Maß- und Integrationstheorie

Matthias Ludewig

Skript | WS 2024/25

Version vom 18.2.2025

Jegliche Fehler bitte bei mir melden =)

Contents

Die erweiterte reelle Achse

Sei $\mathbb{R}$ der angeordnete Körper der reellen Zahlen. Wir setzen

$$\overline{\mathbb{R}} := \{-\infty\} \cup \mathbb{R} \cup \{\infty\}$$

Wir dehnen die Ordnungsrelationen und Rechenoperationen auf $\mathbb{R}$ wie folgt aus:

$$\begin{aligned}
-\infty < a &< \infty && \text{für alle } a \in \mathbb{R} \\
\pm\infty \cdot a &= \pm\infty && \text{für alle } a > 0 \\
\pm\infty \cdot 0 &= 0 && \\
\pm\infty \cdot a &= \mp\infty && \text{für alle } a < 0 \\
a \pm \infty &= \pm\infty && \text{für alle } a \in \mathbb{R}.
\end{aligned}$$

Die Ausdrücke $\infty - \infty$, $\infty + (-\infty)$ etc. sind nicht definiert.

In $\overline{\mathbb{R}}$ hat jede monoton wachsende und jede monoton fallende Folge einen Grenzwert. Dieser Grenzwert ist in $\mathbb{R}$ wenn die Folge beschränkt ist und ∞ oder $-\infty$ falls die Folge unbeschränkt ist.

Wir setzen

$$\begin{aligned}
a \vee b &= \max\{a, b\} \\
a \wedge b &= \min\{a, b\},
\end{aligned}$$

sowie

$$\begin{aligned}
\bigvee_{i=1}^{n} a_i &= \max\{a_1, \dots, a_n\}, & \bigvee_{i=1}^{\infty} a_i &= \sup\{a_i \mid i \in \mathbb{N}\}, \\
\bigwedge_{i=1}^{n} a_i &= \min\{a_1, \dots, a_n\}, & \bigwedge_{i=1}^{\infty} a_i &= \inf\{a_i \mid i \in \mathbb{N}\}.
\end{aligned}$$

Wir setzen weiterhin

$$\begin{aligned}
a^+ &= a \vee 0 && \text{(Positivteil)} \\
a^- &= -(a \wedge 0) && \text{(Negativteil)} \\
|a| &= a^+ + a^- && \text{(Betrag).}
\end{aligned}$$

Wir dehnen die obige Notationen auf Funktionen aus, durch punktweise Anwendung. Sind also $f, g : X \to \overline{\mathbb{R}}$ Funktionen (X eine beliebige Menge), so setzen wir

$$(f \vee g)(x) := f(x) \vee g(x) \qquad \text{etc.}$$

1 Einführung

1.1 Das Maßproblem

Die Bestimmung des *Flächeninhalts* eines Bereichs in der Ebene sowie die Bestimmung des *Volumens* eines dreidimensionalen Körpers ein fundamentale Aufgabe der Mathematik. Ganz am Anfang steht die Konvention, dass die *Länge* des Intervals $[a, b] \subseteq \mathbb{R}$ durch die reellen Zahl $b-a$ gegeben wird. Das d-dimensionale Volumen eines Quaders im $\mathbb{R}^d$ ergibt sich dann als Produkt über seine Seitenlängen,

$$\operatorname{vol}\big([a_1, b_1] \times \cdots \times [a_d, b_d]\big) = \prod_{i=1}^{d}(b_i - a_i). \tag{1.1.1}$$

Dies ist im Wesentlichen ein Axiom, welches als Grundlegender Baustein für alle weiteren Volumenbegriffe fungiert.

Historische Lösungen das Problem, das Volumen eines gegebenen Körpers zu bestimmen, sind größtenteils aus (1.1.1) und/oder geometrischer Anschauung hergleitete Regeln. Fundamental sind hier die Additivität des Volumens, dass also

$$\operatorname{vol}(A \cup B) = \operatorname{vol}(A) + \operatorname{vol}(B) \tag{1.1.2}$$

für disjunkte Teilmengen $A, B, \subseteq \mathbb{R}^d$ gilt, sowie die Beobachtung, dass sich das Volumen einer Menge durch euklidische Bewegungen nicht ändert, dass also

$$\operatorname{vol}(E(A)) = \operatorname{vol}(A) \tag{1.1.3}$$

für jede Euklidische Bewegung

$$E : \mathbb{R}^d \to \mathbb{R}^d, \quad E(x) = Qx + v, \quad Q \in O(d),\ v \in \mathbb{R}^d$$

und beliebige Teilmengen $A \subseteq \mathbb{R}^d$ gilt.

Beispiele 1.1. Weitere wohlbekannte Regeln zur Volumenbestimmung sind

(1) die Regel “Grundfläche mal Höhe“ zur Berechnung des Flächeninhaltes von Dreiecken;

(2) weitere daraus abgeleitete Regeln zur Berechnung des Flächeninhaltes von Trapezen, Parallelogrammen etc.

(3) Die Scherinvarinz des Volumens;

(4) die *Keplersche Fassregel* zur Berechnung des Volumens von Rotationskörpern.

(5) Das Prinzip des Cavalieri.

Hierbei ist allerdings zu bemerken, dass keine dieser Regeln den Status eines mathematischen Satzes hat, ohne eine unabhängige, allgemeingültige Definition der Begriffe "Flächeninhalt" und "Volumen".

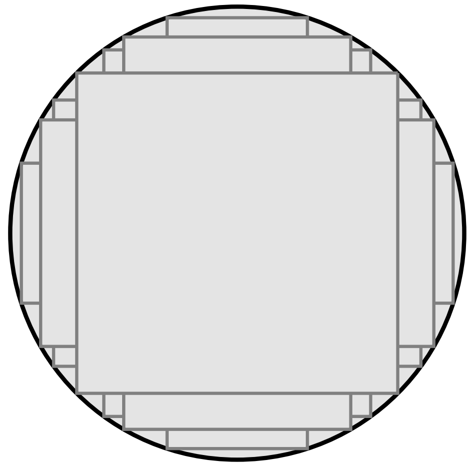
Das Ziel der modernen Maßtheorie ist daher zunächst einmal die theoretische Definition von Inhalten und Volumina, und zwar so, dass die obengenannten Regeln zu mathematischen Sätzen werden.

Konkret ist das Ziel, möglichst vielen Teilmengen $A \subseteq \mathbb{R}^d$ eine nichtnegative Zahl $\mathrm{vol}(A)$, das *Volumen* oder *Maß von* A zuzuordnen. Hierzu sei das Volumen von Quadern (1.1.1) als bekannt angenommen, und wir fordern zudem, dass das zu konstruierende Volumenmaß Additiv und bewegungsinvariant ist. Tatsächlich werden wir die Additivitätsbedingung (1.1.2) durch die stärkere Bedingung der sogenannten σ-Additivität ersetzen, welches besagt, dass sogar

$$\mathrm{vol}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathrm{vol}(A_i) \tag{1.1.4}$$

für jede abzählbare Familie $A_1, A_2, \dots$ von paarweise disjunkten Mengen gilt.

Bemerkung 1.2. Die Relevanz dieser stärkeren Bedingung mag auf den ersten Blick vielleicht nicht ganz offensichtlich sein, ergibt sich aber aus dem mathematischen Wunsch, Volumina approximativ berechnen zu können. Beispielsweise mag es wünschenswert sein, den (als unbekannt angenommenen) Flächeninhalt eines Kreises dadurch zu berechnen, indem man ihn durch immer kleiner werdende (notwendigerweise unendlich viele) Rechtecke ausschöpft.



Insgesamt gelangen wir also zu folgendem

Maßproblem 1.3. Konstruiere eine bewegungsinvariante, σ-additive Abbildung, die möglichst vielen Teilmengen $A \subseteq \mathbb{R}^d$ ein Volumen $\mathrm{vol}(A)$ zuordnet, so dass (1.1.1) für alle achsenparallelen Quader gilt.

Genauer gesagt fordern wir also, das vol die Eigenschaften (1.1.1), (1.1.3) und (3.1.1) erfüllt.

Optimistischerweise mag man zunächst hoffen, dass eine Lösung des Maßproblems existiert, die auf der gesamten Potenzmenge, also allen Teilmengen $A \subseteq \mathbb{R}^d$, definiert ist. Dies ist allerdings nicht der Fall:

Satz 1.4 (Vitali)**.** Jede σ-additive und translationsinvariante Mengenfunktion

$$\mu : \mathcal{P}(\mathbb{R}) \longrightarrow [0, \infty],$$

ist entweder konstant null oder konstant unendlich.

Beweis. Sei $\mu : \mathcal{P}(\mathbb{R}) \to [0, \infty]$ eine σ-additive Funktion mit der Eigenschaft, dass $\mu(A + x) = \mu(A)$ für all $A \subseteq \mathbb{R}$ und $x \in \mathbb{R}$. Wir bemerken zunächst, dass μ monoton ist, also $\mu(A) \leq \mu(B)$ falls $A \subseteq B$. In der Tat, wegen $B = A \cup (B \setminus A)$ gilt aufgrund der Additivität

$$\mu(B) = \mu(A) + \mu(B \setminus A) \geq \mu(A). \tag{1.1.5}$$

Definiere eine Äquivalenzrelation auf $[0,1] \subset \mathbb{R}$ durch

$$x \sim y \qquad \Longleftrightarrow \qquad x - y \in \mathbb{Q}.$$

Sei $V \subset [0,1]$ ein vollständiges Repräsentantensystem für diese Äquivalenzrelation, das heißt V enthält genau ein Element jeder Äquivalenzklasse. (Hier verwenden wir das Auswahlaxiom!) Insbesondere gilt $(V+q) \cap (V+p) = \emptyset$ für alle $p, q \in \mathbb{Q}$. Weiter kann jedes $x \in [0,1]$ kann geschrieben werden als $x = x' + q$ für $x' \in V$ und $q \in \mathbb{Q}$. Dann ist offenbar $q = x - x' \in [-1,1]$. Darum gilt

$$[0,1] \subseteq \bigcup_{q \in [-1,1] \cap \mathbb{Q}}^{\infty} V + q \subseteq [-1,2].$$

Aufgrund von (1.1.5) und der Translationsinvarianz von μ haben wir

$$\begin{aligned}
\mu([0,1]) \leq \mu\left(\bigcup_{q \in [-1,1] \cap \mathbb{Q}} V + q_i \right) \leq \mu([-1,2]) &= \mu([-1,0)) + \mu([0,1)) + \mu([1,2]) \\
&\leq \mu([-1,0]) + \mu([0,1]) + \mu([1,2]) \\
&= 3\mu([0,1])
\end{aligned}$$

Andererseits gilt wegen σ-Additivität und Translationsinvarianz, dass

$$\mu\left(\bigcup_{q \in [-1,1] \cap \mathbb{Q}} V + q \right) = \sum_{q \in [-1,1] \cap \mathbb{Q}} \mu(V+q) = \sum_{q \in [-1,1] \cap \mathbb{Q}} \mu(V).$$

Hieraus folgt, dass der Wert von $\mu\left(\bigcup_{i=1}^{\infty} V + q_i\right)$ entweder Null (falls $\mu(V) = 0$) oder unendlich sein muss (falls $\mu(V) > 0$). Mit der obigen Rechnung folgt also, dass auch $\mu([0,1]) \in \{0, \infty\}$ gelten muss. ■

Korollar 1.5. Es gibt keine Lösung des Maßproblems auf $\mathbb{R}^d$, die auf der gesamten Potenzmenge definiert ist.

Anders ausgedrückt muss eine Lösung vol des Maßproblems notwendigerweise auf einer auf einer echten Teilmenge $\mathcal{A} \subsetneq \mathcal{P}(\mathbb{R}^d)$ von Teilmengen definiert sein.

Beweis. Für $d = 1$ folgt dies direkt aus Satz 1.4. Für $d \geq 2$ bemerkt man, dass eine Lösung vol des Maßproblems auf $\mathbb{R}^d$ auch eine Lösung vol' des Maßproblems auf $\mathbb{R}^m$ für $m < d$ liefert, indem man

$$\mathrm{vol}'(A) = \mathrm{vol}(A \times [0,1]^{d-m})$$

setzt. ■

Bemerkung 1.6. Eine Variante des Maßproblems ist das sogenannte *Inhaltsproblem*, bei dem statt der σ-Additivität noch die Additivitätsbedingung (1.1.2) verlangt wird. Dieses schwächere Problem ist tatsächlich in den Dimensionen $d < 3$ lösbar, für $d = 1, 2$ gibt es also eine Funktion

$$\mu : \mathcal{P}(\mathbb{R}^d) \longrightarrow [0, \infty] ,$$

die (1.1.1), (1.1.2) und (1.1.3) erfüllt. Tatsächlich gibt es sogar *mehrere* Lösungen dieses Problems.

Für $d \geq 3$ allerdings ist auch dieses Problem nicht lösbar. Dies ist die Aussage des *Banach-Tarski-Paradox*, welches wir hier nicht beweisen wollen.

Satz 1.7 (Banach-Tarski)**.** Seien $X, Y \subseteq \mathbb{R}^d$ beschränkte Mengen mit nichtleerem Inneren. Dann existiert eine Zahl $n \in \mathbb{N}$ und eine Zerlegung

$$X = \bigcup_{i=1}^{n} A_i$$

in disjunkte Teilmengen $A_1, \dots, A_n$ von B, sowie Euklidische Bewegungen $E_1, \dots, E_n$ so, dass Y die disjunkte Vereinigung der $E_i(A_i)$ ist,

$$Y = \bigcup_{i=1}^{n} E_i(A_i) .$$

Gäbe es nun eine Lösung μ des Inhaltsproblems, so folgt aus Additivität und Bewegungsinvsarianz, dass für Mengen X und Y wie im obigen Satz stets $\mu(X) = \mu(Y)$ gelten muss. Dies widerspricht aber offenbar der Bedingung (1.1.1).

1.2 Integrationstheorie

Ein mit der Bestimmung von Inhalten und Volumina eng zusammenhängendes Problem ist die Konstruktion eines allgemeinen Integralbegriffes. Axiomatisch gesehen ist Integration zunächst einmal eine lineare Abbildung, welchen geeigneten ("integrierbaren") Funktionen eine reelle Zahl (ihr "Integral") zuordnet, und zwar so, dass das Integral einer nicht-negativen Funktion nicht-negativ ist. Angenommen, wir haben einen solchen Integralbegriff, dann können wir ein Volumenmaß durch die Vorschrift

$$\mathrm{vol}(A) := \int_{\mathbb{R}^d} \mathbf{1}_A(x) dx$$

definieren, wobei

$$\mathbf{1}_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases} \tag{1.2.1}$$

die Indikatorfunktion von A ist. Eine Voraussetzung für diese Definition ist natürlich, dass $\mathbf{1}_A$ zu der Klasse der integrierbaren Funktionen gehört.

Ein wohlbekannter Integralbegriff ist natürlich das Riemann-Integral, welches in der Analysis konstruiert wird, und hier als bekannt vorausgesetzt wird. Alle (stückweise) stetigen, kompakt getragenen Funktionen sind Riemann-integrierbar, eine darüber hinausgehende Charakterisierung gestaltet sich jedoch als schwierig. Insbesondere hat das Riemann-Integral nicht die folgende, intuitiv erwartbare Eigenschaft:

Forderung 1.8. Sei (f_n) eine monoton wachsende Folge von integrierbaren Funktionen mit der Eigenschaft, dass

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} f_n(x)dx < \infty \, .$$

Dann ist die punktweise Grenzfunktion f (falls sie endlich ist) wieder integrierbar, und ihr Integral ist der Grenzwert der Integrale der (f_n).

Beispiel 1.9. Sei (x_n) eine Aufzählung von $\mathbb{Q} \cap [0, 1]$ und setze

$$A_n = \{x_1, \dots, x_n\}.$$

Dann sind die Indikatorfunktionen $\mathbf{1}_{A_n}$ Riemann-integrierbar mit Riemann-Integral Null, aber die Folge $(\mathbf{1}_{A_n})$ konvergiert punktweise monoton gegen die Indikatorfunktion $\mathbf{1}_{\mathbb{Q}\cap[0,1]}$ von $\mathbb{Q} \cap [0, 1]$, ein Standardbeispiel für eine nicht Riemann-Integrierbare Funktion.

Es wird sich herausstellen, dass es eine (im Wesentlichen eindeutige) Erweiterung des Riemannschen Integralbegriffs zu einem Integralbegriff gibt, der Forderung 1.8 erfüllt. Dieses erweiterte Integral lässt wesentlich stärkere Grenzwertsätze zu als das Riemann-Integral, insbesondere den Satz von Lebesgue über dominierte Konvergenz.

2 Lebesgue-Integrale

Sei X eine beliebige Menge. Mit $\mathrm{Fun}(X, \mathbb{R})$ bzw. $\mathrm{Fun}(X, \overline{\mathbb{R}})$ bezeichnen wir die Räume der Funktionen auf X mit Werten in $\mathbb{R}$ bzw. $\overline{\mathbb{R}}$. $\mathrm{Fun}(X, \mathbb{R})$ ist ein Vektorraum mit Punktweiser Addition und skalarer Multiplikation,

$$(f+g)(x) := f(x) + g(x), \qquad (\lambda \cdot f)(x) := \lambda \cdot f(x)$$

für $f, g \in \mathrm{Fun}(X, \mathbb{R})$. Analog sind die Funktionen $f \vee g$ und $f \wedge g$ (punktweises Maximum/Minimum) definiert, sowie der Positivteil f^+, der Negativeteil f^- und der punktweise Betrag $|f| = f^+ + f^-$.

Weiterhin schreiben wir für $f, g \in \mathrm{Fun}(X, \overline{\mathbb{R}})$

$$f \leq g \quad :\Longleftrightarrow \quad \forall x \in X : f(x) \leq g(x).$$

Für eine Folge (f_n) von Funktionen auf X schreiben wir $f_n \nearrow f$, wenn die Folge punktweise monoton wachsend mit punktweisem Grenzwert f.

2.1 Vektorverbände und Elementarintegrale

Definition 2.1. Sei X eine beliebige Menge. Ein Vektorverband (von Funktionen auf X) ist ein Untervektorraum $\mathcal{V} \subseteq \mathrm{Fun}(X, \mathbb{R})$, der zusätzlich abgeschlossen unter den Operationen $\vee$ und $\wedge$ ist.

Mit anderen Worten: Ist $\mathcal{V}$ ein Vektorverband von Funktionen auf X, dann sind $f, g \in \mathcal{V}$ und $\lambda \in \mathbb{R}$ auch die Funktionen

$$f+g, \quad \lambda \cdot f, \quad f \wedge g \quad \text{und} \quad f \vee g$$

in $\mathcal{V}$ enthalten.

Lemma 2.2. Sei $\mathcal{V} \subseteq \mathrm{Fun}(X, \mathbb{R})$ ein Untervektorraum. Dann sind äquivalent:

(i) zu $f, g \in \mathcal{V}$ ist auch $f \vee g \in \mathcal{V}$;

(ii) zu $f, g \in \mathcal{V}$ ist auch $f \wedge g \in \mathcal{V}$;

(iii) zu $f \in \mathcal{V}$ ist auch $|f| \in \mathcal{V}$;

(iv) zu $f \in \mathcal{V}$ ist auch $f^+ \in \mathcal{V}$;

(v) $\mathcal{V}$ ist ein Vektorverband.

Beweis. Übungsaufgabe. ■

Beispiele 2.3. Folgendes sind Beispiele für Vektorverbände:

(1) Der Raum $\mathrm{Fun}(X, \mathbb{R})$ aller reellwertiger Funktionen auf X (dies ist der größtmögliche Vektorverband) und der Raum der konstanten Funktionen auf X (dies ist, abgesehen vom Nullraum, der kleinstmögliche Vektorverband).

(2) Der Raum $C(X)$ der stetigen Funktionen auf einem topologischen Raum X.

(3) Der Raum $\ell^\infty(\mathbb{N})$ aller beschränkten Folgen.

(4) Der Raum $c_0(\mathbb{N})$ aller Nullfolgen.

(5) Der Raum
$$\ell^1(\mathbb{N}) = \left\{ (a_n) \;\middle|\; \sum_{n=1}^{\infty} |a_n| < \infty \right\} \subset c_0$$
aller absolutsummierbaren Folgen.

(6) Der Raum $c_{00}(\mathbb{N}) \subset \ell^1(\mathbb{N})$ aller endlichen Folgen (bzw. der Raum aller Folgen mit nur endlich vielen Gliedern ungleich Null).

Keine Vektorverbände sind:

(N1) Der Raum $\mathrm{Pol}(\mathbb{R})$ aller Polynomfunktionen auf $\mathbb{R}$.

(N2) Der Raum $C^1(\mathbb{R})$ der stetig differenzierbaren Funktionen.

Definition 2.4. Sei X eine Menge und sei $\mathcal{V}$ ein Vektorverband von Funktionen auf X. Eine lineares Funktional $J : \mathcal{V} \to \mathbb{R}$ heißt *Elementarintegral*, falls

(i) J ist *positiv*, d.h., für alle $f \in \mathcal{V}$ mit $f \geq 0$ gilt
$$J(f) \geq 0.$$

(ii) J ist *stetig von unten*, d.h., für jede monoton wachsende Folge (f_n) von Funktionen $f_n \in \mathcal{V}$ mit $f_n \nearrow f$ und $f \in \mathcal{V}$ gilt
$$J(f_n) \nearrow J(f).$$

Ist $J : \mathcal{V} \to \mathbb{R}$ ein Elementarintegral, so heißt das Tripel $(X, \mathcal{V}, J)$ *Integrationsraum*.

Bemerkung 2.5. Die Bedingung (b) aus Definition 2.4 ist äquivalent zu jeder der folgenden Bedingungen:

(ii.1) J ist nullstetig von unten, d.h. für $f_n \nearrow 0$ gilt $J(f_n) \nearrow 0$;

(ii.2) J ist nullstetig von oben, d.h. für $f_n \searrow 0$ gilt $J(f_n) \searrow 0$;

(ii.3) J ist stetig von oben, d.h., für $f_n \searrow f$ mit $f \in \mathcal{V}$ gilt $J(f_n) \searrow J(f)$.

Hier ist jeweils (f_n) eine Folge von Funktionen in $\mathcal{V}$.

Lemma 2.6. Sei $(X, \mathcal{V}, J)$ ein Integrationsraum. Dann gelten:

(a) Für $f, g \in \mathcal{V}$ mit $f \leq g$ gilt $J(f) \leq J(g)$.

(b) Zu $f \in \mathcal{V}$ gilt $|J(f)| \leq J(|f|)$.

Beweis. (a) Gilt $f \leq g$, so ist $g - f \geq 0$, wegen der Positivität und Linearität von J gilt also

$$0 \leq J(g - f) = J(g) - J(f).$$

(b) Da $f_\pm \geq 0$ gilt $J(f_\pm) \geq 0$. Also wegen $f = f^+ - f^-$

$$|J(f)| = |J(f^+) - J(f^-)| \leq |J(f^+)| + |J(f^-)| = J(f^+) + J(f^-) = J(f^+ + f^-) = J(|f|). \qquad \blacksquare$$

Beispiele 2.7. Folgendes sind Beispiele für Elementarintegrale:

(1) Das Funktional

$$J : c_{00}(\mathbb{N}) \longrightarrow \mathbb{R}, \qquad (a_n) \longmapsto J\big((a_n)\big) := \sum_{n=1}^{\infty} a_n$$

ist ein Elementarintegral auf $\mathbb{N}$, definiert auf dem Raum der endlichen Folgen (Beispiel 2.3 (6)). Alternativ kann $c_{00}(\mathbb{N})$ auch durch den größeren Raum $\ell^1(\mathbb{N})$ der endlichen Folgen (Beispiel 2.3 (5)) ersetzt werden. In diesem Fall ist allerdings die Nullstetigkeit von oben vielleicht nicht ganz offensichtlich.

(2) Das Riemann-Integral

$$J(f) = \int_a^b f(x)\,dx$$

ist ein Elementarintegral auf dem Raum $C([a, b])$ der stetigen Funktionen auf $[a, b]$. Die Nullstetigkeit von oben folgt aus dem Satz von Dini 4.9.

2.2 Das erweiterte Integral

Definition 2.8. Sei X eine Menge und sei $\mathcal{V}$ ein Vektorverband von Funktionen auf X. Dann setzen wir

$$\mathcal{V}^* := \left\{F \in \mathrm{Fun}(X, \overline{\mathbb{R}}) \mid \exists (f_n) \subseteq \mathcal{V} : f_n \nearrow F\right\},$$

die Menge aller $\overline{\mathbb{R}}$-wertigen Funktionen, die punktweise monoton von unten durch Funktionen in $\mathcal{V}$ approximiert werden können.

Bemerkung 2.9. Es ist leicht zu sehen, dass zu $F, G \in \mathcal{V}^*$ sind auch stets $F + G, F \vee G$ und $F \wedge G$ in $\mathcal{V}$ enthalten. Für $\lambda \geq 0$ und $F \in \mathcal{V}^*$ ist außerdem $\lambda F \in \mathcal{V}^*$.

Bemerkung 2.10. Zu $F \in \mathcal{V}^*$ ist $-F$ im allgemeinen *nicht* in in $\mathcal{V}^*$ enthalten. $\mathcal{V}^*$ ist daher *kein* Vektorraum. Tatsächlich können Funktionen F von $\mathcal{V}^*$ den Wert ∞ annehmen, aber niemals den Wert $-\infty$.

Ist $(X, \mathcal{V}, J)$ ein Integrationsraum und (f_n) eine monoton wachsende Folge in $\mathcal{V}$, dann ist die Folge $(J(f_n))$ in $\mathbb{R}$ ebenfalls monoton wachsend (Lemma 2.6(a)), und hat darum stets einen Grenzwert in $\overline{\mathbb{R}}$.

Lemma 2.11. Sei $(X, \mathcal{V}, J)$ ein Integrationsraum. Seien (f_n) und (g_n) monoton wachsende Folgen von Funktionen in $\mathcal{V}$. Falls $\lim_n f_n \leq \lim g_n$ (punktweise), dann gilt $\lim_n J(f_n) \leq \lim_n J(g_n)$.

Bemerkung 2.12. Sind die punktweisen Grenzfunktionen $\lim_n f_n$ und $\lim_n g_n$ in $\mathcal{V}$, so folgt die Behauptung aus Lemma 2.6 (a) und der Stetigkeit von unten,

$$\lim_n J(f_n) = J(\lim_n f_n) \leq J(\lim_n g_n) = \lim_n J(g_n).$$

Der entscheidende Punkt ist jedoch, dass das obige Lemma gültig bleibt auch wenn die entsprechenden Grenzwerte nicht in $\mathcal{V}$ enthalten sind (sie liegen notwendigerweise in $\mathcal{V}^*$).

Beweis. Wir nehmen zunächst an, dass $f_n = 0$ für alle n. Dann gilt $\lim_n g_n \geq 0$, also $g_n \wedge 0 \nearrow 0$. Aus Nullstetigkeit von unten folgt daher, dass

$$0 = \lim_n J(g_n \wedge 0) \leq \lim_n J(g_n),$$

wobei wir für jedes n Lemma 2.6 (a) verwendet haben.

Im allgemeinen Fall definieren wir für festes k die monoton wachsende Folge $(h_n^{(k)})_n$

mit

$$h_n^{(k)} = g_n - f_k.$$

Dann gilt

$$\lim_n h_n^{(k)} = \lim_n g_n - f_k \geq \lim_n g_n - \lim_n f_n \geq 0,$$

da (f_n) monoton wachsend ist. Also gilt wegen dem Spezialfall

$$0 \leq J\big(\lim_n h_n^{(k)}\big) = \lim_n J(h_n^{(k)}) = \lim_n J(g_n) - J(f_k), \quad \text{also} \quad J(f_k) \leq \lim_n J(g_n).$$

Nehmen des Limes über k liefert die Behauptung. ■

Definition 2.13. Sei $(X, \mathcal{V}, J)$ ein Integrationsraum. Das *verallgemeinerte Integral* ist die Abbildung $J^* : \mathcal{V}^* \to \overline{\mathbb{R}}$ definiert durch

$$J^*(F) := \lim_n J(f_n),$$

wobei (f_n) eine beliebige monoton wachsende Folge in $\mathcal{V}$ mit $f_n \nearrow F$ ist.

Nach Lemma 2.11 ist die Definition von $J^*(F)$ ist unabhängig von der Wahl der approximierenden Folge (f_n).

Lemma 2.14. Sei $(X, \mathcal{V}, J)$ ein Integrationsraum. Das verallgemeinerte Integral J^* hat folgende Eigenschaften:

(a) Für Funktionen $f \in \mathcal{V}$ gilt $J^*(f) = J(f)$.

(b) Für $F, G \in \mathcal{V}^*$ gilt $J^*(F + G) = J^*(F) + J^*(G)$.

(c) Für $F \in \mathcal{V}^*$ und $\lambda \geq 0$ gilt $J^*(\lambda F) = \lambda J^*(F)$.

(d) Für $F, G \in \mathcal{V}^*$ mit $F \leq G$ gilt $J^*(F) \leq J^*(G)$.

(e) Ist (F_n) eine monoton wachsende Folge in $\mathcal{V}^*$ mit $F_n \nearrow F$, dann gilt $F \in \mathcal{V}^*$ und $\lim_n J^*(F_n) \nearrow J^*(F)$.

Beweis. Aussagen (a) – (c) sind trivial, und (d) folgt aus Lemma 2.11.

(e) Zu F_n wähle eine Folge $(f_m^{(n)})_m$ von Funktionen mit $f_m^{(n)} \nearrow F_n$. Setze

$$f_n := f_n^{(1)} \vee \cdots \vee f_n^{(n)}.$$

Dann ist (f_n) eine monoton wachsende (!) Folge in $\mathcal{V}$, also existiert $\lim_n f_n =: G$ und

ist in $\mathcal{V}^*$ enthalten. Wir behaupten, dass $G = F$. Für alle $m \geq n$ gilt $f_m^{(n)} \leq f_m$, also folgt
$$F_n = \lim_k f_k^{(n)} \leq \lim_m f_m = G.$$
nach Lemma 2.11. Im Grenzwert $n \to \infty$ gilt also $F \leq G$. Andererseits gilt
$$f_n = f_n^{(1)} \vee \cdots \vee f_n^{(n)} \leq F_1 \vee \cdots \vee F_n \leq F_n \vee \cdots \vee F_n = F_n$$
also $G = \lim_n f_n \leq \lim_n F_n = F$ (wieder mit Lemma 2.11).

Zusammen folgt $G = F$, also $F \in \mathcal{V}^*$ und $f_n \nearrow F$. Für das verallgemeinerte Integral gilt also
$$J^*(F) = \lim_n J(f_n) \leq \lim_n J^*(F_n) \leq J^*(F),$$
wobei wir $F_n \leq F$ und (d) benutzt haben. Damit muss in obiger Ungleichung in Wirklichkeit Gleichheit herrschen. ■

2.3 Erweiterung zum vollständigen Lebesgue-Integral

Definition 2.15 (Levifolgen)**.** Sei $(X, \mathcal{V}, J)$ ein Integrationsraum. Eine Funktion $F \in \mathcal{V}^*$ heißt *Levifunktion*, falls
$$J^*(F) < \infty.$$

Nach Definition von $\mathcal{V}^*$ und J^* sind Levifunktionen genau diejenigen Funktionen F für die es eine monoton wachsende Folge (f_n) in $\mathcal{V}$ gibt mit $\lim_n J(f_n) < \infty$.

Definition 2.16. Sei $(X, \mathcal{V}, J)$ ein Integrationsraum.

(a) $(X, \mathcal{V}, J)$ heißt *Lebesgue'scher Integrationsraum*, und J heißt *Lebesgue'sches Integral*, falls $\mathcal{V}$ alle endlichen Levifunktionen enthält (also alle Levifunktionen, die Werte in $\mathbb{R} \subset \overline{\mathbb{R}}$ annehmen).

(b) Eine Lebesgue'scher Integrationsraum $(X, \mathcal{V}, J)$ heißt *vollständig*, falls $\mathcal{V}$ sogar alle Funktionen enthält, die sich von einer Levifunktion nur in deren Unendlichkeitspunkten unterscheiden.

Genauer: Hat $f \in \mathrm{Fun}(X, \mathbb{R})$ die Eigenschaft, dass es eine Levifunktion F gibt mit $f(x) = F(x)$ für alle $x \in X$ mit $F(x) < \infty$, dann gilt $f \in \mathcal{V}$ und $J(f) = J^*(F)$.

Erweiterungssatz 2.17. Jeder Integrationsraum $(X, \mathcal{V}, J)$ besitzt genau eine minimale Erweiterung $(X, \bar{\mathcal{V}}, \bar{J})$ zu einem vollständigen Lebesgue'schen Integrationsraum, nämlich

$$\bar{\mathcal{V}} = \{f \in \mathrm{Fun}(X, \mathbb{R}) \mid \exists \text{ Levifunktionen } F, G : f + G = F\}$$

und

$$\bar{J}(f) = J^*(F) - J^*(G).$$

$(X, \bar{\mathcal{V}}, \bar{J})$ heißt *Lebesgue-Erweiterung* von $(X, \mathcal{V}, J)$.

Tatsächlich werden wir zeigen, dass jede andere Erweiterung $(X, \mathcal{V}', J')$ von $(X, \mathcal{V}, J)$ zu einem vollständigen Lebesgue'schen Integrationsraum schon $(X, \bar{\mathcal{V}}, \bar{J})$ enthält, dass also $\bar{\mathcal{V}} \subseteq \mathcal{V}'$ gilt und J' auf $\bar{\mathcal{V}}$ mit $\bar{J}$ übereinstimmt.

Bemerkung 2.18. Die Idee obiger Konstruktion ist, dass $\bar{\mathcal{V}}$ alle "Differenzen" von Levifunktionen enthält. Allerdings können wir nicht $f = F - G$ schreiben, da F und G beide gleichzeitig den Wert unendlich annehmen könnten und $\infty - \infty$ nicht definiert ist.

Bemerkung 2.19. Der Begriff der Levifunktion ist abhängig von der Wahl des Integrationsraums, also von $\mathcal{V}$ und J. Nach Definition hat Funktion $f \in \bar{\mathcal{V}}$ hat eine Darstellung $f + G = F$ mit Levifunktionen F, G bezüglich $(X, \mathcal{V}, J)$. Um nachzuweisen, dass $(X, \bar{\mathcal{V}}, \bar{J})$ Lebesgue'sch ist, muss jedoch überprüft werden, dass $\bar{\mathcal{V}}$ schon alle Levifunktionen bezüglich $(X, \bar{\mathcal{V}}, \bar{J})$ enthält. Um diese verschiedenen Levifunktionen zu unterscheiden, werden wir in den untenstehenden Beweisen von J-Levifunktionen und $\bar{J}$-Levifunktionen sprechen.

Lemma 2.20. Die Abbildung $\bar{J} : \bar{\mathcal{V}} \to \mathbb{R}$ ist wohldefiniert, also unabhängig von der gewählten Darstellung $f + G = F$ von $f \in \bar{\mathcal{V}}$.

Beweis. Sei $f \in \bar{\mathcal{V}}$ und seien $f + G = F$ und $f + G' = F'$ Darstellungen mit J-Levifunktionen $F, G, F', G' \in \mathcal{V}^*$. Addieren der beiden Gleichungen liefert

$$f + G + F' = f + G' + F, \quad \text{also} \quad G + F' = G' + F,$$

da f endlich ist. Unter Benutzung von Lemma 2.14 (b) folgt

$$J^*(F) - J^*(G) - \big(J^*(F') - J^*(G')\big) = J^*(F + G') - J^*(F' + G) = 0,$$

was zu zeigen war. ∎

Zum weiteren Beweis des Erweiterungssatzes brauchen wir die folgenden beiden Hilfsaussagen.

Lemma 2.21. Sei $f \in \bar{\mathcal{V}}$ und sei $\varepsilon > 0$. Dann besitzt f eine Darstellung $f + G = F$ mit J-Levifunktionen F und G, wobei $G \geq 0$ und $J^*(G) \leq \varepsilon$.

Beweis. Sei $f + G' = F'$ eine beliebige Darstellung von $f \in \bar{\mathcal{V}}$ mit J-Levifunktionen F' und G'. Sei (g_n) eine monoton wachsende Folge in $\mathcal{V}$ mit $g_n \nearrow G'$, also $J(g_n) \nearrow J^*(G')$. Sei n so groß, dass $J^*(G') - J(g_n) \leq \varepsilon$. Dann sind $G = G' - g_n$ und $F = F' - g_n$ wiederum J-Levifunktionen mit $f + G = F$, nach Konstruktion gilt $G \geq 0$ und wegen Lemma 2.14 (a) & (b) gilt

$$J^*(G') = J^*(G + g_n) = J^*(G) + J(g_n),$$

also $J^*(G) = J^*(G') - J(g_n) \leq \varepsilon$, wie gewünscht. ■

Lemma 2.22. Sei (f_n) eine monoton wachsende Folge in $\bar{\mathcal{V}}$ mit $\limsup_n \bar{J}(f_n) < \infty$. Falls der punktweise Grenzwert $f := \lim_n f_n$ endlich ist, gilt $f \in \bar{\mathcal{V}}$ und

$$\lim_n \bar{J}(f_n) = \bar{J}(f).$$

Beweis. Seien $f_n + G_n = F_n$ Darstellungen von f_n mit J-Levifunktionen F_n, G_n so, dass $G_n \geq 0$ und $J^*(G_n) \leq 2^{-n}$ (Lemma 2.21). Dann ist

$$\limsup_n J^*(F_n) \leq \limsup_n \bar{J}(f_n) + \limsup_n J^*(G_n) = \limsup_n \bar{J}(f_n) < \infty. \tag{2.3.1}$$

Setze

$$G := \sum_{k=1}^{\infty} G_k, \qquad G^{(n)} := \sum_{\substack{k=1 \\ k \neq n}}^{\infty} G_k,$$

Es ist leicht zu sehen, dass sowohl G als auch $G^{(n)}$ J-Levifunktionen sind, mit $J^*(G^{(n)}) \leq J^*(G) \leq 1$. Wir addieren $G^{(n)}$ zu $f_n + G_n = F_n$ und erhalten

$$f_n + G = f_n + G_n + G^{(n)} = F_n + G^{(n)} =: F'_n.$$

Dann ist F'_n als Summe der J-Levifunktionen F_n und $G^{(n)}$ ebenfalls eine J-Levifunktion. Da f_n monoton wachsend ist, ist auch (F'_n) monoton wachsend. Die punktweise Grenzfunktion $F := \lim_n F'_n$ ist damit nach Lemma 2.14 (e) in $\mathcal{V}^*$ enthalten und es gilt

$$J^*(F) = \lim_n J^*(F'_n) \leq \limsup_n J^*(F_n) + \limsup_n J^*(G^{(n)}) \leq \limsup_n J^*(F_n) + 1 < \infty,$$

wobei wir (2.3.1) benutzt haben. Also ist F eine Levifunktion. Die punktweise Grenzfunktion $f = \lim_n f_n$ erfüllt nach Konstruktion die Gleichung $f + G = F$ mit J-Levifunktionen F und G. Ist also f endlich, so gilt $f \in \bar{\mathcal{V}}$ und

$$\bar{J}(f) = J^*(F) - J^*(G) = \lim_n \left(J^*(F_n') - J^*(G)\right) = \lim_n \bar{J}(f_n). \qquad \blacksquare$$

Wir kommen nun zum

Beweis des Erweiterungssatzes 2.17. (1) Wir zeigen, dass $(X, \bar{\mathcal{V}}, \bar{J})$ ein Integrationsraum ist.

(a.a) Es ist leicht zu sehen, dass $\bar{\mathcal{V}}$ ein Untervektorverband von $\mathrm{Fun}(X, \mathbb{R})$ ist und dass $\bar{J} : \bar{\mathcal{V}} \to \mathbb{R}$ linear ist.

(b.b) $\bar{J}$ ist positiv: Sei $f \in \bar{\mathcal{V}}$ mit $f \geq 0$. Zu $\varepsilon > 0$ seien F, G J-Levifunktionen mit $f + G = F$, $G \geq 0$ und $J^*(G) \leq \varepsilon$, wie in Lemma 2.21 konstruiert. Dann ist $0 \leq f \leq F$, also $J^*(F) \geq 0$ wegen Lemma 2.14 (d). Es gilt also

$$\bar{J}(f) = J^*(F) - J^*(G) \geq J^*(F) - \varepsilon \geq -\varepsilon.$$

Da $\varepsilon > 0$ beliebig war, muss tatsächlich $\bar{J}(f) \geq 0$ gelten.

(c.c) $\bar{J}$ ist stetig von unten: Sei (f_n) eine monoton wachsende Folge in $\bar{\mathcal{V}}$ mit $f := \lim_n f_n$ in $\bar{\mathcal{V}}$. Da $\bar{J}$ positiv und linear ist, gilt $\bar{J}(f) \geq \bar{J}(f_n)$ für alle n, also

$$\limsup_n \bar{J}(f_n) \leq \bar{J}(f) < \infty.$$

Wegen Lemma 2.22 gilt darum $\lim_n \bar{J}(f_n) = \bar{J}(f)$.

(2) $(X, \bar{\mathcal{V}}, \bar{J})$ ist ein Lebesgue'scher Integrationsraum: Sei also F eine endliche $\bar{J}$-Levifunktion und sei (f_n) eine monoton wachsende Folge von Funktionen in $\bar{\mathcal{V}}$ mit $f_n \nearrow F$. Da F eine $\bar{J}$-Levifunktion ist, gilt

$$\infty > \bar{J}^*(F) = \lim_n \bar{J}(f_n) = \limsup_n \bar{J}(f_n).$$

Nach Lemma 2.22 ist F in $\bar{\mathcal{V}}$ enthalten, was zu zeigen war.

(3) $(X, \bar{\mathcal{V}}, \bar{J})$ ist vollständig: Sei $\bar{F}$ eine $\bar{J}$-Levifunktion und sei $f \in \mathrm{Fun}(X, \mathbb{R})$ eine beliebige Funktion, die mit $\bar{F}$ auf $\{x \in X \mid \bar{F}(x) < \infty\}$ übereinstimmt. Sei (f_n) eine monoton wachsende Folge von Funktionen in $\bar{\mathcal{V}}$ mit $f_n \nearrow \bar{F}$. Wie im Beweis von Lemma 2.22 existieren Darstellungen $f_n + G = F_n$ mit J-Levifunktionen G und F_n, wobei F_n monoton wachsend ist mit $F_n \nearrow F$. Nach Grenzübergang gilt

$$\bar{F} + G = \lim_n f_n + G = \lim_n F_n = F. \tag{2.3.2}$$

wobei F wegen Lemma 2.14 (e) eine J-Levifunktion ist. Wir behaupten, dass

$$f + F + G = 2F \tag{2.3.3}$$

gilt. Für $x \in X$ mit $\bar{F}(x) < \infty$ folgt dies direkt aus (2.3.2), da dort $f(x) = \bar{F}(x)$ gilt. Für $x \in X$ mit $\bar{F}(x) = \infty$ muss wegen (2.3.2) auch $F(x) = \infty$ gelten, also gilt (2.3.3) bei x, da beide Seiten unendlich sind. Damit haben wir die Darstellung (2.3.3) von f, was zeigt, dass f in $\bar{\mathcal{V}}$ enthalten ist. Schließlich gilt

$$\begin{aligned}
\bar{J}(f) &\overset{(2.3.3)}{=} J^*(2F) - J^*(F+G) \\
&\overset{2.14\text{ (b) \& (c)}}{=} J^*(F) - J^*(G) \\
&\overset{2.14\text{ (e)}}{=} \lim_n \left(J^*(F_n) - J^*(G)\right) \\
&= \lim_n \bar{J}(f_n) \\
&= \bar{J}^*(\bar{F}),
\end{aligned}$$

wie gewünscht.

(4) Wir zeigen nun, dass $(X, \bar{\mathcal{V}}, \bar{J})$ in jeder anderen vollständigen Lebesgue-Erweiterung enthalten ist: Sei $(X, \mathcal{V}', J')$ eine beliebige Erweiterung von $(X, \mathcal{V}, J)$ zu einem vollständigen Lebesgue'schen Integrationsraum. Sei $f \in \bar{\mathcal{V}}$ und sei $f + G = F$ eine Darstellung mit J-Levifunktionen G und F. Insbesondere sind G und F auch J'-Levifunktionen. Setze

$$G_0(x) = \begin{cases} G(x) & \text{falls } G(x) < \infty \\ 0 & \text{falls } G(x) = \infty, \end{cases} \qquad F_0(x) = \begin{cases} F(x) & \text{falls } F(x) < \infty \\ f(x) & \text{falls } F(x) = \infty, \end{cases}$$

Dann gehören G_0 und F_0 wegen der Vollständigkeitsannahme zu $\mathcal{V}'$, also ist ebenfalls $f = F_0 - G_0 \in \mathcal{V}'$. Damit gilt $\bar{\mathcal{V}} \subseteq \mathcal{V}'$.

Es bleibt zu zeigen, dass $J'(f) = \bar{J}(f)$. Sei (f_n) eine monoton wachsende Folge in $\mathcal{V}$ mit $f_n \nearrow F$. Dann gilt

$$J'(F_0) \overset{(\star)}{=} (J')^*(F) = \lim_n J'(f_n) \overset{(*)}{=} \lim_n J(f_n) = J^*(F).$$

Hierbei folgt $(\star)$ aus der Vollständigkeit von J' und $(*)$ gilt, da J' auf $\mathcal{V}$ mit J übereinstimmt. Analog gilt $J'(G_0) = J^*(G)$. Insgesamt also

$$J'(f) = J'(F_0) - J'(G_0) = J^*(F) - J^*(G) = \bar{J}(f). \qquad \blacksquare$$

Beispiel 2.23. Die Lebesgue-Erweiterung des Integrationsraums $(\mathbb{N}, c_{00}(\mathbb{N}), \sum)$ aus Beispiel 2.7 (1) ist $(\mathbb{N}, \ell^1(\mathbb{N}), \sum)$.

2.4 Grenzwertsätze

In diesem Abschnitt werden wir einige Grenzwertsätze für allgemeine Lebesgue'sche Integrationsräume zeigen. Der erste Satz zeigt, dass ein Lebesgue'sches Integral unsere Forderung 1.8 erfüllt.

Satz 2.24 (Beppo Levi)**.** Sei $(X, \mathcal{L}, J)$ ein Lebesgue'scher Integrationsraum. Dann gilt:

(1) Ist (f_n) eine monoton wachsenden Folge in $\mathcal{L}$ mit endlichem Grenzwert $f = \lim_n f_n$ und $\lim_{n\to\infty} J(f_n) < \infty$, so folgt $f \in \mathcal{L}$ und $J(f_n) \nearrow J(f)$.

(2) Ist (f_n) eine monoton fallende Folge in $\mathcal{L}$ mit endlichem Grenzwert $f = \lim_n f$ und $\lim_{n\to\infty} J(f_n) > -\infty$, so folgt $f \in \mathcal{L}$ und $J(f_n) \searrow J(f)$.

Beweis. Ist (f_n) eine Folge in $\mathcal{L}$ mit $f_n \nearrow f$ und $J^*(f) = \lim_n J(f_n) < \infty$, so ist f eine Levifunktion. Falls f endlich ist, so gilt $f \in \mathcal{L}$ nach Definition eines Lebesgue'schen Integrationsraumes. Also gilt $J(f_n) \nearrow J(f)$ wegen der Stetigkeitseigenschaften eines Elementarintegrals.

Die Aussage (2) folgt aus der Aussage (1), da für eine monoton fallende Folge (f_n) die Folge $(-f_n)$ monoton wachsend ist. ∎

Definition 2.25. Sei X eine Menge und sei $\mathcal{V}$ ein Vektorverband von Funktionen auf X. Eine Funktion $f \in \mathrm{Fun}(X, \overline{\mathbb{R}})$ heißt *messbar* (oder genauer *$\mathcal{V}$-messbar*), falls eine Folge (f_n) von Funktionen in $\mathcal{V}$ existiert, die punktweise gegen f konvergiert.

Wir schreiben $\mathcal{M} = \mathcal{M}(\mathcal{V})$ für die Menge der $\mathcal{V}$-messbaren Funktionen. $\mathcal{M}$ ist kein Vektorraum, da die Elemente von $\mathcal{M}$ die Werte $\pm\infty$ annehmen können, was die Summen- und Differenzenbildung im Allgemeinen verbietet.

Bemerkung 2.26. Für $f, g \in \mathcal{M}$ sind die Funktionen

$$f \vee g, \quad f \wedge g, \quad -f, \quad f^+, \quad f^-, \quad |f|$$

alle in $\mathcal{M}$ enthalten. Dies folgt direkt durch Anwenden der jeweiligen Operationen auf die approximierenden Folgen, unter Benutzung der Verbandseigenschaft von $\mathcal{V}$. Insbesondere folgt, dass die Menge der *endlichen* messbaren Funktionen (solche, die Werte in $\mathbb{R} \subset \overline{\mathbb{R}}$ annehmen) ein Untervektorraum von $\mathrm{Fun}(X, \mathbb{R})$ ist.

Lemma 2.27. Sei $(X, \mathcal{L}, J)$ ein Lebesgue'scher Integrationsraum und sei $\mathcal{M}$ die zugeörige Menge der messbaren Funktionen. Ist $f \in \mathcal{M}$ mit $f \geq 0$, so folgt $f \in \mathcal{L}^*$.

Beweis. Sei (f_n) eine Folge mit $f_n \to f$. Ohne Beschränkung der Allgemeinheit ist f_n positiv (ansonsten ersetzen wir f_n durch f_n^+). Wir setzen

$$g_n := \bigwedge_{i=n}^{\infty} f_i \ .$$

Dann ist g_n monoton wachsend und $g_n \nearrow \liminf_n f_n = f$. Um die Aussage $f \in \mathcal{L}^*$ zu erhalten, bleibt also zu zeigen, dass die Funktion g_n zu $\mathcal{L}$ gehören. Zunächst ist jedes g_n endlich, da die Folge f_n konvergiert und jede konvergente Folge ein endliches Infimum hat. Weiter gilt $\bigwedge_{i=n}^{m} f_i \in \mathcal{L}$ für alle $m, n \in \mathbb{N}$ und für jedes feste n gilt $\bigwedge_{i=n}^{m} f_i \searrow g_n$ (für $m \to \infty$). Damit folgt $g_n \in \mathcal{L}$ aus Teil (2) des Satzes von Beppo Levi. ■

Lemma 2.28. Sei $(X, \mathcal{L}, J)$ ein Lebesgue'scher Integrationsraum. Für eine endliche $\mathcal{L}$-messbare Funktion f sind äquivalent:

(i) $f \in \mathcal{L}$;

(ii) $|f| \in \mathcal{L}$;

(iii) $J^*(|f|) < \infty$.

Die Aussage (iii) in obigem Lemma ergibt Sinn, da $|f|$ wegen Lemma 2.27 (und Bemerkung 2.26) in $\mathcal{L}^*$ enthalten ist, also $J^*(|f|)$ wohldefiniert ist.

Beweis. (i) $\Rightarrow$ (ii) gilt wegen der Verbandseigenschaft von $\mathcal{L}$ und (ii) $\Rightarrow$ (iii) gilt nach Definition, da $J^* = J$ für Elemente von $\mathcal{L}$.

Für (iii) $\Rightarrow$ (i) sei $f \in \mathcal{M}$ endlich mit $J^*(|f|) < \infty$. Wir zerlegen $f = f^+ - f^-$. Wegen Lemma 2.27 sind $f^+, f^- \in \mathcal{L}^*$. Seien (f_n) und (g_n) Funktionenfolgen in $\mathcal{L}$ mit $f_n \nearrow f^+$ und $g_n \nearrow f^-$. Dann gilt

$$\lim_{n\to\infty} J(f_n) = J^*(f^+) \leq J^*(|f|) < \infty,$$

also ist f^+ eine endliche Levifunktion. Da $(X, \mathcal{L}, J)$ ein Lebesgue'scher Integrationsraum ist, folgt $f^+ \in \mathcal{L}$. Analog ist f^- und damit auch f in $\mathcal{L}$ enthalten. ■

Lemma 2.29. Sei $(X, \mathcal{L}, J)$ ein Lebesgue'scher Integrationsraum und sei $\mathcal{M}$ die zugehörige Menge der messbaren Funktionen. Für eine Folge (f_n) in $\mathcal{M}$ gilt auch

$$\liminf_{n\to\infty} f_n \in \mathcal{M}, \qquad \limsup_{n\to\infty} f_n \in \mathcal{M}\,.$$

Insbesondere ist der punktweise Limes von (f_n), falls er existiert, wieder in $\mathcal{M}$.

Beweis. Es genügt zu zeigen, dass $\limsup_{n\to\infty} f_n$ wieder in $\mathcal{M}$ ist; für die andere Aussage ersetze f_n durch $-f_n$ (siehe Bemerkung 2.26).

Wir nehmen zunächst an, dass f_n monoton wachsend ist. In diesem Fall ist $\limsup_n f_n = \lim_n f_n$. Es gilt außerdem $f_n^+ \nearrow f^+$ und $f_n^- \searrow f^-$. Wegen Lemma 2.27 sind f_n^+ und f_n^- in $\mathcal{L}^*$, also ist $f^+ \in \mathcal{L}^*$ (siehe Lemma 2.14 (e)). Insbesondere ist $f^+ \in \mathcal{M}$.

Um $f^- \in \mathcal{M}$ zu zeigen, wählen wir eine Folge (φ_n) nichtnegativer Funktionen in $\mathcal{L}$ mit $\varphi_n \nearrow f_1^-$. Für festes k sind dann die Funktionen $f_n^- \wedge \varphi_k$ nichtnegative endliche Levifunktionen (also in $\mathcal{L}$ enthalten) und es gilt $f_n^- \wedge \varphi_k \searrow f^- \wedge \varphi_k$. Insbesondere ist $f^- \wedge \varphi_k$ wieder in $\mathcal{L}$ nach Teil (2) des Satzes von Beppo Levi. Schließlich ist

$$\lim_{k\to\infty} f^- \wedge \varphi_k = f^- \wedge f_1^- = f^-\,,$$

also $f^- \in \mathcal{M}$. Insgesamt gilt $f = f^+ - f^- \in \mathcal{M}$.

Ist (f_n) monoton fallend, so gilt $\liminf_n f_n = \lim_n f_n \in \mathcal{M}$. Dies folgt aus der vorigen Aussage, da für eine solche Folge $(-f_n)$ monoton wachsend ist.

Sei nun (f_n) eine beliebige Folge in $\mathcal{M}$. Für ein festes $k \in \mathbb{N}$ setzen wir $g_n := f_k \vee \cdots \vee f_{k+n}$, was wieder in $\mathcal{M}$ enthalten ist (Bemerkung 2.26). Dann gilt $g_n \nearrow \bigvee_{n=k}^\infty f_n$, also gilt wegen dem obigen Spezialfall, dass

$$\bigvee_{n=k}^{\infty} f_n \in \mathcal{M}, \qquad k \in \mathbb{N} \text{ beliebig}\,. \tag{2.4.1}$$

Analog gilt

$$\bigwedge_{n=k}^{\infty} f_n \in \mathcal{M}, \qquad k \in \mathbb{N} \text{ beliebig}\,. \tag{2.4.2}$$

Die allgemeine Aussage folgt nun aus den obigen Spezialfällen zusammen mit der Beobachtung dass

$$\limsup_{n\to\infty} f_n = \bigwedge_{k=1}^{\infty} \bigvee_{n=k}^{\infty} f_n, \qquad \liminf_{n\to\infty} f_n = \bigvee_{k=1}^{\infty} \bigwedge_{n=k}^{\infty} f_n\,. \tag{2.4.3}$$

■

Lemma von Lebesgue-Fatou 2.30. Sei $(X, \mathcal{L}, J)$ ein Lebesgue'scher Integrationsraum. Ist (f_n) eine Folge nichtnegativer Funktionen in $\mathcal{L}$, so gilt $\liminf_n f_n \in \mathcal{L}^*$ und

$$J^*\big(\liminf_n f_n\big) \leq \liminf_n J(f_n)\ .$$

Hierbei beobachten wir, dass wegen Lemma 2.29 und Lemma 2.27 die Folge $\liminf_n f_n$ wieder in $\mathcal{L}^*$ enthalten ist, also ist die linke Seite von (2.30) wohldefiniert.

Beweis. Wir setzen $g_n := \bigwedge_{k=n}^{\infty} f_k$. Dann gilt $g_n \nearrow \liminf_n f_n$. Wegen (2.4.2) sind die Funktionen g_n messbar. Nach Konstruktion gilt

$$0 \leq g_n \leq f_n\ .$$

Wegen Lemma 2.28 ist $g_n \in \mathcal{L}^*$. Dass auch der monotone Grenzwert $\liminf_n f_n$ in $\mathcal{L}^*$ ist folgt entweder aus $J^*(g_n) \leq J^*(f_n) = J(f_n) < \infty$ und Lemma 2.28 oder auch wegen Lemma 2.14(e).

Wegen der obigen Betrachtungen ist also das verallgemeinerte integral wohldefiniert und es gilt

$$J^*(\liminf_n f_n) = \lim_n J(g_n) = \liminf_n J(g_n) \leq \liminf_n J(f_n)\ .$$

Hierbei ist zu beachten, dass der Grenzwert der Folge $(J(f_n))$ nicht unbedingt existiert, auch wenn der Grenzwert von $(J(g_n))$ wegen Monotonie existiert. ∎

Satz von Lebesgue über dominierte Konvergenz 2.31. Sei $(X, \mathcal{L}, J)$ ein Lebesgue'scher Integrationsraum. Sei (f_n) eine Folge von Funktionen in $\mathcal{L}$ mit punktweisem (endlichem) Grenzwert f. Falls eine Funktion $g \in \mathcal{L}$ mit $|f_n| \leq g$ für alle $n \in \mathbb{N}$ existiert, so ist f in $\mathcal{L}$ enthalten, und es gilt

$$J(f) = \lim_n J(f_n).$$

Eine Funktion g wie im Satz heißt *integrierbare Majorante*.

Beweis. Nach Lemma 2.29 ist $|f|$ als punktweiser Grenzwert der Folge $|f_n|$ in $\mathcal{M}$ enthalten, also wegen Lemma 2.27 in $\mathcal{L}^*$. Wegen $|f| \leq g$ gilt also

$$J^*(|f|) \leq J^*(g) = J(g).$$

Nach Lemma 2.28 ist also $f \in \mathcal{L}$.

Zur Berechnung von $J(f)$ wenden wir das Lemma von Lebesgue-Fatou auf die Folgen $(g + f_n)$ und $(g - f_n)$ an. Da nämlich g eine Majorante ist, sind $g + f_n$ und $g - f_n$ nichtnegativ. Also gilt

$$J(g) + J(f) = J(g + f) = J^*(\lim_n (g + f_n)) \leq \liminf_n J(g + f_n) = J(g) + \liminf_n J(f_n)$$
$$J(g) - J(f) = J(g - f) = J^*(\lim_n (g - f_n)) \leq \liminf_n J(g - f_n) = J(g) - \limsup_n J(f_n)\ ,$$

Zusammen gilt

$$\limsup_n J(f_n) \leq J(f) \leq \liminf_n J(f_n)\ ,$$

woraus die Aussage folgt. ■

Bemerkung 2.32. Auf die Existenz der integrierbaren Majorante im Satz von Lebesgue kann nicht verzichtet werden. Beispielsweise sind für die Lebesgue-Vervollständigung des Riemann-Integrals auf $[0, 1]$ die Funktionen

$$f_n(x) = \begin{cases} 0 & \text{if } x \in \{0\} \cup [\frac{1}{n}, 1] \\ n & \text{if } x \in (0, \frac{1}{n}) \end{cases}$$

integrierbar (sogar Riemann-integrierbar) mit Integral eins und der Nullfunktion als punktweisem Grenzwert. Hier existiert dementsprechend keine integrierbare Majorante.

Approximationssatz 2.33. Sei $(X, \mathcal{V}, J)$ ein Integrationsraum und sei $(X, \bar{\mathcal{V}}, \bar{J})$ die minimale Lebesgue-Erweiterung. Dann existiert zu jedem $f \in \bar{\mathcal{V}}$ eine Folge (f_n) von Funktionen in $\mathcal{V}$ mit $\bar{J}(|f - f_n|) \to 0$. Insbesondere gilt $J(f_n) \to \bar{J}(f)$.

Beweis. Sei $f \in \bar{\mathcal{V}}$ und seien F, G J-Levifunktionen mit $f + G = F$. Wir wählen Folgen (F_n), (G_n) in $\mathcal{V}$ mit $F_n \nearrow F$ und $G_n \nearrow G$ und setzen $f_n := F_n - G_n$. Wir wollen zeigen, dass (f_n) die gewünschten Eigenschaften hat. Dazu setzen wir

$$N = \{x \in X \mid F(x) = \infty\} = \{x \in X \mid G(x) = \infty\}\ .$$

Nach Konstruktion gilt für alle $x \in X \setminus N$, dass

$$f_n(x) = F_n(x) - G_n(x) \longrightarrow F(x) - G(x) = f(x)\ .$$

Für jede Funktion $g \in \bar{\mathcal{V}}$ gilt, dass $g \cdot \mathbf{1}_N + F = F$, also ist $g \cdot \mathbf{1}_N \in \bar{\mathcal{V}}$ und es gilt

$$\bar{J}(g \cdot \mathbf{1}_N) = J^*(F) - J^*(F) = 0.$$

Damit gilt auch $g \cdot \mathbf{1}_{X\setminus N} = g - g \cdot \mathbf{1}_N \in \bar{\mathcal{V}}$. Wir können daher aufspalten

$$\bar{J}(|f - f_n|) = \underbrace{\bar{J}(|f - f_n| \cdot \mathbf{1}_N)}_{=0} + \bar{J}(|f - f_n| \cdot \mathbf{1}_{X\setminus N}) = \bar{J}(|f - f_n| \cdot \mathbf{1}_{X\setminus N})\ .$$

Da $f_n \to f$ auf $X \setminus N$ konvergiert die Funktionenfolge $|f - f_n| \cdot \mathbf{1}_{X\setminus N}$ punktweise gegen Null. Wir wenden den Satz von Lebesgue über majorisierte Konvergenz an, um zu zeigen, dass auch das Integral dieser Folge gegen Null konvergiert. Hierzu brauchen wir eine integrierbare Majorante. Wir rechnen

$$\begin{aligned} |f - f_n| \cdot \mathbf{1}_{X\setminus N} &\leq (|f| + |f_n|) \cdot \mathbf{1}_{X\setminus N} \\ &\leq (|f| + |F_n| + |G_n|) \cdot \mathbf{1}_{X\setminus N} \\ &\leq (|f| + |F| + |G|) \cdot \mathbf{1}_{X\setminus N} \,. \end{aligned}$$

Mit F und G sind auf $|F|$ und $|G|$ wieder J-Levifunktionen, die ebenfalls endlich auf $X \setminus N$ und unendlich auf N sind. Aufgrund der Vollständigkeit von $(X, \bar{\mathcal{V}}, \bar{J})$ sind daher auch $|F| \cdot \mathbf{1}_{X\setminus N}$ und $|G| \cdot \mathbf{1}_{X\setminus N}$ in $\bar{\mathcal{V}}$ enthalten. Dies zeigt, dass $2|F| \cdot \mathbf{1}_{X\setminus N} + 2|G| \cdot \mathbf{1}_{X\setminus N}$ die gewünschte integrierbare Majorante ist. ■

3 Maßtheorie

3.1 Prämaße und Elementarintegrale

Definition 3.1. Sei X eine Menge und sei $\mathcal{A} \subseteq \mathcal{P}(X)$ eine nichtleere Menge von Teilmengen von X.

(1) $\mathcal{A}$ heißt *Mengenring* (über X), wenn zu je zwei Mengen $A, B \in \mathcal{A}$ die Mengen $A \cup B$ und $A \setminus B$ wieder in $\mathcal{A}$ enthalten sind.

(2) Ein Mengenring $\mathcal{A}$ heißt *unital* oder *Mengenring mit Eins*, falls $X \in \mathcal{A}$.

(3) Ein Mengenring mit Eins $\mathcal{A}$ heißt *σ-Algebra*, falls $\mathcal{A}$ abgeschlossen bezüglich abzählbarer Vereinigungen ist. Letzteres bedeutet, dass für $A_1, A_2, \dots \in \mathcal{A}$ stets auch die Vereinigung $\bigcup_{n=1}^{\infty} A_n$ in $\mathcal{A}$ ist.

Bemerkung 3.2. Die Terminologie Mengen*ring* erklärt sich viel folgt. Zunächst überlegt man sich, dass $\mathcal{A}$ genau dann ein Mengenring ist, wenn für $A, B \in \mathcal{A}$ auch stets die Mengen $A \triangle B$ und $A \cap B$ wieder in $\mathcal{A}$ sind. Hierbei ist

$$A \triangle B := (A \setminus B) \cup (B \setminus A)$$

die sogenannte *symmetrische Differenz* von A und B. Es ist dann nicht schwer zu zeigen, dass $\mathcal{A}$ mit den Operationen $\triangle$ (als Addition) und $\cap$ (als Multiplikation) ein Ring im

Sinne der Algebra ist. Dieser Ring hat ein Einselement genau dann wenn $\mathcal{A}$ unital im Sinne der obigen Definition ist.

Beispiele 3.3.

(1) Für jede beliebige Menge X ist $\mathcal{P}(X)$ selber eine σ-Algebra (die größte σ-Algebra von Teilmengen von X), sowie auch $\mathcal{A} = \{\emptyset, X\}$ (die kleinste σ-Algebra).

(2) Der Schnitt einer Familie von σ-Algebren von Teilmengen von X ist wieder eine σ-Algebra. Daraus folgt, dass zu einer beliebigen Teilmenge $\mathcal{M} \subseteq \mathcal{P}(X)$ der Schnitt aller σ-Algebren, die $\mathcal{M}$ enthalten, wieder eine σ-Algebra ist. Diese heißt die von $\mathcal{M}$ erzeugte σ-Algebra und wird mit $\sigma(\mathcal{M})$ bezeichnet.

(3) Ist X ein topologischer Raum, dann kann man nach (2) die kleinste σ-Algebra betrachten, die alle offenen Mengen enthält (wegen Axiom (ii) enthält diese auch alle abgeschlossenen Mengen). Diese heißt *Borelsche σ-Algebra*.

Definition 3.4. Sei $\mathcal{A}$ ein Mengenring von Teilmengen einer Menge X. Eine Funktion

$$\mu : \mathcal{A} \longrightarrow [0, \infty],$$

additiv, falls für je zwei disjunkte Mengen $A, B \in \mathcal{A}$ gilt

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

μ heißt *σ-additiv* oder *Prämaß*, falls die stärkere Bedingung (!)

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i), \tag{3.1.1}$$

gilt für eine beliebige abzählbare Familie $A_1, A_2, \ldots$ paarweise disjunkter Elemente von $\mathcal{A}$, deren Vereinigung wieder in $\mathcal{A}$ liegt. Das Tripel $(X, \mathcal{A}, \mu)$ heißt *Prämaßraum*. Elemente von $\mathcal{A}$ heißen *μ-messbar* oder (wenn der Kontext klar ist) *messbar*.

Ein Prämaß μ, welches auf einer σ-Algebra $\mathcal{A}$ definiert ist, heißt *Maß* und das Tripel $(X, \mathcal{A}, \mu)$ heißt dann *Maßraum*.

Ein Prämaß oder Maß heißt *endlich*, falls es nur Werte in $[0, \infty)$ annimmt.

Bemerkung 3.5. Folgende Eigenschaften von Prämaßen über Mengenringen folgen direkt aus der Definition:

(1) Es gilt stets $\mu(\emptyset) = 0$.

(2) Sind A und B beliebige (nicht notwendigerweise disjunkte) Mengen in $\mathcal{A}$, dann gilt

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

(3) Sind $A \subseteq B$ Mengen in $\mathcal{A}$, dann gilt

$$\mu(B) = \mu\big(A \sqcup (B \setminus A)\big) = \mu(A) + \mu(B \setminus A) \geq \mu(A).$$

Beispiel 3.6. Sei X eine beliebige Menge. Dann ist das System

$$\mathcal{E} = \{A \subseteq X \mid \#A < \infty\} \subset \mathcal{P}(X)$$

der endlichen Teilmengen von X ein Mengenring. Falls X nicht endlich ist, so ist $\mathcal{E}$ nicht unital und auch keine σ-Algebra. Die Abbildung

$$\mu : \mathcal{E} \longrightarrow [0, \infty), \qquad \mu(A) := \#A \ ,$$

die jeder Menge in $\mathcal{E}$ die Anzahl ihrer Elemente zuweist, ist ein Prämaß auf X, das *elementare Prämaß auf* X.

Beispiel 3.7. Zu $a = (a_1, \ldots, a_d)$ und $b = (b_1, \ldots, b_d)$ in $\mathbb{R}^d$ mit $a_i < b_i$ betrachten wir die halboffenen Quader

$$[a, b) := [a_1, b_1) \times \cdots [a_d, b_d) \subset \mathbb{R}^d.$$

Sei $\mathcal{Q}_d \subset \mathcal{P}(\mathbb{R}^d)$ das System der endlichen Vereinigungen solcher halboffener Quader. Dann ist $\mathcal{Q}_d$ ein Mengenring und es gibt genau ein Prämaß λ_d auf $\mathcal{Q}_d$ $\lambda_d : \mathcal{Q}_d \to [0, \infty)$ mit der Eigenschaft

$$\lambda_d\big([a, b)\big) = \prod_{i=1}^{d} (b_i - a_i) \ ,$$

siehe Lemma 5.1 für einen formalen Nachweis dieser Aussage.

Zum Nachweis der σ-Additivität sei (A_i) eine disjunkte Folge von Mengen in $\mathcal{Q}_d$, deren Vereinigung wieder in $\mathcal{Q}_d$ liegt. Dann gilt wegen Bemerkung 3.5, dass

$$\sum_{i=1}^{\infty} \lambda_d(A_i) = \lim_{n \to \infty} \sum_{i=1}^{n} \lambda_d(A_i) = \lim_{n \to \infty} \lambda_d \left(\bigcup_{i=1}^{n} A_i \right) \leq \lambda_d(A).$$

Für die umgekehrte Abschätzung wählen wir zu gegebenem $\varepsilon > 0$ eine Menge $A' \in \mathcal{Q}_d$ so, dass $\overline{A'} \subseteq A$ (der Abschluss von A' ist in A enthalten) und so, dass

$$\lambda_d(A') + \varepsilon \geq \lambda_d(A) \ .$$

Weiter wählen wir für jedes $n \in \mathbb{N}$ eine Menge $A_n'' \in \mathcal{Q}_d$ mit $A_n \subset (A_n'')^\circ$ (A_n ist im Inneren von A_n'' enthalten) und so, dass

$$\lambda_d(A_n) \geq \lambda_d(A_n'') - 2^{-n}\varepsilon.$$

Die Existenz solcher Mengen ist leicht einzusehen (!). Nach Konstruktion gilt dann

$$\overline{A'} \subseteq A = \bigcup_{i=1}^{\infty} A_i \subseteq \bigcup_{i=1}^{\infty} (A_i'')^\circ .$$

Also bilden die Mengen $(A_i'')^\circ$ eine offene Überdeckung von $\overline{A'}$. Wegen Kompaktheit wird $\overline{A'}$ schon von endlich vielen Mengen überdeckt, also gilt

$$A' \subseteq \bigcup_{i=1}^{n} A_i''$$

für ein $n \in \mathbb{N}$ groß genug. Wegen Additivität folgt dann

$$\begin{aligned}
\lambda_d(A) - \varepsilon \leq \lambda_d(A') \leq \lambda_d\left(\bigcup_{i=1}^{n} A_i''\right) &= \sum_{i=1}^{n} \lambda_d(A_i'') \\
&\leq \sum_{i=1}^{n} \lambda_d(A_i) + \sum_{i=1}^{n} 2^{-i}\varepsilon \\
&\leq \sum_{i=1}^{\infty} \lambda_d(A_i) + \varepsilon .
\end{aligned}$$

Da $\varepsilon > 0$ beliebig war, folgt die zu zeigende Ungleichung

$$\lambda_d(A) \leq \sum_{i=1}^{\infty} \lambda_d(A_i) .$$

Das folgende ist ein nützliches Kriterium, um nachzuweisen, dass eine gegebene positive, additive Mengenfunktion ein Prämaß ist. Hierbei ist eine Folge (A_n) von Teilmengen einer Menge X *aufsteigend*, falls $A_1 \subseteq A_2 \subseteq A_3 \subseteq \dots$ gilt.

Lemma 3.8. Sei $\mathcal{A} \subseteq \mathcal{P}(X)$ ein Mengenring. Eine additive Abbildung $\mu : \mathcal{A} \to [0, \infty]$ ist genau dann ein Prämaß, wenn für jede aufsteigende Folge (B_n) in $\mathcal{A}$, deren Vereinigung wieder in $\mathcal{A}$ liegt, gilt

$$\mu(B_n) \nearrow \mu\left(\bigcup_{n=1}^{\infty} B_n\right) . \tag{3.1.2}$$

Ist μ zudem endlich, dann kann diese Bedingung durch die äquivalente Bedingung der *Nullstetigkeit von oben* ersetzt werden: Zu jeder absteigenden Folge (B_n) in $\mathcal{A}$ mit $\bigcap_{n=1}^\infty B_n = \emptyset$ gilt

$$\mu(B_n) \searrow 0 \,.$$

Beweis. ($\Longrightarrow$) Sei μ ein Prämaß und sei (B_n) eine aufsteigende Folge von Mengen in $\mathcal{A}$. Wir setzen $A_0 = \emptyset$ und induktiv $A_i = B_i \setminus B_{i-1}$. Dann ist (A_i) eine disjunkte Familie von Teilmengen mit

$$B_n = \bigcup_{i=1}^n A_i \,, \qquad \text{und} \qquad \bigcup_{n=1}^\infty B_n = \bigcup_{i=1}^\infty A_i \,.$$

Wegen der σ-Additivität gilt daher

$$\mu(B_n) = \mu\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n \mu(A_i) \nearrow \sum_{i=1}^\infty \mu(A_i) = \mu\left(\bigcup_{i=1}^\infty A_i\right) = \mu\left(\bigcup_{n=1}^\infty B_n\right).$$

($\Longleftarrow$) Sei umgekehrt μ eine Abbildung wie im Lemma und sei (A_i) eine disjunkte Familie von Teilmengen in $\mathcal{A}$. Dann ist die Familie (B_n) mit $B_n = \bigcup_{i=1}^n A_i$ aufsteigend mit $\bigcup_{n=1}^\infty B_n = \bigcup_{i=1}^\infty A_i$ und also

$$\mu\left(\bigcup_{i=1}^\infty A_i\right) = \mu\left(\bigcup_{n=1}^\infty B_n\right) = \lim_{n\to\infty} \mu(B_n) = \lim_{n\to\infty} \sum_{i=1}^n \mu(A_i) = \sum_{i=1}^\infty \mu(A_i) \,.$$

Für den Zusatz bemerken wir, dass zu einer aufsteigenden Folge (B_n) die Folge (B'_n) mit $B'_n = (\bigcup_{i=1}^\infty B_i) \setminus B_n$ eine absteigende Folge mit leerem Durchschnitt ist. Die Aussage folgt dann aus der Identität

$$\mu(A \setminus B) = \mu(A) - \mu(B),$$

die die Endlichkeit von μ benutzt. ■

Korollar 3.9. Ist μ ein Prämaß auf $\mathcal{A}$ und $A_1 \supseteq A_2 \supseteq \dots$ eine absteigende Folge von Mengen in $\mathcal{A}$ mit $\mu(A_1) < \infty$, so gilt

$$\lim_{n\to\infty} \mu(A_n) = \mu\left(\bigcap_{n=1}^\infty A_n\right)$$

Definition 3.10. Sei $\mathcal{A}$ ein Mengenring über eine Menge X. Dann heißt der von den Indikatorfunktionen der Mengen in $\mathcal{A}$ erzeugte Vektorraum

$$\mathcal{S}(\mathcal{A}) = \left\{ \sum_{i=1}^{n} \lambda_i \mathbf{1}_{A_i} \,\middle|\, n \in \mathbb{N},\ \ \lambda_1, \ldots, \lambda_n \in \mathbb{R},\ \ A_1, \ldots, A_n \in \mathcal{A} \right\}$$

der Raum der *Treppenfunktionen* über $\mathcal{A}$.

Es ist leicht zu sehen, dass $\mathcal{S}(\mathcal{A})$ ein Vektorverband ist.

Satz 3.11. Sei μ ein endliches Prämaß auf einem Mengenring $\mathcal{A}$. Dann ist $(X, \mathcal{S}(\mathcal{A}), J_\mu)$ mit

$$J_\mu \left(\sum_{i=1}^{n} \lambda_i \mathbf{1}_{A_i} \right) = \sum_{i=1}^{n} \lambda_i \, \mu(A_i)$$

ein Elementarintegral.

Beweis. (1) Es ist zunächst zu zeigen, dass J_μ unabhängig von der Wahl der Darstellung der Treppenfunktion ist. Sei $f = \sum_{i=1}^{n} \lambda_i \mathbf{1}_{A_i}$ eine Darstellung von $f \in \mathcal{S}(\mathcal{A})$. Zu jeder Teilmenge $I \subseteq \{1, \ldots, n\}$ setzen wir

$$A_I := \bigcap_{i \in I} A_i \cap \bigcap_{i \notin I} X_0 \setminus A_i \qquad \text{mit} \qquad X_0 := \bigcup_{i=1}^{n} A_i$$

Die Mengen A_I, $I \subseteq \{1, \ldots, n\}$, sind in $\mathcal{A}$ enthalten, paarweise disjunkt und haben Vereinigung X_0. Nach Konstruktion ist A_i die disjunkte Vereinigung

$$A_i = \bigcup_{\substack{I \subseteq \{1, \ldots, n\}, \\ i \in I}} A_I \,,$$

und es gilt

$$f = \sum_{I \subseteq \{1, \ldots, n\}} \lambda_I \mathbf{1}_{A_I} \,, \qquad \text{wobei} \qquad \lambda_I := \sum_{i \in I} \lambda_i \,.$$

also gilt wegen Additivität von μ zunächst

$$\sum_{i=1}^{n} \lambda_i \, \mu(A_i) = \sum_{i=1}^{n} \lambda_i \sum_{\substack{I \subseteq \{1, \ldots, n\} \\ i \in I}} \mu(A_I) = \sum_{I \subseteq \{1, \ldots, n\},} \left(\sum_{i \in I} \lambda_i \right) \mu(A_I) = \sum_{I \subseteq \{1, \ldots, n\},} \lambda_I \, \mu(A_I) \,.$$

Seien $B_r := \{x \in X \mid f(x) = r\}$, $r \neq 0$, die Niveaumengen von f. Dann gilt

$$B_r = \bigcup_{\substack{I \subseteq \{1, \ldots, n\}, \\ \lambda_I = r}} A_I,$$

insbesondere sind die Niveaumengen in $\mathcal{A}$ enthalten. Aus der Additivität von μ folgt, dass

$$\sum_{I\subseteq\{1,\dots,n\},} \lambda_I\,\mu(A_I) = \sum_{r\neq 0} r \sum_{\substack{I\subseteq\{1,\dots,n\},\\ \lambda_I=r}} \mu(A_I) = \sum_{r\neq 0} r\,\mu(B_r)\,.$$

Letzterer Ausdruck ist offenbar unabhängig von der Wahl der Darstellung von f. Damit ist J_μ wohldefiniert.

(2) Wir weisen nun nach, dass J_μ ein Elementarintegral ist. J_μ ist offenbar linear und positiv, es bleibt also die Nullstetigkeit von oben zu zeigen. Sei (f_n) eine Folge in $\mathcal{S}(\mathcal{A})$ mit $f_n \searrow 0$. Zu $\varepsilon > 0$ setze

$$A_n = \{x \in X \mid f_n(x) \geq \varepsilon\}.$$

Da Treppenfunktionen nur endlich viele Werte annehmen, ist A_n eine endliche Vereinigung von Niveaumengen und daher in $\mathcal{A}$ enthalten. Da (f_n) punktweise gegen die Nullfunktion konvergiert ist $\bigcap_{n=1}^{\infty} A_n = \emptyset$. Sei $B = \{x \in X \mid f_1(x) \neq 0\}$ der Träger von f_1. Wegen $f_n \leq f_1$ ist der Träger von f_n in B enthalten. Es gilt also

$$0 \leq f_n \leq \varepsilon \cdot \mathbf{1}_{B\setminus A_n} + \|f_n\|_\infty \cdot \mathbf{1}_{A_n} \leq \varepsilon \cdot \mathbf{1}_{B\setminus A_n} + \|f_1\|_\infty \cdot \mathbf{1}_{A_n}\ ,$$

wobei $\|f_n\|_\infty$ das Maximum von f_n bezeichnet. Insgesamt folgt

$$J_\mu(f_n) \leq J_\mu\big(\varepsilon \cdot \mathbf{1}_{B\setminus A_n} + \|f_1\|_\infty \cdot \mathbf{1}_{A_n}\big) = \varepsilon \cdot \mu(B \setminus A_n) + \|f_1\|_\infty \cdot \mu(A_n)$$

Wegen Lemma 3.8 konvergiert die rechte Seite gegen $\varepsilon \cdot \mu(B)$. Da $\varepsilon > 0$ beliebig war, folgt die Behauptung. ■

Korollar 3.12. Jeder (Prä-)Maßraum $(X, \mathcal{A}, \mu)$ induziert einen Integrationsraum $(X, \mathcal{L}_\mu, J_\mu)$ durch Lebesgue-Vervollständigung des Integrationsraumes $(X, \mathcal{S}(\mathcal{A}_0), J_\mu)$, wobei $\mathcal{A}_0 := \{A \in \mathcal{A} \mid \mu(A) < \infty\}$.

3.2 Maße aus Lebesgue-Integralen

Definition 3.13. Sei $(X, \mathcal{V}, J)$ ein Integationsraum und sei $\mathcal{M}$ die zugehörige Menge der messbaren Funktionen. Eine Teilmenge $A \subseteq X$ heißt *messbar* (genauer, $\mathcal{V}$-messbar), falls $\mathbf{1}_A \in \mathcal{M}$ gilt. Wir schreiben $\mathcal{A}$ für das System der messbaren Mengen.

Lemma 3.14. Sei $(X, \mathcal{V}, J)$ ein Integationsraum. Das zugehörige System $\mathcal{A}$ der messbaren Mengen ist ein Mengenring.

Beweis. Seien A und B messbar, also $\mathbf{1}_A, \mathbf{1}_B \in \mathcal{M}$. Dann sind auch $\mathbf{1}_A + \mathbf{1}_B$ und $\mathbf{1}_A \cdot \mathbf{1}_B$ in $\mathcal{M}$: Eine approximierende Folge ergibt sich als Summe bzw. Produkt der approximierenden Folgen. Genauso sind auch

$$\mathbf{1}_{A\cup B} = \mathbf{1}_A + \mathbf{1}_B - \mathbf{1}_A \cdot \mathbf{1}_B, \qquad \text{und} \qquad \mathbf{1}_{A\setminus B} = \mathbf{1}_A - \mathbf{1}_A \cdot \mathbf{1}_B$$

in $\mathcal{M}$. Also sind $A \cup B$ und $A \setminus B$ messbar. ■

Lemma 3.15. Sei $(X, \mathcal{L}, J)$ ein Lebesgue'scher Integrationsraum und sei $\mathcal{A}$ das zugehörige System der messbaren Mengen. Ist $A_1, A_2, \ldots$ eine abzählbare Familie von messbaren Mengen, dann gilt

$$\bigcup_{n=1}^{\infty} A_n \in \mathcal{A} .$$

Beweis. Sei $A := \bigcup_{n=1}^{\infty} A_n$. Dann ist $\mathbf{1}_A$ der punktweise Limes der messbaren Funktionenfolge $f_n := \mathbf{1}_{A_1\cup\cdots\cup A_n}$. Nach Lemma 2.29 ist $\mathbf{1}_A$ in $\mathcal{M}$ enthalten. ■

Definition 3.16. Ein Lebesgue'scher Integrationsraum $(X, \mathcal{L}, J)$ heißt *Stone'sch*, falls X eine $\mathcal{L}$-messbare Menge ist.

Ist $(X, \mathcal{L}, J)$ ein Lebesgue'scher Integrationsraum mit zugehörigem System der messbaren Mengen $\mathcal{A}$, so gilt $\mathbf{1}_A \in \mathcal{V}^*$ nach Lemma 2.27. Wir können also definieren

$$\mu : \mathcal{A} \longrightarrow [0, \infty], \qquad \mu(A) := J^*(\mathbf{1}_A) .$$

Satz 3.17. Ist $(X, \mathcal{L}, J)$ Stone'sch, dann ist $(X, \mathcal{A}, \mu)$ ein Maßraum.

Beweis. Sei $A_1, A_2, A_3, \ldots$ eine Folge paarweise disjunkter Mengen in $\mathcal{A}$ mit Vereinigung $A = \bigcup_{n=1}^{\infty} A_n$, die nach Lemma 3.15 in $\mathcal{A}$ ist. Dann gilt konvergiert die Funktionenfolge $\mathbf{1}_{A_1\cup\dots\cup A_n} = \mathbf{1}_{A_1} + \dots + \mathbf{1}_{A_n}$ monoton von unten gegen $\mathbf{1}_A$. Also gilt

$$\mu(A) = J^*(\mathbf{1}_A) = \lim_{n\to\infty} J^*(\mathbf{1}_{A_1\cup\dots\cup A_n}) = \lim_{n\to\infty} \sum_{i=1}^{n} J^*(\mathbf{1}_{A_i}) = \sum_{n=1}^{\infty} \mu(A_i)$$

Also ist μ σ-additiv.

Wegen Lemma 3.15 und der Annahme, dass $(X, \mathcal{L}, J)$ Stone'sch ist, ist $\mathcal{A}$ eine σ-Algebra. ∎

Beispiel 3.18. Sei $(X, \mathcal{L}, J)$ ein Stone'scher Integrationsraum und sei f eine messbare Funktion. Dann ist für $a \in \mathbb{R}$ die Menge

$$f^{-1}\big([a,\infty)\big) = \{x \in X \mid f(x) \geq a\}$$

messbar.

Um dies einzusehen, müssen wir $\mathbf{1}_A \in \mathcal{M}$ zeigen, wobei $A := f^{-1}([a,\infty))$. Hierfür sehen wir zunächst, dass mit f auch

$$g := \big((\tfrac{1}{a} \cdot f) \wedge \mathbf{1}_X\big) \vee 0$$

in $\mathcal{M}$ enthalten ist. Hierfür verwenden wir, dass $(X, \mathcal{L}, J)$ Stone'sch ist. Nach Konstruktion ist $g(x) = 1$ genau für $x \in A$ und $0 \leq g(x) < 1$ falls $x \in X \setminus A$. Damit konvergiert die ebenfalls messbare Funktion $g^n = g \cdots g \in \mathcal{M}$ punktweise gegen $\mathbf{1}_A$.

Analog folgt, dass auch die Urbilder beliebiger Intervalle in $\mathcal{A}$ enthalten sind, egal ob das Intervall endlich oder unendlich, offen, abgeschlossen oder halboffen ist.

Schließlich folgt auch, dass die *Niveaumengen*

$$f^{-1}(\{a\}) = f^{-1}\big((-\infty, a]\big) \cap f^{-1}\big([a,\infty)\big)$$

messbar sind.

Sei $(X, \mathcal{L}, J)$ ein Stone'scher Integrationsraum und sei $(X, \mathcal{A}, \mu)$ der zugehörige Maßraum aus Satz 3.17. Die Einschränkung von μ auf

$$\mathcal{A}_0 := \{A \in \mathcal{A} \mid \mu(A) < \infty\}$$

liefert ein endliches Prämaß. Dieser wiederum erzeugt mithilfe von Satz 3.11 einen Integrationsraum $(X, \mathcal{S}(\mathcal{A}_0), J_\mu)$.

Satz 3.19. Falls $(X, \mathcal{L}, J)$ vollständig ist, so ist die Lebesgue-Erweiterung von $(X, \mathcal{S}(\mathcal{A}_0), J_\mu)$ gerade $(X, \mathcal{L}, J)$.

Beweis. Für $A \in \mathcal{A}_0$ gilt nach Definition $J^*(\mathbf{1}_A) < \infty$, also ist $\mathbf{1}_A$ eine endliche $\mathcal{L}$-Levifunktion und daher wegen Vollständigkeit von $(X, \mathcal{L}, J)$ in $\mathcal{L}$ enthalten. Außerdem gilt

$$J_\mu(\mathbf{1}_A) = \mu(A) = J^*(\mathbf{1}_A) = J(\mathbf{1}_A) .$$

Damit gelten $\mathcal{S}(\mathcal{A}_0) \subseteq \mathcal{L}$ und $J_\mu = J|_{\mathcal{S}(\mathcal{A}_0)}$. Wegen Minimalität der Lebesgue-Erweiterung und der Vollständigkeit von $(X, \mathcal{L}, J)$ gilt daher $\overline{\mathcal{S}(\mathcal{A}_0)} \subseteq \mathcal{L}$ und $\bar{J}_\mu = J|_{\overline{\mathcal{S}(\mathcal{A})}}$.

Wir zeigen nun $\mathcal{L} \subseteq \overline{\mathcal{S}(\mathcal{A}_0)}$ und $\bar{J}_\mu|_{\mathcal{L}} = J$. Sei $f \in \mathcal{L}$ und sei $f = f_+ - f_-$ die Aufspaltung in Positiv- und Negativteil. Für $n \in \mathbb{N}$ und $k \in \mathbb{N}_0$ setzen wir

$$A_{k,n} := \left\{ x \in X \;\middle|\; \frac{k}{n} \le f_+(x) < \frac{k+1}{n} \right\} = f_+^{-1}\left(\left[\tfrac{k}{n}, \tfrac{k+1}{n}\right)\right) .$$

Nach Beispiel 3.18 gilt $A_{k,n} \in \mathcal{A}$. Nach Konstruktion gilt zudem $0 \le \frac{k}{n} \cdot \mathbf{1}_{A_{k,n}} \le f_+$ und damit

$$0 \le \frac{k}{n} \cdot \mu(A_{k,n}) \le J(f_+) ,$$

also ist $A_{k,n}$ sogar in $\mathcal{A}_0$, vorausgesetzt, dass $k \ge 1$. Die Funktionenfolge (f_n) definiert durch

$$f_n(x) := \sum_{k=1}^{n^2} \frac{k}{n} \cdot \mathbf{1}_{A_{k,n}}$$

ist dann eine Folge in $\mathcal{S}(\mathcal{A}_0)$ und nach Konstruktion gilt $f_n \nearrow f_+$ und

$$\lim_{n\to\infty} J_\mu(f_n) = \lim_{n\to\infty} J(f_n) \le J(f_+) < \infty .$$

Nach dem Satz von Beppo Levi gilt darum $f_+ \in \overline{\mathcal{S}(\mathcal{A})}$ und

$$\bar{J}_\mu(f_+) = \lim_{n\to\infty} J_\mu(f_n) = \lim_{n\to\infty} J(f_n) = J(f_+).$$

Analog ist $f_- \in \overline{\mathcal{S}(\mathcal{A})}$ und $\bar{J}_\mu(f_-) = J(f_-)$. Damit folgt die gewünschte Aussage. ∎

3.3 Vollständigkeit von Maßen und σ-Endlichkeit

Definition 3.20. Sei $(X, \mathcal{A}, \mu)$ ein Maßraum. Eine (μ-)*Nullmenge* ist eine Menge $N \in \mathcal{A}$ mit $\mu(N) = 0$. Der Maßraum heißt *vollständig*, falls jede Teilmenge $A \subseteq N$ einer Nullmenge $N \in \mathcal{A}$ wieder in $\mathcal{A}$ enthalten ist.

Ist $A \subseteq N$ eine Teilmenge einer Nullmenge mit $A \in \mathcal{A}$, so gilt wegen der Monotonie von Maßen notwendigerweise $\mu(A) = 0$.

Beispiel 3.21. Sei $(X, \mathcal{L}, J)$ ein Stone'scher Integrationsraum und sei $(X, \mathcal{A}, \mu_J)$ der dazugehörige Maßraum. Sei zudem $f \in \mathcal{L}$. Dann gilt $J(|f|) = 0$ genau dann wenn der Träger $A := \{x \in X \mid f(x) \neq 0\}$ eine Nullmenge ist.

($\Longrightarrow$) Nach Beispiel 3.18 ist der Träger (als Komplement der Niveaumenge zur Null) eine messbare Menge. Es gilt nun $n \cdot |f| \wedge \mathbf{1}_X \nearrow \mathbf{1}_A$, also $J^*\big(n \cdot |f| \wedge \mathbf{1}_X\big) \nearrow J^*(\mathbf{1}_A) = \mu_J(A)$. Andererseits ist $J^*\big(n|f| \wedge 1\big) \leq n \cdot J^*(|f|) = 0$. Zusammen folgt $\mu_J(A) = 0$.

($\Longleftarrow$) Wenn A eine Nullmenge ist, so gilt $J^*(n \wedge |f|) \leq J^*(n \cdot \mathbf{1}_A) = n \cdot \mu_J(A) = 0$. Andererseits ist $J(|f|) = \lim_n J^*(n \wedge |f|) = 0$.

Lemma 3.22. Sei $(X, \mathcal{A}, \mu)$ ein Maßraum. Dann existiert es eine eindeutige minimale Erweiterung $(X, \bar{\mathcal{A}}, \bar{\mu})$ zu einem vollständigen Maßraum.

Beweis. Wir setzen

$$\bar{\mathcal{A}} := \{A \subseteq X \mid \exists A_1, A_2 \in \mathcal{A} : A_1 \subseteq A \subseteq A_2 \text{ und } \mu(A_2 \setminus A_1) = 0\}. \tag{3.3.1}$$

Man zeigt, dass $\bar{\mathcal{A}}$ eine σ-Algebra ist. Ein wohldefiniertes Maß $\bar{\mu}$ of $\bar{\mathcal{A}}$ ist dann definiert durch

$$\bar{\mu}(A) := \mu(A_1) = \mu(A_2),$$

für eine beliebige Wahl von Mengen $A_1, A_2 \in \mathcal{A}$ mit $A_1 \subseteq A \subseteq A_2$.

Jede vollständige Erweiterung muss die $(X, \mathcal{A}', \mu')$ muss die σ-Algebra $\bar{\mathcal{A}}$ enthalten, denn für $A \in \bar{\mathcal{A}}$ ist $A \setminus A_1$ eine Nullmenge, also gilt $A \setminus A_1 \in \mathcal{A}'$, und damit auch $A = (A \setminus A_1) \cup A_1 \in \mathcal{A}'$. Die Eindeutigkeit folgt aus der Minimalität. ∎

Lemma 3.23. Sei $(X, \mathcal{L}, J)$ ein vollständiger Stone'scher Integrationsraum. Dann ist der zugehörige Maßraum $(X, \mathcal{A}, \mu)$ vollständig im Sinne von Definition 3.20.

Beweis. Sei N eine Nullmenge und $A \subseteq N$. Dann ist $\mathbf{1}_N$ eine nichtnegative messbare Funktion mit $J(\mathbf{1}_N) = 0 < \infty$, also nach Lemma 2.28 in $\mathcal{L}$ enthalten. Weiter ist $\infty \cdot \mathbf{1}_N$ (als monotoner Grenzwert der Funktionenfolge $f_n := n \cdot \mathbf{1}_N$) eine Levifunktion. Die Indikatorfunktion $\mathbf{1}_A$ unterscheidet sich von der Levifunktion $\infty \cdot \mathbf{1}_N$ nur an deren Unendlichkeitsstellen, ist also nach Vollständigkeit von $(X, \mathcal{L}, J)$ in $\mathcal{L}$ enthalten. Damit ist $A \in \mathcal{A}$, was zu zeigen war. ■

Ein Blick auf den Beweis Approximationssatzes 2.33 zeigt, dass dieser wie folgt verschärft werden kann.

Approximationssatz 3.24. Sei $(X, \mathcal{V}, J)$ ein Integrationsraum und sei $(X, \bar{\mathcal{V}}, \bar{J})$ die minimale Lebesgue-Erweiterung. Dann existiert zu jedem $f \in \bar{\mathcal{V}}$ eine Folge (f_n) von Funktionen in $\mathcal{V}$ mit $\bar{J}(|f - f_n|) \to 0$, und so, dass $f_n(x) \to f(x)$ für alle $x \in X \setminus N$, für eine Nullmenge N.

Beweis. Im Beweis des Approximationssatzes 2.33 wurde zu gegebenem $f \in \bar{\mathcal{V}}$ eine Funktionenfolge (f_n) in $\mathcal{V}$ mit $J(|f - f_n|) \to 0$ konstruiert, die zusätzlich punktweise auf $X \setminus N$ gegen f konvergiert, wobei N die Menge der ∞-Stellen einer Levifunktion F war. Jede solche Menge N ist aber sicherlich eine Nullmenge, da ansonsten $J(F) = \infty$ gelten würde, im Widerspruch dazu, dass F eine Levifunktion ist. ■

Ist $(X, \mathcal{A}, \mu)$ ein Prämaßraum und $\mathcal{A}_0 \subseteq \mathcal{A}$ das Teilsystem der Mengen mit endlichem Maß, so ist die Lebesgue-Vervollständigung des zugehörigen elementaren Integrationsraumes $(X, \mathcal{S}(\mathcal{A}_0), J_\mu)$ ein vollständiger Lebesgue'scher Integrationsraum (siehe Korollar 3.12). Dieser ist aber nicht unbedingt Stone'sch, wir erhalten also nicht unbedingt ein induziertes Maß. Um dies zu erreichen müssen wir die folgende Zusatzannahme machen.

Definition 3.25. Ein Prämaßraum $(X, \mathcal{A}, \mu)$ heißt σ-*endlich*, falls eine Folge $A_1, A_2, \ldots$ von Mengen in $\mathcal{A}$ mit $\mu(A_n) < \infty$ und $\cup_{n=1}^{\infty} A_n = X$ existiert.

Lemma 3.26. Ist der Prämaßraum $(X, \mathcal{A}, \mu)$ σ-endlich, so ist die Lebesgue-Vervollständigung von $(X, \mathcal{S}(\mathcal{A}_0), J_\mu)$ ein Stone'scher Integrationsraum.

Beweis. Ist $A_1, A_2, \ldots$ eine Folge von Mengen in $\mathcal{A}_0$ mit Vereinigung X, so konvergiert die Folge der Funktionen $\mathbf{1}_{A_1 \cup \cdots \cup A_n} \in \mathcal{S}(\mathcal{A}_0)$ punktweise gegen $\mathbf{1}_X$, also ist X eine messbare Menge für das von der Vervollständigung induzierte Maß. ■

Erweiterungssatz für Prämaße 3.27. Für einen σ-endlichen Prämaßraum $(X, \mathcal{A}, \mu)$ sei $(X, \bar{\mathcal{A}}, \bar{\mu})$ das von der Lebesgue-Vervollständigung von $(X, \mathcal{S}(\mathcal{A}_0), J_\mu)$ induzierte Maß. Dann ist $\bar{\mathcal{A}}$ die Vervollständigung der von $\mathcal{A}$ erzeugten σ-Algebra $\sigma(\mathcal{A})$ und $\bar{\mu}$ ist eine Fortsetzung von μ.

Beweis. Für $A \in \mathcal{A}_0$ gilt

$$\bar{\mu}(A) = \bar{J}_\mu(\mathbf{1}_A) = J_\mu(\mathbf{1}_A) = \mu(A)\,,$$

während für A mit $\mu(A) = \infty$ nach wie vor $\bar{\mu}(A) = \infty$ gilt. Also ist $\bar{\mu}$ tatsächlich eine Fortsetzung von μ.

Da $\bar{\mathcal{A}}$ vollständig ist und eine σ-Algebra ist, ist die Vervollständigung von $\sigma(\mathcal{A})$ in $\bar{\mathcal{A}}$ enthalten. Für die umgekehrte Inklusion beweisen wir zunächst die folgenden beiden Hilfsaussagen:

(1) *Sei $A \in \bar{\mathcal{A}}$ eine Menge mit endlichem Maß. Dann existiert für jedes $\varepsilon > 0$ eine Menge $M \in \sigma(\mathcal{A})$ mit $\bar{\mu}(M) < \varepsilon$ und $A \cup M \in \sigma(\mathcal{A})$.*

Zum Beweis wählen wir Levifunktionen F und G mit $\mathbf{1}_A + G = F$ (beachte: $\mathbf{1}_A \in \overline{\mathcal{S}(\mathcal{A})}$). Wir setzen

$$M := \left\{G(x) > \tfrac{1}{2}\right\}$$

Es gilt dann

$$A \cup M = A \cup \left\{G(x) > \tfrac{1}{2}\right\} = \{F(x) > \tfrac{1}{2}\}\,.$$

Um zu sehen, dass beide Superniveaumengen in $\sigma(\mathcal{A})$ enthalten sind, wählen wir Folgen (G_n) und (F_n) in $\mathcal{S}(\mathcal{A}_0)$ mit $G_n \nearrow G$ und $F_n \nearrow F$. Dann gilt

$$M = \left\{G(x) > \tfrac{1}{2}\right\} = \bigcup_{n=1}^{\infty} \left\{G_n(x) > \tfrac{1}{2}\right\}, \qquad \left\{F(x) > \tfrac{1}{2}\right\} = \bigcup_{n=1}^{\infty} \left\{F_n(x) > \tfrac{1}{2}\right\}.$$

Da die Superniveaumengen der Funktionen F_n und G_n in $\mathcal{A}_0$ enthalten sind, sind beide Mengen in $\sigma(\mathcal{A}_0)$ enthalten. Wegen Lemma 2.21 können wir G so wählen, dass $G \geq 0$ und $\bar{J}^*_\mu(G) \leq \varepsilon/2$. Dann gilt $\mathbf{1}_M \leq 2G(x)$, also

$$\mu(M) = \int_X \mathbf{1}_M(x) d\mu(x) \leq 2 \int_X^* G(x) d\mu(x) \leq \varepsilon\,.$$

Dies zeigt die Aussage.

(2) *Für jede Menge $A \in \bar{A}$ eine Nullmenge $N \in \sigma(\mathcal{A})$ existiert mit $A \cup N \in \sigma(\mathcal{A})$.*

Zunächst nehmen wir $\bar{\mu}(A) < \infty$ an und wählen unter Benutzung der obigen Aussage eine Familie (M_n) von Mengen in $\sigma(\mathcal{A})$ mit $A \cup M_n \in \sigma(\mathcal{A})$ und $\bar{\mu}(M_n) \to 0$. Dann ist

$$\bigcap_{n=1}^{\infty} A \cup M_n = A \cup \bigcap_{n=1}^{\infty} M_n \in \sigma(\mathcal{A}) ,$$

wobei $N := \bigcap_{n=1}^{\infty} M_n$ eine Nullmenge ist.

Für beliebiges $A \in \bar{\mathcal{A}}$ sei (A_i) eine Familie von Mengen in $\bar{\mathcal{A}}$ mit $A = \bigcup_{i=1}^{\infty} A_i$ und $\bar{\mu}(A_i) < \infty$. Dann existieren Nullmengen $N_i \in \sigma(\mathcal{A})$ mit $A_i \cup N_i \in \sigma(\mathcal{A})$ und also mit $N := \bigcup_{i=1}^{\infty} N_i$ auch

$$A \cup N = \bigcup_{i=1}^{\infty} A_i \cup N_i \in \sigma(\mathcal{A}) .$$

Wir zeigen schließlich, dass $\bar{\mathcal{A}}$ in der Vervollständigung von $\sigma(\mathcal{A})$ enthalten ist. Sei $A \in \bar{\mathcal{A}}$. Wegen der obigen Hilfsaussage existiert eine Nullmenge $N \in \sigma(\mathcal{A})$ so, dass $A \cup N$ in $\sigma(\mathcal{A})$ enthalten ist. Genauso existiert eine Nullmenge $N' \in \sigma(\mathcal{A})$ mit $(X \setminus A) \cup N' \in \sigma(\mathcal{A})$, also auch

$$X \setminus \big((X \setminus A) \cup N'\big) = A \cap (X \setminus N') \in \sigma(\mathcal{A}) .$$

Damit gilt

$$A_1 := A \cap (X \setminus N') \subseteq A \subseteq A \cup N =: A_2$$

mit $A_1, A_2 \in \sigma(\mathcal{A})$. Damit ist A in der Vervollständigung von $\mathcal{A}$ enthalten. ■

Beispiele 3.28.

(a) Ist $(X, \mathcal{L}, J)$ eine Lebesgue'scher Integrationsraum und ist $A \subseteq X$ messbar, dann ist $(A, \mathcal{L}|_A, J_A)$ mit $\mathcal{L}|_A := \{f|_A \mid f \in \mathcal{L}\}$ und

$$J_A(f) := J(f \cdot \mathbf{1}_A)$$

wieder ein Lebesgue'scher Integrationsraum.

(b) Ist $(X, \mathcal{A}, \mu)$ ein Maßraum und $A \in \mathcal{A}$, dann ist $(A, \mathcal{A} \cap A, \mu_A)$ mit $\mathcal{A} \cap A := \{B \cap A \mid B \in \mathcal{A}\}$ und

$$\mu_A(B) = \mu(B \cap A)$$

wieder ein Maßraum.

Ist A eine abzählbare Vereinigung von Mengen mit endlichem Maß (beispielsweise wenn $(X, \mathcal{A}, \mu)$ σ-endlich ist), so gilt: Ist $(X, \mathcal{L}_\mu, J_\mu)$ der zu $(X, \mathcal{A}, \mu)$ gehörige

Maßraum, so ist $(A, \mathcal{L}_\mu|_A, J_\mu|_A)$ der zu $(A, \mathcal{A} \cap A, \mu_A)$ gehörige Integrationsraum. Mit anderen Worten: Es gilt

$$\int_A f(x) d\mu_A(x) = \int_X f(x) \mathbf{1}_A(x) d\mu(x)$$

für das zugehörige Integral.

Ohne die Voraussetzung an A ist die rechte Seite der obigen Formel im Allgemeinen nicht definiert, da die Funktion $\mathbf{1}_A$ nicht unbedingt $\mathcal{L}_\mu$-messbar sein muss. (Zur Erinnerung: Um das Integral $(X, \mathcal{L}_\mu, J_\mu)$ zu konstruieren, schränken wir uns zunächst auf die Mengen endlichen Maßes ein.)

4 Das Lebesgue-Integral in $\mathbb{R}^d$

Sei λ_d das Prämaß auf dem Mengenring $\mathcal{Q}_d$ der halboffenen Quader in $\mathbb{R}^d$ aus Beispiel 3.7. Satz 3.11 liefert dann ein Elementarintegral $\int_{\mathbb{R}^d}$ auf dem Raum der Treppenfunktionen über $\mathcal{Q}_d$. Wir benutzen die übliche Notation

$$\int_{\mathbb{R}^d} f(x) dx := J(f) \tag{4.0.1}$$

für das zugehörige Integral.

Definition 4.1 (Lebesgue-Integral)**.** Das d-dimensionale *Lebesgue-Integral* ist die Lebesgue-Erweiterung $(\mathbb{R}^d, \mathcal{L}(\mathbb{R}^d), \int_{\mathbb{R}^d})$ des oben beschriebenen Integrationsraumes $(\mathbb{R}^d, \mathcal{S}(\mathcal{Q}_d), \int_{\mathbb{R}^d})$.

Beispiel 4.2. Ist $f : \mathbb{R} \to [0, \infty)$ eine stückweise stetige Funktion, so ist $f \in \mathcal{L}(\mathbb{R}^d)^*$ enthalten, und ihr verallgemeinertes Lebesgue-Integral kann als uneigentliches Riemann-Integral berechnet werden:

$$\int_{\mathbb{R}^n}^* f(x) dx = \lim_{n \to \infty} \int_{-a}^{a} f(x) dx \,.$$

Dies folgt, da f punktweise monoton von unten durch integrierbare Treppenfunktionen approximiert werden kann. Ist das uneigentliche Riemann-Integral endlich, so ist f sogar Lebesgue-integrierbar.

Ohne die Annahme der Positivität der Funktion f ist die Aussage falsch, wie das Beispiel

$$f(x) = \frac{\sin(x)}{x}$$

zeigt. Diese Funktion ist uneigentlich Riemann-integrierbar, aber nicht Lebesgue-integrierbar, weil das verallgemeinerte Integral von $|f|$ unendlich ist.

Da $\mathbb{R}^d$ eine Vereinigung abzählbar vieler Quader ist, ist die Lebesgue-Erweiterung $(X, \mathcal{L}(\mathbb{R}^d), \int_{\mathbb{R}^d})$ ein Stone'scher Integrationsraum. Nach Satz 3.17 erhalten wir also einen Maßraum $(\mathbb{R}^d, \mathcal{A}_d, \lambda_d)$ durch die Definition

$$\lambda_d(A) := \int_{\mathbb{R}^d}^{*} \mathbf{1}_A(x)\,dx\ .$$

ein Maß auf $\mathbb{R}^d$.

Definition 4.3. Das oben beschriebene Maß λ_d heißt *Lebesgue-Maß*.

Der Nachweis der Bewegungsinvarianz des Lebesgue-Maßes gestaltet sich allerdings mit dieser "achsenparallelen Definition" als schwierig. Unten geben wir daher eine andere Konstruktion des Lebesgue-Integrals an, die auf dem ein-dimensionalen Riemann-Integral basiert.

4.1 Das iterierte Riemannintegral

Sei $a > 0$ und $f \in C([-a, a]^d, \mathbb{R})$ eine stetige Funktion. Seien $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_d$ (der i-te Index ausgelassen) so, dass $x_j \in [-a, a]$. Dann ist

$$x_i \longmapsto f(x_1, \ldots, x_i, \ldots, x_d)$$

eine stetige, Funktion von nur einer Variable. Daher existiert das eindimensionale Riemann-Integral über die i-te Variable.

Lemma 4.4. Die resultierende Funktion

$$(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_d) \longmapsto \int_{-a}^{a} f(x_1, \ldots, x_i, \ldots, x_d)\,dx_i.$$

ist stetig.

Beweis. Da zu einer stetigen Funktion f in d Variablen die neue Funktion $\tilde{f}$, die man durch eine beliebige Permutation der Variablen erhält, wieder stetig ist, genügt es, den Fall $i = d$ zu zeigen. Wir schreiben $x = (x', x_d)$ mit $x' \in \mathbb{R}^{d-1}$ und $x_d \in \mathbb{R}$, sowie

$$F(x') := \int_{-a}^{a} f(x', x_d) dx_d, \qquad \forall x' \in \mathbb{R}^{d-1}.$$

Da $[-a, a]^d$ kompakt ist, ist f gleichmäßig stetig. Zu gegebenem $\varepsilon > 0$ existiert also ein $\delta > 0$ so dass $|f(x) - f(y)| \leq \varepsilon/2a$ für alle $x, y \in \mathbb{R}^d$ mit $|x - y| \leq \delta$. Falls daher $|x' - y'| \leq \delta$, so gilt

$$\begin{aligned} |F(x') - F(y')| &= \left| \int_{-a}^{a} f(x', x_d) dx_d - \int_{-a}^{a} f(y', x_d) dx_d \right| \\ &\leq \int_{-a}^{a} \left| f(x', x_d) - f(y', x_d) \right| dx_d \\ &\leq \int_{-a}^{a} \frac{\varepsilon}{2a} dx_d = \varepsilon. \end{aligned}$$

Damit ist F (gleichmäßig) stetig. ∎

Wiederholte Anwendung von Lemma 4.4 erlaubt die folgende Definition:

Definition 4.5. Das *iterierte Riemmann-Integral* einer Funktion $f \in C([-a, a]^d, \mathbb{R})$ ist definiert als die reelle Zahl

$$\int_{[-a,a]^d} f(x) dx := \int_{-a}^{a} \cdots \int_{-a}^{a} f(x_1, \ldots, x_d) dx_1 \cdots dx_d \,. \tag{4.1.1}$$

Die *Maximumsnorm*, gegeben durch

$$\|f\|_\infty := \max_{x \in \mathbb{R}^d} |f(x)| \,,$$

macht $C([-a, a]^d, \mathbb{R})$ zu einem normierten Vektorraum (er istsogar vollständig, also ein Banachraum). Das folgende Lemma sagt, dass das iterierte Riemann-Integral ein beschränktes Funktional bezüglich dieser Norm ist. Es folgt direkt aus der üblichen Standardabschätzung für eindimensionale Riemannintegrale.

Lemma 4.6 (Standardabschätzung)**.** Für alle $f \in C([-a, a]^d, \mathbb{R})$ gilt

$$\left| \int_{[-a,a]^d} f(x) dx \right| \leq (2a)^d \cdot \|f\|_\infty.$$

Hier haben wir die folgende Aussage:

Eine Funktion $f : \mathbb{R}^d \to \mathbb{R}$ heißt *kompakt getragen* bzw. *hat kompakten Träger*, falls es eine kompakte (äquivalent, beschränkte) Menge $K \subseteq \mathbb{R}$ gibt, so dass $f(x) = 0$ gilt für alle $x \in \mathbb{R}^d \setminus K$. Wir schreiben

$$C_c(\mathbb{R}^d) = \{f \in C(\mathbb{R}^d, \mathbb{R}) \mid f \text{ hat kompakten Träger}\}$$

für den Raum der stetigen Funktionen mit kompaktem Träger. Für $a > 0$ ist die Einschränkung $f|_{[-a,a]^d}$ einer Funktion $f \in C_c(\mathbb{R}^d)$ in $C([-a, a]^d, \mathbb{R})$ enthalten, also existiert das iterierte Riemann-Integral dieser Einschränkung.

Definition 4.7. Das *iterierte Riemann-Integral* einer Funktion $f \in C_c(\mathbb{R}^d)$ ist definiert als

$$\int_{\mathbb{R}^d} f(x)dx := \lim_{a\to\infty} \int_{[-a,a]^d} f(x)dx. \tag{4.1.2}$$

Bemerkung 4.8. Sei $f \in C_c(\mathbb{R}^d)$. Es ist leicht zu sehen, dass sich das Integral von $f|_{[-a,a]^d}$ nicht mehr ändert, sobald a so groß ist, dass $f(x) = 0$ für alle x außerhalb von $[-a, a]^d$ (ein solches a existiert, da f kompakten Träger hat). Man muss also in Wirklichkeit keinen Grenzwert bilden, um die rechte Seite von (4.1.2) zu berechnen.

4.2 Beziehung zum Lebesgue-Integral

Um das iterierte Riemann-Integral mit der Lebesgue'schen Integrationstheorie in Verbindung zu bringen, brauchen wir den folgenden

Satz 4.9 (Dini). Sei X ein kompakter Hausdorff-Raum und sei (f_n) eine monoton fallende Folge mit $f_n \searrow 0$. Dann konvergiert (f_n) sogar gleichmässig.

Beweis. Sei $\varepsilon > 0$ gegeben. Setze

$$U_n = \{x \in X \mid |f_n(x)| < \varepsilon\}$$

Wegen Stetigkeit der f_n sind die Mengen U_n offen. Da f_n monoton fällt, ist jeder Punkt von X in einer Mengen U_n enthalten, das System U_n bildet also eine offene Überdeckung von X. Wegen Kompaktheit wird X von endlich vielen der Mengen U_n überdeckt.

Da die Folge (f_n) monoton fallend ist, gilt andererseits $U_n \subseteq U_{n+1}$ für alle $n \in \mathbb{N}$. Da X aber schon von endlich viele Mengen $U_{n_1} \subseteq \cdots \subseteq U_{n_k}$ überdeckt wird, muss in

Wirklichkeit $X = U_{n_k}$ gelten. Das bedeutet aber $\|f_n\|_\infty \leq \varepsilon$ für alle $n \geq n_k$. Da $\varepsilon > 0$ beliebig war, folgt die Aussage. ■

Korollar 4.10. Das iterierte Riemann-Integral

$$J : C_c(\mathbb{R}^d) \longrightarrow \mathbb{R}, \qquad J(f) := \int_{\mathbb{R}^d} f(x)dx$$

ist ein Elementarintegral.

Beweis. Linearität ist offensichtlich, und Positivität folgt aus der Positivität des eindimensionalen Riemann-Integrals, zusammen mit der Beobachtung, dass für positives f die Funktion F von Lemma 4.4 wiederum positiv ist.

Für die Nullstetigkeit von oben sei (f_n) eine Folge in $C_c(\mathbb{R}^d)$ die punktweise gegen monoton von oben gegen Null konvergiert, $f_n \searrow f$. Sei $a > 0$ so groß, dass $f_1(x) = 0$ für alle $x \notin [-a, a]^d$. Da (f_n) monoton von oben gegen Null konvergiert, gilt auch $f_n(x) = 0$ für alle $x \notin [-a, a]^d$. Wenden wir den Satz von Dini auf die eingeschränkte Folge $(f_n|_{[-a,a]^d})$ von stetigen Funktionen auf dem kompakten Hausdorff-Raum $[-a, a]^d$ an, so erhalten wir, dass diese Folge (und damit auch die Folge (f_n)) sogar gleichmäßig gegen Null konvergiert. Also existiert zu $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ so, das $\|f_n\|_\infty \leq \varepsilon$ für alle $n \geq n_0$. Wegen der Standardabschätzung, Lemma 4.6, gilt daher

$$0 \leq \int_{\mathbb{R}^d} f_n(x)dx \leq (2a)^d \cdot \|f_n\|_\infty \leq (2a)^d \cdot \varepsilon$$

sobald $n \geq n_0$. Die Folge der Integrale konvergiert also gegen Null. ■

Lemma 4.11. Seien $\mathcal{K}_d$, $\mathcal{F}_d$ und $\mathcal{O}_d$ die Systeme der kompakten bzw. abgeschlossenen bzw. offenen Teilmengen von $\mathbb{R}^d$. Dann erzeugen $\mathcal{K}_d$, $\mathcal{F}_d$ und $\mathcal{Q}_d$ dieselbe σ-Algebra $\mathcal{B}_d$.

Beweis. Übungsaufgabe. ■

Definition 4.12. Die σ-Algebra $\mathcal{B}_d$ aus Lemma 4.11 heißt die *Borel'sche σ-Algebra* von $\mathbb{R}^d$.

Satz 4.13. Die Lebesgue-Vervollständigung des iterierten Riemann-Integrals stimmt mit dem Lebesgue-Integral aus Definition 4.1 überein. Insbesondere ist die Lebesgue'sche σ-Algebra die Vervollständigung der Borel'schen σ-Algebra.

Beweis. Sei $(\mathbb{R}^d, \overline{C_c(\mathbb{R}^d)}, J)$ die Lebesgue-Vervollständigung des iterierten Riemann-Integrals, sei $(\mathbb{R}^d, \mathcal{A}, \mu)$ der zugehörige Maßraum und sei $\mathcal{A}_0 \subseteq \mathcal{A}$ das Teilsystem der Mengen A mit $\mu(A) < \infty$.

Jede kompakte Menge $A \subseteq \mathbb{R}^d$ tritt als Niveaumenge einer stetigen Funktion auf, beispielsweise von

$$f(x) = \begin{cases} 1 & \text{wenn } x \in A \\ 1 - d(x, A) & \text{wenn } d(x, A) \leq 1 \\ 0 & \text{andernfalls ,} \end{cases} \qquad \text{mit} \qquad d(x, A) = \inf_{y \in A} |x - y| \, .$$

Da Niveaumengen integrierbarer Funktionen messbar sind (Beispiel 3.18), sind damit alle kompakten Mengen in $\mathcal{A}$ enthalten. Da $\mathcal{A}$ eine σ-Algebra ist, ist mit Lemma 4.11 auch $\mathcal{Q}_d \subseteq \mathcal{A}$. Es gilt offenbar sogar $\mathcal{Q}_d \subseteq \mathcal{A}_0$.

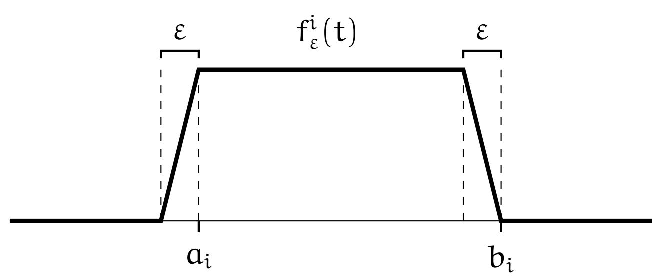
Wir zeigen, dass $\mu(Q) = \lambda_d(Q)$ für alle $Q \in \mathcal{Q}_d$. Hierzu reicht es, die Mengen $Q = [a, b)$ der d-dimensionalen Quader zu betrachten. Zu einem solchen Quader und $\varepsilon > 0$ betrachten wir

$$f_\varepsilon(x) = \prod_{i=1}^{d} f_\varepsilon^i(x_i),$$

wobei $f_i^\varepsilon(t)$ die stückweise affin-lineare Funktion mit den Knoten

$$f_\varepsilon^i(a_i - \varepsilon) = 0, \qquad f_\varepsilon^i(a_i) = 1, \qquad f_\varepsilon^i(b_i - \varepsilon) = 1, \qquad f_\varepsilon^i(b_i) = 0$$

ist.



Das Riemann-Integral von f_ε^i beträgt $b_i - a_i$ (unabhängig von ε) und damit erhalten wir

$$J(f_\varepsilon) = \prod_{i=1}^{d} \int_{-\infty}^{\infty} f_i(t) dt = \prod_{i=1}^{d} (b_i - a_i)$$

für das iterierte Riemann-Integral von f_ε. Andererseits konvergiert f_ε punktweise gegen die Indikatorfunktion von $\mathbf{1}_{[a,b)}$ und damit gilt

$$J(\mathbf{1}_{[a,b)}) = \lim_{\varepsilon\to 0} J(f_\varepsilon) = \prod_{i=1}^{d}(b_i - a_i)$$

nach dem Satz von Lebesgue über die dominierte Konvergenz 2.31 (eine integrierbare Majorante ist hier Koordinatenweise leicht anzugeben). Also gilt

$$\mu\big([a,b)\big) = J(\mathbf{1}_{[a,b)}) = \prod_{i=1}^{d}(b_i - a_i) = \lambda_d\big([a,b)\big)$$

Dies zeigt, dass $(\mathbb{R}^d, \overline{C_c(\mathbb{R}^d)}, J)$ eine vollständige Erweiterung von $(\mathbb{R}^d, \mathcal{L}(\mathbb{R}^d), \int_{\mathbb{R}^d})$ ist.

Es bleibt zu zeigen, dass $\overline{C_c(\mathbb{R}^d)} = \mathcal{L}(\mathbb{R}^d)$ gilt. Wegen der Minimalität der Lebesgue-Erweiterung reicht es hierfür zu zeigen, dass $C_c(\mathbb{R}^d) \subseteq \mathcal{L}(\mathbb{R}^d)$. Sei $f \in C_c(\mathbb{R}^d)$ mit Träger in einem Quader $[a,b)$. Dann kann f punktweise monoton von unten durch in $[a,b)$ getragene Treppenfunktionen in $\mathcal{S}(\mathcal{Q}_d)$ approximiert werden. Jede solche approximierende Folge ist eine Levifolge, da das Integral der Funktion $\|f\|_\infty \cdot \mathbf{1}_{[a,b)}$ eine obere Schranke für die Integrale der approximierenden Funktionen ist. Damit ist $f \in \mathcal{L}(\mathbb{R}^d)$, da $(X, \mathcal{L}(\mathbb{R}^d), \int_{\mathbb{R}^d})$ ein Lebesgue'scher Integrationsraum ist. ∎

Satz 4.14 (Regularität des Lebesguemaßes)**.** Eine Teilmenge $A \subseteq \mathbb{R}^d$ ist Lebesgue-messbar mit $\lambda_d(A) < \infty$ genau dann, wenn zu jedem $\varepsilon > 0$ eine kompakte Menge K und eine offene Menge U existieren mit $K \subseteq A \subseteq U$ sowie $\lambda_d(U) < \infty$ und $\lambda_d(U \backslash K) \leq \varepsilon$.

Insbesondere ist

$$\lambda_d(A) = \sup\left\{\lambda(K) \mid K \subseteq A \text{ kompakt}\right\} = \inf\left\{\lambda(U) \mid U \supseteq A \text{ offen}\right\}.$$

Proof. ($\Longrightarrow$) Sei $A \in \mathcal{A}$ eine Lebesgue-messbare Menge mit $\lambda_d(A) < \infty$, also äquivalent $\mathbf{1}_A \in \mathcal{L}(\mathbb{R}^d)$. Sei $\varepsilon > 0$ mit $\varepsilon < 1$. Da nach Satz 4.13 $\mathcal{L}(\mathbb{R}^d)$ die Lebesgue-Vervollständigung des iterierten Riemann-Integrals auf $C_c(\mathbb{R}^d)$ ist, existieren Levifunktionen $F, G \in C_c(\mathbb{R}^d)^*$ so, dass

$$\mathbf{1}_A + G = F\,,$$

wobei G nach Lemma 2.21 positiv und mit $\int^*_{\mathbb{R}^d} G(x)dx \leq \varepsilon/6$ gewählt werden kann.

Sei (F_n) eine Folge in $C_c(\mathbb{R}^d)$ mit $F_n \nearrow F$. Die Menge

$$U := \{F(x) > 1 - \varepsilon\} = \bigcup_{n=1}^{\infty} \{F_n(x) > 1 - \varepsilon\}$$

ist offen, da die F_n stetig sind. Für $x \in A$ gilt wegen Nichtnegativität von G, dass

$$F(x) = \mathbf{1}_A(x) + G(x) \geq 1 > 1 - \varepsilon \, ,$$

also $x \in U$ und damit $A \subseteq U$. Für $x \in U \setminus A$ gilt

$$1 - \varepsilon < F(x) = \mathbf{1}_A(x) + G(x) = G(x) \, ,$$

also $(1 - \varepsilon) \cdot \mathbf{1}_{U \setminus A} \leq G$ und daher

$$\mu(U \setminus A) = \int^*_{\mathbb{R}^d} \mathbf{1}_{U \setminus A}(x) dx \leq \frac{1}{1 - \varepsilon} \int^*_{\mathbb{R}^d} G(x) dx \leq \frac{1}{1 - \varepsilon} \cdot \frac{\varepsilon}{6} \leq \frac{\varepsilon}{3} \, .$$

Zur Konstruktion von K wählen wir zunächst einen beschränkten, abgeschlossenen Quader Q, der A bis auf eine Menge von Maß $\varepsilon/3$ enthält, also

$$\lambda_d(A \setminus Q) \leq \frac{\varepsilon}{3} \, .$$

Dann ist $Q \setminus A$ messbar mit endlichem Volumen und nach obiger Konstruktion existiert eine offene Menge $V \supseteq (Q \setminus A)$ mit $\lambda_d(U) < \infty$ und $\lambda_d(V \setminus (Q \setminus A)) \leq \varepsilon/3$. Die Menge $K := Q \setminus V$ ist dann kompakt und es gilt

$$K = Q \cap (\mathbb{R}^d \setminus V) \subseteq Q \cap \big(\mathbb{R}^d \setminus (Q \setminus A)\big) = Q \cap A \subseteq A \, .$$

Wegen $Q \setminus K = V$ folgt außerdem

$$\begin{aligned} \lambda_d(A \setminus K) &= \lambda_d(A \setminus Q) + \lambda_d\big((Q \cap A) \setminus K\big) \\ &\leq \frac{\varepsilon}{3} + \lambda_d\big((Q \setminus K) \setminus (Q \setminus A)\big) \\ &\leq \frac{\varepsilon}{3} + \lambda_d\big(V \setminus (Q \setminus A)\big) \leq 2 \cdot \frac{\varepsilon}{3} \, , \end{aligned}$$

also

$$\lambda_d(U \setminus K) = \lambda_d(U \setminus A) + \lambda_d(A \setminus K) \leq \frac{\varepsilon}{3} + 2 \cdot \frac{\varepsilon}{3} = \varepsilon.$$

($\Longleftarrow$) Für $A \subseteq \mathbb{R}^d$ beliebig seien (K_n) und (U_n) Folgen von kompakten bzw. offenen Mengen in $\mathbb{R}^d$ mit $K_n \subseteq A \subseteq U_n$ und $\lambda_d(A_n \setminus K_n) \to 0$. Dann sind

$$A_1 := \bigcup_{n=1}^{\infty} K_n, \qquad A_2 := \bigcap_{n=1}^{\infty} U_n$$

messbar mit $\lambda_d(A_2 \setminus A_1) = 0$ und es gilt $A_1 \subseteq A \subseteq A_2$. Damit ist A wegen Vollständigkeit des Lebesgue-Maßes ebenfalls messbar. Ist zudem $\lambda_d(U_n) < \infty$ für mindestens ein n, so gilt $\lambda_d(A) < \infty$. ■

4.3 Die Transformationsformel

Das Ziel dieses Abschnitts ist der Beweis der fundamentalen

Transformationsformel 4.15. Sei $\Phi : \mathbb{R}^d \to \mathbb{R}^d$ ein C^1-Diffeomorphismus und sei $f \in \mathcal{L}(\mathbb{R}^d)$. Dann gilt

$$\int_{\mathbb{R}^d} f(\Phi(x)) \cdot |\det D\Phi(x)|\, dx = \int_{\mathbb{R}^d} f(x) dx. \tag{4.3.1}$$

Insbesondere ist der Integrand auf der linken Seite in $\mathcal{L}(\mathbb{R}^d)$ enthalten.

Zur Erinnerung: Ein C^1-Diffeomorphismus ist eine stetig differenzierbare und invertierbare Abbildung $\Phi : \mathbb{R}^d \to \mathbb{R}^d$, deren Umkehrfunktion Φ^{-1} ebenfalls stetig differenzierbar ist. Im obigen Satz ist $D\Phi(x) \in \mathrm{Mat}_{d\times d}(\mathbb{R})$ die Jacobimatrix von Φ. Da Φ stetig differenzierbar ist, ist die *Jacobideterminante* $\det D\Phi$ eine stetige Funktion auf $\mathbb{R}^d$. Ihr Produkt mit der stetigen, kompakt getragenen Funktion $f \circ \Phi$ ist dann wieder stetig und kompakt getragen, ihr iteriertes Riemann-Integral also wohldefiniert.

Bemerkung 4.16. Ist der Träger von $f \in C_c(\mathbb{R}^d)$ in einer offenen Menge $V \subseteq \mathbb{R}^d$ enthalten und ist $\Phi : \mathbb{R}^d \supseteq U \to V$ ein Diffeomorphismus, der nur auf einer Teilmenge U definiert ist, dann ist die Funktion $(f \circ \Phi) \cdot \det D\Phi$ eine stetige Funktion auf U und ihre Nullfortsetzung ist eine stetige Funktion mit kompaktem Träger in $\mathbb{R}^d$. Die Transformationsformel (4.3.1) gilt auch in diesem Fall.

Lemma 4.17. Es genügt, die Transformationsformel für alle Funktionen $f \in C_c(\mathbb{R}^d)$ zu beweisen.

Beweis. Sei $F \in C_c(\mathbb{R}^d)^*$ eine endliche Levifunktion und sei (F_n) eine Folge in $C_c(\mathbb{R}^d)$ mit $F_n \nearrow F$. Dann gilt

$$F_n\big(\Phi(x)\big)\big|\det D\Phi(x)\big| \nearrow F\big(\Phi(x)\big)\big|\det D\Phi(x)\big|$$

für alle $x \in \mathbb{R}^d$. Ist die Transformationsformel für Funktionen in $C_c(\mathbb{R}^d)$ bekannt, so gilt

$$\begin{aligned} \int_{\mathbb{R}^d}^* F\big(\Phi(x)\big)\big|\det D\Phi(x)\big|\, dx &= \lim_{n\to\infty} \int_{\mathbb{R}^d} F_n\big(\Phi(x)\big)\big|\det D\Phi(x)\big|\, dx \\ &= \lim_{n\to\infty} \int_{\mathbb{R}^d} F_n(x)\, dx \\ &= \int_{\mathbb{R}^d}^* F(x)\, dx, \end{aligned}$$

also hat $F(\Phi(x))|\det D\Phi(x)|$ endliches verallgemeinertes Integral und ist damit eine

Levifunktion.

Ist nun $f \in \mathcal{L}(\mathbb{R}^d)$ mit einer Darstellung $f + G = F$ durch Levifunktionen, so gilt die Transformationsformel für F und G, und damit auch für f. ■

Lemma 4.18. Die Transformationsformel (4.15) gilt für solche Diffeomorphismen, die durch Permutationen von Koordinaten gegeben sind, d.h.,

$$\Phi(x_1, \dots, x_d) = (x_{\sigma(1)}, \dots, x_{\sigma(d)}) \tag{4.3.2}$$

für eine Permutation $\sigma \in S_d$.

Bemerkung 4.19. Da die Jacobideterminante $\det D\Phi$ in diesem Fall konstant gleich eins ist, sagt das obige Lemma sagt aus, dass das iterierte Integral unabhängig von der Reihenfolge ist, in der die Variablen ausintegriert werden. In anderen Worten, man hat die Identität

$$\int_{-a}^{a} \cdots \int_{-a}^{a} f(x_1, \dots, x_1) dx_{\sigma(d)} \cdots dx_{\sigma(1)} = \int_{-a}^{a} \cdots \int_{-a}^{a} f(x_1, \dots, x_1) dx_1 \cdots dx_d. \tag{4.3.3}$$

für jede Funktion $f \in C([-a, a]^d, \mathbb{R})$. Diese Aussage folgt auch aus dem später zu beweisenden Satz von Fubini.

Beweis. Wegen Lemma 4.17 genügt es, Funktionen f in $C_c(\mathbb{R}^d)$ zu betrachten. In diesem Fall ist die Identität (4.3.3) ist sofort offensichtlich, falls f ein Produkt von eindimensionalen Funktionen ist,

$$f(x_1, \dots, x_d) = f_1(x_1) \cdots f_d(x_d) ,$$

da für solche "Produktfunktionen" gilt

$$\int_{[-a,a]^d} f(x) dx = \int_{-a}^{a} \cdots \int_{-a}^{a} f_1(x_1) \cdots f_d(x_d) dx_1 \cdots dx_d = \prod_{i=1}^{d} \int_{-a}^{a} f_i(x_i) dx_i ,$$

was offenbar invariant unter Vertauschung der Variablen ist. (Hierbei ist a so groß, dass der Träger von allen f_i in $[-a, a]$ enthalten ist.) Aufgrund der Linearität gilt die Aussage des Lemmas weiterhin für Linearkombinationen von solchen Produktfunktionen, also auf den von Produktfunktionen erzeugten Untervektorraum $V \subset C([-a, a]^d, \mathbb{R})$.

Sei

$$\Phi^* : C([-a, a]^d, \mathbb{R}) \longrightarrow C([-a, a]^d, \mathbb{R}), \qquad f \longmapsto f \circ \Phi$$

die Vorverkettungsabbildung mit Φ. Mit dieser Notation ist die Aussage des Lemmas äquivalent zur Gleichheit

$$\int_{[-a,a]^d} \circ \Phi_* = \int_{[-a,a]^d} \tag{4.3.4}$$

linearen Funktionalen auf dem Raum $C([-a,a]^d, \mathbb{R})$. Φ^* ist offenbar linear und beschränkt (sogar norm-erhaltend), da sich das Maximum einer Funktion durch eine Permutation der Argumente nicht ändert. Da $\int_{[-a,a]^d}$ wegen Lemma 4.6 beschränkt ist, ist $\int_{[-a,a]^d} \circ \Phi^*$ als Verkettung von beschränkten linearen Abbildungen ebenfalls beschränkt.

Wir werden nun benutzen, dass der Unterraum V bezüglich der Maximumsnorm dicht in $C([-a,a]^d, \mathbb{R})$ liegt, d.h., jede stetige Funktion $f : [-a,a]^d \to \mathbb{R}$ kann gleichmäßig durch Linearkombinationen von Produktfunktionen approximiert werden. Dies kann man etwas mühsam elementar beweisen, es folgt aber auch direkt aus dem untenstehenden Satz von Stone-Weierstraß aus der Funktionalanalysis, den wir in dieser Vorlesung nicht beweisen wollen (beachte: V ist sogar eine Unteralgebra!).

Aufgrund der Rechnung für Produktfunktionen gilt (4.3.4) auf dem dichten Unterraum V. Es ist nun eine grundlegende Aussage der Funktionalanalysis, dass zwei beschränkte lineare Abbildungen, die auf einem dichten Unterraum übereinstimmen, schon gleich sein müssen. Daher gilt (4.3.4) auf ganz $C([-a,a]^d, \mathbb{R})$, was zu zeigen war. ■

Satz von Stone-Weierstraß 4.20. Sei X ein kompakter Hausdorffraum und sei $V \subseteq C(X, \mathbb{R})$ eine Unteralgebra mit folgenden Eigenschaften:

(i) V ist punktetrennend ist, d.h., zu $x, y \in X$, $x \neq y$, existiert $f \in V$ mit $f(x) \neq f(y)$;

(ii) Zu jedem $x \in X$ existiert ein $f \in V$ mit $f(x) \neq 0$.

Dann ist V dicht bezüglich der Maximumsnorm.

Ein Beweis... findet sich zum Beispiel in [Werner, *Funktionalanalysis*, Satz VIII.4.7] ■

Lemma 4.21. Die Transformationsformel (4.3.1) gilt, falls Φ die Form

$$\Phi(x', x_d) = \big(x', \varphi(x', x_d)\big) \tag{4.3.5}$$

hat, mit einer Funktion $\varphi : \mathbb{R}^d \to \mathbb{R}$.

Im obigen Lemma haben wir Elemente $x \in \mathbb{R}^d$ wieder als $x = (x', x_d)$ mit $x' \in \mathbb{R}^{d-1}$ und $x_d \in \mathbb{R}$ geschrieben.

Beweis. Wir haben

$$D\phi = \begin{pmatrix} 1 & & & 0 \\ & \ddots & & \vdots \\ & & 1 & 0 \\ \frac{\partial\varphi}{\partial x_1} & \dots & \frac{\partial\varphi}{\partial x_{d-1}} & \frac{\partial\varphi}{\partial x_d} \end{pmatrix},$$

also

$$\det D\Phi = \frac{\partial\varphi}{\partial x_d}\,.$$

Sei nun $f \in C_c(\mathbb{R}^d)$ und a so groß, dass sowohl f als auch $f \circ \Phi$ außerhalb von $[-a, a]^d$ verschwinden. Unter Benutzung von Bemerkung 4.19 und der eindimensionalen Substitutionsregel gilt dann

$$\begin{aligned} &\int_{\mathbb{R}^d} (f \circ \Phi) \cdot |\det D\Phi(x)|\, dx \\ &\quad = \int_{-a}^{a} \cdots \int_{-a}^{a} \left(\int_{-a}^{a} f\big(x', \varphi(x', x_d)\big) \cdot \left| \frac{\partial\varphi}{\partial x_d}(x', x_d) \right| dx_d \right) dx_1 \cdots dx_{d-1} \\ &\quad = \int_{-a}^{a} \cdots \int_{-a}^{a} \left(\pm \int_{\varphi(x',-a)}^{\varphi(x',a)} f(x', x_d) dx_d \right) dx_1 \cdots dx_{d-1} \end{aligned}$$

Das Vorzeichen richtet sich danach, ob $\frac{\partial\varphi}{\partial x_d}$ positiv oder negativ ist. Im ersteren Fall ist $\varphi(x', -a) < \varphi(x', a)$ und man hat kein Vorzeichen, während im zweiten Fall $\varphi(x', a) < \varphi(x', -a)$ und ein Vorzeichen auftritt, welches durch Vertauschen der Integrationsgrenzen wieder beseitigt werden kann. Wählt man nun a so groß, dass sowohl der Träger von f als auch der Träger von $f \circ \Phi$ in $[-a, a]^d$ enthalten ist, dann stimmt das innere Integral mit

$$\int_{-a}^{a} f(x', x_d) dx_d$$

überein, und das Resultat folgt. ■

Lemma 4.22. Sei $\Phi : \mathbb{R}^d \to \mathbb{R}^d$ ein Diffeomorphismus und $x_0 \in \mathbb{R}^d$. Nach Permutation der Variablen existiert eine Umgebung U von x_0, so dass

$$\Phi|_U = \Phi_1 \circ \Phi_2|_U\,,$$

wobei Φ_1 die Form (4.3.5) und Φ_2 die Form

$$\Phi_2(x', x_d) = \big(\Phi_2'(x, x_d), x_d\big)$$

hat.

Beweis. Wir schreiben $\Phi(x) = (\Phi'(x), \varphi(x))$ mit Funktionen $\Phi' : \mathbb{R}^d \to \mathbb{R}^{d-1}$ und $\varphi : \mathbb{R}^d \to \mathbb{R}$. Da Φ invertierbar ist, ist der Rang der Matrix $D\Phi'(x_0)$ maximal, also $d-1$. Daher existiert $1 \leq i \leq d$ so, dass

$$\begin{pmatrix} D\Phi'(x_0) \\ e_i^T \end{pmatrix}$$

invertierbar ist, wobei $e_i \in \mathbb{R}^d$ der i-te Einheitsvektor ist. (Ansonsten wäre jeder der Vektoren e_i^T im Erzeugnis der Zeilenvektoren von $D\Phi'(x_0)$ enthalten, was aus Dimensionsgründen nicht möglich ist.) Nach einer Permutation der Variablen können wir annehmen, dass $i = d$ ist. Die obige Matrix ist dann die Jacobimatrix bei $x = x_0$ der Funktion

$$\Phi_2(x) = \big(\Phi'(x), x_d\big)\,.$$

Nach dem Satz über die Umkehrfunktion existiert daher eine offene Umgebung U von x_0, auf der Φ_2 invertierbar ist. Wir setzen $\Phi_1 = \Phi \circ (\Phi_2|_U)^{-1}$ und setzen diese Abbildung fort zu einem Diffeomorphismus auf ganz $\mathbb{R}^d$ (hierzu verkleinern wir U ein wenig, falls nötig). Nach Konstruktion gilt dann

$$\Phi_1(x) = \big(x', \varphi_1(x)\big), \qquad x = (x', x_d),$$

für eine Funktion $\varphi_1 : \mathbb{R}^d \to \mathbb{R}$. Dies schließt den Induktionsschritt ab. ■

Wir können nun die Transformationsformel im allgemeinen Fall beweisen.

Beweis der Transformationsformel 4.15. Wegen Lemma 4.17 genügt es, Funktionen f in $C_c(\mathbb{R}^d)$ zu betrachten.

Der Beweis der Transformationsformel erfolgt per Induktion über die Dimension d. Die Transformationsformel in Dimension $d = 1$ ist einfach die Substitutionsregel für das Riemann-Integral aus der reellen Analysis (bzw. Lemma 4.21). Wir betrachten nun $d \geq 2$ und nehmen an, dass die Transformationsregel in Dimension $d-1$ gilt.

Sei $f \in C_c(\mathbb{R}^d)$. Sei $K \subseteq \mathbb{R}^d$ kompakt so, dass $f \circ \Phi$ außerhalb von K verschwindet. Wir nehmen zunächst an, dass eine offene Menge $U \supset K$ existiert, so dass Φ (evtl. nach Koordinatenpermutation) eine Faktorisierung $\Phi|_U = \Phi_1 \circ \Phi_2|_U$ wie in Lemma 4.22 zulässt. Dann gilt auf U, dass

$$\begin{aligned} f\big(\Phi(x)\big) \cdot |\det D\Phi(x)| &= f(\Phi_1(\Phi_2(x)))\big|\det(D\Phi_1)\big(\Phi_2(x)\big) \cdot \det\big(D\Phi_2(x)\big)\big| \\ &= g\big(\Phi_2(x)\big) \cdot \big|\det\big(D\Phi_2(x)\big)\big|\,, \end{aligned}$$

wobei $g := (f \circ \Phi_1) \cdot |\det D\Phi_1\big|$. Da $\Phi_2 = (\Phi_2'(x', x_d), x_d)$, gilt

$$\det D\Phi_2 = \det \begin{pmatrix} D\Phi_2' & 0 \\ \frac{\partial \Phi_1'}{\partial x_d} & 1 \end{pmatrix} = \det D\Phi_2'\,.$$

Aufgrund der Induktionsvoraussetzung gilt

$$\begin{aligned}\int_{\mathbb{R}^d} f\big(\Phi(x)\big)\cdot|\det D\Phi(x)|\,dx &= \int_{\mathbb{R}^d} g\big(\Phi_2(x)\big)\cdot|\det D\Phi_2(x)|\,dx\\ &= \int_{-\infty}^{\infty}\left(\int_{\mathbb{R}^{d-1}} g\big(\Phi_2'(x',x_d),x_d\big)\big|\det\big(D\Phi_2'(x',x_f)\big)\big|\,dx'\right)dx_d\\ &= \int_{-\infty}^{\infty}\left(\int_{\mathbb{R}^{d-1}} g(x',x_d)\,dx'\right)dx_d\\ &= \int_{\mathbb{R}^d} g(x)\,dx\,.\end{aligned}$$

Wegen Lemma 4.21 gilt die Transformationsformel für Φ_1, also

$$\int_{\mathbb{R}^d} g(x)dx = \int_{\mathbb{R}^d} f\big(\Phi_1(x)\big)|\det D\Phi_1(x)|\,dx = \int_{\mathbb{R}^d} f(x)dx\,.$$

Dies zeigt, dass die Aussage im Spezialfall gilt.

Im Allgemeinen Fall überdecken wir den kompakten Träger von f mit offenen Mengen $U_1,\dots,U_n$, so dass die Einschränkungen $\Phi|_{U_i}$ so faktorisieren wie in Lemma 4.22. Sei $\chi_1,\dots,\chi_n$ eine untergeordnete Partition der Eins für die offene Überdeckung $\Phi(U_1),\dots,\Phi(U_n)$, also eine Ansammlung von stetigen und beschränkten Funktionen auf $\mathbb{R}^d$, so dass $\chi_1+\cdots+\chi_n\equiv 1$ and so, dass der Träger von χ_i in $\Phi(U_i)$ enthalten ist. Aufgrund des Spezialfalles gilt dann

$$\begin{aligned}\int_{\mathbb{R}^d} f\big(\Phi(x)\big)\cdot|\det D\Phi(x)|\,dx &= \sum_{i=1}^{n}\int_{\mathbb{R}^d}\chi_i(x)\cdot f\big(\Phi(x)\big)\cdot|\det D\Phi(x)|\,dx\\ &= \sum_{i=1}^{n}\int_{\mathbb{R}^d}\chi_i\big(\Phi^{-1}(x)\big)\cdot f(x)\,dx = \int_{\mathbb{R}^d} f(x)\,dx\,.\end{aligned}$$

Dies beweist die Transformationsformel im Allgemeinen Fall. ■

Die Transformationsformel für euklidische Bewegungen liefert nun:

Korollar 4.23. Das Lebesguemaß auf $\mathbb{R}^d$ ist bewegungsinvariant.

4.4 Die Kantormenge

Die Kantormenge ist eine spezielle Teilmenge $C \subseteq \mathbb{R}$, die als Schnitt

$$C = \bigcap_{n=0}^{\infty} C_n$$

von abgeschlossenen Teilmengen $C_n \subseteq \mathbb{R}$ definiert ist. Hierbei ist $C_0 = [0,1]$ das Einheitsintervall, und C_1 entsteht aus C_0, indem das mittlere Drittel entfernt wird, also das offene Intervall $U_0 = (\frac{1}{3}, \frac{2}{3})$,

$$C_1 = C_0 \setminus U_0.$$

Die Menge $C_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$ besteht also aus der Vereinigung von zwei abgeschlossenen Intervallen. Die Menge C_2 entsteht dann aus C_1, indem man wieder von jedem der beiden Intervallen das mittlere Drittel entfernt,

$$C_2 = C_1 \setminus U_1 = [0, \tfrac{1}{9}] \cup [\tfrac{2}{9}, \tfrac{1}{3}] \cup [\tfrac{2}{3}, \tfrac{7}{9}] \cup [\tfrac{8}{9}, 1], \qquad \text{mit} \qquad U_1 = (\tfrac{1}{9}, \tfrac{2}{9}) \cup (\tfrac{7}{9}, \tfrac{8}{9}) \,.$$

Dieser Prozess wird induktiv fortgesetzt, um die Menge C_n zu erhalten. Die Menge C_n besteht also aus der disjunkten Vereinigung von 2^n abgeschlossenen Intervallen der Länge $(\frac{1}{3})^n$. Umgekehrt ist die offene Menge $[0,1] \setminus C_n$ die disjunkte Vereinigung der Mengen U_n, wobei jedes U_n aus 2^n offenen Intervallen der Länge $(\frac{1}{3})^{n+1}$ besteht. Wir erhalten

$$\lambda\big([0,1] \setminus C\big) = \lambda\left(\bigcup_{n=1}^{\infty} U_n\right) = \sum_{n=0}^{\infty} \lambda(U_n) = \sum_{n=0}^{\infty} 2^n \cdot \left(\frac{1}{3}\right)^{n+1} = \frac{1}{3} \sum_{n=0}^{\infty} \left(\frac{2}{3}\right)^n = 1,$$

also

$$\lambda(C) = 0 \,,$$

d.h. C ist eine Lebesgue-Nullmenge.

Andererseits ist die Cantormenge gleichmächtig zu $\mathbb{R}$. Dies kann man einsehen, indem man eine Zahl $x \in [0,1]$ in ihrer triadischen Darstellung

$$x = 0, a_1 a_2 a_3 a_4 \cdots = \sum_{n=1}^{\infty} a_n 3^{-n}, \qquad a_i \in \{0, 1, 2\} \,,$$

ausdrückt. Dann ist x in C_1 genau dann, wenn $a_1 \neq 1$, genauer gesagt, wenn x eine triadische Darstellung mit dieser Eigenschaft besitzt (die triadische Darstellung ist, wie alle p-adischen Darstellungen, nicht eindeutig, da beispielsweise $0,022222 = 0,1$). Analog gilt, dass x in C_n enthalten genau dann wenn eine triadische Darstellung für x mit $a_1, \ldots, a_n \in \{0,2\}$ existiert. Im Limes ist x in C genau dann wenn $a_i \in \{0,2\}$, also wenn in einer triadischen Darstellung von x keine Einsen vorkommen.

Damit liefert jede Abbildung $\mathbb{N} \to \{0,2\}$ ein Element der Cantormenge. Die Menge solcher Abbildungen ist überabzählbar (Cantors Diagonalargument), aber man kann C nicht direkt mit dieser Menge identifizieren, da gewisse triadische Darstellungen dieselbe reelle Zahl ausdrücken (periodische Brüche: $0,2222\cdots = 1$). Man erhält aber eine Bijektion

$$C \to [0,1] ,$$

indem man alle Zweien in einer Eins-freien triadischen Entwicklung einer Zahl $x \in C$ durch Einsen austauscht und das Ergebnis als dyadische Entwicklung interpretiert. Wir erhalten also folgendes Ergebnis:

Satz 4.24. Es gibt überabzählbare Nullmengen für das Lebesgue-Maß im $\mathbb{R}^d$.

Bemerkung 4.25. Direkt aus obigem Satz erhält man die Folgerung, dass die Lebesgue-σ-Algebra (also die σ-Algebra $\mathcal{A}$, auf der das Lebesgue-Maß gemäß Definition 4.3 definiert ist) echt größer ist als die Borel-σ-Algebra $\mathcal{B}$, also die vom System der offenen Mengen erzeugte σ-Algebra: $\mathcal{A}$ ist die Vervollständigung von $\mathcal{B}$. Ein Beispiel einer nicht Borel-messbaren Nullmenge in $\mathbb{R}$ ist der Schnitt der Menge V aus dem Beweis von Satz 1.4 mit der Cantormenge C.

Eine Variation C' der Cantormenge erhält man, wenn man in der obigen Konstruktion das mittlere von zwei Fünfteln entfernt. In diesem Fall besteht C_n aus 2^n Intervallen der Länge $(\frac{2}{5})^n$, und U_n besteht aus 2^n Intervallen der Länge $(\frac{1}{5})^{n+1}$. In diesem Fall gilt

$$\lambda(C') = 1 - \sum_{n=0}^{\infty} \lambda(U_n) = 1 - \sum_{n=0}^{\infty} 2^n \cdot \left(\frac{1}{5}\right)^{n+1} = \frac{1}{3} .$$

Diese Menge ist also keine Nullmenge. Allerdings hat C' – genau wie die Cantormenge C – keine inneren Punkte, ist also *nirgendsdicht* und besteht nur aus Rand. Dies kann man beispielsweise einsehen, indem man die 5-adische Darstellung von Zahlen $x \in C'$ betrachtet. Wir erhalten folgende Konsequenz:

Satz 4.26. Es gibt Lebesgue-messbare Mengen ohne innere Punkte mit positivem Maß.

5 Produktmaße und Produktintegrale

5.1 Produktmaße und der Satz von Fubini

Seien $(X, \mathcal{A}, \mu)$ und $(Y, \mathcal{B}, \nu)$ σ-endliche Maßräume.

Lemma 5.1. Das System

$$\mathcal{A} \otimes \mathcal{B} := \left\{ \bigcup_{i=1}^{n} A_i \times B_i \;\middle|\; A_1, \dots, A_n \in \mathcal{A}, B_1, \dots, B_n \in \mathcal{B} \right\}$$

von Teilmengen von $X \times Y$ ist ein unitaler Mengenring und es existiert ein eindeutiges σ-endliches Prämaß $\mu \otimes \nu$ auf $\mathcal{A} \otimes \mathcal{B}$ mit der Eigenschaft, dass

$$(\mu \otimes \nu)(A \times B) = \mu(A) \cdot \nu(B), \qquad A \in \mathcal{A},\ B \in \mathcal{B}\,.$$

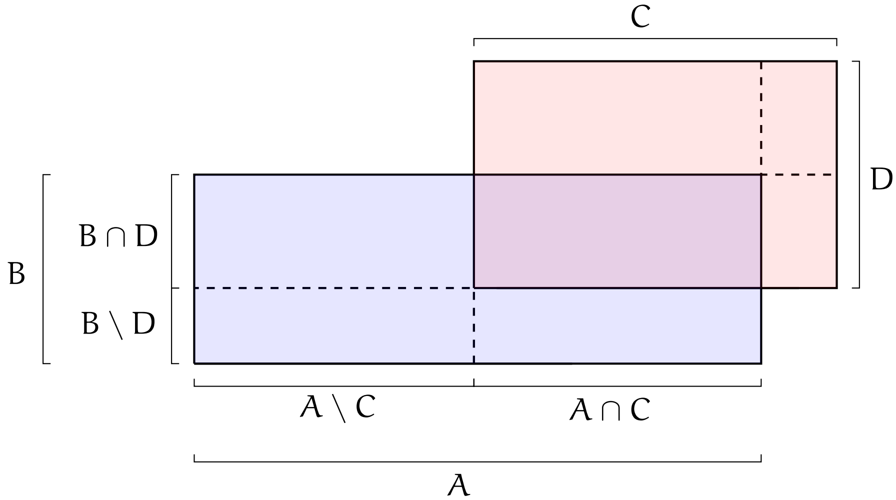
Beweis. $\mathcal{A} \otimes \mathcal{B}$ ist offenbar abgeschlossen unter endlichen Vereinigungen und enthält die Gesamtmenge $X \times Y$. Es bleibt darum zu überprüfen, dass die Differenz von je zwei Mengen aus $\mathcal{A} \otimes \mathcal{B}$ wieder in $\mathcal{A} \otimes \mathcal{B}$ enthalten ist.

Zunächst zeigen wir, dass die Differenz und der Schnitt von zwei elementaren Mengen wieder in $\mathcal{A} \otimes \mathcal{B}$ enthalten ist. Seien also $A, C \in \mathcal{A}$ und $B, D \in \mathcal{B}$. Dann gilt

$$(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D)$$

$$\begin{aligned} \text{und} \qquad (A \times C) \setminus (B \times D) &= \big((A \setminus B) \times (B \setminus D)\big) \\ &\quad \cup \big((A \setminus C) \times (B \cap D)\big) \\ &\quad \cup \big((A \cap C) \times (B \setminus D)\big)\,, \end{aligned}$$

was beides wiederum in $\mathcal{A} \otimes \mathcal{B}$ liegt.



Allgemeiner folgt

$$\left(\bigcup_{i=1}^{n} A_i \times B_i\right) \cap \left(\bigcup_{j=1}^{m} C_j \times D_j\right) = \bigcup_{i=1}^{n}\bigcup_{j=1}^{m} (A_i \cap C_j) \times (C_i \cap D_j)$$

und induktiv sind auch die Schnitte endlich vieler Mengen aus $\mathcal{A} \otimes \mathcal{B}$ wieder in $\mathcal{A} \otimes \mathcal{B}$ enthalten. Die gewünschte Aussage über die Differenz von zwei Mengen folgt nun aus den obigen Beobachtungen zusammen mit der Rechnung

$$\bigcup_{i=1}^{n} A_i \times B_i \Bigg\backslash \bigcup_{j=1}^{m} C_j \times D_j = \bigcup_{i=1}^{n}\bigcap_{j=1}^{m} (A_i \times B_i) \setminus (C_j \times D_j)\ .$$

Wir zeigen nun, dass jede Menge M in $\mathcal{A} \otimes \mathcal{B}$ eine Darstellung als eine endliche Vereinigung paarweise disjunkter Produktmenten geschrieben werden kann. Eine solche Darstellung nennen wir *Zerlegung* von M. Hierzu sei

$$M = \bigcup_{i=1}^{n} A_i \times B_i$$

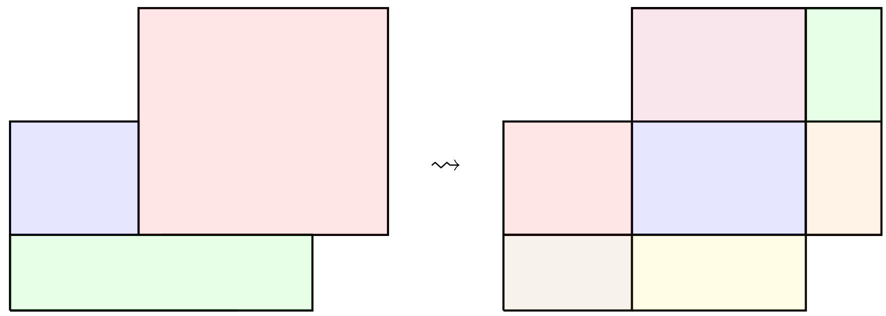
eine beliebige Darstellung von M. Wir konstruieren eine Zerlegung, die diese Darstellung verfeinert: Für eine Teilmenge $I \subseteq \{1, \ldots, n\}$ setzen wir

$$A_I = \bigcap_{i \in I} A_i \cap \bigcap_{i \notin I} X \setminus A_i$$

Dann bilden die Mengen $(A_I)_{I \subseteq \{1,\ldots,\}}$ eine Zerlegung von X und die Analog definierten Mengen $(B_J)_{J \subseteq \{1,\ldots,n\}}$ bilden eine Zerlegung von Y. Es ist dann

$$\bigcup_{i=1}^{N}\bigcup_{I \ni i}\bigcup_{J \ni i} A_I \times B_J = \bigcup_{i=1}^{N}\left(\bigcup_{I \ni i} A_I\right) \times \left(\bigcup_{J \ni i} B_J\right) = \bigcup_{i=1}^{n} A_i \times B_i = M$$

eine Zerlegung von M.



Ist nun $\mu \otimes \nu$ ein Maß mit der gewünschten Eigenschaft, so muss für jede Zerlegung gelten, dass

$$(\mu \otimes \nu)(M) = \sum_{i=1}^{n} (\mu \otimes \nu)(A_i \times B_i) = \sum_{i=1}^{n} \mu(A_i) \cdot \nu(B_i) \, . \tag{5.1.1}$$

Damit gibt es höchstens ein Prämaß mit der gewünschten Eigenschaft.

Wir zeigen nun die Existenz von $\mu \otimes \nu$. Hierzu sei eine Menge $M \in \mathcal{A} \otimes \mathcal{B}$ gegeben. Dann ist für jedes $x \in X$ die Menge

$$M^x = \{ y \in Y \mid (x, y) \in M \}$$

in $\mathcal{B}$ enthalten und die Funktion $x \mapsto \nu(M^x)$ ist eine Treppenfunktion in $\mathcal{S}(\mathcal{A})$. Ist nämlich $M = \bigcup_{i=1}^{n} A_i \times B_i$ eine Zerlegung von M, so ist

$$M^x = \bigcup_{\substack{i=1 \\ x \in A_i}}^{n} B_i$$

eine Zerlegung von M^x (denn ist $y \in B_i \cap B_j$ für Indizes i, j mit $x \in A_i$ und $x \in A_j$, so ist $(x, y) \in (A_i \times B_i) \cap (A_j \times B_j)$ and also $i = j$) und damit einfach

$$\nu(M^\bullet) = \sum_{i=1}^{n} \nu(B_i) \cdot \mathbf{1}_{A_i} \, .$$

Es gilt dann

$$\int_X^* \nu(M^x) d\mu(x) = \sum_{i=1}^{n} \mu(A_i) \cdot \nu(B_i) = \int_Y^* \mu(M^y) d\nu(y) \, ,$$

wobei die Mengen $M^y \subseteq X$, mit $y \in Y$, analog definiert sind. Wir sehen also, dass

$$(\mu \otimes \nu)(M) := \int_X^* \nu(M^x) d\mu(x) = \int_Y^* \mu(M^y) d\nu(y)$$

eine Mengenfunktion mit der Eigenschaft (5.1.1) ist.

Die oben definierte Mengenfunktion $\mu \otimes \nu$ ist offenbar additiv. Zum Nachweis des σ-Additivität sei $M_1, M_2, \ldots$ eine disjunkte Familie von Mengen in $\mathcal{A} \otimes \mathcal{B}$ und sei M deren Vereinigung. Für jedes $x \in X$ ist dann M^x die disjunkte Vereinigung der Mengen M_i^x. Damit gilt unter Benutzung der σ-Additivität von ν

$$\begin{aligned}\sum_{i=1}^{\infty} (\mu \otimes \nu)(M_i) &= \sum_{i=1}^{\infty} \int_X^* \nu(M_i^x) d\mu(x) = \int_X^* \sum_{i=1}^{\infty} \nu(M_i^x) d\mu(x) \\ &= \int_X^* \nu\left(\bigcup_{i=1}^{\infty} M_i^x\right) d\mu(x) = \int_X^* \nu(M_x) d\mu(x) = (\mu \otimes \nu)(M) ,\end{aligned}$$

wobei wir im zweiten Schritt Lemma 2.14 (e) benutzt haben, um Summe und Integral zu vertauschen. (Formal wenden wir die Aussage, dass für monoton steigende Funktionenfolgen der Grenzwert des verallgemeinerten Integrals gleich dem Integral des Grenzwertes ist auf die Folge der Partialsummen an.)

Seien schließlich $X = \bigcup_{i=1}^{\infty} A_i$ und $Y = \bigcup_{j=1}^{\infty} B_i$ Darstellungen von X bzw. Y durch Mengen mit endlichem Maß. Dann gilt

$$X \times Y = \bigcup_{i=1}^{\infty} \bigcup_{j=1}^{\infty} A_i \times B_j ,$$

wobei $(\mu \otimes \nu)(A_i \times B_j) = \mu(A_i) \cdot \nu(B_j) < \infty$. Dies impliziert die σ-Endlichkeit des Maßes $\mu \otimes \nu$. ■

Seien $\mathcal{A}_0 \subseteq \mathcal{A}$ und $\mathcal{B}_0 \subseteq \mathcal{B}$ die Mengenringe der Mengen mit endlichem Maß. Nach dem obigen Lemma ist $\mu \otimes \nu$ ein Elementarintegral auf dem Vektorverband $\mathcal{S}(\mathcal{A}_0 \otimes \mathcal{B}_0)$ von Funktionen auf $X \times Y$. Definieren wir für beliebige Funktionen $f : X \to \mathbb{R}$ und $g : Y \to \mathbb{R}$ die Funktion $f \otimes g$ auf $X \times Y$ durch

$$(f \otimes g)(x, y) = f(x) \cdot g(y)$$

definiert ist, so ist $\mathcal{S}(\mathcal{A}_0 \otimes \mathcal{B}_0)$ einfach das Tensorprodukt der Vektorverbände $\mathcal{S}(\mathcal{A}_0)$ und $\mathcal{S}(\mathcal{B}_0)$,

$$\mathcal{S}(\mathcal{A}_0 \otimes \mathcal{B}_0) = \mathcal{S}(\mathcal{A}_0) \otimes \mathcal{S}(\mathcal{B}_0) .$$

Damit ist $(X \times Y, \mathcal{S}(\mathcal{A}_0 \otimes \mathcal{B}_0), J_{\mu\otimes\nu})$ ein Integrationsraum.

Definition 5.2. Die Lebesgue-Vervollständigung des oben beschriebenen Integrationsraumes heißt das *Produktintegral* von μ und ν, und wir schreiben

$$\bar{J}_{\mu\otimes\nu}(f) =: \int_{X\times Y} f(x, y) d(\mu \otimes \nu)(x, y) .$$

Der zugehörige Maßraum $(X \times Y, \mathcal{A} \hat{\otimes} \mathcal{B}, \mu \otimes \nu)$ ist der *Produktmaßraum* zu μ und ν.

Bemerkung 5.3. Nach dem Erweiterungssatz 3.27 ist $\mathcal{A} \hat{\otimes} \mathcal{B}$ die Vervollständigung der von $\mathcal{A} \otimes \mathcal{B}$ erzeugen σ-Algebra. Insbesondere enthält $\mathcal{A} \hat{\otimes} \mathcal{B}$ beliebige unendliche Vereinigungen von Produktmengen $A \times B$.

Bemerkung 5.4. Sei

$$f = \sum_{i=1}^{n} \lambda_i \cdot \mathbf{1}_{A_i \times B_i} = \sum_{i=1}^{n} \lambda_i \cdot \mathbf{1}_{A_i} \otimes \mathbf{1}_{B_i}$$

eine Funktion in $\mathcal{S}(\mathcal{A}_0 \otimes \mathcal{B}_0)$. Dann ist für jedes $x \in X$ die Funktion $f(x, \cdot)$ in $\mathcal{S}(\mathcal{B}_0)$ enhalten und für jedes $y \in Y$ ist $f(\cdot, y) \in \mathcal{S}(\mathcal{B}_0)$. Es gilt dann

$$\begin{aligned}
\int_{X \times Y} f(x,y) d(\mu \otimes \nu)(x,y) &= \sum_{i=1}^{n} \lambda_i \cdot (\mu \otimes \nu)(A_i \times B_i) \\
&= \sum_{i=1}^{n} \lambda_i \cdot \mu(A_i) \cdot \nu(B_i) \\
&= \sum_{i=1}^{n} \lambda_i \cdot \int_X \mathbf{1}_{A_i}(x) d\mu(x) \cdot \int_Y \mathbf{1}_{B_i}(y) d\nu(y) \\
&= \int_X \int_Y \sum_{i=1}^{n} \lambda_i \cdot \mathbf{1}_{A_i}(x) \cdot \mathbf{1}_{B_i}(y) \, d\mu(x) d\nu(y) \\
&= \int_X \int_Y f(x,y) \, d\nu(y) d\mu(x)
\end{aligned}$$

und analog

$$\int_{X \times Y} f(x,y) d(\mu \otimes \nu)(x,y) = \int_Y \int_X f(x,y) d\mu(x) d\nu(y) \, .$$

Wir wollen die obige Bemerkung auf beliebige $\mu \otimes \nu$-integrierbare Funktionen erweitern.

Lemma 5.5. Sei $F \in \mathcal{S}(\mathcal{A}_0 \otimes \mathcal{B}_0)^*$. Dann gilt für jedes $x \in X$, dass $F(x, \cdot) \in \mathcal{S}(\mathcal{B}_0)^*$ und für jedes $y \in Y$ gilt $F(\cdot, y) \in \mathcal{S}(\mathcal{A}_0)^*$. Weiterhin sind die durch das verallgemeinerte Integral erhaltenen Funktionen

$$x \mapsto \int_Y^* F(x,y) d\nu(y), \qquad \text{und} \qquad y \mapsto \int_X^* F(x,y) d\mu(x)$$

in $\mathcal{S}(\mathcal{A}_0)^*$ bzw. $\mathcal{S}(\mathcal{B}_0)^*$ enthalten und es gilt

$$\int_{X \times Y}^* F(x,y) d(\mu \otimes \nu)(x,y) = \int_X^* \int_Y^* F(x,y) d\nu(y) d\mu(x) = \int_Y^* \int_X^* F(x,y) d\mu(x) d\nu(y) \, .$$

Beweis. Wir zeigen den Fall, dass zunächst bezüglich der y-Variable integiert wird. Die andere Aussage folgt analog.

Sei (F_n) eine Folge in $\mathcal{S}(\mathcal{A}_0 \otimes \mathcal{B}_0)$ mit $F_n \nearrow F$. Dann sind die Funktionen $F_n(x,\cdot)$ in $\mathcal{S}(\mathcal{B}_0)$ enthalten und es gilt natürlich auch $F_n(x,\cdot) \nearrow F(x,\cdot)$, also $F(x,\cdot) \in \mathcal{S}(\mathcal{B}_0)^*$.

Die Funktionen

$$\tilde{F}_n(x) := \int_Y F_n(x,y) d\nu(y)$$

sind Treppenfunktionen über $\mathcal{A}_0$, also Elemente von $\mathcal{S}(\mathcal{A}_0)$. Da $F_n(x,\cdot) \nearrow F(x,\cdot)$ für jedes $x \in X$, gilt nach Definition des verallgemeinerten Integrals, dass

$$\tilde{F}_n(x) = \int_Y F_n(x,y) d\nu(y) \nearrow \int_Y^* F(x,y) d\nu(y) =: \tilde{F}(x)\ .$$

Damit ist $(\tilde{F}_n)$ eine Folge in $\mathcal{S}(\mathcal{A}_0)$, die $\tilde{F}$ monoton von unten approximiert, also gilt $\tilde{F} \in \mathcal{S}(\mathcal{A}_0)$, und

$$\int_X \tilde{F}_n(x) d\mu(x) \nearrow \int_X \tilde{F}(x) d\mu(x) d\mu(x)\ . \tag{5.1.2}$$

Kombinieren wir dies mit der Tatsache, dass die zu beweisende Aussage für Treppenfunktionen gilt (Bemerkung 5.4), so folgt

$$\int_{X\times Y} F_n(x,y) d(\mu\otimes\nu)(x,y) = \int_X \int_Y F_n(x,y) d\nu(y) d\mu(x) \nearrow \int_X \int_Y F(x,y) d\nu(x) d\mu(y)$$

Andererseits konvergiert nach Definition die linke Seite gegen das verallgemeinerte Produktintegral von F. ■

Definition 5.6. Wir sagen, dass zwei auf einem Maßraum $(X, \mathcal{A}, \mu)$ definierte Funktionen f und g *fast überall* (genauer μ-*fast-überall*) übereinstimmen, falls die Menge $\{x \in X \mid f(x) \neq g(x)\}$ eine (μ-)Nullmenge ist.

Satz von Fubini 5.7. Sei f eine bezüglich des Produktmaßes integrierbare Funktion. Dann sind die Funktionen $f(x,\cdot)$ und $f(\cdot,y)$ für fast fast alle $x \in X$ bzw. $y \in Y$ integrierbar und die fast überall definierten Funktionen

$$x \mapsto \int_X f(x,y) d\nu(y), \qquad y \mapsto \int_Y f(x,y) d\mu(x)$$

stimmen bis auf eine Nullmenge mit einer integrierbaren Funktion überein. Es gilt weiterhin

$$\int_{X\times Y} f(x,y) d(\mu\otimes\nu)(x,y) = \int_X \int_Y f(x,y) d\nu(y) d\mu(x) = \int_Y \int_X f(x,y) d\mu(x) d\nu(y)\ .$$

Bemerkung 5.8. Im obigen Satz haben wir die folgende Konvention benutzt: Sei $\tilde{f}$ eine μ-integrierbare Funktion, so dass fast überall

$$\tilde{f}(x) = \int_X f(x,y)d\nu(y)$$

gilt (genauer gesagt, für alle $x \in X \setminus N$, wobei $N \subseteq X$ eine μ-Nullmenge ist, die insbesondere diejenigen Punkte x enthält, für die die Funktion $f(x,\cdot)$ nicht integrierbar und damit die rechte Seite nicht definiert ist). Die Existenz einer solchen Funktion wird vom Satz von Fubini garantiert. Dann ist oben gemeint

$$\int_X \int_Y f(x,y)d\nu(y)d\mu(x) := \int_X \tilde{f}(x)d\mu(x).$$

Diese Definition ist offenbar unabhängig von der Wahl der Fortsetzung $\tilde{f}$, da je zwei Fortsetzungen sich nur auf einer Nullmenge unterscheiden.

Beweis. Sei $f + G = F$ eine Darstellung mit Levi-Funktionen F und G in $\mathcal{S}(\mathcal{A} \otimes \mathcal{B})^*$. Wegen Lemma 5.5 gilt

$$\infty > \int_{X\times Y}^* F(x,y)d(\mu \otimes \nu)(x,y) = \int_X^* \int_Y^* F(x,y)d\nu(y)d\mu(x)\ . \tag{5.1.3}$$

Daher folgt $\mu(N_1) = 0$ mit

$$N_1 := \left\{ x \in X \,\middle|\, \int_Y^* F(x,y)d\nu(y) = \infty \right\}\ ,$$

denn ansonsten wäre das iterierte Integral nicht endlich. Damit ist $F(x,\cdot)$ für fast alle x eine ν-Levifunktion. Analog ist $G(x,\cdot)$ für alle x außerhalb einer Nullmenge N_2 eine ν-Levifunktion. Damit ist

$$f(x,\cdot) + G(x,\cdot) = F(x,\cdot)$$

für alle $x \in X \setminus (N_1 \cup N_2)$ eine Darstellung von $f(x,\cdot)$ durch Levifunktionen, also ist $f(x,\cdot)$ für solche x integrierbar.

Ebenfalls aus (5.1.3) folgt, dass die Funktionen

$$\tilde{F}(x) := \int_Y^* F(x,y)d\nu(y), \qquad \tilde{G}(x) := \int_Y^* G(x,y)d\nu(y)$$

endliches verallgemeinertes Integral haben, also sind beide μ-Levifunktion. Aufgrund der Vollständigkeit des μ-Integrals ist damit

$$\tilde{f} := \tilde{F} \cdot \mathbf{1}_{X\setminus N_1} - \tilde{G} \cdot \mathbf{1}_{X\setminus N_2}$$

integrierbar bezüglich μ, und für alle $x \in X \setminus (N_1 \cup N_2)$ gilt

$$\tilde{f}(x) = \tilde{F}(x) - \tilde{G}(x) = \int_Y f(x,y)d\nu(y) \,.$$

Schließlich gilt

$$\begin{aligned}
\int_{X\times Y} f(x,y)d(\mu\otimes\nu)(x,y) &= \int_{X\times Y}^* F(x,y)d(\mu\otimes\nu)(x,y) - \int_{X\times Y}^* G(x,y)d(\mu\otimes\nu)(x,y) \\
&= \int_X^* \int_Y^* F(x,y)d\nu(y)d\mu(x) - \int_X^* \int_Y^* G(x,y)d\nu(y)d\mu(x) \\
&= \int_X^* \tilde{F}(x)d\mu(x) - \int_X^* \tilde{G}(x)d\mu(x) \\
&= \int_X \tilde{f}(x)d\mu(x) \,,
\end{aligned}$$

wobei wir im zweiten Schitt Lemma 5.5 benutzt haben. ■

Bemerkung 5.9. Induktiv ist natürlich auch für endlich viele Maßräume $(X_i, \mathcal{A}_i, \mu_i)$, $i = 1, \ldots, n$, das n-fache Produktmaß definiert. Da Transpositionen (Zweiervertauschungen) die Gruppe der Permutationen von n Elementen erzeugen, ist das zugehörige iterierte Integral unabhängig von der Integrationsreihenfolge.

Beispiel 5.10. Das Lebesgue-Maß im $\mathbb{R}^d$ ist das d-fache Produkt des ein-dimensionalen Lebesgue-Maßes. Dass das iterierte Riemann-Integral unabhängig von der Integrationsreihenfolge ist, haben wir vorher schon unabhängig nachgewiesen (Lemma 4.18), ist aber natürlich auch eine Konsequenz des Satzes von Fubini.

Beispiel 5.11. Wir betrachten die Funktion $f : [-1,1] \times [-1,1] \to \mathbb{R}$ gegeben durch $f(0,0) = 0$ und

$$f(x,y) = \frac{x^2 - y^2}{(x^2+y^2)^2}, \qquad (x \neq y).$$

Für jedes feste $y \neq 0$ ist $x \mapsto f(x,y)$ eine stetige Funktion auf $[-1,1]$, also inbesondere Lebesgue-integrierbar. Genauer gilt

$$\int_{-1}^1 \int_{-1}^1 f(x,y)dxdy = -\int_{-1}^1 \frac{x}{x^2+y^2}\bigg|_{x=-1}^{x=1} dy = -\int_{-1}^1 \frac{2}{1+y^2}dy = -\pi.$$

Im Gegensatz dazu gilt

$$\int_{-1}^1 \int_{-1}^1 f(x,y)dydx = \int_{-1}^1 \frac{y}{x^2+y^2}\bigg|_{x=-1}^{x=1} dy = \int_{-1}^1 \frac{2}{x^2+1}dy = \pi.$$

Die Integrationsreihenfolge ist also nicht vertauschbar. Das Problem ist hier, dass f nicht bezüglich des Produktmaßes integrierbar ist.

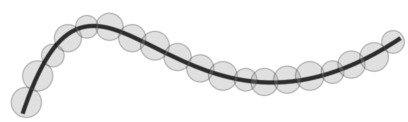
5.2 Die Kofächenformel

Ist $A \subseteq \mathbb{R}^n$ eine Teilmenge, die in einer Hyperebene enthalten ist, so hat diese Menge verschwindendes Lebesgue-Maß, $\lambda_n(A) = 0$. Das unten nun zu definierende Hausdorff-Maß weißt solchen Teilmengen ein möglicherweise nicht triviales Maß zu.

Definition 5.12. Sei $A \subseteq \mathbb{R}^n$ eine beliebige Teilmenge und sei $0 \leq d \leq n$ (nicht notwendigerweise ganzzahlig). Das d-dimensionale *Hausdorff-Maß* von A ist definiert als

$$\mathcal{H}_d(A) := \lim_{\varepsilon \to 0} \inf \left\{ \lambda_d(B^d) \cdot \sum_{i=1}^{\infty} \left(\frac{\operatorname{diam}(D_i)}{2} \right)^d \;\middle|\; \operatorname{diam}(D_i) \leq 2\varepsilon,\ A \subseteq \bigcup_{i=1}^{\infty} D_i \right\}.$$

Hierbei ist $\lambda_d(B^d)$ das Volumen der d-dimensionalen Einheitskugel $B_d \subseteq \mathbb{R}^d$ und das Infimum geht über alle Überdeckungen $(D_i)_{i \in \mathbb{N}}$ von A durch Mengen mit Radius kleiner oder gleich ε.



Das Hausdorff-Maß $\mathcal{H}_d$ ist nicht wirklich ein Maß (es ist nur σ-sub-additiv, nicht σ-additiv), ist dafür aber auf der gesamten Potenzmenge definiert. Um ein Maß zu erhalten, betrachtet man das System

$$\mathcal{C}_{d,n} := \left\{ A \subseteq \mathbb{R}^n \mid \forall B \subseteq \mathbb{R}^n : \mathcal{H}_d(B) = \mathcal{H}_d(B \cap A) + \mathcal{H}_d(B \setminus A) \right\}$$

der *Caratheodory-messbaren Mengen*. Man kann zeigen, dass dies eine σ-Algebra ist, die die Lebesgue-σ-Algebra enthält, und dass die Einschränkung von $\mathcal{H}_d$ auf $\mathcal{C}_{d,n}$ ein (nicht σ-endliches!) Maß auf $\mathbb{R}^n$ ist.

Inesbesondere sind alle Lebesgue-messbaren, beschränkten und kompakt getragenen Funktionen f auf $\mathbb{R}^n$ integrierbar bezüglich des von $\mathcal{H}_d$ erzeugten Integrales.

Definition 5.13. Ist Σ eine abzählbare Vereinigung von Mengen mit endlichem $\mathcal{H}_d$-Maß (beispielsweise eine d-dimensionale Untermannigfaltigkeit, siehe unten), so schreiben wir auch

$$\int_\Sigma f(x)\, d\mathcal{H}_d(x) := \int_{\mathbb{R}^n} f(x)\mathbf{1}_\Sigma(x) d\mathcal{H}_d(x)$$

für das zugehörige Integral über Σ, für geeignete Funktionen f auf Σ (siehe Beispiel 3.28).

Beispiel 5.14. Wir berechnen das Hausdorff-Mass von

$$A = \{(t, 0, \dots, 0) \in \mathbb{R}^n \mid t \in [0, 1]\}\,.$$

Sei $N \in \mathbb{N}$ gegeben und setze $\varepsilon := \frac{1}{2N}$. Wir definieren

$$D_i := B_\varepsilon\big((2i-1)\varepsilon, 0, \dots, 0\big), \qquad i \le N,$$

die Kugel mit Radius ε im $\mathbb{R}^n$, und $D_i = \emptyset$ für $i > N$.

Dann gilt $\operatorname{diam}(D_i) = 2\varepsilon$ und $\lambda_1(B^1) = \lambda_1([-1, 1]) = 2$. Also gilt

$$\lambda_1(B^1) \sum_{i=1}^{\infty} \left(\frac{\operatorname{diam}(D_i)}{2}\right)^1 = 2 \cdot \sum_{i=1}^{N} \underbrace{\varepsilon}_{\frac{1}{2N}} = 1.$$

Da dies unabhängig von ε ist, folgt, dass $\mathcal{H}_1(A) \le 1$. Da umgekehrt die Menge A nicht von Mengen mit Gesamtdurchmesser kleiner als 1 abgedeckt werden kann, folgt insgesamt $\mathcal{H}_1(A) = 1$.

Im Einzelfall ist es schwierig, das Hausdorff-Maß einer allgemeinen Teilmenge zu bestimmen. Die sogenannte *Flächenformel* erlaubt es, das Haudorff-Maß von Untermannigfaltigkeiten mithilfe des gewöhnlichen Lebesgue-Maßes zu berechnen.

Definition 5.15. Sei $\Sigma \subseteq \mathbb{R}^n$ eine Teilmenge. Eine *lokale Parametrisierung* von Σ um einen Punkt $x \in \Sigma$ besteht aus einer offenen Umgebung $V \subseteq \mathbb{R}^n$ von x und einer C^1-Abbildung $\Phi : \mathbb{R}^d \supseteq U \to \mathbb{R}^n$ mit folgenden Eigenschaften:

(i) $\Phi : U \to V \cap \Sigma$ ein Homöomorphismus;

(ii) das Differential $D\Phi(y)$ hat für jeden Punkt $y \in U$ den maximalen Rank d.

Σ heißt C^1-*Untermannigfaltigkeit*, falls um jeden Punkt in Σ eine lokale Parametrisierung existiert.

Untermannigfaltigkeiten entstehen sehr oft als Niveaumengen zu differenzierbaren Funktionen:

Satz vom regulären Wert 5.16. Seien $n \geq m$ in $\mathbb{N}$, $a \in \mathbb{R}$ und sei $u : \mathbb{R}^n \supseteq U \longrightarrow \mathbb{R}^m$ eine C^1-Abbildung so, dass das Differential $Du(x)$ für jeden Punkt $x \in \Sigma := u^{-1}(a)$ den maximalen Rank m besitzt. Dann ist Σ eineUntermannigfaltigkeit des $\mathbb{R}^n$ der Dimension $d = n - m$.

Beispiel 5.17. Die Sphäre vom Radius $r > 0$, gegeben durch

$$S_r^{n-1} := \{x \in \mathbb{R}^n \mid |x| = r\}$$

ist die Niveaumenge zum Wert r der glatten Abbildung

$$u : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}, \qquad u(x) = |x| \, .$$

Im Fall $r = 1$ schreiben auch S^{n-1} statt S_1^{n-1}.

Flächenformel 5.18. Sei Σ eine d-dimensionale Untermannigfaltigkeit von $\mathbb{R}^n$ und sei $\Phi : \mathbb{R}^d \supseteq U \to \mathbb{R}^n$ eine lokale Parametrisierung. Sei $f : \Sigma \supseteq \Phi(U) \to \mathbb{R}$ eine Funktion, die bezüglich des auf Σ eingeschränkten Hausdorff-Maßes integrierbar ist. Dann ist $f \circ \Phi$ Lebesgue-integrierbar und es gilt

$$\int_\Sigma f(x) d\mathcal{H}_d(x) = \int_{\mathbb{R}^d} f\big(\Phi(y)\big) \det\big(D\Phi(y)^T D\Phi(y)\big)^{1/2} dy \, ,$$

Insbesondere gilt die obige Formel für alle stetigen und kompakt getragenen Funktionen $f \in C_c(\Sigma)$.

Hilfreiche Rechnung. Wir geben keinen Beweis der Flächenformel, aber wir rechnen nach, dass die rechte Seite der Flächenformel tatsächlich unabhängig von der Parametrisierung ist. Hierzu seien $\Phi : \mathbb{R}^d \supseteq U \to \mathbb{R}^n$, $\Psi : \mathbb{R}^d \supseteq V \to \mathbb{R}^n$ zwei lokale Parametrisierungen und sei f eine stetige Funktion auf Σ, deren Träger in $\Phi(U) \cap \Psi(V)$

enthalten ist. Es reicht dann zu zeigen, dass

$$\int_{\mathbb{R}^d} f\big(\Phi(y)\big) \det \big(D\Phi(y)^T D\Phi(y)\big)^{1/2} dy = \int_{\mathbb{R}^d} f\big(\Psi(y)\big) \det \big(D\Psi(y)^T D\Psi(y)\big)^{1/2} dy \,.$$

Hierzu bemerken wir, dass $\kappa := \Psi^{-1} \circ \Phi : \tilde{U} \to \tilde{V}$ ein Diffeomorphismus ist, für offene Mengen $\tilde{U} \subseteq U$, $\tilde{V} \subseteq V$ so, dass $f \circ \Phi$ in $\tilde{U}$ und $f \circ \Psi$ in $\tilde{V}$ getragen sind. Somit folgt mit der Transformationsformel 4.15, dass

$$\begin{aligned}
&\int_{\tilde{V}} f\big(\Psi(x)\big) \det \big(D\Psi(x)^T D\Psi(x)\big)^{1/2} dx \\
&\quad = \int_{\tilde{U}} f\Big(\Psi\big(\kappa(x)\big)\Big) \underbrace{\det \Big(D\Psi\big(\kappa(x)\big)^T D\Psi\big(\kappa(x)\big)\Big)^{1/2} \big|\det D\kappa(x)\big|}_{\substack{= |\det D\kappa(x)^T|^{1/2} \det \left(D\Psi(\kappa(x))^T D\Psi(\kappa(x))\right)^{1/2} |\det D\kappa(x)| \\ = \det \left(D\kappa(x)^T D\Psi(\kappa(x))^T D\Psi(\kappa(x)) D\kappa(x)\right)^{1/2} \\ = \det \left(D(\Psi \circ \kappa)(x)^T D(\Psi \circ \kappa)(x)\right)^{1/2} \\ = \det \left(D\Phi(x)^T d\Phi(x)\right)^{1/2}}} dx \\
&\quad = \int_{\tilde{U}} f\big(\Phi(x)\big) \det \big(D\Phi(x)^T D\Phi(x)\big)^{1/2} dx \,. \qquad \blacksquare
\end{aligned}$$

Bemerkung 5.19. Unabhängig von der Definition des Hausdorff-Maßes kann man ein Oberflächenintegral für Funktionen auf Σ durch die rechte Seite der Flächenformel definieren. Der obige Beweis zeigt, dass dieses Integral unabhängig von der Wahl der lokalen Parametrisierung ist (und mit dem Hausdorff-Maß übereinstimmt).

Beispiel 5.20. Wir betrachten $\Sigma = S^1 \subset \mathbb{R}^2$. Hier ist

$$\Phi : (0, 2\pi) \longrightarrow \mathbb{R}^2, \qquad \Phi(t) = \big(\cos(t), \sin(t)\big)$$

eine mögliche Parametrisierung. Das Bild von Φ überdeckt ganz S^1 bis auf den Punkt $(-1, 0)$, eine Nullmenge. Wegen

$$\det \big(D\Phi(t)^T d\Phi(t)\big) = \det \left(\begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix} \begin{pmatrix} -\sin(t) & \cos(t) \end{pmatrix} \right) = 1$$

gilt für alle $\mathcal{H}_1$-integrierbaren Funktionen

$$\int_{S^1} f(x) d\mathcal{H}_1(x) = \int_0^{2\pi} f\big(\cos(t), \sin(t)\big)\, dt \,.$$

Insbesondere gilt für das Hausdorff-Maß von S^1

$$\mathcal{H}_1(S^1) = 2\pi \,.$$

Koflächenformel 5.21. Sei $u : \mathbb{R}^n \to \mathbb{R}$ eine C^1-Abbildung. Dann gilt für jede Lebesgue-integrierbare Funktion $f \in \mathcal{L}(\mathbb{R}^n)$ die Formel

$$\int_{\mathbb{R}^n} f(x)\,|Du(x)|\,dx = \int_{-\infty}^{\infty} \left(\int_{u^{-1}(t)} f(x) d\mathcal{H}_{n-1}(x) \right) dt\,.$$

Bemerkung 5.22. Eine Zahl $t \in \mathbb{R}$ heißt *regulärer Wert* einer Funktion $u : \mathbb{R}^n \to \mathbb{R}$, falls für alle $x \in \mathbb{R}^n$ mit $u(x) = t$ gilt $Du(x) \neq 0$. Nach dem Satz vom regulären Wert ist für reguläre Werte t die Menge $u^{-1}(t)$ eine $(n-1)$-dimensionale Untermannigfaltigkeit des $\mathbb{R}^n$.

Ist u regulär genug (hier braucht man genauer die Regularität C^n), so besagt der *Satz von Sard*, dass die Menge der nicht regulären Werte eine Nullmenge ist. Man kann also in diesem Fall den Integranden der rechten Seite fast überall durch die Flächenformel berechnen.

Beweisskizze. Wir führen den Beweis unter der Zusatzannahme, dass u keine kritischen Punkte hat. Dann sind nach dem Satz vom regulären Wert die Niveauflächen $\Sigma_t := u^{-1}(t)$ jeweils $(n-1)$-dimensionale Untermannigfaltigkeiten.

Fixiere einen Wert t_0 und wähle eine lokale Parametrisierung $\Phi' : \mathbb{R}^{n-1} \supseteq U \to \mathbb{R}^n$ von Σ_{t_0}. Sei dann

$$\Phi : \mathbb{R}^{n-1} \times \mathbb{R} \supseteq U \times (t_0 - \varepsilon, t_0 + \varepsilon) \longrightarrow \mathbb{R}^n$$

die eindeutig definierte Funktion mit den Eigenschaften

$$\frac{\partial \Phi(y,t)}{\partial t} = \frac{Du\big(\Phi(y,t)\big)^T}{\big|Du\big(\Phi(y,t)\big)\big|^2}\,, \qquad \Phi(y,t_0) = \Phi'(y)\,.$$

Mit anderen Worten, $\Phi(y,t)$ ist durch die Lösung einer gewöhnlichen Differentialgleichung gegeben, und die allgemeine Theorie sagt einem, dass $\Phi(y,t)$ eine glatte Abbildung ist (Stichwort: Fluss des Vektorfeldes $Du/|Du|^2$). Für jedes $y \in U$ gilt

$$D\Phi(y,t_0) = \left(D\Phi'(y), \frac{Du\big(\Phi'(y)\big)^T}{\big|Du\big(\Phi'(y)\big)\big|^2} \right)\,.$$

Die Determinante dieser Matrix ist nicht Null, da die Spalten von $D\Phi'(x)$ senkrecht zu $Du(\Phi'(x))^T$ sind (erstere sind tangential, letzterer normal zur Niveaufläche Σ_{t_0}). Damit kann man (nach eventueller Verkleinerung von U) das $\varepsilon > 0$ im Definitionsbereich von Φ so wählen, dass Φ ein Diffeomorphismus auf sein Bild ist.

Weiterhin gilt nach Konstruktion

$$\frac{\partial}{\partial t} u\big(\Phi(y,t)\big) = Du\big(\Phi(y,t)\big) \cdot \frac{\partial \Phi(y,t)}{\partial t} = Du\big(\Phi(y,t)\big) \cdot \frac{Du\big(\Phi(y,t)\big)^T}{\big|Du\big(\Phi(y,t)\big|^2} \equiv 1\,.$$

Wegen $u(\Phi(y,t_0)) = t_0$ folgt damit $u(\Phi(y,t)) = t$. Also ist $\Phi(\cdot,t)$ für jedes feste t eine lokale Parametrisierung von Σ_t und insbesondere sind für jedes t die Spalten von $D_y\Phi(y,t)$ orthogonal zu $Du(\Phi'(y))$.

Wir berechnen die Jacobi-Determinante von Φ. Für einen Punkt (y,t) setzen wir

$$b_n := \frac{Du\big(\Phi(y,t)\big)}{\big|Du\big(\Phi(y,t)\big)\big|}$$

und wählen Vektoren $b_1,\dots,b_{n-1}$, die b_n zu einer Orthonormalbasis des $\mathbb{R}^n$ ergänzen. Da die Spalten von $D_y\Phi(y,t)$ orthogonal zu b_n sind, gilt

$$\frac{\partial\Phi(y,t)}{\partial y_i} = \sum_{j=1}^{d-1} a_{ij}b_j, \qquad i = 1,\dots,n-1\ .$$

für geeignete Koeffizienten a_{ij}. Wir erhalten eine Matrix $A = (a_{ij})$ so, dass $D\Phi(y,t)$ die Matrixdarstellung

$$D\Phi(y,t) \mathrel{\widehat{=}} \begin{pmatrix} A^T & 0 \\ 0 & \big|Du\big(\Phi(y,t)\big)\big|^{-1} \end{pmatrix}$$

bezüglich der Standardbasis $e_1,\dots,e_n$ und der Basis $b_1,\dots,b_n$ besitzt. Damit ist

$$|\det D\Phi(y,t)| = \frac{|\det(A)|}{\big|Du\big(\Phi(y,t)\big)\big|}\ ,$$

wobei

$$\begin{aligned} |\det(A)| = \det(AA^*)^{1/2} &= \det\left(\left\langle \frac{\partial\Phi(y,t)}{\partial y_i}, \frac{\partial\Phi(y,t)}{\partial y_j}\right\rangle\right)^{1/2} \\ &= \det\big(D_y\Phi(y,t)^T D_y\Phi(y,t)\big)^{1/2}\ . \end{aligned}$$

Damit gilt unter Benutzung der Flächenformel 5.18 für Lebesgue-integrierbare Funktionen f mit Träger in $\Phi(U \times (t_0-\varepsilon, t_0+\varepsilon))$, dass

$$\begin{aligned} &\int_{t_0-\varepsilon}^{t_0+\varepsilon} \left(\int_{u^{-1}(t)\cap V} f(x)\, d\mathcal{H}_{n-1}(x)\right) dt \\ &\quad= \int_{t_0-\varepsilon}^{t_0+\varepsilon} \left(\int_U f\big(\Phi(y,t)\big)\cdot \det\big(D_y\Phi(y,t)^T D_y\Phi(y,t)\big)^{1/2} dy\right) dt \\ &\quad= \int_{\mathbb{R}^n} f\big(\Phi(y,a)\big)\cdot \det\big(D_y\Phi(y,t)^T D_y\Phi(y,t)\big)^{1/2} dy\, dt \\ &\quad= \int_{\mathbb{R}^n} f\big(\Phi(y,t)\big)\cdot \big|Du\big(\Phi(y,t)\big)\big| \cdot |\det D\Phi(y,t)|\, d(y,t) \\ &\quad= \int_{\mathbb{R}^n} f(x)\cdot |Du(x)|\, dx \end{aligned}$$

Hier haben wir den Satz von Fubini 5.7 sowie die Transformationsformel 4.3.1 verwendet.

Dies zeigt, dass der Satz für Funktionen gilt, die auf kleinen offenen Umgebungen definiert sind. Für beliebige Funktionen f erfolgt der Beweis wie üblich durch Partition der Eins. ■

5.3 Volumina von Bällen und Sphären

Direkt aus der Koflächenformel folgt das folgende Resultat:

Satz 5.23. Ist $f \in \mathcal{L}(\mathbb{R}^d)$ eine Lebesgue-integrierbare Funktion, so gilt

$$\int_{\mathbb{R}^d} f(x)dx = \int_0^\infty \left(\int_{S_r^{d-1}} f(y) d\mathcal{H}_{d-1}(y) \right) dr = \int_0^\infty \left(\int_{S^{d-1}} f(ry) d\mathcal{H}_{d-1}(y) \right) r^{n-1} dr\,.$$

Beweis. Wir wenden die Koflächenformel auf die Funktion $u(x) = |x|$ aus Beispiel 5.17 an. Hier gilt für $x \neq 0$

$$Du(x) = \frac{x}{|x|}, \qquad \text{also} \qquad |Du(x)| = 1\,.$$

Um die zweite Gleichheit einzusehen, bemerken wir, dass für eine lokale Parametrisierung $\Phi : \mathbb{R}^{d-1} \supset U \to \mathbb{R}^d$ von S^{d-1} die Abbildung $\tilde{\Phi} : \mathbb{R}^{d-1} \supset U \to \mathbb{R}^d$ gegeben durch $\tilde{\Phi}(x) = r\Phi(x)$ eine lokale Parametrisierung von S_r^{d-1} ist, mit $D\tilde{\Phi}(y) = rD\Phi(y)$. Ist nun f eine Funktion mit Träger in $\tilde{\Phi}(U)$, so gilt

$$\begin{aligned}
\int_{S_r^{d-1}} f(x) d\mathcal{H}_{d-1}(x) &= \int_U f\big(\tilde{\Phi}(y)\big) \det\big(D\tilde{\Phi}(y)^T D\tilde{\Phi}(y)\big)^{1/2} dy \\
&= \int_U f\big(r\Phi(y)\big) \underbrace{\det\big(r^2 D\Phi(y)^T D\Phi(y)\big)}_{r^{2(n-1)}\det(D\Phi(y)^T D\Phi(y))}{}^{1/2} dy \\
&= r^{d-1} \int_{S^{d-1}} f(rx) d\mathcal{H}_{d-1}(x)\,.
\end{aligned}$$

Der allgemeine Fall wird wie üblich mit einer Partition der Eins behandelt. ■

Beispiel 5.24. Wir berechnen das Integral der *Gauß'schen Glockenkurve*

$$g(t) = e^{-t^2/2}\,.$$

Zunächst ist die Funktion $f(x_1, x_2) = g(x_1) \cdot g(x_2)$ als Produktfunktion offenbar integrierbar bezüglich des Produktmaßes auf $\mathbb{R}^2$, und es gilt

$$\int_{\mathbb{R}^2} f(x_1, x_2) d(x_1, x_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x_1) \cdot g(x_2) dx_2 dx_1 = \left(\int_{-\infty}^{\infty} g(t) dt \right)^2 .$$

Andererseits hängt der Integrand $f(x_1, x_2) = e^{-x_1^2/2 - x_2^2/2} = e^{-|x|^2/2}$ nur vom Abstand vom Ursprung ab. Mit Satz 1.1.2 gilt daher

$$\int_{\mathbb{R}^2} f(x_1, x_2) d(x_1, x_2) = \int_0^{\infty} \left(\int_{S^1} e^{-|x|^2} d\mathcal{H}_1(x) \right) r dr = 2\pi \cdot \int_0^{\infty} e^{-r^2/2} r\, dr,$$

wobei Beispiel 5.20 benutzt wurde. Die rechte Seite kann wegen Beispiel 4.2 als uneigentliches Riemann-Integral berechnet werden. Wegen $e^{-r^2/2} r = (e^{-r^2})'$ kann diese wiederum mit dem Hauptsatz der Differential- und Integralrechnung berechnet werden und gibt den Wert 1. Insgesamt folgt damit

$$\int_{-\infty}^{\infty} e^{-t^2/2}\, dt = \sqrt{2\pi}\,.$$

Beispiel 5.25. Wir betrachten die Funktion

$$f(x) = f(x_1, \ldots, x_d) = \exp\left(-\frac{1}{2} \sum_{i=1}^{d} x_i^2 \right) .$$

Einerseits ist f eine Produktfunktion, $f(x_1, \ldots, x_d) = \prod_{i=1}^{d} g(x_i)$ mit g die Gauß'sche Glockenkurve aus Beispiel 5.24. Ihr Integral berechnet sich somit als

$$\int_{\mathbb{R}^d} f(x) dx = \left(\int_{-\infty}^{\infty} g(t) dt \right)^d = (2\pi)^{d/2}\,.$$

Andererseits können wir mit Satz 5.23 rechnen

$$\begin{aligned}
\int_{\mathbb{R}^d} f(x) dx &= \int_0^{\infty} \left(\int_{S^{d-1}} e^{-|x|^2/2} dx \right) r^{d-1} dr \\
&= \mathcal{H}_{d-1}(S^{d-1}) \cdot \int_0^{\infty} e^{-r^2/2} r^{d-1} dr \\
&= \mathcal{H}_{d-1}(S^{d-1}) \cdot 2^{\frac{d-2}{2}} \int_0^{\infty} e^{-u} u^{\frac{d-2}{2}} du \\
&= \mathcal{H}_{d-1}(S^{d-1}) \cdot 2^{\frac{d-2}{2}} \cdot \Gamma\left(\tfrac{d}{2}\right) ,
\end{aligned}$$

wobei wir $u = \frac{r^2}{2}$ substituiert haben und $\Gamma(t)$ die Gauß'sche Gammafunktion bezeichnet. Zusammen gilt also

$$\mathcal{H}_{d-1}(S^{d-1}) = \frac{2\pi^{d/2}}{\Gamma\left(\frac{d}{2}\right)} .$$

Insbesondere gilt

$$\mathcal{H}_2(S^2) = 4\pi, \qquad \mathcal{H}_3(S^3) = 2\pi^2, \qquad \mathcal{H}_4(S^4) = \frac{8}{3}\pi^2 .$$

Beispiel 5.26. Wir wenden die Formel von Satz 5.23 auf den Fall an, dass f die Indikatorfunktion des Einheitsballes

$$B^d = \{x \in \mathbb{R}^d \mid |x| \leq 1\}$$

ist. Hier gilt mit Beispiel 5.25

$$\begin{aligned}
\lambda_d(B^d) &= \int_0^1 \left(\int_{S^{d-1}} dy \right) r^{d-1} \, d\mathcal{H}_{d-1}(r) \\
&= \mathcal{H}_{d-1}(S^{d-1}) \cdot \int_0^1 r^{d-1} \, dr \\
&= \frac{\mathcal{H}_{d-1}(S^{d-1})}{d} \\
&= \frac{\pi^{d/2}}{\Gamma\left(\frac{d}{2}+1\right)} .
\end{aligned}$$

Insbesondere gilt also

$$\lambda_2(B^2) = \pi, \qquad \lambda_3(B^3) = \frac{4}{3}\pi, \qquad \lambda_4(B^4) = \frac{1}{2}\pi^2 .$$

5.4 Unendliche Produktmaße

Induktiv ist natürlich auch für endlich viele Maßräume $(X_i, \mathcal{A}_i, \mu_i)$, $i = 1, \ldots, n$, das n-fache Produktmaß definiert. Da Transpositionen (Zweiervertauschungen) die Gruppe der Permutationen von n Elementen erzeugen, ist das zugehörige Produktintegral unabhängig von der Reihenfolge der Integration.

Sei nun $(X_n, \mathcal{A}_n, \mu_n)$, $n \in \mathbb{N}$, eine Familie von Maßräeumen mit $\mu_n(X_n) = 1$ für alle $n \in \mathbb{N}$. Wir setzen

$$X := \prod_{n=1}^{\infty} X_n$$

für den Produktraum und wollen nun ein unendliches Produktmaß definieren.

Lemma 5.27. Das System von Teilmengen

$$\mathcal{A} := \bigotimes_{n=1}^{\infty} \mathcal{A}_n = \left\{ \bigcup_{i=1}^{n} \prod_{n=1}^{\infty} A_n^i \subset X \;\middle|\; A_n^i \in \mathcal{A}_n, \exists n_0 \in \mathbb{N} : \forall n \geq n_0 : A_n^i = X \right\}$$

ist ein Mengenring, und es existiert ein eindeutiges Prämaß $\mu := \bigotimes_{n=1}^{\infty} \mu_n$ auf diesem Mengenring mit der Eigenschaft, dass

$$\mu\left(\prod_{n=1}^{\infty} A_n\right) = \prod_{n=1}^{\infty} \mu_n(A_n)\,. \tag{5.4.1}$$

Da für $A = \prod_{n=1}^{\infty} A_n \in \mathcal{A}$ tatsächlich nur endlich viele Faktoren A_n ungleich X_n sind, sind nur endlich viele Faktoren des Produktes auf der rechten Seite von (5.4.1) ungleich Eins, also ist das Produkt tatsächlich ein endliches Product.

Wir können die endlichen Tensorprodukte $\mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_n$ mit Teilmengen von $\mathcal{A}$ identifizieren, indem wir Mengen $M \subseteq \prod_{i=1}^{n} X_i$ mit den Mengen $M \times \prod_{i=n+1}^{\infty} X_i \subseteq \prod_{i=1}^{\infty} X_i$ identifizieren. Das oben definierte Mengensystem $\mathcal{A}$ ist dann die aufsteigende Vereinigung

$$\mathcal{A} = \prod_{n=1}^{\infty} \mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_n$$

und die oben definierte Mengenfunktion μ ist dann die eindeutige Mengenfunktion, deren Einschränkung auf $\mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_n$ mit dem Produktmaß $\mu_1 \otimes \cdots \otimes \mu_n$ übereinstimmt. Als Vereinigung von unitalen Mengenringen ist damit $\mathcal{A}$ wieder ein unitaler Mengenring, und μ ist ein additives Maß. Die σ-Additivität von μ ist allerdings nicht ohne weiteres klar, da X im Allgemeinen als abzählbare Vereinigung X von Mengen geschrieben kann, die nicht alle in einem endlichen Teilbereich $\mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_n$ der Filtrierung enthalten sind, wie das folgende Beispiel zeigt.

Beispiel 5.28. Wir betrachten die Maßräume $X_n = \{0, 1\}$ mit der gesamten Potenzmenge $\mathcal{A}_n = \mathcal{P}(\{0, 1\})$ als σ-Algebra und beliebigen Maßen μ_n. Dann ist $X = \{0, 1\}^{\mathbb{N}}$ und wir setzen Induktiv

$$M_1 = \{0\} \times \prod_{n=2}^{\infty} X_n \ ,$$
$$M_2 = \{(1,0)\} \times \prod_{n=3}^{\infty} X_n$$
$$M_3 = \{(1,1,0)\} \times \prod_{n=2}^{\infty} X_n$$
$$\dots$$

Dann ist X ist die disjunkte Vereinigung der Mengen M_n und der ein-elementigen Menge

$$M_\infty = \{(1,1,1,1,\dots)\}$$

und es gibt kein $n_0 \in \mathbb{N}$ so, dass alle M_n in $\mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_{n_0}$ enthalten sind.

Beweis der σ-Additivität. Wegen der Additivität von μ reicht es, die σ-Additivität für Zerlegungen der Gesamtmenge nachzuweisen, denn ist $(M_n)_{n\in\mathbb{N}}$ eine beliebige abzählbare Familie von paarweise disjunkten Mengen und M_0 das Komplement der Vereinigung dieser Mengen, so ist $(M_n)_{n\in\mathbb{N}_0}$ eine Zerlegung der Gesamtmenge.

Sei also $(M_n)_{n\in\mathbb{N}}$ eine disjunkte Folge von Menge in $\mathcal{A}$ mit Vereinigung X. Wegen der Additivität und der Monotonie von μ gilt dann

$$\sum_{i=1}^{\infty} \mu(M_i) = \lim_{n\to\infty} \sum_{i=1}^{n} \mu(M_i) = \lim_{n\to\infty} \mu\left(\bigcup_{i=1}^{n} M_i\right) \le \mu\left(\bigcup_{i=1}^{\infty} M_i\right) = 1 \ .$$

Um den Beweis der umgekehrten Ungleichung zu erleichtern, bemerken wir zunächst, dass jedes der M_n in einem der endlichen Produktringe $\mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_m$ enthalten ist. Wie im Beweis von Lemma 5.1 gesehen, kann jede solche Menge als eine endliche disjunkte Zerlegung von Produktmengen geschrieben werden. Wegen der Additivität von μ können wir darum annehmen, dass jedes M_n eine Produktmenge

$$M_n = \prod_{i=1}^{\infty} M_{n,i}, \qquad M_{n,i} \in X_i$$

ist, wobei $M_{n,i} \neq X_i$ nur für endlich viele i. Es gilt dann

$$\begin{aligned}
\sum_{n=1}^{\infty} \mu_n(M_n) &= \sum_{n=1}^{\infty} \int_{X_1} \mathbf{1}_{M_{n,1}}(x_1) \cdot \left(\bigotimes_{i=2}^{\infty} \mu_i\right) \left(\prod_{i=2}^{\infty} M_{n,i}\right) d\mu_1(x_1) \\
&= \int_{X_1} \sum_{n=1}^{\infty} \mathbf{1}_{M_{n,1}}(x_1) \cdot \left(\bigotimes_{i=2}^{\infty} \mu_i\right) \left(\prod_{i=2}^{\infty} M_{n,i}\right) d\mu_1(x_1) \ ,
\end{aligned}$$

wobei wir im zweiten Schritt das Lemma von Lebesgue-Fatou im Maßraum benutzt haben.

Nehmen wir nun an, dass $\sum_{n=1}^{\infty} \mu_n(M_n) <$ gilt, so muss ein x_1 existieren, so dass der Integrand

$$\sum_{n=1}^{\infty} \mathbf{1}_{M_{n,1}}(x_1) \cdot \left(\bigotimes_{i=2}^{\infty} \mu_i\right) \left(\prod_{i=2}^{\infty} M_{n,i}\right) < 1 \ .$$

Schreiben wir die linke Seite der obigen Gleichung als

$$\begin{aligned}
&\sum_{n=1}^{\infty} \mathbf{1}_{M_{n,1}}(x_1) \cdot \int_{X_2} \mathbf{1}_{M_{n,2}}(x_2) \cdot \left(\bigotimes_{i=3}^{\infty} \mu_i\right) \left(\prod_{i=3}^{\infty} M_{n,i}\right) d\mu(x_2) \\
&= \int_{X_2} \sum_{n=1}^{\infty} \mathbf{1}_{M_{n,1}}(x_1) \cdot \mathbf{1}_{M_{n,2}}(x_2) \cdot \left(\bigotimes_{i=3}^{\infty} \mu_i\right) \left(\prod_{i=3}^{\infty} M_{n,i}\right) d\mu(x_2)
\end{aligned}$$

so sehen wir, dass der Integrand wiederum an mindestens einer Stelle x_2 echt kleiner als Eins sein muss. Induktiv erhalten wir so ein Element $x = (x_n)_{n \in \mathbb{N}} \in X$ mit der Eigenschaft, dass für jedes $m \in \mathbb{N}$ die Beziehung

$$\sum_{n=1}^{\infty} \mathbf{1}_{M_{n,1}}(x_1) \cdots \mathbf{1}_{M_{n,m}}(x_m) \cdot \left(\bigotimes_{i=m+1}^{\infty} \mu_i\right) \left(\prod_{i=m+1}^{\infty} M_{n,i}\right) < 1 \ . \tag{5.4.2}$$

gilt. Andererseits ist x aber in einer der Mengen der Zerlegung enthalten, sagen wir in M_{n_0}, und für diese Menge gilt $M_{n_0,i} = X_i$ für alle $i \geq m_0$. Andererseits gilt offenbar

$$1 = \mathbf{1}_{M_{n_0}}(x) = \mathbf{1}_{M_{n_0,1}}(x_1) \cdots \mathbf{1}_{M_{n_0,m_0}}(x_{m_0})$$

und

$$\left(\bigotimes_{i=m_0+1}^{\infty} \mu_i\right) \left(\prod_{i=m_0+1}^{\infty} M_{n_0,i}\right) = \prod_{i=m_0+1}^{\infty} \mu_i(M_{n_0,i}) = \prod_{i=m_0+1}^{\infty} \mu_i(X_i) = 1 \ .$$

Damit ist einer der nichtnegativen Summanden der Summe (5.4.2) gleich Eins, im Widerspruch dazu, dass die Gesamtsumme kleiner ist als Eins. ■

Definition 5.29. Die eindeutige Erweiterung des oben definierten Prämaßes auf die Vervollständigung der von $\mathcal{A}$ erzeugten σ-Algebra heißt das *unendliche Produktmaß*.

Beispiel 5.30 (Unendlicher Münzwurf)**.** Wir betrachten das unendliche Produkt der Maßräume $X_n = \{0, 1\}$, mit dem durch

$$\mu_n(\{0\}) = p_n, \quad \mu_n(\{1\}) = 1 - p_n, \qquad 0 \leq p_n \leq 1$$

gegebenen Maß. Sei $\mathbf{0} = (0, 0, 0, \dots) \in X$. Dann ist die einelementige Menge $\{\mathbf{0}\}$ der absteigende Durchschnitt der Mengen

$$M_n = \underbrace{\{0, \dots, 0\}}_{n} \times \prod_{i=n+1}^{\infty} X_n \,.$$

und damit gilt für das unendliche Produktmaß

$$\mu(\{\mathbf{0}\}) = \mu\left(\bigcap_{n=1}^{\infty} M_n\right) = \lim_{N\to\infty} \mu(M_N) = \lim_{N\to\infty} \prod_{n=1}^{N} p_n \,.$$

Nach einschlägigen Sätzen über unendliche Produkte ist dieser Limes genau dann ungleich Null, falls p_n niemals Null ist und zudem die Summe der Folge $1 - p_n$ konvergiert,

$$\sum_{n=1}^{\infty} (1 - p_n) < \infty \,.$$

Insbesondere gilt: Existiert ein $\varepsilon > 0$ so, dass $\varepsilon \leq p_n \leq 1 - \varepsilon$ für alle $n \in \mathbb{N}$ gilt, so ist $\{x\}$ eine Nullmenge. Ähnlich überlegt man sich, dass auch jede andere einelementige Menge eine μ-Nullmenge ist, und damit auch jede abzählbare Teilmenge von X.

Verwendet man im Rahmen der Wahrscheinlichkeitstheorie den Produktmaßraum zur Beschreibung des Zufallsexperiments "unendlicher Münzwurf" (hier gilt $p_n = 1/2$, so erhält man die Aussage: *Wirft man bis in alle Ewigkeit wiederholt eine Münze, so ist die Wahrscheinlichkeit, stets nur Zahl zu werfen, gleich Null.*

6 Der Satz von Radon-Nikodym

6.1 Signierte Maße

Definition 6.1. Sei X eine Menge und sei $\mathcal{A}$ eine σ-Algebra von Teilmengen von X. Ein *signiertes Maß* ist eine σ-additive Mengenfunktion

$$\mu : \mathcal{A} \longrightarrow \overline{\mathbb{R}}$$

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

Hierbei heißt μ σ-additiv, wenn für jede Folge $A_1, A_2, \dots$ von paarweise disjunkten Teilmengen von X gilt, dass

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

wobei die Reihe in $\overline{\mathbb{R}}$ unbedingt konvergiert.

Hierbei bedeutet unbedingte Konvergenz einer unendlichen Summe, dass die Folge der Partialsummen einer beliebigen Umordnung der Summanden konvergiert, und zwar gegen denselben Grenzwert. Dies ist in diesem Fall nicht gleichbedeutend mit absoluter Konvergenz, da die Summanden die Werte $\pm\infty$ annehmen können.

Insbesondere ist Teil dieser Definition, dass es keine zwei disjunkten Mengen $A, B \in \mathcal{A}$ mit $\mu(A) = \infty$ und $\mu(B) = -\infty$ geben kann, denn für solche Mengen müsste das Maß von $A \cup B$ durch den undefinierten Ausdruck $\mu(A) + \mu(B) = \infty - \infty$ gegeben sein.

Bemerkung 6.2. Wegen der Forderung $\mu(\emptyset) = 0$ impliziert die σ-Additivität die Additivität von μ: Es können einfach alle bis auf endlich viele Mengen leer gewählt werden.

Lemma 6.3. Sei μ ein signiertes Maß auf einer σ-Algebra von Teilmengen von X. Dann gilt:

(1) Sind $A, B \in \mathcal{A}$ mit $A \subseteq B$ und $|\mu(B)| < \infty$, so gilt auch $\mu(A) < \infty$.

(2) Es gibt keine zwei Teilmengen $A, B \in \mathcal{A}$ mit $\mu(A) = \infty$ und $\mu(B) = -\infty$.

(3) Ist $A_1 \supseteq A_2 \supseteq A_3 \supseteq \ldots$ eine absteigende Folge von Mengen mit $|\mu(A_1)| < \infty$, so gilt

$$\mu\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n)\,.$$

Beweis. (1) Seien $A, B \in \mathcal{A}$ mit $A \subseteq B$ und $|\mu(B)| < \infty$. Dann gilt

$$\mu(B) = \mu(A) + \mu(B \setminus A)\,.$$

Da die linke Seite endlich ist, muss auch die rechte Seite endlich sein; insbesondere muss $\mu(A)$ endlich sein.

(2) Angenommen, es gäbe solche Mengen A und B. Wir wissen bereits, dass dann A und B nicht disjunkt sein können. Nun sind aber sind A und $B \setminus A$ disjunkt, also muss $\mu(B \setminus A)$ endlich sein, und genauso $\mu(A \setminus B)$. Dies führt zu dem Widerspruch

$$\infty = \mu(A) = \mu(A) + \mu(B \setminus A) = \mu(A \cup B) = \mu(B) + \mu(A \setminus B) = \mu(B) = -\infty\,.$$

(3) Sei $S = \bigcap_{n=1}^{\infty} A_n$. Für ein beliebiges $m \in \mathbb{N}$ ist

$$A_m = S \cup \bigcup_{n=m}^{\infty} (A_n \setminus A_{n+1})$$

eine disjunkte Zerlegung von A_m. Damit gilt für alle $m \in \mathbb{N}$, dass

$$\mu(A_m) = \mu(S) + \sum_{n=m}^{\infty} \mu(A_n \setminus A_{n+1})$$

Wegen (1) haben A_n, $A_n \setminus A_{n+1}$ und S alle endliches Maß. Also ist die Reihe absolut konvergent, und es folgt $\sum_{n=m}^{\infty} \mu(A_n \setminus A_{n+1}) \to 0$. ■

Definition 6.4. Sei $(X, \mathcal{A}, \mu)$ ein signierter Maßraum. Eine disjunkte Zerlegung $X = X^+ \cup X^-$ heißt *Hahn-Zerlegung* von X, wenn für alle $A \in \mathcal{A}$ die Beziehungen

$$\mu(A \cap X^+) \geq 0 \qquad \text{und} \qquad \mu(A \cap X^-) \leq 0$$

gelten. X^+ und X^- heißen *Positivitätsbereich* bzw. *Negativitätsbereich* von μ.

Beispiel 6.5. Ist $(X, \mathcal{A}, \mu)$ ein (positiver) Maßraum. Für eine integrierbare Funktion g ist

$$\mu_g(A) := \int_A g(x) d\mu(x)$$

endliches Maß auf X und $X^+ = \{f(x) \geq 0\}$, $X^- = \{f(x) < 0\}$ ist eine Hahn'sche Zerlegung von X.

Definition 6.6. Sei $(X, \mathcal{A}, \mu)$ ein signierter Maßraum. Eine Menge $T \in \mathcal{A}$ heißt *Träger* von μ, wenn $\mu(A) = 0$ für alle $A \in \mathcal{A}$ mit $A \cap T = \emptyset$ gilt.

Zerlegungssatz für signierte Maße 6.7. Für jedem signierten Maßraum $(X, \mathcal{A}, \mu)$ existiert eine Hahn-Zerlegung. Sie ist bis auf Nullmengen eindeutig.

Die durch

$$\mu^+(A) := \sup\{\mu(B) \mid B \in \mathcal{A}, B \subseteq A\},$$
$$\mu^-(A) := \sup\{-\mu(B) \mid B \in \mathcal{A}, B \subseteq A\}$$

definierten Mengenfunktionen sind (positive) Maße auf $\mathcal{A}$ mit

$$\mu = \mu^+ - \mu^- .$$

Ist $X = X^+ \cup X^-$ eine Hahn-sche Zerlegung, so gilt

$$\mu^+(A) = \mu(A \cap X^+), \qquad \mu^-(A) = \mu(A \cap X^-) . \tag{6.1.1}$$

Insbesondere haben μ^+ und μ^- Trägermengen X^+ bzw. X^-.

Hierbei ist eine μ-Nullmenge eine Menge $N \in \mathcal{A}$ so, dass für alle messbaren Mengen $A \subseteq N$ gilt $\mu(A) = 0$ (nicht nur für A selber).

Definition 6.8. Die obige Zerlegung $\mu = \mu^+ - \mu^-$ heißt *Jordan-Zerlegung* von μ.

Das Maß

$$|\mu| := \mu_+ + \mu_-$$

heißt *Betrag* von μ und die Zahl

$$\|\mu\| := |\mu|(X)$$

heißt Totalvariation von μ.

Beispiel 6.9. Ist H eine Funktion mit beschränkter Variation, so hatten wir in Lemma 6.31 gesehen, dass sich H als Differenz $H = F - G$ von zwei monoton wachsenden Funktionen schreiben lässt. Für das zugehörige signierte Lebesgue-Stieltjes-Maß μ_H ist dann $\mu_H = \mu_F - \mu_G$ die Jordanzerlegung.

Beweis des Zerlegungssatzes. Ohne Beschränkung der Allgemeinheit sei $-\infty < \mu(A) \leq \infty$, ansonsten ersetzen wir μ durch $-\mu$ (siehe Lemma 6.3 (3)). Wir definieren μ^+ und μ^- durch die angegebenen Formeln und setzen

$$\mathcal{A}_- := \left\{A \in \mathcal{A} \mid \mu^+(A) = 0\right\} = \left\{A \in \mathcal{A} \mid \forall B \subseteq A : \mu(B) \leq 0\right\}.$$

Es ist dann nicht schwer zu sehen, dass $\mathcal{A}_-$ ein Mengenring ist. Tatsächlich ist $\mathcal{A}_-$ sogar ein σ-*Ring*, was bedeutet, dass $\mathcal{A}_-$ abgeschlossen unter abzählbaren Vereinigungen ist. Die Einschränkung von $-\mu$ auf $\mathcal{A}_-$ ist dann positiv und nach wie vor σ-additiv, also ein Prämaß.

Es sei

$$-a := \inf\left\{\mu(A) \mid A \in \mathcal{A}_-\right\} > -\infty.$$

Wir wählen eine Folge von Mengen $A_n \in \mathcal{A}_-$ mit $\mu(A_n) \to -a$ und setzen

$$X^- := \bigcup_{n=1}^{\infty} A_n.$$

Weil $\mathcal{A}_-$ ein σ-Ring ist, gilt dann $X^- \in \mathcal{A}_-$. Da die Einschränkung von $-\mu$ auf $\mathcal{A}_-$ ein Prämaß ist, gilt

$$a \geq -\mu(X^-) = -\mu\left(\bigcup_{n=1}^{\infty} A_n\right) \geq -\mu(A_n) \longrightarrow a\,,$$

also $\mu(X^-) = -a$.

Wir setzen

$$X^+ := X \setminus X^-$$

und behaupten, dass $X = X^+ \cup X^-$ eine Hahn-Zerlegung ist.

Beweis der Behauptung. Angenommen, dies wäre nicht der Fall. Da μ nach Konstruktion auf allen Teilmengen von X^- nichtpositiv ist, muss dann eine Menge $A_1 \subseteq X^+$ mit $\mu(A_1) < 0$ existieren. Es gilt dann aber $\mu^+(A_1) > 0$, denn im Fall $\mu^+(A_1) \leq 0$ wäre $A_1 \in \mathcal{A}_-$ und damit

$$\mu(A_1 \cup X^-) = \mu(A_1) + \mu(X^-) < -a\,,$$

im Widerspruch zur Definition von a. Wegen $\mu^+(A_1) > 0$ existiert (nach Definition von μ^+) eine Menge $B_1 \subseteq A_1$ mit

$$\mu(B_1) \geq \left.\begin{cases} \frac{1}{2}\mu^+(A_1) & \text{falls } \mu^+(A_1) < \infty \\ 1 & \text{falls } \mu^+(A_1) = \infty \end{cases}\right\} \geq 0\,.$$

Wir setzen nun $A_2 := A_1 \setminus B_1 \subseteq X^+$. Dann ist $\mu(A_2) = \mu(A_1) - \mu(B_1) \leq \mu(A_1) < 0$ und wir können die Prozedur mit A_2 anstelle von A_1 wiederholen. Induktive erhalten wir daher Mengenfolgen $A_1, A_2, \ldots$ und $B_1, B_2, \ldots$ mit folgenden Eigenschaften:

- $A_1 \supseteq A_2 \supseteq \cdots$ und $0 > \mu(A_1) \geq \mu(A_2) \geq \ldots$
- $B_n \subseteq A_n$ mit $\mu(B_n) \geq \begin{cases} \frac{1}{2}\mu^+(A_n) & \text{falls } \mu^+(A_n) < \infty \\ 1 & \text{falls } \mu^+(A_n) = \infty \end{cases}$
- $B_n \cap B_m = \emptyset$ für $n \neq m$ und $A_{n+1} = A_n \setminus B_n$.

Wir setzen

$$S := \bigcap_{n=1}^{\infty} A_n = A_1 \setminus \bigcup_{n=1}^{\infty} B_n \,.$$

Nach Lemma 6.3 (3) gilt dann

$$\mu(S) = \lim_{n\to\infty} \mu(A_n) \leq \mu(A_1) < 0 \,,$$

insbesondere ist also $\mu(S)$ endlich. Zudem gilt

$$\left|\sum_{n=1}^{\infty} \mu(B_n)\right| = \left|\mu\left(\bigcup_{n=1}^{\infty} B_n\right)\right| = \big|\mu(A_1 \setminus S)\big| = \big|\mu(A_1) - \mu(S)\big| < \infty \,.$$

Also konvergiert die Reihe absolut und insbesondere gilt $\mu(B_n) \to 0$. Insbesondere gilt $\mu^+(A_n) < \infty$ für alle bis auf endlich viele n, denn ansonsten wäre $\mu(B_n) \leq 1$ für unendlich viele n und die Reihe könnte nicht konvergieren.

Angenommen, es wäre $S \in \mathcal{A}_-$. Dann wäre

$$\mu(S \cup X^-) = \mu(S) + \mu(X^-) < -a \,,$$

im Widerspruch zur Definition von a. Im anderen Fall $\mu^+(S) > 0$ existiert $B \subseteq S$ so, dass $\mu(B) > 0$ gilt. Wegen $B \cap B_n = \emptyset$ und $B \cup B_n \subseteq A_n$ gilt

$$2\mu(B_n) \geq \mu^+(A_n) \geq \mu^+(B \cup B_n) = \mu(B) + \mu(B_n) \geq \mu(B) > 0$$

für alle n mit $\mu^+(A_n) < \infty$, also alle bis auf endlich viele. Dies ist aber ein Widerspruch zu $\mu(B_n) \to 0$. Damit kann die Menge S nicht existieren. $\square$

Ist $X = \tilde{X}^+ \cup \tilde{X}^-$ eine weitere Hahn-Zerlegung, so ist X die disjunkte Vereinigung

$$X = (X^+ \cap \tilde{X}^+) \cup (X^- \cap \tilde{X}^-) \cup (X^+ \cap \tilde{X}^-) \cup (X^- \cap \tilde{X}^+) \,.$$

Für eine Teilmenge A der letzten beiden Mengen gilt sowohl $\mu(A) \leq 0$ als auch $\mu(A) \geq 0$, also $\mu(A) = 0$. Damit sind die beiden letzten Mengen Nullmengen.

Sei nun $X = X^+ \cup X^-$ eine beliebige Hahn-Zerlegung. Wir zeigen, dass die Formeln (6.1.1) gelten. Hierzu bemerken wir, dass für $A, B \in \mathcal{A}$ mit $B \subseteq A$ gilt

$$\mu(B) = \mu(X^+ \cap B) + \underbrace{\mu(X^- \cap B)}_{\leq 0} \leq \mu(X^+ \cap B) \leq \mu(A \cap X^+) \leq \mu^+(A) \,.$$

Nehmen des Supremums über B liefert

$$\mu^+(A) \leq \mu(A \cap X^+) \leq \mu^+(A) \,,$$

also die erste Gleichheit in (6.1.1). Analog zeigt man $\mu^-(A) = \mu(A \cap X^-)$.

Dass μ^+ und μ^- Maße sind, folgt aus (6.1.1). ■

Sei X ein lokalkompakter Hausdorffraum (also ein Hausdorffraum, in dem jeder Punkt in einer offenen Menge mit kompaktem Abschluss enthalten ist). Wir schreiben

$$C_0(X) := \left\{ f : X \to \mathbb{R} \text{ stetig} \mid \forall \varepsilon > 0 \exists K \subseteq X \text{ kompakt} : \|f_{X \setminus K}\|_\infty \leq \varepsilon \right\}$$

für den Raum der stetigen Funktionen auf X, die bei unendlich verschwinden. Hier bei bezeichnet

$$\|f\|_\infty := \sup \left\{ |f(x)| \mid x \in X \right\}$$

die Supremumsnorm, die $C_0(X)$ zu einem normierten Vektorraum macht. Da der Grenzwert einer gleichmässig konvergenten Folge von stetigen Funktionen wieder stetig ist (und auch wieder bei unendlich verschwindet, wenn es die Folgenglieder taten), ist $C_0(X)$ vollständig bezüglich dieser Norm, also ein Banachraum. Tatsächlich ist $C_0(X)$ der Normabschluss des (nicht vollständigen) Raumes $C_c(X)$ der kompakt getragenen Funktionen.

Definition 6.10. Sei $\mathcal{B}_X$ die Borel'sche σ-Algebra von X, also die von den offenen (oder äquivalent den abgeschlossenen) Teilmengen von X erzeugte σ-Algebra. Ein auf $\mathcal{B}_X$ definiertes Maß μ heißt *Borelmaß*, wenn jeder Punkt $x \in X$ eine offene Umgebung U besitzt mit $\mu(U) < \infty$. Ein Borelmaß μ heißt *regulär*, falls

$$\mu(A) = \sup \left\{ \mu(K) \;\middle|\; K \subseteq A \text{ kompakt} \right\} = \inf \left\{ \mu(U) \;\middle|\; U \supseteq A \text{ offen} \right\}$$

für alle $A \in \mathcal{B}_X$ gilt. Ein signiertes Maß $\mu = \mu_+ - \mu_-$ heißt (reguläres) Borelmaß, falls sowohl μ_+ als auch μ_- (reguläre) Borelmaße sind sind. Mit $\mathcal{M}(X)$ bezeichnen wir den Raum der regulären signierten Borelmaße auf X mit endlicher Totalvariation.

Bemerkung 6.11. Man kann zeigen, dass in einem lokalkompakten Raum, dessen Topologie eine abzählbare Basis besitzt (zum Beispiel $\mathbb{R}^d$), jedes Borelmaß regulär ist.

Jedes auf $\mathcal{B}_X$ definierte signierte Borelmaß μ mit endlicher Totalvariation erzeugt ein lineares Funktional

$$J_\mu : C_0(X) \longrightarrow \mathbb{R}$$

auf $C_0(X)$ durch die Definition

$$J_\mu(f) := \int_X f(x)\,d\mu(x) := \int_X f(x)\,d\mu_+(x) - \int_X f(x)\,d\mu_-(x)\ ,$$

wobei $\mu = \mu_+ - \mu_-$ die Jordanzerlegung von μ ist.

Bemerkung 6.12. Für die Wohldefiniertheit von J_μ zeigt man, dass jede positive Funktion $f \in C_0(X)$ monoton von unten durch Treppenfunktionen über $\mathcal{B}_X$ approximiert werden kann. Damit existiert ihr verallgemeinertes Integral, und wegen der endlichen Totalvariation von μ ist f sogar integrierbar.

Zur Erinnerung: Der Dualraum E' eines normierten Raumes E ist definiert als der Raum aller stetigen linearen Funktionale $E \to \mathbb{R}$. E' ist ebenfalls ein normierter Raum, mit der Operatornorm

$$\|\varphi\|_{E'} := \sup\left\{|\varphi(x)| \;\middle|\; x \in E, \|x\|_E \le 1\right\}.$$

E' ist mit dieser Norm stets vollständig, also ein Banachraum. Die Definition ist so gemacht, dass immer

$$|\varphi(x)| \le \|\varphi\|_{E'}\|x\|_E, \qquad \varphi \in E',\ x \in E\ .$$

Satz von Markhov–Kakutani–Riesz 6.13. Die Abbildung

$$\Phi : \mathcal{M}(X) \longrightarrow C_0(X)'\ , \qquad \mu \longrightarrow J_\mu$$

ist ein isometrischer Isomorphismus.

Beweis der Isometrie. Sei μ ein reguläres Borelmaß mit endlicher Totalvariation auf μ. Sei $f \in C_0(X)$ mit $\|f\|_\infty \le 1$. Dann ist

$$|J_\mu(f)| \le J_{|\mu|}(|f|) \le J_{|\mu|}(\mathbf{1}_X) = |\mu|(X) = \|\mu\|\ .$$

Es folgt also, dass $\|J_\mu\| \le \|\mu\|$. Insbesondere ist damit J_μ ein *stetiges* lineares Funktional, also ist Φ zunächst einmal wohldefiniert. Weiter zeigt die Abschätzung, dass die Operatornorm von Φ kleiner oder gleich Eins ist.

Um nachzuweisen, dass Φ isometrisch ist, müssen wir $\|J_\mu\| = \|\mu\|$ nachweisen. Im Fall dass μ positiv und X kompakt ist, können wir dann einfach in der obigen Rechnung $f = \mathbf{1}_X$ setzen.

Im allgemeinen Fall wählen wir zunächst eine Hahn-Zerlegung $X = X^+ \cup X^-$. Zu $\varepsilon > 0$ seinen nun $K^\pm \subseteq X^\pm$ kompakte und $U^\pm \supseteq X^\pm$ offene Teilmengen so, dass

$$\mu(X^\pm \setminus K^\pm) \le \varepsilon \qquad \text{und} \qquad \mu(U^\pm \setminus K^\pm) \le \varepsilon.$$

(diese existieren, da μ als regulär vorausgesetzt ist). Da X also lokalkompakter Raum normal ist existieren kompakt getragene, nichtnegative stetige Funktionen $f^\pm : X \to$

$[0,1]$ so, dass $f^\pm$ konstant Eins auf $K^\pm$ und null of $X \setminus U^\pm$. Da Dann gilt

$$\begin{aligned} J_\mu(f^+ - f^-) &\geq J_{\mu_+}(f^+) + J_{\mu_-}(f^+) - J_{\mu_+}(f^-) - J_{\mu_-}(f^+) \\ &\geq J^*_{\mu_+}(\mathbf{1}_{K^+}) + J^*_{\mu_-}(\mathbf{1}_{K^-}) - 2\varepsilon. \\ &\geq \mu_+(X_+) + \mu_-(X_-) - 4\varepsilon = \|\mu\| - 4\varepsilon \,. \end{aligned}$$

Also gilt auch

$$\|J_\mu\| = \sup\left\{|J_\mu(f)| \,\middle|\, f \in C_0(X)\right\} \geq \|\mu\| - 4\varepsilon \,.$$

Da $\varepsilon > 0$ beliebig war, folgt $\|J_\mu\| = \|\mu\|$. □

Beweis der Surjektivität. Da Φ isometrisch ist, ist Φ automatisch injektiv. Wir zeigen nun, dass Φ surjektiv ist. Sei hierzu $\varphi : C_0(X) \to \mathbb{R}$ ein beliebiges stetiges lineares Funktional. Wir nehmen zunächst an, dass φ positiv ist, dass also $\varphi(f) \geq 0$ für wenn $f \geq 0$. Dann ist φ ein Elementarintegral. Hierzu ist nur die Nullstetigkeit nachzuweisen: Sei $f_n \searrow 0$ eine Folge in $C_0(X)$, die monoton von oben gegen Null konvergiert. Dann gilt sogar $\|f_n\|_\infty \searrow 0$ (dies ist eine Variante des Satzes von Dini 4.9), und damit

$$|\varphi(f_n)| \leq \|\varphi\| \cdot \|f_n\|_\infty \searrow 0 \,.$$

Kann nun X als abzählbare Vereinigung von kompakten Mengen $K_1, K_2, \dots$ geschrieben werden, so ist die Lebesgue-Erweiterung $\bar\varphi$ des Elementarintegrals φ Stone'sch: Sind χ_n kompakt getragene Funktionen, die auf K_n konstant gleich Eins sind (diese existieren, da X als lokalkompakter Raum normal ist), so ist $\chi_1 \wedge \dots \wedge \chi_n$ eine Folge von Funktionen in $C_0(X)$, die punktweise gegen $\mathbf{1}_X$ konvergieren, also ist $\mathbf{1}_X$ messbar. Damit induziert $\bar\varphi$ ein Maß μ auf X, definiert auf einer σ-Algebra $\mathcal{A}$. Dass μ ein reguläres Borelmaß ist, folgt ähnlich wie im Fall des Borelmaßes (siehe Satz 4.14). Nach Satz 3.19 ist die Einschränkung des zu μ gehörigen Lebesgue-Integrals auf $C_0(X)$ gerade φ, also gilt $\varphi = J_\mu$.

Eine Schwierigkeit ist nun, dass die Lebesgue-Erweiterung $\bar\varphi$ des Elementarintegrals φ nicht unbedingt Stone'sch sein muss, falls X keine abzählbare Ausschöpfung durch kompakte Mengen besitzt. In diesem Fall sei (f_n) eine Folge von Funktionen in $C_c(X)$ mit $\|f_n\| = 1$ und $\varphi(f_n) \nearrow \|\varphi\|$ (solch eine Folge existiert, da $C_c(X)$ dicht in $C_0(X)$ liegt). Sei $T \subseteq X$ die Vereinigung der Träger der f_n. Ist dann g eine Funktion mit $\|g\| \leq 1$, die in $X \setminus T$ getragen ist, dann gilt

$$\|\varphi\| \geq \varphi(f_n + g) = \varphi(f_n) + \varphi(g) \to \|\varphi\| + \varphi(g) \,,$$

also $\varphi(g) = 0$. Wir erhalten darum ein Elementarintegral φ_T auf dem Vektorverband $C_0(X)|_T$ von Funktionen auf T, gegeben durch die Vorschrift

$$\varphi_T(f|_T) := \varphi(f) \,.$$

Nach der obigen Rechnung hängt dies nicht von der Wahl der Fortsetzung f ab. Analog zu oben erhält man dann ein Borelmaß μ_T auf T und die Fortsetzung

$$\mu(A) := \mu_T(A \cap T)$$

von μ_T auf ganz X ist ein reguläres Borelmaß auf X. Auch in diesem Fall gilt, dass φ die Einschränkung des zugehörigen Lebesgue-Integrals auf $C_0(X)$ ist.

Es bleibt zu zeigen, dass sich jedes beliebige stetige Funktional $\varphi : C_0(X) \to \mathbb{R}$ als $\varphi = \varphi_+ - \varphi_-$ schreiben lässt. Hierzu definieren wir für positives $f_+ \in C_0(X)$

$$\varphi_+(f_+) := \sup\left\{\varphi(g) \,\middle|\, 0 \le g \le f_+\right\}$$

und allgemein für $f \in C_0(X)$ mit Aufspaltung $f = f^+ - f^-$ in Positiv- und Negativteil

$$\varphi_+(f) := \varphi_+(f^+) - \varphi_+(f^-) \,.$$

Dann ist φ_+ linear, stetig und positiv und genauso auch $\varphi_- := \varphi_+ - \varphi$. Dies liefert die gewünschte Darstellung. ■

Beispiel 6.14. Wir geben ein Beispiel dafür, dass die Lebesgue-Erweiterung das Elementarintegrals $(X, C_0(X), \varphi)$ aus dem obigen Beweis nicht immer Stone'sch sein muss: Sei X eine überabzählbare Menge mit der diskreten Topologie. Hier sind alle kompakten Mengen endlich, also hat jede Funktion aus $C_0(X)$ höchstens abzählbaren Träger. Damit ist $\mathbf{1}_X$ nicht der punktweise Limes von Funktionen in $C_0(X)$.

6.2 Dichtefunktionen

Lemma 6.15. Sei $(X, \mathcal{A}, \mu)$ ein σ-endlicher Maßraum und sei $g \ge 0$ eine $\mathcal{A}$-messbare Funktion auf X. Dann wird durch

$$\mu_g(A) := \int_A^* g(x)\,d\mu(x) = \int_X^* g(x) \cdot \mathbf{1}_A(x)\,d\mu(x)$$

ein Maß auf $\mathcal{A}$ definiert.

Das Maß μ_g ist σ-endlich genau dann, wenn $\mu(\{g(x) = \infty\}) = 0$ gilt.

Beweis. Da g und $\mathbf{1}_A$ messbar ist, ist auch $g \cdot \mathbf{1}_A$ messbar. Wegen $g \cdot \mathbf{1}_A \ge 0$ ist die Funktion nach Lemma 2.27 in $\mathcal{L}(X)^*$ enthalten, also existiert das verallgemeinerte Integral. Die σ-Additivität der resultierenden Mengenfunktion ist offensichtlich, also ist μ_g ein Maß.

($\Longleftarrow$) Zum Beweis des Zusatzes sei zunächst $B := \{g(x) = \infty\}$ eine Nullmenge. Da X σ-endlich ist, existieren Mengen A_n mit $\mu(A_n) < \infty$ und $\bigcup_{n=1}^{\infty} A_n = X$. Für $n \in \mathbb{N}$ setzen wir $B_n = \{n - 1 \le g(x) < n\}$. Die B_n zusammen mit B bilden dann eine

Zerlegung von X in disjunkte messbare Teilmengen und es gilt

$$X = \left(\bigcup_{n=1}^{\infty} A_n \right) \cap \left(B \cup \bigcup_{k=1}^{\infty} B_k \right) = B \cup \bigcup_{k,n=1}^{\infty} A_n \cap B_k \, .$$

Dann gilt

$$\mu_g(A_n \cap B_k) = \int_{A_n \cap B_k}^{*} g(x) d\mu(x) \le k \cdot \int_{A_n}^{*} d\mu(x) \le k \cdot \mu(A_n)$$

und

$$\mu_g(B) = \int_X^{*} \infty \cdot \mathbf{1}_B(x) d\mu(x) = \lim_{n \to \infty} \int_X n \cdot \mathbf{1}_B(x) d\mu(x) = \lim_{n \to \infty} n \cdot \mu(B) = 0.$$

Dies zeigt die σ-Endlichkeit von μ_g.

($\Longrightarrow$) Sei μ_g umgekehrt σ-endlich. Dann existiert eine Zerlegung $X = \bigcup_{n=1}^{\infty} A_n$ in Mengen mit $\mu_g(A_n) < \infty$. Dann gilt für jedes $n \in \mathbb{N}$, dass

$$\int_X^{*} \infty \cdot \mathbf{1}_{A_n \cap B}(x) d\mu(x) \le \int_{A_n \cap B}^{*} g(x) d\mu(x) = \mu_g(A_n \cap B) < \infty \, .$$

Dies ist nur für $\mu(A_n \cap B) = 0$ möglich. Also gilt

$$\mu(B) = \mu\left(\bigcup_{n=1}^{\infty} A_n \cap B \right) = \sum_{n=1}^{\infty} \mu(A_n \cap B) = 0 \, . \qquad \blacksquare$$

Definition 6.16. Seien μ und ν zwei σ-endliche Maße, definiert auf derselben σ-Algebra $\mathcal{A}$ von Teilmengen von X. Eine messbare Funktion $g \ge 0$ heißt *Dichtefunktion von ν bezüglich μ*, falls

$$\nu = \mu_g$$

gilt. Wir schreiben auch

$$g = \frac{d\nu}{d\mu}.$$

Zur Motivation der obigen Schreibweise betrachten wir folgendes Beispiel.

Beispiel 6.17. Sei $\mu = \lambda$, das ein-dimensionale Lebesgue-Maß und $\nu = \lambda_g$ für eine stetige Funktion g. Nach dem Mittelwertsatz der Integralrechnung gilt

$$\nu([a,b]) = \int_a^b g(t) dt = g(\xi) \cdot (b - a) = g(\xi) = \lambda([a,b]) \, .$$

für ein geeignetes $\xi \in [a,b]$. Also gilt

$$g(\xi) = \frac{\nu([a,b])}{\mu([a,b])}$$

und damit

$$g(x) = \lim_{\delta \to 0} \frac{\nu([x,x+\delta])}{\mu([x,x+\delta])}, \quad \text{oder suggestiv auch} \quad \frac{d\nu}{d\mu}(x).$$

Lemma 6.18. Sind g, h zwei Dichtefunktionen von ν bezüglich μ, so gilt $g = h$ μ-fast überall.

Beweis. Sei $X = \bigcup_{n=1}^{\infty} A_n$ eine Zerlegung mit $\nu(A_n) < \infty$. Sei $B := \{g(x) \geq h(x)\}$. Dann gilt

$$\begin{aligned} \int_{A_n \cap B}^{*} \big(g(x) - h(x)\big) d\mu(x) &= \int_{A_n \cap B}^{*} g(x) d\mu(x) - \int_{A_n \cap B}^{*} h(x) d\mu(x) \\ &= \mu_g(A_n \cap B) - \mu_h(A_n \cap B) = 0\,, \end{aligned}$$

da $\mu_g = \mu_h = \nu$. Es folgt, dass auch

$$\int_X^{*} (g-h)_+(x) d\mu(x) = \sum_{n=1}^{\infty} \int_{A_n \cap B}^{*} \big(g(x) - h(x)\big) d\mu(x)^{\infty} = 0\,,$$

also ist der Positivteil $(g-h)_+$ eine nichtnegative Funktion mit Integral Null, der Träger von $(g-h)_+$ ist also eine Nullmenge (siehe Beispiel 3.21). Analog ist auch $(h-g)_+$ eine Nullmenge. Insgesamt ist also $\{g(x) \neq h(x)\}$ eine Nullmenge. ■

Definition 6.19. Seien μ und ν zwei σ-endliche Maße, definiert auf derselben σ-Algebra $\mathcal{A}$ von Teilmengen von X. Das Maß ν heißt *absolut stetig* bezüglich μ, in Zeichen $\nu \ll \mu$, wenn für jedes $A \in \mathcal{A}$ mit $\mu(A) < \infty$ gilt: Für jedes $\varepsilon > 0$ existiert $\delta > 0$ so, dass für alle messbaren Teilmengen $B \subseteq A$ mit $\mu(B) \leq \delta$ gilt $\nu(B) \leq \varepsilon$.

Lemma 6.20. Seien μ und ν zwei σ-endliche Maße auf $\mathcal{A}$. Wir betrachten die folgenden Bedingungen:

(i) ν besitzt eine Dichtefunktion bezüglich μ.

(ii) ν ist absolut stetig bezüglich μ.

(iii) Für alle $B \in \mathcal{A}$ mit $\mu(B) = 0$ gilt $\nu(B) = 0$.

Dann gilt (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii).

Der Satz von Radon-Nikodym 6.22 zeigt schließlich (iii) $\Rightarrow$ (i).

Beweis. (i) $\Rightarrow$ (ii). Sei g eine Dichtefunktion von ν bezüglich μ. Sei A mit $\nu(A) < \infty$. Dann gilt

$$\lim_{n\to\infty} \int_A g(x) \wedge n \, d\mu(x) = \int_A^* g(x) d\nu(x) = \nu(A) < \infty.$$

Daher existiert zu jedem $\varepsilon > 0$ ein $n \in \mathbb{N}$ mit

$$\int_A^* \big(g(x) - g(x) \wedge n\big) \, d\mu(x) \leq \frac{\varepsilon}{2} \, .$$

Sei $\delta = \varepsilon/2n$. Dann gilt für $B \in \mathcal{A}$ mit $B \subseteq A$ und $\mu(B) \leq \delta$, dass

$$\begin{aligned} \nu(B) = \int_B g(x) d\mu(x) &= \int_B \big(g(x) - g(x) \wedge n\big) d\mu(x) + \int_B g(x) \wedge n \, d\mu(x) \\ &\leq \frac{\varepsilon}{2} + n\mu(B) \leq \varepsilon \, . \end{aligned}$$

(ii) $\Rightarrow$ (iii). Sei $N \in \mathcal{A}$ mit $\mu(N) = 0$. Da insbesondere $\mu(N) < \infty$ gilt, existiert zu jedem $\varepsilon > 0$ ein $\delta > 0$ so, dass für $B \subseteq N$ mit $\mu(B) \leq \delta$ gilt $\nu(B) \leq \varepsilon$. Wenn wir $B = N$ wählen, so gilt insbesondere $\mu(N) \leq \delta$ für jedes $\delta > 0$, also auch $\nu(N) \leq \varepsilon$ für jedes $\varepsilon > 0$. Damit gilt $\nu(N) = 0$. ■

6.3 Der Satz von Radon-Nikodym

Definition 6.21. Sei X eine Menge und seien μ, ν Maße auf X, definiert auf derselben σ-Algebra $\mathcal{A}$. μ und ν heißen *orthogonal*, in Zeichen $\mu \perp \nu$, falls es Trägermengen $T \subseteq X$, $S \subseteq X$ von μ, bzw. ν mit $T \cap S = \emptyset$ gibt. Man sagt auch dass μ *singulär* bezüglich μ ist (und umgekehrt).

Satz von Radon–Nikodym 6.22. Seien μ, ν zwei σ-endliche Maße auf X, definiert auf derselben σ-Algebra $\mathcal{A}$. Dann gibt es eine messbare Funktion $f : X \to \mathbb{R}$ und ein eindeutiges zu μ orthogonales Maß ρ mit

$$\nu = \mu_f + \rho \, .$$

Jede andere Funktion f' mit der die obige Gleichheit gilt stimmt mit f fast überall überein.

Beweis der Eindeutigkeit. Seien $\nu = \mu_{f_1} + \rho_1 = \mu_{f_2} + \rho_2$ zwei Darstellungen wie im Satz. Seien T_1, T_2 Trägermengen zu ρ_1, ρ_2 mit $\mu(T_1) = \mu(T_2) = 0$. Dann ist auch $T := T_1 \cup T_2$ eine μ-Nullmenge, sowie eine Trägermenge für sowohl ρ_1 als auch ρ_2. Also gilt für $A \in \mathcal{A}$ beliebig, dass

$$\nu(A \cap T) = \underbrace{\int_{A\cap T} f(x) d\mu_i(x)}_{=0} + \rho_i(A \cap T) = \rho_i(A \cap T) = \rho_i(A) \, .$$

Also gilt $\rho_1 = \rho_2$. Es folgt, dass $f_1 \cdot \mu = f_2 \cdot \mu$ und damit wegen Lemma 6.18 auch μ-fast überall $f_1 = f_2$.

Beweis der Existenz. Wir nehmen zunächst an, dass $\nu(X) < \infty$ gilt. Wir betrachten das System

$$\mathcal{F} := \left\{ g \geq 0 \text{ messbar} \;\middle|\; \mu_g \leq \nu \right\}$$

von Funktionen auf X. Wir behaupten, dass $\mathcal{F}$ ein "maximales Element" enthält, genauer gesagt, ein $f \in \mathcal{F}$ so, dass für jedes $g \in \mathcal{F}$ fast überall $g \leq f$ gilt.

Zum Beweis der Behauptung zeigen wir zunächst die Zwischenbehauptung, dass das Supremum jeder abzählbaren Teilmenge von $\mathcal{F}$ wieder in $\mathcal{F}$ liegt. Hierzu wiederum seien zunächst $g_1, g_2 \in \mathcal{F}$. Dann gilt für $A \in \mathcal{A}$

$$\begin{aligned} \mu_{g_1 \vee g_2}(A) &= \mu_{g_1 \vee g_2}\big(A \cap \{g_1 \leq g_2\}\big) + \mu_{g_1 \vee g_2}\big(A \cap \{g_1 > g_2\}\big) \\ &\leq \mu_{g_2}\big(A \cap \{g_1 \leq g_2\}\big) + \mu_{g_1}\big(A \cap \{g_1 > g_2\}\big) \\ &\leq \nu\big(A \cap \{g_1 \leq g_2\}\big) + \nu\big(A \cap \{g_1 > g_2\}\big) = \nu(A) \, , \end{aligned}$$

also $\mu_{g_1 \vee g_2} \leq \nu$. Ist nun $g_1 \leq g_2 \leq \dots$ eine monoton steigende Folge von Funktionen in $\mathcal{F}$ mit Supremum g, dann gilt für jedes $A \in \mathcal{A}$, dass

$$\mu_g(A) = \lim_n \mu_{g_n}(A) \leq \nu(A) \, ,$$

also $\mu_g \leq \nu$. Ist nun (g_n) eine beliebige Teilfolge mit Supremum g, so ist g auch das Supremum der monoton steigenden Folge $g_1 \vee \dots \vee g_n$, also folgt $\mu_g \leq \nu$ aus der Kombination der obigen beiden Aussagen.

Zur Konstruktion des maximalen Elements sei

$$a := \sup\left\{\mu_g(X) \;\middle|\; g \in \mathcal{F}\right\} \leq \nu(X) < \infty\ .$$

Sei (g_n) eine Folge mit $\mu_{g_n}(X) \nearrow a$ und setze $f := \sup_n g_n$. Dann ist wegen der Zwischenbehauptung $f \in \mathcal{F}$, also $\mu_f \leq \nu$. Weiterhin gilt für ein beliebiges $g \in \mathcal{F}$, dass

$$\begin{aligned} \mu_g(A) + \mu_f(X \setminus A) &\leq \mu_{g\vee f}(A) + \mu_{g\vee f}(X \setminus A) \\ &= \mu_{g\vee f}(X) \\ &\leq a = \mu_f(X) \qquad \text{(weil } g \vee f \in \mathcal{F}) \\ &= \mu_f(A) + \mu_f(X \setminus A) \end{aligned}$$

also $\mu_g \leq \mu_f$ und damit μ-fast überall $g \leq f$, wie behauptet.

Nach der obigen Konstruktion ist $\rho := \nu - \mu_f$ positiv, also ein Maß. Wir zeigen, dass ρ zu μ_f orthogonal ist. Hierzu verwenden wir die signierten Hilfsmaße

$$\rho_n := \rho - \frac{1}{n}\mu = \nu - \mu_{f+\frac{1}{n}}\ .$$

Seien $X = X_n^+ \cup X_n^-$ zugehörige Hahn'sche Zerlegungen. Für $A \in \mathcal{A}$ gilt dann

$$\begin{aligned} \mu_{f+\frac{1}{n}\mathbf{1}_{X_n^+}}(A) &= \mu_f(A) + \frac{1}{n}\mu(A \cap X_n^+) \\ &= \mu_f(A) + \rho(A \cap X_n^+) - \underbrace{\rho_n(A \cap X_n^+)}_{\geq 0} \\ &\leq \mu_f(A) + \rho(A) = \nu(A)\ . \end{aligned}$$

Also gilt $f + \frac{1}{n}\mathbf{1}_{X_n^+} \in \mathcal{F}$. Wegen der Maximalität von f ist dann aber X_n^+ eine μ-Nullmenge. Setzen wir

$$X^+ := \bigcup_{n=1}^{\infty} X_n^+\ , \qquad X^- := \bigcap_{n=1}^{\infty} X_n^-\ ,$$

dann ist X^+ eine μ-Nullmenge und damit $X^- = X \setminus X^+$ eine Trägermenge für μ. Wir zeigen, dass X^+ eine Trägermenge für ρ ist. Für alle $n \in \mathbb{N}$ ist $X^- \subseteq X_n^-$, also

$$\rho_n(X^-) = \rho_n(X^- \cap X_n^-) \leq 0\ ,$$

da ρ_n negativ auf X_n^- ist. Also folgt

$$\frac{1}{n}\mu(X^-) = \rho(X^-) - \rho_n(X^-) \geq \rho(X^-) \geq 0\ ,$$

für alle $n \in \mathbb{N}$, was nur für $\rho(X^-) = 0$ möglich ist. Damit ist X^+ eine Trägermenge für X und somit ist $\rho \perp \mu$.

Ist schließlich ν nicht endlich sondern nur σ-endlich, so wählen wir eine Zerlegung $X = \bigcup_{n=1}^{\infty} A_n$ mit $\nu(A_n) < \infty$ und wenden den obigen Spezialfall auf die Einschränkungen von ν auf die A_n an. "Zusammenbauen" der Resultate liefert das Ergebnis. ■

Satz 6.23 (Spezialfall von Radon-Nikodym)**.** Seien μ und ν zwei σ-endliche Maße auf $\mathcal{A}$. Dann sind äquivalent

(i) ν besitzt eine Dichtefunktion bezüglich μ.

(ii) ν ist absolut stetig bezüglich μ.

(iii) Für alle $B \in \mathcal{A}$ mit $\mu(B) = 0$ gilt $\nu(B) = 0$.

Die Implikationen (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) wurden schon in Lemma 6.20 bewiesen.

Beweis von (iii) $\Longrightarrow$ (i). Nach dem Satz von Radon-Nikodym gibt es eine Darstellung $\nu = \mu_f + \rho$ mit $\rho \perp \mu$. Seien T eine Trägermenge von ρ mit $\mu(T) = 0$. Dann gilt

$$\rho(A) = \rho(A \cap T) = \nu(A \cap T) - \mu_f(A \cap T) = \nu(A \cap T)\ ,$$

da T eine μ-Nullmenge ist. Wegen (iii) ist T aber auch eine ν-Nullmenge. Also folgt sogar $\rho = 0$. Damit gilt $\nu = \mu_f$. ■

6.4 Lebesgue-Stieltjes-Integrale

In diesem Abschnitt klassifizieren wir alle regulären Borelmaße auf der reellen Achse (mit der üblichen Topologie). Im gesamten Abschnitt sei $\mathcal{B}$ die Borelsche σ-Algebra auf $\mathbb{R}$ (also die kleinste σ-Algebra, die von den offenen/halboffenen/abgeschlossenen Intervallen erzeugt wird) und jedes Maß sei auf $\mathcal{B}$ definiert.

Definition 6.24. Eine Funktion $H : \mathbb{R} \to \mathbb{R}$ heißt *Verteilungsfunktion*, falls sie monoton steigend und rechtsseitig stetig ist.

Sei H eine Verteilungsfunktion und sei $f \in C_c([a, b])$. Für eine Zerlegung $Z = \{-\infty = t_0 < t_1 < t_2 < \cdots < t_{N+1} = \infty\}$ des Intervalls $\mathbb{R}$ definieren wir die *untere* und die *obere Zerlegungssumme* als

$$\overline{S}(f, Z) := \sum_{i=0}^{N} \max_{t \in [t_i, t_{i+1}]} f(t) \cdot \big(H(t_{i+1}) - H(t_i)\big)\ ,$$
$$\underline{S}(f, Z) := \sum_{i=0}^{N} \min_{t \in [t_i, t_{i+1}]} f(t) \cdot \big(H(t_{i+1}) - H(t_i)\big)\ .$$

Hierbei ist zu beachten, dass Maximum und Minimum auch auf den unendlichen Randintervallen existieren, da die Funktion kompakt getragen ist. Wir bezeichnen mit $\mathcal{Z}$ die Menge aller Zerlegungen Z wie oben. Ähnlich wie beim Riemann-Integral beweist man folgenden Satz.

Satz 6.25. Für alle $f \in C([a,b])$ gilt

$$\int_{-\infty}^{\infty} f(t)dH(t) := \inf_{Z\in\mathcal{Z}} \overline{S}(f,Z) = \sup_{Z\in\mathcal{Z}} \underline{S}(f,Z)\,. \tag{6.4.1}$$

Diese Zahl heißt *Riemann-Stieltjes-Integral* von f bezüglich H.

Bemerkung 6.26. Zudem gilt: Ist $Z_n = \{-\infty = t_0^{(n)} < t_1^{(n)} < \cdots < t_{N_n+1}^{(n)} = \infty\}$ eine beliebige Folge von Zerlegungen, so, dass der Träger von f stets im Intervall $[t_1^{(n)}, t_{N_n}^{(n)}]$ enthalten ist und so, dass die Feinheit $\sup_{2\le i\le N}(t_i^{(n)} - t_{i-1}^{(n)})$ gegen Null konvergiert, so gilt (6.4.1) mit

$$\int_{-\infty}^{\infty} f(t)dH(t) = \lim_{n\to\infty} \sum_{i=0}^{N_n} f(\xi_i^{(n)}) \cdot \left(H(t_{i+1}^{(n)}) - H(t_i^{(n)})\right).$$

Satz 6.27. Ist H stetig differenzierbar, so gilt

$$\int_{-\infty}^{\infty} f(t)dH(t) = \int_{-\infty}^{\infty} f(t)H'(t)dt.$$

Beweis. Zu einer gegebenen Zerlegung $Z = \{t_0 < \cdots < t_N\} \in \mathcal{Z}$ existieren nach dem Mittelwertsatz der Differentialrechnung $\xi_i \in [t_i, t_{i+1}]$ so, dass

$$H(t_{i+1}) - H(t_i) = H'(\xi_i)\cdot(t_{i+1} - t_i)\,.$$

Ist also (Z_n) eine Folge von Zerlegungen wie in Bemerkung (6.26), so gilt

$$\begin{aligned}\int_{-\infty}^{\infty} f(t)dH(t) &= \lim_{n\to\infty} \sum_{i=0}^{N_n} f(\xi_i^{(n)}) \cdot \left(H(t_{i+1}) - H(t_i)\right)\\ &= \lim_{n\to\infty} \sum_{i=0}^{N_n} f(\xi_i^{(n)}) \cdot H'(\xi_i^{(n)}) \cdot (t_{i+1} - t_i)\\ &= \int_{-\infty}^{\infty} f(t)H'(t)dt\end{aligned}$$

∎

Lemma 6.28. Die Abbildung

$$J_H(f) := \int_{-\infty}^{\infty} f(t)dH(t)$$

ist ein Elementarintegral auf dem Raum $C_c(\mathbb{R})$.

Beweis. J ist positiv, da H monoton wachsend ist. Nullstetigkeit von oben folgt mit dem Satz von Dini zusammen mit der Abschätzung

$$\left|\int_{-\infty}^{\infty} f(t)dH(t)\right| \leq |\overline{S}(f, Z)| \leq \|f\|_\infty \cdot \sum_{i=1}^{N-1} \big(H(t_{i+1}) - H(t_i)\big) = \|f\|_\infty \cdot \big(H(t_N) - H(t_1)\big) \ .$$

Hierbei ist $Z = \{-\infty = t_0 < t_1 < \cdots < t_N = \infty\}$ eine Zerlegung so, dass der Träger von f in $[t_1, t_N]$ enthalten ist. ∎

Mit obigem Lemma ist $(\mathbb{R}, C_c(\mathbb{R}), \int_{-\infty}^{\infty} dH)$ ein elementarer Integrationsraum und unsere Vervollständigungsprozedur liefert einen Lebesgue'schen Integrationsraum sowie ein induziertes reguläres Borelmaß.

Definition 6.29. Die Lebesgue-Vervollständigung des oben beschriebenen Integrationsraumes heißt *Lebesgue-Stieltjes-Integral*, nach wie vor bezeichnet mit

$$\int_{-\infty}^{\infty} f(t)dH(t).$$

Das induzierte reguläre Borelmaß μ_H heißt Lebesgue-Stieltjes-Maß.

Beispiele 6.30.

(a) Ist $H(t) = t$, so ist das zugehörige Riemann-Stieltjes-Integral gleich dem üblichen Riemann-Integral, während das Lebesgue-Stieltjes-Integral gerade das eindimensionale Lebesgue-Integral ist.

(b) Ist

$$H(t) = \begin{cases} 0 & t < 0 \\ 1 & t \geq 0 \end{cases}$$

die sogenannte *Heaviside-Funktion*, so ist das zugehörige Riemann-Stieltjes-Integral gerade die *Dirac-Distribution*

$$\int_{-\infty}^{\infty} f(t)dH(t) = f(0) \ .$$

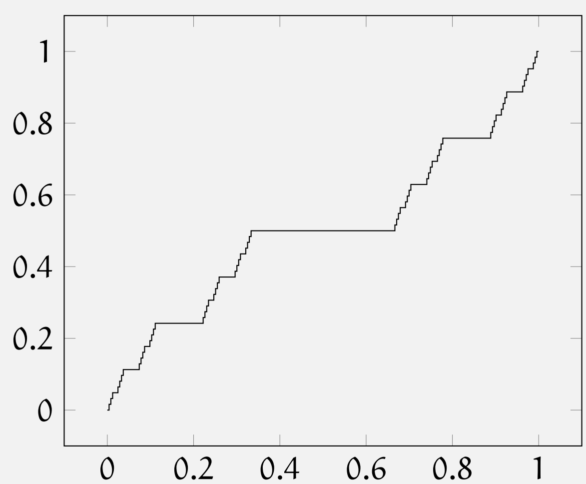
Das induzierte Lebesgue-Stieltjes-Maß ist das *Dirac-Maß*

$$\mu_H(A) = \begin{cases} 1 & 0 \in A \\ 0 & 0 \notin A \ . \end{cases} \qquad A \in \mathcal{B},$$

(c) Ein weiteres Beispiel einer Verteilungsfunktion ist die sogenannte *Teufelstreppe*. Hier setzt man

$$H(t) = \begin{cases} \frac{1}{2} & t \in (\frac{1}{3}, \frac{2}{3}) \\ \frac{1}{4} & t \in (\frac{1}{9}, \frac{2}{9}) \\ \frac{3}{4} & t \in (\frac{7}{9}, \frac{8}{9}) \\ \dots \end{cases} \quad \text{und} \quad H(t) = \begin{cases} 0 & t \leq 0 \\ 1 & t \geq 1, \end{cases}$$

analog zur Definition der Cantormenge $C \subseteq [0, 1]$ (siehe Abschnitt 4.4). Dies definiert H auf $\mathbb{R} \setminus C$. Man überprüft das H auf dieser dichten Teilmenge stetig ist; H besitzt dann eine eindeutige stetige Fortsetzung zu einer monoton steigenden, stetigen Funktion auf ganz $\mathbb{R}$.



Das Riemann-Stieltjes-Integral kann auch für gewisse nicht notwendigerweise monoton steigende Funktionen H definiert werden. Hierfür benötigen wir folgendes Resultat:

Lemma 6.31. Sei $H : \mathbb{R} \to \mathbb{R}$ eine beliebige Funktion. Dann sind die folgende Aussagen äquivalent:

(i) H ist die Differenz zweier monoton steigender Funktionen F und G.

(ii) H hat auf jedem Intervall $[a, b] \subseteq \mathbb{R}$ endliche Variation

$$\operatorname{Var}_a^b H = \sup \left\{ \sum_{i=1}^{N} \left| H(t_i) - H(t_{i-1}) \right| \;\middle|\; a = t_0 < t_1 < \cdots < t_N = b \right\} .$$

Zudem gilt: Im Fall, dass die obigen äquivalenten Bedingungen erfüllt sind, ist H rechtsseitig stetig wenn sowohl F als auch G rechtsseitig stetig (also Verteilungsfunktionen) sind.

Proof. (i) $\Longrightarrow$ (ii) Sei $H = F - G$ mit monoton steigenden Funktionen F und G. Dann gilt für jede Zerlegung $a = t_0 < t_1 < \cdots < t_N = b$

$$\begin{aligned} \operatorname{Var}_a^b H \leq \sum_{i=1}^{N} \left| H(t_i) - H(t_{i-1}) \right| &\leq \sum_{i=1}^{N} \left(\left| F(t_i) - F(t_{i-1}) \right| + \left| G(t_i) - G(t_{i-1}) \right| \right) \\ &= F(b) - F(a) + F(b) - G(a). \end{aligned}$$

(ii) $\Longrightarrow$ (i) Hat H endliche Variation auf $[a, b]$ so auch auf jedem Teilintervall. Für die Funktion

$$F(t) = \operatorname{Var}_a^t H$$

und $t < s$ gilt dann

$$F(s) - F(t) = \operatorname{Var}_a^s H - \operatorname{Var}_a^t H = \operatorname{Var}_t^s H \geq \left| H(s) - H(t) \right| \geq H(s) - H(t) .$$

Insbesondere ist F monoton wachsend und die Funktion $G = F - H$ ist auch monoton wachsend. Damit gilt $H = F - G$ mit monoton wachsenden Funktionen F und G.

Der Zusatz folgt daraus, dass für ein rechtsseitig stetiges H auch die Funktion $F(t) = \operatorname{Var}_a^t H$ rechtsseitig stetig ist. ∎

Satz 6.32. Sei H eine (nicht notwendig monoton steigende) Funktion mit beschränkter Variation und sei $H = F - G$ eine beliebige Zerlegung in monoton steigende Funktionen F und G. Dann ist die Mengenfunktion

$$\mu_H(A) := \mu_F(A) - \mu_G(A)$$

ein wohldefiniertes reguläres signiertes Borelmaß auf $\mathbb{R}$.

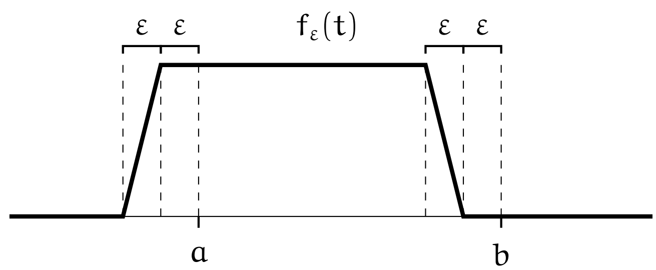
Beweis. Offenbar ist $\mu_H = \mu_F - \mu_G$ ein signiertes Maß. Wir müssen nur nachweisen, dass dies unabhängig von der Wahl der Zerlegung ist. Ist $H = F' - G'$ eine andere Zerlegung, so gilt $F + G' = F' + G$, also

$$\mu_F - \mu_G - (\mu_{F'} - \mu_{G'}) = \mu_{F+G'} - \mu_{G+F'} = 0. \qquad \blacksquare$$

Lemma 6.33. Sei H eine Verteilungsfunktion. Dann gilt

$$\begin{aligned} \mu_H\big([a,b]\big) &= H(b+) - H(a-), \\ \mu_H\big([a,b)\big) &= H(b-) - H(a-) \\ \mu_H\big((a,b]\big) &= H(b+) - H(a+) \\ \mu_H\big((a,b)\big) &= H(b-) - H(a+)\ , \end{aligned} \qquad \text{mit} \qquad \begin{aligned} H(s-) &= \lim_{t \nearrow s} H(t) \\ H(s+) &= \lim_{t \searrow s} H(t). \end{aligned}$$

Proof. Wir behandeln den Fall des halboffenen Intervalles $[a,b)$. Wir betrachten die Schar (f_ε) von affin-linearen Funktionen wie im untenstehenden Bild:



Dann konvergiert f_ε punktweise gegen $\mathbf{1}_{[a,b)}$ und aufgrund des Satzes über majorisierte Konvergenz (eine integrierbare Majorante ist die Funktion $\mathbf{1}_{[a-1,b+1]}$) gilt

$$\mu\big([a,b)\big) = \int_{-\infty}^{\infty} \mathbf{1}_{[a,b)}(x) dH(x) = \lim_{\varepsilon \to 0} \int_{-\infty}^{\infty} f_\varepsilon(t) dH(t)$$

Hierbei ist

$$\int_{-\infty}^{\infty} f_\varepsilon(t) dH(t) = \int_{a-2\varepsilon}^{a-\varepsilon} f_\varepsilon(t) dH(t) + H(b-2\varepsilon) - H(a-\varepsilon) + \int_{b-2\varepsilon}^{b-\varepsilon} f_\varepsilon(t) dH(t)\ .$$

Dabei gilt wegen der Abschätzung aus Lemma 6.28

$$\left| \int_{a-2\varepsilon}^{a-\varepsilon} f_\varepsilon(t) dH(t) \right| \leq H(a-\varepsilon) - H(a-2\varepsilon)\ .$$

Dies konvergiert gegen Null für $\varepsilon \to 0$, denn: Andernfalls existiert ein δ und eine Nullfolge (ε_n), so dass der Zuwachs von H auf jedem Intervall $[a-\varepsilon_n, a-\varepsilon_{n+1}]$ mehr als δ beträgt. Damit wäre H aber links von a unbeschränkt. Analog konvergiert das Integral über $[d-2\varepsilon, d-\varepsilon]$ gegen Null, so dass

$$\lim_{\varepsilon\to 0}\int_{-\infty}^{\infty} f_\varepsilon(t)dH(t) = H(b-) - H(a-)\ .$$

Die Aussage für die Intervalle $[a,b]$, $(a,b]$ und (a,b) werden analog bewiesen. ■

Satz 6.34. Für jedes reguläre Borelmaß μ auf $\mathbb{R}$ existiert eine Verteilungsfunktion H so, dass $\mu = \mu_H$. Eine mögliche Wahl ist

$$H(t) := \begin{cases} \mu\big((0,t]\big) & t \geq 0 \\ -\mu\big((-t,0]\big) & t \leq 0 \end{cases}$$

und jede andere Verteilungsfunktion H' mit $\mu = \lambda_{H'}$ hat die Form $H'(t) = H(t) + a$ für ein $a \in \mathbb{R}$.

Beweis. Sei μ ein reguläres Borelmaß. Wir definieren H durch die obige Formel. Dann ist H offenbar monoton wachsend. Für $t \geq a$ gilt

$$\lim_{s\searrow t} H(s) = \lim_{s\searrow t}\mu\big((a,t]\big) = \mu\left(\bigcap_{s>t}(a,s)\right) = \mu\big((a,t]\big) = H(t),$$

wobei wir für die mittlere Gleichheit Korollar 3.9 verwendet haben. Also ist H rechtsseitig stetig auf $[a,\infty)$ und die rechtsseitige Stetigkeit auf $(-\infty,a)$ folgt analog. Damit ist H eine Verteilungsfunktion.

Die Formeln von Lemma 6.33 zeigen, dass μ auf den halboffenen Intervallen mit μ_H übereinstimmt. Da diese die Borel'sche σ-Algebra erzeugen, müssen μ_H und μ überall übereinstimmen.

Ist nun H' eine beliebige Verteilungsfunktion so, dass $\mu = \lambda_{H'}$ ist, so gilt wegen Lemma 6.33

$$\mu\big((0,t]\big) = H'(t+) - H'(0+),$$

falls $t \geq 0$. Damit gilt $H'(t) = H(t) + H'(0+)$ für die oben definierte Funktion H. Die Aussage für $t < 0$ folgt analog. ■

Definition 6.35. Eine Funktion $H : [a,b] \to \mathbb{R}$ heißt *absolut stetig*, falls zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert, so dass für jedes Familie $(a_1,b_1),\dots,(a_N,b_N) \subseteq [a,b]$ von disjunkten offenen Intervallen in $[a,b]$ mit

$$\lambda\left(\bigcup_{i=1}^{N}(a_i,b_i)\right)=\sum_{i=1}^{N}(b_i-a_i)\le\delta$$

gilt

$$\sum_{i=1}^{N}\left|H(b_i)-H(a_i)\right|\le\varepsilon.$$

Eine Funktion $H:\mathbb{R}\to\mathbb{R}$ heißt (lokal) absolut stetig, falls ihre Einschränkung auf jedes beschränkte Intervall absolut stetig ist.

Bemerkung 6.36. Da die rechte Seite der Abschätzungen in der Definition der absoluten Stetigkeit unabhängig von N ist, gilt mit einer absolut stetigen Funktion H im Grenzwert auch die Abschätzung

$$\lambda\left(\bigcup_{i=1}^{\infty}(a_i,b_i)\right)\le\delta \qquad\Longrightarrow\qquad \sum_{i=1}^{\infty}\left|H(b_i)-H(a_i)\right|\le\varepsilon$$

für unendlich viele disjunkte Intervalle, die in einem gemeinsamen Intervall $[a,b]$ enthalten sind.

Beispiel 6.37. Jede Lipschitz-stetige Funktion (und spezieller jede stetig differenzierbare Funktion) ist absolut stetig.

Beispiel 6.38. Die Teufelstreppe ist nicht absolut stetig. Um dies zu sehen wählen wir $0<\varepsilon<1/2$. Dann Funktionswertesprung von Null auf $1/2$ erreichen wir auf zwei Intervallen der Länge $1/9$, mit Gesamtlänge

$$\frac{2}{9}=\frac{1}{3}\cdot\frac{2}{3}$$

aber bei genauem Hinsehen auch auf vier Intervallen der Länge $1/27$ oder allgemein auf 2^n Intervallen der Länge $1/3^{n+1}$, mit Gesamtlänge

$$\frac{2^n}{3^{n+1}}=\frac{1}{3}\cdot\left(\frac{2}{3}\right)^n\longrightarrow 0.$$

Daher kann es zu keine $\varepsilon<\frac{1}{2}$ ein passendes δ geben.

Lemma 6.39. Ist H eine absolut stetige Verteilungsfunktion, so ist das zugehörige Lebesgue-Stieltjes-Maß μ_H absolut stetig bezüglich des Lebesgue-Maßes.

Beweis. Wir zeigen, dass jede Lebesgue-Nullmenge auch eine Nullmenge für μ_H ist. Sei also $N \subseteq \mathbb{R}$ eine λ-Nullmenge. Wir zeigen, dass $\mu_H(N \cap (a, b)) = 0$ für jedes beschränkte Intervall (a, b).

Unter Benutzung der absoluten Stetigkeit von H sei zu gegebenem $\varepsilon > 0$ ein $\delta > 0$ so, dass

$$\lambda\left(\bigcup_{i=1}^{N}(a_i, b_i)\right) \leq \delta \qquad \Longrightarrow \qquad \sum_{i=1}^{N}\big(H(b_i) - H(a_i)\big) \leq \varepsilon.$$

gilt für beliebige paarweise disjunkte Intervalle (a_i, b_i), die alle in (a, b) enthalten sind.

Wegen der Regularität des Lebesgue-Maßes existiert nun eine offene Menge $U \subseteq (a, b)$, die $N \cap (a, b)$ enthält, und so, dass $\lambda(U) \leq \varepsilon$. Als offene Menge in $\mathbb{R}$ ist U eine disjunkte Vereinigung (möglicherweise abzählbar unendlich vieler) offener Intervalle, $U = \bigcup_{i=1}^{\infty}(a_i, b_i)$. Damit gilt

$$\mu_F\big(N \cap (a, b)\big) \leq \mu_F(U) = \sum_{i=1}^{\infty} \mu_F\big((a_i, b_i)\big) = \sum_{i=1}^{\infty} \big(F(b_i) - F(a_i)\big) \leq \varepsilon,$$

wobei wir in der zweiten Gleichheit Lemma 6.33 benutzt haben und im Anschluss Bemerkung 6.36. Da dies für jedes $\varepsilon > 0$ gilt, folgt $\mu_F(N \cap (a, b)) = 0$. ∎

Hauptsatz der Differential- und Integralrechnung 6.40. Eine stetige Funktion $H : \mathbb{R} \to \mathbb{R}$ ist genau dann absolut stetig, wenn eine lokal integrierbare Funktion $h : \mathbb{R} \to \mathbb{R}$ exisitiert so dass

$$H(t) = H(0) + \int_0^t h(s)ds. \tag{6.4.2}$$

Hier bei bedeutet lokal integrierbar, dass die Einschränkungen $h|_A$ für jede beschränkte Teilmenge $A \subseteq \mathbb{R}$ (Lebesgue-)integrierbar sind.

Beweis. ($\Longrightarrow$) Dass H absolut stetig ist, impliziert insbesonder, dass H auf jedem Teilintervall $[a, b] \subseteq \mathbb{R}$ beschränkte Variation hat. Nach Lemma 6.31 können gilt also $H = F - G$ mit Verteilungsfunktionen F und G. Wir können F und G so wählen, dass sie nach wie vor absolut stetig sind: Dies ist beispielsweise der Fall wenn wir wie im Beweis von Lemma 6.31 explizit $F(t) = \mathrm{Var}_a^t H$ setzen. Seien μ_F und μ_G die zugehörigen Lebesgue-Stieltjes-Maße.

Aus der absoluten Stetigkeit von F und G folgt mit Lemma 6.39 dass μ_F und μ_G absolut stetig bezüglich des Lebesgue-Maßes sind. Daher existieren nach dem Satz von Radon-Nikodym Funktionen f und g mit

$$\mu_F(A) = \int_A f(x)dx, \qquad \mu_G(A) = \int_A g(x)dx.$$

Mit $h := f - g$ gilt dann (6.4.2).

($\Longleftarrow$) Ist h lokal integrierbar, so ist das durch

$$\lambda_{|h|}(A) := \int_A |h(t)|dt$$

gegebene Maß absolut stetig bezüglich des Lebesgue-Maßes (Lemma 6.20). Zu jedem $\varepsilon > 0$ und jedem Intervall $I \subseteq \mathbb{R}$ existiert daher ein $\delta > 0$ so, dass $\lambda(A) \leq \lambda_{|h|}(A)$ für alle Borel-messbaren Mengen $A \subseteq I$. Wegen

$$|H(t) - H(s)| \leq \int_s^t |h(u)|du = \lambda_{|h|}([s,t])$$

ist also die durch (6.4.2) gegebene Funktion absolut stetig. ■

Wir halten folgendes Korollar aus obigem Satz fest:

Korollar 6.41. x Sei H eine Verteilungsfunktion auf $\mathbb{R}$ mit zugehörigem Maß μ_H. Dann sind äquivalent:

(i) H ist absolut stetig;

(ii) $\mu_H \ll \lambda$;

(iii) Es existiert eine nichtnegative lokal integrierbare Funktion $h : \mathbb{R} \to \mathbb{R}$ mit

$$H(t) = H(0) + \int_0^t h(u)du$$

Definition 6.42. Ein Verteilungsfunktion H heißt *diskret*, falls der Wertebereich $\{H(t) \mid t \in \mathbb{R}\}$ endlich oder abzählbar ist.

Das zu einer diskreten Verteilungsfunktion H gehörige Maß μ_H ist eine abzählbare Linearkombination von Diracmaßen (s. Beispiel 8.11 (b)).

Zerlegungssatz 6.43. Jedes Borelmaß μ auf $\mathbb{R}$ hat eine eindeutige Zerlegung

$$\mu = \lambda_h + \delta + \rho,$$

in paarweise orthogonale Maße, wobei λ_h absolutstetig bezüglich des Lebesgue-Maßes ist δ absolutstetig ist.

Beweis. Sei

$$T := \{t \in \mathbb{R} \mid \mu(\{t\}) \neq 0\}.$$

Da μ regulär ist, muss T abzählbar sein. Ein Maß δ mit einer diskreten Verteilungsfunktion wird nun definiert durch

$$\delta(A) = \sum_{t \in T} \mu(\{t\}) \cdot \delta_t.$$

Die Differenz $\mu - \delta$ ist nun wieder ein (positives) Maß, da

$$\mu(A) - \delta(A) = \mu(A \setminus T) + \mu(T) - \delta(T) = \mu(A \setminus T) \geq 0$$

gilt.

Sei nun H eine Verteilungsfunktion so dass $\mu - \delta = \mu_H$. Nach Konstruktion ist μ_H *atomfrei*, erfüllt also $\mu_H(\{t\}) = 0$ für alle $t \in \mathbb{R}$. Mit den Formeln aus Lemma 6.33 gilt dann

$$H(t+) - H(t-) = \lim_{s \searrow t} H(s+) - H(t-) = \lim_{s \searrow t} \mu_H\big([t, s]\big) = \mu_H\left(\bigcap_{s > t} [t, s]\right) = \mu_H(\{t\}) = 0,$$

also ist H stetig.

Nach dem Satz von Radon-Nikodym gilt nun $\mu_H = \lambda_h + \rho$ für eine messbare Funktion h und ein zu λ_h orthogonales Maß ρ. Wegen Korollar 6.41 ist jede Verteilungsfunktion zu λ_h absolut stetig. Damit ist auch die Verteilungsfunktion zu ρ stetig. Dies impliziert, dass sowohl λ_h also auch ρ orthogonal zu δ sind. Denn da beide stetige Verteilungsfunktionen haben, ist die Menge T eine Nullmenge, also ist $\mathbb{R} \setminus T$ eine Trägermenge für λ_h und ρ, während T eine Trägermenge für δ ist. ■

Bemerkung 6.44. Nach Satz 6.34 besitzt das Maß ρ aus dem Zerlegungssatz eine Verteilungsfunktion. Man kann diejenigen Funktionen die als Verteilungsfunktionen von ρ auftreten, charakterisieren als diejenigen Funktionen, die stetig und λ-fast überall differenzierbar sind mit Ableitung Null.

Es folgt, dass jede Verteilungsfunktion $H : \mathbb{R} \to R$ eine Zerlegung in

$$F = F_1 + F_2 + F_3$$

besitzt, so dass gilt:

- F_1 ist absolut stetig. Das zugehörige Maß besitzt dann eine Dichte bezüglich des Lebesgue-Maßes.
- F_2 ist diskret (also lokalkonstant auf $\mathbb{R} \setminus T$, wobei $T \subseteq \mathbb{R}$ abzählbar ist). Das prototypische Beipiel ist die Heaviside-Funktion.
- F_3 ist stetig und fast überall differenzierbar mit Ableitung Null, also lokalkonstant auf dem komplement einer Nullmenge. Das prototypische Beispiel ist die Teufelstreppe.

7 Die Lebesgue-Räume

7.1 Definition der Lebesgue-Räume

Wir sagen, dass eine Funktion auf einem Maß- oder Integrationsraum eine Eigenschaft "fast überall" besitzt, wenn diese Eigenschaft auf dem Komplement einer Nullmenge erfüllt ist.

In diesem Abschnitt sei $(X, \mathcal{A}, \mu)$ ein σ-endlicher Maßraum. Wir schreiben

$$\int_X f(x) d\mu(x)$$

für das zugehörige Integral. Weiterhin sei $\mathcal{M}$ das zugehörige System der messbaren Mengen.

Definition 7.1 ($\mathcal{L}^p$-Räume). Für $1 \leq p < \infty$ definieren wir

$$\mathcal{L}^p(X, \mu) = \left\{ f \in \mathcal{M} \text{ endlich} \mid |f|^p \in \mathcal{L} \right\} .$$

Insbesondere ist $\mathcal{L}^1(X, J) = \mathcal{L}$. Weiterhin definieren wir

$$\mathcal{L}^\infty(X, \mu) = \left\{ f \in \mathcal{M} \text{ fast überall beschränkt} \right\} .$$

Für $f \in \mathcal{L}^p(X, \mu)$ definieren wir die L^p-*Norm* durch

$$\|f\|_{L^p} := \left(\int_X |f|^p d\mu(x) \right)^{1/p}$$

falls $p < \infty$ und

$$\|f\|_{L^\infty} := \operatorname{ess\,sup} |f| = \inf_{\substack{N \subseteq X \\ \text{Nullmenge}}} \sup_{x \in X \setminus N} |f(x)| \,.$$

Lemma 7.2. Für alle $1 \leq p \leq \infty$ gilt die *Minkowskiungleichung*

$$\|f + g\|_{L^p} \leq \|f\|_{L^p} + \|g\|_{L^p} \,.$$

Die Minkowski-Ungleichung besagt, dass die L^p-Normen die Dreiecksungleichung erfüllt. Allerdings sind sie im Fall der Existenz nichttrivialer Nullmengen keine Norm, da beispielsweise für jede Nullmenge die Indikatorfunktion $\mathbf{1}_N$ in $\mathcal{L}^p(X, \mu)$ enthalten ist (für jedes $1 \leq p \leq \infty$) mit $\|f\|_{L^p} = 0$. Wir erhalten also tatsächlich nur eine Halbnorm. Um dies zu reparieren macht man die folgende Definition.

Definition 7.3 (L^p-Räume)**.** Zwei Funktionen $f, h \in \mathcal{M}$ heißen *fast überall gleich*, wenn sie auf dem Komplement einer Nullmenge übereinstimmen. Dies definiert eine Äquivalenzrelation, und wir bezeichnen die Äquivalenzklassen mit $[f]$.

Sei $\mathcal{N} \subseteq \mathcal{M}$ die Menge der Funktionen, die fast überall Null sind. Für $1 \leq p \leq \infty$ schreiben wir

$$L^p(X, \mu) := \mathcal{L}^p(X, \mu) / \big(\mathcal{L}^p(X, \mu) \cap \mathcal{N}\big)$$

für den normierten Raum der Äquivalenzklassen.

Satz 7.4. Für jedes $1 \leq p \leq \infty$ sind die Räume $L^p(X, \mu)$ Banachräume.

Zum Beweis nutzen wir folgendes Lemma.

Lemma 7.5. Für einen normierten Raum E sind folgende Aussagen äquivalent:

(i) E ist vollständig (ein Banachraum), d.h. jede Cauchyfolge in E konvergiert.

(ii) Alle absolut konvergenten Reihen in E konvergieren.

Beweis. Die Implikation $(\Longrightarrow)$ ist trivial, da die die Folge der Partialsummen einer absolut konvergenten Reihe eine Cauchyfolge bildet.

Für die Implikation $(\Longleftarrow)$ sei (x_n) eine Cauchyfolge in E. Wir wählen eine Teilfolge

$(x_{n_k})_k$ so, dass $\|x_{n_{k+1}} - x_{n_k}\| \leq 2^{-k}$. Mit $y_k := x_{n_{k+1}} - x_{n_k}$ gilt dann

$$\sum_{k=1}^{\infty} \|y_k\| \leq \sum_{k=1}^{\infty} 2^{-k} < \infty ,$$

also ist die Reihe $\sum_{k=1}^{\infty} y_k$ absolut konvergent, konvergiert also nach Annahme gegen ein $y \in E$. Damit existiert für jedes $\varepsilon > 0$ ein N_0 so dass für jedes $N \geq N_0$ gilt

$$\varepsilon > \left\| y - \sum_{k=1}^{N} y_k \right\| = \left\| y - (x_{n_N} - x_{n_1}) \right\| = \left\| (y + x_{n_1}) - x_{n_N} \right\| .$$

Damit konvergiert die Cauchyfolge $(x_{n_k})_k$ gegen $x := y + x_{n_1}$. Da eine Cauchyfolge mit einer konvergenten Teilfolge selber konvergiert (und zwar gegen denselben Grenzwert) konvergiert auch (x_n) gegen E. ■

Beweis von Satz 7.4. Wir weisen das Kriterium von Lemma 7.5 nach. Sei $\sum_{n=1}^{\infty} f_n$ eine absolut konvergente Reihe, also

$$\sum_{n=1}^{\infty} \|f_n\|_{L^p} < \infty .$$

Betrachte die punktweise definierte Funktion $F := \sum_{n=1}^{\infty} |f_n|$, die in $\mathcal{L}^*$ liegt, aber den Wert unendlich annehmen kann. Wegen

$$\int_X^* |F(x)|^p \, d\mu(x) = \|F\|_{L^p}^p \leq \left(\sum_{n=1}^{\infty} \|f_n\|_{L^p} \right)^p < \infty$$

ist $|F|^p$ eine Levifunktion. Insbesondere ist die Menge $N = \{F(x) = \infty\}$ eine Nullmenge. Für alle $x \in X$ konvergiert damit die Reihe

$$f(x) := \sum_{n=1}^{\infty} f_n(x) \cdot \mathbf{1}_{X \setminus N}(x)$$

absolut (dies ist einfach die Vollständigkeit von $\mathbb{C}$). Die Funktion f ist als punktweiser Grenzwert von messbaren Funktionen messbar.

Wegen $|f|^p \leq |F|^p$ ist auch das verallgemeinerte Integral von $|f|^p$ endlich, damit ist $|f|^p \in \mathcal{L}$ nach Lemma 2.28, also $f \in \mathcal{L}^p(X, \mu)$.

Es bleibt zu zeigen, dass die Reihe der f_n gegen f konvergiert. Hier gilt für alle $x \in X \setminus N$

$$\left| \sum_{n=1}^{N} f_n(x) - f(x) \right|^p = \left| \sum_{n=N+1}^{\infty} f_n(x) \right|^p \leq \left| \sum_{n=N+1}^{\infty} |f_n(x)| \right|^p \leq |F(x)|^p ,$$

also ist die integrierbare Funktion $|F|^p \cdot \mathbf{1}_{X\setminus N}$ eine Majorante, und es gilt

$$\left\| \sum_{n=1}^{N} f_n(x) - f(x) \right\|_{L^p}^p = \int_X \left| \sum_{n=1}^{N} f_n(x) - f(x) \right|^p d\mu(x) \longrightarrow 0$$

nach dem Satz von Lebesgue über dominierte Konvergenz, da der Integrand punktweise gegen Null konvergiert. ■

7.2 Vollständigkeit

In diesem Abschnitt ist $(X, \mathcal{A}, \mu)$ stets ein σ-endlicher Maßraum.

Lemma 7.6 (Hölder-Ungleichung). Seien $p, q \in [1, \infty]$ mit

$$\frac{1}{p} + \frac{1}{q} = 1 \,. \tag{7.2.1}$$

Dann ist für $f \in \mathcal{L}^p(X, \mu)$, $g \in \mathcal{L}^q(X, \mu)$ die messbare Funktion $f \cdot g$ integrierbar und es gilt

$$\|f \cdot g\|_{L^1} \leq \|f\|_{L^p} \cdot \|g\|_{L^q} \,.$$

Definition 7.7. Zahlen $p, q \in [1, \infty]$ die die Gleichung (7.2.1) heißen *konjugierte Exponenten*.

Bemerkung 7.8. Eine wichtige Verallgemeinerung der Hölder-Ungleichung ist die Ungleichung

$$\|f \cdot g\|_{L^r} \leq \|f\|_{L^p} \cdot \|g\|_{L^q}$$

für $f \in \mathcal{L}^p(X, \mu)$ und $g \in \mathcal{L}^q(X, \mu)$, wobei

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

Denn für solche p, q, r sind die Exponenten $p' = p/r$ und $q' = q/r$ konjugiert, also gilt nach der üblichen Hölderungleichung

$$\|f \cdot g\|_{L^r}^r = \left\| |f \cdot g|^r \right\|_{L^1} \leq \left\| |f|^r \right\|_{L^{p'}} \cdot \left\| |g|^r \right\|_{L^{q'}} = \|f\|_{L^p}^r \cdot \|g\|_{L^q}^r.$$

Darstellungssatz von Riesz 7.9. Seien $p \in (1, \infty]$ und sei q der konjugierte Exponent. Ist $f \in \mathcal{L}^p(X, \mu)$, so wird durch die Vorschrift

$$\varphi_f : g \longmapsto \int_X g(x) \cdot f(x) d\mu(x)$$

ein stetiges lineares Funktional auf $L^q(X, \mu)$ definiert, und die resultierende Abbildung

$$\Phi : L^p(X, \mu) \longrightarrow L^q(X, \mu)' , \qquad f \longmapsto \varphi_f$$

ein isometrischer Isomorphismus.

Bemerkung 7.10. Der Raum $L^2(X, \mu)$ ist sogar ein Hilbertraum mit dem Skalarprodukt

$$\langle f, g \rangle = \int_X f(x) \cdot g(x) dx.$$

Bemerkung 7.11. Der Darstellungssatz von Riesz gilt *nicht* für $p = 1$. Die Abbildung

$$L^1(X, \mu) \longrightarrow L^\infty(X, \mu)'$$

aus dem Satz von Riesz existiert nach wie vor und ist isometrisch (also insbesondere injektiv), aber nicht surjektiv.

Beweis. Nach der Hölder-Ungleichung ist φ_g eine wohldefinierte Abbildung und es gilt

$$|\varphi_f(g)| \leq \int_X |f(x)g(x)| d\mu(x) = \|f \cdot g\|_{L^1} \leq \|f\|_{L^p} \cdot \|g\|_{L^q} .$$

Also ist φ_f stetig mit Norm $\|\varphi_f\| \leq \|f\|_{L^p}$. Zwei Funktionen $f_1, f_2 \in \mathcal{L}^p(X, \mu)$, die fast überall übereinstimmen, induzieren offenbar dasselbe Funktional, also ist Φ eine wohldefinierte lineare Abbildung mit $\|\Phi\| \leq 1$.

Um zu zeigen, dass Φ eine Isometrie ist, müssen wir $\|\varphi_f\| = \|f\|_{L^p}$ für alle $f \in \mathcal{L}^p(X, \mu)$ nachweisen. Wir benutzen hierzu, dass für konjugierte Exponenten stets die Gleichung $(p-1)q = p$ gilt. Konkret sei $f \in \mathcal{L}^p(X, \mu)$ mit $\|f\|_{L^p} = 1$ beliebig. Wir setzen $A := \{f(x) \neq 0\}$ und

$$g := |f|^{p-2} \cdot f \cdot \mathbf{1}_A . \tag{7.2.2}$$

Dann gilt

$$\int_X^* |g(x)|^q d\mu(x) = \int_A^* |f(x)|^{(p-1)q} d\mu(x) = \int_A^* |f(x)|^p d\mu(x) = \int_X^* |f(x)|^p d\mu(x) .$$

Also ist $|g|^q$ integrierbar und $\|g\|_{L^q}^q = \|f\|_{L^p}^p = 1$. Zudem ist

$$\varphi_f(g) = \int_X g(x) \cdot f(x) d\mu(x) = \int_A |f(x)|^p d\mu(x) = 1 ,$$

also folgt $|\varphi_f(g)| = \|f\|_{L^p} \cdot \|g\|_{L^q}$ und damit $\|\varphi_f\| \geq \|f\|_{L^p}$.

Beweis der Surjektivität. Wir betrachten zur Einfachheit nur den Fall $\mu(X) < \infty$. Sei $\varphi \in L^q(X, \mu)'$ ein beliebiges stetiges lineares Funktional. Wir behaupten, dass durch

$$\nu(A) := \varphi(\mathbf{1}_A), \qquad A \in \mathcal{A} ,$$

ein signiertes Maß auf X definiert wird. Hierbei ist $\chi_A \in \mathcal{L}^q(X, \mu)$, da $\mu(X) < \infty$ vorausgesetzt wurde. Offenbar ist ν additiv. Sei $A_1, A_2, \ldots$ eine Familie von disjunkten Teilmengen von X mit Vereinigung A. Dann gilt

$$\begin{aligned} \left\| \mathbf{1}_A - \sum_{i=1}^{n} \mathbf{1}_{A_i} \right\|_{L^q}^q &= \int_X \left(\mathbf{1}_A - \sum_{i=1}^{n} \mathbf{1}_{A_i}(x) \right)^q d\mu(x) \\ &= \int_X \sum_{i=n+1}^{\infty} \mathbf{1}_{A_i}(x) d\mu(x) = \mu \left(\bigcup_{i=n+1}^{\infty} A_i \right) \longrightarrow 0 , \end{aligned}$$

Also konvergiert die Reihe $\sum_{i=1}^{\infty} \mathbf{1}_{A_i}$ in der L^q-Norm gegen $\mathbf{1}_A$. Wegen der Stetigkeit von φ folgt also

$$\sum_{i=1}^{\infty} \nu(A_i) = \sum_{i=1}^{\infty} \varphi(\mathbf{1}_{A_i}) = \varphi \left(L^p\text{-} \sum_{i=1}^{\infty} \mathbf{1}_{A_i} \right) = \varphi(\mathbf{1}_A) = \nu(A) .$$

Dies zeigt die σ-Additivität, also ist ν ein signiertes Maß.

Sei $\nu = \nu^+ - \nu^-$ die Jordan-Zerlegung von ν. Direkt aus der Definition folgt, dass μ-Nullmengen auch Nullmengen für ν^+ und ν^- sind. Also folgt aus dem Spezialfall 6.23 des Satzes von Radon-Nikodym, dass $\nu^\pm = \mu_{f^\pm}$ für nichtnegative messbare Funktionen $f^\pm$ gilt. Ist $X = X^+ \cup X^-$ eine Hahn-Zerlegung für ν, so gilt

$$\int_X f^\pm(x) d\mu(x) = \nu^\pm(X) = \nu^\pm(X^\pm) = \varphi(\mathbf{1}_{X^\pm}) \leq \|\varphi\| \cdot \|\mathbf{1}_{X^\pm}\|_{L^q} < \infty ,$$

also sind f^+ und f^- sogar integrierbar, und damit auch $f := f^+ - f^-$.

Wir behaupten nun, dass für alle beschränkten messbaren Funktionen g gilt, dass

$$\varphi(g) = \int_X g(x) \cdot f(x) d\mu(X) \tag{7.2.3}$$

Nach Konstruktion von f folgt dies für alle Indikatorfunktionen g und damit auch für alle Treppenfunktionen. Sei nun $g \in \mathcal{L}^\infty(X, \mu)$ eine beliebige beschränkte messbare Funktion. Dann existiert eine beschränkte Folge (g_n) von Treppenfunktionen, die punktweise gegen g konvergiert. Dann folgt aus dem Satz von Lebesgue über dominierte Konvergenz, dass

$$\varphi(g_n) = \int_X g_n(x) \cdot f(x) d\mu(x) \longrightarrow \int_X g(x) \cdot f(x) d\mu(x)$$

sowie dass $\|g_n - g\|_{L^q}$ gegen Null konvergiert. Hierbei kann eine konstante Funktion als integrierbare Majorante nehmen, da $\mu(X)$ als endlich vorausgesetzt war. Weil φ stetig bezüglich der L^q-Norm ist, folgt damit auch $\varphi(g_n) \to \varphi(g)$, was die Behauptung beweist.

Es bleibt zu zeigen, dass $f \in \mathcal{L}^p(X, \mu)$, und dass (7.2.3) sogar für alle $g \in L^q(X, \mu)$ gilt. Im Fall dass $q = \infty$, also $p = 1$, sind wir fertig, da wir schon wissen, dass $f \in \mathcal{L}^1(X, \mu)$. Im Fall $q < \infty$ sei g zu gegebenem f die in (7.2.2) definierte Funktion und sei $B_n := \{|g(x)| \leq n\}$. Dann ist $\mathbf{1}_{B_n} \cdot g$ beschränkt und messbar, also gilt mit dem vorigen Schritt

$$\begin{aligned}\int_{B_n} |f(x)|^p d\mu(x) &= \int_X \mathbf{1}_{B_n}(x) \cdot g(x) \cdot f(x) d\mu(x) = \varphi(\mathbf{1}_{B_n} \cdot g) \leq \|\varphi\| \|\mathbf{1}_{B_n} g\|_{L^q} \\ &= \|\varphi\| \cdot \left(\int_{B_n} |g(x)|^q d\mu(x) \right)^{1/q} = \|\varphi\| \cdot \left(\int_{B_n} |f(x)|^p d\mu(x) \right)^{1/q}\end{aligned}$$

Es folgt, dass $\|\mathbf{1}_{B_n} \cdot f\|_{L^p} \leq \|\varphi\| < \infty$. Da andererseits punktweise gilt $\mathbf{1}_{B_n} \cdot |f|^p \nearrow |f|^p$ folgt aus dem Satz von Beppo Levi, dass $f \in \mathcal{L}^p(X)$ mit $\|f\|_{L^p} \leq \|\varphi\|$.

Indem man ein allgemeines $g \in L^q(X, \mu)$ wie eben als Grenzwert von beschränkten Funktionen schreibt, folgt die Darstellung (7.2.3) folgt damit auch für alle $g \in \mathcal{L}^q(X, \mu)$. Damit gilt $\varphi = \varphi_f$. Dies zeigt die Surjektivität.

Der allgemeine Fall $\mu(X) = \infty$ wird wieder auf den Fall endlichen Maßes zurückgeführt. ■

7.3 Konvergenzarten

$(X, \mathcal{A}, \mu)$ sei wieder ein σ-endlicher Maßraum.

Definition 7.12. Sei (f_n) eine Folge von messbaren Funktionen auf X. Sei f eine weitere messbare Funktion auf X. Wir schreiben:

(a) $f_n \xrightarrow{\text{f.ü.}} f \iff \mu(\{f_n(x) \nrightarrow f(x)\}) = 0.$

(Die Menge der Punkte, an denen wo f_n nicht gegen f konvergiert ist eine Nullmenge.)

Wir sagen f_n *konvergiert fast überall gegen* f.

(b) $f_n \xrightarrow{L^p} f \iff \|f_n - f\|_{L^p} \to 0$

Wir sagen f_n *konvergiert gegen* f *im* p*-ten Mittel*.

(c) $f_n \xrightarrow{\mu} f \iff \forall \varepsilon > 0 : \mu(\{|f_n(x) - f(x)| \geq \varepsilon\}) \to 0$

(Das Maß der Punktmengen, wo die Differenz $|f_n - f|$ größer als ε ist, konvergiert gegen Null.)

Wir sagen f_n *konvergiert im Maß gegen* f.

Satz 7.13. Es gelten die folgenden Beziehungen zwischen obigen Konvergenzarten:

(1) Wenn $f_n \xrightarrow{L^p} f$, so existiert eine Teilfolge (f_{n_k}) mit $f_{n_k} \xrightarrow{\text{f.ü.}} f$.

(2) Wenn $f_n \xrightarrow{\text{f.ü.}} f$ und falls (f_n) eine integrierbare Majorante zulässt, so gilt $f_n \xrightarrow{L^p} f$.

(3) Wenn $f_n \overset{\text{f.ü.}}{\nearrow} f$, so gilt $f_n \xrightarrow{L^p} f$.

(4) Wenn $f_n \xrightarrow{L^p} f$, so gilt $f_n \xrightarrow{\mu} f$.

(5) Wenn $f_n \xrightarrow{\text{f.ü.}} f$ und $\mu(X) < \infty$, so gilt $f_n \xrightarrow{\mu} f$.

(6) Wenn $f_n \xrightarrow{\mu} f$ und $\mu(X) < \infty$, so existiert eine Teilfolge (f_{n_k}) mit $f_{n_k} \xrightarrow{\mu} f$.

(7) Wenn $f_n \xrightarrow{L^p} f$ und $\mu(X) < \infty$, so gilt $f_n \xrightarrow{L^q} f$ für alle $q \leq p$.

Wir erhalten also folgendes Diagram der Konvergenzbeziehungen:

$$\begin{tikzcd}
 & f_n \xrightarrow{L^p} f \arrow[dl, Rightarrow] \arrow[dr, Rightarrow, "\text{Teilfolge}"'] & & f_n \xrightarrow{L^\infty} f \arrow[ll, Rightarrow, "\text{falls } \mu(X) < \infty"'] \\
f_n \xrightarrow{\mu} f \arrow[rr, Rightarrow, shift left, "\text{Teilfolge, falls } \mu(X) < \infty"] & & f_n \xrightarrow{\text{f.ü.}} f \arrow[ll, Rightarrow, shift left, "\text{falls } \mu(X) < \infty"] \arrow[ul, Rightarrow, dashed, shift right, "\text{Satz von Beppo Levi / majorisierte Konvergenz}"'] \arrow[ur, Rightarrow, dashed, "\text{Satz von Egoroff}"'] &
\end{tikzcd} \tag{7.3.1}$$

Beweis. (1) Die Konstruktion einer entsprechenden Teilfolge geht ähnlich wie im Beweis von Satz 7.4.

(2) Dies ist des Satz von Lebesgue über die dominierte Konvergenz 2.31.

(3) Dies ist im Wesentlichen der Satz von Beppo-Levi 2.24 (allerdings muss man die Funktionenfolge auf einer Nullmenge modifizieren um ihn wörtlich anzuwenden).

(4) Sei $f_n \xrightarrow{L^p} f$. Und setze $A_n = \{|f_n(x) - f(x)| \geq \varepsilon\}$. Sei zunächst $p < \infty$. Für $\varepsilon > 0$ gilt dann

$$\mu\big(A_n\big) \cdot \varepsilon^p = \int_{A_n} \varepsilon^p d\mu(x) \leq \int_{A_n} \big|f_n(x) - f(x)\big|^p dx \leq \|f_n - f\|^p \longrightarrow 0.$$

Für $p = \infty$ ist nach Definition der L^∞-Norm für jedes $\varepsilon > 0$ die Menge A_n für n groß genug eine Nullmenge. Also ist $\mu(A_n)$ sogar gleich Null für n groß genug.

(5) Die Funktionenfolge f_n konvergiere fast überall gegen f. Zu $\varepsilon > 0$ müssen wir zeigen, dass $\mu(A_n) \to 0$, wobei A_n wie eben definiert ist. Hierzu setzen wir

$$B_n := \bigcup_{i=n}^{\infty} A_i, \qquad B = \bigcap_{n=1}^{\infty} B_n.$$

Wir behaupten, dass $B \subseteq \{x \mid f_n(x) \nrightarrow f(x)\}$. In der Tat, ist $x \in B$, so ist $x \in B_n$ für alle $n \in \mathbb{N}$, also existiert für alle $n \in \mathbb{N}$ ein $i \geq n$ mit $x \in A_i$, also $|f_i(x) - f(x)| \geq \varepsilon$. Damit kann $f_n(x)$ nicht gegen $f(x)$ konvergieren und die Behauptung ist bewiesen.

Es folgt, dass

$$\mu(B) \leq \mu\big(\big\{x \mid f_n(x) \nrightarrow f(x)\big\}\big) = 0,$$

B ist also eine Nullmenge. Wegen $\mu(X) < \infty$ gilt auch $\mu(B_1) < \infty$, also wegen der Abwärtsstetigkeit von μ auch

$$\limsup_{n\to\infty} \mu(A_n) \leq \lim_{n\to\infty} \mu(B_n) = \mu(B) \leq \mu\big(\big\{x \mid f_n(x) \nrightarrow f(x)\big\}\big) = 0.$$

Damit konvergiert $\mu(A_n)$ gegen Null.

(6) Es konvergiere f_n dem Maße nach gegen f. Dann wähle zu jedem $k \in \mathbb{N}$ einen Index $n = n_k$ so groß, dass

$$\mu\Big(\underbrace{\big\{|f_{n_k}(x) - f(x)| \geq \tfrac{1}{k}\big\}}_{=:A_k}\Big) \leq \frac{1}{2^k}$$

Es gilt dann

$$\big\{f_{n_k}(x) \nrightarrow f(x)\big\} \subseteq \bigcap_{l=1}^{\infty}\bigcup_{k=l}^{\infty} A_k,$$

da wenn $f_{n_k}(x) \nrightarrow f(x)$ gilt, die Beziehung $f_{n_k}(x) \geq 1/k$ unendlich oft erfüllt sein muss. Da nach Wahl der Teilfolge $\sum_{k=1}^{\infty} A_k < \infty$ gilt, ist die rechte Seite nach dem Lemma von Borel-Cantelli eine Nullmenge.

(7) Sei $q < p$ und $u \in [0, \infty)$ mit

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

Wegen Bemerkung 7.8 gilt dann für $f \in \mathcal{L}^p(X, \mu)$, dass

$$\|f\|_{L^q} = \|f \cdot \mathbf{1}\|_{L^q} \leq \|f\|_{L^p} \cdot \|\mathbf{1}\|_{L^u} = \|f\|_{L^p} \cdot \mu(X)^{1/u}.$$

Damit ist ist die L^p-Norm stärker als die L^q-Norm. Daraus folgt die Behauptung. ∎

Beispiele 7.14. Wir geben Beispiele dafür, dass die Konvergenzbeziehungen aus Diagram (7.3.1) im allgemeinen nicht umkehrbar sind.

(a) Ein Beispiel für eine Folge (f_n), die dem Maße nach, aber nicht in L^p konvergiert, ist die Folge
$$f_n(x) = \begin{cases} n & x \leq \frac{1}{n} \\ 0 & x > \frac{1}{n}. \end{cases}$$
Diese Folge konvergiert auch fast überall gegen Null (für das Lebesgue-Maß auf $[0, 1]$), ist also auch ein Beispiel für die Notwendigkeit der integrierbaren Majorante im Satz von Lebesgue.

(b) Für $l, k \in \mathbb{N}$ mit $0 \leq k \leq 2^l - 1$ betrachten wir die Menge
$$A_{2^l+k} = \{x \mid k \leq 2^l x < k + 1\} \subseteq [0, 1],$$
wobei wir $[0, 1]$ wiederum mit dem Lebesgue-Maß ausstatten. Dann konvergiert die Folge der Indikatorfunktionen der Mengen A_n dem Maße nach gegen Null (da $\mu(A_n) \to 0$), aber die Mengen sind so gebaut, dass jeder Punkt aus $[0, 1)$ in unendlich vielen A_n enthalten ist. Also konvergiert die Folge der Indikatorfunktionen nur im Punkt $x = 1$ (inbesondere also nicht fast überall).

Die obige Folge konvergiert auch bezüglich jeder L^p-Norm. Es ist nicht schwer, eine fast überall konvergente Teilfolge zu finden.

(c) Ein Beispiel, das die Notwendigkeit der Bedingung $\mu(X) < \infty$ in Satz 7.13(5) veranschaulicht ist die Folge der Indikatorfunktionen der Mengen $[n, n + 1]$, $n \in \mathbb{N}$, von $\mathbb{R}$, ausgestattet mit dem Lebesgue-Maß. Diese Folge konvergiert punktweise gegen Null (sogar überall), nicht aber dem Maße nach.

Satz von Egoroff 7.15. Sei $\mu(X) < \infty$ und sei (f_n) eine Folge von messbaren Funktionen mit $f_n \xrightarrow{\text{f.ü.}} f$. Dann existiert zu jedem $\varepsilon > 0$ eine messbare Menge $X' \subseteq X$ so, dass $\mu(X \setminus X') \leq \varepsilon$ und $f_n \xrightarrow{L^\infty} f$ auf X'. Mit anderen Worten, f_n konvergiert auf X' gleichmässig gegen f.

Beweis. Wir bilden die messbaren Mengen
$$A_n^k := \bigcap_{i=n}^{\infty} \left\{|f_i(x) - f(x)| \leq \tfrac{1}{k}\right\}.$$
Nach Konstruktion ist $A_1^k \subseteq A_2^k \subseteq A_k^3 \subseteq \ldots$.

Falls $f_n(x) \to f(x)$, so gilt für jedes $k \in \mathbb{N}$, dass $x \in A_n^k$ für n groß genug, also

$x \in \bigcup_{n=1}^\infty A_n^k$. Da die Folge (f_n) fast überall gegen f konvergiert, hat also die Menge $X \setminus \bigcup_{n=1}^\infty A_n^k$ das Maß Null. Da die Menge aufsteigend ist, existiert also zu jedem $k \in \mathbb{N}$ ein Index n_k so, dass

$$\mu(X \setminus A_{n_k}^k) \leq \frac{\varepsilon}{2^k}.$$

Wir setzen

$$X' = \bigcap_{k=1}^\infty A_{n_k}^k.$$

Dann gilt

$$\mu(X \setminus X') = \mu\left(\bigcup_{k=1}^\infty X \setminus A_{n_k}^k\right) \leq \sum_{k=1}^\infty \mu(X \setminus A_{n_k}^k) \leq \sum_{k=1}^\infty \frac{\varepsilon}{2^k} = \varepsilon$$

und

$$\sup_{x \in X'} |f_i(x) - f(x)| \leq \frac{1}{k} \qquad \text{für} \qquad i \geq n_k,$$

also konvergiert (f_n) auf X' gleichmässig gegen f. ∎

8 Elemente der Stochastik

8.1 Verpflanzung von Maßen

Definition 8.1. Ein *Messraum* $(X, \mathcal{A})$ besteht aus einer Menge X zusammen mit einer σ-Algebra $\mathcal{A}$ von Teilmengen von X.

Seien $(X, \mathcal{A})$ und $(Y, \mathcal{B})$ zwei Messräume. Eine Abbildung $T : X \to Y$ heißt messbar (genauer $\mathcal{A}$-$\mathcal{B}$-messbar), falls für alle $B \in \mathcal{B}$ gilt, dass $T^{-1}(B) \in \mathcal{A}$.

Bemerkung 8.2. Sei $(X, \mathcal{A}, \mu)$ ein Maßraum. Dann stimmt das System $\mathcal{M}$ der messbaren, reellwertigen Funktionen (im Sinne von Definition 2.25) mit dem System der $\mathcal{A}$-$\mathcal{B}$-messbaren Abbildungen $X \to \mathbb{R}$ überein, wobei $\mathcal{B}$ die Borel'sche σ-Algebra auf $\mathbb{R}$ ist.

Lemma 8.3. Sind $(X, \mathcal{A})$, $(Y, \mathcal{B})$, $(Z, \mathcal{C})$ Messräume und $T : X \to Y$, $S : Y \to Z$ messbar, so ist auch $S \circ T : X \to Z$ messbar.

Beweis. Übung. ■

Lemma 8.4. Es sei $(X, \mathcal{A})$ ein Messraum und $T : X \to Y$ eine beliebige Abbildung.

(a) Das Mengensystem
$$T_*\mathcal{A} := \{B \subseteq Y \mid T^{-1}(B) \in \mathcal{A}\}$$
ist eine σ-Algebra von Teilmengen von Y, genannt die *verpflanzte σ-Algebra*.

(b) Die Abbildung $T : X \to Y$ ist $\mathcal{A}$-$T_*\mathcal{A}$-messbar.

(c) Die σ-Algebra $T_*\mathcal{A}$ ist die größte σ-Algebra von Teilmengen auf X, für die T messbar ist.

(d) Ist μ ein Maß auf $\mathcal{A}$, so wird durch die Formel
$$T_*\mu(B) := \mu\big(T^{-1}(B)\big)$$
ein Maß auf $T_*\mathcal{A}$ definiert. Es heißt das *verpflanzte Maß*.

Beweis. Übung. ■

Transformationsformel 8.5 (Allgemeiner Maße)**.** Seien $(X, \mathcal{A}, \mu)$ ein Maßraum und sei $T : X \to Y$ eine Abbildung, so dass $(Y, T_*\mathcal{A}, T_*\mu)$ ebenfalls ein Maßraum ist. Dann gilt für jede $T_*\mu$-integrierbare Funktion f die Formel

$$\int_Y f(y)\,d(T_*\mu)(y) = \int_X f\big(T(x)\big)\,d\mu(x). \tag{8.1.1}$$

Beweis. Ist $B \in T_*\mathcal{A}$, so gilt

$$\int_Y \mathbf{1}_B(y)\,d(T_*\mu)(y) = T_*\mu(B) = \mu\big(T^{-1}(B)\big) = \int_X \mathbf{1}_{T^{-1}(B)}(x)\,d\mu(x) = \int_X \mathbf{1}_B\big(T(x)\big)\,d\mu(x).$$

Damit gilt (8.1.1) für Indikatorfunktionen. Wegen Linearität gilt die Formel daher auch für beliebige Treppenfunktionen. Wegen Eindeutigkeit der Lebesgue-Erweiterung folgt damit die Gleichheit auch für beliebige integrierbare Funktionen. ■

Bemerkung 8.6. Ist $X = \mathbb{R}^d$ mit der Borel'schen σ-Algebra und dem Lebesgue-Maß und ist T ein Diffeomorphismus, so zeigt die Transformationsformel 4.15, dass $T_*\lambda$

absolut stetig ist bezüglich λ, genauer, dass

$$T_*\lambda = |\det DT^{-1}| \cdot \lambda$$

8.2 Grundbegriffe der Stochastik

Die Stochastik bietet eine mathematische Theorie zur Modellierung von *Zufallsexperimenten* wie Münzwürfen, Rouletteergebnissen bis hin zu komplexen Systemen wie Wettervorhersagen und ähnlichem. Die grundlegende mathematische Definition ist die folgende.

Definition 8.7. Ein *Wahrscheinlichkeitsraum* ist ein Maßraum $(\Omega, \mathcal{A}, \mathbb{P})$ mit $\mathbb{P}(\Omega) = 1$.

Messbare Teilmengen $A \in \mathcal{A}$ werden *Ereignisse* genannt. Die Zahl $\mathbb{P}(A) \in [0,1]$ ist dann die *Wahrscheinlichkeit*, dass das Ereignis A bei dem zu modellierenden Zufallsexperiment eintritt. Heuristisch bedeutet dies, dass bei N-facher Wiederholung des gegebenen Zufallsexperimentes das Ereignis A ungefähr $\mathbb{P}(A) \cdot N$ mal eintritt. Was in diesem Kontext "ungefähr" bedeutet, und wie sich dieser Fehler zu N verhält gehört zu den grundlegenden Fragen der Stochastik.

Außerhalb der Mathematik wird dies typischerweise mit Prozentzahlen angegeben: Man sagt also oft "die Wahrscheinlichkeit beträgt 50%" statt "die Wahrscheinlichkeit ist 0.5".

Beispiel 8.8. Die einfachsten expliziten Beispiele für Zufallsexperimente und die zugehörigen Wahrscheinlichkeitsräume sind die diskreten, wie zum Beispiel n-maliges Würfeln mit einem 6-seitigen Würfel. Hier ist der Wahrscheinlichkeitsraum $\Omega = \{1,2,3,4,5,6\}^n$. Das Maß ist das n-fache Produktmaß, wobei explizit $\mathbb{P}(\{k_1, \ldots, k_n\}) = 1/6^n$ gilt.

Bemerkung 8.9. Die Interpretation des Wahrscheinlichkeit ist oft tückisch. Insbesondere nicht-diskrete Wahrscheinlichkeitsmaße rufen häufig Schwierigkeiten hervor. Zum Beispiel könnte das Zufallsexperiment "Gewicht eines Neugeborenen" durch ein zum Lebesgue-Maß absolut stetiges Wahrscheinlichkeitsmaß modelliert werden. Da Einelementige Mengen stets Lebesguemaß Null haben, ist mathematisch gesehen die Wahrscheinlichkeit Null, dass ein Neugeborenes Baby ein ganz bestimmtes vorher festgelegtes Gewicht besitzt (beispielsweise 3.5kg). Es ist sogar so, dass mit Wahrscheinlichkeit Eins kein jemals geborenes Baby *genau* 3.5 kg bei seiner Geburt gewogen hat. Trotzdem hat natürlich jedes Baby zum Zeitpunkt der Geburt ein wohldefiniertes Körpergewicht. Dass ein Ereignis Wahrscheinlichkeit Null hat, bedeutet also nicht, dass es

nicht eintreten kann.
Eine andere Schwierigkeit betrifft die Interpretation des Wahrscheinlichkeitsbegriffes bei nicht wiederholbaren Zufallsexperimenten. Was ist also die Bedeutung beispielsweise der Aussage, dass ein gewisser Kandidat die Präsidentschaftswahl in den USA mit einer Wahrscheinlichkeit von 68% gewinnt?

Ein fundamentaler Unterschied zwischen Stochastik und Maßtheorie ist jedoch, dass der Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ typischerweise als unbekannt vorausgesetzt wird. Und dies aus gutem Grund: Soll ein komplexeres Problem vollständig mathematisch abgebildet werden, wie beispielsweise das Wetter auf der Erde, so bestünde der Wahrscheinlichkeitsraum beispielsweise aus "sämtlichen möglichen Wetterkonfigurationen" zu einer gegebenen Zeit. Auf diesem Raum ein realistisches Wahrscheinlichkeitsmaß zu bestimmen ist hoffnungslos. Man beschränkt sich daher in der stochastischen Modellierung auf von einem Zufallsexperiment abgeleiteten numerischen Messgrößen:

Definition 8.10. Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum. Eine *Zufallsvariable* (auf Ω) ist eine messbare Abbildung $X : \Omega \to \mathbb{R}$.

Die *Verteilung* einer Zufallsvariable ist das verpflanzte Maß $X_*\mathbb{P}$.

Die Benutzung des Buchstaben X für die Zufallsvariable ist in der Stochastik üblich, während der Wahrscheinlichkeitsraum in den meisten Fällen gar keine Bezeichnung bekommt. Hat man mehrere Zufallsvariablen $X_1, \ldots, X_n$, die dasselbe Zufallsexperiment beschreiben (die also auf dem selben – unbekannten – Wahrscheinlichkeitsraum definiert sind), so kann man diese zu einer $\mathbb{R}^n$-wertigen Zufallsvariable $X = (X_1, \ldots, X_n)$ zusammenfassen; man spricht dann von einem Zufallsvektor. In diesem Fall kann man die gemeinsame Verteilung $X_*\mathbb{P}$ betrachten, welches nun ein Maß auf $\mathbb{R}^n$ ist.

Statt zu versuchen, das Maß $\mathbb{P}$ explizit zu beschreiben, interessiert man sich für die Verteilung von zum Zufallsexperiment gehörenden Zufallsvariablen. Diese werden endweder im Rahmen der mathematischen Modellbildung aus dem Experiment zugrundeliegenden Annahmen hergeleitet oder empirisch durch Messungen approximativ bestimmt (hier sind wir dann im Gebiet der Statistik).

Beispiele 8.11. Hier einige Beispiele für Zufallsvariablen und deren Verteilung.

(1) Wir betrachten das Zufallsexperiment "radioaktiver Zerfall", mit der Zufallsvariable X, die die Zeit angibt, bis zu der ein gegebenes Atom eines radioaktiven Materials zerfallen ist. Die zugehörige Verteilung ist aus physikalischen Gesetzen bestimmbar als die *Exponentialverteilung*, welche die Dichte

$$f_\mu(t) = \begin{cases} \frac{1}{\mu} e^{-t/\mu} & t \geq 0 \\ 0 & t < 0 \end{cases}$$

bezüglich des Lebesgue-Maßes besitzt. Die Zahl μ ist die Halbwertzeit und hängt vom konkreten Element und Isotop ab.

(2) Eine der wichtigsten Verteilungen in der Stochastik ist die *Normalverteilung* (auch *Gaußverteilung* genannt), welche bezüglich des Lebesguemaßes die Dichte

$$f_{\mu,\sigma}(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/\sigma^2}$$

für Parameter $\mu \in \mathbb{R}$ und $\sigma > 0$ besitzt. Sie tritt beispielsweise in der Brown'schen Molekularbewegung auf. Ihre Bedeutung rührt unter anderem aus dem *zentralen Grenzwertsatz* der Stochastik, welcher besagt, dass der normalisierte Mittelwert einer Folge von vielen identisch verteilten, unabhängigen Zufallsvariablen (s.u.) asymptotisch normalverteilt ist.

(3) I m Experiment "Wettervorhersage" sind Beispiele für Zufallsvariablen "Temperatur ..." oder "Luftdruck an einem Ort zu einem gewissen Zeitpunkt". Die Bestimmung der entsprechenden Verteilung ist typischerweise höchst komplex und hängt sowohl von Messdaten als auch von Wettermodellen ab.

Einer Zufallsvariablen ordnet man folgende fundamentale Messgrößen zu:

Definition 8.12. Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und $X : \Omega \to \mathbb{R}$ eine Zufallsvariable. Dann heißt

$$\mathbb{E}[X] := \int_X X(\omega) d\mathbb{P}(\omega) = \int_{\mathbb{R}} x d(X_*\mathbb{P})(x)$$

der *Erwartungswert von* X, wobei wir die Transformationsformel 8.5 benutzt haben.

Die Zahl

$$\mathbb{V}\text{ar}[X] := \mathbb{E}\left[\left(X - \mathbb{E}[X]\right)^2\right],$$

heißt *Varianz* von X und ihre Quadratwurzel ist die *Standardabweichung* von X.

Die *Covarianz* von zwei Zufallsvariablen X, Y ist die Zahl

$$\mathbb{C}\text{ov}(X, Y) := \mathbb{V}\text{ar}[(X - \mathbb{E}[X])(Y - \mathbb{E}(Y))].$$

X und Y heißen *unkorreliert*, falls $\mathbb{C}\text{ov}(X, Y) = 0$.

Maßtheoretisch gesehen ist der Erwartungswert einer Zufallsvariablen also nichts weiter als deren Integral bezüglich des Wahrscheinlichkeitsmaßes $\mathbb{P}$. Von fundamentaler Bedeutung ist jedoch dass die zweite Formal für $\mathbb{E}[X]$ nur von der Verteilung der Zufallsvariable X abhängt, also bestimmt werden kann ohne $\mathbb{P}$ selber zu kennen. Die Varianz ist im Wesentlichen die L^2-Norm der Zufallsvariablen:

$$\mathbb{V}\mathrm{ar}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \|X\|_{L^2}^2 - \mathbb{E}[X]^2.$$

Während natürlich zwei Zufallsvariablen mit unterschiedlichem Erwartungswert oder unterschiedlicher Varianz nicht die gleiche Verteilung besitzen können, ist umgekehrt natürlich nicht gesagt, dass zwei Zufallsvariablen die gleiche Verteilung besitzen nur weil ihr Erwartungswert und ihre Varianz übereinstimmen (dies trifft aber beispielsweise zu, wenn man die Zusatzinformation hat, dass beide normalverteilt sind).

Beispiele 8.13. Erwartungswert und Varianz einer exponentialverteilten Zufallsvariable wie in Beispiel 8.11(1) sind μ bzw. μ^2. Der Erwartungswert einer normalverteilten Zufallsvariable wie in Beispiel 8.11(1) sind μ bzw. σ^2.

Lemma 8.14 (Tschebyschevs Ungleichung)**.** Sei X eine Zufallsvariable. Dann gilt für jedes $C, \alpha > 0$. Dann gilt

$$\mathbb{P}\big(\{|X| \geq C\}\big) \leq \frac{\mathbb{E}[|X|^\alpha]}{C^\alpha}.$$

Insbesondere gilt: Hat X Erwartungswert μ und Varianz σ^2, so gilt für jedes $C > 0$, dass

$$\mathbb{P}\big(\{|X - \mu| \geq C\}\big) \leq \frac{\sigma^2}{C^2}$$

Beweis. Es gilt

$$C^\alpha \cdot \mathbb{P}\big(\{|X| \geq C\}\big) = C^\alpha \cdot \mathbb{P}\big(\{|X|^\alpha \geq C^\alpha\}\big) \leq \int_{\{|X| \geq C\}} |X(\omega)|^\alpha d\mathbb{P}(\omega) = \mathbb{E}\big[|X|^\alpha\big].$$

Der Spezialfall folgt, indem man X durch $X - \mu$ ersetzt und $\alpha = 2$ wählt. ■

8.3 Stochastische Unabhängigkeit

Sind X und Y zwei Zufallsvariablen auf demselben Wahrscheinlichkeitsraum, so ist deren Produkt wieder eine Zufallsvariable. Der zugehörige Erwartungswert ist allerdings im Allgemeinen unmöglich zu bestimmen. Dies führt zu folgendem Begriff:

Definition 8.15. Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum. Eine Familie $(X_i)_{i \in I}$ von Zufallsvariablen auf Ω heißt *(stochastisch) unabhängig*, wenn für jede endliche Teilmenge $\{i_1, \dots, i_n\} \subseteq I$ gilt: Die gemeinsame Verteilung $X_* \mathbb{P}$ des Zufallsvektors $X = (X_{i_1}, \dots, X_{i_n})$ ist das Produktmaß

$$X_* \mathbb{P} = (X_{i_1})_* \mathbb{P} \otimes \cdots \otimes (X_{i_n})_* \mathbb{P}.$$

Anschaulich bedeutet dies, dass die Information, dass Informationen über die Werte einer endlichen Teilmenge der Zufallsvariablen einem nicht hilft, die Werte der anderen vorherzusagen.

Offenbar kann dies nicht in Termen der einzelnen Verteilungen $(X_i)_* \mathbb{P}$ ausgedrückt werden und daher im Allgemeinen (bei unbekanntem Wahrscheinlichkeitsmaß $\mathbb{P}$) auch nicht überprüft werden. In der Praxis ist die Unabhängigkeit daher eine Zusatzannahme, die an das stochastische Modell aufgrund von Plausibiltätsüberlegungen oft gestellt wird.

Beispiele 8.16.

(1) Im Zufallsexperiment "radioaktiver Zerfall" sollten Zufallsvariablen, die den Zerfallszeitpunkt von zwei unterschiedlichen Atomen eines radioaktiven Elements beschreiben, als unabhängig angenommen werden

(2) Im Gegensatz dazu ist es sicherlich nicht realistisch, zwei Zufallsvariablen, die Körpergröße bzw. die Schuhgröße einer zufällig ausgewählten Testperson beschreiben, als unabhängig anzunehmen.

(3) Im Fall, dass X und Y Aktienkurse von zwei Firmen zu einem gegebenen Zeitpunkt beschreiben, ist es auch oft nicht realistisch X und Y als unabhängig anzunehmen: Beide Firmen könnten derselben Branche angehören und daher ähnlich auf Marktbedingungen reagieren. Auch könnten allgemeine wirtschaftliche Faktoren beide Werte beeinflussen, wie zum Beispiel Zinssätze oder makroökonomische Richtlinien.

Beispiel 8.17. Wir betrachten einen Zufallsvektor $X = (X_1, \dots, X_n)$ so, dass die gemeinsame Verteilung $X_* \mathbb{P})$ auf $\mathbb{R}^n$ das zum Lebesguemaß absolut stetige Maß mit der Dichte

$$g(x) = \frac{\sqrt{\det(A)}}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2}\langle x - b, A(x - b)\rangle\right)$$

ist, mit $b \in \mathbb{R}^n$ und A eine symmetrische $(n \times n)$-Matrix mit positiven Eigenwerten. Ein solcher Zufallsvektor heißt *multivariat normalverteilt* mit Erwartungswertvektor b und Kovarianzmatrix $\Sigma = A^{-1}$.

Die Vektoren $X_1, \dots, X_n$ sind unabhängig genau dann wenn A eine Diagonalmatrix ist.

Lemma 8.18. Sind $X_1, \dots, X_n$ unabhängige Zufallsvariablen mit endlichem Erwartungswert, so gilt

$$\mathbb{E}\left[\prod_{i=1}^{\infty} X_i\right] = \prod_{i=1}^{\infty} \mathbb{E}[X_i].$$

Da der Erwartungswert einer Summe von Zufallsvariablen für beliebige Zufallsvariablen die Summe der Erwartungswerte ist, ermöglicht das obige Lemma unter der Zusatzvoraussetzung der Unabhängigkeit, zusätzlich auch den Erwartungswert von beliebigen Polynomen in den Zufallsvariablen $X_1, \dots, X_n$ zu bestimmen. Beispielsweise kann die Varianz der Summe $X_1 + \cdots + X_n$ als

$$\begin{aligned} \mathbb{V}\mathrm{ar}\left[\sum_{i=1}^{n} X_i\right] &= \mathbb{E}\left[\left(\sum_{i=1}^{n} X_i\right)^2\right] - \mathbb{E}\left[\sum_{i=1} X_i\right]^2 \\ &= \sum_{ij=1}^{n} \big(\mathbb{E}[X_i X_j] - \mathbb{E}[X_i]\mathbb{E}[X_j]\big) \qquad (8.3.1) \\ &= \sum_{i=1}^{n} \big(\mathbb{E}[X_i^2] - \mathbb{E}[X_i]^2\big) = \sum_{i=1}^{n} \mathbb{V}\mathrm{ar}[X_i], \end{aligned}$$

also die Summe der Varianzen berechnet werden (die Summanden mit $i \neq j$ kürzen sich aufgrund der Unabhängigkeit heraus).

Beweis. Wir betrachten den Zufallsvektor $X = (X_1, \dots, X_n)$ und die Abbildung

$$T : \mathbb{R}^n \to \mathbb{R}, \qquad (x_1, \dots, x_n) \to x_1 \cdots x_n.$$

Dann gilt $X_1 \cdots X_n = T \circ X$, also

$$(X_1 \cdots X_n)_* \mathbb{P} = T_* X_* \mathbb{P} = T_*\big((X_1)_* \mathbb{P} \otimes \cdots \otimes (X_n)_* \mathbb{P}\big),$$

wobei wir im zweiten Schritt die Unabhängigkeit benutzt haben. Wegen der Transformationsformel 8.5 gilt nun

$$\begin{aligned}
\mathbb{E}\left[\prod_{i=1}^{n} X_i\right] &= \int_{\Omega} \prod_{i=1}^{n} X_i(\omega)\, d\mathbb{P}(\omega) \\
&= \int_{\mathbb{R}} x \, d\Big((X_1 \cdots X_n)_* \mathbb{P}\Big)(x) \\
&= \int_{\mathbb{R}^n} x_1 \cdots x_n \, d\big((X_1)_*\mathbb{P} \otimes \cdots \otimes (X_n)_*\mathbb{P}\big)(x_1, \ldots, x_n) \\
&= \prod_{i=1}^{n} \int_{\mathbb{R}} x_i d(X_i)_*\mathbb{P}(x_i) = \prod_{i=1}^{n} \mathbb{E}[X_i],
\end{aligned}$$

wobei wir den Satz von Fubini angewendet haben, um von der letzten zur vorletzten Zeile zu gelangen. ■

Satz 8.19. Sei $\mu_1, \mu_2, \ldots$ eine Folge von Borel'schen Wahrscheinlichkeitsmaßen auf $\mathbb{R}$. Dann existiert eine Folge $X_1, X_2, \ldots$ unabhängiger Zufallsvariablen so dass X_i die Verteilung μ_i hat.

Beweis. Zunächst existiert offenbar überhaupt eine Zufallsvariable mit Verteilung μ_i. Beispielsweise kann man hier $(\Omega_i, \mathcal{A}_i, \mathbb{P}_i) = (\mathbb{R}, \mathcal{B}, \mu_i)$ nehmen (wobei $\mathcal{B}$ die Borel'sche σ-Algebra auf $\mathbb{R}$ ist). Eine unendliche Folge erhält man nun indem man $\Omega = \prod_{i=1}^{\infty} \Omega_i$ nimmt mit dem unendlichen Produktmaß

$$\mathbb{P} = \bigotimes_{i=1}^{\infty} \mu_i.$$

Die Zufallsvariable X_i ist dann einfach die Projektion auf den i-ten Faktor. ■

Typischerweise spricht man von identisch verteilten Zufallsvariablen:

Definition 8.20. Zwei Zufallsvariablen sind *identisch verteilt*, falls ihre Verteilungen übereinstimmen.

Korollar 8.21. Zu jeder gegebenen Wahrscheinlichkeitsverteilung gibt es eine Folge von identisch verteilten unabhängigen Zufallsvariablen.

Schwaches Gesetz der großen Zahlen. 8.22. Sei $X_1, X_2, \ldots$ eine Folge identisch verteilter unabhängiger Zufallsvariablen mit endlichem Erwartungswert $\mu = \mathbb{E}[X_i]$. Dann konvergiert die Zufallsvariable

$$\bar{X}_n := \frac{1}{n} \sum_{i=1}^{n} X_i$$

in Wahrscheinlichkeit gegen μ.

Hierbei ist *Konvergenz in Wahrscheinlichkeit* der Ausdruck der Stochastik für Konvergenz im Maß.

Bemerkung 8.23. Anschaulich bedeutet das Gesetz der großen Zahlen, dass sich der Durchschnitt der Realisierungen der Zufallsvariablen gegen den Erwartungswert konvergiert.

Bemerkung 8.24. Das (schwerer zu beweisende) *starke* Gesetz der großen Zahlen besagt, $\bar{X}_n$ sogar fast sicher (also maßtheoretisch fast überall) gegen Null konvergiert.

Beweis. Wir führen den Beweis der Einfachheit halber unter der Zusatzannahme, dass die Varianz der X_i endlich ist. Dann gilt wegen (8.3.1) und der Unabhängigkeit der X_i, dass

$$\mathbb{V}\mathrm{ar}\big[\bar{X}_n\big] = \frac{1}{n^2} \mathbb{V}\mathrm{ar}\big[(X_1 + \cdots + X_n)\big] = \frac{1}{n^2} \sum_{i=1}^{n} \mathbb{V}\mathrm{ar}\big[X_i\big] = \frac{\sigma^2}{n},$$

where we wrote $\mathbb{V}\mathrm{ar}[X_i] = \sigma^2$. Wegen der Tschebyschev'schen Ungleichung 8.14 gilt

$$\mathbb{P}\big(\{|\bar{X}_n - \mu| \geq \varepsilon\}\big) \leq \frac{\mathbb{V}\mathrm{ar}[\bar{X}_n]}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}.$$

Dies konvergiert gegen Null. ■

8.4 Stochastische Prozesse

In diesem Abschnitt sei M ein topologischer Raum und sei $\mathcal{B}_M$ die Borel'sche σ-Algebra auch M. Für ein $T \geq 0$ (wir lassen auch $T = \infty$ zu) betrachten wir das unendliche Produkt $M^{[0,T)}$, der mit dem Raum aller Abbildungen $[0, T) \to M$ identifiziert werden kann (eine Familie $(x_t)_{t\in[0,T)}$ entspricht dann der Funktion $t \mapsto x_t$). Der Raum $M^{[0,T)}$

trägt dann die Produkttopologie, welches die kleinste Topologie ist, die alle Projektionen

$$\pi_t : M^{[0,T)} \to M, \qquad (x_t)_{t\in[0,T)} \mapsto x_t,$$

stetig macht. Explizit ist diese Topologie erzeugt von denjenigen Zylindermengen $A = \prod_{t\in[0,T)} A_t$, so dass $A_t \neq M$ für nur endlich viele $t \in [0,T)$. Nach dem *Satz von Tychonov* aus der Topologie ist für kompaktes M auch $M^{[0,T)}$ wieder kompakt.

Die Produkt-σ-Algebra auf dem Product $M^{[0,T)}$ ist die σ-Algebra der Produkttopologie, wir bezeichnen Sie mit $\mathcal{B}_{M^{[0,T)}}$. Da die Projektionen π_t stetig sind, sind sie auch $\mathcal{B}_{M^{[0,T)}}$-$\mathcal{B}_M$-messbar.

Definition 8.25. Ein M-wertiger *stochastischer Prozess* ist ein Familie $(X_t)_{t\in[0,T)}$ von Funktionen $X_t : \Omega \to M$ auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ so, dass die zugehörige Abbildung

$$X : \Omega \longrightarrow M^{[0,T)}, \qquad \omega \longmapsto \big(X_t(\omega)\big)_{t\in[0,T)}$$

$\mathcal{A}$-$\mathcal{B}_{M^{[0,T)}}$-messbar ist.

Da alle Projektionen $\pi_t : M^{[0,T)} \to M$ messbar sind, sind insbesondere alle X_t messbar, also M-wertige Zufallsvariablen.

Eine wichtige Beispielklasse ist die folgende.

Definition 8.26. Ein $\mathbb{R}^n$-wertiger stochastischer Prozess $(X_t)_{t\in[0,T)}$ heißt *Wiener-Prozess* oder *Brown'sche Bewegung*, falls

(i) für alle $0 \leq s < t < T$ ist $X_t - X_s$ ist multivariat normalverteilt mit Mittelwert Null und Kovarianzmatrix $(t-s)\mathrm{id}$;

(ii) der Prozess hat unabhängige Inkremente. Genauer gilt: Für alle Zeiteinteilungen $0 \leq t_0 < t_1 < \cdots < t_n < T$ sind die Zufallsvariablen $X_{t_1} - X_{t_0}, \ldots, X_{t_n} - X_{t_{n-1}}$ unabhängig.

Um die Existenz eines Wienerprozesses nachzuweisen, benutzen wir ein allgemeines Konstruktionswerkzeug für stochastische Prozesse, den Erweiterungssatz von Kolmogorov.

Für $I \subseteq [0,T)$ bezeichnen wir mit $\pi_I : M^{[0,T)} \to M^I$ die Projektion auf die Koordinaten indiziert durch $t \in I$, und für $J \subseteq I$ sei die Projektion $\pi_{I,J} : M^I \to M^J$ genauso definiert. Ist $(X_t)_{t\in[0,T)}$ ein stochastischer Prozess, so gilt offenbar

$$(\pi_J)_*\mathbb{P} = (\pi_{I,J})_*(\pi_I)_*\mathbb{P} \tag{8.4.1}$$

für alle $I \subseteq J \subseteq [0,T)$. Der folgende Satz sagt, dass ein stochastischer Prozess schon durch die Maße $(\pi_I)_*\mathbb{P}$ für *endliches* $I \subseteq [0,T)$ eindeutig bestimmt ist.

Erweiterungssatz von Kolmogorov 8.27. Sei M ein kompakter Hausdorffraum. Für jede endliche Teilmenge $I \subset [0,T)$ sei ein Maß $\mathbb{P}_I$ auf M^I vorgeben so, dass die Familie von Wahrscheinlichkeitsmaßen $(\mathbb{P}_I)$ die Kompatibilitätsbedingung

$$(\pi_{I,J})_*\mathbb{P}_I = \mathbb{P}_J \tag{$*$}$$

erfüllt. Dann existiert ein eindeutiges Wahrscheinlichkeitsmaß $\mathbb{P}$ auf $M^{[0,T)}$ so, dass $(\pi_I)_*\mathbb{P} = \mathbb{P}_I$.

Beweis. Die Eindeutigkeit folgt daher, dass die σ-Algebra $\mathcal{B}_{M^{[0,T)}}$ von der Teilmenge

$$\bigcup_{\substack{I\subset[0,T)\\ \text{endlich}}} \pi_I^{-1}(\mathcal{B}_{M^I}) \subset \mathcal{B}_{M^{[0,T)}}$$

erzeugt wird. Jedes Maß $\mathbb{P}$ auf $\mathcal{B}_{M^{[0,T)}}$ ist also eindeutig durch die Werte auf Elementen dieser Teilmenge festgelegt.

Für jede stetige Funktion $f : M^I \to \mathbb{R}$ erhalten wir eine Funktion $f \circ \pi_I : M^{[0,T)} \to \mathbb{R}$. Eine Funktion dieser Form für $I \subseteq [0,T)$ endlich nennen wir *Zylinderfunktion*; sie ist stetig bezüglich der Produkttopologie auf $M^{[0,T)}$. Sei $V \subseteq C(M^{[0,T)})$ die Menge der Zylinderfunktionen. Sind $I, J \subseteq [0,T)$ endlich und sind $F := f \circ \pi_I$ und $G := g \circ \pi_J$ Zylinderfunktionen, so sind $F + G$ und $F \cdot G$ wieder Zylinderfunktionen: Beide sind von der Form $h \circ \pi_{I\cup J}$ für eine Funktion $h : M_{I\cup J} \to \mathbb{R}$. V ist also eine Unteralgebra von $C(M^{[0,T)})$. Nach dem Satz von Stone-Weierstraß 4.20 ist V dicht:

(i) Zwei Elemente $x = (x_t)_{t\in[0,T)}, y = (y_t)_{t\in[0,T)} \in M^{[0,T)}$ sind genau dann unterschiedlich, wenn $x_{t_0} \neq y_{t_0}$ für ein $t_0 \in [0,T)$. Da M ein Hausdorffraum ist, existert dann eine stetige Funktion $f \in C(M)$ mit $f(x_{t_0}) \neq f(y_{t_0})$. Für die Zylinderfunktion $F = f \circ \pi_{t_0} \in V$ gilt dann $F(x) \neq F(y)$, also ist V punktetrennend.

(ii) Es ist offensichtlich, dass zu jedem $x = (x_t)_{t\in[0,T)}$ eine Zylinderfunktion $F \in V$ existiert, so dass $F(x) \neq 0$.

Durch die Vorschrift

$$\varphi(f \circ \pi_I) := \int_{M^I} f(x)d\mathbb{P}_I(x)$$

wird nun auf V ein positives lineares Funktional definiert. Dies ist wohldefiniert wegen der Kompatibilitätsbedingung $(*)$. Denn: Ist $F = f_I \circ \pi_I = f_J \circ \pi_J$, so gilt auch $f_I \circ \pi_{I\cup J,I} = f_J \circ \pi_{I\cup J,J}$ und damit

$$\int_{M^I} f_I(x)d\mathbb{P}_I(x) = \int_{M^I} \underbrace{f_I\big(\pi_{I\cup J,I}(x)\big)}_{=f_J(\pi_{I\cup J,J}(x))} d\mathbb{P}_{I\cup J}(x)\big)\, d\mathbb{P}_{I\cup J}(x) = \int_{M^J} f_J(x)d\mathbb{P}_J(x)$$

wegen der Transformationsformel 8.5.

Das Funktional ist beschränkt auf V, da

$$|\varphi(f \circ \pi_I)| = \left| \int_{M^I} f(x) d\mathbb{P}_I(x) \right| \leq \|f\|_\infty = \|f \circ \pi_I\|_\infty.$$

Da jedes beschränkte lineare Funktional, das auf einem dichten Teilraum eines Banachraumes definiert ist, eine eindeutige stetige Fortsetzung auf den gesamten Banachraum hat, setzt sich φ eindeutig zu einem stetigen linearen Funktional auf ganz $C(M^{[0,T)})$ fort. Nach dem Satz von Markhov-Kakutani-Riesz 6.13 ist damit φ durch Integration bezüglich eines Maßes $\mathbb{P}$ auf $M^{[0,T)}$ gegeben. ■

Bemerkung 8.28. Für jedes Borel'sche Wahrscheinlichkeitsmaß $\mathbb{P}$ auf $M^{[0,T)}$ (beispielsweise ein durch den Erweiterungssatz von Kolmogorov erhaltenes) können wir einen stochastischen Prozess $(X_t)_{t\in[0,T)}$ auf M definieren, definiert auf dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P}) = (M^{[0,T)}, \mathcal{B}_{M^{[0,T)}}, \mathbb{P})$, indem wir einfach $X_t := \pi_t$, also $X_t(x) := x_t$ setzen. Dies impliziert, dass $I = \{t_1, \ldots, t_N\} \subseteq [0, T)$ die Verteilung der Zufallsvektoren $X_I := (X_{t_1}, \ldots, X_{t_N})$ gegeben ist durch $(X_I)_*\mathbb{P} = \mathbb{P}_I$.

Beispiel 8.29. Um eine kompatible Familie von Maßen $(\mathbb{P}_I)$ wie im Erweiterungssatz von Kolmogorov zu konstruieren, kann man wie folgt vorgehen: Sei $(g_t)_{t>0}$ eine Familie von integrierbaren positiven Funktionen (sogenannten *Kernen*) auf $M \times M$, die die *Markhov-Eigenschaft*

$$\int_M g_t(x,y) \cdot g_s(y,z) d\mu(y) = g_{t+s}(x,z) \tag{8.4.2}$$

bezüglich eines festen Borel-Maßes μ auf M erfüllen und die sich alle zu Eins integrieren. Weiterhin verlangen wir

$$\lim_{t\searrow 0} \int_M g_t(x,y) f(y) d\mu(y) = f(x) \tag{8.4.3}$$

für alle stetigen Funktionen f auf M. Für einen fest gewählten Punkt $x_0 \in M$ und $I = \{t_1, \ldots, t_N\} \subseteq [0, T)$ definierten wir das Maß $\mathbb{P}_I$ auf M^I durch die Dichte

$$g_{t_1,\ldots,t_N}(x_1, \ldots, x_N) := \prod_{i=1}^N g_{t_i - t_{i-1}}(x_i, x_{i-1})$$

bezüglich des Produktmaßes $\mu^{\otimes I}$ (hierbei ist $t_0 = 0$). Im Fall, dass $I = \{0\}$ nehmen wir für P_I das Dirac-δ-Maß bei x_0. Die Kompabitilitätseigenschaft $(*)$ für die Familie $(\mathbb{P}_I)$ folgt dann aus der Markhov-Eigenschaft (8.4.2).

Der Erweiterungssatz von Kolmogorov liefert dann ein Wahrscheinlichkeitsmaß $\mathbb{P}$ auf $M^{[0,\infty)}$ sowie einen stochastischen Prozess $(X_t)_{t>0}$ wie in Bemerkung 8.28. Die Familie $(g_t)_{t>0}$ heißt dann *Übergangskern* des Prozesses.

Lemma 8.30. Der auf $\mathbb{R}^n$ durch den Übergangskern

$$g_t(x,y) := \frac{1}{(2\pi t)^{n/2}} \exp\left(-\frac{|x-y|^2}{2t}\right).$$

wie in Beispiel 8.29 definierte stochastische Prozess ist ein Wienerprozess. Es gilt $X_0 = x_0$ fast sicher.

Bemerkung 8.31. Ein kleines Problem ist hier, dass $\mathbb{R}^n$ zwar ein Hausdorffraum, aber nicht kompakt ist, so dass der Kolmogorov'sche Erweiterungssatz nicht direkt angewendet werden kann. Aber wir können uns hier behelfen, indem wir durch Hinzunahme eines "Punktes bei Unendlich" zur *Einpunktkompaktifizierung* $M := \mathbb{R}^n \cup \{\infty\}$ übergehen. (Dies wird ein kompakter Hausdorffraum indem man $U \subseteq \mathbb{R} \cup \{\infty\}$ für offen erklärt falls entweder $U \subseteq \mathbb{R}^n$ offen oder falls $U = (\mathbb{R} \cup \{\infty\}) \setminus K$ für $K \subseteq \mathbb{R}^n$ kompakt.) Die offensichtliche Einbettung $\mathbb{R}^n \to M$ ist stetig und jedes Wahrscheinlichkeitsmaß auf $\mathbb{R}^n$ liefert eines auf M durch Verpflanzung.

Beweis. Wir bemerken zunächst dass $g_t(x,y)$ nur von der Differenz $x-y$ abhängt und definieren daher

$$\tilde{g}_t(x) := \frac{1}{(2\pi t)^{n/2}} \exp\left(-\frac{|x|^2}{2t}\right).$$

Die Überprüfung der Eigenschaften (8.4.2) und (8.4.3) lassen wir als Übungsaufgabe.

Seien $0 = t_0 < t_1 < \cdots < t_N < T$. Nach Definition ist die gemeinsame Verteilung der Zufallsvariablen $X_{t_1}, \ldots, X_{t_n}$ gegeben durch

$$\int_{\mathbb{R}^n} \cdots \int_{\mathbb{R}^n} f(x_1, \ldots, x_N) \prod_{i=1}^{N} \tilde{g}_{t_i - t_{i-1}}(x_i - x_{i-1}) dx_1 \cdots dx_N.$$

Wir setzen $Y_i = X_{t_i} - X_{t_{i-1}}$ und wollen zeigen, dass die Zufallsvariablen $Y_1, \ldots, Y_N$ normalverteilt und unabhängig sind. Hierzu bemerken wir, dass

$$(Y_1, \ldots, Y_N) = \varphi \circ (X_{t_1}, \ldots, X_{t_N})$$

mit $\varphi : \mathbb{R}^{n\times N} \to \mathbb{R}^{n\times N}$ gegeben durch

$$\varphi(x_1, \ldots, x_n) = (x_1 - x_0, \ldots, x_N - x_{N-1}),$$

da $X_{t_0} = X_0 = x_0$ fast sicher. Nach Transformationsformel gilt daher

$$\begin{aligned}
&\int_{\mathbb{R}^n} \cdots \int_{\mathbb{R}^n} f(y_1, \dots, y_N) d(Y_1, \dots, Y_N)_* \mathbb{P}(y_1, \dots, y_N) \\
&\quad = \int_{\mathbb{R}^n} \cdots \int_{\mathbb{R}^n} f(x_1 - x_0, \dots, x_N - x_{N-1}) d(X_{t_1}, \dots, X_{t_n})_* \mathbb{P}(x_1, \dots, x_N) \\
&\quad = \int_{\mathbb{R}^n} \cdots \int_{\mathbb{R}^n} f(x_1 - x_0, \dots, x_N - x_{N-1}) \prod_{i=1}^{N} \tilde{g}_{t_i - t_{i-1}}(x_i - x_{i-1}) dx_1 \cdots dx_N. \\
&\quad = \int_{\mathbb{R}^n} \cdots \int_{\mathbb{R}^n} f(y_1, \dots, y_N) \prod_{i=1}^{N} \tilde{g}_{t_i - t_{i-1}}(y_i) dy_1 \cdots dy_N.
\end{aligned}$$

Hierbei benutzen die Transformationsformel 4.15 für die obige affin-lineare Funktion φ, deren Jacobimatrix durch die Matrix

$$D\varphi(x) = \begin{pmatrix} 1 & 1 & & & \\ & 1 & 1 & & \\ & & \ddots & \ddots & \\ & & & 1 & 1 \\ & & & & 1 \end{pmatrix}$$

gegeben ist, die Determinante Eins hat. Es gibt also keinen Korrekturfaktor.

Dies zeigt erstens, dass jedes Inkrement $X_{t_i} - X_{t_{i-1}}$ normalverteilt ist und zweitens, dass die Inkremente unabhängig sind. ■

Beispiele 8.32. Ein weiteres Beispiel für eine Familie von Übergangskernen, die die Bedingungen aus Beispiel 8.29 erfüllt, und damit einen stochastischen Prozess liefert ist der *Cauchy-Kern* in $\mathbb{R}^n$, gegeben durch

$$g_t(x, y) = \frac{\Gamma\left(\frac{1}{2}\right)}{\pi^{n/2} \Gamma\left(\frac{n+1}{2}\right)} \cdot \frac{t}{(t^2 + |y - x|^2)^{\frac{n+1}{2}}}.$$

Der erzeugte Prozess heißt *Cauchy-Prozess*. Da $g_t(x - y)$ ebenfalls nur von der Differenz $x - y$ abhängt, zeigt der Beweis von Lemma 8.30, dass dieser Prozess ebenfalls unabhängige Inkremente hat.

Ist $(X_t)_{t \in [0,T)}$ ein stochastischer Prozess in M (definiert auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$), so sind die *Pfade* die Funktionen $X(\omega) : [0, T) \to M$, $\omega \in \Omega$, gegeben durch $t \mapsto X_t(\omega)$. Eine Besonderheit von Wienerprozessen ist, dass diese fast sicher stetig Pfade haben. Um dies einzusehen, benutzen wir folgenden Satz.

Stetigkeitssatz von Kolmogorov 8.33. Sei $(X_t)_{t\in[0,T)}$ ein $\mathbb{R}^n$-wertiger stochastischer Prozess so, dass für alle $0 \leq s < t < T$ und Konstanten $\gamma < \beta$ die Abschätzung

$$\mathbb{E}\left[|X_t - X_s|^\beta\right] \leq c|t-s|^{1+\gamma} \tag{8.4.4}$$

gilt, mit Konstanten $c, \beta, \gamma > 0$. Dann gilt für jede abzählbare und beschränkte Teilmenge $S \subseteq [0,T)$ und jedes $0 \leq \alpha \leq \gamma/\beta$: Die Pfade von $(X_t)_{t\in[0,T)}$ fast sicher α-Hölder-stetig auf S.

Hierbei ist eine Funktion $f : [0,T) \to \mathbb{R}^n$ *α-Hölder-stetig* auf $S \subseteq [0,T)$, falls es zu jedem $t \in [0,T)$ Konstanten $C, \varepsilon > 0$ gibt, so dass

$$\frac{|f(t)-f(s)|}{|t-s|^\alpha} \leq C, \qquad s,t \in S.$$

Beweis. Wegen der Tschebyschev'schen Ungleichung 8.14 gilt für alle $t, s \in [0,T)$ und alle $\alpha > 0$, dass

$$\mathbb{P}\big(|X_t - X_s| \geq |t-s|^\alpha\big) \leq \frac{\mathbb{E}\left[|X_t-X_s|^\beta\right]}{|t-s|^{\beta\alpha}} \leq c|t-s|^{1+\gamma-\beta\alpha},$$

wobei wir (8.4.4) benutzt haben. Sei $t \in [0,T)$ fest.

Wir sagen, dass eine endliche Teilmenge $S' \subseteq S$ *ε-dicht* ist, falls die ε-Umgebung eines jeden Punktes von S einen Punkt von S' enthält (äquivalent: die Vereinigung aller Intervalle $(t-\varepsilon, t+\varepsilon)$, $t \in S'$, enthält S). Schreiben wir $S' = \{t_0 < t_1 < \cdots < t_N\}$, so gilt $|t_i - t_{i-1}| \leq \varepsilon$, und damit

$$\begin{aligned}
\mathbb{P}\Big(\max_{1\leq i\leq N} |X_{t_i} - X_{t_{i-1}}| \geq |t_i - t_{i-1}|^\alpha\Big) &\leq \sum_{i=1}^N \mathbb{P}\big(|X_{t_i} - X_{t_{i-1}}| \geq |t_i - t_{i-1}|^\alpha\big) \\
&\leq \sum_{i=1}^N c|t_i - t_{i-1}|^{1+\gamma-\beta\alpha} \\
&\leq c\sup\{|t_i - t_{i-1}|^{\gamma-\beta/\alpha}\} \cdot \sum_{i=1}^N |t_i - t_{i-1}| \\
&\leq c\varepsilon^{\gamma-\beta/\alpha} \cdot |t_n - t_0|.
\end{aligned}$$

Sei nun $S_1 \subseteq S_2 \subseteq \cdots S$ eine Ausschöpfung durch endliche Teilmengen von S, so dass S_n stets 2^{-n}-dicht ist. Wir setzen

$$A_n = \Big\{\omega \in \Omega \;\Big|\; \max_{1\leq i\leq N} |X_{t_i}(\omega) - X_{t_{i-1}}(\omega)| \geq |t_i - t_{i-1}|^\alpha\Big\},$$

wobei $S_n = \{t_0 < t_1 < \cdots < t_N\}$ eine Aufzählung der Elemente von S_n ist. Ist D eine obere Schranke für den Durchmesser von S, so gilt nach der obigen Rechnung

$$\sum_{n=1}^{\infty} \mathbb{P}(A_n) \leq cD \sum_{n=1}^{\infty} 2^{-n(\gamma-\beta\alpha)} < \infty,$$

falls $\gamma - \beta\alpha > 0$, oder äquivalent $\alpha < \gamma/\beta$. Nach dem Lemma von Borel-Cantelli ist daher die Menge

$$N := \bigcup_{n=1}^{\infty} \bigcap_{i=n}^{\infty} A_i = \{\omega \in \Omega \mid \omega \text{ ist in unendlich vielen } A_n \text{ enthalten}\}$$

eine Nullmenge. Für alle $\omega \in \Omega \setminus N$ ist dann der Pfad $X(\omega)$ α-Hölderstetig auf S. ∎

Beispiel 8.34. Für einen Wiener-Prozess $(X_t)_{t\geq 0}$ gilt

$$\begin{aligned}
\mathbb{E}\left[|X_{s+t} - X_s|^\beta\right] &= \int_{\mathbb{R}^n} |y|^\beta \tilde{g}_t(y)\,dy = \int_{\mathbb{R}^n} |y|^\beta (2\pi t)^{-n/2} e^{-\frac{|y|^2}{2t}}\,dy \\
&= (2\pi t)^{-n/2} \int_0^\infty \int_{S^{n-1}} e^{-r^2/2t} r^{\beta+n-1}\,dr \\
&= (2\pi t)^{-n/2} \cdot \mathcal{H}_{n-1}(S^{n-1}) \cdot \int_0^\infty e^{-r^2/2t} r^{\beta+n-1}\,dr,
\end{aligned}$$

unter Benutzung der Koflächenformel von Satz 5.23. Die Substitution $r^2 = 2ts$ liefert $r\,dr = t\,ds$. Also gilt

$$\begin{aligned}
\mathbb{E}\left[|X_{s+t} - X_s|^\beta\right] &= (2\pi t)^{-n/2} \cdot \mathcal{H}_{n-1}(S^{n-1}) \cdot \int_0^\infty e^{-s}(2ts)^{\beta/2+n/2-1} t\,ds \\
&= \frac{2^{\beta/2-1}}{\pi^{n/2}} \cdot \mathcal{H}_{n-1}(S^{n-1}) \cdot \Gamma\left(\frac{\beta}{2} + \frac{n}{2}\right) \cdot t^{\beta/2}
\end{aligned}$$

oder

$$\mathbb{E}\left[|X_t - X_s|^\beta\right] \leq C|t-s|^{1+\gamma}, \qquad \gamma = \frac{\beta}{2} - 1.$$

Für einen Wienerprozess ist die Abschätzung (8.4.4) also für ein beliebiges β und $\gamma = \beta/2 - 1$ erfüllt. Damit sind die Pfade eines Wienerprozesses zu jedem Zeitpunkt fast sicher α-Hölder-stetig, für jedes

$$\alpha < \frac{\frac{\beta}{2} - 1}{\beta} = \frac{1}{2} - \frac{1}{\beta}.$$

Da β beliebig ist, gilt dies also für jedes $\alpha < 1/2$.

Sei $(X_t)_{t\in[0,T)}$ ein Prozess, der die Vorraussetzungen des Stetigkeitssatzes 8.33 erfüllt. Ist nun $S \subseteq [0,T)$ eine abzählbare Menge von Zeiten, so hat (als abzählbarer Schnitt von Mengen mit vollem Maß) auch die Menge der Pfade, die an allen Punkten von S α-Hölderstetig sind, volles Maß. Dies bedeutet allerdings nicht, dass die Pfade an *jedem Punkt* von $[0,T)$ stetig sind (dies wäre ein Schnitt von überabzählbar vielen Mengen).

Um dennoch einen Prozess zu erhalten, dessen Pfade fast sicher *überall* stetig sind, kann man sich wie folgt behelfen: Sei $S \subseteq [0,T)$ eine abzählbare, dichte Teilmenge. Sei $N \subseteq \Omega$ die $\mathbb{P}$-Nullmenge der Pfade, die nicht α-Hölder-stetig auf S sind. Setze $\Omega' = \Omega \setminus N$ sei $\mathbb{P}'$ die Einschränkung von $\mathbb{P}$ auf Ω' (dies ist wieder ein Wahrscheinlichkeitsmaß). Jeder Pfad in Ω' ist nun stetig auf der dichten Teilmenge $S \subseteq [0,T)$, hat also eine eindeutige stetige Fortsetzung zu ganz $[0,T)$. Dies liefert eine Borel-messbare Abbildung

$$\Omega' \longrightarrow C([0,T),\mathbb{R}^n).$$

Verpflanzung des Maßes $\mathbb{P}'$ zu $C([0,T),\mathbb{R}^n)$ liefert ein Wahrscheinlichkeitsmaß auf $C([0,T),\mathbb{R}^n)$. Definiert man nun $\tilde{X}_t : C([0,T),\mathbb{R}^n) \to \mathbb{R}^n$ durch

$$\tilde{X}_t(\omega) = \omega(t),$$

so definiert dies einen Prozess mit stetigen Pfaden, mit der Eigenschaft, dass für alle $0 \le t_1 < \cdots < t_N < T$ die gemeinsame Verteilungen der Zufallsvektoren $(\tilde{X}_{t_1},\ldots,\tilde{X}_{t_N})$ und $(X_{t_1},\ldots,X_{t_N})$ übereinstimmt. Man benutzt hier folgende Terminologie:

Definition 8.35. Seien $(X_t)_{t\in[0,T)}$ und $(\tilde{X}_t)_{t\in[0,T)}$ M-wertige stochastische Prozesse (möglicherweise auf unterschiedlichen Wahrscheinlichkeitsräumen definiert). Wir sagen, dass $(\tilde{X}_t)_{t\in[0,T)}$ eine *Version von* $(X_t)_{t\in[0,T)}$ ist, falls, beide Prozesse dieselbe Verteilung haben.

Genauer: Wir verlangen, dass für beliebige Zeiten $0 \le t_1 < \cdots < t_N < T$ die zufälligen Vektoren $(X_{t_1},\ldots,X_{t_N})$ und $(\tilde{X}_{t_1},\ldots,\tilde{X}_{t_N})$ dieselben gemeinsamen Verteilungen haben, also dass

$$(X_{t_1},\ldots,X_{t_N})_*\mathbb{P} = (\tilde{X}_{t_1},\ldots,\tilde{X}_{t_N})_*\tilde{\mathbb{P}}$$

gilt.

Die obige Diskussion hat also das folgende Korollar:

Korollar 8.36. Es existiert eine Version eines Wiener-Prozesses im $\mathbb{R}^n$ mit stetigen Pfaden.