Vorlesungsskript ,,Integrationstheorie”

Christoph Kühn

aktuelle Version: 15. Februar 2013

Das Skript wird parallel zur Vorlesung im Wintersemester 2012/13 aufgebaut, Hinweise auf Fehler sind sehr willkommen. Bitte beachten: Sofern der Stoff noch nicht in der Vorlesung dran war, ist noch mit Änderungen und Ergänzungen zu rechnen.

Inhaltsverzeichnis

Die Vorlesung baut auf der Analysis II von Herrn Weth auf, in der Grundlagen der Maßtheorie bereits behandelt wurden (siehe §17 bis §19 dort). Gegeben ein Maß, das auf einer σ-Algebra definiert ist, wollen wir nun Integrale bezüglich dieses Maßes konstruieren. Nimmt man das Maß als gegeben hin, ist die Vorlesung selbsterklärend. Für den nichttrivialen Existenzbeweis für interessante Maße wie dem Lebesgue-Maß wird jedoch auf die Konstruktion durch äußere Maße in der Analysis II verwiesen. Das schnelle Vorankommen in der Integrationstheorie beruht entscheidend auf den maßtheoretischen Vorarbeiten in der Analysis II.

In Kapitel 1 werden einige Begrifflichkeiten und einige oft gebrauchte (und ohne viel Aufwand beweisbare) Resultate aus der Maßtheorie nochmal kurz gebracht.

1 Maßtheorie

Sei Ω eine beliebige Menge und 2^Ω deren Potenzmenge.

Definition 1.1 (σ-Algebra und Maß). *(a) Eine Menge $\mathcal{F}$ von Teilmengen von Ω (also $\mathcal{F} \subset 2^\Omega$) heißt σ-Algebra, wenn sie folgende Eigenschaften besitzt*

(i) $\Omega \in \mathcal{F}$

(ii) $A \in \mathcal{F} \implies A^c := \Omega \setminus A \in \mathcal{F}$

(iii) $A_1, A_2, \ldots \in \mathcal{F} \implies \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{F}$

(b) Eine Abbildung $\mu : \mathcal{F} \to \overline{\mathbb{R}}_+$ heißt Maß, wenn sie folgende Eigenschaften besitzt

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

(ii) Für disjunkte Mengen $A_1, A_2, \ldots \in \mathcal{F}$ gilt $\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n=1}^\infty \mu(A_n)$ (σ-Additivität).

Ein Maß μ heißt endlich, falls $\mu(\Omega) < \infty$ und σ-endlich falls es eine Folge $(A_n)_{n\in\mathbb{N}} \subset \mathcal{F}$ gibt mit $\cup_{n\in\mathbb{N}} A_n = \Omega$ und $\mu(A_n) < \infty$ für alle $n \in \mathbb{N}$.

(c) Das Tripel $(\Omega, \mathcal{F}, \mu)$ nennt man einen Maßraum.

Beispiel 1.2. *Ein Maß mit Masse 1 nennt man Wahrscheinlichkeitsmaß und bezeichnet es zumeist mit P. Elemente $\omega \in \Omega$ nennt man die Ergebnisse des Zufallsexperimentes und Elemente $A \in \mathcal{F}$ die Ereignisse. Interpretation: $\mathcal{F} = \{A \subset \Omega \mid A \text{ ist beobachtbares Ereignis}\}$. Unter der Zusatzbedingung $\mu(\Omega) = 1 < \infty$ folgt Bedingung (b.i) aus (b.ii).*

In der Vorlesung arbeiten wir mit den üblichen Rechenregeln

$$\infty + a = \infty \quad \text{für alle } a \in \mathbb{R} \cup \{\infty\}, \ \ldots$$

Am gewöhnungsbedürftigsten wird vielleicht die Regel

$$0 \cdot \infty = 0$$

sein, die jedoch erst ab Kapitel 2 beim Integral gebraucht wird.

$\infty - \infty$ ist dagegen undefiniert, es muss also darauf geachtet werden, dass dieser Ausdruck nicht auftreten kann.

Definition 1.3. *Seien $A_n, A \subset \Omega$, $n \in \mathbb{N}$. Wir schreiben $A_n \uparrow A$ für $n \uparrow \infty$ falls gilt*

$$A = \bigcup_{n \in \mathbb{N}} A_n \quad \textit{und} \quad A_1 \subset A_2 \subset A_3 \subset \ldots$$

Wir schreiben $A_n \downarrow A$ für $n \uparrow \infty$ falls gilt

$$A = \bigcap_{n \in \mathbb{N}} A_n \quad \textit{und} \quad A_1 \supset A_2 \supset A_3 \supset \ldots$$

Man sagt dann auch „die Folge $(A_n)_{n \in \mathbb{N}}$ steigt gegen A auf bzw. ab”.

Man beachte, dass wegen

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

(de Morgan'sche Regeln ...) mit $(A_n)_{n \in \mathbb{N}}$ auch der abzählbare Schnitt $\bigcap_{n \in \mathbb{N}} A_n$ in der σ-Algebra $\mathcal{F}$ liegt. Zudem ist $A_1 \setminus A_2 = A_1 \cap A_2^c \in \mathcal{F}$.

Proposition 1.4. *Für ein Maß μ und Mengen aus $\mathcal{F}$ gilt folgendes*

(i) $A \subset B \implies \mu(A) \leq \mu(B)$ *(Monotonie)*

(ii) $\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) \le \sum_{n=1}^{\infty} \mu(A_n)$ *σ-Subadditivität*

(iii) $A_n \uparrow A$ *für* $n \uparrow \infty \implies \mu(A_n) \uparrow \mu(A)$ *für* $n \uparrow \infty$ *(Stetigkeit von unten)*

(iv) $\mu(A_1) < \infty$ *und* $A_n \downarrow A$ *für* $n \uparrow \infty \implies \mu(A_n) \downarrow \mu(A)$ *für* $n \uparrow \infty$ *(Stetigkeit von oben)*

Beweis. Ad (i): Es gilt $B = A \cup (B \setminus A)$ und A, $B \setminus A$ sind disjunkt also

$$\mu(B) = \mu(A) + \mu(B \setminus A) \ge \mu(A)$$

(Wähle etwa in Definition 1.1(b.ii) $A_1 = A$, $A_2 = B \setminus A$ und $A_n = \emptyset$ für $n \ge 2$)

Ad (ii): Definiere rekursiv $\widetilde{A}_1 := A_1$ und $\widetilde{A}_{n+1} := A_{n+1} \setminus (\cup_{k=1}^{n} \widetilde{A}_k)$ für $n = 1, 2, \ldots$. Die Mengen $(\widetilde{A}_n)_{n\in\mathbb{N}}$ sind paarweise disjunkt mit $\widetilde{A}_n \subset A_n$ für alle $n \in \mathbb{N}$ sowie $\bigcup_{n\in\mathbb{N}} \widetilde{A}_n = \bigcup_{n\in\mathbb{N}} A_n$ (für letzteres zeige man per vollständiger Induktion, dass $\bigcup_{k=1,\ldots,n} \widetilde{A}_k = \bigcup_{k=1,\ldots,n} A_k$ für alle $n \in \mathbb{N}$). Es folgt

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \mu\left(\bigcup_{n\in\mathbb{N}} \widetilde{A}_n\right) \overset{\text{Definition 1.1(b.ii)}}{=} \sum_{n=1}^{\infty} \mu(\widetilde{A}_n) \overset{\text{Monotonie}}{\le} \sum_{n=1}^{\infty} \mu(A_n).$$

Ad (iii): A lässt sich als abzählbare disjunkte Vereinigung der Mengen A_1 und $A_{n+1} \setminus A_n$, $n \in \mathbb{N}$ schreiben. Es folgt

$$\begin{aligned}
\mu(A) &= \mu\left(A_1 \cup \bigcup_{n\in\mathbb{N}} (A_{n+1} \setminus A_n)\right) \\
&= \lim_{m\to\infty} \left(\mu(A_1) + \sum_{n=1}^{m} \mu(A_{n+1} \setminus A_n)\right) \\
&= \lim_{m\to\infty} \mu(A_1 \cup (A_2 \setminus A_1) \cup \ldots \cup (A_{m+1} \setminus A_m)) \\
&= \lim_{m\to\infty} \mu(A_m).
\end{aligned}$$

(Man beachte, dass nirgendwo der Ausdruck $\mu(A_{n+1}) - \mu(A_n)$ auftaucht, der wegen $\infty - \infty$ undefiniert sein könnte)

Ad (iv): Offenbar gilt $(A_1 \setminus A_n) \uparrow (A_1 \setminus A)$ für $n \uparrow \infty$. Es gilt nämlich $(A_1 \setminus A_n) \subset (A_1 \setminus A_{n+1})$ für alle $n \in \mathbb{N}$ und

$$\bigcup_{n\in\mathbb{N}} (A_1 \setminus A_n) = A_1 \cap \left(\bigcup_{n\in\mathbb{N}} A_n^c\right) = A_1 \cap \left(\bigcap_{n\in\mathbb{N}} A_n\right)^c = A_1 \cap A^c = A_1 \setminus A.$$

Da wegen $\mu(A_1) < \infty$ die Mengen nun alle endliches Maß haben gilt

$$\mu(A_n) = \mu(A_1) - \mu(A_1 \setminus A_n) \quad \text{und} \quad \mu(A) = \mu(A_1) - \mu(A_1 \setminus A)$$

und mit (iii) folgt die Behauptung.

□

Man beachte, dass ohne die Voraussetzung $\mu(A_1) < \infty$ in (iv) die Implikation nicht gelten muss. Betrachte dazu das (für $\Omega \neq \emptyset$ nicht σ-endliche) Maß

$$\mu(A) := \begin{cases} 0 & : \quad \text{für } A = \emptyset \\ \infty & : \quad A \in 2^\Omega \setminus \{\emptyset\} \,. \end{cases}$$

Satz 1.5 (Erzeugung von σ-Algebren)**.** *Sei $\mathcal{E} \subset 2^\Omega$. Es existiert eine kleinste σ-Algebra $\sigma(\mathcal{E})$, die $\mathcal{E}$ umfasst, nämlich*

$$\sigma(\mathcal{E}) := \bigcap_{\mathcal{A} \subset 2^\Omega,\ \mathcal{A} \text{ ist } \sigma\text{-Algebra mit } \mathcal{E} \subset \mathcal{A}} \mathcal{A}. \tag{1.1}$$

"Kleinste" bedeutet, dass jede andere σ-Algebra, die $\mathcal{E}$ umfasst, auch $\sigma(\mathcal{E})$ umfasst.

(Man beachte, dass $A \subset \Omega$ definitionsgemäß Element von $\bigcap \mathcal{A}$ ist, wenn A Element jeder σ-Algebra ist, über die der Schnitt gebildet wird. Es werden also keine Teilmengen von Ω miteinander geschnitten)

Beweis. Die Menge, über die der Schnitt in (4.1) gebildet wird, ist nichtleer, denn die Potenzmenge 2^Ω ist eine σ-Algebra, die $\mathcal{E}$ umfasst. Die Minimalität folgt sofort aus der Konstruktion, da jede σ-Algebra, die $\mathcal{E}$ umfasst, beim Schnitt berücksichtigt wird und somit Obermenge von $\sigma(\mathcal{E})$ ist. Bleibt zu zeigen, dass $\sigma(\mathcal{E})$ tatsächlich eine σ-Algebra ist. Wir prüfen die Eigenschaften (i),(ii) und (iii) von oben.

(i) Für jede σ-Algebra $\mathcal{A}$ gilt $\Omega \in \mathcal{A}$ und damit $\Omega \in \sigma(\mathcal{E})$.

(ii) Sei $A \in \sigma(\mathcal{E})$. Damit ist $A \in \mathcal{A}$ für alle $\mathcal{A}$, über die der Schnitt in (4.1) gebildet wird. Damit gilt aber auch $A^c \in \mathcal{A}$ für die entsprechenden σ-Algebren und folglich $A^c \in \sigma(\mathcal{E})$.

(iii) Seien $A_1, A_2, \ldots \in \sigma(\mathcal{E})$. Es folgt $A_1, A_2, \ldots \in \mathcal{A}$ für alle $\mathcal{A}$, über die der Schnitt in (4.1) gebildet wird. Damit gilt aber auch $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$ für die entsprechenden σ-Algebren und folglich $\bigcup_{n\in\mathbb{N}} A_n \in \sigma(\mathcal{E})$.

□

Proposition 1.6. *Seien $\mathcal{E}_1, \mathcal{E}_2 \subset 2^\Omega$. Es gilt die Implikation*

$$\mathcal{E}_1 \subset \sigma(\mathcal{E}_2) \implies \sigma(\mathcal{E}_1) \subset \sigma(\mathcal{E}_2).$$

Beweis. Sei $\mathcal{E}_1 \subset \sigma(\mathcal{E}_2)$. $\sigma(\mathcal{E}_2)$ ist also eine σ-Algebra, die $\mathcal{E}_1$ umfasst und damit ein Mengensystem, das im Schnitt (4.1) mit $\mathcal{E} = \mathcal{E}_1$ vorkommt. Es folgt $\sigma(\mathcal{E}_1) \subset \sigma(\mathcal{E}_2)$. □

Beispiel 1.7. $\mathcal{E}_1 := \{\{a\} \mid a \in \mathbb{R}\}$ *und* $\mathcal{E}_2 := \{(a,b) \mid a,b \in \mathbb{R},\ a < b\}$. *Es gilt zwar* $\mathcal{E}_1 \not\subset \mathcal{E}_2$, *aber wegen* $\{a\} = \bigcap_{n\in\mathbb{N}}(a - 1/n, a + 1/n) \in \sigma(\mathcal{E}_2)$ *folgt mit Proposition 1.6* $\sigma(\mathcal{E}_1) \subset \sigma(\mathcal{E}_2)$. *Andererseits gilt* $\mathcal{E}_2 \not\subset \sigma(\mathcal{E}_1)$ *(wieso ?)*

Definition 1.8 (Borelsche σ-Algebra auf dem $\mathbb{R}^n$)**.** *Seien $n \in \mathbb{N}$ und $\Omega = \mathbb{R}^n$. Die Borelsche σ-Algebra $\mathcal{B}^n := \mathcal{B}(\mathbb{R}^n)$ ist die von der Menge $\mathcal{E}_1$ der offenen Teilmengen von $\mathbb{R}^n$ erzeugte σ-Algebra, wobei Offenheit einer Menge bzgl. der euklidischen Norm $||x|| := \sqrt{\sum_{k=1}^n x_k^2}$, $x \in \mathbb{R}^n$ zu verstehen ist.*

Bemerkung 1.9. *Die Borelsche σ-Algebra ist für beliebige topologische Räume definiert (also Räume auf denen ein System von offenen Mengen gegeben ist). Sie ist die kleinste σ-Algebra, die die offenen Mengen enthält. Anwendungen findet dies etwa bei Funktionenräumen. Im folgenden betrachten wir die Borelsche σ-Algebra aber nur für $\mathbb{R}^n$, $\overline{\mathbb{R}}$.*

Für $a, b \in \mathbb{R}^n$ mit $a < b$ (d.h. $a_i < b_i\ \forall i = 1, \ldots, n$) definieren wir den offenen **Quader** als das kartesische Produkt

$$(a,b) := (a_1,b_1) \times (a_2,b_2) \times \ldots \times (a_n,b_n)$$

Definiere $\mathcal{E}_2 := \{(a,b) \mid a,b \in \mathbb{R}^n,\ a < b\}$.

Satz 1.10. *Es gilt $\sigma(\mathcal{E}_1) = \sigma(\mathcal{E}_2)$, d.h. die Borelsche σ-Algebra wird auch von den offenen Quadern erzeugt.*

Beweis. 1. Aus $\mathcal{E}_2 \subset \mathcal{E}_1$ folgt $\sigma(\mathcal{E}_2) \subset \sigma(\mathcal{E}_1)$.

2. Sei $\mathcal{O}$ eine offene Teilmenge des $\mathbb{R}^n$. Offenbar gilt

$$\mathcal{O} = \bigcup_{a<b,\ a,b\in\mathbb{Q}^n, (a,b)\subset\mathcal{O}} (a,b) =: \widetilde{\mathcal{O}}.$$

$\widetilde{\mathcal{O}} \subset \mathcal{O}$ ergibt sich direkt aus der Konstruktion von $\widetilde{\mathcal{O}}$. Sei umgekehrt $x \in \mathcal{O}$. Da $\mathcal{O}$ offen ist, existiert ein $\varepsilon > 0$ mit $(x_1-\varepsilon, x_1+\varepsilon) \times \ldots \times (x_n-\varepsilon, x_n+\varepsilon) \subset \mathcal{O}$ und da die rationalen Zahlen dicht liegen, gibt es auch $a_i, b_i \in \mathbb{Q}$ mit $x_i - \varepsilon < a_i < x_i < b_i < x_1 + \varepsilon$. Es folgt $x \in \widetilde{\mathcal{O}}$.

$\widetilde{\mathcal{O}}$ ist als abzählbare Vereinigungen von offenen Quadern Element aus $\sigma(\mathcal{E}_2)$ (jede σ-Algebra, die $\mathcal{E}_2$ umfasst, enthält $\widetilde{\mathcal{O}}$ als Element). Mit Proposition 1.6 folgt $\sigma(\mathcal{E}_1) \subset \sigma(\mathcal{E}_2)$ und damit insgesamt Gleichheit. □

Bemerkung 1.11. *Die Borelsche σ-Algebra wird auch von den abgeschlossenen Quadern (Intervallen) $[a,b]$ oder von den Quadern $(a,b]$ oder von den Quadern $[a,b)$ erzeugt. Mit Proposition 1.6 sieht man dies, indem man etwa $[a,b]$ durch $\cap_{n\in\mathbb{N}}(a-1/n, b+1/n)$ darstellt und umgekehrt (a,b) durch $\cup_{n\in\mathbb{N}}(a+1/n, b-1/n)$.*

Definition 1.12. *Die Borelsche σ-Algebra auf $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ wird durch*

$$\overline{\mathcal{B}} := \{B \subset \overline{\mathbb{R}} \mid (B \cap \mathbb{R}) \in \mathcal{B}(\mathbb{R})\}$$

definiert (Übung: man zeige, dass $\overline{\mathcal{B}}$ tatsächlich eine σ-Algebra ist und von den Intervallen $[-\infty, b]$, $b \in \mathbb{N}$, erzeugt wird)

Satz 1.13 (Spur-σ-Algebra)**.** *Sei $\mathcal{F}$ eine σ-Algebra auf Ω und $A \subset \Omega$ (A muss nicht Element aus $\mathcal{F}$ sein). Dann ist*

$$A \cap \mathcal{F} := \{B \subset \Omega \mid \exists C \in \mathcal{F} \textit{ und } B = A \cap C\} \tag{1.2}$$

eine σ-Algebra auf der Menge A (d.h. die Bedingung (a.i) in Definition 1.1 wird durch $A \in A \cap \mathcal{F}$ ersetzt und (a.ii) durch die Implikation $B \in A \cap \mathcal{F} \Rightarrow (A \setminus B) \in A \cap \mathcal{F}$). $A \cap \mathcal{F}$ wird **Spur-σ-Algebra** *von $\mathcal{F}$ auf A genannt.*

Offenbar ist $\mathcal{B}(\mathbb{R})$ die Spur-σ-Algebra von $\overline{\mathcal{B}}$ auf $\mathbb{R}$ (was natürlich nicht als Definition von $\overline{\mathcal{B}}$ taugt, da verschiedene σ-Algebren auf einer Teilmenge die gleichen Spur-σ-Algebren induzieren können). Umgekehrt kann man aber z.B. $\mathcal{B}(\overline{\mathbb{R}}_+)$ als Spur-σ-Algebra von $\overline{\mathcal{B}}$ auf $\overline{\mathbb{R}}_+ \subset \overline{\mathbb{R}}$ definieren.

Definition 1.14. *Sei Ω' ein weiterer Grundraum und $\mathcal{F}' \subset 2^{\Omega'}$ eine weitere σ-Algebra. Eine Abbildung $f : \Omega \to \Omega'$ heißt $\mathcal{F} - \mathcal{F}'$-messbar, falls die Urbilder messbarer Mengen wieder messbar sind, falls also*

$$f^{-1}(A') := \{\omega \in \Omega \mid f(\omega) \in A'\} \in \mathcal{F}, \quad \forall A' \in \mathcal{F}'.$$

Eine reellwertige Abbildung, die $\mathcal{F} - \mathcal{B}(\mathbb{R})$-messbar ist, nennt man einfach $\mathcal{F}$-messbar, d.h. wenn keine andere σ-Algrbra auf dem Bildraum angegeben ist, wird stillschweigend von der Borelschen σ-Algebra ausgegangen.

Für $f^{-1}(A')$ schreibt man meistens $\{f \in A'\}$, also $\{f > 0\}$ für $f^{-1}((0,\infty))$ oder im Falle von $f = (f_1, f_2)$ auch $\{f_2 \geq f_1\}$ für $\{(f_1, f_2) \in \{(x, y) \in \mathbb{R}^2 \mid y \geq x\}\}$.

Proposition 1.15. *Sei $\mathcal{E}'$ ein Erzeuger von $\mathcal{F}'$, d.h. $\sigma(\mathcal{E}') = \mathcal{F}'$. Wenn $f^{-1}(E') \in \mathcal{F}$ für alle $E' \in \mathcal{E}'$, dann ist f $\mathcal{F} - \mathcal{F}'$-messbar.*

Beweis. Man rechne nach, dass $\mathcal{A}' := \{B' \subset \Omega' \mid f^{-1}(B') \in \mathcal{F}\}$ wiederum eine σ-Algebra ist. Es gilt nämlich $f^{-1}(\Omega') = \Omega \in \mathcal{F}$, $f^{-1}(\Omega' \setminus B') = \Omega \setminus f^{-1}(B')$, $\forall B' \subset \Omega'$ und $f^{-1}(\cup_{n\in\mathbb{N}} B'_n) = \cup_{n\in\mathbb{N}} f^{-1}(B'_n)$, $\forall (B'_n)_{n\in\mathbb{N}} \subset \Omega'$.

Nach Voraussetzung gilt $\mathcal{E}' \subset \mathcal{A}'$. Da $\mathcal{A}'$ eine σ-Algebra ist, gilt $\sigma(\mathcal{A}') = \mathcal{A}'$ und es folgt mit Proposition 1.6, dass $\sigma(\mathcal{E}') \subset \mathcal{A}'$. □

Proposition 1.16. *Jede stetige Funktion $f : \mathbb{R}^n \to \mathbb{R}$ ist $\mathcal{B}^n - \mathcal{B}$-messbar.*

Beweis. Mit Proposition 1.15 klar, da unter einer stetigen Funktion das Urbild einer offenen Menge offen ist. □

Beispiel 1.17 (Zufallsvariable)**.** *In der Wahrscheinlichkeitstheorie ist eine reellwertige Zufallsvariable eine Abbildung $Y : \Omega \to \mathbb{R}$, die $\mathcal{F}$-$\mathcal{B}(\mathbb{R})$-messbar ist, wobei $\mathcal{B}(\mathbb{R})$ die*

Borelsche-σ-Algebra auf $\mathbb{R}$ bezeichnet. Eine Abbildung $Y : \Omega \to \mathbb{R} \cup \{-\infty, \infty\} =: \overline{\mathbb{R}}$, die $\mathcal{F}$-$\overline{\mathcal{B}}$-messbar ist, wird **numerische Zufallsvariable** *genannt.*

Im Fall, dass Ω überabzählbar ist, können i.A. nicht mehr allen Teilmengen von Ω eine Wahrscheinlichkeit zuordnen (wenn zudem das Wahrscheinlichkeitsmaß bestimmte wünschenswerte Eigenschaften haben soll, siehe etwas das Lebesgue-Maß in der Analysis II, das die Gleichverteilung auf $(0,1)$ modelliert und das nicht zu einem Maß auf der gesamten Potenzmenge fortgesetzt werden kann). Für die Erfordernisse der stochastischen Modellierung ist die Menge $\{Y^{-1}(B) \mid B \in \mathcal{B}\}$ allerdings groß genug. Da sie die Menge $\{Y \leq y \mid y \in \mathbb{R}\}$ umfasst, scheint sie alle Informationen über Y zu liefern.

Satz 1.18 (Komposition messbarer Funktionen)**.** *Sei Ω'' ein weiterer Grundraum und $\mathcal{F}'' \subset 2^{\Omega''}$ eine weitere σ-Algebra. Wenn die Abbildung $f : \Omega \to \Omega'$ $\mathcal{F} - \mathcal{F}'$-messbar und die Abbildung $g : \Omega' \to \Omega''$ $\mathcal{F}' - \mathcal{F}''$-messbar ist, dann ist die Komposition $g \circ f : \Omega \to \Omega''$ (d.h. $(g \circ f)(\omega) = g(f(\omega))$) $\mathcal{F} - \mathcal{F}''$-messbar.*

Beweis. Für jede Menge $B'' \in \mathcal{F}''$ gilt

$$\begin{aligned} (g \circ f)^{-1}(B'') &= \{\omega \in \Omega \mid g(f(\omega)) \in B''\} \\ &= \{\omega \in \Omega \mid f(\omega) \in g^{-1}(B'')\} \\ &= f^{-1}(g^{-1}(B'')). \end{aligned} \tag{1.3}$$

Wegen der Messbarkeit von g gilt für alle $B'' \in \mathcal{F}''$, dass $B' := g^{-1}(B'') \in \mathcal{F}'$. Weiter gilt $f^{-1}(B') \in \mathcal{F}$ wegen der Messbarkeit von f. (1.3) ergibt dann die Behauptung. $\square$

Sei $I \neq \emptyset$ eine beliebige Indexmenge und $(f_i)_{i \in I}$ eine Familie reellwertiger Funktionen. Mit

$$\sigma(f_i,\ i \in I) := \sigma(\{f_i^{-1}(B) \mid B \in \mathcal{B}(\mathbb{R}),\ i \in I\})$$

wird die von den f_i, $i \in I$, erzeugte σ-Algebra bezeichnet. $\sigma(f_i,\ i \in I)$ ist also die kleinste σ-Algebra auf Ω, so dass alle Abbildungen $f_i,\ i \in I$, messbar sind.

Offenbar ist

$$\{f_i^{-1}((a,b)) \mid a,b \in \mathbb{R},\ a < b,\ i \in I\}$$

ein Erzeuger von $\sigma(f_i,\ i \in I)$. Dafür reicht es aus zu zeigen, dass für festes $i \in I$ gilt

$$\sigma(f_i^{-1}(\mathcal{E})) = f_i^{-1}(\sigma(\mathcal{E})).$$

(Übung)

Proposition 1.19. *Seien $f_1, f_2, \dots$ $\mathcal{F} - \overline{\mathcal{B}}$-messbare Funktionen, dann sind*

$$f_1 + f_2, \quad f_1 - f_2, \quad f_1 f_2, \quad \min\{f_1, f_2\}, \quad \max\{f_1, f_2\} \tag{1.4}$$

$$\inf_{n\in\mathbb{N}} f_n, \quad \sup_{n\in\mathbb{N}} f_n, \quad \liminf_{n\in\mathbb{N}} f_n, \quad \limsup_{n\in\mathbb{N}} f_n, \ \textit{und } \widetilde{f} \textit{ mit}$$

$$\widetilde{f}(\omega) := \begin{cases} \lim_{n\to\infty} f_n(\omega) & : \quad \textit{wenn der punktweise Limes als Element von } \overline{\mathbb{R}} \textit{ existiert} \\ 0 & : \quad \textit{sonst}. \end{cases}$$

wiederum $\mathcal{F} - \overline{\mathcal{B}}$-messbare Funktionen (wobei wir bei $f_1 + f_2$ und $f_1 - f_2$ zusätzlich voraussetzen, dass für kein $\omega \in \Omega$ der Ausdruck $\infty - \infty$ auftritt). Des weiteren gilt

$$\{f_1 \leq f_2\}, \{f_1 = f_2\}, \{f_1 \geq f_2\} \in \mathcal{F}. \tag{1.5}$$

Beweis. Die Abbildung $\omega \mapsto (f_1(\omega), f_2(\omega))$ ist $\mathcal{F}$-$\mathcal{B}^2$-messbar. Für alle $(a_1, b_1) \times (a_2, b_2)$ gilt nämlich

$$\{(f_1, f_2) \in (a_1, b_1) \times (a_2, b_2)\} = \{f_1 \in (a_1, b_1)\} \cap \{f_2 \in (a_2, b_2)\} \in \mathcal{F},$$

wobei der Schnitt wiederum in $\mathcal{F}$ liegt, da $\mathcal{F}$ eine σ-Algebra ist. Damit sind die Urbilder von Erzeugermengen messbar und mit Proposition 1.15 alle Urbilder von $\mathcal{B}^2$.

Da die Abbildungen $(y_1, y_2) \mapsto y_1 + y_2$ etc. stetig sind, folgt (1.4) aus Proposition 1.16 und Satz 1.18. Aus der Messbarkeit der Abbildung $f_1 - f_2$ folgt dann (1.5), da etwa $\{f_1 \leq f_2\} = \{f_1 - f_2 \in [-\infty, 0]\}$ und $[-\infty, 0] \in \overline{\mathcal{B}}$.

$\{[-\infty, a) \mid a \in \mathbb{R}\}$ ist ein Erzeuger von $\overline{\mathcal{B}}$ (Übungsaufgabe) und es gilt

$$\left(\inf_{n\in\mathbb{N}} f_n\right)^{-1}([-\infty, a)) = \bigcup_{n\in\mathbb{N}} f_n^{-1}([-\infty, a)) \in \mathcal{F}, \quad \forall a \in \mathbb{R}.$$

Damit ist $\inf_{n\in\mathbb{N}} f_n$ messbar. Messbarkeit von sup folgt analog mit $(a,\infty]$ und damit auch die von lim inf und lim sup. Bei $\widetilde{f}$ beachte man, dass der punktweise Limes genau dann existiert wenn lim inf = lim sup. Also gilt

$$\{\widetilde{f} \in B\} = \{\liminf_{n\to\infty} f_n = \limsup_{n\to\infty} f_n\} \cap \{\liminf_{n\to\infty} f_n \in B\}, \quad \forall B \in \overline{\mathcal{B}} \setminus \{0\},$$

woraus mit (1.5) die Messbarkeit von $\widetilde{f}$ folgt (die Herausnahme von $\{0\}$ stört nicht, da $\overline{\mathcal{B}} \setminus \{0\}$ ein Erzeuger von $\overline{\mathcal{B}}$ ist). □

Satz 1.20 (Bildmaß)**.** *Seien $\mathcal{F}$ und $\mathcal{F}'$ σ-Algebren auf Ω bzw. Ω', $f : \Omega \to \Omega'$ eine $\mathcal{F} - \mathcal{F}'$-messbare Abbildung und μ ein Maß auf $(\Omega, \mathcal{F})$. Dann wird durch*

$$\nu(A') := \mu(f^{-1}(A')), \quad \forall A' \in \mathcal{F}'$$

ein Maß auf $(\Omega', \mathcal{F}')$ definiert. ν wird Bildmaß genannt.

Beweis. Es sind lediglich die Eigenschaften (b.i) und (b.ii) in Definition 1.1 nachzuweisen. Es gilt $\mu(f^{-1}(\emptyset)) = \mu(\emptyset) = 0$. Für paarweise disjunkte Mengen A'_n, $n = 1, 2, \ldots$ sind auch die Urbilder $f^{-1}(A'_n)$ disjunkt und es gilt $f^{-1}(\cup_{n\in\mathbb{N}} A'_n) = \cup_{n\in\mathbb{N}} f^{-1}(A'_n)$. Mit der σ-Additivität von μ folgt

$$\mu(f^{-1}(\cup_{n\in\mathbb{N}} A'_n)) = \mu(\cup_{n\in\mathbb{N}} f^{-1}(A'_n)) = \sum_{n=1}^{\infty} \mu(f^{-1}(A'_n)).$$

□

Für $\mu(f^{-1}(A'))$ bzw. $\mu(\{\omega \in \Omega \mid f(\omega) \in A'\})$ schreibt man auch einfach $\mu(f \in A)$, also z.B. $\mu(f > 0)$ oder für $f = (f_1, f_2)$ auch $\mu(f_2 \geq f_1)$.

Beispiel 1.21. *Für eine Zufallvariable: $Y : \Omega \to \mathbb{R}$ bezeichent P_Y das Bildmaß von P under der Abbildung Y also $P_Y(B) := P(Y \in B)$ für alle $B \in \mathcal{B}(\mathbb{R})$.*

Definition 1.22 (Nullmenge)**.** *Eine Menge $N \in \mathcal{F}$ heißt eine μ-Nullmenge, wenn $\mu(N) = 0$.*

Sei E eine Eigenschaft, die in Abhängigkeit von ω wahr oder falsch ist. Wir sagen E gilt μ-fast überall (abkürzend P-f.ü.), wenn die Menge $\{\omega \in \Omega \mid E(\omega)$ ist falsch$\}$ eine μ-Nullmenge ist.

Gilt $\mu(\Omega) = 1$ und interpretiert man μ als Wahrscheinlichkeitsmaß, sagt man auch, dass eine Aussage „fast sicher" (f.s.) gilt, statt „fast überall" (f.ü.).

Definition 1.23 (Vollständigkeit)**.** *Ein Maßraum $(\Omega, \mathcal{F}, \mu)$ heißt vollständig, wenn für alle Mengen $A, B \subset \Omega$ folgende Implikation gilt*

$$A \subset B,\ B \in \mathcal{F}, \mu(B) = 0 \Rightarrow A \in \mathcal{F} \quad (\textit{und damit wegen Monotonie auch } \mu(A) = 0)$$

Vollständigkeit bedeutet also, dass Teilmengen von Nullmengen wiederum Nullmengen sind.

Bemerkung 1.24 (Vervollständigung eines Maßraums)**.** *Jeder Maßraum lässt sich vervollständigen. Zu einem i.A. nicht vollständigen Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mu)$ definiere man das Mengensystem*

$$\widetilde{\mathcal{F}} := \{\widetilde{A} \subset \Omega \mid \exists A_1, A_2 \in \mathcal{F} \textit{ mit } A_1 \subset \widetilde{A} \subset A_2 \textit{ und } \mu(A_2 \setminus A_1) = 0\}$$

und die Abbildung

$$\widetilde{\mu} : \widetilde{\mathcal{F}} \to [0, 1], \quad \widetilde{A} \mapsto \mu(A_1) = \mu(A_2).$$

Es ist zu zeigen

(i) $\widetilde{\mathcal{F}}$ *ist eine σ-Algebra.*

(ii) $\widetilde{\mu}$ *ist wohldefiniert, d.h. $\widetilde{\mu}(\widetilde{A})$ hängt zwar von der Menge $\widetilde{A}$ ab, nicht aber vom Paar $(A_1, A_2) \in \mathcal{F} \times \mathcal{F}$ für das $A_1 \subset \widetilde{A} \subset A_2$ und $\mu(A_2 \setminus A_1) = 0$ gilt.*

(iii) $\widetilde{\mu}$ *ist ein Maß.*

Zudem kann man zeigen, dass jede σ-Algebra, die $\mathcal{F}$ umfasst und die bzgl. eines erweiteren Maßes vollständig ist, auch die σ-Algebra $\widetilde{\mathcal{F}}$ umfasst. Die Fortsetzung ist also minimal (alles Übung, man beachte, dass das Maß nicht endlich und schon nicht einmal σ-endlich sein muss, der Ausdruck $\mu(A_2) - \mu(A_1)$ kann also undefiniert sein).

In der Wahrscheinlichkeitstheorie vervollständigt man den Maßraum (der dann Wahrscheinlichkeitsraum genannt wird) oft. Auch wenn das Ereignis $\{\omega \in \Omega \mid E(\omega) \text{ ist falsch}\}$

nicht in $\mathcal{F}$ liegt (oder man dies zumindest nicht beweisen kann), so ist es im vervollständigten Raum bereits eine Nullmenge, wenn es Teilmenge einer Nullmenge im ursprünglichen Raum ist.

Allerdings braucht man die Vollständigkeit des Maßraumes nur selten. Zumeist kann man von Kandidaten für Nullmengen auch ihre Messbarkeit zeigen. In dieser Vorlesung muss Vollständigkeit nicht gefordert werden. Trotzdem ist Bemerkung 1.24 von Interesse, da es eine positive Antwort für eine einfache Fortsetzbarkeit eines Maßes gibt.

2 Das Integral

Zur Definition des Integrals beginnen wir mit sog. **Elementarfunktionen**, das sind reellwertige messbare Funktionen, die nur endlich viele verschiedene Werte annehmen. Zudem sollen die Funktionen zunächst nichtnegativ sein. Nichtnegative Elementarfunktionen sind also Funktionen mit Darstellung

$$f = \sum_{k=1}^{n} \alpha_k 1_{A_k}, \quad n \in \mathbb{N}, \alpha_1, \dots, \alpha_n \in \mathbb{R}_+,\ A_1, \dots, A_n \in \mathcal{F}. \tag{2.1}$$

Man beachte, dass für eine $\mathcal{F} - \mathcal{B}(\mathbb{R})$-messbare Abbildung f und ein $\alpha \in \mathbb{R}$

$$\{f = \alpha\} \in \mathcal{F}$$

gilt. Die Menge der Funktionen, die wie in (2.1) darstellbar sind, bezeichnen wir mit $\mathbb{E}^+$. Die entsprechende Menge ohne die Einschränkung, dass $\alpha_k \geq 0$ wird mit $\mathbb{E}$ bezeichnet. Die rechte Seite von (2.1) nennt man eine **Normaldarstellung** von f, wenn die Mengen $A_1, A_2, \dots, A_n$ paarweise **disjunkt** sind.

Lemma 2.1. *Seien $f = \sum_{k=1}^n \alpha_k 1_{A_k}$ und $f = \sum_{l=1}^m \beta_l 1_{B_l}$ zwei Normaldarstellungen von $f \in \mathbb{E}^+$ und μ ein Maß. Es gilt*

$$\sum_{k=1}^{n} \alpha_k \mu(A_k) = \sum_{l=1}^{m} \beta_l \mu(B_l).$$

Man beachte, dass $\mu(A_k) = \infty$ nicht ausgeschlossen ist. Wegen der Nichnegativität der α_k sind die Summen aber stets wohldefiniert als Elemente in $\overline{\mathbb{R}}_+$. Man beachte zudem die Konvention $0 \cdot \infty := 0$.

Beweis. Da die Mengen disjunkt sind, muss für alle Paare (k,l) mit $\mu(A_k \cap B_l) > 0$ gelten, dass $\alpha_k = \beta_l$. Außerdem gilt $A_k \subset \cup_{l=1,\dots,m} B_l$ für alle k mit $\alpha_k \neq 0$ und $B_l \subset \cup_{k=1,\dots,n} A_k$ für alle l mit $\beta_l \neq 0$. Es folgt

$$\sum_{k=1}^{n} \alpha_k \mu(A_k) = \sum_{k=1}^{n} \sum_{l=1}^{m} \alpha_k \mu(A_k \cap B_l) = \sum_{l=1}^{m} \sum_{k=1}^{n} \beta_l \mu(A_k \cap B_l) = \sum_{l=1}^{m} \beta_l \mu(B_l).$$

□

Definition 2.2. *Das Elementarintegral für nichtnegative Funktionen ist durch die Abbildung $I : \mathbb{E}^+ \to \overline{\mathbb{R}}_+$ mit*

$$I(f) := \sum_{k=1}^{n} \alpha_k \mu(A_k) \tag{2.2}$$

definiert, wobei $f = \sum_{k=1}^{n} \alpha_k 1_{A_k}$ eine Normaldarstellung von $f \in \mathbb{E}^+$ ist.

(Lemma 2.1 sichert die Wohldefiniertheit von $I(f)$, d.h. die rechte Seite von (2.2) hängt nicht von der konkreten Darstellung von f ab)

Eine Elementarfunktion besitzt immer eine kanonische Darstellung $f = \sum_z z 1_{\{f=z\}}$, wobei die Summe über die endlich vielen Werte läuft, die f annehmen kann. Alternativ kann man das Elementarintegral direkt über die kanonische Darstellung definieren, was Lemma 2.1 überflüssig machen würde. Die Einführung von f wie in (2.1) mag aber die Notation etwas vereinfachen (etwa wenn man eine Darstellung für die Summe zweier Elementarfunktionen aufschreibt, siehe den Beweis des folgenden Lemmas).

Bemerkung 2.3. *Man beachte, dass sich die Eigenschaft ,,elementar" nur auf den Bild- nicht aber auf den Urbildbereich der Funktion bezieht. Im Fall $(\Omega, \mathcal{F}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ müssen die Mengen A_k keine einfachen Intervalle sein. Die Menge der Punkte, die auf ein bestimmtes α_k abgebildet werden, kann eine komplizierte Borelmenge sein. Nur wegen den Vorarbeiten aus der Maßtheorie können wir diesen komplizierten Mengen, eine Masse zuordnen, etwa gemäß dem Lebesgue-Maß, wenn eine gleichmäßige Gewichtung gewünscht ist.*

Lemma 2.4. *Für $f, g \in \mathbb{E}^+$ und $\alpha, \beta \in \mathbb{R}_+$ gilt*

(i) $I(\alpha f + \beta g) = \alpha I(f) + \beta I(g)$

(ii) $f \leq g \implies I(f) \leq I(g)$

Beweis. Seien $f = \sum_{k=1}^{n} \alpha_k 1_{A_k}$ und $g = \sum_{l=1}^{m} \beta_l 1_{B_l}$. Beide Aussagen rechnet man schnell nach, indem man f, g und ggf. $\alpha f + \beta g$ über $1_{A_k \cap B_l}$, $k = 1, \dots, n$, $l = 1, \dots, m$ darstellt. Man beachte, dass $A_{k_1} \cap B_{l_1}$ und $A_{k_2} \cap B_{l_2}$ für $k_1 \neq k_2$ oder $l_1 \neq l_2$ disjunkt sind. Damit erhält man wieder eine Normaldarstellung (An dieser Stelle wird klar, wieso man bei der Normaldarstellung nicht gefordert hat, dass die Vorfaktoren α_k unterschiedlich sind). Zunächst setzen wir $A_{n+1} := (\cup_{k=1,\dots,n} A_k)^c$, $B_{m+1} := (\cup_{k=1,\dots,m} B_k)^c$ und $\alpha_{n+1} = \beta_{m+1} = 0$. Dies ist nützlich, da die Mengen dann eine Partition von Ω bilden. Für diese Normaldarstellung gilt

$$\alpha \sum_{k=1}^{n+1} \alpha_k 1_{A_k} + \beta \sum_{l=1}^{m+1} \beta_l 1_{B_l} = \sum_{k=1}^{n+1} \sum_{l=1}^{m+1} (\alpha \alpha_k + \beta \beta_k) 1_{A_k \cap B_l}$$

und damit

$$\begin{aligned} I(\alpha f + \beta g) &= \sum_{k=1}^{n+1} \sum_{l=1}^{m+1} (\alpha \alpha_k + \beta \beta_l) \mu(A_k \cap B_l) \\ &= \sum_{k=1}^{n+1} \alpha \alpha_k \mu(A_k) + \sum_{l=1}^{m+1} \beta \beta_l \mu(B_l) \\ &= \alpha I(f) + \beta I(g). \end{aligned}$$

Ad (ii): Sei $f \leq g$. Für alle Paare (k, l) mit $\mu(A_k \cap B_l) > 0$ gilt $\alpha_k \leq \beta_l$ und damit

$$I(f) = \sum_{k=1}^{n+1} \alpha_k \mu(A_k) = \sum_{k=1}^{n+1} \sum_{l=1}^{m+1} \alpha_k \mu(A_k \cap B_l) \leq \sum_{k=1}^{n+1} \sum_{l=1}^{m+1} \beta_l \mu(A_k \cap B_l) = \sum_{l=1}^{m+1} \beta_l \mu(B_l) = I(g).$$

□

Definition 2.5 (Integral für nichtnegative Funktionen)**.** *Sei* $f : \Omega \to \overline{\mathbb{R}}_+$ $\mathcal{F} - \mathcal{B}(\overline{\mathbb{R}}_+)$-*messbar. Wir definieren das Integral von* f *bzgl.* μ *durch*

$$\int f \, d\mu := \sup\{I(h) \mid h \in \mathbb{E}^+, \; h \leq f\}.$$

Bemerkung 2.6. *Natürlich würde Definition 2.5 zunächst auch für nicht-messbare Funktionen* f *Sinn ergeben. Einige der im folgenden hergeleiteten Eigenschaften des Integrals würden jedoch verlorengehen.*

Bemerkung 2.7. *Aus Lemma 2.4(ii) folgt, dass für nichtnegative Elementarfunktionen das Integral in Definition 2.5 mit dem Elementarintegral übereinstimmt.*

Satz 2.8. *Für $\overline{\mathbb{R}}_+$-wertige, messbare Funktionen f, g und $\alpha, \beta \in \mathbb{R}_+$ gilt*

(i) $f \le g$ *μ-f.ü.* $\Rightarrow\ \int f\,d\mu \le \int g\,d\mu$

(ii) $f = 0$ *μ-f.ü.* $\Leftrightarrow\ \int f\,d\mu = 0$

(iii) $\int f\,d\mu < \infty \Rightarrow\ f < \infty$ *μ-f.ü.*

(Man beachte die Konvention $0 \cdot \infty := 0$*)*

Beweis. Ad (i): Sei $h \ge 0$ eine Elementarfunktion mit $h \le f$. Es folgt, dass $\widetilde{h} := h1_{\{f \le g\}}$ ebenfalls eine Elementarfunktion ist, für die zudem gilt $\widetilde{h} \le g$ (ohne Ausnahmenullmenge). Wegen $\mu(f > g) = 0$ gilt für die Elementarintegrale offensichtlich $I(\widetilde{h}) = I(h)$. (i) folgt dann aus der Definition.

Ad (ii): Es gilt die Äquivalenz

$$\int f\,d\mu = 0 \Leftrightarrow \big(\forall h \in \mathbb{E}^+, h \le f \implies I(h) = 0\big)\,. \tag{2.3}$$

Zudem ist für Elementarfunktionen $h \in \mathbb{E}^+$ klar, dass $\mu(h > 0) = 0 \Leftrightarrow I(h) = 0$. Für f mit $\mu(f > 0) = 0$ ist die rechte Seite der Äquivalenz (2.3) damit offenbar erfüllt. Umgekehrt kann man aus der rechten Seite der Äquivalenz (2.3) folgern, dass für jedes $n \in \mathbb{N}$ gilt $\mu(f \ge 1/n) = 0$ (betrachte dazu die Elementarfunktion $h = \frac{1}{n}1_{\{f \ge 1/n\}}$). Wegen

$$\{f \ge 1/n\} \uparrow \{f > 0\}, \quad n \uparrow \infty,$$

folgt die Behauptung mit der Stetigkeit von unten von μ (Proposition 1.4(iii)).

Ad (iii): Betrachte die Elementarfunktionen $n1_{\{f=\infty\}}$, $n \in \mathbb{N}$.

□

Bemerkung 2.9. *Nach Satz 2.8(ii) gilt* $\int \infty 1_N\,d\mu = 0$ *für* $\mu(N) = 0$.

Satz 2.10 (Satz von der monotonen Konvergenz, auch Satz von Beppo Levi genannt)**.** *Sei (f_n) eine Folge von $\overline{\mathbb{R}}_+$-wertigen messbaren Funktionen mit $f_1 \le f_2 \le \ldots$ und $f :=$*

$\sup_{n\in\mathbb{N}} f_n = \lim_{n\to\infty} f_n$ *(nach Proposition 1.19 ist die $\overline{\mathbb{R}}_+$-wertige Funktion f wiederum messbar). Es gilt*

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

Beweis. Aus der Monotonie folgt die Existenz des Limes und $\lim_{n\to\infty}\int f_n\,d\mu \le \int f\,d\mu$. Bleibt ≥ 0 zu zeigen. Dazu nehme eine nichtnegative Elementarfunktion h mit $h \le f$. Zu festem $\varepsilon > 0$ betrachte die Elementarfunktionen $h_n := (h-\varepsilon)^+ 1_{\{f_n > h-\varepsilon\}}$, wobei $(h(\omega)-\varepsilon)^+ = \max\{h(\omega)-\varepsilon, 0\}$. Es folgt $h_n \le f_n$ und damit $I(h_n) \le \int f_n\,d\mu$. Andererseits gilt wegen $\{f_n \ge h-\varepsilon\} \uparrow \Omega$ für $n \uparrow \infty$ (man beachte, dass Elementarfunktionen nur endliche Werte annehmen können) und der Stetigkeit von unten von μ

$$I(h_n) = \sum_z (z-\varepsilon)^+ \mu(f_n > h-\varepsilon,\ h = z) \overset{n\to\infty}{\to} \sum_z (z-\varepsilon)^+\mu(h=z)$$

und damit $\lim_{n\to\infty}\int f_n\,d\mu \ge \sum_z (z-\varepsilon)^+\mu(h=z)$. Zudem gilt $\sum_z (z-\varepsilon)^+\mu(h=z) \to \sum_z z\mu(h=z) = I(h)$ für $\varepsilon \downarrow 0$ (im Fall $z>0$ gilt $(z-\varepsilon)^+ > 0$ für ε klein genug und im Fall $z = 0$ gilt $(z-\varepsilon)^+ = 0$ für alle $\varepsilon > 0$, man beachte, dass $\mu(h=z) = \infty$ möglich ist). Es folgt, dass $\lim_{n\to\infty}\int f_n\,d\mu \ge I(h)$ (beachte, dass $(f_n)_{n\in\mathbb{N}}$ nicht von $\varepsilon > 0$ abhängt). Da h eine beliebige nichtnegative Elementarfunktion mit $h \le f$ war, folgt $\lim_{n\to\infty}\int f_n\,d\mu \ge \int f\,d\mu$ und damit die Behauptung. □

Lemma 2.11 (Lemma von Fatou)**.** *Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge $\overline{\mathbb{R}}_+$-wertiger messbarer Funktionen. Es gilt*

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

Beispiel für „<": $f_n := n1_{[0,1/n]}$ und μ Lebesgue-Maß auf $[0,1]$. Es gilt $\int f_n\,d\mu = 1$ für alle $n\in\mathbb{N}$, aber $\int \liminf_{n\to\infty} f_n\,d\mu = \int \infty 1_{\{0\}}\,d\mu = 0$.

Weiteres Beispiel: Verdoppelungsstrategien im Spielkasino.

Beweis. Definiere $g_n := \inf_{m\ge n} f_m$. Die Folge $(g_n)_{n\in\mathbb{N}}$ ist monoton aufsteigend mit $g_n \le f_n$ und $\lim_{n\to\infty} g_n = \liminf_{n\to\infty} f_n$. Mit Satz 2.10 und Satz 2.8(i) folgt

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

□

Satz 2.12 (Linearität). *Für $\overline{\mathbb{R}}_+$-wertige, messbare Funktionen f, g und $\alpha, \beta \in \mathbb{R}_+$ gilt*

$$\int (\alpha f + \beta g)\, d\mu = \alpha \int f\, d\mu + \beta \int g\, d\mu.$$

(Man beachte die Konvention $0 \cdot \infty := 0$)

Beweis. Es gibt Folgen nichnegativer Elementarfunktionen, die monoton gegen f bzw. g aufsteigen. Wähle etwa $f_n := 2^{-n}\lfloor 2^n(f \wedge n)\rfloor$, $g_n := 2^{-n}\lfloor 2^n(g \wedge n)\rfloor$, wobei $\lfloor x\rfloor := \max\{m \in \mathbb{N}_0 \mid m \leq x\}$. Offenbar sind auch $\alpha f_n + \beta g_n$ nichtnegative Elementarfunktionen und die Folge steigt monoton gegen $f + g$ auf. Mit Satz 2.10 und Lemma 2.4(i) folgt

$$\begin{aligned}
\alpha \int f\, d\mu + \beta \int g\, d\mu &\overset{\text{Satz 2.10}}{=} \alpha \lim_{n\to\infty} \int f_n\, d\mu + \beta \lim_{n\to\infty} \int g_n\, d\mu \\
&= \lim_{n\to\infty} \left(\alpha \int f_n\, d\mu + \beta \int g_n\, d\mu \right) \\
&\overset{\text{Lemma 2.4(i)}}{=} \lim_{n\to\infty} \left(\int (\alpha f_n + \beta g_n)\, d\mu \right) \\
&\overset{\text{Satz 2.10}}{=} \int (\alpha f + \beta g)\, d\mu.
\end{aligned}$$

□

Das Integral für beliebige messbare Funktionen kann auf das Integral aus Definition 2.5 für nichtnegative Funktionen zurückgeführt werden. Dazu definieren wir $f^+(\omega) := \max\{f(\omega), 0\}$ und $f^-(\omega) := \max\{-f(\omega), 0\}$.

Definition 2.13. *Falls $\int f^+\, d\mu < \infty$* **oder** *$\int f^-\, d\mu < \infty$, dann sei das Integral definiert durch*

$$\int f\, d\mu := \int f^+\, d\mu - \int f^-\, d\mu \tag{2.4}$$

(mit den üblichen Konventionen $\infty - a := \infty$ für alle $a \in \mathbb{R}$ und $a - \infty := -\infty$ für alle $a \in \mathbb{R}$)

Man beachte, dass für $f \geq 0$ gilt $f^+ = f$ und $f^- = 0$. (2.4) ist also offensichtlich eine Fortsetzung des Integrals für nichtnegative Funktionen, so dass auf ein neues Symbol für das Integral verzichtet wird.

f heißt μ-integrierbar falls $\int f^+ \, d\mu < \infty$ **und** *$\int f^- \, d\mu < \infty$. Der Raum der messbaren, μ-integrierbaren Funktionen wird mit $\mathcal{L}^1(\Omega, \mathcal{F}, \mu)$ bezeichnet .*

Aus $|f| = f^+ + f^-$ und Satz 2.12 folgt

$$\int |f| \, d\mu = \int f^+ \, d\mu + \int f^- \, d\mu. \tag{2.5}$$

f ist also genau dann integrierbar, wenn $\int |f| \, d\mu < \infty$.

Beispiel 2.14 (Summen als Integrale)**.** *Sei $\Omega = \mathbb{N}$, $\mathcal{F} = 2^{\mathbb{N}}$ und μ das* **Zählmaß***, d.h. $\mu(A) := \#A$ (Anzahl der Elemente der Menge A). Eine Folge $(a_n)_{n \in \mathbb{N}}$ reeller Zahlen lässt sich als Abbildung $f : \mathbb{N} \to \mathbb{R}$ interpretieren, also*

$$a_n = f(n).$$

Für nichtnegative Folgen gilt dann

$$\int f \, d\mu = \sum_{n=1}^{\infty} a_n \qquad (:= \lim_{m \to \infty} \sum_{n=1}^{m} a_n). \tag{2.6}$$

Klar, die Elementarfunktionen $f_m := f 1_{\{1,\dots,m\}}$ lassen sich in der Normaldarstellung $f_m = \sum_{n=1}^{m} a_n 1_{\{n\}}$ schreiben und besitzen das (Elementar-)Integral $I(f_m) = \sum_{n=1}^{m} a_n$. Mit $f_m \uparrow f$ für $m \uparrow \infty$ und dem Satz von der monotonen Konvergenz folgt (2.6).

Bei Folgen, die positive und negative Folgenglieder besitzen, gilt (2.6) ebenso, sofern $\int f \, d\mu$ im Sinne von Definition 2.13 existiert. f ist genau dann integriebar, wenn $\sum_{n=1}^{\infty} |a_n| < \infty$. Man beachte jedoch, dass für

$$f(n) = a_n = (-1)^n \frac{1}{n}$$

das Integral nicht erklärt ist, da $\int f^+ \, d\mu = \int f^- \, d\mu = \infty$. Anders als bei Reihen, die man als Grenzwert der Partialsummen definiert, kann es beim Integral nicht darauf ankommen, in welcher Reihenfolge, man die a_n aufsummiert (Folge des Satzes von der monotonen Konvergenz und der Konstruktion in Definition 2.13), so dass das Integral nur existieren kann, wenn die Reihenfolge unerheblich ist.

Analog bezeichnet $\mathcal{L}^p(\Omega, \mathcal{F}, \mu)$ für $p \in [1, \infty)$ den Raum der messbaren Funktionen f, s.d. $|f|^p$ μ-integrierbar ist, also

$$\mathcal{L}^p(\Omega, \mathcal{F}, \mu) := \{f : \Omega \to \mathbb{R} \mid f \text{ ist } \mathcal{F} - \mathcal{B}(\mathbb{R}) - \text{messbar mit } \int |f|^p \, d\mu < \infty\}.$$

$\mathcal{L}^0(\Omega, \mathcal{F}, \mu)$ bezeichnet den Raum der $\mathbb{R}$-wertigen messbaren Funktionen, also

$$\mathcal{L}^0(\Omega, \mathcal{F}, \mu) := \{f : \Omega \to \mathbb{R} \mid f \text{ ist } \mathcal{F} - \mathcal{B}(\mathbb{R}) - \text{messbar}\}.$$

$\mathcal{L}^\infty(\Omega, \mathcal{F}, \mu)$ bezeichnet den Raum der messbaren Funktionen f für die ein $M \in \mathbb{R}_+$ existiert mit $\mu(|f| > M) = 0$ (f heißt dann essentiell beschränkt), also

$$\begin{aligned}\mathcal{L}^\infty(\Omega, \mathcal{F}, \mu) \;&:=\; \{f : \Omega \to \mathbb{R} \mid f \text{ ist } \mathcal{F} - \mathcal{B}(\mathbb{R}) - \text{messbar und es existiert} \\ &\qquad\qquad \text{ein } M \in \mathbb{R}_+, \text{ so dass } \mu(|f| > M) = 0\}.\end{aligned}$$

(Manchmal werden in der Literatur bei $\mathcal{L}^p(\Omega, \mathcal{F}, \mu)$ auch Funktionen mit Werten in $\overline{\mathbb{R}}$ zugelassen. Wegen Satz 2.8(iii) können die Werte ∞ und $-\infty$ jedoch nur auf einer μ-Nullmenge angenommen werden, so dass diese Erweiterung keinen entscheidenden Unterschied macht)

Mit $L^p(\Omega, \mathcal{F}, \mu)$, $p \in [1, \infty)$, $L^0(\Omega, \mathcal{F}, \mu)$ und $L^\infty(\Omega, \mathcal{F}, \mu)$ bezeichnet man die entsprechenden **Quotientenräume** der Äquivalenzklassen, die entstehen, wenn Funktionen, die μ-f.ü. übereinstimmen miteinander identifiziert werden. Also

$$L^p(\Omega, \mathcal{F}, \mu) := \mathcal{L}^p(\Omega, \mathcal{F}, \mu)/\mathcal{N} = \{\overline{f} := f + \mathcal{N} \mid f \in \mathcal{L}^p\}$$

etc., wobei $\mathcal{N} := \{g \in \mathcal{L}^0(\Omega, \mathcal{F}, \mu) \mid \mu(g \neq 0) = 0\}$ und $f + \mathcal{N} := \{g : \Omega \to \mathbb{R} \mid \exists h \in \mathcal{N} \text{ mit } g = f + h\} = \{g : \Omega \to \mathbb{R} \mid g \text{ ist } \mathcal{F} - \mathcal{B}(\mathbb{R}) - \text{messbar und } \mu(f \neq g) = 0\}$.

Bemerkung 2.15. *L^p hat gegenüber $\mathcal{L}^p$ den Vorteil, dass zwischen Funktionen, die μ-fast überall übereinstimmen, nicht mehr unterschieden wird. Es kommt oft vor, dass gewisse Objekte durch ihre charakteristischen Eigenschaften nur bis auf eine Nullmenge eindeutig bestimmt sind (in dieser Vorlesung wir dies etwa die Radon-Nikodym-Ableitung sein). Die Arbeit mit Äquivalenzklassen erlaubt es dann, von* **der** *Radon-Nikodym-Ableitung zu sprechen. Da es aus der Äquivalenzklasse $f + \mathcal{N}$ i.A. keinen kanonischen Repräsentanten gibt, ist dies praktisch. Man wendet dann Operationen auf die Äquivalenzklasse $f + \mathcal{N}$ an, indem man Operationen auf Elemente $g \in f + \mathcal{N}$ anwendet und zeigt, dass das Ergebnis unabhängig von der Wahl des Repräsentanten ist.*

Auf $L^p(\Omega, \mathcal{F}, \mu)$ definiert man die Norm

$$||f||_p := \left(\int |f|^p \, d\mu \right)^{1/p}. \tag{2.7}$$

Auf $L^\infty(\Omega, \mathcal{F}, \mu)$ definiert man die Norm

$$||f||_\infty := \inf\{M \geq 0 \mid \mu(|f| > M) = 0\}. \tag{2.8}$$

Das Infimum wird angenommen, d.h. $\mu(|f| > ||f||_\infty) = 0$ (wieso ?)

(Strenggenommen ist also das f auf der linken Seite eine Äquivalenzklasse und auf der rechten Seite ein Repräsentant aus dieser Äquivalenzklasse. Aus Satz 2.8(i) folgt dann, dass $\int |f|^p \, d\mu$ für alle Repräsentanten der gleichen Äquivalenzklasse übereinstimmt)

Es muss noch gezeigt werden, dass $||f||_p$ für alle $p \in [1, \infty]$ tatsächlich eine Norm ist, was wir nach ein paar Vorarbeiten noch machen werden. Probleme bereitet nur die Dreiecksungleichung. Die Eigenschaft einer Norm, dass $||f||_p = 0 \implies f = 0$ gilt offenbar wegen Satz 2.8(ii), wenn wir f als Äquivalenzklasse mit Nullelement $\mathcal{N}$ auffassen.

Es gelten folgende Eigenschaften des Integrals.

Satz 2.16. *Seien f und g Funktionen s.d. die Integrale nach μ (im Sinne von Definition 2.13) definiert sind.*

(i) (Monotonie) Ist $f \leq g$ μ-f.ü., so ist $\int f \, d\mu \leq \int g \, d\mu$.

(ii) (Linearität) Für alle $\alpha, \beta \in \mathbb{R}$ und f, g integrierbar gilt $\int (\alpha f + \beta g) \, d\mu = \alpha \int f \, d\mu + \beta \int g \, d\mu$.

(iii) (Dreiecksungleichung) Es gilt $|\int f \, d\mu| \leq \int |f| \, d\mu$.

Beweis. Ad (i): Es folgt $f^+ \leq g^+$ und $f^- \leq g^-$ μ-f.ü. und damit mit Satz 2.8(i) $\int f^+ \, d\mu \leq \int g^+ \, d\mu$ und $\int f^- \, d\mu \leq \int g^- \, d\mu$. Es folgt die Behauptung.

Ad(ii) Nun wollen wir für beliebige f, g, α, β die Linearität auf den nichtnegativen Fall zurückführen, der ja schon gezeigt ist. Bekanntlich reicht es, Additivität und Homogenität getrennt zu zeigen. Es gilt

$$(f + g)^+ - (f + g)^- = f + g = f^+ - f^- + g^+ - g^-. \tag{2.9}$$

Aus der Additivität in Satz 2.12 folgt

$$\int (f+g)^+ \, d\mu + \int f^- \, d\mu + \int g^- \, d\mu = \int (f+g)^- \, d\mu + \int f^+ \, d\mu + \int g^+ \, d\mu \quad (2.10)$$

und damit Additivität (man beachte, dass wir bei (ii) zusätzlich Integrierbarkeit vorausgesetzt haben, weswegen alle in (2.10) auftretenden Summanden endlich sind). Für $\alpha \geq 0$ folgt

$$\int \alpha f \, d\mu = \int \alpha f^+ \, d\mu - \int \alpha f^- \, d\mu = \alpha \int f^+ \, d\mu - \alpha \int f^- \, d\mu = \alpha \int f \, d\mu$$

aus der Homogenität in Satz 2.12. Für $\alpha \leq 0$ gilt $(\alpha f)^+ = (-\alpha) f^-$ und $(\alpha f)^- = (-\alpha) f^+$ und damit folgt

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

Ad (iii): Folgt sofort aus (2.5).

□

Proposition 2.17. *Sei μ endlich und $1 \leq p_1 < p_2 < \infty$. Für alle messbaren Funktionen f gilt die Implikation*

$$\int |f|^{p_2} \, d\mu < \infty \implies \int |f|^{p_1} \, d\mu < \infty.$$

Beweis. Es gilt

$$\int |f|^{p_1} \, d\mu \leq \int \left(|f|^{p_2} 1_{\{|f| \geq 1\}} + 1 \right) \, d\mu = \int \left(|f|^{p_2} 1_{\{|f| \geq 1\}} \right) \, d\mu + \mu(\Omega).$$

□

Für das Lebesgue-Maß λ auf $\mathbb{R}$ sieht man dagegen, dass $\mathcal{L}^{p_1}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda) \not\subset \mathcal{L}^{p_2}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ und $\mathcal{L}^{p_2}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda) \not\subset \mathcal{L}^{p_1}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$.

Satz 2.18 (Hölder-Ungleichung)**.** *Seien $p, q > 1$ mit $1/p + 1/q = 1$ (p, q heißen dann konjugiert) und $f \in \mathcal{L}^p(\Omega, \mathcal{F}, \mu)$, $g \in \mathcal{L}^q(\Omega, \mathcal{F}, \mu)$. Die Funktion fg ist dann integrierbar und es gilt*

$$\left| \int fg \, d\mu \right| \leq \left(\int |f|^p \, d\mu \right)^{1/p} \left(\int |g|^q \, d\mu \right)^{1/q}.$$

Im Fall $p = q = 2$ wird dies auch die **Cauchy-Schwarz-Ungleichung** *genannt.*

Beweis. Mit der Youngschen Ungleichung

$$ab \leq \frac{1}{p}a^p + \frac{1}{q}b^q, \quad \forall a, b \geq 0$$

(die sich aus der Konkavität der Logarithmusfunktion ergibt), folgt für alle Zahlen $\alpha, \beta > 0$

$$\frac{1}{\alpha\beta}\left|\int fg\,d\mu\right| \leq \int \frac{|f|}{\alpha}\frac{|g|}{\beta}\,d\mu \leq \int \frac{|f|^p}{p\alpha^p} + \frac{|g|^q}{q\beta^q}\,d\mu = \frac{1}{p\alpha^p}\int |f|^p\,d\mu + \frac{1}{q\beta^q}\int |g|^q\,d\mu.$$

Die Behauptung folgt mit der Wahl von $\alpha := \left(\int |f|^p\,d\mu\right)^{1/p}$ und $\beta := \left(\int |g|^q\,d\mu\right)^{1/q}$ unter Beachtung, dass der Fall $\mu(f \neq 0) = 0$ oder $\mu(g \neq 0) = 0$ trivial ist. □

Satz 2.19 (Minkowski-Ungleichung)**.** *Für $p \in [1, \infty]$ und messbare Funktionen f, g gilt*

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

$||\cdot||_p$ erfüllt also die Dreiecksungleichung. Damit ist gezeigt, dass es tatsächlich eine Norm ist.

Beweis. Sei $p \in (1, \infty)$ (die anderen Fälle folgen sofort). Mit $q := p/(p-1)$ (d.h. $1/p + 1/q = 1$) gilt

$$\begin{aligned}
\int |f+g|^p\,d\mu &\leq \int |f||f+g|^{p-1}\,d\mu + \int |g||f+g|^{p-1}\,d\mu \\
&\overset{\text{Hölder}}{\leq} \left(\left(\int |f|^p\,d\mu\right)^{1/p} + \left(\int |g|^p\,d\mu\right)^{1/p}\right)\left(\int |f+g|^{(p-1)q}\,d\mu\right)^{1/q}.
\end{aligned}$$

Dies ergibt wegen $(p-1)q = p$ und $1 - 1/q = 1/p$ die Behauptung. □

Wichtig für Anwendungen: Summen integrierbarer Funktionen sind integrierbar, Produkte i.A. nicht. Sind die Funktionen quadratintegrierbar, so ist das Produkt integrierbar.

Bemerkung 2.20. *In der Stochastik wird der Erwartungswert einer reellwertigen Zufallsvariablen $Y : \Omega \to \mathbb{R}$ bzgl. eines Wahrscheinlichkeitsmaßes P als Lebesgueintegral $\int Y\,dP =: E_P(Y)$ definiert. Gegenüber dem Riemann-Integral hat dies den Vorteil, dass auf dem Grundraum Ω keine topologische Struktur benötigt wird (wie etwa die euklidische Norm auf dem $\mathbb{R}^n$), um Zufallsvariablen etwa als stetige Abbildungen zu definieren.*

3 Einige Konvergenzbegriffe

Im Folgenden sei μ ein σ-endliches (aber i.A. nicht endliches) Maß.

Definition 3.1. *Sei* $(f_n)_{n\in\mathbb{N}} \subset \mathcal{L}^0(\Omega, \mathcal{F}, \mu)$ *und* $f \in \mathcal{L}^0(\Omega, \mathcal{F}, \mu)$.

(i) Die Folge $(f_n)_{n\in\mathbb{N}}$ *konvergiert im Maße* μ *gegen* f*, wenn für jedes* $\varepsilon > 0$ *und jede Menge* $A \in \mathcal{F}$ *mit* $\mu(A) < \infty$ *gilt*

$$\mu(\{|f_n - f| > \varepsilon\} \cap A) \to 0, \quad n \to \infty.$$

Wird μ *als Wahrscheinlichkeitsmaß interpretiert, dann spricht man von* **stochastischer Konvergenz***.*

(Wenn $\mu(\Omega) < \infty$*, kann der Schnitt mit den Mengen* A *von endlichem Maß auch weggelassen werden, ohne dass sich irgendetwas ändert)*

(ii) $(f_n)_{n\in\mathbb{N}}$ *konvergiert* μ*-fast überall gegen* f*, wenn* $\mu(f_n \not\to f,\ n \to \infty) = 0$.

(iii) Sei $p \in [1, \infty)$. $(f_n)_{n\in\mathbb{N}}$ *konvergiert in* $L^p(\mu)$ *gegen* f*, wenn*

$$\int |f_n - f|^p \, d\mu \to 0, \quad n \to \infty.$$

Bemerkung 3.2. *Ohne die* σ*-Endlichkeit von* μ *wäre Definition 3.1(i) unsinnig, da man dann mit Mengen von endlichem Maß den Gesamtraum nicht ausschöpfen könnte. Für das Maß, das jede nichtleere Menge auf Unendlich abbildet, würde alles konvergieren, auch wenn man bei einem großen Maß eher Äquivalenz zur gleichmäßigen Konvergenz erwarten würde.*

In der Literatur finden sich zwei verschiedene Definitionen der ,,Konvergenz im Maße". Der Schnitt mit Mengen von endlichem Maß wird manchmal auch weggelassen. Im Falle $\mu(\Omega) < \infty$ *fallen die beiden Definitionen natürlich zusammen. Bei Definition 3.1(i) könnte man auch von einer ,,lokalen Konvergenz im Maße" sprechen. Ist* μ *das Lebesgue-Maß auf* $\mathbb{R}$*, dann konvergiert die Folge* $f_n := 1_{[n,n+1)}$ *gegen* $f = 0$*. Obwohl* $\mu(|f_n - f| = 1) = 1$ *für alle* $n \in \mathbb{N}$ *erscheint dies wünschenswert, da* f_n *punktweise gegen* f *konvergiert und diese Konvergenz bei einem endlichen Maß stärker ist.*

Bemerkung 3.3. *Es gilt* $\{f_n \to f\} = \cap_{k\in\mathbb{N}} \cup_{m\in\mathbb{N}} \cap_{n\geq m}\{|f_n - f| \leq 1/k\}$. *Daher liegt die Menge* $\{f_n \not\to f,\ n \to \infty\}$ *in* $\mathcal{F}$.

Die Implikation $(iii) \implies (i)$ folgt aus der Abschätzung

$$\int |f_n - f|^p \, d\mu \geq \varepsilon^p \mu(|f_n - f| > \varepsilon), \quad \forall \varepsilon > 0. \tag{3.1}$$

Die Implikation $(ii) \implies (i)$ zeigen wir in Satz 3.5. Andere Implikationen gelten nicht. Klassische Gegenbeispiele (mit Wahrscheinlichkeitsmaß):

$(ii) \not\Rightarrow (iii)$. Wähle $\Omega = (0,1)$ und $\mu = P$ das Lebesgue-Maß (Gleichverteilung) auf $(0,1)$. Setze $f = 0$ und $f_n(\omega) = n^{1/p} 1_{(0,1/n)}(\omega)$. Für jedes $\omega \in (0,1)$ gilt $f_n(\omega) \to f(\omega)$, $n \to \infty$, also (ii). Es gilt aber

$$\int |f_n - f|^p \, dP = (n^{1/p})^p P((0,1/n)) = n\frac{1}{n} = 1$$

und damit liegt keine $L^p(P)$-Konvergenz vor.

$(iii) \not\Rightarrow (ii)$ Stelle $n \in \mathbb{N}$ durch $n = 2^m + k$, $m \in \mathbb{N}_0$, $k = 0, \ldots, 2^m - 1$ dar und definiere

$$f_n(\omega) = 1_{(\frac{k}{2^m}, \frac{k+1}{2^m}]}(\omega) \quad \text{sowie wieder} \quad f = 0. \tag{3.2}$$

Es gilt

$$\int |f_n - f|^p \, dP = P(f_n = 1) = \frac{1}{2^m}.$$

Da mit $n \to \infty$ auch $m \to \infty$ folgt $L^p(P)$-Konvergenz. Für jedes $\omega \in (0,1)$ gibt es aber unendlich viele n mit $f_n(\omega) = 1$. Damit gibt es keine punktweise Konvergenz ((ii) ist nicht erfüllt).

Satz 3.4. *Wenn eine Folge* $(f_n)_{n\in\mathbb{N}}$ *im Maße* μ *sowohl gegen* f *als auch gegen* g *konvergiert, dann gilt* $f = g$ μ*-f.ü.*

Beweis. Sei $\varepsilon > 0$. Es gilt $\{|f-g| > \varepsilon\} \subset \{|f_n - f| > \varepsilon/2\} \cup \{|f_n - g| > \varepsilon/2\}$ und damit

$$\mu(\{|f-g| > \varepsilon\} \cap A) \leq \mu(\{|f_n - f| > \varepsilon/2\} \cap A) + \mu(\{|f_n - g| > \varepsilon/2\} \cap A).$$

Für alle Mengen $A \in \mathcal{F}$ mit $\mu(A) < \infty$ folgt mit der Konvergenz im Maße $\mu(\{|f-g| > \varepsilon\} \cap A) = 0$. Mit $\{|f-g| > 1/m\} \cap A \uparrow \{|f-g| \neq 0\} \cap A$ und der Stetigkeit von unten folgt $\mu(\{|f-g| \neq 0\} \cap A) = 0$. Wegen der σ-Endlichkeit und der σ-Additivität von μ folgt $\mu(\{|f-g| \neq 0\}) = 0$. □

Die schwächste Konvergenz, die wir hier betrachten liefert also bereits einen fast überall eindeutigen Limes.

Satz 3.5. *Eine Folge $(f_n)_{n\in\mathbb{N}}$, die μ-f.ü. gegen f konvergiert, konvergiert auch im Maße μ gegen f.*

Beweis. Es gilt $\{f_m \not\to f\} = \cup_{k\in\mathbb{N}} \cap_{n\in\mathbb{N}} \cup_{m\geq n} \{|f_m - f| > 1/k\}$. Da die abzählbare Vereinigung von Nullmengen eine Nullmenge ist, ist μ-f.ü. Konvergenz äquivalent dazu, dass für alle $k \in \mathbb{N}$

$$\mu(\cap_{n\in\mathbb{N}} \{|f_m - f| > 1/k \text{ für ein } m \geq n\}) = 0.$$

Dies impliziert, dass für alle A mit $\mu(A) < \infty$

$$\mu(\{|f_m - f| > 1/k \text{ für ein } m \geq n\} \cap A) \downarrow \mu(\cap_{\widetilde{n}\in\mathbb{N}} \{|f_m - f| > 1/k \text{ für ein } m \geq \widetilde{n}\} \cap A) = 0,$$

$n \to \infty$, $\forall k \in \mathbb{N}$ (man beachte, dass die Stetigkeit von oben eine Folge von Mengen mit endlichem Maß voraussetzt)

Aus $\{|f_n - f| > 1/k\} \subset \{|f_m - f| > 1/k \text{ für ein } m \geq n\}$ folgt

$$\mu(\{|f_n - f| > 1/k\} \cap A) \leq \mu(\{|f_m - f| > 1/k \text{ für ein } m \geq n\} \cap A) \downarrow 0,$$

$n \to \infty$, $\forall k \in \mathbb{N}$. Da dies für beliebige $k \in \mathbb{N}$ und beliebige Mengen A mit endlichem Maß gilt, konvergiert $(f_n)_{n\in\mathbb{N}}$ per definitionem im Maße μ gegen f. □

Lemma 3.6 (Erstes Borel-Cantelli-Lemma)**.** *Sei $(A_n)_{n\in\mathbb{N}} \subset \mathcal{F}$ mit $\sum_{n=1}^{\infty} \mu(A_n) < \infty$. Dann gilt $A := \left(\bigcap_{m\in\mathbb{N}} \bigcup_{n\geq m} A_n\right) \in \mathcal{F}$ und $\mu(A) = 0$.*

In der Wahrscheinlichkeitstheorie ist A das Ereignis, dass A_n für unendlich viele n eintritt. Es gilt

$$\omega \in \bigcap_{m\in\mathbb{N}} \bigcup_{n\geq m} A_n \Leftrightarrow \forall m\ \exists n \geq m\ \omega \in A_n \Leftrightarrow \omega \in A_n \text{ für unendlich viele } n \in \mathbb{N}.$$

A_n tritt genau dann unendlich oft ein, wenn es nach jedem m nochmal eintritt.

Beweis. Sei $m \in \mathbb{N}$. Es gilt $A \subset \bigcup\limits_{n\geq m} A_n$ und damit

$$\mu(A) \leq \mu(\bigcup_{n\geq m} A_n) \leq \sum_{n=m}^{\infty} \mu(A_n) \to 0 \quad \text{für } m \to \infty,$$

da $\sum_{n=1}^{\infty} \mu(A_n) < \infty$. Folglich kann die rechte Seite durch Wahl eines großen m beliebig klein gemacht werden. Da die linke Seite nicht von m abhängt, muss sie 0 sein. □

In Anwendungen des Lemmas auf nicht endliche Maße kann man alle A_n mit Mengen B von endlichem Maß schneiden. Macht man dies für alle $B = B_m$, wobei $\cup_{m\in\mathbb{N}} B_m = \Omega$ führt dies zu der schwächeren Bedingung $\sum_{n=1}^{\infty} \mu(A_n \cap B_m) < \infty$ für alle $m \in \mathbb{N}$, die gleichwohl $P\left(\bigcap\limits_{m\in\mathbb{N}} \bigcup\limits_{n\geq m} A_n\right) = 0$ liefert.

Die stochastische Konvergenz ist offenbar **metrisierbar**, d.h. es existiert eine Metrik $d : L^0(\Omega, \mathcal{F}, \mu) \times L^0(\Omega, \mathcal{F}, \mu) \to \mathbb{R}_+$, so dass folgende Äquivalenz gilt

$$f_n \to f,\ n \to \infty \Leftrightarrow d(f_n, f) \to 0,\ n \to \infty.$$

Wähle dazu

$$d(f,g) := \sum_{N=1}^{\infty} \frac{2^{-N}}{1+\mu(A_N)} \int 1_{A_N} \left(|f-g| \wedge 1\right) d\mu, \tag{3.3}$$

wobei $(A_n)_{n\in\mathbb{N}}$ eine Folge ist mit $\mu(A_n) \in (0,\infty)$ und $\cup_{n\in\mathbb{N}} A_n = \Omega$. Natürlich gibt es beliebig viele andere Metriken, die diese Aufgabe auch erfüllen würden.

d ist auf $L^0(\Omega, \mathcal{F}, \mu) \times L^0(\Omega, \mathcal{F}, \mu)$ und nicht auf $\mathcal{L}^0(\Omega, \mathcal{F}, \mu) \times \mathcal{L}^0(\Omega, \mathcal{F}, \mu)$ definiert. Der Übergang zu Äquivalenzklassen ist hier notwendig, da $d(f,g) = 0$ dazu äquivalent ist, dass $f = g$ μ-f.ü. Es ist klar, dass es für d egal ist, welche Repräsentanten aus einer Äquivalenzklasse eingesetzt werden.

Proposition 3.7. *Für einen metrischen Raum (X, d) gilt das folgende* **Teilfolgenkriterium**. *Für alle Folgen $(x_n)_{n\in\mathbb{N}} \subset X$ und $x \in X$ gilt*

$$d(x_n, x) \to 0,\ n \to \infty \Leftrightarrow \forall (n_k)_{k\in\mathbb{N}} \exists (n_{k_l})_{l\in\mathbb{N}} \text{ mit } d(x_{n_{k_l}}, x) \to 0,\ l \to \infty \tag{3.4}$$

d.h. wenn zu jeder Teilfolge eine konvergente Teilfolge mit dem gleichen Limes existiert, dann ist auch die Gesamtfolge konvergent.

Beweis. $\Rightarrow$ ist offensichtlich, da mit der Gesamtfolge auch jede Teilfolge gegen den Grenzwert konvergiert.

$\Leftarrow$ Widerspruchsannahme: $(x_n)_{n\in\mathbb{N}}$ konvergiere nicht gegen x. Man findet dann ein $\varepsilon > 0$ und eine Teilfolge $(n_k)_{k\in\mathbb{N}}$ s.d. $d(x_{n_k}, x) > \varepsilon$ für alle $k \in \mathbb{N}$. Diese Teilfolge kann dann auch keine Teilfolge besitzen, die gegen x konvergiert. □

Satz 3.8. *Seien $(f_n)_{n\in\mathbb{N}} \subset L^0(\Omega, \mathcal{F}, \mu)$ und $f \in L^0(\Omega, \mathcal{F}, \mu)$. Folgende Aussagen sind äquivalent*

(i) $(f_n)_{n\in\mathbb{N}}$ konvergiert im Maße μ gegen f.

(ii) $\forall (n_k)_{k\in\mathbb{N}} \exists (n_{k_l})_{l\in\mathbb{N}}$ so dass $(f_{n_{k_l}})_{l\in\mathbb{N}}$ μ-f.ü. gegen f konvergiert.

Beweis. (ii) $\Longrightarrow$ (i) Mit Satz 3.5 erfüllt $(f_n)_{n\in\mathbb{N}}$ die rechte Seite von (3.4), wenn (3.3) als Metrik gewählt wird, und damit ist auch die linke Seite erfüllt.

(i) $\Longrightarrow$ (ii) Es genügt zu zeigen, dass eine im Maße konvergente Folge eine Teilfolge besitzt, die f.ü. konvergiert. Wir wählen die Folge von Mengen in (3.3) aufsteigend, d.h. $A_1 \subset A_2 \subset \ldots$ Wenn $(f_n)_{n\in\mathbb{N}}$ im Maße konvergiert, dann findet man eine Teilfolge $(n_k)_{k\in\mathbb{N}}$ mit

$$\mu(\{|f_{n_k} - f| > 1/k\} \cap A_k) \le 2^{-k} \quad \forall k \in \mathbb{N}.$$

Aus der Monotonie der Folge $(A_N)_{N\in\mathbb{N}}$ folgt für jedes $N \in \mathbb{N}$

$$\sum_{k=1}^{\infty} \mu(\{|f_{n_k} - f| > 1/k\} \cap A_N) \le (N-1)\mu(A_{N-1}) + \sum_{k=N}^{\infty} 2^{-k} < \infty.$$

Mit Lemma 3.6 folgt, dass $\mu(\{|f_{n_k} - f| > 1/k\} \cap A_N$ für unendlich viele $k) = 0$ dies bedeutet, dass die Teilfolge auf A_N μ-f.ü. gegen f konvergiert. Da $\cup_{N\in\mathbb{N}} A_N = \Omega$ und die abzählbare Vereinigung von Nullmengen eine Nullmenge ist, folgt μ-f.ü. Konvergenz der Teilfolge $(f_{n_k})_{k\in\mathbb{N}}$ auf Ω. □

Bemerkung 3.9. *Die μ-fast überall Konvergenz ist offenbar nicht metrisierbar. Nehme dazu an, $\widetilde{d}$ sei eine Metrik, mit der man die μ-fast überall Konvergenz ableiten kann, d.h.*

$$f_n \to f, \mu\text{-f.s. } n \to \infty \Leftrightarrow \widetilde{d}(f_n, f) \to 0, \; n \to \infty.$$

Für eine Folge $(f_n)_{n\in\mathbb{N}}$, die im Maße μ gegen f konvergiert, existiert mit Satz 3.8 eine μ-f.ü. konvergente Teilfolge. Damit wäre die rechte Seite von (3.4) für $\widetilde{d}$, $(f_n)_{n\in\mathbb{N}}$, f erfüllt und $(f_n)_{n\in\mathbb{N}}$ müsste bereits μ-f.ü. gegen f konvergieren, was i.A. nicht der Fall zu sein braucht (siehe etwa das Gegenbeispiel (3.2)).

Metrisierbarkeit einer Konvergenz ist in folgender Situation hilfreich. Nehme an, eine Funktion g lässt sich durch eine Folge $(g_n)_{n\in\mathbb{N}}$ approximieren, d.h. $g_n \to g$ für $n \to \infty$. Jedes Folgenglied g_n lässt sich wiederum durch eine Folge $(g_{n,m})_{m\in\mathbb{N}}$ approximieren, d.h. $g_{n,m} \to g_n$ für $m \to \infty$ und alle $n \in \mathbb{N}$. Ist die Konvergenz metrisierbar, dann lässt sich g durch $g_{n,m}$ approximieren, in dem Sinne, dass man eine Folge $(k_n)_{n\in\mathbb{N}}$ findet, so dass $g_{n,k_n} \to g$ für $n \to \infty$ und man kann die Folge $(k_n)_{n\in\mathbb{N}}$ mit Hilfe der Metrik konstruieren (wie ?)

Bemerkung 3.10 (Cauchy-Folgen). *$(f_n)_{n\in\mathbb{N}}$ nennt man eine Cauchy-Folge bzgl. d, wenn für alle $\varepsilon > 0$ ein $n_\varepsilon \in \mathbb{N}$ existiert, s.d. $d(f_n, f_m) \leq \varepsilon$ für alle $n, m \geq n_\varepsilon$. Wir schreiben dafür einfach*

$$\lim_{n,m\to\infty} d(f_n, f_m) = 0. \tag{3.5}$$

Offenbar ist (3.5) mit d aus (3.3) dazu äquivalent, dass für alle Mengen $A \in \mathcal{F}$ mit $\mu(A) < \infty$ und alle $\varepsilon > 0$

$$\mu(\{|f_n - f_m| > \varepsilon\} \cap A) \to 0, \quad n, m \to \infty.$$

Eine solche Folge nennt man dann **Cauchy-Folge bzgl. der Konvergenz im Maße**. *Cauchy-Folgen werden gemeinhin nur für metrisierbare Konvergenzen betrachtet.*

Definition 3.11 (Gleichgradige Integrierbarkeit). *Eine Folge $(f_n)_{n\in\mathbb{N}}$ von messbaren Funktionen heißt gleichgradig integrierbar, wenn*

$$\inf_{0\leq g\in\mathcal{L}^1(\Omega,\mathcal{F},\mu)} \sup_{n\in\mathbb{N}} \int |f_n| 1_{\{|f_n|>g\}}\, d\mu = 0. \tag{3.6}$$

(3.6) bedeutet, dass es zu jedem $\varepsilon > 0$ eine Funktion $g \geq 0$ mit $\int g \, d\mu < \infty$ gibt, so dass $\int |f_n| 1_{\{|f_n|>g\}} \, d\mu \leq \varepsilon$ für alle $n \in \mathbb{N}$.

Die folgende Charakterisierung gleichgradiger Integrierbarkeit erweist sich als recht nützlich.

Satz 3.12. *Sei $h \in \mathcal{L}^1(\Omega, \mathcal{F}, \mu)$ mit $h > 0$ (d.h. $\{h > 0\} = \Omega$). Eine Folge $(f_n)_{n\in\mathbb{N}}$ ist genau dann gleichgradig integrierbar, wenn die beiden folgenden Aussagen gelten*

(i) $\sup_{n\in\mathbb{N}} \int |f_n| \, d\mu < \infty$.

(ii) Für jedes $\varepsilon > 0$ existiert ein $\delta > 0$, so dass für alle $A \in \mathcal{F}$ und $n \in \mathbb{N}$ folgende Implikation gilt

$$\int h 1_A \, d\mu \leq \delta \Rightarrow \int |f_n| 1_A \, d\mu \leq \varepsilon.$$

Ist $(f_n)_{n\in\mathbb{N}}$ gleichgradig integrierbar, dann wird das Infimum in (3.6) auch durch Vielfache der Funktion h approximiert, also

$$\lim_{\alpha\to\infty} \sup_{n\in\mathbb{N}} \int |f_n| 1_{\{|f_n|>\alpha h\}} \, d\mu = 0.$$

Bemerkung 3.13. *Eine Funktion $h \in \mathcal{L}^1(\Omega, \mathcal{F}, \mu)$ mit $h > 0$ existiert genau dann, wenn μ σ-endlich ist (wieso ?) Wenn μ endlich ist, also z.B. ein Wahrscheinlichkeitsmaß, dann kann in Satz 3.12 $h = 1$ gewählt werden (konstante Funktionen sind in diesem Fall integrierbar) und (3.6) ist äquivalent zu*

$$\sup_{n\in\mathbb{N}} \int |f_n| 1_{\{|f_n|>M\}} \, d\mu \to 0, \quad M \to \infty.$$

Bemerkung 3.14. *Aus gleichgradiger Integrierbarkeit folgt also, dass alle f_n integrierbar sind und die Erwartungswerte in $n \in \mathbb{N}$ beschränkt sind. Zusätzlich dürfen jedoch auch die Spitzen mit wachsendem n nicht explodieren, was durch (ii) zum Ausdruck kommt.*

Bemerkung 3.15. *Man mache sich klar, dass (i) nicht aus (ii) folgen muss. Wenn es ein $\omega_1 \in \Omega$ gibt mit $\{\omega_1\} \in \mathcal{F}$ und $\mu(\{\omega_1\}) > 0$, dann gilt für die Folge $f_n = n 1_{\{\omega_1\}}$ (ii) aber nicht (i). ,,Typischerweise" folgt jedoch (i) aus (ii). Die σ-Algebra $\mathcal{F}$ darf nur keine Atome haben (eine Menge wie oben, die positive Masse hat und nicht mehr teilbar ist). Dann lässt sich zu festem $\delta > 0$ der Gesamtraum in endlich viele disjunkte Teilmengen A mit $\int h 1_A \, d\mu \leq \delta$ zerlegen (man beachte, dass $\int h \, d\mu < \infty$).*

Beweis. Schritt 1: Wir zeigen zunächst die zweite Aussage des Satzes. Sei $(f_n)_{n\in\mathbb{N}}$ gleichgradig integrierbar, $\varepsilon > 0$ und $g \geq 0$ eine Funktion mit $\sup_{n\in\mathbb{N}} \int |f_n| 1_{\{|f_n|>g\}}\, d\mu \leq \varepsilon/2$. Für alle $\alpha > 0$ gilt

$$\begin{aligned} \int |f_n| 1_{\{|f_n|>\alpha h\}}\, d\mu &\leq \int |f_n| 1_{\{|f_n|>g\}}\, d\mu + \int |f_n| 1_{\{|f_n|\leq g\}\cap\{|f_n|>\alpha h\}}\, d\mu \\ &\leq \int |f_n| 1_{\{|f_n|>g\}}\, d\mu + \int g 1_{\{g>\alpha h\}}\, d\mu. \end{aligned} \tag{3.7}$$

Mit dem Satz von der monotonen Konvergenz gilt $\int g 1_{\{g\leq\alpha h\}}\, d\mu \uparrow \int g\, d\mu$ für $\alpha \uparrow \infty$ und wegen $\int g\, d\mu < \infty$ und Linearität des Integrals $\int g 1_{\{g>\alpha h\}}\, d\mu \to 0$. Also existiert ein α, so dass $\int g 1_{\{g>\alpha h\}}\, d\mu \leq \varepsilon/2$. Mit (3.7) folgt $\int |f_n| 1_{\{|f_n|>\alpha h\}}\, d\mu \leq \varepsilon$ für alle $n \in \mathbb{N}$ und damit die zweite Aussage des Satzes.

Schritt 2: Sei $(f_n)_{n\in\mathbb{N}}$ gleichgradig integrierbar. Wir wollen nun zeigen, dass (i) und (ii) gelten. Für alle $n \in \mathbb{N}$, $A \in \mathcal{F}$ und $\alpha > 0$ gilt

$$\begin{aligned} \int |f_n| 1_A\, d\mu &= \int |f_n| 1_{\{|f_n|>\alpha h\}\cap A}\, d\mu + \int |f_n| 1_{\{|f_n|\leq\alpha h\}\cap A}\, d\mu \\ &\leq \int |f_n| 1_{\{|f_n|>\alpha h\}}\, d\mu + \alpha \int h 1_A\, d\mu. \end{aligned} \tag{3.8}$$

Aus Schritt 1 folgt die Existenz eines α mit $\int |f_n| 1_{\{|f_n|>\alpha h\}}\, d\mu \leq 1$ für alle $n \in \mathbb{N}$. Eigenschaft (i) ergibt sich dann mit der Integrierbarkeit von h bei Betrachtung der Abschätzung (3.8) für $A = \Omega$.

Sei $\varepsilon > 0$. Wiederum aus Schritt 1 folgt die Existenz von α mit $\int |f_n| 1_{\{|f_n|>\alpha h\}}\, d\mu \leq \varepsilon/2$ für alle $n \in \mathbb{N}$. Mit (3.8) folgt Eigenschaft (ii) für $\delta := \frac{\varepsilon}{2\alpha}$.

Schritt 3: Nehme nun an, dass (i) und (ii) gelten. Für alle $\alpha \geq 0$ und $n \in \mathbb{N}$ gilt

$$\int h 1_{\{|f_n|>\alpha h\}}\, d\mu \leq \frac{1}{\alpha} \int |f_n|\, d\mu \tag{3.9}$$

(wobei $1/0 := \infty$). Sei $\varepsilon > 0$ und ein entsprechendes $\delta > 0$ so gewählt, dass die Implikation in (ii) gilt. Wähle nun

$$\alpha := \frac{\sup_{n\in\mathbb{N}} \int |f_n|\, d\mu}{\delta}$$

(was wegen Bedingung (i) endlich ist). Mit (3.9) folgt für obiges α, dass $\int h 1_{\{|f_n|>\alpha h\}}\, d\mu \leq \delta$ für alle $n \in \mathbb{N}$. (ii) angewandt auf die Mengen $A_n = \{|f_n| > \alpha h\}$ liefert $\int |f_n| 1_{\{|f_n|>\alpha h\}}\, d\mu \leq \varepsilon$ für alle $n \in \mathbb{N}$ und damit gleichgradige Integrierbarkeit. □

Korollar 3.16. *Wenn die Folgen $(f_n)_{n\in\mathbb{N}}$ und $(g_n)_{n\in\mathbb{N}}$ gleichgradig integrierbar sind, dann ist auch die Folge $(\alpha f_n + \beta g_n)_{n\in\mathbb{N}}$ mit $\alpha, \beta \in \mathbb{R}$ gleichgradig integrierbar.*

Beweis. Folgt aus der Charakterisierung gleichgradiger Integrierbarkeit durch (i) und (ii) in Satz 3.12. □

Wichtig für Anwendungen ist der folgende Satz.

Satz 3.17. *Sei $p \in [1, \infty)$. Eine Folge $(f_n)_{n\in\mathbb{N}} \subset \mathcal{L}^0(\Omega, \mathcal{F}, \mu)$, die im Maße μ gegen ein $f \in \mathcal{L}^0(\Omega, \mathcal{F}, \mu)$ konvergiert und bei der die Folge $(|f_n|^p)_{n\in\mathbb{N}}$ gleichgradig integrierbar ist, die konvergiert auch in $L^p(\mu)$ gegen f (und es gilt $\lim_{n\to\infty} \int |f_n|^p \, d\mu = \int |f|^p \, d\mu < \infty$).*

Die Umkehrung des Satzes gilt auch. Mit der Umkehrung wir dies dann in Satz 3.21 formuliert und bewiesen.

Beweis. Schritt 1: Unter Konvergenz im Maße und der gleichgradigen Integrierbarkeit der Folge $(|f_n|^p)_{n\in\mathbb{N}}$ folgere man mit Hilfe einer *Teilfolge* $(n_k)_{k\in\mathbb{N}}$, auf der fast überall Konvergenz gilt (Satz 3.8), die Integrierbarkeit von $|f|^p$. Es gilt nämlich

$$\int |f|^p \, d\mu = \int \liminf_{k\to\infty} |f_{n_k}|^p \, d\mu \overset{\text{Fatou}}{\leq} \liminf_{k\to\infty} \int |f_{n_k}|^p \, d\mu < \infty \tag{3.10}$$

($< \infty$ folgt aus Bedingung (i) in Satz 3.12). Wegen (3.10), der Abschätzung

$$|f_n - f|^p \leq (|f_n| + |f|)^p \leq (2 \max\{|f_n|, |f|\})^p \leq 2^p |f_n|^p + 2^p |f|^p$$

und Korollar 3.16 zieht die gleichgradige Integrierbarkeit von $(|f_n|^p)_{n\in\mathbb{N}}$ die gleichgradige Integrierbarkeit von $(|f_n - f|^p)_{n\in\mathbb{N}}$ nach sich.

Schritt 2: Da die $L^p(\mu)$-Konvergenz metrisierbar ist, reicht es aus, $L^p(\mu)$-Konvergenz gegen f auf einer Teilfolge zu zeigen (Wieso ?)

Wir argumentieren wieder mit einer Teilfolge $(n_k)_{k\in\mathbb{N}}$, auf der $f_{n_k} \to f$, μ-f.ü, $k \to \infty$. Sei $\varepsilon > 0$. Mit Schritt 1 gibt es ein $g \in \mathcal{L}^1(\Omega, \mathcal{F}, \mu)$ mit $g \geq 0$, so dass

$$\int |f_{n_k} - f|^p 1_{\{|f_{n_k} - f|^p > g\}} \, d\mu \leq \varepsilon, \quad \forall k \in \mathbb{N}. \tag{3.11}$$

Definiere

$$g_k := |f_{n_k} - f|^p 1_{\{|f_{n_k} - f|^p \leq g\}}.$$

Es gilt $g_k \to 0$ fast überall für $k \to \infty$ und $g - g_k \geq 0$ für alle $k \in \mathbb{N}$. Mit dem Lemma von Fatou, angewandt auf die nichtnegative Folge $(g - g_k)_{k\in\mathbb{N}}$, folgt

$$\begin{aligned}
\limsup_{k\to\infty} \int g_k \, d\mu &= \int g \, d\mu - \liminf_{k\to\infty} \int (g - g_k) \, d\mu \\
&\leq \int g \, d\mu - \int \left(\liminf_{k\to\infty} (g - g_k) \right) d\mu \\
&= 0.
\end{aligned}$$

Mit $|f_{n_k} - f|^p = |f_{n_k} - f|^p 1_{\{|f_{n_k} - f|^p > g\}} + g_k$ folgt aus (3.11)

$$\limsup_{k\to\infty} \int |f_{n_k} - f|^p \, d\mu \leq \limsup_{k\to\infty} \int |f_{n_k} - f|^p 1_{\{|f_{n_k} - f|^p > g\}} \, d\mu + \limsup_{k\to\infty} \int g_k \, d\mu \leq \varepsilon.$$

Da zwar g nicht aber die Teilfolge $(n_k)_{k\in\mathbb{N}}$ von ε abhängt und letzteres beliebig gewählt ist, folgt $\lim_{k\to\infty} \int |f_{n_k} - f|^p \, d\mu = 0$.

Bleibt die Aussage in Klammern zu zeigen. Aus der Minkowski-Ungleichung (Satz 2.19) folgt

$$\left(\int |f_n|^p \, d\mu \right)^{1/p} \leq \left(\int |f|^p \, d\mu \right)^{1/p} + \left(\int |f_n - f|^p \, d\mu \right)^{1/p}$$

und

$$\left(\int |f|^p \, d\mu \right)^{1/p} \leq \left(\int |f_n|^p \, d\mu \right)^{1/p} + \left(\int |f_n - f|^p \, d\mu \right)^{1/p}.$$

Wegen $f_n \in \mathcal{L}^p(\Omega, \mathcal{F}, \mu)$ folgt

$$\left| \left(\int |f_n|^p \, d\mu \right)^{1/p} - \left(\int |f|^p \, d\mu \right)^{1/p} \right| \leq \left(\int |f_n - f|^p \, d\mu \right)^{1/p} \to 0, \quad n \to \infty.$$

□

Korollar 3.18 (Satz von der majorisierten Konvergenz)**.** *Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge, die im Maße μ gegen f konvergiert. Wenn es ein $g \in \mathcal{L}^1(\Omega, \mathcal{F}, \mu)$ gibt mit $|f_n|^p \leq g$ für alle $n \in \mathbb{N}$, konvergiert $(f_n)_{n\in\mathbb{N}}$ auch in $L^p(\Omega, \mathcal{F}, \mu)$ gegen f.*

Beweis. Setzt man die Majorante g in (3.6) ein, sieht man sofort, dass $(|f_n|^p)_{n\in\mathbb{N}}$ gleichgradig integrierbar ist und das Korollar folgt aus Satz 3.17. □

Eine wichtige Eigenschaft von Räumen ist die Vollständigkeit.

Satz 3.19. *Der Raum $L^0(\Omega, \mathcal{F}, \mu)$ ausgestattet mit der Metrik d aus (3.3) ist vollständig, d.h. jede Cauchy-Folge konvergiert (also: aus $\lim_{n,m\to\infty} d(f_n, f_m) = 0$ folgt die Existenz von $f \in L^0(\Omega, \mathcal{F}, \mu)$ mit $\lim_{n\to\infty} d(f_n, f) = 0$).*

Beweis. Der Beweis verläuft ähnlich wie der von (i) $\Longrightarrow$ (ii) in Satz 3.8.

Wir wählen die Folge von Mengen A_k in (3.3) mit $\mu(A_k) < \infty$ aufsteigend, d.h. $A_1 \subset A_2 \subset \dots$. Wenn $\lim_{n,m\to\infty} d(f_n, f_m) = 0$, dann findet man eine Teilfolge $(n_k)_{k\in\mathbb{N}}$ mit

$$\mu(\{|f_{n_{k+1}} - f_{n_k}| > 2^{-k}\} \cap A_k) \leq 2^{-k} \quad \forall k \in \mathbb{N}$$

(Definiere dazu, n_k rekursiv als die kleinste natürliche Zahl $> n_{k-1}$, so dass für alle $n, m \geq n_k$ gilt $\mu(\{|f_n - f_m| > 2^{-k}\} \cap A_k) \leq 2^{-k}$)

Aus der Monotonie der Folge $(A_N)_{N\in\mathbb{N}}$ folgt für jedes $N \in \mathbb{N}$

$$\sum_{k=1}^{\infty} \mu(\{|f_{n_{k+1}} - f_{n_k}| > 2^{-k}\} \cap A_N) \leq (N-1)\mu(A_{N-1}) + \sum_{k=N}^{\infty} 2^{-k} < \infty.$$

Mit Lemma 3.6 folgt, dass

$$\mu(\{|f_{n_{k+1}} - f_{n_k}| > 2^{-k}\} \cap A_N \text{ für unendlich viele } k) = 0, \quad \forall N \in \mathbb{N}$$

und damit wegen $\cup_{N\in\mathbb{N}} A_N = \Omega$

$$\mu(\{|f_{n_{k+1}} - f_{n_k}| > 2^{-k} \text{ für unendlich viele } k\}) = 0. \tag{3.12}$$

Definiere nun

$$f(\omega) := \begin{cases} \lim_{k\to\infty} f_{n_k}(\omega) & : \quad \text{für } \omega \notin \cap_{m\in\mathbb{N}} \cup_{k\geq m} \{|f_{n_{k+1}} - f_{n_k}| > 2^{-k}\} \\ 0 & : \quad \text{sonst}\,. \end{cases}$$

Diese Definition ist möglich, da für $\omega \notin \cap_{m\in\mathbb{N}} \cup_{k\geq m} \{|f_{n_{k+1}} - f_{n_k}| > 2^{-k}\}$ die Folge reeller Zahlen $(f_{n_k}(\omega))_{n\in\mathbb{N}}$ eine Cauchy-Folge ist und damit wegen der Vollständkeit von $\mathbb{R}$ einen Grenzwert $\lim_{k\to\infty} f_{n_k}(\omega)$ besitzt. Für diese ω konvergiert $(f_{n_k}(\omega))_{k\in\mathbb{N}}$ also punktweise gegen $f(\omega)$. Aus (3.12) folgt, dass in obiger Definition von f die Definition mit dem

Limes außerhalb einer μ-Nullmenge greift, die Teilfolge $(f_{n_k})_{k\in\mathbb{N}}$ konvergiert damit μ-f.ü. gegen f. Aus Proposition 1.19 folgt die Messbarkeit von f. Sei nun $A \in \mathcal{F}$ mit $\mu(A) < \infty$. Aus Satz 3.5 folgt

$$\mu(\{|f_{n_k} - f| > \varepsilon\} \cap A) \to 0, \quad k \to \infty$$

und zusammen mit $\mu(\{|f_n - f_m| > \varepsilon\} \cap A) \to 0$ für $n, m \to \infty$ und der Dreiecksungleichung folgt

$$\mu(\{|f_n - f| > \varepsilon\} \cap A) \to 0, \quad n \to \infty,$$

d.h. nicht nur die Teilfolge sondern auch die Gesamtfolge konvergiert im Maße gegen f. □

Der Beweis beruht natürlich auf der Vollständigkeit von $\mathbb{R}$. Es ginge genauso mit Funktionen, die in andere vollständige Räume abbilden. Man beachte jedoch, dass die Vollständigkeit des Maßraums („jede Teilmenge einer Nullmenge ist eine Nullmenge") hier nicht gebraucht wird, da das konstruierte f in jedem Fall messbar ist.

Weiß man von einer Funktion dagegen nur, dass sie außerhalb einer Nullmenge der punktweise Limes messbarer Funktionen ist, dann braucht man die Vollständigkeit des Maßraums um auf die Messbarkeit der Funktion zu schließen.

Satz 3.20 (Fischer-Riesz)**.** *Sei $p \in [1, \infty)$ und sei $(f_n)_{n\in\mathbb{N}} \subset L^p(\Omega, \mathcal{F}, \mu)$ eine Cauchy-Folge bzgl. der Norm $||\cdot||_p$ (vgl. (2.7)), d.h.*

$$\lim_{n,m\to\infty} ||f_n - f_m||_p = 0. \tag{3.13}$$

Dann gibt es ein $f \in L^p(\Omega, \mathcal{F}, \mu)$ mit

$$\lim_{n\to\infty} ||f_n - f||_p = 0.$$

Beweis. Schritt 1: Wir zeigen zunächst, dass aus (3.13) die gleichgradige Integrierbarkeit von $(|f_n|^p)_{n\in\mathbb{N}}$ folgt. Dazu verifizieren wir (i) und (ii) in Satz 3.12. Aus (3.13) folgt

$$\forall \widetilde{\varepsilon} > 0 \exists n_{\widetilde{\varepsilon}} \in \mathbb{N} \quad \int |f_n - f_{n_{\widetilde{\varepsilon}}}|^p \, d\mu \leq \widetilde{\varepsilon} \quad \forall n \geq n_{\widetilde{\varepsilon}}. \tag{3.14}$$

Zunächst wenden wir (3.14) für $\widetilde{\varepsilon} = 1$ an und schließen zusammen mit

$$\int |f_n|^p \, d\mu \leq 2^p \int (|f_1|^p + \ldots + |f_{n_1}|^p) \, d\mu + 2^p \int |f_n - f_{n_1}|^p \, d\mu, \quad \forall n \in \mathbb{N},$$

auf

$$\sup_{n\in\mathbb{N}} \int |f_n|^p \, d\mu \leq 2^p \int (|f_1|^p + \ldots + |f_{n_1}|^p) \, d\mu + 2^p < \infty, \tag{3.15}$$

und damit (i).

Nun wollen wir (ii) zeigen. Sei dazu $\varepsilon > 0$. Wende (3.14) auf

$$\widetilde{\varepsilon} := \frac{\varepsilon}{2^{p+1}}$$

an. Mit dem entsprechenden $n_{\widetilde{\varepsilon}}$ schätzen wir nun ab. Für alle $n \in \mathbb{N}$ und $A \in \mathcal{F}$ gilt die Abschätzung

$$\begin{aligned} \int 1_A |f_n|^p \, d\mu &\leq 2^p \int 1_A |f_{n_{\widetilde{\varepsilon}}}|^p \, d\mu + 2^p \int |f_n - f_{n_{\widetilde{\varepsilon}}}|^p \, d\mu \\ &\leq 2^p \int 1_A (|f_1|^p + \ldots + |f_{n_{\widetilde{\varepsilon}}}|^p) \, d\mu + 2^p \int |f_n - f_{n_{\widetilde{\varepsilon}}}|^p \, d\mu. \end{aligned} \tag{3.16}$$

Zudem gilt mit (3.14)

$$2^p \int |f_n - f_{n_{\widetilde{\varepsilon}}}|^p \, d\mu \leq \frac{\varepsilon}{2}, \quad \forall n \geq n_{\widetilde{\varepsilon}}. \tag{3.17}$$

$2^p \, (|f_1|^p + \ldots + |f_{n_{\widetilde{\varepsilon}}}|^p)$ ist eine integrierbare Zufallsvariable. Da eine einzelne integrierbare Zufallsvariable (als konstante Folge betrachtet) gleichgradig integrierbar ist, folgt mit Satz 3.12 die Existenz eines $\delta > 0$, so dass für alle $A \in \mathcal{F}$ die Implikation

$$\int h 1_A \, d\mu \leq \delta \Rightarrow 2^p \int 1_A (|f_1|^p + \ldots + |f_{n_{\widetilde{\varepsilon}}}|^p) \, d\mu \leq \varepsilon/2. \tag{3.18}$$

gilt. Aus (3.16), (3.17) und (3.18) mit dem dortigen δ folgt die Implikation

$$\int h 1_A \, d\mu \leq \delta \Rightarrow \sup_{n\in\mathbb{N}} \int 1_A |f_n|^p \, d\mu \leq \varepsilon.$$

Da $\varepsilon > 0$ beliebig war, folgt (ii) und damit insgesamt die gleichgradige Integrierbarkeit von $(|f_n|^p)_{n\in\mathbb{N}}$

Schritt 2: Aus (3.13) folgt mit der Abschätzung

$$\mu(|f_n - f_m| > \varepsilon) = \mu(|f_n - f_m|^p > \varepsilon^p) \leq \frac{1}{\varepsilon^p} \int |f_n - f_m|^p, d\mu \quad \text{(Markov-Ungleichung)},$$

dass $(f_n)_{n\in\mathbb{N}}$ auch eine Cauchy Folge im Maße μ ist und wegen Satz 3.19 im Maße μ gegen ein $f \in L^0(\Omega, \mathcal{F}, \mu)$ konvergiert. Mit Schritt 1 und Satz 3.17 folgt somit die $L^p(\mu)$-Konvergenz gegen dieses f und damit die Behauptung (man beachte, dass im dortigen Beweis bereits gezeigt wurde, dass die gleichgradige Integrierbarkeit von $(|f_n|^p)_{n\in\mathbb{N}}$ die Integrierbarkeit von $|f|^p$ impliziert). □

Satz 3.21 (Konvergenzsatz von Vitali). *Sei $(f_n)_{n\in\mathbb{N}} \subset \mathcal{L}^p(\Omega, \mathcal{F}, \mu)$ und $f \in \mathcal{L}^0(\Omega, \mathcal{F}, \mu)$. Folgende Aussagen sind äquivalent*

(i) $(f_n)_{n\in\mathbb{N}}$ konvergiert gegen f in $L^p(\mu)$.

(ii) $(f_n)_{n\in\mathbb{N}}$ konvergiert gegen f im Maße μ und $(|f_n|^p)_{n\in\mathbb{N}}$ ist gleichgradig integrierbar.

(Zudem gilt $f \in \mathcal{L}^p(\Omega, \mathcal{F}, \mu)$, wenn (i) bzw. (ii) erfüllt ist)

Beweis. (ii) $\implies$ (i) ist Satz 3.17.

Ad (i) $\implies$ (ii) Die Konvergenz im Maße folgt mit (3.1). Zudem gilt wegen der Dreiecksungleichung (3.13). Mit Schritt 1 des Beweises von Satz 3.20 folgt die gleichgradige Integrierbarkeit von $(|f_n|^p)_{n\in\mathbb{N}}$.

□

4 Mehrfachintegrale und Produktmaße

4.1 Dynkinscher π-λ-Satz

Wir werden zunächst einen Satz beweisen, mit dem in der Maß- und Wahrscheinlichkeitstheorie sehr oft argumentiert wird.

Definition 4.1. *Ein Mengensystem $\mathcal{A} \subset 2^\Omega$ wird* **Dynkinsystem** *genannt, wenn*

(1) $\Omega \in \mathcal{A}$

(2) $A \in \mathcal{A} \implies A^c \in \mathcal{A}$

(3) für jede Folge $(A_n)_{n\in\mathbb{N}}$ paarweise disjunkter Mengen aus $\mathcal{A}$ gilt $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$

Bemerkung 4.2. *Der Unterschied zu einer σ-Algebra besteht darin, dass (3) nur für eine disjunkte Folge gelten muss. Daher kann man mit $A \cap B = (A^c \cup B^c)^c$ nicht mehr folgern, dass der Schnitt zweier Mengen aus dem Mengensystem wieder drin ist.*

Bemerkung 4.3. *In der Definition könnte man (2) durch*

(2') $A, B \in \mathcal{A}$ mit $A \subset B \implies B \setminus A \in \mathcal{A}$

ersetzen. Es gilt nämlich $B \setminus A = (A \cup B^c)^c$, und die Mengen A, B^c sind disjunkt.

Satz 4.4 (Erzeugung von Dynkin-Systemen)**.** *Sei $\mathcal{E} \subset 2^\Omega$. Es existiert ein kleinstes Dynkin-System $\delta(\mathcal{E})$, das $\mathcal{E}$ umfasst, nämlich*

$$\delta(\mathcal{E}) := \bigcap_{\mathcal{A} \subset 2^\Omega,\ \mathcal{A} \text{ ist Dynkin-System mit } \mathcal{E} \subset \mathcal{A}} \mathcal{A} \tag{4.1}$$

$$:= \{A \subset \Omega \mid A \in \mathcal{A}\ \forall \text{Dynkinsysteme } \mathcal{A} \text{ mit } \mathcal{E} \subset \mathcal{A}\}. \tag{4.2}$$

"Kleinstes" bedeutet, dass jedes andere Dynkin-System, das $\mathcal{E}$ umfasst, auch $\delta(\mathcal{E})$ umfasst.

Beweis. Die Aussage ist völlig analog zur Erzeugung von σ-Algebren in Satz 1.5. Analog zum dortigen Beweis folgt, dass beliegige Schnitte von Dynkin-Systemen wiederum Dynkin-Systeme sind, was die Behauptung ergibt. □

Satz 4.5 (Dynkinscher π-λ-Satz*)**.** *Sei $\mathcal{E} \subset 2^\Omega$. Wenn $\mathcal{E}$* **durchschnittsstabil** *ist, d.h. $A_1, A_2 \in \mathcal{E} \implies A_1 \cap A_2 \in \mathcal{E}$, dann gilt $\delta(\mathcal{E}) = \sigma(\mathcal{E})$.*

Bemerkung 4.6. *Der Satz ist für Anwendungen sehr nützlich. Häufig kann man von einem Mengensystem zwar zeigen, dass es ein Dynkinsystem ist, nicht aber, dass es auch eine σ-Algebra ist. Dies liegt daran, dass man bei einem Dynkinsystem nur abzählbare Vereinigungen von* **disjunkten** *Mengen aus dem Mengensystem betrachten muss. Es müssen also keine Überlappungen berücksichtigt werden. Eine Anwendung von Satz 4.5 wird auch Dynkin-Argument genannt.*

*Ein durchschnittsstabiles Mengensystem wird manchmal auch π-System genannt und ein Dynkinsystem λ-System.

Beweis von Satz 4.5. Da jede σ-Algebra ein Dynkin-System ist, gilt $\delta(\mathcal{E}) \subset \sigma(\mathcal{E})$. Es bleibt zu zeigen, dass $\delta(\mathcal{E})$ eine σ-Algebra ist. Da $\delta(\mathcal{E})$ ein Dynkin-System ist, also insbesondere (3) mit disjunkten Mengen erfüllt, muss nur noch gezeigt werden, dass $\delta(\mathcal{E})$ durchschnittsstabil ist. In diesem Fall kann nämlich eine abzählbare Vereinigung beliebiger Mengen aus $\delta(\mathcal{E})$ auf eine abzählbare Vereinigung disjunkter Mengen aus $\delta(\mathcal{E})$ zurückgeführt werden (setze dazu $\widetilde{A}_1 := A_1$ und $\widetilde{A}_n := A_n \cap A_1^c \cap \ldots \cap A_{n-1}^c \in \mathcal{A}$ für $n \geq 2$. Es gilt $\bigcup_{n\in\mathbb{N}} A_n = \bigcup_{n\in\mathbb{N}} \widetilde{A}_n \in \mathcal{A}$ und die $\widetilde{A}_n$ sind disjunkt). Um Durchschnittsstabilität von $\delta(\mathcal{E})$ zu zeigen, definiere man für jedes $D \in \delta(\mathcal{E})$ die Menge

$$\mathcal{D}_D := \{Q \in 2^\Omega \mid Q \cap D \in \delta(\mathcal{E})\}.$$

(i) Man rechnet leicht nach, dass auch $\mathcal{D}_D$ für jedes $D \in \delta(\mathcal{E})$ ein Dynkinsystem ist.

(ii) Wegen der Durchschnittsstabilität von $\mathcal{E}$ gilt für jedes $E \in \mathcal{E}$, dass $\mathcal{E} \subset \mathcal{D}_E$.

(iii) Aus (i) und der Minimalität des erzeugten Dynkinsystems folgt, dass $\delta(\mathcal{E}) \subset \mathcal{D}_E$ für alle $E \in \mathcal{E}$.

(iv) Mit (iii) gilt für jedes $D \in \delta(\mathcal{E})$ und jedes $E \in \mathcal{E}$, dass $E \cap D \in \delta(\mathcal{E})$. Dies bedeutet aber, dass $\mathcal{E} \subset \mathcal{D}_D$ für alle $D \in \delta(\mathcal{E})$. Wegen (i) zieht dies $\delta(\mathcal{E}) \subset \mathcal{D}_D$ für alle $D \in \delta(\mathcal{E})$ nach sich. D.h. für alle $D, D' \in \delta(\mathcal{E})$ gilt $D \cap D' \in \delta(\mathcal{E})$.

□

Wichtige Folge von Satz 4.5 ist der folgende

Satz 4.7. *Seien μ und ν Maße auf $(\Omega, \mathcal{F})$ und $\mathcal{E}$ ein durchschnittstabiler Erzeuger von $\mathcal{F}$ (d.h. $\sigma(\mathcal{E}) = \mathcal{F}$ und es gilt die Implikation $E_1, E_2 \in \mathcal{E} \implies E_1 \cap E_2 \in \mathcal{E}$). Wenn*

(i) $\mu(E) = \nu(E)$ für alle $E \in \mathcal{E}$ und

(ii) es gibt eine Folge $(E_n)_{n\in\mathbb{N}} \subset \mathcal{E}$ mit $E_n \uparrow \Omega$ und $\mu(E_n) < \infty$ für alle $n \in \mathbb{N}$,

dann gilt $\mu = \nu$ (d.h. $\mu(A) = \nu(A)$ für alle $A \in \mathcal{F}$). Sind μ und ν Wahrscheinlichkeitsmaße, dann kann auf Bedingung (ii) verzichtet werden.

(Man beachte, dass (ii) zwar die σ-Endlichkeit von μ impliziert, jedoch stärker ist, da für die σ-Endlichkeit die die Gesamtmenge approximierenden Mengen aus $\sigma(\mathcal{E})$ und i.A. nicht aus $\mathcal{E}$ sind)

Beweis. Definiere für ein festes $n \in \mathbb{N}$ das Mengensystem

$$\mathcal{A}_n := \{A \in \mathcal{F} \mid \mu(A \cap E_n) = \nu(A \cap E_n)\}$$

Offenbar gilt $\mathcal{E} \subset \mathcal{A}_n$. Zudem ist $\mathcal{A}_n$ ein Dynkin-System. Wegen der σ-Additivität von μ und ν ist nämlich die abzählbare Vereinigung **disjunkter** Mengen aus $\mathcal{A}_n$ wiederum in $\mathcal{A}_n$. Zudem stellt der Schnitt mit der Menge E_n von endlichem Maß sicher, dass die Gesamtmenge und die entsprechenden Komplemente drin sind, da dann die Differenzen $\mu(E_n) - \mu(A \cap E_n)$ und $\nu(E_n) - \nu(A \cap E_n)$ erklärt sind und mit $\mu(A^c \cap E_n)$ bzw. $\nu(A^c \cap E_n)$ übereinstimmen. Mit Satz 4.5 folgt $\mathcal{A}_n \supset \delta(\mathcal{E}) = \sigma(\mathcal{E}) = \mathcal{F}$, d.h. $\mu(A \cap E_n) = \nu(A \cap E_n)$ für alle $A \in \mathcal{F}$. Aus der Stetigkeit von unten von μ und ν folgt $\mu(A) = \lim_{n\to\infty} \mu(A \cap E_n) = \lim_{n\to\infty} \nu(A \cap E_n) = \nu(A)$ für alle $A \in \mathcal{F}$.

Wenn μ und ν Wahrscheinlichkeitsmaße sind, gilt $\mu(\Omega) = \nu(\Omega) = 1$. Man betrachte nun den Erzeuger $\widetilde{\mathcal{E}} := \mathcal{E} \cup \{\Omega\}$. Mit $\mathcal{E}$ ist auch $\widetilde{\mathcal{E}}$ durchschnittsstabil und erfüllt (i). Zudem ist für $\widetilde{\mathcal{E}}$ Bedingung (ii) mit $E_n = \Omega$ für alle $n \in \mathbb{N}$ automatisch erfüllt.

□

Beispiel 4.8. *Sei Y eine reellwertige Zufallsvariable. Offenbar ist das Bildmaß $P_Y(B) := P(Y \in B),\ B \in \mathcal{B}(\mathbb{R})$ (vgl. Beispiel 1.21) durch die Werte für die Mengen $B = (-\infty, y]$, wobei y alle reellen Zahlen durchläuft, bereits eindeutig bestimmt. $\{(-\infty, y] \mid y \in \mathbb{R}\}$ ist nämlich durchschnittsstabil und erzeugt $\mathcal{B}(\mathbb{R})$.*

Die Funktion $F_Y(y) := P_Y((-\infty, y]) = P(Y \leq y),\ y \in \mathbb{R}$ nennt man **Verteilungsfunktion** *der Zufallsvariablen Y.*

Folgendes Beispiel zeigt, dass Bedingung (ii) nicht weggelassen werden kann, wenn μ und ν statt Wahrscheinlichkeitsmaße (mit gleicher Gesamtmasse) nur endliche Maße sein müssen.

Beispiel 4.9. $\Omega = \{1, 2\}$, $\mathcal{E} = \{\{1\}\}$. *Definiere* $\mu(\{1\}) = \nu(\{1\}) = \mu(\{2\}) = 1$ *und* $\nu(\{2\}) = 2$.

Das Beispiel basiert also darauf, dass die Gesamtmasse durch die Erzeugermengen nicht eindeutig festgelegt sein muss. Dieses Problem entfällt, wenn man nur Wahrscheinlichkeitsmaße betrachtet. Für Wahrscheinlichkeitsmaße zeigen wir nun, dass auf die Durchschnittsstabilität nicht verzichtet werden kann

Beispiel 4.10. $\Omega = \{1, 2, 3, 4\}$, $\mathcal{E} = \{\{1, 2\}, \{2, 3\}\}$. *Es gibt mehrere Wahrscheinlichkeitsmaße, die auf* $\mathcal{E}$ *mit der Gleichverteilung übereinstimmen (welche ?).*

4.2 Produktmaß und Satz von Fubini

Seien $(\Omega_1, \mathcal{F}_1, \mu_1)$ und $(\Omega_2, \mathcal{F}_2, \mu_2)$ Maßräume. Auf dem kartesischen Produkt $\Omega_1 \times \Omega_2 := \{(\omega_1, \omega_2) \mid \omega_1 \in \Omega_1,\ \omega_2 \in \Omega_2\}$ wollen wir eine σ-Algebra und ein Maß definieren, so dass das Maß jeder Rechteckmenge $A_1 \times A_2 = \{(\omega_1, \omega_2) \mid \omega_1 \in A_1,\ \omega_2 \in A_2\}$ die Zahl $\mu_1(A_1)\mu_2(A_2)$ zuordnet. Die Menge aller Rechteckmengen bezeichnen wir mit

$$\mathcal{F}_1 \times \mathcal{F}_2 := \{A_1 \times A_2 \mid A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2\}$$

und die von ihr auf der Grundmenge $\Omega_1 \times \Omega_2$ erzeugte σ-Algebra mit

$$\mathcal{F}_1 \otimes \mathcal{F}_2 := \sigma(\mathcal{F}_1 \times \mathcal{F}_2).$$

$\mathcal{F}_1 \times \mathcal{F}_2$ ist (wenn $\mathcal{F}_1 \neq \{\emptyset, \Omega_1\}$ und $\mathcal{F}_2 \neq \{\emptyset, \Omega_2\}$) keine σ-Algebra und damit $\mathcal{F}_1 \times \mathcal{F}_2 \subsetneq \mathcal{F}_1 \otimes \mathcal{F}_2$.

Da bei σ-Algebren auch die Schnitte wieder drin sind, ist $\mathcal{F}_1 \otimes \mathcal{F}_2$ auch die *kleinste* σ-Algebra auf dem Produktraum $\Omega_1 \times \Omega_2$, die die sog. *Zylindermengen* $A_1 \times \Omega_2$, $A_1 \in \mathcal{F}_1$ und $\Omega_1 \times A_2$, $A_2 \in \mathcal{F}_2$ enthält.

Diesen Sachverhalt kann man auch mit den **Projektionsabbildungen**

$$p_i : \Omega_1 \times \Omega_2 \to \Omega_i,\ (\omega_1, \omega_2) \mapsto \omega_i, \quad i = 1, 2,$$

ausdrücken. Es gilt

$$\mathcal{F}_1 \otimes \mathcal{F}_2 = \sigma(p_1, p_2) := \sigma(p_1^{-1}(A_1), p_2^{-1}(A_2), A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2). \tag{4.3}$$

$\mathcal{F}_1 \otimes \mathcal{F}_2$ ist also die kleinste σ-Algebra auf dem Produktraum $\Omega_1 \times \Omega_2$, so dass die Projektionsabbildungen p_i $(\mathcal{F}_1 \otimes \mathcal{F}_2) - \mathcal{F}_i$-messbar sind.

Satz 4.11. *Sei für beide i* $\sigma(\mathcal{E}_i) = \mathcal{F}_i$ *und es existieren Folgen* $(E_{i,n})_{n\in\mathbb{N}} \subset \mathcal{E}_i$ *mit* $E_{i,n} \uparrow \Omega_i$ *für* $n \uparrow \infty$. *Dann gilt*

$$\mathcal{F}_1 \otimes \mathcal{F}_2 = \sigma(\mathcal{E}_1 \times \mathcal{E}_2).$$

Beweis. $\mathcal{F}_1 \otimes \mathcal{F}_2 = \sigma(\mathcal{F}_1 \times \mathcal{F}_2) \supset \sigma(\mathcal{E}_1 \times \mathcal{E}_2)$ ist klar.

Ad $\mathcal{F}_1 \otimes \mathcal{F}_2 \subset \sigma(\mathcal{E}_1 \times \mathcal{E}_2)$. Wegen (4.3) muss für $i \in \{1,2\}$ gezeigt werden, dass $p^{-i}(A_i) \in \sigma(\mathcal{E}_1 \times \mathcal{E}_2)$ für alle $A_i \in \mathcal{F}_i$. Wegen Proposition 1.15 reicht es aus $p^{-i}(E_i) \in \sigma(\mathcal{E}_1 \times \mathcal{E}_2)$ für alle $E_i \in \mathcal{E}_i$ zu zeigen. Sei $E_1 \in \mathcal{E}_1$ (der Fall $i = 2$ unterscheidet sich nur etwas im Aufschrieb). Es gilt

$$p_1^{-1}(E_1) = E_1 \times \Omega_2 = E_1 \times (\cup_{n\in\mathbb{N}} E_{2,n}) = \cup_{n\in\mathbb{N}} \underbrace{(E_1 \times E_{2,n})}_{\in \mathcal{E}_1 \times \mathcal{E}_2,\ \forall n \in \mathbb{N}} \in \sigma(\mathcal{E}_1 \times \mathcal{E}_2).$$

□

Bemerkung 4.12. *Auf die Bedingung, dass* Ω_1 *und* Ω_2 *durch Erzeugermengen approximierbar sind, kann nicht verzichtet werden. Betrachte dazu das Beispiel* $\Omega_1 = \Omega_2 = \{1,2\}$ *und* $\mathcal{E}_1 = \mathcal{E}_2 = \{\{1\}\}$. $\mathcal{F}_1 \otimes \mathcal{F}_2$ *ist dann die Potenzmenge von* $\Omega_1 \times \Omega_2$, *während* $\sigma(\mathcal{E}_1 \times \mathcal{E}_2) = \{\{(1,1)\}, \{(1,2),(2,1),(2,2)\}, \emptyset, \Omega_1 \times \Omega_2\}$.

Korollar 4.13. *Es gilt*

$$\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}) = \mathcal{B}(\mathbb{R}^2)$$

(für die Definition von $\mathcal{B}(\mathbb{R})$ *und* $\mathcal{B}(\mathbb{R}^2)$ *siehe Definition 1.8).*

Beweis. Sei $\mathcal{E}_1 = \mathcal{E}_2 = \{(a,b) \mid a,b \in \mathbb{R}, a < b\}$, wobei mit (a,b) an dieser Stelle das offene Intervall mit Endpunkten a und b gemeint ist. Gemäß Satz 1.10 gilt $\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{E}_i)$ und $\mathcal{B}(\mathbb{R}^2) = \sigma(\mathcal{E}_1 \times \mathcal{E}_2)$. Mit Satz 4.11 folgt die Behauptung. □

Für eine Menge $A \in \mathcal{F}_1 \otimes \mathcal{F}_2$ definiere man die ω_1- und ω_2-Schnitte als

$$A_{\omega_1} := \{\omega_2 \in \Omega_2 \mid (\omega_1, \omega_2) \in A\}, \quad \forall \omega_1 \in \Omega_1$$

und

$$A^{\omega_2} := \{\omega_1 \in \Omega_1 \mid (\omega_1, \omega_2) \in A\}, \quad \forall \omega_2 \in \Omega_2.$$

Für eine Funktion $f : \Omega_1 \times \Omega_2 \to \overline{\mathbb{R}}$ definiert man, die ω_1-Schnitte als die Abbildungen

$$f_{\omega_1} : \Omega_2 \to \overline{\mathbb{R}}, \quad \omega_2 \mapsto f(\omega_1, \omega_2)$$

und die ω_2-Schnitte als die Abbildungen

$$f^{\omega_2} : \Omega_1 \to \overline{\mathbb{R}}, \quad \omega_1 \mapsto f(\omega_1, \omega_2).$$

Lemma 4.14. *Seien μ_1 und μ_2 σ-endlich und $f : \Omega_1 \times \Omega_2 \to \overline{\mathbb{R}}_+$ eine $(\mathcal{F}_1 \otimes \mathcal{F}_2) - \mathcal{B}(\overline{\mathbb{R}}_+)-$messbare Abbildung (eine solche Abbildung nennt man auch produktmessbar). Dann gilt*

(i) Die Abbildungen f_{ω_1} sind für alle $\omega_1 \in \Omega_1$ $\mathcal{F}_2 - \mathcal{B}(\overline{\mathbb{R}}_+)-$messbar und die Abbildungen f^{ω_2} sind für alle $\omega_2 \in \Omega_2$ $\mathcal{F}_1 - \mathcal{B}(\overline{\mathbb{R}}_+)-$messbar.

(ii) Die Abbildung $\omega_1 \mapsto \int f_{\omega_1}\, d\mu_2$ ist $\mathcal{F}_1 - \mathcal{B}(\overline{\mathbb{R}}_+)-$messbar und die Abbildung $\omega_2 \mapsto \int f^{\omega_2}\, d\mu_1$ ist $\mathcal{F}_2 - \mathcal{B}(\overline{\mathbb{R}}_+)-$messbar.

Bemerkung 4.15. *Für $A \subset \Omega_1 \times \Omega_2$ gilt*

$$(1_A)_{\omega_1} = 1_{A_{\omega_1}} \quad \textit{und} \quad (1_A)^{\omega_2} = 1_{A^{\omega_2}}.$$

Aus Lemma 4.14(i) folgt also die Implikation

$$A \in \mathcal{F}_1 \otimes \mathcal{F}_2 \implies A_{\omega_1} \in \mathcal{F}_2\ \forall \omega_1 \in \Omega_1 \textit{ und } A^{\omega_2} \in \mathcal{F}_1\ \forall \omega_2 \in \Omega_2. \tag{4.4}$$

Die Umkehrung von (4.4) gilt aber i.A. nicht. Betrachte als Beispiel $\mathcal{F}_1 = \mathcal{F}_2 = \mathcal{B}(\mathbb{R})$ und die Menge

$$A := \{(\omega_1, \omega_2) \in \mathbb{R}^2 \mid \omega_1 \in M \textit{ und } \omega_1 = \omega_2\},$$

wobei $M \notin \mathcal{B}(\mathbb{R})$. Definiere die Abbildung $d : \mathbb{R} \to \mathbb{R}^2,\ x \mapsto (x, x)$. Es gilt $d^{-1}(A) = M$. Wegen $d^{-1}(A_1 \times A_2) = A_1 \cap A_2$ für alle $A_1, A_2 \in \mathcal{B}(\mathbb{R})$ und Proposition 1.15 ist d jedoch $\mathcal{B}(\mathbb{R}) - (\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}))$-messbar, also kann A nicht in $\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$ liegen. Andererseits bestehen in diesem Beispiel alle Schnitte A_{ω_1} und A^{ω_2} aus maximal einem Punkt, nämlich höchstens aus ω_1 bzw. ω_2. Damit gilt $A_{\omega_1}, A^{\omega_2} \in \mathcal{B}(\mathbb{R})$ für alle $\omega_1, \omega_2 \in \mathbb{R}$.

Beweis. Beide Aussagen kann man mit einem Dynkin-Argument beweisen. Alternativ wollen wir bei Aussage (i) mit einer Spur-σ-Algebra argumentieren. Natürlich reicht es aus, den ersten Teil von (i) zu zeigen. Für festes ω_1 betrachte die Spur-σ-Algebra $(\{\omega_1\} \times \Omega_2) \cap (\mathcal{F}_1 \otimes \mathcal{F}_2)$. Mit Aufgabe 1(ii) folgt, dass diese σ-Algebra vom Mengensystem $(\{\omega_1\} \times \Omega_2) \cap (\mathcal{F}_1 \times \mathcal{F}_2)$ erzeugt wird (schneide $\{\omega_1\} \times \Omega_2$ mit allen Erzeugermengen). Zudem gilt

$$(\{\omega_1\} \times \Omega_2) \cap (\mathcal{F}_1 \times \mathcal{F}_2) = (\{\omega_1\} \times \Omega_2) \cap (\Omega_1 \times \mathcal{F}_2) = \{\widehat{\omega}_1\} \times \mathcal{F}_2,$$

wobei letzteres offensichtlich bereits eine σ-Algebra auf $\{\omega_1\} \times \Omega_2$ ist. Damit gilt

$$(\{\omega_1\} \times \Omega_2) \cap (\mathcal{F}_1 \otimes \mathcal{F}_2) = \{\omega_1\} \times \mathcal{F}_2.$$

Somit ist f eingeschränkt auf die Menge $\{\omega_1\} \times \Omega_2$ $(\{\omega_1\} \times \mathcal{F}_2) - \mathcal{B}(\overline{\mathbb{R}}_+)-$messbar (eine eingeschränkte messbare Abbildung ist messbar bzgl. der entsprechenden Spur-σ-Algebra, dies ergibt sich sofort aus den Definitionen). Dies ist offenbar äquivalent zu (i).

Ad (ii) Diese Aussage beweisen wir nun mit einem Dynkin-Argument. Natürlich reicht es aus, den ersten Teil von (ii) zu zeigen. Zunächst betrachten wir nur den Fall, dass μ_2 endlich ist. Sei

$$\mathcal{D} := \{A \in \mathcal{F}_1 \otimes \mathcal{F}_2 \mid f = 1_A \text{ erfüllt ersten Teil von Bedingung (ii)}\}$$

Da $\int (1_{A_1 \times A_2})_{\omega_1} \, d\mu_2 = 1_{A_1}(\omega_1)\mu_2(A_2)$ für alle $A_1 \in \mathcal{F}_1$, $A_2 \in \mathcal{F}_2$, gilt offensichtlich $\mathcal{F}_1 \times \mathcal{F}_2 \subset \mathcal{D}$. Des weiteren ist $\mathcal{D}$ ein Dynkinsystem, da für disjunkte Mengen $A_1, A_2, A_3, \ldots \in \mathcal{F}_1 \otimes \mathcal{F}_2$ auch die ω_1-Schnitte disjunkt sind,

$$\int (1_{\cup_{n\in\mathbb{N}} A_n})_{\omega_1} \, d\mu_2 = \mu_2((\cup_{n\in\mathbb{N}} A_n)_{\omega_1}) = \sum_{n=1}^{\infty} \mu_2((A_n)_{\omega_1}) = \sum_{n=1}^{\infty} \int (1_{A_n})_{\omega_1} \, d\mu_2.$$

gilt und Summen sowie Limiten messbarer Funktionen messbar sind (zudem gilt $\int (1_{A^c})_{\omega_1} \, d\mu_2 = \mu_2(\Omega_2) - \int (1_A)_{\omega_1} \, d\mu_2$).

Mit der Durchschnittsstabilität des Erzeugers $\mathcal{F}_1 \times \mathcal{F}_2$ und Satz 4.5 folgt $\mathcal{D} \supset \delta(\mathcal{F}_1 \times \mathcal{F}_2) = \sigma(\mathcal{F}_1 \times \mathcal{F}_2) = \mathcal{F}_1 \otimes \mathcal{F}_2$, d.h. der erste Teil von (ii) ist für alle $f = 1_A$, $A \in \mathcal{F}_1 \otimes \mathcal{F}_2$ und μ_2 endlich gezeigt. Sei nun f eine beliebige $(\mathcal{F}_1 \otimes \mathcal{F}_2) - \mathcal{B}(\overline{\mathbb{R}}_+)-$messbare Funktion. Definiere

$$f^{(n)} := \sum_{k=1}^{2^n n} \frac{k-1}{2^n} 1_{\{(k-1)2^{-n} \leq f < k2^{-n}\}}.$$

Es gilt $f^{(n)}_{\omega_1} = \sum_{k=1}^{2^n n} \frac{k-1}{2^n} 1_{\{(k-1)2^{-n} \leq f_{\omega_1} < k2^{-n}\}}$. Wegen der Linearität des Integrals gilt der erste Teil von (ii) für alle $f^{(n)}$. Die Folgen $(f^{(n)}_{\omega_1})_{n\in\mathbb{N}}$ sind offenbar nichtfallend und konvergieren punktweise gegen f. Mit dem Satz von der monotonen Konvergenz und der Tatsache, dass das Supremum von messbaren Funktionen messbar ist, folgt der erste Teil von (ii) auch für f.

Sei μ_2 nun ein σ-endliches Maß und $(A_n)_{n\in\mathbb{N}} \subset \mathcal{F}_2$ eine Folge mit $A_n \uparrow \Omega$ für $n \uparrow \infty$ und $\mu_2(A_n) < \infty$ für alle $n \in \mathbb{N}$. $\mu(\cdot \cap A_n)$ bezeichne das endliche Maß $A \mapsto \mu(A \cap A_n)$ auf $\mathcal{F}_2$. Es gilt

$$\int f_{\omega_1}\, d\mu_2(\cdot \cap A_n) = \int f_{\omega_1} 1_{A_n}\, d\mu_2 \uparrow \int f_{\omega_1}\, d\mu_2, \quad n \uparrow \infty,$$

wobei letzteres aus dem Satz von der monotonen Konvergenz folgt. Da das Supremum von messbaren Funktionen messbar ist, folgt (ii). □

Wegen $(1_A)_{\omega_1} = 1_{A_{\omega_1}}$ und Lemma 4.14(i) gilt $A_{\omega_1} \in \mathcal{F}_2$ und entsprechend $A^{\omega_2} \in \mathcal{F}_1$.

Für jede Menge $A \in \mathcal{F}_1 \otimes \mathcal{F}_2$ definiere man

$$\mu(A) := \int \int 1_A(\omega_1, \omega_2)\mu_2(d\omega_2)\mu_1(d\omega_1) = \int \mu_2(A_{\omega_1})\mu_1(d\omega_1). \tag{4.5}$$

Gemäß Lemma 4.14 ist das iterierte Integral in (4.5) erklärt (wegen der $\mathcal{F}_2$-Messbarkeit von $(1_A)_{\omega_1}$ zunächst die inneren Integrale und dann wegen der $\mathcal{F}_1$-Messbarkeit von $\omega_1 \mapsto \mu_2(A_{\omega_1})$ das äußere Integral).

Satz 4.16. *Das in (4.5) definierte μ ist ein Maß. Wenn μ_1 und μ_2 σ-endlich sind, ist μ auch σ-endlich und es ist das einzige Maß auf $\mathcal{F}_1 \otimes \mathcal{F}_2$, das den Rechteckmengen $A_1 \times A_2$ mit $A_1 \in \mathcal{F}_1$ und $A_2 \in \mathcal{F}_2$ den Wert $\mu_1(A_1)\mu_2(A_2)$ zuordnet. In diesem Fall gilt zudem*

$$\int \int 1_A(\omega_1, \omega_2)\mu_1(d\omega_1)\mu_2(d\omega_2) = \int \int 1_A(\omega_1, \omega_2)\mu_2(d\omega_2)\mu_1(d\omega_1), \quad \forall A \in \mathcal{F}_1 \otimes \mathcal{F}_2.$$

Definition 4.17. *Seien μ_1 und μ_2 σ-endlich. Das in (4.5) definierte Maß μ auf $\mathcal{F}_1 \otimes \mathcal{F}_2$ nennt man das Produktmaß von μ_1 und μ_2 und bezeichnet es mit $\mu_1 \otimes \mu_2$.*

Beweis von Satz 4.16. Es gilt $\mu(\emptyset) = 0$ und für eine disjunkte Folge $(A_n)_{n\in\mathbb{N}} \subset \mathcal{F}_1 \otimes \mathcal{F}_2$ sind auch die ω_1-Schnitte disjunkt mit $(\cup_{n\in\mathbb{N}} A_n)_{\omega_1} = \cup_{n\in\mathbb{N}} (A_n)_{\omega_1}$. Aus der σ-Additivität von μ_2 folgt

$$\sum_{n=1}^{\infty} \mu_2((A_n)_{\omega_1}) = \mu_2((\cup_{n\in\mathbb{N}} A_n)_{\omega_1})$$

für alle $\omega_1 \in \Omega_1$, was mit der Additivität des Integrals und dem Satz von der monotonen Konvergenz

$$\begin{aligned}
\sum_{n=1}^{\infty} \mu(A_n) &= \sum_{n=1}^{\infty} \int \mu_2((A_n)_{\omega_1})\mu_1(d\omega_1) \\
&\overset{\text{Add. \& mon. Konv.}}{=} \int \sum_{n=1}^{\infty} \mu_2((A_n)_{\omega_1})\mu_1(d\omega_1) \\
&= \int \mu_2((\cup_{n\in\mathbb{N}} A_n)_{\omega_1})\mu_1(d\omega_1) \\
&= \mu(\cup_{n\in\mathbb{N}} A_n)
\end{aligned}$$

nach sich zieht. μ ist also ein Maß. Zudem gilt

$$\mu(A_1 \times A_2) = \int \int 1_{A_1}(\omega_1) 1_{A_2}(\omega_2)\mu_2(d\omega_2)\mu_1(d\omega_1) = \mu_1(A_1)\mu_2(A_2). \tag{4.6}$$

Wenn μ_1 und μ_2 σ-endlich sind, können wir Satz 4.7 auf den durchschnittsstabilen Erzeuger $\mathcal{E} = \mathcal{F}_1 \times \mathcal{F}_2 = \{A_1 \times A_2 \mid A_1 \in \mathcal{F}_1,\ A_2 \in \mathcal{F}_2\}$ anwenden (man beachte, dass dann auch die dortige Bedingung (ii) erfüllt ist). Es folgt die Eindeutigkeit. Zudem überträgt sich die σ-Endlichkeit von μ_1 und μ_2 auf μ.

Hätten wir in (4.5) die umgekehrte Integrationsreihenfolge gewählt, wäre für die Rechteckmengen $A = A_1 \times A_2$ offenbar das gleiche herausgekommen. μ wäre wiederum ein Maß gewesen. Da Maße durch ihre Werte auf einem durchschnittsstabilen Erzeuger eindeutig bestimmt sind (Satz 4.7), würde es für alle Mengen aus $\mathcal{F}_1 \otimes \mathcal{F}_2$ mit dem tatsächlichen μ übereinstimmen. □

Bemerkung 4.18 ($\mu_1 \otimes \mu_2$-Nullmengen)**.** *Offenbar ist eine Menge $A \in \mathcal{F}_1 \otimes \mathcal{F}_2$ genau dann eine $\mu_1 \otimes \mu_2$-Nullmenge, wenn A_{ω_1} für μ_1-fast alle ω_1 eine μ_2-Nullmenge ist (bzw. äquivalent, wenn A^{ω_2} für μ_2-fast alle ω_2 eine μ_1-Nullmenge ist). Dies ergibt sich direkt aus der Definition in (4.5), unter Beachtung, dass das äußere Integral bzgl. μ_1 genau dann verschwindet, wenn der Integrand μ_1-fast überall verschwindet.*

Für Indikatorfunktionen 1_A spielt die Integrationsreihenfolge in (4.5) offenbar keine Rolle. Diesen Sachverhalt wollen wir für beliebige $\mathcal{F}_1 \otimes \mathcal{F}_2$-messbare Funktionen, für die das Integral bzgl. $\mu_1 \otimes \mu_2$ erklärt ist, verallgemeinern. Wie üblich in der Integrationstheorie ist es zweckmäßig, mit nichtnegativen Funktionen zu beginnen, da dann alle auftretenden Integrale erklärt sind.

Satz 4.19 (Satz von Fubini für nichtnegative Funktionen). *Seien* $(\Omega_1, \mathcal{F}_1, \mu_1)$ *und* $(\Omega_2, \mathcal{F}_2, \mu_2)$ *σ-endliche Maßräume und sei* $f : \Omega_1 \times \Omega_2 \to \overline{\mathbb{R}}_+$ *eine* $\mathcal{F}_1 \otimes \mathcal{F}_2$*-messbare Funktion. Dann sind die Funktionen*

$$\omega_2 \mapsto \int f^{\omega_2}\, d\mu_1 \quad bzw. \quad \omega_1 \mapsto \int f_{\omega_1}\, d\mu_2$$

$\mathcal{F}_2$*- bzw.* $\mathcal{F}_1$*-messbar und es gilt*

$$\begin{aligned}\int f\, d(\mu_1 \otimes \mu_2) &= \int \int f(\omega_1, \omega_2) \mu_1(d\omega_1) \mu_2(d\omega_2) \\ &= \int \int f(\omega_1, \omega_2) \mu_2(d\omega_2) \mu_1(d\omega_1)\end{aligned}$$

Beweis. Man beachte, dass obige Messbarkeiten bereits in Lemma 4.14(ii) gezeigt wurden.

Für $f = 1_A$ mit $A \in \mathcal{F}_1 \otimes \mathcal{F}_2$ ist die Aussage bereits in Satz 4.16 gezeigt worden (die Gleichheit mit $\int 1_A\, d(\mu_1 \otimes \mu_2)$ folgt unmittelbar aus der Definition von $(\mu_1 \otimes \mu_2)(A)$).

Sei nun f eine nichtnegative Elementarfunktion, d.h. $f = \sum_{k=1}^n \alpha_k 1_{A_k}$, $\alpha_k \in \mathbb{R}_+$. Wegen $f_{\omega_1} = \sum_{k=1}^n \alpha_k 1_{(A_k)_{\omega_1}}$ bzw. $f^{\omega_2} = \sum_{k=1}^n \alpha_k 1_{(A_k)^{\omega_2}}$ folgen die Aussagen aus der Linearität der Integrale (bei den iterierten Integralen jeweils zuerst auf das innere und dann auf das äußere Integral angewandt).

Sei nun f eine beliebige $(\mathcal{F}_1 \otimes \mathcal{F}_2) - \mathcal{B}(\overline{\mathbb{R}}_+)-$messbare Funktion. Die Folge

$$f^{(n)} := \sum_{k=1}^{2^n n} \frac{k-1}{2^n} 1_{\{(k-1)2^{-n} \le f < k 2^{-n}\}} + n 1_{\{f = \infty\}}$$

ist offenbar nicht-fallend und konvergiert punktweise gegen f. Selbiges gilt auch für die entsprechenden ω_i-Schnitte. Damit folgen die Aussagen für f aus dem Satz von der monotonen Konvergenz, für die iterierten Integrale wiederum zweimal angewandt, also

$$\begin{aligned}\int \int f(\omega_1, \omega_2) \mu_1(d\omega_1) \mu_2(d\omega_2) &= \int \lim_{n \to \infty} \int f^{(n)}(\omega_1, \omega_2) \mu_1(d\omega_1) \mu_2(d\omega_2) \\ &= \lim_{n \to \infty} \int \int f^{(n)}(\omega_1, \omega_2) \mu_1(d\omega_1) \mu_2(d\omega_2).\end{aligned}$$

□

Man beachte, dass auch wenn das Integral nach $\mu_1 \otimes \mu_2$ endlich ist, das innere Integral $\int f(\omega_1, \omega_2)\mu_2(d\omega_2)$ auf einer μ_1-Nullmenge trotzdem den Wert ∞ annehmen kann. Diese Beobachtung bereitet etwas Probleme, wenn wir wie bisher das Integral für allgemeine f mit der Zerlegung $f = f^+ - f^-$ auf den nichtnegativen Fall zurückführen wollen. Auch wenn f^+ und f^- beide bezüglich des Produktintegrals integrierbar wären, könnte bei der Differenz der inneren Integrale $\int f^+(\omega_1, \omega_2)\mu_1(d\omega_1) - \int f^-(\omega_1, \omega_2)\mu_1(d\omega_1)$ trotzdem der undefinierte Ausdruck $\infty - \infty$ auftreten. Da dies aber nur auf μ_2-Nullmengen passieren kann (bzw. auf μ_1-Nullmengen bei der umgekehrten Integrationsreihenfolge), kann man für eine leicht abgewandelte Funktion $\widehat{f}$ die Vertauschbarkeit der Integrationsreihenfolge zeigen.

Satz 4.20 (Satz von Fubini)**.** *Seien $(\Omega_1, \mathcal{F}_1, \mu_1)$ und $(\Omega_2, \mathcal{F}_2, \mu_2)$ σ-endliche Maßräume und sei $f \in \mathcal{L}^1(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2, \mu_1 \otimes \mu_2)$. Dann existiert eine $\mathcal{F}_1 \otimes \mathcal{F}_2$-messbare Funktion $\widehat{f}$ mit $\widehat{f} = f$ $(\mu_1 \otimes \mu_2)$-f.ü., so dass alle $\widehat{f}_{\omega_1}$ μ_2-integrierbar, alle $\widehat{f}^{\omega_2}$ μ_1-integrierbar und*

$$\begin{aligned}\int f \, d(\mu_1 \otimes \mu_2) = \int \widehat{f} \, d(\mu_1 \otimes \mu_2) &= \int \int \widehat{f}(\omega_1, \omega_2)\mu_1(d\omega_1)\mu_2(d\omega_2) \\ &= \int \int \widehat{f}(\omega_1, \omega_2)\mu_2(d\omega_2)\mu_1(d\omega_1).\end{aligned}$$

In der obigen Aussage kann die Bedingung $f \in \mathcal{L}^1(\mu_1 \otimes \mu_2)$ durch $f \in \mathcal{L}^0(\mu_1 \otimes \mu_2)$ und die Endlichkeit eines der beiden iterierten Integrale von $|f|$ ersetzt werden (in diesem Fall existiert das andere iterierte Integral also auch, ggf. nach einer Abwandlung der Funktion f auf einer $(\mu_1 \otimes \mu_2)$-Nullmenge).

Beweis. Sei $f \in \mathcal{L}^1(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2, \mu_1 \otimes \mu_2)$. Definiere

$$\widehat{f}(\omega_1, \omega_2) := \begin{cases} f(\omega_1, \omega_2) & : \quad \text{wenn } \int |f_{\omega_1}| \, d\mu_2 < \infty \text{ und } \int |f^{\omega_2}| \, d\mu_1 < \infty \\ 0 & : \quad \text{sonst} \, . \end{cases}$$

Wegen Lemma 4.14(ii) ist $\widehat{f}$ $(\mathcal{F}_1 \otimes \mathcal{F}_2)$-messbar und die Konstruktion stellt sicher, dass

$$\int |\widehat{f}^{\omega_2}| \, d\mu_1 < \infty \text{ und } \int |\widehat{f}_{\omega_1}| \, d\mu_2 < \infty, \quad \forall \omega_1, \omega_2.$$

Da $\mu_1(\{\omega_1 \mid \int |f_{\omega_1}|\, d\mu_2 = \infty\}) = \mu_2(\{\omega_2 \mid \int |f_{\omega_2}|\, d\mu_1 = \infty\}) = 0$, gilt zudem $\widehat{f} = f$ $(\mu_1 \otimes \mu_2)$-f.ü. und damit $\widehat{f} \in \mathcal{L}^1(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2, \mu_1 \otimes \mu_2)$. Wende nun Satz 4.19 auf $\widehat{f}^+$ und $\widehat{f}^-$ an. Es folgt

$$\begin{aligned}
\int f\, d(\mu_1 \otimes \mu_2) &= \int \widehat{f}^+\, d(\mu_1 \otimes \mu_2) - \int \widehat{f}^-\, d(\mu_1 \otimes \mu_2) \\
&= \int \int \widehat{f}^+(\omega_1, \omega_2)\mu_2(d\omega_2)\mu_1(d\omega_1) - \int \int \widehat{f}^-(\omega_1, \omega_2)\mu_2(d\omega_2)\mu_1(d\omega_1) \\
&= \int \left(\int \widehat{f}^+(\omega_1, \omega_2)\mu_2(d\omega_2) - \int \widehat{f}^-(\omega_1, \omega_2)\mu_2(d\omega_2) \right) \mu_1(d\omega_1) \\
&= \int \int (\widehat{f}^+(\omega_1, \omega_2) - \widehat{f}^-(\omega_1, \omega_2))\mu_2(d\omega_2)\mu_1(d\omega_1) \\
&= \int \int \widehat{f}(\omega_1, \omega_2)\mu_2(d\omega_2)\mu_1(d\omega_1)
\end{aligned}$$

und selbiges gilt für das iterierte Integral $\ldots \mu_1(d\omega_1)\mu_2(d\omega_2)$. Man beachte, dass in obigen Gleichungen wegen

$$\int \widehat{f}^+(\omega_1, \omega_2)\mu_2(d\omega_2), \int \widehat{f}^-(\omega_1, \omega_2)\mu_2(d\omega_2) \leq \int |\widehat{f}|(\omega_1, \omega_2)\mu_2(d\omega_2) < \infty$$

der Ausdruck $\infty - \infty$ nicht auftreten kann.

Setze nun $\int \int |f(\omega_1, \omega_2)|\mu_2(d\omega_2)\mu_1(d\omega_1) < \infty$ voraus. Aus Satz 4.19 (Satz von Fubini für nichtnegative Funktionen) angewandt auf die Funktion $|f|$ folgt $f \in \mathcal{L}^1(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2, \mu_1 \otimes \mu_2)$. □

Korollar 4.21 (Umordnung von Reihen)**.** *Sei* $(a_{n,m\in\mathbb{N}})$ *eine doppelt induzierte Folge mit* $\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} |a_{n,m}| < \infty$ *(oder äquivalent* $\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} |a_{n,m}| < \infty$*). Dann gilt*

$$\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{n,m} = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} a_{n,m}$$

Beweis. Wähle $\Omega_i = \mathbb{N}$, $\mathcal{F}_i = 2^{\mathbb{N}}$ und μ_i das Zählmaß für $i = 1, 2$. Die Aussage folgt dann unmittelbar aus Satz 4.20 unter Beachtung, dass $\widehat{f} = f$ gelten muss (da in diesem Beispiel die einzige Nullmenge von $\mu_1 \otimes \mu_2$ die leere Menge ist). □

Eine Anwendung:

Wir wollen das uneigentliche Integral

$$\int_{-\infty}^{\infty} \exp(-z^2)\, dz$$

berechnen. Ein analytischer Ausdruck für die Stammfunktion $\exp(-z^2)$ ist nicht bekannt. Wir betrachten die zweidimensionale Hilfsfunktion $f(x,y) := \exp(-(1+x^2)y^2)y$ und berechnen die iterierten Integrale $\dots dydx$ und $\dots dxdy$. Wir erhalten

$$\begin{aligned}
\int_0^\infty \int_0^\infty \exp(-(1+x^2)y^2)y\,dy\,dx &= \int_0^\infty \left[-\frac{1}{2(1+x^2)}\exp(-(1+x^2)y^2)\right]_{y=0}^{y=\infty} dx \\
&= \int_0^\infty \frac{1}{2(1+x^2)}dx \\
&= \left[\frac{1}{2}\arctan x\right]_{x=0}^{x=\infty} \\
&= \frac{\pi}{4}
\end{aligned}$$

und

$$\begin{aligned}
\int_0^\infty \int_0^\infty \exp(-(1+x^2)y^2)y\,dx\,dy &= \int_0^\infty \exp(-y^2)\int_0^\infty \exp(-(xy)^2)y\,dx\,dy \\
&= \int_0^\infty \exp(-y^2)\int_0^\infty \exp(-z^2)\,dz\,dy \\
&= \left(\int_0^\infty \exp(-z^2)\,dz\right)^2
\end{aligned}$$

(für die 2. Gleichheit benutzt man die Substitution $z := yx$ und damit $dz = ydx$). Mit dem Satz von Fubini für nichtnegative Funktionen (Satz 4.19) kann man die beiden Integrale gleichsetzen (Satz angewandt auf $\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}) = \mathcal{B}(\mathbb{R}^2)$, was mit Proposition 1.16 die Messbarkeit von f sichert, und auf $\mu_1 = \mu_2$ gleich dem Lebesgue-Maß auf $\mathbb{R}$, das σ-endlich ist) und es folgt

$$\int_{-\infty}^\infty \exp(-z^2)\,dz = 2\int_0^\infty \exp(-z^2)\,dz = 2\frac{\sqrt{\pi}}{2} = \sqrt{\pi}.$$

Dies zeigt, dass die Dichte der Normalverteilung mit Erwartungswert μ und Standardabweichung $\sigma > 0$

$$f(z) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(z-\mu)^2}{2\sigma^2}\right), \quad z \in \mathbb{R}$$

tatsächlich eine Dichte ist, also $\int_{-\infty}^\infty f(z)\,dz = 1$ gilt. Beachte jedoch, dass sich die Verteilungsfunktion $F(y) = \int_{-\infty}^y f(z)\,dz$ damit nicht ausrechnen lässt.

Völlig analog zu oben lassen sich Produkträume von endlich vielen Maßräumen konstruieren. Man definiere einfach

$$\mu_1 \otimes \ldots \otimes \mu_n(A) := \int \ldots \int 1_A(\omega_1, \ldots, \omega_n)\mu_1(d\omega_1)\ldots\mu_n(d\omega_1), \quad \forall A \in \mathcal{F}_1 \otimes \ldots \otimes \mathcal{F}_n.$$

Definition 4.22. *Das n-dimensionale Lebesgue-Maß λ^n auf dem $\mathbb{R}^n$ ist definiert als*

$$\lambda^n = \underbrace{\lambda \otimes \ldots \otimes \lambda}_{n\text{-mal}},$$

wobei λ das Lebesgue-Maß auf $\mathbb{R}$ bezeichnet (man beachte, dass λ σ-endlich ist).

Selbst zu einer überabzählbaren Familie von Maßräumen existieren Produktmaße. Deren Existenz lässt sich durch einen abstrakteren Fortsetzungssatz beweisen, was nicht Gegenstand dieser Vorlesung ist.

5 Satz von Radon-Nikodym

Definition 5.1 (Absolutstetigkeit, Äquivalenz und Singularität)**.** *Seien μ und ν Maße auf einer σ-Algebra $\mathcal{F}$. ν heißt absolutstetig bzgl. μ, geschrieben $\nu \ll \mu$, wenn*

$$\mu(A) = 0 \implies \nu(A) = 0, \quad \forall A \in \mathcal{F}.$$

Ist ν absolutstetig bzgl. μ und μ absolutstetig bzgl. ν, dann heißen ν und μ äquivalent, geschrieben $\nu \sim \mu$.

μ und ν heißen zueinander singulär, geschrieben $\mu \perp \nu$, wenn es eine Menge $A \in \mathcal{F}$ gibt mit $\mu(A) = 0$ und $\nu(A^c) = 0$.

Offenbar gilt $\mu \perp \nu$ genau dann wenn $\nu \perp \mu$. Anschaulich bedeutet Äquivalenz, dass die Masse unter beiden Maßen im „gleichen Bereich” liegt, während sie bei zueinander singulären Maßen in völlig unterschiedlichen Bereichen liegt. Auf $\mathbb{R}$ sind z.B. das Lebesgue-Maß und das Zählmaß, das nur die ganzen Zahlen zählt, zueinander singulär (als A kann die Menge der ganzen Zahlen gewählt werden, die eine Lebesgue-Nullmenge ist).

Satz 5.2 (Satz von Radon-Nikodym)**.** *Seien μ und ν σ-endliche Maße auf einer σ-Algebra $\mathcal{F}$. Dann sind äquivalent* *(i)* $\nu \ll \mu$

(ii) Es existiert eine $\overline{\mathbb{R}}_+$*-wertige messbare Funktion* h *mit* $\nu(A) = \int 1_A h \, d\mu$ *für alle* $A \in \mathcal{F}$.

(Umgekehrt induziert in (ii) jede $\overline{\mathbb{R}}_+$*-wertige, messbare Funktion* h *zusammen mit einem gegebenem* μ *ein neues Maß* ν*)*

Die Funktion h *(Radon-Nikodym-Ableitung von* ν *bzgl.* μ *genannt) ist dann* μ*-fast überall endlich und bis auf eine* μ*-Nullmenge eindeutig.*

Beweis der Implikation (ii) $\Longrightarrow$ *(i).* Die Implikation (ii) $\Longrightarrow$ (i) folgt sofort aus der Implikation $\Rightarrow$ in Satz 2.8(ii) (für jede μ-Nullmenge A gilt nämlich $1_A h = 0$ μ-f.ü. und damit $\int h 1_A \, d\mu = 0$).

Zudem kann mit jeder $\overline{\mathbb{R}}_+$-wertigen, messbaren Funktion h über $\nu(A) := \int 1_A h \, d\mu$ ein neues Maß ν definiert werden (die σ-Addivität von ν folgt aus dem Satz von der monotonen Konvergenz und der Additivität des Integrals, also $\int 1_{\cup_{n\in\mathbb{N}} A_n} h \, d\mu = \int \sum_{n=1}^{\infty} 1_{A_n} h \, d\mu = \lim_{m\to\infty} \int \sum_{n=1}^{m} 1_{A_n} h \, d\mu = \lim_{m\to\infty} \sum_{n=1}^{m} \int 1_{A_n} h \, d\mu =: \sum_{n=1}^{\infty} \int 1_{A_n} h \, d\mu$ für alle disjunkten Folgen $(A_n)_{n\in\mathbb{N}}$). Hierfür wird die σ-Endlichkeit von μ nicht gebraucht. Wenn μ jedoch σ-endlich ist und $\mu(h = \infty) = 0$, dann ist auch ν σ-endlich (beachte hierzu, dass $\cup_{n\in\mathbb{N}}\{h \leq n\} = \{h < \infty\}$). □

Der interessante Teil des Satzes ist natürlich die umgekehrte Richtung. Dafür bringen wir hier einen maßtheoretischen Beweis, der auf der sog. Hahnzerlegung beruht, die von unabhängigen Interesse ist (und für sog. signierte Maße später auch zum Beweis von Satz 5.7, der Jordanzerlegung für signierte Maße, benutzt wird). Ein funktionalanalytischer Beweis findet sich z.B. im Lehrbuch von Klenke [3].

Bemerkung 5.3. *Man kann auf die* σ*-Endlichkeit von* μ *verzichten, auf die* σ*-Endlichkeit von* ν *bei der Implikation (i)* $\Longrightarrow$ *(ii) aber nicht (siehe hierzu Aufgabe 9.2 in Brokate und Kersting [2]). In Bauer [1] auf den Seiten 117 und 118 wird gezeigt, wie der Fall für ein beliebiges* μ *auf den Fall, dass* μ σ*-endlich ist, zurückgeführt werden kann. In Anwendungen sind aber eigentlich alle Maße* σ*-endlich, so dass wir auf diese Feinheiten hier verzichten wollen.*

Bemerkung 5.4. *Betrachte den Fall, dass Ω abzählbar ist. In diesem Fall existiert eine abzählbare Familie $(A_i)_{i\in I}$ von Atomen von $\mathcal{F}$, d.h. $A_i \cap A_j = \emptyset$ für $i \neq j$ und*[†]

$$\mathcal{F} = \{A \in 2^\Omega \mid A = \bigcup_{i\in J} A_i \text{ für ein } J \subset I\}. \tag{5.1}$$

Wenn $\nu \ll \mu$, dann erfüllt die Abbildung

$$h(\omega) := \sum_{i\in I,\ \mu(A_i)>0} \frac{\nu(A_i)}{\mu(A_i)} 1_{A_i}(\omega) \tag{5.2}$$

offenbar Bedingung (ii) aus Satz 5.2. Es gilt nämlich

$$\int 1_{A_j} h \, d\mu = \int 1_{A_j} \frac{\nu(A_j)}{\mu(A_j)} \, d\mu = \nu(A_j), \quad \forall j \in I \tag{5.3}$$

(man beachte, dass wegen $\nu \ll \mu$ (5.3) auch für Atome A_j mit $\mu(A_j) = 0$ gilt). Da sich jedes $A \in \mathcal{F}$ als $A = \bigcup_{j\in J} A_j$ für ein $J \subset I$ schreiben lässt und die A_j disjunkt sind, folgt aus der Additivität und dem Satz von der monotonen Konvergenz

$$\int 1_A h \, d\mu = \sum_{j\in J} \int 1_{A_j} h \, d\mu = \sum_{j\in J} \nu(A_j) = \nu(A).$$

h ist also eine Version der Radon-Nikodym-Ableitung von ν bzgl. μ. Wenn wir zusätzlich voraussetzen, dass $\mu(A_i) > 0$ für alle $i \in I$, dann gibt es sogar nur eine Radon-Nikodym-Ableitung, da die leere Menge in diesem Fall die einzige μ-Nullmenge ist.

Die σ-Algebra $\mathcal{B}(\mathbb{R})$ besitzt jedoch keine abzählbare Familie von Atomen. Sie wird zwar (auch) von allen Intervallen (a, b) mit *rationalen* Endpunkten erzeugt (also einem abzählbaren Mengensystem), jedes Intervall (a, b) lässt sich aber weiter teilen. Die Einpunktmengen $\{a\}$ sind zwar Atome, allerdings gibt es überabzählbar viele davon. Die

[†]Sei $\Omega = \{\omega_1, \omega_2, \omega_3, \dots\}$. Wir wollen zeigen, dass es zu der σ-Algebra $\mathcal{F}$ eine abzählbare Familie von Atomen gibt. Definiere $A_1 := \bigcap_{B\in\mathcal{F},\ \omega_1\in B} B$. A_1 lässt sich als *abzählbarer* Schnitt schreiben, indem man alle $\omega = \omega_2, \omega_3, \dots$ durchläuft und schaut, ob man sie von ω_1 trennen kann, d.h. ob es ein $B \in \mathcal{F}$ gibt mit $\omega_1 \in B$ aber $\omega \notin B$. Damit folgt $A_1 \in \mathcal{F}$. Wenn $A_1 = \Omega$, sind wir fertig (In diesem Fall muss $\mathcal{F}$ jedoch trivial sein. Für jedes $B \in \mathcal{F}$ gilt nämlich $A_1 \subset B$ oder $A_1 \subset B^c$). Ansonsten sei n die kleinste natürliche Zahl > 1 mit $\omega_n \notin A_1$. Setze $A_2 := \bigcap_{B\in\mathcal{F},\ \omega_n\in B} B$. Es gilt wieder $A_2 \in \mathcal{F}$. Außerdem $A_1 \cap A_2 = \emptyset$ (sei $\omega \in A_1 \cap A_2$, es gilt für alle $B \in \mathcal{F}$ die Äquivalent $\omega_1 \in B \Leftrightarrow \omega \in B \Leftrightarrow \omega_n \in B$. Dies wäre aber ein Widerspruch zu $\omega_n \notin A_1$). Fährt man so fort, so konstruiert man die Familie von Atomen (Eigenschaft (5.1) lässt sich leicht zeigen).

spezielle Konstruktion der Radon-Nikodym-Ableitung in (5.2) beruht jedoch offenbar auf der Abzählbarkeit der Atome (somit lassen sich Summen bilden und Atome von Masse Null spielen keine Rolle). Trotzdem kann auch im allgemeinen Fall die Existenz einer Radon-Nikodym-Ableitung gezeigt werden. Anschaulich kann man sich h für Maße auf $\mathcal{B}(\mathbb{R})$ ähnlich wie in (5.2) vorstellen, wobei A_i sehr kleine Intervalle sind. h ist dann der in einem gewissen Sinne existierende Limes, wenn die Länge der Intervalle gegen Null geht.

Satz 5.5 (Hahnzerlegung). *Seien ν und ρ endliche Maße auf der σ-Algebra $\mathcal{F}$. Dann gibt es eine Menge $A_\leq \in \mathcal{F}$ mit Komplement $A_\geq := \Omega \setminus A_\leq$, so dass*

$$\nu(A) \leq \rho(A) \quad \forall A \in A_\leq \cap \mathcal{F} \tag{5.4}$$

$$\nu(A) \geq \rho(A) \quad \forall A \in A_\geq \cap \mathcal{F} \tag{5.5}$$

($A_\leq \cap \mathcal{F}$ und $A_\geq \cap \mathcal{F}$ bezeichnen die Spur-σ-Algebren, vgl. (1.2))

Beweis von Satz 5.5. Setze

$$\delta := \nu - \rho$$

(eine Mengenfunktion, die sich als Differenz zweier endlicher Maße schreiben lässt, ist ein endliches signiertes Maß, siehe Definition 5.6). Wir nennen eine Menge $N \in \mathcal{F}$ *negativ*, wenn $\delta(A) \leq 0$ für alle $A \in N \cap \mathcal{F}$, also für alle messbaren Teilmengen von N (selbst wenn $\delta(\Omega) < 0$ ist es nicht von vorneherein klar, dass es eine nichtleere negative Menge gibt).

Schritt 1: Wir zeigen nun folgende Aussage. Zu jeder Menge $A \in \mathcal{F}$ gibt es eine negative Menge $N \in A \cap \mathcal{F}$, für die $\delta(N) \leq \delta(A)$ gilt (und damit auch $\delta(A \setminus N) \geq 0$).

Man kann vom Fall $\delta(A) < 0$ ausgehen, da der Fall $\delta(A) \geq 0$ trivialerweise mit $N = \emptyset$ erfüllt ist. Wir konstruieren N durch sukzessives Herausnehmen von Teilmengen $B_1, B_2, B_3, \ldots$ aus A mit nichtnegativer Masse. Definiere $B_1 := \emptyset$. Sind $B_1, \ldots, B_k$ bereits definiert setze

$$s_k := \sup\{\delta(B) \mid B \in \mathcal{F},\ B \subset A \setminus (B_1 \cup B_2 \cup \ldots \cup B_k)\}.$$

Es gilt $s_k \geq \delta(\emptyset) = 0$. Wähle nun eine Menge $B_{k+1} \subset A \setminus (B_1 \cup B_2 \cup \ldots \cup B_k)$ und $\mu(B_{k+1}) \geq s_k/2$. Setze

$$N := A \setminus \cup_{k \in \mathbb{N}} B_k.$$

Es gilt $\delta(A \setminus N) = \sum_{k=1}^{\infty} \delta(B_k)$ und damit wegen $\delta(B_k) \geq 0$ und $N \subset A$ die eine zu zeigende Bedingung

$$\delta(A) - \delta(N) = \delta(A \setminus N) \geq 0. \tag{5.6}$$

Es folgt zudem $\delta(B_k) \to 0$ für $k \to \infty$. Also auch $s_k \to 0$. Sei nun $B \subset N \subset A$ und $k \in \mathbb{N}$. N ist disjunkt zu $B_1, \dots, B_k$ und damit gilt $B \subset A \setminus (B_1 \cup B_2 \cup \ldots \cup B_k)$. Mit der Definition von s_k folgt $\delta(B) \leq s_k$. Da dies für beliebige $k \in \mathbb{N}$ gilt folgt $\delta(B) \leq 0$. Also ist N eine negative Menge und zusammen mit (5.6) ist die Zwischenaussage bewiesen.

Schritt 2: Sei nun $\alpha := \inf\{\delta(A) \mid A \in \mathcal{F}\}$ und $(A_k)_{k\in\mathbb{N}} \subset \mathcal{F}$ mit $\delta(A_k) \to \alpha$ für $k \to \infty$. Mit Schritt 1 existiert eine Folge $(N_k)_{k\in\mathbb{N}}$ negativer Mengen mit $\delta(N_k) \to \alpha$ für $k \to \infty$. Setze

$$A_{\leq} := \cup_{k\in\mathbb{N}} N_k \quad \text{und} \quad A_{\geq} := \Omega \setminus A_{\leq}.$$

Da die abzählbare Vereinigung negativer Mengen offenbar wieder eine negative Menge ist (wieso ?), ist $A_{\leq}$ negativ, also

$$\delta(A) \leq 0 \quad \forall A \in A_{\leq} \cap \mathcal{F}. \tag{5.7}$$

Damit ist (5.4) also schonmal gezeigt. Zudem folgt

$$\delta(A_{\leq}) = \underbrace{\delta(A_{\leq} \setminus N_k)}_{\leq 0,\ \text{da } A_{\leq} \text{ negativ}} + \delta(N_k) \leq \delta(N_k), \quad \forall k \in \mathbb{N},$$

und damit $\delta(A_{\leq}) = \alpha$. Sei nun $A \subset A_{\geq}$. Es gilt

$$\delta(A) = \delta(A \cup A_{\leq}) - \delta(A_{\leq}) \geq \alpha - \alpha = 0. \tag{5.8}$$

und damit ist auch (5.5) erfüllt und insgesamt folgt die Behauptung. □

Beweis des Satzes von Radon-Nikodym. Implikation (i) $\Longrightarrow$ (ii):

Schritt 1: Seien sowohl ν also auch μ endliche Maße. Definiere

$$\mathcal{G} := \{g : \Omega \to \mathbb{R}_+ \ \mathcal{F}\text{-messbar} \mid \int g 1_A \, d\mu \leq \nu(A), \ \forall A \in \mathcal{F}\}.$$

Die Nullfunktion ist Element von $\mathcal{G}$, was folglich nichtleer ist. Zudem ist $\mathcal{G}$ **maximumsstabil**, d.h. $g_1, g_2 \in \mathcal{G} \implies g_1 \vee g_2 \in \mathcal{G}$. Dies folgt aus $\{g_1 \geq g_2\} \in \mathcal{G}$ und damit

$$\begin{aligned}\int (g_1 \vee g_2) 1_A \, d\mu &= \int g_1 1_{\{g_1 \geq g_2\}} 1_A \, d\mu + \int g_2 1_{\{g_1 < g_2\}} 1_A \, d\mu \\ &\leq \nu(A \cap \{g_1 \geq g_2\}) + \nu(A \cap \{g_1 < g_2\}) \\ &= \nu(A), \quad \forall A \in \mathcal{F}.\end{aligned}$$

Setze

$$\gamma := \sup_{g \in \mathcal{G}} \int g \, d\mu$$

Aus der Definition des Supremums folgt die Existenz einer Folge $(g_n)_{n\in\mathbb{N}} \subset \mathcal{G}$ mit $\sup_{n\in\mathbb{N}} \int g_n \, d\mu = \gamma$. Wegen der Maximumsstabilität sind $\widetilde{g}_n := g_1 \vee g_2 \vee \ldots \vee g_n \in \mathcal{G}$ für alle $n \in \mathbb{N}$. Da zudem $\int \widetilde{g}_n \, d\mu \geq \int g_n \, d\mu$ approximiert auch die Folge $(\widetilde{g}_n)_{n\in\mathbb{N}}$ das Supremum. Die Folge $(g_n)_{n\in\mathbb{N}}$ kann also o.B.d.A. von vorneherein aufsteigend gewählt werden (aufsteigend wird hier immer im Sinne von nichtfallend benutzt). Setze

$$\widehat{g} := \sup_{n\in\mathbb{N}} g_n.$$

$\widehat{g}$ ist als punktweises Supremum von *abzählbar vielen* $\mathcal{F}$-messbaren Funktionen wiederum $\mathcal{F}$-messbar[‡]. Mit dem Satz von der monotonen Konvergenz folgt $\gamma = \sup_{n\in\mathbb{N}} \int g_n \, d\mu = \int \widehat{g} \, d\mu$. Definiere

$$\beta(A) := \nu(A) - \int \widehat{g} 1_A \, d\mu, \ \forall A \in \mathcal{F}.$$

Wiederum wegen monotoner Konvergenz gilt $\int \widehat{g} 1_A \, d\mu = \sup_{n\in\mathbb{N}} \int g_n 1_A \, d\mu \leq \nu(A)$ für alle $A \in \mathcal{F}$ und damit $\widehat{g} \in \mathcal{G}$. Es folgt

$$\beta(A) \geq 0, \ \forall A \in \mathcal{F}. \tag{5.9}$$

[‡] $\widehat{g}$ wird auch das essentielle Supremum von $\mathcal{G}$ bzgl. des Maßes μ genannt. Es ist die kleinste messbare Funktion, die jede Funktion aus $\mathcal{G}$ μ-f.ü. dominiert. Das essentielle Supremum ist bis auf eine μ-Nullmenge eindeutig bestimmt. In der stochastischen Modellierung erweist sich das essentielle Supremum als der geeignete Supremumsbegriff für überabzählbare Mengen von Funktionen (Zufallsvariablen). Gegenüber dem punktweisen Supremum $\sup\{f(\omega) \mid f \in \mathcal{G}\}$ besteht der Vorteil, dass das essentielle Supremum per Konstruktion messbar ist und sich bei einer Abwandlung jedes $f \in \mathcal{G}$ auf einer von f abhängigen Nullmenge nicht verändert.

Wir wollen zeigen, dass $\beta(\Omega) = 0$. Aus $\beta(\Omega) = \beta(A) + \beta(A^c)$ würde dann Gleichheit in (5.9) für alle A folgen und wir hätten gezeigt, dass $\widehat{g}$ eine Radon-Nikodym-Dichte von ν bzgl. μ wäre.

Sei $\varepsilon > 0$. Vermöge $\rho(A) := \int (\widehat{g} + \varepsilon) 1_A \, d\mu$ für alle $A \in \mathcal{F}$ definiere man ein neues endliches Maß, das nach (ii) $\implies$ (i) bzgl. μ absolutstetig ist, d.h. $\rho \ll \mu$.

Die Hahnzerlegung von ν und ρ liefert die Existenz einer Menge $A_\leq$ (und $A_\geq = \Omega \backslash A_\leq$), so dass

$$\begin{aligned} \nu(A) \leq \rho(A) \quad \forall A \in A_\leq \cap \mathcal{F} \\ \nu(A) \geq \rho(A) \quad \forall A \in A_\geq \cap \mathcal{F} \end{aligned}$$

Für alle $A \in \mathcal{F}$ folgt

$$\nu(A) = \nu(A \cap A_\leq) + \nu(A \cap A_\geq) \geq \int \widehat{g} 1_{A \cap A_\leq} \, d\mu + \int (\widehat{g} + \varepsilon) 1_{A \cap A_\geq} \, d\mu = \int (\widehat{g} + \varepsilon 1_{A_\geq}) 1_A \, d\mu$$

und damit $(\widehat{g} + \varepsilon 1_{A_\geq}) \in \mathcal{G}$. Es gilt $\int (\widehat{g} + \varepsilon 1_{A_\geq}) \, d\mu = \gamma + \varepsilon \mu(A_\geq)$. Da γ nicht überschritten werden kann, folgt $\mu(A_\geq) = 0$. Aus $\nu \ll \mu$ und $\rho \ll \mu$ folgt $\nu(A_\geq) = \rho(A_\geq) = 0$ und damit

$$\beta(\Omega) - \varepsilon \mu(\Omega) \ = \ \nu(\Omega) - \rho(\Omega) = \nu(A_\leq) - \rho(A_\leq) \leq 0 \tag{5.10}$$

(man beachte, dass alle Maße endlich sind). Da (5.10) für beliebige $\varepsilon > 0$ gilt und β und μ nicht von ε abhängen, folgt $\beta(\Omega) \leq 0$. Da $\beta(\Omega) \geq 0$ schon gezeigt war, folgt $\beta(\Omega) = 0$ und die Behauptung.

Schritt 2: Seinen ν und μ σ-endliche Maße. Es existiert dann eine disjunkte Folge $(A_n)_{n \in \mathbb{N}}$ mit $\cup_{n \in \mathbb{N}} A_n = \Omega$ sowie $\nu(A_n) < \infty$ und $\mu(A_n) < \infty$ für alle $n \in \mathbb{N}$. Mit Schritt 1 erhalten wir nichtnegative Funktionen h_n mit $h_n 1_{A_n^c} = 0$ und $\nu(A \cap A_n) = \int h_n 1_A \mu$ für alle $A \in \mathcal{F}$. Mit der Additivität des Integrals und monotoner Konvergenz ist $h := \sum_{n=1}^\infty h_n$ dann eine Radon-Nikodym-Ableitung von ν bzgl. μ.

Schritt 3: Seien h und $\widetilde{h}$ Radon-Nikodym-Ableitungen. Wegen $\{h \leq \widetilde{h}\} \in \mathcal{F}$ gilt für alle $n \in \mathbb{N}$

$$\int h 1_{A_n \cap \{h \leq \widetilde{h}\}} \, d\mu = \int \widetilde{h} 1_{A_n \cap \{h \leq \widetilde{h}\}} \, d\mu = \nu(A_n \cap \{h \leq \widetilde{h}\}) < \infty.$$

Es folgt $\int (\widetilde{h} - h) 1_{A_n \cap \{h \leq \widetilde{h}\}}\, d\mu = 0$, was wegen $(\widetilde{h} - h) 1_{A_n \cap \{h \leq \widetilde{h}\}} \geq 0$ $\mu(\{\widetilde{h} > h\} \cap A_n) = 0$ impliziert. Analog folgt $\mu(\{\widetilde{h} < h\} \cap A_n) = 0$ und damit wegen $\cup_{n \in \mathbb{N}} A_n = \Omega$ $\widetilde{h} = h$, μ-f.ü.

Wegen $\int h 1_{A_n}\, d\mu = \nu(A_n) < \infty$ für alle $n \in \mathbb{N}$ und $\cup_{n \in \mathbb{N}} A_n = \Omega$, gilt zudem $\mu(h = \infty) = 0$.

□

Definition 5.6. *[Endliches signiertes Maß] Eine Abbildung $\delta : \mathcal{F} \to \mathbb{R}$, die σ-additiv ist, in dem Sinne, dass*

$$\delta(\cup_{n \in \mathbb{N}} A_n) = \sum_{n=1}^{\infty} \delta(A_n) \quad \textit{für alle Folgen disjunkter Mengen } A_1, A_2, \dots, \tag{5.11}$$

für die der Limes $\sum_{n=1}^{\infty} \delta(A_n)$ existiert, wird endliches signiertes Maß genannt.

Aus der σ-Additivität und der Endlichkeit folgt $\delta(\emptyset) = 0$. Geht man den Beweis der Hahnzerlegung mit der Konstruktion von negativen Mengen durch, dann sieht man, dass es reicht, (5.11) zunächst nur für solche disjunkte Folgen $(A_n)_{n \in \mathbb{N}}$ zu fordern, für die $(\delta(A_n))_{n \in \mathbb{N}}$ absolut summierbar ist, um daraus bereits die absolute Summierbarkeit der Maße für *alle* disjunkte Folgen zu folgern (die schwächere Bedingung liefert nämlich bereits eine Hahnzerlegung $(A_{\leq}, A_{\geq})$ für δ und 0 und mit $\delta(A_{\leq}), \delta(A_{\geq}) \in \mathbb{R}$ folgt die absolute Summierbarkeit der Maße für beliebige disjunkte Mengen).

Alternativ könnte also in obiger Definition gefordert werden, dass $(\delta(A_n))_{n \in \mathbb{N}}$ für alle disjunkten Folgen absolut summierbar ist, ohne dass sich im Ergebnis etwas ändern würde.

Satz 5.7 (Jordanzerlegung für endliche signierte Maße)**.** *Sei δ ein endliches signiertes Maß. Dann gibt es zwei endliche Maße δ^+ und δ^- mit $\delta = \delta^+ - \delta^-$ und $\delta^+ \perp \delta^-$. Durch diese Eigenschaften sind die Maße eindeutig bestimmt und es gilt*

$$\delta^+(A) = \sup_{B \in \mathcal{F},\ B \subset A} \delta(B) \quad \textit{und} \quad \delta^-(A) = -\inf_{B \in \mathcal{F},\ B \subset A} \delta(B). \tag{5.12}$$

Beweis. Schritt 1: Satz 5.5 ist zwar auf $\nu = \delta$ und $\rho = 0$ nicht anwendbar, da δ i.A. nicht $\mathbb{R}_+$-wertig ist. Geht man jedoch durch den Beweis, so sieht man, dass alle Schritte ohne Veränderungen durchgehen.

Schritt 2: Sei also $(A_{\leq}, A_{\geq})$ die Hahnzerlegung von δ und 0, d.h. $\delta(A) \leq 0 \quad \forall A \in A_{\leq} \cap \mathcal{F}$ und $\delta(A) \geq 0 \quad \forall A \in A_{\geq} \cap \mathcal{F}$. Setze nun

$$\delta^+(A) := \delta(A_{\geq} \cap A) \quad \text{und} \quad \delta^-(A) := -\delta(A_{\leq} \cap A).$$

Offensichtlich erfüllen dann δ^+ und δ^- die Bedingungen $\delta^+, \delta^- \geq 0$, $\delta = \delta^+ - \delta^-$ und $\delta^+ \perp \delta^-$.

Um die Eindeutigkeit der Zerlegung und (5.12) zu zeigen, betrachte man zwei endliche Maße μ und ν mit $\delta = \mu - \nu$ und $\mu \perp \nu$. Für messbare Mengen $B \subset A$ folgt

$$\delta(B) \leq \mu(B) \leq \mu(A). \tag{5.13}$$

Zudem existiert wegen der Singularität von μ und ν eine messbare Menge C mit $\nu(C) = \mu(C^c) = 0$. Es folgt

$$\delta(A \cap C) = \mu(A \cap C) = \mu(A). \tag{5.14}$$

(5.13) und (5.14) ergeben zusammen

$$\mu(A) = \sup_{B \in \mathcal{F},\ B \subset A} \delta(B).$$

Analog folgt

$$\nu(A) = -\inf_{B \in \mathcal{F},\ B \subset A} \delta(B)$$

und damit die Behauptung. □

Literatur

[1] Bauer, H. (1992) Maß- und Integrationstheorie, de Gruyter. 2. Auflage.

[2] Brokate, M. und Kersting, G. (2011) Maß und Integral, Birkhäuser.

[3] Klenke, A. (2008) Wahrscheinlichkeitstheorie, Springer, 2. Auflage.