Skript zur Vorlesung Analysis 3

Herbstsemester 2022

Prof. Benjamin Schlein

Diese Notizen sind eine Zusammenfassung der Vorlesung “Analysis 3” und sie sind nur als Hilfsmittel gemeint. Die Notizen wurden vorallem aus dem Buch

- Donald L. Cohn. Measure Theory. Second Edition. Birkhäuser

zusammengestellt. Andere nützliche Referenzen sind

- J. Elstrodt. Mass- und Integrationstheorie. 2. Auflage. Springer
- H. Bauer. Mass- und Integrationstheorie. de Gruyter Lehrbuch.
- O. Forster. Analysis 3. Vieweg studium.

Inhaltsverzeichnis

1 Masstheorie

In der Vorlesung Analysis 1 wurde der Begriff von Riemann'sche Integral eingeführt.

Sei $f : [a; b] \to \mathbb{R}$ eine beschränkte Funktion, definiert auf dem kompakten Intervall $I = [a; b]$. Eine Teilung T von I ist eine endliche Teilmenge $T = \{a = x_0 < x_1 < \cdots < x_n = b\}$. Weiter, eine zu T entsprechende Familie von Representanten ist ein n-Tupel $\xi = \{\xi_1, \dots, \xi_n\}$ mit $\xi_j \in [x_{j-1}; x_j]$ für alle $j = 1, \dots, n$. Für gegebene Teilung T und Familie von Representanten ξ haben wir dann die Riemann'sche Summe

$$S(T, \xi) = \sum_{j=1}^{n} f(\xi_j)(x_j - x_{j-1})$$

definiert. Wir haben die obere Riemann'sche Summe zur Teilung T als

$$\overline{S}(T) = \sup_{\xi} S(T, \xi) = \sum_{j=1}^{n} \sup\{f(x) : x \in [x_{j-1}, x_j]\}(x_j - x_{j-1})$$

und die untere Riemann'sche Summe zu T als

$$\underline{S}(T) = \inf_{\xi} S(T, \xi) = \sum_{j=1}^{n} \inf\{f(x) : x \in [x_{j-1}, x_j]\}(x_j - x_{j-1})$$

definiert. Mit Hilfe von obere und untere Riemann'sche Summe haben wir dann den Begriff von Integrierbarkeit definiert. Wir sagen nämlich, dass f ist auf $[a; b]$ Riemann integrierbar, wenn

$$\sup_T \underline{S}(T) = \inf_T \overline{S}(T)$$

In diesem Fall definieren wir das Riemann'sche Integral von f auf $[a; b]$ durch

$$\int_a^b f(x)dx = \sup_T \underline{S}(T) = \inf_T \overline{S}(T)$$

Wir haben in Analysis 1 gezeigt, dass eine beschränkte Funktion $f : [a; b] \to \mathbb{R}$ mit nur endlich viele Unstetigkeitstellen integrierbar ist. Eine notwendige und hinreichende Bedingung für Integrierbarkeit kann mit Hilfe des Begriffs von Nullmenge bewiesen werden (wir haben aber diese Bedingung in Analysis 1nicht gezeigt). Eine Menge $N \subset \mathbb{R}$ heisst eine Nullmenge, falls für alle $\varepsilon > 0$ eine endliche oder abzählbare Familie $\{J_i\}$ von offenen Intervallen existiert, mit

$$N \subset \bigcup_i J_i \quad \text{und} \quad \sum_i |J_i| \leq \varepsilon$$

wobei $|J|$ die Länge des Intervalls J bezeichnet. Es gilt: eine beschränkte Funktion $f : [a; b] \to \mathbb{R}$ ist genau dann Riemann integrierbar, wenn die Menge aller Unstetigkeitstellen eine Nullmenge ist. Insbesondere ist jede Funktion mit abzählbar viele Unstitigkeitstellen Riemann integrierbar.

Es bleiben trotzdem ziemlich viele Funktionen die nicht Riemann integrierbar sind. Ein einfaches Beispiel ist die Funktion $f : [0;1] \to \mathbb{R}$, definiert durch

$$f(x) = \begin{cases} 1 & \text{falls } x \in \mathbb{Q} \cap [0;1] \\ 0 & \text{sonst} \end{cases}$$

Für eine beliebige Teilung $T = \{0 = x_0 < x_1 < \cdots < x_n = 1\}$, es gilt

$$\sup_{\xi \in [x_{j-1}, x_j]} f(\xi) = 1 \quad \text{und} \quad \inf_{\xi \in [x_{j-1}, x_j]} f(\xi) = 0$$

weil jedes nicht leeres Intervall $[x_{j-1}, x_j]$ sicher ein Punkt $\xi_1 \in \mathbb{Q}$ und ein Punkt $\xi_2 \in \mathbb{R} \backslash \mathbb{Q}$ enthält. Deswegen gilt

$$\overline{S}(T) = 1, \quad \text{und} \quad \underline{S}(T) = 0$$

für jede Teilung T. Es folgt: die Funktion f ist nicht auf $[0;1]$ integrierbar (die Funktion f ist nirgends stetig; die nicht Integrierbarkeit von f folgt auch aus der Bedingung, die wir oben erwähnt haben).

Die Existenz von vielen nicht integrierbaren Funktionen ist die Hauptschwäche des Begriffes von Riemann Integral. In dieser Vorlesung möchten wir ein neues Integral einführen, das sogenannte Lebesgue'sche Integral, das uns erlauben wird, eine grössere Klasse von Funktionen zu integrieren (das Lebesgue'sche Integral einer Riemann integrierbare Funktion wird aber mit ihrem Riemann'sche Integral übereinstimmen). Wir werden sehen, dass der Begriff von Lebesgue'sche Integral einen anderen wichtigen Vorteil hat, verglichen mit dem Riemann'sche Integral. Während das Riemann'sche Integral zunächst nur für Funktionen auf $\mathbb{R}$ definiert ist (und auf Funktionen auf $\mathbb{R}^n$ verallgemeinert werden kann), ist das Lebesgue'sche Integral direkt auf allgemeineren Räumen (sogenannte Massräume) definiert.

Wir erklären, kurz und heuristisch, die Idee des Lebesgue'sche Integrals. Sei $f : [a;b] \to \mathbb{R}$ eine beschränkte Funktion. Nehmen wir an, f ist stetig und nicht negativ. Dann ist f Riemann integrierbar, und

$$\int_a^b f(x)dx = \int_a^b \left[\int_0^\infty \chi(f(x) \geq t) dt \right] dx$$

Hier bezeichnet $\chi(f(x) \geq t)$ die charakteristische Funktion vom Intervall $[o, f(x)]$ (d.h. $\chi(f(x) \geq t) = 1$ falls $f(x) \geq t$ und $\chi(f(x) \geq t) = 0$ falls $f(x) < t$). Man kann dann die zwei Integrale vertauschen (wir haben in Analysis 2 den Theorem von Fubini bewiesen); man bekommt

$$\int_a^b f(x)dx = \int_0^\infty \left[\int_a^b \chi(f(x) \geq t) dx \right] dt$$

Z.B. falls die Funktion f konvex ist, man kann sich leicht überzeugen, dass die Menge $\{x \in [a;b] : f(x) \geq t\}$ ein Intervall ist (wo $\chi(f(x) \geq t) = 1$ gilt), für alle $t \in \mathbb{R}$ (für t gross genug, die Menge ist leer). Es gilt dann

$$\int_a^b \chi(f(x) \geq t) dx = \mu(\{x \in [a;b] : f(x) \geq t\})$$

wobei die rechte Seite die Länge des Intervalls ist, auf welchem $f(x) \geq t$. Damit ist

$$\int_a^b f(x)dx = \int_0^\infty \mu(\{x \in [a;b] : f(x) \geq t\})dt \tag{1}$$

Für nicht konvexe Funktionen, ist $\{x \in [a;b] : f(x) \geq t\}$ i.A. kein Intervall, sondern die Vereinigung von viele (möglicherweise unendlich viele) Intervalle. Trotzdem, mindestens falls die Funktion regulär ist, ist es klar, wie man die gesamte Länge dieser Intervalle bestimmen kann. Die Formel (1) gilt damit für beliebige $f : [a;b] \to \mathbb{R}$ nicht-negativ und stetig. Die rechte Seite von (1) hat nun aber einen Vorteil, verglichen mit der linken Seite. Die Funktion $\mu(\{x \in [a;b] : f(x) \geq t\})$ ist monoton fallend in t. Eine monoton fallende Funktion hat immer höchstens abzählbar viele Unstetigkeiten, und deswegen ist immer Riemann integrierbar. Aus diesem Grund können wir die rechte Seite von (1) benutzten, um das Lebesgue Integral von beliebige Funktionen zu definieren, für welchen $\mu(\{x \in [a;b] : f(x) \geq t\})$ für beliebige $t > 0$ definiert werden soll. Für stetige Funktionen wird das neue Integral wegen (1) mit dem alten Riemann Integral übereinstimmen. Die neue Definition kann aber auf allgemeineren Funktionen angewandt werden. Die einzige Bedingung ist, dass wir die Länge von Mengen der Form $\{x \in [a;b] : f(x) \geq t\}$ definieren können.

Es ist auch klar, dass die rechte Seite von (1) uns auch erlaubt das Integral von Funktionen zu definieren, die auf $\mathbb{R}^n$ oder sogar auf allgemeineren Räume definiert sind. Wichtig ist nur, dass wir $\mu(\{x : f(x) \geq t\})$ definieren können. Ist f auf einer Teilmenge von $\mathbb{R}^n$ definiert, dann ist $\mu(\{x : f(x) \geq t\}))$ das Volumen der Menge $\{x : f(x) \geq t\}) \subset \mathbb{R}^n$. Auf allgemeineren abstrkaten Räume, als wir sehen werden, heisst $\mu(A)$ das Mass der Menge A (Länge und Volumen sind Beispiele von Masse auf $\mathbb{R}$, bzw. auf $\mathbb{R}^n$).

Aus dieser kurzen und heuristischen Diskussion ist hoffentlich klar geworden, dass um den Begriff von Lebesgue'sche Integral einzuführen, brauchen wir das Mass von möglichst allgemeine Mengen zu definieren. Das ist der Zweck von dieser Kapitel.

Bevor wir zu den Definitionen kommen, möchten wir zeigen, dass das Problem der Definition des Masses nicht so trivial ist. Wir untersuchen das Volumen von Teilmengen von $\mathbb{R}^n$. Wir bezeichnen mit $P(\mathbb{R}^n) \equiv 2^{\mathbb{R}}$ die Potenzmenge von $\mathbb{R}^n$, d.h. die Menge aller Teilmengen von $\mathbb{R}^n$. Wir suchen eine Abbildung $\mu : P(\mathbb{R}^n) \to [0;\infty]$ mit den Eigenschaften

i) Monotonie: ist $A \subset B$, so gilt $\mu(A) \leq \mu(B)$.

ii) Euklidische Invarianz: ist $T : \mathbb{R}^n \to \mathbb{R}^n$ eine affine Isometrie und $A \subset \mathbb{R}^n$, so gilt $\mu(TA) = \mu(A)$ (eine affine Isometrie ist eine Abbildung der Form $T(x) = L(x) + b$, wobei $b \in \mathbb{R}^n$ eine Translation darstellt, und $L : \mathbb{R}^n \to \mathbb{R}^n$ eine lineare Abbildung mit $L^T L = 1$ ist).

iii) Normierung: $\mu([0;\ell]^n) = \ell^n$, für all $\ell > 0$.

iv) σ-Additivität: sind $A_1, A_2, \dots \subset \mathbb{R}^n$ abzählbar viele paarweise disjunkte Teilmengen von $\mathbb{R}^n$, so gilt

$$\mu\left(\bigcup_{j=1}^\infty A_j\right) = \sum_{j=1}^\infty \mu(A_j)$$

Man bemerke, dass die Monotonie i) eigentlich aus der σ-Additivität iv) folgt (und braucht deswegen nicht separat angenommen zu werden), weil $A \subset B$ impliziert, dass $B = A \cup (B \cap A^c)$, und also, da A und $B \cap A^c$ disjunkt sind,

$$\mu(B) = \mu(A) + \mu(B \cap A^c) \geq \mu(A)$$

weil $\mu(B \cap A^c) \geq 0$. Überraschend, es existiert keine Abbildung mit den Eigenschaften i)-iv).

Satz 1.1 (Vitali 1905)**.** *Es existiert keine Abbildung $\mu : P(\mathbb{R}^n) \to [0;\infty]$ mit den Eigenscahften i)-iv). Das gilt auch wenn ii) durch die schwächere Bedingung*

ii') Invarianz bzg. Translationen: Für alle $x \in \mathbb{R}^n$ und alle $A \subset \mathbb{R}^n$ gilt $\mu(x + A) = \mu(A)$.

ersetzt wird.

Beweis. Wir nehmen an, es existiere eine Abbildung μ mit den Eigenschaften i), ii'), iii), iv). Auf $A = [0; 1]^n$ definieren wir die Äquivalenzrelation $x \sim y$, falls $x - y \in \mathbb{Q}^n$. Sei nun $M_0 \subset A$ eine Teilmenge, die aus jeder Äquivalenzklasse genau ein Element enthält (hier benutzen wir das Auswahlaxiom). Dann gilt

$$(1) \text{ für jedes } x \in A, \text{ es existiert } y \in M_0 \text{ mit } x \sim y$$

und

$$(2) \text{ sind } x, y \in M_0 \text{ mit } x \sim y, \text{ so muss } x = y \text{ gelten}$$

Da $\mathbb{Q}^n \cap [-1, 1]^n$ abzählbar ist, finden wir eine Bijektion $x : \mathbb{N} \to \mathbb{Q}^n \cap [-1; 1]^n$ mit $x_0 = 0$. Wir definieren $M_j = M_0 + x_j$. Dann gilt

a) Für $j \neq k$ ist $M_j \cap M_k = \emptyset$. In der Tat, falls $M_j \cap M_k \neq \emptyset$, dann könnten wir $y, z \in M_0$ finden, mit $y + x_j = z + x_k$. Das impliziert $y - z = x_k - x_j \in \mathbb{Q}^n$. Also $y \sim z$ und, aus Bemerkung (2) oben, $y = z$. Das gibt aber $x_j = x_k$, was ein Widerspruch zur Injäktivität von x ist.

b) Für alle $j \in \mathbb{N}$ gilt $\mu(M_j) = \mu(M_0) \leq \mu(A) = 1$. Das folgt aus der Eigenschaften ii') (Translationsinvarianz), i) (Monotonie) und iii) (Normierung).

c) Es gilt

$$[0, 1]^n \subset \bigcup_{j \in \mathbb{N}} M_j$$

Sei, in der Tat, $y \in [0, 1]^n$. Dann gibt es aus Bemerkung (1) ein $z \in M_0$ mit $y \sim z$. Deswegen gilt $y - z \in \mathbb{Q}^n \cap [-1, 1]^n$, und es existiert $j \in \mathbb{N}$ mit $x_j = y - z$. Das zeigt, dass $y = z + x_j \in M_j$.

d) Es gilt

$$\bigcup_{j \in \mathbb{N}} M_j \subset [-1, 2]^n$$

In der Tat, für $y \in M_j$ existiert $z \in M_0$ mit $y = z + x_j$. Da $x_j \in [-1, 1]^n$ und $z \in [0, 1]^n$ folgt aber $y \in [-1, 2]^n$.

Wir unterscheiden nun zwei Fälle. Ist $\mu(M_0) = 0$, dann folgt aus c), dass

$$1 = \mu([0;1]^n) \leq \mu\left(\bigcup_{j\in\mathbb{N}} M_j\right) = \sum_{j\in\mathbb{N}} \mu(M_h) = \sum_{j\in\mathbb{N}} \mu(M_0) = 0$$

Gilt dagegen $\mu(M_0) > 0$, so folgt aus d), dass

$$3^n = \mu([-1;2]^n) \geq \mu\left(\bigcup_{j\in\mathbb{N}} M_j\right) = \sum_{j\in\mathbb{N}} \mu(M_j) = \infty$$

In beiden Fällen finden wir ein Widerspruch. $\square$

Man kann sich vorstellen, wir können ein Volumen definieren, indem wir die σ-Additivität durch die schwächere Bedingung

iv') Endliche Additivität: sind $A_1, A_2, \dots, A_k$ endlich viele paarweise disjunkte Teilmengen von $\mathbb{R}^n$, so gilt

$$\mu\left(\bigcup_{j=1}^{k} A_j\right) = \sum_{j=1}^{k} \mu(A_j)$$

ersetzen. Für $n \geq 3$, ist aber auch unter dieser schwächeren Bedingung nicht möglich, ein Volumen zu definieren.

Satz 1.2. *Für $n \geq 3$ gibt es keine Abbildung $\mu : P(\mathbb{R}^n) \to [0;\infty]$ die die Eigenschaften i), ii), iii), iv') erfüllt.*

Bemerkung: Für die Spezialfälle $n = 1, 2$, es ist dagegen möglich ein Volumen auf $\mathbb{R}^n$ zu definieren, mit den Eigenschaften $i), ii), iii), iv')$.

Satz 1.2 und die letzte Bemerkung werden unten im Appendix A bewiesen (das Appendix ist nicht für die Prüfung relevant).

Satz 1.1 und Satz 1.2 zeigen, dass die Definition eines Volumen auf $\mathbb{R}^n$ nicht ganz einfach sein kann. Man findet, dass die beste Lösung um Volumen (insbesondere σ-additive Volumen) einzuführen, ist die Volumenfunktion μ nur auf einer Teilmenge von $P(\mathbb{R}^n)$ zu definieren. Wir werden nicht versuchen, $\mu(A)$ für alle $A \subset \mathbb{R}^n$ zu definieren; dagegen werden wir nur $\mu(A)$ definieren, für A aus einer Klasse von Teilmengen in $\mathbb{R}^n$, nämlich aus einer geeigneten σ-Algebra (damit die Definition nützlich sein kann, müssen wir sicher stellen, dass wir das Volumen von genügend viele Mengen bestimmen können). Der Begriff von σ-Algebra wird im nächsten Abschnitt eingeführt.

1.1 σ-Algebren

Im Gegensatz zur bisherigen Diskussion, wechseln wir nun zu einem abstrakten Setting, wo wir Teilmengen einer beliebigen Menge Ω (und nicht notwendigerweise von $\mathbb{R}^n$) betrachten. Wir werden später das Beispiel $\Omega = \mathbb{R}^n$ speziell untersuchen.

Sei Ω eine beliebige Menge und

$$P(\Omega) = \{A : A \subset \Omega\}$$

die entsprechende Potenzmenge.

Definition 1.3. *Eine Familie $\mathcal{A} \subset P(\Omega)$ von Teilmengen von Ω heisst eine σ-Algebra, wenn*

i) $\Omega \in \mathcal{A}$.

ii) $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$.

iii) Ist $(A_n)_{n\in\mathbb{N}}$ eine Folge von Mengen in $\mathcal{A}$, so ist auch

$$\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$$

Hier bezeichnen wir mit $A^c \equiv \Omega\backslash A = \{x \in \Omega : x \notin A\}$ das Komplement von A in Ω. Die Eigenschaft ii) bedeutet, dass σ-Algebren stabil sind bzg. der Operation $A \to A^c$. Eigenschaft iii) bedeutet dagegen, dass σ-Algebren stabil sind, bzg. abzählbare Vereinigungen.

Bemerkungen: aus der Definition, finden wir sofort die folgenden weitere Eigenschaften von σ-Algebren:

1) $\emptyset \in \mathcal{A}$. Das folgt aus i) und ii), weil $\emptyset = \Omega^c$.

2) σ-Algebren sind stabil bzg. abzählbaren Durchschnitten. Mit anderen Wörter: ist $(A_n)_{n\in\mathbb{N}}$ eine Folge in $\mathcal{A}$, so ist

$$\bigcap_{n\in\mathbb{N}} A_n \in \mathcal{A} \tag{2}$$

Um (2) zu beweisen, benutzen wir die Morgan'sche Regeln (die wir in Analysis 1 schon erwähnt haben; wir lassen den Beweis dieser Regeln als Übung). Wir finden, mit ii) und iii),

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

Aus ii) folgt deswegen, dass

$$\bigcap_{n\in\mathbb{N}} A_n \in \mathcal{A}\,.$$

3) Die Stabilität von σ-Algebren bzg. abzählbare Vereinigungen und abzählbare Durchschnitte impliziert natürlich auch die Stabilität bzg. endlichen Vereinigungen und endlichen Durchschnitten. Ist nämlich $A_1, \dots, A_m$ eine endliche Folge aus $\mathcal{A}$, so können wir $A_{m+1} = A_{m+2} = \cdots = \emptyset \in \mathcal{A}$ definieren. Dann ist

$$\bigcup_{j=1}^{m} A_j = \bigcup_{j\in\mathbb{N}} A_j \in \mathcal{A}$$

wegen iii). Analog zeigt man, dass endliche Durchschnitte von Mengen in $\mathcal{A}$ wieder in $\mathcal{A}$ enthalten sind.

Wir diskutieren nun ein Paar Beispiele von σ-Algebren.

- Die Potenzmenge $P(\Omega)$ ist immer eine σ-Algebra (die grösste σ-Algebra auf Ω).
- Die Familie $\mathcal{A} = \{\emptyset, \Omega\}$ ist auch immer eine σ-Algebra (die kleinste σ-Algebra auf Ω).
- Für $A \subset \Omega$ beliebig ist $\mathcal{A} = \{\emptyset, A, A^c, \Omega\}$ eine σ-Algebra.
- Die Familie
$$\mathcal{A} = \{A \subset \Omega : A \text{ ist abzählbar oder } A^c \text{ ist abzählbar}\}$$
ist eine σ-Algebra. Offenbar ist $\Omega \in \mathcal{A}$, weil $\Omega^c = \emptyset$ eine abzählbare Menge ist (abzählbar bedeutet hier entweder endlich oder abzählbar unendlich; eine Menge mit Null Elemente ist in diesem Sinn abzählbar). Da die Definition symmetrisch bezüglich der Operation $A \to A^c$ ist, ist ii) auch erfüllt. Sei nun $A_1, A_2, \dots$ eine Folge aus $\mathcal{A}$. Wir möchten zeigen, dass
$$\bigcup_{n \geq 1} A_n \in \mathcal{A} \tag{3}$$
Dazu unterscheiden wir zwei Fälle: existiere ein m mit A_m nicht abzählbar, so muss A_m^c abzählbar sein. Deswegen ist
$$\left(\bigcup_{n \geq 1} A_n \right)^c = \bigcap_{n \geq 1} A_n^c \subset A_m^c$$
sicher abzählbar, und (3) ist erfüllt. Falls dagegen A_n abzählbar für alle $n \in \mathbb{N}$, so ist auch
$$\bigcup_{n \geq 1} A_n$$
als abzählbare Vereinigung von abzählbare Mengen wieder abzählbar. Also gilt (3) auch in diesem Fall.
- Ist $\mathcal{A}$ eine σ-Algebra auf Ω, und ist $\Omega' \subset \Omega$ eine beliebige Teilmenge, dann ist
$$\mathcal{A}' = \{\Omega' \cap A : A \in \mathcal{A}\}$$
eine σ-Algebra auf Ω'. Man nennt $\mathcal{A}'$ die Spur von $\mathcal{A}$ auf Ω'. Um zu zeigen, dass $\mathcal{A}'$ eine σ-Algebra ist, man bemerke zunächst, dass $\Omega' = \Omega' \cap \Omega \in \mathcal{A}'$. Weiter, das Komplement von $\Omega' \cap A$ in Ω' ist aus
$$(\Omega' \cap A)^c = \{x \in \Omega' : x \notin \Omega' \cap A\} = \{x \in \Omega' : x \notin A\} = \Omega' \cap A^c$$
gegeben (A^c bezeichnet hier das Komplement von A in Ω). Gilt $A \in \mathcal{A}$, so ist $A^c \in \mathcal{A}$, und deswegen $\Omega' \cap A^c \in \mathcal{A}'$. Schlussendlich, verifizieren wir die Bedingung iii) für $\mathcal{A}'$. Sei $(\Omega' \cap A_1), (\Omega' \cap A_2), \dots$ eine Folge aus $\mathcal{A}'$. Dann ist $A_1, A_2, \dots$ eine Folge aus $\mathcal{A}$, und deswegen
$$\bigcup_{n \geq 1} A_n \in \mathcal{A}$$
Damit gilt
$$\bigcup_{n \geq 1} (\Omega' \cap A_n) = \Omega' \cap \bigcup_{n \geq 1} A_n \in \mathcal{A}'$$

- Seien Ω und Ω' zwei Mengen, $\mathcal{A}'$ eine σ-Algebra auf Ω' und $T : \Omega \to \Omega'$ eine Abbildung. Dann ist
$$T^{-1}(\mathcal{A}') = \{T^{-1}(A') : A' \in \mathcal{A}'\}$$
eine σ-Algebra auf Ω. Beweis: Übung.

Für die Konstruktion von σ-Algebren spielt der folgende Satz eine wichtige Rolle.

Satz 1.4. *Sei Ω eine Menge, und $(\mathcal{A}_i)_{i\in I}$ eine Familie von σ-Algebren auf Ω. Dann ist auch der Durchschnitt*
$$\bigcap_{i\in I} \mathcal{A}_i$$
eine σ-Algebra auf Ω.

Bemerkung: hier braucht die Indexmenge I nicht abzählbar zu sein.

Beweis. Wir bezeichnen
$$\mathcal{A} = \bigcap_{i\in I} \mathcal{A}_i \,.$$
Wir zeigen, dass $\mathcal{A}$ eine σ-Algebra ist. Da $\mathcal{A}_i$ eine σ-Algebra auf Ω ist, gilt $\Omega \in \mathcal{A}_i$ für alle $i \in I$. Damit ist auch $\Omega \in \mathcal{A}$. Sei weiter $A \in \mathcal{A}$. Dann ist $A \in \mathcal{A}_i$ und deswegen auch $A^c \in \mathcal{A}_i$ für alle $i \in I$. Es folgt, dass $A^c \in \mathcal{A}$. Schlussendlich, sei $(A_n)_{n\in\mathbb{N}}$ eine Folge aus $\mathcal{A}$. Für alle $n \in \mathbb{N}$ und alle $i \in I$ gilt also $A_n \in \mathcal{A}_i$. Deswegen ist
$$\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}_i$$
für alle $i \in I$. Das impliziert, dass $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$. □

Für eine beliebige Familie $\mathcal{F}$ von Teilmengen von Ω, wir können nun
$$\sigma(\mathcal{F}) = \bigcap \{\mathcal{A} : \mathcal{A} \text{ ist eine } \sigma\text{-Algebra auf } \Omega \text{ und} \mathcal{F} \subset \mathcal{A}\}$$
definieren. Da $P(\Omega)$ immer eine σ-Algebra ist, ist die rechte Seite sicher wohldefiniert. Offenbar ist $\mathcal{F} \subset \sigma(\mathcal{F})$, und aus Satz 1.4 ist $\sigma(\mathcal{F})$ eine σ-Algebra auf Ω. Weiter: ist $\mathcal{A}$ eine σ-Algebra auf Ω mit $\mathcal{F} \subset \mathcal{A}$, dann gilt $\sigma(\mathcal{F}) \subset \mathcal{A}$. mit anderen Wörter: $\sigma(\mathcal{F})$ ist die kleinste σ-Algebra, die $\mathcal{F}$ enthält. Man nennt $\sigma(\mathcal{F})$ die von $\mathcal{F}$ erzeugten σ-Algebra auf Ω (und $\mathcal{F}$ ist der Erzeuger der σ-Algebra $\sigma(\mathcal{F})$).

Beispiele.

- Ist $\mathcal{A}$ eine σ-Algebra, so gilt $\sigma(\mathcal{A}) = \mathcal{A}$.
- Ist $\mathcal{F} = \{A\}$ für ein einziges $A \subset \Omega$, so gilt $\sigma(\mathcal{F}) = \{\emptyset, A, A^c, \Omega\}$.

Für uns werden die wichtigsten Beispiele von σ-Algebra die Borel σ-Algebren auf $\mathbb{R}^n$ sein, die aus allen offenen Mengen in $\mathbb{R}^n$ erzeugt werden.

Definition 1.5. *Für $n \in \mathbb{N}$, sei*
$$\mathcal{G}_n = \{A \subset \mathbb{R}^n : A \textit{ offen}\}$$
Dann ist $\mathcal{B}(\mathbb{R}^n) = \sigma(\mathcal{G}_n)$ die Borel σ-Algebra auf $\mathbb{R}^n$.

Als Erinnerung: eine Menge $A \subset \mathbb{R}^n$ heisst offen, falls für jede $x \in A$ ein $\varepsilon > 0$ existiert, mit $\{y \in \mathbb{R}^n : \|x - y\| \leq \varepsilon\} \subset A$. Hier ist $\|x\| = (\sum_{j=1}^n |x_j|^2)^{1/2}$ die euklidische Norm von $x = (x_1, \dots, x_n)$ (jede Norm auf $\mathbb{R}^n$ führt aber zu der selben Topologie).

Satz 1.6. *Die Borel σ-Algebra $\mathcal{B}(\mathbb{R})$ auf $\mathbb{R}$ wird auch von den folgenden Systemen von Teilmengen von $\mathbb{R}$ erzeugt:*

a) die Familie aller abgeschlossenen Teilmengen von $\mathbb{R}$.

b) die Familie der Intervalle der Form $(-\infty; b]$, mit $b \in \mathbb{R}$.

c) die Familie der halboffenen Intervalle $(a, b]$ mit $-\infty < a < b < \infty$.

Beweis. Seien $\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3$ die σ-Algebren, die aus den Systemen in a), b), c) erzeugt werden. Wir zeigen die Inklusionen:

- $\mathcal{B}(\mathbb{R}) \supset \mathcal{B}_1$. Da $\mathcal{B}_1$ die kleinste σ-Algebra ist, die alle abgeschlossenen Teilmengen von $\mathbb{R}$ enthält, genügt es zu zeigen, dass $\mathcal{B}(\mathbb{R})$ alle abgeschlossenen Teilmengen von $\mathbb{R}$ enthält. Das ist aber klar, weil $\mathcal{B}(\mathbb{R})$ stabil bezüglich der Operation $A \to A^c$, und weil $\mathcal{B}(\mathbb{R})$ alle offene Teilmengen von $\mathbb{R}$ enthält.

- $\mathcal{B}_1 \supset \mathcal{B}_2$. Die Intervalle $(-\infty, b]$ sind abgeschlossen, und deswegen sicher in $\mathcal{B}_1$. Da $\mathcal{B}_2$ die kleinste σ-Algebra ist, die alle Intervalle der Form $(-\infty, b]$ enthält, muss $\mathcal{B}_1 \supset \mathcal{B}_2$ gelten.

- $\mathcal{B}_2 \supset \mathcal{B}_3$. Es gilt
$$(a, b] = (-\infty, b] \cap (-\infty, a]^c \in \mathcal{B}_2$$
für alle $-\infty < a < b < \infty$. Da $\mathcal{B}_3$ die kleinste σ-Algebra ist, die alle Intervalle der Form $(a; b]$ enthält, muss $\mathcal{B}_2 \supset \mathcal{B}_3$.

- $\mathcal{B}_3 \supset \mathcal{B}(\mathbb{R})$. Zu zeigen: jede offenen Teilmenge von $\mathbb{R}$ ist in $\mathcal{B}_3$ enthalten. Wir bemerken zunächst, dass jede offene Intervall $(a; b) \subset \mathbb{R}$ in $\mathcal{B}_3$ enthalten ist. Das folgt weil
$$(a; b) = \bigcup_{n \in \mathbb{N}} (a; b - (1/n)]$$
und weil $\mathcal{B}_3$, wie jede σ-Algebra, bezüglich abzählbaren Vereinigungen stabil ist. Die Behauptung folgt nun aus Lemma 1.7 (siehe unten), wo wir zeigen, dass jede offene Teilmenge von $\mathbb{R}$ als abzählbare Vereinigung von offenen Intervallen geschrieben werden kann.

□

Lemma 1.7. *Sei $U \subset \mathbb{R}$ offen. Dann gibt es eine Folge I_n von offenen Intervallen in $\mathbb{R}$, mit $U = \bigcup_{n \in \mathbb{N}} I_n$.*

Beweis. Wir definieren die Familie

$$\mathcal{U} = \{I : \; I \text{ ist ein offenes Intervall, enthalten in } U \text{ und maximal}\}$$

Ein Intervall $I \subset U$ heisst maximal, falls $I \subset J \subset U$ für ein Intervall J impliziert, dass $I = J$. Offenbar gilt

$$\bigcup_{I \in \mathcal{U}} I \subset U$$

weil $I \subset U$ für jede $I \in \mathcal{U}$. Anderseits, für ein beliebiges $x \in U$ ist

$$I_x = \bigcup \{I : \; I \text{ ist ein offenes Intervall, mit } I \subset U \text{ und } x \in I\}$$

in $\mathcal{U}$ enthalten. In der Tat, I_x ist nicht leer, weil U offen ist. Weiter, I_x ist, als Vereinigung von offenen Intervallen (die ein gemeinsamer Punkt x enthalten) sicher ein offenes Intervall, und I_x ist (bei Definition) offenbar maximal. Damit haben wir, für alle $x \in U$ ein $I_x \in \mathcal{U}$ gefunden, mit $x \in I_x$. Das zeigt, dass $U \subset \bigcup_{I \in \mathcal{U}} I$ und damit, dass

$$U = \bigcup_{I \in \mathcal{U}} I \,.$$

Wir müssen noch zeigen, dass $\mathcal{U}$ abzählbar ist. Dazu bemerken wir, dass die Intervalle in $\mathcal{U}$ disjunkt sind. Ist in der Tat $x \in I_1 \cap I_2$, für $I_1, I_2 \in \mathcal{U}$, so ist auch $I_1 \cup I_2$ ein offenes Intervall, der in U enthalten ist. Aus der Maximalität von I_1, I_2, muss also $I_1 = I_2 = I_1 \cup I_2$, d.h. $I_1 = I_2$. Wir können also für jede $I \in \mathcal{U}$ eine rationale Zahl $x_I \in \mathbb{Q}$ wählen (weil jedes offenes Intervall eine rationale Zahl enthält). Das definiert eine Abbildung $\mathcal{U} \to \mathbb{Q}$, die injäktiv ist. Damit ist $\mathcal{U}$ abzählbar. $\square$

Analog zu Satz 1.6, bekommen wir in höheren Dimensionen den folgenden Satz. Der Beweis ist analog zum Beweis von Satz 1.6 und ist als Übung gelassen.

Satz 1.8. *Die Borel σ-Algebra $\mathcal{B}(\mathbb{R}^n)$ auf $\mathbb{R}^n$ wird auch von den folgenden Systemen von Teilmengen von $\mathbb{R}^n$ erzeugt:*

a) die Familie aller abgeschlossenen Teilmengen von $\mathbb{R}^n$.

b) die Familie der abgeschlossenen Halb-Ebenen der Form

$$\{(x_1, \dots, x_n) \in \mathbb{R}^n : x_i \leq b\}$$

für ein $i \in \{1, \dots, n\}$ und ein $b \in \mathbb{R}$.

c) die Familie der Quadern der Form

$$\{(x_1, \dots, x_n) \in \mathbb{R}^n : a_i < x_i \leq b_i \text{ für alle } i = 1, \dots, n\}$$

für $-\infty < a_i < b_i < \infty$ für alle $i \in \{1, \dots, n\}$.

Wir werden sehen, die Borel σ-Algebra enthält nicht alle Teilmengen von $\mathbb{R}^n$, d.h. $\mathcal{B}(\mathbb{R}^n) \neq P(\mathbb{R}^n)$. Trotzdem, die Borel σ-Algebra enthält ziemlich viele Teilmengen von $\mathbb{R}^n$ (es ist nicht trivial, obwohl möglich, ein $A \subset \mathbb{R}^n$ zu konstruieren, mit $A \notin \mathcal{B}(\mathbb{R}^n)$). Man findet, die Borel σ-Algebra ist ein guter Kompromiss um Analysis zu machen; sie enthält praktisch jeder Teilmenge, die für die Analysis wichtig ist und, gleichzeitig, sie ist klein genug, dass ihre Teilmengen konstruktiv behandelt werden können (insbesondere, wir werden sehen, sie ist klein genug, um die Probleme mit der Definition eines Volumen zu vermeiden).

Manchmal ist nützlich σ-Algebren aus Systemen von Teilmengen von Ω zu erzeugen, die schon teilweise die Eigenschaften einer σ-Algebra haben. Es spielen dabei Ringe und Algebren eine wichtige Rolle.

Definition 1.9. *Ein System $\mathcal{R}$ von Teilmengen einer Menge Ω heisst ein Ring in Ω falls*

i) $\emptyset \in \mathcal{R}$.

ii) *Für* $A, B \in \mathcal{R}$ *gilt auch* $A \backslash B \in \mathcal{R}$.

iii) *Für* $A, B \in \mathcal{R}$ *gilt auch* $A \cup B \in \mathcal{R}$.

Gilt zusätzlich auch

iv) $\Omega \in \mathcal{R}$.

so heisst der Ring $\mathcal{R}$ eine Algebra.

Hier bezeichnen wir $A\backslash B = \{x \in A : x \notin B\} = A \cap B^c$. Aus Definition ist jede Algebra ein Ring.

Bemerkung. Sei $\mathcal{R}$ ein Ring, und $A, B \in \mathcal{R}$. Dann ist auch $A \cap B = A\backslash(A\backslash B) \in \mathcal{R}$. Also Ringen sind geschlossen bezüglich endliche Vereinigung und endliche Durchschnitte.

Satz 1.10. *Eine Familie $\mathcal{B}$ von Teilmengen einer Menge Ω ist genau dann eine Algebra, wenn*

a) $\Omega \in \mathcal{B}$.

b) $A \in \mathcal{B} \;\Rightarrow\; A^c \in \mathcal{B}$.

c) $A, B \in \mathcal{B} \;\Rightarrow\; A \cup B \in \mathcal{B}$.

Beweis. Sei $\mathcal{B}$ eine Algebra. Dann gelten a) und c) aus Definition, und b) folgt aus der Eigenschaft ii) weil $A^c = \Omega\backslash A$. Gelten die Eigenschaften a), b), c) so gelten auch i), iii). Um ii) zu zeigen, bemerken wir, dass

$$A\backslash B = A \cap B^c = (A^c \cup B)^c$$

□

Der Unterschied zwischen Algebren und σ-Algebren ist also, dass Algebren nur unter endliche Vereinigungen (und Durchschnitte) während σ-Algebren auch unter abzählbare Vereinigungen (und Durchschnitte) stabil sind. Insbesondere ist jede σ-Algebra auch eine Algebra.

Beispiele: ein Paar Beispiele von Algebren und Ringen.

- $\mathcal{R} = \{\emptyset\}$ ist ein Ring, aber keine Algebra. Die kleinste Algebra auf einer Menge Ω besteht aus $\mathcal{A} = \{\emptyset, \Omega\}$ und ist gleichzeitig auch eine σ-Algebra.

- Das System $\mathcal{R} = \{A \subset \Omega : A \text{ endlich ist}\}$ ist immer ein Ring. $\mathcal{R}$ ist nur für endliches Ω auch eine Algebra (in diesem Fall $\mathcal{R} = P(\Omega)$ enthält alle Teilmengen von Ω).

- Das System $\mathcal{A} = \{A \subset \Omega : A \text{ oder } A^c \text{ endlich ist}\}$ ist eine Algebra auf Ω (weil endliche Vereinigungen von endlichen Mengen wieder endlich sind). $\mathcal{A}$ ist nur dann eine σ-Algebra, wenn Ω endlich ist (in diesem Fall ist $\mathcal{A} = P(\Omega)$).

Um zu entscheiden ob eine gegebene Familie von Mengen eine σ-Algebra ist, ist manchmal der Begriff von Dynkin-System nützlich.

Definition 1.11. *Sei Ω eine Menge. Eine Familie $\mathcal{D}$ von Teilmengen von Ω heisst ein Dynkin-System, falls*

i) $\Omega \in \mathcal{D}$.

ii) $D \in \mathcal{D} \Rightarrow D^c \in \mathcal{D}$.

iii) Für jede Folge $(D_n)_{n\in\mathbb{N}}$ von paarweise disjunkten Mengen aus $\mathcal{D}$ ist auch

$$\bigcup_{n\in\mathbb{N}} D_n \in \mathcal{D}\,.$$

Es folgt aus der Definition, dass $\emptyset = \Omega^c \in \mathcal{D}$. Der Unterschied mit dem Begriff von σ-Algebren ist, dass hier iii) nur für paarweise disjunkten Mengen überprüft werden soll. Insbesondere ist jede σ-Algebra ein Dynkin System. Umgekehrt ist dagegen nicht jede Dynkin System eine σ-Algebra. Betrachte zum Beispiel $\Omega = \{1, 2, \dots, 2n\}$, für ein $n \in \mathbb{N}$. Dann ist $\mathcal{D} = \{A \subset \Omega : |A| \text{ gerade ist}\}$ offenbar ein Dynkin System (weil die Vereinigung von zwei disjunkte Mengen mit gerade Kardinalität wieder gerade Kardinalität hat), aber keine σ-Algebra (weil zum Beispiel $\{1, 2\}, \{2, 3\} \in \mathcal{D}$ aber $\{1, 2\} \cup \{2, 3\} = \{1, 2, 3\} \notin \mathcal{D}$).

Lemma 1.12. *Sei $\mathcal{D}$ ein Dynkin System auf einer Menge Ω. Es gilt: sind $A, B \in \mathcal{D}$ mit $B \subset A$, dann ist auch $A \backslash B \in \mathcal{D}$*

Beweis. Da $A^c \cap B = \emptyset$ gilt $A^c \cup B \in \mathcal{D}$. Deswegen ist auch $A \cap B^c = (A^c \cup B)^c \in \mathcal{D}$. □

Der nächsten Satz gibt genügend und hinreichende Bedingungen, damit ein Dynkin System eine σ-Algebra ist.

Satz 1.13. *Sei $\mathcal{D}$ ein Dynkin System. Dann ist $\mathcal{D}$ genau dann eine σ-Algebra, wenn*

$$A, B \in \mathcal{D} \Rightarrow A \cap B \in \mathcal{D} \tag{4}$$

Beweis. Sei $\mathcal{D}$ ein Dynkin System mit (4). Wir müssen zeigen, $\mathcal{D}$ ist eine σ-Algebra. Sei dazu $(A_n)_{n\in\mathbb{N}}$ eine Folge in $\mathcal{D}$. Dann definieren wir $B_1 = A_1$, $B_2 = A_2 \backslash A_1 = A_2 \cap A_1^c$, und, für beliebigen $n \in \mathbb{N}$, $B_n = A_n \cap (A_1 \cup \dots \cup A_{n-1})^c$. Wegen (4) ist $(B_n)_{n\in\mathbb{N}}$ eine Folge von paarweise disjunkten Mengen aus $\mathcal{D}$. Damit ist auch

$$\bigcup_{n\in\mathbb{N}} A_n = \bigcup_{n\in\mathbb{N}} B_n \in \mathcal{D}\,.$$

□

Es gilt: ein beliebiges Durchschnitt von Dynkin Systeme ist wieder ein Dynkin System (der Beweis ist ähnlich wie bei σ-Algebren). Das impliziert, dass für eine beliebige Familie $\mathcal{F}$ von Mengen,

$$\delta(\mathcal{F}) = \bigcap \{\mathcal{D} : \ \mathcal{D} \text{ ein Dynkin System mit } \mathcal{F} \subset \mathcal{D}\}$$

der kleinste Dynkin System ist, der $\mathcal{F}$ enthält. Man sagt $\delta(\mathcal{F})$ ist der Dynkin System, der aus $\mathcal{F}$ erzeugt wird.

Satz 1.14. *Sei Ω eine Menge, und $\mathcal{F} \subset P(\Omega)$ eine Familie von Teilmengen von Ω, mit der Eigenschaft*

$$A, B \in \mathcal{F} \ \Rightarrow \ A \cap B \in \mathcal{F}$$

Dann ist $\delta(\mathcal{F}) = \sigma(\mathcal{F})$.

Beweis. Da $\sigma(\mathcal{F})$ ein Dynkin System ist, der $\mathcal{F}$ enthält, muss $\delta(\mathcal{F}) \subset \sigma(\mathcal{F})$ gelten. Zu zeigen bleibt: $\delta(\mathcal{F})$ ist eine σ-Algebra. Dazu genügt zu zeigen, dass

$$A, B \in \delta(\mathcal{F}) \ \Rightarrow \ A \cap B \in \delta(\mathcal{F})$$

Sei $A \in \delta(\mathcal{F})$ beliebig. Wir setzen

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

Wir behaupten, $\mathcal{D}_A$ ist ein Dynkin System. In der Tat ist $\Omega \in \mathcal{D}_A$ (weil $\Omega \cap A = A \in \delta(\mathcal{F})$). Ist weiter $E \in \mathcal{D}_A$ so gilt $E \cap A \in \delta(\mathcal{F})$. Dann ist auch

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

als Vereinigung von zwei disjunkten Mengen in $\delta(\mathcal{F})$. Damit ist also $E^c \cap A \in \delta(\mathcal{F})$ und $E^c \in \mathcal{D}_A$. Sei schlussendlich $(E_n)_{n\in\mathbb{N}}$ eine Folge von disjunkten Mengen in $\mathcal{D}_A$. Dann ist $E_n \cap A \in \delta(\mathcal{F})$ für alle $n \in \mathbb{N}$. Also ist $(E_n \cap A)$ eine Folge von disjunkten Mengen im Dynkin System $\delta(\mathcal{F})$; es folgt, dass

$$\bigcup_{n\in\mathbb{N}} E_n \cap A = \bigcup_{n\in\mathbb{N}} (E_n \cap A) \in \delta(\mathcal{F})$$

und damit, dass $\bigcup_{n\in\mathbb{N}} E_n \in \mathcal{D}_A$.

Ist nun $A \in \mathcal{F}$, so ist $\mathcal{D}_A$ ein Dynkin System und, nach Voraussetzung, gilt $\mathcal{F} \subset \mathcal{D}_A$ (weil $B \in \mathcal{F}$ impliziert $B \cap A \in \mathcal{F}$ und also $B \cap A \in \delta(\mathcal{F})$). Da $\delta(\mathcal{F})$ der kleinste Dynkin System ist, der $\mathcal{F}$ enthält, muss $\delta(\mathcal{F}) \subset \mathcal{D}_A$. Das impliziert, dass $B \cap A \in \delta(\mathcal{F})$ für alle $B \in \delta(\mathcal{F})$ und alle $A \in \mathcal{F}$. Sei nun $B \in \delta(\mathcal{F})$ beliebig. Dann $\mathcal{D}_B$ ist ein Dynkin System. Da $A \cap B \in \delta(\mathcal{F})$ für alle $A \in \mathcal{F}$, es folgt, dass $\mathcal{F} \subset \mathcal{D}_B$. Damit muss auch $\delta(\mathcal{F}) \subset \mathcal{D}_B$ gelten, für alle $B \in \delta(\mathcal{F})$. Das zeigt, dass $A \cap B \in \delta(\mathcal{F})$, für alle $A, B \in \delta(\mathcal{F})$. □

Neben dem Begriff von Dynkin System, spielt manchmal auch den Begriff von monotone Klasse eine wichtige Rolle, um zu entscheiden, ob einen gegebenen System eine σ-Algebra ist, oder nicht.

Definition 1.15. *Sei Ω eine Menge. Ein System $\mathcal{A}$ von Teilmengen von Ω heisst eine monotone Klasse, falls* *i) Ist $(A_n)_{n\in\mathbb{N}}$ eine Folge aus $\mathcal{A}$, mit $A_{n+1} \supset A_n$ für alle $n \in \mathbb{N}$, so ist*

$$\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$$

ii) Ist $(A_n)_{n\in\mathbb{N}}$ eine Folge aus $\mathcal{A}$, mit $A_{n+1} \subset A_n$ für alle $n \in \mathbb{N}$, so ist

$$\bigcap_{n\in\mathbb{N}} A_n \in \mathcal{A}$$

Jede σ-Algebra ist offenbar eine monotone Klasse. Es gelten weiter die folgenden Eigenschaften:

- Sei $\mathcal{A}$ eine monotone Klasse. Dann ist $\mathcal{A}$ genau dann eine σ-Algebra, wenn $\mathcal{A}$ eine Algebra ist.
- Beliebige Durchschnitte von monotonen Klassen sind wieder monotonen Klassen. Das impliziert: für eine beliebige Familie $\mathcal{F}$ von Teilmengen von Ω es existiert die kleinste monotone Klasse, die $\mathcal{F}$ enthält. Wir bezeichnen die von $\mathcal{F}$ erzeugten monotone Klasse mit $m(\mathcal{F})$.
- Ist $\mathcal{A}$ eine Algebra auf Ω, so gilt $\sigma(\mathcal{A}) = m(\mathcal{A})$.

Die Beweise von diesen Eigenschaften lassen wir als Übung (sie sind ähnlich zu den entsprechenden Beweise für Dynkin Systeme).

1.2 Masse

Sei Ω eine Menge, und $\mathcal{A}$ eine σ-Algebra auf Ω. Man nennt $(\Omega, \mathcal{A})$ ein messbarer Raum. Auf der σ-Algebra $\mathcal{A}$ betrachten wir nun Funktionen, die wir als Masse bezeichnen, die das Inhalt von Mengen in $\mathcal{A}$ messen sollen.

Definition 1.16. *Sei, wie oben $(\Omega, \mathcal{A})$ ein messbarer Raum. Eine Funktion $\mu : \mathcal{A} \to [0; \infty]$ heisst ein Mass auf $\mathcal{A}$ falls*

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

ii) σ-Additivität: ist $(A_n)_{n\in\mathbb{N}}$ eine Folge in $\mathcal{A}$ mit $A_n \cap A_m = \emptyset$, für alle $n \neq m$, so gilt

$$\mu\left(\cup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n)$$

Bemerkungen:

- Die Eigenschaft von σ-Additivität erklärt, warum wir in der Definition von σ-Algebren die Stabilität bezüglich abzählbaren Vereinigungen verlangt haben.
- Man könnte sich vorstellen Funktionen $\mu : \mathcal{A} \to [0; \infty]$ zu betrachten, die nur *endlich additiv* statt σ-additiv sind. In diesem Fall wurde man ii) durch die schwächere Bedingung ersetzen:

ii') Sind $A_1, \dots, A_k$ endlich viele, disjunkten Mengen aus $\mathcal{A}$, dann gilt

$$\mu\left(\cup_{j=1}^k A_j\right) = \sum_{j=1}^k \mu(A_j)$$

Eine Funktion $\mu : \mathcal{A} \to [0;\infty]$ mit den Eigenschaften i) und ii') nennt man ein *endlich additives Mass*. Jede σ-additive Mass ist endlich additiv. Nicht jede endlich additives Mass ist auch σ-additiv. Endlich additive Masse können auf beliebigen Algebren definiert werden (während σ-additive Masse nur auf σ-Algebren definiert werden können). Es war eine wichtige Bemerkung am Ende vom 19. Jahrhundert, dass σ-additive Masse viel besser als endlich additive Masse sind, um Integrale zu definieren und Analysis zu machen (man kann viel stärkere Theoremen beweisen, wenn man σ-Additivität annimmt). In dieser Vorlesung werden wir nur σ-additive Masse betrachten (und wir werden σ-additive Masse einfach als Masse bezeichnen, wie in der Definition oben).

Ist Ω eine Menge, $\mathcal{A}$ eine σ-Algebra auf Ω und μ ein Mass auf $\mathcal{A}$, so heisst $(\Omega, \mathcal{A}, \mu)$ ein Massraum.

Beispiele: hier ein Paar einfache Beispiele von Massen.

- Sei Ω eine beliebige Menge, und $\mathcal{A}$ eine σ-Algebra auf Ω. Für jede $A \in \mathcal{A}$ definieren wir $\mu(A) = n$, falls A genau n Elementen enthält, und $\mu(A) = \infty$, falls A unendlich viele Elementen enthält. Dann ist $\mu : \mathcal{A} \to [0;\infty]$ offenbar ein Mass. Man nennt μ das Zählmass auf dem messbare Raum $(\Omega, \mathcal{A})$.

- Sei Ω eine beliebige nicht leere Menge, und $\mathcal{A} = P(\Omega)$ (oder irgendeine andere σ-Algebra auf Ω). Sei $x \in \Omega$. Für $A \in \mathcal{A}$ definieren wir $\delta_x(A) = 1$ falls $x \in A$ und $\delta_x(A) = 0$ falls $x \notin A$. Dann definiert $\delta_x : \mathcal{A} \to [0;\infty]$ ein Mass auf $(\Omega, \mathcal{A})$, das als Punktmass an der Stelle x bezeichnet wird. Dass δ_x ein Mass ist, kann man wie folgt überprüfen. Sei $(A_n)_{n\in\mathbb{N}}$ eine Folge von disjunkten Mengen in $\mathcal{A}$. Wir unterscheiden zwei Fälle: 1) es gilt $x \notin A_n$ für alle $n \in \mathbb{N}$. Dann ist auch $x \notin \cup_{n\in\mathbb{N}} A_n$. Deswegen ist $\delta_x(A_n) = 0$ für alle n und auch $\delta_x(\cup_{n\in\mathbb{N}} A_n) = 0$, und

  $$\delta_x\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \delta_x(A_n)$$

  2) Es existiere $m \in \mathbb{N}$ mit $x \in A_m$. Dann ist $x \in \cup_{n\in\mathbb{N}} A_n$. Da die Mengen disjunkt sind, muss dann $x \notin A_n$ für alle $n \neq m$. Deswegen gilt

  $$\sum_{n\in\mathbb{N}} \delta_x(A_n) = \delta_x(A_m) = 1 = \delta_x\left(\bigcup_{n\in\mathbb{N}} A_n\right)$$

- Sei Ω eine nicht leere Menge, und $\mathcal{A}$ eine σ-Algebra auf Ω. Für $A \in \mathcal{A}$ setzen wir $\mu(A) = \infty$ falls $A \neq \emptyset$ und $\mu(A) = 0$ falls $A = \emptyset$. Dann ist μ ein Mass.

- Wir werden auf der Borel σ-Algebra $\mathcal{B}(\mathbb{R})$ ein Mass λ definieren, mit der Eigenschaft, dass $\lambda([a;b]) = b - a$ für alle $a \leq b$. Man nennt λ das Lebesgue

Mass auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Allgemeiner werden wir für alle $n \geq 1$ ein Mass $\lambda^{(n)}$ auf der n-dimensionale Borel σ-Algebra $\mathcal{B}(\mathbb{R}^n)$ konstruieren, mit der Eigenschaft, dass $\lambda^{(n)}([a_1, b_1] \times \cdots \times [a_n, b_n]) = \prod_{j=1}^n (b_i - a_i)$. Man nennt $\lambda^{(n)}$ das Lebesgue Mass auf $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$. Das Lebesgue Mass $\lambda^{(n)}$ erfüllt alle Bedingungen, die wir in der Einführung (siehe Seite 4) für ein Volumen verlangt haben (Monotonie, euklidische Invarianz, Normierung, σ-Additivität). In Gegensatz zu Satz 1.1 gibt es hier kein Widerspruch, weil $\lambda^{(n)}$ nicht auf alle Teilmengen von $\mathbb{R}^n$ definiert ist (die Menge M_0 im Beweis von Satz 1.1 ist nicht in der σ-Algebra $\mathcal{B}(\mathbb{R}^n)$ enthalten).

Wir diskutieren einige Eigenschaften von Massen, die aus der allgemeine Definition folgen.

Satz 1.17. *Sei $(\Omega, \mathcal{A}, \mu)$ ein Massraum. Seien $A, B \in \mathcal{A}$ mit $A \subset B$. Dann gilt $\mu(A) \leq \mu(B)$. D.h. Masse sind immer monoton. Gilt weiter $\mu(A) < \infty$, so finden wir $\mu(B \backslash A) = \mu(B) - \mu(A)$.*

Beweis. Es gilt $B = A \cup (B \backslash A)$. Da $A \cap (B \backslash A) = \emptyset$, folgt aus der σ-Additivität, dass

$$\mu(B) = \mu(A) + \mu(B \backslash A) \geq \mu(A)$$

weil $\mu(B \backslash A) \geq 0$. Ist weiter $\mu(A) < \infty$, so können wir $\mu(A)$ links subtrahieren, und wir finden $\mu(B \backslash A) = \mu(B) - \mu(A)$. □

Die σ-Additivität für das Mass von disjunkten Mengen impliziert auch die Subadditivität für Vereinigung von beliebige Teilmengen.

Satz 1.18. *Sei $(\Omega, \mathcal{A}, \mu)$ ein Massraum, $(A_n)_{n \in \mathbb{N}}$ eine Folge in $\mathcal{A}$. Dann gilt*

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

Beweis. Wir definieren eine neue Folge von Mengen in $\mathcal{A}$. Wir setzen $B_1 = A_1$, $B_2 = A_2 \backslash A_1 = A_2 \cap A_1^c$ und, für beliebige $n \in \mathbb{N}$,

$$B_n = A_n \cap \left(\cup_{j=1}^{n-1} A_j\right)^c = A_n \cap A_1^c \cap \cdots \cap A_n^c$$

Da $\mathcal{A}$ stabil bzg. Durchschnitten ist (hier braucht man sogar nur Stabilität bzg. endliche Durchschnitten), ist $B_n \in \mathcal{A}$ für alle $n \in \mathbb{N}$. Die B_n sind nach Konstruktion disjunkt. Für alle $n \in \mathbb{N}$ gilt $B_n \subset A_n$ und, wegen der Monotonie, $\mu(B_n) \leq \mu(A_n)$. Da aber

$$\bigcup_{n \in \mathbb{N}} B_n = \bigcup_{n \in \mathbb{N}} A_n$$

finden wir

$$\mu\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \mu\left(\bigcup_{n \in \mathbb{N}} B_n\right) = \sum_{n \in \mathbb{N}} \mu(B_n) \leq \sum_{n \in \mathbb{N}} \mu(A_n)$$

□

Für das Mass einer Vereinigung von endlich viele (nicht unbedingt disjunkten) Mengen, es ist auch einfach eine explizite Formel zu finden. Sei $(\Omega, \mathcal{A}, \mu)$ ein Massraum, $A, B \in \mathcal{A}$. Dann können wir

$$A \cup B = (A \backslash B) \cup (B \backslash A) \cup (A \cap B)$$

als Vereinigung von drei disjunkten Mengen schreiben. Da

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

es gilt

$$\begin{aligned} \mu(A \backslash B) &= \mu(A) - \mu(A \cap B), \\ \mu(B \backslash A) &= \mu(B) - \mu(A \cap B) \end{aligned}$$

und damit

$$\mu(A \cup B) = \mu(A \backslash B) + \mu(B \backslash A) + \mu(A \cap B) = \mu(A) + \mu(B) - \mu(A \cap B)$$

Man kann analoge Formel für Vereinigung von drei Mengen und, allgemeiner, für die Vereinigung von n Mengen herleiten.

Es ist manchmal wichtig das Mass der Vereinigung oder des Durchschnittes einer monotone Folge von Mengen zu bestimmen.

Satz 1.19. *Sei $(\Omega, \mathcal{A}, \mu)$ ein Massraum.*

a) Ist $(A_n)_{n\geq 1}$ eine Folge in $\mathcal{A}$ mit $A_{n+1} \supset A_n$ für alle $n \in \mathbb{N}$ (eine wachsende Folge), so gilt

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

b) Ist $(A_n)_{n\geq 1}$ eine Folge in $\mathcal{A}$ mit $A_{n+1} \subset A_n$ für alle $n \in \mathbb{N}$ (eine fallende Folge), und es gelte $\mu(A_{n_0}) < \infty$ für irgendein $n_0 \in \mathbb{N}$, so gilt

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

Beweis. a) Sei (A_n) eine Folge in $\mathcal{A}$ mit $A_{n+1} \supset A_n$ für alle $n \geq 1$. Wir setzten $B_1 = A_1$ und $B_i = A_i \backslash A_{i-1} = A_i \cap A_{i-1}^c$ für alle $i \geq 2$. Dann ist $(B_n)_{n\geq 1}$ eine Folge in $\mathcal{A}$ von disjunkten Teilmengen von Ω, mit $\cup_{n\geq 1} B_n = \cup_{n\in\mathbb{N}} A_n$. Deswegen gilt

$$\begin{aligned} \mu\left(\bigcup_{n\geq 1} A_n\right) &= \mu\left(\bigcup_{n\geq 1} B_n\right) = \sum_{n=1}^{\infty} \mu(B_n) \\ &= \lim_{k\to\infty} \sum_{n=1}^{k} \mu(B_n) = \lim_{k\to\infty} \mu\left(\bigcup_{n=1}^{k} B_n\right) = \lim_{k\to\infty} \mu(A_k) \end{aligned}$$

b) Sei $(A_n)_{n\in\mathbb{N}}$ eine Folge in $\mathcal{A}$ mit $A_{n+1} \subset A_n$ und $\mu(A_{n_0}) < \infty$. O.B.d.A. können wir annehmen, dass $n_0 = 1$. Für $n \geq 1$ setzen wir

$$C_n = A_1 \backslash A_n = A_1 \cap A_n^c$$

Dann ist $(C_n)_{n\in\mathbb{N}}$ eine Folge in $\mathcal{A}$ mit $C_{n+1} \supset C_n$ für alle n (d.h. C_n ist eine wachsende Folge), und

$$\bigcup_{n\geq 1} C_n = A_1 \backslash \bigcap_{n\geq 1} A_n$$

Aus Teil a) finden wir

$$\mu\left(A_1 \backslash \cap_{n\geq 1} A_n\right) = \mu\left(\bigcup_{n\geq 1} C_n\right) = \lim_{n\to\infty} \mu(C_n) = \lim_{n\to\infty} \mu(A_1 \backslash A_n)$$

Aus Satz 1.17, und weil $\mu(A_1) < \infty$, erhalten wir die Behauptung. $\square$

In der Tat können die Eigenscahften, die wir im letzten Satz bewiesen haben, als Charakterizierung von σ-additive Masse benutzt werden.

Proposition 1.20. *Sei $(\Omega, \mathcal{A})$ ein messbarer Raum, und μ ein endliches-additives Mass auf $\mathcal{A}$. Dann ist μ ein Mass falls eine der folgenden Bedingungen erfüllt ist.*

a) $\lim_{n\to\infty} \mu(A_n) = \mu(\cup_{n\geq 1} A_n)$ *für alle wachsenden Folge* $(A_n)_{n\geq 1}$ *in* $\mathcal{A}$.

b) $\lim_{n\to\infty} \mu(A_n) = 0$ *für alle fallende Folge* $(A_n)_{n\geq 1}$ *in* $\mathcal{A}$ *mit* $\cap_{n\geq 1} A_n = \emptyset$.

Beweis. Nehmen wir zunächst an, a) sei erfüllt. Wir möchten zeigen, dass μ σ-additiv ist. Sei $(B_n)_{n\geq 1}$ eine Folge in $\mathcal{A}$ von disjunkten Mengen. Wir setzten $A_n = \cup_{j=1}^n B_j$. Dann gilt $\cup_{j=1}^\infty B_j = \cup_{n\geq 1} A_n$. Weiter, A_n ist offenbar eine wachsende Folge in $\mathcal{A}$. Die endliche Additivität impliziert, dass

$$\mu(A_n) = \sum_{j=1}^n \mu(B_j)$$

für alle $n \in \mathbb{N}$. Anderseits, a) impliziert, dass

$$\mu\left(\bigcup_{j=1}^\infty B_j\right) = \mu\left(\bigcup_{n\geq 1} A_n\right) = \lim_{n\to\infty} \mu(A_n) = \lim_{n\to\infty} \sum_{j=1}^n \mu(B_j) = \sum_{j=1}^\infty \mu(B_j)$$

Nehmen wir nun an, dass die Bedingung b) erfüllt ist. Sei wie oben $(B_j)_{j\geq 1}$ eine Folge von disjunkten Mengen in $\mathcal{A}$. Wir setzen $A_k = \cup_{j=k+1}^\infty B_j$. Dann ist $(A_k)_{k\in\mathbb{N}}$ eine fallende Folge in $\mathcal{A}$, mit $\cup_{j=1}^\infty B_j = \cup_{n=1}^\infty A_n$. Aus der endliche Additivität gilt

$$\mu\left(\bigcup_{j=1}^\infty B_j\right) = \sum_{j=1}^k \mu(B_j) + \mu(A_k)$$

für alle $k \in \mathbb{N}$. Insbesondere gilt

$$\mu\left(\bigcup_{j=1}^{\infty} B_j\right) = \sum_{j=1}^{\infty} \mu(B_j) + \lim_{k\to\infty} \mu(A_k)$$

Die σ-Additivität folgt aus der Annahme b), die $\lim_{k\to\infty} \mu(A_k) = 0$ impliziert. Das folgt, weil $\cap_{k=1}^{\infty} A_k = \emptyset$, da die Mengen B_j disjunkt sind. $\square$

Für gewisse Resultate ist es nützlich eine weitere Eigenschaft von Massen zu verlangen, nämlich die σ-Endlichkeit, die wir in der nächsten Definition einführen.

Definition 1.21. *Sei $(\Omega, \mathcal{A}, \mu)$ ein Massraum. Das Mass μ heisst endlich, falls $\mu(\Omega) < \infty$ (in diesem Fall ist $\mu(A) < \infty$ für alle $A \subset \Omega$). Das Mass μ heisst σ-endlich, falls es eine Folge $(B_n)_{n\in\mathbb{N}}$ in $\mathcal{A}$ existiert, mit $\mu(B_n) < \infty$ für alle $n \in \mathbb{N}$ und mit*

$$\bigcup_{n\in\mathbb{N}} B_n = \Omega$$

Endliche Masse spielen eine wichtige Rolle in der Wahrscheinlichkeitstheorie, weil dort $\mu(\Omega) = 1$. Das Lebesgue Mass auf $\mathcal{B}(\mathbb{R})$ ist offenbar nicht endlich, aber σ-endlich (man nehme, zum Beispiel, die Folge $B_n = [n; n+1] \cup [-n-1; -n]$). Dasselbe gilt für das Lebesgue Mass auf $\mathcal{B}(\mathbb{R}^n)$.

Bemerkung: Sei $(\Omega, \mathcal{A}, \mu)$ ein Massraum, und μ ein σ-endliches Mass. Dann existiert auch eine Folge $(B_n)_{n\in\mathbb{N}}$ von disjunkten Mengen in $\mathcal{A}$ mit $\cup_{n\in\mathbb{N}} B_n = \Omega$ und $\mu(B_n) < \infty$ für alle $n \in \mathbb{N}$. Das folgt einfach, indem man $B_n = A_n \backslash (\cup_{j=1}^{n} A_j)$.

1.3 Äusseres Mass

Um interessante Masse zu konstruieren (insbesondere das Lebesgue Mass) ist der Begriff von äussere Mass sehr wichtig.

Definition 1.22. *Sei Ω eine Menge. Ein äusseres Mass auf Ω ist eine Funktion $\mu^* : P(\Omega) \to [0; \infty]$ mit den folgenden Eigenschaften.*

i) $\mu^*(\emptyset) = 0$.

ii) Monotonie: Ist $A \subset B \subset \Omega$, so gilt $\mu^*(A) \leq \mu^*(B)$.

iii) Abzählbare Subadditivität: Ist $(A_n)_{n\in\mathbb{N}}$ eine Folge in $P(\Omega)$, so gilt

$$\mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu^*(A_n)$$

Offenbar ist nicht jede äussere Mass ein Mass (weil die Subadditivität ist schwächer als die σ-Additivität, wie wir in Satz 1.18 bewiesen haben). Anderseits ist auch nicht jedes Mass ein äusseres Mass, weil äussere Masse immer auf der ganzen Potenzmenge $P(\Omega)$ definiert sind. Es gilt: jede Mass, das auf $P(\Omega)$ definiert ist, ist auch ein äusseres Mass.

Die Bedeutung von äussere Masse wird bald klar; wir werden zeigen, dass zu jede äussere Mass μ^* wir eine σ-Algebra $\mathcal{M}_{\mu^*} \subset P(\Omega)$ finden können, so, dass die Einschränkung μ von μ^* auf $\mathcal{M}_{\mu^*}$ ein Mass ist. Bevor wir zur Konstruktion der σ-Algebra $\mathcal{M}_{\mu^*}$ kommen, diskutieren wir ein Paar einfache Beispiele von äussere Masse (das wichtigste Beispiel von äusserem Mass ist das Lebesgue äussere Mass, das wir im nächsten Kapitel untersuchen werden).

Beispiele:

- Sei Ω eine beliebige Menge. Für $A \subset \Omega$ setzen wir $\mu^*(A) = 0$ if $A = \emptyset$ und $\mu^*(A) = 1$ sonst. μ^* ist nicht ein Mass, aber es ist offenbar ein äusseres Mass.

- Sei Ω eine beliebige Menge. Für $A \subset \Omega$ setzen wir $\mu^*(A) = 0$ falls A ist abzählbar und $\mu^*(A) = 1$, falls A is überabzählbar. Dann ist μ^* ein äusseres Mass. Die Monotonie ist klar (weil eine überabzählbare Menge nicht in einer abzählbare Menge enthalten sein kann). Die abzählbare Subadditivität kann wie folgt gezeigt werden. Sei $(A_n)_{n\in\mathbb{N}}$ eine Folge in $P(\Omega)$. Ist A_n abzählbar, für alle $n \in \mathbb{N}$, so ist auch $\cup_{n\in\mathbb{N}} A_n$, als abzählbare Vereinigung von abzählbare Mengen abzählbar, und iii) ist offenbar erfüllt (links und rechte Seite sind Null). Gibt es dagegen $m \in \mathbb{N}$ mit A_m überabzählbar, so ist $\cup_{n\in\mathbb{N}} A_n$ überabzählbar. Dann ist
$$\mu^*\left(\bigcup_{n\in\mathbb{N}} A_n\right) = 1 = \mu^*(A_m) \leq \sum_{n=1}^{\infty} \mu^*(A_n)\,.$$

- Sei Ω eine unendliche Menge. Definiere $\mu^*(A) = 0$ falls $A \subset \Omega$ endlich ist und $\mu^*(A) = 1$ falls $A \subset \Omega$ unendlich ist. Dann ist μ^* kein äusseres Mass, weil μ^* nicht subadditiv ist. Ist nämlich $(A_n)_{n\in\mathbb{N}}$ eine Folge von disjunkten, endlichen und nicht-leeren Teilmengen von Ω, so ist $\mu^*(A_n) = 0$ für alle $n \in \mathbb{N}$ und $\sum_n \mu^*(A_n) = 0$.Anderseits, $\cup_{n\in\mathbb{N}} A_n$ ist eine unendliche Teilmenge, und damit $\mu^*(\cup_{n\in\mathbb{N}} A_n) = 1$.

Jetzt möchten wir lernen, wie man aus einem äusseren Mass ein richtiges σ-additives Mass bekommen kann.

Definition 1.23. *Sei Ω ein Menge, und μ^* ein äusseres Mass auf Ω. Eine Teilmenge $B \subset \Omega$ heisst μ^*-messbar falls*
$$\mu^*(A) = \mu^*(A \cap B) + \mu^*(A \cap B^c)$$
für alle $A \subset \Omega$ gilt. In Wörter: $B \subset \Omega$ ist μ^*-messbar, falls jede Teilmenge von Ω durch B in zwei Teilen zerlegt wird, dessen μ^*-Masse "richtig" summiert werden können.

Bemerkung: aus der Subadditivität von μ^* folgt, dass
$$\mu^*(A) \leq \mu^*(A \cap B) + \mu^*(A \cap B^c)$$
immer gilt. D.h., $B \subset \Omega$ ist genau dann μ^*-messbar, wenn
$$\mu^*(A) \geq \mu^*(A \cap B) + \mu^*(A \cap B^c) \tag{5}$$
für alle $A \subset \Omega$ gilt. Es genügt sogar (5) nur für $A \subset \Omega$ mit $\mu^*(A) < \infty$ zu überprüfen (weil sonst es gibt nichts zu beweisen).

Wir untersuchen nun die Eigenschaften der μ^*-messbare Teilmengen von Ω.

Satz 1.24. *Sei Ω eine Menge und μ^* ein äusseres Mass auf Ω. Ist $B \subset \Omega$ mit $\mu^*(B) = 0$ oder $\mu^*(B^c) = 0$, dann ist B μ^*-messbar.*

Beweis. Nehmen wir an, dass $\mu^*(B) = 0$. Sei $A \subset \Omega$ beliebig. Dann gilt $A \cap B \subset B$ und deswegen $\mu^*(A \cap B) = 0$. Anderseits, da $A \cap B^c \subset A$ finden wir $\mu^*(A) \geq \mu^*(A \cap B^c)$. Das zeigt (5), und (weil A beliebig ist), dass B μ^*-messbar. Der Beweis für den Fall $\mu^*(B^c) = 0$ ist ähnlich. $\square$

Der letzte Satz zeigt, dass $\emptyset, \Omega$ immer μ^*-messbar sind. Im nächsten Theorem zeigen wir, dass die Familie von allen μ^*-messbaren Mengen eine σ-Algebra ist.

Theorem 1.25. *Sei Ω eine Menge, und μ^* ein äusseres Mass auf Ω. Dann ist*

$$\mathcal{M}_{\mu^*} := \{B \subset \Omega : B \textit{ ist } \mu^*\textit{-messbar}\}$$

eine σ-Algebra auf Ω. Die Einschränkung von μ^* auf $\mathcal{M}_{\mu^*}$ ist ein Mass auf $\mathcal{M}_{\mu^*}$.

Beweis. Wir zerlegen den Beweis in drei Schritte.

Schritt 1. $\mathcal{M}_{\mu^*}$ ist eine Algebra.

Wir haben schon in Satz 1.24 bewiesen, dass $\emptyset, \Omega \in \mathcal{M}_{\mu^*}$. Ist $B \in \mathcal{M}_{\mu^*}$, so ist aus (5) auch klar, dass auch $B^c \in \mathcal{M}_{\mu^*}$. Seien nun $B_1, B_2 \in \mathcal{M}_{\mu^*}$. Wir möchten zeigen, dass $B_1 \cup B_2 \in \mathcal{M}_{\mu^*}$. Dazu bemerken wir, dass für ein beliebiges $A \subset \Omega$, die μ^*-Messbarkeit von B_1 impliziert, dass

$$\begin{aligned}
\mu^*(A \cap (B_1 \cup B_2)) &= \mu^*(A \cap (B_1 \cup B_2) \cap B_1) + \mu^*(A \cap (B_1 \cup B_2) \cap B_1^c) \\
&= \mu^*(A \cap B_1) + \mu^*(A \cap B_1^c \cap B_2)
\end{aligned}$$

Zusammen mit $(B_1 \cup B_2)^c = B_1^c \cap B_2^c$, zeigt diese Identität, dass

$$\begin{aligned}
\mu^*(A \cap (B_1 \cup B_2)) &+ \mu^*(A \cap (B_1 \cup B_2)^c) \\
&= \mu^*(A \cap B_1) + \mu^*(A \cap B_1^c \cap B_2) + \mu^*(A \cap B_1^c \cap B_2^c) \\
&= \mu^*(A \cap B_1) + \mu^*(A \cap B_1^c) = \mu^*(A)
\end{aligned}$$

wobei wir zunächst die Messbarkeit von B_2 und nachher die Messbarkeit von B_1 benutzt haben. Das zeigt, dass $B_1 \cup B_2$ ist μ^*-messbar, und also, dass $B_1 \cup B_2 \in \mathcal{M}_{\mu^*}$. Damit ist $\mathcal{M}_{\mu^*}$ eine Algebra.

Schritt 2. $\mathcal{M}_{\mu^*}$ ist eine σ-Algebra.

Wir müssen zeigen, dass $\mathcal{M}_{\mu^*}$ stabil ist, bezüglich abzählbare Vereinigungen. Sei zunächst $(B_n)_{n\geq 1}$ eine Folge von disjunkten μ^*-messbare Teilmengen von Ω. Wir behaupten, dass

$$\mu^*(A) = \sum_{i=1}^{n} \mu^*(A \cap B_i) + \mu^* \left(A \cap \bigcap_{i=1}^{n} B_i^c \right) \tag{6}$$

für alle $n \geq 1$ und alle $A \subset \Omega$. Wir zeigen diese Behauptung durch Induktion über n. Für $n = 1$, sie folgt einfach aus der Messbarkeit von B_1. Nehmen wir nun an die Behauptung sei für $n = k$ korrekt; wir zeigen sie für $n = k+1$. Da B_{k+1} messbar ist, finden wir

$$\begin{aligned}\mu^*\left(A \cap \bigcap_{i=1}^{k} B_i^c\right) &= \mu^*\left(A \cap \bigcap_{i=1}^{k} B_i^c \cap B_{k+1}\right) + \mu^*\left(A \cap \bigcap_{i=1}^{k+1} B_i^c\right) \\ &= \mu^*(A \cap B_{k+1}) + \mu^*\left(A \cap \bigcap_{i=1}^{k+1} B_i^c\right)\end{aligned}$$

weil die B_j disjunkt sind. Aus der Induktion Annahme bekommen wir

$$\mu^*(A) = \sum_{i=1}^{k+1} \mu^*(A \cap B_i) + \mu^*(A \cap \bigcap_{i=1}^{k+1} B_i^c)$$

Damit ist (6) für alle $n \in \mathbb{N}$ bewiesen.

Aus (6) bekommen wir nun die Schranke

$$\mu^*(A) \geq \sum_{i=1}^{n} \mu^*(A \cap B_i) + \mu^*\left(A \cap \bigcap_{i=1}^{\infty} B_i^c\right) \tag{7}$$

weil

$$A \cap \bigcap_{i=1}^{\infty} B_i^c \subset A \cap \bigcap_{i=1}^{n} B_i^c .$$

Da (7) für alle $n \in \mathbb{N}$ gilt, finden wir auch

$$\mu^*(A) \geq \sum_{i=1}^{\infty} \mu^*(A \cap B_i) + \mu^*\left(A \cap \bigcap_{i=1}^{\infty} B_i^c\right) \tag{8}$$

Die abzählbare Subadditivität von μ^* impliziert, dass

$$\mu^*(A) \geq \mu^*\left(\bigcup_{i=1}^{\infty}(A \cap B_i)\right) + \mu^*\left(A \cap \bigcap_{i=1}^{\infty} B_i^c\right) = \mu^*\left(A \cap \bigcup_{i=1}^{\infty} B_i\right) + \mu^*\left(A \cap \left(\bigcup_{i=1}^{\infty} B_i\right)^c\right)$$

für alle $A \subset \Omega$. Das zeigt, dass $\cup_{i=1}^{\infty} B_i$ μ^*-messbar ist.

Wir haben damit gezeigt, dass $\mathcal{M}_{\mu^*}$ stabil ist, bezüglich abzählbaren Vereinigungen von disjunkten Teilmengen. Das zeigt, dass $\mathcal{M}_{\mu^*}$ ein Dynkin System ist. Aus Schritt a) und aus Satz 1.13, $\mathcal{M}_{\mu^*}$ ist auch eine σ-Algebra.

Schritt 3. Die Einschränkung von μ^* auf $\mathcal{M}_{\mu^*}$ ist ein Mass.

Wir müssen die σ-Additivität der Einschränkung von μ^* auf $\mathcal{M}_{\mu^*}$ beweisen. Sei $(B_n)_{n \in \mathbb{N}}$ eine Folge in $\mathcal{M}_{\mu^*}$ von disjunkten Teilmengen. Aus (8), mit $A = \cup_{n=1}^{\infty} B_n$, finden wir

$$\mu^*\left(\bigcup_{i=1}^{\infty} B_i\right) \geq \sum_{i=1}^{\infty} \mu^*(B_i)$$

Die Subadditivität von μ^* impliziert also, dass

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

Das zeigt, dass μ^* ist ein Mass, wann es auf $\mathcal{M}_{\mu^*}$ eingeschränkt ist. $\square$

1.4 Konstruktion des Lebesgue'sche Mass auf $\mathbb{R}^n$

Für $A \subset \mathbb{R}$, sei $\mathcal{C}_A$ die Menge von Folgen $\{(a_i; b_i)\}_{i\in\mathbb{N}}$ von offenen und beschränkten Intervallen, mit

$$A \subset \bigcup_{i\in\mathbb{N}} (a_i; b_i)$$

Wir definieren das Lebesgue äusseres Mass von A durch

$$\lambda^*(A) := \inf \left\{ \sum_{i=0}^{\infty} (b_i - a_i) : \{(a_i; b_i)\} \in \mathcal{C}_A \right\} \tag{9}$$

$\lambda^*(A)$ ist wohldefiniert, weil die Familie $\mathcal{C}_A$ immer nicht leer ist (das Infimum einer Menge, die nur $+\infty$ enthält wird als $+\infty$ definiert).

Satz 1.26. *Das Lebesgue äussere Mass $\lambda^* : P(\mathbb{R}) \to [0; \infty]$, definiert in (9) ist ein äusseres Mass. Das Lebesgue äussere Mass eines Intervall (offen, geschlossen oder halboffen, beschränkt oder unbeschränkt) ist die Länge des Intervalls.*

Beweis. Aus (9) folgt sofort, dass $\lambda^*(\emptyset) = 0$ (man kann die Intervalle beliebig klein wählen). Die Monotonie von λ^* folgt aus der Bemerkung, dass

$$A \subset B \;\Rightarrow\; \mathcal{C}_B \subset \mathcal{C}_A\,.$$

(Weil das Infimum einer grössere Menge immer kleiner als das Infimum einer kleineren Menge ist). Wir zeigen nun die Subadditivität von λ^*. Sei dazu $\{A_n\}_{n\in\mathbb{N}}$ eine Folge von Teilmengen von $\mathbb{R}$. Wir möchten zeigen, dass

$$\lambda^* \left(\bigcup_{n\in\mathbb{N}} A_n \right) \leq \sum_{n\in\mathbb{N}} \lambda^*(A_n)$$

Wir können annehmen, dass

$$\sum_{n\in\mathbb{N}} \lambda^*(A_n) < \infty$$

weil sonst die Aussage trivial ist. Sei $\varepsilon > 0$ fest. Für $n \in \mathbb{N}$ wir finden, aus Definition von $\lambda^*(A_n)$ eine Folge $\{(a_{n,i}; b_{n,i})\}_{i\in\mathbb{N}}$ von offenen, beschränkten Intervalle mit $A_n \subset \bigcup_{i\in\mathbb{N}} (a_{n,i}; b_{n,i})$ und

$$\lambda^*(A_n) \geq \sum_{i=1}^{\infty} (b_{n,i} - a_{n,i}) - \frac{\varepsilon}{2^n}$$

Wir nehmen die Vereinigung von alle Intervalle $(a_{n,i}; b_{n,i})$ über alle $i \in \mathbb{N}$ und $n \in \mathbb{N}$, und bezeichnen die resultierenden Folge mit $\{(a_i; b_i)\}_{i\in\mathbb{N}}$. Dann gilt

$$\bigcup_{n\in\mathbb{N}} A_n \subset \bigcup_{i\in\mathbb{N}} (a_i; b_i)$$

und

$$\sum_{i\in\mathbb{N}} (b_i - a_i) = \sum_{n\in\mathbb{N}} \sum_{i\in\mathbb{N}} (b_{n,i} - a_{n,i}) \leq \sum_{n\in\mathbb{N}} \left[\lambda^*(A_n) + \frac{\varepsilon}{2^n} \right] \leq \sum_{n\in\mathbb{N}} \lambda^*(A_n) + \varepsilon$$

Das impliziert, dass

$$\lambda^*\left(\bigcup_{n\in\mathbb{N}} A_n\right) \leq \sum_{n\in\mathbb{N}} \lambda^*(A_n) + \varepsilon$$

Da $\varepsilon > 0$ beliebig ist, finden wir

$$\lambda^*\left(\bigcup_{n\in\mathbb{N}} A_n\right) \leq \sum_{n\in\mathbb{N}} \lambda^*(A_n)\,.$$

Sei nun $[a;b]$ ein geschlossenes beschränktes Intervall in $\mathbb{R}$. Wir möchten zeigen, dass $\lambda^*([a;b]) = (b-a)$. Offenbar folgt aus der Definition von λ^*, dass

$$\lambda^*([a;b]) \leq (b-a)$$

Wir müssen noch zeigen, dass $\lambda^*([a;b]) \geq (b-a)$. Sei also $\{(a_i;b_i)\}$ eine Folge von offenen und beschränkten Intervalle in $\mathbb{R}$, mit

$$[a;b] \subset \bigcup_{i\in\mathbb{N}} (a_i;b_i).$$

Da $[a;b]$ kompakt ist, es existiert $n \in \mathbb{N}$, mit

$$[a;b] \subset \bigcup_{i=1}^{n} (a_i;b_i)$$

Man kann leicht überprüfen (mit Induktion über $n \in \mathbb{N}$), dass

$$b-a \leq \sum_{i=1}^{n} (b_i - a_i)$$

Das impliziert, dass $b-a \leq \sum_{i=1}^{\infty}(b_i-a_i)$. Da die Folge $\{(a_i;b_i)\}_{i\in\mathbb{N}}$ beliebig ist, es folgt, dass $b-a \leq \lambda^*([a;b])$. Das zeigt, dass $\lambda^*([a;b]) = b-a$, für alle $a<b$. Für ein beliebiges offenes beschränktes Intervall $(a;b)$ bemerken wir, dass

$$\left[a+\frac{1}{n};b-\frac{1}{n}\right] \subset (a;b) \subset [a;b]$$

für alle $n \in \mathbb{N}$. Die Monotonie von λ^* zeigt, dass

$$b-a-\frac{2}{n} = \lambda^*\left(\left[a+\frac{1}{n};b-\frac{1}{n}\right]\right) \leq \lambda^*((a;b)) \leq \lambda^*([a;b]) = b-a$$

Da $n \in \mathbb{N}$ beliebig ist, kriegen wir $\lambda^*((a;b)) = b-a$. Analog kann man den Fall von halboffenen beschränkten Intervalle betrachten. Mit der Monotonie es ist auch einfach zu zeigen, dass $\lambda^*(I) = \infty$ für jede unbeschränkte Intervall $I \subset \mathbb{R}$. □

Definition 1.27. *Wir bezeichnen mit $\mathcal{M}_{\lambda^*}$ die Menge der λ^*-messbare Teilmengen von $\mathbb{R}$ und wir definieren das Lebesgue Mass λ auf $\mathbb{R}$ als die Einschränkung von λ^* auf $\mathcal{M}_{\lambda^*}$, d.h. $\lambda = \lambda^*|_{\mathcal{M}_{\lambda^*}}$. Aus Theorem 1.25 wissen wir, dass $\mathcal{M}_{\lambda^*}$ eine σ-Algebra ist, und, dass λ ein Mass auf $(\mathbb{R}, \mathcal{M}_{\lambda^*})$ ist.*

Damit diese Definition wirklich nützlich sein kann, müssen wir sicher stellen, dass $\mathcal{M}_{\lambda^*}$ gross genug ist (wäre $\mathcal{M}_{\lambda^*} = \{\emptyset, \mathbb{R}\}$, so wäre das Lebesgue Mass nicht so interessant für die Analysis). Das machen wir im nächsten Satz.

Satz 1.28. *Jede Borel Teilmenge von* $\mathbb{R}$ *ist* λ^**-messbar, d.h.* $\mathcal{B}(\mathbb{R}) \subset \mathcal{M}_{\lambda^*}$.

Beweis. Es genügt zu zeigen, dass

$$(-\infty; b] \in \mathcal{M}_{\lambda^*} \tag{10}$$

für alle $b \in \mathbb{R}$. In der Tat, aus (10) folgt, dass $\mathcal{M}_{\lambda^*}$ eine σ-Algebra ist, die alle Mengen der Form $(-\infty; b]$ enthält. Wir haben aber in Satz 1.6 bewiesen, dass $\mathcal{B}(\mathbb{R})$ die kleinste σ-Algebra mit dieser Eigenschaft ist. Das würde also zeigen, dass $\mathcal{B}(\mathbb{R}) \subset \mathcal{M}_{\lambda^*}$. Sei also $B = (-\infty; b]$ für ein $b \in \mathbb{R}$, und $A \subset \mathbb{R}$ mit $\lambda^*(A) < \infty$. Wir möchten zeigen, dass

$$\lambda^*(A) \geq \lambda^*(A \cap B) + \lambda^*(A \cap B^c) \tag{11}$$

(die andere Ungleichung folgt aus der Subadditivität von λ^*). Sei $\varepsilon > 0$ festgewählt. Aus Definition von $\lambda^*(A)$ finden wir eine Folge $\{(a_n; b_n)\}_{n\in\mathbb{N}}$ von offenen beschränkten Intervallen, mit $A \subset \bigcup_{n\in\mathbb{N}} (a_n; b_n)$ und

$$\lambda^*(A) \geq \sum_{n\in\mathbb{N}} (b_n - a_n) - \varepsilon \,.$$

Für alle $n \in \mathbb{N}$ sind $(a_n; b_n) \cap B$ und $(a_n; b_n) \cap B^c$ disjunkten Intervalle in $\mathbb{R}$ (möglicherweise leer), dessen Vereinigung $(a_n; b_n)$ ist. Das impliziert, aus Satz 1.26, dass

$$b_n - a_n = \lambda^*((a_n; b_n)) = \lambda^*((a_n; b_n) \cap B) + \lambda^*((a_n; b_n) \cap B^c)$$

Da

$$A \cap B \subset \bigcup_{n\in\mathbb{N}} (a_n; b_n) \cap B \quad \text{und} \quad A \cap B^c \subset \bigcup_{n\in\mathbb{N}} (a_n; b_n) \cap B^c$$

erhalten wir aus der Subadditivität von λ^*, dass

$$\begin{aligned} \lambda^*(A \cap B) + \lambda^*(A \cap B^c) &\leq \lambda^* \left(\bigcup_{n\in\mathbb{N}} (a_n; b_n) \cap B \right) + \lambda^* \left(\bigcup_{n\in\mathbb{N}} (a_n; b_n) \cap B^c \right) \\ &\leq \sum_{n\in\mathbb{N}} \lambda^*((a_n; b_n) \cap B) + \sum_{n\in\mathbb{N}} \lambda^*((a_n; b_n) \cap B^c) \\ &= \sum_{n\in\mathbb{N}} (b_n - a_n) \leq \lambda^*(A) + \varepsilon \,. \end{aligned}$$

Da $\varepsilon > 0$ beliebig ist, finden wir (11). $\square$

Es folgt aus Satz 1.28, dass $\mathcal{B}(\mathbb{R}) \subset \mathcal{M}_{\lambda^*} \subset P(\mathbb{R})$. Wir werden später zeigen,dass $\mathcal{B}(\mathbb{R}) \neq \mathcal{M}_{\lambda^*}$. Aus dieser Bemerkung folgt, dass das Lebesgue Mass entweder auf $\mathcal{B}(\mathbb{R})$ oder auf $\mathcal{M}_{\lambda^*}$ definiert werden kann. Man spricht von Lebesgue Mass auf $(\mathbb{R}, \mathcal{M}_{\lambda^*})$ und von Lebesgue Mass auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Die zwei Begriffe sind nicht äquivalent, und beide haben Vorteilen. Wir werden hier meistens mit dem Lebesgue Mass auf $(\mathbb{R}, \mathcal{M}_{\lambda^*})$ arbeiten, aber manchmal es ist auch nützlich das Lebesgue Mass auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ zu betrachten.

Analog wie auf $\mathbb{R}$, kann man auch das Lebesgue Mass auf $\mathbb{R}^n$, $n \geq 2$ definieren. Ein n-dimensionalen Intervall ist eine Teilmenge von $\mathbb{R}^n$ der Form

$$I = I_1 \times I_2 \times \cdots \times I_n = \{(x_1, x_2, \ldots, x_n) \subset \mathbb{R}^n : x_1 \in I_1, x_2 \in I_2, \ldots, x_n \in I_n\}$$

wobei $I_1, \ldots, I_n$ Intervalle sind (offen, abgeschlossen, oder halboffen, beschränkt oder unbeschränkt). Wir definieren das Volumen des n-dimensionale Intervall $I = I_1 \times \cdots \times I_n$ als das Produkt der Länge von $I_1, \ldots, I_n$, d.h.

$$\mathrm{vol}(I_1 \times I_2 \times \cdots \times I_n) = \prod_{j=1}^{n} |I_j|$$

Für eine Teilmenge $A \subset \mathbb{R}^n$, wir bezeichnen mit $\mathcal{C}_A$ die Familie aller Folgen $\{R_i\}_{i \in \mathbb{N}}$ von beschränkten und offenen n-dimensionale Intervalle R_i, mit $A \subset \bigcup_{i \in \mathbb{N}} R_i$. Wir definieren das Lebesgue äussere Mass von A durch

$$\lambda_n^*(A) = \inf \left\{ \sum_{i \in \mathbb{N}} \mathrm{vol}(R_i) : \{R_i\}_{i \in \mathbb{N}} \in \mathcal{C}_A \right\} \tag{12}$$

Dann kann man wie im Fall $n = 1$ sich überzeugen, dass λ_n^* ein äusseres Mass auf $\mathbb{R}^n$ ist, und, dass $\lambda_n^*(I) = \mathrm{vol}(I)$ für alle n-dimensionale Intervalle $I \subset \mathbb{R}^n$.

Definition 1.29. *Wir bezeichnen mit $\mathcal{M}_{\lambda_n^*} \subset P(\mathbb{R}^n)$ die Familie der λ_n^*-messbare Teilmengen von $\mathbb{R}^n$, und mit $\lambda_n = \lambda_n^*|_{\mathcal{M}_{\lambda_n^*}}$ die Einschränkung von λ_n^* auf $\mathcal{M}_{\lambda_n^*}$. Aus Theorem 1.25 folgt, dass $\mathcal{M}_{\lambda_n^*}$ eine σ-Algebra ist, und, dass λ_n ein Mass ist. Man nennt λ_n das n-dimensionale Lebesgue Mass (manchmal, falls keine Verwirrung möglich ist, werden wir das Index n weglassen, und das Lebesgue Mass auf $\mathbb{R}^n$ einfach mit λ bezeichnen).*

Wie im Fall $n = 1$ kann man zeigen, dass $\mathcal{B}(\mathbb{R}^n) \subset \mathcal{M}_{\lambda_n^*}$, d.h. jede Borel Menge ist λ_n^*-messbar. Wir werden später zeigen, dass, für alle $n \in \mathbb{N}$, $\mathcal{B}(\mathbb{R}^n) \neq \mathcal{M}_{\lambda_n^*}$. Das bedeutet, dass man für alle $n \in \mathbb{N}$ zwischen Lebesgue Mass auf $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ und Lebesgue Mass auf $(\mathbb{R}^n, \mathcal{M}_{\lambda_n^*}$ unterschieden werden soll.

Wir diskutieren nun einige Eingeschaften von Lebesgue Mass auf $\mathbb{R}^n$.

Regularität. Wir definieren den Begriff von Regularität von Massen, die (mindestens) auf der Borel σ-Algebra $\mathcal{B}(\mathbb{R}^n)$ definiert sind.

Definition 1.30. *Sei $\mathcal{A}$ eine σ-Algebra auf $\mathbb{R}^n$, die $\mathcal{B}(\mathbb{R}^n)$ enthält. Ein Mass μ auf $\mathcal{A}$ heisst regulär, falls*

i) *$\mu(K) < \infty$ für jede kompakte Menge $K \subset \mathbb{R}^n$ (kompakte Menge sind immer abgeschlossen, und deswegen in der Borel σ-Algebra enthalten).*

ii) *Für jede $A \subset \mathcal{A}$ gilt*

$$\mu(A) = \inf\{\mu(U) : U \text{ offen und } A \subset U\}$$

iii) *Für jede offene $U \subset \mathbb{R}^n$ gilt*

$$\mu(U) = \sup\{\mu(K) : K \text{ kompakt und } K \subset U\}$$

Für das Lebesgue Mass gilt offenbar: $\lambda_n(K) < \infty$ für alle kompakte Mengen $K \subset \mathbb{R}^n$. In der Tat ist jede kompakte Menge beschränkt; deswegen finden wir ein kompaktes n-dimensionales Intervall $I \subset \mathbb{R}^n$ mit $K \subset I$. Dann folgt die Behauptung aus $\lambda_n(K) \leq \lambda_n(I) = \mathrm{vol}(I) < \infty$. Dass λ_n auch die Eigenschaften ii) und iii) in der Definition von Regularität erfüllt folgt aus dem nächsten Satz.

Satz 1.31. *Sei $A \in \mathcal{M}_{\lambda_n^*}$. Dann gilt*

a) $\lambda_n(A) = \inf\{\lambda_n(U) : U$ *offen und* $A \subset U\}$.

b) $\lambda_n(A) = \sup\{\lambda_n(K) : K$ *kompakt und* $K \subset A\}$.

Beweis. Wir zeigen zunächst a). Aus der Monotonie von λ_n folgt sofort, dass

$$\lambda_n(A) \leq \inf\{\lambda_n(U) : U \text{ offen und } A \subset U\}\,.$$

Wir müssen die Ungleichung

$$\lambda_n(A) \geq \inf\{\lambda_n(U) : U \text{ offen und } A \subset U\} \tag{13}$$

beweisen. Wir können annehmen, dass $\lambda_n(A) < \infty$, sonst ist die Behauptung klar. Sei $\varepsilon > 0$. Aus Definition von Lebesgue äussere Mass, es existiert eine Folge $\{R_i\}_{i\in\mathbb{N}}$ von offenen n-dimensionale Intervalle mit $A \subset \bigcup_{i\in\mathbb{N}} R_i$ und

$$\lambda_n(A) \geq \sum_{i=1}^n \mathrm{vol}(R_i) - \varepsilon$$

Sei $U = \bigcup_{i\in\mathbb{N}} R_i$. Dann ist U offen und $A \subset U$. Weiter gilt, aus der Subadditivität von λ_n,

$$\lambda_n(U) \leq \sum_{i=1}^n \lambda_n(R_i) = \sum_{i=1}^n \mathrm{vol}(R_i) \leq \lambda_n(A) + \varepsilon$$

Hier haben wir die Tatsache $\lambda(R_i) = \mathrm{vol}(R_i)$ benutzt, die wir nach (12) erwähnt haben. Da $\varepsilon > 0$ beliebig ist, erhalten wir (13).

Nun zeigen wir b). Wegen der Monotonie, es genügt die Ungleichung

$$\lambda_n(A) \leq \sup\{\lambda_n(K) : K \text{ ist kompakt und} K \subset A\} \tag{14}$$

zu zeigen. Sei zunächst A beschränkt. Dann finden wir eine kompakte Menge $C \subset \mathbb{R}^n$ mit $A \subset C$. Sei $\varepsilon > 0$ fest. Aus Teil a) finden wir $U \subset \mathbb{R}^n$ offen, mit $C\backslash A = C \cap A^c \subset U$ und

$$\lambda_n(U) \leq \lambda_n(C\backslash A) + \varepsilon \;\Rightarrow\; \lambda_n(A) \leq \lambda_n(C) - \lambda_n(U) + \varepsilon\,.$$

Wir setzen dann $K = C\backslash U$. Dann ist $K \subset A$ sicher abgeschlossen und beschränkt, und deswegen kompakt. Da $C \subset K \cup U$ finden wir

$$\lambda_n(C) \leq \lambda_n(K) + \lambda_n(U)$$

und deswegen

$$\lambda_n(A) \leq \lambda_n(K) + \varepsilon \tag{15}$$

Da $\varepsilon > 0$ beliebig ist, folgt (14). Sei nun $A \subset \mathbb{R}^n$ nicht beschränkt. Um (14) zu zeigen, konstruieren wir, für ein beliebiges $b \leq \lambda_n(A)$, eine kompakte Menge $K \subset A$ mit $\lambda_n(K) \geq b$. Dazu wählen wir eine wachsende Folge $\{A_j\}_{j\in\mathbb{N}}$ von beschränkten Teilmengen von A mit $A = \bigcup_{j\in\mathbb{N}} A_j$ (wachsend bedeutet, dass $A_{j+1} \supset A_j$ für alle $j \in \mathbb{N}$; z.B. kann man $A_j = A \cap \{x \in \mathbb{R}^n : |x| \leq j\}$ definieren). Dann gilt $\lambda_n(A) = \lim_{j\to\infty} \lambda_n(A_j)$. Da $b \leq \lambda_n(A)$ es existiert $j_0 \in \mathbb{N}$, mit $b \leq \lambda_n(A_{j_0})$. Da A_{j_0} beschränkt ist, finden wir also aus (15) eine kompakte Menge $K \subset A_{j_0} \subset A$ mit $\lambda_n(K) \geq b$. $\square$

Einduetigkeit. Wir zeigen, dass Lebesgue Mass das einzige Mass auf $\mathcal{B}(\mathbb{R}^n)$ ist, das jedem n-dimensionalen Intervall I sein Volumen $\text{vol}(I)$ zuordnet. Dazu brauchen wir das folgende Lemma.

Lemma 1.32. *Sei $A \subset \mathbb{R}^n$ offen. Dann existiert eine Folge $\{Q_\ell\}$ von disjunkten halb-offenen Quadern der Form*

$$\{(x_1, \dots, x_n) \in \mathbb{R}^n : 2^{-k} j_i \leq x_i < 2^{-k}(j_i + 1) \text{ für } i = 1, \dots, n\} \tag{16}$$

mit $k \in \mathbb{N}$, $j_1, \dots, j_n \in \mathbb{Z}$, so, dass $A = \bigcup_{\ell\in\mathbb{N}} Q_\ell$.

Beweis. Für $k \in \mathbb{N}$ wir bezeichnen mit $\mathcal{C}_k$ die Familie der Quadern der Form

$$\{(x_1, \dots, x_n) \in \mathbb{R}^n : 2^{-k} j_i \leq x_i < 2^{-k}(j_i + 1) \text{ für } i = 1, \dots, n\}$$

für beliebig $j_1, \dots, j_n \in \mathbb{Z}$. Wir bemerken, dass, für feste $k \in \mathbb{N}$, $\mathcal{C}_k$ eine abzählbare Partition von $\mathbb{R}^n$ ist (das heisst die Mengen in $\mathcal{C}_k$ sind paarweise disjunkt, und die Vereinigung von allen Mengen in $\mathcal{C}_k$ gibt $\mathbb{R}^n$). Für $k_1 < k_2$ ist jede Quader in $\mathcal{C}_{k_2}$ in genau ein Quader in $\mathcal{C}_{k_1}$ enthalten.

Sei nun $A \subset \mathbb{R}^n$ offen. Wir konstruieren rekursiv eine Familie $\mathcal{D}$ von Quadern der Form (16). Wir beginnen mit $k = 0$; wir nehmen in $\mathcal{D}$ alle Quader in $\mathcal{C}_0$, die in A enthalten sind. Für $k = 1$ addieren wir zu $\mathcal{D}$ alle Quader in $\mathcal{C}_1$, die in A enthalten sind, und die disjunkt sind zu allen $k = 0$ Quadern, die schon in $\mathcal{D}$ sind. Wir gehen rekursiv weiter. Im k-te Schritt, addieren wir zu $\mathcal{D}$ die Quader in $\mathcal{C}_k$, die in A enthalten sind, und die disjunkt sind, zu allen Quader die schon in $\mathcal{D}$ sind. Offenbar enthält $\mathcal{D}$ schlussendlich abzählbar viele paarweise disjunkte Quader der Form (16), die in A enthalten sind. Es bleibt zu zeigen, dass A in der Vereinigung von allen Quadern in $\mathcal{D}$ enthalten ist. Sei also $x \in A$ beliebig. Für jede $k \in \mathbb{N}$ gibt es genau ein Quader $Q_k \in \mathcal{C}_k$, mit $x \in Q_k$. Da A offen ist, ist $Q_k \subset A$ für k gross genug. Sei k_0 die kleinste ganze Zahl mit $Q_{k_0} \subset A$. Dann ist $Q_{k_0} \in \mathcal{D}$, und $x \in \bigcup_{Q\in\mathcal{D}} Q$. $\square$

Wir können nun die Eindeutigkeit vom Lebesgue Mass beweisen.

Satz 1.33. *Sei μ ein Mass auf $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, mit $\mu(I) = vol(I)$ für alle n-dimensionale Intervalle $I = I_1 \times \cdots \times I_n$. Dann ist $\mu = \lambda_n$. Die selbe Aussage gilt auch falls $\mu(Q) = \lambda_n(Q)$ nur für alle halb-offenen Quadern Q der Form (16).*

Beweis. Wir müssen zeigen, dass $\mu(A) = \lambda_n(A)$ für alle $A \in \mathcal{B}(\mathbb{R}^n)$. Sei zunächst $U \subset \mathbb{R}^n$ offen. Dann finden wir aus Lemma 1.32 eine Folge Q_n von disjunkten Quadern der Form (16) mit $U = \bigcup_{j\in\mathbb{N}} Q_j$. Aus der σ-Additivität von μ und λ_n erhalten wir

$$\mu(U) = \sum_{j\in\mathbb{N}} \mu(Q_j) = \sum_{j\in\mathbb{N}} \lambda_n(Q_j) = \lambda(U)$$

weil, nach Annahme, $\mu(Q_i) = \lambda_n(Q_i)$ für alle $i \in \mathbb{N}$.

Sei nun $A \in \mathcal{B}(\mathbb{R}^n)$ beliebig. Für jede offene $U \subset \mathbb{R}^n$ mit $A \subset U$ gilt $\mu(A) \leq \mu(U) = \lambda(U)$. Es folgt, dass

$$\mu(A) \leq \inf\{\lambda_n(U) : U \text{ offen und } A \subset U\}$$

Nach der Regularität von λ_n, finden wir $\mu(A) \leq \lambda_n(A)$, für alle $A \in \mathcal{B}(\mathbb{R}^n)$. Wir müssen noch die umgekehrte Ungleichung beweisen. Sei $A \in \mathcal{B}(\mathbb{R}^n)$ beschränkt. Dann finden wir eine offene Menge $U \subset \mathbb{R}^n$ mit $A \subset U$. In diesem Fall gilt

$$\mu(U) = \mu(A) + \mu(U \backslash A) \leq \lambda_n(A) + \lambda_n(U \backslash A) = \lambda_n(U)$$

Da aber $\mu(U) = \lambda_n(U)$ es muss überall Gleichheit gelten, d.h.

$$\mu(A) + \mu(U \backslash A) = \lambda_n(A) + \lambda_n(U \backslash A)$$

Aus $\mu(A) \leq \lambda_n(A)$ und $\mu(U \backslash A) \leq \lambda_n(U \backslash A)$, erhalten wir $\mu(A) = \lambda_n(A)$ und $\mu(U \backslash A) = \lambda_n(U \backslash A)$.

Sei nun $A \in \mathcal{B}(\mathbb{R}^n)$ beliebig (nicht unbedingt beschränkt). Wir setzen $A_k = A \cap \{x \in \mathbb{R}^n : |x| < k\}$. A_k ist eine wachsende Folge mit $A = \bigcup_{k \in \mathbb{N}} A_k$. Deswegen gilt

$$\mu(A) = \lim_{k \to \infty} \mu(A_k) = \lim_{k \to \infty} \lambda_n(A_k) = \lambda_n(A)$$

weil $A_k \in \mathcal{B}(\mathbb{R}^n)$ für alle $k \in \mathbb{N}$ beschränkt ist. □

Bemerkung: Analog zeigt man: ist μ Mass auf $(\mathbb{R}^n, \mathcal{M}_{\lambda_n^*})$, mit $\mu(I) = \lambda_n(I)$ für alle halb-offenen Quader der Form (16), dann gilt $\mu(A) = \lambda_n(A)$ für alle $A \in \mathcal{M}_{\lambda_n^*}$.

Translationsinvarianz. Im nächsten Lemma zeigen wir, dass das Lebesgue Mass translationsinvariant ist.

Lemma 1.34. *Das Lebesgue Mass λ_n auf $(\mathbb{R}^n, \mathcal{M}_{\lambda_n^*})$ ist translationsinvariant. Das heisst, für $A \in \mathcal{M}_{\lambda_n^*}$ und $x \in \mathbb{R}^n$, gilt $A + x \in \mathcal{M}_{\lambda_n^*}$ und $\lambda_n(A + x) = \lambda_n(A)$. Auch das Lebesgue Mass auf $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ ist translationsinvariant (d.h. für $A \in \mathcal{B}(\mathbb{R}^n)$ und $x \in \mathbb{R}^n$ ist $A + x \in \mathcal{B}(\mathbb{R}^n)$ und $\lambda_n(A + x) = \lambda_n(A)$).*

Beweis. Aus der Definition der äussere Mass λ_n^* gilt offenbar $\lambda_n^*(A + x) = \lambda_n^*(A)$ für alle $A \subset \mathbb{R}^n$ und $x \in \mathbb{R}^n$. Das impliziert, dass, falls $B \subset \mathbb{R}^n$ λ_n^*-messbar ist, so ist auch $B + x$ λ_n^*-messbar, weil, für beliebiges $A \subset \mathbb{R}^n$,

$$\begin{aligned}
\lambda_n^*(A) &= \lambda_n^*(A - x) \\
&= \lambda_n^*((A - x) \cap B) + \lambda_n^*((A - x) \cap B^c) \\
&= \lambda_n^*((A - x) \cap B + x) + \lambda_n^*((A - x) \cap B^c + x) \\
&= \lambda_n^*(A \cap (B + x)) + \lambda_n^*(A \cap (B + x)^c) \,.
\end{aligned}$$

Aus der Translationsinvarianz von λ_n^* folgt dann auch die Invarianz von λ_n auf $\mathcal{M}_{\lambda_n^*}$. Um die zweite Aussage vom Lemma zu beweisen, müssen wir noch zeigen, dass $A+x \in \mathcal{B}(\mathbb{R}^n)$

für alle $A \in \mathcal{B}(\mathbb{R}^n)$ und $x \in \mathbb{R}^n$. Dazu bemerken wir, dass, für ein beliebiges $x \in \mathbb{R}^n$, die Familie

$$\mathcal{B}_x := \{A \subset \mathbb{R}^n : A + x \in \mathcal{B}(\mathbb{R}^n)\}$$

eine σ-Algebra ist (Beweis: Übung), die alle offene Mengen enthält. Deswegen ist $\mathcal{B}(\mathbb{R}^n) \subset \mathcal{B}_x$. Das bedeutet, dass für jede $A \in \mathcal{B}(\mathbb{R}^n)$ ist $A + x \in \mathcal{B}(\mathbb{R}^n)$, und zeigt die Translationsinvarianz von λ_n, definert auf $\mathcal{B}(\mathbb{R}^n)$. □

Wir zeigen im nächsten Satz, dass die Translationsinvarianz, bis auf einer multiplikativen Konstante, das Lebesgue Mass auf $\mathcal{B}(\mathbb{R}^n)$ charakterisiert.

Satz 1.35. *Sei $\mu \not\equiv 0$ ein Mass auf $(\mathbb{R}^n; \mathcal{B}(\mathbb{R}^n))$, mit $\mu(A) < \infty$ für jede beschränkte Menge $A \in \mathcal{B}(\mathbb{R}^n)$, und mit $\mu(A+x) = \mu(A)$ für alle $A \in \mathcal{B}(\mathbb{R}^n)$ und alle $x \in \mathbb{R}^n$. Dann existiert eine Konstante $c > 0$ so, dass $\mu(A) = c\,\lambda_n(A)$ für alle $A \in \mathcal{B}(\mathbb{R}^n)$.*

Bemerkung: Die Annahme $\mu(A+x) = \mu(A)$ ist sinnvoll, weil $A + x \in \mathcal{B}(\mathbb{R}^n)$ für alle $A \in \mathcal{B}(\mathbb{R}^n)$, wie wir im Lemma 1.34 bewiesen haben.

Beweis. Sei $C = \{(x_1, \dots, x_n) : 0 \leq x_i < 1 \text{ für alle } i = 1, \dots, n\}$ und $c = \mu(C)$. Nach Annahme ist $0 < c < \infty$, weil C beschränkt ist (wäre $c = 0$, dann würde $\mu(\mathbb{R}) = 0$, weil $\mathbb{R}$ als Vereinigung von abzählbar viele Mengen der Form $C + z_j$, für geeigneten $z_j \in \mathbb{R}^n$ geschrieben werden kann). Wir definieren dann $\nu(A) := c^{-1}\mu(A)$ für alle $A \in \mathcal{B}(\mathbb{R}^n)$. Dann ist ν translationsinvariant und $\nu(C) = 1$. Sei

$$D = \{(x_1, \dots, x_n) \in \mathbb{R}^n : 2^{-k} j_i \leq x_i < 2^{-k}(j_i + 1) \text{ für } i = 1, \dots, n\} \tag{17}$$

für ein $k \in \mathbb{N}$ und $j_1, \dots, j_n \in \mathbb{Z}$. Dann folgt aus der Translationsinvarianz, dass

$$2^{nk}\nu(D) = \nu(C) = 1 = \lambda_n(C) = 2^{nk}\lambda_n(D)$$

Damit gilt $\nu(D) = \lambda_n(D)$ für alle Quadern der Form (17). Aus Satz 1.33, finden wir $\nu(A) = \lambda_n(A)$ für alle $A \in \mathcal{B}(\mathbb{R}^n)$. □

Bemerkung: Mit Satz 1.33 folgt auch: ist $\mu \not\equiv 0$ ein Mass auf $(\mathbb{R}^n, \mathcal{M}_{\lambda_n^*})$ mit $\mu(A + x) = \mu(A)$ für alle $A \in \mathcal{M}_{\lambda_n^*}$, so existiert $c > 0$ mit $\mu(A) = c\lambda_n(A)$ für alle $A \in \mathcal{M}_{\lambda_n^*}$.

Die Cantor Menge. Wir nennen eine Lebesgue-messbare Menge A mit $\lambda_n(A) = 0$ eine Nullmenge. Aus der Definition von Lebesgue Mass es ist klar, dass jede abzählbare Menge eine Nullmenge ist. Eine natürliche Frage ist: existieren überabzählbare Nullmengen? Die Antwort ist ja. Ein Beispiel ist die Cantor Menge. Da die Cantor Menge eine Quelle von interessanten Beispiele ist, möchten wir sie hier konstruieren und ihre wichtigsten Eigenschaften untersuchen.

Sei $K_0 = [0; 1]$. Wir definieren K_1, indem wir von K_0 das offene Intervall $(1/3; 2/3)$ entfernen; d.h. $K_1 = [0; 1/3] \cup [2/3; 1]$. Weiter wir definieren K_2 indem wir aus den zwei Intervalle $[0; 1/3]$ und $[2/3; 1]$ das offene mittlere Drittel entfernen; d.h. $K_2 = [0; 1/9] \cup [2/9; 1/3] \cup [2/3; 7/9] \cup [8/9; 1]$. Rekursiv definieren wir, für beliebigen $n \in \mathbb{N}$, die Menge K_n indem wir von den Intervallen, aus denen K_{n-1} besteht, immer das mittlere

Drittel entfernen. Damit ist K_n die Vereinigung von 2^n abgeschlossene Intervalle der Länge 3^{-n}. Die Cantor Menge ist dann definiert als

$$K = \bigcap_{n \in \mathbb{N}} K_n. \tag{18}$$

Verschiedene Eigenschaften von der Cantor Menge werden im folgenden Satz gezeigt.

Satz 1.36. *Die Cantor Menge K ist kompakt (und deswegen sicher eine Borel Menge). Sie ist gleichmächtig mit $\mathbb{R}$ (sie ist also überabzählbar), und hat Lebesgue Mass $\lambda(K) = 0$. Weiter, K enthält kein nichtleere offene Menge (d.h. das Inneres von K ist leer).*

Beweis. K ist offenbar beschränkt, und abgeschlossen (als Vereinigung von abgeschlossenen Mengen). K ist also kompakt.

K enthält kein offenes Intervall; ist nämlich $I \subset K$ ein offenes Intervall, dann muss insbesondere $I \subset K_n$ sein, für jede $n \in \mathbb{N}$. Da K_n aus 2^n disjunkten abgeschlossenen Intervalle der Länge 3^{-n} besteht, das impliziert, dass die Länge von I kleiner als 3^{-n} ist. Da $n \in \mathbb{N}$ beliebig ist, die Länge von I muss Null sein, und also $I = \emptyset$.

Es folgt, dass K keine nichtleere offene Menge enthält (ist $A \subset I$ offen, und $x \in A$, dann es gibt ein offenes Intervall I mit $x \in I$ und $I \subset A \subset K$; wir wissen aber, dass K kein offenes Intervall enthält).

Wir zeigen nun, dass das Lebesgue Mass von K Null ist. K_n ist die Vereinigung von 2^n Intervalle der Länge 3^{-n}. Aus der Additivität von λ (und weil das Mass eines Intervall immer seine Länge ist), finden wir $\lambda(K_n) = (2/3)^n$. Aus der Monotonie vom Lebesgue Mass finden wir

$$\lambda(K) \leq \lambda(K_n) = (2/3)^n$$

Da $n \in \mathbb{N}$ beliebig ist, muss $\lambda(K) = 0$ gelten.

Schlussendlich zeigen wir, dass K gleichmächtig mit $\mathbb{R}$ ist. Wir definieren dazu eine Bijektion J zwischen K und der Menge X aller Folgen $(\sigma_j)_{j\in\mathbb{N}}$ mit $\sigma_j \in \{0,1\}$. Die Menge X ist mit $[0;1)$ und damit auch mit $\mathbb{R}$ gleichmächtig; das kann zum Beispiel gezeigt werden, indem man die Binardarstellung von reellen Zahlen betrachtet, d.h. die Abbildung $X \to [0,1)$ definiert durch

$$(\sigma_j)_{j\in\mathbb{N}} \to \sum_{j=0}^{\infty} \sigma_j 2^{-(j+1)}$$

(besser gesagt, um eine injäktive Abbildung zu konstruieren, muss man die Folgen in X mit $\sigma_j = 1$ für alle $j > j_0$ vernachlässigen; die Menge dieser Folge hat aber abzählbare Kardinalität).

Wir definieren die Abbildung $J : K \to X$ wie folgt. Für $x \in K$, muss $x \in K_n$ für alle n sein. Für $n = 1$, die Tatsache $x \in K_1$ bedeutet, dass x in einem von zwei Intervallen ist, $[0;1/3]$ oder $[2/3;1]$. Wir setzen $\sigma_1(x) = 0$, falls $x \in [0;1/3]$ und $\sigma_1(x) = 1$, falls $x \in [2/3;1]$. Für $n = 2$, x kann in einem von vier Intervallen sein. Ist $x \in [0;1/3]$ so kann $x \in [0;1/9]$ oder $x \in [2/9;1/3]$ sein. Ist dagegen $x \in [2/3;1]$, so kann x entweder in $[2/3;7/9]$ oder in $[8/9;1]$ sein. Wir definieren $\sigma_2(x) = 0$, falls $x \in [0;1/9]$ oder $x \in$

$[2/3; 7/9]$ (d.h. falls x in dem linken Drittel von einem Intervall von K_1 ist). Wir setzten dagegen $\sigma_2(x) = 1$ falls $x \in [2/9; 1/3]$ oder falls $x \in [8/9; 1]$ (d.h. falls x im rechten Drittel von einem Intervall von K_2 ist). Wir definieren rekursiv $\sigma_n(x)$ für alle $n \in \mathbb{N}$, indem wir setzen $\sigma_n(x) = 0$, falls x im linken Drittel von einem Intervall in K_{n-1} ist, und $\sigma_n(x) = 1$, falls x im rechten Drittel von einerm Intervall in K_{n-1} ist. Das definiert $\sigma(x) = (\sigma_n(x))_{n\in\mathbb{N}} \in X$ für alle $x \in K$. Die Abbildung $J(x) = (\sigma_n(x))_{n\in\mathbb{N}}$ definiert eine Bijektion zwischen K und X. Die Injäktivität von J kann wie folgt gezeigt werden. Ist $\sigma_n(x) = \sigma_n(y)$ für alle $n \in \mathbb{N}$, so müssen x, y im selben Intervall von K_n sein, für alle $n \in \mathbb{N}$. Da Intervalle in K_n die Länge 3^{-n} haben, muss $|x-y| \leq 3^{-n}$ sein, für alle $n \in \mathbb{N}$. Das zeigt, dass $x = y$. Um die Surjektivität zu beweisen, bemerken wir, dass J invertiert werden kann, indem wir $L : X \to K$ durch

$$(\sigma_j)_{j\in\mathbb{N}} \to x := 2\sum_{j=1}^{\infty} \sigma_j 3^{-j}$$

definieren. □

Existenz von nicht messbare Mengen. Die Tatsache, dass $\mathcal{M}_{\lambda_n^*} \neq P(\mathbb{R}^n)$ folgt aus Satz 1.1, wo wir gezeigt haben, es existiert keine σ-additive und translationsinvariante Abbildung $\mu : P(\mathbb{R}^n) \to [0; \infty]$, mit $\mu([0; \ell]^{\times n}) = \ell^n$ für alle $\ell > 0$. Wäre nun $\mathcal{M}_{\lambda_n^*} = P(\mathbb{R}^n)$, dann würde das Lebesgue Mass genau diese Eigenschaften haben; das würde also die Aussage von Satz 1.1 widersprechen.

Aus dem Beweis von Satz 1.1 es ist auch einfach eine Menge zu konstruieren, die nicht in $\mathcal{M}_{\lambda_n^*}$ enthalten ist. Sei $A = [0; 1]^{\times n}$. Auf A definieren wir die Äquivalenzrelation $x \sim y$ falls $x - y \in \mathbb{Q}^n$. Sei nun $M_0 \subset A$ eine Teilmenge von $[0; 1]^{\times n}$, die aus jeder Äquivalenzklasse genau ein Element enthält. Dann folgt, wie im Beweis von Satz 1.1, dass $M_0 \subset [0; 1]^{\times n}$ nicht Lebesgue messbar ist.

Zur Construction von M_0 haben wir das Auswahlaxiom benutzt. Es wurde eigentlich in den sechsziger Jahre gezeigt, dass das Auswahlaxiom notwendig ist, um die Existenz von nicht-Lebesgue messbare Menge zu zeigen. Ohne Auswahlaxiom kann man also die Aussage $\mathcal{M}_{\lambda_n^*} = P(\mathbb{R}^n)$ nicht widersprechen.

Vollständigkeit. Sei $(\Omega, \mathcal{A}, \mu)$ ein Massraum. Das Mass μ heisst vollständig, falls

$$A \in \mathcal{A}, \mu(A) = 0 \text{ und } B \subset A \;\Rightarrow\; B \in \mathcal{A} \text{ und } \mu(B) = 0$$

Ist $B \in \mathcal{A}$, so folgt offenbar aus der Monotonie, dass $\mu(B) = 0$. Vollständigkeit ist also die Aussage, dass jede Teilmenge einer Nullmenge in der σ-Algebra enthalten ist (so, dass sein Mass berechnet werden kann).

Sei Ω eine Menge, und $\mu^* : P(\Omega) \to [0; \infty]$ ein äusseres Mass auf Ω. Wir bezeichnen mit $\mathcal{M}_{\mu^*}$ die σ-Algebra der μ^*-messbare Mengen, und mit μ die Einschränkung von μ^* auf $\mathcal{M}_{\mu^*}$; wir wissen μ ist ein Mass auf $\mathcal{M}_{\mu^*}$. Wir haben in Satz 1.24 bewiesen, dass $\mu^*(A) = 0$ impliziert, dass $A \in \mathcal{M}_{\mu^*}$. Ist also $A \in \mathcal{M}_{\mu^*}$ mit $\mu(A) = 0$ und $B \subset A$, so gilt, aus der Monotonie des äusseren Mass μ^*, $\mu^*(B) = 0$. Deswegen ist $B \in \mathcal{M}_{\mu^*}$ messbar, mit $\mu(B) = 0$. Mit andere Wörter: μ, definiert auf der σ-Algebra $\mathcal{M}_{\mu^*}$ ist sicher vollständig. Zum Beispiel ist das n-dimensionale Lebesgue Mass λ_n, definiert auf der σ-Algebra $\mathcal{M}_{\lambda_n^*}$ der Lebesgue messbare Mengen vollständig. Wir werden später sehen, dass das Lebesgue Mass, definiert auf der Borel σ-Algebra $\mathcal{B}(\mathbb{R}^n)$, dagegen nicht vollständig ist.

Vollständigkeit eines Masses kann manchmal nützlich sein. Es gibt ein einfaches Verfahren, um ein nicht notwendigerweise vollständiges Massraum zu vervollständigen. Sei $(\Omega, \mathcal{A}, \mu)$ ein Massraum. Die Vervollständigung von $\mathcal{A}$ bezüglich dem Mass μ ist die Familie $\mathcal{A}_\mu \subset P(\Omega)$ die aus allen Teilmengen $A \subset \Omega$ besteht, für die $E, F \in \mathcal{A}$ existieren, mit $E \subset A \subset F$ und $\mu(F \backslash E) = 0$. Bemerke, dass $\mathcal{A} \subset \mathcal{A}_\mu$ (ist $A \in \mathcal{A}$, dann können wir einfach $E = F = A$ wählen).

Existieren $E, F \in \mathcal{A}$ mit $\mu(F \backslash E) = 0$ so muss offenbar $\mu(F) = \mu(E)$ gelten (weil $\mu(F \backslash E) = \mu(F) - \mu(E)$, falls $\mu(E) < \infty$; ist dagegen $\mu(E) = \infty$, so muss auch $\mu(F) = \infty$ sein). Ist $B \subset A$ und $B \in \mathcal{A}$, so gilt

$$\mu(B) \leq \mu(F) = \mu(E)$$

Es folgt, dass

$$\mu(F) = \mu(E) = \sup\{\mu(B) : B \subset A, \; B \in \mathcal{A}\}$$

wobei die rechte Seite nun unabhängig aus der besondere Wahl von $E, F \in \mathcal{A}$, mit $E \subset A \subset F$ ist. Also, für $A \in \mathcal{A}_\mu$, können wir $\bar{\mu}(A) := \mu(E) = \mu(F)$ für eine beliebige Wahl von $E, F \in \mathcal{A}$, mit $E \subset A \subset F$ und $\mu(F \backslash E) = 0$. Wir nennen $\bar{\mu}$ die Vervollständigung von μ.

Satz 1.37. *Sei $(\Omega, \mathcal{A}, \mu)$ ein Massraum. Dann ist $\mathcal{A}_\mu$ eine σ-Algebra auf Ω, und $\bar{\mu}$ ein vollständiges Mass auf $\mathcal{A}_\mu$, dessen Einschränkung auf $\mathcal{A}$ gleich μ ist.*

Beweis. Da $\mathcal{A} \subset \mathcal{A}_\mu$ ist offenbar $\Omega \in \mathcal{A}_\mu$. Ist $A \in \mathcal{A}_\mu$, so finden wir $E, F \in \mathcal{A}$ mit $E \subset A \subset F$ und $\mu(F \backslash E) = 0$. Dann sind $E^c, F^c \in \mathcal{A}$, mit $F^c \subset A^c \subset E^c$, und $\mu(E^c \backslash F^c) = \mu(F \backslash E) = 0$. Damit ist auch $A^c \in \mathcal{A}_\mu$. Sei nun $\{A_n\}_{n \in \mathbb{N}}$ eine Folge aus $\mathcal{A}_\mu$. Für alle $n \in \mathbb{N}$ finden wir dann $E_n, F_n \in \mathcal{A}$ mit $E_n \subset A_n \subset F_n$ und $\mu(F_n \backslash E_n) = 0$. Dann sind $E = \bigcup_{n \in \mathbb{N}} E_n$ und $F = \bigcup_{n \in \mathbb{N}} F_n$ in $\mathcal{A}$, mit $E \subset \bigcup_{n \in \mathbb{N}} A_n \subset F$ und

$$\mu(F \backslash E) = \mu\left(\bigcup_{n \in \mathbb{N}} F_n \backslash \bigcup_{n \in \mathbb{N}} E_n\right) \leq \mu\left(\bigcup_{n \in \mathbb{N}} (F_n \backslash E_n)\right) \leq \sum_{n \in \mathbb{N}} \mu(F_n \backslash E_n) = 0$$

weil

$$\bigcup_{n \in \mathbb{N}} F_n \backslash \bigcup_{n \in \mathbb{N}} E_n \subset \bigcup_{n \in \mathbb{N}} (F_n \backslash E_n) \, .$$

Deswegen ist $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}_\mu$. Das zeigt, dass $\mathcal{A}_\mu$ eine σ-Algebra ist.

Wir zeigen nun die σ-Additivität von $\bar{\mu}$. Sei $\{A_n\}_{n \in \mathbb{N}}$ eine Folge von disjunkten Mengen in $\mathcal{A}_\mu$. Für alle $n \in \mathbb{N}$ finden wir $E_n, F_n \in \mathcal{A}$ mit $E_n \subset A_n \subset F_n$ und $\mu(F_n \backslash E_n) = 0$; es gilt dann $\bar{\mu}(A_n) = \mu(E_n)$. Wir bemerken auch, dass die Mengen E_n paarweis disjunkt sind. Sei $E = \bigcup_{n \in \mathbb{N}} E_n$, und $F = \bigcup_{n \in \mathbb{N}} F_n$. Es gilt $E, F \in \mathcal{A}$, mit $E \subset \bigcup_{n \in \mathbb{N}} A_n \subset F$ und $\mu(F \backslash E) = 0$. Deswegen gilt

$$\bar{\mu}\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \mu(E) = \mu\left(\bigcup_{n \in \mathbb{N}} E_n\right) = \sum_{n \in \mathbb{N}} \mu(E_n) = \sum_{n \in \mathbb{N}} \bar{\mu}(A_n)$$

wobei wir die σ-Additivität von μ benutzt haben sowie die Tatsache, dass die Mengen E_n paarweise disjunkt sind.

Schlussendlich zeigen wir, dass $\bar{\mu}$, definiert auf $\mathcal{A}_\mu$ vollständig ist. Sei $A \in \mathcal{A}_\mu$ mit $\bar{\mu}(A) = 0$, und $B \subset A$. Dann es existieren $E, F \in \mathcal{A}$ mit $E \subset A \subset F$ und $\mu(F \backslash E) = 0$. Es gilt $\mu(F) = \mu(E) = \bar{\mu}(A) = 0$. Für B finden wir also $\emptyset, F \in \mathcal{A}$, mit $\emptyset \subset B \subset F$ und $\mu(F \backslash \emptyset) = \mu(F) = 0$. Deswegen ist $B \in \mathcal{A}_\mu$, und $\bar{\mu}(B) = 0$.

Die Tatsache, dass $\bar{\mu}(A) = \mu(A)$ für alle $A \in \mathcal{A}$ ist klar (in diesem Fall können wir $E = F = A$ wählen). $\square$

Bemerkung: ist $(\Omega, \mathcal{A}, \mu)$ vollständig, dann gilt $\mathcal{A}_\mu = \mathcal{A}$ und $\bar{\mu} = \mu$. In der Tat: $A \in \mathcal{A}_\mu$ impliziert, dass es $E, F \in \mathcal{A}$ existieren, mit $E \subset A \subset F$ und $\mu(F \backslash E) = 0$. Dann ist aber $A \backslash E \subset F \backslash E$; ist μ vollständig, dann folgt $A \backslash E \in \mathcal{A}$, als Teilmenge der Nullmenge $F \backslash E$. Dann ist auch $A = E \cup (A \backslash E) \in \mathcal{A}$. Damit gilt $\mathcal{A}_\mu = \mathcal{A}$, falls$|mu$ vollständig.

Ein wichtiges Beispiel ist im folgenden Satz behandelt.

Satz 1.38. *Das Lebesgue Mass λ_n auf der σ-Algebra $\mathcal{M}_{\lambda_n^*}$ ist die Vervollständigung von Lebesgue Mass λ_n auf der Borel σ-Algebra.*

Beweis. Seien $\mathcal{B}$ die Borel σ-Algebra auf $\mathbb{R}^n$ und λ_b das Lebesgue Mass auf $\mathcal{B}$, $\mathcal{M}$ die σ-Algebra der Lebesgue messbare Mengen und λ_m Lebesgue Mass auf $\mathcal{M}$, und, schlussendlich, $\mathcal{B}_{\lambda_b}$ die Vervollständigung von $\mathcal{B}$ bezüglich λ_b und $\bar{\lambda}_b$ die Vervollständigung von λ_b.

Wir zeigen zunächst, dass $\mathcal{M} \subset \mathcal{B}_{\lambda_b}$. Sei nämlich $A \in \mathcal{M}$ messbar. Wir nehmen zunächst an $\lambda_m(A) < \infty$. Aus der Regularität von Lebesgue Mass, finden wir für alle $k \in \mathbb{N}$ eine kompakte Menge $K_k \subset A$ und eine offene Menge $U_k \supset A$, mit $\lambda_b(K_k) \geq \lambda_m(A) - 1/k$ und $\lambda_b(U_k) \leq \lambda_m(A) + 1/k$. Offenbar sind $K_k, U_k \in \mathcal{B}(\mathbb{R}^n)$ für alle $k \in \mathbb{N}$. Wir setzen $E = \bigcup_{k\in\mathbb{N}} K_k$ und $F = \bigcap_{k\in\mathbb{N}} U_k$. Es gilt $E, F \in \mathcal{B}(\mathbb{R}^n)$ mit $E \subset A \subset F$ und

$$\lambda_b(F \backslash E) \leq \lambda_b(U_k \backslash K_k) = \lambda_m(U_k \backslash K_k) \leq \lambda_m(U_k \backslash A) + \lambda_m(A \backslash K_k) \leq 2/k$$

für alle $k \in \mathbb{N}$ (wir haben hier die Tatsache benutzt, dass die Einschränkung von λ_m auf $\mathcal{B}$ gerade λ_b ist). Deswegen muss $\lambda_b(F \backslash E) = 0$, und $A \in \mathcal{B}_{\lambda_b}$. Weiter ist $\bar{\lambda}_b(A) = \lambda_b(E) = \lambda_b(F) = \lambda_m(E) = \lambda_m(F) = \lambda_m(A)$ (weil $E \subset A \subset F$, und $E, A, F \in \mathcal{M}$).

Ist $A \in \mathcal{M}$, möglicherkweise mit $\lambda_n(A) = \infty$, so finden wir eine wachsende Folge von Mengen $A_k \in \mathcal{M}$ mit $A = \bigcup_{k\in\mathbb{N}} A_k$ und so, dass $\lambda_m(A_k) < \infty$ für alle $k \in \mathbb{N}$. Für jede k finden wir $E_k, F_k \in \mathcal{B}$ mit $E_k \subset A_k \subset F_k$ und $\lambda_b(F_k \backslash E_k) = 0$. Dann setzen wir einfach $E = \bigcup_{k\in\mathbb{N}} E_k$ und $F = \bigcup_{k\in\mathbb{N}} F_k$. Es gilt $E, F \in \mathcal{B}$, $E \subset A \subset F$ und $\lambda_b(F \backslash E) = 0$. Das impliziert, dass $A \in \mathcal{B}_{\lambda_b}$, mit $\bar{\lambda}_b(A) = \lambda_b(E)$. Es folgt auch, dass $\lambda_m(F \backslash E) = 0$ und damit, dass $\lambda_m(F) = \lambda_m(E)$. Da $E \subset A \subset F$, es folgt, dass $\lambda_m(A) = \lambda_m(E) = \lambda_b(E) = \bar{\lambda}_b(A)$.

Wir haben bis jetzt gezeigt, dass $\mathcal{M} \subset \mathcal{B}_{\lambda_b}$, und, dass die Einschränkung von $\bar{\lambda}_b$ auf $\mathcal{M}$ gerade λ_m ist. Es bleibt zu zeigen, dass $\mathcal{B}_{\lambda_b} \subset \mathcal{M}$. Sei also $A \in \mathcal{B}_{\lambda_b}$. Dann es existieren $E, F \in \mathcal{B}$ mit $E \subset A \subset F$ und $\lambda_b(F \backslash E) = 0$. Dann gilt auch $F \backslash E \in \mathcal{M}$ und $\lambda_m(F \backslash E) = 0$. Dann ist aber $A \backslash E \subset F \backslash E$, und da λ_m auf $\mathcal{M}$ vollständig ist, muss $A \backslash E \in \mathcal{M}$ sein. Deswegen ist auch $A = E \cup (A \backslash E) \in \mathcal{M}$. $\square$

Bemerkung: Wir haben noch nicht gezeigt, dass Lebesgue Mass auf $\mathcal{B}(\mathbb{R}^n)$ nicht vollständig ist. Es könnte in Prinzip $\mathcal{M}_{\lambda_n^*} = \mathcal{B}(\mathbb{R}^n)$ gelten; das wäere nicht in Widerspruch mit dem letzten Satz. Wir werden aber in nächsten Kapitel zeigen (mit Hilfe des Auswahlaxiom), dass tatsächlich $\mathcal{B}(\mathbb{R}^n) \neq \mathcal{M}_{\lambda_n^*}$. Damit ist klar, dass $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \lambda_n)$ nicht vollständig ist.

Bemerkung: Obwohl manchmal die Vollständigkeit eines Masses nützlich ist, es gibt auch Nachteile, wenn wir ein Mass durch seine Vervollständigung ersetzen. Zum Beispiel kann die σ-Algebra $\mathcal{A}_\mu$ viel komplizierter sein als die ursprüngliche σ-Algebra $\mathcal{A}$ (das ist zum Beispiel der Fall für die Borel σ-Algebra, vergliechen mit der Vervollständigung $\mathcal{M}_{\lambda_n^*}$). Ein anderer Nachteil: wenn man mit zwei verschiedene Masse μ, ν auf $\mathcal{A}$ zu tun haben, dann sind normalerweise die Vervollständigungen $\mathcal{A}_\mu$ und $\mathcal{A}_\nu$ nicht die selbe σ-Algebra; dann ist schwierig μ und ν zu vergleichen. Es ist deswegen immer besser, wann möglich, Theoremen zu beweisen, ohne Vollständigkeit anzunehmen.

2 Integrationstheorie

Wir kommen nun in diesem zweiten Kapitel zur Konstruktion vom Integral.

2.1 Messbare Funktionen

Sei $(\Omega, \mathcal{A})$ ein messbarer Raum, und $A \in \mathcal{A}$ eine Teilmenge von Ω, die in der σ-Algebra $\mathcal{A}$ enthalten ist. Wir betrachten reelwertige Funktionen, definiert auf A. Unsere Ziel ist das Integral von solchen Funktionen zu definieren. Hier führen wir den Begriff von Messbarkeit von Funktionen ein, der dann wichtig sein wird, um das Integral zu definieren. Es lohnt sich, Funktionen mit Werten in $[-\infty; \infty]$ zu betrachten (d.h. es lohnt sich die Werten $\pm\infty$ zu erlauben); das wird nämlich nützlich sein, um die Messbarkeit und dann die Integrierbarkeit von Grenzwerten von Funktionenfolge zu untersuchen (ein wichtiger Grund, um das Lebesgue Integral einzuführen ist genau die Tatsache, dass es sich besser bezüglich Vertauschung von Limes und Integrale benimmt, vergliechen mit dem Riemann Integral).

Um den Begriff von messbaren Funktionen auf einem messbaren Raum einzuführen, brauchen wir den folgenden Satz.

Satz 2.1. *Sei $(\Omega, \mathcal{A})$ ein messbarer Raum, $A \in \mathcal{A}$, und $f : A \to [-\infty; +\infty]$. Die folgenden Bedingungen sind dann äquivalent:*

a) Für alle $t \in \mathbb{R}$ gilt $\{x \in A : f(x) \leq t\} \in \mathcal{A}$.

b) Für alle $t \in \mathbb{R}$ gilt $\{x \in A : f(x) < t\} \in \mathcal{A}$.

c) Für alle $t \in \mathbb{R}$ gilt $\{x \in A : f(x) \geq t\} \in \mathcal{A}$.

d) Für alle $t \in \mathbb{R}$ gilt $\{x \in A : f(x) > t\} \in \mathcal{A}$.

Beweis. a) $\Rightarrow$ b): wir bemerken, dass

$$\{x \in A : f(x) < t\} = \bigcup_{n \in \mathbb{N}} \{x \in A : f(x) \leq t - 1/n\}$$

Wenn wir a) annehmen, dann ist jede Menge auf der rechten Seite in $\mathcal{A}$. Damit ist auch die linke Seite, als abzählbar Durchschnitt von Mengen in $\mathcal{A}$ in $\mathcal{A}$ enthalten. Das zeigt b)

b) $\Rightarrow$ c): wir bemerken einfach, dass

$$\{x \in A : f(x) \geq t\} = A \backslash \{x \in A : f(x) < t\}$$

Wenn wir b) annehmen, dann ist $A \backslash \{x \in A : f(x) < t\} \in \mathcal{A}$. Damit ist auch die linke Seite in $\mathcal{A}$ enthalten.

c) $\Rightarrow$ d): ähnlich wie a) $\Rightarrow$ b).

d) $\Rightarrow$ a): ähnlich wie b) $\Rightarrow$ c). □

Definition 2.2. *Sei $(\Omega, \mathcal{A})$ ein messbarer Raum, $A \in \mathcal{A}$. Eine Funktion $f : A \to [-\infty; +\infty]$ heisst messbar (bezüglich der σ-Algebra $\mathcal{A}$), wenn eine der Bedingungen a), b), c), d) in Satz 2.1 erfüllt ist (dann sind alle vier Bedingungen erfüllt).*

Ist $\Omega = \mathbb{R}^n$ und $\mathcal{A} = \mathcal{B}(\mathbb{R}^n)$, so heisst jede $\mathcal{A}$-messbare Funktion Borel messbar oder einfach eine Borel Funktion. Ist $\Omega = \mathbb{R}^n$ und $\mathcal{A} = \mathcal{M}_{\lambda_n^*}$, so heisst jede $\mathcal{A}$-messbare Funktion Lebesgue messbar. Da $\mathcal{B}(\mathbb{R}^n) \subset \mathcal{M}_{\lambda_n^*}$ ist jede Borel Funktion auch Lebesgue messbar.

Bemerke, dass die Messbarkeit einer Funktion nur mit der σ-Algebra zu tun hat. Das Mass spielt hier keine Rolle.

Beispiele: wir diskutieren ein Paar Beispiele von messbaren Funktionen.

i) Sei $(\Omega, \mathcal{A})$ ein messbarer Raum, $A \in \mathcal{A}$ und $f : A \to [-\infty; \infty]$ messbar. Sei nun $B \subset A$. Dann ist die Einschränkung f_B von f auf B genau dann messbar, falls $B \in \mathcal{A}$.

ii) Sei $f : \mathbb{R}^n \to \mathbb{R}$ stetig. Dann ist $\{x \in \mathbb{R}^n : f(x) < t\} = f^{-1}((-\infty; t))$ offen, für alle $t \in \mathbb{R}$. Deswegen ist $\{x \in \mathbb{R}^n : f(x) < t\} \in \mathcal{B}(\mathbb{R}^n)$ für alle $t \in \mathbb{R}$. Also: jede stetige Funktion auf $\mathbb{R}^n$ ist Borel messbar.

iii) Sei $I \subset \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}$ monoton steigend. Dann ist $\{x \in I : f(x) < t\}$ entweder leer oder ein Intervall, enthalten in I. Damit ist $\{x \in I : f(x) < t\} \in \mathcal{B}(\mathbb{R})$, und f eine Borel Funktion. Analog ist jede monoton fallende Funktion Borel messbar.

iv) Sei $f : \mathbb{R} \to \mathbb{R}$ durch $f(x) = 1$, falls $x \in \mathbb{Q}$, und $f(x) = 0$ für $x \in \mathbb{R} \backslash \mathbb{Q}$ definiert. Dann ist

$$\{x \in \mathbb{R} : f(x) > t\} = \begin{cases} \mathbb{R}, & \text{falls } t < 0 \\ \mathbb{Q}, & \text{falls } t \in [0; 1) \\ \emptyset, & \text{falls } t \geq 1 \end{cases}$$

Damit ist $\{x \in \mathbb{R} : f(x) > t\} \in \mathcal{B}(\mathbb{R})$ für alle $t \in \mathbb{R}$, und f ist Borel messbar.

v) Das Beispiel iv) ist ein Beispiel einer charakteristische Funktion. Sei $(\Omega, \mathcal{A})$ ein messbarer Raum, und $A \subset \Omega$. Die Funktion χ_A, definiert auf Ω durch $\chi_A(x) = 1$, für $x \in A$, und $\chi_A(x) = 0$ für $x \notin A$, ist die charakteristische Funktion von A. Es gilt χ_A ist $\mathcal{A}$-messbar g.d.w. $A \in \mathcal{A}$.

Messbarkeit einfache Funktionen. Eine wichtige Verallgemeinerung von charakteristische Funktionen sind so genannte einfache Funktionen. Sei $(\Omega, \mathcal{A})$ ein messbarer Raum, Eine Funktion $f : \Omega \to [-\infty; \infty]$ heisst einfach, wenn sie nur endlich viele Werten annimmt. Sind $\{\alpha_1, \dots, \alpha_k\}$ die Werten von der einfachen Funktion f, dann ist f genau dann $\mathcal{A}$-messbar, falls $\{x \in \Omega : f(x) = \alpha_j\} \in \mathcal{A}$ für alle $j = 1, \dots, k$. Beispiele von einfache Funktionen sind Treppenfunktionen auf $\mathbb{R}$. Wir werden einfache Funktionen zur Definition des Integrals benutzen. Wichtig dabei ist, dass man messbare Funktionen durch einfache Funktionen approximieren kann.

Proposition 2.3. *Sei $(\Omega, \mathcal{A})$ ein messbarer Raum, $A \in \mathcal{A}$, und $f : A \to [0; \infty]$ messbar. Dann es existiert eine Folge $\{f_j\}_{j \in \mathbb{N}}$ so, dass $f_j : A \to [0; \infty)$ eine einfache messbare Funktion ist, für alle $j \in \mathbb{N}$,*

$$f_1(x) \leq f_2(x) \leq \dots \tag{19}$$

und $f(x) = \lim_{n\to\infty} f_n(x)$ für alle $x \in A$.

Bemerkung: Ist $f : A \to [-\infty; \infty]$ messbar und nicht notwendigerweise nicht-negativ, so können wir $f = f_+ - f_-$ zerlegen, mit $f_+ = \max(f, 0)$ und $f_- = -\min(f, 0)$. Wir werden unten zeigen (in Lemma 2.5), dass $f_+, f_- : A \to [0; \infty]$ immer messbar sind, falls f messbar ist. Anwendung von Proposition 2.3 an f_+, f_- liefert eine Folge $\{f_n\}_{n \in \mathbb{N}}$ von messbare einfache Funktionen auf $\mathcal{A}$ (mit Werten in $\mathbb{R}$), so, dass $f(x) = \lim_n f_n(x)$ für alle $x \in A$ (aber hier geht die Monotonie (19) verloren).

Beweis. Für alle $n \in \mathbb{N}$ und $k = 1, \dots, n2^n$ sei

$$A_{n,k} = \left\{ x \in A : \frac{(k-1)}{2^n} \leq f(x) < \frac{k}{2^n} \right\}$$

Da f messbar, ist $A_{n,k} \in \mathcal{A}$. Für $n \in \mathbb{N}$ beliebig, definieren wir $f_n(x) := (k-1)2^{-n}$ für alle $x \in A_{n,k}$ und für alle $k = 1, \dots, n2^n$. Wir setzen weiter $f_n(x) := n$ falls $x \in A \backslash \left[\bigcup_{k=1}^{n2^n} A_{n,k} \right]$ (d.h. falls $f(x) \geq n$). Das definiert f_n auf A. f_n ist offenbar einfach und, nach Konstruktion, messbar für alle $n \in \mathbb{N}$. Wenn wir nun n grösser wählen, dann ist die neue Partition definiert durch die Mengen $A_{n+1,k}$, $k = 1, \dots, (n+1)2^{n+1}$ eine Verfeinerung von die Partition definiert durch $A_{n,k}$. Deswegen ist $f_{n+1}(x) \geq f_n(x)$ für alle $x \in A$. Ist nun $x \in A$ mit $f(x) < \infty$, so ist $x \in \bigcup_{k=1}^{n2^n} A_{n,k}$ für alle n gross genug $(n > f(x))$. Dann ist aber, nach Definition, $|f(x) - f_n(x)| \leq 2^{-n}$ für alle n gross genug, und deswegen $\lim_{n\to\infty} f_n(x) = f(x)$. Ist $f(x) = \infty$, dann gilt $f_n(x) = n$ für alle $n \in \mathbb{N}$; also, auch in diesem Fall, gilt $\lim_{n\to\infty} f_n(x) = f(x)$. □

Eigenschaften messbaren Funktionen. Wir beginnen mit einem Satz, der später nützlich sein wird.

Satz 2.4. *Sei $(\Omega, \mathcal{A})$ ein messbarer Raum, $A \in \mathcal{A}$, $f, g : A \to [-\infty; \infty]$ messbar bezüglich $\mathcal{A}$. Dann gilt*

$$\{x \in A : f(x) < g(x)\}, \{x \in A, f(x) \leq g(x)\}, \{x \in A, f(x) = g(x)\} \in \mathcal{A}\,.$$

Beweis. Wir bemerken, dass

$$\{x \in A : f(x) < g(x)\} = \bigcup_{r \in \mathbb{Q}} \{x \in A : f(x) < r\} \cap \{x \in A : g(x) > r\}$$

Damit ist $\{x \in A : f(x) < g(x)\}$, als abzählbare Vereinigung von Mengen in $\mathcal{A}$ wieder in $\mathcal{A}$. Analog ist $\{x \in A : f(x) > g(x)\} \in \mathcal{A}$ (man ersetzt einfach f und g). Deswegen ist auch

$$A \backslash \{x \in A : f(x) > g(x)\} = \{x \in A : f(x) \leq g(x)\} \in \mathcal{A}$$

und

$$\{x \in A : f(x) = g(x)\} = \{x \in A : f(x) \leq g(x)\} \backslash \{x \in A : f(x) < g(x)\} \in \mathcal{A}$$

□

Das Maximum un das Minimum von zwei messbaren Funktionen ist wieder messbar.

Lemma 2.5. *Sei* $(\Omega, \mathcal{A})$ *ein messbarer Raum,* $A \in \mathcal{A}$*, und* $f, g : A \to [-\infty; \infty]$ *messbar bezüglich* $\mathcal{A}$*. Dann sind auch* $\max(f, g), \min(f, g) : A \to [-\infty; \infty]$*, definiert durch*

$$\max(f, g)(x) := \begin{cases} f(x) & \textit{falls } f(x) \geq g(x) \\ g(x) & \textit{falls } g(x) \geq f(x) \end{cases}$$

und

$$\min(f, g)(x) := \begin{cases} g(x) & \textit{falls } f(x) \geq g(x) \\ f(x) & \textit{falls } g(x) \geq f(x) \end{cases}$$

messbar bezüglich $\mathcal{A}$*.*

Bemerkung: es folgt aus diesem Lemma, dass der positive Teil $f_+ = \max(0, f)$ und der negative Teil $f_- = -\min(f, 0)$ einer messbare Funktion messbar sind.

Beweis. Wir bemerken, dass, für alle $t \in \mathbb{R}$,

$$\{x \in A : \max(f, g)(x) \leq t\} = \{x \in A : f(x) \leq t\} \cap \{x \in A : g(x) \leq t\} \in \mathcal{A}$$

weil f, g messbar sind. Analog gilt

$$\{x \in A : \min(f, g)(x) \geq t\} = \{x \in A : f(x) \geq t\} \cap \{x \in A : g(x) \geq t\} \in \mathcal{A}$$

für alle $t \in \mathbb{R}$. □

Nun betrachten wir Folgen $\{f_n\}_{n \in \mathbb{N}}$ von messbaren Funktionen, definiert auf einer gemeinsamen menge A, mit Werten in $[-\infty; \infty]$. Wir können dann auf A die Funktionen $\sup_n f_n, \inf_n f_n, \limsup_{n \to \infty} f_n, \liminf_{n \to \infty} f_n$ definieren (Erinnerung aus Analysis 1: jede Folge auf $[-\infty; \infty]$ besitzt ein Limes Inferior und ein Limes Superior).

Satz 2.6. *Sei* $(\Omega, \mathcal{A})$ *ein messbarer Raum,* $A \in \mathcal{A}$*. Sei* $f_n : A \to [-\infty; \infty]$ *eine Folge von messbaren Funktionen. Dann:*

i) $\sup_n f_n, \inf_n f_n : A \to [-\infty; \infty]$ *sind messbar.*

ii) $\liminf_{n\to\infty} f_n, \limsup_{n\to\infty} f_n : A \to [-\infty;\infty]$ *sind messbar.*

iii) *Sei* $B = \{x \in A : \liminf_{n\to\infty} f_n(x) = \limsup_{n\to\infty} f_n(x)\}$. *Dann ist* $B \in \mathcal{A}$, *und die Funktion* $\lim_n f_n : B \to [-\infty;\infty]$ *ist messbar.*

Beweis. Wir beginnen mit i). Wir bemerken, dass

$$\{x \in A : \sup_n f_n(x) \le t\} = \bigcap_{n\in\mathbb{N}} \{x \in A : f_n(x) \le t\} \in \mathcal{A}$$

als abzählbare Durchschnitt von Mengen in $\mathcal{A}$. Deshalb ist $\sup_n f_n$ messbar. Analog zeigt man, dass $\inf_n f_n$ messbar ist. Um ii) zu zeigen, bemerken wir, dass

$$\{x \in A : \limsup_{n\to\infty} f_n(x) > t\} = \bigcap_{n\in\mathbb{N}} \bigcup_{k\ge n} \{x \in A : f_k(x) > t\}$$

Damit ist $\{x \in A : \limsup_{n\to\infty} f_n(x) \ge t\} \in \mathcal{A}$ für alle $t \in \mathbb{R}$, und $\limsup_{n\to\infty} f_n$ ist messbar. Analog kann man zeigen, dass $\liminf_{n\to\infty} f_n$ messbar ist. Um iii) zu beweisen, benutzen wir Satz 2.4. Er impliziert, dass

$$B = \{x \in A : \liminf_{n\to\infty} f_n(x) = \limsup_{n\to\infty} f_n(x)\} \in \mathcal{A}$$

Die Messbarkeit von $\lim_{n\to\infty} f_n$ folgt dann, weil

$$\{x \in B : \lim_{n\to\infty} f_n(x) < t\} = B \cap \{x \in A : \liminf_{n\to\infty} f_n(x) < t\} \in \mathcal{A}$$

für alle $t \in \mathbb{R}$. $\square$

Wir zeigen nun, dass Summe und Produkte von messbare Funktionen wieder messbar sind. Wir betrachten hier Funktionen mit Werten in $\mathbb{R}$, um Operationen wie $\infty - \infty$, $0 \cdot \infty$, oder ∞/∞ zu vermeiden.

Satz 2.7. *Sei* $(\Omega, \mathcal{A})$ *ein messbarer Raum,* $A \in \mathcal{A}$, $f, g : A \to \mathbb{R}$ *messbar bezüglich* $\mathcal{A}$ *und* $\alpha \in \mathbb{R}$. *Dann sind* αf, $f + g$, $f - g$, $f \cdot g$, f/g *(definiert auf* $\{x \in A : g(x) \ne 0\}$*) messbar.*

Beweis. Für $\alpha = 0$ ist $\alpha f = 0$ messbar. Für $\alpha > 0$ gilt

$$\{x \in A : \alpha f(x) \le t\} = \{x \in A : f(x) \le t/\alpha\} \in \mathcal{A}$$

für alle $t \in \mathbb{R}$. Analog kann man den Fall $\alpha < 0$ betrachten. Um zu zeigen, dass $f + g$ messbar ist, bemerken wir, dass $(f + g)(x) < t$ genau dann wenn es $r \in \mathbb{Q}$ mit $f(x) < r$ und $g(x) < t - r$ existiert. Wir erhalten

$$\{x \in A : f(x) + g(x) < t\} = \bigcup_{r\in\mathbb{Q}} \{x \in A : f(x) < r\} \cap \{x \in A : g(x) < t - r\} \in \mathcal{A}$$

für alle $t \in \mathbb{R}$. Das zeigt, dass $f + g$ messbar ist. $f - g = f + (-1)g$ ist damit auch messbar.

Wir beweisen nun, dass $f \cdot g$ messbar ist. Dazu zeigen wir, dass, falls $h : A \to [-\infty; \infty]$ messbar ist, dann ist auch $h^2 : A \to [0; \infty]$ messbar. Das folgt, weil

$$\{x \in A : h^2(x) \leq t\} = \{x \in A : h(x) \leq \sqrt{t}\} \cap \{x \in A : h(x) \geq -\sqrt{t}\} \in \mathcal{A}$$

für alle $t \geq 0$, und $\{x \in A : h^2(x) \leq t\} = \emptyset$, falls $t < 0$. Die Messbarkeit von fg folgt nun aus der Bemerkung, dass

$$fg = \frac{1}{2}\left[(f+g)^2 - f^2 - g^2\right] .$$

Schlussendlich, zeigen wir, dass f/g messbar ist. Sei $A_0 = \{x \in A : g(x) \neq 0\}$. Satz 2.4 impliziert, dass $A_0 \in \mathcal{A}$. Für alle $t \in \mathbb{R}$ haben wir

$$\begin{aligned} \{x \in A_0 : f(x)/g(x) < t\} = & [\{x \in A : g(x) > 0\} \cap \{x \in A : f(x) < tg(x)\}] \cup \\ & \cup [\{x \in A : g(x) < 0\} \cap \{x \in A : f(x) > tg(x)\}] \end{aligned}$$

Damit ist $\{x \in A : f(x)/g(x) < t\} \in \mathcal{A}$ aus Satz 2.4. □

Bemerkung: Sei $(\Omega, \mathcal{A})$ ein messbarer Raum, $A \in \mathcal{A}$ und $f, g : A \to [0; \infty]$ (hier erlauben wir f, g den Wert $+\infty$ anzunehmen; wir verlangen aber, dass sie positiv sind). Dann sind, für $\alpha > 0$, $\alpha f, f + g : A \to [0 + \infty]$ messbar. Der Beweis ist wie in Satz 2.7.

Bemerkung: Sei $(\Omega, \mathcal{A})$ ein messbarer Raum, $A \in \mathcal{A}$, und $f : A \to [-\infty; \infty]$. Wir wissen schon, dass der positive Teil $f_+ = \max(f, 0)$ und der negative Teil $f_- = -\min(f, 0)$ von f messbar sind. Aus Satz 2.7 folgt, dass auch $|f| = f_+ + f_-$ messbar ist.

Messbarkeit Funktionen mit Werten in messbaren Raum. Die folgende Proposition erlaubt uns den Begriff von Messbarkeit auf Funktionen mit Werten in einem beliebigen messbaren Raum zu verallgemeinern (umgekehrt, sie erklärt warum Messbarkeit von reellwertige Funktionen so definiert wird, wie wir sie oben definiert haben).

Proposition 2.8. *Sei $(\Omega, \mathcal{A})$ ein messbarer Raum, $A \in \mathcal{A}$, und $f : A \to \mathbb{R}$. Die folgenden Bedingungen sind äquivalent.*

i) f ist messbar bezüglich $\mathcal{A}$.

ii) $f^{-1}(U) \in \mathcal{A}$ für alle $U \subset \mathbb{R}$ offen.

iii) $f^{-1}(C) \in \mathcal{A}$ für alle $C \subset \mathbb{R}$ abgeschlossen.

iv) $f^{-1}(B) \in \mathcal{A}$ für alle $B \in \mathcal{B}(\mathbb{R})$.

Beweis. Wir setzen $\mathcal{F} = \{B \subset \mathbb{R} : f^{-1}(B) \in \mathcal{A}\}$. Wir haben $f^{-1}(\mathbb{R}) = A \in \mathcal{A}$, d.h. $\mathbb{R} \in \mathcal{F}$. Ist $B \in \mathcal{F}$ so gilt $f^{-1}(B) \in \mathcal{A}$. Deswegen ist auch $f^{-1}(B^c) = A \backslash f^{-1}(B) \in \mathcal{A}$ und $B^c \in \mathcal{F}$. Sei nun $\{B_n\}_{n \in \mathbb{N}}$ eine Folge in $\mathcal{F}$. Dann gilt $f^{-1}(B_n) \in \mathcal{A}$ für alle $n \in \mathbb{N}$. Deshalb ist auch

$$f^{-1}\left(\bigcup_{n \in \mathbb{N}} B_n\right) = \bigcup_{n \in \mathbb{N}} f^{-1}(B_n) \in \mathcal{A}$$

und $\bigcup_{n \in \mathbb{N}} B_n \in \mathcal{F}$. Wir haben gezeigt, dass $\mathcal{F}$ eine σ-Algebra ist.

Nach Definition, f ist messbar, falls $(-\infty; b] \in \mathcal{F}$ für alle $b \in \mathbb{R}$. Da diese Intervalle die Borel σ-Algebra $\mathcal{B}(\mathbb{R})$ erzeugen, es folgt, dass f ist messbar genau dann wenn $\mathcal{B}(\mathbb{R}) \subset \mathcal{F}$. Diese Bedingung ist aber auch mit der Bedingung $U \in \mathcal{F}$ für alle U offen und der Bedingung $C \in \mathcal{F}$ für alle C abgeschlossen äquivalent. □

Proposition 2.8 erlaubt uns auch die Definition von Messbarkeit auf Funktionen zu erweitern, die Werten in einem beliebigen messbarer Raum haben. Seien nämlich $(\Omega, \mathcal{A})$ und $(\Omega', \mathcal{A}')$ zwei messbare Räume, $A \in \mathcal{A}$. Eine Funktion $f : A \to \Omega'$ heisst messbar bezüglich $\mathcal{A}$ und $\mathcal{A}'$ falls $f^{-1}(B) \in \mathcal{A}$, für alle $B \in \mathcal{A}'$. Wir werden uns aber nur auf Funktionen mit Werten in $\mathbb{R}$ (oder in der Erweiterung $[-\infty; \infty]$) oder in $\mathbb{C}$ einschränken. Deswegen ist für uns der Begriff von Messbarkeit, definiert am Anfang des Kapitels genug (eine $\mathbb{C}$-wertige Funktion wird für uns messbar sein, falls $\operatorname{Re} f$ und $\operatorname{Im} f$ messbar sind).

Existenz messbaren Mengen, die nicht Borel sind. Als Anwendung von Proposition 2.8 zeigen wir nun, dass $\mathcal{B}(\mathbb{R}) \neq \mathcal{M}_{\lambda^*}$. Dazu konstruieren wir eine besondere Funktion, die bezüglich der σ-Algebra $\mathcal{M}_{\lambda^*}$, aber nicht bezüglich $\mathcal{B}(\mathbb{R})$, messbar ist.

Die Cantor Menge $K = \bigcap_{n \in \mathbb{N}} K_n \subset [0; 1]$ wurde in Sektion 1.4 definiert. Wir definieren nun die Cantor Funktion $f : [0; 1] \to [0; 1]$. Für $x \in K_0 \backslash K_1 = (1/3; 2/3)$ setzen wir $f(x) = 1/2$. Für $x \in K_2 \backslash K_1 = (1/9; 2/9) \cup (7/9; 8/9)$ definieren wir $f(x) = 1/4$ für $x \in (1/9; 2/9)$ und $f(x) = 3/4$ für $x \in (7/9; 8/9)$. Im n-te Schritt wird f auf $K_n \backslash K_{n-1}$ definiert. $K_n \backslash K_{n-1}$ besteht aus 2^{n-1} offene Intervalle der Länge 3^{-n}. Die Funktion f nimmt die Werten $(2k - 1)/2^n$ in der k-te Intervall, für $k = 1, \dots, 2^{n-1}$. Das definiert f auf $[0; 1] \backslash K$. Wir erweitern f auf $[0; 1]$ indem wir $f(0) = 0$ und

$$f(x) = \sup\{f(t) : t \in [0; 1] \backslash K \text{ und } t < x\} \tag{20}$$

für alle $x \in (0; 1]$ setzen. f ist offenbar monoton steigend, mit $f(0) = 0$ und $f(1) = 1$. f ist auch stetig. In der Tat, für x in der offene Menge $(0; 1) \backslash K$ finden wir ein kleines Intervall um x, enthalten in $(0; 1) \backslash K$, wo f konstant (und also stetig) ist. Für $x \in K$, dagegen, muss $x \in K_n$ für alle $n \in \mathbb{N}$ gelten. Das bedeutet, dass x innerhalb einem geschlossenen Intervall der Länge 3^{-n} ist, zwischen zwei offene Intervalle der Länge 3^{-n}, wo f die Werten $(2k - 1)/2^n$ und $(2k + 1)/2^n$ annimmt, für geeignete $k \in \mathbb{N}$. Seien t_1 ein Punkt im linken Intervall (mit $f(t_1) = (2k - 1)/2^n$) und t_2 im rechten Intervall (mit $f(t_2) = (2k + 1)/2^n$). Dann ist $(t_1; t_2)$ ein offenes Intervall um x, mit $|f(y) - f(x)| \leq f(t_2) - f(t_1) = 2^{-n}$ für alle $y \in (t_1; t_2)$. Da $n \in \mathbb{N}$ beliebig ist, folgt, dass f stetig ist.

Es folgt aus der Zwischenwertsatz, dass für alle $y \in [0; 1]$ es existiert mindestens ein $x \in [0; 1]$ mit $f(x) = y$. Wir können also $g : [0; 1] \to [0; 1]$ durch

$$g(y) = \inf\{x \in [0; 1] : f(x) = y\}$$

Es gilt: $f(g(y)) = y$ für alle $y \in [0; 1]$. Das folgt aus der Stetigkeit von f; aus Definition finden wir eine Folge $x_n \in [0; 1]$ mit $f(x_n) = y$ für alle $n \in \mathbb{N}$ und $x_n \to g(y)$ für $n \to \infty$. Die Stetigkeit von f impliziert dann, dass $f(g(y)) = \lim_{n \to \infty} f(x_n) = y$. Die Identität $f(g(y)) = y$ zeigt, dass g injektiv ist (ist $g(y_1) = g(y_2)$, so muss $y_1 = f(g(y_1)) = f(g(y_2)) = y_2$). Da f monoton wachsend ist, ist auch g monoton wachsend (für $y_1 < y_2$ und $\varepsilon > 0$ fest, finden wir $x_1, x_2 \in [0; 1]$ mit $f(x_1) = y_1$ und $f(x_2) = y_2$ und $|g(y_1) - x_1| < \varepsilon/2$ und $|g(y_2) - x_2| < \varepsilon/2$; die Monotonie von f impliziert, dass $x_1 \leq x_2$ und damit, dass $g(y_1) \leq g(y_2) + \varepsilon$. Da $\varepsilon > 0$ beliebig ist, muss $g(y_1) \leq g(y_2)$). Deswegen ist g sicher Borel messbar. Schlussendlich bemerken wir, dass die Werten von g in K sind. Ist nämlich $g(y) \in (0; 1) \backslash K$, dann ist $g(y)$ innerhalb einem offenen Interval, auf welchem f konstant ist. Es ist deswegen möglich $\varepsilon > 0$ zu finden, so, dass $f(g(y) - \varepsilon) = f(g(y)) = y$.

Das widerspricht die Defintion von g ($g(y)$ ist nicht minimal). Wir benutzen nun diese Informationen um die Existenz einer Lebesgue messbare Teilmenge von $[0;1]$ zu zeigen, die nicht in $\mathcal{B}(\mathbb{R})$ enthalten ist.

Satz 2.9. *Es gilt* $\mathcal{B}(\mathbb{R}) \neq \mathcal{M}_{\lambda^*}$.

Beweis. Sei $M \subset [0;1]$ nicht Lebesgue messbar (dessen Existenz in Sektion 1.4 bewiesen worden ist). Sei $A = g(M)$. Dann ist $A \subset K$. Da $\lambda(K) = 0$ ist A sicher Lebesgue messbar, mit $\lambda(A) = 0$ (weil Lebesgue Mass definiert auf $\mathcal{M}_{\lambda^*}$ vollständig ist). Wäre A eine Borel Menge, so wäre auch $g^{-1}(A)$ eine Borel Menge, weil g monoton und deswegen Borel messbar ist. Da aber g inektiv ist, ist $g^{-1}(A) = M$, was nicht Lebesgue messbar und deswegen auch nicht Borel sein kann. (Das bedeutet, dass die monoton wachsende Funktion $g : ([0;1], \mathcal{M}_{\lambda^*}) \to ([0;1], \mathcal{M}_{\lambda^*})$ bezüglich der σ-Algebra $\mathcal{M}_{\lambda^*}$ nicht messbar ist). $\square$

Eigenschaften, die fast überall gelten. Sei $(\Omega, \mathcal{A}, \mu)$ ein Massraum. Wir sagen eine Eigenschaft gilt fast überall auf Ω, falls die Menge $F \subset \Omega$ der Punkten, wo die Eigenschaft nicht gilt in einer Nullmenge enthalten ist. Man nennt eine Menge, die in einer Nullmenge enthalten ist, vernachlässigbar; vernachlässigbare Menge brauchen nicht in $\mathcal{A}$ enthalten zu sein, es genug, dass es $N \subset \Omega$ mit $N \in \mathcal{A}$, $\mu(N) = 0$ und $F \subset N$ existiert (ist $(\Omega, \mathcal{A}, \mu)$ vollständig, so muss dann $F \in \mathcal{A}$, mit $\mu(F) = 0$). Allgemeiner, wir sagen eine Eigenschaft gilt fast überall auf $A \subset \Omega$, falls die Menge aller Punkten in A wo die Eigenschaft nicht gilt vernachlässigbar ist.

Satz 2.10. *Sei* $(\Omega, \mathcal{A}, \mu)$ *ein vollständiger Massraum,* $f, g : \Omega \to [-\infty; \infty]$ *so, dass* $f = g$ *fast überall. Dann ist* f *messbar genau dann wenn* g *messbar ist.*

Beweis. Nehmen wir an f sei messbar und sei $N \in \mathcal{A}$, so, dass $\mu(N) = 0$ und $f = g$ auf N^c. Dann ist

$$\{x \in \Omega : g(x) \leq t\} = [\{x \in \Omega : f(x) \leq t\} \cap N^c] \cup [\{x \in \Omega : g(x) \leq t\} \cap N]$$

Aus Messbarkeit von f ist $\{x \in \Omega : f(x) \leq t\} \cap N^c \in \mathcal{A}$. Aus Vollständigkeit von μ ist $\{x \in \Omega : g(x) \leq t\} \cap N \subset N \in \mathcal{A}$ (weil $\mu(N) = 0$). Das zeigt die Messbarkeit von g. $\square$

2.2 Das Integral

Sei $(\Omega, \mathcal{A})$ ein messbarer Raum. Wir erinnern, dass eine einfache Funktion auf Ω eine Funktion ist, die nur endlich viele Werten annimmt. Die einfache Funktion $f : \Omega \to \mathbb{R}$, die die Werten $\{\alpha_1, \ldots, \alpha_n\}$ annimmt ist genau dann messbar, wenn

$$\{x \in \Omega : f(x) = \alpha_j\} \in \mathcal{A}$$

für alle $j = 1, \ldots, n$. Wir bezeichnen mit $\mathcal{S}$ die Menge von allen messbaren einfachen Funktionen auf Ω. Wir bezeichnen weiter mit $\mathcal{S}_+$ die Menge von allen messbaren einfachen nicht-negativen Funktion auf Ω.

Sei nun μ ein Mass auf $(\Omega, \mathcal{A})$ (d.h. $(\Omega, \mathcal{A}, \mu)$ ist ein Massraum). Sei $f \in \mathcal{S}_+$ gegeben aus $f = \sum_{j=1}^n \alpha_j \chi_{A_j}$ für eine endliche Familie von disjunkten $A_j \in \mathcal{A}$. Wir definieren dann das Integral von f bezüglich μ durch

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

Es ist einfach zu sehen, dass $\int f d\mu$ nur aus f abhängt, und nicht aus der besondere Darstellung von f als Kombination von endlich viele charakteristische Funktionen. Ist nämlich $f = \sum_{j=1}^n \alpha_j \chi_{A_j} = \sum_{i=1}^m \beta_i \chi_{B_i}$, mit disjunkten A_j und B_i, so dass

$$\bigcup_{j=1}^n A_j = \bigcup_{i=1}^m B_i = \Omega$$

dann gilt $A_j = \bigcup_{i=1}^m A_j \cap B_i$ für alle $j = 1, \dots, n$. Ist $A_j \cap B_i \neq \emptyset$ dann muss offenbar $a_j = b_i$ sein. Aus der Additivität von μ, finden wir

$$\begin{aligned}\sum_{j=1}^n a_j \mu(A_j) = \sum_{j=1}^n a_j \sum_{i=1}^m \mu(A_j \cap B_i) &= \sum_{j=1}^n \sum_{i=1}^m a_j \mu(A_j \cap B_i) \\ &= \sum_{i=1}^m \sum_{j=1}^n b_i \mu(A_j \cap B_i) = \sum_{i=1}^m b_i \mu(B_i)\end{aligned}$$

wo wir noch $B_i = \bigcup_{j=1}^n A_j \cap B_i$ benutzt haben.

Wir zeigen ein Paar Eigenschaften von dem Integral von Funktionen in $\mathcal{S}_+$.

Proposition 2.11. *Sei $(\Omega, \mathcal{A}, \mu)$ ein Massraum, $f, g \in \mathcal{S}_+$, $\alpha \geq 0$.*

i) $\alpha f \in \mathcal{S}_+$ und

$$\int \alpha f d\mu = \alpha \int f d\mu$$

ii) $f + g \in \mathcal{S}_+$ und

$$\int (f+g) d\mu = \int f d\mu + \int g d\mu$$

iii) Ist $f(x) \leq g(x)$ für alle $x \in \Omega$, dann gilt

$$\int f d\mu \leq \int g d\mu$$

Beweis. i) Sei $f = \sum_{j=1}^n a_j \chi_{A_j}$, mit $a_j \geq 0$ für alle $j = 1, \dots, n$. Dann ist $\alpha f = \sum_{j=1}^n \alpha a_j \chi_{A_j}$ und

$$\int \alpha f d\mu = \sum_{j=1}^n \alpha a_j \mu(A_j) = \alpha \sum_{j=1}^n a_j \mu(A_j) = \alpha \int f d\mu$$

ii) Ist $f = \sum_{j=1}^n a_j \chi_{A_j}$ und $g = \sum_{i=1}^m b_i \chi_{B_i}$ mit $A_i \cap A_j = \emptyset = B_\ell \cap B_m$ für alle $i \neq j$ und $\ell \neq m$, und mit $\bigcup_{j=1}^n A_i = \bigcup_{i=1}^m B_i = \Omega$ so ist $\{A_i \cap B_j\}$ mit $i = 1, \dots, n$ und

$j = 1, \ldots, m$, eine Partition von Ω, die aus disjunkten Mengen besteht. Die Funktion $f + g$ ist eine einfache Funktion, gegeben aus

$$f + g = \sum_{i=1}^{n} \sum_{j=1}^{m} (a_i + b_j) \chi_{A_i \cap B_j}$$

Deswegen ist

$$\begin{aligned}
\int (f + g) d\mu &= \sum_{i=1}^{n} \sum_{j=1}^{m} (a_i + b_j) \mu(A_i \cap B_j) \\
&= \sum_{i=1}^{n} a_i \sum_{j=1}^{m} \mu(A_i \cap B_j) + \sum_{j=1}^{m} b_j \sum_{i=1}^{n} \mu(A_i \cap B_j) \\
&= \sum_{i=1}^{n} a_i \mu(A_i) + \sum_{j=1}^{m} b_j \mu(B_j) = \int f d\mu + \int g d\mu
\end{aligned}$$

iii) Sind nun $f, g \in \mathcal{S}_+$, mit $f(x) \leq g(x)$ für alle $x \in \Omega$, so ist $f - g$ wieder eine messbare einfache Funktion in $\mathcal{S}_+$. Damit gilt

$$\int g d\mu = \int (f + (g - f)) d\mu = \int f d\mu + \int (g - f) d\mu \geq \int f d\mu$$

weil, aus Definition, das Integral von jeder Funktion in $\mathcal{S}_+$ nicht negativ ist. $\square$

Sei nun $f : \Omega \to [0; \infty]$ eine messbare Funktion auf Ω, bezüglich der σ-Algebra $\mathcal{A}$. Wir definieren das Integral von f durch

$$\int f \, d\mu := \sup \left\{ \int g \, d\mu : g \in \mathcal{S}_+ \text{ und } g \leq f \right\} \tag{21}$$

Satz 2.12. *Sei $(\Omega, \mathcal{A}, \mu)$ ein Massraum, $f : \Omega \to [0; \infty]$ messbar bezüglich $\mathcal{A}$ und sei $\{f_n\}_{n \in \mathbb{N}}$ eine Folge aus $\mathcal{S}_+$ mit $f_{n+1}(x) \geq f_n(x)$ und $f(x) = \lim_{n \to \infty} f_n(x)$ für alle $x \in \Omega$. Dann gilt*

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

Beweis. Aus $f_1(x) \leq f_2(x) \leq \cdots \leq f(x)$ es folgt, dass

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

Es bleibt die umgekehrte Ungleichung zu zeigen. Aus Definition vom Integral von f existiert, zu gegebenen $\varepsilon > 0$, eine Funktion $g \in \mathcal{S}_+$ mit $g \leq f$ und

$$\int f d\mu \leq \int g d\mu + \varepsilon \tag{22}$$

Wir definieren die Folge $g_n = \min(f_n, g)$. Dann ist g_n eine Folge von messbaren einfachen Funktionen, mit $g_n(x) \leq g_{n+1}(x)$ für alle $x \in \Omega$ und $n \in \mathbb{N}$, und $g(x) = \lim_{n \to \infty} g_n(x)$ (weil $g(x) \leq f(x)$ für alle $x \in \Omega$). Wir zeigen, dass

$$\int g d\mu = \lim_{n \to \infty} \int g_n d\mu \, . \tag{23}$$

Dann folgt aus (22), dass

$$\int f d\mu \leq \int g d\mu + \varepsilon = \lim_{n\to\infty} \int g_n d\mu + \varepsilon \leq \lim_{n\to\infty} \int f_n d\mu + \varepsilon$$

weil, nach Definition von g_n, $g_n \leq f_n$ (und g_n, f_n sind beide einfache Funktionen; deswegen können wir Proposition 2.11 anwenden). Da $\varepsilon > 0$ beliebig ist, es folgt, dass

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

Wir zeigen nun (23) für ein $g \in \mathcal{S}_+$ und eine wachsende Folge $g_n \in \mathcal{S}_+$ mit $g_n(x) \to g(x)$ für alle $x \in \Omega$. Sei $g = \sum_{j=1}^m \alpha_j \chi_{A_j}$ mit disjunkten A_j und $\varepsilon > 0$. Für $n \in \mathbb{N}$, $j = 1, \dots, m$ setzen wir

$$A_{n,j} = \{x \in A_j : g_n(x) \geq (1-\varepsilon)\alpha_j\}$$

Dann ist $A_{n,j} \in \mathcal{A}$ für alle n und j. Für feste j, die Folge $A_{n,j}$ ist wachsend in n, d.h. $A_{n+1,j} \supset A_{n,j}$ und $\bigcup_{n\in\mathbb{N}} A_{n,j} = A_j$ (weil $g_n \to g$, punktweise). Wir setzen $h_n := \sum_{j=1}^m (1-\varepsilon)\alpha_j \chi_{A_{n,j}}$.Dann ist $h_n \in \mathcal{S}_+$, mit $h_n \leq g_n$ und also

$$\begin{aligned}\lim_{n\to\infty} \int g_n d\mu \geq \lim_{n\to\infty} \int h_n d\mu &= (1-\varepsilon) \lim_{n\to\infty} \sum_{j=1}^m \alpha_j \mu(A_{n,j}) \\ &= (1-\varepsilon) \sum_{j=1}^m \alpha_j \mu(A_j) = (1-\varepsilon) \int g d\mu\end{aligned}$$

Da $\varepsilon > 0$ beliebig ist, finden wir (23). □

Wir erweitern nun die Eigenschaften, die wir in Proposition 2.11 für Funktionen in $\mathcal{S}_+$ bewiesen haben, zu nicht-negativen messbaren Funktionen.

Proposition 2.13. *Sei $(\Omega, \mathcal{A}, \mu)$ ein Massraum. Seien $f, g : \Omega \to [0; \infty]$ messbar bezüglich $\mathcal{A}$, und $\alpha > 0$.*

i) $\alpha f : \Omega \to [0; \infty]$ ist messbar und

$$\int \alpha f d\mu = \alpha \int f d\mu$$

ii) $f + g : \Omega \to [0; \infty]$ ist messbar und

$$\int (f+g) d\mu = \int f d\mu + \int g d\mu$$

iii) Ist $f(x) \leq g(x)$ für alle $x \in \Omega$, so gilt

$$\int f d\mu \leq \int g d\mu$$

Beweis. Wir wissen aus Proposition 2.3, dass es wachsende Folgen $f_n, g_n \in \mathcal{S}_+$ existieren, mit $f = \lim_{n\to\infty} f_n$ und $g = \lim_{n\to\infty} g_n$. Dann ist auch die Folge αf_n in $\mathcal{S}_+$ und wachsend, mit $\alpha f = \lim_{n\to\infty} \alpha f_n$. Deswegen, Satz 2.12 (zusammen mit Proposition 2.11) impliziert, dass

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

Auch $f_n + g_n \in \mathcal{S}_+$ ist eine wachsend Folge, mit $f + g = \lim_{n\to\infty} f_n + g_n$. Deswegen:

$$\begin{aligned} \int (f+g) d\mu = \lim_{n\to\infty} \int (f_n + g_n) d\mu &= \lim_{n\to\infty} \int f_n d\mu + \int g_n d\mu \\ &= \lim_{n\to\infty} \int f_n d\mu + \lim_{n\to\infty} \int g_n d\mu = \int f d\mu + \int g d\mu \,. \end{aligned}$$

Um Teil iii) zu zeigen, bemerken wir einfach aus Definition in (21), dass die Menge der Funktionen $h \in \mathcal{S}_+$ mit $h \leq g$ grösser ist, als die Menge der Funktionen $h \in \mathcal{S}_+$ mit $h \leq f$. Deswegen ist das Supremum grösser, d.h.

$$\int g d\mu \geq \int f d\mu \,.$$

□

Definition 2.14. *Sei* $(\Omega, \mathcal{A}, \mu)$ *ein Massraum, und* $f : \Omega \to [-\infty; \infty]$ *messbar bezüglich* $\mathcal{A}$. *Wir setzen* $f_+ = \max(f, 0)$ *und* $f_- = -\min(f, 0)$. *Die Funktionen* f_+, f_- *sind messbar, aus Lemma 2.5. Wir sagen, dass* f *ist integrierbar, falls*

$$\int f_+ d\mu, \int f_- d\mu < \infty$$

In diesem Fall definieren wir das Integral von f *durch*

$$\int f d\mu = \int f_+ d\mu - \int f_- d\mu \tag{24}$$

Bemerkung:

i) Wir sagen, das Integral von f existiert, falls $\int f_+ d\mu < \infty$ oder $\int f_- d\mu$ endlich ist (sind sie beide endlich, so ist f integrierbar). In diesem Fall können wir auch die Definition (21) für das Integral anwenden (das Integral darf dann plus oder minus Unendlich sein).

ii) Sei $A \in \mathcal{A}$. Wir sagen f ist integrierbar über A, falls $f\chi_A$ integrierbar ist. In diesem Fall definieren wir das Integral von f über A durch

$$\int_A f d\mu = \int f \chi_A d\mu \,.$$

iii) Ist $f : A \to [-\infty; \infty]$ so heisst f integrierbar über A, falls die Funktion $\widetilde{f} : \Omega \to [-\infty; \infty]$ definiert durch $\widetilde{f}(x) = f(x)$ für $x \in A$ und $\widetilde{f}(x) = 0$ für $x \notin A$ integrierbar ist. In diesem Fall ist

$$\int_A f d\mu = \int \widetilde{f} d\mu \,.$$

iv) Ist $\Omega = \mathbb{R}^n$, $\mathcal{A} = \mathcal{M}_{\lambda_n^*}$ oder $\mathcal{A} = \mathcal{B}(\mathbb{R}^n)$ und $\mu = \lambda_n$ das Lebesgue Mass, dann heisst eine integrierbare Funktion einfach Lebesgue integrierbar, ihr Integral $\int f d\lambda_n$ das Lebesgue Integral. Wir schreiben manchmal einfach $\int f dx$ für das Lebesgue Integral.

Beispiele:

1) Sei Ω eine Menge, und $x_0 \in \Omega$. Sei $(\Omega, P(\Omega)), \mu)$ der Massraum, mit $\mu(A) = 1$ falls $x_0 \in A$, $\mu(A) = 0$ falls $x_0 \notin A$. Sei $f : \Omega \to [-\infty; \infty]$ beliebig. Dann ist f messbar, und wir behaupten, dass

$$\int f d\mu = f(x_0) \tag{25}$$

Sei zunächst $f : \Omega \to [0; \infty]$, und $g \in \mathcal{S}_+$, mit $g(x) \leq f(x)$ für alle $x \in \Omega$. Wir schreiben $g = \sum_{j=1}^n \alpha_j \chi_{A_j}$ für eine Familie disjunkten Mengen $A_j \subset \Omega$, mit $\bigcup_{j=1}^n A_j = \Omega$. Dann existiert $j_0 \in \{1, \dots, n\}$ mit $x_0 \in A_{j_0}$ und

$$\int g d\mu = \sum_{j=1}^n \alpha_j \mu(A_j) = \alpha_{j_0} = g(x_0) \leq f(x_0)$$

Anderseits, wenn wir die einfache Funktion $g(x) = f(x_0)$ für $x = x_0$ und $g(x) = 0$ sonst, wir haben

$$\int f d\mu \geq \int g d\mu = f(x_0)\mu(\{x_0\}) = f(x_0)$$

Man kann (25) für allgemeinen $f : \Omega \to [-\infty; \infty]$ zeigen, indem man $f = f_+ - f_-$ zerlegt.

2) Sei $(\Omega, \mathcal{A}, \mu)$ ein Massraum, mit μ ein endliches Mass. Dann ist jede bezüglich $\mathcal{A}$ messbare, beschränkte Funtkion $f : \Omega \to [-\infty; \infty]$ automatisch integrierbar. In der Tat, falls $g : \Omega \to [0; \infty]$ messbar und beschränkt ist, finden wir

$$\int g d\mu \leq C \int d\mu = C\mu(\Omega) < \infty$$

Ist nun $f : \Omega \to [-\infty; \infty]$ messbar und beschränkt, so sind $f_+, f_- : \Omega \to [0; \infty]$ messbar und beschränkt, und deswegen

$$\int f_+ d\mu, \int f_- d\mu < \infty$$

3) Sei $(\Omega, \mathcal{A}, \mu)$ ein Massraum, $A \in \mathcal{A}$ mit $\mu(A) = 0$, und $f : \Omega \to [-\infty; \infty]$ messbar mit $f(x) = 0$ für alle $x \notin A$. Dann ist

$$\int f d\mu = 0$$

Das gilt auch falls, zum Beispiel $f(x) = \infty$ für alle $x \in A$. Es genug die Behauptung für $f : \Omega \to [0; \infty]$ zu zeigen. Unter Annahme, dass $f(x) = 0$ für alle $x \in A^c$ gilt

$f \leq g$ wobei g die einfache Funktion mit $g(x) = \infty$ für $x \in A$ und $g(x) = 0$ für $x \notin A$. Also

$$\int f d\mu \leq \int g d\mu = \sup \left\{ \int h d\mu : h \in \mathcal{S}_+, h \leq g \right\} = 0$$

weil für jede Funktion $h \in \mathcal{S}_+$ mit $h(x) = 0$ für alle $x \notin A$ ist $\int h d\mu = 0$.

4) Sei λ das Lebesgue Mass auf $(\mathbb{R}, \mathcal{M}_{\lambda^*})$. Es gilt

$$\int \chi_{\mathbb{Q}} d\lambda = 0$$

Das folgt aus Beispiel 3).

Wir definieren den Raum

$$\mathcal{L}^1(\Omega, \mathcal{A}, \mu) = \{ f : \Omega \to \mathbb{R} : f \text{ ist auf } \Omega \text{ bezüglich } \mathcal{A} \text{ und } \mu \text{ integrierbar} \} .$$

Wir zeigen im nächsten Satz, dass $\mathcal{L}^1$ ein Vektorraum ist.

Satz 2.15. *Sei $(\Omega, \mathcal{A}, \mu)$ ein Massraum, $f, g \in \mathcal{L}^1(\Omega, \mathcal{A}, \mu)$, und $\alpha \in \mathbb{R}$. Dann gilt*

i) αf ist integrierbar und

$$\int \alpha f d\mu = \alpha \int f d\mu$$

ii) $f + g$ ist integrierbar und

$$\int (f+g) d\mu = \int f d\mu + \int g d\mu$$

iii) Ist $f \leq g$, so gilt

$$\int f d\mu \leq \int g d\mu$$

Beweis. i) Sei $\alpha > 0$, dann gilt $(\alpha f)_+ = \alpha f_+$ und $(\alpha f)_- = \alpha f_-$. Integrierbarkeit von f bedeutet, dass $\int f_+ d\mu$ und $\int f_- d\mu$ endlich sind. Dann sind auch

$$\int (\alpha f)_+ d\mu = \int \alpha f_+ d\mu = \alpha \int f_+ d\mu$$

und

$$\int (\alpha f)_- d\mu = \int \alpha f_- d\mu = \alpha \int f_- d\mu$$

endlich. Damit ist αf integrierbar, und

$$\int \alpha f d\mu = \int (\alpha f)_+ d\mu - \int (\alpha f)_- d\mu = \alpha \left[\int f_+ d\mu - \int f_- d\mu \right] = \alpha \int f d\mu$$

Analog kann man den Fall $\alpha < 0$ behandeln (in diesem Fall ist $(\alpha f)_+ = -\alpha f_-$ und $(\alpha f)_- = -\alpha f_+$). Der Fall $\alpha = 0$ ist trivial.

ii) Wir bemerken, dass $(f+g)_+ \leq f_+ + g_+$ und $(f+g)_- \leq f_- + g_-$. Deswegen gilt (aus Integrierbarkeit von f, g),

$$\int (f+g)_+ d\mu \leq \int (f_+ + g_+) d\mu = \int f_+ d\mu + \int g_+ d\mu < \infty$$

und

$$\int (f+g)_- d\mu \leq \int (f_- + g_-) d\mu = \int f_- d\mu + \int g_- d\mu < \infty$$

Also, $f+g$ ist integrierbar. Nun bemerken wir, dass $f+g = (f+g)_+ - (f+g)_-$ auch als

$$f + g = f_+ + g_+ - (f_- + g_-)$$

geschrieben werden kann (durch einfache Überprüfung der Fällen a) $f(x) \geq 0, g(x) \geq 0$, b) $f(x) \geq 0, g(x) < 0$, c) $f(x) < 0, g(x) \geq 0$ und d) $f(x) < 0, g(x) < 0$). Deswegen gilt

$$(f+g)_+ + (f_- + g_-) = f_+ + g_+ + (f+g)_-$$

und (weil alle Summanden links und recht positiv sind, können wir Proposition 2.13 anwenden)

$$\int (f+g)_+ d\mu + \int f_- d\mu + \int g_- d\mu = \int f_+ d\mu + \int g_+ d\mu + \int (f+g)_- d\mu$$

Das gibt

$$\begin{aligned} \int (f+g) d\mu &= \int (f+g)_+ d\mu - \int (f+g)_- d\mu \\ &= \int f_+ d\mu + \int g_+ d\mu - \int f_- d\mu - \int g_- d\mu = \int f d\mu + \int g d\mu \end{aligned}$$

iii) Ist $f(x) \leq g(x)$ für alle $x \in \Omega$, so ist $g - f$ eine nicht-negative integrierbare Funktion (aus i) und ii)). Deswegen ist, wieder wegen i) und ii),

$$\int g d\mu - \int f d\mu = \int (g - f) d\mu \geq 0$$

□

Es gilt auch eine Dreiecksungleichung für Integrale.

Satz 2.16. *Sei $(\Omega, \mathcal{A}, \mu)$ ein Massraum, $f : \Omega \to [-\infty; \infty]$ messbar bezüglich $\mathcal{A}$. Dann ist f genau dann integrierbar, wenn $|f|$ integrierbar ist. In diesem Fall gilt*

$$\left| \int f d\mu \right| \leq \int |f| d\mu$$

Beweis. f ist genau dann integrierbar, wenn

$$\int f_+ d\mu, \int f_- d\mu < \infty$$

Da $|f| = f_+ + f_-$, diese Bedingung ist mit der Integrierbarkeit von $|f|$ äquivalent. Weiter gilt

$$\left|\int f d\mu\right| = \left|\int f_+ d\mu - \int f_- d\mu\right| \leq \int f_+ d\mu + \int f_- d\mu = \int (f_+ + f_-) d\mu = \int |f| d\mu$$

□

Lemma 2.17. *Sei $(\Omega, \mathcal{A}, \mu)$ ein Massraum, $f, g : \Omega \to [-\infty; \infty]$ messbare Funktionen, mit $f(x) = g(x)$ fast überall. Dann ist f integrierbar, genau dann wenn g integrierbar ist, und in diesem Fall gilt*

$$\int f d\mu = \int g d\mu$$

Beweis. Seien zunächst $f, g : \Omega \to [0; \infty]$ messbar, mit $f(x) = g(x)$ fast überall. Sei weiter $A = \{x \in \Omega : f(x) \neq g(x)\}$ und definiere $h(x) = 0$ für alle $x \in A^c$ und $h(x) = \infty$ für $x \in A$. Wie in Beispiel 3) nach Definition 2.14 finden wir, dass $\int h d\mu = 0$. Da $f(x) \leq g(x) + h(x)$, finden wir

$$\int f d\mu \leq \int g d\mu + \int h d\mu = \int g d\mu$$

Analog kann man die umgekehrte Ungleichung beweisen (ersetze f und g). Das zeigt, dass

$$\int f d\mu = \int g d\mu$$

und insbesondere, dass f genau dann integrierbar ist, wenn g ist. Für allgemeine f, g schreiben wir $f = f_+ - f_-$ und $g = g_+ - g_-$. Es gilt $f_+ = g_+$ und $f_- = g_-$ fast überall. Deswegen ist

$$\int f_+ d\mu = \int g_+ d\mu \quad \text{und} \quad \int f_- d\mu = \int g_- d\mu$$

Das bedeutet, dass f genau dann integrierbar ist, wenn g integrierbar und, dass

$$\int f d\mu = \int g d\mu$$

□

Im nächsten Lemma zeigen wir eine Ungleichung mit vielen wichtigen Anwendungen. In der Wahrscheinlichkeitstheorie, ist diese Ungleichung als die Chebyshev Ungleichung bekannt.

Lemma 2.18. *Sei $(\Omega, \mathcal{A}, \mu)$ ein Massraum, $f : \Omega \to [0; \infty]$ messbar. Für $t > 0$ es gilt*

$$\mu\left(\{x \in \Omega : f(x) \geq t\}\right) \leq \frac{1}{t} \int f d\mu$$

Beweis. Sei $A_t = \{x \in \Omega : f(x) \geq t\}$. Dann gilt $f \geq f\chi_{A_t} \geq t\chi_{A_t}$. Deswegen

$$\int f d\mu \geq \int f\chi_{A_t} d\mu \geq \int t\chi_{A_t} d\mu = t\mu(A_t)\,.$$

□

Das letzte Lemma hat vielen wichtige Folgerungen.

Korollar 2.19. *Sei* $(\Omega, \mathcal{A}, \mu)$ *ein Massraum, und* $f : \Omega \to [-\infty; \infty]$ *messbar, mit*

$$\int |f| \, d\mu = 0$$

Dann ist $f = 0$ *fast überall.*

Proof. Lemma 2.18 für $|f|$ impliziert, dass

$$\mu\left(\{x \in \Omega : |f(x)| \geq 1/n\}\right) \leq n \int |f| d\mu = 0$$

für alle $n \in \mathbb{N}$. Deswegen ist

$$\begin{aligned} \mu\left(\{x \in \Omega : f(x) \neq 0\}\right) &= \mu\left(\bigcup_{n\in\mathbb{N}} \{x \in \Omega : |f(x)| \geq 1/n\}\right) \\ &\leq \sum_{n\in\mathbb{N}} \mu(\{x \in \Omega : |f(x)| \geq 1/n\}) = 0 \end{aligned}$$

□

Korollar 2.20. *Sei* $(\Omega, \mathcal{A}, \mu)$ *ein Massraum, und* $f : \Omega \to [-\infty; \infty]$ *messbar, mit*

$$\int_A f d\mu \geq 0$$

für alle $A \in \mathcal{A}$. *Dann ist* $f \geq 0$ *fast überall.*

Beweis. Sei $A = \{x \in \Omega : f(x) < 0\}$. Dann gilt

$$0 = \int_A f d\mu = \int f \chi_A d\mu = -\int |f \chi_A| d\mu$$

Aus Korollar 2.19 folgt, dass $f\chi_A = 0$ fast überall, und damit, dass $f(x) \geq 0$ fast überall. □

Korollar 2.21. *Sei* $(\Omega, \mathcal{A}, \mu)$ *ein Massraum, und* $f : \Omega \to [-\infty; \infty]$ *integrierbar. Dann ist* $|f(x)| < \infty$ *fast überall.*

Beweis. Aus Lemma 2.18 finden wir

$$\mu(\{x \in \Omega : |f(x)| > n\}) \leq \frac{1}{n} \int |f| d\mu$$

Deswegen

$$\mu(\{x \in \Omega : |f(x)| = \infty\}) \leq \frac{1}{n} \int |f| d\mu$$

für alle n. Das zeigt, dass $\mu(\{x \in \Omega : |f(x)| = \infty\}) = 0$. □

Korollar 2.22. *Sei* $(\Omega, \mathcal{A}, \mu)$ *ein Massraum,* $f : \Omega \to [-\infty; \infty]$ *messbar. Dann ist* f *integrierbar genau dann, wenn ein* $g \in \mathcal{L}^1(\Omega, \mathcal{A}, \mu)$ *(also mit Werten in* $\mathbb{R}$*) existiert so, dass* $f = g$ *fast überall.*

Beweis. Existiert $g \in \mathcal{L}^1(\Omega, \mathcal{A}, \mu)$ mit $f = g$ fast überall, dann ist f integrierbar aus Lemma 2.17. Ist dagegen f integrierbar, dann setzen wir $A = \{x \in \Omega : |f(x)| = \infty\}$. Dann ist $A \in \mathcal{A}$, mit $\mu(A) = 0$ aus Korollar 2.21. Wir setzen $g = f\chi_{A^c}$. Dann gilt $g = f$ fast überall, g hat Werten in $\mathbb{R}$ und, aus Lemma 2.17 ist $g \in \mathcal{L}^1(\Omega, \mathcal{A}, \mu)$. □

2.3 Konvergenzsätze

Ein wichtiges Vorteil vom neuen Integralbegriff ist das Verhalten bezüglich Vertausch von Integral und Limites. In Analysis 1 haben wir gezeigt: ist $\{f_n\}_{n\in\mathbb{N}}$ eine Folge von Funktionen auf einem Intervall $[a;b]$, mit $f_n \to f$ gleichmässig, dann konvergiert das (Riemann) Integral von f_n über $[a;b]$ gegen das Integral von f. In diesem Kapitel möchten wir Theoremen zeigen, die uns erlauben werden Integral und Limes in allgemeinere Situationen und unter schwächere Annahme zu vertauschen.

Wir haben in Satz 2.12 schon ein erstes Resultat in dieser Richtung gezeigt. Dort haben wir angenommen, dass $\{f_n\}$ eine Folge in von einfachen messbaren Funktionen ist. Im folgenden Satz verallgemeinern wir Satz 2.12 zu dem Fall, dass $\{f_n\}$ nicht unbedingt einfach ist.

Satz 2.23 (Satz von der monotonen Konvergenz)**.** *Sei $(\Omega, \mathcal{A}, \mu)$ ein Massraum, $\{f_n\}_{n\in\mathbb{N}}$ eine Folge von messbaren Funktionen $f_n : \Omega \to [0;\infty]$ und $f : \Omega \to [0;\infty]$ eine messbare Funktion so, dass*

i) $f_n(x) \le f_{n+1}(x)$ *für alle $n \in \mathbb{N}$ and*

ii) $\lim_{n\to\infty} f_n(x) = f(x)$

für μ-fast alle $x \in \Omega$. Dann gilt

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

Bemerkung: die Funktionen f_n brauchen nicht integrierbar zu sein. Das Integral ist trotzdem wohldefiniert, weil die Funktionen nicht negativ sind.

Beweis. Wir nehmen zunächst an, dass die Annahme i) und ii) für alle $x \in \Omega$ gelten. Aus der Monotonie des Integrals ist $\int f_n d\mu$ eine monotone (und deswegen konvergente, möglicherweise mit unendlichen Grenzwert) Folge auf $\mathbb{R}$, mit

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

Aus Proposition 2.3 finden wir für alle $n \in \mathbb{N}$ eine Folge $\{g_{n,j}\}_{j\in\mathbb{N}}$ in $\mathcal{S}_+$ mit $g_{n,j}(x) \le g_{n,j+1}(x)$ für alle $j \in \mathbb{N}$ und $f_n(x) = \lim_{j\to\infty} g_{n,j}(x)$ für alle $x \in \Omega$. Wir definieren dann

$$h_n(x) = \max(g_{1,n}(x), g_{2,n}(x), \dots, g_{n,n}(x))$$

Wir finden: $h_n \in \mathcal{S}_+$, $h_n(x) \le f_n(x)$ und $h_n(x) \le h_{n+1}(x)$ für alle $n \in \mathbb{N}$. Ferner, $\lim_{n\to\infty} h_n(x) = f(x)$. Seien, in der Tat, $x \in \Omega$ und $\varepsilon > 0$ festgewählt. Da $f_n(x) \to f(x)$ für $n \to \infty$, finden wir $n_0 \in \mathbb{N}$ gross genug, mit $f_{n_0}(x) \ge f(x) - \varepsilon/2$. Da $g_{n_0,j}(x) \to f_{n_0}(x)$ für $j \to \infty$, finden wir auch $m_0 \in \mathbb{N}$ mit $g_{n_0,m_0}(x) \ge f_{n_0}(x) - \varepsilon/2 \ge f(x) - \varepsilon$. Da $g_{n_0,j}(x)$ wachsend in j ist, können wir annehmen, dass $m_0 \ge n_0$. Dann gilt $h_{m_0}(x) \ge g_{n_0,m_0}(x) \ge f(x) - \varepsilon$. Aus der Monotonie der Folge $\{h_n\}$, finden wir

$$f(x) - \varepsilon \le h_n(x) \le f_n(x) \le f(x)$$

für alle $n > m_0$. Das zeigt, dass $h_n(x) \to f(x)$ für $n \to \infty$. Aus Satz 2.12 und aus Monotonie des Integrals schliessen wir, dass

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

Wenn wir nun annehmen, dass i) und ii) nur μ-fast überall (statt überall) gelten, dann finden wir eine Menge $N \in \mathcal{A}$ mit $\mu(N) = 0$ und so, dass i) und ii) überall auf N^c gelten. Dann ist $f_n \chi_{N^c}$ eine monotone Folge von messbare Funktionen auf Ω, mit $\lim_{x\to\infty} f_n \chi_{N^c}(x) = f\chi_{N^c}(x)$ für alle $x \in \Omega$. Der erste Teil des Beweises impliziert, dass

$$\lim_{n\to\infty} \int f_n \chi_{N^c} d\mu = \int f \chi_{N^c} d\mu$$

Da aber $f_n\chi_{N^c} = f_n$ μ-fast überall (für alle $n \in \mathbb{N}$) und $f\chi_{N^c} = f$ μ-fast überall, es folgt aus Lemma 2.17, dass

$$\int f d\mu = \int f\chi_{N^c} d\mu \quad \text{und} \quad \int f_n d\mu = \int f_n \chi_{N^c} d\mu$$

für alle $n \in \mathbb{N}$, und deswegen, dass

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

□

Eine äquivalente Formulierung von dem Satz der monotonen Konvergenz ist aus dem folgenden Korollar gegeben.

Korollar 2.24 (Theorem von Beppo Levi). *Sei* $(\Omega, \mathcal{A}, \mu)$ *ein Massraum, und* $\{f_n\}_{n\in\mathbb{N}}$ *eine Folge von messbaren Funktionen* $f_n : \Omega \to [0; \infty]$. *Dann gilt*

$$\int \sum_{k=1}^{\infty} f_k \, d\mu = \sum_{k=1}^{\infty} \int f_k d\mu$$

Beweis. Wir definieren einfach $g_n(x) = \sum_{k=1}^n f_k(x)$, und bemerken, dass g_n eine monotone Folge von messbaren Funktionen auf Ω ist. Aus Monotonie, existiert der Limes $g(x) = \lim_{n\to\infty} \sum_{k=1}^n f_k(x) = \sum_{k=1}^\infty f_k(x)$ für alle $x \in \Omega$. Anwendung von Satz 2.23 auf der Folge g_n zeigt die Behauptung. □

Satz 1.17 zeigt die Stetigkeit des Integrals bezüglich monotone Konvergenz von messbaren Funktionen mit Werten in $[0; \infty]$. Die Annahme der Monotonie der Folge f_n in Satz 2.23 ist wichtig. Auch ohne Monotonie finden wir aber, dass das Integral unterhalbstetig ist.

Lemma 2.25 (Lemma von Fatou). *Sei* $(\Omega, \mathcal{A}, \mu)$ *ein Massraum und* $\{f_n\}_{n\in\mathbb{N}}$ *eine Folge von messbaren Funktionen* $f_n : \Omega \to [0; \infty]$. *Dann*

$$\liminf_{n\to\infty} \int f_n d\mu \geq \int \liminf_{n\to\infty} f_n d\mu$$

Beweis. Für $n \in \mathbb{N}$ definieren wir $g_n = \inf_{k \geq n} f_k$. g_n ist messbar aus Satz 2.6. Ferner es gilt $g_n(x) \leq g_{n+1}(x)$ für alle $n \in \mathbb{N}$ und

$$\lim_{n \to \infty} g_n(x) = \liminf_{n \to \infty} f_n(x)$$

nach Definition. Satz 2.23 impliziert, dass

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

weil $g_n \leq f_n$ für alle $n \in \mathbb{N}$. □

Für Anwendungen, der wichtigste Konvergenzsatz ist der folgenden Satz von Lebesgue.

Satz 2.26 (Satz über dominierte Konvergenz)**.** *Sei $(\Omega, \mathcal{A}, \mu)$ ein Massraum, $f : \Omega \to [-\infty; \infty]$ messbar, $f_n : \Omega \to [-\infty; \infty]$ eine Folge messbaren Funktionen und $g : \Omega \to [0; \infty]$ integrierbar. Es gelte*

i) $f(x) = \lim_{n \to \infty} f_n(x)$ *und*

ii) $|f_n(x)| \leq g(x)$,

für μ-fast alle $x \in \Omega$. Dann gilt

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

Bemerkung: die Existenz einer nicht-negative und integrierbare Funktion g, die alle f_n dominiert ist die wesentliche Annahme von diesem Satz, die auch als Satz von der majorisierten Konvergenz bekannt ist.

Beweis. Wir nehmen zuerst an, dass die Annahmen i), ii) und $g(x) < \infty$ für alle $x \in \Omega$ gelten. Die Integrierbarkeit von f_n und f folgt aus der Integrierbar von g ($|f_n(x)| \leq g(x)$ für alle $n \in \mathbb{N}$ und $f_n(x) \to f(x)$ für $n \to \infty$ implizieren auch, dass $|f(x)| \leq g(x)$). Wir setzen $h_n = f_n + g$. Dann ist $h_n : \Omega \to [0; \infty]$ eine Folge messbare Funktionen, mit $(f + g)(x) = \lim_{n \to \infty} h_n(x)$ für alle $x \in \Omega$. Aus Lemma 2.25 folgt, dass

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

und deswegen, dass

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

Anderseits, $k_n = g - f_n$ definiert eine Folge von messbaren Funktionen $k_n : \Omega \to [0; \infty]$ mit $\lim_{n \to \infty} k_n(x) = (g - f)(x)$. Wieder aus Lemma 2.25 finden wir

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

und deswegen, dass

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

Es folgt, dass
$$\int f d\mu = \lim_{n\to\infty} \int f_n d\mu$$

Nehmen wir nun an, dass die Annahme i) und ii) und die Ungleichung $g(x) < \infty$ nur μ-fast überall (statt überall) gelten (die Integrierbarkeit von g impliziert, dass $g(x) < \infty$ fast überall gelten muss). Sei $N \in \mathcal{A}$ mit $\mu(N) = 0$ und so, dass $|f_n(x)| \leq g(x)$ für alle $n \in \mathbb{N}$, $\lim_{n\to\infty} f_n(x) = f(x)$ und $g(x) < \infty$ für alle $x \in N^c$. Dann ist $g\chi_{N^c}$ eine integrierbare Funktion, $f\chi_{N^c}$ eine messbare Funktion und $f_n\chi_{N^c}$ eine Folge messbarer Funktionen, mit $|f_n\chi_{N^c}(x)| \leq g\chi_{N^c}(x)$ für alle $n \in \mathbb{N}$ und $\lim_{n\to\infty} f_n\chi_{N^c}(x) = f\chi_{N^c}(x)$ für alle $x \in \Omega$. Nach dem ersten Teil des Beweises, sind $f\chi_{N^c}$ und $f_n\chi_{N^c}$ integrierbar für alle $n \in \mathbb{N}$ und
$$\lim_{n\to\infty} \int f_n\chi_{N^c} d\mu = \int f\chi_{N^c} d\mu \tag{26}$$
Aus Lemma 2.17, die Integrierbarkeit von $f\chi_{N^c}$ und von $f_n\chi_{N^c}$ impliziert auch, dass f_n und f integrierbar sind und, dass
$$\int f_n d\mu = \int f_n\chi_{N^c} d\mu \quad \text{und} \quad \int f d\mu = \int f\chi_{N^c} d\mu$$
Mit (26) finden wir, dass
$$\lim_{n\to\infty} \int f_n d\mu = \int f d\mu$$
□

2.4 Vergleich mit dem Riemann'schen Integral

Eine wichtige Eigenschaft vom Lebesgue'schen Integral ist die Tatsache, dass es mit dem Riemann'schen Integral übereinstimmt, falls das Riemann'sche Integral existiert. Wir betrachten in diesem Kapitel den Massraum $(\mathbb{R}, \mathcal{M}_{\lambda^*}, \lambda)$, wobei λ das Lebesgue Mass auf $\mathbb{R}$ ist. Wir sagen eine Funktion f definiert auf $\mathbb{R}$ oder auf einer Teilmenge von $\mathbb{R}$ ist Lebesgue integrierbar, wenn sie als Funktion auf dem Massraum $(\mathbb{R}, \mathcal{M}_{\lambda^*}, \lambda)$ integrierbar ist. Wir bezeichnen mit $\int_a^b f(x)dx$ das Riemann Integral von f über $[a;b]$ und mit $\int_{[a;b]} f d\lambda$ das Lebesgue Integral von f uber $[a;b]$.

Satz 2.27. *Sei $-\infty < a < b < \infty$ and $f : [a;b] \to \mathbb{R}$ eine beschränkte Funktion auf $[a;b]$. Dann*

i) f ist über $[a;b]$ Riemann integrierbar genau dann wenn f auf $[a;b]$ λ-fast überall stetig ist.

ii) Ist f über $[a;b]$ Riemann integrierbar, dann ist f auch über $[a;b]$ Lebesgue integrierbar, und
$$\int_a^b f dx = \int_{[a;b]} f d\lambda \tag{27}$$

Beweis. Sei $f : [a;b] \to \mathbb{R}$ Riemann integrierbar. Wir zeigen, dass f fast überall stetig ist, und, dass f Lebesgue integrierbar ist, mit (27). Die Riemann Integrierbarkeit von f bedeutet, dass
$$\sup_T \underline{S}(T) = \inf_T \overline{S}(T)$$

wobei $\overline{S}(T)$ und $\underline{S}(T)$ die obere und, beziehungsweise, die untere Riemann'sche Summe zur Teilung T sind. Ist f integrierbar, so können wir für beliebiges $\varepsilon > 0$ eine Teilung T von $[a;b]$ finden, mit $\overline{S}(T) - \varepsilon \leq \underline{S}(T) \leq \overline{S}(T)$. Wir können deswegen eine Folge T_n von Teilungen von $[a;b]$ finden, mit $\overline{S}(T_n) - \underline{S}(T_n) \leq 1/n$ für alle $n \in \mathbb{N}$. Insbesondere gilt dann

$$\lim_{n\to\infty} \underline{S}(T_n) = \lim_{n\to\infty} \overline{S}(T_n) = \int_a^b f(x)dx$$

O.B.d.A. können wir auch annehmen, dass T_{n+1} eine Verfeinerung von T_n ist (d.h. $T_n \subset T_{n+1}$) für alle $n \in \mathbb{N}$. Für beliebigen $n \in \mathbb{N}$, sei $T_n = \{a = x_0^n < x_1^n < \cdots < x_{m_n}^n = b\}$ und, für alle $j = 1, \dots, m_n$, $I_j^n = (x_{j-1}^n; x_j^n]$,

$$m_j^n = \inf_{x\in I_j^n} f(x), \quad \text{und} \quad M_j^n = \sup_{x\in I_j^n} f(x)$$

Wir definieren dann die Funktionen $\varphi_n, \psi_n : [a;b] \to \mathbb{R}$ durch $\varphi_n(a) = \psi_n(a) = f(a)$ und

$$\varphi_n(x) = \sum_{j=1}^{m_n} m_j^n \chi_{I_j^n}(x)$$

und

$$\psi_n(x) = \sum_{j=1}^{m_n} M_j^n \chi_{I_j^n}(x)$$

für alle $x \in (a;b]$. Es gilt:

- $\varphi_n(x) \leq f(x) \leq \psi_n(x)$ für alle $x \in [a;b]$.
- Die Funktionen φ_n, ψ_n sind einfach und messbar bzg. $\mathcal{B}(\mathbb{R})$ (und deswegen auch bezüglich $\mathcal{M}_{\lambda^*}$), für alle $n \in \mathbb{N}$, mit

$$\int_{[a;b]} \varphi_n d\lambda = \underline{S}(T_n) \quad \text{und} \quad \int_{[a;b]} \psi_n d\lambda = \overline{S}(T_n)$$

- $\varphi_n(x) \leq \varphi_{n+1}(x)$ und $\psi_n(x) \geq \psi_{n+1}(x)$ für alle $x \in [a;b]$ und alle $n \in \mathbb{N}$. D.h. φ_n ist eine wachsende und ψ_n eine fallende Folge von einfachen Funktionen auf $[a;b]$. Das folgt aus der Annahme, dass $T_{n+1} \supset T_n$ eine Verfeinerung von T_n ist.

Da φ_n und ψ_n monoton sind, können wir die Grenzwerten $\varphi(x) = \lim_{n\to\infty} \varphi_n(x)$ und $\psi(x) = \lim_{n\to\infty} \psi_n(x)$ definieren. Da $\varphi_n(x) \leq \psi_n(x)$ für alle $x \in [a;b]$ und $n \in \mathbb{N}$ folgt, dass $\varphi(x) \leq \psi(x)$ für alle $x \in [a;b]$. Aus Satz 2.6 folgt, dass φ, ψ messbar sind. Aus der Beschränkheit von f folgt auch, dass eine Konstante $C > 0$ existiert, mit $|\varphi_n(x)| \leq C$ und $|\psi_n(x)| \leq C$ für alle $x \in [a;b]$. Aus dem Satz über dominierte Konvergenz finden wir, dass

$$\int_{[a;b]} \varphi d\lambda = \lim_{n\to\infty} \int_{[a;b]} \varphi_n d\lambda = \lim_{n\to\infty} \underline{S}(T_n) = \int_a^b f(x)dx$$

und, dass

$$\int_{[a;b]} \psi d\lambda = \lim_{n\to\infty} \int_{[a;b]} \psi_n d\lambda = \lim_{n\to\infty} \overline{S}(T_n) = \int_a^b f(x)dx$$

Da $\varphi(x) \le \psi(x)$ für alle $x \in [a;b]$ erhalten wir, dass

$$\int_{[a;b]} |\psi - \varphi| \, d\lambda = \int_{[a;b]} \psi d\lambda - \int_{[a;b]} \varphi d\lambda = 0$$

und also, dass $\psi(x) = \varphi(x)$ für fast alle $x \in [a;b]$. Wir setzen $A = \{x \in [a;b] : \varphi(x) \neq \psi(x)\}$; dann ist $\lambda(A) = 0$. Wir setzen auch $T = \bigcup_{n\in\mathbb{N}} T_n$; dann ist $\lambda(T) = \lim_{n\to\infty} \lambda(T_n) = 0$ (weil die Folge T_n wachsend ist). Mit $N = A \cup T$, es folgt dass $\lambda(N) = 0$. Wir behaupten nun, dass f auf $[a;b]\backslash N$ stetig ist. Sei nämlich $x \in [a;b]\backslash N$ und $\varepsilon > 0$ fest. Da

$$\lim_{n\to\infty} \varphi_n(x) = \varphi(x) = \psi(x) = \lim_{n\to\infty} \psi_n(x)$$

finden wir n_0 mit $\psi_{n_0}(x) - \varphi_{n_0}(x) \le \varepsilon$. Die Teilung T_{n_0} definiert dann ein Intervall I_0 um x auf dem die Funktionen φ_{n_0} und ψ_{n_0} konstant sind (x ist nicht ein Endpunkt von I_0, aus der Annahme, dass $x \notin T$). Für alle $y \in I_0$ gilt dann

$$|f(y) - f(x)| \le \psi_n(x) - \varphi_n(x) = \varepsilon$$

Das zeigt, dass f λ-fast überall stetig ist.

Da $\varphi_n(x) \le f(x) \le \psi_n(x)$ für alle $x \in [a;b]$ und $n \in \mathbb{N}$, folgt aus $\varphi(x) = \psi(x)$ für fast alle $x \in [a;b]$ auch, dass $\varphi(x) = f(x)$ für fast alle $x \in [a;b]$. Lemma 2.17 impliziert, dass f über $[a;b]$ Lebesgue integrierbar ist, mit

$$\int_{[a;b]} f d\lambda = \int_{[a;b]} \varphi d\lambda = \int_a^b f(x) dx$$

Es bleibt zu zeigen, dass, falls f fast überall stetig ist, so ist f auch Riemann integrierbar. Sei also f fast überall stetig und T_n die Teilung von $[a;b]$ die 2^n halboffenen Intervalle $I_1, \dots, I_{2^n}$ alle mit der selben Länge $(b-a)/2^n$ definiert (d.h. $I_j = (a + 2^{-n}j(b-a), a + 2^{-n}(j+1)(b-a)]$, für alle $j = 1, \dots, 2^n$). Wir setzen dann

$$\varphi_n(x) = \sum_{j=1}^{2^n} \inf_{x\in I_j} f(x) \, \frac{(b-a)}{2^n}, \quad \text{und} \quad \psi_n(x) = \sum_{j=1}^{2^n} \sup_{x\in I_j} f(x) \frac{(b-a)}{2^n}$$

Dann ist φ_n eine wachsende und ψ_n eine fallende Folge von einfachen messbaren Funktionen, mit $\varphi_n(x) \le f(x) \le \psi_n(x)$ für alle $n \in \mathbb{N}$ und $x \in [a;b]$. Nach Definition

$$\int_{[a;b]} \varphi_n d\lambda = \underline{S}(T_n) \quad \text{und} \quad \int_{[a;b]} \psi_n d\lambda = \overline{S}(T_n)$$

Ist f stetig an der Stelle $x \in [a;b]$, so gilt $f(x) = \lim_{n\to\infty} \varphi_n(x) = \lim_{n\to\infty} \psi_n(x)$. Unter der Annahme, dass f fast überall stetig ist, finden wir, dass $f(x) = \lim_{n\to\infty} \varphi_n(x) = \lim_{n\to\infty} \psi_n(x)$ für λ-fast alle $x \in [a;b]$. Das gibt $h_n(x) = \psi_n(x) - \varphi_n(x) \to 0$ für λ-fast alle $x \in [a;b]$. Da ausserdem f beschränkt ist, existiert eine Konstante $C > 0$ mit $|h_n(x)| \le C$ für alle $x \in [a;b]$ und $n \in \mathbb{N}$. Der Satz über dominierte Konvergenz impliziert, dass

$$\lim_{n\to\infty} \overline{S}(T_n) - \underline{S}(T_n) = \lim_{n\to\infty} \int_{[a;b]} (\psi_n - \varphi_n) d\lambda = \lim_{n\to\infty} \int_{[a;b]} h_n d\lambda = 0$$

Da

$$\inf_T \overline{S}(T) \leq \overline{S}(T_n) = \overline{S}(T_n) - \underline{S}(T_n) + \underline{S}(T_n) \leq \overline{S}(T_n) - \underline{S}(T_n) + \sup_T \underline{S}(T)$$

für alle $n \in \mathbb{N}$ gilt, muss auch

$$\inf_T \overline{S}(T) \leq \sup_T \underline{S}(T) + \lim_{n\to\infty} \overline{S}(T_n) - \underline{S}(T_n) = \sup_T \underline{S}(T)$$

Da $\inf_T \overline{S}(T) \geq \sup_T \underline{S}(T)$ immer gilt, es folgt, dass $\inf_T \overline{S}(T) = \sup_T \underline{S}(T)$ und deswegen, dass f integrierbar ist. □

Wir erinnern nun aus Analysis 1 den Begriff von uneigentlichen Riemann'schen Integral. Sei $-\infty < a < b \leq \infty$. Wir sagen die Funktion $f : [a;b) \to \mathbb{R}$ ist über $[a;b)$ uneigentlich Riemann integrierbar, falls f über $[a;z]$ beschränkt und Riemann integrierbar ist, für alle $z \in (a;b)$, und falls der Grenzwert

$$\lim_{z\to b} \int_a^z f(x)dx$$

existiert. In diesem Fall schreiben wir

$$\int_a^b f(x)dx = \lim_{z\to b} \int_a^z f(x)dx$$

Analog definieren wir die uneigentliche Riemann Integrierbarkeit von $f : (a;b] \to \mathbb{R}$ für $-\infty \leq a < b < \infty$. Für $-\infty \leq a < b \leq \infty$ sagen wir, dass $f : (a;b) \to \mathbb{R}$ auf $(a;b)$ uneigentlich Riemann integrierbar ist, falls ein $c \in (a;b)$ existiert, so, dass f auf $(a;c]$ und auf $[c;b)$ uneigentlich Riemann integrierbar ist. In diesem Fall definieren wir

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx$$

Satz 2.28. *Sei $I \subset \mathbb{R}$ ein Intervall (möglicherweise unendlich lang) und $f : I \to \mathbb{R}$ Riemann-integrierbar auf jedem kompakten Teilintervall von I. Dann ist f auf I Lebesgue integrierbar, genau dann wenn $|f|$ uneigentlich Riemann integrierbar ist. In diesem Fall gilt*

$$\int_{[a;b)} f d\lambda = \int_a^b f(x)dx$$

(wobei auf der rechte Seite das uneigentliche Riemann'sche Integral gemeint ist).

Beweis. Wir betrachten den Fall $I = [a;b)$, die andere Fällen sind analog. Sei b_n irgendeine wachsende Folge in $[a;b)$ mit $b_n \to b$ als $n \to \infty$. Wir setzen $I_n = [a;b_n]$. Dann ist $f\chi_{I_n}$ Riemann integrierbar fü alle $n \in \mathbb{N}$. Satz 2.27 impliziert, dass $f\chi_{I_n}$ als Funktion auf $(\mathbb{R}, \mathcal{M}_{\lambda^*}, \lambda)$ integrierbar ist. Insbesondere ist $f\chi_{I_n}$ Borel messbar für alle $n \in \mathbb{N}$. Damit ist auch $f(x) = \lim_{n\to\infty} f\chi_{I_n}(x)$ für alle $x \in [a;b)$ messbar. Die Folge $|f|\chi_{I_n}$ ist monoton wachsend und konvergiert punktweise gegen $|f|$. Der Satz von der monotonen Konvergenz impliziert, dass

$$\int_I |f| d\lambda = \lim_{n\to\infty} \int_{I_n} |f| d\lambda = \lim_{n\to\infty} \int_a^{b_n} |f| d\lambda \tag{28}$$

wobei die zweite Gleichheit wieder aus Satz 2.27 folgt. Mit (28) können wir nun die Äquivalenz von Riemann und Lebesgue Integrierbarkeit beweisen. Ist nämlich f Riemann integrierbar, dann konvergiert die rechte Seite von (28) gegen einem Grenzwert in $\mathbb{R}$. Die Endlichkeit der linken Seite von (28) zeigt die Integrierbarkeit von $|f|$ als eine Funktion auf $(\mathbb{R}, \mathcal{M}_{\lambda^*}, \lambda)$. Satz 2.16 impliziert dann, dass f integrierbar ist. Ist umgekehrt f lebesgue integrierbar, dann ist $|f|$ auch integrierbar, und die linke Seite von (28) ist endlich. Da

$$\lim_{n\to\infty} \int_a^{b_n} f(x)dx = \int_I |f| d\lambda$$

unabhängig aus der Wahl der (monotonen) Folge b_n ist, existiert der Limes

$$\lim_{z\to b} \int_a^z |f(x)|dx = \int_I |f| d\lambda$$

Damit ist $|f|$ Riemann integrierbar. Um zu zeigen, dass Riemann und Lebesgue Integrale übereinstimmen bemerken wir, dass $|f\chi_{I_n}| \leq |f|$ und $f\chi_{I_n} \to f$ punktweise. Also dominierte Konvergenz impliziert, dass

$$\int_I f d\lambda = \lim_{n\to\infty} \int f\chi_{I_n} d\lambda = \lim_{n\to\infty} \int_a^{b_n} f(x)dx = \int_a^b f(x)dx$$

$\square$

Beispiele:

- Die Funktion $f : \mathbb{R} \to \mathbb{R}$ definiert durch $f(x) = e^{-|x|}$ its uneigentlich Riemann und Lebesgue integrierbar, mit

$$\int_{\mathbb{R}} e^{-|x|} d\lambda = \int_{-\infty}^{\infty} e^{-|x|}dx = 2\int_0^\infty e^{-x}dx = -2e^{-x}|_0^\infty = 2$$

- Die Funktion $f : \mathbb{R} \to \mathbb{R}$ durch $f(x) = 1$ für $x = 0$ und $f(x) = (\sin x)/x$ für $x \neq 0$ ist über $\mathbb{R}$ Riemann integrierbar aber nicht Lebesgue integrierbar.

2.5 Produktmasse und das Theorem von Fubini

Seien $(X, \mathcal{A})$ und $(Y, \mathcal{B})$ zwei messbare Räume. Auf der Menge

$$X \times Y = \{(x, y) : x \in X \text{ und } y \in Y\}$$

definieren wir die Produkt σ-Algebra $\mathcal{A} \times \mathcal{B} := \sigma(\mathcal{F})$ als die σ-Algebra, die durch die Familie

$$\mathcal{F} = \{A \times B : A \in \mathcal{A} \text{ und } B \in \mathcal{B}\}$$

erzeugt wird. D.h. $\mathcal{A} \times \mathcal{B}$ ist die kleinste σ-Algebra, die alle Produktmengen der Form $A \times B = \{(x, y) \in X \times Y : x \in A, y \in B\}$ für $A \in \mathcal{A}$ und $B \in \mathcal{B}$ enthält.

Zum Beispiel definiert $\mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R})$ eine σ-Algebra auf $\mathbb{R}^2$.

Proposition 2.29. *Es gilt* $\mathcal{B}(\mathbb{R}^2) = \mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R})$.

Proof. Einerseits gilt $\mathcal{B}(\mathbb{R}^2) \subset \mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R})$, weil $\mathcal{B}(\mathbb{R}^2)$ aus der Mengen $(a_1, b_1] \times (a_2, b_2]$ erzeugt wird, und $(a_1, b_1] \times (a_2, b_2] \in \mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R})$ für alle $a_1 < b_1, a_2 < b_2$ gilt. Es bleibt zu zeigen, dass $\mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R}) \subset \mathcal{B}(\mathbb{R}^2)$. Da $\mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R})$ aus Mengen der Form $A \times B$ mit $A, B \in \mathcal{B}(\mathbb{R})$ erzeugt wird, genügt es zu beweisen, dass $A \times B \in \mathcal{B}(\mathbb{R}^2)$ für alle $A, B \in \mathcal{B}(\mathbb{R})$. Seien $\pi_1 : \mathbb{R}^2 \to \mathbb{R}$ und $\pi_2 : \mathbb{R}^2 \to \mathbb{R}$ die Projektionen $\pi_1(x, y) = x$ und $\pi_2(x, y) = y$. π_1, π_2 sind klar stetig, und deswegen bezüglich $\mathcal{B}(\mathbb{R}^2)$ messbar. Deswegen sind $\pi_1^{-1}(A) = A \times \mathbb{R}$ und $\pi_2^{-1}(B) = \mathbb{R} \times B$ in $\mathcal{B}(\mathbb{R}^2)$ für alle $A, B \in \mathcal{B}(\mathbb{R})$. Deshalb ist $A \times B = (A \times \mathbb{R}) \cap (\mathbb{R} \times B) \in \mathcal{B}(\mathbb{R}^2)$. □

Sei nun $E \subset X \times Y$. Für beliebige $x \in X$ und $y \in Y$, definieren wir die Querschnitte

$$E_x = \{y \in Y : (x, y) \in E\} \subset Y, \quad \text{und} \quad E^y = \{x \in X : (x, y) \in E\} \subset X$$

Ist $f : X \times Y \to [-\infty; \infty]$, dann definieren wir für $x \in X$ und $y \in Y$ beliebig die Querschnitte $f_x : Y \to [-\infty; \infty]$ und $f^y : X \to [-\infty; \infty]$ durch $f_x(y) = f(x, y)$ und $f^y(x) = f(x, y)$.

Lemma 2.30. *Seien $(X, \mathcal{A})$ und $(Y, \mathcal{B})$ zwei messbare Räume.*

a) Ist $E \in \mathcal{A} \times \mathcal{B}$ so gilt $E_x \in \mathcal{B}$ und $E^y \in \mathcal{A}$ für alle $x \in X$ und $y \in Y$.

b) Ist $f : X \times Y \to [-\infty; \infty]$ messbar bezüglich $\mathcal{A} \times \mathcal{B}$, so sind f_x messbar bzg. $\mathcal{B}$ für alle $x \in X$ und f^y messbar bzg. $\mathcal{A}$ für alle $y \in Y$.

Beweis. a) Sei $x \in X$ fest. Wir betrachten die Familie

$$\mathcal{F}_x = \{E \subset X \times Y : E_x \in \mathcal{B}\}$$

Für alle $A \in \mathcal{A}$ und $B \in \mathcal{B}$ gilt $A \times B \in \mathcal{F}_x$, weil $(A \times B)_x = B$, falls $x \in A$, und $(A \times B)_x = \emptyset$, falls $x \notin A$. Insbesondere ist $X \times Y \in \mathcal{F}_x$. Ist $E \in \mathcal{F}_x$ so gilt $E_x \in \mathcal{B}$. Dann ist $(E^c)_x = \{y \in Y : (x, y) \in E^c\} = \{y \in Y : (x, y) \notin E\} = (E_x)^c \in \mathcal{B}$, und deswegen $E^c \in \mathcal{F}_x$. Ist weiter $(E_n)_{n \in \mathbb{N}}$ eine Folge in $\mathcal{F}_x$, dann gilt $(E_n)_x \in \mathcal{B}$ für alle $n \in \mathbb{N}$. Damit ist

$$\left(\bigcup_{n \in \mathbb{N}} E_n \right)_x = \bigcup_{n \in \mathbb{N}} (E_n)_x \in \mathcal{B}$$

und $\bigcup_{n \in \mathbb{N}} E_n \in \mathcal{F}_x$. Das zeigt, dass $\mathcal{F}_x$ eine σ-Algebra ist. Deswegen muss $\mathcal{F}_x \supset \mathcal{A} \times \mathcal{B}$, weil $\mathcal{A} \times \mathcal{B}$ die kleinste σ-Algebra ist, die alle Mengen der Form $A \times B$ enthält. Das bedeutet, dass $E_x \in \mathcal{B}$ für alle $E \in \mathcal{A} \times \mathcal{B}$. Analog ist $E^y \in \mathcal{A}$ für alle $E \in \mathcal{A} \times \mathcal{B}$.

Sei nun $f : X \times Y \to [-\infty; \infty]$ messbar bezüglich $\mathcal{A} \times \mathcal{B}$. Für jede $D \subset [-\infty; \infty]$ haben wir $(f_x)^{-1}(D) = (f^{-1}(D))_x$. Ist f messbar bzg. $\mathcal{A} \times \mathcal{B}$, so gilt $f^{-1}((-\infty; t]) \in \mathcal{A} \times \mathcal{B}$ f ur alle $t \in \mathbb{R}$. Dann ist aber, aus Teil a), $(f_x)^{-1}((-\infty; t]) = (f^{-1}((-\infty; t]))_x \in \mathcal{B}$ für alle $t \in (-\infty; t]$ und $x \in X$.Deswegen ist f_x $\mathcal{B}$-messbar für alle $x \in X$. Analog ist f^y $\mathcal{A}$-messbar für alle $y \in Y$. □

Bemerkung: Mit Hilfe von Lemma 2.30 können wir zeigen, dass die zu Proposition 2.29 äquivalente Behauptung für die σ-Algebra von Lebesgue messbaren Teilmengen von $\mathbb{R}^2$ falsch ist; d.h. $\mathcal{M}_{\lambda_2^*} \neq \mathcal{M}_{\lambda^*} \times \mathcal{M}_{\lambda^*}$. Sei nämlich $M \subset \mathbb{R}$ mit $M \notin \mathcal{M}_{\lambda^*}$. Wir betrachten dann die Menge $\{0\} \times M \subset \mathbb{R}^2$. Es gilt $\lambda_2^*(\{0\} \times M) = 0$. Deswegen ist $\{0\} \times M \in \mathcal{M}_{\lambda_2^*}$. Wäre nun $\{0\} \times M \in \mathcal{M}_{\lambda^*} \times \mathcal{M}_{\lambda^*}$, so würde auch der Querschnitt $(\{0\} \times M)_0 = M \in \mathcal{M}_{\lambda^*}$, in Wiederspruch zur Wahl von M.

Proposition 2.31. *Seien $(X, \mathcal{A}, \mu)$ und $(Y, \mathcal{B}, \nu)$ σ-endliche Massräume, und $E \in \mathcal{A} \times \mathcal{B}$. Dann ist die Funktion $x \to \nu(E_x)$ messbar bezüglich $\mathcal{A}$. Analog ist $y \to \mu(E^y)$ messbar bezüglich $\mathcal{B}$.*

Beweis. Nehmen wir zunächst an, dass das Mass ν endlich ist. Wir definieren

$$\mathcal{D} = \{E \in \mathcal{A} \times \mathcal{B} : x \to \nu(E_x)\ \mathcal{A} - \text{messbar ist}\}$$

Wir zeigen, dass $\mathcal{D}$ ein Dynkin System auf $X \times Y$ ist. Erinnere aus Def. 1.11, dass ein Dynkin System $\mathcal{D}$ auf $X \times Y$ eine Familie von Teilmengen von $X \times Y$ ist, mit den Eigenschaften: i) $X \times Y \in \mathcal{D}$, ii) $D \in \mathcal{D}$ impliziert, dass $D^c \in \mathcal{D}$, iii) $(D_n)_{n \in \mathbb{N}}$ eine Folge von paarweise disjunkten Mengen in $\mathcal{D}$, dann ist auch $\bigcup_{n \in \mathbb{N}} D_n \in \mathcal{D}$.

Für $A \in \mathcal{A}$ und $B \in \mathcal{B}$ gilt $(A \times B)_x = B$ falls $x \in A$ und $(A \times B)_x = \emptyset$ falls $x \notin A$. Deswegen ist $\nu((A \times B)_x) = \nu(B)\chi_A(x)$ messbar bezüglich $\mathcal{A}$. Also gehört jede Produktmenge $A \times B$, mit $A \in \mathcal{A}$ und $B \in \mathcal{B}$, zu $\mathcal{D}$. Insbesondere ist $X \times Y \in \mathcal{D}$. Sei nun $E \in \mathcal{D}$ beliebig. Wir behaupten, dass $E^c \in \mathcal{D}$. Das folgt aus $(E^c)_x = (E_x)^c$ und $\nu((E^c)_x) = \nu((E_x)^c) = \nu(Y) - \nu(E_x)$ (da ν endlich ist). Schlussendlich, sei $\{E_n\}_{n \in \mathbb{N}}$ eine Folge disjunkten Mengen in $\mathcal{D}$. Dann ist $(\bigcup_{n \in \mathbb{N}} E_n)_x = \bigcup_{n \in \mathbb{N}} (E_n)_x$, wobei $(E_n)_x$ eine Folge von disjunkten Mengen in $\mathcal{B}$ ist. Deswegen, aus der σ-Additivität von ν, gilt

$$\nu\left(\left(\bigcup_{n \in \mathbb{N}} E_n\right)_x\right) = \nu\left(\bigcup_{n \in \mathbb{N}} (E_n)_x\right) = \sum_{n \in \mathbb{N}} \nu((E_n)_x)$$

Wir bemerken, dass die Summe $\sum_{n=0}^m \nu((E_n)_x)$ als endliche Summe messbaren Funktionen auch $\mathcal{A}$-messbar ist, für alle $m \in \mathbb{N}$. Damit ist auch $\sum_{n \in \mathbb{N}} \nu((E_n)_x)$ als Limes einer Folge messbaren Funktionen $\mathcal{A}$-messbar. Das zeigt also, dass $\mathcal{D}$ ein Dynkin System ist.

Sei nun $\mathcal{F} = \{A \times B : A \in \mathcal{A}, B \in \mathcal{B}\}$ und bemerke, dass $\mathcal{F}$ abgeschlossen ist, bezüglich endliche Durchschnitte. In der Tat, für $A_1 \times B_1, A_2 \times B_2 \in \mathcal{F}$ gilt auch

$$(A_1 \times B_1) \cap (A_2 \times B_2) = (A_1 \cap A_2) \cap (B_1 \cap B_2) \in \mathcal{F}\,.$$

Nach Definition ist $\mathcal{A} \times \mathcal{B} = \sigma(\mathcal{F})$ die σ-Algebra die aus $\mathcal{F}$ erzeugt wird. Anderseits haben wir oben gezeigt, dass $\mathcal{F} \subset \mathcal{D}$. Da $\mathcal{D}$ ein Dynkin System ist, muss $\delta(\mathcal{F}) \subset \mathcal{D}$, wobei $\delta(\mathcal{F})$ der Dynkin System ist, der aus $\mathcal{F}$ erzeugt wird. Da aber $\mathcal{F}$ abgeschlossen ist, bezüglich endliche Durchschnitte, folgt aus Satz 1.14, dass $\delta(\mathcal{F}) = \sigma(\mathcal{F})$; d.h. der aus $\mathcal{F}$ erzeugten Dynkin System stimmt mit der aus $\mathcal{F}$ erzeugten σ-Algebra überein. Das zeigt, dass $\mathcal{A} \times \mathcal{B} = \mathcal{D}$ (weil $\mathcal{D} \subset \mathcal{A} \times \mathcal{B}$ aus Definition klar ist). Wir haben also gezeigt, dass $x \to \nu(E_x)$ $\mathcal{A}$-messbar ist, für alle $E \in \mathcal{A} \times \mathcal{B}$.

Nehmen wir nun an, dass ν σ-endlich, aber nicht unbedingt endlich, ist. Dann finden wir eine Folge $(D_n)_{n \in \mathbb{N}}$ von disjunkten Teilmengen in $\mathcal{B}$, mit $\nu(D_n) < \infty$ für alle $n \in \mathbb{N}$ und $\bigcup_{n \in \mathbb{N}} D_n = Y$. Wir definieren dann das Mass $\nu_n : \mathcal{B} \to [0; \infty]$ durch $\nu_n(B) := \nu(B \cap D_n)$ für alle $n \in \mathbb{N}$. Für alle $N \in \mathbb{N}$ ist ν_n ein endliches Mass, und für alle $B \in \mathcal{B}$ gilt $\nu(B) = \sum_{n \in \mathbb{N}} \nu_n(B)$. Gemäss was wir oben gezeigt haben, ist $x \to \nu_n(E_x)$ $\mathcal{A}$-messbar, für alle $n \in \mathbb{N}$ und alle $E \in \mathcal{A} \times \mathcal{B}$. Wir schliessen, dass, für jede $E \in \mathcal{A} \times \mathcal{B}$ auch

$$\nu(E_x) = \sum_{n \in \mathbb{N}} \nu_n(E_x) = \lim_{m \to \infty} \sum_{n=0}^m \nu_n(E_x)$$

als Limes einer Folge messbaren Funktionen wieder $\mathcal{A}$-messbar. Analog zeigt man, dass $y \to \mu(E^y)$ $\mathcal{B}$-messbar ist. $\square$

Gegeben zwei Massräume $(X, \mathcal{A}, \mu)$ und $(Y, \mathcal{B}, \nu)$, können wir nun die letzte Proposition benutzen, um das Produktmass $\mu \times \nu$ auf der σ-Algebra $\mathcal{A} \times \mathcal{B}$ zu definieren.

Theorem 2.32. *Seien $(X, \mathcal{A}, \mu)$ und $(Y, \mathcal{B}, \nu)$ zwei σ-endliche Massräume. Dann existiert ein eindeutiges Mass $\mu \times \nu$ auf $\mathcal{A} \times \mathcal{B}$, mit $(\mu \times \nu)(A \times B) = \mu(A)\nu(B)$ für alle $A \in \mathcal{A}, B \in \mathcal{B}$. Das Mass $\mu \times \nu$ ist aus*

$$(\mu \times \nu)(E) = \int_X \nu(E_x) d\mu = \int_Y \mu(E^y) d\nu \tag{29}$$

für alle $E \in \mathcal{A} \times \mathcal{B}$ gegeben.

Um die Eindeutigkeit im Theorem 2.32 zu beweisen, benutzen wir das folgende Resultat.

Proposition 2.33. *Sei $(X, \mathcal{A})$ ein messbarer Raum, und sei $\mathcal{F}$ eine Familie von Teilmengen von X mit $\mathcal{A} = \sigma(\mathcal{F})$ und*

$$A, B \in \mathcal{F} \;\Rightarrow\; A \cap B \in \mathcal{F}\,.$$

Seien μ, ν zwei Masse auf $(X, \mathcal{A})$ mit $\mu(C) = \nu(C)$ für alle $C \in \mathcal{F}$ und es existiere eine Folge $\{C_n\}_{n\in\mathbb{N}}$ in $\mathcal{F}$ mit $C_{n+1} \supset C_n$ für alle $n \in \mathbb{N}$, $X = \bigcup_{n\in\mathbb{N}} C_n$ und $\mu(C_n) < \infty$ und $\nu(C_n) < \infty$ für alle $n \in \mathbb{N}$. Dann gilt $\mu = \nu$.

Beweis. Wir nehmen zunächst an, μ, ν sind endliche Masse, mit $\mu(X) = \nu(X)$. Dann setzen wir

$$\mathcal{D} = \{A \in \mathcal{A} : \mu(A) = \nu(A)\}$$

Nach Annahme ist $X \in \mathcal{D}$. Ist weiter $E \in \mathcal{D}$, dann gilt auch $\mu(E^c) = \mu(X) - \mu(E) = \nu(X) - \nu(E) = \nu(E^c)$, und damit $E^c \in \mathcal{D}$. Schlussendlich, falls $\{E_n\}_{n\in\mathbb{N}}$ eine Folge von disjunkten Mengen in $\mathcal{D}$ ist, dann gilt $\mu(E_n) = \nu(E_n)$ für alle $n \in \mathbb{N}$, und wegen der Additivität von μ und ν, finden wir

$$\mu\left(\bigcup_{n\in\mathbb{N}} E_n\right) = \sum_{n\in\mathbb{N}} \mu(E_n) = \sum_{n\in\mathbb{N}} \nu(E_n) = \nu\left(\bigcup_{n\in\mathbb{N}} E_n\right)$$

D.h. $\bigcup_{n\in\mathbb{N}} E_n \in \mathcal{D}$. Das zeigt, dass $\mathcal{D}$ ein Dynkin System ist. Nach Annahme ist $\mathcal{F} \subset \mathcal{D}$. Damit ist $\delta(\mathcal{F})$, das kleinste Dynkin System, das $\mathcal{F}$ enthält, sicher in $\mathcal{D}$ enthalten, d.h. $\delta(\mathcal{F}) \subset \mathcal{D}$. Anderseits, da $\mathcal{F}$ abgeschlossen ist bezüglich endliche Durchschnitte, finden wir $\delta(\mathcal{F}) = \sigma(\mathcal{F}) = \mathcal{A} \subset \mathcal{D}$. Damit ist $\mathcal{D} = \mathcal{A}$ und $\mu(A) = \nu(A)$ für alle $A \in \mathcal{A}$.

Nun betrachten wir den Fall, dass μ, ν nicht unbedingt endlich sind. Wir nehmen aber an, es existiere eine wachsende Folge $\{C_n\}_{n\in\mathbb{N}}$ mit $C_n \in \mathcal{F}$, $C_n \subset C_{n+1}$ für alle $n \in \mathbb{N}$, $X = \bigcup_{n\in\mathbb{N}} C_n$ und $\mu(C_n) = \nu(C_n) < \infty$ für alle $n \in \mathbb{N}$. Dann definieren wir, für alle $n \in \mathbb{N}$, die Masse $\mu_n(A) := \mu(A \cap C_n)$ und $\nu_n(A) = \nu(A \cap C_n)$. Offenbar sind μ_n und ν_n für alle $n \in \mathbb{N}$ endliche Masse auf $(X, \mathcal{A})$, mit $\mu_n(C) = \nu_n(C)$ für alle $C \in \mathcal{F}$

(weil dann ist auch $C \cap C_n \in \mathcal{F}$). Aus dem ersten Teil des Beweises folgt, dass $\mu_n = \nu_n$ für alle $n \in \mathbb{N}$. Da für alle $A \in \mathcal{A}$, gilt

$$A = \bigcup_{n\in\mathbb{N}} A \cap C_n$$

wobei $A \cap C_n$ eine wachsende Folge ist, finden wir

$$\mu(A) = \lim_{n\to\infty} \mu(A \cap C_n) = \lim_{n\to\infty} \mu_n(A) = \lim_{n\to\infty} \nu_n(A) = \lim_{n\to\infty} \nu(A \cap C_n) = \nu(A)$$

für alle $A \in \mathcal{A}$. □

Beweis von Theorem 2.32. Die Funktionen $x \to \nu(E_x)$ und $y \to \mu(E^y)$ sind messbar (bezüglich $\mathcal{A}$, bezieungsweise, $\mathcal{B}$) aus Proposition 2.31. Deswegen sind die Integrale in (29) wohldefiniert (weil $\nu(E_x)$ und $\mu(E^y)$ nicht-negative Funktionen sind). Wir setzen also

$$(\mu \times \nu)_1(E) := \int_X \nu(E_x) d\mu \quad \text{und} \quad (\mu \times \nu)_2(E) := \int_Y \mu(E^y) d\nu$$

für alle $E \in \mathcal{A} \times \mathcal{B}$. Es gilt offenbar $(\mu \times \nu)_1(\emptyset) = (\mu \times \nu)_2(\emptyset) = 0$. Sei nun $\{E_n\}_{n\in\mathbb{N}}$ eine Folge von disjunkten Mengen in $\mathcal{A} \times \mathcal{B}$. Wir setzen $E = \bigcup_{n\in\mathbb{N}} E_n$. Es gilt

$$E_x = \bigcup_{n\in\mathbb{N}} E_{n,x} \quad \text{und} \quad E^y = \bigcup_{n\in\mathbb{N}} E_n^y$$

Da $\{E_{n,x}\}_{n\in\mathbb{N}}$ eine Folge disjunkten Mengen in $\mathcal{B}$ ist ($y \in E_{n,x} \cap E_{m,x}$ impliziert, dass $(x,y) \in E_n$ und $(x,y) \in E_m$ und also, dass $n = m$, da die E_n paarweise disjunkt sind), folgt aus der σ-Additivität von ν, dass

$$\nu(E_x) = \nu\left(\bigcup_{n\in\mathbb{N}} E_{n,x}\right) = \sum_{n\in\mathbb{N}} \nu(E_{n,x})$$

Analog folgt aus der σ-Additivität von μ, dass

$$\mu(E^y) = \sum_{n\in\mathbb{N}} \mu(E_n^y)$$

Wir schliessen, dass

$$(\mu \times \nu)_1(E) = \int_X \nu(E_x) d\mu = \int_X \sum_{n\in\mathbb{N}} \nu(E_{n,x}) d\mu = \sum_{n\in\mathbb{N}} \int_X \nu(E_{n,x}) d\mu = \sum_{n\in\mathbb{N}} (\mu \times \nu)_1(E_n)$$

und, analog, dass

$$(\mu \times \nu)_2(E) = \sum_{n\in\mathbb{N}} (\mu \times \nu)_2(E_n)$$

In beiden Identitäten haben wir Korollar 2.24 (Beppo Levi) benutzt. Das zeigt, dass $(\mu \times \nu)_1$ und $(\mu \times \nu)_2$ Masse auf $\mathcal{A} \times \mathcal{B}$ definieren.

Ferner bemerken wir, dass, für $A \in \mathcal{A}$ und $B \in \mathcal{B}$,

$$\nu((A \times B)_x) = \chi_A(x)\nu(B) \quad \text{und} \quad \mu((A \times B)^y)\chi_B(y)\mu(A) .$$

Das gibt

$$(\mu \times \nu)_1(A \times B) = \nu(B) \int_X \chi_A d\mu = \mu(A)\nu(B)$$

und

$$(\mu \times \nu)_2(A \times B) = \mu(A) \int_Y \chi_B d\nu = \mu(A)\nu(B)$$

Sei $\mathcal{F} = \{A \times B : A \in \mathcal{A} \text{ und } B \in \mathcal{B}\}$. Dann $\mathcal{F}$ ist abgeschlossen bezüglich endliche Durchschnitte, und, nach Definition, $\mathcal{A} \times \mathcal{B} = \sigma(\mathcal{F})$. Die zwei Masse $(\mu \times \nu)_1$ und $(\mu \times \nu)_2$ stimmen auf $\mathcal{F}$ überein. Wir möchten nun Proposition 2.33 anwenden, um zu zeigen, dass $(\mu \times \nu)_1 = (\mu \times \nu)_2$ gelten muss. Dazu brauchen wir noch zu zeigen, dass eine wachsende Folge $\{C_n\}_{n\in\mathbb{N}}$ in $\mathcal{F}$ existiert, mit $(\mu \times \nu)_1(C_n) = (\mu \times \nu)_2(C_n) < \infty$ für alle $n \in \mathbb{N}$, so, dass $X \times Y = \bigcup_{n\in\mathbb{N}} C_n$. Dazu benutzen wir die Tatsache, dass μ, ν σ-endlich sind. Das impliziert, dass es wachsende Folge $\{A_j\}_{j\in\mathbb{N}}$ in $\mathcal{A}$ und $\{B_j\}_{j\in\mathbb{N}}$ in $\mathcal{B}$ existieren, mit $\mu(A_j) < \infty$ und $\nu(B_j) < \infty$ für alle $j \in \mathbb{N}$, und so, dass $X = \bigcup_{j\in\mathbb{N}} A_j$ und $Y = \bigcup_{j\in\mathbb{N}} B_j$. Wir setzen dann

$$C_n = \bigcup_{j=1}^{n} A_j \times \bigcup_{j=1}^{n} B_j$$

Dann ist $C_n \in \mathcal{F}$ für alle $n \in \mathbb{N}$,

$$(\mu \times \nu)_1(C_n) = (\mu \times \nu)_2(C_n) = \mu\left(\bigcup_{j=1}^{n} A_j\right) \nu\left(\bigcup_{j=1}^{n} B_j\right) = \mu(A_n)\nu(B_n) < \infty$$

und $\bigcup_{n\in\mathbb{N}} C_n = X \times Y$. Damit können wir Proposition 2.33 anwenden; wir finden, dass $(\mu \times \nu)_1 = (\mu \times \nu)_2$. Wir setzen $(\mu \times \nu) = (\mu \times \nu)_1 = (\mu \times \nu)_2$.

Ist schlussendlich $\theta : \mathcal{A} \times \mathcal{B} \to [0; \infty]$ ein anderes Mass mit $\theta(A \times B) = \mu(A)\nu(B)$ für alle $A \in \mathcal{A}$ und $B \in \mathcal{B}$, so folgt wieder aus Proposition 2.33, dass $\theta = \mu \times \nu$. $\square$

Beispiel: Wir haben schon gezeigt, dass $\mathcal{B}(\mathbb{R}^2) = \mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R})$. Auf $\mathcal{B}(\mathbb{R}^2)$ ist das Lebesgue Mass λ_2 definiert. Auf $\mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R})$ können wir durch Theorem 2.32 das Produktmass $\lambda \times \lambda$ definieren. Sei $\mathcal{F} = \{(a_1, b_1] \times (a_2, b_2] : a_1 < b_1, \text{ und } a_2 < b_2\}$. Dann ist $\mathcal{F}$ abgeschlossen bezüglich endliche Durchschnitte, und $\mathcal{B}(\mathbb{R}^2) = \sigma(\mathcal{F})$. Da λ_2 und $\lambda_1 \times \lambda_1$ auf $\mathcal{F}$ übereinstimmen, und weil wir können eine wachsende Folge C_n in $\mathcal{F}$ konstruieren, mit $\lambda_2(C_n) = \lambda \times \lambda(C_n) < \infty$ für alle $n \in \mathbb{N}$ so, dass $\mathbb{R}^2 = \bigcup_{n\in\mathbb{N}} C_n$, es folgt aus Proposition 2.33, dass $\lambda_2 = \lambda \times \lambda$.

Mit Theorem 2.32 haben wir nun aus zwei σ-endliche Massräume $(X, \mathcal{A}, \mu)$, $(Y, \mathcal{B}, \nu)$ ein neues Massraum $(X \times Y, \mathcal{A} \times \mathcal{B}, \mu \times \nu)$ konstruiert. Für eine Funktion $f : X \times Y \to [-\infty; \infty]$ können wir also untersuchen, ob sie $\mathcal{A} \times \mathcal{B}$-messbar ist. Wenn sie messbar ist, können wir untersuchen, ob sie bezüglich das Mass $\mu \times \nu$ integrierbar ist. Wenn sie integrierbar, können wir das Integral $\int_{X\times Y} f d(\mu \times \nu)$ definieren. Eine natürliche Frage ist: kann man das Integral bezüglich das Mass $\mu \times \nu$ durch die Integrale bezüglich μ und ν berechnen? Wegen die Formel (29) können wir erwarten, die Antwort sei, unter gewisse Voraussetzungen, ja. Das wird in den nächsten zwei Sätzen bewiesen.

Satz 2.34 (Satz von Tonelli)**.** *Seien* $(X, \mathcal{A}, \mu)$ *und* $(Y, \mathcal{B}, \nu)$ *zwei* σ*-endliche Massräume. Sei* $f : X \times Y \to [0; \infty]$ *eine* $\mathcal{A} \times \mathcal{B}$*-messbare Funktion. Aus Lemma 2.30 sind die Querschnitte* $f_x : Y \to [0; \infty]$ $\mathcal{B}$*-messbar für alle* $x \in X$ *und die Querschnitte* $f^y : X \to [0; \infty]$ $\mathcal{A}$*-messbar für alle* $y \in Y$*. Deswegen können wir für alle* $x \in X$ *und* $y \in Y$ *die Integrale* $\int_Y f_x d\nu$ *und* $\int_X f^y d\mu$ *definieren (sie können natürlich unendlich sein). Es gilt*

a) Die Funktion $x \to \int_Y f_x d\nu$ *ist* $\mathcal{A}$*-messbar und die Funktion* $y \to \int_X f^y d\mu$ *ist* $\mathcal{B}$*-messbar.*

b) Es gilt

$$\int_{X \times Y} f d(\mu \times \nu) = \int_X \left(\int_Y f_x d\nu \right) d\mu = \int_Y \left(\int_X f^y d\mu \right) d\nu$$

Beweis. Sei zunächst $E \in \mathcal{A} \times \mathcal{B}$, und $f = \chi_E$ die Charakteristische Funktion von E. Dann ist $f_x(y) = \chi_E(x, y) = 1$, falls $(x, y) \in E$ und $f_x(y) = 0$, falls $(x, y) \notin E$. D.h. $f_x(y) = 1$ falls $y \in E_x$ und $f_x(y) = 0$, falls $y \notin E_x$. Mit anderen Wörter ist $f_x = \chi_{E_x}$ die Charakteristische Funktion des Querschnittes E_x. Analog ist $f^y = \chi_{E^y}$. Deswegen gilt

$$\int_Y f_x d\nu = \nu(E_x), \quad \text{und} \quad \int_X f^y d\mu = \mu(E^y).$$

Aus Proposition 2.31 ist also die Funktion $x \to \int_Y f_x d\nu$ $\mathcal{A}$-messbar und die Funktion $y \to \int_X f^y d\mu$ $\mathcal{B}$-messbar. Aus Theorem 2.32 finden wir weiter, dass

$$\int_{X \times Y} f d(\mu \times \nu) = (\mu \times \nu)(E) = \int_X \nu(E_x) d\mu = \int_X \left(\int_Y f_x d\nu \right) d\mu$$

und analog

$$\int_{X \times Y} f d(\mu \times \nu) = (\mu \times \nu)(E) = \int_Y \mu(E^y) d\nu = \int_Y \left(\int_X f^y d\mu \right) d\nu$$

Das zeigt a) und b) für charakteristische Funktionen von Mengen in $\mathcal{A} \times \mathcal{B}$. Einfache $\mathcal{A} \times \mathcal{B}$-messbare nicht-negative Funktionen sind endliche lineare Kombinationen von charakteristische Funktionen von $E \in \mathcal{A} \times \mathcal{B}$. Aus Linearität des Integrals folgt, dass a),b) für beliebige einfache Funktion gelten. Aber eine beliebige $\mathcal{A} \times \mathcal{B}$-messbare Funktion auf $X \times Y$ kann gemäss Proposition 2.3 durch eine monotone Folge nicht-negative messbaren einfache Funktionen approximiert werden. Der Satz von der monotone Konvergenz zeigt, dann, a) und b) für beliebige messbare $f : X \times Y \to [0; \infty]$. □

Integrale von Funktionen auf $X \times Y$ mit Werten in $[-\infty; \infty]$ können auch durch berechnet, indem man zunächst über x und nachher über y integriert (oder zunächst über y und nachher über x), wenn wir annehmen, dass die Funktion f integrierbar ist. Das zeigen wir im nächsten Theorem.

Theorem 2.35 (Theorem von Fubini)**.** *Seien* $(X, \mathcal{A}, \mu)$ *und* $(Y, \mathcal{B}, \nu)$ *zwei* σ*-endliche Massräume. Sei* $f : X \times Y \to [-\infty; \infty]$ $\mathcal{A} \times \mathcal{B}$*-messbar und integrierbar bezüglich* $\mu \times \nu$*. Dann gilt*

a) Die Funktion f_x *ist* ν*-integrierbar für* μ*-fast alle* $x \in X$*. Die Funktion* f^y *ist* μ*-integrierbar für* ν*-fast alle* $y \in Y$*.*

b) Seien

$$I_f(x) = \begin{cases} \int_Y f_x d\nu, & \text{falls } f_x \ \nu\text{-integrierbar ist} \\ 0, & \text{sonst} \end{cases}$$

und

$$J_f(y) = \begin{cases} \int_X f^y d\mu, & \text{falls } f^y \ \mu\text{-integrierbar ist} \\ 0, & \text{sonst} \end{cases}$$

Dann gilt $I_f \in \mathcal{L}^1(X, \mathcal{A}, \mu)$ *und* $J_f \in \mathcal{L}^1(X, \mathcal{B}, \nu)$ *und*

$$\int_{X\times Y} f d(\mu \times \nu) = \int_X I_f d\mu = \int_Y J_f d\nu \tag{30}$$

Bemerkung:

1) Grob gesagt gilt, unter der Annahme, dass f integrierbar ist,

$$\int_{X\times Y} f d(\mu \times \nu) = \int_X \left(\int_Y f_x d\nu \right) d\mu = \int_Y \left(\int_X f^y d\mu \right) d\nu \,,$$

wie im Satz 2.34. Wir müssen aber diese Aussage wie in (30) formulieren, weil die Integrale $\int_X f^y d\mu$ und $\int_Y f_x d\nu$ nur fast-überall und nicht unbedingt überall definiert sind (im Satz 2.34, dagegen, waren die Integrale wegen Nicht-negativität von f überall wohldefiniert).

2) Die Strategie um das Integral $\int f d(\mu \times \nu)$ zu berechnen ist wie folgt. Zunächst benutzt man Satz 2.34, um das Integral von $|f|$ zu berechnen. Ist das Integral von $|f|$ endlich, so ist f integrierbar, und mann kann Theorem 2.35 anwenden, um den Wert von $\int f d(\mu \times \nu)$ zu bestimmen.

Beweis. Die Messbarkeit von f impliziert, dass auch f_+ und f_- $\mathcal{A} \times \mathcal{B}$-messbar sind. Aus Lemma 2.30 folgt, dass $(f_+)_x$ und $(f_-)_x$ $\mathcal{B}$-messbar sind, für alle $x \in X$. Da

$$(f_+)_x(y) = f_+(x,y) = \max(f(x,y), 0) = \max(f_x(y), 0) = (f_x)_+(y)$$

finden wir, dass $(f_x)_+$ und $(f_x)_-$ messbar sind, für alle $x \in X$. Satz 2.34 zeigt, dass die Funktionen $x \to \int_Y (f_+)_x \, d\nu = \int_Y (f_x)_+ \, d\nu$ und $x \to \int_Y (f_-)_x \, d\nu = \int_Y (f_x)_- \, d\nu$ $\mathcal{A}$-messbar sind, und, dass

$$\int_{X\times Y} f_+ d(\mu \times \nu) = \int_X \left(\int_Y (f_x)_+ d\nu \right) d\mu \quad \text{und}$$
$$\int_{X\times Y} f_- d(\mu \times \nu) = \int_X \left(\int_Y (f_x)_- d\nu \right) d\mu$$

Die Integrierbarkeit von f bedeutet, dass die linke Seite der letzten zwei Gleichungen endlich ist. Deswegen muss

$$\int_Y (f_x)_+ d\nu < \infty \quad \text{und} \quad \int_Y (f_x)_- d\nu < \infty$$

für μ-fast alle $x \in X$. D.h. f_x ist μ-fast überall integrierbar. Wir setzten

$$N = \left\{ x \in X : \int_Y (f_x)_+ d\nu = \infty, \quad \text{oder} \quad \int_Y (f_x)_- d\nu = \infty \right\}$$

Wegen Messbarkeit von $x \to \int_Y (f_x)_+ \, d\nu$ und $x \to \int_Y (f_x)_- \, d\nu$ ist $N \in \mathcal{A}$, mit $\mu(N) = 0$. Da auf N^c gilt

$$I_f(x) = \int_Y (f_x)_+ d\nu - \int_Y (f_x)_- d\nu$$

finden wir, dass

$$\begin{aligned}
\int_{X \times Y} f d(\mu \times \nu) &= \int_{X \times Y} f_+ d(\mu \times \nu) - \int_{X \times Y} f_- d(\mu \times \nu) \\
&= \int_X \left(\int_Y (f_x)_+ d\nu \right) d\mu - \int_X \left(\int_Y (f_x)_- d\nu \right) d\mu \\
&= \int_{N^c} \left(\int_Y (f_x)_+ d\nu \right) d\mu - \int_{N^c} \left(\int_Y (f_x)_- d\nu \right) d\mu \\
&= \int_{N^c} \left(\int_Y (f_x)_+ d\nu - \int_Y (f_x)_- d\nu \right) d\mu \\
&= \int_{N^c} I_f \, d\mu = \int_X I_f \, d\mu
\end{aligned}$$

weil $\mu(N) = 0$. Analog zeigt man, dass $\int_{X \times Y} f d(\mu \times \nu) = \int_Y J_f d\nu$. □

Wir haben bis jetzt Produkte von zwei Massräume betrachtet. Man kann die Resultate die wir bis jetzt bewiesen haben einfach auf dem Produkt von mehrere Massräume verallgemeinern. Seien nämlich $(X_1, \mathcal{A}_1, \mu_1), \dots, (X_k, \mathcal{A}_k, \mu_k)$ k σ-endliche Massräume. Dann definieren wir die Produktmenge

$$X_1 \times \dots \times X_k = \{(x_1, \dots, x_k) : x_1 \in X_1, \dots, x_k \in X_k\}$$

Auf $X_1 \times \dots \times X_k$ kann man die σ-Algebra $\mathcal{A}_1 \times \mathcal{A}_k$ definieren, als die kleinste σ-Algebra die alle Mengen der Form $A_1 \times \dots \times A_k$, mit $A_1 \in \mathcal{A}_1, \dots, A_k \in \mathcal{A}_k$, enthält. Wir bemerken, dass

$$\mathcal{A}_1 \times \dots \mathcal{A}_j \times \mathcal{A}_{j+1} \times \dots \times \mathcal{A}_k = (\mathcal{A}_1 \times \dots \times \mathcal{A}_j) \times (\mathcal{A}_{j+1} \times \dots \times \mathcal{A}_k)$$

für alle $j = 1, \dots, k$. Aus diesem Grund kann man ein Mass $\mu_1 \times \dots \times \mu_k$ auf $\mathcal{A}_1 \times \dots \times \mathcal{A}_k$, mit der Eigenschaft, dass

$$(\mu_1 \times \dots \times \mu_k)(A_1 \times \dots \times A_k) = \mu_1(A_1) \dots \mu_k(A_k) \tag{31}$$

für alle $A_1 \in \mathcal{A}_1, \dots, A_k \in \mathcal{A}_k$, durch rekursive Anwendung von Theorem 2.32 konstruieren. Im ersten Schritt definieren wir nämlich $\mu_1 \times \mu_2$ auf $\mathcal{A}_1 \times \mathcal{A}_2$. Im zweiten Schritt definieren wir dann $\mu_1 \times \mu_2 \times \mu_3 := (\mu_1 \times \mu_2) \times \mu_3$ auf $\mathcal{A}_1 \times \mathcal{A}_2 \times \mathcal{A}_3 = (\mathcal{A}_1 \times \mathcal{A}_2) \times \mathcal{A}_3$, und so weiter. Mit Proposition 2.33 kann man dann ähnlich wie in Theorem 2.32 zeigen, dass das konstruierte Mass $\mu_1 \times \dots \times \mu_k$ das einzige Mass ist, das (31) erfüllt. Ist

$f : X_1 \times \cdots \times X_k \to [0;\infty]$ eine $\mathcal{A}_1 \times \ldots \mathcal{A}_k$-messbare Funktion so kann man durch wiederholte Anwendung von Satz 2.34 das Integral

$$\begin{aligned}\int_{X_1\times\cdots\times X_k} f\, d(\mu_1\times\cdots\times\mu_k) &= \int_{X_1\times\cdots\times X_{k-1}} \left(\int_{X_k} f_{x_1,\dots,x_{k-1}} d\mu_k\right) d(\mu_1\times\cdots\times\mu_{k-1}) \\ &= \dots \\ &= \int_{X_1}\left(\int_{X_2}\dots\left(\int_{X_k} f_{x_1,\dots,x_{k-1}} d\mu_k\right)\dots d\mu_{k-1}\right) d\mu_1\end{aligned}$$

berechnen. Analog, falls $f : X_1\times\cdots\times X_k \to [-\infty;\infty]$ bzg. $\mu_1\times\cdots\times\mu_k$ integrierbar ist, so kann man Theorem 2.35 iterativ anwenden, um das Integral $\int_{X_1\times\cdots\times X_k} f d(\mu_1\times\cdots\times\mu_k)$ zu berechnen.

Die wichtigste Anwendung vom Satz von Tonelli und vom Theorem von Fubini ist die Berechnung von Integrale auf $\mathbb{R}^n$. Wir haben schon diskutiert, dass $\mathcal{B}(\mathbb{R}^2) = \mathcal{B}(\mathbb{R})\times\mathcal{B}(\mathbb{R})$ und, dass $\lambda_2 = \lambda\times\lambda$. Allgemeiner gilt $\mathcal{B}(\mathbb{R}^{n+m}) = \mathcal{B}(\mathbb{R}^n)\times\mathcal{B}(\mathbb{R}^m)$ und $\lambda_{n+m} = \lambda_n\times\lambda_m$. Ist $f : \mathbb{R}^{n+m} = \mathbb{R}^n\times\mathbb{R}^m \to [0;\infty]$ $\mathcal{B}(\mathbb{R}^2)$-messbar dann folgt aus Satz 2.34, dass

$$\int_{\mathbb{R}^{n+m}} f d\lambda_{n+m} = \int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^m} f(x,y) d\lambda_m(y)\right) d\lambda_n(x) \tag{32}$$

(Die Notation $d\lambda_n(x)$ und $d\lambda_m(y)$ erklärt welche Variable wird jeweils integriert). Ist $f : \mathbb{R}^n\times\mathbb{R}^m \to [-\infty;\infty]$ λ_{n+m}-integrierbar, so folgt (32) aus Theorem 2.35. Man kann (32) iterieren, und damit die Berechnung von Integrale auf $\mathbb{R}^n$ zur Berechnung von Integrale auf $\mathbb{R}$ zurückführen, die oft durch das Haupttheorem der Integralrechnung berechnet werden können. Für eine Funktion $\mathcal{B}(\mathbb{R}^n)$-messbare Funktion $f : \mathbb{R}^n \to [0;\infty]$ oder für eine Funktion $f : \mathbb{R}^n \to [-\infty;\infty]$ die λ_n-integrierbar ist finden wir

$$\int_{\mathbb{R}^n} f d\lambda_n = \int_{\mathbb{R}}\left(\int_{\mathbb{R}}\dots\left(\int_{\mathbb{R}} f(x_1,\dots,x_n) d\lambda(x_n)\right)\dots d\lambda(x_2)\right) d\lambda(x_1)$$

Beispiel: Wir wollen das Volumen der Einheitskugel $E = \{(x,y,z)\in\mathbb{R}^3 : x^2+y^2+z^2\le 1\}$ berechnen. Das ist aus

$$\lambda_3(E) = \int d\lambda_3\,\chi_E = \int d\lambda_x\left(\int d\lambda_y\left(\int d\lambda_z \chi_E(x,y,z)\right)\right)$$

Nun $\chi_E(x,y,z) = 1$ falls $x^2+y^2+z^2\le 1$ und $\chi_E(x,y,z) = 0$ sonst. Das gibt

$$\chi_E(x,y,z) = \begin{cases} 1 & \text{falls } x^2+y^2\le 1 \text{ und } |z|\le\sqrt{1-x^2-y^2} \\ 0 & \text{sonst}\end{cases}$$

Also

$$\int d\lambda_z\,\chi_E(x,y,z) = 2\chi(x^2+y^2\le 1)\sqrt{1-x^2-y^2}$$

Nun müssen wir

$$\begin{aligned}\int d\lambda_y \left(\int d\lambda_z \chi_E(x,y,z)\right) &= 2\int d\lambda_y \chi(x^2+y^2\leq 1)\sqrt{1-x^2-y^2} \\ &= 2\chi(|x|\leq 1)\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \sqrt{1-x^2-y^2}dy \\ &= 2\chi(|x|\leq 1)\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} dy\sqrt{1-x^2}\sqrt{1-\left(\frac{y}{\sqrt{1-x^2}}\right)^2} \\ &= 2\chi(|x|\leq 1)\,(1-x^2)\int_{-1}^{1} dy\sqrt{1-y^2} = \pi\chi(|x|\leq 1)(1-x^2)\end{aligned}$$

berechnen. Das Volume der Einheitskugel ist also durch

$$\lambda_3(E) = \pi\int d\lambda_x \chi(|x|\leq 1)(1-x^2) = 2\pi\int_0^1 (1-x^2) = 2\pi\left[x-\frac{x^3}{3}\right]_0^1 = \frac{4}{3}\pi$$

Wir haben schon gemerkt, dass im Gegensatz zur Beziehung $\mathcal{B}(\mathbb{R}^n) = \mathcal{B}(\mathbb{R})\times\cdots\times\mathcal{B}(\mathbb{R})$ gilt die analoge Aussage für die σ-Algebra von Lebesgue messbaren Mengen nicht, d.h. $\mathcal{M}_{\lambda_n^*}\neq\mathcal{M}_{\lambda^*}\times\cdots\times\mathcal{M}_{\lambda^*}$. Das bedeutet, man kann Satz 2.34 und Theorem 2.35 nicht direkt anwenden, zur Berechnung von Integrale von $\mathcal{M}_{\lambda_n^*}$-messbaren Funktionen. Oft ist das kein Problem, weil die Funktionen dessen Integrale wir berechnen möchten, normalerweise auch $\mathcal{B}(\mathbb{R}^n)$-messbar sind. In diesem Fall stimmt das Integral von f als Funktion auf dem Massraum $(\mathbb{R}^n,\mathcal{M}_{\lambda_n^*},\lambda_n)$ mit dem Integral von f als Funktion auf $(\mathbb{R}^n,\mathcal{B}(\mathbb{R}^n),\lambda_n)$ überein. Trotzdem wollen wir kurz diskutieren, wie man Fubini zur Berechnung vom Integral einer $\mathcal{M}_{\lambda_{n+m}^*}$-messbare Funktion anwenden kann, obwohl $\mathcal{M}_{\lambda_{n+m}^*}\neq\mathcal{M}_{\lambda_n^*}\times\mathcal{M}_{\lambda_m^*}$.

Satz 2.36. *Sei $E\subset\mathbb{R}^{m+n}$ Lebesgue messbar, d.h. $E\in\mathcal{M}_{\lambda_{n+m}^*}$. Für $x\in\mathbb{R}^n$ sei*

$$E_x = \{y\in\mathbb{R}^m : (x,y)\in E\}$$

Dann existiert $N\subset\mathbb{R}^n$, $N\in\mathcal{M}_{\lambda_n^*}$ mit $\lambda_n(N)=0$ so, dass

a) $E_x\in\mathcal{M}_{\lambda_m^*}$ für alle $x\in N^c$ (d.h. E_x ist Lebesgue messbar für fast alle $x\in\mathbb{R}^n$).

b) Die Funktion

$$f(x) = \begin{cases}\lambda_m(E_x) & \text{für } x\in N^c \\ 0 & \text{für } x\in N\end{cases}$$

ist $\mathcal{M}_{\lambda_n^*}$-messbar.

c) Es gilt

$$\lambda_{n+m}(E) = \int_{\mathbb{R}^n} f d\lambda_n$$

Bemerkung: verglichen mit Lemma 2.30 muss hier E_x nur für fast alle $x\in\mathbb{R}^n$ messbar sein. Das Beispiel $E=\{0\}\times M\subset\in\mathcal{M}_{\lambda_2^*}$, wobei $M\subset\mathbb{R}$ nicht Lebesgue messbar, zeigt, dass wir $E_x\in\mathcal{M}_{\lambda_m^*}$ für alle $x\in\mathbb{R}^n$ nicht erwarten können. Punkt b) ist das Analog zu Proposition 2.31; hier können wir nicht einfach $f(x)=\lambda_m(E_x)$ setzen, weil für $x\in N$ braucht E_x nicht Lebesgue messbar zu sein. Punkt c) ist das Analog von Theorem 2.32.

Beweis. Sei $E \in \mathcal{M}_{\lambda^*_{n+m}}$. Wir nehmen zunächst an, dass $\lambda_{n+m}(E) < \infty$. Da $\mathcal{M}_{\lambda^*_{n+m}}$ die Vervollständigung von $\mathcal{B}(\mathbb{R}^{n+m})$ ist, finden wir $F, G \in \mathcal{B}(\mathbb{R}^{n+m})$ mit $F \subset E \subset G$ und $\lambda_{n+m}(G \backslash F) = 0$ (was insbesondere $\lambda_{n+m}(G) = \lambda_{n+m}(E) = \lambda_{n+m}(F) < \infty$ impliziert). Es gilt dann: für alle $x \in \mathbb{R}$ ist $F_x \subset E_x \subset G_x$. Aus Lemma 2.30 wissen wir, dass $F_x, G_x \in \mathcal{B}(\mathbb{R}^m)$ für alle $x \in \mathbb{R}^n$. Aus Proposition 2.31 impliziert, dass $x \to \lambda_m(F_x)$ und $x \to \lambda_m(G_x)$ $\mathcal{B}(\mathbb{R}^n)$-messbar sind. Aus Theorem 2.32 folgt auch

$$\begin{aligned} \lambda_{n+m}(G) &= \int_{\mathbb{R}^n} \lambda_m(G_x) d\lambda_n, \text{ und} \\ \lambda_{n+m}(F) &= \int_{\mathbb{R}^n} \lambda_m(F_x) d\lambda_n \end{aligned}$$

Aus der Annahme $\lambda_{n+m}(F) = \lambda_{n+m}(G) < \infty$ folgt, dass $\lambda_m(F_x) \leq \lambda_m(G_x) < \infty$ für fast alle $x \in \mathbb{R}^n$. Da $F_x \subset G_x$ für alle $x \in \mathbb{R}^n$ haben wir $\lambda_m(F_x) \leq \lambda_m(G_x)$ für alle $x \in \mathbb{R}^n$. Deshalb

$$\int_{\mathbb{R}^n} |\lambda_m(G_x) - \lambda_m(F_x)| d\lambda_n = 0$$

Das impliziert, dass $\lambda_m(G_x) = \lambda_m(F_x)$ für λ-fast alle $x \in \mathbb{R}^n$. Wir setzen

$$N = \{x \in \mathbb{R}^n : \lambda_m(G_x) = \infty\} \cup \{x \in \mathbb{R}^n : \lambda_m(F_x) < \lambda_m(G_x)\}$$

Dann ist $N \in \mathcal{M}_{\lambda^*_n}$ mit $\lambda_n(N) = 0$. Für $x \in N^c$ gilt $\lambda_m(G_x \backslash F_x) = 0$, und damit E_x in der Vervollständigung von $\mathcal{B}(\mathbb{R}^m)$; d.h. $E_x \in \mathcal{M}_{\lambda^*_m}$. Das zeigt a).

Wir setzen $f(x) = \lambda_m(E_x)$ für $x \in N^c$ und $f(x) = 0$ für $x \in N$. Bemerke, dass $f(x) = \lambda_m(F_x)$ für alle $x \in N^c$. Aus der $\mathcal{B}(\mathbb{R}^m)$-Messbarkeit von $\lambda_m(F_x)$ finden wir, für alle $t > 0$,

$$\{x \in \mathbb{R}^n : f(x) < t\} = N \cup (N^c \cap \{x \in \mathbb{R}^n : \lambda_m(F_x) < t\}) \in \mathcal{M}_{\lambda^*_n}$$

Für $t \leq 0$, gilt analog

$$\{x \in \mathbb{R}^n : f(x) < t\} = N^c \cap \{x \in \mathbb{R}^n : \lambda_m(F_x) < t\} \in \mathcal{M}_{\lambda^*_n}$$

Das zeigt, dass f $\mathcal{M}_{\lambda^*_n}$-messbar ist.

Um c) zu zeigen, bemerken wir, dass

$$\lambda_{m+n}(E) = \lambda_{m+n}(F) = \int_{\mathbb{R}^n} \lambda_m(F_x) d\lambda_n = \int_{\mathbb{R}^n} f d\lambda_n$$

weil $f(x) = \lambda_m(F_x)$ für fast alle $x \in \mathbb{R}^n$.

Um den Fall $\lambda_{n+m}(E) = \infty$ zu betrachten, benutzt man wie üblich die σ-Endlichkeit vom Lebesgue Mass. Wir lassen die Details als Übung. □

Mit Satz 2.36 kann man dann Resultate analog zu Satz 2.34 und Theorem 2.35 beweisen. Sei nämlich $f : \mathbb{R}^{n+m} \to [0; \infty]$ $\mathcal{M}_{\lambda^*_{n+m}}$-messbar. Dann existiert $N \in \mathcal{M}_{\lambda^*_n}$ mit $\lambda_n(N) = 0$ so, dass die Funktion $f_x : \mathbb{R}^m \to [0; \infty]$, definiert durch $f_x(y) = f(x, y)$, für alle $x \in N^c$ $\mathcal{M}_{\lambda^*_m}$-messbar ist. Die Funktion

$$I_f(x) = \begin{cases} \int_{\mathbb{R}^m} f_x d\lambda_m & \text{für alle } x \in N^c \\ 0 & \text{für alle } x \in N \end{cases}$$

ist $\mathcal{M}_{\lambda_n^*}$-messbar und

$$\int_{\mathbb{R}^{n+m}} f d\lambda_{n+m} = \int_{\mathbb{R}^n} I_f d\lambda_n$$

Ist dagegen $f : \mathbb{R}^{n+m} \to [-\infty; \infty]$ $\mathcal{M}_{\lambda_{n+m}^*}$-messbar und bezüglich λ_{n+m}-integrierbar, so existiert $N \in \mathcal{M}_{\lambda_n^*}$ mit $\lambda_n(N) = 0$ so, dass die Funktion $f_x : \mathbb{R}^m \to [-\infty; \infty]$ für alle $x \in N^c$ integrierbar ist. Die Funktion

$$I_f(x) = \begin{cases} \int_{\mathbb{R}^m} f_x d\lambda_m & \text{für alle } x \in N^c \\ 0 & \text{für alle } x \in N \end{cases}$$

ist $\mathcal{M}_{\lambda_n^*}$-messbar und

$$\int_{\mathbb{R}^{n+m}} f d\lambda_{n+m} = \int_{\mathbb{R}^n} I_f d\lambda_n$$

Diese letzte Aussage ist mit der Aussage von Theorem 2.35 identisch (einzige Unterschied: dort war die Nullmenge N nötig, um Integrierbarkeit von f_x zu garantieren; hier braucht man die Nullmenge N schon um die Messbarkeit von f_x zu garantieren). D.h. das Theorem von Fubini kann tats achlich auch benutzt werden, um Integrale von Lebesgue messbare Funktionen anwenden, obwohl $\mathcal{M}_{\lambda_{n+m}^*} \neq \mathcal{M}_{\lambda_n^*} \times \mathcal{M}_{\lambda_m^*}$.

2.6 Transformationsatz

Das Theorem von Fubini erlaubt uns Integrale auf $\mathbb{R}^n$ auf Integrale auf $\mathbb{R}$ zurückzuführen. Manchmal ist aber die Anwendung von Fubini keine gute Idee, weil man verpasst dabei dann die Symmetrien des Integrandes. Das haben wir zum Beispiel in der Berechnung von dem Volumen einer Kugel gesehen. In diesen Fällen ist dagegen oft nützlich Koordinaten einzuführen, die die Symmetrie berücksichtigen. Zum Beispiel, wenn wir das Volumen einer Kugel in drei Dimensionen berechnen wollen, lohnt sich sogenannte sphährische Koordinaten (r, θ, φ) zu benutzen. Die sind aus den Beziehungen

$$x = r \cos\theta \cos\varphi, \quad y = r \cos\theta \sin\varphi, \quad z = r \sin\theta$$

definiert. Man kann dann sich einfach überzeugen, dass die Abbildung $\varphi : (0; \infty) \times [0; \pi] \times [0; 2\pi) \to \mathbb{R}^3 \backslash \{0\}$ bijektiv ist. In sphärische Koordinaten ist dann der Kugel $\{(x, y, z) : x^2 + y^2 + z^2 \leq 1\}$ einfach aus der Bedingung $r \leq 1$ definiert. Aus Analysis 1 erinnern wir, dass wenn wir in einem Riemann Integral Variablen wechseln wollen, so müssen wir die Substitutionregel angewenden. Für das Lebesgue Integral auf $\mathbb{R}^n$ wird die Substitutionsregel durch die folgende Transformationssatz ersetzt.

Satz 2.37 (Transformationssatz)**.** *Sei $n \in \mathbb{N}\backslash\{0\}$, $U, V \subset \mathbb{R}^n$ offen und $\varphi : U \to V$ ein C^1-Diffeomorphismus (d.h. $\varphi \in C^1(U; V)$ ist bijektiv, und $\varphi^{-1} \in C^1(V; U)$). Eine Funktion $f : V \to [-\infty; \infty]$ ist genau dann Lebesgue integrierbar (d.h. f ist $\mathcal{M}_{\lambda_n^*}$ messbar und integrierbar bezüglich dem Lebesgue Mass λ_n), wenn $(f \circ \varphi)|\det(D\varphi)| : U \to [-\infty; \infty]$ Lebesgue integrierbar. Es gilt dann*

$$\int_V f d\lambda_n = \int_U (f \circ \varphi)|\det(D\varphi)| d\lambda_n$$

Bemerkungen:

i) Da φ ein Diffeomorphismus ist, haben wir $0 < |\det(D\varphi(x))| < \infty$ für alle $x \in U$. Deswegen ist das Produkt $f(\varphi(x))|\det(D\varphi(x))|$ wohldefiniert, für alle $x \in U$.

ii) Insbesondere, für $E \in \mathcal{M}_{\lambda_n^*}$ finden wir, dass $\varphi(E) \in \mathcal{M}_{\lambda_n^*}$ und

$$\lambda_n(\varphi(E)) = \int_E |\det(D\varphi)| d\lambda_n$$

iii) Sind $U, V \subset \mathbb{R}$, dann ist $D\varphi(x) = \varphi'(x)$ und $|\det(D\varphi(x))| = |\varphi'(x)|$, was mit der Substitutionsregel für Riemann Integrale auf $\mathbb{R}$ konsistent ist.

Zum Beweis vom Transformationssatz brauchen wir einigen Zwischenresultate. Erstens, möchten wir im nächsten Lemma zeigen, dass Lipschitz-stetige Abbildungen messbare Mengen in messbare Mengen abbilden.

Lemma 2.38. *Sei $E \subset \mathbb{R}^n$ mit $E \in \mathcal{M}_{\lambda_n^*}$ und $f : \mathbb{R}^n \to \mathbb{R}^n$ Lipschitz stetig. Dann ist $f(E) \in \mathcal{M}_{\lambda_n^*}$ und*

$$\lambda_n(f(E)) \leq (2L)^n \lambda_n(E) \tag{33}$$

Hier ist $L > 0$ so gewählt, dass $\|f(x)-f(y)\|_\infty \leq L\|x-y\|_\infty$, mit $\|x\|_\infty = \max_{j=1,\dots,n} |x_j|$.

Beweis. Wir zeigen, dass $f(E)$ messbar ist, für alle $E \in \mathcal{M}_{\lambda_n^*}$. Sei zunächst $E \subset \mathbb{R}^n$ kompakt. Dann ist auch $f(E)$ kompakt, weil stetige Abbildungen kompakte Mengen in kompakten Mengen abbilden. In der Tat, falls y_n eine beliebige Folge in $f(E)$ ist, so existiert $x_n \in E$ mit $y_n = f(x_n)$. Da E kompakt ist, existiert eine Teilfolge x_{n_j} und $x \in E$ mit $x_{n_j} \to x$ für $j \to \infty$; aus Stetigkeit von f folgt, dass $y_{n_j} = f(x_{n_j}) \to f(x) \in f(E)$. D.h. jede Folge in $f(E)$ hat eine in $f(E)$ konvergente Teilfolge. Mit anderen Wörter $f(E)$ ist kompakt (Folgenkompaktheit ist auf metrische Räume, insbesondere auf $\mathbb{R}^n$ mit Überdeckungskompaktheit äquivalent). Sei nun E eine F_σ-Menge, d.h. es existiere eine Folge F_i von abgeschlossenen Mengen, mit $E = \cup_{i\in\mathbb{N}} F_i$. Für jede $i, k \in \mathbb{N}$, definieren wir $F_{i,k} = F_i \cap \{x \in \mathbb{R}^n : \|x\| \leq k\}$. Dann ist $F_{i,k}$ kompakt für alle $i, k \in \mathbb{N}$, und $E = \cup_{i,k\in\mathbb{N}} F_{i,k}$. Dann gilt $f(E) = \cup_{i,k\in\mathbb{N}} f(F_{i,k})$. Da $f(F_{i,k})$ kompakt ist, für alle $i, k \in \mathbb{N}$ ist $f(E)$ auch eine F_σ Menge. Sei nun $E \in \mathcal{M}_{\lambda_n^*}$ beliebig. Aus den Übungen finden wir $N \in \mathcal{M}_{\lambda_n^*}$ mit $\lambda_n(N) = 0$ und B eine F_σ-Menge so, dass $E = B \cup N$. Dann ist $f(E) = f(B) \cup f(N)$. Da $f(B)$ eine F_σ Menge ist (und deswegen sicher Lebesgue messbar), es genügt zu zeigen, dass $f(N) \in \mathcal{M}_{\lambda_n^*}$. Wir zeigen, dass für alle $A \subset \mathbb{R}^n$,

$$\lambda_n^*(f(A)) \leq (2L)^n \lambda_n^*(A) \tag{34}$$

Wenn wir (34) zeigen können, finden wir $\lambda_n^*(f(N)) \leq CL^n \lambda_n^*(N) = 0$. Damit ist $f(N) \in \mathcal{M}_{\lambda_n^*}$. Das zeigt zunächst, dass $f(E) \in \mathcal{M}_{\lambda_n^*}$ für alle $E \in \mathcal{M}_{\lambda_N^*}$. Weiter, da das äussere Mass λ_n^* auf messbaren Mengen mit dem Lebesgue Mass λ_n übereinstimmt, zeigt (34) auch (33).

Offenbar genügt es (34) für $A \subset \mathbb{R}^n$ mit $\lambda_n^*(A) < \infty$ zu zeigen. Sei $\varepsilon > 0$ fest. Unter der Annahme $\lambda_n^*(A) < \infty$ finden wir dann eine Folge $\{R_j\}_{j\in\mathbb{N}}$ von offenen und beschränkten n-dimensionale Intervalle so, dass $A \subset \cup_{j=1}^\infty R_j$ und

$$\lambda_n^*(A) \geq \sum_{j\in\mathbb{N}} \lambda_n^*(R_j) - \varepsilon$$

Für alle $j \in \mathbb{N}$ finden wir dann eine Folge Q_{ji} von disjunkten Würfel mit $R_j = \cup_i Q_{ji}$. Dann gilt $A \subset \cup_{i,j\in\mathbb{N}} Q_{ji}$ und

$$\lambda_n^*(A) \geq \sum_{i,j\in\mathbb{N}} \lambda_n^*(Q_{ji} - \varepsilon$$

Sei nun K_{ji} die Kantenlänge von Q_{ji}. Dann ist $\|x-y\|_\infty \leq K_{ji}$ für alle $x, y \in Q_{ji}$. Deswegen $\|f(x) - f(y)\|_\infty \leq LK_{ji}$ für alle $x, y \in Q_{ji}$. D.h. $f(Q_{ji})$ ist in eine Würfel mit Kantenlänge $2LK_{ji}$ enthalten. Damit ist

$$\lambda_n^*(f(Q_{ji})) \leq (2L)^n K_{ji}^n = (2L)^n \lambda_n^*(Q_{ji})$$

für alle $i, j \in \mathbb{N}$. Das impliziert, dass

$$\lambda_n^*(f(A)) = \lambda_n^* \left(\bigcup_{i,j\in\mathbb{N}} f(Q_{ji}) \right) \leq \sum_{i,j\in\mathbb{N}} \lambda_n^*(f(Q_{ji})) \leq (2L)^n \sum_{i,j\in\mathbb{N}} \lambda_n^*(Q_{ji}) \leq (2L)^n \left(\lambda_n^*(A) + \varepsilon\right)$$

Da $\varepsilon > 0$ beliebig ist, finden wir (34). $\square$

Wir erhalten das folgende Korollar.

Korollar 2.39. *Seien $U, V \subset \mathbb{R}^n$ offen, $N \subset U$ eine Nullmenge und $T : U \to V$ stetig differenzierbar. Dann ist $T(N)$ eine Nullmenge.*

Beweis. Man kann $U \subset \bigcup_{j\in\mathbb{N}} Q_k$ für eine Folge kompakten n-dimensionalen Intervalle Q_k schreiben (das folgt z.B. aus Lemma 1.32). Die Einschränkung von T auf jeder kompakten Menge Q_k ist Lipschitz stetig. Da $T(N) \subset \cup_k T(N \cap Q_k)$ und $\lambda_n(T(N \cap Q_k)) = 0$, finden wir aus der Subadditivität, dass

$$\lambda_n(T(N)) \leq \sum_k \lambda_n(T(N \cap Q_k)) = 0$$

$\square$

Das Korollar impliziert auch, dass Diffeomorphismen Lebesgue messbaren Mengen in Lebesgue messbaren Mengen abbilden.

Satz 2.40. *Seien $U, V \subset \mathbb{R}^n$ offen, $A \subset U$ Lebesgue messbar und $T : U \to V$ ein Diffeomorphismus. Dann ist $T(U)$ Lebesgue messbar.*

Zum Beweis von Satz 2.40 brauchen wir das folgende Lemma.

Lemma 2.41. *Seien $U, V \subset \mathbb{R}^n$ offen, und $T : U \to V$ stetig. Wir bezeichnen mit $\mathcal{B}(U), \mathcal{B}(V)$ die Borel σ-Algebren auf U, V (definiert durch die Spur $\mathcal{B}(U) = \mathcal{B}(\mathbb{R}^n) \cap U$). Sei $A \in \mathcal{B}(V)$. Dann ist $T^{-1}(A) \in \mathcal{B}(U)$.*

Beweis. Es wurde in den Übungen gezeigt, dass der Push-Forward von $\mathcal{B}(U)$ bezüglich T eine σ-Algebra auf V ist (der Push-Forward wird als $\{B \subset V : T^{-1}(B) \in \mathcal{B}(U)\}$ definiert). Da T stetig ist, enthält der Push-Forward alle offene Mengen. Da $\mathcal{B}(V)$ die kleinste σ-Algebra ist, die alle offene Menge enthält, muss $T^{-1}(B) \in \mathcal{B}(U)$ für alle $B \in \mathcal{B}(V)$. $\square$

Es folgt aus dem Lemma, dass für ein Diffeomorphismus $T : U \to V$ und $A \in \mathcal{B}(U)$, es gilt $T(A) \in \mathcal{B}(V)$ (weil T^{-1} stetig ist). Damit können wir Satz 2.40 beweisen.

Beweis von Satz 2.40. Da das Lebesgue Mass auf $\mathcal{M}_{\lambda_n^*}$ die Vervollständigung vom Lebesgue Mass auf $\mathcal{B}(\mathbb{R}^n)$ ist, finden wir $B \in \mathcal{B}(\mathbb{R}^n)$ so, dass $\lambda_n(A \backslash B) = 0$. $T(B) \in \mathcal{B}(\mathbb{R}^n)$ folgt aus Lemma 2.41. $\lambda_n(T(A \backslash B)) = 0$ folgt dagegen aus Korollar 2.39; damit ist $T(A \backslash B) \in \mathcal{M}_{\lambda_n^*}$ und

$$T(A) = T(B) \cup T(A \backslash B) \in \mathcal{M}_{\lambda_n^*}$$

□

Im nächsten Theorem unteresuchen wir, wie Lebesgue Mass unter lineare Abbildungen transformiert.

Theorem 2.42. *Sei $T : \mathbb{R}^n \to \mathbb{R}^n$ linear, und $A \in \mathcal{M}_{\lambda_n^*}$. Dann ist $T(A) \in \mathcal{M}_{\lambda_n^*}$ und*

$$\lambda_n(T(A)) = |det\, T|\, \lambda_n(A)$$

Beweis. Ist $\det T = 0$, dann ist $T(A)$ eine Teilmenge eines $(n-1)$-dimensionalen Unterraum von $\mathbb{R}^n$. In diesem Fall ist $T(A)$ offenbar eine Nullmenge. Wir können also annehmen, dass T ein Automorphismus ist, mit $\det T \neq 0$.

Die Tatsache, dass $T(A) \in \mathcal{M}_{\lambda_n^*}$ folgt dann aus Satz 2.40, weil jede Automosrphismus auch ein Diffeomorphismus ist. Es bleibt zu zeigen, dass $\lambda_n(T(A)) = |\det T| \lambda(A)$. Dazu setzen wir

$$\mu_T(A) := \lambda(T(A))$$

für alle $A \in \mathcal{M}_{\lambda_n^*}$. Dann ist μ_T offenbar translationsinvariant auf $\mathcal{M}_{\lambda_n^*}$. Aus der Bemerkung nach Satz 1.35 folgt, dass eine Konstante $\alpha_T > 0$ existiert, mit $\mu_T(A) = \alpha_T \lambda(A)$ für alle $A \in \mathcal{M}_{\lambda_n^*}$. Um den Satz zu beweisen, brauchen wir nur zu zeigen, dass $\alpha_T = |\det T|$. Wir werden diese Behauptung zunächst für besondere T beweisen.

i) Permutationen. Sei σ eine Permutation der Menge $\{1, \dots, n\}$ und $T : \mathbb{R}^n \to \mathbb{R}^n$ definiert durch $Te_i = e_{\sigma(i)}$, wobei $e_1, \dots, e_n$ die Standard-Basis von $\mathbb{R}^n$ ist. Dann gilt offenbar, dass $T((0;1)^{\times n}) = (0;1)^{\times n}$ und deswegen, dass $\alpha_T = 1 = \det T$.

ii) Skalierungen. Für $t \neq 0$ sei $S(t) : \mathbb{R}^n \to \mathbb{R}^n$ definiert durch $S(t)e_1 = te_1$ und $S(t)e_i = e_i$ für alle $i = 2, \dots, n$. Dann gilt $S(t)(0;1)^{\times n} = (0;t) \times (0;1) \times \cdots \times (0;1)$ und deswegen

$$\lambda_n(S(t)(0;1)^{\times n}) = t = t\lambda_n((0;1)^{\times n})$$

Also, $\alpha_{S(t)} = t = \det\, S(t)$.

iii) Verschiebungen. Für $n \geq 2$, sei $Q : \mathbb{R}^n \to \mathbb{R}^n$ definiert durch $Qe_1 = e_1 + e_2$ und $Qe_j = e_j$ für alle $j = 2, \dots, n$. Dann gilt

$$Q((0;1)^{\times n}) = \{(s; s+t) : s, t \in (0;1)\} \times (0;1)^{\times (n-2)}$$

und deswegen

$$\lambda_n(Q((0;1)^{\times n})) = \lambda_2\left(\{(s; s+t) : s, t \in (0;1)\}\right) = 1 = \det Q$$

Die Tatsache, dass $\lambda_2(\{(s; s+t) \in \mathbb{R}^2 : s, t \in (0;1)\}) = 1$ kann einfach geometrisch überprüft werden, durch Benutzung der Translationsinvarianz von λ_2.

Die Aussage $\alpha_T = \det T$ folgt für allgemeinen lineare Automorphismen T, weil jeder T als endlichen Produkt von Matrizen der Form i),ii),iii) geschrieben werden kann (Gauss'sche Eliminationsverfahren). Ist also $T = T_1 \cdot \dots \cdot T_k$, so gilt offenbar $\alpha_T = \prod_{j=1}^k \alpha_{T_j}$ und $\det T = \prod_{j=1}^k \det T_j$. Die Gleichung $\alpha_T = \det T$ folgt also aus $\alpha_{T_j} = \det T_j$ für alle $j = 1, \dots, k$. □

Das nächste Lemma spielt eine wichtige Rolle im Beweis vom Transformationssatz 2.37.

Lemma 2.43. *Seien $U, V \subset \mathbb{R}^n$ offen und $\varphi : U \to V$ ein C^1-Diffeomorphismus. Sei $W \subset U$ ein kompakter Quader. Dann ist $\varphi(W)$ Lebesgue messbar (weil Homeomorphismen offene Menge auf offene Menge abbilden, und deswegen auch abgeschlossene Menge auf abgeschlossene) und*

$$\lambda_n(\varphi(W)) \leq \int_W |\det(D\varphi)| d\lambda_n$$

Beweis. Da $W \subset \mathbb{R}^n$ kompakt ist, gilt

$$M = \max_{x \in W} \|(D\varphi(x))^{-1}\| < \infty .$$

(weil $D\varphi(x)$ von x stetig abhängt, und $D\varphi(x)$ für alle $x \in W$ invertierbar ist). Sei nun $\varepsilon > 0$ beliebig. Da W kompakt ist, finden wir $\delta > 0$ mit $\|D\varphi(x) - D\varphi(y)\| \leq \varepsilon$ für alle $x, y \in W$ mit $\|x - y\| \leq \delta$; hier bezeichnet $\|.\|$ die euklidische Norm auf $\mathbb{R}^n$ (oder irgendeine Norm auf $\mathbb{R}^n$, weil alle äquivalent sind).

Wir zerlegen nun $W = \bigcup_{j=1}^k Q_j$, wobei Q_j eine endliche Familie von disjunkten Quadern mit Kantenlänge $\ell \leq \delta/\sqrt{n}$ (so, dass $\|x - y\| \leq \delta$ für alle $x, y \in Q_j$, für alle $j = 1, \dots, k$). Wir behaupten nun, dass

$$\lambda_n(\varphi(Q_j)) \leq (1 + 2M\varepsilon\sqrt{n})^n \int_{Q_j} |\det(D\varphi)| d\lambda_n \tag{35}$$

für alle $j = 1, \dots, k$. Da $\varphi(W) = \bigcup_{j=1}^k \varphi(Q_j)$, mit $\varphi(Q_j) \cap \varphi(Q_i) = \emptyset$ für alle $i \neq j$ folgt aus (35) und aus der Linearität des Integrals, dass

$$\begin{aligned}
\lambda_n(\varphi(W)) &= \sum_{j=1}^k \lambda(\varphi(Q_j)) \\
&\leq (1 + 2M\varepsilon\sqrt{n})^n \int_{Q_j} |\det(D\varphi)| d\lambda_n \\
&= (1 + 2M\varepsilon\sqrt{n})^n \int_W |\det(D\varphi)| d\lambda_n
\end{aligned}$$

Da $\varepsilon > 0$ beliebig ist, es folgt, dass

$$\lambda_n(\varphi(W)) \leq \int_W |\det(D\varphi)| d\lambda_n$$

wie behauptet.

Wir müssen noch (35) zeigen. Sei $z \in \mathbb{R}^n$ mit $\overline{Q}_j = z + [0;\ell]^n$. Da $\overline{Q}_j$ kompakt ist, finden wir $\overline{x} \in \overline{Q}_j$ mit

$$|\det(D\varphi(\overline{x}))| = \min_{x \in \overline{Q}_j} |\det(D\varphi(x))|$$

Da $D\varphi$ überall invertierbar ist, gilt $\det(D\varphi(\overline{x})) \neq 0$. Wir definieren nun $\psi : \overline{Q}_j \to \mathbb{R}^n$ durch

$$\psi(y) = (D\varphi(\overline{x}))^{-1}(\varphi(y) - \varphi(z))$$

für alle $y \in \overline{Q}_j$. Dann gilt $\psi(z) = 0$ und, aus der Kettenregel,

$$D\psi(y) = (D\varphi(\overline{x}))^{-1} D\varphi(y).$$

Da $\|D\varphi(y) - D\varphi(x)\| \leq \varepsilon$ für alle $x, y \in \overline{Q}_j$ (nach Wahl von δ), finden wir

$$\|D\psi(y) - 1\| = \|(D\varphi(\overline{x}))^{-1} D\varphi(y) - 1\| \leq \|D\varphi(\overline{x})^{-1}\| \|D\varphi(y) - D\varphi(\overline{x})\| \leq M\varepsilon$$

für alle $y \in \overline{Q}_j$. Nun wenden wir den Mittelwertsatz an. Für ein beliebiges $y \in Q_j$ und für jede Komponente $i = 1, \dots, n$ finden wir ein q_i auf der Geradenstück zwischen y und z mit

$$\psi_i(y) = \psi_i(y) - \psi_i(z) = [D\psi(q_i)(y - z)]_i$$

Deswegen gilt

$$\psi_i(y) - (y - z)_i = [(D\psi(q_i) - 1)(y - z)]_i \, .$$

Da $0 \leq (y - z)_i \leq \ell$ und

$$|[(D\psi(q_i) - 1)(y - z)]_i| \leq \|(D\psi(q_i) - 1)(y - z)\| \leq \|D\psi(q_i) - 1\| \|y - z\| \leq M\varepsilon\sqrt{n}\ell$$

finden wir

$$-M\varepsilon\sqrt{n}\ell \leq \psi_i(y) \leq \ell + M\varepsilon\sqrt{n}\ell$$

für alle $y \in \overline{Q}_j$. Das impliziert, dass

$$\psi(Q_j) \subset [-M\varepsilon\sqrt{n}\ell; \ell + M + \varepsilon\sqrt{n}\ell]^{\times n}$$

und

$$\lambda_n(\psi(Q_j)) \leq (1 + 2M\varepsilon\sqrt{n})^n \ell^n$$

Translationsinvarianz von Lebesgue Mass λ_n zusammen mit Theorem 2.42 implizieren, dass

$$\begin{aligned} \lambda_n(\varphi(Q_j)) &= \lambda_n(D\varphi(\overline{x})(\psi(Q_j))) = |\det D\varphi(\overline{x})| \lambda_n(\psi(Q_j)) \\ &\leq (1 + 2M\varepsilon\sqrt{n})^n |\det(D\varphi(\overline{x}))| \lambda_n(Q_j) \leq (1 + 2M\varepsilon\sqrt{n})^n \int_{Q_j} |\det(D\varphi)| d\lambda_n \end{aligned}$$

In der letzten Ungleichung haben wir benutzt, dass $|\det D\varphi(\overline{x})| \leq |\det(D\varphi(x))|$ für alle $x \in Q_j$. Damit ist (35) gezeigt. $\square$

Wir können nun Lemma 2.43 anwenden, um Satz 2.37 zu beweisen.

Beweis von Satz 2.37. Es genug die Behauptung für $f : V \to [0; \infty]$ nicht-negativ und messbar zu zeigen. Dann kann man eine beliebige messbare Funktion $f : V \to [-\infty; \infty]$ als $f = f_+ - f_-$ schreiben. Der Transformationssatz für f_+ und für f_- impliziert den Transformationssatz für f.

Wir nehmen zunächst zusätzlich an, dass $D\varphi$ und $(D\varphi)^{-1}$ beschränkt sind. In diesem Fall sind φ und φ^{-1} Lipschitz stetig.

Sei $O \subset U$ offen. Aus Lemma 1.32 finden wir eine Folge $(Q_j)_{j\in\mathbb{N}}$ von disjunkten halb-offenen Quadern mit $O = \bigcup_{j\in\mathbb{N}} Q_j$. Dann gilt $\varphi(O) = \bigcup_{j\in\mathbb{N}} \varphi(Q_j)$ und damit, aus Lemma 2.43,

$$\lambda_n(\varphi(O)) \leq \sum_{j\in\mathbb{N}} \lambda_n(\varphi(Q_j)) \leq \sum_{j\in\mathbb{N}} \lambda_n(\varphi(\overline{Q}_j)) \leq \sum_{j\in\mathbb{N}} \int_{Q_j} |\text{det} D\varphi| d\lambda_n = \int_O |\text{det} D\varphi| d\lambda_n$$

Sei nun $E \subset U$ Lebesgue messbar. Aus der Regularität vom Lebesgue Mass, Satz 1.31, finden wir für alle $j \in \mathbb{N}$ eine offene Menge O_j mit $E \subset O_j$ und $\lambda_n(O_j \backslash E) \leq 1/j$. Wir können auch $O_j \subset U$ annehmen (sonst ersetzen wir O_j mit $O_j \cap U$). Da φ Lipschitz stetig ist, sind $\varphi(E)$ und $\varphi(O_j)$ Lebesgue messbar, und

$$\lambda_n(\varphi(E)) \leq \lambda_n(\varphi(O_j)) \leq \int_{O_j} |\det(D\varphi)| d\lambda_n \leq \int_E |\det(D\varphi)| d\lambda_n + M^n \lambda_n(O_j \backslash E)$$

wobei $M = \sup_{x\in U} \|D\varphi(x)\| < \infty$. Da $\lambda_n(O_j \backslash E) \leq 1/j$ und $j \geq 1$ beliebig ist, bekommen wir

$$\lambda_n(\varphi(E)) \leq \int_E |\det(D\varphi)| d\lambda_n$$

für alle Lebesgue messbare $E \in \mathcal{M}_{\lambda_n^*}$.

Sei nun $f = \sum_{j=1}^m \alpha_j \chi_{A_j}$ eine einfache Funktion, mit $A_j \subset V = \varphi(U)$ messbar und disjunkt, $\alpha_j \in [0; \infty)$. Die Mengen $B_j = \varphi^{-1}(A_j) \subset U$ sind auch messbar und disjunkt und

$$\begin{aligned} \int_{\varphi(U)} f d\lambda_n &= \sum_{j=1}^m \alpha_j \lambda_n(A_j) = \sum_{j=1}^m \alpha_j \lambda_n(\varphi(B_j)) \leq \sum_{j=1}^m \alpha_j \int_{B_j} |\det(D\varphi)| d\lambda_n \\ &= \int_U \sum_{j=1}^m \alpha_j \chi_{B_j} |\det(D\varphi)| d\lambda_n = \int_U (f \circ \varphi) \, |\det(D\varphi)| d\lambda_n \end{aligned}$$

Ist nun $f : V = \varphi(U) \to [0; \infty]$ eine beliebige messbare Funktion, so können wir nach Proposition 2.3 eine Folge f_k von nicht-negativen einfachen messbaren Funktionen auf $\varphi(U)$, mit $f_{k+1}(x) \geq f_k(x)$ für alle $k \in \mathbb{N}$ und $f(x) = \lim_{k\to\infty} f_k(x)$ für alle $x \in V = \varphi(U)$. Dann konvergiert die Funktionenfolge $(f_k \circ \varphi)|\det(D\varphi)|$, definiert auf U mit Werten in $[0; \infty]$, punktweise und monoton gegen $g = (f \circ \varphi)|\det(D\varphi)|$. Aus der Satz über die monotone Konvergenz gilt

$$\int_{\varphi(U)} f d\lambda_n = \lim_{k\to\infty} \int_{\varphi(U)} f_k d\lambda_n \leq \lim_{k\to\infty} \int_U (f_k \circ \varphi)|\det(D\varphi)| d\lambda_n = \int (f \circ \varphi)|\det(D\varphi)| d\lambda_n$$

Damit gilt

$$\int_{\varphi(U)} f d\lambda_n \leq \int_U (f \circ \varphi)|\det(D\varphi)| d\lambda_n \tag{36}$$

für alle messbare Funktionen $f : V \to [0;\infty]$.

Mit $\psi = \varphi^{-1}$ und $g = (f \circ \varphi)|\det(D\varphi)|$ finden wir also

$$\int_{\psi(V)} g \, d\lambda_n \leq \int_V (g \circ \psi)|\det(D\psi)| d\lambda_n$$

Da aber $\psi(V) = U$ und $(g \circ \psi)|\det(D\psi)| = f$ finden wir

$$\int_U (f \circ \varphi)|\det(D\varphi)| d\lambda_n \leq \int_V f d\lambda_n$$

Mit (36) folgt, dass

$$\int_U (f \circ \varphi)|\det(D\varphi)| d\lambda_n = \int_V f d\lambda_n$$

für alle messbare Funktionen $f : V \to [0;\infty]$.

Schlussendlich nehmen wir an $\varphi : U \to V$ sei ein Diffeomorphismus, aber nicht unbedingt, dass $D\varphi$ und $D\varphi^{-1}$ beschränkt sind (dann brauchen φ,φ^{-1} nicht unbedingt Lipschitz stetig zu sein). Für $k \in \mathbb{N}$ definieren wir die Mengen

$$U_k = \{x \in U : \|D\varphi(x)\| + \|(D\varphi(x))^{-1}\| < k\}$$

Aus der Stetigkeit von $D\varphi$ ist U_k offen, für alle $k \in \mathbb{N}$. Aus dem ersten Teil des Beweises finden wir, dass

$$\int_{\varphi(U_k)} f d\lambda_n = \int_{U_k} (f \circ \varphi)|\det(D\varphi)| d\lambda_n$$

für alle $k \in \mathbb{N}$. Wir betrachten die monoton wachsende Funktionenfolge $f_k = f\chi_{\varphi(U_k)}$, definiert auf $V = \varphi(U)$, und $g_k = (f \circ \varphi)|\det(D\varphi)|\chi_{U_k}$, definiert auf U. Die Folge f_k konvergiert punktweise zu f, g_k zu $(f \circ \varphi)|\det(D\varphi)|$. Aus der Satz über die monotone Konvergenz finden wir

$$\begin{aligned}\int_{\varphi(U)} f d\lambda_n &= \lim_{k\to\infty} \int_{\varphi(U)} f_k d\lambda_n = \lim_{k\to\infty} \int_{\varphi(U_k)} f d\lambda_n \\ &= \lim_{k\to\infty} \int_{U_k} (f \circ \varphi)|\det(D\varphi)| d\lambda_n = \lim_{k\to\infty} \int_U g_k = \int_U (f \circ \varphi)|\det(D\varphi)| \,.\end{aligned}$$

□

Beispiel:

i) *Polarkoordinaten auf* $\mathbb{R}^2$*:* Wir definieren die Abbildung $\varphi : (0;\infty) \times (0;2\pi) \to \mathbb{R}^2\backslash\{(x,0) : x \geq 0\}$ durch $\varphi(r,\theta) = (r\cos\theta, r\sin\theta)$. Es ist leicht zu überprüfen, dass φ ein C^1-Diffeomorphismum ist. Es gilt:

$$D\varphi(r,\theta) = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$$

und deswegen

$$\det D\varphi(r,\theta) = r$$

Da $\{(x;0) : x \geq 0\}$ eine Nullmenge in $\mathbb{R}^2$ ist, gilt, für eine beliebige Lebesgue messbare Funktion $f : \mathbb{R}^2 \to [-\infty;\infty]$, dass

$$\begin{aligned}\int_{\mathbb{R}^2} f d\lambda_2 &= \int_{(0;\infty)\times(0,2\pi)} f(r\cos\theta, \sin\theta) r d\lambda_2(r,\theta) \\ &= \int_0^\infty \left[\int_0^{2\pi} f(r\cos\theta, r\sin\theta) d\theta\right] r dr\end{aligned}$$

Als Anwendung von Polarkoordinaten auf $\mathbb{R}^2$ berechnen wir das integral

$$I = \int_{\mathbb{R}} e^{-x^2} dx$$

Es gilt, mit Fubini,

$$\begin{aligned}I^2 &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} e^{-y^2} dy\right) e^{-x^2} dx = \int_{\mathbb{R}^2} e^{-(x^2+y^2)} d\lambda_2(x,y) = \int_0^\infty e^{-r^2} r dr \int_0^{2\pi} d\theta \\ &= 2\pi \left[-\frac{1}{2} e^{-r^2}\right]_0^\infty = \pi\end{aligned}$$

Deswegen ist $I = \sqrt{\pi}$.

ii) *Sphärische Koordinaten auf* $\mathbb{R}^3$. Wir setzen $U = (0;\infty) \times (0;\pi) \times (0;2\pi)$ und $V = \mathbb{R}^3 \backslash \{(0,y,z) : y,z \in \mathbb{R}\}$. Wir definieren die Abbildung $\varphi : U \to V$ durch $\varphi(r,\theta,\phi) = (r\sin\theta\cos\phi, r\sin\theta\sin\phi, r\cos\theta)$. Man überprüft einfach, dass φ ein Diffeomorphismus ist. Es gilt

$$D\varphi(r,\theta,\phi) = \begin{pmatrix} \sin\theta\cos\phi & r\cos\theta\cos\phi & -r\sin\theta\sin\phi \\ \sin\theta\sin\phi & r\cos\theta\sin\phi & r\sin\theta\cos\phi \\ \cos\theta & -r\sin\theta & 0 \end{pmatrix}$$

Deswegen finden wir

$$\det(D\varphi) = r^2 \sin\theta$$

Da $\det D\varphi(r,\theta,\varphi) > 0$ auf U und da $\mathbb{R}^3 \backslash V$ eine Nullmenge ist, finden wir, für eine beliebige Lebesgue messbare Funktion $f : \mathbb{R}^3 \to [-\infty;\infty]$,

$$\int_{\mathbb{R}^3} f d\lambda_3 = \int_U f(r\sin\theta\cos\phi, r\sin\theta\sin\phi, r\cos\theta) r^2 \sin\theta d\lambda_3(r,\theta,\phi)$$

Als Anwendung berechnen wir das Volumen eines Kugel mit Radius $R > 0$. In sphärische Koordinaten ist der Kugel von Radius R aus der Bedingung $r \leq R$ definiert. Das Volumen ist also aus

$$\int_U \chi(r \leq R) d\lambda_3(r,\theta,\phi) = 2\pi \int_0^R r^2 dr \int_0^\pi \sin\theta d\theta = \frac{4\pi R^3}{3}$$

gegeben.

3 L^p-Räume und ihre Eigenschaften

Wir führen in diesem Kapitel die L^p-Räume von messbaren Funtkionen auf einem Massraum $(\Omega, \mathcal{A}, \mu)$. Der Einfachheit halber werden hier nur Funktionen mit Werten in $\mathbb{R}$ betrachten. Wir hätten aber auch Funktionen mit Werten in $[-\infty; \infty]$ betrachten können, die fast überall endlich sind.

Wir hätten auch $\mathbb{C}$-wertige oder $\mathbb{R}^m$-wertige Funktionen betrachten können. Sei nämlich $(\Omega, \mathcal{A}, \mu)$ ein Massraum, und $f : \Omega \to \mathbb{R}^n$, Wir sagen f ist $\mathcal{A}$-messbar, falls jede Komponente f_j, $j = 1, \dots, n$, $\mathcal{A}$-messbar ist. Wir sagen die Funktion f ist integrierbar, bezüglich dem Mass μ, falls f_j integrierbar ist, für alle $j = 1, \dots, n$. In diesem Fall definieren wir das Integral

$$\int f d\mu = \left(\int f_1 d\mu, \int f_2 d\mu, \dots, \int f_n d\mu \right)$$

komponentenweise. Man bemerke, dass f genau dann integrierbar ist, falls die reelwertige Funktion $x \to \|f(x)\| = (|f_1(x)|^2 + \dots + |f_n(x)|^2)^{1/2}$ integrierbar ist. Analog, wir sagen eine $\mathbb{C}$-wertige Funktion $f : \Omega \to \mathbb{C}$ ist messbar, falls Re f und Im f messbar sind. f heisst integrierbar, falls Re f und Im f integrierbar sind. Das ist wieder äquivalent zur Bedingung, dass $|f| = (\mathrm{Re}^2(f) + \mathrm{Im}^2(f))^{1/2}$ integrierbar ist. In diesem Fall definieren wir das Integral

$$\int f d\mu = \int \mathrm{Re}\, f d\mu + i \int \mathrm{Im}\, f d\mu$$

Wie gesagt werden wir im folgenden Kapitel nur reelwertige Funktionen untersuchen. Es sollte aber klar sein, wie die Aussage unten zu $\mathbb{C}$- und $\mathbb{R}^n$-wertige Funktionen erweitert werden können.

3.1 Konvergenzbegriffe für Folgen messbarer Funktionen

Wir betrachten ein Massraum $(\Omega, \mathcal{A}, \mu)$, eine Funktionenfolge $f_k : \Omega \to \mathbb{R}$, und eine Funktion $f : \Omega \to \mathbb{R}$. Aus Analysis 1 kennen wir die Begriffe von punktweise und von gleichmässige Konvergenz. Wir sagen nämlich, dass $f_k \to f$ punktweise, falls $f_k(x) \to f(x)$ für alle $x \in \Omega$, und, dass $f_k \to f$ gleichmässig, falls

$$\sup_{x \in \Omega} |f_k(x) - f(x)| \to 0$$

für $k \to \infty$. Gleichmässige Konvergenz impliziert offenbar punktweise konvergenz. Diese beide Begriffe benutzen nicht die Tatsache, dass Ωein Massraum ist.

Wir haben in Analysis 3 schon einen weiteren Begriff von Konvergenz eingeführt. Wir sagen, dass $f_k \to f$ punktweise fast überall, (oder einfach, dass $f_k \to f$ fast überall), falls es ein $N \in \mathcal{A}$ existiert, mit $\mu(N) = 0$ so, dass $f_k(x) \to f(x)$ für alle $x \in N^c$ (ist das Mass μ vollständig, so konvergiert $f_k \to f$ punktweise fast überall genau dann, wenn $\mu(\{x \in \Omega : f_k(x) \not\to f(x)\}) = 0$). Fast überall konvergenz ist offenbar schwächer als punktweise Konvergenz, und deswegen auch als gleichmässige Konvergenz.

Sind alle f_k und f messbare Funktionen, so können wir noch ein Konvergenzbegriff einführen. Wir sagen, dass $f_k \to f$ nach Mass (oder in Wahrscheinlichkeit) falls, für alle

$\delta > 0$,

$$\lim_{k\to\infty} \mu\left(\{x \in \Omega : |f_k(x) - f(x)| > \delta\}\right) = 0$$

Man bemerke, dass die Mengen $\{x \in \Omega : |f_k(x) - f(x)| > \delta\}$ immer in $\mathcal{A}$ sind, wegen der Annahme, dass f und f_k für alle $k \in \mathbb{N}$ messbar sind.

Wir haben in Kapitel 2.2 den Raum

$$\mathcal{L}^1(\Omega, \mathcal{A}, \mu) = \{f : \Omega \to \mathbb{R} : f \text{ ist auf } \Omega \text{ bezüglich } \mathcal{A} \text{ und } \mu \text{ integrierbar}\}$$

definiert. Ist $g \in \mathcal{L}^1(\Omega, \mathcal{A}, \mu)$, so definieren wir

$$\|g\|_1 = \int |g| d\mu$$

Man überprüft einfach, dass $\|\lambda g\|_1 = |\lambda| \|g\|_1$ und $\|g + h\|_1 \leq \|g\|_1 + \|h\|_1$ für alle $\lambda \in \mathbb{R}$ und alle messbare g, h. Trotzdem ist $\|.\|_1$ keine Norm auf $\mathcal{L}_1(\Omega, \mathcal{A}, \mu)$, weil $\|g\|_1 = 0$ nicht impliziert, dass $g = 0$ gelten muss (es impliziert nur, dass $g = 0$ fast überall; wir werden später zu diesem Problem zurückkommen). Man nennt eine Abbildung mit diesen Eigenschaften eine Seminorm auf $\mathcal{L}_1(\Omega, \mathcal{A}, \mu)$. Wir können die Seminorm benutzen um einen neuen Konvergenzbegrieff zu definieren. Ist $f_k \in \mathcal{L}^1(\Omega, \mathcal{A}, \mu)$ für alle $k \in \mathbb{N}$ und $f \in \mathcal{L}^1(\Omega, \mathcal{A}, \mu)$ so sagen wir, dass $f_k \to f$ in $\mathcal{L}^1(\Omega, \mathcal{A}, \mu)$ falls

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

Im nächsten Satz untersuchen wir die Implikationen zwischen den verschiedenen Konvergenzbegriffen.

Satz 3.1. *Sei $(\Omega, \mathcal{A}, \mu)$ ein Massraum, $f : \Omega \to \mathbb{R}$ messbar und $f_k : \Omega \to \mathbb{R}$ messbar,für alle $k \in \mathbb{N}$. Dann gilt*

i) $f_k \to f$ gleichmässig impliziert, dass $f_k \to f$ in Mass.

ii) Gilt $f \in \mathcal{L}^1(\Omega, \mathcal{A}, \mu)$, $\mu(\Omega) < \infty$ und $f_k \to f$ gleichmässig, so gilt $f_k \to f$ in $\mathcal{L}^1(\Omega, \mathcal{A}, \mu)$.

iii) Gilt $f, f_k \in \mathcal{L}^1(\Omega, \mathcal{A}, \mu)$ für alle $k \in \mathbb{N}$ und $f_k \to f$ in $\mathcal{L}^1(\Omega, \mathcal{A}, \mu)$, so gilt $f_k \to f$ in Mass.

iv) Ist $\mu(\Omega) < \infty$ und gilt $f_k \to f$ punktweise fast überall, so gilt $f_k \to f$ in Mass.

v) Gilt $f_k \to f$ in Mass, dann existiert eine Nullmenge $N \in \mathcal{A}$ und eine Teilfolge f_{k_j} mit $f_{k_j}(x) \to f(x)$ für alle $x \in \Omega \backslash N$.

Beweis. i) Aus $f_k \to f$ gleichmässig folgt, dass für alle $\delta > 0$ ein $K \in \mathbb{N}$ existiert, mit $|f_k(x) - f(x)| < \delta$ für alle $k > K$. Das bedeutet, dass

$$\mu\left(\{x \in \Omega : |f_k(x) - f(x)| > \delta\}\right) = 0$$

für alle $k > K$. Das zeigt, dass $f_k \to f$ in Mass.

ii) Wir bemerken, dass

$$\int |f_k| d\mu \leq \int |f| d\mu + \int |f_k - f| d\mu \leq \|f\|_1 + \mu(\Omega) \sup_x |f_k(x) - f(x)|$$

Das zeigt, dass $f_k \in \mathcal{L}^1(\Omega, \mathcal{A}, \mu)$ für alle $k \in \mathbb{N}$. Ferner, zeigt

$$\int |f_k - f| d\mu \leq \mu(\Omega) \sup_{x \in \Omega} |f_k(x) - f(x)| \to 0$$

dass $f_k \to f$ in $\mathcal{L}^1(\Omega, \mathcal{A}, \mu)$.

iii) Aus der Chebyshev-Ungleichung (Lemma 2.18) finden wir, für ein beliebiges $\delta > 0$, dass

$$\mu\left(\{x \in \Omega : |f_k(x) - f(x)| \geq \delta\}\right) \leq \frac{1}{\delta} \int |f_k - f| d\mu \to 0$$

für $k \to \infty$.

iv) Seien $\delta, \varepsilon > 0$ beliebig. Aus dem Satz von Egorov (Aufgabe 4, Blatt 6) folgt, dass ein $F \in \mathcal{A}$ existiert, mit $\mu(\Omega \backslash F) < \varepsilon$ und $f_k \to f$ gleichmässig auf F. Deswegen gilt

$$\mu\left(\{x \in \Omega : |f_k(x) - f(x)| > \delta\}\right) \leq \varepsilon + \mu\left(\{x \in F : |f_k(x) - f(x)| > \delta\}\right)$$

Wie in i) erhalten wir, dass

$$\lim_{k \to \infty} \mu\left(\{x \in \Omega : |f_k(x) - f(x)| > \delta\}\right) \leq \varepsilon$$

Da $\varepsilon > 0$ beliebig war, muss

$$\lim_{k \to \infty} \mu\left(\{x \in \Omega : |f_k(x) - f(x)| > \delta\}\right) = 0$$

v) Wir konstruieren eine monoton steigende Folge k_j rekursiv. Wir setzen $k_0 = 0$. Wenn wir schon $k_0 < k_1 < \cdots < k_{j-1}$ definiert haben definieren wir k_j wie folgt. Da $f_k \to f$ in Mass, gilt

$$\lim_{k \to \infty} \mu\left(\{x \in \Omega : |f_k(x) - f(x)| > 2^{-j}\}\right) = 0$$

Das heisst, es existiert $k_j > k_{j-1}$ so, dass die Menge

$$A_j = \{x \in \Omega : |f_{k_j}(x) - f(x)| > 2^{-j}\}$$

das Mass $\mu(A_j) \leq 2^{-j}$ hat. Wir behaupten nun, dass $f_{k_j} \to f$ punktweise fast überall konvergiert. Wir setzen nämlich

$$N = \bigcap_{n \in \mathbb{N}} \bigcup_{j > n} A_j$$

Einerseits gilt

$$\mu(N) \leq \mu\left(\bigcup_{j > n} A_j\right) \leq \sum_{j > n} 2^{-j} = 2^{-n}$$

für alle $n \in \mathbb{N}$, was offenbar $\mu(N) = 0$ impliziert. Anderseits, für $x \in N^c = \bigcup_{n \in \mathbb{N}} \bigcap_{j > n} A_j^c$ existiert ein $j_0 \in \mathbb{N}$ mit $x \in A_j$ für alle $j > j_0$. Das bedeutet, dass

$$|f_{k_j}(x) - f(x)| \leq 2^{-j}$$

für alle $j > j_0$, und also, dass $f_{k_j}(x) \to f(x)$ für $j \to \infty$. $\square$

Bemerkungen:

1) Ohne Annahme $\mu(\Omega) < \infty$ impliziert gleichmässige Konvergenz i.A. keine $\mathcal{L}^1$-Konvergenz (d.h. ii) braucht wirklich die Bedingung $\mu(\Omega) < \infty$). Betrachte zum Beispiel die Folge $f_k(x) = 1/k$ für alle $x \in \mathbb{R}$. Es gilt offenbar $\sup_x |f_k(x)| \to 0$ für $k \to \infty$, aber $\int_{\mathbb{R}} |f_k| d\lambda = \infty$ für alle $k \in \mathbb{N}$.

2) Auch in iv) ist die Annahme $\mu(\Omega) < \infty$ wichtig. Betrachte z.B. die Folge $f_k = \chi_{[k,\infty]}$. Dann gilt $f_k(x) \to 0$ für $k \to \infty$ für alle $x \in \mathbb{R}$ aber
$$\mu\left(\{x \in \Omega : |f_k(x)| > 1/2\}\right) = \infty$$
für alle $k \in \mathbb{N}$.

3) Konvergenz in Mass impliziert i.A. keine $\mathcal{L}^1$-Konvergenz. Z.B. die Folge $f_k(x) = k\chi_{[0;1/k]}(x)$ für $x \in \mathbb{R}$ ist so, dass
$$\mu\left(\{x \in \mathbb{R} : |f_k(x)| > \delta\}\right) \leq 1/k \to 0$$
für $k \to \infty$. Anderseits gilt
$$\int |f_k| d\lambda = 1$$
für alle $k \in \mathbb{N}$, und deswegen konvergiert f_k nicht zu Null in $\mathcal{L}^1$.

4) Konvergenz fast überall impliziert i.A. keine Konvergenz in $\mathcal{L}^1$. Die Folge $f_k(x) = k\chi_{[0;1/k]}$ in 3) konvergiert fast überall zu Null, aber konvergiert nicht in $\mathcal{L}^1$.

5) $\mathcal{L}^1$-Konvergenz impliziert i.A. keine fast überall punktweise Konvergenz; das wurde in Aufgabenblatt 7, Aufgabe 4)c) gesehen. Im nächsten Lemma geben wir auch ein Gegenbesipiel. Es folgt, dass auch Konvergenz in Mass keine fast überall Konvergenz implizieren kann. Aus v) folgt aber, dass Konvergenz in Mass, und deswegen auch $\mathcal{L}^1$-Konvergenz, die Existenz einer Teilfolge impliziert, die fast überall konvergiert.

Lemma 3.2. *Es existiert eine Folge von Lebesgue integrierbaren Funktionen $f_k : \mathbb{R} \to \mathbb{R}$, die in $\mathcal{L}^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ gegen $f \equiv 0$ konvergiert so, dass, für alle $x \in \mathbb{R}$, die Folge $f_k(x)$ nicht konvergiert.*

Beweis. Wir definieren die Menge $E = \{(j,k) \in \mathbb{Z} \times \mathbb{N} : |j| < k^2\}$. E ist offenbar abzählbar. Deswegen finden wir eine Folge $(a_n)_{n\in\mathbb{N}}$ mit Werten in E so, dass $\{a_n : n \in \mathbb{N}\} = E$. Für $(j,k) \in E$ setzen wir $A_{j,k} = [j/k, (j+1)/k)$. Für $n \in \mathbb{N}$, definieren wir dann $f_n = \chi_{A_{a_n}}$. Einerseits gibt es für alle $\varepsilon > 0$ fest nur endlich viele $n \in \mathbb{N}$ mit $\mathcal{L}^1(A_{a_n}) \geq \varepsilon$ (weil es gibt nur endlich viele $(j,k) \in E$ mit $1/k \geq \varepsilon$). Deswegen gilt $\|f_n\|_1 \to 0$ für $n \to \infty$ (für n gross genug wird $\mathcal{L}^1(A_{a_n}) < \varepsilon$, für beliebiges $\varepsilon > 0$). Anderseits für alle $x \in \mathbb{R}$ gibt es unendlich viele Paare $(j,k) \in E$ mit $x \in A_{j,k}$ und unendlich viele Paare mit $x \notin A_{j,k}$. Das heisst, es gibt unendlich viele $n \in \mathbb{N}$ mit $f_n(x) = 1$ und unendlich viele $n \in \mathbb{N}$ mit $f_n(x) = 0$, und deswegen kann $f_n(x)$ nicht konvergieren. $\square$

Korollar 3.3. *Fast überall punktweise Konvergenz auf* $\mathcal{L}^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ *(oder auch auf* $\mathcal{L}^1(\mathbb{R}, \mathcal{M}_{\lambda^*}, \lambda)$*) ist nicht metrizierbar. D.h. es existiert keine Metrik* d *auf* $\mathcal{L}^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ *mit der Eigenschaft, dass für alle Folgen* $f_k \in \mathcal{L}^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ *und alle* $f \in \mathcal{L}^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$*, es gilt* $f_k(x) \to f(x)$ *fast überall genau dann wenn* $d(f_k, f) \to 0$ *für* $k \to \infty$*.*

Beweis. Nehmen wir an, es existiert so eine Metrik d auf $\mathcal{L}^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$. Die Folge f_k aus Lemma 3.2 konvergiert gegen Null in $\mathcal{L}^1$. Jede Teilfolge f_{j_k} konvergiert also gegen Null in $\mathcal{L}^1$. Aus Teil v) in Satz 3.1 besitzt also jede Teilfolge f_{j_k} eine Teilfolge $f_{j_{k_\ell}}$ die fast überall gegen Null konvergiert. Dann würde $d(f_{j_{k_\ell}}, 0) \to 0$ für $\ell \to \infty$. Das würde aber auch implizieren, dass $d(f_j, 0) \to 0$ für $j \to \infty$ (sonst könnten wir ein $\varepsilon > 0$ und eine Teilfolge f_{j_k} finden, mit $d(f_{j_k}, 0) > \varepsilon$ für alle $k \in \mathbb{N}$; dann hätte f_{j_k} keine Teilfolge $f_{j_{k_\ell}}$ mit $d(f_{j_{k_\ell}}, 0) \to 0$ für $\ell \to \infty$). □

3.2 Die Vektorräume $\mathcal{L}^p(\Omega, \mathcal{A}, \mu)$ und $L^p(\Omega, \mathcal{A}, \mu)$

Sei $(\Omega, \mathcal{A}, \mu)$ ein Massraum, und $1 \leq p < \infty$ eine reelle Zahl. Dann setzen wir

$$\mathcal{L}^p(\Omega, \mathcal{A}, \mu) = \{f : \Omega \to \mathbb{R} : f \text{ ist messbar bzg. } \mathcal{A} \text{ und } |f|^p \text{ integrierbar bezüglich } \mu\}$$

Ist $f \in \mathcal{L}^p(\Omega, \mathcal{A}, \mu)$ und $\lambda \in \mathbb{R}$ so ist offenbar auch $\lambda f \in \mathcal{L}^p(\Omega, \mathcal{A}, \mu)$. Sind $f, g \in \mathcal{L}^p(\Omega, \mathcal{A}, \mu)$, dann ist $f + g$ als Summe von zwei messbare Funktionen auch $\mathcal{A}$-messbar. Ferner, mit

$$\begin{aligned} |f(x) + g(x)|^p &\leq (|f(x)| + |g(x)|)^p \leq (2 \max(|f(x)|, |g(x)|))^p \\ &\leq 2^p \max(|f(x)|^p, |g(x)|^p) \leq 2^p (|f(x)|^p + |g(x)|^p) \end{aligned}$$

folgt aus der Integrierbarkeit von $|f|^p$ und $|g|^p$ auch die Integrierbarkeit von $|f + g|^p$. Das zeigt, dass $\mathcal{L}^p(\Omega, \mathcal{A}, \mu)$ ein Vektorraum ist.

Auf $\mathcal{L}^p(\Omega, \mathcal{A}, \mu)$ definieren wir nun die Abbildung $\|.\|_p : \mathcal{L}^p(\Omega, \mathcal{A}, \mu) \to [0; \infty)$. durch

$$\|f\|_p = \left(\int |f|^p d\mu \right)^{1/p}$$

Wir betrachten auch den Fall $p = \infty$. Wir setzen

$$\mathcal{L}^\infty(\Omega, \mathcal{A}, \mu) = \{f : \Omega \to \mathbb{R} : \exists M > 0 \text{ mit } |f(x)| \leq M \text{ für } \mu\text{-fast alle } x \in \Omega\}$$

Für $f \in \mathcal{L}^\infty(\Omega, \mathcal{A}, \mu)$ definieren wir dann

$$\|f\|_\infty = \inf\{M > 0 : |f(x)| \leq M \text{ für } \mu\text{-fast alle } x \in \Omega\}$$

Wir bemerken, dann, dass $|f(x)| \leq \|f\|_\infty$ für μ-fast alle $x \in \Omega$. In der Tat, für alle $n \in \mathbb{N}$, wir finden eine Nullmenge N_n mit $|f(x)| \leq \|f\|_\infty + 1/n$ für alle $x \in N_n^c$. Wir setzen $N = \bigcup_{n \in \mathbb{N}} N_n$. Dann gilt $N \in \mathcal{A}$ mit $\mu(N) = 0$ und, für alle $x \in N^c$ gilt $|f(x)| \leq \|f\|_\infty + 1/n$ für alle $n \in \mathbb{N}$. Deswegen gilt $|f(x)| \leq \|f\|_\infty$ für alle $x \in N^c$.

Wir können auch Funktionen mit Werten in $\mathbb{R}^n$ oder in $\mathbb{C}$ betrachten. Da eine Funktion $f : \Omega \to \mathbb{R}^n$ genau dann integrierbar ist, falls alle seine Komponenten integrierbar sind, ist $\mathcal{L}^p(\Omega, \mathcal{A}, \mu; \mathbb{R}^n)$ die Menge alle Funktionen $f : \Omega \to \mathcal{A}$ gegeben aus $f(x) = (f_1(x), \ldots, f_n(x))$, wobei f_i $\mathcal{A}$-messbar ist und $|f_i|^p$ integrierbar ist, für alle $i = 1, \ldots, n$. Für $f \in \mathcal{L}^p(\Omega, \mathcal{A}, \mu)$ definieren wir dann $\|f\|_p = \|\|f\|\|_p$, wobei $\|f\| = (|f_1|^2 + \cdots + |f_n|^2)^{1/2}$ die euklidische Norm von f ist.

Wir untersuchen nun die Eigenscahften der Abbildungen $\|.\|_p : \mathcal{L}^p(\Omega, \mathcal{A}, \mu) \to [0; \infty)$. Wir werden das folgende Lemma brauchen.

Lemma 3.4. *Sei $1 < p < \infty$, und $1 < q < \infty$ so, dass $1/p + 1/q = 1$. Seien $x, y \geq 0$. Dann gilt*

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q}$$

Beweis. Wir können annehmen, $x, y > 0$ sonst ist die Behauptung trivial. Die Funktion $f(x) = \log x$, ist konkav auf $(0; \infty)$. Das bedeutet, dass $f(\lambda x + (1-\lambda)y) \geq \lambda f(x) + (1-\lambda)f(y)$ für alle $x, y > 0$ und $0 \leq \lambda \leq 1$. Da $1/p + 1/q = 1$ finden wir also

$$\log\left(\frac{1}{p}x^p + \frac{1}{q}y^q\right) \geq \frac{1}{p}\log x^p + \frac{1}{q}\log y^q = \log x + \log y = \log(xy)$$

Da $\log x$ auf $(0; \infty)$ monoton wachsend ist, folgt, dass

$$\frac{1}{p}x^p + \frac{1}{q}y^q \geq xy$$

wie behauptet. □

Satz 3.5 (Hölder Ungleichung)**.** *Sei $(\Omega, \mathcal{A}, \mu)$ ein Massraum und $p, q \in [1; \infty]$ mit $1/p + 1/q = 1$. Seien $f \in \mathcal{L}^p(\Omega, \mathcal{A}, \mu)$ und $g \in \mathcal{L}^q(\Omega, \mathcal{A}, \mu)$.Dann ist $fg \in \mathcal{L}^1(\Omega, \mathcal{A}, \mu)$ und*

$$\int |fg| d\mu \leq \|f\|_p \|g\|_q$$

Beweis. Ist $p = 1$ und $q = \infty$ so gilt $|g(x)| \leq \|g\|_\infty$ für fast alle $x \in \Omega$ und

$$\int |fg| d\mu = \int |f||g| d\mu \leq \int |f| \|g\|_\infty d\mu = \|g\|_\infty \|f\|_1$$

Analog kann man den Fall $p = \infty$ und $q = 1$ betrachten. Für $p, q \in (1; \infty)$ bemerken wir zunächst, dass es genug, die Behauptung für $f \in \mathcal{L}^p(\Omega, \mathcal{A}, \mu)$ und $g \in \mathcal{L}^q(\Omega, \mathcal{A}, \mu)$ mit $\|f\|_p = \|g\|_q = 1$ zu beweisen. Dann bekommt man

$$\int |fg| d\mu = \|f\|_p \|g\|_q \int \frac{|f|}{\|f\|_p} \frac{|g|}{\|g\|_q} \leq \|f\|_p \|g\|_q$$

weil offenbar $\|(f/\|f\|_p)\|_p = \|(g/\|g\|_q)\|_q = 1$. Seien also $f \in \mathcal{L}^p(\Omega, \mathcal{A}, \mu)$ und $g \in \mathcal{L}^q(\Omega, \mathcal{A}, \mu)$ mit $\|f\|_p = 1$ und $\|g\|_q = 1$. Aus Lemma 3.4 finden wir

$$|f(x)||g(x)| \leq \frac{|f(x)|^p}{p} + \frac{|g(x)|^q}{q}$$

für alle $x \in \Omega$. Deswegen

$$\int |fg| d\mu \leq \frac{1}{p} \int |f|^p d\mu + \frac{1}{q} \int |g|^q d\mu = \frac{\|f\|_p^p}{p} + \frac{\|g\|_q^q}{q} = \frac{1}{p} + \frac{1}{q} = 1\,.$$

□

Mit Hilfe von Hölder's Ungleichung können wir nun zeigen, dass $\|.\|_p$ die Dreiecksungleichung erfüllt.

Satz 3.6 (Minkowski's Ungleichung)**.** *Sei* $(\Omega, \mathcal{A}, \mu)$ *ein Massraum und* $1 \leq p \leq \infty$. *Seien* $f, g \in \mathcal{L}^p(\Omega, \mathcal{A}, \mu)$. *Dann ist*

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

Beweis. Sei zunächst $p = \infty$. Dann setzen wir $N_1 = \{x \in \Omega : |f(x)| > \|f\|_\infty\}$ und $N_2 = \{x \in \Omega : |g(x)| > \|g\|_\infty\}$. N_1, N_2 sind zwei Nullmengen i $\mathcal{A}$, und damit ist auch $N = N_1 \cup N_2$ eine Nullmenge. Für $x \in N^c$ gilt dann

$$|f(x)| + |g(x)| \leq \|f\|_\infty + \|g\|_\infty$$

Das zeigt, dass $\|f + g\|_\infty \leq \|f\|_\infty + \|g\|_\infty$.

Sei nun $1 \leq p < \infty$. Wir haben

$$\begin{aligned} \|f + g\|_p^p = \int |f(x) + g(x)|^p d\mu &= \int |f(x) + g(x)||f(x) + g(x)|^{p-1} d\mu \\ &\leq \int |f(x)||f(x) + g(x)|^{p-1} d\mu + \int |g(x)||f(x) + g(x)|^{p-1} d\mu \\ &(\|f\|_p + \|g\|_p) \cdot \||f + g|^{p-1}\|_q \end{aligned} \tag{37}$$

wobei $1 < q \leq \infty$ so gewählt wurde, dass $1/p + 1/q = 1$. Dann gilt $1/q = 1 - 1/p = (p-1)/p$ und $q = p/(p-1)$. Deshalb

$$\||f + g|^{p-1}\|_q = \left(\int |f + g|^{q(p-1)}\right)^{(p-1)/p} = \left(\int |f + g|^p\right)^{(p-1)/p} = \|f + g\|_p^{p-1}$$

Aus (37) finden wir $\|f + g\|_p \leq \|f\|_p + \|g\|_p$. □

Zusammenfassend, die Abbildung $\|.\|_p : \mathcal{L}^p(\Omega, \mathcal{A}, \mu) \to [0; \infty)$ erfüllt $\|\lambda f\|_p = |\lambda|\|f\|_p$ für alle $\lambda \in \mathbb{R}$ und alle $f \in \mathcal{L}^p(\Omega, \mathcal{A}, \mu)$ und $\|f + g\|_p \leq \|f\|_p + \|g\|_p$ für alle $f, g \in \mathcal{L}^p(\Omega, \mathcal{A}, \mu)$. Eine solche Abbildung heisst eine Seminorm auf $\mathcal{L}^p(\Omega, \mathcal{A}, \mu)$. Die Abbildung $\|.\|_p$ ist aber keine Norm, weil $\|f\|_p = 0$ nicht impliziert, dass $f = 0$ ist. In der Tat impliziert $\|f\|_p = 0$ nur, dass $f = 0$ μ-fast überall.

Damit wir aus $\mathcal{L}^p(\Omega, \mathcal{A}, \mu)$ ein normierten Vektorraum bekommen, identifizieren wir in $\mathcal{L}^p(\Omega, \mathcal{A}.\mu)$ alle Funktionen, die fast überall übereinstimmen. Wir definieren dazu die folgende Äquivalenzrelation auf $\mathcal{L}^p(\Omega, \mathcal{A}, \mu)$. Für $f, g \in \mathcal{L}^p(\Omega, \mathcal{A}, \mu)$ setzen wir $f \sim g$ falls $f(x) = g(x)$ für μ-fast alle $x \in \Omega$. Es ist einfach zu zeigen, dass $\sim$ eine Äquivalenzrelation ist. In der Tat $\sim$ ist offenbar reflexiv ($f \sim f$ für alle $f \in \mathcal{L}^p(\Omega, \mathcal{A}, \mu)$), symmetrisch ($f \sim g$ impliziert auch, dass $g \sim f$, für alle $f, g \in \mathcal{L}^p(\Omega, \mathcal{A}, \mu)$) und transitiv ($f \sim g$ und $g \sim h$ impliziert, dass $f \sim h$, für alle $f, g, h \in \mathcal{L}^p(\Omega, \mathcal{A}, \mu)$). Für alle $f \in \mathcal{L}^p(\Omega, \mathcal{A}, \mu)$ können wir also die Äquivalenzklasse

$$[f] = \{g \in \mathcal{L}^p(\Omega, \mathcal{A}, \mu) : g \sim f\}$$

Äquivalenzklassen sind disjunkt, und $\bigcup_{f \in \mathcal{L}^p(\Omega, \mathcal{A}, \mu)} [f] = \mathcal{L}^p(\Omega, \mathcal{A}, \mu)$.

Wir definieren nun für $1 \leq p \leq \infty$

$$L^p(\Omega, \mathcal{A}, \mu) = \mathcal{L}^p(\Omega, \mathcal{A}, \mu)/\sim$$

als das Quotientenraum von $\mathcal{L}^p(\Omega, \mathcal{A}, \mu)$ bezüglich der Äquivalenzrelation $\sim$. Mit anderen Wörter

$$L^p(\Omega, \mathcal{A}, \mu) = \{[f] : f \in \mathcal{L}^p(\Omega, \mathcal{A}, \mu)\}$$

D.h. Elemente vom Raum $L^p(\Omega, \mathcal{A}, \mu)$ sind nicht Funktionen, sondern Äquivalenzklassen von Funktionen. Trotzdem denkt man oft an den Elementen von $L^p(\Omega, \mathcal{A}, \mu)$ als Funktionen, die eine Äquivalenzklasse darstellen (ein Vertreter einer Äquivalenzklasse). $L^p(\Omega, \mathcal{A}, \mu)$ erbt aus $\mathcal{L}^p(\Omega, \mathcal{A}, \mu)$ die Struktur eines Vektorraumes, durch die Definitionen $[f] + [g] := [f + g]$ und $\lambda[f] := [\lambda f]$.

Auf $L^p(\Omega, \mathcal{A}, \mu)$ definieren wir die Abbildung $\|.\|_p : L^p(\Omega, \mathcal{A}, \mu) \to [0; \infty)$ durch $\|[f]\|_p := \|f\|_p$. Die Abbildung ist zunächst wohldefiniert, weil $f, g \in \mathcal{L}^p(\Omega, \mathcal{A}, \mu)$ mit $f \sim g$ impliziert, dass $f(x) = g(x)$ μ-fast überall und deswegen auch $|f(x)|^p = |g(x)|^p$ μ-fast überall, und also, dass $\|f\|_p = \|g\|_p$. Ferner, $\|[f]\|_p = 0$ impliziert, dass $\|f\|_p = 0$ und deswegen, dass $f = 0$ fast überall. Das heisst, dass $[f] = [0]$. Es gilt offenbar

$$\|\lambda[f]\|_p = \|[\lambda f]\|_p = \|\lambda f\|_p = |\lambda| \|f\|_p = |\lambda| \|[f]\|_p$$

und

$$\|[f] + [g]\|_p = \|[f + g]\|_p = \|f + g\|_p \leq \|f\|_p + \|g\|_p = \|[f]\|_p + \|[g]\|_p$$

Damit definiert nun $\|.\|_p : L^p(\Omega, \mathcal{A}, \mu) \to [0; \infty)$ eine Norm auf $L^p(\Omega, \mathcal{A}, \mu)$. Ein Vektorraum mit einer Norm heisst ein normierter Raum. $L^p(\Omega, \mathcal{A}, \mu)$ ist ein normierter Raum, für alle $1 \leq p \leq \infty$.

Analog können wir auch die normierten Räume $L^p(\Omega, \mathcal{A}, \mu; \mathbb{R}^n)$ definieren, die aus Äquivalenzklassen von $\mathcal{A}$-messbare und im p-te Potenz integrierbare Funktionen $f : \Omega \to \mathbb{R}^m$mit Werten in $\mathbb{R}^m$. Auch die Räume $L^p(\Omega, \mathcal{A}, \mu; \mathbb{C})$ können analog definiert werden.

Bemerkung: Man könnte denken, $\mathcal{L}^p(\Omega, \mathcal{A}, \mu)$ und $L^p(\Omega, \mathcal{A}, \mu)$ auch für $0 < p < 1$ zu definieren. Dann wären $\mathcal{L}_p(\Omega, \mathcal{A}, \mu)$ und $L^p(\Omega, \mathcal{A}, \mu)$ Vektorräume. In diesem Fall würde aber die Dreiecksungleichung $\|f + g\|_p \leq \|f\|_p + \|g\|_p$ nicht gelten (Beweis: Übung), und $\|.\|_p$ wäre dann keine Norm. Das ist der Grund, warum wir nur $p \geq 1$ betrachten.

Die Norm $\|.\|_p$ auf $L^p(\Omega, \mathcal{A}, \mu)$ induziert natürlich eine Metrik, definiert durch $d(f, g) = \|f - g\|_p$ für alle $f, g \in L^p(\Omega, \mathcal{A}, \mu)$. Das definiert also auch ein Begriff von Konvegenz. Eine Folge f_n in $L^p(\Omega, \mathcal{A}, \mu)$ konvergiert gegen $f \in L^p(\Omega, \mathcal{A}, \mu)$ falls $\|f_n - f\|_p \to 0$. Obwohl die Elemente von $L^p(\Omega, \mathcal{A}, \mu)$ Äquivalenzklassen sind, werden wir im folgenden oft vermeiden, die Notation $[f], [g]$ zu benuzten; wir werden dagegen oft die Äquivalenzklasse $[f] \in L^p(\Omega, \mathcal{A}, \mu)$ mit einem Vertreter $f \in \mathcal{L}^p(\Omega, \mathcal{A}, \mu)$ identifizieren (das führt zu einer nicht so ganz genaue Notation; es sollte aber immer klar sein, was wirklich gemeint ist). Wir erinnern, dass ein normierten Raum heisst vollständig, falls die vom Norm induzierten Metrik vollständig ist (das bedeutet, dass jede Cauchy Folge konvergiert).

Theorem 3.7. *Sei $(\Omega, \mathcal{A}, \mu)$ ein Massraum. Dann ist der normierte Raum $L^p(\Omega, \mathcal{A}, \mu)$ vollständig, für alle $1 \leq p \leq \infty$.*

Beweis. Sei $f_n \in L^p(\Omega, \mathcal{A}, \mu)$ eine Cauchy Folge. Wir möchten zeigen, dass ein $f \in L^p(\Omega, \mathcal{A}, \mu)$ existiert, mit $\|f_n - f\|_p \to 0$, für $n \to \infty$. Es genug zu zeigen, es existiert $f \in L^p(\Omega, \mathcal{A}, \mu)$ und eine Teilfolge $f_{n_j} \in L^p(\Omega, \mathcal{A}, \mu)$ mit $\|f_{n_j} - f\|_p \to 0$ für $j \to \infty$. In der Tat, wenn wir eine solche Teilfolge finden können, so können wir

$$\|f_n - f\|_p \leq \|f_n - f_{n_j}\|_p + \|f_{n_j} - f\|_p$$

abschätzen. Gegeben $\varepsilon > 0$ finden wir dann $N_1 > 0$ mit $\|f_{n_j} - f\|_p < \varepsilon/2$ für alle $j > N_1$ und $N_2 > 0$ mit $\|f_n - f_{n_j}\| \leq \varepsilon/2$ für alle $n, j > N_2$ (weil f_n eine Cauchy Folge ist). Also, für $n, j \geq \max(N_1, N_2)$ ist $\|f_n - f\|_p \leq \varepsilon$. Das zeigt, dass $f_n \to f$, für $n \to \infty$.

Es bleibt zu zeigen, dass ein $f \in L^p(\Omega, \mathcal{A}, \mu)$ und eine Teilfolge n_j existieren, mit $\|f_{n_j} - f\|_p \to 0$, für $j \to \infty$. Wir betrachten hier den Fall $1 \leq p < \infty$ (der Fall $p = \infty$ lassen wir als Übung). Wir konstruieren die Teilfolge n_j rekursiv. Wir wählen zunächst n_1 so, dass $\|f_{n_1} - f_n\|_p < 1/2$ für alle $n \geq n_1$ (das ist möglich, weil f_n eine Cauchy Folge ist). Dann wählen wir $n_2 > n_1$ so, dass $\|f_{n_2} - f_n\|_p < 1/4$ für alle $n \geq n_2$. Wenn wir schon $n_1 < n_2 < \cdots < n_k$ definiert haben, wählen wir $n_{k+1} > n_k$ mit $\|f_{n_{k+1}} - f_n\|_p < 2^{-k+1}$ für alle $n \geq n_{k+1}$.

Damit haben wir eine Teilfolge n_j gefunden, mit

$$\|f_{n_k} - f_{n_{k+1}}\|_p \leq 2^{-k}$$

für alle $k \in \mathbb{N}\backslash\{0\}$. Wir definieren nun die monotone Folge von Funktionen

$$F_\ell(x) = |f_{n_1}(x)| + \sum_{k=1}^{\ell} |f_{n_{k+1}}(x) - f_{n_k}(x)|$$

Aus der Dreicksungleichung finden wir

$$\|F_\ell\|_p \leq \|f_{n_1}\|_p + \sum_{k=1}^{\ell} \|f_{n_{k+1}} - f_{n_k}\|_p \leq \|f_{n_1}\|_p + 1$$

Da $F_\ell(x)$ monoton steigend in ℓ ist, für alle $x \in \Omega$, existiert der Limes

$$F(x) := \lim_{\ell \to \infty} F_\ell(x)$$

($F(x)$ darf unendlich sein). Aus der Satz über die Monotone Konvergenz ist $F \in L^p(\Omega, \mathcal{A}, \mu)$. Das impliziert insbesondere, dass $F(x) < \infty$ μ-fast überall. Für alle x mit $F(x) < \infty$ bemerken wir nun, dass

$$f_{n_{k+1}}(x) = f_{n_1}(x) + (f_{n_2}(x) - f_{n_1}(x)) + \cdots + (f_{n_{k+1}}(x) - f_{n_k}(x))$$

wobei die Reihe auf der rechten Seiten absolut konvergent ist und deswegen auch konvergent ist im Limes $k \to \infty$. Also $f(x) := \lim_{k\to\infty} f_{n_{k+1}}(x)$ existiert für fast alle $x \in \Omega$. Wir definieren weiter $f(x) = 0$ für alle $x \in \Omega$, mit $F(x) = \infty$. Dann haben wir $|f_{n_{k+1}}(x)| \leq F(x)$ und $f(x) = \lim_{k\to\infty} f_{n_{k+1}}(x)$ für fast alle $x \in \Omega$. Der Satz über dominierte Konvergenz impliziert, dass $f \in L^p(\Omega, \mathcal{A}, \mu)$. Wir haben $\lim_{k\to\infty} |f_{n_{k+1}}(x) - f(x)| = 0$ für fast alle $x \in \Omega$ und $|f_{n_{k+1}}(x) - f(x)| \leq F(x) + |f(x)|$ für alle $x \in \Omega$, wobei $F + |f| \in L^p(\Omega, \mathcal{A}, \mu)$. Eine weitere Anwendung vom Satz über dominierte Konvergenz impliziert, dass

$$\lim_{k\to\infty} \int |f_{n_{k+1}}(x) - f(x)|^p d\mu = 0$$

und also, dass $\|f_{n_{k+1}} - f\|_p \to 0$. $\square$

Damit haben wir gezeigt, $L^p(\Omega, \mathcal{A}, \mu)$ ist ein vollständiger normierter Raum, für alle $1 \leq p \leq \infty$. Ein vollständiger normierter Raum wird als Banachraum bezeichnet.

Für $p = 2$, kann man auf $L^2(\Omega, \mathcal{A}, \mu)$ auch ein Skalarprodukt definieren. Für $f, g \in L^2(\Omega, \mathcal{A}, \mu)$, wir definieren nämlich

$$\langle f, g \rangle = \int f g d\mu$$

Dann ist $\langle f, g \rangle$ offenbar linear im ersten Argument und symmetrisch (wenn man $\mathbb{C}$-wertige Funktionen betrachtet, dann wäre $\langle f, g \rangle = \int \overline{f} g d\mu$ linear in g aber antilinear in f). Ferner,

$$\langle f, f \rangle = \|f\|_2^2 \geq 0$$

für alle $f \in L^2(\Omega, \mathcal{A}, \mu)$ und verschwindet genau dann, wenn $f = 0$ ist. Damit ist $\langle ., . \rangle$ ein Skalarprodukt, das die Norm $\|.\|_2$ induziert. Ein Vektorraum V, versehen mit einem Skalarprodukt, das eine Norm und damit eine Metrik induziert, bezüglich welche V vollständig ist, heisst ein Hilbertraum (Hilberträume sind immer insbesondere Banachräume). $L^2(\Omega, \mathcal{A}, \mu)$ ist ein Hilbertraum. Für $p \neq 2$, kann die Norm $\|.\|_p$ nicht durch ein Skalarprodukt induziert werden.

3.3 Approximation mit einfache Funktionen

Ein wichtiges Hilfmittel in der Analysis von L^p-Räume ist die Tatsache, dass Funktionen in $L^p(\Omega, \mathcal{A}, \mu)$ durch Folge von einfache messbare Funktionen approximiert werden können.

Proposition 3.8. *Sei $(\Omega, \mathcal{A}, \mu)$ ein Massraum, $1 \leq p \leq \infty$. Die Menge der einfachen Funktionen in $\mathcal{L}^p(\Omega, \mathcal{A}, \mu)$ ist dicht in $\mathcal{L}^p(\Omega, \mathcal{A}, \mu)$ und definiert deswegen eine dichte Teilmenge von $L^p(\Omega, \mathcal{A}, \mu)$.*

Beweis. Wir betrachten hier den Fall reelwertigen Funktionen. Sei zunächst $1 \leq p < \infty$. Für $f \in \mathcal{L}^p(\Omega, \mathcal{A}, \mu)$, sei $f_+ = \max(f, 0)$ und $f_- = \max(0, -f)$ so, dass $f = f_+ - f_-$. Mit Hilfe von Prop. 2.3 finden wir dann monoton wachsende Folgen g_k, h_k von nichtnegative einfache messbare Funktionen, mit $f_+ = \lim_{k\to\infty} g_k$ und $f_- = \lim_{k\to\infty} h_k$. Wir definieren $f_k = g_k - h_k$. Dann ist f_k eine Folge einfache messbare Funktionen, mit $|f_k(x) - f(x)| \to 0$ punktweise und so, dass $|f_k(x)| \leq g_k + h_k \leq f_+ + f_- = |f(x)|$ für alle $x \in \Omega$. Die Tatsache, dass $f \in \mathcal{L}^p(\Omega, \mathcal{A}, \mu)$ und dominierte Konvergenz zeigen, dass $\|f_k - f\|_p \to 0$.

Sei nun $f \in \mathcal{L}^\infty(\Omega, \mathcal{A}, \mu)$ und $\varepsilon > 0$ fest. Wir wählen $a_0 < a_1 < \cdots < a_n$ mit $[-\|f\|_\infty; \|f\|_\infty] \subset \cup_{j=1}^n (a_{j-1}; a_j]$ und so, dass $|a_j - a_{j-1}| < \varepsilon$ für alle $j = 1, \ldots, n$. Sei $A_j = f^{-1}((a_{j-1}; a_j])$ für $j = 1, \ldots, n$. $A_j \in \mathcal{A}$, da f messbar ist. Sei $f_\varepsilon = \sum_{j=1}^n a_j \chi_{A_j}$. Dann ist f_ε eine einfache messbare Funktion auf Ω, mit $\|f - f_\varepsilon\|_\infty \leq \varepsilon$. □

Die Tatsache, dass einfache Funktionen dicht sind, erlaubt uns zu zeigen, dass der Raum $L^p(\Omega, \mathcal{A}, \mu)$ separabel ist, falls $1 \leq p < \infty$ (nicht für $p = \infty$), und falls $\mathcal{A}$ durch eine abzählbare Algebra $\mathcal{A}_0$ erzeugt wird so, dass die Einschränkung von μ auf $\mathcal{A}_0$ σ-endlich ist. Wir erinneren hier die Definition von Separabilität.

Definition 3.9. *Ein normierter Raum (oder, allgemeiner, ein metrischer Raum) heisst separabel, falls er eine abzählbare dichte Teilmenge enthält.*

Theorem 3.10. *Sei $(\Omega, \mathcal{A}, \mu)$ ein Massraum. Wir nehmen an, dass eine abzählbare Algebra $\mathcal{A}_0$ mit $\mathcal{A} = \sigma(\mathcal{A}_0)$ existiert so, dass eine Folge $(C_n)_{n\in\mathbb{N}}$ in $\mathcal{A}_0$ existiert, mit $\Omega = \cup_{n\in\mathbb{N}} C_n$ und $\mu(C_n) < \infty$ für alle $n \in \mathbb{N}$. Dann ist $L^p(\Omega, \mathcal{A}, \mu)$ separabel, für alle $1 \leq p < \infty$.*

Bemerkung. Die Annahme, dass $\mathcal{A} = \sigma(\mathcal{A}_0)$, für eine abzählbare Algebra $\mathcal{A}_0$ so, dass $\mu|_{\mathcal{A}_0}$ σ-endlich ist, ist insbesondere für die Massräume $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \lambda_n)$ erfüllt. Beweis; Übung.

Zum Beweis von Theorem 3.10 werden wir das folgende Lemma benutzen.

Lemma 3.11. *Sei $(\Omega, \mathcal{A}, \mu)$ ein Massraum. Sei $\mathcal{A}_0 \subset \mathcal{A}$ eine Algebra, mit $\sigma(\mathcal{A}_0) = \mathcal{A}$ und so, dass eine Folge $(C_n)_{n\in\mathbb{N}}$ in $\mathcal{A}_0$ existiert, mit $\Omega = \cup_n C_n$ und $\mu(C_n) < \infty$ für alle $n \in \mathbb{N}$. Dann, für alle $\varepsilon > 0$ und alle $A \in \mathcal{A}$ mit $\mu(A) < \infty$, es existiert $A_0 \in \mathcal{A}_0$ mit $\mu(A\Delta A_0) < \varepsilon$.*

Beweis. Wir betrachten zunächst den Fall, dass μ endlich ist. Unter dieser zusätzliche Annahme, definieren wir

$$\mathcal{F} = \{A \in \mathcal{A} : \text{Für alle } \varepsilon > 0 \text{ es existiert } A_0 \in \mathcal{A}_0 \text{ mit } \mu(A\Delta A_0) < \varepsilon\}$$

Es gilt $\mathcal{A}_0 \subset \mathcal{F}$. Insbesondere gilt sicher $\Omega \in \mathcal{F}$ (die Algebra $\mathcal{A}_0$ muss Ω enthalten). Ferner bemerken wir, dass $A^c \Delta A_0^c = A\Delta A_0$. Das bedeutet, dass $A \in \mathcal{F}$ g.d.w. $A^c \in \mathcal{F}$. Wir möchten nun zeigen, dass $\mathcal{F}$ abgeschlossen ist, bezüglich abzählbare Vereinigungen. Sei also $(A_n)_{n\in\mathbb{N}}$ eine Folge in $\mathcal{F}$, $A = \cup_{n\in\mathbb{N}} A_n$ und $\varepsilon > 0$. Da $\mu(A) = \lim_{m\to\infty} \mu(\cup_{n=0}^m A_n)$, wir finden $N \in \mathbb{N}$ so gross, dass $\mu(A \backslash \cup_{n=1}^N A_n) < \varepsilon/2$ (hier ist wichtig, das μ ein endliches Mass ist). Für $j = 1, \dots, N$ benutzen wir dagegen die Tatsache $A_n \in \mathcal{F}$ um ein $B_n \in \mathcal{A}_0$ zu finden, mit $\mu(A_n \Delta B_n) < \varepsilon/2N$. Da $\mathcal{A}_0$ eine Algebra ist, ist dann $B = \cup_{n=1}^N B_n \in \mathcal{A}_0$, und

$$\mu(A\Delta B) \leq \mu(A\backslash \cup_{n=1}^N A_n) + \mu(\cup_{n=1}^N A_n \Delta \cup_{n=1}^N B_n) \leq \varepsilon/2 + \sum_{n=1}^N \mu(A_n \Delta B_n) < \varepsilon$$

weil $(\cup_n A_n)\Delta(\cup_n B_n) \subset \cup_n (A_n \Delta B_n)$. Deswegen ist auch $\cup_{n\in\mathbb{N}} A_n \in \mathcal{F}$. Es folgt, dass $\mathcal{F}$ eine σ-Algebra ist. Da $\mathcal{A}_0 \subset \mathcal{F}$ und $\mathcal{A} = \sigma(\mathcal{A}_0)$, erhalten wir, dass $\mathcal{F} = \mathcal{A}$. Unter der Annahme, dass μ endlich ist, haben wir bewiesen, dass für jede $A \in \mathcal{A}$ und für alle $\varepsilon > 0$ ein $A_0 \in \mathcal{A}_0$ mit $\mu(A\Delta A_0) < \varepsilon$ existiert.

Jetzt betrachten wir ein beliebiges, nicht unbedingt endlich, Mass μ. Sei C_n eine Folge in $\mathcal{A}_0$, mit $\Omega = \cup_n C_n$ und $\mu(C_n) < \infty$ für alle $n \in \mathbb{N}$. O.B.d.A. können wir annehmen, dass die Folge monoton wachsend ist (d.h. $C_n \subset C_{n+1}$); wenn das nicht der Fall ist, ersetzen wir C_n mit $B_n = \cup_{j=1}^n C_j$. Sei nun $A \in \mathcal{A}$, mit $\mu(A) < \infty$. Da $\mu(A) = \lim_{n\to\infty} \mu(A \cap C_n)$, es existiert $N \in \mathbb{N}$ mit $\mu(A \cap C_N) > \mu(A) - \varepsilon/2$; das impliziert auch, dass $\mu(A \backslash (A \cap C_N)) < \varepsilon/2$. Da die Einschränkung von μ auf C_N ein endliches Mass definiert, folgt wie oben, dass ein $B_0 \in \mathcal{A}_0$ existiert so, dass $\mu((A\Delta B_0) \cap C_N) < \varepsilon/2$. Dann ist aber $B_0 \cap C_N \in \mathcal{A}_0$ und

$$\begin{aligned}
\mu(A\Delta(B_0 \cap C_N)) &\leq \mu(A\Delta(A \cap C_N)) + \mu((A \cap C_N)\Delta(B_0 \cap C_N)) \\
&\leq \mu(A\backslash(A \cap C_N)) + \mu((A\Delta B_0) \cap C_N) \\
&< \varepsilon/2 + \varepsilon/2 = \varepsilon
\end{aligned}$$

Die Behauptung folgt, mit $A_0 = B_0 \cap C_N$. $\square$

Beweis von Theorem 3.10. Sei $\mathcal{A}_0$ eine abzählbare Algebra, mit $\mathcal{A} = \sigma(\mathcal{A}_0)$ und so, dass eine Folge $(C_n)_{n\in\mathbb{N}}$ existiert, mit $\Omega = \cup_{n\in\mathbb{N}} C_n$ und $\mu(C_n) < \infty$, für alle $n \in \mathbb{N}$. Wir bezeichnen mit $\mathcal{S}$ die Menge die aus alle einfache Funktionen der Form

$$\sum_{j=1}^{n} d_j \chi_{D_j}$$

besteht, mit $d_j \in \mathbb{Q}$ und $D_j \in \mathcal{A}_0$ so, dass $\mu(D_j) < \infty$. Die Menge $\mathcal{S}$ ist abzählbar, und sie ist in $\mathcal{L}^p(\Omega, \mathcal{A}, \mu)$ enthalten. Wir behaupten, $\mathcal{S}$ definiert eine dichte Teilmenge von $L^p(\Omega, \mathcal{A}, \mu)$. Um diese Behauptung zu zeigen, wählen wir $f \in \mathcal{L}^p(\Omega, \mathcal{A}, \mu)$ und $\varepsilon > 0$. Aus Prop. 3.8 finden wir eine einfache Funktion $g = \sum_{j=1}^n a_j \chi_{A_j} \in \mathcal{L}^p(\Omega, \mathcal{A}, \mu)$, mit geeignete $a_j \in \mathbb{R}$ und $A_j \in \mathcal{A}$ mit $\mu(A_j) < \infty$ so, dass $\|f - g\|_p < \varepsilon/3$. Dann können wir $d_1, \dots, d_n \in \mathbb{Q}$ finden so, dass

$$\|\sum_{j=1}^{n} a_j \chi_{A_j} - \sum_{j=1}^{n} d_j \chi_{A_j}\|_p \leq \sum_{j=1}^{n} |d_j - a_j| \|\chi_{A_j}\|_p = \sum_{j=1}^{n} |d_j - a_j| \mu(A_j)^{1/p} \leq \varepsilon/3$$

Mit Lemma 3.11 finden wir auch Mengen $D_1, \dots, D_n \in \mathcal{A}_0$, mit

$$\mu(A_j \,\Delta\, D_j) < (\varepsilon/3 \sum_{j=1}^{n} |d_j|)^p$$

Das impliziert auch, dass

$$\begin{aligned}\|\sum_{j=1}^{n} d_j \chi_{A_j} - \sum_{j=1}^{n} d_j \chi_{D_j}\|_p &\leq \sum_{j=1}^{n} |d_j| \|\chi_{A_j} - \chi_{D_j}\|_p \\ &= \sum_{j=1}^{n} |d_j| \|\chi_{A_j \Delta D_j}\|_p = \sum_{j=1}^{n} |d_j| \mu(A_j \Delta D_j)^{1/p} < \varepsilon/3\end{aligned}$$

Es folgt, dass $\|f - \sum_{j=1}^n d_j \chi_{D_j}\|_p \leq \varepsilon$ und damit, dass $\mathcal{S}$ dicht in $L^p(\Omega, \mathcal{A}, \mu)$ ist. $\square$

3.4 Dualräume

Sei $(V, \|.\|)$ ein normierter Vektorraum über $\mathbb{R}$. Ein lineare Funktional ist eine lineare Abbildung $\psi : V \to \mathbb{R}$. Das lineare Funktional heisst beschränkt, falls

$$\sup_{x\in V : \|x\| \leq 1} |\psi(x)| = \sup_{x \in V\backslash\{0\}} \frac{|\psi(x)|}{\|x\|} < \infty$$

oder äquivalent, falls ein $C > 0$ existiert, mit $|\psi(x)| \leq C\|x\|$ für alle $x \in V$. Man kann einfach zeigen, dass ein lineares Funktional ist beschränkt, genau dann wenn es stetig ist (Stetigkeit ist, wegen Linearität, mit Stetigkeit an der Stelle $x = 0$ äquivalent, und das ist mit Beschränktheit äquivalent).

Der Raum von allen beschränkten linearen Funktionalen auf V heisst der Dualraum zu V, und wird mit

$$V^* = \{f : V \to \mathbb{R} \text{ linear so, dass } \sup_{\|x\|\leq 1} |f(x)| < \infty\}$$

bezeichnet. Auf V^* ist immer die Norm

$$\|f\|_{V^*} = \sup_{x\in V, \|x\|\leq 1} |f(x)|$$

definiert (Übung: zeigen Sie, dass $\|.\|_{V^*}$ tatsächlich eine Norm ist). Versehen mit dieser Norm ist V^* immer vollständig; mit anderen Wörter ist $(V^*, \|.\|)$ immer ein Banachraum (Beweis: Übung, oder Funktional Analysis).

Bemerke, dass die Definition von V^* von der Wahl der Norm $\|.\|$ auf V abhängt. Ist V endlich dimensional, so sind alle Norme auf V äquivalent, und jede Norm führt deswegen zum selben dualen Raum V^*. In diesem Fall kann man stets V^* mit V identifizieren. Ist nämlich $V \simeq \mathbb{R}^n$, so kann man den Isomorphismus $\phi : V \to V^*$ durch $\phi(v) =: \phi_v : V \to \mathbb{R}$ und

$$\phi_v(w) = \langle v, w\rangle = \sum_{j=1}^n v_i w_i$$

definieren. Ist dim $V = \infty$, so können i.A. V und V^* nicht identifiziert werden (z.B., falls V nicht vollständig ist, kann V^*, als vollständiger Raum, sicher nicht mit V identifiziert werden).

Sei nun $(\Omega, \mathcal{A}, \mu)$ ein Massraum. Es stellt sich die natürliche Frage: was ist der Dualraum von $L^p(\Omega, \mathcal{A}, \mu)$, für $1 \leq p \leq \infty$? Mit anderer Wörter, wie sehen beschränkten Funktionale auf $L^p(\Omega, \mathcal{A}, \mu)$ aus? Um diese Frage zu beantworten, bemerken wir zunächst, dass jede Funktion $f \in L^q(\Omega, \mathcal{A}, \mu)$ ein lineares Funktional auf $L^p(\Omega, \mathcal{A}, \mu)$ definiert. Hier ist $1 \leq q \leq \infty$ so zu wählen, dass $1/p + 1/q = 1$. Wir können nämlich eine Abbildung $\phi : L^q(\Omega, \mathcal{A}, \mu) \to (L^p(\Omega, \mathcal{A}, \mu))^*$ durch $\phi(f) = \phi_f : L^p(\Omega, \mathcal{A}, \mu) \to \mathbb{R}$ und

$$\phi_f(g) = \int fg \, d\mu \tag{38}$$

für alle $f \in L^q(\Omega, \mathcal{A}, \mu)$ und $g \in L^p(\Omega, \mathcal{A}, \mu)$ definieren. Diese Abbildung identifiziert jede $f \in L^q(\Omega, \mathcal{A}, \mu)$ mit einem linearen Funktional auf $L^p(\Omega, \mathcal{A}, \mu)$. Die Abbildung ϕ hat viele wichtige Eigenschaften.

ϕ ist wohldefiniert. Um zu zeigen, dass ϕ wohldefiniert ist, müssen wir überprüfen, dass ϕ_f ein beschränktes lineares Funktionale auf $L^p(\Omega, \mathcal{A}, \mu)$ definiert. Die Linearität $\phi_f(g_1 + \lambda g_2) = \phi_f(g_1) + \lambda\phi_f(g_2)$ folgt offenbar aus der Linearität vom Integral. Es bleibt zu zeigen, dass ϕ_f beschränkt ist. Das folgt aber aus Hölder Ungleichung, da

$$|\varphi_f(g)| = \left| \int fg \, d\mu \right| \leq \|f\|_q \|g\|_p$$

impliziert, dass

$$\|\phi_f\|_{(L^p)^*} = \sup_{g\in L^p : g\neq 0} \frac{|\phi_f(g)|}{\|g\|_p} \leq \|f\|_q < \infty \tag{39}$$

Hier haben wir die Wahl von $1 \leq q \leq \infty$ mit $1/p + 1/q = 1$ benutzt.

ϕ ist linear. Aus der Definition ist offenbar, dass ϕ linear in f ist (d.h. dass $\phi_{f_1+\lambda f_2} = \phi_{f_1} + \lambda \phi_{f_2}$ gilt, für alle $f_1, f_2 \in L^q(\Omega, \mathcal{A}, \mu)$ und alle $\lambda \in \mathbb{R}$).

ϕ ist eine Isometrie. Wir möchten zeigen, dass $\|\phi_f\|_{(L^p)^*} = \|f\|_q$. Wir haben schon in (39) gezeigt, dass $\|\phi_f\|_{(L^p)^*} \leq \|f\|_q$. Es bleibt zu zeigen, dass $\|\phi_f\|_{(L^p)^*} \geq \|f\|_q$. Sei zunächst $1 < p < \infty$ (dann ist auch $1 < q < \infty$). Wir setzen

$$h = \frac{|f|^{q-1}}{\|f\|_q^{q/p}} \operatorname{sgn}(f)$$

wobei $\operatorname{sgn}(f)(x) = 1$, falls $f(x) > 0$, $\operatorname{sgn}(f)(x) = -1$ falls $f(x) < 0$, und $\operatorname{sgn}(f)(x) = 0$ falls $f(x) = 0$. Aus $1/p + 1/q = 1$ folgt, dass $p(q-1) = q$. Also $f \in L^q$ impliziert, dass $h \in L^p$ mit

$$\|h\|_p = \frac{1}{\|f\|_q^{q/p}} \left[\int |f|^{p(q-1)} d\mu \right]^{1/p} = 1$$

Deswegen gilt

$$\|\phi_f\| = \sup_{g \in L^p : \|g\|_p \leq 1} |\phi_f(g)| \geq |\phi_f(h)| = \frac{1}{\|f\|_q^{q/p}} \int |f|^q = \|f\|_q^{q-q/p} = \|f\|_q$$

Seien nun $p = \infty$ und $q = 1$. Für $f \in L^1$ wir definieren $h(x) := \operatorname{sgn} f(x)$. Dann gilt $\|h\|_\infty = 1$ (angenommen $f \neq 0$, als Element von L^1) und

$$\phi_f(h) = \int f h d\mu = \int |f| d\mu = \|f\|_1$$

Deswegen ist auch in diesem Fall $\|\phi_f\| \geq \|f\|_1$. Schlussendlich betrachten wir den Fall $p = 1$ und $q = \infty$. Für $f \in L^\infty$ und $0 < \varepsilon < \|f\|_\infty$ festgewählt, sei

$$A_\varepsilon = \{x \in \Omega : |f(x)| \geq \|f\|_\infty - \varepsilon\}$$

Dann ist $A_\varepsilon \in \mathcal{A}$, mit $\mu(A_\varepsilon) > 0$. Ist μ σ-endlich, so können wir $B_\varepsilon \in \mathcal{A}$ finden, mit $B_\varepsilon \subset A_\varepsilon$ und $0 < \mu(B_\varepsilon) < \infty$. Wir setzen

$$h_\varepsilon = \frac{\chi_{B_\varepsilon}}{\mu(B_\varepsilon)} \operatorname{sgn}(f)$$

Dann es gilt $h_\varepsilon \in L^1$ mit

$$\int |h_\varepsilon| d\mu = \frac{1}{\mu(B_\varepsilon)} \int_{B_\varepsilon} |\operatorname{sgn}(f)| = 1$$

und

$$\phi_f(h_\varepsilon) = \int f h_\varepsilon d\mu = \frac{1}{\mu(B_\varepsilon)} \int_{B_\varepsilon} |f| d\mu \geq \|f\|_\infty - \varepsilon$$

Deswegen ist $\|\phi_f\| \geq \|f\|_\infty - \varepsilon$. Da $\varepsilon > 0$ beliebig ist, folgt, dass $\|\phi_f\| \geq \|f\|_\infty$.

ϕ ist injäktiv. Das folgt aus der Tatsache, dass ϕ isometrisch ist. In der Tat, falls $f, g \in L^q(\Omega, \mathcal{A}, \mu)$ mit $\phi_f = \phi_g$, es folgt, dass $\phi_{f-g} = 0$ und deswegen, dass $\|f - g\|_q = 0$, was auch $f = g$ impliziert.

Wir fassen die Eigenschaften von der Abbildung ϕ im nächsten Theorem zusammen.

Theorem 3.12. *Sei $(\Omega, \mathcal{A}, \mu)$ ein Massraum. Seien $1 \leq p \leq \infty$ und $1 \leq q \leq \infty$ so, dass $1/p + 1/q = 1$. Dann ist die Abbildung $\phi : L^q(\Omega, \mathcal{A}, \mu) \to (L^p(\Omega, \mathcal{A}, \mu))^*$ definiert in (38), eine lineare Isometrie.*

Theorem 3.12 erlaubt uns $L^q(\Omega, \mathcal{A}, \mu)$ mit einem Unterraum von $(L^p(\Omega, \mathcal{A}, \mu))^*$ zu identifizieren, falls $1 \leq p, q \leq \infty$ und $1/p + 1/q = 1$. Die wichtige Frage, die jetzt offen bleibt, ist ob alle beschränkte lineare Funktionale auf $L^p(\Omega, \mathcal{A}, \mu)$ (d.h. alle Elemente von $(L^p(\Omega, \mathcal{A}, \mu))^*$ die Form (38) haben. Mit andere Wörter: ist die Abbildung ϕ surjektiv? In diesem Fall wäre ϕ ein lineares isometrisches Isomorphismus zwischen $L^q(\Omega, \mathcal{A}, \mu)$ und $(L^p(\Omega, \mathcal{A}, \mu))^*$, das uns erlauben würde die zwei Vektorräume zu identifizieren. Wir werden im nächsten Kapitel zeigen (mit Hilfe vom Theorem von Radon-Nikodym), dass, für $1 < p < \infty$, und auch für $p = 1$, falls zusätzlich μ ein σ-endlich Mass ist, ϕ tatsächlich surjektiv ist. Dagegen, für $p = \infty$, ϕ ist nicht surjektiv (es existieren also auf $L^\infty(\Omega, \mathcal{A}, \mu)$ beschränkte lineare Funktionale, die nicht die Form (38), für ein $f \in L^1(\Omega, \mathcal{A}, \mu)$, haben.

3.5 Fourier Transformation

In diesem Abschnitt betrachten wir Funktionen auf dem Massraum $(\mathbb{R}^n, \mathcal{M}_{\lambda_n^*}, \lambda_n)$, mit Werten in $\mathbb{C}$. Wir bezeichnen mit $L^p(\mathbb{R}^n; \mathbb{C})$ der Raum von allen Äquivalenzklassen von Lebesgue messbare Funktionen $f : \mathbb{R}^n \to \mathbb{C}$ so, dass $|f|^p$ integrierbar ist. Bemerke, dass $f \in L^p(\mathbb{R}^n; \mathbb{C})$ genau dann wenn $\operatorname{Re} f, \operatorname{Im} f \in L^p(\mathbb{R}^n)$.

Definition 3.13. *Sei $f \in L^1(\mathbb{R}^n; \mathbb{C})$. Die Fourier Transformierte von f ist die Funktion $\mathcal{F}f : \mathbb{R}^n \to \mathbb{C}$, definiert durch*

$$(\mathcal{F}f)(p) = \frac{1}{(2\pi)^{n/2}} \int f(x) e^{-ip \cdot x} d\lambda_n(x) \,. \tag{40}$$

Bemerke, dass die Fourier Transformierte unabhängig aus der Wahl vom Vertreter der Äquivalenzklasse ist; d.h. $f = g$ λ_n-fast überall impliziert, dass $\mathcal{F}f = \mathcal{F}g$. Aus diesem Grund definiert die Fourier Transformierte eine (lineare) Abbildung auf $L^1(\mathbb{R}^n)$.

Wir beweisen nun ein Paar Eigenschaften der Fourier Transformierte.

Satz 3.14. *Sei $f \in L^1(\mathbb{R}^n; \mathbb{C})$. Dann ist $\mathcal{F}f : \mathbb{R}^n \to \mathbb{C}$ beschränkt, mit*

$$|(\mathcal{F}f)(p)| \leq (2\pi)^{-n/2} \|f\|_1 \,,$$

und stetig.

Beweis. Aus Definition gilt

$$|\mathcal{F}f(p)| \leq \frac{1}{(2\pi)^{n/2}} \left| \int f(x) e^{-ip \cdot x} d\lambda_n(x) \right| \leq \frac{1}{(2\pi)^{n/2}} \int |f(x)| d\lambda_n(x) = \frac{1}{(2\pi)^{n/2}} \|f\|_1$$

Sei p_n eine Folge auf $\mathbb{R}^n$ mit $p_n \to p$. Um die Stetigkeit von $\mathcal{F}f$ zu zeigen, brauchen wir zu beweisen, dass

$$\lim_{n \to \infty} \int f(x) e^{-ip_n \cdot x} d\lambda_n(x) = \int f(x) e^{-ip \cdot x} d\lambda_n(x) \tag{41}$$

Da aber $f(x) e^{-ip_n \cdot x} \to f(x) e^{-ip \cdot x}$ für alle $x \in \mathbb{R}^n$, und $|f(x) e^{-ip_n \cdot x}| \leq |f(x)|$, folgt (41) aus dominierten Konvergenz. $\square$

Eine andere wichtige Eigenschaft der Fourier Transformierte ist aus dem Lemma von Riemann-Lebesgue gegeben.

Lemma 3.15 (Riemann-Lebesgue)**.** *Sei* $f \in L^1(\mathbb{R}^n;\mathbb{C})$*. Dann* $\mathcal{F}f(p) \to 0$ *für* $|p| \to \infty$*.*

Beweis. OBdA können wir annehmen, dass $f \geq 0$ (sonst schreiben wir $f = f_+ - f_-$, und wir beweisen die Behauptung separat für f_+ und für f_-). Wir können auch annehmen, dass f einfach ist. Sonst finden wir aus Prop.3.8 eine Folge von messbare einfache Funktionen f_m, mit $f_m \to f$ in $L^1(\mathbb{R};\mathbb{C})$. Für beliebigen $\delta > 0$ finden wir dann $m \in \mathbb{N}$ mit $\|f_m - f\|_1 \leq \delta$. Dann gilt

$$\begin{aligned}\left|\int f(x)e^{-ip\cdot x}d\lambda_n(x)\right| &\leq \left|\int f_m(x)e^{-ip\cdot x}d\lambda_n(x)\right| + \left|\int (f_m(x)-f(x))e^{-ip\cdot x}d\lambda_n(x)\right| \\ &\leq \left|\int f_m(x)e^{-ip\cdot x}d\lambda(x)\right| + \int |f_m - f|d\lambda_n \\ &\leq \left|\int f_m(x)e^{-ip\cdot x}d\lambda(x)\right| + \delta\end{aligned}$$

Das zeigt, dass es genugt, die Behauptung für einfache Funktionen zu beweisen. Da einfache Funktionen lineare Kombinationen von endlich viele charakteristische Funktionen sind, es genug weiter den Fall $f = \chi_A$ zu betrachten, für $A \in \mathcal{M}_{\lambda_n^*}$. Aus der Regularität von Lebesgue Mass, es genug sogar $f = \chi_U$ für $U \subset \mathbb{R}^n$ offen zu untersuchen. Für jede offene $U \subset \mathbb{R}^n$, finden wir aber eine abzählbare Familie von disjunkten halboffenen Quader der Form

$$Q = \{(x_1,\dots,x_n) \in \mathbb{R}^n : 2^{-k}j_i \leq x_i < 2^{-k}(j_i+1) \text{ für } i = 1,\dots,n\} \tag{42}$$

Es genugt deswegen den Fall $f = \chi_Q$ zu betrachten. In diesem Fall haben wir

$$\begin{aligned}\int e^{ip\cdot x}\chi_Q(x)d\lambda_n(x) &= \prod_{i=1}^n \int_{2^{-k}j_i}^{2^{-k}(j_i+1)} e^{-ip_i\cdot x_i}d\lambda(x_i) \\ &= \prod_{i=1}^n \frac{e^{-ip_i 2^{-k}(j_i+1)} - e^{-ip_i 2^{-k}j_i}}{-ip_i}\end{aligned}$$

Das gibt

$$\left|\int e^{ip\cdot x}\chi_Q(x)d\lambda_n(x)\right| \leq C\frac{1}{|p_1|\dots|p_n|} \to 0$$

für $|p| \to \infty$. $\square$

Wir berechnen die Fourier Transformierte von Gaussische Funktionen.

Lemma 3.16. *Für* $\alpha > 0$ *sei* $g_\alpha : \mathbb{R}^n \to \mathbb{R}$ *definiert durch* $g_\alpha(x) = e^{-\frac{\alpha x^2}{2}}$*. Dann gilt*

$$(\mathcal{F}g_\alpha)(p) = \alpha^{-n/2}e^{-\frac{p^2}{2\alpha}}$$

Beweis. Da $g_\alpha \in L^1(\mathbb{R}^n)$, gilt

$$(\mathcal{F}g_\alpha)(p) = \frac{1}{(2\pi)^{n/2}} \prod_{j=1}^n \int_{-\infty}^\infty dx_j e^{-\frac{\alpha}{2}x_j^2} e^{-ip_j\cdot x_j} = \frac{e^{-\frac{p^2}{2\alpha}}}{(2\pi)^{n/2}} \prod_{j=1}^n \int_{-\infty}^\infty dx_j\, e^{-\frac{\alpha}{2}\left(x_j - \frac{ip_j}{\alpha}\right)^2} \tag{43}$$

Mit $z_j = x_j - ip_i/\alpha$ das Integral ist ein komplexes Linienintegral auf einem Weg im komplexen Ebenen parallel zum reellen Achse. Da die Funktion $e^{-\alpha z_j^2/2}$ auf $\mathbb{C}$ holomorph ist, und für $|\mathrm{Re}\; z_j| \to \infty$ gegen Null strebt, kann man das Integral zurück auf dem reellen Achse schieben. D.h.

$$\int_{-\infty}^\infty dx_i\, e^{-\frac{\alpha}{2}\left(x_i - \frac{ip_i}{\alpha}\right)^2} = \int_{-\infty}^\infty dx_j\, e^{-\frac{\alpha x_j^2}{2}} = \sqrt{\frac{2\pi}{\alpha}}$$

Einsetzen in (43) zeigt die Behauptung. □

Die Fourier Transformation spielt in der Analysis eine sehr wichtige Rolle. Ein Grund dafür ist die Tatsache, dass sie Differentialoperatoren "diagonalisiert".

Lemma 3.17. *Sei $f \in L^1(\mathbb{R}^n;\mathbb{C})$ und $g : \mathbb{R}^n \to \mathbb{C}^n$ definiert durch $g(x) = -ixf(x)$. Ist g integrierbar, so gilt $\mathcal{F}f \in C^1(\mathbb{R}^n;\mathbb{C})$ mit $\nabla(\mathcal{F}f) = \mathcal{F}g$.*

Beweis. Seien $p \in \mathbb{R}^n$, $j \in \{1,\dots,n\}$, $\delta \in \mathbb{R}\backslash\{0\}$ festgewählt. Dann gilt

$$\begin{aligned}\frac{\mathcal{F}f(p+\delta e_j) - \mathcal{F}f(p)}{\delta} &= \int \frac{e^{-ix\cdot(p+\delta e_j)} - e^{-ix\cdot p}}{\delta} f(x) d\lambda_n(x) \\ &= \int \frac{e^{-i\delta x_j} - 1}{\delta} e^{-ix\cdot p} f(x) d\lambda_n(x)\end{aligned}$$

Da $|e^{-i\delta x_j} - 1| \leq |\delta||x_j| \leq |\delta||x|$, finden wir

$$\left| \frac{e^{-i\delta x_j} - 1}{\delta} e^{-ix\cdot p} f(x) \right| \leq |x||f(x)| = |g(x)|$$

Da die rechte Seite integrierbar ist, folgt aus dominierte Konvergenz, dass

$$\lim_{\delta\to 0} \frac{\mathcal{F}f(p+\delta e_j) - \mathcal{F}f(p)}{\delta} = -i \int x_j f(x) e^{-ix\cdot p} d\lambda_n(x)$$

Das zeigt, dass $\mathcal{F}f$ in allen Richtungen partiel differenzierbar ist, und, dass $\partial_j(\mathcal{F}f) = \mathcal{F}g_j$. Aus der Annahme $g \in L^1(\mathbb{R}^n;\mathbb{C})$ folgt, dass $\mathcal{F}g_j$ stetig ist, für alle $j = 1,\dots,n$. Da alle partielle Ableitungen von $\mathcal{F}f$ existieren und stetig sind, es folgt, dass $\mathcal{F}f \in C^1(\mathbb{R}^n;\mathbb{C})$, mit $\nabla(\mathcal{F}f) = \mathcal{F}g$. □

Bemerkung: Sei $f \in L^1(\mathbb{R}^n;\mathbb{C})$. Für ein Multiindex $\alpha \in \mathbb{N}^n$ sei $g_\alpha : \mathbb{R}^n \to \mathbb{C}$ mit $g(x) = (-i)^{|\alpha|} x^\alpha f(x)$. Ist g_α integrierbar (d.h. ist $g_\alpha \in L^1(\mathbb{R}^n;\mathbb{C})$), für alle $\alpha \in \mathbb{N}^n$ mit $|\alpha| \leq k$, so gilt $\mathcal{F}f \in C^k(\mathbb{R}^n;\mathbb{C})$, mit $D^\alpha(\mathcal{F}f) = \mathcal{F}g_\alpha$. Es gibt also eine Beziehung zwischen den Zerfall einer Funktion (damit $x^\alpha f(x)$ integrierbar ist, muss f ins Unendlichen schnell abfallen) und die Regularität der Fourier Transformierte. Es folgt aus Theorem

3.20 unten, dass auch das Gegenteil gilt: um so glatt eine Funktion f ist, desto schneller fällt die Fourier Transformierte $\mathcal{F}f(p)$ für $p \to \infty$ ab.

Seien $f, g \in L^1(\mathbb{R}^n; \mathbb{C})$. Wir definieren die Faltung $f * g$ von $f$und g durch

$$f * g(x) = \int f(x-y)g(y)dy$$

Es wurde in den Übungen bewiesen, dass $f * g$ wohldefiniert ist, und, dass $f * g \in L^1(\mathbb{R}^n; \mathbb{C})$. Wir können also die Fourier Transformierte von $f * g$ berechnen.

Satz 3.18. *Seien $f, g \in L^1(\mathbb{R}^n; \mathbb{C})$. Dann ist*

$$\mathcal{F}(f * g) = (2\pi)^{n/2} \mathcal{F}(f)\mathcal{F}(g)$$

Beweis. Wir haben

$$\begin{aligned} \mathcal{F}(f * g)(p) &= \frac{1}{(2\pi)^{n/2}} \int (f * g)(x) e^{-ip\cdot x} d\lambda_n(x) \\ &= \frac{1}{(2\pi)^{n/2}} \int \left[\int f(x-y)g(y)d\lambda_n(y) \right] e^{-ip\cdot x} d\lambda_n(x) \end{aligned}$$

Mit Fubini (die Funktion $f(x-y)g(y)e^{-ip\cdot x}$ ist auf $\mathbb{R}^{2n}$ integrierbar, weil $f, g \in L^1(\mathbb{R}^n)$) finden wir

$$\begin{aligned} \mathcal{F}(f * g)(p) &= \frac{1}{(2\pi)^{n/2}} \int f(x-y)g(y)e^{-ip\cdot x} d\lambda_{2n}(x,y) \\ &= \frac{1}{(2\pi)^{n/2}} \int f(x)g(y)e^{-ip\cdot(x+y)} d\lambda_{2n}(x,y) \\ &= (2\pi)^{n/2} \left[\frac{1}{(2\pi)^{n/2}} \int f(x)e^{-ip\cdot x} d\lambda_n(x) \right] \left[\frac{1}{(2\pi)^{n/2}} \int g(y)e^{-ip\cdot y} d\lambda_n(y) \right] \\ &= (2\pi)^{n/2} (\mathcal{F}f)(p)(\mathcal{F}g)(p) \end{aligned}$$

wobei wir den Transformationssatz (um die Variable (x, y) mit $(x + y, y)$ zu ersetzen) und noch einmal Fubini benutzt haben. $\square$

Im nächsten Theorem zeigen wir, dass die Fourier Transformation auf $L^2(\mathbb{R}^n)$ isometrisch wirkt.

Theorem 3.19 (Plancherel)**.** *Sei $f \in L^1(\mathbb{R}^n; \mathbb{C}) \cap L^2(\mathbb{R}^n; \mathbb{C})$. Dann ist $\mathcal{F}f \in L^2(\mathbb{R}^n; \mathbb{C})$ und $\|\mathcal{F}f\|_2 = \|f\|_2$.*

Beweis. Die Funktion $e^{-\varepsilon|k|^2/2}$ ist monoton in ε. Das monoton Konvergenz Theorem zeigt, dass

$$\|\mathcal{F}f\|_2^2 = \lim_{\varepsilon \to 0} \int |(\mathcal{F}f)(p)|^2 e^{-\varepsilon \frac{|p|^2}{2}} d\lambda_n(p) \,.$$

(A priori könnten linke und rechte Seite unendlich sein; wir werden aber zeigen, dass sie endlich sind). Wir setzen nun die Definition von $(\mathcal{F}f)(p)$ ein. Da $\overline{f(x)}f(y)e^{-\varepsilon|p|^2/2}$ als

Funktion der drei Variablen $x, y, p \in \mathbb{R}^n$ integrierbar ist (da $f \in L^1$), können wir Fubini anwenden. Wir finden

$$\begin{aligned}
\|\mathcal{F}f\|_2^2 &= \lim_{\varepsilon \to 0} \frac{1}{(2\pi)^3} \int d\lambda_{3n}(x,y,p) \overline{f(x)} f(y) e^{ip\cdot(x-y)} e^{-\varepsilon \frac{p^2}{2}} \\
&= \lim_{\varepsilon \to 0} \frac{1}{(2\pi)^3} \int d\lambda_{2n}(x,y) \overline{f(x)} f(y) \int e^{ip\cdot(x-y)} e^{-\varepsilon \frac{p^2}{2}} d\lambda_n(p) \\
&= \lim_{\varepsilon \to 0} \int \overline{f(x)} \left[\frac{1}{(2\pi\varepsilon)^{n/2}} \int f(y) e^{-\frac{(x-y)^2}{2\varepsilon}} d\lambda_n(y) \right] d\lambda_n(x) \\
&= \lim_{\varepsilon \to 0} \int \overline{f(x)} (f * j_\varepsilon)(x) dx
\end{aligned}$$

wo wir $j(x) = (2\pi)^{-n/2} e^{-x^2/2}$ und $j_\varepsilon(x) = \varepsilon^{-n/2} j(x/\sqrt{\varepsilon})$ definiert haben ($*$ ist hier die Faltung; bemerke, dass j so definiert ist, dass $\|j_\varepsilon\|_1 = 1$, unabhängig aus $\varepsilon > 0$). Wir erhalten

$$|\mathcal{F}f\|_2^2 = \|f\|_2^2 + \lim_{\varepsilon \to 0} \int \overline{f(x)} (f * j_\varepsilon)(x) - f(x)) dx \tag{44}$$

Mit Hölder, wir können abschätzen,

$$\left| \int \overline{f(x)} (f * j_\varepsilon)(x) - f(x)) dx \right| \leq \|f\|_2 \|f - f * j_\varepsilon\|_2$$

D.h. die Behauptung folgt aus (44), falls wir zeigen können, dass

$$\lim_{\varepsilon \to 0^+} \|f - f * j_\varepsilon\|_2 = 0 \tag{45}$$

Um (45) zu zeigen, bemerken wir zunächst, dass für ein beliebiges $g \in L^2(\mathbb{R}^n; \mathbb{C})$,

$$\|f - f * j_\varepsilon\|_2 \leq \|f - g\|_2 + \|g - g * j_\varepsilon\|_2 + \|(f - g) * j_\varepsilon\|_2 \tag{46}$$

Um den letzten Term abzuschätzen, benutzen wir die Young'sche Ungleichung (Beweis: Übung): Seien $1 \leq p, q \leq \infty$ mit $1 + 1/2 = 1/p + 1/q$. Dann gilt $\|h * \varphi\|_2 \leq \|h\|_p \|\varphi\|_q$ für alle $h \in L^p(\mathbb{R}^n; \mathbb{C})$ und alle $\varphi \in L^q(\mathbb{R}^n; \mathbb{C})$. Aus (46) finden wir, da $\|j_\varepsilon\|_1 = \|j\|_1 = 1$ für alle $\varepsilon > 0$,

$$\|f - f * j_\varepsilon\|_2 \leq 2\|f - g\|_2 + \|g - g * j_\varepsilon\|_2$$

Das bedeutet, es genügt (45) für f in einem dichten Teilraum von $L^2(\mathbb{R}^n; \mathbb{C})$ zu zeigen. Zum Beispiel, es genügt (45) für einfache Funktionen in $L^2(\mathbb{R}^n; \mathbb{C})$ zu zeigen. Da einfache Funktionen endliche lineare Kombinationen von charakteristische Funktionen sind, es genügt (45) für charakteristische Funktionen von messbare Mengen in $\mathbb{R}^n$ zu zeigen. Da $\|\chi_A - \chi_B\|_2 = \mu(A\Delta B)^{1/2}$, und da messbare Menge durch offene Menge approximiert werden können (weil λ_n regulär ist), es genügt (45) für charakteristische Funktionen von offenen Mengen zu zeigen. Da offene Menge durch endliche Vereinigungen von Quader approximiert werden können, es genügt (45) für $f = \chi_Q$ zu beweisen, wobei $Q = [a_1; b_1] \times \cdots \times [a_n; b_n]$. In diesem Fall können wir (45) beweisen, indem wir berechnen

$$\begin{aligned}
\|\chi_Q - \chi_Q * j_\varepsilon\|_2^2 &= \int \left| \chi_Q(x) - \int dy \chi_Q(y) j_\varepsilon(x-y) dy \right|^2 dx \\
&= \int \left| \int \left[\chi_Q(x + \varepsilon y) - \chi_Q(x) \right] j(y) dy \right|^2 dx \\
&\leq \int \left| \chi_Q(x + \varepsilon y) - \chi_Q(x) \right|^2 j(y) \, dxdy
\end{aligned}$$

Durch separate Behandlung der Gebiete $|y| \leq R$ und $|y| > R$ erhalten wir

$$\|\chi_Q - \chi_Q * j_\varepsilon\|_2^2 \leq C\varepsilon R + Ce^{-R^2/4}$$

für eine geeignete Konstante $C > 0$, und für alle $R > 0$ (für $|y| < R$ benutzen wir, dass $\chi_Q(x + \varepsilon y) - \chi_Q(x) = 0$, ausser wenn x in einem Gebiet mit Volumen der Ordnung εR ist; für $|y| > R$ benutzen wir dagegen, dass $j(y) \leq e^{-R^2/4} e^{-y^2/4}$). Mit zB. $R = \varepsilon^{-1/2}$, erhalten wir (45). □

Das Theorem von Plancherel erlaubt uns die Fourier Transformation auf $L^2(\mathbb{R}^n)$ zu definieren. Für $f \in L^2(\mathbb{R}^n; \mathbb{C})$ (nicht unbedingt in $L^1(\mathbb{R}^n; \mathbb{C})$), finden wir eine Folge $f_j \in L^1(\mathbb{R}^n; \mathbb{C}) \cap L^2(\mathbb{R}^n; \mathbb{C})$, mit $\|f_j - f\|_2 \to 0$, für $j \to \infty$. Für jede $j \in \mathbb{N}$ können wir dann $\mathcal{F} f_j$ wie in Definition 3.13 definieren (weil $f_j \in L^1(\mathbb{R}^n; \mathbb{C})$). Aus Theorem 3.19 folgt, dass $\mathcal{F} f_j \in L^2(\mathbb{R}^n; \mathbb{C})$ für alle $j \in \mathbb{N}$. Wir bemerken, dass $\|\mathcal{F} f_j - \mathcal{F} f_m\|_2 = \|\mathcal{F}(f_j - f_m)\|_2 = \|f_j - f_m\|_2 \to 0$, für $j, m \to \infty$, weil f_j in $L^2(\mathbb{R}^n; \mathbb{C})$ konvergiert und deswegen eine Cauchy Folge ist. Das zeigt, dass auch $\mathcal{F} f_j$ eine Cauchy Folge auf $L^2(\mathbb{R}^n; \mathbb{C})$ ist. Da $L^2(\mathbb{R}^n; \mathbb{C})$ vollständig ist, ist die Folge $\mathcal{F} f_j$ konvergent. Wir setzen $\mathcal{F} f := \lim_{j \to \infty} \mathcal{F} f_j$.

Bemerke, dass $\mathcal{F} f$ wohldefiniert ist (d.h. der Limes hängt nicht aus der Wahl der Folge f_j). Seien nämlich f_j und g_j zwei Folgen in $L^1(\mathbb{R}^n; \mathbb{C}) \cap L^2(\mathbb{R}^n; \mathbb{C})$ mit $f_j \to f$ und $g_j \to f$ in $L^2(\mathbb{R}^n; \mathbb{C})$. Wir konstruieren eine neue Folge h_j durch $h_{2j} = f_j$ und $h_{2j+1} = g_j$. Dann $h_j \to f$ für $j \to \infty$, ist h_j eine Cauchy Folge. Damit ist auch $\mathcal{F} h_j$ Cauchy, und deswegen konvergent. Aus diesem Grund muss $\lim_{j \to \infty} \mathcal{F} f_j = \lim_{j \to \infty} \mathcal{F} g_j$.

Damit haben wir $\mathcal{F} : L^2(\mathbb{R}^n; \mathbb{C}) \to L^2(\mathbb{R}^n; \mathbb{C})$ definiert. Aus Theorem 3.19 folgt auch, dass $\mathcal{F}$ eine Isometrie ist, d.h., dass $\|\mathcal{F} f\|_2 = \|f\|_2$ für alle $f \in L^2(\mathbb{R}^n; \mathbb{C})$. Ist nämlich $f_j \in L^1(\mathbb{R}^n; \mathbb{C}) \cap L^2(\mathbb{R}^n; \mathbb{C})$ mit $f_j \to f$ in $L^2(\mathbb{R}^n; \mathbb{C})$, so gilt $\|\mathcal{F} f_j\|_2 = \|f_j\|_2$ für alle $j \in \mathbb{N}$, aus Theorem 3.19; $f_j \to f$ und $\mathcal{F} f_j \to \mathcal{F} f$ in $L^2(\mathbb{R}^n; \mathbb{C})$ implizieren, dass

$$\|\mathcal{F} f\|_2 = \lim_{j \to \infty} \|\mathcal{F} f_j\|_2 = \lim_{j \to \infty} \|f_j\|_2 = \|f\|_2 \tag{47}$$

Eigentlich folgt aus (47) auch, dass $\mathcal{F}$ das Skalarprodukt auf $L^2(\mathbb{R}^n; \mathbb{C})$, definiert durch

$$\langle f, g \rangle = \int \bar{f} g d\mu$$

invariant lässt, d.h.

$$\langle \mathcal{F} f, \mathcal{F} g \rangle = \langle f, g \rangle$$

für alle $f, g \in L^2(\mathbb{R}^n; \mathbb{C})$. Das kann durch die Identität

$$\langle f, g \rangle = \frac{1}{2} \left[\|f + g\|_2^2 - i\|f + ig\|_2^2 - (1 - i)\|f\|^2 - (1 - i)\|g\|^2 \right]$$

bewiesen werden (weil Theorem 3.19 zeigt, dass Normen durch $\mathcal{F}$ erhalten bleiben).

Obwohl wir für diese Abbildung auf $L^2(\mathbb{R}^n; \mathbb{C})$ dasselbe Symbol wie für die Abbildung auf $L^1(\mathbb{R}^n; \mathbb{C})$ (siehe Definition 3.13) benutzten, es handelt sich um zwei verschiedenen Abbildungen. In allgemein gilt für $f \in L^2(\mathbb{R}^n; \mathbb{C})$ die Formel (40) nicht (die Formel (40) gilt für ein $f \in L^2(\mathbb{R}^n; \mathbb{C})$ nur falls f auch in $L^1(\mathbb{R}^n; \mathbb{C})$ ist).

Es folgt aus Theorem 3.19, dass $\mathcal{F} : L^2(\mathbb{R}^n;\mathbb{C}) \to L^2(\mathbb{R}^n;\mathbb{C})$ eine Isometrie ist. Wir zeigen nun, dass $\mathcal{F}$ auf $L^2(\mathbb{R}^n;\mathbb{C})$ unitär ist. Für $f \in L^2(\mathbb{R}^n;\mathbb{C})$ definieren wir die inverse Fourier transformierte von f als

$$(\mathcal{F}^* f)(p) := (\mathcal{F}f)(-p)$$

Bemerke, dass $\mathcal{F}^* : L^2(\mathbb{R}^n;\mathbb{C}) \to L^2(\mathbb{R}^n;\mathbb{C})$ das adjungierte Abbildung zu $\mathcal{F}$ ist. Das bedeutet, dass

$$\langle f, \mathcal{F}g \rangle = \langle \mathcal{F}^* f, g \rangle$$

für alle $f, g \in L^2(\mathbb{R}^n;\mathbb{C})$; das erklärt die Notation.

Theorem 3.20. *Sei $f \in L^2(\mathbb{R}^n;\mathbb{C})$. Dann gilt $f = \mathcal{F}^*(\mathcal{F}f)$.*

Beweis. Für $\alpha > 0$, wir setzen $g_\alpha(x) = e^{-\alpha \frac{x^2}{2}}$. Aus Lemma 3.16, finden wir

$$(\mathcal{F}g_\alpha)(p) = \alpha^{-n/2} e^{-\frac{p^2}{2\alpha}}$$

Man bemerke, dass eine weitere Anwendung von Lemma 3.16 ergibt

$$(\mathcal{F}^*(\mathcal{F}g_\alpha))(x) = (\mathcal{F}(\mathcal{F}g_\alpha))(-x) = g_\alpha(x)$$

D.h. die Behauptung gilt für die Funktionen g_α. Nun zeigen wir, sie gilt für alle $f \in L^2(\mathbb{R}^n;\mathbb{C})$. Sei zunächst $f \in L^1(\mathbb{R}^n;\mathbb{C}) \cap L^2(\mathbb{R}^n;\mathbb{C})$. Dann gilt (mit Fubini)

$$\begin{aligned} \int (\mathcal{F}g_\alpha)(y-x) f(y) d\lambda_n(y) &= \int \left[\frac{1}{(2\pi)^{n/2}} \int g_\alpha(k) e^{-ik\cdot(y-x)} d\lambda_n(k) \right] f(y) d\lambda_n(y) \\ &= \int g_\alpha(k) (\mathcal{F}f)(k)\, e^{-ik\cdot x} d\lambda_n(k) \end{aligned} \tag{48}$$

Für $f \in L^2(\mathbb{R}^n;\mathbb{C})$ finden wir nun eine Folge $f_j \in L^1(\mathbb{R}^n;\mathbb{C}) \cap L^2(\mathbb{R}^n;\mathbb{C})$ mit $f_j \to f$ in $L^2(\mathbb{R}^n)$ so, dass f_j die Identität (48) erfüllt, für alle $j \in \mathbb{N}$. Aus Theorem 3.19 folgt auch, dass $\mathcal{F}f_j \to \mathcal{F}f$ in $L^2(\mathbb{R}^n;\mathbb{C})$ für $j \to \infty$. Da $(\mathcal{F}g_\alpha), g_\alpha \in L^2(\mathbb{R}^n;\mathbb{C})$, es folgt, dass

$$\int (\mathcal{F}g_\alpha)(y-x) f(y) d\lambda_n(y) = \int g_\alpha(k) (\mathcal{F}f)(k)\, e^{-ik\cdot x} d\lambda_n(k) \tag{49}$$

auch für alle $f \in L^2(\mathbb{R}^n)$. Für $\alpha \to 0$ konvergiert die linke Seite von (49) gegen f, in $L^2(\mathbb{R}^n)$; das folgt wie in (45), im Beweis von Theorem 3.19 (da $\mathcal{F}g_\alpha$ ist eine skalierte Gauss-Funktion, mit die selbe Form wie j_ε in (45)). Was passiert zur rechten Seite, falls $\alpha \to 0$? Dominierte Konvergenz impliziert, dass $g_\alpha(\mathcal{F}f) \to (\mathcal{F}f)$ in $L^2(\mathbb{R}^n;\mathbb{C})$. Theorem 3.19 impliziert also, dass $\mathcal{F}^*(g_\alpha(\mathcal{F}f)) \to \mathcal{F}^*(\mathcal{F}f)$ für $\alpha \to 0$. Das zeigt, dass $f = \mathcal{F}^*(\mathcal{F}f)$. □

Zusammenfassend: $\mathcal{F} : L^2(\mathbb{R}^n;\mathbb{C}) \to L^2(\mathbb{R}^n;\mathbb{C})$ ist eine unitäre Abbildung.

Wir haben gezeigt, dass die Fourier Transformation auf $L^1(\mathbb{R}^n;\mathbb{C})$ oder auf $L^2(\mathbb{R}^n;\mathbb{C})$ definiert werden kann. Einerseits wird $L^1(\mathbb{R}^n;\mathbb{C})$ auf $L^\infty(\mathbb{R}^n;\mathbb{C})$ abgebildet, mit

$$\|\mathcal{F}f\|_\infty \leq (2\pi)^{-n/2} \|f\|_1$$

für alle $f \in L^1(\mathbb{R}^n;\mathbb{C})$. Anderseits wird $L^2(\mathbb{R}^n;\mathbb{C})$ auf $L^2(\mathbb{R}^n;\mathbb{C})$ abgebildet, mit $\|\mathcal{F}f\|_2 = \|f\|_2$. Sei nun $1 \leq p \leq 2$. Für $f \in L^p(\mathbb{R}^n;\mathbb{C}) \cap L^1(\mathbb{R}^n;\mathbb{C})$ gilt die Hausdorff-Young Ungleichung

$$\|\mathcal{F}f\|_q \leq C\|f\|_p \tag{50}$$

für $2 \leq q \leq \infty$ mit $1/p + 1/q = 1$ und für eine geeignete Konstante $C > 0$ (die nur von p abhängt). Mit (50) kann man nun die Fourier Transformation auf $L^p(\mathbb{R}^n)$ definieren. Ist $f \in L^p(\mathbb{R}^n;\mathbb{C})$, so finden wir eine Folge $f_j \in L^p(\mathbb{R}^n;\mathbb{C}) \cap L^1(\mathbb{R}^n;\mathbb{C})$ mit $f_j \to f$ in $L^p(\mathbb{R}^n;\mathbb{C})$. Dann ist $\mathcal{F}f_j$ wegen (50) eine Cauchy Folge auf $L^q(\mathbb{R}^n;\mathbb{C})$. Deswegen konvergiert $\mathcal{F}f_j$ in $L^q(\mathbb{R}^n;\mathbb{C})$. Wir setzen dann einfach $\mathcal{F}f := \lim_{j\to\infty} \mathcal{F}f_j$. Es ist einfach zu sehen, dass $\mathcal{F}$ damit wohldefiniert ist. Das definiert eine Abbildung $\mathcal{F} : L^p(\mathbb{R}^n;\mathbb{C}) \to L^q(\mathbb{R}^n;\mathbb{C})$, für alle $1 \leq p \leq 2$. Diese neue Abbildung erfüllt die Hausdorff-Young Ungleichung $\|\mathcal{F}f\|_q \leq C\|f\|_p$.

4 Absolute Stetigkeit und das Radon-Nikodym Theorem

4.1 Signierte Masse

Es ist möglich den Begriff von Mass zu verallgemeinern, indem man die Positivität, nämlich die Bedingung, dass $\mu(A) \geq 0$ für alle $A \subset \Omega$, nicht mehr verlangt. Das führt zum Begriff von signiertem Mass.

Definition 4.1. *Sei $(\Omega, \mathcal{A})$ ein messbarer Raum. Eine Funktion $\mu : \mathcal{A} \to [-\infty;\infty]$ heisst ein signiertes Mass auf $(\Omega, \mathcal{A})$ wenn $\mu(\emptyset) = 0$ und wenn, für jede Folge $(A_n)_{n\in\mathbb{N}}$ von disjukten Mengen in $\mathcal{A}$,*

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

Damit $\mu : \mathcal{A} \to [-\infty;\infty]$ ein signiertes Mass ist, muss die Summe $\sum_{n\in\mathbb{N}} \mu(A_n)$ wohldefiniert sein, für jede Folge disjunkter Mengen in $\mathcal{A}$. Insbesonde kann μ nur den Wert $+\infty$ oder den Wert $-\infty$ annehmen, aber nicht beide. Ist $|\mu(\Omega)| < \infty$ so muss $|\mu(A)| < \infty$ endlich sein, für alle $A \subset \Omega$ (sonst würde die Identität $\mu(A) + \mu(A^c) = \mu(\Omega)$ kein Sinn haben). In diesem Fall sagen wir, dass μ ein endliches signiertes Mass ist. Ist anderseits $\mu(\Omega) = +\infty$, so kann es kein $A \in \mathcal{A}$ existieren mit $\mu(A) = -\infty$ (in diesem Fall nimmt μ Werten in $(-\infty;\infty]$). Ist dagegen $\mu(\Omega) = -\infty$, so wird es kein $A \subset \Omega$ existieren mit $\mu(A) = +\infty$ und μ nimmt Werten in $[-\infty;\infty)$.

Beispiel. Sei $(\Omega, \mathcal{A}, \mu)$ ein Massraum (mit einem positiven Mass μ), und sei $f \in \mathcal{L}^1(\Omega, \mathcal{A}, \mu)$ eine reelwertige integrierbare Funktion auf Ω. Dann definiert $\nu(A) = \int_A f d\mu$ ein signiertes Mass auf $\mathcal{A}$ (die abzählbare Additivität folgt aus dominierte Konvergenz). Mit $f_+ = \max(f, 0)$ und $f_- = -\min(f, 0)$ gilt $f = f_+ - f_-$ und $\nu(A) = \nu_+(A) - \nu_-(A)$ für alle $A \in \mathcal{A}$, mit den (positiven) Masse $\nu_\pm$ definiert durch $\nu_\pm(A) = \int_A f_\pm d\mu$.

Es ist einfach zu üeberprüfen, dass ein signiertes Mass viele der Eigenschaften eines positiven Mass hat. Z.B. wenn $(A_n)_{n\in\mathbb{N}}$ eine wachsende Folge von Mengen in $\mathcal{A}$ (d.h. $A_n \subset A_{n+1}$ für alle $n \in \mathbb{N}$), dann gilt

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \lim_{n\to\infty} \mu(A_n)$$

Ist dagegen $(A_n)_{n\in\mathbb{N}}$ eine fallende Folge in $\mathcal{A}$ mit $\mu(A_n)$ endlich für ein $n \in \mathbb{N}$, so gilt

$$\mu\left(\bigcap_{n\in\mathbb{N}} A_n\right) = \lim_{n\to\infty} \mu(A_n)\,. \tag{51}$$

Wie für positive Masse kann man sogar zeigen, dass die letzte zwei Eigenschaften hinreichend sind, um σ-additive Masse zu charakterisieren. Mit anderen Wörter, falls $(\Omega, \mathcal{A})$ ein messbarer Raum ist, und falls $\mu : \mathcal{A} \to [-\infty; \infty]$ so definiert ist, dass $\mu(\emptyset) = 0$ und $\mu(A \cup B) = \mu(A) + \mu(B)$ für alle $A, B \in \mathcal{A}$ mit $A \cap B = \emptyset$ (d.h. μ ist endlich additiv), dann ist μ ein Mass, falls entweder

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

für alle wachsende Folge A_n in $\mathcal{A}$, oder, falls

$$\lim_{n\to\infty} \mu(A_n) = 0 \tag{52}$$

für alle fallende Folge A_n in $\mathcal{A}$, mit $\bigcap_{n=1}^{\infty} A_n = \emptyset$.

Bei signierten Masse wird eine wichtige Rolle von positiven und negativen Mengen in der σ-Algebra gespielt, die wir jetzt definieren möchten.

Definition 4.2. *Sei $(\Omega, \mathcal{A})$ ein messbarer Raum und μ ein signiertes Mass auf $(\Omega, \mathcal{A})$. Wir sagen $A \in \mathcal{A}$ ist eine positive Menge für μ, falls $\mu(B) \geq 0$ für alle $B \subset A$. Wir sagen $A \in \mathcal{A}$ ist eine negative Menge für μ, falls $\mu(B) \leq 0$ für alle $B \subset A$.*

Lemma 4.3. *Sei μ ein signiertes Mass auf ein messbarer Raum $(\Omega, \mathcal{A})$, und sei $A \in \mathcal{A}$ mit $-\infty < \mu(A) < 0$. Dann existiert eine negative Menge $B \in \mathcal{A}$ mit $B \subset A$ und $\mu(B) \leq \mu(A)$.*

Beweis. Sei

$$\delta_1 = \sup\{\mu(E) : E \in \mathcal{A} \text{ und } E \subset A\}$$

und sei $A_1 \in \mathcal{A}$ mit $A_1 \subset A$ und mit

$$\mu(A_1) \geq \min(1, \delta_1/2)$$

Sei nun

$$\delta_2 = \sup\{\mu(E) : E \in \mathcal{A} \text{ und } E \subset A \backslash A_1\}$$

und $A_2 \in \mathcal{A}$ mit $A_2 \subset A \backslash A_1$ mit

$$\mu(A_2) \geq \min(1, \delta_2/2)$$

Wir iterieren diese Konstruktion und definieren

$$\delta_n = \sup\left\{\mu(E) : E \in \mathcal{A} \text{ und } E \subset A \backslash \cup_{j=1}^{n-1} A_j\right\}$$

und dann wählen wir $A_n \in \mathcal{A}$ mit $A_n \subset A \backslash \bigcup_{j=1}^{n} A_j$ und mit

$$\mu(A_n) \geq \min(1, \delta_n/2)$$

Wir definieren dann

$$A_\infty = \bigcup_{n=1}^{\infty} A_n \quad \text{und} \quad B = A \backslash A_\infty$$

Wir behaupten, dass B die gewünschte Eigenschaften hat. In der Tat, $A_n \in \mathcal{A}$ sind disjunkt mit $\mu(A_n) \geq 0$ für alle $n \in \mathbb{N}$. Also

$$\mu(A) = \mu(A_\infty) + \mu(B) \geq \mu(B)$$

Wir müssen die Tatsache zeigen, dass B eine negative Menge ist. Da $-\infty < \mu(A) < 0$, muss $\mu(A_\infty) < \infty$ sein. Da aber $\mu(A_\infty) = \sum_n \mu(A_n)$, es folgt, dass $\mu(A_n) \to 0$. Deswegen muss $\delta_n \to 0$. Für ein beliebiges $E \subset B$ haben wir nun $\mu(E) \leq \delta_n$ für alle $n \in \mathbb{N}$, es muss $\mu(E) \leq 0$ gelten. $\square$

Mit Hilfe des letzten Lemma können wir nun ein beliebiges signiertes Mass in der Differenz zwei positiver Masse zerlegen. Dazu brauchen wir zunächst den Hahn'sche Zerlegungstheorem.

Theorem 4.4 (Hahn'sche Zerlegungstheorem)**.** *Sei $(\Omega, \mathcal{A})$ ein messbarer Raum, und sei μ ein signiertes Mass auf $(\Omega, \mathcal{A})$. Dann existieren $P, N \in \mathcal{A}$ so, dass P eine positive und N eine negative Menge für μ sind und so, dass $P \cup N = \Omega$.*

Beweis. Nehmen wir an μ nimmt den Wert $-\infty$ nicht an (sonst nimmt μ den Wert $+\infty$ nicht an, und man kann analog vorgehen). Sei

$$L = \inf \{\mu(A) : A \text{ eine negative Menge für } \mu \text{ ist}\}$$

Sei nun $(A_n)_{n\in\mathbb{N}}$ eine Folge in $\mathcal{A}$ von negativen Mengen für μ, mit $L = \lim_{n\to\infty} \mu(A_n)$. Sei $N = \cup_{n\in\mathbb{N}} A_n$. Dann ist N offenbar eine negative Menge für μ. Also $L \leq \mu(N) \leq \mu(A_n)$ für alle $n \in \mathbb{N}$. Es folgt, dass $L = \mu(N) > -\infty$ (weil μ den Wert $-\infty$ nicht annimmt). Sei nun $P = N^c$. Wir müssen noch zeigen, dass P eine positive Menge für μ ist. Nehmen wir indirekt an, dass $A \subset P$ in $\mathcal{A}$ existiert, mit $\mu(A) < 0$. Dann existiert aus Lemma 4.3 eine negative Menge $B \subset A$, mit $\mu(B) \leq \mu(A) < 0$. In diesem Fall wäre aber auch $B \cup N$ eine negative Menge für μ, mit

$$\mu(N \cup B) = \mu(N) + \mu(B) < \mu(N) = L$$

in Widerspruch zur Definition von L. $\square$

Bemerkung. Obwohl die Hanh'sche Zerlegung nicht verlangt, dass $P \cap N = \emptyset$, es gilt, nach Definition von positiven und negativen Mengen, dass $\mu(P \cap N) = 0$. Es folgt, dass die Hahn'sche Zerlegung bis auf Nullmengen eindeutig ist. Mit anderen Wörter, wenn P_1, N_1 und P_2, N_2 zwei Hahn'sche Zerlegungen eines signiertes Mass μ sind, dann muss $\mu(P_1 \cap N_2) = \mu(P_2 \cap N_1) = 0$ gelten.

Theorem 4.5 (Jordan'sche Zerlegungstheorem)**.** *Sei $(\Omega, \mathcal{A})$ ein messbarer Raum, und sei μ ein signiertes Mass auf $(\Omega, \mathcal{A})$. Dann existieren zwei (positive) Masse μ_+ und μ_- auf $(\Omega, \mathcal{A})$ so, dass $\mu(A) = \mu_+(A) - \mu_-(A)$ für alle $A \in \mathcal{A}$. Mindestens ein der zwei Masse μ_+, μ_- muss endlich sein.*

Beweis. Sei (P, N) eine Hahn'sche Zerlegung von μ. Wir definieren

$$\mu_+(A) = \mu(A \cap P), \qquad \text{und} \qquad \mu_-(A) = -\mu(A \cap N)$$

für alle $A \in \mathcal{A}$. Dann kann man leicht überprüfen, dass $\mu_\pm$ Masse auf $(\Omega, \mathcal{A})$ sind, mit $\mu(A) = \mu_+(A) - \mu_-(A)$ für alle $A \in \mathcal{A}$. Wenn μ_+ und μ_- beide nicht endlich wären, dann könnte μ beide Werten $\pm\infty$ annehmen; das ist aber nicht möglich. □

Sei μ ein signiertes Mass auf $(\Omega, \mathcal{A})$ und μ_+, μ_- die (positive) Masse, die wir im Beweis von Theorem 4.5 definiert haben. Sei $A \in \mathcal{A}$ beliebig. Für $B \subset A$, $B \in \mathcal{A}$ gilt dann

$$\mu(B) = \mu_+(B) - \mu_-(B) \leq \mu_+(B) \leq \mu_+(A)$$

Anderseits, nach Definition von μ_+, $\mu_+(A) = \mu(A \cap P)$, wobei $A \cap P$ eine messbare Teilmenge von A ist. Das bedeutet, dass

$$\mu_+(A) = \sup\{\mu(B) : B \subset A, B \in \mathcal{A}\} \tag{53}$$

Ähnlich,

$$\mu_-(A) = \sup\{-\mu(B) : B \subset A, B \in \mathcal{A}\} \tag{54}$$

Insbesondere, die Masse μ_+, μ_- hängen nicht von der besondere Wahl der Hahn'sche Zerlegung (P, N), die im Beweis von Theorem 4.5 getroffen wird. Die Masse μ_+, μ_-, definiert durch (53), (54) heissen den positiven und den negativen Teil von μ. Die Zerlegung $\mu = \mu_+ - \mu_-$ heisst die Jordan'sche Zerlegung von μ.

Man kann die Jordan'sche Zerlegung benutzten, um die Variation eines signiertes Mass μ zu definieren.

Definition 4.6. *Sei $(\Omega, \mathcal{A})$ ein messbarer Raum und μ ein signiertes Mass auf Ω. Sei $\mu = \mu_+ - \mu_-$ die Jordan'sche Zerlegung von μ. Die Variation von μ ist dann das (positive) Mass $|\mu| = \mu_+ + \mu_-$ auf $(\Omega, \mathcal{A})$.*

Bemerke, dass falls μ schon ein positives Mass ist, dann ist $\mu_+ = \mu$ und $\mu_- = 0$. Deswegen erhalten wir, in diesem Fall, dass $|\mu| = \mu$. Ferner, es gilt: $|\mu(A)| \leq |\mu|(A)$ für alle $A \in \mathcal{A}$. Übung: $|\mu|$ ist das kleinste Mass mit dieser Eigenschaft. Mit anderer Wörter: ist ν ein Mass auf $(\Omega, \mathcal{A})$, mit $|\mu(A)| \leq \nu(A)$ für alle ν in $\mathcal{A}$, dann gilt sicher $|\mu|(A) \leq \nu(A)$.

Definition 4.7. *Sei $(\Omega, \mathcal{A})$ ein messbarer Raum und μ ein signiertes Mass auf Ω. Wir definieren die totale Variation von μ durch $\|\mu\| := |\mu|(\Omega)$. Bemerke, dass $\|\mu\| < \infty$ genau dann, wenn μ endlich ist.*

Sei $M(\Omega, \mathcal{A})$ die Menge aller endlichen signierten Masse auf dem messbarer Raum $(\Omega, \mathcal{A})$. Es ist dann einfach zu überprüfen, dass $M(\Omega, \mathcal{A})$ ein Vektorraum über $\mathbb{R}$ ist, bezüglich die natürliche Definition von Summe und skalare Multiplikation (bemerke, dagegen, dass die Menge der positiven Masse auf $(\Omega, \mathcal{A})$ kein Vektorraum definiert, weil $-\mu$ kein positives Mass ist, wenn μ ein positives Mass ist). Es ist auch einfach zu sehen, dass die totale Variation $\|.\|$ eine Norm auf $M(\Omega, \mathcal{A})$ definiert; mit anderen Wörter, $0 \leq \|\mu\| < \infty$ für alle $\mu \in M(\Omega, \mathcal{A})$ und $\|\mu\| = 0$ genau dann wenn $\mu = 0$, $\|\lambda\mu\| = |\lambda|\|\mu\|$ für alle $\mu \in M(\Omega, \mathcal{A})$ und alle $\lambda \in \mathbb{R}$, $\|\mu_1 + \mu_2\| \leq \|\mu_1\| + \|\mu_2\|$ (bemerke: damit $\|.\|$ eine Norm ist, ist es wichtig, dass $M(\Omega, \mathcal{A})$ nur endliche Masse enthält, um den Fall $\|\mu\| = \infty$ zu vermeiden).

Proposition 4.8. *Das Vektorraum $M(\Omega, \mathcal{A})$, versehen mit der Norm $\|.\|$ der totale Variation, ist vollständig (d.h. $M(\Omega, \mathcal{A})$ definiert ein Banachraum).*

Beweis. Sei μ_n eine Cauchy Folge in $M(\Omega, \mathcal{A})$. Für ein belibiges $A \in \mathcal{A}$, es gilt: $|\mu_n(A) - \mu_m(A)| = |(\mu_n - \mu_m)(A)| \leq |\mu_n - \mu_m|(A) \leq |\mu_n - \mu_m|(\Omega) = \|\mu_n - \mu_m\| \to 0$, für $n, m \to \infty$. Das bedeutet, $\mu_n(A)$ ist eine Cauchy-Folge auf $\mathbb{R}$. Deswegen können wir $\mu(A) = \lim_{n\to\infty} \mu_n(A)$ definieren. Offenbar gilt: $\mu(\emptyset) = \lim_{n\to\infty} \mu_n(\emptyset) = 0$. Ferner, für $A, B \in \mathcal{A}$ mit $A \cap B = \emptyset$ haben wir: $\mu(A \cup B) = \lim_{n\to\infty} \mu_n(A \cup B) = \lim_{n\to\infty} \mu_n(A) + \mu_n(B) = \mu(A) + \mu(B)$. D.h. μ ist sicher endlich additiv.

Wir müssen noch zeigen, dass μ σ-additiv ist und, dass $\|\mu_n - \mu\| \to 0$, für $n \to \infty$. Dazu bemerken wir zunächst, dass die Konvergenz $\mu_n(A) \to \mu(A)$ gleichmässig in A ist, d.h., dass für alle $\varepsilon > 0$ ein $N \in \mathbb{N}$ existiert, mit $|\mu_n(A) - \mu(A)| < \varepsilon$ für alle $n > N$ und alle $A \in \mathcal{A}$. Um diese Behauptung zu zeigen, benutzten wir wieder die Tatsache, dass μ_n eine Cauchy-Folge ist. Das bedeutet, dass, für $\varepsilon > 0$ festgewählt, ein $N \in \mathbb{N}$ existiert so, dass $\|\mu_n - \mu_m\| < \varepsilon$ für alle $n, m > N$. Deswegen: $|\mu_n(A) - \mu_m(A)| < \varepsilon$ für alle $n, m > N$ und alle $A \in \mathcal{A}$. Insbesondere, im Limes $m \to \infty$ erhalten wir, dass $|\mu_n(A) - \mu(A)| \leq \varepsilon$, für alle $n > N$ und alle $A \in \mathcal{A}$, wie behauptet.

Um die σ-Additivität von μ zu zeigen, benutzen wir nun (52). Sei A_k eine monoton fallende Folge von Mengen in $\mathcal{A}$, mit $\cap_k A_k = \emptyset$. Wir müssen zeigen, dass $\mu(A_k) \to 0$, für $k \to \infty$ (wir wissen schon aus (51), dass $\mu_n(A_k) \to 0$ für $k \to \infty$, und für jede feste $n \in \mathbb{N}$; hier spielt die Tatsache, dass $M(\Omega, \mathcal{A})$ endliche Masse enthält eine wichtige Rolle). Wir wählen also $\varepsilon > 0$. Wir finden dann $N \in \mathbb{N}$ gross genug, damit $|\mu_N(A) - \mu(A)| < \varepsilon/2$ für alle $A \in \mathcal{A}$. Dann finden wir $K \in \mathbb{N}$, mit $|\mu_N(A_k)| \leq \varepsilon/2$ für alle $k > K$ (möglich, weil $|\mu_N(A_k)| \to 0$, für $k \to \infty$). Deswegen, es gilt: $|\mu(A_k)| \leq |\mu_N(A_k)| + |\mu_N(A_k) - \mu(A_k)| \leq \varepsilon$ für alle $k > K$. D.h. $|\mu(A_k)| \to 0$, für $k \to \infty$.

Schlussendlich behaupten wir, dass $\|\mu_n - \mu\| = |\mu_n - \mu|(\Omega) \to 0$. Wir müssen zeigen, dass $(\mu_n - \mu)_+(\Omega) \to 0$, und, dass $(\mu_n - \mu)_-(\Omega) \to 0$, für $n \to \infty$. Gemäss (53) und (54), genügt es zu zeigen, dass $\sup_{A\in\mathcal{A}} |\mu_n(A) - \mu(A)| \to 0$, als $n \to \infty$. Das ist aber genau die Gleichmässigkeit (in A) der Konvergenz $\mu_n(A) \to \mu(A)$, die wir oben bewiesen haben. $\square$

Bemerkung. Ähnlich wie mit signierte Masse kann man auch komplexwertige Masse einführen. Ein komplexes Mass μ auf einem messbarer Raum $(\Omega, \mathcal{A})$ ist eine Abbildung $\mu : \mathcal{A} \to \mathbb{C}$, mit $\mu(\emptyset) = 0$ so, dass $\mu(\cup_n A_n) = \sum_n \mu(A_n)$ für alle Folge A_n von disjunkten Mengen in $\mathcal{A}$. Man kann dann leicht sehen, dass jedes komplexes Mass μ als $\mu = \mu_1 + i\mu_2$ geschrieben werden kann, wobei μ_1, μ_2 zwei endliche signierten Masse auf $(\Omega, \mathcal{A})$ sind (komplexe Masse sind immer endlich, im Gegensatz zu signierten Masse). Nach Jordan'sche Zerlegung, können wir dann μ als $\mu = \mu_{1,+} - \mu_{1,-} + i\mu_{2,+} - i\mu_{2,-}$ schreiben, wobei $\mu_{1,+}, \mu_{1,-}, \mu_{2,+}, \mu_{2,-}$ (positive) Masse auf $(\Omega, \mathcal{A})$ definieren. Ist μ ein komplexes Mass auf $(\Omega, \mathcal{A})$, so definieren wir die entsprechende Variation durch

$$|\mu|(A) = \sup \left\{ \sum_{j=1}^{k} |\mu(A_j)| : A_j \in \mathcal{A}, A = \bigcup_{j=1}^{k} A_j \right\}$$

Man kann dann zeigen, dass $|\mu|$ ein positives Mass auf $(\Omega, \mathcal{A})$ ist. Ähnlich wie für signierte Masse gilt auch: die totale Variation $\|\mu\| = |\mu|(\Omega)$ definiert eine Norm auf dem

Vektorraum $M(\Omega, \mathcal{A}, \mathbb{C})$ der aus allen komplexen Masse auf $(\Omega, \mathcal{A})$ besteht. Bezüglich dieser Norm ist $M(\Omega, \mathcal{A}, \mathbb{C})$ vollständig.

Bemerkung. Man kann Integrale bezüglich signierte (oder komplexe) Masse definieren. Ist $(\Omega, \mathcal{A})$ ein messbarer Raum, f eine beschränkte messbare Funktion auf $(\Omega, \mathcal{A})$ und μ ein signiertes Mass auf $(\Omega, \mathcal{A})$, so können wir

$$\int f d\mu = \int f d\mu_+ - \int f d\mu_- \tag{55}$$

definieren, indem wir die Jordan'sche Zerlegung $\mu = \mu_+ - \mu_-$ von μ benutzten (man schränkt sich besser auf beschränkte Funktionen, damit wir sicher sind, dass die Differenz in (55) wohldefiniert ist).

4.2 Absolut stetige Masse

Wir führen nun den Begriff von absolute Stetigkeit ein. Zunächst betrachten wir hier den Fall von positiven Massen.

Definition 4.9. *Sei $(\Omega, \mathcal{A})$ ein messbarer Raum und μ und ν zwei (positive) Masse auf $(\Omega, \mathcal{A})$. Wir sagen, dass ν bezüglich μ absolut stetig ist, wenn*

$$A \in \mathcal{A},\ \mu(A) = 0 \quad \Rightarrow \quad \nu(A) = 0$$

Wir schreiben in diesem Fall $\nu \ll \mu$. Ein Mass ν auf $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ heisst einfach absolut stetig, wenn es bezüglich dem Lebesgue Mass λ_n absolut stetig ist.

Die folgende Charakterisierung für absolute Stetigkeit wird nützlich sein.

Lemma 4.10. *Sei $(\Omega, \mathcal{A})$ ein messbarer Raum, μ ein positives Mass auf $(\Omega, \mathcal{A})$ und ν ein endliches positives Mass auf $(\Omega, \mathcal{A})$. Dann gilt $\nu \ll \mu$ genau dann wenn für alle $\varepsilon > 0$ ein $\delta > 0$ existiert so, dass*

$$A \in \mathcal{A},\ \mu(A) < \delta \qquad \Rightarrow \quad \nu(A) < \varepsilon \tag{56}$$

Beweis. Nehmen wir zunächst an, dass die Bedingung (56) erfüllt ist. Wir möchten zeigen, dass $\nu \ll \mu$. Sei dazu $A \in \mathcal{A}$, mit $\mu(A) = 0$. Da $\mu(A) < \delta$ für alle $\delta > 0$, gilt also $\nu(A) < \varepsilon$ für alle $\varepsilon > 0$. Deswegen muss $\nu(A) = 0$. Anderseits, sei $\nu \ll \mu$. Nehmen wir indirekt an, dass (56) nicht erfüllt ist. Dann finden wir $\varepsilon > 0$ und, für jede $k \in \mathbb{N}$, ein $A_k \in \mathcal{A}$ mit $\mu(A_k) < 2^{-k}$ und $\nu(A_k) > \varepsilon$. Es folgt, dass, für ein beliebiges $n \in \mathbb{N}$, $\mu(\bigcup_{k\geq n} A_k) \leq \sum_{k\geq n} 2^{-k} = 2^{-n+1}$ während $\nu(\bigcup_{k\geq n} A_k) \geq \nu(A_n) > \varepsilon$. Sei nun $A = \bigcap_{n\geq 1} \bigcup_{k\geq n} A_k$. Dann gilt $A \subset \bigcup_{k\geq n} A_k$ und deswegen $\mu(A) \leq 2^{-n+1}$ für alle $n \in \mathbb{N}$. Das impliziert, dass $\mu(A) = 0$. Anderseits, da die Folge $\bigcup_{k\geq n} A_k$ fallend ist, muss $\nu(A) = \lim_{n\to\infty} \nu(\bigcup_{k\geq n} A_k) \geq \varepsilon$ (weil ν endlich ist). Das widerspricht die Annahme, dass $\nu \ll \mu$. $\square$

Beispiel. Sei $(\Omega, \mathcal{A}, \mu)$ ein Massraum und $f \in \mathcal{L}^1(\Omega, \mathcal{A}, \mu)$ nicht-negativ (d.h. $f(x) \geq 0$ für alle $x \in \Omega$). Durch

$$\nu(A) = \int_A f d\mu$$

definieren wir ein neues positives Mass ν auf $(\Omega, \mathcal{A})$. Ist nun $A \in \mathcal{A}$ mit $\mu(A) = 0$, dann ist $f\chi_A = 0$ μ-fast überall (χ_A ist die charakteristische Funktion von A). Deswegen gilt auch $\nu(A) = 0$. Das bedeutet, $\nu \ll \mu$. Wir zeigen im nächsten Theorem, dass jedes Mass ν mit $\nu \ll \mu$ die Form $\nu(A) = \int_A f d\mu$ hat (mindestens, wenn ν, μ σ-endlich sind).

Theorem 4.11 (Radon-Nikodym Theorem)**.** *Sei $(\Omega, \mathcal{A})$ ein messbarer Raum und seien μ, ν zwei σ-endliche Masse auf $(\Omega, \mathcal{A})$. Ist $\nu \ll \mu$, so existiert eine Funktion $g : \Omega \to [0; \infty)$ messbar, mit*

$$\nu(A) = \int_A g \, d\mu$$

für alle $A \in \mathcal{A}$. Die Funktion g ist eindeutig, bis auf Gleichheit auf einer μ-Nullmenge.

Beweis. Wir betrachten zunächst den Fall, dass μ, ν endliche Masse sind. Sei

$$\mathcal{F} = \left\{ f : \Omega \to [0; \infty] \text{ messbar, s.d. } \int_A f d\mu \leq \nu(A) \quad \text{für alle } A \in \mathcal{A} \right\}$$

Dann ist $\mathcal{F}$ nicht leer, weil die Konstante Funktion $f = 0$ sicher in $\mathcal{F}$ gehört.

Schritt 1. Es existiert $g \in \mathcal{F}$, mit

$$\int g \, d\mu = \sup \left\{ \int f \, d\mu : f \in \mathcal{F} \right\}$$

Um die Behauptung zu zeigen, bemerken wir zunächst, dass

$$f_1, f_2 \in \mathcal{F} \quad \Rightarrow \quad \max\{f_1, f_2\} \in \mathcal{F}$$

In der Tat, für ein beliebiges $A \in \mathcal{A}$, wir können $A_1 = \{x \in A : f_1(x) > f_2(x)\}$ und $A_2 = \{x \in A : f_1(x) \leq f_2(x)\}$ definieren. Dann gilt

$$\int_A \max\{f_1, f_2\} \, d\mu = \int_{A_1} f_1 \, d\mu + \int_{A_2} f_2 \, d\mu \leq \nu(A_1) + \nu(A_2) = \nu(A_1 \cup A_2) = \nu(A)$$

Sei nun $(f_n)_{n \in \mathbb{N}}$ eine Folge in $\mathcal{F}$ mit

$$\lim_{n \to \infty} \int f_n \, d\mu = \sup \left\{ \int f \, d\mu : f \in \mathcal{F} \right\}$$

Dann gilt offenbar auch

$$\lim_{n \to \infty} \int \max(f_1, \ldots, f_n) \, d\mu = \sup \left\{ \int f \, d\mu : f \in \mathcal{F} \right\}$$

Da die Folge $\max(f_1, \ldots, f_n)$ monoton wachsend ist, können wir den Limes

$$g = \lim_{n \to \infty} \max(f_1, \ldots, f_n)$$

punktweise definieren. Das Theorem der monotone Konvergenz zeigt, dass

$$\int_A g \, d\mu = \lim_{n \to \infty} \int_A \max(f_1, \ldots, f_n) \, d\mu \leq \nu(A)$$

weil $\max(f_1, \ldots, f_n) \in \mathcal{F}$ für alle $n \in \mathbb{N}$ und auch, dass

$$\int g \, d\mu = \lim_{n \to \infty} \int \max(f_1, \ldots, f_n) \, d\mu = \sup \left\{ \int f \, d\mu : f \in \mathcal{F} \right\}$$

Schritt 2. Sei g wie im Schritt 1. Wir behaupten, dass

$$\nu(A) = \int_A g \, d\mu$$

für alle $A \in \mathcal{A}$.

Für ein beliebiges $A \in \mathcal{A}$, wir setzen

$$\nu_0(A) = \nu(A) - \int_A g \, d\mu$$

Da $g \in \mathcal{F}$, ist $\nu_0 : \mathcal{A} \to [0; \infty]$ ein Mass auf $\mathcal{A}$. Zu zeigen bleibt, dass $\nu_0 = 0$. Nehmen wir an $\nu_0 \neq 0$. Dann finden wir $\varepsilon > 0$ mit $\nu_0(\Omega) > \varepsilon \mu(\Omega)$ (weil μ ist endlich, nach Annahme). Es ist einfach zu überprüfen, dass $\nu_0 - \varepsilon \mu : \mathcal{A} \to [-\infty; \infty]$ ein signiertes Mass auf $\mathcal{A}$ ist. Sei (P, N) eine Hahn'sche Zerlegung von $\nu_0 - \varepsilon \mu$.

Also, $P, N \in \mathcal{A}$ mit $P \cup N = \Omega$ und so, dass P eine positive und N eine negative Menge für $\nu_0 - \varepsilon \mu$ sind. Wir bemerken sofort, dass $\mu(P) > 0$ sein muss. Wäre nämlich $\mu(P) = 0$, dann müsste auch $\nu(P) = 0$ (aus der absolute Stetigkeit $\nu \ll \mu$) und also $\nu_0(P) = 0$. Deswegen wäre

$$\nu_0(\Omega) - \varepsilon \mu(\Omega) = (\nu_0 - \varepsilon \mu)(N) \leq 0$$

in Widerspruch mit der Annahme, dass $\nu_0(\Omega) > \varepsilon \mu(\Omega)$. Das zeigt, dass $\mu(P) > 0$. Ferner, für ein beliebiges $A \in \mathcal{A}$ gilt $\nu_0(A \cap P) \geq \varepsilon \mu(A \cap P)$. Deswegen

$$\nu(A) = \int_A g \, d\mu + \nu_0(A) \geq \int_A g \, d\mu + \nu_0(A \cap P) \geq \int_A g \, d\mu + \varepsilon \mu(A \cap P) = \int_A (g + \varepsilon \mathbf{1}_P) \, d\mu$$

Das zeigt, dass $g + \varepsilon \mathbf{1}_P \in \mathcal{F}$. Da aber

$$\int g \, d\mu \leq \nu(\Omega) < \infty$$

und

$$\int (g + \varepsilon \mathbf{1}_P) \, d\mu = \int g \, d\mu + \varepsilon \mu(P) > \int g \, d\mu$$

finden wir ein Widerspruch zur Definition von g in Schritt 1. Das zeigt, dass $\nu_0 = 0$ und also, dass

$$\nu(A) = \int_A g \, d\mu$$

für alle $A \in \mathcal{A}$.

Schritt 1 und 2 zeigen das Theorem im Fall, dass ν, μ endlich sind. In allgemein finden wir eine Folge $(A_n)_{n \in \mathbb{N}}$ von disjunkten Mengen in $\mathcal{A}$ mit $\cup_n A_n = \Omega$ und so, dass $\mu(A_n)$ und $\nu(A_n)$ endlich sind, für alle $n \in \mathbb{N}$. Für alle $n \in \mathbb{N}$ finden wir dann eine Funktion $g_n : A_n \to [0; \infty]$ mit

$$\nu(E) = \int_E g_n d\mu$$

für alle $E \in \mathcal{A}$ mit $E \subset A_n$. Wir können $g : \Omega \to [0;\infty]$ durch $g = g_n$ auf A_n, für alle $n \in \mathbb{N}$, definieren. Es ist dann einfach zu überprüfen, dass g die gewünschte Eigenschaften hat.

Schlussendlich zeigen wir die Eindeutigkeit der Funktion g. Wir betrachten zunächst den Fall, dass ν endlich ist. Nehmen wir an $g, h : \Omega \to [0;\infty]$ sind beide $\mathcal{A}$-messbar und so, dass

$$\nu(A) = \int_A g\, d\mu = \int_A h d\mu$$

für alle $A \in \mathcal{A}$. Da ν endlich ist, es folgt, dass

$$\int_A (g - h)\, d\mu = 0$$

für alle $A \in \mathcal{A}$. Wenn wir zunächst $A = \{x \in \Omega : g(x) \geq h(x)\}$ und dann $A = \{x \in \Omega : h(x) > g(x)\}$ wählen, finden wir, dass

$$\int |g - h| d\mu = 0$$

und damit, dass $g = h$ μ-fast überall. Wenn ν σ-endlich aber nicht unbedingt endlich ist, finden wir eine Folge von disjukten Mengen $(A_n)_{n\in\mathbb{N}}$ in $\mathcal{A}$ mit $\Omega = \cup_n A_n$ und so, dass $\nu(A_n) < \infty$ für alle $n \in \mathbb{N}$. Wie oben können wir dann zeigen, dass $g = h$ μ-f.ü. auf A_n, für alle $n \in \mathbb{N}$. Das zeigt auch, dass $g = h$ μ-f.ü. auf Ω. □

Bemerkung. Die Annnahme, dass μ σ-endlich ist, ist wichtig. In der Tat: sei $\Omega = [0;1]$, $\mathcal{A}$ die Borel σ-Algebra auf $[0;1]$, μ das Mass auf Ω, definiert durch $\mu(A) = |A|$ (die Kardinalität von A, wenn A eine endliche Menge ist, und ∞ sonst) und ν das Lebesgue Mass auf Ω. Dann gilt $\nu \ll \mu$ (weil $\mu(A) = 0$ impliziert, dass $A = \emptyset$), aber es existiert keine messbare Funktion $f : \Omega \to [0;\infty)$, mit $\nu(A) = \int_A f d\mu$.

Man kann den Begriff von absolute Stetigkeit auf signierten Masse erweitern. Sei $(\Omega, \mathcal{A})$ ein messbarer Raum, und μ ein positives Mass auf $(\Omega, \mathcal{A})$. Ein signiertes Mass ν auf $(\Omega, \mathcal{A})$ heisst absolut stetig bezüglich μ, wenn seine Variation $|\nu|$ absolut stetig bezüglich μ ist. In diesem Fall schreiben wir (wie im Fall eines positiven Mass ν) $\nu \ll \mu$ (d.h. wir definieren, dass $\nu \ll \mu$ genau dann wenn $|\nu| \ll \mu$). Übung: $\nu \ll \mu$ genau dann wenn $\nu_+, \nu_- \ll \mu$.

Auch für signierte absolut stetige Masse gilt eine Version vom Theorem von Radon-Nikodym. Der Beweis folgt einfach aus Theorem 4.11, mit Hilfe der Zerlegung $\nu = \nu_+ - \nu_-$.

Theorem 4.12 (Radon-Nikodym Theorem - signiertes Mass)**.** *Sei $(\Omega, \mathcal{A})$ ein messbarer Raum und μ ein σ-endliches Masse auf $(\Omega, \mathcal{A})$. Sei $\nu \ll \mu$ ein endliches signiertes Mass auf $(\Omega, \mathcal{A})$, absolut stetig bezüglich μ. Dann existiert eine $\mathbb{R}$-wertige Funktion $g \in \mathcal{L}^1(\Omega, \mathcal{A}, \mu)$ mit*

$$\nu(A) = \int_A g d\mu$$

für alle $A \in \mathcal{A}$. Die Funktion g ist eindeutig, bis auf Gleichheit auf einer μ-Nullmenge.

Bemerkung. Sei $(\Omega, \mathcal{A})$ ein messbarer Raum, μ ein σ-endliches Mass auf $(\Omega, \mathcal{A})$, und ν ein σ-endliches positives Mass oder ein endliches signiertes Mass auf $(\Omega, \mathcal{A})$, mit $\nu \ll \mu$. Eine messbare Funktion $g : \Omega \to \mathbb{R}$ mit der Eigenschaft, dass $\nu(A) = \int_A g d\mu$ für alle $A \in \mathcal{A}$, heisst eine Radon-Nikodym Ableitung von ν bezüglich μ. Man schreibt manchmal $g = d\nu/d\mu$. Es folgt aus Theorem 4.11 und aus Theorem 4.12, dass Radon-Nikodym Ableitungen immer existieren und, dass sie bis auf μ-fast überall Gleichheit eindeutig sind.

Eine wichtige Anwendung vom Theorem von Radon-Nikodym ist ein Beweis der Surjektivität der Abbildung $\phi : L^q(\Omega, \mathcal{A}, \mu) \to (L^p(\Omega, \mathcal{A}, \mu))^*$, die wir in (38) definiert haben, für $1 \leq p < \infty$ (im Fall $p = 1$, unter der zusätzliche Annahme, dass μ σ-endlich ist). Wie wir unter Theorem 3.12 bemerkt haben, die Surjektivität von ϕ erlaubt uns der Dualraum von $L^p(\Omega, \mathcal{A}, \mu)$ (d.h. der Raum von allen beschränkten linearen Funktionalen auf $L^p(\Omega, \mathcal{A}, \mu)$) mit $L^q(\Omega, \mathcal{A}, \mu)$ zu identifizieren, falls $1/p + 1/q = 1$ und $1 \leq p < \infty$.

Theorem 4.13. *Sei $(\Omega, \mathcal{A}, \mu)$ ein Massraum, $1 < p, q < \infty$ so, dass $1/p + 1/q = 1$. Die Abbildung $\phi : L^q(\Omega, \mathcal{A}, \mu) \to (L^p(\Omega, \mathcal{A}, \mu))^*$ definiert wie in (38) durch $\phi(f) = \phi_f$, und $\phi_f(g) = \int fg\, d\mu$, für alle $f \in L^q(\Omega, \mathcal{A}, \mu)$ und alle $g \in L^p(\Omega, \mathcal{A}, \mu)$, ist surjektiv. Mit Theorem 3.12 erhalten wir: ϕ definiert ein lineares isometrisches Isomorphismus zwischen $L^q(\Omega, \mathcal{A}, \mu)$ und $(L^p(\Omega, \mathcal{A}, \mu))^*$, für alle $1 < p, q < \infty$, mit $1/p + 1/q = 1$. Die selbe Aussagen gelten auch für $p = 1$ und $q = \infty$, wenn wir zusätzlich annehmen, dass μ ein σ-endlich Mass ist.*

Beweis. Wir nehmen hier an, dass μ endlich ist, und $1 \leq p < \infty$. Die Verallgemeinerung zu μ beliebig (falls $1 < p < \infty$) oder μ σ-endlich (falls $p = 1$) lassen wir als Übung (oder zur Vorlesung Funktional Analysis).

Sei $\psi \in (L^p(\Omega, \mathcal{A}, \mu))^*$ beliebig. Wir suchen $g \in L^q(\Omega, \mathcal{A}, \mu)$ mit $\psi = \phi(g)$. Wir definieren dazu ein signiertes Mass ν auf $\mathcal{A}$ durch $\nu(A) = \psi(\chi_A)$ für alle $A \in \mathcal{A}$. Hier bezeichnet χ_A die charakteristische Funktion von der Menge A (es gilt $\chi_A \in L^p(\Omega, \mathcal{A}, \mu)$ für alle $A \in \mathcal{A}$, weil wir angenommen haben, dass μ endlich ist). Es gilt offenbar $\nu(\emptyset) = 0$. Um zu zeigen, dass ν ein signiertes Mass ist, müssen wir noch die σ-Additivität beweisen. Sei also $(A_k)_{k \in \mathbb{N}}$ eine Folge disjunkte Menge in $\mathcal{A}$, und $A = \cup_{k \in \mathbb{N}} A_k$. Dann gilt $\sum_{k=1}^n \chi_{A_k} \to \chi_A$ punktweise, als $n \to \infty$. Da ferner $|\chi_A(x) - \sum_{k=1}^n \chi_{A_j}(x)| \leq \chi_A(x)$ impliziert dominierte Konvergenz, dass $\|\chi_A - \sum_{k=1}^n \chi_{A_j}\|_p \to 0$, für $n \to \infty$. Die Stetigkeit und die Linearität von ψ implizieren also, dass $\psi(\chi_A) = \lim_{n \to \infty} \psi(\sum_{k=1}^n \chi_{A_j}) = \lim_{n \to \infty} \sum_{k=1}^n \psi(\chi_{A_k})$. D.h. $\nu(A) = \lim_{n \to \infty} \sum_{k=1}^n \nu(A_k) = \sum_{k=1}^\infty \nu(A_k)$. Das zeigt, dass ν ein signiertes Mass ist.

Offenbar ist $\nu \ll \mu$. Ist in der Tat $A \in \mathcal{A}$ mit $\mu(A)$, so ist $\chi_A(x) = 0$ für μ-fast alle $x \in \Omega$. Das bedeutet, dass $\chi_A = 0$ im Raum $L^p(\Omega, \mathcal{A}, \mu)$ (besser gesagt, χ_A ist in der selbe Äquivalenzklasse wie die Null-Funktion). Aus Theorem 4.12 folgt, dass eine Radon-Nikodym Ableitung $g \in \mathcal{L}^1(\Omega, \mathcal{A}, \mu)$ existiert, mit $\nu(A) = \int_A g\, d\mu$, für alle $A \in \mathcal{A}$. Das bedeutet, dass

$$\psi(\chi_A) = \int \chi_A\, g\, d\mu, \tag{57}$$

für alle $A \in \mathcal{A}$. Wir möchten nun zeigen, dass $g \in L^q(\Omega, \mathcal{A}, \mu)$ und, dass $\psi(f) = \int fg\, d\mu$, für alle $f \in L^p(\Omega, \mathcal{A}, \mu)$ (dann wurde $\psi = \phi(g)$ gelten).

Wir betrachten dazu die Mengen $E_n = \{x \in \Omega : |g(x)| < n\}$. Dann ist $g\chi_{E_n}$ beschränkt, und also $g\chi_{E_n} \in L^q(\Omega, \mathcal{A}, \mu)$ (weil wir angenommen haben, dass μ endlich ist). Wir definieren die Funktionale $\psi_n \in (L^p(\Omega, \mathcal{A}, \mu))^*$ durch $\psi_n(f) = \psi(f\chi_{E_n})$. Wir behaupten, dass

$$\psi_n(f) = \int f g \chi_{E_n} \, d\mu \tag{58}$$

für alle $f \in L^p(\Omega, \mathcal{A}, \mu)$ und alle $n \in \mathbb{N}$. Ist $f = \chi_A$, für ein $A \in \mathcal{A}$, dann gilt $f\chi_{E_n} = \chi_{A \cap E_n}$ und (58) folgt aus (57). Damit gilt (58) sicher für einfache Funktionen in $L^p(\Omega, \mathcal{A}, \mu)$. Da aber einfache Funktionen dicht sind, gilt (58) auch für beliebige $f \in L^p(\Omega, \mathcal{A}, \mu)$.

Aus (58) erhalten wir, dass $\psi_n = \phi(g\chi_{E_n})$. Da wir schon in Theorem 3.12 bewiesen haben, dass ϕ eine Isometrie ist, finden wir

$$\|g\chi_{E_n}\|_q = \|\psi_n\|_{(L^p)^*} \leq \|\psi\|_{(L^p)^*}, \tag{59}$$

weil

$$\|\psi_n\|_{(L^p)^*} = \sup_{\|f\|_p \leq 1} |\psi_n(f)| = \sup_{\|f\|_p \leq 1} |\psi(f\chi_{E_n})| \leq \|\psi\|_{(L^p)^*}$$

da $\|f\chi_{E_n}\|_p \leq \|f\|_p$. Ist $q < \infty$ (d.h. $p > 1$), dann impliziert (59), zusammen mit monotone Konvergenz, dass $g \in L^q(\Omega, \mathcal{A}, \mu)$, mit $\|g\|_q \leq \|\psi\|_{(L^p)^*}$. Für $q = \infty$ kann monotone Konvergenz nicht angewandt werden. Dagegen benutzen wir in diesem Fall, dass $g \in \mathcal{L}^1(\Omega, \mathcal{A}, \mu)$, um zu schliessen, dass $|g(x)| < \infty$ für μ-fast alle $x \in \Omega$ und deswegen, dass $\mu(\Omega \setminus \cup_{n \in \mathbb{N}} E_n) = \mu(\{x \in \Omega : |g(x)| = \infty\}) = 0$. Das impliziert, mit (59) mit $q = \infty$, dass

$$\mu(\{x \in \Omega : |g(x)| > \|\psi\|_{(L^p)^*}\}) = \lim_{n \to \infty} \mu(\{x \in E_n : |g(x)| > \|\psi\|_{(L^p)^*}\}) = 0$$

Auch für $q = \infty$ erhalten wir also, dass $\|g\|_q \leq \|\psi\|_{(L^p)^*}$. Das erlaubt uns den Limes $n \to \infty$ in (58) zu nehmen. Auf der linke Seite benutzen wir, für ein beliebiges $f \in L^p(\Omega, \mathcal{A}, \mu)$, dass

$$|\psi_n(f) - \psi(f)| = |\psi(f\chi_{E_n} - f)| \leq \|\psi\|_{(L^p)^*} \|f\chi_{E_n} - f\|_p \to 0$$

für $n \to \infty$ (aus dominierte Konvergenz, weil $f\chi_{E_n} - f \to 0$ punktweise, und $|f\chi_n - f| \leq |f|$). Auf der rechten Seite benutzen wir wieder dominierte Konvergenz, und die Tatsache, dass $|fg|$ integrierbar ist (aus Hölder Ungleichung, weil wir jetzt wissen, dass $g \in L^q(\Omega, \mathcal{A}, \mu)$). Wir erhalten, dass

$$\psi(f) = \int f g \, d\mu$$

für alle $f \in L^p(\Omega, \mathcal{A}, \mu)$. Das zeigt, dass $\psi = \phi(g)$, und damit, dass ϕ surjektiv ist, im Fall μ endlich. $\square$

4.3 Singuläre Masse

Neben dem Begriff von absolut stetige Masse spielt auch den Begriff von singulären Masse eine wichtige Rolle.

Definition 4.14. *Sei $(\Omega, \mathcal{A})$ ein messbarer Raum. Wir sagen ein positives Mass μ auf $(\Omega, \mathcal{A})$ ist auf einer Menge $E \in \mathcal{A}$ getragen, falls $\mu(E^c) = 0$. Wir sagen ein signiertes Mass μ auf $(\Omega, \mathcal{A})$ ist auf $E \in \mathcal{A}$ getragen, falls $|\mu|(E^c) = 0$ (d.h. wenn $\mu(A) = 0$ für alle $A \in \mathcal{A}$ mit $A \subset E^c$).*

Wir können nun den Begriff von singulären Mass einführen.

Definition 4.15. *Sei $(\Omega, \mathcal{A})$ ein messbarer Raum. Wir sagen zwei (positive oder signierte) Masse μ, ν sind singulär (oder, dass μ ist singulär bezüglich ν oder, dass ν ist singulär bezüglich μ), wenn $E \in \mathcal{A}$ existiert so, dass μ auf E und ν auf E^c getragen sind. Wir schreiben, in diesem Fall, $\mu \perp \nu$. Ein (positives oder signiertes) Mass μ auf $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ heisst einfach singulär, wenn es bezüglich Lebesgue Mass λ_n singulär ist.*

Beispiel. Sei μ ein signiertes Mass auf $(\Omega, \mathcal{A})$ und $\mu = \mu_+ - \mu_-$ die entsprechende Jordan Zerleung. Dann gilt immer $\mu_+ \perp \mu_-$. Das folgt, weil μ_+ und μ_- auf P, beziehungsweise, auf N getragen sind, wobei $\Omega = P \cup N$ die Hahn'sche Zerlegung von μ ist (man kann P und N disjunkt wählen; wenn $P \cap N \neq \emptyset$, kann man z.B. P durch $P \cap N^c$ ersetzen, weil $\mu(P \cap N) = 0$).

Beispiel. Sei $\Omega = \mathbb{R}^n$ und $\mathcal{A} = \mathcal{B}(\mathbb{R}^n)$ die Borel σ-Algebra. Für $x \in \mathbb{R}^n$ beliebig können wir das Mass δ_x durch $\delta_x(A) = 1$, falls $x \in A$, und $\delta_x(A) = 0$ falls $x \notin A$, definieren. Dann ist δ_x auf der Menge $\{x\}$ getragen. Deswegen ist δ_x singulär bezüglich dem Lebesgue Mass. Jedes diskretes Mass, definiert als eine endliche (oder abzählbare) lineare Kombination von Punktmasse der Form δ_x ist singulär (weil jede solches Mass auf abzählbare Mengen getragen ist, und weil das Lebesgue Mass jeder abzählbare Menge verschwindet). Nicht jedes bezüglich λ_n singuläres Mass auf $\mathbb{R}^n$ ist aber diskret. In (20) haben wir die Cantor-Funktion f definiert. Wir haben dort gezeigt, dass $f(x) = 0$ für alle $x \leq 0$ und $f(x) = 1$ für alle $x \geq 1$, dass f monoton steigend ist, dass f stetig ist und auch, dass für alle $x \in (0;1)\backslash K$ eine offene Umgebung von x existiert, wo f konstant ist ($K \subset (0;1)$ ist hier die Cantor Menge, definiert in (18)). In den Übungen haben wir gezeigt, dass jede beschränkte, monoton wachsende, rechtsstetige Funktion, die im Limes $x \to -\infty$ verschwindet, ein Lebesgue-Stieltjes Mass auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ definiert. Aus den Eigenschaften von f folgt dann (Beweis: Übung), dass μ_f auf der Cantor Menge K getragen ist. Da $\lambda(K) = 0$, es folgt, dass μ_f ist singulär bezüglich das Lebesgue Mass λ. Anderseits, $\mu_f(\{x\}) = 0$ für alle $x \in \mathbb{R}$; also μ_f ist nicht diskret.

Theorem 4.16 (Lebesgue Zerlegung)**.** *Sei $(\Omega, \mathcal{A})$ ein messbarer Raum. Sei μ ein positives Mass auf $(\Omega, \mathcal{A})$ und sei ν ein σ-endliches positives Mass oder ein endliches signiertes Mass auf $(\Omega, \mathcal{A})$. Dann es existieren eindeutige positive oder endlich signierte Masse ν_a, ν_s auf $(\Omega, \mathcal{A})$ mit $\nu_a \ll \mu$, $\nu_s \perp \mu$, $\nu = \nu_a + \nu_s$.*

Beweis. Wir betrachten zunächst den Fall, dass ν ein positives endliches Mass ist. Wir definieren dann

$$\mathcal{N}_\mu = \{B \in \mathcal{A} : \mu(B) = 0\}$$

Wir wählen eine Folge $(B_j)_{j \in \mathbb{N}}$ in $\mathcal{N}_\mu$ so, dass

$$\lim_{j \to \infty} \nu(B_j) = \sup\{\nu(B) : B \in \mathcal{N}_\mu\}$$

Sei $N = \cup_{j\in\mathbb{N}} B_j$. Wir bemerken, dass

$$\mu(N) \leq \sum_j \mu(B_j) = 0$$

Deswegen gilt $N \in \mathcal{N}_\mu$ und

$$\nu(N) = \sup\{\nu(B) : B \in \mathcal{N}_\mu\}\,.$$

Wir definieren ν_a und ν_s durch

$$\nu_a(A) = \nu(A \cap N^c), \qquad \nu_s(A) = \nu(A \cap N)$$

Offenbar $\nu = \nu_a + \nu_s$. Da $\mu(N) = 0$ ist $\mu \perp \nu$. Ist nun $A \in \mathcal{A}$ mit $\mu(A) = 0$, dann ist $A \in \mathcal{N}_\mu$ und also $N \cup (A \cap N^c) \in \mathcal{N}_\mu$. Deswegen gilt

$$\nu(N) + \nu(A \cap N^c) = \nu(N \cup (A \cap N^c)) \leq \sup\{\nu(B) : B \in \mathcal{N}_\mu\} = \nu(N)$$

was

$$\nu_a(A) = \nu(A \cap N^c) = 0$$

impliziert. Das zeigt, dass $\nu_a \ll \mu$.

Ist nun ν ein signiertes Mass, so können wir N_μ wie oben definieren, und eine Folge $(B_j)_{j\in\mathbb{N}}$ finden, mit

$$\lim_{j\to\infty} |\nu|(B_j) = \sup\{|\nu|(B) : B \in \mathcal{N}_\mu\}$$

Mit $N = \cup_{j\in\mathbb{N}} B_j$ folgt nun, dass $N \in \mathcal{N}_\mu$, und, dass

$$|\nu|(N) = \sup\{|\nu|(B) : B \in \mathcal{N}_\mu\}\,.$$

Wir können wieder

$$\nu_a(A) = \nu(A \cap N^c), \qquad \nu_s(A) = \nu(A \cap N)$$

definieren. Dann gilt offenbar $\nu = \nu_a + \nu_s$. Da $\mu(N) = 0$ folgt wieder, dass $\nu_s \perp \mu$. Ist weiter $A \in \mathcal{A}$ mit $\mu(A) = 0$, so muss ähnlich wie oben $|\nu|(A \cap N^c) = 0$. Das impliziert aber auch, dass $\nu_a(A) = 0$, und damit, dass $\nu_a \ll \mu$.

Ist dagegen ν ein positives σ-endliches Mass, wir finden eine Familie von Mengen $\{D_j\}_{j\in\mathbb{N}}$ mit $D_i \cap D_j = \emptyset$ für alle $i \neq j$, mit $\cup_{j\in\mathbb{N}} D_j = \Omega$ und mit $\nu(D_j) < \infty$ für alle $j \in \mathbb{N}$. Wie oben, können wir dann für alle $j \in \mathbb{N}$ eine Menge $N_j \subset D_j$ mit $N_j \in \mathcal{A}$, $\mu(N_j) = 0$ und mit

$$\nu(N_j) = \sup\{\nu(B) : B \subset D_j, \mu(B) = 0\}$$

konstruieren. Wir können auch $N = \cup_{j\in\mathbb{N}} N_j$ und

$$\nu_a(A) = \nu(A \cap N^c), \quad \nu_s(A) = \nu(A \cap N)$$

setzen. Dann gilt $\mu(N) = 0$ und deswegen, insbesondere, $\nu_s \perp \mu$. Sei nun A beliebig mit $\mu(A) = 0$. Dann gilt auch $\mu(A \cap D_j \cap N_j^c) = 0$.Deswegen

$$\nu(A \cap D_j \cap N_j^c) + \nu(N_j) = \nu((A \cap D_j \cap N_j^c) \cup N_j) \leq \nu(N_j)$$

und $\nu(A \cap D_j \cap N_j^c) = 0$. Da $N_\ell \subset D_\ell$ ist $N_\ell^c \cap D_j = D_j$ für alle $\ell \neq j$. Deswegen gilt

$$\nu(A \cap N^c) = \sum_{j \in \mathbb{N}} \nu(A \cap D_j \cap N_j^c) = 0$$

und also $\nu_a(A) = 0$. Es folgt, dass $\nu_a \ll \mu$.

Schlussendlich zeigen wir die Eindeutigkeit der Zerlegung. Wir bemerken zunächst, dass, falls μ ein positives Mass und ν ein positives oder ein signiertes Mass auf einem messbaren Raum $(\Omega, \mathcal{A})$ sind, dann gilt: $\nu \ll \mu$ und $\nu \perp \mu$ implizieren, dass $\nu = 0$. In der Tat: da $\nu \perp \mu$, existiert eine Menge $E \in \mathcal{A}$ mit $|\nu|(E) = 0$ und $\mu(E^c) = 0$. Da $\mu(E^c) = 0$ gilt aber auch $|\nu|(E^c) = 0$, weil $\nu \ll \mu$. Das bedeutet, dass $|\nu|(\Omega) = |\nu|(E) + |\nu|(E^c) = 0$. Es folgt dass $|\nu| = 0$ und also auch, dass $\nu = 0$. Nehmen wir nun an, dass ν ein endliches positives oder signiertes Mass ist, mit zwei Zerlegungen $\nu = \nu_a + \nu_s = \nu'_a + \nu'_s$. Dann muss $\nu_a - \nu'_a = \nu'_s - \nu_s$ (die Annahme, dass ν endlich ist, ist wichtig, damit die Masse $\nu_a - \nu'_a$ und $\nu'_s - \nu_s$ wohldefiniert sind). Da $\nu_a - \nu'_a \ll \mu$ und $\nu'_s - \nu_s \perp \mu$, muss also $\nu_a - \nu'_a = \nu'_s - \nu_s = 0$ gelten. Es folgt: $\nu_a = \nu'_a$ und $\nu_s = \nu'_s$. Der Fall ν σ-endliches positives Mass kann analog betrachtet werden, mit Hilfe einer Partition von Ω in Mengen mit endliches Mass. $\square$

4.4 Ableitung von endlichen Borel Masse auf $\mathbb{R}^n$

In diesem Abschnitt bezeichnen wir mit $\mathcal{C}$ die Familie die aus allen Würfel $W \subset \mathbb{R}^n$ der Form

$$W = [a_1; b_1] \times [a_2; b_2] \times \cdots \times [a_n; b_n]$$

besteht, wobei $[a_1; b_1], \dots, [a_n; b_n]$ abgeschlossenen Intervallen in $\mathbb{R}$ sind, mit gemeinsamer Länge $e(W) = (b_1 - a_1) = \cdots = (b_n - a_n) > 0$.

Definition 4.17. *Sei μ ein positives endliches Mass auf $(\mathbb{R}^n; \mathcal{B}(\mathbb{R}^n))$. Wir definieren die obere Ableitung $\overline{D}\mu(x)$ von μ an der Stelle x durch*

$$\overline{D}\mu(x) = \lim_{\varepsilon \to 0^+} \sup \left\{ \frac{\mu(C)}{\lambda_n(C)} : C \in \mathcal{C}, x \in C, e(C) < \varepsilon \right\}$$

Analog, definieren wir die untere Ableitung $\underline{D}\mu(x)$ von μ an der Stelle x durch

$$\underline{D}\mu(x) = \lim_{\varepsilon \to 0^+} \inf \left\{ \frac{\mu(C)}{\lambda_n(C)} : C \in \mathcal{C}, x \in C, e(C) < \varepsilon \right\}$$

Obere und untere Ableitungen sind überall definiert (weil das Supremum in der Definition von $\overline{D}\mu(x)$ und das Infimum in der Definition von $\underline{D}\mu(x)$ monoton in ε sind). Sie sind Funktionen mit Werten in $[0; +\infty]$ ($+\infty$ ist möglich). Wir sagen μ ist differenzierbar an der Stelle $x \in \mathbb{R}$, falls $\overline{D}\mu(x) = \underline{D}\mu(x) < +\infty$. Ist μ an der Stelle $x \in \mathbb{R}^n$ differenzierbar, so definieren wir die Ableitung von μ an der Stelle $x \in \mathbb{R}^n$ durch

$$D\mu(x) = \overline{D}\mu(x) = \underline{D}\mu(x)$$

Lemma 4.18. *Sei μ ein endliches Mass auf $(\mathbb{R}^n; \mathcal{B}(\mathbb{R}^n))$. Dann sind die Funktionen $\overline{D}\mu, \underline{D}\mu$ und $D\mu$ Borel messbar (die Funktion $D\mu$ ist nur auf der messbare Menge $\{x \in \mathbb{R}^n : \overline{D}\mu(x) = \underline{D}\mu(x) < \infty\}$ definiert).*

Beweis. Sei $\mathcal{U}$ die Familie die aus allen offenen Würfel der Form

$$U = (a_1; b_1) \times (a_2; b_2) \times \cdots \times (a_n; b_n)$$

besteht (hier sind $(a_1; b_1), \ldots, (a_n; b_n)$ offene Intervalle in $\mathbb{R}$, mit gemeinsame Länge $e(U) = (b_1 - a_1) = \cdots = (b_n - a_n)$). Für jede $C \in \mathcal{C}$ finden wir dann eine monoton fallende Folge $U_j \in \mathcal{U}$ mit $C = \bigcap_j U_j$. Das impliziert, dass $\mu(C)/\lambda_n(C) = \lim_{j\to\infty} \mu(U_j)/\lambda_n(U_j)$. Ähnlich, für jede $U \in \mathcal{U}$, existiert eine monoton wachsende Folge $C_j \in \mathcal{C}$, mit $U = \bigcup_j C_j$. Dann gilt auch $\mu(U)/\lambda_n(U) = \lim_{j\to\infty} \mu(C_j)/\lambda_n(C_j)$. Es folgt, dass

$$\begin{aligned}
\overline{D}\mu(x) &= \lim_{\varepsilon \to 0^+} \sup \left\{ \frac{\mu(U)}{\lambda_n(U)} : U \in \mathcal{U}, x \in U, e(U) < \varepsilon \right\} \\
\underline{D}\mu(x) &= \lim_{\varepsilon \to 0^+} \inf \left\{ \frac{\mu(U)}{\lambda_n(U)} : U \in \mathcal{U}, x \in U, e(U) < \varepsilon \right\}
\end{aligned}$$

Für ein festes $\varepsilon > 0$, definieren wir nun die Funktionen

$$\begin{aligned}
s_\varepsilon(x) &= \sup \left\{ \frac{\mu(U)}{\lambda_n(U)} : U \in \mathcal{U}, x \in U, e(U) < \varepsilon \right\} \\
t_\varepsilon(x) &= \inf \left\{ \frac{\mu(U)}{\lambda_n(U)} : U \in \mathcal{U}, x \in U, e(U) < \varepsilon \right\}
\end{aligned}$$

Dann gilt $s_\varepsilon(x) > a$ (für ein beliebiges $a \in \mathbb{R}$), genau dann wenn es existiert $U \in \mathcal{U}$ mit $x \in U$ und $e(U) < \varepsilon$ so, dass $\mu(U)/\lambda_n(U) > a$. Das bedeutet, dass

$$\{x \in \mathbb{R}^n : s_\varepsilon(x) > a\} = \bigcup \left\{ U \in \mathcal{U} : e(U) < \varepsilon,\ \frac{\mu(U)}{\lambda_n(U)} > a \right\}$$

Als Vereinigung offener Menge ist also $\{x \in \mathbb{R}^n : s_\varepsilon(x) > a\}$ offen und deswegen Borel messbar, für alle $a \in \mathbb{R}$. Analog ist auch

$$\{x \in \mathbb{R}^n : t_\varepsilon(x) < a\} = \bigcup \left\{ U \in \mathcal{U} : e(U) < \varepsilon,\ \frac{\mu(U)}{\lambda_n(U)} < a \right\}$$

Borel messbar, für alle $a \in \mathbb{R}$. Die Funktionen $\overline{D}\mu(x) = \lim_{\varepsilon\to 0^+} s_\varepsilon(x)$ und $\underline{D}\mu(x) = \lim_{\varepsilon\to 0^+} t_\varepsilon(x)$ sind deswegen als Limes einer Folge messbarer Funktionen auch messbar. Die Messbarkeit von $D\mu$ folgt aus der Bemerkung, dass die Menge der Punkten wo zwei messbare Funktionen übereinstimmen messbar ist (Satz 2.4). $\square$

Das folgende Theorem ist das Hauptresultat dieses Abschnittes.

Theorem 4.19. *Sei μ ein positives endliches Mass auf $(\mathbb{R}^n; \mathcal{B}(\mathbb{R}^n))$. Dann ist μ λ_n-fast überall differenzierbar und die Funktion*

$$x \to \begin{cases} D\mu(x) & \textit{falls } \mu \textit{ differenzierbar an der Stelle } x \\ 0 & \textit{sonst} \end{cases}$$

ist eine Radon-Nikodym Ableitung für den absolut stetigen Teil von μ.

Um Theorem 4.19 zu beweisen, werden wir Vitali-Überdeckungssatz benutzen.

Definition 4.20. *Sei $A \subset \mathbb{R}^n$. Eine Vitali-Überdeckung von A ist eine Familie $\mathcal{V} \subset \mathcal{C}$ so, dass, für alle $x \in A$ und alle $\delta > 0$, ein $W \in \mathcal{V}$ existiert, mit $x \in W$ und $e(W) < \delta$.*

Satz 4.21 (Vitali Überdeckungssatz)**.** *Sei $A \subset \mathbb{R}^n$, $A \neq \emptyset$. Sei $\mathcal{V}$ eine Vitali-Überdeckung von A. Dann existiert eine endliche oder abzählbare Folge $(W_n)_{n\in\mathbb{N}}$ von disjunkten Würfel in $\mathcal{V}$ so, dass $\lambda_n^*(A \backslash \bigcup_{n\in\mathbb{N}} W_n) = 0$ (d.h. die Überdeckung mit Würfel enthält fast alle Punkten von A).*

Beweis. Wir betrachten zunächst den Fall, dass A beschränkt ist. Wir wählen eine beschränkte offene Menge $U \subset \mathbb{R}^n$, mit $A \subset U$. Wir bezeichnen mit $\mathcal{V}_0$ die Menge aller $W \in \mathcal{V}$, mit $W \subset U$. $\mathcal{V}_0$ ist dann wieder eine Vitali-Überdeckung von A. In der Tat, für $x \in A \subset U$ beliebig finden wir $\kappa > 0$ so, dass $B_\kappa(x) \subset U$. Das bedeutet, dass, für jede Würfel W mit $x \in W$ und $e(W) < \kappa/\sqrt{n}$, $W \subset B_\kappa(x) \subset U$ gilt. Da $\mathcal{V}$ eine Vitali-Überdeckung von A ist, finden wir, für jede $\delta > 0$, ein $W \in \mathcal{V}$ mit $x \in W$ und mit $e(W) < \min(\delta, \kappa/\sqrt{n})$. Dann ist auch $W \in \mathcal{V}_0$. Damit ist $\mathcal{V}_0$ auch eine Vitali-Überdeckung von A ist.

Wir setzen
$$\delta_1 = \sup\{e(W) : W \in \mathcal{V}_0\}$$
Da $A \neq \emptyset$ und U beschränkt ist, gilt $0 < \delta_1 < \infty$. Wir können deswegen $W_1 \in \mathcal{V}_0$ finden, mit $e(W_1) > \delta_1/2$. Gilt $A \subset W_1$, dann sind wir fertig. Nehmen wir dagegen an, dass $x \in A \backslash W_1$ existiert. Da W_1 abgeschlossen ist, finden wir $\delta > 0$, so, dass $B_\delta(x) \cap W_1 = \emptyset$. Da $\mathcal{V}_0$ eine Vitali-Überdeckung von A ist, es existieren also $W \in \mathcal{V}_0$, mit $W \cap W_1 = \emptyset$. Wir setzen dann
$$\delta_2 = \sup\{e(W) : W \in \mathcal{V}_0 \text{ und } W \cap W_1 = \emptyset\}$$
und, da $0 < \delta_2 < \infty$, finden wir $W_2 \in \mathcal{V}_0$ mit $W_2 \cap W_1 = \emptyset$ und $e(W_2) > \delta_2/2$. Wir iterieren das Verfahren. Nach Wahl von $\{\delta_j\}_{j=1}^k$ und $\{W_j\}_{j=1}^k$ in $\mathcal{V}_0$ unterscheiden wir zwei Fälle. Gilt $A \subset \cup_{j=1}^k W_j$, dann sind wir fertig. Gilt dagegen $A \not\subset \cup_{j=1}^k W_j$, dann existiert $W \in \mathcal{V}_0$ mit $W \cap \cup_{j=1}^k W_j = \emptyset$ (weil $\cup_{j=1}^k W_j$ abgeschlossen ist). Deswegen, mit
$$\delta_{k+1} = \sup\{e(W) : W \in \mathcal{V}_0, W \cap \cup_{j=1}^k W_j = \emptyset\}$$
gilt $0 < \delta_{k+1} < \infty$ und wir können $W_{k+1} \in \mathcal{V}_0$ finden, mit $W_{k+1} \cap \cup_{j=1}^k W_j = \emptyset$ und $e(W_{k+1}) > \delta_{k+1}/2$.

Wir behaupten nun, dass die konstruierte Folge $\{W_j\}$, die, nach Definition, aus disjunkten Würfeln in $\mathcal{V}_0$ besteht, die gewünschte Eigenschaft $\lambda_n(A \backslash \cup_j W_j) = 0$. Wenn die Konstruktion nach $N < \infty$ Schritten unterbricht, dann gilt $A \subset \cup_{j=1}^N W_j$, und deswegen $\lambda_n(A \backslash \cup_{j=1}^N W_j) = 0$. Nehmen wir nun an, die Konstruktion produziert unendlich viele Würfel W_j aus $\mathcal{V}_0$. Da $\cup_j W_j \subset U$, da U beschränkt ist und da W_j disjunkt sind, muss $\sum_j \lambda_n(W_j) < \infty$. Insbesondere, $\lambda_n(W_j) \to 0$, für $j \to \infty$. Das impliziert auch, dass $\delta_j \to 0$, für $j \to \infty$. Für $k \in \mathbb{N}$ bezeichnen wir nun mit D_k das Würfel mit dem selben Mittelpunkt wie W_k, aber mit $e(D_k) = 5e(W_k)$. Dann gilt $\lambda_n(D_k) = 5^n \lambda_n(W_k)$ und also $\sum_j \lambda_n(D_j) = 5^n \sum_j \lambda_n(W_j) < \infty$.

Wir behaupten nun, dass für alle $N \in \mathbb{N}$,
$$A \backslash \cup_{j=1}^N W_j \subset \cup_{j=N+1}^\infty D_j \tag{60}$$

Das impliziert dann, dass (wir benutzen hier das Lebesgue äusseres Mass λ_n^*, weil wir nicht angenommen haben, dass A messbar ist)

$$\lambda_n^*(A \backslash \cup_{j=1}^\infty W_j) \leq \lambda_n^*(A \backslash \cup_{j=1}^N W_j) \leq \lambda_n^*(\cup_{j=N+1}^\infty D_j) \leq \sum_{j=N+1}^\infty \lambda_n(D_j)$$

Da die rechte Seite gegen Null strebt, für $N \to \infty$ (weil $\sum_j \lambda_n(D_j) < \infty$), und da die linke Seite unabhängig aus N ist, es folgt, dass $\lambda_n^*(A \backslash \cup_{j=1}^\infty W_j) = 0$.

Um (60) zu zeigen, nehme an, dass $x \in A \backslash \cup_{j=1}^N W_j$. Da $\mathcal{V}_0$ eine Vitali-Überdeckung von A ist, finden wir $W \in \mathcal{V}_0$ mit $x \in W$ und mit $W \cap \cup_{j=1}^N W_j = \emptyset$. Dann existiert $\ell > N$, mit $W \cap \cup_{j=1}^\ell W_j \neq \emptyset$. In der Tat, wenn $W \cap \cup_{j=1}^\ell W_j = \emptyset$ für alle $\ell > N$, dann wäre $e(W) \leq \delta_\ell$ für alle $\ell > N$, was aber unmöglich ist, da $\delta_\ell \to 0$, für $\ell \to \infty$. Sei also $\ell_0 > N$ so, dass $W \cap \cup_{j=1}^{\ell_0 - 1} W_j = \emptyset$ aber $W \cap \cup_{j=1}^{\ell_0} W_j \neq \emptyset$. Nach Definition von δ_{ℓ_0} muss dann $e(W) \leq \delta_{\ell_0}$ gelten. Anderseits haben wir W_{ℓ_0} so gewählt, dass $e(W_{\ell_0}) > \delta_{\ell_0}/2$. Da $W \cap W_{\ell_0} \neq \emptyset$ und $e(W_{\ell_0}) > e(W)/2$, es folgt (ein Bild hilft), dass $W \subset D_{\ell_0}$. D.h.

$$x \in W \subset D_{\ell_0} \subset \cup_{j=N+1}^\infty D_j$$

für alle $x \in A \backslash \cup_{j=1}^N W_j$. Das zeigt (60).

Sei nun $A \subset \mathbb{R}^n$, nicht notwendigerweise beschränkt. Wir finden eine Folge die aus disjunkten offene Mengen $U_j \subset \mathbb{R}^n$ besteht, mit $\lambda_n(U_j) < \infty$ für alle $j \in \mathbb{N}$, und mit $\lambda_n^*(\mathbb{R}^n \backslash \cup_j U_j) = 0$. Für alle $j \in \mathbb{N}$ mit $A \cap U_j \neq \emptyset$ finden wir wie oben eine (endliche oder abzählbare) Folge $W_m^j \subset U_j$ von disjunkten Würfel in der Vitali Überdeckung $\mathcal{V}$, so, dass $\lambda_n^*((A \cap U_j) \backslash \cup_m W_m^j) = 0$. Wenn wir alle diese Folge vereinigen, dann kriegen wir eine einzige Folge $\{W_m\}$ von disjunkten Würfel aus $\mathcal{V}$, mit $\lambda_n^*(A \backslash \cup_m W_m) = 0$ (es ist wichtig, dass die U_j disjunkt sind, damit $W_m^j \cap W_r^i = \emptyset$ für alle m, r, falls $i \neq j$). $\square$

Mit Hilfe von Satz 4.21 zeigen wir nun zwei Lemmata, die wir dann zum Beweis von Theorem 4.19 benutzen werden.

Lemma 4.22. *Sei μ ein endliches Mass auf $(\mathbb{R}^n; \mathcal{B}(\mathbb{R}^n))$, $a > 0$. Sei $A \in \mathcal{B}(\mathbb{R}^n)$ mit $\overline{D}\mu(x) \geq a$ für alle $x \in A$. Dann gilt $\mu(A) \geq a\lambda_n(A)$.*

Beweis. O.B.d.A. können wir annehmen, dass $A \neq \emptyset$. Sei U eine offene Menge, mit $A \subset U$. Sei $0 < \delta < a$ und sei

$$\mathcal{V} = \{W \in \mathcal{C} : W \subset U, \mu(W) \geq (a - \delta)\lambda_n(W)\}$$

Aus $\overline{D}\mu(x) \geq a$ für alle $x \in A$ es folgt, dass $\mathcal{V}$ eine Vitali-Überdeckung von A ist. In der Tat,

$$\lim_{\varepsilon \to 0^+} \sup \left\{ \frac{\mu(W)}{\lambda_n(W)} : W \in \mathcal{C}, x \in W, e(W) < \varepsilon \right\} \geq a$$

impliziert, dass für alle $\delta > 0$,

$$\sup \left\{ \frac{\mu(W)}{\lambda_n(W)} : W \in \mathcal{C}, x \in W, e(W) < \varepsilon \right\} \geq a - \delta/2$$

für alle $\varepsilon > 0$ klein genug (abhängig aus δ). D.h. für alle $\varepsilon > 0$ klein genug, finden wir $W \in \mathcal{C}$ mit $x \in W$, $e(W) < \varepsilon$ und mit $\mu(W) < (a-\delta)\lambda_n(W)$. Da U offen ist, und $x \in A \subset U$, folgt aus $x \in W$ und $e(W) < \varepsilon$, dass $W \subset U$, falls $\varepsilon > 0$ klein genug ist. Das zeigt, dass die Familie $\mathcal{V}$ tatsächlich eine Vitali-Überdeckung von A ist.

Mit Satz 4.21 finden wir eine Folge W_j die aus disjunkte Würfel in $\mathcal{V}$ besteht, mit $\lambda_n(A \backslash \cup_j W_j) = 0$. Wir erhalten

$$\mu(U) \geq \mu(\cup_j W_j) = \sum_j \mu(W_j) \geq \sum_j (a-\delta)\lambda_n(W_j) = (a-\delta)\lambda_n(\cup_j W_j) = (a-\delta)\lambda_n(A)$$

für alle offene $U \subset \mathbb{R}^n$, mit $A \subset U$. Aus der Regularität von μ (in den Übungen wurde gezeigt, dass jedes endliches Mass regulär ist) folgt, dass $\mu(A) \geq (a-\delta)\lambda_n(A)$. Da $\delta > 0$ beliebige ist, kriegen wir $\mu(A) \geq \lambda_n(A)$. $\square$

Lemma 4.23. *Sei μ ein endliches Mass auf $(\mathbb{R}^n; \mathcal{B}(\mathbb{R}^n))$, absolut stetig bezüglich λ_n. Sei $a > 0$, $A \in \mathcal{B}(\mathbb{R}^n)$ mit $\underline{D}\mu(x) \leq a$ für alle $x \in A$. Dann gilt $\mu(A) \leq a\lambda_n(A)$.*

Beweis. O.B.d.A. nehmen wir an, dass $A \neq \emptyset$. Sei U offen, mit $A \subset U$. Wie im Beweis von Lemma 4.22, mit Hilfe von Satz 4.21, finden wir, für alle $\delta > 0$, eine Folge von disjunkten Würfel $W_j \in \mathcal{C}$ mit $W_j \subset U$, $\mu(W_j) \leq (a+\delta)\lambda_n(W_j)$ und $\lambda_n(A \backslash \cup_j W_j) = 0$. Da $\mu \ll \lambda_n$ finden wir auch $\mu(A \backslash \cup_j W_j) = 0$. Deswegen

$$(a+\delta)\lambda_n(U) \geq (a+\delta)\sum_j \lambda_n(W_j) \geq \sum_j \mu(W_j) = \mu(\cup_j W_j) = \mu(A)$$

für alle $U \subset \mathbb{R}^n$ offen, mit $A \subset U$. Da λ_n regulär ist (aus Satz 1.31), es folgt, dass $(a+\delta)\lambda_n(A) \geq \mu(A)$. Da $\delta > 0$ beliebig ist, erhalten wir $a\lambda_n(A) \geq \mu(A)$. $\square$

Beweis von Theorem 4.19. Wir betrachten zunächst den Fall, $\mu \perp \lambda_n$. Sei dann $N \in \mathcal{B}(\mathbb{R})$, mit $\lambda_n(N) = 0 = \mu(N^c)$. Für $k \in \mathbb{N}$, sei

$$B_k = \{x \in N^c : \overline{D}\mu(x) \geq 1/k\}$$

Die Menge B_k ist messbar, weil $\overline{D}\mu$ eine Borel messbare Funktion ist (bei Lemma 4.18). Aus Lemma 4.22 erhalten wir, dass $\lambda_n(B_k) \leq k\mu(B_k) \leq k\mu(N^c) = 0$. Also

$$\lambda_n(\{x \in N^c : \overline{D}\mu(x) > 0\}) = 0$$

Da auch $\lambda_n(N) = 0$, es folgt, dass

$$\lambda_n(\{x \in \mathbb{R}^n : \overline{D}\mu(x) > 0\}) = 0$$

Also, $\overline{D}\mu(x) = 0$ λ_n-fast überall. Da $0 \leq \underline{D}\mu(x) \leq \overline{D}\mu(x)$ für alle $x \in \mathbb{R}^n$, wir erhalten, dass μ λ_n-fast überall differenzierbar ist, mit $D\mu(x) = 0$.

Sei nun $\mu \ll \lambda_n$. Für $p, q \in \mathbb{Q}$, mit $p < q$, sei

$$A(p;q) = \{x \in \mathbb{R}^n : \underline{D}\mu(x) \leq p < q \leq \overline{D}\mu(x)\}$$

Da $\underline{D}\mu, \overline{D}\mu$ Borel messbar sind, ist $A(p;q) \in \mathcal{B}(\mathbb{R})$. Aus Lemma 4.22 und Lemma 4.23, erhalten wir

$$q\lambda_n(A(p;q)) \leq \mu(A(p;q)) \leq p\lambda_n(A(p;q))$$

Da μ endlich ist, folgt zunächst, dass $\lambda_n(A(p;q)) < \infty$. Da $p < q$ folgt dann, dass $\lambda_n(A(p;q)) = 0$. Deswegen

$$\lambda_n(\{x \in \mathbb{R}^n : \underline{D}\mu(x) < \overline{D}\mu(x)\}) = \lambda_n(\cup_{p,q\in\mathbb{Q}:p<q} A(p;q)) \leq \sum_{p,q\in\mathbb{Q}:p<q} \lambda_n(A(p;q)) = 0$$

Da aber $\underline{D}\mu(x) \leq \overline{D}\mu(x)$ für alle $x \in \mathbb{R}^n$, es folgt, dass $\overline{D}\mu(x) = \underline{D}\mu(x)$ λ_n-fast überall (das zeigt noch nicht, dass μ fast überall differenzierbar ist, weil wir haben den Fall $\overline{D}\mu(x) = \underline{D}\mu(x) = \infty$ noch nicht ausgeschlossen).

Da $\mu \ll \lambda_n$ angenommen wurde, finden wir aus Theorem 4.11 eine Radon-Nikodym Ableitung $f \in \mathcal{L}^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda_n)$ mit $\mu(A) = \int_A f d\lambda_n$ für alle $A \in \mathcal{B}(\mathbb{R})$. Ist nun $A \in \mathcal{B}(\mathbb{R})$, mit $f(x) \geq a$ für alle $x \in A$, dann gilt offenbar $\mu(A) \geq a\lambda_n(A)$. Ist dagegen $f(x) \leq a$ für alle $x \in A$, dann gilt $\mu(A) \leq a\lambda_n(A)$. Wir können also wie oben argumentieren, um zu zeigen, dass $f(x) \leq \underline{D}\mu(x)$ und, dass $f(x) \geq \overline{D}\mu(x)$ für λ_n-fast alle $x \in \mathbb{R}^n$. In der Tat, für $p, q \in \mathbb{Q}$ mit $p < q$, können wir die Menge

$$A(p;q) = \{x \in \mathbb{R}^n : \underline{D}\mu(x) \leq p < q \leq f(x)\}$$

definieren. Da $f, \underline{D}\mu$ messbar sind, ist $A(p;q) \in \mathcal{B}(\mathbb{R})$. Da $f(x) \geq q$ und $\underline{D}\mu(x) \leq p$ für alle $x \in A(p;q)$, erhalten wir (mit Lemma 4.23)

$$q\lambda_n(A(p;q)) \leq \mu(A(p;q)) \leq p\lambda_n(A(p;q))$$

Das impliziert $\lambda_n(A(p;q)) = 0$ für alle rationale $p < q$ und damit auch, dass

$$\lambda_n(\{x \in \mathbb{R}^n : f(x) > \underline{D}_\mu(x)\}) = 0$$

Der Beweis, dass $\lambda_n(\{x \in \mathbb{R}^n : f(x) < \overline{D}_\mu(x)\}) = 0$, ist ähnlich. Da $f(x)$ λ_n-fast überall endlich ist, es folgt, dass $\underline{D}\mu(x) = \overline{D}\mu(x) = f(x) < \infty$ für λ_n-fast alle$x \in \mathbb{R}^n$.

Schlussendlich betrachten wir den Fall, dass μ ein beliebiges positives endliches Mass auf $(\mathbb{R}; \mathcal{B}(\mathbb{R}))$ ist. In diesem Fall können wir $\mu = \mu_a + \mu_s$ zerlegen, mit endlichen positiven Masse $\mu_a \ll \lambda_n$ und $\mu_s \perp \lambda_n$ (die Existenz (und Eindeutigkeit) der Lebesgue Zerlegung von μ wurde in Theorem 4.16 bewiesen). Nach Definition von Ableitung von μ, es gilt

$$\overline{D}\mu(x) = \overline{D}\mu_a(x) + \overline{D}\mu_s(x), \qquad \text{und} \qquad \underline{D}\mu(x) = \underline{D}\mu_a(x) + \underline{D}\mu_s(x)$$

Sei $f \in \mathcal{L}^1(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \lambda_n)$ die Radon-Nikodym Ableitung von μ_a. Es folgt aus was wir oben bewiesen haben, dass $\overline{D}\mu_a(x) = \underline{D}\mu_a(x) = f(x)$ und, dass $\overline{D}\mu_s(x) = \underline{D}\mu_s(x) = 0$ für λ_n-fast alle $x \in \mathbb{R}^n$. Das impliziert, dass $\overline{D}\mu(x) = \underline{D}\mu(x) = f(x)$ für λ_n-fast alle $x \in \mathbb{R}^n$, und damit, dass μ λ_n-fast überall differenzierbar ist, mit $D\mu(x) = f(x)$. $\square$

Eine erste Anwendung vom letzten Theorem ist der Dichtesatz von Lebesgue.

Definition 4.24. *Sei $E \subset \mathbb{R}^n$ Lebesgue messbar. Ein Punkt $x \in \mathbb{R}^n$ heisst ein Dichtepunkt von E falls, für alle $\varepsilon > 0$ es existiert $\delta > 0$ mit*

$$\left|\frac{\lambda_n(E \cap W)}{\lambda_n(W)} - 1\right| < \varepsilon$$

für alle Würfel $W \in \mathcal{C}$ mit $x \in W$ und $e(W) < \delta$. Ein Punkt $x \in \mathbb{R}^n$ heisst ein Dispersionpunkt von E falls x ein Dichtepunkt von E^c ist.

Satz 4.25 (Lebesgue Dichtesatz)**.** *Sei $E \subset \mathbb{R}^n$ Lebesgue messbar. Dann sind λ_n-fast alle Punkte in E Dichtepunkte von E und λ_n-fast alle Punkte in E^c Dispersionpunkte von E.*

Beweis. Nehmen zunächst an, dass $\lambda_n(E) < \infty$. Wir finden $E_0 \in \mathcal{B}(\mathbb{R})$, mit $E_0 \subset E$ und $\lambda_n(E \backslash E_0) = 0$ (möglich, zB. weil die σ-Algebra der Lebesgue messbare Mengen die Vervollständigung von $\mathcal{B}(\mathbb{R})$ ist). Auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ definieren wir $\mu(A) = \lambda_n(A \cap E_0)$. Es gilt $\mu \ll \lambda_n$. Es folgt aus Theorem 4.19, dass μ differenzierbar und, dass $D\mu(x) = \chi_{E_0}(x)$, für λ_n-fast alle $x \in \mathbb{R}^n$ (weil χ_{E_0}, die charakteristische Funktion von E_0, die Radon-Nikodym Ableitung von μ bezüglich λ_n ist). Es folgt, dass $D\mu(x) = 1$ für λ_n-fast alle $x \in E_0$. Nach Definition der Ableitung $D\mu$, das bedeutet genau, dass x ein Dichtepunkt von E_0 für λ_n-fast alle $x \in E_0$ ist, und, da $\lambda(E \backslash E_0) = 0$, auch, dass x ein Dichtepunkt von E ist, für λ_n-fast alle $x \in E$. $\square$

5 Endliche Borel Masse auf $\mathbb{R}$

In diesem Kapitel werden wir insbesondere endliche positive oder signierte Masse auf dem messbarer Raum $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ betrachten. Wir haben schon in den Übungen gesehen, dass endliche positive Masse auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ durch die entsprechende Verteilungsfunktion $F : \mathbb{R} \to \mathbb{R}$, definiert durch $F(x) = \mu((-\infty; x])$, charakterisiert werden können. Jede beschränkte, monoton wachsende und rechtsstetige Funktion $F : \mathbb{R} \to \mathbb{R}$, mit der Eigenschaft $\lim_{x \to -\infty} F(x) = 0$, ist die Verteilungsfunktion eines positives Masses auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ Um eine analoge Charakterisierung on Verteilungsfunktionen von signierte Masse herzuleiten, müssen wir den Begriff von Funktionen mit beschränkten Variation einführen.

5.1 Funktionen mit beschränkten Variation

Sei F eine reelwertige Funktion, definiert mindestens auf dem Intervall $[a; b]$. Sei $\mathcal{S}_{[a;b]}$ die Menge aller Teilungen $\{t_j\}_{j=0}^n$, mit $a \leq t_0 < t_1 < \cdots < t_n \leq b$. Wir definieren dann die Variation von F auf $[a; b]$ durch

$$V_F[a; b] = \sup \left\{ \sum_{j=1}^n |F(t_j) - F(t_{j-1})| : \{t_j\}_{j=0}^n \in \mathcal{S}_{[a;b]} \right\}$$

Wir sagen, dass F endliche (oder beschränkte) Variation auf $[a; b]$ hat, falls $V_F[a; b] < \infty$. Analog können wir auch $V_F(-\infty; b]$ und $V_F(-\infty; \infty)$ definieren (jetzt werden einfach die Teilungen andere Intervalle überdecken).

Seien $-\infty < a < b < \infty$. Mit der Bemerkung, dass jede Teilung für $(-\infty; b]$ sich in einer Teilung für $(-\infty; a]$ und einer Teilung für $[a; b]$ zerlegen lässt, erhalten wir, dass

$$V_F(-\infty; b] = V_F(-\infty; a] + V_F[a; b] \tag{61}$$

Wir werden auch die Bemerkung benutzen, dass

$$V_F(-\infty; b] = \lim_{a \to -\infty} V_F[a; b] \tag{62}$$

In der Tat, für ein beliebiges $\varepsilon > 0$ finden wir eine Teilung $\{t_j\}_{j=0}^n$ von $(-\infty; b]$ (d.h. $-\infty < t_0 < t_1 < \cdots < t_n \leq b$) mit

$$\sum_{j=1}^{n} |F(t_j) - F(t_{j-1})| \geq V_F(-\infty; b] - \varepsilon$$

Für $a < t_0$ ist aber $\{t_j\}_{j=0}^n$ auch eine Teilung von $[a; b]$. Deswegen gilt

$$V_F(-\infty; b] - \varepsilon \leq V_F[a; b] \leq V_F(-\infty; b]$$

für alle $a < t_0$. Das impliziert (62).

Eine Funktion $F : \mathbb{R} \to \mathbb{R}$ hat beschränkte Variation, falls $V_F(-\infty; \infty) < \infty$. Ist $F : \mathbb{R} \to \mathbb{R}$ eine Funktion mit beschränkten Variation, dann können wir eine neue Funktion $V_F : \mathbb{R} \to \mathbb{R}$ durch $V_F(x) = V_F(-\infty; x]$ definieren. Die Funktion V_F hat einige wichtige Eigenschaften.

Lemma 5.1. *Sei F eine Funktion mit beschränkten Variation. Dann ist V_F beschränkt und monoton und es gilt $\lim_{x \to -\infty} V_F(x) = 0$. Ferner, wenn F rechtsstetig ist, so ist auch V_F rechtsstetig.*

Beweis. Die Monotonie folgt, weil $V_F\, I_1 \leq V_F\, I_2$ für jede zwei Intervalle I_1, I_2 mit $I_1 \subset I_2$ (weil dann ist die Menge der Teilungen von I_2 grösser als die Menge der Teilungen von I_1). Das impliziert auch, dass $V_F(-\infty; x] \leq V_F(-\infty; \infty) < \infty$ für alle $x \in \mathbb{R}$, da F beschränkte Variation auf $\mathbb{R}$ hat. Aus (61) und (62), finden wir

$$\begin{aligned}\lim_{x \to -\infty} V_F(x) &= \lim_{x \to -\infty} V_F(-\infty; x] = \lim_{x \to -\infty} (V_F(-\infty; b] - V_F[x; b]) \\ &= V_F(-\infty; b] - \lim_{x \to -\infty} V_F[x; b] = V_F(-\infty; b] - V_F(-\infty; b] = 0\end{aligned}$$

Sei nun F rechtsstetig. Wir behaupten, dass

$$\lim_{b \to a^+} V_F(-\infty; b] = V_F(-\infty; a]$$

Für $c > a, b$ gilt aber $V_F(-\infty; a] = V_F(-\infty; c] - V_F[a; c]$ und $V_F(-\infty; b] = V_F(-\infty; c] - V_F[b; c]$. Deswegen, es genügt zu zeigen, dass

$$\lim_{b \to a_+} V_F[b; c] = V_F[a; c]$$

Um diese Behauptung zu zeigen, wählen wir $\varepsilon > 0$ fest. Da F rechtsstetig ist, finden wir $\delta > 0$ so, dass

$$|F(t) - F(a)| < \varepsilon/2 \tag{63}$$

für alle $a \leq t < a + \delta$. Anderseits, wir finden eine Teilung $\{t_j\}_{j=0}^n$ von $[a;c]$, mit

$$V_F[a;c] < \sum_{j=1}^n |F(t_j) - F(t_{j-1})| + \varepsilon/2$$

Wir unterscheiden nun verschiedenen Fälle. Ist $t_0 > a$, dann ist $\{t_j\}_{j=0}^n$ auch eine Teilung von $[b;c]$, für alle $b < t_0$. In diesem Fall gilt

$$V_F[a;c] \geq V_F[b;c] \geq \sum_{j=1}^n |F(t_j) - F(t_{j-1})| \geq V_F[a;c] - \varepsilon/2$$

Ist $t_0 = a$ und $t_1 < a + \delta$, dann ist $\{t_j\}_{j=1}^n$ eine Teilung von $[b;c]$, falls $b < t_1$. Mit (63) finden wir dann

$$V_F[a;c] < |F(t_1) - F(a)| + \sum_{j=2}^n |F(t_j) - F(t_{j-1})| + \varepsilon/2 \leq \varepsilon + V_F[b;c]$$

Schlussendlich, wenn $t_0 = a$ aber $t_1 > a + \delta$, benutzen wir die Tatsache, dass $\{a + \delta < t_1 < \cdots < t_n\}$ eine Teilung von $[b;c]$ ist, wenn $b < a + \delta$, um zu schliessen, mit (63), dass

$$\begin{aligned} V_F[a;c] &< |F(t_1) - F(a)| + \sum_{j=2}^n |F(t_j) - F(t_{j-1})| + \varepsilon/2 \\ &\leq |F(a+\delta) - F(a)| + |F(t_1) - F(a+\delta)| + \sum_{j=2}^n |F(t_j) - F(t_{j-1})| + \varepsilon/2 \\ &\leq \varepsilon + V_F[b;c] \end{aligned}$$

In jedem Fall haben wir bewiesen, dass $|V_F[a;c] - V_F[b;c]| < \varepsilon$, falls $b > a$ und $|b - a|$ klein genug ist. $\square$

Eine wichtige Eigenschaft von Funktionen mit beschränkten Variation ist die Tatsache, dass sie immer als Differenz zwei beschränkte und monoton wachsende Funktionen geschrieben werden können.

Lemma 5.2. *Sei $F : \mathbb{R} \to \mathbb{R}$ eine Funktion mit beschränkten Variation. Es existieren dann beschränkte, monoton wachsende Funktionen $F_1, F_2 : \mathbb{R} \to \mathbb{R}$ mit $F = F_1 - F_2$.*

Beweis. Wir setzen $F_1 = (F + V_F)/2$ und $F_2 = (V_F - F)/2$. Dann gilt offenbar, dass $F = F_1 - F_2$, und, dass F_1, F_2 beschränkt sind. Wir müssen noch zeigen, dass F_1, F_2 monoton wachsend sind. Um diese Behauptung zu zeigen, wählen wir $x < y$ in $\mathbb{R}$. Die Ungleichung $F_1(x) \leq F_1(y)$ ist dann äquivalent mit

$$F(x) + V_F(x) \leq F(y) + V_F(y) \qquad \Longleftrightarrow \qquad V_F(y) - V_F(x) \geq F(x) - F(y)$$

d.h. mit $V_F[x;y] \geq F(x) - F(y)$, was offenbar aus Definition von $V_F[x;y]$ gilt (wir haben hier die Identität (61) benutzt). Analog zeigt man die Monotonie von F_2. $\square$

Der Grund warum wir an Funktionen mit beschränkten Variation interessiert sind, ist der folgende Zusammenhang mit endlichen signierten Masse auf $(\mathbb{R};\mathcal{B}(\mathbb{R}))$. Sei nämlich μ ein signiertes Mass auf $(\mathbb{R};\mathcal{B}(\mathbb{R}))$. Wir definieren $F_\mu : \mathbb{R} \to \mathbb{R}$ durch $F_\mu(x) = \mu((-\infty;x])$. Dann ist F_μ offenbar beschränkt (weil $|F_\mu(x)| \leq |\mu|((-\infty;x]) \leq |\mu|(\mathbb{R}) < \infty$ für alle $x \in \mathbb{R}$). Wie schon für positive Masse kann man auch zeigen, dass F_μ rechsstetig ist und, dass $\lim_{x\to-\infty} F_\mu(x) = 0$ (beide Aussage können mit einer Jordan Zerlegung von μ bewiesen werden) . In Gegensatz zum Fall von positiven Masse ist aber F_μ nicht notwendigerweise monoton. Was aber auch für signierte Masse gilt, ist, dass F_μ beschränkte Variation hat. In der Tat, für eine beliebige Teilung $\{t_j\}_{j=0}^n$ von $(-\infty;\infty)$ gilt

$$\sum_{j=1}^n |F_\mu(t_j) - F_\mu(t_{j-1})| = \sum_{j=1}^n |\mu((t_{j-1};t_j])| \leq \sum_{j=1}^n |\mu|((t_{j-1};t_j]) \leq |\mu|(\mathbb{R}) < \infty\,.$$

Satz 5.3. *Die Abbildung $\mu \to F_\mu(x) = \mu((-\infty;x])$ definiert eine Bijektion zwischen der Menge aller endlichen signierten Masse auf $(\mathbb{R};\mathcal{B}(\mathbb{R}))$ und der Menge alle rechtsstetige Funktionen mit beschränkten Variation, die im Limes $x \to -\infty$ verschwinden.*

Beweis. Wir haben schon bewiesen, dass für jedes endliches signiertes Mass μ auf $(\mathbb{R},\mathcal{B}(\mathbb{R}))$, die Funktion $F_\mu(x) = \mu((-\infty;x])$ die gewünschte Eigenschaften besitzt. Wir überprüfen jetzt die Injektivität der Abbildung. Seien μ,ν zwei endliche signierte Masse auf $(\mathbb{R},\mathcal{B}(\mathbb{R}))$ mit $F_\mu = F_\nu$. Seien $\mu = \mu^+ - \mu^-$ und $\nu = \nu^+ - \nu^-$ die Jordan'sche Zerlegungen von μ,ν. Dann gilt $F_\mu = F_{\mu^+} - F_{\mu^-}$ und $F_\nu = F_{\nu^+} - F_{\nu^-}$. Deswegen $F_{\mu^+} - F_{\mu^-} = F_{\nu^+} - F_{\nu^-}$, und $F_{\mu^+} + F_{\nu^-} = F_{\nu^+} + F_{\mu^-}$. $F_{\mu^+} + F_{\nu^-}$ ist aber die Verteilungsfunktion vom positiven endlichen Mass $\mu^+ + \nu^-$. $F_{\nu^+} + F_{\mu^-}$ ist dagegen die Verteilungsfunktion vom positiven endlichen Mass $\nu^+ + \mu^-$. Es wurde in den Übungen bewiesen, dass die Verteilungsfunktionen eines positiven Mass eindeutig ist. Das impliziert, dass $\mu^+ + \nu^- = \nu^+ + \mu^-$ und deswegen, dass $\mu = \mu_+ - \mu_- = \nu_+ - \nu_- = \nu$.

Wir müssen noch zeigen, dass die Abbildung surjektiv ist. Für eine beliebige rechtsstetige Funktion F mit beschränkten Variation und mit $F(x) \to 0$ für $x \to -\infty$ müssen wir also ein endliches signiertes Mass μ auf $(\mathbb{R},\mathcal{B}(\mathbb{R}))$ finden mit $F = F_\mu$. Um diese Behauptung zu zeigen, benutzten wir Lemma 5.2 um beschränkte und monoton wachsende Funktionen F_1, F_2 zu finden, mit $F = F_1 - F_2$. Vom Beweis von Lemma 5.2 erinnern wir, dass $F_1 = (F + V_F)/2$ und $F_2 = (V_F - F)/2$. Aus Lemma 5.1 folgt, dass F_1, F_2 auch rechtsstetig sind, und, dass sie $F_1(x), F_2(x) \to 0$, für $x \to -\infty$, erfüllen. Wir wissen deswegen aus den Übungen, dass F_1, F_2 die Verteilungsfunktionen von zwei endliche positive Masse auf $(\mathbb{R},\mathcal{B}(\mathbb{R}))$ sind. Mit anderer Wörter, es existieren μ_1, μ_2 endliche positive Masse auf $(\mathbb{R},\mathcal{B}(\mathbb{R}))$, mit $F_1 = F_{\mu_1}$, $F_2 = F_{\mu_2}$. Dann ist aber $F = F_\mu$, mit dem signierten endlichen Mass $\mu = \mu_1 - \mu_2$. □

5.2 Absolut stetige Funktionen

Einige Eigenschaften eines signiertes Mass μ auf $(\mathbb{R};\mathcal{B}(\mathbb{R}))$ lassen sich durch Untersuchung der entsprechende Funktion $F_\mu(x) = \mu((-\infty;x])$ herleiten. Ein Beispiel ist die absolute Stetigkeit von μ bezüglich dem Lebesgue Mass λ. Um die Verteilungsfunktion von absolut stetige Masse auf $(\mathbb{R};\mathcal{B}(\mathbb{R}))$ zu untersuchen, führen wir den Begriff von absolut stetige Funktionen ein.

Eine Funktion $F : \mathbb{R} \to \mathbb{R}$ heisst absolut stetig, falls für alle $\varepsilon > 0$ ein $\delta > 0$ existiert so, dass für jede endliche Folge $\{(s_j, t_j)\}_{j=1}^n$ von disjunkten, offenen Intervalle in $\mathbb{R}$ mit $\sum_{j=1}^n (t_j - s_j) < \delta$ es gilt

$$\sum_{j=1}^n |F(t_j) - F(s_j)| < \varepsilon$$

Absolute Stetigkeit impliziert offenbar gleichmässige Stetigkeit und deswegen auch Stetigkeit (die Cantor Funktion ist dagegen ein Beispiel einer Funktion die gleichmässig stetig ist, aber nicht absolut stetig). Absolut stetige Funktionen haben immer beschränkte Variation auf kompakten Mengen (aber nicht notwendigerweise auf nicht-kompakte Mengen, zB. $F(x) = x$ ist absolut stetig, aber hat auf $\mathbb{R}$ keine beschränkte Variation).

Lemma 5.4. *Sei $F : \mathbb{R} \to \mathbb{R}$ absolut stetig und mit beschränkte Variation. Dann ist $V_F : \mathbb{R} \to \mathbb{R}$ auch absolut stetig.*

Beweis. Sei $\varepsilon > 0$ festgewählt. Da F absolut stetig ist, finden wir $\delta > 0$ so, dass, für jede endliche Folge $\{(s_j, t_j)\}_{j=1}^n$ von offenen Intervallen mit $\sum_{j=1}^n (t_j - s_j) < \delta$,

$$\sum_{j=1}^n |F(t_j) - F(s_j)| < \varepsilon. \tag{64}$$

Sei also $\{(s_j, t_j)\}_{j=1}^n$ eine solche Folge. Dann

$$\begin{aligned}\sum_{j=1}^n |V_F(t_j) - V_F(s_j)| &= \sum_{j=1}^n V_F[s_j; t_j] \\ &= \sum_{j=1}^n \sup \left\{ \sum_\ell |F(u_\ell^j) - F(u_{\ell-1}^j)| : s_j \leq u_0^j < \cdots < u_{m_j}^j \leq t_j \right\}\end{aligned}$$

Da aber jede Teilung $\{u_\ell^j\}$ von $[s_j; t_j]$ das Intervall in kleinere disjunkte Intervalle zerlegt, mit gesamte Länge kleiner oder gleich $(t_j - s_j)$, und weil $\sum_j (t_j - s_j) < \delta$, es folgt aus (64), dass

$$\sum_{j=1}^n |V_F(t_j) - V_F(s_j)| < \varepsilon$$

Das zeigt, dass V_F absolut stetig ist. □

Theorem 5.5. *Sei μ ein endliches signiertes Mass auf $(\mathbb{R}; \mathcal{B}(\mathbb{R}))$ und sei $F_\mu : \mathbb{R} \to \mathbb{R}$ definiert durch $F_\mu(x) = \mu((-\infty; x])$. Dann ist F_μ absolut stetig genau dann wenn $\mu \ll \lambda$.*

Beweis. Wir nehmen zunächst an, dass $\mu \ll \lambda$. Wir möchten dann zeigen, dass die Funktion F_μ absolut stetig ist. Sei dazu $\varepsilon > 0$ festgewählt. Aus Lemma 4.10 finden wir $\delta > 0$ so, dass $A \in \mathcal{B}(\mathbb{R})$ mit $\lambda(A) < \delta$ impliziert, dass $|\mu|(A) \leq \varepsilon$. Sei nun $\{(s_j; t_j)\}_{j=1}^n$ eine endliche Familie disjunkter offener Intervalle, mit $\sum_{j=1}^n (t_j - s_j) < \delta$. Dann gilt

$$\sum_{j=1}^n |F_\mu(t_j) - F_\mu(s_j)| = \sum_{j=1}^n |\mu((s_j; t_j])| \leq \sum_{j=1}^n |\mu|((s_j; t_j]) = |\mu|(\cup_j (s_j; t_j])) \leq \varepsilon$$

weil $\cup(s_j;t_j]$ eine Borel messbare Menge ist, mit $\lambda(\cup_j(s_j;t_j]) = \sum_j(t_j - s_j) < \delta$. Das zeigt, dass F_μ absolut stetig ist.

Nehmen wir nun an, dass die Funktion F_μ absolut stetig ist. Wir behaupten, dass $\mu \ll \lambda$. Da F_μ absolut stetig ist, nach Lemma 5.4, ist auch die Variation V_F absolut stetig. Deswegen sind auch die monoton wachsende Funktionen $F_1 = (V_F + F)/2$, $F_2 = (V_F - F)/2$, die im Beweis von Lemma 5.2 eingeführt wurden, absolut stetig. Seien nun μ_1, μ_2 die endliche positive Masse mit $F_1 = F_{\mu_1}$ und $F_2 = F_{\mu_2}$. Dann gilt $\mu = \mu_1 - \mu_2$. Es genügt also zu zeigen, dass $\mu_1, \mu_2 \ll \lambda$. Mit anderer Wörter, o.B.d.A. können wir annehmen, dass μ ein endliches positives Mass auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.Wir müssen dann zeigen, dass wenn $F_\mu(x) = \mu((-\infty;x])$ absolut stetig ist, dann ist $\mu \ll \lambda$. Wir wenden wieder Lemma 4.10 an. Sei $\varepsilon > 0$ festgewählt. Da F_μ absolut stetig ist, finden wir $\delta > 0$ so, dass, für jede endliche Familie $\{(s_j;t_j)\}_{j=1}^n$ von offenen disjunkten Intervalle, mit $\sum_j(t_j - s_j) < \delta$, es gilt $\sum_j |F_\mu(t_j) - F_\mu(s_j)| < \varepsilon$. Das bedeutet, dass $\mu(\cup_j(s_j;t_j]) < \varepsilon$. Sei nun $A \in \mathcal{B}(\mathbb{R})$, mit $\lambda(A) < \delta$. Wegen Regularität vom Lebesgue Mass, finden wir eine offene Menge U mit $A \subset U$ und $\lambda(U) < \delta$. Jede offene Menge $U \subset \mathbb{R}$ kann aber als abzählbare Vereinigung von disjunkte offene Intervalle $\{(s_j;t_j)\}_{j\in\mathbb{N}}$ geschrieben werden. Es gilt

$$\sum_{j=1}^n (t_j - s_j) = \lambda(\cup_{j=1}^n (s_j;t_j)) \le \lambda(U) < \delta$$

Es folgt, dass

$$\mu(\cup_{j=1}^n (s_j;t_j)) \le \mu(\cup_{j=1}^n (s_j;t_j]) \le \varepsilon$$

für alle $n \in \mathbb{N}$. Deswegen gilt auch (im Limes $n \to \infty$) $\mu(A) \le \mu(U) \le \varepsilon$. Das zeigt, dass $\mu \ll \lambda$. □

Eine nützliche Anwendung vom letzten Theorem ist die folgende Charakterisierung von absolut stetige Funktionen mit beschränkten Variation, die im Limes $x \to -\infty$ verschwinden.

Proposition 5.6. *Eine Funktion $F : \mathbb{R} \to \mathbb{R}$ ist genau dann absolut stetig, mit beschränkten Variation und mit $F(x) \to 0$, für $x \to -\infty$, wenn sie die Form*

$$F(x) = \int_{-\infty}^x f(t)dt \tag{65}$$

für ein $f \in \mathcal{L}^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ hat.

Beweis. Sei zunächst F eine Funktion der Form (65), für ein $f \in \mathcal{L}^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$. Das signierte Mass $\mu(A) = \int_A f d\lambda$ ist absolut stetig bezüglich bezüglich λ. Ferner, $F_\mu = F$. Es folgt aus Satz 5.3 und aus Theorem 5.5, dass F beschränkte Variation hat, absolut stetig ist, und $F(x) \to 0$ für $x \to -\infty$ erfüllt. Anderseits, sei F eine absolut stetige Funktion mit beschränkte Variation und so, dass $F(x) \to 0$, für $x \to -\infty$. Aus Satz 5.3 und Theorem 5.5 finden wir ein endliches Mass μ, absolut stetig bezüglich λ, mit $F = F_\mu$. Aus Theorem 4.12 finden wir $f \in \mathcal{L}^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ mit

$$\mu(A) = \int_A f d\lambda$$

für alle $A \in \mathcal{A}$. Insbesondere,

$$F(x) = F_\mu(x) = \mu((-\infty; x]) = \int_{-\infty}^{x} f(t)dt$$

$\square$

5.3 Differenzierbarkeit von Funktionen mit beschränkter Variation

Schlussendlich benutzen wir der Zusammenhang mit Masse, um die Differenzierbarkeit von Funktionen $F : \mathbb{R} \to \mathbb{R}$ zu diskutieren. Wir bemerken zunächst, dass Differenzierbarkeit eines endlichen Mass μ auf $(\mathbb{R}; \mathcal{B}(\mathbb{R}))$ die Differenzierbarkeit der entsprechende Verteilungsfunktion impliziert.

Lemma 5.7. *Sei μ ein endliches positives Mass auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ und sei $F_\mu : \mathbb{R} \to \mathbb{R}$ definiert durch $F_\mu(x) = \mu((-\infty; x])$. Ist μ differenzierbar an der Stelle $x_0 \in \mathbb{R}$ (in der Sinne von Def. 4.17), dann ist F_μ differenzierbar an der Stelle x_0, und $F'(x_0) = D\mu(x_0)$.*

Beweis. Die Tatsache, dass μ differenzierbar an der Stelle x_0 ist, impliziert offenbar, dass $\mu(\{x_0\}) = 0$. Es folgt, dass

$$\lim_{x \to x_0^-} F(x) = \lim_{x \to x_0^-} \mu((-\infty; x]) = F(x_0) - \lim_{x \to x_0^-} \mu((x; x_0]) = F(x_0) - \mu(\{x_0\}) = F(x_0),$$

d.h. F ist stetig an der Stelle x_0 (erinnere, dass F immer rechtsstetig ist). Wir haben

$$\frac{F(x) - F(x_0)}{x - x_0} = \frac{\mu((x_0; x])}{x - x_0} = \frac{\mu([x_0; x])}{\lambda([x_0; x])}$$

falls $x_0 < x$ und

$$\frac{F(x) - F(x_0)}{x - x_0} = \frac{\mu((x; x_0])}{\lambda((x; x_0])} = \lim_{k \to \infty} \frac{\mu([x + 1/k; x_0])}{\lambda([x + 1/k; x_0])}$$

Es folgt, dass F differenzierbar ist, mit $F'(x_0) = D\mu(x_0)$, falls μ differenzierbar ist. $\square$

Wir benutzen Lemma 5.7 zusammen mit Theorem 4.19 um zu zeigen, dass monotone Funktionen auf $\mathbb{R}$ fast überall differenzierbar sind.

Theorem 5.8. *Sei $F : \mathbb{R} \to \mathbb{R}$ monoton wachsend. Dann ist F λ-fast überall differenzierbar.*

Um Theorem 5.8 zu zeigen, brauchen wir zunächst ein Paar einfache Eigenschaften von monotone Funktionen.

Lemma 5.9. *Sei $F : \mathbb{R} \to \mathbb{R}$ monoton wachsend. Dann*

i) Es existieren die einseitige Grenzwerte $F(x^+) = \lim_{y \to x^+} F(y)$ und $F(x^-) = \lim_{y \to x^-} F(y)$.

ii) Die Menge $D = \{x \in \mathbb{R} : F \text{ unstetig a.d.S. } x\}$ der Unstetigkeitstellen von F ist endlich oder abzählbar unendlich.

iii) Die Funktion $G : \mathbb{R} \to \mathbb{R}$ definiert durch $G(x) = F(x_+)$ ist monoton wachsend und rechtsstetig, und $G(x) = F(x)$ für alle x, wo F stetig ist.

Beweis. i) F monoton wachsend impliziert, dass $F(x^+) = \inf\{F(t) : t > x\}$, $F(x^-) = \sup\{F(t) : t < x\}$. Insbesondere existieren die einseitige Grenzwerten, für alle $x \in \mathbb{R}$.

ii) Wegen Monotonie gilt auch $F(x^-) \leq F(x) \leq F(x^+)$ für alle $x \in \mathbb{R}$. Das bedeutet, dass F stetig an der Stelle x ist, g.d.w. $F(x^-) = F(x^+)$. Sei $D \subset \mathbb{R}$ die Menge der Unstetigkeitstellen von F. Für $x \in D$, wählen wir $r_x \in \mathbb{Q}$ mit $F(x^-) < r_x < F(x^+)$. Für $x, y \in D$ mit $x < y$ gilt dann (aus Monotonie von F), dass $r_x < r_y$. Die Abbildung $\varphi : D \to \mathbb{Q}$, definiert durch $x \to \varphi(x) = r_x$ ist deswegen injektiv. Die Abzählbarkeit von D folgt also aus der Abzählbarkeit von $\mathbb{Q}$.

iii) Mit $G(x) = F(x^+) = \sup\{F(t) : t < x\}$ ist einfach zu zeigen, dass G monoton und rechtsstetig ist. Aus $F(x^-) \leq F(x) \leq F(x^+)$ folgt auch, dass $F(x) = G(x)$ g.d.w. F ist stetig a.d.S. x. □

Beweis von Theorem 5.8. Nehmen wir zunächst an, dass F beschränkt, monoton wachsend und rechtsstetig ist, mit $\lim_{x\to-\infty} F(x) = 0$. In diesem Fall finden wir ein positives endliches Mass μ auf $(\mathbb{R}; \mathcal{B}(\mathbb{R}))$, mit $F(x) = \mu((-\infty; x])$ für alle $x \in \mathbb{R}$. Es folgt aus Theorem 4.19 und aus Lemma 5.7, dass F λ-fast überall differenzierbar ist.

Nehmen wir nun an, dass F beschränkt, monoton wachsend, mit $\lim_{x\to-\infty} F(x) = 0$. In diesem Fall setzen wir $G(x) = F(x^+)$. Aus Lemma 5.9 folgt, dass G beschränkt (weil F beschränkt ist), monoton wachsend und rechtsstetig, mit $\lim_{x\to-\infty} G(x) = 0$ (weil F diese Eigenschaft hat). Wir haben oben bewiesen, dass G λ-fast überall differenzierbar ist. Sei nun $x_0 \in \mathbb{R}$ so, dass F stetig a.d.S. x_0 ist. Dann gilt $F(x_0) = G(x_0)$ und also

$$\frac{G(x^-) - G(x_0)}{x - x_0} \leq \frac{F(x) - F(x_0)}{x - x_0} \leq \frac{G(x) - G(x_0)}{x - x_0} \tag{66}$$

für alle $x \neq x_0$. Here we defined $G(x^-) = \lim_{x\to x^-} G(x)$. Ist G differenzierbar a.d.S. x_0, dann konvergieren linke und rechte Seite von (66) gegen $G'(x_0)$, für $x \to x_0$. In diesem Fall ist also F auch a.d.S. x_0 differenzierbar, mit $F'(x_0) = G(x_0)$. F kann also nur nicht differenzierbar sein, an den Unstetigkeitstellen von F und an den Stellen wo G nicht differenzierbar ist. Es folgt, dass F λ-fast überall differenzierbar ist.

Schlussendlich, sei $F : \mathbb{R} \to \mathbb{R}$ eine beliebige monoton wachsende Funktion. Um zu zeigen, dass F λ-fast überall differenzierbar ist, gengt es zu zeigen, dass F auf $(a; b)$ λ-fast überall differenzierbar ist, für alle $-\infty < a < b < \infty$ (weil die Vereinigung von abzählbar viele Nullmenge wieder eine Nullmenge ist). Die Funktion F ist aber auf $(a; b)$ λ-fast überall differenzierbar, falls die Funktion $g : \mathbb{R} \to \mathbb{R}$, definiert durch

$$g(x) = \begin{cases} 0, & \text{falls } x \leq a \\ F(x) - F(a), & \text{falls } a < x < b \\ F(b) - F(a), & \text{falls } x \geq b \end{cases}$$

λ-fast überall diferenzierbar ist. Das folgt aber aus oben, weil g beschränkt ist, monoton wachsend, mit $\lim_{x\to-\infty} g(x) = 0$. □

Korollar 5.10. *Sei $F : \mathbb{R} \to \mathbb{R}$ eine Funktion mit beschränkte Variation. Dann ist F λ-fast überall differenzierbar.*

Beweis. Aus Lemma 5.2 kann jede Funktion F mit beschränkten Variation als $F = F_1 - F_2$ geschrieben werden, wo F_1, F_2 monoton wachsend sind. Es folgt aus Theorem 5.8, dass F_1, F_2 λ-fast überall differenzierbar sind. Das impliziert auch, dass F λ-fast überall differenzierbar ist. $\square$

Wenn alle Funktionen mit beschränkten Variation fast überall differenzierbar sind, es stellt sich dann die Frage, was charakterisiert absolut stetige Funktionen? Wir zeigen: Absolut stetige Funktionen sind genau die Funktionen die fast überall differenzierbar sind, und die sich durch Integration der Ableitung rekonstruieren lassen (das gilt für allgemeine Funktionen mit beschränkte Variation nicht). Um diese Aussage zu beweisen, brauchen wir zunächst das folgende Lemma, das auch (wie Theorem 5.8) aus Lemma 5.7 und Theorem 4.19 folgt.

Lemma 5.11. *Sei $f \in \mathcal{L}^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ und $F : \mathbb{R} \to \mathbb{R}$ definiert durch $F(x) = \int_{-\infty}^x f(t)dt$. Dann ist F differenzierbar, mit $F'(x) = f(x)$, für λ-fast alle $x \in \mathbb{R}$.*

Beweis. Wir nehmen zunächst an, dass $f \geq 0$. Dann definieren wir das positives endliches Mass $\mu(A) = \int_A f d\lambda$. Dann gilt $F(x) = \mu((-\infty; x])$. Es folgt aus Theorem 4.19 und aus Lemma 5.7, dass $D\mu(x) = F'(x) = f(x)$ für λ-fast alle $x \in \mathbb{R}$. Für allegemein $f \in \mathcal{L}^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$, folgt die Behauptung durch die Zerlegen $f = f_+ - f_-$. $\square$

Wir erhalten damit die folgende Charakterisierung von absolut stetige Funktionen auf kompakten Intervalle.

Theorem 5.12. *Eine Funktion $F : [a; b] \to \mathbb{R}$ ist absolut stetig genau dann, wenn F λ-fast überall differenzierbar ist, F' integrierbar ist, und F erfüllt*

$$F(x) = F(a) + \int_a^x F'(t)dt$$

Beweis. Nehmen wir zunächst an, dass F absolut stetig ist. Dann hat F beschränkte Variation (weil F in einem kompakten Intervall definiert ist). Wir setzen

$$G(x) = \begin{cases} 0 & \text{for } x \leq a \\ F(x) - F(a) & \text{for } a < x < b \\ F(b) - F(a) & \text{for } x \geq b \end{cases}$$

Dann ist G eine absolut stetige Funktion, mit beschränkten Variation und mit $G(x) \to 0$, für $x \to -\infty$. Es folgt aus Prop. 5.6, dass $f \in \mathcal{L}^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ existiert, mit $G(x) = \int_{-\infty}^x f(t)dt$ für alle $x \in \mathbb{R}$. Das impliziert, dass

$$F(x) = F(a) + \int_a^x f(t)dt$$

für alle $x \in [a; b]$. Aus Theorem 5.11 es folgt, dass F differenzierbar ist, mit $F'(x) = f(x)$, für λ-fast alle $x \in [a; b]$. Das zeigt, dass F' integrierbar (die Tatsache, dass F'

messbar ist, folgt aus der Definition von F' als Limes messbarer Funktionen), und, dass $F(x) = F(a) + \int_a^x F'(t)dt$, wie behauptet.

Anderseits, nehmen wir an, dass F λ-fast überall differenzierbar ist, mit F' integrierbar und

$$F(x) - F(a) = \int_a^x F'(t)dt$$

Dann folgt aus Prop. 5.6, dass F absolut stetig ist. $\square$

A Endlich additive Volumen auf $\mathbb{R}^n$

In Satz 1.1 im Skript haben wir gezeigt, dass kein σ-additives Volumen auf der Potenzmenge $P(\mathbb{R}^n)$ definiert werden kann.

Man könnte sich vorstellen, das Problem ist aus der Bedingung der σ-Additivität gegeben. Man könnte deswegen versuchen, ein Volumen auf $P(\mathbb{R}^n)$ zu definieren, das nur endlich additiv, statt σ-additiv, ist. D.h. man könnte versuchen eine Funktion $\mu : P(\mathbb{R}^n) \to [0; \infty]$ zu finden, mit den folgenden Eigenschaften:

i) Monotonie: ist $A \subset B$, so gilt $\mu(A) \leq \mu(B)$ (folgt eigentlich aus iv') unten).

ii) Euklidische Invarianz: ist $T : \mathbb{R}^n \to \mathbb{R}^n$ eine affine Isometrie, und $A \subset \mathbb{R}^n$, so gilt $\mu(TA) = \mu(A)$.

iii) Normierung: es gilt $\mu([0; \ell]^n) = \ell^n$ für alle $\ell > 0$.

iv') Endliche Additivität: sind $A_1, \dots, A_k \subset \mathbb{R}^n$ disjunkt, so gilt

$$\mu\left(\bigcup_{j=1}^k A_j\right) = \sum_{j=1}^k \mu(A_j)$$

Der nächste Satz besagt aber, dass auch unter die schwächere Annahme iv') von endlichen Additivität, für $n \geq 3$ kein solches Volumen existiert.

Satz A.1. *Für $n \geq 3$ gibt es keine Abbildung $\mu : P(\mathbb{R}^n) \to [0; \infty]$ die die Eigenschaften i), ii), iii), iv') erfüllt.*

Wir beweisen Satz A.1 im Fall $n = 3$ (der Fall $n \geq 4$ kann analog behandelt werden). Dazu verwenden wir die nächste Proposition.

Proposition A.2 (Hausdorff, 1914)**.** *Es existiert eine Folge $(v_n)_{n\in\mathbb{N}}$ so, dass, die Menge*

$$K = \{x \in \mathbb{R}^3 : |x| < 1, x \neq \lambda v_n \text{ für alle } n \in \mathbb{N} \text{ und } \lambda \in \mathbb{R}\},$$

die folgende Eigenschaft hat: Es existieren disjunkte Teilmengen $K_1, K_2, K_3, K_4, K_5 \subset K$ und zwei Rotationen $\varphi, \psi \in SO(3)$ mit

$$K = K_1 \cup K_2 \cup K_3 \cup K_4 \cup K_5$$

und auch

$$K = K_1 \cup \varphi(K_2) = K_3 \cup \psi(K_4).$$

Wir zeigen zunächst wie Proposition A.2 benutzt werden kann, um Satz A.1 zu beweisen, im Fall $n = 3$.

Beweis von Satz A.1. Wir nehmen an, es existiert eine Abbildung $\mu : P(\mathbb{R}^3) \to [0; \infty]$ mit den Eigenschaften i), ii), iii), iv') und wir suchen ein Widerspruch. Sei $B_1 = \{x \in \mathbb{R}^3 : |x| < 1\}$. Es gilt $[-1/2; 1/2]^3 \subset B_1 \subset [-1; 1]^3$, und deswegen, aus i), ii) und iii),

$$1 \leq \mu(B_1) \leq 8$$

Sei nun $(v_n)_{n\in\mathbb{N}}$ die Folge von Vektoren aus Proposition A.2. Wie in Proposition A.2 definieren wir die Menge

$$K = \{x \in \mathbb{R}^3 : |x| < 1 \text{ und } x \neq \lambda v_n \text{ für alle } n \in \mathbb{N} \text{ und } \lambda \in \mathbb{R}\}$$

Wir definieren auch

$$M = \{x \in \mathbb{R}^3 : |x| < 1 \text{ und } x = \lambda v_n \text{ für ein } n \in \mathbb{N} \text{ und ein } \lambda \in \mathbb{R}\}$$

so, dass $B_1 = K \cup M$ und $K \cap M = \emptyset$.

Wir behaupten, dass $\mu(M) = 0$. Aus Proposition A.2, können wir offenbar annehmen, dass $|v_n| = 1$ für alle $n \in \mathbb{N}$ und $v_n \neq v_k$ und $v_n \neq -v_k$ für alle $n \neq k$. Für jede $n, k \in \mathbb{N}$ existiert dann genau eine Ebene, die aus den Vektoren v_n und v_k gespannt wird; wir bezeichnen diese Ebene mit $E_{n,k}$ (jede Vektor auf $E_{n,k}$ hat die Form $\lambda_1 v_n + \lambda_2 v_k$, für geeignete $\lambda_1, \lambda_2 \in \mathbb{R}$). Wir definieren

$$S := \{x \in \mathbb{R}^3 : |x| = 1, x \perp E_{n,k} \text{ für } n, k \in \mathbb{N}\}$$

als die Menge alle Punkten auf dem Rand von B_1, die zu einer Ebene $E_{n,k}$ senkrecht stehen. Für jede $n, k \in \mathbb{N}$ gibt es genau zwei Punkten auf $\partial B_1 = \{x \in \mathbb{R}^3 : |x| = 1\}$ die senkrecht auf $E_{n,k}$ stehen. Deswegen ist S sicher abzählbar. Da aber ∂B_1 überabzählbar ist, finden wir sicher ein $d \in \partial B_1 \backslash S$. Mit d definieren wir den Kreis

$$T = \{x \in \mathbb{R}^3 : |x| = 1 \text{ und } x \cdot d = 0\}$$

Da $d \notin S$, es existieren sicher keine $n, k \in \mathbb{N}$ mit $T \subset E_{n,k}$. Das impliziert, dass $T \cap E_{n,k}$ aus gerade zwei Punkten besteht, für alle $n, k \in \mathbb{N}$. Damit ist die Menge

$$T \cap \bigcup_{n\neq k} E_{n,k} = \bigcup_{n\neq k} (T \cap E_{n,k})$$

abzählbar. Da aber T überabzählbar ist, finden wir $e \in T$ so, dass $e \notin E_{n,k}$ für alle $n, k \in \mathbb{N}$.

Für $t \in [-1; 1]$ setzen wir nun $M_t = M + te$ (wir verschieben M um den Vektor te). Wir bemerken, dass $M_t \cap M_s = \emptyset$ für alle $t \neq s$. Ist, in der Tat, $x \in M_t \cap M_s$ für $t \neq s$, so gibt es $\lambda_1, \lambda_2 \in \mathbb{R}$ und $n, k \in \mathbb{N}$ mit

$$\lambda_1 v_n + te = \lambda_2 v_k + se \;\Rightarrow\; e = \frac{\lambda_2}{t-s} v_k - \frac{\lambda_1}{t-s} v_n \in E_{n,k}$$

in Widerspruch zur Annahme $e \notin E_{n,k}$ für alle $n, k \in \mathbb{N}$. Das zeigt, dass die Menge M_t paarweise disjunkt sind. Aus der Translationsinvarianz von μ folgt auch, dass $\mu(M_t) = \mu(M)$ für alle $t \in [-1; 1]$.

Für alle $\ell \in \mathbb{N}$, wir betrachten

$$N = \bigcup_{n=1}^{\ell} M_{2^{-n}}$$

Offenbar gilt $N \subset [-2; 2]^3$, und deswegen

$$\mu(N) \leq 4^3 = 64$$

Anderseits, weil die Mengen $M_{2^{-n}}$ disjunkt sind, impliziert die endliche Additivität, dass

$$\mu(N) = \sum_{n=1}^{\ell} \mu(M_{2^{-n}}) = \sum_{n=1}^{\ell} \mu(M) = \ell\mu(M)$$

Das impliziert, dass $\ell\mu(M) \leq 64$ für alle $\ell \in \mathbb{N}$. Deswegen muss $\mu(M) = 0$.

Aus $\mu(M) = 0$ folgt nun, dass $\mu(K) = \mu(B_1)$ und damit $1 \leq \mu(K) \leq 8$. Aus Proposition A.2 folgt aber auch, dass

$$\mu(K) = \mu(K_1) + \mu(K_2) + \mu(K_3) + \mu(K_4) + \mu(K_5)$$

und auch, dass (mit der Annahme von Rotationsinvarianz von μ)

$$\mu(K) = \mu(K_1) + \mu(K_2) = \mu(K_3) + \mu(K_4)$$

Das impliziert, dass

$$\mu(K) \geq 2\mu(K)$$

was ein Widerspruch ist. $\square$

Es bleibt nun Proposition A.2 zu beweisen. Dazu brauchen wir ein bisschen Vorbereitung. Wir definieren die zwei Rotationen

$$\varphi = \begin{pmatrix} \frac{1}{3} & \frac{-2\sqrt{2}}{3} & 0 \\ \frac{2\sqrt{2}}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{und} \quad \psi = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{3} & \frac{-2\sqrt{2}}{3} \\ 0 & \frac{2\sqrt{2}}{3} & \frac{1}{3} \end{pmatrix}$$

mit den Inversen

$$\varphi^{-1} = \begin{pmatrix} \frac{1}{3} & \frac{2\sqrt{2}}{3} & 0 \\ \frac{-2\sqrt{2}}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{und} \quad \psi^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{3} & \frac{2\sqrt{2}}{3} \\ 0 & \frac{-2\sqrt{2}}{3} & \frac{1}{3} \end{pmatrix}$$

Offenbar sind $\varphi, \psi \in SO(3) = \{F \in \mathbb{R}^{3\times3} : F^T F = 1, \text{ und } \det F = 1\}$.

Wir betrachten die Gruppe der Rotationen, die aus $\varphi, \varphi^{-1}, \psi, \psi^{-1}$ erzeigt wird. Um den Überblick über die Elemente dieser Gruppe zu behalten, führen wir den Begriff von Wörter ein. Wir sagen ein k-Tupel $(\rho_1, \ldots, \rho_k)$ ist ein Wort der Länge k falls $\rho_j \in \{\varphi, \varphi^{-1}, \psi, \psi^{-1}\}$ für alle $j = 1, \ldots, k$, und falls $\rho_j \rho_{j+1} \neq 1$ für alle $1 \leq j < k$. Die leere Menge $\emptyset$ ist ein Wort der Länge 0. Wir bezeichnen mit $\mathcal{W}$ die Menge von allen Wörter. Sind $w_1 = (\rho_1, \ldots, \rho_k), w_2 = (\rho'_1, \ldots, \rho'_n) \in \mathcal{W}$ zwei Wörter, so definieren wir das Produkt $w_1 \cdot w_2$ als das Wort, das aus dem $(k+n)$ Tupel $(\rho_1, \ldots, \rho_k, \rho'_1, \ldots, \rho'_n)$ entsteht, nach Abkürzung von allen Paaren $\varphi\varphi^{-1}, \varphi^{-1}\varphi, \psi\psi^{-1}, \psi^{-1}\psi$. Zum Beispiel:

$$(\psi, \varphi, \varphi) \cdot (\varphi^{-1}, \psi, \varphi^{-1}) = (\psi, \varphi, \psi, \varphi^{-1})$$

Für jede Wort $w = (\rho_1, \ldots, \rho_k)$ können wir die Inverse $w^{-1} = (\rho_k^{-1}, \rho_{k-1}^{-1}, \ldots, \rho_1^{-1})$ definieren so, dass $w \cdot w^{-1} = w^{-1} \cdot w = \emptyset$. Zum Beispiel:

$$(\psi, \varphi, \varphi^{-1}, \psi)^{-1} = (\psi^{-1}, \varphi, \varphi^{-1}, \psi^{-1})$$

Es folgt, dass $(\mathcal{W}, \cdot)$ eine Gruppe ist, mit $\emptyset$ als Identität. Wir definieren nun eine Abbildung $F : \mathcal{W} \to SO(3)$ durch $F(\emptyset) = 1$ und $F(\rho_1, \ldots, \rho_k) = \rho_1 \ldots \rho_k$ für jede Wort der Länge $k \geq 1$. Es ist einfach zu sehen, dass F ein Gruppenhomomorphismus ist, d.h.

$$F(w_1) \cdot F(w_2) = F(w_1 \cdot w_2)$$

für alle $w_1, w_2 \in \mathcal{W}$ gilt. Das impliziert, dass $F(\mathcal{W})$ ist eine Untergruppe von $SO(3)$.

Lemma A.3. *Sei $G = F(\mathcal{W})$. Dann ist $F : \mathcal{W} \to G$ bijektiv.*

Beweis. Ist $F(w_1) = F(w_2)$, so folgt aus der Tatsache, dass F ein Gruppenhomomorphismus ist, dass

$$1 = F(w_1)F(w_2)^{-1} = F(w_1)F(w_s^{-1}) = F(w_1 w_2^{-1})$$

Es genugt also zu zeigen, dass

$$w \neq \emptyset \;\Rightarrow\; F(w) \neq 1$$

Wir unterscheiden zwei Fällen. Fall 1: Ist $w = (\rho_1, \ldots, \rho_k) \neq \emptyset$ ein Wort mit $\rho_k = \varphi$ oder $\rho_k = \varphi^{-1}$, dann zeigen wir, dass $F(w)e_1 \neq e_1$. Fall 2: Ist $w = (\rho_1, \ldots, \rho_k) \neq \emptyset$ ein Wort mit $\rho_k = \psi$ oder $\rho_k = \psi^{-1}$, so zeigen wir, dass $F(w)e_3 \neq e_3$. Die Beweise für die zwei Fällen sind sehr ähnlich; wir betrachten nur den Fall 1.

Sei $w = (\rho_1, \ldots, \rho_k)$ ein Wort der Länge $k \geq 1$ mit $\rho_k = \varphi$ oder $\rho_k = \varphi^{-1}$. Wir behaupten zunächst, dass es $a, b, c \in \mathbb{Z}$ mit

$$F(w)e_1 = \frac{1}{3^k}(a, b\sqrt{2}, c) \tag{67}$$

existieren. Wir zeigen (67) durch Induktion über k. Für $k = 1$, wir haben

$$\varphi e_1 = (1/3, 2\sqrt{2}/3, 0), \quad \text{and} \quad \varphi^{-1}e_1 = (1/3, -2\sqrt{2}/3, 0) \tag{68}$$

und (67) gilt mit $a = 1, b = \pm 1, c = 0$. Wir nehmen nun an, dass Gleichung (67) für $k = n$ gilt, und wir zeigen sie für $k = n+1$. Sei $w = (\rho_1, \ldots, \rho_n, \rho_{n+1}) \in \mathcal{W}$ ein Wort der Länge $(n+1)$ mit $\rho_{n+1} = \varphi$ oder $\rho_{n+1} = \varphi^{-1}$. Wir unterscheiden vier Fällen, gemäss der Wert von ρ_1. Sei zunächst $\rho_1 = \varphi$. Dann ist $w = \varphi \cdot w'$ für ein Wort w' der Länge n. Aus der Induktion Annahme, finden wir

$$F(w)e_1 = F(\varphi \cdot w')e_1 = \varphi F(w')e_1 = 3^{-n}\,\varphi\,(a, b\sqrt{2}, c) = 3^{-(n+1)}((a-4b), (2a+b)\sqrt{2}, 3c) \tag{69}$$

und damit $F(w)e_1$ hat wieder die Form (67), mit $a, b, c \in \mathbb{Z}$ ersetzt durch die neuen ganzen Zahlen $a' = a - 4b$, $b' = 2a + b$ und $c' = 3c$. Ist $\rho_1 = \varphi^{-1}$, so kann man ähnlich argumentieren (man findet $a' = a + 4b$, $b' = 2a - b$ und $c' = 3c$). Ist $\rho_1 = \psi$, dann schreiben wir $w = \psi \cdot w'$ für ein Wort w' der Länge n.Die Induktion Annahme impliziert, dass

$$F(w)e_1 = \psi F(w')e_1 = 3^{-n}\psi(a, b\sqrt{2}, c) = 3^{-(n+1)}(3a, (b-2c)\sqrt{2}, 4b + c) \tag{70}$$

Auch in diesem Fall, hat also $F(w)e_1$ die Form (67), mit a, b, c ersetzt durch die ganzen Zahlen $a' = 3a$, $b' = b - 2c$, $c' = 4b + c$. Ist $\rho_1 = \psi^{-1}$ so findet man analog $a' = 3a$, $b' = b + 2c$ und $c' = -4b + c$. Das zeigt die Behauptung (67). Für jede Wort $w = (\rho_1, \ldots, \rho_k)$, mit $\rho_k = \varphi$ oder $\rho_k = \varphi^{-1}$, wir bezeichnen mit $a(w), b(w), c(w) \in \mathbb{Z}$ die ganzen Zahlen in der Darstellung (67) von $F(w)e_1$.

Um zu schliessen, dass $F(w)e_1 \neq e_1$, zeigen wir jetzt, dass $b(w)$ nicht durch 3 teilbar ist, für jede Wort w der Länge $k \geq 1$ (deswegen muss $b \neq 0$, und $F(w)e_1 \neq e_1$). Wir zeigen diese Behauptung durch Induktion über die Länge k des Wortes w. Für $k = 1$ folgt die Behauptung aus (68). Sei nun $w = (\rho_1, \ldots, \rho_k)$ ein Wort der Länge $k \geq 2$. Wir unterscheiden verschiedene Fällen. Ist $\rho_1 = \varphi$, $\rho_2 = \psi$, so gilt $w = (\varphi, \psi) \cdot w'$ für ein Wort w' der Länge $k - 2$. Aus Induktion Annahme folgt, dass $b(\psi \cdot w')$ nicht durch 3 teilbar ist. Aus (70) folgt, dass $a(\psi \cdot w') = 3a(w')$ immer durch 3 teilbar ist. Deswegen ist, aus (69), $b(\varphi \cdot \psi \cdot w) = 2a(\psi \cdot w) + b(\psi \cdot w)$ nicht durch 3 teilbar. Analog kann man alle Fälle $(\rho_1, \rho_2) = (\varphi^\sigma, \psi^\tau)$ oder $(\rho_1, \rho_2) = (\psi^\tau, \varphi^\sigma)$, für $\sigma, \tau \in \{\pm 1\}$ behandeln. Sei nun $(\rho_1, \rho_2) = (\varphi, \varphi)$. Dann ist $w = (\varphi, \varphi) \cdot w'$ für ein Wort w' der Länge $k - 2$. Aus (69) finden wir, dass

$$a(\varphi \cdot w') + b(\varphi \cdot w') = a(w') - 4b(w') + 2a(w') + b(w') = 3(a(w') - b(w'))$$

durch 3 teilbar ist. Aus der Induktion Annahme ist $b(\varphi \cdot w')$ dagegen nicht durch 3 teilbar. Wir schliessen (wieder mit (69), dass

$$b((\varphi, \varphi) \cdot w') = 2a(\varphi \cdot w') + b(\varphi \cdot w') = 2(a(\varphi \cdot w') + b(\varphi \cdot w')) - b(\varphi \cdot w')$$

nicht durch 3 teilbar ist. Ist $(\rho_1, \rho_2) = (\psi, \psi)$, so schreiben wir $w = (\psi, \psi) \cdot w'$für ein Wort der Länge $k - 2$ (der nicht mit ψ^{-1} beginnt). Aus (70) bemerken wir, dass

$$b(\psi \cdot w') - c(\psi \cdot w') = (b(w') - 2c(w') - 4b(w') - c(w')) = -3(b(w') + c(w'))$$

durch 3 teilbar ist. Die Induktion Annahme impliziert, dass $b(\psi \cdot w')$ nicht durch 3 teilbar ist. Damit ist (mit (70))

$$b((\psi, \psi) \cdot w') = b(\psi \cdot w') - 2c(\psi \cdot w') = 2(b(\psi \cdot w') - c(\psi \cdot w')) - b(\psi \cdot w')$$

nicht durch 3 teilbar. Die Fälle $(\rho_1, \rho_2) = (\varphi^{-1}, \varphi^{-1})$ und $(\rho_1, \rho_2) = (\psi^{-1}, \psi^{-1})$ können analog behandelt werden. Das zeigt, dass $b(w)$ nicht durch 3 teilbar ist, für alle Wörter w der Länge $k \geq 1$. $\square$

Wir kommen nun zum Beweis von Proposition A.2.

Beweis von Proposition A.2. Wir bezeichnen mit $\mathcal{W}_\sigma \subset \mathcal{W}$ die Menge der Wörter, die mit $\sigma \in \{\varphi, \varphi^{-1}, \psi, \psi^{-1}\}$ anfangen. Wir setzten auch $\mathcal{W}_e = \{\emptyset\}$. Dann ist

$$\mathcal{W} = \mathcal{W}_e \cup \mathcal{W}_\varphi \cup \mathcal{W}_{\varphi^{-1}} \cup \mathcal{W}_\psi \cup \mathcal{W}_{\psi^{-1}}$$

Sei nun $w \in \mathcal{W} \backslash \mathcal{W}_\varphi$. Dann fängt w nicht mit φ an, und also $w' = \varphi^{-1} \cdot w \in \mathcal{W}_{\varphi^{-1}}$ (weil es keine Kürzung gibt). Deshalb gibt es für jede $w \in \mathcal{W} \backslash \mathcal{W}_\varphi$ ein eindeutiges $w' \in \mathcal{W}_{\varphi^{-1}}$ mit $w = \varphi \cdot w'$, und

$$\mathcal{W} = \mathcal{W}_\varphi \cup \varphi \cdot (\mathcal{W}_{\varphi^{-1}})$$

Analog gilt

$$\mathcal{W} = \mathcal{W}_\psi \cup \psi \cdot (\mathcal{W}_{\psi^{-1}})$$

Sei nun $G_\sigma = F(\mathcal{W}_\sigma) \subset SO(3)$. Da F ein Gruppenhomomorphismus ist, gilt

$$G = F(\mathcal{W}) = G_e \cup G_\varphi \cup G_{\varphi^{-1}} \cup G_\psi \cup G_{\psi^{-1}}$$

und auch

$$G = G_\varphi \cup \varphi(G_{\varphi^{-1}}), \quad G = G_\psi \cup \psi(G_{\psi^{-1}}) \tag{71}$$

Für $g \in G$ sei nun $r_g = \{x \in \mathbb{R}^3 : gx = x\}$. Für jede $g \in G \backslash \{1\}$ ist r_g eine Gerade durch Null (r_g ist die Rotationsachse von g). Wir definieren

$$K = \{x \in \mathbb{R}^3 : |x| < 1\} \backslash \bigcup_{g \in G \backslash \{1\}} r_g \tag{72}$$

gegeben ist. Auf K definieren wir eine Äquivalenzrelation. Für $x, y \in K$ setzten wir $x \sim y$ falls es existiert $g \in G$ mit $x = gy$. Mit der Auswahlaxiom finden wir eine Menge $A \subset K$ die aus jeder Äquivalenzklasse genau ein Element enthält. Für jede $x \in K$ existiert dann genau ein $y \in A$ und ein $g \in G$ mit $x = gy$. Die Eindeutigkeit von y folgt aus der Tatsache, dass Aenthält nur ein Element aus jeder Äquivalenzklasse. Die Eindeutigkeit von $g \in G$ kann wie folgt bewiesen werden. Seien $g, g' \in G$ mit $gy = g'y$. Dann ist $g^{-1}g'y = y$, und deswegen $y \in r_{g^{-1}g'}$. Da $y \in K$ muss deswegen $g^{-1}g' = 1$ und also $g = g'$. Wir haben also bewiesen, dass

$$K = A \cup G_\varphi A \cup G_{\varphi^{-1}} A \cup G_\psi A \cup G_{\psi^{-1}} A$$

wobei die fünf Mengen disjunkt sind. Aus (71), folgt auch, dass

$$K = G_\varphi A \cup \varphi(G_{\varphi^{-1}} A) = G_\psi A \cup \psi(G_{\psi^{-1}} A)$$

Schlussendlich bemerken wir, dass die Gruppe G abzählbar ist. Das folgt, weil

$$\mathcal{W} = \bigcup_{k \geq 0} \{ \text{ Menge der Wörter mit der Länge } k \ \}$$

als abzählbare Vereinigung von endlichen Mengen sicher abzählbar ist. Deswegen können wir eine Bijektion $\mathbb{N} \ni n \to g_n \in G$ finden. Für jede $n \in \mathbb{N} \backslash \{0\}$ wählen wir nun ein Vektor $v_n \in r_{g_n}$. Dann folgt aus (72), dass

$$K = \{x \in \mathbb{R}^3 : |x| < 1, \text{ und } x \neq v_n \text{ für alle } n \in \mathbb{N}\}$$

und die Proposition ist bewiesen. □

Im Fall $n = 1, 2$ ist dagegen möglich ein endlich additives Volumen zu konstruieren, mit den Eigenschaften i)-iv'). Wir möchten nun diese Aussage beweisen, für den Fall $n = 1$ (und der Einfachkeit halber nur für Teilmengen von $[0; 1)$). Wir werden dazu ein wichtiges Theorem aus der Funktionalanalysis verwenden, das Hahn-Banach theorem.

Theorem A.4 (Hahn-Banach)**.** *Sei V ein Vektorraum über $\mathbb{R}$ und $p : V \to \mathbb{R}$ ein (nicht unbedingt lineares) Funktional auf V mit*

- *Subadditivität: $p(x + y) \leq p(x) + p(y)$, für alle $x, y \in V$.*
- *$p(\lambda x) = \lambda p(x)$ für alle $\lambda > 0$.*

Sei W ein linearer Unterraum von V und $f : W \to \mathbb{R}$ ein lineares Funktional auf W, mit $f(x) \leq p(x)$ für alle $x \in W$. Dann existiert ein lineares Funktional $F : V \to \mathbb{R}$ mit $F(x) = f(x)$ für alle $x \in W$ und $F(x) \leq p(x)$ für alle $x \in V$ (man nennt F eine Erweiterung von f auf V).

Der Beweis von Theorem A.4 wird in der Vorlesung Funktionalanalysis diskutiert (bemerke: der Beweis braucht das Auswahlaxiom). Hier möchten wir das Theorem anwenden, um den folgenden Satz zu zeigen.

Satz A.5 (Banach, 1923)**.** *Sei $P([0; 1))$ die Potenzmenge des Intervalls $[0; 1) \subset \mathbb{R}$. Es existiert eine Funktion $\mu : P([0; 1)) \to [0; \infty)$ mit*

a) μ ist translationsinvariant: für jede $E \subset [0; 1)$ und $x \in [-1; 1]$ gilt $\mu(E+x) = \mu(E)$. Hier bezeichnet $E + x$ die Menge

$$E+x = \{z \in [0; 1) : z = x+y \text{ oder } z = x+y-1 \text{ oder } z = x+y+1 \text{ für ein } y \in E\}$$

(D.h. die Menge E wird mit "periodische Randbedingungen" verschoben).

b) Normierung: $\mu([a; b]) = b - a$ für alle $0 \leq a < b < 1$.

c) Endliche Additivität: sind $E_1, E_2 \subset [0; 1)$ disjunkt, so gilt

$$\mu(E_1 \cup E_2) = \mu(E_1) + \mu(E_2)$$

(Die endliche Additivität folgt dann durch Induktion).

Beweis. Wir definieren die Menge

$$V := \{f : \mathbb{R} \to \mathbb{R} : \ f \text{ ist beschränkt, und } f(x + 1) = f(x) \text{ für alle } x \in \mathbb{R}\}$$

aller beschränkten periodischen Funktionen mit Periode 1. V ist in natürliche Weise ein Vektorraum über $\mathbb{R}$.

Für $f \in V$, $n \in \mathbb{N}\backslash\{0\}$ und $\alpha_1, \dots, \alpha_n \in \mathbb{R}$, setzen wir

$$M(f, \alpha_1, \dots, \alpha_n) = \sup_{x\in\mathbb{R}} \frac{1}{n} \sum_{j=1}^{n} f(x + \alpha_j)$$

Für $f \in V$ definieren wir weiter

$$p(f) = \inf_{n \in \mathbb{N}\setminus\{0\}, \alpha_1, \ldots, \alpha_n \in \mathbb{R}} M(f, \alpha_1, \ldots, \alpha_n)$$

Offenbar gilt $p(\lambda f) = \lambda p(f)$ für jede $\lambda \geq 0$.

Wir behaupten nun, p ist subadditiv. Um diese Behauptung zu beweisen, wählen wir $\varepsilon > 0$ fest. Wir finden dann $p, q \in \mathbb{N}\setminus\{0\}$ und $\alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q \in \mathbb{R}$ mit

$$p(f) \geq M(f, \alpha_1, \ldots, \alpha_p) - \varepsilon, \quad \text{und} \quad p(g) \geq M(g, \beta_1, \ldots, \beta_q) - \varepsilon$$

Wir betrachten nun die Folge $\gamma_1, \ldots, \gamma_n$, mit $n = pq$ und so, dass jede γ_j die Summe eines α_i mit einem β_ℓ ist. Dann gilt

$$\begin{aligned} p(f+g) &\leq M(f+g, \gamma_1, \ldots, \gamma_n) \\ &= \sup_x \frac{1}{pq} \sum_{i=1}^{p} \sum_{j=1}^{q} (f+g)(x + \alpha_i + \beta_j) \\ &\leq \sup_x \frac{1}{q} \sum_{j=1}^{1} \frac{1}{p} \sum_{i=1}^{p} f(x + \beta_j + \alpha_i) + \sup_x \frac{1}{p} \sum_{i=1}^{p} \frac{1}{q} \sum_{j=1}^{q} f(x + \alpha_i + \beta_j) \end{aligned}$$

Da aber

$$\frac{1}{p} \sum_{i=1}^{p} f(x + \beta_j + \alpha_i) \leq M(f, \alpha_1, \ldots, \alpha_p)$$

für alle $x, \beta_j \in \mathbb{R}$, und

$$\frac{1}{q} \sum_{j=1}^{q} f(x + \alpha_i + \beta_j) \leq M(g, \beta_1, \ldots, \beta_q)$$

für alle $x, \alpha_i \in \mathbb{R}$, wir finden

$$p(f+g) \leq M(f, \alpha_1, \ldots, \alpha_p) + M(g, \beta_1, \ldots, \beta_q) \leq p(f) + p(g) + 2\varepsilon$$

Da $\varepsilon > 0$ beliebig war, erhalten wir $p(f+g) \leq p(f) + p(g)$ für alle $f, g \in V$.

Sei nun

$$W = \{f \in V : f \text{ auf } [0;1) \text{ höchstens endlich viele Unstetigkeitstellen hat}\}$$

W ist ein linearer Unterraum von V. Für $f \in W$ definieren wir

$$I(f) = \int_0^1 f(x)\, dx$$

als das Riemann'sche Integral von f auf $[0;1)$ ($I(f)$ ist wohldefiniert, weil jede Funktion in f Riemann integrierbar ist). $I : W \to \mathbb{R}$ ist offenbar ein lineares Funktional auf W.

Wir behaupten, dass $I(f) \leq p(f)$ für alle $f \in W$. In der Tat, aus der Periodizität von $f \in W$ gilt

$$\int_0^1 f(x+\alpha) dx = \int_0^1 f(x)$$

für alle $\alpha \in \mathbb{R}$. Deshalb gilt

$$\begin{aligned} I(f) &= \int_0^1 f(x)dx = \frac{1}{n}\sum_{j=1}^{n}\int_0^1 f(x+\alpha_j)dx \\ &= \int_0^1 \left[\frac{1}{n}\sum_{j=1}^{n} f(x+\alpha_j)\right]dx \leq \left[\sup_x \frac{1}{n}\sum_{j=1}^{n} f(x+\alpha_j)\right]\int_0^1 dx = M(f,\alpha_1,\ldots,\alpha_n) \end{aligned}$$

für alle $n \in \mathbb{N}$ und alle $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$. Das impliziert, dass $I(f) \leq p(f)$ für alle $f \in W$.

Wir können nun das Hahn-Banach Theorem anwenden. Aus Theorem A.4 folgt, dass es ein lineares Funktional $F : V \to \mathbb{R}$ existiert, mit $F(f) = I(f)$ für alle $f \in W$, und $F(f) \leq p(f)$ für alle $f \in V$. Wir bemerken, dass das lineare Funktional F die folgenden Eigenschaften hat.

1) F ist translationsinvariant. Für $x_0 \in [0;1)$ sei $T_{x_0} : V \to V$ durch $(T_{x_0}f)(x) = f(x+x_0)$ definiert. Dann gilt $F(T_{x_0}f) = T(f)$ für alle $f \in V$.

   Um diese Aussage zu beweisen, setzen wir $g(x) = f(x+x_0) - f(x)$. Dann ist $g \in V$. Weiter wählen wir $n \in \mathbb{N}$ und $\alpha_1 = 0$, $\alpha_2 = x_0$, $\ldots$, $\alpha_n = (n-1)x_0$. Dann gilt

$$p(g) \leq M(g,\alpha_1,\ldots,\alpha_n) = \frac{1}{n}\sup_x\left[f(x+nx_0) - f(x)\right]$$

   Da f beschränkt ist, muss $p(g) \leq C/n$ gelten. Da weiter $n \in \mathbb{N}$ beliebig ist, muss $p(g) \leq 0$. Analog kann man zeigen, dass$p(-g) \leq 0$. Dann muss aber einerseits

$$F(g) \leq p(g) \leq 0$$

   und anderseits

$$-F(g) = F(-g) \leq p(-g) \leq 0$$

   Das bedeutet, dass $F(g) = 0$, und damit, dass $F(T_{x_0}f) = F(f)$.

2) F ist positiv, d.h. $F(f) \geq 0$ für jede $f \in V$ mit $f(x) \geq 0$ für alle $x \in \mathbb{R}$.

   Um diese Aussage zu beweisen, gehen wir wie folgt vor. Ist $f(x) \leq 0$ für alle $x \in \mathbb{R}$, dann gilt $p(f) \leq 0$ (nach Definition von p). Dann muss $F(f) \leq p(f) \leq 0$. Ist nun $f \in V$ mit $f(x) \geq 0$ für alle $x \in \mathbb{R}$, so ist $-f(x) \leq 0$ für alle $x \in \mathbb{R}$, und deswegen $F(-f) \leq 0$. Da aber $F(-f) = -F(f)$, finden wir, dass $F(f) \geq 0$, wie behauptet.

Wir benutzen nun das Funktional F, um das Mass μ auf $P([0;1))$ zu definieren. Für $E \subset [0;1)$ beliebig, betrachten wir die charakteristische Funktion $\chi_E : [0;1) \to \mathbb{R}$ definiert durch $\chi_E(x) = 1$ für $x \in E$ und $\chi_E(x) = 0$ für $x \in [0;1)\backslash E$. Wir betrachten auch die periodische Fortsetzung $\widetilde{\chi}_E : \mathbb{R} \to \mathbb{R}$ von χ_E, definiert durch $\widetilde{\chi}_E(x) = \chi_E(x-n)$ für $x \in [n;n+1)$, für alle $n \in \mathbb{Z}$. Dann ist $\widetilde{\chi}_E \in V$ für jede $E \subset [0;1)$. Wir setzten

$$\mu(E) := F(\widetilde{\chi}_E)$$

Dann gilt:

- $\mu(E) \geq 0$ für alle $E \subset [0;1)$. Das folgt aus der Eigenschaft 2) (Positivität) von F, die wir oben bewiesen haben, weil $\widetilde{\chi}_E(x) \geq 0$ für alle $x \in \mathbb{R}$.

- Sind $E_1, E_2 \subset [0;1)$ disjunkt, so ist $\mu(E_1 \cup E_2) = \mu(E_1) + \mu(E_2)$. Das folgt aus der Linearität von F weil, für disjunkten $E_1, E_2 \subset [0;1)$ ist $\chi_{E_1 \cup E_2} = \chi_{E_1} + \chi_{E_2}$ und damit auch $\widetilde{\chi}_{E_1 \cup E_2} = \widetilde{\chi}_{E_1} + \widetilde{\chi}_{E_2}$.

- $\mu(E + x_0) = \mu(E)$ für jede $E \subset [0;1)$ und $x_0 \in (-1;1)$. Mit der Definition von der Menge $E + x_0$, die wir in Satz A.5 gegeben haben, gilt $\widetilde{\chi}_{E+x_0}(x) = \widetilde{\chi}_E(x - x_0) = (T_{-x_0}\widetilde{\chi}_E)(x)$. Damit folgt die Translationsinvarianz von μ aus der Translationsinvarianz von F, die wir im Punkt 1) oben gezeigt haben.

- Es gilt $\mu([a;b]) = b - a$ für alle $0 \leq a < b < 1$. In der Tat, die charakteristische Funktion $\widetilde{\chi}_{[a;b]}$ hat in $[0;1)$ nur zwei Unstetigkeitstellen. Deswegen ist $\widetilde{\chi}_{[a;b]} \in W$, und
$$F(\widetilde{\chi}_{[a;b]}) = \int_0^1 \chi_{[a;b]}(x)dx = b - a$$
wobei wir das Riemann'sche Integral von $\chi_{[a;b]}$ berechnet haben.

Damit hat μ alle gewünschten Eigenschaften. □

Bemerkungen:

- Im Beweis von Satz A.5 haben wir das Mass μ konstruiert, indem wir zunächst einen neuen Integralbegriff eingeführt haben. Das Funktional $F : V \to \mathbb{R}$ ist nämlich eine Erweiterung des Riemann'sche Integrals auf beliebigen Funktionen. Falls man beliebigen Funktionen integrieren kann, dann kann man das Mass einer beliebigen Teilmenge E einfach als das integral der charakteristische Funktion χ_E definieren. Bei der Einführung des Lebesgue Integral in Kapitel 1 und 2 werden wir anders vorgehen. Wir werden dort zunächst das Lebesgue Mass definieren, und dann das Mass benutzen um das Integral zu konstruieren. Vergleichen mit dem Integral F, das wir im Beweis von Satz A.5 eingeführt haben, werden nicht alle Funktionen Lebesgue integrierbar sein (weil das Lebesgue Mass nur auf eine echte Teilmenge von $P(\mathbb{R})$ definiert ist). Dafür hat das Lebesgue Integral viel bessere Eigenschaften (z.B. beim Vertausch von Limes und Integral), weil das entsprechende Lebesgue Mass σ-additiv ist (während das zu F entsprechende Mass μ, das in Satz A.5 definiert wird) nur endlich additiv ist (sie kann gar nicht σ-additiv sein; das wäre ein Widerspruch zu Satz 1.1 sein).

- Sei $S : V \to V$ durch $(Sf)(x) = f(1-x)$ eine Spiegelung von f um $x = 1/2$. Man kann im Beweis von Satz A.5 das Funktional F durch
$$J(f) := \frac{1}{2}(F(f) + F(Sf))$$
ersetzen. Da $I(Sf) = I(f)$ für jede $f \in W$ (aus Eigenschaften des Riemann'sche Integral), gilt wieder $J(f) = I(f)$ für jede $f \in W$. J ist auch translationsinvariant und positiv, wie F. Mit J kann man dann ein Mass $\widetilde{\mu}$ auf $P([0;1))$ definieren, die alle Eigenschaften von μ hat, und zusätzlich auch invariant bezüglich der Spiegelung

S ist (d.h. $\widetilde{\mu}(S(E)) = \widetilde{\mu}(E)$). Auf dem Intervall $[0;1)$ sind euklidische Transformation genau Translationen (verstanden als "periodische" Translationen, im Sinn der Definition in Satz A.5) und die Spiegelung S. Damit hat das konstruierte Mass $\widetilde{\mu}$ alle Eigenschaften i), ii), iii), iv), die am Anfang dieser Appendix erwähnt werden.