Skript zur Vorlesung

Mass und Integral

Urs Lang

Frühjahrssemester 2018

ETH Zürich

Version vom 16. Juli 2018

Literatur

[Rudin] W. Rudin, Real and Complex Analysis, Third Edition, McGraw-Hill 1987.

[Salamon] D. A. Salamon, Measure and Integration, EMS 2016.

Inhaltsverzeichnis

1 Messbarkeit, Masse

Für eine Menge X bezeichne $2^X = \{A : A \subset X\}$ die Potenzmenge.

1.1 Definition (σ-Algebra, messbarer Raum, messbare Menge)
Sei X eine Menge. Ein System $\mathscr{M} \subset 2^X$ von Teilmengen von X heisst eine *σ-Algebra* in X, falls gilt:

(i) $X \in \mathscr{M}$,
(ii) aus $A \in \mathscr{M}$ folgt $A^c = X \setminus A \in \mathscr{M}$,
(iii) ist $(A_i)_{i\in\mathbb{N}}$ eine Folge in $\mathscr{M}$, so ist $\bigcup_{i=1}^{\infty} A_i \in \mathscr{M}$.

Ist in X eine σ-Algebra $\mathscr{M}$ gewählt, so heisst X (genauer das Paar $(X, \mathscr{M})$) ein *messbarer Raum* und jedes Element von $\mathscr{M}$ eine *messbare Menge*.

1.2 Bemerkungen
Sei $\mathscr{M}$ eine σ-Algebra in X. Dann gilt:

(1) $\emptyset \in \mathscr{M}$, da $\emptyset = X^c$.
(2) Aus $A_1, A_2, \ldots, A_n \in \mathscr{M}$ folgt $\bigcup_{i=1}^{n} A_i \in \mathscr{M}$.
(3) Ist $(A_i)_{i\in\mathbb{N}}$ eine Folge in $\mathscr{M}$, so ist $\bigcap_{i=1}^{\infty} A_i \in \mathscr{M}$, da

$$\bigcap_{i=1}^{\infty} A_i = \Big(\bigcup_{i=1}^{\infty} A_i^c\Big)^c;$$

ebenso ist $\bigcap_{i=1}^{n} A_i \in \mathscr{M}$ für alle $n \in \mathbb{N}$.
(4) Aus $A, B \in \mathscr{M}$ folgt $A \setminus B \in \mathscr{M}$, da $A \setminus B = A \cap B^c$.

Der Präfix σ steht für abzählbare Vereinigungen. Gilt für ein Mengensystem $\mathscr{M} \subset 2^X$ neben 1.1(i) und 1.1(ii) noch 1.2(2), so heisst $\mathscr{M}$ eine *Algebra*.

1.3 Satz (erzeugte σ-Algebra)
Sei X eine Menge. Ist $\mathscr{E} \subset 2^X$, so existiert eine kleinste σ-Algebra $\mathscr{M}$ in X, die $\mathscr{E}$ enthält.

$\mathscr{M}$ heisst die *von $\mathscr{E}$ erzeugte σ-Algebra.*

Beweis: Sei Σ die Menge aller σ-Algebren $\mathscr{M}'$ in X, die $\mathscr{E}$ enthalten; Σ ist nicht leer, da $2^X \in \Sigma$. Setze $\mathscr{M} := \bigcap_{\mathscr{M}'\in\Sigma} \mathscr{M}'$. Dann ist $\mathscr{E} \subset \mathscr{M}$, $\mathscr{M} \subset \mathscr{M}'$ für alle $\mathscr{M}' \in \Sigma$, und $\mathscr{M}$ ist eine σ-Algebra. $\square$

Man vergleiche 1.1 mit:

1.4 Definition (Topologie, topologischer Raum, offene Menge)
Sei X eine Menge. Ein System $\mathscr{O} \subset 2^X$ von Teilmengen von X heisst eine *Topologie* in X, falls gilt:

(i) $\emptyset \in \mathscr{O}$ und $X \in \mathscr{O}$,
(ii) aus $V_1, V_2, \ldots, V_n \in \mathscr{O}$ folgt $\bigcap_{i=1}^n V_i \in \mathscr{O}$,
(iii) ist $(V_i)_{i\in I}$ eine (endliche, abzählbar-unendliche oder überabzählbare) Teilfamilie von $\mathscr{O}$, so ist $\bigcup_{i\in I} V_i \in \mathscr{O}$.

Ist in X eine Topologie gewählt, so heisst X (genauer das Paar $(X, \mathscr{O})$) ein *topologischer Raum* und jedes Element von $\mathscr{O}$ eine *offene Menge*.

Eine *abgeschlossene Menge* in einem topologischen Raum X ist das Komplement einer offenen Menge. Der *Abschluss* $\bar{A}$ einer Menge $A \subset X$ ist der Schnitt aller abgeschlossenen Mengen, die A enthalten; $\bar{A}$ ist abgeschlossen.

Wichtige Beispiele von topologischen Räumen sind *metrische Räume*. Zur Erinnerung: Eine Funktion $d\colon X \times X \to [0, \infty)$ heisst eine *Metrik* auf der Menge X, falls gilt:

(i) $d(x, y) = 0$ genau dann, wenn $x = y$,
(ii) $d(x, y) = d(y, x)$ für alle $x, y \in X$,
(iii) $d(x, z) \leq d(x, y) + d(y, z)$ für alle $x, y, z \in X$.

Ist auf der Menge X eine Metrik d gewählt, so heisst X (genauer das Paar (X, d)) ein *metrischer Raum*. Die von der Metrik induzierte Topologie eines metrischen Raumes X besteht aus allen Mengen $V \subset X$, die sich als Vereinigung von (beliebig vielen) offenen Bällen $B(x, r) = \{y \in X : d(x, y) < r\}$ darstellen lassen.

Ein weiteres für die Masstheorie wichtiges Beispiel eines topologischen Raumes ist die Menge

$$\bar{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\} =: [-\infty, \infty]$$

der *erweiterten reellen Zahlen*, wobei die Topologie aus allen Vereinigungen von Mengen der Gestalt (a, b) oder $(a, \infty] := (a, \infty) \cup \{\infty\}$ oder $[-\infty, b) := \{-\infty\} \cup (-\infty, b)$ besteht $(a, b \in \mathbb{R})$.

1.5 Definition (Borelmenge)
Sei X ein topologischer Raum. Die kleinste σ-Algebra in X, die alle offenen Mengen enthält, wird mit $\mathscr{B}$ bezeichnet; die Elemente von $\mathscr{B}$ heissen *Borelmengen*.

Die σ-Algebra $\mathscr{B}$ der Borelmengen ist also gerade die von der Topologie $\mathscr{O} \subset 2^X$ erzeugte σ-Algebra (vgl. Satz 1.3). Zu $\mathscr{B}$ gehören alle abgeschlossenen Mengen, alle abzählbaren Vereinigungen von abgeschlossenen Mengen (sog. *F_σ-Mengen*), sowie alle abzählbaren Durchschnitte von offenen Mengen (sog. *G_δ-Mengen*[1]).

[1] Diese Bezeichnungen stammen von F. Hausdorff; F steht für *fermé*, G für Gebiet, σ für Vereinigung (Summe) und δ für Durchschnitt.

1.6 Definition (messbare Abbildung)
Sei X ein messbarer Raum, Y ein topologischer Raum (z.B. $\mathbb{R}$, $\bar{\mathbb{R}}$ oder ein metrischer Raum). Eine Abbildung $f\colon X \to Y$ heisst *messbar*, falls $f^{-1}(V)$ messbar ist für jede offene Menge $V \subset Y$.

Offenbar ist die Menge $A \subset X$ genau dann messbar, wenn ihre *charakteristische Funktion* $\chi_A\colon X \to \mathbb{R}$ messbar ist;

$$\chi_A(x) := \begin{cases} 1 & \text{falls } x \in A, \\ 0 & \text{falls } x \in X \setminus A. \end{cases}$$

Sind X, Y zwei topologische Räume, so heisst $f\colon X \to Y$ *Borel-messbar* (oder eine *Borelfunktion*, insb. für $Y = \mathbb{R}, \bar{\mathbb{R}}$), falls $f^{-1}(V)$ eine Borelmenge ist für jede offene Menge $V \subset Y$.

Vergleiche 1.6 mit:

1.7 Definition (stetige Abbildung)
Seien X, Y zwei topologische Räume. Eine Abbildung $f\colon X \to Y$ heisst *stetig*, falls $f^{-1}(V)$ offen ist für jede offene Menge $V \subset Y$.

Jede stetige Abbildung ist Borel-messbar. Sind X ein messbarer Raum, Y und Z topologische Räume, $f\colon X \to Y$ messbar und $g\colon Y \to Z$ stetig, so ist $h := g \circ f\colon X \to Z$ messbar: Ist $W \subset Z$ offen, so ist $g^{-1}(W)$ offen, also $h^{-1}(W) = f^{-1}(g^{-1}(W))$ messbar. Der folgende Satz enthält eine allgemeinere Aussage:

1.8 Satz
Seien $(X, \mathscr{M})$ ein messbarer Raum, Y ein topologischer Raum und $f\colon X \to Y$ eine Abbildung. Dann gilt:

(1) $\{A \subset Y : f^{-1}(A) \in \mathscr{M}\}$ *ist eine σ-Algebra in Y.*
(2) *Ist f messbar und $B \subset Y$ eine Borelmenge, so ist $f^{-1}(B) \in \mathscr{M}$.*
(3) *Ist f messbar und $g\colon Y \to Z$ eine Borel-messbare Abbildung in einen topologischen Raum Z, so ist $h := g \circ f\colon X \to Z$ messbar.*
(4) *Ist $Y = \bar{\mathbb{R}}$ und $f^{-1}((\alpha, \infty]) \in \mathscr{M}$ für jedes $\alpha \in \mathbb{R}$, so ist f messbar.*

Beweis: (1) folgt aus den Relationen $f^{-1}(Y) = X$, $f^{-1}(Y \setminus A) = X \setminus f^{-1}(A)$ und $f^{-1}(\bigcup_{i=1}^{\infty} A_i) = \bigcup_{i=1}^{\infty} f^{-1}(A_i)$.

(2): Die in (1) beschriebene σ-Algebra in Y enthält wegen der Messbarkeit von f alle offenen Mengen und somit auch alle Borelmengen.

(3): Ist $W \subset Z$ offen, so ist $g^{-1}(W)$ ein Borelmenge; somit ist $h^{-1}(W) = f^{-1}(g^{-1}(W))$ in $\mathscr{M}$ gemäss (2).

(4): Setze $\Omega := \{A \subset \bar{\mathbb{R}} : f^{-1}(A) \in \mathscr{M}\}$. Sei $\alpha \in \mathbb{R}$. Wähle $\alpha_1, \alpha_2, \ldots < \alpha$ so, dass $\alpha_i \to \alpha$ für $i \to \infty$. Dann ist $[-\infty, \alpha) = \bigcup_{i=1}^{\infty} [-\infty, \alpha_i] = \bigcup_{i=1}^{\infty} (\alpha_i, \infty]^c$. Nach Voraussetzung ist $(\alpha_i, \infty] \in \Omega$ für alle i, und Ω ist eine σ-Algebra gemäss (1); somit ist $[-\infty, \alpha)$ in Ω, also auch $(\alpha, \beta) = [-\infty, \beta) \cap (\alpha, \infty]$. Jede offene Menge in $\bar{\mathbb{R}}$ ist abzählbare Vereinigung von Mengen der Gestalt $(\alpha, \infty]$ oder (α, β) oder $[-\infty, b)$, also in Ω. Damit ist f messbar. □

1.9 Lemma
Seien X ein messbarer Raum, $u, v\colon X \to \mathbb{R}$ zwei messbare Funktionen und $\phi\colon \mathbb{R}^2 \to Y$ eine stetige Abbildung in einen topologischen Raum Y. Dann ist $h\colon X \to Y$, $h(x) = \phi(u(x), v(x))$, messbar.

Beweis: Für $f\colon X \to \mathbb{R}^2$, $f(x) = (u(x), v(x))$, gilt $h = \phi \circ f$. Wegen der Stetigkeit von ϕ genügt es, die Messbarkeit von f zu zeigen. Für jedes offene Rechteck $R = I \times J \subset \mathbb{R}^2$ (I, J zwei offene Intervalle) ist

$$f^{-1}(R) = \{x \in X : u(x) \in I,\ v(x) \in J\} = u^{-1}(I) \cap v^{-1}(J)$$

messbar, da $u^{-1}(I)$ und $v^{-1}(J)$ messbar sind. Jede offene Menge $V \subset \mathbb{R}^2$ ist Vereinigung einer Folge solcher Rechtecke R_i. Wegen $f^{-1}(V) = f^{-1}(\bigcup_i R_i) = \bigcup_i f^{-1}(R_i)$ ist $f^{-1}(V)$ messbar. □

1.10 Satz
Sei X ein messbarer Raum.

(1) *Für messbare Funktionen $f, g\colon X \to \mathbb{R}$ sind auch $f + g$, fg, $|f|$, $\min\{f, g\}$ und $\max\{f, g\}$ messbar.*
(2) *Für messbare Funktionen $f_1, f_2, \ldots : X \to \bar{\mathbb{R}}$ sind auch $\inf_j f_j$, $\sup_j f_j$, $\liminf_{j\to\infty} f_j$ und $\limsup_{j\to\infty} f_j$ messbar.*

Beweis: Wir zeigen zuerst (2). Sei $g := \sup_j f_j$, $h := \limsup_{j\to\infty} f_j$. Wegen $g^{-1}((\alpha, \infty]) = \bigcup_{j=1}^{\infty} f_j^{-1}((\alpha, \infty])$ und 1.8(4) ist g messbar; analog ist $\inf_j f_j$ messbar. Dann ist auch $h = \inf_{i\geq 1}(\sup_{j\geq i} f_j)$ messbar, ebenso $\liminf_{j\to\infty} f_j$.

(1): Die Messbarkeit von $f + g$ bzw. fg folgt aus 1.9 (mit $\phi(s, t) = s + t$ bzw. $\phi(s, t) = st$). Wegen (2) sind $\min\{f, g\}$ und $\max\{f, g\}$ messbar, somit auch

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

und $|f| = f^+ + f^-$. □

Damit ist auch der Limes jeder punktweise konvergenten Folge von messbaren Funktionen $f_j\colon X \to \bar{\mathbb{R}}$ messbar.

1.11 Definition (Treppenfunktion)
Sei X ein messbarer Raum. Wir nennen eine Funktion $s\colon X \to \mathbb{R}$, die nur endlich viele verschiedene Werte annimmt, eine *Treppenfunktion*. In [Rudin] heissen Treppenfunktionen *simple functions*.

Man beachte, dass hier die Werte $-\infty, \infty$ nicht zugelassen sind. Ist $s\colon X \to \mathbb{R}$ eine Treppenfunktion mit paarweise verschiedenen Werten $\alpha_1, \dots, \alpha_n$, und setzen wir $A_i := s^{-1}\{\alpha_i\}$, so ist $s = \sum_{i=1}^n \alpha_i \chi_{A_i}$, und s ist genau dann messbar, wenn jedes A_i messbar ist.

1.12 Satz (Approximation durch Treppenfunktionen)
Seien X ein messbarer Raum und $f\colon X \to [0,\infty] \subset \bar{\mathbb{R}}$ eine messbare Funktion. Dann existiert eine Folge von messbaren Treppenfunktionen $s_n\colon X \to \mathbb{R}$, so dass
(1) $0 \le s_1 \le s_2 \le \cdots \le f$ *und*
(2) $s_n(x) \to f(x)$ $(n \to \infty)$ *für jedes* $x \in X$.

Beweis: Für $n \in \mathbb{N}$ sei $\varphi_n\colon [0,\infty] \to \mathbb{R}$ die grösste Funktion mit Werten in $2^{-n}\mathbb{Z}$ und der Eigenschaft, dass $\varphi_n(t) \le \inf\{t, n\}$ für alle $t \in [0,\infty]$; φ_n ist eine Borelfunktion mit $t - 2^{-n} < \varphi_n(t) \le t$ für $t \in [0,n]$. Es gilt $0 \le \varphi_1(t) \le \varphi_2(t) \le \cdots \le t$ und $\varphi_n(t) \to t$ $(n \to \infty)$ für $t \in [0,\infty]$. Die Funktionen $s_n := \varphi_n \circ f$ sind messbar nach 1.8(3) und erfüllen (1) und (2). □

1.13 Definition (Mass, Massraum, reelles Mass)
Sei $X = (X, \mathscr{M})$ ein messbarer Raum. Eine Funktion μ auf $\mathscr{M}$ mit Werten in $[0,\infty]$ oder in $\mathbb{R}$ heisst σ*-additiv*, falls

$$\mu\big(\textstyle\bigcup_{i=1}^\infty A_i\big) = \displaystyle\sum_{i=1}^\infty \mu(A_i)$$

für jede Folge paarweise disjunkter Mengen $A_1, A_2, \ldots \in \mathscr{M}$.
Ein *Mass* μ auf X ist eine σ-additive Funktion $\mu\colon \mathscr{M} \to [0,\infty]$ mit $\mu(A) < \infty$ für wenigstens ein $A \in \mathscr{M}$. Ist auf $X = (X, \mathscr{M})$ ein Mass μ gewählt, so heisst X (genauer das Tripel $(X, \mathscr{M}, \mu)$) ein *Massraum*.
Ein *reelles Mass* λ auf X ist eine σ-additive Funktion $\lambda\colon \mathscr{M} \to \mathbb{R}$.

Statt σ-additiv sagt man auch *abzählbar additiv*, statt reelles Mass auch *signiertes Mass*. Beachte, dass ein Mass μ (wie oben definiert) nur dann auch ein reelles Mass ist, wenn $\mu(A) < \infty$ für alle $A \in \mathscr{M}$. In [Rudin] heissen Masse *positive measures*.

1.14 Satz
Sei μ ein Mass auf $(X, \mathscr{M})$. Dann gilt:

(1) $\mu(\emptyset) = 0$.
(2) $\mu(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n \mu(A_i)$ *für paarweise disjunkte* $A_1, \ldots, A_n \in \mathscr{M}$.
(3) $\mu(A) \leq \mu(B)$ *für* $A \subset B$, $A, B \in \mathscr{M}$.
(4) $\mu(A_n) \to \mu(\bigcup_{i=1}^\infty A_i)$ $(n \to \infty)$ *falls* $A_1 \subset A_2 \subset \ldots$, $A_1, A_2, \ldots \in \mathscr{M}$.
(5) $\mu(A_n) \to \mu(\bigcap_{i=1}^\infty A_i)$ $(n \to \infty)$ *falls* $A_1 \supset A_2 \supset \ldots$, $A_1, A_2, \ldots \in \mathscr{M}$ *und* $\mu(A_1) < \infty$.

Masse sind also auch (*endlich*) *additiv* (Eigenschaft (2)) und *monoton* (Eigenschaft (3)).

Beweis: (1): Wähle $A \in \mathscr{M}$ mit $\mu(A) < \infty$ und setze $A_1 := A$, $A_2 = A_3 = \ldots := \emptyset$ in der Definition der σ-Additivität von μ.

(2) folgt dann aus (1) und der σ-Additivität von μ.

(3): Da $B = A \cup (B \setminus A)$ und $A \cap (B \setminus A) = \emptyset$, folgt $\mu(B) = \mu(A) + \mu(B \setminus A) \geq \mu(A)$ aus (2).

(4): Sei $A := \bigcup_{i=1}^\infty A_i$. Setze $B_1 := A_1$ und $B_i := A_i \setminus A_{i-1}$ für $i = 2, 3, \ldots$. Dann ist $B_i \in \mathscr{M}$, $B_i \cap B_j = \emptyset$ für $i \neq j$, $A_n = \bigcup_{i=1}^n B_i$ und $A = \bigcup_{i=1}^\infty B_i$. Es gilt $\mu(A_n) = \sum_{i=1}^n \mu(B_i)$, $\mu(A) = \sum_{i=1}^\infty \mu(B_i)$ und somit $\mu(A_n) \to \mu(A)$ $(n \to \infty)$ nach der Definition der Summe einer unendlichen Reihe.

(5): Sei $A := \bigcap_{i=1}^\infty A_i$. Setze $C_n := A_1 \setminus A_n$. Dann gilt $C_1 \subset C_2 \subset \ldots$, $\mu(C_n) = \mu(A_1) - \mu(A_n)$ und $A_1 \setminus A = \bigcup_{i=1}^\infty C_i$. Mit (4) folgt $\mu(A_1) - \mu(A) = \mu(A_1 \setminus A) = \lim_{n\to\infty} \mu(C_n) = \mu(A_1) - \lim_{n\to\infty} \mu(A_n)$ und somit (5). □

Die Konstruktion von interessanten Massen erfordert einen gewissen Aufwand; wir kommen in Kapitel 3 ausführlich darauf zurück. Ein paar sehr einfache Beispiele lassen sich aber leicht angeben. In beiden der folgenden Beispiele ist X eine beliebige Menge und $\mathscr{M} = 2^X$.

1.15 Beispiele

(1) Es sei $\mu(A) := \#A \in [0, \infty]$ die Anzahl Punkte in $A \subset X$; μ heisst das *Zählmass* auf X.
(2) Sei x ein fest gewählter Punkt in X und
$$\mu(A) := \begin{cases} 1 & \text{falls } x \in A, \\ 0 & \text{sonst;} \end{cases}$$
μ heisst das *Dirac-Mass* im Punkt x und wird oft mit δ_x bezeichnet.

Ist μ das Zählmass auf $\mathbb{N}$ und $A_n := \{n, n+1, n+2, \ldots\}$, so ist $\bigcap_{i=1}^\infty A_i = \emptyset$ aber $\mu(A_n) = \infty$ für alle n. Dies zeigt, dass die Bedingung $\mu(A_1) < \infty$ in 1.14(5) nicht weggelassen werden kann.

2 Integration

Für das Rechnen in $[0, \infty]$ verwenden wir die folgenden Konventionen:

$$a + \infty = \infty + a = \infty \quad \text{für } 0 \le a \le \infty,$$
$$a \cdot \infty = \infty \cdot a = \begin{cases} \infty & \text{falls } 0 < a \le \infty, \\ 0 & \text{falls } a = 0. \end{cases}$$

Man beachte, dass die Kürzungsregeln $a+b = a+c \Rightarrow b = c$ bzw. $ab = ac \Rightarrow b = c$ nur dann gelten, wenn $a < \infty$ bzw. $0 < a < \infty$.

Für $0 \le a_1 \le a_2 \le \dots$, $0 \le b_1 \le b_2 \le \dots$, $a_n \to a$, $b_n \to b$ folgt $a_n + b_n \to a + b$, $a_n b_n \to ab$. Mit 1.12 (Approximation durch Treppenfunktionen) und 1.10(2) (Messbarkeit von $\sup_j f_j$) folgt daraus: Summen und Produkte von messbaren Funktionen mit Werten in $[0, \infty]$ sind messbar. Insbesondere ist für eine messbare $\bar{\mathbb{R}}$-wertige Funktion f auch $|f| = f^+ + f^-$ messbar.

Im Folgenden bezeichnet $(X, \mathscr{M}, \mu)$ einen beliebigen Massraum.

2.1 Definition (Integration von nichtnegativen Funktionen)
Ist $s \colon X \to [0, \infty)$ eine messbare Treppenfunktion mit paarweise verschiedenen Werten $\alpha_1, \dots, \alpha_n$, $s = \sum_{i=1}^n \alpha_i \chi_{A_i}$, und ist $E \in \mathscr{M}$, so setzen wir

$$\int_E s \, d\mu := \sum_{i=1}^n \alpha_i \mu(A_i \cap E)$$

(hier wird die Konvention $0 \cdot \infty = 0$ benützt.) Ist $f \colon X \to [0, \infty]$ messbar, und ist $E \in \mathscr{M}$, so definieren wir

$$\int_E f \, d\mu := \sup \int_E s \, d\mu,$$

wobei das Supremum über alle messbaren Treppenfunktionen s mit $0 \le s \le f$ genommen wird. $\int_E f \, d\mu$ heisst das *Lebesgue-Integral* von f über E bzgl. μ.

Man beachte, dass $\int_E f \, d\mu \in [0, \infty]$. Ist f selbst eine Treppenfunktion, so sind die beiden Definitionen konsistent, und für eine beliebige Darstellung $f = \sum_{i=1}^n \alpha_i \chi_{A_i}$ mit (nicht notwendigerweise paarweise verschiedenen) $\alpha_1, \dots, \alpha_n \in [0, \infty)$ und (möglicherweise leeren) Mengen $A_i \in \mathscr{M}$ gilt offensichtlich $\int_E f \, d\mu = \sum_{i=1}^n \alpha_i \mu(A_i \cap E)$.

2.2 Satz
Seien $f, g \colon X \to [0, \infty]$ messbar und $A, B, E \in \mathscr{M}$. Dann gilt:

(1) *Aus $f \le g$ folgt $\int_E f \, d\mu \le \int_E g \, d\mu$.*

(2) *Aus* $A \subset B$ *folgt* $\int_A f\,d\mu \le \int_B f\,d\mu$.
(3) *Für* $c \in [0,\infty)$ *ist* $\int_E cf\,d\mu = c\int_E f\,d\mu$.
(4) *Ist* $f(x) = 0$ *für alle* $x \in E$, *so ist* $\int_E f\,d\mu = 0$ *(auch wenn* $\mu(E) = \infty$*).*
(5) *Ist* $\mu(E) = 0$, *so ist* $\int_E f\,d\mu = 0$ *(auch wenn* $f(x) = \infty$ *für alle* $x \in E$*).*
(6) $\int_E f\,d\mu = \int_X \chi_E f\,d\mu$.

Beweis: Alle Eigenschaften folgen direkt aus der Definition. □

In 2.1 haben wir vorausgesetzt, dass f auf ganz X definiert sei. Für eine messbare Funktion $f\colon E \to [0,\infty]$ auf einer messbaren Menge $E \subset X$ verwendet man, dass $(E, \mathscr{M}_E, \mu|_{\mathscr{M}_E})$ ein Massraum ist, wobei $\mathscr{M}_E := \{A \in \mathscr{M} : A \subset E\}$.

2.3 Proposition
Seien $s,t\colon X \to [0,\infty)$ *zwei messbare Treppenfunktionen.*
(1) $\varphi(E) := \int_E s\,d\mu$ *definiert ein Mass auf* $\mathscr{M}$.
(2) $\int_X s + t\,d\mu = \int_X s\,d\mu + \int_X t\,d\mu$.

Beweis: (1): Sei $s = \sum_{i=1}^n \alpha_i \chi_{A_i}$. Ist $E = \bigcup_{k=1}^\infty E_k$ für paarweise disjunkte $E_k \in \mathscr{M}$, so folgt $\varphi(E) = \sum_{i=1}^n \alpha_i \mu(A_i \cap E) = \sum_{i=1}^n \alpha_i \sum_{k=1}^\infty \mu(A_i \cap E_k) = \sum_{k=1}^\infty \sum_{i=1}^n \alpha_i \mu(A_i \cap E_k) = \sum_{k=1}^\infty \varphi(E_k)$. Es gilt auch $\varphi(\emptyset) = 0$ (und somit $\varphi \not\equiv \infty$).

(2): Sei s wie oben, und sei $t = \sum_{j=1}^m \beta_j \chi_{B_j}$. Für $E_{ij} := A_i \cap B_j$ gilt dann $\int_{E_{ij}} s + t\,d\mu = (\alpha_i + \beta_j)\mu(E_{ij}) = \alpha_i \mu(E_{ij}) + \beta_j \mu(E_{ij}) = \int_{E_{ij}} s\,d\mu + \int_{E_{ij}} t\,d\mu$. Da X disjunkte Vereinigung aller E_{ij} ist, folgt mit (1) die Behauptung. □

2.4 Satz (Lebesgue, monotone Konvergenz)
Sind $f_1, f_2, \ldots : X \to [0,\infty]$ *messbar und gilt* $f_1(x) \le f_2(x) \le \ldots$ *und* $f_n(x) \to f(x)$ $(n \to \infty)$ *für alle* $x \in X$, *so ist* f *messbar und*

$$\int_X f_n\,d\mu \to \int_X f\,d\mu \quad (n \to \infty).$$

Beweis: Wegen $\int_X f_n\,d\mu \le \int_X f_{n+1}\,d\mu$ existiert ein $\alpha \in [0,\infty]$, so dass

$$\int_X f_n\,d\mu \to \alpha \quad (n \to \infty).$$

Da $f = \sup_n f_n$, ist f messbar. Da $\int_X f_n\,d\mu \le \int_X f\,d\mu$ für alle n, folgt

$$\alpha \le \int_X f\,d\mu.$$

Seien $0 \leq s \leq f$ eine messbare Treppenfunktion und $0 < c < 1$ eine Konstante. Setze

$$E_n := \{x : f_n(x) \geq cs(x)\}$$

für $n = 1, 2, \ldots$; E_n ist messbar, $E_1 \subset E_2 \subset \ldots$ und $X = \bigcup_{n=1}^{\infty} E_n$ (aus $f(x) = 0$ folgt $x \in E_1$, aus $f(x) > 0$ folgt $cs(x) < f(x)$ und somit $x \in E_n$ für ein n). Es gilt

$$\int_X f_n \, d\mu \geq \int_{E_n} f_n \, d\mu \geq c \int_{E_n} s \, d\mu =: c\varphi(E_n).$$

Für $n \to \infty$ folgt mit 2.3(1) und 1.14(4), dass

$$\alpha \geq c\varphi(X) = c \int_X s \, d\mu.$$

Da dies für alle solchen s und c gilt, erhalten wir die noch fehlende Ungleichung $\alpha \geq \int_X f \, d\mu$. □

2.5 Satz

Sind $f_1, f_2, \ldots : X \to [0, \infty]$ messbar und ist $f(x) = \sum_{n=1}^{\infty} f_n(x)$ für alle $x \in X$, so ist f messbar und

$$\int_X f \, d\mu = \sum_{n=1}^{\infty} \int_X f_n \, d\mu.$$

Beweis: Wähle monotone Folgen (s_i'), (s_i'') von messbaren Treppenfunktionen, so dass $s_i' \to f_1$, $s_i'' \to f_2$ (s. 1.12). Für $s_i := s_i' + s_i''$ gilt $s_i \to f_1 + f_2$. Mit 2.3(2) und 2.4 folgt

$$\int_X f_1 + f_2 \, d\mu = \int_X f_1 \, d\mu + \int_X f_2 \, d\mu.$$

Setze $g_N := f_1 + \cdots + f_N$. Mit Induktion über N folgt

$$\int_X g_N \, d\mu = \sum_{n=1}^{N} \int_X f_n \, d\mu$$

für alle N. Die Folge (g_N) konvergiert monoton gegen f, mit 2.4 ergibt sich daher die Behauptung. □

Ist μ das Zählmass auf einer abzählbaren Menge, so entspricht 2.5 der Tatsache, dass

$$\sum_i \sum_j a_{ij} = \sum_j \sum_i a_{ij}$$

für $a_{ij} \in [0, \infty]$.

2.6 Satz (Lemma von Fatou)
Sind $f_1, f_2, \ldots : X \to [0, \infty]$ messbar, so gilt

$$\int_X \liminf_{n\to\infty} f_n \, d\mu \leq \liminf_{n\to\infty} \int_X f_n \, d\mu.$$

Beweis: Sei $g_k(x) := \inf_{i\geq k} f_i(x)$. Die Folge (g_k) konvergiert monoton gegen $\liminf_{n\to\infty} f_n$; mit 2.4 folgt

$$\int_X \liminf_{n\to\infty} f_n \, d\mu = \lim_{k\to\infty} \int_X g_k \, d\mu.$$

Da ausserdem $g_k \leq f_k$ und somit $\int_X g_k \, d\mu \leq \int_X f_k \, d\mu$, folgt der Satz. □

Der nächste Satz verallgemeinert 2.3(1).

2.7 Satz
Sei $f\colon X \to [0, \infty]$ messbar, und sei $\varphi(E) := \int_E f \, d\mu$ für $E \in \mathscr{M}$. Dann ist φ ein Mass auf $\mathscr{M}$, und

$$\int_X g \, d\varphi = \int_X g f \, d\mu$$

für jede messbare Funktion $g\colon X \to [0, \infty]$.

Die obige Identität wird manchmal einfach als $d\varphi = f \, d\mu$ geschrieben, die Symbole $d\varphi$ und $d\mu$ haben dabei aber keine eigene Bedeutung. In Kapitel 6 werden wir eine wichtige Umkehrung von 2.7 kennenlernen, den Satz von Radon–Nikodym.

Beweis: Sei $E = \bigcup_{j=1}^{\infty} E_j$ für paarweise disjunkte $E_j \in \mathscr{M}$. Dann gilt

$$\chi_E f = \sum_{j=1}^{\infty} \chi_{E_j} f, \quad \varphi(E) = \int_X \chi_E f \, d\mu, \quad \varphi(E_j) = \int_X \chi_{E_j} f \, d\mu.$$

Mit 2.5 folgt

$$\varphi(E) = \sum_{j=1}^{\infty} \varphi(E_j).$$

Wegen $\varphi(\emptyset) = 0$ ist φ also ein Mass. Die behauptete Identität gilt für $g = \chi_E$ ($E \in \mathscr{M}$), da

$$\int_X \chi_E \, d\varphi = \varphi(E) = \int_X \chi_E f \, d\mu,$$

und somit auch für jede messbare Treppenfunktion $g\colon X \to [0, \infty]$. Das allgemeine Resultat folgt dann mit 1.12 (Approximation durch Treppenfunktionen) und 2.4 (monotone Konvergenz). □

$(X, \mathscr{M}, \mu)$ bezeichnet weiterhin einen beliebigen Massraum.

2.8 Definition (Integration von reellen Funktionen)
Eine Funktion $f\colon X \to \mathbb{R}$ heisst *Lebesgue-integrabel* (bzgl. μ), falls f messbar und

$$\int_X |f|\, d\mu < \infty$$

ist. Die Menge aller solchen Funktionen bezeichnen wir mit $L^1(\mu)$. (Die Bedeutung des Exponenten 1 wird in Kapitel 5 klar werden.) Ist $f \in L^1(\mu)$ und $f = f^+ - f^-$ die Zerlegung in Positivteil $f^+ = \max\{f, 0\}$ und Negativteil $f^- = -\min\{f, 0\}$, so definiert

$$\int_X f\, d\mu := \int_X f^+\, d\mu - \int_X f^-\, d\mu$$

das *Lebesgue-Integral* von f (bzgl. μ).

Man beachte, dass $|f|$, f^+ und f^- messbar sind und dass $0 \leq f^+ \leq |f|$ und $0 \leq f^- \leq |f|$; somit gilt $\int f^+ \leq \int |f| < \infty$ und $\int f^- \leq \int |f| < \infty$. Ist allgemeiner $f\colon X \to \bar{\mathbb{R}}$ messbar und wenigstens eines der Integrale $\int_X f^+\, d\mu$ und $\int_X f^-\, d\mu$ endlich, so kann man immer noch

$$\int_X f\, d\mu := \int_X f^+\, d\mu - \int_X f^-\, d\mu$$

erklären, wobei jetzt $\int_X f\, d\mu$ möglicherweise ∞ oder $-\infty$ ist.

2.9 Satz
Seien $f, g \in L^1(\mu)$ und $\alpha, \beta \in \mathbb{R}$. Dann gilt:

(1) *$\alpha f + \beta g \in L^1(\mu)$ und $\int_X \alpha f + \beta g\, d\mu = \alpha \int_X f\, d\mu + \beta \int_X g\, d\mu$.*
(2) *Aus $f \leq g$ folgt $\int_X f\, d\mu \leq \int_X g\, d\mu$.*
(3) *$\left|\int_X f\, d\mu\right| \leq \int_X |f|\, d\mu$.*

Beweis: (1): $\alpha f + \beta g$ ist messbar (s. 1.10(1)), und es gilt

$$\begin{aligned}\int_X |\alpha f + \beta g|\, d\mu &\leq \int_X |\alpha||f| + |\beta||g|\, d\mu \\ &= |\alpha| \int_X |f|\, d\mu + |\beta| \int_X |g|\, d\mu < \infty.\end{aligned}$$

Somit ist $\alpha f + \beta g \in L^1(\mu)$. Sei $h := f + g$. Dann ist $h^+ - h^- = f^+ - f^- + g^+ - g^-$, also $h^+ + f^- + g^- = f^+ + g^+ + h^-$ und $\int h^+ + \int f^- + \int g^- = \int f^+ + \int g^+ + \int h^-$ nach 2.5. Da alle diese Integrale endlich sind, folgt

$$\textstyle\int h = \int h^+ - \int h^- = \int f^+ - \int f^- + \int g^+ - \int g^- = \int f + \int g.$$

Es bleibt noch zu zeigen, dass $\int \alpha f = \alpha \int f$. Für $\alpha \geq 0$ ist $\int \alpha f = \int \alpha f^+ - \int \alpha f^- = \alpha \int f^+ - \alpha \int f^- = \alpha \int f$ nach 2.2(3). Wegen $(-f)^+ = f^-$, $(-f)^- = f^+$ gilt ausserdem $\int(-f) = \int f^- - \int f^+ = -\int f$.

(2): $f \leq g$ impliziert $f^+ \leq g^+$ und $f^- \geq g^-$.

(3): $\left|\int f\right| = \alpha \int f = \int \alpha f \leq \int |f|$ für ein $\alpha \in \{-1, 1\}$. □

2.10 Satz (Lebesgue, beschränkte Konvergenz)
Seien $f_1, f_2, \ldots : X \to \mathbb{R}$ messbare Funktionen mit $f_n(x) \to f(x)$ $(n \to \infty)$ für alle $x \in X$. Existiert eine Funktion $g \in L^1(\mu)$, so dass $|f_n(x)| \leq g(x)$ für alle n und alle $x \in X$, so ist $f \in L^1(\mu)$ und es gilt

$$\int_X |f_n - f| \, d\mu \to 0 \quad \textit{und} \quad \int_X f_n \, d\mu \to \int_X f \, d\mu \quad (n \to \infty).$$

Beweis: Die Funktion f ist (als punktweiser Limes der f_n) messbar, und $|f| \leq g$. Somit ist $f \in L^1(\mu)$. Da $|f_n - f| \leq |f_n| + |f| \leq 2g$, können wir 2.6 (Lemma von Fatou) auf die Funktionen $2g - |f_n - f|$ anwenden. Es folgt

$$\begin{aligned}
\int_X 2g \, d\mu &\leq \liminf_{n\to\infty} \int_X 2g - |f_n - f| \, d\mu \\
&= \int_X 2g \, d\mu + \liminf_{n\to\infty} \left(- \int_X |f_n - f| \, d\mu \right) \\
&= \int_X 2g \, d\mu - \limsup_{n\to\infty} \int_X |f_n - f| \, d\mu.
\end{aligned}$$

Da $\int 2g < \infty$, folgt $\limsup_{n\to\infty} \int |f_n - f| \leq 0$, also $\int |f_n - f| \to 0$. Mit 2.9(3) folgt ausserdem $\left|\int f_n - \int f\right| = \left|\int (f_n - f)\right| \leq \int |f_n - f| \to 0$. □

Es folgen einige Bemerkungen zur Rolle von Nullmengen (Mengen vom Mass 0).

2.11 Definition (Nullmengen-Terminologie)
Sei P eine Eigenschaft, die ein Punkt $x \in X$ haben kann, beispielsweise „$f(x) > 0$“ für eine gegebene Funktion f oder „$(f_n(x))$ konvergiert“ für eine gegebene Folge von Funktionen f_n. Ist $X = (X, \mathscr{M}, \mu)$ und $E \in \mathscr{M}$, so bedeutet die Aussage „P gilt (μ-)*fast überall* auf E“ oder „P gilt für (μ-)*fast alle* x in E“, dass eine Menge $N \in \mathscr{M}$ mit $\mu(N) = 0$ existiert, so dass alle $x \in E \setminus N$ die Eigenschaft P haben.

Sind z.B. f, g messbare Funktionen und ist $\mu(\{x : f(x) \neq g(x)\}) = 0$, so sagen wir „$f = g$ (μ-)fast überall“. Dies definiert eine Äquivalenzrelation $\sim$.

Transitivität ($f \sim g$ und $g \sim h$ impliziert $f \sim h$) folgt aus der Tatsache, dass die Vereinigung zweier Nullmengen wieder eine Nullmenge ist. Für $f \sim g$ und $E \in \mathscr{M}$ gilt dann

$$\int_E f\,d\mu = \int_E g\,d\mu$$

(da $E = (E \setminus N) \cup (E \cap N)$, $f = g$ auf $E \setminus N$ und $\mu(E \cap N) = 0$), d.h. „Nullmengen sind bei der Integration vernachlässigbar".

Ein Mass μ heisst *vollständig*, wenn jede Teilmenge einer Nullmenge messbar (und somit wieder eine Nullmenge) ist.

2.12 Satz (Vervollständigung von Massen)
Sei $(X, \mathscr{M}, \mu)$ ein Massraum und $\mathscr{M}^*$ die Menge aller $E \subset X$, für die $A, B \in \mathscr{M}$ existieren, so dass $A \subset E \subset B$ und $\mu(B \setminus A) = 0$; setze in dieser Situation $\mu(E) := \mu(A)$. Dann ist $\mathscr{M}^*$ eine σ-Algebra, und die so erweiterte Funktion μ ist wohldefiniert und ein vollständiges Mass auf $\mathscr{M}^*$.

Beweis: Wir zeigen zuerst, dass $\mu(E)$ für jede Menge $E \in \mathscr{M}^*$ wohldefiniert ist. Es sei $A \subset E \subset B$, $A' \subset E \subset B'$, $\mu(B \setminus A) = \mu(B' \setminus A') = 0$. Da $A \setminus A' \subset E \setminus A' \subset B' \setminus A'$, gilt $\mu(A \setminus A') = 0$ und somit $\mu(A) = \mu(A \cap A')$. Ebenso ist $\mu(A') = \mu(A \cap A')$, also $\mu(A) = \mu(A')$.

Wir zeigen als Nächstes, dass $\mathscr{M}^*$ eine σ-Algebra ist:

(i) $X \in \mathscr{M}^*$, da $X \in \mathscr{M}$ und $\mathscr{M} \subset \mathscr{M}^*$.

(ii) Aus $A \subset E \subset B$ folgt $B^c \subset E^c \subset A^c$, und $A^c \setminus B^c = A^c \cap B = B \setminus A$. Mit $E \in \mathscr{M}^*$ ist also auch $E^c \in \mathscr{M}^*$.

(iii) Aus $A_i \subset E_i \subset B_i$ und $E := \bigcup_{i=1}^\infty E_i$, $A := \bigcup_{i=1}^\infty A_i$, $B := \bigcup_{i=1}^\infty B_i$ folgt $A \subset E \subset B$ und

$$B \setminus A = \textstyle\bigcup_{i=1}^\infty (B_i \setminus A) \subset \bigcup_{i=1}^\infty (B_i \setminus A_i).$$

Da abzählbare Vereinigungen von Nullmengen wieder Nullmengen sind, ist mit $E_i \in \mathscr{M}^*$ $(i = 1, 2, \dots)$ also auch $E \in \mathscr{M}^*$.

Wir zeigen noch die σ-Additivität von μ auf $\mathscr{M}^*$. Sind die E_i in (iii) paarweise disjunkt, dann auch die A_i; es folgt

$$\mu(E) = \mu(A) = \sum_{i=1}^\infty \mu(A_i) = \sum_{i=1}^\infty \mu(E_i).$$

Somit ist μ ein Mass auf $\mathscr{M}^*$, und μ ist offensichtlich vollständig. □

2.13 Bemerkung (erweiterter Messbarkeitsbegriff)
$X = (X, \mathscr{M}, \mu)$ bezeichne weiterhin einen beliebigen Massraum. Da Nullmengen bei der Integration vernachlässigbar sind, ist es zweckmässig, auch

Funktionen (mit Werten in $\mathbb{R}$ oder $\bar{\mathbb{R}}$) *messbar auf* X zu nennen, die nur auf einer Menge $E \in \mathscr{M}$ mit $\mu(E^c) = 0$ definiert und dort messbar sind. Ist f eine solche auf E definierte Funktion und setzt man $f(x) := 0$ für alle $x \in E^c$, so erhält man eine messbare Funktion auf X im ursprünglichen Sinn; ist μ vollständig, so kann man f sogar auf ganz beliebige Weise erweitern. Das Integral von f über $A \in \mathscr{M}$ ist unabhängig von der Definition von f auf E^c, man kann daher oft darauf verzichten, überhaupt eine Fortsetzung zu wählen.

Der folgende Satz ergänzt Satz 2.5.

2.14 Satz
Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge von (μ-fast überall definierten) reellen messbaren Funktionen auf X, so dass

$$\sum_{n=1}^{\infty} \int_X |f_n| \, d\mu < \infty.$$

Dann konvergiert die Reihe $\sum_{n=1}^{\infty} f_n(x) =: f(x)$ für μ-fast alle x, und die dadurch definierte Funktion f ist in $L^1(\mu)$ und erfüllt

$$\int_X f \, d\mu = \sum_{n=1}^{\infty} \int_X f_n \, d\mu.$$

Beweis: Sei f_n auf S_n definiert, $\mu(S_n^c) = 0$. Setze $\varphi(x) := \sum_{n=1}^{\infty} |f_n(x)|$ für $x \in S = \bigcap_{n=1}^{\infty} S_n$. Dann ist $\mu(S^c) = 0$. Nach Satz 2.5 und der Voraussetzung gilt

$$\int_S \varphi \, d\mu = \sum_{n=1}^{\infty} \int_S |f_n| \, d\mu < \infty.$$

Dies impliziert $\mu(E^c) = 0$ für $E := \{x \in S : \varphi(x) < \infty\}$. Nach der Definition von φ konvergiert die Reihe $\sum_{n=1}^{\infty} f_n(x) = f(x)$ absolut für alle $x \in E$, und $|f(x)| \leq \varphi(x)$ für alle $x \in E$. Für $g_n := f_1 + \ldots + f_n$ folgt $|g_n(x)| \leq \varphi(x)$ und $g_n(x) \to f(x)$ für alle $x \in E$. Mit Satz 2.10 (beschränkte Konvergenz) folgt $f \in L^1(\mu)$ auf E und

$$\sum_{i=1}^{n} \int_E f_i \, d\mu = \int_E g_n \, d\mu \to \int_E f \, d\mu.$$

Da $\mu(E^c) = 0$, folgt der Satz. $\square$

Man beachte, dass man die Konvergenz der Reihe auch dann nur *fast überall* erhalten würde, wenn die f_n *überall* auf X definiert wären.

Wir erwähnen noch drei weitere Situationen, in denen man Folgerungen nur fast überall erhält.

2.15 Satz

(1) *Sei* $f\colon X \to [0,\infty]$ *messbar,* $E \in \mathscr{M}$*, und* $\int_E f\,d\mu = 0$*. Dann ist* $f = 0$ *fast überall auf* E*.*

(2) *Sei* $f \in L^1(\mu)$ *und* $\int_E f\,d\mu = 0$ *für alle* $E \in \mathscr{M}$*. Dann ist* $f = 0$ *fast überall auf* X*.*

(3) *Sei* $f \in L^1(\mu)$ *und* $\left|\int_X f\,d\mu\right| = \int_X |f|\,d\mu$*. Dann gibt es ein* $\alpha \in \{-1,1\}$*, so dass* $\alpha f = |f|$ *fast überall auf* X*.*

Satz 2.15(3) beschreibt den Gleichheitsfall in 2.9(3).

Beweis: (1): Für $A_n := \{x \in E : f(x) > 1/n\}$, $n = 1, 2, \ldots$, folgt

$$\frac{1}{n}\mu(A_n) \le \int_{A_n} f\,d\mu \le \int_E f\,d\mu = 0,$$

also $\mu(\{x \in E : f(x) > 0\}) = \mu(\bigcup_{n=1}^\infty A_n) = 0$.

(2): Wähle $E := \{x : f(x) \ge 0\}$; dann ist $\int_E f^+\,d\mu = \int_E f\,d\mu = 0$, also $f^+ = 0$ fast überall auf X nach (1). Ebenso verschwindet f^- fast überall.

(3): Untersuche den Beweis von 2.9(3). Mit der Voraussetzung folgt $\int_X |f| - \alpha f\,d\mu = 0$ für ein $\alpha \in \{-1,1\}$. Da $|f| - \alpha f \ge 0$, folgt mit (1) die Behauptung. $\square$

3 Konstruktion von Massen

Für die Konstruktion von Massen ist der folgende Begriff nützlich.

3.1 Definition (äusseres Mass, messbare Menge)

Sei X eine Menge. Ein *äusseres Mass* ν auf X ist eine Funktion $\nu\colon 2^X \to [0,\infty]$, so dass

(i) $\nu(\emptyset) = 0$,

(ii) $\nu(A) \le \nu(B)$ für $A \subset B \subset X$, und

(iii) $\nu(\bigcup_{i=1}^\infty A_i) \le \sum_{i=1}^\infty \nu(A_i)$ für $A_1, A_2, \ldots \subset X$.

Eine Menge $A \subset X$ heisst ν*-messbar*, falls

$$\nu(D) = \nu(D \cap A) + \nu(D \setminus A)$$

für jede Menge $D \subset X$.

Ein äusseres Mass ν ist also *monoton* (Eigenschaft (ii)) und σ*-subadditiv* (Eigenschaft (iii)), mit (i) also auch (*endlich*) *subadditiv*. Damit ist eine Menge $A \subset X$ bereits ν-messbar, wenn

$$\nu(D) \geq \nu(D \cap A) + \nu(D \setminus A)$$

für alle $D \subset X$ mit $\nu(D) < \infty$.

3.2 Satz (Carathéodory)
Sei X eine Menge und ν ein äusseres Mass auf X. Dann ist

$$\mathscr{M} := \{A \subset X : A \textit{ ist } \nu\textit{-messbar}\}$$

eine σ-Algebra, und die Einschränkung $\mu := \nu|_{\mathscr{M}}$ ist ein vollständiges Mass.

Beweis: Offensichtlich ist $X \in \mathscr{M}$, da $\nu(D) = \nu(D \cap X) + \nu(D \setminus X)$ für alle $D \subset X$. Mit $A \in \mathscr{M}$ ist auch $A^c \in \mathscr{M}$, da

$$\nu(D) = \nu(D \cap A) + \nu(D \setminus A) = \nu(D \setminus A^c) + \nu(D \cap A^c)$$

für alle $D \subset X$. Aus $A, B \in \mathscr{M}$ folgt $A \cup B \in \mathscr{M}$, da

$$\begin{aligned}
\nu(D) &= \nu(D \cap A) + \nu(D \setminus A) \\
&= \nu(D \cap A) + \nu((D \setminus A) \cap B) + \nu((D \setminus A) \setminus B) \\
&\geq \nu(D \cap (A \cup B)) + \nu(D \setminus (A \cup B))
\end{aligned}$$

für alle $D \subset X$. Somit ist $\mathscr{M}$ eine Algebra (s. 1.2).
Seien $A_1, A_2, \ldots \in \mathscr{M}$ paarweise disjunkt. Wir zeigen, dass $B := \bigcup_{i=1}^{\infty} A_i \in \mathscr{M}$ und $\nu(B) = \sum_{i=1}^{\infty} \nu(A_i)$; dann ist $\mathscr{M}$ eine σ-Algebra und $\nu|_{\mathscr{M}}$ ein Mass. Setze $B_k := \bigcup_{i=1}^{k} A_i$. Für alle $D \subset X$ und $k = 2, 3, \ldots$ gilt

$$\begin{aligned}
\nu(D \cap B_k) &= \nu(D \cap B_k \cap A_k) + \nu(D \cap B_k \setminus A_k) \\
&= \nu(D \cap A_k) + \nu(D \cap B_{k-1}),
\end{aligned}$$

also $\nu(D \cap B_k) = \sum_{i=1}^{k} \nu(D \cap A_i)$. Da $B_k \in \mathscr{M}$ und $B_k \subset B$, folgt

$$\nu(D) = \nu(D \cap B_k) + \nu(D \setminus B_k) \geq \sum_{i=1}^{k} \nu(D \cap A_i) + \nu(D \setminus B).$$

Da dies für alle k gilt und ν σ-subadditiv ist, folgt schliesslich

$$\nu(D) \geq \sum_{i=1}^{\infty} \nu(D \cap A_i) + \nu(D \setminus B) \geq \nu(D \cap B) + \nu(D \setminus B) \geq \nu(D),$$

d.h. $B \in \mathscr{M}$. Für $D = B$ ergibt die letzte Zeile gerade $\nu(B) = \sum_{i=1}^{\infty} \nu(A_i)$.
Das Mass $\mu = \nu|_{\mathscr{M}}$ ist vollständig: Jede Menge $A \subset X$ mit $\nu(A) = 0$ ist ν-messbar, da $\nu(D) \leq \nu(D \cap A) + \nu(D \setminus A) \leq \nu(D)$ für alle $D \subset X$. $\square$

Ist umgekehrt $(X, \mathscr{M}, \mu)$ ein Massraum, so wird durch

$$\nu(A) := \inf\{\mu(B) : A \subset B \in \mathscr{M}\}$$

ein äusseres Mass erklärt, und jede Menge $A \in \mathscr{M}$ ist ν-messbar (Übung).

Sei X eine Menge und $\mathscr{A} \subset 2^X$ eine Algebra (s. 1.2). Ein *Prämass* μ auf $\mathscr{A}$ ist eine Funktion $\mu\colon \mathscr{A} \to [0, \infty]$, $\mu \not\equiv \infty$, die in folgendem Sinn *σ-additiv* ist: Sind $A_1, A_2, \ldots \in \mathscr{A}$ paarweise disjunkt, *und ist auch* $\bigcup_{i=1}^\infty A_i \in \mathscr{A}$, so gilt

$$\mu\big(\textstyle\bigcup_{i=1}^\infty A_i\big) = \displaystyle\sum_{i=1}^\infty \mu(A_i).$$

Jedes Prämass $\mu\colon \mathscr{A} \to [0, \infty]$ erfüllt $\mu(\emptyset) = 0$ und ist (endlich) additiv und monoton (dies folgt wie in 1.14). Ein Prämass μ auf einer Algebra $\mathscr{A}$ heisst *σ-endlich*, falls eine Folge von Mengen $E_1, E_2, \ldots \in \mathscr{A}$ mit $\mu(E_i) < \infty$ und $\bigcup_{i=1}^\infty E_i = X$ existiert.

3.3 Satz (Erweiterungssatz, Carathéodory-Hahn)
Sei X eine Menge, $\mathscr{A} \subset 2^X$ eine Algebra und μ ein Prämass auf $\mathscr{A}$.

(1) *Definiere $\nu\colon 2^X \to [0, \infty]$ durch*

$$\nu(A) := \inf\big\{\textstyle\sum_{k=1}^\infty \mu(B_k) : B_k \in \mathscr{A},\ A \subset \bigcup_{k=1}^\infty B_k\big\};$$

ν ist ein äusseres Mass.

(2) *Sei $\mathscr{M}$ die σ-Algebra der ν-messbaren Mengen (s. 3.2). Das Mass $\nu|_{\mathscr{M}}$ ist eine Erweiterung von μ, d.h. $\mathscr{A} \subset \mathscr{M}$ und $\nu|_{\mathscr{A}} = \mu$.*

(3) *Ist zusätzlich μ σ-endlich, so ist die in (2) angegebene Erweiterung eindeutig in folgendem Sinn: Ist $\mathscr{M}'$ eine σ-Algebra mit $\mathscr{A} \subset \mathscr{M}' \subset \mathscr{M}$ und ist μ' ein Mass auf $\mathscr{M}'$ mit $\mu'|_{\mathscr{A}} = \mu$, so gilt $\mu' = \nu|_{\mathscr{M}'}$.*

Beweis: (1): Offensichtlich ist ν monoton und $\nu(\emptyset) = 0$. Seien $A_1, A_2, \ldots \subset X$, $A := \bigcup_{i=1}^\infty A_i$. Zu $\epsilon > 0$ wähle $B_{ik} \in \mathscr{A}$, so dass $A_i \subset \bigcup_{k=1}^\infty B_{ik}$ und $\sum_{k=1}^\infty \mu(B_{ik}) \le \nu(A_i) + 2^{-i}\epsilon$; dann folgt

$$\nu(A) \le \sum_{i,k=1}^\infty \mu(B_{ik}) \le \sum_{i=1}^\infty \nu(A_i) + \epsilon.$$

Somit ist ν σ-subadditiv.

(2): Sei $A \in \mathscr{A}$. Es gilt $\nu(A) \le \mu(A)$. Ist $A \subset \bigcup_{k=1}^\infty B_k$ für paarweise disjunkte $B_k \in \mathscr{A}$, setze $A_k := A \cap B_k$. Dann sind die A_k in $\mathscr{A}$ und paarweise disjunkt, und $A = \bigcup_{k=1}^\infty A_k$. Somit gilt $\mu(A) = \sum_{k=1}^\infty \mu(A_k) \le \sum_{k=1}^\infty \mu(B_k)$. Dies impliziert $\mu(A) \le \nu(A)$. Damit ist $\nu|_{\mathscr{A}} = \mu$.

Es bleibt noch die ν-Messbarkeit von $A \in \mathscr{A}$ zu zeigen. Sei $D \subset X$ mit $\nu(D) < \infty$. Zu $\epsilon > 0$ wähle $B_k \in \mathscr{A}$ mit $D \subset \bigcup_{k=1}^\infty B_k$ und $\sum_{k=1}^\infty \mu(B_k) \leq \nu(D) + \epsilon$. Es gilt $\mu(B_k \cap A) + \mu(B_k \setminus A) = \mu(B_k)$, also

$$\sum_{k=1}^\infty \mu(B_k \cap A) + \sum_{k=1}^\infty \mu(B_k \setminus A) \leq \nu(D) + \epsilon.$$

Da $D \cap A \subset \bigcup_{k=1}^\infty (B_k \cap A)$ und $D \setminus A \subset \bigcup_{k=1}^\infty (B_k \setminus A)$, folgt $\nu(D \cap A) + \nu(D \setminus A) \leq \nu(D) + \epsilon$.

(3): Sei μ' ein Mass auf der σ-Algebra $\mathscr{M}'$, wobei $\mathscr{A} \subset \mathscr{M}' \subset \mathscr{M}$ und $\mu'|_{\mathscr{A}} = \mu$. Sei $A \in \mathscr{M}'$. Für $B_k \in \mathscr{A}$ mit $A \subset \bigcup_{k=1}^\infty B_k$ gilt $\mu'(A) \leq \sum_{k=1}^\infty \mu'(B_k) = \sum_{k=1}^\infty \mu(B_k)$. Dies zeigt, dass $\mu' \leq \nu$ auf $\mathscr{M}'$. Wähle paarweise disjunkte $E_i \in \mathscr{A}$ mit $\mu(E_i) < \infty$ und $X = \bigcup_{i=1}^\infty E_i$. Dann sind $E_i \cap A, E_i \setminus A \in \mathscr{M}'$, also gilt $\mu'(E_i \cap A) \leq \nu(E_i \cap A)$ und $\mu'(E_i \setminus A) \leq \nu(E_i \setminus A)$. Mit

$$\mu'(E_i \cap A) + \mu'(E_i \setminus A) = \mu'(E_i) = \mu(E_i) = \nu(E_i) = \nu(E_i \cap A) + \nu(E_i \setminus A)$$

und $\mu(E_i) < \infty$ folgt $\mu'(E_i \cap A) = \nu(E_i \cap A)$. Summation über i gibt $\mu'(A) = \nu(A)$. $\square$

Wir konstruieren jetzt das Lebesgue-Mass auf $\mathbb{R}^n$ durch Erweiterung des Elementarinhalts von endlichen Vereinigungen von Quadern. Ein *Intervall* $I \subset \mathbb{R}$ ist eine zusammenhängende Menge, d.h. eine Menge von einem der Typen

$$\begin{array}{ll} (a,b), & -\infty \leq a \leq b \leq \infty, \\ [a,b), & -\infty < a \leq b \leq \infty, \\ (a,b], & -\infty \leq a \leq b < \infty, \\ [a,b], & -\infty < a \leq b < \infty. \end{array}$$

Dann ist $L(I) := b - a \in [0, \infty]$ die *Länge* von I. Ein *Quader* $Q \subset \mathbb{R}^n$ ist eine Menge der Gestalt $Q = I_1 \times I_2 \times \ldots \times I_n$ für Intervalle $I_1, \ldots, I_n$. Dann ist $V(Q) := L(I_1) \cdot L(I_2) \cdot \ldots \cdot L(I_n)$ das *Volumen* von Q (wobei wieder $0 \cdot \infty = 0$).

3.4 Proposition
Sei $\mathscr{A} := \{A \subset \mathbb{R}^n : A$ ist Vereinigung endlich vieler Quader$\}$; $\mathscr{A}$ ist eine Algebra. Jedes $A \in \mathscr{A}$ lässt sich als Vereinigung von endlich vielen, paarweise disjunkten Quadern schreiben. Ist $A = \bigcup_{i=1}^k Q_i$ eine solche disjunkte Vereinigung, setze $\mu(A) := \sum_{i=1}^k V(Q_i)$. Dies definiert ein Prämass μ auf $\mathscr{A}$.

Beweis: Man sieht leicht, dass $\mathscr{A}$ eine Algebra ist und jedes $A \in \mathscr{A}$ als Vereinigung endlich vieler, paarweise disjunkter Quader geschrieben werden kann (Übung).

Um zu zeigen, dass μ wohldefiniert ist, betrachtet man für zwei verschiedene disjunkte Darstellungen $A = \bigcup_{i=1}^k Q_i = \bigcup_{j=1}^l Q'_j$ die aus allen $Q_i \cap Q'_j$ bestehende gemeinsame Verfeinerung; es gilt

$$\sum_{i=1}^k V(Q_i) = \sum_{i=1}^k \sum_{j=1}^l V(Q_i \cap Q'_j) = \sum_{j=1}^l V(Q'_j).$$

Offensichtlich ist $\mu(\emptyset) = 0$. Es bleibt noch die σ-Additivität von μ zu zeigen. Seien $A_1, A_2, \ldots \in \mathscr{A}$ paarweise disjunkt, und sei $A := \bigcup_{i=1}^\infty A_i$ in $\mathscr{A}$. Da μ offenbar (endlich) additiv und daher monoton ist, gilt

$$\mu(A) \geq \mu\big(\textstyle\bigcup_{i=1}^k A_i\big) = \displaystyle\sum_{i=1}^k \mu(A_i)$$

für alle k, also $\mu(A) \geq \sum_{i=1}^\infty \mu(A_i)$. Für den Beweis der umgekehrten Ungleichung nehmen wir o.B.d.A. an, dass alle A_i beschränkte Quader sind und dass $A \in \mathscr{A}$ abgeschlossen ist. Zu $\epsilon > 0$ wähle offene Quader $U_i \supset A_i$ mit $\mu(U_i) \leq \mu(A_i) + 2^{-i}\epsilon$. Für $l > 0$ betrachte $A^l := A \cap [-l, l]^n$. Da A^l kompakt ist und da $A^l \subset \bigcup_{i=1}^\infty U_i$, gilt $A^l \subset \bigcup_{i=1}^k U_i$ für ein $k = k(l)$. Es folgt

$$\mu(A^l) \leq \sum_{i=1}^k \mu(U_i) \leq \sum_{i=1}^k \mu(A_i) + \epsilon \leq \sum_{i=1}^\infty \mu(A_i) + \epsilon.$$

Da $\mu(A^l) \to \mu(A)$ $(l \to \infty)$, folgt die noch fehlende Ungleichung. $\square$

3.5 Definition (Lebesgue-Mass)
Seien $\mathscr{A}$ und μ wie in 3.4. Das wie in 3.3(1) erklärte zugehörige äussere Mass ν heisst *äusseres Lebesgue-Mass* auf $\mathbb{R}^n$. Die wieder mit μ bezeichnete Erweiterung $\nu|_{\mathscr{M}}$ von μ gemäss 3.3(2) heisst *Lebesgue-Mass* auf $\mathbb{R}^n$; $\mathscr{M}$ ist die σ-Algebra der *Lebesgue-messbaren* Teilmengen von $\mathbb{R}^n$.

Man beachte, dass das äussere Lebesgue-Mass ν einfach durch

$$\nu(A) = \inf\big\{\textstyle\sum_{i=1}^\infty V(Q_i) : Q_i \text{ kompakter Quader, } A \subset \bigcup_{i=1}^\infty Q_i\big\}$$

charakterisiert ist. Jede offene Menge ist abzählbare Vereinigung von Quadern und daher Lebesgue-messbar. Es folgt, dass jede Borelmenge Lebesgue-messbar ist. Dies ergibt sich auch aus dem folgenden allgemeinen Satz.

3.6 Satz (Carathéodory's Kriterium)
Sei ν ein äusseres Mass auf einem metrischen Raum $X = (X, d)$. Dann sind äquivalent:

(1) *Alle Borelmengen sind ν-messbar.*
(2) $\nu(A)+\nu(B) = \nu(A\cup B)$ *für alle* $A, B \subset X$ *mit* $d(A, B) := \inf\{d(x, y) : x \in A,\ y \in B\} > 0$.

Beweis: Aus (1) folgt (2): Seien $A, B \subset X$ mit $d(A, B) > 0$. Dann existiert eine offene Menge $U \supset A$ mit $U \cap B = \emptyset$. Da U ν-messbar ist, folgt

$$\nu(A \cup B) = \nu((A \cup B) \cap U) + \nu((A \cup B) \setminus U) = \nu(A) + \nu(B).$$

(2) impliziert (1): Sei $\mathscr{M}$ die σ-Algebra der ν-messbaren Mengen. Es genügt zu zeigen, dass

$$\nu(D) \geq \nu(D \cap A) + \nu(D \setminus A)$$

falls $D \subset X$, $\nu(D) < \infty$ und $A \subset X$ abgeschlossen; dann ist $A \in \mathscr{M}$, also $\mathscr{B} \subset \mathscr{M}$. Für $i = 1, 2, \ldots$ sei $A_i := \{x \in X : d(x, A) < 1/i\}$. Es gilt $d(D \cap A, D \setminus A_i) > 0$ und somit

$$\nu(D) \geq \nu((D \cap A) \cup (D \setminus A_i)) = \nu(D \cap A) + \nu(D \setminus A_i)$$

gemäss (2). Es bleibt noch zu zeigen, dass $\nu(D\setminus A_i) \to \nu(D\setminus A)$ $(i \to \infty)$. Da A abgeschlossen ist, folgt $D\setminus A = (D\setminus A_i)\cup\bigcup_{j=i}^{\infty} E_j$ für $E_j := D\cap(A_j\setminus A_{j+1})$. Es gilt

$$\nu(D \setminus A_i) \leq \nu(D \setminus A) \leq \nu(D \setminus A_i) + \sum_{j=i}^{\infty} \nu(E_j);$$

es genügt zu zeigen, dass $\sum_{j=1}^{\infty} \nu(E_j) < \infty$. Für $k \geq j + 2$ ist $d(E_j, E_k) > 0$. Mit (2) folgt für alle $n \in \mathbb{N}$, dass

$$\sum_{l=1}^{n} \nu(E_{2l-1}) = \nu\big(\textstyle\bigcup_{l=1}^{n} E_{2l-1}\big) \leq \nu(D)$$

und ebenso $\sum_{l=1}^{n} \nu(E_{2l}) \leq \nu(D)$. Da $\nu(D) < \infty$, folgt der Satz. □

Eigenschaft (2) gilt offensichtlich für das äussere Lebesgue-Mass: In der Charakterisierung von ν können o.B.d.A. alle Quader mit Durchmesser $\leq \delta$ gewählt werden für ein beliebiges $\delta > 0$. Für $A, B \subset X$ mit $d(A, B) > 0$ wähle man $\delta < d(A, B)/2$.

3.7 Definition (Borel-reguläres äusseres Mass)
Ein äusseres Mass ν auf einem topologischen Raum X heisst *Borel-regulär*, falls jede Borelmenge ν-messbar ist und jede Menge $A \subset X$ in einer Borelmenge B mit $\nu(A) = \nu(B)$ enthalten ist.

3.8 Satz
Sei ν das äussere Lebesgue-Mass auf $\mathbb{R}^n$. Für jede Menge $A \subset \mathbb{R}^n$ gilt

$$\nu(A) = \inf\{\nu(U) : A \subset U,\, U \subset \mathbb{R}^n \text{ offen}\},$$

insbesondere existiert eine G_δ-Menge B, so dass $A \subset B$ und $\nu(A) = \nu(B)$. Somit ist ν Borel-regulär.

Beweis: Sei $A \subset \mathbb{R}^n$. Zu $\epsilon > 0$ wähle kompakte Quader Q_i mit $A \subset \bigcup_{i=1}^\infty Q_i$ und $\sum_{i=1}^\infty V(Q_i) \le \nu(A) + \epsilon$. Wähle dann offene Quader $U_i \supset Q_i$, so dass $V(U_i) \le V(Q_i) + 2^{-i}\epsilon$. Dann ist $U := \bigcup_{i=1}^\infty U_i$ offen, $A \subset U$, und

$$\nu(U) \le \sum_{i=1}^\infty \nu(U_i) = \sum_{i=1}^\infty V(U_i) \le \sum_{i=1}^\infty V(Q_i) + \epsilon \le \nu(A) + 2\epsilon.$$

Dies zeigt die erste Behauptung.
Um eine G_δ-Menge B mit $A \subset B$ und $\nu(A) = \nu(B)$ zu finden, wähle offene Mengen U_k, so dass $A \subset U_k$ und $\nu(U_k) \to \nu(A)$ $(k \to \infty)$. Setze $B := \bigcap_{k=1}^\infty U_k$. Für alle k gilt $A \subset B \subset U_k$, also $\nu(A) \le \nu(B) \le \nu(U_k) \to \nu(A)$, d.h. $\nu(A) = \nu(B)$. □

3.9 Satz (Einschränkung äusserer Masse)
Sei ν ein äusseres Mass auf einem topologischen Raum X, $A \subset X$. Definiere $\nu \llcorner A \colon 2^X \to [0,\infty]$ durch $(\nu \llcorner A)(B) := \nu(A \cap B)$.

(1) *$\nu \llcorner A$ ist ein äusseres Mass, und jede ν-messbare Menge ist auch $(\nu \llcorner A)$-messbar.*
(2) *Ist ν Borel-regulär und A ν-messbar mit $\nu(A) < \infty$, so ist $\nu \llcorner A$ Borel-regulär.*

Beweis: (1): Es ist klar, dass $\nu \llcorner A$ ein äusseres Mass ist. Ist $C \subset X$ ν-messbar, so gilt für jedes $D \subset X$

$$\begin{aligned}(\nu \llcorner A)(D) = \nu(A \cap D) &= \nu((A \cap D) \cap C) + \nu((A \cap D) \setminus C) \\ &= \nu(A \cap (D \cap C)) + \nu(A \cap (D \setminus C)) \\ &= (\nu \llcorner A)(D \cap C) + (\nu \llcorner A)(D \setminus C);\end{aligned}$$

C ist also auch $(\nu \llcorner A)$-messbar.

(2): Sei A ν-messbar mit $\nu(A) < \infty$. Wähle eine Borelmenge B mit $A \subset B$ und $\nu(A) = \nu(B)$. Da A ν-messbar und $\nu(A) < \infty$ ist, folgt

$$\nu(B \setminus A) = \nu(B) - \nu(B \cap A) = \nu(A) - \nu(A) = 0.$$

Sei jetzt $C \subset X$; wir zeigen, dass C in einer Borelmenge D mit $(\nu \llcorner A)(D) = (\nu \llcorner A)(C)$ enthalten ist. Wähle eine Borelmenge E, so dass $B \cap C \subset E$ und $\nu(B \cap C) = \nu(E)$. Dann ist $D := E \cup B^c$ eine Borelmenge, $C \subset (B \cap C) \cup B^c \subset D$ und $A \cap D \subset B \cap D \subset E$. Es folgt

$$\nu(A \cap D) \leq \nu(E) = \nu(B \cap C) \leq \nu(A \cap C) + \nu((B \setminus A) \cap C) = \nu(A \cap C)$$

und somit $(\nu \llcorner A)(D) = (\nu \llcorner A)(C)$. □

3.10 Satz (Approximationseigenschaften)
Sei ν ein Borel-reguläres äusseres Mass auf einem metrischen Raum X und $A \subset X$ eine ν-messbare Menge.

(1) *Ist $\nu(A) < \infty$, so existiert für jedes $\epsilon > 0$ eine abgeschlossene Menge $C \subset A$ mit $\nu(A \setminus C) < \epsilon$, insbesondere gibt es eine F_σ-Menge $F \subset A$ mit $\nu(A \setminus F) = 0$.*

(2) *Ist $A \subset \bigcup_{i=1}^\infty V_i$ für offene Mengen $V_i \subset X$ mit $\nu(V_i) < \infty$, so existiert für jedes $\epsilon > 0$ eine offene Menge $U \supset A$ mit $\nu(U \setminus A) < \epsilon$, insbesondere gibt es eine G_δ-Menge $G \supset A$ mit $\nu(G \setminus A) = 0$.*

Man beachte, dass die Folgerung von (2) nicht allgemein unter der Voraussetzung $\nu(A) < \infty$ in (1) gilt: Ist ν das Zählmass auf $\mathbb{R}$ und $A \subset X$ eine endliche Menge, so ist $\nu(U \setminus A) = \infty$ für jede offene Umgebung U von A. Aus (1) und der Borel-Regularität von ν folgt, dass für jede ν-messbare Menge $A \subset X$ mit $\nu(A) < \infty$ eine F_σ-Menge F und eine Borelmenge G existieren, so dass $F \subset A \subset G$ und

$$\nu(G \setminus F) = \nu(G \setminus A) + \nu(A \setminus F) = 0.$$

Umgekehrt gilt: Ist ν ein äusseres Mass auf einer Menge X, $A \subset X$, und existieren ν-messbare Mengen $F, G \subset X$ mit $F \subset A \subset G$ und $\nu(G \setminus F) = 0$, so ist A ν-messbar. Dies entspricht gerade der Vollständigkeit des Masses $\mu = \nu|_{\mathscr{M}}$ (s. 3.2): Da $\mu(G \setminus F) = 0$ und $A \setminus F \subset G \setminus F$, ist $A \setminus F \in \mathscr{M}$, also $A = F \cup (A \setminus F) \in \mathscr{M}$.

Beweis von 3.10: (1): Nach 3.9 ist $\nu_A := \nu \llcorner A$ ein (endliches) Borel-reguläres äusseres Mass. Sei $\mathscr{A}$ die Menge aller $A' \subset X$, so dass für jedes $\epsilon > 0$ eine abgeschlossene Menge C und eine offene Menge U mit $C \subset A' \subset U$ und

$\nu_A(U \setminus C) < \epsilon$ existieren. Wir zeigen zunächst, dass $\mathscr{A}$ eine σ-Algebra ist. Offensichtlich ist $X \in \mathscr{A}$, da X abgeschlossen und offen ist. Ist $A' \in \mathscr{A}$, und gilt $\nu_A(U \setminus C) < \epsilon$ für Mengen $C \subset A' \subset U$ wie oben, so folgt $U^c \subset (A')^c \subset C^c$ und $\nu_A(C^c \setminus U^c) = \nu_A(U \setminus C) < \epsilon$; somit ist $(A')^c \in \mathscr{A}$. Sind $A_1, A_2, \ldots \in \mathscr{A}$, und gilt $\nu_A(U_i \setminus C_i) < 2^{-i}\epsilon$ für abgeschlossene bzw. offene Mengen $C_i \subset A_i \subset U_i$, so ist $C^k := \bigcup_{i=1}^k C_i$ für $k \in \mathbb{N}$ abgeschlossen, $U := \bigcup_{i=1}^\infty U_i$ ist offen, und es gilt $C^k \subset \bigcup_{i=1}^\infty A_i \subset U$. Da ausserdem $\bigcap_{k=1}^\infty (U \setminus C^k) = U \setminus \bigcup_{i=1}^\infty C_i$, folgt mit 3.2 und 1.14(5) (hier wird $\nu_A < \infty$ verwendet)

$$\lim_{k\to\infty} \nu_A(U \setminus C^k) = \nu_A\big(U \setminus \textstyle\bigcup_{i=1}^\infty C_i\big) \leq \displaystyle\sum_{i=1}^\infty \nu_A(U_i \setminus C_i) < \epsilon;$$

somit ist $\bigcup_{i=1}^\infty A_i \in \mathscr{A}$. Die σ-Algebra $\mathscr{A}$ enthält alle abgeschlossenen Mengen in X, da jede solche Menge gleich dem Durchschnitt ihrer offenen $1/k$-Umgebungen ist für $k \in \mathbb{N}$. Somit umfasst $\mathscr{A}$ alle Borelmengen. Nun existiert eine Borelmenge B mit $A \subset B$ und $\nu(B \setminus A) = 0$, weiter ist $B \setminus A$ in einer Borelmenge D mit $\nu(D) = 0$ enthalten. Für $E := B \setminus D$ gilt dann $E \subset A$ und $\nu(A \setminus E) = 0$. Da $E \in \mathscr{A}$, folgt die erste Aussage in (1) für A.

Um eine F_σ-Menge $F \subset A$ mit $\nu(A \setminus F) = 0$ zu finden, wähle abgeschlossene Mengen $C_j \subset A$ mit $\nu(A \setminus C_j) \to 0$. Setze $F := \bigcup_{j=1}^\infty C_j$. Für alle j gilt $\nu(A \setminus F) \leq \nu(A \setminus C_j)$, also $\nu(A \setminus F) = 0$.

(2): Nach (1) existiert für jedes i eine abgeschlossene Menge $C_i \subset V_i \setminus A$, so dass $\nu(V_i \setminus A \setminus C_i) < 2^{-i}\epsilon$. Dann ist $A \subset \bigcup_{i=1}^\infty (V_i \setminus C_i) =: U$ und $\nu(U \setminus A) \leq \sum_{i=1}^\infty \nu(V_i \setminus C_i \setminus A) < \epsilon$.

Um eine G_δ-Menge $G \supset A$ mit $\nu(G \setminus A) = 0$ zu finden, wähle offene Mengen $U_j \supset A$ mit $\nu(U_j \setminus A) \to 0$. Setze $G := \bigcap_{j=1}^\infty U_j$. Dann folgt $\nu(G \setminus A) \leq \nu(U_j \setminus A) \to 0$. □

Wir stellen die Eigenschaften des Lebesgue-Masses auf $\mathbb{R}^n$ zusammen.

3.11 Satz (Eigenschaften des Lebesgue-Masses)
Seien $\mathscr{A}$, ν, $\mu = \nu|_{\mathscr{M}}$ wie in 3.5.

(1) *μ ist ein vollständiges Mass mit $\mu(Q) = V(Q)$ für jeden Quader $Q \subset \mathbb{R}^n$.*

(2) *Es gilt $\mathscr{A} \subset \mathscr{B} \subset \mathscr{M}$, ebenso $\mathscr{B} \subset \mathscr{M}'$ für jede σ-Algebra $\mathscr{M}' \supset \mathscr{A}$. Für alle $A \in \mathscr{M}$ gilt*

$$\mu(A) = \sup\{\mu(C) : C \subset A,\ C \text{ kompakt}\},$$
$$\mu(A) = \inf\{\mu(U) : A \subset U,\ U \text{ offen}\},$$

und es existieren eine F_σ-Menge $F \subset A$ und eine G_δ-Menge $G \supset A$, so dass $\mu(G \setminus F) = 0$.

(3) *μ ist translationsinvariant, d.h. für alle $A \in \mathscr{M}$ und $x \in \mathbb{R}^n$ ist $x+A \in \mathscr{M}$ und $\mu(x+A) = \mu(A)$. Ebenso ist $\nu(x+A) = \nu(A)$ für alle $A \subset X$ und $x \in \mathbb{R}^n$.*

(4) *Ist μ' ein translationsinvariantes Mass auf einer σ-Algebra $\mathscr{M}'$ mit $\mathscr{A} \subset \mathscr{M}' \subset \mathscr{M}$, und ist μ' endlich auf kompakten Mengen, so existiert eine Konstante c, so dass $\mu'(A) = c\mu(A)$ für alle $A \in \mathscr{M}'$. Ist ν' ein translationsinvariantes, Borel-reguläres äusseres Mass auf $\mathbb{R}^n$, und ist ν' endlich auf kompakten Mengen, so gilt $\nu' = c\nu$ für eine Konstante c.*

(5) *Für jede lineare Abbildung $T\colon \mathbb{R}^n \to \mathbb{R}^n$ existiert ein $\triangle(T) \in \mathbb{R}$, so dass $T(A) \in \mathscr{M}$ und $\mu(T(A)) = \triangle(T)\mu(A)$ für alle $A \in \mathscr{M}$. Insbesondere gilt $\mu(T(A)) = \mu(A)$, falls T eine Drehung ist, d.h. $T \in O(n)$.*

Eigenschaften (3) und (5) zeigen, dass μ invariant ist unter Bewegungen, d.h. $\mu(x + T(A)) = \mu(A)$ für alle $T \in O(n)$ und $x \in \mathbb{R}^n$.

Beweis: (1) ist klar nach Konstruktion (s. 3.2, 3.3(2)).

(2): Die Identität $\mu(A) = \inf\{\mu(U) : A \subset U,\, U \text{ offen}\}$ und die Existenz einer G_δ-Menge $G \supset A$ mit $\mu(G \setminus A) = 0$ folgen aus (3.8 und) 3.10(2). Aus 3.10(1) folgt ausserdem, dass für $k = 1, 2, \ldots$ kompakte Mengen $C_k \subset A_k := A \cap [-k,k]^n$ mit $\mu(A_k \setminus C_k) < 1/k$ existieren. Da $\mu(A_k) \to \mu(A)$ (1.14(4)), gilt $\mu(C_k) \to \mu(A)$. Weiter existieren F_σ-Mengen $F_k \subset A_k$ mit $\mu(A_k \setminus F_k) = 0$. Für $F = \bigcup_{k=1}^\infty F_k$ folgt $\mu(A \setminus F) \le \sum_{k=1}^\infty \mu(A_k \setminus F_k) = 0$.

(3) ist klar, da $\mu|_{\mathscr{A}}$ translationsinvariant ist.

(4): Sei $c := \mu'([0,1)^n)$. Mit der Translationsinvarianz von μ' folgt leicht, dass jeder Quader der Gestalt $Q = Q(k,x) = x + [0, 2^{-k})^n$ mit $x \in \mathbb{R}^n$ und $k \in \mathbb{N} \cup \{0\}$ Mass $\mu'(Q) = c\mu(Q)$ hat. (Zerlege $[0,1)^n$ in 2^{kn} Quader dieser Gestalt für ein festes k.) Jede offene Menge $U \subset \mathbb{R}^n$ lässt sich als Vereinigung abzählbar vieler, paarweise disjunkter Quader dieser Gestalt darstellen. (Nenne $Q(k,x)$ *diadisch* falls $x \in (2^{-k}\mathbb{Z})^n$. Nimm alle diadischen $Q(k,x) \subset U$, die nicht in einem grösseren diadischen $Q(k',x') \subset U$ enthalten sind.) Damit folgt $\mu'(A) = c\mu(A)$ für alle $A \in \mathscr{A}$; zuerst für alle offenen Quader, dann mittels 1.14(5) für alle beschränkten Quader, dann für alle $A \in \mathscr{A}$. Da μ σ-endlich ist, folgt mit der Eindeutigkeitsaussage aus 3.3(3), angewandt auf $(1/c)\mu'$, dass $\mu' = c\mu|_{\mathscr{M}'}$.
Für ν' folgt ganz analog, dass $\nu'(U) = c\nu(U)$ für jede offene Menge U. Mit 3.10(2) und der Borel-Regularität folgt $\nu' = c\nu$.

(5): Ist $\dim(T(\mathbb{R}^n)) < n$, so ist $\mu(T(\mathbb{R}^n)) = 0$. Somit gilt die Behauptung für $\triangle(T) = 0$. Sei jetzt T regulär. Dann ist T ein Homöomorphismus, T bildet also Borelmengen auf Borelmengen ab. Definiere μ' auf $\mathscr{B}$ durch $\mu'(B) :=$

$\mu(T(B))$. Für $x \in \mathbb{R}^n$ und $B \in \mathscr{B}$ gilt

$$\mu'(x+B) = \mu(T(x+B)) = \mu(Tx + T(B)) = \mu(T(B)) = \mu'(B);$$

μ' ist also translationsinvariant. Da $\mathscr{A} \subset \mathscr{B} \subset \mathscr{M}$, folgt aus (4), dass $\mu(T(B)) = \mu'(B) = \triangle(T)\mu(B)$ für eine Konstante $\triangle(T)$. Für $A \in \mathscr{M}$ wähle $F, G \in \mathscr{B}$, so dass $F \subset A \subset G$ und $\mu(G \setminus F) = 0$, s. (2). Dann ist $T(F) \subset T(A) \subset T(G)$ und

$$\mu(T(G) \setminus T(F)) = \mu(T(G \setminus F)) = \triangle(T)\mu(G \setminus F) = 0.$$

Da μ vollständig ist, folgt $T(A) \in \mathscr{M}$ und

$$\mu(T(A)) = \mu(T(G)) = \triangle(T)\mu(G) = \triangle(T)\mu(A).$$

Ist $T \in O(n)$, wähle $B = \{x \in \mathbb{R}^n : |x| < 1\}$; dann ist $T(B) = B$, also $\mu(T(B)) = \mu(B)$, d.h. $\triangle(T) = 1$. $\square$

Ist $(\mathbb{R}^n, \mathscr{M}, \mu)$ der Lebesgue-Massraum, so schreiben wir $L^1(\mathbb{R}^n)$ anstelle von $L^1(\mu)$. Ist $E \subset \mathbb{R}^n$ Lebesgue-messbar und $(E, \mathscr{M}_E, \mu_E)$ der entsprechende Massraum ($\mathscr{M}_e = \{A \in \mathscr{M} : A \subset E\}$, $\mu_e = \mu|\mathscr{M}_e$), so schreiben wir $L^1(E)$ statt $L^1(\mu_E)$ und $\int_E f(x)\,dx$ statt $\int_E f\,d\mu_e$, im Fall $n = 1$ und $E = [a,b]$ (oder $E = (a,b), [a,b), (a,b]$) wie üblich $\int_a^b f(x)\,dx$. Ist f beschränkt und Riemann-integrabel über $[a,b]$, dann ist $f \in L^1([a,b])$ und das Riemann-Integral R-$\int_a^b f(x)\,dx$ stimmt mit dem Lebesgue-Integral $\int_a^b f(x)\,dx$ überein. (Dies folgt leicht mit Satz 2.4 (monotone Konvergenz).)

Wir haben noch keine Antwort auf folgende Fragen: Ist jede Lebesgue-messbare Menge eine Borelmenge? Ist jede Teilmenge von $\mathbb{R}^n$ Lebesgue-messbar? Die Antwort ist in beiden Fällen negativ, sogar für $n = 1$. Die erste Frage lässt sich mit einer Kardinalitätsüberlegung beantworten. Sei c die Kardinalität von $\mathbb{R}$ (äquivalent: von $2^{\mathbb{N}}$). Die Topologie von $\mathbb{R}^n$ hat eine abzählbare Basis und erzeugt $\mathscr{B}$; man kann daraus schliessen, dass $\mathscr{B}$ die Kardinalität c hat. Andererseits existiert eine Cantormenge $E \subset \mathbb{R}$ mit $\mu(E) = 0$. Da μ vollständig ist, ist jede der 2^c Teilmengen von E Lebesgue-messbar. Da $2^c > c$, sind die meisten Teilmengen von E keine Borelmengen. Der folgende Satz beantwortet die zweite Frage.

3.12 Satz

Ist $A \subset \mathbb{R}$ und ist jede Teilmenge von A Lebesgue-messbar, so ist A eine Lebesgue-Nullmenge. Jede Menge mit positivem Lebesgue-Mass enthält daher nicht-messbare Teilmengen.

Beweis: Betrachte die folgende Äquivalenzrelation auf $\mathbb{R}$: $r \sim s \Leftrightarrow r - s \in \mathbb{Q}$. Wähle $E \subset \mathbb{R}$ so, dass E genau einen Punkt aus jeder Äquivalenzklasse enthält (Auswahlaxiom). Dann gilt:

(1) $(r+E) \cap (s+E) = \emptyset$ für $r, s \in \mathbb{Q}$, $r \neq s$.
(2) Jedes $x \in \mathbb{R}$ liegt in $r + E$ für ein $r \in \mathbb{Q}$.

(Beweis von (1): Sei $x \in (r+E) \cap (s+E)$; dann ist $x = r + y = s + z$ für $y, z \in E$, somit $y - z = s - r \in \mathbb{Q} \setminus \{0\}$, im Widerspruch zur Wahl von E. Beweis von (2): Für jedes $x \in \mathbb{R}$ gibt es ein eindeutig bestimmtes $y \in E$, so dass $x \sim y$. Dann ist $r := x - y \in \mathbb{Q}$ und $x = r + y \in r + E$.) Fixiere im Moment $t \in \mathbb{Q}$ und setze $A_t := A \cap (t + E)$; A_t ist messbar nach Voraussetzung. Sei $C \subset A_t$ kompakt, und sei $D := \bigcup_{r \in \mathbb{Q} \cap [0,1]} (r + C)$. D ist beschränkt, also ist $\mu(D) < \infty$. Da $C \subset t + E$, zeigt (1), dass die $r + C$ paarweise disjunkt sind. Somit gilt $\mu(D) = \sum_r \mu(r + C)$. Da $\mu(r + C) = \mu(C)$, folgt $\mu(C) = 0$. Dies gilt für alle kompakten $C \subset A_t$, also ist $\mu(A_t) = 0$ wegen 3.11(2). Wegen (2) ist ausserdem $A = \bigcup_{t \in \mathbb{Q}} A_t$, somit $\mu(A) = 0$. □

3.13 Bemerkung (Nachtrag zu 3.11)
Für die Konstante $\triangle(T) \geq 0$ in 3.11(5) folgt $\triangle(T) = |\det T|$. Jede lineare Abbildung $T \colon \mathbb{R}^n \to \mathbb{R}^n$ lässt sich als Verkettung endlich vieler linearer Abbildungen der folgenden drei Typen schreiben:

(I) T permutiert die Standard-Basisvektoren $e_1, \ldots, e_n$.
(II) $Te_1 = \alpha e_1$, $Te_k = e_k$ für $k = 2, \ldots, n$.
(III) $Te_1 = e_1 + e_2$, $Te_k = e_k$ für $k = 2, \ldots, n$.

Wegen dem Multiplikationssatz für Determinanten genügt es, $\triangle(T) = |\det T|$ für Abbildungen T der Typen (I), (II), (III) einzeln zu zeigen. Für Typ (I) und (II) ist dies leicht. Ist T vom Typ (III), so ist $\det(T) = 1$ und für $Q := [0,1)^n$ ist

$$T(Q) = \{(\xi_1, \ldots, \xi_n) : \xi_1 \leq \xi_2 < \xi_1 + 1,\ 0 \leq \xi_k < 1 \text{ für } k = 2, \ldots, n\}.$$

Für $A_1 := \{(\xi_1, \ldots, \xi_n) \in T(Q) : \xi_2 < 1\}$ und $A_2 := T(Q) \setminus A_1$ gilt $A_1 \cup (A_2 - e_2) = Q$ und $A_1 \cap (A_2 - e_2) = \emptyset$, somit $\triangle(T) = \mu(A_1 \cup A_2) = \mu(A_1) + \mu(A_2) = \mu(A_1) + \mu(A_2 - e_2) = \mu(Q) = 1$.

Wir konstruieren weitere Beispiele von (äusseren) Massen. Für eine Teilmenge C eines metrischen Raumes X bezeichnet

$$\operatorname{diam} C := \sup\{d(x,y) : x, y \in C\}$$

den *Durchmesser* von C.

3.14 Definition (Hausdorff-Masse)
Sei X ein metrischer Raum, $0 \le s < \infty$. Für $A \subset X$ definiere

$$\begin{aligned}\mathscr{H}^s_\delta(A) &:= \inf\bigl\{\textstyle\sum_{k=1}^\infty (\operatorname{diam} C_k)^s : C_k \subset X,\ \operatorname{diam} C_k \le \delta,\ A \subset \bigcup_{k=1}^\infty C_k\bigr\},\\ \mathscr{H}^s(A) &:= \lim_{\delta \downarrow 0} \mathscr{H}^s_\delta(A) = \sup_{\delta > 0} \mathscr{H}^s_\delta(A).\end{aligned}$$

Die so definierte Funktion $\mathscr{H}^s \colon 2^X \to [0,\infty]$ heisst s-dimensionales *äusseres Hausdorff-Mass.*

Eine entsprechende übliche Bezeichnung für das n-dimensionale äussere Lebesgue-Mass ist $\mathscr{L}^n$. In 3.14 werden die Konventionen

$$0^0 = 1 \quad \text{und} \quad (\operatorname{diam} \emptyset)^s = 0 \text{ für alle } s \ge 0$$

sowie $\inf \emptyset = \infty$ verwendet. Offensichtlich ist $\mathscr{H}^0$ gerade das Zählmass (s. 1.15(1)), und auf $\mathbb{R}$ stimmt $\mathscr{H}^1$ mit dem äusseren Lebesgue-Mass $\mathscr{L}^1$ überein.

3.15 Satz (Eigenschaften der Hausdorff-Masse)
(1) *$\mathscr{H}^s$ ist ein Borel-reguläres äusseres Mass für alle $s \ge 0$.*
(2) *Ist $0 \le s < t < \infty$, $A \subset X$ und $\mathscr{H}^s(A) < \infty$, so ist $\mathscr{H}^t(A) = 0$.*
(3) *$\mathscr{H}^s$ ist isometrie-invariant. Für $n \in \mathbb{N}$ und $X = \mathbb{R}^n$ existiert eine Konstante $0 < c < \infty$, so dass $\mathscr{H}^n = c\mathscr{L}^n$.*
(4) *Ist $A \subset X$ und $f \colon A \to Y$ eine λ-Lipschitz-Abbildung in einen metrischen Raum Y (d.h. $d(f(x), f(x')) \le \lambda d(x,x')$ für alle $x, x' \in A$), so gilt $\mathscr{H}^s(f(A)) \le \lambda^s \mathscr{H}^s(A)$.*

Die *Hausdorff-Dimension* $\dim_H(A)$ einer nicht-leeren Menge $A \subset X$ ist die Zahl

$$\begin{aligned}\dim_H(A) &= \inf\{s \ge 0 : \mathscr{H}^s(A) = 0\}\\ &= \inf\{s \ge 0 : \mathscr{H}^s(A) < \infty\}\\ &= \sup\{s \ge 0 : \mathscr{H}^s(A) > 0\};\end{aligned}$$

nach (2) stimmen diese Grössen überein. In (3) gilt für die in 3.14 gewählte Normierung $\mathscr{H}^n = (2^n/\alpha_n)\mathscr{L}^n$, wobei $\alpha_n := \mathscr{L}^n(B(0,1))$.

Beweis von 3.15: (1): Offensichtlich sind $\mathscr{H}^s_\delta$ $(0 < \delta \le \infty)$ und $\mathscr{H}^s$ äussere Masse. Für $A, B \subset X$ mit $d(A,B) > \delta$ gilt $\mathscr{H}^s_\delta(A \cup B) = \mathscr{H}^s_\delta(A) + \mathscr{H}^s_\delta(B)$. Daraus folgt $\mathscr{H}^s(A \cup B) = \mathscr{H}^s(A) + \mathscr{H}^s(B)$ falls $d(A,B) > 0$, nach 3.6 (Carathéodory's Kriterium) sind also alle Borelmengen $\mathscr{H}^s$-messbar. Sei $A \subset X$ eine Menge mit $\mathscr{H}^s(A) < \infty$. Wähle eine Nullfolge $\delta_1 > \delta_2 > \ldots$, so dass $\mathscr{H}^s_{\delta_i}(A) < \infty$. Für $i = 1, 2, \dots$ existiert dann jeweils eine Überdeckung $(C_{ik})_{k\in\mathbb{N}}$ von A durch abgeschlossene Mengen C_{ik} mit $\operatorname{diam} C_{ik} \le \delta_i$ und $\sum_{k=1}^\infty(\operatorname{diam} C_{ik})^s \le \mathscr{H}^s_{\delta_i}(A) + \delta_i$. Die Borelmenge $B := \bigcap_{i=1}^\infty \bigcup_{k=1}^\infty C_{ik}$ enthält A, und es gilt

$$\begin{aligned}\mathscr{H}^s(B) = \lim_{i\to\infty} \mathscr{H}^s_{\delta_i}(B) &\le \liminf_{i\to\infty} \sum_{k=1}^\infty (\operatorname{diam} C_{ik})^s \\ &\le \lim_{i\to\infty} \big(\mathscr{H}^s_{\delta_i}(A) + \delta_i\big) = \mathscr{H}^s(A).\end{aligned}$$

Dies zeigt, dass $\mathscr{H}^s$ Borel-regulär ist.

(2): Ist $(C_k)_{k\in\mathbb{N}}$ eine Überdeckung von A mit $\operatorname{diam} C_k \le \delta$, so gilt $(\operatorname{diam} C_k)^t \le \delta^{t-s}(\operatorname{diam} C_k)^s$ für alle k. Daraus folgt leicht die Behauptung.

(3): Für einen Würfel $C \subset \mathbb{R}^n$ gilt $(\operatorname{diam} C)^n = \sqrt{n}^n V(C)$. Daraus folgt $\mathscr{H}^n \le \sqrt{n}^n \mathscr{L}^n$, insbesondere ist $\mathscr{H}^n$ endlich auf kompakten Mengen. Die Behauptung folgt mit 3.11(4).

(4): Ist $(C_k)_{k\in\mathbb{N}}$ eine Überdeckung von A mit $\operatorname{diam} C_k \le \delta$, so ist $\operatorname{diam} f(C_k \cap A) \le \lambda \operatorname{diam} C_k \le \lambda\delta$ und somit

$$\mathscr{H}^s_{\lambda\delta}(f(A)) \le \sum_{k=1}^\infty (\lambda \operatorname{diam} C_k)^s = \lambda^s \sum_{k=1}^\infty (\operatorname{diam} C_k)^s.$$

Da dies für alle solchen Überdeckungen gilt, folgt $\mathscr{H}^s_{\lambda\delta}(f(A) \le \lambda^s \mathscr{H}^s_\delta(A)$ und daraus die Behauptung. □

Wir können 3.2, 3.9 und 3.10 auf $\mathscr{H}^s$ anwenden. Man beachte aber, dass $\mathscr{H}^s$ nur in speziellen Fällen lokal-endlich ist (z.B. $X = \mathbb{R}^n$, $s = n$ oder X diskret, $s = 0$). Ist $E \subset X$ eine $\mathscr{H}^s$-messbare Menge mit $\mathscr{H}^s(E) < \infty$, so ist jedoch $\mathscr{H}^s \llcorner E$ ein endliches Borel-reguläres äusseres Mass (3.9(2)), für welches die Voraussetzungen in 3.10 also immer erfüllt sind.

4 Radon-Masse

4.1 Definition (lokal-kompakter Hausdorff-Raum)
Sei X ein topologischer Raum.

(1) Eine Menge $U \subset X$ heisst eine *Umgebung* von $x \in X$, falls U eine offene Menge enthält, die x enthält. X heisst ein *Hausdorff-Raum*, falls zu je zwei verschiedenen Punkten disjunkte Umgebungen existieren.

(2) X heisst *kompakt*, wenn jede offene Überdeckung von X eine endliche Teilüberdeckung besitzt. X heisst *lokal-kompakt*, falls jeder Punkt eine kompakte Umgebung besitzt.

$\mathbb{R}^n$ ist ein lokal-kompakter Hausdorff-Raum. Jeder metrische Raum ist ein Hausdorff-Raum. Ein kompakter Teilraum eines Hausdorff-Raumes X ist abgeschlossen in X.

4.2 Definition (Radon-Mass)
Sei X ein lokal-kompakter Hausdorff-Raum, $\mathscr{M}$ eine σ-Algebra in X. Ein Mass μ auf $\mathscr{M}$ heisst ein *Radon-Mass*, falls gilt:

(i) $\mathscr{B} \subset \mathscr{M}$.
(ii) Ist $C \subset X$ kompakt (also abgeschlossen, also in $\mathscr{B}$), so ist $\mu(C) < \infty$.
(iii) Ist $U \subset X$ offen, so ist

$$\mu(U) = \sup\{\mu(C) : C \subset U,\, C \text{ kompakt}\}.$$

(iv) Ist $A \in \mathscr{M}$, so ist

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

Gemäss 3.11(2) ist das Lebesgue-Mass auf $\mathbb{R}^n$ ein Radon-Mass; (iii) gilt dann sogar für alle $U \in \mathscr{M}$. Man vergleiche den folgenden Approximationssatz mit 3.10:

4.3 Satz (Approximationseigenschaften von Radon-Massen)
Sei X ein lokal-kompakter Hausdorff-Raum, μ ein Radon-Mass auf $(X, \mathscr{M})$ und $A \in \mathscr{M}$.

(1) *Ist $\mu(A) < \infty$, so existiert für jedes $\epsilon > 0$ eine kompakte Menge $C \subset A$ mit $\mu(A \setminus C) < \epsilon$.*
(2) *Ist $A = \bigcup_{i=1}^\infty A_i$ für messbare Mengen A_i mit $\mu(A_i) < \infty$, so existiert für jedes $\epsilon > 0$ eine offene Menge $U \supset A$ mit $\mu(U \setminus A) < \epsilon$.*

Beweis: (1): Wähle (mittels 4.2(iv)) offene Mengen V und W, so dass $A \subset V$ und $\mu(V \setminus A) < \epsilon/2$ sowie $V \setminus A \subset W$ und $\mu(W) < \epsilon/2$. Wähle (mittels 4.2(iii)) eine kompakte Menge $C \subset V$, so dass $\mu(V \setminus C) < \epsilon/2$. Dann ist $C \setminus W$ ein kompakte Teilmenge von $V \setminus W \subset A$, und

$$\mu(A \setminus (C \setminus W)) \leq \mu(V \setminus C) + \mu(W) < \epsilon.$$

(2): Wähle (mittels 4.2(iv)) offene Mengen $U_i \supset A_i$ mit $\mu(U_i \setminus A_i) < 2^{-i}\epsilon$, und setze $U := \bigcup_{i=1}^\infty U_i$. □

Ist X zusätzlich *σ-kompakt*, d.h. $X = \bigcup_{i=1}^\infty D_i$ für kompakte Mengen D_i, so lässt sich 4.3 wie folgt vereinfachen bzw. verschärfen:

4.4 Satz
Sei $(X, \mathscr{M}, \mu)$ wie in 4.3, aber X zusätzlich σ-kompakt, und sei $A \in \mathscr{M}$.

(1) *Für jedes* $\epsilon > 0$ *existiert eine abgeschlossene Menge* $C \subset A$ *mit* $\mu(A \setminus C) < \epsilon$*, und*

$$\mu(A) = \sup\{\mu(C) : C \subset A,\ C \text{ kompakt}\}.$$

(2) *Für jedes* $\epsilon > 0$ *existiert eine offene Menge* $U \supset A$ *mit* $\mu(U \setminus A) < \epsilon$.

Beweis: Da X σ-kompakt ist und kompakte Mengen endliches Mass haben, folgt (2) aus 4.3(2).

(1): Sei $A \in \mathscr{M}$ und $\epsilon > 0$. Nach (2) existiert eine offene Menge $U \supset A^c$ mit $\mu(U \setminus A^c) < \epsilon$. Für $C := U^c$ gilt dann $C \subset A$ und $A \setminus C = U \setminus A^c$. Die zweite Aussage von (1) gilt nach 4.3(1) unter der Zusatzvoraussetzung $\mu(A) < \infty$. Ist $\mu(A) = \infty$, $X = \bigcup_{i=1}^\infty D_i$, $D_1 \subset D_2 \subset \ldots$ kompakt, wähle kompakte Mengen $C_i \subset A \cap D_i$ mit $\mu(C_i) \geq \mu(A \cap D_i) - 1$. Da $\bigcup_{i=1}^\infty (A \cap D_i) = A$, folgt $\mu(A \cap D_i) \to \mu(A) = \infty$ (1.14(4)) und somit $\mu(C_i) \to \mu(A)$. $\square$

Für den nächsten Satz zitieren wir ein Resultat aus der Topologie:

4.5 Satz (Lemma von Urysohn)
Sei X *ein lokal-kompakter Hausdorff-Raum,* $V \subset X$ *offen,* $C \subset V$ *kompakt. Dann existiert eine stetige Funktion* $f\colon X \to [0,1]$ *mit kompaktem Träger* $\operatorname{Tr} f$ *(Abschluss von* $\{x : f(x) \neq 0\}$*), so dass* $f \geq \chi_C$ *und* $\operatorname{Tr} f \subset V$.

Beweis: S. [Rudin, 2.12]. $\square$

Ist X ein topologischer Raum, so bezeichnen wir mit $C_c(X)$ den Vektorraum der stetigen reellen Funktionen auf X mit kompaktem Träger.

4.6 Satz (Lusin)
Sei μ *ein Radon-Mass auf einem lokal-kompakten Hausdorff-Raum* X*,* f *eine reelle messbare Funktion auf* X*,* $\mu(A) < \infty$*,* $f(x) = 0$ *für* $x \in X \setminus A$*, und sei* $\epsilon > 0$*. Dann existiert eine Funktion* $g \in C_c(X)$*, so dass*

$$\mu(\{x : f(x) \neq g(x)\}) < \epsilon.$$

Beweis: Nimm zuerst an, dass $0 \leq f < 1$ und A kompakt sei. Wähle messbare Treppenfunktionen $0 \leq s_1 \leq s_2 \leq \ldots \leq f$ wie im Beweis von 1.12, so dass $s_n(x) \to f(x)$ $(n \to \infty)$ für alle $x \in X$ und $s_n(X) \subset 2^{-n}\mathbb{Z}$. (Sei $\varphi_n\colon [0,1) \to [0,1)$ die grösste Funktion mit Werten in $2^{-n}\mathbb{Z}$, so dass $\varphi_n(t) \leq t$ für alle t. Setze $s_n := \varphi_n \circ f$.) Definiere $t_1 := s_1$ und $t_n := s_n - s_{n-1}$ für $n = 2, 3, \ldots$. Dann ist $2^n t_n$ die charakteristische Funktion einer Menge $T_n \subset A$, und $f(x) = \sum_{n=1}^\infty t_n(x)$ $(x \in X)$. Wähle eine offene Menge $V \supset A$

mit kompaktem Abschluss $\overline{V}$. Es existieren kompakte Mengen C_n und offene Mengen V_n, so dass $C_n \subset T_n \subset V_n \subset V$ und $\mu(V_n \backslash C_n) < 2^{-n}\epsilon$ (4.3). Nach 4.5 gibt es stetige Funktionen $h_n \colon X \to [0,1]$, so dass $\chi_{C_n} \le h_n \le \chi_{V_n}$. Definiere

$$g(x) := \sum_{n=1}^{\infty} 2^{-n} h_n(x) \quad (x \in X).$$

Diese Reihe konvergiert gleichmässig auf X, somit ist g stetig. Der Träger von g liegt in $\overline{V}$, ist also kompakt. Es gilt $2^{-n}h_n(x) = t_n(x)$ ausserhalb $V_n \backslash C_n$, somit $g(x) = f(x)$ ausserhalb $\bigcup_{n=1}^{\infty}(V_n \setminus C_n)$, wobei $\mu(\bigcup_{n=1}^{\infty}(V_n \setminus C_n)) < \epsilon$. Dies zeigt die Behauptung im Fall $0 \le f < 1$ und A kompakt. Daraus folgt die Behauptung für f beschränkt und A kompakt, die Kompaktheitsvoraussetzung an A kann ausserdem wegen $\mu(A) < \infty$ mittels 4.3(1) eliminiert werden. Für eine allgemeine reelle (messbare) Funktion f setze $B_n := \{x : |f(x)| > n\}$; dann gilt $\bigcap_{n=1}^{\infty} B_n = \emptyset$, also $\mu(B_n) \to 0$ (1.14(5), $B_1 \subset A$). Da f mit der beschränkten Funktion $(1 - \chi_{B_n})f$ ausserhalb B_n übereinstimmt, folgt das allgemeine Resultat. $\square$

4.7 Definition
Eine $\mathbb{R}$-wertige (oder $\bar{\mathbb{R}}$-wertige) Funktion f auf einem topologischen Raum X heisst *von unten halbstetig* bzw. *von oben halbstetig*, falls $\{x \in X : f(x) > \alpha\}$ bzw. $\{x \in X : f(x) < \alpha\}$ offen ist für jedes $\alpha \in \mathbb{R}$.

Offensichtlich ist $f \colon X \to \bar{\mathbb{R}}$ ist genau dann stetig, wenn f sowohl von unten als auch von oben halbstetig ist. Die charakteristische Funktion einer offenen bzw. abgeschlossenen Menge ist von unten bzw. von oben halbstetig. Das Supremum einer Familie von Funktionen, die von unten halbstetig sind, ist von unten halbstetig. Das Infimum einer Familie von Funktionen, die von oben halbstetig sind, ist von oben halbstetig.

4.8 Satz (Vitali-Carathéodory)
Sei μ ein Radon-Mass auf einem lokal-kompakten Hausdorff-Raum X. Sei $f \in L^1(\mu)$ und $\epsilon > 0$. Dann existieren eine von oben beschränkte und von oben halbstetige Funktion u und eine von unten beschränkte und von unten halbstetige Funktion v auf X, so dass $u \le f \le v$ und

$$\int_X v - u \, d\mu < \epsilon.$$

Beweis: Zuerst nehmen wir $f \ge 0$ und $f \not\equiv 0$ an. Dann ist f punktweiser Limes einer Folge messbarer Treppenfunktionen $0 = s_0 \le s_1 \le s_2 \ldots$. Setze $t_n := s_n - s_{n-1}$ ($n \in \mathbb{N}$); es gilt $f = \sum_{n=1}^{\infty} t_n$. Es existieren messbare Mengen E_i (nicht notwendigerweise disjunkt) und Konstanten $c_i > 0$, so dass $f(x) = \sum_{i=1}^{\infty} c_i \chi_{E_i}(x)$ $(x \in X)$. Es folgt $\sum_{i=1}^{\infty} c_i \mu(E_i) = \int_X f\,d\mu < \infty$. Wähle kompakte Mengen C_i und offene Mengen V_i, so dass $C_i \subset E_i \subset V_i$ und $c_i \mu(V_i \setminus C_i) < 2^{-i-1}\epsilon$ $(i \in \mathbb{N})$. Setze dann

$$u := \sum_{i=1}^{N} c_i \chi_{C_i}, \quad v := \sum_{i=1}^{\infty} c_i \chi_{V_i},$$

wobei N so gewählt sei, dass $\sum_{i=N+1}^{\infty} c_i \mu(E_i) < \epsilon/2$. Dann ist n von oben halbstetig, v von unten halbstetig, $u \le f \le v$, und

$$v - u \le \sum_{i=1}^{\infty} c_i(\chi_{V_i} - \chi_{C_i}) + \sum_{i=N+1}^{\infty} c_i \chi_{E_i},$$

somit $\int_X v - u\,d\mu < \epsilon$. Im allgemeinen Fall schreiben wir $f = f^+ - f^-$ und wählen $u_1 \le f^+ \le v_1$, $u_2 \le f^- \le v_2$ wie oben, mit $\int_X v_1 - u_1\,d\mu < \epsilon/2$, $\int_X v_2 - u_2\,d\mu < \epsilon/2$. Dann haben $u := u_1 - v_2$ und $v := v_1 - u_2$ die gewünschten Eigenschaften, und $\int_X v - u\,d\mu = \int_X v_1 - u_1\,d\mu + \int_X v_2 - u_2\,d\mu < \epsilon$. □

4.9 Satz (Zerlegung der Eins)
Seien X ein lokal-kompakter Hausdorff-Raum, $V_1, V_2, \ldots, V_n \subset X$ offen und $C \subset \bigcup_{i=1}^n V_i$ kompakt. Dann existieren stetige Funktionen $h_i \colon X \to [0,1]$ $(i = 1, 2, \ldots, n)$ mit kompaktem Träger $\mathrm{Tr}\, h_i \subset V_i$ und $\sum_{i=1}^n h_i(x) = 1$ für alle $x \in C$.

Beweis: Jedes $x \in C$ besitzt eine offene Umgebung W_x mit kompaktem Abschluss $\overline{W}_x \subset V_i$ für ein $i = i(x)$. (Verwende z.B. 4.5 für $\{x\} \subset V_i$ und setze $W_x := \{x : f(x) \ne 0\}$.) Es existieren $x_1, \ldots, x_m$, so dass $C \subset \bigcup_{j=1}^m W_{x_j}$. Für $1 \le i \le n$ sei D_i die Vereinigung aller $\overline{W}_{x_j}$, die in V_i liegen. Nach 4.5 gibt es stetige Funktionen $g_i \colon X \to [0,1]$ mit $g_i \ge \chi_{D_i}$ und kompaktem Träger $\mathrm{Tr}\, g_i \subset V_i$. Setze

$$\begin{aligned} h_1 &= g_1, \\ h_2 &= (1 - g_1) g_2, \\ &\vdots \\ h_n &= (1 - g_1)(1 - g_2) \ldots (1 - g_{n-1}) g_n. \end{aligned}$$

Es gilt $\mathrm{Tr}\, h_i \subset V_i$ und $h_1 + h_2 + \ldots + h_n = 1 - (1 - g_1)(1 - g_2) \ldots (1 - g_n)$ (Beweis durch Induktion). Wegen $C \subset \bigcup_{i=1}^n D_i$ ist für jedes $x \in C$ wenigstens ein $g_i(x)$ gleich 1, somit $\sum_{i=1}^n h_i(x) = 1$. □

Im Folgenden bezeichnet $C_c(X)$ wieder den Vektorraum aller stetigen reellen Funktionen auf X mit kompaktem Träger. Ist μ ein Radonmass auf X, so wird durch

$$\Lambda f := \int_X f\, d\mu$$

eine lineare Abbildung $\Lambda\colon C_c(X) \to \mathbb{R}$ definiert; Λ heisst ein *lineares Funktional* auf $C_c(X)$. (Nach Satz 2.9 gilt genauso: Für ein allgemeines Mass μ auf einer Menge X definiert $\Lambda f := \int_X f\, d\mu$ ein lineares Funktional auf $L^1(\mu)$.) Ein lineares Funktional $\Lambda\colon C_c(X) \to \mathbb{R}$ heisst *positiv*, falls $\Lambda f \geq 0$ für alle $f \in C_c(X)$ mit $f \geq 0$.

4.10 Satz (Darstellungssatz von Riesz)
Sei X ein lokal-kompakter Hausdorff-Raum, und sei Λ ein positives lineares Funktional auf $C_c(X)$. Dann existieren eine σ-Algebra $\mathscr{M}$ in X mit $\mathscr{B} \subset \mathscr{M}$ und ein eindeutig bestimmtes Radonmass μ auf $\mathscr{M}$, so dass

$$\Lambda f = \int_X f\, d\mu \quad \textit{für alle } f \in C_c(X).$$

Beweis: (I) Wir zeigen zuerst die Eindeutigkeit von μ. Jedes Radonmass ist aufgrund von 4.2(iii) und (iv) durch seine Werte auf kompakten Mengen bestimmt. Sei μ' ein weiteres Radonmass auf $\mathscr{M}$ wie im Satz. Sei $C \subset X$ kompakt und $\epsilon > 0$. Nach 4.2(ii) und (iv) existiert eine offene Menge $V \supset C$ mit $\mu'(V) < \mu'(C) + \epsilon$. Sei $f \in C_c(X)$ wie in 4.5. Dann gilt

$$\begin{aligned} \mu(C) &= \int_X \chi_C\, d\mu \leq \int_X f\, d\mu = \Lambda f = \int_X f\, d\mu' \\ &\leq \int_X \chi_V\, d\mu' = \mu'(V) < \mu'(C) + \epsilon. \end{aligned}$$

Es folgt $\mu(C) \leq \mu'(C)$ und analog $\mu'(C) \leq \mu(C)$. Dies zeigt die Eindeutigkeit von μ.

(II) Nun definieren wir eine Funktion $\nu\colon 2^X \to [0, \infty]$ so, dass

$$\nu(V) = \sup\{\Lambda f : 0 \leq f \leq 1,\ \operatorname{Tr} f \subset V\}$$

für jede offene Menge V und

$$\nu(A) = \inf\{\nu(V) : A \subset V,\ V \text{ offen}\}$$

für jede Menge A. Ist $f \leq g$, so gilt

$$\Lambda f \leq \Lambda f + \Lambda(g - f) = \Lambda g$$

wegen der Positivität von Λ. Daraus folgt, dass die beiden Definitionen von $\nu(A)$ für offene Mengen A konsistent sind und dass ν monoton ist. Ausserdem gilt $\nu(\emptyset) = 0$. Wir zeigen noch, dass ν σ-subadditiv ist; dann ist ν ein äusseres Mass auf X. Es genügt zu zeigen, dass

$$\nu\big(\textstyle\bigcup_{i=1}^\infty A_i\big) \leq \displaystyle\sum_{i=1}^\infty \nu(A_i)$$

für Mengen $A_i \subset X$ mit $\nu(A_i) < \infty$. Sei $\epsilon > 0$. Dann existieren offene Mengen $V_i \supset A_i$, so dass $\nu(V_i) < \nu(A_i) + 2^{-i}\epsilon$. Setze $V := \bigcup_{i=1}^\infty V_i$. Sei $f \in C_c(X)$ mit $0 \leq f \leq 1$ und $\operatorname{Tr} f \subset V$. Es gilt $\operatorname{Tr} f \subset \bigcup_{i=1}^n V_i$ für ein $n \in \mathbb{N}$. Nach 4.9 existieren Funktionen $h_i \in C_c(X)$, so dass $0 \leq h_i \leq 1$ und $\sum_{i=1}^n h_i(x) = 1$ für alle $x \in \operatorname{Tr} f$ und $\operatorname{Tr} h_i \subset V_i$. Es folgt

$$\Lambda f = \Lambda\big(\textstyle\sum_{i=1}^n h_i f\big) = \displaystyle\sum_{i=1}^n \Lambda(h_i f) \leq \sum_{i=1}^n \nu(V_i) < \sum_{i=1}^\infty \nu(A_i) + \epsilon.$$

Da dies für alle $f \in C_c(X)$ mit $0 \leq f \leq 1$ und $\operatorname{Tr} f \subset V$ gilt, und da $\bigcup_{i=1}^\infty A_i \subset V$, folgt

$$\nu\big(\textstyle\bigcup_{i=1}^\infty A_i\big) \leq \nu(V) \leq \displaystyle\sum_{i=1}^\infty \nu(A_i) + \epsilon.$$

Somit ist ν σ-subadditiv und ein äusseres Mass.

(III) Wir zeigen: *Ist C kompakt, so ist $\nu(C) < \infty$ und*

$$\nu(C) = \inf\{\Lambda f : \chi_C \leq f \leq 1\}.$$

Ist $f \in C_c(X)$ mit $\chi_C \leq f \leq 1$, setze $V_\alpha := \{x : f(x) > \alpha\}$ für $0 < \alpha < 1$. Dann ist $C \subset V_\alpha$, und für alle $g \in C_c(X)$ mit $0 \leq g \leq 1$ und $\operatorname{Tr} g \subset V_\alpha$ gilt $\alpha g \leq f$. Somit ist

$$\nu(C) \leq \nu(V_\alpha) = \sup\{\Lambda g : 0 \leq g \leq 1,\ \operatorname{Tr} g \subset V_\alpha\} \leq \alpha^{-1}\Lambda f.$$

Für $\alpha \to 1$ folgt $\nu(C) \leq \Lambda f$, insbesondere $\nu(C) < \infty$. Sei $\epsilon > 0$. Da $\nu(C) < \infty$, existiert eine offene Menge $V \supset C$ mit $\nu(V) < \mu(C) + \epsilon$. Nach 4.5 existiert ein $f \in C_c(X)$, so dass $\chi_C \leq f \leq 1$ und $\operatorname{Tr} f \subset V$. Dann gilt

$$\Lambda f \leq \nu(V) < \nu(C) + \epsilon.$$

Da ausserdem $\nu(C) \leq \Lambda f$, wie oben gezeigt, folgt die Behauptung.

(IV) Wir zeigen: *Für jede offene Menge V gilt*

$$\nu(V) = \sup\{\nu(C) : C \subset V,\, C \text{ kompakt}\}.$$

Sei $\alpha \in \mathbb{R}$, $\alpha < \nu(V)$. Es existiert ein $f \in C_c(X)$ mit $0 \le f \le 1$, $C := \operatorname{Tr} f \subset V$ und $\Lambda f > \alpha$. Für jede offene Menge $W \supset C$ gilt $\nu(W) \ge \Lambda f$. Somit ist

$$\nu(C) = \inf\{\nu(W) : C \subset W,\, W \text{ offen}\} \ge \Lambda f > \alpha.$$

Dies zeigt die Behauptung.

Die erste Aussage von (III), die soeben gezeigte Behauptung und die Definition von $\nu(A)$ für $A \subset X$ entsprechen gerade den Eigenschaften 4.2(ii), (iii) und (iv) der Radon-Masse. Ist X ein lokal-kompakter *metrischer* Raum, so folgt leicht mit 3.6 (Carathéodory's Kriterium), dass die σ-Algebra $\mathscr{M}$ der ν-messbaren Mengen alle Borelmengen enthält; nach 3.2 ist $\mu := \nu|_{\mathscr{M}}$ dann das gesuchte Radonmass. Um den Beweis in diesem Fall zu beenden, kann man die folgenden Teile (V) bis (VIII) überspringen.

(V) Sei $\mathscr{M}_e$ die Menge aller $A \subset X$ mit $\nu(A) < \infty$ und

$$\nu(A) = \sup\{\nu(C) : C \subset A,\, C \text{ kompakt}\},$$

und sei $\mathscr{M}$ die Menge aller $A \subset X$ mit $A \cap C \in \mathscr{M}_e$ für jede kompakte Menge C. Wir zeigen: *Sind $A_1, A_2, \ldots \in \mathscr{M}_e$ paarweise disjunkt und ist $A := \bigcup_{i=1}^\infty A_i$, so gilt*

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

Im Fall $\nu(A) < \infty$ ist auch $A \in \mathscr{M}_e$. Wir zeigen zuerst, dass $\nu(C_1 \cup C_2) = \nu(C_1) + \nu(C_2)$ für disjunkte Mengen C_1, C_2. Sei $\epsilon > 0$. Nach 4.5 existiert $f \in C_c(X)$, so dass $\chi_{C_1} \le f \le 1 - \chi_{C_2}$. Nach (III) gibt es ein $g \in C_c(X)$ mit $\chi_{C_1 \cup C_2} \le g \le 1$ und $\Lambda g < \nu(C_1 \cup C_2) + \epsilon$. Es folgt $\chi_{C_1} \le fg$, $\chi_{C_2} \le (1-f)g$ und somit

$$\nu(C_1) + \nu(C_2) \le \Lambda(fg) + \Lambda((1-f)g) = \Lambda(g) < \nu(C_1 \cup C_2) + \epsilon.$$

Seien jetzt $A_1, A_2, \ldots$ und A wie oben, o.B.d.A. $\nu(A) < \infty$. Sei $\epsilon > 0$. Da $A_i \in \mathscr{M}_e$, existieren kompakte Mengen $B_i \subset A_i$ mit $\nu(B_i) > \nu(A_i) - 2^{-i}\epsilon$. Setze $C_n := \bigcup_{i=1}^n B_i$; es gilt

$$\nu(A) \ge \nu(C_n) = \sum_{i=1}^n \nu(B_i) > \sum_{i=1}^n \nu(A_i) - \epsilon.$$

Die behauptete Identität folgt. Im Fall $\nu(A) < \infty$ ist ausserdem $\nu(A) \le \sum_{i=1}^N \nu(A_i) + \epsilon < \nu(C_N) + 2\epsilon$ für ein $N \in \mathbb{N}$, also $A \in \mathscr{M}_e$.

(VI) Wir zeigen: *Ist $A \in \mathscr{M}_e$ und $\epsilon > 0$, so existieren eine kompakte Menge C und eine offene Menge V, so dass $C \subset A \subset V$ und $\nu(V \setminus C) < \epsilon$.* Da $A \in \mathscr{M}_e$, und nach Definition von $\nu(A)$, existieren solche C und V mit

$$\nu(V) - \epsilon/2 < \nu(A) < \nu(C) + \epsilon/2.$$

Da $V \setminus C$ offen ist, ist $V \setminus C \in \mathscr{M}_e$ gemäss (IV). Mit (V) folgt $\nu(V \setminus C) = \nu(V) - \nu(C) < \epsilon$.

(VII) Wir zeigen: *Aus $A, B \in \mathscr{M}_e$ folgt $A \setminus B$, $A \cup B$, $A \cap B \in \mathscr{M}_e$.* Sei $\epsilon > 0$. Nach (VI) existieren C_i und V_i, so dass $C_1 \subset A \subset V_1$, $C_2 \subset B \subset V_2$ und $\nu(V_i \setminus C_i) < \epsilon$, $i = 1, 2$. Da

$$A \setminus B \subset V_1 \setminus C_2 \subset (V_1 \setminus C_1) \cup (C_1 \setminus V_2) \cup (V_2 \setminus C_2),$$

folgt $\nu(A \setminus B) \leq \epsilon + \nu(C_1 \setminus V_2) + \epsilon$. Da $C_1 \setminus V_2$ kompakte Teilmenge von $A \setminus B$ ist, folgt $A \setminus B \in \mathscr{M}_e$. Dann ist auch $A \cap B = A \setminus (A \setminus B)$ in $\mathscr{M}_e$, und mit (V) auch $A \cup B = (A \setminus B) \cup B$.

(VIII) Wir zeigen: *$\mathscr{M}$ ist eine σ-Algebra, die alle Borelmengen enthält, und die Einschränkung $\mu := \nu|_{\mathscr{M}}$ ist ein Radonmass.* Sei $C \subset X$ kompakt. Ist $A \in \mathscr{M}$, so gilt $C, A \cap C \in \mathscr{M}_e$ und somit $A^c \cap C = C \setminus (A \cap C) \in \mathscr{M}_e$ nach (VII). Also ist $A^c \in \mathscr{M}$. Sind $A_1, A_2, \dots \in \mathscr{M}$, $A := \bigcup_{i=1}^{\infty} A_i$, so sei $B_1 := A_1 \cap C$ und

$$B_n := (A_n \cap C) \setminus \bigcap_{i=1}^{n-1} B_i \quad (n = 2, 3, \dots).$$

Dann ist $(B_n)_{n \in \mathbb{N}}$ eine disjunkte Folge in $\mathscr{M}_e$ nach (VII), und $A \cap C = \bigcup_{n=1}^{\infty} B_n$, also $A \cap C \in \mathscr{M}_e$ gemäss (V). Somit ist $A \in \mathscr{M}$. Ist $A \subset X$ abgeschlossen, so ist $A \cap C$ kompakt und somit in $\mathscr{M}_e$. Also ist $A \in \mathscr{M}$, d.h. $\mathscr{B} \subset \mathscr{M}$. Sei $A \in \mathscr{M}$ mit $\mu(A) < \infty$. Wir zeigen, dass dann $A \in \mathscr{M}_e$; aus (V) folgt damit die σ-Additivität von μ auf $\mathscr{M}$. Sei $\epsilon > 0$. Es existiert eine offene Menge $V \supset A$ mit $\mu(V) < \infty$. Nach (IV) ist $V \in \mathscr{M}_e$, und nach (VI) gibt es eine kompakte Menge $B \subset V$, so dass $\mu(V \setminus B) < \epsilon$. Da $A \cap B \in \mathscr{M}_e$, existiert eine kompakte Menge $C \subset A \cap B$ mit $\mu(A \cap B) < \mu(C) + \epsilon$. Wegen $A \subset (A \cap B) \cup (V \setminus B)$ folgt

$$\mu(A) \leq \mu(A \cap B) + \mu(V \setminus B) < \mu(C) + 2\epsilon.$$

Somit ist $A \in \mathscr{M}_e$.

(IX) Es bleibt noch zu zeigen, dass $\Lambda f = \int_X f \, d\mu$ für alle $f \in C_c(X)$. Sei $C := \operatorname{Tr} f$, $f(X) \subset [a, b]$. Sei $\epsilon > 0$. Wähle y_i, $i = 0, 1, \dots, n$, so dass $y_i - y_{i-1} < \epsilon$ und

$$y_0 < a < y_1 < \dots < y_n = b.$$

Setze $A_i := \{x \in C : y_{i-1} < f(x) \leq y_i\}$ $(i = 1, \dots, n)$. Da f stetig ist, ist f Borel-messbar; die A_i sind also paarweise disjunkte Borelmengen mit $\bigcup_{i=1}^n A_i = C$. Es existieren offene Mengen $V_i \supset A_i$, so dass

$$\mu(V_i) < \mu(A_i) + \epsilon/n$$

und $f(x) < y_i + \epsilon$ für alle $x \in V_i$. Wähle Funktionen h_i wie in 4.9. Dann ist $f = \sum_{i=1}^n h_i f$, und nach (III) gilt $\mu(C) \leq \Lambda(\sum_{i=1}^\infty h_i) = \sum_{i=1}^\infty \Lambda h_i$. Da $h_i f \leq (y_i + \epsilon) h_i$ und $y_i - \epsilon < f(x)$ auf A_i, folgt

$$\begin{aligned}
\Lambda f &= \sum_{i=1}^n \Lambda(h_i f) \leq \sum_{i=1}^n (y_i + \epsilon)\Lambda h_i \\
&= \sum_{i=1}^\infty (|a| + y_i + \epsilon)\Lambda h_i - |a| \sum_{i=1} \Lambda h_i \\
&\leq \sum_{i=1}^\infty (|a| + y_i + \epsilon)(\mu(A_i) + \epsilon/n) - |a|\mu(C) \\
&= \sum_{i=1}^\infty (y_i - \epsilon)\mu(A_i) + \frac{\epsilon}{n} \sum_{i=1}^n (|a| + y_i + \epsilon) + 2\epsilon\mu(C) \\
&\leq \int_X f\, d\mu + \epsilon\big(|a| + b + \epsilon + 2\mu(C)\big).
\end{aligned}$$

Dies zeigt $\Lambda f \leq \int_X f\, d\mu$. Damit ist auch

$$-\Lambda f = \Lambda(-f) \leq \int_X (-f)\, d\mu = -\int_X f\, d\mu,$$

also $\Lambda f = \int_X f\, d\mu$. □

5 L^p-Räume

Eine reelle Funktion φ auf einem Intervall (a, b), wobei $-\infty \leq a < b \leq \infty$, heisst *konvex*, falls

$$\varphi((1 - \lambda)x + \lambda y) \leq (1 - \lambda)\varphi(x) + \lambda\varphi(y)$$

für alle $x, y \in (a, b)$ und $\lambda \in [0, 1]$. Dies ist äquivalent zur Forderung, dass

$$\frac{\varphi(t) - \varphi(s)}{t - s} \leq \frac{\varphi(u) - \varphi(t)}{u - t}$$

für $a < s < t < u < b$. Mit dem Mittelwertsatz folgt: Eine reelle differenzierbare Funktion φ ist genau dann konvex auf (a, b), wenn $\varphi'(s) \leq \varphi'(t)$ für $a < s < t < b$. Ist φ konvex auf (a, b), so ist φ stetig auf (a, b). Dies verwendet, dass (a, b) offen ist: Ist $\varphi(x) = 0$ auf $[0, 1)$ und $\varphi(1) = 1$, so ist φ konvex auf $[0, 1]$ aber nicht stetig.

5.1 Satz (Jensen-Ungleichung)
Sei $(\Omega, \mathscr{M}, \mu)$ ein Massraum mit $\mu(\Omega) = 1$. Ist $f \in L^1(\mu)$ mit $a < f(x) < b$ für alle $x \in \Omega$, und ist φ konvex auf (a, b), so gilt

$$\varphi\left(\int_\Omega f \, d\mu\right) \leq \int_\Omega \varphi \circ f \, d\mu$$

Die Fälle $a = -\infty$ oder $b = \infty$ sind zugelassen. Möglicherweise ist $\varphi \circ f$ nicht in $L^1(\mu)$; der Beweis wird zeigen, dass in diesem Fall das Integral von $\varphi \circ f$ in dem nach Definition 2.8 erklärten Sinn existiert und gleich ∞ ist.

Beweis: Sei $t := \int_\Omega f \, d\mu$. Dann ist $a < t < b$. Da φ konvex ist, gilt für alle $u \in (t, b)$

$$\beta := \sup_{s \in (a,t)} \frac{\varphi(t) - \varphi(s)}{t - s} \leq \frac{\varphi(u) - \varphi(t)}{u - t}$$

und somit $\varphi(u) - \varphi(t) \geq \beta(u - t)$; diese Ungleichung gilt aufgrund der Definition von β aber auch für alle $u \in (a, t)$. Es folgt

$$\varphi(f(x)) \geq g(x) := \varphi(t) + \beta(f(x) - t)$$

für alle $x \in \Omega$. Da φ stetig ist, ist $\varphi \circ f$ messbar. Mit $\int_\Omega 1 \, d\mu = 1$ folgt

$$\int_\Omega (\varphi \circ f)^+ \, d\mu - \int_\Omega (\varphi \circ f)^- \, d\mu \geq \int_\Omega g^+ \, d\mu - \int_\Omega g^- \, d\mu = \int_\Omega g \, d\mu = \varphi(t),$$

also die Behauptung. □

Sei z.B. $\varphi(x) = e^x$. Dann gilt

$$\exp\left(\int_\Omega f \, d\mu\right) \leq \int_\Omega e^f \, d\mu.$$

Ist Ω eine endliche Menge, $\Omega = \{p_1, \ldots, p_n\}$, und ist $\mu(\{p_i\}) = 1/n$, $f(p_i) = x_i$, so folgt

$$\exp\left(\frac{1}{n}(x_1 + \ldots + x_n)\right) \leq \frac{1}{n}(e^{x_1} + \ldots + e^{x_n})$$

für $x_1, \ldots, x_n \in \mathbb{R}$. Setzt man noch $y_i := e^{x_i}$, so erhält man die Ungleichung zwischen geometrischem und arithmetischem Mittel,

$$(y_1 y_2 \ldots y_n)^{1/n} \leq \frac{1}{n}(y_1 + y_2 + \ldots + y_n)$$

für $y_1, \ldots, y_n \geq 0$. Dementsprechend heissen die linke bzw. rechte Seite von

$$\exp\left(\int_\Omega \log g \, d\mu\right) \leq \int_\Omega g \, d\mu$$

auch *geometrisches* bzw. *arithmetisches Mittel* der positiven Funktion g. Wählt man oben $\mu(\{p_i\}) = \alpha_i > 0$, wobei $\sum_{i=1}^n \alpha_i = 1$, so erhält man

$$y_1^{\alpha_1} y_2^{\alpha_2} \ldots y_n^{\alpha_n} \leq \alpha_1 y_1 + \alpha_2 y_2 + \ldots + \alpha_n y_n.$$

5.2 Definition (konjugierte Exponenten)
Sind p, q positive reelle Zahlen mit $p + q = pq$, d.h.

$$\frac{1}{p} + \frac{1}{q} = 1,$$

so heissen p und q ein Paar von *konjugierten Exponenten*. Ein wichtiger Spezialfall ist $p = q = 2$. Es gilt $1 < p < \infty$ und $1 < q < \infty$. Für $p \to 1$ folgt $q \to \infty$; 1 und ∞ werden deshalb auch als ein Paar konjugierter Exponenten betrachtet. Oft wird der zu p konjugierte Exponent auch mit p' bezeichnet.

5.3 Satz (Hölder- und Minkowski-Ungleichung)
Seien p und q konjugierte Exponenten, $1 < p < \infty$. Sei X ein Massraum mit Mass μ, und seien f, g messbare Funktionen auf X mit Werten in $[0, \infty]$. Dann gelten

(1) *die Hölder-Ungleichung*

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

(2) *die Minkowski-Ungleichung*

$$\left(\int_X (f+g)^p \, d\mu\right)^{1/p} \leq \left(\int_X f^p \, d\mu\right)^{1/p} + \left(\int_X g^p \, d\mu\right)^{1/p}.$$

Im Spezialfall $p = q = 2$ heisst (1) auch *Schwarz'sche Ungleichung.*

Beweis: (1) Seien A und B die beiden Faktoren auf der rechten Seite. Im Fall $A = 0$ ist $f = 0$ fast überall auf X (2.15(1)), also $fg = 0$ fast überall; dann gilt (1). Im Fall $A > 0$ und $B = \infty$ gilt (1) trivialerweise. Sei jetzt $0 < A < \infty$, $0 < B < \infty$. Setze $F := f/A$ und $G := g/B$. Dann ist

$$\int_X F^p \, d\mu = \frac{1}{A^p} \int_X f^p \, d\mu = 1, \qquad \int_X G^q \, d\mu = 1.$$

Sei $x \in X$ mit $0 < F(x) < \infty$ und $0 < G(x) < \infty$. Dann existieren $s, t \in \mathbb{R}$, so dass $F(x) = e^{s/p}$, $G(x) = e^{t/q}$. Da $\frac{1}{p} + \frac{1}{q} = 1$, ist

$$(*) \qquad \exp\left(\frac{s}{p} + \frac{t}{q}\right) \leq \frac{1}{p} e^s + \frac{1}{q} e^t$$

wegen der Konvexität von exp. Es folgt

$$F(x)G(x) \leq \frac{1}{p} F(x)^p + \frac{1}{q} G(x)^q$$

für alle $x \in X$ (vgl. auch die Ungleichung vor 5.2). Integration gibt

$$\int_X FG \, d\mu \leq \frac{1}{p} + \frac{1}{q} = 1,$$

und (1) folgt.

(2): Aus (1) folgt

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

Da $(p-1)q = p$, folgt durch Addition dieser Ungleichungen

$$(**) \qquad \textstyle\int (f+g)^p \, d\mu \leq \left(\textstyle\int (f+g)^p \, d\mu\right)^{1/q} \Big(\left(\textstyle\int f^p \, d\mu\right)^{1/q} + \left(\textstyle\int g^p \, d\mu\right)^{1/p} \Big).$$

Wir nehmen an, dass $\int (f+g)^p \, d\mu > 0$ und $\int f^p \, d\mu, \int g^p \, d\mu < \infty$; sonst gilt (2) trivialerweise. Die Konvexität von t^p für $0 < t < \infty$ zeigt, dass

$$\left(\frac{f+g}{2}\right)^p \leq \frac{1}{2}(f^p + g^p).$$

Damit ist auch $\int (f+g)^p \, d\mu < \infty$, und (2) folgt aus $(**)$, da $1 - \frac{1}{q} = \frac{1}{p}$. $\square$

Wir untersuchen noch den Gleichheitsfall in (1). Im Beweis gilt $\int FG\,d\mu = 1$ genau dann, wenn $F(x)G(x) = \frac{1}{p}F(x)^p + \frac{1}{q}G(x)^q$ für fast alle x. In $(*)$ gilt Gleichheit genau dann, wenn $s = t$. Also ist „$F^p = G^q$ fast überall" eine notwendige und hinreichende Bedingung für Gleichheit in $\int FG\,d\mu \le 1$, immer vorausgesetzt, dass $\int F^p\,d\mu = \int G^q\,d\mu = 1$. Für die ursprünglichen Funktionen f und g folgt: *Sind beide Faktoren auf der rechten Seite von* (1) *endlich, so gilt Gleichheit in* (1) *genau dann, wenn Konstanten* α *und* β *existieren, nicht beide gleich* 0*, so dass* $\alpha f^p = \beta g^p$ *fast überall.*

Im Folgenden bezeichnet X wieder einen beliebigen Massraum mit Mass μ.

5.4 Definition (L^p)
Ist $0 < p < \infty$ und f eine reelle messbare Funktion auf X, setze

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

und bezeichne mit $L^p(\mu)$ die Menge aller f mit $\|f\|_p < \infty$. Wir nennen $\|f\|_p$ die L^p-*Norm* von f; wir kommen in 5.8 auf diese Sprechweise zurück.

Ist μ das Lebesgue-Mass auf $\mathbb{R}^n$, so schreiben wir $L^p(\mathbb{R}^n)$ statt $L^p(\mu)$. Ist μ das Zählmass auf einer Menge A, so bezeichnet man den L^p-Raum üblicherweise mit $l^p(A)$, im Fall $A = \mathbb{N}$ auch einfach mit l^p. Ein Element von l^p kann als reelle Folge $x = (\xi_n)_{n\in\mathbb{N}}$ betrachtet werden, mit

$$\|x\|_p = \left(\textstyle\sum_{n=1}^{\infty}|\xi_n|^p\right)^{1/p}.$$

5.5 Definition (essentielles Supremum, L^∞)
Sei g eine messbare Funktion auf X mit Werten in $[0,\infty]$. Definiere das *essentielle Supremum* von g bzgl. μ durch

$$\operatorname{ess\,sup} g := \inf\{\alpha \in [0,\infty] : g(x) \le \alpha \text{ für } \mu\text{-fast alle } x \in X\};$$

es gilt $g(x) \le \operatorname{ess\,sup} g$ für μ-fast alle $x \in X$ (da $\{x : g(x) > \operatorname{ess\,sup} g\} = \bigcup_{n=1}^{\infty}\{x : g(x) > \operatorname{ess\,sup} g + \frac{1}{n}\}$ Mass null hat). Ist f ein reelle messbare Funktion auf X, setze

$$\|f\|_\infty := \operatorname{ess\,sup}|f|$$

und bezeichne mit $L^\infty(\mu)$ die Menge aller f mit $\|f\|_\infty < \infty$. Die Elemente von $L^\infty(\mu)$ heissen auch *essentiell beschränkte* Funktionen auf X.

Die Ungleichung $|f(x)| \le \lambda$ gilt genau dann für fast alle $x \in X$, wenn $\lambda \ge \|f\|_\infty$. Ist μ das Lebesgue-Mass auf $\mathbb{R}^n$, so schreiben wir $L^\infty(\mathbb{R}^n)$ statt $L^\infty(\mu)$. Für eine Menge A ist $l^\infty(A)$ die Menge aller *beschränkten* Funktionen auf A (hier bedeutet essentiell beschränkt dasselbe wie beschränkt, da das Zählmass einer nicht-leeren Menge positiv ist), und $l^\infty = l^\infty(\mathbb{N})$.

5.6 Satz
Sind p und q konjugierte Exponenten, $1 \leq p \leq \infty$, und ist $f \in L^p(\mu)$ und $g \in L^q(\mu)$, so ist $fg \in L^1(\mu)$ und

$$\|fg\|_1 \leq \|f\|_p \|g\|_q.$$

Beweis: Für $1 < p < \infty$ ist dies gerade die Hölder-Ungleichung, angewandt auf $|f|$ und $|g|$. Im Fall $p = \infty$ gilt

$$|f(x)g(x)| \leq \|f\|_\infty |g(x)|$$

für fast alle x, und Integration gibt die Behauptung. □

5.7 Satz
Seien $1 \leq p \leq \infty$ und $f, g \in L^p(\mu)$. Dann ist $f + g \in L^p(\mu)$ und

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

Beweis: Für $1 < p < \infty$ folgt dies aus der Minkowski-Ungleichung, da

$$\int_X |f + g|^p \, d\mu \leq \int_X (|f| + |g|)^p \, d\mu.$$

Für $p = 1$ oder $p = \infty$ folgt die Behauptung einfach aus der Ungleichung $|f + g| \leq |f| + |g|$. □

5.8 Bemerkung (L^p-Räume)
Fixiere p, $1 \leq p \leq \infty$. Für $f \in L^p(\mu)$ und $\alpha \in \mathbb{R}$ gilt offensichtlich $\alpha f \in L^p(\mu)$ und $\|\alpha f\|_p = |\alpha| \, \|f\|_p$. Zusammen mit 5.7 zeigt dies, dass $L^p(\mu)$ ein reeller Vektorraum ist. Es gilt $\|f\|_p = 0$ genau dann, wenn $f(x) = 0$ für fast alle $x \in X$; $\|\cdot\|_p$ definiert daher im Allgemeinen nur eine *Halbnorm* auf $L^p(\mu)$. Dementsprechend definiert $d(f, g) := \|f - g\|_p$ im Allgemeinen nur eine *Pseudometrik* auf $L^p(\mu)$; es gilt $d(f, g) = 0$ genau dann, wenn $f(x) = g(x)$ für fast alle x. Um eine Norm bzw. eine Metrik zu erhalten, betrachtet man wieder die schon nach 2.11 eingeführte Äquivalenzrelation: $f \sim g \Leftrightarrow f = g$ fast überall. Dann definieren

$$\|[f]\|_p := \|f\|_p \quad \text{und} \quad d([f], [g]) := d(f, g) = \|f - g\|_p$$

eine Norm und die zugehörige Metrik auf dem Raum der Äquivalenzklassen von Funktionen in $L^p(\mu)$. Es ist üblich, hierfür keine neue Notation einzuführen und eine Funktion $f \in L^p(\mu)$ stillschweigend als ihre Äquivalenzklasse aufzufassen, wenn der Kontext dies verlangt.

Ist $(f_n)_{n\in\mathbb{N}}$ eine Folge in $L^p(\mu)$ und gilt $\lim_{n\to\infty}\|f_n - f\|_p = 0$ für ein $f \in L^p(\mu)$, so sagen wir (f_n) *konvergiert nach f in $L^p(\mu)$*. Existiert für jedes $\epsilon > 0$ ein $N \in \mathbb{N}$, so dass $\|f_n - f_m\|_p < \epsilon$ für $n, m > N$, so heisst (f_n) eine *Cauchyfolge in $L^p(\mu)$*.

5.9 Satz (Vollständigkeit von $L^p(\mu)$)
Für $1 \le p \le \infty$ ist $L^p(\mu)$ (aufgefasst als metrischer Raum der Äquivalenzklassen von Funktionen in $L^p(\mu)$) ein vollständiger metrischer Raum, d.h. jede Cauchyfolge in $L^p(\mu)$ konvergiert (in $L^p(\mu)$).

Beweis: Sei zuerst $1 \le p < \infty$. Sei $(f_n)_{n\in\mathbb{N}}$ eine Cauchyfolge in $L^p(\mu)$. Es existiert eine Teilfolge (f_{n_i}), $n_1 < n_2 < \dots$, so dass

$$\|f_{n_{i+1}} - f_{n_i}\|_p < 2^{-i} \quad (i = 1, 2, \dots).$$

Setze dann

$$g_k := \sum_{i=1}^{k} |f_{n_{i+1}} - f_{n_i}|, \quad g := \sum_{i=1}^{\infty} |f_{n_{i+1}} - f_{n_i}|.$$

Nach 5.7 gilt $\|g_k\|_p \le \sum_{i=1}^k \|f_{n_{i+1}} - f_{n_i}\|_p < 1$. Mit 2.6 (Lemma von Fatou), angewandt auf die Folge $({g_k}^p)$, folgt $\|g\|_p \le 1$. Insbesondere ist $g(x) < \infty$ fast überall, so dass die Reihe

$$f_{n_1}(x) + \sum_{i=1}^{\infty} \big(f_{n_{i+1}}(x) - f_{n_i}(x)\big)$$

absolut konvergiert für fast alle $x \in X$. Für jedes solche x bezeichne $f(x)$ die Summe dieser Reihe. Da

$$f_{n_1} + \sum_{i=1}^{k-1} (f_{n_{i+1}} - f_{n_i}) = f_{n_k},$$

gilt $\lim_{i\to\infty} f_{n_i}(x) = f(x)$ fast überall. (Insbesondere ist f eine reelle messbare Funktion auf X im Sinn von 2.13.) Wir zeigen, dass $f_n \to f$ in $L^p(\mu)$. Sei $\epsilon > 0$. Es gibt ein $N \in \mathbb{N}$, so dass $\|f_n - f_m\|_p < \epsilon$ für $m, n > N$. Für jedes $m > N$ gilt somit

$$\int_X |f - f_m|^p \, d\mu \le \liminf_{i\to\infty} \int_X |f_{n_i} - f_m|^p \, d\mu \le \epsilon^p$$

nach 2.6. Dies zeigt, dass $f - f_m \in L^p(\mu)$, also $f \in L^p(\mu)$, und dass $\|f - f_m\|_p \to 0$ für $m \to \infty$. Dies beweist den Satz für $1 \le p < \infty$.

Der Beweis für $p = \infty$ ist einfacher. Sei $(f_n)_{n\in\mathbb{N}}$ eine Cauchyfolge in $L^\infty(\mu)$. Für $m, n \in \mathbb{N}$ sei

$$A_{m,n} := \big\{x \in X : |f_n(x) - f_m(x)| \leq \|f_n - f_m\|_\infty\big\}$$

und $A := \bigcap_{m,n\in\mathbb{N}} A_{m,n}$. Dann ist $A^c = \bigcup_{m,n}(A_{m,n})^c$ eine Nullmenge. Für alle $x \in A$ ist $(f_n(x))_{n\in\mathbb{N}}$ eine Cauchyfolge; sei $f(x)$ ihr Grenzwert. Für alle $m \in \mathbb{N}$ und $x \in A$ gilt

$$|f(x) - f_m(x)| \leq \sup_{n\geq m} |f_n(x) - f_m(x)| \leq \sup_{n\geq m} \|f_n - f_m\|_\infty.$$

Daraus folgt, dass $f - f_m$ auf A beschränkt ist (für m gross genug), also insbesondere $f - f_m \in L^\infty(\mu)$ und somit $f \in L^\infty(\mu)$, und dass $f_m \to f$ gleichmässig auf A, also insbesondere $f_m \to f$ in $L^\infty(\mu)$. □

Dieser Beweis zeigt auch:

5.10 Satz
Ist $1 \leq p \leq \infty$, und ist $(f_n)_{n\in\mathbb{N}}$ eine Cauchyfolge in $L^p(\mu)$ mit Grenzwert f, so enthält (f_n) eine Teilfolge, die fast überall punktweise gegen f konvergiert.

Eine Teilmenge D eines topologischen Raumes X heisst *dicht* in X, falls der Abschluss von D ganz X ist.

5.11 Satz
Sei S die Menge aller messbaren Treppenfunktionen s auf X mit

$$\mu\big(\{x \in X : s(x) \neq 0\}\big) < \infty.$$

Für $1 \leq p < \infty$ ist S dicht in $L^p(\mu)$.

Beweis: Es ist klar, dass $S \subset L^p(\mu)$. Sei $f \in L^p(\mu)$, $f \geq 0$, und sei $(s_n)_{n\in\mathbb{N}}$ wie in 1.12. Da $0 \leq s_n \leq f$, folgt $s_n \in L^p(\mu)$ und somit $s_n \in S$. Da $|f - s_n|^p \leq f^p$, folgt mit 2.10 (beschränkte Konvergenz), dass $\|f - s_n\|_p \to 0$ $(n \to \infty)$. Somit ist f im L^p-Abschluss von S. Der allgemeine Fall (f reell) folgt aus diesem Spezialfall mittels Zerlegung $f = f^+ - f^-$. □

5.12 Satz
Sei X ein lokal-kompakter Hausdorff-Raum, und sei μ ein Radonmass auf X. Für $1 \leq p < \infty$ ist $C_c(X)$ dicht in $L^p(\mu)$.

Beweis: Definiere S wie in 5.11. Für $s \in S$ und $\epsilon > 0$ existiert nach 4.6 (Lusin) ein $g \in C_c(X)$, so dass $g(x) = s(x)$ ausserhalb einer Menge B vom Mass $< \epsilon$, und o.B.d.A. ist $|g| \le \|s\|_\infty$ auf X. Mit $|g(x) - s(x)| \le |g(x)| + |s(x)|$ folgt

$$\|g - s\|_p = \left(\int_B |g - s|^p \, d\mu \right)^{1/p} \le \big(\epsilon (2\|s\|_\infty)^p\big)^{1/p} = 2\epsilon^{1/p} \|s\|_\infty.$$

Da S dicht ist in $L^p(\mu)$, folgt der Satz. □

Wir betrachten den Zusammenhang zwischen $L^p(\mathbb{R}^k)$ und $C_c(\mathbb{R}^k)$ etwas genauer. Für jedes $p \in [1, \infty]$ haben wir eine Metrik auf $C_c(\mathbb{R}^k)$; der Abstand zwischen f und g ist $\|f - g\|_p$. Hier braucht man nicht Äquivalenzklassen von Funktionen zu betrachten; für $f \ne g$ ist $\{f \ne g\}$ eine nicht-leere offene Menge, die also positives Lebesgue-Mass hat. Für $f \in C_c(\mathbb{R}^k)$ gilt ausserdem auch

$$\|f\|_\infty = \sup_{x \in \mathbb{R}^k} |f(x)|.$$

Für $1 \le p < \infty$ ist $C_c(\mathbb{R}^k)$ dicht in $L^p(\mathbb{R}^k)$ nach 5.12, und $L^p(\mathbb{R}^k)$ ist vollständig gemäss 5.9. Es gilt also: *$L^p(\mathbb{R}^k)$ ist die Vervollständigung von $C_c(\mathbb{R}^k)$ mit der L^p-Metrik.* Für $p = 1$ und $k = 1$ ergibt sich folgende Aussage, die zeigt, dass das Lebesgue-Integral die „richtige“ Verallgemeinerung des Riemann-Integrals ist: Versieht man $C_c(\mathbb{R})$ mit der durch

$$d(f, g) := \int_{-\infty}^{\infty} |f(t) - g(t)| \, dt$$

definierten Metrik, so ist die Vervollständigung dieses metrischen Raumes gerade der Raum der (Äquivalenzklassen von) Lebesgue-integrablen Funktionen auf $\mathbb{R}$. Für $p = \infty$ ist die L^∞-Vervollständigung von $C_c(\mathbb{R}^k)$ nicht $L^\infty(\mathbb{R}^k)$, sondern $C_0(\mathbb{R}^k)$, der Raum der stetigen Funktionen auf $\mathbb{R}^k$, die „im Unendlichen verschwinden“:

5.13 Definition ($C_0(X)$)
Eine reelle Funktion f auf einem lokal-kompakten Hausdorff-Raum X *verschwindet im Unendlichen*, falls für jedes $\epsilon > 0$ eine kompakte Menge C existiert, so dass $|f(x)| < \epsilon$ ausserhalb C; die Menge aller stetigen solchen Funktionen wird mit $C_0(X)$ bezeichnet.

Es ist klar, dass $C_c(X) \subset C_0(X)$ und dass jedes $f \in C_0(X)$ beschränkt ist.

5.14 Satz
Ist X ein lokal-kompakter Hausdorff-Raum, so ist $C_0(X)$ die Vervollständigung von $C_c(X)$ bzgl. der durch die Supremumsnorm

$$\|f\| = \sup_{x \in X} |f(x)|$$

definierten Metrik.

Beweis: Wir zeigen zuerst, dass $C_c(X)$ dicht in $C_0(X)$ ist. Sei $f \in C_0(X)$ und $\epsilon > 0$. Dann existiert eine kompakte Menge C, so dass $|f(x)| < \epsilon$ ausserhalb C. Nach 4.5 (Lemma von Urysohn) gibt es eine Funktion $g \in C_c(X)$ mit $\chi_C \leq g \leq 1$. Setze $h := fg$; dann ist $h \in C_c(X)$ und

$$\|f - h\| = \sup_{x \in X} |f(x)|\,|1 - g(x)| < \epsilon.$$

Es ist noch zu zeigen, dass $C_0(X)$ vollständig ist. Sei $(f_n)_{n\in\mathbb{N}}$ eine Cauchyfolge in $C_0(X)$. Dann konvergiert (f_n) gleichmässig gegen eine Funktion f, und f ist stetig. Sei $\epsilon > 0$. Dann existieren ein n und eine kompakte Menge C, so dass $\|f_n - f\| < \epsilon/2$ und $|f_n(x)| < \epsilon/2$ ausserhalb C, also $|f(x)| < \epsilon$ ausserhalb C. Dies zeigt, dass $f \in C_0(X)$. □

Es folgen noch einige Bemerkungen zu Banach- und Hilberträumen.

5.15 Definition (Banachraum, Hilbertraum)
Ein normierter Raum (wir betrachten hier nur den Fall von $\mathbb{R}$-Vektorräumen), der vollständig ist bzgl. der von der Norm induzierten Metrik, heisst ein *Banachraum*. Ein normierter Raum Y heisst ein *Prähilbertraum*, wenn seine Norm von einem Skalarprodukt (einer symmetrischen, positiv definiten Bilinearform) induziert wird, d.h. wenn $\|y\| = \sqrt{\langle y, y\rangle}$ für alle $y \in Y$. Dann gilt die Schwarz'sche Ungleichung

$$|\langle y, z\rangle| \leq \|y\|\|z\| \quad (y, z \in Y).$$

Ein vollständiger Prähilbertraum heisst ein *Hilbertraum*.

Satz 5.9 besagt also, dass $L^p(\mu)$ für $1 \leq p \leq \infty$ ein Banachraum ist. Auf $L^2(\mu)$ definiert

$$\langle f, g\rangle := \int_X fg\, d\mu$$

ein Skalarprodukt, da $fg \in L^1(\mu)$ nach 5.6, und es gilt

$$\langle f, f\rangle = \int_X f^2\, d\mu = \int_X |f|^2\, d\mu = \|f\|_2{}^2.$$

Somit ist $L^2(\mu)$ ein Hilbertraum. Die Schwarz'sche Ungleichung

$$|\langle f, g\rangle| \leq \|f\|_2\|g\|_2$$

folgt auch als Spezialfall von 5.6:

$$|\langle f, g\rangle| = \left|\int fg\,d\mu\right| \le \int |fg|\,d\mu = \|fg\|_1 \le \|f\|_2\|g\|_2.$$

Wir halten noch einige grundlegende Eigenschaften von Hilberträumen fest. Sei Y ein Hilbertraum mit Norm $\|\cdot\|$ und Skalarprodukt $\langle\cdot,\cdot\rangle$. Für festes $z \in Y$ ist die Abbildung $y \mapsto \langle y, z\rangle$ gleichmässig stetig, da

$$\big|\langle y, z\rangle - \langle y', z\rangle\big| = |\langle y - y', z\rangle| \le \|y - y\|\|z\|.$$

Ebenso ist $y \mapsto \|y\|$ gleichmässig stetig, da

$$\big|\|y\| - \|y'\|\big| \le \|y - y'\|.$$

Eine Menge $E \subset Y$ heisst *konvex*, falls mit $y, y' \in E$ auch $(1-t)y + ty'$ in E ist für alle $t \in [0,1]$.

5.16 Satz
Jede nicht-leere, abgeschlossene konvexe Menge $E \subset Y$ enthält ein eindeutig bestimmtes Element minimaler Norm.

Beweis: Für $y, y' \in Y$ gilt die „Parallelogramm-Identität“

$$\|y + y'\|^2 + \|y - y'\|^2 = 2\|y\|^2 + 2\|y'\|^2.$$

Sei $\delta := \inf\{\|y\| : y \in E\}$. Für $y, y' \in E$ ist $\frac{1}{2}(y + y') \in E$, also

$$\|y - y'\|^2 \le 2\|y\|^2 + 2\|y'\|^2 - 4\delta^2.$$

Für $\|y\| = \|y'\| = \delta$ folgt $y = y'$; dies beweist die Eindeutigkeitsaussage. Um die Existenz eines $y \in E$ mit $\|y\| = \delta$ zu zeigen, wähle $y_1, y_2, \ldots \in E$ mit $\lim_{i\to\infty} \|y_i\| = \delta$. Für $i, j \to \infty$ geht die rechte Seite der Ungleichung $\|y_i - y_j\|^2 \le 2\|y_i\|^2 + 2\|y_j\|^2 - 4\delta^2$ gegen 0; somit ist (y_i) eine Cauchyfolge. Da Y vollständig ist, konvergiert diese. Ihr Grenzwert y liegt in E, da E abgeschlossen ist, und $\|y\| = \lim_{i\to\infty} \|y_i\| = \delta$ wegen der Stetigkeit der Norm.

□

5.17 Satz
Ist L ein stetiges lineares Funktional auf Y, so existiert ein eindeutig bestimmtes Element $z \in Y$ mit der Eigenschaft, dass $Ly = \langle y, z\rangle$ für alle $y \in Y$.

Beweis: Ist $Ly = 0$ für alle y, setze $z := 0$.

Im andern Fall sei $M := \{y : Ly = 0\}$. Da L linear ist, ist M ein Unterraum, und da L stetig ist, ist M abgeschlossen. Sei y so, dass $Ly \neq 0$, d.h. $y \notin M$. Aus 5.16 folgt, dass ein $z' \in y + M$ mit minimaler Norm existiert. Da $y \notin M$, ist $z' \neq 0$. Wir zeigen, dass $\langle y', z' \rangle = 0$ für alle $y' \in M$ (d.h. dass z' im orthogonalen Komplement $M^\perp$ von M liegt): Sei o.B.d.A. $\|y'\| = 1$. Für alle $\alpha \in \mathbb{R}$ gilt $\|z'\|^2 \leq \|z' - \alpha y'\|^2$, also $0 \leq -2\alpha\langle y', z'\rangle + \alpha^2$; für $\alpha = \langle y', z'\rangle$ folgt $\langle y', z'\rangle = 0$. Sei jetzt $y \in Y$ beliebig und $y' := (Ly)z' - (Lz')y$. Da $Ly' = 0$, ist $y' \in M$, also $\langle y', z'\rangle = 0$. Damit folgt

$$Ly = \frac{1}{\|z'\|^2}\langle (Ly)z', z'\rangle = \frac{1}{\|z'\|^2}\langle (Lz')y, z'\rangle = \langle y, z\rangle$$

für $z := (Lz'/\|z'\|^2)z'$.

Um die Eindeutigkeit von z zu zeigen, nehmen wir an, dass $\langle y, z\rangle = \langle y, w\rangle$ für alle $y \in Y$. Dann ist $\langle y, z - w\rangle = 0$ für alle $y \in Y$, also insbesondere $\|z - w\|^2 = \langle z - w, z - w\rangle = 0$, d.h. $z = w$. □

6 Absolute Stetigkeit

Sei $\mathscr{M}$ eine σ-Algebra in einer Menge X. Wir betrachten zunächst ein *reelles* (oder *signiertes*) Mass λ auf $\mathscr{M}$, d.h. eine σ-additive Funktion $\lambda\colon \mathscr{M} \to \mathbb{R}$. Es gilt also

$$\lambda(E) = \sum_{i=1}^{\infty} \lambda(E_i)$$

für alle $E \in \mathscr{M}$ und jede *messbare Partition* $(E_i)_{i\in\mathbb{N}}$ von E (d.h. $E_i \in \mathscr{M}$, $E_i \cap E_j = \emptyset$ für $i \neq j$ und $E = \bigcup_{i=1}^{\infty} E_i$). Da $\lambda(E) < \infty$, konvergieren die Reihe auf der rechten Seite und alle ihre Umordnungen; die Reihe konvergiert deshalb auch absolut. Für λ gelten auch alle in Satz 1.14 genannten Eigenschaften ausser der Monotonie (3). Wir suchen ein (möglichst kleines) *Mass* μ auf $\mathscr{M}$ mit der Eigenschaft, dass

$$|\lambda(E)| \leq \mu(E)$$

für alle $E \in \mathscr{M}$. Für jedes solche μ müsste

$$\mu(E) = \sum_{i=1}^{\infty} \mu(E_i) \geq \sum_{i=1}^{\infty} |\lambda(E_i)|$$

gelten, für alle $E \in \mathscr{M}$ und jede messbare Partition (E_i) von E.

6.1 Definition (totales Variationsmass)
Sei λ ein reelles Mass auf $(X, \mathscr{M})$. Definiere $|\lambda| \colon \mathscr{M} \to [0, \infty]$ durch

$$|\lambda|(E) = \sup\Big\{\textstyle\sum_{i=1}^{\infty} |\lambda(E_i)| : (E_i) \text{ messbare Partition von } E\Big\};$$

$|\lambda|$ heisst *totale Variation* von λ oder genauer *totales Variationsmass* von λ (der erste Begriff bezeichnet häufig auch die Zahl $|\lambda|(X)$).

6.2 Satz
Die totale Variation $|\lambda|$ eines reellen Masses λ ist ein Mass.

Beweis: Sei $(E_i)_{i\in\mathbb{N}}$ eine messbare Partition von $E \in \mathscr{M}$. Wähle $t_i \in \mathbb{R}$ mit $t_i < |\lambda|(E_i)$. Dann besitzt jedes E_i eine messbare Zerlegung (A_{ij}), so dass

$$\sum_{j=1}^{\infty} |\lambda(A_{ij})| > t_i.$$

Da $(A_{ij})_{i,j\in\mathbb{N}}$ eine Zerlegung von E ist, folgt

$$\sum_i t_i \le \sum_{i,j} |\lambda(A_{ij})| \le |\lambda|(E).$$

Da dies für alle solchen t_i gilt, ist also $\sum_i |\lambda|(E_i) \le |\lambda|(E)$.
Um die umgekehrte Ungleichung zu zeigen, sei $(A_j)_{j\in\mathbb{N}}$ eine messbare Zerlegung von E. Für festes j ist $(A_j \cap E_i)$ eine Partition von A_j, und für festes i ist $(A_j \cap E_i)$ eine Partition von E_i. Es folgt

$$\textstyle\sum_j |\lambda(A_j)| = \sum_j \big|\sum_i \lambda(A_j \cap E_i)\big| \le \sum_{i,j} |\lambda(A_j \cap E_i)| \le \sum_i |\lambda|(E_i).$$

Da dies für alle solchen (A_j) gilt, ist $|\lambda|(E) \le \sum_i |\lambda|(E_i)$. Somit ist $|\lambda|$ σ-additiv. Es gilt auch $|\lambda|(\emptyset) = 0$. □

Offensichtlich ist $|\lambda|$ gerade das kleinste Mass auf $\mathscr{M}$ mit der Eigenschaft, dass $|\lambda(E)| \le |\lambda|(E)$ für alle $E \in \mathscr{M}$.

6.3 Satz
Für jedes reelle Mass λ auf $(X, \mathscr{M})$ gilt $|\lambda|(X) < \infty$.

Somit ist λ beschränkt: Für alle $E \in \mathscr{M}$ gilt

$$|\lambda(E)| \le |\lambda|(E) \le |\lambda|(X) < \infty.$$

Insbesondere ist das Mass $|\lambda|$ auch ein reelles Mass.

Beweis: Wir nehmen $|\lambda|(X) = \infty$ an. Sei $t := 2(1 + |\lambda(X)|)$. Da $t < |\lambda|(X)$, existiert eine messbare Partition $(E_i)_{i\in\mathbb{N}}$ von X, so dass $\sum_{i=1}^{\infty} |\lambda(E_i)| > t$. Setze $A := \bigcup\{E_i : \lambda(E_i) \geq 0\}$ und $B := X \setminus A$. Dann gilt $|\lambda(A)|+|\lambda(B)| > t$, o.B.d.A. $|\lambda(A)| > t/2 \geq 1$, und

$$|\lambda(B)| = |\lambda(X) - \lambda(A)| \geq |\lambda(A)| - |\lambda(X)| > \frac{t}{2} - |\lambda(X)| = 1.$$

Damit ist $X = A \cup B$ eine Partition mit $|\lambda(A)|, |\lambda(B)| > 1$. Da $|\lambda|$ additiv ist, gilt ausserdem $|\lambda|(A) = \infty$ oder $|\lambda|(B) = \infty$, o.B.d.A. $|\lambda|(B) = \infty$. Setze dann $A_1 := A$ und wiederhole das obige Argument mit $B_1 := B$ anstelle von X. Dies gibt eine Zerlegung $B_1 = A_2 \cup B_2$ mit $|\lambda(A_2)| > 1$ und $|\lambda|(B_2) = \infty$. Durch Iteration dieses Verfahrens erhält man paarweise disjunkte $A_1, A_2, \ldots$ mit $|\lambda(A_i)| > 1$ und $\sum_{i=1}^{\infty} \lambda(A_i) = \lambda\big(\bigcup_{i=1}^{\infty} A_i\big) \in \mathbb{R}$. Dies ist ein Widerspruch; die Reihe kann wegen $|\lambda(A_i)| > 1$ nicht konvergieren. Somit ist $|\lambda|(X) < \infty$. □

Sind λ, μ zwei reelle Masse auf der σ-Algebra $\mathscr{M}$, und ist $c \in \mathbb{R}$, so definieren

$$(\lambda + \mu)(E) := \lambda(E) + \mu(E),$$
$$(c\lambda)(E) := c\lambda(E)$$

($E \in \mathscr{M}$) reelle Masse $\lambda + \mu$ und $c\lambda$. Die Menge aller reellen Masse auf $\mathscr{M}$ bildet somit einen Vektorraum, $\|\lambda\| := |\lambda|(X)$ definiert ausserdem eine Norm auf diesem Raum. (Man kann zeigen, dass dies sogar einen Banachraum gibt.)

6.4 Definition (Jordan-Zerlegung)
Sei λ ein reelles Mass mit totaler Variation $|\lambda|$. Dann definieren

$$\lambda^+ := \frac{1}{2}(|\lambda| + \lambda) \quad \text{und} \quad \lambda^- := \frac{1}{2}(|\lambda| - \lambda)$$

zwei endliche Masse auf $\mathscr{M}$, die *positive Variation* und die *negative Variation* von λ. Dann heisst

$$\lambda = \lambda^+ - \lambda^-$$

die *Jordan-Zerlegung* von λ.

Es gilt auch $|\lambda| = \lambda^+ + \lambda^-$.

6.5 Definition (absolut stetige bzw. zueinander singuläre Masse)
Sei λ ein Mass oder ein reelles Mass auf einer σ-Algebra $\mathscr{M}$, und sei μ ein Mass auf $\mathscr{M}$. Wir nennen λ *absolut stetig* bzgl. μ, und schreiben

$$\lambda \ll \mu,$$

falls $\lambda(E) = 0$ für jedes $E \in \mathscr{M}$ mit $\mu(E) = 0$. Existiert ein $A \in \mathscr{M}$, so dass $\lambda(E) = \lambda(A \cap E)$ für alle $E \in \mathscr{M}$, so heisst λ *auf A konzentriert*; äquivalent gilt $\lambda(E) = 0$ für $E \cap A = \emptyset$. Sind λ_1, λ_2 zwei Masse oder reelle Masse auf $\mathscr{M}$, und existieren zwei disjunkte Mengen A_1 und A_2, so dass λ_i auf A_i konzentriert ist, $i = 1, 2$, so nennen wir λ_1 und λ_2 *zueinander singulär* und schreiben

$$\lambda_1 \perp \lambda_2.$$

6.6 Proposition
Seien $\lambda, \lambda_1, \lambda_2$ Masse oder reelle Masse auf einer σ-Algebra $\mathscr{M}$, und sei μ ein Mass auf $\mathscr{M}$. Dann gilt:

(1) *Ist λ auf A konzentriert, so ist $|\lambda|$ auf A konzentriert.*
(2) *Aus $\lambda_1 \perp \lambda_2$ folgt $|\lambda_1| \perp |\lambda_2|$.*
(3) *Gilt $\lambda_1 \perp \mu$ und $\lambda_2 \perp \mu$, so ist $\lambda_1 + \lambda_2 \perp \mu$.*
(4) *Gilt $\lambda_1 \ll \mu$ und $\lambda_2 \ll \mu$, so ist $\lambda_1 + \lambda_2 \ll \mu$.*
(5) *Aus $\lambda \ll \mu$ folgt $|\lambda| \ll \mu$.*
(6) *Gilt $\lambda_1 \ll \mu$ und $\lambda_2 \perp \mu$, so ist $\lambda_1 \perp \lambda_2$.*
(7) *Ist $\lambda \ll \mu$ und $\lambda \perp \mu$, so ist $\lambda = 0$.*

Beweis: (1): Ist $E \cap A = \emptyset$, und ist (E_j) eine messbare Zerlegung von E, so ist $E_j \cap A = \emptyset$ für alle j. Also gilt $|\lambda|(E) = 0$.

(2) folgt aus (1).

(3): Es existieren disjunkte Mengen A_1 und B_1, so dass λ_1 auf A_1 und μ auf B_1 konzentriert ist, und es existieren disjunkte Mengen A_2 und B_2, so dass λ_2 auf A_2 und μ auf B_2 konzentriert ist. Dann ist $\lambda_1 + \lambda_2$ auf $A := A_1 \cup A_2$ konzentriert, μ auf $B := B_1 \cap B_2$, und es gilt $A \cap B = \emptyset$.

(4) ist klar.

(5): Sei $\mu(E) = 0$, und sei (E_j) eine messbare Partition von E. Dann ist $\mu(E_j) = 0$ und somit $\lambda(E_j) = 0$ für alle j. Daraus folgt $|\lambda|(E) = 0$.

(6): Da $\lambda_2 \perp \mu$, ist λ_2 auf einer Menge A mit $\mu(A) = 0$ konzentriert. Da $\lambda_1 \ll \mu$, ist $\lambda_1(E) = 0$ für jedes $E \subset A$. Somit ist λ_1 auf A^c konzentriert.

(7): Aus (6) folgt $\lambda \perp \lambda$, also $\lambda = 0$. $\square$

6.7 Lemma
Ist μ ein σ-endliches Mass auf einem Massraum $(X, \mathscr{M})$, so existiert eine Funktion $w \in L^1(\mu)$, so dass $0 < w(x) < 1$ für alle $x \in X$.

Beweis: Nach Voraussetzung existieren $E_n \in \mathscr{M}$, $n = 1, 2, \ldots$, mit $\mu(E_n) < \infty$ und $\bigcup_{n=1}^{\infty} E_n = X$. Setze $w_n(x) := 0$ für $x \in E_n^c$ und

$$w_n(x) := 2^{-n} \frac{1}{1 + \mu(E_n)}$$

für $x \in E_n$. Dann hat $w := \sum_{n=1}^{\infty} w_n$ die gewünschten Eigenschaften. $\square$

Das Lemma ermöglicht es, μ durch das *endliche* Mass $\tilde{\mu}$ mit $d\tilde{\mu} = w\,d\mu$ zu ersetzen (d.h. $\tilde{\mu}(E) = \int_E w\,d\mu$, vgl. 2.7), wobei $\tilde{\mu}$ wegen der strikten Positvität von w genau dieselben Nullmengen wie μ besitzt.

6.8 Satz (Lebesgue-Radon-Nikodym)
Sei μ ein σ-endliches Mass auf $\mathscr{M}$, und sei λ ein reelles Mass auf $\mathscr{M}$.

(1) *Es existiert ein eindeutig bestimmtes Paar reeller Masse λ_a und λ_s auf $\mathscr{M}$, so dass*

$$\lambda = \lambda_a + \lambda_s, \quad \lambda_a \ll \mu \quad \textit{und} \quad \lambda_s \perp \mu.$$

Ist $\lambda \geq 0$, so sind auch $\lambda_a, \lambda_s \geq 0$.

(2) *Es existiert ein eindeutig bestimmtes Element $h \in L^1(\mu)$, so dass*

$$\lambda_a(E) = \int_E h\,d\mu$$

für alle $E \in \mathscr{M}$.

Das Paar (λ_a, λ_s) heisst die *Lebesgue-Zerlegung* von λ bzgl. μ. Die Eindeutigkeit dieser Zerlegung folgt leicht: Ist (λ_a', λ_s') ein weiteres solches Paar, so gilt $\lambda_a' - \lambda_a = \lambda_s - \lambda_s'$, $\lambda_a' - \lambda_a \ll \mu$ (6.6(4)), $\lambda_s - \lambda_s' \perp \mu$ (6.6(3)), also $\lambda_a' = \lambda_a$ und $\lambda_s' = \lambda_s$ (6.6(7)). Die Aussage von 6.8(2) ist der *Satz von Radon-Nikodym*. Die Eindeutigkeit der Äquivalenzklasse $h \in L^1(\mu)$ folgt sofort aus Satz 2.15(2). Die Funktion h heisst *Radon-Nikodym-Ableitung* von λ_a bzgl. μ; man schreibt dafür auch einfach $d\lambda_a = h\,d\mu$. Umgekehrt gilt: Ist $h \in L^1(\mu)$ beliebig, so definiert $\lambda(E) := \int_E h\,d\mu$ ein reelles Mass auf $\mathscr{M}$ (vgl. Satz 2.7), und $\lambda \ll \mu$.

Beweis von 6.8: Sei zuerst $\lambda \geq 0$. Wähle $w \in L^1(\mu)$ wie in 6.7. Dann definiert $d\varphi := d\lambda + w\,d\mu$ ein endliches Mass φ auf $\mathscr{M}$. Für $f \in L^2(\varphi)$ gilt nach der Schwarz'schen Ungleichung

$$\left|\int_X f\,d\lambda\right| \leq \int_X |f|\,d\lambda \leq \int_X |f|\,d\varphi \leq \|f\|_{L^2(\varphi)}\,\varphi(X)^{1/2}.$$

Da $\varphi(X) < \infty$, folgt daraus, dass

$$f \mapsto \int_X f\,d\lambda$$

ein *stetiges* lineares Funktional auf $L^2(\varphi)$ ist. Nach 5.17 existiert ein $g \in L^2(\varphi)$, so dass

$$(*) \qquad \int_X f\,d\lambda = \langle f, g\rangle_{L^2(\varphi)} = \int_X fg\,d\varphi$$

für alle $f \in L^2(\varphi)$. Für $f = \chi_E$, $E \in \mathscr{M}$, folgt $\lambda(E) = \int_E g\, d\varphi$. Da $0 \leq \lambda \leq \varphi$, gilt also

$$0 \leq \int_E g\, d\varphi \leq \varphi(E)$$

für alle $E \in \mathscr{M}$. Dies impliziert, dass $0 \leq g(x) \leq 1$ für φ-fast alle $x \in X$. Mit $(*)$ und $d\varphi = d\lambda + w\, d\mu$ folgt

$$(**) \qquad \int_X (1-g) f\, d\lambda = \int_X fgw\, d\mu$$

für alle nichtnegativen $f \in L^2(\varphi)$ (s. 2.7). Setze $A := \{x : 0 \leq g(x) < 1\}$, $B := \{x : g(x) = 1\}$, und definiere Masse λ_a und λ_s auf $\mathscr{M}$ durch

$$\lambda_a(E) := \lambda(A \cap E), \quad \lambda_s(E) := \lambda(B \cap E).$$

Für $f = \chi_B$ wird $(**)$ zu $0 = \int_B w\, d\mu$. Da $w(x) > 0$ für alle x, folgt $\mu(B) = 0$. Somit ist $\lambda_s \perp \mu$. Für $n \in \mathbb{N}$ und $E \in \mathscr{M}$ ist $f = (1 + g + \ldots + g^n)\chi_E$ in $L^2(\varphi)$, da $0 \leq g \leq 1$ φ-fast überall und $\varphi(E) < \infty$; $(**)$ wird dann zu

$$\int_E (1 - g^{n+1})\, d\lambda = \int_E g(1 + g + \ldots + g^n) w\, d\mu.$$

Für $x \in B$ ist $1 - g^{n+1}(x) = 0$, und für $x \in A$ konvergiert $1 - g^{n+1}(x)$ monoton gegen 1 $(n \to \infty)$. Somit ist

$$\lim_{n\to\infty} \int_E (1 - g^{n+1})\, d\lambda = \lambda(A \cap E) = \lambda_a(E).$$

Andererseits folgt mit 2.4 (monotone Konvergenz)

$$\lim_{n\to\infty} \int_E g(1 + g + \ldots + g^n) w\, d\mu = \int_E h\, d\mu$$

für eine nichtnegative messbare Funktion h. Es gilt also $\lambda_a(E) = \int_E h\, d\mu$ für alle $E \in \mathscr{M}$. Da $\int_X h\, d\mu = \lambda_a(X) < \infty$, ist $h \in L^1(\mu)$. Da offensichtlich $\lambda_a \ll \mu$, ist der Satz im Fall $\lambda \geq 0$ bewiesen.

Für ein allgemeines reelles Mass λ wendet man den vorangehenden Fall auf λ^+ und λ^- an (s. 6.4). □

Ist λ nicht ein reelles Mass sondern ein σ-endliches *Mass*, so gilt 6.8 praktisch immer noch. Man kann dann $X = \bigcup_{n\in\mathbb{N}} X_n$ schreiben, wobei $\mu(X_n) < \infty$ und $\lambda(X_n) < \infty$ für $n \in \mathbb{N}$. Die Lebesgue-Zerlegungen der endlichen Masse $\lambda \llcorner X_n$ geben eine Lebesgue-Zerlegung von λ, und es existiert immer noch eine messbare Funktion h (wobei jetzt $h \geq 0$), so dass $\lambda_a(E) = \int_E h\, d\mu$. Im Allgemeinen ist h aber nicht in $L^1(\mu)$, obwohl $\int_{X_n} h\, d\mu < \infty$ für alle n. Ohne die Voraussetzung der σ-Endlichkeit müssen (1) und (2) nicht mehr gelten.

6.9 Satz
Sei μ ein Mass auf $\mathscr{M}$, und sei λ ein reelles Mass auf $\mathscr{M}$. Dann sind die folgenden beiden Bedingungen äquivalent:

(1) $\lambda \ll \mu$.
(2) *Für jedes $\epsilon > 0$ gibt es ein $\delta > 0$, so dass $|\lambda(E)| < \epsilon$ für alle $E \in \mathscr{M}$ mit $\mu(E) < \delta$.*

Ist λ ein unbeschränktes *Mass*, so muss aus (1) nicht (2) folgen; beispielsweise sei μ das Lebesgue-Mass auf $(0,1)$ und

$$\lambda(E) = \int_E t^{-1} dt$$

für alle Lebesgue-messbaren $E \subset (0,1)$.

Beweis: Aus (2) folgt (1): Ist $\mu(E) = 0$, so ist $\mu(E) < \delta$ für alle $\delta > 0$, also $|\lambda(E)| < \epsilon$ für alle $\epsilon > 0$, also $\lambda(E) = 0$.

(1) impliziert (2): Wir nehmen an, (2) sei falsch. Dann existieren ein $\epsilon > 0$ und Mengen $E_n \in \mathscr{M}$, $n \in \mathbb{N}$, so dass $\mu(E_n) < 2^{-n}$ und $|\lambda(E_n)| \geq \epsilon$. Dann ist auch $|\lambda|(E_n) \geq \epsilon$. Setze

$$A_n := \bigcup_{i=n}^{\infty} E_i, \quad A := \bigcap_{n=1}^{\infty} A_n.$$

Es folgt $\mu(A_n) < 2^{-n+1}$, $A_n \supset A_{n+1}$, und somit $\mu(A) = 0$ und

$$|\lambda|(A) = \lim_{n\to\infty} |\lambda|(A_n) \geq \epsilon > 0,$$

da $|\lambda|(A_n) \geq |\lambda|(E_n)$. Somit ist $|\lambda|$ nicht absolut stetig bzgl. μ, also λ nicht absolut stetig bzgl. μ (6.6(5)). □

6.10 Satz
Sei λ ein reelles Mass auf $(\mathscr{M}, X)$.

(1) *Es existiert eine messbare Funktion h mit Werten in $\{-1,1\}$, so dass $d\lambda = h\, d|\lambda|$.*
(2) *Es existiert eine messbare Zerlegung $X = A \cup B$, so dass*

$$\lambda^+(E) = \lambda(A \cap E) \quad \textit{und} \quad \lambda^-(E) = -\lambda(B \cap E)$$

für alle $E \in \mathscr{M}$.

Hier ist $\lambda = \lambda^+ - \lambda^-$ die Jordan-Zerlegung (s. 6.4). Das Paar (A,B) heisst eine *Hahn-Zerlegung* von X bzgl. λ. Man beachte, dass $\lambda(E) \geq 0$ für $E \subset A$ und $\lambda(E) \leq 0$ für $E \subset B$, d.h. „A trägt alle positive Masse von λ und B trägt alle negative Masse von λ“.

Beweis: (1): Es gilt $\lambda \ll |\lambda|$, nach 6.8 existiert somit ein $h \in L^1(|\lambda|)$, so dass $d\lambda = h\, d|\lambda|$. Sei $A_r := \{x : |h(x)| < r\}$, $r > 0$, und sei (E_j) eine messbare Zerlegung von A_r. Dann ist

$$\sum_j |\lambda(E_j)| = \sum_j \left| \int_{E_j} h\, d|\lambda| \right| \leq \sum_j r|\lambda|(E_j) = r|\lambda|(A_r),$$

also $|\lambda|(A_r) \leq r|\lambda|(A_r)$. Für $r < 1$ folgt $|\lambda|(A_r) = 0$. Somit ist $|h| \geq 1$ fast überall. Andererseits gilt für alle $E \in \mathscr{M}$

$$\left| \int_E h\, d|\lambda| \right| = |\lambda(E)| \leq |\lambda|(E),$$

und daraus folgt, dass $|h| \leq 1$ fast überall.

(2): Sei h wie in (1). Setze $A := \{x : h(x) = 1\}$ und $B := X \setminus A$. Da $\lambda^+ = \frac{1}{2}(|\lambda| + \lambda)$ und $d\lambda = h\, d|\lambda|$, folgt

$$\lambda^+(E) = \int_E \frac{1+h}{2}\, d|\lambda| = \int_E \chi_A\, d|\lambda| = \int_{E\cap A} h\, d|\lambda| = \lambda(E \cap A).$$

Da $\lambda(E) = \lambda(E \cap A) + \lambda(E \cap B)$ und $\lambda = \lambda^+ - \lambda^-$, folgt $\lambda^-(E) = \lambda^+(E) - \lambda(E) = \lambda(E \cap A) - \lambda(E) = -\lambda(E \cap B)$. □

Aus 6.10(2) folgt eine Minimalitätseigenschaft der Jordan-Zerlegung $\lambda = \lambda^+ - \lambda^-$ eines reellen Masses λ: *Gilt* $\lambda = \mu_1 - \mu_2$ *für Masse* μ_1, μ_2, *so ist* $\lambda^+ \leq \mu_1$ *und* $\lambda^- \leq \mu_2$*:*

$$\begin{aligned} \lambda^+(E) &= \lambda(E \cap A) \leq \mu_1(E \cap A) \leq \mu_1(E), \\ \lambda^-(E) &= -\lambda(E \cap B) \leq \mu_2(E \cap B) \leq \mu_2(E). \end{aligned}$$

7 Differentiation

Der folgende einfach zu beweisende Satz motiviert die nachfolgenden Definitionen.

7.1 Satz
Sei λ *ein reelles Mass auf* $(\mathbb{R}, \mathscr{B})$, μ *das Lebesgue-Mass auf* $\mathbb{R}$. *Definiere* $f\colon \mathbb{R} \to \mathbb{R}$ *durch* $f(x) := \lambda((-\infty, x))$. *Für* $x \in \mathbb{R}$ *und* $\alpha \in \mathbb{R}$ *sind folgende Aussagen äquivalent:*

(1) *f ist differenzierbar in x und $f'(x) = \alpha$.*
(2) *Für jedes $\epsilon > 0$ existiert ein $\delta > 0$, so dass*

$$\left| \frac{\lambda(I)}{\mu(I)} - \alpha \right| < \epsilon$$

für alle offenen Intervalle $I \subset \mathbb{R}$ mit $x \in I$ und $L(I) < \delta$.

Dies legt nahe, die Ableitung von λ (bzgl. μ) an der Stelle x als den Limes von $\lambda(I)/\mu(I)$ für $I \to \{x\}$ zu definieren. In höheren Dimensionen kann man statt Intervalle offene Bälle $B(x,r) = \{y \in \mathbb{R}^k : |y - x| < r\}$ $(r > 0)$ verwenden.

7.2 Definition ($D\lambda$, $M\lambda$)
Sei λ ein reelles Mass auf $(\mathbb{R}^k, \mathscr{B})$, μ das Lebesgue-Mass auf $\mathbb{R}^k$. Für $x \in \mathbb{R}^k$ und $r > 0$ sei

$$(Q_r\lambda)(x) := \frac{\lambda(B(x,r))}{\mu(B(x,r))}.$$

Dann definiert

$$(D\lambda)(x) := \lim_{r \to 0} (Q_r\lambda)(x)$$

die *symmetrische Ableitung* von λ in den Punkten $x \in \mathbb{R}^k$, wo dieser Limes existiert. Die *Maximalfunktion* $M\lambda\colon \mathbb{R}^k \to [0, \infty]$ von λ ist definiert durch

$$(M\lambda)(x) := \sup_{0<r<\infty} (Q_r|\lambda|)(x).$$

Man sieht leicht, dass $M\lambda$ von unten halbstetig und somit messbar ist: Sei $\alpha \geq 0$, $E := \{M\lambda > \alpha\}$ und $x \in E$. Dann existiert ein $r > 0$, so dass $|\lambda|(B(x,r)) = t\mu(B(x,r))$ für ein $t > \alpha$, und es existiert ein $\delta > 0$, so dass $tr^k > \alpha(r + \delta)^k$. Für $y \in B(x,\delta)$ ist $B(x,r) \subset B(y, r + \delta)$; es folgt

$$|\lambda|(B(y, r + \delta)) \geq |\lambda|(B(x,r)) = t\mu(B(x,r)) > \alpha\mu(B(y, r + \delta)).$$

Dies zeigt, dass $B(x,\delta) \subset E$; also ist E offen.

7.3 Lemma ($3r$-Überdeckung)
Sei W die Vereinigung endlich vieler offener Bälle $B(x_i, r_i) \subset \mathbb{R}^k$, $i = 1, \ldots, N$. Dann existiert $S \subset \{1, \ldots, N\}$, so dass

$$B(x_i, r_i) \cap B(x_j, r_j) = \emptyset \quad \textit{für } i, j \in S,\ i \neq j,$$

und $W \subset \bigcup_{i \in S} B(x_i, 3r_i)$.

Beweis: 0.B.d.A. seien die $B_i = B(x_i, r_i)$ so geordnet, dass $r_1 \geq r_2 \geq \ldots \geq r_N$. Setze $i_1 := 1$ und entferne alle B_j mit $j > i_1$, die B_{i_1} schneiden. Sind noch Bälle B_j mit $j > i_1$ übrig, so sei B_{i_2} der erste davon; streiche dann alle B_j mit $j > i_2$, die B_{i_2} schneiden. Sind noch Bälle B_j mit $j > i_2$ übrig, so sei B_{i_3} der erste davon; streiche dann alle B_j mit $j > i_3$, die B_{i_3} schneiden. Fahre so weiter, bis für ein n keine Bälle B_j mit $j > i_n$ mehr übrig sind. Setze $S := \{i_1, \ldots i_n\}$. Die B_i mit $i \in S$ sind paarweise disjunkt nach Konstruktion. Jeder Ball B_j mit $j \notin S$ trifft ein B_i mit $i < j$ und $i \in S$; wegen $r_i \geq r_j$ ist dann $B_j \subset B(x_i, 3r_i)$. Dies zeigt, dass $W \subset \bigcup_{i \in S} B(x_i, 3r_i)$. □

7.4 Satz

Sei λ ein reelles Mass auf $(\mathbb{R}^k, \mathscr{B})$, μ das Lebesgue-Mass auf $\mathbb{R}^k$ und $\alpha > 0$. Dann ist

$$\alpha\mu\big(\{x \in \mathbb{R}^k : (M\lambda)(x) > \alpha\}\big) \leq 3^k \|\lambda\|,$$

wobei $\|\lambda\| := |\lambda|(\mathbb{R}^k)$.

Beweis: Sei C eine kompakte Teilmenge der offenen Menge $\{M\lambda > \alpha\}$. Jedes $x \in C$ ist Zentrum eines offenen Balles B mit $|\lambda|(B) > \alpha\mu(B)$. Wegen der Kompaktheit von C und Lemma 7.3 existieren paarweise disjunkte solche Bälle $B_1, \ldots, B_n$, so dass

$$\alpha\mu(C) \leq 3^k \alpha \sum_{i=1}^{n} \mu(B_i) \leq 3^k \sum_{i=1}^{n} |\lambda|(B_i) \leq 3^k \|\lambda\|.$$

Da jede offene Menge in $\mathbb{R}^k$ Vereinigung einer aufsteigenden Folge kompakter Mengen ist, folgt die Behauptung. □

Ist $f \in L^1(\mathbb{R}^k)$ und $\alpha > 0$, so gilt

$$\alpha\mu\big(\{|f| > \alpha\}\big) \leq \|f\|_1,$$

da $\alpha\mu(E) \leq \int_E |f|\, d\mu \leq \int_{\mathbb{R}^k} |f|\, d\mu = \|f\|_1$ für $E = \{|f| > \alpha\}$.

7.5 Definition (schwach $L^1(\mathbb{R}^k)$, Mf)

Eine messbare Funktion f gehört zu *schwach* $L^1(\mathbb{R}^k)$, falls

$$\alpha \mapsto \alpha\mu\big(\{|f| > \alpha\}\big)$$

auf $(0, \infty)$ beschränkt ist. Jeder Funktion $f \in L^1(\mathbb{R}^k)$ ordnen wir ihre *Maximalfunktion* $Mf \colon \mathbb{R}^k \to [0, \infty]$ zu,

$$(Mf)(x) := \sup_{0<r<\infty} \frac{1}{\mu(B(x,r))} \int_{B(x,r)} |f|\, d\mu.$$

Ordnen wir $f \in L^1(\mathbb{R}^k)$ das durch $d\lambda = f\,d\mu$ definierte reelle Mass λ zu, so folgt mit Satz 6.10(1), dass $h\,d|\lambda| = d\lambda = f\,d\mu$ für eine messbare Funktion h mit Werten in $\{-1, 1\}$, also $d|\lambda| = hf\,d\mu$. Da $|\lambda|$ und μ Masse sind, gilt $hf \geq 0$ μ-fast überall und somit $d|\lambda| = |f|\,d\mu$. Demzufolge ist $Mf = M\lambda$. Satz 7.4 besagt dann, dass der „Maximaloperator“ M den Raum $L^1(\mathbb{R}^k)$ in schwach $L^1(\mathbb{R}^k)$ abbildet: *Für jedes $f \in L^1(\mathbb{R}^k)$ und jedes $\alpha > 0$ ist*

$$\alpha\,\mu\big(\{Mf > \alpha\}\big) \leq 3^k\|f\|_1.$$

7.6 Definition (Lebesgue-Punkt)
Sei $f \in L^1(\mathbb{R}^k)$. Ein Punkt $x \in \mathbb{R}^k$ heisst ein *Lebesgue-Punkt* von f, falls

$$\lim_{r\to 0} \frac{1}{\mu(B(x,r))} \int_{B(x,r)} |f(y) - f(x)|\,d\mu(y) = 0.$$

Ist f stetig an der Stelle x, so ist x ein Lebesgue-Punkt.

7.7 Satz (Lebesgue)
Ist $f \in L^1(\mathbb{R}^k)$, so ist fast jeder Punkt $x \in \mathbb{R}^k$ ein Lebesgue-Punkt von f.

Beweis: Für $x \in \mathbb{R}^k$ und $r > 0$ sei

$$(T_r f)(x) := \frac{1}{\mu(B(x,r))} \int_{B(x,r)} |f(y) - f(x)|\,d\mu(y)$$

und $(Tf)(x) := \limsup_{r\to 0}(T_r f)(x)$. Es ist zu zeigen, dass Tf fast überall null ist. Sei $n \in \mathbb{N}$. Nach Satz 5.12 existiert ein $g \in C_c(\mathbb{R}^k)$, so dass $\|f-g\|_1 < 1/n$. Setze $h := f - g$. Wegen

$$|f(y) - f(x)| \leq |h(y)| + |h(x)| + |g(y) - g(x)|$$

ist $T_r f \leq Mh + |h| + T_r g$ für alle $r > 0$. Da $Tg = 0$ aufgrund der Stetigkeit von g, folgt $Tf \leq Mh + |h|$. Für $\alpha > 0$ ist somit

$$\{Tf > 2\alpha\} \subset E_{\alpha,n} := \{Mh > \alpha\} \cup \{|h| > \alpha\}.$$

Es gilt $\alpha\mu(\{Mh > \alpha\}) \leq 3^k\|h\|_1$ und $\alpha\mu(\{|h| > \alpha\}) \leq \|h\|_1$ (s. Bemerkungen nach 7.5 und 7.4). Da $\|h\|_1 < 1/n$, folgt

$$\alpha\mu(E_{\alpha,n}) \leq (3^k + 1)/n.$$

Es gilt $\{Tf > 2\alpha\} \subset \bigcap_{n=1}^{\infty} E_{\alpha,n}$; dieser Schnitt ist eine Nullmenge. Das Lebesgue-Mass ist vollständig, $\{Tf > 2\alpha\}$ ist also messbar und hat Mass null. Da dies für alle $\alpha > 0$ gilt, ist $Tf = 0$ fast überall. □

7.8 Satz
Sei λ ein reelles Mass auf $(\mathbb{R}^k, \mathscr{B})$, und sei λ absolut stetig bzgl. dem Lebesgue-Mass μ, mit Radon-Nikodym-Ableitung f. Dann ist $D\lambda = f$ μ-fast überall, und

$$\lambda(E) = \int_E (D\lambda)\, d\mu$$

für jede Borelmenge $E \subset \mathbb{R}^k$.

Beweis: Nach 6.8(2) (Radon-Nikodym) ist

$$\lambda(E) = \int_E f\, d\mu$$

für $E \in \mathscr{B}$. In jedem Lebesgue-Punkt x von f gilt somit

$$\frac{\lambda(B(x,r))}{\mu(B(x,r))} = \frac{1}{\mu(B(x,r))} \int_{B(x,r)} f\, d\mu \to f(x) \quad (r \to 0),$$

d.h. der Limes $(D\lambda)(x)$ existiert und ist gleich $f(x)$. Mit 7.7 folgt der Satz. □

7.9 Satz
Sei x ein Lebesgue-Punkt von f, und sei $(E_i)_{i\in\mathbb{N}}$ eine Folge von Borelmengen in $\mathbb{R}^k$ mit der Eigenschaft, dass $E_i \subset B(x, r_i)$ und $\mu(E_i) \geq \alpha\mu(B(x, r_i))$ für alle $i \in \mathbb{N}$, wobei $r_i \to 0$ $(i \to \infty)$ und $\alpha > 0$. Dann gilt

$$\lim_{i\to\infty} \frac{1}{\mu(E_i)} \int_{E_i} |f - f(x)|\, d\mu = 0 \quad \textit{und} \quad f(x) = \lim_{i\to\infty} \frac{1}{\mu(E_i)} \int_{E_i} f\, d\mu.$$

Beweis: Für $i \in \mathbb{N}$ ist

$$\frac{\alpha}{\mu(E_i)} \int_{E_i} |f - f(x)|\, d\mu \leq \frac{1}{\mu(B(x, r_i))} \int_{B(x,r_i)} |f - f(x)|\, d\mu.$$

Da x ein Lebesgue-Punkt von f ist, konvergiert für $i \to \infty$ die rechte Seite und somit auch die linke Seite gegen 0. Die zweite behauptete Identität folgt mit $f(x) - |f(y) - f(x)| \leq f(y) \leq f(x) + |f(y) - f(x)|$ aus der ersten. □

7.10 Definition (Dichte)
Sei E eine Lebesgue-messbare Teilmenge von $\mathbb{R}^k$. Dann definiert

$$\Theta(E, x) = \lim_{r\to 0} \frac{\mu(E \cap B(x,r))}{\mu(B(x,r))}$$

die *Dichte* von E in den Punkten $x \in \mathbb{R}^k$, wo dieser Limes existiert.

Aus 7.8 oder 7.9 für $f = \chi_E$ folgt, dass $\Theta(E, x) = 1$ für fast alle $x \in E$ und $\Theta(E, x) = 0$ für fast alle $x \in E^c$.

Wir wenden uns nun dem Hauptsatz der Differential- und Integralrechnung für das Lebesgue-Integral zu. Der folgende Satz bildet den ersten Teil:

7.11 Satz
Ist $f \in L^1(\mathbb{R})$ und $F(x) = \int_{-\infty}^x f\, d\mu$ für $x \in \mathbb{R}$, so gilt $F'(x) = f(x)$ in jedem Lebesgue-Punkt von f, also insbesondere für fast alle x.

Beweis: Sei $(\delta_i)_{i\in\mathbb{N}}$ eine Nullfolge, $\delta_i > 0$. Ist x ein Lebesgue-Punkt von f und $E_i = [x, x+\delta_i]$, so stimmt $f(x)$ nach 7.9 mit der rechtsseitigen Ableitung von F an der Stelle x überein. Mit $E_i = [x-\delta_i, x]$ folgt dasselbe für die linksseitige Ableitung. □

Der zweite Teil des Hauptsatzes der Differential- und Integralrechnung besagt, dass

$$(*) \quad f(x) - f(a) = \int_a^x f'(t)\, dt \quad \text{für alle } x \in [a, b]$$

unter geeigneten Voraussetzungen an $f\colon [a, b] \to \mathbb{R}$. In der elementaren Version dieses Satzes setzt man voraus, dass f stetig differenzierbar sei. Unter welchen schwächeren Voraussetzungen an f gilt $(*)$ noch? Wir betrachten zwei Beispiele.

Sei zuerst $f(x) = x^2 \sin(x^{-2})$ für $x \neq 0$, $f(0) = 0$. Dann ist f *überall* differenzierbar, aber

$$\int_0^1 |f'(t)|\, dt = \infty,$$

also $f' \notin L^1([0, 1])$.

Das zweite Beispiel zeigt insbesondere, dass $(*)$ selbst dann nicht gelten muss, wenn f stetig auf $[a, b]$ und fast überall differenzierbar ist mit $f' \in L^1([a, b])$. Wähle $1 =: \delta_0 > \delta_1 > \ldots$ mit $\delta_n \to 0$ $(n \to \infty)$. Setze $E_0 := [0, 1]$. Sei E_n für $n \geq 0$ so konstruiert, dass E_n die Vereinigung von 2^n paarweise disjunkten abgeschlossenen Intervallen der Länge $2^{-n}\delta_n$ ist. Entferne aus jedem dieser 2^n Intervalle ein zentriertes offenes Intervall, so dass jedes der verbleibenden 2^{n+1} Intervalle die Länge $2^{-(n+1)}\delta_{n+1}$ hat; sei dann E_{n+1} die Vereinigung dieser 2^{n+1} Intervalle. Diese induktive Konstruktion liefert eine Folge kompakter Mengen $E_0 \subset E_1 \subset E_2 \ldots$ mit Lebesgue-Mass $\mu(E_n) = \delta_n$. Dann ist $E := \bigcap_{n=1}^\infty E_n$ kompakt, und $\mu(E) = 0$. (Für $\delta_n = (2/3)^n$ ist E die übliche Cantor-1/3-Menge.) Setze

$$g_n := \frac{1}{\delta_n}\chi_{E_n} \quad \text{und} \quad f_n(x) := \int_0^x g_n(t)\, dt.$$

Es gilt $f_n(0) = 0$, $f_n(1) = 1$, und f_n ist monoton auf $[0,1]$ und konstant auf jedem Intervall in ${E_n}^c$. Ist I eines der 2^n Teilintervalle von E_n, so gilt

$$\int_I g_n(t)\,dt = 2^{-n} = 2 \cdot 2^{-(n+1)} = \int_I g_{n+1}(t)\,dt.$$

Daraus folgt, dass $f_{n+1}(x) = f_n(x)$ für $x \notin E_n$ und

$$|f_n(x) - f_{n+1}(x)| \le \int_I |g_n - g_{n+1}|\,dt \le 2 \cdot 2^{-n}$$

für $x \in E_n$ (wähle I so, dass $x \in I$). Somit konvergiert (f_n) gleichmässig gegen eine stetige monotone Funktion f mit $f(0) = 0$, $f(1) = 1$ und $f'(x) = 0$ für alle $x \notin E$. Da $\mu(E) = 0$, ist $f' = 0$ fast überall, $(*)$ gilt also nicht.

Wir nehmen jetzt an, für eine Funktion $f\colon [a,b] \to \mathbb{R}$ sei $f' \in L^1$ und es gelte

$$f(x) - f(a) = \int_a^x f'(t)\,dt \quad \text{für alle } x \in [a,b].$$

Sei dann λ das durch $d\lambda = f'\,d\mu$ definierte reelle Mass. Es gilt $|\lambda| \ll \mu$ (6.6(5)). Aus Satz 6.9 folgt: Für jedes $\epsilon > 0$ existiert ein $\delta > 0$, so dass $|\lambda|(E) < \epsilon$ für jede Vereinigung $E = \bigcup_{i=1}^n (\alpha_i, \beta_i)$ von paarweise disjunkten Intervallen mit $\mu(E) = \sum_{i=1}^n (\beta_i - \alpha_i) < \delta$. Da

$$\sum_{i=1}^n |f(\beta_i) - f(\alpha_i)| = \sum_{i=1}^n \left| \int_{\alpha_i}^{\beta_i} f'(t)\,dt \right| = \sum_{i=1}^n |\lambda((\alpha_i, \beta_i))| \le |\lambda|(E),$$

ist jede solche Funktion f notwendigerweise absolut stetig, wie unten definiert.

7.12 Definition (absolut stetige Funktion)
Eine Funktion $f\colon [a,b] \to \mathbb{R}$ heisst *absolut stetig* auf $[a,b]$, falls gilt: Für jedes $\epsilon > 0$ existiert ein $\delta > 0$, so dass

$$\sum_{i=1}^n |f(\beta_i) - f(\alpha_i)| < \epsilon$$

für jedes $n \in \mathbb{N}$ und jede Familie von paarweise disjunkten Intervallen $(\alpha_1, \beta_1), \ldots, (\alpha_n, \beta_n)$ in $[a,b]$ mit

$$\sum_{i=1}^n (\beta_i - \alpha_i) < \delta.$$

Jedes solche f ist offensichtlich stetig ($n = 1$).

7.13 Satz
Sei $f\colon [a,b] \to \mathbb{R}$ stetig und monoton steigend. Dann sind äquivalent:

(1) *f ist absolut stetig auf $[a,b]$.*
(2) *f ist fast überall in $[a,b]$ differenzierbar, $f' \in L^1$, und*

$$f(x) - f(a) = \int_a^x f'(t)\,dt \quad \text{für alle } x \in [a,b].$$

(3) *f bildet Nullmengen auf Nullmengen ab.*

Beachte, dass die oben definierte (Cantor-Lebesgue-)Funktion f die Nullmenge $E \subset [0,1]$ auf ganz $[0,1]$ abbildet!

Beweis: Die Implikation (2) $\Rightarrow$ (1) gilt gemäss der Bemerkung vor Definition 7.12 sogar ohne die Stetigkeits- und Monotonievoraussetzung an f.

Aus (1) folgt (3): Sei $\mathscr{M}$ die σ-Algebra der Lebesgue-messbaren Teilmengen von $\mathbb{R}$. Sei $E \subset [a,b]$, $E \in \mathscr{M}$ und $\mu(E) = 0$. Wir müssen zeigen: Gilt (1), so ist $f(E) \in \mathscr{M}$ und $\mu(f(E)) = 0$. Seien o.B.d.A. $a, b \notin E$. Sei $\epsilon > 0$; dann gibt es ein passendes $\delta > 0$ wie in 7.12. Es existiert eine offene Menge V mit $\mu(V) < \delta$ und $E \subset V \subset [a,b]$; V ist Vereinigung von abzählbar vielen paarweise disjunkten Intervallen (α_i, β_i). Da $\sum_i (\beta_i - \alpha_i) < \delta$, folgt

$$\sum_i |f(\beta_i) - f(\alpha_i)| \leq \epsilon.$$

Wegen der Monotonie von f ist $f(E) \subset f(V) \subset \bigcup_i [f(\alpha_i), f(\beta_i)]$; das Lebesgue-Mass dieser Vereinigung ist $\leq \epsilon$. Dies zeigt, dass $f(E)$ in Borelmengen mit beliebig kleinem Mass enthalten ist. Mit der Vollständigkeit von μ folgt $f(E) \in \mathscr{M}$ und $\mu(f(E)) = 0$.

(3) impliziert (2): Definiere $g(x) := x + f(x)$ für $x \in [a,b]$. Hat das f-Bild eines Intervalls der Länge l die Länge l', so hat das g-Bild dieses Intervalls die Länge $l + l'$. Daraus folgt, dass unter der Voraussetzung (3) auch g Nullmengen auf Nullmengen abbildet. Sei $E \subset [a,b]$, $E \in \mathscr{M}$. Dann ist $E = E_0 \cup F$ für eine Nullmenge E_0 und eine F_σ-Menge F (3.11(2)); F ist Vereinigung abzählbar vieler kompakter Mengen. Da g stetig ist und (3) erfüllt, folgt die Lebesgue-Messbarkeit von $g(E) = g(E_0) \cup g(F)$. Definiere dann

$$\lambda(E) := \mu(g(E))$$

für $E \subset [a,b]$, $E \in \mathscr{M}$. Da g injektiv ist, ist λ ein (endliches) Mass, ausserdem gilt $\lambda \ll \mu$. Nach 6.8(2) (Radon-Nikodym) ist

$$d\lambda = h\,d\mu$$

für ein $h \in L^1(\mu)$. Für $E = [a,x]$ folgt $g(E) = [g(a), g(x)]$ und somit

$$g(x) - g(a) = \mu(g(E)) = \lambda(E) = \int_E h \, d\mu = \int_a^x h(t) \, dt.$$

Es gilt also

$$f(x) - f(a) = \int_a^x h(t) - 1 \, dt$$

für alle $x \in [a,b]$. Nach Satz 7.11 folgt

$$f'(x) = h(x) - 1$$

für fast alle $x \in [a,b]$. Somit ist (2) erfüllt. □

7.14 Satz

Sei $f\colon [a,b] \to \mathbb{R}$ absolut stetig. Definiere

$$F(x) := \sup \sum_{j=1}^{k} |f(t_j) - f(t_{j-1})| \quad \text{für } x \in [a,b],$$

wobei das Supremum über alle k und alle (t_i) mit $a = t_0 < t_1 < \ldots < t_k = x$ genommen wird. Dann sind die Funktionen F, $F+f$, $F-f$ monoton steigend und absolut stetig auf $[a,b]$.

Ist $F(b) < \infty$ für eine (nicht notwendigerweise absolut stetige) Funktion $f\colon [a,b] \to \mathbb{R}$, so hat f *beschränkte Variation* auf $[a,b]$ und $F(b)$ ist die *totale Variation* von f auf $[a,b]$.

Beweis: Für $a \le x < y \le b$ und $a = t_0 < t_1 < \ldots < t_k = x$ gilt

$$F(y) \ge |f(y) - f(x)| + \sum_{j=1}^{k} |f(t_j) - f(t_{j-1})|.$$

Somit ist $F(y) \ge |f(y) - f(x)| + F(x)$, also $F(y) \ge f(y) - f(x) + F(x)$ und $F(y) \ge f(x) - f(y) + F(x)$. Dies zeigt die Monotonie von F, $F+f$ und $F-f$. Da die Summe zweier absolut stetiger Funktionen absolut stetig ist, bleibt noch zu zeigen, dass F absolut stetig auf $[a,b]$ ist. Sei $\epsilon > 0$; dann existiert ein $\delta > 0$ wie in 7.12. Seien $(\alpha_1, \beta_1), \ldots, (\alpha_n, \beta_n)$ paarweise disjunkte Intervalle in $[a,b]$ mit $\sum_{i=1}^n (\beta_i - \alpha_i) < \delta$. Für jedes i gilt

$$F(\beta_i) - F(\alpha_i) = \sup \sum_{j=1}^{k_i} |f(t_{i,j}) - f(t_{i,j-1})|,$$

wobei das Supremum über alle k_i und $(t_{i,j})$ mit $\alpha_i = t_{i,0} < t_{i,1} < \ldots < t_{i,k_i} = \beta_i$ genommen wird. Für alle solchen k_i und $(t_{i,j})$, $i = 1, \ldots, n$, ist

$$\sum_{i=1}^{n} \sum_{j=1}^{k_i} (t_{i,j} - t_{i,j-1}) = \sum_{i=1}^{n} (\beta_i - \alpha_i) < \delta$$

und somit

$$\sum_{i=1}^{n} \sum_{j=1}^{k_i} |f(t_{i,j}) - f(t_{i,j-1})| < \epsilon$$

nach der Wahl von δ. Es folgt $\sum_{i=1}^{n} (F(\beta_i) - F(\alpha_i)) \leq \epsilon$. □

7.15 Satz
Für eine Funktion $f\colon [a, b] \to \mathbb{R}$ sind äquivalent:

(1) *f ist absolut stetig auf $[a, b]$.*
(2) *f ist fast überall in $[a, b]$ differenzierbar, $f' \in L^1$, und*

$$f(x) - f(a) = \int_a^x f'(t)\, dt \quad \text{für alle } x \in [a, b].$$

Beweis: Wir haben schon gesehen, dass (1) aus (2) folgt. Gilt (1), so sei F die wie in 7.14 definiert. Setze $f_1 := \frac{1}{2}(F + f)$, $f_2 := \frac{1}{2}(F - f)$, und wende die Implikation (1) ⇒ (2) von 7.13 auf f_1 und f_2 an. Da $f = f_1 - f_2$, folgt (2). □

7.16 Satz
Ist $f\colon [a, b] \to \mathbb{R}$ überall *auf $[a, b]$ differenzierbar und $f' \in L^1$, so ist*

$$f(x) - f(a) = \int_a^x f'(t)\, dt \quad \text{für alle } x \in [a, b].$$

Beweis: Es genügt, dies für $x = b$ zu zeigen. Sei $\epsilon > 0$. Nach Satz 4.8 (Vitali-Carathéodory) existiert eine von unten halbstetige Funktion g auf $[a, b]$, so dass $g > f'$ und $\int_a^b g(t)\, dt < \int_a^b f'(t)\, dt + \epsilon$. (4.8 gibt eigentlich nur $g \geq f'$, aber wegen $\mu([a, b]) < \infty$ kann man leicht auch $g > f'$ erreichen.) Für $\eta > 0$ und $x \in [a, b]$ sei

$$F_\eta(x) := \int_a^x g(t)\, dt - f(x) + f(a) + \eta(x - a).$$

Sei η zunächst fest. Für jedes $x \in [a, b)$ existiert ein $\delta_x > 0$ mit

$$g(t) > f'(x) \quad \text{und} \quad \frac{f(t) - f(x)}{t - x} < f'(x) + \eta$$

für alle $t \in (x, x+\delta_x)$, da g von unten halbstetig und $g(x) > f'(x)$ ist. Für jedes solche t folgt

$$\begin{aligned} F_\eta(t) - F_\eta(x) &= \int_x^t g(s)\,ds - (f(t) - f(x)) + \eta(t-x) \\ &> (t-x)f'(x) - (t-x)(f'(x)+\eta) + \eta(t-x) = 0. \end{aligned}$$

Da $F_\eta(a) = 0$ und F_η *stetig* ist, existiert ein maximales $x \in [a,b]$ mit $F_\eta(x) = 0$. Im Fall $x < b$ ist also $F_\eta(t) > 0$ für $t \in (x,b]$. In beiden Fällen gilt $F_\eta(b) \geq 0$. Da dies für alle $\eta > 0$ zutrifft, folgt

$$f(b) - f(a) \leq \int_a^b g(t)\,dt < \int_a^b f'(t)\,dt + \epsilon.$$

Da $\epsilon > 0$ beliebig war, folgt $f(b) - f(a) \leq \int_a^b f'(t)\,dt$. Dieselbe Ungleichung gilt für $-f$ anstelle von f; damit gilt Gleichheit. □

8 Produkt-Räume

In diesem Kapitel untersuchen wir die Integration von Funktionen auf kartesischen Produkten $X \times Y = \{(x,y) : x \in X,\ y \in Y\}$. Wir bilden zuerst das Produkt zweier messbarer Räume.

8.1 Definition (Produkt-σ-Algebra)
Seien $(X, \mathscr{S})$ und $(Y, \mathscr{T})$ zwei messbare Räume. Die *Produkt-σ-Algebra* $\mathscr{S} \times \mathscr{T}$ in $X \times Y$ ist die kleinste σ-Algebra in $X \times Y$, die alle Mengen der Gestalt $A \times B$ mit $A \in \mathscr{S}$ und $B \in \mathscr{T}$ enthält.

8.2 Satz
Sei $E \in \mathscr{S} \times \mathscr{T}$. Für jedes $x \in X$ ist $E_x := \{y \in Y : (x,y) \in E\}$ in $\mathscr{T}$, und für jedes $y \in Y$ ist $E^y := \{x \in X : (x,y) \in E\}$ in $\mathscr{S}$.

Beweis: Sei Ω die Menge aller $E \in \mathscr{S} \times \mathscr{T}$ mit $E_x \in \mathscr{T}$ für jedes $x \in X$. Wir wollen zeigen, dass $\Omega = \mathscr{S} \times \mathscr{T}$. Ist $E = A \times B$ für $A \in \mathscr{S}$ und $B \in \mathscr{T}$, so ist $E_x = B$ falls $x \in A$ und $E_x = \emptyset$ falls $x \notin A$; somit ist $E \in \Omega$. Da $\mathscr{T}$ eine σ-Algebra ist, gilt ausserdem:

(i) $X \times Y \in \Omega$.
(ii) Für $E \in \Omega$ ist $(E^c)_x = (E_x)^c \in \mathscr{T}$, also $E^c \in \Omega$.
(iii) Für $E = \bigcup_{i=1}^\infty E_i$, $E_i \in \Omega$, ist $E_x = \bigcup_{i=1}^\infty (E_i)_x \in \mathscr{T}$, also $E \in \Omega$.

Dies zeigt, dass Ω eine σ-Algebra ist. Da $\{A \times B : A \in \mathscr{S},\ B \in \mathscr{T}\} \subset \Omega \subset \mathscr{S} \times \mathscr{T}$, folgt $\Omega = \mathscr{S} \times \mathscr{T}$ nach Definition 8.1. Dies zeigt eine Hälfte des Satzes, die andere folgt analog. □

8.3 Definition (monotone Klasse)
Eine *monotone Klasse* $\mathscr{M}$ ist ein System von Mengen mit den folgenden Eigenschaften:

(i) Aus $A_1, A_2, \ldots \in \mathscr{M}$, $A_1 \subset A_2 \subset \ldots$, folgt $\bigcup_{i=1}^{\infty} A_i \in \mathscr{M}$.
(ii) Aus $B_1, B_2, \ldots \in \mathscr{M}$, $B_1 \supset B_2 \supset \ldots$, folgt $\bigcap_{i=1}^{\infty} B_i \in \mathscr{M}$.

Im Folgenden nennen wir eine Menge $Q \subset X \times Y$ *elementar*, und schreiben $Q \in \mathscr{E}$, falls Q die Vereinigung von endlichen vielen, paarweise disjunkten Mengen der Gestalt $A \times B$ ist, wobei $A \in \mathscr{S}$ und $B \in \mathscr{T}$.

8.4 Satz
Die Produkt-σ-Algebra $\mathscr{S} \times \mathscr{T}$ ist die kleinste monotone Klasse, die alle elementaren Mengen enthält.

Beweis: Sei $\mathscr{M}$ die kleinste monotone Klasse, die $\mathscr{E}$ enthält; die Existenz von $\mathscr{M}$ zeigt man wie im Beweis von 1.3. Da $\mathscr{S} \times \mathscr{T}$ eine monotone Klasse ist, gilt $\mathscr{M} \subset \mathscr{S} \times \mathscr{T}$. Aus den Identitäten

$$
\begin{aligned}
(A_1 \times B_1) \cap (A_2 \times B_2) &= (A_1 \cap A_2) \times (B_1 \cap B_2), \\
(A_1 \times B_1) \setminus (A_2 \times B_2) &= [(A_1 \setminus A_2) \times B_1] \cup [(A_1 \cap A_2) \times (B_1 \setminus B_2)]
\end{aligned}
$$

folgt, dass mit $P, Q \in \mathscr{E}$ auch $P \cap Q$, $P \setminus Q$ und $P \cup Q = (P \setminus Q) \cup Q$ in $\mathscr{E}$ sind. Für eine Menge $P \subset X \times Y$ bezeichne $\Omega(P)$ die Menge aller $Q \subset X \times Y$, so dass $P \setminus Q, Q \setminus P, P \cup Q \in \mathscr{M}$. Man sieht leicht:

(1) $Q \in \Omega(P)$ gilt genau dann, wenn $P \in \Omega(Q)$.
(2) Da $\mathscr{M}$ eine monotone Klasse ist, ist jedes $\Omega(P)$ eine monotone Klasse.

Fixiere $P \in \mathscr{E}$. Nach den oben gezeigten Eigenschaften von $\mathscr{E}$ ist $\mathscr{E} \subset \Omega(P)$; mit (2) folgt $\mathscr{M} \subset \Omega(P)$. Fixiere jetzt $Q \in \mathscr{M}$. Für $P \in \mathscr{E}$ ist $Q \in \mathscr{M} \subset \Omega(P)$, also $P \in \Omega(Q)$ nach (1). Somit gilt $\mathscr{E} \subset \Omega(Q)$, und mit (2) folgt $\mathscr{M} \subset \Omega(Q)$. Wir haben also gezeigt: Sind $P, Q \in \mathscr{M}$, so sind auch $P \setminus Q$ und $P \cup Q$ in $\mathscr{M}$. Nun folgt leicht, dass $\mathscr{M}$ eine σ-Algebra ist:

(i) $X \times Y \in \mathscr{E} \subset \mathscr{M}$.
(ii) Mit $Q \in \mathscr{M}$ ist auch $Q^c = (X \times Y) \setminus Q \in \mathscr{M}$.
(iii) Sind $P_1, P_2, \ldots \in \mathscr{M}$, und ist $P = \bigcup_{i=1}^{\infty} P_i$, so ist $Q_n := \bigcup_{i=1}^{n} P_i \in \mathscr{M}$ für alle $n \in \mathbb{N}$, also $P = \bigcup_{n=1}^{\infty} Q_n \in \mathscr{M}$ wegen der Monotonie von $\mathscr{M}$.

Da $\mathscr{E} \subset \mathscr{M} \subset \mathscr{S} \times \mathscr{T}$, folgt $\mathscr{M} = \mathscr{S} \times \mathscr{T}$. $\square$

8.5 Satz
Sei $f\colon X \times Y \to \bar{\mathbb{R}}$ eine $(\mathscr{S} \times \mathscr{T})$-messbare Funktion. Für jedes $x \in X$ ist $f_x := f(x, \cdot)\colon Y \to \bar{\mathbb{R}}$ $\mathscr{T}$-messbar, und für jedes $y \in Y$ ist $f^y := f(\cdot, y)\colon X \to \bar{\mathbb{R}}$ $\mathscr{S}$-messbar.

Beweis: Sei $V \subset \bar{\mathbb{R}}$ offen. Dann ist $Q := f^{-1}(V) \in \mathscr{S} \times \mathscr{T}$, und für alle $x \in X$ ist die Menge

$$f_x^{-1}(V) = \{y : f(x,y) \in V\} = \{y : (x,y) \in Q\} = Q_x$$

in $\mathscr{T}$ gemäss 8.2. □

8.6 Satz
Seien $(X, \mathscr{S}, \mu)$ und $(Y, \mathscr{T}, \nu)$ zwei σ-endliche Massräume, und sei $Q \in \mathscr{S} \times \mathscr{T}$. Setze

$$\varphi(x) := \nu(Q_x) = \nu\big(\{y : (x,y) \in Q\}\big) \quad \textit{für } x \in X,$$
$$\psi(y) := \mu(Q^y) = \mu\big(\{x : (x,y) \in Q\}\big) \quad \textit{für } y \in Y.$$

Dann gilt: φ ist $\mathscr{S}$-messbar, ψ ist $\mathscr{T}$-messbar, und

$$(*) \qquad \int_X \varphi\, d\mu = \int_Y \psi\, d\nu.$$

Nach 8.2 sind φ und ψ definiert. Gleichung $(*)$ kann auch in der Form

$$\int_X \int_Y \chi_Q(x,y)\, d\nu(y)\, d\mu(x) = \int_Y \int_X \chi_Q(x,y)\, d\mu(x)\, d\nu(y)$$

geschrieben werden.

Beweis: Sei Ω die Menge aller $Q \in \mathscr{S} \times \mathscr{T}$, für die die Aussage des Satzes gilt. Wir zeigen zunächst:

(1) Ist $Q = A \times B$ für $A \in \mathscr{S}$ und $B \in \mathscr{T}$, so ist $Q \in \Omega$.
(2) Ist $Q = \bigcup_{i=1}^\infty Q_i$ für $Q_1 \subset Q_2 \subset \dots$, $Q_i \in \Omega$, so ist $Q \in \Omega$.
(3) Ist $Q = \bigcup_i Q_i$ für eine abzählbare Familie (Q_i) paarweise disjunkter $Q_i \in \Omega$, so ist $Q \in \Omega$.
(4) Ist $Q = \bigcap_{i=1}^\infty Q_i$, wobei $A \times B \supset Q_1 \supset Q_2 \supset \dots$, $Q_i \in \Omega$, $\mu(A) < \infty$ und $\nu(B) < \infty$, so ist $Q \in \Omega$.

Für $Q = A \times B$ wie in (1) gilt

$$\nu(Q_x) = \nu(B)\chi_A(x) \quad \text{und} \quad \mu(Q^y) = \mu(A)\chi_B(y).$$

Somit sind beide Integrale in $(*)$ gleich $\mu(A)\nu(B)$.
Sei jetzt Q wie in (2). Setze $\varphi_i(x) := \nu((Q_i)_x)$ für $x \in X$ und $\psi_i(y) := \mu((Q_i)^y)$ für $y \in Y$. Es gilt $\varphi_1(x) \le \varphi_2(x) \le \dots$ und $\varphi_i(x) \to \varphi(x)$ $(i \to \infty)$ für alle $x \in X$ und ebenso $\psi_1(y) \le \psi_2(y) \le \dots$ und $\psi_i(y) \to \psi(y)$ $(i \to \infty)$

für alle $y \in Y$. Da nach Voraussetzung die Aussage des Satzes für φ_i und ψ_i gilt, folgt $Q \in \Omega$ aus 2.4 (monotone Konvergenz).

Sei Q wie in (3). Ist die Familie (Q_i) endlich, so folgt $Q \in \Omega$ wegen $\chi_Q = \sum_i \chi_{Q_i}$. Zusammen mit (2) impliziert dies $Q \in \Omega$ im allgemeinen Fall.

(4) folgt auf ähnliche Weise wie (2), mit 2.10 (beschränkte Konvergenz) anstelle von 2.4.

Nach Voraussetzung existieren messbare Zerlegungen $(X_m)_{m\in\mathbb{N}}$ und $(Y_n)_{n\in\mathbb{N}}$ von X bzw. Y mit $\mu(X_m) < \infty$ und $\nu(Y_n) < \infty$ für alle $m, n \in \mathbb{N}$. Für $Q \in \mathscr{S} \times \mathscr{T}$ sei $Q_{mn} := Q \cap (X_m \times Y_n)$; sei dann $\mathscr{M}$ die Menge aller $Q \in \mathscr{S} \times \mathscr{T}$ mit $Q_{mn} \in \Omega$ für jedes Paar (m, n). Aus (2) und (4) folgt, dass $\mathscr{M}$ eine monotone Klasse ist. Wegen (1) und (3) ist $\mathscr{E} \subset \mathscr{M}$; da $\mathscr{M} \subset \mathscr{S} \times \mathscr{T}$, gilt $\mathscr{M} = \mathscr{S} \times \mathscr{T}$ nach Satz 8.4. Für $Q \in \mathscr{S} \times \mathscr{T}$ ist somit $Q_{mn} \in \Omega$ für jedes jedes Paar (m, n). Da $Q = \bigcup_{m,n=1}^{\infty} Q_{m,n}$, und da die Q_{mn} paarweise disjunkt sind, folgt $Q \in \Omega$ aus (3). $\square$

8.7 Definition (Produktmass)
Seien $(X, \mathscr{S}, \mu)$ und $(Y, \mathscr{T}, \nu)$ wie in Satz 8.6. Für $Q \in \mathscr{S} \times \mathscr{T}$ sei

$$(\mu \times \nu)(Q) := \int_X \nu(Q_x)\, d\mu(X) = \int_Y \mu(Q^y)\, d\nu(Y);$$

$\mu \times \nu$ heisst das *Produktmass* von μ und ν.

Dass $\mu \times \nu$ wirklich ein Mass auf $\mathscr{S} \times \mathscr{T}$ definiert, folgt sofort aus Satz 2.5. Beachte, dass $\mu \times \nu$ auch σ-endlich ist.

8.8 Satz (Fubini)
Seien $(X, \mathscr{S}, \mu)$ und $(Y, \mathscr{T}, \nu)$ zwei σ-endliche Massräume, und sei f eine $(\mathscr{S} \times \mathscr{T})$-messbare Funktion auf $X \times Y$.

(1) *Ist $0 \le f \le \infty$, so gilt: $x \mapsto \int_Y f(x,y)\, d\nu(y)$ ist $\mathscr{S}$-messbar, $y \mapsto \int_X f(x,y)\, d\mu(x)$ ist $\mathscr{T}$-messbar, und*

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

(2) *Ist f reell und wenigstens eines der Integrale $\int_X \int_Y |f(x,y)|\, d\nu(y)\, d\mu(x)$ und $\int_Y \int_X |f(x,y)|\, d\mu(x)\, d\nu(y)$ endlich, so ist $f \in L^1(\mu \times \nu)$.*

(3) *Ist f reell und $f \in L^1(\mu \times \nu)$, so ist $f(x, \cdot)$ in $L^1(\nu)$ für μ-fast alle x und $x \mapsto \int_Y f(x,y)\, d\nu(y)$ in $L^1(\mu)$, ebenso ist $f(\cdot, y)$ in $L^1(\mu)$ für ν-fast alle y und $y \mapsto \int_X f(x,y)\, d\mu(x)$ in $L^1(\nu)$, und es gilt $(*)$.*

Kurz gesagt: Die Reihenfolge der Integration darf für $(\mathscr{S} \times \mathscr{T})$-messbare Funktionen f vertauscht werden, wenn $f \geq 0$ (s. (1)) oder wenn wenigstens eines der Doppelintegrale von $|f|$ endlich ist (s. (2) und (3)).

Beweis: (1): Nach Satz 8.5 sind die Funktionen $\varphi(x) := \int_Y f(x,y)\,d\nu(y)$ und $\phi(y) := \int_X f(x,y)\,d\mu(x)$ erklärt. Sei $Q \in \mathscr{S} \times \mathscr{T}$ und $f = \chi_Q$. Nach Definition 8.7 entspricht (1) dann gerade der Aussage von Satz 8.6. Damit gilt (1) für alle $(\mathscr{S} \times \mathscr{T})$-messbaren Treppenfunktionen $s \geq 0$. Im allgemeinen Fall existiert eine Folge solcher Funktionen s_n, so dass $0 \leq s_1 \leq s_2 \leq \ldots$ und $s_n(x,y) \to f(x,y)$ für alle $(x,y) \in X \times Y$. Sei $\varphi_n(x) := \int_Y s_n(x,y)\,d\nu(y)$. Satz 2.4 (monotone Konvergenz), angewandt auf $(Y, \mathscr{T}, \nu)$, zeigt dass $\varphi_n(x) \to \varphi(x)$ $(n \to \infty)$ für alle $x \in X$. Da diese Konvergenz monoton ist, und da

$$\int_X \varphi_n\,d\mu = \int_{X\times Y} s_n\,d(\mu \times \nu)$$

gemäss dem anfangs diskutierten Spezialfall, folgt durch Anwenden von 2.4 auf diese beiden Integrale die erste Gleichung in $(*)$. Die zweite folgt durch Vertauschen der Rollen von x und y.

(2) folgt durch Anwenden von (1) auf $|f|$.

(3): Sei $\varphi_1(x) := \int_Y f^+(x,y)\,d\nu(y)$ und $\varphi_2(x) := \int_Y f^-(x,y)\,d\nu(y)$. Da (1) für f^+ gilt, und da $f \in L^1(\mu \times \nu)$, folgt

$$\int_X \varphi_1\,d\mu = \int_{X\times Y} f^+\,d(\mu \times \nu) \leq \int_{X\times Y} |f|\,d(\mu \times \nu) < \infty,$$

also $\varphi_1 \in L^1(\mu)$. Ebenso ist $\varphi_2 \in L^1(\mu)$. Für μ-fast alle x gilt deshalb

$$\int_Y |f(x,\cdot)|\,d\nu = \int_Y f^+(x,\cdot) + f^-(x,\cdot)\,d\nu = \varphi_1(x) + \varphi_2(x) < \infty,$$

also $f(x,\cdot) \in L^1(\nu)$, und

$$\varphi(x) := \int_Y f(x,y)\,d\nu(y) = \varphi_1(x) - \varphi_2(x).$$

Somit ist $\varphi \in L^1(\mu)$. Mit $\int_X \varphi_1\,d\mu = \int_{X\times Y} f^+\,d(\mu \times \nu)$ und $\int_X \varphi_2\,d\mu = \int_{X\times Y} f^-\,d(\mu \times \nu)$ folgt die erste Ungleichung in $(*)$. Dies zeigt eine Hälfte von (3). Die andere Hälfte folgt analog. □

Wir betrachten drei Beispiele, die zeigen, dass die verschiedenen Voraussetzungen in 8.8 nicht weggelassen werden können.

Sei zuerst $X = Y = [0,1]$, $\mu = \nu$ das Lebesgue-Mass auf $[0,1]$. Wähle $0 = \delta_1 < \delta_2 < \dots$ mit $\delta_n \to 1$ $(n \to \infty)$. Für $n \in \mathbb{N}$ sei g_n eine reelle stetige Funktion mit $\operatorname{Tr} g_n \subset (\delta_n, \delta_{n+1})$ und $\int_0^1 g_n(t)\,dt = 1$. Setze dann

$$f(x,y) := \sum_{n=1}^{\infty} (g_n(x) - g_{n+1}(x))g_n(y);$$

für jeden Punkt (x,y) ist höchstens ein Summand ungleich null. Eine einfache Rechnung zeigt, dass

$$\int_0^1 \int_0^1 f(x,y)\,dy\,dx = 1 \neq 0 = \int_0^1 \int_0^1 f(x,y)\,dx\,dy.$$

In diesem Beispiel ist f stetig ausser im Punkt $(1,1)$, aber beide Doppelintegrale von $|f|$ sind unendlich.

Für das zweite Beispiel sei $X = Y = [0,1]$, μ das Lebesgue-Mass auf $[0,1]$, ν das Zählmass auf $[0,1]$, $f(x,y) = 1$ für $x = y$ und $f(x,y) = 0$ für $x \neq y$. Dann ist $\int_X f(x,y)\,d\mu(x) = 0$ für alle $y \in Y$ und $\int_Y f(x,y)\,d\nu(y) = 1$ für alle $x \in X$, somit

$$\int_Y \int_X f(x,y)\,d\mu(x)\,d\nu(y) = 0 \neq 1 = \int_X \int_Y f(x,y)\,d\nu(y)\,d\mu(x).$$

In diesem Beispiel ist ν nicht σ-endlich. Beachte, dass f $(\mathscr{S} \times \mathscr{T})$-messbar ist, wobei $\mathscr{S}$ die σ-Algebra der Lebesgue-messbaren Teilmengen von $[0,1]$ und $\mathscr{T}$ die Potenzmenge von $[0,1]$ bezeichnen: Es gilt $f = \chi_D$ für $D = \{(x,y) : 0 \le x = y \le 1\}$, und $D = \bigcap_{n=1}^{\infty} Q_n$ für die $(\mathscr{S} \times \mathscr{T})$-messbaren Mengen $Q_n = \bigcup_{j=1}^{n} [\frac{j-1}{n}, \frac{j}{n}]^2$.

Im dritten Beispiel ist wieder $(X,\mathscr{S},\mu) = (Y,\mathscr{T},\nu) = [0,1]$ mit dem Lebesgue-Mass. Unter Voraussetzung der Kontinuumshypothese kann man eine Menge $Q \subset [0,1]^2$ mit folgender Eigenschaft konstruieren: Für jedes $x \in [0,1]$ enthält Q_x alle bis auf abzählbar viele Punkte von $[0,1]$, und für jedes $y \in [0,1]$ ist Q^y abzählbar. Für $f := \chi_Q$ folgt, dass $f(x,\cdot)$ und $f(\cdot,y)$ Borel-messbar sind und dass $\int_0^1 f(x,y)\,dy = 1$ für alle x und $\int_0^1 f(x,y)\,dx = 0$ für alle y. Somit ist

$$\int_0^1 \int_0^1 f(x,y)\,dy\,dx = 1 \neq 0 = \int_0^1 \int_0^1 f(x,y)\,dx\,dy.$$

In diesem Beispiel ist f offenbar nicht $(\mathscr{S} \times \mathscr{T})$-messbar.

Das Produkt von zwei vollständigen Massräumen $(X,\mathscr{S},\mu)$ und $(Y,\mathscr{T},\nu)$ ist im Allgemeinen nicht vollständig: Ist $A \in \mathscr{S}$, $A \neq \emptyset$, $\mu(A) = 0$ und $B \subset Y$, $B \notin \mathscr{T}$, so ist $A \times B \subset A \times Y$ und $(\mu \times \nu)(A \times Y) = 0$, aber $A \times B \notin \mathscr{S} \times \mathscr{T}$ (für $x \in A$ müsste $(A \times B)_x = B$ in $\mathscr{T}$ sein gemäss 8.2).

8.9 Satz
Sei μ_k das Lebesgue-Mass auf $\mathbb{R}^k$, $k \geq 1$. Ist $k = r + s$, $r, s \geq 1$, so ist μ_k die Vervollständigung von $\mu_r \times \mu_s$.

Beweis: Seien $\mathscr{B}_k$ und $\mathscr{M}_k$ die σ-Algebren der Borelmengen bzw. Lebesgue-messbaren Mengen in $\mathbb{R}^k$. Jeder Quader in $\mathbb{R}^k$ ist in $\mathscr{M}_r \times \mathscr{M}_s$, und die von der Menge aller Quader erzeugte σ-Algebra ist $\mathscr{B}_k$. Somit ist $\mathscr{B}_k \subset \mathscr{M}_r \times \mathscr{M}_s$. Für $C \in \mathscr{M}_r$ und $D \in \mathscr{M}_s$ sind $C \times \mathbb{R}^s$ und $\mathbb{R}^r \times D$ in $\mathscr{M}_k$ (dies folgt aus 3.11(2)); somit ist auch $C \times D = (C \times \mathbb{R}^s) \cap (\mathbb{R}^r \times D)$ in $\mathscr{M}_k$. Es folgt $\mathscr{M}_r \times \mathscr{M}_s \subset \mathscr{M}_k$. Da $\mu_r \times \mu_s$ translationsinvariant ist und jedem Quader in $\mathbb{R}^k$ denselben Wert wie μ_k zuordnet, gilt
$$\mu_r \times \mu_s = \mu_k|_{(\mathscr{M}_r \times \mathscr{M}_s)}$$
gemäss 3.11(4). Für jedes $A \in \mathscr{M}_k$ existieren $F, G \in \mathscr{B}_k$ mit $F \subset A \subset G$ und $\mu_k(G \setminus F) = 0$ (3.11(2)). Da $\mathscr{B}_k \subset \mathscr{M}_r \times \mathscr{M}_s$, ist $\mathscr{M}_k$ die Vervollständigung von $\mathscr{M}_r \times \mathscr{M}_s$ bzgl. $\mu_r \times \mu_s$ (s. 2.12). $\square$

Wir zeigen in 8.11 noch eine Variante des Satzes von Fubini, die vor allem im Zusammenhang mit 8.9 wichtig ist. Der Beweis verwendet 8.8 und das folgende Lemma.

8.10 Lemma
Sei $(X, \mathscr{M}, \mu)$ ein Massraum, und sei $\mathscr{M}^*$ die Vervollständigung von $\mathscr{M}$ bzgl. μ. Ist $f\colon X \to \bar{\mathbb{R}}$ eine $\mathscr{M}^*$-messbare Funktion, so existiert eine $\mathscr{M}$-messbare Funktion $g\colon X \to \bar{\mathbb{R}}$ mit $|g| \leq |f|$ und $g(x) = f(x)$ für alle $x \in X \setminus Q$, wobei $Q \in \mathscr{M}$ und $\mu(Q) = 0$.

Beweis: Sei zunächst $f \geq 0$. Dann existieren $\mathscr{M}^*$-messbare Treppenfunktionen $0 \leq s_1 \leq s_2 \leq \ldots$, so dass $s_n(x) \to f(x)$ $(n \to \infty)$ für alle $x \in X$. Es gilt $f = \sum_{n=1}^{\infty}(s_{n+1} - s_n)$ und somit $f = \sum_{i=1}^{\infty} c_i \chi_{E_i}$ für Konstanten $c_i > 0$ und Mengen $E_i \in \mathscr{M}^*$. Nach Definition von $\mathscr{M}^*$ (s. 2.12) gibt es Mengen $A_i, B_i \in \mathscr{M}$ mit $A_i \subset E_i \subset B_i$ und $\mu(B_i \setminus A_i) = 0$. Dann ist
$$g := \sum_{i=1}^{\infty} c_i \chi_{A_i}$$
$\mathscr{M}$-messbar, $g \leq f$, und $g(x) = f(x)$ für alle x ausserhalb $\bigcup_{i=1}^{\infty}(E_i \setminus A_i)$. Da $E_i \subset B_i$ und $\mu(B_i \setminus A_i) = 0$, gilt für $Q := \bigcup_{i=1}^{\infty}(B_i \setminus A_i)$ die Behauptung im Fall $f \geq 0$. Das allgemeine Resultat folgt leicht. $\square$

8.11 Satz
Seien $(X, \mathscr{S}, \mu)$ und $(Y, \mathscr{T}, \nu)$ zwei vollständige und σ-endliche Massräume, und sei $(\mathscr{S} \times \mathscr{T})^*$ die Vervollständigung von $\mathscr{S} \times \mathscr{T}$ bzgl. $\mu \times \nu$. Sei f eine

$(\mathscr{S} \times \mathscr{T})^*$*-messbare Funktion auf* $X \times Y$. *Dann gelten* 8.8(1), (2) *und* (3) *für das auf* $(\mathscr{S} \times \mathscr{T})^*$ *fortgesetzte Mass* $\mu \times \nu$, *mit folgendem Unterschied: In* (1) *ist die* $\mathscr{T}$*-Messbarkeit von* $f(x, \cdot)$ *und damit die Existenz von* $\int_Y f(x, y)\, d\nu(y)$ *nur für* μ*-fast alle* x *gesichert, ebenso ist* $\int_X f(x, y)\, d\mu(x)$ *im Allgemeinen nur* ν*-fast überall erklärt.*

Beweis: Sei f wie im Satz. Nach Lemma 8.10 ist $f = g + h$ für eine $(\mathscr{S} \times \mathscr{T})$-messbare Funktion g mit $|g| \leq |f|$ und eine Funktion h mit $h(x, y) = 0$ für alle $(x, y) \in (X \times Y) \setminus Q$, wobei $Q \in \mathscr{S} \times \mathscr{T}$ und $(\mu \times \nu)(Q) = 0$. Dann gilt

$$\int_X \nu(Q_x)\, d\mu(x) = (\mu \times \nu)(Q) = 0,$$

also $\nu(Q_x) = 0$ für μ-fast alle x. Für jedes solche x ist $h(x, \cdot) = 0$ ν-fast überall. Ebenso gilt: Für ν-fast alle y ist $h(\cdot, y) = 0$ μ-fast überall. Mit der Vollständigkeit von μ und ν folgt die $\mathscr{T}$-Messbarkeit von $h(x, \cdot)$ für μ-fast alle x sowie die $\mathscr{S}$-Messbarkeit von $h(\cdot, y)$ für ν-fast alle y. Nun folgt der Satz durch Anwenden von 8.8 auf g; alle drei auftretenden Integrale stimmen mit denjenigen für f überein. $\square$