$ eine vollständige Maßfortsetzung von μ gegeben.

Beweis. Nach dem Satz 1.11 von Carathéodory ist $\mu^*|_{\overline{\sigma}(\mathscr{A})}$ eine Maßfortsetzung von μ auf die σ-Algebra $\overline{\sigma}(\mathscr{A}) \supset \sigma(\mathscr{A})$. Nach dem Eindeutigkeitssatz 1.12 ist die Fortsetzung auf $\sigma(\mathscr{A})$ eindeutig.

Seien nun $N \in \overline{\sigma}(\mathscr{A})$ und $A \in \mathscr{P}(X)$ mit $A \subset N$ und $\mu^*(N) = 0$. Nach Satz 1.10 ist μ^* ein äußeres Maß und damit monoton, es gilt somit $\mu^*(A) = 0$. Für $B \in \mathscr{P}(X)$ folgt aus der Subadditivität und Monotonie von μ^*

$$\mu^*(B) \le \mu^*(B \cap A) + \mu^*(B \cap A^c) \le \mu^*(A) + \mu^*(B \cap A^c) = \mu^*(B \cap A^c) \le \mu^*(B).$$

Also gilt überall Gleichheit, d.h. $A \in \overline{\sigma}(\mathscr{A})$. Damit ist $\mu^*|_{\overline{\sigma}(\mathscr{A})}$ vollständig. □

1.15 Bemerkung. Sei μ endlicher, σ-additiver Inhalt auf einer Algebra $\mathscr{A}$.

a) Man kann zeigen, dass $\mu^*|_{\overline{\sigma}(\mathscr{A})}$ die Vervollständigung des Maßes $\mu^*|_{\sigma(\mathscr{A})}$ ist.

b) Definiere die Halbmetrik (!)

$$d_{\mu^*}(A,B) := \mu^*(A \triangle B) \quad \text{auf } \mathscr{P}(X) \times \mathscr{P}(X)\,,$$

wobei $A \triangle B := (A \backslash B) \cup (B \backslash A)$ die symmetrische Differenz von A und B ist. Dann gilt

$$\overline{\sigma}(\mathscr{A}) = \{B \in \mathscr{P}(X) : \text{ Für alle } \varepsilon > 0 \text{ existiert ein } A \in \mathscr{A} \text{ mit } d_{\mu^*}(A,B) < \varepsilon\}.$$

(d.h. $\overline{\sigma}(\mathscr{A})$ ist der Abschluss von $\mathscr{A}$ bzgl. d_{μ^*} in $\mathscr{P}(X)$).

c) Die in b) definierte symmetrische Differenz $A \triangle B$ entspricht der Addition von Mengen in folgendem Sinn: Falls $\mathscr{A}$ ein (Mengen-)Ring ist, so wird durch die Addition $A \triangle B$ und die Multiplikation $A \cap B$ eine Ringstruktur im algebraischen Sinn über dem endlichen Körper $\mathbb{F}_2 = \{0, 1\}$ definiert. Dieser Ring ist kommutativ, d.h. die Multiplikation ist kommutativ, und jedes Element ist idempotent, d.h. $A \cap A = A$. Der Ring enthält genau dann ein Einselement, falls $X \in \mathscr{A}$, d.h. falls $\mathscr{A}$ eine Mengenalgebra ist.

d) Im Satz 1.11 von Carathéodory wird das zum Inhalt μ gehörige äußere Maß μ^* betrachtet. Im Beweis werden jedoch nur die Eigenschaften aus der Definition 1.9 eines äußeren Maßes verwendet. Damit sieht man, dass folgende Aussage gilt: Sei ν ein äußeres Maß (im Sinne von 1.9) auf $\mathscr{P}(X)$, und seien $\mathscr{M} \subset \mathscr{P}(X)$ die ν-messbaren Mengen. Dann ist $(X, \mathscr{M}, \nu|_{\mathscr{M}})$ ein Maßraum.

b) Das Lebesgue-Maß

1.16 Definition. a) Sei $X \neq \emptyset$ eine Menge. Dann heißt ein Mengensystem $\tau \subset \mathscr{P}(X)$ eine Topologie auf X, falls $\emptyset, X \in \tau$ und falls endliche Durchschnitte und beliebige Vereinigungen von Mengen aus τ wieder in τ liegen. In diesem Fall heißt (X, τ) ein topologischer Raum und die Mengen in τ heißen offen.

b) Sei (X, τ) ein topologischer Raum. Dann heißt die von den offenen Mengen erzeugte σ-Algebra $\mathscr{B}(X) := \sigma(\tau)$ die Borel-σ-Algebra zu X.

1.17 Lemma. *Die Borel-σ-Algebra $\mathscr{B}(\mathbb{R}^n)$ ist die von den halboffenen Intervallen $\mathbb{I}_n = \{(a,b] : a,b \in \mathbb{R}^n\}$ erzeugte σ-Algebra.*

Beweis. (i) Wegen $(a, b + (\frac{1}{n}, \dots, \frac{1}{n})^t) \in \mathscr{B}(\mathbb{R}^n)$ und

$$(a,b] = \bigcap_{n \in \mathbb{N}} \left(a, b + \left(\frac{1}{n}, \dots, \frac{1}{n}\right)^t\right)$$

ist offensichtlich $\sigma(\mathbb{I}_n) \subset \mathscr{B}(\mathbb{R}^n)$.

(ii) Sei $U \subset \mathbb{R}^n$ offen. Zu jedem Punkt $q \in U \cap \mathbb{Q}^n$ sei $\varepsilon_q := \inf\{|q-x|_\infty : x \in \mathbb{R}^n \setminus U\}$ der Abstand von q zu $\mathbb{R}^n \setminus U$ bzgl. der $|\cdot|_\infty$-Norm. Dann ist $I_q := \{x \in \mathbb{R}^n : q_j - \varepsilon_q < x_j < q_j + \varepsilon_q\} \subset U$. Wie oben sieht man $I_q \in \sigma(\mathbb{I}_n)$. Da $\mathbb{Q}$ dicht in $\mathbb{R}$ ist, gilt

$$U = \bigcup_{q \in U \cap \mathbb{Q}^n} I_q.$$

Dies ist aber eine abzählbare Vereinigung, und somit ist $U \in \sigma(\mathbb{I}_n)$. Also ist $\mathscr{B}(\mathbb{R}^n) \subset \sigma(\mathbb{I}_n)$. □

1.18 Lemma. *a) Jedes Maß auf $\mathscr{B}(\mathbb{R}^n)$, welches auf den halboffenen Intervallen $\mathbb{I}_n$ endliche Werte annimmt, ist bereits durch die Werte auf $\mathbb{I}_n$ eindeutig festgelegt.*

b) Jedes endliche Maß $\mu\colon \mathscr{B}(\mathbb{R}) \to [0,\infty)$ (insbesondere jedes Wahrscheinlichkeitsmaß) ist bereits durch seine Verteilungsfunktion

$$F_\mu\colon \mathbb{R} \to [0,\infty),\ x \mapsto \mu\big((-\infty,x]\big) = \mu(\{z \in \mathbb{R} : z \le x\}),$$

eindeutig festgelegt.

Beweis. a) Da $\mathbb{I}_n$ ein $\cap$-stabiles Erzeugendensystem von $\mathscr{B}(\mathbb{R}^n)$ ist, welche die Folge $(-N,N]^n \nearrow \mathbb{R}^n$ enthält, folgt die Aussage sofort aus dem Eindeutigkeitssatz 1.12.

b) Bei endlichen Maßen μ ist der Wert auf $\mathbb{I}_1$ durch $\mu((a,b]) = F(b) - F(a)$ für $a, b \in \mathbb{R}$ mit $a < b$ bereits eindeutig festgelegt, damit folgt die Aussage aus a). □

Teil b) des obigen Lemmas gilt auch für die n-dimensionale Verteilungsfunktion

$$F_\mu\colon \mathbb{R} \to [0,\infty),\ x \mapsto \mu\big((-\infty,x_1] \times \cdots \times (-\infty,x_n]\big) = \mu(\{z \in \mathbb{R}^n : z \le x\}),$$

allerdings ist die Darstellung von $\mu((a,b])$ durch F_μ komplizierter.

Wir werden im Folgenden auch die Borel-σ-Algebra auf $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty,+\infty\}$ benötigen. Diese wird definiert durch $\mathscr{B}(\overline{\mathbb{R}}) := \sigma(\{(a,\infty] \subset \overline{\mathbb{R}} : a \in \mathbb{R}\})$. Das wichtigste Maß der Analysis ist das Lebesgue-Maß, das mit Hilfe des elementargeometrischen Inhalts (Beispiel 1.7) definiert wird.

1.19 Lemma. *a) Das Mengensystem $\mathbb{A}_n$ ist ein Ring.*

b) Der elementargeometrische Inhalt $\lambda\colon \mathbb{A}_n \to [0,\infty)$ ist ein endlicher σ-additiver Inhalt auf $\mathbb{A}_n$.

Beweis. a) Seien $A, B \in \mathbb{A}_n$. Nach Definition von $\mathbb{A}_n$ und $\mathbb{I}_n$ gilt $A \cap B \in \mathbb{A}_n$. Wir zeigen $A \setminus B \in \mathbb{A}_n$.

Seien dazu zunächst $A, B \in \mathbb{I}_n$, d.h. $A = (a, b]$, $B = (c, d]$ mit $a, b, c, d \in \mathbb{R}^n$. Wegen $A \setminus B = A \setminus (A \cap B)$ und $A \cap B \in \mathbb{I}_n$ sei o.E. $B \subset A$. Dann ist $A \setminus B$ die Vereinigung von $3^n - 1$ disjunkten Intervallen, bei denen jede Seite eines der Intervalle $(a_i, c_i], (c_i, d_i], (d_i, b_i]$ ist. Somit ist $A \setminus B \in \mathbb{A}_n$. Setzt man $A = B$, so erhält man $\emptyset \in \mathbb{A}_n$.

Seien nun $A = \dot{\bigcup}_{k=1}^{K} A_k, B = \dot{\bigcup}_{\ell=1}^{L} B_\ell \in \mathbb{A}_n$ mit $A_k, B_\ell \in \mathbb{I}_n$. Dann ist $A \setminus B = \dot{\bigcup}_{k=1}^{K} \bigcap_{\ell=1}^{L} (A_k \setminus B_\ell)$. Da nach dem bereits Bewiesenen $A_k \setminus B_\ell \in \mathbb{A}_n$ und da endliche Durchschnitte und disjunkte Vereinigungen von Mengen aus $\mathbb{A}_n$ wieder in $\mathbb{A}_n$ liegen, folgt $A \setminus B \in \mathbb{A}_n$. Wegen $A \cup B = (A \setminus B) \,\dot\cup\, (A \cap B) \,\dot\cup\, (B \setminus A)$ folgt $A \cup B \in \mathbb{A}_n$.

b) Durch ähnliche Zerlegung wie in a) sieht man, dass $\lambda \colon \mathbb{A}_n \to [0, \infty)$ durch die Bedingung $\lambda(\dot{\bigcup}_{k=1}^{K} A_k) = \sum_{k=1}^{K} \lambda(A_k)$ für $A_1, \dots, A_k \in \mathbb{I}_n$ bereits eindeutig festgelegt, wohldefiniert und additiv, d.h. ein Inhalt auf $\mathbb{A}_n$ ist.

Für den Beweis der σ-Additivität verwenden wir die Stetigkeit von oben (Bemerkung 1.6). Sei $(A_k)_{k \in \mathbb{N}} \subset \mathbb{A}_k$ mit $A_k \supset A_{k+1}$ und $\bigcap_{k \in \mathbb{N}} A_k = \emptyset$, und sei $\varepsilon > 0$. Zu jedem $A_k = \dot{\bigcup}_{\ell=1}^{L_k} A_{k\ell}$ mit $A_{k\ell} \in \mathbb{I}_n$ wähle $B_{k\ell} \in \mathbb{I}_n$ mit $\overline{B_{k\ell}} \subset A_{k\ell}$ so, dass für $B_k := \dot{\bigcup}_{\ell=1}^{L_k} B_{k\ell}$ die Abschätzung $\lambda(A_k \setminus B_k) \le 2^{-k}\varepsilon$ gilt. Dies ist möglich durch Vergrößerung der linken Intervallgrenzen.

Setzt man nun $C_k := B_1 \cap \dots \cap B_k$, so gilt $C_k \in \mathbb{A}_n$ und $\overline{C_k} \subset \overline{B_k} \subset A_k$ sowie $\overline{C_k} \supset \overline{C_{k+1}}$ und $\bigcap_{k \in \mathbb{N}} \overline{C_k} \subset \bigcap_{k \in \mathbb{N}} A_k = \emptyset$. Da die Mengen $\overline{C_k}$ kompakt sind, folgt $C_k = \emptyset \, (k \ge k_0)$ für ein hinreichend großes k_0. Für $k \ge k_0$ gilt somit

$$\lambda(A_k) = \lambda(A_k \setminus C_k) \le \sum_{j=1}^{k} \lambda(A_k \setminus B_j) \le \sum_{j=1}^{k} \lambda(A_j \setminus B_j) \le \sum_{j=1}^{k} 2^{-j}\varepsilon < \varepsilon.$$

Also folgt $\lambda(A_k) \to 0 \ (k \to \infty)$, und λ ist σ-additiv. □

1.20 Definition (Lebesgue-Maß)**.** Sei $\lambda \colon \mathbb{A}_n \to [0, \infty)$ der elementargeometrische Inhalt. Die nach Satz 1.14 existierende eindeutige Maßfortsetzung auf $\sigma(\mathbb{A}_n) = \mathscr{B}(\mathbb{R}^n)$ heißt das Lebesgue-Maß auf $\mathscr{B}(\mathbb{R}^n)$ und wird wieder mit λ bezeichnet, ebenso wie die Maßfortsetzung auf $\overline{\sigma}(\mathbb{A}_n)$. Die λ^*-messbaren Mengen $A \in \overline{\sigma}(\mathbb{A}_n)$ heißen die Lebesgue-Mengen des $\mathbb{R}^n$.

1.21 Bemerkung. Die Mächtigkeit (Kardinalität) von $\mathscr{B}(\mathbb{R})$ ist dieselbe wie die von $\mathbb{R}$, aber die Kardinalität der Lebesgue-messbaren Mengen ist $2^{|\mathbb{R}|}$ und damit größer. Das letzte sieht man, indem man eine Menge C mit $|C| = |\mathbb{R}|$ und $\lambda(C) = 0$ angibt (z.B. die Cantor-Menge). Dann ist jede Teilmenge von C Lebesgue-messbar. Es gibt also i.A. sehr viel mehr Mengen in der Vervollständigung $\overline{\sigma}(\mathscr{A})$ als in $\sigma(\mathscr{A})$.

1.22 Satz (Regularität des Lebesgue-Maßes)**.** *Sei $M \subset \mathbb{R}^n$ Lebesgue-messbar und beschränkt. Dann existiert eine kompakte Menge K und eine offene Menge U mit $K \subset M \subset U$ und $\lambda(U \setminus K) < \varepsilon$.*

Beweis. (i) Da M beschränkt ist, gilt $\lambda(M) < \infty$. Nach Definition des äußeren Maßes existieren Mengen $A_k \in \mathbb{A}_n$ mit $M \subset \bigcup_{k\in\mathbb{N}} A_k$ und

$$\lambda\Big(\bigcup_{k\in\mathbb{N}} A_k\Big) \le \sum_{k\in\mathbb{N}} \lambda(A_k) < \lambda(M) + \frac{\varepsilon}{4}.$$

Da die Mengen in $\mathbb{A}_n$ endliche Vereinigung von Intervallen sind, dürfen wir $A_k \in \mathbb{I}_n$ annehmen. Zu A_k wählen wir ein offenes Intervall $U_k \subset \mathbb{R}^n$ mit $A_k \subset U_k$ und $\lambda(U_k \setminus A_k) < \frac{1}{4}\varepsilon 2^{-k}$. Dann ist die Menge $U := \bigcup_{k\in\mathbb{N}} U_k$ offen mit $M \subset U$ und $\lambda(U \setminus M) < \frac{\varepsilon}{2}$.

(ii) Sei $Q = (-N, N)^n$ ein Würfel mit $\overline{M} \subset Q$. Wir wenden (i) auf die messbare und beschränkte Menge $Q \setminus M$ an und erhalten eine offene Menge $V \supset Q \setminus M$ mit $\lambda(V) < \lambda(Q \setminus M) + \frac{\varepsilon}{2}$. O.E. sei dabei $V \subset Q$ (sonst ersetze V durch $V \cap Q$).

Wir erhalten

$$\lambda(Q \setminus V) = \lambda(Q) - \lambda(V) \ge \lambda(Q) - \lambda(Q \setminus M) - \frac{\varepsilon}{2} = \lambda(M) - \frac{\varepsilon}{2}.$$

Die Menge $K := Q \setminus V \subset Q$ ist offensichtlich beschränkt.

Es gilt $K = Q \setminus V \subset M$ wegen $Q \setminus M \subset V \subset Q$. Damit ist $\overline{K} \subset \overline{M} \subset Q$. Wegen

$$\overline{K} = \overline{Q \cap V^c} \subset \overline{V^c} = V^c$$

folgt daher $\overline{K} \subset Q \cap V^c = K$, d.h. K ist abgeschlossen und damit kompakt.

Insgesamt erhalten wir $K \subset M \subset U$ und $\lambda(U \setminus K) \le \lambda(U \setminus M) + \lambda(M \setminus K) < \varepsilon$. □

1.23 Satz. *Sei $M \subset \mathbb{R}^n$ Lebesgue-messbar. Dann existieren Borelmengen $F, G \in \mathscr{B}(\mathbb{R}^n)$ mit $F \subset M \subset G$ und $\lambda(G \setminus M) = \lambda(M \setminus F) = 0$.*

Beweis. (i) Sei M beschränkt. Dann existieren nach Satz 1.22 abgeschlossene Mengen F_k und offene Mengen G_k mit $F_k \subset M \subset G_k$ und $\lambda(G_k \setminus F_k) < \frac{1}{k}$. Wählt man nun $F := \bigcup_{k\in\mathbb{N}} F_k$ und $G := \bigcap_{k\in\mathbb{N}} G_k$, so folgt $F \subset M \subset G$ und

$$\lambda(G \setminus M) \le \lambda(G_k \setminus M) < \frac{1}{k}$$

für alle $k \in \mathbb{N}$, d.h. $\lambda(G \setminus M) = 0$ und genauso $\lambda(M \setminus F) = 0$.

(ii) Falls M unbeschränkt ist, wähle eine abzählbare Folge $(A_k)_{k\in\mathbb{N}} \subset \mathscr{B}(\mathbb{R}^n)$ beschränkter disjunkter Mengen mit $\mathbb{R}^n = \bigcup_{k\in\mathbb{N}} A_k$ (etwa durch $A_k := [-(k+1), k+1]^n \setminus [-k,k]^n$). Zu $M_k := M \cap A_k$ existieren nach (i) $F_k, G_k \in \mathscr{B}(\mathbb{R}^n)$ mit $F_k \subset M_k \subset G_k$ und $\lambda(G_k \setminus F_k) = 0$. Setze nun $F := \bigcup_{k\in\mathbb{N}} F_k$ und $G := \bigcup_{k\in\mathbb{N}} G_k$. □

1.24 Definition. Sei $(X, \mathscr{A}, \mu)$ ein Maßraum. Eine Menge $N \subset X$ heißt μ-Nullmenge, falls $N \in \mathscr{A}$ mit $\mu(N) = 0$. Man sagt, eine Eigenschaft gilt μ-fast überall, falls die Menge aller $x \in X$, für welche diese Eigenschaft nicht gilt, eine μ-Nullmenge ist, z.B. $f = 0$ μ-fast überall ist definiert als $\mu(\{x \in X : f(x) \neq 0\}) = 0$. Man schreibt abkürzend μ-f.ü., im Englischen a.e.=almost everywhere. Speziell falls μ ein Wahrscheinlichkeitsmaß ist, spricht man von fast sicher geltenden Eigenschaften, abgekürzt f.s., im Englischen a.s.=almost sure.

1.25 Bemerkung. a) Falls das Maß vollständig ist, ist die Teilmenge einer Nullmenge wieder eine Nullmenge. Nach Satz 1.14 ist jede Teilmenge einer Lebesgue-messbaren Nullmenge wieder eine Nullmenge. Aber nicht jede Teilmenge einer Borelmenge $N \in \mathscr{B}(\mathbb{R}^n)$ mit $\lambda(N) = 0$ ist Borel-messbar!

b) Abzählbare Vereinigungen von Nullmengen sind Nullmengen.

c) Für das Lebesgue-Maß wissen wir bereits, dass einzelne Punkte, abzählbare Punktmengen, lineare Unterräume und Graphen von Funktionen Nullmengen sind. Die Cantormenge ist eine überabzählbare Lebesgue-Nullmenge.

2. Das allgemeine Lebesgue-Integral: erste Eigenschaften

2.1 Worum geht's? Wir kennen bereits den Begriff des allgemeinen Lebesgue-Integrals, es bleiben aber noch einige wichtige Beweise zu führen. In diesem Abschnitt werden die Begriffe messbare Funktion, Integral kurz wiederholt, die klassischen Konvergenzsätze bewiesen sowie Maße mit Dichten betrachtet.

a) Messbare Abbildungen und das Lebesgue-Integral

Falls $(X, \mathscr{A})$ und $(S, \mathscr{S})$ zwei Messräume sind, so heißt eine Abbildung $f\colon X \to S$ messbar (oder genauer $\mathscr{A}$-$\mathscr{S}$-messbar), falls $f^{-1}(\mathscr{S}) \subset \mathscr{A}$ gilt, d.h. mit der Bezeichnung aus Bemerkung 1.3 b), falls $\sigma(f) \subset \mathscr{A}$ gilt. In vielen Fällen wird bei uns $(S, \mathscr{S}) = (\mathbb{R}, \mathscr{B}(R))$ oder $(S, \mathscr{S}) = (\overline{\mathbb{R}}, \mathscr{B}(\overline{\mathbb{R}}))$ sein. Falls X ein topologischer Raum ist, wählt man kanonisch die Borel-σ-Algebra $\mathscr{B}(X)$ auf X.

An dieser Stelle sei auch noch einmal auf die Sprechweise der Stochastik hingewiesen: Falls $(\Omega, \mathscr{A}, P)$ ein Wahrscheinlichkeitsraum ist, so heißt eine messbare Abbildung $X\colon \Omega \to \mathbb{R}$ eine Zufallsvariable.

Folgende Aussage ist bereits aus Analysis II bekannt: Seien $(X, \mathscr{A})$ und $(S, \mathscr{S})$ Messräume und $\mathscr{S} = \sigma(\mathscr{E})$. Dann ist $f\colon X \to S$ genau dann $\mathscr{A}$-$\mathscr{S}$-messbar, wenn $f^{-1}(\mathscr{E}) \subset \mathscr{A}$. Speziell für $S = \overline{\mathbb{R}}$ reicht es zu zeigen, dass $\{x \in X : f(x) \le a\} \in \mathscr{A}$ für alle $a \in \mathbb{R}$.

2.2 Satz. *Sei $(X, \mathscr{A})$ ein Messraum.*

a) Falls X topologischer Raum ist und $\mathscr{A} = \mathscr{B}(X)$, so sind alle stetigen Funktionen $f\colon X \to \mathbb{R}$ messbar.

b) Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge $\mathscr{A}$-messbarer Funktionen $f\colon X \to \overline{\mathbb{R}}$. Dann sind auch die Funktionen $\inf_{n\in\mathbb{N}} f_n$, $\sup_{n\in\mathbb{N}} f_n$, $\liminf_{n\to\infty} f_n$ und $\limsup_{n\to\infty} f_n$ $\mathscr{A}$-messbar.

c) Der Grenzwert einer punktweise konvergenten Folge $(f_n)_{n\in\mathbb{N}}$ $\mathscr{A}$-messbarer Funktionen ist $\mathscr{A}$-messbar.

d) Seien $f, g\colon X \to \mathbb{R}$ $\mathscr{A}$-messbar und $F\colon \mathbb{R}^2 \to \mathbb{R}$ stetig, so ist auch $h\colon X \to \mathbb{R}$, $h(x) := F(f(x), g(x))$ $\mathscr{A}$-messbar. Insbesondere sind mit f und g auch $\max\{f, g\}$, $\min\{f, g\}$, $f \pm g$ und $f \cdot g$ $\mathscr{A}$-messbar, ebenso $|f|^r$ für $r > 0$ und f^r für $r \in \mathbb{N}$.

Beweis. Dies wird wörtlich wie im Fall $X = \mathbb{R}^n$ bewiesen, der bereits in Analysis II behandelt wurde. □

Wir erinnern an die Definition des allgemeinen Lebesgue-Integrals: Sei $(X, \mathscr{A}, \mu)$

ein Maßraum. Eine Stufenfunktion ist eine Funktion der Form $f = \sum_{i=1}^k c_i \chi_{A_i}$ mit $c_i \in \mathbb{R}$ und $A_i \in \mathscr{A}$, wobei χ_A wieder die charakteristische Funktion der Menge A bezeichne.

Für eine Stufenfunktion $f \geq 0$ ist das Integral durch $\int f d\mu := \sum_{i=1}^k c_i \mu(A_i) \in [0, \infty]$ definiert. Für messbare Funktionen $f \geq 0$ setzt man

$$\int f d\mu := \sup \Big\{ \int s d\mu : s \text{ Stufenfunktion mit } 0 \leq s \leq f \Big\} \in [0, \infty].$$

Sei schließlich $f \colon X \to \overline{\mathbb{R}}$ eine messbare Funktion. Man zerlegt $f = f_+ - f_-$ mit $f_+ := \max\{f, 0\}$ und $f_- := -\min\{0, f\}$ und definiert

$$\int f d\mu := \int f_+ d\mu - \int f_- d\mu,$$

falls beide Integrale auf der rechten Seite endlich sind. In diesem Fall heißt f μ-integrierbar. Die Menge aller μ-integrierbaren Funktionen $f \colon X \to \overline{\mathbb{R}}$ wird mit $\mathscr{L}^1(\mu) = \mathscr{L}^1(X, \mu)$ oder (insbesondere falls μ das Lebesgue-Maß ist) mit $\mathscr{L}^1(X)$ bezeichnet. Für $A \in \mathscr{A}$ setzt man $\int_A f d\mu := \int \chi_A f d\mu$.

Für festes $Y \in \mathscr{A}$ sei $\mathscr{A} \cap Y := \{A \cap Y : A \in \mathscr{A}\}$ die Spur-σ-Algebra über der Menge Y und $\mu|_Y := \mu|_{\mathscr{A} \cap Y}$ das Spurmaß. Dann schreibt man $\mathscr{L}^1(Y) := \mathscr{L}^1(Y, \mu|_Y)$. Insbesondere ist damit für $Y \in \mathscr{B}(\mathbb{R}^n)$ der Raum $\mathscr{L}^1(Y)$ aller über Y Lebesgue-integrierbaren Funktionen definiert.

Elementare Eigenschaften des Lebesgue-Integrals sind uns bereits bekannt. So ist z.B. die Abbildung $\mathscr{L}^1(\mu) \to \mathbb{R}$, $f \mapsto \int f d\mu$ monoton, d.h. aus $f \leq g$ folgt $\int f d\mu \leq \int g d\mu$, und homogen, d.h. $\int \alpha f d\mu = \alpha f d\mu$ für $\alpha \in \mathbb{R}$.

2.3 Bemerkung. Wir kennen bereits einige Aussagen über Integrale, in welchen Nullmengen auftauchen. Sei $(X, \mathscr{A}, \mu)$ ein Maßraum, und seien $f, g \colon X \to \overline{\mathbb{R}}$ messbar. Wir schreiben wieder $f = g$ μ-fast überall, falls $\mu(\{x \in X \colon f(x) \neq g(x)\}) = 0$.

Dann gilt:

(1) Ist $N \in \mathscr{A}$ eine μ-Nullmenge, so ist $\int_N f d\mu = 0$.

(2) Ist $f \in \mathscr{L}^1(\mu)$ und sind $A, N \in \mathscr{A}$ mit $\mu(N) = 0$, so ist $\int_{A \cup N} f d\mu = \int_A f d\mu$.

(3) Ist $f \in \mathscr{L}^1(\mu)$ und gilt $f = g$ μ-fast überall, so ist auch $g \in \mathscr{L}^1(\mu)$ und $\int f d\mu = \int g d\mu$.

(4) Ist $f \geq 0$ mit $\int f d\mu = 0$, so ist $f = 0$ μ-fast überall.

(5) Ist $f \in \mathscr{L}^1(\mu)$, so ist $\mu(\{x \in X : f(x) \in \{+\infty, -\infty\}\}) = 0$.

Dabei wurde lediglich Aussage (5) noch nicht in Analysis II bewiesen. Für den Beweis von (5) sei etwa $A \in \mathscr{A}$ mit $\mu(A) > 0$ und $f(x) = \infty$ $(x \in A)$. Zu $r > 0$ betrachte die Stufenfunktion $s_r := r\chi_A$. Dann gilt $0 \leq s_r \leq f_+$ und somit $\int f_+ d\mu \geq \int s_r d\mu = r\mu(A) \to \infty$ $(r \to \infty)$. Also ist f nicht integrierbar.

Im Folgenden sei $(X, \mathscr{A}, \mu)$ ein Maßraum.

2.4 Bemerkung. Sei $f\colon X \to \overline{\mathbb{R}}$ messbar. Dann gilt $f \in \mathscr{L}^1(\mu)$ genau dann, wenn $|f| \in \mathscr{L}^1(\mu)$. Falls $g \in \mathscr{L}^1(\mu)$ mit $|f| \le g$ μ-fast überall, so folgt $f \in \mathscr{L}^1(\mu)$ und $\int |f| d\mu \le \int g d\mu$ (Majorantenkriterium).

2.5 Satz (Maße durch Dichten)**.** *Sei $f\colon X \to \overline{\mathbb{R}}$ messbar mit $f \ge 0$. Definiere $\varphi_f\colon \mathscr{A} \to [0,\infty]$ durch*

$$\varphi_f(A) := \int_A f d\mu \quad (A \in \mathscr{A}).$$

Dann ist φ_f ein Maß.

Beweis. (i) Für charakteristische Funktionen $f = \chi_B$ mit $B \in \mathscr{A}$ gilt $\varphi_f(A) = \int_A \chi_B d\mu = \mu(A \cap B)$. Damit folgt die σ-Additivität aus der von μ.

(ii) Falls f eine Stufenfunktion ist, folgt die σ-Additivität aus (i) und der Linearität des Integrals.

(iii) Sei nun $f\colon X \to [0,\infty]$ messbar und s eine Stufenfunktion mit $0 \le s \le f$. Für $A = \dot{\bigcup}_{j=1}^{\infty} A_j$ mit $A_j \in \mathscr{A}$ gilt

$$\varphi_s(A) = \sum_{j=1}^{\infty} \varphi_s(A_j) = \sum_{j=1}^{\infty} \int_{A_j} s d\mu \le \sum_{j=1}^{\infty} \int_{A_j} f d\mu = \sum_{j=1}^{\infty} \varphi_f(A_j).$$

Nimmt man das Supremum über alle Stufenfunktionen, so erhält man

$$\varphi_f(A) \le \sum_{j=1}^{\infty} \varphi_f(A_j). \tag{2-1}$$

Falls ein $j \in \mathbb{N}$ existiert mit $\varphi_f(A_j) = \infty$, so folgt wegen $A \supset A_j$ auch $\varphi_f(A) = \infty$, und wir erhalten Gleichheit in (2-1). Sei also $\varphi_f(A_j) < \infty$ für alle $j \in \mathbb{N}$. Sei $\varepsilon > 0$. Da A_1 und A_2 disjunkt sind, existiert nach Definition des Integrals als Supremum eine Stufenfunktion s mit $0 \le s \le f$ und

$$\int_{A_i} s d\mu \ge \int_{A_i} f d\mu - \varepsilon \quad (i = 1, 2).$$

Damit

$$\begin{aligned}\varphi_f(A_1 \dot\cup A_2) &\ge \int_{A_1 \dot\cup A_2} s d\mu = \int_{A_1} s d\mu + \int_{A_2} s d\mu \\ &\ge \int_{A_1} f d\mu + \int_{A_2} f d\mu - 2\varepsilon = \varphi_f(A_1) + \varphi_f(A_2) - 2\varepsilon.\end{aligned}$$

Da ε beliebig war, folgt $\varphi_f(A_1 \dot\cup A_2) \ge \varphi_f(A_1) + \varphi_f(A_2)$. Iterativ erhält man

$$\varphi_f(A_1 \dot\cup \ldots \dot\cup A_n) \ge \sum_{i=1}^n \varphi_f(A_i).$$

Wegen $A \supset A_1 \dot\cup \ldots \dot\cup A_n$ folgt

$$\varphi_f(A) \ge \sum_{i=1}^n \varphi_f(A_i)$$

für alle $n \in \mathbb{N}$. Nimmt man nun $n \to \infty$, erhält man Gleichheit in (2-1). □

2.6 Korollar. *Sei $f \in \mathscr{L}^1(X, \mu)$ und seien $(A_n)_{n\in\mathbb{N}} \subset \mathscr{A}$ disjunkt mit $A := \dot\bigcup_{n\in\mathbb{N}} A_n$. Dann gilt*

$$\int_A f d\mu = \sum_{n=1}^\infty \int_{A_n} f d\mu.$$

Beweis. Seien $\varphi_\pm$ die zu $f_\pm$ gehörigen Maße laut Satz 2.5. Da $\int_X f_\pm d\mu < \infty$, sind beide Maße endlich. Es gilt

$$\begin{aligned}\varphi_f(A) &= \int_A f d\mu = \int_A f_+ d\mu - \int_A f_- d\mu = \varphi_+(A) - \varphi_-(A) \\ &= \sum_{n=1}^\infty \varphi_+(A_n) - \sum_{n=1}^\infty \varphi_-(A_n) = \sum_{n=1}^\infty (\varphi_+(A_n) - \varphi_-(A_n)) \\ &= \sum_{n=1}^\infty \int_{A_n} f d\mu.\end{aligned}$$

□

b) Konvergenzsätze

Im Folgenden sei $(X, \mathscr{A}, \mu)$ ein Maßraum.

2.7 Satz (von Lebesgue über monotone Konvergenz)**.** *Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge messbarer Funktionen mit $0 \le f_1 \le f_2 \le \ldots$, und sei $f \colon X \to [0, \infty]$ definiert durch $f(x) := \lim_{n\to\infty} f_n(x) \quad (x \in X)$. Dann gilt*

$$\lim_{n\to\infty} \int f_n d\mu = \int f d\mu = \int \lim_{n\to\infty} f_n d\mu.$$

Beweis. Nach Satz 2.2 b) ist $f = \sup f_n$ messbar, also existiert $\int f d\mu \in [0, \infty]$. Sei $\alpha := \lim_{n\to\infty} \int f_n d\mu \in [0, \infty]$. Wegen $\int f_n d\mu \le \int f d\mu$ folgt $\alpha \le \int f d\mu$.

Sei $0 < c < 1$ und s Stufenfunktion mit $0 \le s \le f$. Definiere $A_n := \{x \in X : f_n(x) \ge cs(x)\}$. Da $(f_n)_n$ monoton ist, folgt $A_1 \subset A_2 \subset \dots$. Wegen $cs \le f_n$ gilt

$$\int f_n d\mu \ge \int_{A_n} f_n d\mu \ge c \int_{A_n} s d\mu.$$

Wegen $\lim_n f_n = f$ und $c < 1$ gilt $\bigcup_{n\in\mathbb{N}} A_n = X$, und aus der σ-Additivität des Maßes $A \mapsto \int_A s d\mu$ (siehe Satz 2.5) folgt

$$\alpha = \lim_{n\to\infty} \int f_n d\mu \ge c \int s d\mu.$$

Nimmt man nun das Supremum über alle Stufenfunktionen s mit $0 \le s \le f$, so erhält man

$$\lim_{n\to\infty} \int f_n d\mu \ge c \int f d\mu.$$

Da $c < 1$ beliebig war, folgt daraus die Abschätzung $\alpha \ge \int f d\mu$ und damit die Behauptung. □

2.8 Satz (Addititivität des Integrals). *Seien $f_1, f_2 \in \mathscr{L}^1(\mu)$. Dann ist $f_1 + f_2 \in \mathscr{L}^1(\mu)$ und*

$$\int (f_1 + f_2) d\mu = \int f_1 d\mu + \int f_2 d\mu.$$

Beweis. (i) Seien zunächst $f_1, f_2 \ge 0$. Da f_i messbar ist, existieren Stufenfunktionen $(s_n^{(i)})_{n\in\mathbb{N}}$ für $i = 1, 2$ mit $0 \le s_n^{(i)}$ und $s_n^{(i)}(x) \nearrow f_i(x) \quad (x \in X)$. Für $s_n := s_n^{(1)} + s_n^{(2)}$ gilt $s_n(x) \nearrow f(x) := f_1(x) + f_2(x) \quad (x \in X)$. Da das Integral über Stufenfunktionen linear ist, folgt mit dem Satz über monotone Konvergenz:

$$\int f d\mu = \lim_{n\to\infty} \int s_n d\mu = \lim_{n\to\infty} \Big(\int s^{(1)} d\mu + \int s^{(2)} d\mu \Big) = \int f_1 d\mu + \int f_2 d\mu.$$

(ii) Sei nun $f_1 \ge 0$ und $f_2 \le 0$. Setze $A := \{f_1 + f_2 \ge 0\}$ und $B := \{f_1 + f_2 < 0\}$. Nach Fall (i) gilt auf A

$$\int_A (f_1 + f_2) d\mu + \int_A (-f_2) d\mu = \int_A [(f_1 + f_2) + (-f_2)] d\mu = \int_A f_1 d\mu.$$

Unter Verwendung der Homogenität folgt $\int_A f d\mu = \int_A f_1 d\mu + \int_A f_2 d\mu$. Auf der Menge B folgt die Behauptung analog mit der Darstellung $-f_2 = f_1 + [-(f_1 + f_2)]$.

(iii) Im allgemeinen Fall zerlegen wir $X = \dot{\bigcup}_{i=1}^{4} A_i$ mit

$$\begin{aligned} A_1 &:= \{x \in X : f_1(x) \geq 0,\ f_2(x) \geq 0\}, \\ A_2 &:= \{x \in X : f_1(x) \geq 0,\ f_2(x) < 0\}, \\ A_3 &:= \{x \in X : f_1(x) < 0,\ f_2(x) \geq 0\}, \\ A_4 &:= \{x \in X : f_1(x) < 0,\ f_2(x) < 0\}. \end{aligned}$$

Auf jeder dieser Mengen gilt $\int_{A_i} f d\mu = \int_{A_i} f_1 d\mu + \int_{A_i} f_2 d\mu$ nach Fall (i) oder (ii). Summation über i liefert die Behauptung. □

2.9 Satz. *Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge nichtnegativer messbarer Funktionen und $f\colon X \to [0,\infty]$ definiert durch $f := \sum_{n=1}^{\infty} f_n$. Dann gilt*

$$\int f d\mu = \sum_{n=1}^{\infty} \int f_n d\mu.$$

Falls $f_n \in \mathscr{L}^1(\mu)$ $\quad (n \in \mathbb{N})$ und die Summe auf der rechten Seite konvergiert, so gilt $f \in \mathscr{L}^1(\mu)$, und die Reihe $\sum_{n=1}^{\infty} f_n(x)$ konvergiert für μ-fast alle $x \in X$.

Beweis. Wende monotone Konvergenz auf die Partialsummen an. Die letzte Aussage folgt aus Bemerkung 2.3 (5). □

2.10 Satz (Lemma von Fatou). *Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge nichtnegativer messbarer Funktionen und $f\colon X \to [0,\infty]$ definiert durch $f := \liminf_{n\to\infty} f_n$. Dann gilt*

$$\int f d\mu \leq \liminf_{n\to\infty} \int f_n d\mu.$$

Beweis. Definiere $g_n\colon X \to [0,\infty]$ durch

$$g_n(x) := \inf\{f_k(x) : k \geq n\} \quad (x \in X,\ n \in \mathbb{N}).$$

Die Funktionen g_n sind messbar nach Satz 2.2, die Folge $(g_n)_n$ ist monoton wachsend mit $g_n(x) \nearrow f(x) \quad (x \in X)$. Nach dem Satz von der monotonen Konvergenz folgt

$$\lim_{n\to\infty} \int g_n d\mu = \int f d\mu.$$

Wegen $g_n \leq f_n$ gilt andererseits

$$\lim_{n\to\infty} \int g_n d\mu = \liminf_{n\to\infty} \int g_n d\mu \leq \liminf_{n\to\infty} \int f_n d\mu,$$

was die Behauptung zeigt. □

Der folgende Satz ist einer der wichtigsten Sätze der Integrationstheorie und heißt auch Satz von Lebesgue über dominierte Konvergenz.

2.11 Satz (von Lebesgue über majorisierte Konvergenz). *Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge messbarer Funktionen $f_n\colon X \to \overline{\mathbb{R}}$, und der Grenzwert $f(x) := \lim_{n\to\infty} f_n(x) \in \overline{\mathbb{R}}$ existiere μ-fast überall. Weiter existiere eine Funktion $g \in \mathscr{L}^1(\mu)$ mit $|f_n(x)| \le g(x)$ für μ-fast alle $x \in X$ und alle $n \in \mathbb{N}$. Dann ist $f \in \mathscr{L}^1(\mu)$ und*

$$\lim_{n\to\infty} \int f_n d\mu = \int f d\mu = \int \lim_{n\to\infty} f_n d\mu.$$

Beweis. Durch Änderung auf einer Nullmenge können wir annehmen, dass $\lim_n f_n$ überall existiert und $|f_n| \le g$ überall gilt. Als punktweiser Grenzwert messbarer Funktionen ist f messbar, und nach dem Majorantenkriterium (Bemerkung 2.4) ist $f \in \mathscr{L}^1(\mu)$.

Wegen $f_n + g \ge 0$ ist das Lemma von Fatou anwendbar, und es folgt

$$\int (f+g)d\mu = \int (\liminf_{n\to\infty} f_n + g)d\mu \le \liminf_{n\to\infty} \int (f_n + g)d\mu.$$

Wegen Additivität (Satz 2.8) folgt $\int f d\mu \le \liminf_n \int f_n d\mu$.

Da auch $g - f_n \ge 0$, folgt wieder mit Fatou

$$\int (g-f)d\mu \le \liminf_{n\to\infty} \int (g - f_n)d\mu$$

und damit $-\int f d\mu \le \liminf_n \int (-f_n)d\mu$, d.h.

$$\int f d\mu \ge \limsup_{n\to\infty} \int f_n d\mu.$$

Insgesamt haben wir also

$$\int f d\mu \le \liminf_{n\to\infty} \int f_n d\mu \le \limsup_{n\to\infty} \int f_n d\mu \le \int f d\mu,$$

was überall Gleichheit und damit die Konvergenz von $\int f_n d\mu$ gegen $\int f d\mu$ impliziert. □

2.12 Satz (Parameterabhängigkeit von Integralen). *Sei $U \subset \mathbb{R}^n$ offen, $a \in U$ und $f\colon X \times U \to \mathbb{R}$ eine Funktion. Für alle $y \in U$ sei $f(\cdot, y)\colon X \to \mathbb{R}$ μ-integrierbar. Definiere*

$$g\colon U \to \mathbb{R}, \quad g(y) := \int f(\cdot, y)d\mu = \int f(x,y)d\mu(x).$$

a) Für alle $x \in X$ sei $f(x, \cdot)\colon U \to \mathbb{R}$ stetig an der Stelle a, und es existiere eine μ-integrierbare Funktion $h\colon X \to [0, \infty]$ mit

$$|f(x, y)| \leq h(x) \quad ((x, y) \in X \times U).$$

Dann ist auch g stetig an der Stelle a.

b) Für alle $x \in X$ sei $f(x, \cdot)\colon U \to \mathbb{R}$ stetig differenzierbar in U, und es existiere eine μ-integrierbare Funktion $h\colon X \to [0, \infty]$ mit

$$\Big|\frac{\partial f}{\partial y_i}(x, y)\Big| \leq h(x) \quad ((x, y) \in X \times U,\ i = 1, \dots, n).$$

Dann ist auch g stetig differenzierbar in U, und es gilt

$$\frac{\partial g}{\partial y_i}(a) = \int \frac{\partial f}{\partial y_i}(x, y) d\mu(x) \quad (i = 1, \dots, n).$$

Beweis. a) Für jede Folge $(y_k)_{k\in\mathbb{N}} \subset U$ mit $y_k \to a$ gilt $f(\cdot, y_k) \to f(\cdot, a)$ punktweise (bzgl. $x \in X$). Nach Voraussetzung ist h eine integrierbare Majorante von $(f(\cdot, y_k))_{k\in\mathbb{N}}$, und die Aussage folgt aus dem Satz über majorisierte Konvergenz 2.11.

b) Sei $1 \leq i \leq n$ und $(t_k)_{k\in\mathbb{N}} \subset \mathbb{R}$ eine Folge mit $t_k \to 0$, $t_k \neq 0$. Dann konvergiert

$$\frac{f(\cdot, a + t_k e_i) - f(\cdot, a)}{t_k} \to \frac{\partial f}{\partial y_i}(\cdot, a)$$

punktweise bzgl. $x \in X$. Nach dem Mittelwertsatz der Differentialrechnung ist die Funktionenfolge auf der linken Seite majorisiert durch h. Damit folgt die Differenzierbarkeit von g nach dem Satz von der majorisierten Konvergenz und die Stetigkeit von g' nach Teil a). $\square$

3. Produktmaße und der Satz von Fubini

3.1 Worum geht's? Der Satz von Fubini ist bereits aus der Analysis II bekannt, dort wurde er aber nur unter relativ starken Voraussetzungen bewiesen. Der Beweis für den allgemeinen Fall soll nun nachgeholt werden.

Zunächst werden Produkte von Messräumen und von Maßen betrachtet, welche selber von Interesse sind. Insbesondere sind die Borelmengen des $\mathbb{R}^n$ das Produkt der eindimensionalen Borelmengen. Der hier gewählte Zugang lässt sich auf unendliche Produkte übertragen, bei welchen aber schnell topologische Fragen auftauchen, die wir noch nicht behandeln können.

a) Produkt-σ-Algebren

Seien $(X_i, \mathscr{A}_i)$, $1 \leq i \leq n$, endlich viele Messräume. Dazu betrachte das kartesische Produkt

$$X := X_1 \times \ldots \times X_n = \prod_{i=1}^n X_i.$$

3.2 Definition. Die Menge

$$\mathscr{Z} := \{A_1 \times \ldots \times A_n : A_i \in \mathscr{A}_i, 1 \leq i \leq n\} \subset \mathscr{P}(X)$$

heißt *Gesamtheit der Zylindermengen* (bzgl. der $\mathscr{A}_i$).

$$\mathscr{A}_1 \otimes \ldots \otimes \mathscr{A}_n := \bigotimes_{i=1}^n \mathscr{A}_i := \sigma(\mathscr{Z})$$

heißt die *Produkt-σ-Algebra der $\mathscr{A}_i$*.

3.3 Bemerkung. a) $\mathcal{Z}$ ist $\cap$-stabil. Denn $(A_1 \times \ldots \times A_n) \cap (B_1 \times \ldots \times B_n) = (A_1 \cap B_1) \times \ldots \times (A_n \cap B_n)$.

b) $\otimes$ ist assoziativ: $\mathscr{A}_1 \otimes \mathscr{A}_2 \otimes \mathscr{A}_3 = (\mathscr{A}_1 \otimes \mathscr{A}_2) \otimes \mathscr{A}_3 = \mathscr{A}_1 \otimes (\mathscr{A}_2 \otimes \mathscr{A}_3)$. Zu zeigen ist wegen Symmetrie nur die erste Gleichheit und diese bedeutet definitionsgemäß $\sigma(\{A_1 \times A_2 \times A_3\}) = \sigma(\sigma(\{A_1 \times A_2\}) \times A_3)$. Dabei ist "$\subset$" klar. Zum Beweis der anderen Inklusion betrachte ein festes $A_3 \in \mathscr{A}_3$. Für $B \in \mathscr{A}_1 \otimes \mathscr{A}_2$ gilt dann

$$B \times A_3 \subset \sigma(\{A_1 \times A_2\}) \times A_3 = \sigma(\{A_1 \times A_2 \times A_3 : A_1 \in \mathscr{A}_1, A_2 \in \mathscr{A}_2\})$$

Damit liegt aber auch jedes $B \times A_3$ mit $B \in \mathscr{A}_1 \otimes \mathscr{A}_2$ und $A_3 \in \mathscr{A}_3$ in der von allen $A_1 \times A_2 \times A_3$, $A_i \in \mathscr{A}_i$, erzeugten σ-Algebra. Diese ist nach Definition $\mathscr{A}_1 \otimes \mathscr{A}_2 \otimes \mathscr{A}_3$, was „$\supset$" zeigt.

c) Sei $\operatorname{pr}_i : X_1 \times \ldots \times X_n \to (X_i, \mathscr{A}_i)$ die i-te Koordinatenprojektion. Dann gilt

$$\bigotimes_{i=1}^n \mathscr{A}_i = \sigma(\operatorname{pr}_i : 1 \leq i \leq n) := \sigma\Big(\bigcup_{i=1}^n \operatorname{pr}_i^{-1}(\mathscr{A}_i)\Big).$$

Die Produkt-σ-Algebra ist somit die kleinste σ-Algebra auf X, die alle Koordinatenprojektionen messbar macht. Denn für $A_i \in \mathscr{A}_i$ folgt $\operatorname{pr}_i^{-1}(A_i) = X_1 \times \ldots \times A_i \times \ldots \times X_n$, $1 \leq i \leq n$. Somit ist $A_1 \times \ldots \times A_n = \bigcap_{i=1}^n \operatorname{pr}_i^{-1}(A_i) \in \sigma(\operatorname{pr}_i : 1 \leq i \leq n)$. Damit gilt "$\subset$".
"$\supset$" folgt sofort aus $\operatorname{pr}_i^{-1}(A_i) \in \bigotimes_{i=1}^n \mathscr{A}_i$ wie eben gesehen.

d) Die Eigenschaft in c) kann verwendet werden, um das Produkt von beliebig vielen σ-Algebren zu definieren: Sei $I \neq \emptyset$ eine Indexmenge, und zu $i \in I$ sei $(X_i, \mathscr{A}_i)$ ein Messraum. Dann wird auf $X := \prod_{i\in I} X_i$ die Produkt-σ-Algebra definiert durch $\bigoplus_{i\in I} \mathscr{A}_i := \sigma(\{\operatorname{pr}_i : i \in I\})$.

3.4 Satz. *Zu $1 \leq i \leq n$ sei jeweils $\mathscr{E}_i \subset \mathscr{A}_i$ ein Erzeuger von $\mathscr{A}_i$ (also $\sigma(\mathscr{E}_i) = \mathscr{A}_i$) derart, dass es stets eine Folge $(E_{ik})_{k\in\mathbb{N}}$ in $\mathscr{E}_i$ gibt mit $E_{ik} \nearrow X_i$ für $k \to \infty$. Dann gilt*

$$\sigma(\{E_1 \times \ldots \times E_n : E_i \in \mathscr{E}_i, 1 \leq i \leq n\}) = \bigotimes_{i=1}^n \mathscr{A}_i.$$

Beweis. "$\subset$" ist klar und für die umgekehrte Inklusion ist zu zeigen, dass $A_1 \times \ldots \times A_n \in \sigma(\{E_1 \times \ldots \times E_n\})$ gilt für alle $A_i \in \mathscr{A}_i$. Dazu wiederum reicht es, $X_1 \times \ldots \times A_i \times \ldots \times X_n \in \sigma(\{E_1 \times \ldots \times E_n\})$ zu verifizieren. Hierfür betrachte zu gegebenem $E_i \in \mathscr{E}_i$

$$F_k := E_{1k} \times \ldots \times E_{i-1,k} \times E_i \times E_{i+1,k} \times \ldots \times E_{nk} \in \sigma(\{E_1 \times \ldots \times E_n\}).$$

Dann ist auch $X_1 \times \ldots \times E_i \times \ldots \times X_n = \bigcup_{k\in\mathbb{N}} F_k$ in $\sigma(\{E_1 \times \ldots \times E_n\})$. Wegen

$$\begin{aligned} X_1 \times \ldots \times A_i \times \ldots \times X_n &\in X_1 \times \ldots \times \sigma(\mathscr{E}_i) \times \ldots \times X_n \\ &= \sigma(\{X_1 \times \ldots \times E_i \times \ldots \times X_n\}) \end{aligned}$$

folgt die Behauptung. □

3.5 Satz. *Es gilt*

$$\mathscr{B}(\mathbb{R}^n) = \bigotimes_{i=1}^n \mathscr{B}(\mathbb{R}).$$

Beweis. Dies folgt aus Satz 3.4 mit $\mathscr{E}_i = \{(a_i, b_i] : a_i, b_i \in \mathbb{R}\}$. Man beachte $(a, b] = \prod_{i=1}^n (a_i, b_i]$ und $\sigma(\mathbb{I}_n) = \mathscr{B}(\mathbb{R}^n)$. □

b) Die Sätze von Tonelli und Fubini

Seien $(X_i, \mathscr{A}_i, \mu_i)$, $1 \leq i \leq n$, nun endlich viele Maßräume.

3.6 Definition. Ein Maß $\mu : \bigotimes_{i=1}^n \mathscr{A}_i \to [0, \infty]$ heißt *Produktmaß der* μ_i, wenn

$$\mu(A_1 \times \ldots \times A_n) = \prod_{i=1}^n \mu_i(A_i) \quad \text{für alle } A_i \in \mathscr{A}_i, 1 \leq i \leq n$$

(mit der Konvention $0 \cdot \infty = 0$).

Wir konstruieren das Produktmaß nur für zwei Faktoren, Produkte höherer Ordnung definiert man dann induktiv in Verbindung mit geeigneten Assoziativitätsüberlegungen.

Achtung: Die Existenz und Eindeutigkeit des Produktmaßes ist nur für das Produkt σ-endlicher Maßräume gewährleistet!

3.7 Lemma. *Seien $(X_i, \mathscr{A}_i, \mu_i)$, $i = 1, 2$, σ-endliche Maßräume und sei $f : X_1 \times X_2 \to [0, \infty]$, $(x, y) \mapsto f(x, y)$, $\mathscr{A}_1 \otimes \mathscr{A}_2$-messbar. Dann gilt:*

(i) Für jedes $x \in X_1$ ist $f(x, \cdot) : X_2 \to [0, \infty]$, $y \mapsto f(x, y)$, $\mathscr{A}_2$-messbar.

(ii) Die Abbildung $X_1 \to [0, \infty]$, $x \mapsto \int f(x, y) d\mu_2(y)$ ist $\mathscr{A}_1$-messbar.

Beweis. (1) Sei zunächst μ_2 endlich. Dazu betrachten wir das Teilmengensystem der Produkt-σ-Algebra

$$\mathscr{D} := \{A \in \mathscr{A}_1 \otimes \mathscr{A}_2 : \chi_A \text{ hat die Eigenschaften (i) und (ii)}\}.$$

Wir zeigen, daß $\mathscr{D}$ ein Dynkin-System ist, wobei wir die Kriterien aus Bemerkung 1.3 c) nachprüfen.

- $X = X_1 \times X_2 \in \mathscr{D}$, da $\chi_X = 1$.
- $\mathscr{D} \ni A_l \nearrow A$, dann auch $A \in \mathscr{D}$: Es gilt $\chi_{A_l} \nearrow \chi_A$ und damit folgt (i), da der punktweise Limes messbarer Funktionen wieder messbar ist, und (ii) mit monotoner Konvergenz.
- $A, B \in \mathscr{D}$, $A \subset B$, dann auch $B \setminus A \in \mathscr{D}$: Es ist $\chi_{B\setminus A} = \chi_B - \chi_A$ und (i) ist klar. (ii) folgt aus der Endlichkeit von μ_2, denn damit ist $\int \chi_{B\setminus A}(x, \cdot) d\mu_2 = \int \chi_B(x, \cdot) d\mu_2 - \int \chi_A(x, \cdot) d\mu_2$ („$\infty - \infty$“ kann nicht auftreten).

Außerdem umfaßt $\mathscr{D}$ den $\cap$-stabilen Erzeuger $\mathscr{Z}$ der Zylindermengen, denn $A = A_1 \times A_2 \in \mathscr{Z}$, dann ist $\chi_A(x, y) = \chi_{A_1}(x)\chi_{A_2}(y)$, woraus sich (i) ergibt und mit $\int \chi_A(x, \cdot) d\mu_2 = \int \chi_{A_1}(x)\chi_{A_2} d\mu_2 = \chi_{A_1}(x)\mu_2(A_2)$ folgt (ii).

Das Dynkin-Lemma besagt nun $\mathscr{D} = \mathscr{A}_1 \otimes \mathscr{A}_2$, was gleichbedeutend damit ist, daß die Behauptung für alle χ_A mit $A \in \mathscr{A}_1 \otimes \mathscr{A}_2$ richtig ist. Wegen Linearität folgt die Behauptung für alle Stufenfunktionen, und mit den üblichen Konvergenzargumenten (punktweise Limiten messbarer Funktionen sind messbar und Satz von der monotonen Konvergenz) daher für alle $[0, \infty]$-wertigen produktmessbaren Funktionen.

(2) Sei nun μ_2 σ-endlich. Wähle $\mathscr{A}_2 \ni S_k \nearrow X_2$ mit $\mu_2(S_k) < \infty$ und betrachte statt f nun $f_k := f \cdot \chi_{X_1 \times S_k}$, das bezüglich der zweiten Komponente auf dem endlichen Maßraum $(X_2; \mathscr{A}_2; \mu_2(S_k \cap \cdot\,))$ lebt.

Nach (1) stimmt die Behauptung für die f_k und damit wegen $f_k \nearrow f$ punktweise auch für f, mit den gleichen Konvergenzargumenten wie oben. $\square$

3.8 Lemma. *Seien $(X_i, \mathscr{A}_i, \mu_i)$, $i = 1, 2$, σ-endliche Maßräume. Dann existiert ein eindeutiges Produktmaß $\mu_1 \otimes \mu_2$ auf $\mathscr{A}_1 \otimes \mathscr{A}_2$. Explizit ist es gegeben durch*

$$(\mu_1 \otimes \mu_2)(A) = \int \left(\int \chi_A(x, \cdot\,) d\mu_2 \right) d\mu_1(x) = \int \left(\int \chi_A(\,\cdot\,, y) d\mu_1 \right) d\mu_2(y).$$

Beweis. (i) $\mu_1 \otimes \mu_2$ ist ein Maß: Klar ist $\mu_1 \otimes \mu_2(\emptyset) = \int \left(\int \chi_\emptyset(x, \cdot\,) d\mu_2 \right) d\mu_1(x) = \int 0 \, d\mu_1 = 0$.

$\mu_1 \otimes \mu_2$ ist additiv wegen $\chi_{A_1 \dot\cup A_2} = \chi_{A_1} + \chi_{A_2}$ und damit

$$\begin{aligned}(\mu_1 \otimes \mu_2)(A_1 \,\dot\cup\, A_2) &= \int \left(\int \chi_{A_1}(x, \cdot\,) + \chi_{A_2}(x, \cdot\,) d\mu_2 \right) d\mu_1(x) \\ &= (\mu_1 \otimes \mu_2)(A_1) + (\mu_1 \otimes \mu_2)(A_2)\end{aligned}$$

wegen Linearität der Integrale.

Für die σ-Additivität ist somit noch die Stetigkeit von unten zu zeigen (Satz 1.6). Sei daher $\mathscr{A}_1 \otimes \mathscr{A}_2 \ni A_l \nearrow A$. Dann folgt wegen $\chi_{A_l} \nearrow \chi_A$

$$\begin{aligned}\mu_1 \otimes \mu_2(A_l) &= \int \left(\int \chi_{A_l}(x, \cdot\,) d\mu_2 \right) d\mu_1(x) \\ &\nearrow \int \left(\int \chi_A(x, \cdot\,) d\mu_2 \right) d\mu_1(x) = \mu_1 \otimes \mu_2(A)\end{aligned}$$

durch zweimalige Anwendung monotoner Konvergenz.

(ii) $\mu_1 \otimes \mu_2$ ist Produktmaß von $\mu_1 \otimes \mu_2$:

$$\mu_1 \otimes \mu_2(A_1 \times A_2) = \int \left(\int \chi_{A_1}(x) \chi_{A_2} d\mu_2 \right) d\mu_1(x)$$

$$= \int \chi_{A_1}(x)\mu_2(A_2)d\mu_1(x) = \mu_2(A_2)\mu_1(A_1).$$

(iii) Zur Eindeutigkeit: $\mu_1 \otimes \mu_2$ liegt durch $\mu_1 \otimes \mu_2(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2)$ fest auf dem $\cap$-stabilen Erzeuger $\mathscr{Z}$ von $\mathscr{A}_1 \otimes \mathscr{A}_2$. Da die beiden Maße σ-endlich sind, gibt es $\mathscr{A}_1 \ni S_k \nearrow X_1$, $\mu_1(S_k) < \infty$, sowie $\mathscr{A}_2 \ni T_k \nearrow X_2$, $\mu_2(T_k) < \infty$. Damit ergibt sich $S_k \times T_k \nearrow X_1 \times X_2$ mit $\mu_1 \otimes \mu_2(S_k \times T_k) = \mu_1(S_k)\mu_2(T_k) < \infty$ und die Eindeutigkeit folgt aus dem Eindeutigkeitssatz 1.12.

(iv) Genauso zeigt man $\mu_1 \otimes \mu_2(A) = \int \left(\int \chi_A(\,\cdot\,, y)d\mu_1\right) d\mu_2(y)$. $\square$

3.9 Bemerkung. a) $(\mu_1 \otimes \mu_2) \otimes \mu_3 = \mu_1 \otimes (\mu_2 \otimes \mu_3) =: \mu_1 \otimes \mu_2 \otimes \mu_3$ usw. für höhere Produkte. Denn die Gleichung stimmt für beliebige $A = A_1 \times A_2 \times A_3 \in \mathscr{Z}$ (iteriertes Ausrechnen der Integrale) und gibt ein Produktmaß auf $\mathscr{A}_1 \otimes \mathscr{A}_2 \otimes \mathscr{A}_3$. Der Eindeutigkeitssatz für Maße liefert wie vorhin die Behauptung.

b) Für das Lebesgue-Maß $\lambda_n \colon \mathscr{B}(\mathbb{R}^n) \to [0, \infty]$ folgt insbesondere

$$\lambda_n = \bigotimes_{j=1}^{n} \lambda_1.$$

Denn beide Seiten dieser Gleichung stimmen auf $\mathbb{I}_n$ überein.

c) Die Voraussetzung der σ-Endlichkeit ist für die Aussage von Lemma 3.8 wesentlich: Seien $X_1 = X_2 = [0, 1]$ beide versehen mit den Borelmengen und $\mu_1 = \lambda|_{[0,1]}$ sowie $\mu_2 = \zeta$ das Zählmaß, das auf dem überabzählbaren Intervall nicht σ-finit ist. Wir betrachten die Diagonale

$$\Delta := \{(x, x) : x \in [0, 1]\} = f^{-1}\{0\} \in \mathscr{B}(\mathbb{R}) \otimes \mathscr{B}(\mathbb{R}),$$

wobei $f : [0, 1]^2 \to \mathbb{R}$, $(x, y) \mapsto x - y$, stetig und damit $\mathscr{B}(\mathbb{R}^2) = \mathscr{B}(\mathbb{R}) \otimes \mathscr{B}(\mathbb{R})$-messbar ist. Aber nun berechnet sich (kurz λ für $\lambda|_{[0,1]}$)

$$\int \Big(\int \chi_\Delta(x, y)d\lambda(x)\Big)d\zeta(y) = \int \lambda(\{y\})d\zeta(y) = \int 0 \; d\zeta(y) = 0$$

im Gegensatz zu

$$\int \Big(\int \chi_\Delta(x, y)d\zeta(y)\Big)d\lambda(x) = \int \zeta(\{x\})d\lambda(x) = \int 1 \; d\lambda(x) = 1.$$

3.10 Satz (von Tonelli)**.** *Seien $(X_i; \mathscr{A}_i; \mu_i)$, $i = 1, 2$ zwei σ-endliche Maßräume sowie $\mu := \mu_1 \otimes \mu_2$ das Produktmaß auf der Produkt-σ-Algebra $\mathscr{A}_1 \otimes \mathscr{A}_2$. Sei $f : X_1 \times X_2 \to [0, \infty]$ $\mathscr{A}_1 \otimes \mathscr{A}_2$-messbar. Dann gilt in $[0, \infty]$ die Gleichheit*

$$\int f d\mu = \int \Big(\int f(x, y)d\mu_1(x)\Big)d\mu_2(y) = \int \Big(\int f(x, y)d\mu_2(y)\Big)d\mu_1(x).$$

Beweis. Nach Lemma 3.7 existieren alle Integrale in $[0,\infty]$. Nach Lemma 3.8 gilt die Behauptung für Indikatorfunktionen χ_A, $A \in \mathscr{A}_1 \otimes \mathscr{A}_2$. Mittels Linearität folgt die Behauptung für Stufenfunktionen, schließlich mit monotoner Konvergenz für alle produktmessbaren $[0,\infty]$-wertigen Funktionen. □

3.11 Satz (von Fubini). *Seien $(X_i;\mathscr{A}_i;\mu_i)$, $i=1,2$ zwei σ-endliche Maßräume sowie $\mu := \mu_1 \otimes \mu_2$ das Produktmaß auf der Produkt-σ-Algebra $\mathscr{A}_1 \otimes \mathscr{A}_2$. Sei $f : X_1 \times X_2 \to \overline{\mathbb{R}}$ mit $f \in \mathscr{L}^1(\mu)$, d.h. f sei integrierbar bzgl. des Produktmaßes.*

a) Es gilt $N := \{x \in X_1 : \int |f(x,y)| d\mu_2(y) = \infty\} \in \mathscr{A}_1$ mit $\mu_1(N) = 0$.

b) Definiere $g : X_1 \to \overline{\mathbb{R}}$ durch

$$g(x) := \begin{cases} \int f(x,y) d\mu_2(y), & x \notin N, \\ 0, & x \in N. \end{cases}$$

Dann ist $g \in \mathscr{L}^1(\mu_1)$, und es gilt

$$\int f d\mu = \int g d\mu_1 = \int \Big(\int f(x,y) d\mu_2(y) \Big) d\mu_1(x).$$

Die entsprechende Aussage gilt mit vertauschten Rollen von μ_1 und μ_2. Insbesondere ist

$$\int \Big(\int f(x,y) d\mu_2(y) \Big) d\mu_1(x) = \int \Big(\int f(x,y) d\mu_1(x) \Big) d\mu_2(y).$$

Beweis. Lemma 3.7 für $|f|$ liefert die $\mathscr{A}_1$-Messbarkeit von $x \mapsto \int |f(x,y)| d\mu_2(y)$, damit ist $N \in \mathscr{A}_1$. Wegen $|f| \in \mathscr{L}^1(\mu)$ folgt aus dem Satz von Tonelli

$$\int \Big(\int |f(x,y)| d\mu_2(y) \Big) d\mu_1(x) = \int |f| d\mu < \infty$$

und damit $\mu_1(N) = 0$.

Also ist auch $\mu(N \times X_2) = \mu_1(N)\mu_2(X_2) = 0$. Betrachte nun $\widetilde{f} := f \cdot \chi_{(N\times X_2)^c} = f \cdot \chi_{N^c \times X_2}$, dann ist $\widetilde{f} = f$ μ-fast überall. Nun folgt für $f \geq 0$ (und damit auch $\widetilde{f} \geq 0$)

$$\int g d\mu_1 = \int \Big(\int \widetilde{f}(x,\cdot) d\mu_2 \Big) d\mu_1(x) = \int \widetilde{f} d\mu = \int f d\mu,$$

wobei die mittlere Gleichheit wieder nach Tonelli folgt. Für ein beliebiges komplexwertiges f folgt dann die Behauptung durch die Zerlegung $f = f_+ - f_-$. □

In Anwendungen wird der Satz von Fubini häufig folgendermaßen verwendet: Man zeigt zunächst, dass die Funktion f messbar ist (häufig als stetige Funktion oder als Stufenfunktion oder als punktweise Grenzwert messbarer Funktionen). Dann berechnet man das iterierte Integral $\int\int |f(x,y)| d\mu_1(x) d\mu_2(y)$, wobei die Reihenfolge beliebig gewählt werden kann. Falls dieses Integral endlich ist, folgt nach dem Satz von Tonelli $f \in \mathscr{L}^1(\mu_1 \otimes \mu_2)$, und man kann den Satz von Fubini anwenden, indem man jetzt $\int\int f(x,y) d\mu_1(x) d\mu_2(y)$ wiederum in einer beliebigen Reihenfolge berechnen darf.

Die Folgerungen aus dem Satz von Fubini sind bereits aus der Analysis II bekannt, etwa das Prinzip von Cavalieri und die Tatsache, dass Graphen von messbaren Funktionen Lebesgue-Nullmengen sind.

4. Der Transformationssatz

4.1 Worum geht's? Der Transformationssatz ist ein weiterer wichtiger Satz der Lebesgueschen Integrationstheorie. Die Anwendungen des Transformationssatzes sind bereits aus der Analysis II bekannt und werden hier nicht mehr besprochen. Als Beispiel sei hier nur an die Definition des Flächeninhalts für eine m-dimensionale Fläche des $\mathbb{R}^n$ und den Nachweis der Wohldefiniertheit erinnert. Wir wissen bereits, dass der Transformationssatz für lineare Abbildungen gilt. Dies wird jetzt verwendet, um den allgemeinen Satz zu beweisen.

4.2 Definition und Satz. *Sei $(X, \mathscr{A}, \mu)$ ein Maßraum, $(Y, \mathscr{B})$ ein Messraum und $\Phi\colon X \to Y$ messbar. Dann ist*

$$\mu \circ \Phi^{-1}\colon \mathscr{B} \to [0, \infty],\ (\mu \circ \Phi^{-1})(B) := \mu(\Phi^{-1}(B))$$

ein Maß, das Bildmaß von μ unter Φ.

Beweis. Die Eigenschaften eines Maßes übertragen sich unmittelbar bei Verwendung des Urbilds. □

In der Sprache der Stochastik sieht die letzte Definition folgendermaßen aus: Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum, und sei $X\colon \Omega \to \mathbb{R}$ eine Zufallsvariable. Dann ist $\mathbb{P} \circ X^{-1}\colon \mathscr{B}(\mathbb{R}) \to [0, 1]$ ein Wahrscheinlichkeitsmaß, die sogenannte (Wahrscheinlichkeits-)Verteilung von X. Bekannte Verteilungen sind die Normalverteilung und die Gleichverteilung auf einem Intervall oder auf einer endlichen Menge. Nach Lemma 1.18 ist die Wahrscheinlichkeitsverteilung von X bereits eindeutig durch die Verteilungsfunktion

$$F_X\colon \mathbb{R} \to [0, 1],\ z \mapsto F_X(z) := \mathbb{P}(\{\omega \in \Omega : X(\omega) \leq z\}) =: \mathbb{P}(X \leq z)$$

festgelegt.

4.3 Lemma (Transformationslemma)**.** *Seien $(X, \mathscr{A}, \mu)$ ein Maßraum, $(Y, \mathscr{B})$ ein Messraum und $\Phi\colon X \to Y$ messbar. Sei weiter $f\colon Y \to \overline{\mathbb{R}}$ $\mathscr{B}$-messbar. Dann ist f genau dann $\mu \circ \Phi^{-1}$-integrierbar, wenn $f \circ \Phi$ μ-integrierbar ist, und in diesem Fall ist*

$$\int (f \circ \Phi) d\mu = \int f d(\mu \circ \Phi^{-1}).$$

Beweis. Für $f = \chi_A$ mit $A \in \mathscr{A}$ stimmt dies nach Definition des Bildmaßes und wegen $\chi_A(\Phi(x)) = \chi_{\Phi^{-1}(A)}(x)$, daher wegen Linearität des Integrals auch für Stufenfunktionen $f \geq 0$, mit monotoner Konvergenz für alle messbaren Funktionen $f \geq 0$. Der allgemeine Fall folgt wieder durch die Zerlegung $f = f_+ - f_-$. □

4.4 Beispiel (Maße mit Dichten)**.** Sei $(X, \mathscr{A}, \mu)$ ein Maßraum und $f\colon X \to [0, \infty]$ messbar. Dann ist nach Satz 2.5 durch

$$\nu(A) := \int_A f d\mu \quad (A \in \mathscr{A})$$

ein Maß gegeben. Eine messbare Funktion $g\colon X \to \overline{\mathbb{R}}$ ist genau dann ν-integrierbar, wenn $g \cdot f$ μ-integrierbar ist, und in diesem Fall ist

$$\int g d\nu = \int g \cdot f d\mu.$$

Dies folgt genauso wie im Beweis des Transformationslemmas 4.3. Man sagt, das Maß ν besitzt die μ-Dichte f und schreibt $\nu = f \cdot \mu$. In symbolischer Schreibweise $d\nu = f d\mu$. Die Dichte f heißt auch die Radon-Nikodym-Ableitung von ν nach μ, symbolisch $f = \frac{d\nu}{d\mu}$.

4.5 Beispiel (Riemann-Stieltjes-Integrale)**.** Seien $a, b \in \mathbb{R}$ mit $a < b$ und $g\colon [a, b] \to \mathbb{R}$ eine monoton wachsende und rechtsstetige Funktion. Dann existiert genau ein Borelmaß $\mu_g\colon \mathscr{B}([a, b]) \to [0, \infty]$ mit

$$\mu_g((s, t]) = g(t) - g(s) \quad (a \leq s \leq t \leq b)$$

und $\mu_g(\{a\}) = 0$.

Um dies zu sehen, setzen wir g konstant auf $(-\infty, a]$ und $[b, \infty)$ fort. Durch die obige Gleichung ist μ_g bereits auf $\mathbb{I}_n$ und damit auf $\mathbb{A}_n$ eindeutig festgelegt. Man rechnet nach, dass die Rechts-Stetigkeit von g gerade die σ-Additivität von μ_g auf $\mathbb{A}_n$ bedeutet. Wegen $\mu_g(\mathbb{R}) = \mu_g([a, b]) = g(b) - g(a)$ ist μ_g endlich. Damit existiert genau eine Maßfortsetzung von $\mu_g|_{\mathbb{A}_n}$ zu einem Borelmaß $\mu_g|_{\mathscr{B}(\mathbb{R})}$.

Das zu μ_g gehörige Integral

$$\int f dg := \int f d\mu_g$$

heißt auch das Riemann-Stieltjes-Integral von f bzgl. g. Die Theorie der Riemann-Stieltjes-Integrale kann analog zum Riemann-Integral auch unabhängig von der Lebesgue-Theorie entwickelt werden, indem man Zerlegungssummen der Form

$$\sum_{k=1}^{K} f(y_k)\big[g(x_{k+1}) - g(x_k)\big]$$

mit $y_k \in (x_k, x_{k+1}]$ betrachtet.

Eine kompakte Ausschöpfung und das übliche σ-Additivitätsargument zeigt, dass die obige Aussage auch für monoton wachsende und rechtsstetige $g\colon \mathbb{R} \to \mathbb{R}$ gilt. Wählt man $g(x) = x$, erhält man das Riemann-Integral als Spezialfall des Riemann-Stieltjes-Integrals.

4.6 Lemma. *Sei $U \subset \mathbb{R}^n$ offen. Dann existieren paarweise disjunkte Intervalle $I_k \in \mathbb{I}_n$ $(k \in \mathbb{N})$ mit $U = \dot{\bigcup}_{k\in\mathbb{N}} I_k$.*

Beweis. Wir wissen bereits aus dem Beweis von Lemma 1.17, dass $U = \bigcup_{\ell\in\mathbb{N}} U_\ell$ mit offenen Intervallen $U_\ell = (a_\ell, b_\ell)$ gilt. Wegen $U_\ell = \bigcup_{m\in\mathbb{N}}(a_\ell, b_\ell - \frac{1}{m}]$ ist U_ℓ und damit auch U abzählbare Vereinigung von halboffenen Intervallen, d.h. $U = \bigcup_{k\in\mathbb{N}} I_k$ mit $I_k \in \mathbb{I}_n$.

Setze nun $A_1 := I_1$ und induktiv $A_m := I_m \setminus (A_1 \cup \dots \cup A_{m-1})$. Da $\mathbb{A}_n$ ein Ring ist, gilt $A_m \in \mathbb{A}_n$ $(m \in \mathbb{N})$. Somit ist jedes A_m disjunkte Vereinigung von endlich vielen Intervallen $I_{km} \in \mathbb{I}_n$. Insgesamt erhält man $U = \dot{\bigcup}_{k,m} I_{km}$. □

Wir wissen bereits aus Analysis II, dass das Lebesgue-Maß $\lambda = \lambda_n \colon \mathscr{B}(\mathbb{R}^n) \to [0,\infty]$ bewegungsinvariant ist, und dass für invertierbare lineare Abbildungen $T \in GL(n,\mathbb{R})$ die Gleichheit

$$\lambda(T(B)) = |\det T| \cdot \lambda(B) \quad (B \in \mathscr{B}(\mathbb{R}^n)) \tag{4-1}$$

gilt. Die Verallgemeinerung dieser Aussage auf C^1-Diffeomorphismen wird Transformationssatz oder auch Substitutionssatz oder Substitutionsregel für das n-dimensionale Lebesgue-Integral genannt.

4.7 Satz (Transformationssatz). *Seien $U, V \subset \mathbb{R}^n$ offen und $\Phi \colon U \to V$ ein C^1-Diffeomorphismus. Eine messbare Funktion $f \colon V \to \mathbb{R}$ ist über $V = \Phi(U)$ genau dann λ-integrierbar, wenn $(f \circ \Phi) \cdot |\det \Phi'| \colon U \to \mathbb{C}$ λ-integrierbar über U ist. In diesem Fall gilt*

$$\int_{\Phi(U)} f(y)dy = \int_U f\big(\Phi(x)\big) \cdot |\det \Phi'(x)| dx.$$

Beweis. **(i)** Wir zeigen zunächst, dass für jeden abgeschlossenen Würfel $A = \overline{A} \subset U$ der Kantenlänge $d > 0$ die Ungleichung

$$\lambda(\Phi(A)) \le \int_A |\det \Phi'(x)| dx \tag{4-2}$$

gilt. Wir verwenden dazu die Maximumnorm $|x|_\infty := \max\{|x_1|, \dots, |x_n|\}$ im $\mathbb{R}^n$ und die zugehörige Operatornorm $\|\cdot\|_\infty$. Da Φ C^1-Diffeomorphismus ist, existiert $M := \max_{x\in A} \|\Phi'(x)^{-1}\|_\infty < \infty$. Sei $\varepsilon > 0$ gegeben. Als stetige Funktion auf einer kompakten Menge ist Φ' auf A gleichmäßig stetig, d.h. es existiert ein $\delta > 0$ mit

$$\|\Phi'(x) - \Phi'(y)\|_\infty \le \frac{\varepsilon}{M} \quad (|x-y|_\infty \le \delta).$$

O.E. sei dabei $\delta = d/N$ mit einem $N \in \mathbb{N}$. Wir überdecken den Würfel A durch N^n abgeschlossene Würfel A_j, $j = 1, \dots, N^m$ mit Kantenlänge δ so, dass diese sich nur in Nullmengen schneiden.

Zu $j = 1, \dots, N^n$ wähle $a_j \in A_j$ mit $|\det \Phi'(a_j)| = \min_{x \in A_j} |\det \Phi'(x)|$. Zu festem j sei

$$h(x) := (h_1(x), \dots, h_n(x))^T := \Phi(x) - \Phi'(a_j)x \quad (x \in A_j).$$

Nach dem Mittelwertsatz der Differentialrechnung existiert zu $x \in A_j$ ein $c_i \in A_j$ mit

$$\begin{aligned} |h_i(x) - h_i(a_j)| &= |h_i'(c_i)(x - a_j)| = |(\Phi_i'(c_i) - \Phi_i'(a_j))(x - a_j)| \\ &\le \|\Phi'(c_i) - \Phi'(a_j)\|_\infty |x - a_j|_\infty \\ &\le \frac{\varepsilon}{M}\,\delta \quad (i = 1, \dots, n). \end{aligned}$$

Nimmt man das Maximum über alle $i = 1, \dots, n$, so erhält man

$$|\Phi(x) - \Phi'(a_j)x - \Phi(a_j) + \Phi'(a_j)a_j|_\infty \le \frac{\varepsilon\delta}{M} \quad (x \in A_j)$$

und somit

$$\Phi(A_j) \subset \Phi(a_j) - \Phi'(a_j)a_j + \Phi'(a_j)(A_j) + [-\tfrac{\varepsilon\delta}{M}, \tfrac{\varepsilon\delta}{M}]^n.$$

Nach Definition von M gilt

$$[-\tfrac{\varepsilon\delta}{M}, \tfrac{\varepsilon\delta}{M}]^n = \Phi'(a_j)\Phi'(a_j)^{-1}\big([-\tfrac{\varepsilon\delta}{M}, \tfrac{\varepsilon\delta}{M}]^n\big) \subset \Phi'(a_j)\big([-\varepsilon\delta, \varepsilon\delta]^n\big).$$

Somit ist

$$\Phi(A_j) \subset \Phi(a_j) - \Phi'(a_j)a_j + \Phi'(a_j)\big(A_j + [-\varepsilon\delta, \varepsilon\delta]^n\big),$$

und mit der Translationsinvarianz des Lebesgue-Maßes und (4-1) folgt

$$\begin{aligned} \lambda(\Phi(A_j)) &\le \lambda\Big(\Phi'(a_j)\big(A_j + [-\varepsilon\delta, \varepsilon\delta]^n\big)\Big) \\ &= |\det \Phi'(a_j)|\lambda\big(A_j + [-\varepsilon\delta, \varepsilon\delta]^n\big) \\ &= |\det \Phi'(a_j)|\,(\delta + 2\varepsilon\delta)^n \\ &= |\det \Phi'(a_j)|\,\lambda(A_j)(1 + 2\varepsilon)^n. \end{aligned}$$

Wir summieren über alle $j = 1, \dots, N^n$ und erhalten

$$\begin{aligned} \lambda(\Phi(A)) &\le \sum_{j=1}^{N^n} \lambda(\Phi(A_j)) \le (1 + 2\varepsilon)^n \sum_{j=1}^{N^n} |\det \Phi'(a_j)|\lambda(A_j) \\ &\le (1 + 2\varepsilon)^n \sum_{j=1}^{N^n} \int_{A_j} |\det \Phi'(x)|dx = (1 + 2\varepsilon)^n \int_A |\det \Phi'(x)|dx. \end{aligned}$$

Hierbei wurde die Definition der a_j verwendet. Da $\varepsilon > 0$ beliebig war, folgt (4-2).

(ii) Wir zeigen nun, dass (4-2) für alle Borelmengen $A \in \mathscr{B}(U)$ gilt.

(1) Da jedes Intervall $(a, b] \subset U$ durch eine Folge endlicher Vereinigungen von Würfeln approximiert werden kann, gilt die Ungleichung (4-2) auch für $A = (a, b]$. (Man beachte hierbei, dass beide Seiten von (4-2) Maße und damit stetig von unten sind.)

(2) Sei nun $A \subset U$ offen. Nach Lemma 4.6 ist $A = \dot{\bigcup}_{k\in\mathbb{N}} I_k$ mit $I_k \in \mathbb{I}_n$. Somit gilt

$$\lambda(\Phi(A)) \le \sum_{k\in\mathbb{N}} \lambda(\Phi(I_k)) \le \sum_{k\in\mathbb{N}} \int_{I_k} |\det \Phi'(x)|dx = \int_A |\det \Phi'(x)|dx.$$

(3) Sei schließlich $A \in \mathscr{B}(U)$, wobei o.E. $\overline{A} \subset U$ kompakt sei. (Sonst betrachte $A = \bigcup_{k\in\mathbb{N}}(A \cap U_k)$ mit $U_k := \{x \in \mathbb{R}^n : |x| \le k, \operatorname{dist}(x, \mathbb{R}^n \setminus U) \ge \frac{1}{k}\}$.) Da das Lebesgue-Maß regulär ist, existiert eine Folge $A_k \supset A$, $k \in \mathbb{N}$, mit A_k offen und $\lambda(A_k \setminus A) \to 0$ $(k \to \infty)$, wobei o.E. $\overline{A}_k \subset U$ kompakt sei. Es folgt

$$\lambda(\Phi(A)) \le \lambda(\Phi(A_k)) \le \int_{A_k} |\det \Phi'(x)|dx \to \int_A |\det \Phi'(x)|dx \quad (k \to \infty).$$

Hier wurden $\int_{A_k} |\det \Phi'(x)|dx \le \lambda(A_k) \max_{x\in\overline{A}_k} |\det \Phi'(x)| < \infty$ und majorisierte Konvergenz verwendet.

(iii) Zum Beweis des Transformationssatzes sei zunächst $f = \chi_B$ mit $B \in \mathscr{B}(V)$. Dann gilt mit $A := \Phi^{-1}(B)$

$$\begin{aligned}
\int_V \chi_B(y)dy &= \lambda(B) = \lambda(\Phi(A)) \\
&\le \int_A |\det \Phi'(x)|dx = \int_U \chi_A(x)|\det \Phi'(x)|dx \\
&= \int_U \chi_B(\Phi(x))|\det \Phi'(x)|dx.
\end{aligned}$$

Wegen Linearität gilt

$$\int_V f(y)dy \le \int_U f(\Phi(x))|\det \Phi'(x)|dx$$

für alle Stufenfunktionen $f \colon V \to [0, \infty)$ und mit dem Satz von der monotonen Konvergenz für alle messbaren Funktionen $f \colon V \to [0, \infty]$. Wendet man dies auch auf Φ^{-1} anstelle von Φ an, so erhält man für messbares $f \ge 0$

$$\begin{aligned}
\int_V f(y)dy &\le \int_U (f \circ \Phi)(x)|\det \Phi'(x)|dx \\
&\le \int_V (f \circ \Phi \circ \Phi^{-1})(y)|\det \Phi'(\Phi^{-1}(y))| \cdot |\det(\Phi^{-1})'(y)|dy
\end{aligned}$$

$$= \int_V f(y)dy.$$

Dabei wurde $(\Phi^{-1})'(y) = \big(\Phi'(\Phi^{-1}(y))\big)^{-1}$ verwendet.

Die Aussagen des Transformationssatzes ergibt sich nun durch die Zerlegung $f = f_+ - f_-$ und die Anwendung des bisher Bewiesenen auf $f_\pm$. □

4.8 Bemerkung. a) Im Beweis des Transformationssatzes wurde ein Spezialfall der Sardschen Ungleichung gezeigt: Sei $U \subset \mathbb{R}^n$ offen, $\Phi\colon U \to \mathbb{R}^n$ stetig differenzierbar und $K \subset U$ kompakt. Dann gilt

$$\lambda(\Phi(K)) \le \int_K |\det \Phi'(x)| dx.$$

b) Aus dem Transformationssatz folgt wiederum, dass das Lebesgue-Maß bewegungsinvariant ist. Ebenso folgt für jedes $h \in \mathscr{L}^1(\mathbb{R}^n)$ und $\alpha \in \mathbb{R} \setminus \{0\}$ sofort die Gleichung

$$\int h(\alpha x)dx = |\alpha|^{-n} \int h(x)dx.$$

5. Der Satz von Radon-Nikodym

5.1 Worum geht's? Wir wissen bereits, dass durch $A \mapsto \varphi_f(A) := \int_A f d\mu$ für festes messbares $f \geq 0$ ein Maß definiert wird. In diesem Falle heißt f eine Dichte von φ_f bzgl. μ. Das Maß φ_f ist absolutstetig bzgl. μ im Sinne, dass aus $\mu(A) = 0$ auch $\varphi_f(A) = 0$ folgt. Der Satz von Radon-Nikodym besagt, dass aus der Bedingung der Absolutstetigkeit auch schon die Existenz einer Dichte folgt. Dieser wichtige Satz der Maßtheorie hat unter anderem Anwendungen in der Stochastik, da damit die Existenz bedingter Erwartungswerte gezeigt werden kann.

Im Folgenden sei stets $(X, \mathscr{A}, \mu)$ ein Maßraum. Zu $f\colon X \to [0, \infty]$, f messbar, sei

$$\mu_f(A) := \int_A f d\mu \quad (A \in \mathscr{A}).$$

Man schreibt mit $\varphi := \mu_f$ auch $(f\mu)(A) := \mu_f(A)$ $(A \in \mathscr{A})$ oder $d\mu_f = f d\mu$ sowie

$$\frac{d\mu_f}{d\mu} := f.$$

Die Funktion f heißt eine (Radon-Nikodym-)Dichte von μ_f bezüglich μ. Das folgende Lemma zeigt, dass die Dichte bereits fast überall eindeutig festgelegt ist.

5.2 Lemma. *a) Falls $f, g \in \mathscr{L}^1(\mu)$ mit $f, g \geq 0$ und $\mu_f = \mu_g$, so folgt $f = g$ μ-fast überall.*

b) Falls $f\colon X \to [0, \infty]$ messbar ist, so gilt für alle messbaren $h\colon X \to [0, \infty]$

$$\int h d\mu_f = \int (hf) d\mu. \tag{5-1}$$

Falls $h\colon X \to \overline{\mathbb{R}}$ messbar ist, so gilt $h \in \mathscr{L}^1(\mu_f) \Leftrightarrow h \cdot f \in \mathscr{L}^1(\mu)$, und in diesem Fall gilt wieder (5-1).

c) Falls $f, h\colon X \to [0, \infty]$ messbar sind, so gilt $(\mu_f)_h = \mu_{f \cdot h}$.

Beweis. a) Sei $N := \{x \in X : f(x) > g(x)\}$ und $h := \chi_N f - \chi_N g$. Dann ist $h\colon X \to [0, \infty]$ messbar, und $\int h d\mu = \int_N f d\mu - \int_N g d\mu = \mu_f(N) - \mu_g(N) = 0$. Nach Bemerkung 2.3 (4) folgt $h = 0$ μ-fast überall, d.h. $\mu(N) = 0$. Analog folgt $\mu(\{x \in X : f(x) < g(x)\}) = 0$.

b) Das war Beispiel 4.4.

c) Für $A \in \mathscr{A}$ gilt $(\mu_f)_h(A) = \int_A h d\mu_f = \int \chi_A h d\mu_f = \int \chi_A h f d\mu = \int_A (fh) d\mu = \mu_{fh}(A)$, wobei b) verwendet wurde. □

Die nächste Aussage wird später im Beweis benötigt, ist aber auch für sich interessant.

5.3 Lemma. *Sei $(X, \mathscr{A}, \mu)$ ein Maßraum. Dann ist μ genau dann σ-endlich, falls ein $h \in \mathscr{L}^1(\mu)$ existiert mit $0 < h(x) < \infty$ $(x \in X)$.*

Beweis. (i) Sei μ σ-endlich, d.h. es existieren $(X_n)_{n\in\mathbb{N}} \subset \mathscr{A}$ mit $X_n \nearrow X$ $[:\Leftrightarrow X_n \subset X_{n+1}, \bigcup_{n\in\mathbb{N}} X_n = X]$ und $\mu(X_n) < \infty$. Wähle $(c_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ mit $0 < c_n \le 2^{-n}$ und $c_n\mu(X_n) \le 2^{-n}$ $(n \in \mathbb{N})$. Dann ist $h(x) := \sum_{n\in\mathbb{N}} c_n \chi_{X_n}$ messbar mit $0 < h \le 1$ und $\int h d\mu \le \sum_{n\in\mathbb{N}} c_n\mu(X_n) \le 1$.

(ii) Sei $h \in \mathscr{L}^1(\mu)$ mit $0 < h < \infty$. Dann ist $X = \{x \in X : h(x) > 0\} = \bigcup_{n\in\mathbb{N}} X_n$ mit $X_n := \{x \in X : h(x) \ge \frac{1}{n}\}$. Wegen $h \in \mathscr{L}^1(\mu)$ ist

$$\mu(X_n) = \int_{X_n} 1 d\mu \le \int_{X_n} nh(x)d\mu(x) \le n \int h d\mu < \infty.$$

□

5.4 Definition. Sei $\varphi\colon \mathscr{A} \to [0,\infty]$ ein Maß. Dann heißt φ absolutstetig bzgl. μ, falls für alle $A \in \mathscr{A}$ mit $\mu(A) = 0$ gilt: $\varphi(A) = 0$. Man schreibt $\varphi \ll \mu$.

5.5 Lemma. *Sei $\varphi\colon \mathscr{A} \to [0,\infty)$ ein endliches Maß. Dann gilt $\varphi \ll \mu$ genau dann, wenn*

$$\forall \varepsilon > 0\, \exists \delta > 0\, \forall A \in \mathscr{A},\, \mu(A) < \delta : \varphi(A) < \varepsilon. \tag{5-2}$$

Beweis. (i) Es gelte (5-2). Dann folgt für $A \in \mathscr{A}$ mit $\mu(A) = 0$ aus (5-2) $\varphi(A) < \varepsilon$ für jedes $\varepsilon > 0$ und damit $\varphi(A) = 0$, d.h. $\varphi \ll \mu$.

(ii) Die Rückrichtung wird hier nicht bewiesen. □

Das folgende Lemma betrachtet die Differenz zweier Maße. Diese Differenz ist selbst wieder σ-additiv, kann aber auch negative Werte annehmen und ist daher selbst kein Maß. Man spricht hier auch von einem signierten Maß. Die folgende Aussage besagt, dass man den Grundraum X zerlegen kann in $X = X_0 \cup X_0^c$, wobei die Differenz der Maße auf der Spur-σ-Algebra $\mathscr{A} \cap X_0$ nur Werte in $[0,\infty)$ annimmt und dort damit selbst ein Maß ist.

5.6 Lemma. *Seien $\varphi_1, \varphi_2\colon \mathscr{A} \to [0,\infty)$ zwei endliche Maße und sei $\tau := \varphi_1 - \varphi_2$. Dann existiert ein $X_0 \in \mathscr{A}$ mit $\tau(X_0) \ge \tau(X)$ und $\tau(A) \ge 0$ $(A \in \mathscr{A},\, A \subset X_0)$.*

Beweis. (i) Sei $\varepsilon > 0$. Wir zeigen zunächst:

$$\exists\, X_\varepsilon \in \mathscr{A} \text{ mit } \tau(X_\varepsilon) \ge \tau(X) \text{ und } \tau(A) > -\varepsilon\ (A \in \mathscr{A},\, A \subset X_\varepsilon). \tag{5-3}$$

Falls $\tau(X) \le 0$, wähle $X_\varepsilon := \emptyset$. Sei also $\tau(X) > 0$. Der Beweis verwendet die Spur-σ-Algebra $\mathscr{A} \cap E := \{A \in \mathscr{A} : A \subset E\}$ für Mengen $E \in \mathscr{A}$.

Falls $\tau(A) > -\varepsilon$ $(A \in \mathscr{A})$, wähle $X_\varepsilon := X$. Sonst existiert ein $A_1 \in \mathscr{A}$ mit $\tau(A_1) \le -\varepsilon$. Für das Komplement $A_1^c := X \setminus A_1$ gilt dann

$$\tau(A_1^c) = \tau(X) - \tau(A_1) \ge \tau(X) + \varepsilon > \tau(X).$$

Nun geht man zur Spur-σ-Algebra $\mathscr{A}_1 := \mathscr{A} \cap A_1^c$ über.

Falls $\tau(A) > -\varepsilon$ $(A \in \mathscr{A}_1)$, so wähle $X_\varepsilon := A_1^c$. Sonst existiert ein $A_2 \in \mathscr{A}_1$ mit $\tau(A_2) \le -\varepsilon$. Wegen $A_1 \cap A_2 = \emptyset$ folgt

$$\tau((A_1 \cup A_2)^c) = \tau(X) - \tau(A_1) - \tau(A_2) \ge \tau(X) + 2\varepsilon > \tau(X).$$

Setze nun $\mathscr{A}_2 := \mathscr{A} \cap (A_1 \dot\cup A_2)^c$.

Falls dieses Verfahren nach N Schritten abbricht, so wähle $X_\varepsilon := (A_1 \cup \dots \cup A_N)^c$. Es gilt dann $\tau(X_\varepsilon) > \tau(X)$ und $\tau(A) > -\varepsilon$ $(A \in \mathscr{A} \cap X_\varepsilon)$, d.h. (5-3).

Sonst existiert eine Folge $(A_n)_{n\in\mathbb{N}} \subset \mathscr{A}$ von paarweise disjunkten Mengen mit $\tau((A_1 \dot\cup \dots \dot\cup A_n)^c) > \tau(X)$ und $\tau(A_n) \le -\varepsilon$ $(n \in \mathbb{N})$. Für $A := \dot{\bigcup}_{n\in\mathbb{N}} A_n$ gilt dann $\tau(A) = \varphi_1(A) - \varphi_2(A) > -\infty$. Unter Verwendung der Stetigkeit von unten bzw. oben von φ_1 und φ_2 folgt aber auch

$$\tau(A) = \lim_{N\to\infty} \sum_{n=1}^N (\varphi_1(A_n) - \varphi_2(A_n)) = \lim_{N\to\infty} \sum_{n=1}^N \tau(A_n) \le \lim_{N\to\infty} (-N\varepsilon) = -\infty,$$

Widerspruch.

(ii) Setze in (i) $\varepsilon := \frac{1}{n}$ und erhalte Mengen $X_{1/n}$ für $n \in \mathbb{N}$. Dabei kann man $X_{1/(n+1)} \subset X_{1/n}$ annehmen (sonst ersetze in (5-3) die Menge X durch $X_{1/n}$). Sei $X_0 := \bigcap_{n\in\mathbb{N}} X_{1/n}$. Da τ stetig von oben ist, folgt $\tau(X_0) \ge \tau(X)$, und für jedes $A \in \mathscr{A} \cap X_0$ ist $\tau(A) > -\frac{1}{n}$ $(n \in \mathbb{N})$ und damit $\tau(A) \ge 0$. □

5.7 Satz (von Radon-Nikodym)**.** *Sei $(X, \mathscr{A}, \mu)$ ein σ-endlicher Maßraum und sei $\varphi\colon \mathscr{A} \to [0,\infty]$ ein Maß mit $\varphi \ll \mu$. Dann besitzt φ eine Dichte bzgl. μ, d.h. es existiert eine messbare Funktion $f\colon X \to [0,\infty]$ mit*

$$\varphi(A) = \int_A f d\mu \quad (A \in \mathscr{A}).$$

Die Dichte f ist μ-fast überall eindeutig bestimmt.

Beweis. Wir beweisen den Satz nur für den Fall, dass μ und φ beide endlich sind; der allgemeine Fall folgt dann unter Verwendung geeignet konstruierter Zerlegungen und mit Lemma 5.3. Definiere

$$U := \Big\{ g\colon X \to [0,\infty] \,|\, g \text{ messbar}, \; \forall\, A \in \mathscr{A} : \int_A g d\mu \le \varphi(A) \Big\}.$$

Wegen $0 \in U$ ist U nicht leer. Falls $g_1, g_2 \in U$, so ist auch $g := \max\{g_1, g_2\} \in U$, denn mit $C := \{x \in X : g_1(x) \ge g_2(x)\}$ gilt

$$\int_A g d\mu = \int_{A\cap C} g_1 d\mu + \int_{A\setminus C} g_2 d\mu \le \varphi(A \cap C) + \varphi(A \setminus C) = \varphi(A) \quad (A \in \mathscr{A}).$$

Sei $\alpha := \sup\{\int g d\mu : g \in U\}$. Wegen $\int g d\mu \le \varphi(X)$ ist $\alpha \le \varphi(X) < \infty$. Wähle eine Folge $(\widetilde{g}_n)_{n\in\mathbb{N}} \subset U$ mit $\int \widetilde{g}_n d\mu \nearrow \alpha$. Für $g_n := \max\{\widetilde{g}_1, \dots, \widetilde{g}_n\}$ gilt dann nach Obigem $g_n \in U$ und wegen $g_n \ge \widetilde{g}_n$ auch $\int g_n d\mu \nearrow \alpha$.

Als monoton wachsende Folge besitzt $(g_n)_{n\in\mathbb{N}}$ einen messbaren Grenzwert $f := \lim_{n\to\infty} g_n \colon X \to [0,\infty]$. Nach dem Satz über monotone Konvergenz gilt $\int f d\mu = \lim_{n\to\infty} \int g_n d\mu = \alpha$, und

$$\int_A f d\mu = \lim_{n\to\infty} \int_A g_n d\mu \le \varphi(A) \quad (A \in \mathscr{A}),$$

d.h. es gilt $f \in U$.

Wir zeigen nun, dass f eine Dichte von φ bzgl. μ ist. Da $f \in U$, definiert

$$\varphi_1(A) := \varphi(A) - \int_A f d\mu \quad (A \in \mathscr{A})$$

ein endliches Maß mit $\varphi_1 \ll \mu$. Zu zeigen ist $\varphi_1 = 0$, d.h. $\varphi_1(X) = 0$.

Angenommen, $\varphi_1(X) > 0$. Da $\varphi_1 \ll \mu$, ist auch $\mu(X) > 0$. Für $s := \frac{\varphi_1(X)}{2\mu(X)} > 0$ folgt $\varphi_1(X) > s\mu(X)$. Wir wenden Lemma 5.6 auf φ_1 und $\varphi_2 := s\mu$ an und erhalten ein $X_0 \in \mathscr{A}$ mit

$$\varphi_1(X_0) - s\mu(X_0) \ge \varphi_1(X) - s\mu(X) > 0$$

und $\varphi_1(A) \ge s\mu(A)$ $(A \in \mathscr{A} \cap X_0)$. Setze $f_0 := f + s\chi_{X_0}$. Dann gilt

$$\begin{aligned} \int_A f_0 d\mu &= \int_A f d\mu + s\mu(A \cap X_0) \le \int_A f d\mu + \varphi_1(A \cap X_0) \\ &\le \int_A f d\mu + \varphi_1(A) = \varphi(A) \quad (A \in \mathscr{A}). \end{aligned}$$

Daher ist $f_0 \in U$. Wegen $\varphi_1 \ll \mu$ und $\varphi_1(X_0) > s\mu(X_0)$ ist $\mu(X_0) > 0$. Damit

$$\int f_0 d\mu = \int f d\mu + s\mu(X_0) = \alpha + s\mu(X_0) > \alpha,$$

Widerspruch zur Definition von α. Somit ist f eine Dichte von φ bzgl. μ.

Seien nun $f, g\colon X \to [0,\infty]$ mit $\int_A f d\mu = \int_A g d\mu = \varphi(A)$ $(A \in \mathscr{A})$. Wegen $\varphi(X) < \infty$ gilt $f, g \in \mathscr{L}^1(\mu)$. Nach Lemma 5.2 a) folgt $f = g$ μ-fast überall. □

Eine wichtige Anwendung des Satzes von Radon-Nikodym ist der Satz vom bedingten Erwartungswert, den wir in der Sprache der Stochastik formulieren.

5.8 Satz (vom bedingten Erwartungswert)**.** *Sei $(\Omega, \mathscr{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und sei $X \in \mathscr{L}^1(\mathbb{P})$. Sei $\mathscr{A}_0 \subset \mathscr{A}$ eine Unter-σ-Algebra. Dann existiert ein $\mathscr{A}_0$-messbares $X_0 \in \mathscr{L}^1(\mathbb{P})$ mit*

$$\int_A X_0 d\mathbb{P} = \int_A X d\mathbb{P} \quad (A \in \mathscr{A}_0).$$

Die Funktion X_0 ist $\mathbb{P}$-fast sicher eindeutig bestimmt und heißt bedingter Erwartungswert von X unter der Bedingung $\mathscr{A}_0$, in Zeichen $X_0 =: E(X|\mathscr{A}_0)$.

Beweis. Sei zunächst $X \geq 0$. Setze $\mu(A) := \mathbb{P}|_{\mathscr{A}_0}$ und $\varphi(A) := \int_A X d\mathbb{P}$ $(A \in \mathscr{A}_0)$. Dann sind μ und φ endliche Maße auf $\mathscr{A}_0$ mit $\varphi \ll \mu$. Nach dem Satz von Radon-Nikodym besitzt φ eine ($\mathbb{P}$-fast überall eindeutige) Dichte bzgl. μ, d.h. es existiert eine $\mathscr{A}_0$-messbare Funktion $X_0 \geq 0$ mit

$$\int_A X_0 d\mathbb{P} = \varphi(A) = \int_A X d\mathbb{P} \quad (A \in \mathscr{A}_0).$$

Für allgemeines $X \in \mathscr{L}^1(\mathbb{P})$ zerlegt man $X = X_+ - X_-$ und erhält zwei Dichten $X_{0,+}$ und $X_{0,-}$. Setze dann $X_0 := X_{0,+} - X_{0,-}$. □

5.9 Bemerkung. a) Der Satz gilt analog für messbare $X \colon \Omega \to [0, \infty]$.

b) Speziell für die triviale σ-Algebra $\mathscr{A}_0 = \{\emptyset, \Omega\}$ erhält man eine konstante Funktion $X_0 = c$ mit Wert $c = c\mathbb{P}(\Omega) = \int_\Omega X_0 d\mathbb{P} = \int_\Omega X d\mathbb{P}$. Der Wert $E(X) := \int_\Omega X d\mathbb{P}$ heißt der Erwartungswert von X.

c) Sei in der Situation des Satzes $A \in \mathscr{A}$. Dann heißt $\mathbb{P}(A|\mathscr{A}_0) := E(\chi_A|\mathscr{A}_0)$ die bedingte Wahrscheinlichkeit von A unter der Bedingung $\mathscr{A}_0$. Falls $\mathscr{A}_0 = \sigma(Y)$ für ein $Y \in \mathscr{L}^1(\mathbb{P})$, so schreibt man auch $E(X|Y)$ bzw. $\mathbb{P}(X|Y)$.

d) Falls sogar $X \in \mathscr{L}^2(\mathbb{P})$ gilt (d.h. falls $|X|^2 \in \mathscr{L}^1(\mathbb{P})$), dann kann man zeigen, dass X_0 das Element im Unterraum $\mathscr{L}^2(\mathbb{P}|_{\mathcal{A}_0})$ ist, welches den Abstand $\|X - X_0\|_2$ minimiert. Dies erlaubt auch einen alternativen Beweis des Satzes.

6. Die L^p-Räume

6.1 Worum geht's? Der Raum $\mathscr{L}^1(\mu)$ oder allgemeiner die Räume $\mathscr{L}^p(\mu)$ für $1 \le p \le \infty$ können mit einer Halbmetrik versehen werden. Durch eine geeignete Äquivalenzklassenbildung erhält man einen metrischen Raum $L^p(\mu)$, der sogar - ohne Zusatzbedingung an das Maß - ein Banachraum ist. Am interessantesten ist hierbei wieder das Lebesgue-Maß, d.h. die Räume $L^p(U)$ mit $U \subset \mathbb{R}^n$. In diesen Räumen liegen für $p < \infty$ die stetigen Funktionen und sogar die Testfunktionen dicht, wie man mit Hilfe der Faltung zeigen kann.

a) Definition und erste Eigenschaften

Im folgenden sei $(X, \mathscr{A}, \mu)$ ein Maßraum.

6.2 Definition. Eine Funktion $f\colon X \to \mathbb{C}$ heißt integrierbar, falls $\operatorname{Re} f \in \mathscr{L}^1(\mu)$ und $\operatorname{Im} f \in \mathscr{L}^1(\mu)$ gilt. Wir schreiben $f \in \mathscr{L}^1(\mu; \mathbb{C})$ oder $f \in \mathscr{L}^1(X; \mathbb{C})$ bzw. umgekehrt auch $f \in \mathscr{L}^1(\mu; \mathbb{R})$ für reellwertige Funktionen. Für $f \in \mathscr{L}^1(\mu; \mathbb{C})$ wird $\int f d\mu$ definiert durch $\int f d\mu := \int \operatorname{Re} f d\mu + i \int \operatorname{Im} f d\mu$.

6.3 Bemerkung. a) Die Messbarkeit einer Funktion $f\colon X \to \mathbb{C}$ ist wie üblich als Borel-Messbarkeit definiert (also $\mathscr{A}$-$\mathscr{B}(\mathbb{C})$-Messbarkeit). Insbesondere folgt aus der Stetigkeit der Abbildungen $z \mapsto \operatorname{Re} z$, $(z_1, z_2) \mapsto z_1 + z_2$, $(z_1, z_2) \mapsto z_1 \cdot z_2$ und $z \mapsto |z|$:

Eine Funktion $f\colon X \to \mathbb{C}$ ist genau dann messbar, wenn $\operatorname{Re} f, \operatorname{Im} f\colon X \to \mathbb{R}$ beide messbar sind. Sind $f, g\colon X \to \mathbb{C}$ messbar, so auch $f \pm g$, $f \cdot g$ und $|f|$.

b) Sei $f\colon X \to \mathbb{C}$ messbar. Dann gilt $f \in \mathscr{L}^1(\mu; \mathbb{C})$ genau dann, wenn $|f| \in \mathscr{L}^1(\mu; \mathbb{R})$ gilt. Das folgt sofort aus dem Majorantenkriterium (Bemerkung 2.4) und den Ungleichungen $|\operatorname{Re} f| \le |f|$, $|\operatorname{Im} f| \le |f|$ und $|f| \le |\operatorname{Re} f| + |\operatorname{Im} f|$.

c) Die obigen Sätze der Integrationstheorie (mit Ausnahme der Sätze, bei denen Monotonie eine Rolle spielt) gelten analog für komplexwertige Funktionen.

6.4 Definition ($\mathscr{L}^p$-Räume)**.** a) Sei $1 \le p < \infty$. Definiere $\mathscr{L}^p(\mu; \mathbb{C})$ als die Menge aller messbaren Funktionen $f\colon X \to \mathbb{C}$ mit

$$\|f\|_p := \Big(\int |f|^p d\mu \Big)^{1/p} < \infty.$$

b) Für $p = \infty$ wird $\mathscr{L}^\infty(\mu; \mathbb{C})$ definiert als die Menge aller messbarer Funktionen $f\colon X \to \mathbb{C}$, für welche es ein $C_f > 0$ gibt mit $\mu(\{x \in X : |f(x)| > C_f\}) = 0$. Man spricht von μ-fast überall beschränkten Funktionen. Für $f \in \mathscr{L}^\infty(\mu;\mathbb{C})$ definiert man

$$\|f\|_\infty := \inf\Big\{C \in \mathbb{R} : \mu(\{x \in X : |f(x)| > C\}) = 0\Big\}.$$

c) Für Funktionen $f\colon X \to \overline{\mathbb{R}}$ werden die entsprechenden Funktionenräume mit $\mathscr{L}^p(\mu;\mathbb{R})$ bezeichnet.

Im Folgenden schreiben wir kurz $\mathscr{L}^p(\mu)$ für $\mathscr{L}^p(\mu;\mathbb{K})$, wobei $\mathbb{K} \in \{\mathbb{R},\mathbb{C}\}$.

6.5 Satz. *a)* (Höldersche Ungleichung) *Für* $1 \le p,q \le \infty$ *mit* $\frac{1}{p} + \frac{1}{q} = 1$ *gilt*

$$\|f \cdot g\|_1 \le \|f\|_p \cdot \|g\|_q \quad (f \in \mathscr{L}^p(\mu),\ g \in \mathscr{L}^q(\mu)).$$

b) (Minkowskische Ungleichung) *Für* $1 \le p \le \infty$ *gilt*

$$\|f+g\|_p \le \|f\|_p + \|g\|_p \quad (f,g \in \mathscr{L}^p(\mu)).$$

c) Für $1 \le p \le \infty$ ist der Raum $\mathscr{L}^p(\mu)$ ein Vektorraum, und $\|\cdot\|_p$ definiert eine Seminorm (Halbnorm) auf $\mathscr{L}^p(\mu)$, d.h. es gilt

$$\begin{aligned}
\|f\|_p &\ge 0 && (f \in \mathscr{L}^p(\mu)),\\
\|\alpha f\|_p &= |\alpha| \cdot \|f\|_p && (\alpha \in \mathbb{K},\ f \in \mathscr{L}^p(\mu)),\\
\|f+g\|_p &\le \|f\|_p + \|g\|_p && (f,g \in \mathscr{L}^p(\mu)).
\end{aligned}$$

Beweis. a),b): Der Beweis der Hölderschen Ungleichung aus der Youngschen Ungleichung sowie der Minkowski-Ungleichung aus der Hölder-Ungleichung ist bereits aus Analysis II bekannt.

c) Die Messbarkeit von αf und $f+g$ ist klar, die Homogenität von $\|\cdot\|_p$ folgt für $1 \le p < \infty$ aus der Linearität des Integrals und ist offensichtlich für $p = \infty$. Die Dreiecksungleichung ist gerade die Minkowski-Ungleichung, welche auch zeigt, dass $f + g \in \mathscr{L}^p(\mu)$ gilt. □

6.6 Definition und Satz (L^p-Räume)**.** *Sei $1 \le p \le \infty$. Definiere auf $\mathscr{L}^p(\mu)$ die Äquivalenzrelation $\sim$ durch*

$$f \sim g \quad :\Longleftrightarrow \quad \mu\big(\{x \in X : f(x) \ne g(x)\}\big) = 0$$

(d.h. durch Gleichheit μ-fast überall). Die Menge der Äquivalenzklassen $\{[f] : f \in \mathscr{L}^p(\mu)\}$ wird mit $L^p(\mu)$ bezeichnet (auch: $L_p(\mu)$).

Auf $L^p(\mu)$ wird repräsentantenweise eine Vektorraumstruktur definiert: $\alpha[f] := [\alpha f]$ und $[f] + [g] := [f+g]$ für $\alpha \in \mathbb{K}$ und $f,g \in \mathscr{L}^p(\mu)$. Analog definiert man

$$\|[f]\|_p := \|f\|_p \quad (f \in \mathscr{L}^p(\mu)).$$

Damit wird $(L^p(\mu), \|\cdot\|_p)$ zu einem normierten Vektorraum.

Beweis. Die Wohldefiniertheit der Skalarmultiplikation und der Addition in $L^p(\mu)$ ist klar, ebenso übertragen sich die Eigenschaften einer Seminorm von $\mathscr{L}^p(\mu)$ auf $L^p(\mu)$. Sei nun $[f] \in L^p(\mu)$ mit $\|[f]\|_p = 0$. Dann gilt nach Definition $\int |f|^p d\mu = 0$ und damit nach Bemerkung 2.3 $f = 0$ μ-fast überall, d.h. $[f] = 0$ in $L^p(\mu)$. Somit ist $\|\cdot\|_p$ eine Norm auf $L^p(\mu)$. □

6.7 Bemerkung. Im Folgenden wird für $f \in \mathscr{L}^p(\mu)$ in der Schreibweise nicht mehr zwischen $[f]$ und f unterschieden. Wichtig ist bei dieser Betrachtung, dass es für „Funktionen" $f \in L^p(\mu)$ im Allgemeinen keinen Sinn macht, vom Wert $f(x)$ an einer Stelle $x \in X$ zu sprechen. Denn falls $\mu(\{x\}) = 0$, so kann man den Repräsentanten f an der Stelle x ändern, ohne die Äquivalenzklasse zu ändern.

6.8 Satz. *Sei $(X, \mathscr{A}, \mu)$ ein Maßraum und $1 \le p \le \infty$. Dann ist $(L^p(\mu), \|\cdot\|_p)$ ein Banachraum.*

Beweis. Nach Satz 6.6 ist $L^p(\mu)$ ein normierter Raum, es ist also nur noch die Vollständigkeit zu zeigen.

(i) Sei $1 \le p < \infty$ und $(f_k)_{k\in\mathbb{N}} \subset L^p(\mu)$ eine Cauchyfolge. Wähle eine Teilfolge $(f_{k_j})_{j\in\mathbb{N}}$ mit $\|f_{k_{j+1}} - f_{k_j}\|_p \le 2^{-j}$ („Turbo-Teilfolge") und definiere

$$g_\ell := \sum_{j=1}^{\ell} |f_{k_{j+1}} - f_{k_j}|, \quad g := \sum_{j=1}^{\infty} |f_{k_{j+1}} - f_{k_j}|.$$

Aus der Minkowskischen Ungleichung folgt $\|g_\ell\|_p \le \sum_{j=1}^{\ell} 2^{-j} \le 1$, und nach dem Satz von der monotonen Konvergenz gilt

$$\|g\|_p^p = \int |g|^p d\mu = \int \lim_{\ell\to\infty} |g_\ell|^p d\mu = \lim_{\ell\to\infty} \int |g_\ell|^p d\mu \le 1.$$

Nach Bemerkung 2.3 ist g μ-fast überall endlich und damit ist die Reihe

$$f(x) := f_{k_1}(x) + \sum_{j=1}^{\infty} (f_{k_{j+1}(x)} - f_{k_j}(x))$$

für μ-fast alle $x \in X$ absolut konvergent. Somit gilt $f = \lim_{j\to\infty} f_{k_j}$ μ-fast überall. Setze noch $f = 0$, falls die obige Reihe nicht absolut konvergent ist, so ist f messbar als punktweiser Grenzwert messbarer Funktionen.

Wir zeigen $\|f - f_k\|_p \to 0 \quad (k \to \infty)$. Zu $\varepsilon > 0$ wähle $N \in \mathbb{N}$ mit $\|f_k - f_\ell\|_p < \varepsilon$ für alle $k, \ell \ge N$. Nach dem Lemma von Fatou ist für $k \ge N$

$$\int |f - f_k|^p d\mu = \int \lim_{j\to\infty} |f_{k_j} - f_k|^p d\mu \le \liminf_{j\to\infty} \int |f_{k_j} - f_k|^p d\mu \le \varepsilon^p.$$

Insbesondere ist $f - f_k \in L^p(\mu)$ und damit $f = f_k + (f - f_k) \in L^p(\mu)$, und $f_k \to f$ in $L^p(\mu)$.

(ii) Sei nun $p = \infty$. Definiere $N_k := \{|f_k| > \|f_k\|_\infty\}$ und $N_{k,\ell} := \{|f_k - f_\ell| > \|f_k - f_\ell\|_\infty\}$. Dann sind all diese Mengen Nullmengen, also auch ihre Vereinigung N. Außerhalb von N ist die Folge $(f_k)_{k\in\mathbb{N}}$ eine Cauchyfolge bzgl. Supremumsnorm von beschränkten Funktionen. Also konvergiert diese Folge in $X \setminus N$ gleichmäßig gegen eine beschränkte Funktion f. Setzt man noch $f = 0$ auf N, so ist $f \in L^\infty(\mu)$ und $f_k \to f$ in $L^\infty(\mu)$. □

6.9 Korollar. *Sei $1 \le p \le \infty$ und $(f_n)_{n\in\mathbb{N}} \subset L^p(\mu)$ eine Folge mit $f_n \to f \in L^p(\mu) \quad (n \to \infty)$. Dann besitzt $(f_n)_n$ eine Teilfolge, welche μ-fast überall gegen f konvergiert.*

Beweis. Das wurde im Beweis von Satz 6.8 mit gezeigt. □

6.10 Korollar. *Durch*

$$\langle f, g\rangle_2 := \int f\, \overline{g}\, d\mu \quad (f, g \in L^2(\mu))$$

ein Skalarprodukt auf $L^2(\mu)$ erklärt. Der Raum $(L^2(\mu), \langle\cdot,\cdot\rangle_2)$ ist ein Hilbertraum.

Beweis. Die Wohldefiniertheit des Skalarprodukts folgt aus der Hölderschen Ungleichung, die Eigenschaften eines Skalarprodukts sind klar, und die Vollständigkeit wurde in Satz 6.8 gezeigt. □

In gewisser Weise ist $L^2(\mu)$ „der“ Hilbertraum schlechthin. Beispiele sind $\mathbb{K}^n$ als $L^2(\zeta_n)$, wobei ζ_n das Zählmaß auf der Menge $\{1, \dots, n\}$ bezeichnet, ℓ^2 als $L^2(\zeta_\mathbb{N})$ mit dem Zählmaß $\zeta_\mathbb{N}$ auf der Menge $\mathbb{N}$, sowie $L^2(U)$ für $U \subset \mathbb{R}^n$ offen.

b) Die Faltung

Im folgenden beziehen sich die Begriffe wie Messbarkeit und Integrierbarkeit stets auf das Lebesgue-Maß. Dabei heißt messbar zunächst Borel-messbar, falls nötig, verwenden wir auch die Vervollständigung des Lebesgue-Maßes, also Lebesgue-messbare Mengen.

Sei $U \subset \mathbb{R}^n$ offen. Der Träger einer Funktion $f\colon U \to \mathbb{K}$ ist definiert als $\operatorname{supp} f := \overline{\{x \in U : f(x) \neq 0\}}$. Weiter sei

$$C_b(U) := \{f\colon U \to \mathbb{K} \mid f \text{ stetig und beschränkt}\},$$

$$C_c^k(U) := \{f\colon U \to \mathbb{K} \,|\, f \text{ } k\text{-mal stetig differenzierbar, } \operatorname{supp} f \text{ kompakt}\}$$

für $k \in \mathbb{N} \cup \{\infty\}$. Für $k = \infty$ schreibt man auch $\mathscr{D}(U) := C_c^\infty(U)$ und spricht vom Raum der Raum der Testfunktionen auf U.

6.11 Definition. Seien $f, g\colon \mathbb{R}^n \to \mathbb{K}$ messbar. Definiere

$$N_{f,g} := \Big\{x \in \mathbb{R}^n : \int |f(y)| \cdot |g(x-y)| dy = \infty\Big\}$$

und das Faltungsprodukt $f * g\colon \mathbb{R}^n \to \mathbb{K}$ durch

$$(f*g)(x) := \begin{cases} \int f(y)g(x-y)dy, & x \notin N_{f,g}, \\ 0, & x \in N_{f,g}. \end{cases}$$

6.12 Bemerkung. Definiere die Abbildung $d\colon \mathbb{R}^2 \to \mathbb{R}$ durch $d(x,y) = x-y$. Dann ist d stetig und damit messbar. Also ist die Abbildung $(x,y) \mapsto f(y)g(x-y) = (f \cdot (g \circ d))(x,y)$ ebenfalls messbar. Nach Lemma 3.7 ii) ist $N_{f,g} \in \mathscr{B}(\mathbb{R}^n)$ und $f * g$ messbar.

6.13 Lemma. *Seien $f, g\colon \mathbb{R}^n \to \mathbb{K}$ messbar.*

a) Es ist $f * g = g * f$, d.h. die Faltung ist kommutativ.

b) Es gilt $\{f * g \neq 0\} \subset \{f \neq 0\} + \{g \neq 0\} = \{y_1 + y_2 : f(y_1) \neq 0,\ g(y_2) \neq 0\}$.

c) Falls $f, g \in L^1(\mathbb{R}^n)$, so ist $\lambda(N_{f,g}) = 0$ und $f * g \in L^1(\mathbb{R}^n)$ mit

$$\|f * g\|_1 \leq \|f\|_1 \cdot \|g\|_1.$$

d) Sei $1 \leq p, q \leq \infty$ mit $\frac{1}{p} + \frac{1}{q} = 1$. Für $f \in L^p(\mathbb{R}^n)$ und $g \in L^q(\mathbb{R}^n)$ ist dann $\lambda(N_{f,g}) = 0$ und $f * g \in L^\infty(\mathbb{R}^n)$ mit

$$\|f * g\|_\infty \leq \|f\|_p \cdot \|g\|_q.$$

e) Falls $f, g, h \in L^1(\mathbb{R}^n)$, so ist $(f*g)*h = f*(g*h)$, d.h. die Faltung ist assoziativ.

Beweis. a) folgt sofort aus dem Transformationssatz mit der Transformation $y \mapsto x - y$, einmal angewendet auf den Integranden $|f(\cdot)g(x-\cdot)|$ und einmal auf den Integranden $f(\cdot)g(x-\cdot)$.

b) Falls $x \notin \{f \neq 0\} + \{g \neq 0\}$ und $y \in \mathbb{R}^n$ beliebig, so ist $f(y)g(x-y) = 0$ wegen $x = y + (x-y)$. Also ist $(f*g)(x) = 0$.

c) Nach dem Satz von Tonelli und mit der Translationsinvarianz des Lebesgue-Integrals gilt

$$\int \Big(\int |f(y)| \cdot |g(x-y)| dy \Big) dx = \int |f(y)| dy \int |g(x-y)| dx = \|f\|_1 \cdot \|g\|_1.$$

Nach Bemerkung 2.3 a) folgt $\lambda(N_{f,g}) = 0$ und

$$\begin{aligned} \|f * g\|_1 &= \int \Big| \int f(y) g(x-y) dy \Big| dx \\ &\le \int \Big(\int |f(y)| \cdot |g(x-y)| dy \Big) dx = \|f\|_1 \cdot \|g\|_1. \end{aligned}$$

d) Nach der Hölderschen Ungleichung gilt für jedes $x \in \mathbb{R}^n$

$$|(f*g)(x)| \le \int |f(y)| \cdot |g(x-y)| dy = \big\| f(\cdot) g(x - \cdot) \big\|_1 \le \|f\|_p \cdot \|g\|_q,$$

wobei wieder die Translationsinvarianz des Lebesgue-Maßes verwendet wurde.

e) folgt mit Fubini. □

6.14 Definition. Sei $U \subset \mathbb{R}^n$ messbar. Eine Funktion $f \colon U \to \mathbb{K}$ heißt lokal integrierbar in U, falls f messbar ist und $f \cdot \chi_K \in L^1(U)$ für alle kompakten Teilmengen $K \subset U$ gilt. Man schreibt auch $f \in L^1_{\text{loc}}(U)$.

6.15 Bemerkung. a) Sei $(X, \mathscr{A}, \mu)$ ein endlicher Maßraum. Dann gilt $L^q(\mu) \subset L^p(\mu)$ für $1 \le p \le q \le \infty$, da für alle $f \in L^q(\mu)$ gilt

$$\begin{aligned} \|f\|_p^p &= \int |f|^p d\mu = \int_{\{|f|>1\}} |f|^p d\mu + \int_{\{|f|\le 1\}} |f|^p d\mu \\ &\le \int_{\{|f|>1\}} |f|^q d\mu + \mu(X) \le \|f\|_q^q + \mu(X) < \infty. \end{aligned}$$

b) Insbesondere folgt $L^p(U) \subset L^1_{\text{loc}}(U)$ für alle $U \subset \mathbb{R}^n$ messbar und $1 \le p \le \infty$.

6.16 Lemma. *a) Sei $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ und $g \in L^1(\mathbb{R}^n)$ mit $\operatorname{supp} g$ kompakt. Dann ist auch $f * g \in L^1_{\text{loc}}(\mathbb{R}^n)$.*

b) Sei $f \in L^1(\mathbb{R}^n)$ und $g \in C_b(\mathbb{R}^n)$, dann ist $f * g$ stetig. Das Gleiche gilt, falls $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ und $g \in C_c(\mathbb{R}^n)$.

Beweis. a) Sei $K \subset \mathbb{R}^n$ kompakt und $L := K - \operatorname{supp} g$. Als stetiges Bild der kompakten Menge $K \times \operatorname{supp} g$ unter der Abbildung $(x, y) \mapsto x - y$ ist L wieder kompakt.

Für $x \in K$ und $y \in \mathbb{R}^n$ ist $f(x-y)g(y) = (f \cdot \chi_L)(x-y)g(y)$. Wegen $f \cdot \chi_L \in L^1(\mathbb{R}^n)$ ist

$$\chi_K \cdot (f * g) = \chi_K \cdot [(f \cdot \chi_L) * g] \in L^1(\mathbb{R}^n)$$

nach Lemma 6.13 c).

b) folgt aus dem Satz über die Parameterabhängigkeit des Integrals 2.12 a). $\square$

6.17 Satz. *Sei $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ und $g \in C_c^k(\mathbb{R}^n)$. Dann ist $f * g \in C^k(\mathbb{R}^n)$, und für die partiellen Ableitungen $\partial^\alpha := (\frac{\partial}{\partial x_1})^{\alpha_1} \dots (\frac{\partial}{\partial x_n})^{\alpha_n}$ gilt*

$$\partial^\alpha (f * g) = f * (\partial^\alpha g) \quad (|\alpha| \le k).$$

Insbesondere ist $f * g \in C^\infty(\mathbb{R}^n)$, falls $g \in \mathscr{D}(\mathbb{R}^n)$.

Beweis. Nach Lemma 6.16 ist $f * g \in L^1_{\text{loc}}(\mathbb{R}^n)$. Da die Differenzierbarkeit eine lokale Aussage ist, kann man $f \in L^1(\mathbb{R}^n)$ annehmen. Dann folgt die Aussage aus Satz 2.12 b) über Parameterabhängigkeit der Integrale, da $\partial^\alpha g$ als stetige Funktion mit kompaktem Träger beschränkt ist. $\square$

c) Vollständigkeits- und Dichtheitsaussagen

6.18 Satz (von Lusin)**.** *Sei $U \subset \mathbb{R}^n$ offen mit $\lambda(U) < \infty$ und $f \colon U \to \mathbb{K}$ messbar. Zu $\varepsilon > 0$ existiert eine kompakte Menge $K \subset U$ mit $\lambda(U \setminus K) < \varepsilon$, so dass $f|_K$ stetig ist.*

Beweis. O.E. sei $\mathbb{K} = \mathbb{R}$. Wähle zu festem $i \in \mathbb{N}$ Mengen $B_{ij} \subset \mathbb{R}$ $(j \in \mathbb{N})$ mit $(B_{ij})_j$ disjunkt, B_{ij} messbar, $\dot{\bigcup}_{j\in\mathbb{N}} B_{ij} = \mathbb{R}$ und

$$\operatorname{diam}(B_{ij}) := \sup_{s,t \in B_{ij}} |s-t| < \frac{1}{i}\,.$$

Setze

$$A_{ij} := f^{-1}(B_{ij}) \in \mathcal{B}(\mathbb{R}^n).$$

Dann ist $\dot{\bigcup}_{j\in\mathbb{N}} A_{ij} = U$. Wir verwenden die Regularität des Lebesgue-Maßes (Satz 1.22):

$$\lambda(A) = \sup\{\lambda(K) : K \subset A,\ K \text{ kompakt}\} \quad (A \in \mathscr{B}(\mathbb{R}^n)).$$

Somit können wir kompakte Mengen $K_{ij} \subset A_{ij}$ wählen mit $\lambda(A_{ij} \setminus K_{ij}) < \varepsilon \cdot 2^{-(i+j)}$. Dann ist

$$\lambda\Big(U \setminus \bigcup_{j\in\mathbb{N}} K_{ij}\Big) = \lambda\Big(\dot{\bigcup_{j\in\mathbb{N}}} (A_{ij} \setminus K_{ij})\Big) < \frac{\varepsilon}{2^i}\,.$$

Somit existiert ein $N(i) \in \mathbb{N}$ mit

$$\lambda\Big(U \setminus \dot{\bigcup_{j=1}^{N(i)}} K_{ij}\Big) < \frac{\varepsilon}{2^i}\,.$$

Definiere $D_i := \dot{\bigcup\limits_{j=1}^{N(i)}} K_{ij}$. Dann ist D_i kompakt.

Wähle nun für alle i, j ein $b_{ij} \in B_{ij}$ und definiere

$$g_i \colon D_i \to \mathbb{R},\; g_i(x) := b_{ij} \text{ für } x \in K_{ij} \quad (j = 1, \dots, N(i)).$$

Da die Mengen K_{ij} disjunkt und kompakt sind, haben sie einen positiven Abstand, d.h. g_i ist stetig.

Es gilt

$$|f(x) - g_i(x)| \le \frac{1}{i} \quad (x \in D_i). \tag{6-1}$$

Setzt man schließlich $K := \bigcap_{i=1}^{\infty} D_i$, so ist K kompakt, und es gilt

$$\lambda(U \setminus K) \le \sum_{i=1}^{\infty} \lambda(U \setminus D_i) < \varepsilon.$$

Andererseits konvergiert wegen (6-1) die Folge $(g_i)_i$ gleichmäßig auf K gegen f. Damit ist $f|_K$ stetig. $\square$

Wir werden im Folgenden den Fortsetzungssatz von Tietze verwenden, den wir hier nicht beweisen wollen.

6.19 Satz (Fortsetzungssatz von Tietze)**.** *Seien X ein metrischer Raum und $M \subset X$ abgeschlossen, $a, b \in \mathbb{R}$ mit $a < b$ und $f \colon M \to [a, b]$ stetig. Dann existiert eine stetige Fortsetzung $F \colon X \to [a, b]$ von f.*

6.20 Satz. *Sei $U \subset \mathbb{R}^n$ offen, $1 \le p < \infty$. Dann ist $C_c(U)$ dicht in $L^p(U)$.*

Beweis. O.E. sei $f \in L^p(U; \mathbb{R})$ mit $f \ge 0$ (sonst verwende man eine Zerlegung). Sei $\varepsilon > 0$.

Wegen

$$U = \bigcup_{k \in \mathbb{N}} U_k \quad \text{ mit } U_k := \Big\{x \in U : |x| < k,\, \operatorname{dist}(x, \partial U) > \frac{1}{k}\Big\}$$

gilt $\int_U |f - f \cdot \chi_{U_k}|^p d\lambda \to 0 \quad (k \to \infty)$. Wähle ein $k \in \mathbb{N}$ mit $\|f - f_k\|_p < \frac{\varepsilon}{4}$ für $f_k := f \cdot \chi_{U_k}$.

Zu f_k existiert eine Folge $(s_n)_{n\in\mathbb{N}}$ von Stufenfunktionen mit $s_n \nearrow f_k$ punktweise. Wegen $0 \le s_n \le f_k$ ist $s_n \in L^p(U)$. Wegen $(f_k - s_n)^p \le f_k^p$ folgt mit majorisierter Konvergenz $s_n \to f_k$ in $L^p(U)$. Wähle ein $n_0 \in \mathbb{N}$ mit

$$\|f_k - s_{n_0}\|_p < \frac{\varepsilon}{4}\,.$$

Da $\lambda(U_k) < \infty$, existiert nach dem Satz von Lusin 6.18 eine kompakte Menge $K \subset U_k$ mit $s_{n_0}|_K$ stetig und

$$\lambda(U_k \setminus K) < \Big(\frac{\varepsilon}{4\|s_{n_0}\|_\infty}\Big)^p.$$

Setze nun

$$\widetilde{\varphi}(x) := \begin{cases} s_{n_0}(x), & x \in K,\\ 0, & x \in U \setminus U_k. \end{cases}$$

Da K und $U \setminus U_k$ disjunkt und beide abgeschlossen (in der Relativtopologie von U) sind, existiert nach dem Satz von Tietze (Satz 6.19) eine Fortsetzung $\varphi \in C(U)$ mit $|\varphi(x)| \le \|s_{n_0}\|_\infty \quad (x \in U)$. Wegen $\varphi = 0$ auf $U \setminus U_k$ ist $\operatorname{supp}\varphi \subset \overline{U_k} \subset U$, d.h. $\varphi \in C_c(U)$.

Es gilt

$$\begin{aligned}\|s_{n_0} - \varphi\|_p^p &= \int_U |s_{n_0} - \varphi|^p d\lambda = \int_{U_k\setminus K} |s_{n_0} - \varphi|^p d\lambda\\ &\le 2^p \|s_{n_0}\|_\infty^p\, \lambda(U_k \setminus K) < \Big(\frac{\varepsilon}{2}\Big)^p.\end{aligned}$$

Damit ist $\|f - \varphi\|_p \le \|f - f_k\|_p + \|f_k - s_{n_0}\|_p + \|s_{n_0} - \varphi\|_p < \varepsilon.$ □

6.21 Satz. *Sei $U \subset \mathbb{R}^n$ offen. Dann liegt $C_c^\infty(U)$ dicht in $L^p(U)$ für alle $1 \le p < \infty$.*

Beweis. Der Beweis wird unter Verwendung des Friedrichschen Glättungsoperators geführt. Seien $f \in L^p(U)$ und $\varepsilon > 0$.

Nach Satz 6.20 existiert ein $g \in C_c(U)$ mit $\|f - g\|_p < \frac{\varepsilon}{2}$.

Wähle $\varphi \in \mathscr{D}(\mathbb{R}^n)$ mit $\varphi \ge 0$, $\int \varphi(x)dx = 1$ und $\operatorname{supp}\varphi \subset \{x \in \mathbb{R}^n : |x| \le 1\}$. Setze dann $\varphi_\eta(x) := \eta^{-n}\varphi(\frac{x}{\eta})$ für $\eta > 0$. Dann gilt $\operatorname{supp}\varphi_\eta \subset \{|x| \le \eta\}$.

Nach Satz 6.17 gilt $g * \varphi_\eta \in C^\infty(\mathbb{R}^n)$. Nach Lemma 6.13 b) ist für hinreichend kleine η die Menge

$$K := \operatorname{supp} g \cup \operatorname{supp}(g * \varphi_\eta) \subset U$$

kompakt, d.h. $g * \varphi_\eta \in \mathscr{D}(U)$.

Für $x \in K$ gilt

$$|g * \varphi_\eta(x) - g(x)| = \Big| \int_{\mathbb{R}^n} \varphi_\eta(x-y)\big(g(x) - g(y)\big)dy \Big|$$

$$\leq \Big(\sup_{|x-y|\leq\eta} |g(x)-g(y)| \Big) \int_{\mathbb{R}^n} \varphi_\eta(x-y)dy$$
$$= \sup_{|x-y|\leq\eta} |g(x)-g(y)|.$$

Da K kompakt und $g \in C(K)$, ist g auf K gleichmäßig stetig. Damit existiert ein $\eta > 0$ mit

$$\sup_{|x-y|\leq\eta} |g(x)-g(y)| < \frac{\varepsilon}{2} \Big(\frac{1}{\lambda(K)} \Big)^{1/p}.$$

Damit folgt

$$\int_U |g * \varphi_\eta(x) - g(x)|^p dx \leq \lambda(K) \sup_{x\in K} |g * \varphi_\eta(x) - g(x)|^p < \Big(\frac{\varepsilon}{2}\Big)^p.$$

Damit erhalten wir $\|f - (g * \varphi_\eta)\|_p \leq \|f-g\|_p + \|g - (g*\varphi_\eta)\|_p < \varepsilon.$ □

Literatur

[1] H. Bauer: Maß- und Integrationstheorie. De Gruyter, Berlin 1990.

[2] L. Breiman: Probability. Addison-Wesley, Reading 1968.

[3] K. L. Chung: A Course in Probability Theory. 2nd edition. Academic Press, New York 1974.

[4] R. Denk, R. Racke: Kompendium der Analysis, Band 2: Maß- und Integrationstheorie, Funktionalanalysis, Partielle Differentialgleichungen. Springer Spektrum, Wiesbaden 2012.

[5] W. Hackenbroch: Integrationstheorie. Teubner, Stuttgart 1987.

[6] P. R. Halmos: Measure Theory. Van Nostrand Reinhold, New York 1969.

[7] D. Hoffmann, F.-W. Schäfke: Integrale. BI-Wiss.-Verl., Mannheim 1992.

[8] J. C. Oxtoby: Maß und Kategorie. Springer, Berlin 1971.

[9] W. Rudin: Reelle und komplexe Analysis. Oldenburg-Verlag, München 1999.

[10] W. Rudin: Functional analysis. McGraw-Hill, New York 1973.

Index

äußeres Maß, 5
  zu Inhalt, 5

absolutstetig, 36
Algebra, 1

bewegungsinvariant, 31
Bildmaß, 29
Borel-σ-Algebra, 9

Dynkin-Lemma, 2
Dynkin-System, 1

elementargeometrischer Inhalt, 4
endlicher Inhalt, 3

Faltung, 44
fast überall, 13
fast sicher, 13
Friedrichscher Glättungsoperator, 48

Höldersche Ungleichung, 41
Halbnorm, 41

Inhalt, 3
  monoton, 4
  subadditiv, 5
  subtraktiv, 5
integrierbar, 15, 40

Koordinatenprojektion, 23

Lebesgue-Maß, 11
Lebesgue-Mengen, 11
Lemma von Fatou, 19
lokal integrierbar, 45
$L^p(\mu)$, 40

Maß, 3
Maße durch Dichten, 16
Maßraum, 3
Majorantenkriterium, 16
messbar (bzgl. äußeres Maß), 5
messbare Abbildung, 14
Messraum, 1
Minkowskische Ungleichung, 41

normal, 3
Nullmenge, 13

offen, 9

Produkt-σ-Algebra, 22
Produktmaß, 24

reguläres Maß, 12
Riemann-Stieltjes-Integral, 30
Ring, 1

Sardsche Ungleichung, 34
Satz
  Eindeutigkeitssatz, 7
  Fortsetzungssatz, 8
  Fortsetzungssatz von Tietze, 47
  Lemma von Fatou, 19
  Parameterabhängigkeit von Integralen, 20
  Transformationssatz, 31
  vom bedingten Erwartungswert, 39
  von Carathéodory, 6
  von Fubini, 27
  von Lebesgue über majorisierte Konvergenz, 20
  von Lebesgue über monotone Konvergenz, 17
  von Lusin, 46
  von Radon-Nikodym, 37
  von Tonelli, 26
Seminorm, 41
σ-additiv, 3
σ-Algebra, 1
σ-endlich, 3
σ-finit, 3
Spurmaß, 15
Stufenfunktion, 15
symmetrische Differenz, 9

Testfunktion, 44
Topologie, 9
topologischer Raum, 9
Träger, 43
Transformationslemma, 29

Verteilungsfunktion, 10
Vervollständigung eines Maßes, 8
vollständiges Maß, 8

Wahrscheinlichkeitsmaß, 3

Zufallsvariable, 14
Zylindermengen, 22