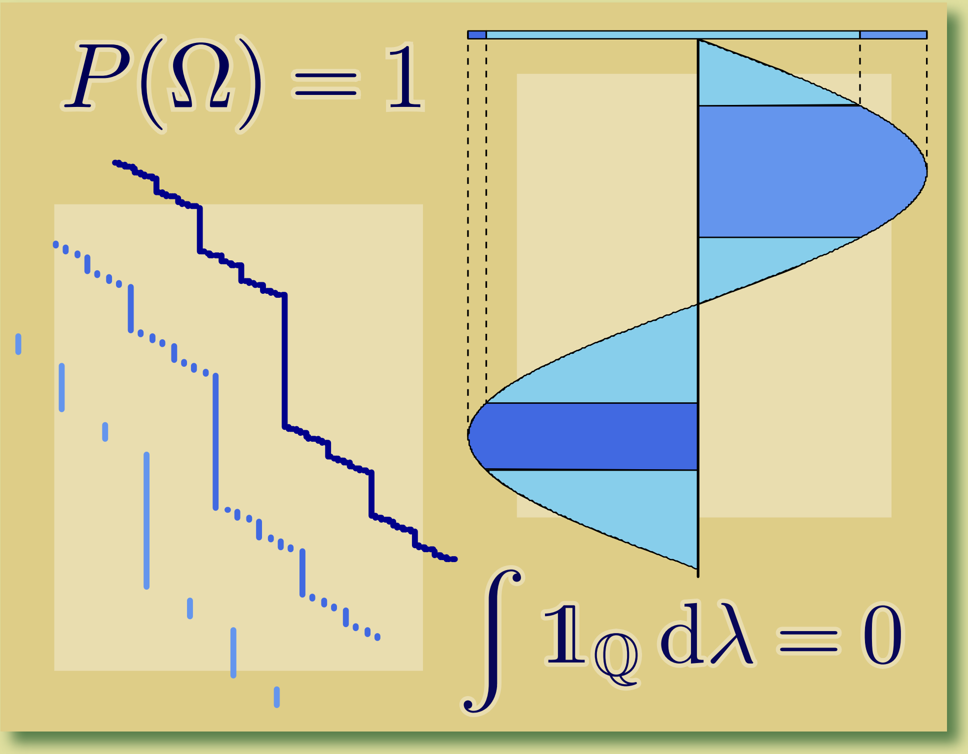
Axel Munk

Maß- und Integrationstheorie



Universitätsdrucke Göttingen

Axel Munk

Maß- und Integrationstheorie

erschienen in der Reihe der Universitätsdrucke
im Universitätsverlag Göttingen 2010

Axel Munk

Maß- und Integrationstheorie

Universitätsverlag Göttingen
2010

Bibliographische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliographie; detaillierte bibliographische Daten sind im Internet über <http://dnb.ddb.de> abrufbar.

Anschrift des Autors
Prof. Dr. Axel Munk
Georg-August-Universität Göttingen
Institut für Mathematische Stochastik
Goldschmidtstr. 7
37077 Göttingen

Satz und Layout: Steffi Greiner, Alexander Hartmann
Umschlaggestaltung: Jutta Pabst

http://univerlag.uni-goettingen.de
ISBN: 978-3-941875-85-2

Vorwort

Die Maßtheorie entstand aus dem Bedürfnis, Flächen oder Volumina von Körpern präzise messen zu können, die komplizierter aufgebaut sind als einfache Grundgebilde wie Quader oder Kugeln, deren Inhalte sich mit elementargeometrischen Überlegungen oder durch Symmetriebetrachtungen leicht erschließen lassen. Der Leitgedanke ist stets, allgemeinere Körper durch einfache Systeme wie Quader geeignet zu approximieren und diese Approximation sukzessive zu verfeinern. Schnell wird klar, dass nicht jeder beliebigen Menge ein sinnvoller Inhalt zugeordnet werden kann. Dies führt auf den Begriff des Maßraums, ein System von geeigneten Mengen mit einer Funktion, welche diesen Volumina zuordnet. Gewichtet man die Volumenbildung über solchen Mengensystemen gemäß dieser Funktionen, so führt dies auf die Idee des Lebesgue-Integrals, eine Art Mittelung dieser Funktion. Analog wird das allgemeine Maßintegral konstruiert, welches geeigneten Mengensystemen über allgemeinen Grundräumen einen Inhalt zuordnet. Hieraus ergibt sich sofort ein grundlegender Zusammenhang zur Wahrscheinlichkeitstheorie: Erwartungswerte von Zufallsvariablen können als Mittelungen dieser Funktionen verstanden werden, wobei die Volumenmessung der Wahrscheinlichkeit eines durch die Zufallsvariable induzierten Ereignisses entspricht.

Dieses Skript hat zum Ziel, die notwendigen Begriffe und Techniken der Maß- und Integrationstheorie bereitzustellen, um insbesondere die Grundlagen für eine moderne Wahrscheinlichkeitstheorie zu legen. Dabei können viele vor allem auch für die Analysis wichtige Aspekte aus Platzgründen nicht behandelt werden. Es wurde jedoch insbesondere Wert auf den Zusammenhang und die Unterschiede von Lebesgue- und Riemann-Integral gelegt. Insbesondere das Riemann-Integral, welches lange Zeit als überholt betrachtet wurde, bildet die Grundlage für moderne Integrationstheorien der stochastischen Analysis und hat hierdurch eine Renaissance erfahren.

Dieses Skript ist für einen zweistündigen Kurs konzipiert. Die mit ★ gekennzeichneten Abschnitte gehen über den Stoff einer zweistündigen Vorlesung hinaus und sind für das weitere Verständnis nicht notwendig.

Ich möchte mich an dieser Stelle sehr herzlich bei meiner Kollegin Prof. Anja Sturm für wertvolle Hinweise und Korrekturen und bei Alexander Hartmann, Benjamin Heuer und Achim Wübker für die ausgezeichnete Hilfe bei der Ausarbeitung dieses Skripts bedanken.

Göttingen, August 2010 — Axel Munk

Inhaltsverzeichnis

Motivation

Die Begriffe Fläche und Volumen sind jedem von uns vertraut und wir gehen im Alltag intuitiv damit um. Aufgrund dieser Vertrautheit entstand eine präzise und allgemeine Definition erst verhältnismäßig spät, etwa zu Beginn des 20. Jahrhunderts. An eine sinnvolle Verwendung des Begriffs Volumen stellen wir zunächst folgende Forderungen:

F1 Einem Gebilde wird eine nichtnegative Zahl zugeordnet: sein Volumen.

F2 Zwei kongruente Gebilde haben das gleiche Volumen.

F3 Besteht ein Gebilde aus mehreren disjunkten Einzelgebilden, so ist das Volumen des Gebildes gleich der Summe der Volumina der Einzelgebilde.

Formal fassen wir die Gebilde als Teilmengen des $\mathbb{R}^n$ auf und bezeichnen die Potenzmenge mit $\mathscr{P}(\mathbb{R}^n) := \{A \text{ Menge} \mid A \subseteq \mathbb{R}^n\}$. Forderung F1 bedeutet, dass wir eine Funktion

$$\begin{aligned} \iota\colon \mathscr{P}(\mathbb{R}^n) &\to [0,\infty] := [0,\infty) \cup \{\infty\} \\ A &\mapsto \iota(A) \end{aligned}$$

suchen, die einer Teilmenge $A \subseteq \mathbb{R}^n$ ihr Volumen $\iota(A)$ zuordnet. Um die zweite Forderung zu formalisieren, müssen wir den Begriff der Kongruenz definieren.

Definition 0.1 (Kongruenz)**.** Wir bezeichnen zwei Mengen $A, B \in \mathscr{P}(\mathbb{R}^n)$ als *kongruent* (engl.: congruent), falls es eine orthogonale Matrix $U \in \mathbb{R}^{n\times n}$ und $v \in \mathbb{R}^n$ gibt, sodass mit

$$U(A) + v := \{Ux + v \mid x \in A\}$$

gilt:

$$B = U(A) + v.$$

Die orthogonale Matrix U bewirkt dabei eine Drehung oder eine Spiegelung, der Vektor v eine Verschiebung der Menge A. Zusammen bewegen sie die Menge A durch den Raum $\mathbb{R}^n$.

Mit dieser Definition können wir F2 für die Funktion $\iota\colon \mathscr{P}(\mathbb{R}^n) \to [0,\infty]$ formulieren:

$$\iota(A) = \iota(B), \text{ falls } A \text{ und } B \text{ kongruent sind.}$$

Diese Eigenschaft heißt naheliegenderweise *Bewegungsinvarianz*. Für die Formalisierung der dritten Forderung erinnern wir daran, dass zwei Teilmengen A und B disjunkt heißen, wenn $A \cap B = \varnothing$ gilt. Damit erhalten wir eine weitere Forderung an unsere Funktion, die *paarweise Additivität*:

$$\iota(A \cup B) = \iota(A) + \iota(B), \quad \text{falls } A \text{ und } B \text{ disjunkt sind.}$$

Mittels eines Induktionsargumentes kann man aus dieser direkt die *endliche Additivität* folgern, wodurch F3 erfüllt ist:

$$\iota\left(\bigcup_{i=1}^{k} A_i\right) = \sum_{i=1}^{k} \iota(A_i), \quad \text{falls die } A_i \text{ paarweise disjunkt sind.}$$

Formal genügt die Nullfunktion $\iota \equiv 0$ all unseren Anforderungen. Um diese auszuschließen, fügen wir eine letzte Forderung, die *Normiertheit*, hinzu: Das Einheitsintervall $[0,1]$ soll die Länge 1 haben, das Einheitsquadrat $[0,1]^2$ die Fläche 1 und so weiter:

$$\iota([0,1]^n) = 1 \text{ für alle } n \in \mathbb{N}.$$

Das Inhaltsproblem

Aus dieser Formalisierung ergibt sich das sogenannte Inhaltsproblem: Gesucht ist eine Funktion $\iota\colon \mathscr{P}(\mathbb{R}^n) \to [0,\infty]$, welche endlich additiv, bewegungsinvariant und normiert ist. Es stellt sich heraus, dass dieses Problem für $n \in \{1,2\}$ nicht eindeutig lösbar und für $n \geqslant 3$ gar völlig unlösbar ist. Unsere scheinbar banalen Forderungen an die Funktion ι führen also zu überraschend tiefen Schwierigkeiten. Die Ursache hierfür liegt darin, dass wir F1 - F3 für *alle* Teilmengen des $\mathbb{R}^n$ fordern. Der folgende Satz unterstreicht, wie bereits das Verhalten von Teilmengen im $\mathbb{R}^3$ unserer Intuition zuwider laufen kann.

Satz 0.2 (Banach-Tarsky-Paradoxon, 1924)**.** *Seien A und B beschränkte Teilmengen des $\mathbb{R}^3$ und das Innere von A sowie das Innere von B nicht leer, so gibt es ein $m \in \mathbb{N}$ und paarweise disjunkte Mengen $A_i \subseteq \mathbb{R}^3$ für alle $i \in \{1,\dots,m\}$ sowie paarweise disjunkte Mengen $B_i \subseteq \mathbb{R}^3$ für alle $i \in \{1,\dots,m\}$, sodass gilt:*

$$A = \bigcup_{i=1}^{m} A_i, \quad B = \bigcup_{i=1}^{m} B_i$$

und A_i ist kongruent zu B_i für alle $i \in \{1,\dots,m\}$.

Anschaulich bedeutet dies, dass es zum Beispiel möglich ist, eine Kugel vom Radius 1 so in endlich viele Stücke zu zerlegen und wieder zusammenzusetzen, dass 1000 Kugeln vom Radius 1000 entstehen. Die einzelnen Teile sind zwar nicht konstruktiv bestimmbar, doch ihre Existenz kann aus dem Auswahlaxiom A.1 abgeleitet werden. Für eine weiterführende Diskussion und einen Beweis vergleiche [11], Seite 27. Vorkenntnisse in Algebra sind hierfür hilfreich.

Das Maßproblem

Obwohl bereits das Inhaltsproblem unlösbar ist, wollen wir unsere Forderungen noch ver-schärfen. Es wird sich herausstellen, dass diese Verschärfung nicht zu einer echten Einschränkung führt, für den Aufbau einer starken Theorie jedoch unabdingbar ist.

Betrachten wir eine krummlinig begrenzte Fläche im $\mathbb{R}^2$. Um den Flächeninhalt zu bestimmen, können wir die Fläche durch disjunkte Rechtecke approximieren. Gemäß unserer Intuition sollte eine Verfeinerung durch immer mehr und immer kleinere Rechtecke die gesuchte Fläche immer besser approximieren und im Grenzwert genau den gesuchten Flächeninhalt ergeben. Für die Formalisierung erweitern wir dazu die endliche Additivität zu einer abzählbaren Additivität, der sogenannten *σ-Additivität.* Wir gelangen zu folgender Fragestellung:

Problem 0.3 (Maßproblem)**.** Gesucht ist eine „Maßfunktion" $\mu\colon \mathscr{P}(\mathbb{R}^n) \to [0,\infty]$ mit folgenden Eigenschaften:

M1 *σ-Additivität*: Seien $A_i \in \mathscr{P}(\mathbb{R}^n)$ für alle $i \in \mathbb{N}$ paarweise disjunkt, so gilt:

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

M2 *Bewegungsinvarianz*: Sind $A, B \in \mathscr{P}(\mathbb{R}^n)$ kongruent, so ist $\mu(A) = \mu(B)$.

M3 *Normiertheit*: $\mu([0,1]^n) = 1$ für alle $n \in \mathbb{N}$.

Satz 0.4. *Das Maßproblem 0.3 ist für alle $n \in \mathbb{N}$ unlösbar.*

Beweis. Wir beginnen mit dem eindimensionalen Fall $n = 1$ und führen einen Widerspruchsbeweis. Sei $\mu\colon \mathscr{P}(\mathbb{R}) \to [0,\infty]$ eine Funktion mit den Eigenschaften M1 - M3. Wir betrachten die Quotientengruppe $\mathbb{R}/\mathbb{Q} := \{[x] \mid x \in \mathbb{R}\}$ der Äquivalenzklassen $[x]$ bezüglich der Äquivalenzrelation

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

und ein Repräsentantensystem R der Nebenklassen, von dem wir ohne Einschränkung $R \subseteq [0,1]$ annehmen können. Die Existenz von R ist durch das Auswahlaxiom gewährleistet. Wir erhalten die abzählbare disjunkte Vereinigung

$$\mathbb{R} = \bigcup_{q\in\mathbb{Q}} (q + R).$$

Ist $\mu(R) = 0$, so folgt aus der σ-Additivität und der Bewegungsinvarianz:

$$\mu(\mathbb{R}) = \sum_{q\in\mathbb{Q}} \mu(q + R) = 0.$$

Da aus der σ-Additivität $\mu(A) \leqslant \mu(B)$ für alle $A \subseteq B$ folgt, gilt $\mu([0,1]) = 0$ im Widerspruch zu (M3). Ist hingegen $\mu(R) > 0$, so ist mit der Translationsinvarianz von μ

$$\infty = \sum_{q\in\mathbb{Q}\cap[0,1]} \mu(q + R) \leqslant \mu([0,2]) \leqslant \mu([0,1]) + \mu([1,2]) = 2,$$

sodass wir wiederum einen Widerspruch erhalten. Den Fall $n > 1$ beweist man völlig analog mit Hilfe der Quotientengruppe $\mathbb{R}^n/\mathbb{Q}^n$.

Um den Begriff des Volumens dennoch mathematisch sinnvoll zu formalisieren, eröffnen sich uns prinzipiell drei Möglichkeiten:

1. Zum Beweis der Unlösbarkeit des Maßproblems haben wir das Auswahlaxiom verwendet. Daher könnte man das Auswahlaxiom als Axiom der Mengenlehre ablehnen. Aber das Auswahlaxiom sowie die äquivalenten Aussagen des Wohlordnungssatzes und des Zornschen Lemmas besitzen in der Mathematik eine bedeutende Stellung und erlauben es, viele Beweise sehr einfach zu führen. Deshalb erscheint eine Ablehnung des Auswahlaxioms nicht sinnvoll.

2. Abstriche bei den Forderungen an unsere Volumenfunktion wären denkbar. Bewegungsinvarianz und Additivität sind jedoch so elementar für unsere Vorstellung eines Volumens, dass wir daran festhalten werden. Die Normiertheit dient lediglich dazu, die unliebsame Nullfunktion als Volumenfunktion auszuschließen. Die Festlegung irgendeines anderen Volumens $a > 0$ für den Einheitswürfel macht die Probleme jedoch auch nicht lösbar.

3. Es bleibt schließlich, den Definitionsbereich einzuschränken. Wir wissen, aufgrund der Überlegungen in diesem Abschnitt, dass es uns nicht gelingen wird, ein Maß auf der ganzen Potenzmenge zu finden. Wir werden jedoch sehen, dass das Mengensystem, auf dem wir ein Maß erhalten, groß genug sein wird, um alle vernünftigen Mengen zu enthalten.

Kapitel 1

Maßräume

1.1 Allgemeine Maßräume

Notation 1.1. Sind zwei Mengen A und B disjunkt, so schreiben wir für ihre Vereinigung auch $A + B := A \cup B$. Ist I eine Indexmenge und sind die Mengen $\{A_i\}_{i\in I}$ paarweise disjunkt, so schreiben wir analog $\sum_{i\in I} A_i := \bigcup_{i\in I} A_i$.

Definition 1.2. Sei $\Omega \neq \varnothing$ eine Menge. Wir bezeichnen $\mathscr{A} \subseteq \mathscr{P}(\Omega)$ als *σ-Algebra* (engl.: σ-algebra, σ-field) über Ω, falls folgende Bedingungen erfüllt sind:

1. $\Omega \in \mathscr{A}$.
2. Es ist $A \in \mathscr{A}$ genau dann, wenn $A^c \in \mathscr{A}$.
3. Gilt $A_n \in \mathscr{A}$ für alle $n \in \mathbb{N}$, so folgt $\bigcup_{n\in\mathbb{N}} A_n \in \mathscr{A}$.

Definition 1.3 (Kolmogorov-Axiome, 1933)**.** Sei $\mathscr{A}$ eine σ-Algebra. $P\colon \mathscr{A} \to \mathbb{R}$ heißt *Wahrscheinlichkeitsmaß* (W-Maß, engl.: probability measure) auf $\mathscr{A}$, falls folgende Bedingungen erfüllt sind:

1. Für alle $A \in \mathscr{A}$ gilt $P(A) \geqslant 0$ (Nichtnegativität).
2. $P(\Omega) = 1$ (Normierung).
3. Sei $\{A_n\}_{n\in\mathbb{N}}$ ein paarweise disjunkte Familie von Elementen von $\mathscr{A}$, das heißt für alle $i \neq j$; $i, j \in \mathbb{N}$ gilt $A_i \cap A_j = \varnothing$, dann gilt
$$P\left(\sum_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} P(A_n) \quad (\sigma\text{-Additivität}).$$

Bemerkung 1.4. Falls wir in Definition 1.3 auf die Normierung von P verzichten, so heißt P *Maß* (vgl. Definition 1.25).

Definition 1.5. Seien $\Omega \neq \emptyset$ eine Menge, $\mathscr{A}$ eine σ-Algebra auf Ω und P ein Wahrscheinlichkeitsmaß auf $\mathscr{A}$. $(\Omega, \mathscr{A})$ heißt *Messraum* (engl.: measurable space). $(\Omega, \mathscr{A}, P)$ heißt *Wahrscheinlichkeitsraum* (W-Raum, engl.: probability space).

Definition 1.6. Ein Mengensystem M heißt $\cap$*-stabil* (engl.: closed under countable intersection), wenn es abgeschlossen bezüglich endlicher Schnitte ist, also falls gilt:

$$A, B \in M \Rightarrow A \cap B \in M.$$

Ein Mengensystem M heißt $\cup$*-stabil* (engl.: closed under countable union), wenn es abgeschlossen bezüglich endlicher Vereinigungen ist, also falls gilt:

$$A, B \in M \Rightarrow A \cup B \in M.$$

Definition 1.7. Sei $\Omega \neq \emptyset$ eine Menge. $\mathscr{A}_0 \subseteq \mathscr{P}(\Omega)$ heißt *Algebra* oder λ*-System* (engl.: algebra), falls folgende Bedingungen erfüllt sind:

1. $\Omega \in \mathscr{A}_0$.
2. Es ist $A \in \mathscr{A}_0$ genau dann, wenn $A^c \in \mathscr{A}_0$.
3. $\mathscr{A}_0$ ist $\cup$*-stabil*.

Eine Algebra unterscheidet sich also dadurch von einer σ-Algebra, dass sie nur bezüglich endlicher, nicht aber abzählbarer, Vereinigungen abgeschlossen sein muss.

Bemerkung 1.8. Folgende Implikationen ergeben sich leicht aus der Definition einer Algebra:

1. $\mathscr{A}_0$ ist Algebra $\Rightarrow \emptyset \in \mathscr{A}_0$.
2. $\mathscr{A}_0$ ist $\cap$*-stabil*.
3. $\mathscr{A}$ ist σ-Algebra $\Rightarrow \mathscr{A}$ ist Algebra.
4. Die Umkehrung von 3. gilt im Allgemeinen nicht.
   Gegenbeispiel: $\Omega = \mathbb{N}$, $\mathscr{A}_0 = \{A \subseteq \Omega \mid A \text{ endlich oder } A^c \text{endlich}\}$. [Übung]

Definition 1.9. Sei $\Omega \neq \emptyset$ eine Menge. Wir bezeichnen $\mathscr{D} \subseteq \mathscr{P}(\Omega)$ als *Dynkin-System* (engl.: Dynkin system), falls

1. $\Omega \in \mathscr{D}$,
2. $D, E \in \mathscr{D}, D \subseteq E \Rightarrow E \backslash D \in \mathscr{D}$,
3. $D_n \in \mathscr{D}$ für alle $n \in \mathbb{N}$ paarweise disjunkt $\Rightarrow \sum_{n \in \mathbb{N}} D_n \in \mathscr{D}$.

Lemma 1.10.

1. *Jede σ-Algebra ist ein Dynkin-System.*
2. *Umkehrung: Jedes $\cap$-stabile Dynkin-System ist eine σ-Algebra.*

Beweis.

1. Sei $\mathscr{A}$ eine σ-Algebra. Dann ist $\Omega \in \mathscr{A}$ und $\mathscr{A}$ ist abgeschlossen bezüglich abzählbarer Vereinigungen, insbesondere auch Vereinigungen disjunkter Mengen.
   Seien nun $D, E \in \mathscr{A}$ mit $D \subseteq E$. Da $\mathscr{A}$ nach Bemerkung 1.8 $\cap$-stabil ist, gilt:
   $$E \backslash D = E \cap D^c \in \mathscr{A}.$$
2. Sei $\mathscr{D}$ ein $\cap$-stabiles Dynkin-System. Es bleibt zu zeigen, dass $\mathscr{D}$ abgeschlossen ist bezüglich abzählbarer Vereinigungen. Seien $D_1, D_2, \ldots \in \mathscr{D}$. Dann gilt:
   $$\bigcup_{n \in \mathbb{N}} D_n = D_1 + \sum_{n=2}^{\infty} (\underbrace{D_n \cap D_1^c \cap \cdots \cap D_{n-1}^c}_{\text{in } \mathscr{D} \text{ paarweise disjunkt}}) \in \mathscr{D}.$$

Lemma 1.11. *Seien $\Omega \neq \varnothing$ eine Menge, $I \neq \varnothing$ eine Indexmenge und $\mathscr{A}_i$ eine σ-Algebra über Ω für alle $i \in I$. Dann ist auch*

$$\bigcap_{i \in I} \mathscr{A}_i = \{A \in \mathscr{A}_i \text{ für alle } i \in I\}$$

eine σ-Algebra über Ω. Diese Aussage gilt analog auch für Algebren und Dynkin-Systeme. [Übung]

Satz 1.12. *Für eine Menge $\Omega \neq \varnothing$ sei $M \subseteq \mathscr{P}(\Omega)$. Dann sind*

$$\begin{aligned}
\sigma(M) &:= \bigcap \{\mathscr{A}\, \sigma\text{-Algebra} \mid M \subseteq \mathscr{A} \subseteq \mathscr{P}(\Omega)\}\,, \\
\alpha(M) &:= \bigcap \{\mathscr{A}_0 \text{ Algebra} \mid M \subseteq \mathscr{A}_0 \subseteq \mathscr{P}(\Omega)\}\,, \\
\delta(M) &:= \bigcap \{\mathscr{D} \text{ Dynkin-System} \mid M \subseteq \mathscr{D} \subseteq \mathscr{P}(\Omega)\}
\end{aligned}$$

wiederum σ-Algebra, Algebra beziehungsweise Dynkin-System über Ω.

Beweis. Für $\sigma(M)$ wende Lemma 1.11 auf folgende Indexmenge an:

$$I := \{\mathscr{A}\, \sigma\text{-Algebra} \mid M \subseteq \mathscr{A} \subseteq \mathscr{P}(\Omega)\}\,.$$

Wähle analog Indexmengen für $\alpha(M)$ und $\delta(M)$.

Definition 1.13. In Satz 1.12 heißt M *Erzeuger* (engl.: generator) oder *Erzeugendensystem*, $\sigma(M)$, $\alpha(M)$ und $\delta(M)$ heißen die *von M erzeugte σ-Algebra, Algebra* beziehungsweise *von M erzeugtes Dynkin-System.*

Bemerkung 1.14. Aus der Definition folgt direkt, dass $\sigma(M)$ die „kleinste“ σ-Algebra ist, welche M enthält. Das heißt:

$$\mathscr{A}\ \sigma\text{-Algebra mit } M \subseteq \mathscr{A} \subseteq \mathscr{P}(\Omega) \Rightarrow \sigma(M) \subseteq \mathscr{A}.$$

Wir sagen auch: $\sigma(M)$ ist minimal. Analog sind $\alpha(M)$ und $\delta(M)$ die kleinste Algebra beziehungweise das kleinste Dynkin-System, welche M enthalten.

Definition 1.15. Sei $\mathscr{A}$ eine σ-Algebra. Gibt es ein abzählbares Mengensystem $M \subseteq \mathscr{A}$ mit $\sigma(M) = \mathscr{A}$, so heißt $\mathscr{A}$ *abzählbar erzeugt* (engl.: countably generated). Entsprechend definieren wir abzählbar erzeugte Algebren und Dynkin-Systeme.

Beispiel 1.16. Sei $M := \{\{\omega\} \mid \omega \in \Omega\}$. Dann sind

1. $\sigma(M) = \{A \subseteq \Omega \mid A \text{ abzählbar oder } A^c \text{ abzählbar}\}$,
2. $\alpha(M) = \{A \subseteq \Omega \mid A \text{ endlich oder } A^c \text{ endlich}\}$,
3. $\delta(M) = \sigma(M)$.

[Übung] (Tipp: $M \cup \{\varnothing\}$ ist $\cap$-stabil, Lemma 1.18)

Regeln 1.17. *Sei $\Omega \neq \varnothing$ eine Menge und seien $M, M_1, M_2 \subseteq \mathscr{P}(\Omega)$. Dann gilt:*

1. $M \subseteq \sigma(M)$,
2. $M_1 \subseteq M_2 \Rightarrow \sigma(M_1) \subseteq \sigma(M_2)$,
3. $\sigma(M) = \sigma(\sigma(M))$,
4. $\alpha(M) \subseteq \sigma(M), \delta(M) \subseteq \sigma(M)$,
5. $M_1 \subseteq \sigma(M_2), M_2 \subseteq \sigma(M_1) \Rightarrow \sigma(M_1) = \sigma(M_2)$.

Beweis.

1. Diese Aussage folgt direkt aus der Definition von $\sigma(M)$.
2. Sei $M_1 \subseteq M_2$. Dann gilt $\sigma(M_2) \supseteq M_2 \supseteq M_1$. Aus der Minimalität von $\sigma(M_1)$ folgt somit $\sigma(M_1) \subseteq \sigma(M_2)$.
3. $\sigma(M)$ ist bereits eine σ-Algebra und außerdem minimal.
4. Nach Bemerkung 1.8 und Lemma 1.10 ist $\sigma(M)$ auch eine Algebra und ein Dynkin-System. Aus der Minimalität von $\alpha(M)$ und $\delta(M)$ folgt somit die Behauptung.

5. Aus $M_1 \subseteq \sigma(M_2)$ folgt mit 2. und 3., dass $\sigma(M_1) \subseteq \sigma(\sigma(M_2)) = \sigma(M_2)$ und umgekehrt.

Lemma 1.18. *Ist $M \subseteq \mathscr{P}(\Omega)$ $\cap$-stabil, so gilt*

$$\delta(M) = \sigma(M).$$

Beweis. Mit Lemma 1.10 und Regel 1.17.4. bleibt zu zeigen, dass $\delta(M)$ $\cap$-stabil ist. Seien $A \in \delta(M)$ und $\mathscr{D}_A := \{B \subseteq \Omega \mid B \cap A \in \delta(M)\}$. Dann ist $M \subseteq \mathscr{D}_A$ für alle $A \in M$, da M $\cap$-stabil ist. Es gilt:

- $\Omega \cap A = A \in \delta(M)$.
- Seien $D, E \in \mathscr{D}_A$ mit $D \subseteq E$. Dann ist $(E \backslash D) \cap A = (E \cap A) \backslash (D \cap A) \in \delta(M)$, da $(D \cap A) \subseteq (E \cap A)$ und beide Schnitte in $\delta(M)$ liegen.
- Seien $D_1, D_2, \ldots \in \mathscr{D}_A$ paarweise disjunkt. Dann ist

$$\left(\sum_{n \in \mathbb{N}} D_n \right) \cap A = \sum_{n \in \mathbb{N}} (D_n \cap A) \in \delta(M).$$

Folglich ist $\mathscr{D}_A$ ein Dynkin-System über Ω. Daraus folgt mit der Minimalität von $\delta(M)$, dass $\delta(M) \subseteq \mathscr{D}_A$ für alle $A \in M$. Außerdem gilt für alle $B \in \delta(M)$ und $A \in M$, dass $B \cap A \in \delta(M)$ und ebenso natürlich $A \cap B \in \delta(M)$. Somit ist $M \subseteq \mathscr{D}_B$ für alle $B \in \delta(M)$. Das wiederum impliziert, dass $\delta(M) \subseteq \mathscr{D}_B$ für alle $B \in \delta(M)$. Folglich ist $\delta(M)$ $\cap$-stabil.

Notation 1.19. Seien $x := (x_1, \ldots, x_k), y := (y_1, \ldots, y_k) \in \mathbb{R}^k$. Schreibe $x < y$ oder $x \leqslant y$, falls $x_i < y_i$ beziehungsweise $x_i \leqslant y_i$ für alle $i \in \{1, \ldots, k\}$.

Seien $x, y \in \mathbb{R}^k$ mit $x \leqslant y$. Analog zu Intervallen in $\mathbb{R}$ schreibe

$$(x, y] := \bigtimes_{i=1}^{k} (x_i, y_i]$$

für Quader im $\mathbb{R}^k$. Entsprechend definieren wir (x, y), $[x, y)$ und $[x, y]$.

Definition 1.20 (Borel-σ-Algebra)**.** Für jedes $k \in \mathbb{N}$ bezeichnen wir das System der k-dimensionalen offenen Mengen mit $\mathscr{O}^k := \left\{A \text{ Menge} \,\middle|\, A \subseteq \mathbb{R}^k \text{ offen}\right\}$. $\mathscr{B}^k := \sigma(\mathscr{O}^k)$ heißt *σ-Algebra der Borel-Mengen* oder *Borel-σ-Algebra* (engl.: Borel σ-algebra).

Bemerkung 1.21. Sei $k \in \mathbb{N}$. Wir werden später sehen, dass sich auf $\mathscr{B}^k$ ein Maß mit den im Maßproblem 0.3 geforderten Eigenschaften definieren lässt. Aus der Unlösbarkeit des Maßproblems folgt somit $\mathscr{B}^k \neq \mathscr{P}(\mathbb{R}^k)$.

Satz 1.22. *Weitere Erzeuger von $\mathscr{B}^k$ sind:*

$$\begin{aligned}
\mathscr{A}^k &:= \left\{A \subseteq \mathbb{R}^k \ \textit{abgeschlossen}\right\},\\
\mathscr{I}^k &:= \left\{(x,y] \,\middle|\, x,y \in \mathbb{R}^k, x \leqslant y\right\},\\
\mathscr{I}_0^k &:= \left\{(x,y) \,\middle|\, x,y \in \mathbb{R}^k, x \leqslant y\right\},\\
\mathscr{J}_\infty^k &:= \left\{(-\infty,x] \,\middle|\, x \in \mathbb{R}^k\right\}.
\end{aligned}$$

Beweis. Es reicht, folgende Inklusionen zu zeigen:

1. $\mathscr{A}^k \subseteq \sigma(\mathscr{O}^k) = \mathscr{B}^k$,
2. $\mathscr{J}_\infty^k \subseteq \sigma(\mathscr{A}^k)$,
3. $\mathscr{I}^k \subseteq \sigma(\mathscr{J}_\infty^k)$,
4. $\mathscr{I}_0^k \subseteq \sigma(\mathscr{I}^k)$,
5. $\mathscr{O}^k \subseteq \sigma(\mathscr{I}_0^k)$.

Mit den Regeln 1.17.2. und 1.17.3. folgt dann die Behauptung.

1. Sei $A \in \mathscr{A}^k$. Dann ist $A^c \in \mathscr{O}^k$ und somit ist $\mathscr{A}^k \subseteq \sigma(\mathscr{O}^k)$.
2. Es ist $\mathscr{J}_\infty^k \subseteq \mathscr{A}^k \subseteq \sigma(\mathscr{A}^k)$.
3. Zur Erläuterung der Grundidee des folgenden Beweises betrachten wir Abbildung 1.1. Um den Quader $\big((x_1,x_2),(y_1,y_2)\big]$ im $\mathbb{R}^2$ aus Elementen von $\mathscr{J}_\infty^k$ zu konstruieren, betrachten wir den Quader $\big((-\infty,-\infty),(y_1,y_2)\big]$ und schneiden die beiden Quader $\big((-\infty,-\infty),(x_1,y_2)\big]$ und $\big((-\infty,-\infty),(y_1,x_2)\big]$ weg.

   Für den allgemeinen Fall betrachten wir $x = (x_1,\ldots,x_k)$, $y = (y_1,\ldots,y_k) \in \mathbb{R}^k$ mit $x \leqslant y$. Für alle $m \in \{1,\ldots,k\}$ definiere

   $$z^m := (z_1^m,\ldots,z_k^m) \quad \text{mit} \quad z_n^m = \begin{cases} x_m, & \text{falls } n = m\\ y_m, & \text{falls } n \neq m \end{cases},$$

   sowie $I_m := (-\infty, z^m] \in \mathscr{J}_\infty^k$ für alle $m \in \{1,\ldots,k\}$. Dann gilt:

   $$(x,y] = (-\infty,y] \setminus \bigcup_{m=1}^{k} I_m \in \sigma(\mathscr{J}_\infty^k) \quad \text{für alle } (x,y] \in \mathscr{I}^k.$$

   Folglich ist $\mathscr{I}^k \subseteq \sigma(\mathscr{J}_\infty^k)$.
4. Seien $x,y \in \mathbb{R}^k$ wie oben. Definiere die Folge $(y^n)_{n\in\mathbb{N}}$ mit

   $$y^n := \big(\max\{x_1, y_1 - 1/n\},\ldots,\max\{x_k, y_k - 1/n\}\big) \in \mathscr{I}^k.$$

   Somit gilt

   $$\bigcup_{n=1}^{\infty} (x,y^n] = (x,y) \in \sigma(\mathscr{I}^k).$$

   Folglich ist $\mathscr{I}_0^k \subseteq \sigma(\mathscr{I}^k)$.

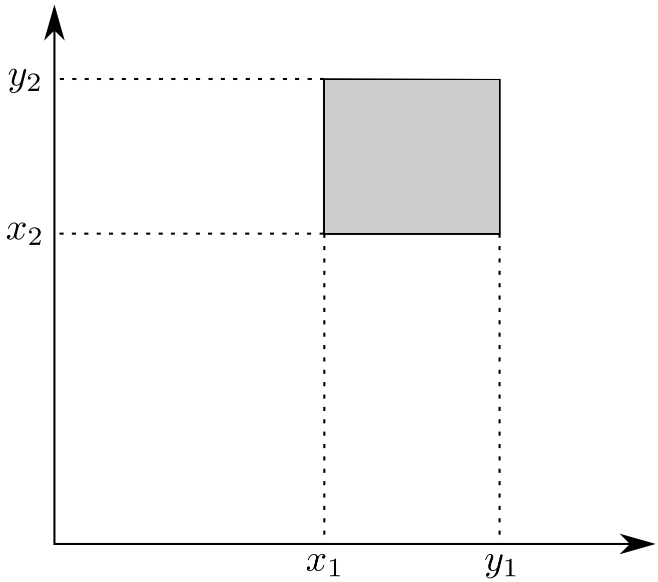


Abbildung 1.1: Der Quader $\big((x_1, x_2), (y_1, y_2)\big]$ im $\mathbb{R}^2$

5. Sei $A \in \mathscr{O}^k$. Wir definieren den offenen Würfel mit Kantenlänge $2r$ und Mittelpunkt q als
$$B_r^\circ(q) := \left\{x \in \mathbb{R}^k \;\middle|\; \|x - q\|_\infty < r\right\}.$$
Da A offen ist, gibt es für alle $a \in A$ ein $q_a \in \mathbb{Q}^k$ und $r_a \in \mathbb{Q}$ mit $a \in B_{r_a}^\circ(q_a) \subseteq A$. Somit gilt:
$$A = \bigcup_{(q,r) \in \mathbb{Q}^k \times \mathbb{Q}} B(q, r) \quad \text{mit} \quad B(q, r) := \begin{cases} B_r^\circ(q) & B_r^\circ(q) \subseteq A \\ \varnothing & \text{sonst} \end{cases}.$$
Folglich ist $\mathscr{O}^k \subseteq \sigma(\mathscr{I}_0^k)$.

Im Folgenden beschäftigen wir uns mit der Frage, ob eine auf einem Mengensystem M gegebene Mengenfunktion $\mu\colon M \to [0, \infty]$ immer geeignet auf die von M erzeugte σ-Algebra $\sigma(M)$ fortgesetzt werden kann. Dabei interessiert uns insbesondere, ob eine solche Fortsetzung eindeutig ist.

Betrachten wir zum Beispiel die borelsche σ-Algebra $\mathscr{B}^k$ mit der Menge der halboffenen Quader $\mathscr{I}^k$ als Erzeuger und die Volumenbildung als Mengenfunktion, so ermöglicht eine solche Fortsetzung eine Approximation komplizierter Mengen durch Quader.

Bemerkung 1.23. Wir können Definition 1.20 auf metrische Räume verallgemeinern. Sei (Ω, d) ein metrischer Raum. Wir bezeichnen die vom System der offenen Mengen erzeugte σ-Algebra als *Borel-σ-Algebra.* Ist der metrische Raum (Ω, d) separabel, so wird die Borel-σ-Algebra durch das System der offenen Bälle erzeugt. [Übung]

Zur Erinnerung: Ein metrischer Raum heißt separabel, falls er eine abzählbare dichte Teilmenge besitzt.

Definition 1.24. Sei $M \subseteq \mathscr{P}(\Omega)$ ein Mengensystem. $\mu\colon M \to [0,\infty]$ heißt *Mengenfunktion* (engl.: set function). μ heißt außerdem

1. *stetig* (engl.: continuous) oder *atomfrei*, falls $\mu(\{a\}) = 0$ für alle $\{a\} \in M$ mit $a \in \Omega$,
2. *endlich* (engl.: finite), falls $\mu(A) < \infty$ für alle $A \in M$,
3. *σ-endlich* (engl.: σ-finite), falls es $A_1 \subseteq A_2 \subseteq \ldots \in M$ gibt, sodass $\bigcup_{i\in\mathbb{N}} A_i = \Omega$ und $\mu(A_i) < \infty$ für alle $i \in \mathbb{N}$,
4. *endlich-additiv* (engl.: finitely additive), falls aus $A_1, \ldots, A_n \in M$ paarweise disjunkt und $\sum_{i=1}^n A_i \in M$ folgt, dass $\mu\left(\sum_{i=1}^n A_i\right) = \sum_{i=1}^n \mu(A_i)$,
5. *σ-additiv* (engl.: σ-additive), falls aus $A_1, A_2, \ldots \in M$ paarweise disjunkt und $\sum_{i\in\mathbb{N}} A_i \in M$ folgt, dass $\mu\left(\sum_{i\in\mathbb{N}} A_i\right) = \sum_{i\in\mathbb{N}} \mu(A_i)$.

Definition 1.25. Sei $M \subseteq \mathscr{P}(\Omega)$ ein Mengensystem. $\mu\colon M \to [0,\infty]$ heißt *Maß* (engl.: measure), falls M σ-Algebra, μ σ-additiv und $\mu(\varnothing) = 0$. Ist μ ein Maß, so heißt (Ω, M, μ) *Maßraum* (engl.: measure space). Ein Maßraum (Ω, M, μ) heißt *σ-endlich* (engl.: σ-finite) oder *σ-kompakt*, falls μ σ-endlich ist.

Beispiel 1.26. Sei $(\Omega, \mathscr{A})$ ein Messraum.

1. Für $\omega \in \Omega$ heißt
$$\delta_\omega\colon \mathscr{A} \to \{0,1\}, \qquad A \mapsto \begin{cases} 1 & \omega \in A \\ 0 & \text{sonst} \end{cases}$$
Dirac-Maß oder *Einpunktverteilung in ω.*
2. $\mu\colon \mathscr{A} \to \mathbb{N} \cup \{\infty\}$ mit $A \mapsto |A|$ heißt *Zählmaß.*
3. Seien μ_n Maße auf $(\Omega, \mathscr{A})$ für alle $n \in \mathbb{N}$ und sei $(b_n)_{n\in\mathbb{N}}$ eine Folge nichtnegativer Zahlen. Dann ist
$$\mu(A) := \sum_{n\in\mathbb{N}} b_n \mu_n(A)$$
ebenfalls ein Maß.
4. Seien beispielsweise $\mu_n := \delta_{\omega_n}$ und $b_n > 0$ für alle $n \in \mathbb{N}$ mit $\sum_{n\in\mathbb{N}} b_n = 1$. $\mu := \sum_{n\in\mathbb{N}} b_n \delta_{\omega_n}$ heißt *diskretes* auf $\{\omega_1, \omega_2, \ldots\}$ konzentriertes Wahrscheinlichkeitsmaß.

5. In 4. seien $\omega_n := n$ für alle $n \in \mathbb{N}_0$ und $b_n := \frac{\lambda^n}{n!}\mathrm{e}^{-\lambda}$ mit dem sogenannten *Intensitätsparameter* $\lambda > 0$.

$$\mu\colon \mathscr{P}(\mathbb{N}_0) \to [0,\infty), \qquad \mathscr{A} \mapsto \sum_{n\in\mathbb{N}_0} \frac{\lambda^n}{n!}\mathrm{e}^{-\lambda}\mathbb{1}_A(n)$$

heißt *Poisson-Maß*. Hierbei definieren wir für $A \in \mathscr{P}(\Omega)$ die *Indikatorfunktion*

$$\mathbb{1}_A\colon \Omega \to \{0,1\}\,, \qquad \omega \mapsto \begin{cases} 1 & \omega \in A \\ 0 & \text{sonst} \end{cases}.$$

Regeln 1.27. *Seien $(\Omega, \mathscr{A}, \mu)$ ein Maßraum und A, B, $A_1, A_2, \ldots \in \mathscr{A}$. Dann gilt:*

1. $A \subseteq B \Rightarrow \mu(A) \leqslant \mu(B)$ *(Monotonie),*
2. $A \subseteq B$, $\mu(A) < \infty \Rightarrow \mu(B\backslash A) = \mu(B) - \mu(A)$,
3. $\mu\left(\bigcup_{i\in\mathbb{N}} A_i\right) \leqslant \sum_{i\in\mathbb{N}} \mu(A_i)$ *(σ-Subadditivität),*
4. $A_n \nearrow A \Rightarrow \mu(A) = \lim_{n\to\infty} \mu(A_n)$ *(Stetigkeit von unten),*
5. $A_n \searrow A$, $\mu(A_1) < \infty \Rightarrow \mu(A) = \lim_{n\to\infty} \mu(A_n)$ *(Stetigkeit von oben).*

Hierbei bedeutet $A_n \nearrow A$, dass $A_n \subseteq A_{n+1}$ für alle $n \in \mathbb{N}$ und $\displaystyle\bigcup_{n\in\mathbb{N}} A_n = A$

und $A_n \searrow A$, dass $A_n \supseteq A_{n+1}$ für alle $n \in \mathbb{N}$ und $\displaystyle\bigcap_{n\in\mathbb{N}} A_n = A$.

Beweis. 1. und 2. folgen direkt aus der σ-Additivität von μ.

3. Seien $B_1 := A_1$ und $B_n := A_n \cap A_1^c \cap \cdots \cap A_{n-1}^c$ für alle $n > 1$. Dann sind $B_1, B_2, \ldots$ paarweise disjunkt in $\mathscr{A}$ und es gilt:

$$\sum_{i=1}^n B_i = \bigcup_{i=1}^n A_i \text{ für alle } n \in \mathbb{N},$$

woraus folgt: $\sum_{i\in\mathbb{N}} B_i = \bigcup_{i\in\mathbb{N}} A_i$. Da μ monoton ist, gilt:

$$\mu\left(\bigcup_{i\in\mathbb{N}} A_i\right) = \mu\left(\sum_{i\in\mathbb{N}} B_i\right) = \sum_{i\in\mathbb{N}} \mu(B_i) \leqslant \sum_{i\in\mathbb{N}} \mu(A_i).$$

4. Wähle $B_1, B_2, \ldots$ wie oben. Es gilt:

$$\mu(A) = \mu\left(\bigcup_{i\in\mathbb{N}} A_i\right) = \mu\left(\sum_{i\in\mathbb{N}} B_i\right) = \lim_{n\to\infty} \mu\left(\sum_{i=1}^n B_i\right)$$
$$= \lim_{n\to\infty} \mu\left(\bigcup_{i=1}^n A_i\right) = \lim_{n\to\infty} \mu(A_n).$$

5. [Übung]

Bemerkung 1.28. In Regel 1.27.5. kann auf $\mu(A_1) < \infty$ nicht verzichtet werden. [Übung]

Lemma 1.29. *Seien $\mathscr{A}_0$ eine Algebra und $\mu\colon \mathscr{A}_0 \to [0,\infty)$ eine additive Mengenfunktion. Sei μ ferner stetig von unten an allen $A \in \mathscr{A}_0$ oder stetig von oben an $\varnothing$. Dann ist μ σ-additiv.*

Beweis. Seien $(A_n)_{n\in\mathbb{N}}$ eine Folge disjunkter Mengen in $\mathscr{A}_0$ und $A := \sum_{n\in\mathbb{N}} A_n \in \mathscr{A}_0$. Definiere

$$B_n := \sum_{i=1}^{n} A_i, \quad C_n := A \backslash B_n$$

für alle $n \in \mathbb{N}$.

Sei zunächst μ stetig von unten. Da $B_n \nearrow A$, folgt:

$$\mu(A) = \lim_{n\to\infty} \mu(B_n) = \lim_{n\to\infty} \sum_{i=1}^{n} \mu(A_i) = \sum_{i\in\mathbb{N}} \mu(A_i).$$

Sei nun μ stetig von oben an $\varnothing$. Da μ additiv ist, gilt für alle $n \in \mathbb{N}$:

$$\mu(A) = \mu(C_n) + \sum_{i=1}^{n} \mu(A_i).$$

Da $C_n \searrow \varnothing$ und μ endlich ist, folgt:

$$\mu(A) = \lim_{n\to\infty} \mu(C_n) + \lim_{n\to\infty} \sum_{i=1}^{n} \mu(A_i) = \sum_{i\in\mathbb{N}} \mu(A_i).$$

Satz 1.30 (Eindeutigkeitssatz)**.** *Sei Ω eine Menge. Sei $M \subseteq \mathscr{P}(\Omega)$ ein $\cap$-stabiles Mengensystem. Definiere die σ-Algebra $\mathscr{A} := \sigma(M)$ auf Ω. Seien μ_1, μ_2 Maße auf $(\Omega, \mathscr{A})$, welche eingeschränkt auf M σ-endlich sind. Außerdem gelte*

$$\mu_1(A) = \mu_2(A) \text{ für alle } A \in M. \tag{1.1}$$

Dann ist $\mu_1 = \mu_2$ auf ganz $\mathscr{A}$.

Beweis. Sei $B \in M$ mit $\mu_1(B) = \mu_2(B) < \infty$. Das Mengensystem

$$\mathscr{D}_B := \{A \in \mathscr{A} \mid \mu_1(B \cap A) = \mu_2(B \cap A)\}$$

ist ein Dynkin-System, denn:

- $\mu_1(B \cap \Omega) = \mu_1(B) = \mu_2(B) = \mu_2(B \cap \Omega)$ impliziert, dass $\Omega \in \mathscr{D}_B$.

- Seien $D, E \in \mathscr{D}_B$ mit $D \subseteq E$. Dann ist

$$\mu_1\big((E \backslash D) \cap B\big) = \mu_1(E \cap B) - \mu_1(D \cap B) = \mu_2(E \cap B) - \mu_2(D \cap B)$$
$$= \mu_2\big((E \backslash D) \cap B\big).$$

  Folglich ist $E \backslash D \in \mathscr{D}_B$.

- Seien $D_1, D_2, \ldots \in \mathscr{D}_B$ paarweise disjunkt. Dann ist

$$\mu_1\left(\left(\sum_{n\in\mathbb{N}} D_n\right) \cap B\right) = \mu_1\left(\sum_{n\in\mathbb{N}} (D_n \cap B)\right) = \sum_{n\in\mathbb{N}} \mu_1(D_n \cap B)$$
$$= \sum_{n\in\mathbb{N}} \mu_2(D_n \cap B) = \mu_2\left(\sum_{n\in\mathbb{N}} (D_n \cap B)\right) = \mu_2\left(\left(\sum_{n\in\mathbb{N}} D_n\right) \cap B\right).$$

  Folglich ist die Vereinigung der D_n in $\mathscr{D}_B$.

Da M $\cap$-stabil ist und (1.1) gilt, folgt $M \subseteq \mathscr{D}_B \Rightarrow \delta(M) \subseteq \mathscr{D}_B$. Mit Lemma 1.18 folgt:

$$\mathscr{A} = \sigma(M) = \delta(M) \subseteq \mathscr{D}_B. \tag{1.2}$$

Da μ_1 und μ_2 σ-endlich auf M sind und wegen (1.1) gibt es eine Folge $(A_n)_{n\in\mathbb{N}}$ in M mit $A_n \nearrow \Omega$ und $\mu_1(A_n) = \mu_2(A_n) < \infty$ für alle $n \in \mathbb{N}$. Wegen (1.2) gilt $\mathscr{A} \subseteq \mathscr{D}_{A_n}$ für alle $n \in \mathbb{N}$ und mit der Stetigkeit von unten folgt

$$\mu_1(A) = \lim_{n\to\infty} \mu_1(A \cap A_n) = \lim_{n\to\infty} \mu_2(A \cap A_n) = \mu_2(A) \text{ für alle } A \in \mathscr{A}.$$

Korollar 1.31.

1. *Seien P_1, P_2 Wahrscheinlichkeitsmaße auf $\mathscr{B}^k$ mit $P_1\big((-\infty, x]\big) = P_2\big((-\infty, x]\big)$ für alle $x \in \mathbb{R}^k$. Dann ist $P_1 = P_2$.*

2. *Es gibt nur (höchstens) ein Maß λ^k auf $\mathscr{B}^k$ mit*

$$\lambda^k\big((x, y]\big) = \prod_{i=1}^{k} (y_i - x_i) \quad \textit{für alle } (x, y] \in \mathscr{I}^k.$$

Nachdem wir die Eindeutigkeit der Fortsetzung eines Maßes gewährleistet haben, interessiert uns nun die Existenz einer solchen Fortsetzung. Wir werden insbesondere die als Semiringe bezeichneten Mengensysteme kennenlernen und zeigen, dass auf Semiringen S definierte Mengenfunktionen mit bestimmten Eigenschaften sich immer zu Maßen auf den erzeugten σ-Algebren $\sigma(S)$ fortsetzen lassen.

Definition 1.32. Sei $\Omega \neq \varnothing$ eine Menge. $S \subseteq \mathscr{P}(\Omega)$ heißt *Semiring* (engl.: semiring), falls

1. $\varnothing \in S$,
2. S ist $\cap$-stabil,
3. für $A, B \in S$ mit $A \subseteq B$ gibt es paarweise disjunkte Mengen $C_1, \dots, C_n \in S$ mit $B \backslash A = \sum_{i=1}^{n} C_i$.

Beispiel 1.33.

1. Sei $\Omega := \mathbb{R}^k$. Dann ist $\mathscr{I}^k := \{(x,y] \mid x, y \in \mathbb{R}^k, x \leqslant y\}$ ein Semiring.
2. Seien $(\Omega_1, \mathscr{A}_1)$, $(\Omega_2, \mathscr{A}_2)$ Messräume. $S := \{A_1 \times A_2 \mid A_1 \in \mathscr{A}_1, A_2 \in \mathscr{A}_2\}$ ist ein Semiring über $\Omega := \Omega_1 \times \Omega_2$ (der *Semiring der Rechtecke*).

Definition 1.34. Sei $\Omega \neq \varnothing$ eine Menge. Eine Mengenfunktion $\mu^* : \mathscr{P}(\Omega) \to [0, \infty]$ heißt *äußeres Maß* (engl.: outer measure, exterior measure), falls gilt:

1. $\mu^*(\varnothing) = 0$,
2. $A \subseteq B \Rightarrow \mu^*(A) \leqslant \mu^*(B)$,
3. $\mu^*\left(\bigcup_{n \in \mathbb{N}} A_n\right) \leqslant \sum_{n \in \mathbb{N}} \mu^*(A_n)$.

Satz 1.35. *Sei $\Omega \neq \varnothing$ eine Menge. Für $M \subseteq \mathscr{P}(\Omega)$ mit $\varnothing \in M$ sei $\mu \colon M \to [0, \infty]$ mit $\mu(\varnothing) = 0$ eine Mengenfunktion. Für $A \subseteq \Omega$ sei*

$$\mu^*(A) := \inf \left\{ \sum_{n \in \mathbb{N}} \mu(A_n) \;\middle|\; (A_n)_{n \in \mathbb{N}} \textit{ Folge in } M \textit{ mit } A \subseteq \bigcup_{m \in \mathbb{N}} A_m \right\}.$$

Solche $(A_n)_{n \in \mathbb{N}}$ heißen Überdeckungsfolgen. *Falls keine Überdeckungsfolge existiert, setze $\mu^*(A) := \infty$. Dann ist μ^* ein äußeres Maß, das* von μ induzierte *äußere Maß.*

Beweis.

1. Offensichtlich gilt: $\mu^*(\varnothing) = 0$.
2. Sei $A \subseteq B \subseteq \Omega$. Dann überdeckt jede Überdeckungsfolge von B auch A. Folglich ist μ^* monoton.
3. Subadditivität: Ohne Einschränkung sei $\mu^*(A_n) < \infty$ für alle $n \in \mathbb{N}$. Dann gibt es für alle $\varepsilon > 0$ und für alle $n \in \mathbb{N}$ eine Folge $(B_{n,k})_{k \in \mathbb{N}}$ in M, sodass $A_n \subseteq \bigcup_{k \in \mathbb{N}} B_{n,k}$ und $\sum_{k \in \mathbb{N}} \mu(B_{n,k}) \leqslant \mu^*(A_n) + \varepsilon/2^n$ für alle $n \in \mathbb{N}$.
   Die Doppelfolge $(B_{n,k})_{n,k \in \mathbb{N}}$ ist eine Überdeckungsfolge in M für $\bigcup_{n \in \mathbb{N}} A_n$. Somit gilt:

$$\mu^*\left(\bigcup_{n \in \mathbb{N}} A_n\right) \leqslant \sum_{n \in \mathbb{N}} \sum_{k \in \mathbb{N}} \mu(B_{n,k}) \leqslant \sum_{n \in \mathbb{N}} \mu^*(A_n) + \varepsilon.$$

Satz 1.36 (Carathéodory). *Sei $\mu^*: \mathscr{P}(\Omega) \to [0,\infty]$ ein äußeres Maß.*

$$\mathscr{A}(\mu^*) := \{A \subseteq \Omega \mid \mu^*(A \cap E) + \mu^*(A^c \cap E) = \mu^*(E) \text{ für alle } E \subseteq \Omega\}$$

sei das System der μ^*-messbaren *Mengen. Dann gilt:*

1. *$\mathscr{A}(\mu^*)$ ist eine σ-Algebra über Ω (die σ-Algebra der μ^*-messbaren Mengen).*
2. *Die Einschränkung von μ^* auf $\mathscr{A}(\mu^*)$ ist ein Maß.*

Beweis.

1.
   - Da $E \cap \Omega = E$ und $E \cap \Omega^c = \varnothing$ für alle $E \subseteq \Omega$, ist $\Omega \in \mathscr{A}(\mu^*)$.
   - Ist $A \in \mathscr{A}(\mu^*)$, so ist wegen der Symmetrie der Gleichung
   $$\mu^*(E) = \mu^*(A \cap E) + \mu^*(A^c \cap E) \tag{1.3}$$
   in A und A^c auch $A^c \in \mathscr{A}(\mu^*)$.
   - Seien $A, B \in \mathscr{A}(\mu^*)$. Dann gilt (1.3) für alle $E \subseteq \Omega$. Da sowohl $B \cap E \subseteq \Omega$ als auch $B^c \cap E \subseteq \Omega$, können wir E in (1.3) jeweils durch diese Mengen ersetzen und erhalten folgende beiden Gleichungen:
   $$\begin{aligned}\mu^*(B \cap E) &= \mu^*(A \cap B \cap E) + \mu^*(A^c \cap B \cap E),\\ \mu^*(B^c \cap E) &= \mu^*(A \cap B^c \cap E) + \mu^*(A^c \cap B^c \cap E).\end{aligned}$$
   Diese wiederum eingesetzt in (1.3) mit B statt A ergibt:
   $$\mu^*(E) = \mu^*(A \cap B \cap E) + \mu^*(A^c \cap B \cap E) + \mu^*(A \cap B^c \cap E) + \mu^*(A^c \cap B^c \cap E).$$
   Ersetzen wir hierin E durch $E \cap (A \cup B) \subseteq \Omega$, so erhalten wir:
   $$\mu^*\big(E \cap (A \cup B)\big) = \mu^*(A \cap B \cap E) + \mu^*(A^c \cap B \cap E) + \mu^*(A \cap B^c \cap E). \tag{1.4}$$
   Zusammen mit der vorhergehenden Gleichung ergibt sich:
   $$\mu^*(E) = \mu^*\big(E \cap (A \cup B)\big) + \mu^*\big(E \cap (A \cup B)^c\big).$$
   Folglich ist $A \cup B \in \mathscr{A}(\mu^*)$.
   Seien nun $A_1', A_2', \ldots \in \mathscr{A}(\mu^*)$. Definiere die Mengen $A := \bigcup_{i \in \mathbb{N}} A_i'$ sowie $B_n := \bigcup_{i=1}^n A_i'$ für alle $n \in \mathbb{N}$. Wir definieren induktiv $A_1 := A_1' \in \mathscr{A}(\mu^*)$ und
   $$A_n := A_n' \cap \left(\bigcup_{i=1}^{n-1} A_i\right)^c = \left((A_n')^c \cup \bigcup_{i=1}^{n-1} A_i\right)^c \in \mathscr{A}(\mu^*).$$

Nach Konstruktion gilt $A = \sum_{n\in\mathbb{N}} A_n$ und $B_n = \sum_{i=1}^{n} A_i$ für alle $n \in \mathbb{N}$. Aus (1.4) ergibt sich induktiv für alle $E \subseteq \Omega$ und für alle $n \in \mathbb{N}$:

$$\mu^*(E \cap B_n) = \sum_{i=1}^{n} \mu^*(E \cap A_i).$$

Nach dem bisher Gezeigten gilt $B_n \in \mathscr{A}(\mu^*)$ für alle $n \in \mathbb{N}$ und

$$E \cap B_n^c \supseteq E \cap A^c, \quad \text{also} \quad \mu^*(E \cap B_n^c) \geqslant \mu^*(E \cap A^c)$$

für alle $E \subseteq \Omega$. Folglich gilt für alle $n \in \mathbb{N}$:

$$\mu^*(E) = \mu^*(E \cap B_n) + \mu^*(E \cap B_n^c) \geqslant \sum_{i=1}^{n} \mu^*(E \cap A_i) + \mu^*(E \cap A^c).$$

Mit der σ-Subadditivität von μ^* folgt:

$$\mu^*(E) \geqslant \sum_{i\in\mathbb{N}} \mu^*(E \cap A_i) + \mu^*(E \cap A^c) \geqslant \mu^*(E \cap A) + \mu^*(E \cap A^c).$$

Wenden wir außerdem die σ-Subadditivität auf die Mengenfolge

$$E \cap A, E \cap A^c, \varnothing, \varnothing, \ldots$$

an, so ergibt sich sogar Gleichheit:

$$\mu^*(E) = \sum_{i\in\mathbb{N}} \mu^*(E \cap A_i) + \mu^*(E \cap A^c) = \mu^*(E \cap A) + \mu^*(E \cap A^c). \tag{1.5}$$

Also ist $A \in \mathscr{A}(\mu^*)$ und $\mathscr{A}(\mu^*)$ eine σ-Algebra.

2. Wählen wir in (1.5) $E = A$, so ergibt sich:

$$\mu^*(A) = \sum_{i\in\mathbb{N}} \mu^*(A_i).$$

Folglich ist die Einschränkung des äußeren Maßes μ^* auf $\mathscr{A}(\mu^*)$ σ-additiv und somit ein Maß.

Satz 1.37 (Maßfortsetzungssatz)**.** *Seien $S \subseteq \mathscr{P}(\Omega)$ ein Semiring und $\mu\colon S \to [0,\infty]$ eine Mengenfunktion mit*

1. *$\mu(\varnothing) = 0$,*
2. *μ ist endlich additiv,*
3. *μ ist σ-subadditiv auf S, das heißt für eine Folge $(A_n)_{n\in\mathbb{N}}$ in S mit $\bigcup_{n\in\mathbb{N}} A_n \in S$ ist $\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) \leqslant \sum_{n\in\mathbb{N}} A_n$.*

Dann gibt es ein Maß $\tilde{\mu}$ auf $\sigma(S)$ mit $\mu(A) = \tilde{\mu}(A)$ für alle $A \in S$.

Beweis. Wir zeigen zunächst die Monotonie von μ. Seien $A, B \in S$ mit $A \subseteq B$. Da S ein Semiring ist, gibt es disjunkte Mengen $C_1, \dots, C_n \in S$ mit $\sum_{i=1}^n C_i = B \backslash A$. Es folgt:

$$B = A + \sum_{i=1}^{n} C_i.$$

Da μ additiv und nichtnegativ ist, gilt:

$$\mu(B) = \mu(A) + \sum_{i=1}^{n} \mu(C_i) \geqslant \mu(A).$$

Sei nun μ^* das von μ induzierte Maß und $\mathscr{A}(\mu^*)$ die σ-Algebra der μ^*-messbaren Mengen (siehe Satz 1.36). Wir zeigen nun, dass $S \subseteq \mathscr{A}(\mu^*)$ ist. Seien $A \in S$ und $E \subseteq \Omega$. μ^* ist aufgefasst als äußeres Maß σ-subadditiv und daher gilt:

$$\mu^*(E) \leqslant \mu^*(A \cap E) + \mu^*(A^c \cap E).$$

Außerdem gibt es nach Definition von μ^* für alle $\varepsilon > 0$ eine Folge $(A_n)_{n \in \mathbb{N}}$ in S mit $E \subseteq \bigcup_{n \in \mathbb{N}} A_n$ und

$$\sum_{n \in \mathbb{N}} \mu(A_n) \leqslant \mu^*(E) + \varepsilon. \tag{1.6}$$

Da S ein Semiring ist, gilt: $B_n := A \cap A_n \in S$ für alle $n \in \mathbb{N}$. Außerdem gibt es für jedes $n \in \mathbb{N}$ paarweise disjunkte Mengen $C_{n,1}, \dots, C_{n,m_n} \in S$ mit

$$A_n \cap A^c = A_n \backslash B_n = \sum_{i=1}^{m_n} C_{n,i}.$$

Für alle $n \in \mathbb{N}$ folgt:

$$A_n = B_n + \sum_{i=1}^{m_n} C_{n,i}. \tag{1.7}$$

Es gilt:

$$A \cap E \subseteq \bigcup_{n \in \mathbb{N}} B_n,$$
$$A^c \cap E \subseteq \bigcup_{n \in \mathbb{N}} \sum_{i=1}^{m_n} C_{n,i}.$$

Mit der σ-Subadditivität von μ^* und der Additivität von μ und (1.7) sowie (1.6) folgt:

$$\begin{aligned}\mu^*(A \cap E) + \mu^*(A^c \cap E) &\leqslant \sum_{n \in \mathbb{N}} \mu(B_n) + \sum_{n \in \mathbb{N}} \sum_{i=1}^{m_n} \mu(C_{n,i}) \\ &= \sum_{n \in \mathbb{N}} \left(\mu(B_n) + \sum_{i=1}^{m_n} \mu(C_{n,i}) \right) = \sum_{n \in \mathbb{N}} \mu(A_n) \leqslant \mu^*(E) + \varepsilon.\end{aligned}$$

Folglich gilt:

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

und somit $S \subseteq \mathscr{A}(\mu^*)$. Da $\mathscr{A}(\mu^*)$ eine σ-Algebra ist, gilt: $\sigma(S) \subseteq \mathscr{A}(\mu^*)$.

Definiere $\tilde{\mu} := \mu^*|_{\sigma(S)}$. Es bleibt zu zeigen:

$$\tilde{\mu}(A) = \mu(A)$$

für alle $A \in S$. Nach Definition von μ^* gilt:

$$\tilde{\mu}(A) = \mu^*(A) \leqslant \mu(A)$$

für alle $A \in S$. Außerdem gilt für alle Folgen $(A_n)_{n\in\mathbb{N}}$ in S mit $A \subseteq \bigcup_{n\in\mathbb{N}} A_n$ wegen der Monotonie und der σ-Subadditivität von μ:

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

Es folgt:

$$\mu(A) \leqslant \inf\left\{\sum_{n\in\mathbb{N}} \mu(A_n) \;\middle|\; A_n \in S \text{ für alle } n \in \mathbb{N} \text{ und } A \subseteq \bigcup_{n\in N} A_n\right\} = \mu^*(A) = \tilde{\mu}(A).$$

Somit folgt: $\tilde{\mu}(A) = \mu(A)$ für alle $A \in S$.

Bemerkung 1.38. Nach Satz 1.30 ist die Fortsetzung auf $\sigma(S)$ eindeutig, falls das Maß μ σ-endlich ist.

Abschließend wollen wir noch einen wichtigen Begriff einführen.

Definition 1.39. Sei $(\Omega, \mathscr{A}, \mu)$ ein Maßraum. Wir bezeichnen eine Menge $S \in \mathscr{A}$ als *Träger* (engl.: support) von μ, falls $\mu(S^c) = 0$. Ist (Ω, d) ein metrischer Raum und $\mathscr{A}$ die Borel-σ-Algebra (siehe Bemerkung 1.23), so heißt S *topologischer Träger* (engl.: topological support) von μ, falls S die kleinste abgeschlossene Menge mit der obigen Eigenschaft ist, das heißt, falls S abgeschlossen ist, $\mu(S^c) = 0$ gilt und für alle abgeschlossenen Teilmengen $T \subseteq S$ aus $\mu(T^c) = 0$ stets $T = S$ folgt. Hierbei sollte beachtet werden, dass nach Definition der Borel-σ-Algebra die offene Menge T^c stets messbar ist.

Satz 1.40. *Seien (Ω, d) ein separabler metrischer Raum, ausgestattet mit der Borel-σ-Algebra, und μ ein Maß auf diesem Raum. Dann existiert ein eindeutiger topologischer Träger von μ.* [Übung]

1.2 Maße auf $\mathscr{B}^k$

Im Folgenden werden wir uns mit Maßen auf der Borel-σ-Algebra $\mathscr{B}^k$ beschäftigen und Verteilungen kennenlernen. Wir beginnen mit dem eindimensionalen Fall $k = 1$ und werden später auf beliebige $k \in \mathbb{N}$ verallgemeinern.

Definition 1.41. Eine monoton wachsende, rechtsstetige Funktion $G\colon \mathbb{R} \to \mathbb{R}$ heißt *maßdefinierend*. G heißt *Verteilungsfunktion* (engl.: distribution function), wenn außerdem gilt:

$$\lim_{x\to\infty} G(x) = 1 \quad \text{und} \quad \lim_{x\to-\infty} G(x) = 0.$$

Notation 1.42. Sei $A \subseteq \mathbb{R}^k$ eine Menge. Wir bezeichnen den Abschluss von A in $\mathbb{R}^k$ mit $\overline{A}$ und das Innere von A mit A°.

Satz 1.43. *Sei G eine maßdefinierende Funktion. Dann existiert genau ein Maß μ_G auf $\mathscr{B}^1$ mit*

$$\mu_G((a,b]) = G(b) - G(a) \qquad \textit{für alle } (a,b] \in \mathscr{I}^1. \tag{1.8}$$

Falls G eine Verteilungsfunktion ist, so ist μ_G ein Wahrscheinlichkeitsmaß und heißt das Lebesgue-Stieltjes-Maß *zu G.*

Beweis. Da G monoton wachsend ist, definiert Gleichung (1.8) auf dem Semiring $\mathscr{I}^1$ eine nichtnegative, endlich-additive Mengenfunktion mit $\mu_G(\varnothing) = 0$. Es bleibt, zu zeigen, dass μ_G σ-subadditiv auf $\mathscr{I}^1$ ist.
Sei $(A_n)_{n\in\mathbb{N}}$ mit $A_n := (x_n, y_n]$ für alle $n \in \mathbb{N}$ eine Folge in $\mathscr{I}^1$ mit

$$A := (x,y] := \bigcup_{n\in\mathbb{N}} A_n \in \mathscr{I}^1.$$

Da G rechtsstetig ist, gibt es für alle $\varepsilon > 0$ ein $\delta > 0$ mit $\delta < y - x$, sodass

$$0 \leqslant \mu_G((x, x+\delta]) = G(x+\delta) - G(x) \leqslant \frac{\varepsilon}{2}.$$

Sei $A' := (x+\delta, y]$. Es gilt:

$$\mu_G(A) \leqslant \mu_G(A') + \frac{\varepsilon}{2}. \tag{1.9}$$

Analog gibt es für alle $n \in \mathbb{N}$ ein $\delta_n > 0$, sodass mit $A'_n := (x_n, y_n + \delta_n]$ gilt:

$$\mu_G(A'_n) \leqslant \mu_G(A_n) + \frac{\varepsilon}{2^{n+1}}. \tag{1.10}$$

Da $\{(x_n, y_n + \delta_n)\}_{n\in\mathbb{N}}$ eine offene Überdeckung des kompakten Intervalls $[x+\delta, y]$ ist, gibt es nach dem Satz von Heine-Borel ein $m \in \mathbb{N}$, sodass gilt:

$$A' \subseteq [x+\delta, y] \subseteq \bigcup_{n=1}^{m} {A'_n}^\circ.$$

Damit können wir die endliche Subadditivität von μ_G ausnutzen und mit der Monotonie von G ergibt sich:

$$\mu_G(A') \leqslant \mu_G\left(\bigcup_{n=1}^{m} A'_n\right) \leqslant \sum_{n=1}^{m} \mu_G(A'_n) \overset{(1.10)}{\leqslant} \sum_{n\in\mathbb{N}} \mu_G(A_n) + \frac{\varepsilon}{2}.$$

Lässt man ε gegen null gehen, folgt mit (1.9)

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

Da $\mathscr{I}^1$ $\cap$-stabil ist und μ_G eingeschränkt auf I^1 σ-endlich ist, folgt die Eindeutigkeit von μ_G aus Satz 1.30.

Satz 1.44. *Sei P ein Wahrscheinlichkeitsmaß. Dann ist $F(x) := P\big((-\infty, x]\big)$ eine Verteilungsfunktion. Zu jeder Verteilungsfunktion F existiert genau ein Wahrscheinlichkeitsmaß auf $\mathscr{B}^1$ mit $P\big((a,b]\big) = F(b) - F(a)$ und umgekehrt.*

Beweis. Betrachte die Abbildung

$$\begin{aligned} \psi\colon \{F \mid F \text{ Verteilungsfunktion}\} &\to \big\{P\colon \mathscr{B}^1 \to [0,1] \,\big|\, P \text{ Wahrscheinlichkeitsmaß}\big\} \\ F &\mapsto P_F, \end{aligned}$$

wobei P_F das zu F gehörige Maß nach Satz 1.43 sei. Beachte, dass P_F ein Wahrscheinlichkeitsmaß ist, da F eine Verteilungsfunktion ist. Des Weiteren definiere die Abbildung

$$\begin{aligned} \phi\colon \big\{P\colon \mathscr{B}^1 \to [0,1] \,\big|\, P \text{ Wahrscheinlichkeitsmaß}\big\} &\to \{F \mid F \text{ Verteilungsfunktion}\} \\ P &\mapsto F_P, \end{aligned}$$

wobei

$$F_P\colon \mathbb{R} \to [0,1], \quad x \mapsto P\big((-\infty, x]\big).$$

Da P ein Wahrscheinlichkeitsmaß ist, ist dies in der Tat eine Verteilungsfunktion:

- Seien $a, b \in \mathbb{R}$ mit $a \leqslant b$. Dann gilt:

  $$\begin{aligned} F_P(b) - F_P(a) &= P\big((-\infty, b]\big) - P\big((-\infty, a]\big) \\ &= P\big((-\infty, b] \setminus (-\infty, a]\big) = P\big((a,b]\big) \geqslant 0. \end{aligned}$$

  Es folgt also $F_P(a) \leqslant F_P(b)$. Somit ist F_P monoton steigend.

- Seien $x \in \mathbb{R}$ und $(x_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{R}$ mit $x_n \searrow x$. Da P endlich ist, folgt mit der Stetigkeit von oben (Regeln 1.27):

  $$\begin{aligned} \lim_{n\to\infty} F_P(x_n) = \lim_{n\to\infty} P\big((-\infty, x_n]\big) &= P\left(\bigcap_{n\in\mathbb{N}} (-\infty, x_n]\right) \\ &= P\big((-\infty, x]\big) = F_P(x). \end{aligned}$$

  Folglich ist F_P rechtsstetig.

- Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{R}$ mit $\lim_{n\to\infty} x_n = \infty$. Für alle $n \in \mathbb{N}$ definieren wir $y_n := \min\{x_k \mid k \geqslant n\}$. Dann gilt $y_n \leqslant x_n$ für alle $n \in \mathbb{N}$, $y_n \nearrow \infty$ und mit der Stetigkeit von unten von P sowie der Monotonie von F_P folgt:

$$\begin{aligned} 1 \geqslant \lim_{n\to\infty} F_P(x_n) \geqslant \lim_{n\to\infty} F_P(y_n) = \lim_{n\to\infty} P\big((-\infty, y_n]\big) \\ = P\left(\bigcup_{n\in\mathbb{N}} (-\infty, y_n]\right) = P(\mathbb{R}) = 1. \end{aligned}$$

  Somit gilt: $\lim_{x\to\infty} F_P(x) = 1$.

- Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{R}$ mit $\lim_{n\to\infty} x_n = -\infty$. Für alle $n \in \mathbb{N}$ definiere $y_n := \max\{x_k \mid k \geqslant n\}$. Dann gilt $y_n \geqslant x_n$ für alle $n \in \mathbb{N}$, $y_n \searrow -\infty$ und mit der Stetigkeit von oben von P sowie der Monotonie von F_P folgt:

$$\begin{aligned} 0 \leqslant \lim_{n\to\infty} F_P(x_n) \leqslant \lim_{n\to\infty} F_P(y_n) = \lim_{n\to\infty} P\big((-\infty, y_n]\big) \\ = P\left(\bigcap_{n\in\mathbb{N}} (-\infty, y_n]\right) = P(\varnothing) = 0. \end{aligned}$$

  Somit gilt: $\lim_{x\to-\infty} F_P(x) = 0$.

Um die Aussage des Satzes zu beweisen, reicht es zu zeigen, dass ψ und ϕ zueinander invers sind. Sei Q ein Wahrscheinlichkeitsmaß auf $\mathscr{B}^1$. Für alle $(a,b] \in \mathscr{I}^1$ gilt:

$$\begin{aligned} \psi\big(\phi(Q)\big)\big((a,b]\big) &= P_{F_Q}\big((a,b]\big) = F_Q(b) - F_Q(a) \\ &= Q\big((-\infty, b]\big) - Q\big((-\infty, a]\big) = Q\big((a,b]\big). \end{aligned}$$

Nach Bemerkung 1.38 ist somit $\psi\big(\phi(Q)\big) = Q$.

Sei nun G eine Verteilungsfunktion. Für alle $x \in \mathbb{R}$ gilt:

$$\phi\big(\psi(G)\big)(x) = F_{P_G}(x) = P_G\big((-\infty, x]\big) = G(x) - \lim_{y\to-\infty} G(y) = G(x).$$

Folglich ist $\phi\big(\psi(G)\big) = G$. Somit sind ψ und ϕ zueinander invers.

Bemerkung 1.45.

1. Das zu $G(x) = x$ gehörige Maß $\lambda^1((x,y]) := y - x$ heißt *Borel-Lebesgue-Maß.*

2. Sei G definiert wie in Satz 1.43. Dann gilt $\mu_G(\{x\}) = G(x) - G(x-)$, wobei $G(x-) := \lim_{y\to x, y<x} G(y)$ der linksseitige Limes ist. Folglich ist G genau dann stetig, wenn μ_G stetig ist.

3. μ_G in Satz 1.43 ist ein *Borel-Maß*, das heißt ein Maß auf der Borel-σ-Algebra, mit $\mu_G(K) < \infty$ für alle kompakten Teilmengen K. Ferner ist μ_G σ-endlich.

Beispiel 1.46.

1. Die Verteilungsfunktion der *Normalverteilung* mit Parameter $(\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}_+$ ist
$$G(x) := \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x} \exp\left(-\frac{1}{2\sigma^2}(t-\mu)^2\right) \mathrm{d}t. \quad \text{[Übung]}$$

2. Sei $\lambda|_{[a,b]}$ das standardisierte Lebesgue-Maß auf $[a, b]$. Die zugehörige Verteilungsfunktion ist die *Gleichverteilung* auf dem Intervall $[a, b]$:
$$G(x) := \begin{cases} 0 & x \leqslant a \\ (x-a)/(b-a) & a < x < b \\ 1 & x > b \end{cases}.$$

3. Die Verteilungsfunktion des Dirac-Maßes δ_{x_0} bei x_0 ist
$$G(x) := \begin{cases} 0 & x < x_0 \\ 1 & x \geqslant x_0 \end{cases}.$$

4. Die Verteilungsfunktion der *Binomialverteilung* mit Parametern $p \in [0, 1]$ und $n \in \mathbb{N}$ ist
$$G(x) := \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \mathbb{1}_{(-\infty,x)}(k).$$
Das zugehörige Maß ist
$$B_{n,p} := \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \delta_k.$$
Für $n = 1$ ergibt sich als Spezialfall die Bernoulli-Verteilung.

5. Die Verteilungsfunktion der *Bernoulli-Verteilung* mit Parameter p ist
$$G(x) := \begin{cases} 0 & x < 0 \\ 1-p & 0 \leqslant x < 1 \\ 1 & x \geqslant 1 \end{cases}.$$
Das zugehörige Maß ist $B(p) := B_{1,p} = (1-p)\delta_0 + p\delta_1$.

6. Die Verteilungsfunktion der *hypergeometrischen Verteilung* mit Parametern $N, M, n \in \mathbb{N}$, wobei $M \leqslant N$ und $n \leqslant N$, ist
$$G(x) := \sum_{k=0}^{n} \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}} \mathbb{1}_{(-\infty,x)}(k).$$

7. Die Verteilungsfunktion der *Gammaverteilung* mit Parametern $p, b > 0$ ist
$$G(x) := \int_{-\infty}^{x} \frac{b^p}{\Gamma(p)} t^{p-1} \mathrm{e}^{-bt} \mathbb{1}_{[0,\infty)}(t)\,\mathrm{d}t,$$
wobei
$$\Gamma(x) := \int_0^\infty t^{x-1}\mathrm{e}^{-t}\,\mathrm{d}t$$
die *Gammafunktion* ist.

8. Die Verteilungsfunktion der *Betaverteilung* mit Parametern $p, q > 0$ ist
$$G(x) := \int_{-\infty}^{x} \frac{1}{B(p,q)} t^{p-1}(1-t)^{q-1} \mathbb{1}_{[0,1)}(t)\,\mathrm{d}t,$$
wobei
$$B(p,q) := \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)} = \int_0^1 t^{p-1}(1-t)^{q-1}\,\mathrm{d}t$$
die *Betafunktion* ist.

9. Die Verteilungsfunktion der *Cauchy-Verteilung* mit Parameter $\lambda > 0$ ist
$$G(x) := \int_{-\infty}^{x} \frac{1}{\lambda\pi} \frac{1}{1+\left(\frac{t}{\lambda}\right)^2}\,\mathrm{d}t. \quad \text{[Übung]}$$

10. Die *empirische Verteilungsfunktion* zu n Beobachtungen $x_1, \ldots, x_n$ ist
$$F_n(t) := \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{(-\infty,t]}(x_i).$$
Das zugehörige Maß heißt *empirisches Maß* und ist definiert als
$$\mu_n(A) := \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_A(x_i).$$

Bemerkung 1.47. Die Beobachtungen $x_1, \ldots, x_n$ sind Werte von Zufallsvariablen X_i (siehe Definition 2.17). Folglich ist auch die empirische Verteilungsfunktion F_n eine Zufallsgröße, da sie von den Zufallsgrößen x_i abhängt. Es handelt sich also um eine „zufällige“ Verteilungsfunktion. Im nächsten Kapitel erfahren wir mehr über Zufallsvariablen.

Im Folgenden werden wir uns mit der Konstruktion von Maßen auf $\mathscr{B}^k$ in höheren Dimensionen beschäftigen.

Definition 1.48. $G\colon \mathbb{R}^k \to \mathbb{R}$ heißt *maßdefinierende Funktion*, falls gilt:

1. G ist rechtsstetig, das heißt für alle $x = (x^1, \ldots, x^k) \in \mathbb{R}^k$ existiert für jede Folge $(x_n)_{n\in\mathbb{N}}$ mit $x_n = (x_n^1, \ldots, x_n^k) \in \mathbb{R}^k$ und $x_n^j \searrow x^j$ für alle $j \in \{1, \ldots, k\}$ der Grenzwert $G(x) = \lim_{n\to\infty} G(x_n)$.

2. Verallgemeinerte Monotonie: Für alle $a, b \in \mathbb{R}^k$ mit $a \leqslant b$ gilt:

$$\Delta_a^b G := \sum_{\varepsilon \in \{0,1\}^k} (-1)^{k-\varepsilon_1-\cdots-\varepsilon_k} G\left(b_1^{\varepsilon_1} a_1^{1-\varepsilon_1}, \ldots, b_k^{\varepsilon_k} a_k^{1-\varepsilon_k}\right) \geqslant 0.$$

3. G heißt *k-dimensionale Verteilungsfunktion*, falls außerdem gilt:

$$\lim_{x_i \to -\infty} G(x) = 0 \text{ für alle } i \in \{1, \ldots, k\} \text{ und } \lim_{x_1 \to \infty, \ldots, x_k \to \infty} G(x) = 1.$$

Satz 1.49. *Sei G eine maßdefinierende Funktion. Dann gibt es genau ein Maß μ_G auf $\mathscr{B}^k$ mit*

$$\mu_G\big((a, b]\big) = \Delta_a^b G$$

für alle $(a, b] \in \mathscr{I}^k$. Dieses heißt Lebesgue-Stieltjes-Maß *zu G.*

Beweis. Wir zeigen, dass die obige Vorschrift eine σ-additive Mengenfunktion auf $\mathscr{I}^k$ definiert. Sei zunächst $\varnothing = (a, b] \in \mathscr{I}^k$, das heißt es gibt ein $i \in \{1, \ldots, k\}$ mit $a_i = b_i$. Dann gilt:

$$\begin{aligned}
\Delta_a^b G = & \sum_{\varepsilon \in \{0,1\}^k} (-1)^{k-\varepsilon_1-\cdots-\varepsilon_k} G\left(b_1^{\varepsilon_1} a_1^{1-\varepsilon_1}, \ldots, a_i, \ldots, b_k^{\varepsilon_k} a_k^{1-\varepsilon_k}\right) \\
= & \sum_{\varepsilon \in \{0,1\}^k, \varepsilon_i = 0} (-1)^{k-\varepsilon_1-\cdots-\varepsilon_{i-1}-\varepsilon_{i+1}-\cdots-\varepsilon_k} G\left(b_1^{\varepsilon_1} a_1^{1-\varepsilon_1}, \ldots, a_i, \ldots, b_k^{\varepsilon_k} a_k^{1-\varepsilon_k}\right) \\
& - \sum_{\varepsilon \in \{0,1\}^k, \varepsilon_i = 1} (-1)^{k-\varepsilon_1-\cdots-\varepsilon_{i-1}-\varepsilon_{i+1}-\cdots-\varepsilon_k} G\left(b_1^{\varepsilon_1} a_1^{1-\varepsilon_1}, \ldots, a_i, \ldots, b_k^{\varepsilon_k} a_k^{1-\varepsilon_k}\right) \\
= & 0.
\end{aligned}$$

Somit definiert $\Delta_a^b G$ eine Mengenfunktion.

Wir zeigen nun die Additivität von μ_G auf $\mathscr{I}^k$. Hierfür zeigen wir die Additivität zunächst für spezielle, sogenannte reguläre, Zerlegungen von Quadern. Eine *reguläre Zerlegung* von $(a, b]$ erhält man durch Zerlegen der Kanten $K_i := (a_i, b_i]$ in n_i Teilintervalle $T_{i,j} := (t_{i,j-1}, t_{i,j}]$ für $j \in \{1, \ldots, n_i\}$ mit $a_i = t_{i,0} < t_{i,1} < \cdots < t_{i,n_i} = b_i$. Dies liefert uns eine Zerlegung des Quaders $(a, b]$ in $\prod_{i=1}^k n_i$ Teilquader der Form

$$B_{j_1, \ldots, j_k} := \bigtimes_{i=1}^k T_{i, j_i}$$

mit $1 \leqslant j_i \leqslant n_i$ für alle $i \in \{1, \ldots, k\}$. Sei $(a, b]$ regulär zerlegt. Es ist zu zeigen:

$$\mu_G\big((a, b]\big) = \sum_{j_1, \ldots, j_k} \mu_G(B_{j_1, \ldots, j_k}).$$

Es gilt:

$$\begin{aligned}
&\sum_{j_1,\dots,j_k} \mu_G(B_{j_1,\dots,j_k}) \\
&= \sum_{j_1,\dots,j_k} \sum_{\varepsilon\in\{0,1\}^k} (-1)^{k-\varepsilon_1-\dots-\varepsilon_k} G\left(t_{1,j_1-1}^{\varepsilon_1} t_{1,j_1}^{1-\varepsilon_1}, \dots, t_{k,j_k-1}^{\varepsilon_k} t_{k,j_k}^{1-\varepsilon_k}\right) \\
&= \sum_{\varepsilon\in\{0,1\}^k} (-1)^{k-\varepsilon_1-\dots-\varepsilon_k} \sum_{j_1,\dots,j_k} G\left(t_{1,j_1-1}^{\varepsilon_1} t_{1,j_1}^{1-\varepsilon_1}, \dots, t_{k,j_k-1}^{\varepsilon_k} t_{k,j_k}^{1-\varepsilon_k}\right).
\end{aligned}$$

Betrachte ein $c \in \{1,\dots,k\}$, ein $\varepsilon_c \in \{0,1\}$ und ein $m_c \in \{1,\dots,n_c\}$, sodass

$$\left(t_{1,j_1-1}^{\varepsilon_1} t_{1,j_1}^{1-\varepsilon_1}, \dots, t_{c,m_c-1}^{\varepsilon_c} t_{c,m_c}^{1-\varepsilon_c}, \dots, t_{k,j_k-1}^{\varepsilon_k} t_{k,j_k}^{1-\varepsilon_k}\right)$$

für alle $j_i \in \{1,\dots,n_i\}$ mit $i \neq c$ kein Eckpunkt des Quaders $(a,b]$ ist. Das heißt, es gilt:

$$t_{c,j_c-1}^{\varepsilon_c} t_{c,j_c}^{1-\varepsilon_c} \notin \{a_c, b_c\}.$$

Ohne Einschränkung seien $c = 1$, $\varepsilon_1 = 0$ und $t_{1,m_1} \neq b_1$. Folglich ist $m_1 < n_1$. In der obigen Summe entspricht dies einem Summanden

$$\sum_{\varepsilon\in\{0,1\}^k, \varepsilon_1=0} (-1)^{k-\varepsilon_2-\dots-\varepsilon_k} \sum_{j_2,\dots,j_k} G\left(t_{1,m_1}, t_{2,j_2-1}^{\varepsilon_2} t_{2,j_2}^{1-\varepsilon_2}, \dots, t_{k,j_k-1}^{\varepsilon_k} t_{k,j_k}^{1-\varepsilon_k}\right).$$

Wegen $m_1 + 1 \leqslant n_1$ erhalten wir außerdem einen Summanden:

$$-\sum_{\varepsilon\in\{0,1\}^k, \varepsilon_1=1} (-1)^{k-\varepsilon_2-\dots-\varepsilon_k} \sum_{j_2,\dots,j_k} G\left(t_{1,(m_1+1)-1}, t_{2,j_2-1}^{\varepsilon_2} t_{2,j_2}^{1-\varepsilon_2}, \dots, t_{k,j_k-1}^{\varepsilon_k} t_{k,j_k}^{1-\varepsilon_k}\right).$$

Je zwei solcher Summanden heben sich gegenseitig weg. Da dieser Effekt immer dann auftritt, wenn für ein $\varepsilon \in \{0,1\}^k$ der Punkt

$$\left(t_{1,j_1-1}^{\varepsilon_1} t_{1,j_1}^{1-\varepsilon_1}, \dots, t_{k,j_k-1}^{\varepsilon_k} t_{k,j_k}^{1-\varepsilon_k}\right)$$

kein Eckpunkt des Quaders $(a,b]$ ist, bleiben nur die Summanden, welche Eckpunkte beinhalten, übrig. Wir erhalten:

$$\begin{aligned}
\sum_{j_1,\dots,j_k} \mu_G(B_{j_1,\dots,j_k}) &= \sum_{\varepsilon\in\{0,1\}^k} (-1)^{k-\varepsilon_1-\dots-\varepsilon_k} \sum_{j_1,\dots,j_k} G\left(t_{1,j_1-1}^{\varepsilon_1} t_{1,j_1}^{1-\varepsilon_1}, \dots, t_{k,j_k-1}^{\varepsilon_k} t_{k,j_k}^{1-\varepsilon_k}\right) \\
&= \sum_{\varepsilon\in\{0,1\}^k} (-1)^{k-\varepsilon_1-\dots-\varepsilon_k} G\left(a_1^{\varepsilon_1} b_1^{1-\varepsilon_1}, \dots, a_k^{\varepsilon_k} b_k^{1-\varepsilon_k}\right) \\
&= \mu_G\big((a,b]\big).
\end{aligned}$$

Sei nun $(a,b] = \bigcup_{p=1}^n (a^p, b^p]$ eine Zerlegung in Quader $A_p = \times_{i=1}^k I_{i,p}$. Für jeden Index $i \in \{1,\dots,k\}$ definiere $I_i := \bigcup_{p=1}^n I_{i,p}$. Wir können paarweise disjunkte Teilintervalle $T_{i,1},\dots,T_{i,n_i}$ der $I_{i,p}$ auswählen, sodass die Menge der Randpunkte dieser

Teilintervalle der Menge der Randpunkte der $I_{i,p}$ entspricht und dass gilt:

$$I_i = \bigcup_{p=1}^{n_i} T_{i,p}.$$

Die Quader $B_{j_1,\ldots,j_k} = \bigtimes_{i=1}^{k} T_{i,j_i}$ bilden somit eine reguläre Zerlegung von $(a,b]$. Außerdem gilt für alle $p \in \{1,\ldots,n\}$:

$$A_p = \bigcup_{B_{j_1,\ldots,j_k} \subseteq A_p} B_{j_1,\ldots,j_k}.$$

Somit wird auch jedes A_p regulär zerlegt und es gilt:

$$\mu_G\big((a,b]\big) = \sum_{j_1,\ldots,j_k} \mu_G(B_{j_1,\ldots,j_k}) = \sum_{p=1}^{n} \sum_{B_{j_1,\ldots,j_k} \subseteq A_p} \mu_G(B_{j_1,\ldots,j_k}) = \sum_{p=1}^{n} \mu_G(A_p).$$

Folglich ist μ_G additiv auf $\mathscr{I}^k$. Insbesondere folgt aus $(a,b] \subseteq \bigcup_{p=1}^{n} A_p$ mit $A_p \in \mathscr{I}^k$ und $n \in \mathbb{N}$, dass gilt:

$$\mu_G\big((a,b]\big) \leqslant \sum_{p=1}^{n} \mu_G(A_p),$$

weil wir durch Ergänzen von $(a,b]$ um Quader Q_q, $q \in \{1,\ldots,r\}$ die Gleichheit

$$(a,b] \cup \bigcup_{q=1}^{r} Q_q = \bigcup_{p=1}^{n} A_p$$

erreichen und die Additivität von μ_G ausnutzen können:

$$\begin{aligned}\mu_G\big((a,b]\big) \leqslant \mu_G\big((a,b]\big) + \sum_{q=1}^{r} \mu_G(Q_q) &= \mu_G\left((a,b] \cup \bigcup_{q=1}^{r} Q_q\right)\\ &= \mu_G\left(\bigcup_{p=1}^{n} A_p\right) = \sum_{p=1}^{n} \mu_G(A_p). \qquad (1.11)\end{aligned}$$

Wir zeigen nun, dass μ_G auf $\mathscr{I}^k$ σ-subadditiv ist. Sei $(a,b] \subseteq \bigcup_{p=1}^{\infty}(a^p,b^p]$. Es ist zu zeigen:

$$\mu_G\big((a,b]\big) \leqslant \sum_{p=1}^{\infty} \mu_G\big((a^p,b^p]\big).$$

Sei $\varepsilon > 0$. Für $\delta > 0$ definiere

$$B := \left\{x \in \mathbb{R}^k \;\middle|\; x_i \in (a_i+\delta, b_i] \text{ für alle } i \in \{1,\ldots,k\}\right\}.$$

Da G rechtsstetig ist, können wir δ so klein wählen, dass $\mu_G(B) > \mu_G\big((a,b]\big) - \varepsilon$ gilt. Beachte, dass gilt:

$$\overline{B} = \left\{x \in \mathbb{R}^k \;\middle|\; x_i \in [a_i+\delta, b_i] \text{ für alle } i \in \{1,\ldots,k\}\right\} \subseteq (a,b].$$

Ferner erhalten wir für jedes $p \in \mathbb{N}$ eine Menge

$$B_p := \left\{ x \in \mathbb{R}^k \,\middle|\, x_i \in (a_i^p, b_i^p + \delta_p] \text{ für alle } i \in \{1, \ldots, k\} \right\}$$

mit $\delta_p > 0$ und

$$\mu_G(B_p) < \mu_G\big((a^p, b^p]\big) + \frac{\varepsilon}{2^p}.$$

Beachte, dass gilt:

$$(a^p, b^p] \subseteq B_p^\circ = \left\{ x \in \mathbb{R}^k \,\middle|\, x_i \in (a_i^p, b_i^p + \delta_p) \text{ für alle } i \in \{1, \ldots, k\} \right\}.$$

Somit gilt:

$$\overline{B} \subseteq (a, b] \subseteq \bigcup_{p=1}^{\infty} (a^p, b^p] \subseteq \bigcup_{p=1}^{\infty} B_p^\circ.$$

Nach dem Satz von Heine-Borel ist $\overline{B}$ kompakt. Somit gibt es ein $N \in \mathbb{N}$ mit

$$B \subseteq \overline{B} \subseteq \bigcup_{p=1}^{N} B_p^\circ \subseteq \bigcup_{p=1}^{N} B_p.$$

Mit (1.11) erhalten wir:

$$\mu_G\big((a, b]\big) - \varepsilon < \mu_G(B) \leqslant \sum_{p=1}^{N} \mu_G(B_p) < \sum_{p=1}^{\infty} \mu_G\big((a^p, b^p]\big).$$

Da dies für alle $\varepsilon > 0$ gilt, folgt die σ-Subadditivität von μ_G auf $\mathscr{I}^k$. Da $\mathscr{I}^k$ ein Semiring ist und $\mu_G\big((a, b]\big) < \infty$ für beschränkte Quader $(a, b]$ gilt, gibt es nach dem Maßfortsetzungssatz (Satz 1.37) genau ein Maß auf $\mathscr{B}^k$, welches μ_G fortsetzt.

Das Rechnen mit verallgemeinerter Monotonie erfordert oft viel Kombinatorik. Das folgende Lemma soll das Arbeiten mit verallgemeinerter Monotonie vereinfachen.

Lemma 1.50. *Sei*

$$G\colon \left\{ (a, b] \subseteq \mathbb{R}^k \,\middle|\, a \in \big(\mathbb{R} \cup \{-\infty\}\big)^k, b \in \mathbb{R}^k \right\} \to \mathbb{R}$$

eine endlich-additive Mengenfunktion. Definiere die Funktion

$$F_G\colon \mathbb{R} \to \mathbb{R}, \quad x \mapsto G\big((-\infty, x]\big).$$

Dann gilt für alle $a, b \in \mathbb{R}^k$ mit $a \leqslant b$:

$$\Delta_a^b F_G = G\big((a, b]\big).$$

Beweis. Zunächst betrachten wir eine für die Rechnung günstige Zerlegung der auftretenden Mengen. Definiere zu $\varepsilon \in \{0,1\}^k$ die Menge:

$$R_\varepsilon := \bigtimes_{i=1}^{k} (-\infty, b_i^\varepsilon a_i^{1-\varepsilon}] \subseteq (-\infty, b].$$

Außerdem definiere zu $\delta \in \{0,1\}^k$ und $i \in \{1,\ldots,k\}$ das Intervall:

$$z_{\delta,i} := \begin{cases} (-\infty, a_i] & \text{falls } \delta_i = 0 \\ (a_i, b_i] & \text{falls } \delta_i = 1 \end{cases}.$$

Nun können wir die Quader, in die wir R_ε zerlegen wollen, definieren:

$$Z_\delta := \bigtimes_{i=1}^{k} z_{\delta,i}.$$

Wir zeigen nun:

$$R_\varepsilon = \sum_{\delta \leqslant \varepsilon} Z_\delta.$$

Sei zunächst $x \in R_\varepsilon$. Dann gilt für alle $i \in \{1,\ldots,k\}$: $x_i \in (-\infty, b_i^{\varepsilon_i} a_i^{1-\varepsilon_i}]$. Also gilt für alle $i \in \{1,\ldots,k\}$ entweder $x_i \in (-\infty, a_i]$ oder $x_i \in (a_i, b_i]$. Folglich gibt es ein $\delta \in \{0,1\}^k$ mit $x \in Z_\delta$. Da aus $\delta_i = 1$ $x_i \in (a_i, b_i]$ und somit $\varepsilon_i = 1$ folgt, gilt stets $\delta \leqslant \varepsilon$. Somit gilt:

$$R_\varepsilon \subseteq \bigcup_{\delta \leqslant \varepsilon} Z_\delta.$$

Ist umgekehrt $\delta \leqslant \varepsilon$ und $x \in Z_\delta$, so gilt, falls $\delta_i = 0$: $x_i \in (-\infty, a_i] \subseteq (-\infty, b_i^{\varepsilon_i} a_i^{1-\varepsilon_i}]$, und aus $\delta_i = 1$ folgt stets $\varepsilon_i = 1$ und somit ebenso $x_i \in (a_i, b_i] \subseteq (-\infty, b_i^{\varepsilon_i} a_i^{1-\varepsilon_i}]$. Also gilt $x \in R_\varepsilon$ und wir erhalten:

$$R_\varepsilon \supseteq \bigcup_{\delta \leqslant \varepsilon} Z_\delta.$$

Es bleibt die Disjunktheit der Vereinigung zu zeigen. Seien $\delta^1, \delta^2 \in \{0,1\}^k$ verschieden. Dann gibt es ein $i \in \{1,\ldots,k\}$ mit $\delta_i^1 \neq \delta_i^2$. Ohne Einschränkung können wir $\delta_i^1 = 0$ und $\delta_i^2 = 1$ annehmen und aus $(-\infty, a_i] \cap (-a_i, b_i] = \varnothing$ folgt die geforderte Disjunktheit.

Nun zeigen wir noch

$$\sum_{\varepsilon \geqslant \delta} (-1)^{k-\varepsilon_1-\cdots-\varepsilon_k} = \begin{cases} 0 & \text{falls } \delta \neq (1,\ldots,1) \\ 1 & \text{falls } \delta = (1,\ldots,1) \end{cases}.$$

Sei $\delta = (1,\ldots,1)$. Es folgt:

$$\sum_{\varepsilon \geqslant \delta} (-1)^{k-\varepsilon_1-\cdots-\varepsilon_k} = \sum_{\varepsilon=(1,\ldots,1)} (-1)^{k-\varepsilon_1-\cdots-\varepsilon_k} = (-1)^{k-k} = 1.$$

Sei $\delta \neq (1,\dots,1)$ mit $\delta_i = 0$ für ein $i \in \{1,\dots,k\}$. Wir definieren die Vektoren $\delta' := (\delta_1,\dots,\delta_{i-1},\delta_{i+1},\dots,\delta_k)$ sowie $\varepsilon' := (\varepsilon_1,\dots,\varepsilon_{i-1},\varepsilon_{i+1},\dots,\varepsilon_k)$. Es gilt:

$$\begin{aligned}\sum_{\varepsilon \geqslant \delta}(-1)^{k-\varepsilon_1-\dots-\varepsilon_k} &= \sum_{\varepsilon\geqslant\delta,\varepsilon_i=0}(-1)^{k-\varepsilon_1-\dots\varepsilon_{i-1}-0-\varepsilon_{i+1}-\dots-\varepsilon_k}\\ &\quad+\sum_{\varepsilon\geqslant\delta,\varepsilon_i=1}(-1)^{k-\varepsilon_1-\dots\varepsilon_{i-1}-1-\varepsilon_{i+1}-\dots-\varepsilon_k}\\ &=\sum_{\varepsilon'\geqslant\delta'}(-1)^{k-\varepsilon'_1-\dots-\varepsilon'_{k-1}}-\sum_{\varepsilon'\geqslant\delta'}(-1)^{k-\varepsilon'_1-\dots-\varepsilon'_{k-1}}=0.\end{aligned}$$

Mit diesen Vorüberlegungen erhalten wir:

$$\begin{aligned}\Delta_a^b F_G &= \sum_{\varepsilon\in\{1,0\}^k}(-1)^{k-\varepsilon_1-\dots-\varepsilon_k}F_G(b_1^{\varepsilon_1}a_1^{1-\varepsilon_1},\dots,b_k^{\varepsilon_k}a_k^{1-\varepsilon_k})\\ &=\sum_{\varepsilon\in\{1,0\}^k}(-1)^{k-\varepsilon_1-\dots-\varepsilon_k}G(R_\varepsilon)\\ &=\sum_{\varepsilon\in\{1,0\}^k}(-1)^{k-\varepsilon_1-\dots-\varepsilon_k}G\left(\sum_{\delta\leqslant\varepsilon}Z_\delta\right)\\ &=\sum_{\varepsilon\in\{1,0\}^k}\sum_{\delta\leqslant\varepsilon}(-1)^{k-\varepsilon_1-\dots-\varepsilon_k}G(Z_\delta)\\ &=\sum_{\delta\in\{0,1\}^k}\sum_{\varepsilon\geqslant\delta}(-1)^{k-\varepsilon_1-\dots-\varepsilon_k}G(Z_\delta)\\ &=\sum_{\delta\in\{0,1\}^k}G(Z_\delta)\sum_{\varepsilon\geqslant\delta}(-1)^{k-\varepsilon_1-\dots-\varepsilon_k}=G(Z_{(1,\dots,1)})=G\big((a,b]\big).\end{aligned}$$

Bemerkung 1.51. Beachte, dass die in Lemma 1.50 an die Mengenfunktion G gestellten Voraussetzungen im Allgemeinen von Maßen nicht erfüllt werden, jedoch von endlichen Maßen und somit auch von Wahrscheinlichkeitsmaßen.

Satz 1.52. *Sei P ein Wahrscheinlichkeitsmaß. Dann ist $F(x) := P\big((-\infty,x]\big)$ eine Verteilungsfunktion. Zu jeder Verteilungsfunktion F existiert genau ein Wahrscheinlichkeitsmaß auf $\mathscr{B}^k$ mit $P\big((a,b]\big) = \Delta_a^b F$ und umgekehrt.*

Beweis. Betrachte die Abbildung

$$\begin{aligned}\psi\colon \{F \mid F \text{ Verteilungsfunktion}\} &\to \big\{P\colon \mathscr{B}^k \to [0,1] \,\big|\, P \text{ Wahrscheinlichkeitsmaß}\big\}\\ F &\mapsto P_F,\end{aligned}$$

wobei P_F das zu F gehörige Maß nach Satz 1.49 sei. Beachte, dass P_F ein Wahrscheinlichkeitsmaß ist, da F eine Verteilungsfunktion ist. Des Weiteren definiere die Abbildung

$$\begin{aligned}\phi\colon \big\{P\colon \mathscr{B}^k \to [0,1] \,\big|\, P \text{ Wahrscheinlichkeitsmaß}\big\} &\to \{F \mid F \text{ Verteilungsfunktion}\}\\ P &\mapsto F_P,\end{aligned}$$

wobei
$$F_P\colon \mathbb{R}^k \to [0,1], \quad x \mapsto P\big((-\infty, x]\big).$$
F_P ist in der Tat eine Verteilungsfunktion, denn es gilt:

- Seien $a, b \in \mathbb{R}^k$ mit $a \leqslant b$. Da P ein Wahrscheinlichkeitsmaß ist, können wir Lemma 1.50 anwenden und erhalten:
$$\Delta_a^b F_P = P\big((a,b]\big) \geqslant 0.$$
- Seien $x = (x^1, \dots, x^k) \in \mathbb{R}^k$ und $(x_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{R}^k$, sodass für alle $j \in \{1, \dots, k\}$ gilt: $x_n = (x_n^1, \dots, x_n^k)$ und $x_n^j \searrow x^j$. Da P endlich ist, folgt mit der Stetigkeit von oben (Regeln 1.27):
$$\begin{aligned}\lim_{n\to\infty} F_P(x_n) = \lim_{n\to\infty} P\big((-\infty, x_n]\big) &= P\left(\bigcap_{n\in\mathbb{N}} (-\infty, x_n]\right)\\ &= P\big((-\infty, x]\big) = F_P(x).\end{aligned}$$
Folglich ist F_P rechtsstetig.
- Sei $(x_n)_{n\in\mathbb{N}} = \big((x_n^1, \dots, x_n^k)\big)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{R}^k$ mit $\lim_{n\to\infty} x_n^j = \infty$ für alle $j \in \{1, \dots, k\}$. Zu $n \in \mathbb{N}$ und $j \in \{1, \dots, k\}$ definiere $y_n^j := \min\left\{x_k^j \,\middle|\, k \geqslant n\right\}$. Dann ist $y_n \leqslant x_n$ für alle $n \in \mathbb{N}$, $y_n^j \nearrow \infty$ für alle $j \in \{1, \dots, k\}$ und mit der Stetigkeit von unten von P folgt:
$$1 \geqslant \lim_{n\to\infty} P\big((-\infty, x_n]\big) \geqslant \lim_{n\to\infty} P\big((-\infty, y_n]\big) = P\left(\bigcup_{n\in\mathbb{N}} (-\infty, y_n]\right) = P(\mathbb{R}^k) = 1.$$
Somit gilt: $\lim_{x_1\to\infty,\dots,x_k\to\infty} F_P(x) = 1$.
- Sei $(x_n)_{n\in\mathbb{N}} = \big((x_n^1, \dots, x_n^k)\big)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{R}^k$ und sei $\lim_{n\to\infty} x_n^j = -\infty$ für ein $j \in \{1, \dots, k\}$. Für alle $n \in \mathbb{N}$ definiere $y_n^j := \max\left\{x_k^j \,\middle|\, k \geqslant n\right\}$. Es gilt $y_n^j \geqslant x_n^j$ für alle $n \in \mathbb{N}$, $y_n^j \searrow -\infty$ und mit der Stetigkeit von oben von P folgt:
$$\begin{aligned}0 \leqslant \lim_{n\to\infty} P\big((-\infty, x_n]\big) &\leqslant \lim_{n\to\infty} P\big(\mathbb{R}^{j-1} \times (-\infty, y_n^j] \times \mathbb{R}^{k-j}\big)\\ &\leqslant P\left(\bigcap_{n\in\mathbb{N}} \mathbb{R}^{j-1} \times (-\infty, y_n^j] \times \mathbb{R}^{k-j}\right) = P(\varnothing) = 0.\end{aligned}$$
Somit gilt $\lim_{x_j\to-\infty} F_P(x) = 0$ für alle $j \in \{1, \dots, k\}$.

Um die Aussage des Satzes zu beweisen, reicht es zu zeigen, dass ψ und ϕ zueinander invers sind. Sei Q ein Wahrscheinlichkeitsmaß auf $\mathscr{B}^k$. Für alle $(a,b] \in \mathscr{I}^k$ gilt:
$$\psi\big(\phi(Q)\big)\big((a,b]\big) = P_{F_Q}\big((a,b]\big) = \Delta_a^b F_Q = Q\big((-\infty, b]\big) - Q\big((-\infty, a]\big) = Q\big((a,b]\big).$$

Nach Bemerkung 1.38 ist somit $\psi\big(\phi(Q)\big) = Q$.

Sei nun G eine Verteilungsfunktion. Für alle $x \in \mathbb{R}$ gilt:

$$\phi\big(\psi(G)\big)(x) = F_{P_G}(x) = P_G\big((-\infty, x]\big) = G(x) - \lim_{y \to -\infty} G(y) = G(x).$$

Folglich ist $\phi\big(\psi(G)\big) = G$. Somit sind ψ und ϕ zueinander invers.

Lemma 1.53. *Seien $G_i \colon \mathbb{R} \to \mathbb{R}$ Verteilungsfunktionen für alle $i \in \{1, \dots, k\}$. Dann ist*

$$G^{(k)}(x) := \prod_{i=1}^{k} G_i(x_i) \tag{1.12}$$

eine k-dimensionale Verteilungsfunktion. Außerdem gilt:

$$\Delta_a^b G^{(k)} = \prod_{i=1}^{k} \big(G_i(b_i) - G_i(a_i)\big) \qquad \text{für alle } a, b \in \mathbb{R}^k \text{ mit } a \leqslant b. \tag{1.13}$$

Beweis. Wir zeigen zunächst induktiv die Gültigkeit von (1.13). Für $k = 1$ gilt:

$$\Delta_a^b G^{(1)} = \sum_{\varepsilon_1 \in \{0,1\}} (-1)^{1-\varepsilon_1} G^{(1)}\left(b_1^{\varepsilon_1} a_1^{1-\varepsilon_1}\right) = G^{(1)}(b_1) - G^{(1)}(a_1) = G_1(b_1) - G_1(a_1).$$

Folglich ist die Behauptung für $k = 1$ erfüllt. Sein nun die Behauptung gezeigt für ein $k \in \mathbb{N}$. In $k + 1$ Dimensionen gilt:

$$\begin{aligned}
&\Delta_a^b G^{(k+1)} \\
&= \sum_{\varepsilon \in \{0,1\}^{k+1}} (-1)^{k+1-\varepsilon_1-\dots-\varepsilon_{k+1}} G^{(k+1)}\left(b_1^{\varepsilon_1} a_1^{1-\varepsilon_1}, \dots, b_{k+1}^{\varepsilon_{k+1}} a_{k+1}^{1-\varepsilon_{k+1}}\right) \\
&= \sum_{\varepsilon \in \{0,1\}^{k+1}, \varepsilon_{k+1}=0} (-1)^{k-\varepsilon_1-\dots-\varepsilon_k} (-1) G^{(k)}\left(b_1^{\varepsilon_1} a_1^{1-\varepsilon_1}, \dots, b_k^{\varepsilon_k} a_k^{1-\varepsilon_k}\right) G_{k+1}(a_{k+1}) \\
&\quad + \sum_{\varepsilon \in \{0,1\}^{k+1}, \varepsilon_{k+1}=1} (-1)^{k-\varepsilon_1-\dots-\varepsilon_k} G^{(k)}\left(b_1^{\varepsilon_1} a_1^{1-\varepsilon_1}, \dots, b_k^{\varepsilon_k} a_k^{1-\varepsilon_k}\right) G_{k+1}(b_{k+1}).
\end{aligned}$$

Mit der Induktionsvoraussetzung folgt:

$$\begin{aligned}
\Delta_a^b G^{(k+1)} &= \big(-G_{k+1}(a_{k+1})\big) \prod_{j=1}^{k} \big(G_j(b_j) - G_j(a_j)\big) \\
&\quad + G_{k+1}(b_{k+1}) \prod_{j=1}^{k} \big(G_j(b_j) - G_j(a_j)\big) \\
&= \prod_{j=1}^{k+1} \big(G_j(b_j) - G_j(a_j)\big).
\end{aligned}$$

Damit ist (1.13) bewiesen.

Wir prüfen nun, ob $G := G^{(k)}$ aus (1.12) die Eigenschaften einer k-dimensionalen Verteilungsfunktion erfüllt.

- *Rechtsstetigkeit:* Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{R}^k$ mit $x_n \searrow x \in \mathbb{R}^k$, das heißt $x_n^i \searrow x^i$ für alle $i \in \{1,\ldots,k\}$. Es gilt:

$$\lim_{n\to\infty} G(x_n) = \lim_{n\to\infty} G(x_n^1,\ldots,x_n^k) = \lim_{n\to\infty} \prod_{i=1}^k G_i(x_n^i) = \prod_{i=1}^k \lim_{n\to\infty} G_i(x_n^i).$$

  Hierbei existieren die Limiten, da $0 \leqslant G_i \leqslant 1$ für alle $i \in \{1,\ldots,k\}$ gilt. Mit der Rechtsstetigkeit der G_i folgt:

$$\lim_{n\to\infty} G(x_n) = \prod_{i=1}^k \lim_{n\to\infty} G_i(x_n^i) = \prod_{i=1}^k G_i(x^i) = G(x).$$

  Also ist G rechtsstetig.

- *Verallgemeinerte Monotonie:* Seien $a, b \in \mathbb{R}^k$ mit $a \leqslant b$. Da alle G_i isoton sind, gilt mit (1.13):

$$\Delta_a^b G = \prod_{i=1}^k \big(G_i(b_i) - G_i(a_i)\big) \geqslant 0.$$

- Da $\lim_{x_i\to\infty} G_i(x_i) = 1$ für alle $i \in \{1,\ldots,k\}$, gilt mit (1.13):

$$\lim_{x\to\infty} G(x) = \lim_{x\to\infty} \left(\prod_{i=1}^k G_i(x_i)\right) = \prod_{i=1}^k \lim_{x_i\to\infty} G_i(x_i) = \prod_{i=1}^k 1 = 1.$$

- Da $\lim_{x_j\to-\infty} G_j(x_j) = 0$ für alle $j \in \{1,\ldots,k\}$, gilt mit (1.13):

$$\lim_{x_j\to-\infty} G(x) = \lim_{x_j\to-\infty} \left(\prod_{i=1}^k G_i(x_i)\right) = \lim_{x_j\to-\infty} G_j(x_j) \prod_{i\neq j} G_i(x_i) = 0$$

  für alle $j \in \{1,\ldots,k\}$.

Folglich ist G eine k-dimensionale Verteilungsfunktion.

Beispiel 1.54.

1. Falls in Lemma 1.53 $G_i = \mathrm{id}_{\mathbb{R}}$ für alle $i \in \{1,\ldots,k\}$, so definiert

$$G(x) := \prod_{i=1}^k x_i$$

   das *Borel-Lebesgue-Maß* λ^k in $\mathbb{R}^k$ mit

$$\lambda^k\big((a,b]\big) = \Delta_a^b G = \prod_{i=1}^k (b_i - a_i) \qquad \text{für alle } a, b \in \mathbb{R}^k \text{ mit } a \leqslant b.$$

2. Die k-dimensionale Normalverteilung mit Erwartungswert $\mu \in \mathbb{R}^k$ und positiv definiter und symmetrischer Kovarianzmatrix $\Sigma \in \mathbb{R}^{k\times k}$ (siehe Definition 3.17) ist gegeben durch

$$G(x) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_k} \frac{1}{\sqrt{2\pi}^k} (\det \Sigma)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(y-\mu)^t \Sigma^{-1}(y-\mu)\right) \, \mathrm{d}y.$$

Dies ist eine k-dimensionale Verteilungsfunktion. [Übung]$^\bigstar$Beachte, dass aus der positiven Definitheit von Σ die Existenz ihrer Inversen folgt.

Tipp: Wenden Sie den Transformationssatz für mehrdimensionale Riemann-Integrale an und verwenden Sie die Hauptachsentransformation $U\Sigma U^t = D$ mit einer orthogonalen Matrix U und einer Diagonalmatrix D.

3. Die Verteilungsfunktion F der *Cantor-Verteilung* über dem Intervall $(0,1)$ definiert man iterativ, und zwar im ersten Schritt über $(1/3, 2/3)$, im zweiten Schritt über $(1/9, 2/9)$ und $(7/9, 8/9)$ und allgemein im k-ten Schritt über den mittleren Dritteln der verbliebenen 2^{k-1} Intervalle. Den Funktionswert wählen wir dabei über jedem dieser offenen Intervalle konstant gleich dem arithmetischen Mittel der Funktionswerte über den unmittelbar benachbarten Intervallen, über denen F bereits erklärt ist. Wegen $F(0) = 0$ und $F(1) = 1$ ergibt sich im ersten Schritt (siehe Abbildung 1.2)

$$F(x) = \frac{1}{2} \quad \text{für } \frac{1}{3} < x < \frac{2}{3}$$

und im zweiten Schritt (siehe Abbildung 1.3)

$$F(x) = \begin{cases} 1/4 & \text{für } 1/9 < x < 2/9 \\ 3/4 & \text{für } 7/9 < x < 8/9 \end{cases}$$

und allgemein für $k \in \mathbb{N}$

$$F(x) = \sum_{i=1}^{k-1} \frac{\delta_i}{2^i} + \frac{1}{2^k} \quad \text{für} \quad x \in \left(\sum_{i=1}^{k-1} 2\frac{\delta_i}{3^i} + \frac{1}{3^k}, \sum_{i=1}^{k-1} 2\frac{\delta_i}{3^i} + \frac{2}{3^k}\right),$$

wobei $(\delta_1, \ldots, \delta_{k-1})$ alle $(k-1)$-Tupel mit den Komponenten $\delta_i \in \{0,1\}$ für $i \in \{1, \ldots, k-1\}$ durchläuft. Dies definiert F für $k \to \infty$ bereits auf einer dichten Teilmenge des Intervalls $(0,1)$. Durch rechtsstetiges (sogar stetiges) Fortsetzen erhält man die *Cantor-Verteilung*. Dies ist eine Verteilungsfunktion. [Übung]

Eine alternative Definition der Cantor-Verteilung lautet folgendermaßen: Sei $z \in [0,1]$. Wir können z in der 3-adischen Darstellung schreiben als

$$z = \sum_{i=0}^{\infty} z_i 3^{-i}.$$

Damit diese Darstellung eindeutig wird, fordern wir außerdem, dass im Fall $z \neq 0$ für alle $n \in \mathbb{N}$ ein $k \geqslant n$ mit $z_k \neq 0$ existiert. Anschaulich verbietet man, dass die Darstellung abbricht, da zum Beispiel gilt:

$$\sum_{i=1}^{\infty} 2 \cdot 3^{-i} = 1.$$

Mit dieser Forderung folgt für alle $z \in [0,1]$ direkt $z_0 = 0$. Fordert man stattdessen, dass es für alle $n \in \mathbb{N}$ ein $k \geqslant n$ mit $z_k \neq 2$ gibt, so ist die 3-adische Darstellung ebenfalls eindeutig. Für die folgende Definition ist es egal, auf welche Weise man die Darstellung eindeutig macht, weshalb wir annehmen, die Darstellung erfülle die erste Forderung. Definiere nun zu $z \in [0,1]$

$$k_z := \inf \{k \in \mathbb{N} \mid z_k = 1\},$$

also die Stelle in der 3-adischen Darstellung von z, an der das erste Mal eine 1 als Ziffer auftritt. Tritt keine 1 auf, so gilt: $k_z = \infty$. Nun lässt sich die Cantor-Verteilung folgendermaßen definieren:

$$C\colon [0,1] \to \mathbb{R}, \quad z \mapsto 1 + \sum_{i=0}^{k_z} \frac{z_i - 1}{2^{i+1}}.$$

Die Menge $\{z \in [0,1] \mid k_z = \infty\}$ heißt *Cantor-Menge*.

Dass die beiden Definitionen äquivalent sind, resultiert daraus, dass für alle $k \in \mathbb{N}$ und $z_i \in \{0,2\}$ mit $i < k$ sowie $z_k = 1$ gilt:

$$\begin{aligned} M(z_1, \ldots, z_k) &:= \left\{ \sum_{i=1}^{\infty} z_i 3^{-i} \;\middle|\; z_i \in \{0,1,2\} \text{ für alle } i > k \right\} \\ &= \left(\sum_{i=1}^{k-1} \frac{z_i}{3^i} + \frac{1}{3^k}, \sum_{i=1}^{k-1} \frac{z_i}{3^i} + \frac{2}{3^k} \right), \end{aligned}$$

und dass für alle $z \in M(z_1, \ldots, z_k)$ gilt:

$$C(z) = 1 + \sum_{i=0}^{k-1} \frac{z_i - 1}{2^{i+1}} = \sum_{i=1}^{k-1} \frac{z_i}{2^{i+1}} + \frac{1}{2} - \sum_{i=1}^{k-1} \frac{1}{2^{i+1}} = \sum_{i=1}^{k-1} \frac{z_i/2}{2^i} + \frac{1}{2^k} = F(z).$$

Wenn μ ein Maß auf einer σ-Algebra $\mathscr{A}$ ist, so liegt es intuitiv nahe, dass jede Teilmenge einer μ-Nullmenge ebenfalls eine μ-Nullmenge sein sollte. Wegen der Monotonie von μ wird es natürlich nicht vorkommen, dass einer solchen Teilmenge ein Maß größer null zugeordnet wird. Es ist aber durchaus möglich, dass nicht jede Teilmenge einer μ-Nullmenge in $\mathscr{A}$ enthalten ist und folglich gar nicht gemessen werden kann. Es stellt sich jedoch heraus, dass wir jeden Maßraum problemlos mittels Ergänzung der σ-Algebra um diese Teilmengen vervollständigen können.

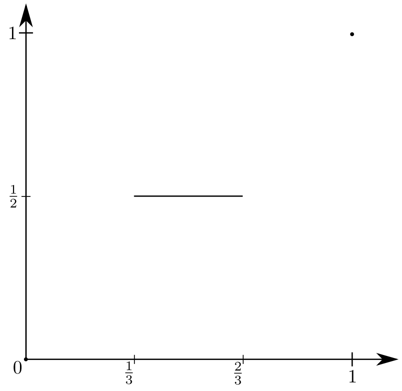


Abbildung 1.2: erste Iteration der Cantor-Verteilung

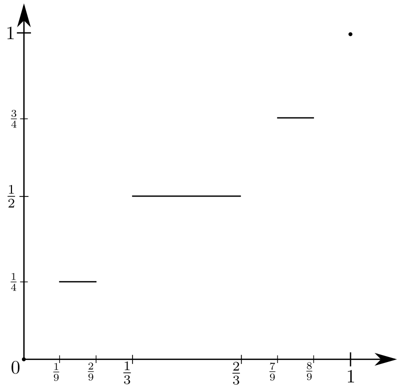


Abbildung 1.3: zweite Iteration der Cantor-Verteilung

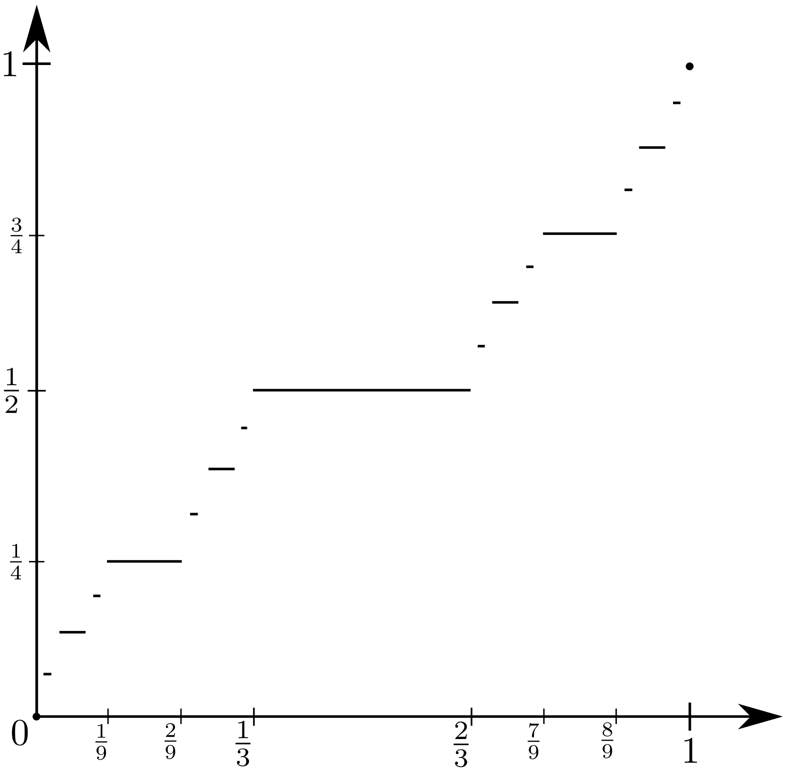


Abbildung 1.4: vierte Iteration der Cantor-Verteilung

Definition 1.55. Ein Maßraum $(\Omega, \mathscr{A}, \mu)$ heißt *vollständig* (engl.: complete), falls für alle μ-Nullmengen $A \in \mathscr{A}$ gilt:

$$B \subseteq A \Rightarrow B \in \mathscr{A}.$$

Lemma 1.56. *Sei $(\Omega, \mathscr{A}, \mu)$ ein Maßraum. Das System*

$$\mathscr{A}_\mu := \{A \subseteq \Omega \mid \text{es gibt } E, F \in \mathscr{A} \text{ mit } E \subseteq A \subseteq F \text{ und } \mu(F \backslash E) = 0\}$$

ist eine σ-Algebra $\mathscr{A} \subseteq \mathscr{A}_\mu$ und

$$\begin{aligned} \tilde{\mu} \colon \mathscr{A}_\mu &\to [0, \infty] \\ A &\mapsto \sup \{\mu(B) \mid B \in \mathscr{A}, B \subseteq A\} \end{aligned}$$

ist ein Maß auf $\mathscr{A}_\mu$, welches μ fortsetzt. Der Maßraum $(\Omega, \mathscr{A}_\mu, \tilde{\mu})$ ist vollständig und heißt Vervollständigung *von $(\Omega, \mathscr{A}, \mu)$.*

Beweis. Für $\mathscr{A}_\mu$ gilt:

- Für alle $A \in \mathscr{A}$ gilt $A \subseteq A \subseteq A$ mit $\mu(A \backslash A) = 0$. Somit ist $\mathscr{A} \subseteq \mathscr{A}_\mu$ und da $\mathscr{A}$ eine σ-Algebra ist, folgt $\Omega \in \mathscr{A}_\mu$.
- Seien $E, F \in \mathscr{A}$ mit $\mu(F \backslash E) = 0$ und sei $A \in \mathscr{A}_\mu$ mit $E \subseteq A \subseteq F$. Da $\mathscr{A}$ eine σ-Algebra ist, gilt $E^c, F^c \in \mathscr{A}$ sowie $F^c \subseteq A^c \subseteq E^c$. Wegen

  $$\mu(E^c \backslash F^c) = \mu\big((\Omega \backslash E) \backslash (\Omega \backslash F)\big) = \mu(F \backslash E) = 0$$

  folgt: $A^c \in \mathscr{A}_\mu$.
- Seien $A_1, A_2, \ldots \in \mathscr{A}_\mu$, $E_n, F_n \in \mathscr{A}$, sodass $E_n \subseteq A_n \subseteq F_n$ und $\mu(F_n \backslash E_n) = 0$ für alle $n \in \mathbb{N}$. Definiere $A := \bigcup_{n \in \mathbb{N}} A_n, E := \bigcup_{n \in \mathbb{N}} E_n$ und $F := \bigcup_{n \in \mathbb{N}} F_n$. Dann sind, da $\mathscr{A}$ eine σ-Algebra ist, $E, F \in \mathscr{A}$ mit $E \subseteq A \subseteq F$ und da μ ein Maß ist, folgt $\mu(F \backslash E) \leqslant \sum_{n \in \mathbb{N}} \mu(F_n \backslash E_n) = 0$. Also ist $A \in \mathscr{A}_\mu$.

Folglich ist $\mathscr{A}_\mu$ eine σ-Algebra über Ω. Für $\tilde{\mu}$ gilt:

- Sei $A \in \mathscr{A}$. Wegen der Monotonie des Maßes μ gilt:

  $$\tilde{\mu}(A) = \sup \{\mu(B) \mid B \in \mathscr{A}, B \subseteq A\} = \mu(A).$$

  Damit folgt insbesondere $\tilde{\mu}(\varnothing) = 0$.
- Seien $A, B \in \mathscr{A}_\mu$ mit $A \subseteq B$. Dann ist

  $$\{\mu(C) \mid C \in \mathscr{A}, C \subseteq A\} \subseteq \{\mu(C) \mid C \in \mathscr{A}, C \subseteq B\}$$

  und es folgt:

  $$\tilde{\mu}(A) = \sup \{\mu(C) \mid C \in \mathscr{A}, C \subseteq A\} \leqslant \sup \{\mu(C) \mid C \in \mathscr{A}, C \subseteq B\} = \tilde{\mu}(B).$$

  Also ist $\tilde{\mu}$ monoton.

- Seien nun $A_1, A_2, \ldots \in \mathscr{A}_\mu$ paarweise disjunkt. Weiterhin seien $E_n, F_n \in \mathscr{A}$ mit $E_n \subseteq A_n \subseteq F_n$ und $\mu(F_n \backslash E_n) = 0$ für alle $n \in \mathbb{N}$. Definiere $A := \sum_{n\in\mathbb{N}} A_n$, $E := \sum_{n\in\mathbb{N}} E_n$ und $F := \bigcup_{n\in\mathbb{N}} F_n$. Beachte, dass somit aus der Monotonie von $\tilde{\mu}$ und $\tilde{\mu}|_{\mathscr{A}} = \mu$ bereits $\tilde{\mu}(A) = \mu(E)$ und $\tilde{\mu}(A_n) = \mu(E_n)$ für alle $n \in \mathbb{N}$ folgt. Es gilt:
$$\tilde{\mu}(A) = \mu(E) = \sum_{n\in\mathbb{N}} \mu(E_n) = \sum_{n\in\mathbb{N}} \tilde{\mu}(A_n).$$
Folglich ist $\tilde{\mu}$ σ-additiv.

Somit haben wir gezeigt, dass $\tilde{\mu}$ ein Maß ist, welches μ auf $\mathscr{A}_\mu$ fortsetzt. Sei nun $N \in \mathscr{A}_\mu$ eine $\tilde{\mu}$-Nullmenge. Seien $A \subseteq N$ und $E, F \in \mathscr{A}$ mit $E \subseteq N \subseteq F$ und $\mu(F \backslash E) = 0$. Nach dem bisher Gezeigten ist F eine $\tilde{\mu}$-Nullmenge mit $\emptyset \subseteq A \subseteq F$ und $\mu(F \backslash \emptyset) = 0$, also gilt $A \in \mathscr{A}_\mu$. Somit ist $(\Omega, \mathscr{A}_\mu, \tilde{\mu})$ eine Vervollständigung von $(\Omega, \mathscr{A}, \mu)$.

Bemerkung 1.57. Der Maßraum $(\mathbb{R}^k, \mathscr{B}^k, \lambda^k)$ ist nicht vollständig. Die Fortsetzung von λ^k auf die Vervollständigung von $(\mathbb{R}^k, \mathscr{B}^k, \lambda^k)$ heißt *Lebesgue-Maß*. Die Vervollständigung von $\mathscr{B}^k$ heißt *σ-Algebra der Lebesgue-Mengen.*

Übungsaufgaben:

Übung 1: Beweisen Sie Punkt 4 von Bemerkung 1.8.

Übung 2: Beweisen Sie Lemma 1.11.

Übung 3: Beweisen Sie Beispiel 1.16.

Übung 4: Beweisen Sie Punkt 5 der Regeln 1.27.

Übung 5: Beweisen Sie Bemerkung 1.28.

Übung 6: Zeigen Sie, dass es sich bei der k-dimensionalen Normalverteilung um eine k-dimensionale Verteilungsfunktion handelt (siehe 1.54).

Übung 7: Zeigen Sie, dass es sich bei Cantor-Verteilung um eine Verteilungsfunktion handelt (siehe 1.54).

Übung 8: Beweisen Sie das *Approximationslemma*: Sei $(\Omega, \mathscr{A}, \mu)$ ein endlicher Maßraum und $\mathscr{A}_0$ eine Algebra mit $\sigma(\mathscr{A}_0) = \mathscr{A}$. Sei $A \triangle B := (A \backslash B) \cup (B \backslash A)$ die symmetrische Differenz von A und B. Dann gibt es zu jedem $A \in \mathscr{A}$ und $\varepsilon > 0$ ein $A_0 \in \mathscr{A}_0$ mit $\mu(A \triangle A_0) < \varepsilon$.

Tipp: Für den Nachweis, dass eine Eigenschaft E für jede Menge A einer σ-Algebra $\mathscr{A}$ über Ω gilt, bilde man das Mengensystem M aller Teilmengen von Ω, für welche E erfüllt ist und zeige:

(a) M enthält einen Erzeuger von $\mathscr{A}$,

(b) M ist eine σ-Algebra.

Lernziel-Kontrolle:

Frage 1: Was sind σ-Algebren und welchen Zweck erfüllen sie?

Frage 2: Was ist das Maßproblem und welche Paradoxien können sich aus diesem ergeben?

Frage 3: Was ist ein Dynkin-System und warum wird dieser Begriff eingeführt?

Frage 4: Was besagt der Satz von Carathéodory?

Frage 5: Was ist ein äußeres Maß und auf welchen Mengen kann es definiert werden?

Frage 6: Was besagt der Maßfortsetzungssatz?

Frage 7: Welcher Zusammenhang besteht zwischen Wahrscheinlichkeitsmaßen und Verteilungsfunktionen auf $\mathbb{R}^1$, welcher auf $\mathbb{R}^k$?

Frage 8: Kennen Sie Beispiele von Verteilungsfunktionen und den zugehörigen Maßen?

Frage 9: Welche Eigenschaften muss eine maßdefinierende Funktion auf $\mathbb{R}^k$ erfüllen?

Frage 10: Wie und weshalb vervollständigt man Maßräume?

Kapitel 2

Messbare Abbildungen und Zufallsvariablen

Im Folgenden betrachten wir Abbildungen zwischen Maßräumen. Insbesondere werden wir Zufallsvariablen kennen lernen.

Regeln 2.1. *Seien Ω und Ω' Mengen und $f\colon \Omega \to \Omega'$ eine Abbildung. Diese induziert eine Urabbildung*

$$\begin{aligned} f^{-1}\colon \mathscr{P}(\Omega') &\to \mathscr{P}(\Omega) \\ A' &\mapsto \left\{\omega \in \Omega \,\middle|\, f(\omega) \in A'\right\}. \end{aligned}$$

Seien J eine Indexmenge und $A'_j \subseteq \Omega'$ für alle $j \in J$. Die folgenden Rechenregeln bilden das Fundament dieses Abschnitts:

$$f^{-1}\left(\bigcap_{j\in J} A'_j\right) = \bigcap_{j\in J} f^{-1}(A'_j), \tag{2.1}$$

$$f^{-1}(A'^c) = \left(f^{-1}(A')\right)^c, \tag{2.2}$$

$$f^{-1}\left(\bigcup_{j\in J} A'_j\right) = \bigcup_{j\in J} f^{-1}(A'_j). \tag{2.3}$$

Beweis.

1. Es gilt:

$$\begin{aligned} f^{-1}\left(\bigcap_{j\in J} A'_j\right) &= \left\{\omega \in \Omega \,\middle|\, f(\omega) \in \bigcap_{j\in J} A'_j\right\} = \left\{\omega \in \Omega \,\middle|\, f(\omega) \in A'_j \ \forall\, j \in J\right\} \\ &= \bigcap_{j\in J} \left\{\omega \in \Omega \,\middle|\, f(\omega) \in A'_j\right\} = \bigcap_{j\in J} f^{-1}(A'_j). \end{aligned}$$

2. $f^{-1}(A'^c) = \{\omega \in \Omega \mid f(\omega) \in A'^c\} = \{\omega \in \Omega \mid f(\omega) \in A'\}^c = \big(f^{-1}(A')\big)^c$.

3. Die Behauptung folgt aus den ersten beiden Aussagen mit

$$\bigcup_{j \in J} A'_j = \left(\bigcap_{j \in J} A'^{\,c}_j\right)^c .$$

Lemma 2.2. *Sei $f\colon \Omega \to \Omega'$ eine Abbildung. Dann gilt:*

1. Ist $\mathscr{A}'$ eine σ-Algebra über Ω', so ist

$$f^{-1}(\mathscr{A}') := \left\{ f^{-1}(A') \in \mathscr{P}(\Omega) \,\middle|\, A' \in \mathscr{A}' \right\}$$

eine σ-Algebra über Ω.

2. Ist $\mathscr{A}$ eine σ-Algebra über Ω, so ist

$$\mathscr{A}_f := \left\{ A' \in \mathscr{P}(\Omega') \,\middle|\, f^{-1}(A') \in \mathscr{A} \right\}$$

eine σ-Algebra über Ω'.

Beweis. Wir prüfen mithilfe der Regeln 2.1, ob die Eigenschaften einer σ-Algebra erfüllt sind.

1.
   - Da $\Omega' \in \mathscr{A}'$ und $f^{-1}(\Omega') = \Omega$, ist $\Omega \in f^{-1}(\mathscr{A}')$.
   - Sei $A \in f^{-1}(\mathscr{A}')$. Dann gibt es $A' \in \mathscr{A}'$ mit $A = f^{-1}(A')$. Da auch $A'^c \in \mathscr{A}'$ ist und $f^{-1}(A'^c) = \big(f^{-1}(A')\big)^c = A^c$ gilt, folgt $A^c \in f^{-1}(\mathscr{A}')$.
   - Seien $A_1, A_2, \ldots \in f^{-1}(\mathscr{A}')$. Dann gibt es Mengen $A'_1, A'_2, \ldots \in \mathscr{A}'$ mit $A_n = f^{-1}(A'_n)$ für alle $n \in \mathbb{N}$. Es gilt:

$$\bigcup_{n \in \mathbb{N}} A_n = \bigcup_{n \in \mathbb{N}} f^{-1}(A'_n) = f^{-1}\left(\bigcup_{n \in \mathbb{N}} A'_n\right) \in f^{-1}(\mathscr{A}').$$

2.
   - Da $\Omega \in \mathscr{A}$ und $f^{-1}(\Omega') = \Omega$, ist $\Omega' \in \mathscr{A}_f$.
   - Sei $A' \in \mathscr{A}_f$. Dann ist $f^{-1}(A') \in \mathscr{A}$ und somit $f^{-1}(A'^c) = \big(f^{-1}(A')\big)^c \in \mathscr{A}$. Folglich ist $A'^c \in \mathscr{A}_f$.
   - Seien $A'_1, A'_2, \ldots \in \mathscr{A}_f$. Dann sind $f^{-1}(A'_1), f^{-1}(A'_2), \ldots \in \mathscr{A}$. Es gilt:

$$f^{-1}\left(\bigcup_{n \in \mathbb{N}} A'_n\right) = \bigcup_{n \in \mathbb{N}} f^{-1}(A'_n) \in \mathscr{A}.$$

   Folglich ist

$$\bigcup_{n \in \mathbb{N}} A'_n \in \mathscr{A}_f.$$

Bemerkung 2.3. Seien $(\Omega, \mathscr{A}), (\Omega', \mathscr{A}')$ Messräume und $f\colon \Omega \to \Omega'$ eine Abbildung. $f(\mathscr{A}) := \{f(A) \in \mathscr{P}(\Omega') \mid A \in \mathscr{A}\}$ ist im Allgemeinen keine σ-Algebra über Ω'.

Definition 2.4. Seien $(\Omega, \mathscr{A}), (\Omega', \mathscr{A}')$ Messräume. Eine Abbildung $f\colon \Omega \to \Omega'$ heißt $(\mathscr{A}, \mathscr{A}')$-*messbar* (engl.: measurable), falls $f^{-1}(\mathscr{A}') \subseteq \mathscr{A}$.

Bemerkung 2.5.

- Eine Abbildung ist genau dann messbar, wenn die Urbilder messbarer Mengen messbar sind. Dies verhält sich analog zu stetigen Abbildungen (Urbilder offener Mengen sind offen).
- $\mathscr{A}_f$ ist die größte σ-Algebra, sodass f gerade noch $(\mathscr{A}, \mathscr{A}_f)$-messbar ist. Folglich ist f genau dann $(\mathscr{A}, \mathscr{A}')$-messbar, wenn $\mathscr{A}' \subseteq \mathscr{A}_f$.
- Die Borel-σ-Algebra (vgl. Bemerkung 1.23) ist die kleinste σ-Algebra, bezüglich welcher alle reellwertigen, stetigen Funktionen messbar sind.

Lemma 2.6. *Seien $(\Omega, \mathscr{A}), (\Omega', \mathscr{A}')$ Messräume. Sei $M' \subseteq \mathscr{P}(\Omega')$ mit $\sigma(M') = \mathscr{A}'$. Dann ist $f\colon \Omega \to \Omega'$ genau dann $(\mathscr{A}, \mathscr{A}')$-messbar, wenn $f^{-1}(M') \subseteq \mathscr{A}$.*

Beweis. Sei f $(\mathscr{A}, \mathscr{A}')$-messbar, also $f^{-1}(\mathscr{A}') \subseteq \mathscr{A}$. Da $M \subseteq \mathscr{A}'$, folgt insbesondere $f^{-1}(M') \subseteq \mathscr{A}$.

Sei nun $f^{-1}(M') \subseteq \mathscr{A}$. Nach der Definition von $\mathscr{A}_f$ ist $M' \subseteq \mathscr{A}_f$. Da $\mathscr{A}_f$ eine σ-Algebra ist, folgt: $\sigma(M') \subseteq \mathscr{A}_f$. Mit Bemerkung 2.2.5. folgt die Behauptung.

Beispiel 2.7. Betrachte eine Abbildung $f\colon \Omega \to \mathbb{R}^k$ zwischen den Messräumen $(\Omega, \mathscr{A})$ und $(\mathbb{R}^k, \mathscr{B}^k)$. Diese ist genau dann $(\mathscr{A}, \mathscr{B}^k)$-messbar, wenn $f^{-1}(\mathscr{I}^k) \subseteq \mathscr{A}$ ist.

Bemerkung 2.8. Lemma 2.6 besagt, dass es für die Messbarkeit bezüglich $\mathscr{A}'$ ausreichend ist, Messbarkeit bezüglich eines Erzeugers von $\mathscr{A}'$ zu zeigen.

Lemma 2.9. *Sei $(\Omega, \mathscr{A})$ ein Messraum.*

1. *Seien $f_j\colon \Omega \to \mathbb{R}$ für alle $j \in \{1, \ldots, k\}$ und $f := (f_1, \ldots, f_k)$. So ist f genau dann $(\mathscr{A}, \mathscr{B}^k)$-messbar, wenn f_j $(\mathscr{A}, \mathscr{B}^1)$-messbar ist für alle $j \in \{1, \ldots, k\}$.*
2. *Sei $f\colon \mathbb{R}^k \to \mathbb{R}^n$ stetig. Dann ist f* Borel-messbar*, das heißt $(\mathscr{B}^k, \mathscr{B}^n)$-messbar.*
3. *$\mathbb{1}_A$ ist genau dann $(\mathscr{A}, \mathscr{B}^1)$-messbar, wenn $A \in \mathscr{A}$.*
4. *Seien $(\Omega_1, \mathscr{A}_1)$, $(\Omega_2, \mathscr{A}_2)$ und $(\Omega_3, \mathscr{A}_3)$ Messräume und seien $f_1\colon \Omega_1 \to \Omega_2$ sowie $f_2\colon \Omega_2 \to \Omega_3$ $(\mathscr{A}_1, \mathscr{A}_2)$-messbar beziehungsweise $(\mathscr{A}_2, \mathscr{A}_3)$-messbar. Dann ist $f_2 \circ f_1$ $(\mathscr{A}_1, \mathscr{A}_3)$-messbar.*

$$\begin{tikzcd}
(\Omega_1, \mathscr{A}_1) \arrow[r, "f_1"] \arrow[dr, "f_2 \circ f_1"'] & (\Omega_2, \mathscr{A}_2) \arrow[d, "f_2"] \\
& (\Omega_3, \mathscr{A}_3)
\end{tikzcd}$$

Beweis.

1. Sei f $(\mathscr{A}, \mathscr{B}^k)$-messbar. Dann ist $f^{-1}(\mathscr{B}^k) \subseteq \mathscr{A}$. Insbesondere enthält $\mathscr{B}^k$ Mengen der Form

$$M_i\big((a,b]\big) := \big\{x \in \mathbb{R}^k \,\big|\, x_i \in (a,b]\big\} \quad \text{mit } a, b \in \mathbb{R}; a \leqslant b; i \in \{1, \ldots, k\}\,.$$

Es gilt:

$$\begin{aligned} f^{-1}\Big(M_i\big((a,b]\big)\Big) &= f_1^{-1}(\mathbb{R}) \cap \cdots \cap f_i^{-1}\big((a,b]\big) \cap \cdots \cap f_k^{-1}(\mathbb{R}) \\ &= \Omega \cap \cdots \cap f_i^{-1}\big((a,b]\big) \cap \cdots \cap \Omega \\ &= f_i^{-1}\big((a,b]\big). \end{aligned}$$

Folglich ist $f_i^{-1}(\mathscr{I}^1) \subseteq \mathscr{A}$. Also ist f_i für alle $i \in \{1, \ldots, k\}$ $(\mathscr{A}, \mathscr{B}^1)$-messbar.

Sei nun $f_i^{-1}(\mathscr{I}^1) \subseteq \mathscr{A}$ für $i \in \{1, \ldots, k\}$. Betrachte einen Quader $(x, y] \in \mathscr{I}^k$. Es ist

$$f^{-1}\big((x,y]\big) = \bigcap_{i=1}^{k} f_i^{-1}\big((x_i, y_i]\big).$$

Da für alle $i \in \{1, \ldots, k\}$ $f_i^{-1}\big((x_i, y_i]\big) \in \mathscr{A}$ gilt und außerdem $\mathscr{A}$ $\cap$-stabil ist, folgt: $f^{-1}\big((x,y]\big) \in \mathscr{A}$. Somit ist $f^{-1}(\mathscr{I}^k) \subseteq \mathscr{A}$.

2. Sei f stetig. Dann sind Urbilder offener Mengen offen und es gilt $f^{-1}(\mathscr{O}^n) \subseteq \mathscr{O}^k$. Also ist f $(\mathscr{B}^k, \mathscr{B}^n)$-messbar.

3. $\mathbb{1}_A$ hat nur die Urbilder $\varnothing$, A, A^c und Ω. Folglich ist $\mathbb{1}_A$ genau dann messbar, wenn $\{\varnothing, A, A^c, \Omega\} \subseteq \mathscr{A}$, was genau dann erfüllt ist, wenn $A \in \mathscr{A}$.

4. Seien f_1, f_2 messbar und $A \in \mathscr{A}_3$. Dann ist $(f_2 \circ f_1)^{-1}(A) = f_1^{-1}\big(f_2^{-1}(A)\big) \in \mathscr{A}_1$.

Definition 2.10. Seien Ω eine Menge, J eine Indexmenge und $(\Omega_j, \mathscr{A}_j)_{j \in J}$ eine Familie von Messräumen sowie $(f_j)_{j \in J}$ eine Familie von Abbildungen mit $f_j \colon \Omega \to \Omega_j$ für alle $j \in J$. Dann heißt

$$\sigma\,(f_j \mid j \in J) := \sigma\left(\bigcup_{j \in J} f_j^{-1}(\mathscr{A}_j)\right)$$

die *von $(f_j)_{j \in J}$ erzeugte σ-Algebra.*

Lemma 2.11. *Seien Ω eine Menge, $(\Omega_0, \mathscr{A}_0)$ ein Messraum, $f \colon \Omega_0 \to \Omega$, J eine Indexmenge und $(\Omega_j, \mathscr{A}_j)_{j \in J}$ eine Familie von Messräumen sowie $(f_j)_{j \in J}$ eine Familie von Abbildungen mit $f_j \colon \Omega \to \Omega_j$ für alle $j \in J$. Dann ist f genau dann $\big(\mathscr{A}_0, \sigma\,(f_j \mid j \in J)\big)$-messbar, wenn $f_j \circ f$ für alle $j \in J$ $(\mathscr{A}_0, \mathscr{A}_j)$-messbar ist.*

Definition 2.12. Seien $(\Omega_j, \mathscr{A}_j)_{j\in\{1,\dots,n\}}$ Messräume und $\Omega := \times_{j=1}^n \Omega_j$. Definiere für alle $j \in \{1,\dots,n\}$ die Projektion auf Ω_j:

$$\pi_j \colon \Omega \to \Omega_j, \quad \omega \mapsto \omega_j.$$

Die von den Projektionen erzeugte σ-Algebra

$$\bigotimes_{j=1}^n \mathscr{A}_j := \sigma(\pi_1,\dots,\pi_n)$$

heißt *Produkt-σ-Algebra.*

Satz 2.13. *Für alle $i \in \{1,\dots,n\}$ sei M_i ein Erzeuger der σ-Algebra $\mathscr{A}_i$ auf Ω_i, welcher eine Folge $(N_i^k)_{k\in\mathbb{N}}$ von Mengen mit $N_i^k \nearrow \Omega_i$ enthält. Dann erzeugt das Mengensystem*

$$E := \left\{ \bigtimes_{i=1}^n N_i \;\middle|\; N_i \in M_i \text{ für alle } i \in \{1,\dots,n\} \right\}$$

die σ-Algebra $\mathscr{A} := \bigotimes_{i=1}^n \mathscr{A}_i$.

Beweis. Wir zeigen zunächst $\mathscr{A} \subseteq \sigma(E)$. Für $j \in \{1,\dots,n\}$ sei $\pi_j \colon \times_{i=1}^n \Omega_i \to \Omega_j$ die Projektion auf die j-te Komponente. Es gilt

$$\sigma(\pi_j) = \left\{\pi_j^{-1}(A) \,\middle|\, A \in \mathscr{A}_j\right\} = \left\{ \bigtimes_{i=1}^{j-1} \Omega_i \times A \times \bigtimes_{i=j+1}^n \Omega_i \;\middle|\; A \in \mathscr{A}_j \right\}.$$

Weiterhin gilt für alle $N \in M_j$

$$\left(\bigtimes_{i=1}^{j-1} N_i^k \times N \times \bigtimes_{i=j+1}^n N_i^k \right) \nearrow \left(\bigtimes_{i=1}^{j-1} \Omega_i \times N \times \bigtimes_{i=j+1}^n \Omega_i \right) = \pi_j^{-1}(N) \in \sigma(E).$$

Aus Lemma 2.6 folgt somit die $(\sigma(E), \mathscr{A}_j)$-Messbarkeit von π_j, also $\sigma(\pi_j) \subseteq \sigma(E)$. Nach Definition der Produkt-σ-Algebra folgt $\mathscr{A} \subseteq \sigma(E)$.

Nun zeigen wir $\sigma(E) \subseteq \mathscr{A}$. Es reicht $E \subseteq \mathscr{A}$ zu zeigen. Sei $N_1 \times \dots \times N_n \in E$. Es gilt

$$\bigtimes_{i=1}^n N_i = \bigcap_{i=1}^n \pi_i^{-1}(N_i) \in \sigma(\pi_1,\dots,\pi_n) = \mathscr{A}.$$

Somit folgt $\sigma(E) \subseteq \mathscr{A}$ und insgesamt $\sigma(E) = \mathscr{A}$ wie gewünscht.

Beispiel 2.14. Wähle in Satz 2.13 für alle $i \in \{1,\dots,n\}$ $(\Omega_i, \mathscr{A}_i) := (\mathbb{R}, \mathscr{B}^1)$ sowie $M_i := \mathscr{I}^1$. Dann gilt:

$$\mathscr{I}^n = \left\{ \bigtimes_{i=1}^n N_i \;\middle|\; N_i \in \mathscr{I}^1 \right\}.$$

Da $\mathscr{I}^n$ ein Erzeuger für $\mathscr{B}^n$ ist, gilt somit:

$$\mathscr{B}^n = \bigotimes_{i=1}^n \mathscr{B}^1.$$

Definition 2.15. Seien $(\Omega, \mathscr{A})$ ein Messraum, $\Omega_0 \subseteq \Omega$ mit $\Omega_0 \neq \emptyset$. Man erhält die natürliche Einbettung

$$\iota\colon \Omega_0 \to \Omega, \quad \omega \mapsto \omega.$$

Dann heißt

$$\sigma(\iota) = \iota^{-1}(\mathscr{A}) = \{\Omega_0 \cap A \mid A \in \mathscr{A}\}$$

die *Spur-σ-Algebra* (engl.: trace σ-field, trace σ-algebra).

Die folgende Definition ermöglicht es uns, ein Maß mittels einer messbaren Abbildung von einem Maßraum auf einen Messraum zu transportieren.

Definition 2.16. Sei $(\Omega, \mathscr{A}, \mu)$ ein Maßraum und $(\Omega', \mathscr{A}')$ ein Messraum. Betrachte eine $(\mathscr{A}, \mathscr{A}')$-messbare Abbildung $f\colon \Omega \to \Omega'$. Dann definiert

$$\begin{aligned} \mu^f\colon \mathscr{A}' &\to [0, \infty] \\ A' &\mapsto \mu\big(f^{-1}(A')\big) \end{aligned}$$

ein Maß auf $(\Omega', \mathscr{A}')$ und wird *Bildmaß* (engl.: pushforward measure) von μ unter f genannt.

Definition 2.17. Sei $(\Omega, \mathscr{A}, P)$ ein Wahrscheinlichkeitsraum. Wir bezeichnen eine $(\mathscr{A}, \mathscr{B}^k)$-messbare Abbildung $X\colon \Omega \to \mathbb{R}^k$ als *Zufallsvariable* (engl.: random variable), falls $k = 1$. Für $k > 1$ heißt X *Zufallsvektor* (engl.: random vector). Das Maß P^X heißt *Verteilung* (engl.: distribution) von X.

Satz 2.18. *Zu jeder k-dimensionalen Verteilungsfunktion F existiert ein $\mathbb{R}^k$-wertiger Zufallsvektor X mit $F_X = F$, wobei*

$$F_X\colon \mathbb{R}^k \to [0,1], \quad x \mapsto P^X\big((-\infty, x]\big)$$

die Verteilungsfunktion von X bezeichnet.

Beweis. Wähle $X = \mathrm{id}_{\mathbb{R}^k}\colon (\mathbb{R}^k, \mathscr{B}^k, \mu_F) \to (\mathbb{R}^k, \mathscr{B}^k)$, wobei μ_F das zu F gehörige Lebesgue-Stieltjes-Maß aus Satz 1.49 ist.

Notation 2.19. Falls F_X die Verteilungsfunktion von X mit Verteilung P^X ist, so schreiben wir $X \sim F_X$ oder $X \sim P^X$.

Bemerkung 2.20. Die Konstruktion in Satz 2.18 hat primär theoretische Bedeutung. Andere Konstruktionen sind oft nützlicher.

Beispiel 2.21. Sei F eine eindimensionale Verteilungsfunktion mit Verteilung P. Definiere die *linksstetige verallgemeinerte Inverse* von F,

$$F^{-1}\colon [0,1] \to \overline{\mathbb{R}}, \quad t \mapsto \inf\{x \in \mathbb{R} \mid F(x) \geqslant t\}.$$

Dann ist $F^{-1}(t) \leqslant x$ genau dann, wenn $t \leqslant F(x)$. Damit ist F^{-1} messbar und es gilt für alle $x \in \mathbb{R}$:

$$P\big(\{t \in [0,1] \mid F^{-1}(t) \leqslant x\}\big) = F(x).$$

Also liefert auch diese Konstruktion eine Zufallsvariable, wie sie in Satz 2.18 gesucht wird. In höheren Dimensionen ist diese Konstruktion aufwändiger.

Bemerkung 2.22 (stochastische Modellbildung). Anhand des folgenden einfachen Beispiels wollen wir die Situation zusammenfassen. Für $p \in [0,1]$ und $n \in \mathbb{N}$ sowie

$$X\colon (\Omega, \mathscr{A}) \to \Big(\{0,\dots,n\}, \mathscr{P}\big(\{0,\dots,n\}\big)\Big)$$

erklären wir die Verteilung von X als Binomialverteilung mit Parametern (n,p) (vergleiche Beispiel 1.46), das heißt:

$$P^X := B_{n,p} = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \delta_k.$$

Somit ist $X \sim B_{n,p}$ und wir bestimmen Wahrscheinlichkeiten mittels

$$P\Big(X^{-1}\big(\{k\}\big)\Big) = P\big(\{X = k\}\big) = P^X\big(\{k\}\big) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Die Existenz dieser Zufallsvariablen X als messbare Abbildung von einem geeigneten abstrakten Raum $(\Omega, \mathscr{A})$ in $\big(\{0,\dots,n\}, \mathscr{P}(\{0,\dots,n\})\big)$ ist hier trivial – wähle zum Beispiel $(\Omega, \mathscr{A}) = \big(\{0,\dots,n\}, \mathscr{P}(\{0,\dots,n\})\big)$ – und durch die Endlichkeit des Wertebereiches stets gewährleistet. Für das Beispiel der multivariaten Normalverteilung (vergleiche Beispiel 1.54) können wir Satz 2.18 bemühen. Die zugehörige Verteilungsfunktion hatten wir in Beispiel 1.46 angegeben. Hierüber könnten wir mit Beispiel 2.21 analog argumentieren. Beachte, dass in der stochastischen Modellbildung der Grundraum $(\Omega, \mathscr{A})$ meistens keine Rolle spielt. Er wird deshalb oft unterdrückt. Das heißt, wir setzen stillschweigend voraus, dass auf einem geeigneten Raum eine Zufallsvariable X mit Verteilung P^X existiert (und verweisen dabei gedanklich auf Satz 2.18).

Man könnte nun den Eindruck gewinnen, hier werde ein Formalismus entwickelt, nur um ihn dann wieder zu vergessen. In komplizierteren Situationen – wenn X beispielsweise Werte in Funktionenräumen annimmt (stochastische Prozesse) – kann diesem Vorgehen aber durchaus eine wichtige Bedeutung zukommen.

Definition 2.23. Sei $X = (X_1, \dots, X_k)$ ein Zufallsvektor. Für $j \in \{1,\dots,k\}$ heißt P^{X_j} *Marginalverteilung* (engl.: marginal distribution) von X.

Beispiel 2.24. Seien $(\Omega_i, \mathscr{A}_i)$ für $i \in \{1,2,3\}$ Messräume, $f_i\colon \Omega_i \to \Omega_{i+1}$ für $i \in \{1,2\}$ messbar und μ ein Maß auf $\mathscr{A}_1$. Das Bilden von Bildmaßen ist in folgendem Sinne transitiv:

$$\mu^{f_2 \circ f_1} = (\mu^{f_1})^{f_2}.$$

Betrachten wir zum Beispiel einen Zufallsvektor $X = (X_1, \dots, X_k)$ und die Projektionen $\pi_j\colon \mathbb{R}^k \to \mathbb{R}$ auf die j-te Koordinate, so folgt aus $X_j = \pi_j \circ X$ für die Marginalverteilung P^{X_j}:

$$P^{X_j} = P^{\pi_j \circ X} = (P^X)^{\pi_j}.$$

Somit erhalten wir die Marginalverteilung als Bildmaß von P^X unter π_j. Den zugrundeliegenden Wahrscheinlichkeitsraum $(\Omega, \mathscr{A}, P)$ benötigen wir hierfür also nicht.

Beispiel 2.25. Für $a \in \mathbb{R}^k$ betrachte die Translation

$$T_a\colon (\mathbb{R}^k, \mathscr{B}^k, \lambda^k) \to (\mathbb{R}^k, \mathscr{B}^k, \lambda^k)$$
$$x \mapsto x + a.$$

Diese ist stetig und somit messbar. Für alle $(x, y] \in \mathscr{I}^k$ gilt:

$$\lambda^k\Big(T_a^{-1}\big((x,y]\big)\Big) = \lambda^k\big((x-a, y-a]\big)$$
$$= \prod_{i=1}^k \big((y_i - a_i) - (x_i - a_i)\big) = \prod_{i=1}^k (y_i - x_i) = \lambda^k\big((x,y]\big).$$

Da $\sigma(\mathscr{I}^k) = \mathscr{B}^k$, folgt aus der Eindeutigkeit der Fortsetzung

$$(\lambda^k)^{T_a} = \lambda^k.$$

Das Borel-Lebesgue-Maß ist also translationsinvariant.

Bemerkung 2.26. Das Borel-Lebesgue-Maß λ^k auf $(\mathbb{R}^k, \mathscr{B}^k)$ ist durch seine Translationsinvarianz und Normierung, also

$$(\lambda^k)^{T_a} = \lambda^k \quad \text{und} \quad \lambda^k\big((0,1]^k\big) = 1,$$

bereits festgelegt.

Definition 2.27. Definiere $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$.

$$\overline{\mathscr{B}}^1 := \big\{B \subseteq \overline{\mathbb{R}} \,\big|\, B \cap \mathbb{R} \in \mathscr{B}^1\big\}$$

heißt *σ-Algebra der $\overline{\mathbb{R}}$-Borel-Mengen*. Wir erhalten:

$$\overline{\mathscr{B}}^1 = \mathscr{B}^1 \cup \big\{B \cup \{\infty\} \,\big|\, B \in \mathscr{B}^1\big\} \cup \big\{B \cup \{-\infty\} \,\big|\, B \in \mathscr{B}^1\big\} \cup \big\{B \cup \{-\infty, \infty\} \,\big|\, B \in \mathscr{B}^1\big\}.$$

Definition 2.28. Sei $(\Omega, \mathscr{A})$ ein Messraum. Eine $(\mathscr{A}, \overline{\mathscr{B}}^1)$-messbare Funktion heißt *messbar numerisch* (measurable numeric).

Um mit messbaren numerischen Funktionen sinnvoll rechnen zu können, setzen wir

$$\frac{1}{\infty} := \frac{1}{-\infty} := 0, \quad \frac{1}{0} := \infty, \quad 0 \cdot \infty := 0 \cdot (-\infty) := 0.$$

Notation 2.29. Seien $f, g\colon \Omega \to \overline{\mathbb{R}}$ Abbildungen. Schreibe:

$$\{f > a\} := \{\omega \in \Omega \mid f(\omega) > a\},$$
$$\{f = g\} := \{\omega \in \Omega \mid f(\omega) = g(\omega)\}$$

sowie analog für ähnliche Mengen.

Satz 2.30. *Sei* $f\colon (\Omega, \mathscr{A}) \to (\overline{\mathbb{R}}, \overline{\mathscr{B}})$ *eine Abbildung.* f *ist genau dann messbar numerisch, wenn* $\{f \leqslant a\} \in \mathscr{A}$ *für alle* $a \in \mathbb{R}$.

Beweis. Mit Lemma 2.6 bleibt zu zeigen, dass das Mengensystem

$$\overline{M} := \{[-\infty, a] \mid a \in \mathbb{R}\}$$

ein Erzeuger von $\overline{\mathscr{B}}$ ist. Es gilt:

$$(a, b] = [-\infty, b] \backslash [-\infty, a].$$

Somit ist $\mathscr{I}^1 \subseteq \sigma(\overline{M})$, also auch $\mathscr{B}^1 \subseteq \sigma(\overline{M})$. Außerdem gilt:

$$\{\infty\} = \bigcap_{n \in \mathbb{N}} [-\infty, n]^c \quad \text{und} \quad \{-\infty\} = \bigcap_{n \in \mathbb{N}} [-\infty, -n].$$

Folglich ist $\sigma(\overline{M}) = \overline{\mathscr{B}}$.

Lemma 2.31. *Sei* $f\colon (\Omega, \mathscr{A}) \to (\overline{\mathbb{R}}, \overline{\mathscr{B}})$ *eine Abbildung. Folgende Aussagen sind äquivalent:*

1. f *ist messbar numerisch,*
2. $\{f \geqslant a\} \in \mathscr{A}$ *für alle* $a \in \mathbb{R}$,
3. $\{f > a\} \in \mathscr{A}$ *für alle* $a \in \mathbb{R}$,
4. $\{f \leqslant a\} \in \mathscr{A}$ *für alle* $a \in \mathbb{R}$,
5. $\{f < a\} \in \mathscr{A}$ *für alle* $a \in \mathbb{R}$.

Beweis. Mit Satz 2.30 bleibt zu zeigen, dass die letzten vier Aussagen äquivalent sind. Für alle $a \in \mathbb{R}$ gilt:

$$\{f > a\} = \bigcup_{n \in \mathbb{N}} \{f \geqslant a + 1/n\}, \quad \{f \leqslant a\} = \{f > a\}^c,$$
$$\{f < a\} = \bigcup_{n \in \mathbb{N}} \{f \leqslant a - 1/n\}, \quad \{f \geqslant a\} = \{f < a\}^c.$$

Da $\mathscr{A}$ eine σ-Algebra ist, folgt die Behauptung.

Lemma 2.32. *Seien* $f, g\colon (\Omega, \mathscr{A}) \to (\overline{\mathbb{R}}, \overline{\mathscr{B}})$ *messbar numerisch. Dann sind die Mengen*

$$\{f < g\}, \{f \leqslant g\}, \{f = g\}, \{f \neq g\}$$

messbar*, das heißt in* $\mathscr{A}$.

Beweis. Mengen der Form $\{f < a\}$ sind Urbilder von Mengen der Form $[-\infty, a) \in \overline{\mathscr{B}}$.

- $\{f < g\} = \bigcup_{q \in \mathbb{Q}} \big(\{f < q\} \cap \{q < g\}\big) \in \mathscr{A}$, da $\mathbb{Q}$ dicht in $\mathbb{R}$ liegt.

- $\{f \leqslant g\} = \{g < f\}^c \in \mathscr{A}$.
- $\{f = g\} = \{f \leqslant g\} \cap \{g \leqslant f\} \in \mathscr{A}$.
- $\{f \neq g\} = \{f = g\}^c \in \mathscr{A}$.

Lemma 2.33. *Seien* $f, g\colon (\Omega, \mathscr{A}) \to (\overline{\mathbb{R}}, \overline{\mathscr{B}})$ *messbar numerisch.*

1. *Dann sind auch*
$$f + g,\ f - g \text{ und } f \cdot g$$
messbar numerisch. Hierbei setzen wir voraus, dass Konstellationen wie $(\infty - \infty)$ *nicht auftreten.*

2. *Ferner ist*
$$(f, g)\colon \Omega \to \overline{R}^2, \quad \omega \mapsto \big(f(\omega), g(\omega)\big)$$
$(\mathscr{A}, \overline{\mathscr{B}} \otimes \overline{\mathscr{B}})$*-messbar.*

[Übung]

Lemma 2.34. *Seien* $f, f_1, f_2, \ldots : (\Omega, \mathscr{A}) \to (\overline{\mathbb{R}}, \overline{\mathscr{B}})$ *messbare numerische Funktionen. Dann sind auch folgende Funktionen messbar numerisch:*

1. $\sup_{n\in\mathbb{N}} f_n$, $\inf_{n\in\mathbb{N}} f_n$,
2. $\limsup_{n\in\mathbb{N}} f_n = \inf_{n\in\mathbb{N}} \sup_{k\geqslant n} f_k$, $\liminf_{n\in\mathbb{N}} f_n = \sup_{n\in\mathbb{N}} \inf_{k\geqslant n} f_k$,
3. $f^+ = \max\{f, 0\}$ *(Positivteil),* $f^- = -\min\{f, 0\}$ *(Negativteil),* $|f|$.

Beweis.

1. Da für alle $n \in \mathbb{N}$ f_n messbar numerisch ist, gilt für alle $a \in \mathbb{R}$:

$$\left\{\sup_{n\in\mathbb{N}} f_n \leqslant a\right\} = \bigcap_{n\in\mathbb{N}} \{f_n \leqslant a\} \in \mathscr{A}.$$

   Folglich ist $\sup_{n\in\mathbb{N}} f_n$ messbar numerisch. Mit Lemma 2.33 ist dann auch

$$\inf_{n\in\mathbb{N}} f_n = -\sup_{n\in\mathbb{N}}(-f_n)$$

   messbar numerisch.

2. Wie gerade gezeigt wurde, ist $\sup_{n\in\mathbb{N}} f_n$ messbar numerisch und damit natürlich auch $\sup_{k=n}^{\infty} f_k$. Folglich sind auch $\inf_{n\in\mathbb{N}} \sup_{k\geqslant n} f_k$ und $\sup_{n\in\mathbb{N}} \inf_{k\geqslant n} f_k$ messbar numerisch.

3. Da f messbar numerisch ist, gilt für alle $a \in \mathbb{R}$:

$$\left\{f^+ \leqslant a\right\} = \{\max\{f, 0\} \leqslant a\} = \{f \leqslant a\} \cap \{0 \leqslant a\} \in \mathscr{A}.$$

   Folglich ist f^+ messbar numerisch. Mit Lemma 2.33 ist dann auch $f^- = f^+ - f$ und somit auch $|f| = f^+ + f^-$ messbar numerisch.

Übungsaufgaben:

Übung 1: Beweisen Sie Lemma 2.11.

Übung 2: Beweisen Sie Lemma 2.33.

Übung 3: Sei (Θ, d) ein separabler metrischer Raum und sei $(\Omega, \mathscr{A})$ ein Messraum. Sei ferner $f\colon \Omega \times \Theta \to \overline{\mathbb{R}}$. Für alle $\theta \in \Theta$ sei $\omega \mapsto f(\omega, \theta)$ $(\mathscr{A}, \overline{\mathscr{B}})$-messbar. Für alle $\omega \in \Omega$ sei $\theta \mapsto f(\omega, \theta)$ stetig. Dann ist $\omega \mapsto \sup_{\theta \in \Theta} f(\omega, \theta)$ $(\mathscr{A}, \overline{\mathscr{B}})$-messbar.

Übung 4: Sei $f\colon \mathbb{R} \to \mathbb{R}$ differenzierbar. Beweisen oder widerlegen Sie: Die Ableitung $f'\colon \mathbb{R} \to \mathbb{R}$ ist borel-messbar.

Übung 5: Beweisen oder widerlegen Sie: Eine monotone Funktion $f\colon \mathbb{R} \to \mathbb{R}$ ist borel-messbar.

Lernziel-Kontrolle:

Frage 1: Wann ist eine Abbildung messbar?

Frage 2: Kennen Sie Beispiele für nicht messbare Abbildungen?

Frage 3: Welcher Zusammenhang besteht zwischen Messbarkeit und Stetigkeit einer Abbildung?

Frage 4: Was ist eine Zufallsvariable und was ist deren Verteilung?

Frage 5: Was ist ein Bildmaß?

Frage 6: Wie lauten die Rechenregeln für messbare numerische Funktionen?

Frage 7: Seien $(\Omega_1, \mathscr{A}_1)$ und $(\Omega_2, \mathscr{A}_2)$ Messräume. Unter welchen Bedingungen an $\mathscr{A}_1$ und $\mathscr{A}_2$ sind sämtliche Abbildungen $f\colon \Omega_1 \to \Omega_2$ messbar?

Frage 8: Wie ist eine Produkt-σ-Algebra definiert und wie erhält man einen Erzeuger?

Kapitel 3

Das Maßintegral

In diesem Kapitel beschäftigen wir uns mit dem Begriff des Maßintegrals. Unser Ziel ist es, für messbare Funktionen $f\colon \Omega \to \mathbb{R}$ Integrale zu definieren. Hierbei werden wir uns mehrfach des Prinzips der **algebraischen Induktion** bedienen, welches den Aufbau des allgemeinen Maßintegrals bestimmt. Viele Beweise von Aussagen über das Maßintegral verlaufen nach diesem Prinzip:

1. Betrachte zunächst nur Elementarfunktionen, das heißt bestimmte Treppenfunktionen, und weise die Aussage für diese nach.

2. Betrachte dann nichtnegative Funktionen, die sich durch Elementarfunktionen approximieren lassen, und weise die Aussage auch für diese nach.

3. Betrachte abschließend allgemeine messbare Funktionen und zerlege diese in ihren Positiv- und Negativteil: $f = f^+ - f^-$.

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

Definition 3.1. $\mathscr{E} := \{f\colon \Omega \to \mathbb{R} \mid f \geqslant 0, f \text{ messbar}, |f(\Omega)| < \infty\}$ heißt die Menge der *Elementarfunktionen* (engl.: simple functions) über Ω. Ist es notwendig, zwischen Elementarfunktionen auf verschiedenen Grundräumen zu unterscheiden, so schreiben wir $\mathscr{E}(\Omega)$ für die Elementarfunktionen auf Ω.

Definition 3.2. Seien $f \in \mathscr{E}$ und $\alpha_1, \ldots, \alpha_n \in [0, \infty)$. Sind $A_1, \ldots, A_n \in \mathscr{A}$ paarweise disjunkt mit $\sum_{j=1}^n A_j = \Omega$, so heißt

$$f = \sum_{j=1}^n \alpha_j \mathbb{1}_{A_j}$$

Normaldarstellung von f.

Bemerkung 3.3. Die folgende Form einer Normaldarstellung existiert für jedes $f \in \mathscr{E}$. Seien $f(\Omega) = \{\alpha_1, \dots, \alpha_n\} \subseteq [0, \infty)$ und $A_j := f^{-1}(\{\alpha_j\})$ für alle $j \in \{1, \dots, n\}$. Dann ist

$$f = \sum_{j=1}^{n} \alpha_j \mathbb{1}_{A_j}$$

eine Normaldarstellung von f.

Lemma 3.4. *Für zwei verschiedene Normaldarstellungen*

$$f = \sum_{i=1}^{n} \alpha_i \mathbb{1}_{A_i} = \sum_{j=1}^{m} \beta_j \mathbb{1}_{B_j} \tag{3.1}$$

gilt:

$$\sum_{i=1}^{n} \alpha_i \mu(A_i) = \sum_{j=1}^{m} \beta_j \mu(B_j).$$

Beweis. Es ist $\Omega = \sum_{i=1}^{n} A_i = \sum_{j=1}^{m} B_j$. Falls $\mu(A_i \cap B_j) \neq 0$, so gilt: $A_i \cap B_j \neq \varnothing$. Mit (3.1) folgt:

$$\alpha_i = \beta_j. \tag{3.2}$$

Mit $\mu(A_i) = \sum_{j=1}^{m} \mu(A_i \cap B_j)$ und (3.2) folgt:

$$\sum_{i=1}^{n} \alpha_i \mu(A_i) = \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_i \mu(A_i \cap B_j) = \sum_{j=1}^{m} \sum_{i=1}^{n} \beta_j \mu(A_i \cap B_j) = \sum_{j=1}^{m} \beta_j \mu(B_j).$$

Wir definieren nun den Begriff des Maßintegrals für Elementarfunktionen.

Definition 3.5. Sei $f = \sum_{j=1}^{n} \alpha_j \mathbb{1}_{A_j} \in \mathscr{E}$.

$$\mu(f) := \int f \,\mathrm{d}\mu := \sum_{j=1}^{n} \alpha_j \mu(A_j)$$

heißt das μ*-Integral* (engl.: μ-integral) von f.

Regeln 3.6 (Eigenschaften des μ-Integrals)**.** *Seien $f, g \in \mathscr{E}$. Dann gilt:*

1. *$\int \mathbb{1}_A \,\mathrm{d}\mu = \mu(A)$ für alle $A \in \mathscr{A}$ (Inhaltstreue).*
2. *$\int \alpha f \,\mathrm{d}\mu = \alpha \int f \,\mathrm{d}\mu$ für alle $\alpha \geqslant 0$ (Homogenität).*
3. *$\int f + g \,\mathrm{d}\mu = \int f \,\mathrm{d}\mu + \int g \,\mathrm{d}\mu$ (Additivität).*
4. *Falls $f \leqslant g$, so ist $\int f \,\mathrm{d}\mu \leqslant \int g \,\mathrm{d}\mu$ (Monotonie).*

Beweis.

1. Für alle $A \in \mathscr{A}$ ist $\mathbb{1}_A = 1 \cdot \mathbb{1}_A + 0 \cdot \mathbb{1}_{A^c}$ eine Normaldarstellung. Mit der Definition des Integrals folgt:

$$\int \mathbb{1}_A \, \mathrm{d}\mu = 1 \cdot \mu(A) + 0 \cdot \mu(A^c).$$

2. Für alle $\alpha \geqslant 0$ gilt:

$$\int \alpha f \, \mathrm{d}\mu = \sum_{j=1}^{n} \alpha\alpha_j \mu(A_j) = \alpha \sum_{j=1}^{n} \alpha_j \mu(A_j) = \alpha \int f \, \mathrm{d}\mu.$$

3. Schreibe f und g in Normaldarstellung:

$$f = \sum_{i=1}^{n} \alpha_i \mathbb{1}_{A_i}, \quad g = \sum_{j=1}^{m} \beta_j \mathbb{1}_{B_j}.$$

Dann hat $f + g$ die Normaldarstellung

$$f + g = \sum_{i,j} (\alpha_i + \beta_j) \mathbb{1}_{A_i \cap B_j}.$$

Folglich gilt:

$$\mu(f + g) = \sum_{i,j} (\alpha_i + \beta_j) \mu(A_i \cap B_j) = \sum_{i=1}^{n} \alpha_i \mu(A_i) + \sum_{j=1}^{m} \beta_j \mu(B_j) = \mu(f) + \mu(g).$$

4. Sei $f \leqslant g$. Dann ist wegen

$$f = \sum_{i,j} \alpha_i \mathbb{1}_{A_i \cap B_j}, \quad g = \sum_{i,j} \beta_j \mathbb{1}_{A_i \cap B_j}$$

$\alpha_i \leqslant \beta_j$ für alle $i \in \{1, \ldots, n\}, j \in \{1, \ldots, m\}$ mit $A_i \cap B_j \neq \varnothing$. Folglich ist

$$\mu(f) = \sum_{i,j} \alpha_i \mu(A_i \cap B_j) \leqslant \sum_{i,j} \beta_j \mu(A_i \cap B_j) = \mu(g).$$

Beispiel 3.7.

1. *Erwartungswert:* Sei $(\Omega, \mathscr{A}, P)$ ein Wahrscheinlichkeitsraum. Sei $X \colon \Omega \to [0, \infty)$ eine Zufallsvariable mit $X(\Omega) = \{x_1, \ldots, x_n\}$ (also $X \in \mathscr{E}$) und $A_j := X^{-1}(x_j)$ für alle $j \in \{1, \ldots, n\}$. Dann existiert der *Erwartungswert* von X und beträgt

$$E_P(X) := \int X \, \mathrm{d}P = \sum_{j=1}^{n} x_j P(A_j) = \sum_{j=1}^{n} x_j P(X = x_j).$$

2. *Summation ist Integration:* Seien $\Omega = \{1, \dots, n\}$, $X(i) := x_i \in \mathbb{R}$ für alle $i \in \Omega$ und $\mu = \sum_{i=1}^n \delta_i$ ein Zählmaß. Dann gilt:
$$\sum_{i=1}^n x_i = \int X \,\mathrm{d}\mu.$$

Wir wollen nun den Integralbegriff auf geeignete Funktionen ausweiten.

Definition 3.8.
$$\mathscr{E}^* := \left\{ f : \Omega \to \overline{\mathbb{R}} \,\middle|\, f \geqslant 0, f \text{ messbar} \right\}$$
heißt der *Raum der nichtnegativen, messbaren, numerischen Funktionen.*

Satz 3.9. *Sei $f \in \mathscr{E}^*$. Dann gibt es eine isotone Folge $(u_n)_{n\in\mathbb{N}}$ in $\mathscr{E}$, welche punktweise aufsteigend gegen f konvergiert. Wir schreiben in diesem Fall: $u_n \nearrow f$.*

Beweis. Wir wollen die Identität $\mathrm{id}_{[0,\infty)}$ durch Elementarfunktionen isoton approximieren. Dazu definieren wir
$$\alpha_n : \mathbb{R} \to \mathbb{R}, \qquad x \mapsto \begin{cases} 0 & x < 0 \\ j/2^n & j/2^n \leqslant x < (j+1)/2^n \text{ für alle } j \in \{0, \dots, n \cdot 2^n - 1\} \\ n & x \geqslant n \end{cases}.$$

Dann ist $\alpha_n \in \mathscr{E}$ und $\alpha_n(x) \nearrow x$ und für alle $x \in [0, \infty)$. Sei $u_n := \alpha_n \circ f$ für alle $n \in \mathbb{N}$. Es gilt:
$$u_n - f = (\alpha_n \circ f) - (\mathrm{id} \circ f) = (\alpha_n - \mathrm{id}) \circ f.$$
Da $(\alpha_n - \mathrm{id}) \nearrow 0$, gilt: $u_n \nearrow f$. Auf kompakten Teilmengen, oder falls f beschränkt ist, ist diese Konvergenz sogar gleichmäßig.

Lemma 3.10. *Seien $(u_n)_{n\in\mathbb{N}}$ eine isotone Folge in $\mathscr{E}$ und $v \in \mathscr{E}$ mit $v \leqslant \lim_{n\to\infty} u_n$. Dann gilt:*
$$\int v \,\mathrm{d}\mu \leqslant \lim_{n\to\infty} \int u_n \,\mathrm{d}\mu.$$

Beweis. Da $v \in \mathscr{E}$, schreibe $v = \sum_{j=1}^m \alpha_j \mathbb{1}_{A_j}$. Seien $c \in (0,1)$ und $B_n := \{u_n \geqslant cv\}$ für alle $n \in \mathbb{N}$. Dann gilt wegen der Regeln 3.6.2. und 3.6.4.:
$$\mu(u_n) \geqslant c\mu(v\mathbb{1}_{B_n}) \quad \text{für alle } n \in \mathbb{N}. \tag{3.3}$$

Nach Definition von B_n gilt: $B_n \nearrow \Omega$, also auch $A_j \cap B_n \nearrow A_j$ für alle $j \in \{1, \dots, m\}$. Mit der Stetigkeit von unten folgt:
$$\mu(v) = \sum_{j=1}^m \alpha_j \mu(A_j) = \lim_{n\to\infty} \sum_{j=1}^m \alpha_j \mu(A_j \cap B_n) = \lim_{n\to\infty} \mu(v\mathbb{1}_{B_n}).$$

Mit (3.3) folgt:
$$\lim_{n\to\infty} \mu(u_n) \geqslant c\mu(v) \quad \text{für alle } c \in (0,1).$$

Korollar 3.11. *Seien* $(u_n)_{n\in\mathbb{N}}$ *und* $(v_n)_{n\in\mathbb{N}}$ *monoton steigende Folgen in* $\mathscr{E}$*, sodass gilt:*

$$\lim_{n\to\infty} u_n = \lim_{n\to\infty} v_n.$$

Dann folgt:

$$\lim_{n\to\infty} \int u_n \,\mathrm{d}\mu = \lim_{n\to\infty} \int v_n \,\mathrm{d}\mu.$$

Beweis. Für alle $k \in \mathbb{N}$ gilt:

$$v_k \leqslant \lim_{n\to\infty} u_n, \quad u_k \leqslant \lim_{n\to\infty} v_n.$$

Mit Lemma 3.10 folgt:

$$\int v_k \,\mathrm{d}\mu \leqslant \lim_{n\to\infty} \int u_n \,\mathrm{d}\mu, \quad \int u_k \,\mathrm{d}\mu \leqslant \lim_{n\to\infty} \int v_n \,\mathrm{d}\mu.$$

Durch Limesbildung für $k \to \infty$ folgt die Behauptung.

Wir weiten nun den Begriff des Integrals auf Funktionen in $\mathscr{E}^*$ aus.

Definition 3.12. Seien $f \in \mathscr{E}^*$ und $(u_n)_{n\in\mathbb{N}}$ eine Folge in $\mathscr{E}$ mit $u_n \nearrow f$.

$$\mu(f) := \int f \,\mathrm{d}\mu := \lim_{n\to\infty} \int u_n \,\mathrm{d}\mu$$

heißt das μ-*Integral* von f.

Bemerkung 3.13.

1. Wegen Satz 3.9 und Korollar 3.11 ist $\mu(f)$ wohldefiniert.
2. Alle Eigenschaften des μ-Integrals auf $\mathscr{E}$ (siehe Regeln 3.6) übertragen sich auf das μ-Integral auf $\mathscr{E}^*$. Für $f, g \in \mathscr{E}^*$ mit $u_n \nearrow f$ und $v_n \nearrow g$ gilt beispielsweise:

$$\mu(f+g) = \lim_{n\to\infty} \mu(u_n + v_n) = \lim_{n\to\infty} \mu(u_n) + \lim_{n\to\infty} \mu(v_n) = \mu(f) + \mu(g).$$

3. Für $f \in \mathscr{E}^*$ gilt:

$$\int f \,\mathrm{d}\mu = \lim_{n\to\infty} \left(\sum_{j=0}^{n\cdot 2^n - 1} \frac{j}{2^n} \mu\left(\frac{j}{2^n} \leqslant f < \frac{j+1}{2^n} \right) + n \cdot \mu(f \geqslant n) \right).$$

Satz 3.14 (Beppo Levi, Monotone Konvergenz)**.** *Sei* $(f_n)_{n\in\mathbb{N}}$ *eine isotone Folge in* $\mathscr{E}^*$*. Dann gilt:*

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

Beweis. Es ist $f := \lim_{n\to\infty} f_n \geqslant f_m$ für alle $m \in \mathbb{N}$. Folglich ist

$$\limsup_{n\to\infty} \mu(f_n) \leqslant \mu(f). \tag{3.4}$$

Wir wollen nun f in geeigneter Weise durch Treppenfunktionen annähern. Für alle $n \in \mathbb{N}$ sei $(u_{n,k})_{k\in\mathbb{N}}$ eine Folge in $\mathscr{E}$, sodass für $k \to \infty$ gilt: $u_{n,k} \nearrow f_n$. Weiterhin sei $v_k := \sup\{u_{1,k}; u_{2,k}; \dots ; u_{k,k}\}$ für alle $k \in \mathbb{N}$. Da $(v_k)_{k\in\mathbb{N}}$ in $\mathscr{E}$ isoton ist mit $v_k \leqslant f_k$ für alle $k \in \mathbb{N}$, gilt:

$$\lim_{k\to\infty} v_k \leqslant f. \tag{3.5}$$

Andererseits ist $u_{n,k} \leqslant v_k$ für alle $n \leqslant k$. Folglich gilt für alle $n \in \mathbb{N}$:

$$\lim_{k\to\infty} u_{n,k} = f_n \leqslant \lim_{k\to\infty} v_k.$$

und somit

$$f \leqslant \lim_{k\to\infty} v_k.$$

Mit (3.5) folgt Gleichheit:

$$f = \lim_{k\to\infty} v_k.$$

Da $(v_k)_{k\in\mathbb{N}}$ in $\mathscr{E}$ isoton und das Integral wohldefiniert ist, folgt mit (3.4):

$$\mu(f) = \lim_{k\to\infty} \mu(v_k) \leqslant \liminf_{n\to\infty} \mu(f_n) \leqslant \mu(f).$$

Korollar 3.15. *Sei $(f_n)_{n\in\mathbb{N}}$ eine isotone Folge in $\mathscr{E}^*$. Dann gilt:*

$$\int \sum_{n\in\mathbb{N}} f_n \,\mathrm{d}\mu = \sum_{n\in\mathbb{N}} \int f_n \,\mathrm{d}\mu.$$

Beweis. Sei $(u_k)_{k\in\mathbb{N}}$ mit $u_k := \sum_{n=1}^{k} f_n$ für alle $k \in \mathbb{N}$. $(u_k)_{k\in\mathbb{N}}$ ist eine isotone Folge in $\mathscr{E}$. Mit dem Satz über die monotone Konvergenz folgt:

$$\begin{aligned} \int \sum_{n\in\mathbb{N}} f_n \,\mathrm{d}\mu &= \int \lim_{k\to\infty} u_k \,\mathrm{d}\mu = \lim_{k\to\infty} \int u_k \,\mathrm{d}\mu \\ &= \lim_{k\to\infty} \int \sum_{n=1}^{k} f_n \,\mathrm{d}\mu = \lim_{k\to\infty} \sum_{n=1}^{k} \int f_n \,\mathrm{d}\mu = \sum_{n\in\mathbb{N}} \int f_n \,\mathrm{d}\mu. \end{aligned}$$

Wir erweitern nun den Begriff des Integrals auf allgemeine messbare numerische Funktionen.

Definition 3.16. Eine messbare Abbildung $f\colon \Omega \to \overline{\mathbb{R}}$ heißt *μ-integrierbar*, falls

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

$$\mu(f) := \int f \,\mathrm{d}\mu := \int f^+ \,\mathrm{d}\mu - \int f^- \,\mathrm{d}\mu$$

heißt *μ-Integral* von f. Wir schreiben auch

$$\int f \,\mathrm{d}\mu = \int_\Omega f(\omega)\mu(\mathrm{d}\omega).$$

Definition 3.17. Sei $(\Omega, \mathscr{A}, P)$ ein Wahrscheinlichkeitsraum und sei $X\colon \Omega \to \overline{\mathbb{R}}$ eine Zufallsvariable, das heißt $(\mathscr{A}, \overline{\mathscr{B}})$-messbar. Unter der Annahme, dass die hier auftretenden Integrale existieren, definieren wir den *Erwartungswert* (engl.: expected value, expectation value, mean) von X als

$$\mathrm{E}X := \int X \,\mathrm{d}P$$

und die *Varianz* (engl.: variance) von X als

$$\mathrm{Var}(X) := \mathrm{E}\big((X - \mathrm{E}X)^2\big).$$

$\sigma_X := \sqrt{\mathrm{Var}(X)}$ heißt *Streuung* oder *Standardabweichung* (engl.: standard deviation) von X. Sei $Y\colon \Omega \to \overline{\mathbb{R}}$ eine weitere Zufallsvariable auf $(\Omega, \mathscr{A}, P)$. Wir definieren die *Kovarianz* von X und Y als

$$\mathrm{Cov}(X, Y) := \mathrm{E}\big((X - \mathrm{E}X)(Y - \mathrm{E}Y)\big).$$

Analog definieren wir für einen Zufallsvektor $X = (X_1, \dots, X_k)$ den *Erwartungswertvektor* als

$$\mathrm{E}X := (\mathrm{E}X_1, \dots, \mathrm{E}X_k).$$

Etwas genauer definieren wir $X_i := \pi_i \circ X$, wobei $\pi_i\colon \mathbb{R}^k \to \mathbb{R}$ die Projektion auf die i-te Koordinate ist, und erhalten:

$$\mathrm{E}X_i = \int \pi_i \circ X \,\mathrm{d}P.$$

Weiterhin definieren wir die *Kovarianzmatrix* von X als

$$\mathrm{Cov}(X) := \big(\mathrm{Cov}(X_i, X_j)\big)_{i,j=1}^{k}.$$

Bemerkung 3.18.

1. f heißt *quasi-integrierbar*, falls $\int f^+ \, \mathrm{d}\mu < \infty$ oder $\int f^- \, \mathrm{d}\mu < \infty$.

2. Sei X eine quasi-integrierbare Zufallsvariable. Ist $\int X^- \, \mathrm{d}P = \infty$, so schreiben wir für den Erwartungswert: $\mathrm{E}X = -\infty$. Analog schreiben wir $\mathrm{E}X = \infty$, falls $\int X^+ \, \mathrm{d}P = \infty$.

3. Sei $(\Omega, \mathscr{A}, P)$ ein Wahrscheinlichkeitsraum und $X\colon \Omega \to \overline{\mathbb{R}}$ eine Zufallsvariable. Dann existiert der Erwartungswert $\mathrm{E}X$, falls X^+ und X^- P-integrierbar sind, das heißt falls $\mathrm{E}\,|X| < \infty$.

4. Die Eigenschaften des μ-Integrals (siehe Regeln 3.6) übertragen sich auf das μ-Integral für messbare numerische Funktionen. Beachte, dass nun auch für $\alpha < 0$ gilt:
$$\int \alpha f \, \mathrm{d}\mu = \alpha \int f \, \mathrm{d}\mu.$$

Beispiel 3.19. Es gelten die folgenden Aussagen. [Übung]

1. Für $\omega \in \Omega$ sei δ_ω das Dirac-Maß. Eine Abbildung $f\colon \Omega \to \overline{\mathbb{R}}$ ist genau dann δ_ω-integrierbar, wenn $|f(\omega)| < \infty$. In diesem Fall gilt: $\int f \, \mathrm{d}\delta_\omega = f(\omega)$.

2. Seien $(\mu_n)_{n\in\mathbb{N}}$ eine Folge von Maßen, $\mu := \sum_{n\in\mathbb{N}} \mu_n$ und $f\colon \Omega \to \overline{\mathbb{R}}$ messbar. So ist f genau dann μ-integrierbar, wenn $\sum_{n\in\mathbb{N}} \int |f| \, \mathrm{d}\mu_n < \infty$. In diesem Fall ist
$$\int f \, \mathrm{d}\mu = \sum_{n\in\mathbb{N}} \int f \, \mathrm{d}\mu_n.$$
Der Beweis lässt sich mit algebraischer Induktion führen.

3. Sei $(\Omega, \mathscr{A}) := (\mathbb{N}, \mathscr{P}(\mathbb{N}))$ und μ das Zählmaß. Sei ferner $(a_n)_{n\in\mathbb{N}}$ eine reelle Folge. Betrachte die Abbildung $f\colon \mathbb{N} \to \mathbb{R}$ mit $f(n) := a_n$ für alle $n \in \mathbb{N}$. f ist genau dann μ-integrierbar, wenn $\sum_{n\in\mathbb{N}} |a_n| < \infty$. In diesem Fall ist
$$\int f \, \mathrm{d}\mu = \sum_{n\in\mathbb{N}} a_n.$$

4. Seien $(x_{n,k})_{n,k\in\mathbb{N}}$ eine nichtnegative Doppelfolge und $(x_k)_{k\in\mathbb{N}}$ eine nichtnegative Folge mit $x_{n,k} \nearrow x_k$. Dann gilt:
$$\lim_{n\to\infty} \sum_{k\in\mathbb{N}} x_{n,k} = \sum_{k\in\mathbb{N}} x_k.$$

5. Seien $\Omega := \{\omega_1, \omega_2, \ldots\}$ abzählbar, $\mathscr{A} := \mathscr{P}(\Omega)$ und $(p_i)_{i\in\mathbb{N}}$ eine nichtnegative Folge mit $\sum_{i\in\mathbb{N}} p_i = 1$. Dann ist
$$P := \sum_{i\in\mathbb{N}} p_i \delta_{\omega_i}$$

ein Wahrscheinlichkeitsmaß auf $\mathscr{A}$.

Sei $X\colon \Omega \to \overline{\mathbb{R}}$ eine Zufallsvariable. So existiert der Erwartungswert $\mathrm{E}X$ genau dann, wenn X P-integrierbar ist, das heißt wenn

$$\int |X| \, \mathrm{d}P = \sum_{i \in \mathbb{N}} |X(\omega_i)| \, p_i < \infty.$$

In diesem Fall ist

$$\mathrm{E}X = \sum_{i \in \mathbb{N}} X(\omega_i) p_i.$$

Wir erinnern uns, dass für eine messbare Abbildung $f\colon (\Omega, \mathscr{A}, \mu) \to (\Omega', \mathscr{A}')$ das Bildmaß definiert ist als $\mu^f(A') = \mu \circ f^{-1}(A')$ für alle $A' \in \mathscr{A}'$.

Satz 3.20 (Transformationssatz für Integrale)**.** *Betrachte einen Maßraum $(\Omega, \mathscr{A}, \mu)$, einen Messraum $(\Omega', \mathscr{A}')$ sowie eine messbare Abbildung $f\colon \Omega \to \Omega'$.*

1. *Sei $h\colon \Omega' \to \overline{\mathbb{R}}$ in $\mathscr{E}^*$ Dann gilt:*

$$\int h \, \mathrm{d}\mu^f = \int h \circ f \, \mathrm{d}\mu. \tag{3.6}$$

2. *Eine messbare Abbildung $h\colon \Omega' \to \overline{\mathbb{R}}$ ist genau dann μ^f-integrierbar, wenn $h \circ f$ μ-integrierbar ist. In diesem Fall gilt (3.6).*

Beweis.

1. Wir führen den Beweis mittels algebraischer Induktion.

   - Sei zunächst $h = \sum_{j=1}^n \alpha_j \mathbb{1}_{A'_j} \in \mathscr{E}$ mit $A'_j \in \mathscr{A}'$. Dann gilt:

$$\begin{aligned}\mu^f(h) &= \sum_{j=1}^n \alpha_j \mu^f(A'_j) = \sum_{j=1}^n \alpha_j \mu\big(f^{-1}(A'_j)\big) \\ &= \sum_{j=1}^n \alpha_j \mu\left(\mathbb{1}_{f^{-1}(A'_j)}\right) = \mu\left(\sum_{j=1}^n \alpha_j \mathbb{1}_{f^{-1}(A'_j)}\right) = \mu(h \circ f).\end{aligned}$$

     Hierbei gilt die letzte Gleichheit, da $f(x) \in A'_j$ genau dann erfüllt ist, wenn $x \in f^{-1}(A'_j)$ gilt.

   - Sei nun $h \in \mathscr{E}^*$. Sei $(u_n)_{n \in \mathbb{N}}$ eine Folge in $\mathscr{E}(\Omega')$ mit $u_n \nearrow h$. Dann ist $(u_n \circ f)_{n \in \mathbb{N}}$ eine Folge in $\mathscr{E}(\Omega)$. Aus $u_n \nearrow h$ folgt: $u_n \circ f \nearrow h \circ f$. Mit der Definition des Integrals und dem bereits Gezeigten ergibt sich:

$$\mu^f(h) = \lim_{n \to \infty} \mu^f(u_n) = \lim_{n \to \infty} \mu(u_n \circ f) = \mu(h \circ f).$$

2. Aus dem ersten Teil des Beweises wissen wir:

$$\mu^f(h^+) = \mu(h^+ \circ f), \quad \mu^f(h^-) = \mu(h^- \circ f).$$

Außerdem gilt:

$$(h \circ f)^+ = h^+ \circ f, \quad (h \circ f)^- = h^- \circ f.$$

h ist genau dann μ^f-integrierbar, wenn

$$\mu^f(h^+) = \mu\big((h \circ f)^+\big) < \infty, \quad \mu^f(h^-) = \mu\big((h \circ f)^-\big) < \infty,$$

was genau dann der Fall ist, wenn $h \circ f$ μ-integrierbar ist.

Beispiel 3.21 (Erwartungswert). Sei in Satz 3.20.2. $\Omega' := \mathbb{R}$, $f := X : \Omega \to \mathbb{R}$ eine Zufallsvariable und $h := \mathrm{id}_{\mathbb{R}}$ die Identität auf $\mathbb{R}$. Dann ist X genau dann P-integrierbar, wenn $\mathrm{id}_{\mathbb{R}}$ P^X-integrierbar ist. In diesem Fall gilt:

$$\mathrm{E}X = \int_\Omega X \,\mathrm{d}P = \int_\Omega \mathrm{id}_{\mathbb{R}} \circ X \,\mathrm{d}P = \int_{\mathbb{R}} \mathrm{id}_{\mathbb{R}} \,\mathrm{d}P^X = \int_{\mathbb{R}} t\, P^X(\mathrm{d}t).$$

Hieraus wird ersichtlich, dass $\mathrm{E}X$ nicht vom zugrundeliegenden Wahrscheinlichkeitsraum abhängt, sondern nur vom Bildmaß P^X. Zur Veranschaulichung werden wir nun den zweifachen Würfelwurf betrachten und den Erwartungswert der Augensumme auf zwei verschiedene Weisen bestimmen: Seien also $\Omega := \{1, \dots, 6\}^2$ und P die Gleichverteilung auf Ω, also $P(\{(i,j)\}) = 1/36$ für alle $i, j \in \{1, \dots, 6\}$. Sei ferner

$$X : \Omega \to \Omega' := \{2, \dots, 12\}, \qquad (i,j) \mapsto i + j.$$

- Dann ergibt sich direkt aus der Definition des Erwartungswertes:

$$\mathrm{E}X = \sum_{(i,j)\in\Omega} i + j\, P\big(\{(i,j)\}\big) = 7.$$

- Nun bestimmen wir den Erwartungswert über das Bildmaß. Wir berechnen $P^X(\{x\})$ für alle $x \in \Omega'$ und erhalten Tabelle 3.1.

x	2	3	4	5	6	7	8	9	10	11	12
$P^X(\{x\})$	$1/36$	$2/36$	$3/36$	$4/36$	$5/36$	$6/36$	$5/36$	$4/36$	$3/36$	$2/36$	$1/36$

Tabelle 3.1: Wahrscheinlichkeiten der Augensummen beim zweifachen Münzwurf

Dies lässt sich zu folgender Formel zusammenfassen:

$$P^X(\{x\}) = \frac{1}{36}\big(6 - |x - 7|\big) \qquad \text{für alle } x \in \Omega'.$$

Nach dem Transformationssatz für Integrale erhalten wir:

$$\mathrm{E}X = \int_{\Omega'} t\, P^X(\mathrm{d}t) = \sum_{t=2}^{12} \frac{t}{36}\big(6 - |t - 7|\big) = 7.$$

- Sei X eine normalverteilte Zufallsvariable mit Mittelwert $\mu \in \mathbb{R}$ und Varianz $\sigma^2 > 0$, das heißt

$$P^X\big((-\infty, t]\big) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{1}{2}\cdot\frac{(x-\mu)^2}{\sigma^2}}\, \mathrm{d}x.$$

Dann gilt:

$$\begin{aligned}
\mathrm{E}X &= \int_{-\infty}^{\infty} x \cdot \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{1}{2}\cdot\frac{(x-\mu)^2}{\sigma^2}}\, \mathrm{d}x \\
&= \int_{-\infty}^{\infty} \mu \cdot \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{1}{2}\cdot\frac{(x-\mu)^2}{\sigma^2}}\, \mathrm{d}x + \int_{-\infty}^{\infty} (x-\mu) \cdot \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{1}{2}\cdot\frac{(x-\mu)^2}{\sigma^2}}\, \mathrm{d}x \\
&= \mu \cdot \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{1}{2}\cdot\frac{(x-\mu)^2}{\sigma^2}}\, \mathrm{d}x + \left[\frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{1}{2}\cdot\frac{(x-\mu)^2}{\sigma^2}}\right]_{-\infty}^{\infty} \\
&= \mu \cdot P^X(\mathbb{R}) + 0 = \mu.
\end{aligned}$$

- Sei X eine Cauchy-verteilte Zufallsvariable mit *Zentrum* $t \in \mathbb{R}$ und *Skalenparameter* $s > 0$, das heißt

$$P^X\big((-\infty, t]\big) = \int_{-\infty}^{t} \frac{1}{\pi} \cdot \frac{s}{s^2 + (x-t)^2}\, \mathrm{d}x.$$

Dann gilt:

$$\begin{aligned}
&\int_{-\infty}^{\infty} \left(x \cdot \frac{1}{\pi} \cdot \frac{s}{s^2 + (x-t)^2}\right)^{+} \mathrm{d}x \\
=&\int_{0}^{\infty} x \cdot \frac{1}{\pi} \cdot \frac{s}{s^2 + (x-t)^2}\, \mathrm{d}x \\
=&\frac{1}{2}\int_{0}^{\infty} 2 \cdot (x-t) \cdot \frac{1}{\pi} \cdot \frac{s}{s^2 + (x-t)^2}\, \mathrm{d}x + \int_{0}^{\infty} t \cdot \frac{1}{\pi} \cdot \frac{s}{s^2 + (x-t)^2}\, \mathrm{d}x \\
=&\frac{1}{2}\int_{0}^{\infty} 2 \cdot (x-t) \cdot \frac{1}{\pi} \cdot \frac{s}{s^2 + (x-t)^2}\, \mathrm{d}x + \frac{t}{2} \\
=&\frac{1}{2} \cdot \frac{s}{\pi} \int_{s^2+t^2}^{\infty} \frac{1}{z}\, \mathrm{d}z + \frac{t}{2} \\
=&\frac{s}{2\pi}\big[\log(z)\big]_{s^2+t^2}^{\infty} + \frac{t}{2} = \infty,
\end{aligned}$$

wobei wir $z = s^2 + (x-t)^2$ substituiert haben. Analog gilt:

$$\int_{-\infty}^{\infty} \left(x \cdot \frac{1}{\pi} \cdot \frac{s}{s^2 + (x-t)^2}\right)^{-} \mathrm{d}x = -\infty.$$

Folglich existiert $\mathrm{E}X$ nicht.

Übungsaufgaben:

Übung 1: Beweisen Sie die fünf Teilaussagen aus Beispiel 3.19.

Lernziel-Kontrolle:

Frage 1: Was verbirgt sich hinter dem Begriff der algebraischen Induktion?

Frage 2: Überlegen Sie sich eine Darstellung einer Elementarfunktion, welche nicht die Normaldarstellung ist.

Frage 3: Was besagt der Satz von der monotonen Konvergenz?

Frage 4: Geben Sie ein Beispiel für eine messbare, aber nicht (quasi-)integrierbare Funktion an.

Frage 5: Was besagt der Transformationssatz für Integrale und worin liegt seine besondere Bedeutung für die Wahrscheinlichkeitstheorie?

Kapitel 4

Konvergenzsätze

4.1 Fast überall bestehende Eigenschaften

Für viele Belange in der Maßtheorie, beispielsweise die Integration, ist das Verhalten von Funktionen auf Nullmengen irrelevant. Zum Beispiel ändert sich das Integral über eine Funktion nicht, wenn wir diese auf einer Nullmenge abändern.

Im Folgenden betrachten wir nur vollständige Maßräume.

Definition 4.1. Sei $(\Omega, \mathscr{A}, \mu)$ ein Maßraum. Weiterhin sei E eine Aussage, sodass für alle $\omega \in \Omega$ definitiert ist, ob E für ω zutrifft oder nicht. Wir sagen: E gilt *μ-fast-überall* (engl.: μ-almost-everywhere), falls es eine μ-Nullmenge N gibt, sodass E für jedes $\omega \in N^c$ zutrifft. Wir schreiben E $[\mu]$.

Beispiel 4.2.

1. Seien $(\Omega, \mathscr{A}, \mu)$ ein Maßraum und $f, g \colon \Omega \to \overline{\mathbb{R}}$. Es ist $f = g$ $[\mu]$ genau dann, wenn es eine μ -Nullmenge N gibt, sodass $f(\omega) = g(\omega)$ für alle $\omega \in N^c$. Sind f und g außerdem messbar, so ist genau dann $f = g$ $[\mu]$, wenn $\mu\big(\{f \neq g\}\big) = 0$.

2. Betrachte $(\Omega, \mathscr{A}) := (\mathbb{R}, \mathscr{B})$ und die Funktion $f \colon \mathbb{R} \to [0, \infty)$, $x \mapsto x^2$. Dann ist $f \neq 0$ $[\lambda]$, aber $f = 0$ $[\delta_0]$.

Satz 4.3. *Seien $(\Omega, \mathscr{A}, \mu)$ ein Maßraum und $f, g \colon \Omega \to \overline{\mathbb{R}}$ messbar numerische Funktionen mit $f = g$ $[\mu]$. f ist genau dann μ-integrierbar, wenn g μ-integrierbar ist und es gilt:*

$$\int f \,\mathrm{d}\mu = \int g \,\mathrm{d}\mu.$$

Beweis. Aus $\{f^+ \neq g^+\} \cup \{f^- \neq g^-\} = \{f \neq g\} \in \mathscr{A}$ erhalten wir $f^+ = g^+$ $[\mu]$ und $f^- = g^-$ $[\mu]$. Wir können daher ohne Einschränkung $f \geqslant 0$ und $g \geqslant 0$ annehmen. Sei $h_n := n \cdot \mathbb{1}_{\{f \neq g\}} \in \mathscr{E}$. Dann ist $\mu(h_n) = n \cdot \mu\big(\{f \neq g\}\big) = 0$. Wegen $h_n \nearrow h$ mit $h := \infty \cdot \mathbb{1}_{\{f \neq g\}}$ folgt, dass $\mu(h) = 0$. Es gilt: $g \leqslant f + h$ und $f \leqslant g + h$. Wegen $\mu(h) = 0$ folgt die Behauptung.

Satz 4.4 (Markov-Ungleichung). *Seien $(\Omega, \mathscr{A}, \mu)$ ein Maßraum und $f\colon \Omega \to \overline{\mathbb{R}}$ integrierbar mit $f \geqslant 0$ und $t > 0$. Dann gilt:*

$$\mu\big(\{f \geqslant t\}\big) \leqslant \frac{1}{t} \int_{\{f \geqslant t\}} f \,\mathrm{d}\mu \leqslant \frac{1}{t} \int f \,\mathrm{d}\mu.$$

Beweis. Es gilt:

$$\mathbb{1}_{\{f \geqslant t\}} \leqslant \frac{1}{t} f \cdot \mathbb{1}_{\{f \geqslant t\}} \leqslant \frac{1}{t} f.$$

Mittels Integration folgt die Behauptung.

Korollar 4.5. *Seien $(\Omega, \mathscr{A}, P)$ ein Wahrscheinlichkeitsraum, $t > 0$, $\varepsilon > 0$ und $X\colon \Omega \to \overline{\mathbb{R}}$ eine Zufallsvariable.*

1. *Es gilt:*

$$P\big(|X| \geqslant t\big) \leqslant \frac{1}{t} \mathrm{E}\,|X|\,.$$

2. *Es gilt die Tschebyschev-Ungleichung:*

$$P\big(|X - \mathrm{E}X| \geqslant \varepsilon\big) \leqslant \frac{1}{\varepsilon^2} \mathrm{Var}(X).$$

Seien $(\Omega, \mathscr{A}, \mu)$ ein Maßraum und $f\colon \Omega \to \overline{\mathbb{R}}$ integrierbar. Dann gilt:

3. *Falls $f \geqslant 0$, so ist genau dann $\int f \,\mathrm{d}\mu = 0$, wenn $f = 0$ $[\mu]$.*

4. *Es ist $\mu\big(|f| = \infty\big) = 0$. Das heißt, $|f| < \infty$ $[\mu]$.*

Beweis.

1. Die Behauptung folgt aus der Markov-Ungleichung (Satz 4.4) mit $f := |X|$.
2. Dies folgt ebenso aus der Markov-Ungleichung mit $f := |X - \mathrm{E}X|^2$ und $t := \varepsilon^2$.
3. Sei zunächst $f = 0$ $[\mu]$. Wegen Satz 4.3 ist f integrierbar und es gilt:

$$\int f \,\mathrm{d}\mu = \int 0 \,\mathrm{d}\mu = 0. \tag{4.1}$$

Gelte nun (4.1). Es gilt:

$$\{f > 0\} = \bigcup_{n \in \mathbb{N}} \left\{ f \geqslant \frac{1}{n} \right\}.$$

Mit Satz 4.4 folgt:

$$\mu\big(\{f > 0\}\big) \leqslant \sum_{n \in \mathbb{N}} \mu \left(\left\{ f \geqslant \frac{1}{n} \right\} \right) \leqslant \sum_{n \in \mathbb{N}} n \int f \,\mathrm{d}\mu = 0.$$

4. Für alle $n \in \mathbb{N}$ gilt nach der Markov-Ungleichung:

$$\mu\big(\{|f| = \infty\}\big) \leqslant \mu\big(\{|f| \geqslant n\}\big) \leqslant \frac{1}{n} \int |f| \, \mathrm{d}\mu.$$

Damit gilt:

$$\mu\big(\{|f| = \infty\}\big) \leqslant \lim_{n\to\infty} \frac{1}{n} \int |f| \, \mathrm{d}\mu = 0.$$

4.2 Konvergenzsätze

Satz 4.6 (Lemma von Fatou)**.** *Seien $(\Omega, \mathscr{A}, \mu)$ ein Maßraum und $(f_n)_{n\in\mathbb{N}}$ eine Folge nichtnegativer messbar numerischer Funktionen $f_n\colon \Omega \to \overline{\mathbb{R}}$. Dann gilt:*

$$\int \liminf_{n\to\infty} f_n \, \mathrm{d}\mu \leqslant \liminf_{n\to\infty} \int f_n \, \mathrm{d}\mu.$$

Beweis. Für $n \in \mathbb{N}$ definiere $g_n := \inf_{k\geqslant n} f_k$. Somit ist $(g_n)_{n\in\mathbb{N}}$ eine isotone Folge und es gilt: $\liminf_{n\to\infty} f_n = \lim_{n\to\infty} g_n$. Mit $g_n \leqslant f_n$ für alle $n \in \mathbb{N}$ und dem Satz über monotone Konvergenz folgt:

$$\mu\left(\liminf_{n\to\infty} f_n\right) = \lim_{n\to\infty} \mu(g_n) \leqslant \liminf_{n\to\infty} \mu(f_n).$$

Satz 4.7 (Dominierte Konvergenz, Satz von Lebesgue)**.** *Seien $(\Omega, \mathscr{A}, \mu)$ ein Maßraum und $f, f_1, f_2, \ldots\colon \Omega \to \overline{\mathbb{R}}$ messbar numerisch mit $f = \lim_{n\to\infty} f_n$ $[\mu]$. Gibt es eine integrierbare Funktion $g\colon \Omega \to \overline{\mathbb{R}}$ mit $|f_n| \leqslant g$ $[\mu]$ für alle $n \in \mathbb{N}$, so ist f integrierbar und es gilt:*

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

Sofern μ-fast-überall eine Majorante g existiert, können wir also Limesbildung (punktweise Konvergenz) mit Integration vertauschen.

Beweis. Definiere

$$N := \left\{f \neq \lim_{n\to\infty} f_n\right\} \cup \bigcup_{n\in\mathbb{N}} \{|f_n| > g\} \cup \{g = \infty\} \in \mathscr{A}.$$

Dies ist nach Voraussetzung eine μ-Nullmenge. Für alle $n \in \mathbb{N}$ gilt $|f_n \cdot \mathbb{1}_{N^c}| \leqslant g$ sowie $|f \cdot \mathbb{1}_{N^c}| \leqslant g$. Folglich sind $\tilde{f} := f \cdot \mathbb{1}_{N^c}$ und $\tilde{f}_n := f_n \cdot \mathbb{1}_{N^c}$ für alle $n \in \mathbb{N}$ integrierbar. Es gilt:

$$\tilde{f} = \lim_{n\to\infty} \tilde{f}_n = \liminf_{n\to\infty} \tilde{f}_n, \qquad \left|\tilde{f}_n\right| \leqslant \tilde{g} := g \cdot \mathbb{1}_{N^c} < \infty.$$

Somit sind $(\tilde{g} + \tilde{f}_n)_{n\in\mathbb{N}}$ und $(\tilde{g} - \tilde{f}_n)_{n\in\mathbb{N}}$ nichtnegative Folgen in $\mathscr{E}^*$ und es folgt mit dem Lemma von Fatou (Satz 4.6) einerseits:

$$\mu(\tilde{g} + \tilde{f}) = \mu\big(\liminf_{n\to\infty}(\tilde{g} + \tilde{f}_n)\big) \leqslant \liminf_{n\to\infty} \mu(\tilde{g} + \tilde{f}_n) = \mu(\tilde{g}) + \liminf_{n\to\infty} \mu(\tilde{f}_n)$$

und wir erhalten:

$$\mu(\tilde{f}) \leqslant \liminf_{n\to\infty} \mu(\tilde{f}_n).$$

Andererseits gilt:

$$\mu(\tilde{g} - \tilde{f}) = \mu\big(\liminf_{n\to\infty}(\tilde{g} - \tilde{f}_n)\big) \leqslant \liminf_{n\to\infty} \mu(\tilde{g} - \tilde{f}_n) = \mu(\tilde{g}) - \limsup_{n\to\infty} \mu(\tilde{f}_n)$$

und wir erhalten:

$$\mu(\tilde{f}) \geqslant \limsup_{n\to\infty} \mu(\tilde{f}_n).$$

Diese beiden Abschätzungen ergeben:

$$\lim_{n\to\infty} \mu(f_n) = \lim_{n\to\infty} \mu(\tilde{f}_n) = \mu(\tilde{f}) = \mu(f).$$

Korollar 4.8.

1. *Seien* $(\Omega, \mathscr{A}, P)$ *ein Wahrscheinlichkeitsraum und* $X, X_1, X_2, \ldots : \Omega \to \mathbb{R}$ *Zufallsvariablen mit* $|X_n| \leqslant c \in \mathbb{R}$ *für alle* $n \in N$ *und* $\lim_{n\to\infty} X_n = X\ [P]$*. Dann ist* $\mathrm{E}X = \lim_{n\to\infty} \mathrm{E}X_n$*.*

2. *Seien* $(\Omega, \mathscr{A}, \mu)$ *ein Maßraum,* $U \subseteq \mathbb{R}$ *offen und* $f\colon U \times \Omega \to \mathbb{R}$ *eine Abbildung mit folgenden Eigenschaften:*

   (a) *Die Abbildungen* $\omega \mapsto f(t, \omega)$ *sind für alle* $t \in U$ μ*-integrierbar.*

   (b) *Die Abbildungen* $t \mapsto f(t, \omega)$ *sind für alle* $\omega \in \Omega$ *differenzierbar.*

   (c) *Es gibt eine* μ*-integrierbare Abbildung* $h\colon \Omega \to \overline{\mathbb{R}}$ *mit* $\left|\frac{\partial}{\partial t} f(t,\omega)\right| \leqslant h(\omega)$ *für alle* $(t, \omega) \in U \times \Omega$*.*

   Dann ist die Abbildung

$$\begin{aligned}\phi\colon U &\to \mathbb{R}\\ t &\mapsto \int f(t,\omega)\, \mu(\mathrm{d}\omega)\end{aligned}$$

   differenzierbar und es gilt:

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi(t) = \int \frac{\partial}{\partial t} f(t,\omega)\, \mu(\mathrm{d}\omega).$$

3. *Seien* $(X_{n,k})_{n,k\in\mathbb{N}}$ *eine reelle Doppelfolge und* $(X_k)_{k\in\mathbb{N}}$ *eine reelle Folge, sodass für alle* $k \in \mathbb{N}$ *gilt:* $\lim_{n\to\infty} X_{n,k} = X_k$. *Sei außerdem* $(Y_k)_{k\in\mathbb{N}}$ *eine nichtnegative Folge mit* $\sum_{k\in\mathbb{N}} Y_k < \infty$ *und* $\sup_{n\in\mathbb{N}} |X_{n,k}| \leqslant Y_k$ *für alle* $k \in \mathbb{N}$. *Dann gilt:*

$$\lim_{n\to\infty} \sum_{k\in\mathbb{N}} X_{n,k} = \sum_{k\in\mathbb{N}} X_k.$$

[Übung]

Lemma 4.9 (Lemma von Pratt). *Sei* $(\Omega, \mathscr{A}, \mu)$ *ein Maßraum. Seien* $(f_n)_{n\in\mathbb{N}}$, $(g_n)_{n\in\mathbb{N}}$ *und* $(h_n)_{n\in\mathbb{N}}$ μ*-fast-überall konvergente Folgen von integrierbaren Funktionen auf* Ω. *Seien* f *und* h *integrierbare Funktionen mit* $\lim_{n\to\infty} f_n = f$ *und* $\lim_{n\to\infty} h_n = h$ μ*-fast-überall. Weiterhin sei* g *messbar mit* $\lim_{n\to\infty} g_n = g$ μ*-fast-überall. Es gelte:*

$$f_n \leqslant g_n \leqslant h_n\ [\mu] \quad \text{für alle } n \in \mathbb{N} \tag{4.2}$$

sowie $\lim_{n\to\infty} \int f_n \,\mathrm{d}\mu = \int f \,\mathrm{d}\mu$ *und* $\lim_{n\to\infty} \int h_n \,\mathrm{d}\mu = \int h \,\mathrm{d}\mu$. *Dann ist* g *integrierbar und es gilt:*

$$\lim_{n\to\infty} \int g_n \,\mathrm{d}\mu = \int g \,\mathrm{d}\mu.$$

[Übung]

Satz 4.10 (Egorov). *Seien* $(\Omega, \mathscr{A}, P)$ *ein Wahrscheinlichkeitsraum,* $X, X_1, X_2, \ldots$ *Zufallsvariablen mit* $\lim_{n\to\infty} X_n = X\ [P]$. *Dann gibt es für alle* $\varepsilon > 0$ *ein* $A \in \mathscr{A}$ *mit* $P(A^c) \geqslant 1 - \varepsilon$ *und*

$$\lim_{n\to\infty} \sup_{\omega\in A} |X_n(\omega) - X(\omega)| = 0.$$

[Übung]

Übungsaufgaben:

Übung 1: Beweisen Sie Korollar 4.8.

Übung 2: Beweisen Sie den Satz von Egorov (Satz 4.10).

Lernziel-Kontrolle:

Frage 1: Was besagt der Begriff der μ-fast-überall bestehenden Gleichheit und was ist seine Bedeutung für das μ-Integral?

Frage 2: Was besagt der Satz von der dominierten Konvergenz und wie beweist man ihn?

Frage 3: Kennen Sie ein Beispiel für eine Funktionenfolge $(f_n)_{n\in\mathbb{N}}$, für welche gilt:

$$\lim_{n\to\infty} \int f_n \,\mathrm{d}\lambda \neq \int \lim_{n\to\infty} f_n \,\mathrm{d}\lambda?$$

Frage 4: Was besagt die Markov-Ungleichung?

Kapitel 5

Lebesgue vs. Riemann

In diesem Kapitel werden wir uns dem Vergleich zwischen Lebesgue- und Riemann-Integral widmen. Um das Riemann-Integral definieren zu können, benötigen wir die folgenden Begriffe.

Definition 5.1. Eine *Zerlegung* (engl.: decomposition) eines Intervalls $[a,b]$ mit $a \leqslant b \in \mathbb{R}$ ist eine Menge $Z := \{x_i\}_{i=0}^k$ mit $a = x_0 < x_1 < \cdots < x_k = b$. Wir schreiben $|Z| := \max_{i=1}^k |x_i - x_{i-1}|$. Eine Zerlegung Z' mit $Z \subseteq Z'$ bezeichnen wir als *Verfeinerung* (engl.: refinement) von Z.

Sei $f\colon [a,b] \to \mathbb{R}$ eine Funktion. Die *Untersumme von f bezüglich Z* (engl.: lower sum) ist definiert als

$$U(f,Z) := \sum_{i=1}^k \inf_{x \in [x_{i-1},x_i)} f(x)(x_i - x_{i-1}).$$

Die *Obersumme von f bezüglich Z* (engl.: upper sum) ist definiert als

$$O(f,Z) := \sum_{i=1}^k \sup_{x \in [x_{i-1},x_i)} f(x)(x_i - x_{i-1}).$$

Die *Untersumme von f* ist definiert als $U(f) := \sup_Z U(f,Z)$, die *Obersumme von f* als $O(f) := \inf_Z O(f,Z)$.

Definition 5.2. Sei $[a,b]$ mit $a \leqslant b \in \mathbb{R}$ ein Intervall. Wir bezeichnen eine Funktion $f\colon [a,b] \to \mathbb{R}$ als *Riemann-integrierbar* (engl.: Riemann integrable), falls $O(f) = U(f)$. Das *eigentliche Riemann-Integral* (engl.: proper Riemann integral) von f ist definiert als

$$\int_a^b f(x)\,\mathrm{d}x := O(f) = U(f).$$

Der nächste Satz ermöglicht eine einfachere Berechnung des eigentlichen Riemann-Integrals.

Satz 5.3. *Seien $[a,b]$ mit $a \leqslant b \in \mathbb{R}$ ein Intervall und $f\colon [a,b] \to \mathbb{R}$ Riemann-integrierbar. Dann gilt für jede Folge von Verfeinerungen $Z_0 \subseteq Z_1 \subseteq Z_2 \subseteq \cdots$ mit $\lim_{k\to\infty} |Z_k| = 0$:*

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{k\to\infty} U(f, Z_k) = \lim_{k\to\infty} O(f, Z_k).$$

Beweis. Nach Definition gilt bereits $\lim_{k\to\infty} U(f, Z_k) \leqslant U(f)$. Aus $\lim_{k\to\infty} |Z_k| = 0$ folgt:

$$\overline{\bigcup_{k\in\mathbb{N}} Z_k} = [a,b]. \tag{5.1}$$

Sei nun $Z := \{x_i\}_{i=0}^n$ eine Zerlegung von $[a,b]$. Definiere $\gamma_i := \inf_{x\in[x_{i-1},x_i]} f(x)$. Aus (5.1) folgt, dass es für alle $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt, sodass es zu jedem $x_i \in Z$ ein $y_{j(i)} \in Z_N = \{y_j\}_{j=0}^k$ gibt mit $y_{j(i)-1} \leqslant x_i \leqslant y_{j(i)}$ und:

$$\left|y_{j(i)-1} - y_{j(i)}\right| \leqslant \frac{\varepsilon}{n\gamma_i}.$$

Es gilt:

$$\begin{aligned}
&U(f,Z) - U(f,Z_N)\\
&= \sum_{i=1}^n \inf_{x\in[x_{i-1},x_i)} f(x)(x_i - x_{i-1}) - \sum_{j=1}^k \inf_{x\in[y_{j-1},y_j)} f(x)(y_j - y_{j-1})\\
&= \sum_{i=1}^n \left(\inf_{x\in[x_{i-1},x_i)} f(x)(x_i - x_{i-1}) - \sum_{l=j(i-1)+1}^{j(i)} \inf_{x\in[y_{l-1},y_l)} f(x)(y_l - y_{l-1}) \right)\\
&\leqslant \sum_{i=1}^n \left(\inf_{x\in[x_{i-1},x_i)} f(x)(x_i - x_{i-1}) - \sum_{l=j(i-1)+1}^{j(i)-1} \inf_{x\in[x_{i-1},x_i)} f(x)(y_l - y_{l-1}) \right)\\
&\leqslant \sum_{i=1}^n \inf_{x\in[x_{i-1},x_i)} f(x)(y_{j(i)} - y_{j(i)-1}) \leqslant \sum_{i=1}^n \inf_{x\in[x_{i-1},x_i)} f(x)\frac{\varepsilon}{n\gamma_i} \leqslant \varepsilon.
\end{aligned}$$

Somit folgt $\lim_{k\to\infty} U(f,Z_k) \geqslant U(f,Z) - \varepsilon$ für alle $\varepsilon > 0$ und alle Zerlegungen Z. Also folgt $\lim_{k\to\infty} U(f,Z_k) \geqslant \sup_Z U(f,Z)$ und damit

$$\lim_{k\to\infty} U(f,Z_k) = \sup_Z U(f,Z) = U(f) = \int_a^b f(x)\,\mathrm{d}x.$$

Definition 5.4. Seien $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ reelle Folgen mit $a_n \leqslant b_n$ für alle $n \in \mathbb{N}$ und $\lim_{n\to\infty} a_n = -\infty$ sowie $\lim_{n\to\infty} b_n = +\infty$. $f\colon \mathbb{R} \to \mathbb{R}$ heißt *Riemann-integrierbar*, falls für alle reellen Folgen $(c_n)_{n\in\mathbb{N}}$ und $(d_n)_{n\in\mathbb{N}}$ mit $c_n \leqslant d_n$ für alle

$n \in \mathbb{N}$ und $\lim_{n\to\infty} c_n = -\infty$ sowie $\lim_{n\to\infty} d_n = +\infty$ gilt:

$$\lim_{n\to\infty} \int_{a_n}^{b_n} f(x)\,\mathrm{d}x = \lim_{n\to\infty} \int_{c_n}^{d_n} f(x)\,\mathrm{d}x \in \mathbb{R}.$$

Dies gewährleistet die Wohldefiniertheit des Grenzwertes. Wir definieren das *uneigentliche Riemann-Integral* (engl.: improper Riemann integral) von f als

$$\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x := \lim_{n\to\infty} \int_{a_n}^{b_n} f(x)\,\mathrm{d}x.$$

★**Lemma 5.5.** *In der obigen Definition des uneigentlichen Riemann-Integrals reicht es, monotone reelle Folgen zu betrachten. $f\colon \mathbb{R} \to \mathbb{R}$ ist also genau dann Riemann-integrierbar, wenn für jede monoton wachsende reelle Folge $(b_n)_{n\in\mathbb{N}}$ und jede monoton fallende reelle Folge $(a_n)_{n\in\mathbb{N}}$ mit $\lim_{n\to\infty} a_n = -\infty$, $\lim_{n\to\infty} b_n = \infty$ und $b_n - a_n \geqslant 0$ für alle $n \in \mathbb{N}$ die Folge $(R_n)_{n\in\mathbb{N}}$ mit $R_n := \int_{a_n}^{b_n} f(x)\,\mathrm{d}x$ konvergiert und der Grenzwert unabhängig von $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ ist.*

Beweis. Ist f Riemann-integrierbar, so folgt die Konvergenz und Wohldefiniertheit der Folge $(R_n)_{n\in\mathbb{N}}$ direkt aus der Definition.

Sei nun f nicht Riemann-integrierbar. Existieren Folgen $(a_n)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$, $(c_n)_{n\in\mathbb{N}}$ und $(d_n)_{n\in\mathbb{N}}$ mit $\lim_{n\to\infty} a_n = -\infty = \lim_{n\to\infty} c_n$, $\lim_{n\to\infty} b_n = \infty = \lim_{n\to\infty} d_n$ sowie $a_n \leqslant b_n$ und $c_n \leqslant d_n$ für alle $n \in \mathbb{N}$, sodass gilt:

$$\lim_{n\to\infty} \int_{a_n}^{b_n} f(x)\,\mathrm{d}x = a \neq c = \lim_{n\to\infty} \int_{c_n}^{d_n} f(x)\,\mathrm{d}x,$$

so können wir mit Hilfe eines Reißverschlussverfahrens Folgen $(x_n)_{n\in\mathbb{N}}$ und $(y_n)_{n\in\mathbb{N}}$ mit $x_{2n} := a_n$, $x_{2n+1} := c_n$, $y_{2n} := b_n$ und $y_{2n+1} := d_n$ konstruieren. Es sind $\lim_{n\to\infty} x_n = -\infty$, $\lim_{n\to\infty} y_n = \infty$ und $x_n \leqslant y_n$, aber die Folge $\left(\int_{x_n}^{y_n} f(x)\,\mathrm{d}x\right)_{n\in\mathbb{N}}$ konvergiert nicht. Somit finden wir immer Folgen $(e_n)_{n\in\mathbb{N}}$ und $(f_n)_{n\in\mathbb{N}}$ mit $e_n \leqslant f_n$ für alle $n \in \mathbb{N}$, $\lim_{n\to\infty} e_n = -\infty$ und $\lim_{n\to\infty} f_n = \infty$, sodass $(R_n)_{n\in\mathbb{N}} := \left(\int_{e_n}^{f_n} f(x)\,\mathrm{d}x\right)_{n\in\mathbb{N}}$ nicht konvergiert. Es treten folgende Fälle auf:

1. $\liminf_{n\to\infty} R_n = \limsup_{n\to\infty} R_n \in \{-\infty, +\infty\}$,

2. $\liminf_{n\to\infty} R_n \neq \limsup_{n\to\infty} R_n$.

In beiden Fällen können wir Teilfolgen $(R'_n)_{n\in\mathbb{N}}$ und $(R''_n)_{n\in\mathbb{N}}$ von $(R_n)_{n\in\mathbb{N}}$ wählen, bei denen man über immer größere Intervalle integriert, mit $\lim_{n\to\infty} R'_n = \liminf_{n\to\infty} R_n$ und $\lim_{n\to\infty} R''_n = \limsup_{n\to\infty} R_n$. Zu diesen Teilfolgen gehören wiederum monotone Folgen, welche durch die Integrationsgrenzen bestimmt sind. Wir finden also antitone reelle Folgen $(e'_n)_{n\in\mathbb{N}}$ und $(e''_n)_{n\in\mathbb{N}}$ mit $\lim_{n\to\infty} e'_n = \lim_{n\to\infty} e''_n = -\infty$ und isotone reelle Folgen $(f'_n)_{n\in\mathbb{N}}$ und $(f''_n)_{n\in\mathbb{N}}$ mit $\lim_{n\to\infty} f'_n = \lim_{n\to\infty} f''_n = \infty$ und $e'_n \leqslant f'_n$

sowie $e''_n \leqslant f''_n$ und

$$\lim_{n\to\infty} \int_{e'_n}^{f'_n} f(x)\,\mathrm{d}x = \lim_{n\to\infty} R'_n = \liminf_{n\to\infty} R_n,$$
$$\lim_{n\to\infty} \int_{e''_n}^{f''_n} f(x)\,\mathrm{d}x = \lim_{n\to\infty} R''_n = \limsup_{n\to\infty} R_n.$$

Da die Folgen $(R'_n)_{n\in\mathbb{N}}$ und $(R''_n)_{n\in\mathbb{N}}$ nicht gegen den gleichen Grenzwert konvergieren, folgt die Behauptung.

★Lemma 5.6. *Eine Funktion $f\colon \mathbb{R} \to [0,\infty)$ ist genau dann Riemann-integrierbar, wenn die Folge $(R_n)_{n\in\mathbb{N}}$ mit $R_n := \int_{-n}^{n} f(x)\,\mathrm{d}x$ konvergiert. Dies ist genau dann der Fall, wenn sie beschränkt ist.*

Beweis. Aus $f \geqslant 0$ folgt, dass $(R_n)_{n\in\mathbb{N}}$ eine isotone Folge ist. Somit konvergiert sie genau dann, wenn sie außerdem beschränkt ist.

Ist f Riemann-integrierbar, so folgt mit $\lim_{n\to\infty} n = \infty$, $\lim_{n\to\infty} -n = -\infty$ und $n \geqslant -n$ für alle $n \in \mathbb{N}$ die Konvergenz von $(R_n)_{n\in\mathbb{N}}$ direkt aus der Definition des uneigentlichen Riemann-Integrals.

Sei nun $(R_n)_{n\in\mathbb{N}}$ konvergent. Seien $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ monotone reelle Folgen mit $\lim_{n\to\infty} b_n = \infty$, $\lim_{n\to\infty} a_n = -\infty$ und $b_n \geqslant a_n$. Ohne Einschränkung können wir $a_n \leqslant 0$ und $b_n \geqslant 0$ für alle $n \in \mathbb{N}$ vorraussetzen. Somit gibt es für alle $n \in \mathbb{N}$ ein $k_n \in \mathbb{N}$ mit $b_n \in [k_n, k_n + 1)$. Da $(b_n)_{n\in\mathbb{N}}$ monoton steigend ist, ist auch $(k_n)_{n\in\mathbb{N}}$ monoton steigend und es gilt:

$$\int_0^{k_n} f(x)\,\mathrm{d}x \leqslant \int_0^{b_n} f(x)\,\mathrm{d}x < \int_0^{k_n+1} f(x)\,\mathrm{d}x.$$

Durch Grenzwertbildung erhält man:

$$\begin{aligned}\lim_{n\to\infty} \int_0^{k_n} f(x)\,\mathrm{d}x = \lim_{n\to\infty} \int_0^{n} f(x)\,\mathrm{d}x \leqslant \lim_{n\to\infty} \int_0^{b_n} f(x)\,\mathrm{d}x &\leqslant \lim_{n\to\infty} \int_0^{n} f(x)\,\mathrm{d}x \\ &= \lim_{n\to\infty} \int_0^{k_n+1} f(x)\,\mathrm{d}x.\end{aligned}$$

Analog ergibt sich:

$$\lim_{n\to\infty} \int_{-n}^{0} f(x)\,\mathrm{d}x \leqslant \lim_{n\to\infty} \int_{a_n}^{0} f(x)\,\mathrm{d}x \leqslant \lim_{n\to\infty} \int_{-n}^{0} f(x)\,\mathrm{d}x.$$

Mit Lemma 5.5 folgt die Behauptung.

Definition 5.7. Sei $(\Omega, \mathscr{A}, \mu)$ ein Maßraum. Sei $A \in \mathscr{A}$ mit $A \neq \varnothing$ und sei $\iota\colon A \to \Omega$ die Einbettung (vgl. 2.15). Außerdem sei $\mu_A := \mu|_{\sigma(\iota)}$ die Einschränkung von μ auf die durch A gegebene Spur-σ-Algebra. Weiterhin sei $f\colon A \to \overline{\mathbb{R}}$ messbar numerisch sowie μ_A-integrierbar.

$$\int_A f \,\mathrm{d}\mu := \int f \,\mu_A$$

heißt *μ-Integral* von f über A.

Im Spezialfall des Lebesgue-Maßes $\mu = \lambda$ erhalten wir auf diese Weise für eine Menge $B \in \mathscr{B}$ das *Lebesgue-Integral* von f über B:

$$\int_B f(x) \,\mathrm{d}x := \int_{\mathbb{R}} f(x) \,\lambda_B(\mathrm{d}x).$$

Lemma 5.8. *Unter den Voraussetzungen von Definition 5.7 gilt:*

$$\int_B f \,\mathrm{d}\lambda_B = \int_{\mathbb{R}} f \cdot \mathbb{1}_B \,\mathrm{d}\lambda.$$

Beweis. Wir beweisen die Aussage durch algebraische Induktion. Sei zunächst $f \in \mathscr{E}$ mit der Normaldarstellung $f = \sum_{i=1}^n \alpha_i \cdot \mathbb{1}_{A_i}$. Dann ist $B = \sum_{i=1}^n A_i$. Folglich ist $0 \cdot \mathbb{1}_{B^c} + \sum_{i=1}^n \alpha_i \cdot \mathbb{1}_{A_i}$ eine Normaldarstellung von $\mathbb{1}_B \cdot f$ und es gilt:

$$\int_B f \,\mathrm{d}\lambda_B = \sum_{i=1}^n \alpha_i \lambda_B(A_i) = 0 \cdot \lambda(B^c) + \sum_{i=1}^n \alpha_i \lambda(A_i) = \int_{\mathbb{R}} f \cdot \mathbb{1}_B \,\mathrm{d}\lambda.$$

Somit gilt die Aussage für Elementarfunktionen. Seien nun $f \in \mathscr{E}^*$ und $(u_n)_{n\in\mathbb{N}}$ eine monoton gegen f steigende Folge in $\mathscr{E}$. Dann ist $(u_n \cdot \mathbb{1}_B)_{n\in\mathbb{N}}$ eine monoton gegen $f \cdot \mathbb{1}_B$ steigende Folge in $\mathscr{E}$ und es gilt:

$$\int_B f \,\mathrm{d}\lambda_B = \lim_{n\to\infty} \int_B u_n \,\mathrm{d}\lambda_B = \lim_{n\to\infty} \int u_n \cdot \mathbb{1}_B \,\mathrm{d}\lambda = \int_{\mathbb{R}} f \cdot \mathbb{1}_B \,\mathrm{d}\lambda.$$

Sei nun f λ_B-integrierbar. Es gilt:

$$\begin{aligned}\int_B f \,\mathrm{d}\lambda_B = \int_B f^+ \,\mathrm{d}\lambda_B - \int_B f^- \,\mathrm{d}\lambda_B &= \int f^+ \cdot \mathbb{1}_B \,\mathrm{d}\lambda - \int f^- \cdot \mathbb{1}_B \,\mathrm{d}\lambda \\ &= \int (f \cdot \mathbb{1}_B)^+ \,\mathrm{d}\lambda - \int (f \cdot \mathbb{1}_B)^- \,\mathrm{d}\lambda = \int_{\mathbb{R}} f \cdot \mathbb{1}_B \,\mathrm{d}\lambda.\end{aligned}$$

Satz 5.9. *Sei $[a,b]$ mit $a \leqslant b \in \mathbb{R}$ ein Intervall. Jede $([a,b] \cap \mathscr{B}, \mathscr{B})$-messbare Riemann-integrierbare Funktion $f\colon [a,b] \to \mathbb{R}$ ist auch Lebesgue-integrierbar und die Integrale stimmen überein.*

Beweis. Sei $Z := \{x_i\}_{i=0}^n$ eine Zerlegung von $[a,b]$. Für alle $j \in \{1, \ldots, n\}$ definiere

$$A_j := [x_{j-1}, x_j), \quad \gamma_j := \inf_{x \in A_j} f(x), \quad \delta_j := \sup_{x \in A_j} f(x)$$

sowie die Funktionen

$$u_Z := \sum_{j=1}^{n} \gamma_j \mathbb{1}_{A_j}, \quad o_Z := \sum_{j=1}^{n} \delta_j \mathbb{1}_{A_j}.$$

Diese Funktionen sind messbar, da A_j für alle $j \in \{1, \dots, n\}$ abgeschlossen ist, und sogar Lebesgue-integrierbar, da f beschränkt ist, und somit gilt $|\gamma_j| < \infty$ und $|\delta_j| < \infty$ für alle $j \in \{1, \dots, n\}$. Es gilt:

$$U(f, Z) = \int u_Z \, \mathrm{d}\lambda_{[a,b]}, \quad O(f, Z) = \int o_Z \, \mathrm{d}\lambda_{[a,b]}.$$

Sei $Z_0 \subseteq Z_1 \subseteq Z_2 \subseteq \cdots$ eine Folge von Verfeinerungen mit $\lim_{n\to\infty} |Z_n| = 0$. Dann ist $(u_{Z_n})_{n\in\mathbb{N}}$ monoton wachsend und $(o_{Z_n})_{n\in\mathbb{N}}$ monoton fallend. Somit ist die Folge $(o_{Z_n} - u_{Z_n})_{n\in\mathbb{N}} \geqslant 0$ monoton fallend. Also existiert der punktweise Limes $q := \lim_{n\to\infty}(o_{Z_n} - u_{Z_n})$. Nach dem Lemma von Fatou (Satz 4.6) gilt:

$$0 \leqslant \int q \, \mathrm{d}\lambda_{[a,b]} \leqslant \liminf_{n\to\infty} \int o_{Z_n} - u_{Z_n} \, \mathrm{d}\lambda_{[a,b]} = \liminf_{n\to\infty} \big(O(f, Z_n) - U(f, Z - n)\big) = 0.$$

Folglich ist $q = 0 \ [\lambda_{[a,b]}]$. Da $u_{Z_n} \leqslant f \leqslant o_{Z_n}$ ist, gilt:

$$\lim_{n\to\infty} u_{Z_n} = \lim_{n\to\infty} o_{Z_n} = f \quad [\lambda_{[a,b]}].$$

Aus der Definition des Riemann-Integrals, der Beschränktheit von f und dem Satz über dominierte Konvergenz folgt:

$$\int_a^b f(x) \, \mathrm{d}x = \lim_{n\to\infty} U(f, Z_n) = \lim_{n\to\infty} \int u_{Z_n} \, \mathrm{d}\lambda_{[a,b]} = \int f \, \mathrm{d}\lambda_{[a,b]}.$$

Dass die Umkehrung im Allgemeinen nicht zutrifft, sehen wir an folgendem Beispiel.

Beispiel 5.10 (Dirichletsche Sprungfunktion). Betrachte die Funktion

$$\begin{aligned} f\colon [0,1] &\to [0,1] \\ x &\mapsto \begin{cases} 1 & x \in \mathbb{Q} \cap [0,1] \\ 0 & \text{sonst} \end{cases}. \end{aligned}$$

Diese ist nicht Riemann-integrierbar, da $O(f) = 1$, aber $U(f) = 0$. Sie ist jedoch Lebesgue-integrierbar, da $f = 0 \ [\lambda_{[0,1]}]$, und es gilt: $\int f \, \mathrm{d}\lambda_{[0,1]} = 0$.

Satz 5.11. *Sei $f\colon \mathbb{R} \to [0, \infty)$ Borel-messbar und Riemann-integrierbar auf jedem kompakten Intervall. f ist genau dann Lebesgue-integrierbar, wenn das uneigentliche Riemann-Integral von f existiert. In diesem Fall stimmen die Integrale überein.*

Beweis. Nach Lemma 5.6 ist f genau dann Riemann-integrierbar, wenn die Folge $\left(\int_{-n}^{n} f(x)\,\mathrm{d}x\right)_{n\in\mathbb{N}}$ konvergiert. Es gilt: $f \cdot \mathbb{1}_{[-n,n]} \nearrow f$. Somit ist wegen Satz 5.9 $f \cdot \mathbb{1}_{[-n,n]}$ Lebesgue-integrierbar und mit dem Satz über die monotone Konvergenz (Satz 3.14) folgt:

$$\lim_{n\to\infty} \int_{-n}^{n} f(x)\,\mathrm{d}x = \lim_{n\to\infty} \int f \cdot \mathbb{1}_{[-n,n]}\,\mathrm{d}\lambda = \int f\,\mathrm{d}\lambda < \infty.$$

Beispiel 5.12. Betrachte die *Sinc-Funktion*

$$\begin{aligned} f\colon \mathbb{R} &\to \mathbb{R} \\ x &\mapsto \begin{cases} \frac{\sin(x)}{x} & x \neq 0 \\ 1 & x = 0 \end{cases}. \end{aligned}$$

Diese ist zwar Riemann-integrierbar mit $\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = \pi$, aber nicht Lebesgue-integrierbar. [Übung]

Lemma 5.13. *Seien $[a,b]$ mit $a \leqslant b \in \mathbb{R}$ ein Intervall und $f\colon [a,b] \to \mathbb{R}$ eine Funktion. Dann ist die Menge*

$$C_f := \{x \in [a,b] \mid f \textit{ ist stetig an } x\}$$

eine Borel-Menge. [Übung]

Satz 5.14. *Sei $[a,b]$ mit $a \leqslant b \in \mathbb{R}$ ein Intervall. Eine Funktion $f\colon [a,b] \to \mathbb{R}$ ist genau dann Riemann-integrierbar, wenn sie beschränkt und außerdem $\lambda_{[a,b]}$-fast-überall stetig ist.*

Beweis. Für eine Zerlegung $Z := \{x_i\}_{i=0}^{n}$ von $[a,b]$ definieren wir die Funktionen u_Z und o_Z wie im Beweis von Satz 5.9. Sei $Z_0 \subseteq Z_1 \subseteq Z_2 \subseteq \cdots$ eine Folge von Verfeinerungen mit $\lim_{n\to\infty} |Z_n| = 0$. Dann ist $(u_{Z_n})_{n\in\mathbb{N}}$ monoton wachsend und $(o_{Z_n})_{n\in\mathbb{N}}$ monoton fallend. Da $(u_{Z_n})_{n\in\mathbb{N}}$ und $(o_{Z_n})_{n\in\mathbb{N}}$ sich gegenseitig beschränken, konvergieren sie punktweise gegen Grenzfunktionen u beziehungsweise o.

Wir zeigen zunächst, dass für jedes $x \in [a,b]$ mit $x \notin Z_n$ für alle $n \in \mathbb{N}$ gilt, dass f genau dann stetig an x ist, wenn $u(x) = f(x) = o(x)$ ist.

Sei f stetig an x. Dann gibt es für alle $\varepsilon = 1/m > 0$ ein $\delta > 0$, sodass aus $|y - x| < \delta$ folgt: $|f(y) - f(x)| < 1/m$. Wähle $N_m \in \mathbb{N}$ mit $|Z_{N_m}| < \delta$. Seien $z_k, z_{k+1} \in Z_{N_m}$, sodass $x \in (z_k, z_{k+1}) \subseteq (x - \delta, x + \delta)$. Somit gilt:

$$\begin{aligned} \left|o_{Z_{N_m}}(x) - f(x)\right| &= \left|\sup_{y\in[z_k,z_{k+1})} f(y) - f(x)\right| \leqslant \frac{1}{m}, \\ \left|u_{Z_{N_m}}(x) - f(x)\right| &= \left|\inf_{y\in[z_k,z_{k+1})} f(y) - f(x)\right| \leqslant \frac{1}{m}. \end{aligned}$$

Folglich ist $\left|o_{Z_{N_m}}(x) - u_{Z_{N_m}}(x)\right| \leqslant \frac{2}{m}$. Für $m \to \infty$ erhalten wir: $u(x) = f(x) = o(x)$.

Sei nun $u(x) = o(x)$. Dann gibt es für alle $\varepsilon > 0$ ein $N_\varepsilon \in \mathbb{N}$ mit

$$o_{Z_{N_\varepsilon}}(x) - f(x) < \varepsilon, \quad f(x) - u_{Z_{N_\varepsilon}}(x) < \varepsilon.$$

Seien $z_k, z_{k+1} \in Z_{N_\varepsilon}$ mit $x \in (z_k, z_{k+1})$. Wähle $\delta > 0$, sodass $(x-\delta, x+\delta) \subseteq (z_k, z_{k+1})$. Für alle $y \in (x - \delta, x + \delta)$ gilt somit:

$$f(y) < \sup_{w \in (x-\delta, x+\delta)} f(w) < \sup_{w \in [z_k, z_{k+1})} f(w) = o_{Z_{N_\varepsilon}}(x) < \varepsilon + f(x),$$
$$f(y) > \inf_{w \in (x-\delta, x+\delta)} f(w) > \inf_{w \in [z_k, z_{k+1})} f(w) = u_{Z_{N_\varepsilon}}(x) > f(x) - \varepsilon.$$

Folglich ist $|f(y) - f(x)| < \varepsilon$.

Sei f Riemann-integrierbar. Nach Satz 5.3 folgt:

$$\int o \, \lambda_{[a,b]} = \int_a^b f(x) \, \mathrm{d}x = \int u \, \lambda[a, b]$$

und somit:

$$\int o - u \, \lambda[a, b] = 0.$$

Wegen $o - u \geqslant 0$ folgt $\lambda_{[a,b]}$-fast-überall: $o = u$. Nach dem oben Gezeigten ist die Funktion f $\lambda_{[a,b]}$-fast-überall stetig. Angenommen, f wäre unbeschränkt. Dann gäbe es ein $x \in [a, b]$, sodass für jedes offene Intervall $A \subseteq [a, b]$ mit $x \in A$ gilt: $\sup_{y \in A} |f(y)| = \infty$. Wir betrachten nun eine Folge von Zerlegungen $Z_1 \subseteq Z_2 \subseteq \cdots$ mit $\lim_{n\to\infty} |Z_n| = 0$ und $x \notin Z_n$ für alle $n \in \mathbb{N}$. Somit gilt: $U(f, Z_n) = -\infty$ für alle $n \in \mathbb{N}$ oder $O(f, Z_n) = \infty$ für alle $n \in \mathbb{N}$. Dies ist ein Widerspruch dazu, dass f Riemann-integrierbar ist.

Sei nun f beschränkt und $\lambda_{[a,b]}$-fast-überall stetig. Dann ist $u = f = o$ $\lambda_{[a,b]}$-fast-überall. Da die σ-Algebra der Lebesgue-Mengen (siehe Bemerkung 1.57) vollständig ist, ist f somit messbar und als beschränkte Funktion auf $[a, b]$ Lebesgue-integrierbar. Es folgt:

$$\int o \, \lambda_{[a,b]} = \int f \, \lambda_{[a,b]} = \int u \, \lambda_{[a,b]}.$$

Da dies unabhängig von der gewählten Zerlegung gilt, ist f Riemann-integrierbar.

Übungsaufgaben:

Übung 1: Beweisen Sie Beispiel 5.12.

Übung 2: Beweisen Sie Lemma 5.13.

Lernziel-Kontrolle:

Frage 1: Warum verwendet man nicht immer das Lebesgue-Integral anstelle des Riemann-Integrals?

Frage 2: Was ist anschaulich der Unterschied zwischen den beiden Integralbegriffen und wo liegt die besondere Stärke in der Definition des Lebesgue-Integrals?

Frage 3: Kennen Sie ein Beispiel für eine nicht Riemann-integrierbare Funktion? (Etwas schwieriger: Finden Sie eine Funktion, welche weder Riemann-, noch Lebesgue-integrierbar ist.)

Frage 4: Sei $f\colon \mathbb{R} \to \mathbb{R}$ eine stetige Funktion, sodass für alle reellen Folgen $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ mit $\lim_{n\to\infty} a_n = -\infty$ sowie $\lim_{n\to\infty} b_n = \infty$ die durch die Integrale $\int_{a_n}^{b_n} f(x)\,\mathrm{d}x$ gegebene Folge konvergiert. Ist f uneigentlich Riemann-integrierbar?

Frage 5: Sei $f\colon \mathbb{R} \to \mathbb{R}$ eine stetige Funktion. Es gebe Folgen $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ mit $\lim_{n\to\infty} a_n = -\infty$ sowie $\lim_{n\to\infty} b_n = \infty$, sodass die durch $\int_{a_n}^{b_n} f(x)\,\mathrm{d}x$ gegebene Folge divergiert. Lassen sich immer entsprechende Folgen finden, sodass die Integralfolgen konvergieren, aber verschiedene Grenzwerte auftreten?

Kapitel 6

L^p-Räume

Definition 6.1. Seien $(\Omega, \mathscr{A}, \mu)$ ein Maßraum, $p > 0$ und $f\colon \Omega \to \mathbb{R}$ messbar. f heißt *p-fach μ-integrierbar*, falls gilt:

$$\int |f|^p \, \mathrm{d}\mu < \infty.$$

Die Menge dieser Funktionen bezeichnen wir mit $L^p(\Omega, \mathscr{A}, \mu)$ oder kurz L^p.

Lemma 6.2. *L^p ist ein $\mathbb{R}$-Vektorraum.*

Beweis. Es ist Abgeschlossenheit bezüglich skalarer Multiplikation und Addition zu zeigen. Seien $f \in L^p$ und $\alpha \in \mathbb{R}$. Dann gilt:

$$\int |\alpha f|^p \, \mathrm{d}\mu = |\alpha|^p \int |f|^p \, \mathrm{d}\mu < \infty.$$

Also ist L^p abgeschlossen bezüglich skalarer Multiplikation. Seien $f, g \in L^p$. Dann gilt:

$$\begin{aligned}\int |f+g|^p \, \mathrm{d}\mu &\leqslant \int \big(|f| + |g|\big)^p \, \mathrm{d}\mu \leqslant \int \big(2 \max\{|f|, |g|\}\big)^p \, \mathrm{d}\mu \\ &\leqslant \int 2^p \big(|f|^p + |g|^p\big) \, \mathrm{d}\mu = \int 2^p \, |f|^p \, \mathrm{d}\mu + \int 2^p \, |g|^p \, \mathrm{d}\mu < \infty.\end{aligned}$$

Also ist L^p abgeschlossen bezüglich Addition.

Definition 6.3. $L^\infty(\Omega, \mathscr{A}, \mu) := \{f\colon \Omega \to \mathbb{R} \text{ messbar} \mid f \text{ ist beschränkt } [\mu]\}$.

Lemma 6.4. *L^∞ ist ein $\mathbb{R}$-Vektorraum.*

Beweis. Es ist Abgeschlossenheit bezüglich skalarer Multiplikation und Addition zu zeigen. Seien $f, g \in L^\infty$ mit $|f| < m \in \mathbb{R}$ $[\mu]$ und $|g| < n \in \mathbb{R}$ $[\mu]$ sowie $\alpha \in \mathbb{R}$. Dann ist

$$|\alpha f| < |\alpha| \, m \quad [\mu].$$

Also ist L^∞ abgeschlossen bezüglich skalarer Multiplikation. Außerdem ist

$$|f+g| < m+n \quad [\mu].$$

Also ist L^∞ abgeschlossen bezüglich Addition.

Notation 6.5. Seien $0 < p < \infty$ und $f \in L^p$. Wir schreiben:

$$\|f\|_p := \left(\int |f|^p \, \mathrm{d}\mu\right)^{\frac{1}{p}}$$

sowie

$$\|f\|_\infty := \inf\left\{k \in [0,\infty] \,\middle|\, \mu\big(|f| > k\big) = 0\right\}.$$

$\|f\|_\infty$ heißt auch *μ-essentielles Supremum von f.*

Lemma 6.6 (Young-Ungleichung)**.** *Seien $p, q \in (1,\infty)$ mit $1/p + 1/q = 1$ und $x, y \geqslant 0$. Dann gilt:*

$$xy \leqslant \frac{x^p}{p} + \frac{y^q}{q}.$$

Gleichheit gilt genau dann, wenn $x^p = y^q$.

Beweis. Wegen $x \geqslant 0$, $y \geqslant 0$ gibt es $s, t \in \overline{\mathbb{R}}$ mit

$$x = \exp\left(\frac{s}{p}\right), \qquad y = \exp\left(\frac{t}{q}\right).$$

Da die Exponentialfunktion auf $\overline{\mathbb{R}}$ konvex ist, gilt:

$$xy = \exp\left(\frac{s}{p}\right) \cdot \exp\left(\frac{t}{q}\right) = \exp\left(\frac{s}{p} + \frac{t}{q}\right) \leqslant \frac{1}{p} \cdot \exp(s) + \frac{1}{q} \cdot \exp(t) = \frac{1}{p} \cdot x^p + \frac{1}{q} \cdot y^q. \tag{6.1}$$

Da die Exponentialfunktion sogar strikt konvex ist, gilt in (6.1) Gleichheit genau dann, wenn $s = t$ ist, also wenn $x^p = y^q$ ist.

Satz 6.7 (Hölder-Ungleichung)**.** *Seien $p, q \in (1,\infty)$ mit $1/p + 1/q = 1$ und $f \in L^p$, $g \in L^q$ messbar numerisch. Dann gilt:*

$$\|fg\|_1 \leqslant \|f\|_p \cdot \|g\|_q.$$

Gleichheit ist genau dann erfüllt, wenn gilt:

$$\frac{|f|^p}{\|f\|_p^p} = \frac{|g|^q}{\|g\|_q^q} \quad [\mu]. \tag{6.2}$$

Beweis. Wir verwenden die Young-Ungleichung (Lemma 6.6) mit $x := \frac{|f|}{\|f\|_p}$ und $y := \frac{|g|}{\|g\|_q}$:

$$xy = \frac{|f|}{\|f\|_p} \cdot \frac{|g|}{\|g\|_q} \leqslant \frac{1}{p} \cdot \frac{|f|^p}{\|f\|_p^p} + \frac{1}{q} \cdot \frac{|g|^q}{\|g\|_q^q}. \tag{6.3}$$

Nun integrieren wir bezüglich μ:

$$\int \frac{|f|}{\|f\|_p} \cdot \frac{|g|}{\|g\|_q} \,\mathrm{d}\mu = \frac{\|fg\|_1}{\|f\|_p \cdot \|g\|_q} \leqslant \int \frac{1}{p} \cdot \frac{|f|^p}{\|f\|_p^p} + \frac{1}{q} \cdot \frac{|g|^q}{\|g\|_q^q} \,\mathrm{d}\mu = \frac{1}{p} + \frac{1}{q} = 1. \tag{6.4}$$

Damit folgt die Ungleichung. Gleichheit: [Übung]

Bemerkung 6.8. Der Spezialfall $p = q = 2$ in der Hölder-Ungleichung ist genau die Cauchy-Schwarz-Ungleichung in L^2.

Lemma 6.9. *Seien $f \in L^1$ und $g \in L^\infty$ messbar numerisch. Dann gilt:*

$$\|fg\|_1 \leqslant \|f\|_1 \cdot \|g\|_\infty . \quad \text{[Übung]}$$

Satz 6.10 (Minkowski-Ungleichung)**.** *Seien $p \in [1, \infty]$ und $f, g \in L^p$. Es gilt:*

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

Beweis. Sei zunächst $p \in (1, \infty)$. Dann gibt es $q \in (1, \infty)$ mit $1/p + 1/q = 1$. Folglich gilt: $(p-1) \cdot q = p$ und $1/q = 1 - 1/p$. Wegen $\|f + g\|_p \leqslant \||f| + |g|\|_p$ können wir ohne Einschränkung $f \geqslant 0$ und $g \geqslant 0$ annehmen. Mit der Hölder-Ungleichung gilt:

$$\begin{aligned}
\int (f+g)^p \,\mathrm{d}\mu &= \int f \cdot (f+g)^{p-1} \,\mathrm{d}\mu + \int g \cdot (f+g)^{p-1} \,\mathrm{d}\mu \\
&\leqslant \|f\|_p \cdot \left\|(f+g)^{p-1}\right\|_q + \|g\|_p \cdot \left\|(f+g)^{p-1}\right\|_q \\
&= \left(\|f\|_p + \|g\|_p\right) \left(\int (f+g)^{q\cdot(p-1)} \,\mathrm{d}\mu\right)^{\frac{1}{q}} \\
&= \left(\|f\|_p + \|g\|_p\right) \left(\int (f+g)^{p} \,\mathrm{d}\mu\right)^{1-\frac{1}{p}} \\
&= \left(\|f\|_p + \|g\|_p\right) \left(\int (f+g)^{p} \,\mathrm{d}\mu\right) \left(\int (f+g)^{p} \,\mathrm{d}\mu\right)^{-\frac{1}{p}} .
\end{aligned}$$

Damit folgt der Satz für $p \in (1, \infty)$. Aus der Dreiecksungleichung für den Betrag folgt durch μ-Integration die Behauptung für $p = 1$. Für $p = \infty$ gilt für alle $\varepsilon > 0$ nach Definition des μ-essentiellen Supremums sowie der Dreiecksungleichung des Betrages:

$$\mu\big(|f+g| > \|f\|_\infty + \|g\|_\infty + \varepsilon\big) \leqslant \mu\left(|f| > \|f\|_\infty + \frac{\varepsilon}{2}\right) + \mu\left(|g| > \|g\|_\infty + \frac{\varepsilon}{2}\right) = 0.$$

Somit gilt:

$$\|f + g\|_\infty \leqslant \|f\|_\infty + \|g\|_\infty + \varepsilon$$

für alle $\varepsilon > 0$. Damit ist die Behauptung bewiesen.

Bemerkung 6.11. Sei $p \in [1,\infty]$. Für $f, g \in L^p$ gelten nach dem bisher gezeigten folgende Aussagen:

1. $\|f\|_p \geqslant 0$,
2. $\|0\|_p = 0$,
3. $\|\alpha f\|_p = |\alpha| \cdot \|f\|_p$ für alle $\alpha \in \mathbb{R}$,
4. $\|f+g\|_p \leqslant \|f\|_p + \|g\|_p$.

Folglich ist $\left(L^p, \|\cdot\|_p\right)$ ein halbnormierter $\mathbb{R}$-Vektorraum.

Definition 6.12. Sei $p \in [1,\infty]$. Wir definieren die Relation $\sim$ auf L^p folgendermaßen:

$$f \sim g :\Leftrightarrow \|f-g\|_p = 0.$$

Dass es sich dabei um eine Äquivalenzrelation handelt, folgt direkt aus Bemerkung 6.11. Für die Menge der Äquivalenzklassen schreiben wir:

$$\mathscr{L}^p := {}^{L^p}\!/\!_{\sim}.$$

Lemma 6.13. *$\left(\mathscr{L}^p, \|\cdot\|_p\right)$ ist ein normierter $\mathbb{R}$-Vektorraum.* [Übung]

Definition 6.14. Sei $p \in [1,\infty]$. Eine Folge $(f_n)_{n\in\mathbb{N}}$ in $\mathscr{L}^p$ *konvergiert im p-ten Mittel* (engl.: converges in p-th order mean) gegen $f \in \mathscr{L}^p$, falls

$$\lim_{n\to\infty} \|f_n - f\|_p = 0.$$

Wir schreiben: $f_n \xrightarrow{\mathscr{L}^p} f$.

Beispiel 6.15. An folgendem Beispiel wollen wir erläutern, dass Konvergenz im p-ten Mittel im Allgemeinen nicht hinreichend für punktweise Konvergenz ist. Betrachte den Maßraum $([0,1), \mathscr{B}_{[0,1)}, \lambda)$. Zu jedem $n \in \mathbb{N}$ gibt es genau ein Paar $(k_n, j_n) \in \mathbb{N}^2$ mit $0 \leqslant j_n < 2^{k_n}$ und $n = 2^{k_n} + j_n$. Beachte, dass aus $n \to \infty$ folgt: $k_n \to \infty$. Definiere

$$A_n := A_{k_n, j_n} := \left[j_n \cdot 2^{-k_n}, (j_n+1) \cdot 2^{-k_n}\right) \in \mathscr{B}_{[0,1)}.$$

Für jedes $n \in \mathbb{N}$ erhält man eine messbare Abbildung $f_n := \mathbb{1}_{A_n}$. Es gilt:

$$\int f_n^p \,\mathrm{d}\lambda = \int f_n \,\mathrm{d}\lambda = 2^{-k_n} \xrightarrow{n\to\infty} 0.$$

Es folgt: $f_n \xrightarrow{\mathscr{L}^p} 0$. Für alle $k \in \mathbb{N}$ gilt:

$$\sum_{j=0}^{2^k-1} A_{k,j} = [0,1).$$

Somit gibt es für alle $\omega \in [0,1)$ und alle $k \in \mathbb{N}$ ein $0 \leqslant j < 2^k$ mit $\omega \in A_{k,j}$. Außerdem gilt für alle $0 \leqslant i < 2^k$ mit $i \neq j$: $\omega \notin A_{k,i}$. Also konvergiert f_n nirgendwo punktweise.

Definition 6.16. Sei $X \in \mathscr{L}^p$ eine Zufallsvariable und sei $k \in \mathbb{N}$. Dann heißen

- $m_k := \mathrm{E}X^k$ das *k-te Moment* (engl.: k-th moment),
- $\mathrm{E}\,|X|^k$ das *k-te absolute Moment* (engl.: k-th absolute moment),
- $\mathrm{E}(X - m_1)^k$ das *k-te zentrierte Moment* (engl.: k-th central moment),
- $\mathrm{E}\,|X - m_1|^k$ das *k-te absolute zentrierte Moment* (engl.: k-th absolute central moment).

Wir kennen das erste Moment bereits als Erwartungswert von X und das zweite zentrierte Moment als die Varianz von X.

Lemma 6.17. *Seien μ ein endliches Maß und $1 \leqslant r \leqslant s$. Dann gilt: $\mathscr{L}^s \subseteq \mathscr{L}^r$. Ist μ außerdem ein Wahrscheinlichkeitsmaß, X eine Zufallsvariable und $\mathrm{E}\,|X|^s < \infty$, so ist $\mathrm{E}\,|X|^r < \infty$.*

Beweis. Seien $f\colon \Omega \to \mathbb{R}$ messbar und $\omega \in \Omega$. Aus $|f(\omega)| \leqslant 1$ folgt $|f(\omega)|^r \leqslant 1$, aus $|f(\omega)| > 1$ folgt $|f(\omega)|^r \leqslant |f(\omega)|^s$. Somit gilt für alle $\omega \in \Omega$: $|f(\omega)|^r \leqslant 1 + |f(\omega)|^s$. Sei $f \in \mathscr{L}^s$. Es gilt:

$$\int |f|^r \,\mathrm{d}\mu \leqslant \int 1 + |f|^s \,\mathrm{d}\mu = \mu(\Omega) + \|f\|_s^s < \infty.$$

Daraus folgt: $f \in \mathscr{L}^r$.

Bemerkung 6.18. Im nächsten Abschnitt werden wir Jensens Ungleichung (Satz 6.29) kennenlernen. Wenden wir diese auf die Funktion $\phi\colon x \mapsto x^2$ an, sehen wir, dass für jede Zufallsvariable X mit $\mathrm{E}X^2 < \infty$ gilt: $\mathrm{E}\,|X| < \infty$. Damit ist $\mathrm{E}X^2 < \infty$ genau dann, wenn $\mathrm{Var}(X) < \infty$.

Beispiel 6.19 (Korrelationskoeffizient). Seien X, Y Zufallsvariablen mit endlicher Varianz.

$$\rho_{X,Y} := \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}$$

mit $\mathrm{Cov}(X,Y) := \mathrm{E}\big((X - \mathrm{E}X)(Y - \mathrm{E}Y)\big)$ heißt *Korrelationskoeffizient* von X und Y. Gilt $\rho_{X;Y} = 0$, so heißen X und Y *unkorreliert* Es ist $|\rho_{X,Y}| \leqslant 1$. $\rho_{X,Y} = 1$ gilt genau dann, wenn es ein $a > 0$ gibt, sodass

$$X - \mathrm{E}X = a \cdot (Y - \mathrm{E}Y) \quad [P].$$

$\rho_{X,Y} = -1$ gilt genau dann, wenn es ein $a < 0$ gibt, sodass

$$X - \mathrm{E}X = a \cdot (Y - \mathrm{E}Y) \quad [P].$$

[Übung]

Damit ist $\rho_{X,Y}$ ein Maß für die lineare Abhängigkeit von Zufallsvariablen.

6.1 Konvexe Funktionen und Jensens Ungleichung

Ziel dieses Abschnitts ist der Beweis von Jensens Ungleichung (Satz 6.29). Hierfür wollen wir zunächst genauer auf konvexe Funktionen eingehen.

Definition 6.20. Eine Teilmenge $A \subseteq \mathbb{R}^k$ heißt *konvex* (engl.: convex), falls für je zwei Elemente $x, y \in A$ die Verbindungsstrecke zwischen x und y ganz in A liegt, das heißt, falls für alle $\lambda \in [0,1]$ gilt: $\lambda x + (1-\lambda)y \in A$. Eine Funktion $\phi\colon \mathbb{R}^k \to \overline{\mathbb{R}}$ heißt *konvex*, falls die Menge

$$K_\phi := \big\{(x,y) \in \mathbb{R}^k \times \mathbb{R} \,\big|\, y \geqslant \phi(x)\big\} \subseteq \mathbb{R}^{k+1}$$

konvex ist.

Satz 6.21. *Eine Funktion $\phi\colon \mathbb{R}^k \to \mathbb{R} \cup \{\infty\}$ ist genau dann konvex, wenn für je zwei Elemente $x, y \in \mathbb{R}^k$ und alle $\lambda \in [0,1]$ gilt:*

$$\phi\big(\lambda x + (1-\lambda)y\big) \leqslant \lambda\phi(x) + (1-\lambda)\phi(y). \tag{6.5}$$

Beweis. Sei zunächst ϕ konvex. Betrachten wir Punkte der Form $\big(x, \phi(x)\big) \in K_\phi$, so liefert die Konvexität von K_ϕ für alle $x, y \in \mathbb{R}^k$ und $\lambda \in [0,1]$:

$$\Big(\lambda x + (1-\lambda)y, \phi\big(\lambda x + (1-\lambda)y\big)\Big) \leqslant \lambda\big(x, \phi(x)\big) + (1-\lambda)\big(y, \phi(y)\big).$$

Somit gilt insbesondere (6.5).

Ist ϕ nicht konvex, dann gibt es $(a,b), (c,d) \in K_\phi$ und $\lambda \in [0,1]$ mit

$$\lambda(a,b) + (1-\lambda)(c,d) \notin K_\phi.$$

Nach Definition von K_ϕ folgt:

$$\lambda\phi(a) + (1-\lambda)\phi(c) \leqslant \lambda b + (1-\lambda)d < \phi\big(\lambda a + (1-\lambda)c\big),$$

im Widerspruch zu (6.5).

Lemma 6.22. *Sei $\phi\colon \mathbb{R}^k \to \mathbb{R} \cup \{\infty\}$ konvex. Dann ist $\phi^{-1}(\mathbb{R})$ konvex.*

Beweis. Seien $x, y \in \phi^{-1}(\mathbb{R})$ sowie $\lambda \in [0,1]$. Nach Satz 6.21 gilt:

$$\phi\big(\lambda x + (1-\lambda)y\big) \leqslant \lambda\phi(x) + (1-\lambda)\phi(y) < \infty.$$

Also ist $\lambda x + (1-\lambda)y \in \phi^{-1}(\mathbb{R})$.

Satz 6.23. *Sei $\phi\colon \mathbb{R}^k \to \mathbb{R} \cup \{\infty\}$ konvex. Setze $M := \big(\phi^{-1}(\mathbb{R})\big)^\circ$. Dann ist $\phi|_M$ stetig.*

Beweis. Sei $A \subseteq \mathbb{R}^k$ offen. Eine Funktion $f\colon A \to \mathbb{R}$ ist genau dann stetig, wenn sie in jeder Komponente stetig ist, das heißt, wenn für alle $i \in \{1, \dots, k\}$ und alle $x_j \in \mathbb{R}$ mit $j \in \{1, \dots, k\} \setminus \{i\}$ die Funktion

$$g\colon A \cap \big(\{x_1\} \times \cdots \times \{x_{i-1}\} \times \mathbb{R} \times \{x_{i+1}\} \times \cdots \times \{x_k\}\big) \to \mathbb{R}, \quad x_i \mapsto f(x_1, \dots, x_k)$$

stetig ist. g ist zudem automatisch konvex, wenn bereits f konvex ist, denn mit Satz 6.21 gilt für alle $\lambda \in [0, 1]$:

$$\begin{aligned} g\big(\lambda a + (1-\lambda)b\big) &= f\big(x_1, \dots, x_{i-1}, \lambda a + (1-\lambda)b, x_{i+1}, \dots, x_k\big) \\ &\leqslant \lambda f(x_1, \dots, x_{i-1}, a, x_{i+1}, \dots, x_k) + (1-\lambda) f(x_1, \dots, x_{i-1}, b, x_{i+1}, \dots, x_k) \\ &= \lambda g(a) + (1-\lambda) g(b). \end{aligned}$$

Somit lässt sich der mehrdimensionale Fall auf den eindimensionalen Fall $k = 1$ zurückführen. Wir zeigen eine etwas stärkere Eigenschaft, nämlich dass für eine offene Menge $A \subseteq \mathbb{R}$ eine konvexe Funktion $\psi\colon A \to \mathbb{R}$ in jedem Punkt zumindest einseitig differenzierbar ist. Sei $a \in A$. Da A offen ist, gibt es ein $\delta > 0$ mit $(a - \delta, a + \delta) \subseteq A$. Definiere $B := (a - \delta, a) \cup (a, a + \delta)$. Für jede Stelle $x \in B$ betrachten wir die Steigung der Sekante durch die Punkte $\big(x, \psi(x)\big)$, $\big(a, \psi(a)\big)$. Wir erhalten die Funktion:

$$s\colon B \to \mathbb{R}, \quad x \mapsto \frac{\psi(a) - \psi(x)}{a - x}.$$

Wir zeigen nun, dass diese Funktion monoton wächst. Seien $x, y \in B$ mit $x < y$. Es treten drei ähnlich behandelbare Fälle auf:

- Gilt $x < y < a$, so gibt es wegen des Zwischenwertsatzes ein $\lambda \in (0, 1)$ mit $y = \lambda x + (1-\lambda)a$. Da ψ konvex ist, folgt mit Satz 6.21:

  $$\psi(y) = \psi\big(\lambda x + (1-\lambda)a\big) \leqslant \lambda \psi(x) + (1-\lambda)\psi(a).$$

  Wir multiplizieren diese Ungleichung zunächst mit (-1) und addieren $\psi(a)$. Es ergibt sich die äquivalente Ungleichung:

  $$\psi(a) - \psi(y) \geqslant \psi(a) - \lambda\psi(x) - (1-\lambda)\psi(a) = \lambda\big(\psi(a) - \psi(x)\big).$$

  Division durch $\lambda(a - x) > 0$ liefert:

  $$\frac{\psi(a) - \psi(y)}{\lambda(a - x)} \geqslant \frac{\psi(a) - \psi(x)}{a - x}.$$

  Insgesamt erhalten wir:

  $$\begin{aligned} s(y) = \frac{\psi(a) - \psi(y)}{a - y} &= \frac{\psi(a) - \psi(y)}{a - \lambda x + (1-\lambda)a} \\ &= \frac{\psi(a) - \psi(y)}{\lambda(a - x)} \geqslant \frac{\psi(a) - \psi(x)}{a - x} = s(x). \end{aligned}$$

- Gilt $a < x < y$, so erhalten wir ein $\lambda \in (0,1)$ mit $x = \lambda a + (1-\lambda)y$. Da ψ konvex ist, folgt mit Satz 6.21:

$$\psi(x) \leqslant \lambda\psi(a) + (1-\lambda)\psi(y).$$

Durch Umformen erhalten wir die Ungleichung:

$$\psi(a) - \psi(x) \geqslant (1-\lambda)\big(\psi(a) - \psi(y)\big).$$

Division durch $(1-\lambda)(a-y) < 0$ liefert nun:

$$\frac{\psi(a) - \psi(x)}{(1-\lambda)(a-y)} \leqslant \frac{\psi(a) - \psi(y)}{a-y}.$$

Insgesamt erhalten wir:

$$\begin{aligned} s(x) = \frac{\psi(a) - \psi(x)}{a-x} &= \frac{\psi(a) - \psi(x)}{a - \lambda a + (1-\lambda)y} \\ &= \frac{\psi(a) - \psi(x)}{(1-\lambda)(a-y)} \leqslant \frac{\psi(a) - \psi(y)}{a-y} = s(y). \end{aligned}$$

- Gilt $x < a < y$, so erhalten wir ein $\lambda \in (0,1)$ mit $a = \lambda x + (1-\lambda)y$. Insbesondere folgt:

$$y = \frac{a - \lambda x}{1-\lambda}$$

und wir erhalten den Zusammenhang:

$$-\frac{1-\lambda}{\lambda}(a-y) = -\frac{1-\lambda}{\lambda}\left(a - \frac{a-\lambda x}{1-\lambda}\right) = a - x. \tag{6.6}$$

Da ψ konvex ist, folgt mit Satz 6.21:

$$\psi(a) \leqslant \lambda\psi(x) + (1-\lambda)\psi(y).$$

Durch Umformen erhalten wir die Ungleichung:

$$(1-\lambda)\big(\psi(a) - \psi(y)\big) \leqslant -\lambda\big(\psi(a) - \psi(x)\big).$$

Division durch $(1-\lambda)(a-y) < 0$ liefert nun:

$$\frac{\psi(a) - \psi(y)}{a-y} \geqslant \frac{\psi(a) - \psi(x)}{-\frac{1-\lambda}{\lambda}(a-y)}.$$

Zusammen mit (6.6) erhalten wir:

$$s(y) = \frac{\psi(a) - \psi(y)}{a-y} \geqslant \frac{\psi(a) - \psi(x)}{-\frac{1-\lambda}{\lambda}(a-y)} = \frac{\psi(a) - \psi(x)}{a-x} = s(x).$$

Somit ist die Funktion s monoton wachsend. Insbesondere ist ihre Einschränkung $s_- := s|_{(a-\delta,a)}$ monoton wachsend und nach oben beschränkt. Aus der Analysis ist bekannt, dass in diesem Fall der Grenzwert $\lim_{x\to a} s_-(x)$ existiert. Entsprechend ist die Funktion $s_+ := s|_{(a,a+\delta)}$ monoton wachsend und nach unten beschränkt, somit existiert der Grenzwert $\lim_{x\to a} s_+(x)$. Nach Definition von s gilt also:

$$\lim_{x\to a} s_-(x) = \lim_{x\nearrow a} \frac{\psi(a)-\psi(x)}{a-x} \leqslant \lim_{x\searrow a} \frac{\psi(a)-\psi(x)}{a-x} = \lim_{x\to a} s_+(x).$$

Insbesondere folgt die Stetigkeit von ψ an a, denn es gilt:

$$\lim_{x\nearrow a} \psi(a)-\psi(x) = \lim_{x\to a} s_-(x)\cdot \lim_{x\nearrow a}(a-x) = 0$$

sowie analog:

$$\lim_{x\searrow a} \psi(a)-\psi(x) = \lim_{x\to a} s_-(x)\cdot \lim_{x\searrow a}(a-x) = 0.$$

Insgesamt folgt:

$$\lim_{x\to a} f(x) = f(a).$$

Dies zeigt die Stetigkeit von f an a.

Bemerkung 6.24. Der Beweis des Satzes 6.23 zeigt außerdem, dass wir auf sinnvolle Weise Tangenten an eine konvexe Funktion $\psi\colon A \to \mathbb{R}$ anlegen können. Achtung: $A \subseteq \mathbb{R}$ muss hierfür offen sein. Hierzu können wir für eine Stelle $a \in A$ jede affine Gerade durch $\big(a,\psi(a)\big)$ benutzen, deren Steigung m die folgende Bedingung erfüllt:

$$\lim_{x\nearrow a} \frac{\psi(a)-\psi(x)}{a-x} \leqslant m \leqslant \lim_{x\searrow a} \frac{\psi(a)-\psi(x)}{a-x}.$$

Im Allgemeinen k-dimensionalen Fall erhalten wir entsprechend durch Betrachtung der partiellen Ableitungen Bedingungen, welche die Konstruktion von Tangentialhyperebenen ermöglichen.

Satz 6.25. *Sei $\phi\colon \mathbb{R} \to \mathbb{R}\cup\{\infty\}$ eine Funktion, sodass die Menge $M := \phi^{-1}(\mathbb{R})$ offen und $\phi|_M$ zweimal differenzierbar ist. Ist $\phi''(x) \geqslant 0$ für alle $x \in M$, so ist ϕ konvex.*

Beweis. Wir führen einen indirekten Beweis. Sei ϕ nicht konvex. Dann gibt es nach Satz 6.21 $x, y \in M$ mit $x < y$ und $\lambda \in (0,1)$, sodass gilt:

$$\phi\big(\lambda x + (1-\lambda)y\big) > \lambda\phi(x) + (1-\lambda)\phi(y).$$

Durch Subtraktion von $\phi(x)$ und Multiplikation mit $\big((1-\lambda)(y-x)\big)^{-1} > 0$ erhalten wir:

$$\frac{\phi\big(\lambda x + (1-\lambda)y\big) - \phi(x)}{\big(\lambda x + (1-\lambda)y\big) - x} > \frac{\phi(y)-\phi(x)}{y-x}$$

sowie durch Subtraktion von $\phi(y)$ und Multiplikation mit $\big(-\lambda(y-x)\big)^{-1} < 0$:

$$\frac{\phi(y) - \phi\big(\lambda x + (1-\lambda)y\big)}{y - \big(\lambda x + (1-\lambda)y\big)} < \frac{\phi(y)-\phi(x)}{y-x}.$$

Da ϕ stetig auf $[x, y]$ und differenzierbar auf (x, y) ist, folgt mit dem Mittelwertsatz, dass es Zwischenstellen $\xi \in \big((x, \lambda x + (1-\lambda)y\big)$ und $\eta \in \big(\lambda x + (1-\lambda)y, y\big)$ gibt mit

$$\begin{aligned}\phi'(\xi) &= \frac{\phi\big(\lambda x + (1-\lambda)y\big) - \phi(x)}{\big(\lambda x + (1-\lambda)y\big) - x},\\ \phi'(\eta) &= \frac{\phi(y) - \phi\big(\lambda x + (1-\lambda)y\big)}{y - \big(\lambda x + (1-\lambda)y\big)}.\end{aligned}$$

Insbesondere gilt: $\xi < \eta$ und $\phi'(\xi) > \phi'(\eta)$. Daraus folgt:

$$\frac{\phi'(\eta) - \phi'(\xi)}{\eta - \xi} < 0.$$

Da ϕ' stetig auf $[\xi, \eta]$ und differenzierbar auf (ξ, η) ist, folgt mit dem Mittelwertsatz, dass es eine Zwischenstelle $\zeta \in (\xi, \eta)$ gibt mit $\phi''(\zeta) < 0$.

Lemma 6.26. *Der Abschluss einer konvexen Menge ist konvex.*

Beweis. Sei A eine konvexe Menge. Zu $x, y \in \overline{A}$ gibt es Folgen $(x_n)_{n\in\mathbb{N}}$ und $(y_n)_{n\in\mathbb{N}}$ in A mit $\lim_{n\to\infty} x_n = x$ und $\lim_{n\to\infty} y_n = y$. Aus der Konvexität von A folgt für alle $n \in \mathbb{N}$ und $\lambda \in [0,1]$: $z_n := \lambda x_n + (1-\lambda)y_n \in A$. Somit ist $(z_n)_{n\in\mathbb{N}}$ eine konvergente Folge in A und es folgt:

$$\lambda x + (1-\lambda)y = \lim_{n\to\infty} z_n \in \overline{A}.$$

Lemma 6.27 (Seperating Hyperplane Theorem)**.** *Sei $A \subseteq \mathbb{R}^k$ konvex und $b \in \overline{A}^c$. Dann gibt es ein $n \in \mathbb{R}^k \backslash \{0\}$, sodass die affine Hyperebene*

$$H_n(b) := \left\{x \in \mathbb{R}^k \,\middle|\, \langle x - b, n\rangle = 0\right\}$$

durch b mit Normalenvektor n die Mengen A und $\{b\}$ voneinander trennt in dem Sinne, dass gilt:

$$A \subseteq H_n^+(b) := \left\{x \in \mathbb{R}^k \,\middle|\, \langle x - b, n\rangle \geqslant 0\right\}$$

sowie

$$\{b\} \subseteq H_n^-(b) := \left\{x \in \mathbb{R}^k \,\middle|\, \langle x - b, n\rangle \leqslant 0\right\}.$$

[Übung]

Bemerkung 6.28. Bemerkung 6.24 lässt vermuten, dass wir Lemma 6.27 sogar noch verbessern können. In der Tat kann man mittels Tangentialhyperebenen sogar Punkte im Rand von $\overline{A}$ von A trennen. Es geht sogar noch allgemeiner: Zwei konvexe Mengen $A, B \subseteq \mathbb{R}^k$ lassen sich genau dann durch eine Hyperebene trennen, wenn $A^\circ \cap B^\circ = \varnothing$ ist. Siehe hierzu [7].

Satz 6.29 (Jensens Ungleichung)**.** *Sei $(\Omega, \mathscr{A}, P)$ ein Wahrscheinlichkeitsraum. Außerdem sei $\phi\colon \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ eine konvexe Funktion. Definiere $M := \big(\phi^{-1}(\mathbb{R})\big)^{\circ}$. Weiterhin seien $f, \phi \circ f \in \mathscr{L}^1(\Omega, \mathscr{A}, P)$ mit $f(\Omega) \subseteq M$. Dann gilt:*

$$\int \phi \circ f \,\mathrm{d}P \geqslant \phi\left(\int f \,\mathrm{d}P\right).$$

Beweis. Bevor wir den eigentlichen Beweis angehen, ist eine Vorüberlegung nötig. Betrachte die Menge der durch ϕ majorisierten affinen Funktionen

$$F := \{L \mid L\colon \mathbb{R} \to \mathbb{R} \text{ ist affin und } L(x) \leqslant \phi(x) \text{ für alle } x \in \mathbb{R}\}.$$

Wir überzeugen uns von der Gültigkeit der folgenden Gleichung.

$$\phi(x) = \sup_{L \in F} L(x) \quad \text{für alle } x \in M. \tag{6.7}$$

Für alle $x \in M$ gilt nach Definition von F: $\phi(x) \geqslant L(x)$ für alle $L \in F$. Somit ist $\phi(x) \geqslant \sup_{L \in F} L(x)$. Seien $x \in M$ und $\varepsilon > 0$. Setze $b := \big(x, \phi(x) - \varepsilon\big)$. Wir zeigen nun: $b \notin \overline{K}_\phi$. Angenommen, es würde $b \in \overline{K}_\phi$ gelten. Dann gäbe es eine Folge $\big((x_j, y_j)\big)_{j \in \mathbb{N}}$ in K_ϕ mit $\lim_{j \to \infty} x_j = x$ und $\lim_{j \to \infty} y_j = \phi(x) - \varepsilon$. Nach Definition von K_ϕ gilt für alle $j \in \mathbb{N}$: $y_j \geqslant \phi(x_j)$. Nach Satz 6.23 ist ϕ an x stetig. Somit erhalten wir den folgenden Widerspruch:

$$\phi(x) - \varepsilon = \lim_{j \to \infty} y_j \geqslant \lim_{j \to \infty} \phi(x_j) = \phi(x).$$

Also ist $b \notin \overline{K}_\phi$. Nach Lemma 6.27 erhalten wir eine Hyperebene $H_n(b)$ (in diesem Fall eine Gerade), welche K_ϕ und $\{b\}$ voneinander trennt. Beachte, dass es sich bei $H_n(b)$ um den Graphen einer affinen Funktion L' handelt, da ϕ in einer Umgebung von $x \in M$ endlich ist. Da $K_\phi \subseteq H_n^+(b)$, ist ϕ eine Majorante von L'. Somit folgt: $\sup_{LinF} L(x) \geqslant L'(x) = \phi(x) - \varepsilon$. Also ist $\sup_{LinF} L(x) \geqslant \phi(x)$. Insgesamt erhalten wir (6.7).

Wir zeigen nun Jensens Ungleichung. Für alle $L' \in F$ gilt: $L' \circ f \leqslant \sup_{L \in F} L \circ f$. Mit der Monotonie des Integrals (siehe die Regeln 3.6 sowie Bemerkung 3.18) folgt:

$$\int L' \circ f \,\mathrm{d}P \leqslant \int \sup_{L \in F} L \circ f \,\mathrm{d}P.$$

Mit (6.7) und $f(\Omega) \subseteq M$ folgt:

$$\phi \circ f = \sup_{L \in F} L \circ f.$$

Insgesamt folgt:

$$\sup_{L \in F} \int L \circ f \,\mathrm{d}P \leqslant \int \sup_{L \in F} L \circ f \,\mathrm{d}P = \int \phi \circ f \,\mathrm{d}P.$$

Außerdem gilt für jede affine Funktion L mit $L(x) = mx + a$:

$$\int L \circ f \,\mathrm{d}P = m \int f \,\mathrm{d}P + a \int 1 \,\mathrm{d}P = L\left(\int f \,\mathrm{d}P\right).$$

Wir erhalten:

$$\int \phi \circ f \,\mathrm{d}P \geqslant \sup_{L \in F} \int L \circ f \,\mathrm{d}P = \sup_{L \in F} L\left(\int f \,\mathrm{d}P\right) = \phi\left(\int f \,\mathrm{d}P\right).$$

Beispiel 6.30. Sei X eine Zufallsvariable. Da die Exponentialfunktion konvex ist, folgt aus Satz 6.29 mit $\phi = \exp$:

$$\mathrm{E}(\exp \circ X) \geqslant \exp(\mathrm{E}X).$$

6.2 ★Vollständigkeit der $\mathscr{L}^p$-Räume

Satz 6.31 (Fischer, Riesz). *Sei $p \in [1, \infty]$. Dann ist $\left(\mathscr{L}^p, \|\cdot\|_p\right)$ ein Banachraum, also ein vollständiger normierter Vektorraum.*

Beweis. Nach Lemma 6.13 ist nur noch die Vollständigkeit zu zeigen. Wir betrachten zunächst den Fall $p = \infty$. Sei $(\tilde{f}_n)_{n \in \mathbb{N}}$ eine Cauchyfolge in $\mathscr{L}^\infty$. Nach Definition des μ-essentiellen Supremums findet man für alle $n \in \mathbb{N}$ eine Funktion f_n in der Äquivalenzklasse von $\tilde{f}_n$, für die das μ-essentielle Supremum gleich dem Supremum ist. Aus

$$\lim_{n,m \to \infty} \|f_n - f_m\|_\infty = 0$$

folgt $\lim_{n,m \to \infty} |f_n(\omega) - f_m(\omega)| = 0$ für alle $\omega \in \Omega$. Somit liegt punktweise Konvergenz auf ganz Ω vor. Sei $f := \lim_{n \to \infty} f_n \in \mathscr{L}^\infty$. Für alle $\varepsilon > 0$ wähle $N(\varepsilon) \in \mathbb{N}$, sodass $|f_n(\omega) - f_m(\omega)| < \varepsilon/2$ für alle $n, m > N(\varepsilon)$ und alle $\omega \in \Omega$. Wähle außerdem für jedes $\omega \in \Omega$ ein $n_\omega > N(\varepsilon)$ mit $|f(\omega) - f_{n_\omega}(\omega)| < \varepsilon/2$. Es gilt:

$$\|f - f_n\|_\infty = \sup_{\omega \in \Omega} |f(\omega) - f_n(\omega)| \leqslant \sup_{\omega \in \Omega} \big(|f(\omega) - f_{n_\omega}(\omega)| + |f_{n_\omega}(\omega) - f_n(\omega)|\big) < \varepsilon$$

für alle $n > N(\varepsilon)$. Somit gilt:

$$\lim_{n \to \infty} \|f - f_n\|_\infty = 0.$$

Sei nun $p < \infty$ und sei außerdem $(f_n)_{n \in \mathbb{N}}$ eine Cauchyfolge in $\mathscr{L}^p$, das heißt $\lim_{m,n \to \infty} \|f_n - f_m\|_p = 0$. Obwohl die Folge $\big(f_n(\omega)\big)_{n \in \mathbb{N}}$ für kein $\omega \in \Omega$ konvergent sein muss, können wir jedoch eine Teilfolge von $(f_n)_{n \in \mathbb{N}}$ konstruieren, welche μ-fast-überall punktweise konvergiert. Da $(f_n)_{n \in \mathbb{N}}$ eine Cauchyfolge ist, können wir eine Teilfolge $(f_{n(k)})_{k \in \mathbb{N}}$ finden, sodass $\left\|f_m - f_{n(k)}\right\|_p < (1/2)^k$ für alle $m \geqslant n(k)$. Definiere die Folge $(g_k)_{k \in \mathbb{N}}$ durch

$$g_k := \left|f_{n(1)}\right| + \sum_{j=1}^{k} \left|f_{n(j+1)} - f_{n(j)}\right|$$

für alle $k \in \mathbb{N}$. Es gilt:

$$\begin{aligned}\|g_k^p\|_1 = \|g_k\|_p^p &= \left\| \left|f_{n(1)}\right| + \sum_{j=1}^{k} \left|f_{n(j+1)} - f_{n(j)}\right| \right\|_p^p \\ &\leqslant \left(\left\|f_{n(1)}\right\|_p + \sum_{j=1}^{k} \left\|f_{n(j+1)} - f_{n(j)}\right\|_p \right)^p \leqslant \left(\left\|f_{n(1)}\right\|_p + \sum_{j=1}^{k} \left(\frac{1}{2}\right)^k \right)^p < \infty.\end{aligned}$$

Beachte, dass $(g_k)_{k\in\mathbb{N}}$ isoton ist. Definiere $g := \lim_{k\to\infty} g_k$. Nach dem Satz über die monotone Konvergenz folgt mit der obigen Abschätzung:

$$\int g^p \,\mathrm{d}\mu = \int \lim_{k\to\infty} g_k^p \,\mathrm{d}\mu = \lim_{k\to\infty} \int g_k^p \,\mathrm{d}\mu < \infty.$$

Somit ist $g \in \mathscr{L}^p$. Insbesondere ist g^p μ-fast-überall beschränkt und folglich konvergiert die Reihe

$$\sum_{j\in\mathbb{N}} \left|f_{n(j+1)}(\omega) - f_{n(j)}(\omega)\right|.$$

für μ-fast-alle $\omega \in \Omega$. Damit konvergiert auch die Reihe

$$f_{n(1)}(\omega) + \sum_{j\in\mathbb{N}} \big(f_{n(j+1)}(\omega) - f_{n(j)}(\omega)\big)$$

für μ-fast-alle $\omega \in \Omega$. Beachte, dass die k-te Partialsumme dieser Reihe $f_{n(k+1)}(\omega)$ entspricht. Folglich ist $\big(f_{n(k)}(\omega)\big)_{k\in\mathbb{N}}$ konvergent für μ-fast-alle $\omega \in \Omega$. Sei $A \in \mathscr{A}$ die Menge der Punkte, an welchen Konvergenz vorliegt. Definiere

$$f\colon \Omega \to \mathbb{R}, \quad \omega \mapsto \begin{cases} \lim_{k\to\infty} f_{n(k)}(\omega) & \text{falls } \omega \in A \\ 0 & \text{sonst} \end{cases}.$$

Als Grenzwert messbarer Funktionen ist f messbar. Es bleibt zu zeigen: $f_n \xrightarrow{\mathscr{L}^p} f$. Für alle $\varepsilon > 0$ sei $r \in \mathbb{N}$ groß genug, sodass für alle $s, t \geqslant n(r)$ gilt:

$$\|f_s - f_t\|_p < \varepsilon.$$

Somit gilt für alle $k \geqslant r$ und $m > n(r)$:

$$\left\|f_m - f_{n(k)}\right\|_p < \varepsilon.$$

Mit dem Lemma von Fatou (Satz 4.6) erhalten wir:

$$\int |f - f_m|^p \,\mathrm{d}\mu = \int \liminf_{k\to\infty} \left|f_{n(k)} - f_m\right|^p \,\mathrm{d}\mu \leqslant \liminf_{k\to\infty} \int \left|f_{n(k)} - f_m\right|^p \,\mathrm{d}\mu \leqslant \varepsilon^p$$

für alle $m > n(r)$. Somit ist $f - f_m \in \mathscr{L}^p$ für alle $m > n(r)$. Folglich ist auch $f = f - f_m + f_m \in \mathscr{L}^p$ und es gilt:

$$\lim_{m\to\infty} \|f - f_m\|_p = 0.$$

Bemerkung 6.32. Ein Paar von Zahlen $p, q \geqslant 1$ mit $1/p + 1/q = 1$ nennen wir auch *duale Zahlen*, da $\mathscr{L}^p$ den Dualraum von $\mathscr{L}^q$ darstellt. Siehe hierzu [9], Seiten 230 f.

Beispiel 6.33. Betrachte den Maßraum $(\mathbb{N}, \mathscr{P}(\mathbb{N}), \mu)$ mit dem Zählmaß μ. Da zu jeder Abbildung $f\colon \mathbb{N} \to \mathbb{R}$ genau eine Folge $(x_i)_{i\in\mathbb{N}}$ mit $f(i) = x_i$ gehört, gilt:

$$\mathscr{L}^p = L^p = l^p = \left\{ x = (x_i)_{i\in\mathbb{N}} \,\middle|\, \|x\|_p < \infty \right\}.$$

Außerdem gilt:

$$\|x\|_p^p = \sum_{n\in\mathbb{N}} |x_n|^p \qquad \|x\|_\infty = \sup_{n\in\mathbb{N}} |x_n|\,.$$

Nach Satz 6.31 ist l^p ein Banachraum.

Übungsaufgaben:

Übung 1: Beweisen Sie in der Hölder-Ungleichung (Satz 6.7) den Fall der Gleichheit.

Übung 2: Beweisen Sie Lemma 6.9.

Übung 3: Beweisen Sie Lemma 6.13.

Übung 4: Beweisen Sie Beispiel 6.19.

Übung 5: Zeigen Sie Jensens Ungleichung für k-dimensionale Zufallsvektoren. Sei dazu $\phi\colon \mathbb{R}^k \to \mathbb{R}$ konvex und sei $X = (X_1, \ldots, X_k)$ ein Zufallsvektor mit Werten in $M := \left(\phi^{-1}(\mathbb{R})\right)^\circ$, sodass $\mathrm{E}X$ und $\mathrm{E}(\phi \circ X)$ existieren. Dann gilt:

$$\mathrm{E}(\phi \circ X) \geqslant \phi(\mathrm{E}X).$$

Übung 6: Zeigen Sie, dass für eine positive Zufallsvariable X mit $\mathrm{E}X < \infty$ und $\mathrm{E}(\log \circ X) < \infty$ gilt:

$$\mathrm{E}(\log \circ X) \leqslant \log(\mathrm{E}X).$$

Lernziel-Kontrolle:

Frage 1: Was besagt die Hölder-Ungleichung?

Frage 2: Welche Beziehung besteht zwischen der Cauchy-Schwarz-Ungleichung und der Hölder-Ungleichung?

Frage 3: Impliziert fast-überall bestehende Konvergenz L^p-Konvergenz oder umgekehrt?

Frage 4: Für welche p ist L^p ein Hilbertraum?

Frage 5: Folgt aus $p' < p$ immer $L^p \subseteq L^{p'}$? Was heißt dies für die Existenz der Momente von Zufallsvariablen?

Frage 6: Was besagt Jensens Ungleichung und wie ist die Beweisidee?

Kapitel 7

Maße mit Dichten

7.1 Der Satz von Radon-Nikodym

Viele Maße und Verteilungen können durch Integrale bezüglich Standardmaßen wie beispielsweise dem Lebesgue-Maß über Funktionen einfach beschrieben werden. Der Satz von Radon-Nikodym gibt Aufschluss darüber, wann dies möglich ist.

Definition 7.1. Seien $(\Omega, \mathscr{A}, \mu)$ ein Maßraum und $f\colon \Omega \to \overline{\mathbb{R}}$ messbar numerisch mit $f \geqslant 0$. Dann heißt

$$\nu\colon \mathscr{A} \to [0, \infty), \quad A \mapsto \int_A f \,\mathrm{d}\mu$$

das *Maß mit Dichte f bezüglich μ* (engl.: measure with densitiy f with respect to μ). Wir schreiben: $\nu = f\mu$.

Bemerkung 7.2.

1. ν ist ein Maß auf $\mathscr{A}$, denn
   - $\nu(\varnothing) = \int_\varnothing f \,\mathrm{d}\mu = 0$.
   - $\nu \geqslant 0$, da $f \geqslant 0$.
   - Für $\{A_j\}_{j\in\mathbb{N}} \subseteq \mathscr{A}$ paarweise disjunkt gilt nach dem Satz über monotone Konvergenz:

$$\begin{aligned}\nu\left(\sum_{j\in\mathbb{N}} A_j\right) = \int_{\sum_{j\in\mathbb{N}} A_j} f \,\mathrm{d}\mu &= \int \sum_{j\in\mathbb{N}} \mathbb{1}_{A_j} \cdot f \,\mathrm{d}\mu \\ &= \sum_{j\in\mathbb{N}} \int \mathbb{1}_{A_j} \cdot f \,\mathrm{d}\mu = \sum_{j\in\mathbb{N}} \nu(A_j).\end{aligned}$$

2. Gilt $\mu(f) = 1$, so ist ν ein Wahrscheinlichkeitsmaß.

3. Im Spezialfall $(\Omega, \mathscr{A}, \mu) = (\mathbb{R}^k, \mathscr{B}^k, \lambda^k)$ heißt f die *Lebesgue-Dichte von* ν. Gilt insbesondere $\nu = P^X$ für einen Zufallsvektor $X \in \mathbb{R}^k$, so heißt f *Wahrscheinlichkeitsdichte von* X *bezüglich* λ^k.

Satz 7.3. *Seien* $f, g \in \mathscr{E}^*$ *und* $\nu := f\mu$. *Dann ist* g *genau dann* ν*-integrierbar, wenn* $g \cdot f$ μ*-integrierbar ist. In diesem Fall gilt:*

$$\int g \,\mathrm{d}\nu = \int g \cdot f \,\mathrm{d}\mu.$$

Beweis. Sei zunächst $g \in \mathscr{E}$ und $g = \sum_{i=1}^n \alpha_i \mathbb{1}_{A_i}$ eine Normaldarstellung. Es gilt:

$$\int g \,\mathrm{d}\nu = \sum_{i=1}^n \alpha_i \nu(A_i) = \sum_{i=1}^n \alpha_i \int f \cdot \mathbb{1}_{A_i} \,\mathrm{d}\mu = \int f \cdot \sum_{i=1}^n \alpha_i \mathbb{1}_{A_i} \,\mathrm{d}\mu = \int f \cdot g \,\mathrm{d}\mu.$$

Seien nun $g \in \mathscr{E}^*$ und $(u_k)_{k\in\mathbb{N}}$ eine monoton gegen g steigende Folge in $\mathscr{E}$. Dann ist $(f \cdot u_k)_{k\in\mathbb{N}}$ eine monoton gegen $f \cdot g$ steigende Folge und mit dem Satz über monotone Konvergenz folgt:

$$\int g \,\mathrm{d}\nu = \lim_{k\to\infty} \int u_k \,\mathrm{d}\nu = \lim_{k\to\infty} \int f \cdot u_k \,\mathrm{d}\mu = \int f \cdot g \,\mathrm{d}\mu.$$

Beispiel 7.4.

- Seien $\Omega := \{\omega_1, \omega_2, \ldots\}$ abzählbar, $\mu := \sum_{i\in\mathbb{N}} \delta_{\omega_i}$ das Zählmaß, $f\colon \Omega \to \mathbb{R}$ mit $f \geqslant 0$ und $\sum_{i\in\mathbb{N}} f(\omega_i) = 1$. Dann ist $P := f\mu$ ein diskretes Wahrscheinlichkeitsmaß mit „Zähldichte“ f bezüglich des Zählmaßes.
- Sei in Definition 7.1 $\mu = \lambda$ das Lebesgue-Maß auf $(\mathbb{R}, \mathscr{B})$. Dann ist

  $$\phi\colon \mathbb{R} \to \mathbb{R}, \quad x \mapsto \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}x^2}$$

  die Normalverteilungsdichte.

Satz 7.5 (Eindeutigkeit der Dichte)**.** *Für* $f, g \in \mathscr{E}^*$ *gilt: Aus* $f = g$ $[\mu]$ *folgt* $f\mu = g\mu$, *das heißt* $\int_A f \,\mathrm{d}\mu = \int_A g \,\mathrm{d}\mu$ *für alle* $A \in \mathscr{A}$. *Sind außerdem* f *und* g μ*-integrierbar, so gilt auch die Umkehrung.*

Beweis. Die erste Aussage folgt direkt aus Satz 4.3. Seien nun f und g μ-integrierbar mit $f\mu = g\mu$. Definiere $N := \{f > g\}$ und $h := (f - g) \cdot \mathbb{1}_N$. Dann gilt:

$$\int h \,\mathrm{d}\mu = \int_N f - g \,\mathrm{d}\mu = 0.$$

Mit $h \geqslant 0$ folgt: $h = 0$ $[\mu]$ und somit $\mu(N) = 0$. Analog folgt $\mu(\{f < g\}) = 0$ und damit $f = g$ $[\mu]$.

Definition 7.6. In Satz 7.5 heißt f *Version* oder *Festlegung* der Dichte g und umgekehrt.

Problem 7.7. Seien μ und ν zwei Maße auf einem Messraum $(\Omega, \mathscr{A})$. Wir stellen uns die Frage, unter welchen Umständen ν eine μ-Dichte besitzt, das heißt, wann es $f \geqslant 0$ gibt, sodass

$$\nu(A) = \int_A f \, \mathrm{d}\mu \tag{7.1}$$

für alle $A \in \mathscr{A}$ gilt.

Lemma 7.8. *Eine notwendige Bedingung für die Existenz einer Dichte f, welche (7.1) erfüllt, ist, dass jede μ-Nullmenge auch eine ν-Nullmenge ist.*

Beweis. Sei $A \in \mathscr{A}$ mit $\mu(A) = 0$. Dann gilt:

$$\nu(A) = \int f \cdot \mathbb{1}_A \, \mathrm{d}\mu = 0.$$

Die in Lemma 7.8 beschriebene Eigenschaft wird sich als wesentlich für die Lösung von Problem 7.7 herausstellen.

Definition 7.9. Seien μ und ν zwei Maße auf einem Messraum $(\Omega, \mathscr{A})$. ν heißt *absolut stetig* (engl.: absolutely continuous) bezüglich μ, falls jede μ-Nullmenge auch eine ν-Nullmenge ist. Wir schreiben auch $\nu \ll \mu$ oder sagen: μ *dominiert* ν.

Das folgende Lemma motiviert den Begriff der absoluten Stetigkeit.

Lemma 7.10. *Seien μ und ν zwei Maße auf einem Messraum $(\Omega, \mathscr{A})$. Weiterhin sei ν endlich. Dann sind äquivalent:*

1. $\nu \ll \mu$.
2. *Für alle $\varepsilon > 0$ gibt es ein $\delta > 0$, sodass für alle $A \in \mathscr{A}$ mit $\mu(A) < \delta$ gilt: $\nu(A) < \varepsilon$.*

Beweis. Sei zunächst $\nu \ll \mu$. Angenommen, es gäbe ein $\varepsilon > 0$, sodass es für alle $\delta > 0$ ein $A_\delta \in \mathscr{A}$ gibt mit $\mu(A_\delta) < \delta$, aber $\nu(A_\delta) \geqslant \varepsilon$. Dann gäbe es insbesondere zu jedem $n \in \mathbb{N}$ ein $E_n \in \mathscr{A}$ mit $\mu(E_n) < 1/n^2$, aber $\nu(E_n) \geqslant \varepsilon$. Definiere

$$E := \bigcap_{k \in \mathbb{N}} \bigcup_{n \geqslant k} E_n \in \mathscr{A}.$$

Für alle $k \in \mathbb{N}$ gilt aufgrund der Monotonie von μ

$$\mu(E) = \mu\left(\bigcap_{k \in \mathbb{N}} \bigcup_{n \geqslant k} E_n\right) \leqslant \mu\left(\bigcup_{n \geqslant k} E_n\right).$$

Somit erhalten wir mit der σ-Subadditivität des Maßes μ

$$0 \leqslant \mu(E) \leqslant \lim_{k\to\infty} \mu\left(\bigcup_{n\geqslant k} E_n\right) \leqslant \lim_{k\to\infty} \sum_{n=k}^{\infty} \frac{1}{n^2} = 0.$$

Da ν endlich ist, folgt mit der Stetigkeit von oben

$$\nu(E) = \lim_{k\to\infty} \nu\left(\bigcup_{n\geqslant k} E_n\right) \geqslant \varepsilon.$$

Dies ist jedoch ein Widerspruch zu $\nu \ll \mu$. Folglich gilt 2.

Gelte nun 2. Sei $A \in \mathscr{A}$ eine μ-Nullmenge. Dann gilt für alle $\delta > 0$: $\mu(A) < \delta$. Somit folgt für alle $\varepsilon > 0$: $\nu(A) < \varepsilon$. Also ist $\nu(A) = 0$.

Lemma 7.11. *Seien μ, ν Maße auf dem Messraum $(\Omega, \mathscr{A})$ mit $\nu(A) \leqslant \mu(A)$ für alle $A \in \mathscr{A}$. Sei $p > 0$. Dann folgt aus $f \in \mathscr{L}^p(\Omega, \mathscr{A}, \mu)$, dass $f \in \mathscr{L}^p(\Omega, \mathscr{A}, \nu)$ und*

$$\int |f|^p \,\mathrm{d}\nu \leqslant \int |f|^p \,\mathrm{d}\mu.$$

Beweis. Sei $f \in \mathscr{L}^p(\Omega, \mathscr{A}, \mu)$. Dann gibt es eine Folge $(u_n)_{n\in\mathbb{N}}$ in $\mathscr{E}$ mit Normaldarstellungen

$$u_n = \sum_{i=1}^{m_n} \alpha_{i,n} \cdot \mathbb{1}_{A_{i,n}}$$

und $u_n \nearrow |f|^p$. Es gilt:

$$\int u_n \,\mathrm{d}\nu = \sum_{i=1}^{m_n} \alpha_{i,n} \cdot \nu(A_{i,n}) \leqslant \sum_{i=1}^{m_n} \alpha_{i,n} \cdot \mu(A_{i,n}) = \int u_n \,\mathrm{d}\mu.$$

Somit folgt:

$$\int |f|^p \,\mathrm{d}\nu = \lim_{n\to\infty} \int u_n \,\mathrm{d}\nu \leqslant \lim_{n\to\infty} \int u_n \,\mathrm{d}\mu = \int |f|^p \,\mathrm{d}\mu.$$

Lemma 7.12. *Seien μ und ν endliche Maße auf einem Messraum $(\Omega, \mathscr{A})$. Sei $\nu \ll \mu$. Dann gibt es eine messbare Funktion $g\colon \Omega \to [0,1)$ mit*

$$\int f \cdot (1-g) \,\mathrm{d}\nu = \int f \cdot g \,\mathrm{d}\mu$$

für alle $f \in \mathscr{L}^2(\Omega, \mathscr{A}, \mu + \nu)$.

Beweis. Sei $f \in \mathscr{L}^2(\Omega, \mathscr{A}, \mu + \nu)$. Da μ und ν endlich sind, ist $f \in \mathscr{L}^1(\Omega, \mathscr{A}, \mu + \nu)$. Mit Lemma 7.11 folgt $f \in \mathscr{L}^1(\Omega, \mathscr{A}, \nu)$. Definiere

$$H\colon \mathscr{L}^2(\Omega, \mathscr{A}, \mu + \nu) \to \mathbb{R}, \quad f \mapsto \int f \, \mathrm{d}\nu.$$

Diese Funktion ist linear. Mit der Hölder-Ungleichung für $p = q = 2$ folgt:

$$\begin{aligned}|H(f)| = \left|\int f \cdot \mathrm{id} \, \mathrm{d}\nu\right| &\leqslant \sqrt{\int |f|^2 \, \mathrm{d}\nu} \cdot \sqrt{\nu(\Omega)} \\ &\leqslant \sqrt{\int |f|^2 \, \mathrm{d}(\mu + \nu} \cdot \sqrt{\nu(\Omega)} = \|f\|_2 \cdot \sqrt{\nu(\Omega)}.\end{aligned}$$

Somit ist H ein beschränktes lineares Funktional auf $\mathscr{L}^2(\Omega, \mathscr{A}, \mu + \nu)$. Mit dem aus der Funktionalanalysis bekannten Satz von Riesz (siehe Satz A.2) folgt, dass es ein $h \in \mathscr{L}^2(\Omega, \mathscr{A}, \mu + nu)$ gibt, sodass für alle $f \in \mathscr{L}^2(\Omega, \mathscr{A}, \mu + \nu)$ gilt:

$$H(f) = \int f \cdot h \, \mathrm{d}(\mu + \nu). \tag{7.2}$$

Wir zeigen nun, dass $(\mu + \nu)$-fast-überall $h \geqslant 0$ gilt. Angenommen, es gäbe ein $A \in \mathscr{A}$ mit $h|_A < 0$, aber $(\mu + \nu)(A) > 0$. Dann wäre

$$0 < (\mu + \nu)(A) = \int \mathbb{1}_A \, \mathrm{d}(\mu + \nu) = H(\mathbb{1}_A) = \int h \cdot \mathbb{1}_A \, \mathrm{d}(\mu + \nu) < 0.$$

Das ist ein Widerspruch. Somit gilt $(\mu + \nu)$-fast-überall $h \geqslant 0$. Aus der Definition von H und (7.2) folgt:

$$\int f \cdot (1 - h) \, \mathrm{d}\nu = \int f \cdot h \, \mathrm{d}\mu \tag{7.3}$$

für alle $f \in \mathscr{L}^2(\Omega, \mathscr{A}, \mu + \nu)$. Durch Einsetzen von $f = \mathbb{1}_{\{h \geqslant 1\}}$ in (7.3) erhalten wir:

$$0 \leqslant \mu(\{h \geqslant 1\}) = \int \mathbb{1}_{\{h \geqslant 1\}} \, \mathrm{d}\mu \leqslant \int h \cdot \mathbb{1}_{\{h \geqslant 1\}} \, \mathrm{d}\mu = \int \mathbb{1}_{\{h \geqslant 1\}} \cdot (1 - h) \, \mathrm{d}\nu \leqslant 0.$$

Somit ist $\{h \geqslant 1\}$ eine μ-Nullmenge und wegen $\nu \ll \mu$ auch eine ν-Nullmenge. Definiere $g := h \cdot \mathbb{1}_{\{h < 1\}}$. Nach dem bisher Gezeigten gilt: $g(\Omega) \subseteq [0, 1)$, $g = h$ μ- und ν-fast-überall und somit nach (7.3):

$$\int f \cdot (1 - g) \, \mathrm{d}\nu = \int f \cdot g \, \mathrm{d}\mu$$

für alle $f \in \mathscr{L}^2(\Omega, \mathscr{A}, \mu + \nu)$.

Satz 7.13 (Radon-Nikodym für endliche Maße)**.** *Seien μ und ν zwei endliche Maße auf einem Messraum $(\Omega, \mathscr{A})$ mit $\nu \ll \mu$. Dann hat ν eine Dichte f bezüglich μ und wir schreiben:*

$$f = \frac{\mathrm{d}\nu}{\mathrm{d}\mu}.$$

Beweis. Sei $h \in \mathscr{E}^*$ beschränkt und $g\colon \Omega \to [0,1)$ wie in Lemma 7.12. Da g und h beschränkt sind und $(\mu + \nu)$ ein endliches Maß ist, gilt:

$$h \cdot \sum_{i=0}^{n-1} g^i \in \mathscr{L}^2(\Omega, \mathscr{A}, \mu + \nu)$$

für alle $n \in \mathbb{N}$. Mit Lemma 7.12 erhalten wir:

$$\int h \cdot (1-g) \cdot \sum_{i=0}^{n-1} g^i \,\mathrm{d}\nu = \int h \cdot g \cdot \sum_{i=0}^{n-1} g^i \,\mathrm{d}\mu.$$

Mit $0 \leqslant g(\omega) < 1$ für alle $\omega \in \Omega$ erhalten wir:

$$\int (1-g^n) \cdot h \,\mathrm{d}\nu = \int \frac{g}{1-g} \cdot (1-g^n) \cdot h \,\mathrm{d}\mu.$$

Die Funktionenfolge $\big((1-g^n) \cdot h\big)_{n\in\mathbb{N}}$ steigt gegen h auf. Nach dem Satz von der monotonen Konvergenz folgt:

$$\int h \,\mathrm{d}\nu = \int \frac{g}{1-g} \cdot h \,\mathrm{d}\mu.$$

Insbesondere folgt für $h \equiv 1$, dass $g/(1-g) \in \mathscr{E}^*$ ist. Setze $f := g/(1-g)$.

Ist h unbeschränkt, so definiere $h_n := \min\{h, n\}$ für alle $n \in \mathbb{N}$. Diese sind beschränkt und es gilt: $h_n \nearrow h$. Mit dem Satz von der monotonen Konvergenz folgt:

$$\int h \,\mathrm{d}\nu = \lim_{n\to\infty} \int h_n \,\mathrm{d}\nu = \lim_{n\to\infty} \int f \cdot h_n \,\mathrm{d}\mu = \int f \cdot h \,\mathrm{d}\mu.$$

Satz 7.14 (Radon-Nikodym für σ-endliche Maße)**.** *Seien μ und ν zwei σ-endliche Maße auf einem Messraum $(\Omega, \mathscr{A})$ mit $\nu \ll \mu$. Dann hat ν eine reellwertige Dichte f bezüglich μ und wir schreiben:*

$$f = \frac{\mathrm{d}\nu}{\mathrm{d}\mu}.$$

Außerdem ist f eindeutig in dem Sinne, dass für jede weitere Dichte $\tilde{f}$ von ν bezüglich μ gilt, dass $\tilde{f} = f$ $[\mu]$.

Beweis. Seien $(A_n)_{n\in\mathbb{N}}$ und $(B_n)_{n\in\mathbb{N}}$ Folgen paarweise disjunkter Mengen in $\mathscr{A}$, sodass gilt:

$$\bigcup_{n\in\mathbb{N}} A_n = \bigcup_{n\in\mathbb{N}} B_n = \Omega$$

sowie $\mu(A_n) < \infty$ und $\mu(B_n) < \infty$ für alle $n \in \mathbb{N}$. Definiere

$$\mathscr{C} := \{A_m \cap B_n \mid m, n \in \mathbb{N}\} \subseteq \mathscr{A}.$$

Wir können $\mathscr{C}$ zu einer Folge $(C_n)_{n\in\mathbb{N}}$ paarweise disjunkter Mengen anordnen. Für alle $n \in \mathbb{N}$ definiere

$$\mu_n\colon \mathscr{A} \to [0,\infty), \quad A \mapsto \mu(A \cap C_n),$$
$$\nu_n\colon \mathscr{A} \to [0,\infty), \quad A \mapsto \nu(A \cap C_n).$$

Somit sind μ_n und ν_n endliche Maße auf $(\Omega, \mathscr{A})$ mit $\nu_n \ll \mu_n$ für alle $n \in \mathbb{N}$. Mit Satz 7.13 erhalten wir Dichten $f_n = {}^{\mathrm{d}\nu_n}/_{\mathrm{d}\mu_n}$. Für alle messbaren Funktionen $g\colon \Omega \to [0,\infty]$ gilt:

$$\int g \,\mathrm{d}\nu_n = \int g \cdot f_n \,\mathrm{d}\mu_n. \tag{7.4}$$

Definiere

$$f\colon \Omega \to [0,\infty), \quad \omega \mapsto \sum_{n\in\mathbb{N}} f_n(\omega) \cdot \mathbb{1}_{C_n}(\omega).$$

f ist messbar und es gilt für alle messbaren Funktionen $g\colon \Omega \to [0,\infty]$:

$$\begin{aligned}\int g \,\mathrm{d}\nu = \sum_{n\in\mathbb{N}} \int g \cdot \mathbb{1}_{C_n} \,\mathrm{d}\nu = \sum_{n\in\mathbb{N}} \int g \,\mathrm{d}\nu_n = \sum_{n\in\mathbb{N}} \int g \cdot f_n \,\mathrm{d}\mu_n \\ = \sum_{n\in\mathbb{N}} \int g \cdot f \cdot \mathbb{1}_{C_n} \,\mathrm{d}\mu = \int g \cdot f \,\mathrm{d}\mu.\end{aligned}$$

Somit ist f eine Dichte von ν bezüglich μ.

Seien nun f und $\tilde{f}$ Dichten von ν bezüglich μ. Angenommen, es gäbe ein $A \in \mathscr{A}$ mit $\mu(A) > 0$ und $f(\omega) > \tilde{f}(\omega)$ für alle $\omega \in A$. Es gibt ein $n \in \mathbb{N}$ mit $\mu(A \cap C_n) > 0$. Sei $B := A \cap C_n$. Dann ist $\nu(B) < \infty$, $0 < \mu(B)$ und $(f - \tilde{f})|_B > 0$. Es folgt:

$$0 < \int_B f - \tilde{f} \,\mathrm{d}\mu = \nu(B) - \nu(B) = 0.$$

Dies ist ein Widerspruch und folglich ist $f \leqslant \tilde{f}$ μ-fast-überall. Analog folgt $f \geqslant \tilde{f}$ μ-fast-überall. Somit gilt die gewünschte Eindeutigkeit.

Dieser funktionalanalytische Beweis geht auf J. von Neumann zurück. Wir haben uns hier an Hewitt und Stromberg [9], Seite 320, orientiert. Einen rein maßtheoretischen Beweis findet man in Bauer, *Maß- und Integrationstheorie* [1].

7.2 ★Mehr zu Radon-Nikodym

Der Satz von Radon-Nikodym kann unter bestimmten Voraussetzungen an den Messraum $(\Omega, \mathscr{A})$ für nicht notwendigerweise σ-endliche Maße verallgemeinert werden.

Lemma 7.15. *Seien μ sowie ν Maße auf einem Messraum $(\Omega, \mathscr{A})$ mit $\mu(\Omega) < \infty$ und $\nu \ll \mu$. Dann gibt es ein $E \in \mathscr{A}$, sodass gilt:*

1. *Für alle $A \in \mathscr{A}$ mit $A \subseteq E$ ist $\nu(A) \in \{0, \infty\}$.*

2. *Für alle $A \in \mathscr{A}$ mit $A \subseteq E$ ist $\mu(A) = 0$, falls $\nu(A) = 0$.*

3. *ν ist σ-endlich auf E^c.*

Beweis. Definiere

$$\mathscr{C} := \{C \in \mathscr{A} \mid \text{Aus } A \in \mathscr{A} \text{ mit } A \subseteq C \text{ folgt } \nu(A) \in \{0, \infty\}.\}.$$

Offenbar ist $\varnothing \in \mathscr{C}$. Definiere

$$\alpha := \sup\{\mu(C) \mid C \in \mathscr{C}\}.$$

Da μ endlich ist, gilt: $\alpha < \infty$. Somit gibt es eine aufsteigende Folge $(C_n)_{n \in \mathbb{N}}$ in $\mathscr{C}$ mit $\lim_{n\to\infty} \mu(C_n) = \alpha$. Setze $D := \bigcup_{n\in\mathbb{N}} C_n$. Aus der σ-Additivität von μ folgt: $\mu(D) = \alpha$.

Wir zeigen zunächst $D \in \mathscr{C}$. Sei $A \in \mathscr{A}$ mit $A \subseteq D$, so gilt:

$$\nu(A) = \nu(A \cap C_1) + \sum_{n=2}^{\infty} \nu\big(A \cap (C_n \cap C_{n-1}^c)\big).$$

Es ist $\nu\big(A \cap (C_n \cap C_{n-1}^c)\big) \in \{0, \infty\}$. Somit folgt: $\nu(A) \in \{0, \infty\}$.

Wir zeigen nun, dass D^c die folgende Eigenschaft besitzt: Für alle $A \in \mathscr{A}$ mit $A \subseteq D^c$ und $\nu(A) > 0$ gibt es ein $B \in \mathscr{A}$ mit $B \subseteq A$ und

$$0 < \nu(B) < \infty. \tag{7.5}$$

Betrachte ein solches $A \in \mathscr{A}$. Falls $\nu(A) < \infty$, so ist (7.5) trivial erfüllt. Sei also $\nu(A) = \infty$. Angenommen, für jede Teilmenge $B \subseteq A$ mit $B \in \mathscr{A}$ wäre $\nu(B) \in \{0, \infty\}$. Dann wäre $A \cup D \in \mathscr{C}$. Aus $\nu \ll \mu$ und $\nu(A) > 0$ folgt: $\mu(A) > 0$. Allerdings gilt wegen $A \subseteq D^c$: $\mu(A \cup D) = \mu(A) + \mu(D) > \alpha$. Dies ist ein Widerspruch. Somit gibt es eine Menge $B \in \mathscr{A}$ mit $B \subseteq A$, welche (7.5) erfüllt.

Als nächstes zeigen wir, dass ν auf D^c σ-endlich ist. Definiere

$$\mathscr{F} := \{A \in \mathscr{A} \mid A \subseteq D^c \text{ und } \nu \text{ ist } \sigma\text{-endlich auf } A\}.$$

Offenbar ist $\varnothing \in \mathscr{F}$. Definiere

$$\beta := \sup\{\mu(A) \mid A \in \mathscr{F}\}.$$

Da μ endlich ist, gilt: $\beta < \infty$. Somit gibt es eine aufsteigende Folge $(H_n)_{n\in\mathbb{N}}$ in $\mathscr{F}$ mit $\lim_{n\to\infty} \mu(H_n) = \beta$. Setze $H := \bigcup_{n\in\mathbb{N}} H_n$. Aus der σ-Additivität von μ folgt: $\mu(H) = \beta$. H ist eine abzählbare Vereinigung von Mengen, auf welchen ν σ-endlich ist. Somit ist auch ν auf H σ-endlich. Also gilt: $H \in \mathscr{F}$. Angenommen, es wäre $\nu(H^c \cap D^c) > 0$. Wegen der oben gezeigten Eigenschaft von D^c erhielten wir ein $A \in \mathscr{A}$ mit $A \subseteq H^c \cap D^c$ und $\nu(A) \in (0, \infty)$. Somit wäre $H \cup A \in \mathscr{F}$. Es gilt aber:

$$\mu(H \cup A) = \mu(H) + \mu(A) > \mu(H) = \beta.$$

Dies ist ein Widerspruch zur Definition von β. Somit ist $\nu(H^c \cap D^c) = 0$ und ν σ-endlich auf D^c.

Um die gesuchte Menge E zu definieren, betrachte

$$\mathscr{G} := \{A \in \mathscr{A} \mid A \subseteq D \text{ und } \nu(A) = 0\}.$$

Offenbar ist $\varnothing \in \mathscr{G}$. Definiere

$$\gamma := \sup\{\mu(A) \mid A \in \mathscr{G}\}.$$

Da μ endlich ist, gilt: $\gamma < \infty$. Somit gibt es eine aufsteigende Folge $(G_n)_{n\in\mathbb{N}}$ in $\mathscr{G}$ mit $\lim_{n\to\infty} \mu(G_n) = \gamma$. Setze $G := \bigcup_{n\in\mathbb{N}} G_n$. Aus der σ-Additivität von μ folgt: $\mu(G) = \gamma$. Definiere

$$E := D \cap G^c.$$

Wegen $\nu(G) = 0$ ist ν σ-endlich auf $D^c \cup G = E^c$. Somit erfüllt E 3. Die Gültigkeit von 1. folgt aus $E \subseteq D$. Angenommen, 2. wäre nicht erfüllt. Dann gäbe es ein $A \subseteq E$ mit $A \in \mathscr{A}$, $\nu(A) = 0$ und $\mu(A) > 0$. Somit wäre $G \cup A \in \mathscr{G}$. Dies stünde jedoch im Widerspruch zu

$$\mu(G \cup A) = \mu(G) + \mu(A) > \mu(G) = \gamma.$$

Somit ist 2. erfüllt und E die gesuchte Menge.

Definition 7.16. Sei $(\Omega, \mathscr{A}, \mu)$ ein Maßraum. Es gebe ein $\mathscr{Z} \subseteq \mathscr{A}$ mit

1. $\mu(Z) \in [0, \infty)$ für alle $Z \in \mathscr{Z}$.

2. Die Mengen in $\mathscr{Z}$ sind paarweise disjunkt und es gilt:

$$\bigcup_{Z\in\mathscr{Z}} Z = \Omega.$$

3. Sei $A \in \mathscr{A}$ mit $\mu(A) < \infty$. Dann gilt:

$$\mu(A) = \sum_{Z\in\mathscr{Z}} \mu(A \cap Z).$$

   Falls $\mathscr{Z}$ überabzählbar ist, definieren wir

$$\sum_{Z\in\mathscr{Z}} \mu(A \cap Z) := \sup\left\{ \sum_{Z\in\mathscr{Z}_0} \mu(A \cap Z) \;\middle|\; \mathscr{Z}_0 \subseteq \mathscr{Z} \text{ ist endlich} \right\}.$$

4. Sei $B \subseteq \Omega$ mit $B \cap Z \in \mathscr{A}$ für alle $Z \in \mathscr{Z}$. Dann ist $B \in \mathscr{A}$.

In diesem Fall heißen $(\Omega, \mathscr{A}, \mu)$ und μ *zerlegbar* (engl.: decomposable). $\mathscr{Z}$ heißt *Zerlegung* (engl.: decomposition) von $(\Omega, \mathscr{A}, \mu)$.

Satz 7.17 (Radon-Nikodym). *Sei $(\Omega, \mathscr{A}, \mu)$ ein zerlegbarer Maßraum und $\mathscr{Z}$ eine Zerlegung. Sei ν ein Maß auf $(\Omega, \mathscr{A})$ mit $\nu \ll \mu$. Dann existiert eine messbare Funktion $f\colon \Omega \to [0, \infty]$ (welche auf den $Z \in \mathscr{Z}$, auf denen ν σ-endlich ist, reellwertig gewählt werden kann) mit den folgenden Eigenschaften:*

1. *Für alle $A \in \mathscr{A}$, auf denen μ σ-endlich ist, gilt:*
$$\nu(A) = \int_A f \,\mathrm{d}\mu.$$

2. *Für alle messbaren Funktionen $g\colon \Omega \to [0, \infty]$, sodass μ auf $\{\omega \in \Omega \mid g(\omega) > 0\}$ σ-endlich ist, gilt:*
$$\int g \,\mathrm{d}\nu = \int g \cdot f \,\mathrm{d}\mu.$$

3. *Für alle (nicht zwangsläufig positiven) $g \in \mathscr{L}^1(\Omega, \mathscr{A}, \nu)$, für welche μ auf der Menge $\{\omega \in \Omega \mid g(\omega) \neq 0\}$ σ-endlich ist, gilt: $g \cdot f \in \mathscr{L}^1(\Omega, \mathscr{A}, \mu)$ sowie*
$$\int g \,\mathrm{d}\nu = \int g \cdot f \,\mathrm{d}\mu.$$

4. *f ist eindeutig in dem Sinne, dass für jede messbare Funktion $\tilde{f}\colon \Omega \to [0, \infty]$ mit*
$$\nu(A) = \int_A \tilde{f} \,\mathrm{d}\mu$$
für alle $A \in \mathscr{A}$ mit $\mu(A) < \infty$ gilt: $f \cdot \mathbb{1}_E = \tilde{f} \cdot \mathbb{1}_E$ $[\mu]$ für alle $E \in \mathscr{A}$, auf welchen μ σ-endlich ist.

Beweis. Für jedes $Z \in \mathscr{Z}$ ist die Einschränkung $\nu|_Z$ von ν auf Z absolut stetig bezüglich der Einschränkung $\mu|_Z$ von μ auf Z. Wegen $\mu(Z) < \infty$ können wir Lemma 7.15 anwenden und erhalten Mengen $D_Z, E_Z \in \mathscr{A}$ mit
$$D_Z \cap E_Z = \varnothing, \quad D_Z \cup E_Z = Z,$$
sodass E_Z die Eigenschaften 1. und 2. aus Lemma 7.15 hat und ν σ-endlich auf D_Z ist. Ist ν bereits σ-endlich auf Z, so wählen wir $D_Z = Z$. Da ν σ-endlich auf D_Z ist und μ endlich auf D_Z ist, können wir Satz 7.14 anwenden und erhalten eine Dichte $f_Z\colon Z \to [0, \infty)$ von $\nu|_Z$ bezüglich $\mu|_Z$. Definiere
$$f\colon \Omega \to [0, \infty], \quad \omega \mapsto \begin{cases} f_Z(\omega) & \text{falls } \omega \in D_Z \\ \infty & \text{falls } \omega \in E_Z \end{cases}. \tag{7.6}$$

Somit ist f reellwertig auf $Z \in \mathscr{Z}$, falls $Z = D_Z$, falls also ν auf Z σ-endlich ist. Offenbar ist f messbar. Definiere
$$D := \bigcup_{Z \in \mathscr{Z}} D_Z, \quad E := \bigcup_{Z \in \mathscr{Z}} E_Z.$$

Wir zeigen nun, dass f die gewünschten Eigenschaften hat. Sei $A \in \mathscr{A}$ mit $\mu(A) < \infty$. Da Z eine Zerlegung ist, gilt:

$$\mu(A) = \sum_{Z \in \mathscr{Z}_0} \mu(A \cap Z), \tag{7.7}$$

wobei $\mathscr{Z}_0$ eine höchstens abzählbare Teilmenge von $\mathscr{Z}$ ist. Dies liefert:

$$\mu\left(A \cap \bigcup_{Z \in \mathscr{Z} \cap \mathscr{Z}_0^c} Z\right) = 0.$$

Aus $\nu \ll \mu$ folgt:

$$\nu\left(A \cap \bigcup_{Z \in \mathscr{Z} \cap \mathscr{Z}_0^c} Z\right) = 0.$$

und somit

$$\nu(A) = \sum_{Z \in \mathscr{Z}_0} \nu(A \cap Z). \tag{7.8}$$

Aus (7.6) folgt für alle $Z \in \mathscr{Z}$:

$$\nu(A \cap Z) = \nu(A \cap D_Z) + \nu(A \cap E_Z) = \int_{A \cap D_Z} f \,\mathrm{d}\mu + \nu(A \cap E_Z). \tag{7.9}$$

Nach Lemma 7.15 ist $\nu(A \cap E_Z) \in \{0, \infty\}$ und $\nu(A \cap E_Z) = 0$ genau dann, wenn $\mu(A \cap E_Z) = 0$ ist. Somit gilt:

$$\nu(A \cap E_Z) = \int_{A \cap E_Z} \infty \,\mathrm{d}\mu = \int_{A \cap E_Z} f \mathrm{d}\mu. \tag{7.10}$$

Zusammen mit (7.8) und (7.9) ergibt sich:

$$\nu(A) = \sum_{Z \in \mathscr{Z}_0} \nu(A \cap Z) = \sum_{Z \in \mathscr{Z}_0} \int_{A \cap Z} f \,\mathrm{d}\mu = \int_A f \,\mathrm{d}\mu. \tag{7.11}$$

Sei nun $A \in \mathscr{A}$ eine Menge, auf welcher μ σ-endlich ist. Sei $(A_n)_{n \in \mathbb{N}}$ eine Folge in $\mathscr{A}$ mit $A_n \nearrow A$ und $\mu(A_n) < \infty$ für alle $n \in \mathbb{N}$. Somit gilt nach (7.11) und dem Satz von der monotonen Konvergenz:

$$\nu(A) = \lim_{n \to \infty} \nu(A_n) = \lim_{n \to \infty} \int_{A_n} f \,\mathrm{d}\mu = \int_A f \,\mathrm{d}\mu.$$

Also gilt 1.

Wir erhalten 2. mittels algebraischer Induktion analog zum Beweis von Satz 7.3. Durch Zerlegen von f in Positiv- und Negativteil folgt 3.

Es bleibt die Eindeutigkeit zu zeigen. Betrachte $\tilde{f}$ wie in 4. Für alle $Z \in \mathscr{Z}$ und jedes $A \in \mathscr{A}$ mit $A \subseteq D_Z$ gilt:

$$\nu(A) = \int_A f \,\mathrm{d}\mu = \int_A \tilde{f} \,\mathrm{d}\mu.$$

Nach Satz 7.14 folgt: $f(\omega) = \tilde{f}(\omega)$ für μ-fast-alle $\omega \in D_Z$. Angenommen, es gäbe ein $A \in \mathscr{A}$ mit $A \subseteq E_Z$, $\mu(A) > 0$ und

$$\tilde{f}(\omega) < \infty = f(\omega)$$

für alle $\omega \in A$. Nach Wahl der E_Z würde $\nu(A) > 0$ folgen. Für alle $n \in \mathbb{N}$ definiere

$$A_n := \left\{\omega \in A \,\middle|\, \tilde{f}(\omega) < n\right\}.$$

Somit gilt: $A_n \nearrow A$. Nach der Stetigkeit des Maßes ν von unten folgt:

$$0 < \nu(A) = \lim_{n\to\infty} \nu(A_n).$$

Folglich gibt es ein $n \in \mathbb{N}$ mit $\nu(A_n) > 0$. Aus $A_n \subseteq E_Z$ folgt: $\nu(A_n) = \infty$. Einsetzen in 4. liefert:

$$\infty = \nu(A_n) = \int_{A_n} \tilde{f} \,\mathrm{d}\mu \leqslant n \cdot \mu(A_n) < \infty.$$

Dies ist ein Widerspruch und es folgt: $\tilde{f} = f$ μ-fast überall auf allen $Z \in \mathscr{Z}$.

Korollar 7.18 (Radon-Nikodym)**.** *Seien μ und ν zwei Maße auf einem Messraum $(\Omega, \mathscr{A})$, wobei μ σ-endlich sei. ν hat genau dann eine Dichte f bezüglich μ, wenn $\nu \ll \mu$. In diesem Fall ist f μ-fast-überall eindeutig und wir schreiben:*

$$f = \frac{\mathrm{d}\nu}{\mathrm{d}\mu}.$$

Außerdem ist ν genau dann σ-endlich, wenn f μ-fast-überall endlich ist.

Beweis. Sei zunächst $\nu \ll \mu$. Da μ σ-endlich ist, gibt es eine Folge $(A_n)_{n\in\mathbb{N}}$ paarweise disjunkter Mengen in $\mathscr{A}$ mit $\mu(A_n) < \infty$ für alle $n \in \mathbb{N}$ und $\Omega = \bigcup_{n\in\mathbb{N}} A_n$. Dies ist eine Zerlegung von $(\Omega, \mathscr{A}, \mu)$. Somit können wir Satz 7.17 anwenden. Beachte, dass die Einschränkungen in Satz 7.17, 1. bis (3., nicht zum Tragen kommen, da μ bereits auf ganz Ω σ-endlich ist. Folglich ist die durch den Satz gegebene Funktion f eine Dichte von ν bezüglich μ. Die Eindeutigkeit folgt ebenfalls aus dem Satz.

Habe nun ν eine Dichte f bezüglich μ. Sei $N \in \mathscr{A}$ eine μ-Nullmenge. Es gilt:

$$\nu(N) = \int_N f \,\mathrm{d}\mu = 0.$$

Somit ist $\nu \ll \mu$. Der Rest ist eine Übung. [Übung]

7.3 Rechnen mit Dichten

Lemma 7.19. *Seien $(\Omega, \mathscr{A}, \mu)$ ein Maßraum, $\phi, f \in \mathscr{E}^*$ und $\nu := f\mu$. ϕ ist genau dann ν-integrierbar, wenn $\phi \cdot f$ μ-integrierbar ist. In diesem Fall gilt für alle $A \in \mathscr{A}$:*

$$\int_A \phi \,\mathrm{d}\nu = \int_A \phi \cdot f \,\mathrm{d}\mu.$$

[Übung]

Regeln 7.20. *Seien μ, ν und τ σ-endliche Maße auf einem Messraum $(\Omega, \mathscr{A})$. Dann gilt:*

1. *Seien $\mu \ll \nu$, $A \in \mathscr{A}$ und f integrierbar. Dann gilt:*

$$\int_A f \,\mathrm{d}\mu = \int_A f \frac{\mathrm{d}\mu}{\mathrm{d}\nu} \,\mathrm{d}\nu. \quad \textit{(Erste Kettenregel)}$$

Beweis. Nach dem Satz von Radon-Nikodym gibt es eine Dichte ${}^{\mathrm{d}\mu}/_{\mathrm{d}\nu}$ mit $\mu = {}^{\mathrm{d}\mu}/_{\mathrm{d}\nu}\nu$. Nach Lemma 7.19 gilt für alle $A \in \mathscr{A}$:

$$\int_A f \,\mathrm{d}\mu = \int_A f^+ \,\mathrm{d}\mu - \int_A f^- \,\mathrm{d}\mu = \int_A f^+ \frac{\mathrm{d}\mu}{\mathrm{d}\nu} \,\mathrm{d}\nu - \int_A f^- \frac{\mathrm{d}\mu}{\mathrm{d}\nu} \,\mathrm{d}\nu = \int_A f \frac{\mathrm{d}\mu}{\mathrm{d}\nu} \,\mathrm{d}\nu.$$

2. *Falls $\tau \ll \nu \ll \mu$, so gilt:*

$$\frac{\mathrm{d}\tau}{\mathrm{d}\mu} = \frac{\mathrm{d}\tau}{\mathrm{d}\nu} \cdot \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \quad [\mu]. \quad \textit{(Zweite Kettenregel)}$$

3. *Falls $\nu \ll \mu$ und $\tau \ll \mu$, so gilt:*

$$\frac{\mathrm{d}(\nu + \tau)}{\mathrm{d}\mu} = \frac{\mathrm{d}\nu}{\mathrm{d}\mu} + \frac{\mathrm{d}\tau}{\mathrm{d}\mu} \quad [\mu].$$

[Übung]

Beispiel 7.21. Sei X eine Zufallsvariable mit Verteilung P und $g\colon \mathbb{R} \to \mathbb{R}$ eine messbare Funktion, sodass $\mathrm{E}(g \circ X)$ existiert. Dann folgt aus Regel 7.20.1 für eine μ-Dichte p von P:

$$\mathrm{E}(g \circ X) = \int g(x)p(x)\,\mu(\mathrm{d}x).$$

Bemerkung 7.22. In Regel 7.20.2 heißt ${}^{\mathrm{d}\tau}/_{\mathrm{d}\mu}$ der *Dichtequotient* von τ und μ. Für zwei Maße μ und ν kann stets ein Dichtequotient bezüglich des dominierenden Maßes $\mu + \nu$ angegeben werden.

Der folgende Satz gibt eine überraschend einfache Möglichkeit an, den sogenannten Totalvariationsabstand zwischen zwei Wahrscheinlichkeitsmaßen P und Q mittels Dichten zu berechnen. Beachte, dass $P+Q$ stets ein dominierendes Maß für P und Q ist.

Satz 7.23 (Scheffés Lemma). *Sei (Ω, A, μ) ein Maßraum. Seien $P = f\mu$, $Q = g\mu$ und $P_n = f_n\mu$ für alle $n \in \mathbb{N}$ Wahrscheinlichkeitsmaße auf $\mathscr{A}$ mit μ-Dichten f, g beziehungsweise f_n. Dann gilt:*

1.
$$\sup_{A\in\mathscr{A}} |P(A) - Q(A)| = \frac{1}{2}\int |f-g| \,\mathrm{d}\mu.$$

2. Konvergiert $(f_n)_{n\in\mathbb{N}}$ μ-fast-überall gegen f, so folgt:
$$\lim_{n\to\infty}\int |f_n - f| \,\mathrm{d}\mu = 0.$$

Beweis.

1. Da P und Q Wahrscheinlichkeitsmaße sind, gilt:
$$0 = \int f - g \,\mathrm{d}\mu = \int (f-g)^+ \,\mathrm{d}\mu - \int (f-g)^- \,\mathrm{d}\mu.$$
Folglich gilt:
$$\int (f-g)^+ \,\mathrm{d}\mu = \int (f-g)^- \,\mathrm{d}\mu = \frac{1}{2}\int |f-g| \,\mathrm{d}\mu.$$
Für alle $A \in \mathscr{A}$ gilt:
$$\begin{aligned} |P(A) - Q(A)| &= \left|\int (f-g)^+ \cdot \mathbb{1}_A \,\mathrm{d}\mu - \int (f-g)^- \cdot \mathbb{1}_A \,\mathrm{d}\mu\right| \\ &\leqslant \int (f-g)^+ \,\mathrm{d}\mu = \frac{1}{2}\int f - g \cdot \mathbb{1}_A \,\mathrm{d}\mu. \end{aligned}$$
Hierbei herrscht Gleichheit für $A := \{f - g > 0\} \in \mathscr{A}$. Somit folgt die Behauptung.

2. Für alle $n \in \mathbb{N}$ gilt:
$$0 \leqslant (f - f_n)^+ \leqslant f^+ = f \quad \text{sowie} \quad (f-f_n)^+ \xrightarrow{n\to\infty} 0 \quad [\mu].$$
Mit dem Satz von der dominierten Konvergenz folgt:
$$\int |f - f_n| \,\mathrm{d}\mu = 2\int (f-f_n)^+ \,\mathrm{d}\mu \xrightarrow{n\to\infty} 0.$$

7.4 Signierte Maße und Lebesgue-Zerlegung

Definition 7.24. Zwei Maße μ und ν auf einem Messraum $(\Omega, \mathscr{A})$ heißen *singulär* (engl.: singular), wenn es ein $A \in \mathscr{A}$ gibt, sodass gilt:

$$\mu(A) = \nu(A^c) = 0.$$

Wir schreiben auch $\mu \perp \nu$.

Satz 7.25 (Lebesgue-Zerlegung)**.** *Seien μ und ν zwei Maße auf einem Messraum $(\Omega, \mathscr{A})$. Ist ν σ-endlich, dann gibt es eindeutig bestimmte Maße ν_a und ν_s mit folgenden Eigenschaften:*

1. $\nu_a \ll \mu$,
2. $\nu_s \perp \mu$,
3. $\nu = \nu_a + \nu_s$.

Beweis. Wir zeigen zunächst die Existenz der Zerlegung. Definiere

$$N_\mu := \{A \in \mathscr{A} \mid \mu(A) = 0\}, \quad \alpha := \sup\{\nu(A) \mid A \in N_\mu\}.$$

Nach der Definition des Supremums können wir eine Folge $(A_n)_{n\in\mathbb{N}}$ in N_μ finden, sodass $\lim_{n\to\infty} \nu(A_n) = \alpha$. Definiere die Folge $(N_n)_{n\in\mathbb{N}}$ in N_μ mit

$$N_n := \bigcup_{i=1}^{n} A_i \quad \text{sowie} \quad N := \bigcup_{i\in\mathbb{N}} A_i \in N_\mu.$$

Nach Definition gilt: $N_n \nearrow N$ und aus der Stetigkeit von unten folgt: $\nu(N) = \alpha$. Für alle $A \in \mathscr{A}$ setze

$$\nu_a(A) := \nu(A \cap N^c), \quad \nu_s(A) := \nu(A \cap N).$$

Dann ist $\nu = \nu_a + \nu_s$. Außerdem gilt: $\nu_s \perp \mu$, da $\nu_s(N^c) = \nu(N^c \cap N) = 0$ und $\mu(N) = 0$. Es bleibt zu zeigen, dass $\nu_a \ll \mu$.

Sei ν zunächst endlich. Sei $A \in N_\mu$. Es ist $N + (A \cap N^c) \in N_\mu$ und somit gilt:

$$\alpha \geqslant \nu\big(N + (A \cap N^c)\big) = \nu(N) + \nu(A \cap N^c) = \alpha + \nu_a(A).$$

Da ν endlich ist, gilt: $\alpha \in [0, \infty)$ und folglich: $\nu_a(A) = 0$.

Sei nun ν σ-endlich. Somit gibt es eine Folge $(B_n)_{n\in\mathbb{N}}$ in $\mathscr{A}$ mit $B_n \nearrow \Omega$ und $\nu(B_n) < \infty$ für alle $n \in \mathbb{N}$. Betrachte die Folgen $(\nu^n)_{n\in\mathbb{N}}$ und $(\nu_a^n)_{n\in\mathbb{N}}$ endlicher Maße

$$\nu^n \colon \mathscr{A} \to [0, \infty), \quad A \mapsto \nu(A \cap B_n),$$
$$\nu_a^n \colon \mathscr{A} \to [0, \infty), \quad A \mapsto \nu_a(A \cap B_n).$$

Dann gilt nach dem oben Gezeigten, dass $\nu_a^n \ll \mu$ für alle $n \in \mathbb{N}$. Mit $\nu = \sup_{n\in\mathbb{N}} \nu^n$ und $\nu_a = \sup_{n\in\mathbb{N}} \nu_a^n$ folgt, dass für alle $A \in N_\mu$ gilt: $\nu_a^n(A) = 0$ für alle $n \in \mathbb{N}$ und somit $\nu_a(A) = 0$.

Wir zeigen nun die Eindeutigkeit der Zerlegung. Sei $\nu = \nu_a' + \nu_s'$ eine weitere Lebesgue-Zerlegung von ν. Aufgrund der Singularität von ν_s und ν_s' bezüglich μ gibt es μ-Nullmengen N und N' mit

$$\nu_s(A) = \nu_s(A \cap N) \quad \text{und} \quad \nu_s'(A) = \nu_s'(A \cap N') \tag{7.12}$$

für alle $A \in \mathscr{A}$. Betrachte $N_0 := N \cup N'$. Dies ist ebenfalls eine μ-Nullmenge und aus $\nu_a \ll \mu$ sowie $\nu_a' \ll \mu$ folgt für alle $A \in \mathscr{A}$:

$$\nu_a(A \cap N_0) = \nu_a'(A \cap N_0) = 0.$$

Mit (7.12) folgt:

$$\begin{aligned}\nu_s(A) = \nu_s(A \cap N) = \nu_s(A \cap N_0 \cap N) = \nu_s(A \cap N_0) = \nu(A \cap N_0)\\ = \nu_s'(A \cap N_0) = \nu_s'(A \cap N_0 \cap N') = \nu_s'(A \cap N') = \nu_s'(A).\end{aligned}$$

Somit folgt: $\nu_s = \nu_s'$ und aus $\nu = \nu_a + \nu_s = \nu_a' + \nu_s'$ folgt $\nu_a = \nu_a'$. Also ist die Zerlegung eindeutig.

Bemerkung 7.26. In Satz 7.25 heißen ν_a der *absolut stetige* oder *reguläre Anteil* von ν bezüglich μ und ν_s der *singuläre Anteil* von ν bezüglich μ.

Definition 7.27. Sei $(\Omega, \mathscr{A})$ ein Messraum. Eine Funktion $\phi\colon \mathscr{A} \to (-\infty, \infty]$ oder $\phi\colon \mathscr{A} \to [-\infty, \infty)$ mit $\phi(\varnothing) = 0$ und

$$\phi\left(\sum_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \phi(A_n)$$

für alle Folgen $(A_n)_{n\in\mathbb{N}}$ disjunkter Mengen in $\mathscr{A}$ heißt *signiertes Maß* (engl.: signed measure).

Beispiel 7.28. Seien $(\Omega, \mathscr{A}, \mu)$ ein Maßraum und f messbar mit $f^- \in L^1(\mu)$. Dann ist

$$\phi\colon \mathscr{A} \to (-\infty, \infty], \quad A \mapsto \int_A f \,\mathrm{d}\mu$$

ein signiertes Maß. [Übung]

Lemma 7.29. *Für jedes signierte Maß gilt die Stetigkeit von unten und die Stetigkeit von oben.*

Beweis. Der Beweis läuft völlig analog zu dem für Maße. Siehe die Regeln 1.27.

Satz 7.30 (Jordan-Hahn-Zerlegung). *Sei ϕ ein signiertes Maß auf einem Messraum $(\Omega, \mathscr{A})$. Dann gibt es eine bis auf Nullmengen eindeutige Zerlegung $\Omega = \Omega^+ \cup \Omega^-$, sodass gilt:*

1. *$\phi(A) \geqslant 0$ für alle $A \in \mathscr{A}$, $A \subseteq \Omega^+$,*
2. *$\phi(A) \leqslant 0$ für alle $A \in \mathscr{A}$, $A \subseteq \Omega^-$.*

Insbesondere sind

$$\phi^+ : \mathscr{A} \to [0, \infty], \quad A \mapsto \phi(A \cap \Omega^+)$$

und

$$\phi^- : \mathscr{A} \to [0, \infty], \quad A \mapsto -\phi(A \cap \Omega^-)$$

Maße auf $(\Omega, \mathscr{A})$. Darüber hinaus ist eines der beiden Maße endlich und es gilt: $\phi = \phi^+ - \phi^-$. Desweiteren nennen wir $|\phi| := \phi^+ + \phi^-$ das Total-Variationsmaß *auf $(\Omega, \mathscr{A})$.*

Beweis. Wir zeigen zunächst die Existenz der Zerlegung. Ohne Einschränkung sei $\phi(A) < \infty$ für alle $A \in \mathscr{A}$. Andernfalls betrachte das signierte Maß $-\phi$. Definiere

$$B^+ := \{A \in \mathscr{A} \mid \phi(A) \geqslant 0\}, \quad B^- := \{A \in \mathscr{A} \mid \phi(A) \leqslant 0\}$$

sowie

$$\alpha := \sup_{A \in B^+} \phi(A) \geqslant 0.$$

Somit finden wir eine Folge $(A_k)_{k \in \mathbb{N}}$ in B^+ mit $\lim_{k \to \infty} \phi(A_k) = \alpha$. Für alle $n \in \mathbb{N}$ definiere

$$P_n := \bigcup_{k=1}^{n} A_k \in \mathscr{A} \quad \text{sowie} \quad P := \bigcup_{n \in \mathbb{N}} P_n \in \mathscr{A}.$$

Außerdem ist $\phi(P_n) \geqslant \phi(A_n)$ für alle $n \in \mathbb{N}$. Aus der Stetigkeit des signierten Maßes von unten folgt:

$$\infty > \alpha = \phi(P) = \lim_{n \to \infty} \phi(P_n) \geqslant 0.$$

Angenommen, es gäbe ein $A \in \mathscr{A}$, sodass $\phi(P \cap A) < 0$ ist. Dann würde gelten:

$$\phi(P \backslash A) = \phi(P) - \phi(P \cap A) > \alpha.$$

Dies ist ein Widerspruch zur Definition von α. Folglich ist $P \cap A \in B^+$ für alle $A \in \mathscr{A}$. Analog nehmen wir an, es gäbe ein $A \in \mathscr{A}$, sodass $\phi(P^c \cap A) > 0$. Dann würde gelten:

$$\phi\big((P^c \cap A) \cup P\big) = \phi(P^c \cap A) + \phi(P) = \phi(P^c \cap A) + \alpha > \alpha.$$

Dies widerspricht jedoch der Definition von α und es ist $P^c \cap A \in B^-$ für alle $A \in \mathscr{A}$.

Setze $\Omega^+ := P$ und $\Omega^- := P^c$. Dann ist $\Omega = \Omega^+ \cup \Omega^-$ die gewünschte Zerlegung von Ω. Mit $\mathscr{A}^+ := \mathscr{A}|_{\Omega^+}$ und $\mathscr{A}^- := \mathscr{A}|_{\Omega^-}$ ist außerdem $\phi^+ := \phi|_{\mathscr{A}^+}$ ein endliches Maß auf $(\Omega^+, \mathscr{A}^+)$ und $\phi^- := -\phi|_{\mathscr{A}^-}$ ein Maß auf $(\Omega^-, \mathscr{A}^-)$.

Um die Eindeutigkeit zu zeigen, betrachten wir zwei Zerlegungen $\Omega = \Omega_1^+ \cup \Omega_1^-$ und $\Omega = \Omega_2^+ \cup \Omega_2^-$. Sei $A \in \mathscr{A}$. Da $A \cap \Omega_1^+ \cap \Omega_2^-$ eine Teilmenge von Ω_1^+ und Ω_2^- ist, gilt: $\phi(A \cap \Omega_1^+ \cap \Omega_2^-) = 0$. Analog folgt: $\phi(A \cap \Omega_1^- \cap \Omega_2^+) = 0$. Somit gilt:

$$\phi(A \cap \Omega_1^+) = \phi(A \cap \Omega_1^+ \cap \Omega_2^+) = \phi(A \cap \Omega_2^+)$$

sowie

$$\phi(A \cap \Omega_1^-) = \phi(A \cap \Omega_1^- \cap \Omega_2^-) = \phi(A \cap \Omega_2^-).$$

Damit ist die Zerlegung bis auf Nullmengen eindeutig.

Bemerkung 7.31.

1. Seien ϕ ein σ-endliches signiertes Maß und f eine ϕ^+- und ϕ^--integrierbare Funktion. Definiere:

$$\int f \,\mathrm{d}\phi := \int f \,\mathrm{d}\phi^+ - \int f \,\mathrm{d}\phi^-.$$

Es gilt:

$$\left|\int f \,\mathrm{d}\phi\right| \leqslant \int |f| \,\mathrm{d}\,|\phi|. \qquad \text{[Übung]}$$

2. Sei $\phi = \phi^+ - \phi^-$ ein σ-endliches signiertes Maß auf $(\mathbb{R}, \mathscr{B})$ mit zugehöriger maßdefinierender Funktion $F = F^+ - F^-$. Sei g eine ϕ^+- und ϕ^--integrierbare Funktion. Das *Lebesgue-Stieltjes-Integral* ist definiert durch

$$\int g \,\mathrm{d}F = \int g \,\mathrm{d}F^+ - \int g \,\mathrm{d}F^-.$$

Ist $g\colon [a, b] \to \mathbb{R}$ mit $a, b \in \mathbb{R}$ stetig, so heißt es auch das *Riemann-Stieltjes-Integral*.

7.5 Funktionen von beschränkter Variation

Monotone Funktionen haben gute Eigenschaften, welche sich zum Teil auf Differenzen monotoner Funktionen verallgemeinern lassen. Oftmals ist es jedoch schwierig, einer Funktion anzusehen, ob sie sich als solche Differenz schreiben lässt. Eine Möglichkeit hierfür liefert die Untersuchung der sogenannten Totalvariation, wie wir in Satz 7.35 sehen werden.

Definition 7.32. Seien $a, b \in \mathbb{R}$ mit $a \leqslant b$ und $F\colon [a, b] \to \mathbb{R}$ eine Funktion.

$$\mathrm{TV}(F) := \mathrm{TV}_{[a,b]}(F) := \sup_{n\in\mathbb{N}} \left\{ \sum_{k=1}^{n} |F(x_k) - F(x_{k-1})| \;\middle|\; a = x_0 < x_1 < \cdots < x_n = b \right\}$$

heißt *Totalvariation* (engl.: total variation) von F. F heißt *von beschränkter Variation* (engl.: of bounded variation), falls $\mathrm{TV}(F) < \infty$. Für ein Intervall $I \subseteq \mathbb{R}$ definieren wir außerdem:

$$\mathrm{BV}(I) := \{F\colon I \to \mathbb{R} \mid \mathrm{TV}(F) < \infty\}.$$

Definition 7.33. Sei $I \subseteq \mathbb{R}$ ein Intervall. Eine Funktion $F\colon I \to \mathbb{R}$ heißt *absolut stetig* (engl.: absolutely continuous), falls es für alle $\varepsilon > 0$ ein $\delta > 0$ gibt, sodass für alle Auswahlen disjunkter Teilintervalle $(c_1, d_1), \ldots, (c_n, d_n) \subseteq I$ aus

$$\sum_{k=1}^{n} (d_k - c_k) < \delta$$

folgt:

$$\sum_{k=1}^{n} |F(d_k) - F(c_k)| < \varepsilon.$$

Lemma 7.34.

1. *Jede Lipschitz-stetige Funktion ist auch absolut stetig.*
2. *Für $a, b, c \in \mathbb{R}$ mit $a \leqslant b \leqslant c$ gilt: $TV_{[a,c]}(F) = TV_{[a,b]}(F) + TV_{[b,c]}(F)$.*
3. *Jede absolut stetige Funktion ist auch von beschränkter Variation.*

[Übung]

Satz 7.35. *Seien $a, b \in \mathbb{R}$ mit $a < b$. Falls $F \in \mathrm{BV}\big([a, b]\big)$, so gibt es zwei isotone Funktionen $F_1, F_2\colon [a, b] \to \mathbb{R}$ mit $F_1(x) = \mathrm{TV}_{[a,x]}(F)$ für alle $x \in [a, b]$ sowie $F = F_1 - F_2$. Umgekehrt gilt für zwei isotone Funktionen $F_1, F_2\colon [a, b] \to \mathbb{R}$, dass $F = F_1 - F_2 \in \mathrm{BV}\big([a, b]\big)$. Insbesondere ist jede isotone Funktion $G\colon [a, b] \to \mathbb{R}$ von beschränkter Variation.*

Beweis. Sei zunächst $F \in \mathrm{BV}\big([a, b]\big)$. Definiere $F_1[a, b] \to \mathbb{R}$ mit $x \mapsto \mathrm{TV}_{[a,x]}(F)$ sowie $F_2[a, b] \to \mathbb{R}$ mit $x \mapsto F_1(x) - F(x)$. Seien $x, y \in [a, b]$ mit $x \leqslant y$. Dann gilt wegen Lemma 7.34:

$$F_1(y) = \mathrm{TV}_{[a,y]}(F) = \mathrm{TV}_{[a,x]}(F) + \mathrm{TV}_{[x,y]}(F) \geqslant \mathrm{TV}_{[a,x]}(F) = F_1(x).$$

Also ist F_1 isoton.

Angenommen, F_2 wäre nicht isoton. Dann gäbe es $x, y \in [a, b]$ mit $x \leqslant y$, aber $F_2(x) > F_2(y)$. Somit gilt:

$$F_1(x) - F(x) = \mathrm{TV}_{[a,x]}(F) - F(x) > \mathrm{TV}_{[a,y]}(F) - F(y) = F_1(y) - F(y).$$

Mit Lemma 7.34 erhält man:

$$F(y) - F(x) > \mathrm{TV}_{[x,y]}(F)$$

Dies steht im Widerspruch zur Definition der Totalvariation. Folglich ist F_2 isoton.

Außerdem gilt:

$$F_1 - F_2 = F_1 - (F_1 - F) = F.$$

Seien nun $F_1, F_2\colon [a,b] \to \mathbb{R}$ isotone Funktionen und $F := F_1 - F_2$. Es gilt:

$$\begin{aligned}
&\mathrm{TV}_{[a,b]}(F) \\
=&\sup_{n\in\mathbb{N}}\left\{\sum_{k=1}^{n} |F(x_k) - F(x_{k-1})| \,\middle|\, a = x_0 < x_1 < \cdots < x_n = b\right\} \\
=&\sup_{n\in\mathbb{N}}\left\{\sum_{k=1}^{n} \left|F_1(x_k) - F_1(x_{k-1}) - \big(F_2(x_k) - F_2(x_{k-1})\big)\right| \,\middle|\, a = x_0 < \cdots < x_n = b\right\} \\
\leqslant&\sup_{n\in\mathbb{N}}\left\{\sum_{k=1}^{n} |F_1(x_k) - F_1(x_{k-1})| + \left|\big(F_2(x_k) - F_2(x_{k-1})\big)\right| \,\middle|\, a = x_0 < \cdots < x_n = b\right\} \\
\leqslant&\sup_{n\in\mathbb{N}}\left\{\sum_{k=1}^{n} |F_1(x_k) - F_1(x_{k-1})| \,\middle|\, a = x_0 < \cdots < x_n = b\right\} \\
&+\sup_{n\in\mathbb{N}}\left\{\sum_{k=1}^{n} |F_2(x_k) - F_2(x_{k-1})| \,\middle|\, a = x_0 < \cdots < x_n = b\right\} \\
=&\mathrm{TV}_{[a,b]}(F_1) + \mathrm{TV}_{[a,b]}(F_2).
\end{aligned}$$

Es bleibt zu zeigen, dass jede isotone Funktion $G\colon [a,b] \to \mathbb{R}$ von beschränkter Variation ist. Es gilt:

$$\begin{aligned}
\mathrm{TV}_{[a,b]}(G) &= \sup_{n\in\mathbb{N}}\left\{\sum_{k=1}^{n} |G(x_k) - G(x_{k-1})| \,\middle|\, a = x_0 < x_1 < \cdots < x_n = b\right\} \\
&= \sup_{n\in\mathbb{N}}\left\{\sum_{k=1}^{n} \big(G(x_k) - G(x_{k-1})\big) \,\middle|\, a = x_0 < x_1 < \cdots < x_n = b\right\} \\
&= G(b) - G(a) < \infty.
\end{aligned}$$

7.6 ⋆Der Hauptsatz der Integralrechnung

Wir erinnern uns, dass für eine differenzierbare Funktion $F\colon [a,b] \to \mathbb{R}$ der Hauptsatz der Differenzial- und Integralrechnung gilt:

$$\int_a^b F'(x)\,\mathrm{d}x = F(b) - F(a).$$

Da die Funktionen, welche uns in der Maß- und Wahrscheinlichkeitstheorie begegnen, oft nicht überall differenzierbar sind, soll dieser Satz nun verallgemeinert werden.

Für den Beweis des Satzes von Lebesgue (Satz 7.39) werden wir die folgenden beiden Sätze benötigen.

Satz 7.36. *Seien $a < b \in \mathbb{R}$ und $f\colon (a,b) \to \mathbb{R}$ eine Funktion. Definiere*

$$f'_-(x) := \lim_{y \nearrow x} \frac{f(y) - f(x)}{y - x}, \quad f'_+(x) := \lim_{y \searrow x} \frac{f(y) - f(x)}{y - x},$$

falls die Grenzwerte existieren. Dann gibt es nur abzählbar viele $x \in (a,b)$, sodass $f'_-(x)$ und $f'_+(x)$ existieren, aber nicht gleich sind.

Beweis. Definiere

$$A := \left\{x \in (a,b) \,\middle|\, f'_-(x) \text{ und } f'_+(x) \text{ existieren, } f'_-(x) > f'_+(x)\right\}$$

sowie

$$B := \left\{x \in (a,b) \,\middle|\, f'_-(x) \text{ und } f'_+(x) \text{ existieren, } f'_-(x) < f'_+(x)\right\}.$$

Zu jedem $x \in A$ sei $r_x \in \mathbb{Q}$ mit $f'_-(x) > r_x > f'_+(x)$. Desweiteren wähle $s_x, t_x \in \mathbb{Q}$ mit $a < s_x < x < t_x < b$ sowie

$$\frac{f(y) - f(x)}{y - x} > r_x \quad \text{für alle } y \in (s_x, x), \tag{7.13}$$

$$\frac{f(y) - f(x)}{y - x} < r_x \quad \text{für alle } y \in (x, t_x). \tag{7.14}$$

Aus (7.13) und (7.14) folgt:

$$f(y) - f(x) < r_x(y - x) \quad \text{für alle } y \in (s_x, t_x) \text{ mit } y \neq x. \tag{7.15}$$

Somit erhalten wir eine Abbildung

$$\phi\colon A \to \mathbb{Q}^3, \quad x \mapsto (r_x, s_x, t_x).$$

Es bleibt zu zeigen, dass ϕ injektiv ist. Angenommen, es gäbe $x, y \in A$ mit $x \neq y$, aber $\phi(x) = \phi(y)$. Dann gilt: $x, y \in (s_x, t_x) = (s_y, t_y)$. Aus (7.15) folgt:

$$f(y) - f(x) < r_x(y - x),$$
$$f(x) - f(y) < r_y(x - y).$$

Da $r_x = r_y$, führt Addition dieser beiden Ungleichungen auf $0 < 0$. Dies ist ein Widerspruch. Folglich ist ϕ injektiv und somit A abzählbar. Analog folgt die Abzählbarkeit von B.

Definition 7.37. Sei $E \subseteq \mathbb{R}$. Eine Familie $\mathscr{V}$ von abgeschlossenen Intervallen positiver Länge in $\mathbb{R}$ heißt *Vitali-Überdeckung* (engl.: Vitali covering) von E, falls es für alle $x \in E$ und alle $\varepsilon > 0$ ein $V \in \mathscr{V}$ mit $x \in V$ und $\lambda(V) < \varepsilon$ gibt.

Satz 7.38 (Vitalis Überdeckungssatz). *Seien $E \subseteq \mathbb{R}$ und $\mathscr{V} \neq \varnothing$ eine Vitali-Überdeckung von E. Dann gibt es eine Folge $(V_n)_{n\in\mathbb{N}}$ paarweise disjunkter Intervalle in $\mathscr{V}$ mit*

$$\lambda\left(E \cap \left(\bigcup_{n\in\mathbb{N}} V_n\right)^c\right) = 0. \tag{7.16}$$

Falls $\lambda(E) < \infty$, gibt es für alle $\varepsilon > 0$ eine endliche Menge $\{V_1, \dots, V_m\} \subseteq \mathscr{V}$ mit

$$\lambda\left(E \cap \left(\bigcup_{n=1}^{m} V_n\right)^c\right) < \varepsilon.$$

Beweis. Sei zunächst $\lambda(E) < \infty$. Betrachte eine offene Menge $W \subseteq \mathbb{R}$ mit $E \subseteq W$ und $\lambda(W) < \infty$. Definiere eine Vitali-Überdeckung

$$\mathscr{V}_0 := \{V \in \mathscr{V} \mid V \subseteq W\}$$

von E. Finden wir eine endliche Menge $\{V_1, \dots, V_m\} \subseteq \mathscr{V}_0$ mit $E \subseteq \bigcup_{k=1}^m V_k$, so folgt die Behauptung. Andernfalls definieren wir rekursiv eine Folge $(V_n)_{n\in\mathbb{N}}$ in $\mathscr{V}_0$, welche (7.16) erfüllt. Wähle $V_1 \in \mathscr{V}_0$. Für $V_1, \dots, V_n \in \mathscr{V}_0$ paarweise disjunkt definiere

$$A_n := \bigcup_{k=1}^{n} V_k, \quad B_n := W \cap A_n^c.$$

Offenbar ist A_n abgeschlossen, B_n offen und $B_n \cap E \neq \varnothing$. Definiere

$$\delta_n := \sup\{\lambda(V) \mid V \in \mathscr{V}_0, V \subseteq B_n\}. \tag{7.17}$$

Wähle $V_{n+1} \in \mathscr{V}_0$ so, dass $V_{n+1} \subseteq B_n$ und $\lambda(V_{n+1}) > 1/2\delta_n$. Dies liefert eine Folge $(V_n)_{n\in\mathbb{N}}$ paarweise disjunkter Mengen in $\mathscr{V}_0$. Definiere $A := \bigcup_{n\in\mathbb{N}} V_n$. Wir wollen $\lambda(E \cap A^c) = 0$ zeigen. Für alle $n \in \mathbb{N}$ sei U_n das abgeschlossene Intervall mit demselben Mittelpunkt wie V_n, aber fünffacher Länge. Dann gilt:

$$\lambda\left(\bigcup_{n\in\mathbb{N}} U_n\right) \leqslant \sum_{n\in\mathbb{N}} \lambda(U_n) = 5\sum_{n\in\mathbb{N}} \lambda(V_n) = 5\lambda(A) \leqslant 5\lambda(W) < \infty. \tag{7.18}$$

Es folgt:

$$\lim_{k\to\infty} \lambda\left(\bigcup_{n=k}^{\infty} U_n\right) \leqslant \lim_{k\to\infty} \sum_{n=k}^{\infty} \lambda(U_n) = 0.$$

Somit reicht es, $E \cap A^c \subseteq \bigcup_{n=k}^{\infty} U_n$ für alle $k \in \mathbb{N}$ zu zeigen. Seien $k \in \mathbb{N}$ und $x \in E \cap A^c$. Dann ist $x \in E \cap A_k^c \subseteq B_k$. Da B_k offen ist, gibt es ein $\varepsilon > 0$ mit $(x - \varepsilon, x + \varepsilon) \subseteq B_k$. Da $\mathscr{V}_0$ eine Vitali-Überdeckung ist, gibt es ein $V \in \mathscr{V}_0$ mit $x \in V$ und $\lambda(V) < 1/3\varepsilon$. Folglich ist $V \subseteq (x - \varepsilon, x + \varepsilon) \subseteq B_k$. Nach Wahl von V_{n+1} ist $\delta_n < 2\lambda(V_{n+1})$ und aus (7.18) folgt: $\lim_{n\to\infty} \lambda(V_n) = 0$. Da $\mathscr{V}_0$ eine Vitali-Überdeckung ist, gibt es somit ein $n \in \mathbb{N}$ mit $\delta_n < \lambda(V)$. Aus (7.17) folgt: $V \nsubseteq B_n$. Definiere $q := \min\{n \in \mathbb{N} \mid V \nsubseteq B_n\}$. Somit ist $k < q$. Da q minimal gewählt wurde, gilt:

$$V \cap A_q \neq \varnothing, \quad V \cap A_{q-1} = \varnothing$$

und somit

$$V \cap V_q \neq \varnothing. \tag{7.19}$$

Weiterhin folgt aus $V \subseteq B_{q-1}$ die Ungleichung

$$\lambda(V) \leqslant \delta_{q-1} < 2\lambda(V_q). \tag{7.20}$$

Aus (7.19) folgt die Existenz eines $y \in V \cap V_q$. Mit (7.20) erhalten wir:

$$V \subseteq \left[y - 2\lambda(V_q), y + 2\lambda(V_q)\right] \subseteq U_q \subseteq \bigcup_{n=k}^{\infty} U_n.$$

Somit ist $x \in \bigcup_{n=k}^{\infty} U_n$ und es folgt wie gewünscht $\lambda(E \cap A^c) = 0$. Für $\varepsilon > 0$ wähle ein $k \in \mathbb{N}$, sodass

$$\sum_{n=k}^{\infty} \lambda(V_n) < \varepsilon.$$

Es gilt:

$$E \cap A_k^c \subseteq (E \cap A^c) \cup \bigcup_{n=k+1}^{\infty} V_n.$$

Somit folgt:

$$\lambda(E \cap A_k^c) \leqslant \lambda\left(\bigcup_{n=k+1}^{\infty} V_n\right) < \infty.$$

Die Menge $\{V_1, \ldots, V_k\}$ liefert also das gewünschte Ergebnis.

Sei nun $\lambda(E) = \infty$. Für alle $z \in \mathbb{Z}$ definiere $E_z := E \cap (z, z+1)$ sowie

$$\mathscr{V}_z := \{V \in \mathscr{V} \mid V \subseteq (z, z+1)\}.$$

$\mathscr{V}_z$ ist eine Vitali-Überdeckung von E_z. Da $\lambda(E_z) < \infty$, finden wir Folgen $(V_n^z)_{n\in\mathbb{N}}$ paarweise disjunkter Intervalle in $\mathscr{V}_z$ mit

$$\lambda\left(E_z \cap \left(\bigcup_{n\in\mathbb{N}} V_n^z\right)^c\right) = 0 \quad \text{für alle } z \in \mathbb{Z}.$$

Sei $V := \{V_n^z \mid n \in \mathbb{N}, z \in \mathbb{Z}\} \subseteq \mathscr{V}$. Nach Definition von E_z sind die Mengen in V paarweise disjunkt und als abzählbare Vereinigung abzählbarer Mengen ist V abzählbar. Somit können wir eine Folge $(T_n)_{n\in\mathbb{N}}$ disjunkter Intervalle in V mit $V = \{T_n\}_{n\in\mathbb{N}}$ finden. Es gilt:

$$E \cap \left(\bigcup_{n\in\mathbb{N}} T_n\right)^c \subseteq \mathbb{Z} \cup \bigcup_{z\in\mathbb{Z}} \left(E_z \cap \left(\bigcup_{n\in\mathbb{N}} V_n^z\right)^c\right).$$

Mit der σ-Additivität von λ folgt:

$$\lambda\left(E \cap \left(\bigcup_{n\in\mathbb{N}} T_n\right)^c\right) \leqslant \lambda(\mathbb{Z}) + \sum_{z\in\mathbb{Z}} 0 = 0.$$

Satz 7.39 (Lebesgue)**.** *Seien $a, b \in \mathbb{R}$ mit $a < b$ und $F\colon [a,b] \to \mathbb{R}$ isoton. Dann existiert λ-fast-überall eine Ableitung F' von F, die $\mathscr{B}|_{[a,b]}$-messbar ist.*

Beweis. Definiere

$$E := \left\{ x \in [a,b) \;\middle|\; \liminf_{h\to 0} \frac{F(x+h)-F(x)}{h} < \limsup_{h\to 0} \frac{F(x+h)-F(x)}{h} \right\}.$$

Wir zeigen zunächst: $\lambda(E) = 0$. Zu $u, v \in \mathbb{Q}$ mit $u < v$ definiere

$$E_{u,v} := \left\{ x \in E \;\middle|\; \liminf_{h\to 0} \frac{F(x+h)-F(x)}{h} < u < v < \limsup_{h\to 0} \frac{F(x+h)-F(x)}{h} \right\}.$$

Es gilt:

$$E = \bigcup_{u,v\in\mathbb{Q},0<u<v} E_{u,v}.$$

Es reicht also, $\lambda(E_{u,v}) = 0$ für alle $u, v \in \mathbb{Q}$ mit $0 < u < v$ zu zeigen. Angenommen, es würden $u, v \in \mathbb{Q}$ mit $0 < u < v$ existieren, sodass $\lambda(E_{u,v}) = \alpha > 0$. Wähle $\varepsilon > 0$ mit

$$\varepsilon < \frac{\alpha(v-u)}{u+2v}.$$

Es gibt eine offene Menge $U \supseteq E_{u,v}$ mit $\lambda(U) < \alpha + \varepsilon$. Für alle $x \in E_{u,v}$ gibt es ein $h > 0$, sodass $[x, x+h] \subseteq U \cap [a,b]$. Nach Definition von $E_{u,v}$ kann h sogar so gewählt werden, dass außerdem gilt:

$$F(x+h) - F(x) < uh. \tag{7.21}$$

Definiere

$$\mathscr{V} := \{[x, x+h] \mid x \in E_{u,v}, h > 0, [x, x+h] \subseteq U \cap [a,b], F(x+h) - F(x) < uh\}.$$

Dies ist eine Vitali-Überdeckung von $E_{u,v}$. Nach Satz 7.38 gibt es eine endliche Teilmenge $\{[x_i, x_i + h_i]\}_{i=1}^m \subseteq \mathscr{V}$ paarweise disjunkter Intervalle mit

$$\lambda\left(E_{u,v} \cap \left(\bigcup_{i=1}^m [x_i, x_i + h_i] \right)^c \right) < \varepsilon.$$

Definiere

$$V := \bigcup_{i=1}^m (x_i, x_i + h_i).$$

Es gilt:

$$\lambda(E_{u,v} \cap V^c) < \varepsilon. \tag{7.22}$$

Aus $V \subseteq U$ folgt:

$$\sum_{i=1}^m h_i = \lambda(V) \leqslant \lambda(U) < \alpha + \varepsilon.$$

Mit (7.21) erhalten wir:

$$\sum_{i=1}^{m} \big(F(x_i + h_i) - F(x_i)\big) < u \sum_{i=1}^{m} h_i < u(\alpha + \varepsilon). \tag{7.23}$$

Außerdem gibt es für alle $y \in E_{u,v} \cap V$ ein $k > 0$ mit $[y, y + k] \subseteq V$ und

$$F(y + k) - F(y) > vk. \tag{7.24}$$

Dies liefert wiederum eine Vitali-Überdeckung

$$\mathscr{V}' := \{[y, y + k] \mid y \in E_{u,v} \cap V, k > 0, [y, y + k] \subseteq V, F(y + k) - F(y) > vk\}$$

von $E_{u,v} \cap V$ und somit eine endliche Teilmenge $\{[y_j, y_j + k_j]\}_{j=1}^{n} \subseteq \mathscr{V}'$ paarweise disjunkter Intervalle mit

$$\lambda\left(E_{u,v} \cap V \cap \left(\bigcup_{j=1}^{n} [y_j, y_j + k_j]\right)^c\right) < \varepsilon.$$

Zusammen mit (7.22) erhalten wir:

$$\alpha = \lambda(E_{u,v}) = \lambda(E_{u,v} \cap V^c) + \lambda(E_{u,v} \cap V) < \varepsilon + \left(\varepsilon + \sum_{j=1}^{n} k_j\right). \tag{7.25}$$

Dies ergibt zusammen mit (7.24):

$$v(\alpha - 2\varepsilon) < v \sum_{j=1}^{n} k_j < \sum_{j=1}^{n} \big(F(y_j + k_j) - F(y_j)\big). \tag{7.26}$$

Da $\sum_{j=1}^{n} [y_j, y_j + k_j] \subseteq \sum_{i=1}^{m} [x_i, x_i + h_i]$ und F isoton ist, gilt:

$$\sum_{j=1}^{n} \big(F(y_j + k_j) - F(y_j)\big) \leqslant \sum_{i=1}^{m} \big(F(x_i + h_i) - F(x_i)\big). \tag{7.27}$$

Zusammen mit (7.23) und (7.26) erhalten wir:

$$v(\alpha - 2\varepsilon) < u(\alpha + \varepsilon).$$

Dies liefert

$$\alpha(v - u) < \varepsilon(u + 2v),$$

im Widerspruch zur Wahl von ε. Somit ist $\lambda(E) = 0$ und folglich existiert $F'_+(x)$ λ-fast-überall auf $[a, b]$. Analog erhalten wir die Existenz von $F'_-(x)$ λ-fast-überall auf $[a, b]$. Nach Satz 7.36 folgt die Existenz von $F'(x)$ λ-fast-überall auf $[a, b]$.

Es bleibt zu zeigen, dass $G := \{x \in (a,b) \mid F'(x) = \infty\}$ eine λ-Nullmenge ist. Sei $\beta > 0$. Für alle $x \in G$ gibt es ein $h > 0$ mit $[x, x+h] \subseteq (a,b)$ und

$$F(x+h) - F(x) > \beta h. \tag{7.28}$$

Betrachte die Vitali-Überdeckung

$$\mathscr{V}'' := \{[x, x+h] \mid x \in G, h > 0, [x, x+h] \subseteq (a,b), F(x+h) - F(x) > \beta h\}$$

von G. Nach Satz 7.38 können wir eine Folge $\big([x_t, x_t + h_t]\big)_{t\in\mathbb{N}}$ paarweise disjunkter Intervalle in $\mathscr{V}''$ finden, sodass gilt:

$$\lambda\left(G \cap \left(\bigcup_{t\in\mathbb{N}} [x_t, x_t + h_t]\right)^c\right) = 0.$$

Zusammen mit (7.28) ergibt sich:

$$\beta\lambda(G) \leqslant \beta \sum_{t\in\mathbb{N}} h_t < \sum_{t\in\mathbb{N}} \big(F(x_t + h_t) - F(x_t)\big) \leqslant F(b) - F(a).$$

Somit folgt:

$$\beta\lambda(G) < F(b) - F(a)$$

für alle $\beta \in \mathbb{R}$. Somit ist $\lambda(G) = 0$.

Satz 7.40. *Seien $a, b \in \mathbb{R}$ mit $a < b$.*

1. *Jedem endlichen signierten Maß ϕ auf dem Messraum $\big([a,b], \mathscr{B}|_{[a,b]}\big)$ entspricht genau eine maßdefinierende Funktion $F \in \mathrm{BV}\big([a,b]\big)$ und umgekehrt.*

2. *Sei $F\colon [a,b] \to \mathbb{R}$ absolut stetig. Nach Lemma 7.34 und Satz 7.35 gibt es isotone Funktionen $F_1, F_2\colon [a,b] \to \mathbb{R}$ sodass $F = F_1 - F_2$. F_1 und F_2 sind absolut stetig und F ist λ-fast-überall differenzierbar mit $F' = F_1' - F_2'$ λ-fast-überall.*

Beweis.

1. Sei zunächst ϕ ein endliches signiertes Maß auf dem Messraum $\big([a,b], \mathscr{B}|_{[a,b]}\big)$. Wir betrachten die Jordan-Hahn-Zerlegung (siehe Satz 7.30) $\phi = \phi^+ - \phi^-$ mit zugehörigen maßdefinierenden Funktionen $F = F^+ - F^-$. Aus Satz 7.35 folgt: $F \in \mathrm{BV}\big([a,b]\big)$.

   Sei nun $F \in \mathrm{BV}\big([a,b]\big)$ rechtsstetig. Nach Satz 7.35 gibt es zwei isotone Funktionen $F_1, F_2\colon [a,b] \to \mathbb{R}$ mit $F = F_1 - F_2$. Beachte, dass nach Konstruktion von F_1 in Satz 7.35 diese beiden Funktionen ebenfalls rechtsstetig sind. Somit handelt es sich um maßdefinierende Funktionen und $F = F_1 - F_2$ induziert ein endliches signiertes Maß auf dem Messraum $\big([a,b], \mathscr{B}|_{[a,b]}\big)$.

2. Nach Satz 7.39 sind F_1 und F_2 und somit auch F λ-fast-überall differenzierbar. Es bleibt zu zeigen, dass F_1 und F_2 absolut stetig sind. Sei $\varepsilon > 0$. Wähle $\delta > 0$ klein genug, dass für alle Auswahlen disjunkter Teilintervalle $(a_1, b_1), \dots, (a_p, b_p) \subseteq [a, b]$ aus
$$\sum_{i=1}^{p} (a_i - b_i) < \delta$$
folgt:
$$\sum_{i=1}^{p} |F(b_i) - F(a_i)| < \frac{\varepsilon}{2}.$$
Sei $(c_1, d_1), \dots, (c_n, d_n) \subseteq [a, b]$ eine Auswahl disjunkter Teilintervalle. Nach der Definition der Totalvariation gibt es für alle $k \in \{1, \dots, n\}$ eine Zerlegung $Z = \left\{z_{n,k}^i\right\}_{i=0}^{m_k}$ des Intervalls $[c_k, d_k]$ mit
$$\mathrm{TV}_{[c_k,d_k]}(F) < \sum_{i=1}^{m_k} \left|F(z_{n,k}^i) - F(z_{n,k}^{i-1})\right| + \frac{\varepsilon}{2n}.$$
Es gilt:
$$\sum_{k=1}^{n} \sum_{i=1}^{m_k} \left(z_{n,k}^i - z_{n,k}^{i-1}\right) = \sum_{k=1}^{n} (d_k - c_k) < \delta$$
und somit:
$$\sum_{k=1}^{n} \sum_{i=1}^{m_k} \left|F(z_{n,k}^i) - F(z_{n,k}^{i-1})\right| \leqslant \frac{\varepsilon}{2}.$$
Es folgt:
$$\begin{aligned}\sum_{k=1}^{n} |F_1(d_k) - F_1(c_k)| &= \sum_{k=1}^{n} \left|\mathrm{TV}_{[a,d_k]}(F) - \mathrm{TV}_{[a,c_k]}(F)\right| = \sum_{k=1}^{n} \mathrm{TV}_{[c_k,d_k]}(F) \\ &\leqslant \sum_{k=1}^{n} \sum_{i=1}^{m_k} \left|F(z_{n,k}^i) - F(z_{n,k}^{i-1})\right| + \sum_{k=1}^{n} \frac{\varepsilon}{2n} \leqslant \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.\end{aligned}$$
Somit ist F_1 absolut stetig und daher auch $F_2 = F_1 - F$.

Satz 7.41 (Hauptsatz der Integralrechnung)**.** *Seien $a, b \in \mathbb{R}$ mit $a < b$.*

1. *Sei $F\colon [a,b] \to \mathbb{R}$ absolut stetig. Dann existiert die Ableitung F' λ-fast-überall und ist λ-integrierbar mit*
$$F(x) - F(a) = \int_{[a,x]} F' \,\mathrm{d}\lambda$$
für alle $x \in [a,b]$.

2. *Sei $f\colon [a,b] \to \mathbb{R}$ λ-integrierbar mit*
$$F(x) - F(a) = \int_{[a,x]} f \,\mathrm{d}\lambda$$
für alle $x \in [a,b]$, so ist F absolut stetig und es ist $F' = f$ λ-fast-überall.

Der Beweis des Hauptsatzes erfordert erheblich mehr Vorarbeit, was den Rahmen dieses Skripts sprengen würde. Der vollständige Beweis kann zum Beispiel in [9], Seite 285, nachgelesen werden.

Man beachte, dass mit den uns zur Verfügung stehenden Mitteln die λ-fast-überall bestehende Existenz von F' bereits gezeigt werden kann. Hierzu zerlege man F in eine Differenz zweier monotoner Funktionen, für welche der Satz von Lebesgue die gewünschte Aussage liefert.

7.7 Die Lebesgue-Zerlegung einer Verteilungsfunktion

Analog zu Satz 7.25 lassen sich auch Verteilungfunktionen auf geeignete Weise zerlegen, wobei wir anstelle eines allgemeinen Maßes μ das Lebesgue-Maß betrachten. Dies führt auf die sogenannte Lebesgue-Zerlegung von Verteilungsfunktionen. Der folgende Satz zeigt, dass der absolut stetige Anteil in der Zerlegung aus Satz 7.25 genau einem absolut stetigen Anteil in der Verteilungsfunktion entspricht.

Satz 7.42. *Sei $F\colon \mathbb{R} \to [0,1]$ eine Verteilungsfunktion. Dann ist F genau dann absolut stetig, wenn das durch F bestimmte Wahrscheinlichkeitsmaß P_F absolut stetig bezüglich λ ist, wenn also λ-fast-überall die Radon-Nikodym-Dichte*
$$f = \frac{\mathrm{d}P_F}{\mathrm{d}\lambda}$$
existiert und λ-fast-überall $F' = f$ gilt.

Beweis. Sei zunächst $P_F \ll \lambda$. Nach Lemma 7.10 gibt es für alle $\varepsilon > 0$ ein $\delta > 0$, sodass für alle $A \in \mathscr{B}$ mit $\lambda(A) < \delta$ gilt: $P_F(A) < \varepsilon$. Sei $(a_1, b_1), \ldots, (a_n, b_n) \subseteq \mathbb{R}$ eine Auswahl disjunkter Intervalle. Mit
$$\lambda\left(\bigcup_{i=1}^{n} [a_i, b_i)\right) = \sum_{i=1}^{n} (b_i - a_i) < \delta$$

folgt:

$$\sum_{i=1}^{n}\big(F(b_i)-F(a_i)\big)=\sum_{i=1}^{n}P_F\big([a_i,b_i)\big)=P_F\left(\bigcup_{i=1}^{n}[a_i,b_i)\right)<\varepsilon.$$

Somit ist F absolut stetig.

Sei nun F absolut stetig. Für $\varepsilon>0$ wähle $\delta>0$, sodass für alle Auswahlen disjunkter Intervalle $(a_1,b_1),\dots,(a_n,b_n)\subseteq\mathbb{R}$ mit

$$\sum_{i=1}^{n}(b_i-a_i)<\delta$$

gilt:

$$\sum_{i=1}^{n}\big(F(b_i)-F(a_i)\big)<\varepsilon.$$

Sei A eine λ-Nullmenge. Dann gibt es eine disjunkte Überdeckung von A durch halboffene Intervalle $[a_1,b_1),[a_2,b_2),\dots\subseteq\mathbb{R}$ mit

$$\sum_{i\in\mathbb{N}}(b_i-a_i)<\delta.$$

Beachte hierfür, dass das Lebesgue-Maß auf der borelschen σ-Algebra definiert ist durch

$$\lambda(B)=\inf\left\{\sum_{n\in\mathbb{N}}\lambda\big([a_n,b_n)\big)\;\middle|\;B\subseteq\bigcup_{n\in\mathbb{N}}[a_n,b_n)\right\}.$$

Die obige Aussage folgt somit aus der Definition des Infimums. Für alle $n\in\mathbb{N}$ folgt:

$$\sum_{i=1}^{n}F(b_i)-F(a_i)<\varepsilon.$$

Somit gilt für alle $\varepsilon>0$:

$$P_F(A)\leqslant\sum_{i\in\mathbb{N}}P_F\big([a_i,b_i)\big)=\sum_{i\in\mathbb{N}}F(b_i)-F(a_i)\leqslant\varepsilon.$$

Es folgt: $P_F(A)=0$.

Definition 7.43. Wir bezeichnen eine Verteilungsfunktion $F\colon\mathbb{R}\to[0,1]$ als *stetig singulär* (engl.: continuously singular), falls λ-fast-überall $F'=0$ gilt. Wir bezeichnen F als *diskret* (engl.: discrete), falls das zugehörige Wahrscheinlichkeitsmaß diskret ist (vergleiche Beispiel 1.26.4).

Lemma 7.44. *Jede Verteilungsfunktion $F\colon\mathbb{R}\to[0,1]$ hat höchstens abzählbar viele Sprungstellen.* [Übung]

Entsprechend gilt: Jede isotone Funktion $F\colon\mathbb{R}\to\mathbb{R}$ hat höchstens abzählbar viele Sprungstellen auf einer kompakten Teilmenge von $\mathbb{R}$.

Satz 7.45 (Lebesgue-Zerlegung einer Verteilungsfunktion).
Jede Verteilungsfunktion $F\colon \mathbb{R} \to [0,1]$ *besitzt genau eine Darstellung*

$$F = a_1 F_d + a_2 F_{cs} + a_3 F_{ac},$$

wobei $a_1, a_2, a_3 \geqslant 0$ *mit* $a_1 + a_2 + a_3 = 1$, F_d *eine diskrete,* F_{cs} *eine stetig singuläre und* F_{ac} *eine absolut stetige Verteilungsfunktion ist.*

Beweis. Sei D die Menge der Unstetigkeitsstellen von F. Diese ist nach Lemma 7.44 höchstens abzählbar. Definiere

$$a_1 := \begin{cases} 0 & D = \varnothing \\ \sum_{x \in D} \big(F(x) - F(x-)\big) & \text{sonst} \end{cases}.$$

Für $a_1 \neq 0$ erhalten wir eine diskrete Verteilungsfunktion

$$F_d\colon \mathbb{R} \to [0,1], \quad y \mapsto \frac{1}{a_1} \sum_{x \in D} P\big(\{x\}\big) \cdot \mathbb{1}_{[x,\infty)}(y).$$

Betrachte eine stetige Verteilungsfunktion $F_c := F - a_1 F_d$. Nach Satz 7.39 ist F_c λ-fast-überall differenzierbar und F_c' λ-integrierbar. Definiere

$$a_3 := \int F_c' \,\mathrm{d}\lambda.$$

Falls $a_3 = 0$, setze $F_{cs} := F_c$ und $a_2 := 1 - a_1$. Andernfalls definiere

$$F_{ac}(y) := \frac{1}{a_3} \int_{(-\infty,y]} F_c' \,\mathrm{d}\lambda$$

und $a_2 := 1 - a_1 - a_3$ sowie

$$F_{cs} := \frac{1}{a_2}(F_c - a_3 F_{ac}).$$

Nach dem Hauptsatz der Integralrechnung (Satz 7.41) ist F_{ac} absolut stetig und es gilt:

$$\begin{aligned} \int F_{cs}' \,\mathrm{d}\lambda = \int \frac{1}{a_2}(F_c' - a_3 F_{ac}') \,\mathrm{d}\lambda = \frac{a_3}{a_2} - \frac{1}{a_2} \int \left(\int_{(-\infty,y]} F_c' \,\mathrm{d}\lambda \right)' \lambda(\mathrm{d}y) \\ = \frac{a_3}{a_2} - \frac{1}{a_2} \int F_c' \,\mathrm{d}\lambda = \frac{a_3}{a_2} - \frac{a_3}{a_2} = 0. \end{aligned}$$

Somit ist F_{cs} stetig singulär.

Es bleibt die Eindeutigkeit der Zerlegung zu zeigen. Seien $a_1 F_d + a_2 F_{cs} + a_3 F_{ac}$ und $a_1' F_d' + a_2' F_{cs}' + a_3' F_{ac}'$ zwei Lebesgue-Zerlegungen von F. Nach Definition 7.43 liefert $a_1 F_d$ ein endliches diskretes Maß ν_s. Ferner liefert $a_2 F_{cs} + a_3 F_{ac}$ ein endliches, stetiges Maß μ. Folglich gilt: $\nu_s \perp \mu$. Nun können wir Satz 7.25 mit $\nu_a := \mu$ anwenden und erhalten die Eindeutigkeit von $a_1 F_d$ sowie von $a_2 F_{cs} + a_3 F_{ac}$. Es gilt also $a_1 F_d = a_1' F_d'$ und $a_2 F_{cs} + a_3 F_{ac} = a_2' F_{cs}' + a_3' F_{ac}'$. Aus dem Hauptsatz der Integralrechnung folgt:

$$a_3 F_{ac}(y) = \int_{(-\infty,y]} F - a_1 F_d \,\mathrm{d}\lambda = a_3' F_{ac}'(y).$$

Somit folgt auch $a_2 F_{cs} = a_2' F_{cs}'$. Da F_d, F_{cs}, F_{ac}, F_d', F_{cs}' und F_{ac}' Verteilungsfunktionen sind, folgt: $a_1 = a_1'$, $a_2 = a_2'$ und $a_3 = a_3'$. Somit ist die Zerlegung eindeutig.

Übungsaufgaben:

Übung 1: Beweisen Sie den Fall der σ-Endlichkeit des Maßes ν im Satz von Radon-Nikodym (Korollar 7.18).

Übung 2: Beweisen Sie, dass man Lemma 7.10 nicht auf σ-endliche Maße ν verallgemeinern kann.

Übung 3: Beweisen Sie die Regeln 7.20.

Übung 4: Beweisen Sie Beispiel 7.28.

Übung 5: Beweisen Sie Punkt 1 in Bemerkung 7.31.

Übung 6: Beweisen Sie Lemma 7.34.

Übung 7: Beweisen Sie Lemma 7.44.

Übung 8: Es bezeichne $\mathcal{N}_{\mu,\sigma^2}$ das Normalverteilungsmaß mit Erwartungswert μ und Varianz σ^2. Bestimmen Sie einen Dichtequotienten von $\frac{\mathrm{d}\mathcal{N}_{\mu_1,\sigma_1^2}}{\mathrm{d}\mathcal{N}_{\mu_2,\sigma_2^2}}$.

Übung 9: Seien P und Q Wahrscheinlichkeitsmaße auf einem Messraum $(\Omega, \mathscr{A})$. Wir definieren die *Kullback-Leibler-Divergenz* von P und Q als

$$KL(P,Q) := \begin{cases} \int \log\left(\frac{\mathrm{d}P}{\mathrm{d}Q}\right) \mathrm{d}P & \text{falls } P \ll Q \\ \infty & \text{sonst} \end{cases}.$$

Die Kullback-Leibler-Divergenz liefert eine Art Abstand zwischen Wahrscheinlichkeitsmaßen. Sie ist jedoch insbesondere nicht symmetrisch und daher keine Metrik. Zeigen Sie, dass stets gilt: $KL(P,Q) \geqslant 0$.

Lernziel-Kontrolle:

Frage 1: Ist jede differenzierbare Funktion absolut stetig? Wenn ja, wie erhält man ε und δ?

Frage 2: Ist jede absolut stetige Funktion auch Lipschitz-stetig?

Frage 3: Folgt aus der Monotonie einer Verteilungsfunktion bereits absolute Stetigkeit?

Frage 4: Was besagt der Satz von Radon-Nikodym?

Frage 5: Vergleichen sie die Radon-Nikodym-Dichte mit der Ableitung.

Frage 6: Zu welchem Zweck führt man den Begriff der Dichte ein?

Frage 7: Ist die Cantorsche Verteilungsfunktion absolut stetig?

Kapitel 8

Produkträume und der Satz von Fubini

Das Ziel dieses Kapitels ist es, auf geeignete Weise Produkte von Maßräumen zu konstruieren, sodass auf natürliche Weise ein Produktmaß entsteht. Dies erlaubt auch mehrdimensionale Integration auf den eindimensionalen Fall zurückzuführen.

Problem 8.1. Seien $(\Omega_j, \mathscr{A}_j, \mu_j)$ für alle $j \in \{1, \ldots, n\}$ Maßräume, $\Omega := \times_{j=1}^{n} \Omega_j$ der Produktraum, $\pi_j : \Omega \to \Omega_j$ die kanonischen Projektionen und

$$\mathscr{A} := \bigotimes_{j=1}^{n} \mathscr{A}_j = \sigma(\pi_1, \ldots, \pi_n) = \sigma\left(\bigcup_{j=1}^{n} \pi_j^{-1}(\mathscr{A}_j)\right)$$

die Produkt-σ-Algebra. Es stellt sich die Frage, ob wir ein Maß μ auf dem Messraum $(\Omega, \mathscr{A})$ finden können, welches folgende Eigenschaft besitzt:

$$\mu(A_1 \times \cdots \times A_n) = \prod_{j=1}^{n} \mu_j(A_j) \tag{8.1}$$

für alle $A_1 \times \cdots \times A_n \in \mathscr{A}$.

Satz 8.2. *Gelten die Voraussetzungen von Problem 8.1 und sind die Maße $\mu_1, \ldots, \mu_n$ zusätzlich σ-endlich, so gibt es höchstens ein Maß μ auf $(\Omega, \mathscr{A})$, mit der Eigenschaft (8.1).*

Beweis. Definiere $M_j := \{A_j \in \mathscr{A}_j \mid \mu_j(A_j) < \infty\}$ für alle $j \in \{1, \ldots, n\}$. Da die μ_j σ-endlich sind, gibt es in jedem M_j eine gegen Ω_j aufsteigende Folge von Mengen. Somit ist nach Satz 2.13 die Menge

$$M := \{A_1 \times \cdots \times A_n \mid A_j \in M_j \ \forall\, j \in \{1, \ldots, n\}\}$$

ein Erzeuger von $\bigotimes_{j=1}^{n} \mathscr{A}_j$. Um zu zeigen, dass M außerdem $\cap$-stabil ist, betrachte $A_1 \times \cdots \times A_n, B_1 \times \cdots \times B_n \in M$. Dann ist $A_i \cap B_i \in \mathscr{A}_i$ und $\mu_i(A_i \cap B_i) \leqslant \mu_i(A_i) < \infty$

für alle $i \in \{1, \dots, n\}$. Folglich ist

$$(A_1 \times \cdots \times A_n) \cap (B_1 \times \cdots \times B_n) = A_1 \cap B_1 \times \cdots \times A_n \cap B_n \in M.$$

Also ist M ein $\cap$-stabiler Erzeuger von $\bigotimes_{j=1}^n \mathscr{A}_j$ und mit dem Eindeutigkeitssatz (Satz 1.30) folgt die Behauptung, da wir für μ (8.1) fordern.

Definition 8.3. Seien Ω_1 und Ω_2 nichtleere Mengen, $\Omega := \Omega_1 \times \Omega_2$, $Q \subseteq \Omega$ und $(\omega_1, \omega_2) \in \Omega$. Dann heißen

$$\begin{aligned} \omega_1 Q &:= \{\omega \in \Omega_2 \mid (\omega_1, \omega) \in Q\}, \\ \omega_2 Q &:= \{\omega \in \Omega_1 \mid (\omega, \omega_2) \in Q\} \end{aligned}$$

ω_1- beziehungweise ω_2-*Schnitt* von Q.

Lemma 8.4. *Seien $(\Omega_1, \mathscr{A}_1)$ und $(\Omega_2, \mathscr{A}_2)$ Messräume. Seien weiterhin $Q \in \mathscr{A}_1 \otimes \mathscr{A}_2$ und $(\omega_1, \omega_2) \in \Omega_1 \times \Omega_2$. Dann gilt: $\omega_1 Q \in \mathscr{A}_2$ und $\omega_2 Q \in \mathscr{A}_1$.*

Beweis. Definiere $\mathscr{A} := \{Q \subseteq \Omega \mid \omega_1 Q \in \mathscr{A}_2\}$. Dies ist eine σ-Algebra, denn es gilt:

- $\omega_1 \varnothing = \varnothing \in \mathscr{A}_2$. Somit ist $\varnothing \in \mathscr{A}$.
- Sei $Q \in \mathscr{A}$. Dann ist $\omega_1 Q \in \mathscr{A}_2$ und somit $\omega_1 Q^c = (\omega_1 Q)^c \in \mathscr{A}_2$. Folglich ist $Q^c \in \mathscr{A}$.
- Seien $Q_1, Q_2, \ldots \in \mathscr{A}$. Dann ist $\omega_1 Q_i \in \mathscr{A}_2$ für alle $i \in \mathbb{N}$ und somit

  $$\omega_1 \bigcup_{i \in \mathbb{N}} Q_i = \bigcup_{i \in \mathbb{N}} \omega_1 Q_i \in \mathscr{A}_2.$$

  Folglich ist $\bigcup_{i \in \mathbb{N}} Q_i \in \mathscr{A}$.

Ferner gilt:

$$S := \{A_1 \times A_2 \mid A_1 \in \mathscr{A}_1, A_2 \in \mathscr{A}_2\} \subseteq \mathscr{A},$$

denn für alle $A_1 \times A_2 \in S$ gilt:

- Falls $\omega_1 \in A_1$, so ist $\omega_1(A_1 \times A_2) = A_2 \in \mathscr{A}_2$.
- Falls $\omega_1 \notin A_1$, so ist $\omega_1(A_1 \times A_2) = \varnothing \in \mathscr{A}_2$.

Folglich gilt: $\sigma(S) = \mathscr{A}_1 \otimes \mathscr{A}_2 \subseteq \mathscr{A}$. Hiermit folgt der erste Teil der Behauptung. Für den ω_2-Schnitt führt man den Beweis analog.

Lemma 8.5. *Seien $(\Omega_1, \mathscr{A}_1, \mu_1)$ und $(\Omega_2, \mathscr{A}_2, \mu_2)$ σ-endliche Maßräume. Betrachte $Q \in \mathscr{A}_1 \otimes \mathscr{A}_2$. Dann ist die Funktion*

$$f_Q^1 \colon \Omega_1 \to [0, \infty], \quad \omega_1 \mapsto \mu_2(\omega_1 Q)$$

$(\mathscr{A}_1, \overline{\mathscr{B}})$-messbar und die Funktion

$$f_Q^2 \colon \Omega_2 \to [0, \infty], \quad \omega_2 \mapsto \mu_1(\omega_2 Q)$$

$(\mathscr{A}_2, \overline{\mathscr{B}})$-messbar.

Beweis. Sei zunächst μ_2 endlich. Betrachte das Mengensystem

$$\mathscr{D} := \left\{ D \in \mathscr{A}_1 \otimes \mathscr{A}_2 \,\middle|\, f^1_D \text{ ist } \mathscr{A}_1\text{-messbar} \right\}.$$

Wir prüfen zunächst nach, dass es sich dabei um ein Dynkin-System auf $\Omega := \Omega_1 \times \Omega_2$ handelt.

- $f^1_\Omega = \mu_2(\Omega_2)$ ist konstant und somit messbar. Folglich ist $\Omega \in \mathscr{D}$.
- Sei $D \in \mathscr{D}$. Es gilt: $f^1_{D^c} = f^1_\Omega - f^1_D$. Somit ist $D^c \in \mathscr{D}$.
- Seien $D_1, D_2, \ldots \in \mathscr{D}$ paarweise disjunkte Mengen. Für $D := \sum_{n\in\mathbb{N}} D_n$ gilt: $f^1_D = \sum_{n\in\mathbb{N}} f^1_{D_n}$. Somit ist $D \in \mathscr{D}$.

Außerdem gilt für $A_1 \times A_2 \in \mathscr{A}_1 \otimes \mathscr{A}_2$:

$$f^1_{A_1\times A_2} = \mu_2(A_2) \cdot \mathbb{1}_{A_1}.$$

Folglich ist

$$S := \{A_1 \times A_2 \mid A_1 \in \mathscr{A}_1, A_2 \in \mathscr{A}_2\} \subseteq \mathscr{D}.$$

Bekanntlich ist S ein $\cap$-stabiler Erzeuger von $\mathscr{A}_1 \otimes \mathscr{A}_2$. Daraus folgt:

$$\sigma(S) = \delta(S) \subseteq \mathscr{D} \subseteq \mathscr{A}_1 \otimes \mathscr{A}_2 = \sigma(S).$$

Somit ist die Behauptung für endliche Maße gezeigt.

Sei nun μ_2 σ-endlich. Demnach gibt es eine Folge $(B_n)_{n\in\mathbb{N}}$ von Mengen aus $\mathscr{A}_2$ mit $B_n \nearrow \Omega_2$ und $\mu_2(B_n) < \infty$ für alle $n \in \mathbb{N}$. Wir erhalten endliche Maße

$$\mu_{2,n} \colon \mathscr{A}_2 \to [0, \infty), \quad A_2 \mapsto \mu_2(A_2 \cap B_n).$$

Somit sind für $Q \in \mathscr{A}_1 \otimes \mathscr{A}_2$ die Funktionen $\omega_1 \mapsto \mu_{2,n}(\omega_1 Q)$ $\mathscr{A}_1$-messbar. Weiterhin gilt wegen der Stetigkeit von unten:

$$\sup_{n\in\mathbb{N}} \mu_{2,n}(\omega_1 Q) = \mu_2(\omega_1 Q).$$

Damit ist nach Lemma 2.34 auch f^1_Q $\mathscr{A}_1$-messbar.

Für f^2_Q folgt die Aussage analog.

Satz 8.6. *Seien $(\Omega_1, \mathscr{A}_1, \mu_1)$ und $(\Omega_2, \mathscr{A}_2, \mu_2)$ σ-endliche Maßräume. Dann gelten folgende Aussagen:*

1. *Es gibt genau ein σ-endliches Maß μ auf dem Messraum $(\Omega_1 \times \Omega_2, \mathscr{A}_1 \otimes \mathscr{A}_2)$ mit*

$$\mu(A_1 \times A_2) = \mu_1(A_1) \cdot \mu_2(A_2) \quad \text{für alle } A_1 \times A_2 \in \mathscr{A}_1 \otimes \mathscr{A}_2. \tag{8.2}$$

2. *Für alle $Q \in \mathscr{A}_1 \otimes \mathscr{A}_2$ gilt:*

$$\mu(Q) = \int \mu_2(\omega_1 Q)\, \mu_1(\mathrm{d}\omega_1) = \int \mu_1(\omega_2 Q)\, \mu_2(\mathrm{d}\omega_2).$$

μ heißt das Produktmaß *von μ_1 und μ_2. Wir schreiben: $\mu = \mu_1 \otimes \mu_2$.*

Beweis. Aus Lemma 8.5 folgt, dass für alle $Q \in \mathscr{A}_1 \otimes \mathscr{A}_2$ die Abbildung

$$f_Q^1 \colon \Omega_1 \to [0,\infty], \quad \omega_1 \mapsto \mu_2(\omega_1 Q)$$

messbar ist. Betrachte die Abbildung

$$\mu \colon \mathscr{A}_1 \otimes \mathscr{A}_2 \to [0,\infty], \quad Q \mapsto \int \mu_2(\omega_1 Q)\, \mu_1(\mathrm{d}\omega_1).$$

Es gilt:

- $\mu(\varnothing) = \int \mu_2(\varnothing)\, \mathrm{d}\mu_1 = 0.$
- Seien $Q_1, Q_2, \ldots \in \mathscr{A}_1 \otimes \mathscr{A}_2$ paarweise disjunkt und $Q := \sum_{n\in\mathbb{N}} Q_n$. Dann gilt wegen monotoner Konvergenz:

$$\begin{aligned}\mu(Q) = \int \mu_2\left(\omega_1 \sum_{n\in\mathbb{N}} Q_n\right) \mu_1(\mathrm{d}\omega_1) &= \int \mu_2\left(\sum_{n\in\mathbb{N}} \omega_1 Q_n\right) \mu_1(\mathrm{d}\omega_1) \\ &= \sum_{n\in\mathbb{N}} \int \mu_2\left(\omega_1 Q_n\right) \mu_1(\mathrm{d}\omega_1) = \sum_{n\in\mathbb{N}} \mu(Q_n).\end{aligned}$$

Somit ist μ ein Maß auf $\mathscr{A}_1 \otimes \mathscr{A}_2$. Für alle $A_1 \times A_2 \in \mathscr{A}_1 \otimes \mathscr{A}_2$ gilt:

$$\mu(A_1 \times A_2) = \int \mu_2(A_2) \mathbb{1}_{A_1}(\omega_1)\, \mu_1(\mathrm{d}\omega_1) = \mu_1(A_1) \cdot \mu_2(A_2).$$

Analog können wir ein Maß

$$\tilde{\mu} \colon \mathscr{A}_1 \otimes \mathscr{A}_2 \to [0,\infty], \quad Q \mapsto \int \mu_1(\omega_2 Q)\, \mu_2(\mathrm{d}\omega_2)$$

definieren. Aus Satz 8.2 folgt $\tilde{\mu} = \mu$.

Beispiel 8.7 (Lebesgue-Maß). Für alle $k, s \in \mathbb{N}$ ist $\lambda^k \otimes \lambda^s = \lambda^{k+s}$. Dies folgt mit der Eindeutigkeit des Produktmaßes aus der Tatsache, dass λ^{k+s} die Gleichung (8.2) erfüllt.

Satz 8.8 (Tonelli)**.** *Seien $(\Omega_1, \mathscr{A}_1, \mu_1)$ und $(\Omega_2, \mathscr{A}_2, \mu_2)$ σ-endliche Maßräume und*

$$f \colon \Omega_1 \times \Omega_2 \to [0,\infty]$$

messbar numerisch. Definiere

$$\begin{aligned} f_{\omega_1} &\colon \Omega_2 \to [0,\infty], \quad \omega_2 \mapsto f(\omega_1,\omega_2), \\ f_{\omega_2} &\colon \Omega_1 \to [0,\infty], \quad \omega_1 \mapsto f(\omega_1,\omega_2). \end{aligned}$$

Dann ist die Funktion

$$f_1\colon \Omega_1 \to [0,\infty], \quad \omega_1 \mapsto \int f\omega_1 \,\mathrm{d}\mu_2$$

$(\mathscr{A}_1, \overline{\mathscr{B}})$*-messbar und die Funktion*

$$f_2\colon \Omega_2 \to [0,\infty], \quad \omega_2 \mapsto \int f\omega_2 \,\mathrm{d}\mu_1$$

$(\mathscr{A}_2, \overline{\mathscr{B}})$*-messbar und es gilt:*

$$\begin{aligned}\int f \,\mathrm{d}(\mu_1 \otimes \mu_2) &= \int \left(\int f(\omega_1, \omega_2)\, \mu_1(\mathrm{d}\omega_1) \right) \mu_2(\mathrm{d}\omega_2) \\ &= \int \left(\int f(\omega_1, \omega_2)\, \mu_2(\mathrm{d}\omega_2) \right) \mu_1(\mathrm{d}\omega_1). \end{aligned} \tag{8.3}$$

Beweis. Wir führen den Beweis mittels algebraischer Induktion. Definiere

$$(\Omega, \mathscr{A}, \mu) := (\Omega_1 \times \Omega_2, \mathscr{A}_1 \otimes \mathscr{A}_2, \mu_1 \otimes \mu_2).$$

Sei zunächst $f \in \mathscr{E}(\Omega)$ mit Normaldarstellung

$$f = \sum_{i=1}^{n} \alpha_i \mathbb{1}_{Q_i}.$$

So gilt:

$$f_{\omega_2}(\omega_1) = \sum_{i=1}^{n} \alpha_i \mathbb{1}_{\omega_2 Q_i}(\omega_1)$$

und folglich:

$$f_2(\omega_2) = \int f(\omega_1, \omega_2)\, \mu_1(\mathrm{d}\omega_1) = \sum_{i=1}^{n} \alpha_i \mu_1(\omega_2 Q_i).$$

Nach Satz 8.6 ist f_2 somit $\mathscr{A}_2$-messbar und Integration nach μ_2 liefert:

$$\int f_2 \,\mathrm{d}\mu_2 = \sum_{i=1}^{n} \alpha_i \mu(Q_i) = \int f \,\mathrm{d}\mu.$$

Seien nun $f \in \mathscr{E}^*(\Omega)$ und $(u^n)_{n\in\mathbb{N}}$ eine gegen f aufsteigende Folge in $\mathscr{E}(\Omega)$. Dies liefert eine Folge $(u^n_{\omega_2})_{n\in\mathbb{N}}$ in $\mathscr{E}(\Omega_1)$, die gegen f_{ω_2} aufsteigt. Wir erhalten eine Folge $(\phi_n)_{n\in\mathbb{N}}$ von $\mathscr{A}_2$-messbaren Funktionen

$$\phi_n\colon \Omega_2 \to [0,\infty), \quad \omega_2 \mapsto \int u^n_{\omega_2} \,\mathrm{d}\mu_1,$$

welche wie $(u^n_{\omega_2})_{n\in\mathbb{N}}$ isoton ist. Es gilt:

$$\sup_{n\in\mathbb{N}} \phi_n = f_2.$$

Also ist auch f_2 $\mathscr{A}_2$-messbar und nach dem Satz von der monotonen Konvergenz folgt:

$$\int f_2 \, \mathrm{d}\mu_2 = \sup_{n \in \mathbb{N}} \int \phi_n \, \mathrm{d}\mu_2.$$

Außerdem gilt:

$$\int \phi_n \, \mathrm{d}\mu_2 = \int u^n \, \mathrm{d}\mu.$$

Da $u^n \nearrow f$, folgt:

$$\int \phi_n \, \mathrm{d}\mu_2 \nearrow \int f \, \mathrm{d}\mu$$

und somit

$$\int f_2 \, \mathrm{d}\mu_2 = \int f \, \mathrm{d}\mu.$$

Durch analoge Vorgehensweise für f_{ω_1} folgt die Behauptung.

Satz 8.9 (Fubini)**.** *Seien $(\Omega_1, \mathscr{A}_1, \mu_1)$ und $(\Omega_2, \mathscr{A}_2, \mu_2)$ σ-endliche Maßräume und*

$$f \colon \Omega_1 \times \Omega_2 \to \overline{\mathbb{R}}.$$

$\mu_1 \otimes \mu_2$-integrierbar. Dann gilt:

1. *Die Abbildungen*
$$g_{\omega_1} \colon \Omega_2 \to \overline{\mathbb{R}}, \quad \omega_2 \mapsto f(\omega_1, \omega_2)$$
sind für μ_1-fast-alle ω_1 μ_2-integrierbar. Die Abbildungen
$$h_{\omega_2} \colon \Omega_1 \to \overline{\mathbb{R}}, \quad \omega_1 \mapsto f(\omega_1, \omega_2)$$
sind für μ_2-fast-alle ω_2 μ_1-integrierbar.

2. *Die Abbildung*
$$g \colon \Omega_1 \to \overline{\mathbb{R}}, \quad \omega_1 \mapsto \int g_{\omega_1} \, \mathrm{d}\mu_2$$
ist μ_1-integrierbar. Die Abbildung
$$h \colon \Omega_2 \to \overline{\mathbb{R}}, \quad \omega_2 \mapsto \int h_{\omega_2} \, \mathrm{d}\mu_1$$
ist μ_2-integrierbar.

und es gilt analog zu (8.3):

$$\int f \, \mathrm{d}(\mu_1 \otimes \mu_2) = \int g \, \mathrm{d}\mu_1 = \int h \, \mathrm{d}\mu_2. \tag{8.4}$$

Beweis. Definiere $\mu := \mu_1 \otimes \mu_2$.

1. Betrachte die Abbildung $|f|$. Diese ist nichtnegativ und messbar numerisch, weshalb wir den Satz von Tonelli anwenden können:
$$\int\int |g_{\omega_1}| \, \mathrm{d}\mu_2 \, \mu_1(\mathrm{d}\omega_1) = \int\int |h_{\omega_2}| \, \mathrm{d}\mu_1 \, \mu_2(\mathrm{d}\omega_2) = \int |f| \, \mathrm{d}\mu < \infty.$$
Somit ist
$$\int |g_{\omega_1}| \, \mathrm{d}\mu_2 < \infty \quad [\mu_1].$$
Also ist g_{ω_1} μ_2-integrierbar für μ_1-fast-alle ω_1. Analog folgt die Behauptung für h_{ω_2}.

2. Betrachte die Abbildungen f^+ und f^-. Diese sind nichtnegativ und messbar numerisch, weshalb wir den Satz von Tonelli anwenden können:
$$\begin{aligned}\int f \, \mathrm{d}\mu = \int f^+ \, \mathrm{d}\mu - \int f^- \, \mathrm{d}\mu &= \int\int g^+_{\omega_1} \mathrm{d}\mu_2 \, \mu_1(\mathrm{d}\omega_1) - \int\int g^-_{\omega_1} \mathrm{d}\mu_2 \, \mu_1(\mathrm{d}\omega_1) \\ &= \int g \, \mu_1(\mathrm{d}\omega_1).\end{aligned}$$
Analog folgt die Behauptung für h.

Satz 8.10. *Seien $(\Omega_i, \mathscr{A}_i, \mu_i)$ σ-endliche Maßräume für alle $i \in \{1, \ldots, n\}$. Dann gibt es genau ein Maß*
$$\mu \colon \left(\bigtimes_{i=1}^n \Omega_i, \bigotimes_{i=1}^n \mathscr{A}_i \right) \to [0, \infty]$$
mit
$$\mu\left(\bigtimes_{i=1}^n A_i \right) = \prod_{i=1}^n \mu_i(A_i) \quad \textit{für alle} \quad \bigtimes_{i=1}^n A_i \in \bigotimes_{i=1}^n \mathscr{A}_i.$$
Schreibe $\bigotimes_{i=1}^n \mu_i := \mu$. Wir definieren den Produktraum *durch*
$$\bigotimes_{i=1}^n (\Omega_i, \mathscr{A}_i, \mu_i) := \left(\bigtimes_{i=1}^n \Omega_i, \bigotimes_{i=1}^n \mathscr{A}_i, \bigotimes_{i=1}^n \mu_i \right).$$

[Übung]

Satz 8.11. *Seien $\bigotimes_{i=1}^n (\Omega_i, \mathscr{A}_i, \mu_i)$ ein Produktraum und*
$$f \colon \bigtimes_{i=1}^n \Omega_i \to [0, \infty]$$
$\bigotimes_{i=1}^n \mu_i$-integrierbar. Dann gilt für jede Permutation σ der Zahlen $\{1, \ldots, n\}$:
$$\int f \, \mathrm{d}\left(\bigotimes_{i=1}^n \mu_i \right) = \int \cdots \int f(\omega_1, \ldots, \omega_n) \, \mu_{\sigma(1)}(\mathrm{d}\omega_{\sigma(1)}) \cdots \mu_{\sigma(n)}(\mathrm{d}\omega_{\sigma(n)}).$$

Beweis. Nach dem Satz von Fubini können wir benachbarte Integrationen vertauschen. Da jede Permutation eine Verknüpfung von endlich vielen Transpositionen ist, folgt die Behauptung.

Beispiel 8.12 (Darstellung des Erwartungswertes). Sei X eine Zufallsvariable mit Verteilungsfunktion F. Der Erwartungswert von X existiert genau dann, wenn

$$\int_0^\infty \big(1 - F(x)\big)\,\mathrm{d}x < \infty \quad \text{und} \quad \int_{-\infty}^0 F(x)\,\mathrm{d}x < \infty.$$

In diesem Fall gilt:

$$\mathrm{E}X = \int_0^\infty \big(1 - F(x)\big)\,\mathrm{d}x - \int_{-\infty}^0 F(x)\,\mathrm{d}x.$$

Beweis. Definiere $A := \big\{(x,y) \in \mathbb{R}^2 \,\big|\, x \geqslant 0, 0 \leqslant y \leqslant x\big\}$. xA und yA bezeichnen dann den x- beziehungsweise y-Schnitt von A (siehe Definition 8.3). Nach dem Satz von Tonelli gilt:

$$\begin{aligned}\int_0^\infty \big(1 - F(y)\big)\,\mathrm{d}y &= \int_0^\infty P^X\big((y,\infty)\big)\,\lambda(\mathrm{d}y) = \int_0^\infty \int \mathbb{1}_{yA}(x)\,P^X(\mathrm{d}x)\,\lambda(\mathrm{d}y)\\ &= \int \mathbb{1}_A\,P^X \otimes \lambda = \int \mathbb{1}_{[0,\infty)}(x) \int \mathbb{1}_{xA}(y)\,\lambda(\mathrm{d}y)\,P^X(\mathrm{d}x)\\ &= \int \mathbb{1}_{[0,\infty)}(x)\lambda\big([0,x)\big)\,P^X(\mathrm{d}x) = \int_0^\infty x\,P^X(\mathrm{d}x).\end{aligned}$$

Analog gilt:

$$-\int_{-\infty}^0 F(y)\,\mathrm{d}y = \int_{-\infty}^0 x\,P^X(\mathrm{d}x). \quad \text{[Übung]}$$

Für den folgenden Satz ist es sinnvoll, sich die Definition des Lebesgue-Stieltjes-Integrals ins Gedächtnis zu rufen (siehe Bemerkung 7.31).

Satz 8.13 (Partielle Integration). *Seien $a, b \in \mathbb{R}$ mit $a < b$ und $F, G\colon \mathbb{R} \to \mathbb{R}$ maßdefinierende Funktionen ohne gemeinsame Unstetigkeitsstellen auf $(a,b]$. Sei μ das von F erzeugte Maß und sei ν das von G erzeugte Maß. Dann gilt:*

$$\int_{(a,b]} G\,\mathrm{d}\mu = F(b)G(b) - F(a)G(a) - \int_{(a,b]} F\,\mathrm{d}\nu.$$

Beweis. G und $G - G(a)$ erzeugen dasselbe Maß ν. Wir können ohne Einschränkung $F(a) = G(a) = 0$ annehmen, da gilt:

$$\begin{aligned}
\int_{(a,b]} G \, \mathrm{d}\mu &= F(b)G(b) - F(a)G(a) - \int_{(a,b]} F \, \mathrm{d}\nu \\
\Leftrightarrow \int_{(a,b]} G \, \mathrm{d}\mu - G(a)\big(F(b) - F(a)\big) &= F(b)\big(G(b) - G(a)\big) - \int_{(a,b]} F \, \mathrm{d}\nu \\
\Leftrightarrow \quad \int_{(a,b]} G \, \mathrm{d}\mu + \int_{(a,b]} -G(a) \, \mathrm{d}\mu &= F(b)\big(G(b) - G(a)\big) \\
&\quad + F(a)\big(G(a) - G(a)\big) - \int_{(a,b]} F \, \mathrm{d}\nu \\
\Leftrightarrow \quad \int_{(a,b]} G - G(a) \, \mathrm{d}\mu &= F(b)\big(G(b) - G(a)\big) \\
&\quad + F(a)\big(G(a) - G(a)\big) - \int_{(a,b]} F \, \mathrm{d}\nu.
\end{aligned}$$

Somit reicht es, für die Formel für die partielle Integration für $G - G(a)$ zu zeigen. Verfahre analog für $F - F(a)$. Definiere

$$A := \left\{(x, y) \in \mathbb{R}^2 \;\middle|\; a < y \leqslant x \leqslant b\right\}, \quad B := \left\{(x, y) \in \mathbb{R}^2 \;\middle|\; a < x \leqslant y \leqslant b\right\}.$$

Dann gilt:

$$A \cup B = (a, b]^2, \quad A \cap B = \left\{(x, x) \in \mathbb{R}^2 \;\middle|\; x \in (a, b]\right\}.$$

Mit dem Satz von Tonelli erhalten wir:

$$(\mu \otimes \nu)(A) = \int_{(a,b]} \nu\big((a, x]\big) \, \mu(\mathrm{d}x) = \int_{(a,b]} G \, \mathrm{d}\mu, \tag{8.5}$$

$$(\mu \otimes \nu)(B) = \int_{(a,b]} \mu\big((a, y]\big) \, \nu(\mathrm{d}y) = \int_{(a,b]} F \, \mathrm{d}\nu, \tag{8.6}$$

$$(\mu \otimes \nu)(A \cap B) = \int_{(a,b]} \nu\big(\{x\}\big) \, \mu(\mathrm{d}x). \tag{8.7}$$

Aus Lemma 7.44 und der Tatsache, dass F und G keine gemeinsamen Unstetigkeitsstellen haben, folgt, dass die Funktion

$$h \colon (a, b] \to \mathbb{R}, \quad x \mapsto \nu(\{x\})$$

μ-fast-überall verschwindet. Somit gilt in (8.7):

$$(\mu \otimes \nu)(A \cap B) = 0. \tag{8.8}$$

Folglich erhalten wir mit (8.5), (8.6) und (8.8):

$$\begin{aligned}
F(b)G(b) = \mu\big((a, b]\big)\nu\big((a, b]\big) &= (\mu \otimes \nu)\big((a, b]^2\big) \\
&= (\mu \otimes \nu)(A \cup B) = \int_{(a,b]} G \, \mathrm{d}\mu + \int_{(a,b]} F \, \mathrm{d}\nu.
\end{aligned}$$

Beispiel 8.14.

1. Betrachte das Integral von Dirichlet:

$$\lim_{t\to\infty}\int_0^t \frac{\sin x}{x}\,\mathrm{d}x = \frac{\pi}{2}. \qquad \text{[Übung]}$$

   Hinweis: Verwende den Satz von Fubini sowie

$$S(t) := \int_0^t \frac{\sin x}{x}\,\mathrm{d}x = \int_0^t \sin x \int_0^\infty \mathrm{e}^{-ux}\,\mathrm{d}u\,\mathrm{d}x. \tag{8.9}$$

2. Seien X und Y Zufallsvariablen mit Verteilungsfunktionen F_X beziehungsweise F_Y und F die Verteilungsfunktion von (X,Y). Dann gilt:

$$\mathrm{Cov}(X,Y) = \int\int \big(F(x,y) - F_X(x)F_Y(y)\big)\,\mathrm{d}x\mathrm{d}y.$$

   Beweis. Es gilt:

$$\mathrm{Cov}(X,Y) = \mathrm{E}\big((X-\mathrm{E}X)(Y-\mathrm{E}Y)\big) = \mathrm{E}(X\cdot Y) - \mathrm{E}X\cdot\mathrm{E}Y. \tag{8.10}$$

   Zunächst berechnen wir $\mathrm{E}(X\cdot Y)$:

$$\begin{aligned}
&\int_\Omega X\cdot Y\,\mathrm{d}P = \int_{\mathbb{R}^2} xy\,P^{(X,Y)}(\mathrm{d}(x,y))\\
&= \int_{\mathbb{R}^2}\left(\int_{\mathbb{R}}\big(\mathbb{1}_{[0,x]}(z_1) - \mathbb{1}_{[x,0]}(z_1)\big)\,\mathrm{d}z_1\right)\\
&\quad\cdot\left(\int_{\mathbb{R}}\big(\mathbb{1}_{[0,y]}(z_2) - \mathbb{1}_{[y,0]}(z_2)\big)\,\mathrm{d}z_2\right)P^{(X,Y)}(\mathrm{d}(x,y))\\
&= \int_{\mathbb{R}}\int_{\mathbb{R}}\int_{\mathbb{R}^2}\mathbb{1}_{[x,0]}(z_1)\cdot\mathbb{1}_{[y,0]}(z_2) - \mathbb{1}_{[x,0]}(z_1)\cdot\mathbb{1}_{[0,y]}(z_2) - \mathbb{1}_{[0,x]}(z_1)\cdot\mathbb{1}_{[y,0]}(z_2)\\
&\quad + \mathbb{1}_{[0,x]}(z_1)\cdot\mathbb{1}_{[0,y]}(z_2)\,P^{(X,Y)}(\mathrm{d}(x,y))\,\mathrm{d}z_1\mathrm{d}z_2\\
&= \int_{\mathbb{R}}\int_{\mathbb{R}}\mathbb{1}_{\{z_1<0,z_2<0\}}F(z_1,z_2) - \mathbb{1}_{\{z_1<0,z_2>0\}}\big(F_X(z_1) - F(z_1,z_2)\big)\\
&\quad - \mathbb{1}_{\{z_1>0,z_2<0\}}\big(F_Y(z_2) - F(z_1,z_2)\big)\\
&\quad + \mathbb{1}_{\{z_1>0,z_2>0\}}\big(1 - F_X(z_1) - F_Y(z_2) + F(z_1,z_2)\big)\,\mathrm{d}z_2\mathrm{d}z_2.
\end{aligned}$$

Mit Beispiel 8.12 erhalten wir:

$$\begin{aligned}
&\mathrm{E}X \cdot \mathrm{E}Y \\
&= \left(\int_0^\infty \big(1 - F_X(z_1)\big)\,\mathrm{d}z_1 - \int_{-\infty}^0 F_X(z_1)\,\mathrm{d}z_1 \right) \\
&\quad \cdot \left(\int_0^\infty \big(1 - F_Y(z_2)\big)\,\mathrm{d}z_2 - \int_{-\infty}^0 F_Y(z_2)\,\mathrm{d}z_2 \right) \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbb{1}_{\{z_1 > 0, z_2 > 0\}} \big(1 - F_X(z_1)\big)\big(1 - F_Y(z_2)\big) \\
&\quad - \mathbb{1}_{\{z_1 > 0, z_2 \leqslant 0\}} \big(1 - F_X(z_1)\big) F_Y(z_2) - \mathbb{1}_{\{z_1 \leqslant 0, z_2 > 0\}} F_X(z_1)\big(1 - F_Y(z_2)\big) \\
&\quad + \mathbb{1}_{\{z_1 \leqslant 0, z_2 \leqslant 0\}} F_X(z_1) F_Y(z_2)\,\mathrm{d}z_1 \mathrm{d}z_2.
\end{aligned}$$

Mit (8.10) folgt die Behauptung.

Problem 8.15. Seien $(\Omega_i, \mathscr{A}_i, P_i)$ Wahrscheinlichkeitsräume für alle $i \in \mathbb{N}$. Bisher haben wir nur Produkträume endlich vieler Maßräume betrachtet. Um auch Folgen von Zufallsvariablen untersuchen zu können, brauchen wir ein Wahrscheinlichkeitsmaß P auf einem Messraum $(\Omega, \mathscr{A})$ mit

$$P\left(A_1 \times \cdots \times A_n \times \bigtimes_{i=n+1}^{\infty} \Omega_i \right) = \prod_{i=1}^{n} P_i(A_i), \tag{8.11}$$

wobei

$$\Omega := \bigtimes_{i=1}^{\infty} \Omega_i := \{(\omega_i)_{i \in \mathbb{N}} \mid \omega_i \in \Omega_i \ \forall\, i \in \mathbb{N}\}$$

$$\mathscr{A} := \bigotimes_{i=1}^{\infty} \mathscr{A}_i := \sigma(\pi_i, i \in \mathbb{N}).$$

Satz 8.16 (Fortsetzungssatz)**.** *In der Situation von Problem 8.15 gelten die folgenden Aussagen:*

1. *Es gibt genau ein Wahrscheinlichkeitsmaß auf* $(\Omega, \mathscr{A})$*, welches (8.11) erfüllt.*
2. *Für alle* $n \in \mathbb{N}$ *und alle* $B \in \mathscr{A}_1 \otimes \cdots \otimes \mathscr{A}_n$ *gilt:*
$$P\left(B \times \bigtimes_{i=n+1}^{\infty} \Omega_i \right) = \left(\bigotimes_{i=1}^{n} P_i \right)(B).$$

P heißt Produktmaß *der* $(P_i)_{i \in \mathbb{N}}$*. Wir schreiben* $\bigotimes_{i=1}^{\infty} P_i := P$*. Der Wahrscheinlichkeitsraum*

$$\bigotimes_{i=1}^{\infty} (\Omega_i, \mathscr{A}_i, P_i) := \left(\bigtimes_{i=1}^{\infty} \Omega_i, \bigotimes_{i=1}^{\infty} \mathscr{A}_i, \bigotimes_{i=1}^{\infty} P_i \right)$$

heißt Produktraum.

Beweis. Für alle $n \in \mathbb{N}$ definiere die Algebren

$$\mathscr{A}_n^* := \alpha\left(\left\{A_1 \times \cdots \times A_n \times \bigtimes_{i=n+1}^{\infty} \Omega_i \,\middle|\, A_j \in \mathscr{A}_j \;\forall\, j \in \{1,\ldots,n\}\right\}\right)$$

sowie die Algebra

$$\begin{aligned} \mathscr{A}^* &:= \bigcup_{n\in\mathbb{N}} \mathscr{A}_n^* \\ &= \alpha\left(\left\{A_1 \times \cdots \times A_n \times \bigtimes_{i=n+1}^{\infty} \Omega_i \,\middle|\, n \in \mathbb{N} \text{ und } A_j \in \mathscr{A}_j \;\forall\, j \in \{1,\ldots,n\}\right\}\right). \end{aligned}$$

Auf $\mathscr{A}^*$ verhält sich P wie ein Produktmaß endlich vieler Wahrscheinlichkeitsräume und ist somit eine endlich-additive Mengenfunktion mit $P(\varnothing) = 0$. Wegen Lemma 1.29 und Satz 1.37 reicht es zu zeigen, dass P stetig von oben an $\varnothing$ ist. Betrachte $A_1 \supseteq A_2 \supseteq \cdots \in \mathscr{A}^*$ und $\varepsilon > 0$, sodass $\varepsilon \leqslant P(A_j)$ für alle $j \in \mathbb{N}$. Es ist zu zeigen: $\bigcap_{j\in\mathbb{N}} A_j \neq \varnothing$. Wir erinnern uns an die Definition des ω_1-Schnittes (Definition 8.3). Definiere

$$B_j := \left\{\omega_1 \in \Omega_1 \,\middle|\, \bigotimes_{i=2}^{\infty} P_i(\omega_1 A_j) > \frac{\varepsilon}{2}\right\}.$$

Für alle $j \in \mathbb{N}$ gilt:

$$\begin{aligned} \varepsilon \leqslant P(A_j) &= \int \bigotimes_{i=2}^{\infty} P_i(\omega_1 A_j)\, P_1(\mathrm{d}\omega_1) \\ &= \int_{B_j} \bigotimes_{i=2}^{\infty} P_i(\omega_1 A_j)\, P_1(\mathrm{d}\omega_1) + \int_{B_j^c} \bigotimes_{i=2}^{\infty} P_i(\omega_1 A_j)\, P_1(\mathrm{d}\omega_1) \leqslant P_1(B_j) + \frac{\varepsilon}{2}. \end{aligned}$$

Folglich ist $P_1(B_j) \geqslant \varepsilon/2$ für alle $j \in \mathbb{N}$. Da P_1 σ-additiv ist, ist es stetig von oben an $\varnothing$ und somit ist $\bigcap_{j\in\mathbb{N}} B_j \neq \varnothing$. Also gibt es ein $\omega_1' \in \bigcap_{j\in\mathbb{N}} B_j$, sodass für alle $j \in \mathbb{N}$ gilt:

$$\bigotimes_{i=2}^{\infty} P_i(\omega_1' A_j) \geqslant \frac{\varepsilon}{2}.$$

Für $\omega_1' A_1 \supseteq \omega_1' A_2 \supseteq \cdots \in \mathscr{A}^*$ können wir das gleiche Verfahren für $\bigtimes_{i=2}^{\infty} \Omega_i$ statt Ω, $(\omega_1' A_j)_{j\in\mathbb{N}}$ statt $(A_j)_{j\in\mathbb{N}}$ und $\varepsilon/2$ statt ε anwenden und erhalten ein $\omega_2' \in \Omega_2$ mit

$$\bigotimes_{i=3}^{\infty} P_i\big((\omega_1', \omega_2') A_j\big) \geqslant \frac{\varepsilon}{4}$$

für alle $j \in \mathbb{N}$. Iterativ erhalten wir eine Folge $(\omega_n')_{n\in\mathbb{N}}$ mit $\omega_n' \in \Omega_n$ für alle $n \in \mathbb{N}$ und

$$\bigotimes_{i=n+1}^{\infty} P_i\big((\omega_1', \ldots, \omega_n') A_j\big) \geqslant \frac{\varepsilon}{2^n}$$

für alle $j \in \mathbb{N}$. Wir zeigen nun:

$$(\omega'_k)_{k\in\mathbb{N}} \in \bigcap_{j=1}^{\infty} A_j.$$

Für $j \in \mathbb{N}$ betrachte $A_j \in \mathscr{A}^*$ und wähle $n \in \mathbb{N}$, sodass $A_j \in \mathscr{A}_n^*$. Aus

$$\bigotimes_{i=n+1}^{\infty} P_i\big((\omega'_1, \ldots, \omega'_n)A_j\big) > 0$$

folgt, dass $(\omega'_1, \ldots, \omega'_n)A_j$ nicht leer ist. Somit gibt es ein $(\omega_k)_{k\in\mathbb{N}} \in (\omega'_1, \ldots, \omega'_n)A_j$ mit $(\omega_1, \ldots, \omega_n) = (\omega'_1, \ldots, \omega'_n)$. Wegen $A_j \in \mathscr{A}_n^*$ folgt $(\omega'_k)_{k\in\mathbb{N}} \in A_j$ für alle $j \in \mathbb{N}$ wie behauptet.

Übungsaufgaben:

Übung 1: Beweisen Sie Satz 8.10.

Übung 2: Vervollständigen Sie den Beweis von Beispiel 8.12.

Übung 3: Beweisen Sie Punkt 1 von Beispiel 8.14.

Lernziel-Kontrolle:

Frage 1: Wozu dient die Betrachtung von Produktmaßen?

Frage 2: Was besagt der Satz von Fubini und worin besteht seine besondere Bedeutung?

Frage 3: Welche praktische Bedeutung haben unendliche Produkträume?

Anhang A

Mathematische Hilfsmittel

Axiom A.1 (Auswahlaxiom)**.** *Sei M ein Mengensystem und sei*

$$N := \bigcup_{A \in M} A.$$

Es gibt eine Abbildung $\Phi\colon M \to N$ mit $\Phi(A) \in A$ für alle $A \in M$.

Satz A.2 (Darstellungssatz von Riesz)**.** *Sei H ein Hilbertraum und H' der Raum der beschränkten linearen Funktionale. Dann ist die Abbildung*

$$\phi\colon H \to H', \quad x \mapsto \langle x, \cdot \rangle$$

eine komplex-konjugiert lineare Isometrie.

Für einen Beweis von Satz A.2 siehe zum Beispiel [8], Seite 81.

Literaturverzeichnis

[1] Heinz Bauer. *Maß- und Integrationstheorie*. De Gruyter, 1998.

[2] Patrick Billingsley. *Probability and Measure*. Wiley & Sons, 3rd edition, 1995.

[3] Yuan Shih Chow and Henry Teicher. *Probability Theory*. Springer Verlag, 2003.

[4] Richard Mansfield Dudley. *Real Analysis and Probability*. Cambridge University Press, 2002.

[5] Olav Kallenberg. *Foundations of Modern Probability*. Springer Verlag, 2nd edition, 2002.

[6] Achim Klenke. *Wahrscheinlichkeitstheorie*. Springer Verlag, 2008.

[7] Ralph Tyrrell Rockafellar. *Convex Analysis*. Princeton University Press, 1972.

[8] Reinhold Meise und Dietmar Vogt. *Einführung in die Funktionalanalysis*. Friedr. Vieweg & Sohn Verlagsgesellschaft mbH, Braunschweig/Wiesbaden, 1992.

[9] Edwin Hewitt und Karl Stromberg. *Real and Abstract Analysis*. Springer Verlag, 1969.

[10] Peter Gänssler und Winfried Stute. *Wahrscheinlichkeitstheorie*. Springer Verlag, 1977.

[11] Stan Wagon. *The Banach-Tarsky Paradox*. Cambridge University Press, 1985.

[12] Dirk Werner. *Funktionalanalysis*. Springer Verlag, 5th edition, 2007.

Index

Die Maß- und Integrationstheorie ist ein unabdingbares Werkzeug für viele Bereiche der Mathematik, wie etwa die Funktionalanalysis, die reelle Analysis oder die Wahrscheinlichkeitstheorie. Sie sollte deshalb in der Mathematikausbildung im zweiten Studienjahr gelehrt werden. In diesem Skript wird eine knappe, aber vollständige Darstellung der wichtigsten Techniken, Resultate und Beispiele gegeben, wobei ein besonderer Akzent auf die für die Wahrscheinlichkeitstheorie relevanten Konzepte gelegt wird. Dies beinhaltet insbesondere eine ausführliche Behandlung des Lebesgueschen Maßintegrals, aber auch des Riemannschen Integralbegriffs, welcher für die stochastische Analysis von grundlegender Bedeutung ist.

GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

ISBN: 978-3-941875-85-2

Universitätsdrucke Göttingen