MASS- UND INTEGRATIONSTHEORIE

Wolfgang König

Vorlesungsskript

Universität Leipzig

Wintersemester 2007/08

Inhaltsverzeichnis

Vorwort

Dies ist das Vorlesungsskript zu einer vierstündigen Vorlesung über Maß- und Integrationstheorie, gehalten an der Universität Leipzig im Wintersemester 2007/08. Der Stoff umfasst die mehrdimensionale Riemann- und Lebesgue-Integration bis zum Gauß'schen Integralsatz sowie eine umfassende Einführung in die allgemeine Maß- und Integrationstheorie. Die Vorlesung wurde im dritten Fachsemester gehalten, und eine zweisemestrige Analysisausbildung wurde vorausgesetzt, insbesondere mehrdimensionale Differenziation und das eindimensionale Riemann-Integral.

In Kapitel 1 folgen wir dem Beginn des Lehrbuches *Analysis 3* von Otto Forster [Fo3] und widmen uns dem mehrdimensionalen Riemann-Integral stetiger Funktionen. Insbesondere werden Approximationstechniken, der Transformationssatz, Differenzialoperatoren und Volumenberechnungen behandelt. Dann erfolgt ein Sprung in Kapitel 2 in die allgemeine Maßtheorie, deren Abstraktheit wir zunächst mit einer expliziten Präsentation des zweidimensionalen Lebesgue-Maßes abzumildern versuchen. Im weiteren Verlauf behandeln wir unter Anderem den Existenz- und Eindeutigkeitssatz von Carathéodory und Beweistechniken mit Hilfe von Dynkin-Systemen, Algebren und Inhalten. In Kapitel 3 führen wir messbare Funktionen und ihre Integration ein und stellen die fundamentalen Begriffe, Objekte und Techniken bereit wie z. B. die berühmten Konvergenzsätze, Bildmaße und Dichten sowie Produktmaße. Kapitel 4 ist dem Gauß'schen Integralsatz im $\mathbb{R}^d$ und Anwendungen gewidmet. Dazu untersuchen wir zunächst Untermannigfaltigkeiten und die Integration auf ihnen. Eine der Anwendungen des Gauß'schen Satzes betrifft auch die klassische Version des Stokes'schen Integralsatzes, dessen volle Wirksamkeit wir leider im Rahmen dieser Vorlesung nicht zeigen können.

Etliche Übungsaufgaben und Beispiele sind in den Text eingearbeitet. Die wichtigsten Beweise sind voll ausformuliert worden, aber Wiederholungen von Standardargumenten wurden weitgehend vermieden.

Leipzig, im Januar 2008

Kapitel 1

Das mehrdimensionale Riemann-Integral

In diesem Kapitel behandeln wir das Riemann-Integral für stetige Funktionen $\mathbb{R}^d \to \mathbb{R}$ mit kompaktem Träger und stellen etliche wichtige Regeln und Techniken bereit. Insbesondere klären wir in den Abschnitten 1.1–1.3 viele fundamentale Eigenschaften dieses Integralbegriffs wie Monotonie, Translationsinvarianz, Stetigkeit, Verhalten unter linearen und, allgemeiner, differenzierbaren Transformationen und vieles mehr. Ferner benutzen wir Teilungen der Eins, um gewisse Approximationen durchzuführen. Um in Abschnitt 1.6 auch die Volumina einiger kompakter Körper berechnen zu können, erweitern wir in Abschnitt 1.5 den Integralbegriff auf halbstetige Funktionen, die wir mit Hilfe monotoner Grenzwerte stetiger Funktionen mit kompaktem Träger einführen. In diesem Kapitel halten wir uns eng an [Fo3, Kap. 1–5].

1.1 Das Haar'sche Maß

In diesem Abschnitt betrachten wir die Abbildung, die stetige Funktionen mit kompaktem Träger auf ihr Riemann-Integral abbildet. Unser Hauptergebnis lautet, dass dieses Funktional bis auf konstante Vielfache durch die Eigenschaften der Linearität, Monotonie und Translationsinvarianz eindeutig festgelegt wird.

Wir erinnern an die Definition des Riemann-Integrals für Funktionen auf einem kompakten Intervall. Eine Funktion $f\colon [a,b] \to \mathbb{R}$ heißt *Riemann-integrierbar* oder kurz *integrierbar*, wenn für jedes $\varepsilon > 0$ zwei Treppenfunktionen, d. h. stückweise konstante Funktionen $\varphi, \psi\colon [a,b] \to \mathbb{R}$ mit $\varphi \le f \le \psi$ existieren, so dass

$$\int_a^b (\psi(x) - \varphi(x))\,\mathrm{d}x < \varepsilon.$$

Hierbei ist $\int_a^b \varphi(x)\,\mathrm{d}x = \sum_{i=1}^n \varphi(t_i)(t_i - t_{i-1})$, wenn $a = t_0 < t_1 < t_2 < \cdots < t_n = b$ eine Unterteilung ist, so dass φ auf $(t_{i-1}, t_i]$ konstant ist, analog für ψ und $\psi - \varphi$. Eine äquivalente Formulierung ist, dass f genau dann Riemann-integrierbar ist, falls eine Zahl $I \in \mathbb{R}$ existiert, so dass für jedes $\varepsilon > 0$ ein $\delta > 0$ existiert, so dass für jede Unterteilung $a = t_0 < t_1 < t_2 < \cdots <$

$t_n = b$ mit Feinheit $\max_{i=1}^n (t_i - t_{i-1}) < \delta$ und für jede Wahl von Stützstellen $\xi_i \in (t_{i-1}, t_i]$ gilt:

$$\Big| I - \sum_{i=1}^n f(\xi_i)(t_i - t_{i-1}) \Big| < \varepsilon.$$

Man schreibt dann

$$I = \int_a^b f(x)\,\mathrm{d}x$$

und nennt I das *Riemann-Integral* von f. Die Summe $\sum_{i=1}^n f(\xi_i)(t_i - t_{i-1})$ heißt eine *Riemann-Summe* zu dieser Unterteilung und diesen Stützstellen. In diesem Kapitel lassen wir oft den Zusatz 'Riemann' weg.

In der zweisemestrigen Vorlesung *Analysis A* zeigt man, dass integrierbare Funktionen immer beschränkt sind. Ferner sieht man, dass stetige Funktionen integrierbar sind. In diesem Abschnitt wollen wir den Riemann-Integralbegriff auf Funktionen $\mathbb{R}^d \to \mathbb{R}$ ausdehnen, wobei wir uns allerdings zunächst auf stetige Funktionen beschränken wollen. Einen Schritt in diese Richtung machten wir schon am Ende der Vorlesung *Analysis A*, wo wir Funktionen auf kompakten Quadern betrachteten. Dabei beweisen wir unter Anderem Folgendes.

Satz 1.1.1 (Satz von Fubini für Riemann-Integrale stetiger Funktionen). *Es sei $Q = \prod_{i=1}^d [a_i, b_i]$ ein kompakter Quader im $\mathbb{R}^d$. Ferner sei $f\colon Q \to \mathbb{R}$ stetig. Dann gilt für jede Permutation $i_1, \dots, i_d$ der Zahlen $1, \dots, d$, dass die folgend auftretenden Integrale existieren mit*

$$\int_{a_d}^{b_d} \Big(\cdots \Big(\int_{a_1}^{b_1} f(x_1, \dots, x_d)\,\mathrm{d}x_1 \Big) \cdots \Big) \mathrm{d}x_d = \int_{a_{i_d}}^{b_{i_d}} \Big(\cdots \Big(\int_{a_{i_1}}^{b_{i_1}} f(x_1, \dots, x_d)\,\mathrm{d}x_{i_1} \Big) \cdots \Big) \mathrm{d}x_{i_d}. \tag{1.1.1}$$

Dies folgt etwa aus [Fo2, Satz 9.3]. Eine wichtige Vorbereitung des Beweises ist [Fo2, Satz 9.1], der besagt, dass die Abbildung $x \mapsto \int_c^d f(x, y)\,\mathrm{d}y$ auf dem Intervall $[a, b]$ stetig ist, wenn f auf dem Quader $[a, b] \times [c, d]$ stetig ist.

Also können wir leicht das Riemann-Integral erklären für stetige Funktionen, die außerhalb eines kompakten Quaders verschwinden. Wir brauchen die folgenden fundamentalen Begriffe.

Definition 1.1.2 (Träger einer Funktion, $\mathcal{C}_\mathrm{c}(\mathbb{R}^d)$). *(i) Der* Träger *(engl.* support*) einer Funktion $f\colon \mathbb{R}^d \to \mathbb{R}$ ist definiert als die Menge*

$$\mathrm{supp}\,(f) = \overline{\{x \in \mathbb{R}^d \colon f(x) \neq 0\}},$$

also der Abschluss der Menge, auf der f nicht verschwindet.

(ii) Mit $\mathcal{C}_\mathrm{c}(\mathbb{R}^d)$ bezeichnen wir die Menge aller stetigen Funktionen $\mathbb{R}^d \to \mathbb{R}$ mit kompaktem Träger.

Der Träger einer Funktion ist also das Komplement der größten offenen Menge, auf der die Funktion verschwindet. Natürlich ist der Träger genau dann kompakt, wenn er beschränkt ist.

Nun können wir das Integral jeder Funktion in $\mathcal{C}_\mathrm{c}(\mathbb{R}^d)$ leicht definieren, indem wir setzen:

$$\int_{\mathbb{R}^d} f(x)\,\mathrm{d}x = \int_Q f(x)\,\mathrm{d}x, \qquad f \in \mathcal{C}_\mathrm{c}(\mathbb{R}^d), \tag{1.1.2}$$

falls $\operatorname{supp}(f)$ in dem Quader Q enthalten ist. Das Integral $\int_Q f(x)\,\mathrm{d}x$ ist auf natürliche Weise erklärt als die linke Seite von (1.1.1), falls $Q = \prod_{i=1}^d [a_i, b_i]$. Man kann sich leicht davon überzeugen, dass die Definition in (1.1.2) nicht von der Wahl des Quaders Q abhängt, solange $\operatorname{supp}(f) \subset Q$ gilt. Statt $\int_{\mathbb{R}^d} f(x)\,\mathrm{d}x$ schreibt man auch $\int_{\mathbb{R}^d} f(x)\,\mathrm{d}^d x$ oder $\int f(x)\,\mathrm{d}x$ oder $\int_{\mathbb{R}^d} f(x_1,\dots,x_d)\,\mathrm{d}x_1 \dots \mathrm{d}x_d$ oder $\int_{\mathbb{R}^d} f(y)\,\mathrm{d}y$ und so weiter.

Bemerkung 1.1.3 (Produktfunktionen). Falls f von der Form $f = \bigotimes_{i=1}^d \varphi_i$ ist, also $f(x_1,\dots,x_d) = \prod_{i=1}^d \varphi_i(x_i)$, für stetige Funktionen $\varphi_1,\dots,\varphi_d\colon \mathbb{R} \to \mathbb{R}$ mit kompaktem Träger, so ist $f \in \mathcal{C}_\mathrm{c}(\mathbb{R}^d)$ und es gilt

$$\int_{\mathbb{R}^d} f(x)\,\mathrm{d}x = \prod_{i=1}^d \int_{\mathbb{R}} \varphi_i(x_i)\,\mathrm{d}x_i.$$

(Übungsaufgabe). Man nennt f auch das *Tensorprodukt* der Funktionen $\varphi_1,\dots,\varphi_d$. $\diamond$

Wir können das Integral als ein Funktional auf $\mathcal{C}_\mathrm{c}(\mathbb{R}^d)$ auffassen. Wir sammeln seine wichtigsten Eigenschaften:

Satz 1.1.4 (Eigenschaften des Riemann-Integrals). *Das durch* (1.1.2) *definierte Funktional* $I\colon \mathcal{C}_\mathrm{c}(\mathbb{R}^d) \to \mathbb{R}$, $I(f) = \int_{\mathbb{R}^d} f(x)\,\mathrm{d}x$, *ist linear, monoton und translationsinvariant.*

Dabei heißt ein Funktional $J\colon \mathcal{C}_\mathrm{c}(\mathbb{R}^d) \to \mathbb{R}$

linear, falls $J(\alpha f + \beta g) = \alpha J(f) + \beta J(g)$ für alle $\alpha, \beta \in \mathbb{R}, f, g \in \mathcal{C}_\mathrm{c}(\mathbb{R}^d)$,
monoton oder *isoton*, falls für $f, g \in \mathcal{C}_\mathrm{c}(\mathbb{R}^d)$ gilt: $f \le g \Longrightarrow J(f) \le J(g)$,
translationsinvariant, falls $J(\theta_y f) = J(f)$ für jedes $f \in \mathcal{C}_\mathrm{c}(\mathbb{R}^d), y \in \mathbb{R}^d$.

Dabei ist der *Verschiebungsoperator* $\theta_y\colon \mathcal{C}_\mathrm{c}(\mathbb{R}^d) \to \mathcal{C}_\mathrm{c}(\mathbb{R}^d)$ definiert durch $\theta_y f(x) = f(x-y)$. Der Beweis von Satz 1.1.4 ist eine leichte Übungsaufgabe: Man führt die Linearität und Monotonie auf die bekannten Eigenschaften des eindimensionalen Riemann-Integrals auf Kompakta zurück und benutzt die Substitutionsregel für den Nachweis der Translationsinvarianz.

Wir werden als Nächstes zeigen, dass I das *einzige* lineare, monotone und translationsinvariante Funktional $\mathcal{C}_\mathrm{c}(\mathbb{R}^d) \to \mathbb{R}$ ist, zumindest bis auf konstante Vielfache. Dazu benötigen wir ein paar Vorbereitungen, zunächst eine gewisse Stetigkeitseigenschaft. Wir erinnern daran, dass gleichmäßige Konvergenz von Funktionen in $\mathcal{C}_\mathrm{c}(\mathbb{R}^d)$ die Konvergenz bezüglich der Supremumsnorm $\|f\|_\infty = \sup_{x\in\mathbb{R}^d} |f(x)|$ ist: f_k konvergiert genau dann gleichmäßig gegen f, wenn $\lim_{k\to\infty} \|f_k - f\|_\infty = 0$.

Lemma 1.1.5. *Es sei* $J\colon \mathcal{C}_\mathrm{c}(\mathbb{R}^d) \to \mathbb{R}$ *ein lineares und monotones Funktional. Es sei* $(f_k)_{k\in\mathbb{N}}$ *eine Folge in* $\mathcal{C}_\mathrm{c}(\mathbb{R}^d)$*, deren Träger alle in einem gewissen kompakten Quader* $Q \subset \mathbb{R}^d$ *enthalten sind. Die Folge der* f_k *konvergiere gleichmäßig gegen ein* $f \in \mathcal{C}_\mathrm{c}(\mathbb{R}^d)$*. Dann gilt* $\lim_{k\to\infty} J(f_k) = J(f)$*.*

Beweis. Wir betrachten eine stetige Funktion $\Phi\colon \mathbb{R}^d \to [0,1]$ mit kompaktem Träger und $\Phi|_Q = 1$. (Eine solche Funktion erhält man z. B. als $\Phi(x_1,\dots,x_d) = \prod_{i=1}^d \varphi_i(x_i)$, wobei $\varphi_i\colon \mathbb{R} \to [0,1]$

stetig mit kompaktem Träger ist und $\varphi_i|_{[a_i,b_i]} = 1$, wenn $Q = \prod_{i=1}^d [a_i, b_i]$.) Da der Träger von $f_k - f$ in Q enthalten ist, haben wir daher

$$-\|f_k - f\|_\infty \Phi \leq f_k - f \leq \|f_k - f\|_\infty \Phi \qquad \text{in } \mathbb{R}^d,$$

also wegen der Linearität und Monotonie von J auch

$$-\|f_k - f\|_\infty J(\Phi) \leq J(f_k) - J(f) \leq \|f_k - f\|_\infty J(\Phi),$$

und dies heißt, dass $|J(f_k) - J(f)| \leq \|f_k - f\|_\infty J(\Phi)$. Daraus folgt sofort die Aussage. □

Im Folgenden wollen wir beliebige Funktionen in $\mathcal{C}_\mathrm{c}(\mathbb{R}^d)$ gleichmäßig approximieren mit Hilfe spezieller Funktionen besonders einfacher Gestalt. Dazu betrachten wir zunächst die reskalierte Version einer gewissen Zackenfunktion.

Bemerkung 1.1.6 (Teilung der Eins). Betrachte $\psi \in \mathcal{C}_\mathrm{c}(\mathbb{R})$, definiert durch $\psi(t) = (1 - |t|)_+$, dann ist der Graph von ψ ein Dreieck mit Grundfläche $[-1, 1]$ und Höhe Eins. Die d-dimensionale Variante definieren wir durch $\Psi = \psi^{\otimes d} \in \mathcal{C}_\mathrm{c}(\mathbb{R})$ als das d-fache Tensorprodukt, also $\Psi(x_1, \dots, x_d) = \prod_{i=1}^d \psi(x_i)$ (siehe Bemerkung 1.1.3). Dann ist der Graph von Ψ eine Pyramide mit Grundfläche $[-1, 1]^d$ und Höhe Eins. Für $\varepsilon > 0$ sei $\Psi_\varepsilon(x) = \Psi(x/\varepsilon)$, also ist insbesondere $\Psi_1 = \Psi$. Der Träger $\operatorname{supp}(\Psi_\varepsilon) = [-\varepsilon, \varepsilon]^d$ ist der Würfel mit Mittelpunkt im Ursprung und Seitenlänge 2ε.[1] Mit Hilfe von Bemerkung 1.1.3 und einer elementaren Substitution sieht man, dass $\int_{\mathbb{R}^d} \Psi_\varepsilon(x)\,\mathrm{d}x = (\int_\mathbb{R} \psi(x/\varepsilon)\,\mathrm{d}x)^d = \varepsilon^d$.

Wir erinnern an den Verschiebeoperator $\theta_y f(x) = f(x-y)$. Wenn man ψ und $\theta_1\psi$ addiert, so erhält man eine Funktion, die im ganzen Intervall $[0, 1]$ konstant gleich Eins ist. Addiert man $\theta_2\psi$ hinzu, ist die resultierende Funktion konstant gleich Eins im Intervall $[1, 2]$, und im Intervall $[0, 1]$ hat sich nichts geändert, da der Träger von $\theta_2\psi$ ja gleich $[1, 3]$ ist. Diese Überlegung kann man iterieren und erhält, dass $\sum_{k\in\mathbb{Z}} \theta_k\psi$ die konstante Einsfunktion in $\mathbb{R}$ ist. Durch Multiplikation erhält man die entsprechende Aussage für die d-dimensionale Version:

$$\sum_{k\in\mathbb{Z}^d} \theta_k\Psi(x_1, \dots, x_d) = \prod_{i=1}^d \sum_{k_i\in\mathbb{Z}} \theta_{k_i}\psi(x_i) = \prod_{i=1}^d 1 = 1, \qquad x_1, \dots, x_d \in \mathbb{R},$$

d. h. die Summe der Verschiebungen $\theta_k\Psi$ mit $k \in \mathbb{Z}^d$ ist konstant gleich Eins. Man sagt, dass die Familie dieser Funktionen eine *Teilung der Eins* darstellt. Insbesondere haben wir auch $1 = \sum_{k\in\mathbb{Z}^d} \theta_{k\varepsilon}\Psi_\varepsilon$ für jedes $\varepsilon > 0$. Die Graphen der Funktionen $\theta_{k\varepsilon}\Psi_\varepsilon$ mit $k \in \mathbb{Z}^d$ bilden spitze Pyramiden auf Würfeln der Seitenlänge 2ε um $k\varepsilon$. Überlappungen der Träger zweier Funktionen $\theta_{k\varepsilon}\Psi_\varepsilon$ und $\theta_{k'\varepsilon}\Psi_\varepsilon$ sind nur nichttrivial, wenn k und k' Nachbarn sind in dem Sinne, dass $\max_{i=1}^d |k_i - k_i'| \leq 1$. ◇

Diese Teilung der Eins wird das Gerüst unserer Approximation einer beliebigen Funktion in $\mathcal{C}_\mathrm{c}(\mathbb{R}^d)$ bilden, also betrachten wir erst einmal, wie ein beliebiges lineares translationsinvariantes Funktional auf ihre Skalierung wirkt:

Lemma 1.1.7. *Seien ψ, Ψ und Ψ_ε definiert wie in Bemerkung 1.1.6. Dann gilt für jedes lineare translationsinvariante Funktional $J\colon \mathcal{C}_\mathrm{c}(\mathbb{R}) \to \mathbb{R}$ und jedes $\varepsilon > 0$:*

$$J(\Psi_{\varepsilon/2}) = 2^{-d} J(\Psi_\varepsilon), \qquad \varepsilon > 0.$$

[1] Man kann diesen Würfel auch ansehen als den 'Ball' um Null mit 'Radius' ε bezüglich der Maximumsnorm $\|x\|_\infty = \max_{i=1}^d |x_i|$, aber diese Sprechweise könnte verwirren.

Beweis. Wir behandeln zunächst den Fall $d = 1$. Nach scharfem Hinsehen oder mit Hilfe einer geeigneten Skizze sieht man, dass gilt:

$$\psi_\varepsilon = \frac{1}{2}\theta_{-\varepsilon/2}\psi_{\varepsilon/2} + \psi_{\varepsilon/2} + \frac{1}{2}\theta_{\varepsilon/2}\psi_{\varepsilon/2}, \qquad \varepsilon > 0,$$

wobei wir an den Verschiebungsoperator $\theta_y f(x) = f(x - y)$ erinnern. Daraus folgt mit der Linearität und Verschiebungsinvarianz, dass gilt:

$$J(\psi_\varepsilon) = \frac{1}{2}J(\theta_{-\varepsilon/2}\psi_{\varepsilon/2}) + J(\psi_{\varepsilon/2}) + \frac{1}{2}J(\theta_{\varepsilon/2}\psi_{\varepsilon/2}) = 2J(\psi_{\varepsilon/2}),$$

also ist die Behauptung im Fall $d = 1$ nun schon bewiesen. Für allgemeines d benutzen wir die obige Darstellung in jeder einzelnen Dimension und multiplizieren aus:

$$\begin{aligned}
\Psi_\varepsilon(x_1, \dots, x_d) = \prod_{i=1}^d \psi_\varepsilon(x_i) &= \prod_{i=1}^d \Big(\frac{1}{2}\theta_{-\varepsilon/2}\psi_{\varepsilon/2}(x_i) + \psi_{\varepsilon/2}(x_i) + \frac{1}{2}\theta_{\varepsilon/2}\psi_{\varepsilon/2}(x_i)\Big)\\
&= \prod_{i=1}^d \sum_{\alpha_i \in \{-1,0,1\}} \Big(\frac{1}{2}\Big)^{|\alpha_i|} \theta_{\alpha_i\varepsilon/2}\psi_{\varepsilon/2}(x_i)\\
&= \sum_{\alpha_1,\dots,\alpha_d \in \{-1,0,1\}} \Big(\prod_{i=1}^d \Big(\frac{1}{2}\Big)^{|\alpha_i|}\Big) \prod_{i=1}^d \theta_{\alpha_i\varepsilon/2}\psi_{\varepsilon/2}(x_i)\\
&= \sum_{\alpha_1,\dots,\alpha_d \in \{-1,0,1\}} \Big(\prod_{i=1}^d \Big(\frac{1}{2}\Big)^{|\alpha_i|}\Big) \theta_{\alpha\varepsilon/2}\Psi_{\varepsilon/2}(x),
\end{aligned}$$

wobei wir $\alpha = (\alpha_1, \dots, \alpha_d) \in \{-1, 0, 1\}^d$ und $x = (x_1, \dots, x_d)$ schrieben. Nun ergibt die Linearität und Verschiebungsinvarianz:

$$J(\Psi_\varepsilon) = J(\Psi_{\varepsilon/2}) \sum_{\alpha_1,\dots,\alpha_d \in \{-1,0,1\}} \Big(\prod_{i=1}^d \Big(\frac{1}{2}\Big)^{|\alpha_i|}\Big) = J(\Psi_{\varepsilon/2}) \prod_{i=1}^d \sum_{\alpha \in \{-1,0,1\}} \Big(\frac{1}{2}\Big)^{|\alpha|} = J(\Psi_{\varepsilon/2})2^d,$$

und das beendet den Beweis. □

Mehrmalige Anwendung von Lemma 1.1.7 ergibt natürlich, dass $J(\Psi_{2^{-n}}) = 2^{-nd}J(\Psi)$ für jedes $n \in \mathbb{N}$.

Mit Hilfe der reskalierten Teilung der Eins werden wir nun beliebige Funktionen in $\mathcal{C}_c(\mathbb{R}^d)$ gleichmäßig approximieren.

Lemma 1.1.8. *Seien ψ, Ψ und Ψ_ε definiert wie in Bemerkung 1.1.6. Dann gilt für jedes $f \in \mathcal{C}_c(\mathbb{R}^d)$:*

$$\lim_{\varepsilon \downarrow 0} \Big\| f - \sum_{k \in \mathbb{Z}^d} f(k\varepsilon)\theta_{k\varepsilon}\Psi_\varepsilon \Big\|_\infty = 0.$$

(Da f kompakten Träger hat, ist die obige Summe für jedes $\varepsilon > 0$ endlich.)

Beweis. Sei ein kleines $\eta > 0$ vorgegeben. Da f kompakten Träger hat, ist f gleichmäßig stetig, also gibt es ein $\varepsilon > 0$ mit $|f(x) - f(y)| < \eta$ für alle $x, y \in \mathbb{R}^d$ mit $\|x - y\|_\infty < \varepsilon$.[2] Wir zeigen

[2]Hier benutzen wir, dass $f\colon \mathbb{R}^d \to \mathbb{R}$ stetig ist, wenn man den $\mathbb{R}^d$ mit der Maximumsnorm $\|x\|_\infty = \max_{i=1}^d |x_i|$ ausstattet. Da alle Normen auf dem $\mathbb{R}^d$ äquivalent sind (explizit für diese Norm und die Euklid'sche Norm $\|\cdot\|$ gilt: $\|\cdot\|_\infty \leq \|\cdot\| \leq \sqrt{d}\|\cdot\|_\infty$), durften wir uns eine 'angenehme' Norm aussuchen.

nun, dass gilt:

$$\|G_\varepsilon\|_\infty < \eta, \qquad \text{wobei } G_\varepsilon = f - \sum_{k\in\mathbb{Z}^d} f(k\varepsilon)\theta_{k\varepsilon}\Psi_\varepsilon,$$

und dies beendet den Beweis.

Für jedes $x \in \mathbb{R}^d$ sei $A(x) = \{k \in \mathbb{Z}^d : x \in \operatorname{supp}(\theta_{k\varepsilon}\Psi_\varepsilon)\}$ (dies ist eine endliche Menge), dann ist

$$G_\varepsilon(x) = \sum_{k\in\mathbb{Z}^d} (f(x) - f(k\varepsilon))\theta_{k\varepsilon}\Psi_\varepsilon(x) = \sum_{k\in A(x)} (f(x) - f(k\varepsilon))\theta_{k\varepsilon}\Psi_\varepsilon(x).$$

Für $k \in A(x)$ gilt $\|x - k\varepsilon\|_\infty \le \varepsilon$, da der Träger von Ψ_ε gleich dem Würfel mit Seitenlänge 2ε ist. Also können wir für $k \in A(x)$ abschätzen: $|f(x) - f(k\varepsilon)| \le \eta$. Wenn wir dies oben einsetzen, erhalten wir

$$|G_\varepsilon(x)| \le \sum_{k\in A(x)} |f(x) - f(k\varepsilon)|\theta_{k\varepsilon}\Psi_\varepsilon(x) \le \eta \sum_{k\in A(x)} \theta_{k\varepsilon}\Psi_\varepsilon(x) \le \eta,$$

und der Beweis ist beendet. □

Nun können wir endlich das angekündigte Resultat über die Eindeutigkeit linearer monotoner translationsinvarianter Funktionale beweisen:

Satz 1.1.9. *Es sei $J : \mathcal{C}_c(\mathbb{R}^d) \to \mathbb{R}$ ein lineares monotones translationsinvariantes Funktional. Dann gibt es ein $c \in \mathbb{R}$ mit $J(f) = c\int_{\mathbb{R}^d} f(x)\,\mathrm{d}x$ für jedes $f \in \mathcal{C}_c(\mathbb{R}^d)$.*

Beweis. Wir zeigen die Aussage mit $c = J(\Psi)$, wobei Ψ in Bemerkung 1.1.6 definiert wurde. Wir setzen $I(f) = \int_{\mathbb{R}^d} f(x)\,\mathrm{d}x$ und müssen zeigen, dass $J = cI$. Aus der Bemerkung nach Lemma 1.1.7 folgt, dass $J(\Psi_{2^{-n}}) = cI(\Psi_{2^{-n}})$ für jedes $n \in \mathbb{N}$. Für beliebiges $f \in \mathcal{C}_c(\mathbb{R}^d)$ setzen wir

$$f_n = \sum_{k\in\mathbb{Z}^d} f(k2^{-n})\theta_{k2^{-n}}\Psi_{2^{-n}}.$$

Aus der Linearität und Translationsinvarianz von J und I folgt $J(f_n) = cI(f_n)$ für jedes $n \in \mathbb{N}$. Die Träger der Funktionen f_n sind alle in einem gemeinsamen Kompaktum enthalten, das nur von f abhängt. Nach Lemma 1.1.8 konvergieren die f_n gleichmäßig auf $\mathbb{R}^d$ gegen f. Aus Lemma 1.1.5 folgt dann, dass $J(f) = cI(f)$. □

Definition 1.1.10 (Haar'sches Maß). *Jedes lineare monotone und translationsinvariante Funktional $\mathcal{C}_c(\mathbb{R}^d) \to \mathbb{R}$ heißt ein* Haar'sches Maß.

Wir haben also die Eindeutigkeit des Haar'schen Maßes bis auf Vielfache gezeigt.

Der Begriff 'Maß' ist an dieser Stelle für uns nur ein Wort, da wir die allgemeine Theorie der Maße und der Integrale bezüglich ihrer erst noch kennenlernen werden. Hinter dieser Begriffsbildung steht die Idee, dass hinter jedem linearen monotonen und translationsinvarianten Funktional ein Maß steht, so dass dieses Funktional gleich dem Integral bezüglich diesen Maßes ist. In diesem Fall wird es sich als das Lebesgue-Maß herausstellen.

1.2 Die Transformationsformel

In diesem Abschnitt formulieren und beweisen wir die Transformationsformel für d-dimensionale Riemann-Integrale stetiger Funktionen mit kompaktem Träger, eine Verallgemeinerung der bekannten Substitutionsregel für eindimensionale Riemann-Integrale. Für lineare Koordinationswechsel kann die Transformationsformel leicht aus der axiomatischen Charakterisierung des Integrals (siehe Lemma 1.1.4) abgeleitet werden, für differenzierbare Wechsel werden wir geeignet lokal approximieren.

Mit $\mathrm{GL}(d,\mathbb{R})$ bezeichnen wir die Menge aller reellen invertierbaren $d \times d$-Matrizen. Wir machen keinen Unterschied zwischen einer reellen $(d \times d)$-Matrix $A \in \mathbb{R}^{d\times d}$ und der von ihr induzierten linearen bijektiven Abbildung $A\colon \mathbb{R}^d \to \mathbb{R}^d$. Für $f \in \mathcal{C}_\mathrm{c}(\mathbb{R}^d)$ liegt $f \circ A$ auch wieder in $\mathcal{C}_\mathrm{c}(\mathbb{R}^d)$.

Lemma 1.2.1. *Für jedes lineare monotone und translationsinvariante Funktional $J\colon \mathcal{C}_\mathrm{c}(\mathbb{R}^d) \to \mathbb{R}$ und jedes $A \in \mathrm{GL}(d,\mathbb{R})$ ist auch die Abbildung $f \mapsto J(f \circ A)$ ein lineares monotones und translationsinvariantes Funktional.*

Beweis. Die Linearität und Monotonie sind sehr leicht zu zeigen. Für jedes $f \in \mathcal{C}_\mathrm{c}(\mathbb{R}^d)$ und $x, y \in \mathbb{R}^d$ ist

$$\begin{aligned}((\theta_y f) \circ A)(x) &= (\theta_y f)(Ax) = f(Ax - y) = f\big(A(x - A^{-1}y)\big) = (f \circ A)(x - A^{-1}y)\\ &= \theta_{A^{-1}y}(f \circ A)(x),\end{aligned}$$

also gilt $(\theta_y f) \circ A = \theta_{A^{-1}y}(f \circ A)$. Daraus folgt

$$J\big((\theta_y f) \circ A\big) = J\big(\theta_{A^{-1}y}(f \circ A)\big) = J(f \circ A),$$

und dies zeigt die Translationsinvarianz. □

Damit wollen wir nun zeigen, dass das Riemann-Integral unter orthogonalen Koordinatentransformationen invariant bleibt. Mit $O(d)$ bezeichnen wir die Menge der orthogonalen $d \times d$-Matrizen, d. h. die Menge der Matrizen $A \in \mathbb{R}^{d\times d}$, die $AA^\mathrm{T} = E = A^\mathrm{T}A$ erfüllen bzw. $\|Ax\| = \|x\|$ für jedes $x \in \mathbb{R}^d$. (Hier ist E die $(d \times d)$-Einheitsmatrix und $\|\cdot\|$ die Euklid'sche Norm.)

Satz 1.2.2 (Bewegungsinvarianz des Riemann-Integrals). *Für jedes $A \in O(d)$ und jedes $f \in \mathcal{C}_\mathrm{c}(\mathbb{R}^d)$ gilt*

$$\int_{\mathbb{R}^d} f(Ax)\,\mathrm{d}x = \int_{\mathbb{R}^d} f(x)\,\mathrm{d}x.$$

Beweis. Die beiden Seiten der behaupteten Gleichung definieren nach Lemma 1.1.4 bzw. 1.2.1 jeweils ein lineares monotones und translationsinvariantes Funktional auf $\mathcal{C}_\mathrm{c}(\mathbb{R}^d)$. Also gibt es nach Satz 1.1.9 ein $c \in \mathbb{R}$ mit $\int_{\mathbb{R}^d} f(Ax)\,\mathrm{d}x = c\int_{\mathbb{R}^d} f(x)\,\mathrm{d}x$ für alle $f \in \mathcal{C}_\mathrm{c}(\mathbb{R}^d)$. Um den Wert von c zu bestimmen, werten wir beide Seiten in der Funktion $f_0 \in \mathcal{C}_\mathrm{c}(\mathbb{R}^d)$ aus, die gegeben ist durch $f_0(x) = (1 - \|x\|)_+$. Da A orthogonal ist, gilt $f_0(Ax) = f_0(x)$ für jedes $x \in \mathbb{R}^d$, also auch $\int_{\mathbb{R}^d} f_0(Ax)\,\mathrm{d}x = \int_{\mathbb{R}^d} f_0(x)\,\mathrm{d}x \neq 0$. Also ist $c = 1$. □

Bemerkung 1.2.3. Nun können wir einen neuen Beweis des Satzes von Fubini (Satz 1.1.1) führen: Wenn $i_1, \dots, i_d$ eine Permutation von $1, \dots, d$ ist, dann wird die Abbildung $(x_1, \dots, x_d) \mapsto (x_{i_1}, \dots, x_{i_d})$ durch eine orthogonale Matrix A beschrieben, also folgt (1.1.1) aus Satz 1.2.2. $\diamond$

Nun wollen wir $\int_{\mathbb{R}^d} f(Ax)\,\mathrm{d}x$ für eine beliebige Matrix $A \in \mathrm{GL}(d, \mathbb{R})$ identifizieren. Zunächst sieht man im Spezialfall einer Diagonalmatrix $A = \mathrm{Diag}(\alpha_1, \dots, \alpha_d)$ mit Diagonalelementen $\alpha_1, \dots, \alpha_d > 0$ sofort durch d-malige Anwendung der Substitutionsregel für eindimensionale Riemann-Integrale, dass gilt:

$$\int_{\mathbb{R}^d} f(\alpha_1 x_1, \dots, \alpha_d x_d)\,\mathrm{d}x = \frac{1}{\alpha_1 \dots \alpha_d} \int_{\mathbb{R}^d} f(x)\,\mathrm{d}x. \tag{1.2.1}$$

Um den Fall einer beliebigen Matrix $A \in \mathrm{GL}(d, \mathbb{R})$ auf diesen simplen Fall zurück zu führen, benötigen wir das Folgende. Mit $\mathrm{Diag}(\alpha_1, \dots, \alpha_d)$ bezeichnen wir die Diagonalmatrix mit den Zahlen $\alpha_1, \dots, \alpha_d \in \mathbb{R}$ auf der Diagonalen.

Lemma 1.2.4. *Für jede Matrix $A \in \mathrm{GL}(d, \mathbb{R})$ gibt es orthogonale Matrizen S_1 und S_2 und $\alpha_1, \dots, \alpha_d > 0$, so dass $A = S_1 \mathrm{Diag}(\alpha_1, \dots, \alpha_d) S_2$.*

Beweis. Die Matrix $A^{\mathrm{T}} A$ ist positiv definit, lässt sich also durch eine orthogonale Transformation auf Diagonalgestalt mit positiven Diagonaleinträgen $\alpha_1^2, \dots, \alpha_d^2 > 0$ bringen, d. h. es gibt eine orthogonale Matrix S mit $S^{\mathrm{T}} A^{\mathrm{T}} A S = \mathrm{Diag}(\alpha_1^2, \dots, \alpha_d^2)$. Zur Abkürzung setzen wir $D = \mathrm{Diag}(\alpha_1, \dots, \alpha_d)$ (also $D^2 = \mathrm{Diag}(\alpha_1^2, \dots, \alpha_d^2)$) und gehen davon aus, dass $\alpha_1, \dots, \alpha_d > 0$. Die Matrix $S_1 = ASD^{-1}$ ist orthogonal, denn es gilt

$$S_1^{\mathrm{T}} S_1 = (D^{-1} S^{\mathrm{T}} A^{\mathrm{T}})(ASD^{-1}) = D^{-1}(S^{\mathrm{T}} A^{\mathrm{T}} AS) D^{-1} = D^{-1} \mathrm{Diag}(\alpha_1^2, \dots, \alpha_d^2) D^{-1} = E.$$

Die Matrix $S_2 = S^{-1}$ ist ebenfalls orthogonal, und wir haben $S_1 D S_2 = ASD^{-1} D S^{-1} = A$. $\square$

Damit können wir nun die Transformationsformel für bijektive lineare Koordinationswechsel beweisen:

Satz 1.2.5. *Für jede Matrix $A \in \mathrm{GL}(d, \mathbb{R})$ und jede Funktion $f \in \mathcal{C}_{\mathrm{c}}(\mathbb{R}^d)$ gilt*

$$\int_{\mathbb{R}^d} f(Ax) |\det(A)|\,\mathrm{d}x = \int_{\mathbb{R}^d} f(y)\,\mathrm{d}y.$$

Symbolisch kann man schreiben: $\mathrm{d}(Ax) = |\det(A)|\,\mathrm{d}x$.

Beweis. Wie in Lemma 1.2.4 schreiben wir $A = S_1 D S_2$ mit orthogonalen Matrizen S_1 und S_2 und $D = \mathrm{Diag}(\alpha_1, \dots, \alpha_d)$ mit $\alpha_1, \dots, \alpha_d > 0$. Wegen $|\det(S_1)| = |\det(S_2)| = 1$ und $\det(D) > 0$ ist dann $|\det(A)| = \det(D)$. Setzen wir zur Abkürzung $I(f) = \int_{\mathbb{R}^d} f(x)\,\mathrm{d}x$, dann kann man Satz 1.2.2 formulieren als $I(f) = I(f \circ S_1)$, und (1.2.1) schreibt sich als $\det(D) I(f \circ D) = I(f)$, jeweils für jedes $f \in \mathcal{C}_{\mathrm{c}}(\mathbb{R}^d)$. Setzen wir dies zusammen, ergibt sich für jedes $f \in \mathcal{C}_{\mathrm{c}}(\mathbb{R}^d)$:

$$I(f) = I(f \circ S_1) = \det(D) I(f \circ S_1 \circ D) = \det(D) I(f \circ S_1 \circ D \circ S_2) = |\det(A)| I(f \circ A),$$

und das beendet den Beweis. $\square$

Man mag sich wundern über die Betragstriche um die Determinante von A, die ja in der Theorie der Riemann-Integrale über kompakte Intervalle nicht auftauchen. Der Grund ist, dass in unserer aktuellen Schreibweise $\int_{\mathbb{R}} f(x)\,\mathrm{d}x$ eine Regel wie $\int_a^b f(-x)\,\mathrm{d}x = -\int_{-a}^{-b} f(y)\,\mathrm{d}y$ nicht ausgedrückt werden kann, denn wir integrieren immer 'von links nach rechts'. Statt $-\int_{-a}^{-b} f(y)\,\mathrm{d}y$ schreiben wir also immer $\int_{\mathbb{R}} f(y)\,\mathrm{d}y$ und meinen damit $\int_{-b}^{-a} f(y)\,\mathrm{d}y$.

Bemerkung 1.2.6 (Alternativbeweis). Man kann Satz 1.2.5 auch beweisen (Übungsaufgabe), indem man ihn zunächst auf elementare Weise nur für Matrizen beweist, die sich in nur einer Zeile von der Einheitsmatrix unterscheiden, und dann eine beliebige invertierbare Matrix als ein Produkt solcher Matrizen schreibt. $\diamond$

Beispiel 1.2.7 (Volumen eines Parallelepipeds). Der von den Vektoren $a_1, \dots, a_d \in \mathbb{R}^d$ aufgespannte *Parallelepiped* oder auch *Parallelflach* oder *Parallelotop* oder *Spat* ist die Menge

$$P(a_1, \dots, a_d) = \Big\{ \sum_{i=1}^d \lambda_i a_i \colon \lambda_1, \dots, \lambda_d \in [0,1] \Big\}.$$

Sei $A \in \mathbb{R}^{d\times d}$ die Matrix mit den Spalten $a_1, \dots, a_d$, dann kann man als eine Anwendung von Satz 1.2.5 zeigen, dass das Volumen von $P(a_1, \dots, a_d)$ gerade gleich $|\det(A)|$ ist. Hierbei definieren wir das Volumen einer kompakten Menge $K \subset \mathbb{R}^d$ als $\int_K 1\,\mathrm{d}y$ oder auch als $\int_{\mathbb{R}^d} \mathbb{1}_K(x)\,\mathrm{d}x$, wobei $\mathbb{1}_K(x) = 1$ für $x \in K$ und $\mathbb{1}_K(x) = 0$ sonst die Indikatorfunktion auf K ist, haben allerdings solche Integrale für allgemeine kompakte Mengen K noch nicht definiert.[3] Allerdings ist die Definition dieses Integrals für $K = [0,1]^d$ klar, denn es handelt sich dann um ein Riemann-Integral einer stetigen Funktion über einen achsenparallelen Quader. Außerdem ist

$$P(A) = P(a_1, \dots, a_d) = \big\{ A\lambda \colon \lambda \in [0,1]^d \big\} = A([0,1]^d)$$

gleich dem Bild dieses Quaders unter A. Also können wir das Volumen von $P(A)$ im Geiste von Satz 1.2.5 definieren als

$$\int_{P(A)} 1\,\mathrm{d}x = \int_{\mathbb{R}^d} \mathbb{1}_{A([0,1]^d)}(x)\,\mathrm{d}x = \int_{\mathbb{R}^d} \mathbb{1}_{[0,1]^d}(A^{-1}x)\,\mathrm{d}x = \frac{1}{|\det(A^{-1})|} \int_{\mathbb{R}^d} \mathbb{1}_{[0,1]^d}(y)\,\mathrm{d}y = |\det(A)|,$$

falls A invertierbar ist. $\diamond$

Nun wollen wir Satz 1.2.5 verallgemeinern auf beliebige stetig differenzierbare Funktionen an Stelle von A. Dafür benötigen wir ein paar Vorbereitungen. Zunächst bezeichnen wir für eine offene Menge $U \subset \mathbb{R}^d$ mit $\mathcal{C}_{\mathrm{c}}(U)$ die Menge aller stetigen Funktionen $U \to \mathbb{R}$ mit Träger in U. Wenn man eine Funktion $f \in \mathcal{C}_{\mathrm{c}}(U)$ zu einer Funktion $\widetilde{f} \colon \mathbb{R}^d \to \mathbb{R}$ trivial mit $\widetilde{f}(x) = 0$ für $x \in \mathbb{R}^d \setminus U$ fortsetzt, dann ist $\widetilde{f} \in \mathcal{C}_{\mathrm{c}}(\mathbb{R}^d)$, und wir können natürlicherweise $\int_U f(x)\,\mathrm{d}x = \int_{\mathbb{R}^d} \widetilde{f}(x)\,\mathrm{d}x$ setzen. Seien nun U und V zwei offene Teilmengen des $\mathbb{R}^d$. Eine bijektive Abbildung $\varphi \colon U \to V$ heißt *$\mathcal{C}^1$-invertierbar*, falls sowohl φ als auch die Umkehrabbildung $\varphi^{-1} \colon V \to U$ stetig differenzierbar sind. Die Funktionalmatrix $D\varphi = (\frac{\partial \varphi_i}{\partial x_j})_{i,j=1,\dots,d}$ ist dann in jedem $a \in U$ invertierbar mit $D\varphi(a)^{-1} = D\varphi^{-1}(\varphi(a))$. Für jedes $f \in \mathcal{C}_{\mathrm{c}}(V)$ ist $f \circ \varphi \in \mathcal{C}_{\mathrm{c}}(U)$.

Satz 1.2.8 (Transformationssatz für stetige Funktionen mit kompaktem Träger). *Seien U und V offene Teilmengen des $\mathbb{R}^d$ und $\varphi \colon U \to V$ eine $\mathcal{C}^1$-invertierbare Abbildung. Dann gilt für jedes $f \in \mathcal{C}_{\mathrm{c}}(V)$:*

$$\int_U f(\varphi(x))|\det D\varphi(x)|\,\mathrm{d}x = \int_V f(y)\,\mathrm{d}y. \tag{1.2.2}$$

[3] Dies machen wir erst in Abschnitt 1.6.

Symbolisch kann man schreiben: $\mathrm{d}(\varphi(x)) = |\det D\varphi(x)|\,\mathrm{d}x$.

Beweis. Wir führen die Aussage durch lokale Approximation auf Satz 1.2.5 zurück, den Spezialfall für lineare Abbildungen φ mit $U = V = \mathbb{R}^d$. Ein kleiner Überblick über den Beweis ist wie folgt. Die Funktion f wird nach Lemma 1.1.8 mit Hilfe von $f_\varepsilon = \sum_{k\in\mathbb{Z}^d} \theta_{k\varepsilon}\Psi_\varepsilon$ approximiert, wobei wir an die in Bemerkung 1.1.6 eingeführten Zackenfunktionen Ψ_ε und den Verschiebeoperator $\theta_y f(x) = f(x-y)$ erinnern. Wegen Linearität des Integrals muss man also die Transformationsformel nur für die Funktionen $h_{k\varepsilon,\varepsilon} = \theta_{k\varepsilon}\Psi_\varepsilon$ beweisen, und zwar reicht es, sie nur approximativ zu beweisen. Der Fehler muss für $\varepsilon \downarrow 0$ gleichmäßig in den k verschwinden, die einen nichttrivialen Beitrag geben (hier kommt die Kompaktheit des Trägers von f in's Spiel). Um eine approximierende Transformationsformel für $h_{k\varepsilon,\varepsilon}$ zu erhalten, reicht es, die transformierende Funktion φ auf dem Träger von $h_{k\varepsilon,\varepsilon}$ (dem Würfel um $k\varepsilon$ mit Seitenlänge 2ε) mit einer linearen Abbildung zu approximieren. Dies macht man natürlich mit Hilfe der Taylorformel, d. h. mit Hilfe der Differentialmatrix $D\varphi$ im Mittelpunkt dieses Trägers, $k\varepsilon$. Für die linearisierte Version von φ gibt Satz 1.2.5 die exakte Gültigkeit des Transformationssatzes. Es ist dann eine Sache der Technik zu zeigen, dass alle Approximationsfehler im Grenzwert $\varepsilon \downarrow 0$ verschwinden.

Kommen wir zu den Details.

Schritt 1: Vorbereitungen. Es sei $\|x\|_\infty = \max_{i=1}^d |x_i|$ die Maximumsnorm des Vektors $x = (x_1,\dots,x_d) \in \mathbb{R}^d$, dann gilt $\frac{1}{\sqrt{d}}\|x\| \leq \|x\|_\infty \leq \|x\|$. Es sei $W(a,\varepsilon) = \{x \in \mathbb{R}^d : \|x-a\|_\infty \leq \varepsilon\}$ der abgeschlossene Würfel um a mit Seitenlänge 2ε. Sei nun $f \in \mathcal{C}_\mathrm{c}(V)$. Wir dürfen davon ausgehen, dass f in $\mathcal{C}_\mathrm{c}(\mathbb{R}^d)$ liegt und sein Träger $L = \mathrm{supp}\,(f)$ in V liegt. Die Menge $K = \varphi^{-1}(L)$ ist dann eine kompakte Teilmenge von U, und $f\circ\varphi$ verschwindet außerhalb von K. Da K von ∂U und L von ∂V jeweils einen positiven Abstand haben, können wir ein $\varepsilon_1 > 0$ wählen, so dass die abgeschlossene ε_1-Umgebung von K (in Maximumsnorm) in U liegt und die abgeschlossene ε_1-Umgebung von L in V. Nennen wir diese Umgebungen K' bzw. L', dann sind sie kompakt, und daher sind $D\varphi$ auf K' und $D\varphi^{-1}$ auf L' beschränkt, d. h., es gibt ein $C > 0$ mit

$$\|D\varphi(a)\xi\|_\infty \leq C\|\xi\|_\infty \quad \text{und} \quad \|D\varphi^{-1}(b)\xi\|_\infty \leq C\|\xi\|_\infty, \qquad \xi \in \mathbb{R}^d, a \in K', b \in L'.$$

Nach dem Mittelwertsatz gilt für alle $a, x \in U$ mit $[a,x] \subset U$:

$$\varphi(x) - \varphi(a) = \int_0^1 D\varphi(a + t(x-a))\,\mathrm{d}t\,(x-a), \tag{1.2.3}$$

wobei wir $\varphi(x)-\varphi(a)$ und $x-a$ jeweils als einen Spaltenvektor und das Integral komponentenweise als eine Matrix auffassen. Also folgt $\varphi(W(a,\varepsilon)) \subset W(\varphi(a), C\varepsilon)$ für alle $\varepsilon \in (0,\varepsilon_1]$ und alle $a \in K$. Analog haben wir $\varphi^{-1}(W(b,\varepsilon)) \subset W(\varphi^{-1}(b), C\varepsilon)$ für alle $\varepsilon \in (0,\varepsilon_1]$ und alle $b \in L$.

Schritt 2: Approximation durch lineare Abbildungen. Nun approximieren wir φ lokal in jeder Umgebung $W(a,\varepsilon_1)$ eines $a \in K$ mit Hilfe der affin-linearen Funktion

$$\lambda_a\colon \mathbb{R}^d \to \mathbb{R}^d, \qquad \lambda_a(x) = \varphi(a) + D\varphi(a)(x-a).$$

Mit Hilfe von (1.2.3) erhalten wir die Abschätzung

$$\begin{aligned} \|\varphi(x) - \lambda_a(x)\|_\infty &= \Big\| \int_0^1 \big(D\varphi(a + t(x-a)) - D\varphi(a)\big)\,\mathrm{d}t\,(x-a)\Big\|_\infty \\ &\leq \omega_1(\|x-a\|_\infty)\|x-a\|_\infty, \qquad a \in K, x \in W(a,\varepsilon_1), \end{aligned} \tag{1.2.4}$$

wobei $\omega_1\colon [0,\varepsilon_1] \to [0,\infty)$ eine geeignete monotone Funktion ist mit $\lim_{\varepsilon\downarrow 0}\omega_1(\varepsilon) = 0$. (Eine solche Funktion finde man als Übungsaufgabe mit Hilfe der gleichmäßigen Stetigkeit der Komponenten von $D\varphi$ auf K'.)

Schritt 3: approximative Transformationsformel. Man sieht leicht, dass die Zackenfunktionen aus Bemerkung 1.1.6 erfüllen:

$$\|\Psi(x) - \Psi(x')\|_\infty \leq d\|x - x'\|_\infty, \qquad x, x' \in \mathbb{R}^d,$$

also auch

$$\|h_{b,\varepsilon}(x) - h_{b,\varepsilon}(x')\|_\infty \leq \frac{d}{\varepsilon}\|x - x'\|_\infty, \qquad x, x' \in \mathbb{R}^d, \varepsilon > 0, \tag{1.2.5}$$

wobei wir $h_{b,\varepsilon}(x) = \theta_b \Psi_\varepsilon(x)$ setzten. Für $h_{b,\varepsilon}$ haben wir nun eine approximative Transformationsformel: Wir setzen $\varepsilon_2 = \varepsilon_1/C$ und behaupten, dass für alle $b \in L$ und alle $\varepsilon \in (0, \varepsilon_2]$ gilt:

$$\Big| \int_U h_{b,\varepsilon}(\varphi(x)) |\det(D\varphi(x))| \,\mathrm{d}x - \int_V h_{b,\varepsilon}(y)\,\mathrm{d}y \Big| \leq \omega_2(\varepsilon)\varepsilon^d, \tag{1.2.6}$$

wobei $\omega_2\colon (0, \varepsilon_2] \to [0, \infty)$ eine geeignete monotone Funktion mit $\lim_{\varepsilon\downarrow 0} \omega_2(\varepsilon) = 0$ ist. Um dies zu beweisen, setzen wir zunächst $a = \varphi^{-1}(b)$ und schreiben h an Stelle von $h_{b,\varepsilon}$. Wir kombinieren (1.2.4) und (1.2.5), um zu erhalten:

$$|h(\varphi(x)) - h(\lambda_a(x))| \leq \frac{d}{\varepsilon}\omega_1(\|x - a\|_\infty)\|x - a\|_\infty, \qquad x \in W(a, \varepsilon_1).$$

Wir erinnern uns, dass das Bild des Würfels $W(b, \varepsilon)$ unter φ^{-1} im Würfel $W(\varphi^{-1}(b), C\varepsilon)$ liegt und dass $\|D\varphi^{-1}(b)\xi\|_\infty \leq C\|\xi\|_\infty$ für jedes $\xi \in \mathbb{R}^d$ gilt. Also liegt der Träger von $h \circ \varphi = \theta_b\Psi_\varepsilon \circ \varphi$ im Würfel $W(a, C\varepsilon)$, der wegen $C\varepsilon \leq \varepsilon_1$ in $W(a, \varepsilon_1)$, also in K' liegt. Analog sehen wir, dass der Träger von $h \circ \lambda_a$ ebenfalls in $W(a, C\varepsilon)$, also auch in K' liegt. Daher gilt

$$|h(\varphi(x)) - h(\lambda_a(x))| \leq \frac{d}{\varepsilon}\omega_1(C\varepsilon)C\varepsilon = dC\omega_1(C\varepsilon), \qquad x \in U.$$

Da die Abbildung $x \mapsto |\det(D\varphi(x))|$ auf K' gleichmäßig stetig ist, gilt

$$\Big|h(\varphi(x)|\det(D\varphi(x))| - h(\lambda_a(x))|\det(D\varphi(a))|\Big| \leq \widetilde{\omega}(\varepsilon), \qquad x \in U,$$

wobei $\widetilde{\omega}$ unabhängig von $x \in U$ und von $a = \varphi^{-1}$ gewählt werden kann mit $\lim_{\varepsilon\downarrow 0} \widetilde{\omega}(\varepsilon) = 0$. Also haben wir

$$\Big| \int_U h(\varphi(x))|\det(D\varphi(x))|\,\mathrm{d}x - \int_U h(\lambda_a(x))|\det(D\varphi(a))|\,\mathrm{d}x\Big| \leq \int_{W(a,C\varepsilon)} \widetilde{\omega}(\varepsilon)\,\mathrm{d}x = \widetilde{\omega}(\varepsilon)(2C\varepsilon)^d.$$

Da aber Satz 1.2.5 auf h angewendet werden kann und die Beziehung

$$\int_U h(\lambda_a(x))|\det(D\varphi(a))|\,\mathrm{d}x = \int_V h(y)\,\mathrm{d}y$$

liefert, ist (1.2.6) bewiesen mit $\omega_2(\varepsilon) = (2C)^d\widetilde{\omega}(\varepsilon)$.

Schritt 4: Beweisende. Nun approximieren wir f mit den Funktionen

$$f_\varepsilon = \sum_{k\in\mathbb{Z}^d} f(k\varepsilon)\theta_{k\varepsilon}\Psi_\varepsilon = \sum_{k\in\mathbb{Z}^d} f(k\varepsilon)h_{k\varepsilon,\varepsilon}, \qquad \varepsilon > 0,$$

die ja nach Lemma 1.1.8 für $\varepsilon \downarrow 0$ gleichmäßig gegen f konvergieren. Außerdem gilt $\operatorname{supp}(f_\varepsilon) \subset L'$ für jedes $\varepsilon \in (0, \varepsilon_1]$, daher lässt Lemma 1.1.5 folgen, dass $\lim_{\varepsilon\downarrow 0} \int_V f_\varepsilon(y)\,\mathrm{d}y = \int_V f(y)\,\mathrm{d}y$.

Also konvergieren auch die Funktionen $x \mapsto g_\varepsilon(x) = f_\varepsilon(\varphi(x))|\det(D\varphi(x))|$ gegen die Funktion $g(x) = f(\varphi(x))|\det(D\varphi(x))|$, und wir haben $\lim_{\varepsilon\downarrow 0}\int_U g_\varepsilon(x)\,\mathrm{d}x = \int_U g(x)\,\mathrm{d}x$. Also ist nur noch zu zeigen, dass die Differenz

$$\Delta_\varepsilon = \int_U g_\varepsilon(x)\,\mathrm{d}x - \int_V f_\varepsilon(y)\,\mathrm{d}y$$

für $\varepsilon \downarrow 0$ verschwindet. Dies sehen wir mit Hilfe von (1.2.6) wie folgt:

$$\begin{aligned} |\Delta_\varepsilon| &\leq \sum_{k\in\mathbb{Z}^d\colon k\varepsilon\in L} |f(k\varepsilon)| \Big| \int_U h_{k\varepsilon,\varepsilon}(\varphi(x))|\det(D\varphi(x))|\,\mathrm{d}x - \int_V h_{k\varepsilon,\varepsilon}(y)\,\mathrm{d}y\Big| \\ &\leq \#\{k\in\mathbb{Z}^d\colon k\varepsilon\in L\}\sup_{y\in V}|f(y)|\omega_2(\varepsilon)\varepsilon^d. \end{aligned} \tag{1.2.7}$$

Da L kompakt ist, ist der Zählterm $\#\{k\in\mathbb{Z}^d\colon k\varepsilon\in L\}$ höchstens von der Ordnung ε^{-d}. Wegen $\lim_{\varepsilon\downarrow 0}\omega_2(\varepsilon) = 0$ konvergiert die rechte Seite von (1.2.7) für $\varepsilon \downarrow 0$ gegen Null. Das beendet den Beweis. □

Die Bedingung, dass f kompakten Träger hat, ist für viele interessante Anwendungen zu einschränkend, also wollen wir nun den Transformationssatz erweitern auf stetige nichtnegative Funktionen $V \to \mathbb{R}$, für die beide Seiten der Gleichung (1.2.2) sinnvoll sind. Dies müssen wir zunächst genauer erklären:

Definition 1.2.9. *Es sei $V \subset \mathbb{R}^d$ offen und $f\colon V \to [0,\infty)$ stetig. Dann definieren wir*

$$\int_V f(y)\,\mathrm{d}y = \lim_{n\to\infty}\int_V f_n(y)\,\mathrm{d}y \in [0,\infty], \tag{1.2.8}$$

wobei $(f_n)_{n\in\mathbb{N}}$ eine Folge von Funktionen in $\mathcal{C}_\mathrm{c}(V)$ ist mit $f_n \uparrow f$ für $n\to\infty$. Wir nennen f integrierbar*, falls die rechte Seite von* (1.2.8) *endlich ist.*

Dies ist ein Spezialfall des Integrals für halbstetige Funktionen, das wir in Abschnitt 1.5 behandeln werden. Insbesondere werden wir in Lemma 1.5.3 zeigen, dass diese Definition nicht von der Wahl der Folge $(f_n)_{n\in\mathbb{N}}$ abhängt, und in Lemma 1.5.8 (siehe auch Lemma 1.5.4(ii)), dass überhaupt eine solche Folge existiert. Für stetige Funktionen $f\colon V\to\mathbb{R}$, die eventuell beide Vorzeichen annehmen, definieren wir $\int_V f(y)\,\mathrm{d}y$ durch Zerlegung in Positiv- und Negativteil $f = f_+ - f_-$ als $\int_V f_+(y)\,\mathrm{d}y - \int_V f_-(y)\,\mathrm{d}y$, falls f_+ und f_- jeweils im Sinne von Definition 1.2.9 integrierbar sind. Es ist nicht schwer zu sehen, dass im Spezialfall $d = 1$, wenn V ein Intervall ist, dieser Integralbegriff identisch ist mit dem des uneigentlichen Riemann-Integrals für nichtnegative stetige Funktionen. Falls V ein Produkt von Intervallen ist, so ist das in (1.2.8) definierte Integral gleich dem iterierten uneigentlichen Riemann-Integral, wobei man mit Hilfe von Lemma 1.5.3 zeigen kann, dass es nicht auf die Reihenfolge der Grenzwerte der einzelnen eindimensionalen Integrale ankommt. Da wir später noch viel allgemeinere Funktionenklassen integrieren werden und für sie auch den Transformationssatz beweisen werden, verzichten wir an dieser Stelle auf formale Beweise dieser Aussagen.

Nun ist es leicht, den Transformationssatz 1.2.8 auf beliebige stetige Funktionen zu erweitern, wenn sie im Sinne von Definition 1.2.9 integrierbar sind:

Korollar 1.2.10 (Transformationssatz für stetige Funktionen). *Seien U und V offene Teilmengen des $\mathbb{R}^d$ und $\varphi\colon U\to V$ eine $\mathcal{C}^1$-invertierbare Abbildung. Dann gilt für jede stetige*

Funktion $f\colon V \to [0,\infty)$, so dass f_+ und f_- im Sinne von Definition 1.2.9 integrierbar sind:

$$\int_U f(\varphi(x))|\det D\varphi(x)|\,\mathrm{d}x = \int_V f(y)\,\mathrm{d}y.$$

Beweis. Wir approximieren beide Seiten der Gleichung mit Integralen über approximierende Funktionenfolgen in $\mathcal{C}_\mathrm{c}(\mathbb{R}^d)$ und wenden Satz 1.2.8 auf jede der approximierenden Funktionen an. Genauer machen wir Folgendes. Wegen Zerlegung in Positiv- und Negativteil reicht es, f als nichtnegativ voraus zu setzen. Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge in $\mathcal{C}_\mathrm{c}(V)$, die monoton steigend gegen f konvergiert. Dann bilden die Funktionen $g_n(x) = f_n(\varphi(x))|\det D\varphi(x)|$ eine Familie in $\mathcal{C}_\mathrm{c}(U)$, die monoton gegen $g(x) = f(\varphi(x))|\det D\varphi(x)|$ konvergieren. Für jedes $n \in \mathbb{N}$ liefert Satz 1.2.8 die Gültigkeit der behaupteten Formel für f_n an Stelle von f. Nun liefert der Grenzübergang $n \to \infty$ nach Definition 1.2.9 auch die Formel für f. □

Beispiel 1.2.11 (Ebene Polarkoordinaten). Wir betrachten das Komplement $V = \mathbb{R}^2\setminus S$ des Halbstrahls $S = \{\binom{r}{0} : r \in [0,\infty)\}$. Dann bildet $\Phi\colon U \to V$, definiert durch $\Phi(r,\varphi) = \binom{r\cos\varphi}{r\sin\varphi} = \binom{x}{y}$ die Menge $U = (0,\infty) \times (0,2\pi)$ bijektiv auf V ab und ist sogar $\mathcal{C}^1$-invertierbar. (Dies zeigt man, indem man die Umkehrabbildung angibt und ihr die stetige Differenzierbarkeit ansieht.) Man nennt (r,φ) die *Polarkoordinaten* des Punktes (x,y). Die Funktionalmatrix ist

$$D\Phi(r,\varphi) = \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix}. \tag{1.2.9}$$

Also ist $|\det D\Phi(r,\varphi)| = r$. Nach Satz 1.2.10 gilt für jede stetige, im Sinne von Definition 1.2.9 integrierbare Funktion $f\colon V \to \mathbb{R}$:

$$\int_V f\tbinom{x}{y}\,\mathrm{d}x\mathrm{d}y = \int_U f\tbinom{r\cos\varphi}{r\sin\varphi}|\det D\Phi(r,\varphi)|\,\mathrm{d}(r,\varphi) = \int_0^{2\pi}\int_0^\infty r f\tbinom{r\cos\varphi}{r\sin\varphi}\,\mathrm{d}r\mathrm{d}\varphi.$$

Insbesondere wenn f rotationssymmetrisch ist, d. h. von der Form $f\binom{x}{y} = g(\|\binom{x}{y}\|)$ für eine (im eventuell uneigentlichen Riemann-Sinne) integrierbare Funktion $g\colon [0,\infty) \to \mathbb{R}$, so folgt

$$\int_{\mathbb{R}^2} f\tbinom{x}{y}\,\mathrm{d}x\mathrm{d}y = 2\pi\int_0^\infty r g(r)\,\mathrm{d}r,$$

wobei wir das Integral über $V = \mathbb{R}^2\setminus S$ dem Integral über $\mathbb{R}^2$ gleichgesetzt haben, also das Integral über den Halbstrahl ignoriert haben. (Wir werden später begründen, warum dies möglich ist, und wir werden d-dimensionale Varianten kennenlernen; siehe Lemma 3.6.13 und auch Beispiel 1.2.12 für $d = 3$.) Insbesondere können wir den Wert des uneigentlichen Integrals von e^{-x^2} über $\mathbb{R}$ berechen:

$$\begin{aligned}\Big(\int_{\mathbb{R}} \mathrm{e}^{-x^2}\,\mathrm{d}x\Big)^2 &= \int_{\mathbb{R}}\int_{\mathbb{R}} \mathrm{e}^{-x^2}\mathrm{e}^{-y^2}\,\mathrm{d}x\mathrm{d}y = \int_V \mathrm{e}^{-(x^2+y^2)}\,\mathrm{d}(x,y) = 2\pi\int_0^\infty r\mathrm{e}^{-r^2}\,\mathrm{d}r\\ &= -2\pi\frac{1}{2}\mathrm{e}^{-r^2}\Big|_0^\infty = \pi.\end{aligned}$$

Zwar gibt es für $x \mapsto \mathrm{e}^{-x^2}$ keine explizite Stammfunktion, aber der Trick des Quadrierens des Integrals führt auf eine zweidimensionale Variante, für die nach Ausnutzung der Rotationssymmetrie leicht eine Stammfunktion gesehen werden kann. ◇

Beispiel 1.2.12 (Räumliche Polarkoordinaten). Sei H die Halbebene $S\times\mathbb{R} = \{(r,0,z)\colon r \geq 0, z \in \mathbb{R}\}$ im $\mathbb{R}^3$. Wir betrachten $V = \mathbb{R}^3 \setminus H$ und $U = (0,\infty) \times (0,\pi) \times (0,2\pi)$. Die Abbildung $\Phi\colon U \to V$, definiert durch

$$\Phi(r,\vartheta,\varphi) = \begin{pmatrix} r\sin\vartheta\cos\varphi \\ r\sin\vartheta\sin\varphi \\ r\cos\vartheta \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

ist $\mathcal{C}^1$-invertierbar. Man nennt (r,ϑ,φ) die *Polarkoordinaten* des Punktes (x,y,z). Die Funktionalmatrix ist

$$D\Phi(r,\vartheta,\varphi) = \begin{pmatrix} \sin\vartheta\cos\varphi & r\cos\vartheta\cos\varphi & -r\sin\vartheta\sin\varphi \\ \sin\vartheta\sin\varphi & r\cos\vartheta\sin\varphi & r\sin\vartheta\cos\varphi \\ \cos\vartheta & -r\sin\vartheta & 0 \end{pmatrix},$$

also gilt $|\det D\Phi(r,\varphi)| = r^2\sin\vartheta$. Nach Satz 1.2.10 gilt für jede stetige, im Sinne von Definition 1.2.9 integrierbare Funktion $f\colon V \to \mathbb{R}$:

$$\begin{aligned}\int_V f(x,y,z)\,\mathrm{d}x\mathrm{d}y\mathrm{d}z &= \int_U f(\Phi(r,\vartheta,\varphi))|\det D\Phi(r,\vartheta,\varphi)|\,\mathrm{d}(r,\vartheta,\varphi) \\ &= \int_0^{2\pi}\int_0^{\pi}\int_0^{\infty} r^2\sin\vartheta f\big(r\sin\vartheta\cos\varphi, r\sin\vartheta\sin\varphi, r\cos\vartheta\big)\,\mathrm{d}r\mathrm{d}\vartheta\mathrm{d}\varphi.\end{aligned}$$

Insbesondere wenn f rotationssymmetrisch ist, d. h. von der Form $f(x) = g(\|x\|)$ für eine (im eventuell uneigentlichen Riemann-Sinne) integrierbare Funktion $g\colon [0,\infty) \to \mathbb{R}$, so folgt

$$\int_{\mathbb{R}^3} \widehat{f}(\|x\|)\,\mathrm{d}x = \int_0^{2\pi}\int_0^{\pi}\int_0^{\infty} r^2 g(r)\,\mathrm{d}r\mathrm{d}\vartheta\mathrm{d}\varphi = 4\pi\int_0^{\infty} r^2 g(r)\,\mathrm{d}r,$$

wobei wir wieder das Integral über die Halbebene H ignoriert haben. ◇

Beispiel 1.2.13 (Jacobi-Transformation). Definiere $J\colon \mathbb{R}^2 \to \mathbb{R}^2$ durch $J\binom{u}{v} = \binom{u(1-v)}{uv}$. Dann ist J eine $\mathcal{C}^1$-invertierbare Abbildung von $(0,\infty)\times(0,1)$ auf $(0,\infty)^2$. Daher liefert eine Anwendung von Satz 1.2.10 auf die Abbildung $(x,y) \mapsto x^{p-1}\mathrm{e}^{-x}y^{q-1}\mathrm{e}^{-y}$ die folgende Identifikation des *Euler'schen Beta-Integrals*:

$$B(p,q) = \int_0^1 (1-t)^{p-1}t^{q-1}\,\mathrm{d}t = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}, \qquad p,q \in (0,\infty),$$

wobei Γ die Gamma-Funktion bezeichnet. (Übungsaufgabe.)

◇

1.3 Partielle Integration

In diesem Abschnitt formulieren wir mehrdimensionale Versionen der Regel von der partiellen Integration. Dafür ist es zweckmäßig, Differentialoperatoren einzuführen.

Zunächst führen wir ein paar natürliche Bezeichnungen ein. Sei $U \subset \mathbb{R}^d$ eine offene Menge. Mit $\mathcal{C}(U)$ bzw. $\mathcal{C}^n(U)$ bezeichnen wir die Menge der stetigen bzw. n Mal stetig differenzierbaren Funktionen $U \to \mathbb{R}$, wobei auch $n = \infty$ eingeschlossen ist. Funktionen in $\mathcal{C}^\infty(U)$ nennt man auch manchmal *glatt*. Statt $\mathcal{C}(U)$ kann man auch $\mathcal{C}^0(U)$ schreiben. Der Index 'c' bedeutet wie bisher, dass wir zusätzlich einen kompakten Träger in U fordern. Es ist klar, dass für offene Mengen $V \subset U \subset \mathbb{R}^d$ jede Funktion in $\mathcal{C}^n_\mathrm{c}(V)$ trivial mit Null zu einer Funktion in $\mathcal{C}^n_\mathrm{c}(U)$ fortgesetzt werden kann.

1.3.1 Glatte Teilung der Eins

Als Verstärkung der Approximationstechniken aus Abschnitt 1.2 (siehe insbesondere auch Bemerkung 1.1.6) konstruieren wir nun unendlich oft differenzierbare Funktionen mit gewissen Eigenschaften. Betrachten wir $g\colon \mathbb{R} \to \mathbb{R}$, definiert durch

$$g(t) = \begin{cases} \exp\left(-\frac{1}{1-t^2}\right), & \text{falls } |t| < 1, \\ 0 & \text{sonst,} \end{cases}$$

dann liegt g in $\mathcal{C}_\mathrm{c}^\infty(\mathbb{R})$.[4] Die Funktion $G(t) = \sum_{k\in\mathbb{Z}} g(t-k)$ ist beliebig oft differenzierbar in $\mathbb{R}$, überall ungleich Null und erfüllt $G(t) = G(t-k)$ für jedes $t \in \mathbb{R}$ und $k \in \mathbb{Z}$. Setzt man $h(t) = g(t)/G(t)$ für $t \in \mathbb{R}$, so ist $h \in \mathcal{C}_\mathrm{c}^\infty(\mathbb{R})$ mit $\operatorname{supp}(h) = [-1, 1]$ und $\sum_{k\in\mathbb{Z}} h(t-k) = 1$ für alle $t \in \mathbb{R}$. Also bildet die Familie der $\theta_k h$ mit $k \in \mathbb{Z}$ eine glatte Teilung der Eins in $\mathbb{R}$, wobei wir wieder an den Verschiebungsoperator $\theta_y f(x) = f(x-y)$ erinnern. Somit spielt h die selbe Rolle wie ψ in Bemerkung 1.1.6, aber ist glatt. Nun bilden wir noch Tensorprodukte und reskalieren mit einem kleinen Parameter: Für $k \in \mathbb{Z}^d$ und $\varepsilon > 0$ sei

$$\alpha_{k,\varepsilon}(x) = \prod_{i=1}^d h\Big(\frac{x_i}{\varepsilon} - k_i\varepsilon\Big) = \bigotimes_{i=1}^d (\theta_{k_i\varepsilon} h)(x/\varepsilon) = \theta_{k\varepsilon} h_\varepsilon^{\otimes d}(x), \qquad x \in \mathbb{R}^d,$$

wobei wir $h_\varepsilon(x) = h(x/\varepsilon)$ setzten. Dann ist $\alpha_{k,\varepsilon} \in \mathcal{C}_\mathrm{c}^\infty(\mathbb{R}^d)$ und $\sum_{k\in\mathbb{Z}^d} \alpha_{k,\varepsilon} = 1$. Ferner ist der Träger von $\alpha_{k,\varepsilon}$ gleich dem Würfel $W(k\varepsilon, \varepsilon) = \{x \in \mathbb{R}^d \colon \|x - k\varepsilon\|_\infty \leq \varepsilon\}$.

Mit dieser Teilung der Eins kann man jede stetige Funktion außerhalb eines Kompaktum glatt auf Null 'herunterdrücken':

Lemma 1.3.1. *Sei $U \subset \mathbb{R}^d$ offen und $K \subset U$ kompakt. Sei $n \in \mathbb{N} \cup \{\infty\}$. Dann gibt es zu jeder n Mal stetig differenzierbaren Funktion $f\colon U \to \mathbb{R}$ eine Funktion $\beta \in \mathcal{C}_\mathrm{c}^n(U)$ mit $\beta|_K = f$.*

Beweis. Zunächst behandeln wir nur den Fall, dass $f = 1$ auf K. Wir wählen $\varepsilon > 0$ so klein, dass jeder abgeschlossene Würfel der Seitenlänge 2ε, der K trifft, ganz in U enthalten ist. Sei $P = \{k \in \mathbb{Z}^d \colon \operatorname{supp}(\alpha_{k,\varepsilon}) \cap K \neq \emptyset\}$ die Menge der Multiindizes k, so dass der Träger von $\alpha_{k,\varepsilon}$ das Kompaktum K trifft. Wir zeigen, dass $\beta = \sum_{k\in P} \alpha_{k,\varepsilon}$ die gewünschten Eigenschaften hat. Zunächst sieht man, dass der Träger von β gleich der Vereinigung der Träger der $\alpha_{k,\varepsilon}$ mit $k \in P$ ist, also selber auch in U enthalten ist, d. h. $\beta \in \mathcal{C}_\mathrm{c}^\infty(U)$. Außerdem ist klar, dass $\beta(x) = 1$ für jedes $x \in K$. Also ist das Lemma bewiesen für den Fall, dass $f = 1$ auf K. Im Fall einer allgemeinen n Mal stetig differenzierbaren Funktion $f\colon K \to \mathbb{R}$ multiplizieren wir die obige Funktion β mit f und erhalten eine Funktion aus $\mathcal{C}_\mathrm{c}^n(U)$, die auf K mit f übereinstimmt. □

Mit Hilfe von Lemma 1.3.1 kann man auch einen Beweis der Existenz von Funktionenfolgen $(f_n)_{n\in\mathbb{N}}$ führen, wie sie in Definition 1.2.9 vorausgesetzt werden, siehe Lemma 1.5.4(ii). Außerdem können wir mit Hilfe der Funktionen $\alpha_{k,\varepsilon}$ zeigen, dass eine stetige Funktion schon Null ist, wenn ihre Integrale gegen alle Funktionen aus $\mathcal{C}_\mathrm{c}^\infty(U)$ verschwindet:

Lemma 1.3.2. *Sei $U \subset \mathbb{R}^d$ offen und $h \in \mathcal{C}(U)$ mit der Eigenschaft, dass $\int_U h(x)g(x)\,\mathrm{d}x = 0$ für jedes $g \in \mathcal{C}_\mathrm{c}^\infty(U)$. Dann ist $h(x) = 0$ für jedes $x \in U$.*

---

[4]Dies beweist man ähnlich wie in [Fo1, Bsp. (22.2)], wo gezeigt wurde, dass die Taylorreihe der Funktion $x \mapsto \mathrm{e}^{-1/x^2}$ mit Entwicklungspunkt Null konstant gleich Null ist.

Beweis. Angenommen, für ein $a \in U$ gilt $h(a) \neq 0$, also etwa $h(a) > 0$. Wegen der Stetigkeit von h gibt es dann eine offene Umgebung $V \subset U$ von a und ein $\delta > 0$ mit $h(x) \geq \delta$ für $x \in V$. Nach (dem Beweis von) Lemma 1.3.1 gibt es eine nichtnegative Funktion $g \in \mathcal{C}_c^\infty(V)$ mit $g(a) = 1$. Dann haben wir

$$\int_U h(x)g(x)\,\mathrm{d}x = \int_V h(x)g(x)\,\mathrm{d}x \geq \delta \int_V g(x)\,\mathrm{d}x > 0,$$

und dies beendet den Beweis. □

1.3.2 Partielle Integration im $\mathbb{R}^d$

Wir formulieren nun die Regeln der partiellen Integration für stetige Funktionen $\mathbb{R}^d \to \mathbb{R}$ mit kompaktem Träger. Wegen der Kompaktheit des Trägers ist dies eine sehr leichte Aufgabe. Wesentlich schwieriger wird die Sache, wenn Randterme zu beachten sind und das Integrationsgebiet nicht einfach durch einen Quader ausgedrückt werden kann. Diesen interessanten Fall werden wir später im Zusammenhang mit dem *Gauß'schen Integralsatz* behandeln.

Wir erinnern an die Notation $D_i = \frac{\partial}{\partial x_i}$ für die partielle Ableitung nach der i-ten Koordinaten. Mit $D_i^2 = \frac{\partial^2}{\partial x_i^2}$ wird die zweite Ableitung nach x_i bezeichnet.

Lemma 1.3.3 (Partielle Integration). *Sei $U \subset \mathbb{R}^d$ offen und $i \in \{1, \ldots, d\}$. Dann gelten*

(i) $\int_U D_i\varphi(x)\,\mathrm{d}x = 0$ *für jedes* $\varphi \in \mathcal{C}_c^1(U)$ *und*

(ii) $\int_U D_i f(x)g(x)\,\mathrm{d}x = -\int_U f(x)D_i g(x)\,\mathrm{d}x$ *für jedes* $f \in \mathcal{C}^1(U)$ *und jedes* $g \in \mathcal{C}_c^1(U)$.

Beweis. Man kann (ii) leicht aus (i) folgern, denn $fg \in \mathcal{C}_c^1(U)$ und $D_i(fg) = D_i f \cdot g + f \cdot D_i g$. Um (i) zu zeigen, kann man ohne Einschränkung annehmen, dass $U = \mathbb{R}^d$ und $i = d = 1$. Dann wählt man $R > 0$ so groß, dass $\operatorname{supp}(\varphi) \subset [-R, R]$ und hat nach dem Hauptsatz der Differential- und Integralrechnung

$$\int_\mathbb{R} D_1\varphi(x)\,\mathrm{d}x = \int_{-R}^{R} \varphi'(x)\,\mathrm{d}x = \varphi\Big|_{-R}^{R} = \varphi(R) - \varphi(-R) = 0.$$

□

Beispiel 1.3.4. Seien $U \subset \mathbb{R}^d$ offen und $f \in \mathcal{C}^2(U)$ und $g \in \mathcal{C}_c^2(U)$. Dann gilt für jedes $i \in \{1, \ldots, d\}$:

$$\int_U D_i^2 f(x)g(x)\,\mathrm{d}x = -\int_U D_i f(x)D_i g(x)\,\mathrm{d}x = \int_U f(x)D_i^2 g(x)\,\mathrm{d}x.$$

Summation über $i \in \{1, \ldots, d\}$ ergibt

$$\int_U (\Delta f)g\,\mathrm{d}x = -\int_U \langle \nabla f, \nabla g\rangle\,\mathrm{d}x = \int_U f(\Delta g)\,\mathrm{d}x,$$

wobei wir an den Laplace-Operator $\Delta = \sum_{i=1}^d D_i^2$ und den Gradienten $\nabla = (D_1, \ldots, D_d)$ erinnern. ◇

1.3.3 Lineare Differentialoperatoren und ihre Adjungierte

Wir erinnnern an die kompakte Notation

$$|\alpha| = \alpha_1 + \cdots + \alpha_d \qquad \text{und} \qquad D^\alpha = D_1^{\alpha_1} \dots D_d^{\alpha_d}$$

für $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}^d$. Dann ist ein Operator der Gestalt

$$L = \sum_{\alpha \in \mathbb{N}^d \colon |\alpha| \leq n} a_\alpha D^\alpha, \qquad n \in \mathbb{N},$$

ein *linearer Differentialoperator* der Ordnung n in einer offenen Menge $U \subset \mathbb{R}^d$, wobei wir einfach annahmen, dass die Funktionen a_α in $\mathcal{C}^\infty(U)$ liegen. Dann definiert L eine lineare Abbildung $\mathcal{C}^m(U) \to \mathcal{C}^{m-n}(U)$ für $m \in \{n, n+1, \dots\} \cup \{\infty\}$, wobei $\infty - n = \infty$. Wenn L die Ordnung $n = 0$ hat, so ist L der Multiplikationsoperator mit der Funktion a_0. Lineare Differentialoperatoren können auf natürliche Weise mit Skalaren multipliziert werden und mit einander addiert werden, bilden also einen Vektorraum. Ferner kann man ebenfalls auf natürliche Weise eine Multiplikation erklären, die Komposition: Wenn $L_2 \colon \mathcal{C}^m(U) \to \mathcal{C}^{m-n_2}(U)$ die Ordnung n_2 hat und $L_1 \colon \mathcal{C}^{m-n_2}(U) \to \mathcal{C}^{m-n_1-n_2}(U)$ die Ordnung n_1 (natürlich unter der Voraussetzung $m \geq n_1 + n_2$), so ist $L_1 \circ L_2 \colon \mathcal{C}^m(U) \to \mathcal{C}^{m-n_1-n_2}(U)$. Es ist klar, dass $L_1 \circ L_2$ ein linearer Differentialoperator der Ordnung $n_1 + n_2$ ist. Dies zeigt man per Induktion, so dass es reicht, den Fall $L_1 = D_i$ und $L_2 = aD^\alpha$ zu betrachten, wobei $i \in \{1, \dots, d\}$, $\alpha \in \mathbb{N}^d$ und $a \in \mathcal{C}_\infty(U)$ seien. Dann ist nämlich für jedes $f \in \mathcal{C}^{|\alpha|+1}(U)$

$$(L_1 \circ L_2)f = L_1(L_2 f) = D_i(aD^\alpha f) = (D_i a)(D^\alpha f) + aD_i D^\alpha f = \big(aD_i D^\alpha + (D_i a)D^\alpha\big)f,$$

also ist $L_1 \circ L_2$ ein linearer Differentialoperator der Ordnung $|\alpha| + 1$.

Bemerkung 1.3.5 (Nicht-Kommutativität). Man beachte, dass die Komposition von linearen Differentialoperatoren nicht kommutativ ist. Zum Beispiel ist der *Kommutator*

$$[L_1, L_2] = L_1 \circ L_2 - L_2 \circ L_1$$

im Allgemeinen nicht der Null-Operator. Wenn etwa $L_1 = D_i$ und $L_2 = x_j$ (Multiplikation mit der Funktion $x \mapsto x_j$), dann ist

$$[L_1, L_2]f = L_1(L_2 f) - L_2(L_1 f) = D_i(x_j f) - x_j D_i f = \delta_{i,j} f + x_j D_i f - x_j D_i f = \delta_{i,j} f,$$

wobei $\delta_{i,j}$ das Kroneckersymbol bezeichnet. In Kurzschreibweise: $[D_i, x_j] = \delta_{i,j}$. ◇

Nun können wir mit Hilfe der partiellen Integration aus Lemma 1.3.3 den Begriff eines adjungierten linearen Differentialoperators einführen:

Satz 1.3.6. *Sei $U \subset \mathbb{R}^d$ offen und L ein linearer Differentialoperator der Ordnung $n \in \mathbb{N}$ in U. Dann gibt es genau einen linearen Differentialoperator L^* der Ordnung n mit*

$$\int_U (L^* f) g \,\mathrm{d}x = \int_U f (Lg) \,\mathrm{d}x, \qquad f \in \mathcal{C}^n(U), g \in \mathcal{C}_\mathrm{c}^n(U). \tag{1.3.1}$$

Man nennt L^* den zu L adjungierten linearen Differentialoperator. *Es gelten die Rechenregeln*

$$(\lambda L)^* = \lambda L^*, \qquad (L_1 + L_2)^* = L_1^* + L_2^*, \qquad (L_1 \circ L_2)^* = L_2^* \circ L_1^* \tag{1.3.2}$$

für $\lambda \in \mathbb{R}$ und lineare Differentialoperatoren L, L_1 und L_2.

Beweis. Zunächst zeigen wir die Eindeutigkeit. Falls M_1 und M_2 zwei lineare Differentialoperatoren der Ordnung n sind, die die Eigenschaft in (1.3.1) an Stelle von L^* haben, so gilt für jedes $f \in \mathcal{C}^n(U)$:

$$\int_U (M_1 f - M_2 f) g \,\mathrm{d}x = \int_U f(Lg - Lg)\,\mathrm{d}x = 0, \qquad g \in \mathcal{C}^n_{\mathrm{c}}(U).$$

Nach Lemma 1.3.2 folgt, dass $M_1 f - M_2 f = 0$. Also ist $M_1 = M_2$.

Leicht ist zu zeigen, dass, falls $\lambda \in \mathbb{R}$ und lineare Differentialoperatoren L, L_1 und L_2 gegeben sind, auch $(\lambda L)^*$, $(L_1 + L_2)^*$ und $(L_1 \circ L_2)^*$ gegeben sind und die Formeln in (1.3.2) erfüllen. Dies folgt nämlich daraus, dass für alle $f \in \mathcal{C}^n(U)$ und $g \in \mathcal{C}^n_{\mathrm{c}}(U)$ gilt:

$$\begin{aligned}
\int (\lambda L^* f) g &= \lambda \int (L^* f) g = \lambda \int f L g = \int f(\lambda L g),\\
\int [(L_1^* + L_2^*) f] g &= \int (L_1^* f) g + \int (L_2^* f) g = \int f(L_1 g) + \int f(L_2 g) = \int f[(L_1 + L_2) g],\\
\int (L_1^* \circ L_2^* f) g &= \int (L_2^* f)(L_1 g) = \int f(L_2 \circ L_1 g).
\end{aligned}$$

Für einen Operator $L = a$ der Ordnung Null mit $a \in \mathcal{C}^\infty(U)$ ist $L^* = a$, denn $\int (af)g = \int f(ag)$. Für $L = D_i$ ist $L^* = -D_i$ nach Lemma 1.3.3(ii). Da sich jeder lineare Differentialoperator durch Addition, skalare Multiplikation und Komposition aus diesen speziellen Operatoren zusammensetzen lässt, erhalten wir die Existenz des Adjungierten zu jedem linearen Differentialoperator. □

Beispiel 1.3.7. (i) Für $a_1, \dots, a_d \in \mathcal{C}^\infty(U)$ ist

$$\Big(\sum_{i=1}^d a_i D_i\Big)^* = \sum_{i=1}^d (a_i D_i)^* = \sum_{i=1}^d (-D_i) a_i = -\sum_{i=1}^d a_i D_i - \sum_{i=1}^d (D_i a_i).$$

Man beachte, dass in dieser Kurzschreibweise die Klammerung entscheidend ist. Zum Beispiel ist $(D_i a_i)$ der Multiplikationsoperator mit der Funktion $D_i a_i$, aber $D_i a_i$ ist der Operator $f \mapsto D_i(a_i f)$, und man muss die Produktregel anwenden.

(ii) Für $n \in \mathbb{N}$ und $c_\alpha \in \mathbb{R}$ für $\alpha \in \mathbb{N}^d$ mit $|\alpha| \leq n$ ist

$$\Big(\sum_{\alpha \in \mathbb{N}^d \colon |\alpha| \leq n} c_\alpha D^\alpha\Big)^* = \sum_{\alpha \in \mathbb{N}^d \colon |\alpha| \leq n} (-1)^{|\alpha|} c_\alpha D^\alpha.$$

Insbesondere gilt $\Delta^* = \Delta$, und man sagt, der Laplace-Operator sei *selbstadjungiert* (siehe auch Beispiel 1.3.4). ◇

1.4 Der Laplace-Operator in krummlinigen Koordinaten

Wir wollen nun den Transformationssatz benutzen, um den Laplace-Operator in krummlinigen Koordinaten darzustellen. Seien $\Omega, \Omega' \subset \mathbb{R}^d$ offen und $\Phi = (\Phi_1, \dots, \Phi_d) \colon \Omega' \to \Omega$ eine $\mathcal{C}^2$-invertierbare Abbildung, also eine bijektive Abbildung, so dass sowohl Φ als auch Φ^{-1} zweimal stetig differenzierbar sind. Wir betrachten die Matrix

$$G(y) = \big(D\Phi(y)\big)^{\mathrm{T}}\big(D\Phi(y)\big) = \sum_{\nu=1}^{d}\Big(\frac{\partial \Phi_\nu(y)}{\partial y_i}\,\frac{\partial \Phi_\nu(y)}{\partial y_j}\Big)_{i,j=1,\dots,d},$$

wobei wir daran erinnern, dass $D\Phi = (\frac{\partial \Phi_\nu}{\partial y_j})_{\nu,j=1,\dots,d}$ die Differentialmatrix von Φ ist. Nach dem Determinantenmultiplikationssatz ist $\sqrt{\det(G)} = |\det(D\Phi)|$, also kann der Transformationssatz (Satz 1.2.8) auch formuliert werden als

$$\int_{\Omega'} u(\Phi(y))\sqrt{\det(G(y))}\,\mathrm{d}y = \int_{\Omega} u(x)\,\mathrm{d}x, \qquad u \in \mathcal{C}_{\mathrm{c}}(\Omega). \tag{1.4.1}$$

Wir fassen hier alle Vektoren $a, b \in \mathbb{R}^d$ als Zeilenvektoren auf und schreiben das Skalarprodukt als $\langle a, b\rangle = ab^{\mathrm{T}}$. Zunächst benötigen wir die folgende Hilfsaussage.

Lemma 1.4.1. *Für alle $u, v \in \mathcal{C}^1(\Omega)$ gilt*

$$\big\langle (\nabla u)\circ\Phi, (\nabla v)\circ\Phi\big\rangle = \big\langle \nabla(u\circ\Phi)G^{-1}, \nabla(v\circ\Phi)\big\rangle,$$

ausgeschrieben:

$$\sum_{i=1}^{d}\frac{\partial u}{\partial x_i}(\Phi(y))\,\frac{\partial v}{\partial x_i}(\Phi(y)) = \sum_{k,l=1}^{d} G^{-1}_{k,l}(y)\frac{\partial(u\circ\Phi)}{\partial y_k}(y)\frac{\partial(v\circ\Phi)}{\partial y_l}(y), \qquad y \in \Omega',$$

wobei $G^{-1} = (G^{-1}_{k,l})_{k,l=1,\dots,d}$.

Beweis. Die Kettenregel ergibt

$$\nabla(u\circ\Phi) = [(\nabla u)\circ\Phi]\,D\Phi,$$

wobei ∇u als ein Zeilenvektor aufgefasst wird, und eine analoge Formel für v statt u. Also haben wir wegen $G^{-1} = (D\Phi)^{-1}((D\Phi)^{\mathrm{T}})^{-1}$:

$$\begin{aligned}
\big\langle \nabla(u\circ\Phi)G^{-1}, \nabla(v\circ\Phi)\big\rangle &= \big\langle [(\nabla u)\circ\Phi]\,(D\Phi)(D\Phi)^{-1}((D\Phi)^{\mathrm{T}})^{-1}, [(\nabla v)\circ\Phi]\,D\Phi\big\rangle \\
&= [(\nabla u)\circ\Phi]((D\Phi)^{\mathrm{T}})^{-1}\Big([(\nabla v)\circ\Phi]\,D\Phi\Big)^{\mathrm{T}} \\
&= [(\nabla u)\circ\Phi]((D\Phi)^{\mathrm{T}})^{-1}(D\Phi)^{\mathrm{T}}[(\nabla v)\circ\Phi]^{\mathrm{T}} \\
&= \big\langle (\nabla u)\circ\Phi, (\nabla v)\circ\Phi\big\rangle.
\end{aligned}$$

□

Wir definieren nun die Transformation Δ^Φ des Laplace-Operators $\Delta = \sum_{i=1}^{d}\frac{\partial^2}{\partial x_i^2}$ bezüglich Φ durch

$$\Delta^\Phi u = \big(\Delta(u\circ\Phi^{-1})\big)\circ\Phi, \qquad u \in \mathcal{C}^2(\Omega'). \tag{1.4.2}$$

In Termen der oben definierten Matrix $G = (D\Phi)^{\mathrm{T}}(D\Phi)$ und ihrer Inversen können wir Δ^Φ wie folgt identifizieren:

Satz 1.4.2.

$$\begin{aligned}\Delta^\Phi &= \frac{1}{\sqrt{\det(G)}} \sum_{k,l=1}^d \frac{\partial}{\partial y_l}\Big(G^{-1}_{k,l}\sqrt{\det(G)}\frac{\partial}{\partial y_k}\Big) \\ &= \sum_{k,l=1}^d G^{-1}_{k,l}\frac{\partial^2}{\partial y_k\,\partial y_l} + \frac{1}{\sqrt{\det(G)}}\sum_{k,l=1}^d \frac{\partial(G^{-1}_{k,l}\sqrt{\det(G)})}{\partial y_l}\,\frac{\partial}{\partial y_k}.\end{aligned} \tag{1.4.3}$$

Beweis. Seien $u \in \mathcal{C}^2(\Omega)$ und $v \in \mathcal{C}^2_\mathrm{c}(\Omega)$, und wir setzen $\widetilde{u} = u \circ \Phi$ und $\widetilde{v} = v \circ \Phi$. Nach der Transformationsformel in der Form (1.4.1) haben wir

$$\int_\Omega (\Delta u)v\,\mathrm{d}x = \int_{\Omega'} \Delta u(\Phi(y))v(\Phi(y))\sqrt{\det(G(y))}\,\mathrm{d}y = \int_{\Omega'}(\Delta^\Phi\widetilde{u})\widetilde{v}\sqrt{\det(G)}\,\mathrm{d}y.$$

Andererseits gilt nach Beispiel 1.3.4 und Lemma 1.4.1:

$$\begin{aligned}\int_\Omega (\Delta u)v\,\mathrm{d}x &= -\int_\Omega \langle\nabla u, \nabla v\rangle\,\mathrm{d}x \\ &= -\int_{\Omega'} \big(\langle\nabla u, \nabla v\rangle\big)\circ\Phi\sqrt{\det(G)}\,\mathrm{d}y \\ &= -\int_{\Omega'}\sum_{k,l=1}^d G^{-1}_{k,l}\frac{\partial\widetilde{u}}{\partial y_k}\frac{\partial\widetilde{v}}{\partial y_l}\sqrt{\det(G)}\,\mathrm{d}y \\ &= -\int_{\Omega'}\sum_{k,l=1}^d \Big(G^{-1}_{k,l}\sqrt{\det(G)}\frac{\partial\widetilde{u}}{\partial y_k}\Big)\frac{\partial\widetilde{v}}{\partial y_l}\,\mathrm{d}y \\ &= \int_{\Omega'}\sum_{k,l=1}^d \frac{\partial}{\partial y_l}\Big(G^{-1}_{k,l}\sqrt{\det(G)}\frac{\partial\widetilde{u}}{\partial y_k}\Big)\widetilde{v}\,\mathrm{d}y,\end{aligned}$$

wobei wir im letzten Schritt noch einmal partielle Integration anwendeten (siehe Satz 1.3.3). Wenn man die beiden Rechnungen kombiniert, erhält man

$$\int_{\Omega'}(\Delta^\Phi\widetilde{u})\widetilde{v}\sqrt{\det(G)}\,\mathrm{d}y = \int_{\Omega'}\sum_{k,l=1}^d \frac{\partial}{\partial y_l}\Big(G^{-1}_{k,l}\sqrt{\det(G)}\frac{\partial\widetilde{u}}{\partial y_k}\Big)\widetilde{v}\,\mathrm{d}y.$$

Da $\widetilde{v} \in \mathcal{C}^2_\mathrm{c}(\Omega')$ beliebig ist, impliziert Lemma 1.3.2, dass

$$\Delta^\Phi\widetilde{u} = \frac{1}{\sqrt{\det(G)}}\sum_{k,l=1}^d \frac{\partial}{\partial y_l}\Big(G^{-1}_{k,l}\sqrt{\det(G)}\frac{\partial}{\partial y_k}\Big)\widetilde{u}.$$

Da $\widetilde{u} \in \mathcal{C}(\Omega')$ beliebig ist, ist der Beweis der ersten Gleichung beendet. Die zweite folgt direkt aus einer Anwendung der Produktregel. □

Beispiel 1.4.3 (Ebene Polarkoordinaten). Wie in Beispiel 1.2.11 betrachten wir das Komplement $\Omega = \mathbb{R}^2 \setminus S$ des Halbstrahls $S = \{\binom{r}{0} : r \in [0,\infty)\}$. Dann bildet $\Phi\colon \Omega' \to \Omega$, definiert durch $\Phi(r,\varphi) = r\binom{\cos\varphi}{\sin\varphi} = \binom{x}{y}$ die Menge $\Omega' = (0,\infty)\times(0,2\pi)$ bijektiv auf Ω ab und ist sogar $\mathcal{C}^1$-invertierbar. Die Funktionalmatrix ist

$$D\Phi(r,\varphi) = \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix}, \quad \text{also } G(r,\varphi) = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix} \quad \text{und } G^{-1}(r,\varphi) = \begin{pmatrix} 1 & 0 \\ 0 & r^{-2} \end{pmatrix}. \tag{1.4.4}$$

Insbesondere ist $\sqrt{\det(G(r,\varphi))} = r$. Für den Laplace-Operator in ebenen Polarkoordinaten ergibt sich daher (man benutze die zweite Zeile von (1.4.3) mit $y_1 = r$ und $y_2 = \varphi$)

$$\Delta^\Phi = \frac{1}{r}\frac{\partial}{\partial r}\Big(r\frac{\partial}{\partial r}\Big) + \frac{1}{r}\frac{\partial}{\partial\varphi}\Big(\frac{1}{r}\frac{\partial}{\partial\varphi}\Big) = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial\varphi^2}.$$

Sei zum Beispiel $u\colon \mathbb{R}^2\setminus\{0\}\to\mathbb{R}$ eine Funktion, die in Polarkoordinaten gegeben wird durch

$$\widetilde{u}(r,\varphi) = u(\Phi(r,\varphi)) = r^m\cos(m\varphi)$$

mit einem $m\in\mathbb{Z}$, dann ist

$$\Delta^\Phi\widetilde{u} = \frac{1}{r}\frac{\partial}{\partial r}\Big(r\frac{\partial r^m}{\partial r}\Big)\cos(m\varphi) + r^{m-2}\frac{\partial^2\cos(m\varphi)}{\partial\varphi^2} = m^2r^{m-2}\cos(m\varphi) - r^{m-2}m^2\cos(m\varphi) = 0,$$

also erweist sich u als harmonisch. $\diamond$

Beispiel 1.4.4 (Räumliche Polarkoordinaten). Wie in Beispiel 1.2.12 sei H die Halbebene $S\times\mathbb{R} = \{(r,0,z)\colon r\geq 0, z\in\mathbb{R}\}$ im $\mathbb{R}^3$. Wir betrachten $\Omega = \mathbb{R}^3\setminus H$ und $\Omega' = (0,\infty)\times(0,\pi)\times(0,2\pi)$. Die Abbildung $\Phi\colon\Omega'\to\Omega$, definiert durch

$$\Phi(r,\vartheta,\varphi) = \begin{pmatrix} r\sin\vartheta\cos\varphi \\ r\sin\vartheta\sin\varphi \\ r\cos\vartheta \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

ist $\mathcal{C}^1$-invertierbar mit Funktionalmatrix

$$D\Phi(r,\vartheta,\varphi) = \begin{pmatrix} \sin\vartheta\cos\varphi & r\cos\vartheta\cos\varphi & -r\sin\vartheta\sin\varphi \\ \sin\vartheta\sin\varphi & r\cos\vartheta\sin\varphi & r\sin\vartheta\cos\varphi \\ \cos\vartheta & -r\sin\vartheta & 0 \end{pmatrix}.$$

Also ist

$$G(r,\vartheta,\varphi) = \mathrm{Diag}\big(1, r^2, r^2\sin^2\vartheta\big) \qquad \text{und} \qquad G^{-1}(r,\vartheta,\varphi) = \mathrm{Diag}\big(1, r^{-2}, (r\sin\vartheta)^{-2}\big).$$

und $\sqrt{\det(G(r,\vartheta,\varphi))} = r^2\sin\vartheta$. Für den Laplace-Operator in räumlichen Polarkoordinaten ergibt sich daher

$$\begin{aligned}\Delta^\Phi &= \frac{1}{r^2\sin\vartheta}\Big[\frac{\partial}{\partial r}\Big(r^2\sin\vartheta\frac{\partial}{\partial r}\Big) + \frac{\partial}{\partial\vartheta}\Big(\sin\vartheta\frac{\partial}{\partial\vartheta}\Big) + \frac{\partial}{\partial\varphi}\Big(\frac{1}{\sin\vartheta}\frac{\partial}{\partial\varphi}\Big)\Big] \\ &= \frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\Big[\frac{\partial^2}{\partial\vartheta^2} + \cot\vartheta\frac{\partial}{\partial\vartheta} + \frac{1}{\sin^2\vartheta}\frac{\partial^2}{\partial\varphi^2}\Big].\end{aligned}$$

$\diamond$

1.5 Integrale halbstetiger Funktionen

Wir wollen auch Integrale stetiger Funktionen $f\colon K \to \mathbb{R}$ über kompakte Mengen $K \subset \mathbb{R}^d$ berechnen, also Integrale der Form $\int_K f(x)\,\mathrm{d}x$. Solche Integrale werden wir natürlich als $\int_{\mathbb{R}^d} \widetilde{f}(x)\,\mathrm{d}x$ definieren, wobei $\widetilde{f}\colon \mathbb{R}^d \to \mathbb{R}$ die triviale Fortsetzung von f mit $\widetilde{f}(x) = 0$ für $x \in \mathbb{R}^d \backslash K$ ist. Insbesondere würden wir dann das Volumen von K berechnen können, wenn wir $f \equiv 1$ auf K wählen; dann ist $\widetilde{f} = \mathbb{1}_K\colon \mathbb{R}^d \to \{0,1\}$ die sogenannte *charakteristische Funktion* oder *Indikatorfunktion* von K:

$$\mathbb{1}_A(x) = \begin{cases} 1, & \text{falls } x \in A, \\ 0, & \text{falls } x \in \mathbb{R}^d \setminus A, \end{cases} \qquad A \subset \mathbb{R}^d.$$

Allerdings ist es für uns noch unklar, ob die Funktion $\widetilde{f}$ Riemann-integrierbar ist oder wie wir dieses Integral sonst auf natürliche Weise erklären wollen. Tatsächlich ist $\mathbb{1}_K$ für viele kompakte Mengen K *nicht* Riemann-integrierbar. Das bringt uns auf die Idee, $\widetilde{f}$ bzw. $\mathbb{1}_K$ durch eine Folge von Funktionen aus $\mathcal{C}_\mathrm{c}(\mathbb{R}^d)$ zu approximieren, etwa mit Hilfe eines gleichmäßigen Grenzwertes. Demonstrieren wir zunächst, dass für diese Idee es ausreichend ist, einen *monotonen* Grenzwert vorauszusetzen:

Satz 1.5.1 (Dini). *Sei $K \subset \mathbb{R}^d$ kompakt und $f, f_0, f_1, f_2, \ldots : K \to \mathbb{R}$ stetige Funktionen, sodass $f_n(x) \leq f_{n+1}(x)$ für jedes $x \in K$ und jedes $n \in \mathbb{N}$ und $\lim_{n\to\infty} f_n(x) = f(x)$. Dann ist die Konvergenz von f_n gegen f sogar gleichmäßig auf K.*

Beweis. Wir betrachten $g_n = f - f_n$, dann konvergieren die nichtnegativen Funktionen g_n fallend punktweise gegen Null. Sei $\varepsilon > 0$. Zu jedem $x \in K$ gibt es dann ein $N_x \in \mathbb{N}$ mit $g_n(x) < \varepsilon/2$ für alle $n \geq N_x$. Da g_{N_x} stetig in x ist, gibt es ein $\delta_x > 0$ mit $|g_{N_x}(y) - g_{N_x}(x)| < \varepsilon/2$ für alle $y \in K$ mit $\|y - x\| < \delta_x$. Insbesondere ist $g_n(y) < \varepsilon$ für alle $y \in K$ mit $\|y - x\| < \delta_x$ und alle $n \geq N_x$.

Da K kompakt ist, genügen endlich viele der Kugeln $U_x = \{y \in K\colon \|y - x\| < \delta_x\}$, um K zu überdecken, sagen wir etwa $K \subset U_{x_1} \cup \cdots \cup U_{x_k}$ für geeignete $x_1, \ldots, x_k \in K$. Wenn wir $N = \max\{N_{x_1}, \ldots, N_{x_k}\}$ setzen, dann haben wir für alle $n \geq N$, dass $g_n(y) < \varepsilon$ für jedes $y \in K$. Dies heißt aber, dass $\sup_{y\in K} |f(y) - f_n(y)| \leq \varepsilon$ für alle $n \geq N$. □

Also reicht monotone Konvergenz, um Integrale zu approximieren. Wir schreiben $f_n \uparrow f$, wenn $f, f_0, f_1, f_2, \ldots : \mathbb{R}^d \to \mathbb{R}$ Funktionen sind mit $f_n \leq f_{n+1}$ für jedes $n \in \mathbb{N}$ und $\lim_{n\to\infty} f_n = f$ punktweise.

Korollar 1.5.2. *Seien $f, f_0, f_1, f_2, \cdots \in \mathcal{C}_\mathrm{c}(\mathbb{R}^d)$ mit $f_n \uparrow f$. Dann gilt*

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} f_n(x)\,\mathrm{d}x = \int_{\mathbb{R}^d} f(x)\,\mathrm{d}x. \tag{1.5.1}$$

Beweis. Wegen der Monotonie ist der Träger von jedem f_n in der Vereinigung der Träger von f_0 und dem von f enthalten. Auf diesem Kompaktum können wir den Satz von Dini anwenden und erhalten gleichmäßige Konvergenz der f_n gegen f. Dann folgt die Konvergenz der Integrale aus Lemma 1.1.5 in Kombination mit Satz 1.1.4. □

Wir lassen uns durch Korollar 1.5.2 nun leiten und definieren für *jede* Funktion f, die sich als monotoner Grenzwert von Funktionen f_n aus $\mathcal{C}_\mathrm{c}(\mathbb{R}^d)$ schreiben lässt, das Integral von f durch (1.5.1), egal ob f selber in $\mathcal{C}_\mathrm{c}(\mathbb{R}^d)$ liegt oder nicht und sogar egal, ob f reellwertig ist oder die Werte ∞ oder $-\infty$ annimmt. Sei also $\mathcal{H}^\uparrow(\mathbb{R}^d)$ bzw. $\mathcal{H}^\downarrow(\mathbb{R}^d)$ die Menge aller Funktionen $f\colon \mathbb{R}^d \to \mathbb{R}\cup\{\infty\}$ bzw. $f\colon \mathbb{R}^d \to \mathbb{R}\cup\{-\infty\}$, die sich als monoton steigender bzw. monoton fallender Grenzwert von Funktionen $f_0, f_1, f_2, \dots \in \mathcal{C}_\mathrm{c}(\mathbb{R}^d)$ schreiben lassen. Dann definieren wir das Integral $\int_{\mathbb{R}^d} f(x)\,\mathrm{d}x$ durch (1.5.1). Dieser Grenzwert existiert in jedem Falle wegen Monotonie, aber wir behalten im Auge, dass er in $\mathbb{R}\cup\{\infty\}$ bzw. in $\mathbb{R}\cup\{-\infty\}$ liegt. Falls f schon in $\mathcal{C}_\mathrm{c}(\mathbb{R}^d)$ liegt, so impliziert Korollar 1.5.2, dass diese Definition nicht von der Wahl der approximierenden Funktionen f_n abhängt. Im allgemeinen Fall müssen wir das nachholen:

Lemma 1.5.3. *Sei $f \in \mathcal{H}^\uparrow(\mathbb{R}^d)$ und $(f_n)_{n\in\mathbb{N}}$ und $(g_n)_{n\in\mathbb{N}}$ zwei Folgen in $\mathcal{C}_\mathrm{c}(\mathbb{R}^d)$ mit $f_n \uparrow f$ und $g_n \uparrow f$. Dann gilt*

$$\lim_{n\to\infty}\int_{\mathbb{R}^d} f_n(x)\,\mathrm{d}x = \lim_{n\to\infty}\int_{\mathbb{R}^d} g_n(x)\,\mathrm{d}x.$$

Beweis. Aus Symmetriegründen reicht es, für jedes $k\in\mathbb{N}$ zu zeigen, dass

$$\int_{\mathbb{R}^d} f_k(x)\,\mathrm{d}x \le \lim_{n\to\infty}\int_{\mathbb{R}^d} g_n(x)\,\mathrm{d}x.$$

Dies folgt aber aus einer Betrachtung der Funktion $h_n = \min\{f_k, g_n\} \in \mathcal{C}_\mathrm{c}(\mathbb{R}^d)$, denn es gilt $h_n \uparrow f_k$ und $h_n \le g_n$, also nach Korollar 1.5.2 auch

$$\int_{\mathbb{R}^d} f_k(x)\,\mathrm{d}x = \lim_{n\to\infty}\int_{\mathbb{R}^d} h_n(x)\,\mathrm{d}x \le \lim_{n\to\infty}\int_{\mathbb{R}^d} g_n(x)\,\mathrm{d}x.$$

□

Um zu sehen, dass diese erweiterte Definition des Integralbegriffes sinnvoll ist, müssen wir prüfen, ob für eine Funktion, die sowohl in $\mathcal{H}^\uparrow(\mathbb{R}^d)$ als auch in $\mathcal{H}^\downarrow(\mathbb{R}^d)$ liegt, die beiden Definitionen des Integrals $\int_{\mathbb{R}^d} f(x)\,\mathrm{d}x$ miteinander übereinstimmen, d. h. ob die Grenzwerte bei fallender und bei steigender Approximation mit Funktionen aus $\mathcal{C}_\mathrm{c}(\mathbb{R}^d)$ miteinander übereinstimmen. Dies wird in Korollar 1.5.9 nachgeholt werden, wo wir zeigen, dass $\mathcal{H}^\uparrow(\mathbb{R}^d)\cap\mathcal{H}^\downarrow(\mathbb{R}^d) = \mathcal{C}_\mathrm{c}(\mathbb{R}^d)$ gilt.

Nun interessiert uns, ob wir mit dieser Erweiterung der integrierbaren Funktionen auch die oben genannten Funktionen erreicht haben, d. h. ob wir nun Integrale der Form $\int_K f(x)\,\mathrm{d}x$ behandeln können.

Lemma 1.5.4. *(i) Sei $K\subset\mathbb{R}^d$ kompakt und $f\colon K\to[0,\infty)$ stetig. Mit $\widetilde{f}\colon \mathbb{R}^d\to[0,\infty)$ bezeichnen wir die triviale Fortsetzung von f, d. h. $\widetilde{f}(x)=0$ für $x\in\mathbb{R}^d\setminus K$. Dann liegt $\widetilde{f}$ in $\mathcal{H}^\downarrow(\mathbb{R}^d)$. Insbesondere ist die folgende Definition sinnvoll:*

$$\int_K f(x)\,\mathrm{d}x = \int_{\mathbb{R}^d}\widetilde{f}(x)\,\mathrm{d}x. \tag{1.5.2}$$

(ii) Sei $U\subset\mathbb{R}^d$ offen und $f\colon U\to[0,\infty)$ stetig. Mit $\widetilde{f}\colon\mathbb{R}^d\to[0,\infty)$ bezeichnen wir die triviale Fortsetzung von f, d. h. $\widetilde{f}(x)=0$ für $x\in\mathbb{R}^d\setminus U$. Dann liegt $\widetilde{f}$ in $\mathcal{H}^\uparrow(\mathbb{R}^d)$. Insbesondere ist die folgende Definition sinnvoll:

$$\int_U f(x)\,\mathrm{d}x = \int_{\mathbb{R}^d}\widetilde{f}(x)\,\mathrm{d}x. \tag{1.5.3}$$

Beweis. Übungsaufgabe. Man benutze z. B. die Teilung der Eins aus dem Beweis von Lemma 1.3.1. Alternativ kann man auch Lemma 1.5.8 (siehe weiter unten) benutzen. $\square$

Im Abschnitt 1.6 werden wir Lemma 1.5.4 benutzen, um einige Integrale über Kompakta zu berechnen. Jetzt geben wir eine Charakterisierung der Funktionen in $\mathcal{H}^\uparrow(\mathbb{R}^d)$ bzw. $\mathcal{H}^\downarrow(\mathbb{R}^d)$. Es stellt sich heraus, dass dies im Wesentlichen genau die halbstetigen Funktionen sind.

Definition 1.5.5 (halbstetige Funktion). *(i) Eine Funktion $f\colon \mathbb{R}^d \to \mathbb{R} \cup \{\infty\}$ heißt in $x \in \mathbb{R}^d$* halbstetig von unten *oder* unterhalbstetig, *falls zu jedem $c \in \mathbb{R}$ mit $c < f(x)$ ein $\delta > 0$ existiert mit $c < f(y)$ für jedes $y \in \mathbb{R}^d$ mit $\|y - x\| < \delta$.*

(ii) Eine Funktion $f\colon \mathbb{R}^d \to \mathbb{R} \cup \{-\infty\}$ heißt in $x \in \mathbb{R}^d$ halbstetig von oben *oder* oberhalbstetig, *falls $-f$ unterhalbstetig ist.*

Natürlich nennen wir eine Funktion halbstetig (von unten bzw. oben), wenn sie in jedem Punkt dies ist. Man sieht leicht, dass eine Funktion, die in einem Punkt gleichzeitig von oben und von unten halbstetig ist, dort stetig ist. Wir sammeln ein paar Charakterisierungen der Halbstetigkeit, die Beweise sind Übungsaufgaben:

Lemma 1.5.6 (Charakterisierungen der Halbstetigkeit). *(i) Die Indikatorfunktion $\mathbb{1}_A$ einer Menge $A \subset \mathbb{R}^d$ ist genau dann von unten halbstetig, wenn A offen ist, und sie ist genau dann von oben halbstetig, wenn A abgeschlossen ist.*

(ii) Eine Funktion $f\colon \mathbb{R}^d \to \mathbb{R} \cup \{\infty\}$ ist genau dann von unten halbstetig in $x \in \mathbb{R}^d$, wenn $\liminf_{y \to x} f(y) \geq f(x)$ gilt.

(iii) Eine Funktion $f\colon \mathbb{R}^d \to \mathbb{R} \cup \{\infty\}$ ist genau dann von unten halbstetig in $x \in \mathbb{R}^d$, wenn gilt:

$$\lim_{\varepsilon \downarrow 0} \inf_{y \in B_\varepsilon(x)} f(y) = f(x),$$

wobei $B_\varepsilon(x) = \{y \in \mathbb{R}^d \colon \|y - x\| < \varepsilon\}$ die offene ε-Kugel um x ist.

(iv) Eine Funktion $f\colon \mathbb{R}^d \to \mathbb{R} \cup \{\infty\}$ ist genau dann von unten halbstetig, wenn die Niveaumengen $\{f \leq s\} = \{x \in \mathbb{R}^d \colon f(x) \leq s\}$ für jedes $s \in \mathbb{R}$ abgeschlossen sind.

Bemerkung 1.5.7. Jede oberhalbstetige Funktion $f\colon \mathbb{R}^d \to \mathbb{R} \cup \{-\infty\}$ nimmt auf jeder kompakten Menge $K \subset \mathbb{R}^d$ ihr Maximum $\max_{x \in K} f(x)$ an (Übungsaufgabe). $\diamond$

Damit können wir nun $\mathcal{H}^\uparrow(\mathbb{R}^d)$ charakterisieren. Wir schreiben kurz $\{f \geq a\}$ für die Menge $\{x \in \mathbb{R}^d \colon f(x) \geq a\}$, analog für andere Ungleichungen oder Gleichungen.

Lemma 1.5.8 (Charakterisierung von $\mathcal{H}^\uparrow(\mathbb{R}^d)$).

$$\mathcal{H}^\uparrow(\mathbb{R}^d) = \big\{f\colon \mathbb{R}^d \to \mathbb{R} \cup \{\infty\} \colon f \text{ ist unterhalbstetig und } \{f < 0\} \text{ ist beschränkt}\big\}. \tag{1.5.4}$$

Beweis. '$\subset$': Man sieht leicht, dass das punktweise Supremum jeder Familie von unterhalbstetigen Funktionen ebenfalls von unten halbstetig ist. Jedes $f \in \mathcal{H}^\uparrow(\mathbb{R}^d)$ ist insbesondere das Supremum von stetigen Funktionen, also daher unterhalbstetig. Außerdem gibt es ein $f_0 \in \mathcal{C}_\mathrm{c}(\mathbb{R}^d)$ mit $f_0 \leq f$, woraus auch folgt, dass f höchstens in dem kompakten Träger von f_0 negativ sein kann.

'$\supset$': Sei $f\colon \mathbb{R}^d \to \mathbb{R}\cup\{\infty\}$ eine unterhalbstetige Funktion, so dass $\{f<0\}$ eine beschränkte Menge ist. Dann ist $K = \overline{\{f<0\}}$ kompakt. Wir konstruieren nun eine Folge in $\mathcal{C}_c(\mathbb{R}^d)$, die monoton steigend gegen f konvergiert.

Nach Bemerkung 1.5.7 ist f von unten beschränkt, d. h. es gibt ein $M>0$ mit $f(x)\geq -M$ für alle $x\in\mathbb{R}^d$. Mit J bezeichnen wir die Menge aller rationalen Tupel (ε,a,c) in $(0,\infty)\times\mathbb{R}^d\times[-M,\infty)\cap\mathbb{Q}^{d+2}$ mit $f(x)\geq c$ für alle x in $B_\varepsilon(a)$, der ε-Umgebung von a. Dann ist J abzählbar. Zu jedem $j=(\varepsilon,a,c)\in J$ gibt es eine Funktion $g_j\in\mathcal{C}_c(\mathbb{R}^d)$ mit den Eigenschaften

$$\begin{array}{rlll} (i) & g_j(x)=c & \text{für alle} & x\in B_{\varepsilon/2}(a),\\ (ii) & g_j(x)\leq c & \text{für alle} & x\in B_\varepsilon(a),\\ (iii) & g_j(x)=-M & \text{für alle} & x\in K\setminus B_\varepsilon(a),\\ (iv) & g_j(x)\leq 0 & \text{für alle} & x\in\mathbb{R}^d\setminus(K\cup B_\varepsilon(a)). \end{array}$$

Aus der Konstruktion ersieht man, dass $f\geq g_j$ für jedes $j\in J$ gilt. Mit einigem technischen Aufwand kann man auch noch zeigen, dass $f=\sup\{g_j\colon j\in J\}$ gilt. (Anleitung: Für gegebenes $x\in\mathbb{R}^d$ wähle man eine geeignete Folge $(\varepsilon_n,a_n,c_n)_{n\in\mathbb{N}}$ in J mit $\varepsilon_n\downarrow 0$, $a_n\to x$ und $c_n\uparrow f(x)$.) Da J abzählbar ist, kann man $J=\{g_{j_0},g_{j_1},\dots\}$ aufzählen. Dann liegt $f_n=\max\{g_{j_0},\dots,g_{j_n}\}$ in $\mathcal{C}_c(\mathbb{R}^d)$, und wir haben $f_n\uparrow f$ für $n\to\infty$. Das heißt, dass $f\in\mathcal{H}^\uparrow(\mathbb{R}^d)$. □

Korollar 1.5.9.

$$\mathcal{H}^\uparrow(\mathbb{R}^d)\cap\mathcal{H}^\downarrow(\mathbb{R}^d)=\mathcal{C}_c(\mathbb{R}^d).$$

Beweis. Der Übergang von f zu $-f$ in Lemma 1.5.8 zeigt, dass

$$\mathcal{H}^\downarrow(\mathbb{R}^d)=\big\{f\colon\mathbb{R}^d\to\mathbb{R}\cup\{-\infty\}\colon f \text{ ist oberhalbstetig und } \{f>0\} \text{ ist beschränkt}\big\}.$$

Daraus folgt die Aussage leicht. □

Bemerkung 1.5.10. Der Sinn dieses Abschnittes über Integrale halbstetiger Funktionen ist, möglichst schnell interessante Integrale behandeln zu können, siehe Abschnitt 1.6. Mit Hilfe eines völlig anderen, viel allgemeineren Zugangs, der allgemeinen Maßtheorie, werden wir später viel größere Funktionenklassen integrieren lernen, aber das nimmt einige Zeit in Anspruch. Ein alternativer Zugang zu dieser Funktionenklasse (siehe [Fo3]) läuft über einen weiteren Approximationsschritt mit Grenzwerten von Funktionenfolgen aus $\mathcal{H}^\uparrow(\mathbb{R}^d)$ bzw. $\mathcal{H}^\downarrow(\mathbb{R}^d)$, aber diesen Weg verfolgen wir in diesem Skript nicht. ◇

Da wir also später viel größere Funktionenklassen behandeln werden, werden wir uns jetzt bei der Herleitung der Integrationsregeln für Funktionen aus $\mathcal{H}^\uparrow(\mathbb{R}^d)$ bzw. $\mathcal{H}^\downarrow(\mathbb{R}^d)$ kurz fassen.

Satz 1.5.11. *Für alle $f,g\in\mathcal{H}^\uparrow(\mathbb{R}^d)$ und $\lambda\in[0,\infty)$ liegt $f+\lambda g$ in $\mathcal{H}^\uparrow(\mathbb{R}^d)$, und es gilt*

$$\int (f(x)+\lambda g(x))\,\mathrm{d}x=\int f(x)\,\mathrm{d}x+\lambda\int g(x)\,\mathrm{d}x.$$

Ferner gilt $\int f(x)\,\mathrm{d}x\leq\int g(x)\,\mathrm{d}x$, falls $f\leq g$.

Der Beweis ist eine einfache Übungsaufgabe. Man beachte, dass alle in Satz 1.5.11 auftretenden Integrale auch gleich ∞ sein können. Dabei benutzen wir die Konventionen $a+\infty=\infty+a=\infty$ für jedes $a\in\mathbb{R}\cup\{\infty\}$ sowie $\lambda\cdot\infty=\infty$ für jedes $\lambda>0$ und $0\cdot\infty=0$. Eine analoge

Version von Satz 1.5.11 für $\mathcal{H}^\uparrow(\mathbb{R}^d)$ gibt es selbstverständlich auch, wir formulieren sie hier aber nicht.

Die Translationsinvarianz des Integrals (siehe Satz 1.1.4) und den Transformationssatz für lineare Funktionen (Satz 1.2.5) kann man leicht auf Integrale über Funktionen aus $\mathcal{H}^\uparrow(\mathbb{R}^d)$ ausdehnen:

Satz 1.5.12. *Seien $A \in \mathrm{GL}(d,\mathbb{R})$ und $b \in \mathbb{R}^d$, dann liegt für jede Funktion $f \in \mathcal{H}^\uparrow(\mathbb{R}^d) \cup \mathcal{H}^\downarrow(\mathbb{R}^d)$ die Funktion $x \mapsto f(Ax+b)$ ebenfalls in $\mathcal{H}^\uparrow(\mathbb{R}^d) \cup \mathcal{H}^\downarrow(\mathbb{R}^d)$, und es gilt*

$$\int_{\mathbb{R}^d} f(Ax+b)\,\mathrm{d}x = \frac{1}{|\det(A)|}\int_{\mathbb{R}^d} f(x)\,\mathrm{d}x.$$

Ferner benötigen wir noch eine Version des Satzes 1.1.1 für Integrale über Funktionen aus $\mathcal{H}^\uparrow(\mathbb{R}^d)$:

Satz 1.5.13 (Satz von Fubini für Funktionen aus $\mathcal{H}^\uparrow(\mathbb{R}^d)$). *Sei $f \in \mathcal{H}^\uparrow(\mathbb{R}^d)$, und sei $k \in \{1,\dots,d-1\}$. Dann liegt für jedes $(x_{k+1},\dots,x_d) \in \mathbb{R}^{d-k}$ die Funktion $(x_1,\dots,x_k) \mapsto f(x)$ in $\mathcal{H}^\uparrow(\mathbb{R}^k)$, also ist das Integral*

$$F(x_{k+1},\dots,x_d) = \int_{\mathbb{R}^k} f(x)\,\mathrm{d}x_1\dots\mathrm{d}x_k, \qquad x=(x_1,\dots,x_d),$$

wohldefiniert. Ferner liegt F in $\mathcal{H}^\uparrow(\mathbb{R}^{d-k})$, und es gilt

$$\int_{\mathbb{R}^{d-k}} F(x_{k+1},\dots,x_d)\,\mathrm{d}x_{k+1}\dots\mathrm{d}x_d = \int_{\mathbb{R}^d} f(x)\,\mathrm{d}x.$$

Der Beweis besteht aus einer Anwendung des Satzes 1.1.1 auf approximierende Funktionen aus $\mathcal{C}_\mathrm{c}(\mathbb{R}^d)$; wir verzichten auf eine Ausführung der Details.

1.6 Berechnung einiger Volumina

Nun haben wir das theoretische Rüstzeug, um Integrale der Form $\int_K f(x)\,\mathrm{d}x$ zu berechnen, wobei $K \subset \mathbb{R}^d$ kompakt und $f\colon K \to \mathbb{R}$ stetig ist, siehe Lemma 1.5.4. Wir benutzen dies, um in diesem Abschnitt Volumina einiger Körper zu berechnen. Dabei definieren wir das *Volumen* der kompakten Menge $K \subset \mathbb{R}^d$ durch

$$\mathrm{Vol}(K) = \mathrm{Vol}_d(K) = \int_K 1\,\mathrm{d}x = \int_{\mathbb{R}^d} \mathbb{1}_K(x)\,\mathrm{d}x. \tag{1.6.1}$$

Nach Lemma 1.5.4 ist diese Definition sinnvoll. Das zweidimensionale Volumen Vol_2 wird auch *Fläche* genannt und das eindimensionale Vol_1 *Länge*. Zunächst berechnen wir die Volumina von Kreuzprodukten:

Lemma 1.6.1. *Sei $k \in \{1,\dots,d-1\}$, und seien $K_1 \subset \mathbb{R}^k$ und $K_2 \subset \mathbb{R}^{d-k}$ kompakt. Dann gilt*

$$\mathrm{Vol}_d(K_1 \times K_2) = \mathrm{Vol}_k(K_1)\cdot \mathrm{Vol}_{d-k}(K_2).$$

Beweis. Wir erinnern uns, dass $K_1 \times K_2$ ebenfalls kompakt ist, also ist $\mathrm{Vol}_d(K_1 \times K_2)$ wohldefiniert. Offensichtlich ist

$$\mathbb{1}_{K_1\times K_2}(x) = \mathbb{1}_{K_1}(x_1,\dots,x_k)\mathbb{1}_{K_2}(x_{k+1},\dots,x_d), \qquad x=(x_1,\dots,x_d).$$

Also folgt die Aussage leicht aus dem Satz 1.5.13 von Fubini. □

Beispiel 1.6.2 (Volumen eines Quaders). Es ist klar, dass $\mathrm{Vol}_1([a,b]) = b - a$ für alle $-\infty < a < b < \infty$. Also folgt aus Lemma 1.6.1:

$$\mathrm{Vol}_d\Big(\prod_{i=1}^d [a_i, b_i]\Big) = \prod_{i=1}^d (b_i - a_i), \qquad -\infty < a_i < b_i < \infty \,\forall i = 1, \dots, d.$$

Also ist das Volumen des Quaders $\prod_{i=1}^d [a_i, b_i]$ genau sein elementargeometrischer Inhalt. ◇

Beispiel 1.6.3 (Volumen eines Zylinders). Das Volumen eines Zylinders $Z = B \times [0,h]$ mit kompakter Basismenge $B \subset \mathbb{R}^{d-1}$ und Höhe $h > 0$ ist nach Lemma 1.6.1 gegeben als $\mathrm{Vol}_d(Z) = h \cdot \mathrm{Vol}_{d-1}(B)$. ◇

Wir untersuchen nun, wie sich Volumina unter linearen Abbildungen transformieren.

Lemma 1.6.4. *Seien $A \in \mathbb{R}^{d\times d}$ eine Matrix und $b \in \mathbb{R}^d$ ein Vektor. Ferner sei $K \subset \mathbb{R}^d$ kompakt. Dann ist das Volumen der Menge $AK + b = \{Ax + b \colon x \in K\}$ gegeben durch*

$$\mathrm{Vol}_d(AK + b) = |\det(A)|\mathrm{Vol}_d(K).$$

Insbesondere ändert sich das Volumen unter orthogonalen Transformationen nicht.

Beweis. Wenn A regulär ist, folgt die Aussage sofort aus Satz 1.5.12. Wenn A nicht regulär ist (also $\det(A) = 0$), so ist $AK+b$ in einer Hyperebene enthalten, die durch eine orthogonale Transformation (die das Volumen nicht ändert) in die Hyperebene $\{x \in \mathbb{R}^d \colon x_n = 0\}$ überführt werden kann. Also ist $AK + b$ in einem Zylinder der Höhe Null enthalten, hat also nach Beispiel 1.6.3 das Volumen Null. □

Beispiel 1.6.5 (Volumen eines Parallelepipeds). Wir erinnern an das Volumen eines Parallelepipeds, das wir in Beispiel 1.2.7 ausrechneten. ◇

Bemerkung 1.6.6 (Verhalten unter Streckungen). Für jedes Kompaktum $K \subset \mathbb{R}^d$ und jedes $r > 0$ gilt $\mathrm{Vol}_d(rK) = r^d \mathrm{Vol}_d(K)$, denn $rK = \{rx \colon x \in K\}$ ist das Bild von K unter der Multiplikation mit rE, wobei E die $d \times d$-Einheitsmatrix ist. ◇

Satz 1.6.7 (Cavalieri'sches Prinzip). *Sei $K \subset \mathbb{R}^d$ kompakt, und für $t \in \mathbb{R}$ sei*

$$K_t = \{(x_1, \dots, x_{d-1}) \in \mathbb{R}^{d-1} \colon (x_1, \dots, x_{d-1}, t) \in K\}$$

der t-Schnitt von K. Dann gilt

$$\mathrm{Vol}_d(K) = \int_{\mathbb{R}} \mathrm{Vol}_{d-1}(K_t)\,\mathrm{d}t.$$

Beweis. Dies folgt direkt aus dem Satz 1.5.13 von Fubini, da

$$\mathbb{1}_{K_t}(x_1, \dots, x_{d-1}) = \mathbb{1}_K(x_1, \dots, x_{d-1}, t).$$

□

Auf ähnliche Weise zeigt man das Folgende.

Lemma 1.6.8. *Sei $A \subset \mathbb{R}^d$ kompakt und $f\colon A \to [0,\infty)$ stetig. Dann hat die kompakte Menge*

$$K = \{(x,y) \in A \times \mathbb{R} \colon 0 \le y \le f(x)\}$$

das Volumen $\mathrm{Vol}_{d+1}(K) = \int_A f(x)\,\mathrm{d}x$.

Beweis. Übungsaufgabe. □

Beispiel 1.6.9 (Volumen eines Kegels). Sei $B \subset \mathbb{R}^{d-1}$ kompakt und $h > 0$. Wir betrachten den *Kegel* mit Basis B und Höhe h:

$$C_h(B) = \big\{((1-\lambda)y, \lambda h) \in \mathbb{R}^d \colon y \in B, \lambda \in [0,1]\big\}.$$

Die Schnitte sind

$$C_h(B)_t = \begin{cases} (1 - \frac{t}{h})B, & \text{falls } t \in [0,h], \\ \emptyset, & \text{falls } t < 0 \text{ oder } t > h. \end{cases}$$

Dann folgt aus dem Cavalieri'schen Prinzip und Beispiel 1.6.6:

$$\mathrm{Vol}_d(C_h(B)) = \int_0^h \mathrm{Vol}_{d-1}\Big(\Big(1 - \frac{t}{h}\Big)B\Big)\,\mathrm{d}t = \mathrm{Vol}_{d-1}(B) \int_0^h \Big(1 - \frac{t}{h}\Big)^{d-1}\,\mathrm{d}t = \frac{h}{d}\mathrm{Vol}_{d-1}(B).$$

Das Volumen des Kegels ist also der d-te Teil des Volumens des Zylinders mit der selben Basis und der selben Höhe. ◇

Beispiel 1.6.10 (Volumen eines Simplex). Unter dem von den Vektoren $a_1, \dots, a_d \in \mathbb{R}^d$ aufgespannten *Simplex* versteht man die Menge

$$S(a_1, \dots, a_d) = \Big\{ \sum_{i=1}^d \lambda_i a_i \colon \lambda_1, \dots, \lambda_d \in [0,\infty), \sum_{i=1}^d \lambda_i \le 1 \Big\}.$$

Es sei A die $d \times d$-Matrix, deren Spalten die $a_1, \dots, a_d$ sind. Wir werden nun die Formel

$$\mathrm{Vol}_d\big(S(a_1, \dots, a_d)\big) = \frac{1}{d!} |\det(A)|$$

zeigen. Man beachte, dass also das Volumen eines Simplex gleich dem $d!$-ten Teil des Volumens des zugehörigen Parallelotops ist, siehe Beispiel 1.2.7.

Da $S(a_1, \dots, a_d) = A(S(\mathrm{e}_1, \dots, \mathrm{e}_d))$ ist, wobei $\mathrm{e}_1, \dots, \mathrm{e}_d$ die kanonische Standardbasis des $\mathbb{R}^d$ ist, genügt es nach Lemma 1.6.4 zu zeigen, dass $\mathrm{Vol}_d(S(\mathrm{e}_1, \dots, \mathrm{e}_d)) = \frac{1}{d!}$ ist. Wir benutzen dafür eine Induktion über d und sehen, dass $S(\mathrm{e}_1) = [0,1]$, also stimmt die Behauptung für $d = 1$. Für allgemeines d mache man sich klar, dass

$$S(\mathrm{e}_1, \dots, \mathrm{e}_d) = C_1(S(\mathrm{e}_1, \dots, \mathrm{e}_{d-1}))$$

der Kegel mit Höhe 1 über der Basis $S(\mathrm{e}_1, \dots, \mathrm{e}_{d-1}) \subset \mathbb{R}^{d-1}$ ist, siehe die Notation aus Beispiel 1.6.9. Also erhalten wir aus der Formel in diesem Beispiel und der Induktionsvoraussetzung:

$$\mathrm{Vol}_d(S(\mathrm{e}_1, \dots, \mathrm{e}_d)) = \frac{1}{d}\mathrm{Vol}_{d-1}(S(\mathrm{e}_1, \dots, \mathrm{e}_{d-1})) = \frac{1}{d}\frac{1}{(d-1)!} = \frac{1}{d!}.$$

◇

Beispiel 1.6.11 (Volumen einer Kugel). Es sei $K_d(r) = \{x \in \mathbb{R}^d \colon \|x\| \leq r\}$ die d-dimensionale Kugel mit Radius $r > 0$. Wegen $\mathrm{Vol}_d(K_d(r)) = r^d \mathrm{Vol}_d(K_d(1))$ genügt es, das Volumen $\tau_d = \mathrm{Vol}_d(K_d)$ der Einheitskugel $K_d = K_d(1)$ zu berechnen. Wir benutzen wieder eine Induktion über d und sehen zunächst, dass $\tau_1 = 2$, denn $K_1 = [-1, 1]$. Um mit Hilfe des Cavalieri'schen Prinzips (Satz 1.6.7) das Volumen von K_d auf das von K_{d-1} zurück zu führen, schauen wir uns die Schnitte von K_d an:

$$(K_d)_t = \begin{cases} K_{d-1}\big(\sqrt{1-t^2}\big), & \text{falls } t \in [-1,1], \\ \emptyset, & \text{falls } |t| > 1. \end{cases}$$

Also haben wir

$$\tau_d = \int_{\mathbb{R}} \mathrm{Vol}_{d-1}((K_d)_t)\,\mathrm{d}t = \int_{-1}^{1} \mathrm{Vol}_{d-1}\big(K_{d-1}\big(\sqrt{1-t^2}\big)\big)\,\mathrm{d}t = \tau_{d-1} \int_{-1}^{1} \sqrt{1-t^2}^{\,d-1}\,\mathrm{d}t.$$

Es bleibt das letzte Integral zu berechnen. Mit Hilfe einer Substitution $t = \cos(x)$ sieht man, dass

$$c_d := \int_{-1}^{1} \sqrt{1-t^2}^{\,d-1}\,\mathrm{d}t = \int_0^{\pi} (\sin x)^d\,\mathrm{d}x = 2\int_0^{\pi/2} (\sin x)^d\,\mathrm{d}x.$$

Dies wurde in der Vorlesung *Analysis A* berechnet, siehe [Fo1, Beispiel (19.22)]:

$$c_{2d} = \pi \prod_{n=1}^{d} \frac{2n-1}{2n} \qquad \text{und} \qquad c_{2d+1} = 2\prod_{n=1}^{d} \frac{2n}{2n+1}.$$

Da also $c_d c_{d-1} = \frac{2\pi}{d}$ für jedes $d \in \mathbb{N}$ gilt, haben wir die Zweischritt-Rekursionsformel $\tau_d = \tau_{d-1} c_d = \tau_{d-2} c_d c_{d-1} = \frac{2\pi}{d}\tau_{d-2}$. Daraus kann man explizit die Werte

$$\tau_{2d} = \frac{\pi^d}{d!} \qquad \text{und} \qquad \tau_{2d+1} = \frac{2^{d+1}}{1 \cdot 3 \cdot \dots \cdot (2d+1)}\,\pi^d$$

erhalten. Eine einheitliche Formel für gerade und ungerade Dimensionen erhalten wir mit Hilfe der Gammafunktion, die nämlich für jedes $k \in \mathbb{N}$ erfüllt:

$$\Gamma(k+1) = k! \qquad \text{und} \qquad \Gamma(k+\tfrac{3}{2}) = \Gamma(\tfrac{1}{2}) \prod_{n=0}^{k} \frac{2n+1}{2} = \sqrt{\pi} \prod_{n=0}^{k} \frac{2n+1}{2}.$$

Daher ist

$$\boxed{\tau_d = \frac{\pi^{d/2}}{\Gamma(1+\frac{d}{2})}, \qquad d \geq 1.} \tag{1.6.2}$$

◇

Beispiel 1.6.12 (Volumen eines Ellipsoids). Ein *Ellipsoid* mit Halbachsen $\alpha_1, \dots, \alpha_d \in (0, \infty)$ ist die Menge

$$E(\alpha_1, \dots, \alpha_d) = \Big\{(x_1, \dots, x_d) \in \mathbb{R}^d \colon \sum_{i=1}^{d} \Big(\frac{x_i}{\alpha_i}\Big)^2 \leq 1\Big\}.$$

Also ist $E(\alpha_1, \dots, \alpha_d) = A(K_d)$ das Bild der Einheitskugel K_d unter der linearen Abbildung $A = \mathrm{Diag}(\alpha_1, \dots, \alpha_d)$. Nach Bemerkung 1.6.6 gilt $\mathrm{Vol}_d(E(\alpha_1, \dots, \alpha_d)) = \tau_d \prod_{i=1}^{d} \alpha_i$, siehe (1.6.2). ◇

Lemma 1.6.13 (Volumen von Rotationskörpern). *Sei* $-\infty < a < b < \infty$ *und* $f\colon [a,b] \to [0,\infty)$ *stetig. Dann hat der dreidimensionale Rotationskörper*

$$K = \{(x,y,z) \in [a,b] \times \mathbb{R}^2 \colon y^2 + z^2 \leq f(x)^2\}$$

das Volumen $\mathrm{Vol}_3(K) = \pi \int_a^b f(x)^2 \,\mathrm{d}x$.

Beweis. Übungsaufgabe. □

Kapitel 2

Einführung in die Maßtheorie

Sind wir in Kapitel 1 von der Integration von Funktionen ausgegangen, so werden wir in diesem Kapitel das Messen von Mengen zu Grunde legen. Wir haben schon in Abschnitt 1.5 das Maß (dort 'Volumen' genannt) von beliebigen kompakten Mengen definiert und bestimmt; jetzt aber wollen wir noch viel allgemeinere Mengen messen. Mehr noch, wir werden die gesamte Theorie auf dem Begriff des Maßes einer Menge aufbauen.

In Abschnitt 2.1 werden wir zunächst an einem konkreten Beispiel die Vorgehensweise demonstrieren, dem zweidimensionalen Lebesgue-Maß. Danach geben wir eine Einführung in die allgemeine Maßtheorie. In Abschnitt 2.2 erläutern wir die fundamentalen Begriffe und Konzepte wie σ-Algebren, Messbarkeit und Maße und deren Eigenschaften. Das wichtige Problem der Konstruktion von Maßen mit gegebenen Eigenschaften wird in Abschnitt 2.3 diskutiert; hierbei werden Dynkin-Systeme, Algebren und Prämaße eine wichtige Rolle spielen, und das Hauptergebnis wird der Existenz- und Eindeutigkeitssatz von Carathéodory sein. In Abschnitt 2.4 werden wir Maße mit Hilfe von Verteilungsfunktionen charakterisieren, wobei ein alternativer Existenzbeweis des Lebesgue-Maßes abfallen wird.

2.1 Motivation und Einführung: das Lebesgue-Maß auf $\mathbb{R}^2$

In diesem Abschnitt werden wir das Problem des Messens von Teilmengen am Beispiel des $\mathbb{R}^2$ genauer betrachten. Die folgenden Überlegungen können leicht auf den $\mathbb{R}^d$ verallgemeinert werden; sie werden uns in der allgemeinen Theorie als Leitfaden dienen. In diesem Abschnitt lassen wir uns von Abschnitt 2 des Skriptes [G05] inspirieren.

Es sei $\mathcal{P}(\mathbb{R}^2)$ die Potenzmenge von $\mathbb{R}^2$. Das Ziel ist, auf einer möglichst großen Teilmenge $\mathcal{F}$ von $\mathcal{P}(\mathbb{R}^2)$ eine Abbildung $\lambda\colon \mathcal{F} \to [0,\infty]$ zu definieren, so dass für jedes $A \in \mathcal{F}$ die Zahl $\lambda(A)$ in möglichst natürlichem Sinne das Maß von A genannt werden kann. (Eine solche Menge $\mathcal{F}$ nennt man meist ein *Mengensystem.*) Insbesondere sollte für Rechtecke A die Zahl $\lambda(A)$ auch immer der elementargeometrische Inhalt sein, d. h., wir fordern, dass

$$\lambda\big([a,b] \times [c,d]\big) = (b-a)(d-c), \qquad a<b, c<d. \tag{2.1.1}$$

Damit können wir also λ schon einmal auf dem Mengensystem $\mathcal{R}$ aller Mengen der Form

$$[a,b] \times [c,d], \qquad (a,b) \times [c,d], \qquad (a,b] \times (c,d), \qquad [a,b) \times (c,d) \qquad \text{usw.} \tag{2.1.2}$$

sinnvoll definieren, wobei $a \leq b$ und $c \leq d$ sind. Man beachte, dass $\mathcal{R}$ alle offenen, abgeschlossenen oder halboffenen beschränkten achsenparallelen Rechtecke enthält, insbesondere auch die leere Menge, Punkte und achsenparallele Strecken. Im Folgenden nennen wir jede Menge aus $\mathcal{R}$ kurz ein Rechteck. Für jede Menge $R \in \mathcal{R}$ kann man durch ein offensichtliche Erweiterung von (2.1.1) den Wert von $\lambda(R)$ definieren: Jedes Rechteck mit Seiten von a bis b und von c bis d hat das Maß $(b-a)(d-c)$, egal ob es offen, abgeschlossen oder halboffen ist. Insbesondere haben wir $\lambda(\emptyset) = 0$. Man sieht auch leicht, dass λ *additiv* ist: Falls sich ein Rechteck $R \in \mathcal{R}$ als eine paarweise disjunkte Vereinigung von Rechtecken $R_1, \dots, R_n \in \mathcal{R}$ darstellen lässt (also $R_i \cap R_j = \emptyset$ für $i \neq j$ und $R = \bigcup_{i=1}^n R_i$), dann gilt $\lambda(R) = \sum_{i=1}^n \lambda(R_i)$.

Nun wollen wir die Abbildung λ fortsetzen auf die Menge $\mathcal{E}$ aller *elementaren* Mengen, d. h. die Menge aller endlichen Vereinigungen von paarweise disjunkten Mengen aus $\mathcal{R}$. Dieses Mengensystem ist nicht mehr so leicht zu überblicken, hat aber immer noch ein paar angenehme Abgeschlossenheitseigenschaften bezüglich der elementaren Operationen:

Lemma 2.1.1. *Für je zwei Mengen $A, B \in \mathcal{E}$ liegen auch $A \cup B$, $A \cap B$, $A \setminus B$ und $A \triangle B$*[1] *in $\mathcal{E}$.*

Beweis. Wenn $A = \bigcup_{i=1}^n R_i$ und $B = \bigcup_{j=1}^m Q_j$ zwei elementare Mengen sind mit paarweise disjunkten Rechtecken $R_1, \dots, R_n \in \mathcal{R}$ bzw. paarweise disjunkten Rechtecken $Q_1, \dots, Q_m \in \mathcal{R}$, so ist $A \cap B = \bigcup_{i=1}^n \bigcup_{j=1}^m (R_i \cap Q_j)$ eine Vereinigung paarweise disjunkter Rechtecke in $\mathcal{R}$, gehört also zu $\mathcal{E}$. Ferner sieht man, dass die Differenzmenge eines Rechteckes R, das A umfasst, und A immer eine elementare Menge ist, denn $R \setminus A = \bigcap_{i=1}^n (R \setminus R_i)$, und $R \setminus R_i$ ist leicht als elementar zu erkennen. Also ist $A \cup B = R \setminus [(R \setminus A) \cap (R \setminus B)]$ elementar, wenn $R \in \mathcal{R}$ sowohl A als auch B umfasst. Die Beweise, dass $A \setminus B$ und $A \triangle B$ elementar sind, sind Übungsaufgaben. □

Insbesondere sehen wir, dass $\mathcal{E}$ genau die Menge aller endlichen Vereinigungen von (nicht notwendigerweise paarweise disjunkten) Rechtecken in $\mathcal{R}$ ist. Nun können wir λ auf $\mathcal{E}$ fortsetzen, indem wir setzen:

$$\lambda(A) = \sum_{i=1}^n \lambda(R_i), \qquad \text{falls } A = \bigcup_{i=1}^n R_i \text{ mit paarweise disjunkten } R_1, \dots, R_n \in \mathcal{R}.$$

Da die Darstellung von A als Vereinigung paarweise disjunkter Rechtecke nicht eindeutig ist, muss man noch den Beweis führen, dass verschiedene solche Darstellungen auf den selben Wert von $\lambda(A)$ führen (Übungsaufgabe). Offensichtlich ist das so definierte λ wirklich eine Fortsetzung des oben auf $\mathcal{R}$ definierten λ. Somit ist $\lambda\colon \mathcal{E} \to [0, \infty)$ wohldefiniert und ordnet jedem Rechteck seinen elementargeometrischen Inhalt zu. Es stellt sich heraus, dass dieses λ nicht nur die oben erwähnte Additivität geerbt hat, sondern sogar eine weitaus stärkere Eigenschaft hat, die sogenannte *σ-Additivität*:

Lemma 2.1.2. *Sei $(A_n)_{n \in \mathbb{N}}$ eine Folge elementarer Mengen, so dass auch die abzählbare Vereinigung $A = \bigcup_{n \in \mathbb{N}} A_n$ in $\mathcal{E}$ liegt. Dann gilt*

$$\lambda(A) \leq \sum_{n \in \mathbb{N}} \lambda(A_n). \tag{2.1.3}$$

Falls die Mengen A_n zusätzlich noch paarweise disjunkt sind, gilt sogar

$$\lambda(A) = \sum_{n \in \mathbb{N}} \lambda(A_n). \tag{2.1.4}$$

[1] Die *symmetrische Differenz* von A und B ist definiert als die Menge $A \triangle B = (A \cup B) \setminus (A \cap B) = (A \setminus B) \cup (B \setminus A)$.

Die Eigenschaft in (2.1.4) heißt σ-Additivität und wird auch manchmal *abzählbare Additivität* genannt. Man beachte, dass fast alle der Mengen A_n auch als die leere Menge gewählt werden können, so dass (2.1.4) insbesondere auch die (einfache) Additivität von λ auf $\mathcal{E}$ beinhaltet. Die Eigenschaft in (2.1.3) wird die *abzählbare Subadditivität* oder *σ-Subadditivität* genannt.

Beweis von Lemma 2.1.2. Wir zeigen zunächst die (einfache) Additivität. Wenn $A_1, \dots, A_N$ paarweise disjunkte elementare Mengen mit Darstellungen $A_i = \bigcup_{j=1}^{n_i} R_{i,j}$ sind (wobei die Rechtecke $R_{i,1}, \dots, R_{i,n_i}$ paarweise disjunkt seien), so ist $A = \bigcup_{i=1}^N A_i = \bigcup_{i=1}^N \bigcup_{j=1}^{n_i} R_{i,j}$ eine Vereinigung von paarweise disjunkten Rechtecken. Die oben erwähnte Additivität von λ auf $\mathcal{R}$ ergibt, dass

$$\lambda(A) = \sum_{i=1}^N \sum_{j=1}^{n_i} \lambda(R_{i,j}) = \sum_{i=1}^N \lambda\Big(\bigcup_{j=1}^{n_i} R_{i,j}\Big) = \sum_{i=1}^N \lambda(A_i),$$

also ist λ auch auf $\mathcal{E}$ additiv.

An dieser Stelle sollte man sich als Übungsaufgabe davon überzeugen, dass $\lambda\colon \mathcal{E} \to [0,\infty)$ auch subadditiv und monoton ist: Für jedes $n \in \mathbb{N}$ und alle (nicht notwendigerweise paarweise disjunkten) elementaren Mengen $A_1, \dots, A_n$ gilt $\lambda(\bigcup_{i=1}^n A_i) \leq \sum_{i=1}^n \lambda(A_i)$, und für alle elementaren Mengen $A \subset B$ gilt $\lambda(A) \leq \lambda(B)$.

Nun zeigen wir die abzählbare Additivität. Zunächst zeigen wir '$\geq$' in (2.1.4). Sei $(A_n)_{n\in\mathbb{N}}$ eine Folge paarweise disjunkter elementarer Mengen. Da auch ihre Vereinigung A als elementar vorausgesetzt wurde, ist nach Lemma 2.1.1 für jedes $N \in \mathbb{N}$ auch $B_N = A \setminus (\bigcup_{n=1}^N A_n)$ elementar. Die (einfache) Additivität ergibt, dass

$$\lambda(A) = \lambda(B_N) + \sum_{n=1}^N \lambda(A_n) \geq \sum_{n=1}^N \lambda(A_n).$$

Der Grenzübergang $N \to \infty$ lässt folgen, dass in (2.1.4) '$\geq$' gilt.

Der Beweis von '$\leq$' in (2.1.4) ist schwieriger und benutzt ein Kompaktheitsargument. Aus dem Folgenden folgt auch die σ-Subadditivität von λ, also (2.1.3). Sei $(A_n)_{n\in\mathbb{N}}$ eine Folge elementarer Mengen, und sei $\varepsilon > 0$. Zunächst zeigen wir, dass eine kompakte Menge $K \subset A$ existiert mit $\lambda(K) \geq \lambda(A) - \frac{\varepsilon}{2}$. Um dies zu zeigen, wählen wir paarweise disjunkte Rechtecke $R_1, \dots, R_n$ mit $A = \bigcup_{i=1}^n R_i$ und für jedes i ein *kompaktes* Rechteck $Q_i \subset R_i$ mit $\lambda(Q_i) \geq \lambda(R_i) - \frac{\varepsilon}{2n}$. Dann erfüllt $K = \bigcup_{i=1}^n Q_i$ die Bedingung.

Auf ähnliche Weise zeigt man, dass zu jeder elementaren Menge B und jedem $\delta > 0$ eine offene elementare Menge G existiert mit $B \subset G$ und $\lambda(G) \leq \lambda(B) + \delta$. Dies wenden wir an auf die Mengen A_n: Für jedes $n \in \mathbb{N}$ gibt es also eine offene Menge G_n mit $A_n \subset G_n$ und $\lambda(G_n) \leq \lambda(A_n) + \varepsilon 2^{-n+1}$. Die oben gewählte kompakte Menge K wird von der Vereinigung der G_n überdeckt, denn K liegt in A, und A wird von den A_n, also auch von den G_n, überdeckt. Nach dem Satz von Heine-Borel gibt es eine endliche Teilüberdeckung, d. h. es gibt ein $N \in \mathbb{N}$ und (paarweise verschiedene) $r_1, \dots, r_N \in \mathbb{N}$, so dass $K \subset \bigcup_{n=1}^N G_{r_n}$. Also sehen wir mit Hilfe der oben erwähnten Monotonie und Subadditivität, dass gilt:

$$\begin{aligned}\lambda(A) &\leq \lambda(K) + \frac{\varepsilon}{2} \leq \sum_{i=1}^N \lambda(G_{r_i}) + \frac{\varepsilon}{2} \leq \sum_{n\in\mathbb{N}} \lambda(G_n) + \frac{\varepsilon}{2} \leq \sum_{n\in\mathbb{N}} \big(\lambda(A_n) + \varepsilon 2^{-n+1}\big) + \frac{\varepsilon}{2} \\ &\leq \sum_{n\in\mathbb{N}} \lambda(A_n) + \varepsilon.\end{aligned}$$

Da $\varepsilon > 0$ beliebig war, haben wir '$\leq$' in (2.1.4) sowie (2.1.3) gezeigt. $\square$

Nun wollen wir λ sogar auf die Potenzmenge von $\mathbb{R}^2$ fortsetzen. Um unendliche Maßzahlen zu vermeiden, beschränken wir uns zunächst auf $E = [0,1]^2$. Mit $\mathcal{R}_E$ und $\mathcal{E}_E$ bezeichen wir die Mengen aus $\mathcal{R}$ bzw. $\mathcal{E}$, die Teilmengen von E sind, und mit λ_E bezeichnen wir die Einschränkung von λ auf $\mathcal{E}_E$. Dann definieren wir das *äußere Maß* der Menge $A \subset E$ als die Zahl

$$\lambda_E^*(A) = \inf\Big\{\sum_{n\in\mathbb{N}} \lambda_E(R_n) \colon R_n \in \mathcal{R}_E \text{ für jedes } n \in \mathbb{N}, A \subset \bigcup_{n\in\mathbb{N}} R_n\Big\}. \tag{2.1.5}$$

Man kann sich leicht davon überzeugen, dass in dieser Definition an Stelle von Rechtecken R_n auch elementare Mengen R_n eingesetzt werden können.

Wir werden später sehen, dass dieser Erweiterungsschritt ein wenig zu groß ist: Die Abbildung $\lambda_E^* \colon \mathcal{P}(E) \to [0,1]$ hat nicht die ersehnte σ-Additivitätseigenschaft, die wir gerne von einem Maß haben wollen. Daher werden wir λ_E^* auf eine geeignete Teilmenge von $\mathcal{P}(E)$ einschränken, die Menge der *Lebesgue-messbaren Teilmengen von E*, die allerdings immer noch viel größer ist als $\mathcal{E}_E$. Zunächst zeigen wir, dass λ_E^* wirklich eine Fortsetzung von $\lambda_E \colon \mathcal{E}_E \to [0,1]$ ist:

Lemma 2.1.3. *Für jedes $A \in \mathcal{E}_E$ gilt $\lambda_E^*(A) = \lambda_E(A)$.*

Beweis. Sei $A \in \mathcal{E}_E$. Da in der Infimumsbildung in (2.1.5) auch die Vereinigung endlich vieler Rechtecke eingeschlossen ist (man nehme $R_n = \emptyset$ für fast alle n), ist klar, dass $\lambda_E^*(A) \leq \lambda_E(A)$ gilt. Falls $(R_n)_{n\in\mathbb{N}}$ eine Folge von Rechtecken ist mit $A \subset \bigcup_{n\in\mathbb{N}} R_n$, so gilt auf Grund der Monotonie und σ-Subadditivität von λ_E auf $\mathcal{E}_E$ (siehe Lemma 2.1.2), dass $\lambda_E(A) \leq \sum_{n\in\mathbb{N}} \lambda_E(R_n)$ gilt. Also folgt auch, dass $\lambda_E^*(A) \geq \lambda_E(A)$ gilt. $\square$

Nun zeigen wir die σ-Subadditivität des äußeren Maßes.

Lemma 2.1.4 (σ-Subadditivität von λ_E^*). *Sei $(A_n)_{n\in\mathbb{N}}$ eine Folge in $\mathcal{P}(E)$ und $A \in \mathcal{P}(E)$ mit $A \subset \bigcup_{n\in\mathbb{N}} A_n$. Dann gilt $\lambda_E^*(A) \leq \sum_{n\in\mathbb{N}} \lambda_E^*(A_n)$.*

Insbesondere ist λ_E^* *monoton* (auch *isoton* genannt), d. h. für alle $A, B \in \mathcal{P}(E)$ mit $A \subset B$ gilt $\lambda_E^*(A) \leq \lambda_E(B)$. Dies sieht man durch eine Anwendung auf $A_1 = B$ und $A_n = \emptyset$ für $n \geq 2$.

Beweis von Lemma 2.1.4. Sei $\varepsilon > 0$. Zu jedem $n \in \mathbb{N}$ gibt es per Definition eine höchstens abzählbare Familie $(R_{n,k})_{k\in\mathbb{N}}$ von Rechtecken mit $A_n \subset \bigcup_k R_{n,k}$ und $\sum_k \lambda_E^*(R_{n,k}) \leq \lambda_E(A_n) + \varepsilon 2^{-n}$. Dann liegt A in der abzählbaren Vereinigung der Rechtecke $R_{n,k}$, und wir haben

$$\lambda_E^*(A) \leq \sum_{n,k} \lambda_E(R_{n,k}) = \sum_{n\in\mathbb{N}} \Big(\sum_k \lambda_E(R_{n,k})\Big) \leq \sum_{n\in\mathbb{N}} \Big(\lambda_E^*(A_n) + \frac{\varepsilon}{2^n}\Big) \leq \varepsilon + \sum_{n\in\mathbb{N}} \lambda_E^*(A_n).$$

Da die linke Seite nicht von ε abhängt, ist die Aussage bewiesen. $\square$

Nun kommen wir endlich zum vorläufigen Hauptobjekt dieses Abschnittes: Wir nennen jede Menge, die mit Mengen aus $\mathcal{E}_E$ beliebig genau im Sinne des äußeren Maßes approximiert werden kann, messbar, und die Einschränkung von λ_E^* auf diese Mengen das Lebesgue-Maß auf E.

Definition 2.1.5 (Lebesgue-messbar, Lebesgue-Maß). *(i) Eine Menge $A \in \mathcal{P}(E)$ heißt* Lebesgue-messbar*, falls zu jedem $\varepsilon > 0$ ein $B \in \mathcal{E}_E$ existiert mit $\lambda_E^*(A \triangle B) \leq \varepsilon$. Sei $\mathcal{F}_E$ die Menge der Lebesgue-messbaren Teilmengen von E.*

(ii) Die Einschränkung λ_E von λ_E^* auf $\mathcal{F}_E$ heißt das Lebesgue-Maß auf E.

Man sieht leicht, dass jede elementare Menge Lebesgue-messbar ist, d. h. $\mathcal{E}_E \subset \mathcal{F}_E$. Außerdem folgt aus Lemma 2.1.3 auch leicht, dass das λ_E aus Definition 2.1.5(ii) eine Fortsetzung von λ_E ist. Die fundamentalen Eigenschaften des Mengensystems $\mathcal{F}_E$ und der Abbildung $\lambda_E\colon \mathcal{F}_E \to [0,1]$ werden nun zusammengefasst:

Satz 2.1.6. *(i) Das Mengensystem $\mathcal{F}_E$ ist abgeschlossen gegen Bildung von Differenzen, symmetrischer Differenz und abzählbarer Vereinigungs- und Durchschnittbildung.*

(ii) Die Abbildung $\lambda_E\colon \mathcal{F}_E \to [0,1]$ ist σ-additiv.

Aussage (i) sagt, dass für $A, B, A_1, A_2, A_3, \dots \in \mathcal{F}_E$ auch wieder $A \setminus B$, $A \triangle B$, $\bigcup_{n\in\mathbb{N}} A_n$ und $\bigcap_{n\in\mathbb{N}} A_n$ in $\mathcal{F}_E$ liegen. Aussage (ii) sagt, dass für jede Familie paarweiser disjunkter Mengen $A_1, A_2, A_3, \dots \in \mathcal{F}_E$ gilt: $\lambda_E(\bigcup_{n\in\mathbb{N}} A_n) = \sum_{n\in\mathbb{N}} \lambda_E(A_n)$. In der später benutzten Sprechweise heißt dies, dass $\mathcal{F}$ eine σ-Algebra ist und λ_E ein Maß darauf. Der Beweis von Satz 2.1.6 wird über die folgenden Lemmata verteilt.

Lemma 2.1.7 (Abgeschlossenheit von $\mathcal{F}_E$ gegen endliche Operationen). *Seien $n \in \mathbb{N}$ und $A, B, A_1, A_2, \dots, A_n$ in $\mathcal{F}_E$. Dann liegen auch folgende Mengen in $\mathcal{F}_E$:*

$$\text{(i) } E \setminus A, \qquad \text{(ii) } \bigcup_{i=1}^n A_i, \qquad \text{(iii) } \bigcap_{i=1}^n A_i, \qquad \text{(iv) } A \setminus B, \qquad \text{(v) } A \triangle B.$$

Beweis. (i) folgt aus $(E \setminus A) \triangle (E \setminus B) = A \triangle B$ und der Definition der Lebesgue-Messbarkeit.

Aussage (ii) für $n = 2$ (den Fall eines allgemeinen n beweist man dann durch Induktion) folgt aus $(A_1 \cup A_2) \triangle (B_1 \cup B_2) \subset (A_1 \triangle B_1) \cup (A_2 \triangle B_2)$ mit Hilfe der Monotonie und Subadditivität von λ_E.

Wegen $\bigcap_{i=1}^n A_i = E \setminus \bigcup_{i=1}^n (E \setminus A_i)$ folgt (iii) aus (i) und (ii).

(iv) ist klar wegen $A \setminus B = A \cap (E \setminus B)$ und (i) und (iii).

(v) folgt wegen $A \triangle B = (A \setminus B) \cup (B \setminus A)$ aus (ii) und (iv). □

Lemma 2.1.8 (Endliche Additivität von λ_E). *Seien $n \in \mathbb{N}$ und $A_1, \dots, A_n \in \mathcal{F}_E$ paarweise disjunkt, dann gilt $\lambda_E(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n \lambda_E(A_i)$.*

Beweis. Wir dürfen uns auf $n = 2$ beschränken. Wegen der Subadditivität von λ_E muss nur noch gezeigt werden, dass $\lambda_E(A_1 \cup A_2) \geq \lambda_E(A_1) + \lambda_E(A_2)$ gilt. Sei $\varepsilon > 0$, dann können wir elementare Mengen B_1, B_2 wählen mit $\lambda_E^*(A_i \triangle B_i) < \varepsilon$ für $i \in \{1, 2\}$. Wir zeigen zunächst, dass für beliebige $A, B \in \mathcal{F}_E$ gilt:

$$|\lambda_E^*(A) - \lambda_E^*(B)| \leq \lambda_E^*(A \triangle B). \tag{2.1.6}$$

Um dies zu sehen, benutzen wir die Subadditivität von λ_E^* und die Beziehung $A \subset (A \triangle B) \cup B$, um $\lambda_E^*(A) \leq \lambda_E^*(A \triangle B) + \lambda_E^*(B)$ zu erhalten. Daraus folgt die behauptete Ungleichung ohne Betragstriche, und aus einer Vertauschung der Rollen von A und B folgt die Ungleichung selber.

Also haben wir $|\lambda_E(A_i) - \lambda_E(B_i)| < \varepsilon$ für $i \in \{1, 2\}$.

Die elementaren Mengen B_1 und B_2 sind i. Allg. nicht disjunkt, aber wir können abschätzen:

$$\lambda_E(B_1 \cap B_2) \leq \lambda_E\big((A_1 \triangle B_1) \cup (A_2 \triangle B_2)\big) \leq \lambda_E^*(A_1 \triangle B_1) + \lambda_E^*(A_2 \triangle B_2) < 2\varepsilon,$$

wobei die erste Ungleichung aus der Inklusion $B_1 \cap B_2 \subset (A_1 \triangle B_1) \cup (A_2 \triangle B_2)$ folgt, die man als eine kleine Übungsaufgabe beweist (hier geht die Disjunktheit von A_1 und A_2 ein). Die Menge $B = B_1 \cup B_2$ ist ebenfalls elementar, und $A \triangle B$ liegt in $(A_1 \triangle B_1) \cup (A_2 \triangle B_2)$, wobei wir $A = A_1 \cup A_2$ setzten. Die Subadditivität ergibt, dass $\lambda_E^*(A \triangle B) < 2\varepsilon$ gilt. Also folgt aus (2.1.6), dass auch $|\lambda_E^*(A) - \lambda_E^*(B)| < 2\varepsilon$ ist. Nun können wir zusammensetzen und erhalten, indem wir auch die Additivität von λ_E auf $\mathcal{E}_E$ benutzten (siehe Lemma 2.1.2),

$$\begin{aligned}\lambda_E(A) &\geq \lambda_E(B) - 2\varepsilon = \lambda_E(B_1) + \lambda_E(B_2) - \lambda_E(B_1 \cap B_2) - 2\varepsilon > \lambda_E(B_1) + \lambda_E(B_2) - 4\varepsilon \\ &> \lambda_E(A_1) + \lambda_E(A_2) - 6\varepsilon.\end{aligned}$$

Da die linke Seite nicht von ε abhängt, ist der Beweis beendet. □

Lemma 2.1.9 (Abgeschlossenheit von $\mathcal{F}_E$ gegen abzählbare Operationen). *Seien die Mengen $A_1, A_2, \dots \in \mathcal{F}_E$, dann liegen auch $\bigcup_{n\in\mathbb{N}} A_n$ und $\bigcap_{n\in\mathbb{N}} A_n$ in $\mathcal{F}_E$.*

Beweis. Wir brauchen nur die Vereinigung zu betrachten, denn $\bigcap_{n\in\mathbb{N}} A_n = (\bigcup_{n\in\mathbb{N}} A_n^{\mathrm{c}})^{\mathrm{c}}$, und $\mathcal{F}_E$ ist abgeschlossen gegen Komplementbildung nach Lemma 2.1.7.

Zunächst sehen wir, dass wir die Menge $A = \bigcup_{n\in\mathbb{N}} A_n$ auch als *disjunkte* Vereinigung $\bigcup_{n\in\mathbb{N}} A_n'$ schreiben können mit $A_1' = A_1$ und $A_n' = A_n \setminus \bigcup_{i=1}^{n-1} A_i$ für $n \geq 2$. Sei $\varepsilon > 0$. Da die endliche Vereinigung $\bigcup_{i=1}^n A_i'$ nach Lemma 2.1.7(ii) in $\mathcal{F}_E$ liegt, gibt es eine Menge B_n in $\mathcal{E}_E$ mit $\lambda_E^*((\bigcup_{i=1}^n A_i') \triangle B_n) < \frac{\varepsilon}{2}$. Außerdem ist

$$A \triangle B_n \subset \Big(\Big(\bigcup_{i=1}^n A_i'\Big) \triangle B_n\Big) \cup \bigcup_{i=n+1}^\infty A_i'.$$

Auf Grund der σ-Subadditivität des äußeren Maßes erhalten wir daraus

$$\lambda_E^*(A \triangle B_n) \leq \lambda_E^*\Big(\Big(\bigcup_{i=1}^n A_i'\Big) \triangle B_n\Big) + \sum_{i=n+1}^\infty \lambda_E^*(A_i') \leq \frac{\varepsilon}{2} + \sum_{i=n+1}^\infty \lambda_E^*(A_i').$$

Der letzte Term verschwindet für $n \to \infty$, denn die gesamte Reihe konvergiert: Für jedes $N \in \mathbb{N}$ haben wir auf Grund der Additivität von λ_E

$$\sum_{i=1}^N \lambda_E^*(A_i') = \sum_{i=1}^N \lambda_E(A_i') = \lambda_E\Big(\bigcup_{i=1}^N A_i'\Big) \leq 1,$$

also auch $\sum_{i\in\mathbb{N}} \lambda_E^*(A_i') < \infty$. Also kann man n so groß wählen, dass die elementare Menge B_n erfüllt: $\lambda_E^*(A \triangle B_n) \leq \varepsilon$. Daher ist A Lebesgue-messbar, und der Beweis ist beendet. □

Mit dem folgenden Lemma wird der Beweis von Satz 2.1.6 beendet:

Lemma 2.1.10. *Die Mengenfunktion $\lambda_E \colon \mathcal{F}_E \to [0, 1]$ ist σ-additiv.*

Beweis. Sei $(A_n)_{n\in\mathbb{N}}$ eine Folge paarweise disjunkter Mengen in $\mathcal{F}_E$, und wir setzen $A = \bigcup_{n\in\mathbb{N}} A_n$. Nach Lemma 2.1.9 ist A auch Lebesgue-messbar. Wir müssen zeigen, dass $\lambda_E(A) = \sum_{n\in\mathbb{N}} \lambda_E(A_n)$ gilt. Wegen der σ-Subadditivität (siehe Lemma 2.1.4) gilt schon '$\leq$', so dass wir nur noch '$\geq$' zeigen müssen. Mit Hilfe der endlichen Additivität (siehe Lemma 2.1.8) und Monotonie von λ_E sehen wir, dass für jedes $N \in \mathbb{N}$ gilt:

$$\sum_{n=1}^N \lambda_E(A_n) = \lambda_E\Big(\bigcup_{n=1}^N A_n\Big) \leq \lambda_E(A).$$

Der Grenzübergang $N \to \infty$ lässt die Behauptung folgen. $\square$

Also haben wir nun auf dem Einheitsquadrat $E = [0,1]^2$ ein Mengensystem $\mathcal{F}_E$ definiert, das gegen alle abzählbaren elementaren Operationen abgeschlossen ist, und auf $\mathcal{F}_E$ eine Abbildung λ_E, die kompatibel mit diesen Mengenoperationen ist und die wir deswegen ein Maß auf E nennen dürfen.

Es bleibt noch, dieses Maß auf den ganzen $\mathbb{R}^2$ zu erweitern. Dies machen wir auf die folgende Weise. Für $i, j \in \mathbb{Z}$ setzen wir

$$E_{i,j} = [i, i+1] \times [j, j+1] \qquad \text{und} \qquad E^0_{i,j} = [i, i+1) \times [j, j+1),$$

dann ist $\mathbb{R}^2$ gleich der Vereinigung der $E_{i,j}$ und auch gleich der Vereinigung der $E^0_{i,j}$, wobei wir bemerken, dass letztere disjunkt ist. Mit $\mathcal{F}_{i,j}$ bezeichnen wir die Menge der Lebesgue-messbaren Teilmengen von $E_{i,j}$, wobei man sich leicht davon überzeugen kann, dass alles Bisherige auch für diese Mengen an Stelle von E durchgeführt werden kann. Insbesondere haben wir auf $E_{i,j}$ ein Lebesgue-Maß $\lambda_{i,j}$. Wir definieren nun eine Menge $A \subset \mathbb{R}^2$ als *Lebesgue-messbar*, wenn für alle $i, j \in \mathbb{Z}$ die Menge $A \cap E_{i,j}$ in $\mathcal{F}_{i,j}$ liegt. Sei $\mathcal{F}$ die Menge der Lebesgue-messbaren Teilmengen des $\mathbb{R}^2$. Auf $\mathcal{F}$ definieren wir nun das *zweidimensionale Lebesgue-Maß* $\lambda \colon \mathcal{F} \to [0, \infty]$ durch

$$\lambda(A) = \sum_{i,j \in \mathbb{Z}} \lambda_{i,j}(A \cap E_{i,j}).$$

Bemerkung 2.1.11 (Vollständigkeit des Lebesgue-Maßes auf $\mathcal{F}$). Als eine Übungsaufgabe zeigt man, dass jede Teilmenge einer Menge $A \in \mathcal{F}$ mit $\lambda(A) = 0$ ebenfalls wieder in $\mathcal{F}$ liegt. Wir werden in Bemerkung 2.2.11 diese Eigenschaft die Vollständigkeit des Mengensystems $\mathcal{F}$ nennen. Mit Hilfe dieser Eigenschaft zeigt man als Übungsaufgabe, dass eine Menge $A \subset \mathbb{R}^2$ genau dann Lebesgue-messbar ist, wenn für alle $i, j \in \mathbb{Z}$ die Menge $A \cap E^0_{i,j}$ in $\mathcal{F}_{i,j}$ liegt, und dass auch gilt:

$$\lambda(A) = \sum_{i,j \in \mathbb{Z}} \lambda_{i,j}(A \cap E^0_{i,j}).$$

$\diamond$

Es lässt sich auch leicht zeigen, dass das Mengensystem $\mathcal{F}$ und die Mengenfunktion $\lambda \colon \mathcal{F} \to [0, \infty]$ ebenfalls die Kompatibilitätseigenschaften von $\mathcal{F}_E$ und λ_E besitzen:

Satz 2.1.12. *(i) Das Mengensystem $\mathcal{F}$ ist abgeschlossen gegen Bildung von Differenzen, symmetrischer Differenz und abzählbarer Vereinigungs- und Durchschnittbildung.*

(ii) Die Abbildung $\lambda \colon \mathcal{F} \to [0, \infty]$ ist σ-additiv.

In der Sprechweise der allgemeinen Maßtheorie (siehe Abschnitt 2.2) heißt dies, dass $\mathcal{F}$ eine σ-Algebra ist und λ ein Maß darauf.

Wir bemerken noch, dass das System $\mathcal{F}$ sehr unhandlich und riesengroß ist, aber dass es nicht $\mathcal{P}(\mathbb{R}^2)$ umfasst, d. h., dass es Mengen gibt, die nicht Lebesgue-messbar sind. In Beispiel 2.2.14 wird mit Hilfe des Auswahlaxioms eine solche Menge konstruiert. Als Faustregel kann man sagen, dass alle Mengen, die mehr oder weniger explizit konstruiert werden können durch abzählbare Operationen, in $\mathcal{F}$ liegen. Zur Konstruktion nicht Lebesgue-messbarer Mengen ist bisher immer das Auswahlaxiom benutzt worden.

2.2 σ-Algebren und Maße

Nach der Motivation an Hand des zweidimensionalen Lebesgue-Maßes wollen wir die Grundbegriffe der allgemeinen Maßtheorie einführen. Es sei Ω eine nichtleere, eventuell sehr große Menge. Wir möchten auf Ω ein Maß einführen, um die Teilmengen von Ω zu messen. In anderen Worten, wir möchten eine Abbildung μ auf der Menge der Teilmengen von Ω mit Werten in $[0, \infty]$ festlegen, so dass $\mu(A)$ in einem vernünftigen Sinn das Maß einer Menge A angibt. Zu den Grundanforderungen an μ gehört sicher die Forderung, dass für jede Folge von paarweise disjunkten Teilmengen $A_1, A_2, \dots$ gelten soll: $\mu(\bigcup_{n\in\mathbb{N}} A_n) = \sum_{n\in\mathbb{N}} \mu(A_n)$. Diese Eigenschaft nennt man die *σ-Additivität*; sie ist eine der Grundpfeiler der Maßtheorie.

Allerdings definiert man ein Maß fast nie auf der *gesamten* Menge aller Teilmengen von Ω (der sogenannten *Potenzmenge* $\mathcal{P}(\Omega)$ von Ω), sondern auf Teilmengensystemen, die kompatibel sind mit der σ-Additivität und auch noch mit dem Übergang zu Komplementen. Die betrachteten Systeme müssen abgeschlossen sein gegenüber allen elementaren *abzählbaren* Mengenoperationen:

Definition 2.2.1 (σ-Algebra). *Ein Mengensystem $\mathcal{F}$ auf Ω (also eine Teilmenge $\mathcal{F}$ von $\mathcal{P}(\Omega)$) heißt eine* σ-Algebra, *falls gelten:*

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

(ii) falls $A \in \mathcal{F}$, so gilt auch $A^{\mathrm{c}} \in \mathcal{F}$,

(iii) falls $A_1, A_2, A_3, \dots \in \mathcal{F}$, so gilt auch $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{F}$.

Die Elemente von $\mathcal{F}$ heißen messbare Mengen *und das Paar $(\Omega, \mathcal{F})$ ein* messbarer Raum.

Also sind σ-Algebren geeignete Definitionsbereiche für Maße, siehe Definition 2.2.7 weiter unten. Die folgenden Bemerkungen sind offensichtlich oder Übungsaufgaben.

Bemerkung 2.2.2. (a) Die Potenzmenge und $\{\emptyset, \Omega\}$ sind zwei σ-Algebren, und zwar die beiden trivialen. Alle anderen σ-Algebren liegen zwischen diesen beiden.

(b) Jede σ-Algebra $\mathcal{F}$ ist nicht nur gegen Komplementbildung und abzählbare Vereinigung abgeschlossen, sondern auch gegen abzählbare Schnittbildung, d. h. auch $\bigcap_{n\in\mathbb{N}} A_n$ liegt in $\mathcal{F}$, wenn $A_1, A_2, A_3, \dots \in \mathcal{F}$, denn $\bigcap_{n\in\mathbb{N}} A_n = (\bigcup_{n\in\mathbb{N}} A_n^{\mathrm{c}})^{\mathrm{c}}$.

(c) Beliebige Schnitte von σ-Algebren sind ebenfalls σ-Algebren, aber im Allgemeinen sind Vereinigungen von σ-Algebren selber keine.

(d) Für jede Menge Ω ist das System der Mengen $A \subset \Omega$, so dass A oder A^{c} abzählbar ist, eine σ-Algebra.

(e) Wenn Ω und Ω' zwei Mengen sind und $f\colon \Omega \to \Omega'$ eine Abbildung, dann ist für jede σ-Algebra $\mathcal{F}'$ auf Ω' das System aller Urbilder $f^{-1}(A') = \{\omega \in \Omega \colon f(\omega) \in A'\}$ mit $A' \in \mathcal{F}'$ eine σ-Algebra auf Ω. Im Allgemeinen ist aber für eine σ-Algebra $\mathcal{F}$ auf Ω das System der Bilder $f(A)$ mit $A \in \mathcal{F}$ keine σ-Algebra.

$\diamond$

Im Allgemeinen gibt es sehr viele σ-Algebren auf Ω, und viele σ-Algebren sind nicht nur selber gigantisch groß, sondern auch kaum explizit zu beschreiben. Daher ist es oft nützlich, Erzeugendensysteme gut zu kennen:

Lemma 2.2.3 (erzeugte σ-Algebra). *Zu jedem Mengensystem $\mathcal{C}$ auf Ω gibt es genau eine kleinste σ-Algebra $\sigma(\mathcal{C})$, die $\mathcal{C}$ enthält, d. h. eine σ-Algebra, die $\mathcal{C}$ enthält und die in jeder σ-Algebra enthalten ist, die $\mathcal{C}$ enthält. Wir nennen $\sigma(\mathcal{C})$ die* von $\mathcal{C}$ erzeugte σ-Algebra.

Beweis. Wir können $\sigma(\mathcal{C})$ angeben als[2]

$$\sigma(\mathcal{C}) = \bigcap\{\mathcal{A}\colon \mathcal{A} \text{ ist eine } \sigma\text{-Algebra mit } \mathcal{C} \subset \mathcal{A}\}.$$

Diese Definition ist sinnvoll (also ein nichtleerer Schnitt), denn mindestens $\mathcal{P}(\Omega)$ ist in dem Mengensystem enthalten, über das der Schnitt gebildet wird. Man verifiziert leicht, dass dieses Mengensystem die beiden Eigenschaften besitzt. □

Es ist klar, dass immer $\mathcal{C} \subset \sigma(\mathcal{C})$ gilt und dass Gleichheit genau dann gilt, wenn $\mathcal{C}$ eine σ-Algebra ist. Ferner gilt $\sigma(\mathcal{C}_1) \subset \sigma(\mathcal{C}_2)$, falls $\mathcal{C}_1 \subset \mathcal{C}_2$.

Bemerkung 2.2.4. Wenn für jedes i aus einer beliebigen Indexmenge I das System $\mathcal{F}_i$ eine σ-Algebra ist, so ist ja (wie schon in Bemerkung 2.2.2(c) erwähnt) die Vereinigung $\bigcup_{i\in I} \mathcal{F}_i$ im Allgemeinen keine σ-Algebra. Die kleinste σ-Algebra, die diese Vereinigung enthält, wird mit

$$\bigvee_{i\in I} \mathcal{F}_i = \sigma\Big(\bigcup_{i\in I} \mathcal{F}_i\Big)$$

bezeichnet. Im Fall $I = \{1, \dots, n\}$ schreibt man auch $\mathcal{F}_1 \vee \dots \vee \mathcal{F}_n$. ◇

Die wichtigste σ-Algebra auf $\mathbb{R}$ bzw. $\mathbb{R}^d$ ist die folgende.

Beispiel 2.2.5 (Borel-σ-Algebra). (a) Auf $\mathbb{R}$ betrachten wir die von den offenen Mengen erzeugte σ-Algebra, also $\mathcal{B} = \sigma(\mathcal{O})$, wobei $\mathcal{O}$ die Menge aller offenen Teilmengen von $\mathbb{R}$ ist. Die σ-Algebra $\mathcal{B}$ heißt die *Borel-σ-Algebra* auf $\mathbb{R}$. Es gilt auch

$$\begin{aligned}\mathcal{B} &= \sigma(\{(a,b)\colon -\infty \le a < b \le \infty\}\\ &= \sigma(\{(-\infty, t]\colon t \in \mathbb{R}\})\\ &= \sigma(\{(-\infty, t]\colon t \in \mathbb{Q}\}),\end{aligned}$$

d. h. $\mathcal{B}$ wird auch von dem System der offenen Intervalle, dem der links unbeschränkten und rechts beschränkten, abgeschlossenen Intervalle (mit rationalem Randpunkt) erzeugt (Übungsaufgaben). Es gibt kein direktes Kriterium, an Hand dessen man von einer gegebenen Teilmenge von $\mathbb{R}$ erkennen kann, ob sie zu $\mathcal{B}$ gehört oder nicht.[3]

(b) Auf dem $\mathbb{R}^d$ betrachten wir die von den offenen Mengen erzeugte σ-Algebra $\mathcal{B}_d$, die die *Borel-σ-Algebra* auf $\mathbb{R}^d$ genannt wird. Auch $\mathcal{B}_d$ wird von einer Reihe anderer Mengensysteme erzeugt, wie etwa vom System der offenen Kuben $\prod_{i=1}^d (a_i, b_i)$ mit $-\infty \le a_i < b_i \le \infty$ für $i = 1, \dots, d$, dem der abgeschlossenen Mengen oder auch von dem der kompakten Mengen (Übungsaufgabe).

◇

[2] Wir schreiben $\bigcap\{\mathcal{A}_i\colon i \in I\}$ an Stelle von $\bigcap_{i\in I} \mathcal{A}_i = \{M \subset \Omega\colon M \in \mathcal{A}_i\, \forall i \in I\}$.

[3] Dies ist etwa bei dem Begriff der Offenheit einer Menge anders: Falls zu jedem Punkt der Menge noch eine ganze Umgebung dazu gehört, hat die Menge diese Eigenschaft.

Es sei bemerkt, dass die Borel-σ-Algebra echt kleiner ist als die in Definition 2.1.5 eingeführte σ-Algebra der Lebesgue-messbaren Mengen. Genauer gesagt, ist die letztere die Vervollständigung der ersteren, siehe Definition 2.2.11 und Bemerkung 2.2.13. (Der Beweis dieser Aussage ist eine Übungsaufgabe.)

Bemerkung 2.2.6 (Spur-σ-Algebra). Wenn $\mathcal{F}$ eine σ-Algebra über Ω ist und $\widetilde{\Omega} \subset \Omega$ eine beliebige Teilmenge von Ω, so sieht man leicht, dass

$$\widetilde{\mathcal{F}} = \{A \cap \widetilde{\Omega} \colon A \in \mathcal{F}\}$$

eine σ-Algebra über $\widetilde{\Omega}$ ist. Wir nennen $\widetilde{\mathcal{F}}$ die *Spur-σ-Algebra* von $\mathcal{F}$ über $\widetilde{\Omega}$. Falls $\widetilde{\Omega}$ selber messbar ist, so besteht $\widetilde{\mathcal{F}}$ aus allen messbaren Teilmengen von $\widetilde{\Omega}$. Beispielsweise werden wir die Spur-σ-Algebra der Borel-σ-Algebra auf einer Teilmenge I von $\mathbb{R}^d$ mit $\mathcal{B}_I$ bezeichnen. $\diamond$

Der zweite zentrale Begriff der Maßtheorie ist das Maß:

Definition 2.2.7 (Maß, Maßraum, Wahrscheinlichkeitsraum). *Es sei $(\Omega, \mathcal{F})$ ein messbarer Raum. Es sei $\mu\colon \mathcal{F} \to [0, \infty]$ eine Abbildung.*

(a) μ heißt ein Maß*, falls gelten:*

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

*(ii) μ ist σ-*additiv*, d. h. für alle Folgen $(A_n)_{n\in\mathbb{N}}$ von paarweise disjunkten Mengen $A_n \in \mathcal{F}$ gilt $\mu(\bigcup_{n\in\mathbb{N}} A_n) = \sum_{n\in\mathbb{N}} \mu(A_n)$.*

In diesem Fall heißt das Tripel $(\Omega, \mathcal{F}, \mu)$ ein Maßraum.

*(b) μ heißt σ-*endlich*, falls eine aufsteigende Folge $(\Omega_n)_{n\in\mathbb{N}}$ von messbaren Teilmengen von Ω existiert mit $\Omega = \bigcup_{n\in\mathbb{N}} \Omega_n$ und $\mu(\Omega_n) < \infty$ für jedes $n \in \mathbb{N}$.*

(c) μ heißt endlich*, falls $\mu(\Omega) < \infty$.*

(d) μ heißt ein Wahrscheinlichkeitsmaß*, falls μ ein Maß ist mit $\mu(\Omega) = 1$. In diesem Fall heißt das Tripel $(\Omega, \mathcal{F}, \mu)$ ein* Wahrscheinlichkeitsraum*, und wir schreiben meist $\mathbb{P}$ statt μ.*

Man nennt $\mu(\Omega)$ die *Masse* des Maßes μ. Die Bedingung $\mu(\emptyset) = 0$ soll ausschließen, dass die konstante Abbildung $\mathcal{F} \to \{\infty\}$ ein Maß ist. Man sieht leicht, dass positive Vielfache eines Maßes wieder Maße sind. Also kann man ein beliebiges endliches Maß $\mu \neq 0$ leicht zu einem Wahrscheinlichkeitsmaß $\mathbb{P} = \mu/\mu(\Omega)$ normieren. Außerdem zeigt man als eine Übungsaufgabe, dass abzählbare Summen von Maßen ebenfalls Maße sind.

Wir erwähnen schon einmal das natürlichste Maß auf der Borel-σ-Algebra:

Beispiel 2.2.8 (Lebesgue-Borel-Maß). Auf dem messbaren Raum $(\mathbb{R}^d, \mathcal{B}_d)$ (siehe Beispiel 2.2.5) wollen wir ein Maß λ_d betrachten, das jedem Kubus $Q = \prod_{i=1}^d [a_i, b_i]$ seinen elementargeometrischen Inhalt $\lambda_d(Q) = \prod_{i=1}^d (b_i - a_i)$ zuordnet. Ein solches Maß nennen wir das *Lebesgue-Borel-Maß* auf $(\mathbb{R}^d, \mathcal{B}_d)$. Allerdings müssen wir die Frage nach der Existenz und Eindeutigkeit noch ein wenig zurück stellen, siehe Korollar 2.4.3 und Beispiel 3.8.7. (Es wird sich später herausstellen, dass das in Abschnitt 2.1 konstruierte Lebesgue-Maß die Vervollständigung des Lebesgue-Borel-Maßes ist, siehe Bemerkung 2.2.13.) $\diamond$

Die Existenz und Eindeutigkeit von Maßen mit gewissen gewünschten Eigenschaften behandeln wir in Abschnitt 2.3. Wir sammeln zunächst ein paar fundamentale Eigenschaften von Maßen:

Lemma 2.2.9 (Eigenschaften von Maßen). *Es sei μ ein Maß auf einer σ-Algebra $\mathcal{F}$. Dann hat μ die folgenden Eigenschaften.*

(a) μ ist stetig von unten und von oben*, d. h.*

  (i) für jede Folge $(A_n)_{n\in\mathbb{N}}$ in $\mathcal{F}$ mit $A_n \uparrow A$[4] für ein $A \in \mathcal{F}$ gilt $\lim_{n\to\infty} \mu(A_n) = \mu(A)$,

  (ii) für jede Folge $(A_n)_{n\in\mathbb{N}}$ in $\mathcal{F}$ mit $\mu(A_n) < \infty$ für ein $n \in \mathbb{N}$ und $A_n \downarrow A$ für ein $A \in \mathcal{F}$ gilt $\lim_{n\to\infty} \mu(A_n) = \mu(A)$.

(b) μ ist σ-subadditiv*, d. h. für jede Folge $(A_n)_{n\in\mathbb{N}}$ in $\mathcal{F}$ gilt $\mu(\bigcup_{n\in\mathbb{N}} A_n) \leq \sum_{n\in\mathbb{N}} \mu(A_n)$.*

(c) μ ist monoton *oder* isoton*, d. h. für je zwei Mengen $A, B \in \mathcal{F}$ mit $A \subset B$ gilt $\mu(A) \leq \mu(B)$.*

Beweis. (a) Wir behandeln nur den Fall $A_n \uparrow A$. Mit $A_0 = \emptyset$ definieren wir $B_n = A_n \setminus A_{n-1}$, dann sind die Mengen $B_1, B_2, B_3, \dots$ paarweise disjunkt, und ihre Vereinigung ist A. Also gilt

$$\mu(A) = \mu\Big(\bigcup_{n\in\mathbb{N}} B_n\Big) = \sum_{n\in\mathbb{N}} \mu(B_n) = \lim_{N\to\infty} \sum_{n=1}^{N} \mu(B_n) = \lim_{N\to\infty} \mu\Big(\bigcup_{n=1}^{N} B_n\Big) = \lim_{N\to\infty} \mu(A_N).$$

(b) und (c) Übungsaufgabe. □

Es ist eine der Stärken der Maßtheorie, dass auch die Theorie der Reihen als ein Spezialfall in dem allgemeinen Konzepts von Definition 2.2.7 eingeschlossen ist:

Beispiel 2.2.10 (diskrete Maße, Dirac-Maß). Es sei Ω beliebig und $\widetilde{\Omega} \subset \Omega$ eine höchstens abzählbare Teilmenge. Mit einer Abbildung $p\colon \widetilde{\Omega} \to (0, \infty]$ definieren wir

$$\mu(A) = \sum_{\omega\in A\cap\widetilde{\Omega}} p(\omega), \qquad A \subset \Omega.$$

Dann ist μ ein Maß auf der Potenzmenge von Ω, das auf der diskreten Teilmenge $\widetilde{\Omega}$ konzentriert[5] ist. Es gilt dann $\mu(\{\omega\}) = p(\omega) > 0$ für alle $\omega \in \widetilde{\Omega}$, und man nennt ω ein *Atom* von μ. Das Maß μ (oder auch seine Einschränkung auf die Potenzmenge von $\widetilde{\Omega}$) heißt ein *diskretes Maß*. Falls $p(\omega) < \infty$ für alle $\omega \in \widetilde{\Omega}$, so ist μ σ-endlich. Falls $\sum_{\omega\in\widetilde{\Omega}} p(\omega) = 1$, so ist μ ein diskretes Wahrscheinlichkeitsmaß. Falls $p(\omega) = 1$ für jedes $\omega \in \widetilde{\Omega}$, nennen wir μ das *Zählmaß auf* $\widetilde{\Omega}$; dann zählt $\mu(A) = \#(A \cap \widetilde{\Omega})$ die Anzahl der Punkte aus $\widetilde{\Omega}$ in A.

Falls $\widetilde{\Omega}$ einelementig ist, also $\widetilde{\Omega} = \{\omega_0\}$ für ein $\omega_0 \in \Omega$, so heißt μ das *Dirac-Maß* in ω, und wir schreiben $\mu = \delta_{\omega_0}$. Es gilt also

$$\delta_{\omega_0}(A) = \begin{cases} 1, & \text{falls } \omega_0 \in A, \\ 0 & \text{sonst.} \end{cases} \tag{2.2.1}$$

[4] Wir schreiben $A_n \uparrow A$, falls $A_n \subset A_{n+1}$ für jedes $n \in \mathbb{N}$ und $A = \bigcup_{n\in\mathbb{N}} A_n$, und wir schreiben $A_n \downarrow A$, falls $A_{n+1} \subset A_n$ für jedes $n \in \mathbb{N}$ und $A = \bigcap_{n\in\mathbb{N}} A_n$.

[5] Wir sagen, ein Maß ist *konzentriert* auf einer messbaren Menge $\widetilde{\Omega} \subset \Omega$, falls $\mu(\Omega \setminus \widetilde{\Omega}) = 0$ gilt.

Ein beliebiges diskretes Maß μ wie oben kann man dann auch als eine Linearkombination von Dirac-Maßen auffassen, denn es gilt $\mu = \sum_{\omega\in\widetilde{\Omega}} p(\omega)\delta_\omega$. Wenn man z. B. $\widetilde{\Omega} = \Omega = \mathbb{N}$ wählt, so kann man für nichtnegative Koeffizienten a_n die (eventuell unendliche) Summe $\sum_{n\in\mathbb{N}} a_n$ als die Masse $\mu(\mathbb{N})$ auffassen, wenn $\mu = \sum_{n\in\mathbb{N}} a_n\delta_n$.

Jede aus der diskreten Wahrscheinlichkeitstheorie bekannte Verteilung lässt sich auffassen als ein diskretes Maß auf $(\mathbb{R}, \mathcal{P}(\mathbb{R}))$. Die wichtigsten Beispiele sind die folgenden:

geometrische Verteilung	$\displaystyle\sum_{n\in\mathbb{N}_0} p(1-p)^n\delta_n,$	Parameter $p \in (0,1)$,
Poisson-Verteilung	$\displaystyle\sum_{n\in\mathbb{N}_0} \frac{\alpha^n}{n!}e^{-\alpha}\delta_n,$	Parameter $\alpha > 0$,
Binomial-Verteilung	$\displaystyle\sum_{k=0}^{n} \binom{n}{k} p^k(1-p)^{n-k}\delta_k,$	Parameter $n \in \mathbb{N}, p \in (0,1)$.

◇

Ein oft benutzter Begriff im Zusammenhang mit Maßen ist der folgende.

Definition 2.2.11 (Nullmengen, Vollständigkeit). *Es sei $(\Omega, \mathcal{F}, \mu)$ ein Maßraum. Eine Menge $A \subset \Omega$ (auch wenn sie nicht in $\mathcal{F}$ liegt) heißt eine* μ-Nullmenge, *falls eine messbare Menge $F \in \mathcal{F}$ existiert mit $\mu(F) = 0$ und $A \subset F$. Der Maßraum $(\Omega, \mathcal{F}, \mu)$ heißt* vollständig, *falls $\mathcal{F}$ alle μ-Nullmengen enthält.*

Als eine einfache Übungsaufgabe zeigt man leicht:

Lemma 2.2.12. *Abzählbare Vereinigungen von Nullmengen sind Nullmengen.*

Bemerkung 2.2.13 (Vervollständigung). Jeder Maßraum $(\Omega, \mathcal{F}, \mu)$ lässt sich sehr leicht vervollständigen. Das Mengensystem

$$\mathcal{G} = \{A \subset \Omega \colon \text{Es gibt ein } B \in \mathcal{F}, \text{ so dass } A\triangle B \text{ eine } \mu\text{-Nullmenge ist}\}$$

kann man leicht als eine σ-Algebra erkennen. Ferner lässt sich μ leicht auf $\mathcal{G}$ erweitern: Für $A \in \mathcal{G}$ existiert ja ein $B \in \mathcal{F}$, so dass $A\triangle B$ eine μ-Nullmenge ist. Nun setzt man $\overline{\mu}(A) = \mu(B)$. Als Übungsaufgabe verifiziert man leicht, dass $\overline{\mu}$ wohldefiniert ist (d. h. dass diese Definition nicht von der Wahl von B abhängt) und dass $(\Omega, \mathcal{G}, \overline{\mu})$ ein vollständiger Maßraum ist. Man nennt ihn die *Vervollständigung* von $(\Omega, \mathcal{F}, \mu)$. Es hat manchmal Vorteile, mit einer Vervollständigung zu arbeiten, aber man darf nicht vergessen, dass die σ-Algebra $\mathcal{G}$ vom Maß μ abhängt. ◇

Warum konstruieren wir nicht alle Maße, die uns interessieren, gleich auf der Potenzmenge von Ω, sondern machen uns die Mühe, auf σ-Algebren einzuschränken, die ja ein wenig widerspenstig sind? Ein berühmtes Beispiel von Vitali (1905) gibt eine überzeugende Antwort:

Beispiel 2.2.14 (Potenzmenge ist zu groß). Es gibt kein translationsinvariantes Maß auf der Potenzmenge von $\mathbb{R}$, das jedem Intervall gerade seine Länge zuordnet, d. h. man kann das Lebesgue-Borel-Maß aus Beispiel 2.2.8 nicht auf $\mathcal{P}(\mathbb{R})$ erweitern. (Der Beweis der Translationsinvarianz des Lebesgue-Borel-Maßes ist eine Übungsaufgabe.)

Nehmen wir an, dass $\lambda \colon \mathcal{P}(\mathbb{R}) \to [0,\infty]$ ein solches Maß sei, und betrachten wir die Äquivalenzrelation $\sim$ auf $\mathbb{R}$, die definiert wird durch $x \sim y \iff x - y \in \mathbb{Q}$. Unter Benutzung des

Auswahlaxioms kann man eine Menge $A \subset [0,1]$ konstruieren, die mit jeder Äquivalenzklasse genau ein Element gemeinsam hat. Daher ist $\mathbb{R}$ gleich der abzählbaren disjunkten Vereinigung $\mathbb{R} = \bigcup_{q\in\mathbb{Q}}(q + A)$, wobei $q + A$ die um q verschobene Menge A ist. Wegen der abzählbaren Additivität ist also

$$\infty = \lambda(\mathbb{R}) = \sum_{q\in\mathbb{Q}} \lambda(q + A) = \sum_{q\in\mathbb{Q}} \lambda(A),$$

wobei wir die Translationsinvarianz von λ benutzt haben. Nun zeigen wir aber, dass $\lambda(A) = 0$ gilt, womit ein Widerspruch hergestellt ist. Dies geht so. Für die Menge

$$C = \bigcup_{q\in\mathbb{Q}\cap[0,1]} (q + A) \subset [0,2]$$

gilt

$$2 = \lambda([0,2]) \geq \lambda(C) = \sum_{q\in\mathbb{Q}\cap[0,1]} \lambda(q + A) = \sum_{q\in\mathbb{Q}\cap[0,1]} \lambda(A).$$

Daraus folgt $\lambda(A) = 0$, und der Widerspruch ist komplett. $\diamond$

2.3 Konstruktion von Maßen

Wie wir schon in Beispiel 2.2.8 gesehen haben, ist es eine wichtige Aufgabe, die Existenz und Eindeutigkeit von Maßen zu sichern, die auf gewissen Teilsystemen bestimmte Eigenschaften aufweisen. Das Hauptmittel ist hierbei der *Satz von Carathéodory* (siehe Satz 2.3.7), der besagt, dass Mengenfunktionen, die auf geeigneten Teilsystemen gewisse Struktureigenschaften haben, auf die davon erzeugte σ-Algebra erweitert werden können. Betrachten wir zunächst geeignete Teilmengensysteme:

Definition 2.3.1 (durchschnittstabil, Dynkin-System). *Es sei $\mathcal{C} \subset \mathcal{P}(\Omega)$ ein nichtleeres Mengensystem.*

(a) $\mathcal{C}$ heißt durchschnittstabil, *falls für je zwei Mengen $A, B \in \mathcal{C}$ auch $A \cap B \in \mathcal{C}$ gilt.*

(b) $\mathcal{C}$ heißt ein Dynkin-System, *falls gelten:*

(i) $\Omega \in \mathcal{C}$,

(ii) falls $A \in \mathcal{C}$, so gilt auch $A^{\mathrm{c}} \in \mathcal{C}$,

(iii) falls $A_1, A_2, A_3, \dots \in \mathcal{C}$ paarweise disjunkt sind, so gilt auch $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{C}$.

Ein Dynkin-System unterscheidet sich also nur darin von einer σ-Algebra, dass die Abgeschlossenheit gegen abzählbare Vereinigung nur für Folgen *paarweise disjunkter* Teilmengen gefordert wird. Insbesondere ist jede σ-Algebra ein Dynkin-System, aber nicht umgekehrt.[6] Dieser kleine Unterschied macht es in vielen Fällen viel einfacher, die Eigenschaft eines Dynkin-Systems zu verifizieren als die einer σ-Algebra. Der Zusammenhang ist allerdings sehr eng:

Lemma 2.3.2. *Jedes durchschnittstabile Dynkin-System ist eine σ-Algebra.*

[6] Falls Ω eine (endliche) gerade Anzahl > 2 von Elementen enthält, ist das System der Teilmengen mit gerader Elementanzahl ein Dynkin-System, aber keine σ-Algebra.

Beweis. Es sei $\mathcal{C}$ ein durchschnittstabiles Dynkin-System. Seien $A_1, A_2, A_3, \dots \in \mathcal{C}$, dann müssen wir zeigen, dass auch $A = \bigcup_{n\in\mathbb{N}} A_n$ in $\mathcal{C}$ liegt.

Mit $A_0 = \emptyset$ betrachten wir die Mengen $B_n = A_n \setminus (A_1 \cup \dots \cup A_{n-1})$, die ja offensichtlich paarweise disjunkt sind. Offensichtlich ist die Vereinigung der B_n gleich der Vereinigung der A_n. Nun zeigen wir mit einer Vollständigen Induktion nach n, dass B_n und $A_1 \cup \dots \cup A_n$ zu $\mathcal{C}$ gehören, was also den Beweis beendet.

Für $n = 1$ ist nichts zu zeigen. Für $n \geq 2$ gilt

$$B_n = A_n \cap \big(A_1 \cup \dots \cup A_{n-1}\big)^{\mathrm{c}}.$$

Nach Induktionsvoraussetzung ist $A_1 \cup \dots \cup A_{n-1}$ in $\mathcal{C}$, also auch sein Komplement, also auch der Schnitt mit A_n, das heißt B_n liegt in $\mathcal{C}$. Ferner ist

$$A_1 \cup \dots \cup A_n = (A_1 \cup \dots \cup A_{n-1}) \cup B_n$$

eine disjunkte Vereinigung von zwei Mengen, die in $\mathcal{C}$ liegen, also liegt auch $A_1 \cup \dots \cup A_n$ in $\mathcal{C}$. Damit ist die Induktion beendet und damit auch der Beweis des Lemmas. □

Analog zu Lemma 2.2.3 erzeugt jedes Mengensystem $\mathcal{C}$ auf Ω auch ein Dynkin-System, das wir mit $d(\mathcal{C})$ bezeichnen wollen, also

$$d(\mathcal{C}) = \bigcap\{\mathcal{A} \subset \mathcal{P}(\Omega) \colon \mathcal{A} \text{ ist ein Dynkin-System mit } \mathcal{C} \subset \mathcal{A}\}. \tag{2.3.1}$$

Ein weiterer wichtiger Zusammenhang ist der folgende.

Satz 2.3.3. *Für jedes durchschnittstabile Mengensystem $\mathcal{C}$ ist $d(\mathcal{C}) = \sigma(\mathcal{C})$.*

Beweis. Da jede σ-Algebra ein Dynkin-System ist, ist klar, dass $d(\mathcal{C}) \subset \sigma(\mathcal{C})$. Wenn wir zeigen, dass $d(\mathcal{C})$ eine σ-Algebra ist, folgt also Gleichheit. Nach Lemma 2.3.2 reicht es zu zeigen, dass $d(\mathcal{C})$ durchschnittstabil ist. Dazu betrachten wir das Mengensystem

$$\mathcal{A} = \{A \subset \Omega \colon A \cap C \in d(\mathcal{C}) \text{ für alle } C \in \mathcal{C}\}.$$

Da $\mathcal{C}$ durchschnittstabil ist, gilt $\mathcal{C} \subset \mathcal{A}$. Nun zeigen wir, dass $\mathcal{A}$ ein Dynkin-System ist, indem wir die drei definierenden Eigenschaften nachweisen. Die erste ist klar. Für jedes $A \in \mathcal{A}$ ist für jedes $C \in \mathcal{C}$ auch die Menge $A^{\mathrm{c}} \cap C = (C^{\mathrm{c}} \cup (A \cap C))^{\mathrm{c}}$ in $d(\mathcal{C})$, wie man leicht sieht, also liegt auch A^{c} in $\mathcal{A}$. Und wenn $A_1, A_2, \dots \in \mathcal{A}$ paarweise disjunkt sind, so liegt für jedes $C \in \mathcal{C}$ auch die Menge $(\bigcup_{n\in\mathbb{N}} A_n) \cap C = \bigcup_{n\in\mathbb{N}}(A_n \cap C)$ in $d(\mathcal{C})$, also $\bigcup_{n\in\mathbb{N}} A_n$ in $\mathcal{A}$. Also ist $\mathcal{A}$ ein Dynkin-System. Insbesondere gilt $d(\mathcal{C}) \subset \mathcal{A}$.

Nun betrachten wir das System

$$\widetilde{\mathcal{A}} = \{A \subset \Omega \colon A \cap A' \in d(\mathcal{C}) \text{ für alle } A' \in d(\mathcal{C})\}.$$

Wegen $d(\mathcal{C}) \subset \mathcal{A}$ gilt $\mathcal{C} \subset \widetilde{\mathcal{A}}$. Wie bei $\mathcal{A}$ zeigt man, dass $\widetilde{\mathcal{A}}$ ebenfalls ein Dynkin-System ist. Also folgt auch $d(\mathcal{C}) \subset \widetilde{\mathcal{A}}$. Dies ist aber äquivalent dazu, dass $d(\mathcal{C})$ durchschnittstabil ist. □

Unser Ziel in diesem Abschnitt ist, Mengenfunktionen auf gewissen Mengensystemen zu Maßen auf σ-Algebren zu erweitern. Dazu brauchen wir zunächst geeignete Mengensysteme, die einfacher zu behandeln sind als σ-Algebren:

Definition 2.3.4 (Algebra). *Ein nichtleeres Mengensystem $\mathcal{A} \subset \mathcal{P}(\Omega)$ heißt eine* Algebra*, falls für alle $A, B \in \mathcal{A}$ auch A^c und $A \cap B$ und $A \cup B$ in $\mathcal{A}$ liegen.*

Einer Algebra fehlt also nur die Abgeschlossenheit gegen *abzählbare* Vereinigungsbildung zu einer σ-Algebra.[7] Die Bedeutung des Begriffes der Algebra liegt darin, dass Algebren (im Gegensatz zu σ-Algebren) oft sehr explizit angegeben werden können. Zum Beispiel ist das System aller endlichen Vereinigungen von Intervallen eine Algebra über $\mathbb{R}$. Ein zweidimensionales Analogon ist das Mengensystem $\mathcal{E}_E$ der elementaren Teilmengen des Einheitsquadrates $E = [0,1]^2$, das wir in Abschnitt 2.1 einführten.

Analog zur Algebra gibt es auch eine nur *endlich* additive Version von Maßen:

Definition 2.3.5 (Inhalt, Prämaß). *(a) Ein* Inhalt *μ auf einer Algebra $\mathcal{A}$ ist eine Abbildung $\mu \colon \mathcal{A} \to [0, \infty]$ mit den Eigenschaften $\mu(\emptyset) = 0$ und $\mu(A \cup B) = \mu(A) + \mu(B)$ für alle disjunkten Mengen $A, B \in \mathcal{A}$.*

(b) Ein σ-additiver Inhalt heißt ein Prämaß.

Einem Inhalt fehlt also nur die *abzählbare* Additivität zu einem Maß, vorausgesetzt, dass sein Definitionsbereich eine σ-Algebra ist. Wenn der Definitionsbereich eines Prämaßes nicht nur eine Algebra, sondern sogar eine σ-Algebra ist, so ist es schon ein Maß. Es ist klar, dass jeder Inhalt μ *monoton* ist (d. h. $\mu(A) \leq \mu(B)$ für alle $A, B \in \mathcal{A}$ mit $A \subset B$), *endlich additiv* (d. h. $\mu(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n \mu(A_i)$ für paarweise disjunkte $A_1, \dots, A_n \in \mathcal{A}$) und *endlich subadditiv* (d. h. $\mu(\bigcup_{i=1}^n A_i) \leq \sum_{i=1}^n \mu(A_i)$ für beliebige $A_1, \dots, A_n \in \mathcal{A}$). Die Fortsetzung von λ_E auf die Algebra $\mathcal{E}_E$ der elementaren Teilmengen von $E = [0,1]^2$ in Abschnitt 2.1 erwies sich in Lemma 2.1.2 als ein Prämaß.

Die σ-Additivität ist sogar äquivalent zur Stetigkeit in $\emptyset$:

Lemma 2.3.6. *Es sei μ ein endlicher Inhalt auf einer Algebra $\mathcal{A}$. Dann sind äquivalent:*

(i) μ ist σ-additiv (d. h. ein Prämaß),

(ii) μ ist stetig in der leeren Menge, d. h. für alle Folgen von Mengen $A_1, A_2, \dots \in \mathcal{A}$ mit $A_n \downarrow \emptyset$ gilt $\lim_{n\to\infty} \mu(A_n) = 0$.

Beweis. (i)$\Longrightarrow$(ii): Sei $A_n \downarrow \emptyset$. Wir definieren $B_n = A_n \setminus A_{n+1}$ für $n \in \mathbb{N}$. Dann sind die Mengen B_n paarweise disjunkt, und es gilt $A_n = \bigcup_{m=n}^\infty B_m$ für alle $n \in \mathbb{N}$. Also ist $\mu(A_n) = \sum_{m=n}^\infty \mu(B_m)$. Da diese Summe konvergiert, gilt $\lim_{n\to\infty} \mu(A_n) = 0$.

(ii)$\Longrightarrow$(i): Es sei $(B_n)_{n\in\mathbb{N}}$ eine Folge von paarweise disjunkten Mengen in $\mathcal{A}$, so dass $B = \bigcup_{n\in\mathbb{N}} B_n$ in $\mathcal{A}$ liegt. Für $n \in \mathbb{N}$ betrachte $A_{n+1} = \bigcup_{m=n+1}^\infty B_m = B \setminus (B_1 \cup \dots \cup B_n) \in \mathcal{A}$. Dann sind die Mengen $B_1, \dots, B_n, A_{n+1}$ paarweise disjunkt und in $\mathcal{A}$. Wegen der endlichen Additivität gilt

$$\mu(B) = \sum_{m=1}^n \mu(B_m) + \mu(A_{n+1}).$$

Wegen $A_{n+1} \downarrow \emptyset$ gilt $\lim_{n\to\infty} \mu(A_{n+1}) = 0$. Also folgt $\mu(B) = \sum_{m=1}^\infty \mu(B_m)$, d. h. μ ist σ-additiv. □

Nun kommen wir endlich zum Hauptergebnis dieses Abschnittes:

[7] Zum Beispiel ist das System der Teilmengen A von Ω, so dass A oder A^c endlich sind, eine Algebra, aber für unendliches Ω keine σ-Algebra.

Satz 2.3.7 (Carathéodory). *Es sei μ ein σ-endliches Prämaß auf einer Algebra $\mathcal{A}$. Dann gibt es genau ein Maß $\widetilde{\mu}$ auf $\sigma(\mathcal{A})$, das μ erweitert, d. h. das auf $\mathcal{A}$ mit μ übereinstimmt.*

Der Existenzteil dieses Satzes wird am Ende dieses Abschnittes bewiesen werden. Wir isolieren zunächst die Eindeutigkeit, die direkt aus dem folgenden Resultat folgt, das wir noch oft benutzen werden.

Satz 2.3.8 (Eindeutigkeitssatz). *Es sei $\mathcal{F}$ eine σ-Algebra auf Ω und $\mathcal{C}$ ein durchschnittstabiles Erzeugendensystem von $\mathcal{F}$. Falls zwei auf $\mathcal{F}$ definierte Maße μ, ν auf $\mathcal{C}$ mit einander übereinstimmen und eine Folge $(\Omega_n)_{n\in\mathbb{N}}$ in $\mathcal{C}$ existiert mit $\Omega_n \uparrow \Omega$ und $\mu(\Omega_n) = \nu(\Omega_n) < \infty$ für jedes $n \in \mathbb{N}$, so gilt $\mu = \nu$ auf $\mathcal{F}$.*

Beweis. Zunächst behandeln wir den Spezialfall, wo $\mu(\Omega) = \nu(\Omega) < \infty$. Wir betrachten das Mengensystem

$$\widetilde{\mathcal{F}} = \{A \in \mathcal{F} : \mu(A) = \nu(A)\}$$

und brauchen nur zu zeigen, dass $\mathcal{F} \subset \widetilde{\mathcal{F}}$. Hierzu reicht es zu zeigen, dass $\widetilde{\mathcal{F}}$ ein Dynkin-System ist, denn dann gilt nach Satz 2.3.3 schon $\mathcal{F} = \sigma(\mathcal{C}) = d(\mathcal{C}) \subset \widetilde{\mathcal{F}}$, und $\widetilde{\mathcal{F}}$ enthält offensichtlich $\mathcal{C}$. Dass $\widetilde{\mathcal{F}}$ ein Dynkin-System ist, rechnet man leicht an Hand der Definition nach: Offensichtlich ist $\Omega \in \widetilde{\mathcal{F}}$, und für jedes $D \in \widetilde{\mathcal{F}}$ gilt $\mu(D^{\mathrm{c}}) = \mu(\Omega) - \mu(D) = \nu(\Omega) - \nu(D) = \nu(D^{\mathrm{c}})$, also liegt auch D^{c} in $\widetilde{\mathcal{F}}$. Und für jede Folge $(D_n)_{n\in\mathbb{N}}$ von paarweise disjunkten Mengen in $\widetilde{\mathcal{F}}$ gilt

$$\mu\Big(\bigcup_{n\in\mathbb{N}} D_n\Big) = \sum_{n\in\mathbb{N}} \mu(D_n) = \sum_{n\in\mathbb{N}} \nu(D_n) = \nu\Big(\bigcup_{n\in\mathbb{N}} D_n\Big),$$

also liegt $\bigcup_{n\in\mathbb{N}} D_n$ auch in $\widetilde{\mathcal{F}}$. Damit ist der Spezialfall $\mu(\Omega) = \nu(\Omega) < \infty$ bewiesen.

Im allgemeinen Fall betrachten wir für $n \in \mathbb{N}$ die endlichen Maße μ_n und ν_n, die definiert sind durch $\mu_n(A) = \mu(A \cap \Omega_n)$ bzw. $\nu_n(A) = \nu(A \cap \Omega_n)$ für alle $A \in \mathcal{F}$. Auf μ_n und ν_n können wir die oben bewiesene Aussage anwenden und erhalten, dass $\mu_n = \nu_n$ für jedes $n \in \mathbb{N}$. Durch Grenzübergang $n \to \infty$ erhält man $\mu(A) = \nu(A)$ für jedes $A \in \mathcal{F}$, also stimmen μ und ν überein. □

Korollar 2.3.9. *Falls zwei endliche Maße μ, ν mit $\mu(\Omega) = \nu(\Omega)$ auf einem durchschnittstabilen Erzeugendensystem mit einander überein stimmen, so sind sie schon gleich.*

Beispiel 2.3.10. Wir betrachten das Mengensystem $\mathcal{C} = \{\{1,2\},\{2,3\}\}$ auf $\Omega = \{1,2,3,4\}$, dann kann man leicht sehen, dass $\sigma(\mathcal{C}) = \mathcal{P}(\Omega)$ gilt. Aber $\mathcal{C}$ ist nicht durchschnittstabil. Tatächlich stimmen die beiden Wahrscheinlichkeitsmaße $\frac{1}{2}\delta_1 + \frac{1}{2}\delta_3$ und $\frac{1}{2}\delta_2 + \frac{1}{2}\delta_4$ auf $\mathcal{C}$ mit einander überein, d. h., die Werte auf $\mathcal{C}$ legen das Wahrscheinlichkeitsmaß nicht eindeutig fest.

Dass die Bedingung $\mu(\Omega) = \nu(\Omega)$ nicht weggelassen werden kann, sieht man an dem Beispiel $\Omega = \{1,2\}$ und $\mathcal{C} = \{1\}$ und $\nu = \delta_1$ und $\mu = \delta_1 + \delta_2$. ◇

Nun kommen wir zum Beweis der Existenzaussage in Satz 2.3.7. Wir zerlegen ihn in mehrere Schritte, die für sich genommen ebenfalls interessant sind. Ein Überblick ist wie folgt: Zu dem Prämaß μ definiert man auf der Potenzmenge von Ω ein äußeres Mass μ^* und den Begriff der

Messbarkeit bezüglich μ^*. Die Menge aller μ^*-messbaren Mengen stellt sich als eine σ-Algebra heraus die $\sigma(\mathcal{A})$ enthält, und μ^* als ein Maß darauf. Daraus folgt, dass die Einschränkung von μ^* auf $\sigma(\mathcal{A})$ das gesuchte Maß $\widetilde{\mu}$ ist.

Definition 2.3.11 (Äußeres Maß). *Eine Mengenfunktion $\mu^*\colon \mathcal{P}(\Omega) \to [0,\infty]$ heißt ein* äußeres Maß*, falls gelten:*

$$\textit{(i) } \mu^*(\emptyset) = 0, \qquad \textit{(ii) } \mu^* \textit{ ist monoton}, \qquad \textit{(iii) } \mu^* \textit{ ist } \sigma\textit{-subadditiv}. \tag{2.3.2}$$

Sehr viele Mengenfunktionen können zu einem äußeren Maß fortgesetzt werden (siehe auch das Beispiel des Lebesgue-Maßes in (2.1.5)):

Lemma 2.3.12. *Sei $\mathcal{C} \subset \mathcal{P}(\Omega)$ ein beliebiges Mengensystem mit $\emptyset \in \mathcal{C}$ und $\mu\colon \mathcal{C} \to [0,\infty]$ eine monotone Mengenfunktion mit $\mu(\emptyset) = 0$. Definiere $\mu^*\colon \mathcal{P}(\Omega) \to [0,\infty]$ durch*

$$\mu^*(A) = \inf\Big\{\sum_{n\in\mathbb{N}} \mu(A_n)\colon A_1, A_2, \dots \in \mathcal{C}, A \subset \bigcup_{n\in\mathbb{N}} A_n\Big\},$$

wobei $\inf\emptyset = \infty$*. Dann ist μ^* ein äußeres Maß, das μ fortsetzt (d. h., $\mu^*(A) = \mu(A)$ für jedes $A \in \mathcal{C}$).*

Beweis. Wir prüfen die drei Eigenschaften aus (2.3.2). (i) ist klar. Sei $A \subset B$, und sei $(B_n)_{n\in\mathbb{N}}$ eine Familie in $\mathcal{C}$ mit $B \subset \bigcup_{n\in\mathbb{N}} B_n$. Dann ist $\mu^*(A) \le \sum_{n\in\mathbb{N}} \mu(B_n)$, denn die B_n überdecken ja auch A. Der Übergang zum Infimum über alle $(B_n)_{n\in\mathbb{N}}$ lässt folgen, dass $\mu^*(A) \le \mu^*(B)$, also ist μ^* monoton.

Nun zeigen wir (iii). Sei $(A_n)_{n\in\mathbb{N}}$ eine Familie in $\mathcal{P}(\Omega)$ und $A \subset \bigcup_{n\in\mathbb{N}} A_n$. Wir müssen zeigen, dass $\mu^*(A) \le \sum_{n\in\mathbb{N}} \mu(A_n)$. Ohne Einschränkung kann man annehmen, dass $\mu(A_n) < \infty$ für jedes $n \in \mathbb{N}$. Sei $\varepsilon > 0$. Wir wählen für jedes $n \in \mathbb{N}$ eine Überdeckung $(B_{n,i})_{i\in\mathbb{N}}$ von A_n mit $\sum_{i\in\mathbb{N}} \mu(B_{n,i}) \le \mu^*(A_n) + \varepsilon 2^{-n}$. Dann ist die Vereinigung der Mengen $B_{n,i}$ mit $n, i \in \mathbb{N}$ eine abzählbare Überdeckung von A, und wir haben

$$\mu^*(A) \le \sum_{i,n\in\mathbb{N}} \mu(B_{n,i}) \le \sum_{n\in\mathbb{N}} \Big(\mu^*(A_n) + \varepsilon 2^{-n}\Big) = \varepsilon + \sum_{n\in\mathbb{N}} \mu^*(A_n).$$

Da $\varepsilon > 0$ beliebig ist, ist das Ziel erreicht. □

Nun definieren wir Messbarkeit bezüglich eines äußeren Maßes.

Definition 2.3.13 (μ^*-messbare Mengen). *Sei μ^* ein äußeres Maß. Eine Menge $A \in \mathcal{P}(\Omega)$ heißt* μ^*-messbar*, falls*

$$\mu^*(A \cap E) + \mu^*(A^{\mathrm{c}} \cap E) = \mu^*(E), \qquad E \in \mathcal{P}(\Omega).$$

Mit $\mathcal{M}(\mu^*)$ bezeichnen wir die Menge aller μ^*-messbaren Mengen.

Zum Zusammenhang zwischen diesem Messbarkeitsbegriff mit dem Lebesgue'schen aus Definition 2.1.5 siehe Bemerkung 2.3.16 und Satz 2.3.17.

Bemerkung 2.3.14. Man sieht leicht, dass eine Menge $A \in \mathcal{P}(\Omega)$ genau dann μ^*-messbar ist, wenn

$$\mu^*(A \cap E) + \mu^*(A^{\mathrm{c}} \cap E) \le \mu^*(E), \qquad E \in \mathcal{P}(\Omega),$$

denn wegen Subadditivität gilt ja '$\ge$' sowieso immer. ◇

Die folgende Aussage ist die abstrakte Form des Lemmas 2.1.6.

Lemma 2.3.15. *Wenn μ^* ein äußeres Maß ist, so ist $\mathcal{M}(\mu^*)$ eine σ-Algebra und die Einschränkung von μ^* auf $\mathcal{M}(\mu^*)$ ein Maß.*

Beweis. Zunächst prüfen wir nach, dass $\mathcal{M}(\mu^*)$ eine Algebra ist. Es ist klar, dass $\Omega \in \mathcal{M}(\mu^*)$. Dass Komplemente von Mengen aus $\mathcal{M}(\mu^*)$ auch in $\mathcal{M}(\mu^*)$ liegen, ersieht man sofort aus der Definition. Dass für $A, B \in \mathcal{M}(\mu^*)$ auch $A \cap B$ in $\mathcal{M}(\mu^*)$ liegt, sieht man mit Hilfe der Bemerkung 2.3.14 daraus, dass für jedes $E \in \mathcal{P}(\Omega)$ gilt:

$$\begin{aligned}
&\mu^*\big((A \cap B) \cap E\big) + \mu^*\big((A \cap B)^{\mathrm{c}} \cap E\big) \\
&\qquad = \mu^*(A \cap B \cap E) + \mu^*\big((A^{\mathrm{c}} \cap B \cap E) \cup (A^{\mathrm{c}} \cap B^{\mathrm{c}} \cap E) \cup (A \cap B^{\mathrm{c}} \cap E)\big) \\
&\qquad \le \mu^*(A \cap B \cap E) + \mu^*(A^{\mathrm{c}} \cap B \cap E) + \mu^*(A^{\mathrm{c}} \cap B^{\mathrm{c}} \cap E) + \mu^*(A \cap B^{\mathrm{c}} \cap E) \\
&\qquad = \mu^*(B \cap E) + \mu^*(B^{\mathrm{c}} \cap E) = \mu^*(E).
\end{aligned}$$

Als Nächstes zeigen wir, dass μ^* auf $\mathcal{M}(\mu^*)$ σ-additiv ist. Die endliche Additivität folgt daraus, dass für disjunkte Mengen $A, B \in \mathcal{M}(\mu^*)$ gilt:

$$\mu^*(A \cup B) = \mu^*(A \cap (A \cup B)) + \mu^*(A^{\mathrm{c}} \cap (A \cup B)) = \mu^*(A) + \mu^*(B).$$

Daher folgt für jede Familie $(A_n)_{n \in \mathbb{N}}$ paarweise disjunkter Mengen $A_n \in \mathcal{P}(\Omega)$ und jedes $N \in \mathbb{N}$:

$$\mu^*\Big(\bigcup_{n \in \mathbb{N}} A_n\Big) \ge \mu^*\Big(\bigcup_{n=1}^{N} A_n\Big) = \sum_{n=1}^{N} \mu^*(A_n).$$

Indem man $N \to \infty$ lässt, sieht man, dass $\mu^*(\bigcup_{n \in \mathbb{N}} A_n) \ge \sum_{n \in \mathbb{N}} \mu^*(A_n)$ gilt. Die umgekehrte Ungleichung gilt, da μ^* ein äußeres Maß ist.

Zuletzt müssen wir noch zeigen, dass $\mathcal{M}(\mu^*)$ eine σ-Algebra ist. Da $\mathcal{M}(\mu^*)$ eine Algebra ist, also insbesondere durchschnittstabil, müssen wir nach Lemma 2.3.3 nur noch zeigen, dass $\mathcal{M}(\mu^*)$ ein Dynkin-System ist. Sei also $(A_n)_{n \in \mathbb{N}}$ eine Familie paarweise disjunkter Mengen $A_n \in \mathcal{P}(\Omega)$, dann müssen wir zeigen, dass $A = \bigcup_{n \in \mathbb{N}} A_n$ in $\mathcal{M}(\mu^*)$ liegt, also dass gilt:

$$\mu^*(A \cap E) + \mu^*(A^{\mathrm{c}} \cap E) = \mu^*(E), \qquad E \in \mathcal{P}(\Omega).$$

Sei $E \in \mathcal{P}(\Omega)$. Mit der Abkürzung $B_N = \bigcup_{n=1}^{N} A_n$ gilt für jedes $N \in \mathbb{N}$:

$$\begin{aligned}
\mu^*(E \cap B_{N+1}) &= \mu^*\big((E \cap B_{N+1}) \cap B_N\big) + \mu^*\big((E \cap B_{N+1}) \cap B_N^{\mathrm{c}}\big) \\
&= \mu^*(E \cap B_N) + \mu^*(E \cap A_{N+1}).
\end{aligned}$$

Per Induktion leitet man daraus her, dass $\mu^*(E \cap B_N) = \sum_{i=1}^{N} \mu^*(E \cap A_i)$ für jedes $N \in \mathbb{N}$ gilt. Wegen der Monotonie von μ^* folgt

$$\mu^*(E) = \mu^*(E \cap B_N) + \mu^*(E \cap B_N^{\mathrm{c}}) \ge \mu^*(E \cap B_N) + \mu^*(E \cap A^{\mathrm{c}}) = \sum_{i=1}^{N} \mu^*(E \cap A_i) + \mu^*(E \cap A^{\mathrm{c}}).$$

Indem wir $N \to \infty$ gehen lassen, erhalten wir mit der σ-Subadditivität von μ^*, dass gilt:

$$\mu^*(E) \ge \sum_{i=1}^{\infty} \mu^*(E \cap A_i) + \mu^*(E \cap A^{\mathrm{c}}) \ge \mu^*(E \cap A) + \mu^*(E \cap A^{\mathrm{c}}).$$

Also ist das Ziel erreicht. $\square$

Beweis der Existenzaussage in Satz 2.3.7 von Carathéodory. Es sei das äußere Maß μ^* wie in Lemma 2.3.12 mit $\mathcal{C} = \mathcal{A}$ definiert. Aus Lemma 2.3.15 wissen wir, dass μ^* ein Maß auf der σ-Algebra $\mathcal{M}(\mu^*)$ ist, das μ fortsetzt. Also reicht es zu zeigen, dass $\sigma(\mathcal{A})$ in $\mathcal{M}(\mu^*)$ enthalten ist, denn dann wäre die Einschränkung $\widetilde{\mu}$ von μ^* auf $\sigma(\mathcal{A})$ ein Maß, das die geforderten Eigenschaften hat. Da $\mathcal{M}(\mu^*)$ eine σ-Algebra ist, reicht es zu zeigen, dass $\mathcal{A}$ in $\mathcal{M}(\mu^*)$ enthalten ist, und dies machen wir nun.

Sei also $A \in \mathcal{A}$, und sei $E \in \mathcal{P}(\Omega)$ mit $\mu^*(E) < \infty$. Sei $\varepsilon > 0$, dann gibt es in $\mathcal{A}$ eine Familie $(E_n)_{n\in\mathbb{N}}$, die E überdeckt und $\sum_{n\in\mathbb{N}} \mu(E_n) \leq \mu^*(E) + \varepsilon$ erfüllt. Wir setzen $B_n = E_n \cap A \in \mathcal{A}$. Somit wird $E \cap A$ von der Vereinigung der B_n überdeckt und $E \cap A^{\mathrm{c}}$ von der Vereinigung der Mengen $E_n \setminus A = E_n \setminus B_n$. Da μ^* abzählbar subadditiv und μ additiv sind und da $\mu^* \leq \mu$ auf $\mathcal{A}$ gilt, haben wir

$$\begin{aligned}
\mu^*(E \cap A) + \mu^*(E \cap A^{\mathrm{c}}) &\leq \mu^*\Big(\bigcup_{n\in\mathbb{N}} B_n\Big) + \mu^*\Big(\bigcup_{n\in\mathbb{N}} (E_n \setminus B_n)\Big) \\
&\leq \sum_{n\in\mathbb{N}} \mu^*(B_n) + \sum_{n\in\mathbb{N}} \mu^*(E_n \setminus B_n) \\
&\leq \sum_{n\in\mathbb{N}} \big(\mu(B_n) + \mu(E_n \setminus B_n)\big) \\
&= \sum_{n\in\mathbb{N}} \mu(E_n) \leq \mu^*(E) + \varepsilon.
\end{aligned}$$

Also haben wir $\mu^*(E \cap A) + \mu^*(E \cap A^{\mathrm{c}}) \leq \mu^*(E)$ für jedes $E \in \mathcal{P}(\Omega)$ und damit $A \in \mathcal{M}(\mu^*)$. Dies beendet den Beweis. $\square$

Bemerkung 2.3.16 (Zusammenhang zwischen μ^*- und Lebesgue-Messbarkeit). Als Übungsaufgabe zeigt man, dass jede im Sinn der Definition 2.1.5 Lebesgue-messbare Menge auch im Sinn der Definition 2.3.13 λ^*-messbar ist. Die umgekehrte Richtung folgt aus Satz 2.3.17. Dies zeigt, dass die Herangehensweise von Abschnitt 2.1 tatsächlich ein Spezialfall der allgemeinen Theorie in diesem Abschnitt ist. $\diamond$

Satz 2.3.17 (Approximationssatz). *Es sei $\mathcal{A}$ eine Algebra und μ ein σ-endliches Maß auf der σ-Algebra $\sigma(\mathcal{A})$. Dann gibt es zu jedem $A \in \sigma(\mathcal{A})$ mit $\mu(A) < \infty$ und jedem $\varepsilon > 0$ ein $B \in \mathcal{A}$ mit $\mu(A \triangle B) < \varepsilon$.*

Beweis. Im Beweis der Existenzaussage im Satz 2.3.7 von Carathéodory sahen wir, dass das auf $\mathcal{M}(\mu^*)$ definierte äußere Maß μ^* auf der σ-Algebra $\sigma(\mathcal{A})$ (die ja in $\mathcal{M}(\mu^*)$ enthalten ist) mit μ übereinstimmt. Seien nun $A \in \sigma(\mathcal{A})$ mit $\mu(A) < \infty$ und $\varepsilon > 0$ gegeben, dann ist $\mu(A) = \mu^*(A)$. Nach Definition des Infimums gibt es eine Familie $(A_n)_{n\in\mathbb{N}}$ in $\mathcal{A}$ mit $A \subset \bigcup_{n\in\mathbb{N}} A_n$ und $\mu(A) \leq \sum_{n\in\mathbb{N}} \mu(A_n) \leq \mu(A) + \varepsilon/2$. Unter Ausnutzung der σ-Subadditivität von μ ergibt sich dann

$$\mu\Big(A \triangle \bigcup_{i=1}^{n} A_i\Big) \leq \mu\Big(\bigcup_{i=1}^{\infty} A_i \setminus A\Big) + \mu\Big(\bigcup_{i=n+1}^{\infty} A_i\Big) \leq \frac{\varepsilon}{2} + \mu\Big(\bigcup_{i=n+1}^{\infty} A_i\Big).$$

Mit $n \to \infty$ konvergiert der letzte Term gegen Null wegen der Stetigkeit des Maßes μ in der leeren Menge. Also ist $\mu(A \triangle \bigcup_{i=1}^{n} A_i) < \varepsilon$ für alle genügend großen n. Daher erfüllt $B = \bigcup_{i=1}^{n} A_i$ die Forderung, und dies beendet den Beweis. $\square$

2.4 Verteilungsfunktionen

In diesem Abschnitt wollen wir den Satz 2.3.7 von Carathéodory benutzen, um viele Maße auf der reellen Achse zu konstruieren. Wir erinnern daran (siehe Beispiel 2.2.5), dass die Borel-σ-Algebra $\mathcal{B}$ auf $\mathbb{R}$ insbesondere von den Mengen $(-\infty, t]$ mit $t \in \mathbb{R}$ erzeugt wird. Wir zeigen in diesem Abschnitt, dass Maße μ auf $(\mathbb{R}, \mathcal{B})$ eindeutig festgelegt werden durch die Abbildung $t \mapsto \mu((-\infty, t])$. Im Fall eines Wahrscheinlichkeitsmaßes μ nennt man diese Abbildung die *Verteilungsfunktion* von μ. Im Folgenden finden wir hinreichende und notwendige Eigenschaften von Funktionen $F\colon \mathbb{R} \to \mathbb{R}$, so dass es ein Maß μ gibt mit $F(t) = \mu((-\infty, t])$ für alle $t \in \mathbb{R}$. Zunächst eine einfache Beobachtung.

Lemma 2.4.1. *Für jedes Maß μ auf $(\mathbb{R}, \mathcal{B})$ mit $\mu((s,t]) < \infty$ für alle $s < t$ ist die Abbildung $t \mapsto \mu((s,t])$ monoton steigend und rechtsseitig stetig.*

Beweis. Die Monotonie ist klar, und die rechtsseitige Stetigkeit folgt aus der Darstellung $(-\infty, t] = \bigcap_{n\in\mathbb{N}}(-\infty, t + \frac{1}{n}]$ sowie der Stetigkeit des Maßes μ von oben, siehe Lemma 2.2.9(a). □

Tatsächlich erweisen sich diese beiden Eigenschaften schon als entscheidend:

Satz 2.4.2. *Zu jeder monoton steigenden, rechtsseitig stetigen Abbildung $F\colon \mathbb{R} \to \mathbb{R}$ gibt es genau ein Maß μ auf $(\mathbb{R}, \mathcal{B})$ mit $F(t) - F(s) = \mu((s,t])$ für alle $s, t \in \mathbb{R}$ mit $s < t$.*

Beweis. Es sei $\mathcal{J}$ die Menge aller Intervalle $(s,t]$ mit $-\infty \le s \le t < \infty$ und aller Intervalle (s, ∞) mit $-\infty \le s \le \infty$. Ferner sei $a(\mathcal{J})$ die von $\mathcal{J}$ erzeugte Algebra, das heißt die kleinste Algebra, die $\mathcal{J}$ enthält. Man kann leicht zeigen, dass $a(\mathcal{J})$ aus allen endlichen disjunkten Vereinigungen der Intervalle in $\mathcal{J}$ besteht. Wir definieren

$$\begin{aligned} \mu((s,t]) &= F(t) - F(s), & -\infty \le s \le t < \infty, \\ \mu((s,\infty)) &= F(\infty) - F(s), & -\infty \le s < \infty, \end{aligned}$$

wobei wir $F(\infty) = \lim_{t\to\infty} F(t) \in (\infty, \infty]$ und $F(-\infty) = \lim_{s\to-\infty} F(s) \in [-\infty, \infty)$ setzen. Natürlich können wir μ sofort leicht auf $a(\mathcal{J})$ erweitern: Eine disjunkte Vereinigung von Intervallen dieser Form erhält als μ-Wert natürlich die Summe der μ-Werte der Intervalle. Es ist evident, dass μ ein Inhalt auf $a(\mathcal{J})$ ist. Offensichtlich ist μ auch σ-endlich, denn $\mathbb{R} = \bigcup_{n\in\mathbb{N}}(-n, n]$, und $\mu((-n,n]) = F(n) - F(-n) < \infty$.

Nach dem Satz von Carathéodory genügt es zu zeigen, dass μ ein Prämaß auf $a(\mathcal{J})$ ist. Diese Eigenschaft zeigen wir nur für den Fall $-\infty < F(-\infty) < F(\infty) < \infty$, der allgemeine Fall ist eine Übungsaufgabe. Nach Lemma 2.3.6 müssen wir zeigen, dass $\mu(A_n) \to 0$ für jede Folge $(A_n)_{n\in\mathbb{N}}$ in $a(\mathcal{J})$ mit $A_n \downarrow \emptyset$ gilt. Dies beweisen wir indirekt: Sei $(A_n)_{n\in\mathbb{N}}$ eine fallende Folge in $a(\mathcal{J})$ mit $a = \lim_{n\to\infty} \mu(A_n) > 0$, dann werden wir zeigen, dass $\bigcap_{n\in\mathbb{N}} A_n \neq \emptyset$ gilt.

(i) Zunächst zeigen wir, dass für jedes Intervall $I \in \mathcal{J}$ und jedes $\varepsilon > 0$ ein Intervall $I' \in \mathcal{J}$ und ein kompaktes Intervall L existieren mit

$$I' \subset L \subset I \qquad \text{und} \qquad \mu(I') \ge \mu(I) - \varepsilon. \tag{2.4.1}$$

Im Fall $I = (s,t]$ mit $s, t \in \mathbb{R}$ wählen wir $I' = (s', t]$ und $L = [(s+s')/2, t]$, wobei $s' \in (s,t)$ so gewählt wurde, dass $F(s') \le F(s) + \varepsilon$ gilt. Dann ist offensichtlich (2.4.1) erfüllt. Im Fall

$I = (-\infty, t]$ wählen wir ein $s' \in (-\infty, t]$ mit $F(s') \leq F(-\infty) + \varepsilon$ und setzen $I' = (s', t]$ und $L = [s' - 1, t]$. Die Fälle $I = (s, \infty)$ und $I = \mathbb{R}$ gehen analog.

(ii) Jedes $A_n \in a(\mathcal{J})$ ist eine endliche Vereinigung von Intervallen in $\mathcal{J}$. Aus (i) folgt daher, dass wir $B_n \in a(\mathcal{J})$ und eine kompakte Menge $K_n \subset \mathbb{R}$ finden können mit $B_n \subset K_n \subset A_n$ und $\mu(B_n) \geq \mu(A_n) - a2^{-n-1}$.

(iii) Es gilt für jedes $n \in \mathbb{N}$

$$A_n \setminus (B_1 \cap \cdots \cap B_n) = (A_1 \cap \cdots \cap A_n) \setminus (B_1 \cap \cdots \cap B_n) \subset \bigcup_{i=1}^n (A_i \setminus B_i)$$

und daher auch

$$\begin{aligned}\mu(B_1 \cap \cdots \cap B_n) &\geq \mu(A_n) - \mu\Big(\bigcup_{i=1}^n (A_i \setminus B_i)\Big)\\ &\geq \mu(A_n) - \sum_{i=1}^n \mu(A_i \setminus B_i) \geq a - \sum_{i=1}^n a2^{-i-1} \geq a/2.\end{aligned}$$

Daher ist $B_1 \cap \cdots \cap B_n \neq \emptyset$ für jedes $n \in \mathbb{N}$. Wegen $B_n \subset K_n$ ist auch die Menge $\overline{K}_n = K_1 \cap \cdots \cap K_n$ nicht leer für jedes $n \in \mathbb{N}$. Da $(\overline{K}_n)_{n\in\mathbb{N}}$ eine absteigende Folge nichtleerer kompakter Mengen ist, ist ihr Schnitt $\bigcap_{n\in\mathbb{N}} \overline{K}_n$ nicht leer. Wegen $K_n \subset A_n$ ist auch $\bigcap_{n\in\mathbb{N}} A_n$ nicht leer, und das war zu zeigen. □

Nachzutragen ist noch, dass μ im Falle $F(\infty) - F(-\infty) = 1$ ein *Wahrscheinlichkeits*maß ist. In diesem Fall kann man ohne Einschränkung voraus setzen, dass $F(-\infty) = 0$ gilt, und man nennt F die *Verteilungsfunktion* von μ. Wir haben dann $\mu((-\infty, t]) = F(t)$ für jedes $t \in \mathbb{R}$. Ferner haben wir endlich einen Existenz- und Eindeutigkeitsbeweis für das Lebesgue-Borel-Maß, siehe Beispiel 2.2.8:

Korollar 2.4.3 (Lebesgue-Borel-Maß). *Es gibt genau ein Maß λ auf $(\mathbb{R}, \mathcal{B})$ mit $\lambda((s,t]) = t - s$ für alle $s < t$. Dieses Maß λ heißt das* Lebesgue-Borel-Maß *auf $\mathbb{R}$.*

Beweis. Wir wenden Satz 2.4.2 auf $F(t) = t$ an. □

Das d-dimensionale Lebesgue-Borel-Maß kann man aus einer d-dimensionalen Version von Satz 2.4.2 erhalten, die nicht wirklich schwieriger zu beweisen ist als die eindimensionale Version. Wir werden in Beispiel 3.8.7 das d-dimensionale Lebesgue-Borel-Maß als ein Produktmaß erhalten.

Korollar 2.4.4 (stetige Dichten). *Es sei $f\colon \mathbb{R} \to [0, \infty)$ eine stetige nichtnegative Funktion. Dann existiert genau ein Maß μ_f auf $(\mathbb{R}, \mathcal{B})$ mit $\mu_f((a,b]) = \int_a^b f(x)\,\mathrm{d}x$ für alle $a < b$. (Hierbei sei $\int_a^b f(x)\,\mathrm{d}x$ ein Riemann-Integral.) Wir sagen, μ_f habe die* Dichte f.

Beweis. Wir wenden Satz 2.4.2 auf $F(t) = \int_0^t f(x)\,\mathrm{d}x$ an. □

Das Beispiel aus Korollar 2.4.4 kann man stark verallgemeinern; siehe Abschnitt 3.6.

Kapitel 3

Integration messbarer Funktionen

Nachdem wir nun das Messen von Mengen formalisiert haben, wollen wir auch begrifflich fassen, was die Integration von Funktionen sein soll. In Abschnitt 3.1 stellen wir die Funktionen vor, die mit der messbaren Struktur kompatibel sind, die messbaren Funktionen. In Abschnitt 3.2 erklären wir, was das Integral einer messbaren Funktion bezüglich eines Maßes sein soll. Den Zusammenhang zwischen dem Riemann- und dem Lebesgue-Integral untersuchen wir in Abschnitt 3.3. Die fundamentalen Grenzwertsätze zur Vertauschung von Grenzwert und Integral sind der Gegenstand des Abschnitts 3.4. Die wichtigen $\mathcal{L}^p$-Räume und ihre Eigenschaften sind der Inhalt des Abschnitts 3.5, und die Integration bezüglich Maße, die durch messbare Abbildungen transportiert bzw. transformiert werden, werden in Abschnitt 3.6 behandelt, wo wir auch den allgemeinen Transformationssatz für Lebesgue-Integrale nebst Anwendungen betrachten. Die abschließenden Abschnitte 3.7 und 3.8 behandeln Produkträume und -maße und Integration auf ihnen.

3.1 Messbare Abbildungen

Bei Abbildungen zwischen messbaren Räumen muss die Kompatibilität mit den jeweiligen σ-Algebren beachtet werden.

Definition 3.1.1 (messbare Abbildung). *Es seien $(\Omega, \mathcal{F})$ und $(\Omega', \mathcal{F}')$ zwei messbare Räume. Eine Abbildung $f\colon \Omega \to \Omega'$ heißt $\mathcal{F}$-$\mathcal{F}'$-*messbar *oder auch nur kurz* messbar*, falls für jedes $A \in \mathcal{F}'$ das Urbild $f^{-1}(A) = \{\omega \in \Omega \colon f(\omega) \in A\}$ in $\mathcal{F}$ liegt.*

Eine Abbildung $f\colon \Omega \to \Omega'$ heißt also messbar, falls die Urbild-σ-Algebra

$$f^{-1}(\mathcal{F}') = \{f^{-1}(A) \colon A \in \mathcal{F}'\}$$

eine Teilmenge von $\mathcal{F}$ ist. Im Fall $(\Omega', \mathcal{F}') = (\mathbb{R}, \mathcal{B})$ sprechen wir auch einfach von $\mathcal{F}$-Messbarkeit oder von Borel-Messbarkeit.

Bemerkung 3.1.2. Wenn über Ω a priori keine σ-Algebra existiert, so kann man mit Hilfe eines messbaren Raums $(\Omega', \mathcal{F}')$ und einer Abbildung $f\colon \Omega \to \Omega'$ über Ω eine σ-Algebra erklären, indem man $\mathcal{F} = \sigma(f) = f^{-1}(\mathcal{F}')$ setzt. Diese σ-Algebra ist die kleinste σ-Algebra über Ω, so dass f bezüglich ihr messbar ist. Man nennt $\sigma(f)$ auch die *durch f auf Ω erzeugte* σ-Algebra. $\diamond$

Für das Urbild $f^{-1}(A)$ werden wir manchmal auch abkürzend $\{f \in A\}$ schreiben. Darüber hinaus werden wir auch folgende Konventionen benutzen:

$$\begin{aligned}\{f < t\} &= \{\omega \in \Omega\colon f(\omega) < t\},\\ \{f = g\} &= \{\omega \in \Omega\colon f(\omega) = g(\omega)\},\\ \{f = 0\} &= \{\omega \in \Omega\colon f(\omega) = 0\}\end{aligned}$$

und so weiter.

Es ist im Folgenden bequem, den Wertebereich $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty\} \cup \{-\infty\} = [-\infty, \infty]$ zuzulassen. Eine Funktion $f\colon \Omega \to \overline{\mathbb{R}}$ belegen wir auch mit der unsinnigen, aber üblichen Bezeichnung *numerisch*. Auf $\overline{\mathbb{R}}$ betrachten wir die σ-Algebra $\overline{\mathcal{B}}$, die von $\mathcal{B}$, $\{\infty\}$ und $\{-\infty\}$ erzeugt wird. Eine $\mathcal{F}$-$\overline{\mathcal{B}}$-messbare Funktion ist dann also eine messbare numerische Funktion.

Bemerkung 3.1.3 (fast überall, fast sicher). Eine wichtige, oft benutzte Begriffsbildung im Zusammenhang mit Eigenschaften in Bezug auf Elemente in Ω ist die folgende. Wir sagen, eine Eigenschaft $E(\omega)$ gilt μ-*fast überall* oder kurz μ-*f. ü.*, falls es eine μ-Nullmenge N gibt, so dass die Menge derjenigen ω, für die $E(\omega)$ nicht gilt, in N enthalten ist; siehe Definition 2.2.11. (Man beachte, dass nicht vorausgesetzt wird, dass die Eigenschaft $E(\omega)$ messbar ist.) Falls also z. B. zwei messbare Funktionen f und g sich nur auf einer μ-Nullmenge von einander unterscheiden, d. h. wenn $\mu(\{\omega\colon f(\omega) \neq g(\omega)\}) = 0$ (die Messbarkeit der Menge $\{\omega\colon f(\omega) \neq g(\omega)\}$ folgt aus Lemma 3.1.7(b)), so schreiben wir $f = g$ μ-f. ü. In dem Fall, dass $\mu = \mathbb{P}$ ein Wahrscheinlichkeitsmaß ist, sagen wir $\mathbb{P}$-*fast sicher* und schreiben $\mathbb{P}$-f. s. ◇

Es ist schlicht unmöglich, alle Urbildmengen $f^{-1}(A)$ auf Zugehörigkeit zu $\mathcal{F}$ zu prüfen. Glücklicherweise kann man sich stark einschränken:

Lemma 3.1.4. *Eine Abbildung $f\colon \Omega \to \Omega'$ ist genau dann $\mathcal{F}$-$\mathcal{F}'$-messbar, wenn es ein Erzeugendensystem $\mathcal{C}$ von $\mathcal{F}'$ gibt mit $f^{-1}(\mathcal{C}) \subset \mathcal{F}$.*

Beweis. Man weist leicht als Übungsaufgabe nach, dass das System

$$\mathcal{A}_f = \{A \subset \Omega'\colon f^{-1}(A) \in \mathcal{F}\}$$

eine σ-Algebra ist. Da nach Voraussetzung $\mathcal{C} \subset \mathcal{A}_f$ gilt, gilt auch $\mathcal{F}' = \sigma(\mathcal{C}) \subset \mathcal{A}_f$. Dies aber bedeutet, dass $f^{-1}(\mathcal{F}') \subset \mathcal{F}$ gilt. □

Insbesondere ist eine Abbildung $f\colon \Omega \to \mathbb{R}$ schon Borel-messbar, wenn für alle $t \in \mathbb{R}$ gilt: $\{\omega \in \Omega\colon f(\omega) \leq t\} = f^{-1}((-\infty, t]) \in \mathcal{F}$, denn die Intervalle $(-\infty, t]$ bilden ein Erzeugendensystem von $\mathcal{B}$.

Messbarkeit wird unter nahezu allen nur denkbaren sinnvollen Operationen erhalten. Wir sammeln diese Ergebnisse in den folgenden Lemmata.

Lemma 3.1.5. *Hintereinanderausführungen von messbaren Funktionen sind messbar. Ausführlich: Falls $(\Omega_i, \mathcal{F}_i)$ mit $i = 1, 2, 3$ drei messbare Räume sind und $f_1\colon \Omega_1 \to \Omega_2$ sowie $f_2\colon \Omega_2 \to \Omega_3$ messbar sind, so ist auch $f_2 \circ f_1\colon \Omega_1 \to \Omega_3$ messbar.*

Beweis. Dies folgt leicht aus $(f_2 \circ f_1)^{-1}(A) = f_2^{-1}(f_1^{-1}(A)) \in \mathcal{F}_1$ für $A \in \mathcal{F}_3$. □

Lemma 3.1.6. *Es sei $(\Omega, \mathcal{F})$ ein messbarer Raum und $(f_n)_{n\in\mathbb{N}}$ eine Folge messbarer numerischer Funktionen auf Ω. Dann sind $\liminf_{n\to\infty} f_n$ und $\limsup_{n\to\infty} f_n$ ebenfalls messbare numerische Funktionen. Falls alle f_n reellwertig sind und $\liminf_{n\to\infty} f_n$ ebenfalls, so ist $\liminf_{n\to\infty} f_n$ Borel-messbar, analog für $\limsup_{n\to\infty} f_n$. Insbesondere sind (punktweise) Grenzwerte von messbaren Funktionen messbar.*

Beweis. Wir zeigen die Aussage nur im reellwertigen Fall für $\liminf_{n\to\infty} f_n$. Es genügt zu zeigen, dass die Menge $\{\omega \in \Omega \colon \liminf_{n\to\infty} f_n(\omega) \leq t\}$ für jedes $t \in \mathbb{R}$ in $\mathcal{F}$ liegt. Es ist $\liminf_{n\to\infty} f_n = \lim_{n\to\infty} \inf_{m\geq n} f_m$, wobei dieser Grenzwert monoton steigend ist. Also können wir schreiben:

$$\begin{aligned}\{\omega \colon \underline{f}(\omega) \leq t\} &= \bigcap_{n\in\mathbb{N}} \Big\{\omega \colon \inf_{m\geq n} f_m(\omega) \leq t\Big\} = \bigcap_{n\in\mathbb{N}} \bigcap_{k\in\mathbb{N}} \Big\{\omega \colon \inf_{m\geq n} f_m(\omega) < t + \frac{1}{k}\Big\} \\ &= \bigcap_{n\in\mathbb{N}} \bigcap_{k\in\mathbb{N}} \bigcup_{m\in\mathbb{N}\colon m\geq n} \Big\{\omega \colon f_m(\omega) < t + \frac{1}{k}\Big\} \in \mathcal{F}. \qquad \square\end{aligned}$$

Messbarkeit erhält sich auch unter allen algebraischen Verknüpfungen:

Lemma 3.1.7. *(a) Jede konstante numerische Funktion ist messbar.*

(b) Sind f und g messbare numerische Funktionen, so sind auch $f + g$ und fg messbar.

(c) Sind f und g messbare numerische Funktionen und $g(\omega) \neq 0$ für jedes ω, so ist f/g messbar.

(d) Die Indikatorfunktion $\mathbb{1}_A$ einer Menge $A \subset \Omega$ ist genau dann messbar, wenn A messbar ist.

Hierbei ist die *Indikatorfunktion* $\mathbb{1}_A \colon \Omega \to \mathbb{R}$ auf einer Menge $A \subset \Omega$ definiert durch

$$\mathbb{1}_A(\omega) = \begin{cases} 1, & \text{falls } \omega \in A, \\ 0 & \text{sonst.} \end{cases}$$

Beweis von Lemma 3.1.7. Die Aussagen (a) und (d) sind klar. Offensichtlich ist die Menge

$$\{f < g\} = \bigcup_{q\in\mathbb{Q}} \{f < q\} \cap \{q < g\} = \bigcup_{q\in\mathbb{Q}} f^{-1}(([-\infty, q)) \cap g^{-1}(([-\infty, q))$$

messbar, wenn f und g messbar sind, und leicht sieht man auch die Messbarkeit der Mengen $\{f \leq g\} = \bigcap_{q\in\mathbb{Q}\cap(0,\infty)} \{f < g + q\}$, $\{f = g\}$ und $\{f \neq g\}$ ein. Auch sieht man leicht die Messbarkeit von $\alpha + \tau g$ für alle $\alpha, \tau \in \mathbb{R}$. Die Messbarkeit von $f + g$ folgt dann aus der Messbarkeit der Menge $\{f + g \leq t\} = \{f \leq t - g\}$ für jedes $t \in \mathbb{R}$. Dass fg messbar ist, sieht man an der Darstellung $fg = \frac{1}{4}(f + g)^2 - \frac{1}{4}(f - g)^2$, denn auch die Messbarkeit von Quadraten messbarer Funktionen ist leicht zu sehen. Der Beweis von (c) ist eine Übungsaufgabe. $\square$

Bemerkung 3.1.8 (Borel-σ-Algebren). Die Borel-σ-Algebra kann natürlicherweise auf jedem *topologischen* Raum Ω definiert werden als die von dem System der offenen Mengen erzeugte σ-Algebra. Dies erzwingt auf natürliche Weise eine Kompatibilität mit der Stetigkeit: Jede stetige Abbildung zwischen messbaren Räumen, die mit der entsprechenden Borel-σ-Algebra versehen sind, ist messbar. Dies folgt direkt aus der Tatsache, dass die Urbilder offener Mengen unter stetigen Abbildungen offen sind, in Kombination mit Lemma 3.1.4. $\diamond$

Lemma 3.1.9. *Ist f eine messbare numerische Funktion, so sind $f^+ = \max\{f, 0\}$, $f^- = \max\{-f, 0\}$ und $|f|$ messbar.*

Beweis. Man kombiniere Lemma 3.1.5, Lemma 3.1.7 und Bemerkung 3.1.8. $\square$

Definition 3.1.10 (einfache Funktion). *Es sei $(\Omega, \mathcal{F})$ ein messbarer Raum. Eine* einfache Funktion *ist eine endliche Linearkombination von Indikatorfunktionen auf messbaren Mengen, d. h. eine Funktion der Form $\sum_{i=1}^n a_i \mathbb{1}_{A_i}$ mit $n \in \mathbb{N}$, $a_1, \dots, a_n \in \mathbb{R}$ und $A_1, \dots, A_n \in \mathcal{F}$.*

Man kann auch leicht eine Darstellung mit paarweise disjunkten Mengen $A_1, \dots, A_n \in \mathcal{F}$ erreichen. Die Menge der einfachen Funktionen ist offensichtlich abgeschlossen gegen (punktweise) Addition, Multiplikation, Maximum- und Minimumbildung. Mit ihnen kann man alle nichtnegativen messbaren numerischen Funktionen approximieren durch monotone Grenzwertbildung:

Lemma 3.1.11. *Für jede nichtnegative messbare numerische Funktion f gibt es eine monoton wachsende Folge von einfachen nichtnegativen Funktionen $f_1, f_2, f_3, \dots$ mit $f_n \uparrow f$ für $n \to \infty$.*

Beweis. Wir wählen z. B.

$$f_n = \sum_{k=1}^{n2^n} (k-1)2^{-n} \mathbb{1}_{\{(k-1)2^{-n} \le f < k2^{-n}\}} + n\mathbb{1}_{\{f \ge n\}}.$$

(Man beachte, dass wir den Bildbereich der Funktion f zerlegt haben.) $\square$

Mit diesem Ergebnis kann man die Menge aller nichtnegativen messbaren numerischen Funktionen wie folgt charakterisieren:

Satz 3.1.12. *Es sei $(\Omega, \mathcal{F})$ ein messbarer Raum und Γ eine Menge von nichtnegativen messbaren numerischen Funktionen mit den folgenden Eigenschaften:*

(i) Falls $f, g \in \Gamma$ und $a, b \in (0, \infty)$, so gilt $af + bg \in \Gamma$.

(ii) Falls $f_1, f_2, f_3, \dots \in \Gamma$ mit $f_n \uparrow f$ für $n \to \infty$, so gilt $f \in \Gamma$.

(iii) Γ enthält alle Indikatorfunktionen $\mathbb{1}_A$ mit $A \in \mathcal{F}$.

Dann ist Γ die Menge aller nichtnegativen messbaren numerischen Funktionen.

Beweis. Aus (i) und (iii) folgt, dass Γ alle nichtnegativen einfachen Funktionen enthält. Mit (ii) und Lemma 3.1.11 folgt die Aussage. $\square$

Satz 3.1.12 wird routinemäßig für diverse Beweise über Eigenschaften messbarer Funktionen oder verwandter Objekte eingesetzt werden. Wenn gezeigt werden soll, dass alle nichtnegativen messbaren numerischen Funktionen eine gewisse Eigenschaft besitzen, so reicht es also zu zeigen, dass alle Indikatorfunktionen auf messbaren Mengen diese Eigenschaft besitzen und dass diese Eigenschaft sich erhält unter Bildung von Linearkombinationen und monotonen Grenzwerten. Die letztere Eigenschaft wird sehr oft mit Hilfe des Monotonen Konvergenzsatzes (Satz 3.4.1) geprüft. Ein sehr klares Beispiel für diesen Beweisweg ist der Beweis von Satz 3.6.1.

3.2 Integration

In diesem Abschnitt sei ein beliebiger Maßraum $(\Omega, \mathcal{F}, \mu)$ vorgegeben. Wir wollen nun das Integral $\int_\Omega f \, \mathrm{d}\mu$ von messbaren Funktionen $f\colon \Omega \to \mathbb{R}$ bezüglich μ einführen. Dies machen wir in drei Schritten:

(i) für einfache Funktionen durch eine offensichtliche Formel,

(ii) für nichtnegative Funktionen durch monotonen Grenzübergang,

(iii) für beliebige Funktionen durch Zerlegung in Positiv- und Negativteil.

Dieses Vorgehen ist fundamental in der Integrationstheorie, und es wird sich später an vielen Stellen zeigen, dass es auch ein effektives Beweisgerüst liefert für eine Vielzahl von Aussagen über Integrale messbarer Funktionen (siehe auch die Bemerkung nach Satz 3.1.12).

Definition 3.2.1 (Definition des Integrals).

(i) Das Integral einer nichtnegativen einfachen Funktion $f = \sum_{i=1}^n a_i \mathbb{1}_{A_i}$ mit $A_1, \dots, A_n \in \mathcal{F}$ und $a_1, \dots, a_n > 0$ wird definiert als

$$\int f \, \mathrm{d}\mu = \sum_{i=1}^n a_i \mu(A_i).$$

(ii) Das Integral einer nichtnegativen messbaren numerischen Funktion $f\colon \Omega \to [0, \infty]$ wird definiert als

$$\int f \, \mathrm{d}\mu = \lim_{n\to\infty} \int f_n \, \mathrm{d}\mu \in [0, \infty], \tag{3.2.1}$$

wobei $(f_n)_{n\in\mathbb{N}}$ eine Folge nichtnegativer einfacher Funktionen sei mit $f_n \uparrow f$.

(iii) Sei $f\colon \Omega \to \overline{\mathbb{R}}$ messbar und numerisch, so dass $\int f^+ \, \mathrm{d}\mu < \infty$ und $\int f^- \, \mathrm{d}\mu < \infty$. In diesem Fall nennen wir f integrierbar bezüglich *μ, und das Integral von f wird definiert als*

$$\int f \, \mathrm{d}\mu = \int f^+ \, \mathrm{d}\mu - \int f^- \, \mathrm{d}\mu$$

Wir schreiben $\mathcal{L}^1(\Omega, \mathcal{F}, \mu)$ (oder auch kurz $\mathcal{L}^1(\mu)$ oder einfach nur $\mathcal{L}^1$) für die Menge aller bezüglich μ integrierbaren Funktionen.

Es sind etliche Bemerkungen notwendig:

Bemerkung 3.2.2. (a) Um zu zeigen, dass Definition 3.2.1(i) sinnvoll ist, muss man zeigen, dass der Ausdruck $\sum_{i=1}^n a_i \mu(A_i)$ nicht von der gewählten Darstellung $f = \sum_{i=1}^n a_i \mathbb{1}_{A_i}$ abhängt. Dies kann man leicht tun, nachdem man sich überlegt hat, dass man ohne Einschränkung davon ausgehen kann, dass die Mengen $A_1, \dots, A_n$ paarweise disjunkt sind (Übungsaufgabe bzw. siehe [Ba91, Lemma 10.2] oder [Kl06, Lemma 4.1]).

(b) Eine ähnliche, aber etwas schwierigere Aufgabe hat man bei Definition 3.2.1(ii): Man muss zeigen, dass der Ausdruck $\lim_{n\to\infty} \int f_n \, \mathrm{d}\mu$ nicht von der gewählten Folge $(f_n)_{n\in\mathbb{N}}$ abhängt. (Die Existenz einer solchen Folge wird ja durch Lemma 3.1.11 garantiert.) Dies wird in

Lemma 3.2.3 behandelt. Wir weisen darauf hin, dass dieser Punkt die eigentliche Schwierigkeit ist beim Beweis der folgenden Resultate in Lemma 3.2.6 und den Sätzen 3.4.1 und 3.4.6. Man könnte auch das Integral einer nichtnegativen messbaren Funktion definieren als *Supremum* aller Integrale messbarer nichtnegativer einfacher Funktionen, die nicht größer sind als die betrachtete Funktion, und verschiebt die oben genannte Schwierigkeit.

(c) Statt $\int f\,\mathrm{d}\mu$ schreibt man auch oft ausführlicher $\int_\Omega f(\omega)\,\mu(\mathrm{d}\omega)$ oder $\int_\Omega f(\omega)\,\mathrm{d}\mu(\omega)$.

(d) Für nichtnegative numerische messbare Funktionen f ist $\int f\,\mathrm{d}\mu$ also immer definiert, aber es kann gleich ∞ sein. Man definiert manchmal das Integral einer beliebigen numerischen messbaren Funktion f auch dann noch, wenn $\int f^+\,\mathrm{d}\mu = \infty$ und $\int f^-\,\mathrm{d}\mu < \infty$ (dann mit Wert $\int f\,\mathrm{d}\mu = \infty$) oder wenn $\int f^+\,\mathrm{d}\mu < \infty$ und $\int f^-\,\mathrm{d}\mu = \infty$ (dann mit Wert $\int f\,\mathrm{d}\mu = -\infty$).

(e) Falls $f \geq 0$ und $\mu(\{\omega\colon f(\omega) = \infty\}) > 0$, so gilt $\int f\,\mathrm{d}\mu = \infty$, wie man leicht als eine Übungsaufgabe zeigt.

(f) Falls $\mu = \mathbb{P}$ ein Wahrscheinlichkeitsmaß ist, so schreiben wir auch $\mathbb{E}(f)$ (oder $\mathbb{E}_{\mathbb{P}}(f)$, falls $\mathbb{P}$ betont werden soll) statt $\int f\,\mathrm{d}\mathbb{P}$ und nennen $\mathbb{E}(f)$ den *Erwartungswert* von f.

◇

Nun folgt also die in Bemerkung 3.2.2(b) angekündigte Rechtfertigung dafür, dass Definition 3.2.1(ii) sinnvoll ist: Wir zeigen, dass der Wert der linken Seite von (3.2.1) nicht von der Folge $(f_n)_{n\in\mathbb{N}}$ abhängt.

Lemma 3.2.3. *Seien $f\colon \Omega \to [0,\infty]$ eine messbare Funktion und $(f_n)_{n\in\mathbb{N}}$ und $(g_n)_{n\in\mathbb{N}}$ zwei Folgen einfacher nichtnegativer Funktionen mit $f_n \uparrow f$ und $g_n \uparrow f$ für $n \to \infty$. Dann gilt $\lim_{n\to\infty} \int f_n\,\mathrm{d}\mu = \lim_{n\to\infty} \int g_n\,\mathrm{d}\mu$.*

Beweis. Wegen der Symmetrie der Aussage in f_n und g_n müssen wir nur die Ungleichung $\lim_{n\to\infty} \int f_n\,\mathrm{d}\mu \leq \lim_{n\to\infty} \int g_n\,\mathrm{d}\mu$ zeigen. Dafür reicht es zu zeigen, dass für jedes $k \in \mathbb{N}$ gilt: $\int f_k\,\mathrm{d}\mu \leq \lim_{n\to\infty} \int g_n\,\mathrm{d}\mu$. Es sei $f_k = \sum_{i=1}^N a_i \mathbb{1}_{A_i}$ mit $a_1, \dots, a_N \geq 0$ und messbaren, paarweise disjunkten Mengen $A_1, \dots, A_N$. Sei $\varepsilon > 0$. Wir betrachten die Menge $B_n = \{g_n \geq (1-\varepsilon)f_k\}$. Wegen $g_n \uparrow f$ gilt $B_n \uparrow \Omega$ für $n \to \infty$. Es ist leicht zu sehen, dass das Integral für einfache Funktionen monoton und linear ist. Also gilt

$$\int g_n\,\mathrm{d}\mu \geq \int \mathbb{1}_{B_n} g_n\,\mathrm{d}\mu \geq \int \mathbb{1}_{B_n}\big((1-\varepsilon)f_k\big)\,\mathrm{d}\mu = \int \sum_{i=1}^N (1-\varepsilon)a_i \mathbb{1}_{A_i\cap B_n}\,\mathrm{d}\mu$$
$$= \sum_{i=1}^N (1-\varepsilon)a_i\mu(A_i \cap B_n) \overset{n\to\infty}{\longrightarrow} \sum_{i=1}^N (1-\varepsilon)a_i\mu(A_i) = (1-\varepsilon)\int f_k\,\mathrm{d}\mu.$$

Nun können wir $\varepsilon \downarrow 0$ lassen und erhalten die gewünschte Aussage. □

Lemma 3.2.4 (Eigenschaften des Integrals). *Seien $f, g \in \mathcal{L}^1(\mu)$.*

(i) (Monotonie.) *Falls $f \leq g$, so gilt $\int f\,\mathrm{d}\mu \leq \int g\,\mathrm{d}\mu$.*

(ii) (Dreiecksungleichung.) *Es gilt $|\int f\,\mathrm{d}\mu| \leq \int |f|\,\mathrm{d}\mu$.*

(iii) (Linearität.) *Für alle $\alpha, \beta \in \mathbb{R}$ ist $\alpha f + \beta g \in \mathcal{L}^1(\mu)$, und es gilt $\int (\alpha f + \beta g)\,\mathrm{d}\mu = \alpha \int f\,\mathrm{d}\mu + \beta \int g\,\mathrm{d}\mu$. Dies gilt auch noch, wenn höchstens eines der Integrale $\int f\,\mathrm{d}\mu$ und $\int g\,\mathrm{d}\mu$ einen der Werte ∞ oder $-\infty$ annimmt.*

Beweis. (i). Für nichtnegative einfache Funktionen kann man die Ungleichung leicht nachrechnen, nachdem man Darstellungen für f und g gewählt hat, die die selben paarweise disjunkten messbaren Mengen benutzen. Für nichtnegative messbare Funktionen f und g wählen wir Folgen einfacher nichtnegativer Funktionen $(f_n)_{n\in\mathbb{N}}$ und $(g_n)_{n\in\mathbb{N}}$ mit $f_n \uparrow f$ und $g_n \uparrow g$. Die Folge $(f_n \wedge g_n)_{n\in\mathbb{N}}$[1] besteht dann ebenfalls aus einfachen Funktionen, die monoton steigend gegen f konvergieren. Wegen $f_n \wedge g_n \leq g_n$ gilt $\int f_n \wedge g_n\,\mathrm{d}\mu \leq \int g_n\,\mathrm{d}\mu$ für jedes $n \in \mathbb{N}$. Der Grenzübergang $n \to \infty$ ergibt

$$\int f\,\mathrm{d}\mu = \lim_{n\to\infty} \int f_n \wedge g_n\,\mathrm{d}\mu \leq \lim_{n\to\infty} \int g_n\,\mathrm{d}\mu = \int g\,\mathrm{d}\mu.$$

Für allgemeine messbare f und g mit $f \leq g$ haben wir $f^+ \leq g^+$ und $f^- \geq g^-$ und daher die entsprechenden Ungleichungen für die Integrale der Positiv- und Negativteile. Durch Zusammensetzen erhält man die Aussage.

(ii). Für nichtnegative messbare Funktionen f und g und nichtnegative α und β folgt die Beziehung $\int (\alpha f + \beta g)\,\mathrm{d}\mu = \alpha \int f\,\mathrm{d}\mu + \beta \int g\,\mathrm{d}\mu$ (also die Linearität) leicht aus der Linearität des Integrals für einfache Funktionen sowie der Definition des Integrals durch Grenzübergang, d. h. monotone Approximation. Dies wenden wir an auf die Zerlegung $f = f^+ - f^-$ (also $|f| = f^+ + f^-$) und erhalten die Dreiecksungleichung.

(iii). Wie im Beweis von (ii) erwähnt, ist die Linearität klar für nichtnegative f, g, α und β. Seien nun $f, g \in \mathcal{L}^1(\mu)$ und α und $\beta \in \mathbb{R}$ beliebig. Um die Linearität zu zeigen, reicht es zu zeigen, dass (die folgenden Integrale existieren und) (a) $\int (f+g)\,\mathrm{d}\mu = \int f\,\mathrm{d}\mu + \int g\,\mathrm{d}\mu$ gilt, dass (b) für $\alpha \geq 0$ gilt: $\int (\alpha f)\,\mathrm{d}\mu = \alpha \int f\,\mathrm{d}\mu$, und dass (c) $\int (-f)\,\mathrm{d}\mu = -\int f\,\mathrm{d}\mu$ gilt. (a) zeigt man, indem man zerlegt: $(f+g)^+ - (f+g)^- = f + g = f^+ - f^- + g^+ - g^-$, also hat man $(f+g)^+ + f^- + g^- = (f+g)^- + f^+ + g^+$ und daher die entsprechende Beziehung für die Integrale, und daher

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

Dei beiden anderen Aussagen (b) und (c) zeigt man wiederum durch Zerlegung in Positiv- und Negativteil. Der Beweis des Zusatzes ist eine Übungsaufgabe. □

Wir ziehen ein paar Folgerungen:

Korollar 3.2.5. *(i) Falls $g\colon \Omega \to [0,\infty]$ nichtnegativ und integrierbar ist und f eine numerische messbare Funktion mit $|f| \leq g$, so ist auch f integrierbar. Insbesondere ist eine messbare numerische Funktion genau dann integrierbar, wenn ihr Betrag integrierbar ist.*

(ii) Die Menge $\mathcal{L}^1(\mu)$ ist ein Vektorraum.

Für eine messbare Menge $A \in \mathcal{F}$ und ein $f \in \mathcal{L}^1$ setzt man $\int_A f\,\mathrm{d}\mu = \int (f\mathbb{1}_A)\,\mathrm{d}\mu$. Die Wohldefiniertheit dieses Integral folgt aus Korollar 3.2.5(i), denn $|f\mathbb{1}_A| \leq |f|$.

Noch zwei wichtige Eigenschaften der Integrale messbarer Funktionen sind die folgenden.

[1] Wir schreiben $x \wedge y = \min\{x, y\}$ und $x \vee y = \max\{x, y\}$ für das Minimum bzw. das Maximum zweier reeller Zahlen x und y.

Lemma 3.2.6. *Es seien $f, g \in \mathcal{L}^1$. Dann gelten:*

(a) Falls $f \le g$ μ-f. ü., so gilt $\int f \, \mathrm{d}\mu \le \int g \, \mathrm{d}\mu$.

(b) Falls $f \ge g$ μ-f. ü. und $\int f \, \mathrm{d}\mu = \int g \, \mathrm{d}\mu$, so gilt $f = g$ μ-f. ü.

Beweis. (a) Durch Zerlegung in Positiv- und Negativteil können wir annehmen, dass f und g nichtnegativ sind. Es sei $(f_n)_{n\in\mathbb{N}}$ eine Folge von einfachen Funktionen mit $f_n \uparrow f$. Für die (messbare) Menge $D = \{\omega \in \Omega \colon f(\omega) \le g(\omega)\}$ gelten $\mu(D^{\mathrm{c}}) = 0$ und $\mathbb{1}_D f_n \uparrow \mathbb{1}_D f \le \mathbb{1}_D g$. Man beachte, dass für jede einfache Funktion h gilt: $\int_D h \, \mathrm{d}\mu = \int h \, \mathrm{d}\mu$, da ja D^{c} eine μ-Nullmenge ist. Insbesondere ist $\int_D f_n \, \mathrm{d}\mu = \int f_n \, \mathrm{d}\mu$ für jedes $n \in \mathbb{N}$, und der Grenzübergang $n \to \infty$ lässt $\int_D f \, \mathrm{d}\mu = \int f \, \mathrm{d}\mu$ folgen. Analog gilt $\int_D g \, \mathrm{d}\mu = \int g \, \mathrm{d}\mu$. Wegen $\mathbb{1}_D f \le \mathbb{1}_D g$ folgt die Aussage.

(b) Nach Voraussetzung gelten $f - g \ge 0$ μ-f. ü. und $\int (f-g) \, \mathrm{d}\mu = 0$. Es genügt also, den Fall $g = 0$ zu betrachten. Für $n \in \mathbb{N}$ sei $A_n = \{f \ge \frac{1}{n}\}$. Dann gilt $f \ge \frac{1}{n}\mathbb{1}_{A_n}$. Aus $\int f \, \mathrm{d}\mu = 0$ folgt somit $0 = \int \frac{1}{n}\mathbb{1}_{A_n} \, \mathrm{d}\mu = \frac{1}{n}\mu(A_n)$, also $\mu(A_n) = 0$. Da $\{f > 0\} = \bigcup_{n\in\mathbb{N}} A_n$ eine abzählbare Vereinigung von Nullmengen ist, ist $\{f > 0\}$ nach Lemma 2.2.12 selber eine. □

Beispiel 3.2.7 (Diskrete Integrale). Sei $\mu = \sum_{\omega\in\Omega} \alpha(\omega)\delta_\omega$ ein diskretes Maß; siehe Beispiel 2.2.10. D. h., Ω ist eine höchstens abzählbare Menge, δ_ω ist das Dirac-Maß (Einpunktmaß) in ω, und $\alpha(\omega)$ sind nichtnegative Zahlen. Dann ist eine Abbildung $f \colon \Omega \to \mathbb{R}$ genau dann integrierbar bezüglich μ, wenn die Reihe $\sum_{\omega\in\Omega} |f(\omega)|\alpha(\omega)$ konvergiert, und in diesem Fall ist $\int f \, \mathrm{d}\mu = \sum_{\omega\in\Omega} f(\omega)\alpha(\omega)$. Wir sehen also, dass die Theorie der absolut konvergenten Reihen ein Spezialfall der Integrationstheorie ist.

Ein konkretes Beispiel ist das Zählmaß $\mu = \sum_{n\in\mathbb{N}} \delta_n$ auf dem messbaren Raum $(N, \mathcal{P}(\mathbb{N}))$, das $\mu(A) = \#A$ für jede endliche Menge $A \subset \mathbb{N}$ erfüllt und $\mu(A) = \infty$, falls $A \subset \mathbb{N}$ nicht endlich ist. Hier ist der Raum $\mathcal{L}^1(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu)$ nichts Anderes als der Folgenraum $\ell^1(\mathbb{N})$, die Menge aller absolut summierbaren reellen Folgen. ◇

3.3 Riemann- versus Lebesgue-Integral

Das Lebesgue-Maß λ_d auf der Borel-σ-Algebra $\mathcal{B}_d$ auf dem $\mathbb{R}^d$ wurde in Bemerkung 2.2.8 als dasjenige Maß eingeführt, das den achsenparallelen Quadern ihren elementargeometrischen Inhalt zuordnet. Seine Existenz wurde in Korollar 2.4.3 bewiesen, zumindest für $d = 1$. Mit $\mathcal{B}_d^*$ wollen wir die Vervollständigung von $\mathcal{B}_d$ bezüglich dieses Maßes bezeichen und mit dem selben Symbol λ_d die Erweiterung von λ_d auf $\mathcal{B}_d^*$, siehe Definition 2.2.11 und Bemerkung 2.2.13. Es ist eine Übungsaufgabe zu zeigen, dass $\mathcal{B}_d^*$ identisch ist mit der σ-Algebra $\mathcal{F}$ der Lebesgue-integrierbaren Mengen, die in Abschnitt 2.1 eingeführt wurde, und dass das Maß λ_d auf $\mathcal{B}_d^*$ identisch ist mit dem dort eingeführten Lebesgue-Maß, siehe Bemerkung 2.1.11. Die Integration bezüglich des Lebesgue-Maßes λ_d auf $(\mathbb{R}^d, \mathcal{B}_d)$ wird *Lebesgue-Integration* genannt, wobei wir hierbei das vervollständigte Lebesgue-Maß meinen.

Der größte und wichtigste Unterschied zum Riemann-Integral ist, dass die Unterteilungen beim Riemann-Integral im Definitionsbereich gemacht werden und beim Lebesgue-Integral im Wertebereich. Das Riemann-Integral respektiert die Geometrie des Definitionsbereichs, indem es als Grenzprozess aus einer Summe vieler schmaler, *senkrechter* Streifen entsteht, während das Lebesgue-Integral über *waagerechte* Streifen definiert ist. Insbesondere macht letzteres keine Annahmen über den Definitionsbereich, was es universeller macht.

Das Riemann-Integral kommt nicht von einem Maß her, denn dieser Integralbegriff ist nicht kompatibel mit abzählbaren Operationen. Zum Beispiel ist $\mathbb{1}_{\{\omega\}}$ für jedes $\omega \in [0,1]$ Riemann-integrierbar, aber nicht die abzählbare Summe $\mathbb{1}_{\mathbb{Q}\cap[0,1]} = \sum_{\omega\in\mathbb{Q}\cap[0,1]} \mathbb{1}_{\{\omega\}}$, die allerdings ihrerseits wieder Lebesgue-integrierbar ist. Trotzdem gibt es enge Zusammenhänge zwischen dem Riemann- und dem Lebesgue-Integral:

Satz 3.3.1. *Jede auf einem kompakten Intervall Riemann-integrierbare Funktion ist Lebesgue-integrierbar, und die beiden Integrale stimmen überein.*

Beweis. (a) Sei $I = [a,b] \subset \mathbb{R}$ das kompakte Intervall, und sei $(t^{(n)})_{n\in\mathbb{N}}$ eine Folge von Unterteilungen $t^{(n)} = \{a = t_0^{(n)} < t_1^{(n)} < \cdots < t_n^{(n)} = b\}$ des Intervalls I, so dass die Feinheit $\eta_n = \max_{i=1}^n (t_i^{(n)} - t_{i-1}^{(n)})$ gegen Null konvergiert. Wir dürfen annehmen, dass $t^{(n+1)}$ eine Verfeinerung von $t^{(n)}$ ist, d. h. dass $t^{(n+1)} \subset t^{(n)}$ gilt. Da f Riemann-integrierbar über I ist, konvergieren die Unter- und die Obersumme

$$U(f,t^{(n)}) = \sum_{i=1}^n (t_i^{(n)} - t_{i-1}^{(n)}) \inf_{[t_{i-1}^{(n)},t_i^{(n)})} f \qquad \text{bzw.} \qquad O(f,t^{(n)}) = \sum_{i=1}^n (t_i^{(n)} - t_{i-1}^{(n)}) \sup_{[t_{i-1}^{(n)},t_i^{(n)})} f$$

für $n \to \infty$ gegen $\int_a^b f(x)\,\mathrm{d}x$. Insbesondere ist f beschränkt. Nach Addition einer Konstanten können wir davon ausgehen, dass f nichtnegativ ist.

Wir betrachten nun die einfachen Funktionen

$$\begin{aligned} g_n &= f(b)\mathbb{1}_{\{b\}} + \sum_{i=1}^n \Big(\inf_{[t_{i-1}^{(n)},t_i^{(n)})} f \Big) \mathbb{1}_{[t_{i-1}^{(n)},t_i^{(n)})}, \\ h_n &= f(b)\mathbb{1}_{\{b\}} + \sum_{i=1}^n \Big(\sup_{[t_{i-1}^{(n)},t_i^{(n)})} f \Big) \mathbb{1}_{[t_{i-1}^{(n)},t_i^{(n)})}. \end{aligned}$$

Da $t^{(n+1)}$ eine Verfeinerung von $t^{(n)}$ ist, gelten $g_n \leq g_{n+1} \leq f \leq h_{n+1} \leq h_n$ für jedes $n \in \mathbb{N}$. Also existieren messbare nichtnegative Funktionen g und h mit $g_n \uparrow g$ und $h_n \downarrow h$ sowie $g \leq f \leq h$. Nach Konstruktion gilt für die Lebesgue-Integrale

$$\int_I g\,\mathrm{d}\lambda = \lim_{n\to\infty} \int_I g_n\,\mathrm{d}\lambda = \lim_{n\to\infty} U(f,t^{(n)}) = \lim_{n\to\infty} O(f,t^{(n)}) = \lim_{n\to\infty} \int_I h_n\,\mathrm{d}\lambda = \int_I h\,\mathrm{d}\lambda.$$

Nach Lemma 3.2.6(b) ist λ-fast überall $g = h$. Für jedes $\alpha \in \mathbb{R}$ haben wir

$$\{f \leq \alpha\} = \big(\{g \leq \alpha\} \cap \{g = h\}\big) \cup \big(\{f \leq \alpha\} \cup \{g \neq h\}\big),$$

und dies ist die Vereinigung einer Menge aus $\mathcal{B}(I)$ (der Spur-σ-Algebra auf I) und einer Teilmenge einer Nullmenge. Also liegt $\{f \leq \alpha\}$ in der Vervollständigung $\mathcal{B}(I)^*$ von $\mathcal{B}(I)$ bezüglich des Lebesgue-Maßes. Also ist f messbar bezüglich $\mathcal{B}(I)^*$. Die Definition des Integrals impliziert, dass $\int_I f\,\mathrm{d}\mu = \lim_{n\to\infty} \int_I g_n\,\mathrm{d}\mu = \int_a^b f(x)\,\mathrm{d}x$ gilt, und das beendet den Beweis. $\square$

Als Übungsaufgabe leitet man daraus den folgenden Zusammenhang zwischen uneigentlichen Riemann-Integralen und dem Lebesgue-Integral her:

Lemma 3.3.2. *Jede auf einem unbeschränkten Intervall I messbare nichtnegative Funktion, die über jedes kompakte Teilintervall Riemann-integrierbar ist, ist genau dann Lebesgue-integrierbar über I, wenn ihr uneigentliches Riemann-Integral über I existiert, und die beiden Integrale stimmen dann überein.*

Dass man nicht auf die Nichtnegativität verzichten kann, sieht man an dem Beispiel $[1,\infty) \ni x \mapsto \frac{1}{x}\sin x$: In einer beliebten Übungsaufgabe zeigt man in der Analysis-Grundvorlesung, dass das uneigentliche Riemann-Integral $\int_1^\infty \frac{1}{x}\sin x\,\mathrm{d}x$ existiert, nicht aber $\int_1^\infty |\frac{1}{x}\sin x|\,\mathrm{d}x$. Für das uneigentliche Riemann-Integral gilt eben *nicht* die Regel (siehe Korollar 3.2.5), dass eine Funktion genau dann integrierbar ist, wenn ihr Betrag integrierbar ist.

Satz 3.3.1 und Lemma 3.3.2 implizieren, dass die meisten Riemann-Integrale, die in der Praxis auftauchen, auch als Lebesgue-Integrale aufgefasst werden können und dass insbesondere die bekannten Integrationstechniken auch für diese Integrale gelten. Wir zitieren hier ohne Beweis (dafür siehe etwa [E05, Abschnitt VI.6.1]) die folgende schöne Charakterisierung der Riemann-Integrierbarkeit: Eine beschränkte Lebesgue-messbare Funktion ist auf einem kompakten Intervall genau dann Riemann-integrierbar, wenn die Menge ihrer Unstetigkeitsstellen eine Lebesgue-Nullmenge ist, und die beiden Integrale stimmen dann überein.

3.4 Die fundamentalen Konvergenzsätze

In diesem Abschnitt sei $(\Omega, \mathcal{F}, \mu)$ wieder ein gegebener Maßraum. Man steht oft vor dem Problem, ein Integral mit einem Grenzwert zu vertauschen, also eine Regel wie $\int \lim_{n\to\infty} f_n\,\mathrm{d}\mu = \lim_{n\to\infty}\int f_n\,\mathrm{d}\mu$ anzuwenden. Dies geht natürlich nicht immer; zum Beispiel konvergieren die 'Zackenfunktionen' $f_n\colon [0,1] \to [0,\infty)$, deren Graph der Polygonzug durch $(0,0), (\frac{1}{n}, n), (\frac{2}{n}, 0)$ und $(1,0)$ ist, gegen $f = 0$, aber ihre Integrale sind alle gleich Eins.

Wir geben die zwei berühmten hinreichenden Kriterien, unter denen man den Grenzwert mit dem Integral vertauschen darf. Man beachte, dass die Voraussetzungen extrem schwach sind im Vergleich etwa zum Riemann-Integralbegriff. Wir erinnern daran, dass die Notation $f_n \uparrow f$ bedeutet, dass $(f_n)_{n\in\mathbb{N}}$ monoton steigt und gegen f konvergiert.

Satz 3.4.1 (Monotoner Konvergenzsatz). *Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge nichtnegativer, messbarer numerischer Funktionen mit $f_n \uparrow f$ μ-f. ü. für $n \to \infty$. Dann gilt*

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

Man beachte, dass wir nicht vorausgesetzt haben, dass $\int f\,\mathrm{d}\mu$ endlich ist. Die Bedingung $f_n \geq 0$ ist nötig, wie man an dem Beispiel $\Omega = \mathbb{R}$, $\mu =$ Lebesgue-Maß und $f_n = -\frac{1}{n}$ sieht.

Beweis. Durch eine Modifikation auf Nullmengen (die nach Lemma 3.2.6(a) nichts an den Werten der Integrale ändert) können wir annehmen, dass $f_n(\omega) \uparrow f(\omega)$ sogar für *alle* $\omega \in \Omega$ gilt. Wegen der Monotonie des Integrals gilt natürlich $\int f_n\,\mathrm{d}\mu \leq \int f\,\mathrm{d}\mu$ für jedes $n \in \mathbb{N}$. Wegen Monotonie der Folge $(f_n)_{n\in\mathbb{N}}$ existiert der Grenzwert $\lim_{n\to\infty}\int f_n\,\mathrm{d}\mu$ in $[0,\infty]$. Wir müssen also nur noch zeigen, dass er nicht kleiner als $\int f\,\mathrm{d}\mu$ ist.

Wir wählen für jedes $n \in \mathbb{N}$ eine Folge $(f_{n,m})_{m\in\mathbb{N}}$ nichtnegativer einfacher Funktionen mit $f_{n,m} \uparrow f_n$ für $m \to \infty$. Wir betrachten die einfache Funktion

$$g_{n,m} = \max\{f_{1,m}, \dots, f_{n,m}\}.$$

Es gilt auch $g_{n,m} \uparrow f_n$ für $m \to \infty$ für jedes $n \in \mathbb{N}$, denn die Monotonie in m ist klar, und es gilt $f_n \geq g_{n,m} \geq f_{n,m}$. Außerdem ist offensichtlich $g_{n,m}$ monoton in n, d. h. $g_{n+1,m} \geq g_{n,m}$ für alle

n, m. Daraus folgt sehr einfach, dass $g_{n,n} \uparrow f$ für $n \to \infty$. Damit ergibt sich

$$\int f \,\mathrm{d}\mu = \lim_{n\to\infty} \int g_{n,n} \,\mathrm{d}\mu \le \lim_{n\to\infty} \int f_n \,\mathrm{d}\mu,$$

denn $g_{n,n} \le f_n$. □

Dieser Satz ist das Kernstück vieler Beweise von Aussagen über nichtnegative messbare Funktionen, siehe auch die Bemerkungen im Anschluss an Satz 3.1.12. Ein gutes Beispiel für diese Beweisstruktur ist der Beweis des Satzes 3.6.1. Man beachte, dass der Monotone Konvergenzsatz per Definition schon für Folgen *einfacher* Funktionen f_n gilt; er erweitert diese Aussage auf beliebige nichtnegative messbare Funktionen.

Eine Variante des Monotonen Konvergenzsatzes ist der folgende Satz (Beweis als Übungsaufgabe):

Korollar 3.4.2 (Satz von Beppo-Lévy). *Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge integrierbarer Funktionen mit $f_n \uparrow f$ μ-f. ü. für $n \to \infty$ und $\sup_{n\in\mathbb{N}} \int f_n \,\mathrm{d}\mu < \infty$. Dann ist auch f integrierbar, und es gilt*

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

Weitere Folgerungen aus dem Satz von der monotonen Konvergenz:

Korollar 3.4.3. *Für beliebige nichtnegative messbare Funktionen $f_1, f_2, f_3, \dots$ gilt*

$$\int \sum_{n\in\mathbb{N}} f_n \,\mathrm{d}\mu = \sum_{n\in\mathbb{N}} \int f_n \,\mathrm{d}\mu.$$

Insbesondere gilt für paarweise disjunkte messbare Mengen $A_1, A_2, A_3, \dots$ und jede nichtnegative messbare Funktion f:

$$\int_{\bigcup_{n\in\mathbb{N}} A_n} f \,\mathrm{d}\mu = \sum_{n\in\mathbb{N}} \int_{A_n} f \,\mathrm{d}\mu.$$

Satz 3.4.4 (Lemma von Fatou). *Es sei $(f_n)_{n\in\mathbb{N}}$ eine beliebige Folge nichtnegativer messbarer numerischer Funktionen. Dann gilt*

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

Beweis. Nach Definition des Limes Inferior konvergiert $\inf_{m\colon m\ge n} f_m$ monoton steigend gegen $\liminf_{n\to\infty} f_n$. Nach dem Monotonen Konvergenzsatz gilt

$$\int \liminf_{n\to\infty} f_n \,\mathrm{d}\mu = \lim_{n\to\infty} \int \inf_{m\colon m\ge n} f_m \,\mathrm{d}\mu \le \lim_{n\to\infty} \inf_{m\colon m\ge n} \int f_m \,\mathrm{d}\mu = \liminf_{n\to\infty} \int f_n \,\mathrm{d}\mu.$$

□

Bemerkung 3.4.5. Es ist leicht, Beispiele zu geben, in denen die Ungleichung im Lemma von Fatou strikt ist. Etwa für $(\Omega, \mathcal{F}, \mu) = (\mathbb{R}, \mathcal{B}, \lambda)$ kann man 'Zackenfunktionen' wählen, deren Graph die lineare Interpolation der Punkte $(0, 0)$, $(\frac{1}{n}, n)$ und $(\frac{2}{n}, 0)$ ist. ◇

Der wohl nützlichste und stärkste Satz über die Vertauschung von Grenzwert und Integral ist der folgende.

Satz 3.4.6 (Satz von Lebesgue, Satz von der majorisierten Konvergenz). *Es sei $(f_n)_{n\in\mathbb{N}}$ eine Folge von μ-integrierbaren Funktionen. Es existiere eine nichtnegative Funktion $g \in \mathcal{L}^1(\mu)$ mit $|f_n(\omega)| \le g(\omega)$ für alle $\omega \in \Omega$ und alle $n \in \mathbb{N}$. Falls $f(\omega) = \lim_{n\to\infty} f_n(\omega)$ für μ-fast alle $\omega \in \Omega$ existiert, so gilt*

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

Beweis. Durch Änderung auf einer Nullmenge können wir erreichen, dass $f_n(\omega) \to f(\omega)$ für *alle* $\omega \in \Omega$ gilt. Es gilt $f_n + g \ge 0$, und das Lemma von Fatou impliziert

$$\int f \,\mathrm{d}\mu = \int (f+g)\,\mathrm{d}\mu - \int g \,\mathrm{d}\mu \le \liminf_{n\to\infty} \int (f_n + g)\,\mathrm{d}\mu - \int g\,\mathrm{d}\mu \le \liminf_{n\to\infty} \int f_n \,\mathrm{d}\mu.$$

Wenn wir den selben Trick auf $g - f_n \ge 0$ anwenden, erhalten wir die komplementäre Ungleichung $\int f \,\mathrm{d}\mu \ge \limsup_{n\to\infty} \int f_n \,\mathrm{d}\mu$. □

Der Spezialfall eines diskreten Maßraums enthält aus der Analysis bekannte Sätze über Doppelfolgen:

Bemerkung 3.4.7 (Zahlenfolgen). Es ist eine der Stärken der Maßtheorie, dass viele bekannte Sätze über reelle Zahlenfolgen einen Spezialfall darstellen. Wendet man zum Beispiel den Satz von Lebesgue auf das Zählmaß auf $\mathbb{N}$ an, also $(\Omega, \mathcal{F}, \mu) = (\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu)$ mit dem Zählmaß $\mu = \sum_{n\in\mathbb{N}} \delta_n$ (siehe Beispiel 2.2.10), dann erhält man das folgende Resultat. Wir erinnern daran (siehe Beispiel 3.2.7), dass $\mathcal{L}^1(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu)$ nichts anderes ist als der bekannte Folgenraum $\ell^1(\mathbb{N})$. Der Satz von Lebesgue besagt dann, dass für jede Doppelfolge $(a_{i,n})_{i,n\in\mathbb{N}}$, die die Bedingungen

$$\begin{aligned} a_i &= \lim_{n\to\infty} a_{i,n} \text{ existiert für alle } i \in \mathbb{N}, \\ |a_{i,n}| &\le b_i \,\forall i, n \in \mathbb{N} \text{ mit } \sum_{i\in\mathbb{N}} b_i < \infty, \end{aligned}$$

erfüllt, die Summation mit dem Grenzwert vertauschbar ist, d. h. $\lim_{n\to\infty} \sum_{i\in\mathbb{N}} a_{i,n} = \sum_{i\in\mathbb{N}} a_i$ gilt. ◇

Wir ziehen noch zwei Folgerungen aus dem Satz von Lebesgue, deren Beweise Übungsaufgaben sind:

Korollar 3.4.8 (Stetigkeitslemma). *Es sei $f\colon \Omega \times \mathbb{R} \to \mathbb{R}$ eine Funktion und $x_0 \in \mathbb{R}$, so dass gelten:*

(a) Für jedes $x \in \mathbb{R}$ ist $f(\cdot, x) \in \mathcal{L}^1(\mu)$,

(b) für fast alle $\omega \in \Omega$ ist $f(\omega, \cdot)$ stetig in x_0,

(c) es gibt eine Umgebung $\mathcal{U}$ von x_0, so dass die Abbildung $\sup_{x\in\mathcal{U}} |f(\cdot, x)|$ in $\mathcal{L}^1(\mu)$ liegt.

Dann ist die Abbildung $\int f(\omega, \cdot)\,\mu(\mathrm{d}\omega)$ stetig in x_0.

Korollar 3.4.9 (Differenziationslemma). *Es sei $I \subset \mathbb{R}$ ein nichttriviales Intervall und $f\colon \Omega \times I \to \mathbb{R}$ eine Abbildung, so dass gelten:*

(a) Für jedes $x \in I$ ist $f(\cdot, x) \in \mathcal{L}^1(\mu)$,

(b) für fast alle $\omega \in \Omega$ ist $f(\omega, \cdot)$ differenzierbar (wir bezeichnen die Ableitung mit f'),

(c) die Abbildung $\sup_{x \in I} |f'(\cdot, x)|$ liegt im $\mathcal{L}^1(\mu)$.

Dann liegt $f'(\cdot, x)$ für jedes $x \in I$ in $\mathcal{L}^1(\mu)$, und die Funktion $F(x) = \int f(\omega, x)\, \mu(\mathrm{d}\omega)$ ist differenzierbar mit $F'(x) = \int f'(\omega, x)\, \mu(\mathrm{d}\omega)$.

Bemerkung 3.4.10 (Momenten erzeugende Funktion I). Wir betrachten die sogenannte *Momenten erzeugende Funktion* $M(t) = \int e^{tx}\, \mu(\mathrm{d}x)$ eines Maßes μ auf $(\mathbb{R}, \mathcal{B})$. Es sei vorausgesetzt, dass der Definitionsbereich $I = \{t \in \mathbb{R}\colon M(t) < \infty\}$ ein nichtleeres Inneres hat. Als eine Anwendung des Differenziationslemmas sieht man, dass M im Innern von I beliebig oft differenzierbar ist mit $M^{(k)}(t) = \int x^k e^{tx}\, \mu(\mathrm{d}x)$ für jedes $k \in \mathbb{N}_0$ und $t \in I^\circ$. Wenn insbesondere $0 \in I^\circ$, so existieren alle *Momente* $\int x^k\, \mu(\mathrm{d}x)$, $k \in \mathbb{N}_0$, von μ, und daher rührt der Name der Funktion M. Mit Hilfe der zweiten Ableitung erhält man als Übungsaufgabe insbesondere einen Beweis für die Konvexität der Abbildung $\log M$, siehe auch Bemerkung 3.5.8. $\diamond$

3.5 Die L^p-Räume

Es sei $(\Omega, \mathcal{F}, \mu)$ ein beliebiger Maßraum. In diesem Abschnitt führen wir die zur Potenz $p \geq 1$ integrierbaren Funktionen ein und leiten ein paar wichtige Ungleichungen her.

Für eine messbare numerische Funktion $f\colon \Omega \to \overline{\mathbb{R}}$ und $p \in [1, \infty]$ definieren wir

$$\|f\|_p = \begin{cases} \left(\int_\Omega |f|^p \,\mathrm{d}\mu\right)^{\frac{1}{p}}, & \text{falls } p < \infty, \\ \sup\{a \geq 0\colon \mu(|f| \geq a) > 0\}, & \text{falls } p = \infty, \end{cases}$$

und

$$\mathcal{L}^p(\Omega, \mathcal{F}, \mu) = \mathcal{L}^p(\mu) = \{f\colon f \text{ ist messbar und numerisch mit } \|f\|_p < \infty\}.$$

Dies ist offensichtlich eine Erweiterung der Definition von $\mathcal{L}^1$. Man überlegt sich als Übungsaufgabe, dass $\|f\|_\infty = \min\{c \in [0, \infty]\colon |f| \leq c\ \mu\text{-f.ü.}\}$, d. h. $\|f\|_\infty = c$ gilt genau dann, wenn es eine Nullmenge N gibt mit $\sup_{N^c} |f| = c$, d. h. $\|f\|_\infty$ ist das Supremum von $|f|$ außerhalb einer μ-Nullmenge. Man sieht auch leicht, dass $(\mathcal{L}^p, \|\cdot\|_p)$ ein Vektorraum ist. Für $p < \infty$ folgt dies aus $|f + g|^p \leq (2 \max\{f, g\})^p \leq 2^p(|f|^p + |g|^p)$ mit Korollar 3.2.5(i), und für $p = \infty$ mit der soeben erwähnten Charakterisierung von $\|\cdot\|_\infty$.

Eines unserer Ziele ist es, $(\mathcal{L}^p, \|\cdot\|_p)$ als einen normierten Raum aufzufassen. Immerhin ist $\|\cdot\|_p$ positiv homogen (d. h. $\|\alpha f\|_p = |\alpha| \|f\|_p$ für $\alpha \in \mathbb{R}$ und $f \in \mathcal{L}^p(\mu)$), und in Satz 3.5.10 werden wir die Gültigkeit der Dreiecksungleichung zeigen. Allerdings ist $\|\cdot\|_p$ nicht definit, denn aus Lemma 3.2.6 folgt: $\|f\|_p = 0 \iff f = 0$ μ-f. ü. (Man sagt, dass $\|\cdot\|_p$ eine *Halbnorm* auf dem $\mathcal{L}^p(\mu)$ ist.) Dies bringt uns auf die Idee, zu dem folgenden Quotientenraum überzugehen: Mit

$$\mathcal{N} = \{f \in \mathcal{L}^p(\mu)\colon f = 0\ \mu\text{-fast überall}\}$$

definieren wir

$$L^p(\Omega, \mathcal{F}, \mu) = L^p(\mu) = \mathcal{L}^p(\Omega, \mathcal{F}, \mu)/\mathcal{N},$$

also ist $L^p(\mu)$ die Menge aller Äquivalenzklassen $[f] = f + \mathcal{N} = \{f + g \colon g \in \mathcal{N}\}$ von $\mathcal{L}^p(\mu)$ bezüglich der Äquivalenzrelation 'Gleichheit μ-fast überall'. Auf dem Quotientenraum $L^p(\mu)$ werden die üblichen Operationen (Addition, skalare Multiplikation) vertreterweise erklärt, und daher ist der $L^p(\mu)$ ein Vektorraum. Für eine Klasse $[f] \in L^p(\mu)$ setzen wir $\|[f]\|_p = \|f\|_p$, und wiederum Lemma 3.2.6 sagt uns, dass diese Definition sinnvoll ist, d. h. nicht vom gewählten Vertreter f abhängt. Die Normeigenschaften sind ebenfalls leicht aus den Halbnormeigenschaften von $\|\cdot\|_p$ abzuleiten. Also ist der $L^p(\mu)$ ein normierter Raum. Im wichtigen Spezialfall $p = 2$ wird durch $\langle [f], [g] \rangle = \int fg \,\mathrm{d}\mu$ ein Skalarprodukt auf dem $L^2(\mu)$ definiert, das offensichtlich die Norm $\|\cdot\|_2$ erzeugt, also ist der $L^2(\mu)$ ein Vektorraum mit Skalarprodukt. Nun muss nur noch die Vollständigkeit des Raumes $L^p(\mu)$ gezeigt werden, und der Beweis des folgenden Satzes ist komplett.

Satz 3.5.1 (Riesz-Fischer). *Für jedes $p \in [1, \infty]$ ist $(L^p(\mu), \|\cdot\|_p)$ ein Banachraum. Im Spezialfall $p = 2$ ist $(L^2(\mu), \langle \cdot, \cdot \rangle)$ ein Hilbertraum.*

Ein *Banachraum* ist ein vollständiger normierter Vektorraum, und ein *Hilbertraum* ist ein vollständiger Vektorraum mit Skalarprodukt. Den Beweis von Satz 3.5.1 werden wir nicht ausformulieren, insbesondere die Vollständigkeit werden wir hier nicht beweisen (siehe dazu etwa [Kl06, Abschnitte 6 und 7] oder eine Funktionalanalysis-Vorlesung). Die Dreiecksungleichung für $\|\cdot\|_p$ wird in Satz 3.5.10 bewiesen, und daraus folgen auch die Vektorraumeigenschaften.

Bemerkung 3.5.2 (Der Raum $\ell^p(\mathbb{N})$). In Satz 3.5.1 ist als Spezialfall auch der Folgenraum

$$\ell^p(\mathbb{N}) = \Big\{ (x_n)_{n\in\mathbb{N}} \colon \sum_{n\in\mathbb{N}} |x_n|^p < \infty \Big\}, \qquad p \in [1, \infty),$$

aller Folgen reeller Zahlen, deren p-te Potenzen absolut summierbar sind, enthalten und natürlich auch der Raum $\ell^\infty(\mathbb{N})$ aller beschränkten Folgen. Wenn man nämlich μ als das Zählmaß $\sum_{n\in\mathbb{N}} \delta_n$ auf $\mathbb{N}$ wählt, entpuppt sich $\ell^p(\mathbb{N})$ als der $L^p(\mu)$. Man beachte, dass dieses Zählmaß kein endliches Maß ist. $\diamond$

Wir wollen im Folgenden die wichtigsten allgemeinen Ungleichungen im Zusammenhang mit Integralen und $\mathcal{L}^p$-Räumen vorstellen. Wir wählen einen Beweisweg, der über die Jensen'sche Ungleichung läuft, eine der wichtigsten Ungleichungen der Maß- und Wahrscheinlichkeitstheorie. Wir erinnern daran (siehe Bemerkung 3.2.2(f)), dass wir $\mathbb{E}(X)$ statt $\int X \,\mathrm{d}\mathbb{P}$ schreiben, wenn $\mathbb{P}$ ein Wahrscheinlichkeitsmaß ist.

Satz 3.5.3 (Jensen'sche Ungleichung). *Es sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum, und es sei $X \colon \Omega \to \mathbb{R}$ integrierbar. Ferner sei $I \subset \mathbb{R}$ ein Intervall und $\varphi \colon I \to \mathbb{R}$ eine konvexe Funktion. Es gelte $\mathbb{P}(X \notin I) = 0$. Dann gilt*

$$\mathbb{E}(\varphi \circ X) \geq \varphi\big(\mathbb{E}(X)\big).$$

Beweis. Da konvexe Funktionen immerhin im Inneren ihres Definitionsbereichs stetig sind, ist $\varphi \circ X$ messbar. Es ist zugelassen, dass $\mathbb{E}(\varphi \circ X) = \infty$ gilt.

In dem Fall, dass $\mathbb{E}(X)$ im Innern von I liegt, gibt es eine Gerade unterhalb des Graphen von φ, die durch den Punkt $(\mathbb{E}(X), \varphi(\mathbb{E}(X)))$ geht und eine Steigung α hat:

$$\varphi(y) \geq \alpha(y - \mathbb{E}(X)) + \varphi(\mathbb{E}(X)), \qquad y \in I.$$

Nun ersetzen wir y durch $X(\omega)$ und integrieren über ω, um zu erhalten:

$$\mathbb{E}(\varphi \circ X) = \int \varphi(X(\omega))\,\mathbb{P}(\mathrm{d}\omega) \geq \alpha \int (X(\omega) - \mathbb{E}(X))\,\mathbb{P}(\mathrm{d}\omega) + \int \varphi(\mathbb{E}(X))\,\mathbb{P}(\mathrm{d}\omega) = \varphi\big(\mathbb{E}(X)\big).$$

Im anderen Fall, wo $\mathbb{E}(X)$ ein Randpunkt von I ist, gilt $X = \mathbb{E}(X)$ $\mathbb{P}$-fast sicher, und die Behauptung gilt trivialerweise. $\square$

Daraus sieht man leicht als Übungsaufgabe, dass für endliche Maßräume bei wachsendem p die $\mathcal{L}^p$- und die L^p-Räume absteigend in einander enthalten sind:

Korollar 3.5.4. *Es sei $(\Omega, \mathcal{F}, \mu)$ ein endlicher Maßraum. Dann gelten für $1 \leq p \leq p' \leq \infty$ die Inklusionen $\mathcal{L}^{p'}(\mu) \subset \mathcal{L}^p(\mu)$ und $L^{p'}(\mu) \subset L^p(\mu)$. Falls $\mu(\Omega) = 1$, so ist die Abbildung $[1, \infty] \to [0, \infty]$, $p \mapsto \|f\|_p$ für jede messbare numerische Funktion f monoton nichtfallend.*

Bemerkung 3.5.5. Bei den Folgenräumen $\ell^p(\mathbb{N})$ (siehe Bemerkung 3.5.2) sind die Inklusionen umgekehrt. Genauer gilt Folgendes. Falls ein Maß μ die Eigenschaft hat, dass für ein $a > 0$ gilt: $\mu(A) \notin (0, a)$ für alle $A \in \mathcal{F}$, so gelten die Inklusionen $\mathcal{L}^{p'}(\mu) \subset \mathcal{L}^p(\mu)$ und $L^{p'}(\mu) \subset L^p(\mu)$ für $1 \leq p' \leq p \leq \infty$ (Übungsaufgabe). $\diamond$

Bemerkung 3.5.6 (Entropien). Die *relative Entropie* eines Wahrscheinlichkeitsmaßes $\mathbb{P}$ bezüglich eines Wahrscheinlichkeitsmaßes $\widetilde{\mathbb{P}}$ definieren wir durch

$$H\big(\mathbb{P} \mid \widetilde{\mathbb{P}}\big) = \int \log \frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\widetilde{\mathbb{P}}}\,\mathrm{d}\mathbb{P},$$

falls die Dichte $f = \frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\widetilde{\mathbb{P}}}$ existiert (für diesen Begriff siehe Definition 3.6.6; dies heißt, dass $\mathbb{P}(A) = \int_A f\,\mathrm{d}\widetilde{\mathbb{P}}$ für jede messbare Menge A), und $H(\mathbb{P} \mid \widetilde{\mathbb{P}}) = \infty$ sonst. Als eine einfache Übungsaufgabe wendet man die Jensen'sche Ungleichung auf die konvexe Abbildung $x \mapsto x \log x$ an, um zu zeigen, dass die Entropie nichtnegativ ist. Mit einer anderen Anwendung der Jensen'schen Ungleichung kann man sehen, dass $H(\cdot \mid \widetilde{\mathbb{P}})$ eine konvexe Abbildung ist. $\diamond$

Eine weitere wichtige, berühmte Ungleichung ist die folgende, die wir nun als eine Anwendung der Jensen'schen Ungleichung beweisen. Wie zu Beginn sei $(\Omega, \mathcal{F}, \mu)$ ein beliebiger Maßraum.

Satz 3.5.7 (Hölder'sche Ungleichung). *Es seien $p, q \in [1, \infty]$ mit $\frac{1}{p} + \frac{1}{q} = 1$. Ferner seien $f \in \mathcal{L}^p(\mu)$ und $g \in \mathcal{L}^q(\mu)$. Dann gelten $fg \in \mathcal{L}^1(\mu)$ und*

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

Beweis. Der Fall $p = \infty$ ist sehr leicht; wir behandeln nur den Fall $1 < p < \infty$. Wir dürfen auch voraussetzen, dass $\|f\|_p > 0$ und $\|g\|_q > 0$. Wir betrachten die Funktionen $\widetilde{f} = (|f|/\|f\|_p)^p$ und $\widetilde{g} = (|g|/\|g\|_q)^q$. Dann gilt $\int \widetilde{f}\,\mathrm{d}\mu = \int \widetilde{g}\,\mathrm{d}\mu = 1$. Daher ist das Maß $\mathbb{P} = \widetilde{f}\,\mathrm{d}\mu$ ein Wahrscheinlichkeitsmaß (wobei wir $\mathbb{P}$ durch $\mathbb{P}(A) = \int_A \widetilde{f}\,\mathrm{d}\mu$ definieren). Ferner sieht man, dass

$$\int \frac{|fg|}{\|f\|_p \|g\|_q}\,\mathrm{d}\mu = \int \widetilde{f}^{1/p} \widetilde{g}^{1/q}\,\mathrm{d}\mu = \int \Big(\frac{\widetilde{g}}{\widetilde{f}}\Big)^{1/q} \widetilde{f}\,\mathrm{d}\mu = \mathbb{E}\Big[\Big(\frac{\widetilde{g}}{\widetilde{f}}\Big)^{1/q}\Big],$$

wobei wir $\mathbb{E}$ für den Erwartungswert bezüglich $\mathbb{P}$ geschrieben haben. (Man beachte auch, dass der Quotient $\mathbb{P}$-fast sicher wohldefiniert ist.) Nun benutzen wir die Konvexität der Abbildung $x \mapsto x^q$ und erhalten mit Hilfe der Jensen'schen Ungleichung:

$$\int \frac{|fg|}{\|f\|_p \|g\|_q} \, \mathrm{d}\mu = \mathbb{E}\Big[\Big(\frac{\widetilde{g}}{\widetilde{f}}\Big)^{1/q}\Big] \leq \mathbb{E}\Big[\frac{\widetilde{g}}{\widetilde{f}}\Big]^{1/q} = \Big(\int \widetilde{g} \, \mathrm{d}\mu\Big)^{1/q} = 1,$$

also die Hölder'sche Ungleichung. (Genau genommen, dürfen wir die Jensen'sche Ungleichung nur anwenden, wenn wir wissen, dass der Erwartungswert von $(\widetilde{g}/\widetilde{f})^{1/q}$ endlich ist, aber indem man den obigen Beweis zunächst nur für nichtnegative einfache Funktionen durchführt und dann zu monotonen Grenzwerten übergeht, erhält man den allgemeinen Fall.) □

Es sei noch bemerkt, dass ein weiterer, sehr weit verbreiteter Beweis der Hölder'schen Ungleichung über die *Young'sche Ungleichung* läuft (Übungsaufgabe), die besagt, dass für $p, q > 1$ mit $1 = \frac{1}{p} + \frac{1}{q}$ und alle $x, y \geq 0$ gilt: $xy \leq \frac{x^p}{p} + \frac{y^q}{q}$.

Bemerkung 3.5.8 (Momenten erzeugende Funktion II). Wie in Bemerkung 3.4.10 betrachten wir wieder die Momenten erzeugende Funktion $M(t) = \int \mathrm{e}^{tx} \, \mu(\mathrm{d}x) \in [0, \infty]$ eines Maßes μ auf $(\mathbb{R}, \mathcal{B})$. Es sei $I = \{t \in \mathbb{R} \colon M(t) < \infty\}$ nicht leer. Eine Anwendung der Hölder'schen Ungleichung zeigt, dass I ein Intervall ist und dass die Abbildung $\log M$ konvex auf I ist (Übungsaufgabe). ◇

Der symmetrische Spezialfall der Hölder-Ungleichung ist ebenfalls berühmt:

Korollar 3.5.9 (Cauchy-Schwarz-Ungleichung). *Für je zwei Funktionen $f, g \in \mathcal{L}^2(\mu)$ gelten $fg \in \mathcal{L}^1(\mu)$ und $\|fg\|_1 \leq \|f\|_2 \cdot \|g\|_2$.*

Als eine Anwendung der Hölder'schen Ungleichung zeigen wir nun die Dreiecksungleichung für $\|\cdot\|_p$:

Satz 3.5.10 (Minkowski-Ungleichung). *Es sei $p \in [1, \infty]$. Dann gilt für alle $f, g \in \mathcal{L}^p(\mu)$:*

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

Beweis. Die Fälle $p = \infty$ und $p = 1$ sind leicht bzw. schon bekannt, so dass wir nur den Fall $1 < p < \infty$ behandeln. Die Dreiecksungleichung impliziert, dass $\|f + g\|_p \leq \||f| + |g|\|_p$, sodass wir annehmen dürfen, dass $f \geq 0$ und $g \geq 0$. Wegen $(f + g)^p \leq [2 \max\{f, g\}]^p \leq 2^p (f^p + g^p)$ ist schon klar, dass $\|f + g\|_p < \infty$. Wir schreiben

$$\int (f + g)^p \, \mathrm{d}\mu = \int f (f + g)^{p-1} \, \mathrm{d}\mu + \int g (f + g)^{p-1} \, \mathrm{d}\mu$$

und wenden die Hölder'sche Ungleichung (mit einem $q > 1$, so dass $\frac{1}{p} + \frac{1}{q} = 1$) auf die beiden Integrale an, um zu erhalten:

$$\int (f + g)^p \, \mathrm{d}\mu \leq \|f\|_p \|(f + g)^{p-1}\|_q + \|g\|_p \|(f + g)^{p-1}\|_q = \big(\|f\|_p + \|g\|_p\big) \Big(\int (f + g)^p \, \mathrm{d}\mu\Big)^{1/q}.$$

Elementare Umformungen lassen die Aussage folgen. □

3.6 Bildmaße und Dichten

Mit Hilfe von messbaren Abbildungen kann man aus gegebenen Maßen weitere Maße konstruieren. Wir stellen die beiden wichtigsten Prinzipien hier vor; es handelt sich um *Transportieren* mittels einer messbaren Funktion bzw. um das *Transformieren* mittels einer nichtnegativen messbaren Funktion, der *Dichte*. Insbesondere werden wir in diesem Abschnitt auch die allgemeine Transformationsformel für Lebesgue-integrierbare Funktionen behandeln.

Bildmaße

Es sei $(\Omega, \mathcal{F}, \mu)$ ein Maßraum und $(\Omega', \mathcal{F}')$ ein weiterer messbarer Raum. Ferner sei $f\colon \Omega \to \Omega'$ eine messbare Funktion. Wir definieren $\mu \circ f^{-1}\colon \mathcal{F}' \to [0, \infty]$ durch

$$\mu \circ f^{-1}(A) = \mu\big(f^{-1}(A)\big), \qquad A \in \mathcal{F}', \tag{3.6.1}$$

wobei $f^{-1}(A)$ wie immer das Urbild von A unter f bezeichnet.

Satz 3.6.1. *(a) $\mu \circ f^{-1}$ ist ein Maß. Falls μ ein Wahrscheinlichkeitsmaß ist, so auch $\mu \circ f^{-1}$.*

(b) Für jede nichtnegative messbare numerische Funktion $\phi\colon \Omega' \to [0, \infty]$ gilt

$$\int_{\Omega'} \phi \, \mathrm{d}(\mu \circ f^{-1}) = \int_{\Omega} (\phi \circ f) \, \mathrm{d}\mu. \tag{3.6.2}$$

(c) Für jede messbare Funktion $\phi\colon \Omega' \to \mathbb{R}$ gilt

$$\phi \in \mathcal{L}^1(\Omega', \mathcal{F}', \mu \circ f^{-1}) \qquad \Longleftrightarrow \qquad \phi \circ f \in \mathcal{L}^1(\Omega, \mathcal{F}, \mu),$$

und in diesem Fall gilt (3.6.2) *ebenfalls.*

Beweis. (a) folgt aus den allgemein gültigen Rechenregeln für die Abbildung $f^{-1}\colon \mathcal{F}' \to \mathcal{F}$.

Um die Aussage (b) zu beweisen, werden wir Satz 3.1.12 anwenden und müssen also zeigen, dass die Gleichung (3.6.2) (i) sich auf Linearkombinationen überträgt, (ii) sich auf monotone Limiten überträgt und (iii) für Indikatorfunktionen auf messbaren Mengen gilt. Diese drei Dinge sind leicht zu sehen: (i) stimmt, da beide Seiten der Gleichung (3.6.2) linear in ϕ sind, (ii) folgt direkt aus dem Monotonen Konvergenzsatz (Satz 3.4.1), und (iii) ist nach Definition (3.6.1) vorausgesetzt.

Aussage (c) ergibt sich direkt aus einer Anwendung von (b) auf $|\phi|$, und die Formel (3.6.2) für allgemeine integrierbare Funktionen zeigt man leicht durch Zerlegung in Positiv- und Negativteil. □

Definition 3.6.2 (Bildmaß). *Das Maß $\mu \circ f^{-1}$ heißt das* durch f induzierte Maß *oder das* Bildmaß *von μ unter f. Es sind auch die Schreibweisen μf^{-1} und $f(\mu)$ gebräuchlich.*

Bemerkung 3.6.3. Der Beweis von Satz 3.6.1 ist ein Paradebeispiel für ein allgemeines Beweisverfahren für Aussagen über messbare Funktionen. Wir werden in Zukunft solche Beweise nicht mehr voll ausformulieren, sondern Wendungen benutzen wie ‘beweist man auf die übliche Weise’. Umgangsprachlich spricht man auch manchmal vom ‘Hochziehen’. ◇

Falls das Bildmaß $\mu\circ f^{-1}$ σ-endlich ist, so auch μ, aber die Umkehrung gilt im Allgemeinen nicht (Übungsaufgabe).

Beispiel 3.6.4. Sei μ ein Maß auf dem $\mathbb{Z}^2$ und $f\colon \mathbb{Z}^2\to\mathbb{Z}$ definiert durch $f(x,y)=x+y$. Dann ist dass Bildmaß von μ unter f gegeben durch

$$\mu\circ f^{-1}(\{z\})=\sum_{x\in\mathbb{Z}}\mu(\{(x,z-x)\})=\sum_{x\in\mathbb{Z}}\mu(\{(z-x,x)\}).$$

◇

Beispiel 3.6.5. Sei $A\in\mathrm{GL}(\mathbb{R},d)$ eine reguläre Matrix und λ das Lebesgue-Maß auf dem $\mathbb{R}^d$. Dann ist das Bildmaß von λ unter der Abbildung $A\colon\mathbb{R}^d\to\mathbb{R}^d$ gegeben durch $\lambda\circ A^{-1}=|\det(A)|^{-1}\lambda$. Dies folgt daraus, dass für jeden Quader $Q=\prod_{i=1}^d[a_i,b_i]$ der Spat $A^{-1}(Q)$ das Volumen $|\det(A^{-1})|\prod_{i=1}^d(b_i-a_i)=|\det(A)|^{-1}\lambda(Q)$ besitzt, siehe Beispiel 1.2.7, wo wir A statt A^{-1} schrieben. Also stimmen die beiden Maße $\lambda\circ A^{-1}$ und $|\det(A)|^{-1}\lambda$ auf der Menge aller Quader überein, und dies ist ein durchschnittsstabiles Erzeugendensystem der Borel-σ-Algebra. Nach dem Eindeutigkeitssatz (Satz 2.3.8) ist also $\lambda\circ A^{-1}=|\det(A)|^{-1}\lambda$. ◇

Man beachte, dass aus Bemerkung 3.6.5 die Erweiterung des Transformationssatzes 1.2.5 für lineare Transformationen auf das Lebesgue-Integral folgt, wie wir im Beweis des Satzes 3.6.11 erläutern werden.

Dichten

Nun kommt die nach Korollar 2.4.4 angekündigte starke Verallgemeinerung des Begriffs der Dichte. Wir erinnern uns, dass $\mathcal{B}$ die Borel-σ-Algebra auf $\mathbb{R}$ bezeichnet.

Definition 3.6.6 (Dichten). *Es seien μ und ν zwei Maße auf einem messbaren Raum $(\Omega,\mathcal{F})$. Eine $\mathcal{F}$-$\mathcal{B}$-messbare Abbildung $\varphi\colon\Omega\to[0,\infty)$ heißt eine* Dichte *von ν bezüglich μ, wenn $\nu(A)=\int_A\varphi\,\mathrm{d}\mu$ für alle $A\in\mathcal{F}$ gilt. Wir schreiben dann auch $\varphi=\frac{\mathrm{d}\nu}{\mathrm{d}\mu}$ und $\nu=\varphi\,\mathrm{d}\mu$.*

Wenn φ eine Dichte von ν bezüglich μ ist, dann ist jede messbare Abbildung $f\colon\Omega\to\mathbb{R}$ genau dann ν-integrierbar, wenn $f\varphi$ integrierbar bezüglich μ ist, und dann gilt $\int f\,\mathrm{d}\nu=\int f\varphi\,\mathrm{d}\mu$. Dies gilt per Definition für Indikatorfunktionen f und wird standardmäßig auf beliebige messbare Funktionen erweitert (siehe Bemerkung 3.6.3).

Lemma 3.6.7 (Eindeutigkeit der Dichte). *Es seien μ und ν zwei Maße auf einem messbaren Raum $(\Omega,\mathcal{F})$. Der Maßraum $(\Omega,\mathcal{F},\nu)$ sei σ-endlich. Falls eine Dichte von ν bezüglich μ existiert, so ist sie eindeutig bis auf Gleichheit μ-f. ü.*

Beweis. Seien φ_1 und φ_2 zwei Dichten von ν bezüglich μ. Wir betrachten die messbare Menge $A=\{\varphi_1>\varphi_2\}=\{\omega\in\Omega\colon\varphi_1(\omega)>\varphi_2(\omega)\}$. Nach Voraussetzung existiert eine Folge messbarer Mengen $\Omega_n\subset\Omega$ mit $\nu(\Omega_n)<\infty$ und $\Omega_n\uparrow\Omega$ für $n\to\infty$. Dann gilt für jedes $n\in\mathbb{N}$

$$\int\mathbb{1}_{A\cap\Omega_n}(\varphi_1-\varphi_2)\,\mathrm{d}\mu=\mu\circ\varphi_1^{-1}(A\cap\Omega_n)-\mu\circ\varphi_2^{-1}(A\cap\Omega_n)=\nu(A\cap\Omega_n)-\nu(A\cap\Omega_n)=0.$$

Im Grenzwert $n\to\infty$ erhält man wegen $\mathbb{1}_A(\varphi_1-\varphi_2)\ge 0$ mit dem Monotonen Konvergenzsatz (Satz 3.4.1), dass $\int\mathbb{1}_A(\varphi_1-\varphi_2)\,\mathrm{d}\mu=0$. Wegen $\mathbb{1}_A(\varphi_1-\varphi_2)\ge 0$ folgt mit Lemma 3.2.6(b), dass

$\mathbb{1}_A(\varphi_1 - \varphi_2) = 0$ μ-f. ü., also $\varphi_1 \le \varphi_2$ μ-f. ü. Die komplementäre Ungleichung $\varphi_1 \ge \varphi_2$ μ-f. ü. folgt analog. $\square$

Definition 3.6.8 (Absolutstetigkeit, Äquivalenz). *Wir nennen ein Maß ν* absolutstetig *bezüglich eines Maßes μ, wenn jede μ-Nullmenge auch eine ν-Nullmenge ist, also falls aus $\mu(A) = 0$ folgt: $\nu(A) = 0$. In diesem Fall schreiben wir $\nu \ll \mu$. Falls $\nu \ll \mu$ und $\mu \ll \nu$ gelten, so nennen wir ν und μ* äquivalent *zu einander und schreiben $\nu \approx \mu$. Falls eine Menge $A \in \mathcal{F}$ existiert mit $\mu(A) = 0$ und $\nu(A^c) = 0$, so nennen wir μ* singulär *zu ν.*

Wenn ν eine Dichte φ bezüglich μ hat, so sind offensichtlich alle μ-Nullmengen auch ν-Nullmengen, denn für jede messbare Menge A mit $\mu(A) = 0$ gilt $\nu(A) = \int_A \varphi \, \mathrm{d}\mu = \int \varphi \mathbb{1}_A \, \mathrm{d}\mu = 0$, da der Integrand μ-fast sicher Null ist. Der berühmte *Satz von Radon-Nikodym* dreht die Aussage um: Falls $\nu \ll \mu$, so existiert eine Dichte von ν bezüglich μ. Diese Dichte wird oft die *Radon-Nikodym-Ableitung* oder *Radon-Nikodym-Dichte* von ν bezüglich μ genannt.

Satz 3.6.9 (Radon-Nikodym). *Seien μ und ν zwei σ-endliche Maße auf $(\Omega, \mathcal{F})$ mit $\nu \ll \mu$. Dann existiert eine Dichte von ν bezüglich μ.*

Beweis. Zunächst beweisen wir die Aussage unter der stärkeren Voraussetzung, dass μ ein endliches Maß ist und $\nu \le \mu$ gilt, also $\nu(A) \le \mu(A)$ für jedes $A \in \mathcal{F}$. Ein elegantes Argument, das den *Riesz'schen Darstellungssatz*[2] benutzt, geht wie folgt. Zunächst sieht man mit Hilfe eines Standardbeweises (siehe Bemerkung 3.6.3), dass wir wegen $\nu \le \mu$ auch $\int g \, \mathrm{d}\nu \le \int g \, \mathrm{d}\mu$ für jede nichtnegative messbare numerische Funktion $g \colon \Omega \to [0, \infty]$ haben. Daraus folgt für jedes $f \in \mathcal{L}^2(\mu)$:

$$\Big| \int_\Omega f \, \mathrm{d}\nu \Big| \le \int_\Omega |f| \, \mathrm{d}\nu \le \Big(\int f^2 \, \mathrm{d}\nu \Big)^{1/2} \nu(\Omega)^{1/2} \le \sqrt{\nu(\Omega)} \Big(\int f^2 \, \mathrm{d}\mu \Big)^{1/2} = \sqrt{\nu(\Omega)} \|f\|_{L^2(\mu)},$$

wobei wir die Cauchy-Schwarz-Ungleichung (siehe Korollar 3.5.9) benutzten. Wegen $\nu(\Omega) \le \mu(\Omega) < \infty$ zeigt diese Ungleichungskette, dass $\mathcal{L}^2(\mu) \subset \mathcal{L}^2(\nu) \subset \mathcal{L}^1(\nu)$ gilt und dass durch $f \mapsto \int_\Omega f \, \mathrm{d}\nu$ eine lineare und stetige Abbildung $\mathcal{L}^2(\mu) \to \mathbb{R}$ definiert wird.

Nach dem Riesz'schen Darstellungssatz existiert ein $h \in \mathcal{L}^2(\mu)$ mit $\int_\Omega f \, \mathrm{d}\nu = \langle f, h \rangle = \int_\Omega f h \, \mathrm{d}\mu$ für alle $f \in \mathcal{L}^2(\mu)$. Wenn wir $f = \mathbb{1}_A$ für $A \in \mathcal{F}$ wählen, sehen wir, dass $\nu(A) = \langle \mathbb{1}_A, h \rangle = \int_A h \, \mathrm{d}\mu$ gilt, h also eine Dichte von ν bezüglich μ ist. Also ist die Aussage im Fall $\nu \le \mu$ und $\mu(\Omega) < \infty$ bewiesen.

Nun beweisen wir die Aussage im Fall endlicher Maße ν und μ mit $\nu \ll \mu$. Nach dem Vorherigen existieren die Dichten $g = \frac{\mathrm{d}\mu}{\mathrm{d}(\mu+\nu)}$ und $h = \frac{\mathrm{d}\nu}{\mathrm{d}(\mu+\nu)}$. Die Menge $N = \{g = 0\}$ ist wegen $\mu(N) = \int_N g \, \mathrm{d}(\mu + \nu) = 0$ eine μ-Nullmenge, also auch eine ν-Nullmenge. Die Funktion $f \colon \Omega \to [0, \infty)$, definiert durch $f(\omega) = h(\omega)/g(\omega)$ für $\omega \in \Omega \setminus N$ und $f(\omega) = 0$ sonst, ist nichtnegativ und messbar und erfüllt für jedes $A \in \mathcal{F}$:

$$\nu(A) = \nu(A \cap N^c) = \int_{A \cap N^c} h \, \mathrm{d}(\mu + \nu) = \int_{A \cap N^c} f g \, \mathrm{d}(\mu + \nu) = \int_{A \cap N^c} f \, \mathrm{d}\mu = \int_A f \, \mathrm{d}\mu,$$

ist also eine Dichte von ν bezüglich μ.

[2] Der Riesz'sche Darstellungssatz besagt, dass zu jeder stetigen linearen Abbildung $F \colon H \to \mathbb{R}$ auf einem reellen Hilbertraum H ein $h \in H$ existiert mit $F(f) = \langle f, h \rangle$ für jedes $f \in H$.

Nun zeigen wir die Aussage für alle endlichen Maße μ und alle Maße ν mit $\nu \ll \mu$. Wir betrachten

$$\alpha = \sup\big\{\mu(B) \colon B \in \mathcal{F}, \nu(B) < \infty\big\} \in [0, \infty).$$

Dann gibt es eine wachsende Folge $(B_n)_{n\in\mathbb{N}}$ in $\mathcal{F}$ mit $\nu(B_n) < \infty$ für jedes $n \in \mathbb{N}$ und $\mu(B_n) \uparrow \alpha$ und $B_n \uparrow B$ mit $\mu(B) = \alpha$. Man sieht, dass für jede Menge $A \in \mathcal{F}$ mit $A \subset B^{\mathrm{c}}$ entweder $\mu(A) = \nu(A) = 0$ gilt oder $\mu(A) = 0$ und $\nu(A) = \infty$, denn falls $\nu(A) < \infty$, so ist $\alpha + \mu(A) = \lim_{n\to\infty} \mu(B_n \cup A) \leq \alpha$, denn $\nu(B_n \cup A) < \infty$. Die Mengen $C_n = B_n \setminus B_{n-1}$ (wir haben $B_0 = \emptyset$ gesetzt) sind disjunkt, haben endliches ν-Maß und schöpfen B aus. Wir setzen $\nu_n = \mathbb{1}_{C_n}\,\mathrm{d}\nu$ für $n \in \mathbb{N}$ und $\nu_{B^{\mathrm{c}}} = \mathbb{1}_{B^{\mathrm{c}}}\,\mathrm{d}\nu$, dann sind die ν_n endliche Maße auf $\mathcal{F}$ mit $\nu_n \ll \mu$. Nach dem zweiten Schritt dieses Beweises gibt es die Dichten $f_n = \frac{\mathrm{d}\nu_n}{\mathrm{d}\nu}$. Außerdem ist $\infty\mathbb{1}_{B^{\mathrm{c}}}$ eine Dichte von $\nu_{B^{\mathrm{c}}}$ bezüglich μ, und es folgt

$$\nu = \sum_{n\in\mathbb{N}} \nu_n + \nu_{B^{\mathrm{c}}} = \Big(\sum_{n\in\mathbb{N}} f_n + \infty\mathbb{1}_{B^{\mathrm{c}}}\Big)\,\mathrm{d}\mu,$$

also ist der Term in den Klammern eine Dichte von ν bezüglich μ.

Den letzten Schritt des Beweises (Erweiterung auf beliebige σ-endliche Maße ν und μ mit $\nu \ll \mu$) führen wir nicht mehr aus; dies ist analog zu dem entsprechenden Beweisteil im Eindeutigkeitssatz, Satz 2.3.8. □

Diskrete Maße (siehe Beispiel 2.2.10) haben keine Dichten bezüglich des Lebesgue-Maßes. Ein Maß kann durchaus aus einem diskreten und einem absolutstetigen Anteil bestehen. Es folgen die wichtigsten Beispiele mit Dichten:

Beispiel 3.6.10 (Lebesgue-Dichten). Wichtige Beispiele von Wahrscheinlichkeitsmaßen auf $\mathbb{R}$ mit Lebesgue-Dichten sind die folgenden:

Normalverteilung $\mathcal{N}(a, \sigma^2)$	$\varphi_{a,\sigma^2}(x) = \dfrac{1}{\sqrt{2\pi\sigma^2}} \exp\Big\{-\dfrac{(x-a)^2}{2\sigma^2}\Big\}$	Parameter $a \in \mathbb{R}, \sigma > 0$,
Exponentialverteilung	$\varphi_\alpha(x) = \alpha e^{-\alpha x}\mathbb{1}_{[0,\infty)}(x)$,	Parameter $\alpha > 0$,
Cauchy-Verteilung	$\varphi_c(x) = \dfrac{c}{\pi(c^2 + x^2)}$,	Parameter $c > 0$.

◇

Nun formulieren wir die allgemeine Version des Transformationssatzes, die den Satz 1.2.8 auf Lebesgue-integrierbare Funktionen erweitert:

Satz 3.6.11 (Transformationsformel für Lebesgue-integrierbare Funktionen). *Seien U und V zwei offene Mengen im $\mathbb{R}^d$ und $\varphi \colon U \to V$ eine $\mathcal{C}^1$-invertierbare Abbildung. Dann ist eine Funktion $f \colon V \to \overline{\mathbb{R}}$ genau dann Lebesgue-integrierbar über V, wenn die Abbildung $f \circ \varphi |\det(D\varphi)|$ über U integrierbar ist, und dann gilt*

$$\int_U f(\varphi(x))|\det(D\varphi(x))|\,\lambda(\mathrm{d}x) = \int_V f(y)\,\lambda(\mathrm{d}y). \tag{3.6.3}$$

Das Bildmaß der Abbildung φ^{-1} unter λ (also das Maß $A \mapsto \lambda(\varphi(A))$) ist gegeben durch $|\det(D\varphi(x))|\,\lambda(\mathrm{d}x)$, d. h. als das Maß mit Lebesgue-Dichte $x \mapsto |\det(D\varphi(x))|$.

Insbesondere gilt für jede reguläre Matrix A und für jede Lebesgue-integrierbare Funktion $f\colon \mathbb{R}^d \to \overline{\mathbb{R}}$:

$$\int_{\mathbb{R}^d} f(Ax)|\det(A)|\,\lambda(\mathrm{d}x) = \int_{\mathbb{R}^d} f(y)\,\lambda(\mathrm{d}y). \tag{3.6.4}$$

Nachzutragen ist noch, dass für nichtnegative Funktionen f die Beziehung (3.6.3) ohne weiteren Zusatz gilt, aber beide Seiten können ∞ sein. Das Einzige, was im Beweis des Satzes 3.6.11 zu zeigen ist, ist die Identifikation des Bildmaßes, denn dann folgt die Aussage in (3.6.3) nach der Standardargumentation schrittweise für einfache Funktionen, nichtnegative messbare Funktionen und integrierbare Funktionen.

Beweisskizze für Satz 3.6.11. Der Beweis von (3.6.4) besteht aus einer Anwendung der Beziehung $\int g\,\mathrm{d}(\lambda \circ A^{-1}) = \int g|\det(A)|^{-1}\,\mathrm{d}\lambda$ (siehe Beispiel 3.6.5) auf $g = f \circ A$ und einem Übergang von A zu A^{-1}. Die Erweiterung auf beliebige $\mathcal{C}^1$-invertierbare Abbildungen φ wird ähnlich wie im Beweis des Satzes 1.2.8 durch lokale Approximation durchgeführt. Bezüglich φ wird die selbe Strategie verfolgt. Die Funktion f kann als eine Indikatorfunktion auf einer Menge A angenommen werden, und A wird angenähert durch eine abzählbare Vereinigung immer kleinerer achsenparalleler Quader. Die Details sind sehr ähnlich jenen des Beweises des Satzes 1.2.8; siehe etwa [K02, Kap. 9]. □

Als eine Anwendung können wir den Begriff einer Normalverteilung stark erweitern:

Beispiel 3.6.12 (mehrdimensionale Normalverteilung). Die *Standardnormalverteilung* $\mathcal{N}$ auf dem messbaren Raum $(\mathbb{R}^d, \mathcal{B}_d)$ ist definiert als das Maß, das die Dichte

$$\varphi(x) = \frac{\mathrm{e}^{-\frac{1}{2}\|x\|^2}}{(2\pi)^{d/2}} = (2\pi)^{-d/2}\exp\Big\{-\frac{1}{2}\sum_{i=1}^d x_i^2\Big\}, \qquad x = (x_1, \dots, x_d) \in \mathbb{R}^d,$$

besitzt. Jedes Bildmaß $\mathcal{N} \circ f^{-1}$ mit einer affinen Abbildung $f\colon \mathbb{R}^d \to \mathbb{R}^d$, d. h. $f(x) = Ax + b$ mit einer $d \times d$-Matrix A und einem Vektor $b \in \mathbb{R}^d$, heißt eine *Normalverteilung*. Dieses Maß $\mathcal{N} \circ f^{-1}$ besitzt genau dann eine Dichte, wenn A invertierbar ist, und zwar ist dann eine Dichte gegeben durch

$$\varphi_{b,\Sigma}(x) = \frac{1}{\sqrt{(2\pi)^d \det \Sigma}}\exp\Big\{-\frac{1}{2}(x-b)^{\mathrm{T}}\Sigma^{-1}(x-b)\Big\}, \qquad x \in \mathbb{R}^d,$$

wobei $\Sigma = AA^{\mathrm{T}}$. Falls nämlich A nicht invertierbar ist, ist das Bild $f(\mathbb{R}^d)$ in einer affinen Hyperebene enthalten, besitzt also das Maß Null unter λ_d, aber es gilt $\mathcal{N} \circ f^{-1}(f(\mathbb{R}^d)) = 1$. Also hat $\mathcal{N} \circ f^{-1}$ keine Lebesgue-Dichte. Wenn andererseits A invertierbar ist, so ist auch $f\colon \mathbb{R}^d \to \mathbb{R}^d$ bijektiv, und (3.6.4) (mit Σ an Stelle von A) impliziert die Aussage.

Daher schreibt man oft auch $\mathcal{N}(b, \Sigma)$ an Stelle von $\mathcal{N} \circ f^{-1}$, und dann ist $\mathcal{N} = \mathcal{N}(0, I)$, wobei I die $d \times d$-Einheitsmatrix ist. ◇

Mit Hilfe von Satz 3.6.11 können wir nun endlich die letzten Unklarheiten aus den Beispielen 1.2.11 und 1.2.12 beseitigen. Wir geben eine d-dimensionale Version dieser Beispiele.

Lemma 3.6.13 (d-dimensionale Polarkoordinaten). *Es seien $d \geq 2$ und $U = (0,\infty) \times (0,\pi)^{d-2} \times (0, 2\pi)$ und $V = \mathbb{R}^d \setminus H_d$ der 'geschlitzte' $\mathbb{R}^d$, wobei $H_d = \{y \in \mathbb{R}^d : y_{d-1} \geq 0, y_d = 0\}$ eine halbe Hyperebene ist. Dann wird durch $\Phi(r, \varphi_1, \dots, \varphi_{d-1}) = y = (y_1, \dots, y_d)$ mit*

$$y_k = r\Big(\prod_{i=1}^{k-1}\sin(\varphi_i)\Big) \times \begin{cases} \cos(\varphi_k), & \textit{falls } k < d,\\ 1 & \textit{sonst,}\end{cases}$$

eine $\mathcal{C}^1$-invertierbare Abbildung $\Phi\colon U \to V$ definiert mit Funktionaldeterminante

$$\det\big(D\Phi(r,\varphi_1,\dots,\varphi_{d-1})\big) = r^{d-1}\prod_{i=1}^{d-2}\sin^{d-1-i}(\varphi_i).$$

Eine Funktion $f\colon \mathbb{R}^d \to \overline{\mathbb{R}}$ ist genau dann über den $\mathbb{R}^d$ Lebesgue-integrierbar, wenn die Abbildung $(r,\varphi_1,\dots,\varphi_{d-1}) \mapsto f(\Phi(r,\varphi_1,\dots,\varphi_{d-1}))r^{d-1}\prod_{i=1}^{d-2}\sin^{d-1-i}(\varphi_i)$ über V integrierbar ist, und in diesem Fall gilt

$$\int_{\mathbb{R}^d} f(x)\,\lambda(\mathrm{d}x) = \int_U f(\Phi(r,\varphi_1,\dots,\varphi_{d-1}))r^{d-1}\prod_{i=1}^{d-2}\sin^{d-1-i}(\varphi_i)\,\lambda\big(\mathrm{d}(r,\varphi_1,\dots,\varphi_{d-1})\big). \tag{3.6.5}$$

Beweisskizze. Man beachte, dass $y_k = (\sum_{i=k}^d y_i^2)^{1/2}\cos(\varphi_k)$ für $k=1,\dots,d-1$ gilt, während $y_d = (y_{d-1}^2 + y_d^2)^{1/2}\sin(\varphi_{d-1})$ gilt. Insbesondere ist $\|y\| = r$. Mit Hilfe dieser Angaben kann man eine Umkehrabbildung $V \to U$ angeben und leicht die stetige Differenzierbarkeit von Φ und Φ^{-1} prüfen. Die zweite Aussage folgt dann aus dem Transformationssatz 3.6.11, wobei wir beachten, dass die Menge V durch den $\mathbb{R}^d$ ersetzt wurde. Doch dies ist statthaft, da H_d eine Lebesgue-Nullmenge ist. □

Eine sehr hilfreiche Folgerung kann über rotationssymmetrische Funktionen gezogen werden:

Korollar 3.6.14 (Integrale rotationssymmetrischer Funktionen). *Sei $g\colon [0,\infty) \to \mathbb{R}$ messbar, dann ist die Abbildung $x \mapsto g(\|x\|)$ genau dann über den $\mathbb{R}^d$ Lebesgue-integrierbar, wenn die Abbildung $r \mapsto r^{d-1}g(r)$ über $[0,\infty)$ Lebesgue-integrierbar ist, und dann gilt*

$$\int_{\mathbb{R}^d} g(\|x\|)\,\lambda_d(\mathrm{d}x) = d\tau_d\int_{[0,\infty)} r^{d-1}g(r)\,\lambda_1(\mathrm{d}r), \tag{3.6.6}$$

wobei $\tau_d = \pi^{d/2}/\Gamma(1+\frac{d}{2})$ das Volumen der d-dimensionalen Kugel mit Radius Eins ist, siehe Beispiel 1.6.11.

Der Beweis besteht aus einer Anwendung von Lemma 3.6.13 auf die Abbildung $f(x) = g(\|x\|)$ und einer Auswertung der Integrale:

$$\int_{(0,\pi)^{d-2}\times(0,2\pi)}\prod_{i=1}^{d-2}\sin^{d-1-i}(\varphi_i)\,\lambda\big(\mathrm{d}(\varphi_1,\dots,\varphi_{d-1})\big) = 2\pi\prod_{i=1}^{d-2}\Big(\int_0^\pi \sin^{d-1-i}(\varphi)\,\mathrm{d}\varphi\Big) = d\tau_d$$

(Übungsaufgabe). Es sei bemerkt, dass die Zahl $d\tau_d$ genau die Oberfläche der Einheitskugel ist, wie wir in Beispiel 4.2.13 berechnen werden.

3.7 Produkte messbarer Räume

In der Wahrscheinlichkeitstheorie möchte man auch gerne ganze *Tupel* oder gar unendliche *Folgen* von zufälligen Experimenten (sogenannte stochastische Prozesse) behandeln und in die allgemeine Maßtheorie einbetten. Das natürlichste mathematische Konstrukt hierfür ist das *Produkt* der einzelnen Mengen, die die einzelnen Experimente beschreiben. In diesem Abschnitt führen wir die natürlichste σ-Algebra auf diesem Produkt ein, und zwar die kleinste, die es zulässt, dass die einzelnen Experimente in dem Produktraum genauso gut beschrieben werden können wie für sich alleine.

Es sei I eine beliebige Indexmenge, und für jedes $i \in I$ sei $(E_i, \mathcal{E}_i)$ ein messbarer Raum. Zur besseren Anschauung stelle man sich $I = \mathbb{N}$ vor, aber das Folgende ist nicht beschränkt auf abzählbare Indexmengen. Auf der Produktmenge $\prod_{i\in I} E_i$ wollen wir eine natürliche σ-Algebra definieren. Dazu betrachten wir die Mengen der Form

$$\mathcal{A}^{(j)}(A_j) = \Big\{(x_i)_{i\in I} \in \prod_{i\in I} E_i \colon x_j \in A_j\Big\} \qquad \text{mit } A_j \in \mathcal{E}_j, j \in I. \tag{3.7.1}$$

In Worten: $\mathcal{A}^{(j)}(A_j)$ ist die Menge derjenigen Elemente im Produktraum, deren j-te Komponente eine Bedingung erfüllt, nämlich dass sie in A_j liegt. Diese Mengen wählen wir als Erzeuger der gesuchten σ-Algebra:

Definition 3.7.1 (Produkt-σ-Algebra). *Die σ-Algebra auf dem Produkt $\prod_{i\in I} E_i$, die von den Mengen $\mathcal{A}^{(j)}(A_j)$ mit $A_j \in \mathcal{E}_j$ und $j \in I$ erzeugt wird, heißt die* Produkt-σ-Algebra *auf $\prod_{i\in I} E_i$ und wird mit $\bigotimes_{i\in I} \mathcal{E}_i$ bezeichnet. Den messbaren Raum $(\prod_{i\in I} E_i, \bigotimes_{i\in I} \mathcal{E}_i)$ nennt man den* Produktraum *der $(E_i, \mathcal{E}_i)$. Falls alle $(E_i, \mathcal{E}_i)$ gleich sind, also $(E_i, \mathcal{E}_i) = (E, \mathcal{E})$, so schreibt man auch $(E^I, \mathcal{E}^{\otimes I})$ für den Produktraum. Im Fall $I = \{1, \dots, n\}$ schreiben wir auch $\mathcal{E}^{\otimes n}$ statt $\mathcal{E}^{\otimes I}$.*

Bemerkung 3.7.2. Die Produkt-σ-Algebra ist die kleinste σ-Algebra auf dem Produkt, so dass die Projektionen

$$\pi_j \colon \prod_{i\in I} E_i \to E_j, \qquad \pi_j\big((x_i)_{i\in I}\big) = x_j,$$

messbar sind. Für die Mengen in (3.7.1) gilt $\mathcal{A}^{(j)}(A_j) = \pi_j^{-1}(A_j)$. Also haben wir auch auch $\bigotimes_{i\in I} \mathcal{E}_i = \bigvee_{i\in I} \pi_i^{-1}(\mathcal{E}_i) = \bigvee_{i\in I} \sigma(\pi_i)$; siehe Bemerkung 3.1.2. ◇

Bemerkung 3.7.3 (durchschnittstabile Erzeuger des Produktraums). Die Mengen der Form (3.7.1) bilden kein *durchschnittstabiles* Erzeugendensystem. Insbesondere ist der Eindeutigkeitssatz (Satz 2.3.8) nicht auf dieses System anwendbar. Daher benutzt man oft lieber das Erzeugendensystem, das aus den *endlichen Durchschnitten* von Mengen der Form (3.7.1) besteht. Genauer: Für eine endliche Teilmenge J von I und Mengen $A_j \in \mathcal{E}_j$ für $j \in J$ setzen wir

$$\mathcal{A}^{(J)}((A_j)_{j\in J}) = \bigcap_{j\in J} \mathcal{A}^{(j)}(A_j) = \{(x_i)_{i\in I} \colon x_j \in A_j \text{ für alle } j \in J\}. \tag{3.7.2}$$

Das Mengensystem $\mathcal{D}$ aller solcher $\mathcal{A}^{(J)}((A_j)_{j\in J})$ ist offensichtlich ein durchschnittstabiles Erzeugendensystem der Produkt-σ-Algebra auf $\prod_{i\in I} E_i$. Die Mengen $\mathcal{A}^{(J)}((A_j)_{j\in J})$ nennt man auch *endlichdimensionale Zylinder*. Auf $\mathcal{D}$ kann man Satz 2.3.8 problemlos anwenden. Im Spezialfall von Folgenräumen (also $I = \mathbb{N}_0$) und $J = \{0, 1, \dots, n\}$ hat $\mathcal{A}^{(J)}$ die Gestalt

$$\mathcal{A}^{(J)}((A_j)_{j\in J}) = A_0 \times A_1 \times \cdots \times A_n \times \prod_{i=n+1}^{\infty} E_i.$$

◇

Beispiel 3.7.4. Als eine Übungsaufgabe mache man sich klar, dass die Borel-σ-Algebra $\mathcal{B}_d$ auf dem $\mathbb{R}^d$ gleich dem d-fachen Produkt der Borel-σ-Algebra $\mathcal{B}$ auf $\mathbb{R}$ ist, also $\mathcal{B}_d = \mathcal{B}^{\otimes d}$ (siehe Beispiel 2.2.5). ◇

Beispiel 3.7.5 (Unendliche Folgenräume). Für die mathematische Modellierung eines unendlich oft wiederholten Versuchs benötigt man eine geeignete σ-Algebra auf dem unendlichen Folgenraum $E^{\mathbb{N}}$, wobei $(E, \mathcal{E})$ ein gegebener messbarer Raum ist, der einen einzelnen Versuch beschreibt. Eines der einfachsten und meist bemühten Beispiele ist der Münzwurf, wo man $E = \{K, Z\}$ setzt und die Potenzmenge $\mathcal{E} = \mathcal{P}(\{K, Z\})$ wählt. Die Produkt-σ-Algebra $(\mathcal{P}(\{K, Z\}))^{\otimes \mathbb{N}}$ besitzt das durchschnittstabile Erzeugendensystem, das aus den Zylindermengen $\{(x_n)_{n\in\mathbb{N}} \colon x_1 \in A_1, \dots, x_m \in A_m\}$ mit $m \in \mathbb{N}$ und $A_1, \dots, A_m \in \mathcal{P}(\{K, Z\})$ besteht. Diese Mengen legen Bedingungen fest, die die ersten m Komponenten der betrachteten Folge betreffen. Ein Maß auf dem Produktraum $(\{K, Z\}^{\mathbb{N}}, (\mathcal{P}(\{K, Z\}))^{\otimes \mathbb{N}})$ ist eindeutig festgelegt, wenn es auf allen diesen Zylindermengen festgelegt ist.

Man beachte, das z. B. die Mengen ‘Fast alle Würfe zeigen Kopf’ oder ‘Unendliche viele Würfe sind Kopf’ in der Produkt-σ-Algebra enthalten sind, denn es gilt z. B.

$$\begin{aligned}\big\{(x_n)_{n\in\mathbb{N}} \in \{K, Z\}^{\mathbb{N}} \colon x_n = K \text{ für fast alle } n \in \mathbb{N}\big\} &= \bigcup_{M\in\mathbb{N}} \bigcap_{N\in\mathbb{N}\colon N\geq M} \big\{(x_n)_{n\in\mathbb{N}} \colon x_N = K\big\} \\ &= \bigcup_{M\in\mathbb{N}} \bigcap_{N\in\mathbb{N}\colon N\geq M} \pi_N^{-1}(\{K\}),\end{aligned}$$

und die Menge auf der rechten Seite ist messbar als abzählbare Vereinigung eines abzählbaren Schnitts des Urbilds von $\{K\}$ unter der Projektion π_N.

Die Produkt-σ-Algebra $(\mathcal{P}(\{K, Z\}))^{\otimes \mathbb{N}}$ ist tatsächlich echt kleiner als die Menge $\mathcal{P}(\{K, Z\}^{\mathbb{N}})$ *aller* Teilmengen des Folgenraums. Dies ist nicht leicht einzusehen, man muss das Beispiel von Vitali (siehe Beispiel 2.2.14) geeignet modifizieren. Man kann zeigen, dass die Produkt-σ-Algebra von der Mächtigkeit des Kontinuums ist, d. h. dass es eine bijektive Abbildung $(\mathcal{P}(\{K, Z\}))^{\otimes \mathbb{N}} \to \mathbb{R}$ gibt, während die Potenzmenge $\mathcal{P}(\{K, Z\}^{\mathbb{N}})$ von der Mächtigkeit von $\mathcal{P}(\mathbb{R})$ ist. ◇

Beispiel 3.7.6 (Räume reeller Zahlenfolgen). Die Produkt-σ-Algebra $\mathcal{B}^{\otimes \mathbb{N}}$ auf der Menge aller Zahlenfolgen $\mathbb{R}^{\mathbb{N}}$ enthält unter Anderem auch die Menge aller beschränkten Folgen, die Menge aller Nullfolgen und die Menge aller konvergenten Folgen (Übungsaufgaben). Die Messbarkeit der Menge aller beschränkten Folgen sieht man so:

$$\big\{(x_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}} \colon \exists M \in \mathbb{N}\, \forall N \in \mathbb{N} \colon |x_N| \leq M\big\} = \bigcup_{M\in\mathbb{N}} \bigcap_{N\in\mathbb{N}} \pi_N^{-1}([-M, M]),$$

und die Menge auf der rechten Seite ist eine abzählbare Vereinigung eines abzählbares Schnittes von Urbildern messbarer Mengen unter Projektionen. ◇

3.8 Produktmaße

Auf der Produkt-σ-Algebra möchten wir nun auch Maße definieren, die in geeignetem Sinn als das Produkt von Maßen auf den einzelnen Faktoren aufgefasst werden können. Dies tun wir in diesem Abschnitt für endliche viele Faktoren. Es seien also zwei messbare Räume $(\Omega_1, \mathcal{F}_1)$ und $(\Omega_2, \mathcal{F}_2)$ gegeben. Wir erinnern uns, dass die Produkt-σ-Algebra $\mathcal{F}_1 \otimes \mathcal{F}_2$ auf $\Omega_1 \times \Omega_2$ durch das durchschnittstabile Mengensystem

$$\mathcal{C} = \{A_1 \times A_2 \colon A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2\} \tag{3.8.1}$$

erzeugt wird. Unsere Idee ist, ein Maß μ zu definieren, dass die natürliche Form $\mu(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2)$ auf $\mathcal{C}$ annimmt. Um aber $\mu(A)$ für eine beliebige messbare Menge A im Produktraum zu definieren, werden wir (ähnlich wie bei zweidimensionaler Riemann-Integration) die Menge A zunächst entlang aller 'waagerechten Schnitte' integrieren und die einzelnen Ergebnisse dann 'senkrecht aufintegrieren'. (Alternativ kann man natürlich auch die andere Reihenfolge wählen und zunächst entlang der 'senkrechten Schnitte' integrieren.) Um dies zu formalisieren, führen wir die *Schnitte* einer beliebigen Menge $A \subset \Omega_1 \times \Omega_2$ ein als die Mengen

$$\begin{aligned} A^{(2)}_{\omega_2} &= \{\omega_1 \in \Omega_1 \colon (\omega_1, \omega_2) \in A\}, \qquad && \omega_2 \in \Omega_2, \\ A^{(1)}_{\omega_1} &= \{\omega_2 \in \Omega_1 \colon (\omega_1, \omega_2) \in A\}, \qquad && \omega_1 \in \Omega_1. \end{aligned}$$

Allgemeiner nennen wir für eine Abbildung $f\colon \Omega_1 \times \Omega_2 \to \mathbb{R}$ die beiden Abbildungen $f^{(2)}_{\omega_2}(\cdot) = f(\cdot, \omega_2)$ und $f^{(1)}_{\omega_1}(\cdot) = f(\omega_1, \cdot)$ die *Schnitte* von f. Man sieht leicht, dass $1\!\!1_{A^{(2)}_{\omega_2}} = (1\!\!1_A)^{(2)}_{\omega_2}$ und eine analoge Formel für $A^{(1)}_{\omega_1}$ gelten. Unser Ziel ist also, zunächst das μ_1-Maß von $A^{(2)}_{\omega_2}$ zu bilden und das Ergebnis dann bezüglich $\mu_2(\mathrm{d}\omega_2)$ zu integrieren (oder auch mit vertauschten Rollen von '1' und '2'). Um den ersten Schritt machen zu können, müssen wir uns um Messbarkeit kümmern:

Satz 3.8.1. *Ist $f\colon \Omega_1 \times \Omega_2 \to \mathbb{R}$ eine messbare Funktion, so sind alle Schnitte messbar bezüglich der jeweiligen σ-Algebra, d. h. für jedes $\omega_2 \in \Omega_2$ ist $f^{(2)}_{\omega_2}$ $\mathcal{F}_1$-messbar, und für jedes $\omega_1 \in \Omega_1$ ist $f^{(1)}_{\omega_1}$ $\mathcal{F}_2$-messbar.*

Beweis. Wir zeigen nur die erste Aussage, und zwar zunächst nur für alle Funktionen der Form $f = 1\!\!1_A$ mit $A \in \mathcal{F}_1 \otimes \mathcal{F}_2$. (Der Rest des Beweises folgt dann wie gewohnt etwa mit Hilfe von Satz 3.1.12.) Betrachten wir das Mengensystem

$$\mathcal{G} = \{A \in \mathcal{F}_1 \otimes \mathcal{F}_2 \colon A^{(2)}_{\omega_2} \in \mathcal{F}_1 \text{ für alle } \omega_2 \in \Omega_2\}.$$

Wir möchten gerne zeigen, dass $\mathcal{G} = \mathcal{F}_1 \otimes \mathcal{F}_2$, denn das beendet den Beweis. Zunächst sehen wir, dass $\mathcal{C} \subset \mathcal{G}$ (siehe (3.8.1)), denn

$$(A_1 \times A_2)^{(2)}_{\omega_2} = \begin{cases} A_1, & \text{falls } \omega_2 \in A_2, \\ \emptyset & \text{sonst.} \end{cases} \tag{3.8.2}$$

Wegen $(A^{\mathrm{c}})^{(2)}_{\omega_2} = (A^{(2)}_{\omega_2})^{\mathrm{c}}$ für alle $A \subset \Omega_1 \times \Omega_2$ und $(\bigcup_{n\in\mathbb{N}} B_n)^{(2)}_{\omega_2} = \bigcup_{n\in\mathbb{N}} (B_n)^{(2)}_{\omega_2}$ für alle Folgen von Mengen $B_1, B_2, B_3, \dots \subset \Omega_1 \times \Omega_2$ ist $\mathcal{G}$ eine σ-Algebra. Also gilt $\mathcal{G} = \mathcal{F}_1 \otimes \mathcal{F}_2$. □

Nun können wir Produktmaße auf $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2)$ definieren. Dazu legen wir einen Kandidaten zunächst auf dem durchschnittstabilen Erzeugendensystem $\mathcal{C}$ aus (3.8.1) mit einer offensichtlichen Formel fest und erweitern ihn dann auf die Produkt-σ-Algebra:

Satz 3.8.2 (Produktmaße). *Seien $(\Omega_1, \mathcal{F}_1, \mu_1)$ und $(\Omega_2, \mathcal{F}_2, \mu_2)$ zwei σ-endliche Maßräume. Dann gibt es genau ein Maß μ auf dem Produktraum $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2)$, so dass gilt:*

$$\mu(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2), \qquad A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2.$$

Das Maß μ heißt das Produktmaß *von μ_1 und μ_2 und wird mit $\mu = \mu_1 \otimes \mu_2$ bezeichnet.*

Bemerkung 3.8.3 (Produkte unendlich vieler Maße). Das Produktmaß $\mu = \otimes_{i\in I}\mu_i$ von *un*endlich vielen Maßen μ_i mit $i \in I$ wird analog definiert (wenn es existiert) durch die Forderung

$$\mu\Big(\mathcal{A}^{(J)}((A_j)_{j\in J})\Big) = \prod_{i\in J}\mu_i(A_i), \qquad J \subset I \text{ endlich}, A_i \in \mathcal{E}_i \text{ für } i \in J;$$

siehe (3.7.2). Die Existenz von μ werden wir in der Vorlesung *Wahrscheinlichkeitstheorie* mit Hilfe des Satzes 2.3.7 von Carathéodory zumindest im abzählbaren Fall unter gewissen weiteren Einschränkungen beweisen; siehe den *Satz von Ionescu Tulcea.* Hier geben wir hier einen konstruktiven expliziten Beweis für endliche Produkte. ◇

Beweis von Satz 3.8.2. Die Eindeutigkeit ergibt sich sofort aus dem Eindeutigkeitssatz 2.3.8. Die Existenz zeigen wir nun konstruktiv mit Hilfe von Satz 3.8.1: Sei $A \in \mathcal{F}_1 \otimes \mathcal{F}_2$. Nach Satz 3.8.1 ist $A^{(2)}_{\omega_2} \in \mathcal{F}_1$ für alle $\omega_2 \in \mathcal{F}_2$, also dürfen wir $\mu_1(A^{(2)}_{\omega_2})$ bilden. Nun zeigen wir, dass die Abbildung $\omega_2 \mapsto \mu_1(A^{(2)}_{\omega_2})$ eine $\mathcal{F}_2$-messbare numerische Funktion ist. Da μ_1 σ-endlich ist, gibt es eine Folge $(B_n)_{n\in\mathbb{N}}$ von messbaren Teilmengen von Ω_1 mit $B_n \uparrow \Omega_1$ und $\mu_1(B_n) < \infty$ für jedes $n \in \mathbb{N}$. Wegen $\mu_1(A^{(2)}_{\omega_2}) = \lim_{n\to\infty}\mu_1(A^{(2)}_{\omega_2} \cap B_n)$ genügt es daher zu zeigen, dass für jedes $n \in \mathbb{N}$ die Abbildung $\omega_2 \mapsto \mu_1(A^{(2)}_{\omega_2} \cap B_n)$ eine $\mathcal{F}_2$-messbare Funktion ist. Man kann elementar zeigen, dass das Mengensystem

$$\{A \in \mathcal{F}_1 \otimes \mathcal{F}_2 \colon \omega_2 \mapsto \mu_1(A^{(2)}_{\omega_2} \cap B_n) \text{ ist } \mathcal{F}_2\text{-messbar}\}$$

ein Dynkin-System ist, das $\mathcal{C}$ enthält. Also ist es gleich $\mathcal{F}_1 \otimes \mathcal{F}_2$. Insgesamt haben wir also die Messbarkeit der Abbildung $\omega_2 \mapsto \mu_1(A^{(2)}_{\omega_2})$ gezeigt.

Nun können wir definieren:

$$\mu(A) = \int_{\Omega_2} \mu_1(A^{(2)}_{\omega_2})\,\mu_2(\mathrm{d}\omega_2), \qquad A \in \mathcal{F}_1 \otimes \mathcal{F}_2.$$

Eine Anwendung von Korollar 3.4.3 zeigt, dass μ ein Maß ist. Und mit Hilfe von (3.8.2) sieht man leicht, dass μ auf $\mathcal{C}$ die gewünschte Gestalt besitzt. □

Um problemlos bezüglich $\mu_1 \otimes \mu_2$ integrieren zu können, brauchen wir noch zwei Vertauschungssätze. Es seien $(\Omega_1, \mathcal{F}_1, \mu_1)$ und $(\Omega_2, \mathcal{F}_2, \mu_2)$ zwei σ-endliche Maßräume.

Satz 3.8.4 (Satz von Tonelli). *Es sei $f\colon \Omega_1\times\Omega_2 \to [0,\infty]$ eine messbare numerische Funktion. Dann sind die Abbildungen $\omega_1 \mapsto \int f^{(1)}_{\omega_1}\,\mathrm{d}\mu_2$ und $\omega_2 \mapsto \int f^{(2)}_{\omega_2}\,\mathrm{d}\mu_1$ messbar, und es gilt*

$$\int f\,\mathrm{d}(\mu_1 \otimes \mu_2) = \int_{\Omega_1}\Big(\int_{\Omega_2} f^{(1)}_{\omega_1}(\omega_2)\,\mu_2(\mathrm{d}\omega_2)\Big)\mu_1(\mathrm{d}\omega_1) = \int_{\Omega_2}\Big(\int_{\Omega_1} f^{(2)}_{\omega_2}(\omega_1)\,\mu_1(\mathrm{d}\omega_1)\Big)\mu_2(\mathrm{d}\omega_2).$$

Beweis. Diese Aussage ist nach Konstruktion richtig für alle Funktionen der Form $f = \mathbb{1}_A$ mit $A \in \mathcal{F}_1 \otimes \mathcal{F}_2$. Die allgemeine Aussage folgt wie gewohnt mit Hilfe von Satz 3.1.12. □

Mit Hilfe der üblichen Zerlegung in Positiv- und Negativteil erweitert man den Satz von Tonelli leicht auf alle integrierbare Funktionen $\Omega_1 \times \Omega_2 \to \mathbb{R}$ und erhält den berühmten Satz von Fubini, der vereinfacht sich so ausdrücken lässt: Falls $f \in \mathcal{L}^1(\mu_1 \otimes \mu_2)$, so kann man die Integrationsreihenfolge vertauschen:

$$\int f\,\mathrm{d}(\mu_1 \otimes \mu_2) = \int\Big(\int f(\omega_1,\omega_2)\,\mu_1(\mathrm{d}\omega_1)\Big)\mu_2(\mathrm{d}\omega_2) = \int\Big(\int f(\omega_1,\omega_2)\,\mu_2(\mathrm{d}\omega_2)\Big)\mu_1(\mathrm{d}\omega_1).$$

Satz 3.8.5 (Satz von Fubini). *Sei* $f \in \mathcal{L}^1(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2, \mu_1 \otimes \mu_2)$*. Dann liegt für* μ_1*-fast alle* $\omega_1 \in \Omega_1$ *die Funktion* $f^{(1)}_{\omega_1}$ *in* $\mathcal{L}^1(\Omega_2, \mathcal{F}_2, \mu_2)$*, und für* μ_2*-fast alle* $\omega_2 \in \Omega_2$ *liegt die Funktion* $f^{(2)}_{\omega_2}$ *in* $\mathcal{L}^1(\Omega_1, \mathcal{F}_1, \mu_1)$*. Setzt man*

$$f_1(\omega_1) = \begin{cases} \int_{\Omega_2} f^{(1)}_{\omega_1}(\omega_2)\, \mu_2(\mathrm{d}\omega_2), & \textit{falls } f^{(1)}_{\omega_1} \in \mathcal{L}^1, \\ 0 & \textit{sonst,} \end{cases}$$

und $f_2(\omega_2)$ *analog, so gilt*

$$\int f \,\mathrm{d}(\mu_1 \otimes \mu_2) = \int f_1(\omega_1)\, \mu_1(\mathrm{d}\omega_1) = \int f_2(\omega_2)\, \mu_2(\mathrm{d}\omega_2).$$

Bemerkung 3.8.6 (endlich viele Faktoren). Die obigen Betrachtungen können problemlos auf endlich viele Faktoren ausgedehnt werden, wobei man sich klar machen muss, dass die Produkte beliebig umgeklammert werden dürfen. Wenn also $(\Omega_i, \mathcal{F}_i, \mu_i)$ für jedes $i \in \{1, \dots, n\}$ ein σ-endlicher Maßraum ist, so definiert man rekursiv den Produktraum $\bigotimes_{i=1}^n (\Omega_i, \mathcal{F}_i, \mu_i) = (\prod_{i=1}^n \Omega_i, \bigotimes_{i=1}^n \mathcal{F}_i, \bigotimes_{i=1}^n \mu_i)$ und meint damit das Produkt $\prod_{i=1}^n \Omega_i = (\dots((\Omega_1 \times \Omega_2) \times \Omega_3) \times \dots \times \Omega_n)$ und analoge Formeln für die Produkte der σ-Algebren und die der Maße. Es ist dann klar, dass $\prod_{i=1}^n \Omega_i = (\Omega_1 \times \dots \times \Omega_k) \times (\Omega_{k+1} \times \dots \times \Omega_n)$ für jedes $k \in \{1, \dots, n\}$ gilt und analoge Formeln für die Produkte der σ-Algebren und die der Maße. Falls alle Faktoren $(\Omega_i, \mathcal{F}_i, \mu_i) = (\Omega, \mathcal{F}, \mu)$ übereinstimmen, schreiben wir natürlich Ω^n sowie $\mathcal{F}^{\otimes n}$ und $\mu^{\otimes n}$ oder auch Ω^I sowie $\mathcal{F}^{\otimes I}$ und $\mu^{\otimes I}$, falls $I = \{1, \dots, n\}$. ◇

Beispiel 3.8.7 (d-dimensionales Lebesgue-Maß). Im wichtigen Spezialfall, wo (siehe Beispiel 2.2.5) $(\Omega_i, \mathcal{F}_i, \mu_i) = (\mathbb{R}, \mathcal{B}, \lambda)$ für $i \in \{1, \dots, d\}$, identifizieren wir das d-fache Produkt mit $(\mathbb{R}^d, \mathcal{B}_d, \lambda^{\otimes d})$ (siehe auch Beispiel 3.7.4) und führen das Lebesgue-Maß auf dem $\mathbb{R}^d$ ein als das d-fache Produkt des eindimensionalen Lebesgue-Maßes. Es ist klar, dass $\lambda^{\otimes d}$ übereinstimmt mit den eindeutig bestimmten Maß λ_d auf $\mathcal{B}_d$, das jedem Quader $Q = \prod_{i=1}^d [a_i, b_i]$ seinen elementargeometrischen Inhalt $\lambda_d(Q) = \prod_{i=1}^d (b_i - a_i)$ zuweist. ◇

Beispiel 3.8.8 (Bernoulli-Maß). Für eine endliche Menge I heißt das Maß $(p\delta_1 + (1-p)\delta_0)^{\otimes I}$ auf $\{0,1\}^I$ mit der Potenzmenge (siehe Beispiel 2.2.10) das *Bernoulli-Maß* zum Parameter $p \in (0,1)$. Es beschreibt eine endliche Folge von unabhängigen Versuchen, die jeweils mit Wahrscheinlichkeit p Erfolg haben und sonst nicht.

Wenn zum Beispiel I die Menge aller Kanten zwischen benachbarten Punkten in dem Gitter $\mathbb{Z}^d \cap [-n, n]^d$ bezeichnet, dann kann $(p\delta_1 + (1-p)\delta_0)^{\otimes I}$ als ein stochastisches Modell für einen porösen Stein dienen, den man sich als in regelmäßige Kammern unterteilt vorstellt, die jeweils mit ihrer Nachbarkammer verbunden sind durch eine Wand, die mit Wahrscheinlichkeit p durchlässig ist und sonst nicht. Dies ist der Ausgangspunkt des Forschungszweiges der *Perkolation*, die untersucht, ob solche auf zufällige Weise poröse Steine eine Flüssigkeit durchsickern lassen oder nicht. Am interessantesten wird diese Frage natürlich, wenn wir statt eine endliche Box $\mathbb{Z}^d \cap [-n, n]^d$ den gesamten Raum $\mathbb{Z}^d$ betrachten, aber *un*endliche Produktmaße werden wir erst später im Zusammenhang mit dem Satz von Ionescu Tulcea behandeln. ◇

Kapitel 4

Der Gauß'sche Integralsatz

In diesem Kapitel behandeln wir den Gauß'schen Integralsatz, einen der wichtigsten Sätze der Integralrechnung im $\mathbb{R}^d$. Er erlaubt, das Volumenintegral der Divergenz eines Vektorfeldes durch ein Oberflächenintegral zu ersetzen. Damit ist er das d-dimensionale Analogon des Fundamentalsatzes der Integral- und Differenzialrechnung. Als eine Anwendung beweisen wir in Abschnitt 4.4 auch die klassische Variante des Stokes'schen Integralsatzes. Aus Zeit- bzw. Platzgründen ist die Darstellung dieser berühmten und anwendungsreichen Sätze in diesem Skript extrem kurz gehalten; ihre weit reichenden Auswirkungen werden erst im Zusammenhang mit der Theorie der exakten Differenzialformen durchsichtig, die allerdings nicht im Rahmen dieser Vorlesung enthalten ist.

Bevor wir den Gauß'schen Satz in Abschnitt 4.3 formulieren und beweisen, müssen wir in Abschnitt 4.2 erläutern, was ein Integral über gekrümmte Flächen im Raum sein soll. Diese gekrümmten Flächen werden zunächst in Abschnitt 4.1 eingeführt. Wir halten uns in den Abschnitten 4.1–4.3 an [Fo3, Par. 14 und 15] und in 4.4 an [H03, Abschnitt 6.4].

Im gesamten Kapitel seien $k, d \in \mathbb{N}$ mit $k \leq d$ fest vorgegeben.

4.1 Untermannigfaltigkeiten

In diesem Abschnitt erklären wir, was eine k-dimensionale Fläche im d-dimensionalen Raum sein soll, eine sogenannte Untermannigfaltigkeit. Kurz gesagt, ist dies eine Menge, die lokal als Nullstellenmenge eines Vektors von $d-k$ differenzierbaren Funktionen dargestellt werden kann, deren Funktionalmatrix maximalen Rang hat. Wir benutzen die Abkürzung $\partial_j = \frac{\partial}{\partial x_j}$ für die Ableitung nach der j-ten Komponente. Außerdem gehen wir dazu über, die Lebesgue-Integration $\lambda_d(\mathrm{d}x)$ als $\mathrm{d}^d x$ zu schreiben. Für offene Mengen U' und U'' schreiben wir kurz $\mathcal{C}^\alpha(U' \to U'')$ für die Menge der α Mal stetig differenzierbaren Abbildungen $U' \to U''$.

Definition 4.1.1. *Eine Teilmenge M des $\mathbb{R}^d$ heißt eine k-*dimensionale Untermannigfaltigkeit der Klasse $\mathcal{C}^\alpha$ *(mit $\alpha \in \mathbb{N}$), falls zu jedem Punkt $a \in M$ eine offene Umgebung U und ein $f \in \mathcal{C}^\alpha(U \to \mathbb{R}^{d-k})$ existiert, so dass $M \cap U = \{x \in U : f(x) = 0\}$ gilt und die Differenzialmatrix $Df = (\partial_j f_i)_{i=1,\dots,d-k; j=1,\dots,d}$ im Punkt a den maximalen Rang $d-k$ besitzt.*

Wir denken uns die Abbildung $f = (f_1, \dots, f_{d-k})^{\mathrm{T}}$ als eine Spalte angeordnet. Die Differenzialmatrix Df ist also eine $(d-k) \times d$-Matrix. Sie hat genau dann maximalen Rang, wenn die Gradienten $\nabla f_1, \dots, \nabla f_{d-k}$ linear unabhängig sind. Wenn wir die Klasse $\mathcal{C}^\alpha$ nicht gesondert erwähnen, wird α gleich Eins gesetzt.

Beispiel 4.1.2. Die $(d-1)$-dimensionalen Untermannigfaltigkeiten des $\mathbb{R}^d$ heißen auch *Hyperflächen*. Sie sind also lokal die Nullstellengebilde einer Funktion mit nicht-verschwindendem Gradienten. Ein Beispiel ist die $(d-1)$-*dimensionale Einheitssphäre*

$$S_{d-1} = \{x \in \mathbb{R}^d \colon \|x\| = 1\}.$$

Mit $f(x) = \|x\| - 1$ ist nämlich $S_{d-1} = \{x \in \mathbb{R}^d \colon f(x) = 0\}$, und der Gradient $\nabla f(x) = 2x$ ist überall in S_{d-1} ungleich Null. ◇

Mit Hilfe des Satzes über implizite Funktionen zeigen wir nun, dass eine k-dimensionale Untermannigfaltigkeit sich lokal als Graph einer Abbildung von k Variablen darstellen lässt.

Lemma 4.1.3 (Untermannigfaltigkeit als Graph). *Sei $M \subset \mathbb{R}^d$ eine k-dimensionale Untermannigfaltigkeit der Klasse $\mathcal{C}^\alpha$, und sei $a = (a_1, \dots, a_d) \in M$. Dann gibt es nach eventueller Umnummerierung der Koordinaten offene Umgebungen $U' \subset \mathbb{R}^k$ von $a' = (a_1, \dots, a_k)$ und $U'' \subset \mathbb{R}^{d-k}$ von $a'' = (a_{k+1}, \dots, a_d)$ sowie eine Funktion $g \in \mathcal{C}^\alpha(U' \to U'')$, so dass*

$$M \cap (U' \times U'') = \{(x', x'') \in U' \times U'' \colon g(x') = x''\} \qquad \text{mit } x' = (x_1, \dots, x_k), x'' = (x_{k+1}, \dots, x_d). \tag{4.1.1}$$

Beweis. Nach Definition gibt es eine offene Umgebung U von a und ein $f \in \mathcal{C}^\alpha(U \to \mathbb{R}^{d-k})$ mit $M \cap U = \{x \in U \colon f(x) = 0\}$, und $Df(a)$ hat maximalen Rang $d-k$. Also hat diese Matrix $d-k$ linear unabhängige Spalten, die nach eventueller Umnummerierung gerade die letzten $d-k$ Spalten sind. Nach eventueller Verkleinerung von U können wir also davon ausgehen, dass $(\partial_j f_i(x))_{i=1,\dots,d-k; j=k+1,\dots,d}$ für jedes $x \in U$ regulär ist, wobei $f = (f_1, \dots, f_{d-k})^{\mathrm{T}}$.

Nun wenden wir den Satz über implizite Funktionen an (siehe etwa [Fo2, Par. 8, Satz 2]) und erhalten offene Umgebungen U' von a' und U'' von a'' mit $U' \times U'' \subset U$ sowie ein $g \in \mathcal{C}^1(U' \to U'')$, so dass (4.1.1) gilt. Dass g wirklich in $\mathcal{C}^\alpha$ liegt, sieht man an der Darstellung $Dg = \frac{\partial g}{\partial x'} = -\frac{\partial f}{\partial x'}(\frac{\partial f}{\partial x''})^{-1}$, da die beiden Matrizen auf der rechten Seite jeweils $(\alpha - 1)$ Mal stetig differenzierbar sind. □

Beispiel 4.1.4. Wir schreiben die Einheitssphäre S_{d-1} aus Beispiel 4.1.2 lokal als einen Graphen in der Umgebung eines Punktes $a \in S_{d-1}$. Ohne Beschränkung der Allgemeinheit sei $a_d > 0$. Dann setzen wir $U' = \{x' \in \mathbb{R}^{d-1} \colon \|x'\| < 1\}$ und $g(x') = \sqrt{1 - \|x'\|}$. Dann ist $S_{d-1} \cap (U' \times (0, \infty)) = \{(x', x_n) \in U' \times (0, \infty) \colon g(x') = x_n\}$. Falls $a_d < 0$, gehen wir zu $-g$ über. ◇

Nun zeigen wir, dass sich k-dimensionale Untermannigfaltigkeiten lokal wie k-dimensionale Ebenen im $\mathbb{R}^d$ verhalten, die differenzierbar gekrümmt wurden. Für offene Mengen $U, V \subset \mathbb{R}^d$ und $\alpha \in \mathbb{N}$ nennen wir eine Funktion $F \colon U \to V$ einen *Diffeomorphismus der Klasse* $\mathcal{C}^\alpha$ oder kurz $\mathcal{C}^\alpha$-*invertierbar*, falls F bijektiv ist mit $F \in \mathcal{C}^\alpha(U \to V)$ und $F^{-1} \in \mathcal{C}^\alpha(V \to U)$. Wir schreiben E_k für die k-dimensionale Ebene $\{x \in \mathbb{R}^d \colon x_{k+1} = \dots = x_d = 0\}$.

Lemma 4.1.5 (Untermannigfaltigkeit als gekrümmte Ebene). *Eine Menge $M \subset \mathbb{R}^d$ ist genau dann eine k-dimensionale Untermannigfaltigkeit der Klasse $\mathcal{C}^\alpha$, wenn zu jedem $a \in M$ eine offene Umgebung $U \subset \mathbb{R}^d$ von a und eine $\mathcal{C}^\alpha$-invertierbare Abbildung $F \colon U \to V$ existiert (mit einem offenen $V \subset \mathbb{R}^d$), so dass $F(M \cap U) = E_k \cap V$ gilt.*

Beweis. '$\Longrightarrow$': Wir wenden Lemma 4.1.3 an und benutzen die Bezeichnungen von dort. Wir setzen $U = U' \times U''$ und definieren

$$F\colon U \to \mathbb{R}^d \qquad \text{durch} \qquad F(x', x'') = \big(x', x'' - g(x')\big)$$

Dann ist $V = F(U)$ offen und $F\colon U \to V$ ein Diffeomorphismus mit $F(M \cap U) = E_k \cap V$.

'$\Longleftarrow$': Wenn $F = (F_1, \dots, F_d)\colon U \to V$ eine $\mathcal{C}^\alpha$-invertierbare Abbildung mit $F(M \cap U) = E_k \cap V$ ist, so gilt auch $M \cap U = \{x \in U\colon F_{k+1}(x) = \dots = F_d(x) = 0\}$. Da die $d \times d$-Differenzialmatrix DF regulär ist, hat auch die Matrix ihrer letzten $d-k$ Spalten maximalen Rang $d-k$, und dies ist die Differenzialmatrix von $(F_{k+1}, \dots, F_d)$. Nach Definition ist M also eine k-dimensionale Untermannigfaltigkeit der Klasse $\mathcal{C}^\alpha$. □

Wir erinnern an den Begriff der Relativtopologie. Eine Teilmenge A einer gegebenen Menge $M \subset \mathbb{R}^d$ heißt *offen in M* oder *offen relativ zu M*, wenn es eine in $\mathbb{R}^d$ offene Menge U gibt, so dass $A = M \cap U$. Analog sind natürlich auch alle anderen topologischen Begriffe wir 'abgeschlossen', 'kompakt', 'Rand'und 'stetig' definiert. Man beachte, dass Untermannigfaltigkeiten M einer Dimension $< d$ im $\mathbb{R}^d$ nicht offen sind, so dass die auf M induzierte Relativtopologie nicht gleich der des $\mathbb{R}^d$ ist. Als Übungsaufgabe überlegt man sich, dass eine Menge $K \subset M$ genau dann kompakt relativ zu M ist, wenn K kompakt im $\mathbb{R}^d$ ist.

Definition 4.1.6 (Homöomorphismus, Immersion). *(i) Eine bijektive Abbildung f zwischen topologischen Räumen heißt ein* Homöomorphismus*, falls sowohl f als auch f^{-1} stetig sind.*

(ii) Sei $T \subset \mathbb{R}^k$ offen. Eine Abbildung $\varphi \in \mathcal{C}^1(T \to \mathbb{R}^d)$ heißt eine Immersion*, falls der Rang der Differenzialmatrix $D\varphi$ in ganz T den maximalen Rang k besitzt.*

Es stellt sich heraus, dass lokale Bilder von Immersionen Untermannigfaltigkeiten sind:

Lemma 4.1.7 (Untermannigfaltigkeit als Bild eines Homöomorphismus). *Sei $\Omega \subset \mathbb{R}^k$ offen und $\varphi \in \mathcal{C}^1(\Omega \to \mathbb{R}^d)$ eine Immersion der Klasse $\mathcal{C}^\alpha$. Dann gibt es zu jedem $c \in \Omega$ eine offene Umgebung $T \subset \Omega$, so dass $\varphi(T)$ eine k-dimensionale Untermannigfaltigkeit der Klasse $\mathcal{C}^\alpha$ und $\varphi\colon T \to \varphi(T)$ ein Homöomorphismus ist.*

Beweis. Da der Rang der Differenzialmatrix $D\varphi(c)$ gleich k ist, kann man nach eventueller Umnummerierung der Koordinaten im $\mathbb{R}^d$ annehmen, dass die ersten k Spalten von $D\varphi(c)$ eine reguläre $k \times k$-Matrix bilden. Dies ist die Differenzialmatrix von $(\varphi_1, \dots, \varphi_k)$, wenn $\varphi = (\varphi_1, \dots, \varphi_d)$. Nach dem Satz von der Umkehrabbildung (siehe etwa [Fo2, Par. 8, Satz 3]) gibt es daher eine offene Umgebung $T \subset \Omega$ von c und eine offene Menge $V \subset \mathbb{R}^k$, so dass $(\varphi_1, \dots, \varphi_k)\colon T \to V$ ein $\mathcal{C}^\alpha$-Diffeomorphismus ist. Wir definieren nun eine Abbildung $\Phi = (\Phi_1, \dots, \Phi_d)\colon T \times \mathbb{R}^{d-k} \to V \times \mathbb{R}^{d-k}$ durch $\Phi_i(t) = \varphi_i(t) + t_i \mathbb{1}_{j>k}$ für $t = (t_1, \dots, t_k)$. Dann ist Φ offensichtlich ein $\mathcal{C}^\alpha$-Diffeomorphismus mit $\Phi(T \times \{0\}) = \varphi(T)$. Also folgt aus Lemma 4.1.5, dass $\varphi(T)$ eine k-dimensionale Untermannigfaltigkeit der Klasse $\mathcal{C}^\alpha$ ist und T durch φ homöomorph auf $\varphi(T)$ abgebildet wird. □

Beispiel 4.1.8 (Kugeloberfläche). Sei $k = 2$, $d = 3$ und $\Omega = (0, \pi) \times \mathbb{R}$ und

$$\varphi\binom{t_1}{t_2} = \begin{pmatrix} \cos(t_2)\sin(t_1) \\ \sin(t_2)\sin(t_1) \\ \cos(t_1) \end{pmatrix} \qquad \text{für } \binom{t_1}{t_2} \in \Omega.$$

Dann ist

$$D\varphi\begin{pmatrix} t_1 \\ t_2 \end{pmatrix} = \frac{\partial\varphi}{\partial(t_1,t_2)}\begin{pmatrix} t_1 \\ t_2 \end{pmatrix} = \begin{pmatrix} \cos(t_2)\cos(t_1) & -\sin(t_{})\sin(t_1) \\ \sin(t_2)\cos(t_1) & \cos(t_2)\sin(t_1) \\ -\sin(t_1) & 0 \end{pmatrix}.$$

Man sieht leicht, dass diese Matrix für jedes $\binom{t_1}{t_2} \in \Omega$ den Rang Zwei besitzt, also ist φ eine Immersion. Man beachte, dass das Bild

$$M := \varphi(\Omega) = S_2 \setminus \{(0,0,1),(0,0,-1)\} \subset \mathbb{R}^3$$

die Einheitssphäre S_2 im $\mathbb{R}^3$ ohne den Nordpol $(0,0,1)$ und ohne den Südpol $(0,0,-1)$ ist. Allerdings wird Ω durch φ nicht homöomorph auf die Untermannigfaltigkeit M abgebildet, φ ist nicht einmal injektiv. Aber die offene Teilmenge $T = (0,\pi) \times (0,2\pi)$ von Ω wird von φ homöomorph auf das Bild $\varphi(T)$ abgebildet, das die offene Teilmenge der S_2 ist, die durch Wegnahme des Meridians des Längengrades Null aus M entsteht. Um M komplett mit solchen Abbildungen (sogenannten *Karten*, siehe unten) abzudecken, muss man also mindestens zwei benutzten. Allerdings wird durch eine einzige Karte schon 'fast' ganz M abgedeckt; nur eine eindimensionale Menge bleibt übrig. $\diamond$

Wir werden im Folgenden Homöomorphismen benutzen, um Untermannigfaltigkeiten lokal als Bild einer Teilmenge des $\mathbb{R}^k$ darzustellen:

Lemma 4.1.9 (Parameterdarstellungen durch Karten). *Eine Teilmenge M des $\mathbb{R}^d$ ist genau dann eine k-dimensionale Untermannigfaltigkeit der Klasse $\mathcal{C}^\alpha$, wenn zu jedem $a \in M$ eine in M offene Umgebung $V \subset M$ von a, eine offene Teilmenge T des $\mathbb{R}^k$ und eine Immersion $\varphi \in \mathcal{C}^\alpha(T \to V)$, die T homöomorph auf V abbildet, existieren.*

Der Homöomorphismus $\varphi\colon T \to V$ heißt eine lokale *Parameterdarstellung* oder *Karte* der Untermannigfaltigkeit M.

Beweis. '$\Longleftarrow$': Dies folgt aus Lemma 4.1.7.

'$\Longrightarrow$': Wir wenden Lemma 4.1.3 an und benutzen die Bezeichnungen von dort; d. h., M ist in einer Umgebung von a der Graph einer Funktion $g\colon U' \to U''$. Wenn man setzt: $V = M \cap (U' \times U'')$ und $T = U'$ und $\varphi\colon T \to \mathbb{R}^d$ durch $\varphi(t) = (t, g(t))$ definiert, so ist φ eine $\mathcal{C}^\alpha$-Immersion, die T homöomorph auf V abbildet. $\square$

Später wird wichtig werden, wie man verschiedene Karten miteinander vergleicht:

Lemma 4.1.10 (Parametertransformationen). *Sei $M \subset \mathbb{R}^d$ eine k-dimensionale Untermannigfaltigkeit der Klasse $\mathcal{C}^\alpha$, und seien $\varphi_i\colon T_i \to V_i$ für $i \in \{1,2\}$ zwei Karten der Klasse $\mathcal{C}^\alpha$, die T_i homöomorph auf $V_i \subset M$ abbilden. Setze $V = V_1 \cap V_2$. Dann ist $W_i = \varphi_i^{-1}(V)$ für $i \in \{1,2\}$ eine offene Teilmenge von T_i, und $\tau = \varphi_2^{-1} \circ \varphi_1\colon W_1 \to W_2$ ist ein $\mathcal{C}^\alpha$-Diffeomorphismus.*

Beweis. Als Urbild der offenen Menge V unter der stetigen Abbildung φ_i ist W_i offen. Nach Konstruktion ist τ bijektiv.

Sei nun $c_1 \in W_1$ beliebig, und wir setzen $a = \varphi_1(c_1)$ sowie $c_2 = \varphi_2^{-1}(a) = \tau(c_1)$. Nach Lemma 4.1.5 gibt es eine offene Umgebung $U \subset \mathbb{R}^d$ von a, eine offene Menge $U' \subset \mathbb{R}^d$ und einen $\mathcal{C}^\alpha$-Diffeomorphismus $F\colon U \to U'$ mit $F(M \cap U) = E_k \cap U'$, wobei wir daran erinnern, dass E_k die Ebene $\mathbb{R}^k \times \{0\}$ im $\mathbb{R}^d$ ist. Wir dürfen annehmen, dass $M \cap U \subset V$ gilt. Sei

$W_i' = \varphi_i^{-1}(M \cap U)$ für $i \in \{1,2\}$. Auf W_i hat $F \circ \varphi_i$ die Form $(g^{(i)}, 0, \dots, 0)$ mit einer geeigneten Funktion $g^{(i)} = (g_1^{(i)}, \dots, g_k^{(i)})$. Da DF invertierbar ist und $D\varphi_i$ den Rang k besitzt, ist auch die Differenzialmatrix von $g^{(i)}$ invertierbar, denn sie hat auch den Rang k, und sie ist eine $k \times k$-Matrix. Also ist $g^{(i)} \colon W_i' \to E_k \cap U'$ ein $\mathcal{C}^\alpha$-Diffeomorphismus, wobei wir $E_k \cap U'$ als eine offene Menge des $\mathbb{R}^d$ auffassen. Auf W_1 ist

$$\tau = \varphi_2^{-1} \circ \varphi_1 = (F \circ \varphi_2)^{-1} \circ (F \circ \varphi_1) = (g^{(2)})^{-1} \circ g^{(1)},$$

also liefert τ eine $\mathcal{C}^\alpha$-invertierbare Abbildung von W_1' auf W_2'. Da $c_1 \in W_1$ beliebig ist, folgt die Behauptung. □

In diesem Abschnitt haben wir also Untermannigfaltigkeiten auf drei Arten dargestellt: Als Nullstellenmenge, als Graph und als homöomorphes Bild einer Menge $\subset \mathbb{R}^k$. Alle drei Darstellungsarten werden im Folgenden benutzt werden.

4.2 Integration auf Untermannigfaltigkeiten

Wir wollen nun erklären, was das Integral $\int_M f(x)\,\mathbb{S}(\mathrm{d}x)$ einer Funktion $f\colon M \to \mathbb{R}$ auf einer Untermannigfaltigkeit M sein soll. Dies machen wir mit Hilfe von lokalen Karten, die wir geeignet zusammensetzen. Das Integral $\int_M f(x)\,\mathbb{S}(\mathrm{d}x)$ wird dann mit Hilfe des Transformationssatzes durch die Karte auf eine offene Teilmenge des $\mathbb{R}^k$ transportiert, und das geeignete Maß $\mathbb{S}$ auf M wird durch die Determinante der Funktionalmatrix der Karte gegeben; hierbei soll die Notation '$\mathbb{S}$' für das Flächenmaß an 'surface' erinnern. Kommen wir zu den Details.

Sei $M \subset \mathbb{R}^d$ eine k-dimensionale Untermannigfaltigkeit, und sei $\varphi\colon T \to V$ eine Karte, wobei $T \subset \mathbb{R}^k$ und $V \subset M$ zwei offene Mengen seien. Ganz ähnlich zu unserem Vorgehen in Abschnitt 1.4 führen wir nun die Matrix

$$G(t) = \big(D\varphi(t)\big)^{\mathrm{T}}\big(D\varphi(t)\big) = \sum_{\nu=1}^{k} \Big(\frac{\partial \varphi_\nu(t)}{\partial t_i}\,\frac{\partial \varphi_\nu(t)}{\partial t_j}\Big)_{i,j=1,\dots,k} = (g_{i,j}(t))_{i,j=1,\dots,k}$$

ein, wobei wir daran erinnern, dass $D\varphi = (\frac{\partial \varphi_\nu}{\partial t_j})_{\nu,j=1,\dots,k}$ die $k \times k$-Differenzialmatrix von φ ist. Die Matrix $G(t)$ heißt die *Gram'sche Matrix* von M bezüglich der Karte φ, und $g(t) = \det G(t)$ heißt die zugehörige *Gram'sche Determinante*. Bevor wir das Integral über M einführen, geben wir zunächst eine alternative Darstellung der Gram'schen Determinante und untersuchen, wie sie sich unter Parameterwechsel verhält.

Lemma 4.2.1 (Darstellung der Gram'schen Determinante). *Sei $\varphi = (\varphi_1, \dots, \varphi_d)\colon T \to V \subset M$ eine Karte der k-dimensionalen Untermannigfaltigkeit $M \subset \mathbb{R}^d$ und $g\colon T \to \mathbb{R}$ die zugehörige Gram'sche Determinante. Dann gilt*

$$g = \sum_{1 \le i_1 < i_2 < \dots < i_k \le d} \Big(\det \frac{\partial(\varphi_{i_1}, \dots, \varphi_{i_k})}{\partial(t_1, \dots, t_k)}\Big)^2.$$

Insbesondere ist die Gram'sche Determinante stets positiv.

Beweis. Zunächst zeigt man, dass für zwei Matrizen $A, B \in \mathbb{R}^{d\times k}$ gilt:

$$\det(A^{\mathrm{T}}B) = \sum_{1 \le i_1 < i_2 < \dots < i_k \le d} \det(A_{i_1,\dots,i_k}) \det(B_{i_1,\dots,i_k}), \tag{4.2.1}$$

wobei die $k\times k$-Matrix $A_{i_1,\dots,i_k}$ genau aus den Zeilen von A mit den Nummern $i_1,\dots,i_k$ besteht, analog für B. Diese Formel beweist man elementar (Übungsaufgabe), indem man zeigt, dass die rechte Seite für festes A für Matrizen B gilt, deren Spalten Einheitsvektoren sind, und dass die rechte Seite linear und alternierend in den Spalten von B ist. Wenn man die Formel (4.2.1) anwendet auf $A=B=D\varphi$ (Erinnerung: $G=(D\varphi)^{\mathrm T}D\varphi$), erhält man die Aussage. □

Lemma 4.2.2 (Gram-Determinante unter Parameterwechsel). *Seien $\varphi\colon T\to V\subset M$ und $\widetilde\varphi\colon\widetilde T\to\widetilde V\subset M$ zwei Karten der k-dimensionalen Untermannigfaltigkeit $M\subset\mathbb R^d$ und $g\colon T\to\mathbb R$ und $\widetilde g\colon\widetilde V\to\mathbb R$ die zugehörigen Gram'schen Determinanten. Es gelte $V\cap\widetilde V\neq\emptyset$. Ohne Beschränkung der Allgemeinheit sei $V=\widetilde V$. Wie in Lemma 4.1.10 sei $\tau\colon\widetilde T\to T$ eine $\mathcal C^\alpha$-invertierbare Abbildung mit $\widetilde\varphi=\varphi\circ\tau$. Dann gilt*

$$\widetilde g(y)=|\det(D\tau(y))|^2\,g(\tau(y)),\qquad y\in\widetilde T.$$

Beweis. Übungsaufgabe. □

Nun kommen wir endlich zur Definition des Integrals $\int_M f(x)\,\mathbb S(\mathrm dx)$.

Definition 4.2.3 (Integral über eine Untermannigfaltigkeit). *Seien $M\subset\mathbb R^d$ eine k-dimensionale Untermannigfaltigkeit und $f\colon M\to\mathbb R$ messbar.*

(i) Falls eine Karte $\varphi\colon T\to V\subset M$ existiert, so dass f nur auf V lebt (d. h. $f=0$ in $M\setminus V$), so heißt f integrierbar *über M, falls die Abbildung $t\mapsto f(\varphi(t))\sqrt{g(t)}$ bezüglich des k-dimensionalen Lebesgue-Maßes über T integrierbar ist. In diesem Fall setzen wir*

$$\int_M f(x)\,\mathbb S(\mathrm dx)=\int_T f(\varphi(t))\sqrt{g(t)}\,\mathrm d^kt,\tag{4.2.2}$$

wobei g die zu φ gehörige Gram'sche Determinante ist.

(ii) Falls endlich viele Karten $\varphi_i\colon T_i\to V_i$, $i=1,\dots,m$ existieren mit $\bigcup_{i=1}^m V_i=M$, so wählen wir eine der Überdeckung $(V_i)_{i=1,\dots,m}$ untergeordnete Teilung der Eins $(\alpha_i)_{i=1,\dots,m}$ (d. h., $\alpha_i\colon M\to[0,1]$ mit $\operatorname{supp}(\alpha_i)\subset V_i$ und $\sum_{i=1}^m\alpha_i=1$ auf M, so dass $\alpha_i\circ\varphi_i$ über jede kompakte Teilmenge von T_i integrierbar ist). Die Funktion f heißt dann integrierbar über *M, falls die Einschränkung von f auf V_i für jedes $i\in\{1,\dots,m\}$ im Sinne von (i) integrierbar ist, und man setzt*

$$\int_M f(x)\,\mathbb S(\mathrm dx)=\sum_{i=1}^m\int_M\alpha_i(x)f(x)\,\mathbb S(\mathrm dx).\tag{4.2.3}$$

Bemerkung 4.2.4. (i) Dass die rechte Seite von (4.2.2) nicht von der Karte φ abhängt, folgt leicht aus Lemma 4.2.2 (Übungsaufgabe), also ist Definition 4.2.3(i) sinnvoll.

(ii) Eine geeignete Teilung der Eins in Definition 4.2.3(ii) erhält man z. B. durch $\alpha_i=\mathbb 1_{W_i}$, wobei $W_i=V_i\setminus(V_1\cup\dots\cup V_{i-1})$.

(iii) Dass die Abbildung $t\mapsto\alpha_i(\varphi_i(t))f(\varphi_i(t))\sqrt{g(t)}$ unter den Voraussetzungen in Definition 4.2.3(ii) über T_i integrierbar ist (dies braucht man für die Wohldefiniertheit der rechten Seite von (4.2.3)), überlegt man sich als Übungsaufgabe.

(iv) Um zu sehen, dass Definition 4.2.3(ii) sinnvoll ist, muss man prüfen, dass die rechte Seite von (4.2.3) unabhängig von der Wahl der Überdeckung $(V_i)_{i=1,\dots,m}$ durch Karten und der Teilung der Eins $(\alpha_i)_{i=1,\dots,m}$ ist. Falls $\widetilde{\varphi}\colon \widetilde{T}_l \to \widetilde{V}_l$ für $l = 1,\dots,p$ Karten sind mit $\bigcup_{j=1}^p \widetilde{V}_j = M$ und $(\beta_j)_{j=1,\dots,p}$ eine ihr untergeordnete Teilung der Eins ist, so summiert sich der Term $\int_M \alpha_i(x)\beta_j(x)f(x)\,\mathbb{S}(\mathrm{d}x)$ über $j \in \{1,\dots,p\}$ auf zu $\int_M \alpha_i(x)f(x)\,\mathbb{S}(\mathrm{d}x)$ und über $i \in \{1,\dots,m\}$ zu $\int_M \beta_j(x)f(x)\,\mathbb{S}(\mathrm{d}x)$, da im ersten Fall alle Funktionen auf $M \setminus V_i$ verschwinden und im zweiten alle Funktionen auf $M \setminus \widetilde{V}_j$. Summiert man diese Summen über $i \in \{1,\dots,m\}$ bzw. über $j \in \{1,\dots,p\}$, ergibt sich

$$\sum_{i=1}^m \int_M \alpha_i(x)f(x)\,\mathbb{S}(\mathrm{d}x) = \sum_{j=1}^p \int_M \beta_j(x)f(x)\,\mathbb{S}(\mathrm{d}x),$$

und dies war zu zeigen.

(v) Wir haben in Definition 4.2.3 nicht den allgemeinsten Fall behandelt, sondern nur den, dass M von endlich vielen Karten überdeckt werden kann. Dies ist aber zum Beispiel immer schon dann der Fall, wenn M kompakt ist, denn nach Lemma 4.1.9 kann man M immer mit (eventuell überabzählbar vielen) Karten überdecken, und unter der Voraussetzung der Kompaktheit reichen endlich viele zum Überdecken aus; siehe auch die Bemerkung vor Definition 4.1.6, nach der M auch im $\mathbb{R}^d$ kompakt ist. Natürlich muss auch M nicht unbedingt kompakt sein, wenn nur der Träger der zu integrierenden Funktion f es ist.

◇

Man nennt $\mathbb{S}(\mathrm{d}x)$ das *k-dimensionale Flächenelement* oder das *k-dimensionale Volumen*, eine Bezeichnung, dem wir gleich in Beispiel 4.2.5 Substanz geben werden. Wenn wir die Definition 4.2.3 im Sinne von Bemerkung 4.2.4(v) erweitern würden, könnten wir $\mathbb{S}(\mathrm{d}x)$ als ein Maß auf M etablieren, aber wir führen das hier nicht durch.

Bemerkung 4.2.5 (k-dimensionales Volumen). Sei $M \subset \mathbb{R}^d$ eine k-dimensionale Untermannigfaltigkeit. Eine Teilmenge A von M heißt eine integrierbare Teilmenge, wenn $\mathbb{1}_A$ über M integrierbar ist. In diesem Fall nennt man

$$\mathrm{Vol}_k(A) = \int_M \mathbb{1}_A(x)\,\mathbb{S}(\mathrm{d}x)$$

das k-dimensionale Volumen oder den k-dimensionalen Flächeninhalt von A. Eine Funktion $f\colon M \to \mathbb{R}$ heißt über A integrierbar, wenn $f\mathbb{1}_A$ über M integrierbar ist, und man setzt

$$\int_A f(x)\,\mathbb{S}(\mathrm{d}x) = \int_M f(x)\mathbb{1}_A(x)\,\mathbb{S}(\mathrm{d}x).$$

Die Menge A heißt eine k-dimensionale Nullmenge, falls $\mathrm{Vol}_k(A) = 0$. ◇

Beispiel 4.2.6 (Kurvenlängen). Die Länge einer differenzierbaren Kurve ist ihr eindimensionales Volumen. Sei nämlich $I \subset \mathbb{R}$ ein offenes Intervall und $\varphi = (\varphi_1,\dots,\varphi_d) \in \mathcal{C}^1(I \to \mathbb{R}^d)$ mit $\varphi'(t) \neq 0$ für jedes $t \in I$, so dass φ das Intervall I homöomorph auf $\varphi(I)$ abbildet. Dann ist $\varphi(I)$ eine eindimensionale Untermannigfaltigkeit des $\mathbb{R}^d$ und φ eine globale Karte. Die Gram'sche

Determinante errechnet sich nach Lemma 4.2.1 als

$$g(t) = \sum_{i=1}^{d} \varphi_i'(t)^2 = \|\varphi'(t)\|^2,$$

also gilt für jedes kompakte Teilintervall J von I, dass $\mathrm{Vol}_1(\varphi(J)) = \int_J \|\varphi'(t)\| \,\mathrm{d}t$. $\diamond$

Beispiel 4.2.7 (Kugeloberfläche). Die 2-Sphäre S_2 im $\mathbb{R}^3$ hat die Karte

$$\Phi\colon (0,\pi)\times(0,2\pi)\to\mathbb{R}^3, \qquad \Phi\begin{pmatrix}\vartheta\\ \varphi\end{pmatrix} = \begin{pmatrix}\cos(\varphi)\sin(\vartheta)\\ \sin(\varphi)\sin(\vartheta)\\ \cos(\vartheta)\end{pmatrix},$$

siehe Beispiel 4.1.8. Die Gram'sche Determinante ist

$$g\begin{pmatrix}\vartheta\\ \varphi\end{pmatrix} = \det\begin{pmatrix}1 & 0\\ 0 & \sin^2(\vartheta)\end{pmatrix} = \sin^2(\vartheta).$$

Da der Nullmeridian eine zweidimensionale Nullmenge ist, gilt für jede integrierbare Funktion $f\colon S_2\to\mathbb{R}$:

$$\int_{S_2} f(x)\,\mathbb{S}(\mathrm{d}x) = \int_{(0,2\pi)}\int_{(0,\pi)} f\Big(\Phi\begin{pmatrix}\vartheta\\ \varphi\end{pmatrix}\Big)\sin(\vartheta)\,\mathrm{d}\vartheta\mathrm{d}\varphi.$$

Insbesondere ist die Oberfläche der Einheitskugel also gegeben durch

$$\mathrm{Vol}_2(S_2) = \int_0^{2\pi}\int_0^{\pi}\sin(\vartheta)\,\mathrm{d}\vartheta\mathrm{d}\varphi = \int_0^{2\pi} 2\,\mathrm{d}\varphi = 4\pi.$$

$\diamond$

Beispiel 4.2.8 (Rotationsflächen). Sei $I\subset\mathbb{R}$ ein offenes Intervall und $f\in\mathcal{C}^1(I\to(0,\infty))$. Die Menge

$$M = \{(x,y,z)\in\mathbb{R}^3\colon z\in I, x^2+y^2 = f(z)^2\}$$

ist eine zweidimensionale Untermannigfaltigkeit des $\mathbb{R}^3$, die bis auf eine Nullmenge durch die Karte

$$\Phi\colon I\times(0,2\pi)\to\mathbb{R}^3, \qquad \Phi\begin{pmatrix}t\\ \varphi\end{pmatrix} = \begin{pmatrix}f(t)\cos(\varphi)\\ f(t)\sin(\varphi)\\ t\end{pmatrix} = \begin{pmatrix}x\\ y\\ z\end{pmatrix}$$

dargestellt wird. Wir haben

$$D\Phi\begin{pmatrix}t\\ \varphi\end{pmatrix} = \begin{pmatrix}f'(t)\cos(\varphi) & -f(t)\sin(\varphi)\\ f'(t)\sin(\varphi) & f(t)\cos(\varphi)\\ 1 & 0\end{pmatrix},$$

also

$$g\begin{pmatrix}t\\ \varphi\end{pmatrix} = \det\begin{pmatrix}1+f'(t)^2 & 0\\ 0 & f(t)^2\end{pmatrix} = f(t)^2(1+f'(t)^2).$$

Daher ist das zweidimensionale Volumen der Rotationsfläche M gegeben durch

$$\mathrm{Vol}_2(M) = \int_{(0,2\pi)}\int_I f(t)\sqrt{1+f'(t)^2}\,\mathrm{d}t\mathrm{d}\varphi = 2\pi\int_I f(t)\sqrt{1+f'(t)^2}\,\mathrm{d}t,$$

falls das Integral existiert. $\diamond$

Beispiel 4.2.9. Sei $T \subset \mathbb{R}^{d-1}$ offen und $F \in \mathcal{C}^1(T \to \mathbb{R})$. Ihr Graph

$$M = \{x \in T \times \mathbb{R} \colon F(x_1, \dots, x_{d-1}) = x_d\}, \qquad \text{wobei } x = (x_1, \dots, x_d),$$

ist eine Hyperfläche im $\mathbb{R}^d$ mit Parameterdarstellung $\Phi \colon T \to M$, gegeben durch $\Phi(t) = (t, F(t))$. Die Funktionalmatrix ist $D\Phi(t) = \binom{E}{\nabla F(t)}$, wobei E die $(d-1)$-dimensionale Einheitsmatrix ist. Die Gram'sche Determinante berechnet sich mit Hilfe von Lemma 4.2.1 zu $g(t) = 1 + \|\nabla F(t)\|^2$. Also ist

$$\mathrm{Vol}_{d-1}(M) = \int_T \sqrt{1 + \|\nabla F(t)\|^2}\, \mathrm{d}^{d-1}t,$$

vorausgesetzt, das Integral existiert. ◇

Beispiel 4.2.10 (Obere Halbsphäre). Die obere Halbsphäre vom Radius $r \in (0, \infty)$,

$$M = \{x \in \mathbb{R}^d \colon \|x\| = r, x_d > 0\},$$

lässt sich darstellen als Graph der Funktion $F \colon T \to \mathbb{R}$, $F(t) = \sqrt{r^2 - \|t\|^2}$, wobei $T = \{t \in \mathbb{R}^{d-1} \colon \|t\| < r\}$ die offene Kugel im $\mathbb{R}^{d-1}$ mit Radius r ist. Daher ist $\nabla F(t) = -t/F(t)$, also nach Beispiel 4.2.9

$$g(t) = 1 + \frac{\|t\|^2}{F(t)^2} = \frac{r^2}{r^2 - \|t\|^2}.$$

Für jede integrierbare Funktion $f \colon M \to \mathbb{R}$ erhält man deshalb

$$\int_M f(x)\, \mathbb{S}(\mathrm{d}x) = \int_{\|t\|^2 < r^2} f\big(t, \sqrt{r^2 - \|t\|^2}\big) \frac{r}{\sqrt{r^2 - \|t\|^2}}\, \mathrm{d}^{d-1}t.$$

◇

Wie nicht anders zu erwarten ist, skaliert sich das Volumen einer k-dimensionalen Menge mit der k-ten Potenz des Skalierungsfaktors:

Lemma 4.2.11 (Skalierung). *Sei $r \in (0, \infty)$ und $M \subset \mathbb{R}^d$ eine k-dimensionale Untermannigfaltigkeit. Dann ist $rM = \{rm \colon m \in M\}$ ebenfalls eine k-dimensionale Untermannigfaltigkeit. Eine Funktion $f \colon rM \to \mathbb{R}$ ist genau dann über rM integrierbar, wenn die Abbildung $x \mapsto f(rx)$ über M integrierbar ist, und dann gilt*

$$\int_{rM} f(y)\, \mathbb{S}(\mathrm{d}y) = r^k \int_M f(rx)\, \mathbb{S}(\mathrm{d}x).$$

Insbesondere ist für jede integrierbare Teilmenge A von M auch die Menge rA integrierbar mit $\mathrm{Vol}_k(rA) = r^k \mathrm{Vol}_k(A)$.

Beweis. Übungsaufgabe. Man benutze eine Karte und zeige, dass die Gram'sche Determinante die geeignete Skalierungseigenschaft besitzt. □

Nun integrieren wir Funktionen über Kugeln, indem wir sie über die Sphären integrieren und alle diese Sphärenintegrale aufintegrieren. Dadurch erhalten wir eine neue Version des Lemmas 3.6.13 über d-dimensionale Kugelkoordinaten:

Lemma 4.2.12 (Integral über Kugelschalen). *: Sei $f\colon \mathbb{R}^d \to \mathbb{R}$ integrierbar. Dann ist für Lebesgue-fast jedes $r \in (0,\infty)$ die Funktion f über die Sphäre $\{x \in \mathbb{R}^d\colon \|x\| = r\}$ integrierbar, und es gilt*

$$\int_{\mathbb{R}^d} f(x)\,\mathrm{d}^d x = \int_{(0,\infty)} \Big(\int_{\|x\|=r} f(x)\,\mathbb{S}(\mathrm{d}x)\Big)\,\mathrm{d}r = \int_{(0,\infty)} \Big(\int_{S_{d-1}} f(rx)\,\mathbb{S}(\mathrm{d}x)\Big) r^{d-1}\,\mathrm{d}r,$$

wobei S_{d-1} die Einheitssphäre im $\mathbb{R}^d$ ist.

Beweis. Wir bringen einen Beweis ohne Zuhilfenahme von Lemma 3.6.13. Sei $H_\pm = \{x \in \mathbb{R}^d\colon \pm x_d > 0\}$ der obere bzw. untere Halbraum. Da sich f bis auf eine Nullmenge als Summe von $f1\!\!1_{H_+}$ und $f1\!\!1_{H_-}$ schreiben lässt, genügt es vorauszusetzen, dass $f1\!\!1_{H_-} = 0$ ist, wir brauchen also nur die obere Halbebene zu betrachten.

Sei $U = \{y \in \mathbb{R}^{d-1}\colon \|y\| < 1\}$ die offene Einheitskugel im $\mathbb{R}^{d-1}$, und sei $\Phi\colon U \times (0,\infty) \to H_+$ definiert durch $\Phi(y,r) = (ry, r\sqrt{1-\|y\|^2})$. Dann ist Φ eine $\mathcal{C}^1$-invertierbare Abbildung mit Differenzialmatrix

$$D\Phi\binom{y}{r} = \frac{r^{d-1}}{\sqrt{1-\|y\|^2}},$$

wie man als eine Übungsaufgabe zeigt. Nach dem Transformationssatz für Lebesgue-integrierbare Funktionen (Satz 3.6.11) gilt

$$\begin{aligned}\int_{H+} f(x)\,\mathbb{S}(\mathrm{d}x) &= \int_{U\times(0,\infty)} f\big(ry, r\sqrt{1-\|y\|^2}\big)\frac{r^{d-1}}{\sqrt{1-\|y\|^2}}\,\mathrm{d}^{d-1}y\mathrm{d}r\\ &= \int_{(0,\infty)} \Big(\int_U f\big(ry, r\sqrt{1-\|y\|^2}\big)\frac{r^{d-1}}{\sqrt{1-\|y\|^2}}\,\mathrm{d}^{d-1}y\Big)\,\mathrm{d}r,\end{aligned}$$

wobei wir im zweiten Schritt den Satz 3.8.5 von Fubini benutzten. Nach Beispiel 4.2.10 ist das Integral in der Klammer aber gerade gleich $\int_{\|x\|=r} f(x)\,\mathbb{S}(\mathrm{d}x)$, und dies beendet den Beweis. □

Nun berechnen wir endlich die Oberfläche der Einheitskugel, wie wir schon im Anschluss an Korollar 3.6.14 (Integrale rotationssymmetrischer Funktionen) ankündigten.

Beispiel 4.2.13 (Oberfläche der Einheitskugel). Sei $K_d = \{x \in \mathbb{R}^d\colon \|x\| \le 1\}$ die abgeschlossene Einheitskugel im $\mathbb{R}^d$ und $S_{d-1} = \partial K_d$ die $(d-1)$-dimensionale Einheitssphäre. In Beispiel 1.6.11 berechneten wir das Volumen von K_d zu

$$\tau_d = \mathrm{Vol}_d(K_d) = \frac{\pi^{d/2}}{\Gamma(\frac{d}{2}+1)}.$$

Wir wollen nun die Oberfläche von K_d berechnen, also das $(d-1)$-dimensionale Volumen $\omega_d = \mathrm{Vol}_{d-1}(S_{d-1})$ von S_{d-1}. Wendet man Lemma 4.2.12 auf $1\!\!1_{K_d}$ an, erhält man

$$\tau_d = \int_{\mathbb{R}^d} 1\!\!1_{K_d}(x)\,\mathrm{d}^d x = \int_0^1 \Big(\int_{S_{d-1}} \mathbb{S}(\mathrm{d}x)\Big) r^{d-1}\,\mathrm{d}r = \frac{\omega_d}{d}.$$

Also haben wir

$$\omega_d = d\tau_d = \frac{d\pi^{d/2}}{\Gamma(\frac{d}{2}+1)} = \frac{2\pi^{d/2}}{\Gamma(\frac{d}{2})}.$$

Insbesondere sind $\omega_2 = 2\pi$, $\omega_3 = 4\pi$ und $\omega_4 = 2\pi^2$. ◇

Natürlich kann man Lemma 4.2.12 auch benutzen, um einen zweiten Beweis von Korollar 3.6.14 zu erhalten, aber dies formulieren wir nicht aus.

4.3 Der Gauß'sche Integralsatz

In diesem Abschnitt behandeln wir den Gauß'schen Integralsatz, einen der wichtigsten Sätze der Integralrechnung im $\mathbb{R}^d$. Als ein d-dimensionales Analogon des Fundamentalsatzes der Integral- und Differenzialrechnung ersetzt er das Volumenintegral über die Divergenz eines Vektorfeldes durch ein Oberflächenintegral.

Zunächst müssen wir uns mit Tangential- und Normalenvektoren beschäftigen. Ein Tangentialvektor an einer Untermannigfaltigkeit ist, vereinfacht gesagt, ein Tangentialvektor an einer in der Untermannigfaltigkeit verlaufenden Kurve.

Definition 4.3.1 (Tangentialvektor). *Seien $M \subset \mathbb{R}^d$ eine Untermannigfaltigkeit und $a \in M$. Ein Vektor $v \in \mathbb{R}^d$ heißt ein* Tangentialvektor an M *in a, wenn es eine $\mathcal{C}^1$-Kurve $\psi\colon (-\varepsilon, \varepsilon) \to M$ (mit einem $\varepsilon > 0$) gibt mit $\psi(0) = a$ und $\psi'(0) = v$. Die Menge aller Tangentialvektoren an M in a heißt der* Tangentialraum *und wird mit T_aM bezeichnet.*

Es stellt sich heraus, dass der Tangentialraum gleich dem Spaltenraum der Differenzialmatrix einer Karte ist oder alternativ gleich dem orthogonalen Komplement der Spalten der Differenzialmatrix einer Abbildung f wie in Definition 4.1.1:

Lemma 4.3.2 (T_aM als Vektorraum). *Seien $M \subset \mathbb{R}^d$ eine k-dimensionale Untermannigfaltigkeit und $a \in M$. Dann gelten:*

(i) T_aM ist ein k-dimensionaler Untervektorraum des $\mathbb{R}^d$.

(ii) Seien $\varphi\colon \Omega \to V$ eine Karte von M (also $\Omega \subset \mathbb{R}^k$ offen und $V \subset M$ offen in M) und $c \in \Omega$ mit $\varphi(c) = a$. Dann ist eine Basis von T_aM gegeben durch die Vektoren $\partial_1\varphi(c), \dots, \partial_k\varphi(c)$.

(iii) Sei $U \subset \mathbb{R}^d$ eine offene Umgebung von a, und sei $f = (f_1, \dots, f_{d-k}) \in \mathcal{C}^1(U \to \mathbb{R}^{d-k})$ mit $M \cap U = \{x \in U\colon f(x) = 0\}$, so dass die $(d-k) \times d$-Matrix $Df(a)$ den maximalen Rang $d-k$ hat. Dann ist T_aM gleich dem Vektorraum

$$T = \{v \in \mathbb{R}^d\colon \langle v, \nabla f_i\rangle = 0 \text{ für alle } i = 1, \dots, d-k\}.$$

Beweis. Es sei T_1 der von $\partial_1\varphi(c), \dots, \partial_k\varphi(c)$ aufgespannte Vektorraum. Wir werden zeigen, dass $T_1 \subset T_aM \subset T$ gelten. Da T_1 und T jeweils k-dimensionale Vektorräume sind, sind dann alle drei Mengen gleich, und es folgen alle drei Aussagen des Lemmas.

(a) Beweis der Inklusion $T_1 \subset T_aM$: Sei $v \in T_1$, also $v = \sum_{i=1}^k \lambda_i\partial_i\varphi(c)$ für geeignetes $\lambda = (\lambda_1, \dots, \lambda_k) \in \mathbb{R}^k$. Mit einem genügend kleinen $\varepsilon > 0$ definieren wir eine Kurve $\psi\colon (-\varepsilon, \varepsilon) \to M$ durch $\psi(t) = \varphi(c + t\lambda)$. Dann gelten $\psi(0) = \varphi(c) = a$ und (nach der Kettenregel) $\psi'(0) = \sum_{i=1}^k \lambda_i\partial_i\varphi(c) = v$. Nach Definition ist also $v \in T_aM$.

(b) Beweis der Inklusion $T_aM \subset T$: Sei $v \in T_aM$, also $v = \psi'(0)$ für eine $\mathcal{C}^1$-Kurve $\psi\colon (-\varepsilon, \varepsilon) \to M$ mit $\psi(0) = a$. Da die Kurve in M verläuft, gilt $f_j(\psi(t)) = 0$ für alle $t \in (-\varepsilon, \varepsilon)$ und alle $j \in \{1, \dots, d-k\}$. Differenziation nach t ergibt

$$0 = \sum_{i=1}^d \partial_i f_j(\psi(0))\psi_i'(0) = \langle \nabla f_j(a), \psi'(0)\rangle = \langle v, \nabla f_j(a)\rangle,$$

also $v \in T$. □

Nun kommen wir zu Normalenvektoren.

Definition 4.3.3 (Normalenvektoren). *Seien $M \subset \mathbb{R}^d$ eine k-dimensionale Untermannigfaltigkeit und $a \in M$. Ein Vektor $v \in \mathbb{R}^d$ heißt ein* Normalenvektor von M *in a, wenn er auf T_aM senkrecht steht, also wenn für jedes $w \in T_aM$ gilt: $\langle v, w\rangle = 0$. Mit N_aM bezeichnen wir die Menge der Normalenvektoren von M in a.*

Insbesondere ist N_aM ein $(d-k)$-dimensionaler Untervektorraum des $\mathbb{R}^d$, und nach Lemma 4.3.2(iii) bilden $\nabla f_1(a), \dots, \nabla f_{d-k}(a)$ eine Basis von N_aM.

Im Gauß'schen Integralsatz werden wir nur noch $(d-1)$-dimensionale Untermannigfaltigkeiten betrachten, und zwar die Ränder von geeigneten kompakten Mengen:

Definition 4.3.4 (Kompaktum mit glattem Rand). *Wir sagen, ein Kompaktum $A \subset \mathbb{R}^d$ habe einen* glatten Rand, *wenn zu jedem Randpunkt $a \in \partial A$ eine offene Umgebung $U \subset \mathbb{R}^d$ und eine Abbildung $\psi \in \mathcal{C}^1(U \to \mathbb{R})$ existieren mit $A \cap U = \{x \in U : \psi(x) \leq 0\}$ und $\nabla\psi(x) \neq 0$ für jedes $x \in U$.*

Man beachte, dass Definition 4.3.4 insbesondere fordert, dass A in der Nähe seines Randes ein wenig 'Fleisch' hat. Zum Beispiel ist eine differenzierbare Kurve im $\mathbb{R}^2$ kein Kompaktum mit glattem Rand.

Lemma 4.3.5. *In der Situation von Definition 4.3.4 gilt $\partial A \cap U = \{x \in U : \psi(x) = 0\}$. Insbesondere ist der Rand von A eine kompakte $(d-1)$-dimensionale Untermannigfaltigkeit des $\mathbb{R}^d$.*

Beweis. (a) Die Inklusion $\partial A \cap U \subset \{x \in U : \psi(x) = 0\}$ sieht man wie folgt. Falls für ein $x \in U$ gilt: $\psi(x) < 0$, so gibt es wegen Stetigkeit von ψ eine Umgebung $V \subset U$ von x mit $\psi(y) < 0$ für $y \in V$. Insbesondere ist $V \subset A$, also ist x kein Randpunkt von A.

(b) Die umgekehrte Inklusion beweist man wie folgt. Sei $a \in U$ mit $\psi(a) = 0$. Wir müssen zeigen, dass a ein Randpunkt von A ist. Sei $v = \nabla\psi(a) \neq 0$. Eine Taylorentwicklung ergibt

$$\psi(a+\xi) = \psi(a) + \langle \nabla\psi(a), \xi\rangle + o(\xi) = \langle v, \xi\rangle + o(\xi), \qquad \xi \to 0.$$

Setzt man speziell $\xi = tv$, erhält man $\psi(a+tv) = t\|v\|^2 + o(t)$ für $t \to 0$. Es gibt daher ein $\varepsilon > 0$ mit $\psi(a+tv) > 0$ für $t \in (0, \varepsilon)$ und $\psi(a+tv) < 0$ für $t \in (-\varepsilon, 0)$. Also liegt für $t \in (0,\varepsilon)$ der Vektor $a + t\nabla\psi(a)$ in A^{c}, und für $t \in (-\varepsilon, 0)$ liegt er in A. Daher können wir a sowohl mit Punkten aus A^{c} approximieren als auch mit Punkten aus A, d. h., $a \in \partial A$. $\square$

Insbesondere ist $N_a\partial A$ eindimensional. Wir zeichnen einen Normalenvektor aus: den, der normiert ist und aus dem Kompaktum hinaus zeigt.

Lemma 4.3.6 (äußerer Einheits-Normalenvektor). *Sei $A \subset \mathbb{R}^d$ ein Kompaktum mit glattem Rand, und sei $a \in \partial A$. Dann gibt es genau einen Vektor $\nu(a) \in \mathbb{R}^d$ mit den folgenden Eigenschaften:*

(i) $\nu(a)$ steht senkrecht auf $T_a(\partial A)$,

(ii) $\|\nu(a)\| = 1$,

(iii) Es gibt ein $\varepsilon > 0$ mit $a + t\nu(a) \in A^{\mathrm{c}}$ für alle $t \in (0, \varepsilon)$.

Dieser Vektor $\nu(a)$ heißt äußerer Einheits-Normalenvektor *von A in a.*

Beweis. *Existenz:* Sei U eine offene Umgebung von a und $\psi \in \mathcal{C}^1(U \to \mathbb{R})$ mit $\nabla\psi \neq 0$ und $A \cap U = \{x \in U : \psi(x) \le 0\}$. Dann hat $\nu(a) = \nabla\psi(a)/\|\nabla\psi(a)\|$ die drei Eigenschaften (i) – (iii).

Eindeutigkeit: Der Normalenvektorraum $N_a\partial A$ ist eindimensional mit Basis $\{\nabla\psi(a)\}$, also gibt es wegen der Eigenschaft (i) ein $\lambda \in \mathbb{R} \setminus \{0\}$ mit $\nu(a) = \lambda\nabla\psi(a)$. Wegen der Normierung in (ii) ist $|\lambda| = 1/\|\nabla\psi(a)\|$, und wegen (iii) ist $\lambda > 0$. Also ist $\nu(a)$ durch (i) – (iii) eindeutig festgelegt. □

Beispiel 4.3.7 (Kugel). Die Kugel $A = \{x \in \mathbb{R}^d : \|x\| \le r\}$ mit Radius $r > 0$ hat einen glatten Rand, denn in Definition 4.3.4 kann man $\psi(x) = \|x\|^2 - r^2$ wählen. Der äußere Einheits-Normalenvektor in einem Randpunkt a ist $\nu(a) = a/r$. ◇

Die folgende Überlegung wird essenziell im Beweis des Gauß'schen Satzes sein.

Beispiel 4.3.8. Sei $A \subset \mathbb{R}^d$ ein Kompaktum mit glattem Rand und $a = (a_1, \dots, a_d) \in \partial A$ ein Randpunkt. In einer offenen Umgebung von a kann ∂A als Graph einer Funktion von $d-1$ Variablen dargestellt werden. Nach eventueller Umnummerierung gibt es also eine Umgebung U' von $a' = (a_1, \dots, a_{d-1})$, ein Intervall I mit $a_d \in I$ sowie ein $g \in \mathcal{C}^1(U' \to \mathbb{R})$, so dass $A \cap (U' \times I) = \{(x', x_d) \in U' \times I : x_d \le g(x')\}$ (bzw. '$\ge$' an Stelle von '$\le$'). Mit $\psi(x) = x_d - g(x')$ können wir (wie im Beweis von Lemma 4.3.6) den äußeren Einheits-Normalenvektor identifizieren als

$$\nu(a) = \frac{\nabla\psi(a)}{\|\nabla\psi(a)\|} = \frac{(-\nabla g(a'), 1)}{\sqrt{1 + \|\nabla g(a')\|^2}}, \qquad \text{wobei } \nabla g = (\partial_1 g, \dots, \partial_{d-1} g).$$

◇

Das folgende (recht technische) Lemma ist schon ein Spezialfall des Gauß'schen Integralsatzes, in dem A schon global in der angenehmen Form von Beispiel 4.3.8 gegeben ist.

Lemma 4.3.9. *Seien $U' \subset \mathbb{R}^{d-1}$ offen, $I = (\alpha, \beta)$ ein beschränktes offenes Intervall und $g \in \mathcal{C}^1(U' \to I)$. Wir setzen*

$$\begin{aligned} A &= \{(x', x_d) \in U' \times I : x_d \le g(x')\}, \\ M &= \{(x', x_d) \in U' \times I : x_d = g(x')\}. \end{aligned}$$

Dann gilt für jedes $f \in \mathcal{C}^1(U' \times I \to \mathbb{R})$ mit kompaktem Träger in $U' \times I$ und für jedes $i \in \{1, \dots, d\}$:

$$\int_A \partial_i f(x)\,\mathrm{d}^d x = \int_M f(x)\nu_i(x)\,\mathbb{S}(\mathrm{d}x), \tag{4.3.1}$$

wobei (mit der Notation $x = (x', x_d)$)

$$\nu_i(x) = \big(1 + \|\nabla g(x')\|^2\big)^{-1/2} \times \begin{cases} -\partial_i g(x'), & \textit{falls } i \le d-1, \\ 1, & \textit{falls } i = d. \end{cases}$$

Beweis. Nach Beispiel 4.2.9 besitzt das Flächenelement $\mathbb{S}$ von M bezüglich der Parameterdarstellung $x' \mapsto (x', g(x'))$ die Form

$$\mathbb{S}(\mathrm{d}x) = \sqrt{1 + \|\nabla g(x')\|^2}\,\mathrm{d}^{d-1}x'.$$

Im Beweis von (4.3.1) müssen die Fälle $i < d$ und $i = d$ unterscheiden.

1. Fall: $i < d$: Wir definieren $F\colon U' \times I \to \mathbb{R}$ durch $F(x', z) = \int_\alpha^z f(x', x_d)\,\mathrm{d}x_d$. Dann gelten

$$\partial_z F(x', z) = f(x', z) \qquad \text{und} \qquad \partial_{x_i} F(x', z) = \int_\alpha^z \partial_{x_i} f(x', x_d)\,\mathrm{d}x_d.$$

Daraus folgt mit der Kettenregel

$$\partial_{x_i} \int_\alpha^{g(x')} f(x', x_d)\,\mathrm{d}x_d = \partial_{x_i} F(x', g(x')) = \int_\alpha^{g(x')} \partial_{x_i} f(x', x_d)\,\mathrm{d}x_d + f(x', g(x'))\partial_{x_i} g(x').$$

Da die Funktion $x' \mapsto \int_\alpha^{g(x')} f(x', x_d)\,\mathrm{d}x_d$ einen kompakten Träger in U' besitzt, ergibt eine partielle Integration (siehe Lemma 1.3.3), dass

$$\int_{U'} \partial_{x_i} \Big(\int_\alpha^{g(x')} f(x', x_d)\,\mathrm{d}x_d \Big)\,\mathrm{d}^{d-1}x' = 0.$$

Damit erhalten wir mit Hilfe des Satzes 3.8.5 von Fubini

$$\begin{aligned}
\int_A \partial_i f(x)\,\mathrm{d}^d x &= \int_{U'} \Big(\int_\alpha^{g(x')} \partial_{x_i} f(x', x_d)\,\mathrm{d}x_d \Big)\,\mathrm{d}^{d-1}x' \\
&= \int_{U'} \partial_{x_i} \Big(\int_\alpha^{g(x')} f(x', x_d)\,\mathrm{d}x_d \Big)\,\mathrm{d}^{d-1}x' - \int_{U'} f(x', g(x'))\partial_{x_i} g(x')\,\mathrm{d}^{d-1}x' \\
&= \int_M f(x)\nu_i(x)\,\mathbb{S}(\mathrm{d}x).
\end{aligned}$$

2. Fall: $i = d$: Da für jedes $x' \in U'$ die Funktion $x_d \mapsto f(x', x_d)$ einen kompakten Träger in $I = (\alpha, \beta)$ hat, folgt aus dem Fundamentalsatz der Integral- und Differenzialrechnung

$$\int_\alpha^{g(x')} \partial_{x_d} f(x', x_d)\,\mathrm{d}x_d = f(x', g(x')).$$

Also folgt

$$\begin{aligned}
\int_A \partial_d f(x)\,\mathrm{d}^d x &= \int_{U'} \Big(\int_\alpha^{g(x')} \partial_{x_d} f(x', x_d)\,\mathrm{d}x_d \Big)\,\mathrm{d}^{d-1}x' = \int_{U'} f(x', g(x'))\,\mathrm{d}^{d-1}x' \\
&= \int_M f(x)\nu_d(x)\,\mathbb{S}(\mathrm{d}x),
\end{aligned}$$

und dies beendet den Beweis. □

Aus (4.3.1) erhalten wir schon die Aussage des Gauß'schen Integralsatzes, indem wir, wenn ein $\mathcal{C}^1$-Vektorfeld $F = (F_1, \dots, F_d)$ gegeben ist, (4.3.1) anwenden auf $f = F_i$ und über $i \in \{1, \dots, d\}$ aufsummieren. Dann entsteht auf der linken Seite das Integral über A von der Divergenz von F und auf der rechten Seite das Oberflächenintegral über ∂A des Skalarproduktes von F mit ν, siehe Beispiel 4.3.8. Den Beweis des Gauß'schen Integralsatzes werden wir mit Hilfe einer Teilung der Eins auf Lemma 4.3.9 zurückführen. Dabei werden wir das Folgende benötigen.

Lemma 4.3.10 (Lebesgue'sches Lemma). *Sei $A \subset \mathbb{R}^d$ kompakt und $(U_i)_{i \in I}$ eine offene Überdeckung von A. Dann gibt es ein $\lambda > 0$, eine* Lebesgue'sche Zahl*, mit der folgenden Eigenschaft: Jede Menge $K \subset \mathbb{R}^d$, die A trifft (d. h. $K \cap A \neq \emptyset$) und deren Durchmesser nicht größer als λ ist, ist ganz in einer der Mengen U_i enthalten.*

Beweis. Zu jedem $a \in A$ wählen wir ein $r_a > 0$ und ein $i \in I$ mit $B(a, r_a) \subset U_i$. Da auch die Familie $(B(a, r_a/2))_{a\in A}$ eine offene Überdeckung von A ist, gibt es endlich viele Punkte $a_1, \dots, a_m \in A$, so dass auch $B(a_1, r_{a_1}/2), \dots, B(a_m, r_{a_m}/2)$ noch A überdecken. Sei $\lambda = \frac{1}{2} \min\{r_1, \dots, r_m\}$. Sei nun $K \subset \mathbb{R}^d$ mit $K \cap A \neq \emptyset$ und $\operatorname{Diam}(K) \leq \lambda$. Wähle ein $a \in K \cap A$. Dann gibt es ein $k \in \{1, \dots, m\}$ und ein $i \in I$ mit $a \in B(a_k, r_{a_k}/2) \subset B(a_k, r_{a_k}) \subset U_i$. Wegen $\operatorname{Diam}(K) \leq \lambda \leq r_{a_k}/2$ folgt $K \subset B(a_k, r_{a_k}) \subset U_i$, also hat λ die geforderte Eigenschaft. □

Wir kommen nun endlich zum Gauß'schen Integralsatz. Wir erinnern daran, dass die *Divergenz* eines $\mathcal{C}^1$-Vektorfeldes $F \colon U \to \mathbb{R}^d$ (mit einem offenen $U \subset \mathbb{R}^d$) die Funktion $\operatorname{div} F = \sum_{i=1}^d \partial_i F_i \colon U \to \mathbb{R}$ ist. Die Abbildung, die jedem Randpunkt eines Kompaktums mit glattem Rand seinen äußeren Einheits-Normalenvektor zuordnet, nennen wir auch das *äußere Einheits-Normalenfeld.*

Satz 4.3.11 (Gauß'scher Integralsatz). *Sei $A \subset \mathbb{R}^d$ ein Kompaktum mit glattem Rand, $\nu \colon \partial A \to \mathbb{R}^d$ das äußere Einheits-Normalenfeld und $U \subset \mathbb{R}^d$ eine offene Menge, die A enthält. Dann gilt für jedes $\mathcal{C}^1$-Vektorfeld $F \colon U \to \mathbb{R}^d$:*

$$\int_A \operatorname{div} F(x)\,\mathrm{d}^d x = \int_{\partial A} \langle F(x), \nu(x)\rangle\,\mathbb{S}(\mathrm{d}x). \tag{4.3.2}$$

Beweis. Nach Beispiel 4.3.8 ist ∂A in der Umgebung jedes Punktes der Graph einer Funktion von $d-1$ Variablen. Daher gibt es eine offene Überdeckung $(U_i)_{i\in I}$ von A, so dass für jedes $i \in I$ entweder U_i in $A \setminus \partial A$ liegt oder (nach eventueller Umnummerierung der Koordinaten) U_i die Gestalt $U_i = U' \times (\alpha, \beta)$ hat und ein $g \in \mathcal{C}^1(U' \to (\alpha, \beta))$ existiert mit $U_i \cap A = \{(x', x_d) \in U' \times (\alpha, \beta) \colon x_d \leq g(x')\}$ (bzw. '$\geq$' an Stelle von '$\leq$').

Sei $\lambda > 0$ eine Lebesgue'sche Zahl wie im Lebesgue'schen Lemma 4.3.10. Wir setzen $\varepsilon = \lambda/2\sqrt{d}$ und betrachten die glatte Teilung der Eins $(\alpha_{p\varepsilon})_{p\in\mathbb{Z}^d}$, die wir am Anfang des Abschnittes 1.3.1 konstruierten. Der Träger jedes $\alpha_{p\varepsilon}$ ist ein Würfel der Seitenlänge 2ε, hat also einen Durchmesser $\leq \lambda$. Sei $P = \{p \in \mathbb{Z}^d \colon \operatorname{supp}(\alpha_{p\varepsilon}) \cap A \neq \emptyset\}$, dann ist P endlich, und es gilt $F = \sum_{p\in P} \alpha_{p\varepsilon} F$. Da offensichtlich beide Seiten der Behauptung (4.3.2) linear in F sind, reicht es also, (4.3.2) nur für die Funktionen $\alpha_{p\varepsilon} F$ zu beweisen.

Nach Konstruktion ist $\operatorname{supp}(\alpha_{p\varepsilon} F)$ in einer der Mengen U_i enthalten. Falls $U_i \subset A \setminus \partial A$, so sind beide Seiten von (4.3.2) Null für $\alpha_{p\varepsilon} F$ an Stelle von F, denn eine partielle Integration (siehe Lemma 1.3.3) zeigt, dass die linke Seite verschwindet (denn der Träger des Integranden ist eine kompakte Teilmenge des Innern von A), und der Integrand der rechten Seite ist Null auf dem Integrationsbereich ∂A. Im anderen Fall aber ist (4.3.2) (für $\alpha_{p\varepsilon} F$ an Stelle von F) eine Folgerung aus Lemma 4.3.9, angewendet auf die Komponenten von $\alpha_{p\varepsilon} F$. □

Beispiel 4.3.12. Die Divergenz des Feldes $F(x) = x$ ist $\operatorname{div} F(x) = \sum_{i=1}^d \partial_i x_i = d$. Für jedes Kompaktum $A \subset \mathbb{R}^d$ mit glattem Rand gilt daher $\int_A \operatorname{div} F(x)\,\mathrm{d}^d x = d\operatorname{Vol}_d(A)$, also folgt aus dem Gauß'schen Integralsatz

$$\operatorname{Vol}_d(A) = \frac{1}{d} \int_{\partial A} \langle x, \nu(x)\rangle\,\mathbb{S}(\mathrm{d}x).$$

Speziell für die Einheitskugel $A = K_d$ gilt nach Beispiel 4.3.7 $\nu(x) = x$ für alle Randpunkte x, also $\langle x, \nu(x)\rangle = \|x\|^2 = 1$. Daher folgt

$$\mathrm{Vol}_d(K_d) = \frac{1}{d}\int_{\partial K_d} \mathbb{S}(\mathrm{d}x) = \frac{1}{d}\mathrm{Vol}_{d-1}(S_{d-1}),$$

was wir schon aus Beispiel 4.2.13 wussten. ◇

Beispiel 4.3.13 (Interpretation). Das Skalarprodukt $\langle F(x), \nu(x)\rangle$ im Integral auf der rechten Seite von (4.3.2) ist gleich der Länge der Projektion des Vektors $F(x)$ auf den Normalenvektor $\nu(x)$. Man interpretiert daher das Maß $\langle F(x), \nu(x)\rangle\, \mathbb{S}(\mathrm{d}x)$ als den durch das Oberflächenelement $\mathbb{S}(\mathrm{d}x)$ aus dem Randpunkt x austretenden Fluss des Vektorfeldes $F(x)$. Das Integral $\int_{\partial A}\langle F(x), \nu(x)\rangle\, \mathbb{S}(\mathrm{d}x)$ ist daher der Gesamtfluss des Vektorfeldes F durch die Oberfläche von A.

Wenn z. B. F die Strömung einer inkompressiblen Flüssigkeit ist, so fordert man oft, dass $\mathrm{div}\ F(x) = 0$ ist für jedes $x \in \mathbb{R}^d$. Man nennt dann F *quellenfrei*. Dann ist also das Integral $\int_A \mathrm{div}\ F(x)\, \mathrm{d}^d x$ gleich Null für jedes Kompaktum A mit glattem Rand. Nach dem Gauß'schen Integralsatz heißt dies, dass der Gesamtfluss der Flüssigkeit durch die Oberfläche von A gleich Null ist. Dies erklärt den Begriff der Quellenfreiheit. Eine ähnliche Interpretation gilt für das elektrische Feld F in einem ladungsfreien Raum. ◇

Beispiel 4.3.14 (Archimedisches Prinzip). Ein Körper A befinde sich in einer Flüssigkeit der konstanten Dichte $c \in (0, \infty)$, deren Oberfläche mit der Ebene $\{x \in \mathbb{R}^3 \colon x_3 = 0\}$ zusammenfalle. Im Punkt $x \in \partial A$ übt die Flüssigkeit auf den Körper den Druck der Größe $cx_3\nu(x)$ aus, wobei $\nu(x)$ der äußere Einheits-Normalenvektor von A in x ist (und x_3 negativ: Der Druck ist nach innen gerichtet!). Die gesamte Auftriebskraft $K = (K_1, K_2, K_3)$ ist also gegeben durch $K_i = \int_{\partial A} cx_3\nu_i(x)\, \mathbb{S}(\mathrm{d}x)$. Mit dem Gauß'schen Satz kann man umformen zu $K_i = \int_A c\partial_i x_3\, \mathrm{d}^3 x$, also $K_1 = K_2 = 0$ und $K_3 = c\int_A \mathrm{d}^3 x = c\mathrm{Vol}_3(A)$. Der Körper erfährt also einen Auftrieb in x_3-Richtung, dessen Betrag gleich dem Gewicht der verdrängten Flüssigkeit ist. ◇

Für die folgende wichtige Anwendung erinnern wir daran, dass die Richtungsableitung einer Funktion $f \in \mathcal{C}^1(U \to \mathbb{R})$ (mit einem offenen $U \subset \mathbb{R}^d$) in Richtung eines Vektors $v \in \mathbb{R}^d$ gegeben ist als

$$\partial_v f(x) = \langle \nabla f(x), v\rangle = \sum_{i=1}^d \partial_i f(x) v_i.$$

Satz 4.3.15 (Green'sche Formel). *Sei $U \subset \mathbb{R}^d$ offen, $A \subset U$ ein Kompaktum mit glattem Rand und $\nu \colon \partial A \to \mathbb{R}^d$ das äußere Einheits-Normalenfeld. Dann gilt für je zwei Funktionen $f, g \in \mathcal{C}^2(U \to \mathbb{R})$:*

$$\int_A (f\Delta g - g\Delta f)\, \mathrm{d}^d x = \int_{\partial A} \big(f\,\partial_\nu g - g\,\partial_\nu f\big)\, \mathbb{S}(\mathrm{d}x). \qquad (4.3.3)$$

Beweis. Wir wenden den Gauß'schen Integralsatz an auf das Vektorfeld $F = f\nabla g - g\nabla f$. Da $\mathrm{div}\ (f\nabla g) = f\Delta g + \langle \nabla f, \nabla g\rangle$ gilt und eine analoge Formel für $\mathrm{div}\ (g\nabla f)$, folgt $\mathrm{div}\ F = f\Delta g - g\Delta f$. Auf ∂A gilt $\langle F, \nu\rangle = f\langle \nabla g, \nu\rangle - g\langle \nabla f, \nu\rangle = f\,\partial_\nu g - g\,\partial_\nu f$. Daher folgt (4.3.3) aus dem Gauß'schen Integralsatz. □

4.4 Der Stokes'sche Integralsatz

Als eine weitere Anwendung des Gauß'schen Integralsatzes beweisen wir nun die klassische Variante des Stokes'schen Integralsatzes, eine Aussage über das Integral der Rotation eines Vektorfeldes über ein zweidimensionales Kompaktum mit glattem Rand. Dies ist nur die Spitze des Eisberges der Gültigkeit dieses Satzes, und weit reichende Verallgemeinerungen, insbesondere in höheren Dimensionen, würden sichtbar, wenn man den Differenzialformenkalkül einführte, doch gehört dies nicht mehr zum Umfang dieser Vorlesung.

Wir erinnern daran, dass die *Rotation* eines $\mathcal{C}^1$-Vektorfeldes $a\colon \mathbb{R}^3 \to \mathbb{R}^3$ definiert ist als

$$\operatorname{rot} a = \begin{pmatrix} \partial_2 a_3 - \partial_3 a_2 \\ \partial_3 a_1 - \partial_1 a_3 \\ \partial_1 a_2 - \partial_2 a_1 \end{pmatrix}. \tag{4.4.1}$$

Ferner führen wir für eine $\mathcal{C}^2$-Abbildung $\varphi = (\varphi_1, \varphi_2, \varphi_3)\colon \mathbb{R}^2 \to \mathbb{R}^3$, $(u, v) \mapsto \varphi(u, v)$, das sogenannte *Dachprodukt* ein:

$$\varphi_u \wedge \varphi_v = \begin{pmatrix} \partial_u\varphi_2\,\partial_v\varphi_3 - \partial_u\varphi_3\,\partial_v\varphi_2 \\ \partial_u\varphi_3\,\partial_v\varphi_1 - \partial_u\varphi_1\,\partial_v\varphi_3 \\ \partial_u\varphi_1\,\partial_v\varphi_2 - \partial_u\varphi_2\,\partial_v\varphi_1 \end{pmatrix}. \tag{4.4.2}$$

Hier soll der Index 'u' an die Ableitung nach u erinnern, aber die Indices 1, 2 und 3 bezeichnen die Komponenten. Wir werden im Folgenden (wie üblich) auch $a(\varphi)$ an Stelle von $a\circ\varphi$ schreiben.

Satz 4.4.1 (Klassischer Stokes'scher Satz). *Es sei $\Omega \subset \mathbb{R}^3$ offen und $a \in \mathcal{C}^1(\Omega \to \mathbb{R}^3)$ ein $\mathcal{C}^1$-Vektorfeld. Ferner sei $A \subset \mathbb{R}^2$ ein Kompaktum mit glattem Rand und $\nu\colon \partial A \to S_1$ das äußere Einheits-Normalenfeld. Außerdem sei $\varphi = (\varphi_1, \varphi_2, \varphi_3) \in \mathcal{C}^2(A \to \Omega)$ eine Fläche der Klasse $\mathcal{C}^2$ in Ω. Dann gilt*

$$\int_A \big\langle (\operatorname{rot} a)\circ\varphi, \varphi_u \wedge \varphi_v \big\rangle \,\mathrm{d}u\mathrm{d}v = \int_{\partial A} \langle a\circ\varphi, \mathrm{d}\varphi\rangle, \tag{4.4.3}$$

wobei wir setzten:

$$\langle a\circ\varphi, \mathrm{d}\varphi\rangle = \langle a\circ\varphi, \varphi_u\rangle\,\mathrm{d}u + \langle a\circ\varphi, \varphi_v\rangle\,\mathrm{d}v, \qquad \textit{wobei } \varphi_u = (\partial_u\varphi_1, \partial_u\varphi_2, \partial_u\varphi_3)^{\mathrm{T}}. \tag{4.4.4}$$

Beweis. Wir wenden den Gauß'schen Integralsatz an auf das Vektorfeld

$$F(u, v) = \big(\langle a\circ\varphi, \varphi_v\rangle, -\langle a\circ\varphi, \varphi_u\rangle\big).$$

Dann errechnet sich nämlich die Divergenz von F zu

$$\begin{aligned} \operatorname{div} F &= \partial_u\langle a\circ\varphi, \varphi_v\rangle - \partial_v\langle a\circ\varphi, \varphi_u\rangle = \big\langle \partial_u\varphi, Da(\varphi)\partial_v\varphi\big\rangle - \big\langle \partial_v\varphi, Da(\varphi)\partial_u\varphi\big\rangle \\ &= \sum_{1\le j<k\le 3} \big(\partial_j a_k(\varphi) - \partial_k a_j(\varphi)\big)\big(\partial_u\varphi_j\,\partial_v\varphi_k - \partial_u\varphi_k\,\partial_v\varphi_j\big) \\ &= \big\langle (\operatorname{rot} a)\circ\varphi, \varphi_u\wedge\varphi_v\big\rangle, \end{aligned}$$

wobei der letzte Schritt aus einem Vergleich mit (4.4.1) und (4.4.2) folgt. Die linke Seite der Behauptung (4.4.3) ist also identisch mit der linken Seite (4.3.2) im Gauß'schen Satz.

Nun identifizieren wir die rechte Seite, und zwar nur unter der Annahme, dass ∂A durch eine einzige $\mathcal{C}^1$-Kurve γ parametrisiert wird. Sei also I ein Intervall und $\gamma = (\gamma_1, \gamma_2)\colon I \to \partial A$ eine $\mathcal{C}^1$-Kurve, die das Innere von A immer zu ihrer Linken hat. Der Vektor $(\gamma_2'(t), -\gamma_1'(t))$ steht senkrecht auf dem Tangentenvektor $\gamma'(t)$, und wegen der der Voraussetzung, dass γ entgegen dem Uhrzeigersinn um A herum verläuft, weist er aus A hinaus. Ohne Beschränkung der Allgemeinheit können wir voraussetzen, dass $(\gamma_2'(t), -\gamma_1'(t))$ die Länge Eins besitzt, denn sonst parametrisieren wir auf Bogenlänge. Also ist $\nu(x) = (\gamma_2'(t), -\gamma_1'(t))$ der äußere Einheits-Normalenvektor von A im Punkt $x = \gamma(t) = (u, v)$. Daher haben wir

$$\begin{aligned}\int_{\partial A} \langle F(x), \nu(x)\rangle \,\mathbb{S}(\mathrm{d}x) &= \int_I \Big[\big\langle a(\varphi(\gamma(t))), \varphi_v(\gamma(t))\big\rangle \gamma_2'(t) + \big\langle a(\varphi(\gamma(t))), \varphi_u(\gamma(t))\big\rangle \gamma_1'(t)\Big]\,\mathrm{d}t\\ &= \int_{\partial A} \langle a\circ\varphi, \varphi_u\rangle\,\mathrm{d}u + \langle a\circ\varphi, \varphi_v\rangle\,\mathrm{d}v = \int_{\partial A} \langle a\circ\varphi, \mathrm{d}\varphi\rangle,\end{aligned}$$

also ist die rechte Seite in (4.4.3) identisch mit der rechten Seite (4.3.2) im Gauß'schen Satz. Damit haben wir Satz 4.4.1 als einen Spezialfall des Gauß'schen Satzes erkannt. □

Bemerkung 4.4.2 (Interpretation). Nehmen wir an, dass $\varphi\colon A \to \varphi(A)$ regulär und injektiv ist, also eine Immersion. Dann ist φ eine Karte der zweidimensionalen Untermannigfaltigkeit $M = \varphi(A)$. Diese Fläche kann man orientieren, indem man ihr das Einheits-Normalenfeld $N\circ\varphi^{-1}$ zuordnet, wobei $N = \frac{\varphi_u\wedge\varphi_v}{\|\varphi_u\wedge\varphi_v\|}$. Man kann den Ausdruck $\int_{\partial A}\langle a(x), \mathrm{d}x\rangle$ z. B. durch die rechte Seite $\int_{\partial A}\langle a\circ\varphi, \mathrm{d}\varphi\rangle$ von (4.4.3) definieren; er hängt nicht von der Karte φ ab. Man nennt $\int_{\partial A}\langle a(x), \mathrm{d}x\rangle$ die *Zirkulation* des Vektorfeldes a längs der geschlossenen Kurve ∂M. Also kann man den Stokes'schen Satz 4.4.1 auch wie folgt formulieren: Der Fluss des Vektorfeldes rot a durch die orientierte Fläche M ist gleich seiner Zirkulation $\int_{\partial A}\langle a(x), \mathrm{d}x\rangle$ längs des Randes von M. Falls rot a überall in Ω gleich Null ist, ist die Zirkulation $\int_{\partial A}\langle a(x), \mathrm{d}x\rangle$ von a längs jeder geschlossenen Kurve in Ω gleich Null, die ein Flächenstück in Ω berandet. Dies ist der Grund dafür, dass man ein Vektorfeld a mit rot $a = 0$ auch *wirbelfrei* nennt. ◇

Bemerkung 4.4.3 (Vektorpotenzial). Ein Vektorfeld $b \in \mathcal{C}(\Omega \to \mathbb{R}^3)$ (mit einem offenen $\Omega \subset \mathbb{R}^3$) besitzt per Definition das *Vektorpotenzial* $a \in \mathcal{C}^1(\Omega \to \mathbb{R}^3)$, falls $b = \operatorname{rot} a$ gilt.

Diese Begriffsbildung ist analog zu der eines *Potenzials* $u \in \mathcal{C}^1(\Omega \to \mathbb{R})$ von b, das durch $b = \nabla u$ definiert wird. Auf Grund des Schwarz'schen Satzes ($\partial_i\partial_j u = \partial_j\partial_i u$ für alle i, j, falls $u \in \mathcal{C}^2$) muss notwendiger Weise rot $\nabla u = 0$ gelten. In Analysis-Grundvorlesungen sieht man, dass diese Bedingung in einfach zusammenhängenden Gebieten Ω schon hinreichend für die Existenz eines Potenzials u ist. Ferner ist es bis auf additive Konstante eindeutig bestimmt.

Ähnliches gilt für Vektorpotenziale. Falls b ein Vektorpotenzial $a \in \mathcal{C}^2(\Omega \to \mathbb{R}^3)$ besitzt, so ist div b = div rot $a = 0$ auf Grund des Schwarz'schen Satzes, d. h., b ist quellenfrei. Mit Hilfe eines Approximationsargumentes kann man die Folgerung div b = div rot $a = 0$ auch unter der schwächeren Voraussetzung $a \in \mathcal{C}^1(\Omega \to \mathbb{R}^3)$ ziehen. Ferner unterscheiden sich, falls Ω einfach zusammenhängend ist, zwei Vektorpotenziale additiv genau um das Gradientenfeld $\nabla\phi$ einer Funktion $\phi \in \mathcal{C}^2(\Omega \to \mathbb{R})$. Schließlich besitzt in einem einfach zusammenhängenden Gebiet Ω ein quellenfreies Vektorfeld $b = (b_1, b_2, b_3) \in \mathcal{C}^1(\Omega \to \mathbb{R}^3)$ das Vektorpotenzial

$$a(x) = \Big(\int_{z_0}^{x_3} b_2(x_1, x_2, t)\,\mathrm{d}t - \int_{y_0}^{x_2} b_3(x_1, t, z_0)\,\mathrm{d}t, -\int_{z_0}^{x_3} b_1(x_1, x_2, t)\,\mathrm{d}t, 0\Big),$$

wobei diese Formel allerdings nur für einen achsenparallelen Quader Ω gilt. ◇

Bibliographie

[AE01] H. Amann und J. Escher, *Analysis III*, Birkhäuser, Berlin 2001.

[Ba91] H. Bauer, *Maß- und Integrationstheorie,* 2. Auflage, Walter de Gruyter, Berlin–New York 1991.

[Co80] D.L. Cohn, *Measure Theory,* Birkhäuser, Basel 1980.

[D94] J.L. Doob, *Measure Theory*, Springer, Berlin 1994.

[E05] J. Elstrodt, *Maß- und Integrationstheorie*, 4. Auflage, Springer, Berlin 2005.

[Fo1] O. Forster, *Analysis 1*, 8. Auflage Vieweg, Braunschweig/Wiesbaden 2006.

[Fo2] O. Forster, *Analysis 2*, 5. Auflage Vieweg, Braunschweig/Wiesbaden 1999.

[Fo3] O. Forster, *Analysis 3*, Vieweg, Braunschweig/Wiesbaden 1984.

[G05] Jürgen Gärtner, Skript zur Vorlesung *Maß- und Integrationstheorie*, `ftp://ftp.math.tu-berlin.de/pub/Lehre/Masstheorie/SS05/skript.pdf`, TU Berlin 2005.

[H03] S. Hildebrandt, *Analysis 2*, Springer, Berlin 2003.

[Kl06] A. Klenke, *Wahrscheinlichkeitstheorie*, Springer, Berlin 2006.

[K02] K. Königsberger, *Analysis 2*, 4. Auflage, Springer, Berlin 2002.

Index