Analysis III

Wintersemester 07/08

Ernst Kuwert

Mathematisches Institut
Universität Freiburg

Inhaltsverzeichnis

1 Inhaltsfunktionen

In diesem Kapitel wird der Jordansche Inhaltsbegriff für Mengen $E \subset \mathbb{R}^n$ erklärt. Dabei wird das Volumen zunächst für endliche Vereinigungen von achsenparallelen Quadern, sogenannte Figuren, elementargeometrisch definiert. Dann werden beliebige, beschränkte Mengen $E \subset \mathbb{R}^n$ von außen und von innen durch solche Figuren approximiert. Es stellt sich heraus, dass auf diese Weise zwar für hinreichend reguläre Mengen ein vernünftiger Inhaltsbegriff gegeben ist, dass aber andererseits mit diesem Zugang nicht einmal allen beschränkten, offenen Mengen sinnvoll ein Volumen zugeordnet werden kann. Parallel zu der konkreten Diskussion im $\mathbb{R}^n$ werden die Begriffe Halbring, Ring, Inhalt allgemein definiert.

Definition 1.1 (Quader) *Eine Menge $I \subset \mathbb{R}$ heißt Intervall, wenn es $a, b \in \mathbb{R}$ mit $a \leq b$ gibt, so dass gilt:*

$$(a, b) \subset I \subset [a, b]. \tag{1.1}$$

Ein achsenparalleler n-dimensionaler Quader (im Folgenden kurz als Quader bezeichnet) ist ein kartesisches Produkt $Q = I_1 \times \ldots \times I_n \subset \mathbb{R}^n$ von Intervallen. Sind $a_j \leq b_j$ die Intervallgrenzen von I_j, so ist das elementargeometrische Volumen von Q definiert als

$$\lambda^n(Q) = \prod_{j=1}^{n} (b_j - a_j) \geq 0. \tag{1.2}$$

Auch die leere Menge ist ein Intervall, denn es gilt $\emptyset = (a, a)$ für jedes $a \in \mathbb{R}$. Die Grenzen eines nichtleeren Intervalls I sind eindeutig bestimmt als $\inf I$ und $\sup I$. Die leere Menge ist auch ein n-dimensionaler Quader wegen $\emptyset = \emptyset \times \ldots \times \emptyset$.

Wir wollen nun die Menge aller Quader im $\mathbb{R}^n$ betrachten. Allgemein bezeichnen wir die Potenzmenge einer Menge X, also die Menge aller Teilmengen von X, mit 2^X. Eine Teilmenge von 2^X wird in der Maßtheorie als *System von Mengen* bezeichnet, um den Ausdruck „Menge von Mengen“ zu vermeiden.

Definition 1.2 (Halbring) *Ein Mengensystem $\mathcal{P} \subset 2^X$ heißt Halbring über X, wenn gilt:*

(i) $\emptyset \in \mathcal{P}$

(ii) $P, Q \in \mathcal{P} \quad \Rightarrow \quad P \cap Q \in \mathcal{P}$

(iii) $P, Q \in \mathcal{P} \quad \Rightarrow \quad P \backslash Q = \bigcup_{i=1}^{k} P_i$ *mit endlich vielen, paarweise disjunkten* $P_i \in \mathcal{P}$.

Satz 1.1 (Halbring $\mathcal{P}^n$) *Das System $\mathcal{P}^n$ der Quader im $\mathbb{R}^n$ ist ein Halbring.*

Beweis: Wir zeigen die Aussage zunächst für $n = 1$, also für das System aller Intervalle. Die leere Menge ist ein Intervall, denn es gilt $\emptyset = (a, a)$ für $a \in \mathbb{R}$ beliebig. Seien $I, J \subset \mathbb{R}$ Intervalle mit Grenzen $a \leq b$ bzw. $c \leq d$. Für $I \cap J \neq \emptyset$ ist $\max(a, c) \leq \min(b, d)$, und

$$(\max(a, c), \min(b, d)) \subset I \cap J \subset [\max(a, c), \min(b, d)].$$

Also ist $I \cap J$ ein Intervall. $I \backslash J$ läßt sich als Vereinigung von höchstens zwei disjunkten Intervallen darstellen. Wegen $I \backslash J = I \backslash (I \cap J)$ können wir dazu $J \subset I$ annehmen. Setze

$$I' = \{x \in I \backslash J : x \leq c\} \quad I'' = \{x \in I \backslash J : x \geq d\}.$$

Falls $I' \cap I'' \neq \emptyset$, so ist $c = d \in I \backslash J$, also $J = \emptyset$ bzw. $I \backslash J = I$. Andernfalls gilt die disjunkte Zerlegung $I \backslash J = I' \cup I''$, wobei

$$(a, c) \subset I' \subset [a, c], \quad (d, b) \subset I'' \subset [d, b].$$

Dies zeigt die Behauptung im Fall $n = 1$.

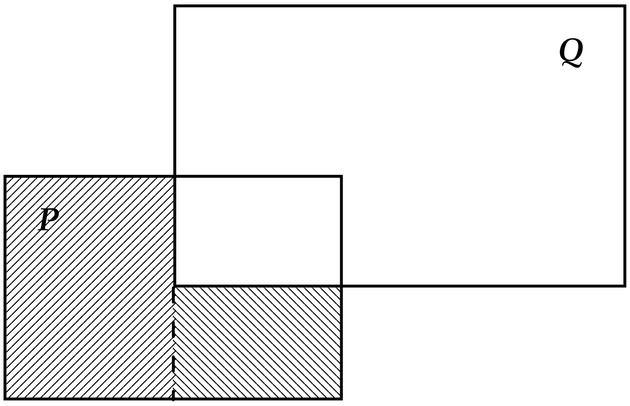
Für Quader $P = I_1 \times \ldots \times I_n$ und $Q = J_1 \times \ldots \times J_n$ im $\mathbb{R}^n$ gilt

$$P \cap Q = (I_1 \cap J_1) \times \ldots \times (I_n \cap J_n),$$

also ist $P \cap Q$ ein Quader. Nun gibt es zu $x \in P \backslash Q$ ein kleinstes $k \in \{1, \ldots, n\}$ mit $x_k \notin J_k$. Daraus ergibt sich die paarweise disjunkte Zerlegung

$$P \backslash Q = \bigcup_{k=1}^{n} (I_1 \cap J_1) \times \ldots \times (I_{k-1} \cap J_{k-1}) \times (I_k \backslash J_k) \times I_{k+1} \times \ldots \times I_n.$$

Wie oben gezeigt, ist $I_k \backslash J_k$ Vereinigung von endlich vielen (höchstens zwei) disjunkten Intervallen. Damit ist auch $P \backslash Q$ als Vereinigung von endlich vielen (höchstens $2n$) paarweise disjunkten Quadern dargestellt. $\square$



Der zweite Teil des Beweises von Satz 1.1 zeigt ganz allgemein Folgendes:

Folgerung 1.1 (Produktmengen) *Für $i = 1, \ldots, n$ sei $\mathcal{P}_i$ ein Halbring über X_i. Dann ist das System*

$$\mathcal{P} = \{P_1 \times \ldots \times P_n : P_i \in \mathcal{P}_i\}$$

der Produktmengen ein Halbring über $X_1 \times \ldots \times X_n$.

Definition 1.3 (Inhalt) *Sei $\mathcal{P}$ ein Halbring über X. Eine Funktion $\lambda : \mathcal{P} \longrightarrow [0, \infty]$ mit $\lambda(\emptyset) = 0$ heißt Inhalt, wenn für jedes $P \in \mathcal{P}$ gilt:*

(1.3) *Ist $P = \bigcup_{i=1}^{k} P_i$ mit $P_1, \ldots, P_k \in \mathcal{P}$ paarweise disjunkt, so folgt $\lambda(P) = \sum_{i=1}^{k} \lambda(P_i)$.*

Satz 1.2 (Elementarinhalt von Quadern) *Das elementargeometrische Volumen*

$$\lambda^n(Q) = \prod_{i=1}^{n} (b_i - a_i)$$

ist ein Inhalt auf dem Halbring $\mathcal{P}^n$ der Quader im $\mathbb{R}^n$.

BEWEIS: Offensichtlich ist $\lambda(\emptyset) = 0$. Um die Eigenschaft (1.3) zu zeigen, verwenden wir Induktion über die Dimension n. Für $n = 1$ sind die charakteristischen Funktionen von $P, P_1, \ldots, P_k$ Riemann-integrierbar und wir erhalten

$$\chi_P = \sum_{i=1}^{k} \chi_{P_i} \quad \Rightarrow \quad \lambda_1(P) = \int_{\mathbb{R}} \chi_P = \sum_{i=1}^{k} \int_{\mathbb{R}} \chi_{P_i} = \sum_{i=1}^{k} \lambda_1(P_i).$$

Sei nun die Aussage für den Inhalt λ_{n-1} im $\mathbb{R}^{n-1}$ schon bewiesen. Betrachte für den Quader $P = I_1 \times \ldots \times I_n \in \mathcal{P}^n$ den y-Schnitt

$$P_y = \{x \in \mathbb{R}^{n-1} : (x, y) \in P\} = \begin{cases} I_1 \times \ldots \times I_{n-1} & \text{falls } y \in I_n \\ \emptyset & \text{sonst.} \end{cases}$$

Dann gilt $P_y = \bigcup_{i=1}^{k} (P_i)_y$ und es folgt induktiv

$$\lambda^n(P) = \int_{\mathbb{R}} \lambda^{n-1}(P_y)\, dy = \int_{\mathbb{R}} \sum_{i=1}^{k} \lambda^{n-1}((P_i)_y)\, dy = \sum_{i=1}^{k} \int_{\mathbb{R}} \lambda^{n-1}((P_i)_y)\, dy = \sum_{i=1}^{k} \lambda^n(P_i).$$

Dies beweist den Satz. □

Unser Ziel ist nun, die Definition der Inhaltsfunktion auf allgemeinere Mengen zu erweitern. Als erstes betrachten wir dabei endliche Vereinigungen von Quadern, sogenannte Figuren. Wie in Satz 1.3 gezeigt wird, ist das System der Figuren unter endlichen Mengenoperationen abgeschlossen, und es ist auch das kleinste Mengensystem mit dieser Eigenschaft, das die Quader enthält. Dafür führen wir die folgenden Begriffe ein.

Definition 1.4 (Ring) *Ein Mengensystem $\mathcal{R} \subset 2^X$ heißt Ring über X, falls gilt:*

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

(ii) $A, B \in \mathcal{R} \Rightarrow A \backslash B \in \mathcal{R}$

(iii) $A, B \in \mathcal{R} \Rightarrow A \cup B \in \mathcal{R}$.

Für zwei Mengen A, B in einem Ring $\mathcal{R}$ liegt auch der Duchschnitt wieder in $\mathcal{R}$, denn es gilt $A \cap B = A \backslash (A \backslash B) \in \mathcal{R}$.

Definition 1.5 (erzeugter Ring) *Für $\mathcal{E} \subset 2^X$ sei*

$$\varrho(\mathcal{E}) = \bigcap \{\mathcal{R} : \mathcal{R} \text{ ist Ring in } X \text{ mit } \mathcal{E} \subset \mathcal{R}\}. \tag{1.4}$$

Dann ist $\varrho(\mathcal{E})$ ein Ring mit $\mathcal{E} \subset \varrho(\mathcal{E})$, und heißt der von $\mathcal{E}$ erzeugte Ring. Es gilt

$$\mathcal{R} \text{ Ring mit } \mathcal{E} \subset \mathcal{R} \quad \Rightarrow \quad \varrho(\mathcal{E}) \subset \mathcal{R}. \tag{1.5}$$

Kurz gesagt ist $\varrho(\mathcal{E})$ also der kleinste Ring, der $\mathcal{E}$ enthält. Diese Beschreibung ist allerdings nicht konstruktiv. Eine explizite Angabe der Elemente des erzeugten Rings leistet aber der folgende Satz.

Satz 1.3 (Konstruktion des erzeugten Rings) *Sei $\mathcal{P}$ ein Halbring über X und $\mathcal{F}$ das System aller endlichen, disjunkten Vereinigungen $F = \bigcup_{i=1}^k P_i$ von Mengen $P_i \in \mathcal{P}$. Dann ist $\mathcal{F}$ der von $\mathcal{P}$ erzeugte Ring.*

Beweis: Jeder Ring $\mathcal{R}$ mit $\mathcal{P} \subset \mathcal{R}$ enthält nach Definition $\mathcal{F}$. Es ist also nur zu zeigen, dass $\mathcal{F}$ ein Ring ist. Offenbar ist $\emptyset \in \mathcal{F}$. Wir zeigen nun durch Induktion über k

$$P \backslash \bigcup_{i=1}^k Q_i \in \mathcal{F} \quad \text{für } P, Q_1, \ldots, Q_k \in \mathcal{P}.$$

Für $k = 1$ gilt das nach Definition 1.2. Ist die disjunkte Zerlegung $P \backslash \bigcup_{i=1}^k Q_i = \bigcup_{j=1}^l P_j$ für ein $k \in \mathbb{N}$ schon gefunden, so folgt

$$P \backslash \bigcup_{i=1}^{k+1} Q_i = \Big(\bigcup_{j=1}^l P_j\Big) \backslash Q_{k+1} = \bigcup_{j=1}^l (P_j \backslash Q_{k+1}).$$

Die $P_j \backslash Q_{k+1}$, $j = 1, \ldots, l$, sind paarweise disjunkt und ihrerseits Vereinigung von endlich vielen disjunkten Mengen aus $\mathcal{P}$ nach Definition 1.2, womit die Induktionsbehauptung gezeigt ist. Sind nun $E = \bigcup_{i=1}^k P_i$, $F = \bigcup_{j=1}^l Q_j$ mit $P_i \in \mathcal{P}$ disjunkt sowie $Q_j \in \mathcal{P}$ disjunkt, so folgt die disjunkte Zerlegung

$$E \backslash F = \bigcup_{i=1}^k \Big(P_i \backslash \bigcup_{j=1}^l Q_j\Big),$$

und damit leicht $E \backslash F \in \mathcal{F}$. Schließlich ergibt sich $E \cup F = (E \backslash F) \cup F \in \mathcal{F}$. $\square$

Satz 1.4 (Fortsetzung auf den erzeugten Ring) *Sei λ ein Inhalt auf dem Halbring $\mathcal{P}$ und $\mathcal{F}$ der von $\mathcal{P}$ erzeugte Ring. Dann gibt es genau einen Inhalt $\overline{\lambda} : \mathcal{F} \to [0, \infty]$ mit $\overline{\lambda}|_{\mathcal{P}} = \lambda$.*

Beweis: Ist $F = \bigcup_{i=1}^k P_i$ mit $P_i \in \mathcal{P}$ paarweise disjunkt wie in Satz 1.3, so muss für die Fortsetzung notwendig gelten

$$\overline{\lambda}(F) = \sum_{i=1}^k \lambda(P_i). \tag{1.6}$$

Deshalb ist die Fortsetzung eindeutig bestimmt. Wir wollen $\overline{\lambda}$ durch (1.6) definieren. Dazu ist zu zeigen, dass die rechte Seite nicht von der Wahl der Zerlegung abhängt. Sei $F = \bigcup_{j=1}^l Q_j$ eine andere disjunkte Darstellung mit $Q_j \in \mathcal{P}$. Dann folgt

$$\sum_{j=1}^l \lambda(Q_j) = \sum_{j=1}^l \sum_{i=1}^k \lambda(P_i \cap Q_j) = \sum_{i=1}^k \sum_{j=1}^l \lambda(P_i \cap Q_j) = \sum_{i=1}^k \lambda(P_i).$$

Somit ist $\overline{\lambda}$ wohldefiniert. Für eine disjunkte Vereinigung $F = \bigcup_{i=1}^k F_i$ mit $F, F_i \in \mathcal{F}$ schreiben wir $F_i = \bigcup_{j=1}^{m_i} P_{i,j}$ mit $P_{i,1}, \ldots, P_{i,m_i} \in \mathcal{P}$ paarweise disjunkt und erhalten

$$\overline{\lambda}(F) = \sum_{i=1}^k \sum_{j=1}^{m_i} \lambda(P_{i,j}) = \sum_{i=1}^k \overline{\lambda}(F_i).$$

Also ist $\overline{\lambda}$ der gesuchte Inhalt. $\square$

Folgerung 1.2 (Monotonie und Subadditivität) *Sei λ ein Inhalt auf dem Halbring $\mathcal{P}$ über X. Dann gilt die Implikation*

$$P \subset \bigcup_{i=1}^{k} P_i \textit{ mit } P, P_1, \ldots, P_k \in \mathcal{P} \quad \Rightarrow \quad \lambda(P) \leq \sum_{i=1}^{k} \lambda(P_i).$$

Beweis: Nach Satz 1.4 können wir annehmen, dass $\mathcal{P}$ ein Ring ist. Es gilt zunächst die Monotonie

$$P, Q \in \mathcal{P},\, Q \supset P \quad \Rightarrow \quad \lambda(Q) = \lambda(P) + \underbrace{\lambda(Q \backslash P)}_{\geq 0} \geq \lambda(P).$$

Für $P, P_1, \ldots, P_k$ wie in der Voraussetzung erhalten wir weiter die Subadditivität

$$\lambda\big(\bigcup_{i=1}^{k} P_i\big) = \lambda\big(\bigcup_{i=1}^{k} P_i \backslash (\bigcup_{j=1}^{i-1} P_j)\big) = \sum_{i=1}^{k} \lambda\big(P_i \backslash (\bigcup_{j=1}^{i-1} P_j)\big) \leq \sum_{i=1}^{k} \lambda(P_i).$$

Also folgt $\lambda(P) \leq \lambda(\bigcup_{i=1}^{k} P_i) \leq \sum_{i=1}^{k} \lambda(P_i)$. $\square$

Den vom System $\mathcal{P}^n$ aller Quader $P \subset \mathbb{R}^n$ erzeugten Ring bezeichnen wir mit $\mathcal{F}^n$. Seine Elemente sind die Figuren, das heißt endliche Vereinigungen von Quadern. Für den nach Satz 1.4 fortgesetzten Inhalt schreiben wir $\lambda^n : \mathcal{F}^n \to [0, \infty]$.

Definition 1.6 (Jordanscher Inhalt) *Für eine beschränkte Menge $E \subset \mathbb{R}^n$ definieren wir den äußeren bzw. inneren Jordanschen Inhalt durch*

$$\begin{aligned} \overline{\mathrm{vol}}^n(E) &= \inf\{\lambda^n(F) : F \textit{ Figur mit } F \supset E\} \\ \underline{\mathrm{vol}}^n(E) &= \sup\{\lambda^n(F) : F \textit{ Figur mit } F \subset E\}. \end{aligned}$$

Bei Gleichheit $\overline{\mathrm{vol}}^n(E) = \underline{\mathrm{vol}}^n(E) =: \mathrm{vol}^n(E)$ heißt E quadrierbar, und $\mathrm{vol}^n(E)$ heißt Jordanscher Inhalt von E.

Lemma 1.1 *Die beschränkte Menge $E \subset \mathbb{R}^n$ ist genau dann quadrierbar, wenn es zu jedem $\varepsilon > 0$ Figuren $F_1 \subset E \subset F_2$ gibt mit $\lambda^n(F_2 \backslash F_1) < \varepsilon$.*

Beweis: Für $F_1', F_2' \in \mathcal{F}^n$ mit $F_1' \subset E \subset F_2'$ gilt nach Satz 1.4

$$\lambda^n(F_1') \leq \lambda^n(F_1') + \lambda^n(F_2' \backslash F_1') = \lambda^n(F_2').$$

Indem wir das Supremum über F_1' sowie das Infimum über F_2' bilden, folgt für $F_1, F_2 \in \mathcal{F}^n$

$$F_1 \subset E \subset F_2 \quad \Rightarrow \quad \lambda^n(F_1) \leq \underline{\mathrm{vol}}^n(E) \leq \overline{\mathrm{vol}}^n(E) \leq \lambda^n(F_2) < \infty, \tag{1.7}$$

und hieraus ergibt sich die Behauptung. $\square$

Satz 1.5 (Jordaninhalt) *Für das System $\mathcal{Q}^n$ der quadrierbaren Mengen im $\mathbb{R}^n$ und den Jordanschen Inhalt vol^n gilt:*

(i) *$\mathcal{Q}^n$ ist ein Ring, der den Ring $\mathcal{F}^n$ der Figuren enthält.*

(ii) $\mathrm{vol}^n : \mathcal{Q}^n \to [0, \infty)$ *ist ein Inhalt mit* $\mathrm{vol}^n|_{\mathcal{F}^n} = \lambda^n$.

Beweis: Es ist klar, dass $\emptyset \in \mathcal{Q}^n$ und $\mathrm{vol}^n(\emptyset) = 0$, vgl. den Beweis von Satz 1.1. Für (i) ist zu zeigen, dass mit $E, E' \in \mathcal{Q}^n$ auch $E \backslash E'$ und $E \cup E'$ in $\mathcal{Q}^n$ liegen. Wähle nach Lemma 1.1 Figuren $F_1 \subset E \subset F_2$ bzw. $F_1' \subset E' \subset F_2'$ mit

$$\lambda^n(F_2 \backslash F_1) < \frac{\varepsilon}{2} \quad \text{und} \quad \lambda^n(F_2' \backslash F_1') < \frac{\varepsilon}{2}.$$

Offenbar gilt

$$F_1 \backslash F_2' \subset E \backslash E' \subset F_2 \backslash F_1' \qquad \text{und} \qquad F_1 \cup F_1' \subset E \cup E' \subset F_2 \cup F_2'.$$

Man prüft durch Fallunterscheidung nach:

$$\begin{aligned}
(F_2 \backslash F_1') \backslash (F_1 \backslash F_2') &\subset (F_2 \backslash F_1) \cup (F_2' \backslash F_1'),\\
(F_2 \cup F_2') \backslash (F_1 \cup F_1') &\subset (F_2 \backslash F_1) \cup (F_2' \backslash F_1').
\end{aligned}$$

Damit erhalten wir nach Folgerung 1.2

$$\lambda^n\big((F_2 \backslash F_1') \backslash (F_1 \backslash F_2')\big) \leq \lambda^n(F_2 \backslash F_1) + \lambda^n(F_2' \backslash F_1') < \varepsilon.$$

Analog ergibt sich $\lambda^n\big((F_2 \cup F_2') \backslash (F_1 \cup F_1')\big) < \varepsilon$, und es folgt $E \backslash E'$, $E \cup E' \in \mathcal{Q}^n$ nach Lemma 1.1. Seien nun $E, E' \in \mathcal{Q}^n$ disjunkt. Für Figuren $F_1 \subset E \subset F_2$ bzw. $F_1' \subset E' \subset F_2'$ folgt dann

$$\begin{aligned}
\lambda^n(F_1) + \lambda^n(F_1') &= \lambda^n(F_1 \cup F_1') \quad \text{(nach Satz 1.4)}\\
&\leq \underline{\mathrm{vol}}^n(E \cup E')\\
&= \overline{\mathrm{vol}}^n(E \cup E') \quad (\text{da } E \cup E' \in \mathcal{Q}^n)\\
&\leq \lambda^n(F_2 \cup F_2')\\
&\leq \lambda^n(F_2) + \lambda^n(F_2').
\end{aligned}$$

Bilde das Supremum über alle F_1, F_1' und das Infimum über alle F_2, F_2'. Dies ergibt

$$\mathrm{vol}^n(E \cup E') = \mathrm{vol}^n(E) + \mathrm{vol}^n(E'),$$

und damit die endliche Additivität der Funktion vol^n. Für $F \in \mathcal{F}^n$ wählen wir schließlich $F_1 = F_2 = F$ in (1.7) und erhalten $\lambda^n(F) = \underline{\mathrm{vol}}^n(F) = \overline{\mathrm{vol}}^n(F)$. Damit ist der Satz bewiesen. $\square$

Der Jordansche Inhalt setzt somit den Elementarinhalt $\lambda^n : \mathcal{F}^n \to [0, \infty]$ auf den Ring $\mathcal{Q}^n$ der quadrierbaren Mengen fort. Natürlich stellt sich nun die Frage, wie groß $\mathcal{Q}^n$ eigentlich ist. Hierzu betrachten wir zunächst zwei negative Beispiele.

Beispiel 1.1 Die Menge $E = \mathbb{Q} \cap [0, 1] \subset \mathbb{R}$ ist nicht quadrierbar. Es gilt nämlich:

$$\begin{array}{llcccl}
E \subset F & \text{mit } F \in \mathcal{F}^1 & \Rightarrow & [0,1] = \overline{E} \subset \overline{F} & \Rightarrow & \overline{\mathrm{vol}}^1(E) = 1,\\
F \subset E & \text{mit } F \in \mathcal{F}^1 & \Rightarrow & \mathrm{int}(F) = \emptyset & \Rightarrow & \underline{\mathrm{vol}}^1(E) = 0.
\end{array}$$

Ein einzelner Punkt $\{x\} \subset \mathbb{R}$ ist offensichtlich quadrierbar mit $\mathrm{vol}^1(\{x\}) = 0$. Also ist das System der quadrierbaren Mengen zwar unter endlichen Vereinigungen abgeschlossen, nicht jedoch unter abzählbar unendlichen Vereinigungen.

Beispiel 1.2 Es gibt sogar eine offene Menge $U \subset (0,1)$, die nicht quadrierbar ist. Wähle dazu eine Abzählung $\mathbb{Q} \cap (0,1) = \{q_1, q_2, \ldots\}$ und setze für $\varepsilon \in (0, 1/2)$

$$U = \bigcup_{k=1}^{\infty} U_k \quad \text{mit } U_k = (q_k - 2^{-k}\varepsilon, q_k + 2^{-k}\varepsilon) \cap (0,1).$$

Wir behaupten $\underline{\mathrm{vol}}^1(U) \leq 2\varepsilon < 1 = \overline{\mathrm{vol}}^1(U)$. Für die linke Ungleichung betrachten wir eine beliebige Figur $F \subset U$. Ist F kompakt, so gilt nach dem Überdeckungssatz von Heine-Borel schon $F \subset \bigcup_{k=1}^m U_k$ für ein $m < \infty$, und es folgt

$$\lambda^1(F) \leq \sum_{k=1}^{m} \lambda^1(U_k) \leq \sum_{k=1}^{\infty} 2^{1-k}\varepsilon = 2\varepsilon.$$

Ist F nicht kompakt, so wählen wir eine Folge kompakter Figuren $F_j \subset F$ mit $\lambda^1(F_j) \nearrow \lambda^1(F)$, womit die Ungleichung $\underline{\mathrm{vol}}^1(U) \leq 2\varepsilon$ gezeigt ist. Jede Figur $F \supset U$ erfüllt andererseits $\overline{F} \supset \overline{U} = [0,1]$, und dies bedeutet $\overline{\mathrm{vol}}^1(U) = 1$. Somit ist U ist nicht quadrierbar.

Um die quadrierbaren Mengen allgemein zu charakterisieren, benötigen wir das folgende Verfahren, das beliebige Mengen durch Figuren in einem Gitter approximiert. Dieses Lemma wird an mehreren Stellen der Vorlesung eingesetzt.

Lemma 1.2 (Approximation durch Gitterfiguren) *Betrachte die Würfelfamilie $\mathcal{W}_k = \{Q_{k,m} = 2^{-k}(m + [0,1]^n) : m \in \mathbb{Z}^n\}$ für $k \in \mathbb{N}_0$ und definiere für $E \subset \mathbb{R}^n$ die Mengen*

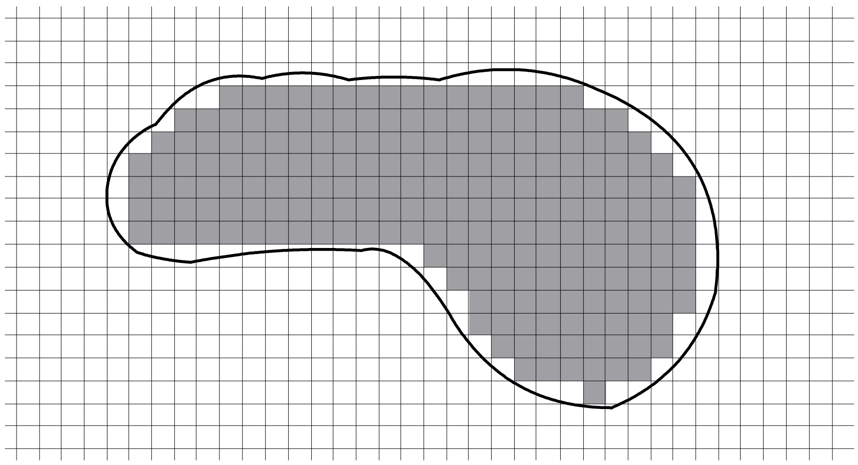
$$F_k(E) = \bigcup \{Q \in \mathcal{W}_k : Q \subset E\}, \quad F^k(E) = \bigcup \{Q \in \mathcal{W}_k : Q \cap E \neq \emptyset\}.$$

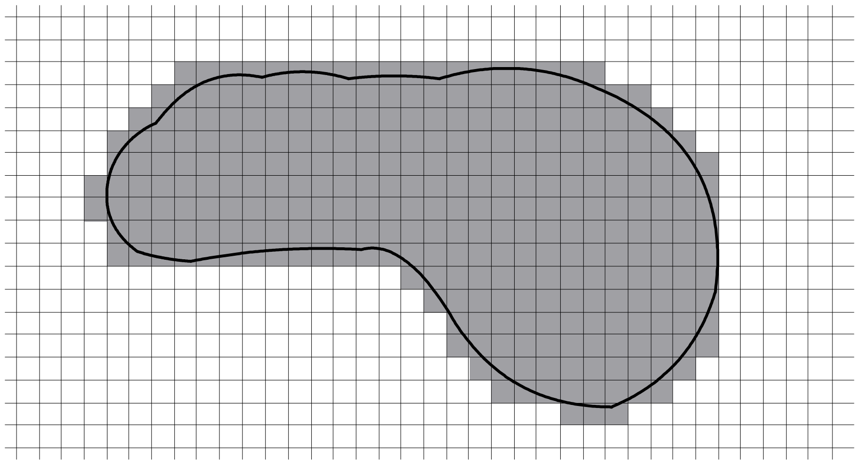
Dann ist $F_k(E)$ bzw. $F^k(E)$ abgeschlossene Vereinigung von abzählbar vielen Würfeln mit paarweise disjunktem Inneren, und es gelten folgende Aussagen:

(i) $F_1(E) \subset F_2(E) \subset \ldots \subset E$ *und* $F^1(E) \supset F^2(E) \supset \ldots \supset E$.

(ii) $$F_k(E) \supset \{x \in \mathbb{R}^n : \mathrm{dist}(x, \mathbb{R}^n \backslash E) > 2^{-k}\sqrt{n}\}$$
$$F^k(E) \subset \{x \in \mathbb{R}^n : \mathrm{dist}(x, E) \leq 2^{-k}\sqrt{n}\}.$$

(iii) $\mathrm{int}(E) \subset \bigcup_{k=1}^{\infty} F_k(E) \subset E$ *sowie* $\overline{E} \supset \bigcap_{k=1}^{\infty} F^k(E) \supset E$.





Beweis: Die Würfel in $\mathcal{W}_k$ sind kompakt, haben paarweise disjunktes Inneres und jede beschränkte Menge wird nur von endlich vielen Würfeln in $\mathcal{W}_k$ getroffen. Insbesondere sind $F_k(E), F^k(E)$ abgeschlossen. Nun ist $Q_{k,m}$ die Vereinigung der 2^n Teilwürfel $Q_{k+1,2m+l}$ mit $l \in \{0,1\}^n$, und es gilt

$$\begin{aligned} Q_{k,m} \subset E \quad &\Rightarrow \quad Q_{k+1,2m+l} \subset E \qquad \text{für alle } l \in \{0,1\}^n, \\ Q_{k+1,2m+l} \cap E \neq \emptyset \quad &\Rightarrow \quad Q_{k,m} \cap E \neq \emptyset \qquad \text{wobei } l \in \{0,1\}^n. \end{aligned}$$

Daraus folgt Aussage (i). Sei nun $x \in \mathbb{R}^n$ beliebig. Dann gilt $x \in Q$ für ein $Q \in \mathcal{W}_k$, und es folgt wegen $\operatorname{diam}(Q) = 2^{-k}\sqrt{n}$

$$\operatorname{dist}(x, \mathbb{R}^n \backslash E) > 2^{-k}\sqrt{n} \quad \Rightarrow \quad Q \subset E \quad \Rightarrow \quad x \in F_k(E).$$

Ist andererseits $x \in F^k(E)$, so gilt $x \in Q$ für ein $Q \in \mathcal{W}_k$ mit $Q \cap E \neq \emptyset$. Also gilt

$$x \in F^k(E) \quad \Rightarrow \quad \operatorname{dist}(x, E) \leq \operatorname{diam}(Q) \leq 2^{-k}\sqrt{n},$$

womit beide Behauptungen in (ii) gezeigt sind. Behauptung (iii) folgt. $\square$

Satz 1.6 (Quadrierbarkeitskriterium) *Für eine beschränkte Menge $E \subset \mathbb{R}^n$ gilt:*

$$E \text{ quadrierbar} \quad \Leftrightarrow \quad \overline{\mathrm{vol}}^n(\partial E) = 0.$$

Beweis: Sei E quadrierbar. Dann gibt es nach Lemma 1.1 zu $\varepsilon > 0$ Figuren $F_1 \subset E \subset F_2$ mit $\lambda^n(F_2 \backslash F_1) < \varepsilon$. Es sei oBdA F_1 offen und F_2 kompakt, andernfalls schreibe F_1, F_2 als Vereinigung endlich vieler Quader und gehe zu den offenen bzw. abgeschlossenen Quadern über. Dabei ändert sich $\lambda(F_2 \backslash F_1) = \lambda(F_2) - \lambda(F_1)$ nicht. Es folgt

$$\partial E = \overline{E} \backslash \mathrm{int}(E) \subset F_2 \backslash F_1 \quad \Rightarrow \quad \overline{\mathrm{vol}}^n(\partial E) \leq \lambda^n(F_2 \backslash F_1) < \varepsilon.$$

Sei jetzt umgekehrt $\overline{\mathrm{vol}}^n(\partial E) = 0$, das heißt zu $\varepsilon > 0$ gibt es eine Figur $F \supset \partial E$ mit $\lambda^n(F) < \varepsilon$. Indem wir F als endliche Vereinigung von Quadern schreiben und deren Kanten etwas vergrößern, können wir annehmen, dass F offen und damit $\mathbb{R}^n \backslash F$ abgeschlossen ist. Damit erhalten wir

$$\inf_{\mathbb{R}^n \backslash F} f > 0 \quad \text{für } f(x) = \max\{\mathrm{dist}(x, E), \mathrm{dist}(x, \mathbb{R}^n \backslash E)\}.$$

Denn f ist stetig und $f(x) \to \infty$ mit $|x| \to \infty$, da E beschränkt, also nimmt f das Infimum in einem Punkt $x_0 \in \mathbb{R}^n \backslash F$ an. Wäre $f(x_0) = 0$ so folgt $x_0 \in \overline{E} \cap \overline{\mathbb{R}^n \backslash E} = \partial E \subset F$, ein Widerspruch.

Jetzt verwende Lemma 1.2 und erhalte Figuren $F_k(E) \subset E \subset F^k(E)$, wobei nach (ii)

$$F^k(E) \backslash F_k(E) \subset \{x \in \mathbb{R}^n : f(x) \leq 2^{-k}\sqrt{n}\}.$$

Es folgt $F^k(E) \backslash F_k(E) \subset F$ und folglich $\lambda^n(F^k(E) \backslash F_k(E)) \leq \lambda^n(F) < \varepsilon$ für k hinreichend groß. Nach Lemma 1.1 ist damit E quadrierbar. □

2 Äußere Maße und Maße

Im Unterschied zum Inhalt ist der Definitionsbereich eines Maßes eine σ-Algebra, also abgeschlossen unter unter abzählbaren Mengenoperationen, und das Maß ist darauf σ-additiv, das heißt es gilt für abzählbar viele, paarweise disjunkte Mengen A_i

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

Hier steht der Buchstabe σ für„abzählbar unendlich", in Abgrenzung zu „endlich".

Wir beginnen mit einigen Vorbemerkungen zum Umgang mit den Symbolen ∞ (genauer: $+\infty$) und $-\infty$. Auf der erweiterten Zahlengeraden $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty, -\infty\}$ sind die Ordnungsrelation $-\infty < a < \infty$ für $a \in \mathbb{R}$ und der Konvergenzbegriff auf naheliegende Weise gegeben.

Definition 2.1 (Konvergenz in $\overline{\mathbb{R}}$) *Eine Folge $s_k \in \overline{\mathbb{R}}$ ($k \in \mathbb{N}$) konvergiert gegen $s \in \overline{\mathbb{R}}$, falls eine der folgenden Alternativen gilt:*

(i) *$s \in \mathbb{R}$, und für jedes $\varepsilon > 0$ gilt $s_k \in (s - \varepsilon, s + \varepsilon) \subset \mathbb{R}$ für k hireichend groß.*

(ii) *$s = \infty$, und für jedes $r \in \mathbb{R}$ gilt $s_k \in (r, \infty]$ für k hinreichend groß.*

(iii) *$s = -\infty$, und für jedes $r \in \mathbb{R}$ gilt $s_k \in [-\infty, r)$ für k hinreichend groß.*

Eine Folge $s_k \in \mathbb{R}$ ist genau dann in $\overline{\mathbb{R}}$ konvergent, wenn sie entweder in $\mathbb{R}$ konvergiert oder bestimmt divergiert gegen ∞ bzw. gegen $-\infty$. Der Grenzwert einer monoton wachsenden Folge, bzw. einer Reihe mit nichtnegativen Gliedern, ist also immer existent. Die Addition wird wie folgt auf $\overline{\mathbb{R}}$ fortgesetzt:

$+$	$-\infty$	$\mathbb{R}$	$+\infty$
$-\infty$	$-\infty$	$-\infty$	$-$
$\mathbb{R}$	$-\infty$	$\mathbb{R}$	$+\infty$
$+\infty$	$-$	$+\infty$	$+\infty$

Schließlich verwenden wir die Vereinbarungen

$$\sup \emptyset = -\infty \quad \text{und} \quad \inf \emptyset = +\infty.$$

Diese sind konsistent mit der Tatsache, dass für Mengen $A, B \subset \mathbb{R}$ stets gilt:

$$A \subset B \quad \Rightarrow \quad \sup A \leq \sup B \quad \text{sowie} \quad \inf A \geq \inf B.$$

Definition 2.2 (äußeres Maß) *Sei X eine Menge. Eine Funktion $\mu : 2^X \to [0, \infty]$ mit $\mu(\emptyset) = 0$ heißt äußeres Maß auf X, falls gilt:*

$$A \subset \bigcup_{i=1}^{\infty} A_i \quad \Rightarrow \quad \mu(A) \leq \sum_{i=1}^{\infty} \mu(A_i). \tag{2.1}$$

Oft spricht man der Kürze halber auch von einem Maß, wenn keine Misverständnisse entstehen können (vgl. Definition 2.6). Für eine endliche Überdeckung folgt aus Definition 2.2, indem wir $A_i = \emptyset$ für $i > k$ setzen,

$$A \subset \bigcup_{i=1}^{k} A_i \quad \Rightarrow \quad \mu(A) \leq \sum_{i=1}^{k} \mu(A_i).$$

Insbesondere ergibt sich

$$(2.2) \qquad A \subset B \quad \Rightarrow \quad \mu(A) \leq \mu(B) \quad \text{(Monotonie von } \mu\text{)}.$$

Weiter haben wir

$$(2.3) \qquad \mu\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mu(A_i) \quad (\sigma\text{-Subadditivität}).$$

Umgekehrt folgt aus (2.2) und (2.3) offensichtlich die Eigenschaft (2.1).

Beispiel 2.1 Der äußere Jordansche Inhalt $\overline{\text{vol}}^1$ ist kein äußeres Maß auf $\mathbb{R}$. Mit der Abzählung $\mathbb{Q} \cap [0,1] = \bigcup_{k=1}^{\infty} \{q_k\}$ gilt nämlich, vgl. Beispiel 1.1,

$$\overline{\text{vol}}^1(\mathbb{Q} \cap [0,1]) = 1 > 0 = \sum_{k=1}^{\infty} \overline{\text{vol}}^1(\{q_k\}).$$

Auch der innere Jordansche Inhalt $\underline{\text{vol}}^1$ ist kein äußeres Maß. Denn mit $E_k = [0,1] \backslash \bigcup_{i=k}^{\infty} \{q_i\}$ gilt, da $\{q_k, q_{k+1}, \ldots\}$ dicht in $[0,1]$ ist,

$$[0,1] = \bigcup_{k=1}^{\infty} E_k \quad \text{aber} \quad \underline{\text{vol}}^1([0,1]) = 1 > 0 = \sum_{k=1}^{\infty} \underline{\text{vol}}^1(E_k).$$

Definition 2.3 (Messbarkeit) *Sei μ äußeres Maß auf X. Die Menge $A \subset X$ heißt μ-messbar, falls gilt*

$$(2.4) \qquad \mu(S) \geq \mu(S \cap A) + \mu(S \backslash A) \quad \textit{für alle } S \subset X.$$

Die umgekehrte Ungleichung in (2.4) folgt mit $S = (S \cap A) \cup (S \backslash A)$ aus (2.3), also gilt

$$(2.5) \qquad A \quad \text{messbar} \quad \Rightarrow \quad \mu(S) = \mu(S \cap A) + \mu(S \backslash A) \quad \forall S \subset X.$$

Wir wollen die Definitionen nun an ein paar einfachen Beispielen betrachten.

Beispiel 2.2 Für einen Punkt $x \in X$ ist das zugehörige Diracmaß gegeben durch

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

Es gilt $\delta_x(A) \in \{0,1\}$ und $\delta_x(\emptyset) = 0$ per Definition. Ist eine Überdeckung $A \subset \bigcup_{k=1}^{\infty} A_k$ gegeben und ist $x \in A$, so folgt $x \in A_k$ für (mindestens) ein k. Hieraus folgt die Eigenschaft (2.1), denn im Fall $x \notin A$ gilt ohnehin $\delta_x(A) = 0$. Alle Mengen $A \subset X$ sind messbar bzgl. δ_x. Ist nämlich $x \notin S$ so sind beide Seiten in (2.4) Null, und für $x \in S$ liegt x in genau einer der Mengen $S \cap A$ bzw. $S \backslash A$.

Beispiel 2.3 Sei $N_n = \{1, 2, \ldots, n\}$ für $n \in \mathbb{N}$. Auf einer Menge X definieren wir das Zählmaß $\mathrm{card} : 2^X \to [0, +\infty]$ durch

$$\mathrm{card}(A) = \begin{cases} \sup\{n \in \mathbb{N} : \text{ es gibt eine injektive Abbildung } \varphi : N_n \to A\} & \text{falls } A \neq \emptyset \\ 0 & \text{falls } A = \emptyset. \end{cases}$$

Um die Maßeigenschaft (2.1) für card zu zeigen, erinnern wir an das Schubfachprinzip

$$\varphi : N_n \to N_m \text{ injektiv} \quad \Rightarrow \quad n \leq m.$$

Dies wird durch Induktion über m, jeweils für alle $n \in \mathbb{N}$, bewiesen. Der Induktionsanfang $m = 1$ ist trivial. Ist eine injektive Abbildung $\varphi : N_n \to N_{m+1}$ gegeben, wobei $m \geq 1$ und oBdA $n \geq 2$, so erhalten wir die Injektion

$$\varphi' : N_{n-1} \to N_m,\ \varphi'(i) = \begin{cases} \varphi(i) & \text{falls } \varphi(i) \neq m+1, \\ \varphi(n) & \text{falls } \varphi(i) = m+1. \end{cases}$$

Es folgt induktiv $n - 1 \leq m$, also $n \leq m + 1$. Insbesondere gilt $\mathrm{card}(N_n) = n$ für $n \in \mathbb{N}$. Ist $\mathrm{card}(A) = n \in \mathbb{N}$, so gibt es eine Injektion $\varphi : N_n \to A$ und diese ist sogar bijektiv, denn andernfalls gäbe es eine Injektion $\varphi' : N_{n+1} \to A$. Wir behaupten nun zunächst

$$\mathrm{card}(A_1 \cup A_2) \leq \mathrm{card}(A_1) + \mathrm{card}(A_2).$$

Dazu können wir annehmen, dass $\mathrm{card}(A_i) = \nu_i \in \mathbb{N}$ für $i = 1, 2$, das heißt es gibt Bijektionen $\varphi_i : N_{\nu_i} \to A_i$. Ist eine injektive Abbildung $\varphi : N_k \to A_1 \cup A_2$ gegeben, so definieren wir

$$\Phi : N_k \to N_{\nu_1+\nu_2},\ \Phi(i) = \begin{cases} \varphi_1^{-1}(\varphi(i)) & \text{falls } \varphi(i) \in A_1 \\ \nu_1 + \varphi_2^{-1}(\varphi(i)) & \text{sonst.} \end{cases}$$

Φ ist wohldefiniert und injektiv, also folgt aus dem Schubfachprinzip $k \leq \nu_1 + \nu_2$, was zu zeigen war. Daraus ergibt sich weiter die σ-Subadditivität

$$\mathrm{card}\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mathrm{card}(A_i).$$

Denn ist die rechte Seite endlich, so gibt es ein $k \in \mathbb{N}$ mit $A_i = \emptyset$ für alle $i > k$, und die endliche Ungleichung $\mathrm{card}(\bigcup_{i=1}^{k} A_i) \leq \sum_{i=1}^{k} \mathrm{card}(A_i)$ folgt aus dem oben behandelten Fall $k = 2$ sofort durch Induktion. Aus der Definition folgt außerdem die Monotonie

$$A \subset B \quad \Rightarrow \quad \mathrm{card}(A) \leq \mathrm{card}(B).$$

Damit ist gezeigt, dass card ein äußeres Maß ist.

Wir behaupten weiter, dass alle Mengen $A \subset X$ card-messbar sind. Wähle dazu für $S \subset X$, oBdA $\mathrm{card}(S) < \infty$, Bijektionen $\varphi_1 : N_{\nu_1} \to S \cap A$ und $\varphi_2 : N_{\nu_2} \to S \backslash A$ und setze

$$\varphi : N_{\nu_1+\nu_2} \to S,\ \varphi(i) = \begin{cases} \varphi_1(i) & \text{falls } i \in \{1, \ldots, \nu_1\} \\ \varphi_2(i - \nu_1) & \text{falls } i \in \{\nu_1 + 1, \ldots, \nu_1 + \nu_2\}. \end{cases}$$

Die Abbildung φ ist wohldefiniert und injektiv, also schließen wir

$$\mathrm{card}(S) \geq \nu_1 + \nu_2 = \mathrm{card}(S \cap A) + \mathrm{card}(S \backslash A).$$

Beispiel 2.4 Auf jeder Menge X erhalten wir ein blödes Maß β durch

$$\beta(A) = \begin{cases} 0 & \text{falls } A = \emptyset \\ 1 & \text{sonst.} \end{cases}$$

Der Nachweis von (2.1) für β ist trivial. Es sind nur $\emptyset$ und X β-messbar, denn mit der Wahl $S = X$ in (2.4) folgt, falls $A \subset X$ β-messbar ist,

$$1 \geq \beta(X) = \beta(A) + \beta(X \backslash A).$$

Beispiel 2.5 Für eine Familie μ_λ, $\lambda \in \Lambda$ von äußeren Maßen auf X sei

$$\mu(A) = \sup_{\lambda \in \Lambda} \mu_\lambda(A).$$

Dann ist μ ein äußeres Maß auf X, denn für $A \subset \bigcup_{i=1}^\infty A_i$ und festes $\lambda \in \Lambda$ gilt

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

Bilde nun auf der linken Seite das Supremum über alle $\lambda \in \Lambda$ und erhalte (2.1). Es ist im allgemeinen nicht klar, welche Mengen bzgl. μ messbar sind.

Die folgenden beiden Konstruktionen von Maßen werden öfters benutzt.

Satz 2.1 (Bildmaß) *Seien X, Y Mengen und $f : X \to Y$. Für ein gegebenes äußeres Maß $\mu : 2^X \to [0, \infty]$ erhält man ein äußeres Maß $f(\mu)$ auf Y durch*

$$f(\mu) : 2^Y \to [0, \infty],\ f(\mu)(B) = \mu(f^{-1}(B)).$$

$f(\mu)$ heißt Bildmaß von μ unter f, und es gilt für alle $B \subset Y$

$$f^{-1}(B)\ \mu\text{-messbar} \quad \Rightarrow \quad B\ f(\mu)\text{-messbar}\,. \tag{2.6}$$

Beweis: Für $B \subset \bigcup_{i=1}^\infty B_i$ gilt $f^{-1}(B) \subset \bigcup_{i=1}^\infty f^{-1}(B_i)$ und folglich nach Definition 2.1

$$f(\mu)(B) = \mu\big(f^{-1}(B)\big) \leq \sum_{i=1}^\infty \mu\big(f^{-1}(B_i)\big) = \sum_{i=1}^\infty f(\mu)(B_i).$$

Außerdem ist trivialerweise $f(\mu)(\emptyset) = \mu(f^{-1}(\emptyset)) = \mu(\emptyset) = 0$. Nun ist nach Definition von $f(\mu)$ die Menge $B \subset Y$ genau dann $f(\mu)$-messbar, wenn gilt:

$$\mu(f^{-1}(T)) \geq \mu(f^{-1}(T \cap B)) + \mu(f^{-1}(T \backslash B)) \quad \text{für alle } T \subset Y.$$

Da $f^{-1}(T \cap B) = f^{-1}(T) \cap f^{-1}(B)$ sowie $f^{-1}(T \backslash B) = f^{-1}(T) \backslash f^{-1}(B)$, ist dies äquivalent zu

$$\mu(f^{-1}(T)) \geq \mu(f^{-1}(T) \cap f^{-1}(B)) + \mu(f^{-1}(T) \backslash f^{-1}(B)) \quad \text{für alle } T \subset Y.$$

Dagegen ist $f^{-1}(B)$ genau dann μ-messbar, wenn

$$\mu(S) \geq \mu(S \cap f^{-1}(B)) + \mu(S \backslash f^{-1}(B)) \quad \text{für alle } S \subset X.$$

Also folgt die Behauptung (2.6), indem wir $S = f^{-1}(T)$ setzen. □

Satz 2.2 (Einschränkung) *Sei $\mu : 2^X \to [0,\infty]$ ein äußeres Maß auf X. Für eine gegebene Menge $M \subset X$ erhält man ein äußeres Maß $\mu \llcorner M$ auf X durch*

$$\mu \llcorner M : 2^X \to [0,\infty],\ \mu \llcorner M(A) = \mu(A \cap M).$$

$\mu \llcorner M$ heißt Einschränkung von μ auf M, und es gilt

$$A \ \mu\text{-messbar} \quad \Rightarrow \quad A \ \mu \llcorner M\text{-messbar}. \tag{2.7}$$

Beweis: Aus der Definition folgt sofort, dass $\mu \llcorner M$ ein äußeres Maß ist. Weiter gilt für $A \subset X$ μ-messbar und $S \subset X$ beliebig

$$\begin{aligned}
\mu \llcorner M(S) &= \mu(S \cap M) \\
&\geq \mu((S \cap M) \cap A) + \mu((S \cap M) \backslash A) \quad (\text{da } A \ \mu\text{-messbar}) \\
&= \mu((S \cap A) \cap M) + \mu((S \backslash A) \cap M) \\
&= \mu \llcorner M(S \cap A) + \mu \llcorner M(S \backslash A).
\end{aligned}$$

Dies beweist Behauptung (2.7). □

Definition 2.4 (Nullmenge) *Sei μ äußeres Maß auf X. Die Menge $N \subset X$ heißt μ-Nullmenge, falls $\mu(N) = 0$.*

Proposition 2.1 (Messbarkeit von Nullmengen) *Sei μ äußeres Maß auf X. Dann gilt:*

$$N \ \text{Nullmenge} \quad \Rightarrow \quad N \ \mu\text{-messbar} \tag{2.8}$$

$$N_1, N_2, \ldots \ \text{Nullmengen} \quad \Rightarrow \quad \bigcup_{k=1}^{\infty} N_k \ \text{Nullmenge}. \tag{2.9}$$

Beweis: Sei $\mu(N) = 0$. Für $S \subset X$ folgt aus der Monotonie $\mu(S \cap N) \leq \mu(N) = 0$, also

$$\mu(S) \geq \mu(S \backslash N) = \mu(S \cap N) + \mu(S \backslash N).$$

Die zweite Aussage gilt nach Definition 2.2. □

Wir wollen nun allgemein die Struktur des Systems $\mathcal{M}$ der μ-messbaren Mengen untersuchen. Auf jeden Fall enthält $\mathcal{M}$ alle Nullmengen $N \subset X$, und damit auch deren Komplemente $X \backslash N$, wie unten in Satz 2.3 gezeigt. Es kann sein, dass keine anderen Mengen messbar sind, zum Beispiel ist $\mathcal{M} = \{\emptyset, X\}$ in Beispiel 2.4. Aber in den relevanten Fällen erwarten wir doch, dass es viele weitere messbare Mengen gibt. Jedenfalls ist das System $\mathcal{M}$ unter abzählbaren Vereinigungen und Durchschnitten abgeschlossen, im Gegensatz zum Beispiel zu den quadrierbaren Mengen $\mathcal{Q}^n$ aus Satz 1.5. Dies soll nun gezeigt werden.

Definition 2.5 (σ-Algebra) *Ein Mengensystem $\mathcal{A} \subset 2^X$ heißt σ-Algebra, wenn gilt:*

(i) $X \in \mathcal{A}$

(ii) $A \in \mathcal{A} \Rightarrow X \backslash A \in \mathcal{A}$

(iii) $A_i \in \mathcal{A}$ *für* $i = 1, 2, \ldots \quad \Rightarrow \quad \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$.

Eine σ-Algebra $\mathcal{A}$ ist auch unter abzählbaren Durchschnitten abgeschlossen, das heißt es gilt

$$(2.10) \qquad A_i \in \mathcal{A} \quad \text{für } i = 1, 2, \ldots \quad \Rightarrow \quad \bigcap_{i=1}^{\infty} A_i \in \mathcal{A}.$$

Dies folgt sofort aus der Darstellung $\bigcap_{i=1}^{\infty} A_i = X \backslash \left(\bigcup_{i=1}^{\infty} X \backslash A_i\right)$. Außerdem stellen wir fest: jede σ-Algebra ist ein Ring, denn nach Definition gilt $\emptyset = X \backslash X \in \mathcal{A}$ sowie

$$A, B \in \mathcal{A} \quad \Rightarrow \quad A \backslash B = A \cap (X \backslash B) \in \mathcal{A}.$$

Lemma 2.1 *Seien $A_1, A_2, \ldots, A_k \subset X$ paarweise disjunkt und μ-messbar. Dann gilt*

$$\mu(S \cap \bigcup_{i=1}^{k} A_i) = \sum_{i=1}^{k} \mu(S \cap A_i) \quad \textit{für alle } S \subset X.$$

Beweis: Für $k = 1$ ist die Aussage trivial, und für $k \geq 2$ folgt induktiv, da A_k μ-messbar,

$$\begin{aligned}
\mu(S \cap \bigcup_{i=1}^{k} A_i) &= \mu\big((S \cap \bigcup_{i=1}^{k} A_i) \cap A_k\big) + \mu\big((S \cap \bigcup_{i=1}^{k} A_i) \backslash A_k\big) \\
&= \mu(S \cap A_k) + \mu(S \cap \bigcup_{i=1}^{k-1} A_i) \\
&= \sum_{i=1}^{k} \mu(S \cap A_i).
\end{aligned}$$

□

Satz 2.3 (äußeres Maß ⇒ Maß) *Sei $\mu : X \to [0, \infty]$ ein äußeres Maß. Dann ist das System $\mathcal{M}$ der μ-messbaren Mengen eine σ-Algebra, und es gilt*

$$\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i) \quad \textit{für abzählbar viele, paarweise disjunkte } A_i \in \mathcal{M}.$$

Beweis: Es gilt $X \in \mathcal{M}$, denn für jede Menge $S \subset X$ ist

$$\mu(S \cap X) + \mu(S \backslash X) = \mu(S) + \mu(\emptyset) = \mu(S).$$

Mit $A \in \mathcal{M}$ folgt auch $X \backslash A \in \mathcal{M}$, denn für $S \subset X$ gilt

$$\mu\big(S \cap (X \backslash A)\big) + \mu\big(S \backslash (X \backslash A)\big) = \mu(S \backslash A) + \mu(S \cap A) = \mu(S).$$

Als nächstes zeigen wir, dass $A \cup B \in \mathcal{M}$ für $A, B \in \mathcal{M}$, und zwar gilt für $S \subset X$ beliebig

$$\begin{aligned}
\mu\big(S \cap (A \cup B)\big) + \mu\big(S \backslash (A \cup B)\big) &\leq \mu(S \cap A) + \mu\big((S \backslash A) \cap B\big) + \mu\big((S \backslash A) \backslash B\big) \\
&= \mu(S \cap A) + \mu(S \backslash A) \quad (\text{ da } B \in \mathcal{M}) \\
&= \mu(S) \quad (\text{da } A \in \mathcal{M}).
\end{aligned}$$

Hieraus folgt auch $A \cap B = X \backslash ((X \backslash A) \cup (X \backslash B)) \in \mathcal{M}$ und $A \backslash B = A \cap (X \backslash B) \in \mathcal{M}$. Per Induktion erhalten wir, dass $\mathcal{M}$ unter endlichen Vereinigungen und Durchschnitten abgeschlossen ist. Jetzt zeigen wir die noch fehlende Behauptung

$$A_i \in \mathcal{M} \quad \text{für } i = 1, 2, \ldots \quad \Rightarrow \quad A = \bigcup_{i=1}^{\infty} A_i \in \mathcal{M}.$$

Wir können dazu annehmen, dass $A_i \cap A_j = \emptyset$ für $i \neq j$, andernfalls betrachten wir $\tilde{A}_i = A_i \backslash (A_1 \cup \ldots \cup A_{i-1})$. Für $S \subset X$ beliebig folgt, da $\bigcup_{i=1}^{k} A_i \in \mathcal{M}$,

$$\mu(S) = \mu(S \cap \bigcup_{i=1}^{k} A_i) + \mu(S \backslash \bigcup_{i=1}^{k} A_i) \geq \sum_{i=1}^{k} \mu(S \cap A_i) + \mu(S \backslash A).$$

Im zweiten Schritt wurden dabei Lemma 2.1 und die Monotonie von μ benutzt. Mit $k \to \infty$ erhalten wir schließlich wegen der σ-Subadditivität von μ

$$\mu(S) \geq \sum_{i=1}^{\infty} \mu(S \cap A_i) + \mu(S \backslash A) \geq \mu\big(\bigcup_{i=1}^{\infty} (S \cap A_i)\big) + \mu(S \backslash A) = \mu(S \cap A) + \mu(S \backslash A).$$

Also ist $A = \bigcup_{i=1}^{\infty} A_i$ μ-messbar. Schließlich folgt, indem wir in Lemma 2.1 $S = X$ wählen,

$$\lim_{k \to \infty} \mu\big(\bigcup_{i=1}^{k} A_i\big) = \sum_{i=1}^{\infty} \mu(A_i) \geq \mu\big(\bigcup_{i=1}^{\infty} A_i\big) \geq \lim_{k \to \infty} \mu\big(\bigcup_{i=1}^{k} A_i\big),$$

wobei wieder die σ-Subadditivität und die Monotonie benutzt wurden. benutzt wurden. Damit ist der Satz bewiesen. □

An dieser Stelle fühen wir den Begriff des Maßes ein, wie er in der Wahrscheinlichkeitstheorie üblicherweise definiert wird; dabei ist eine σ-Algebra ad hoc gegeben.

Definition 2.6 (Maß) *Sei $\mathcal{A} \subset 2^X$ eine σ-Algebra. Eine Funktion $\mu : \mathcal{A} \to [0, \infty]$ mit $\mu(\emptyset) = 0$ heißt Maß auf $\mathcal{A}$, falls für jede paarweise disjunkte Folge $A_i \in \mathcal{A}$ gilt:*

$$\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i) \qquad (\sigma\text{-Additivität}). \tag{2.11}$$

Das Tripel $(X, \mathcal{A}, \mu)$ wird dann auch als Maßraum bezeichnet.

Für jedes äußere Maß μ auf X erhält man nach Satz 2.3 kanonisch ein Maß, indem man μ auf die σ-Algebra der μ-messbaren Mengen einschränkt.

Satz 2.4 (Stetigkeitseigenschaften von Maßen) *Sei μ ein Maß auf $\mathcal{A}$. Dann gelten für Mengen $A_1, A_2, \ldots$ in $\mathcal{A}$ folgende Aussagen:*

(i) $A_1 \subset A_2 \subset \ldots \quad \Rightarrow \quad \mu\big(\bigcup_{i=1}^{\infty} A_i\big) = \lim_{k \to \infty} \mu(A_k)$ *(Stetigkeit von unten)*

(ii) $A_1 \supset A_2 \supset \ldots,\ \mu(A_1) < \infty \quad \Rightarrow \quad \mu\big(\bigcap_{i=1}^{\infty} A_i\big) = \lim_{k \to \infty} \mu(A_k)$ *(Stetigkeit von oben).*

BEWEIS: Für (i) setze $\tilde{A}_k = A_k \backslash \bigcup_{i=1}^{k-1} A_i$ und berechne unter Verwendung von (2.11)

$$\mu\Big(\bigcup_{i=1}^{\infty} A_i\Big) = \mu\Big(\bigcup_{i=1}^{\infty} \tilde{A}_i\Big) = \sum_{i=1}^{\infty} \mu(\tilde{A}_i) = \lim_{k\to\infty} \mu\Big(\bigcup_{i=1}^{k} \tilde{A}_i\Big) = \lim_{k\to\infty} \mu(A_k).$$

Für (ii) betrachte die aufsteigende Folge $A'_k = A_1 \backslash A_k$. Es gilt

$$\mu(A_1) = \mu(A_1 \cap A_k) + \mu(A_1 \backslash A_k) = \mu(A_k) + \mu(A'_k).$$

Daraus folgt wegen (i)

$$\mu(A_1) - \lim_{k\to\infty} \mu(A_k) = \lim_{k\to\infty} \mu(A'_k) = \mu\Big(\bigcup_{i=1}^{\infty} A'_i\Big) = \mu\Big(A_1 \backslash \bigcap_{i=1}^{\infty} A_i\Big) = \mu(A_1) - \mu\Big(\bigcap_{i=1}^{\infty} A_i\Big).$$

□

Beispiel 2.6 Die Bedingung $\mu(A_1) < \infty$ in (iii) kann natürlich durch $\mu(A_k) < \infty$ für ein k ersetzt werden. Sie kann aber nicht ganz weggelassen werden. Mit $A_k = \{k, k+1, \dots\} \subset \mathbb{N}$ gilt zum Beispiel $\operatorname{card}(A_k) = \infty$ für alle $k \in \mathbb{N}$, aber $\operatorname{card}(\bigcap_{i=1}^{\infty} A_i) = \operatorname{card}(\emptyset) = 0$.

3 Der Fortsetzungssatz von Caratheodory

Der elementargeometrische Inhalt liefert eine vernünftige Definition des n-dimensionalen Volumens auf der Klasse der Quader $\mathcal{P}^n$ oder der Klasse der Figuren $\mathcal{F}^n$. Wir stehen nun vor der Aufgabe, diesen Inhalt zu einem äußeren Maß auf das System aller Teilmengen des $\mathbb{R}^n$ fortzusetzen, wobei die Mengen in $\mathcal{P}^n$ natürlich messbar sein sollen. In diesem Kapitel beschreiben wir in abstraktem Rahmen die Konstruktion der Maßfortsetzung nach Caratheodory und beantworten auch die Frage, in welchem Umfang die Fortsetzung eindeutig bestimmt ist. Die Spezialisierung auf das Volumenmaß im $\mathbb{R}^n$, genauer das Lebesguemaß, folgt im nächsten Kapitel.

Definition 3.1 (Fortsetzung) *Sei λ ein Inhalt auf dem Halbring $\mathcal{P} \subset 2^X$. Ein äußeres Maß μ auf X heißt Fortsetzung von λ, falls folgende zwei Bedingungen gelten:*

(i) $\mu|_{\mathcal{P}} = \lambda$, *das heißt* $\mu(P) = \lambda(P)$ *für alle* $P \in \mathcal{P}$

(ii) $\mathcal{P} \subset \mathcal{M}$, *also alle Mengen in* $\mathcal{P}$ *sind* μ*-messbar.*

Nach Satz 2.3 muss ein Inhalt, der zu einem äußeren Mass fortsetzbar ist, auf dem Halbring $\mathcal{P}$ notwendig σ-additiv sein. Hierfür wird die folgende Bezeichnung eingeführt.

Definition 3.2 (Prämaß) *Sei $\mathcal{P} \subset 2^X$ ein Halbring. Eine Funktion $\lambda : \mathcal{P} \to [0,\infty]$ mit $\lambda(\emptyset) = 0$ heißt Prämaß, wenn gilt: ist $P \in \mathcal{P}$ und $P = \bigcup_{i=1}^\infty P_i$ mit $P_i \in \mathcal{P}$ paarweise disjunkt, so folgt*

$$\lambda(P) = \sum_{i=1}^\infty \lambda(P_i) \qquad (\sigma\text{-Additivität auf } \mathcal{P}).$$

Lemma 3.1 *Ist $\lambda : \mathcal{P} \to [0,\infty]$ ein Prämaß, $\mathcal{R}$ der von $\mathcal{P}$ erzeugte Ring und $\overline{\lambda} : \mathcal{R} \to [0,\infty]$ der eindeutig bestimmte Inhalt mit $\overline{\lambda}|_{\mathcal{P}} = \lambda$ (siehe Satz 1.4), so ist $\overline{\lambda}$ ebenfalls ein Prämaß.*

Beweis: Sei $F = \bigcup_{i=1}^\infty F_i$ mit $F, F_i \in \mathcal{R}$ und F_i paarweise disjunkt. Nach Satz 1.3 gibt es dann paarweise disjunkte Zerlegungen $F = \bigcup_{j=1}^k P_j$ und $F_i = \bigcup_{\nu=1}^{\nu_i} P_{i,\nu}$ mit P_j, $P_{i,\nu} \in \mathcal{P}$. Es folgt die disjunkte Zerlegung

$$P_j = \bigcup_{i=1}^\infty (P_j \cap F_i) = \bigcup_{i=1}^\infty \bigcup_{\nu=1}^{\nu_i} (P_j \cap P_{i,\nu}).$$

Da λ Prämass, folgt hieraus $\lambda(P_j) = \sum_{i=1}^\infty \sum_{\nu=1}^{\nu_i} \lambda(P_j \cap P_{i,\nu}) = \sum_{i=1}^\infty \overline{\lambda}(P_j \cap F_i)$, und somit

$$\overline{\lambda}(F) = \sum_{j=1}^k \lambda(P_j) = \sum_{i=1}^\infty \sum_{j=1}^k \overline{\lambda}(P_j \cap F_i) = \sum_{i=1}^\infty \overline{\lambda}(F_i).$$

$\square$

Satz 3.1 (Caratheodory-Fortsetzung) *Sei $\lambda : \mathcal{P} \to [0,\infty]$ ein Prämaß auf dem Halbring $\mathcal{P} \subset 2^X$. Definiere $\mu : 2^X \to [0,\infty]$ durch*

$$\mu(E) = \inf\{\sum_{i=1}^\infty \lambda(P_i) : P_i \in \mathcal{P},\, E \subset \bigcup_{i=1}^\infty P_i\}. \tag{3.1}$$

Dann ist μ eine Fortsetzung von λ. Man bezeichnet μ als das durch λ induzierte, äußere Maß oder als Caratheodory-Fortsetzung von λ.

Beweis: Wir nehmen zunächst an, dass $\mathcal{P}$ sogar ein Ring ist und zeigen in diesem Fall die Aussage in drei Schritten.

Schritt 1: *μ ist äußeres Maß*
Mit $P_i = \emptyset$ für alle $i \in \mathbb{N}$ folgt $\mu(\emptyset) = 0$ nach (3.1). Sei $E \subset \bigcup_{i=1}^\infty E_i$ mit $E, E_i \subset X$ und oBdA $\mu(E_i) < \infty$. Wähle Überdeckungen $E_i \subset \bigcup_{j=1}^\infty P_{i,j}$ mit $P_{i,j} \in \mathcal{P}$, so dass zu gegebenem $\varepsilon > 0$ gilt:

$$\sum_{j=1}^\infty \lambda(P_{i,j}) < \mu(E_i) + 2^{-i}\varepsilon \quad \text{für alle } i \in \mathbb{N}.$$

Dann folgt $E \subset \bigcup_{i,j=1}^\infty P_{i,j}$ und somit

$$\mu(E) \leq \sum_{i,j=1}^\infty \lambda(P_{i,j}) \leq \sum_{i=1}^\infty \big(\mu(E_i) + 2^{-i}\varepsilon\big) = \sum_{i=1}^\infty \mu(E_i) + \varepsilon.$$

Mit $\varepsilon \searrow 0$ folgt $\mu(E) \leq \sum_{i=1}^\infty \mu(E_i)$.

Schritt 2: $\mu(P) = \lambda(P)$ für $P \in \mathcal{P}$.
Die Ungleichung $\mu(P) \leq \lambda(P)$ folgt direkt aus (3.1), indem wir $P_1 = P$ und $P_i = \emptyset$ für $i \geq 2$ wählen. Für die umgekehrte Ungleichung $\lambda(P) \leq \mu(P)$ ist zu zeigen:

$$P \subset \bigcup_{i=1}^\infty P_i \text{ mit } P_i \in \mathcal{P} \quad \Rightarrow \quad \lambda(P) \leq \sum_{i=1}^\infty \lambda(P_i).$$

Wir betrachten die paarweise disjunkten Mengen $Q_i = (P_i \backslash \bigcup_{j=1}^{i-1} P_j) \cap P \in \mathcal{P}$ und schließen

$$\lambda(P) = \lambda(\bigcup_{i=1}^\infty Q_i) = \sum_{i=1}^\infty \lambda(Q_i) \leq \sum_{i=1}^\infty \lambda(P_i).$$

Schritt 3: *Jedes $P \in \mathcal{P}$ ist μ-messbar*
Sei $P \in \mathcal{P}$ und $S \subset X$ beliebig. Zu $\varepsilon > 0$ wähle $P_i \in \mathcal{P}$ für $i = 1, 2, \ldots$, so dass $S \subset \bigcup_{i=1}^\infty P_i$ und $\sum_{i=1}^\infty \lambda(P_i) \leq \mu(S) + \varepsilon$. Es folgt $S \cap P \subset \bigcup_{i=1}^\infty (P_i \cap P)$ sowie $S \backslash P \subset \bigcup_{i=1}^\infty (P_i \backslash P)$. Aus (3.1) erhalten wir

$$\mu(S \cap P) + \mu(S \backslash P) \leq \sum_{i=1}^\infty \lambda(P_i \cap P) + \sum_{i=1}^\infty \lambda(P_i \backslash P) = \sum_{i=1}^\infty \lambda(P_i) \leq \mu(S) + \varepsilon.$$

Mit $\varepsilon \searrow 0$ folgt Schritt 3 und der Satz ist bewiesen, wenn $\mathcal{P}$ ein Ring ist.

Sei nun $\mathcal{P}$ lediglich ein Halbring, und $\mathcal{R}$ der von $\mathcal{P}$ erzeugte Ring. Nach Satz 1.4 gibt es einen eindeutig bestimmten Inhalt $\overline{\lambda} : \mathcal{R} \to [0, \infty]$ mit $\overline{\lambda}|_{\mathcal{P}} = \lambda$, und $\overline{\lambda}$ ist ein Prämaß nach Lemma 3.1. Wir zeigen:

$$\text{(3.2)} \qquad \text{Mit } \overline{\mu}(E) = \inf\{\sum_{i=1}^\infty \overline{\lambda}(F_i) : F_i \in \mathcal{R},\, E \subset \bigcup_{i=1}^\infty F_i\} \quad \text{gilt} \quad \overline{\mu} = \mu.$$

Für beliebiges $E \subset X$ ergibt sich die Ungleichung $\overline{\mu}(E) \le \mu(E)$ direkt aus den Definitionen wegen $\overline{\lambda}|_{\mathcal{P}} = \lambda$. Ist andererseits $E \subset \bigcup_{i=1}^{\infty} F_i$ mit $F_i \in \mathcal{R}$, so gibt es nach Satz 1.3 eine Darstellung $F_i = \bigcup_{\nu=1}^{\nu_i} P_{i,\nu}$ mit paarweise disjunkten $P_{i,\nu} \in \mathcal{P}$, und es folgt

$$\mu(E) \le \sum_{i=1}^{\infty} \sum_{\nu=1}^{\nu_i} \lambda(P_{i,\nu}) = \sum_{i=1}^{\infty} \overline{\lambda}(F_i).$$

Durch Bilden des Infimums über alle Überdeckungen $E \subset \bigcup_{i=1}^{\infty} F_i$ mit $F_i \in \mathcal{R}$ folgt die umgekehrte Ungleichung $\mu(E) \le \overline{\mu}(E)$. Nun ist, wie oben gezeigt, $\overline{\mu}$ Fortsetzung von $\overline{\lambda}$ und damit folgt aus (3.2) die Behauptung des Satzes. □

Natürlich schließt sich die Frage der Charakterisierung der μ-messbaren Mengen an. Laut Satz 3.1 sind zunächst alle Mengen in $\mathcal{P}$ μ-messbar, aber dann nach Satz 2.3 auch alle abzählbaren Vereinigungen und Durchschnitte dieser Mengen, und dann wiederum deren abzählbare Vereinigungen und Durchschnitte, und so fort. Um dies umfassend zu beschreiben, ist der folgende Begriff zweckmäßig.

Definition 3.3 (erzeugte σ-Algebra) *Für ein Mengensystem $\mathcal{E} \subset 2^X$ sei*

$$\sigma(\mathcal{E}) = \bigcap \{\mathcal{A} : \mathcal{A} \text{ ist } \sigma\text{-Algebra in } X \text{ mit } \mathcal{E} \subset \mathcal{A}\}. \tag{3.3}$$

Dann ist $\sigma(\mathcal{E})$ eine σ-Algebra, die $\mathcal{E}$ enthält, und heißt die von $\mathcal{E}$ erzeugte σ-Algebra. Es gilt

$$\mathcal{A} \;\; \sigma\text{-Algebra mit } \mathcal{E} \subset \mathcal{A} \quad \Rightarrow \quad \sigma(\mathcal{E}) \subset \mathcal{A}. \tag{3.4}$$

Beispiel 3.1 Ist $E \subset X$ und $\mathcal{E} = \{E\}$, so gilt $\sigma(\mathcal{E}) = \{\emptyset, E, X \backslash E, X\}$.

Während wir für den erzeugten Ring eine einfache konstruktive Beschreibung in Satz 1.3 angeben konnten, ist das für die erzeugte σ-Algebra im Allgemeinen nicht möglich. Die Anwendung der Definition stützt sich deshalb stets auf die Eigenschaft (3.4), dass nämlich $\sigma(\mathcal{E})$ – salopp gesprochen – die kleinste σ-Algebra ist, die $\mathcal{E}$ enthält.

Lemma 3.2 *Sei $\lambda : \mathcal{P} \to [0, \infty]$ ein Prämaß. Ist das äußere Maß $\tilde{\mu}$ eine Fortsetzung von λ, so ist jede Menge $A \in \sigma(\mathcal{P})$ $\tilde{\mu}$-messbar.*

Beweis: Nach Satz 2.3 ist das System $\tilde{\mathcal{M}}$ der $\tilde{\mu}$-messbaren Mengen eine σ-Algebra, die nach Definition $\mathcal{P}$ enthält. Also folgt $\sigma(\mathcal{P}) \subset \tilde{\mathcal{M}}$ aus (3.4). □

Definition 3.4 (Regularität) *Eine äußeres Maß $\mu : 2^X \to [0, \infty]$ heißt regulär, wenn es zu jeder Menge $D \subset X$ eine μ-messbare Menge $E \subset X$ gibt mit $D \subset E$ und $\mu(E) = \mu(D)$.*

Proposition 3.1 (Regularität der Caratheodory-Fortsetzung) *Sei $\mu : 2^X \to [0, \infty]$ die Caratheodory-Fortsetzung des Prämaßes $\lambda : \mathcal{P} \to [0, \infty]$. Dann gibt es zu jeder Menge $D \subset X$ eine Menge $E \in \sigma(\mathcal{P})$ mit $E \supset D$ und $\mu(E) = \mu(D)$. Insbesondere ist μ ein reguläres Maß.*

Beweis: Im Fall $\mu(D) = \infty$ können wir $E = X$ wählen. Ist $\mu(D) < \infty$, so gibt es nach Definition von μ in (3.1) zu jedem $\nu \in \mathbb{N}$ eine Überdeckung

$$D \subset \bigcup_{i=1}^{\infty} P_i^{\nu} =: E^{\nu} \;\text{ mit }\; P_i^{\nu} \in \mathcal{P} \;\text{ und }\; \sum_{\nu=1}^{\infty} \lambda(P_i^{\nu}) < \mu(D) + \frac{1}{\nu}.$$

Wähle $E = \bigcap_{\nu=1}^\infty E^\nu$. Dann ist $E \in \sigma(\mathcal{P})$ mit $E \supset D$, und für jedes $\nu \in \mathbb{N}$ gilt

$$\mu(D) \leq \mu(E) \leq \mu(E^\nu) \leq \sum_{i=1}^\infty \mu(P_i^\nu) = \sum_{i=1}^\infty \lambda(P_i^\nu) \leq \mu(D) + \frac{1}{\nu} < \infty,$$

das heißt $\mu(E) = \mu(D)$. □

Für gewisse Eigenschaften der Caratheodory-Fortsetzung wird folgende Bedingung benötigt.

Definition 3.5 (σ-endliches Prämaß) *Ein Prämaß λ auf einem Halbring $\mathcal{P} \subset 2^X$ heißt σ-endlich, wenn es eine Folge $P_n \in \mathcal{P}$ gibt mit $X = \bigcup_{n=1}^\infty P_n$ und $\lambda(P_n) < \infty$ für alle $n \in \mathbb{N}$.*

Das Zählmaß auf X ist genau dann σ-endlich, wenn X abzählbar ist, und das Prämass λ aus Beispiel 3.2 unten ist nicht σ-endlich.

Lemma 3.3 *Sei λ ein σ-endliches Prämaß auf X mit Caratheodory-Fortsetzung μ. Dann hat jede μ-messbare Menge $D \subset X$ eine Darstellung $D = \bigcup_{n=1}^\infty D_n$ für μ-messbare D_n mit $\mu(D_n) < \infty$ und $D_1 \subset D_2 \subset \ldots$.*

Beweis: Sei $X = \bigcup_{n=1}^\infty P_n$ wie in Definition 3.5. Setze dann $D_n = \bigcup_{k=1}^n D \cap P_n$. □

Wir können nun im σ-endlichen Fall die bezüglich der Caratheodory-Fortsetzung messbaren Mengen genau charakterisieren. Die im Zusatz gegebene Präzisierung werden wir beim Beweis des Cavalierischen Prinzips in Satz 7.1 verwenden.

Satz 3.2 (Messbarkeit bzgl. Caratheodory-Fortsetzung) *Sei $\lambda : \mathcal{P} \to [0,\infty]$ ein σ-endliches Prämaß auf X mit Caratheodory-Fortsetzung μ. Eine Menge $D \subset X$ ist genau dann μ-messbar, wenn eine der folgenden, äquivalenten Bedingungen gilt:*

(i) *Es gibt ein $E \in \sigma(\mathcal{P})$ mit $E \supset D$ und $\mu(E\backslash D) = 0$.*

(ii) *Es gibt ein $C \in \sigma(\mathcal{P})$ mit $C \subset D$ und $\mu(D\backslash C) = 0$.*

Zusatz: *Im Fall $\mu(D) < \infty$ kann $E = \bigcap_{\nu=1}^\infty E^\nu$ gewählt werden, wobei $E^\nu = \bigcup_{i=1}^\infty P_i^\nu$ paarweise disjunkte Vereinigung von Mengen $P_i^\nu \in \mathcal{P}$ mit $E^1 \supset E^2 \supset \ldots$ und $\mu(E^1) < \infty$.*

Beweis: Es ist klar, dass jede der beiden Bedingungen die μ-Messbarkeit impliziert. Für $\mu(D) < \infty$ wählen wir E wie im Beweis von Proposition 3.1. Es folgt

$$\mu(D) = \mu(E) = \mu(E \cap D) + \mu(E\backslash D) = \mu(D) + \mu(E\backslash D) \quad \Rightarrow \quad \mu(E\backslash D) = 0.$$

Für beliebiges messbares D sei $D = \bigcup_{n=1}^\infty D_n$ wie in Lemma 3.3. Wie bewiesen gibt es $E_n \supset D_n$ mit $E_n \in \sigma(\mathcal{P})$ und $\mu(E_n\backslash D_n) = 0$. Für $E = \bigcup_{n=1}^\infty E_n \supset D$ folgt $E \in \sigma(\mathcal{P})$ und

$$\mu(E\backslash D) \leq \mu(\bigcup_{n=1}^\infty E_n\backslash D_n) \leq \sum_{n=1}^\infty \mu(E_n\backslash D_n) = 0.$$

Damit ist Aussage (i) gezeigt. Für (ii) wähle mit Proposition 3.1 eine μ-Nullmenge $N \in \sigma(\mathcal{P})$ mit $N \supset E\backslash D$ für E wie in (i), und setze $C = E\backslash N \in \sigma(\mathcal{P})$. Es folgt

$$C = E\backslash N \subset E\backslash(E\backslash D) = D \quad \text{und} \quad \mu(D\backslash C) = \mu(D \cap N) = 0.$$

Ist $\mu(D) < \infty$, so liefert die Konstruktion im Beweis von Proposition 3.1 die Darstellung $E = \bigcap_{\nu=1}^{\infty} E^\nu$ mit $E^\nu = \bigcup_{i=1}^{\infty} P_i^\nu$ für $P_i^\nu \in \mathcal{P}$, wobei $\mu(E^1) \leq \mu(D) + 1 < \infty$. Das System aller abzählbaren Vereinigungen von Mengen in $\mathcal{P}$ ist abgeschlossen unter endlichen Durchschnitten, denn es gilt $(\bigcup_{i=1}^{\infty} P_i) \cap (\bigcup_{i=1}^{\infty} Q_i) = \bigcup_{i=1}^{\infty} P_i \cap Q_i$. Indem wir zu $\tilde{E}^\nu = E^1 \cap \ldots \cap E^\nu$ übergehen, können wir deshalb $E^1 \supset E^2 \supset \ldots$ annehmen. Zusätzlich können wir dann nach Satz 1.3 voraussetzen, dass für festes ν die Mengen P_i^ν, $i \in \mathbb{N}$, paarweise disjunkt sind. Damit ist auch der Zusatz bewiesen. □

Nachdem die Maßfortsetzung konstruiert ist und auch die messbaren Mengen identifiziert sind, stellt sich zu guter letzt die Frage, inwieweit das gefundene Maß eindeutig bestimmt ist bzw. durch welche Eigenschaften die gegebene Konstruktion ausgezeichnet ist. Direkt aus der Definition ergibt sich die folgende Charakterisierung:

Lemma 3.4 (Maximalität der Caratheodory-Fortsetzung) *Sei μ die Caratheodory-Fortsetzung des Prämaßes $\lambda : \mathcal{P} \to [0, \infty]$ auf X. Dann gilt für jede andere Fortsetzung $\tilde{\mu}$ von λ*

$$\tilde{\mu}(E) \leq \mu(E) \quad \text{für alle } E \subset X.$$

Beweis: Es gilt die Implikation

$$E \subset \bigcup_{i=1}^{\infty} P_i \text{ mit } P_i \in \mathcal{P} \quad \Rightarrow \quad \tilde{\mu}(E) \leq \sum_{i=1}^{\infty} \tilde{\mu}(P_i) = \sum_{i=1}^{\infty} \lambda(P_i).$$

Bilden des Infimums über alle solchen Überdeckungen liefert wegen (3.1) die Behauptung. □

Im nicht σ-endlichen Fall kann es Fortsetzungen geben, die auf der erzeugten σ-Algebra verschieden sind. Hier ist ein Standardbeispiel.

Beispiel 3.2 Betrachte auf $X \neq \emptyset$ das Prämaß $\lambda : \mathcal{P} = \{\emptyset\} \to [0, \infty]$, $\lambda(\emptyset) = 0$. Offenbar ist λ nicht σ-endlich. Definiere nun für $t \in [0, \infty]$, vgl. Beispiel 2.4,

$$\mu_t : 2^X \to [0, \infty],\ \mu_t(E) = \begin{cases} 0 & \text{falls } E = \emptyset \\ t & \text{sonst.} \end{cases}$$

Dann sind alle μ_t Fortsetzungen von λ, die jedoch auf der von $\mathcal{P}$ erzeugten σ-Algebra $\sigma(\mathcal{P}) = \{\emptyset, X\}$ nicht gleich sind. Die Caratheodory-Fortsetzung ist μ_∞.

Satz 3.3 (Eindeutigkeit der Maßfortsetzung) *Sei λ ein σ-endliches Prämaß auf einem Halbring $\mathcal{P}$ über X. Ist $\mu : 2^X \to [0, \infty]$ die Caratheodory-Fortsetzung von λ und $\tilde{\mu}$ eine beliebige Fortsetzung, so gilt*

$$E \ \mu\text{-messbar} \quad \Rightarrow \quad E \ \tilde{\mu}\text{-messbar und } \tilde{\mu}(E) = \mu(E).$$

Beweis: Seien $\mathcal{M}$ bzw. $\tilde{\mathcal{M}}$ die jeweiligen Systeme messbarer Mengen für μ bzw. $\tilde{\mu}$. Nach Satz 3.2 (ii) gibt es zu $E \in \mathcal{M}$ ein $C \in \sigma(\mathcal{P})$ und eine μ-Nullmenge N mit $E = C \cup N$. Da $\tilde{\mu}(N) \leq \mu(N) = 0$ nach Lemma 3.4, ist N $\tilde{\mu}$-messbar nach Proposition 2.1. Aber C ist ebenfalls $\tilde{\mu}$-messbar nach Lemma 3.2, also ist E $\tilde{\mu}$-messbar.

Wir zeigen $\tilde{\mu}(E) = \mu(E)$ erst unter der zusätzlichen Annahme, dass $E \subset P$ für ein $P \in \mathcal{P}$ mit $\lambda(P) < \infty$. Und zwar gilt dann nach Lemma 3.4

$$\begin{aligned} \tilde{\mu}(E) + \tilde{\mu}(P \backslash E) & \leq \mu(E) + \mu(P \backslash E) \\ & = \mu(P) \quad (\text{da } E \ \mu\text{-messbar}) \\ & = \tilde{\mu}(P) \quad (\text{da } \mu|_{\mathcal{P}} = \tilde{\mu}|_{\mathcal{P}} = \lambda) \\ & \leq \tilde{\mu}(E) + \tilde{\mu}(P \backslash E). \end{aligned}$$

Die Behauptung folgt, da in allen Abschätzungen Gleichheit gelten muss. Sei nun $E \in \mathcal{M}$ beliebig. Nach Lemma 3.3 gilt $E = \bigcup_{n=1}^{\infty} E_n$ für μ-messbare E_n mit $\mu(E_n) < \infty$ und $E_1 \subset E_2 \subset \ldots$. Es folgt wegen Satz 2.4 (i)

$$\tilde{\mu}(E) = \lim_{n \to \infty} \tilde{\mu}(E_n) = \lim_{n \to \infty} \mu(E_n) = \mu(E).$$

□

Zum Schluss des Kapitels wollen wir die Beziehung zwischen äußeren Maßen und Maßen aufklären. Aus einem äußeren Maß μ erhalten wir das Maß $\lambda = \mu|_{\mathcal{M}(\mu)}$ durch Einschränkung auf die μ-messbaren Mengen, und λ ist nach Proposition 2.1 *vollständig*, das heißt es gilt

$$N \subset A \text{ für ein } A \in \mathcal{A} \text{ mit } \lambda(A) = 0 \quad \Rightarrow \quad N \in \mathcal{A}. \tag{3.5}$$

Umgekehrt können wir ein gegebenes Maß λ auf einer σ-Algebra $\mathcal{A} \subset 2^X$ mit Satz 3.1 zu einem äußeren Maß λ^C auf X fortsetzen, und nach Proposition 3.1 ist λ^C regulär. Wir wollen jetzt sehen, dass diese Zuordnungen invers sind, jedenfalls im σ-endlichen Fall.

Definition 3.6 *Ein äußeres Maß $\mu : 2^X \to [0, \infty]$ heißt σ-endlich, wenn X abzählbare Vereingung $X = \bigcup_{i=1}^{\infty} X_i$ von μ-messbaren Mengen X_i mit $\mu(X_i) < \infty$ ist.*

Nach Definition 3.5 ist μ genau dann σ-endlich, wenn das Maß $\mu|_{\mathcal{M}(\mu)}$ σ-endlich ist.

Satz 3.4 (äußere Maße vs. Maße) *Die durch Einschränkung bzw. Fortsetzung gegebenen Abbildungen zwischen den σ-endlichen, regulären äußeren Maßen und den σ-endlichen, vollständigen Maßen auf X sind zueinander invers und insbesondere bijektiv.*

Beweis: Wir zeigen als erstes $\lambda = \mu|_{\mathcal{M}(\mu)}$ für $\mu = \lambda^C$. Nach Satz 3.2 ist $D \subset X$ genau dann λ^C-messbar, wenn es $C, E \in \mathcal{A}$ gibt mit $C \subset D \subset E$ und $\lambda^C(E \backslash D) = \lambda^C(D \backslash C) = 0$. Es folgt $E \backslash C \in \mathcal{A}$ und nach Satz 3.1

$$\lambda(E \backslash C) = \lambda^C(E \backslash C) \leq \lambda^C(E \backslash D) + \lambda^C(D \backslash C) = 0.$$

Da $\mathcal{A}$ vollständig nach Voraussetzung, ist $D \backslash C \in \mathcal{A}$ und damit $D = C \cup (D \backslash C) \in \mathcal{A}$, das heißt es gilt $\mathcal{M}(\lambda^C) = \mathcal{A}$ und $\lambda^C = \lambda$ auf $\mathcal{A}$ nach Satz 3.1.

Jetzt zeigen wir $\mu = \lambda^C$ für $\lambda = \mu|_{\mathcal{M}(\mu)}$. Es gilt $\mathcal{M}(\mu) = \mathcal{M}(\lambda^C)$ nach Satz 3.1 und Satz 3.3. Aber μ ist regulär nach Voraussetzung, sowie λ^C regulär nach Proposition 3.1, und beide äußere Maße stimmen überein auf $\mathcal{M}(\mu) = \mathcal{M}(\lambda^C)$, also sind die Maße gleich. □

4 Das n-dimensionale Lebesguemaß

Wir spezialisieren nun die allgemeinen Ausführungen des vorigen Kapitels und betrachten das Lebesguemaß auf dem n-dimensionalen Euklidischen Raum. Als erstes wird untersucht, in welcher Beziehung das System der Lebesgue-messbaren Mengen zu der gegebenen Topologie auf $\mathbb{R}^n$, also zu dem System der offenen Mengen, steht. Insbesondere ergibt sich daraus eine Charakterisierung der Lebesgue-messbaren Mengen. Zweitens wird die Invarianz des Lebesguemaßes unter Isometrien, das heißt Translationen und orthogonalen Transformationen, bewiesen, sowie allgemeiner eine Transformationsformel unter linearen Abbildungen. Zum Schluss geben wir ein Beispiel einer nicht Lebesgue-messbaren Menge in $\mathbb{R}$.

Lemma 4.1 *Auf dem Halbring $\mathcal{P}^n$ der (achsenparallelen) Quader im $\mathbb{R}^n$ ist der elementargeometrische Inhalt $\lambda^n : \mathcal{P}^n \to [0, \infty)$ ein Prämaß.*

Beweis: Sei $P = \bigcup_{i=1}^{\infty} P_i$ mit $P, P_i \in \mathcal{P}^n$ und $P_i \cap P_j = \emptyset$ für $i \neq j$. Nun ist λ^n ein Inhalt auf dem Ring der Figuren, siehe Satz 1.4, also gilt aufgrund der Monotonie, Folgerung 1.2,

$$\sum_{i=1}^{\infty} \lambda^n(P_i) = \lim_{k\to\infty} \sum_{i=1}^{k} \lambda^n(P_i) = \lim_{k\to\infty} \lambda^n\big(\bigcup_{i=1}^{k} P_i\big) \leq \lambda^n(P).$$

Für die umgekehrte Ungleichung wähle zu $\varepsilon > 0$ offene Quader $Q_i \supset P_i$ und einen kompakten Quader $Q \subset P$, so dass gilt

$$\sum_{i=1}^{\infty} \lambda^n(Q_i) < \sum_{i=1}^{\infty} \lambda^n(P_i) + \frac{\varepsilon}{2}, \quad \lambda^n(P) < \lambda^n(Q) + \frac{\varepsilon}{2}.$$

Nach Heine-Borel wird Q durch endliche viele Quader $Q_1, \ldots, Q_k$ überdeckt, und mit Folgerung 1.2 schließen wir

$$\lambda^n(P) < \lambda^n(Q) + \frac{\varepsilon}{2} \leq \sum_{i=1}^{k} \lambda^n(Q_i) + \frac{\varepsilon}{2} < \sum_{i=1}^{\infty} \lambda^n(P_i) + \varepsilon.$$

Daraus folgt die Behauptung. $\square$

Definition 4.1 (Lebesguemaß) *Das Lebesguemaß einer Menge $E \subset \mathbb{R}^n$ ist definiert durch*

$$\mathcal{L}^n(E) = \inf\big\{\sum_{i=1}^{\infty} \lambda^n(P_i) : P_i \text{ Quader, } E \subset \bigcup_{i=1}^{\infty} P_i\big\}. \tag{4.1}$$

Damit ist $\mathcal{L}^n$ die Caratheodory-Fortsetzung des auf $\mathcal{P}^n$ definierten Elementarinhalts λ^n.

Definition 4.2 (Borelmenge) *Die vom System $\mathcal{O}^n$ der offenen Teilmengen des $\mathbb{R}^n$ erzeugte σ-Algebra heißt Borelalgebra $\mathcal{B}^n$, ihre Elemente heißen Borelmengen.*

Lemma 4.2 *$\mathcal{B}^n$ ist die vom Halbring $\mathcal{P}^n$ der Quader erzeugte σ-Algebra.*

Beweis: Wir zeigen als erstes, dass jeder Quader eine Borelmenge ist. Ein Intervall $I \subset \mathbb{R}$ ist entweder offen oder läßt sich als abzählbarer Schnitt $I = \bigcap_{k=1}^{\infty} U_k$ von offenen Intervallen U_k schreiben und liegt damit in $\mathcal{B}^1$, zum Beispiel gilt $[a,b) = \bigcap_{k=1}^{\infty} (a - \frac{1}{k}, b)$. Für einen Quader $Q = I_1 \times \ldots \times I_n \subset \mathbb{R}^n$ schreiben wir $I_j = \bigcap_{k=1}^{\infty} U_{j,k}$ und erhalten $Q = \bigcap_{k=1}^{\infty} (U_{1,k} \times \ldots \times U_{n,k}) \in \mathcal{B}^n$. Daraus folgt $\mathcal{F}^n \subset \mathcal{B}^n$ und folglich $\sigma(\mathcal{F}^n) \subset \mathcal{B}^n$.

Nun ist andererseits nach Lemma 1.2 jede offene Menge $\Omega \subset \mathbb{R}^n$ als Vereinigung von abzählbar vielen kompakten Würfeln darstellbar. Also gilt $\mathcal{O}^n \subset \sigma(\mathcal{F}^n)$ und somit $\mathcal{B}^n \subset \sigma(\mathcal{F}^n)$. $\square$

Satz 4.1 ($\mathcal{L}^n$-Messbarkeit der Borelmengen) *Für das Lebesguemaß auf $\mathbb{R}^n$ gilt:*

(i) *Alle Borelmengen sind Lebesgue-messbar.*

(ii) *Zu $E \subset \mathbb{R}^n$ gibt es eine Borelmenge $B \supset E$ mit $\mathcal{L}^n(B) = \mathcal{L}^n(E)$.*

(iii) *$\mathcal{L}^n(K) < \infty$ für alle $K \subset \mathbb{R}^n$ kompakt.*

Beweis: Sei $\mathcal{M}^n$ das System der $\mathcal{L}^n$-messbaren Mengen. Dann gilt $\mathcal{P}^n \subset \mathcal{M}^n$ nach Satz 3.1, also $\mathcal{B}^n = \sigma(\mathcal{P}^n) \subset \mathcal{M}^n$ nach Lemma 4.2 und Satz 2.3.

Aussage (ii) gilt nach Proposition 3.1. Da $\mathcal{L}^n = \lambda^n$ auf Quadern, gilt für beliebiges $a > 0$ $\mathcal{L}^n([-a,a]^n) = \lambda^n([-a,a]^n) = (2a)^n < \infty$, und (iii) folgt. $\square$

Lemma 4.3 (Approximationslemma) *Für eine beliebige Menge $E \subset \mathbb{R}^n$ gilt:*

(i) $\mathcal{L}^n(E) = \inf\{\mathcal{L}^n(U) : U$ *offen,* $U \supset E\}$,

(ii) $\mathcal{L}^n(E) = \sup\{\mathcal{L}^n(K) : K$ *kompakt,* $K \subset E\}$, *falls E $\mathcal{L}^n$-messbar.*

Beweis: Offensichtlich gilt $\mathcal{L}^n(E) \leq \inf\{\mathcal{L}^n(U) : U \text{ offen}, U \supset E\}$. Für die umgekehrte Ungleichung können wir $\mathcal{L}^n(E) < \infty$ annehmen. Nach Definition des Lebesguemaßes in (4.1) gibt es zu $\varepsilon > 0$ eine Überdeckung $E \subset \bigcup_{i=1}^{\infty} P_i$ mit Quadern P_i, so dass gilt:

$$\sum_{i=1}^{\infty} \lambda^n(P_i) < \mathcal{L}^n(E) + \varepsilon.$$

Wir können annehmen, dass die P_i offen sind. Also ist $U = \bigcup_{i=1}^{\infty} P_i$ offen, es gilt $U \supset E$ und

$$\mathcal{L}^n(U) \leq \sum_{i=1}^{\infty} \lambda^n(P_i) < \mathcal{L}^n(E) + \varepsilon.$$

In (ii) ist klar, dass $\mathcal{L}^n(E) \geq \sup\{\mathcal{L}^n(K) : K \text{ kompakt}, K \subset E\}$. Wir zeigen die umgekehrte Ungleichung zunächst für E beschränkt. Wähle $K_0 \subset \mathbb{R}^n$ kompakt mit $E \subset K_0$. Nach (i) gibt es zu $\varepsilon > 0$ eine offene Menge $U \supset K_0 \backslash E$ mit

$$\mathcal{L}^n(U) < \mathcal{L}^n(K_0 \backslash E) + \varepsilon = \mathcal{L}^n(K_0) - \mathcal{L}^n(E) + \varepsilon.$$

Dabei wurde benutzt, dass E $\mathcal{L}^n$-messbar ist. Nun ist $K := K_0 \backslash U \subset K_0 \backslash (K_0 \backslash E) = E$ kompakt, und wegen der Subadditivität folgt

$$\mathcal{L}^n(K) \geq \mathcal{L}^n(K_0) - \mathcal{L}^n(U) > \mathcal{L}^n(E) - \varepsilon.$$

Für E beliebig betrachte $E_j = E \cap \{x \in \mathbb{R}^n : |x| \leq j\}$. E_j ist beschränkt und Lebesgue-messbar, also folgt aus obigem

$$\mathcal{L}^n(E_j) \leq \sup\{\mathcal{L}^n(K) : K \text{ kompakt}, K \subset E_j\} \leq \sup\{\mathcal{L}^n(K) : K \text{ kompakt}, K \subset E\}.$$

Aber $\mathcal{L}^n(E_j) \to \mathcal{L}^n(E)$ mit $j \to \infty$ nach Satz 2.4. Damit ist (ii) bewiesen. □

Die $\mathcal{L}^n$-messbaren Mengen können nun wie folgt charakterisiert werden.

Satz 4.2 (Messbarkeit bzgl. $\mathcal{L}^n$) *Eine Menge $D \subset \mathbb{R}^n$ ist genau dann $\mathcal{L}^n$-messbar, wenn eine der beiden (äquivalenten) Bedingungen gilt:*

(i) *Es gibt eine Borelmenge $E \supset D$ mit $\mathcal{L}^n(E \backslash D) = 0$.*

(ii) *Es gibt eine Borelmenge $C \subset D$ mit $\mathcal{L}^n(D \backslash C) = 0$.*

Es kann $E = \bigcap_{i=1}^{\infty} U_i$ mit U_i offen, $C = \bigcup_{j=1}^{\infty} A_i$ mit A_i abgeschlossen gewählt werden.

Beweis: Die Äquivalenz von (i) bzw. (ii) mit der Messbarkeit von D wurde bereits in Satz 3.2 bewiesen. Wir führen die Konstruktion von E bzw. C aber explizit durch, um die Zusatzaussage zu zeigen. Schreibe $D = \bigcup_{j=1}^{\infty} D_j$ mit $D_j = \{x \in D : j-1 \leq |x| < j\}$ für $j \in \mathbb{N}$. Nach Lemma 4.3 gibt es offene Mengen $U_{i,j}$ und kompakte Mengen $K_{i,j}$ mit $U_{i,j} \supset D_j \supset K_{i,j}$ und

$$\mathcal{L}^n(U_{i,j}) < \mathcal{L}^n(D_j) + 2^{-j}/i, \quad \mathcal{L}^n(K_{i,j}) > \mathcal{L}^n(D_j) - 2^{-j}/i.$$

Dann ist $U_i = \bigcup_{j=1}^{\infty} U_{i,j}$ offen, $A_i = \bigcup_{j=1}^{\infty} K_{i,j}$ abgeschlossen(!) und es gilt $U_i \supset D \supset A_i$. Mit $E = \bigcap_{i=1}^{\infty} U_i$ bzw. $C = \bigcup_{i=1}^{\infty} A_i$ gelten für beliebiges $i \in \mathbb{N}$ die Abschätzungen

$$\mathcal{L}^n(E \backslash D) \leq \mathcal{L}^n(U_i \backslash D) \leq \sum_{j=1}^{\infty} \mathcal{L}^n(U_{i,j} \backslash D_j) \leq \sum_{j=1}^{\infty} (\mathcal{L}^n(U_{i,j}) - \mathcal{L}^n(D_j)) \leq \frac{1}{i},$$

$$\mathcal{L}^n(D \backslash C) \leq \mathcal{L}^n(D \backslash A_i) \leq \sum_{j=1}^{\infty} \mathcal{L}^n(D_j \backslash K_{i,j}) \leq \sum_{j=1}^{\infty} (\mathcal{L}^n(D_j) - \mathcal{L}^n(K_{i,j})) \leq \frac{1}{i}.$$

Dabei wurde die Messbarkeit von D_j benutzt. Mit $i \to \infty$ folgen die Behauptungen. □

Satz 4.3 (Lebesguemaß vs. Jordaninhalt)

(i) *Für $E \subset \mathbb{R}^n$ gilt $\underline{\mathrm{vol}}^n(E) \leq \mathcal{L}^n(E) \leq \overline{\mathrm{vol}}^n(E)$.*

(ii) *Ist E quadrierbar, so ist E auch $\mathcal{L}^n$-messbar und es gilt $\mathcal{L}^n(E) = \mathrm{vol}^n(E)$.*

Beweis: Nach dem Maßfortsetzungssatz stimmen der Elementarinhalt λ^n und das Lebesguemaß auf Quadern überein, also wegen Satz 1.3 auch auf allen Figuren. Für $F_1, F_2 \in \mathcal{F}^n$ mit $F_1 \subset E \subset F_2$ folgt

$$\lambda^n(F_1) = \mathcal{L}^n(F_1) \leq \mathcal{L}^n(E) \leq \mathcal{L}^n(F_2) = \lambda^n(F_2).$$

Aussage (i) folgt, indem wir das Supremum über alle $F_1 \subset E$ bzw. das Infimum über alle $F_2 \supset E$ bilden. Ist nun E quadrierbar, so gilt nach Satz 1.6 $\overline{\mathrm{vol}}^n(\partial E) = 0$, also $\mathcal{L}^n(\partial E) = 0$. E ist also Vereinigung der offenen Menge $\mathrm{int}(E)$ und der $\mathcal{L}^n$-Nullmenge $E \cap \partial E$, und damit $\mathcal{L}^n$-messbar. □

Wir wollen jetzt untersuchen, wie sich das Lebesguemaß unter affinlinearen Abbildungen transformiert. Dafür ist der folgende Begriff nützlich.

Definition 4.3 (Borelmaß) *Ein äußeres Maß μ auf $\mathbb{R}^n$ heißt Borelmaß, falls gilt:*

(i) *Alle Borelmengen sind μ-messbar.*

(ii) *$\mu(K) < \infty$ für jede kompakte Menge $K \subset \mathbb{R}^n$.*

Beispiel 4.1 Das Lebesguemaß $\mathcal{L}^n$ ist ein Borelmaß nach Satz 4.1. Mit Satz 2.2 ist dann auch $\mathcal{L}^n \llcorner E$ ein Borelmaß, für jede Menge $E \subset \mathbb{R}^n$.

Ein Maß μ auf $\mathbb{R}^n$ heißt translationsinvariant, wenn mit $E + a = \{x + a : x \in E\}$ gilt:

$$\mu(E + a) = \mu(E) \qquad \text{für alle } a \in \mathbb{R}^n,\ E \subset \mathbb{R}^n.$$

Aus der Translationsinvarianz des Elementarinhalts $\lambda^n : \mathcal{P}^n \to [0, \infty)$ und der Definition des Lebesguemaßes folgt sofort, das $\mathcal{L}^n$ ein translationsinvariantes Maß ist. In Satz 4.4 unten wird gezeigt, dass das Borelmaß $\mathcal{L}^n$ durch die Eigenschaft der Translationsinvarianz bis auf Normierung eindeutig charakterisiert ist.

Lemma 4.4 *Ist μ translationsinvariantes Borelmaß auf $\mathbb{R}^n$, so ist jede Koordinatenhyperebene $H = \{x \in \mathbb{R}^n : x_i = c\}$ eine μ-Nullmenge.*

Beweis: Sei $Q = [0,1]^n$ und $F = \{x \in Q : x_i = 0\}$. Für $a \in \mathbb{R}^n$ ist $F + a$ abgeschlossen, also μ-messbar. Es folgt für jede endliche Menge $\{s_1, \ldots, s_k\} \subset [0,1]$

$$k\,\mu(F) = \sum_{j=1}^{k} \mu(s_j e_i + F) = \mu\left(\bigcup_{j=1}^{k} s_j e_i + F\right) \leq \mu(Q) < \infty.$$

Da k beliebig groß gewählt werden kann, ist $\mu(F) = 0$. Aber H ist Vereinigung von abzählbar vielen Translationen von F, und somit $\mu(H) = 0$. $\square$

Satz 4.4 (Charakterisierung durch Translationsinvarianz) *Sei μ ein translationsinvariantes Borelmaß auf $\mathbb{R}^n$. Dann gilt*

$$\mu(E) = \theta\,\mathcal{L}^n(E) \text{ für alle } \mathcal{L}^n\text{-messbaren } E \subset \mathbb{R}^n, \text{ wobei } \theta = \mu([0,1]^n).$$

Beweis: Setze $Q_{k,j} = 2^{-k}(j + [0,1]^n)$ für $k \in \mathbb{N}_0$ und $j \in \mathbb{Z}^n$. Dann ist $[0,1]^n$ Vereinigung der 2^{nk} abgeschlossenen Teilwürfel $\{Q_{k,j} : j \in J_k\}$, wobei $J_k = \{j = (j_1, \ldots, j_n) \in \mathbb{Z}^n : 0 \leq j_i \leq 2^k - 1\}$, mit paarweise disjunktem Inneren. Aus Lemma 4.4 folgt

$$\mu([0,1])^n = \sum_{j \in J_k} \mu(Q_{k,j}), \quad \mathcal{L}^n([0,1]^n) = \sum_{j \in J_k} \mathcal{L}^n(Q_{k,j}).$$

Die Translationsinvarianz impliziert $\mu(Q_{k,j}) = \mu(Q_{k,0})$ und $\mathcal{L}^n(Q_{k,j}) = \mathcal{L}^n(Q_{k,0})$ für alle $j \in \mathbb{Z}^n$, also

$$\theta = \frac{\mu([0,1]^n)}{\mathcal{L}^n([0,1]^n)} = \frac{\mu(Q_{k,0})}{\mathcal{L}^n(Q_{k,0})} = \frac{\mu(Q_{k,j})}{\mathcal{L}^n(Q_{k,j})} \quad \text{für alle } j \in \mathbb{Z}^n.$$

Daraus folgt mit Lemma 1.2 wobei wieder Lemma 4.4 benutzt wird,

$$\mu(U) = \theta\,\mathcal{L}^n(U) \quad \text{für alle offenen } U \subset \mathbb{R}^n.$$

Insbesondere gilt die Behauptung des Satzes für alle Quader, und damit für alle $\mathcal{L}^n$-messbaren Mengen aufgrund der Eindeutigkeit der Maßfortsetzung, siehe Satz 3.3. $\square$

Wir zeigen als nächstes die Messbarkeit von Bildmengen, wobei wir uns im Hinblick auf die spätere Anwendung im Transformationssatz für Diffeomorphismen nicht auf lineare Abbildungen beschränken.

Lemma 4.5 *Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^n$ Lipschitzstetig, mit Konstante Λ bzgl. der Maximumsnorm $\|\cdot\|$. Dann gilt*

$$\mathcal{L}^n(f(E)) \leq \Lambda^n \, \mathcal{L}^n(E) \quad \textit{für alle } E \subset U.$$

Beweis: Wir können $\mathcal{L}^n(E) < \infty$ annehmen. Setze

$$Q(x_0, \varrho) = \{x \in \mathbb{R}^n : \|x - x_0\| < \varrho\} \quad \text{für } x_0 \in \mathbb{R}^n,\ \varrho > 0.$$

Nach Voraussetzung gilt $\|f(x) - f(x_0)\| \leq \Lambda \, \|x - x_0\|$ für $x, x_0 \in U$, also

$$Q = Q(x_0, \varrho) \subset U \quad \Rightarrow \quad f(Q) \subset Q(f(x_0), \Lambda\varrho).$$

Nach Lemma 4.3 gibt es nun eine offene Menge $V \supset E$ mit $\mathcal{L}^n(V) < \mathcal{L}^n(E) + \varepsilon$, wobei oBdA $V \subset U$, und weiter eine Ausschöpfung $V = \bigcup_{j=1}^{\infty} Q_j$ durch Würfel Q_j mit paarweise disjunktem Inneren, siehe Lemma 1.2. Damit folgt

$$\mathcal{L}^n(f(E)) \leq \mathcal{L}^n(f(V)) \leq \sum_{j=1}^{\infty} \mathcal{L}^n(f(Q_j)) \leq \Lambda^n \sum_{j=1}^{\infty} \mathcal{L}^n(Q_j) \leq \Lambda^n (\mathcal{L}^n(E) + \varepsilon).$$

Mit $\varepsilon \to 0$ folgt die Behauptung. $\square$

Satz 4.5 ($\mathcal{L}^n$-Messbarkeit von Bildmengen) *Ist $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^n$ lokal Lipschitzstetig, zum Beispiel $f \in C^1(U, \mathbb{R}^n)$, so gilt:*

(i) *$N \subset U$ Nullmenge $\quad \Rightarrow \quad f(N)$ Nullmenge.*

(ii) *$E \subset U$ $\mathcal{L}^n$-messbar $\quad \Rightarrow \quad f(E)$ $\mathcal{L}^n$-messbar.*

Beweis: f ist auf kompakten Teilmengen von U Lipschitzstetig, deshalb folgt Aussage (i) direkt aus Lemma 4.5. Für (ii) können wir annehmen, dass E beschränkt ist, andernfalls betrachten wir $E_j = \{x \in E : |x| \leq j\}$. Nach Satz 4.2 gibt es dann kompakte Mengen K_j und eine $\mathcal{L}^n$-Nullmenge N mit $E = \bigcup_{j=1}^{\infty} K_j \cup N$. Da $f(K_j)$ kompakt und $\mathcal{L}^n(f(N)) = 0$ nach Lemma 4.5, ist $f(E)$ $\mathcal{L}^n$-messbar. $\square$

Satz 4.6 (Bewegungsinvarianz von $\mathcal{L}^n$) *Für $S \in \mathbb{O}(\mathbb{R}^n)$ und $a \in \mathbb{R}^n$ gilt*

$$\mathcal{L}^n(S(E) + a) = \mathcal{L}^n(E) \textit{ für alle } E \subset \mathbb{R}^n.$$

Beweis: Die Translationsinvarianz von $\mathcal{L}^n$ ist schon bekannt, also können wir $a = 0$ annehmen. Wir setzen zunächst nur $S \in GL_n(\mathbb{R})$ voraus und betrachten mit $T = S^{-1}$ das Bildmaß

$$\mu = T(\mathcal{L}^n) : 2^{\mathbb{R}^n} \to [0, \infty],\ \mu(E) = \mathcal{L}^n(T^{-1}(E)) \quad \Rightarrow \quad \mathcal{L}^n(S(E)) = \mu(E).$$

Wir behaupten, dass μ ein translationsinvariantes Borelmaß ist. Ist $B \subset \mathbb{R}^n$ Borelmenge und damit $\mathcal{L}^n$-messbar nach Satz 4.1, so ist $T^{-1}(B) = S(B)$ ebenfalls $\mathcal{L}^n$-messbar wegen Satz 4.5 und damit B μ-messbar nach Satz 2.1. Für $K \subset \mathbb{R}^n$ kompakt ist auch $T^{-1}(K) = S(K)$ kompakt, also $\mu(K) < \infty$. Damit ist gezeigt, dass μ ein Borelmaß ist. Für die Translationsinvarianz berechnen wir für $b \in \mathbb{R}^n$ und $E \subset \mathbb{R}^n$ beliebig

$$\mu(E+b) = \mathcal{L}^n(S(E+b)) = \mathcal{L}^n(S(E)+S(b)) = \mathcal{L}^n(S(E)) = \mu(E).$$

Aus Satz 4.4 folgt nun $\mu(E) = \theta(S)\,\mathcal{L}^n(E)$ für alle $\mathcal{L}^n$-messbaren $E \subset \mathbb{R}^n$, wobei

$$\theta(S) = \mu([0,1]^n) = \mathcal{L}^n(S([0,1]^n)) \in [0,\infty).$$

Für nicht notwendig messbares $E \subset \mathbb{R}^n$ schließen wir mit Lemma 4.3

$$\begin{aligned}
\mu(E) &= \mathcal{L}^n(S(E)) \\
&= \inf\{\mathcal{L}^n(V) : V \supset S(E) \text{ offen}\} \\
&= \inf\{\mathcal{L}^n(S(U)) : U \supset E \text{ offen}\} \\
&= \theta(S)\inf\{\mathcal{L}^n(U) : U \supset E \text{ offen}\}.
\end{aligned}$$

Daraus folgt, wieder mit Lemma 4.3,

$$\mathcal{L}^n(S(E)) = \theta(S)\,\mathcal{L}^n(E) \quad \text{für alle } E \subset \mathbb{R}^n. \tag{4.2}$$

Ist nun sogar $S \in \mathbb{O}(\mathbb{R}^n)$, so setzen wir in (4.2) als Testmenge $E = B_1(0) = \{x \in \mathbb{R}^n : |x| < 1\}$ ein und erhalten

$$\theta(S)\,\mathcal{L}^n(B_1(0)) = \mathcal{L}^n(S(B_1(0))) = \mathcal{L}^n(B_1(0)).$$

Wegen $\mathcal{L}^n(B_1(0)) \in (0,\infty)$ folgt $\theta(S) = 1$ für $S \in \mathbb{O}(\mathbb{R}^n)$, und der Satz ist bewiesen. □

Der Elementarinhalt λ^n ist nur für achsenparallele Quader bzw. Figuren definiert worden. Deshalb ist aus der Definition 4.1 von $\mathcal{L}^n$ nicht unmittelbar ersichtlich, dass das Lebesguemaß unabhängig von der Wahl des Euklidischen Koordinatensystems ist, sondern dies folgt erst aus dem vorangegangenen Satz 4.6. Für die Transformationsformel unter beliebigen linearen Abbildungen $S \in \mathbb{R}^{n\times n}$ benötigen wir die folgende Hilfsaussage aus der Linearen Algebra.

Lemma 4.6 (Polarzerlegung) *Zu jedem $S \in GL_n(\mathbb{R})$ gibt es eine Diagonalmatrix Λ mit Einträgen $\lambda_i > 0$ und $T_1, T_2 \in \mathbb{O}(\mathbb{R}^n)$, so dass $S = T_1\Lambda T_2$.*

Beweis: Die Matrix $S^{\mathrm{T}}S$ ist symmetrisch und hat positive Eigenwerte, denn für $v \in \mathbb{R}^n\backslash\{0\}$ gilt $\langle S^{\mathrm{T}}Sv, v\rangle = |Sv|^2 > 0$. Also gibt es ein $T \in \mathbb{O}(\mathbb{R}^n)$ und eine Diagonalmatrix Λ mit Einträgen $\lambda_1,\ldots,\lambda_n > 0$, so dass gilt:

$$S^{\mathrm{T}}S = T\Lambda^2T^{-1}.$$

$R = T\Lambda T^{-1}$ ist dann symmetrisch mit $R^2 = S^{\mathrm{T}}S$, und $Q = SR^{-1}$ ist orthogonal wegen

$$Q^{\mathrm{T}}Q = (R^{-1})^{\mathrm{T}}S^{\mathrm{T}}SR^{-1} = R^{-1}R^2R^{-1} = \mathrm{E}_n.$$

Es folgt $S = QR = QT\Lambda T^{-1} = T_1\Lambda T_2$ für $T_1 = QT, T_2 = T^{-1} \in \mathbb{O}(\mathbb{R}^n)$ wie verlangt. □

Satz 4.7 (Lineare Transformationsformel) *Für eine lineare Abbildung $S \in \mathbb{R}^{n \times n}$ gilt*

$$\mathcal{L}^n(S(E)) = |\det(S)|\, \mathcal{L}^n(E) \quad \text{für alle } E \subset \mathbb{R}^n.$$

Beweis: Ist $\det(S) = 0$, so liegt $S(E)$ in einer Hyperebene und die Behauptung ist richtig. Für $\det(S) \neq 0$ haben wir aus dem Beweis von Satz 4.6 bereits die Aussage (4.2) zur Verfügung und müssen dort nur noch zeigen:

$$\theta(S) = |\det(S)|.$$

Dies stimmt für eine Diagonalmatrix Λ mit positiven Einträgen $\lambda_i > 0$, $i = 1, \ldots, n$, denn

$$\theta(\Lambda) = \mathcal{L}^n(\Lambda([0,1]^n) = \mathcal{L}^n([0,\lambda_1] \times \ldots \times [0,\lambda_n]) = \prod_{i=1}^n \lambda_i = |\det(\Lambda)|.$$

Für S beliebig sei $S = T_1 \Lambda T_2$ mit einer Diagonalmatrix Λ und $T_1, T_2 \in \mathbb{O}(\mathbb{R}^n)$ wie in Lemma 4.6. Aus den schon bekannten Aussagen für orthogonale sowie Diagonalmatrizen folgt

$$\theta(S) = \mathcal{L}^n(T_1 \Lambda T_2([0,1]^n)) = |\det(\Lambda)| = |\det(S)|,$$

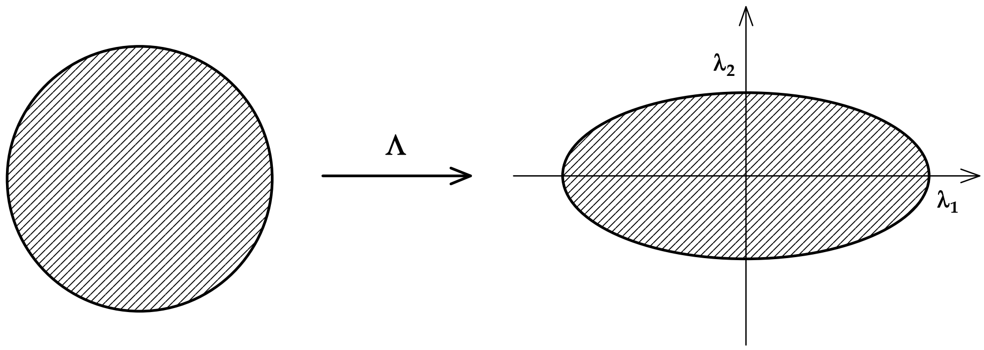
und der Satz ist bewiesen. $\square$

Beispiel 4.2 (Volumen eines Ellipsoids) Für $\lambda_1, \ldots, \lambda_n > 0$ ist die Menge

$$E = \Big\{x \in \mathbb{R}^n : \left(\frac{x_1}{\lambda_1}\right)^2 + \ldots + \left(\frac{x_n}{\lambda_n}\right)^2 < 1\Big\}$$

ein Ellipsoid mit den Halbachsen $\lambda_i > 0$. Mit $B_1(0) = \{x \in \mathbb{R}^n : |x| < 1\}$ gilt $E = \Lambda(B_1(0))$, wobei $\Lambda \in \mathrm{Gl}_n(\mathbb{R})$ die Diagonalmatrix mit den Einträgen λ_i ist. Aus Satz 4.7 folgt

$$\mathcal{L}^n(E) = \mathcal{L}^n(\Lambda(E)) = (\lambda_1 \cdot \ldots \cdot \lambda_n)\ \mathcal{L}^n(B_1(0)).$$



Zum Ende des Kapitels geben wir ein Standardbeispiel für eine nicht messbare Menge an.

Beispiel 4.3 (Vitali 1905) Es gibt eine Menge $S \subset [0,1]$, die nicht $\mathcal{L}^1$-messbar ist. Betrachte dazu auf $[0,1]$ die Äquivalenzrelation

$$x \sim y \quad \Leftrightarrow \quad x - y \in \mathbb{Q}.$$

Mit dem Auswahlaxiom der Mengenlehre erhalten wir ein Repräsentantensystem $S \subset [0,1]$ für die Relation $\sim$, d. h. zu jedem $y \in [0,1]$ gibt es genau ein $x \in S$ mit $x \sim y$. Sei nun $q_1, q_2, \ldots$ eine Abzählung von $\mathbb{Q} \cap [-1,1]$. Dann gilt

$$(q_j + S) \cap (q_k + S) = \emptyset \qquad \text{für } j \neq k.$$

Denn andernfalls gibt es $x_1, x_2 \in S$ mit $q_j + x_1 = q_k + x_2$, also $x_2 - x_1 = q_j - q_k \in \mathbb{Q}$. Nach Definition von S folgt dann $x_1 = x_2$, also $q_j = q_k$ im Widerspruch zur Annahme. Zweitens behaupten wir

$$[0,1] \subset \bigcup_{k=1}^{\infty} (q_k + S) \subset [-1,2].$$

Die rechte Inklusion ist trivial. Die linke Inklusion folgt aus der Definition von S, denn zu $y \in [0,1]$ gibt es ein $x \in S$ mit $y - x =: q \in \mathbb{Q} \cap [-1,1]$.

Nun ist $\mathcal{L}^1(q+S) = \mathcal{L}^1(S)$ wegen der Translationsinvarianz von $\mathcal{L}^1$. Wäre S und damit jede der Mengen $q_k + S$ $\mathcal{L}^1$-meßbar, so folgt aus den obigen beiden Aussagen und der σ-Additivität, siehe Satz 2.4,

$$\sum_{k=1}^{\infty} \mathcal{L}^1(S) = \sum_{k=1}^{\infty} \mathcal{L}^1(q_k + S) = \mathcal{L}^1\left(\bigcup_{k=1}^{\infty} (q_k + S)\right) \in [1,3].$$

Das ist aber unmöglich, da links die Zahl $\mathcal{L}^1(S) \in [0,\infty)$ unendlich oft addiert wird.

5 Das Lebesgueintegral

Ziel dieses Kapitels ist die Definition des Lebesgueintegrals bezüglich eines Maßes μ auf einer σ-Algebra $\mathcal{A}$ über X. Es stellt sich heraus, dass das Integral für eine große Klasse von Funktionen erklärt werden kann. Zum Beispiel ist das Integral einer nichtnegativen Funktion f schon dann in $[0,\infty]$ definiert, wenn f messbar bezüglich $\mathcal{A}$ ist. Dies ist eine überaus milde Regularitätsbedingung.

Die erweiterte Zahlengerade $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ wurde zu Beginn von Kapitel 2 eingeführt. Dort wurden die Ordnungsrelation, der Konvergenzbegriff und die Addition in $\overline{\mathbb{R}}$ erklärt. Für die Multiplikation und Division wird zusätzlich zu den Regeln in $\mathbb{R}$ folgendes vereinbart:

$$\begin{aligned}
s \cdot (\pm\infty) = (\pm\infty) \cdot s \quad &= \quad \begin{cases} \pm\infty & \text{falls } s \in (0,\infty] \\ 0 & \text{falls } s = 0 \\ \mp\infty & \text{falls } s \in [-\infty, 0), \end{cases} \\
\frac{1}{t} \quad &= \quad 0 \qquad \text{für } t = \pm\infty.
\end{aligned}$$

Die Multiplikation $\overline{\mathbb{R}} \times \overline{\mathbb{R}} \to \overline{\mathbb{R}}$ ist damit unstetig in den vier Punkten $\{(0,\pm\infty), (\pm\infty, 0)\}$, aber die Vereinbarung $0 \cdot (\pm\infty) = 0$ erweist sich bei der Definition des Integrals als praktisch. Nicht definiert ist nach wie vor die Division durch Null.

Definition 5.1 (messbare Funktion) *Eine Funktion $f : X \to \overline{\mathbb{R}}$ heißt messbar bezüglich einer σ-Algebra $\mathcal{A} \subset 2^X$ oder kurz $\mathcal{A}$-messbar, falls gilt:*

(i) *$f^{-1}(U) \in \mathcal{A}$ für jede offene Menge $U \subset \mathbb{R}$.*

(ii) *$f^{-1}\{\infty\} \in \mathcal{A}$ und $f^{-1}\{-\infty\} \in \mathcal{A}$.*

Es ist etwas irritierend, dass man hier von messbar redet, bevor ein Maß gegeben ist. Ist μ ein äußeres Maß auf X, so nennen wir eine Funktion $f : X \to \overline{\mathbb{R}}$ μ-messbar, wenn sie messbar ist bezüglich der σ-Algebra $\mathcal{M}$ der μ-messbaren Mengen; dies ist ein wichtiger Spezialfall. Allgemein sprechen wir kurz von messbar statt $\mathcal{A}$-messbar, wenn die σ-Algebra aus dem Kontext eindeutig hervorgeht.

Eine reellwertige Funktion kann natürlich als Funktion nach $\overline{\mathbb{R}}$ aufgefasst werden. Die Frage ihrer Messbarkeit reduziert sich dann auf Bedingung (i) in Definition 5.1. Funktionen mit Werten in $\overline{\mathbb{R}}$ werden auch als numerische Funktionen bezeichnet.

Das nächste Lemma ist nützlich, um die Messbarkeit von Funktionen nachzuweisen.

Lemma 5.1 (Messbarkeitskriterium für Funktionen) *Sei $\mathcal{A} \subset 2^X$ eine σ-Algebra und $f : X \to \overline{\mathbb{R}}$. Dann sind folgende Aussagen äquivalent:*

(i) *f ist $\mathcal{A}$-messbar.*

(ii) *$\{f < s\} = \{x \in X : f(x) \in [-\infty, s)\} \in \mathcal{A}$ für alle $s \in \mathbb{R}$.*

(iii) *$\{f \leq s\} = \{x \in X : f(x) \in [-\infty, s]\} \in \mathcal{A}$ für alle $s \in \mathbb{R}$.*

(iv) $\{f > s\} = \{x \in X : f(x) \in (s,\infty]\} \in \mathcal{A}$ *für alle* $s \in \mathbb{R}$.

(v) $\{f \geq s\} = \{x \in X : f(x) \in [s,\infty)\} \in \mathcal{A}$ *für alle* $s \in \mathbb{R}$.

Beweis: Aus (i) folgt (ii) wegen $\{f < s\} = f^{-1}(-\infty,s) \cup f^{-1}\{-\infty\}$. Aus den folgenden Gleichungen ergibt sich, dass (ii) bis (v) untereinander äquivalent sind:

$$\{f \leq s\} = \bigcap_{k=1}^{\infty} \{f < s + \frac{1}{k}\}, \qquad \{f > s\} = X \backslash \{f \leq s\},$$
$$\{f \geq s\} = \bigcap_{k=1}^{\infty} \{f > s - \frac{1}{k}\}, \qquad \{f < s\} = X \backslash \{f \geq s\}.$$

Es gelte nun eine und damit jede der Aussagen (ii) bis (v). Für ein kompaktes Intervall $[a,b]$ ist dann $f^{-1}([a,b]) = \{f \geq a\} \cap \{f \leq b\} \in \mathcal{A}$. Da sich nach Lemma 1.2 jede offene Menge $U \subset \mathbb{R}$ als abzählbare Vereinigung $U = \bigcup_{k=1}^{\infty} I_k$ von kompakten Intervallen darstellen lässt, ist $f^{-1}(U) = \bigcup_{k=1}^{\infty} f^{-1}(I_k) \in \mathcal{A}$. Ferner haben wir

$$f^{-1}\{\infty\} = \bigcap_{k=1}^{\infty} \{f > k\} \quad \text{und} \quad f^{-1}\{-\infty\} = \bigcap_{k=1}^{\infty} \{f < -k\}.$$

Also ist f $\mathcal{A}$-messbar nach Definition 5.1. □

Wie man leicht sieht, ist das Mengensystem $\{E \subset \mathbb{R} : f^{-1}(E) \in \mathcal{A}\}$ eine σ-Algebra. Ist daher f $\mathcal{A}$-messbar, so ist $f^{-1}(B) \in \mathcal{A}$ für jede Borelmenge $B \subset \mathbb{R}$.

In folgendem Satz sind die Grenzfunktionen punktweise definiert, zum Beispiel ist

$$\liminf_{k\to\infty} f_k : X \to \overline{\mathbb{R}},\ (\liminf_{k\to\infty} f_k)(x) = \liminf_{k\to\infty} f_k(x).$$

Punktweise Konvergenz ist eine sehr schwache Form der Konvergenz, unter der sich Eigenschaften wie Stetigkeit oder Riemann-Integrierbarkeit nicht notwendig auf den Grenzwert übertragen.

Satz 5.1 (Grenzwerte messbarer Funktionen) *Sei $f_k : X \to \overline{\mathbb{R}}$ eine Folge von $\mathcal{A}$-messbaren Funktionen. Dann sind auch folgende Funktionen $\mathcal{A}$-messbar:*

$$\inf_{k\in\mathbb{N}} f_k,\ \sup_{k\in\mathbb{N}} f_k,\ \liminf_{k\to\infty} f_k,\ \limsup_{k\to\infty} f_k.$$

Beweis: Für $s \in \mathbb{R}$ gilt

$$\{\inf_k f_k \geq s\} = \bigcap_{k=1}^{\infty} \{f_k \geq s\}, \quad \{\sup_k f_k \leq s\} = \bigcap_{k=1}^{\infty} \{f_k \leq s\}.$$

Nach Lemma 5.1 sind $\inf_k f_k$ und $\sup_k f_k$ also $\mathcal{A}$-messbar, und damit auch die Funktionen

$$\liminf_{k\to\infty} f_k = \sup_{k\in\mathbb{N}} (\inf_{l\geq k} f_l), \quad \limsup_{k\to\infty} f_k = \inf_{k\in\mathbb{N}} (\sup_{l\geq k} f_l).$$

□

Für $f : X \to \overline{\mathbb{R}}$ ist der Positiv- bzw. Negativanteil $f^{\pm} : X \to [0,\infty]$ definiert durch

$$f^+ = \max(f,0) \geq 0 \quad \text{und} \quad f^- = \max(-f,0) = -\min(f,0) \geq 0. \tag{5.1}$$

Es gilt also $f = f^+ - f^-$ und $|f| = f^+ + f^-$.

Satz 5.2 (Messbarkeit und Rechenoperationen) *Seien $f, g : X \to \overline{\mathbb{R}}$ $\mathcal{A}$-messbar. Dann sind auch folgende Funktionen $\mathcal{A}$-messbar, falls sie definiert sind:*

$$f + g,\ \alpha f \text{ für } \alpha \in \mathbb{R},\ f^{\pm},\ \max(f, g),\ \min(f, g), |f|,\ fg,\ f/g.$$

Beweis: Wir nehmen zunächst an, dass f, g nur Werte in $\mathbb{R}$ annehmen. Es gilt

$$\{f + g < t\} = \bigcup_{r,s \in \mathbb{Q},\, r+s<t} \{f < r\} \cap \{g < s\}, \quad \{-f < t\} = \{f > -t\}.$$

Nach Lemma 5.1 sind also $f + g$ und $-f$ messbar, ebenso αf für $\alpha \in \mathbb{R}$.

Für jedes $\varphi \in C^0(\mathbb{R})$ ist die Verkettung $\varphi \circ f$ messbar, denn für $U \subset \mathbb{R}$ offen ist $\varphi^{-1}(U)$ offen (Analysis II, Satz 1.4), und folglich $(\varphi \circ f)^{-1}(U) = f^{-1}(\varphi^{-1}(U))$ messbar. Damit ergibt sich die Messbarkeit der Funktionen $f^{\pm}$, indem wir $\varphi(s) = \max(\pm s, 0)$ wählen. Weiter sind dann die Funktionen

$$|f| = f^+ + f^-,\ \max(f, g) = \frac{1}{2}(f + g + |f - g|) \text{ und } \min(f, g) = \frac{1}{2}(f + g - |f - g|)$$

messbar. Nun ist $f^2 = \varphi \circ f$ mit $\varphi(s) = s^2$, also folgt die Messbarkeit von f^2 und von

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

Schließlich ist auch $1/g$ messbar, denn

$$\{1/g < s\} = \begin{cases} \{1/s < g < 0\} & s < 0 \\ \{g < 0\} & s = 0 \\ \{g < 0\} \cup \{g > 1/s\} & s > 0. \end{cases}$$

Nimmt nun f (bzw. g) den Wert ∞ oder $-\infty$ an, so betrachte die abgeschnittene Funktion

$$f_k(x) = \begin{cases} k & f(x) \geq k \\ -k & f(x) \leq -k \\ f(x) & \text{sonst.} \end{cases}$$

Die Funktionen f_k, g_k sind messbar. Man prüft nach, dass die Funktionen

$$f_k + g_k,\ \alpha f_k,\ f_k^{\pm},\ \max(f_k, g_k),\ \min(f_k, g_k), |f_k|,\ f_k g_k,\ f_k/g_k$$

punktweise gegen die entsprechenden Funktionen für f und g konvergieren, auch im Fall des (unstetigen) Produkts. Also folgt die allgemeine Behauptung aus Satz 5.1. □

Folgerung 5.1 *Für zwei $\mathcal{A}$-messbare Funktionen $f, g : X \to \overline{\mathbb{R}}$ sind die Mengen $\{f < g\}$, $\{f \leq g\}$, $\{f = g\}$ und $\{f \neq g\}$ Elemente von $\mathcal{A}$.*

Beweis: Es gilt $\{f < g\} = \bigcup_{s \in \mathbb{Q}} \big(\{f < s\} \cap \{g > s\}\big) \in \mathcal{A}$. Die weiteren Aussagen folgen wie in Lemma 5.1. □

Definition 5.2 (Treppenfunktion) *Eine Funktion $\varphi : X \to \mathbb{R}$ heißt $\mathcal{A}$-Treppenfunktion, wenn sie $\mathcal{A}$-messbar ist und nur endlich viele Funktionswerte annimmt. Nach Satz 5.2 ist die Menge $\mathcal{T}_{\mathcal{A}}$ der $\mathcal{A}$-Treppenfunktionen ein $\mathbb{R}$-Vektorraum. Wir setzen*

$$\mathcal{T}_{\mathcal{A}}^{+} = \{\varphi \in \mathcal{T}_{\mathcal{A}} : \varphi \geq 0\}.$$

Beispiel 5.1 Für $E \subset X$ heißt die Funktion

$$\chi_E : X \to \mathbb{R},\ \chi_E(x) = \begin{cases} 1 & \text{falls } x \in E \\ 0 & \text{sonst} \end{cases}$$

charakteristische Funktion von E (alternativ: Indikatorfunktion $\mathbf{1}_E$). Diese Funktion ist genau dann eine $\mathcal{A}$-Treppenfunktion, wenn $E \in \mathcal{A}$.

Der folgende Approximationssatz wird es erlauben, Aussagen für Treppenfunktionen auf beliebige messbare Funktionen zu übertragen.

Satz 5.3 (Approximation durch Treppenfunktionen) *Zu jeder $\mathcal{A}$-messbaren Funktion $f : X \to [0, \infty]$ gibt es eine Folge $f_k \in \mathcal{T}_{\mathcal{A}}^{+}$ mit*

$$f_0 \leq f_1 \leq \ldots \quad \text{und} \quad \lim_{k\to\infty} f_k(x) = f(x) \text{ für alle } x \in X.$$

Beweis: Wir setzen $f_0 = 0$ und definieren für $k \geq 1$ induktiv $E_k = \{f_{k-1} + \frac{1}{k} \leq f\}$ sowie

$$f_k = f_{k-1} + \frac{1}{k}\chi_{E_k} \quad \Rightarrow \quad f_k = \sum_{j=1}^{k} \frac{1}{j}\chi_{E_j}.$$

Die f_k sind Treppenfunktionen mit $f_0 \leq f_1 \leq \ldots$ und $f_k \leq f$ für alle $k \in \mathbb{N}_0$: ist $x \in E_k$, so gilt $f_k(x) = f_{k-1}(x) + \frac{1}{k} \leq f(x)$ nach Definition, für $x \notin E_k$ folgt $f_k(x) = f_{k-1}(x) \leq f(x)$ per Induktion. Insbesondere gilt $\lim_{k\to\infty} f_k(x) \leq f(x)$. Da die Reihe $\sum_{k=1}^{\infty} \frac{1}{k}$ divergiert, gibt es im Fall $\lim_{k\to\infty} f_k(x) < \infty$ unendlich viele $k \in \mathbb{N}$ mit $x \notin E_k$ bzw. $f_{k-1}(x) > f(x) - \frac{1}{k}$, und es folgt $\lim_{k\to\infty} f_k(x) \geq f(x)$. $\square$

Ab jetzt sei ein Maß μ auf der σ-Algebra $\mathcal{A} \subset 2^X$ gegeben. Statt *$\mathcal{A}$-messbar* oder *$\mathcal{A}$-Treppenfunktion* verwenden wir nun *μ-messbar* bzw. *μ-Treppenfunktion* und $\mathcal{T}^{+}(\mu)$, wenn klar ist, auf welcher σ-Algebra das Maß μ gegeben ist. Als Vorstufe zum Lebesgueintegral wird für $\varphi \in \mathcal{T}^{+}(\mu)$ das Integral $I(\varphi)$ auf elementare Weise, also ohne Grenzwertprozess, folgendermaßen definiert: sei $\{s_1, \ldots, s_l\} \subset [0, \infty)$ die Wertemenge von φ. Dann ist

$$I(\varphi) = \sum_{i=1}^{l} s_i\, \mu(\{\varphi = s_i\}). \tag{5.2}$$

Die Summe ist wohldefiniert, eventuell gleich $+\infty$.

Lemma 5.2 (Eigenschaften des Integrals auf $\mathcal{T}^{+}(\mu)$) *Für $\varphi, \psi \in \mathcal{T}^{+}(\mu)$ und $\alpha, \beta \in [0, \infty)$ gelten folgende Aussagen:*

(i) $I(\alpha\varphi + \beta\psi) = \alpha\, I(\varphi) + \beta\, I(\psi)$.

(ii) $\varphi \leq \psi \quad \Rightarrow \quad I(\varphi) \leq I(\psi)$.

Beweis: Ist $\{s_1, \ldots, s_l\}$ die Wertemenge von φ und $\alpha > 0$, so ist $\{\alpha s_1, \ldots, \alpha s_l\}$ die Wertemenge von $\alpha\varphi$ und es folgt

$$I(\alpha\varphi) = \sum_{i=1}^{l} \alpha s_i\, \mu(\{\alpha\varphi = \alpha s_i\}) = \alpha \sum_{i=1}^{l} s_i\, \mu(\{\varphi = s_i\}) = \alpha I(\varphi).$$

Für $\alpha = 0$ gilt $I(\alpha\varphi) = \alpha I(\varphi)$ nach Definition. Seien weiter $\{t_1, \ldots, t_m\}$ und $\{r_1, \ldots, r_n\}$ die Wertemengen von ψ sowie $\varphi + \psi$. Dann berechnen wir

$$\begin{aligned}
I(\varphi + \psi) &= \sum_{k=1}^{n} r_k\, \mu(\{\varphi + \psi = r_k\}) \\
&= \sum_{k=1}^{n} r_k\, \mu\Big(\bigcup_{s_i + t_j = r_k} \{\varphi = s_i\} \cap \{\psi = t_j\} \Big) \\
&= \sum_{k=1}^{n} r_k \sum_{s_i + t_j = r_k} \mu(\{\varphi = s_i\} \cap \{\psi = t_j\}) \\
&= \sum_{k=1}^{n} \sum_{s_i + t_j = r_k} (s_i + t_j)\, \mu(\{\varphi = s_i\} \cap \{\psi = t_j\}) \\
&= \sum_{i=1}^{l} \sum_{j=1}^{m} (s_i + t_j)\, \mu(\{\varphi = s_i\} \cap \{\psi = t_j\}) \\
&= \sum_{i=1}^{l} s_i \sum_{j=1}^{m} \mu(\{\varphi = s_i\} \cap \{\psi = t_j\}) + \sum_{j=1}^{m} t_j \sum_{i=1}^{l} \mu(\{\varphi = s_i\} \cap \{\psi = t_j\}) \\
&= \sum_{i=1}^{l} s_i\, \mu(\{\varphi = s_i\}) + \sum_{j=1}^{m} t_j\, \mu(\{\psi = t_j\}) \\
&= I(\varphi) + I(\psi).
\end{aligned}$$

Damit ist Behauptung (i) gezeigt. In (ii) gilt $\psi - \varphi \in \mathcal{T}^+(\mu)$, und aus (i) folgt

$$I(\psi) = I(\varphi + (\psi - \varphi)) = I(\varphi) + I(\psi - \varphi) \geq I(\varphi).$$

□

Die Definition des Lebesgueintegrals geschieht jetzt in zwei Schritten.

Definition 5.3 (Lebesgueintegral) *Sei μ ein Maß auf X und f eine μ-messbare Funktion. Ist $f : X \to [0, \infty]$, so setzen wir*

$$\int f\, d\mu = \sup\{I(\varphi) : \varphi \in \mathcal{T}^+(\mu),\, \varphi \leq f\}.$$

In diesem Zusammenhang heißt φ Unterfunktion von f. Ist $f : X \to [-\infty, \infty]$ und sind die Integrale von $f^{\pm}$ wie eben definiert nicht beide unendlich, , so setzen wir weiter

$$\int f\, d\mu = \int f^+\, d\mu - \int f^-\, d\mu \in [-\infty, \infty].$$

Die beiden Schritte der Definition sind kompatibel, denn für $f \geq 0$ gilt $f^+ = f$, $f^- = 0$, und für die Nullfunktion ist das in (i) definierte Integral gleich Null.

Folgerung 5.2 (Integral für nichtnegative Treppenfunktionen) *Für $f \in \mathcal{T}^+(\mu)$ gilt*

$$\int f\, d\mu = I(f) = \sum_{s \in f(X)} s\, \mu\big(\{f = s\}\big).$$

Beweis: f ist selbst Unterfunktion von f, also ist $\int f\, d\mu \geq I(f)$. Nach Lemma 5.2(ii) haben wir andererseits $I(\varphi) \leq I(f)$ für jede Unterfunktion. □

Beispiel 5.2 Für das Lebesguemaß $\mathcal{L}^1$ auf $\mathbb{R}$ gilt: $\chi_{\mathbb{Q}} \in \mathcal{T}^+(\mathcal{L}^1)$ und

$$\int \chi_{\mathbb{Q}}\, d\mathcal{L}^1 = 0 \cdot \mathcal{L}^1(\mathbb{R}\backslash\mathbb{Q}) + 1 \cdot \mathcal{L}^1(\mathbb{Q}) = 0.$$

Definition 5.4 (Integrierbarkeit) *Eine Funktion $f : X \to \overline{\mathbb{R}}$ heißt integrierbar bzgl. μ, wenn sie μ-messbar ist und wenn gilt:*

$$\int f\, d\mu \in \mathbb{R} \quad \textit{oder äquivalent} \quad \int f^+\, d\mu + \int f^-\, d\mu < \infty.$$

Beispiel 5.3 Nach Beispiel 2.3 ist card : $2^X \to [0,\infty]$ für jede Menge X ein Maß; wir betrachten hier $X = \mathbb{N}_0$. Eine Funktion $f : \mathbb{N}_0 \to \mathbb{R}$ ist genau dann bzgl. card auf $\mathbb{N}_0$ integrierbar, wenn die Reihe $\sum_{k=0}^\infty f(k)$ absolut konvergiert, und dann gilt

$$\int f\, d\mathrm{card} = \sum_{k=0}^\infty f(k). \tag{5.3}$$

Insbesondere kann eine nur bedingt konvergente Reihe wie $\log 2 = \sum_{k=1}^\infty (-1)^k/k$ nicht als Lebesgueintegral bezüglich card aufgefasst werden. Wir zeigen nun (5.3) zunächst für Funktionen $f : \mathbb{N}_0 \to [0,\infty]$. Dazu betrachten wir

$$f_n : \mathbb{N}_0 \to \mathbb{R},\ f_n(k) = \begin{cases} f(k) & \text{falls } k \leq n \\ 0 & \text{sonst} \end{cases}$$

Die f_n sind Unterfunktionen von f mit $I(f_n) = \sum_{k=0}^n f(k)$. Also folgt

$$\int f\, d\mathrm{card} \geq \lim_{n\to\infty} I(f_n) = \sum_{k=0}^\infty f(k).$$

Für die umgekehrte Ungleichung sei ohne Einschränkung $\sum_{k=0}^\infty f(k) < \infty$, also insbesondere $f(k) \to 0$ mit $k \to \infty$. Ist dann φ Unterfunktion von f, so ist $\varphi(k) \neq 0$ nur für endlich viele k und folglich $\varphi \leq f_n$ für n hinreichend groß. Es folgt

$$I(\varphi) \leq I(f_n) = \sum_{k=0}^n f(k) \leq \sum_{k=0}^\infty f(k), \quad \text{also} \quad \int f\, d\mathrm{card} \leq \sum_{k=0}^\infty f(k).$$

Die Äquivalenz von Integrierbarkeit und absoluter Konvergenz folgt aus

$$\int f^+ \, d\mathrm{card} + \int f^- \, d\mathrm{card} = \sum_{k=0}^{\infty} f^+(k) + \sum_{k=0}^{\infty} f^-(k) = \sum_{k=0}^{\infty} |f(k)|,$$

und weiter erhält man

$$\int f \, d\mathrm{card} = \int f^+ \, d\mathrm{card} - \int f^- \, d\mathrm{card} = \sum_{k=0}^{\infty} f^+(k) - \sum_{k=0}^{\infty} f^-(k) = \sum_{k=0}^{\infty} f(k).$$

Definition 5.5 (Nullmenge für ein Maß) *Sei μ ein Maß auf einer σ-Algebra $\mathcal{A} \subset 2^X$. Eine Menge $N \subset X$ heißt μ-Nullmenge, falls gilt:*

$$N \in \mathcal{A} \quad \text{und} \quad \mu(N) = 0. \tag{5.4}$$

Es kommt in der Maßtheorie oft vor, dass eine Aussage nur für Punkte außerhalb einer μ-Nullmenge gebraucht wird oder gezeigt werden kann. Man sagt, die Aussage $A[x]$ ist wahr für μ-fast-alle $x \in M$ oder μ-fast-überall auf M, falls es eine μ-Nullmenge N gibt mit

$$\{x \in M : A[x] \text{ ist falsch}\} \subset N.$$

Dabei wird nicht verlangt, dass $\{x \in X : A[x] \text{ ist falsch}\}$ selbst zu $\mathcal{A}$ gehört.

Satz 5.4 (Monotonie des Integrals) *Sei μ Maß auf X und $f, g : X \to \overline{\mathbb{R}}$ seien μ-messbar. Ist $f \leq g$ μ-fast-überall und $\int f \, d\mu > -\infty$, so existiert auch das Integral von g und es gilt*

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

Die Aussage mit $\geq$ gilt entsprechend wenn $\int f \, d\mu < \infty$.

Beweis: Seien zunächst f und g nichtnegativ. Ist $\varphi \in \mathcal{T}^+(\mu)$ eine Unterfunktion von f, so ist $\psi := \chi_{\{f \leq g\}} \varphi$ eine Unterfunktion von g, und es gilt

$$\mu(\{\varphi = s\}) = \mu\big(\{\varphi = s\} \cap \{f \leq g\}\big) + \mu\big(\{\varphi = s\} \cap \{f > g\}\big) \leq \mu\big(\{\psi = s\}\big).$$

Es folgt $I(\varphi) \leq I(\psi) \leq \int g \, d\mu$, also $\int f \, d\mu \leq \int g \, d\mu$. Für f, g beliebig ist nach Voraussetzung $f^+ \leq g^+$ sowie $g^- \leq f^-$ μ-fast-überall, also

$$\int f^+ \, d\mu \leq \int g^+ \, d\mu \quad \text{und} \quad \int g^- \, d\mu \leq \int f^- \, d\mu < \infty,$$

und somit

$$\int f \, d\mu = \int f^+ \, d\mu - \int f^- \, d\mu \leq \int g^+ \, d\mu - \int g^- \, d\mu = \int g \, d\mu.$$

□

Bemerkung 5.1 Sei μ vollständiges Maß auf $\mathcal{A}$, das heißt jede Teilmenge einer μ-Nullmenge liegt in $\mathcal{A}$ und ist damit selbst eine μ-Nullmenge, vgl. Definition 3.5. Ist dann f μ-messbar und $f = g$ μ-fast-überall, so ist auch g μ-messbar, denn

$$\{g < s\} = \big(\{f < s\} \cap \{f = g\}\big) \cup \big(\{g < s\} \cap \{f \neq g\}\big) \in \mathcal{A}.$$

Weiter folgt dann $\int g \, d\mu = \int f \, d\mu$, wenn das Integral von f existiert.

Lemma 5.3 (Tschebyscheff-Ungleichung) *Für eine μ-messbare Funktion $f : X \to [0, \infty]$ mit $\int f \, d\mu < \infty$ gilt*

$$\mu(\{f \geq s\}) \leq \begin{cases} \dfrac{1}{s} \displaystyle\int f \, d\mu & \text{für } s \in (0, \infty), \\ 0 & \text{für } s = \infty. \end{cases}$$

Beweis: Für $s \in (0, \infty)$ ist die Funktion $s\chi_{\{f \geq s\}}$ eine Unterfunktion von f, also folgt

$$s \, \mu(\{f = \infty\}) \leq s \, \mu(\{f \geq s\}) = I(s\chi_{\{f \geq s\}}) \leq \int f \, d\mu.$$

□

Folgerung 5.3 *Die Funktion $f : X \to \overline{\mathbb{R}}$ sei μ-messbar.*

(i) *Ist $\int f \, d\mu < \infty$, so ist $\{f = \infty\}$ eine μ-Nullmenge.*

(ii) *Ist $f \geq 0$ und $\int f \, d\mu = 0$, so ist $\{f > 0\}$ eine μ-Nullmenge.*

Beweis: Da f bezüglich der σ-Algebra $\mathcal{A}$ von μ messbar ist, sind die fraglichen Mengen in $\mathcal{A}$. Aussage (i) folgt mit $s = \infty$ aus Lemma 5.3, angewandt auf f^+. In (ii) schließen wir $\mu(\{f \geq s\}) = 0$ für $s > 0$ aus Lemma 5.3, also $\mu(\{f > 0\}) = 0$ mit Satz 2.4(i). □

Um die Linearität des Integrals zu zeigen, brauchen wir den folgenden, zentralen Satz zur Vertauschung von punktweisem Grenzwert und Integral. Mit dieser Problematik werden wir uns im nächsten Kapitel noch ausführlicher beschäftigen.

Satz 5.5 (Satz über monotone Konvergenz von B. Levi) *Sei μ ein Maß auf X und $f_k : X \to [0, \infty]$ eine Folge von μ-messbaren Funktionen mit $f_1 \leq f_2 \leq \ldots$. Definiere $f : X \to [0, \infty]$ durch $f(x) = \lim_{k \to \infty} f_k(x)$. Dann gilt*

$$\int f \, d\mu = \lim_{k \to \infty} \int f_k \, d\mu.$$

Beweis: Die Funktion f ist μ-messbar nach Satz 5.1. Mit Satz 5.4 gilt

$$0 \leq \int f_1 \, d\mu \leq \int f_2 \, d\mu \leq \ldots \leq \int f \, d\mu, \quad \text{also} \quad \lim_{k \to \infty} \int f_k \, d\mu \leq \int f \, d\mu.$$

Sei φ eine Unterfunktion von f mit Wertemenge $\{s_1, \ldots, s_m\}$. Setze $E_i = \{\varphi = s_i\}$ für $i = 1, \ldots, m$ und betrachte für einen Parameter $\theta \in (0, 1)$ die Mengen

$$E_{i,k} = E_i \cap \{f_k \geq \theta s_i\}.$$

Die Funktion $\sum_{i=1}^m \theta s_i \, \chi_{E_{i,k}}$ ist eine Unterfunktion von f_k, und mit Lemma 5.2 folgt

$$\sum_{i=1}^m \theta s_i \, \mu(E_{i,k}) = I\left(\sum_{i=1}^m \theta s_i \, \chi_{E_{i,k}}\right) \leq \int f_k \, d\mu.$$

Nun gilt $E_i = \bigcup_{k=1}^\infty E_{i,k}$, denn für $s_i > 0$ ist $\lim_{k \to \infty} f_k(x) = f(x) \geq s_i > \theta s_i$ für alle $x \in E_i$. Da außerdem $E_{i,1} \subset E_{i,2} \subset \ldots$, folgt aus der Stetigkeit des Maßes von unten, siehe 2.4(i), die Gleichung $\mu(E_i) = \lim_{k \to \infty} \mu(E_{i,k})$ und somit

$$\theta \, I(\varphi) = \sum_{i=1}^m \theta s_i \, \mu(E_i) = \lim_{k \to \infty} \sum_{i=1}^m \theta s_i \, \mu(E_{i,k}) \leq \lim_{k \to \infty} \int f_k \, d\mu.$$

Mit $\theta \nearrow 1$ und Bildung des Supremums über alle Unterfunktionen φ folgt

$$\int f \, d\mu \leq \lim_{k\to\infty} \int f_k.$$

$\square$

Satz 5.6 (Linearität des Integrals) *Sei μ ein Maß auf X. Für μ-messbare Funktionen $f, g : X \to \overline{\mathbb{R}}$ und $\alpha, \beta \in \mathbb{R}$ sei $\alpha \int f \, d\mu + \beta \int g \, d\mu$ in $\overline{\mathbb{R}}$ definiert. Dann ist $\alpha f + \beta g$ außerhalb einer μ-Nullmenge N definiert, und mit $\alpha f + \beta g := 0$ auf N gilt*

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

Beweis: Wir betrachten erst den Fall $\beta = 0$. Ist $\alpha > 0$, $f \geq 0$ und φ Unterfunktion von f, so ist $\alpha\varphi$ Unterfunktion von αf und es folgt aus Lemma 5.2(ii)

$$\alpha \, I(\varphi) = I(\alpha\varphi) \leq \int (\alpha f) \, d\mu \quad \text{also} \quad \alpha \int f \, d\mu \leq \int (\alpha f) \, d\mu.$$

Anwendung auf $1/\alpha$ und αf (statt α und f) ergibt die gewünschte Gleichheit. Für $\alpha > 0$ und f wie in der Behauptung gilt $(\alpha f)^+ = \alpha f^+$ und $(\alpha f)^- = \alpha f^-$. Es folgt

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

Schließlich folgt aus $(-f)^+ = f^-$ und $(-f)^- = f^+$ die Gleichung

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

Es reicht, als zweites den Fall $\alpha = \beta = 1$ zu behandeln. Sind $f, g : X \to [0, \infty]$, so gibt es nach Satz 5.3 Folgen $\varphi_k, \psi_k \in \mathcal{T}^+(\mu)$ mit $\varphi_1 \leq \varphi_2 \leq \ldots$ und $\psi_1 \leq \psi_2 \leq \ldots$, so dass $\varphi_k \to f$ und $\psi_k \to g$ punktweise auf X. Es folgt

$$\varphi_1 + \psi_1 \leq \varphi_2 + \psi_2 \leq \ldots \quad \text{und} \quad \varphi_k + \psi_k \to f + g \text{ punktweise auf } X.$$

Mit dem Satz über monotone Konvergenz und Lemma 5.2(ii) erhalten wir

$$\int (f+g) \, d\mu = \lim_{k\to\infty} \int (\varphi_k + \psi_k) \, d\mu = \lim_{k\to\infty} \left(\int \varphi_k \, d\mu + \int \psi_k \, d\mu \right) = \int f \, d\mu + \int g \, d\mu.$$

Seien schließlich $f, g; X \to \overline{\mathbb{R}}$ wie in der Behauptung. Indem wir evtl. zu $-f$ und $-g$ übergehen und den ersten Fall mit $\alpha = -1$ anwenden, gilt oBdA

$$\int f \, d\mu + \int g \, d\mu > -\infty, \quad \text{also} \quad \int f^- \, d\mu + \int g^- \, d\mu < \infty.$$

Nach Folgerung 5.3(i) ist dann $\mu(\{f = -\infty\} \cup \{g = -\infty\}) = 0$, und somit ist auch die Teilmenge $\{(f,g) = \pm(\infty, -\infty)\} \in \mathcal{A}$ eine μ-Nullmenge. Aus $(f+g)^- \leq f^- + g^-$, der Monotonie des Integrals und dem schon bewiesenen nichtnegativen Fall folgt

$$\int (f+g)^- \, d\mu \leq \int (f^- + g^-) \, d\mu = \int f^- \, d\mu + \int g^- \, d\mu < \infty.$$

Also ist das Integral von $f+g$ definiert. Nun gilt $(f+g)^+ - (f+g)^- = f+g = f^+ + g^+ - (f^- + g^-)$, beziehungsweise

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

Da auf beiden Seiten nichtnegative Funktionen stehen, folgt durch Integration

$$\int (f+g)^+ d\mu + \int f^-\, d\mu + \int g^-\, d\mu = \int (f+g)^- d\mu + \int f^+\, d\mu + \int g^+\, d\mu. \tag{5.5}$$

Ist nun $\int f^+\, d\mu + \int g^+\, d\mu < \infty$, so sind alle beteiligten Integrale endlich und wir erhalten

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

Für $\int f^+\, d\mu + \int g^+\, d\mu = \infty$ folgt $\int (f+g)^+\, d\mu = \infty$ wieder aus (5.5), und in der Behauptung sind beide Seiten gleich ∞. Damit ist der Satz bewiesen. □

Definition 5.6 (Integration über Teilmengen) *Sei μ ein Maß auf X und $E \subset X$ sei μ-messbar. Dann setzen wir, wenn das rechte Integral existiert,*

$$\int_E f\, d\mu = \int f \chi_E\, d\mu.$$

f heißt auf E integrierbar, wenn die Funktion $f\chi_E$ integrierbar ist.

Wegen $(f\chi_E)^\pm = f^\pm \chi_E \leq f^\pm$ existiert das Integral von f über E auf jeden Fall dann, wenn das Integral von f über ganz X existiert. Insbesondere ist das Integral von f über E stets definiert, wenn f nichtnegativ ist.

Beispiel 5.4 Betrachte für $\alpha \in \mathbb{R}$ die Funktion $f : \mathbb{R}^n \to \mathbb{R}$, $f(x) = |x|^{-\alpha}$. Wir behaupten:

$$\int_{\mathbb{R}^n \backslash B_1(0)} f\, d\mathcal{L}^n < \infty \Leftrightarrow \alpha > n, \quad \text{und} \quad \int_{B_1(0)} f\, d\mathcal{L}^n < \infty \Leftrightarrow \alpha < n.$$

Zum Beweis vergleichen wir f mit der Funktion

$$g = \sum_{k \in \mathbb{Z}} 2^{-k\alpha}\, \chi_{A_k} \quad \text{mit } A_k = \{2^k \leq |x| < 2^{k+1}\}.$$

Es gelten die Abschätzungen

$$2^{-\alpha} g \leq f \leq g \text{ für } \alpha \geq 0 \quad \text{bzw.} \quad 2^{-\alpha} g \geq f \geq g \text{ für } \alpha \leq 0.$$

Wegen der Monotonie des Integrals reicht es also aus, die Aussagen für g zu zeigen. Da $A_k = 2^k A_0$, folgt aus der Transformation von $\mathcal{L}^n$ unter Streckungen, vgl. Satz 4.6,

$$\mathcal{L}^n(A_k) = (2^k)^n\, \mathcal{L}^n(A_0) = 2^{nk}\, \gamma_n \quad \text{mit } \gamma_n = \mathcal{L}^n(A_0) \in (0, \infty).$$

Da die Folge der Partialsummen $\sum_{k=0}^{l} 2^{-k\alpha}\chi_{A_k}$ punktweise auf $\mathbb{R}^n$ gegen $g\chi_{\mathbb{R}^n\setminus B_1(0)}$ konvergiert, folgt aus dem Satz über monotone Konvergenz

$$\int_{\mathbb{R}^n\setminus B_1(0)} g\,d\mathcal{L}^n = \sum_{k=0}^{\infty} \int 2^{-k\alpha}\chi_{A_k}\,d\mathcal{L}^n = \gamma_n \sum_{k=0}^{\infty} 2^{(n-\alpha)k} = \begin{cases} \gamma_n \dfrac{1}{1-2^{n-\alpha}} & \text{falls } \alpha > n \\ \infty & \text{sonst.} \end{cases}$$

Ganz entsprechend erhalten wir auf $B_1(0)$

$$\int_{B_1(0)} g\,d\mathcal{L}^n = \sum_{k=-1}^{-\infty} \int 2^{-k\alpha}\chi_{A_k}\,d\mathcal{L}^n = \gamma_n \sum_{k=-1}^{-\infty} 2^{(n-\alpha)k} = \begin{cases} \gamma_n \dfrac{1}{2^{n-\alpha}-1} & \text{falls } \alpha < n \\ \infty & \text{sonst.} \end{cases}$$

Damit ist die Behauptung von Beispiel 5.4 gezeigt.

Das folgende Integrierbarkeitskriterium (iii) wird sehr oft benutzt, meistens ohne extra erwähnt zu werden.

Satz 5.7 (Majorantenkriterium) *Sei $f : X \to \overline{\mathbb{R}}$ μ-messbar.*

(i) *f integrierbar $\Leftrightarrow$ $|f|$ integrierbar.*

(ii) *Es gilt $|\int f\,d\mu| \leq \int |f|\,d\mu$, wenn das Integral von f existiert.*

(iii) *Ist $g : X \to [0,\infty]$ μ-messbar mit $|f| \leq g$ μ-fast-überall und $\int g\,d\mu < \infty$, so ist f integrierbar.*

Beweis: Es gilt $|f| = f^+ + f^-$, und aus der Linearität des Integrals, Satz 5.6, folgt

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

Mit Definition 5.4 folgt (i). Ist das Integral von f definiert, so gilt weiter

$$\Big|\int f\,d\mu\Big| = \Big|\int f^+\,d\mu - \int f^-\,d\mu\Big| \leq \int f^+\,d\mu + \int f^-\,d\mu = \int |f|\,d\mu.$$

Ist schließlich g wie in (iii), so folgt $\int |f|\,d\mu \leq \int g\,d\mu < \infty$ aus Satz 5.4 und dann mit (i) die Integrierbarkeit von f. □

Mit dem Majorantenkriterium und Beispiel 5.4 folgt, dass allgemein Funktionen integrierbar sind, die durch eine geeignete Potenz $|x|^{-\alpha}$ abgeschätzt sind, genauer:

Beispiel 5.5 Ist $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ $\mathcal{L}^n$-messbar und gilt für eine Konstante $C \in [0,\infty)$

$$\begin{aligned} |f(x)| &\leq C\,|x|^{-\alpha} \text{ fast überall in } B_\varepsilon(0) \text{ mit } \alpha < n, \quad \text{bzw.} \\ |f(x)| &\leq C\,|x|^{-\alpha} \text{ fast überall in } \mathbb{R}^n\setminus B_R(0) \text{ mit } \alpha > n, \end{aligned}$$

so ist f auf $B_\varepsilon(0)$ bzw. auf $\mathbb{R}^n\setminus B_R(0)$ integrierbar.

6 Konvergenzsätze und L^p-Räume

Sei $f_k : X \to \mathbb{R}$ eine Folge von Funktionen, die gegen eine Funktion $f : X \to \mathbb{R}$ konvergiert. Welche Voraussetzungen sind an die Qualität der Konvergenz $f_k \to f$ zu stellen, damit der Grenzübergang mit dem Integral vertauscht werden kann, das heißt damit gilt:

$$\int f \, d\mu = \lim_{k\to\infty} \int f_k \, d\mu.$$

Die Konvergenzsätze von B. Levi und H. Lebesgue geben mit der monotonen Konvergenz bzw. der majorisierten Konvergenz hinreichende Bedingungen an, die wesentlich schwächer sind als die beim Riemannintegral benötigte gleichmäßige Konvergenz. Der Satz über majorisierte Konvergenz liefert (unter anderem) folgendes zentrale Resultat von Riesz-Fischer: der Raum der integrierbaren Funktionen – modulo Gleichheit fast überall – ist mit der Integralnorm ein Banachraum.

Die punktweise Konvergenz $f_k \to f$ reicht im allgemeinen nicht aus, um das Integral mit dem Grenzwert zu vertauschen.

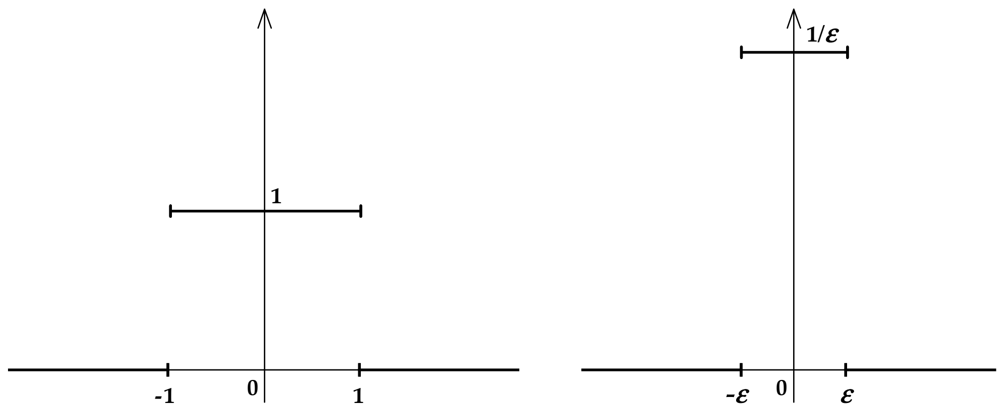
Beispiel 6.1 Betrachte auf $\mathbb{R}$ die charakteristischen Funktionen $f_\varepsilon = \frac{1}{2\varepsilon}\chi_{[-\varepsilon,\varepsilon]}$. Es gilt

$$\lim_{\varepsilon\searrow 0} f_\varepsilon(x) = f_0(x) \text{ mit } f_0(x) = \begin{cases} 0 & \text{für } x \neq 0 \\ \infty & \text{für } x = 0. \end{cases}$$

Andererseits haben wir $\int f_\varepsilon \, d\mathcal{L}^1 = \frac{1}{2\varepsilon}\mathcal{L}^1([-\varepsilon,\varepsilon]) = 1$ für alle $\varepsilon > 0$, das heißt

$$\int f_0 \, d\mathcal{L}^1 = 0 < 1 = \lim_{\varepsilon\searrow 0} \int f_\varepsilon \, d\mathcal{L}^1.$$

Für festes $x \neq 0$ gilt: $\sup_{\varepsilon>0} f_\varepsilon(x) = \dfrac{1}{2|x|}$.



Eine für die Vertauschung hinreichende, zusätzliche Bedingung liefert der Satz über monotone Konvergenz, der bereits im vorigen Kapitel für den Nachweis der Linearität des Integrals benötigt wurde. Wir wiederholen den Satz hier unverändert.

Satz 5.5 (über monotone Konvergenz von B. Levi) Sei $f_k : X \to [0,\infty]$ eine Folge von μ-messbaren Funktionen mit $f_1 \leq f_2 \leq \ldots$. Definiere $f : X \to [0,\infty]$ durch $f(x) = \lim_{k\to\infty} f_k(x)$. Dann gilt

$$\int f \, d\mu = \lim_{k\to\infty} \int f_k \, d\mu.$$

Als Anwendung wollen wir kurz Maße besprechen, die durch Integration gegen eine Dichtefunktion gegeben sind.

Satz 6.1 (Maß mit Dichtefunktion) *Sei μ ein Maß auf der σ-Algebra $\mathcal{A} \subset 2^X$. Für eine μ-messbare Funktion $\theta : X \to [0,\infty]$ ist dann*

$$\lambda : \mathcal{A} \to [0,\infty],\ \lambda(A) = \int_A \theta \, d\mu = \int \theta \chi_A \, d\mu$$

ein Mass, das mit $\mu \llcorner \theta$ (oder $\mu \odot \theta$) bezeichnet wird. Es gelten folgende Aussagen:

(1) $\mu(A) = 0 \quad \Rightarrow \quad \mu\llcorner\theta(A) = 0.$

(2) $\int f \, d(\mu\llcorner\theta) = \int f\theta \, d\mu$ *für alle μ-messbaren $f : X \to [0,\infty]$.*

Beweis: Ist $A = \bigcup_{i=1}^\infty A_i$ mit $A_i \in \mathcal{A}$ paarweise disjunkt, so folgt mit monotoner Konvergenz

$$\lambda(A) = \int \theta\chi_A \, d\mu = \int \sum_{i=1}^\infty \theta\chi_{A_i} \, d\mu = \sum_{i=1}^\infty \int \theta\chi_{A_i} \, d\mu = \sum_{i=1}^\infty \lambda(A_i).$$

Wegen $\chi_\emptyset \equiv 0$ ist $\lambda(\emptyset) = 0$, das heißt λ ist ein Maß auf $\mathcal{A}$. Für $\mu(A) = 0$, also $\theta\chi_A$ μ-fast-überall Null, folgt aus Satz 5.4

$$\mu\llcorner\theta(A) = \int \theta\chi_A \, d\mu = 0.$$

Weiter gilt für $A \in \mathcal{A}$

$$\int \chi_A \, d(\mu\llcorner\theta) = \mu\llcorner\theta(A) = \int \chi_A\theta \, d\mu.$$

Damit folgt (2) für alle Treppenfunktionen $\varphi \in \mathcal{T}^+(\mu)$. Für eine beliebige μ-messbare Funktion $f \geq 0$ wähle eine Folge von Treppenfunktionen $\varphi_k \in \mathcal{T}^+(\mu)$ mit $\varphi_k \nearrow f$ nach Satz 5.3, und schließe wieder mit monotoner Konvergenz

$$\int f \, d(\mu\llcorner\theta) = \lim_{k\to\infty} \int \varphi_k \, d(\mu\llcorner\theta) = \lim_{k\to\infty} \int \varphi_k\theta \, d\mu = \int f\theta \, d\mu.$$

$\square$

Bemerkung 6.1 Ist zusätzlich $\theta : X \to [0,\infty]$ integrierbar bezüglich μ, so lässt sich Aussage (1) von Satz 6.1 wie folgt verschärfen: zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$, so dass gilt:

$$A \in \mathcal{A},\ \mu(A) < \delta \quad \Rightarrow \quad (\mu\llcorner\theta)(A) < \varepsilon. \tag{6.1}$$

Denn mit $\theta_k = \min(\theta, k)$ gilt die Abschätzung

$$(\mu\llcorner\theta)(A) = \int_A \theta \, d\mu = \int_A (\theta - \theta_k) \, d\mu + \int \theta_k \, d\mu \leq \int (\theta - \theta_k) \, d\mu + k\,\mu(A).$$

Nach Satz 5.5 können wir $k \in \mathbb{N}$ hinreichend groß wählen, so dass

$$\int (\theta - \theta_k)\, d\mu = \int \theta\, d\mu - \int \theta_k\, d\mu < \varepsilon/2.$$

Die Bemerkung folgt für $\mu(A) < \varepsilon/(2k) =: \delta$.

Das folgende Lemma ist von unabhängigem Interesse, zum Beispiel in der Variationsrechnung. Es besagt unter anderem, dass das Integral bezüglich punktweiser Konvergenz nichtnegativer Funktionen unterhalbstetig ist.

Satz 6.2 (Lemma von Fatou) *Sei $f_k : X \to [0,\infty]$ eine Folge von μ-messbaren Funktionen. Für $f : X \to \overline{\mathbb{R}}$, $f(x) = \liminf_{k\to\infty} f_k(x)$ gilt dann*

$$\int f\, d\mu \le \liminf_{k\to\infty} \int f_k\, d\mu.$$

Beweis: Für die Folge $g_k = \inf_{j \ge k} f_j$ gilt $g_{k+1} \ge g_k$ und $\lim_{k\to\infty} g_k = f$. Mit Satz 5.5 folgt, da andererseits $g_k \le f_k$ für alle $k \in \mathbb{N}$,

$$\int f\, d\mu = \lim_{k\to\infty} \int g_k\, d\mu \le \liminf_{k\to\infty} \int f_k\, d\mu.$$

□

Satz 6.3 (Satz über majorisierte Konvergenz von Lebesgue) *Sei $f_1, f_2, \ldots$ eine Folge von μ-messbaren Funktionen, und $f(x) = \lim_{k\to\infty} f_k(x)$ für μ-fast-alle $x \in X$. Es gebe eine integrierbare Funktion $g : X \to [0,\infty]$ mit $\sup_k |f_k(x)| \le g(x)$ für μ-fast-alle x. Dann ist f integrierbar bzgl. μ, und es gilt*

$$\int f\, d\mu = \lim_{k\to\infty} \int f_k\, d\mu.$$

Es gilt sogar $\|f - f_k\|_{L^1(\mu)} = \int |f - f_k|\, d\mu \to 0$, vgl. Definition 6.1.

Beweis: Die Folge $2g - |f - f_k| \ge 0$ konvergiert punktweise fast überall gegen g. Mit Satz 6.2 erhalten wir

$$\begin{aligned}
\limsup_{k\to\infty} \left|\int f\, d\mu - \int f_k\, d\mu\right| &\le \limsup_{k\to\infty} \int |f - f_k|\, d\mu \\
&= \int 2g\, d\mu - \liminf_{k\to\infty} \int (2g - |f - f_k|)\, d\mu \\
&\le \int 2g\, d\mu - \int \liminf_{k\to\infty} (2g - |f - f_k|)\, d\mu \\
&= 0.
\end{aligned}$$

□

Als erste Anwendung wollen wir das eindimensionale Riemannintegral mit dem Lebesgueintegral bzgl. des Maßes $\mathcal{L}^1$ vergleichen. Sei $I = [a,b]$ ein kompaktes Intervall und $f : I \to \mathbb{R}$

beschränkt. Für eine durch Unterteilungspunkte $a = x_0 \leq \ldots \leq x_N = b$ gegebene Zerlegung $\mathcal{Z}$ von $[a, b]$ in Teilintervalle $I_j = [x_{j-1}, x_j]$ werden Ober- und Untersumme wie folgt gebildet:

$$\overline{S}_{\mathcal{Z}}(f) = \sum_{j=1}^{N} (\sup_{I_j} f)\,(x_j - x_{j-1}) \quad \text{bzw.} \quad \underline{S}_{\mathcal{Z}}(f) = \sum_{j=1}^{N} (\inf_{I_j} f)\,(x_j - x_{j-1}).$$

Für zwei Zerlegungen $\mathcal{Z}_1$ und $\mathcal{Z}_2$ mit gemeinsamer Verfeinerung $\mathcal{Z}_1 \cup \mathcal{Z}_2$ sieht man leicht

$$\underline{S}_{\mathcal{Z}_1}(f) \leq \underline{S}_{\mathcal{Z}_1 \cup \mathcal{Z}_2}(f) \leq \overline{S}_{\mathcal{Z}_1 \cup \mathcal{Z}_2}(f) \leq \overline{S}_{\mathcal{Z}_2}(f).$$

f heißt Riemannintegrierbar mit Integral $\int_a^b f(x)\,dx = S$, wenn gilt:

$$\sup_{\mathcal{Z}} \underline{S}_{\mathcal{Z}}(f) = \inf_{\mathcal{Z}} \overline{S}_{\mathcal{Z}}(f) = S.$$

Aus der Vorlesung Analysis I sind hinreichende Kriterien für die Riemannintegrierbarkeit bekannt, zum Beispiel Stetigkeit auf $[a, b]$. Es blieb aber die Frage offen, welche Funktionen – im Sinne einer Charakterisierung durch punktweise Eigenschaften – Riemannintegrierbar sind; dies können wir nun beantworten. Das Ergebnis ist analog zur Charakterisierung der Quadrierbarkeit in Satz 1.6. Ein positiver Nebeneffekt ist, dass wir so die Integrationsregeln aus Analysis I auch für das Lebesgueintegral zur Verfügung haben, jedenfalls wenn die Funktionen Riemannintegrierbar sind.

Satz 6.4 (Riemannintegrierbarkeit) *Sei $f : I \to \mathbb{R}$ eine beschränkte Funktion auf dem kompakten Intervall $I = [a, b]$. Dann gilt:*

$$f \text{ Riemannintegrierbar} \quad \Leftrightarrow \quad \mathcal{L}^1(\{x \in I : f \text{ ist nicht stetig in } x\}) = 0.$$

In diesem Fall ist f auch Lebesgueintegrierbar, und die Integrale stimmen überein.

Beweis: Für eine Zerlegung $\mathcal{Z}$ mit Teilintervallen $I_j = [x_{j-1}, x_j]$, $1 \leq j \leq N$, definieren wir die Riemannschen Treppenfunktionen

$$\overline{f}_{\mathcal{Z}}(x) = \max_{I_j \ni x} \sup_{I_j} f \geq \limsup_{y \to x} f(y) \quad \text{und} \quad \underline{f}_{\mathcal{Z}}(x) = \min_{I_j \ni x} \inf_{I_j} f \leq \liminf_{y \to x} f(y).$$

Sei $N_f(s) = \{x \in I : \limsup_{y \to x} f(y) - \liminf_{y \to x} f(y) \geq s\}$ für $s > 0$. Sind $\mathcal{Z}_1, \mathcal{Z}_2$ beliebige Zerlegungen, so folgt $\overline{f}_{\mathcal{Z}_2}(x) - \underline{f}_{\mathcal{Z}_1}(x) \geq s$ für alle $x \in N_f(s)$, und hieraus mit Lemma 5.3

$$\overline{S}_{\mathcal{Z}_2}(f) - \underline{S}_{\mathcal{Z}_1}(f) = \int_I \left(\overline{f}_{\mathcal{Z}_2} - \underline{f}_{\mathcal{Z}_1}\right) d\mathcal{L}^1 \geq s\,\mathcal{L}^1(N_f(s)).$$

Ist f Riemannintegrierbar, so bilden wir das Infimum über alle $\mathcal{Z}_1, \mathcal{Z}_2$ und schließen $\mathcal{L}^1(N_f(s)) = 0$ für alle $s > 0$, womit die eine Richtung der Behauptung gezeigt ist.

Sei nun f $\mathcal{L}^1$-fast-überall stetig, und $\mathcal{Z}_i$ eine beliebige Folge von Zerlegungen mit Feinheit $\delta_i := \max_{1 \leq j \leq N_i} |x_{i,j} - x_{i,j-1}| \to 0$. Ist f stetig in x, so folgt

$$\overline{f}_{\mathcal{Z}_i}(x) \leq \sup_{|y-x| \leq \delta_i} f(y) \searrow f(x) \quad \text{und} \quad \underline{f}_{\mathcal{Z}_i}(x) \geq \inf_{|y-x| \leq \delta_i} f(y) \nearrow f(x) \text{ mit } i \to \infty.$$

Also konvergieren $\overline{f}_{\mathcal{Z}_i}$, $\underline{f}_{\mathcal{Z}_i}$ punktweise $\mathcal{L}^1$-fast-überall auf I gegen f, insbesondere ist f $\mathcal{L}^1$-messbar nach Satz 5.1. Wegen $|\overline{f}_{\mathcal{Z}_i}|, |\underline{f}_{\mathcal{Z}_i}| \leq \sup_I |f| < \infty$ folgt aus Satz 6.3

$$\overline{S}_{\mathcal{Z}_i}(f) = \int_I \overline{f}_{\mathcal{Z}_i}\, d\mathcal{L}^1 \to \int_I f\, d\mathcal{L}^1 \quad \text{und} \quad \underline{S}_{\mathcal{Z}_i}(f) = \int_I \underline{f}_{\mathcal{Z}_i}\, d\mathcal{L}^1 \to \int_I f\, d\mathcal{L}^1.$$

Also ist f Riemannintegrierbar mit Riemannintegral $\int_a^b f(x)\, dx = \int_I f\, d\mathcal{L}^1$. □

Ein uneigentliches Riemannintegral kann dann und nur dann als Lebesgueintegral aufgefasst werden, wenn es absolut konvergiert. Dies zeigt man leicht mit Definition 5.4, Satz 6.4 und einem Konvergenzsatz. Zum Beispiel ist das Integral $\int_0^\infty \frac{\sin x}{x}\, dx$ nicht als Lebesgueintegral definiert, vergleiche auch Beispiel 5.3 und Beispiel 6.2 unten.

Wir kommen nun zu einer zweiten Anwendung des Satzes von Lebesgue, nämlich der Frage der Stetigkeit und Differenzierbarkeit von Integralen, deren Integrand von einem Parameter abhängt.

Satz 6.5 (Stetigkeit von Parameterintegralen) *Sei X ein metrischer Raum, μ ein Maß auf Y und $f : X \times Y \to \mathbb{R}$ mit $f(x, \cdot)$ integrierbar bzgl. μ für alle $x \in X$. Betrachte*

$$F : X \to \mathbb{R},\ F(x) = \int f(x, y)\, d\mu(y).$$

Es sei $f(\cdot, y)$ stetig in $x_0 \in X$ für μ-fast-alle $y \in Y$. Weiter gebe es eine μ-integrierbare Funktion $g : Y \to [0, \infty]$, so dass für alle $x \in X$ gilt:

$$|f(x, y)| \leq g(y) \ \text{ für alle } y \in Y \backslash N_x, \ \text{ mit einer } \mu\text{-Nullmenge } N_x.$$

Dann ist F stetig in x_0.

Beweis: Zu jeder Folge $x_k \to x_0$ gibt es nach Voraussetzung eine μ-Nullmenge N, so dass für alle $y \in Y \backslash N$ gilt:

$$f(x_k, y) \to f(x_0, y) \quad \text{und} \quad |f(x_k, y)| \leq g(y) \text{ für alle } k \in \mathbb{N}.$$

Aus dem Konvergenzsatz von Lebesgue, Satz 6.3, folgt

$$F(x) = \int f(x, y)\, d\mu(y) = \lim_{k \to \infty} \int f(x_k, y)\, d\mu(y) = \lim_{k \to \infty} F(x_k).$$

□

Satz 6.6 (Differentiation unter dem Integralzeichen) *Sei I offenes Intervall, μ ein Maß auf Y und $f : I \times Y \to \mathbb{R}$ mit $f(x, \cdot)$ integrierbar bzgl. μ für alle $x \in I$. Setze*

$$F : I \to \mathbb{R},\ F(x) = \int f(x, y)\, d\mu(y).$$

Es sei $f(\cdot, y)$ in $x_0 \in I$ differenzierbar für μ-fast-alle $y \in Y$, und es gebe eine μ-integrierbare Funktion $g : Y \to [0, \infty]$, so dass für alle $x \in I$ gilt:

$$\frac{|f(x, y) - f(x_0, y)|}{|x - x_0|} \leq g(y) \ \text{ für alle } y \in Y \backslash N_x,$$

mit einer geeigneten μ-Nullmenge N_x. Dann folgt

$$F'(x_0) = \int \frac{\partial f}{\partial x}(x_0, y)\, d\mu(y).$$

Beweis: Zu jeder Folge $x_k \to x_0$ gibt es nach Voraussetzung eine μ-Nullmenge $N \subset Y$, so dass für alle $y \in Y \backslash N$ gilt:

$$\lim_{k \to \infty} \frac{f(x_k, y) - f(x_0, y)}{x_k - x_0} = \frac{\partial f}{\partial x}(x_0, y),$$

$$\left| \frac{f(x_k, y) - f(x_0, y)}{x_k - x_0} \right| \leq g(y).$$

Aus dem Konvergenzsatz von Lebesgue, Satz 6.3, folgt mit $k \to \infty$

$$\frac{F(x_k) - F(x_0)}{x_k - x_0} = \int \frac{f(x_k, y) - f(x_0, y)}{x_k - x_0}\, d\mu(y) \to \int \frac{\partial f}{\partial x}(x_0, y)\, d\mu(y).$$

□

Wir wollen noch eine Version von Satz 6.6 formulieren, die etwas leichter zu handhaben ist.

Folgerung 6.1 *Sei $U \subset \mathbb{R}^n$ offen, μ ein Maß auf Y und $f : U \times Y \to \mathbb{R}$ mit $f(x, \cdot)$ integrierbar bezüglich μ für alle $x \in U$. Betrachte*

$$F : U \to \mathbb{R},\ F(x) = \int f(x, y)\, d\mu(y).$$

Es gebe eine μ-Nullmenge $N \subset Y$, so dass für alle $y \in Y \backslash N$ gilt: $f(\cdot, y) \in C^1(U)$ und

$$|D_x f(x, y)| \leq g(y) \quad \text{mit einer integrierbaren Funktion } g : Y \to [0, \infty].$$

Dann ist $F \in C^1(U)$ und es gilt für alle $x \in U$

$$\frac{\partial F}{\partial x_i}(x) = \int \frac{\partial f}{\partial x_i}(x, y)\, d\mu(y).$$

Beweis: Nach Voraussetzung gilt für alle $y \in Y$ mit Ausnahme der μ-Nullmenge N

$$\frac{|f(x + he_i, y) - f(x, y)|}{|h|} \leq \int_0^1 \left| \frac{\partial f}{\partial x_i}(x + the_i, y) \right| dt \leq g(y).$$

Nach Satz 6.6 ist die Funktion F in jedem Punkt $x \in U$ nach x_i partiell differenzierbar, und die partielle Ableitung ist durch Differentiation unter dem Integralzeichen gegeben. Aber nach Satz 6.5 sind die partiellen Ableitungen stetig auf U, also ist $F \in C^1(U)$. □

Beispiel 6.2 Als Beispiel berechnen wir den Grenzwert

$$\int_0^\infty \frac{\sin x}{x}\, dx := \lim_{r \to \infty} \int_0^r \frac{\sin x}{x}\, dx.$$

Dazu bilden wir das Parameterintegral

$$F : [0, \infty) \to \mathbb{R},\ F(t) = \int_0^\infty e^{-tx} \frac{\sin x}{x}\, dx.$$

Der Integrand $f(t,x) = e^{-tx}\frac{\sin x}{x}$ ist bzgl. $x \in (0,\infty)$ integrierbar für jedes $t > 0$, sowie nach $t \in (0,\infty)$ stetig differenzierbar für alle $x > 0$, und erfüllt für $t > \tau$ die Abschätzung

$$\left|\frac{\partial f}{\partial t}(t,x)\right| \leq e^{-\tau x} =: g(x).$$

Also gilt nach Folgerung 6.1 für $t > \tau$, und damit für alle $t > 0$,

$$\begin{aligned}
F'(t) &= -\int_0^\infty e^{-tx}\sin x\,dx \\
&= \left[e^{-tx}\cos x\right]_{x=0}^{x=\infty} - \int_0^\infty (-t)e^{-tx}\cos x\,dx \\
&= -1 + t^2\int_0^\infty e^{-tx}\sin x\,dx, \qquad \text{also} \\
F'(t) &= -\frac{1}{1+t^2}.
\end{aligned}$$

Nun gilt $\lim_{t\nearrow\infty} f(t,x) = 0$ mit integrierbarer Majorante e^{-x}, also $\lim_{t\nearrow\infty} F(t) = 0$ nach Satz 6.3 und somit $F(t) = \pi/2 - \arctan t$ für alle $t > 0$.

Weiter ist $\lim_{t\searrow 0} f(t,x) = \frac{\sin x}{x}$, aber hier gibt es *keine* Majorante, denn $\int_0^\infty \left|\frac{\sin x}{x}\right|\,dx = \infty$. Wir müssen also subtiler argumentieren. Für $t \geq 0$ und $0 < r_1 < r_2 < \infty$ gilt

$$\int_{r_1}^{r_2} e^{-tx}\frac{\sin x}{x}\,dx = \operatorname{Im}\int_{r_1}^{r_2} e^{(i-t)x}\frac{dx}{x} = \operatorname{Im}\left[\frac{e^{(i-t)x}}{(i-t)x}\right]_{x=r_1}^{x=r_2} + \operatorname{Im}\int_{r_1}^{r_2}\frac{e^{(i-t)x}}{(i-t)}\frac{dx}{x^2}.$$

Wegen $|i-t| \geq 1$ erhalten wir für $t \geq 0$ die Abschätzung

$$\left|\int_{r_1}^{r_2} e^{-tx}\frac{\sin x}{x}\,dx\right| \leq \frac{1}{r_1} + \frac{1}{r_2} + \int_{r_1}^{r_2}\frac{dx}{x^2} \leq \frac{2}{r_1}.$$

Mit dem Cauchykriterium folgt die Existenz des Integrals $F(0)$, und weiter

$$|F(0) - F(t)| \leq \underbrace{\left|\int_0^{r_1}(1-e^{-tx})\frac{\sin x}{x}\,dx\right|}_{\to 0 \text{ mit } t\searrow 0} + \frac{4}{r_1}.$$

Somit ist F stetig in $t = 0$, und es folgt schließlich

$$\int_0^\infty \frac{\sin x}{x}\,dx = F(0) = \lim_{t\searrow 0} F(t) = \lim_{t\searrow 0}(\pi/2 - \arctan t) = \pi/2.$$

Wir führen jetzt die L^p-Räume ein und zeigen, dass sie Banachräume sind. Dies ist die wohl wichtigste Konsequenz der Konvergenzsätze.

Definition 6.1 (L^p-Raum) *Für μ-messbares $f : X \to \overline{\mathbb{R}}$ und $1 \leq p \leq \infty$ setzen wir*

$$\|f\|_{L^p(\mu)} = \begin{cases} \left(\int |f|^p\,d\mu\right)^{1/p} & \text{für } 1 \leq p < \infty \\ \inf\left\{s > 0 : \mu(\{|f| > s\}) = 0\right\} & \text{für } p = \infty. \end{cases}$$

Auf $\mathcal{L}^p(\mu) = \{f : X \to \overline{\mathbb{R}} : f$ *μ-messbar,* $\|f\|_{L^p(\mu)} < \infty\}$ *betrachten wir die Äquivalenzrelation*

$$f \sim g \quad \Leftrightarrow \quad f(x) = g(x) \text{ für } \mu\text{-fast-alle } x \in X,$$

und definieren den L^p*-Raum durch* $L^p(\mu) = \mathcal{L}^p(\mu)/\sim$.

Wir schreiben $\|\cdot\|_{L^p}$ statt $\|\cdot\|_{L^p(\mu)}$, wenn sich das Maß aus dem Kontext ergibt. Es ist üblich, die Elemente von $L^p(\mu)$ wieder als Funktionen zu bezeichnen. Etwas allgemeiner kann auch der Begriff der L^p-Funktion auf einer messbaren Teilmenge definiert werden.

Definition 6.2 *Für* $E \subset X$ *μ-messbar und* $f : E \to \overline{\mathbb{R}}$ *sei* $f_0 : X \to \overline{\mathbb{R}}$ *die Fortsetzung mit* $f_0(x) = 0$ *für alle* $x \in X \backslash E$. *Wir setzen dann*

$$\mathcal{L}^p(E) = \{f : E \to \overline{\mathbb{R}} : f_0 \in \mathcal{L}^p(X)\},$$

und $L^p(E,\mu) = \mathcal{L}^p(E,\mu)/\sim$ *(analog zu Definition 6.1).*

Proposition 6.1 *Für* $1 \leq p \leq \infty$ *ist* $(L^p(\mu), \|\cdot\|_{L^p(\mu)})$ *ein normierter Vektorraum. Insbesondere gelten für* $\lambda \in \mathbb{R}$ *und* $f, g \in L^p(\mu)$ *folgende Aussagen:*

(1) $\|f\|_{L^p} = 0 \quad \Rightarrow \quad f = 0$ *μ-fast-überall.*

(2) $f \in L^p(\mu),\ \lambda \in \mathbb{R} \quad \Rightarrow \quad \lambda f \in L^p(\mu)$ *und* $\|\lambda f\|_{L^p} = |\lambda|\,\|f\|_{L^p}$.

(3) $f, g \in L^p(\mu) \quad \Rightarrow \quad f + g \in L^p(\mu)$ *und* $\|f + g\|_{L^p} \leq \|f\|_{L^p} + \|g\|_{L^p}$.

Beweis: Sei zunächst $1 \leq p < \infty$. Dann ist $\|f\|_{L^p}$ wohldefiniert nach Satz 5.4, und Folgerung 5.3 impliziert Aussage (1). Weiter folgt Behauptung (2) aus der Linearität des Integrals, siehe Satz 5.6. Da die Funktion $t \mapsto t^p$ auf $[0,\infty)$ konvex ist, gilt nun

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

Nach dem Majorantenkriterium ist mit $f, g \in L^p(\mu)$ also auch $f + g \in L^p(\mu)$. Die Dreiecksungleichung wird unten in Satz 6.8 mithilfe der Hölderschen Ungleichung gezeigt.

Im Fall $p = \infty$ ist (1) klar. Für (2) und (3) verwenden wir, wobei $\lambda > 0$ angenommen wird,

$$\{|\lambda f| > \lambda s\} = \{|f| > s\} \quad \text{und} \quad \{|f + g| > s_1 + s_2\} \subset \big(\{|f| > s_1\} \cup \{|g| > s_2\}\big).$$

Daraus ergibt sich leicht die Behauptung. □

Lemma 6.1 (Youngsche Ungleichung) *Für* $1 < p, q < \infty$ *mit* $\frac{1}{p} + \frac{1}{q} = 1$ *und* $x, y \geq 0$ *gilt*

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q}.$$

Beweis: Sei $y \geq 0$ fest und $f(x) = \frac{1}{p}x^p + \frac{1}{q}y^q - xy$. Dann gilt

$$f'(x) = x^{p-1} - y \begin{cases} < 0 & \text{für } x < y^{\frac{1}{p-1}} \\ > 0 & \text{für } x > y^{\frac{1}{p-1}} \end{cases}$$

Also gilt für alle $x \geq 0$

$$f(x) \geq f\big(y^{\frac{1}{p-1}}\big) = \frac{1}{p}\, y^{\frac{p}{p-1}} + \frac{1}{q}\, y^{\frac{p}{p-1}} - y^{\frac{p}{p-1}} = 0.$$

$\square$

Satz 6.7 (Höldersche Ungleichung) *Für μ-messbare $f, g : X \to \overline{\mathbb{R}}$ gilt*

$$\Big| \int f g \, d\mu \Big| \leq \|f\|_{L^p} \|g\|_{L^q}, \qquad \textit{falls } 1 \leq p, q \leq \infty \textit{ mit } \frac{1}{p} + \frac{1}{q} = 1.$$

Beweis: Wir können $\|f\|_{L^p} = \|g\|_{L^q} = 1$ und $f, g \geq 0$ annehmen. Aus Lemma 6.1 folgt

$$\int f g \, d\mu \leq \int \left(\frac{1}{p} f^p + \frac{1}{q} g^q \right) d\mu = 1 = \|f\|_{L^p} \, \|g\|_{L^q}.$$

Der Fall $p = 1$, $q = \infty$ ergibt sich direkt aus Satz 5.4. $\square$

Die Höldersche Ungleichung hat als Spezialfall für $p = q = 2$ die Ungleichung von Cauchy-Schwarz. Wir können nun den noch fehlenden Beweis der Dreiecksungleichung in $L^p(\mu)$ nachtragen.

Satz 6.8 (Minkowski-Ungleichung) *Für $f, g \in L^p(\mu)$ mit $1 \leq p \leq \infty$ gilt*

$$\|f + g\|_{L^p} \leq \|f\|_{L^p} + \|g\|_{L^p}.$$

Beweis: Indem wir im letzten Schritt die Höldersche Ungleichung anwenden, folgt

$$\begin{aligned}
\|f + g\|_{L^p(\mu)}^p &= \int |f + g|^p \, d\mu \\
&\leq \int |f| \, |f + g|^{p-1} \, d\mu + \int |g| \, |f + g|^{p-1} \, d\mu \\
&\leq \|f\|_{L^p} \|f + g\|_{L^p}^{p-1} + \|g\|_{L^p} \|f + g\|_{L^p}^{p-1}.
\end{aligned}$$

Kürzen liefert die Behauptung. $\square$

Das folgende Lemma stellt den wesentlichen Schritt im Beweis der Vollständigkeit von $L^p(\mu)$ dar.

Lemma 6.2 *Sei $1 \leq p < \infty$ und $f_k = \sum_{j=1}^k u_j$ mit $u_j \in L^p(\mu)$. Falls $\sum_{j=1}^\infty \|u_j\|_{L^p} < \infty$, so gelten folgende Aussagen:*

(1) *Es gibt eine μ-Nullmenge N, so dass $f(x) = \lim_{k \to \infty} f_k(x)$ für alle $x \in X \backslash N$ existiert.*

(2) *Mit $f := 0$ auf N gilt $f \in L^p(\mu)$.*

(3) $\|f - f_k\|_{L^p} \to 0$.

Beweis: Wir betrachten die Funktionen

$$g_k = \sum_{j=1}^k |u_j| \quad \text{und} \quad g = \sum_{j=1}^\infty |u_j|.$$

Es gilt $g_1 \leq g_2 \leq \ldots$ und $g_k(x) \to g(x) \in [0,\infty]$ mit $k \to \infty$ für alle $x \in X$. Aus dem Satz über monotone Konvergenz, Satz 5.5, und der Minkowski-Ungleichung folgt

$$\|g\|_{L^p} = \lim_{k\to\infty} \|g_k\|_{L^p} \leq \sum_{j=1}^{\infty} \|u_j\|_{L^p} < \infty.$$

Wegen Folgerung 5.3 ist $N := \{g = \infty\}$ eine μ-Nullmenge. Für $x \in X\backslash N$ ist die reelle Reihe $\sum_{j=1}^{\infty} u_j(x)$ absolut konvergent, also ist $\big(f_k(x)\big)_{k\in\mathbb{N}}$ eine Cauchyfolge in $\mathbb{R}$. Damit existiert der Grenzwert $f(x) = \lim_{k\to\infty} f_k(x)$, und es gilt weiter

$$|f_k|^p \leq g^p \in L^1(\mu) \quad \text{ sowie } |f - f_k|^p \leq 2^{p-1}\big(|f|^p + |f_k|^p\big) \leq 2^p\, g^p.$$

Der Satz über majorisierte Konvergenz, Satz 6.3, liefert $f \in L^p(\mu)$ und $\|f - f_k\|_{L^p} \to 0$. □

Satz 6.9 (Riesz-Fischer) *$(L^p(\mu), \|\cdot\|_{L^p})$ ist vollständig, also ein Banachraum.*

Beweis: Sei $f_k \in L^p(\mu)$ eine gegebene Cauchyfolge bezüglich der Norm $\|\cdot\|_{L^p}$. Es reicht aus, eine Teilfolge anzugeben, die in $L^p(\mu)$ konvergiert. Wir betrachten zuerst den Fall $1 \leq p < \infty$, und können nach evtl. Wahl einer Teilfolge annehmen:

$$\|f_{k+1} - f_k\|_{L^p} \leq 2^{-k} \text{ für alle } k \in \mathbb{N}.$$

Mit $f_0 := 0$ gilt $f_k = \sum_{j=1}^{k} u_j$ für $u_j = f_j - f_{j-1}$. Die Voraussetzungen von Lemma 6.2 sind erfüllt, also konvergiert f_k in $L^p(\mu)$ sowie punktweise μ-fast-überall gegen eine Funktion $f \in L^p(\mu)$. Dies beweist den Satz für $1 \leq p < \infty$.

Sei nun $p = \infty$. Wegen $|\,\|f_k\|_{L^\infty} - \|f_l\|_{L^\infty}| \leq \|f_k - f_l\|_{L^\infty}$ existiert $\lim_{k\to\infty} \|f_k\|_{L^\infty}$. Die folgenden Mengen haben μ-Maß Null:

$$N_k = \{|f_k| > \|f_k\|_{L^\infty}\} \quad \text{ sowie } \quad N_{k,l} = \{|f_k - f_l| > \|f_k - f_l\|_{L^\infty}\}.$$

Also ist $N = \bigcup_{k=1}^{\infty} N_k \cup \bigcup_{k,l=1}^{\infty} N_{k,l}$ ebenfalls eine Nullmenge. Für $x \in X\backslash N$ gilt nun

$$|f_k(x) - f_l(x)| \leq \|f_k - f_l\|_{L^\infty} < \varepsilon \quad \text{ für } k, l \geq k(\varepsilon),$$

insbesondere ist $f(x) = \lim_{k\to\infty} f_k(x)$ definiert. Weiter haben wir für $x \in X\backslash N$

$$|f(x)| = \lim_{k\to\infty} |f_k(x)| \leq \lim_{k\to\infty} \|f_k\|_{L^\infty}, \text{ und}$$

$$|f_k(x) - f(x)| = \lim_{l\to\infty} |f_k(x) - f_l(x)| \leq \varepsilon \text{ für } k \geq k(\varepsilon).$$

Dies bedeutet $\|f\|_{L^\infty} \leq \lim_{k\to\infty} \|f_k\|_{L^\infty} < \infty$ und $\|f_k - f\|_{L^\infty} \to 0$ mit $k \to \infty$. □

Ein Teilergebnis des Beweises, das oft benutzt wird, ist die

Folgerung 6.2 *Konvergiert f_k gegen f in $L^p(\mu)$, so konvergiert eine Teilfolge f_{k_j} punktweise μ-fast-überall gegen f.*

Auf die Wahl der Teilfolge in Folgerung 6.2 kann für $p < \infty$ im allgemeinen nicht verzichtet werden. Betrachte dazu folgendes

Beispiel 6.3 Jedes $n \in \mathbb{N}$ besitzt eine eindeutige Darstellung $n = 2^k + j$ mit $k \in \mathbb{N}_0$, $0 \leq j < 2^k$. Definiere damit eine Folge von Funktionen $f_n : [0,1] \to \mathbb{R}$ durch

$$f_n(x) = \begin{cases} 1 & \text{falls } j\,2^{-k} \leq x \leq (j+1)2^{-k} \\ 0 & \text{sonst.} \end{cases}$$

Es folgt $\int_0^1 f_n(x)\,dx = 2^{-k} < 2/n \to 0$ mit $n \to \infty$. Andererseits gilt $\limsup_{n\to\infty} f_n(x) = 1$ für alle $x \in [0,1)$, denn zu $x \in [0,1)$, $k \in \mathbb{N}$ können wir $j \in \{0,1,\ldots,2^k-1\}$ wählen mit $j\,2^{-k} \leq x < (j+1)\,2^{-k}$, also $f_n(x) = 1$ für $n = 2^k + j$. Also konvergiert die Folge nicht punktweise $\mathcal{L}^1$-fast-überall gegen Null.

Wir betrachten nun speziell das n-dimensionale Lebesguemaß. Im Gegensatz zu einem allgemeinen Maßraum haben wir auf $\mathbb{R}^n$ eine Metrik zur Verfügung. Eine naheliegende Frage ist dann, ob L^p-Funktionen durch stetige Funktionen approximiert werden können. Ziel ist eine Version für L^p-Funktionen auf einer offenen Teilmenge $\Omega \subset \mathbb{R}^n$, siehe Definition 6.2. Wie allgemein üblich schreiben wir kurz $L^p(\Omega)$ statt $L^p(\Omega, \mathcal{L}^n)$.

Definition 6.3 *Der Träger einer Funktion $f : \Omega \to \mathbb{R}$ ist die Menge*

$$\operatorname{spt} f = \overline{\{x \in \Omega : f(x) \neq 0\}}.$$

Der Raum der stetigen Funktionen mit kompaktem Träger in Ω wird mit $C^0_c(\Omega)$ bezeichnet.

Für $K \subset \Omega$ kompakt sei $\operatorname{dist}(\cdot, K) : \mathbb{R}^n \to [0,\infty)$, $\operatorname{dist}(x,K) = \inf_{z\in K} |x-z|$ die Abstandsfunktion von K. Wir brauchen die folgenden zwei Tatsachen:

$$\operatorname{dist}(\cdot, K) \text{ ist Lipschitzstetig mit Konstante Eins.} \tag{6.2}$$

$$\operatorname{dist}(\mathbb{R}^n\backslash\Omega, K) = \inf_{x\in\mathbb{R}^n\backslash\Omega} \operatorname{dist}(x,K) > 0. \tag{6.3}$$

Satz 6.10 (Dichtheit von $C^0_c(\Omega)$ in $L^p(\Omega)$) *Sei $\Omega \subset \mathbb{R}^n$ offen und $1 \leq p < \infty$. Dann gibt es zu jedem $f \in L^p(\Omega)$ eine Folge $f_k \in C^0_c(\Omega)$ mit $\|f - f_k\|_{L^p(\Omega)} \to 0$.*

Beweis: Wir zeigen die Aussage zuerst für $f = \chi_E$, wobei $E \subset \Omega$ eine $\mathcal{L}^n$-messbare Menge ist mit $\mathcal{L}^n(E) < \infty$. Nach Satz 4.3 existiert $K \subset E$ kompakt mit $\mathcal{L}^n(E\backslash K) < \varepsilon/2$. Setze

$$f_\varrho : \Omega \to [0,1],\ f_\varrho(x) = \left(1 - \frac{\operatorname{dist}(x,K)}{\varrho}\right)^+.$$

Nach (6.2), (6.3) ist $f_\varrho \in C^0(\mathbb{R}^n)$ und $\operatorname{spt} f_\varrho = \{x : \operatorname{dist}(x,K) \leq \varrho\}$ kompakte Teilmenge von Ω für $\varrho > 0$ hinreichend klein. Da $f_\varrho = f$ auf K, folgt

$$\begin{aligned} \int_\Omega |f_\varrho - f|^p\,d\mathcal{L}^n &\leq 2^{p-1} \int_{\Omega\backslash K} (|f_\varrho|^p + |f|^p)\,d\mathcal{L}^n \\ &\leq \mathcal{L}^n(\{0 < \operatorname{dist}(\cdot,K) \leq \varrho\}) + \mathcal{L}^n(E\backslash K) \\ &< \varepsilon \quad \text{für } \varrho > 0 \text{ hinreichend klein.} \end{aligned}$$

Sei nun $f \in L^p(\Omega)$ beliebig. Wir können $f \geq 0$ annehmen, sonst betrachte f^+ und f^-. Nach Satz 5.3 gibt es eine Folge $f_1 \leq f_2 \leq \ldots$ von nichtnegativen $\mathcal{L}^n$-Treppenfunktionen auf $\mathbb{R}^n$

mit $\lim_{k\to\infty} f_k(x) = f_0(x)$ für alle $x \in \mathbb{R}^n$; dabei bezeichnet f_0 die Fortsetzung von f mit $f \equiv 0$ auf $\mathbb{R}^n\backslash\Omega$. Es folgt $f_k \to f_0$ in $L^p(\Omega)$ mit dem Satz über majorisierte Konvergenz. Da jede Treppenfunktion endliche Linearkombination von charakteristischen Funktionen ist, folgt die Behauptung des Satzes. $\square$

Bemerkung 6.2 Es bezeichne $BC^0(\Omega)$ den Raum der beschränkten, stetigen Funktionen $f : \Omega \to \mathbb{R}$ auf der offenen Menge $\Omega \subset \mathbb{R}^n$. Für $f \in BC^0(\Omega)$ gilt

$$\|f\|_\Omega := \sup_{x\in\Omega} |f(x)| = \|f\|_{L^\infty(\Omega)},$$

denn für $s < \|f\|_\Omega$ ist $\{|f| > s\}$ nichtleer und offen, also ist $\mu(\{|f| > s\}) > 0$ und folglich $\|f\|_{L^\infty} \geq \|f\|_\Omega$; die umgekehrte Ungleichung ist trivial. Die Projektion von $BC^0(\Omega)$ nach $L^\infty(\Omega)$, die f seine fast-überall-Äquivalenzklasse zuordnet, ist damit isometrisch und injektiv. Nun ist $BC^0(\Omega)$ mit der Supremumsnorm ein Banachraum, vgl. Analysis II, Kapitel 10, Satz 1.2. Konvergiert $f_n \in BC^0(\Omega)$ bezüglich der L^∞-Norm gegen $f \in L^\infty(\Omega)$, so ist f_n eine Cauchyfolge bezüglich der Supremumsnorm und konvergiert gleichmäßig gegen eine Repräsentantin $\tilde{f} \in BC^0(\Omega)$ von f. Also ist $BC^0(\Omega)$ ein abgeschlossener, echter Unterraum von $L^\infty(\Omega)$, und insbesondere nicht dicht in $L^\infty(\Omega)$.

Die Entstehung des Konzepts des Lebesgueintegrals ist eng verknüpft mit dem Problem der Darstellbarkeit einer periodischen Funktion durch ihre Fourierreihe. Sei $L^2(I;\mathbb{C})$ der Raum der $\mathcal{L}^1$-messbaren Funktionen $f : I = (-\pi,\pi) \to \mathbb{C}$, also $u = \operatorname{Re} f$ und $v = \operatorname{Im} f$ sind $\mathcal{L}^1$-messbar, mit

$$\|f\|_{L^2} = \left(\int_{-\pi}^{\pi} |f(x)|^2\,dx\right)^{1/2} < \infty.$$

Dabei schreiben wir im folgenden $\int_{-\pi}^{\pi} f(x)\,dx$ statt $\int_I f(x)\,d\mathcal{L}^1(x)$. Nach Satz 6.9 ist $L^2(I;\mathbb{C})$ bezüglich der L^2-Norm vollständig, also ein Hilbertraum mit dem hermiteschen Skalarprodukt

$$\langle f, g\rangle_{L^2} = \int_{-\pi}^{\pi} f(x)\overline{g(x)}\,dx.$$

Die Funktionen $w_k(x) = \frac{1}{\sqrt{2\pi}} e^{ikx}$, $k \in \mathbb{Z}$, bilden ein Orthonormalsystem und spannen den Raum $\mathbb{P}$ der trigonometrischen Polynome auf. Das n-te Fourierpolynom von f ist definiert als die Orthogonalprojektion von f auf den Raum $\mathbb{P}_n$ der trigonometrischen Polynome vom Grad höchstens n, also

$$f_n = \sum_{k=-n}^{n} \langle f, w_k\rangle_{L^2} w_k = \sum_{k=-n}^{n} \hat{f}(k) e^{ikx}.$$

Dabei bezeichnet $\hat{f}(k)$ die Fourierkoeffizienten von f:

$$\hat{f}(k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-ikx}\,dx = \frac{1}{\sqrt{2\pi}}\langle f, w_k\rangle_{L^2}.$$

Die Folge f_n heißt Fourierreihe von f. Wegen $f - f_n \perp_{L^2} \mathbb{P}_n$ gilt für alle $p \in \mathbb{P}_n$ die Gleichung

$$\|f-p\|_{L^2}^2 = \|(f-f_n) + (f_n - p)\|_{L^2}^2 = \|f-f_n\|_{L^2}^2 + \|f_n - p\|_{L^2}^2. \tag{6.4}$$

Insbesondere folgt mit $p \equiv 0$ die Besselsche Ungleichung

$$(6.5) \qquad 2\pi \sum_{k=-\infty}^{\infty} |\hat{f}(k)|^2 = 2\pi \lim_{n\to\infty} \sum_{k=-n}^{n} |\hat{f}(k)|^2 = \lim_{n\to\infty} \|f_n\|_{L^2}^2 \leq \|f\|_{L^2}^2 < \infty.$$

Für $f \in L^2(I;\mathbb{C})$ und $m \geq n$ folgt

$$\|f_m - f_n\|_{L^2}^2 = 2\pi \sum_{k=n+1}^{m} |\hat{f}(k)|^2 < \varepsilon \quad \text{für } n \geq N,$$

das heißt f_n ist eine L^2-Cauchyfolge und konvergiert nach Satz 6.9, dem Satz von Riesz-Fischer, gegen eine gewisse L^2-Funktion. Die entscheidende Frage lautet: ist dieser Grenzwert die Funktion f? Wegen (6.4) haben wir die folgende Bestapproximations-Eigenschaft der Fourierreihe:

$$(6.6) \qquad \|f - f_n\|_{L^2} = \min_{p \in \mathbb{P}_n} \|f - p\|_{L^2}.$$

Die L^2-Konvergenz der Fourierreihe gegen f folgt somit aus nachstehendem Resultat:

Satz 6.11 *Der Raum $\mathbb{P}$ der trigonometrischen Polynome ist dicht in $L^2(I;\mathbb{C})$.*

Beweis: Nach Satz 6.10 ist $C_c^0(I;\mathbb{C})$ dicht in $L^2(I;\mathbb{C})$. Weiter kann jede C_c^0-Funktion bezüglich der L^2-Norm (sogar gleichmäßig) durch stückweise konstante Funktionen approximiert werden. Wir zeigen nun, dass die Fourierreihe f_n einer beliebigen stückweise konstanten Funktion f punktweise – bis auf eventuell in den Sprungstellen – gegen f konvergiert. Dann konvergiert f_n auch bezüglich der L^2-Norm gegen f, denn die Cauchyfolge f_n ist in $L^2(I;\mathbb{C})$ konvergent und nach Folgerung 6.2 kann der Grenzwert nur die Funktion f sein. Damit ist die Behauptung des Lemma dann bewiesen.

Die Funktion f sei mit Periode 2π auf $\mathbb{R}$ fortgesetzt. Nach Definition von f_n gilt für $x_0 \in I$

$$f_n(x_0) = \sum_{k=-n}^{n} \hat{f}(k)\, e^{ikx_0} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \sum_{k=-n}^{n} e^{-ik(x-x_0)}\, dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x_0+\xi) \sum_{k=-n}^{n} e^{-ik\xi}\, d\xi.$$

Dabei wurde $x = x_0 + \xi$ substituiert und die Periodizität ausgenutzt. Nun ist $(f_0)_n = f_0$ für die konstante Funktion $f_0(x) \equiv 1$, da $f_0 \in \mathbb{P}_n$. Es folgt

$$1 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{k=-n}^{n} e^{-ik\xi}\, d\xi.$$

Weiter berechnen wir für $\xi \in \mathbb{R}$

$$\sum_{k=-n}^{n} e^{-ik\xi} = e^{-in\xi} \sum_{k=0}^{2n} e^{ik\xi} = e^{-in\xi}\, \frac{e^{i(2n+1)\xi} - 1}{e^{i\xi} - 1} = \frac{e^{i(n+1)\xi} - e^{-in\xi}}{e^{i\xi} - 1}.$$

Danit erhalten wir die Darstellung

$$\begin{aligned} f_n(x_0) - f(x_0) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} (f(x_0+\xi) - f(x_0)) \sum_{k=-n}^{n} e^{-ik\xi}\, d\xi \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{f(x_0+\xi) - f(x_0)}{e^{i\xi} - 1}\, e^{i(n+1)\xi}\, d\xi - \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{f(x_0+\xi) - f(x_0)}{e^{i\xi} - 1}\, e^{-in\xi}\, d\xi. \end{aligned}$$

Ist f stückweise konstant und x_0 keine Sprungstelle von f, so verschwindet $f(x_0+\xi)-f(x_0)$ nahe bei $\xi = 0$. Auf der rechten Seite stehen dann die Fourierkoeffizienten $\hat{F}(-n-1)-\hat{F}(n)$ der beschränkten Funktion $F(\xi) = (f(x_0+\xi)-f(x_0))/(e^{i\xi}-1)$, und aus der Besselschen Ungleichung (6.5) folgt $\lim_{n\to\infty} f_n(x_0) = f(x_0)$. □

Es bezeichne nun $\ell^2(\mathbb{C})$ den Raum aller komplexen Folgen $c = (c_k)_{k\in\mathbb{Z}}$ mit

$$\|c\|_{\ell^2}^2 = 2\pi \sum_{k\in\mathbb{Z}} |c_k|^2 < \infty.$$

Der Raum $\ell^2(\mathbb{C})$ ist vollständig und damit ein Hilbertraum. Dies folgt aus dem Satz von Riesz-Fischer, angewandt auf des Zahlmaß auf $\mathbb{Z}$; es kann natürlich auch ad hoc gezeigt werden. Das bewiesene Resultat kann damit folgendermaßen formuliert werden:

Satz 6.12 (Vollständigkeit der trigonometrischen Polynome) *Für $f \in L^2(I;\mathbb{C})$ konvergiert die Fourierreihe bezüglich der L^2-Norm gegen f, und die Abbildung*

$$\mathcal{F} : (L^2(I;\mathbb{C}), \|\cdot\|_{L^2}) \longrightarrow (\ell^2(\mathbb{C}), \|\cdot\|_{\ell^2}), \quad \mathcal{F}(f) = (\hat{f}(k))_{k\in\mathbb{Z}},$$

ist eine (bijektive) Isometrie von Hilberträumen.

Beweis: Die Konvergenz $f_n \to f$ in $L^2(I;\mathbb{C})$ gilt nach Satz 6.11 und (6.6); insbesondere folgt hieraus die Injektivität der Abbildung $\mathcal{F}$. Die Surjektivität ergibt sich aus Satz 6.9, denn für $c \in \ell^2(\mathbb{C})$ ist die Folge $f_n = \sum_{k=-n}^{n} c_k e^{-ikx}$ eine Cauchyfolge in $L^2(I;\mathbb{C})$. Schließlich erhalten wir aus (6.4) mit $p \equiv 0$ und der gezeigten L^2-Konvergenz $f_n \to f$, vgl. (6.5),

$$2\pi \sum_{k=-\infty}^{\infty} |\hat{f}(k)|^2 = \lim_{n\to\infty} \|f_n\|_{L^2}^2 = \lim_{n\to\infty} \big(\|f\|_{L^2}^2 - \|f-f_n\|_{L^2}^2\big) = \|f\|_{L^2}^2.$$

□

Das Resultat ist ein Spezialfall des Spektralsatzes für selbstadjungierte Operatoren. Dieser verallgemeinert die aus der Linearen Algebra bekannte Diagonalisierbarkeit symmetrischer Matrizen, und wird in verschiedenen Varianten in der Funktionalanalysis bewiesen. Der Operator ist hier $-\frac{d^2}{dx^2}$, wir können ihn als Endomorphismus des Raums $C^\infty_{per}(I)$ dejenigen Funktionen auffassen, deren 2π-periodische Fortsetzung glatt ist:

$$H : C^\infty_{per}(I) \to C^\infty_{per}(I),\ Hf = -\frac{d^2 f}{dx^2}.$$

Mit partieller Integration zeigt man $\langle Hf, g\rangle_{L^2} = \langle f, Hg\rangle_{L^2}$ für alle $f,g \in C^\infty_{per}(I)$, sowie $\langle Hf, f\rangle_{L^2} \geq 0$. Die w_k sind Eigenfunktionen von H zu den Eigenwerten $\lambda_k = k^2$:

$$Hw_k = k^2 w_k \quad \text{für } k \in \mathbb{Z}.$$

Die w_k bilden eine Basis von $\mathbb{P}$, aber nicht von $C^\infty_{per}(I)$. Der Satz besagt, dass der von den Eigenfunktionen von H aufgespannte Raum $\mathbb{P}$ aber dicht in $L^2(I;\mathbb{C})$ ist.

Satz 6.13 (Egorov) *Sei μ ein Maß auf $\mathcal{A} \subset 2^X$ mit $\mu(X) < \infty$. Konvergiert die Folge f_k messbarer Funktionen punktweise μ-fast-überall gegen f, so gibt es zu $\delta > 0$ eine Menge $A \in \mathcal{A}$ mit $\mu(X\backslash A) < \delta$, so dass $f_k|_A$ gleichmäßig gegen $f|_A$ konvergiert.*

BEWEIS: Für jedes $\varepsilon > 0$ bilden die Mengen $C_j^\varepsilon = \bigcup_{k=j}^\infty \{|f - f_k| \geq \varepsilon\} \in \mathcal{A}$ eine absteigende Folge. Ist $f(x) = \lim_{k\to\infty} f_k(x)$, so gilt $x \notin C_j^\varepsilon$ für j groß, also folgt nach Satz 2.4(ii)

$$\lim_{j\to\infty} \mu(C_j^\varepsilon) = \mu\left(\bigcap_{j=1}^\infty C_j^\varepsilon\right) = 0.$$

Hier wurde die Bedingung $\mu(X) < \infty$ benutzt. Wähle nun eine Folge $\varepsilon_\nu \searrow 0$, und bestimme zu jedem $\nu \in \mathbb{N}$ ein $j_\nu \in \mathbb{N}$ mit $\mu(C_{j_\nu}^{\varepsilon_\nu}) < 2^{-\nu}\delta$, also

$$\mu\left(\bigcup_{\nu=1}^\infty C_{j_\nu}^{\varepsilon_\nu}\right) < \delta.$$

Mit $A := X \backslash \big(\bigcup_{\nu=1}^\infty C_{j_\nu}^{\varepsilon_\nu}\big) \in \mathcal{A}$ folgt $\sup_{x\in A} |f_k(x) - f(x)| \leq \varepsilon_\nu$ für $k \geq j_\nu$. $\square$

Satz 6.14 (Konvergenzsatz von Vitali) *Sei μ ein Maß auf $\mathcal{A} \subset 2^X$ und $1 \leq p < \infty$. Die Folge $f_n \in L^p(\mu)$ konvergiere punktweise μ-fast-überall gegen f. Dann sind folgende Aussagen (a) und (b) äquivalent:*

(a) *$f \in L^p(\mu)$ und $\|f_n - f\|_{L^p} \to 0$.*

(b) *Mit $\lambda(A) = \limsup_{n\to\infty} \int_A |f_n|^p \, d\mu$ gilt:*

  (1) *Zu $\varepsilon > 0$ gibt es ein $\delta > 0$ mit $\lambda(A) < \varepsilon$ für alle $A \in \mathcal{A}$ mit $\mu(A) < \delta$.*

  (2) *Zu $\varepsilon > 0$ gibt es ein $E \in \mathcal{A}$ mit $\mu(E) < \infty$ und $\lambda(X\backslash E) < \varepsilon$.*

Eine Folge mit (1) und (2) heißt gleichgradig integrierbar.

BEWEIS: Es gelte $f_n \to f$ in $L^p(\mu)$. Die Minkowski-Ungleichung ergibt dann für jedes $A \in \mathcal{A}$

$$\big|\|f_n\|_{L^p(A)} - \|f\|_{L^p(A)}\big| \leq \|f_n - f\|_{L^p(A)} \leq \|f_n - f\|_{L^p} \to 0,$$

also folgt $\lim_{n\to\infty} \|f\|_{L^p(A)} = \|f\|_{L^p(A)}$ und somit

$$\lambda(A) = \int_A |f|^p \, d\mu.$$

Eigenschaft (1) gilt nach Bemerkung 6.1. Mit $E_\delta = \{|f|^p > \delta\} \in \mathcal{A}$ gilt

$$\mu(E_\delta) \leq \frac{1}{\delta} \int |f|^p \, d\mu < \infty.$$

Für $\delta \searrow 0$ konvergiert $\chi_{X\backslash E_\delta}|f|^p$ punktweise gegen Null mit Majorante $|f|^p$, also folgt für $\delta > 0$ hinreichend klein

$$\lambda(X\backslash E_\delta) = \int_{X\backslash E_\delta} |f|^p \, d\mu < \varepsilon.$$

Seien jetzt umgekehrt (1) und (2) vorausgesetzt. Zu $\varepsilon > 0$ sei $E \in \mathcal{A}$ wie in (2) und $\delta > 0$ wie in (1) gewählt. Da $\mu(E) < \infty$, gibt es nach dem Satz von Egorov ein $A \in \mathcal{A}$ mit $A \subset E$ und $\mu(E\backslash A) < \delta$, so dass $f_n \to f$ gleichmäßig auf A. Nun gilt

$$|f_n - f|^p \leq \chi_A |f - f_n|^p + 2^{p-1}(\chi_{X\backslash E} + \chi_{E\backslash A})(|f|^p + |f_n|^p).$$

Ferner folgt aus dem Lemma von Fatou, Satz 6.2, für alle $B \in \mathcal{A}$

$$\int_B |f|^p \, d\mu \leq \liminf_{n\to\infty} \int_B |f_n|^p \, d\mu \leq \lambda(B).$$

Also erhalten wir mit $n \to \infty$ wegen (2) und (1)

$$\limsup_{n\to\infty} \|f_n - f\|_{L^p} \leq 2^p\big(\lambda(X\backslash E) + \lambda(E\backslash A)\big) < 2^{p+1}\varepsilon.$$

□

Bedingung (1) schließt eine Konzentration des Integrals wie in Beispiel 6.1 aus, und Bedingung (2) verhindert zum Beispiel eine Konzentration bei Unendlich wie im Beispiel

$$f_n : \mathbb{R} \to \mathbb{R},\ f_n = \chi_{[n,n+1]}.$$

7 Der Satz von Fubini

Mit dem Cavalierischen Prinzip kann das Volumen von Körpern durch Zerlegung in parallele Schnitte berechnet werden. Allgemeiner können mit dem Satz von Fubini mehrdimensionale Integrale bzw. allgemeiner Integrale in Produkträumen auf Integrationen in den Faktoren zurückgeführt werden. Gleichzeitig liefert der Satz eine optimale Aussage zur Vertauschbarkeit der Reihenfolge der Integrationen bei iterierten Integralen.

Die Idee der Volumenberechnung durch Zerlegung in parallele Schnitte wurde schon von Archimedes benutzt und ist in folgender Fassung als Cavalierisches Prinzip bekannt: gilt für die Schnitte in jeder Höhe y von zwei Körpern $K, L \subset \mathbb{R}^3$ das Verhältnis $\mathcal{L}^2(K_y) = \theta\,\mathcal{L}^2(L_y)$, so folgt auch $\mathcal{L}^3(K) = \theta\,\mathcal{L}^3(L)$. Mit diesem Argument lässt sich das Volumen einer Kugel folgendermaßen berechnen.

Beispiel 7.1 Betrachte in $\mathbb{R}^3 = \mathbb{R}^2 \times \mathbb{R}$ die folgenden Mengen und die Maße ihrer zweidimensionalen Schnitte in Höhe $y \in (0,1)$; dabei sei $\pi = \mathcal{L}^2(\{x \in \mathbb{R}^2 : |x| < 1\})$.

Zylinder:	$Z = \{(x,y) : \|x\| < 1,\, 0 < y < 1\}$	$\mathcal{L}^2(Z_y) = \pi$,
Kegel:	$K = \{(x,y) : \|x\| < y,\, 0 < y < 1\}$	$\mathcal{L}^2(K_y) = \pi\, y^2$,
Halbkugel:	$H = \{(x,y) : \|x\| < \sqrt{1-y^2},\, 0 < y < 1\}$	$\mathcal{L}^2(H_y) = \pi\,(1-y^2)$.

Durch Vergleich mit einem Quader gleicher Höhe und gleichen Querschnitts folgt aus dem Cavalierischen Prinzip $\mathcal{L}^3(Z) = \pi$. Betrachte weiter die Pyramide

$$P = \{y(x,1) : x \in (-1,1)^2,\, 0 < y < 1\}.$$

Sechs kongruente Exemplare setzen sich zu einem Würfel der Kantenlänge zwei zusammen, also gilt $6\,\mathcal{L}^3(P) = 8$. Wegen $\mathcal{L}^2(P_y) = 4y^2$ folgt aus dem Cavalierischen Prinzip $\mathcal{L}^3(K) = \frac{\pi}{4}\,\mathcal{L}^3(P) = \frac{\pi}{3}$. Schließlich erhalten wir, wieder mit Cavalieri,

$$\mathcal{L}^3(Z) = \mathcal{L}^3(K) + \mathcal{L}^3(H) \quad \Rightarrow \quad \mathcal{L}^3(H) = \frac{2\pi}{3}.$$

Insbesondere folgt $\mathcal{L}^3(K) : \mathcal{L}^3(H) : \mathcal{L}^3(Z) = 1 : 2 : 3$.

Wir wollen nun allgemein Produktmnaße einführen, wobei wir uns an der Konstruktion des Lebesguemaßes aus Kapitel 4 orientieren können. Wir arbeiten hier mit äußeren Maßen, für beliebige Maße verweisen wir auf das Buch von Elstrodt.

Definition 7.1 (Produktmenge) *Seien α_i äußere Maße auf X_i für $i = 1,\ldots,n$. Eine Menge $P \subset X_1 \times \ldots \times X_n$ heißt Produktmenge, wenn es α_i-messbare Mengen $A_i \subset X_i$ gibt mit $P = A_1 \times \ldots \times A_n$.*

Die Darstellung $P = A_1 \times \ldots \times A_n$ ist eindeutig, außer wenn $P = \emptyset$. In der nachfolgenden Definition des Produktonhalts λ ist die Vereinbarung $0 \cdot \infty = 0$ relevant.

Lemma 7.1 *Seien α_i äußere Maße auf X_i für $i = 1,\ldots,n$. Dann gilt:*

(1) *Das System $\mathcal{P}$ der Produktmengen ist ein Halbring.*

(2) *Die Funktion $\lambda : \mathcal{P} \to [0,\infty]$, $\lambda(A_1 \times \ldots \times A_n) = \prod_{i=1}^n \alpha_i(A_i)$, ist ein Prämaß auf $\mathcal{P}$.*

Beweis: Aussage (1) wurde in Folgerung 1.1 bewiesen. Ist $P = A_1 \times \ldots \times A_n = \emptyset$, so ist $A_i = \emptyset$ für mindestens ein i und folglich $\lambda(P) = 0$. Die σ-Additivität zeigen wir durch Induktion über n, ähnlich wie in Satz 1.2. Setze dazu $\lambda'(A_1 \times \ldots \times A_{n-1}) = \prod_{i=1}^{n-1} \alpha_i(A_i)$ für α_i-messbare Mengen $A_i \subset X_i$, $i = 1, \ldots, n-1$. Betrachte nun eine n-fache Produktmenge $P = A_1 \times \ldots \times A_n$. Für $y \in X_n$ ist der y-Schnitt von P gegeben durch

$$P_y = \{x \in X_1 \times \ldots \times X_{n-1} : (x,y) \in P\} = \begin{cases} A_1 \times \ldots \times A_{n-1} & \text{falls } y \in A_n \\ \emptyset & \text{sonst.} \end{cases}$$

Ist $P = \bigcup_{j=1}^\infty P_j$ paarweise disjunkte Vereinigung von Produktmengen P_j, so erhalten wir wegen $P_y = \bigcup_{j=1}^\infty (P_j)_y$ aus der Induktionsannahme mit dem Satz über monotone Konvergenz, Satz 5.5,

$$\lambda(P) = \int \lambda'(P_y)\, d\alpha_n(y) = \int \sum_{j=1}^\infty \lambda'((P_j)_y)\, d\alpha_n(y) = \sum_{j=1}^\infty \int \lambda'((P_j)_y)\, d\alpha_n(y) = \sum_{j=1}^\infty \lambda(P_j).$$

Dies beweist Aussage (2). $\square$

Definition 7.2 (Produktmaß) *Seien α_i äußere Maße auf X_i für $i = 1, \ldots, n$. Das Produktmaß $\alpha_1 \times \ldots \times \alpha_n$ einer Menge $E \subset X_1 \times \ldots \times X_n$ ist wie folgt definiert:*

$$(\alpha_1 \times \ldots \times \alpha_n)(E) = \inf\{\sum_{j=1}^\infty \lambda(P_j) : P_j \textit{ Produktmenge, } E \subset \bigcup_{j=1}^\infty P_j\}.$$

Damit ist $\alpha_1 \times \ldots \times \alpha_n$ die Caratheodory-Fortsetzung des Produktinhalts λ aus Lemma 7.1.

Das Lebesguemaß auf $\mathbb{R}^n = \mathbb{R}^{n_1} \times \ldots \times \mathbb{R}^{n_k}$ ist ein wichtiges Beispiel für ein Produktmaß.

Lemma 7.2 *Es gilt $\mathcal{L}^n = \mathcal{L}^{n_1} \times \ldots \times \mathcal{L}^{n_k}$ für $n = n_1 + \ldots + n_k$.*

Beweis: Jeder Quader $P \subset \mathbb{R}^n$ ist das Produkt von Quadern $P_j \subset \mathbb{R}^{n_j}$ mit Inhalt $|P| = \mathcal{L}^{n_1}(P_1) \cdot \ldots \cdot \mathcal{L}^{n_k}(P_k)$. Die Ungleichung $\mathcal{L}^{n_1} \times \ldots \times \mathcal{L}^{n_k} \leq \mathcal{L}^n$ folgt damit aus den Definitionen 4.1 und 7.2. Umgekehrt ist zu zeigen, dass für jede Überdeckung $E \subset \bigcup_{i=1}^\infty A_{i,1} \times \ldots \times A_{i,k}$ einer Menge $E \subset \mathbb{R}^n$ mit $\mathcal{L}^{n_j}$-messbaren Mengen $A_{i,j} \subset \mathbb{R}^{n_j}$ gilt:

$$\mathcal{L}^n(E) \leq \sum_{i=1}^\infty \mathcal{L}^{n_1}(A_{i,1}) \cdot \ldots \cdot \mathcal{L}^{n_k}(A_{i,k}).$$

Indem wir aus E eine $\mathcal{L}^n$-Nullmenge entfernen, können wir $\mathcal{L}^n(A_{i,1} \times \ldots \times A_{i,k}) > 0$ für alle $i \in \mathbb{N}$ annehmen. Zu $\theta > 1$ gibt es dann Quader $P_{i,j}^\nu \subset \mathbb{R}^{n_j}$ mit $A_{i,j} \subset \bigcup_{\nu=1}^\infty P_{i,j}^\nu$, so dass gilt:

$$\sum_{\nu=1}^\infty |P_{i,j}^\nu| \leq \theta\, \mathcal{L}^{n_j}(A_{i,j}) \quad \text{für alle } i \in \mathbb{N} \text{ und } j = 1, \ldots, k.$$

Da $A_{i,1} \times \ldots A_{i,k} \subset \bigcup_{\nu_1,\ldots,\nu_k=1}^\infty P_{i,1}^{\nu_1} \times \ldots \times P_{i,k}^{\nu_k}$, folgt die Bahauptung aus der Abschätzung

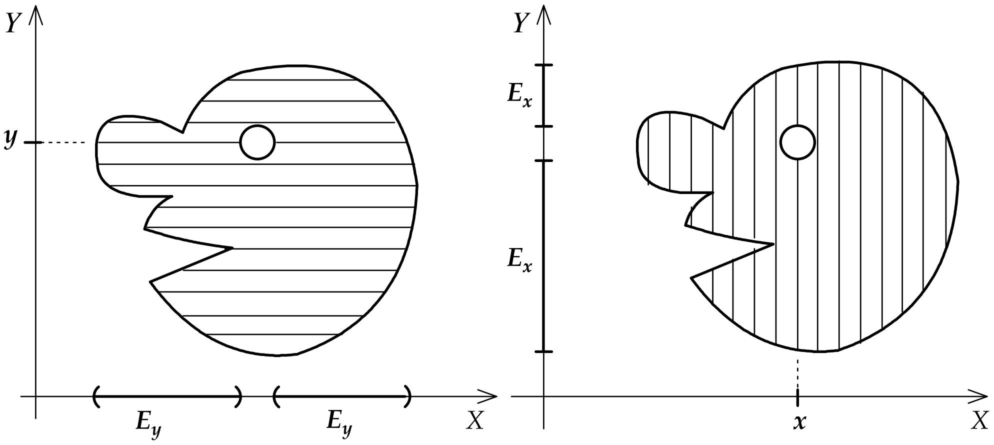
$$\begin{aligned} \sum_{i=1}^\infty \sum_{\nu_1,\ldots,\nu_k=1}^\infty |P_{i,1}^{\nu_1} \times \ldots \times P_{i,k}^{\nu_k}| &= \sum_{i=1}^\infty \left(\sum_{\nu_1=1}^\infty |P_{i,1}^{\nu_1}|\right) \cdot \ldots \cdot \left(\sum_{\nu_k=1}^\infty |P_{i,k}^{\nu_k}|\right) \\ &\leq \theta^k \sum_{i=1}^\infty \mathcal{L}^{n_1}(A_{i,1}) \cdot \ldots \cdot \mathcal{L}^{n_k}(A_{i,k}). \end{aligned}$$

$\square$

Satz 7.1 (Cavalierisches Prinzip) *Seien α und β σ-endliche äußere Maße auf X bzw. Y, und $D \subset X \times Y$ sei $\alpha \times \beta$-messbar. Dann ist $D_y = \{x \in X : (x,y) \in D\}$ α-messbar für β-fast-alle $y \in Y$, die Funktion $y \mapsto \alpha(D_y)$ ist β-messbar und es gilt*

$$(\alpha \times \beta)(D) = \int_Y \alpha(D_y)\, d\beta(y) = \int_Y \int_X \chi_D(x,y)\, d\alpha(x)\, d\beta(y).$$

Eine entsprechende Aussage gilt, wenn erst das β-Maß von $D_x = \{y \in Y : (x,y) \in D\}$ gebildet und dann bezüglich α über X integriert wird.



Beweis: Wir zeigen den Satz erst für $(\alpha \times \beta)(D) < \infty$. Nach Satz 3.2 gibt es eine Menge $E \supset D$ mit $E \in \sigma(\mathcal{P})$ und $(\alpha \times \beta)(E \backslash D) = 0$. Genauer besagt der dortige Zusatz: es kann $E = \bigcap_{\nu=1}^\infty E^\nu$ gewählt werden, wobei $E^\nu = \bigcup_{i=1}^\infty P_\nu^i$ paarweise disjunkte Vereinigung von Produktmengen P_i^ν mit $E^1 \supset E^2 \supset \ldots$ und $(\alpha \times \beta)(E^1) < \infty$. Wir wollen zunächst folgende drei Aussagen beweisen:

(7.1) $\quad E_y = \{x \in X : (x,y) \in E\}$ ist α-messbar für alle $y \in Y$,

(7.2) $\quad$ die Funktion $f_E : Y \to [0,\infty]$, $y \mapsto \alpha(E_y)$ ist β-messbar,

(7.3) $\quad \gamma(E) := \int_Y f_E\, d\beta = (\alpha \times \beta)(E)$.

Sei $\mathcal{E}$ das System aller $\alpha \times \beta$-messbaren Mengen mit (7.1), (7.2) und (7.3). Für eine Produktmenge $A \times B$ gilt $(A \times B)_y = A$, falls $y \in B$, und $(A \times B)_y = \emptyset$ sonst. Daraus folgt $A \times B \in \mathcal{E}$, und zwar genauer

$$f_{A \times B} = \alpha(A)\, \chi_B \quad \text{und } \gamma(A \times B) = \alpha(A)\beta(B) = (\alpha \times \beta)(A \times B).$$

Als nächstes betrachte eine disjunkte Vereinigung $E = \bigcup_{i=1}^\infty E_i$ mit $E_i \in \mathcal{E}$. Dann ist $E_y = \bigcup_{i=1}^\infty (E_i)_y$ α-messbar, $f_E = \sum_{i=1}^\infty f_{E_i}$ ist β-messbar, und aus dem Satz über monotone Konvergenz folgt

$$\gamma(E) = \int_Y \sum_{i=1}^\infty f_{E_i}\, d\beta = \sum_{i=1}^\infty \int_Y f_{E_i}\, d\beta = \sum_{i=1}^\infty (\alpha \times \beta)(E_i) = (\alpha \times \beta)(E).$$

Schließlich sei $E^1 \supset E^2 \supset \ldots$ eine absteigende Folge mit $E^\nu \in \mathcal{E}$ und $(\alpha \times \beta)(E^1) < \infty$. Für $E = \bigcap_{\nu=1}^\infty E^\nu$ ist $E_y = \bigcap_{\nu=1}^\infty (E^\nu)_y$ α-messbar für alle $y \in Y$, und wegen $\int_Y f_{E^1} d\beta = (\alpha \times \beta)(E^1) < \infty$ gilt $\alpha((E^1)_y) = f_{E^1}(y) < \infty$ für β-fast-alle $y \in Y$. Aus Satz 2.4 folgt

$$f_E(y) = \alpha(E_y) = \lim_{\nu \to \infty} \alpha((E^\nu)_y) = \lim_{\nu \to \infty} f_{E^\nu}(y) \quad \text{für } \beta\text{-fast-alle } y \in Y.$$

Insbesondere ist f_E β-messbar, und wegen $f_{E^\nu} \leq f_{E^1}$ liefert der Satz von Lebesgue

$$\gamma(E) = \int_Y f_E \, d\beta = \lim_{\nu\to\infty} \int_Y f_{E^\nu} \, d\beta = \lim_{\nu\to\infty} (\alpha \times \beta)(E^\nu) = (\alpha \times \beta)(E).$$

Insgesamt sind nun (7.1), (7.2) und (7.3) für E wie oben gezeigt. Durch Anwendung auf eine $\alpha \times \beta$-Nullmenge N statt D erhalten wir eine Menge $C \supset N$ mit $C \in \mathcal{E}$ und $(\alpha \times \beta)(C) = 0$, also folgt

$$0 = (\alpha \times \beta)(C) = \int_Y \alpha(C_y) \, d\beta(y) \quad \Rightarrow \quad \alpha(N_y) \leq \alpha(C_y) = 0 \text{ für } \beta\text{-fast-alle } y \in Y.$$

Daraus folgt mit $N = E \backslash D$ für β-fast-alle $y \in Y$: die Menge $D_y = E_y \backslash N_y$ ist α-messbar und es gilt $f_D(y) = f_E(y)$. Insbesondere ist f_D β-messbar und

$$\int_Y f_D \, d\beta = \int_Y f_E \, d\beta = (\alpha \times \beta)(E) = (\alpha \times \beta)(D).$$

Ist D lediglich $\alpha \times \beta$-messbar, so verwenden wir die Vorausetzung an α und β. Nach Lemma 3.3 gilt $D = \bigcup_{n=1}^\infty D_n$ für $\alpha \times \beta$-messbare D_n mit $(\alpha \times \beta)(D_n) < \infty$ und $D_1 \subset D_2 \subset \ldots$. Also ist $D_y = \bigcup_{n=1}^\infty (D_n)_y$ messbar bezüglich α, es gilt $f_{D_1} \leq f_{D_2} \leq \ldots$ und

$$f_D(y) = \alpha(D_y) = \lim_{n\to\infty} \alpha((D_n)_y) = \lim_{n\to\infty} f_{D_n}(y).$$

Folglich ist f_D messbar bzgl. β und der Satz über monotone Konvergenz impliziert

$$\int_Y f_D \, d\beta = \lim_{n\to\infty} \int_Y f_{D_n} \, d\beta = \lim_{n\to\infty} (\alpha \times \beta)(D_n) = (\alpha \times \beta)(D).$$

Damit ist der Satz bewiesen. $\square$

Das folgende Beispiel zeigt, dass auf die Bedingung der σ-Endlichkeit der Maße im allgemeinen nicht verzichtet werden kann.

Beispiel 7.2 Für $D = \{(x,y) \in [0,1] \times [0,1] : x = y\} \subset \mathbb{R} \times [0,1]$ gilt

$$\int_{\mathbb{R}} \operatorname{card}(D_x) \, d\mathcal{L}^1(x) = 1 \neq 0 = \int_{[0,1]} \mathcal{L}^1(D_y) \, d\operatorname{card}(y).$$

Mit $I_k = [\frac{k-1}{n}, \frac{k}{n}]$ gilt $D = \bigcap_{n=1}^\infty \left(\bigcup_{k=1}^n I_k \times I_k\right)$, also ist D messbar bezüglich $\mathcal{L}^1 \times \operatorname{card}$. Aus Satz 7.1 folgt $(\mathcal{L}^1 \times \operatorname{card})(D) = \infty$.

Wir kommen nun zu Anwendungen des Cavalierischen Prinzips.

Beispiel 7.3 Wir berechnen das Lebesguemaß $\alpha_n = \mathcal{L}^n(B_1(0))$ der n-dimensionalen Kugel $B_1(0) = \{z \in \mathbb{R}^n : |z| < 1\}$. Aus Lemma 7.2 und Satz 7.1 folgt mit der Substitution $y = \cos\vartheta$

$$\alpha_n = \int_{-1}^1 \mathcal{L}^{n-1}(\{x \in \mathbb{R}^{n-1} : |x| < \sqrt{1-y^2}\}) \, dy = \alpha_{n-1} \int_{-1}^1 (1-y^2)^{\frac{n-1}{2}} \, dy = \alpha_{n-1} \underbrace{\int_0^\pi \sin^n \vartheta \, d\vartheta}_{=:A_n}.$$

Durch partielle Integration ergibt sich die Rekursionsformel $A_n = \frac{n-1}{n} A_{n-2}$ für $n \geq 2$, wobei $A_0 = \pi$ und $A_1 = 2$. Es folgt

$$A_{2k} = \frac{2k-1}{2k} \cdot \ldots \cdot \frac{1}{2} \cdot A_0 = \pi \prod_{j=1}^{k} \frac{2j-1}{2j} \quad \text{und} \quad A_{2k+1} = \frac{2k}{2k+1} \cdot \ldots \cdot \frac{2}{3} \cdot A_1 = 2 \prod_{j=1}^{k} \frac{2j}{2j+1}.$$

Also gilt $A_{2k+1} A_{2k} = \frac{2\pi}{2k+1}$ bzw. $A_{2k} A_{2k-1} = \frac{\pi}{k}$, und somit

$$\begin{aligned} \alpha_{2k} &= (A_{2k} A_{2k-1}) \cdot \ldots \cdot (A_2\, A_1)\alpha_0 = \frac{\pi^k}{k!}, \\ \alpha_{2k+1} &= (A_{2k+1} A_{2k}) \cdot \ldots \cdot (A_3\, A_2)\alpha_1 = \frac{\pi^k}{(k+\frac{1}{2})(k-\frac{1}{2}) \cdot \ldots \cdot \frac{1}{2}}. \end{aligned}$$

Beispiel 7.4 Für $A \subset \mathbb{R}^n$ sei $K(A) = \{y(x,1) \in \mathbb{R}^n \times \mathbb{R} = \mathbb{R}^{n+1} : 0 < y < 1,\, x \in A\}$; dies ist der Kegel mit Spitze 0 über der Basis $A \times \{1\} \subset \mathbb{R}^{n+1}$. Wir behaupten: Ist A messbar bzgl. $\mathcal{L}^n$, so ist $K(A)$ messbar bzgl. $\mathcal{L}^{n+1}$ und es gilt

$$\mathcal{L}^{n+1}(K(A)) = \frac{1}{n+1}\, \mathcal{L}^n(A).$$

Dazu beachten wir zunächst $K(\mathbb{R}^n) = \mathbb{R}^n \times (0,1)$ sowie

$$K\Big(\bigcup_{i=1}^{\infty} A_i\Big) = \bigcup_{i=1}^{\infty} K(A_i) \quad \text{und} \quad K(\mathbb{R}^n \backslash A) = (\mathbb{R}^n \times (0,1)) \backslash K(A).$$

Also ist das System $\{A \subset \mathbb{R}^n : K(A) \text{ ist Borelmenge}\}$ eine σ-Algebra. Es ist leicht zu sehen, dass für $\Omega \subset \mathbb{R}^n$ offen auch $K(\Omega)$ offen in $\mathbb{R}^{n+1}$ ist. Also ist für jede Borelmenge $B \subset \mathbb{R}^n$ der Kegel $K(B)$ eine Borelmenge, insbesondere $\mathcal{L}^{n+1}$-messbar, und aus Satz 7.1 folgt

$$\mathcal{L}^{n+1}(K(B)) = \int_0^1 \mathcal{L}^n(\{yx : x \in B\})\, dy = \int_0^1 y^n\, \mathcal{L}^n(B)\, dy = \frac{1}{n+1}\, \mathcal{L}^n(B).$$

Ist nun $A \subset \mathbb{R}^n$ lediglich $\mathcal{L}^n$-messbar, so gibt es nach Satz 4.2 Borelmengen $B_{1,2} \subset \mathbb{R}^n$ mit $B_1 \subset A \subset B_2$ und $\mathcal{L}^n(B_2 \backslash B_1) = 0$. Es folgt $K(B_1) \subset K(A) \subset K(B_2)$ und $\mathcal{L}^{n+1}(K(B_2) \backslash K(B_1)) = \mathcal{L}^{n+1}(K(B_2 \backslash B_1)) = 0$. Also ist $K(A)$ $\mathcal{L}^{n+1}$-messbar und die Formel für das Volumen des Kegels gilt auch für A.

Der folgende Satz über Integrale auf Produkträumen besitzt zahlreiche Anwendungen.

Satz 7.2 (Fubini) *Seien α und β σ-endliche äußere Maße auf X bzw. Y und $f : X \times Y \to \overline{\mathbb{R}}$ sei $\alpha \times \beta$-messbar. Ist das Integral $\int f\, d(\alpha \times \beta)$ definiert, so gilt:*

$f(\cdot, y)$ *ist α-messbar und* $\int_X f(x,y)\, d\alpha(x)$ *ist definiert für β-fast-alle* $y \in Y$,
$f(x, \cdot)$ *ist β-messbar und* $\int_Y f(x,y)\, d\beta(y)$ *ist definiert für α-fast-alle* $x \in X$,

$y \mapsto \int_X f(x,y)\, d\alpha(x)$ *ist β-messbar und* $\int_Y \int_X f(x,y)\, d\alpha(x)\, d\beta(y)$ *ist definiert,*
$x \mapsto \int_Y f(x,y)\, d\beta(y)$ *ist α-messbar und* $\int_X \int_Y f(x,y)\, d\beta(y)\, d\alpha(x)$ *ist definiert,*

und vor allem

$$\int_{X \times Y} f\, d(\alpha \times \beta) = \int_Y \int_X f(x,y)\, d\alpha(x)\, d\beta(y) = \int_X \int_Y f(x,y)\, d\beta(y)\, d\alpha(x).$$

Zusatz: *Für $\alpha \times \beta$-messbares f ist die Voraussetzung erfüllt, wenn $f \geq 0$ oder wenn*

$$\int_Y \int_X |f(x,y)|\, d\alpha(x)\, d\beta(y) < \infty \text{ bzw. } \int_X \int_Y |f(x,y)|\, d\beta(y)\, d\alpha(x) < \infty.$$

Beweis: Sei $\varphi \in \mathcal{T}^+(\alpha \times \beta)$ eine Treppenfunktion mit Funktionswerten $\{s_1, \ldots, s_k\}$. Dann gilt mit $E_i = \{\varphi = s_i\}$

$$\varphi(\cdot, y) = \sum_{i=1}^k s_i\, \chi_{(E_i)_y} \text{ und } \int_X \varphi(x,y)\, d\alpha(x) = \sum_{i=1}^k s_i\, \alpha((E_i)_y).$$

Für φ folgen alle Behauptungen damit aus Satz 7.1. Betrachte als nächstes eine $\alpha\times\beta$-messbare Funktion $f : X \times Y \to [0,\infty]$. Nach Satz 5.3 gibt es eine Folge von Treppenfunktionen $0 \leq \varphi_1 \leq \varphi_2 \leq \ldots$ mit $\varphi_k(x,y) \nearrow f(x,y)$ für alle $(x,y) \in X \times Y$. Dann gilt $\varphi_k(\cdot,y) \nearrow f(\cdot,y)$ für alle $y \in Y$, insbesondere ist $f(\cdot,y)$ messbar bezüglich α. Weiter liefert der Satz über monotone Konvergenz $\int_X \varphi_k(x,y)\, d\alpha(x) \nearrow \int_X f(x,y)\, d\alpha(x)$ für alle $y \in Y$. Folglich ist die Funktion $y \mapsto \int_X f(x,y)\, d\alpha(x)$ messbar bezüglich β, und es folgt wieder mit dem Satz über monotone Konvergenz

$$\begin{aligned}
\int_Y \int_X f(x,y)\, d\alpha(x)\, d\beta(y) &= \lim_{k\to\infty} \int_Y \int_X \varphi_k(x,y)\, d\alpha(x)\, d\beta(y) \\
&= \lim_{k\to\infty} \int_{X\times Y} \varphi_k\, d(\alpha \times \beta) \\
&= \int_{X\times Y} f\, d(\alpha\times\beta).
\end{aligned}$$

Da sich bei vertauschter Reihenfolge der Integrationen bzgl. x und y ganz analog argumentieren lässt, ist der Satz für $f \geq 0$ bewiesen. Sei schließlich nur das Integral von f definiert, also zum Beispiel $\int f^-\, d(\alpha\times\beta) < \infty$. Wie gezeigt gilt dann

$$\int_Y \int_X f^-(x,y)\, d\alpha(x)\, d\beta(y) = \int_X \int_Y f^-(x,y)\, d\beta(y)\, d\alpha(x) = \int_{X\times Y} f^-\, d(\alpha\times\beta) < \infty,$$

und es gibt eine β-Nullmenge $N \subset Y$ mit $\int_X f^-(x,y)\, d\alpha(x) < \infty$ für alle $y \in Y \backslash N$. Durch Abänderung zu $f = 0$ auf $X \times N$ können wir $N = \emptyset$ annehmen; beachte, dass hierdurch die Voraussetzungen und Behauptungen des Satzes nicht berührt werden. Folglich ist das Integral $\int_X f(x,y)\, d\alpha(x)$ für alle $y \in Y$ definiert, die Funktion

$$y \mapsto \int_X f(x,y)\, d\alpha(x) = \int_X f^+(x,y)\, d\alpha(x) - \int_X f^-(x,y)\, d\alpha(x)$$

ist messbar bezüglich β, und es gilt

$$\int_Y \left(\int_X f(x,y)\, d\alpha(x) \right)^- d\beta(y) \leq \int_Y \int_X f^-(x,y)\, d\alpha(x)\, d\beta(y) < \infty.$$

Somit ist auch das Integral der Funktion $y \mapsto \int_X f(x,y)\, d\alpha(x)$ definiert, und aus der Linearität des Integrals, Satz 5.6, folgt

$$\begin{aligned}
\int_Y \int_X f(x,y)\, d\alpha(x)\, d\beta(y) &= \int_Y \int_X (f^+(x,y) - f^-(x,y))\, d\alpha(x)\, d\beta(y) \\
&= \int_Y \int_X f^+(x,y)\, d\alpha(x)\, d\beta(y) - \int_Y \int_X f^-(x,y)\, d\alpha(x)\, d\beta(y) \\
&= \int_{X\times Y} f^+\, d(\alpha\times\beta) - \int_{X\times Y} f^-\, d(\alpha\times\beta) \\
&= \int_{X\times Y} f\, d(\alpha\times\beta).
\end{aligned}$$

Damit ist der Satz von Fubini bewiesen. Der Zusatz ist trivial für $f \geq 0$. Andernfalls gilt, wenn zum Beispiel das erste Integral endlich ist,

$$\int_{X\times Y} |f|\, d(\alpha\times\beta) = \int_Y \int_X |f(x,y)|\, d\alpha(x)\, d\beta(y) < \infty,$$

und f ist integrierbar bezüglich $\alpha\times\beta$ nach dem Majorantenkriterium, Satz 5.7. $\square$

Beispiel 7.5 Ein Beispiel, wo die iterierten Integrale existieren aber verschieden sind, ist

$$\int_{-1}^{1} \int_{-1}^{1} \frac{x^2-y^2}{(x^2+y^2)^2}\, dy\, dx = -\int_{-1}^{1} \int_{-1}^{1} \frac{x^2-y^2}{(x^2+y^2)^2}\, dx\, dy = \pi.$$

Beachte dabei, dass der Integrand gegeben ist durch $f(x,y) = \frac{\partial}{\partial x}\frac{\partial}{\partial y} \arctan\frac{x}{y}$ für $y \neq 0$. In diesem Fall folgt aus dem Satz von Fubini, dass das Integral bezüglich des Produktmaßes nicht existiert. Es kommt aber auch vor, dass die iterierten Integrale gleich sind, und dennoch das Integral bezüglich des Produktmaßes nicht definiert ist:

$$\int_{-1}^{1} \int_{-1}^{1} \frac{xy}{(x^2+y^2)^2}\, dx\, dy = \int_{-1}^{1} \int_{-1}^{1} \frac{xy}{(x^2+y^2)^2}\, dy\, dx = 0,$$

aber das $\mathcal{L}^2$-Integral über $[-1,1]^2$ existiert nicht wegen

$$\int_{[0,1]^2} \frac{xy}{(x^2+y^2)^2}\, d\mathcal{L}^2(x,y) = \int_0^1 \int_0^1 \frac{xy}{(x^2+y^2)^2}\, dx\, dy == \frac{1}{2}\int_0^1 \left(\frac{1}{y} - \frac{y}{1+y^2}\right) dy = \infty.$$

Beispiel 7.6 Sei μ ein σ-endliches äußeres Maß auf X und $f : X \to [0,\infty]$ sei μ-messbar. Ist die Funktion $\varphi : [0,\infty] \to [0,\infty]$ stetig mit $\varphi(0) = 0$, sowie auf $(0,\infty)$ stetig differenzierbar mit $\varphi'(t) \geq 0$, so gilt

$$\int_X \varphi(f(x))\, d\mu(x) = \int_0^\infty \varphi'(t)\, \mu(\{x \in X : f(x) > t\})\, dt.$$

Betrachte dazu den Subgraph $E = \{(x,t) \in X\times[0,\infty) : t < f(x)\}$. Die Funktionen $(x,t) \mapsto t$, $(x,t) \mapsto f(x)$ und $(x,t) \mapsto \varphi'(t)$ sind $\mu\times\mathcal{L}^1$-messbar. Folglich ist auch E messbar bezüglich $\mu\times\mathcal{L}^1$ sowie die Funktion $(x,t) \mapsto \varphi'(t)\chi_E(x,t)$. Aus dem Hauptsatz der Differential- und

Integralrechnung und dem Satz von Fubini folgt

$$\begin{aligned}
\int_X \varphi(f(x))\,d\mu(x) &= \int_X \int_0^\infty \varphi'(t)\,\chi_E(x,t)\,dt\,d\mu(x)\\
&= \int_E \varphi'(t)\,d(\mu\times\mathcal{L}^1)(x,t)\\
&= \int_0^\infty \int_X \varphi'(t)\,\chi_E\,(x,t)\,d\mu(x)\,dt\\
&= \int_0^\infty \varphi'(t)\,\mu(\{x\in X : f(x)>t\})\,dt.
\end{aligned}$$

Für $\varphi(t)=t^p$ mit $p>0$ und $f:X\to\mathbb{R}$ μ-messbar folgt zum Beispiel

$$\int_X |f|^p\,d\mu = \int_0^\infty p\,t^{p-1}\mu(\{x : |f(x)|>t\})\,dt.$$

Eine wichtige Anwendung des Satzes von Fubini ist die partielle Integration.

Satz 7.3 (Partielle Integration) *Sind für $f,g\in C^1(\mathbb{R}^n)$ die Funktionen $(\partial_j f)g$, $f(\partial_j g)$ und fg integrierbar, so folgt*

$$\int_{\mathbb{R}^n} (\partial_j f)g\,dx = -\int_{\mathbb{R}^n} f(\partial_j g)\,dx.$$

BEWEIS: Es reicht aus, die folgende Behauptung zu beweisen:

$$\int_{\mathbb{R}^n} \partial_j f\,dx = 0 \quad \text{für alle } f\in C^1(\mathbb{R}^n) \text{ mit } f,\partial_j f\in L^1(\mathbb{R}^n). \tag{7.4}$$

Denn aus der Voraussetzung und der Hölderschen Ungleichung folgt $fg\in L^1(\mathbb{R}^n)$ sowie $\partial_j(fg)=(\partial_j f)g+f(\partial_j g)\in L^1(\mathbb{R}^n)$, also ergibt sich mit (7.4)

$$0=\int_{\mathbb{R}^n}\partial_j(fg)\,dx=\int_{\mathbb{R}^n}(\partial_j f)g\,dx+\int_{\mathbb{R}^n} f(\partial_j g)\,dx,$$

und die Folgerung ist bewiesen. Wir zeigen (7.4) zunächst unter der Annahme, dass es ein $R>0$ gibt mit $f(x)=0$ für $|x_j|\geq R$. Der Satz von Fubini ergibt für $x=(x',x_n)\in\mathbb{R}^{n-1}\times\mathbb{R}$

$$\int_{\mathbb{R}^n}\partial_j f(x)\,dx=\int_{\mathbb{R}}\int_{\mathbb{R}^{n-1}}\partial_j f(x',x_n)\,dx'\,dx_n=\int_{\mathbb{R}^{n-1}}\int_{\mathbb{R}}\partial_j f(x',x_n)\,dx_n\,dx'.$$

Für $j=n$ ist das rechte Integral Null nach dem Hauptsatz der Differential- und Integralrechnung, und für $1\leq j\leq n-1$ verschwindet das mittlere Integral nach Induktion. Ist f beliebig wie in (7.4), so wähle $\varphi\in C^\infty(\mathbb{R})$ mit $0\leq\varphi\leq 1$ sowie

$$\varphi(t)=\begin{cases}1 & \text{für } t\leq 1,\\ 0 & \text{für } t\geq 2,\end{cases}$$

und definiere $\eta_R(x)=\varphi(\frac{x_j}{R})$. Dann ist $(\eta_R f)(x)=0$ für $|x_j|\geq 2R$, und es folgt

$$\int_{\mathbb{R}^n}\eta_R(\partial_j f)\,dx=-\int_{\mathbb{R}^n}(\partial_j\eta_R)f\,dx.$$

Es ist $\eta_R(x) = 1$ für $|x_j| \leq R$, insbesondere $\eta_R \to 1$ und $\partial_j \eta_R \to 0$ punktweise auf $\mathbb{R}^n$ mit $R \to \infty$. Außerdem haben wir mit $C = \max |\varphi'|$ die Abschätzungen

$$|\eta_R(\partial_j f)| \leq |\partial_j f| \in L^1(\mathbb{R}^n) \quad \text{und} \quad |(\partial_j \eta_R) f)| \leq \frac{C}{R} |f| \in L^1(\mathbb{R}^n).$$

Behauptung (7.4) folgt mit $R \to \infty$ aus dem Konvergenzsatz von Lebesgue. □

Folgerung 7.1 (Partielle Integration) *Sei Für $f \in C^1_c(\Omega)$ und $g \in C^1(\Omega)$ gilt*

$$\int_{\mathbb{R}^n} (\partial_j f) g \, dx = -\int_{\mathbb{R}^n} f(\partial_j g) \, dx \quad \textit{für } j = 1, \ldots, n.$$

Ist $X \in C^1(\Omega, \mathbb{R}^n)$ so gilt außerdem

$$\int_\Omega \langle \operatorname{grad} f, X \rangle \, dx = -\int_\Omega f(\operatorname{div} X) \, dx.$$

Beweis: Nach Voraussetzung ist $fg \in C^1_c(\Omega)$, also $\int_\Omega \partial_j (fg) \, dx = 0$. Die zweite Fassung folgt durch Ausschreiben in Koordinaten aus der ersten. □

Die Verallgemeinerung des Satzes von Fubini auf kartesische Produkte mit endlich vielen (statt nur zwei) Faktoren soll hier nicht ausgeführt werden. Man zeigt analog zu Lemma 7.2, dass in einem endlichen Produkt von Maßen beliebig Klammern gesetzt oder weggelassen werden können. Der Satz von Fubini wird durch Induktion über die Anzahl der Faktoren verallgemeinert, wobei die Reihenfolge der Integrationen bezüglich der einzelnen Variablen beliebig gewählt werden kann.

8 Der Transformationssatz

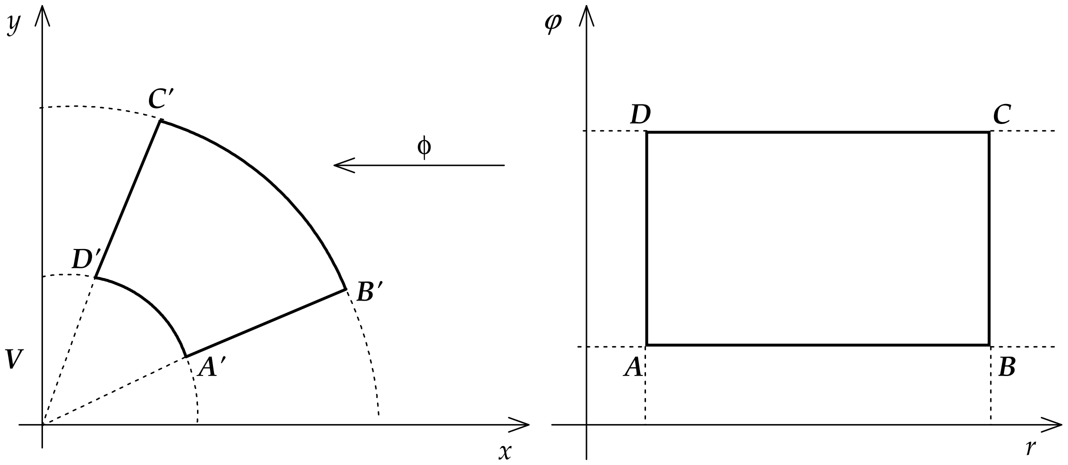
Neben dem Satz von Fubini ist die Transformation auf geeignete Koordinaten das zweite wichtige Hilfsmittel zur Berechnung von Maßen beziehungsweise Integralen im $\mathbb{R}^n$. Nach einigen Beispielen behandeln wir in einem Exkurs die Umrechnung der Differentialoperatoren Gradient, Divergenz und Laplace auf beliebige krummlinige Koordinaten. Diese Formeln kommen bei diversen Problemen in Geometrie und Physik zur Anwendung.

Zu Anfang des Kapitels erinnern wir an folgenden Begriff:

Definition 8.1 (Diffeomorphismus) *Eine Abbildung $\phi : U \to V$ zwischen offenen Mengen $U, V \subset \mathbb{R}^n$ heißt C^1-Diffeomorphismus, wenn ϕ bijektiv ist und ϕ, ϕ^{-1} stetig differenzierbar sind.*

Beispiel 8.1 (Polarkoordinaten im $\mathbb{R}^2$) Betrachte die glatte Abbildung

$$\phi : (0,\infty) \times (0, 2\pi) = U \to V = \mathbb{R}^2 \backslash \{(x,0) : x \geq 0\},\ \phi(r,\varphi) = (r\,\cos\varphi, r\,\sin\varphi).$$



Die Umkehrabbildung von ϕ lautet mit $r = \sqrt{x^2+y^2}$

$$\phi^{-1}(x,y) = \begin{cases} (r,\ \arccos\frac{x}{r}) & \text{falls } y \geq 0, \\ (r,\ 2\pi - \arccos\frac{x}{r}) & \text{falls } y < 0. \end{cases}$$

Für $x < 0$ gilt alternativ $\phi^{-1}(x,y) = (r, \frac{\pi}{2} + \arccos\frac{y}{r})$, insbesondere ist ϕ^{-1} glatt auf ganz V und somit ϕ diffeomorph.

Unser Beweis des Transformationssatzes stützt sich auf folgende infinitesimale Fassung. Dabei verwenden wir für $x \in \mathbb{R}^n$ und $\varrho > 0$ die Bezeichnung

$$Q(x,\varrho) = \{y \in \mathbb{R}^n : \|y - x\| \leq \varrho\} \qquad \text{mit } \|x\| = \max_{1 \leq k \leq n} |x_k|,$$

das heißt $Q(x,\varrho)$ ist der achsenparallele Würfel mit Mittelpunkt x und Kantenlänge 2ϱ.

Lemma 8.1 *Sei $U \subset \mathbb{R}^n$ offen, $x_0 \in U$ und $\phi : U \to \mathbb{R}^n$ mit $D\phi(x_0) \in \mathrm{Gl}_n(\mathbb{R})$. Gegeben sei eine Folge $Q_j = Q(x_j, \varrho_j) \subset U$ mit $\varrho_j \to 0$ und $x_0 \in Q_j$ für alle j. Dann gilt*

$$\limsup_{j\to\infty} \frac{\mathcal{L}^n\big(\phi(Q_j)\big)}{\mathcal{L}^n(Q_j)} \le |\det D\phi(x_0)|.$$

Beweis: Wir können nach geeigneten Translationen $x_0 = 0$ und $\phi(0) = 0$ annehmen, außerdem sei zunächst $D\phi(0) = \mathrm{Id}$. Nach Definition der Differenzierbarkeit gilt dann

$$0 = \lim_{x\to 0} \frac{\|\phi(x) - \big(\phi(0) + D\phi(0)x\big)\|}{\|x\|} = \lim_{x\to 0} \frac{\|\phi(x) - x\|}{\|x\|}.$$

Hier können wir die Maximumsnorm verwenden wegen $\|x\| \le |x| \le \sqrt{n}\,\|x\|$. Sei nun $\varepsilon > 0$. Für $x \in Q_j$ gilt $\|x\| \le \|x - x_j\| + \|x_j - 0\| \le 2\varrho_j$, also folgt für j hinreichend groß

$$\|\phi(x) - x\| \le \varepsilon\,\|x\| \le 2\varepsilon\varrho_j \quad \text{für alle } x \in Q_j.$$

Dies impliziert die Abschätzung

$$\|\phi(x) - \phi(x_j)\| \le \|\phi(x) - x\| + \|x - x_j\| + \|x_j - \phi(x_j)\| \le (1 + 4\varepsilon)\varrho_j,$$

und damit weiter

$$\frac{\mathcal{L}^n\big(\phi(Q_j)\big)}{\mathcal{L}^n(Q_j)} \le (1 + 4\varepsilon)^n.$$

Mit $j \to \infty$ und $\varepsilon \searrow 0$ ergibt sich die Behauptung des Lemmas im Fall $D\phi(0) = \mathrm{Id}$. Allgemein sei $S = D\phi(0)$ und $\phi_0 = S^{-1} \circ \phi$, also $D\phi_0(0) = \mathrm{Id}$. Aus Satz 4.7 folgt

$$\begin{aligned}
\limsup_{j\to\infty} \frac{\mathcal{L}^n\big(\phi(Q_j)\big)}{\mathcal{L}^n(Q_j)} &= \limsup_{j\to\infty} \frac{\mathcal{L}^n\big(S(\phi_0(Q_j))\big)}{\mathcal{L}^n(Q_j)} \\
&= |\det S|\, \limsup_{j\to\infty} \frac{\mathcal{L}^n\big(\phi_0(Q_j)\big)}{\mathcal{L}^n(Q_j)} \\
&\le |\det S| = |\det D\phi(0)|.
\end{aligned}$$

□

Es ist praktisch, in der Transformationsformel statt $d\mathcal{L}^n$ die klassischen Bezeichungen dx bzw. dy zu verwenden, um zwischen Bild und Urbild zu unterscheiden.

Satz 8.1 (Transformationsformel) *Seien $U, V \subset \mathbb{R}^n$ offen und $\phi : U \to V$ ein C^1-Diffeomorphismus. Ist $A \subset U$ $\mathcal{L}^n$-messbar, so ist auch $\phi(A)$ $\mathcal{L}^n$-messbar und es gilt*

$$\mathcal{L}^n(\phi(A)) = \int_A |\det D\phi(x)|\, dx. \tag{8.1}$$

Weiter gilt für jede $\mathcal{L}^n$-messbare Funktion $f : V \to \overline{\mathbb{R}}$

$$\int_V f(y)\, dy = \int_U f(\phi(x)))\, |\det D\phi(x)|\, dx, \tag{8.2}$$

falls eines der Integrale definiert ist.

Beweis: Wir betrachten auf der σ-Algebra $\mathcal{A}$ der $\mathcal{L}^n$-messbaren Mengen $A \subset U$ die Maße

$$\lambda = \mathcal{L}^n \llcorner |\det D\phi| \quad \text{sowie} \quad \mu = \psi(\mathcal{L}^n), \text{ wobei } \psi = \phi^{-1} : V \to U.$$

Nach Definition des Bildmaßes ist

$$\mu(A) = \mathcal{L}^n\big(\psi^{-1}(A)\big) = \mathcal{L}^n(\phi(A)).$$

Für jede $\mathcal{L}^n$-messbare Menge $A \subset U$ ist $\psi^{-1}(A)$ ebenfalls $\mathcal{L}^n$-messbar nach Satz 4.5, und damit A μ-messbar nach Satz 2.1. Die Einschränkung des äußeren Maßes μ auf $\mathcal{A}$ ist somit ein Maß. Die Maßeigenschaft von λ gilt nach Definition des Maßes mit Dichte in Satz 6.1. Weiter wurde in den Sätzen 6.1 und 4.5 bewiesen:

$$N \subset U \text{ mit } \mathcal{L}^n(N) = 0 \quad \Rightarrow \quad \lambda(N) = \mu(N) = 0. \tag{8.3}$$

Wir zeigen nun $\lambda \geq \mu$. Indem wir U durch offene, relativ kompakte Mengen ausschöpfen, können wir dabei $\lambda(U) < \infty$ und $\mu(U) < \infty$ voraussetzen. Nach Satz 4.2 ist jede $\mathcal{L}^n$-messbare Menge von der Form $A = E \backslash N$, wobei $\mathcal{L}^n(N) = 0$ und $E = \bigcap_{i=1}^\infty U_i$ mit offenen $U_1 \supset U_2 \supset \ldots$. Wegen Satz 2.4(ii) (hier brauchen wir $\lambda(U), \mu(U) < \infty$) reicht es daher aus, die Ungleichung auf allen offenen Mengen nachzuweisen. Aber nach Lemma 1.2 ist jede offene Menge als abzählbare Vereinigung von kompakten, achsenparallelen Würfeln in U mit paarweise disjunktem Inneren darstellbar. Damit bleibt zu zeigen:

$$\lambda(Q_0) \geq \mu(Q_0) \text{ für alle } Q_0 = Q(p_0, \varrho_0) \subset U. \tag{8.4}$$

Angenommen es ist $\lambda(Q_0) < \mu(Q_0)$ für ein $Q_0 = Q(p_0, \varrho_0)$. Wähle dann ein $\theta \in [0,1)$ mit $\lambda(Q_0) < \theta\mu(Q_0)$, und zerlege Q_0 durch Halbierung der Kanten in 2^n kompakte Teilwürfel $Q_{0,1}, \ldots, Q_{0,2^n}$. Wäre $\lambda(Q_{0,i}) \geq \theta\mu(Q_{0,i})$ für alle i, so folgt durch Summation

$$\lambda(Q_0) = \sum_{i=1}^{2^n} \lambda(Q_{0,i}) \geq \theta \sum_{i=1}^{2^n} \mu(Q_{0,i}) = \theta\mu(Q_0),$$

ein Widerspruch. Hierbei wurde benutzt, dass die Ränder der $Q_{0,i}$ nach (8.3) jeweils Nullmengen sind. Unter den $Q_{0,i}$ gibt es also einen Würfel Q_1 mit $\lambda(Q_1) < \theta\mu(Q_1)$. Bestimme so induktiv eine Schachtelung $Q_0 \supset Q_1 \supset \ldots$ mit

$$\lambda(Q_j) < \theta\mu(Q_j) \quad \text{für alle } j \in \mathbb{N}. \tag{8.5}$$

Für $x_0 = \bigcap_{j=1}^\infty Q_j \in U$ gilt wegen der Stetigkeit der Funktion $\det D\phi$ mit $j \to \infty$

$$\left| \frac{\lambda(Q_j)}{\mathcal{L}^n(Q_j)} - |\det D\phi(x_0)| \right| = \frac{1}{\mathcal{L}^n(Q_j)} \left| \int_{Q_j} \big(|\det D\phi(x)| - |\det D\phi(x_0)|\big)\, dx \right| \to 0.$$

Zusammen mit Lemma 8.1 folgt aber nun

$$\liminf_{j\to\infty} \frac{\lambda(Q_j)}{\mu(Q_j)} = \liminf_{j\to\infty} \Big(\frac{\lambda(Q_j)}{\mathcal{L}^n(Q_j)} \cdot \frac{\mathcal{L}^n(Q_j)}{\mu(Q_j)} \Big) \geq 1.$$

ein Widerspruch zu (8.5). Dies zeigt (8.4) und damit die Ungleichung $\lambda \geq \mu$.

Für $f : V \to \overline{\mathbb{R}}$ gilt $\{f \circ \phi < s\} = \phi^{-1}\big(\{f < s\}\big)$, also ist mit f auch $f \circ \phi$ messbar bezüglich $\mathcal{L}^n$, und umgekehrt. Es folgt weiter für $\mathcal{L}^n$-messbares $f : V \to [0, \infty]$

$$\int_U f(\phi(x))\,|\det D\phi(x)|\,dx \geq \int_V f(y)\,dy. \tag{8.6}$$

Und zwar ergibt sich (8.6) erst für $f = \chi_B$, indem wir $A = \psi(B)$ in der Ungleichung $\lambda(A) \geq \mu(A)$ setzen, dann für nichtnegative $\mathcal{L}^n$-Treppenfunktionen und schließlich für beliebiges, messbares f durch Approximation von unten nach Satz 5.3. Um schließlich für $f \geq 0$ die Gleichheit in (8.2) zu zeigen, wenden wir (8.6) auf $\psi : V \to U$ statt ϕ und auf $g = f \circ \phi |\det D\phi|$ an. Dies liefert wegen $1 = \det D(\phi \circ \psi)(y) = \big(\det D\phi\big)\big(\psi(y)\big) \det D\psi(y)$

$$\int_V f(y)\,dy = \int_V g\big(\psi(y)\big)|\det D\psi(y)|\,dy \geq \int_U g(x)\,dx = \int_U f\big(\phi(x)\big)|\det D\phi(x)|\,dx.$$

Damit ist (8.2) für $f \geq 0$ bewiesen, und (8.1) folgt durch Wahl von $f = \chi_{\phi(A)}$. Für beliebige f zerlegen wir $f = f^+ - f^-$. $\square$

Beispiel 8.2 (Gauß-Integral) Für $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = e^{-(x^2+y^2)}$ folgt mit Fubini

$$\int_{\mathbb{R}^2} f\,d\mathcal{L}^2 = \int_{\mathbb{R}} e^{-x^2} \left(\int_{\mathbb{R}} e^{-y^2}\,dy\right) dx = \left(\int_{\mathbb{R}} e^{-x^2}\,dx\right)^2.$$

Für Polarkoordinaten $\phi : (0, \infty) \times (0, 2\pi) \to \mathbb{R}^2 \backslash \{(x, 0) : x \geq 0\}$ gilt $\det D\phi(r, \theta) = r$. Da $\{(x, 0) : x \geq 0\}$ eine $\mathcal{L}^2$-Nullmenge ist, folgt aus dem Transformationssatz, und anschließend dem Satz von Fubini,

$$\int_{\mathbb{R}^2} f\,d\mathcal{L}^2 = \int_{(0,\infty)\times(0,2\pi)} e^{-r^2} r\,d\mathcal{L}^2(r, \theta) = \int_0^\infty e^{-r^2} r \left(\int_0^{2\pi} d\theta\right) dr = \pi.$$

Also gilt $\int_{\mathbb{R}} e^{-x^2}\,dx = \sqrt{\pi}$.

Beispiel 8.3 (Lineare Transformationsformel) Betrachte den Spezialfall von Satz 8.1, dass der Diffeomorphismus $\phi : U \to V$ Einschränkung einer linearen Abbildung ist, das heißt $\phi(x) = Sx$ mit $S \in \mathrm{GL}_n(\mathbb{R})$. Dann gilt $D\phi(x) = S$ für alle $x \in U$, und die Formeln aus dem Transformationssatz lauten, vgl. auch Satz 4.7,

$$\mathcal{L}^n(S(D)) = |\det S|\,\mathcal{L}^n(D) \quad \text{bzw.} \quad \int_V f(y)\,d\mathcal{L}^n(y) = |\det S| \int_U f(Sx)\,d\mathcal{L}^n(x).$$

Das nächste Ziel ist die Umrechnung von Differentialoperatoren in neue Koordinaten. Es ist praktisch, dazu folgenden Begriff einzuführen. Für einen C^k-Diffeomorphismus $\phi : U \to V$ zwischen offenen Mengen $U, V \subset \mathbb{R}^n$ definieren wir die *Gramsche Matrix* $g \in C^{k-1}(U, \mathbb{R}^{n\times n})$, $g = (g_{ij})$, durch

$$g(x) = D\phi(x)^T D\phi(x) \quad \text{bzw.} \quad g_{ij}(x) = \Big\langle \frac{\partial\phi}{\partial x_i}(x), \frac{\partial\phi}{\partial x_j}(x)\Big\rangle. \tag{8.7}$$

Nach Definition ist $g(x)$ symmetrisch und strikt positiv definit, denn für $v \in \mathbb{R}^n$, $v \neq 0$ gilt

$$\langle g(x)v, v\rangle = \langle D\phi(x)^T D\phi(x)v, v\rangle = |D\phi(x)v|^2 > 0, \quad \text{da } D\phi(x) \in \mathrm{Gl}_n(\mathbb{R}).$$

Insbesondere ist $g(x)$ invertierbar. Wir setzen

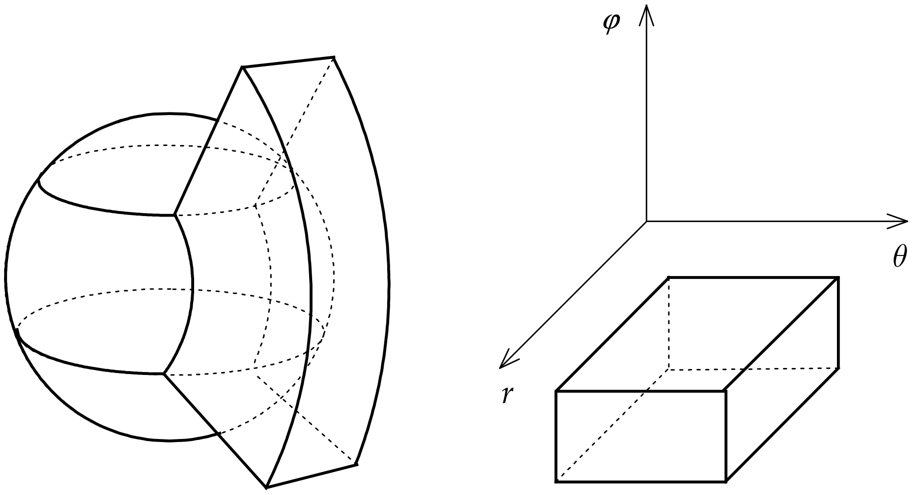
$$g^{ij}(x) = \big(g(x)^{-1}\big)_{ij}, \qquad \text{also } \sum_{j=1}^{n} g^{ij} g_{jk} = \delta_{ik}. \tag{8.8}$$

Wegen $\det g = \det\big(D\phi^T D\phi\big) = |\det D\phi|^2$ können wir die Transformationsformel alternativ wie folgt schreiben:

$$\int_V f(y)\,dy = \int_U f\big(\phi(x)\big)\sqrt{\det g(x)}\,dx. \tag{8.9}$$

Beispiel 8.4 (Polarkoordinaten im $\mathbb{R}^3$) Betrachte für $U = (0,\infty)\times(0,\pi)\times(0,2\pi)$ und $V = \mathbb{R}^3\backslash\{(x,0,z) : x \geq 0\}$ die Polarkoordinatenabbildung

$$\phi : U \to V,\ \phi(r,\theta,\varphi) = (r\sin\theta\cos\varphi,\ r\sin\theta\sin\varphi,\ r\cos\theta).$$



Die Jacobimatrix von ϕ lautet

$$D\phi(r,\theta,\varphi) = \begin{pmatrix} \sin\theta\,\cos\varphi & r\,\cos\theta\,\cos\varphi & -r\,\sin\theta\,\sin\varphi \\ \sin\theta\,\sin\varphi & r\,\cos\theta\,\sin\varphi & r\,\sin\theta\,\cos\varphi \\ \cos\theta & -r\,\sin\theta & 0 \end{pmatrix}.$$

Für die Gramsche Matrix ergibt sich

$$\big(g_{ij}(r,\theta,\varphi)\big)_{1\leq i,j\leq 3} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2\sin^2\theta \end{pmatrix}.$$

ϕ ist diffeomorph, und zwar kann die Umkehrabbildung explizit bestimmt werden. Das Lebesguemaß der Menge $\phi(E)$ für $E = [r_1,r_2]\times[\theta_1,\theta_2]\times[\varphi_1,\varphi_2]$ ist nach dem Transformationssatz und Fubini

$$\mathcal{L}^3\big(\phi(E)\big) = \int_{r_1}^{r_2}\int_{\theta_1}^{\theta_2}\int_{\varphi_1}^{\varphi_2} r^2\sin\theta\ d\varphi\,d\theta\,dr = \frac{r_2^3 - r_1^3}{3}(\cos\theta_1 - \cos\theta_2)(\varphi_2 - \varphi_1).$$

Sei $\phi : U \to V$ ein Diffeomorphismus und $v : V \to \mathbb{R}$ eine Funktion. Die Darstellung von $v = v(y)$ in den x-Koordinaten ist die Verkettung $u = v \circ \phi : U \to \mathbb{R}$, $u(x) = v\big(\phi(x)\big)$. Zum Beispiel lautet die Polarkoordinatendarstellung einer Funktion $v = v(x, y, z)$

$$u(r, \theta, \varphi) = v(r \sin\theta \sin\varphi, r \sin\theta \cos\varphi, r \cos\theta).$$

In der Physik ist es üblich, die neue Funktion mit demselben Buchstaben zu bezeichnen, da sie nach wie vor dieselbe physikalische Größe repräsentiert. Zum Beispiel kann eine Temperaturverteilung T entweder als Funktion $T = T(x, y, z)$ der Euklidischen Koordinaten oder als Funktion $T = T(r, \theta, \varphi)$ der Polarkoordinaten aufgefasst werden. Für Mathematiker ist das etwas verwirrend, und wir wollen die Unterscheidung ziwschen v und $u = v \circ \phi$ lieber beibehalten. Für die Darstellung eines Vektorfelds $Y : V \to \mathbb{R}^n$ in den x-Koordinaten muss $Y \circ \phi$ in der Basis $\frac{\partial \phi}{\partial x_1}, \ldots, \frac{\partial \phi}{\partial x_n}$ entwickelt werden:

$$Y \circ \phi = \sum_{i=1}^{n} X_i \frac{\partial \phi}{\partial x_i} = \sum_{i=1}^{n} X_i \, D\phi \cdot e_i = D\phi \cdot X.$$

Das eindeutig bestimmte Vektorfeld $X : U \to \mathbb{R}^n$ mit dieser Eigenschaft lautet

$$X = \sum_{i,j=1}^{n} g^{ij} \Big\langle Y \circ \phi, \frac{\partial \phi}{\partial x_i} \Big\rangle e_j. \tag{8.10}$$

Denn mit der Formel (8.10) für X berechnen wir wegen $\sum_{j=1}^n g^{ij} g_{jk} = \delta_{ik}$

$$\begin{aligned}
\Big\langle D\phi \cdot X, \frac{\partial \phi}{\partial x_k} \Big\rangle &= \sum_{i,j=1}^{n} g^{ij} \Big\langle Y \circ \phi, \frac{\partial \phi}{\partial x_i} \Big\rangle \Big\langle \frac{\partial \phi}{\partial x_j}, \frac{\partial \phi}{\partial x_k} \Big\rangle \\
&= \sum_{i=1}^{n} \Big\langle Y \circ \phi, \frac{\partial \phi}{\partial x_i} \Big\rangle \sum_{j=1}^{n} g^{ij} g_{jk} \\
&= \Big\langle Y \circ \phi, \frac{\partial \phi}{\partial x_k} \Big\rangle.
\end{aligned}$$

Da die Vektoren $\frac{\partial \phi}{\partial x_k}$ eine Basis bilden, folgt $Y \circ \phi = D\phi \cdot X$ wie gewünscht. Natürlich ist auch die Formel $X = (D\phi)^{-1}(Y \circ \phi)$ richtig, erweist sich aber als weniger praktisch, weil (g^{ij}) oft einfacher zu berechnen ist als $(D\phi)^{-1}$.

Satz 8.2 (Umrechnung von Differentialoperatoren) *Sei $\phi \in C^1(U, V)$ ein Diffeomorphismus zwischen den offenen Mengen $U, V \subset \mathbb{R}^n$ mit Gramscher Matrix (g_{ij}).*

(1) *Für $v \in C^1(V)$ gilt mit $u = v \circ \phi$*

$$(\operatorname{grad} v) \circ \phi = D\phi \cdot \operatorname{grad}_g u, \qquad \textit{wobei } \operatorname{grad}_g u := \sum_{i,j=1}^{n} g^{ij} \frac{\partial u}{\partial x_i} e_j.$$

(2) *Für $Y \in C^1(V, \mathbb{R}^n)$ gilt mit $Y \circ \phi = D\phi \cdot X$*

$$(\operatorname{div} Y) \circ \phi = \operatorname{div}_g X := \frac{1}{\sqrt{\det g}} \sum_{j=1}^{n} \frac{\partial}{\partial x_j} \big(\sqrt{\det g} X_j\big).$$

(3) *Ist $\phi \in C^2(U,V)$ und $v \in C^2(V)$, so folgt wieder mit $u = v \circ \phi$*

$$(\Delta v) \circ \phi = \mathrm{div}_g\, \mathrm{grad}_g u = \frac{1}{\sqrt{\det g}} \sum_{i,j=1}^n \frac{\partial}{\partial x_i}\Big(\sqrt{\det g}\, g^{ij} \frac{\partial u}{\partial x_j}\Big).$$

Beweis: Nach der Definition des Gradienten und der Kettenregel gilt

$$\Big\langle (\mathrm{grad}\, v) \circ \phi, \frac{\partial \phi}{\partial x_i} \Big\rangle = (Dv) \circ \phi \cdot D\phi \cdot e_i = D(v \circ \phi) \cdot e_i = \frac{\partial u}{\partial x_i}.$$

Also folgt (1) aus Gleichung (8.10), und allgemein gilt für ein Vektorfeld $X : U \to \mathbb{R}^n$

$$\langle D\phi \cdot \mathrm{grad}_g u, D\phi \cdot X \rangle = \sum_{j=1}^n \Big\langle D\phi \cdot \mathrm{grad}_g u, \frac{\partial \phi}{\partial x_j} \Big\rangle X_j = \sum_{j=1}^n \frac{\partial u}{\partial x_j} X_j. \tag{8.11}$$

Aussage (2) führen wir durch partielle Integration auf die Umrechnungsformel für den Gradienten zurück. Sei $\psi \in C^1(V)$ beliebig und $\varphi = \psi \circ \phi$. Dann folgt mit partieller Integration, dem Transformationssatz, Behauptung (1), (8.11) und der Definition von $\mathrm{div}_g X$

$$\begin{aligned}
\int_V \psi\, \mathrm{div}\, Y\, dy &= -\int_V \langle \mathrm{grad}\, \psi, Y \rangle\, dy \\
&= -\int_U \langle (\mathrm{grad} \psi) \circ \phi, Y \circ \phi \rangle \sqrt{\det g}\, dx \\
&= -\int_U \langle D\phi \cdot \mathrm{grad}_g \varphi, D\phi \cdot X \rangle \sqrt{\det g}\, dx \\
&= -\int_U \sum_{j=1}^n \frac{\partial \varphi}{\partial x_j} X_j \sqrt{\det g}\, dx \\
&= \int_U \varphi\, \mathrm{div}_g X \sqrt{\det g}\, dx \\
&= \int_V \psi\, (\mathrm{div}_g X) \circ \phi^{-1}\, dy.
\end{aligned}$$

Da ψ beliebig ist, folgt $\mathrm{div}\, Y = (\mathrm{div}_g X) \circ \phi^{-1}$ und damit Aussage (2) aus nachfolgendem Lemma 8.2. Schließlich ergibt sich (3) durch Kombination von (1) und (2), und zwar gilt

$$(\Delta v) \circ \phi = (\mathrm{div}\, \mathrm{grad}\, v) \circ \phi \quad \text{mit } (\mathrm{grad}\, v) \circ \phi = D\phi \cdot \mathrm{grad}_g u \text{ nach (1)},$$

also folgt weiter wegen (2) und $g^{ji} = g^{ij}$

$$(\Delta v) \circ \phi = \mathrm{div}_g\, \mathrm{grad}_g u = \frac{1}{\sqrt{\det g}} \sum_{i,j=1}^n \frac{\partial}{\partial x_i}\Big(\sqrt{\det g}\; g^{ij} \frac{\partial u}{\partial x_j}\Big).$$

□

Wir tragen nun das Lemma nach, das im Beweis gebraucht wurde.

Lemma 8.2 *Sei $U \subset \mathbb{R}^n$ offen und $u \in C^0(U)$ mit $\int_U u\varphi\, dx = 0$ für alle $\varphi \in C_c^\infty(U)$. Dann ist u die Nullfunktion.*

Beweis: Angenommen es gibt ein $x \in U$ mit $u(x) > 0$. Dann gibt es $\varrho > 0$ und $\delta > 0$ mit $u(x) \geq \delta$ auf $B_\varrho(x) \subset U$. Wähle eine Funktion $\eta \in C^\infty_c(U)$ mit $\operatorname{spt} \eta \subset B_\varrho(x)$, $\eta \geq 0$ und $\int_U \eta(x)\,dx > 0$. Aus der Monotonie des Integrals folgt

$$0 = \int_U u(x)\eta(x)\,dx \geq \int_U \delta\eta(x)\,dx > 0,$$

ein Widerspruch. $\square$

Beispiel 8.5 (Polarkoordinatendarstellung des Laplaceoperators) Für Polarkoordinaten im $\mathbb{R}^3$, siehe Beispiel 8.4, ergeben sich folgende Formeln:

$$\begin{aligned}
\operatorname{grad}_g u &= \frac{\partial u}{\partial r}\, e^r + \frac{1}{r^2}\frac{\partial u}{\partial \theta}\, e^\theta + \frac{1}{r^2 \sin^2\theta}\,\frac{\partial u}{\partial \varphi}\, e^\varphi,\\
\operatorname{div}_g X &= \frac{1}{r^2}\,\frac{\partial}{\partial r}(r^2 X^r) + \frac{1}{\sin\theta}\,\frac{\partial}{\partial\theta}\big((\sin\theta)\, X^\theta\big) + \frac{\partial X^\varphi}{\partial\varphi},\\
\Delta_g u &= \frac{1}{r^2}\,\frac{\partial}{\partial r}\Big(r^2\frac{\partial u}{\partial r}\Big) + \frac{1}{r^2\sin\theta}\,\frac{\partial}{\partial\theta}\Big((\sin\theta)\,\frac{\partial u}{\partial\theta}\Big) + \frac{1}{r^2\sin^2\theta}\,\frac{\partial^2 u}{\partial\varphi^2}.
\end{aligned}$$

Hier bezeichnet e^r, e^θ, e^φ die Standardbasis im (r,θ,φ)-Raum, und X^r, X^θ, X^φ sind die zugehörigen Koordinaten von X. Für die Funktion $v(x,y,z) = (x^2+y^2+z^2)^{-1/2}$ ist $u(r) = 1/r$ und somit $\Delta v = 0$ auf $\mathbb{R}^3 \backslash \{0\}$.

Die Gramsche Matrix $g : U \to \mathbb{R}^{n\times n}$ eines Diffeomorphismus $\phi \in C^1(U,V)$ definiert in jedem Punkt $x \in U$ ein Skalarprodukt und eine zugehörige Euklidische Norm, und zwar gilt

$$g(x)(v,w) = \langle D\phi(x)v, D\phi(x)w\rangle \quad \text{und} \quad \|v\|_{g(x)} = \sqrt{g(x)(v,v)} = |D\phi(x)v|.$$

Ein solches ortsabhängiges Skalarprodukt nennt man eine Riemannsche Metrik, genauer spricht man hier von der durch ϕ *induzierten Metrik*. Mit g kann nicht nur das Maß des Bildes berechnet werden, sondern auch die Bogenlänge von Kurven $\phi \circ \gamma : [a,b] \to V$:

$$L(\phi\circ\gamma) = \int_a^b \Big|\frac{d}{dt}\phi\big(\gamma(t)\big)\Big|\,dt = \int_a^b |D\phi\big(\gamma(t)\big)\gamma'(t)|\,dt = \int_a^b \sqrt{g\big(\gamma(t)\big)\big(\gamma'(t),\gamma'(t)\big)}\,dt.$$

In seinem Habilitationsvortrag *Über die Hypothesen, die der Geometrie zugrundeliegen* hat Bernhard Riemann 1854 die Idee einer Geometrie für solche ortsabhängigen Skalarprodukte entworfen, die später Grundlage der Allgemeinen Relativitätstheorie wurde.

9 Das Flächenmaß auf Untermannigfaltigkeiten

Hauptresultat dieses Abschnitts ist die Definition des Flächenmaßes für n-dimensionale C^1-Untermannigfaltigkeiten des $\mathbb{R}^{n+k}$. Dazu wird zunächst der Flächeninhalt für Immersionen ad hoc definiert und an einigen Beispielen motiviert. Diese Definition wird dann mittels lokaler Parameterdarstellungen auf Untermannigfaltigkeiten übertragen. Als Anwendung wird die Aufspaltung von Gebietsintegralen bezüglich Radial- und Winkelkoordinaten, die sogenannte Zwiebelformel, behandelt.

Zur Definition des Flächeninhalts muss zunächst geklärt werden, was unter einer Fläche zu verstehen ist. Am einfachsten ist der Begriff der parametrisierten Fläche oder Immersion. Sei $U \subset \mathbb{R}^n$ offen. Eine Abbildung $f \in C^1(U, \mathbb{R}^{n+k})$ heißt Immersion, wenn gilt:

$$\operatorname{rang} Df(x) = n \quad \text{bzw. äquivalent} \quad \ker Df(x) = \{0\} \quad \text{für alle } x \in U.$$

Die Vektoren $\frac{\partial f}{\partial x_1}(x), \ldots, \frac{\partial f}{\partial x_n}(x)$ bilden dann eine Basis von $\operatorname{Bild} Df(x) \subset \mathbb{R}^{n+k}$. Die Zahl n heißt Dimension, die Zahl k Kodimension von f. Wir definieren analog zu (8.7) im vorigen Kapitel die Gramsche Matrix oder induzierte Metrik

$$g(x) = Df(x)^T Df(x) \quad \text{bzw.} \quad g_{ij}(x) = \Big\langle \frac{\partial f}{\partial x_i}(x), \frac{\partial f}{\partial x_j}(x) \Big\rangle.$$

Offenbar ist (g_{ij}) für eine beliebige Abbildung $f \in C^1(U, \mathbb{R}^{n+k})$ definiert und positiv semidefinit. Die Matrix ist genau dann strikt positiv definit und damit invertierbar, wenn f eine Immersion ist, denn es gilt $\langle g(x)v, v\rangle = |Df(x)v|^2$, also $\ker g(x) = \ker Df(x)$.

Definition 9.1 (Flächenformel) *Sei $f \in C^1(U, \mathbb{R}^{n+k})$ eine n-dimensionale Immersion mit Gramscher Matrix g, und $E \subset U$ sei $\mathcal{L}^n$-messbar. Der (n-dimensionale) Flächeninhalt von f auf E ist definiert durch*

$$A_E(f) = \int_E Jf(x)\, dx \qquad \textit{mit } Jf = \sqrt{\det g}.$$

Die Funktion Jf heißt Flächenintegrand oder Jacobische von f.

Wir wollen einige Spezialfälle betrachten.

Beispiel 9.1 Sei $S : \mathbb{R}^n \to \mathbb{R}^{n+k}$ eine lineare Abbildung vom Rang n. Da das Maß $\mathcal{L}^n$ invariant unter $\mathbb{O}(n)$ ist, gibt es auf $\operatorname{Bild} S$ genau ein Maß λ mit $\lambda(T(E)) = \mathcal{L}^n(E)$ für jede Isometrie $T : \mathbb{R}^n \to \operatorname{Bild} S$. Wähle nun $R \in \mathbb{R}^{n\times n}$ symmetrisch und strikt positiv definit, so dass $S^T S = R^2$, vergleiche Lemma 4.6. Für $Q = SR^{-1}$ folgt

$$\langle Qv, Qw\rangle = \langle (R^{-1})^T S^T S R^{-1} v, w\rangle = \langle v, w\rangle,$$

das heißt $Q : \mathbb{R}^n \to \operatorname{Bild} S$ ist eine Isometrie, und wir erhalten wegen $\sqrt{\det(S^T S)} = \det R$

$$A_E(S) = (\det R)\mathcal{L}^n(E) = \mathcal{L}^n(R(E)) = \lambda(Q(R(E)) = \lambda(S(E)).$$

Für lineare Abbildungen liefert die Flächenformel also jedenfalls den richtigen Wert.

Beispiel 9.2 Eine eindimensionale Immersion $f : I = (a, b) \to \mathbb{R}^{1+k}$, $f = f(t)$, heißt auch reguläre Kurve. Die induzierte Metrik ist dann die 1×1-Matrix $\langle f'(t), f'(t)\rangle = |f'(t)|^2$, und nach Definition 9.1 ist die Länge der Kurve

$$L(f) = \int_a^b |f'(t)|\, dt.$$

Analog zu Beispiel 9.4 wird hier die Länge des Bildes mit der Vielfachheit gezählt, mit der es durchlaufen wird. Zum Beispiel gilt $L_{[0,3\pi]}(f) = 3\pi$ für $f(t) = (\cos t, \sin t)$.

Beispiel 9.3 Eine zweidimensionale Immersion $f : U \to \mathbb{R}^3$, $f = f(x, y)$, heißt auch reguläre Fläche. Die induzierte Metrik lautet

$$(g_{ij}) = \begin{pmatrix} \left|\frac{\partial f}{\partial x}\right|^2 & \left\langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right\rangle \\ \left\langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right\rangle & \left|\frac{\partial f}{\partial y}\right|^2 \end{pmatrix}.$$

Also lautet die Jacobische

$$Jf = \sqrt{\left|\frac{\partial f}{\partial x}\right|^2 \left|\frac{\partial f}{\partial y}\right|^2 - \left\langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right\rangle^2} = \left|\frac{\partial f}{\partial x} \times \frac{\partial f}{\partial y}\right|.$$

Die zweite Gleichung folgt dabei mit etwas Rechnung aus der Definition des Kreuzprodukts. Betrachte als Beispiel Polarkoordinaten auf der Sphäre, also

$$f : U = (0, \pi) \times (0, 2\pi) \to \mathbb{S}^2 \subset \mathbb{R}^3,\ f(\theta, \varphi) = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta).$$

Mit Beispiel 8.4 im vorigen Kapitel gilt $Jf(\theta, \varphi) = \sin\theta$, insbesondere ist f eine reguläre Fläche mit Flächeninhalt

$$A(f) = \int_0^\pi \int_0^{2\pi} \sin\theta\, d\varphi d\theta = 4\pi.$$

Beispiel 9.4 Für $k = 0$, also $f \in C^1(U, \mathbb{R}^n)$ mit $U \subset \mathbb{R}^n$ offen, sprechen wir statt vom n-dimensionalen Flächeninhalt besser vom Volumen $\mathrm{vol}_E(f)$. Die Abbildung f ist genau dann eine Immersion, wenn $\det Df(x) \neq 0$ für alle $x \in U$, das heißt f ist lokal diffeomorph. Es gilt $Jf = \sqrt{\det(Df^T Df)} = |\det Df|$. Ist zusätzlich $f : U \to f(U)$ injektiv und damit diffeomorph, so besagt der Transformationssatz

$$\mathrm{vol}_E(f) = \int_E |\det Df(x)|\, dx = \mathcal{L}^n(f(E)).$$

Im allgemeinen ist $\mathrm{vol}_E(f)$ nicht das $\mathcal{L}^n$-Maß des Bildes $f(E)$, sondern die Bildpunkte werden mit der Vielfachheit gezählt, mit der sie angenommen werden.

Beispiel 9.5 Für $u \in C^1(U, \mathbb{R}^k)$ ist die zugehörige Graphenabbildung

$$f : U \to \mathbb{R}^{n+k},\ f(x) = (x, u(x)),$$

eine Immersion. Denn bezeichnet $P : \mathbb{R}^{n+k} = \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^n$ die Projektion auf die ersten n Koordinaten, so gilt $P \circ f = \mathrm{id}_U$, also nach Kettenregel $P \cdot Df(x) = D(P \circ f)(x) = \mathrm{Id}_{\mathbb{R}^n}$, das heißt $Df(x)$ ist injektiv. Man berechnet

$$\left\langle \frac{\partial f}{\partial x_i}, \frac{\partial f}{\partial x_j}\right\rangle = \left\langle \left(e_i, \frac{\partial u}{\partial x_i}\right), \left(e_j, \frac{\partial u}{\partial x_j}\right)\right\rangle = \delta_{ij} + \left\langle \frac{\partial u}{\partial x_i}, \frac{\partial u}{\partial x_j}\right\rangle,$$

das heißt es gilt $(Df)^T Df = \mathrm{Id}_{\mathbb{R}^n} + (Du)^T Du$ und

$$A(f) = \int_U \sqrt{\det\big(\mathrm{Id}_{\mathbb{R}^n} + (Du)^T Du\big)}\,dx.$$

In Kodimension $k = 1$, also u reellwertig, gibt es zu $x \in U$ eine Orthonormalbasis $v_1, \ldots, v_n$ mit $Du(x)v_1 = |Du(x)|$ und $Du(x)v_j = 0$ für $j \geq 2$. Das ist trivial im Fall $Du(x) = 0$, für $Du(x) \neq 0$ wähle $v_1 = Du(x)/|Du(x)|$ und ergänze zu einer Orthonormalbasis. Es folgt

$$\begin{aligned}\big\langle \big(\mathrm{Id}_{\mathbb{R}^n} + Du(x)^T Du(x)\big)v_i, v_j\big\rangle &= \delta_{ij} + (Du(x)v_i)\,(Du(x)v_j)\\ &= \begin{cases} 1 + |Du(x)|^2 & \text{für } i = j = 1,\\ \delta_{ij} & \text{sonst.}\end{cases}\end{aligned}$$

Somit ergibt für einen Graphen der Kodimension Eins die klassische Formel

$$A(f) = \int_U \sqrt{1 + |Du(x)|^2}\,dx \qquad \text{für } f : U \to \mathbb{R}^{n+1},\ f(x) = (x, u(x)). \tag{9.1}$$

Als geometrische Größe sollte der Flächeninhalt nicht von der Wahl der Parametrisierung abhängen. Dies wollen wir sofort überprüfen.

Satz 9.1 (Invarianz des Flächeninhalts) *Sei $\phi \in C^1(U, V)$ ein Diffeomorphismus und $f \in C^1(V, \mathbb{R}^{n+k})$ eine Immersion. Dann gilt*

$$A_{\phi(E)}(f) = A_E(f \circ \phi) \qquad \textit{für jede } \mathcal{L}^n\textit{-messbare Menge } E \subset U.$$

Beweis: Seien g bzw. h die induzierten Metriken von $f \circ \phi$ bzw. f. Es gilt

$$D(f \circ \phi)(x)^T D(f \circ \phi)(x) = D\phi(x)^T Df(\phi(x))^T Df(\phi(x)) D\phi(x),$$

beziehungsweise

$$g(x) = D\phi(x)^T h(\phi(x)) D\phi(x). \tag{9.2}$$

Mit den Rechenregeln für die Determinante ergibt sich

$$J(f \circ \phi)(x) = (Jf)(\phi(x))\,|\det D\phi(x)|. \tag{9.3}$$

Die Behauptung folgt damit aus dem Transformationssatz, Satz 8.1. $\square$

Wir wollen nun einen globaleren Standpunkt einnehmen und Flächen betrachten, die nicht bzw. nicht a priori durch eine einzige Parametrisierung gegeben sind. Unser Ziel ist es, auf diesen Flächen – genauer: Untermannigfaltigkeiten – ein n-dimensionales Flächenmaß zu definieren. Wir erinnern vorher an die relevanten Begriffe aus Analysis II, Kapitel 9.2.

Definition 9.2 (Untermannigfaltigkeit) *Sei $n \in \mathbb{N}$, $k \in \mathbb{N}_0$. Eine Menge $M \subset \mathbb{R}^{n+k}$ heißt n-dimensionale Untermannigfaltigkeit des $\mathbb{R}^{n+k}$ der Klasse C^1, falls gilt: zu jedem $p \in M$ gibt es eine offene Umgebung $W \subset \mathbb{R}^{n+k}$ und einen C^1-Diffeomorphismus $\phi : W \to \phi(W) \subset \mathbb{R}^{n+k}$ mit*

$$\phi(M \cap W) = \mathbb{R}^n \cap \phi(W) = \{z \in \phi(W) : z_{n+1} = \ldots = z_{n+k} = 0\}.$$

Wir haben diese Eigenschaft auch als lokale Plättbarkeit bezeichnet, und folgende alternative Charakterisierungen bewiesen.

Satz 9.2 (Untermannigfaltigkeitskriterien) *Sei $n \in \mathbb{N}$, $k \in \mathbb{N}_0$. Für $M \subset \mathbb{R}^{n+k}$ sind folgende Aussagen äquivalent:*

(1) *M ist eine n-dimensionale C^1-Untermannigfaltigkeit.*

(2) Niveaumengenkriterium: *Zu jedem $p \in M$ gibt es eine offene Umgebung $W \subset \mathbb{R}^{n+k}$ und eine Funktion $h \in C^1(W, \mathbb{R}^k)$ mit $\operatorname{rang} Dh(q) = k$ für alle $q \in W$, so dass*

$$M \cap W = \{q \in W : h(q) = 0\}.$$

(3) Graphenkriterium: *Zu jedem $p \in M$ gibt es nach geeigneter Umnummerierung der Koordinaten eine offene Umgebung $U \times V \subset \mathbb{R}^n \times \mathbb{R}^k$ und eine Funktion $u \in C^1(U, V)$, so dass gilt:*

$$M \cap (U \times V) = \big(\{(x, u(x)) : x \in U\}\big).$$

Beispiel 9.6 Die Sphäre $\mathbb{S}^n = \{p \in \mathbb{R}^{n+1} : |p| = 1\}$ ist eine C^1-Untermannigfaltigkeit von $\mathbb{R}^{n+1}$ der Dimension n, denn für jedes $p \in \mathbb{S}^n$ können wir im Niveaumengenkriterium $W = \mathbb{R}^{n+1} \backslash \{0\}$ und $h(q) = |q|^2 - 1$ wählen:

$$Dh(q) = 2q \neq 0 \text{ auf } W \quad \text{ und } \quad \mathbb{S}^n \cap W = \mathbb{S}^n = \{q \in W : h(q) = 0\}.$$

Wir brauchen nun einige topologische Tatsachen. Jede Menge $M \subset \mathbb{R}^{n+k}$ ist, versehen mit dem Abstand $d(p, q) = |p - q|$ für $p, q \in M$, ein metrischer Raum. Eine Menge $K \subset M$ ist genau dann kompakt in M, wenn sie kompakt in $\mathbb{R}^{n+k}$ ist. Dies ist offensichtlich, wenn wir den Begriff der Folgenkompaktheit verwenden.

Satz 9.3 (σ-Kompaktheit von Untermannigfaltigkeiten) *Jede n-dimensionale Untermannigfaltigkeit $M \subset \mathbb{R}^{n+k}$ ist als abzählbare Vereinigung $M = \bigcup_{i=1}^{\infty} K_i$ von kompakten Mengen darstellbar.*

Beweis: Zuerst konstruieren wir eine abzählbare, dichte Menge $P \subset M$. Betrachte die Würfel $Q_{j,k} = 2^{-k}(j + [0,1]^{n+k})$ mit $j \in \mathbb{Z}^{n+k}$, $k \in \mathbb{N}_0$, und wähle beliebige Punkte $p_{j,k} \in M \cap Q_{j,k}$, sofern dieser Schnitt nichtleer ist. Da jedes $p \in M$ in Würfeln $Q_{j,k}$ mit beliebig kleiner Kantenlänge liegt, ist die Menge P aller $p_{j,k}$ dicht in M. Als nächstes zeigen wir: ist $\phi : W \to \phi(W)$ Plättung bei p und $\overline{B_r(p)} \subset W$, so ist $M \cap \overline{B_r(p)}$ kompakt. Denn es gilt

$$\phi(M \cap \overline{B_r(p)}) = \phi(M \cap W) \cap \phi\big(\overline{B_r(p)}\big) = \mathbb{R}^n \cap \phi(W) \cap \phi\big(\overline{B_r(p)}\big) = \mathbb{R}^n \cap \phi\big(\overline{B_r(p)}\big).$$

Diese Menge ist kompakt, denn das Bild von $\overline{B_r(p)}$ unter der stetigen Abbildung ϕ ist kompakt, und $\mathbb{R}^n$ ist abgeschlossene Teilmenge des $\mathbb{R}^{n+k}$. Damit ist auch $M \cap \overline{B_r(p)} = \phi^{-1}\big(\mathbb{R}^n \cap \phi\big(\overline{B_r(p)}\big)$ kompakt, was zu zeigen war. Wir definieren nun für jedes $p \in M$

$$r(p) = \sup\{r > 0 : M \cap \overline{B_r(p)} \text{ ist kompakt}\} \in (0, \infty].$$

Wäre $r(q) + |p - q| < r(p)$, so können wir $r > r(q)$ und $s < r(p)$ wählen mit $r + |p - q| < s$ und $M \cap \overline{B_s(p)}$ kompakt. Es folgt $\overline{B_r(q)} \subset \overline{B_s(p)}$ und weiter

$$M \cap \overline{B_r(q)} = \big(M \cap \overline{B_s(p)}\big) \cap \overline{B_r(q)}.$$

Aber die rechte Seite ist kompakt im Widerspruch zu $r > r(q)$. Konvergiert also eine Folge $p_i \in M$ gegen $p \in M$, so folgt $\liminf_{i\to\infty} r(p_i) \geq r(p) > 0$, insbesondere $p \in M \cap \overline{B_{r(p_i)/2}(p_i)}$ für i hinreichend groß. Somit ist M Vereinigung der abzählbar vielen, kompakten Mengen $K(p) = M \cap \overline{B_{r(p)/2}(p)}$ mit $p \in P$. □

Definition 9.3 (lokale Parametrisierung) *Sei $M \subset \mathbb{R}^{n+k}$ eine n-dimensionale Untermannigfaltigkeit der Klasse C^1. Eine lokale Parametrisierung von M ist eine injektive Immersion $f : U \to M \subset \mathbb{R}^{n+k}$, wobei $U \subset \mathbb{R}^n$ offen, der Klasse C^1.*

Wir beabsichtigen, die Untermannigfaltigkeit M durch Bildgebiete von Parametrisierungen zu überdecken und das Maß in jedem einzelnen Bildgebiet mit der Flächenformel zu berechnen. Vorab sei daran erinnert, dass die offenen Mengen in M genau die Mengen $M \cap \Omega$ sind mit Ω offen im $\mathbb{R}^{n+k}$.

Lemma 9.1 *Für jede Untermannigfaltigkeit $M \subset \mathbb{R}^{n+k}$ der Klasse C^1 gibt es lokale C^1-Parametrisierungen $f_i : U_i \to M$, wobei $i \in \mathbb{N}$, so dass $M = \bigcup_{i=1}^\infty f(U_i)$.*

Beweis: Zu jedem $p \in M$ gibt es eine lokale Plättung $\phi : W \to \phi(W)$ mit $p \in W$. Dann ist $U = \mathbb{R}^n \cap \phi(W)$ offen in $\mathbb{R}^n$, und $f = \phi^{-1}|_{\mathbb{R}^n \cap \phi(W)}$ ist eine lokale Parametrisierung von M mit $p \in f(U)$. Außerdem ist das Bildgebiet $f(U) = M \cap W$ offen in M. Eine kompakte Menge $K \subset M$ wird also durch endlich viele Bildgebiete überdeckt. Die Behauptung folgt mit Satz 9.3. □

Die folgenden Eigenschaften von lokalen Parametrisierungen sind wesentlich.

Satz 9.4 *Für eine C^1-Untermannigfaltigkeit $M \subset \mathbb{R}^{n+k}$ gelten folgende Aussagen.*

(1) *Ist $f : U \to M$ eine lokale Parametrisierung von M, so ist $V = f(U)$ offen in M und $f^{-1} : V \to U$ ist stetig, also $f : U \to V$ homeomorph.*

(2) *Sind $f_i : U_i \to f(U_i) = V_i$ für $i = 1, 2$ lokale C^1-Parametrisierungen von M, so ist $f_2^{-1} \circ f_1 : f_1^{-1}(V_1 \cap V_2) \to f_2^{-1}(V_1 \cap V_2)$ ein C^1-Diffeomorphismus.*

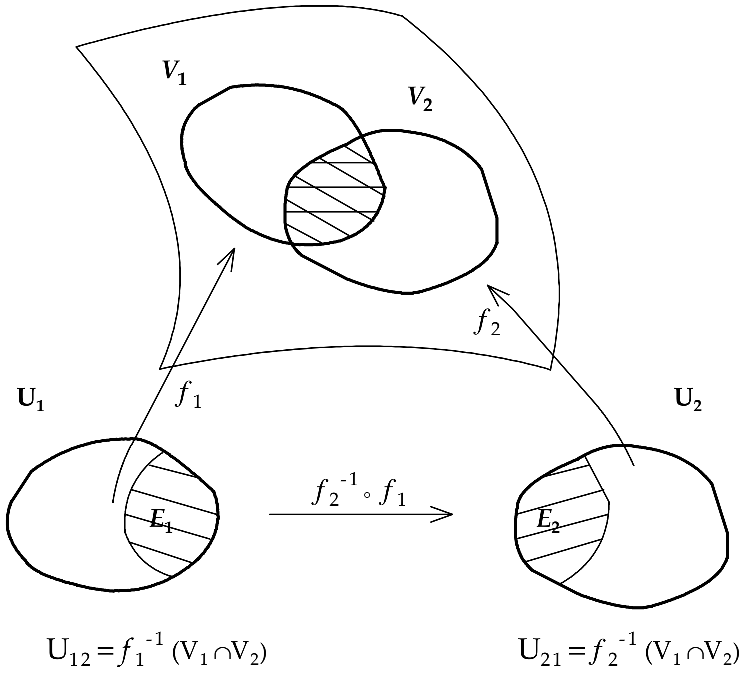
Beweis: Wir zeigen (1) zunächst unter der Annahme $V \subset W$, wobei $W \subset \mathbb{R}^{n+k}$ offen und $\phi(M \cap W) = \mathbb{R}^n \cap \phi(W)$ für einen C^1-Diffeomorphismus $\phi : W \to \phi(W)$. Dann ist die Abbildung $\phi \circ f : U \to \mathbb{R}^n \subset \mathbb{R}^{n+k}$ definiert, injektiv und es gilt

$$\operatorname{rang} D(\phi \circ f)(x) = \operatorname{rang}\big(D\phi(f(x))\, Df(x)\big) = n \quad \text{für alle } x \in U.$$

Also ist $Y = (\phi \circ f)(U) \subset \mathbb{R}^n$ offen, und $\phi \circ f : U \to Y$ ist ein C^1-Diffeomorphismus. Folglich ist $Z = (Y \times \mathbb{R}^k) \cap \phi(W)$ offen in $\mathbb{R}^{n+k}$ mit $\mathbb{R}^n \cap Z = Y$. Wegen $V = \phi^{-1}(Y)$ folgt

$$V = \phi^{-1}(\mathbb{R}^n \cap Z) = \phi^{-1}\big(\mathbb{R}^n \cap \phi(W)\big) \cap \phi^{-1}(Z) = M \cap \phi^{-1}(Z),$$

wobei zuletzt $\mathbb{R}^n \cap \phi(W) = \phi(M \cap W)$ benutzt wurde. Somit ist V offen in M, und die Abbildung $f^{-1} = (\phi \circ f)^{-1}\, \phi|_V$ ist stetig. Für V beliebig wählen wir zu $p \in V$ eine Plättung $\phi : W \to \phi(W)$ mit $p \in W$, und wenden das obige Argument an auf $\tilde{U} = U \cap \phi^{-1}(W)$ sowie $\tilde{f} = f|_{\tilde{U}}$. Dann ist $\tilde{V} = f(\tilde{U})$ offen in M mit $p \in \tilde{V}$, und $f^{-1}|_{\tilde{V}} = (f|_{\tilde{U}})^{-1}$ ist auf $\tilde{V}$ stetig.



Für (2) können wir ebenfalls annehmen, dass $V_i \subset W$ für $i = 1, 2$, wobei $\phi : W \to \phi(W)$ Diffeomorphismus mit $\phi(M \cap W) = \mathbb{R}^n \cap \phi(W)$. Dann sind die Abbildungen

$$\phi \circ f_i : f_i^{-1}(V_1 \cap V_2) \to \phi(V_1 \cap V_2)$$

diffeomorph, also auch $f_2^{-1} \circ f_1 = (\phi \circ f_2)^{-1} \circ (\phi \circ f_1)$. $\square$

Eine injektive Immersion $f : U \to \mathbb{R}^{n+k}$ heißt Einbettung, wenn sie offene Teilmengen von U in relativ offene Teilmengen von $f(U)$ abbildet. Dies ist äquivalent dazu, dass die Umkehrabbildung $f^{-1} : f(U) \to U$ stetig ist, also $f : U \to f(U)$ ein Homeomorphismus. Es gibt einfache Beispiele von injektiven Immersionen, die keine Einbettungen sind. Nach Satz 9.4(1) ist eine n-dimensionale, injektive Immersion mit Bild in einer n-dimensionalen Untermannigfaltigkeit automatisch eine Einbettung. Die Stetigkeit von f^{-1} muss also nicht explizit geprüft werden.

Ist M beliebige Teilmenge des $\mathbb{R}^{n+k}$, und gibt es zu jedem $p \in M$ eine n-dimensionale C^1-Einbettung $f : U \to f(U) \subset M$ mit $f(U)$ offen in M bezüglich der Relativtopologie, so ist M eine n-dimensionale Untermannigfaltigkeit der Klasse C^1. Dies ist ein viertes Kriterium, um eine Menge als Untermannigfaltigkeit zu identifizieren. Wie gesagt ist aber die Verifikation der Voraussetzungen, dass $f(U)$ offen in M ist und f eine Einbettung, unter Umständen schwierig; deshalb verzichten wir auf den Beweis.

Satz 9.5 (Definition des Flächenmaßes) *Sei $M \subset \mathbb{R}^{n+k}$ eine n-dimensionale Untermannigfaltigkeit der Klasse C^1. Dann heißt $E \subset M$ messbar, falls gilt:*

$$f^{-1}(E) \text{ ist } \mathcal{L}^n\text{-messbar für jede lokale Parametrisierung } f : U \to M.$$

Das System $\mathcal{M}$ der messbaren Teilmengen von M ist eine σ-Algebra, und enthält die Borelmnegen in M. Weiter gibt es auf $\mathcal{M}$ genau ein Maß μ_M, so dass für jede lokale Parametri-

sierung $f : U \to M$ und jede messbare Menge $E \subset f(U)$ gilt:

$$\mu_M(E) = \int_{f^{-1}(E)} Jf(x)\,dx.$$

Beweis: Trivialerweise gilt $\emptyset \in \mathcal{M}$. Für jede lokale Parametrisierung $f : U \to M$ gilt

$$f^{-1}(M \backslash E) = U \backslash f^{-1}(E) \quad \text{und} \quad f^{-1}\Big(\bigcup_{i=1}^{\infty} E_i\Big) = \bigcup_{i=1}^{\infty} f^{-1}(E_i).$$

Also ist $\mathcal{M}$ eine σ-Algebra. Ist $V \subset M$ offen, so ist $f^{-1}(V)$ offen im $\mathbb{R}^n$ aufgrund der Stetigkeit von f, das heißt $\mathcal{M}$ enthält die offenen Teilmengen von M und damit alle Borelmengen. Sei nun $M = \bigcup_{i=1}^{\infty} M_i$ eine disjunkte Zerlegung in Mengen $M_i \in \mathcal{M}$, so dass $M_i \subset V_i$ für lokale Parametrisierungen $f_i : U_i \to f(U_i) = V_i$. Eine solche Zerlegung existiert, denn nach Lemma 9.1 gibt es lokale Parametrisierungen $f_i : U_i \to f(U_i) = V_i$ mit $M = \bigcup_{i=1}^{\infty} V_i$, und wir können die Borelmengen $M_i = V_i \backslash \bigcup_{j=1}^{i-1} V_j$ wählen. Aus den verlangten Eigenschaften folgt für das gesuchte Maß μ_M und alle messbaren $E \subset M$

$$\mu_M(E) = \sum_{i=1}^{\infty} \mu_M(E \cap M_i) = \sum_{i=1}^{\infty} \int_{f_i^{-1}(E \cap M_i)} Jf_i(x)\,dx. \tag{9.4}$$

Dies beweist die Eindeutigkeit. Andererseits wird durch (9.4) ein Maß auf $\mathcal{M}$ definiert: ist $E = \bigcup_{j=1}^{\infty} E_j$ mit E_j messbar und paarweise disjunkt, so folgt nämlich

$$\mu_M(E) = \sum_{i=1}^{\infty} \int_{f_i^{-1}(E \cap M_i)} Jf_i(x)\,dx = \sum_{i,j=1}^{\infty} \int_{f_i^{-1}(E_j \cap M_i)} Jf_i(x)\,dx = \sum_{j=1}^{\infty} \mu_M(E_j).$$

Ist $f : U \to V = f(U)$ irgendeine lokale Parametrisierung und $E \subset V$ messbar, so ist $\phi_i = f_i^{-1} \circ f : f^{-1}(V \cap V_i) \to f_i^{-1}(V \cap V_i)$ ein C^1-Diffeomorphismus zwischen offenen Mengen nach Satz 9.4, und aus Satz 9.1 folgt

$$\begin{aligned}
\mu_M(E) &= \sum_{i=1}^{\infty} \int_{f_i^{-1}(E \cap M_i)} Jf_i(y)\,dy \\
&= \sum_{i=1}^{\infty} \int_{\phi_i(f^{-1}(E \cap M_i))} Jf_i(y)\,dy \\
&= \sum_{i=1}^{\infty} \int_{f^{-1}(E \cap M_i)} J(f_i \circ \phi_i)(x)\,dx \\
&= \int_{f^{-1}(E)} Jf(x)\,dx.
\end{aligned}$$

$\square$

In der Regel werden nur endlich viele Parametrisierungen benötigt, um ein Flächenintegral auszurechnen. Man kann sogar zeigen, dass jede kompakte Untermannigfaltigkeit bis auf eine μ_M-Nullmenge durch eine einzige Parametrisierung erfasst werden kann, aber diese Aussage hat eher theoretischen Charakter.

Folgerung 9.1 (Oberflächenintegral) *Sei $M \subset \mathbb{R}^{n+k}$ eine n-dimensionale Untermannigfaltigkeit der Klasse C^1, und $M = \bigcup_{i=1}^{\infty} M_i$ eine paarweise disjunkte, messbare Zerlegung, so dass $M_i \subset V_i$ für lokale Parametrisierungen $f_i : U_i \to V_i$. Für eine messbare Funktion $u : M \to \overline{\mathbb{R}}$ gilt die Formel*

$$\int_M u\, d\mu_M = \sum_{i=1}^{\infty} \int_{f_i^{-1}(M_i)} u(f_i(x)) Jf_i(x)\, dx,$$

wenn entweder u nichtnegativ ist oder wenn u bezüglich μ_M integrierbar ist.

BEWEIS: Die Aussage gilt nach Satz 9.5, wenn u eine charakteristische Funktion ist, und damit auch für messbare Treppenfunktionen. Für $u \geq 0$ folgt die Behauptung durch Approximation von unten mit Treppenfunktionen u_k aus dem Satz über monotone Konvergenz, der rechts erst auf die einzelnen Integrale und dann auf die Reihe angewandt wird. Für integrierbares u zerlegen wir in u^+ und u^-. □

Die gegebene Definition des Flächenmaßes ist gut geeignet, um auf einer festen Untermannigfaltigkeit M das Maß von Teilmengen bzw. Integrale von Funktionen zu berechnen. Dagegen ist die Konstruktion weiniger praktisch, wenn Aussagen über Folgen von Untermannigfaltigkeiten benötigt werden. Es kann auf $\mathbb{R}^{n+k}$ ein Maß $\mathcal{H}^n$ definiert werden – das n-dimensionale Hausdorffmaß –, so dass die Einschränkung auf jede n-dimensionale Untermannigfaltigkeit M das jeweilige Flächenmaß μ_M ergibt. Die Flächenformel ergibt sich bei diesem Zugang als Satz. Wir sind hier von der Flächenformel als Definition ausgegangen, um schneller Beispiele ausrechnen zu können. Die Nachteile unseres Zugangs zeigen sich auch bei folgender Aussage, die eigentlich eine Trivialität sein sollte.

Lemma 9.2 (Transformation des Flächeninhalts unter Ähnlichkeiten) *Sei $T : \mathbb{R}^{n+k} \to \mathbb{R}^{n+k}$ eine Ähnlichkeitsabbildung, das heißt es gibt $\lambda > 0$, $Q \in \mathbb{O}(n+k)$ und $a \in \mathbb{R}^{n+k}$ mit $T(p) = \lambda\, Q(p+a)$. Ist $M \subset \mathbb{R}^{n+k}$ eine C^1-Untermannigfaltigkeit, so ist $N = T(M)$ auch eine C^1-Untermannigfaltigkeit, und für die zugehörigen Maße μ_M bzw. μ_N gilt:*

(1) *Ist $A \subset M$ messbar, so ist $T(A) \subset N$ messbar und*

$$\mu_N(T(A)) = \lambda^n\, \mu_M(A).$$

(2) *Ist $u : N \to \overline{\mathbb{R}}$ μ_N-messbar, so ist $u \circ T : M \to \overline{\mathbb{R}}$ μ_M-messbar und es gilt, sofern eines der Integrale existiert,*

$$\int_N u(q)\, d\mu_N(q) = \lambda^n \int_M u(T(p))\, d\mu_M(p).$$

BEWEIS: Zu $q \in N$ wähle eine lokale Plättung $\phi : W \to \phi(W)$ von M mit $\phi^{-1}(q) \in W$. Dann ist $\phi \circ T^{-1} : T(W) \to \phi(W)$ eine lokale Plättung von N mit $q \in T(W)$, also ist N eine C^1-Untermannigfaltigkeit. Ist $f : U \to M$ lokale Parametrisierung, so auch $T \circ f : U \to N$ und umgekehrt. Wegen $(T \circ f)^{-1}\big(T(A)\big) = f^{-1}(A)$ ist $A \subset M$ messbar genau wenn $T(A) \subset N$ messbar ist. Zum Beweis von (1) können wir mittels Zerlegung wie in (9.4) annehmen, dass $A \subset f(U)$ für eine lokale Parametrisierung $f : U \to M$. Nun gilt wegen $DT(x) = \lambda Q$

$$D(T \circ f)(x)^T D(T \circ f)(x) = Df(x)^T DT(f(x))^T DT(f(x)) Df(x) = \lambda^2\, Df(x)^T Df(x).$$

Mit Satz 9.5 folgt

$$\mu_N(T(A)) = \int_{(T\circ f)^{-1}(T(A))} J(T\circ f)(x)\,dx = \lambda^n \int_{f^{-1}(A)} Jf(x)\,dx = \lambda^n\,\mu_M(A).$$

Für $u = \chi_B$ mit $B \subset N$ μ_N-messbar folgt (2) aus (1). Durch Approximation mit Treppenfunktionen von unten, siehe 5.3, ergibt sich (2) dann für $u \geq 0$, und schließlich durch Zerlegung in u^+ und u^- für beliebige μ_N-messbare Funktionen u. $\square$

Satz 9.6 (Zwiebelformel) *Für $u \in L^1(\mathbb{R}^{n+1})$ ist $u|_{\partial B_r} \in L^1(\mu_{\partial B_r})$ für fast alle $r > 0$, wobei $\partial B_r = \partial B_r(0)$, und es gilt*

$$\int_{\mathbb{R}^{n+1}} u(p)\,dp = \int_0^\infty \int_{\partial B_r} u(p)\,d\mu_{\partial B_r}(p)\,dr = \int_0^\infty r^n \int_{\mathbb{S}^n} u(r\omega)\,d\mu_{\mathbb{S}^n}(\omega)\,dr.$$

Beweis: Sei $f : U \to V \subset \mathbb{S}^n$ lokale Parametrisierung und $C(V) = \{r\omega : \omega \in V,\, r > 0\}$ der offene Kegel über V. Betrachte den Diffeomorphismus

$$\phi : (0,\infty)\times U \to C(V),\ \phi(r,x) = rf(x).$$

Mit $g_{ij}(x) = \langle \frac{\partial f}{\partial x_i}, \frac{\partial f}{\partial x_j}\rangle$ lautet die Gramsche Matrix von ϕ

$$D\phi(r,x)^T D\phi(r,x) = \begin{pmatrix} 1 & 0 \\ 0 & r^2 g(x)\end{pmatrix} \in \mathbb{R}^{(n+1)\times(n+1)}.$$

Ist $E = f(A)$ für $\mathcal{L}^n$-messbares $A \subset U$ und $C(E)$ der Kegel über E, so folgt aus dem Transformationssatz und Fubini

$$\begin{aligned}
\int_{C(E)} u(p)\,dp &= \int_{(0,\infty)\times A} u\big(rf(x)\big)\,r^n\sqrt{\det g(x)}\,d\mathcal{L}^{n+1}(r,x)\\
&= \int_0^\infty r^n \int_A u\big(rf(x)\big)\,\sqrt{\det g(x)}\,dx\,dr\\
&= \int_0^\infty r^n \int_E u(r\omega)\,d\mu_{\mathbb{S}^n}(\omega)\,dr\\
&= \int_0^\infty \int_{\{r\omega:\omega\in E\}} u(p)\,d\mu_{\partial B_r}(p)\,dr,
\end{aligned}$$

wobei zuletzt Lemma 9.2 benutzt wurde. Wähle nun eine disjunkte Zerlegung $\mathbb{S}^n = \bigcup_{j=1}^N E_j$ mit $E_j \subset V_j$, wobei $f_j : U_j \to V_j$ lokale Parametrisierungen sind. Durch Addition folgt die Behauptung. $\square$

Beispiel 9.7 Mit $u = \chi_{B_1(0)}$ folgt für das Maß $\omega_n = \mu_{\mathbb{S}^n}(\mathbb{S}^n)$ der n-dimensionalen Sphäre

$$\alpha_{n+1} = \mathcal{L}^{n+1}(B_1(0)) = \int_0^1 \mu_{\partial B_r}(\partial B_r)\,dr = \int_0^1 \omega_n r^n\,dr = \frac{\omega_n}{n+1},$$

also zum Beispiel $\omega_1 = 2\pi,\ \omega_2 = 4\pi$ und $\omega_3 = 2\pi^2$, vgl. Beispiel 7.3.

10 Der Integralsatz von Gauß

In diesem Abschnitt beweisen wir den Integralsatz von Gauß (Englisch: *divergence theorem*), der eine mehrdimensionale Verallgemeinerung des Hauptsatzes der Differential- und Integralrechnung ist. Aussage des Satzes ist, unter geeigneten technischen Voraussetzungen, die Formel

$$\int_\Omega \operatorname{div} X \, d\mathcal{L}^n = \int_{\partial\Omega} \langle X, \nu \rangle \, d\mu.$$

Dabei ist X ein Vektorfeld, ν die äußere Einheitsnormale auf dem Rand von Ω und μ das Flächenmaß auf $\partial\Omega$. In der Flüssigkeitsdynamik und der Elektrodynamik wird die Divergenz als Quellenstärke und das Randintegral als Fluss des Vektorfelds durch $\partial\Omega$ interpretiert. Aus dem Satz folgen dann entsprechende Erhaltungsgesetze.

Zuerst behandeln wir heuristisch den einfachen Fall, dass das zugrundeliegende Gebiet ein n-dimensionaler Quader Q ist, das heißt $Q = (a_1, b_1) \times \ldots \times (a_n, b_n)$. Die äußere Einheitsnormale sollte außerhalb der niederdimensionalen Kanten wie folgt gegeben sein:

$$\nu(x) = \begin{cases} -e_i & \text{für } x \in \partial Q \text{ mit } x_i = a_i \\ e_i & \text{für } x \in \partial Q \text{ mit } x_i = b_i. \end{cases}$$

Wir setzen $Q_i = (a_1, b_1) \times \ldots \times \widehat{(a_i, b_i)} \times \ldots \times (a_n, b_n) \subset \mathbb{R}^{n-1}$, wobei das Dach bedeutet, dass der Faktor wegzulassen ist. Für ein hinreichend glattes Vektorfeld $X : \overline{Q} \to \mathbb{R}^n$ berechnen wir mit dem Satz von Fubini und dem Hauptsatz der Differential- und Integralrechnung

$$\begin{aligned}
\int_Q \operatorname{div} X \, d\mathcal{L}^n &= \sum_{i=1}^n \int_Q \frac{\partial X_i}{\partial x_i} \, d\mathcal{L}^n \\
&= \sum_{i=1}^n \int_{Q_i} \int_{a_i}^{b_i} \frac{\partial X_i}{\partial x_i}(x_1, \ldots, x_i, \ldots, x_n) \, dx_i \; dx_1 \ldots \widehat{dx_i} \ldots dx_n \\
&= \sum_{i=1}^n \int_{Q_i} X_i(x_1, \ldots, b_i, \ldots, x_n) \, dx_1 \ldots \widehat{dx_i} \ldots dx_n \\
&\quad - \sum_{i=1}^n \int_{Q_i} X_i(x_1, \ldots, a_i, \ldots, x_n) \, dx_1 \ldots \widehat{dx_i} \ldots dx_n \\
&= \sum_{i=1}^n \int_{\{x_i = b_i\}} \langle X(x), e_i \rangle \, d\mu(x) - \sum_{i=1}^n \int_{\{x_i = a_i\}} \langle X(x), e_i \rangle \, d\mu(x) \\
&= \int_{\partial Q} \langle X, \nu \rangle \, d\mu.
\end{aligned}$$

Aber wir wollen natürlich die Aussage nicht nur für Quader zur Verfügung haben. Eine geeignete Klasse von Gebieten wird in folgendem Satz definiert. Der Beweis ist analog zur Dskussion der Untermannigfaltigkeitskriterien, siehe Satz 9.2 beziehungsweise Analysis II, Satz 9.2.

Satz 10.1 (Kriterien für C^1-Rand) *Für $\Omega \subset \mathbb{R}^n$ offen sind äquivalent:*

(1) Plättung: *zu jedem* $p \in \partial\Omega$ *gibt es eine offene Umgebung* $W \subset \mathbb{R}^n$ *und einen* C^1*-Diffeomorphismus* $\phi : W \to \phi(W)$ *mit*

$$\phi(\Omega \cap W) = \mathbb{H}^n \cap \phi(W), \quad \textit{wobei } \mathbb{H}^n = \mathbb{R}^{n-1} \times (-\infty, 0).$$

(2) Subniveau: *zu jedem* $p \in \partial\Omega$ *gibt es eine offene Umgebung* $W \subset \mathbb{R}^n$ *und eine Funktion* $h \in C^1(W)$ *mit* $Dh(q) \neq 0$ *für alle* $q \in W$*, so dass*

$$\Omega \cap W = \{q \in W : h(q) < 0\}.$$

(3) Subgraph: *zu jedem* $p \in \partial\Omega$ *gibt es eine offene Menge* $U \subset \mathbb{R}^{n-1}$*, ein offenes Intervall* $I \subset \mathbb{R}$ *und eine* C^1*-Funktion* $u : U \to I$*, so dass nach geeigneter Umnummerierung der Koordinaten gilt:*

$$\Omega \cap (U \times I) = \{(x, y) \in U \times I : y < u(x)\}.$$

Die Menge Ω *hat* C^1*-Rand, wenn eines (und damit jedes) der drei Kriterien erfüllt ist.*

In diesem Kapitel arbeiten wir ausschließlich mit dem Subgraphenkriterium (3). Anschaulich hat eine Menge C^1-Rand, wenn ihr Rand eine $(n-1)$-dimensionale Untermannigfaltigkeit ist, und die Menge lokal auf einer Seite des Randes liegt. Dies wird in folgendem Lemma präzisiert.

Lemma 10.1 *In der Situation von Satz 10.1(3) gilt*

$$\begin{aligned} \partial\Omega \cap (U \times I) &= \{(x, y) \in U \times I : y = u(x)\}, \\ (\mathbb{R}^n \backslash \overline{\Omega}) \cap (U \times I) &= \{(x, y) \in U \times I : y > u(x)\}. \end{aligned}$$

Insbesondere ist $\partial\Omega$ *eine* $(n-1)$*-dimensionale* C^1*-Untermannigfaltigkeit des* $\mathbb{R}^n$ *nach dem Graphenkriterium in Satz 9.2.*

Beweis: Mit der Funktion $F : U \times I \to \mathbb{R}$, $F(x, y) = y - u(x)$, gilt nach Voraussetzung $\Omega \cap (U \times I) = \{F < 0\}$. Aus der Stetigkeit von F folgt $\partial\Omega \cap (U \times I) \subset \{F = 0\}$. Ist andererseits $F(x, y) = 0$, so folgt für $\varepsilon > 0$ hinreichend klein $(x, y - \varepsilon) \in U \times I$ und

$$F(x, y - \varepsilon) = F(x, y) - \varepsilon < 0.$$

Dies zeigt $(x, y) \in \overline{\Omega}$, und wegen $(x, y) \notin \Omega$ folgt $(x, y) \in \partial\Omega \cap (U \times I)$. Damit ist die erste Behauptung bewiesen, und die zweite folgt wegen $\overline{\Omega} = \Omega \cup \partial\Omega$. $\square$

Beispiel 10.1 Der Halbraum $\mathbb{H}^n = \mathbb{R}^{n-1} \times (-\infty, 0)$ hat C^1-Rand, denn in Satz 10.1(3) können wir $U = \mathbb{R}^{n-1}$, $I = (-\infty, \infty)$ und $u : U \to I$, $u(x) \equiv 0$, wählen, und zwar für jeden Punkt $p \in \partial\mathbb{H}^n$. Dagegen hat $\Omega = \{x \in \mathbb{R}^n : x_n \neq 0\}$ keinen C^1-Rand, obwohl der Rand $\partial\Omega = \mathbb{R}^{n-1} \times \{0\}$ eine $(n-1)$-dimensionale Untermannigfaltigkeit ist. Denn für eine lokale Beschreibung als Subgraph wie in Satz 10.1(3) würde mit Lemma 10.1 folgen:

$$\{(x, y) \in U \times I : y > u(x)\} = (\mathbb{R}^n \backslash \overline{\Omega}) \cap (U \times I) = \emptyset,$$

ein Widerspruch wegen I offen und $u(x) \in I$ für $x \in U$.

Als nächstes erinnern wir an das Konzept des Tangentialraums. Ein Vektor $v \in \mathbb{R}^n$ heißt Tangentialvektor von $M \subset \mathbb{R}^n$ im Punkt $p \in M$, wenn es eine Abbildung $\gamma : (-\delta, \delta) \to M$ gibt mit $\gamma(0) = p$ und $\gamma'(0) = v$. Ist M eine m-dimensionale Untermannigfaltigkeit, so ist die Menge T_pM aller Tangentialvektoren in $p \in M$ ein m-dimensionaler Unterraum, der Tangentialraum von M in p, siehe Kapitel 9 in Analysis II. Die Hyperebene $T_p(\partial\Omega)$ besitzt genau zwei Einheitsnormalen, von denen wir jetzt eine auszeichnen.

Lemma 10.2 (Definition der äußeren Normale) *Sei $\Omega \subset \mathbb{R}^n$ offen mit C^1-Rand. Dann gibt es zu $p \in \partial\Omega$ genau einen Vektor $\nu(p) \in \mathbb{R}^n$ mit folgenden Eigenschaften:*

(1) $\nu(p) \perp T_p(\partial\Omega)$ *und* $|\nu(p)| = 1$,

(2) $p + t\nu(p) \notin \Omega$ *für* $t > 0$ *hinreichend klein.*

Das Vektorfeld $\nu : \partial\Omega \to \mathbb{R}^n$, $p \mapsto \nu(p)$, ist stetig und heißt äußere Normale von Ω.

Beweis: Wähle mit Satz 10.1(3) zu $p \in \partial\Omega$ nach eventueller Umnummerierung der Koordinaten eine Darstellung $\Omega \cap (U \times I) = \{(x,y) \in U \times I : y < u(x)\}$. Wir zeigen die Existenz und Eindeutigkeit der äußeren Normale gleich für alle $q \in \partial\Omega \cap (U \times I)$. Nach Lemma 10.1 ist die Graphenabbildung $f : U \to \mathbb{R}^n$, $f(x) = (x, u(x))$, eine lokale Parametrisierung von $\partial\Omega$. Der Tangentialraum $T_q(\partial\Omega)$ im Punkt $q = (x, u(x))$ hat die Basis

$$\big(e_i, \frac{\partial u}{\partial x_i}(x)\big) = \frac{d}{dt} f(x + te_i)|_{t=0} \quad \text{für } i = 1, \ldots, n-1.$$

Definiere nun auf $\partial\Omega \cap (U \times I)$ das Vektorfeld

$$\nu(q) = \frac{(-Du(x), 1)}{\sqrt{1 + |Du(x)|^2}} \quad \text{für } q = (x, u(x)). \tag{10.1}$$

Damit erfüllt $\nu(q)$ die Bedingung (1). Wir berechnen $q + t\nu(q) = (x(t), y(t)$ mit

$$x(t) = x - t\,\frac{Du(x)}{\sqrt{1 + |Du(x)|^2}} \quad \text{und} \quad y(t) = u(x) + t\,\frac{1}{\sqrt{1 + |Du(x)|^2}}.$$

Daraus folgt

$$\frac{d}{dt}\big(y(t) - u(x(t))\big)|_{t=0} = \frac{1}{\sqrt{1 + |Du(x)|^2}} + Du(x)\frac{Du(x)}{\sqrt{1 + |Du(x)|^2}} = \sqrt{1 + |Du(x)|^2} > 0.$$

Mit Lemma 10.1 folgt $q + t\,\nu(q) \notin \Omega$, sowie $q - t\nu(q) \in \Omega$, für $t > 0$ hinreichend klein, das heißt $\nu(q)$ ist der eindeutig bestimmte Vektor mit den Eigenschaften (1) und (2). Außerdem ergibt sich aus (10.1) die Stetigkeit von ν. $\square$

Definition 10.1 ($C^1(\overline{\Omega})$-Raum) *Für $\Omega \subset \mathbb{R}^n$ bezeichnen wir mit $C^1(\overline{\Omega})$ den Unterraum aller Funktionen $f \in C^1(\Omega)$, für die f und Df stetige Fortsetzungen auf $\overline{\Omega}$ besitzen. Es ist dabei üblich, die Fortsetzung von f wieder mit f zu bezeichnen.*

Wir zeigen nun eine lokale Fassung des Satzes von Gauß, die den Kern des Beweises ausmacht.

Lemma 10.3 *Sei $\Omega = \{(x,y) \in U \times I : y < u(x)\}$, wobei $U \subset \mathbb{R}^{n-1}$ offen, $I = (a,b) \subset \mathbb{R}$ und $u \in C^1(U,I)$. Hat das Vektorfeld $X \in C^1(\overline{\Omega}, \mathbb{R}^n)$ kompakten Träger in $U \times I$, so gilt*

$$\int_\Omega \operatorname{div} X \, d\mathcal{L}^n = \int_{\partial\Omega} \langle X, \nu \rangle \, d\mu.$$

Beweis: Da $X(x,a) = 0$ für alle $x \in U$, folgt mit Fubini

$$\int_\Omega \frac{\partial X_n}{\partial y} \, d\mathcal{L}^n = \int_U \int_a^{u(x)} \frac{\partial X_n}{\partial y}(x,y) \, dy \, dx = \int_U X_n\big(x, u(x)\big) \, dx. \tag{10.2}$$

Weiter behaupten wir für $i = 1, \ldots, n-1$

$$\int_\Omega \frac{\partial X_i}{\partial x_i} \, d\mathcal{L}^n = -\int_U X_i\big(x, u(x)\big) \frac{\partial u}{\partial x_i}(x) \, dx. \tag{10.3}$$

Aus (10.2) und (10.3) folgt das Lemma, denn wir haben

$$\begin{aligned}
\int_\Omega \operatorname{div} X \, d\mathcal{L}^n &= -\sum_{i=1}^{n-1} \int_U X_i\big(x, u(x)\big) \frac{\partial u}{\partial x_i}(x) \, dx + \int_U X_n\big(x, u(x)\big) \, dx \\
&= \int_U \Big\langle X(x, u(x)), \frac{(-Du(x), 1)}{\sqrt{1 + |Du(x)|^2}} \Big\rangle \sqrt{1 + |Du(x)|^2} \, dx \\
&= \int_{\partial\Omega} \langle X, \nu \rangle \, d\mu.
\end{aligned}$$

Dabei haben wir im letzten Schritt Folgerung 9.1 zur Berechnung des Oberflächenintegrals, Formel (9.1) für die Jacobische von Graphen und Formel (10.1) für die äußere Normale benutzt. Um Gleichung (10.3) zu verifizieren, wählen wir eine Abschneidefunktion $\eta \in C^\infty(\mathbb{R})$ mit $\eta(t) = 1$ für $t \leq -2$ und $\eta(t) = 0$ für $t \geq -1$, und setzen $\eta_\varepsilon(t) = \eta(t/\varepsilon)$. Es folgt $\eta_\varepsilon(y - u(x)) = 0$ für $y \geq u(x) - \varepsilon$ und

$$\lim_{\varepsilon \searrow 0} \eta_\varepsilon\big(y - u(x)\big)\big) = \begin{cases} 1 & \text{falls } y < u(x), \\ 0 & \text{sonst.} \end{cases}$$

Mit zweimaliger partieller Integration, siehe Satz 7.3, sehen wir für $i = 1, \ldots, n-1$

$$\begin{aligned}
\int_\Omega \frac{\partial X_i}{\partial x_i}(x,y) \eta_\varepsilon(y - u(x)) \, d\mathcal{L}^n(x,y) &= \int_\Omega X_i(x,y) \eta_\varepsilon'(y - u(x)) \frac{\partial u}{\partial x_i}(x) \, d\mathcal{L}^n(x,y) \\
&= -\int_\Omega \frac{\partial X_i}{\partial y}(x,y) \eta_\varepsilon(y - u(x)) \frac{\partial u}{\partial x_i}(x) \, d\mathcal{L}^n(x,y).
\end{aligned}$$

Da die beteiligten Funktionen beschränkt sind, folgt mit Lebesgue für $\varepsilon \searrow 0$

$$\begin{aligned}
\int_\Omega \frac{\partial X_i}{\partial x_i} \, d\mathcal{L}^n &= -\int_\Omega \frac{\partial X_i}{\partial y}(x,y) \frac{\partial u}{\partial x_i}(x) d\mathcal{L}^n(x,y) \\
&= -\int_U \Big(\int_a^{u(x)} \frac{\partial X_i}{\partial y}(x,y) \, dy \Big) \frac{\partial u}{\partial x_i}(x) \, dx \\
&= -\int_U X_i(x, u(x)) \frac{\partial u}{\partial x_i}(x) \, dx.
\end{aligned}$$

Das ist (10.3), also ist das Lemma bewiesen. □

Wir müssen schließlich das Resultat globalisieren. Das entscheidende Hilfsmittel ist dabei eine sogenannte Teilung der Eins, die nun konstruiert werden soll.

Lemma 10.4 (Teilung der Eins) *Sei W_λ, $\lambda \in \Lambda$, eine offene Überdeckung der kompakten Menge $K \subset \mathbb{R}^n$. Dann gibt es eine untergeordnete Teilung der Eins, das heißt es gibt eine endliche Familie von Funktionen $\chi_j \in C^\infty_c(\mathbb{R}^n)$, $j \in J$, so dass gilt:*

(1) $\sum_{j\in J} \chi_j(p) = 1$ *für alle* $p \in K$.

(2) *Für jedes* $j \in J$ *gibt es ein* $\lambda = \lambda(j)$ *mit* $\operatorname{spt}\chi_j \subset W_\lambda$.

Beweis: Wähle eine beschränkte offene Menge $\Omega \supset K$, und bestimme zu jedem $p \in \overline{\Omega}$ einen Ball $B_{r(p)}(p)$ wie folgt: für $p \in K$ wähle $\lambda(p) \in \Lambda$ mit $p \in W_{\lambda(p)}$, und weiter $r(p) > 0$ mit $\overline{B_{2r(p)}(p)} \subset \big(W_{\lambda(p)} \cap \Omega\big)$. Zu $p \in \overline{\Omega}\backslash K$ wähle $r(p) > 0$ mit $\overline{B_{2r(p)}(p)} \cap K = \emptyset$. Endlich viele Kugeln $B_{r(p_j)}(p_j)$, $1 \leq j \leq N$, überdecken $\overline{\Omega}$. Wähle nun $\tilde{\chi}_j \in C^\infty_c(\mathbb{R}^n)$ mit

$$\tilde{\chi}_j = \begin{cases} 1 & \text{auf } B_{r(p_j)}(p_j) \\ 0 & \text{auf } \mathbb{R}^n \backslash B_{2r(p_j)}(p_j). \end{cases}$$

Es folgt $\sum_{j=1}^N \tilde{\chi}_j \geq 1$ auf $\overline{\Omega}$. Die Funktionen $\chi_j = \tilde{\chi}_j / \big(\sum_{j=1}^N \tilde{\chi}_j\big)$ mit $j \in J := \{j : p_j \in K\}$ sind glatt in Ω mit $\operatorname{spt}\chi_j \subset\subset \big(W_{\lambda(p_j)} \cap \Omega\big)$, und wegen $\tilde{\chi}_j|_K = 0$ für $j \notin J$ gilt $\sum_{j\in J}\chi_j \equiv 1$ auf K. □

Satz 10.2 (Integralsatz von Gauß) *Sei $\Omega \subset \mathbb{R}^n$ eine offene, beschränkte Menge mit C^1-Rand und äußerer Normale $\nu : \partial\Omega \to \mathbb{R}^n$. Dann gilt für ein Vektorfeld $X \in C^1(\overline{\Omega}, \mathbb{R}^n)$*

$$\int_\Omega \operatorname{div} X \, d\mathcal{L}^n = \int_{\partial\Omega} \langle X, \nu\rangle \, d\omega.$$

Beweis: Wähle nach Satz 10.1(3) zu jedem $p \in \partial\Omega$ eine Umgebung W_p, in der Ω bezüglich geeigneter Koordinaten als Subgraph dargestellt ist. Für $p \in \Omega$ setze einfach $W_p = \Omega$. Die Mengen W_p bilden eine offene Überdeckung von $\overline{\Omega}$. Wähle mit Lemma 10.4 eine untergeordnete Teilung der Eins $\chi_1, \ldots, \chi_N \in C^\infty_c(\mathbb{R}^n)$. Liegt $\operatorname{spt}\chi_{\mathrm{j}}$ in einer Randumgebung $W_p = U \times I$ wie in 10.1(3), so folgt aus Lemma 10.3

$$\int_\Omega \operatorname{div}(\chi_j X)\, dx = \int_{\partial\Omega} \langle \chi_j X, \nu\rangle \, d\mu.$$

Ist $\operatorname{spt}\chi_{\mathrm{j}} \subset W_p = \Omega$, so folgt einfach mit partieller Integration, Satz 7.3,

$$\int_\Omega \operatorname{div}(\chi_j X)\, dx = 0 = \int_{\partial\Omega} \langle \chi_j X, \nu\rangle \, d\mu.$$

Durch Addition erhalten wir wie gewünscht

$$\int_\Omega \operatorname{div} X\, dx = \sum_{j=1}^N \int_\Omega \operatorname{div}(\chi_j X)\, dx = \sum_{j=1}^N \int_{\partial\Omega} \langle \chi_j X, \nu\rangle \, d\mu = \int_{\partial\Omega} \langle X, \nu\rangle \, d\mu.$$

□

Der Satz von Gauß wird oft für Gebiete benötigt, die nicht C^1-Rand haben, zum Beispiel Polyeder. Der gegebene Beweis kann auf Gebiete ausgedehnt werden, deren Rand lokal ein Lipschitzgraph ist*.

*siehe H.W. Alt, Lineare Funktionalanalysis.

Beispiel 10.2 Wählen wir im Satz von Gauß als Vektorfeld $X(x) = x$, so folgt

$$\mathcal{L}^n(\Omega) = \frac{1}{n}\int_\Omega \operatorname{div} X \, d\mathcal{L}^n = \frac{1}{n}\int_{\partial\Omega} \langle x, \nu(x)\rangle \, d\mu(x).$$

Insbesondere ergibt sich wieder $\alpha_n = \omega_{n-1}/n$, vergleiche Beispiel 9.7.

Folgerung 10.1 (Greensche Formeln) *Sei $\Omega \subset \mathbb{R}^n$ offen und beschränkt mit C^1-Rand. Dann gilt für $u \in C^1(\overline{\Omega})$ und $v \in C^2(\overline{\Omega})$*

$$\int_\Omega (u\Delta v + \langle Du, Dv\rangle) \, d\mathcal{L}^n = \int_{\partial\Omega} u \frac{\partial v}{\partial \nu} \, d\mu.$$

Weiter folgt für $u.v \in C^2(\overline{\Omega})$

$$\int_\Omega (u\Delta v - v\Delta u) \, d\mathcal{L}^n = \int_{\partial\Omega} \Big(u\frac{\partial v}{\partial \nu} - v\frac{\partial u}{\partial \nu}\Big) \, d\mu.$$

Beweis: Die erste Aussage folgt aus dem Satz von Gauß wegen $\operatorname{div}(uDv) = \langle Du, Dv\rangle + u\Delta v$. Die zweite Aussage ergibt sich aus der ersten durch Vertauschen von u und v. □

Beispiel 10.3 (Mittelwerteigenschaft harmonischer Funktionen) Sei $u \in C^2(\Omega)$ eine harmonische Funktion, das heißt $\Delta u = 0$ auf Ω. Aus dem Satz von Gauß folgt dann für $x_0 \in \Omega$ und $0 < r < \operatorname{dist}(x_0, \partial\Omega)$

$$\int_{\partial B_r(x_0)} \frac{\partial u}{\partial \nu} \, d\mu = \int_{B_r(x_0)} \Delta u \, dx = 0.$$

Betrachte weiter die auf $\mathbb{R}^n\backslash\{x_0\}$ harmonische Funktion $v(x) = \gamma(|x - x_0|)$ mit

$$\gamma(\varrho) = \begin{cases} \dfrac{\varrho^{2-n}}{2-n} & \text{für } n \geq 3, \\ \log \varrho & \text{für } n = 2. \end{cases}$$

Aus der zweiten Greenschen Formel folgt wegen $\gamma'(\varrho) = \varrho^{1-n}$

$$r^{1-n}\int_{\partial B_r(x_0)} u \, d\mu = \varepsilon^{1-n}\int_{\partial B_\varepsilon(x_0)} u \, d\mu.$$

Nun gilt aber für $\varepsilon \searrow 0$

$$\begin{aligned} \Big|\frac{1}{\omega_{n-1}\varepsilon^{n-1}}\int_{\partial B_\varepsilon(x_0)} u(x) \, d\mu(x) - u(x_0)\Big| &= \Big|\frac{1}{\omega_{n-1}\varepsilon^{n-1}}\int_{\partial B_\varepsilon(x_0)} \big(u(x) - u(x_0)\big) \, d\mu(x)\Big| \\ &\leq \sup_{|x-x_0|=\varepsilon} |u(x) - u(x_0)| \to 0. \end{aligned}$$

Dies beweist die sphärische Mittelwerteigenschaft harmonischer Funktionen:

$$u(x_0) = \frac{1}{\omega_{n-1}r^{n-1}}\int_{\partial B_r(x_0)} u(x) \, d\mu(x) \quad \text{für } 0 < r < \operatorname{dist}(x_0, \partial\Omega).$$

Mit Satz 9.6 folgt auch die Mittelwertformel auf Kugeln:

$$\begin{aligned}
\frac{1}{\alpha_n r^n}\int_{B_r(x_0)} u(x)\,dx &= \frac{1}{\alpha_n r^n}\int_0^r \int_{\partial B_\varrho(x_0)} u(x)\,d\mu_\varrho(x)\,d\varrho\\
&= \frac{u(x_0)}{\alpha_n r^n}\int_0^r \omega_{n-1}\varrho^{n-1}\,d\varrho\\
&= u(x_0).
\end{aligned}$$

Beispiel 10.4 (Strömungen) Sei $v \in C^1(G\times(a,b),\mathbb{R}^n)$, $v = v(x,t)$, ein evtl. zeitabhängiges Vektorfeld auf dem Gebiet $G\subset\mathbb{R}^n$. Für $x\in G$ sei $\phi^x : (a_x,b_x)\to G$ die maximale Lösung des Anfangswertproblems

$$\frac{d\phi^x}{dt}(t) = v(\phi^x(t),t) \quad \text{und} \quad \phi^x(0) = x.$$

Anschaulich interpretieren wir v als Geschwindigkeitsfeld einer Strömung und $\phi^x(t)$ als Bahnkurve eines Partikels, das zur Zeit $t=0$ an der Stelle x ist. Zusammenfassen aller Bahnkurven ergibt eine C^1-Abbildung, den Fluss von v,

$$\phi : \{(x,t)\in G\times(a,b) : a_x < t < b_x\}\to G,\ \phi(x,t) = \phi^x(t).$$

Die Mengen $G_t = \{x\in G : t\in(a_x,b_x)\}$ sind offen, und die Abbildungen $\phi_t : G_t\to G_{-t}$, $\phi_t(x) = \phi(x,t)$, sind Diffeomorphismen mit

$$\phi_s\circ\phi_t = \phi_{s+t} \quad \text{auf } G_t\cap G_{s+t}.$$

Zur Zeit $t=0$ gilt $\frac{\partial}{\partial t}(D_x\phi) = D_x\big(\frac{\partial\phi}{\partial t}\big) = D_x v$, insbesondere wegen $D_x\phi(x,0) = \mathrm{Id}_{\mathbb{R}^n}$

$$\frac{\partial}{\partial t}(\det D_x\phi)|_{t=0} = \operatorname{tr} D_x v = \operatorname{div} v.$$

Sei nun $\varrho\in C^1(G\times(a,b))$ die evtl. auch zeitabhängige Dichteverteilung der Strömung. Ist $\Omega\subset\subset G$, so gilt $\Omega\subset G_t$ für $|t|$ hinreichend klein. Es bezeichne $m_\Omega(t)$ die im Gebiet $\phi_t(\Omega)$ enthaltene Masse. Mit dem Transformationssatz und Differentiation unter dem Integral folgt

$$\begin{aligned}
m_\Omega'(0) &= \frac{d}{dt}\int_{\phi_t(\Omega)} \varrho(y,t)\,dy|_{t=0}\\
&= \frac{d}{dt}\int_\Omega \varrho(\phi_t(x),t)\,\det D_x\phi(x,t)\,dx|_{t=0}\\
&= \int_\Omega \Big(D_x\varrho\cdot v + \frac{\partial\varrho}{\partial t} + \varrho\operatorname{div} v\Big)\,dx\\
&= \int_\Omega \Big(\frac{\partial\varrho}{\partial t} + \operatorname{div}(\varrho v)\Big)\,dx.
\end{aligned}$$

Für beliebige $t\in(a,b)$ mit $\Omega\subset\subset G_t$ ergibt sich aus dem Fall $t=0$

$$m_\Omega'(t) = \frac{d}{ds}\int_{\phi_s(\phi_t(\Omega))} \varrho(y,s+t)\,dy|_{s=0} = \int_{\phi_t(\Omega)} \Big(\frac{\partial\varrho}{\partial t} + \operatorname{div}(\varrho v)\Big)\,dx.$$

Gilt daher auf $G\times(a,b)$ die Kontinuitätsgleichung

$$\frac{\partial\varrho}{\partial t} + \operatorname{div}(\varrho v) = 0,$$

so ist die Funktion $m_\Omega(t)$ für alle $\Omega \subset\subset G$ zeitlich konstant. Wählen wir $\varrho \equiv 1$, so ist $m_\Omega(t)$ das Volumen von $\phi_t(\Omega)$, und der Satz von Gauß impliziert

$$m'_\Omega(0) = \int_{\partial\Omega} \langle v, \nu \rangle d\mu.$$

Das Flächenelement $d\mu$ trägt mit der Normalkomponente $\langle v, \nu \rangle$ zur Änderung des eingeschlossenen Volumens bei, was eine sehr anschauliche Deutung des Gaußschen Satzes ist.

Beispiel 10.5 (Integralsatz von Cauchy) Sei $\Omega \subset \mathbb{C} \cong \mathbb{R}^2$ beschränktes Gebiet mit C^1-Rand. Die Funktion $f \in C^1(\overline{\Omega}, \mathbb{C})$, $f = u + iv$, sei holomorph, das heißt es gelten die Cauchy-Riemannschen Differentialgleichungen

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad \text{und} \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}.$$

Sei $\gamma : [0, L] \to \partial\Omega$ die Parametrisierung einer Komponente von $\partial\Omega$ nach der Bogenlänge, so dass Ω links von γ liegt, das heißt die äußere Normale längs $\gamma(s) = \big(x(s), y(s)\big)$ ist $\nu(s) = \big(y'(s), -x'(s)\big)$. Aus der Definition des komplexen Kurvenintegrals folgt

$$\begin{aligned} \int_\gamma f\,dz &= \int_0^L (ux' - vy')\,ds + i\int_0^L (uy' + vx')\,ds \\ &= \int_0^L \langle(-v, -u), (y', -x')\rangle\,ds + i\int_0^L \langle(u, -v), (y', -x')\rangle\,ds. \end{aligned}$$

Also impliziert der Satz von Gauß

$$\begin{aligned} \int_{\partial\Omega} f(z)\,dz &= \int_\Omega \operatorname{div}(-v, -u)\,dxdy + i\int_\Omega \operatorname{div}(u, -v)\,dxdy \\ &= -\int_\Omega \Big(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\Big)\,dxdy + i\int_\Omega \Big(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}\Big)\,dxdy \\ &= 0. \end{aligned}$$

11 Faltung und Fouriertransformation

Es werden die Faltung und die Fouriertransformation für Funktionen im $\mathbb{R}^n$ als Anwendungen der Integrationstheorie behandelt. Die Faltung ordnet zwei gegebenen Funktionen eine dritte Funktion durch gewichtete Mittelung zu. Dieses Verfahren kann unter anderem dazu benutzt werden, gegebene Funktionen zu glätten. Die zentrale Aussage zur Fouriertransformation ist der Satz von Plancherel, der die Rückberechnung einer Funktion aus ihrer Fouriertransformierten erlaubt. Als Anwendung berechnen wir die Lösung der Wärmeleitungsgleichung zu gegebenen Anfangsdaten im $\mathbb{R}^n$, und diskutieren das entsprechende Problem für die Wellengleichung.

In diesem Kapitel haben wir es ausschließlich mit dem n-dimensionalen Lebesguemaß zu tun, und wir schreiben statt $d\mathcal{L}^n(x)$ stets einfach dx.

Satz 11.1 (Definition der Faltung) *Sei $f \in L^p(\mathbb{R}^n)$, $1 \leq p \leq \infty$, und $g \in L^1(\mathbb{R}^n)$. Die Faltung von f mit g ist die $\mathcal{L}^n$-fast-überall definierte Funktion*

$$f * g : \mathbb{R}^n \to \overline{\mathbb{R}},\ (f * g)(x) = \int_{\mathbb{R}^n} f(x-y)g(y)\,dy.$$

Es gilt $f * g \in L^p(\mathbb{R}^n)$ sowie $\|f * g\|_{L^p} \leq \|f\|_{L^p}\,\|g\|_{L^1}$.

Beweis: Die Funktion $F : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $F(x,y) = f(x-y)g(y)$ ist messbar bezüglich $\mathcal{L}^{2n} = \mathcal{L}^n \times \mathcal{L}^n$. Denn $f_0(x,y) = f(x)$ und $g_0(x,y) = g(y)$ sind $\mathcal{L}^{2n}$-messbar, und wegen $f(x-y) = (f_0 \circ T)(x,y)$ mit $T(x,y) = (x-y,y)$ ist $(x,y) \mapsto f(x-y)g(y)$ ebenfalls $\mathcal{L}^{2n}$-messbar nach Satz 4.5. Wir zeigen die Behauptung nun zunächst für $f,g \geq 0$. Nach dem Satz von Fubini, Satz 7.2, ist die Funktion $(f*g)(x) = \int_{\mathbb{R}^n} F(x,y)\,dy$ messbar, und mit der Hölderschen Ungleichung folgt

$$\begin{aligned}
\|f * g\|_{L^p}^p &= \int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n} f(x-y)g(y)^{\frac{1}{p}}\, g(y)^{\frac{p-1}{p}}\,dy\Big)^p dx\\
&\leq \int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n} f(x-y)^p\, g(y)\,dy\Big)\Big(\int_{\mathbb{R}^n} g(y)\,dy\Big)^{p-1} dx\\
&= \int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n} f(x-y)^p\,dx\Big) g(y)\,dy \cdot \Big(\int_{\mathbb{R}^n} g(y)\,dy\Big)^{p-1}\\
&= \|f\|_{L^p}^p\,\|g\|_{L^1}^p < \infty.
\end{aligned}$$

Für $f \in L^p(\mathbb{R}^n)$, $g \in L^1(\mathbb{R}^n)$ beliebig folgt durch Zerlegung in $f^\pm$ bzw. $g^\pm$, dass die Funktion $f * g$ für $\mathcal{L}^n$-fast-alle $x \in \mathbb{R}^n$ definiert und endlich ist, sowie $\mathcal{L}^n$-messbar. Der Satz ergibt sich nun aus der Abschätzung

$$\|f * g\|_{L^p}^p \leq \int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n} |f(x-y)|\,|g(y)|\,dy\Big)^p dx \leq \|f\|_{L^p}^p\,\|g\|_{L^1}^p.$$

$\square$

Die Faltung ist kommutativ, denn mit der Substitution $x - y = z$ folgt

$$(f * g)(x) = \int_{\mathbb{R}^n} f(x-y)g(y)\,dy = \int_{\mathbb{R}^n} f(z)g(x-z)\,dz = (g * f)(x). \tag{11.1}$$

Lemma 11.1 *Sei $f \in L^p(\mathbb{R}^n)$ mit $1 \leq p \leq \infty$, und $\tau_h : \mathbb{R}^n \to \mathbb{R}^n$, $\tau_h(x) = x + h$. Dann gelten folgende Aussagen.*

(i) $f \circ \tau_h \in L^p(\mathbb{R}^n)$ *und* $\|f \circ \tau_h\|_{L^p} = \|f\|_{L^p}$.

(ii) $f \circ \tau_h \to f$ *in* $L^p(\mathbb{R}^n)$ *für* $h \to 0$, *falls* $p < \infty$.

BEWEIS: Aussage (i) ist trivial. Wir zeigen (ii) zunächst unter der Annahme $f \in C^0_c(\mathbb{R}^n)$. Dann gilt $\operatorname{osc}(f,\delta) := \sup_{|x-y|\leq\delta} |f(x) - f(y)| \searrow 0$ für $\delta \searrow 0$, folglich

$$\|f \circ \tau_h - f\|_{C^0(\mathbb{R}^n)} = \sup_{x\in\mathbb{R}^n} |f(x+h) - f(x)| \leq \operatorname{osc}(f, |h|) \to 0 \quad \text{mit } h \to 0.$$

Wähle $R > 0$ mit $\operatorname{spt} f \subset B_R(0)$. Dann gilt $\operatorname{spt}(f \circ \tau_h) \subset B_{R+|h|}(0)$, also für $|h| < 1$

$$\|f \circ \tau_h - f\|_{L^p} \leq \|f \circ \tau_h - f\|_{C^0(\mathbb{R}^n)}\, \mathcal{L}^n(B_{R+1}(0))^{\frac{1}{p}} \to 0 \quad \text{mit } h \to 0.$$

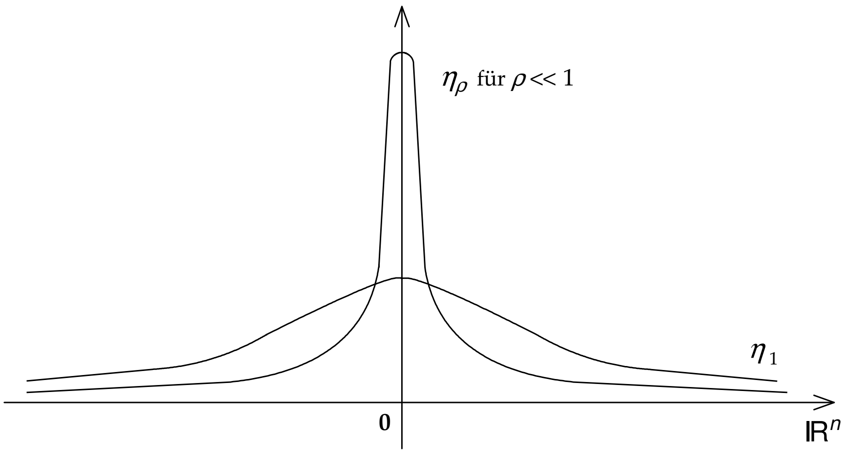
Sei nun $f \in L^p(\mathbb{R}^n)$ beliebig. Nach Satz 6.10 ist $C^0_c(\mathbb{R}^n)$ dicht in $L^p(\mathbb{R}^n)$, das heißt zu $\varepsilon > 0$ gibt es ein $f_\varepsilon \in C^0_c(\mathbb{R}^n)$ mit $\|f - f_\varepsilon\|_{L^p} < \varepsilon/2$, und es folgt mit Aussage (i)

$$\|f \circ \tau_h - f\|_{L^p} \leq \|(f - f_\varepsilon) \circ \tau_h\|_{L^p} + \|f_\varepsilon \circ \tau_h - f_\varepsilon\|_{L^p} + \|f_\varepsilon - f\|_{L^p} < \|f_\varepsilon \circ \tau_h - f_\varepsilon\|_{L^p} + \varepsilon.$$

Somit gilt $\limsup_{h\to 0} \|f \circ \tau_h - f\|_{L^p} \leq \varepsilon$, und Behauptung (ii) folgt mit $\varepsilon \searrow 0$. □

Satz 11.2 (Approximation durch Faltung) *Sei* $f \in L^p(\mathbb{R}^n)$ *mit* $1 \leq p < \infty$. *Ist* $\eta \in L^1(\mathbb{R}^n)$ *mit* $\int_{\mathbb{R}^n} \eta(z)\,dz = 1$, *so folgt für* $\eta_\varrho(x) = \varrho^{-n}\eta(x/\varrho)$

$$\|f * \eta_\varrho\|_{L^p} \leq \|f\|_{L^p}\, \|\eta\|_{L^1} \quad \textit{und} \quad f * \eta_\varrho \to f \ \textit{in } L^p(\mathbb{R}^n).$$



BEWEIS: Durch Substitution sieht man $\|\eta_\varrho\|_{L^1} = \|\eta\|_{L^1}$, daher gilt $f * \eta_\varrho \in L^p(\mathbb{R}^n)$ und $\|f * \eta_\varrho\|_{L^p} \leq \|f\|_{L^p}\, \|\eta\|_{L^1}$ nach Satz 11.1. Weiter folgt mit der Substitution $y = \varrho z$

$$(f * \eta_\varrho)(x) = \int_{\mathbb{R}^n} f(x-y)\eta_\varrho(y)\,dy = \int_{\mathbb{R}^n} f(x - \varrho z)\eta(z)\,dz.$$

Wir schätzen mit der Hölderschen Ungleichung und dem Satz von Fubini wie folgt ab:

$$\begin{aligned}
\int_{\mathbb{R}^n} |(f * \eta_\varrho)(x) - f(x)|^p \, dx &= \int_{\mathbb{R}^n} \Big| \int_{\mathbb{R}^n} \big(f(x - \varrho z) - f(x)\big)\eta(z)\, dz \Big|^p \, dx \\
&\leq \int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n} |f(x-\varrho z) - f(x)|\, |\eta(z)|^{\frac{1}{p}}\, |\eta(z)|^{\frac{p-1}{p}}\, dz \Big)^p \, dx \\
&\leq \|\eta\|_{L^1}^{p-1} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |f(x-\varrho z) - f(x)|^p \, |\eta(z)|\, dz\, dx \\
&= \|\eta\|_{L^1}^{p-1} \int_{\mathbb{R}^n} |\eta(z)| \int_{\mathbb{R}^n} |f(x-\varrho z) - f(x)|^p \, dx\, dz \\
&= \|\eta\|_{L^1}^{p-1} \int_{\mathbb{R}^n} |\eta(z)|\, \|f \circ \tau_{-\varrho z} - f\|_{L^p}^p \, dz.
\end{aligned}$$

Im letzten Integral geht der Integrand punktweise gegen Null mit $\varrho \to 0$ nach Lemma 11.1(ii). Außerdem gilt die Abschätzung

$$|\eta(z)|\, \|f \circ \tau_{-\varrho z} - f\|_{L^p}^p \leq 2^p \, \|f\|_{L^p}^p \, |\eta(z)| \in L^1(\mathbb{R}^n).$$

Also konvergiert das Integral gegen Null nach dem Satz über majorisierte Konvergenz. □

Die Approximation ist besonders nützlich, wenn die Funktion η glatt gewählt wird. Wir erinnern an die Multiindexnotation für Ableitungen von Funktionen im $\mathbb{R}^n$, und zwar setzt man für $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}_0^n$

$$|\alpha| = \alpha_1 + \ldots + \alpha_n \quad \text{und} \quad D^\alpha = \partial_1^{\alpha_1} \ldots \partial_n^{\alpha_n}.$$

Satz 11.3 (Glättung) *Sei $\eta \in C^k(\mathbb{R}^n)$ mit $\|D^\alpha \eta\|_{C^0(\mathbb{R}^n)} \leq K$ für $|\alpha| \leq k$. Für $f \in L^1(\mathbb{R}^n)$ ist dann $f * \eta \in C^k(\mathbb{R}^n)$ und es gilt*

$$D^\alpha (f * \eta) = f * (D^\alpha \eta), \ \textit{insbesondere} \ \|D^\alpha (f * \eta)\|_{C^0(\mathbb{R}^n)} \leq K \, \|f\|_{L^1}.$$

Beweis: Nach der Substitution $z = x - y$ haben wir

$$(f * \eta)(x) = \int_{\mathbb{R}^n} F(x,z)\, dz \quad \text{mit } F(x,z) = f(z)\eta(x-z).$$

Im Fall $k = 0$ gilt $F(x,\cdot) \in L^1(\mathbb{R}^n)$ für alle $x \in \mathbb{R}^n$, die Funktion $F(\cdot, z)$ ist stetig für $\mathcal{L}^n$-fast-alle $z \in \mathbb{R}^n$ und wir haben die Abschätzung

$$\sup_{x \in \mathbb{R}^n} |F(x,z)| \leq K \, |f(z)| \in L^1(\mathbb{R}^n).$$

Also ist $\eta * f$ stetig nach Satz 6.5 und es gilt $\|f * \eta\|_{C^0(\mathbb{R}^n)} \leq K \, \|f\|_{L^1}$. Im Fall $k = 1$ gilt $F(\cdot, z) \in C^1(\mathbb{R}^n)$ für $\mathcal{L}^n$-fast-alle $z \in \mathbb{R}^n$ sowie

$$\sup_{x \in \mathbb{R}^n} \Big| \frac{\partial F}{\partial x_j}(x,z) \Big| \leq K \, |f(z)| \in L^1(\mathbb{R}^n).$$

Aus Folgerung 6.1 folgt $f * \eta \in C^1(\mathbb{R}^n)$ und $\partial_j (f * \eta) = \int_{\mathbb{R}^n} f(z) \partial_j \eta(x - z)\, dz = (f * \partial_j \eta)(z)$, insbesondere gilt die Abschätzung $\|\partial j (f * \eta)\|_{C^0(\mathbb{R}^n)} \leq K \, \|f\|_{L^1}$. Die Aussage für eine Ableitung D^α mit Ordnung $|\alpha| \leq k$ ergibt sich in offensichtlicher Weise durch Induktion. □

Wir können jetzt das Dichteresultat aus Satz 6.10 verschärfen.

Satz 11.4 (Dichtheit von $C_c^\infty(\Omega)$ in $L^p(\Omega)$) *Sei $\Omega \subset \mathbb{R}^n$ offen und $1 \leq p < \infty$. Dann gibt es zu $f \in L^p(\Omega)$ eine Folge $f_k \in C_c^\infty(\Omega)$ mit $\|f - f_k\|_{L^p(\Omega)} \to 0$.*

Beweis: Wegen Satz 6.10 können wir $f \in C_c^0(\Omega)$ annehmen. Wähle $\eta \in C_c^\infty(\mathbb{R}^n)$ mit $\eta \geq 0$, $\int \eta(z)\,dz = 1$ und $\operatorname{spt}\eta \subset \overline{B_1(0)}$. Mit $\eta_\varrho(x) = \varrho^{-n}\eta(\frac{x}{\varrho})$ gilt dann $f * \eta_\varrho \in C^\infty(\mathbb{R}^n)$ nach Satz 11.3 und $f * \eta_\varrho \to f$ in $L^p(\mathbb{R}^n)$ nach Satz 11.2. Weiter gilt für $x \in \mathbb{R}^n$ mit $\operatorname{dist}(x, \operatorname{spt} f) > \varrho$

$$(f * \eta_\varrho)(x) = \int_{\mathbb{R}^n} f(x-y)\eta_\varrho(y)\,dy = 0,$$

denn für $|y| \geq \varrho$ ist $\eta_\varrho(y) = \varrho^{-n}\eta(y/\varrho) = 0$ und für $|y| \leq \varrho$ ist $f(x-y) = 0$. Also folgt

$$\operatorname{spt}(f * \eta_\varrho) \subset \{x \in \mathbb{R}^n : \operatorname{dist}(x, \operatorname{spt} f) \leq \varrho\}, \tag{11.2}$$

das heißt $f * \eta_\varrho \in C_c^\infty(\Omega)$ für $\varrho > 0$ hinreichend klein, womit der Satz bewiesen ist. □

Das folgende Ergebnis verallgemeinert Lemma 8.2. Es ist dabei sinnvoll, die Aussage mit lokal integrierbaren Funktionen zu formulieren.

Definition 11.1 *Sei $\Omega \subset \mathbb{R}^n$ offen und $1 \leq p \leq \infty$. Die messbare Funktion $f : \Omega \to \mathbb{R}$ liegt in $L^p_{\mathrm{loc}}(\Omega)$, falls $\chi_K f \in L^p(\Omega)$ ist für alle kompakten Mengen $K \subset \Omega$.*

Satz 11.5 (Fundamentallemma der Variationsrechnung) *Sei $\Omega \subset \mathbb{R}^n$ offen. Für die Funktion $f \in L^1_{\mathrm{loc}}(\Omega)$ gelte*

$$\int_\Omega f(x)\varphi(x)\,dx \geq 0 \ \textit{ für alle } \varphi \in C_c^\infty(\Omega) \textit{ mit } \varphi \geq 0.$$

Dann folgt $f(x) \geq 0$ für $\mathcal{L}^n$-fast-alle $x \in \Omega$.

Beweis: Zu zeigen ist, dass $E = \{x \in \Omega : f(x) < 0\}$ eine Nullmenge ist. Nach Satz 4.3 gilt

$$\mathcal{L}^n(E) = \sup\{\mathcal{L}^n(K) : K \subset E \text{ kompakt}\}.$$

Sei $K \subset E$ kompakt und $\eta \in C^\infty(\mathbb{R}^n)$ mit $\eta \geq 0$, $\operatorname{spt}\eta \subset \overline{B_1(0)}$ und $\int \eta(z)\,dz = 1$. Dann folgt $\eta_\varrho * \chi_K \to \chi_K$ in $L^1(\mathbb{R}^n)$, insbesondere $\eta_{\varrho_i} * \chi_K \to \chi_K$ punktweise $\mathcal{L}^n$-fast-überall für eine teilfolge nach Folgerung 6.2. Wegen $|\eta_{\varrho_i} * \chi_K| \leq 1$ liefert der Konvergenzsatz von Lebesgue

$$0 \leq \lim_{i\to\infty} \int f(x)(\eta_{\varrho_i} * \chi_K)(x)\,dx = \int f(x)\chi_K(x)\,dx.$$

Aber $f < 0$ auf K, also $\mathcal{L}^n(K) = 0$ nach Folgerung 5.3 und damit $\mathcal{L}^n(\{f < 0\}) = 0$. □

Faltungen mit singulären Integralkernen spielen bei partiellen Differentialgleichungen eine große Rolle. Ein prominentes Beispiel ist das Newtonpotential.

Beispiel 11.1 (Newtonpotential) Sei $G \in C^\infty(\mathbb{R}^n\backslash\{0\})$. Für eine Funktion $f \in L^\infty(\Omega)$, mit $\Omega \subset \mathbb{R}^n$ offen und beschränkt, betrachten wir die Faltung

$$u : \mathbb{R}^n \to \mathbb{R},\ u(x) = \int_\Omega G(x-y)f(y)\,dy.$$

Es ist $u = G * f$, wobei f durch Null auf $\mathbb{R}^n$ fortgesetzt wird. Es gelte zunächst

$$|G(z)| \leq C\,|z|^{-p} \quad \text{für ein } p < n, \tag{11.3}$$

das heißt $u(x)$ ist für alle $x \in \mathbb{R}^n$ definiert nach Beispiel 5.5. Um zu zeigen, dass u stetig ist, wählen wir $\varphi \in C^\infty(\mathbb{R}^n)$ mit $\varphi(z) = 0$ für $|z| \leq 1/2$, $\varphi(z) = 1$ für $|z| \geq 1$ und $|D^k\varphi| \leq C_k$, und definieren den abgeschnittenen Kern

$$G_\varepsilon = \varphi_\varepsilon G \quad \text{ mit } \varphi_\varepsilon(z) = \varphi\Big(\frac{z}{\varepsilon}\Big).$$

Aus (11.3) folgt die Abschätzung

$$|G(z) - G_\varepsilon(z)| \leq C\chi_{B_\varepsilon(0)}|z|^{-p}. \tag{11.4}$$

Die Funktion $u_\varepsilon : \mathbb{R}^n \to \mathbb{R}$, $u_\varepsilon(x) = \int_\Omega G_\varepsilon(x-y)f(y)\,dy$, ist glatt, und es folgt

$$|u(x) - u_\varepsilon(x)| \leq C \int_{B_\varepsilon(x)} |x-y|^{-p}|f(y)|\,dy \leq \frac{C}{n-p}\,\|f\|_{L^\infty}\,\varepsilon^{n-p} \to 0.$$

Somit ist u stetig als gleichmäßiger Grenzwert der u_ε. Nun gelte zusätzlich

$$|DG(z)| \leq C\,|z|^{-p-1} \quad \text{ für ein } p < n-1. \tag{11.5}$$

Dann folgt wegen $(D\varphi_\varepsilon)(z) = \frac{1}{\varepsilon}D\varphi(\frac{z}{\varepsilon})$

$$|DG(z) - DG_\varepsilon(z)| \leq C\chi_{B_\varepsilon(0)}\Big(|z|^{-p-1} + \frac{1}{\varepsilon}|z|^{-p}\Big) \leq C\chi_{B_\varepsilon(0)}|z|^{-p-1}. \tag{11.6}$$

Mit $v_j(x) = \int_\Omega (\partial_j G)(x-y)f(y)\,dy$ ergibt sich

$$\begin{aligned}
\big|v_j(x) - \partial_j u_\varepsilon(x)\big| &\leq \int_\Omega |(\partial_j G)(x-y) - (\partial_j G_\varepsilon)(x-y)|\,|f(y)|\,dy \\
&\leq \frac{C}{n-p-1}\,\|f\|_{L^\infty}\,\varepsilon^{n-p-1} \to 0.
\end{aligned}$$

Also konvergiert $\partial_j u_\varepsilon$ gleichmäßig auf $\mathbb{R}^n$ gegen v_j, und es folgt $\partial_j u = v_j \in C^0(\mathbb{R}^n)$. Wir spezialisieren jetzt auf das Newtonpotential

$$G(z) = \begin{cases} \dfrac{|z|^{2-n}}{(2-n)\omega_{n-1}} & \text{für } n \geq 3, \\ \frac{1}{2\pi}\log|z| & \text{für } n = 2. \end{cases}$$

Im Fall $n \geq 3$ sind (11.3) und (11.5) mit $p = n-2$ erfüllt, im Fall $n = 2$ gilt (11.3) für jedes $p > 0$ und (11.5) für $p = 0$. Somit ist das Newtonpotential von f in $C^1(\mathbb{R}^n)$, und

$$\partial_j u(x) = \int_\Omega H(x-y)f(y)\,dy \quad \text{ mit } H(z) = \partial_j G(z) = \frac{1}{\omega_{n-1}}\frac{z_j}{|z|^n}.$$

Auf $\mathbb{R}^n\backslash\overline{\Omega}$ kann weiter unter dem Integral differenziert werden, also ist u auf $\mathbb{R}^n\backslash\overline{\Omega}$ glatt und

$$\Delta u(x) = \int_\Omega (\Delta G)(x-y)\,f(y)\,dy = 0 \quad \text{ für } x \in \mathbb{R}^n\backslash\overline{\Omega}.$$

Im Innern von Ω kann aber wegen der Singularität so nicht argumentiert werden. Das Ergebnis wäre sogar falsch, denn u ist nicht harmonisch auf Ω, sondern löst dort die Poissongleichung $\Delta u = f$, was wir nun zeigen wollen. Wie oben approximieren wir $\partial_j u = v_j$ durch

$$v_{j,\varepsilon} : \mathbb{R}^n \to \mathbb{R},\ v_{j,\varepsilon}(x) = \int_\Omega H_\varepsilon(x-y)f(y)\,dy, \quad \text{ wobei } H_\varepsilon = \varphi_\varepsilon H.$$

Da (11.3) für H mit $p = n-1$ erfüllt ist, konvergiert $v_{j,\varepsilon}$ gleichmäßig auf $\mathbb{R}^n$ gegen $v_j = \partial_j u$. Das Problem ist nun, dass die Ungleichung in (11.5) nur mit $p = n-1$ richtig ist, das heißt wir haben nur die Abschätzung

$$|DH(z) - DH_\varepsilon(z)| \leq C\chi_{B_\varepsilon(0)}\Big(|z|^{-n} + \frac{1}{\varepsilon}|z|^{1-n}\Big) \leq C\chi_{B_\varepsilon(0)}|z|^{-n}. \tag{11.7}$$

Um die fehlende Integrierbarkeit zu kompensieren, setzen wir zusätzlich voraus, dass f Hölderstetig mit einem Exponenten $\alpha \in (0,1]$ ist, also

$$[\,f\,]_\alpha := \sup_{x,y\in\Omega,\, x\neq y} \frac{|f(x)-f(y)|}{|x-y|^\alpha} < \infty. \tag{11.8}$$

Sei $U \subset \mathbb{R}^n$ offen und beschränkt mit C^1-Rand, so dass $\Omega \subset\subset U$, und $x \in \Omega$. Dann gilt, wobei im letzten Schritt der Satz von Gauß auf das Vektorfeld $y \mapsto H_\varepsilon(x-y)e_i$ angewandt wird,

$$\begin{aligned}
\partial_i v_{j,\varepsilon}(x) &= \int_\Omega (\partial_i H_\varepsilon)(x-y)\, f(y)\, dy \\
&= \int_U (\partial_i H_\varepsilon)(x-y)\,\big(f(y)-f(x)\big)\, dy + f(x)\int_U (\partial_i H_\varepsilon)(x-y)\, dy \\
&= \int_U (\partial_i H_\varepsilon)(x-y)\,\big(f(y)-f(x)\big)\, dy - f(x)\int_{\partial U} H_\varepsilon(x-y)\langle \nu(y), e_i\rangle\, d\mu(y).
\end{aligned}$$

Nun ist $H = \partial_j G$, und $H_\varepsilon(x-y) = \partial_j G(x-y)$ für $y \in \partial U$ und $\varepsilon > 0$ klein. Setze

$$w_{ij}(x) = \int_U (\partial^2_{ij} G)(x-y)\big(f(y)-f(x)\big)\, dy - f(x)\int_{\partial U} (\partial_j G)(x-y)\langle \nu(y), e_i\rangle\, d\mu(y).$$

Die $\partial_i v_{j,\varepsilon}$ konvergieren gleichmäßig auf Ω gegen w_{ij}, denn aus (11.7) und (11.8) folgt

$$\begin{aligned}
\big|w_{ij}(x) - \partial_i v_{j,\varepsilon}(x)\big| &\leq \int_U \big|\partial_i H(x-y) - \partial_i H_\varepsilon(x-y)\big|\,\big|f(y)-f(x)\big|\, dy \\
&\leq \frac{C}{\alpha}\,[\,f\,]_\alpha\, \varepsilon^\alpha \to 0.
\end{aligned}$$

Somit gilt $\partial^2_{ij} u = w_{ij} \in C^0(\Omega)$. Insbesondere folgt, da $\Delta G = 0$ auf $\mathbb{R}^n\backslash\{0\}$,

$$\begin{aligned}
\Delta u(x) &= -f(x)\int_{\partial U} \langle DG(x-y), \nu(y)\rangle\, d\mu(y) \\
&= -f(x)\int_{\partial B_\varepsilon(x)} \langle DG(x-y), \nu(y)\rangle\, d\mu(y) \\
&= f(x).
\end{aligned}$$

Damit ist die Poisson-Gleichung $\Delta u = f$ in Ω verifiziert. Es gibt stetige Funktionen f, für die keine zweimal differenzierbare Lösung der Gleichung $\Delta u = f$ existiert, die Voraussetzung der Hölderstetigkeit ist also nicht überflüssig.

Wir kommen zum Abschluss der Vorlesung zur Fourierintegraldarstellung, die wir zuerst heuristisch als kontinuierlichen Grenzwert der Entwicklung in eine Fourierreihe herleiten. Sei $f \in C^\infty_c(\mathbb{R})$ und $N > 0$ so groß, dass $\operatorname{spt} f \subset (-N\pi, N\pi)$. Die normierten Basisfunktionen mit Periode $2\pi N$ lauten

$$w^N_k(x) = \frac{1}{\sqrt{2\pi N}} e^{i\frac{k}{N}x} \quad \text{für } k \in \mathbb{Z}.$$

Nach Satz 6.12 wird die Funktion f auf $(-N\pi, N\pi)$ durch ihre Fourierreihe dargestellt, das heißt für jedes $x \in (-N\pi, N\pi)$ gilt

$$f(x) = \sum_{k=-\infty}^{\infty} \frac{1}{2\pi N} e^{i\frac{k}{N}x} \int_{\mathbb{R}} f(y) e^{-i\frac{k}{N}y}\, dy =: \sum_{k=-\infty}^{\infty} F\left(\frac{k}{N}\right) \frac{1}{N},$$

wobei

$$F(p) = \frac{1}{2\pi} e^{ipx} \int_{\mathbb{R}} f(y) e^{-ipy}\, dy.$$

Falls nun $F(p)$ für $p \to \pm\infty$ hinreichend schnell gegen Null geht, sollte die Riemannsche Summe $\sum_{k=-\infty}^{\infty} F\left(\frac{k}{N}\right) \frac{1}{N}$ für $N \to \infty$ gegen das Integral $\int_{\mathbb{R}} F(p)\, dp$ konvergieren. Dies motiviert für $f \in C_c^\infty(\mathbb{R})$ die Darstellung

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{f}(p) e^{ipx}\, dp \quad \text{mit} \quad \hat{f}(p) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(y) e^{-ipy}\, dy.$$

Der geschilderte Ansatz kann zu einem vollen Beweis ausgebaut werden, wir gehen aber anders vor. Alle im Folgenden auftretenden Funktionen sind $\mathbb{C}$-wertig.

Definition 11.2 *Die Fouriertransformierte von $f \in L^1(\mathbb{R}^n; \mathbb{C})$ ist die Funktion*

$$\hat{f} : \mathbb{R}^n \to \mathbb{C}, \quad \hat{f}(p) = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} f(x) e^{-i\langle p, x\rangle}\, dx.$$

Die inverse Fouriertransformierte von $g \in L^1(\mathbb{R}^n; \mathbb{C})$ ist

$$\check{g} : \mathbb{R}^n \to \mathbb{C}, \quad \check{g}(x) = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} g(p) e^{i\langle p, x\rangle}\, dp.$$

Bevor wir die Frage angehen, inwiefern diese beiden Transformationen invers sind, notieren wir einige wichtige Eigenschaften.

Satz 11.6 (Fouriertransformation auf L^1) *Für $f, g \in L^1(\mathbb{R}^n; \mathbb{C})$ gilt:*

(1) $\hat{f} \in C^0(\mathbb{R}^n; \mathbb{C})$ *und* $\|\hat{f}\|_{C^0(\mathbb{R}^n)} \leq (2\pi)^{-\frac{n}{2}} \|f\|_{L^1}$.

(2) $\widehat{f * g} = (2\pi)^{\frac{n}{2}} \hat{f}\hat{g}$.

(3) $\langle \hat{f}, g \rangle_{L^2(\mathbb{R}^n)} = \langle f, \check{g} \rangle_{L^2(\mathbb{R}^n)}$.

Beweis: Für $F(p, x) = f(x) e^{-i\langle p, x\rangle}$ ist $F(p, \cdot) \in L^1(\mathbb{R}^n; \mathbb{C})$ für alle $p \in \mathbb{R}^n$ sowie $F(\cdot, x) \in C^0(\mathbb{R}^n; \mathbb{C})$ für $\mathcal{L}^n$-fast-alle $x \in \mathbb{R}^n$, und es gilt

$$\sup_{p \in \mathbb{R}^n} |F(x, p)| = |f(x)| \in L^1(\mathbb{R}^n).$$

Also folgt die Stetigkeit von $\hat{f}$ aus der entsprechenden Aussage für Parameterintegrale, Satz 6.5. Die Abschätzung in (1) ist offensichtlich. Für (2) berechnen wir mit Fubini

$$\begin{aligned}
\widehat{f * g}(p) &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(y) g(x - y)\, dy\, e^{-i\langle p, x\rangle}\, dx \\
&= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} f(y) e^{-i\langle p, y\rangle} \int_{\mathbb{R}^n} g(x - y) e^{-i\langle p, x - y\rangle}\, dx\, dy \\
&= \int_{\mathbb{R}^n} f(y) e^{-i\langle p, y\rangle} \hat{g}(p)\, dy \\
&= (2\pi)^{\frac{n}{2}} \hat{f}(p)\hat{g}(p).
\end{aligned}$$

Auch (3) folgt aus dem Satz von Fubini, und zwar gilt

$$\begin{aligned}
\int_{\mathbb{R}^n} \hat{f}(p)\overline{g(p)}\,dp &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(x) e^{-i\langle p,x\rangle}\,dx\,\overline{g(p)}\,dp \\
&= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} f(x)\overline{\int_{\mathbb{R}^n} g(p) e^{i\langle p,x\rangle}\,dp}\,dx \\
&= \int_{\mathbb{R}^n} f(x)\overline{\check{g}(x)}\,dx.
\end{aligned}$$

$\square$

Wir berechnen nun zwei Beispiele – das erste werden wir im Beweis von Satz 11.7 verwenden.

Beispiel 11.2 Für die Gaußfunktion

$$G : \mathbb{R}^n \to \mathbb{R},\ G(x) = e^{-|x|^2/2},$$

gilt $\hat{G} = G = \check{G}$. Um dies zunächst für $n = 1$ zu zeigen, berechnen wir mit Differentiation unter dem Integral und partieller Integration

$$\begin{aligned}
\frac{d}{dp}\int_{\mathbb{R}} e^{-x^2/2} e^{-ixp}\,dx &= \int_{\mathbb{R}} (-ix) e^{-x^2/2} e^{-ixp}\,dx \\
&= \int_{\mathbb{R}} \left(\frac{d}{dx}\, e^{-x^2/2}\right) i e^{-ixp}\,dx \\
&= (-p)\int_{\mathbb{R}} e^{-x^2/2} e^{-ixp}\,dx,
\end{aligned}$$

das heißt $\big(e^{p^2/2}\hat{G}(p)\big)' \equiv 0$. Aber nach Beispiel 8.2 gilt

$$\hat{G}(0) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-x^2/2}\,dx = 1,$$

also folgt $\hat{G} = G$. Für n beliebig folgt mit Fubini für $x = (x', x_n) \in \mathbb{R}^{n-1} \times \mathbb{R}$ induktiv

$$\hat{G}(p) = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^{n-1}} e^{-|x'|^2/2 - i\langle x',p'\rangle} \int_{\mathbb{R}} e^{-(x_n)^2/2 - ix_np_n}\,dx_n\,dx' = e^{-|p'|^2/2} e^{-(p_n)^2/2} = G(p).$$

Schließlich ergibt sich $\check{G}(x) = \hat{G}(-x) = G(x)$ wie behauptet.

Beispiel 11.3 Für $a > 0$ gilt

$$\widehat{\chi_{[-a,a]}}(p) = \begin{cases} \sqrt{\dfrac{2}{\pi}}\,\dfrac{\sin(pa)}{p} & \text{für } p \neq 0 \\ \sqrt{\dfrac{2}{\pi}}\,a & \text{für } p = 0. \end{cases}$$

Das zweite Beispiel zeigt, dass die Fouriertransformierte einer L^1-Funktion im allgemeinen nicht wieder in L^1 liegt. Aufgrund dieser Asymmetrie kann das Problem der Fourierintegraldarstellung im L^1-Kontext nicht befriedigend behandelt werden. Der richtige Raum ist $L^2(\mathbb{R}^n;\mathbb{C})$, denn bezüglich der L^2-Norm ist die Fouriertransformation sogar isometrisch.

Satz 11.7 (Plancherel) *Für $f \in (L^1 \cap L^2)(\mathbb{R}^n;\mathbb{C})$ gilt $\|\hat{f}\|_{L^2} = \|\check{f}\|_{L^2} = \|f\|_{L^2}$.*

Beweis: Es reicht die Aussage für $\hat{f}$ zu beweisen, da $\check{f}(x) = \hat{f}(-x)$. Sei $G(z) = e^{-|z|^2/2}$ und $G_\varrho(z) = \varrho^{-n} G\left(\frac{z}{\varrho}\right)$ für $\varrho > 0$. Dann folgt aus Beispiel 11.2

$$\frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} G(\varrho p) e^{-i\langle p,z\rangle}\,dp = \frac{\varrho^{-n}}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} G(q) e^{-i\langle q,\frac{z}{\varrho}\rangle}\,dq = \varrho^{-n}\,\hat{G}\left(\frac{z}{\varrho}\right) = G_\varrho(z).$$

Wir berechnen nun mit dem Satz über monotone Konvergenz und dem Satz von Fubini

$$\begin{aligned}
\|\hat{f}\|_{L^2}^2 &= \lim_{\varrho \searrow 0} \int_{\mathbb{R}^n} |\hat{f}(p)|^2 G(\varrho p)\,dp \\
&= \frac{1}{(2\pi)^n} \lim_{\varrho \searrow 0} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(x)\overline{f(y)} e^{-i\langle p,x-y\rangle}\,dy\,dx\,G(\varrho p)\,dp \\
&= \frac{1}{(2\pi)^n} \lim_{\varrho \searrow 0} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} G(\varrho p) e^{-i\langle p,x-y\rangle}\,dp\,f(x)\overline{f(y)}\,dy\,dx \\
&= \frac{1}{(2\pi)^{\frac{n}{2}}} \lim_{\varrho \searrow 0} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(x)\overline{f(y)} G_\varrho(x-y)\,dy\,dx \\
&= \frac{1}{(2\pi)^{\frac{n}{2}}} \lim_{\varrho \searrow 0} \int_{\mathbb{R}^n} f(x)(\overline{f} * G_\varrho)(x)\,dx.
\end{aligned}$$

Aus Beispiel 8.2 folgt mit Fubini $\int_{\mathbb{R}^n} G(z)\,dz = (2\pi)^{\frac{n}{2}}$. Aber nach Satz 11.2 konvergiert dann $(2\pi)^{-n/2}\,\overline{f} * G_\varrho$ gegen $\overline{f}$ in $L^2(\mathbb{R}^n;\mathbb{C})$, und es folgt $\|\hat{f}\|_{L^2} = \|f\|_{L^2}$. □

Der Satz von Plancherel erlaubt es, die Fouriertransformation auf den Raum $L^2(\mathbb{R}^n;\mathbb{C})$ fortzusetzen, obwohl das in der Definition gegebene Integral nicht notwendig konvergiert.

Satz 11.8 (Fouriertransformation auf L^2) *Es gibt eindeutig bestimmte Abbildungen $\mathcal{F},\ \mathcal{F}^* : L^2(\mathbb{R}^n;\mathbb{C}) \to L^2(\mathbb{R}^n;\mathbb{C})$ mit folgenden Eigenschaften:*

(1) $\mathcal{F}f = \hat{f}$ *und* $\mathcal{F}^* f = \check{f}$ *für alle* $f \in (L^1 \cap L^2)(\mathbb{R}^n;\mathbb{C})$,

(2) $\|\mathcal{F}f\|_{L^2} = \|f\|_{L^2} = \|\mathcal{F}^* f\|_{L^2}$ *für alle* $f \in L^2(\mathbb{R}^n;\mathbb{C})$.

Weiter gelten für $f, g \in L^2(\mathbb{R}^n;\mathbb{C})$ folgende Aussagen:

(3) $\langle \mathcal{F}f, \mathcal{F}g\rangle_{L^2} = \langle f, g\rangle_{L^2} = \langle \mathcal{F}^* f, \mathcal{F}^* g\rangle_{L^2}$.

(4) $\langle \mathcal{F}f, g\rangle_{L^2} = \langle f, \mathcal{F}^* g\rangle_{L^2}$.

(5) $\mathcal{F}^*\mathcal{F} = \mathcal{F}\mathcal{F}^* = \mathrm{Id}_{L^2(\mathbb{R}^n)}$.

Beweis: Der Raum $(L^1 \cap L^2)(\mathbb{R}^n;\mathbb{C})$ ist dicht in $L^2(\mathbb{R}^n;\mathbb{C})$, denn für $f \in L^2(\mathbb{R}^n;\mathbb{C})$ ist $\chi_{B_R(0)} f \in L^1(\mathbb{R}^n;\mathbb{C})$ nach Cauchy-Schwarz, und es gilt $\chi_{B_R(0)} f \to f$ in $L^2(\mathbb{R}^n;\mathbb{C})$ mit $R \to \infty$. Die Eindeutigkeit und Existenz der Abbildungen $\mathcal{F}$ bzw. $\mathcal{F}^*$ mit (1) und (2) folgt damit leicht aus Satz 11.7. In dem komplexen Skalarproduktraum $L^2(\mathbb{R}^n;\mathbb{C})$ gilt die Polarisationsformel

$$\langle f, g\rangle_{L^2} = \frac{1}{4}\left(\|f+g\|_{L^2}^2 - \|f-g\|_{L^2}^2 + i\|f+ig\|_{L^2}^2 - i\|f-ig\|_{L^2}^2\right).$$

Also folgt (3) aus (2). Da in (4) beide Seiten stetig bzgl. der L^2-Norm sind, folgt die Aussage aus Satz 11.6(3) durch Approximation. Schließlich ergibt sich aus (4) und (3)

$$\langle \mathcal{F}^*\mathcal{F}f, g\rangle_{L^2} = \langle \mathcal{F}f, \mathcal{F}g\rangle_{L^2} = \langle f, g\rangle_{L^2},$$

also $\mathcal{F}^*\mathcal{F}f = f$. Die Gleichung $\mathcal{F}\mathcal{F}^*f = f$ folgt analog. □

Im folgenden schreiben wir $\hat{f}$ bzw. $\check{g}$ auch dann, wenn f, g nur in $L^2(\mathbb{R}^n)$ liegen. Die Fouriertransformation spielt eine wichtige Rolle in der Theorie linearer partieller Differentialgleichungen. Dies basiert vor allem darauf, dass Ableitungsoperatoren in Multiplikationsoperatoren transformiert werden. Es ist oft praktisch, dabei im sogenannten Raum der schnell fallenden Funktionen oder Schwartz-Raum zu operieren:

$$\mathcal{S}(\mathbb{R}^n;\mathbb{C}) = \{f \in C^\infty(\mathbb{R}^n;\mathbb{C}) : x^\alpha D^\beta f \text{ ist beschränkt für alle } \alpha, \beta \in \mathbb{N}_0^n\} \tag{11.9}$$

Wie üblich ist hier $x^\alpha = x_1^{\alpha_1} \cdot \ldots \cdot x_n^{\alpha_n}$ und $D^\beta = \partial_1^{\beta_1} \ldots \partial_n^{\beta_n}$. Für $f \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$ gilt

$$|f(x)| \leq C_N\,(1 + |x|^2)^{-N/2} \qquad \text{für jedes } N \in \mathbb{N},$$

insbesondere ist $f \in L^p(\mathbb{R}^n;\mathbb{C})$ für alle $p \in [1,\infty]$. Außerdem sind mit $f \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$ auch $\partial_j f$ und $x_j f$ in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$ für $1 \leq j \leq n$. Im Gegensatz zum Raum $C_c^\infty(R^n;\mathbb{C})$ ist der Schwartz-Raum $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$ invariant unter der Fouriertransformation, wie wir jetzt zeigen.

Satz 11.9 (Fouriertransformation auf $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$) *Mit f ist auch $\hat{f} \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$ und*

(1) $\widehat{\partial_j f}(p) = i\,p_j \hat{f}(p)$ *sowie* $\widehat{x_j f}(p) = i\,\partial_j \hat{f}(p)$.

(2) $f(x) = \dfrac{1}{(2\pi)^{\frac{n}{2}}} \displaystyle\int_{\mathbb{R}^n} \hat{f}(p) e^{i\langle p,x\rangle}\,dp$ *für alle $x \in \mathbb{R}^n$.*

Beweis: Die Funktion $\hat{f}$ ist beschränkt nach Satz 11.6(1). Wir berechnen mit partieller Integration, siehe Satz 7.3,

$$\widehat{\partial_j f}(p) = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} (\partial_j f)(x)\,e^{-i\langle p,x\rangle}\,dx = -\frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} f(x)\,\frac{\partial}{\partial x_j} e^{-i\langle p,x\rangle}\,dx = i\,p_j\,\hat{f}(p).$$

Weiter folgt durch Herausziehen der Ableitung aus dem Integral

$$\widehat{x_j f}(p) = \frac{i}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} f(x) \frac{\partial}{\partial p_j} e^{-i\langle p,x\rangle}\,dx = \frac{i}{(2\pi)^{\frac{n}{2}}} \frac{\partial}{\partial p_j} \int_{\mathbb{R}^n} f(x) e^{-i\langle p,x\rangle}\,dx = i\,\partial_j \hat{f}(p).$$

Insbesondere sind auch die Funktionen $p_j \hat{f}$ und $\partial_j \hat{f}$ beschränkt, wieder nach Satz 11.6(1). Durch Induktion über $|\alpha| + |\beta|$ verifiziert man leicht die allgemeine Formel

$$\widehat{D^\alpha x^\beta f} = i^{|\alpha|+|\beta|}\,p^\alpha D^\beta \hat{f},$$

und erhält $\hat{f} \in \mathcal{S}(\mathbb{R}^n)$. Aus Satz 11.8 folgt Behauptung (2) zunächst für $\mathcal{L}^n$-fast-alle $x \in \mathbb{R}^n$, wegen der Stetigkeit beider Seiten also sogar für alle $x \in \mathbb{R}^n$. Damit ist der Satz bewiesen. □

Wir wollen zum Schluss des Kapitels noch zwei Anwendungen auf die Lösung von linearen partiellen Differentialgleichungen mit konstanten Koeffizienten behandeln. Typischerweise wird dabei zunächst mit einem Fourieransatz eine Lösungsformel ermittelt, wobei die Eigenschaften der Fouriertransformation formal angewandt werden. In einem zweiten Schritt wird dann geprüft, unter welchen Voraussetzungen an die Daten die Formel eine gültige Lösung liefert. Dieser Punkt kann technisch anspruchsvoll sein – je nach Gleichung – und wird hier nur angedeutet.

Beispiel 11.4 Betrachte das Anfangswertproblem für die Wärmeleitungsgleichung im $\mathbb{R}^n$:

$$\begin{aligned} \partial_t u - \Delta u = 0 \qquad & \text{auf } \mathbb{R}^n \times (0,\infty), \\ u(\cdot,0) = f \qquad & \text{auf } \mathbb{R}^n. \end{aligned}$$

Wir bilden die Fouriertransformierte $\hat{u}(\cdot,t) = \widehat{u(\cdot,t)}$ bezüglich der räumlichen Variablen. Mit Satz 11.9 erhalten wir für $\hat{u}$ das Anfangswertproblem

$$\begin{aligned} 0 = \widehat{\partial_t u} - \widehat{\Delta u} &= \partial_t \hat{u} + |p|^2 \hat{u} \quad \text{auf } \mathbb{R}^n \times (0,\infty), \\ \hat{f} &= \hat{u}(0,\cdot) \quad \text{auf } \mathbb{R}^n. \end{aligned}$$

Es folgt $\hat{u}(p,t) = \hat{f}(p)\, e^{-t|p|^2}$. Aber nach Beispiel 11.2 gilt für $G_{\sqrt{2t}}(x) = (2t)^{-\frac{n}{2}}\, G\left(\frac{x}{\sqrt{2t}}\right)$

$$\widehat{G_{\sqrt{2t}}}(p) = \hat{G}(\sqrt{2t}\, p) = G(\sqrt{2t}\, p) = e^{-t|p|^2}.$$

Aus Satz 11.6(2) folgt nun $u = (2\pi)^{-\frac{n}{2}}\, f * G_{\sqrt{2t}}$, das heißt

$$u(x,t) = \frac{1}{(4\pi t)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4t}}\, f(y)\, dy.$$

Es ist leicht zu sehen, dass die Formel für beschränkte und stetige Anfangsdaten f eine Lösung $u \in C^\infty(\mathbb{R}^n \times (0,\infty)) \cap C^0(\mathbb{R}^n \times [0,\infty))$ des Anfangswertproblems liefert.

Beispiel 11.5 Als zweites betrachten wir das Anfangswertproblem für die Wellengleichung:

$$\begin{aligned} \partial_t^2 u - \Delta u &= 0 \quad \text{auf } \mathbb{R}^n \times (0,\infty), \\ u(\cdot,0) &= f \quad \text{auf } \mathbb{R}^n, \\ \partial_t u(\cdot,0) &= g \quad \text{auf } \mathbb{R}^n. \end{aligned}$$

Wir führen den analogen Fourieransatz durch und erhalten das Anfangswertproblem

$$\begin{aligned} 0 = \widehat{\partial_t^2 u} - \widehat{\Delta u} &= \partial_t^2 \hat{u} + |p|^2 \hat{u} \quad \text{auf } \mathbb{R}^n \times (0,\infty), \\ \hat{f} &= \hat{u}(0,\cdot) \quad \text{auf } \mathbb{R}^n, \\ \hat{g} &= \partial_t \hat{u}(0,\cdot). \end{aligned}$$

Es folgt diesmal für die Fouriertransformierte

$$\hat{u}(p,t) = \hat{f}(p)\, \cos t|p| + \hat{g}(p)\, \frac{\sin t|p|}{|p|}.$$

Angenommen, es gibt eine Funktion $R : \mathbb{R}^n \times (0,\infty) \to \mathbb{R}$ mit $\hat{R} = (2\pi)^{-\frac{n}{2}}\, \frac{\sin(t|p|)}{|p|}$. Dann liefert formale Anwendung von Satz 11.6(2)

$$\begin{aligned} \widehat{R * g} &= (2\pi)^{\frac{n}{2}}\, \hat{R}\, \hat{g} = \hat{g}(p)\, \frac{\sin t|p|}{|p|}, \\ \partial_t \widehat{(R * f)} &= \partial_t \left(\widehat{R * f}\right) = \partial_t \left(\hat{f}(p)\, \frac{\sin(t|p|)}{|p|}\right) = \hat{f}(p) \cos(t|p|). \end{aligned}$$

Die Lösung des Anfangswertproblems wäre dann $u = \partial_t(R * f) + R * g$. Für $n = 1$ können wir $R = \frac{1}{2}\chi_{[-t,t]}$ wählen nach Beispiel 11.3 und erhalten

$$u(x,t) = \frac{1}{2}\left(f(x+t) + f(x-t) + \int_{x-t}^{x+t} g(y)\,dy\right).$$

Dies liefert eine gültige Lösung $u \in C^2(\mathbb{R} \times (0,\infty)) \cap C^1(\mathbb{R} \times [0,\infty))$, falls $f \in C^2(\mathbb{R})$ und $g \in C^1(\mathbb{R})$. Für $n = 2$ behaupten wir

$$R(x,t) = \frac{1}{2\pi}\frac{\chi_{B_t(0)}}{\sqrt{t^2 - |x|^2}}.$$

Denn mit Fubini folgt, zunächst für $p = (0, |p|) \neq 0$,

$$\begin{aligned}
2\pi\,\hat{R}(p,t) &= \frac{1}{2\pi}\int_{\{|x|<t\}} \frac{e^{-i\langle p,x\rangle}}{\sqrt{t^2-|x|^2}}\,dx \\
&= \frac{t}{2\pi}\int_{\{|y|<1\}} \frac{e^{-i\langle tp,y\rangle}}{\sqrt{1-|y|^2}}\,dy \\
&= \frac{t}{2\pi}\int_{-1}^{1} e^{-it|p|y_2}\int_{-\sqrt{1-y_2^2}}^{\sqrt{1-y_2^2}} \frac{1}{\sqrt{1-y_2^2-y_1^2}}\,dy_1\,dy_2 \\
&= \frac{t}{2\pi}\int_{-1}^{1} e^{-it|p|y_2}\int_{-1}^{1} \frac{1}{\sqrt{1-\eta^2}}\,d\eta\,dy_2 \\
&= \frac{t}{2}\int_{-1}^{1} e^{-it|p|y_2}\,dy_2 \\
&= \frac{\sin|p|t}{|p|}.
\end{aligned}$$

Dann gilt aber $\hat{R}(p,t) = \frac{1}{2\pi}\frac{\sin|p|t}{|p|}$ für alle $p \neq 0$, da beide Seiten invariant unter Drehungen sind. Wir haben also für $n = 2$ die Formel

$$u(t,x) = \partial_t\left(\frac{t}{2\pi}\int_{\{|y|<1\}} \frac{f(x-ty)}{\sqrt{1-|y|^2}}\,dy\right) + \frac{t}{2\pi}\int_{\{|y|<1\}} \frac{g(x-ty)}{\sqrt{1-|y|^2}}\,dy.$$

Damit die Formel eine Lösung $u \in C^2(\mathbb{R} \times (0,\infty)) \cap C^1(\mathbb{R} \times [0,\infty))$ definiert, müssen wir für $n = 2$ mehr voraussetzen, nämlich $g \in C^2(\mathbb{R}^n)$ und $f \in C^3(\mathbb{R}^n)$. Beachte auch, dass $R(\cdot,t)$ für $n = 2$ nicht mehr in $L^2(\mathbb{R}^n)$ liegt. Für $n \geq 3$ ist die Gleichung $\hat{R}(p,t) = (2\pi)^{-\frac{n}{2}}\frac{\sin|p|t}{|p|}$ gar nicht mehr durch eine Funktion lösbar; trotzdem kann mit dem Ansatz eine Lösung gefunden werden. Der zweite Schritt – die Rechtfertigung der Lösungsformel – ist für die Wellengleichung deutlich subtiler als für die Wärmeleitungsgleichung.

12 Anhang

Satz 12.1 (Heine-Borel) *Sei (X,d) metrischer Raum. Für $M \subset X$ sind äquivalent:*

(1) *M ist folgenkompakt: jede Folge $x_k \in M$ hat einen Häufungspunkt $x \in M$.*

(2) *M ist überdeckungskompakt: ist $M \subset \bigcup_{i \in I} U_i$ mit $U_i \subset X$ offen, so gibt es $I' \subset I$ endlich mit $M \subset \bigcup_{i \in I'} U_i$.*

Beweis: Sei M überdeckungskompakt. Wäre M nicht folgenkompakt, so gibt es eine Folge $x_k \in M$ mit folgender Eigenschaft: zu jedem $x \in M$ gibt es eine offene Umgebung U_x mit $x_k \notin U_x$ für hinreichend große k. Nach Voraussetzung wird M durch endlich viele Mengen $U_{x_1}, \ldots, U_{x_N}$ überdeckt, also folgt $x_k \notin M$ für k groß, ein Widerspruch.

Sei nun umgekehrt M folgenkompakt, und $\bigcup_{i \in I} U_i$ eine offene Überdeckung von M. Als erstes zeigen wir: es gibt ein $\varrho > 0$, so dass für alle $x \in M$ ein $i = i(x)$ existiert mit $B_\varrho(x) \subset U_i$. Andernfalls gibt es zu $k \in \mathbb{N}$ ein $x_k \in M$ mit $B_{1/k}(x_k) \backslash U_i \neq \emptyset$ für alle $i \in I$. Nach Übergang zu einer Teilfolge gilt $x_k \to x \in M$. Aber $x \in U_{i_0}$ für ein $i_0 \in I$, also $B_{1/k}(x_k) \subset U_{i_0}$ für k hinreichend groß, Widerspruch.

Gäbe es nun keine endliche Teilüberdeckung, so finden wir induktiv $x_1, x_2, \ldots$ in M mit $x_k \notin \bigcup_{i=1}^{k-1} B_\varrho(x_i)$. Es folgt $d(x_k, x_i) \geq \varrho > 0$ für $k > i$, das heißt die Folge x_k besitzt keine konvergente Teilfolge, im Widerspruch zur Voraussetzung. $\square$