Vorlesungsskript zur Maßtheorie

Emil Wiedemann
Universität Ulm, WS 2020/21

Vorbemerkungen

Dies ist ein Skript zur Vorlesung über Maßtheorie, die ich erstmals im Wintersemester 2019/20 an der Universität Ulm gehalten habe. Die Vorlesung richtete sich an Studierende der Mathematik, Wirtschaftsmathematik und der mathematischen Biometrie im dritten Semester, und umfaßte zwei Semesterwochenstunden. Anders als an den meisten deutschen Universitäten ist die Maßtheorie in Ulm also nicht in eine verpflichtende vierstündige Vorlesung Analysis III integriert. Eine weitere Ulmer Besonderheit besteht in der Angst der Studierenden vor der Maßtheorie, die ich andernorts so nicht wahrgenommen habe und die sich wohl nur aus Tradition erklären läßt.

Das Grundproblem der Maßtheorie besteht, wie der Name sagt, in der *Messung* von Mengen. Jeder Teilmenge des $\mathbb{R}^3$ zum Beispiel sollte dabei ihr Volumen zugeordnet werden. Es stellt sich leider (vielleicht überraschenderweise) heraus, daß dies in dieser Allgemeinheit nicht möglich ist, weswegen wir spezielle Teilmengen des $\mathbb{R}^n$ betrachten müssen (die die sogenannte Borel-σ-Algebra bilden), die in konsistenter Weise meßbar sind, d.h. denen sinnvoll ein Volumen zuordbar ist. Dieser zunächst lästige Theorieaufwand führt uns dann aber zum Lebesgue-Integral mit seinen wundervollen Konvergenzeigenschaften, die in vielen Anwendungsfällen die Vertauschung von Limes und Integral erlauben.

Könnte man das Problem der Meßbarkeit von Teilmengen des $\mathbb{R}^n$ noch als rein akademisch abtun (da alle „vernünftigen“ Mengen Lebesgue-meßbar sind), wird die praktische Bedeutung der Meßbarkeit und damit der σ-Algebra in der Wahrscheinlichkeitstheorie offensichtlich: Offenbar hat die Aussage „Die Körpergröße einer zufällig ausgewählten Person liegt im Intervall $[174{,}13756879\,\mathrm{cm}; 174{,}13756880\,\mathrm{cm}]$“ keinen Sinn, da die Körpergröße niemals so genau bestimmt werden kann; im mathematischen Modell würde man ein solches Intervall daher nicht als meßbar betrachten. Im Kontext stochastischer Prozesse, wie sie etwa in der Finanzmathematik auftreten, hat die σ-Algebra die Bedeutung der zu einem Zeitpunkt *verfügbaren Information*: Das Ereignis „am 23.10. wird es regnen“, interpretiert als Teilmenge eines geeigneten Wahrscheinlichkeitsraums, ist am 22.10. nicht meßbar, am 24.10. allerdings schon.

Die genannten Beispiele verdeutlichen, daß es sich bei der Maßtheorie um eine zentrale Grundlage sowohl der Analysis als auch der Wahrscheinlichkeitstheorie und verwandter Gebiete handelt. Nicht umsonst gehört die Maßtheorie an jeder kontinentaleuropäischen Universität zur Grundausbildung im Mathematikstudium.

Die Maßtheorie ist weitgehend unabhängig von den Lehrveranstaltungen des ersten Studienjahres; lediglich Grundkenntnisse über Mengen, Abbildungen und reelle Zahlen, wie sie in Analysis I erworben wurden [**2**], werden vorausgesetzt. Dieses Skript folgt im wesentlichen dem Buch von Bauer [**1**].

Inhaltsverzeichnis

KAPITEL 1

Maßräume

Ein Maßraum ist ein Tripel $(\Omega, \mathcal{A}, \mu)$, wobei Ω eine Menge, $\mathcal{A}$ eine sogenannte σ-Algebra, die die zu messenden Teilmengen von Ω enthält, und μ ein Maß ist, also eine Abbildung, die jeder in $\mathcal{A}$ enthaltenen Teilmenge $A \subset \Omega$ eine nichtnegative reelle Zahl $\mu(A)$ (oder den Wert $+\infty$) zuordnet, die als das „Maß" (z.B. Länge, Flächeninhalt, Volumen, Wahrscheinlichkeit) von A interpretiert wird. Es sind also zunächst einmal die Begriffe „σ-Algebra" und „Maß" zu klären.

1.1. σ-Algebren

Definition 1.1 (σ-Algebra). Sei Ω eine Menge und $\mathcal{A} \subset \mathcal{P}(\Omega)$[1] eine Menge von Teilmengen von Ω. Dann heißt $\mathcal{A}$ *σ-Algebra* (auf Ω), falls gilt[2]:

(1) $\Omega \in \mathcal{A}$,
(2) $A \in \mathcal{A} \quad \Rightarrow \quad A^c \in \mathcal{A}$,
(3) Die abzählbare[3] Vereinigung von Elementen von $\mathcal{A}$ ist wieder in $\mathcal{A}$, d.h.

$$A_n \in \mathcal{A} \quad \forall n \in \mathbb{N} \quad \Rightarrow \quad \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}.$$

Die Elemente von $\mathcal{A}$ heißen *meßbare Mengen*.

Beispiel 1.2. (1) Auf jeder Menge kann man zwei triviale σ-Algebren definieren, nämlich $\mathcal{A}_1 := \{\varnothing, \Omega\}$ sowie $\mathcal{A}_2 := \mathcal{P}(\Omega)$. Man sagt, $\mathcal{A}_1$ sei die *gröbste* und $\mathcal{A}_2$ die *feinste* σ-Algebra auf Ω. Für das Studium endlicher Wahrscheinlichkeitsräume verwendet man meist $\mathcal{A}_2 = \mathcal{P}(\Omega)$ als σ-Algebra; ist etwa $\Omega = \{1,2,3,4,5,6\}$, so kann man Mengen wie $\{2\}$, $\{4,5,6\}$ oder $\{1,6\}$ sinnvoll eine Wahrscheinlichkeit zuordnen (denken Sie an den Wurf eines Würfels).

(2) Für den Würfelwurf könnte man aber auch $\mathcal{A} := \{\varnothing, \{1,3,5\}, \{2,4,6\}, \{1,2,3,4,5,6\}\}$ als σ-Algebra wählen; dann wären nur gerade und ungerade Seiten unterscheidbar, nicht aber die genaue Augenzahl. Ein solches Modell wäre sinnvoll, wenn die Seiten des Würfels nicht durchnumeriert, sondern nur in zwei unterschiedlichen Farben eingefärbt wären.

(3) Sei Ω eine Menge, dann ist

$$\mathcal{A} := \{A \subset \Omega : A \text{ oder } A^c \text{ abzählbar}\}$$

eine σ-Algebra: Die ersten beiden Axiome sind offensichtlich erfüllt, für das dritte seien $A_i \in \mathcal{A}$ für $i \in \mathbb{N}$. Existiert dann $i \in \mathbb{N}$, sodaß A_i^c abzählbar, so ist $\left(\bigcup_{j=1}^{\infty} A_j\right)^c \subset A_i^c$ abzählbar; ist hingegen A_i^c überabzählbar für alle $i \in \mathbb{N}$, so sind alle A_i nach Definition von $\mathcal{A}$ abzählbar, und die abzählbare Vereinigung abzählbarer Mengen ist selbst abzählbar.

[1] Wie in **[2]** bezeichnen wir mit $\mathcal{P}(\Omega)$ die Potenzmenge von Ω, also die Menge aller Teilmengen von Ω.
[2] Wie üblich meint A^c das Komplement von A in Ω, also $A^c = \Omega \setminus A$.
[3] Wie in **[2]** bezeichnen wir eine Menge als abzählbar, wenn sie endlich oder abzählbar unendlich ist.

Aus den Axiomen einer σ-Algebra folgt unmittelbar, daß für eine σ-Algebra $\mathcal{A}$ stets $\varnothing \in \mathcal{A}$. Sind zudem $A_i \in \mathcal{A}$ für alle $i \in \mathbb{N}$, so ist auch $\bigcap_{i=1}^{\infty} A_i \in \mathcal{A}$, denn

$$\bigcap_{i=1}^{\infty} A_i = \left(\bigcup_{i=1}^{\infty} A_i^c\right)^c,$$

und die σ-Algebra ist stabil unter Komplementbildung und abzählbarer Vereinigung.

Eine besonders wichtige Klasse von σ-Algebren erhält man, indem man sie aus vorgegebenen Teilmengensystemen *erzeugt*. So wie der von gegebenen Vektoren erzeugte Vektorraum der kleinste Vektorraum ist, der diese Vektoren enthält, so ist die von Teilmengen von Ω erzeugte σ-Algebra die kleinste σ-Algebra, die diese Mengen enthält. Dies wird durch folgende Definition präzisiert:

Definition 1.3. Seien Ω eine Menge und $\mathcal{E} \subset \mathcal{P}(\Omega)$ eine Menge von Teilmengen von Ω. Dann heißt

$$\sigma(\mathcal{E}) := \bigcap\{\mathcal{A} : \mathcal{A}\ \sigma\text{-Algebra und } \mathcal{E} \subset \mathcal{A}\}$$

die von $\mathcal{E}$ *erzeugte* σ-Algebra.

Um die Wohldefiniertkeit der erzeugten σ-Algebra zu erkennen, benötigen wir

Proposition 1.4. *Durchschnitte von σ-Algebren sind wieder σ-Algebren.*

Beweis. Seien $\mathcal{A}_i$ σ-Algebren für jedes $i \in I$, wobei I eine beliebige Indexmenge ist. Da $\Omega \in \mathcal{A}_i$ für jedes $i \in I$, ist auch $\Omega \in \bigcap_{i\in I} \mathcal{A}_i$. Sei weiter $A \in \bigcap_{i\in I} \mathcal{A}_i$, dann ist $A \in \mathcal{A}_i$ für alle i, dann ist $A^c \in \mathcal{A}_i$ für alle i, und daher ist auch $A^c \in \bigcap_{i\in I} \mathcal{A}_i$. Seien schließlich $A_j \in \mathcal{A}_i$ für jedes $j \in \mathbb{N}$ und $i \in I$, so ist auch $\bigcup_{j=1}^{\infty} A_j \in \mathcal{A}_i$ für alle $i \in I$, also ist $\bigcup_{j=1}^{\infty} A_j \in \bigcap_{i\in I} \mathcal{A}_i$. $\square$

Beachte außerdem $\mathcal{E} \subset \sigma(\mathcal{E})$ für jede Familie $\mathcal{E}$ von Teilmengen von Ω (da der Durchschnitt in Definition 1.3 über eine nichtleere Menge gebildet wird, schließlich ist $\mathcal{P}(\Omega)$ eine σ-Algebra, die $\mathcal{E}$ als Teilmenge enthält). Außerdem hat $\sigma(\mathcal{E})$ per definitionem die folgende Minimalitätseigenschaft: Ist $\mathcal{A}$ eine σ-Algebra mit $\mathcal{E} \subset \mathcal{A}$, so ist $\sigma(\mathcal{E}) \subset \mathcal{A}$. Dies bestätigt unsere ursprüngliche Motivation, daß $\sigma(\mathcal{E})$ die kleinste (bzw. gröbste) σ-Algebra ist, die die Mengen aus $\mathcal{E}$ enthält.

1.2. Dynkin-Systeme

Definition 1.5. Sei Ω eine Menge und $\mathcal{D} \subset \mathcal{P}(\Omega)$ eine Menge von Teilmengen von Ω. Dann heißt $\mathcal{D}$ Dynkin-*System* (auf Ω), falls gilt:

(1) $\Omega \in \mathcal{D}$,
(2) $A \in \mathcal{D} \quad \Rightarrow \quad A^c \in \mathcal{D}$,
(3) Die abzählbare Vereinigung paarweise disjunkter Elemente von $\mathcal{D}$ ist wieder in $\mathcal{D}$, d.h.

$$A_n \in \mathcal{D} \quad \forall n \in \mathbb{N} \quad \wedge \quad A_i \cap A_j = \varnothing \quad \text{für } i \neq j \quad \Rightarrow \quad \bigcup_{n=1}^{\infty} A_n \in \mathcal{D}.$$

Offenbar ist jede σ-Algebra ein Dynkin-System. Die Umkehrung gilt nur für durchschnittsstabile Dynkin-Systeme: Ein Dynkin-System $\mathcal{D}$ heißt *durchschnittsstabil*, wenn aus $A, B \in \mathcal{D}$ folgt $A \cap B \in \mathcal{D}$.

Satz 1.6. Ein Dynkin-System ist eine σ-Algebra genau dann, wenn es durchschnittsstabil ist.

Beweis. Da jede σ-Algebra durchschnittsstabil ist, ist nur zu zeigen, daß ein durchschnittsstabiles Dynkin-System $\mathcal{D}$ sogar eine σ-Algebra ist. Die ersten beiden Axiome der

σ-Algebra sind nach Definition des DYNKIN-Systems erfüllt. Seien also $A_n \in \mathcal{D}$ für $n \in \mathbb{N}$ und setze für $n \in \mathbb{N} \cup \{0\}$

$$B_n := A_{n+1} \setminus \left(\bigcup_{i=1}^{n} A_i \right).$$

Dann ist $\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=0}^{\infty} B_i$. In der Tat, sei $x \in \bigcup_{i=1}^{\infty} A_i$ und n die kleinste natürliche Zahl mit $x \in A_n$, so ist $x \in B_{n-1} \subset \bigcup_{i=0}^{\infty} B_i$. Ist umgekehrt $x \in B_n$ für ein $n \in \mathbb{N} \cup \{0\}$, so ist $x \in A_{n+1}$.

Es genügt also zu zeigen $\bigcup_{i=0}^{\infty} B_i \in \mathcal{D}$. Beachte zunächst, daß die B_n paarweise disjunkt sind: Ist nämlich $m < n$ und $x \in B_n$, so ist nach Definition $x \notin B_m$. Nach dem dritten Axiom des DYNKIN-Systems genügt es also zu zeigen, daß $B_n \in \mathcal{D}$ für alle $n \in \mathbb{N} \cup \{0\}$.

Dazu zeigen wir zunächst, daß ein durchschnittsstabiles DYNKIN-System mit zwei Mengen auch deren Differenz und Vereinigung enthält. Seien also $C, D \in \mathcal{D}$. Dann ist

$$C \setminus D = C \cap D^c,$$

und $C \setminus D \in \mathcal{D}$ folgt aufgrund der Durchschnittsstabilität und der Stabilität unter Komplementbildung. Außerdem ist $C \cup D \in \mathcal{D}$ wegen $C \cup D = C \cup (D \setminus C)$, wobei C und $D \setminus C$ disjunkt sind.

Daraus folgt nun aber sofort, daß $B_n \in \mathcal{D}$ für alle $n \in \mathbb{N} \cup \{0\}$. □

Wie für σ-Algebren kann man zu jeder Teilmenge $\mathcal{E} \subset \mathcal{P}(\Omega)$ den Durchschnitt aller DYNKIN-Systeme bilden, die $\mathcal{E}$ enthalten (vgl. Definition 1.3). Da $\mathcal{P}(\Omega)$ ein solches DYNKIN-System ist, ist insbesondere $\mathcal{E}$ ein Element dieses Durchschnitts. Wie in Proposition 1.4 zeigt man, daß dieser Durchschnitt selbst ein DYNKIN-System ist, das als das von $\mathcal{E}$ erzeugte DYNKIN-System bezeichnet wird. Man schreibt dafür $\delta(\mathcal{E})$.

SATZ 1.7. Sei Ω eine Menge und $\mathcal{E}$ eine durchschnittsstabile Teilmenge von $\mathcal{P}(\Omega)$, d.h. aus $A, B \in \mathcal{E}$ folgt $A \cap B \in \mathcal{E}$. Dann gilt $\delta(\mathcal{E}) = \sigma(\mathcal{E})$.

BEWEIS. Da jede σ-Algebra ein DYNKIN-System ist, ist $\delta(\mathcal{E}) \subset \sigma(\mathcal{E})$. Für die andere Inklusion genügt es zu zeigen, daß $\delta(\mathcal{E})$ eine σ-Algebra ist, denn dann nimmt $\delta(\mathcal{E})$ an der Durchschnittsbildung in der Definition von $\sigma(\mathcal{E})$ teil. Nach Satz 1.6 genügt es dafür, die Durchschnittsstabilität von $\delta(\mathcal{E})$ zu zeigen.

Betrachte dazu ein beliebiges Element $D \in \delta(\mathcal{E})$ sowie die Menge

$$\mathcal{D}_D := \{A \in \mathcal{P}(\Omega) : A \cap D \in \delta(\mathcal{E})\}.$$

Wir behaupten, daß $\mathcal{D}_D$ ein DYNKIN-System ist, das $\mathcal{E}$ enthält; dann nämlich sind wir fertig, da daraus $\delta(\mathcal{E}) \subset \mathcal{D}_D$ folgt, was (da D beliebig war) genau die Durchschnittsstabilität von $\delta(\mathcal{E})$ besagt.

Zunächst haben wir $\Omega \in \mathcal{D}_D$, da $\Omega \cap D = D \in \delta(\mathcal{E})$. Ist $A \in \mathcal{D}_D$, so ist $A \cap D \in \delta(\mathcal{E})$. Aber

$$A^c \cap D = D \setminus (A \cap D) = D \cap (A \cap D)^c = (D^c \cup (A \cap D))^c \in \delta(\mathcal{E}),$$

da D^c und $A \cap D$ disjunkt sind.

Seien außerdem $A_n \in \mathcal{D}_D$ für alle $n \in \mathbb{N}$ und A_n paarweise disjunkt. Dann gilt

$$\left(\bigcup_{n=1}^{\infty} A_n \right) \cap D = \bigcup_{n=1}^{\infty} (A_n \cap D) \in \delta(\mathcal{E}),$$

da $A_n \cap D \in \delta(\mathcal{E})$ und die Mengen $A_n \cap D$ paarweise disjunkt sind.

Es ist noch zu zeigen $\mathcal{E} \subset \mathcal{D}_D$. Beachte dazu zuerst $\mathcal{E} \subset \mathcal{D}_E$ für jedes $E \in \mathcal{E}$ aufgrund der Durchschnittsstabilität von $\mathcal{E}$. Da $\mathcal{D}_E$ DYNKIN ist, folgt sogar $\delta(\mathcal{E}) \subset \mathcal{D}_E$ für jedes $E \in \mathcal{E}$. Das bedeutet, daß $E \cap D \in \delta(\mathcal{E})$ für alle $D \in \delta(\mathcal{E})$ und $E \in \mathcal{E}$; dies wiederum bedeutet ja gerade $\mathcal{E} \subset \mathcal{D}_D$, was zu zeigen war. □

1.3. Ringe

DEFINITION 1.8. Sei Ω eine Menge und $\mathcal{R} \subset \mathcal{P}(\Omega)$ eine Menge von Teilmengen von Ω. Dann heißt $\mathcal{R}$ *Ring* (auf Ω), falls gilt:

(1) $\varnothing \in \mathcal{R}$,
(2) $A, B \in \mathcal{R} \quad \Rightarrow \quad A \smallsetminus B \in \mathcal{R}$,
(3) Die endliche Vereinigung von Elementen von $\mathcal{R}$ ist wieder in $\mathcal{R}$, d.h.

$$A, B \in \mathcal{R} \quad \Rightarrow \quad A \cup B \in \mathcal{R}.$$

Offenbar ist jede σ-Algebra ein Ring: Sei nämlich $\mathcal{A}$ σ-Algebra, so folgt aus $\Omega \in \mathcal{A}$, daß auch $\Omega^c = \varnothing \in \mathcal{A}$; mit zwei Mengen ist auch deren Differenz in $\mathcal{A}$, denn $A \smallsetminus B = A \cap B^c \in \mathcal{A}$. Schließlich ist auch das dritte Ringaxiom erfüllt, denn eine σ-Algebra ist ja stabil unter abzählbaren, insbesondere also auch unter endlichen Vereinigungen.

BEISPIEL 1.9. Sei Ω eine Menge, dann ist

$$\mathcal{R} := \{A \in \mathcal{P}(\Omega) : A \text{ endlich}\}$$

ein Ring auf Ω, wie man sofort nachprüft. Allerdings ist $\mathcal{R}$ keine σ-Algebra, sofern Ω selbst nicht endlich ist.

Ein maßtheoretischer Ring kann als algebraischer Ring aufgefaßt werden (wie Sie ggf. in der Algebra lernen, ist dort ein Ring eine Struktur ähnlich einem Körper, nur daß in einem Ring keine multiplikative Inverse und noch nicht einmal ein neutrales Element der Multiplikation existieren muß). Dazu nimmt man als Addition die *symmetrische Differenz* $A \Delta B := (A \smallsetminus B) \cup (B \smallsetminus A)$ und als Multiplikation den Durchschnitt zweier Mengen.

1.4. Inhalte und Prämaße

Nachdem wir nun festgelegt haben, was gemessen werden soll (nämlich die in einer σ-Algebra enthaltenen Mengen), müssen wir uns überlegen, *wie* diese Mengen gemessen werden sollen. Wir benötigen dazu ein *Maß*. Als Vorstufe dazu definieren wir den Begriff des *Prämaßes*:

DEFINITION 1.10 (Inhalt und Prämaß). Seien Ω eine Menge und $\mathcal{R}$ ein Ring auf Ω. Eine Abbildung $\mu : \mathcal{R} \to \mathbb{R}_0^+ \cup \{+\infty\}$ heißt *Prämaß* auf $\mathcal{R}$, wenn gilt

(1) $\mu(\varnothing) = 0$,
(2) $\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$ für jede Familie $(A_n)_{n \in \mathbb{N}}$ von paarweise disjunkten Mengen aus $\mathcal{R}$, deren Vereinigung wieder in $\mathcal{R}$ liegt.[4]

Die Abbildung μ heißt *Inhalt*, wenn die zweite Bedingung für *endliche* disjunkte Vereinigungen gilt.

Es ist also jedes Prämaß ein Inhalt, aber nicht notwendigerweise umgekehrt.

BEMERKUNG 1.11. Die zweite Eigenschaft im Falle unendlicher Vereinigungen heißt *σ-Additivität* und, falls nur endliche Vereinigungen zugelassen sind, schlicht Additivität. Zur Bedeutung des Wertes $+\infty$: Man kann dies einfach als Symbol verstehen, für das vereinbart wird $+\infty + a = a + \infty = +\infty$ für alle $a \in \mathbb{R}$ und $+\infty + (+\infty) = \infty$. Setzen wir außerdem $\sum_{n=1}^{\infty} a_n = +\infty$ für jede Summe mit $a_n \in \mathbb{R}_0^+$ oder $a_n = +\infty$, falls mindestens ein $a_n = +\infty$ oder falls alle $a_n \in \mathbb{R}_0^+$ und die Reihe divergiert, so ist die σ-Additivität auch dann erklärt, wenn ein oder mehrere der Mengen A_n unendliches Prämaß haben. Beachte dabei, daß keine Ambiguitäten der Form $\infty - \infty$ auftreten können, da alle Terme nichtnegativ sind. Wir schreiben manchmal $\mu(A) < \infty$ statt $\mu(A) \in \mathbb{R}$.

[4] Es ist ja nicht notwendigerweise der Fall, daß die abzählbare Vereinigung von Mengen aus $\mathcal{R}$ wieder in $\mathcal{R}$ liegt.

Beispiel 1.12. (1) Seien Ω und $\mathcal{R}$ beliebig, so heißt für jedes $x \in \Omega$ die Abbildung $\delta_x : \mathcal{R} \to \mathbb{R}$,

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

das in x konzentrierte *Dirac-Maß*. Man überzeugt sich leicht davon, daß das Dirac-Maß ein Prämaß ist: Die erste Bedingung ist klar, für die zweite beachte, daß für paarweise disjunkte A_i das Element x entweder in genau einer oder in gar keiner der Mengen A_i enthalten ist; im ersten Falle sind beide Seiten in der σ-Additivätsbedingung eins, im zweiten Falle null.

(2) Sei Ω überabzählbar und

$$\mathcal{A} := \{A \subset \Omega : A \text{ oder } A^c \text{ abzählbar}\}$$

die aus Beispiel 1.2 bekannte σ-Algebra, die insbesondere ein Ring ist. Da Ω überabzählbar ist, ist für jedes $A \in \mathcal{A}$ genau eine der Mengen A und A^c abzählbar. Definiere für $A \in \mathcal{A}$ eine Abbildung durch $\mu(A) = 0$, falls A abzählbar, und $\mu(A) = 1$, falls A^c abzählbar. Dann ist μ ein Prämaß: Die erste Bedingung ist wieder klar; für die zweite seien die Mengen A_i aus $\mathcal{A}$ paarweise disjunkt. Sind alle A_i abzählbar, so auch ihre Vereinigung, und beide Seiten in der σ-Additivätsbedingung sind null. Ist mindestens ein A_i überabzählbar, so auch die Vereinigung aller A_i, und die linke Seite ist eins; Da für dieses i aber A_i^c abzählbar ist und (wegen der Disjunktheit) $A_j \subset A_i^c$ für jedes $j \neq i$, gilt $\mu(A_i) = 1$ und $\mu(A_j) = 0$ für jedes $j \neq i$, und somit ist auch die rechte Seite eins.

Proposition 1.13 (Eigenschaften von Inhalten). *Sei $\mathcal{R}$ ein Ring, μ ein Inhalt auf $\mathcal{R}$, und $A, B, A_n \in \mathcal{R}$. Dann gilt*

1. $\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B)$,
2. $A \subset B \implies \mu(A) \leq \mu(B)$,
3. $A \subset B \quad \wedge \quad \mu(A) < \infty \implies \mu(B \setminus A) = \mu(B) - \mu(A)$,
4. *für alle $n \in \mathbb{N}$ gilt* $\mu\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n \mu(A_i)$,
5. *für jede Familie $(A_n)_{n\in\mathbb{N}}$ von paarweise disjunkten Mengen aus $\mathcal{R}$, deren Vereinigung wieder in $\mathcal{R}$ liegt, gilt* $\sum_{n\in\mathbb{N}} \mu(A_n) \leq \mu\left(\bigcup_{n\in\mathbb{N}} A_n\right)$.

Beweis. (1) Mit den mengentheoretischen Identitäten

$$A \cup B = A \cup (B \setminus A) \quad \text{und} \quad B = (A \cap B) \cup (B \setminus A)$$

haben wir aufgrund der Additivität von μ

$$\mu(A \cup B) = \mu(A) + \mu(B \setminus A) \quad \text{und} \quad \mu(B) = \mu(A \cap B) + \mu(B \setminus A).$$

Die Behauptung ergibt sich nun aus der Addition dieser beiden Gleichungen, sofern $\mu(B \setminus A) < \infty$. Falls aber $\mu(B \setminus A) = +\infty$, so ist aber nach den eben verwendeten Identitäten $\mu(B) = \mu(A \cup B) = +\infty$, und somit gilt die Behauptung in der Form $+\infty = +\infty$.

(2) Falls $A \subset B$, so gilt $B = A \cup (B \setminus A)$ und daher aufgrund der Additivität $\mu(B) = \mu(A) + \mu(B \setminus A) \geq \mu(A)$.

(3) Dies folgt aus dem Beweis der vorhergehenden Behauptung durch Subtraktion von $\mu(A)$ auf beiden Seiten.

(4) Setze $B_1 := A_1$, $B_2 = A_2 \setminus A_1$, ..., $B_n = A_n \setminus \left(\bigcup_{i=1}^{n-1} A_i\right)$. Dann sind die B_i alle in $\mathcal{R}$ aufgrund der Ringaxiome, und sind nach Konstruktion paarweise disjunkt. Für jedes $n \in \mathbb{N}$ können wir wegen der Additivität also schließen

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

Andererseits ist $\bigcup_{i=1}^n B_i = \bigcup_{i=1}^n A_i$ und $B_i \subset A_i$, woraus mit (2) die Behauptung folgt.

(5) Sind nun die $A_i \in \mathcal{R}$ paarweise disjunkt, so ist wegen der Additivität für jedes $n \in \mathbb{N}$

$$\sum_{i=1}^n \mu(A_i) = \mu\left(\bigcup_{i=1}^n A_i\right) \leq \mu\left(\bigcup_{i=1}^\infty A_i\right),$$

wobei wir für die letzte Ungleichung wieder (2) benutzt haben. Die Behauptung folgt dann im Limes $n \to \infty$.

□

PROPOSITION 1.14. *Sei μ ein Inhalt auf einem Ring $\mathcal{R}$. Für jede absteigende Folge $A_1 \supset A_2 \supset \ldots \supset A_n \supset \ldots$ von Mengen $A_n \in \mathcal{R}$ mit $\bigcap_{n\in\mathbb{N}} A_n = \varnothing$ gelte, daß $\lim_{n\to\infty} \mu(A_n) = 0$. Dann ist μ sogar ein Prämaß.*

BEWEIS. Es ist nur die σ-Additivität zu zeigen. Seien dazu $(B_n)_{n\in\mathbb{N}}$ paarweise disjunkte Mengen aus $\mathcal{R}$, deren Vereinigung $B = \bigcup_{n\in\mathbb{N}} B_n$ ebenfalls in $\mathcal{R}$ liegt. Setze $A_n := \bigcup_{k=n+1}^\infty B_k$, so ist $A_n \supset A_{n+1}$ für jedes $n \in \mathbb{N}$, und $\bigcap_{n\in\mathbb{N}} A_n$ ist die Menge aller $x \in \Omega$, die in unendlich vielen B_k liegen. Da aber die B_k paarweise disjunkt sind, folgt $\bigcap_{n\in\mathbb{N}} A_n = \varnothing$.

Nach Voraussetzung gilt also $\lim_{n\to\infty} \mu(A_n) = 0$. Wegen der endlichen Additivität von μ ist andererseits

$$\mu(B \setminus A_n) = \mu\left(\bigcup_{k=1}^n B_k\right) = \sum_{k=1}^n \mu(B_k), \tag{1.1}$$

außerdem ist nach Proposition 1.13(3) – da $\lim_{n\to\infty} \mu(A_n) = 0$ und daher $\mu(A_n) < \infty$ für hinreichend große n – $\mu(B \setminus A_n) = \mu(B) - \mu(A_n)$ für hinreichend große n. Nehmen wir also in (1.1) auf beiden Seiten den Limes $n \to \infty$, folgt $\mu(B) = \sum_{k=1}^\infty \mu(B_k)$, und dies ist genau die gewünschte σ-Additivität. □

1.5. Maße

DEFINITION 1.15 (Maß). Sei Ω eine Menge und $\mathcal{A}$ eine σ-Algebra auf Ω. Ein Prämaß auf $\mathcal{A}$ heißt *Maß*. Ein Maß μ auf $\mathcal{A}$ heißt *endlich*, wenn $\mu(\Omega) < \infty$.

Aus Proposition 1.13(2) folgt sofort: Ist μ endlich, so gilt $\mu(A) < \infty$ für alle $A \in \mathcal{A}$.

BEISPIEL 1.16. (1) Die beiden Prämaße aus Beispiel 1.12 sind sogar Maße (sofern im ersten Beispiel $\mathcal{R}$ eine σ-Algebra ist).

(2) Sei Ω eine Menge versehen mit der σ-Algebra $\mathcal{A} = \mathcal{P}(\Omega)$, so definiert

$\mu(A)$ = Anzahl der Elemente von A

ein Maß, wie man nachprüfen möge (unendlichen Mengen wird dabei natürlich der Wert $+\infty$ zugeordnet). Man bezeichnet es als das *Zählmaß* auf Ω.

Äquivalent könnten wir definieren: Eine Abbildung $\mu : \mathcal{A} \to \mathbb{R}_0^+ \cup \{+\infty\}$ auf einer σ-Algebra $\mathcal{A}$ ist ein Maß, wenn es σ-additiv ist und $\mu(\varnothing) = 0$. Ein Tripel $(\Omega, \mathcal{A}, \mu)$, bestehend aus einer Menge, einer σ-Algebra auf dieser Menge, und einem Maß auf dieser σ-Algebra, heißt *Maßraum*.

Ein Maß, für das $\mu(\Omega) = 1$, heißt *Wahrscheinlichkeitsmaß* oder, wenn keine probabilistische Konnotation evoziert werden soll, *normiertes Maß*.

Eines der wichtigsten Beispiele für ein Maß ist das LEBESGUE-Maß auf $\mathbb{R}^n$, das wir im folgenden Kapitel konstruieren. Dann wird auch klarer werden, warum wir zunächst gesondert über Inhalte und Prämaße gesprochen haben.

KAPITEL 2

Das LEBESGUE-Maß

Bisher haben wir unsere Begriffe rein mengentheoretisch eingeführt. Dementsprechend kann jede beliebige Menge, sofern eine σ-Algebra und ein Maß darauf definiert werden, als Maßraum betrachtet werden. In diesem Kapitel wollen wir speziell auf der Menge der reellen Zahlen, und allgemeiner auf $\mathbb{R}^n$, ein Maß konstruieren, das die grundlegende Frage beantwortet, wie man für „möglichst viele" Teilmengen von $\mathbb{R}^n$ sinnvoll das n-dimensionale Volumen definieren kann.

2.1. Der LEBESGUE-Inhalt

Wir gehen vom simplen Axiom aus, daß das Volumen eines Quaders das Produkt seiner Seitenlängen sein soll.

Seien $a, b \in \mathbb{R}^n$ mit $a_i \leq b_i$ für alle $i = 1, \dots, n$. Dann heißt die Produktmenge $[a, b) := \Pi_{i=1}^n [a_i, b_i)$ *halboffener Quader*. Beachte, daß $[a, b) = \varnothing$, falls $a_i = b_i$ für ein $i \in \{1, \dots, n\}$. Bezeichne mit $\mathcal{R} \subset \mathcal{P}(\mathbb{R}^n)$ die Menge aller endlichen Vereinigungen halboffener Quader. (Die Bezeichnung suggeriert, daß $\mathcal{R}$ ein Ring sei; dies werden wir bald zeigen.)

LEMMA 2.1. *Seien $[a, b)$ und $[c, d)$ zwei halboffene Quader in $\mathbb{R}^n$.*

(1) *$[a, b) \cap [c, d)$ ist wieder ein halboffener Quader,*
(2) *$[c, d) \setminus [a, b) \in \mathcal{R}$,*
(3) *Jedes Element von $\mathcal{R}$ ist die endliche Vereinigung paarweise disjunkter halboffener Quader.*

BEWEIS. (1) Der halboffene Quader $[e, f)$ sei definiert durch $e_i := \max\{a_i, c_i\}$ und $f_i := \max\{e_i, \min\{b_i, d_i\}\}$ für jedes $i = 1, \dots, n$. Dann ist offenbar $[e, f) = [a, b) \cap [c, d)$.

(2) Da $[a, b) \cap [c, d)$ wieder ein Quader ist und $[c, d) \setminus [a, b) = [c, d) \setminus ([a, b) \cap [c, d)$, dürfen wir ohne Einschränkung annehmen $[a, b) \subset [c, d)$. Außerdem nehmen wir an $[a, b) \neq \varnothing$, da die Behauptung sonst trivial ist.

Für jedes $i \in \{1, \dots, n\}$ zerlegen wir das Intervall $[c_i, d_i)$ in die drei disjunkten Intervalle $[c_i, a_i)$, $[a_i, b_i)$, $[b_i, c_i)$ und setzen aus notationellen Gründen $x_i^1 := c_i$, $x_i^2 := a_i$, $x_i^3 := b_i$, $x_i^4 := d_i$. Dann gibt es genau 3^n halboffene Quader der Form $\Pi_{i=1}^n [x_i^{j(i)}, x_i^{j(i)+1})$, wobei für jedes i der Index $j = j(i)$ beliebig aus der Menge $\{1, 2, 3\}$ gewählt werden kann. Die (disjunkte) Vereinigung dieser Quader ist $[c, d)$, und der Quader $[a, b)$ ist genau $\Pi_{i=1}^n [x_i^2, x_i^3)$. Damit ist $[c, d) \setminus [a, b)$ die (disjunkte) endliche Vereinigung aller Teilquader außer $\Pi_{i=1}^n [x_i^2, x_i^3)$.

(3) Sei $A \in \mathcal{R}$, also läßt sich A schreiben als

$$A = \bigcup_{k=1}^m [a^k, b^k) = \bigcup_{k=1}^m \left[[a^k, b^k) \setminus \bigcup_{j=1}^{k-1} [a^j, b^j) \right],$$

wobei letztere Vereinigung über disjunkte Mengen der Gestalt $Q \setminus \bigcup_{j=1}^l Q_j$ gebildet wird mit halboffenen Quadern $Q, Q_1, \dots, Q_l$. Nun ist aber

$$Q \setminus \bigcup_{j=1}^l Q_j = \bigcap_{j=1}^l (Q \setminus Q_j).$$

Nach (2) ist jede der Mengen $Q \setminus Q_j$ endliche disjunkte Vereinigung halboffener Quader, $Q \setminus Q_j = \bigcup_{\alpha}^{M(j)} R_j^{\alpha}$, und daher[1]

$$Q \setminus \bigcup_{j=1}^{l} Q_j = \bigcap_{j=1}^{l} \bigcup_{\alpha=1}^{M} R_j^{\alpha} = \bigcup (R_1^{\alpha_1} \cap \ldots \cap R_l^{\alpha_l}),$$

wobei die Vereinigung über alle Tupel $(\alpha_1, \ldots, \alpha_l)$ mit $\alpha_j \in \{1, \ldots, M(j)\}$ läuft. Die Mengen unter der Vereinigung sind paarweise disjunkt, und sind nach (1) selbst wieder halboffene Quader. Damit ist alles gezeigt.

□

Satz 2.2. $\mathcal{R}$ ist ein Ring auf $\mathbb{R}^n$.

Beweis. Offenbar ist $\varnothing$ ein halboffener Quader und damit Element von $\mathcal{R}$. Außerdem ist mit zwei Elementen von $\mathcal{R}$, die ja endliche Vereinigungen halboffener Quader sind, auch die Vereinigung wieder in $\mathcal{R}$. Es bleibt also zu zeigen, daß mit $A, B \in \mathcal{R}$ auch $A \setminus B \in \mathcal{R}$.

Nun gibt es halboffene Quader Q_i, R_j, sodaß $A = \bigcup_{i=1}^{k} Q_i$ und $B = \bigcup_{j=1}^{m} R_j$. Also ist

$$A \setminus B = \bigcup_{i=1}^{k} \left(\bigcap_{j=1}^{m} (Q_i \setminus R_j) \right).$$

Es bleibt also zu zeigen, daß $\bigcap_{j=1}^{m}(Q_i \setminus R_j) \in \mathcal{R}$ für alle i und j. Nach Lemma 2.1(2) ist $Q_i \setminus R_j \in \mathcal{R}$. Also bleibt nur noch zu zeigen, daß der Durchschnitt zweier (und damit endlich vieler) Elemente von $\mathcal{R}$ wieder in $\mathcal{R}$ ist.

Aber der Durchschnitt zweier Vereinigungen von Quadern ist die Vereinigung von Durchschnitten dieser Quader[2], und diese Durchschnitte sind nach Lemma 2.1(1) wieder in $\mathcal{R}$. □

Satz 2.3 (Existenz und Eindeutigkeit des Lebesgue-Inhalts). Es existiert genau ein Inhalt λ auf $\mathcal{R}$, der jedem halboffenen Quader $\Pi_{i=1}^{n}[a_i, b_i)$ den Wert $\Pi_{i=1}^{n}(b_i - a_i)$ zuordnet. Dieser Inhalt ist reellwertig.

Dieser Inhalt heißt Lebesgue-*Inhalt*.

Beweis. Nach Lemma 2.1(3) kann jede Menge $A \in \mathcal{R}$ dargestellt werden als $A = \bigcup_{k=1}^{m} Q_k$, wobei jedes Q_k ein halboffener Quader ist und die Q_k paarweise disjunkt sind. Ist also λ ein Inhalt auf $\mathcal{R}$, der jedem Quader das Produkt seiner Seitenlängen zuordnet, so muß aufgrund der Additivität gelten

$$\lambda(A) = \sum_{k=1}^{m} \lambda(Q_k) = \sum_{k=1}^{m} \Pi_{i=1}^{n}(b_i^k - a_i^k) \tag{2.1}$$

(hier sind a_i^k, b_i^k die Randkoordinaten des k-ten Quaders). Da also jeder Inhalt auf $\mathcal{R}$ bereits durch seine Werte auf den Quadern eindeutig bestimmt ist, kann es höchstens einen Inhalt mit der geforderten Eigenschaft geben; dadurch ist bereits die Eindeutigkeit gezeigt.

Um zu zeigen, daß durch (2.1) tatsächlich ein Inhalt definiert ist, sind erst die Wohldefiniertheit und dann die beiden Inhaltsaxiome (leere Menge und Additivität) nachzuprüfen.

Zur Wohldefiniertheit: Ein Element $A \in \mathcal{R}$ kann ja im allgemeinen auf mehrere Arten in disjunkte Quader zerlegt werden, sodaß wir zunächst zeigen müssen, daß (2.1) für jede Zerlegung dasselbe Ergebnis liefert. Wir gehen in mehreren Schritten vor:

a) Bezeichne für einen Quader $Q = \Pi_{i=1}^{n}[a_i, b_i)$ die Zahl $\mu(Q) := \Pi_{i=1}^{n}(b_i - a_i)$ das Produkt der Seitenlängen. Sind für jedes $i = 1, \ldots, n$ Hyperebenen $\{x_i = c_{ij}\}$

[1] Wir verwenden hier das Distributivgesetz $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.

[2] Wir verwenden wieder das Distributivgesetz aus der vorigen Fußnote.

gegeben mit $j = 1, \dots, l(i)$, $a_i = c_{i1} < \dots < c_{il(i)} = b_i$, so zerfällt Q in disjunkte Quader der Gestalt

$$Q = \bigcup Q_k,$$

wobei Q_k das Produkt von Intervallen der Form $[c_{ij}, c_{i(j+1)})$ ist. Nun ist wegen $b_i - a_i = \sum_{j=1}^{l(i)-1}(c_{i(j+1)} - c_{ij})$

$$\mu(Q) = \prod_{i=1}^{n}(b_i - a_i) = \prod_{i=1}^{n} \sum_{j=1}^{l(i)-1} (c_{i(j+1)} - c_{ij}) = \sum \prod_{i=1}^{n}(c_{i(j_i+1)} - c_{ij_i}),$$

wobei die Summe über alle Tupel $(j_1, \dots, j_n)$ genommen wird mit $j_i \in \{1, \dots, l(i) - 1\}$ für alle $i = \{1, \dots, n\}$ (Distributivgesetz). Dies ist aber gerade die Summe über $\mu(Q_k)$ für alle durch k indizierten Teilquader, d.h. $\mu(Q) = \sum \mu(Q_k)$.[3]

b) Allgemeiner gilt: Ist ein Quader $Q = \bigcup Q_k$ dargestellt als *irgendeine* endliche disjunkte Vereinigung halboffener Quader, so gilt $\mu(Q) = \sum \mu(Q_k)$. In der Tat, für jedes k sei $Q_k := \prod_{i=1}^{n}[a_i^k, b_i^k)$, dann definieren für jedes $i = 1, \dots, n$ die Punkte a_i^k eine Zerlegung des Intervalls $[a_i, b_i)$. Durch Schneiden von Q mit allen Hyperebenen der Form $\{x_i = a_i^k\}$ entsteht eine disjunkte Zerlegung von Q wie in Schritt a) in Quader Q'_j; außerdem zerfällt jedes Q_k in einige der Quader Q'_j. Es gilt daher nach a)

$$\mu(Q) = \sum_j \mu(Q'_j) \quad \text{sowie} \quad \mu(Q_k) = \sum_{Q'_j \subset Q_k} \mu(Q'_j),$$

woraus sich die Behauptung ergibt.

c) Sei nun $A \in \mathcal{R}$ gegeben durch zwei unterschiedliche disjunkte Zerlegungen in Quader:

$$A = \bigcup_{j=1}^{M} Q_j = \bigcup_{l=1}^{N} S_l.$$

Dann können wir für jedes $j = 1, \dots M$ schreiben $Q_j = Q_j \cap A = \bigcup_{l=1}^{N}(Q_j \cap S_l)$, wobei diese Vereinigung disjunkt ist. Nach b) gilt daher

$$\mu(Q_j) = \sum_{l=1}^{N} \mu(Q_j \cap S_l).$$

Analog haben wir für alle $l = 1, \dots N$

$$\mu(S_l) = \sum_{j=1}^{M} \mu(Q_j \cap S_l).$$

Da aber $\sum_{j=1}^{M} \sum_{l=1}^{N} \mu(Q_j \cap S_l) = \sum_{l=1}^{N} \sum_{j=1}^{M} \mu(Q_j \cap S_l)$, folgt $\sum_{l=1}^{N} \mu(S_l) = \sum_{j=1}^{M} \mu(Q_j)$, und damit ist gezeigt, daß die Definition

$$\lambda(A) := \sum_{j=1}^{M} \mu(Q_j)$$

eine wohldefinierte reellwertige Abbildung bestimmt.

Es bleibt zu zeigen, daß λ ein Inhalt auf $\mathcal{R}$ ist. Es ist klar, daß $\lambda(\emptyset) = 0$, da die leere Menge als Quader aufgefaßt werden kann, der eine Seitenlänge gleich null hat. Zur Additivität: Sind A und B zwei disjunkte Mengen in $\mathcal{R}$, so können wir

---

[3] Dieser Beweisschritt ist zugegebenermaßen schwer verdaulich. Zeichnen Sie sich (für $n = 2$ oder $n = 3$) eine Skizze, um die Trivialität der Aussage zu erkennen.

schreiben $A = \bigcup_{i=1}^N Q_i$ und $B = \bigcup_{j=1}^M S_j$, wobei (Q_i) und (S_j) paarweise disjunkte Familien halboffener Quader sind. Daher ist

$$\lambda(A \cup B) = \lambda\left(\bigcup_{i=1}^N Q_i \cup \bigcup_{j=1}^M S_j\right) = \sum_{i=1}^N \mu(Q_i) + \sum_{j=1}^M \mu(S_j) = \lambda(A) + \lambda(B).$$

□

2.2. Das LEBESGUE-Prämaß

Hier zeigen wir, daß der LEBESGUE-Inhalt sogar ein Prämaß auf $\mathcal{R}$ ist. Es wird als (n-dimensionales) LEBESGUE-Prämaß bezeichnet.

SATZ 2.4. Mit den Bezeichnungen aus dem vorigen Abschnitt gilt: λ ist ein Prämaß auf $\mathcal{R}$.

BEWEIS. Es ist nur noch die σ-Additivität zu zeigen. Dafür genügt es nach Proposition 1.14 zu zeigen: Ist $(A_k) \subset \mathcal{R}$ eine absteigende Folge mit leerem Durchschnitt, so ist $\lim_{k\to\infty} \lambda(A_k) = 0$.

Sei also (A_k) eine absteigende Folge in $\mathcal{R}$. Angenommen, $\delta := \lim_{k\to\infty} \lambda(A_k) = \inf_{k\in\mathbb{N}} \lambda(A_k) > 0$. Wenn wir zeigen können, daß daraus $\bigcap_{k\in\mathbb{N}} A_k \neq \varnothing$ folgt, sind wir fertig.

Nun ist $A_k = \bigcup_{j=1}^m Q_j$ die endliche Vereinigung paarweise disjunkter Quader. Verkleinerung dieser Quader Q_j (z.B. ersetze man $Q_j = \prod_{i=1}^n [a_i, b_i)$ durch $Q_j' = \prod_{i=1}^n [a_i, b_i')$ mit $b_i' = b_i - \epsilon$ für hinreichend kleines $\epsilon > 0$) ergibt eine neue Menge $A_k' \in \mathcal{R}$ dergestalt, daß $\overline{A_k'} \subset A_k$[4] und daß

$$\lambda(A_k) - \lambda(A_k') < \frac{\delta}{2^k}. \tag{2.2}$$

Setze $B_k := \bigcap_{j=1}^k A_j'$, so ist (B_k) eine absteigende Folge in $\mathcal{R}$ mit $\overline{B_k} \subset \overline{A_k'} \subset A_k$.

Behauptung: $B_k \neq \varnothing$ für alle $k \in \mathbb{N}$.

Ist die Behauptung wahr, so ist $(\overline{B_k})$ eine absteigende Folge nichtleerer kompakter Mengen, und daraus folgt $\bigcap_{k=1}^\infty \overline{B_k} \neq \varnothing$[5] und somit auch $\bigcap_{k=1}^\infty A_k \neq \varnothing$, da $A_k \supset \overline{B_k}$. Dies aber war zu zeigen.

Beweis der Behauptung. Wir zeigen durch vollständige Induktion:

$$\lambda(B_k) \geq \lambda(A_k) - \delta(1 - 2^{-k}) \tag{2.3}$$

für alle $k \in \mathbb{N}$. Zunächst ist $B_1 = A_1'$, aber $\lambda(A_1') \geq \lambda(A_1) - \frac{\delta}{2}$ gemäß (2.2). Gelte nun (2.3) für ein k, so beachten wir zunächst $B_{k+1} = A_{k+1}' \cap B_k$ und daher nach Proposition 1.13(1)

$$\lambda(B_{k+1}) = \lambda(A_{k+1}') + \lambda(B_k) - \lambda(A_{k+1}' \cup B_k). \tag{2.4}$$

Nun ist nach Induktionsannahme $\lambda(B_k) \geq \lambda(A_k) - \delta(1 - 2^{-k})$, während nach (2.2) gilt $\lambda(A_{k+1}') \geq \lambda(A_{k+1}) - 2^{-k-1}\delta$; zudem haben wir nach Konstruktion $A_{k+1}' \cup B_k \subset A_k$. Zusammen ergibt sich damit aus (2.4)

$$\lambda(B_{k+1}) \geq \lambda(A_{k+1}) - 2^{-k-1}\delta + \lambda(A_k) - \delta(1 - 2^{-k}) - \lambda(A_k) = \lambda(A_{k+1}) - \delta(1 - 2^{-k-1}),$$

womit (2.3) gezeigt ist.

Aus (2.3) folgt aber unmittelbar die Behauptung, daß B_n nichtleer ist: Denn nach Annahme ist $\inf_{k\in\mathbb{N}} \lambda(A_k) = \delta$, also gilt $\lambda(B_k) \geq \delta - \delta(1 - 2^{-k}) = 2^{-k}\delta > 0$, also $B_k \neq \varnothing$. □

[4] Mit einem Querstrich bezeichnen wir wie in **[2]** den Abschluß einer Teilmenge des $\mathbb{R}^n$; in diesem Falle nehmen wir einfach die rechten Intervallenden b_i' zur Menge dazu.

[5] Denn wähle für jedes $k \in \mathbb{N}$ ein $x_k \in \overline{B_k}$; da insbesondere $x_k \in \overline{B_1}$, existiert eine konvergente Teilfolge $(x_{k_j})_{j\in\mathbb{N}}$ (Folgenkompaktheit!) mit Grenzwert x. Sei $N \in \mathbb{N}$ beliebig, dann liegen (wegen der absteigenden Bedingung) fast alle Glieder dieser Teilfolge in $\overline{B_N}$, und es folgt, da diese Menge abgeschlossen ist, $x \in \overline{B_N}$. Da N beliebig war, folgt $\bigcap_{k=1}^\infty \overline{B_k} \neq \varnothing$.

2.3. Das LEBESGUE-Maß

Der Ring der endlichen Vereinigungen halboffener Quader ist noch keine σ-Algebra (z.B. da $\mathbb{R}^n$ keine solche Vereinigung sein kann). Mithin ist also das LEBESGUE-Prämaß kein Maß. Wir zeigen nun aber, daß das LEBESGUE-Prämaß in eindeutiger Weise zu einem Maß fortgesetzt werden kann.

2.3.1. Der Fortsetzungssatz von CARATHÉODORY.

SATZ 2.5 (CARATHÉODORY). Sei Ω eine Menge, $\mathcal{R} \subset \mathcal{P}(\Omega)$ ein Ring und μ ein Prämaß auf $\mathcal{R}$. Dann existiert ein Maß $\tilde{\mu}$ auf $\sigma(\mathcal{R})$ (der von $\mathcal{R}$ erzeugten σ-Algebra), sodaß $\tilde{\mu} \restriction_{\mathcal{R}} = \mu$.

BEWEIS. *Schritt 1: Definition von $\tilde{\mu}$.* Sei A eine Teilmenge von Ω. Wir definieren $\tilde{\mu}(A)$ als das Infimum über alle Summen der Gestalt $\sum_{k=1}^{\infty} \mu(A_k)$, wobei $(A_k)_{k\in\mathbb{N}} \subset \mathcal{R}$ eine Familie ist, für die gilt $\bigcup_{k\in\mathbb{N}} A_k \supset A$. Existiert keine solche Familie, so setzen wir $\tilde{\mu}(A) = \infty$.

Schritt 2: Eigenschaften von $\tilde{\mu}$. Nach Definition ist $\tilde{\mu} \geq 0$. Weiter gilt

(1) $\tilde{\mu}(\varnothing) = 0$,
(2) $A_1 \subset A_2 \quad \Rightarrow \quad \tilde{\mu}(A_1) \leq \tilde{\mu}(A_2)$,
(3) $\tilde{\mu}\left(\bigcup_{k\in\mathbb{N}} B_k\right) \leq \sum_{k\in\mathbb{N}} \tilde{\mu}(B_k)$ für jede Familie $(B_k)_{k\in\mathbb{N}} \subset \mathcal{P}(\Omega)$.

In der Tat: Für (1) bemerken wir, daß $\varnothing = \bigcup_{k\in\mathbb{N}} \varnothing$ mit $\varnothing \in \mathcal{R}$; (1) folgt also aus $\mu(\varnothing) = 0$. Eigenschaft (2) folgt aus der einfachen Beobachtung, daß, falls $A_1 \subset A_2$, jede Überdeckung von A_2 durch Elemente von $\mathcal{R}$ automatisch auch A_1 überdeckt; das Infimum in der Definition von $\tilde{\mu}(A_1)$ ist also nicht größer als das in der Definition von $\tilde{\mu}(A_2)$. Schließlich zu (3): Ist ein $\tilde{\mu}(B_k) = \infty$, so ist die Aussage trivial. Seien also $\tilde{\mu}(B_k) < \infty$ für alle $k \in \mathbb{N}$. Dann existiert zu jedem $k \in \mathbb{N}$ und jedem $\epsilon > 0$ eine Familie $(C_m^k)_{m\in\mathbb{N}} \subset \mathcal{R}$, sodaß $B_k \subset \bigcup_{m\in\mathbb{N}} C_m^k$ und

$$\sum_{m\in\mathbb{N}} \mu(C_m^k) \leq \tilde{\mu}(B_k) + 2^{-k}\epsilon.$$

Nun gilt aber $\bigcup_{k\in\mathbb{N}} B_k \subset \bigcup_{k,m\in\mathbb{N}} C_m^k$ und daher

$$\tilde{\mu}\left(\bigcup_{k\in\mathbb{N}} B_k\right) \leq \sum_{k,m\in\mathbb{N}} \mu(C_m^k) \leq \sum_{k\in\mathbb{N}} \tilde{\mu}(B_k) + \epsilon,$$

wobei wir im letzten Schritt $\sum_{k\in\mathbb{N}} 2^{-k} = 1$ verwendet haben. Da $\epsilon > 0$ beliebig war, folgt (3).

Schritt 3: Verhalten von $\tilde{\mu}$ auf $\mathcal{R}$. Wir zeigen noch die folgenden beiden Eigenschaften:

$$\tilde{\mu}(A) \geq \tilde{\mu}(A \cap B) + \tilde{\mu}(A \cap B^c) \tag{2.5}$$

für alle $A \in \mathcal{P}(\Omega)$ und alle $B \in \mathcal{R}$; und

$$\tilde{\mu} \restriction_{\mathcal{R}} = \mu. \tag{2.6}$$

Zum Nachweis von (2.5) können wir wieder annehmen $\tilde{\mu}(A) < \infty$, da die Behauptung sonst trivial ist. Sei $(A_k) \subset \mathcal{R}$ eine Folge, die A überdeckt, und $B \in \mathcal{R}$; dann gilt wegen $A_k = (A_k \cap B) \cup (A_k \cap B^c)$ und der Additivität von μ

$$\sum_{k\in\mathbb{N}} \mu(A_k) = \sum_{k\in\mathbb{N}} \mu(A_k \cap B) + \sum_{k\in\mathbb{N}} \mu(A_k \cap B^c).$$

(Da alle Summanden nichtnegativ sind, ist die Umordnung der Summationsreihenfolge unproblematisch.) Nach Definition ist aber $\sum_{k\in\mathbb{N}} \mu(A_k \cap B) \geq \tilde{\mu}(A \cap B)$ (man beachte, daß der Schnitt zweier Elemente eines Rings wieder im Ring liegt) und ebenso $\sum_{k\in\mathbb{N}} \mu(A_k \cap B^c) \geq \tilde{\mu}(A \cap B^c)$; andererseits können für jedes $\epsilon > 0$ die A_k derart gewählt werden, daß $\sum_{k\in\mathbb{N}} \mu(A_k) \leq \tilde{\mu}(A) - \epsilon$. Es folgt (2.5).

Zum Beweis von (2.6) wähle man die Überdeckung $A = \bigcup_{k\in\mathbb{N}} A_k$ mit $A_1 = A$ und $A_k = \varnothing$ für $k \geq 2$; dies liefert bereits $\tilde{\mu}(A) \leq \mu(A)$. Für die entgegengesetzte Ungleichung beachte man: Ist $(A_k)_{k\in\mathbb{N}}$ eine Überdeckung von $A \in \mathcal{R}$ aus Elementen von $\mathcal{R}$, so ist $(A \cap A_k)_{k\in\mathbb{N}}$ ebenfalls eine solche Überdeckung, und es gilt

$$\mu(A) = \mu\left(\bigcup_{k\in\mathbb{N}} (A \cap A_k)\right) \le \sum_{k=1}^{\infty} \mu(A \cap A_k) \le \sum_{k=1}^{\infty} \mu(A_k),$$

und das Infimum auf beiden Seiten über alle zulässigen Überdeckungen liefert $\mu(A) \le \tilde{\mu}(A)$ wie gewünscht.

Schritt 4: (2.5) *definiert eine σ-Algebra.* Genauer gesagt: Sei

$$\mathcal{A} := \{B \in \mathcal{P}(\Omega) : \quad \tilde{\mu}(A) \ge \tilde{\mu}(A \cap B) + \tilde{\mu}(A \cap B^c) \quad \text{für alle } A \in \mathcal{P}(\Omega)\},$$

so ist $\mathcal{A}$ eine σ-Algebra. Die ersten beiden Axiome sind leicht zu überprüfen: Mit der Wahl $B = \Omega$ erhält man aus (2.5) die (wahre) Ungleichung $\tilde{\mu}(A) \ge \tilde{\mu}(A)$, also folgt $\Omega \in \mathcal{A}$. Da außerdem (2.5) invariant unter Vertauschung von B und B^c ist, ist mit B auch $B^c \in \mathcal{A}$.

Für das Weitere beachten wir, daß $\tilde{\mu}(A) \le \tilde{\mu}(A \cap B) + \tilde{\mu}(A \cap B^c)$ für alle $A, B \in \mathcal{P}(\Omega)$: Denn wenn wir in Eigenschaft (3) von Schritt 2 speziell $B_1 = A \cap B$, $B_2 = A \cap B^c$ und $B_k = \emptyset$ für alle $k \ge 3$ setzen, erhalten wir genau (2.5) mit umgedrehtem Ungleichheitszeichen.

Für die Abgeschlossenheit von $\mathcal{A}$ unter abzählbaren Vereinigungen gehen wir schrittweise vor: Zunächst betrachten wir endliche Vereinigungen. Seien also $B_1, B_2 \in \mathcal{A}$, sodaß insbesondere

$$\tilde{\mu}(A) = \tilde{\mu}(A \cap B_2) + \tilde{\mu}(A \cap B_2^c)$$

für alle $A \in \mathcal{P}(\Omega)$. Ersetzen wir darin A durch $A \cap B_1$ und dann durch $A \cap B_1^c$, so erhalten wir

$$\tilde{\mu}(A \cap B_1) = \tilde{\mu}(A \cap B_1 \cap B_2) + \tilde{\mu}(A \cap B_1 \cap B_2^c)$$

sowie

$$\tilde{\mu}(A \cap B_1^c) = \tilde{\mu}(A \cap B_1^c \cap B_2) + \tilde{\mu}(A \cap B_1^c \cap B_2^c).$$

Summation beider Gleichungen und Verwendung von $B_1 \in \mathcal{A}$ ergibt dann

$$\tilde{\mu}(A) = \tilde{\mu}(A \cap B_1 \cap B_2) + \tilde{\mu}(A \cap B_1 \cap B_2^c) + \tilde{\mu}(A \cap B_1^c \cap B_2) + \tilde{\mu}(A \cap B_1^c \cap B_2^c).$$

Ersetzen wir hierin A durch $A \cap (B_1 \cup B_2)$, so folgt

$$\tilde{\mu}(A \cap (B_1 \cup B_2)) = \tilde{\mu}(A \cap B_1 \cap B_2) + \tilde{\mu}(A \cap B_1 \cap B_2^c) + \tilde{\mu}(A \cap B_1^c \cap B_2). \tag{2.7}$$

Die letzten beiden Identitäten zusammen ergeben schließlich

$$\tilde{\mu}(A) = \tilde{\mu}(A \cap (B_1 \cup B_2)) + \tilde{\mu}(A \cap B_1^c \cap B_2^c),$$

also (unter Beachtung von $B_1^c \cap B_2^c = (B_1 \cup B_2)^c$) wie gewünscht $B_1 \cup B_2 \in \mathcal{A}$.

Als nächstes zeigen wir, daß $\mathcal{A}$ abgeschlossen ist unter abzählbaren, paarweise disjunkten Vereinigungen. Seien also $(B_k)_{k\in\mathbb{N}} \subset \mathcal{A}$ paarweise disjunkt. Aus (2.7) folgt, unter Verwendung der Disjunktheit,

$$\tilde{\mu}(A \cap (B_1 \cup B_2)) = \tilde{\mu}(A \cap B_1) + \tilde{\mu}(A \cap B_2),$$

und damit mittels vollständiger Induktion

$$\tilde{\mu}\left(A \cap \bigcup_{k=1}^{m} B_k\right) = \sum_{k=1}^{m} \tilde{\mu}(A \cap B_k) \tag{2.8}$$

für jedes $m \in \mathbb{N}$.

Nun setze $C_m := \bigcup_{k=1}^m B_k$, dann ist, wie bereits gezeigt, $C_m \in \mathcal{A}$ als endliche Vereinigung von Elementen aus $\mathcal{A}$. Also gilt für alle $A \in \mathcal{P}(\Omega)$ unter Berücksichtigung von (2.8)

$$\begin{aligned}\tilde{\mu}(A) &= \tilde{\mu}(A \cap C_m) + \tilde{\mu}(A \cap C_m^c) = \sum_{k=1}^m \tilde{\mu}(A \cap B_k) + \tilde{\mu}(A \cap C_m^c) \\ &\geq \sum_{k=1}^m \tilde{\mu}(A \cap B_k) + \tilde{\mu}\left(A \cap (\bigcup_{k\in\mathbb{N}} B_k)^c\right),\end{aligned}$$

wo wir im letzten Schritt $C_m \subset \bigcup_{k\in\mathbb{N}} B_k$ und Eigenschaft (2) aus Schritt 2 ausgenutzt haben. Da $m \in \mathbb{N}$ beliebig war, folgt sogar

$$\tilde{\mu}(A) \geq \sum_{k=1}^\infty \tilde{\mu}(A \cap B_k) + \tilde{\mu}\left(A \cap (\bigcup_{k\in\mathbb{N}} B_k)^c\right) \geq \tilde{\mu}\left(A \cap \bigcup_{k\in\mathbb{N}} B_k\right) + \tilde{\mu}\left(A \cap (\bigcup_{k\in\mathbb{N}} B_k)^c\right), \tag{2.9}$$

wobei wir im letzten Schritt Eigenschaft (3) aus Schritt 2 verwendet haben.

Damit ist gezeigt, daß $\mathcal{A}$ ein DYNKIN-System ist. Es ist sogar durchschnittsstabil, was wegen $B_1 \cap B_2 = (B_1^c \cup B_2^c)^c$ aus der Stabilität unter Komplementbildung und endlichen Vereinigungen folgt. Nach Satz 1.6 folgt, daß $\mathcal{A}$ eine σ-Algebra ist.

Schritt 5: Ende des Beweises. Wir haben gezeigt, daß $\mathcal{A}$ eine σ-Algebra ist, die $\mathcal{R}$ enthält; damit ist $\mathcal{A} \supset \sigma(\mathcal{R})$. Setzt man noch in (2.9) $A = \bigcup_{k\in\mathbb{N}} B_k$, so erhalten wir

$$\tilde{\mu}\left(\bigcup_{k\in\mathbb{N}} B_k\right) \geq \sum_{k\in\mathbb{N}} \tilde{\mu}(B_k).$$

Zusammen mit Eigenschaft (3) aus Schritt 2 ergibt sich die gewünschte σ-Additivität. Also ist $\tilde{\mu}$ ein Maß auf $\mathcal{A}$ (und damit auch auf $\sigma(\mathcal{R})$), und nach (2.6) ist die Einschränkung dieses Maßes auf $\mathcal{R}$ gerade μ, was zu zeigen war. ☐

2.3.2. Eindeutigkeit. Ein Inhalt μ auf einem Ring $\mathcal{R}$ heißt ***σ-endlich***, wenn es eine Folge $(A_k)_{k\in\mathbb{N}} \subset \mathcal{R}$ gibt mit $\bigcup_{k\in\mathbb{N}} A_n = \Omega$ und $\mu(A_k) < \infty$ für alle $k \in \mathbb{N}$. Man beachte, daß $\sup_{k\in\mathbb{N}} \mu(A_k)$ hierfür nicht endlich zu sein braucht. Insbesondere ist der LEBESGUE-Inhalt auf dem Ring der endlichen Vereinigungen halboffener Quader σ-endlich: Es ist etwa $\mathbb{R}^n = \bigcup_{k\in\mathbb{N}}[a_k, b_k)$ mit $a_k = (-k, \dots, -k)$, $b_k = (k, \dots, k)$, und es gilt $\mu([a_k, b_k)) = (2k)^n < \infty$.

SATZ 2.6 (Eindeutigkeit der Fortsetzung eines σ-endlichen Prämaßes). *Sei Ω eine Menge, $\mathcal{R} \subset \mathcal{P}(\Omega)$ ein Ring und μ ein σ-endliches Prämaß auf $\mathcal{R}$. Dann existiert genau ein Maß $\tilde{\mu}$ auf $\sigma(\mathcal{R})$, sodaß $\tilde{\mu} \restriction_{\mathcal{R}} = \mu$.*

BEWEIS. Die Existenz folgt aus dem eben gezeigten Satz von CARATHÉODORY. Seien also μ_1, μ_2 zwei Maße auf $\sigma(\mathcal{R})$, die auf $\mathcal{R}$ mit μ übereinstimmen. Wir wollen zeigen $\mu_1 = \mu_2$. Sei dazu $\mathcal{R}_{<\infty}$ die Menge aller $E \in \mathcal{R}$ mit $\mu_1(E) = \mu_2(E) < \infty$, und betrachte für ein $E \in \mathcal{R}_{<\infty}$ das System

$$\mathcal{D}_E := \{D \in \sigma(\mathcal{R}) : \mu_1(D \cap E) = \mu_2(D \cap E)\}.$$

Wir wollen zeigen, daß $\mathcal{D}_E$ DYNKIN ist: Nach Wahl von E gilt $\Omega \in \mathcal{D}_E$. Ist $D \in \mathcal{D}_E$, so gilt

$$\begin{aligned}\mu_1(E \cap D^c) &= \mu_1(E \setminus E \cap D) = \mu_1(E) - \mu_1(E \cap D) \\ &= \mu_2(E) - \mu_2(E \cap D) = \mu_2(E \setminus E \cap D) = \mu_2(E \cap D^c),\end{aligned}$$

also $D^c \in \mathcal{D}_E$. Seien außerdem $(D_k)_{k\in\mathbb{N}} \subset \mathcal{D}_E$ paarweise disjunkt, so gilt

$$\begin{aligned}\mu_1(E \cap \bigcup_{k\in\mathbb{N}} D_k) &= \mu_1(\bigcup_{k\in\mathbb{N}}(E \cap D_k)) = \sum_{k\in\mathbb{N}} \mu_1(E \cap D_k) \\ &= \sum_{k\in\mathbb{N}} \mu_2(E \cap D_k) = \mu_2(\bigcup_{k\in\mathbb{N}}(E \cap D_k)) = \mu_2(E \cap \bigcup_{k\in\mathbb{N}} D_k),\end{aligned}$$

wobei wir die σ-Additivität der Maße μ_1, μ_2 ausgenutzt haben. Also ist $\bigcup_{k\in\mathbb{N}} D_k \in \mathcal{D}_E$, und somit ist $\mathcal{D}_E$ DYNKIN.

Nun ist $\mathcal{R}$ als Ring durchschnittsstabil, und wegen $\mu_1 \restriction_{\mathcal{R}} = \mu_2 \restriction_{\mathcal{R}}$ ist $\mathcal{R} \subset \mathcal{D}_E$. Nach Satz 1.7 gilt also $\delta(\mathcal{R}) = \sigma(\mathcal{R})$ und damit, da $\mathcal{D}_E$ ein DYNKIN-System ist, das $\mathcal{R}$ enthält, $\sigma(\mathcal{R}) \subset \mathcal{D}_E$. Andererseits ist nach Definition $\mathcal{D}_E \subset \sigma(\mathcal{R})$, also insgesamt $\mathcal{D}_E = \sigma(\mathcal{R})$; das bedeutet: $\mu_1(E \cap A) = \mu_2(E \cap A)$ für alle $A \in \sigma(\mathcal{R})$.

Wegen der Voraussetzung der σ-Endlichkeit gibt es nun $(E_k)_{k\in\mathbb{N}} \subset \mathcal{R}_{<\infty}$ mit $\bigcup_{k\in\mathbb{N}} E_k = \Omega$. Wir setzen nun

$$F_1 := E_1, \quad F_k := E_k \smallsetminus \bigcup_{j=1}^{k-1} E_j \quad \text{für } k \geq 2.$$

Dann sind die F_k paarweise disjunkt, es gilt $F_k \subset E_k$, $F_k \in \mathcal{R}_{<\infty}$, und $\bigcup_{k\in\mathbb{N}} F_k = \Omega$. Nach dem eben Gezeigten ist also $\mu_1(F_k \cap A) = \mu_2(F_k \cap A)$ für alle $k \in \mathbb{N}$ und $A \in \sigma(\mathcal{R})$. (Denn $\sigma(\mathcal{R}) = \mathcal{D}_{F_k}$.)

Sei nun $A \in \sigma(\mathcal{R})$ beliebig, so ist $A = \bigcup_{k\in\mathbb{N}}(F_k \cap A)$ und daher wegen der σ-Additivität

$$\mu_1(A) = \sum_{k\in\mathbb{N}} \mu_1(F_k \cap A) = \sum_{k\in\mathbb{N}} \mu_2(F_k \cap A) = \mu_2(A).$$

□

2.3.3. LEBESGUE-Maß und BOREL-σ-Algebra. Wir befassen uns wieder mit dem Ring $\mathcal{R}$ der endlichen Vereinigungen halboffener Quader in $\mathbb{R}^n$.

DEFINITION 2.7. Sei $\mathcal{R}$ der Ring der endlichen Vereinigungen halboffener Quader in $\mathbb{R}^n$. Dann heißt $\mathcal{B} := \sigma(\mathcal{R})$ die BOREL-*σ-Algebra* auf $\mathbb{R}^n$. Die (gemäß Satz 2.6) eindeutig bestimmte Fortsetzung λ des LEBESGUE-Prämaßes auf $\mathcal{B}$ heißt (n-dimensionales) LEBESGUE-*Maß*.

Eine Menge $B \in \mathcal{B}$ heißt LEBESGUE-*meßbar*. Eine Menge $B \subset \mathbb{R}^n$ heißt LEBESGUE-*Nullmenge*, falls es eine meßbare Menge $B' \supset B$ gibt mit $\lambda(B') = 0$. Ist aus dem Kontext klar, daß mit dem LEBESGUE-Maß gearbeitet wird, sprechen wir einfach von meßbaren Mengen bzw. Nullmengen.

BEMERKUNG 2.8. In der Literatur findet man meist folgende Terminologie: Das hier eingeführte Maß wird als LEBESGUE-BOREL-Maß bezeichnet; der Begriff LEBESGUE-Maß dagegen ist für die eindeutig bestimmte *Vervollständigung* von λ reserviert. Man erinnere sich an die Übung: Ist $B \in \mathcal{B}$ eine Nullmenge und $A \subset B$, so würde man erwarten, daß auch A eine Nullmenge ist (so haben wir es ja soeben auch definiert); allerdings ist nicht klar, ob überhaupt $A \in \mathcal{B}$. Die Vervollständigung von λ ist dann diejenige Fortsetzung von λ, die Teilmengen von Nullmengen ebenfalls das Maß null zuordnet. Für Details siehe, wie gesagt, die Übungen. In der Praxis spielt die Unterscheidung zwischen LEBESGUE-BOREL-Maß und LEBESGUE-Maß keine Rolle.

Für jedes $B \in \mathcal{B}$ definiert $\mathcal{B}_B = \{A \cap B : A \in \mathcal{B}\}$ eine σ-Algebra, und $\lambda \restriction_{\mathcal{B}_B}$ ist ein Maß auf $\mathcal{B}_B$ (Übung). Mann nennt diese Einschränkung das LEBESGUE-*Maß auf* B.

SATZ 2.9. Die BOREL-σ-Algebra $\mathcal{B}$ wird von den offenen Teilmengen des $\mathbb{R}^n$ erzeugt, ebenso von den abgeschlossenen und von den kompakten. Insbesondere sind alle offenen, abgeschlossenen oder kompakten Mengen meßbar.

Dieser Satz stellt eine Verbindung zwischen Topologie und Meßbarkeit dar, die in der Theorie topologischer Maßräume verallgemeinert wird (siehe [**1**, Kapitel IV]).

BEWEIS. Zunächst zeigen wir, daß die offenen, abgeschlossenen und kompakten Mengen jeweils dieselbe σ-Algebra erzeugen. Seien dazu $\mathcal{A}_o$, $\mathcal{A}_a$ und $\mathcal{A}_k$ die entsprechenden σ-Algebren. Da jede kompakte Menge abgeschlossen ist, gilt $\mathcal{A}_k \subset \mathcal{A}_a$. Es ist aber jede abgeschlossene Teilmenge des $\mathbb{R}^n$ die abzählbare Vereinigung kompakter Mengen (denn für jede abgeschlossene Menge $A \subset \mathbb{R}^n$ ist $A = \bigcup_{k\in\mathbb{N}}(A \cap \overline{B_k(0)})$). Daraus folgt, daß jede σ-Algebra, die alle kompakten Mengen enthält, sogar alle abgeschlossenen Mengen enthält, und somit $\mathcal{A}_a = \mathcal{A}_k$. Außerdem sind die abgeschlossenen Mengen genau die Komplemente offener Mengen, und daher enthält jede σ-Algebra, die alle abgeschlossenen Mengen enthält, auch alle offenen Mengen und umgekehrt. Es folgt, wie behauptet, $\mathcal{A}_o = \mathcal{A}_a = \mathcal{A}_k$.

Es genügt also zu zeigen $\mathcal{A}_o = \mathcal{B}$. Um $\mathcal{A}_o \supset \mathcal{B}$ zu zeigen, genügt es, jeden halboffenen Quader $[a,b) \subset \mathbb{R}^n$ als abzählbaren Durchschnitt offener Mengen darzustellen (denn σ-Algebren sind stabil unter abzählbaren Durchschnitten). Dazu bemerken wir

$$\prod_{k=1}^{n}[a_k,b_k) = \bigcap_{j\in\mathbb{N}}\prod_{k=1}^{n}\left(a_k - \frac{1}{j}, b_k\right).$$

Für die andere Inklusion genügt es, daß jede offene Teilmenge des $\mathbb{R}^n$ die abzählbare Vereinigung halboffener Quader ist. In der Tat, sei $A \subset \mathbb{R}^n$ offen, dann existiert für jedes $x \in A$ eine Umgebung in A und daher insbesondere ein halboffener Quader $x \in Q_x \subset A$ mit rationalen Eckpunkten. Die Menge der halboffenen Quader mit rationalen Ecken ist aber abzählbar. Also ist $\bigcup_{x\in A} Q_x$ tatsächlich eine abzählbare Vereinigung, und diese ergibt genau A. □

Die Berechnung des Maßes konkreter Mengen stellen wir hintan, da sich im nächsten Kapitel herausstellen wird, daß die Volumenberechnung gängiger geometrischer Objekte mit dem LEBESGUE-Maß auf das gleiche Ergebnis führt wie mit den elementareren Integrationsmethoden aus Analysis II [**2**]. Wir bemerken hier nur folgendes:

PROPOSITION 2.10. *Jede abzählbare Teilmenge von $\mathbb{R}^n$ ist eine* LEBESGUE-*Nullmenge.*

BEWEIS. Sei $(x_k)_{k\in\mathbb{N}}$ eine Abzählung der fraglichen Menge, wobei ohne Einschränkung die x_k paarweise verschieden seien. Dann genügt es zu zeigen, daß für jedes $x \in \mathbb{R}^n$ gilt $\lambda(\{x\}) = 0$, denn aus der σ-Additivität von λ folgt dann $\lambda\left(\bigcup_{k\in\mathbb{N}}\{x_k\}\right) = \sum_{k\in\mathbb{N}}\lambda(\{x_k\}) = 0$. (Beachte, daß $\{x\}$ kompakt und damit meßbar ist.)

Sei also $x \in \mathbb{R}^n$. Dann gilt $\{x\} \subset [x,y)$ für jedes y, dessen Koordinaten je größer sind als die von x. Insbesondere können wir $y = x + (\epsilon,\dots,\epsilon)$ wählen und erhalten so

$$\lambda(\{x\}) \le \lambda([x,y)) = \epsilon^n,$$

und da $\epsilon > 0$ beliebig war, folgt die Behauptung. □

Insbesondere erhalten wir, daß $\mathbb{Q}^n$ eine n-dimensionale Nullmenge ist. In diesem Sinne ist, obwohl $\mathbb{Q}$ dicht in $\mathbb{R}$ liegt, $\mathbb{Q}$ eine „kleine“ Teilmenge von $\mathbb{R}$.

Die Frage, ob umgekehrt jede Nullmenge abzählbar ist, muß verneint werden: Ein Beispiel für eine überabzählbare Nullmenge ist die CANTOR-Menge (Übung).

SATZ 2.11 (Translationsinvarianz). Sei $B \in \mathcal{B}$ und $x \in \mathbb{R}^n$. Dann gilt $\lambda(B) = \lambda(x+B)$, wobei $x + B := \{x + y : y \in B\}$.

BEWEIS. Offenbar gilt die Aussage für alle $B \in \mathcal{R}$ (nach Definition des LEBESGUE-Inhalts – das Produkt der Seitenlängen eines Quaders ist invariant unter Translation). Definieren wir also auf $\mathcal{B}$ ein weiteres Maß λ_x durch die Vorschrift $\lambda_x(B) := \lambda(x+B)$ (man prüft sofort, daß λ_x ein Maß ist), so gilt auf $\mathcal{R}$ $\lambda = \lambda_x$, und mit der Eindeutigkeitsaussage von Satz 2.6 gilt diese Gleichheit sogar auf $\mathcal{B}$. □

Selbstverständlich ist die Translationsinvarianz etwas, das man vom LEBESGUE-Maß erwartet, ebenso wie die Bewegungsinvarianz (s. dazu Korollar 3.30 im nächsten Kapitel).

Es ist nun überhaupt nicht offensichtlich, ob es Teilmengen von $\mathbb{R}^n$ gibt, die *nicht* meßbar sind. Akzeptiert man das Auswahlaxiom, so kann man aber eine solche Menge konstruieren (bzw. „konstruieren“):

SATZ 2.12 (VITALI). $\mathcal{B} \neq \mathcal{P}(\mathbb{R}^n)$.

Beweis. Wir betrachten nur den eindimensionalen Fall, $n = 1$; der höherdimensionale Fall geht analog. Definiere auf $\mathbb{R}$ die Äquivalenzrelation $x \equiv y$ genau dann, wenn $x - y \in \mathbb{Q}$. Bezeichne mit I die Menge der Äquivalenzklassen und wähle[6] für jedes $i \in I$ ein $x_i \in i$ mit $x_i \in [0,1]$ (die Einschränkung auf das Intervall $[0,1]$ läßt sich stets durch Addition einer geeigneten rationalen Zahl bewerkstelligen). Angenommen, die Menge $V := \{x_i : i \in I\} \subset [0,1]$ wäre meßbar. Dann wäre $\lambda(V) \leq 1$.

Andererseits ist $[0,1] \subset W := \bigcup_{q \in \mathbb{Q} \cap [-1,1]} (q + V)$: In der Tat, sei $x \in [0,1]$, dann existiert ein $i \in I$ mit $x \in i$, und somit ist $x = q + x_i \in q + V$ für ein $q \in \mathbb{Q} \cap [-1,1]$. Außerdem sind $p + V$ und $q + V$ disjunkt für $p, q \in \mathbb{Q}$, $p \neq q$: Gibt es nämlich $x \in p + V \cap q + V$, so gibt es $x_i, x_j \in V$ mit $x_i - x_j \in \mathbb{Q}$, woraus folgt $i = j$; aber dann folgt auch $p = q$.

Nun ist einerseits $[0,1] \subset W \subset [-1,2]$, also $1 \leq \lambda(W) \leq 3$. Mit der σ-Additivität und der Translationsinvarianz des Lebesgue-Maßes folgt andererseits

$$\lambda(W) = \sum_{q \in \mathbb{Q} \cap [-1,1]} \lambda(q + V) = \sum_{q \in \mathbb{Q} \cap [-1,1]} \lambda(V),$$

was nur 0 oder ∞ sein kann. Dies ist der gewünschte Widerspruch zu $\lambda(W) \in [1,3]$ und damit zur Meßbarkeit von V. □

Tatsächlich haben wir damit sogar gezeigt: *Es existiert kein translationsinvariantes Maß auf $\mathcal{P}(\mathbb{R})$, das jedem Intervall seine Länge zuordnet!* Die Unmöglichkeit, jeder Teilmenge des $\mathbb{R}^n$ in sinnvoller Weise ein Volumen zuzuordnen, ist also keine Defizienz des Lebesgue-Maßes, sondern eine generelle Unumgänglichkeit.

Man kann zeigen, daß das Lebesgue-Maß bis auf eine multiplikative Konstante das einzige translationsinvariante Maß auf $\mathcal{B}$ ist. Ein Maß mit dieser Eigenschaft heißt Haar-Maß; ein Haar-Maß kann in viel allgemeineren Maßräumen konstruiert werden (nämlich in lokalkompakten topologischen Gruppen).

[6] mit dem Auswahlaxiom!

KAPITEL 3

Integrationstheorie

Ist f eine Zufallsvariable, so möchte man (falls existent) deren Erwartungswert, Varianz etc. bestimmen. Diese Größen kann man als Integrale (über das zugrundeliegende Wahrscheinlichkeitsmaß) ausdrücken. Integriert man hingegen bezüglich des LEBESGUE-Maßes, so erhält man (für „schöne", z.B. stetige, Integranden) das Integral aus Analysis I oder II zurück, mit dem man Flächeninhalte oder Volumina bestimmen konnte. Integration bezüglich allgemeiner Maße bildet daher einen sehr allgemeinen Rahmen zur Beschreibung und Lösung diverser Probleme in Stochastik und Analysis.

Auf einer etwas technischeren Ebene sind es die extrem nützlichen Konvergenzsätze, die den hier vorgestellten, von LEBESGUE entwickelten Integralbegriff rechtfertigen.

3.1. Meßbare Abbildungen

3.1.1. Meßbarkeit zwischen allgemeinen Maßräumen.

DEFINITION 3.1 (Meßbare Abbildung). Seien $(\Omega, \mathcal{A}, \mu)$, $(\Omega', \mathcal{A}', \mu')$ Maßräume. Eine Abbildung $f : \Omega \to \Omega'$ heißt *meßbar*, wenn für jedes $A' \in \mathcal{A}'$ gilt

$$f^{-1}(A') \in \mathcal{A}.$$

Kurz: Eine Abbildung ist meßbar, wenn das Urbild jeder meßbaren Menge wieder meßbar ist.

BEMERKUNG 3.2. Interpretation: Damit f eine Zufallsvariable ist, muß für jedes Ereignis A' die Aussage „f liegt in A'" interpretierbar sein.

Man beachte die Analogie zur topologischen Charakterisierung der Stetigkeit: Eine Funktion ist stetig, wenn das Urbild jeder offenen Menge wieder offen ist.

Man beachte, daß die Meßbarkeit *nicht* von den Maßen μ, μ' abhängt, sondern nur von den entsprechenden σ-Algebren.

BEISPIEL 3.3. Konstante Funktionen sind meßbar, denn für jede Menge $A' \in \mathcal{A}'$ ist $f^{-1}(A')$ entweder $\varnothing$ oder Ω.

PROPOSITION 3.4. *Seien $(\Omega, \mathcal{A}, \mu)$, $(\Omega', \mathcal{A}', \mu')$ Maßräume und $\mathcal{E}' \subset \mathcal{P}(\Omega')$ so, daß $\sigma(\mathcal{E}') = \mathcal{A}'$. Eine Abbildung $f : \Omega \to \Omega'$ ist meßbar genau dann, wenn für jedes $E' \in \mathcal{E}'$ gilt*

$$f^{-1}(E') \in \mathcal{A}.$$

BEWEIS. Wenn f meßbar ist, so ist die Aussage wegen $E' \in \mathcal{A}'$ klar. Gelte umgekehrt $f^{-1}(E') \in \mathcal{A}$ für alle $E' \in \mathcal{E}'$. Definiere die Menge

$$\mathcal{Q}' := \{Q' \subset \mathcal{P}(\Omega') : f^{-1}(Q') \in \mathcal{A}\},$$

so ist $\mathcal{Q}'$, wie man leicht nachprüft, eine σ-Algebra, die $\mathcal{E}'$ enthält. Damit ist nach Voraussetzung $\mathcal{Q}' \supset \sigma(\mathcal{E}') = \mathcal{A}'$, was zu zeigen war.

□

PROPOSITION 3.5 (Komposition meßbarer Abbildungen). *Seien $(\Omega, \mathcal{A}, \mu)$, $(\Omega', \mathcal{A}', \mu')$, $(\Omega'', \mathcal{A}'', \mu'')$ Maßräume und $f : \Omega \to \Omega'$, $g : \Omega' \to \Omega''$ meßbar. Dann ist auch $g \circ f : \Omega \to \Omega''$ meßbar.*

Beweis. Sei $A'' \in \mathcal{A}''$, so ist $g^{-1}(A'') \in \mathcal{A}'$, und damit $f^{-1}(g^{-1}(A'')) \in \mathcal{A}$. Die Behauptung folgt nun aus $f^{-1} \circ g^{-1} = (g \circ f)^{-1}$. □

Definition 3.6 (Bildmaß). Seien $(\Omega, \mathcal{A}, \mu)$, $(\Omega', \mathcal{A}', \mu')$ Maßräume und $f : \Omega \to \Omega'$ meßbar, so definiert

$$\mu \circ f^{-1} : \mathcal{A}' \to \mathbb{R}_0^+ \cup \{+\infty\}, \quad A' \mapsto \mu(f^{-1}(A'))$$

ein Maß auf $\mathcal{A}'$, das *Bildmaß* von μ unter f.

Bemerkung 3.7. Ist f eine Zufallsvariable und μ ein Wahrscheinlichkeitsmaß, so beschreibt $\mu \circ f^{-1}(A')$ die Wahrscheinlichkeit, daß f einen Wert in A' annimmt; lies: $\mu \circ f^{-1}(A') = \mu(\{f \in A'\})$.

3.1.2. Meßbarkeit reellwertiger Funktionen. Wir spezialisieren nun auf den Fall $(\Omega', \mathcal{A}') = (\mathbb{R}, \mathcal{B})$, also auf Funktionen, die nach $\mathbb{R}$ versehen mit der Borel-σ-Algebra abbilden. Es wird nützlich sein, $\mathbb{R}$ um die beiden Punkte $\pm\infty$ zu erweitern und mit der σ-Algebra zu versehen, die genau die Mengen B, $B \cup \{+\infty\}$, $B \cup \{-\infty\}$, $B \cup \{+\infty\} \cup \{-\infty\}$ mit $B \in \mathcal{B}$ enthält.[1] Wir bezeichnen diese erweiterte Zahlengerade mit $\bar{\mathbb{R}}$. Per Konvention vereinbaren wir $+\infty + \infty = +\infty$, $-\infty - \infty = -\infty$, $a \cdot (\pm\infty) = \pm\infty$ für $a > 0$, $a \cdot (\pm\infty) = \mp\infty$ für $a < 0$, $0 \cdot (\pm\infty) = 0$, $(\pm\infty) \cdot (\pm\infty) = +\infty$, $(\pm\infty) \cdot (\mp\infty) = -\infty$. Nicht definiert ist dagegen $+\infty - \infty$. Wir ordnen $\bar{\mathbb{R}}$ so an, daß $-\infty < a < +\infty$ für alle $a \in \mathbb{R}$.

Beispiel 3.8.
- Sei $(\Omega, \mathcal{A}) = (\mathbb{R}^n, \mathcal{B})$ und $f : \mathbb{R}^n \to \mathbb{R}$ stetig. Dann ist f meßbar, denn das Urbild jeder offenen Menge ist offen (insbesondere meßbar), und die offenen Mengen erzeugen die Borel-σ-Algebra $\mathcal{B}$. Verwende nun Proposition 3.4. Ebenso ist jede stetige Funktion, die auf einer meßbaren Teilmenge $B \subset \mathbb{R}^n$ definiert ist, meßbar (bzgl. der eingeschränkten Borel-σ-Algebra $\mathcal{B}_B$, siehe Diskussion vor Satz 2.9).
- Seien $(A_i)_{i=1}^N \subset \mathcal{A}$ eine endliche Familie meßbarer Teilmengen von Ω und $\alpha_i \in \mathbb{R}_0^+$, $i = 1, \dots, N$. Die *charakteristische Funktion* von A_i ist definiert als

  $$\chi_{A_i}(x) = \begin{cases} 1 & \text{falls } x \in A_i, \\ 0 & \text{sonst.} \end{cases}$$

  Dann heißt $\sum_{i=1}^N \alpha_i \chi_{A_i}$ *Elementarfunktion* und ist meßbar: Jedes χ_{A_i} ist meßbar, da die Urbilder meßbarer Teilmengen von $\mathbb{R}$ nur $\varnothing$, Ω, A_i oder A_i^c sein können. Daher ist auch jede Elementarfunktion meßbar, da, wie wir bald sehen werden, Linearkombinationen meßbarer Funktionen wieder meßbar sind.

Satz 3.9. Eine Funktion $f : \Omega \to \bar{\mathbb{R}}$ ist meßbar genau dann, wenn für jedes $\alpha \in \mathbb{R}$ die Menge $\{x \in \Omega : f(x) \geq \alpha\}$ meßbar ist.

Beweis. Ist f meßbar, so ist auch die Menge $\{f \geq \alpha\}$ als Urbild des (meßbaren) Intervalls $[\alpha, +\infty]$ wieder meßbar. Sei umgekehrt $f^{-1}([\alpha, +\infty])$ meßbar für alle α. Nach Proposition 3.4 genügt es für die Meßbarkeit von f zu zeigen, daß die Intervalle $[\alpha, +\infty]$ die Borel-σ-Algebra auf $\bar{\mathbb{R}}$ erzeugen. Falls $\mathbb{R} \ni \beta \geq \alpha$, so ist $[\alpha, \beta) = [\alpha, +\infty] \setminus [\beta, +\infty]$ im Erzeugnis der fraglichen Intervalle; da aber halboffene Intervalle der Form $[\alpha, \beta)$ mit $\alpha \leq \beta$, $\alpha, \beta \in \mathbb{R}$ die Borel-σ-Algebra auf $\mathbb{R}$ erzeugen, müssen wir nur noch zeigen, daß auch die einelementigen Mengen $\{\pm\infty\}$ im Erzeugnis der Intervalle $[\alpha, +\infty]$ liegen. Es gilt aber

$$\{+\infty\} = \bigcap_{\alpha \in \mathbb{Q}} [\alpha, +\infty], \quad \{-\infty\} = \left(\bigcup_{\alpha \in \mathbb{Q}} [\alpha, +\infty] \right)^c.$$

[1] Die Ontologie der Objekte $\{\pm\infty\}$ spielt hier keine Rolle, solange man gewiß ist, daß die angegebenen Mengen eine σ-Algebra bilden und man die Konventionen $a \pm \infty = \pm\infty$, $-\infty < a < +\infty$ etc. für jedes reelle a beachtet.

□

In ähnlicher Weise zeigt man: f ist meßbar genau dann, wenn für alle $\alpha \in \mathbb{R}$ die Menge $\{f > \alpha\}$ meßbar ist, und analog für $\{f \le \alpha\}$ und $\{f < \alpha\}$.

PROPOSITION 3.10. *Seien $f, g : \Omega \to \bar{\mathbb{R}}$ meßbar, so sind die Mengen $\{f < g\}$, $\{f \le g\}$, $\{f = g\}$ und $\{f \ne g\}$ meßbar*[2].

BEWEIS. Wegen $\{f < g\} = \bigcup_{\alpha \in \mathbb{Q}} (\{f \le \alpha\} \cap \{g > \alpha\})$ folgt die Meßbarkeit von $\{f < g\}$ aus dem vorigen Satz (mit Nachbemerkung). Dann ist aber $\{f \ne g\} = \{f < g\} \cup \{g < f\}$ und damit auch $\{f = g\} = \{f \ne g\}^c$ meßbar. Schließlich folgt auch die Meßbarkeit von $\{f \le g\} = \{f < g\} \cup \{f = g\}$. □

SATZ 3.11. Seien $f, g : \Omega \to \bar{\mathbb{R}}$ meßbar, so auch $f \pm g$ (falls überall definiert) sowie fg.

BEWEIS. Ist f meßbar und $\beta, \gamma \in \mathbb{R}$, so ist auch $\beta + \gamma f$ meßbar, denn für jedes $\alpha \in \mathbb{R}$ ist $\{\beta + \gamma f \ge \alpha\} = \{f \ge \gamma^{-1}(\alpha - \beta)\}$ (oder $\le$, falls $\gamma < 0$; der Fall $\gamma = 0$ ist trivial).

Für $\alpha \in \mathbb{R}$ ist daher $\{f - g \ge \alpha\} = \{f \ge \alpha + g\}$, was nach der vorigen Proposition meßbar ist. Somit ist $f - g$ meßbar. Ebenso ist $f + g = f - (-1)g$ meßbar.

Sind f und g reellwertig, so gilt

$$fg = \frac{1}{4}(f+g)^2 - \frac{1}{4}(f-g)^2,$$

also genügt es, die Meßbarkeit von f^2 zu zeigen. Es ist aber für $\alpha \in \mathbb{R}$

$$\{f^2 \ge \alpha\} = \{f \ge \sqrt{\alpha}\} \cup \{f \le -\sqrt{\alpha}\},$$

was nach der vorangehenden Proposition wieder meßbar ist.

Es verbleibt noch der Fall, daß f oder g den Wert $+\infty$ oder $-\infty$ annehmen kann. (Per Konvention ist $+\infty \cdot +\infty = +\infty$, $-\infty \cdot (-\infty) = +\infty$, $+\infty \cdot (-\infty) = -\infty$, $\pm\infty \cdot 0 = 0$.) Betrachte dazu die paarweise disjunkten Mengen $\Omega_\pm := \{fg = \pm\infty\}$, $\Omega_0 := \{fg = 0\}$ und $\Omega' := (\Omega_+ \cup \Omega_- \cup \Omega_0)^c$, deren Vereinigung Ω ist. Diese Mengen sind sämtlich meßbar, da

$$\begin{aligned}\{fg = +\infty\} = (\{f = +\infty\} \cap \{g > 0\}) \cup (\{g = +\infty\} \cap \{f > 0\}) \\ \cup (\{f = -\infty\} \cap \{g < 0\}) \cup (\{g = -\infty\} \cap \{f < 0\})\end{aligned}$$

und analog für Ω_- und Ω_0. Nach dem eben Gezeigten ist die Einschränkung von fg auf Ω' meßbar, und ebenso die Einschränkungen von fg auf $\Omega_\pm$ und Ω_0.

Ist nun $B \in \mathcal{B}$, so ist

$$\{fg \in B\} = \{fg \restriction_{\Omega_\pm} \in B\} \cup \{fg \restriction_{\Omega_0} \in B\} \cup \{fg \restriction_{\Omega'} \in B\},$$

und dementsprechend ist $\{fg \in B\} = (A_\pm \cap \Omega_\pm) \cup (A_0 \cap \Omega_0) \cup (A' \cap \Omega')$ für gewisse $A_\pm, A_0, A' \in \mathcal{A}$, also gilt $\{fg \in B\} \in \mathcal{A}$, was zu zeigen war. □

SATZ 3.12. Sei $(f_k)_{k \in \mathbb{N}} : \Omega \to \bar{\mathbb{R}}$ eine Familie meßbarer Funktionen, so sind auch die Funktionen $\sup_{k \in \mathbb{N}} f_k$, $\inf_{k \in \mathbb{N}} f_k$, $\limsup_{k \to \infty} f_k$, $\liminf_{k \to \infty} f_k$ meßbar.

BEWEIS. Wegen $\{\sup_{k \in \mathbb{N}} f_k \le \alpha\} = \bigcap_{k \in \mathbb{N}} \{f_k \le \alpha\}$ für alle $\alpha \in \mathbb{R}$ ist $\sup_{k \in \mathbb{N}} f_k$ nach Satz 3.9 meßbar, ebenso $\inf_{k \in \mathbb{N}} f_k = -\sup_{k \in \mathbb{N}} (-f_k)$.

Die Aussagen über $\limsup, \liminf$ folgen aus

$$\limsup_{k \to \infty} f_k = \inf_{k \in \mathbb{N}} \sup_{m \ge k} f_m$$

bzw. $\limsup_{k \to \infty} f_k = -\liminf_{k \to \infty} (-f_k)$ und dem eben Bewiesenen. □

KOROLLAR 3.13. Ist eine Folge $(f_k)_{k \in \mathbb{N}} : \Omega \to \bar{\mathbb{R}}$ meßbarer Funktionen punktweise konvergent, so ist auch die Grenzfunktion meßbar.

[2] Wie teilweise schon zuvor schreiben wir kurz $\{f < g\}$ statt $\{x \in \Omega : f(x) < g(x)\}$.

Wir definieren für eine Funktion $f:\Omega\to\bar{\mathbb{R}}$ den *Positiv-* bzw. *Negativteil* durch

$$f^+ := \max\{f,0\}, \quad f^- := -\min\{f,0\},$$

sodaß $f = f^+ - f^-$ sowie $|f| = f^+ + f^-$.

KOROLLAR 3.14. Eine Funktion $f:\Omega\to\bar{\mathbb{R}}$ ist meßbar genau dann, wenn f^+ und f^- meßbar sind.

BEWEIS. Ist f meßbar, so auch f^+ und f^- nach Satz 3.12; sind f^+ und f^- meßbar, so nach Satz 3.11 auch $f = f^+ - f^-$ (beachte, daß die Differenz stets definiert ist, da höchstens eine der Funktionen f^+ und f^- ungleich null ist). □

Es folgt unmittelbar, daß mit f auch $|f|$ meßbar ist. Die Umkehrung gilt nicht (warum?).

3.2. Integration von Elementarfunktionen

Auch in diesem Abschnitt sei wieder $\mathcal{A}$ eine σ-Algebra auf einer Menge Ω.

DEFINITION 3.15 (Elementarfunktion). Eine *Elementarfunktion* ist eine Funktion $\Omega\to\mathbb{R}$ der Gestalt $\sum_{k=1}^m \alpha_k\chi_{A_k}$, wobei $\alpha_k\in\mathbb{R}_0^+$ und A_k meßbar für alle $k=1,\dots,m$ (die charakteristische Funktion χ_{A_k} wurde in Beispiel 3.8 eingeführt). Die Menge der Elementarfunktionen werde mit E bezeichnet.

Nach Beispiel 3.8 sind Elementarfunktionen meßbar. Man sieht leicht, daß E abgeschlossen unter Multiplikation ist. Außerdem ist mit f,g auch $\max\{f,g\}$ und $\min\{f,g\}$ in E.

Offenbar besteht E aus genau den meßbaren Funktionen $f:\Omega\to\mathbb{R}_0^+$, die nur endlich viele paarweise verschiedene Werte $\alpha_1,\dots,\alpha_n$ annehmen; dann läßt sich ein solches f in eindeutiger Weise schreiben als $f=\sum_{k=1}^n\alpha_k\chi_{\{f=\alpha_k\}}$, wobei $\Omega=\bigcup_{k=1}^n\{f=\alpha_k\}$ eine disjunkte Vereinigung meßbarer Mengen ist.

DEFINITION 3.16 (Integral von Elementarfunktionen). Sei μ ein Maß auf $\mathcal{A}$. Für $f=\sum_{k=1}^n\alpha_k\chi_{\{f=\alpha_k\}}$ definieren wir

$$\int_\Omega f(x)d\mu(x) := \sum_{k=1}^n \alpha_k\mu(\{f=\alpha_k\}).$$

PROPOSITION 3.17. *Seien $(\Omega,\mathcal{A},\mu)$ ein Maßraum und $f,g\in E$, $\alpha\in\mathbb{R}_0^+$, $A\in\mathcal{A}$. Dann gilt*

(1) $\int_\Omega \chi_A(x)d\mu(x) = \mu(A)$;
(2) $\int_\Omega (\alpha f)(x)d\mu(x) = \alpha\int_\Omega f(x)d\mu(x)$;
(3) $\int_\Omega (f+g)(x)d\mu(x) = \int_\Omega f(x)d\mu(x) + \int_\Omega g(x)d\mu(x)$;
(4) *Falls $f\le g$, so gilt* $\int_\Omega f(x)d\mu(x) \le \int_\Omega g(x)d\mu(x)$.

BEWEIS. Die ersten beiden Eigenschaften folgen sofort aus der Definition. Sind nun $f=\sum_{k=1}^n\alpha_k\chi_{\{f=\alpha_k\}}$ und $g=\sum_{j=1}^l\beta_j\chi_{\{g=\beta_j\}}$, wobei die α_k und die β_j jeweils paarweise verschieden sind, so ist $f+g$ auf den Mengen $C_{k,j} := \{f=\alpha_k\}\cap\{g=\beta_j\}$ konstant gleich $\gamma_{k,j} := \alpha_k+\beta_j$.

Für jedes $\gamma\in\mathbb{R}_0^+$ setzen wir $C_\gamma := \{f+g=\gamma\}$; dann ist C_γ nur für endlich viele γ nichtleer, und jedes nichtleere C_γ ist die disjunkte Vereinigung von Mengen aus der

Familie $(C_{k,j})_{k,j}$. Wegen der Additivität von μ ist $\mu(C_\gamma) = \sum_{C_{j,k}\in C_\gamma} \mu(C_{j,k})$ und daher

$$\begin{aligned}\int_\Omega (f+g)(x)d\mu(x) &= \sum_{\gamma\in\mathbb{R}_0^+} \gamma\mu(C_\gamma) = \sum_{k,j} \gamma_{k,j}\mu(C_{k,j}) \\ &= \sum_{k,j} \alpha_k\mu(C_{k,j}) + \sum_{k,j} \beta_j\mu(C_{k,j}) \\ &= \sum_k \alpha_k \sum_j \mu(\{f=\alpha_k\}\cap\{g=\beta_j\}) + \sum_j \beta_j \sum_k \mu(\{f=\alpha_k\}\cap\{g=\beta_j\}) \\ &= \int_\Omega f(x)d\mu(x) + \int_\Omega g(x)d\mu(x),\end{aligned}$$

wobei wir im letzten Schritt wieder die Additivität von μ benutzt haben.

Ist zusätzlich $f \le g$, so gilt ebenso

$$\begin{aligned}\int_\Omega f(x)d\mu(x) &= \sum_{k,j} \alpha_k\mu(\{f=\alpha_k\}\cap\{g=\beta_j\}) \\ &\le \sum_{k,j} \beta_j\mu(\{f=\alpha_k\}\cap\{g=\beta_j\}) = \int_\Omega g(x)d\mu(x),\end{aligned}$$

denn für alle k, j, für die $\mu(\{f=\alpha_k\}\cap\{g=\beta_j\}) > 0$, ist $\{f=\alpha_k\}\cap\{g=\beta_j\} \neq \varnothing$, und daher $\alpha_k \le \beta_j$ wegen $f \le g$. □

3.3. Integration nichtnegativer meßbarer Funktionen

Mithilfe der Elementarfunktionen erweitern wir nun den Integralbegriff auf meßbare nichtnegative Funktionen. Die Idee ist, jede meßbare nichtnegative Funktion $\Omega \to \bar{\mathbb{R}}$ geeignet durch Elementarfunktionen zu approximieren und das Integral entsprechend als Grenzwert der Integrale der Elementarfunktionen zu definieren.

Es sei nach wie vor E der Raum der Elementarfunktionen.

Satz 3.18 (Approximation durch Elementarfunktionen). Sei $f : \Omega \to \bar{\mathbb{R}}$ meßbar und $f \ge 0$. Dann existiert eine Folge $(f_k)_{k\in\mathbb{N}} \subset E$ mit $f_k \le f_{k+1}$ für alle $k \in \mathbb{N}$, und $f = \lim_{k\to\infty} f_k = \sup_{k\in\mathbb{N}} f_k$.

Beweis. Für $k \in \mathbb{N}$ und $j = 0, \dots, k2^k - 1$ setze

$$A_{k,j} := \{j2^{-k} \le f < (j+1)2^{-k}\}, \quad A_{k,k2^k} := \{f \ge k\}$$

und entsprechend

$$f_k := \sum_{j=1}^{k2^k} j2^{-k}\chi_{A_{k,j}}.$$

Wir zeigen, daß f_k wie gewünscht ist. Nach Definition ist offenbar $f_k \le f$ für alle $k \in \mathbb{N}$. Zur Monotonie der Folge (f_k): Sei $x \in \Omega$, dann gibt es ein $A_{k,j} \ni x$, das bedeutet $j2^{-k} \le f(x) < (j+1)2^{-k}$ und nach Definition von f_k also $f_k(x) = j2^{-k}$. Andererseits ist wegen $2j2^{-(k+1)} \le f(x) < (2j+2)2^{-(k+1)}$ auch $x \in A_{k+1,2j} \cup A_{k+1,2j+1}$.

Ist $x \in A_{k+1,2j}$, so ist $f_{k+1}(x) = 2j2^{-(k+1)} = j2^{-k} = f_k(x)$. Ist andererseits $x \in A_{k+1,2j+1}$, so ist $f_{k+1}(x) = (2j+1)2^{-(k+1)} > j2^{-k} = f_k(x)$. In jedem Falle gilt also $f_{k+1}(x) \ge f_k(x)$.

Zweitens ist die Approximationseigenschaft zu zeigen. Sei dazu wieder $x \in \Omega$ und $\epsilon > 0$. Wir müssen zeigen, daß es ein $k \in \mathbb{N}$ gibt mit $f_k(x) > f(x) - \epsilon$. Wähle dazu einfach k so groß, daß $2^{-k} < \epsilon$. Ist $x \in A_{k,j}$, so ist dann $f(x) < (j+1)2^{-k} < f_k(x) + \epsilon$ wie gewünscht. □

Definition 3.19. Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und $f : \Omega \to \bar{\mathbb{R}}$ meßbar und nichtnegativ. Ist $(f_k)_{k\in\mathbb{N}} \subset E$ eine monoton steigende Folge von Elementarfunktionen mit $\lim_{k\to\infty} f_k = f$, so heißt

$$\int_\Omega f d\mu := \lim_{k\to\infty} \int_\Omega f_k d\mu$$

das *Integral* von f bezüglich μ.

Man beachte, daß die Folge der Integrale von f_k nach Proposition 3.17(4) monoton steigt und der Limes daher stets als reelle Zahl oder $+\infty$ existiert. Weniger offensichtlich ist dagegen die Frage, ob diese Definition des Integrals von der Wahl der approximierenden Folge abhängt. Daß dies nicht der Fall ist, soll nun gezeigt werden.

LEMMA 3.20. *Sei $(f_k)_{k\in\mathbb{N}} \subset E$ eine monoton steigende Folge von Elementarfunktionen und f eine Elementarfunktion mit $f \le \sup_{k\in\mathbb{N}} f_k$. Dann gilt auch*

$$\int_\Omega f d\mu \le \sup_{k\in\mathbb{N}} \int_\Omega f_k d\mu.$$

Insbesondere gilt: Sind $(f_k), (g_j)$ zwei monoton steigende Folgen von Elementarfunktionen mit $\sup_{k\in\mathbb{N}} f_k = \sup_{j\in\mathbb{N}} g_j$, so gilt auch

$$\sup_{k\in\mathbb{N}} \int_\Omega f_k d\mu = \sup_{j\in\mathbb{N}} \int_\Omega g_j d\mu.$$

BEWEIS. Sei $f = \sum_{l=1}^m \alpha_l \chi_{\{f=\alpha_l\}}$ mit paarweise verschiedenen $\alpha_l \ge 0$ und $\alpha \in (0,1)$ beliebig. Betrachte für jedes $k \in \mathbb{N}$ die (meßbare) Menge $A_k := \{f_k \ge \alpha f\}$. Da (f_k) monoton steigt, gilt $A_k \subset A_{k+1}$ für alle $k \in \mathbb{N}$. Nach Definition gilt $f_k \ge \alpha f \chi_{A_k}$ und daher nach Proposition 3.17

$$\int_\Omega f_k d\mu \ge \alpha \int_\Omega f \chi_{A_k} d\mu.$$

Nun ist aber $\bigcup_{k\in\mathbb{N}} A_k = \Omega$, da $f \le \sup_k f_k$ und $\alpha < 1$. Nach Übungsaufgabe 1 (Blatt 4) folgt daraus

$$\int_\Omega f d\mu = \sum_{l=1}^m \alpha_l \mu(\{f = \alpha_l\}) = \sup_{k\in\mathbb{N}} \sum_{l=1}^m \alpha_l \mu(\{f = \alpha_l\} \cap A_k) = \sup_{k\in\mathbb{N}} \int_\Omega f \chi_{A_k} d\mu.$$

Zusammen ergibt sich

$$\int_\Omega f d\mu \le \frac{1}{\alpha} \sup_{k\in\mathbb{N}} \int_\Omega f_k d\mu.$$

Da $\alpha < 1$ beliebig war, folgt die Behauptung.

Zum „insbesondere"-Teil: da für alle $k, j \in \mathbb{N}$ gilt $f_k \le \sup_j g_j$ und $g_j \le \sup_k f_k$, folgt aus dem bereits Bewiesenen

$$\sup_k \int_\Omega f_k d\mu \le \sup_j \int_\Omega g_j d\mu$$

und analog

$$\sup_j \int_\Omega g_j d\mu \le \sup_k \int_\Omega f_k d\mu.$$

□

Damit ist unsere Definition des Integrals gerechtfertigt.

BEMERKUNG 3.21. Unsere explizite Konstruktion der approximierenden Elementarfunktionen im Beweis von Satz 3.18 zeigt einen entscheidenden Unterschied zwischen dem hier eingeführten LEBESGUE-Integral und dem aus der Analysis bekannten RIEMANN-Integral auf: Während bei RIEMANN der Definitionsbereich in kleine Stücke geteilt wird, teilt LEBESGUE stattdessen den *Wertebereich* in kleine Intervalle. Dadurch ist sein Integralbegriff besser an die zu integrierende Funktion angepaßt. Die Vorzüge des LEBESGUE-Integrals zeigen sich später in den Konvergenzsätzen.

PROPOSITION 3.22 (Linearität und Monotonie). *Seien $\alpha \in \mathbb{R}_0^+$ und $f, g \geq 0$ meßbar. Dann gilt*

$$\int_\Omega (\alpha f) d\mu = \alpha \int_\Omega f d\mu$$

und

$$\int_\Omega (f+g) d\mu = \int_\Omega f d\mu + \int_\Omega g d\mu.$$

Ist außerdem $f \leq g$, so gilt

$$\int_\Omega f d\mu \leq \int_\Omega g d\mu.$$

BEWEIS. Seien (f_k), (g_j) Folgen von Elementarfunktionen, die monoton gegen f bzw. g konvergieren. Dann sind (αf_k), $(f_k + g_k)$ Folgen von Elementarfunktionen, die monoton gegen αf bzw. $f+g$ konvergieren. Daher gilt nach Definition des Integrals

$$\int_\Omega (\alpha f) d\mu = \sup_k \int_\Omega (\alpha f_k) d\mu = \alpha \sup_k \int_\Omega f_k d\mu = \alpha \int_\Omega f d\mu,$$

wobei wir die Linearität des Integrals für Elementarfunktionen benutzt haben (Proposition 3.17). Analog argumentiert man für $\int_\Omega (f+g) d\mu$.

Ist nun zusätzlich $f \leq g$, so haben wir $f_k \leq \sup_j g_j$ für jedes $k \in \mathbb{N}$, und damit nach Lemma 3.20

$$\int_\Omega f_k d\mu \leq \sup_j \int_\Omega g_j d\mu = \int_\Omega g d\mu.$$

Nimmt man noch das Supremum über alle $k \in \mathbb{N}$, so folgt die Behauptung. □

3.4. Integrierbare Funktionen

Die Forderung der Nichtnegativität ist scheinbar eine große Einschränkung. Diese kann jedoch in denkbar einfacher Weise aufgehoben werden. Wir erinnern uns an den *Positiv- bzw. Negativteil* einer meßbaren Funktion: $f^+ = \max\{f, 0\}$, $f^- = -\min\{f, 0\}$. Offenbar sind f^+ und f^- nichtnegativ, können also gemäß unserer bisherigen Theorie integriert werden. Nach Korollar 3.14 ist außerdem f meßbar genau dann, wenn f^+ und f^- meßbar sind.

DEFINITION 3.23 (Integrierbarkeit). Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und $f : \Omega \to \bar{\mathbb{R}}$ meßbar. Dann heißt f *μ-integrierbar*, wenn $\int_\Omega f^+ d\mu < \infty$ und $\int_\Omega f^- d\mu < \infty$. In diesem Falle setzt man

$$\int_\Omega f(x) d\mu(x) := \int_\Omega f^+(x) d\mu(x) - \int_\Omega f^-(x) d\mu(x).$$

Die erwarteten Eigenschaften des Integrals folgen damit sofort aus denen des Integrals nichtnegativer Funktionen (Proposition 3.22):

SATZ 3.24 (Linearität, Monotonie, Dreiecksungleichung). Seien $\alpha \in \mathbb{R}$ und f, g integrierbar. Dann sind auch αf, $f+g$ (falls überall definiert) und $|f|$ integrierbar, und es gilt

$$\int_\Omega (\alpha f) d\mu = \alpha \int_\Omega f d\mu,$$

$$\int_\Omega (f+g) d\mu = \int_\Omega f d\mu + \int_\Omega g d\mu,$$

und

$$\left|\int_\Omega f d\mu\right| \leq \int_\Omega |f| d\mu.$$

Ist außerdem $f \leq g$, so gilt

$$\int_\Omega f d\mu \leq \int_\Omega g d\mu.$$

BEWEIS. Ist $\alpha \geq 0$, so ist $(\alpha f)^\pm = \alpha f^\pm$, woraus mit Proposition 3.22 die Behauptung bereits folgt. Ist dagegen $\alpha < 0$, so ist $(\alpha f)^\pm = -\alpha f^\mp$, und die Behauptung folgt analog.

Die Aussage über $f+g$ erhalten wir wie folgt: Einerseits ist $f+g = (f+g)^+ - (f+g)^-$, wobei $\int_\Omega (f+g)^+ d\mu < \infty$ wegen $(f+g)^+ \leq f^+ + g^+$ (was man sich überlegen möge) und analog $\int_\Omega (f+g)^- d\mu < \infty$. Andererseits $f+g = (f^+ + g^+) - (f^- + g^-)$. Umstellen gibt

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

und nach Integration mithilfe von Proposition 3.22

$$\int_\Omega (f+g)^+ d\mu + \int_\Omega f^- d\mu + \int_\Omega g^- d\mu = \int_\Omega (f+g)^- d\mu + \int_\Omega f^+ d\mu + \int_\Omega g^+ d\mu.$$

Daraus folgt schließlich

$$\begin{aligned}\int_\Omega (f+g) d\mu &= \int_\Omega (f+g)^+ d\mu - \int_\Omega (f+g)^- d\mu \\ &= \int_\Omega f^+ d\mu - \int_\Omega f^- d\mu + \int_\Omega g^+ d\mu - \int_\Omega g^- d\mu \\ &= \int f d\mu + \int g d\mu.\end{aligned}$$

Die Monotonie erhalten wir durch $f^+ \leq g^+$ und $f^- \geq g^-$ und die Monotonie des Integrals für nichtnegative Funktionen (wiederum Proposition 3.22). Schließlich stellen wir fest, daß einerseits $|f| = f^+ + f^-$, was integrierbar ist, und $\pm f \leq |f|$, sodaß aus der Monotonie des Integrals die Dreiecksungleichung folgt.

□

BEISPIEL 3.25.
- Sei $\Omega \subset \mathbb{R}^n$ kompakt und f stetig in Ω. Wir werden bald sehen, daß das Integral $\int_\Omega f d\lambda$ (hier ist wie üblich λ das LEBESGUE-Maß auf Ω) mit dem aus der Analysis-Vorlesung bekannten Integral übereinstimmt. Jetzt werde nur die Integrierbarkeit gezeigt: Als stetige Funktion ist f meßbar. Da f stetig auf einem Kompaktum ist, ist es auch beschränkt durch $M > 0$, also $|f| \leq M$. Nun ist $\int_\Omega M d\lambda = M\lambda(\Omega) < \infty$, also ist M integrierbar. Wegen $f^\pm \leq |f| \leq M$ sind dann auch die Integrale über $f^\pm$ endlich, also ist f tatsächlich integrierbar.
- Sei Ω beliebig, $\mathcal{A} = \mathcal{P}(\Omega)$ und $x \in \Omega$. Das in x konzentrierte DIRAC-Maß δ_x war ja definiert durch $\delta_x(A) = 1$ falls $x \in A$ und $\delta_x(A) = 0$ sonst. Sei $f : \Omega \to \bar{\mathbb{R}}$ beliebig[3] mit $f(x) \in \mathbb{R}$. Wir behaupten, daß f integrierbar ist mit

$$\int_\Omega f d\delta_x = f(x).$$

  In der Tat, sei (f_k) eine steigende Folge von Elementarfunktionen, die gegen $f^\pm$ konvergiert. Für jede Elementarfunktion $g = \sum_{l=1}^m \alpha_l \chi_{\{g=\alpha_l\}}$ ist

$$\int_\Omega g d\delta_x = \sum_{l=1}^m \alpha_l \delta_x(\{g = \alpha_l\}),$$

[3] Im Falle $\mathcal{A} = \mathcal{P}(\Omega)$ muß die Meßbarkeit nicht gefordert werden, da jede Funktion meßbar ist.

und $\delta_x(\{g = \alpha_l\}) = 1$ genau dann, wenn $g(x) = \alpha_l$. Es folgt $\int_\Omega g d\delta_x = g(x)$. Demnach ist

$$\int_\Omega f^\pm d\delta_x = \lim_{k\to\infty} \int_\Omega f_k d\delta_x = \lim_{k\to\infty} f_k(x) = f^\pm(x) < \infty$$

und entsprechend $\int_\Omega f d\delta_x = f^+(x) - f^-(x) = f(x)$.

- Sei $\Omega = \mathbb{N}$, $\mathcal{A} = \mathcal{P}(\Omega)$ und μ das in Beispiel 1.16 eingeführte Zählmaß auf $\mathcal{P}(\Omega)$. Ist $f : \mathbb{N} \to \mathbb{R}$ nichtnegativ, so ist eine monotone Approximation durch die Elementarfunktionen

  $$f_k := f\chi_{\{n\in\mathbb{N}: n\leq k\}}$$

  gegeben. Wir haben aber

  $$\int_\mathbb{N} f_k d\mu = \sum_{n=1}^{k} f(n),$$

  also $\int_\mathbb{N} f d\mu = \sum_{n=1}^\infty f(n)$.

  Dementsprechend ist ein (nicht notwendig nichtnegatives) $f : \mathbb{N} \to \mathbb{R}$ integrierbar bzgl. des Zählmaßes genau dann, wenn die Reihe $\sum_{n\in\mathbb{N}} f(n)$ absolut konvergiert, und $\int_\mathbb{N} f d\mu = \sum_{n=1}^\infty f(n)$.
- Sei $X : \mathbb{R} \to \mathbb{R}$ eine Zufallsvariable (d.h. eine meßbare Funktion) bezüglich der Borel-σ-Algebra. Die Verteilung von X sei durch eine glatte *Wahrscheinlichkeitsdichtefunktion* $f : \mathbb{R} \to \mathbb{R}_0^+$ mit $\int_\mathbb{R} f d\lambda = 1$ gegeben, das heißt

  $$\mathbb{P}(X \in A) = \int_\mathbb{R} \chi_A(x) f(x) d\lambda(x) =: f\lambda(A)$$

  für jedes meßbare $A \subset \mathbb{R}$. (Dann ist $f\lambda$ das Bildmaß von $\mathbb{P}$ unter X: $f\lambda = \mathbb{P} \circ X^{-1}$.) Dann lassen sich Erwartungswert und Varianz von X als Integrale über das Maß $f\lambda$ schreiben:

  $$\mathbb{E}[X] = \int_\mathbb{R} x f(x) d\lambda(x), \quad \mathrm{Var}[X] = \int_R x^2 f(x) d\lambda(x) - \mathbb{E}[X]^2.$$

  Allgemeiner gilt, auch wenn die Verteilung von X nicht durch eine Dichte gegeben ist:

  $$\mathbb{E}[X] = \int_\mathbb{R} x d\mu(x), \quad \mathrm{Var}[X] = \int_R x^2 d\mu(x) - \mathbb{E}[X]^2,$$

  wobei $\mu := \mathbb{P} \circ X^{-1}$.

Bemerkung 3.26. Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und $A \in \mathcal{A}$. Dann kann man das Integral auf A einschränken, indem man setzt

$$\int_A f d\mu := \int_\Omega \chi_A f d\mu.$$

Andererseits könnte man auch die Einschränkung des Maßes μ auf die Menge A betrachten und

$$\int_A f d\mu := \int_A f \restriction_A d\mu \restriction_A$$

setzen. Man kann zeigen, daß beide Definitionen äquivalent sind [**1**, §12].

3.5. Das LEBESGUE-Integral

Im wichtigen Spezialfall, daß Ω eine meßbare Teilmenge von $\mathbb{R}^n$ und $\mu = \lambda$ ist, heißt $\int_\Omega f d\lambda$ das LEBESGUE-*Integral* der integrierbaren Funktion $f : \Omega \to \bar{\mathbb{R}}$. Wir haben in Satz 3.24 gezeigt, daß das LEBESGUE-Integral linear und monoton ist. Sei nun $\Omega = \mathbb{R}^n$, dann ist das LEBESGUE-Maß außerdem translationsinvariant (Satz 2.11). Aus der Translationseigenschaft des Integrals (Übung) erhält man somit

PROPOSITION 3.27. *Das* LEBESGUE-*Integral auf* $\mathbb{R}^n$ *ist translationsinvariant, das heißt: Ist* $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ *integrierbar bzgl.* λ *und* $a \in \mathbb{R}^n$*, so definiert auch* $T_a f(x) := f(x-a)$ *eine integrierbare Funktion, und es gilt*

$$\int_{\mathbb{R}^n} T_a f d\lambda = \int_{\mathbb{R}^n} f d\lambda.$$

Schränken wir uns auf die Menge $C_c(\mathbb{R}^n)$ der stetigen Funktionen $\mathbb{R}^n \to \mathbb{R}$ mit kompaktem Träger ein, so sieht man leicht (nämlich wie im ersten Beispiel von 3.25), daß solche Funktionen integrierbar sind. Also definiert

$$f \mapsto \int_{\mathbb{R}^n} f d\lambda$$

ein lineares, monotones, translationsinvariantes Funktional $C_c(\mathbb{R}^n) \to \mathbb{R}$. Aus Analysis II [**2**, Satz 8.11] wissen wir aber[4], daß dieses Funktional bis auf eine multiplikative Konstante dem RIEMANN-Integral gleicht. Diese multiplikative Konstante ist aber gleich eins, da das RIEMANN- und das LEBESGUE-Integral der charakteristischen Funktion des Einheitswürfels übereinstimmen (nämlich gleich eins sind). Es folgt:

SATZ 3.28 (Übereinstimmung der Integralbegriffe auf stetigen Funktionen). Ist $f \in C_c(\mathbb{R}^n)$, so gilt

$$\int_{\mathbb{R}^n} f d\lambda = \int_{\mathbb{R}^n} f(x) dx$$

(wobei links das LEBESGUE- und rechts das RIEMANN-Integral gemeint ist).

Oft wird darum einfach $\int_{\mathbb{R}^n} f dx$ statt $\int_{\mathbb{R}^n} f d\lambda$ geschrieben. Statt dx ist auch die Notation $d^n x$ in Gebrauch, in der die Dimension des Integrationsbereichs explizit vorkommt.

Der Satz bleibt auch z.B. für stetige Funktionen auf kompakten Teilmengen des $\mathbb{R}^n$, für stückweise stetige Funktionen etc. gültig. Man kann also das LEBESGUE-Integral als Erweiterung des Integralbegriffs von RIEMANN auffassen.

Man kann zeigen, daß für allgemeine LEBESGUE-integrierbare Funktionen die Transformationsformel gültig bleibt. Wir erinnern uns aus der Analysis: Eine Abbildung ϕ von einer offenen Teilmenge $\Omega \subset \mathbb{R}^n$ nach $\mathbb{R}^n$ heißt C^1-*Diffeomorphismus*, wenn $\phi : \Omega \to \phi(\Omega)$ bijektiv ist und sowohl ϕ als auch ϕ^{-1} stetig differenzierbar sind.

SATZ 3.29 (Transformationsformel). Sei $\Omega \subset \mathbb{R}^n$ offen und $\phi : \Omega \to \phi(\Omega) \subset \mathbb{R}^n$ ein C^1-Diffeomorphismus. Eine Funktion $f : \phi(\Omega) \to \bar{\mathbb{R}}$ ist integrierbar bzgl. des LEBESGUE-Maßes genau dann, wenn $|\det D\phi| f \circ \phi$ auf Ω integrierbar ist bzgl. des LEBESGUE-Maßes. In diesem Falle gilt

$$\int_\Omega |\det D\phi(x)| f(\phi(x)) d\lambda(x) = \int_{\phi(\Omega)} f d\lambda.$$

KOROLLAR 3.30 (Bewegungsinvarianz des LEBESGUE-Maßes). Sei $T \in O(n)$ eine orthogonale lineare Abbildung. Dann gilt für jede meßbare Menge $A \subset \mathbb{R}^n$

$$\lambda(TA) = \lambda(A),$$

wobei wir geschrieben haben $TA := \{Tx : x \in A\}$.

[4] Falls Sie die Analysis nicht bei mir gehört haben, nehmen Sie die Aussage bitte einfach hin.

BEWEIS. Seien $T \in O(n)$, $A \in \mathcal{B}$ gegeben. Die lineare Abbildung $\phi : x \mapsto T^{-1}x$ ist ein Diffeomorphismus von $\mathbb{R}^n$ nach $\mathbb{R}^n$ mit $|\det D\phi| \equiv 1$. Wende nun die Transformationsformel auf dieses ϕ und auf $f := \chi_A$ an, dann folgt

$$\lambda(TA) = \int_{\mathbb{R}^n} \chi_A(T^{-1}x)d\lambda(x) = \int_{\mathbb{R}^n} \chi_A d\lambda = \lambda(A),$$

wobei zu beachten ist $\chi_A \circ T^{-1} = \chi_{TA}$. □

Besonders wichtige Eigenschaften der LEBESGUE-Integration, die RIEMANN nicht hat und die für die Entwicklung der Mathematik des 20. Jahrhunderts (Wahrscheinlichkeitstheorie, Funktionalanalysis, partielle Differentialgleichungen etc.) fundamental waren, beruhen auf den Konvergenzsätzen, die wir im nächsten Abschnitt besprechen.

3.6. Konvergenzsätze

Aus der Analysis wissen wir, daß im Falle gleichmäßiger Konvergenz Integration und Limesbildung vertauscht werden dürfen, genauer: Ist $[a, b] \subset \mathbb{R}$ ein kompaktes Intervall und konvergiert eine Folge $(f_k)_{k\in\mathbb{N}}$ stetiger Funktionen gleichmäßig gegen f, so gilt auch $\int_a^b f_k dx \to \int_a^b f dx$. Konvergiert eine Funktionenfolge dagegen nur punktweise, so konvergieren die Integrale im allgemeinen *nicht*.

Hauptziel dieses Abschnitts ist der Satz von der dominierten Konvergenz, demzufolge unter milden Zusatzbedingungen die punktweise Konvergenz eben doch für die Konvergenz der Integrale ausreicht. Der Satz hat weitreichende Anwendungen in Analysis und Stochastik.

3.6.1. Monotone Konvergenz und Lemma von FATOU. Vorbereitend zunächst ein einfacherer, aber dennoch bisweilen für sich nützlicher Satz:

SATZ 3.31 (von der monotonen Konvergenz (LEVI)). Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und $(f_k)_{k\in\mathbb{N}}$ eine Folge nichtnegativer meßbarer Funktionen $f_k : \Omega \to \bar{\mathbb{R}}$, sodaß $f_k \le f_{k+1}$ für alle $k \in \mathbb{N}$. Dann gilt

$$\lim_{k\to\infty} \int_\Omega f_k d\mu = \int_\Omega \lim_{k\to\infty} f_k d\mu. \tag{3.1}$$

Hierbei ist auch der Wert $+\infty$ als Limes zulässig.

BEWEIS. Beachte zunächst, daß $f := \lim_{k\to\infty} f_k$ aufgrund der Monotonie als Abbildung $\Omega \to \bar{\mathbb{R}}$ existiert. Zudem ist dieser Limes nichtnegativ und nach Satz 3.12 meßbar, also ist die rechte Seite von (3.1) als reelle Zahl oder $+\infty$ wohldefiniert.

Behauptung: Es existiert eine monotone Folge $(v_k)_{k\in\mathbb{N}}$ von Elementarfunktionen, sodaß $\sup_{k\in\mathbb{N}} v_k = f$ und $v_k \le f_k$ für alle $k \in \mathbb{N}$.

Wenn die Behauptung wahr ist, so gilt $\int_\Omega f d\mu = \sup_{k\in\mathbb{N}} \int_\Omega v_k d\mu$ nach Definition des Integrals, und außerdem $\int_\Omega v_k d\mu \le \int_\Omega f_k d\mu$ für alle $k \in \mathbb{N}$. Wegen $f_k \le f$ ist aber schließlich $\sup_{k\in\mathbb{N}} \int_\Omega f_k d\mu \le \int_\Omega f d\mu$, und der Satz folgt.

Beweis der Behauptung: Nach Satz 3.18 gibt es zu jedem $k \in \mathbb{N}$ eine Folge $(u_{k,m})_{m\in\mathbb{N}}$ von Elementarfunktionen, die monoton gegen f_k konvergiert. Setze nun

$$v_m := \max\{u_{1,m}, u_{2,m}, \dots, u_{m,m}\}.$$

Dann ist v_m eine Elementarfunktion, und da für alle $k, m \in \mathbb{N}$ gilt $u_{k,m} \le u_{k,m+1}$, ist $v_m \le v_{m+1}$ für alle $m \in \mathbb{N}$. Wegen $u_{k,m} \le f_k$ und der Monotonie der Folge $(f_k)_{k\in\mathbb{N}}$ folgt zudem $v_m \le f_m$ für alle $m \in \mathbb{N}$.

Es bleibt noch die Konvergenz gegen f zu zeigen. Wegen $v_m \le f_m \le f$ ist klar, daß $\sup_{m\in\mathbb{N}} v_m \le f$. Nach Definition von v_m gilt ferner $v_m \ge u_{k,m}$ für alle $k \le m$ und deshalb für jedes $k \in \mathbb{N}$

$$f_k = \sup_{m\in\mathbb{N}} u_{k,m} \le \sup_{m\in\mathbb{N}} v_m.$$

Da $k \in \mathbb{N}$ beliebig war, folgt $\sup_{m\in\mathbb{N}} v_m \geq f$. □

BEMERKUNG 3.32. Durch Umdrehen des Vorzeichens ist klar, daß der Satz auch im Falle nichtpositiver, monoton fallender Folgen gültig bleibt.

Für den Satz von der dominierten Konvergenz brauchen wir noch ein vorbereitendes Resultat:

SATZ 3.33 (Lemma von FATOU). Sei $(f_k)_{k\in\mathbb{N}}$ eine Folge nichtnegativer meßbarer Funktionen $\Omega \to \bar{\mathbb{R}}$. Dann gilt

$$\int_\Omega \liminf_{k\to\infty} f_k \leq \liminf_{k\to\infty} \int_\Omega f_k d\mu.$$

BEWEIS. Nach Satz 3.12 sind mit den f_k auch $f := \liminf_{k\to\infty} f_k$ sowie, für $l \in \mathbb{N}$, $g_l := \inf_{k\geq l} f_k$ meßbar. Offenbar steigen die g_l monoton mit l, und wegen $\liminf_{k\to\infty} = \lim_{l\to\infty} \inf_{k\geq l}$ ist $\lim_{l\to\infty} g_l = f$. Wir können also den Satz von der monotonen Konvergenz verwenden und erhalten so

$$\int_\Omega f d\mu = \lim_{l\to\infty} \int_\Omega g_l d\mu.$$

Da aber nach Definition $g_l \leq f_k$ für alle $k \geq l$, folgt mit der Monotonie des Integrals

$$\int_\Omega g_l d\mu \leq \inf_{k\geq l} \int_\Omega f_k d\mu,$$

und die Behauptung folgt, wenn wir auf beiden Seiten den Limes $l \to \infty$ bilden. □

Daß im Lemma von FATOU die Ungleichung nicht durch eine Gleichheit ersetzt werden kann, sieht man anhand des folgenden Beispiels (vgl. [**2**, Beispiel 5.27]):

Die Folge $f_k : (0,\infty) \to \mathbb{R}$ gegeben durch

$$f_k(x) = \begin{cases} k & \text{falls } x \in (0, \frac{1}{k}), \\ 0 & \text{sonst} \end{cases}$$

konvergiert punktweise gegen null, aber das Integral von f_k bzgl. des LEBESGUE-Maßes ist eins für alle $k \in \mathbb{N}$, also gilt das Lemma von FATOU in der Form $0 \leq 1$.

3.6.2. Fast überall bestehende Eigenschaften. Wir erinnern uns: Eine *μ-Nullmenge* ist eine Teilmenge einer meßbaren Menge $A \in \mathcal{A}$ mit $\mu(A) = 0$. In der Übung zeigen Sie: Das Integral einer Funktion über eine meßbare Nullmenge ist stets null. Bei der Integration scheint es also auf Nullmengen nicht anzukommen. Für die Integrationstheorie genügt es also, wenn eine Funktion eine bestimmte Eigenschaft nur außerhalb einer Nullmenge hat.

Wir sagen deshalb, eine Funktion habe eine Eigenschaft *μ-fast überall*, falls es eine μ-Nullmenge N gibt, sodaß $f(x)$ die Eigenschaft hat für alle $x \in \Omega \setminus N$.

Beispiele: Zwei Funktionen $f, g : \Omega \to \bar{\mathbb{R}}$ sind μ-fast überall gleich, wenn $\{f \neq g\}$ eine μ-Nullmenge ist; f ist positiv μ-fast überall, falls $\{f \leq 0\}$ eine μ-Nullmenge ist; eine Funktionenfolge $(f_k)_{k\in\mathbb{N}}$ konvergiert μ-fast überall gegen f, falls es eine Nullmenge N gibt, sodaß $\lim_{k\to\infty} f_k(x) = f(x)$ für alle $x \in \Omega \setminus N$, usw.

Man beachte unbedingt, daß „fast überall" von μ abhängt: Zum Beispiel sind die beiden Funktionen $f \equiv 0$, $g = \sin$ auf dem Maßraum $\mathbb{R}$ mit der BOREL-σ-Algebra LEBESGUE-fast überall verschieden, aber δ_0-fast überall gleich.

PROPOSITION 3.34. *Sei $f : \Omega \to \bar{\mathbb{R}}$ meßbar und nichtnegativ. Dann ist $\int_\Omega f d\mu = 0$ genau dann, wenn $f = 0$ μ-fast überall.*

BEWEIS. Falls $f = 0$ μ-fast überall, so gibt es eine meßbare Nullmenge $N \subset \Omega$, sodaß $f(x) = 0$ für alle $x \in \Omega \setminus N$. Wegen $f = f\chi_N + f\chi_{\Omega\setminus N}$ ist dann

$$\int_\Omega f d\mu = \int_{\Omega\setminus N} f d\mu + \int_N f d\mu = 0,$$

denn der erste Term ist null (da $f\restriction_{\Omega\setminus N} = 0$) und der zweite Term ebenfalls (s. Übung).

Sei umgekehrt $\int_\Omega f d\mu = 0$ und $N := \{f > 0\}$. Da f nichtnegativ ist, haben wir $\Omega \setminus N = \{f = 0\}$. Da f meßbar ist, ist auch N meßbar, und wir haben

$$0 = \int_\Omega f d\mu = \int_{\{f=0\}} f d\mu + \int_N f d\mu = \int_N f d\mu.$$

Sei $m \in \mathbb{N}$ beliebig, dann ist $N_m := \{f \geq \frac{1}{m}\} \subset N$, also insbesondere

$$\mu(N_m) = \int_{N_m} 1 d\mu = m \int_{N_m} \frac{1}{m} d\mu \leq m \int_{N_m} f d\mu \leq m \int_N f d\mu = 0.$$

Nun ist aber $N = \bigcup_{m\in\mathbb{N}} N_m$ und deshalb $\mu(N) \leq \sum_{m\in\mathbb{N}} \mu(N_m) = 0$, wie behauptet. □

Proposition 3.35. *Seien $f, g : \Omega \to \bar{\mathbb{R}}$ meßbar und $f = g$ μ-fast überall. Ist f integrierbar bzgl. μ, so auch g, und dann gilt*

$$\int_\Omega f d\mu = \int_\Omega g d\mu.$$

Beweis. Unter den angegebenen Voraussetzungen ist $f^\pm = g^\pm$ fast überall, denn die Menge $\{f^\pm \neq g^\pm\}$ ist eine Teilmenge der Nullmenge $N := \{f \neq g\}$. Nach Übung ist

$$\int_N f^\pm d\mu = \int_N g^\pm d\mu = 0,$$

also ist

$$\int_\Omega g^\pm d\mu = \int_{\Omega\setminus N} g^\pm d\mu = \int_{\Omega\setminus N} f^\pm d\mu = \int_\Omega f^\pm d\mu < \infty$$

nach Voraussetzung der Integrierbarkeit von f, also ist auch g integrierbar, und zugleich folgt die Gleichheit der Integrale. □

Proposition 3.36. *Ist $f : \Omega \to \bar{\mathbb{R}}$ integrierbar bzgl. μ, so nimmt f μ-fast überall reelle Werte an.*

Beweis. Wäre dies nicht der Fall, so wäre $\{f = +\infty\} \cup \{f = -\infty\}$ eine meßbare Menge positiven Maßes. O.B.d.A. sei $\mu(\{f = +\infty\}) > 0$. Dann aber hätten wir

$$\int_\Omega f^+ d\mu \geq \int_{\{f=+\infty\}} +\infty d\mu = +\infty \cdot \mu(\{f = +\infty\}) = +\infty,$$

im Widerspruch zur Integrierbarkeit von f. □

Proposition 3.37. *Sei $\mathcal{F}$ eine Menge von meßbaren Funktionen $\Omega \to \bar{\mathbb{R}}$. Für $f, g \in \mathcal{F}$ setze $f \equiv g$, falls $f = g$ μ-fast überall. Dann ist dadurch eine Äquivalenzrelation auf $\mathcal{F}$ definiert.*

Beweis. Es ist klar, daß $f = f$ fast überall, und daß $f = g$ fast überall genau dann, wenn $g = f$ fast überall. Gelte $f = g$ fast überall und $g = h$ fast überall. Dann gibt es meßbare Nullmengen N_1 und N_2, sodaß $f(x) = g(x)$ für alle $\Omega \setminus N_1$ und $g(x) = h(x)$ für alle $x \in \Omega \setminus N_2$. Da die abzählbare Vereinigung von Nullmengen wieder eine Nullmenge ist (da $\mu(\bigcup_{k\in\mathbb{N}} N_k) \leq \sum_{k\in\mathbb{N}} \mu(N_k) = 0$), ist auch $N_1 \cup N_2$ eine Nullmenge, und für alle $x \in \Omega \setminus (N_1 \cup N_2)$ ist $f(x) = h(x)$. Also gilt $f = h$ fast überall. □

3.6.3. Dominierte Konvergenz.

Satz 3.38 (von der dominierten/majorisierten Konvergenz (Lebesgue)). Sei $(f_k)_{k\in\mathbb{N}}$ eine Folge meßbarer Funktionen $\Omega \to \bar{\mathbb{R}}$, die μ-fast überall (punktweise) konvergiert. Falls eine integrierbare Funktion $g : \Omega \to \bar{\mathbb{R}}$ existiert mit $|f_k| \leq g$ für alle $k \in \mathbb{N}$, so konvergiert $(f_k)_{k\in\mathbb{N}}$ μ-fast überall gegen eine integrierbare Funktion f, und es gilt

$$\lim_{k\to\infty} \int_\Omega f_k d\mu = \int_\Omega f d\mu.$$

BEWEIS. Wir bezeichnen mit N_1 eine meßbare Nullmenge dergestalt, daß $\lim_{k\to\infty} f_k(x)$ in $\bar{\mathbb{R}}$ existiert für alle $x \in \Omega \setminus N_1$. Nach Proposition 3.36 existiert zudem eine meßbare Nullmenge N_2, sodaß $g(x) < \infty$ für alle $x \in \Omega \setminus N_2$. Wir setzen nun

$$f(x) := \begin{cases} \lim_{k\to\infty} f_k(x) & \text{falls } x \in \Omega \setminus (N_1 \cup N_2), \\ 0 & \text{sonst.} \end{cases}$$

Dann ist (nach Korollar 3.13 und wegen der Meßbarkeit von $N_1 \cup N_2$) f meßbar, und $\lim_{k\to\infty} f_k = f$ fast überall. Außerdem gilt $|f| \le g$ und damit auch $f^{\pm} \le g$, sodaß aus der Integrierbarkeit von g diejenige von f folgt. Mit dem selben Argument sieht man, daß auch jedes f_k integrierbar ist. Also sind alle im Satz auftretenden Integrale wohldefiniert.

Für jedes $k \in \mathbb{N}$ ist die Funktion $g_k := |f - f_k|$ integrierbar, ebenso die Funktion $h := |f| + g$. Zudem ist $g_k \le |f| + |f_k| \le h$ für jedes $k \in \mathbb{N}$, sodaß $(h - g_k)_{k\in\mathbb{N}}$ eine Folge nichtnegativer meßbarer Funktionen ist. Also ist nach dem Lemma von FATOU

$$\int_\Omega \liminf_{k\to\infty} (h - g_k) d\mu \le \liminf_{k\to\infty} \int_\Omega (h - g_k) d\mu = \int_\Omega h d\mu - \limsup_{k\to\infty} \int_\Omega g_k d\mu.$$

Nun ist aber wegen der Konvergenz der f_k gegen f fast überall auch $\lim_{k\to\infty} g_k = 0$ fast überall und somit $\liminf_{k\to\infty}(h - g_k) = h$ fast überall. Nach Proposition 3.35 folgt

$$\int_\Omega \liminf_{k\to\infty} (h - g_k) d\mu = \int_\Omega h d\mu,$$

und aus den letzten beiden Identitäten erhalten wir $\limsup_{k\to\infty} \int_\Omega g_k d\mu = 0$, mithin

$$\lim_{k\to\infty} \int_\Omega |f - f_k| d\mu = 0.$$

Die Behauptung folgt schließlich aus der Ungleichung

$$\left| \int_\Omega f d\mu - \int_\Omega f_k d\mu \right| = \left| \int_\Omega f - f_k d\mu \right| \le \int_\Omega |f - f_k| d\mu.$$

□

3.7. LEBESGUE-Räume

Wir untersuchen nun wichtige Klassen von Funktionen, nämlich für $1 \le p < \infty$ diejenigen meßbaren Funktionen f, für die $|f|^p$ integrierbar ist, und für $p = \infty$ die wesentlich beschränkten Funktionen. Für μ ein Wahrscheinlichkeitsmaß und $p = 2$ erhält man so z.B. die Zufallsvariablen mit endlicher Varianz. Wir zeigen, daß diese Mengen von Funktionen – sofern man μ-fast überall gleiche Funktionen miteinander identifiziert – einen normierten Vektorraum bilden. Mithilfe der Konvergenzsätze kann man sogar zeigen, daß diese Räume vollständig, also *Banachräume*, sind. Solche Räume werden systematisch in der Funktionalanalysis untersucht und sind grundlegend für fast alle Bereiche der (im weitesten Sinne) angewandten Mathematik, wie partielle Differentialgleichungen, mathematische Physik, Numerik, Stochastik und Finanzmathematik.

Wie gewohnt sei im gesamten Abschnitt $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und $\bar{\mathbb{R}}$ mit der BOREL-σ-Algebra versehen.

DEFINITION 3.39. Sei $1 \le p < \infty$.

(1) Die Menge $\mathcal{L}^p(\Omega)$ ist definiert als die Menge aller meßbaren Funktionen $f : \Omega \to \bar{\mathbb{R}}$ mit der Eigenschaft

$$\int_\Omega |f|^p d\mu < \infty.$$

In diesem Falle setzen wir

$$\|f\|_p := \left(\int_\Omega |f|^p d\mu \right)^{1/p}.$$

(2) Die Menge $\mathcal{L}^\infty(\Omega)$ ist definiert als die Menge aller meßbaren Funktionen $f : \Omega \to \bar{\mathbb{R}}$, die *wesentlich beschränkt* sind, d.h.: Es existiert eine reelle Zahl $M > 0$ dergestalt, daß $|f| \le M$ μ-fast überall. In diesem Falle setzen wir

$$\|f\|_\infty := \inf\{M > 0 : |f| \le M \ \mu\text{-fast überall}\}.$$

Bemerkung 3.40. (1) Ist f meßbar, so ist auch $|f|^p$ meßbar und nichtnegativ, also ist das Integral in obiger Definition stets in $\bar{\mathbb{R}}$ wohldefiniert.

(2) Die Notation suggeriert, daß es sich bei $\|\cdot\|_p$ und $\|\cdot\|_\infty$ um Normen handele. Dies ist nicht der Fall, da aus $\|f\|_p = 0$ nicht folgt $f = 0$: Wähle dafür eine Funktion, die zwar μ-fast überall gleich null ist, aber eben nicht überall. Zur Lösung dieses Problems identifizieren wir bald durch Äquivalenzklassenbildung Funktionen miteinander, die fast überall übereinstimmen.

(3) Die Notation $\|\cdot\|_\infty$ erklärt sich aus folgendem Sachverhalt (Übung): Ist $\mu(\Omega) < \infty$ und $f \in \mathcal{L}^\infty(\Omega)$, so ist auch $f \in \mathcal{L}^p(\Omega)$ für jedes $1 \le p < \infty$, und es gilt

$$\lim_{p\to\infty} \|f\|_p = \|f\|_\infty.$$

3.7.1. Die Ungleichungen von Hölder und Minkowski. Sei $1 \le p \le \infty$. Im Falle $1 < p < \infty$ heißt die eindeutig bestimmte Zahl $1 < p' < \infty$ mit $\frac{1}{p} + \frac{1}{p'} = 1$ der *zu p konjugierte Exponent.* Für $p = 1$ setzen wir außerdem $p' = +\infty$ und für $p = +\infty$ entsprechend $p' = 1$. Offenbar gilt stets $(p')' = p$.

Satz 3.41 (Hölder-Ungleichung). Sei $1 \le p \le \infty$, $f \in \mathcal{L}^p(\Omega)$, $g \in \mathcal{L}^{p'}(\Omega)$. Dann ist $fg \in \mathcal{L}^1(\Omega)$, und es gilt

$$\|fg\|_1 \le \|f\|_p \|g\|_{p'}.$$

Beweis. Nach Satz 3.11 ist mit f und g auch das Produkt fg meßbar.

Sei zunächst $p = 1$ (also $p' = \infty$) und N eine meßbare Nullmenge, sodaß gilt $|g(x)| \le \|g\|_\infty$ für alle $x \in \Omega \setminus N$. Dann folgt

$$\|fg\|_1 = \int_\Omega |fg| d\mu = \int_{\Omega\setminus N} |f||g| d\mu \le \|g\|_\infty \int_{\Omega\setminus N} |f| d\mu = \|f\|_1 \|g\|_\infty,$$

wie behauptet. Der Fall $p = \infty$ folgt durch Vertauschung der Rollen von f und g.

Sei nun $1 < p < \infty$, so ist auch $1 < p' < \infty$. Im Falle $\|f\|_p = 0$ oder $\|g\|_{p'} = 0$ folgt gemäß Proposition 3.34 $f = 0$ bzw. $g = 0$ fast überall, und dann sind beide Seiten der zu zeigenden Ungleichung null. Wir dürfen also annehmen $\|f\|_p, \|g\|_{p'} > 0$.

Betrachte die Funktion $(-1, +\infty) \to \mathbb{R}$, $s \mapsto (1+s)^{1/p}$, die wegen $p > 1$ konkav ist; daher ist sie nicht größer als ihre lineare Approximation im Punkt $s = 0$, welche durch $1 + \frac{s}{p}$ gegeben ist:

$$(1+s)^{1/p} \le 1 + \frac{s}{p}$$

bzw. mit $t := 1 + s$ und der Definition von p'

$$t^{1/p} \le \frac{t}{p} + \frac{1}{p'}$$

für alle $t > 0$. Sind $x, y > 0$ reelle Zahlen, so folgt daraus mit der Wahl $t := \frac{x}{y}$

$$x^{1/p} y^{1/q} \le \frac{x}{p} + \frac{y}{p'}.$$

Die Wahl $x := \frac{|f|^p}{\|f\|_p^p}$, $y := \frac{|g|^{p'}}{\|g\|_{p'}^{p'}}$ liefert nun

$$|fg| \le \|f\|_p^{1-p} \|g\|_{p'} \frac{1}{p} |f|^p + \|f\|_p \|g\|_{p'}^{1-p'} \frac{1}{p'} |g|^{p'}.$$

Integration beider Seiten ergibt schließlich

$$\|fg\|_1 \le \frac{1}{p}\|f\|_p\|g\|_{p'} + \frac{1}{p'}\|f\|_p\|g\|_{p'} = \|f\|_p\|g\|_{p'}.$$

□

KOROLLAR 3.42. Sei $\mu(\Omega) < \infty$ und $f \in \mathcal{L}^p(\Omega)$ für ein $1 \le p \le \infty$. Dann ist $f \in \mathcal{L}^1(\Omega)$, und es gilt

$$\|f\|_1 \le \mu(\Omega)^{1/p'}\|f\|_p.$$

BEWEIS. Wir wenden die HÖLDER-Ungleichung mit p und p' an:

$$\|f\|_1 = \int_\Omega 1 \cdot |f| d\mu \le \|1\|_{p'}\|f\|_p = \mu(\Omega)^{1/p'}\|f\|_p.$$

□

Eine wichtige Anwendung der HÖLDER-Ungleichung ist die folgende, nach MINKOWSKI benannte Ungleichung, die nichts anderes ist als die Dreiecksungleichung für $\|\cdot\|_p$. Insbesondere folgt aus ihr, daß die Summe zweier Funktionen in $\mathcal{L}^p$ wieder in $\mathcal{L}^p$ ist.

SATZ 3.43 (MINKOWSKI-Ungleichung). Sei $1 \le p \le \infty$ und $f, g \in \mathcal{L}^p(\Omega)$ dergestalt, daß $f + g$ überall in Ω definiert ist. Dann gilt

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

Mit der Voraussetzung, daß $f+g$ überall definiert ist, soll ausgeschlossen werden, daß an einzelnen Punkten $f(x) = +\infty$ und $g(x) = -\infty$, in welchem Falle die Summe nicht definiert wäre. Nach Proposition 3.36 kann dies aber nur auf einer Nullmenge passieren, die für den Wert von $\|\cdot\|_p$ ohnehin unerheblich ist. Wir kommen bald darauf zurück.

BEWEIS. Für $p = \infty$ seien N_1 und N_2 zwei Nullmengen, sodaß $|f(x)| \le \|f\|_\infty$ für alle $x \ne N_1$ und $|g(x)| \le \|g\|_\infty$ für alle $x \ne N_2$. Für $x \in \Omega \setminus (N_1 \cup N_2)$ gilt dann nach Dreiecksungleichung

$$|(f + g)(x)| \le |f(x)| + |g(x)| \le \|f\|_\infty + \|g\|_\infty,$$

woraus bereits $\|f + g\|_\infty \le \|f\|_\infty + \|g\|_\infty$ folgt.

Für $p = 1$ beachte, wieder nach Dreiecksungleichung, $|f + g| \le |f| + |g|$, und die Behauptung folgt mit der Monotonie und Linearität des Integrals.

Sei nun also $1 < p < \infty$. Wir verwenden die Abschätzung

$$|f + g|^p \le (|f| + |g|)^p \le [2\max\{|f|, |g|\}]^p = 2^p \max\{|f|^p, |g|^p\} \le 2^p(|f|^p + |g|^p),$$

die sofort (wegen der vorausgesetzten Integrierbarkeit von $|f|^p$ und $|g|^p$) impliziert, daß $f + g \in \mathcal{L}^p$.

Nun haben wir aber

$$\int_\Omega |f + g|^p d\mu \le \int_\Omega |f||f + g|^{p-1} d\mu + \int_\Omega |g||f + g|^{p-1} d\mu.$$

Wenden wir nun auf die beiden Integrale jeweils die HÖLDER-Ungleichung mit den konjugierten Exponenten p und $p' = \frac{p}{p-1}$ an:

$$\int_\Omega |f + g|^p d\mu \le \|f\|_p\|f + g\|_p^{p-1} + \|g\|_p\|f + g\|_p^{p-1},$$

woraus (nach Kürzen von $\|f + g\|_p^{p-1}$ auf beiden Seiten) bereits die Behauptung folgt. □

3.7.2. Der normierte Raum $L^p(\Omega)$. Unsere Funktionenmengen $\mathcal{L}^p(\Omega)$ sind bisher durch zwei Mängel aufgefallen: Die Summe zweier Funktionen in $\mathcal{L}^p$ ist unter Umständen nicht wohldefiniert (wenn z.B. $f(x) = +\infty$ und $g(x) = -\infty$ an einem Punkt $x \in \Omega$), und $\|f\|_p = 0$ impliziert nicht $f = 0$ (denn $\|f\|_p = 0$ kann auch dann der Fall sein, wenn $f = 0$ nur fast überall, aber nicht überall gilt). Beide Probleme lassen sich leicht in der folgenden Weise beheben:

Definition 3.44. Sei $1 \leq p \leq \infty$. Betrachte auf $\mathcal{L}^p(\Omega)$ die Äquivalenzrelation (siehe Proposition 3.37)

$$f \equiv g \text{ genau dann, wenn } f = g \ \mu\text{-fast überall.}$$

Dann ist der Lebesgue-*Raum mit Exponent p* definiert als die Menge der Äquivalenzklassen bezüglich dieser Relation. Man bezeichnet ihn mit $L^p(\Omega)$.

Soll verdeutlicht werden, bezüglich welchen Maßes gearbeitet wird, schreibt man auch $L^p(\Omega; \mu)$.

Die Definition ist gewöhnungsbedürftig. Üblicherweise faßt man ein Element $f \in L^p(\Omega)$ als eine Abbildung $\Omega \to \mathbb{R}$ auf, die nur μ-fast überall definiert ist. So sind etwa die Funktionen $\chi_{\mathbb{Q}}$ und 0 als Abbildungen $\mathbb{R} \to \mathbb{R}$ Lebesgue-fast überall gleich, also werden sie als Elemente von $L^p(\mathbb{R})$ nicht unterschieden. Wir sprechen deshalb oft von einer „Funktion" in L^p, obwohl es sich strenggenommen um eine Äquivalenzklasse von Funktionen, nicht aber selbst um eine Funktion handelt.

Ist μ das Lebesgue-Maß, so ist für eine L^p-Funktion eine Aussage der Form „$f(x) = y$" nicht sinnvoll, denn man kann f auf der Nullmenge $\{x\}$ beliebig abändern und erhält wieder dieselbe L^p-Funktion.

Auf L^p kann man nun in natürlicher Weise die üblichen Rechenoperationen sowie das Integral definieren: Ist $[f] \in L^p$ die zu $f \in \mathcal{L}^p$ gehörige Äquivalenzklasse und $\alpha \in \mathbb{R}$, so setzt man $\alpha[f] := [\alpha f]$; für $f, g \in \mathcal{L}^p$ ist gemäß Proposition 3.36 $f, g \neq \pm\infty$ μ-fast überall, und wir dürfen daher annehmen, daß f und g nur reelle Werte annehmen. Dann ist $f + g$ überall wohldefiniert und wir setzen $[f] + [g] = [f + g]$. Ist $f \in \mathcal{L}^p$ und $g \in \mathcal{L}^{p'}$, so ist nach Hölder $fg \in \mathcal{L}^1$, und entsprechend ist $[f][g] := [fg] \in L^1$. Man überzeugt sich leicht, daß diese Definitionen von der Wahl des Repräsentanten unabhängig sind (sind zum Beispiel $[f] = [\tilde{f}]$ und $[g] = [\tilde{g}]$, so ist $f = \tilde{f}$ und $g = \tilde{g}$ fast überall, und dann ist auch $f + g = \tilde{f} + \tilde{g}$ fast überall, somit $[f + g] = [\tilde{f} + \tilde{g}]$, also $[f] + [g] = [\tilde{f}] + [\tilde{g}]$).

Schließlich setzen wir

$$\int_\Omega [f] d\mu := \int_\Omega f d\mu.$$

Diese Definition ist nach Proposition 3.35 ebenfalls unabhängig von der Wahl des Repräsentanten. Damit ist aber auch $\|[f]\|_p$ definiert für jedes $[f] \in L^p$.

Wir verzichten künftig auf die eckigen Klammern, die die Äquivalenzklassenbildung kennzeichnen.

Wir erinnern uns an den Begriff des *normierten Raums*: Dabei handelt es sich um einen (hier stets reellen) Vektorraum X mit einer Norm $\|\cdot\| : X \to \mathbb{R}_0^+$, sodaß $\|x\| = 0$ genau dann, wenn $x = 0$, $\|\alpha x\| = |\alpha| \|x\|$ für alle $\alpha \in \mathbb{R}$ und $x \in X$, und $\|x + y\| \leq \|x\| + \|y\|$ für alle $x, y \in X$.

Eine Folge $(x_k)_{k \in \mathbb{N}} \subset X$ konvergiert gegen $x \in X$ genau dann, wenn $\lim_{k \to \infty} \|x_k - x\| = 0$, und heißt Cauchy, wenn zu jedem $\epsilon > 0$ ein $N \in \mathbb{N}$ existiert, sodaß für alle $k, m \geq N$ gilt $\|x_k - x_m\| < \epsilon$. Ein normierter Raum heißt *vollständig*, wenn in ihm jede Cauchyfolge konvergiert. Ein vollständiger normierter Raum heißt *Banachraum*.

Lemma 3.45. *Sei X normierter Raum. Dann ist X vollständig genau dann, wenn in X jede absolut konvergente Reihe konvergent ist.*

Beweis. Sei X vollständig und $(x_k)_{k\in N} \subset X$ absolut konvergent, also $\sum_{k=1}^{\infty} \|x_k\| < \infty$. Sei $\epsilon > 0$, dann gilt für hinreichend große $l, m \in \mathbb{N}$:

$$\left\| \sum_{k=1}^{l} x_k - \sum_{k=1}^{m} x_k \right\| = \left\| \sum_{k=l+1}^{m} x_k \right\| \le \sum_{k=l+1}^{m} \|x_k\| < \epsilon$$

aufgrund der Konvergenz. Also ist die Folge $s_m := \sum_{k=1}^{m} x_k$ der Partialsummen Cauchy und damit nach Voraussetzung konvergent.

Sei umgekehrt jede absolut konvergente Reihe konvergent und $(x_k)_{k\in N}$ Cauchy. Für $s \in \mathbb{N}$ wähle $N_s \in \mathbb{N}$ so groß, daß $\|x_l - x_m\| < 2^{-s}$ für alle $l, m \ge N_s$. O.B.d.A. ist die Folge $(N_s)_{s\in\mathbb{N}}$ streng monoton wachsend. Die Teilfolge $(x_{N_s})_{s\in\mathbb{N}}$ erfüllt damit

$$\|x_{N_{s+1}} - x_{N_s}\| < 2^{-s}$$

für alle $s \in \mathbb{N}$. Setzen wir $y_s := x_{N_{s+1}} - x_{N_s}$, so ist also $\sum_{s=1}^{\infty} \|y_s\| < \infty$.

Nach Voraussetzung konvergiert nun aber $\sum_{s=1}^{\infty} y_s$ gegen ein $y \in X$, also haben wir

$$0 = \lim_{l\to\infty} \left\| y - \sum_{s=1}^{l} y_s \right\| = \|y - (x_{N_{l+1}} - x_{N_1})\|,$$

wobei wir die Definition der y_s verwendet haben. Dies bedeutet aber, daß die Teilfolge $(x_{N_s})_{s\in\mathbb{N}}$ konvergiert (nämlich gegen $y + x_{N_1}$). Man sieht aber leicht, daß eine Cauchyfolge, die eine konvergente Teilfolge besitzt, konvergent ist, was zu zeigen war. □

Satz 3.46. Sei $1 \le p \le \infty$. Dann ist $L^p(\Omega)$ mit der Norm $\|\cdot\|_p$ ein Banachraum.

Beweis. Da für $\alpha \in \mathbb{R}$ und $f, g \in L^p(\Omega)$ auch $\alpha f \in L^p(\Omega)$ sowie $f + g \in L^p$ (nach Minkowski!), ist L^p ein Vektorraum. Für $p < \infty$ ist nach Proposition 3.34 $\|f\|_p = 0$ genau dann, wenn $f = 0$ (als L^p-Funktion, d.h. bis auf eine Nullmenge!); für $p = \infty$ folgt dies direkt aus der Definition von $\mathcal{L}^\infty$. Nach Definition von $\|\cdot\|_p$ ist die Eigenschaft $\|\alpha f\|_p = |\alpha|\|f\|_p$ klar. Die Dreiecksungleichung $\|f+g\|_p \le \|f\|_p + \|g\|_p$ ist genau die Minkowski-Ungleichung.

Es bleibt die Vollständigkeit zu beweisen. Sei zunächst $p < \infty$. Nach vorigem Lemma genügt es dafür zu zeigen, daß jede absolut konvergente Reihe konvergent ist. Sei also $\sum_{k=1}^{\infty} \|f_k\|_p < \infty$. Da f_k (genauer gesagt: jeder Repräsentant von f_k in $\mathcal{L}^p$) meßbar ist für alle $k \in \mathbb{N}$, so auch $g := \sum_{k=1}^{\infty} |f_k|$ gemäß Satz 3.12 und erst recht $g_k := \sum_{l=1}^{k} |f_l|$. Nach Dreiecks- bzw. Minkowski-Ungleichung ist mit f_l auch $g_k \in L^p(\Omega)$ für jedes $k \in \mathbb{N}$, wobei $\|g_k\|_p \le \sum_{l=1}^{k} \|f_l\|_p < \infty$.

Nun sind die g_k (für jede Wahl von Repräsentanten für f_l) nichtnegativ und monoton steigend, und es gilt $g = \sup_{k\in\mathbb{N}} g_k$, weshalb mit dem Satz von der monotonen Konvergenz gilt

$$\int_\Omega g^p d\mu = \sup_{k\in\mathbb{N}} \int_\Omega g_k^p d\mu \le \left(\sum_{l=1}^{\infty} \|f_l\|_p \right)^p < \infty.$$

Nach Proposition 3.36 ist daher $g = \sum_{l=1}^{\infty} |f_l| < \infty$ μ-fast überall. Da $\mathbb{R}$ vollständig ist, ist die absolut konvergente Reihe $f(x) := \sum_{l=1}^{\infty} f_l(x)$ konvergent für fast alle $x \in \Omega$, und nach Satz 3.12 ist f meßbar.

Es bleibt zu zeigen $f \in L^p(\Omega)$ und $f = \sum_{l=1}^{\infty} f_l$ bzgl. der Norm in L^p. Die erste Aussage folgt sofort wegen

$$\int_\Omega |f|^p d\mu \le \int_\Omega \left(\sum_{l=1}^{\infty} |f_l| \right)^p d\mu = \int_\Omega g^p d\mu < \infty.$$

Für die L^p-Konvergenz der Reihe setzen wir

$$h_k := \left| f - \sum_{l=1}^{k} f_l \right|^p = \left| \sum_{l=k+1}^{\infty} f_l \right|^p,$$

und wir müssen zeigen $\int_\Omega h_k d\mu \to 0$ mit $k \to \infty$. Wir wissen bereits (nach Definition von f!), daß $h_k \to 0$ fast überall, und außerdem gilt

$$0 \le h_k = \left| \sum_{l=k+1}^{\infty} f_l \right|^p \le \left(\sum_{l=1}^{\infty} |f_l| \right)^p = g^p$$

für alle $x \in \Omega$. Da aber g^p eine integrierbare Majorante für die Folge $(h_k)_{k\in\mathbb{N}}$ ist, können wir den Satz von der dominierten Konvergenz anwenden und erhalten in der Tat $\int_\Omega h_k d\mu \to 0$.

Betrachten wir schließlich den Fall $p = \infty$. Sei also $(f_k)_{k\in\mathbb{N}}$ Cauchy in L^∞ und bezeichne mit f_k gleichzeitig[5] einen beliebigen Repräsentanten in $\mathcal{L}^\infty$. Für $l, m \in \mathbb{N}$ sei $N_{l,m}$ eine meßbare Nullmenge, sodaß $|f_l(x) - f_m(x)| \le \|f_l - f_m\|_\infty$ für alle $x \in \Omega \setminus N_{l,m}$. Dann ist die abzählbare Vereinigung $N := \bigcup_{l,m\in\mathbb{N}} N_{l,m}$ wieder eine Nullmenge, und für alle $x \in \Omega \setminus N$ gilt $|f_l(x) - f_m(x)| \le \|f_l - f_m\|_\infty$ für alle $l, m \in \mathbb{N}$. Wir dürfen außerdem annehmen, daß auf $\Omega \setminus N$ alle Funktionen f_k endliche Werte annehmen.

Daher ist für jedes $x \in \Omega \setminus N$ die Folge $(f_k(x))_{k\in\mathbb{N}}$ Cauchy in $\mathbb{R}$ und damit konvergent. Definieren wir

$$f(x) := \begin{cases} \lim_{k\to\infty} f_k(x), & x \in \Omega \setminus N \\ 0, & x \in N, \end{cases}$$

so ist f nach Korollar 3.13 meßbar.

Sei $\epsilon > 0$. Dann existiert $M \in \mathbb{N}$, sodaß für alle $l, m \ge M$ und für alle $x \in \Omega \setminus N$ gilt

$$|f_m(x) - f_l(x)| \le \|f_m - f_l\|_\infty < \epsilon.$$

Lassen wir in dieser Ungleichung $m \to \infty$, so erhalten wir für jedes $l \ge M$ und jedes $x \in \Omega \setminus N$

$$|f(x) - f_l(x)| \le \epsilon.$$

Hieraus folgt einerseits die Beschränktheit von f (denn für $x \in \Omega \setminus N$ ist $|f(x)| \le \|f_M\|_\infty + \epsilon$, und auf N ist $f = 0$), und außerdem die gewünschte Konvergenz $f_l \to f$ in $L^\infty(\Omega)$. □

[5] Sie finden, dies sei schlechter Stil oder gar falsch? Sie haben vollkommen recht. Aber es ist üblich und sehr praktisch, in der Notation nicht zwischen einer L^p-Funktion und ihren Repräsentanten in $\mathcal{L}^p$ zu unterscheiden.

KAPITEL 4

Produktmaße

Das kartesische Produkt zweier (oder mehrerer) Maßräume kann in natürlicher Weise mit einer Produkt-σ-Algebra und einem Produktmaß versehen werden. Zum Beispiel ist $\mathbb{R}^2$ das Produkt $\mathbb{R} \times \mathbb{R}$, und es stellt sich heraus, daß das Produkt der eindimensionalen LEBESGUE-Maße auf $\mathbb{R}$ genau das zweidimensionale LEBESGUE-Maß auf $\mathbb{R}^2$ ist. Von größter Bedeutung ist der Satz von FUBINI, der es erlaubt, sukzessive und in beliebiger Reihenfolge zu integrieren.

Wir beschränken uns hier der Übersichtlichkeit halber auf Produkte zweier Maßräume, bemerken aber, daß die Theorie problemlos auf beliebige endliche Produkte übertragbar ist.

4.1. Produkte von Maßräumen

Seien $(\Omega_1, \mathcal{A}_1, \mu_1)$ und $(\Omega_2, \mathcal{A}_2, \mu_2)$ zwei Maßräume. Wir kennen bereits das kartesische Produkt der Mengen Ω_1 und Ω_2, nämlich die Menge der geordneten Paare

$$\Omega_1 \times \Omega_2 := \{(x, y) : x \in \Omega_1, y \in \Omega_2\}.$$

Wir schreiben im folgenden $\Omega := \Omega_1 \times \Omega_2$.

4.1.1. Produkt-σ-Algebren.

DEFINITION 4.1 (Produkt-σ-Algebra). Die Produkt-σ-Algebra $\mathcal{A}_1 \otimes \mathcal{A}_2$ ist definiert als die von den Mengen $A_1 \times A_2$ erzeugte σ-Algebra mit $A_1 \in \mathcal{A}_1$ und $A_2 \in \mathcal{A}_2$.

PROPOSITION 4.2 (Projektionen). *Für $i = 1, 2$ seien die* Projektionen $p_i : \Omega_1 \times \Omega_2 \to \Omega_i$ *definiert durch $p_1(x, y) = x$ und $p_2(x, y) = y$. Dann ist $\mathcal{A} := \mathcal{A}_1 \otimes \mathcal{A}_2$ die kleinste σ-Algebra, bezüglich derer p_1 und p_2 beide meßbar sind.*

BEWEIS. Sei $A_1 \in \mathcal{A}_1$, so ist $p_1^{-1}(A_1) = A_1 \times \Omega_2 \in \mathcal{A}$, also ist p_1 meßbar. Ebenso ist p_2 meßbar. Ist nun $\mathcal{A}'$ irgendeine σ-Algebra auf $\Omega_1 \times \Omega_2$, bezüglich der p_1 und p_2 meßbar sind, so müssen alle Mengen der Form $A_1 \times \Omega_2$ und $\Omega_1 \times A_2$ für $A_1 \in \mathcal{A}_1$, $A_2 \in \Omega_2$ in $\mathcal{A}'$ liegen, und damit auch $A_1 \times A_2 = (A_1 \times \Omega_2) \cap (\Omega_1 \times A_2)$. Also ist $\mathcal{A}$ als die von Mengen dieser Gestalt erzeugte σ-Algebra eine Teilmenge von $\mathcal{A}'$. □

KOROLLAR 4.3. Sei $\mathcal{B}$ die BOREL-σ-Algebra auf $\mathbb{R}$. Dann ist $\mathcal{B} \otimes \mathcal{B}$ die BOREL-σ-Algebra auf $\mathbb{R}^2$.

BEWEIS. Seien $A_1, A_2 \in \mathcal{B}$, so liegt $A_1 \times A_2$ nach Definition in $\mathcal{B} \otimes \mathcal{B}$, sodaß diese Produkt-$\sigma$-Algebra insbesondere alle halboffenen Quader und damit die BOREL-σ-Algebra auf $\mathbb{R}^2$ enthält.

Umgekehrt genügt es gemäß voriger Proposition zu zeigen, daß p_1, p_2 bezüglich der BOREL-σ-Algebra auf $\mathbb{R}^2$ meßbar sind. Nach Proposition 3.4 genügt es dafür zu zeigen, daß das Urbild jedes halboffenen Intervalls unter p_1 bzw. p_2 meßbar ist. Das Urbild eines solchen Intervalls $[a, b)$ unter p_1 ist aber $[a, b) \times \mathbb{R}$, was in der BOREL-σ-Algebra auf $\mathbb{R}^2$ enthalten ist. Analog argumentiert man für p_2. □

4.1.2. Produktmaße. Wir wollen ein Maß finden, das jeder Menge der Form $A_1 \times A_2$ mit $A_1 \in \mathcal{A}_1$ und $A_2 \in \mathcal{A}_2$ den Wert $\mu_1(A_1)\mu_2(A_2)$ zuordnet. Zunächst einige Vorbereitungen:

Für $A \in \mathcal{A} := \mathcal{A}_1 \otimes \mathcal{A}_2$ und für $x \in \Omega_1$, $y \in \Omega_2$ setzen wir

$$A_x := \{z \in \Omega_2 : (x,z) \in A\} \subset \Omega_2, \quad A_y := \{z \in \Omega_1 : (z,y) \in A\} \subset \Omega_1.$$

Man nennt A_x den *x-Schnitt von A*.

Lemma 4.4. *Sei $A \in \mathcal{A}$ und $x \in \Omega_1$, $y \in \Omega_2$. Dann ist $A_x \in \mathcal{A}_2$ und $A_y \in \mathcal{A}_1$.*

Beweis. Wir zeigen nur den ersten Teil der Aussage, der zweite geht analog. Wir behaupten, daß die Mengen $A \subset \Omega$ mit der Eigenschaft $A_x \in \mathcal{A}_2$ eine σ-Algebra auf $\Omega_1 \times \Omega_2$ bilden. In der Tat, Ω hat diese Eigenschaft, denn $\Omega_x = \Omega_2$; Ist A dergestalt, daß $A_x \in \mathcal{A}_2$, so gilt auch $(\Omega \setminus A)_x = \Omega_2 \setminus A_x \in \mathcal{A}_2$; und sind A_k Mengen mit der fraglichen Eigenschaft ($k \in \mathbb{N}$), so gilt

$$\Big(\bigcup_{k\in\mathbb{N}} A_k\Big)_x = \bigcup_{k\in\mathbb{N}} (A_k)_x \in \mathcal{A}_2.$$

Weiterhin enthält diese σ-Algebra offenbar alle Mengen $A_1 \times A_2$ mit $A_1 \in \mathcal{A}_1$, $A_2 \in \mathcal{A}_2$, denn $(A_1 \times A_2)_x$ ist A_2, falls $x \in A_1$, und $\varnothing$ sonst.

Es folgt nach Definition der Produkt-σ-Algebra, daß $\mathcal{A}$ in der eben beschriebenen σ-Algebra enthalten ist, und damit die Behauptung. □

Wir erinnern uns: Ein Maß μ auf Ω heißt σ-endlich, wenn es eine Folge meßbarer Mengen Ω_k gibt mit $\mu(\Omega_k) < \infty$ und $\bigcup_{k\in\mathbb{N}} \Omega_k = \Omega$.

Lemma 4.5. *Seien μ_1 und μ_2 σ-endliche Maße auf Ω_1 bzw. Ω_2. Für jedes $A \in \mathcal{A} := \mathcal{A}_1 \otimes \mathcal{A}_2$ ist die Funktion $\Omega_1 \to \bar{\mathbb{R}}$, $x \mapsto \mu_2(A_x)$ meßbar bzgl. $\mathcal{A}_1$.*

Ebenso ist die Funktion $\Omega_2 \to \bar{\mathbb{R}}$, $y \mapsto \mu_1(A_y)$ meßbar bzgl. $\mathcal{A}_2$.

Beweis. Wir zeigen wieder nur die erste Aussage. Nehmen wir zunächst an $\mu_2(\Omega_2) < \infty$. Sei $\mathcal{D}$ die Menge aller $D \in \mathcal{A}$, für die $x \mapsto \mu_2(D_x)$ meßbar ist. Wir zeigen, daß $\mathcal{D}$ Dynkin ist: Es ist nämlich $\mu_2(\Omega_x) = \mu_2(\Omega_2)$ für alle $x \in \Omega_1$; ist $x \mapsto \mu_2(D_x)$ meßbar, so auch

$$x \mapsto \mu_2((\Omega \setminus D)_x) = \mu_2(\Omega_2 \setminus D_x) = \mu_2(\Omega_2) - \mu_2(D_x);$$

und ist $x \mapsto \mu_2((D_k)_x)$ meßbar für paarweise disjunkte D_k, $k \in \mathbb{N}$, so auch

$$x \mapsto \mu_2\Big((\bigcup_{k\in\mathbb{N}} D_k)_x\Big) = \mu_2\Big(\bigcup_{k\in\mathbb{N}} (D_k)_x\Big) = \sum_{k\in\mathbb{N}} \mu_2((D_k)_x),$$

wobei wir Satz 3.12 verwenden.

Seien weiter $A_1 \in \mathcal{A}_1$, $A_2 \in \mathcal{A}_2$, dann ist $A_1 \times A_2 \in \mathcal{D}$, denn

$$x \mapsto \mu_2((A_1 \times A_2)_x) = \mu_2(A_2)\chi_{A_1}(x)$$

ist meßbar. Das System der Mengen der Gestalt $A_1 \times A_2$ ist aber durchschnittsstabil und erzeugt $\mathcal{A}_1 \otimes \mathcal{A}_2$. Gemäß Satz 1.7 ist aber $\mathcal{A}_1 \otimes \mathcal{A}_2$ auch das von den Mengen $A_1 \times A_2$ erzeugte Dynkin-System, woraus $\mathcal{A}_1 \otimes \mathcal{A}_2 \subset \mathcal{D}$ folgt. Dies ist gerade die Behauptung.

Wir verzichten nun auf die Annahme $\mu_2(\Omega_2) < \infty$ und verlangen nur noch σ-Endlichkeit. Sei also $(\tilde{\Omega}_k)_{k\in\mathbb{N}}$ eine aufsteigende Folge meßbarer Mengen mit $\mu_2(\tilde{\Omega}_k) < \infty$ und $\bigcup_{k\in\mathbb{N}} \tilde{\Omega}_k = \Omega_2$. Für jedes $k \in \mathbb{N}$ ist $\tilde{\mu}_k := \mu_2(\cdot \cap \tilde{\Omega}_k)$ ein endliches Maß, sodaß nach dem bereits Bewiesenen die Abbildung $x \mapsto \tilde{\mu}_k(A_x)$ meßbar ist für jedes $A \in \mathcal{A}$. Nach Übung 9 (Blatt 3) ist aber

$$\mu_2(A_x) = \sup_{k\in\mathbb{N}} \tilde{\mu}_k(A_x),$$

was nach Satz 3.12 meßbar ist. □

Nun können wir endlich das Produktmaß definieren:

Definition 4.6 (Produktmaß). Seien $(\Omega_1, \mathcal{A}_1, \mu_1)$ und $(\Omega_2, \mathcal{A}_2, \mu_2)$ zwei σ-endliche Maßräume. Dann ist auf $\mathcal{A}_1 \otimes \mathcal{A}_2$ das *Produktmaß* $\mu_1 \otimes \mu_2$ definiert durch

$$(\mu_1 \otimes \mu_2)(A) := \int_{\Omega_1} \mu_2(A_x) d\mu_1(x).$$

Nach vorigem Lemma ist der Integrand meßbar bzgl. $\mathcal{A}_1$ und nichtnegativ, also ist das Produktmaß wohldefiniert. Es handelt sich tatsächlich um ein Maß: Klar ist $(\mu_1 \otimes \mu_2)(\varnothing) = 0$. Ist $(A_k)_{k\in\mathbb{N}}$ eine Familie paarweise disjunkter Mengen in $\mathcal{A}_1 \otimes \mathcal{A}_2$, so sind für jedes $x \in \Omega_1$ auch die $(A_k)_x$ paarweise disjunkt, und somit erhalten wir mit der σ-Additivität von μ_2 und dem Satz von der monotonen Konvergenz

$$\begin{aligned}(\mu_1 \otimes \mu_2)\left(\bigcup_{k\in\mathbb{N}} A_k\right) &= \int_{\Omega_1} \sum_{k\in\mathbb{N}} \mu_2((A_k)_x) d\mu_1(x) \\ &= \sum_{k\in\mathbb{N}} \int_{\Omega_1} \mu_2((A_k)_x) d\mu_1(x) = \sum_{k\in\mathbb{N}} (\mu_1 \otimes \mu_2)(A_k).\end{aligned}$$

Satz 4.7 (Eigenschaften des Produktmaßes). Für $A_1 \in \mathcal{A}_1$ und $A_2 \in \mathcal{A}_2$ gilt

$$(\mu_1 \otimes \mu_2)(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2). \tag{4.1}$$

Aus der σ-Endlichkeit von μ_1 und μ_2 folgt die von $\mu_1 \otimes \mu_2$.

Durch Eigenschaft (4.1) ist $\mu_1 \otimes \mu_2$ als Maß auf $\mathcal{A}_1 \otimes \mathcal{A}_2$ bereits eindeutig bestimmt, und es gilt für jedes $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$

$$(\mu_1 \otimes \mu_2)(A) = \int_{\Omega_2} \mu_1(A_y) d\mu_2(y). \tag{4.2}$$

Beweis. Die Eigenschaft (4.1) folgt sofort aus $(A_1 \times A_2)_x = A_2$, falls $x \in A_1$, und $(A_1 \times A_2)_x = \varnothing$ falls $x \notin A_1$.

Die σ-Endlichkeit des Produktmaßes folgt damit leicht: Sind nämlich $(\Omega_{1,k})_{k\in\mathbb{N}}$ und $(\Omega_{2,k})_{k\in\mathbb{N}}$ Folgen meßbarer Mengen mit $\bigcup_{k\in\mathbb{N}} \Omega_{1,k} = \Omega_1$, $\bigcup_{k\in\mathbb{N}} \Omega_{2,k} = \Omega_2$, sowie $\mu_1(\Omega_{1,k}), \mu_2(\Omega_{2,k}) < \infty$ für alle $k \in \mathbb{N}$, so ist $(\Omega_{1,k} \times \Omega_{2,k})_{k\in\mathbb{N}}$ eine Familie meßbarer Mengen, deren Vereinigung $\Omega_1 \times \Omega_2$ ergibt und die jeweils endliches $\mu_1 \otimes \mu_2$-Maß haben.

Die Eindeutigkeit eines Maßes mit der Eigenschaft (4.1) folgt aus Satz 2.6, wenn man bemerkt, daß (4.1) das Produktmaß bereits auf dem Ring[1] der endlichen disjunkten Vereinigungen von Mengen der Form $A_1 \times A_2$ ($A_1 \in \mathcal{A}_1$, $A_2 \in \mathcal{A}_2$) eindeutig festlegt.

Ähnlich wie zu Beginn dieses Beweises sieht man, daß auch das Integral in (4.2) die Eigenschaft (4.1) besitzt. Da diese Eigenschaft aber ein Maß eindeutig bestimmt, folgt bereits (4.2).

□

4.2. Die Sätze von Tonelli und Fubini

Wir untersuchen schließlich, wie eine Funktion bezüglich eines Produktmaßes integriert werden kann. Es stellt sich dabei in den Sätzen von Tonelli und Fubini heraus, daß Funktionen mehrerer Variabler sukzessive und in beliebiger Reihenfolge integriert werden dürfen.

Wieder seien $(\Omega_1, \mathcal{A}_1, \mu_1)$ und $(\Omega_2, \mathcal{A}_2, \mu_2)$ Maßräume.

Satz 4.8 (Tonelli). Seien μ_1 und μ_2 σ-endlich und $f : \Omega_1 \times \Omega_2 \to \bar{\mathbb{R}}$ nichtnegativ und meßbar bzgl. $\mathcal{A}_1 \otimes \mathcal{A}_2$. Dann sind die Abbildungen

$$x \mapsto \int_{\Omega_2} f(x,y) d\mu_2(y) \quad \text{und} \quad y \mapsto \int_{\Omega_1} f(x,y) d\mu_1(x)$$

[1] Daß es sich dabei tatsächlich um einen Ring handelt, folgt wie in Satz 2.2. Mit diesem Argument könnte man die Existenz eines eindeutig bestimmten Maßes mit der Eigenschaft (4.1) übrigens auch aus dem Fortsetzungssatz von Carathéodory folgern. Wir ziehen es allerdings vor, das Produktmaß durch eine explizite Formel zu definieren.

meßbar bzgl. $\mathcal{A}_1$ bzw. $\mathcal{A}_2$, und es gilt

$$\begin{aligned}\int_{\Omega_1\times\Omega_2} f d(\mu_1\otimes\mu_2) &= \int_{\Omega_2}\left(\int_{\Omega_1} f(x,y)d\mu_1(x)\right)d\mu_2(y)\\ &= \int_{\Omega_1}\left(\int_{\Omega_2} f(x,y)d\mu_2(y)\right)d\mu_1(x).\end{aligned}$$

Beweis. Schreibe $\Omega := \Omega_1\times\Omega_2$, $\mathcal{A} := \mathcal{A}_1\otimes\mathcal{A}_2$ und $\mu := \mu_1\otimes\mu_2$. Betrachte zunächst eine Elementarfunktion $f = \sum_{j=1}^k \alpha_j\chi_{A_j}$ mit $\alpha_j \geq 0$ und $A_j \in \mathcal{A}$. Dann ist offenbar $f(\cdot,y) = \sum_{j=1}^k \alpha_j\chi_{(A_j)_y}$ und daher

$$\int_{\Omega_1} f(x,y)d\mu_1(x) = \sum_{j=1}^k \alpha_j\mu_1((A_j)_y).$$

Dieser Ausdruck (als Funktion von y betrachtet) ist gemäß Lemma 4.5 meßbar bzgl. $\mathcal{A}_2$, und es gilt

$$\begin{aligned}\int_{\Omega_2}\left(\int_{\Omega_1} f(x,y)d\mu_1(x)\right)d\mu_2(y) &= \sum_{j=1}^k \alpha_j\int_{\Omega_2}\mu_1((A_j)_y)d\mu_2(y)\\ &= \sum_{j=1}^k \alpha_j\mu(A_j) = \int_\Omega f d\mu\end{aligned}$$

nach (4.2).

Sei nunmehr f meßbar bzgl. $\mathcal{A}$ und nichtnegativ und f_k eine monotone Folge von Elementarfunktionen, die gegen f konvergiert. Dann ist, für jedes $y \in \Omega_2$, $f_k(\cdot,y)$ eine monotone Folge von Elementarfunktionen, die gegen $f(\cdot,y)$ konvergiert[2], und daher haben wir

$$\int_{\Omega_1} f(x,y)d\mu_1(x) = \sup_{k\in\mathbb{N}}\int_{\Omega_1} f_k(x,y)d\mu_1(x).$$

Nach Lemma 4.5 sind die Integrale unter dem Supremum jeweils meßbare Funktionen von y (bzgl. $\mathcal{A}_2$), also ist die linke Seite nach Satz 3.12 ebenfalls meßbar bzgl. $\mathcal{A}_2$. Der Satz von der monotonen Konvergenz liefert nun

$$\begin{aligned}\int_{\Omega_2}\left(\int_{\Omega_1} f(x,y)d\mu_1(x)\right)d\mu_2(y) &= \sup_{k\in\mathbb{N}}\int_{\Omega_2}\left(\int_{\Omega_1} f_k(x,y)d\mu_1(x)\right)d\mu_2(y)\\ &= \sup_{k\in\mathbb{N}}\int_\Omega f_k d\mu,\end{aligned}$$

wobei wir im letzten Schritt den ersten Beweisteil verwendet haben. Dies ist aber gleich $\int_\Omega f d\mu$ nach Definition des Integrals.

Die übrigen Aussagen erhält man analog, indem man $f(x,\cdot)$ betrachtet. □

Satz 4.9 (Fubini). Seien μ_1 und μ_2 σ-endlich und $f : \Omega_1\times\Omega_2 \to \bar{\mathbb{R}}$ integrierbar bzgl. $\mu_1\otimes\mu_2$. Dann sind für μ_1-fast alle $x\in\Omega_1$ und für μ_2-fast alle $y\in\Omega_2$ die Abbildungen

$$\Omega_2 \to \bar{\mathbb{R}},\quad y\mapsto f(x,y)\quad \text{und}\quad \Omega_1\to\bar{\mathbb{R}},\quad x\mapsto f(x,y)$$

integrierbar bzgl. μ_2 bzw. μ_1. Ferner sind die (fast überall definierten) Abbildungen

$$x\mapsto \int_{\Omega_2} f(x,y)d\mu_2(y)\quad\text{und}\quad y\mapsto\int_{\Omega_1} f(x,y)d\mu_1(x)$$

[2] Dies zeigt zugleich die Meßbarkeit von $f(\cdot,y)$ bzgl. $\mathcal{A}_1$ und analog von $f(x,\cdot)$ bzgl. $\mathcal{A}_2$.

integrierbar bzgl. μ_1 bzw. μ_2, und es gilt

$$\begin{aligned}\int_{\Omega_1\times\Omega_2} f d(\mu_1\otimes\mu_2) &= \int_{\Omega_2}\left(\int_{\Omega_1} f(x,y)d\mu_1(x)\right)d\mu_2(y)\\ &= \int_{\Omega_1}\left(\int_{\Omega_2} f(x,y)d\mu_2(y)\right)d\mu_1(x).\end{aligned}$$

BEWEIS. Sei f integrierbar bzgl. μ. Nach dem Satz von TONELLI gilt

$$\int_{\Omega_2}\left(\int_{\Omega_1}|f(x,y)|d\mu_1(x)\right)d\mu_2(y) = \int_{\Omega}|f|d\mu < \infty,$$

also ist $y \mapsto \int_{\Omega_1}|f(x,y)|d\mu_1(x)$ integrierbar bzgl. μ_2 und damit nach Proposition 3.36 μ_2-fast überall endlich. Das bedeutet insbesondere, daß für μ_2-fast alle $y \in \Omega_2$ die Funktion $f(\cdot,y)$ integrierbar bzgl. μ_1 ist. Wegen

$$\int_{\Omega_2}\left|\int_{\Omega_1} f(x,y)d\mu_1(x)\right|d\mu_2(y) \le \int_{\Omega_2}\left(\int_{\Omega_1}|f(x,y)|d\mu_1(x)\right)d\mu_2(y) < \infty$$

ist aber auch die Funktion $y \mapsto \int_{\Omega_1} f(x,y)d\mu_1(x)$ integrierbar bzgl. μ_2. Schließlich haben wir

$$\int_{\Omega_2}\left(\int_{\Omega_1} f(x,y)^{\pm}d\mu_1(x)\right)d\mu_2(y) = \int_{\Omega} f^{\pm}d\mu < \infty,$$

woraus sich mit $f = f^+ - f^-$ die Behauptung ergibt.

Für $f(x,\cdot)$ argumentiert man wieder analog, um die übrigen Aussagen zu erhalten.

$\square$

Literaturverzeichnis

[1] H. Bauer. Maß- und Integrationstheorie. *De Gruyter, Berlin/New York,* 2. Aufl. 1992.
[2] E. Wiedemann. Analysis. Vorlesungsskript Uni Ulm 2018/19.