Vorlesungsnotizen

Höhere Analysis

Wintersemester 2017/18 und Wintersemester 2022/23

Thomas Schmidt

Version: 1. Februar 2023

Höhere Analysis: Inhaltsverzeichnis

Warnung: Diese Notizen umfassen nicht den vollständigen Vorlesungsstoff!
Falls Sie Fehler (jeglicher Art) finden oder sonstige Hinweise haben, bitte ich Sie, mir dies entweder persönlich oder unter thomas.schmidt.math@uni-hamburg.de mitzuteilen.

Vorwort zur Thematik

Die Vorlesung „Höhere Analysis“ beschäftigt sich fast ausschließlich mit der

Integrationstheorie für Funktionen mehrerer Variablen.

In der Vorlesung wird die Theorie in folgendem Umfang und folgender Reihenfolge entwickelt:

(1) Ein erstes kürzeres Thema ist die **Integration längs Kurven**, die eng mit der Stammfunktionsbildung bei Vektorfeldern und 1-Formen auf $\mathbb{R}^N$ verbunden ist. Dieses Kapitel bewegt sich damit in gewisser Weise zwischen der Integrationstheorie in einer und in mehreren Variablen.

(2) Ein Schwerpunktthema der Vorlesung ist der **maßtheoretische Zugang zur Integration**. Dieser liefert eine sehr weitgehende Integrationstheorie für Funktionen mehrerer Variablen und führt zum überlegenen Lebesgueschen Integralbegriff, zu allgemeinen Konzepten von Volumina und Flächeninhalten sowie zu wichtigen Integralsätzen.

(3) Der letzte große Themenkomplex der Vorlesung beschäftigt sich mit **Untermannigfaltigkeiten** des $\mathbb{R}^N$ (dies sind, grob gesprochen, reguläre nieder-dimensionalen Flächen), mit dem Kalkül der **Differentialformen** (eines speziellen Typs von Abbildungen) und schließlich mit der Integration von Differentialformen über Untermannigfaltigkeiten und weiteren Integralsätzen.

Kapitel 1

Kurvenintegrale und Stammfunktionen

1.1 Kurvenlänge, (Um-)Parametrisierung nach der Bogenlänge

In diesem ersten Abschnitt der höheren Analysis werden Kurven als Objekte eigenständigen Interesses untersucht. Zunächst wird die Länge einer Kurve und damit ihre wohl grundlegendste einzelne Kennzahl eingeführt:

Definition (**Totalvariation, Kurvenlänge**)**.** *Seien I ein Intervall und $\mathcal{X}$ ein normierter Raum über*[1] $\mathbb{K}$.

(I) *Die **(Total-) Variation** $\operatorname{Var}_I(f)$ einer Funktion $f\colon I \to \mathcal{X}$ über I wird erklärt als*

$$\operatorname{Var}_I(f) := \sup\left\{ \sum_{i=1}^{m} \|f(t_i)-f(t_{i-1})\|_{\mathcal{X}} \;:\; m \in \mathbb{N}_0\,,\; t_0<t_1<t_2<\ldots<t_m \text{ in } I \right\} \in [0,\infty]\,,$$

*und im Fall $\operatorname{Var}_I(f) < \infty$ nennt man f eine **Funktion von beschränkter**[2] **Variation** auf I oder kurz eine* BV*-Funktion auf I.*

(II) *Eine stetige Funktionen $c\colon I \to \mathcal{X}$ bezeichnet man als **Kurve**, **Weg** oder **Pfad** in $\mathcal{X}$ und modifiziert für solche Funktionen die gerade eingeführte Terminologie: Man nennt*

$$\mathrm{L}(c) := \operatorname{Var}_I(c) \in [0,\infty]$$

*die **Länge der Kurve** c und bezeichnet c im Fall $\mathrm{L}(c) < \infty$ als Kurve endlicher Länge oder **rektifizierbare Kurve**.*

Bemerkungen (zu Totalvariation und Kurvenlänge)**.**

(1) Bei der Totalvariation $\operatorname{Var}_I(f)$ und der Kurvenlänge $\mathrm{L}(c)$ handelt es sich per Definition um das Supremum über die Länge aller Streckenzüge, deren Ecken im Durchlaufsinn auf $\operatorname{Bild}(f)$

[1] Das Symbol $\mathbb{K}$ wird in diesen Notizen stets als Platzhalter für einen der beiden vollständigen Körper $\mathbb{R}$, $\mathbb{C}$ verwendet.

[2] Tatsächlich ist es treffender, von endlicher statt von beschränkter Variation zu sprechen. Nichtsdestotrotz sind die letztere Sprechweise und das Kürzel BV allgemein üblich, so dass es sich empfiehlt, dieser Konvention zu folgen.

beziehungsweise Bild(c) liegen. Man spricht dabei auch von ins Bild einbeschriebenen Streckenzügen und erkennt die Kurvenlänge $\mathrm{L}(c)$ im Licht dieser Interpretation als geometrisch sinnvolle Länge des (eventuell mehrfach durchlaufenen) Bildes von c. Die Variation $\mathrm{Var}_I(f)$ wurde allgemeiner für unstetige f erklärt und berücksichtigt neben der Länge von Bild(f) zusätzlich auch solche Längen, die in Unstetigkeitsstellen von f in gerader Linie „überflogen“ werden.

(2) Fügt man zu $t_0 < t_1 < t_2 < \ldots < t_{m-1} < t_m$ in I zusätzliche Zwischenstellen hinzu, geht also zu $\widetilde{t}_0 < \widetilde{t}_1 < \widetilde{t}_2 < \ldots < \widetilde{t}_{\widetilde{m}-1} < \widetilde{t}_{\widetilde{m}}$ in I mit $\widetilde{m} \geq m$ und $\{t_0, t_1, \ldots, t_m\} \subset \{\widetilde{t}_0, \widetilde{t}_1, \ldots, \widetilde{t}_{\widetilde{m}}\}$ über, so gilt gemäß der Dreiecksungleichung

$$\sum_{i=1}^{m} \|f(t_i)-f(t_{i-1})\|_{\mathcal{X}} \leq \sum_{i=1}^{\widetilde{m}} \|f(\widetilde{t}_i)-f(\widetilde{t}_{i-1})\|_{\mathcal{X}} \,,$$

die Länge des zugehörigen Streckenzugs wird also bei Verfeinerung der Zwischenstellen höchstens größer. Diese „Monotonie bei Verfeinerung“ ist beim Arbeiten mit der Definition des Öfteren nützlich und ist letztlich auch die Erklärung dafür, dass in dieser die Bildung des Supremums reicht und nicht etwa ein anderer Grenzprozess erforderlich ist.

(3) **Grundeigenschaften** der Totalvariation und der Kurvenlänge werden hier teils für beide Bildungen, teils nur für die allgemeinere Variation angegeben. Für Intervalle I, J, Stellen $a < t_* < b$ in I, normierte Räume $\mathcal{X}$, $\mathcal{Y}$, Funktionen $f, \widetilde{f} \colon I \to \mathcal{X}$ und $g \colon f(I) \to \mathrm{Y}$ sowie eine Kurve $c \colon I \to \mathcal{X}$ gelten

(a) der Konstanzsatz

$$\mathrm{Var}_I(f) = 0 \iff f \text{ ist konstant auf } I \,,$$

(b) die Halbnormeigenschaften (mit Konvention $0 \cdot \infty = 0$)

$$\mathrm{Var}_I(sf) = |s|\mathrm{Var}_I(f) \text{ für } s \in \mathbb{K} \qquad \text{und} \qquad \mathrm{Var}_I(f+\widetilde{f}) \leq \mathrm{Var}_I(f) + \mathrm{Var}_I(\widetilde{f}) \,,$$

(c) die **Zerlegungsadditivität**

$$\mathrm{Var}_I(f) = \mathrm{Var}_{I_{\leq t_*}}(f) + \mathrm{Var}_{I_{\geq t_*}}(f) \qquad \text{bzw.} \qquad \mathrm{L}\big(c|_{[a,b]}\big) = \mathrm{L}\big(c|_{[a,t_*]}\big) + \mathrm{L}\big(c|_{[t_*,b]}\big)$$

(wobei $\mathrm{L}\big(c|_{(a,b)}\big) = \mathrm{L}\big(c|_{[a,b]}\big)$ mit stetigem c durch Hinzunahme/Weglassen von a und b nicht beeinflusst wird, während sich die Variation unstetiger Funktionen ändern kann),

(d) die **Parametrisierungsinvarianz**

$$\mathrm{Var}_J(f \circ \tau) = \mathrm{Var}_I(f) \qquad \text{bzw.} \qquad \mathrm{L}(c \circ \tau) = \mathrm{L}(c)$$

für surjektive, monotone Parametertransformationen $\tau \colon J \to I$ (die aufgrund der Surjektivität, der Monotonie und der Intervall-Eigenschaft von I automatisch stetig sind),

(e) die untere Schranke

$$\mathrm{Var}_{[a,b]}(f) \geq \|f(b)-f(a)\|_{\mathcal{X}} \,,$$

die die gerade Verbindung als die kürzeste identifiziert,

(f) die Dehnungs-Abschätzungen und oberen Schranken

$$\mathrm{Var}_I(f) \le (\mathrm{Lip}\, f)\, \mathrm{Länge}(I) \qquad \text{und allgemeiner} \qquad \mathrm{Var}_I(g \circ f) \le (\mathrm{Lip}\, g)\, \mathrm{Var}_I(f)\,,$$

wobei $\mathrm{Lip}\, f$ und $\mathrm{Lip}\, g$ die optimalen Lipschitz-Konstanten von f und g bezeichnen und im Fall nicht Lipschitz-stetiger Funktionen als ∞ zu verstehen sind.

Ein wichtiger Spezialfall von (3f) ist die **Isometrieinvarianz**

$$\mathrm{Var}_I(h \circ f) = \mathrm{Var}_I(f) \qquad \text{bzw.} \qquad \mathrm{L}(h \circ c) = \mathrm{L}(c)$$

für Isometrien $h\colon f(I) \to \mathcal{Y}$ bzw. $h\colon c(I) \to \mathcal{Y}$ (d.h. für Abbildungen h mit $\|h(\widetilde{x}) - h(x)\|_{\mathcal{Y}} = \|\widetilde{x} - x\|_{\mathcal{X}}$ für alle $x, \widetilde{x} \in f(I)$ bzw. $x, \widetilde{x} \in c(I)$). Diese Eigenschaft ergibt sich durch Anwendung von (3f) auf die Abbildungen $g := h$ und $g := h^{-1}$ mit Lipschitz-Konstante 1 und beinhaltet als weitere Spezialfälle die **Translationsinvarianz** $\mathrm{Var}_I(x + f) = \mathrm{Var}_I(f)$ für $x \in \mathcal{X}$, die **Rotationsinvarianz** $\mathrm{Var}_I(Tf) = \mathrm{Var}_I(f)$ für lineare Isometrien[3] $T\colon \mathcal{X} \to \mathcal{X}$ und die diese zusammenfassende **Bewegungsinvarianz** $\mathrm{Var}_I(x + Tf) = \mathrm{Var}_I(f)$ für $x \in \mathcal{X}$ und lineare Isometrien $T\colon \mathcal{X} \to \mathcal{X}$.

(4) Im allgemeinen kann eine Kurve kompakte Teilintervalle beliebig kleiner positiver Länge auf unendlich lange Stücke im Bild abbilden. Ein typisches Beispiel $c\colon \mathbb{R} \to \mathbb{R}$ ist durch $c(0) := 0$ und $c(t) := t \cos \frac{1}{t}$ für $t \ne 0$ gegeben; es gelten dann $\sum_{i=m}^{n} \big|c\big(\frac{1}{i\pi}\big) - c\big(\frac{1}{(i+1)\pi}\big)\big| = \sum_{i=m}^{n} \big(\frac{1}{i\pi} + \frac{1}{(i+1)\pi}\big) \underset{n\to\infty}{\longrightarrow} \infty$ für alle $m \in \mathbb{N}$ und folglich $\mathrm{L}\big(c|_{[0,\delta]}\big) = \mathrm{L}\big(c|_{(0,\delta)}\big) = \infty$ für alle $\delta \in \mathbb{R}_{>0}$, die **Beispielkurve** c **hat** also **lokal bei 0 unendliche Länge**.

Ein gewisses Defizit des gerade gegebenen Beispiels besteht darin, dass die Beispielkurve c in $\mathbb{R}$ Teile ihres Bildes immer wieder durchläuft. In höheren Dimensionen ist dies für den demonstrierten Effekt aber keineswegs notwendig, und als anschaulicheres Beispiel kann man die zugehörige Graphenkurve $\gamma\colon \mathbb{R} \to \mathbb{R}^2$ mit $\gamma(t) = (t, c(t))$ für $t \in \mathbb{R}$ heranziehen; auch bei dieser gilt $\mathrm{L}\big(\gamma|_{[0,\delta]}\big) = \mathrm{L}\big(\gamma|_{(0,\delta)}\big) = \infty$ für alle $\delta \in \mathbb{R}_{>0}$.

Zur konkreten Berechnung der Kurvenlänge — jedenfalls im differenzierbaren Fall — ist folgendes Resultat sehr nützlich:

Satz (Darstellung der **Kurvenlänge als Riemann-Integral**)**.** *Seien $a < b$ in $\mathbb{R}$ und $\mathcal{X}$ ein normierter Raum. Dann ist jede C^1-Kurve[4] $c\colon [a,b] \to \mathcal{X}$ rektifizierbar mit*

$$\boxed{\mathrm{L}(c) = \int_a^b \|c'(t)\|_{\mathcal{X}}\, \mathrm{d}t\,.} \tag{$*$}$$

Bemerkungen (zur Integralformel für die Kurvenlänge)**.**

(1) Um die Formel ($*$) physikalisch zu interpretieren, kann man sich vorstellen, dass $c(t)$ die Position eines Teilchens zur Zeit t angibt und die Abbildung c dementsprechend die gesamte Bewegung dieses Teilchens über alle Zeiten $t \in [a,b]$ beschreibt. Die Ableitung $c'(t)$ entspricht

[3] Im Fall $\mathcal{X} = \mathbb{K}^N$ ist aus der linearen Algebra bekannt, dass die linearen Isometrien $\mathcal{X} \to \mathcal{X}$ gerade den orthogonalen beziehungsweise unitären Matrizen in $\mathbb{K}^{N\times N}$ entsprechen.

[4] Die C^1-Eigenschaft auf dem kompakten Intervall $[a,b]$ ist hier so zu verstehen, dass die rechtsseitige Ableitung $c'(a)$, die linksseitige Ableitung $c'(b)$ und auf (a,b) natürlich die übliche beidseitige Ableitung c' existieren und sich insgesamt eine stetige Ableitungsfunktion $c'\colon [a,b] \to \mathcal{X}$ ergibt.

dann dem Geschwindigkeitsvektor, $\|c'(t)\|_{\mathcal{X}}$ der skalaren Geschwindigkeit zur Zeit t und die Kurvenlänge der Länge der zwischen den Zeitpunkten a und b insgesamt zurückgelegten Wegstrecke. Bei dieser Interpretation besagt ($*$) nichts anderes, als dass man die Länge der Wegstrecke durch Integration der skalaren Geschwindigkeit berechnen kann.

(2) Der Satz bleibt allgemeiner für *Stückweise*-C^1-Kurven $c\colon [a,b] \to \mathcal{X}$ mit beschränkter Ableitung richtig, also für stetige Abbildungen $c \in \mathrm{C}^0([a,b],\mathcal{X})$ mit $c \in \mathrm{C}^1((a,b) \setminus S, \mathcal{X})$ und $\sup_{(a,b)\setminus S} \|c'\|_{\mathcal{X}} < \infty$ für eine endliche Menge S von „Knickstellen“. (Dabei kann man auf $S \cup \{a,b\}$ beliebige Werte von $\|c'\|$ festsetzen kann, um wie in der ursprünglichen Definition des Integrals einen auf ganz $[a,b]$ definierten Integranden zu erhalten. Tatsächlich sind diese Werte aber irrelevant und ihre Festlegung nicht wirklich nötig.)

Noch allgemeiner kann man Kurven $c \in \mathrm{C}^0(I,\mathcal{X})$, $c|_{I\setminus\{s_1,s_2,\ldots,s_n\}} \in \mathrm{C}^1(I\setminus\{s_1,s_2,\ldots,s_n\},\mathcal{X})$ auf beliebigen Definitionsintervallen I und mit (potentiell) unbeschränkter Ableitung c' erlauben. Dann folgt zwar i.A. nicht mehr die Rektifizierbarkeit von c auf I, aber die Formel ($*$) bleibt in der Form $\mathrm{L}(c) = \sum_{i=0}^{n} \int_{s_i}^{s_{i+1}} \|c'(t)\|_{\mathcal{X}}\,\mathrm{d}t$ mit möglicherweise uneigentlichen Riemann-Integralen auf der rechten Seite richtig, wobei s_0 als der linke und s_{n+1} als der rechte eventuell uneigentliche Randpunkt von I zu verstehen ist.

(3) Im **wichtigsten Fall $\boldsymbol{\mathcal{X} = \mathbb{R}^N}$** reduziert sich ($*$) zu

$$\boxed{\mathrm{L}(c) = \int_a^b |c'(t)|\,\mathrm{d}t = \int_a^b \sqrt{c_1'(t)^2 + c_2'(t)^2 + \ldots + c_N'(t)^2}\,\mathrm{d}t\,.}$$

Beispiel (zur Berechnung der Kurvenlänge mit der Integralformel)**.** Für die Parabel $P\colon \mathbb{R} \to \mathbb{R}^2$ mit $P(t) := (t, \frac{1}{2}t^2)$ und $P'(t) = (1,t)$ für $t \in \mathbb{R}$ erhält man

$$\mathrm{L}(P|_{(0,b)}) = \int_0^b |(1,t)|\,\mathrm{d}t = \int_0^b \sqrt{1+t^2}\,\mathrm{d}t = \frac{1}{2}b\sqrt{1+b^2} + \frac{1}{2}\,\mathrm{Arsinh}\,b\,,$$

wobei die Berechnung des Integrals (enteder direkt oder nach Substituation $t = \sinh\gamma$) mit Hilfe von partieller Integration und „Verschlucken“ der Ursprungsintegrals gelingt.

Um den vorausgehenden Satz zur Kurvenlänge zu beweisen, soll eine Charakterisierung des Riemann-Integrals benutzt werden, die in der Analysis II nicht im Detail besprochen wurde und hier nachgetragen wird. Um die Charakterisierung bequem formulieren zu können, wird erst ein zugehöriger Begriff geprägt (der übrigens auch im folgenden Abschnitt noch nützlich sein wird):

Definition (Limes der Zwischensummen bei gegen Null strebender Zerlegungsfeinheit)**.** *Seien $a < b$ in $\mathbb{R}$ und $f\colon [a,b] \to \mathbb{K}$ eine beschränkte Funktion. Dann heißt $\eta \in \mathbb{K}$ der **Limes der Zwischensummen** $\sum_{i=1}^{m} f(\tau_i)(t_i - t_{i-1})$ von f über $[a,b]$ **bei gegen Null strebender Zerlegungsfeinheit** $\max_{i\in\{1,2,\ldots,m\}}(t_i - t_{i-1})$, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ mit folgender Eigenschaft gibt: Für alle $m \in \mathbb{N}$, alle Zerlegungen $a = t_0 < t_1 < t_2 < \ldots < t_{m-1} < t_m = b$ von $[a,b]$ mit $\max_{i\in\{1,2,\ldots,m\}}(t_i - t_{i-1}) < \delta$ und alle Zwischenstellen $\tau_1 \in (t_0,t_1)$, $\tau_2 \in (t_1,t_2)$, $\ldots$, $\tau_m \in (t_{m-1},t_m)$ gilt*

$$\left|\eta - \sum_{i=1}^{m} f(\tau_i)(t_i - t_{i-1})\right| < \varepsilon\,.$$

Bemerkung (zum Limesbegriff bei Zwischensummen). Statt mit ε und δ kann der Limesbegriff der vorausgehenden Definition auch mit Folgen charakterisiert werden. Dass der Limes mit Wert η existiert bedeutet, dass für alle Folgen $(m_k)_{k\in\mathbb{N}}$ in $\mathbb{N}$, $(t_{0;k}, t_{1;k}, t_{2;k}, \ldots, t_{m_k;k})_{k\in\mathbb{N}}$, $(\tau_{1;k}, \tau_{2;k}, \ldots, \tau_{m_k;k})_{k\in\mathbb{N}}$ mit $a = t_{0;k} < t_{1;k} < \ldots < t_{m_k;k} = b$ und $\tau_{1;k} \in (t_{0;k}, t_{1;k})$, $\ldots$, $\tau_{m_k;k} \in (t_{m_k-1;k}, t_{m_k;k})$ aus der Konvergenz $\lim_{k\to\infty} \max_{i\in\{1,2,\ldots,m_k\}}(t_{i;k} - t_{i-1;k}) = 0$ der Zerlegungsfeinheiten die Konvergenz $\lim_{k\to\infty} \sum_{i=1}^{m_k} f(\tau_{i;k})(t_{i;k}-t_{i-1;k}) = \eta$ der zugehörigen Zwischensummen folgt.

Satz (**Charakterisierung des Riemann-Integrals mit beliebigen Zwischensummen**). *Seien $a < b$ in $\mathbb{R}$, $\eta \in \mathbb{K}$ und $f\colon [a,b] \to \mathbb{K}$ eine beschränkte Funktion. Genau dann ist f über $[a,b]$ eigentlich Riemann-integrierbar mit $\int_a^b f = \eta$, wenn η der Limes der Zwischensummen $\sum_{i=1}^m f(\tau_i)(t_i-t_{i-1})$ bei gegen Null strebender Zerlegungsfeinheit ist.*

Der Beweis der Charakterisierung ist völlig elementar, gestaltet sich in den Details aber etwas technisch. Der Vollständigkeit halber wird die Argumentation hier ausgeführt:

Beweis der Zwischensummen-Charakterisierung. Es reicht, den Fall $\mathbb{K} = \mathbb{R}$ zu behandeln, denn der Fall $\mathbb{K} = \mathbb{C}$ folgt daraus durch Aufspaltung in Real- und Imaginärteil.

Zum Beweis der Rück-Richtung sei η der Limes der Zwischensummen $\sum_{i=1}^m f(\tau_i)(t_i-t_{i-1})$ bei gegen Null strebender Zerlegungsfeinheit. Zu einem beliebigen $\varepsilon \in \mathbb{R}_{>0}$ kann dann *eine* Zerlegung $a = t_0 < t_1 < \ldots < t_m = b$ von $[a,b]$ fixiert werden, so dass für alle Zwischenstellen $\tau_1 \in (t_0,t_1)$, $\ldots$, $\tau_m \in (t_{m-1},t_m)$ stets $\left|\eta - \sum_{i=1}^m f(\tau_i)(t_i-t_{i-1})\right| < \varepsilon$ gilt (denn tatsächlich gilt dies ja sogar für *alle* Zerlegungen ab einer gewissen Feinheit δ und alle Zwischenstellen). Da $f(\tau_i)$ für $\tau_i \in (t_{i-1},t_i)$ beliebig nahe an $c_i := \sup_{(t_{i-1},t_i)} f$ kommen kann, folgt hieraus insbesondere $\sum_{i=1}^m c_i(t_i-t_{i-1}) \le \eta+\varepsilon$. Weil $\sum_{i=1}^m c_i(t_i-t_{i-1})$ die Obersumme zur Oberfunktion g zu f auf $[a,b]$ mit $g \equiv c_i$ auf (t_{i-1},t_i) (und beliebigen Werten $g(t_i)$) ist, ergibt sich auch für das Oberintegral ${}^*\!\!\int_a^b f$ von f die Abschätzung ${}^*\!\!\int_a^b f \le \eta+\varepsilon$. Aufgrund der Beliebigkeit von ε ist damit ${}^*\!\!\int_a^b f \le \eta$ gezeigt, und mit einer völlig analogen Argumentation sieht man für das Unterintegral ${}_*\!\!\int_a^b f \ge \eta$ ein. Per Definition bedeutet dies, dass f über $[a,b]$ Riemann-integrierbar ist und $\int_a^b f = \eta$ gilt.

Zum Beweis der Hin-Richtung existiere $\int_a^b f = \eta$, und $\varepsilon \in \mathbb{R}_{>0}$ sei wieder beliebig. Nach Definition des Riemann-Integrals existieren dann eine Zerlegung $a = x_0 < x_1 < \ldots < x_{\ell-1} < x_\ell = b$ von $[a,b]$ und eine Oberfunktion g zu f auf $[a,b]$ mit $g \equiv c_j$ auf (x_{j-1},x_j) und mit Obersumme $\mathrm{I}_a^b(g) = \sum_{j=1}^\ell c_j(x_j-x_{j-1}) < \eta+\varepsilon$. Sei nun $M := \sup_{[a,b]} |f| < \infty$, und sei $\delta^* \in \mathbb{R}_{>0}$ mit $\delta^*\ell M < \varepsilon$ gewählt. Im nächsten Schritt wird für alle Zerlegungen $a = t_0 < t_1 < \ldots < t_m = b$ von $[a,b]$ der Feinheit $\max_{i\in\{1,2,\ldots,m\}}(t_i-t_{i-1}) < \delta^*$ und alle Zwischenstellen $\tau_1 \in (t_0,t_1)$, $\ldots$, $\tau_m \in (t_{m-1},t_m)$ eine Abschätzung für die zugehörigen Zwischensummen nachgewiesen. Dabei hilft die Beobachtung, dass es in dieser Situation höchstens $(\ell-1)$ Intervalle des Typs (t_{i-1},t_i) gibt, die eine oder mehrere der zuvor fixierten Stellen $x_1, x_2, \ldots, x_{\ell-1}$ enthalten, während alle anderen Intervalle (t_{i-1},t_i) in einem Intervall des Typs (x_{j-1},x_j) enthalten sind, was die Abschätzung $f(\tau_i) \le c_j$ mit sich bringt. Bezeichnet h eine Hilfsfunktion mit $h \equiv f(\tau_i)$ auf (t_{i-1},t_i), so gilt also auf höchstens $(\ell-1)$ Intervallen der Länge höchstens δ^* nur die Abschätzung $h \le g+2M$ (denn $h \le \sup_{[a,b]} f = M$ und $g \ge f \ge -M$), während auf dem Rest von $[a,b]$ (eventuell mit Ausnahme der Stellen t_i selbst) die bessere Abschätzung $h \le g$ gilt. Dies impliziert

$$\sum_{i=1}^m f(\tau_i)(t_i-t_{i-1}) = \mathrm{I}_a^b(h) \le \mathrm{I}_a^b(g) + 2\delta^*(\ell-1)M < \eta+3\varepsilon\,.$$

Mit einer analogen Argumentation findet man außerdem ein $\delta_* \in \mathbb{R}_{>0}$, so dass für Zerlegungen $a = t_0 < t_1 < \ldots < t_m = b$ der Feinheit $< \delta_*$ und Zwischenstellen τ_i stets $\sum_{i=1}^m f(\tau_i)(t_i-t_{i-1}) > \eta-3\varepsilon$ gilt. Ist die Feinheit sogar $< \delta := \min\{\delta^*,\delta_*\}$, so gilt also $|\eta - \sum_{i=1}^m f(\tau_i)(t_i-t_{i-1})| < 3\varepsilon$. Aufgrund der Beliebigkeit von ε ist damit nachgewiesen, dass η der Limes von $\sum_{i=1}^m f(\tau_i)(t_i-t_{i-1})$ bei gegen Null strebender Zerlegungsfeinheit ist. □

Definition & Bemerkung (zum **Riemann-Integral Vektor-wertiger Funktionen**). *Der Limes der Zwischensummen $\sum_{i=1}^m f(\tau_i)(t_i-t_{i-1})$ bei gegen Null strebender Zerlegungsfeinheit kann analog für beschränkte Funktionen $f\colon [a,b] \to \mathcal{X}$* ***mit Werten in einem normierten Raum*** *$\mathcal{X}$ erklärt werden (wozu man in der Definition lediglich den Betrag durch die Norm zu ersetzen hat) und bietet eine kanonische Möglichkeit, das Riemann-Integral auch in diesem Fall zu definieren: Man nennt f Riemann-integrierbar, wenn der Limes in $\mathcal{X}$ existiert, und schreibt dann $\int_a^b f \in \mathcal{X}$ für seinen Wert. Im wichtigsten Fall $\mathcal{X} = \mathbb{K}^N$ ist diese Bildung äquivalent zur* *komponentenweisen Bildung des Integrals, d.h. für* $\mathbb{R}^N$*-wertiges* $f = (f_1, f_2, \dots, f_N)$ *gilt*

$$\int_a^b f = \left(\int_a^b f_1, \int_a^b f_2, \dots, \int_a^b f_N \right) \in \mathbb{K}^N$$

(*falls eine Seite der Gleichung definiert ist*).

Nach diesen Vorbereitungen erfolgt jetzt der Beweis des vorletzten Satzes:

Beweis des Satzes über die Darstellung der Kurvenlänge als Riemann-Integral. Zu einem fixierten $\varepsilon \in \mathbb{R}_{>0}$ wählt man zuerst ein $\delta \in \mathbb{R}_{>0}$, so dass einerseits für alle Zerlegungen $a = x_0 < x_1 < x_2 < \ldots < x_\ell = b$ von $[a,b]$ und alle Zwischenstellen $\xi_1 \in (x_0, x_1), \dots, \xi_\ell \in (x_{\ell-1}, x_\ell)$ die Implikation

$$\max_{j \in \{1,2,\dots,\ell\}} (x_j - x_{j-1}) < \delta \implies \left| \int_a^b \|c'(t)\|_{\mathcal{X}} \, \mathrm{d}t - \sum_{j=1}^{\ell} \|c'(\xi_j)\|_{\mathcal{X}} (x_j - x_{j-1}) \right| < \varepsilon$$

gilt und andererseits für alle $x, \widetilde{x} \in [a,b]$ die Implikation

$$|\widetilde{x} - x| < \delta \implies \|c'(\widetilde{x}) - c'(x)\|_{\mathcal{X}} < \varepsilon \,.$$

Die Möglichkeit, ein solches δ zu wählen, ist zum einen durch die zuvor bewiesene Charakterisierung des Riemann-Integrals und zum anderen die gleichmäßige Stetigkeit der stetigen Funktion c' auf dem Kompaktum $[a,b]$ garantiert.

Als Nächstes folgt der Hauptteil des Arguments, wobei die gerade erfolgte Wahl von δ erst in einem Moment wieder eine Rolle spielen wird. Aus der früheren Bemerkung (3f) und dem Schrankensatz erhält man direkt die Abschätzung $\mathrm{L}(c) \le (\operatorname{Lip} c)(b-a) = \big(\max_{[a,b]} \|c'\|_{\mathcal{X}}\big)(b-a) < \infty$ und somit die Rektifizierbarkeit von c. Nach Definition von $\mathrm{L}(c)$ lässt sich jetzt eine Zerlegung $a = t_0 < t_1 < t_2 < \ldots < t_m = b$ von $[a,b]$ mit

$$\mathrm{L}(c) - \varepsilon < \sum_{i=1}^{m} \|c(t_i) - c(t_{i-1})\|_{\mathcal{X}} \le \mathrm{L}(c)$$

wählen. Da die auftretende Summe bei Verfeinerung der Zerlegung höchstens größer wird, kann ohne Einschränkung $\max_{i \in \{1,2,\dots,m\}} (x_i - x_{i-1}) < \delta$ angenommen werden, und insbesondere gilt mit obiger Ungleichung auch

$$T_1 := \left| \mathrm{L}(c) - \sum_{i=1}^{m} \|c(t_i) - c(t_{i-1})\|_{\mathcal{X}} \right| < \varepsilon \,.$$

Seien jetzt $\tau_1 \in (t_0, t_1), \dots, \tau_m \in (t_{m-1}, t_m)$ beliebig fixiert. Gemäß Schrankensatz und Wahl von δ gilt $\|[c(t_i) - c'(\tau_i)t_i] - [c(t_{i-1}) - c'(\tau_i)t_{i-1}]\|_{\mathcal{X}} \le \big(\sup_{(t_{i-1}, t_i)} \|c' - c'(\tau_i)\|_{\mathcal{X}}\big)(t_i - t_{i-1}) \le \varepsilon(t_i - t_{i-1})$ für alle $i \in \{1, 2, \dots, m\}$. Mit der Dreiecksungleichung und der umgekehrten Dreiecksungleichung ergibt sich hieraus

$$\begin{aligned} T_2 &:= \left| \sum_{i=1}^{m} \|c(t_i) - c(t_{i-1})\|_{\mathcal{X}} - \sum_{i=1}^{m} \|c'(\tau_i)\|_{\mathcal{X}} (t_i - t_{i-1}) \right| \\ &\le \sum_{i=1}^{m} \|c(t_i) - c(t_{i-1}) - c'(\tau_i)(t_i - t_{i-1})\|_{\mathcal{X}} \le \sum_{i=1}^{m} \varepsilon (t_i - t_{i-1}) = (b-a)\varepsilon \,. \end{aligned}$$

Schließlich garantiert die Wahl von δ auch

$$T_3 := \left| \int_a^b \|c'(t)\|_{\mathcal{X}} \,\mathrm{d}t - \sum_{i=1}^m \|c'(\tau_i)\|_{\mathcal{X}}(t_i - t_{i-1}) \right| < \varepsilon \,,$$

und insgesamt erhält man

$$\left| \mathrm{L}(c) - \int_a^b \|c'(t)\|_{\mathcal{X}} \,\mathrm{d}t \right| \le T_1 + T_2 + T_3 \le (2+b-a)\varepsilon \,.$$

Wegen der Beliebigkeit von ε ist damit die Behauptung $\mathrm{L}(c) = \int_a^b \|c'(t)\|_{\mathcal{X}} \,\mathrm{d}t$ nachgewiesen. □

In vielen Zusammenhängen ist die genaue (Parametrisierung einer) Kurve $c\colon I \to \mathcal{X}$, vor allem die Durchlaufgeschwindigkeit $|c'|$, nebensächlich. Das eigentliche Objekt des Interesses ist nur das Bild der Kurve c, eventuell samt Durchlaufsinn, und dies ist ja auch genau das, was man in $\mathcal{X} = \mathbb{R}^2$ häufig durch eine Skizze des Bildes und Pfeile verdeutlicht. Mit anderen Worten liegt es oft nahe, Kurven mit gleichem Bild (und Durchlaufsinn) als äquivalent zu betrachten. Der genaue Begriff einer solchen Äquivalenz wird nun geprägt:

Definition (**Umparametrisierung von Kurven**)**.** *Seien I und J Intervalle, $\mathcal{X}$ ein normierter Raum sowie $c\colon I \to \mathcal{X}$ und $\widetilde{c}\colon J \to \mathcal{X}$ Kurven in $\mathcal{X}$. Gilt entweder $c = \widetilde{c}\circ\tau$ für eine surjektive, monotone Abbildung $\tau\colon I \to J$ oder $\widetilde{c} = c\circ\tau$ für eine surjektive, monotone Abbildung $\tau\colon J \to I$, so heißen $\widetilde{c}$ eine **Umparametrisierung** von c und τ die zugehörige **Parametertransformation**. Bei wachsendem τ spricht man von einer **Orientierung-erhaltenden** und bei fallendem τ von einer **Orientierung-umkehrenden** Umparametrisierung und Parametertransformation.*

Bemerkungen (zur Umparametrisierung von Kurven)**.**

(1) Als surjektive, monotone Abbildungen zwischen Intervallen sind die betrachteten Parametertransformationen τ automatisch stetig. Anders als in großen Teilen der Literatur wird für τ aber nur schwache Monotonie gefordert. Dies führt einerseits dazu, dass der nächste Satz in wünschenswerter Weise für beliebige rektifizierbare Kurven formuliert werden kann, aber andererseits auch dazu, dass solche τ Konstanzintervalle haben können und die Umkehrabbildung τ^{-1} nicht existieren muss. Die explizite Nennung der beiden Kompositionsreihenfolgen $c = \widetilde{c}\circ\tau$ und $\widetilde{c} = c\circ\tau$ in der Definition stellt sicher, dass die Umparametrisierung trotz dieser Problematik zumindest eine symmetrische Relation zwischen Kurven vermittelt.

(2) Ist $\widetilde{c}$ eine **Umparametrisierung** von c, so haben die Kurven c und $\widetilde{c}$ **dasselbe Bild** (das im Orientierung-erhaltenden Fall auch im selben Durchlaufsinn durchlaufen wird) und **dieselbe Länge**, sie sind aber i.A. **mit unterschiedlichen Geschwindigkeiten parametrisiert**.

(3) Genau genommen liefert das oben definierte Konzept der Umparametrisierung keine transitive Relation und deshalb keine Äquivalenzrelation auf Kurven. Dieses Problem besteht allerdings nur, wenn die betrachteten Kurven mit zwischenzeitlichem „Stehenbleiben“ in verschiedenen Punkten durchlaufen werden. Genauer denke man an eine injektive Kurve $\widehat{c}$ in $\mathcal{X}$ und zwei Umparametrisierungen $\widehat{c}\circ\tau$, $\widehat{c}\circ\widetilde{\tau}$ von $\widehat{c}$, so dass τ einen Wert $x_0 \in \widehat{c}(I)$ und $\widetilde{\tau}$ einen Wert $\widetilde{x}_0 \in \widehat{c}(I)$ auf einem Konstanzintervall positiver Länge annehmen, die jeweils andere Transformation diesen Wert aber nur einmal realisiert. In dieser Situation sind $\widehat{c}\circ\tau$ und $\widehat{c}\circ\widetilde{\tau}$ formal keine Umparametrisierungen voneinander, obwohl man sie mit

Blick sowohl auf die geometrische Situation als auch auf die Transitivität gerne als solche betrachten möchte. Man kann dieses Problem beseitigen, indem man eine Kurve $\widetilde{c}$ auch dann Umparametrisierung einer Kurve c nennt, wenn es nur eine weitere Kurve $\widehat{c}$ und surjektive, monotone Parametertransformationen τ und $\widetilde{\tau}$ zwischen Intervallen mit $c = \widehat{c} \circ \tau$ sowie $\widetilde{c} = \widehat{c} \circ \widetilde{\tau}$ gibt. Für dieses Konzept der Umparametrisierung im etwas weiteren Sinne lässt sich dann tatsächlich zeigen (zum Beispiel, indem man auf solche $\widehat{c}$ wie im nächsten Satz reduziert), dass es eine Äquivalenzrelation zwischen beliebigen Kurven vermittelt. Für alles Folgende ist diese Verfeinerung der Definition aber tatsächlich irrelevant, und es reicht, mit Umparametrisierungen im etwas engeren Sinn der obigen Definition zu arbeiten.

Wie oben erklärt, interessiert man sich statt für einzelne Kurven oft für (Äquivalenz-)Klassen von Kurven, die durch Orientierung-erhaltende Umparametrisierungen auseinander hervorgehen. Um in solchen Klassen einen ausgezeichneten Repräsentanten angeben zu können, wird als Nächstes eine besonders kanonische Parametrisierung von Kurven(klassen) definiert. Im Wesentlichen handelt es sich dabei um Parametrisierungen mit konstanter Durchlaufgeschwindigkeit 1:

Definition (**Parametrisierung nach der Bogenlänge**). *Sei I ein Intervall, sei $\mathcal{X}$ ein normierter Raum, und sei $c\colon I \to \mathcal{X}$ eine auf kompakten Teilintervallen von I rektifizierbare Kurve in $\mathcal{X}$.*

(I) *Eine Funktion $\lambda\colon I \to \mathbb{R}$ heißt **Bogenlängenfunktion** zu c, wenn*

$$\mathrm{L}\big(c|_{[a,b]}\big) = \lambda(b) - \lambda(a) \qquad \text{für alle Stellen } a < b \text{ in } I$$

gilt.

(II) *Die Kurve c heißt **nach der Bogenlänge parametrisiert**, wenn*

$$\mathrm{L}\big(c|_{[a,b]}\big) = b - a \qquad \text{für alle Stellen } a < b \text{ in } I$$

gilt.

Bemerkungen (zur Parametrisierung nach der Bogenlänge).

(1) Die Bogenlängenfunktion zu einer gegebenen Kurve c ist (nur) bis auf Addition einer reellen Konstante eindeutig. Dies ergibt sich direkt aus der Definition.

(2) Offensichtlich ist eine Kurve $c\colon I \to \mathcal{X}$ genau dann nach der Bogenlänge parametrisiert, wenn die identische Abbildung $I \to \mathbb{R}, t \mapsto t$ (bzw. jede Abbildung $I \to \mathbb{R}, t \mapsto t+\text{const}$) eine Bogenlängenfunktion zu c ist. Mit der Darstellungsformel für die Kurvenlänge und dem HDI sieht man zudem, dass die Parametrisierung nach der Bogenlänge für eine (Stückweise-) C^1-Kurve $c\colon I \to \mathcal{X}$ gerade Einheits-Durchlaufgeschwindigkeit $\|c'\|_{\mathcal{X}} \equiv 1$ auf (dem Definitionsbereich von c' in) I bedeutet.

Nach diesen Vorbereitungen kann gezeigt werden, dass eine Äquivalenzklasse von Kurven (im Sinn Orientierung-erhaltender Umparametrisierungen wie zuvor) unter schwachen Voraussetzungen einen nach der Bogenlänge parametrisierten Repräsentanten enthält.

Satz & Definition (Existenz einer **Umparametrisierung nach der Bogenlänge**). *Seien I ein Intervall, $\mathcal{X}$ ein normierter Raum und $c\colon I \to \mathcal{X}$ eine Kurve in $\mathcal{X}$. Ist c auf kompakten Teilintervallen von I rektifizierbar, so gibt es eine* stetige *Bogenlängenfunktion λ zu c und eine nach der Bogenlänge parametrisierte Umparametrisierung $\widehat{c}$ von c, die aus c durch die Orientierung-erhaltende Parametertransformation λ gemäß $c = \widehat{c} \circ \lambda$ auf I hervorgeht. Man nennt $\widehat{c}$ die (bis auf Parameterverschiebung eindeutige) Umparametrisierung von c nach der Bogenlänge.*

Beweis. Sei ein beliebiges $t_* \in I$ fixiert. Aus der Zerlegungsadditivität der Kurvenlänge folgt, dass durch die Festlegungen $\lambda(t) := \mathrm{L}\big(c|_{[t_*,t]}\big)$ für $t \in I_{\geq t_*}$ und $\lambda(t) := -\mathrm{L}\big(c|_{[t,t_*]}\big)$ für $t \leq I_{\leq t_*}$ eine monoton wachsende Bogenlängenfunktion $\lambda\colon I \to \lambda(I)$ gegeben ist, die auf ihr Bild $\lambda(I)$ natürlich auch surjektiv abbildet. Entscheidend ist nun der Nachweis der Stetigkeit von λ. Dazu seien zunächst beliebige Stellen $a < b$ in $I_{\geq t_*}$ und $\varepsilon \in \mathbb{R}_{>0}$ fixiert. Wegen gleichmäßiger Stetigkeit von c auf dem Kompaktum $[a,b]$ gibt es ein $\delta > 0$, so dass für $t, \widetilde{t} \in [a,b]$ mit $|\widetilde{t}-t| < \delta$ stets $\|c(\widetilde{t})-c(t)\|_{\mathcal{X}} < \varepsilon$ gilt. Nach Definition der Kurvenlänge gibt es zudem $t_0 < t_1 < t_2 < \ldots < t_m$ in $[a,b]$ mit

$$\sum_{i=1}^{m} \|c(t_i)-c(t_{i-1})\|_{\mathcal{X}} > \mathrm{L}\big(c|_{[a,b]}\big) - \varepsilon\,,$$

wobei $t_1-t_0 < \delta$ angenommen werden kann (denn die Näherungssumme an die Kurvenlänge wird durch Einfügen zusätzlicher Stellen höchstens größer). Es folgt

$$\begin{aligned}\lambda(t_1)-\lambda(a) &= \mathrm{L}\big(c|_{[a,t_1]}\big) = \mathrm{L}\big(c|_{[a,b]}\big) - \mathrm{L}\big(c|_{[t_1,b]}\big)\\ &< \varepsilon + \|c(t_1)-c(t_0)\|_{\mathcal{X}} + \sum_{i=2}^{m} \|c(t_i)-c(t_{i-1})\|_{\mathcal{X}} - \mathrm{L}\big(c|_{[t_1,b]}\big) < 2\varepsilon\,,\end{aligned}$$

weil aus $0 < t_1-t_0 < \delta$ und der Wahl von δ die Abschätzung $\|c(t_1)-c(t_0)\|_{\mathcal{X}} < \varepsilon$ folgt, und weil $\mathrm{L}\big(c|_{[t_1,b]}\big)$ per Definition mindestens so groß wie $\sum_{i=2}^{m} \|c(t_i)-c(t_{i-1})\|_{\mathcal{X}}$ ist. Da λ monoton wachsend ist, reicht dies zum Nachweis rechtsseitiger Stetigkeit von λ auf $I_{\geq t_*}$. Die Nachweise der linksseitigen Stetigkeit von λ auf $I_{\geq t_*}$ und der rechts- und linksseitigen Stetigkeit von λ auf $I_{\leq t_*}$ verlaufen analog. Insgesamt ist λ auf ganz I stetig, nach dem Zwischenwertsatz ist $\lambda(I)$ ein Intervall, und somit ist λ eine zulässige Parametertransformation. Nun wird $\widehat{c}\colon \lambda(I) \to \mathcal{X}$ für $s \in \lambda(I)$ durch

$$\widehat{c}(s) := c(t) \text{ für ein } t \in I \text{ mit } \lambda(t) = s$$

festgelegt. Dies ist wohldefiniert, weil im Fall $\lambda(t_1) = \lambda(t_2)$ mit $t_1 < t_2$ stets $\mathrm{L}\big(c|_{[t_1,t_2]}\big) = \lambda(t_2)-\lambda(t_1) = 0$ und damit auch $c(t_1) = c(t_2)$ gilt (und tatsächlich die Konstanzintervalle von λ und c übereinstimmen). Offensichtlich gilt $\widehat{c} \circ \lambda = c$. Für $s_1 = \lambda(t_1) < s_2 = \lambda(t_2)$ gilt wegen Monotonie stets $t_1 < t_2$, und dann folgt mit $\|\widehat{c}(s_2)-\widehat{c}(s_1)\|_{\mathcal{X}} = \|c(t_2)-c(t_1)\|_{\mathcal{X}} \leq \mathrm{L}\big(c|_{[t_1,t_2]}\big) = \lambda(t_2)-\lambda(t_1) = s_2-s_1$ Stetigkeit von $\widehat{c}$, so dass $\widehat{c}$ als Kurve und Orientierung-erhaltende Umparametrisierung von c bezeichnet werden darf. Über die Parametrisierungsinvarianz der Kurvenlänge folgt mit $\mathrm{L}\big(\widehat{c}|_{[s_1,s_2]}\big) = \mathrm{L}\big(c|_{[t_1,t_2]}\big) = \lambda(t_2)-\lambda(t_1) = s_2 - s_1$ weiter, dass $\widehat{c}$ nach der Bogenlänge parametrisiert ist. □

Verfahren (Berechnung der Bogenlängen-Umparametrisierung). *Um die Umparametrisierung einer gegebenen* C^1*-Kurve* $c\colon I \to \mathcal{X}$ *nach der Bogenlänge konkret zu berechnen, bestimmt man, soweit möglich, sukzessive …*

(1) *eine* ***Bogenlängenfunktion*** $\lambda\colon I \to \mathbb{R}$ *zu* c *durch Integralberechnung*

$$\lambda(t) = \int \|c'(t)\|_{\mathcal{X}}\,\mathrm{d}t \qquad \text{für } t \in I$$

(wobei jede Stammfunktion λ *zu* $\|c'\|$ *eine Bogenlängenfunktion zu* c *ist und man alle solchen durch Addition reeller Konstanten zu diesem* λ *erhält),*

(2) *eine* ***Umkehrabbildung*** $\lambda^{-1}\colon \lambda(I) \to I$ *zu* λ *(oder, falls* λ *Konstanzintervalle hat, zumindest eine Rechtsinverse* λ^{-1} *mit* $\lambda(\lambda^{-1}(s)) = s$ *für* $s \in \lambda(I)$*),*

(3) *die im Wesentlichen eindeutige **Umparametrisierung** $\widehat{c}$ von c **nach der Bogenlänge** als $\widehat{c}(s) = c(\lambda^{-1}(s))$ für $s \in \lambda(I)$.*

Auch wenn c durch elementare Funktionen gegeben ist, kann hier aber keinesfalls ein elementares Ergebnis garantiert werden. Sowohl die zu bildende Stamm- als auch die Umkehrfunktion lassen sich in vielen Fällen nicht elementar angeben.

Beispiel (Bogenlängen-Umparametrisierung von Kreislinien)**.** Für $r, \omega \in \mathbb{R}_{>0}$ und $x_0, y_0 \in \mathbb{R}$ sei eine Kurve $c\colon \to \mathbb{R}^2$ durch

$$c(t) := (x_0{+}r\cos(\omega t), y_0{+}r\sin(\omega t)) \qquad \text{für } t \in \mathbb{R}$$

gegeben. Dann verläuft obiges Verfahren wie folgt:

(1) Wegen $|c'| \equiv rw$ erhält man eine Bogenlängenfunktion λ als $\lambda(t) = r\omega t{+}s_0$ mit Integrationskonstante $s_0 \in \mathbb{R}$ (die man für die meisten Zwecke auch gleich Null setzen kann).

(2) Durch Auflösen nach t ergibt sich die zugehörige Umkehrfunktion λ^{-1} als $\lambda^{-1}(s) = \frac{s-s_0}{r\omega}$.

(3) Einsetzen führt auf die Bogenlängen-Umparametrisierung $\widehat{c}$ von c mit

$$\widehat{c}(s) = c(\lambda(s)) = \Big(x_0{+}r\cos\frac{s-s_0}{r}, y_0{+}r\sin\frac{s-s_0}{r}\Big) \qquad \text{für } s \in \mathbb{R}\,.$$

1.2 Kurvenintegrale

Definitionen (absolute und orientierte Kurvenintegrale). *Sei $a < b$ in $\mathbb{R}$, seien $\mathcal{X}$, $\mathcal{Y}$ und $\mathcal{Z}$ normierte Räume, sei $c\colon [a,b] \to \mathcal{X}$ eine Kurve in $\mathcal{X}$, und sei $f\colon c([a,b]) \to \mathcal{Y}$ eine beschränkte Funktion auf der Spur von c. Unter Verwendung von Zwischensummen zu $m \in \mathbb{N}$, zu Zerlegungen $a = t_0 < t_1 < t_2 < \ldots < t_{m-1} < t_m = b$ von $[a,b]$ und zu Zwischenstellen $\tau_1 \in (t_0,t_1)$, $\tau_2 \in (t_1,t_2)$, $\ldots$, $\tau_m \in (t_{m-1},t_m)$ trifft man folgende Festlegungen, bei denen die Grenzübergänge analog zu den in Abschnitt 1.1 definierten zu verstehen sind:*

(1) *Existiert der Limes der Zwischensummen $\sum_{i=1}^m f(c(\tau_i))\|c(t_i){-}c(t_{i-1})\|_{\mathcal{X}}$ bei gegen Null strebender Zerlegungsfeinheit in $\mathcal{Y}$, so heißt sein Wert das **absolute Kurvenintegral** von f über c und wird mit*

$$\int_c f(x)\,\|\mathrm{d}x\|_{\mathcal{X}} = \int_a^b (f\circ c)\,\|\mathrm{d}c\|_{\mathcal{X}} = \int_a^b f(c(t))\,\|\mathrm{d}c(t)\|_{\mathcal{X}} \in \mathcal{Y}$$

bezeichnet.

(2) *Sei $*\colon \mathcal{Y}{\times}\mathcal{X} \to \mathcal{Z}$ eine bilineare, stetige Abbildung. Existiert der Limes der Zwischensummen $\sum_{i=1}^m f(c(\tau_i)) * \big[c(t_i){-}c(t_{i-1})\big]$ bei gegen Null strebender Zerlegungsfeinheit in $\mathcal{Z}$, so heißt sein Wert das **orientierte Kurvenintegral** von f über c **bezüglich des Produkts** $*$ und wird mit*

$$\int_c f(x) * \mathrm{d}x = \int_a^b (f\circ c) * \mathrm{d}c = \int_a^b f(c(t)) * \mathrm{d}c(t) \in \mathcal{Z}$$

bezeichnet.

Bemerkungen (zu Spezialfällen des *absoluten* Kurvenintegrals)**.**

(A) Spezialisiert man die Definition auf eine rektifizierbare Kurve c in $\mathcal{X}$ und den konstanten Integranden $f \equiv 1$ mit Werten in $\mathcal{Y} = \mathbb{R}$, so erhält man mit

$$\boxed{\int_c 1 \,\|\mathrm{d}x\|_{\mathcal{X}} = \mathrm{L}(c)}$$

das Konzept der Kurvenlänge zurück. Dies ist insoweit plausibel, dass beide Seiten der Gleichung über die gleichen Näherungssummen definiert sind. Um einzusehen, dass in diesem Fall auch zwischen der Supremumbildung aus der Definition der Kurvenlänge und der Limesbildung aus der Definition des Kurvenintegrals kein Unterschied besteht, braucht man aber noch ein Zusatzargument oder demnächst folgende Theorie.

(B) Im wichtigsten Fall $\mathcal{X} = \mathbb{R}^N$ bildet man das absolute Kurvenintegral praktisch immer bezüglich der 2-Norm und verwendet analoge Notationen mit $|\mathrm{d}x|$, $|\mathrm{d}c|$, $|\mathrm{d}c(t)|$ anstelle von $\|\mathrm{d}x\|_{\mathcal{X}}$, $\|\mathrm{d}c\|_{\mathcal{X}}$, $\|\mathrm{d}c(t)\|_{\mathcal{X}}$.

Bemerkungen (zu Spezialfällen des *orientierten* Kurvenintegrals)**.**

(C) Im Fall $\mathcal{X} = \mathcal{Y} = \mathcal{Z} = \mathbb{C}$ und mit der komplexen Multiplikation als Produkt $*$ ergibt sich das komplexe Kurvenintegral

$$\int_c h(z)\,\mathrm{d}z \in \mathbb{C}\,,$$

das für Kurven c in $\mathbb{C}$ und $\mathbb{C}$-wertige Funktionen h einer komplexen Variablen erklärt ist. Dieser Integralbegriff ist in der Funktionentheorie, der Theorie von Funktionen komplexer Variablen, von zentraler Bedeutung.

(D) Im Fall $\mathcal{X} = \mathcal{Y} = \mathbb{R}^N$, $\mathcal{Z} = \mathbb{R}$ mit dem Euklidischen Skalarprodukt $\cdot$ des $\mathbb{R}^N$ als Produkt $*$ ergibt sich das **orientierte Kurvenintegral**

$$\int_c V(x) \cdot \mathrm{d}x \in \mathbb{R}$$

eines Vektorfelds $V\colon D \to \mathbb{R}^N$ auf $D \subset \mathbb{R}^N$ über eine Kurve $c\colon [a,b] \to D$.

(E) Im Fall $\mathcal{X} = \mathbb{R}^N$, $\mathcal{Y} = \big(\mathbb{R}^N\big)^* := \mathcal{L}(\mathbb{R}^N, \mathbb{R})$, $\mathcal{Z} = \mathbb{R}$ mit der Auswertung $\langle\,\cdot\,,\,\cdot\,\rangle$ von Linearformen auf Vektoren als Produkt $*$ ergibt sich das **Kurvenintegral**

$$\int_c \Lambda := \int_c \langle \Lambda(x), \mathrm{d}x\rangle \in \mathbb{R}$$

einer sogenannten **1-Form** $\Lambda\colon D \to \big(\mathbb{R}^N\big)^*$ auf $D \subset \mathbb{R}^N$ über eine Kurve $c\colon [a,b] \to D$.

Die für den weiteren Verlauf der Vorlesung **wichtigsten Spezialfälle** (D) **und** (E) **entsprechen einander** durch Transponieren, genauer durch $\Lambda = V^{\mathrm{T}}$, wenn man Linearformen in $(\mathbb{R}^N)^*$ wie üblich mit den diese darstellenden Zeilenvektoren in $\mathbb{R}^{1\times N}$ identifiziert (so dass die Auswertung $\langle\,\cdot\,,\,\cdot\,\rangle$ nichts anderes ist als die Matrix-Multiplikation eines Zeilenvektors aus $\mathbb{R}^{1\times N}$ mit einem Spaltenvektor aus $\mathbb{R}^N = \mathbb{R}^{N\times 1}$).

Bemerkung (zu Kurvenintegralen mit Belegung)**.** Ist eine beschränkte Belegung $\omega\colon [a,b] \to \mathbb{R}$ einer Kurve $c\colon [a,b] \to \mathcal{X}$ gegeben, so erklärt man das absolute Kurvenintegral $\int_a^b (f \circ c)\,\omega \|\mathrm{d}c\|_{\mathcal{X}}$ einer beschränkten Funktion f über c bezüglich der Belegung ω als Limes von $\sum_{i=1}^m f(c(\tau_i))\omega(\tau_i)\|c(t_i){-}c(t_{i-1})\|_{\mathcal{X}}$ bei gegen Null strebender Zerlegungsfeinheit. Das orientierte Kurvenintegral $\int_a^b (f \circ c) * \omega\,\mathrm{d}c$ bezüglich der Belegung ω ist analog zu verstehen. Wenn ω nicht nach c faktorisiert, sind diese Bildungen echte Verallgemeinerungen der zuvor gegebenen Definition.

Eigenschaften (von Kurvenintegralen)**.** Seien $a < b$ in $\mathbb{R}$, seien $\mathcal{X}$, $\widetilde{\mathcal{X}}$, $\mathcal{Y}$ und $\mathcal{Z}$ normierte Räume über $\mathbb{K}$, sei $c\colon [a,b] \to \mathcal{X}$ eine Kurve in $\mathcal{X}$, seien $f,g\colon c([a,b]) \to \mathcal{Y}$ beschränkte Funktionen, und sei $*\colon \mathcal{Y} \times \mathcal{X} \to \mathcal{Z}$ eine $\mathbb{K}$-bilineare und stetige Abbildung. Dann gelten folgende Eigenschaften von Kurvenintegralen, die für den Spezialfall der Kurvenlänge größtenteils schon bekannt sind, und bei denen hier unterstellt wird, dass alle beteiligten Integrale (bei Ungleichungen) oder zumindest die Integrale auf der rechten Seite (bei Gleichungen) a priori existieren:

(0) Komponenten-Regel (im Fall $\mathcal{Y} = \mathbb{K}^N$ für $\mathbb{K}^N$-wertiges $f = (f_1, f_2, \ldots, f_N)$)

$$\int_c f(x)\,\|\mathrm{d}x\|_{\mathcal{X}} = \left(\int_c f_1(x)\,\|\mathrm{d}x\|_{\mathcal{X}}, \int_c f_2(x)\,\|\mathrm{d}x\|_{\mathcal{X}}, \ldots, \int_c f_N(x)\,\|\mathrm{d}x\|_{\mathcal{X}} \right),$$

(1) **Linearität** im Integranden (für $s, \widetilde{s} \in \mathbb{K}$)

$$\int_c \big[s f(x) + \widetilde{s} g(x) \big]\,\|\mathrm{d}x\|_{\mathcal{X}} = s \int_c f(x)\,\|\mathrm{d}x\|_{\mathcal{X}} + \widetilde{s} \int_c g(x)\,\|\mathrm{d}x\|_{\mathcal{X}}\,,$$

$$\int_c \big[s f(x) + \widetilde{s} g(x) \big] * \mathrm{d}x = s \int_c f(x) * \mathrm{d}x + \widetilde{s} \int_c g(x) * \mathrm{d}x\,,$$

(2) **Zerlegungsadditivität** (mit $a < t_* < b$)

$$\int_a^b (f \circ c)\,\|\mathrm{d}c\|_{\mathcal{X}} = \int_a^{t_*} (f \circ c)\,\|\mathrm{d}c\|_{\mathcal{X}} + \int_{t_*}^b (f \circ c)\,\|\mathrm{d}c\|_{\mathcal{X}}\,,$$

$$\int_a^b (f \circ c) * \mathrm{d}c = \int_a^{t_*} (f \circ c) * \mathrm{d}c + \int_{t_*}^b (f \circ c) * \mathrm{d}c\,,$$

(3) **Parametrisierungsinvarianz** (beim *orientierten* Kurvenintegral aber *keine* Invarianz unter Orientierungsumkehr; beim orientierten Integral ist die Orientierung damit relevant)

$$\int_{c \circ \tau} f(x)\,\|\mathrm{d}x\|_{\mathcal{X}} = \int_c f(x)\,\|\mathrm{d}x\|_{\mathcal{X}}\,,$$

$$\int_{c \circ \tau} f(x) * \mathrm{d}x = \mathrm{Or}(\tau) \int_c f(x) * \mathrm{d}x$$

für surjektive, monotone Parametertransformationen $\tau\colon [\alpha, \beta] \to [a,b]$ mit $\mathrm{Or}(\tau) := 1$ im Orientierung-erhaltenden und mit $\mathrm{Or}(\tau) := -1$ im Orientierung-umkehrenden Fall,

(4) Monotonie des absoluten Kurvenintegrals im Integranden (für f, g mit Werten in $\mathcal{Y} = \mathbb{R}$)

$$f \leq g \text{ auf } c([a,b]) \implies \int_c f(x)\,\|\mathrm{d}x\|_{\mathcal{X}} \leq \int_c g(x)\,\|\mathrm{d}x\|_{\mathcal{X}}\,,$$

(5) **Dreiecksungleichung** (auch fundamentale Abschätzung oder Schrankensatz genannt)

$$\left\| \int_c f(x)\,\|\mathrm{d}x\|_{\mathcal{X}} \right\|_{\mathcal{Y}} \leq \int_c \|f(x)\|_{\mathcal{Y}}\,\|\mathrm{d}x\|_{\mathcal{X}} \leq \Big(\sup_{c([a,b])} \|f\|_{\mathcal{Y}} \Big) \mathrm{L}(c)\,,$$

$$\left\| \int_c f(x) * \mathrm{d}x \right\|_{\mathcal{Z}} \leq K_* \int_c \|f(x)\|_{\mathcal{Y}}\,\|\mathrm{d}x\|_{\mathcal{X}} \leq K_* \Big(\sup_{c([a,b])} \|f\|_{\mathcal{Y}} \Big) \mathrm{L}(c)$$

mit der Operatornorm K_* des Produkts $*$, d.h. der kleinstmöglichen Konstante K_* in der Produktabschätzung $\|y * x\|_{\mathcal{Z}} \leq K_* \|y\|_{\mathcal{Y}} \|x\|_{\mathcal{X}}$, für konkrete Produkte $*$ meist mit $K_* = 1$,

(6) **Isometrieinvarianz** des absoluten Kurvenintegrals (für eine Isometrie $h\colon c([a,b]) \to \widetilde{\mathcal{X}}$)

$$\int_{h\circ c} f(h^{-1}(\widetilde{x}))\,\|\mathrm{d}\widetilde{x}\|_{\widetilde{\mathcal{X}}} = \int_c f(x)\,\|\mathrm{d}x\|_{\mathcal{X}}\,,$$

Translationsinvarianz des orientierten Kurvenintegrals (für $x_0 \in \mathcal{X}$)

$$\int_{x_0+c} f(x-x_0) * \mathrm{d}x = \int_c f(x) * \mathrm{d}x$$

und **Skalierungsverhalten** (für $s \in \mathbb{K} \setminus \{0\}$)

$$\int_{sc} f\Big(\frac{x}{s}\Big)\,\|\mathrm{d}x\|_{\mathcal{X}} = |s| \int_c f(x)\,\|\mathrm{d}x\|_{\mathcal{X}}\,, \qquad \int_{sc} f\Big(\frac{x}{s}\Big) * \mathrm{d}x = s\int_c f(x) * \mathrm{d}x\,.$$

All diese Eigenschaften lassen sich recht problemlos an den Definitionen und den zugehörigen Zwischensummen ablesen. Beim Nachweis der Zerlegungsadditivität ist zusätzlich zu argumentieren, dass Verfeinerung an der Zwischenstelle t_* die Zwischensummen nur geringfügig verändert.

Zur konkreten Berechnung von Kurvenintegralen dienen folgende Formeln:

Satz (Darstellung der **Kurvenintegrale als Riemann-Integrale**). *Sei $a<b$ in $\mathbb{R}$, seien $\mathcal{X}$, $\mathcal{Y}$ und $\mathcal{Z}$ normierte Räume, sei $c\colon [a,b] \to \mathcal{X}$ eine C^1-Kurve in $\mathcal{X}$, sei $f\colon c([a,b]) \to \mathcal{Y}$ eine stetige Funktion, und sei $*\colon \mathcal{Y}\times\mathcal{X} \to \mathcal{Z}$ bilinear und stetig.*

(I) *Bei vollständigem $\mathcal{Y}$ existiert das absolute Kurvenintegral*

$$\boxed{\int_c f(x)\,\|\mathrm{d}x\|_{\mathcal{X}} = \int_a^b f(c(t))\|c'(t)\|_{\mathcal{X}}\,\mathrm{d}t} \in \mathcal{Y}\,.$$

(II) *Bei vollständigem $\mathcal{Z}$ existiert das orientierte Kurvenintegral*

$$\boxed{\int_c f(x) * \mathrm{d}x = \int_a^b f(c(t)) * c'(t)\,\mathrm{d}t} \in \mathcal{Z}\,.$$

Beweisskizze. Die Existenz der Riemann-Integrale auf den rechten Seiten wird hier als bekannt vorausgesetzt, im Wesentlichen wird das dafür relevante Argument aber im Beweis des nachfolgenden Satzes ausgeführt. Der Beweis der Formeln und damit auch der Existenz der Integrale auf der linken Seite verläuft weitgehend analog zum in Abschnitt 1.1 gegebenen Beweis der Formel $\mathrm{L}(c) = \int_a^b \|c'(t)\|_{\mathcal{X}}\,\mathrm{d}t$ für die Kurvenlänge und gestaltet sich sogar etwas einfacher, da jetzt beide Seiten als Limites bei gegen Null strebender Zerlegungsfeinheit betrachtet werden können. Tatsächlich kann die Differenz

$$\Bigg|\sum_{i=1}^m f(c(\tau_i))\|c(t_i)-c(t_{i-1})\|_{\mathcal{X}} - \sum_{i=1}^m f(c(\tau_i))\|c'(\tau_i)\|_{\mathcal{X}}(t_i-t_{i-1})\Bigg|$$

der für Teil (I) relevanten Näherungssummen genau wie im früheren Beweis über den Schrankensatz abgeschätzt werden, wenn nur die Terme $f(c(\tau_i))$ gleichzeitig durch $\sup_{[a,b]}|f\circ c| < \infty$ kontrolliert und herausgezogen werden. Man erhält damit, dass diese Differenz bei ausreichend kleiner Zerlegungsfeinheit beliebig klein wird, und insgesamt ergibt sich die Behauptung des Teils (I). Teil (II) behandelt man ganz analog mit Hilfe der entsprechenden Näherungssummen. □

Bemerkungen (zur den Darstellungsformeln für Kurvenintegrale)**.**

(1) Wie die entsprechende Formel für die Kurvenlänge bleiben auch die Darstellungsformeln des vorausgehenden Satzes für nur *Stückweise*-C^1-Kurven c richtig. Bei unbeschränktem c' sind die rechten Seiten dabei wieder im Sinn uneigentlicher Integrale zu interpretieren.

(2) Speziell für nach der Bogenlänge parametrisierte C^1-Kurven $c\colon [a,b] \to \mathcal{X}$ und stetige $f\colon c([a,b]) \to \mathcal{Y}$ besagt Teil (I) des Satzes

$$\int_c f(x)\,\|\mathrm{d}x\|_{\mathcal{X}} = \int_a^b f(c(t))\mathrm{d}t \qquad\qquad (\text{wobei } b = a{+}\mathrm{L}(c) \text{ gilt})\,.$$

In den Übungen soll gezeigt werden, dass die letzte Formel sogar für beliebige nach der Bogenlänge parametrisierte Kurven $c\colon [a,b] \to \mathcal{X}$ ohne irgendeine Differenzierbarkeitsannahme gilt; dies erfordert allerdings eine erneute und technisch etwas anspruchsvollere Argumentation mit Näherungssummen.

Auch die Existenzaussage des vorigen Satzes lässt sich verallgemeinern; sie gilt allgemein für rektifizierbare Kurven, ohne dass man eine Differenzierbarkeitsannahme an diese benötigt:

Satz (Existenz von Kurvenintegralen über lediglich rektifizierbare Kurven)**.** *Sei $a < b$ in $\mathbb{R}$, seien $\mathcal{X}$, $\mathcal{Y}$ und $\mathcal{Z}$ normierte Räume, sei $c\colon [a,b] \to \mathcal{X}$ eine rektifizierbare Kurve in $\mathcal{X}$, sei $f\colon c([a,b]) \to \mathcal{Y}$ eine stetige Funktion, und sei schließlich $*\colon \mathcal{Y}\times\mathcal{X} \to \mathcal{Z}$ bilinear und stetig. Dann existiert bei vollständigem $\mathcal{Y}$ das absolute Kurvenintegral $\int_c f(x)\,\|\mathrm{d}x\|_{\mathcal{X}} \in \mathcal{Y}$, und bei vollständigem $\mathcal{Z}$ existiert das orientierte Kurvenintegral $\int_c f(x) * \mathrm{d}x \in \mathcal{Z}$.*

Beweisskizze. Es reicht, die Existenz des „normalen" Riemann-Integrals $\int_a^b f(c(t))\,\mathrm{d}t \in \mathcal{Y}$ und des orientierten Kurvenintegrals $\int_a^b f(c(t)) * \mathrm{d}c(t) \in \mathcal{Z}$ nachzuweisen, denn der Fall des absoluten Kurvenintegrals $\int_a^b f(c(t))\,\|\mathrm{d}c(t)\| \in \mathcal{Y}$ kann mittels Umparametrisierung nach der Bogenlänge und der in der vorausgehenden Bemerkung (2) erwähnten Übungsaufgabe darauf zurückgeführt werden.

Zuerst wird das Integral $\int_a^b f(c(t))\,\mathrm{d}t$ behandelt, das im hier betrachteten $\mathcal{Y}$-wertigen Fall in Abschnitt 1.1 über Zwischensummen definiert wurde. Es bezeichne W_δ die Menge aller Werte solcher Zwischensummen $\sum_{i=1}^m f(c(\tau_i))(t_i{-}t_{i-1})$ zu Zerlegungen $a = t_0 < t_1 < t_2 < \ldots < t_m = b$ der Feinheit kleiner als δ und zu Zwischenstellen $\tau_i \in (t_{i-1},t_i)$. Zu gegebenem $\varepsilon \in \mathbb{R}_{>0}$ lässt sich wegen gleichmäßiger Stetigkeit von $f\circ c$ auf $[a,b]$ jetzt $\delta \in \mathbb{R}_{>0}$ so wählen, dass für $s,\widetilde{s} \in [a,b]$ mit $|\widetilde{s}{-}s| < 2\delta$ stets $\|f(c(\widetilde{s})){-}f(c(s))\|_{\mathcal{Y}} < \varepsilon$ gilt. Für Werte $w,\widetilde{w} \in W_\delta$ von Zwischensummen wird nun $\|\widetilde{w}{-}w\|_{\mathcal{Y}} \le \varepsilon(b{-}a)$ begründet. Dazu ist es praktisch, $w = \mathrm{I}_a^b(g)$ und $\widetilde{w} = \mathrm{I}_a^b(\widetilde{g})$ in der naheliegenden Weise als elementare Integrale von Treppenfunktionen $g,\widetilde{g}\colon [a,b] \to \mathcal{Y}$ zu schreiben. Für jedes $t \in [a,b]$ (außer, eventuell, endlich vielen Sprungstellen) gelten dann $g(t) = f(c(s))$ und $\widetilde{g}(t) = f(c(\widetilde{s}))$ für gewisse $s,\widetilde{s} \in [a,b]$ mit $|s{-}t| < \delta$, $|\widetilde{s}{-}t| < \delta$, so dass sich wie behauptet $\|\widetilde{w}{-}w\|_{\mathcal{Y}} \le (\sup_{[a,b]} \|\widetilde{g}{-}g\|_{\mathcal{Y}})(b{-}a) \le \varepsilon(b{-}a)$ ergibt. Damit ist gezeigt, dass zu jedem $\varepsilon \in \mathbb{R}_{>0}$ ein $\delta \in \mathbb{R}_{>0}$ mit $\operatorname{diam} W_\delta \le \varepsilon(b{-}a)$ existiert. Um die Existenz des Integrals zu sichern, reicht es nun, für Folgen $(\delta_k)_{k\in\mathbb{N}}$ (von Zerlegungsfeinheiten) in $\mathbb{R}_{>0}$ und $(w_k)_{k\in\mathbb{N}}$ (von Zwischensummenwerten) in $\mathcal{Y}$ mit $w_k \in W_{\delta_k}$ für alle $k \in \mathbb{N}$ und $\lim_{k\to\infty} \delta_k = 0$ zu verifizieren, dass $\lim_{k\to\infty} w_k \in \mathcal{Y}$ existiert (denn durch Argumentation mit Teilfolgen ist dann auch klar, dass der Wert des Limes nicht von den betrachteten Folgen abhängt). Die zuvor nachgewiesene Eigenschaft der W_δ bedeutet aber nichts anderes, als dass solche Folgen $(w_k)_{k\in\mathbb{N}}$ Cauchy-Folgen in $\mathcal{Y}$ sind, mit der Vollständigkeit von $\mathcal{Y}$ folgen daher die Existenz des Limes $\lim_{k\to\infty} w_k$ und des Integrals $\int_a^b f(c(t))\,\mathrm{d}t$ in $\mathcal{Y}$.

Das orientierte Kurvenintegral $\int_a^b f(c(t)) * \mathrm{d}c(t)$ lässt sich mit einer sehr ähnlichen Argumentation behandeln, bei der man natürlich mit der Menge W_δ der Werte von Zwischensummen $\sum_{i=1}^m f(c(\tau_i)) * (c(t_i){-}c(t_{i-1}))$ zu Feinheiten kleiner δ arbeitet. Wählt man δ genau wie zuvor, so erhält man für $w,\widetilde{w} \in W_\delta$ jetzt die Abschätzung $\|\widetilde{w}{-}w\|_{\mathcal{Z}} \le K_*\varepsilon\mathrm{L}(c)$ mit der Operatornorm K_* des Produkts $*$; dies ergibt sich daraus, dass auch hier — aber übrigens nicht[5] beim absoluten Kurvenintegral — der Wert $w = \mathrm{I}_a^b(g*\mathrm{d}c) := \sum_{i=1}^m f(c(\tau_i))*(c(t_i){-}c(t_{i-1}))$ als elementares Integral einer Treppenfunktion g mit $g \equiv f(c(\tau_i))$ auf (t_{i-1},t_i) und analog $\widetilde{w} = \mathrm{I}_a^b(\widetilde{g} * \mathrm{d}c)$ geschrieben werden kann. Somit gibt es also zu jedem $\varepsilon \in \mathbb{R}_{>0}$ ein $\delta \in \mathbb{R}_{>0}$ mit $\operatorname{diam} W_\delta \le K_*\varepsilon\mathrm{L}(c)$. Ausgehend von dieser Abschätzung und der Vollständigkeit von $\mathcal{Z}$ ergibt sich die Existenz des Integrals genau wie zuvor. ☐

[5] Näherungssummen $\sum_{i=1}^m f(c(\tau_i))\|c(t_i){-}c(t_{i-1})\|$ an das absolute Kurvenintegral kann man i.A. nicht als elementare Integrale einer Funktion g mit $g \equiv f(c(\tau_i))$ auf (t_{i-1},t_i) auffassen. Der Grund hierfür ist, dass der Wert solcher Summen auch von Zerlegungsstellen t_i abhängt, bei denen der Wert von g sich nicht ändert, solche Stellen in der Funktion g aber nicht kodiert sind.

1.3 Stammfunktionen zu Vektorfeldern und 1-Formen

In diesem Abschnitt wird die Frage behandelt, unter welchen Umständen Stammfunktionen zu Vektorfeldern und 1-Formen existieren und wie sie gegebenenfalls berechnet werden können. Als Erstes wird dazu das Konzept der Stammfunktion im Fall von Vektorfeldern präzisiert:

Definition (Stammfunktionen zu Vektorfeldern). *Sei D offen in $\mathbb{R}^N$. Man bezeichnet $f\colon D \to \mathbb{R}$ als **Stamm- oder Potentialfunktion** eines Vektorfelds $V\colon D \to \mathbb{R}^N$ auf D, wenn f auf D total differenzierbar ist und V das Gradientenfeld von f ist, also $\nabla f = V$ auf D gilt.*

Unter einer 1-Form, daran sei hier erinnert, versteht man eine Abbildung $\Lambda\colon D \to \big(\mathbb{R}^N\big)^*$ von einem Definitionsbereich $D \subset \mathbb{R}^N$ in den Dualraum $\big(\mathbb{R}^N\big)^* := \mathcal{L}(\mathbb{R}^N, \mathbb{R})$, den Raum der (stetigen) Linearformen auf $\mathbb{R}^N$. Insbesondere kann die totale Ableitung $f'\colon D \to \mathcal{L}(\mathbb{R}^N, \mathbb{R})$ einer $\mathbb{R}$-wertigen Funktion $f\colon D \to \mathbb{R}$ auf offenem $D \subset \mathbb{R}^N$ als 1-Form auf D aufgefasst werden, und wird bei dieser Betrachtung vorzugsweise als **Differential $\mathbf{d}\boldsymbol{f} := \boldsymbol{f'}$** von f bezeichnet. Da 1-Formen Λ auf $D \subset \mathbb{R}^N$ mit Vektorfeldern V auf D identifiziert werden können, lässt sich das Konzept der Stammfunktion in die Sprache von 1-Formen übertragen, und man erhält:

Definition (Stammfunktionen zu 1-Formen, exakte 1-Formen). *Sei D offen in $\mathbb{R}^N$. Man nennt $f\colon D \to \mathbb{R}$ eine **Stammfunktion** einer 1-Form $\Lambda\colon D \to \big(\mathbb{R}^N\big)^*$, wenn f auf D total differenzierbar ist und Λ das Differential von f ist, also $\mathrm{d}f = \Lambda$ auf D gilt. Besitzt eine 1-Form auf D eine Stammfunktion, so bezeichnet man sie als **exakt** (auf D).*

Das nächste grundlegende Resultat ermöglicht es, bekannte Stammfunktionen zu Vektorfeldern und 1-Formen zur Berechnung ihrer (orientierten) Kurvenintegrale zu nutzen. Mit den folgenden Sätzen wird klar werden, dass auch darüber hinaus ein enger Zusammenhang zwischen der Stammfunktionsfrage und der orientierten Integration über Kurven besteht.

Satz (HDI für Kurvenintegrale). *Sei D offen in $\mathbb{R}^N$, und sei $f \in \mathrm{C}^1(D)$ die Stammfunktion eines Vektorfelds V beziehungsweise einer 1-Form Λ auf D. Dann gilt*

$$\int_c V(x) \cdot \mathrm{d}x = f(c(b)) - f(c(a)) \qquad \textit{beziehungsweise} \qquad \int_c \Lambda = f(c(b)) - f(c(a))$$

für alle Stückweise-C^1-Kurven $c\colon [a,b] \to D$ in D (d.h. ab jetzt immer $c \in \mathrm{C}^0([a,b], D)$ mit $c \in \mathrm{C}^1((a,b) \setminus S, \mathbb{R}^N)$ und $\sup_{(a,b)\setminus S} |c'| < \infty$ für eine endliche Menge S von „Knickstellen“).

Beweis. Aus den Darstellungsformeln des Abschnitts 1.2, der Kettenregel und dem „normalen“ HDI ergibt sich

$$\int_c V(x) \cdot \mathrm{d}x = \int_a^b V(c(t)) \cdot c'(t)\,\mathrm{d}t = \int_a^b \nabla f(c(t)) \cdot c'(t)\,\mathrm{d}t = \int_a^b (f \circ c)'(t)\,\mathrm{d}t = f(c(b)) - f(c(a))$$

beziehungsweise

$$\int_c \Lambda = \int_a^b \langle \Lambda(c(t)), c'(t) \rangle\,\mathrm{d}t = \int_a^b \langle \mathrm{d}f(c(t)), c'(t) \rangle\,\mathrm{d}t = \int_a^b (f \circ c)'(t)\,\mathrm{d}t = f(c(b)) - f(c(a))$$

(wobei die Integrale trotz Definitionslücken von c' in eventuellen Knickstellen Sinn machen). $\square$

Bemerkung (zum HDI für Kurvenintegrale)**.** Der HDI für Kurvenintegrale bleibt allgemeiner für die Integration über beliebige rektifizierbare Kurven $c\colon [a,b] \to D$ in D richtig. Dies sieht man durch gleichmäßige Approximation solcher allgemeinen Kurven c mit Streckenzügen ein:

Beweis zur Bemerkung. Es reicht zu zeigen, dass zu V wie im Satz, nur rektifzierbarem $c\colon [a,b] \to D$ und beliebigem $\varepsilon > 0$ eine Stückweise-C^1-Kurve $s\colon [a,b] \to D$ mit unverändertem Ausgangs- und Endpunkt $s(a) = c(a)$, $s(b) = c(b)$ und $\left|\int_s V(x)\boldsymbol{\cdot}\mathrm{d}x - \int_c V(x)\boldsymbol{\cdot}\mathrm{d}x\right| < 2\varepsilon \mathrm{L}(c)$ existiert. Ist dies gezeigt, so ergibt sich zusammen mit $\int_s V(x)\boldsymbol{\cdot}\mathrm{d}x = f(s(b)) - f(s(a)) = f(c(b)) - f(c(a))$ (nach bereits gezeigter Version des Satzes für s) auch $\left|f(c(b)) - f(c(a)) - \int_c V(x)\boldsymbol{\cdot}\mathrm{d}x\right| < 2\varepsilon\mathrm{L}(c)$, und mit der Beliebigkeit von ε und $\mathrm{L}(c) < \infty$ folgt die Behauptung $\int_c V(x)\boldsymbol{\cdot}\mathrm{d}x = f(c(b)) - f(c(a))$ auch für nur rektifizierbares c.

Zur Konstruktion von s fixiert man eine beschränkte konvexe Menge $K \subset \mathbb{R}^N$ mit $c([a,b]) \subset K$ und findet wegen der gleichmäßigen Stetigkeit von V auf K ein $\gamma > 0$, so dass für $x, \widetilde{x} \in K$ mit $|\widetilde{x} - x| < \gamma$ stets $|V(\widetilde{x}) - V(x)| < \varepsilon$ gilt. Weiterhin gibt es wegen der gleichmäßigen Stetigkeit von c auch ein $\delta > 0$, so dass für $t, \widetilde{t} \in [a,b]$ mit $|\widetilde{t} - t| < \delta$ stets $|c(\widetilde{t}) - c(t)| < \gamma$ eintritt. Schließlich sei eine beliebige Zerlegung $a = x_0 < x_1 < \ldots < x_\ell = b$ von $[a,b]$ der Feinheit $< \delta$ fixiert, und s sei eine Stückweise-C^1-Parametrisierung des Streckenzugs von x_0 über die Stützstellen $x_1, x_2, \ldots, x_{\ell-1}$ nach x_ℓ mit $s(x_j) = c(x_j)$ für alle $j \in \{0, 1, 2, \ldots, \ell{-}1, \ell\}$. Sei $a = t_0 < t_1 < \ldots < t_m = b$ mit $\{x_0, x_1, \ldots, x_\ell\} \subset \{t_0, t_1, \ldots, t_m\}$ eine beliebige Verfeinerung der fixierten Zerlegung, und für $i \in \{1, 2, \ldots, m\}$ bezeichne $j_i \in \{1, 2, \ldots, \ell\}$ den Index mit $t_i \in (x_{j_i - 1}, x_{j_i}]$. Dann folgt per Teleskopsumme $\sum_{j_i = j}(s(t_i) - s(t_{i-1})) = s(x_j) - s(x_{j-1}) = c(x_j) - c(x_{j-1}) = \sum_{j_i = j}(c(t_i) - c(t_{i-1}))$ (wobei $\sum_{j_i = j}$ für die Summe über alle Indizes i mit $j_i = j$ steht), und insgesamt ergibt sich

$$\sum_{i=1}^m V(c(x_{j_i}))\boldsymbol{\cdot}(s(t_i) - s(t_{i-1})) = \sum_{i=1}^m V(c(x_{j_i}))\boldsymbol{\cdot}(c(t_i) - c(t_{i-1}))\,.$$

Jetzt werden für beliebige Zwischenstellen $\tau_i \in (t_{i-1}, t_i) \subset (x_{j_i - 1}, x_{j_i})$ die getroffenen Wahlen von δ und γ ausgenutzt. Einerseits ergibt sich aus $|\tau_i - x_{j_i}| < \delta$ erst $|c(\tau_i) - c(x_{j_i})| < \gamma$, dann $|V(c(\tau_i)) - V(c(x_{j_i}))| < \varepsilon$ gilt. Andererseits folgen aus $|x_{j_i} - x_{j_i - 1}| < \delta$ erst $|c(x_{j_i}) - c(x_{j_i - 1})| < \gamma$ und dann, da $s(\tau_i)$ auf der in K enthaltenen Strecke von $c(x_{j_i - 1})$ nach $c(x_{j_i})$ liegt, auch $|s(\tau_i) - c(x_{j_i})| < \gamma$ und $|V(s(\tau_i)) - V(c(x_{j_i}))| < \varepsilon$. Anwendung der so erhaltenen ε-Abschätzungen auf den beiden Seiten der zuvor hergeleiteten Gleichheit ergibt

$$\left|\sum_{i=1}^m V(s(\tau_i))\boldsymbol{\cdot}(s(t_i) - s(t_{i-1})) - \sum_{i=1}^m V(c(\tau_i))\boldsymbol{\cdot}(c(t_i) - c(t_{i-1}))\right| \le \varepsilon\left(\sum_{i=1}^m |s(t_i) - s(t_{i-1})| + \sum_{i=1}^m |c(t_i) - c(t_{i-1})|\right) \le 2\varepsilon\mathrm{L}(c)\,.$$

Schließlich existieren die orientierten Kurvenintegrale $\int_s V(x)\boldsymbol{\cdot}\mathrm{d}x$ und $\int_c V(x)\boldsymbol{\cdot}\mathrm{d}x$, letzteres gemäß dem abschließenden Existenzsatz des Abschnitts 1.2, und sind die Limites der gerade abgeschätzten Zwischensummen bei gegen Null strebender Zerlegungsfeinheit. Da die vorausgehenden Überlegungen Zerlegungen beliebig kleiner Feinheit einschließen, ergibt sich im Limes die noch benötigte Hilfsbehauptung

$$\left|\int_s V(x)\boldsymbol{\cdot}\mathrm{d}x - \int_c V(x)\boldsymbol{\cdot}\mathrm{d}x\right| \le 2\varepsilon\mathrm{L}(c)\,. \qquad \square$$

Satz (**Stammfunktionen und Wegunabhängigkeit von Kurvenintegralen**)**.** *Sei D offen in $\mathbb{R}^N$, und sei $V \in \mathrm{C}^0(D, \mathbb{R}^N)$ ein stetiges Vektorfeld beziehungsweise $\Lambda \in \mathrm{C}^0\big(D, \big(\mathbb{R}^N\big)^*\big)$ eine stetige 1-Form. Dann sind folgende Aussagen* ***äquivalent***:

(I) *V beziehungsweise Λ* ***besitzt*** *eine* ***Stammfunktion*** *auf D.*

(II) ***Kurvenintegrale*** *von V beziehungsweise Λ sind* ***wegunabhängig****, d.h. es gilt*

$$\int_c V(x)\boldsymbol{\cdot}\mathrm{d}x = \int_{\widetilde{c}} V(x)\boldsymbol{\cdot}\mathrm{d}x \qquad \text{beziehungsweise} \qquad \int_c \Lambda = \int_{\widetilde{c}} \Lambda$$

für alle Stückweise-C^1-Kurven $c\colon [a,b] \to D$ und $\widetilde{c}\colon \big[\widetilde{a}, \widetilde{b}\big] \to D$ in D ***mit gleichen Endpunkten*** *$\widetilde{c}(\widetilde{a}) = c(a)$, $\widetilde{c}(\widetilde{b}) = c(b)$.*

(III) ***Kurvenintegrale*** *von V beziehungsweise Λ* ***über*** **geschlossene** ***Kurven verschwinden****, d.h. es gilt*

$$\int_\gamma V(x)\boldsymbol{\cdot}\mathrm{d}x = 0 \qquad \text{beziehungsweise} \qquad \int_\gamma \Lambda = 0$$

für alle Stückweise-C^1-Kurven $\gamma\colon [a,b] \to D$ in D mit $\gamma(b) = \gamma(a)$.

Sind die Aussagen (I), (II), (III) erfüllt und ist D zusammenhängend, so kann man die bis auf eine additive Konstante eindeutige Stammfunktion f zu V beziehungsweise Λ als ***unbestimmtes Wegintegral*** *berechnen, d.h. für $x \in D$ ist*

$$f(x) = \int_{c_x} V(y) \cdot \mathrm{d}y + \mathrm{const} \qquad \textit{beziehungsweise} \qquad f(x) = \int_{c_x} \Lambda + \mathrm{const}\,,$$

wobei $x_0 \in D$ ein beliebig fixierter Punkt in D und c_x eine beliebige Stückweise-C^1-Kurve von x_0 zu x in D ist (und eine solche Kurve c_x existiert auch, weil offene zusammenhängende Teilmengen von $\mathbb{R}^N$ auch wegzusammenhängend und Streckenzug-zusammenhängend sind). Ist D nicht zusammenhängend, so hat die Stammfunktion zumindest auf jeder Zusammenhangskomponente von D die vorausgehende Form des unbestimmten Wegintegrals.

Vektorfelder mit einer der äquivalenten Eigenschaften des Satzes sind auch als **konservative Vektorfelder** bekannt, vor allem denkt man bei der Prägung dieses Begriffs aber an die Eigenschaft (III).

Beweis. Die Implikation (I) $\Longrightarrow$ (III) sowie die Aussagen über die Eindeutigkeit und Form der Stammfunktion ergeben sich direkt aus dem vorigen Satz.

Um die Implikation (III) $\Longrightarrow$ (II) einzusehen, bildet man aus den Wegen c und $\widetilde{c}$ mit gleichen Endpunkten den zusammengesetzten, geschlossenen Weg $\gamma\colon [a{-}b, \widetilde{b}{-}\widetilde{a}] \to D$ mit $\gamma(t) := c(t{+}b)$ für $t \in [a{-}b, 0]$ und $\gamma(t) := \widetilde{c}(\widetilde{b}{-}t)$ für $t \in [0, \widetilde{b}{-}\widetilde{a}]$. Mit der Zerlegungsadditivität und der Parametrisierungsinvarianz orientierter Kurvenintegrale folgt (weil $\gamma\big|_{[a-b,0]}$ Orientierung-erhaltende Umparametrisierung von c und $\gamma\big|_{[0,\widetilde{b}-\widetilde{a}]}$ Orientierung-*umkehrende* Umparametrisierung von $\widetilde{c}$ ist)

$$0 = \int_\gamma V(x) \cdot \mathrm{d}x = \int_{a-b}^0 (V \circ \gamma)\,\mathrm{d}\gamma + \int_0^{\widetilde{b}-\widetilde{a}} (V \circ \gamma)\,\mathrm{d}\gamma = \int_c V(x) \cdot \mathrm{d}x - \int_{\widetilde{c}} V(x) \cdot \mathrm{d}x$$

beziehungsweise Analoges für die 1-Form Λ. Damit ist klar, dass die Kurvenintegrale über c und $\widetilde{c}$ übereinstimmen müssen.

Es bleibt zu zeigen, dass (I) aus (II) folgt. Der Einfachheit halber wird dies nur für den Fall von 1-Formen durchgeführt, mit Vektorfeldern ließe sich aber natürlich auch argumentieren. O.E. sei D als zusammenhängend angenommen, $x_0 \in D$ sei fixiert, und f sei wie im Satz durch das unbestimmte Wegintegral definiert. Aufgrund der Voraussetzung (II) hängen die Integrale dabei nicht von den gewählten Kurven c_x ab, so dass ein solches f wohldefiniert ist. Bei festem $a \in D$ liegt nun für $x \in D$ mit $|x{-}a| \ll 1$ die gerade Verbindungsstrecke von a zu x in D und wird durch $s\colon [0,1] \to D\,,\ t \mapsto a{+}t(x{-}a)$ parametrisiert. Wählt man c_x als den Weg, der sich durch Durchlaufen von erst c_a und dann s ergibt, so folgt mit der Definition von f, der Darstellungsformel des Abschnitts 1.2 und der fundamentalen Abschätzung für Integrale

$$\begin{aligned} |f(x){-}f(a){-}\langle\Lambda(a), x{-}a\rangle| &= \left| \int_s \Lambda - \langle\Lambda(a), x{-}a\rangle \right| \\ &= \left| \int_0^1 \big[\langle\Lambda(a{+}t(x{-}a)), x{-}a\rangle - \langle\Lambda(a), x{-}a\rangle\big]\,\mathrm{d}t \right| \\ &\le \Big(\sup_{\mathrm{Bild}(s)} \|\Lambda{-}\Lambda(a)\|_{(\mathbb{R}^N)^*} \Big) |x{-}a|\,. \end{aligned}$$

Da Bild(s) in der abgeschlossenen Kugel mit Radius $|x-a|$ um a enthalten und Λ stetig ist, ergibt sich aus dieser Abschätzung

$$\lim_{x\to a} \frac{|f(x)-f(a)-\langle \Lambda(a), x-a\rangle|}{|x-a|} = 0\,,$$

d.h. f ist an der Stelle a total differenzierbar mit $f'(a) = \Lambda(a)$. Insgesamt ist f also auf ganz D total differenzierbar mit $\mathrm{d}f = \Lambda$, und Λ besitzt f als Stammfunktion. □

Ungünstigerweise lassen sich alle drei äquivalenten Bedingungen des letzten Satzes nicht ohne Weiteres prüfen, die Frage, wie man *konkret* über die Existenz von Stammfunktionen entscheidet, ist also im Moment noch offen. Eine weitreichende Antwort wird jedoch im nächsten grundlegenden Satz folgen, sie benötigt aber noch eine Bildung mit Ableitungen:

Definition (**Rotationsmatrizen, äußere Ableitungen, geschlossene 1-Formen**). *Sei x ein innerer Punkt von $D \subset \mathbb{R}^N$.*

(I) *Existiert für ein Vektorfeld $V\colon D \to \mathbb{R}^n$ die Jacobi-Matrix $\mathrm{D}V(x) \in \mathbb{R}^{N\times N}$ in x, so definiert man die* ***Rotationsmatrix*** $\operatorname{Rot} V(x) \in \mathbb{R}^{N\times N}_{\mathrm{asym}}$ *von V in x als den doppelten antisymmetrischen Anteil von $\mathrm{D}V(x)$, d.h.*

$$\operatorname{Rot} V(x) := \mathrm{D}V(x) - \mathrm{D}V(x)^{\mathrm{T}}\,.$$

(II) *Existiert für eine 1-Form $\Lambda\colon D \to \left(\mathbb{R}^N\right)^*$ die totale Ableitung $\Lambda'(x) \in \mathcal{L}(\mathbb{R}^N, (\mathbb{R}^N)^*)$ in x, so definiert man die* ***äußere Ableitung*** $\mathrm{d}\Lambda(x) \in \mathcal{L}^2_{\mathrm{asym}}(\mathbb{R}^N, \mathbb{R})$ *von Λ in x als den doppelten antisymmetrischen Anteil von Λ', aufgefasst als Element des Raums $\mathcal{L}^2_{\mathrm{asym}}(\mathbb{R}^N, \mathbb{R})$ der (stetigen) antisymmetrischen $\mathbb{R}$-Bilinearformen $\mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$, d.h.*

$$\langle \mathrm{d}\Lambda(x); v, w\rangle := \langle \Lambda'(x)v, w\rangle - \langle \Lambda'(x)w, v\rangle = \langle \partial_w \Lambda(x), v\rangle - \langle \partial_v \Lambda(x), w\rangle \qquad \text{für } v, w \in \mathbb{R}^N\,.$$

Gilt $\mathrm{d}\Lambda \equiv 0$ auf ganz D, so bezeichnet man die 1-Form Λ als ***geschlossen****.*

Definition (**Sternförmigkeit**). *Sei A eine Teilmenge eines $\mathbb{K}$-Vektorraums. Man nennt $x_0 \in A$ einen* ***Sternpunkt*** *für A, wenn für jedes $x \in A$ die Strecke $\{x_0 + \lambda(x-x_0) : \lambda \in [0,1]\}$ von x_0 nach x in A enthalten ist. Die Menge A heißt* ***sternförmig****, wenn es einen Sternpunkt für A gibt.*

Nach diesen Vorbereitungen kann der bereits angekündigte, wichtigste Satz dieses Vorlesungskapitels formuliert werden.

Hauptsatz (Kriterien für die **Existenz von Stammfunktionen**). *Sei D offen in $\mathbb{R}^N$, und sei $V \in \mathrm{C}^1(D, \mathbb{R}^N)$ ein stetig differenzierbares Vektorfeld beziehungsweise $\Lambda \in \mathrm{C}^1\big(D, \left(\mathbb{R}^N\right)^*\big)$ eine stetige differenzierbare 1-Form auf D.*

(I) ***Notwendiges Kriterium:*** *Besitzt V beziehungsweise Λ eine Stammfunktion auf D, so gilt $\operatorname{Rot} V \equiv 0$ beziehungsweise $\mathrm{d}\Lambda \equiv 0$ auf D.*

(II) ***Hinreichendes Kriterium/Poincaré-Lemma:*** *Ist D sternförmig und gilt $\operatorname{Rot} V \equiv 0$ beziehungsweise $\mathrm{d}\Lambda \equiv 0$ auf D, so besitzt V beziehungsweise Λ eine Stammfunktion auf D.*

Bemerkung (zur den Kriterien für die Existenz von Stammfunktionen)**. Bei sternförmigem $\boldsymbol{D}$** besagen beide Teile des Hauptsatzes zusammen, dass die sogenannten **Integrabilitätsbedingungen**

$$\boxed{\operatorname{Rot} V \equiv 0 \text{ auf } D} \qquad \text{beziehungsweise} \qquad \boxed{\mathrm{d}\Lambda \equiv 0 \text{ auf } D}$$

ein zugleich **notwendiges und hinreichendes Kriterium für** die **Existenz einer Stammfunktion** zu $V \in \mathrm{C}^1(D, \mathbb{R}^N)$ beziehungsweise $\Lambda \in \mathrm{C}^1\big(D, (\mathbb{R}^N)^*\big)$ sind. Komponentenweise ausgeschrieben lauten die Integrabilitätsbedingungen $\partial_j V_i = \partial_i V_j$ auf D beziehungsweise $\langle \partial_j \Lambda, \mathrm{e}_i \rangle = \langle \partial_i \Lambda, \mathrm{e}_j \rangle$ auf D für alle $i, j \in \{1, 2, \ldots, N\}$ (wobei $\mathrm{e}_1, \mathrm{e}_2, \ldots, \mathrm{e}_N$ die kanonische Basis des $\mathbb{R}^N$ bezeichnet), und mit anderen Worten gelten die Äquivalenzen

$$\begin{aligned} V \text{ Gradientenfeld auf } D &\iff V \text{ rotationsfrei auf } D\,, \\ \Lambda \text{ exakt auf } D &\iff \Lambda \text{ geschlossen auf } D\,. \end{aligned}$$

Tatsächlich ist Teil (I) des Hauptsatzes im Wesentlichen bereits aus der Analysis bekannt und wird nach kurzem Sortieren der Begriffe offensichtlich:

Beweis zu Teil (I) *des Hauptsatzes.* Im Fall $V = \nabla f$, dass V eine Stammfunktion f besitzt, ist die Hesse-Matrix $\mathrm{D}V = \mathrm{D}(\nabla f) = \mathrm{H}f$ von f bekanntlich symmetrisch, und ihr antisymmetrischer Anteil $\operatorname{Rot} V = \mathrm{D}V - \mathrm{D}V^{\mathrm{T}}$ verschwindet auf D.

Im Fall $\Lambda = \mathrm{d}f$, dass Λ eine Stammfunktion f besitzt, gilt $\langle \Lambda'(x)w, v \rangle = \langle \partial_w \mathrm{d}f(x), v \rangle = \big[\partial_w (f')(x)\big] v = f''(x)(v, w)$ für $x \in D$, $v, w \in \mathbb{R}^N$, wobei f'' die totale Ableitung zweiter Ordnung von f ist. Aus der bekannten Symmetrie der Bilinearform $f''(x)$ folgt dann $\mathrm{d}\Lambda \equiv 0$ auf D. □

Um Teil (I) des Hauptsatzes herzuleiten, setzt man — was sich nach dem vorigen Satz aufdrängt — die gesuchte Stammfunktion als unbestimmtes Wegintegral an und verifiziert, dass sie tatsächlich die gewünschte Ableitung hat:

Beweis zu Teil (II) *des Hauptsatzes.* Nach Translation lässt sich o.E. annehmen, dass $0 \in D$ ein Sternpunkt für D ist. Man definiert $f \colon D \to \mathbb{R}$ durch

$$f(x) := \int_0^1 V(tx) \cdot x \, \mathrm{d}t \qquad \text{beziehungsweise} \qquad f(x) := \int_0^1 \langle \Lambda(tx), x \rangle \, \mathrm{d}t$$

für $x \in D$. Es folgt eine Rechnung mit der Definition der Richtungsableitung, Differentiation unter dem Integral (hier erlaubt, da $\mathrm{D}V$ beziehungsweise Λ' auf D beschränkt ist), den Produkt- und Kettenregeln und dem HDI. Außerdem gehen, jeweils beim Übergang von der ersten zur zweiten Zeile der Rechnung, die Integrabilitätsbedingungen ein, gemäß denen die antisymmetrischen Anteile verschwinden und $\mathrm{D}V$ beziehungsweise Λ' symmetrisch ist. Für $x \in D$, $w \in \mathbb{R}^N$ ergibt sich tatsächlich

$$\begin{aligned} \partial_w f(x) &= \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0} \int_0^1 V(t(x+sw)) \cdot (x+sw) \, \mathrm{d}t = \int_0^1 \big[\mathrm{D}V(tx)tw \cdot x + V(tx) \cdot w\big] \, \mathrm{d}t \\ &= \int_0^1 \big[\mathrm{D}V(tx)x \cdot tw + V(tx) \cdot w\big] \, \mathrm{d}t = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}\big(V(tx) \cdot tw\big) \, \mathrm{d}t = V(x) \cdot w \end{aligned}$$

beziehungsweise äquivalent

$$\begin{aligned} \partial_w f(x) &= \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0} \int_0^1 \langle \Lambda(t(x+sw)), (x+sw) \rangle \, \mathrm{d}t = \int_0^1 \big[\langle \Lambda'(tx)tw, x \rangle + \langle \Lambda(tx), w \rangle\big] \, \mathrm{d}t \\ &= \int_0^1 \big[\langle \Lambda'(tx)x, tw \rangle + \langle \Lambda(tx), w \rangle\big] \, \mathrm{d}t = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t} \langle \Lambda(tx), tw \rangle \, \mathrm{d}t = \langle \Lambda(x), w \rangle\,. \end{aligned}$$

Da die erhaltene Ableitung stetig von x abhängt, ist hiermit totale Differenzierbarkeit von f mit $\nabla f = V$ beziehungsweise $\mathrm{d}f = \Lambda$ auf D gezeigt und f als die gesuchte Stammfunktion identifiziert. □

Bemerkung und Beispiel (**Ableitung der Argumentfunktion**)**.** Die **Voraussetzung der Sternförmigkeit** von D in Teil (II) des Hauptsatzes (dem Poincaré-Lemma) **kann** abgeschwächt, aber **keineswegs** ganz **fallen gelassen werden**. Ein typisches und wichtiges Beispiel, an dem dies deutlich wird, ist die sogenannte **Ableitung der Argumentfunktion**

$$\Theta\colon \mathbb{R}^2 \setminus \{0\} \to \mathbb{R}^2\,,\ (x,y) \mapsto \frac{(-y,x)}{x^2+y^2}\,.$$

Man rechnet nach, dass das Vektorfeld Θ den Integrabilitätsbedingungen $\operatorname{Rot}\Theta \equiv 0$ auf ganz $\mathbb{R}^2 \setminus \{0\}$ genügt, kann sich aber auch überlegen, dass es keine Stammfunktion zu Θ auf ganz $\mathbb{R}^2 \setminus \{0\}$ gibt. Tatsächlich ist jede sogenannte Argumentfunktion auf offenem $D \subset \mathbb{R}^2 \setminus \{0\}$, d.h. jede stetige Auswahl eines Polarwinkels zu den Punkten in D, eine Stammfunktion zu Θ auf D. Auf ganz $\mathbb{R}^2 \setminus \{0\}$ (und manchen anderen nicht sternförmigen Gebieten) existiert aber keine Stammfunktion. Die entscheidende Problematik erkennt man anhand des sogenannten Hauptzweigs der Argumentfunktion

$$\operatorname{Arg}\colon \mathbb{R}^2 \setminus (\mathbb{R}_{\le 0} \times \{0\}) \to (-\pi,\pi)\,,\ (r\cos\varphi, r\sin\varphi) \mapsto \varphi\,,$$

denn Arg nimmt knapp oberhalb der negativen x-Achse Werte knapp kleiner als π und knapp unterhalb der negativen x-Achse Werte knapp größer als $-\pi$ an und kann deshalb nicht stetig auf ganz $\mathbb{R}^2 \setminus \{0\}$ fortgesetzt werden. In ganz ähnlicher Weise scheitert die Fortsetzung auch bei anderer Auswahl der Polarwinkel. Daher existieren auf ganz $\mathbb{R}^2 \setminus \{0\}$ wie bereits erwähnt keine Argumentfunktion und keine Stammfunktion zu Θ. Etwas mehr zu diesem Themenkreis und zur Klasse der offenen Mengen D, auf denen die Existenz einer Argumentfunktion und generell von Stammfunktionen gesichert ist, folgt unten und in den Übungen.

Nachdem die Stammfunktionsfrage auf Seite der Theorie weitgehend beantwortet ist, wird nun ein naheliegendes Verfahren zur konkreten Berechnung von Stammfunktionen beschrieben:

Verfahren (**Berechnung von Stammfunktionen** zu Vektorfeldern und 1-Formen)**.** *Um über die Existenz einer Stammfunktion zu einem gegebenen* C^1*-Vektorfeld* V *beziehungsweise zu einer gegebenen* C^1*-1-Form* Λ *auf* D *zu entscheiden und diese Stammfunktion gegebenenfalls zu berechnen, …*

(1) ***prüft*** *man* ***als Erstes immer****, ob die* ***Integrabilitätsbedingungen*** $\operatorname{Rot} V \equiv 0$ *beziehungsweise* $\mathrm{d}\Lambda \equiv 0$ *auf* D *erfüllt sind. Sind diese Bedingungen nämlich verletzt, so gibt es keine Stammfunktion und man kann sich jede weitere Rechnung sparen. Sind sie erfüllt, so …*

(2) ***berechnet man****, soweit nötig,* ***Stammfunktionen*** f_i *der Komponentenfunktionen* V_i *beziehungsweise* Λ_i ***bezüglich der*** *zugehörigen* ***Einzelvariablen*** x_i

$$\int V_1(x)\,\mathrm{d}x_1 \ \textit{bzw.}\ \int \langle\Lambda(x),\mathrm{e}_1\rangle\,\mathrm{d}x_1 =: f_1(x) + C_1(x_2,x_3,x_4\ldots,x_N)\,,$$
$$\int V_2(x)\,\mathrm{d}x_2 \ \textit{bzw.}\ \int \langle\Lambda(x),\mathrm{e}_2\rangle\,\mathrm{d}x_2 =: f_2(x) + C_2(x_1,x_3,x_4\ldots,x_N)\,,$$
$$\vdots$$
$$\int V_N(x)\,\mathrm{d}x_N \ \textit{bzw.}\ \int \langle\Lambda(x),\mathrm{e}_N\rangle\,\mathrm{d}x_N =: f_N(x) + C_N(x_1,x_2,,\ldots,x_{N-1})\,,$$

wobei die Integrationskonstante bei der x_i-Integration von allen anderen Variablen, aber eben nicht *von x_i abhängt. Es gilt nun, die C_i mit diesen Abhängigkeiten so zu bestimmen, dass alle Ergebnisse $f_1(x)+C_1(x_2, x_3, x_4 \ldots, x_N) = f_2(x)+C_2(x_1, x_3, x_4 \ldots, x_N) = \ldots = f_N(x)+C_N(x_1, x_2, , \ldots, x_{N-1})$ übereinstimmen (was bei erfüllten Integrabilitätsbedingungen möglich ist).*

(3) *Das* ***gemeinsame Ergebnis*** *der Integrationen* ***liefert*** *dann* ***eine Stammfunktion*** *zu V beziehungsweise Λ — und zwar global auf einem sternförmigem Definitionsbereich D, auf einem nicht sternförmigem D eventuell nur lokal.*

Zum Abschluss dieses Kapitels wird noch aufgezeigt, dass der Hauptsatz auch auf allgemeineren Definitionsbereichen D gilt. Dies erfordert allerdings etwas begrifflichen Aufwand.

Ausblick (**Stammfunktionen auf allgemeineren** als sternförmigen **Definitionsbereichen**). Die Voraussetzung der Sternförmigkeit in Teil (II) des vorausgehenden Hauptsatzes, also im Poincaré-Lemma, ist unnötig stark und kann zu einer schwächeren, in vielerlei Hinsicht optimalen, rein topologischen Bedingung abgeschwächt werden: Tatsächlich sind die **Integrabilitätsbedingungen** auch dann **hinreichend** für die Existenz einer Stammfunktion, wenn der offene **Definitionsbereich** $D \subset \mathbb{R}^N$ **nur einfach zusammenhängend** ist, das bedeutet per Definition, dass **jeder geschlossene Weg** $c\colon [a, b] \to D$ (durch Schleifen in D) **homotop zu einem konstanten Weg** ist, und dies wiederum ist so definiert, dass es zu c eine **Homotopie** genannte stetige Abbildung $H\colon [0, 1] \times [a, b] \to D$ mit $H(0, \cdot\,) = c$ auf $[a, b]$, mit $H(1, \cdot\,) \equiv \text{const}$ auf $[a, b]$ und mit $H(\,\cdot\,, b) = H(\,\cdot\,, a)$ auf $[0, 1]$ gibt. **Anschaulich** kann man sich diesen Begriff so vorstellen, **dass jede geschlossene Kurve** $c = H(0, \cdot\,)$ durch die geschlossenen Kurven $H(\lambda, \cdot\,)$ mit $\lambda \in (0, 1)$ **zu einer konstanten Kurve** $H(1, \cdot\,) = \text{const}$ stetig deformiert werden kann, ohne dabei D zu verlassen. Mit noch etwas anderen Worten kann jede „Schlinge" c in D zusammengezogen werden. Einfacher Zusammenhang ist weder stärker noch schwächer als Zusammenhang, sondern vielmehr ein davon unabhängiger Begriff: Teilmengen von $\mathbb{R}$ sind stets einfach zusammenhängend; der Begriff trivialisiert sich also in diesem Fall. Für Mengen D in der Ebene $\mathbb{R}^2$ bedeutet einfacher Zusammenhang, dass D keine „Löcher" aufweist, die in D umlaufen werden können, im Anschauungsraum $\mathbb{R}^3$ gestaltet sich das Konzept aber schon subtiler, denn $\mathbb{R}^3$ ohne einen Punkt ist einfach zusammenhängend, $\mathbb{R}^3$ ohne eine Gerade dagegen nicht.

Beweis des Poincaré-Lemmas auf nur einfach zusammenhängendem, offenen $D \subset \mathbb{R}^N$. Sei $V \in \mathrm{C}^1(D, \mathbb{R}^N)$ ein rotationsfreies Vektorfeld beziehungsweise $\Lambda \in \mathrm{C}^1\big(D, (\mathbb{R}^N)^*\big)$ eine geschlossene 1-Form auf D. Nach dem Satz über Stammfunktionen und Wegunabhängigkeit genügt es zu zeigen, dass das orientierte Kurvenintegral von V beziehungsweise Λ über alle geschlossenen Stückweise-C^1-Kurven $c\colon [a, b] \to D$ verschwindet. Sei dazu $H\colon [0, 1] \times [a, b] \to D$ eine Homotopie, die eine solche Kurve $c = H(0, \cdot\,)$ mit einer konstanten Kurve $H(1, \cdot\,)$ bei $H(\,\cdot\,, b) = H(\,\cdot\,, a)$ verbindet. Bei ausreichender Differenzierbarkeit der Homotopie H, zum Beispiel im C^2-Fall, ließe sich nun direkt eine Rechnung ähnlich der unten Folgenden ansetzen. Da H aber im Allgemeinen nur stetig ist, wird eine Approximation von H nötig, die hier durch Streckenzüge implementiert wird. Weil das Bild von H kompakt und D offen ist, folgt als Erstes $\delta := \operatorname{dist}(\operatorname{Bild}(H), \mathbb{R}^N \setminus D) > 0$. Wegen gleichmäßiger Stetigkeit von H gibt es ein $K \in \mathbb{N}$, so dass für $\lambda, \widetilde{\lambda} \in [0, 1]$ und $t, \widetilde{t} \in [a, b]$ mit $|\widetilde{\lambda}-\lambda|+|\widetilde{t}-t| < \frac{1}{K}$ stets $|H(\widetilde{\lambda}, \widetilde{t})-H(\lambda, t)| < \frac{\delta}{4}$ eintritt. Insbesondere kann für jedes $i \in \{1, 2, \ldots, K-1\}$ die geschlossene Kurve $H\big(\frac{i}{K}, \cdot\,\big)$ durch einen über $[a, b]$ parametrisierten geschlossenen Streckenzug $c_{i/K}$ approximiert werden, so dass $\sup_{[a,b]} \big|c_{i/K}-H\big(\frac{i}{K}, \cdot\,\big)\big| \le \frac{\delta}{4}$ gilt (dazu $[a, b]$ in Intervalle der Länge $<\frac{1}{K}$ zerlegen und die Werte von $H\big(\frac{i}{K}, \cdot\,\big)$ in Zerlegungsstellen durch Strecken verbinden). Vereinbart man noch $c_0 := c$ und $c_1 := H(1, \cdot\,)$, so kann man für $i \in \{1, 2, \ldots, K-1, K\}$ und $\lambda \in \big[\frac{i-1}{K}, \frac{i}{K}\big]$ durch $c_\lambda := c_{(i-1)/K}+(\lambda K-i+1)(c_{i/K}-c_{(i-1)/K})$ eine geschlossene Stückweise-C^1-Kurve $c_\lambda\colon [a, b] \to \mathbb{R}^N$ mit $\sup_{[a,b]} \big|c_\lambda-H\big(\lambda, \cdot\,\big)\big| \le \frac{3}{4}\delta$ erhalten. Insgesamt ist c_λ damit für jedes $\lambda \in [0, 1]$ definiert, verhält sich konsistent zu den früheren $c_{i/K}$ und *verläuft ganz in D*. Für $i \in \{1, 2, \ldots, K-1, K\}$ und $\lambda \in \big(\frac{i-1}{K}, \frac{i}{K}\big)$ ergibt sich durch Rechnung mit Ketten-/Produktregeln, den Integrabilitätsbedingungen und dem HDI jetzt

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\lambda}\int_{c_\lambda} V(x)\cdot \mathrm{d}x &= \frac{\mathrm{d}}{\mathrm{d}\lambda}\int_a^b \big[(V\circ c_\lambda)\cdot c_\lambda'\big] \\
&= K\int_a^b \big[(\mathrm{D}V\circ c_\lambda)(c_{i/K}-c_{(i-1)/K})\cdot c_\lambda' + (V\circ c_\lambda)\cdot (c_{i/K}'-c_{(i-1)/K}')\big] \\
&= K\int_a^b \big[(\mathrm{D}V\circ c_\lambda)c_\lambda'\cdot (c_{i/K}-c_{(i-1)/K}) + (V\circ c_\lambda)\cdot (c_{i/K}'-c_{(i-1)/K}')\big] \\
&= K\int_a^b \big[(V\circ c_\lambda)\cdot (c_{i/K}-c_{(i-1)/K})\big]' \\
&= KV(c_\lambda(b))\cdot (c_{i/K}(b)-c_{(i-1)/K}(b)) - KV(c_\lambda(a))\cdot (c_{i/K}(a)-c_{(i-1)/K}(a)) = 0
\end{aligned}
$$

beziehungsweise analog

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\int_{c_\lambda}\Lambda = 0\,.$$

Da das Parameter-abhängige Integral gemäß einem Satz der Analysis stetig vom Parameter $\lambda\in[0,1]$ abhängt, hat somit $\int_{c_\lambda} V(x)\cdot \mathrm{d}x$ beziehungsweise $\int_{c_\lambda}\Lambda$ für alle $\lambda\in[0,1]$ denselben Wert; das kann man mit dem Konstanzsatz aus dem Verschwinden der Ableitung nach $\lambda\in[0,1]\setminus\big\{0,\frac{1}{K},\frac{2}{K},\ldots,\frac{K-1}{K},1\big\}$ (trotz eventueller Nicht-Differenzierbarkeit für die endlich vielen Parameterwahlen $\lambda=\frac{i}{K}$) schließen. Insbesondere folgt

$$\int_c V(x)\cdot \mathrm{d}x = \int_{H(1,\,\cdot\,)} V(x)\cdot \mathrm{d}x = 0 \qquad \text{beziehungsweise} \qquad \int_c \Lambda = \int_{H(1,\,\cdot\,)}\Lambda = 0$$

für die beliebig vorgegebene Kurve $c=c_0$, weil $c_1 = H(1,\,\cdot\,)$ konstant ist. Wie eingangs erläutert, ist hiermit nachgewiesen, dass V ein Gradientenfeld beziehungsweise Λ exakt ist. $\square$

Kapitel 2

Allgemeine Maß- und Integrationstheorie

Einleitung: Das Maßproblem

Dieses Kapitel beschäftigt sich, ganz grob gesprochen, mit Konzepten des Messens von Mengen, insbesondere mit Messungen

- des ***k*-dimensionalen Inhalts** von Teilmengen von $\mathbb{R}^N$ mit $0 \le k \le N \in \mathbb{N}$

und spezieller mit Messungen

- der Anzahl der Elemente einer Menge ($k = 0$),
- der Länge einer Kurve ($k = 1$),
- des Inhalts einer Fläche ($k = 2$)
- oder des Volumens eines Raumbereichs ($k = 3$).

Die Messung ordnet Teilmengen von $\mathbb{R}^N$ dabei eine (möglicherweise unendliche) nichtnegative Kennzahl zu. Als mathematisches Modell bieten sich somit Abbildungen

$$\mu \colon \mathcal{P}(\mathbb{R}^N) \to [0, \infty]$$

auf der Potenzmenge $\mathcal{P}(\mathbb{R}^N)$ von $\mathbb{R}^N$, also der Menge aller Teilmengen von $\mathbb{R}^N$, an. Folgende Grundeigenschaften von μ erscheinen plausibel:

- $\mu(\emptyset) = 0$,
- (Mengen-)**Monotonie**: $\mu(A) \le \mu(\widetilde{A})$ für $A, \widetilde{A} \in \mathcal{P}(\mathbb{R}^N)$ mit $A \subset \widetilde{A}$,
- **Additivität**[1]: $\mu(A \cup \widetilde{A}) = \mu(A) + \mu(\widetilde{A})$ für *disjunkte* $A, \widetilde{A} \in \mathcal{P}(\mathbb{R}^N)$.
- Additivität bei abzählbarer[2] Vereinigung, genannt **σ-Additivität**: $\mu(\bigcup_{i=1}^\infty A_i) = \sum_{i=1}^\infty \mu(A_i)$ für *disjunkte*[3] $A_1, A_2, A_3, \ldots \in \mathcal{P}(\mathbb{R}^N)$,

[1] Genauer sollte man von Additivität bei endlicher Vereinigung oder kurz von endlicher Additivität sprechen.

[2] Additivität bei *über*abzählbarer Vereinigung ist für $k > 0$ keine sinnvolle Forderung: Dann soll nämlich einerseits $\mu(\{x\}) = 0$ für alle Einermengen $\{x\}$ mit $x \in \mathbb{R}^N$ gelten. Andererseits soll es aber auch Mengen $A \in \mathcal{P}(\mathbb{R}^N)$ mit $\mu(A) > 0$ geben. Da jedes A disjunkte Vereinigung von Einermengen ist, würde überabzählbare Additivität zu einem Widerspruch führen.

[3] Das Adjektiv ‘disjunkt’ wird hier und im Folgenden stets im Sinn von ‘paarweise disjunkt’ gebraucht.

- **Translationsinvarianz**[4]: $\mu(x+A) = \mu(A)$ für $x \in \mathbb{R}^N$ und $A \in \mathcal{P}(\mathbb{R}^N)$,
- **Rotationsinvarianz**[5]: $\mu(TA) = \mu(A)$ für $T \in \mathcal{O}(\mathbb{R}^N)$ und $A \in \mathcal{P}(\mathbb{R}^N)$.

Dabei sind Monotonie und Additivität als Spezialfälle in σ-Additivität enthalten, und Translations- und Rotationsinvarianz werden zusammenfassend als **Bewegungsinvarianz** bezeichnet.

Leider kann man für $k = N$ und als Folge davon auch für $k \in \{1, 2, \dots, N-1\}$ nicht alle diese Anforderungen an k-dimensionale Inhalte aufrechterhalten. Dies zeigt folgendeer Satz von G. Vitali (∼1905):

Satz (über die **Unlösbarkeit des Maßproblems**). *Es gibt keine σ-additive und translationsinvariante Abbildung $\mu\colon \mathcal{P}(\mathbb{R}^N) \to [0, \infty]$ mit $0 < \mu\big([0, 1]^N\big) < \infty$.*

Beweis. Aus der Annahme, dass es doch ein solches μ gibt, wird ein Widerspruch hergeleitet. Dazu rechnet man zuerst mit σ-Additivität, Translationsinvarianz und Monotonie nach:

$$\mu\big([0, 2)^N\big) = \sum_{x \in \{0,1\}^N} \mu\big(x+[0, 1)^N\big) = \sum_{x \in \{0,1\}^N} \mu\big([0, 1)^N\big) \le 2^N \mu\big([0, 1]^N\big) < \infty\,.$$

Jetzt sei A ein Repräsentantensystem des Faktors $\mathbb{R}^N/\mathbb{Q}^N$ mit $A \subset [0, 1]^N$, d. h. die Teilmenge A von $\mathbb{R}^N$ wird durch Auswahl eines Repräsentanten in $[0, 1]^N$ aus jeder Restklasse $x+\mathbb{Q}^N$ mit $x \in \mathbb{R}^N$ gebildet. Dann ist $(q+A)_{q \in \mathbb{Q}^N}$ eine abzählbare Familie disjunkter Mengen mit

$$[0, 1]^N \subset \bigcup_{q \in \mathbb{Q}^N} (q+A)\,.$$

Unter Verwendung von Monotonie, σ-Additivität und Translationsinvarianz folgt

$$0 < \mu\big([0, 1]^N\big) \le \sum_{q \in \mathbb{Q}^N} \mu(q+A) = \sum_{q \in \mathbb{Q}^N} \mu(A)\,,$$

und deshalb muss $\mu(A) > 0$ gelten. Andererseits ist

$$[0, 2)^N \supset \bigcup_{q \in \mathbb{Q}^N \cap [0,1)^N} (q+A)\,,$$

und man erhält

$$\infty > \mu\big([0, 2)^N\big) \ge \sum_{q \in \mathbb{Q}^N \cap [0,1)^N} \mu(A)\,.$$

Da auf der rechten Seite der letzten Formel unendlich viele gleiche Summanden stehen, folgt $\mu(A) = 0$, und der gewünschte Widerspruch ist erreicht. □

Die Menge A im vorigen Beweis wird mit Hilfe des Auswahlaxioms gebildet, und tatsächlich sind auch andere Mengen, die zu derartigen Widersprüchen führen, ähnliche pathologische Konstruktionen und können nicht explizit angegeben. Ein weiteres Resultat in dieser Richtung geht auf S. Banach und A. Tarski (∼1924) zurück, wird hier *ohne Beweis* angegeben und verschärft die Situation, insofern, dass in Dimensionen 3 oder höher schon endliche Additivität, nicht erst σ-Additivität ein Problem darstellt:

[4] Dabei ist $x+A := \{x+a : a \in A\}$.

[5] $\mathcal{O}(\mathbb{R}^N)$ bezeichnet die Gruppe der orthogonalen $(N\times N)$-Matrizen beziehungsweise der linearen Isometrien von $\mathbb{R}^N$ auf sich, wozu insbesondere Rotationen und Spiegelungen zählen. Die Notation $TA := \{Ta : a \in A\}$ steht für das Bild von A unter (Multiplikation mit/Anwendung von) T.

Satz (über das **Banach-Tarski-Paradoxon**). *Seien A und $\widetilde{A}$ beliebige beschränkte Teilmengen mit nicht-leerem Innern in $\mathbb{R}^N$ mit $N \geq 3$. Dann gibt es ein $m \in \mathbb{N}$, Teilmengen $A_1, A_2, \ldots, A_m$ von $\mathbb{R}^N$, Vektoren $x_1, x_2, \ldots, x_m \in \mathbb{R}^N$ und lineare Isometrien $T_1, T_2, \ldots, T_m \in \mathcal{O}(\mathbb{R}^N)$, so dass $A = \bigcup_{i=1}^m A_i$ die* disjunkte *Vereinigung von $A_1, A_2, \ldots, A_m$ und $\widetilde{A} = \bigcup_{i=1}^m (x_i + T_i A_i)$ die* disjunkte *Vereinigung von $x_1+T_1A_1, x_2+T_2A_2, \ldots, x_m+T_mA_m$ ist. Insbesondere gibt es für $N \geq 3$ keine endlich additive und bewegungsinvariante Abbildung $\mu \colon \mathcal{P}(\mathbb{R}^N) \to [0,\infty]$ mit $0 < \mu\big(\overline{\mathrm{B}}_1^N(0)\big) < \infty$ (denn auf zwei disjunkten Einheitskugeln $\widetilde{A} = \overline{\mathrm{B}}_1^N(x) \mathbin{\dot\cup} \overline{\mathrm{B}}_1^N(y)$ mit $|y-x| > 2$ müsste μ einerseits den doppelten, andererseits denselben Wert wie auf der Einheitskugel $A = \overline{\mathrm{B}}_1^N(0)$ selbst haben).*

Speziell für $N = 3$ und den angesprochenen Fall $A = \overline{\mathrm{B}}_1^3(0)$ und $\widetilde{A} = \overline{\mathrm{B}}_1^3(x) \mathbin{\dot\cup} \overline{\mathrm{B}}_1^3(y)$ zeigt der hier nicht ausgeführte Beweis des Paradoxons übrigens, dass $m = 5$ gewählt werden kann.

Es gibt zwei Wege zur Umgehung des Maßproblems und der Problematik pathologischer, mit dem Auswahlaxiom gebildeter Mengen: Man kann …

- entweder den Definitionsbereich von μ einschränken und statt auf ganz $\mathcal{P}(\mathbb{R}^N)$ nur auf gewissen Teilsystemen $\mathscr{A} \subset \mathcal{P}(\mathbb{R}^N)$ arbeiten, die aber dennoch alle "vernünftigen" Mengen enthalten,
- oder die Forderung der σ-Additivität abschwächen zu ***σ*-Subadditivität**: $\mu(\bigcup_{i=1}^\infty A_i) \leq \sum_{i=1}^\infty \mu(A_i)$ für alle $A_1, A_2, A_3, \ldots \in \mathcal{P}(\mathbb{R}^N)$.

Im Folgenden wird im Wesentlichen der erste (und üblichere) Weg beschritten. Es gibt aber Zusammenhänge zwischen den beiden Vorgehensweisen, weshalb der zweite Weg in Abschnitt 2.4 noch einmal angesprochen wird.

Außerdem erweist es sich als sinnvoll, eine Theorie zu entwickeln, die auch die **Messungen von Wahrscheinlichkeiten** erlaubt. Hierbei soll für eine Menge A von möglichen Zufallsereignissen die Kennzahl $\mu(A) \in [0,1]$ die Wahrscheinlichkeit angeben, dass eines der Ereignisse aus A eintritt. Allerdings sind Mengen von Ereignissen im Allgemeinen keine Teilmengen von $\mathbb{R}^N$, und Bewegungsinvarianz von μ macht in diesem Kontext keinen Sinn. Deshalb wird im Folgenden eine **allgemeine Theorie *σ*-additiver Abbildungen**

$$\mu \colon \mathscr{A} \to [0,\infty]$$

auf Teilmengen $\mathscr{A}$ von $\mathcal{P}(\Omega)$ mit einer **beliebigen Grundmenge Ω** entwickelt. Im Vordergrund steht zumeist aber der **Modellfall des *k*-dimensionalen Inhalts** von Teilmengen von $\mathbb{R}^N$.

2.1 Allgemeine Maße (auf σ-Algebren)

Bevor nun eine präzise Definition und erste Eigenschaften der beschriebenen Messvorschriften $\mu \colon \mathscr{A} \to [0,\infty]$ angegeben werden, wird zunächst die passende Klasse von Definitionsbereichen $\mathscr{A} \subset \mathcal{P}(\Omega)$ eingeführt:

Definition (**σ-Algebren, Messräume**). *Sei Ω eine Menge. Eine Teilmenge $\mathscr{A}$ der Potenzmenge $\mathcal{P}(\Omega)$ von Ω heißt eine **σ-Algebra** über Ω, wenn gelten:*

- *$\emptyset \in \mathscr{A}$,*
- *$\mathscr{A}$ ist abgeschlossen unter Komplementbildung, d.h. für $A \in \mathscr{A}$ ist stets auch $\Omega \setminus A \in \mathscr{A}$,*

- *$\mathscr{A}$ ist abgeschlossen unter abzählbarer Vereinigung, d.h. für $A_1, A_2, A_3, \ldots \in \mathscr{A}$ ist stets auch $\bigcup_{i=1}^{\infty} A_i \in \mathscr{A}$.*

*Ist $\mathscr{A}$ eine σ-Algebra über Ω, so heißt das Paar $(\Omega, \mathscr{A})$ ein **Messraum**, die Mengen in $\mathscr{A}$ werden auch ($\mathscr{A}$-)**messbare Mengen** und die Mengen in $\mathcal{P}(\Omega) \setminus \mathscr{A}$ nicht-($\mathscr{A}$-)messbare Mengen genannt.*

Bemerkungen (zu σ-Algebren)**.** Elementare Beobachtungen sind:

(1) Jede σ-Algebra $\mathscr{A}$ über Ω ist abgeschlossen unter abzählbarem Durchschnitt, und es gilt stets $\Omega \in \mathscr{A}$.

(2) $\{\emptyset, \Omega\}$ und $\mathcal{P}(\Omega)$ sind (die extremen Beispiele von) σ-Algebren über Ω.

(3) Ist $(\mathscr{A}_i)_{i \in I}$ eine Familie von σ-Algebren über Ω (mit beliebiger Indexmenge I), so ist $\bigcap_{i \in I} \mathscr{A}_i$ wiederum eine σ-Algebra über Ω.

Definition (**erzeugte $\boldsymbol{\sigma}$-Algebren**)**.** *Sei $\mathscr{E}$ eine beliebige Teilmenge von $\mathcal{P}(\Omega)$. Nach der vorigen Bemerkung ist der Durchschnitt aller σ-Algebren $\mathscr{A}$ über Ω, für die $\mathscr{E} \subset \mathscr{A}$ gilt, selbst eine σ-Algebra. Man nennt diese σ-Algebra die **von $\mathscr{E}$ erzeugte $\boldsymbol{\sigma}$-Algebra** und bezeichnet sie mit $\sigma(\mathscr{E})$.*

Bemerkung. Nach Definition ist $\mathscr{E} \subset \sigma(\mathscr{E})$ und es gilt die Implikation

$$\mathscr{A} \ \sigma\text{-Algebra mit } \mathscr{E} \subset \mathscr{A} \implies \sigma(\mathscr{E}) \subset \mathscr{A} \,.$$

In diesem Sinne ist $\sigma(\mathscr{E})$ die kleinste σ-Algebra, die $\mathscr{E}$ enthält.

Es folgt die bereits ausführlich motivierte Definition von Messvorschriften $\mu \colon \mathscr{A} \to [0, \infty]$:

Definition (**Maße, Maßräume**)**.** *Sei $(\Omega, \mathscr{A})$ ein Messraum. Eine Abbildung*

$$\mu \colon \mathscr{A} \to [0, \infty]$$

*heißt ein **Maß** auf $(\Omega, \mathscr{A})$, wenn $\mu(\emptyset) = 0$ gilt und μ im folgenden Sinne σ-additiv ist:* Für disjunkte $A_1, A_2, A_3, \ldots \in \mathscr{A}$ *gilt stets*

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

*Ist μ ein Maß auf $(\Omega, \mathscr{A})$, so heißt das Tripel $(\Omega, \mathscr{A}, \mu)$ ein **Maßraum**.*

Definition (**Wahrscheinlichkeitsmaße**)**.** *Ist $(\Omega, \mathscr{A}, \mu)$ ein Maßraum mit $\mu(\Omega) = 1$, so heißt μ ein **Wahrscheinlichkeitsmaß** und $(\Omega, \mathscr{A}, \mu)$ ein **Wahrscheinlichkeitsraum**.*

Bemerkung (zu Linearkombinationen von Maßen)**.** Ist $(\mu_i)_{i \in I}$ eine beliebige Familie von Maßen auf $(\Omega, \mathscr{A})$, und $(\omega_i)_{i \in I}$ eine Familie von Parametern in $[0, \infty]$, so definiert die Vorschrift[6]

$$\left(\sum_{i \in I} \omega_i \mu_i\right)(A) := \sum_{i \in I} \omega_i \mu_i(A) \qquad \text{für } A \in \mathscr{A}$$

ein weiteres Maß $\sum_{i \in I} \omega_i \mu_i$ auf $(\Omega, \mathscr{A})$. Dabei wird die **in der Maßtheorie stets sinnvolle Konvention $\boldsymbol{0 \cdot \infty := 0 =: \infty \cdot 0}$** unterstellt.

[6] Eine möglicherweise überabzählbare Summe $\sum_{i \in I} a_i$ nichtnegativer Zahlen a_i kann und soll hier als $\sup \big\{ \sum_{i \in E} a_i \ : \ E \text{ ist endliche Teilmenge von } I \big\}$ verstanden werden. Ein endlicher Wert kann sich dabei aber nur ergeben, wenn alle bis auf höchstens abzählbar viele Summanden Null sind.

Beispiele (von elementaren Maßen auf der ganzen Potenzmenge)**.** Für eine beliebige Menge Ω lassen sich verschiedene Beispiele von Maßen auf $(\Omega, \mathcal{P}(\Omega))$ problemlos angeben: Man definiert zuerst

- das **Dirac-Maß** δ_x zu $x \in \Omega$ durch $\delta_x(A) := \mathbb{1}_A(x) := \begin{cases} 1 & \text{falls } x \in A \\ 0 & \text{falls } x \notin A \end{cases}$

für $A \in \mathcal{P}(\Omega)$. Mit der vorigen Bemerkung bildet man daraus (wobei stets $A \in \mathcal{P}(\Omega)$)

- das **Zählmaß** $\xi := \sum_{x\in\Omega} \delta_x$, das $\xi(A) = \#A =$ Anzahl der Elemente von A erfüllt und dem **0-dimensionalen Inhalt aus der Einleitung** entspricht,
- das Nullmaß $0 \cdot \xi$ mit $(0 \cdot \xi)(A) = 0$,
- das ∞-Maß $\infty \cdot \xi$ mit $(\infty \cdot \xi)(A) = \begin{cases} \infty & \text{falls } A \neq \emptyset \\ 0 & \text{falls } A = \emptyset \end{cases}$
- und allgemeiner zu jeder Familie $(\omega_x)_{x\in\Omega}$ von Parametern in $[0, \infty]$ ein **diskretes Maß** $\omega := \sum_{x\in\Omega} \omega_x \delta_x$ mit $\omega(\{x\}) = \omega_x$ und $\omega(A) = \sum_{x\in A} \omega_x$. Ist $\sum_{x\in\Omega} \omega_x = 1$, so ist ω ein diskretes Wahrscheinlichkeitsmaß.

Man prüft leicht nach, dass diskrete Maße (und folglich alle vorausgehenden Beispiele) sich nicht nur σ-additiv, sondern sogar überabzählbar-additiv verhalten. Ein Beispiel eines Maßes auf $(\Omega, \mathcal{P}(\Omega))$, das die letzte stärkere Eigenschaft nicht mehr hat, ist (bei überabzählbarem Ω)

- das Überabzählbarkeitsmaß $\aleph(A) := \begin{cases} 0 & \text{falls } A \text{ endlich oder abzählbar} \\ \infty & \text{falls } A \text{ überabzählbar} \end{cases}$.

Alle hier angegebenen Beispiele von Maßen sind sehr elementar und erfordern weder die Einschränkung auf ein Teilsystem von $\mathcal{P}(\Omega)$ noch überhaupt die Entwicklung ernsthafter Theorie. Die Konstruktion von interessanteren Maßen wie dem k-dimensionalen Inhalt mit $k > 0$ ist aber etwas aufwendiger und wird deshalb erst im nächsten Abschnitt erfolgen.

Nichtsdestotrotz werden jetzt schon einige Operationen mit Maßen und und allgemeine Eigenschaften eingeführt.

Definitionen & Bemerkungen (zur Fortsetzung und Einschränkung von Maßen)**.** *Seien $(\Omega, \mathscr{A})$ ein Messraum und $X \in \mathcal{P}(\Omega)$. Dann ist*

$$\mathscr{A}|X := \{A \cap X \,:\, A \in \mathscr{A}\}$$

eine σ-Algebra über X, genannt die **Spur-σ-Algebra** *von $\mathscr{A}$ auf X. Maße auf $\mathscr{A}|X$ kann man stets auf $\mathscr{A}$ erweitern:*

(I) *Ist μ ein Maß auf $(X, \mathscr{A}|X)$, so erhält man daraus ein Maß μ^0 auf $(\Omega, \mathscr{A})$ durch*

$$\mu^0(A) := \mu(A \cap X) \qquad \text{für } A \in \mathscr{A}\,.$$

Ist $X \in \mathscr{A}$, so ist $\mathscr{A}|X = \mathscr{A} \cap \mathcal{P}(X)$ und es gelten zusätzlich:

(II) *Die Einschränkung $\mu|_{\mathscr{A}|X}$ eines Maßes μ auf $(\Omega, \mathscr{A})$ ist ein Maß auf $(X, \mathscr{A}|X)$.*

(III) *In (I) ist das μ^0 ist eine Fortsetzung von μ, also $\mu^0|_{\mathscr{A}|X} = \mu$.*

(IV) *Ist μ ein Maß auf $(\Omega, \mathscr{A})$, so erhält man ein weiteres Maß $\mathbb{1}_X \cdot \mu$ auf $(\Omega, \mathscr{A})$ durch $(\mathbb{1}_X \cdot \mu)(A) := \mu(A \cap X)$ für $A \in \mathscr{A}$, mit anderen Worten ist $\mathbb{1}_X \cdot \mu = \big(\mu|_{\mathscr{A}|X}\big)^0$ die Fortsetzung der Einschränkung von μ. Der Hintergrund der Schreibweise $\mathbb{1}_X \cdot \mu$ wird im späteren Abschnitt 2.6 klarer.*

Satz (über **Stetigkeitseigenschaften von Maßen**)**.** *Sei $(\Omega, \mathscr{A}, \mu)$ ein Maßraum und seien $A_1, A_2, A_3, \ldots \in \mathscr{A}$.*

(I) *Dann gilt*

$$\mu\bigg(\bigcup_{i=1}^{\infty} A_i\bigg) = \lim_{m\to\infty} \mu\bigg(\bigcup_{i=1}^{m} A_i\bigg).$$

(II) *Ist $\mu(A_i) < \infty$ für ein $i \in \mathbb{N}$, so gilt außerdem*

$$\mu\bigg(\bigcap_{i=1}^{\infty} A_i\bigg) = \lim_{m\to\infty} \mu\bigg(\bigcap_{i=1}^{m} A_i\bigg).$$

Bemerkungen (zu den Stetigkeitseigenschaften)**.**

(1) Im Falle einer aufsteigenden Folge $A_1 \subset A_2 \subset A_3 \subset \ldots$ vereinfacht sich die erste Aussage zu $\mu\big(\bigcup_{i=1}^{\infty} A_i\big) = \lim_{m\to\infty} \mu(A_m)$.

(2) Und im Falle einer absteigenden Folge $A_1 \supset A_2 \supset A_3 \supset \ldots$ mit $\mu(A_i) < \infty$ für ein $i \in \mathbb{N}$ gibt die zweite Aussage $\mu\big(\bigcap_{i=1}^{\infty} A_i\big) = \lim_{m\to\infty} \mu(A_m)$.

(3) Die Notwendigkeit der Voraussetzung $\mu(A_1) < \infty$ im zweiten Teil erkennt man am Beispiel $(\Omega, \mathscr{A}, \mu) = (\mathbb{N}, \mathcal{P}(\mathbb{N}), \xi)$ mit $A_i = \{m \in \mathbb{N} : m \geq i\}$.

Beweis des Satzes. Man betrachtet die disjunkt gemachten Mengen

$$B_i := A_i \setminus \bigcup_{j=1}^{i-1} A_j\,.$$

Mit der σ-Additivität von μ erhält man

$$\mu\bigg(\bigcup_{i=1}^{\infty} B_i\bigg) = \sum_{i=1}^{\infty} \mu(B_i) = \lim_{m\to\infty} \sum_{i=1}^{m} \mu(B_i) = \lim_{m\to\infty} \mu\bigg(\bigcup_{i=1}^{m} B_i\bigg).$$

Da $\bigcup_{i=1}^{m} B_i = \bigcup_{i=1}^{m} A_i$ für alle $m \in \mathbb{N} \cup \{\infty\}$ gilt, zeigt diese die erste Behauptung. Die zweite Behauptung ergibt sich durch Anwendung der ersten auf die Komplemente $A_{i_0} \setminus A_i$, wenn ein $i_0 \in \mathbb{N}$ mit $\mu(A_{i_0}) < \infty$ fixiert wird. □

2.2 Das Lebesgue-Maß (auf der Borel-σ-Algebra)

Dieser Abschnitt beschäftigt sich mit dem Problem des N-dimensionalen Inhalts von Teilmengen von $\mathbb{R}^N$, also mit dem Fall $k = N$ der Kapiteleinleitung. Wie zuvor erläutert, soll der N-dimensionale Inhalt durch ein Maß auf einer "großen" σ-Algebra über $\mathbb{R}^N$ modelliert werden, die alle "vernünftigen" Mengen enthält. Die Definition eines solche Maßes muss gewissermaßen alle Formeln für Inhaltsberechnungen (von Dreiecken, Quadern, Pyramiden, Kreisen, Kugeln und allen anderen geometrischen Figuren) enthalten und kann deshalb nicht so einfach hingeschrieben werden. Unter den verschiedenen möglichen Konstruktionsverfahren wird hier eines vorgestellt, das die Messvorschrift zunächst nur auf einem Halbring, einem "kleinen" Mengensystem, definiert und dann abstrakt auf eine σ-Algebra fortsetzt:

Definition (**Halbringe**[7])**.** *Sei Ω eine Menge. Eine Teilmenge $\mathcal{H}$ der Potenzmenge $\mathcal{P}(\Omega)$ heißt ein Halbring über Ω, wenn gelten:*

- $\emptyset \in \mathcal{H}$,
- *zu $A, B \in \mathcal{H}$ gibt es stets ein $m \in \mathbb{N}$ und* disjunkte $H_1, H_2, \ldots, H_m \in \mathcal{H}$ *mit* $A \setminus B = \bigcup_{i=1}^m H_i$,
- *$\mathcal{H}$ ist abgeschlossen unter endlichem Durchschnitt, d.h. für $A, \widetilde{A} \in \mathcal{H}$ gilt stets $A \cap \widetilde{A} \in \mathcal{H}$.*

Bemerkung. Jede σ-Algebra über Ω ist auch ein Halbring über Ω.

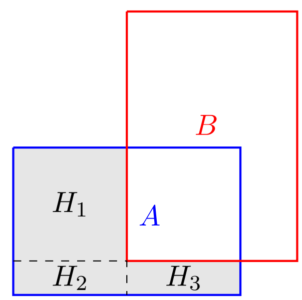


Abb. 1: $A \setminus B = H_1 \cup H_2 \cup H_3$ im Halbring $\mathcal{I}_2$

Beispiel und Definition. Für $a, b \in \mathbb{R}^N$ wird im Folgenden $a \leq b$ notiert, wenn $a_i \leq b_i$ für $i = 1, 2, \ldots, n$ gilt. Sind $a, b \in \mathbb{R}^N$ mit $a \leq b$, so wird der halboffene Quader $[a, b)$ als

$$\begin{aligned} [a,b) &:= [a_1, b_1) \times [a_2, b_2) \times \ldots \times [a_N, b_N) \\ &= \{x \in \mathbb{R}^N \,:\, a \leq x < b\} \subset \mathbb{R}^N . \end{aligned}$$

definiert. Mit diesen Bezeichnungen, von denen später auch naheliegende Abwandlungen verwendet werden, ist

$$\mathcal{I}_N := \big\{[a,b) \,:\, a, b \in \mathbb{R}^N \text{ mit } a \leq b\big\}$$

ein Halbring über $\mathbb{R}^N$. Im Folgenden wird vor allem dieser **Halbring der halboffenen Quader** relevant sein.

Definition (**Prämaße**)**.** *Sei $\mathcal{H}$ ein Halbring über Ω. Eine Abbildung*

$$\eta \colon \mathcal{H} \to [0, \infty]$$

heißt ein Prämaß, wenn $\eta(\emptyset) = 0$ gilt und η im folgenden Sinne σ-additiv ist: Für disjunkte $H_1, H_2, H_3, \ldots \in \mathcal{H}$ *mit* $\bigcup_{i=1}^\infty H_i \in \mathcal{H}$ *gilt stets*

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

Bemerkung (zu Prämaßen und Maßen)**.** Prämaße und Maße unterscheiden sich nur durch die Bauart des Definitionsbereichs (weshalb man hier auch $\bigcup_{i=1}^\infty H_i \in \mathcal{H}$ explizit voraussetzen muss). Daraus ergibt sich:

- Jedes Maß ist auch ein Prämaß und
- jedes Prämaß, das auf einer σ-Algebra definiert ist, ist bereits ein Maß.

[7] Die Terminologie "Halbring" wird durch den folgenden Sachverhalt motiviert: Als (Mengen-)Ringe bezeichnet man Mengensysteme, die zusätzlich zu den Eigenschaften des Halbrings abgeschlossen unter symmetrischer Differenz $A \Delta B := (A \setminus B) \cup (B \setminus A)$ sind. Dies ist insofern konsistent mit dem Begriff des algebraischen Rings, dass Mengen-Ringe mit der symmetrischen Differenz als Addition und dem Durchschnitt als Multiplikation die Struktur eines kommutativen Rings im Sinne der Algebra aufweisen.

Auf den halboffenen Quadern aus $\mathcal{I}_N$ ist anschaulich klar, wie man den N-dimensionalen Inhalt zu definieren hat:

Satz & Definition (Lebesguesches Prämaß). *Durch*

$$\lambda^N\big([a,b)\big) := \prod_{i=1}^{N}(b_i - a_i) \qquad \textit{für } [a,b) \in \mathcal{I}_N\,.$$

wird ein Prämaß $\lambda^N : \mathcal{I}_N \to [0,\infty]$ definiert. Man nennt λ^N das ***(N-dimensionale) Lebesgue-Prämaß****.*

Zum Beweis des Satzes ist folgendes Lemma über Halbringe nützlich:

Lemma. *Sei $\mathcal{H}$ ein Halbring. Zu $\ell \in \mathbb{N}$ und $A, A_1, A_2, \dots, A_\ell \in \mathcal{H}$ gibt es stets ein $m \in \mathbb{N}$ und* disjunkte *$H_1, H_2, \dots, H_m \in \mathcal{H}$ mit*

$$A \setminus \bigcup_{j=1}^{\ell} A_j = \bigcup_{i=1}^{m} H_i$$

Beweis des Lemmas. Man beweist das Lemma durch Induktion nach $\ell \in \mathbb{N}$. Detaillierter soll dies in einer Übungsaufgabe ausgearbeitet werden. □

Beweis des Satzes. Man zeigt sukzessive:

(1) *Endliche Additivität*: Für disjunkte $Q_1, Q_2, \dots, Q_\ell \in \mathcal{I}_N$ mit $\bigcup_{j=1}^{\ell} Q_j \in \mathcal{I}_N$ gilt

$$\lambda^N\Big(\bigcup_{j=1}^{\ell} Q_j\Big) = \sum_{j=1}^{\ell} \lambda^N(Q_j)\,,,$$

(2) σ-*Superadditivität*: Für disjunkte $Q_1, Q_2, Q_3, \dots \in \mathcal{I}_N$ mit $\bigcup_{j=1}^{\infty} Q_j \in \mathcal{I}_N$ gilt

$$\lambda^N\Big(\bigcup_{j=1}^{\infty} Q_j\Big) \geq \sum_{j=1}^{\infty} \lambda^N(Q_j)\,,$$

(3) σ-*Subadditivität*: Für beliebige $Q_1, Q_2, Q_3, \dots \in \mathcal{I}_N$ mit $\bigcup_{i=1}^{\infty} Q_i \in \mathcal{I}_N$ gilt

$$\lambda^N\Big(\bigcup_{i=1}^{\infty} Q_i\Big) \leq \sum_{i=1}^{\infty} \lambda^N(Q_i)\,.$$

Aus (2) und (3) ergeben sich die gewünschte σ-Additivität von λ^N und somit die Behauptung des Satzes. Die Beweise von (1), (2) und (3) folgen:

• *Nachweis von* (1): Durch Verfeinerung der Zerlegung und Hinzufügen leerer Quader kann man auf den Fall

$$(Q_j)_{j\in\{1,2,\dots,\ell\}} = \Big(\big[a_{1;\gamma_1-1}, a_{1;\gamma_1}\big) \times \big[a_{2;\gamma_2-1}, a_{2;\gamma_2}\big) \times \ldots \times \big[a_{N;\gamma_N-1}, a_{N;\gamma_N}\big)\Big)_{\gamma\in\{1,2,\dots,k\}^N}$$

reduzieren, wobei $\ell = k^N$ ist und $a_{i;0} \leq a_{i;1} \leq a_{i;2} \leq \ldots \leq a_{i;k}$ in $\mathbb{R}$ für $i = 1, 2, \dots, N$ gelten. In dieser Situation gilt

$$\bigcup_{j=1}^{\ell} Q_j = \big[a_{1;0}, a_{1;k}\big) \times \big[a_{2;0}, a_{2;k}\big) \times \ldots \times \big[a_{N;0}, a_{N;k}\big)\,,$$

und man rechnet die endliche Additivität (1) mit dem allgemeinen Distributivgesetz folgendermaßen nach:

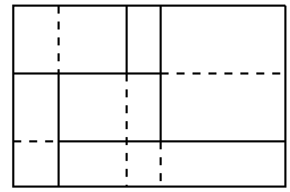


Abb. 2: Verfeinerung einer Quaderzerlegung

$$\lambda^N\Big(\big[a_{1;0}, a_{1;k}\big)\times\big[a_{2;0}, a_{2;k}\big)\times\ldots\times\big[a_{N;0}, a_{N;k}\big)\Big) = \prod_{i=1}^{N}(a_{i;k}-a_{i;0})$$
$$= \prod_{i=1}^{N}\sum_{\gamma_i=1}^{N}(a_{i;\gamma_i}-a_{i;\gamma_i-1}) = \sum_{\gamma\in\{1,2,\ldots,k\}^N}\prod_{i=1}^{N}(a_{i;\gamma_i}-a_{i;\gamma_i-1})$$
$$= \sum_{\gamma\in\{1,2,\ldots,k\}^N}\lambda^N\Big(\big[a_{1;\gamma_1-1}, a_{1;\gamma_1}\big) \times \big[a_{2;\gamma_2-1}, a_{2;\gamma_2}\big) \times \ldots \times \big[a_{N;\gamma_N-1}, a_{N;\gamma_N}\big)\Big)\,.$$

- *Nachweis von* (2): Seien $Q_1, Q_2, Q_3, \ldots \in \mathcal{I}_N$ disjunkt und

$$Q := \bigcup_{j=1}^{\infty} Q_j \in \mathcal{I}_N\,.$$

Zu jedem $\ell \in \mathbb{N}$ liefert das vorausgehende Lemma über Halbringe ein $m \in \mathbb{N}$ und disjunkte $H_1, H_2, \ldots, H_m \in \mathcal{I}_N$ mit

$$Q \setminus \bigcup_{j=1}^{\ell} Q_j = \bigcup_{i=1}^{m} H_i\,.$$

Es folgt, dass $Q_1, Q_2, \ldots, Q_\ell, H_1, H_2, \ldots, H_m$ alle disjunkt sind mit

$$Q = \bigcup_{j=1}^{\ell} Q_j \cup \bigcup_{i=1}^{m} H_i\,.$$

Wegen der endlichen Additivität (1) gilt also

$$\lambda^N(Q) = \sum_{j=1}^{\ell}\lambda^N(Q_j) + \sum_{i=1}^{k}\lambda^N(H_i) \geq \sum_{j=1}^{\ell}\lambda^N(Q_j)$$

für alle $\ell \in \mathbb{N}$, und Grenzübergang $\ell \to \infty$ zeigt

$$\lambda^N(Q) \geq \sum_{j=1}^{\infty}\lambda^N(Q_j)\,.$$

- *Nachweis von* (3): Seien $Q_1, Q_2, Q_3, \ldots \in \mathcal{I}_N$ mit $\bigcup_{i=1}^{\infty} Q_i \in \mathcal{I}_N$. Dies bedeutet, dass sich

$$\bigcup_{i=1}^{\infty} Q_i = [a, b) \qquad \text{und} \qquad Q_i = [a_i, b_i)$$

mit $a \le b$ und $a_i \le b_i$ in $\mathbb{R}^N$ schreiben lässt. Sei jetzt $\varepsilon \in \mathbb{R}_{>0}$. Dann gibt es ein $\widetilde{b} < b$ und zu jedem $i \in \mathbb{N}$ ein $\widetilde{a}_i < a_i$, so dass

$$\lambda^N\big([a,\widetilde{b})\big) \ge \lambda^N\Bigg(\bigcup_{i=1}^{\infty} Q_i\Bigg) - \varepsilon \qquad \text{und} \qquad \lambda^N\big([\widetilde{a}_i,b_i)\big) \le \lambda^N(Q_i) + \frac{\varepsilon}{2^i}$$

gelten. Es ist

$$[a,\widetilde{b}] \subset \bigcup_{i=1}^{\infty} Q_i \subset \bigcup_{i=1}^{\infty} (\widetilde{a}_i, b_i)\,,$$

also bilden die $(\widetilde{a}_i, b_i)$ eine offene Überdeckung des Kompaktums $[a,\widetilde{b}]$, die nach Definition der Überdeckungskompaktheit eine endliche Teilüberdeckung $\big((\widetilde{a}_{i_j}, b_{i_j})\big)_{j=1,2,\ldots,\ell}$ enthält. Setzt man $A_j := [\widetilde{a}_{i_j}, b_{i_j}) \cap [a,\widetilde{b}) \in \mathcal{I}_N$ für $j = 1,2,\ldots,\ell$, so lässt sich gemäß dem Lemma über Halbringe

$$A_j \setminus (A_1 \cup A_2 \cup \ldots \cup A_{j-1}) = \bigcup_{k=1}^{m_j} H_{j;k}$$

mit disjunkten $H_{j;1}, H_{j;2}, \ldots, H_{j;m_j} \in \mathcal{I}_N$ schreiben. Mit der endlichen Additivität (1) folgt[8]

$$\sum_{k=1}^{m_j} \lambda^N(H_{j;k}) \le \lambda^N(A_j)\,.$$

Wegen

$$[a,\widetilde{b}) = \bigcup_{j=1}^{\ell} A_j = \bigcup_{j=1}^{\ell} \bigcup_{k=1}^{m_j} H_{j;k}$$

gibt die endliche Additivität außerdem

$$\lambda^N\big([a,\widetilde{b})\big) = \sum_{j=1}^{\ell} \sum_{k=1}^{m_j} \lambda^N(H_{j;k})\,.$$

Schließlich setzt man die obigen (Un-)Gleichungen zusammen:

$$\begin{aligned}\lambda^N\Bigg(\bigcup_{i=1}^{\infty} Q_i\Bigg) &\le \lambda^N\big([a,\widetilde{b})\big) + \varepsilon = \sum_{j=1}^{\ell}\sum_{k=1}^{m_j} \lambda^N(H_{j;k}) + \varepsilon \le \sum_{j=1}^{\ell} \lambda^N(A_j) + \varepsilon \\ &\le \sum_{i=1}^{\infty} \lambda^N\big([\widetilde{a}_i, b_i)\big) + \varepsilon \le \sum_{i=1}^{\infty} \lambda^N(Q_i) + \sum_{i=1}^{\infty} \frac{\varepsilon}{2^i} + \varepsilon = \sum_{i=1}^{\infty} \lambda^N(Q_i) + 2\varepsilon\,.\end{aligned}$$

Da $\varepsilon \in \mathbb{R}_{>0}$ beliebig war, folgt die Behauptung von (3). □

Bemerkung (zum Beweis, dass λ^N ein Prämaß ist)**.** Die obige Vorgehensweise zum Nachweis der σ-Subadditivität (3) wird als $\boldsymbol{\frac{\varepsilon}{2^i}}$**-Trick** bezeichnet und ist in der Maßtheorie oft nützlich.

[8] Tatsächlich muss man zur Anwendung von (1) noch $A_j \setminus \bigcup_{k=1}^{m_j} H_{j;k}$ als endliche Vereinigung disjunkter Mengen aus $\mathcal{I}_N$ schreiben können. Auch dies ist gemäß dem Lemma aber möglich.

Wie schon angekündigt, soll λ^N nun vom Halbring $\mathcal{I}_N$ auf eine $\mathcal{I}_N$ enthaltende σ-Algebra fortgesetzt werden. Dazu verwendet man folgenden Satz über die Fortsetzung eines Prämaßes zu einem Maß auf der erzeugten σ-Algebra.

Satz (Maßfortsetzungssatz). *Seien Ω eine Menge und $\eta\colon \mathcal{H} \to [0,\infty]$ ein Prämaß auf einem Halbring $\mathcal{H}$ über Ω. Dann gelten folgende Aussagen.*

(I) ***Existenz*** *einer Fortsetzung: Es gibt ein Maß μ auf $(\Omega, \sigma(\mathcal{H}))$ mit $\mu|_{\mathcal{H}} = \eta$.*

(II) ***Eindeutigkeit*** *der Fortsetzung: Bei σ-endlichem η gibt es höchstens ein solches Maß μ.*

Der *Beweis* dieses Satzes wird erst im späteren Abschnitt 2.4 erfolgen.

Bei der Formulierung von Teil des Satzes (II) ging bereits die folgende Definition ein, die unten noch genauer diskutiert wird:

Definition (endliche und $\boldsymbol{\sigma}$-endliche Mengen und Maße). *Seien Ω eine Menge und η ein Prämaß auf einem Halbring $\mathcal{H}$ über Ω.*

- *Eine Menge $H \in \mathcal{H}$ heißt **$\boldsymbol{\eta}$-endlich**, wenn $\eta(H) < \infty$ gilt.*
- *Eine Menge $A \in \mathcal{P}(\Omega)$ heißt **$\boldsymbol{\eta}$-$\boldsymbol{\sigma}$-endlich**, wenn es η-endliche $H_1, H_2, H_3, \ldots \in \mathcal{H}$ mit $A \subset \bigcup_{i=1}^{\infty} H_i$ gibt.*
- *Ist Ω selbst η-(σ-)endlich, so nennt man η ein **($\boldsymbol{\sigma}$-)endliches Prämaß** oder, falls $\mathcal{H}$ sogar eine σ-Algebra ist, ein **($\boldsymbol{\sigma}$-)endliches Maß**.*

Bemerkungen (zum Maßfortsetzungssatz und der Voraussetzung der σ-Endlichkeit)**.**

(1) Als **entscheidende Anwendung** garantiert der Maßfortsetzungssatz die **Fortsetzbarkeit des Lebesgue-Prämaßes $\boldsymbol{\lambda^N}$ zu einem Maß** auf $(\mathbb{R}^N, \sigma(\mathcal{I}_N))$. Dieses Maß ordnet Mengen in $\sigma(\mathcal{I}_N)$ einen sinnvollen N-dimensionalen Inhalt zu und **löst im Fall $\boldsymbol{k = N}$ das** in der Einleitung des Kapitels beschriebene **Problem des $\boldsymbol{k}$-dimensionalen Inhalte von Teilmengen von $\mathbb{R}^{\boldsymbol{N}}$**. Es verbleibt allerdings das Problem, die σ-Algebra $\sigma(\mathcal{I}_N)$ und die in diesem Sinne behandelbaren/messbaren Mengen besser zu verstehen. Dieses Restproblem wird in Folge angegangen.

(2) Im **Fall $\boldsymbol{k < n}$** lässt sich für das Problem des k-dimensionalen Inhalts auf $\mathbb{R}^N$ **keine $\boldsymbol{\sigma}$-Endlichkeit** erwarten. Deshalb ist dieser Fall **nicht mit dem Maßfortsetzungssatz zu behandeln** und wird auf den späteren Abschnitt 2.11 verschoben.

(3) Insgesamt sollte man die σ-Endlichkeitsvoraussetzung in Teil (II) des Satzes als technische Voraussetzung einordnen. Nichtsdestotrotz kann auf diese Voraussetzung nicht verzichtet werden, wie man an folgenden beiden Beispiel(klass)en erkennt: Erstens wird durch σ-Endlichkeit ausgeschlossen, dass $\mathcal{H}$ gar nicht ganz Ω trifft, also $\bigcup_{H\in\mathcal{H}} H \neq \Omega$ ist und μ auf (Teilmengen von) $\Omega \setminus \bigcup_{H\in\mathcal{H}} H$ beliebig gewählt werden kann. Zweitens liegt auch beim Null-Prämaß $\eta \equiv 0$ auf dem Halbring $\mathcal{H} = \{\{x\} : x \in \Omega\} \cup \{\emptyset\}$ der Einermengen über einer überabzählbaren Menge Ω keine σ-Endlichkeit vor. Auch dieser Fall, in dem das Nullmaß und das Überabzählbarkeitsmaß zwei verschiedene Fortsetzungen auf $\sigma(\mathcal{H}) = \{A \in \mathcal{P}(\Omega) : A \text{ oder } \Omega \setminus A \text{ höchstens abzählbar}\}$ geben, wird in Teil (II) des Satzes durch die Voraussetzung der σ-Endlichkeit ausgeschlossen.

Als Nächstes folgt die angekündigte Untersuchung von $\sigma(\mathcal{I}_N)$. Diese kann in einem topologischen Rahmen, der hier kurz angerissen wird, erfolgen. Man kann Topologien aber auch vermeiden und im Folgenden stets an einen metrischen Raum Ω oder eine Teilmenge Ω von $\mathbb{R}^n$ denken, dann mit dem System $\mathcal{T}$ aller im üblichen Sinn offenen Teilmengen von Ω.

Eingeschobene Definition (Topologien, topologische Räume). *Sei Ω eine Menge. Man nennt eine Teilmenge $\mathcal{T}$ von $\mathcal{P}(\Omega)$ eine **Topologie** auf Ω und $(\Omega, \mathcal{T})$ einen **topologischen Raum**, wenn gelten:*

- $\emptyset \in \mathcal{T}$ *und* $\Omega \in \mathcal{T}$,
- $\mathcal{T}$ *ist abgeschlossen unter endlichem Durchschnitt, d.h. für* $O, U \in \mathcal{T}$ *ist stets auch* $O \cap U \in \mathcal{T}$,
- $\mathcal{T}$ *ist abgeschlossen unter beliebiger Vereinigung, d.h. aus* $O_i \in \mathcal{T}$ *für alle* i *in einer beliebigen Indexmenge* I *folgt stets* $\bigcup_{i \in I} O_i \in \mathcal{T}$.

*Die Mengen in $\mathcal{T}$ nennt man dann die **($\mathcal{T}$-)offenen Mengen** und die Mengen $A \in \mathcal{P}(\Omega)$ mit $\Omega \setminus A \in \mathcal{T}$ die **($\mathcal{T}$-) abgeschlossenen** Mengen.*

Bemerkung (zu topologischen Räumen)**.** Trotz der formalen Ähnlichkeit der Definitionen erfahren Topologien in der Mathematik eine gänzlich andere Interpretation als σ-Algebren. Man nutzt eine Topologie, um mit Hilfe von Umgebungen viele Konzepte der Analysis zu verallgemeinern: Beispielsweise machen Inneres, Äußeres, Rand, Abschluss und Kompaktheit von Mengen, Konvergenz von Folgen und Stetigkeit von Funktionen allgemein in/zwischen topologischen Räumen Sinn.

Definition (Borel-σ-Algebra). *Sei $(\Omega, \mathcal{T})$ ein topologischer Raum. Die von den offenen Mengen erzeugte σ-Algebra $\sigma(\mathcal{T})$ über Ω wird mit $\mathcal{B}(\Omega)$ bezeichnet*[9] *und heißt die **σ-Algebra der Borelschen Teilmengen** von Ω oder kurz die **Borel-σ-Algebra** über Ω.*

Zunächst sei jetzt eine technische Beobachtung eingeschoben, die später nützlich sein wird. Eine Vorstellung von $\mathcal{B}(\Omega)$ vermittelt erst die darauf folgende Serie von Bemerkungen.

Bemerkung (zur Borel-σ-Algebren auf Teilmengen)**.** Ist $(\Omega, \mathcal{T})$ ein topologischer Raum und versieht man eine Teilmenge X von Ω mit der Spurtopologie $\mathcal{T}|X := \{O \cap X \,:\, O \in \mathcal{T}\}$, so ist $\mathcal{B}(X)$ gleich der Spur-σ-Algebra $\mathcal{B}(\Omega)|X$. Folglich gilt $\mathcal{B}(X) = \mathcal{B}(\Omega) \cap \mathcal{P}(X)$ genau dann, wenn X selbst in $\mathcal{B}(\Omega)$ ist.

Beweis der Gleichheit $\mathcal{B}(X) = \mathcal{B}(\Omega)|X$. Da definitionsgemäß jede Menge aus $\mathcal{T}|X$ die Form $O \cap X$ mit $O \in \mathcal{T} \subset \mathcal{B}(\Omega)$ hat, gelten $\mathcal{T}|X \subset \mathcal{B}(\Omega)|X$ und $\mathcal{B}(X) = \sigma(\mathcal{T}|X) \subset \mathcal{B}(\Omega)|X$. Zudem enthält die σ-Algebra $\{B \in \mathcal{B}(\Omega) \,:\, B \cap X \in \mathcal{B}(X)\}$ aus einem ähnlichen Grund wie zuvor erst $\mathcal{T}$ und dann auch $\mathcal{B}(\Omega)$. Deshalb gilt auch $\mathcal{B}(\Omega)|X \subset \mathcal{B}(X)$. ☐

Bemerkungen und Beispiele (zu/von **Borel-Mengen**)**.**

(1) Per Definition ist $\mathcal{T} \subset \mathcal{B}(\Omega)$, mit anderen Worten: **Jede offene Menge ist Borelsch**.

(2) Da σ-Algebren unter Komplementbildung abgeschlossen sind, gilt auch: **Jede abgeschlossene Menge ist Borelsch**.

(3) Da σ-Algebren unter abzählbarem Durchschnitt und abzählbarer Vereinigung abgeschlossen sind, folgt weiter: Abzählbare Vereinigungen von abgeschlossenen Mengen, genannt **F_σ-Mengen**, und abzählbare Durchschnitte von offenen Mengen, genannt **G_δ-Mengen**, sind Borelsch.

(4) Aus dem gleichen Grund sind auch $\mathrm{F}_{\sigma\delta}$-Mengen und $\mathrm{G}_{\delta\sigma}$-Mengen, d.h. abzählbare Durchschnitte von F_σ-Mengen und abzählbare Vereinigungen von G_δ-Mengen, die analog zu verstehenden $\mathrm{F}_{\sigma\delta\sigma}$- und $\mathrm{G}_{\delta\sigma\delta}$-Mengen sowie weiter iterierte Bildungen alle stets Borelsch.

[9] Die Notation $\mathcal{B}(\Omega)$ ist etwas ungenau, da $\mathcal{B}(\Omega)$ nicht nur von Ω, sondern auch von $\mathcal{T}$ maßgeblich abhängt. Meist ist die zugrunde liegende Topologie $\mathcal{T}$ jedoch aus dem Kontext klar.

(5) Im Falle $\Omega = \mathbb{R}^N$ sind somit alle “vernünftigen” Mengen (und viele mehr) Borelsch. Insbesondere sind halboffene Quader wie in der Definition von $\mathcal{I}_N$ sowohl F_σ als auch G_δ und damit Borelsch, d.h. es gilt $\mathcal{I}_N \subset \mathcal{B}(\mathbb{R}^N)$. Als Nächstes wird diese Inklusion noch verschärft:

Satz (über die Erzeugung der Borel-σ-Algebra von $\mathbb{R}^N$)**.** *Halboffene Quader erzeugen die Borel-σ-Algebra über $\mathbb{R}^N$, d. h. für den zu Beginn des Abschnitts definierten Halbring $\mathcal{I}_N$ über $\mathbb{R}^N$ gilt*

$$\sigma(\mathcal{I}_N) = \mathcal{B}(\mathbb{R}^N)\,.$$

Beweis. Es reicht,

$$\mathcal{I}_N \subset \sigma(\mathcal{T}) \qquad \text{und} \qquad \mathcal{T} \subset \sigma(\mathcal{I}_N) \tag{$*$}$$

zu zeigen. Sind diese Inklusionen nämlich gezeigt, so gelten nach der Bemerkung zur erzeugten σ-Algebra auch $\sigma(\mathcal{I}_N) \subset \sigma(\mathcal{T})$ und $\sigma(\mathcal{T}) \subset \sigma(\mathcal{I}_N)$, also $\sigma(\mathcal{I}_N) = \sigma(\mathcal{T}) = \mathcal{B}(X)$.

Zum Beweis der ersten Inklusion in ($*$) betrachtet man einen halboffenen Quader $[a,b) \in \mathcal{I}_N$ mit $a \le b$ in $\mathbb{R}^N$. Man schreibt

$$[a,b) = \bigcap_{i=1}^{\infty} \big(a_1{-}\tfrac{1}{i}, b_1\big) \times \big(a_2{-}\tfrac{1}{i}, b_2\big) \times \ldots \times \big(a_N{-}\tfrac{1}{i}, b_N\big)$$

als abzählbaren Durchschnitt offener Mengen und erhält $[a,b) \in \sigma(\mathcal{T})$.

Zum Beweis der zweiten Inklusion in ($*$) sei O eine nicht-leere offene Menge in $\mathbb{R}^N$. Dann hat $x \in O$ positiven Abstand $\mathrm{dist}(x, \mathbb{R}^N \setminus O) \in \mathbb{R}_{>0}$ zum Komplement von O. Für $q \in O \cap \mathbb{Q}^N$ sei $r_q := \frac{1}{2\sqrt{N}}\mathrm{dist}(q, \mathbb{R}^N \setminus O) \in \mathbb{R}_{>0}$ und

$$Q_q := \big[q_1{-}r_q, q_1{+}r_q\big) \times \big[q_2{-}r_q, q_2{+}r_q\big) \times \ldots \times \big[q_N{-}r_q, q_N{+}r_q\big) \in \mathcal{I}_N\,.$$

Es reicht nun,

$$O = \bigcup_{q \in O \cap \mathbb{Q}^N} Q_q \in \sigma(\mathcal{I}_N)$$

zu verifizieren. Zum Beweis dieser Gleichheit überlegt man einerseits, dass der Abstand der Punkte in Q_q von q maximal[10] $r_q\sqrt{N} = \frac{1}{2}\mathrm{dist}(q, \mathbb{R}^N \setminus O)$ beträgt, diese Punkte also noch in O liegen. Andererseits argumentiert man wie folgt: Zu jedem $x \in O$ lässt sich ein zugehöriges $q \in O \cap \mathbb{Q}^N$ mit $(1{+}2\sqrt{N})\,|x{-}q| \le \mathrm{dist}(x, \mathbb{R}^N \setminus O)$ finden. Man erhält dann die Abschätzung

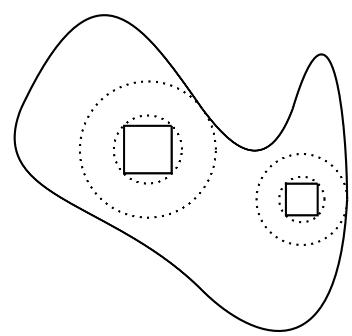


Abb. 3: Zwei Quadrate des Typs Q_q in einer offenen Menge in $\mathbb{R}^2$

$$r_q = \frac{1}{2\sqrt{N}}\mathrm{dist}(q, \mathbb{R}^N \setminus O) \ge \frac{1}{2\sqrt{N}}\big[\mathrm{dist}(x, \mathbb{R}^N \setminus O) - |x{-}q|\big] \ge |x{-}q|\,,$$

und folglich gilt $x \in Q_q$. □

Bemerkungen (zu Varianten des letzten Satzes)**.** Dieselbe Beweisidee zeigt:

(1) In jedem separablen[11] metrischen Raum wird die Borel-σ-Algebra vom System aller offenen (oder auch aller abgeschlossenen) Kugeln erzeugt.

[10] Für $x \in Q_q$ gilt ja $|x{-}q| = \sqrt{\sum_{i=1}^N (x_i{-}q_i)^2} \le \sqrt{\sum_{i=1}^N (r_q)^2} = r_q\sqrt{N}$.

[11] Definition: Ein topologischer Raum heißt separabel, wenn er eine abzählbare, dichte Teilmenge enthält. Dabei heißt eine Teilmenge eines topologischen Raumes dicht, wenn ihr Äußeres die leere Menge ist.

(2) Über Intervallen I in $\mathbb{R}$ gilt $\sigma(\mathcal{I}_1 \cap \mathcal{P}(I)) = \mathcal{B}(I)$.

Nachdem $\sigma(\mathcal{I}_N)$ als die Borel-σ-Algebra $\mathcal{B}(\mathbb{R}^N)$ identifiziert wurde, wird nun endlich die schon mehrfach angekündigte Fortsetzung des Lebesgue-Prämaßes zu einem Maß durchgeführt:

Korollar & Definition (**Lebesgue-Maß**). *Es gibt ein eindeutig bestimmtes Maß $\mathscr{L}^N$ auf $(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N))$, das auf $\mathcal{I}_N$ mit dem Lebesgueschen Prämaß λ^N übereinstimmt. Man nennt $\mathscr{L}^N$ das **(N-dimensionale) Lebesgue-Maß** (auf Borel-Mengen in $\mathbb{R}^N$).*

Beweis. Da $\mathbb{R}^N = \bigcup_{k=1}^\infty [-k,k)^N$ die Vereinigung der λ^N-endlichen Quader $[-k,k)^N$ ist, ist λ^N ein σ-endliches Prämaß. Nach dem Maßfortsetzungssatz kann λ^N somit auf genau eine Weise zu einem Maß $\mathscr{L}^N$ auf $(\mathbb{R}^N, \sigma(\mathcal{I}_N))$ mit $\mathscr{L}^N|_{\mathcal{I}_N} = \lambda^N$ fortgesetzt werden. Gemäß dem vorigen Satz ist $\sigma(\mathcal{I}_N) = \mathcal{B}(\mathbb{R}^N)$. □

Man erwartet natürlich und zu Recht, dass $\mathscr{L}^N$ die schon in der Einleitung des Kapitels erwähnte Bewegungsinvarianz aufweist, die Rotations- *und* Translationsinvarianz zusammenfasst. Für den Moment werden aber nur die Translationsinvarianz und zusätzlich ein natürliches Skalierungsverhalten von $\mathscr{L}^N$ nachgewiesen, der Beweis der Rotationsinsvarianz bietet sich in einem anderen Kontext an und wird auf den späteren Abschnitt 2.11 vertagt.

Satz (über die **Translationsinvarianz von $\mathscr{L}^N$**). *Seien $A \in \mathcal{B}(\mathbb{R}^N)$ und $x \in \mathbb{R}^N$. Dann ist $x{+}A \in \mathcal{B}(\mathbb{R}^N)$, und es gilt*

$$\mathscr{L}^N(x{+}A) = \mathscr{L}^N(A)\,.$$

Beweis. Sei $x \in \mathbb{R}^N$ für den restlichen Beweis fixiert. Man prüft nach, dass

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

eine σ-Algebra ist, die alle offenen Mengen enthält. Mit der Bemerkung zur erzeugten σ-Algebra folgt daraus $\mathcal{B}_x = \mathcal{B}(\mathbb{R}^N)$, also $x{+}A \in \mathcal{B}(\mathbb{R}^N)$ für alle $A \in \mathcal{B}(\mathbb{R}^N)$.

Durch $\mathscr{L}_x(A) := \mathscr{L}^N(x{+}A)$ für $x \in \mathcal{B}(\mathbb{R}^N)$ kann jetzt eine Abbildung $\mathscr{L}_x \colon \mathcal{B}(\mathbb{R}^N) \to [0,\infty]$ definiert werden. Man rechnet nach, dass $\mathscr{L}_x$ ein Maß ist, und findet für $[a,b) \in \mathcal{I}_N$:

$$\mathscr{L}_x([a,b)) = \lambda^N(x{+}[a,b)) = \prod_{i=1}^N \big((x_i{+}b_i){-}(x_i{+}a_i)\big) = \prod_{i=1}^N (b_i{-}a_i) = \lambda^N([a,b)).$$

Also stimmt $\mathscr{L}_x$ auf $\mathcal{I}_N$ mit λ^N überein, und gemäß dem Eindeutigkeitsteil des Maßfortsetzungssatzes folgt $\mathscr{L}_x = \mathscr{L}^N$. Dies ist die Behauptung. □

Analoge Argumente ergeben zwar nicht die Rotationsinsvarianz von $\mathscr{L}^N$, aber übrigens schon die von $\mathcal{B}(\mathbb{R}^N)$.

Satz (über das **Skalierungsverhalten von $\mathscr{L}^N$**). *Seien $A \in \mathcal{B}(\mathbb{R}^N)$ und $r \in \mathbb{R}_{\geq 0}$. Dann ist $rA \in \mathcal{B}(\mathbb{R}^N)$, und es gilt*

$$\mathscr{L}^N(rA) = r^N \mathscr{L}^N(A)\,.$$

Beweis. Für $r \in \mathbb{R}_{>0}$ beweist man dies weitgehend analog zum vorigen Satz mit dem Hilfs-Maß $\mathcal{L}_r(A) := r^{-N}\mathscr{L}^N(rA)$. Im Fall $r = 0$ ist nur $\mathscr{L}^N(\{0\}) = 0$ zu zeigen, und dies ergibt sich aus der für alle $\varepsilon \in \mathbb{R}_{>0}$ gültigen Abschätzung

$$\mathscr{L}^N(\{0\}) \leq \mathscr{L}^N([0,\varepsilon)^N) = \lambda^N([0,\varepsilon)^N) = \varepsilon^N\,. \qquad \square$$

Mit ähnlichen Argumenten lässt sich folgende universelle Eigenschaft von $\mathscr{L}^N$ nachweisen:

Satz (über die **Eindeutigkeit von $\mathscr{L}^N$**)**.** *Jedes translationsinvariante Maß auf* $(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N))$ *mit* $\mu(\,(0,1)^N) < \infty$ *hat die Form* $\mu = \gamma\,\mathscr{L}^N$ *mit einem* $\gamma \in \mathbb{R}_{\geq 0}$. *Insbesondere ist* $\mathscr{L}^N$ *das einzige durch* $\mathscr{L}^N(\,(0,1)^N) = 1$ *normierte, translationsinvariante Maß auf* $(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N))$ (*und auch das einzige durch* $\mathscr{L}^N([0,1)^N) = 1$, $\mathscr{L}^N(\,(0,1]^N) = 1$ *oder* $\mathscr{L}^N([0,1]^N) = 1$ *normierte*).

Beweisskizze. Man überlegt sich zunächst (ein ähnliches Argument kam schon in der Einleitung dieses Kapitels vor), dass mit $\mu(\,(0,1)^N)$ auch $\gamma := \mu([0,1)^N)$ endlich ist. Durch Zerlegung von $[0,1)^N$ in $k_1 k_2 \cdot \ldots \cdot k_N$ maßgleiche Teilquader ergibt sich dann, dass μ und $\gamma\,\mathscr{L}^N$ auf allen Quadern der Form $[0,\frac{1}{k_1}) \times [0,\frac{1}{k_2}) \times \ldots \times [0,\frac{1}{k_N})$ mit $k_i \in \mathbb{N}$ übereinstimmen. Daraus folgt die Übereinstimmung von μ und $\gamma\,\mathscr{L}^N$ auf allen Quadern $[a,b)$ mit $a \leq b$ in $\mathbb{Q}^N$. Diese Quader bilden einen Halbring, der $\mathcal{B}(\mathbb{R}^N)$ erzeugt. Gemäß dem Eindeutigkeitsteil des Maßfortsetzungssatzes gilt deshalb die Gleichheit $\mu = \gamma\,\mathscr{L}^N$ auf allen Mengen in $\mathcal{B}(\mathbb{R}^N)$. □

Die Vorgehensweise zur Konstruktion des Lebesgue-Maßes $\mathscr{L}^N$ lässt sich auf andere Maße übertragen. Beispielsweise funktioniert sie auch für folgende Klasse von Maßen auf Intervallen in $\mathbb{R}$ (die gemäß Abschnitt 2.12 tatsächlich schon alle „gutartigen" Maße auf Intervallen enthält).

Satz & Definition (**Lebesgue-Stieltjes-Maße**)**.** *Seien* I *ein offenes Intervall in* $\mathbb{R}$ *und* $F\colon I \to \mathbb{R}$ *eine nichtfallende Funktion. Dann gibt es ein eindeutig bestimmtes Maß* $\mathscr{L}^1_F$ *auf* $(I, \mathcal{B}(I))$ *mit*

$$\mathscr{L}^1_F([a,b)) = F(b-) - F(a-) \qquad \text{für } a \leq b \text{ in } I\,.$$

Man nennt $\mathscr{L}^1_F$ *das* ***Lebesgue-Stieltjes-Maß zu F*** (*auf Borel-Mengen in* $\mathbb{R}^N$).

Bemerkungen (zu Lebesgue-Stieltjes-Maßen)**.**

(1) Die linksseitigen Grenzwerte $F(a-)$ und $F(b-)$ existieren für monotones F stets in $\mathbb{R}$. Ist F linksseitig stetig, so kann man sie durch die Funktionswerte $F(a)$ und $F(b)$ ersetzen.

(2) Gemäß den Stetigkeitseigenschaften von Maßen erhält man für $a \leq b$ in I auch

$$\mathscr{L}^1_F((a,b)) = F(b-) - F(a+)\,, \quad \mathscr{L}^1_F((a,b]) = F(b+) - F(a+)\,, \quad \mathscr{L}^1_F([a,b]) = F(b+) - F(a-)\,.$$

Beispiele (von Lebesgue-Stieltjes-Maßen)**.** Zwei Beispiele im Falle $I = \mathbb{R}$ sind:

(1) Für $F(x) := \gamma x + \xi$ (mit $\gamma \in \mathbb{R}_{\geq 0}$, $\xi \in \mathbb{R}$) ist $\mathscr{L}^1_F = \gamma\,\mathscr{L}^1$ ein Vielfaches des Lebesgue-Maßes.

(2) Für $F := \mathbb{1}_{(p,\infty)}$ ist[12] $\mathscr{L}^1_F = \delta_p$ das Dirac-Maß zu $p \in \mathbb{R}$.

Beweisskizze zum Satz über die Existenz der Lebesgue-Stieltjes-Maße. Man zeigt zunächst, dass $\lambda^1_F([a,b)) := F(b-) - F(a-)$ ein Prämaß λ^F_1 auf dem Halbring $\mathcal{H} := \{[a,b) \,:\, a \leq b \text{ in } I\}$ über I definiert. Dazu verfährt analog zum Fall des Lebesgue-Prämaßes, kann die Argumentation im vorliegenden 1-dimensionalen Fall teils sogar vereinfachen und benötigt nur beim Nachweis der σ-Subadditivität folgende einfache Zusatzbeobachtung: Zu $[a,b) \in \mathcal{H}$ und $\varepsilon \in \mathbb{R}_{>0}$ gibt es immer ein $\widetilde{b} \in [a,b)$ mit $F(\widetilde{b}-) \geq F(b-) - \varepsilon$ und folglich $\lambda^1_F([a,\widetilde{b})) \geq \lambda^1_F([a,b)) - \varepsilon$.

Da man das offene Intervall I als abzählbare Vereinigung von Intervallen aus $\mathcal{H}$ schreiben kann, ist λ^1_F ein σ-endliches Prämaß. Gemäß dem Maßfortsetzungssatz besitzt λ^1_F also genau eine Fortsetzung zu einem Maß auf $(I, \sigma(\mathcal{H}))$, und es bleibt nur $\sigma(\mathcal{H}) = \mathcal{B}(I)$ zu zeigen. Letzteres folgt aber aus der Gleichheit $\sigma(\mathcal{I}_1 \cap \mathcal{P}(I)) = \mathcal{B}(I)$ der früheren Bemerkung (2) und den Beobachtungen $\mathcal{H} \subset \mathcal{I}_1 \cap \mathcal{P}(I) \subset \sigma(\mathcal{H})$ (wobei $\mathcal{H}$ und $\mathcal{I}_1 \cap \mathcal{P}(I)$ sich überhaupt nur bei von oben beschränktem I um die Intervalle unterscheiden, die bis an den rechten Randpunkt von I heranreichen). □

[12] Genauer gesprochen ist $\mathscr{L}^1_F$ die Einschränkung des Dirac-Maßes δ_p von $\mathcal{P}(\mathbb{R})$ auf $\mathcal{B}(\mathbb{R})$.

2.3 Nullmengen und Vervollständigung von Maßräumen

Definition (**Nullmengen**)**.** *Sei* $(\Omega, \mathscr{A}, \mu)$ *ein Maßraum. Eine Menge* $M \in \mathscr{A}$ *heißt* **(*μ*-)*Nullmenge***, *wenn* $\mu(M) = 0$ *gilt.*

Grundeigenschaften (von Nullmengen)**.** Sei $(\Omega, \mathscr{A}, \mu)$ ein Maßraum.

(1) Modifikation durch eine Nullmenge ändert das Maß nicht, d.h. für $A \in \mathscr{A}$ und eine μ-Nullmenge $M \in \mathscr{A}$ gilt

$$\mu(A \cup M) = \mu(A) = \mu(A \setminus M)\,.$$

Beweis. Wegen Monotonie ist $\mu(M \setminus A) = 0$, und per Additivität ergibt sich $\mu(A \cup M) = \mu(A) + \mu(M \setminus A) = \mu(A)$. Genauso folgt $\mu(A) = \mu(A \setminus M) + \mu(A \cap M) = \mu(A \setminus M)$. □

(2) Abzählbare Vereinigungen $\bigcup_{i=1}^{\infty} M_i$ von μ-Nullmengen $M_i \in \mathscr{A}$ sind μ-Nullmengen.

Beweis. Für endliche Vereinigungen ergibt sich dies induktiv aus (1). Mit einer Stetigkeitseigenschaft aus Abschnitt 2.1 folgt $\mu\big(\bigcup_{i=1}^{\infty} M_i\big) = \lim_{m\to\infty} \mu\big(\bigcup_{i=1}^{m} M_i\big) = 0$. □

Speziell für das Lebesgue-Maß $\mathscr{L}^N$ sind alle endlichen und alle abzählbaren Teilmengen von $\mathbb{R}^N$ Nullmengen. Es gibt aber auch überabzählbare Lebesgue-Nullmengen:

Beispiele (von Lebesgue-Nullmengen)**.**

(1) **Jeder *echte* affine Unterraum von $\mathbb{R}^N$**, also jeder k-dimensionale affine Unterraum von $\mathbb{R}^N$ mit $k < N$, **ist eine $\mathscr{L}^N$-Nullmenge.**

Teilbeweis. Man bemerkt zuerst, dass ein Unterraum abgeschlossen, also Borelsch ist. Da ein k-dimensionaler Unterraum mit $k < N$ stets in einem $(N-1)$-dimensionalen Unterraum, einer Hyperebene, enthalten ist, lässt sich o.E. $k = N-1$ annehmen. Tatsächlich wird jetzt aber nur eine *Achsen*-Hyperebene $H = \{x \in \mathbb{R}^N : x_\gamma = 0\}$ mit $\gamma \in \{1, 2, \ldots, N\}$ behandelt. Für allgemeine Hyperebenen folgt die Nullmengen-Eigenschaft mit der Translations- und Rotationsinvarianz von $\mathscr{L}^N$, wobei der Beweis für letztere noch aussteht und erst im späteren Abschnitt 2.11 nachgetragen wird.

Zur Behandlung von H sei $i \in \mathbb{N}$ beliebig. Aus

$$\mathscr{L}^N(H \cap [-i,i)^N) \le \lambda^N(\{x \in [-i,i)^N : 0 \le x_\gamma < \varepsilon\}) = (2i)^{N-1}\varepsilon$$

für alle $\varepsilon \in \mathbb{R}_{>0}$ schließt man, dass $H \cap [-i,i)^N$ eine $\mathscr{L}^N$-Nullmenge ist. Nach Grundeigenschaft (2) ist dann auch H selbst eine $\mathscr{L}^N$-Nullmenge. □

(2) Für $N \ge 2$ liefert (1) natürlich Beispiele überabzählbarer $\mathscr{L}^N$-Nullmengen (denn jeder Unterraum positiver Dimension ist überabzählbar). Es gibt aber auch **überabzählbare $\mathscr{L}^1$-Nullmengen**, und konkret ist die **Cantor-Menge** der mittleren Drittel eine solche; dies wird in Rahmen der Übungen gezeigt.

Oft erkennt man eine Menge T als μ-Nullmenge, weil $T \subset M$ für eine bereits bekannte μ-Nullmenge $M \in \mathscr{A}$ gilt. Genau genommen ist dieser Schluss aber nur dann erlaubt, wenn man Messbarkeit $T \in \mathscr{A}$ nachweisen kann. Bei nicht messbarem $T \notin \mathscr{A}$ — und das kann durchaus vorkommen — ist $\mu(T)$ dagegen gar nicht definiert, und T ist keine μ-Nullmenge im obigen Sinn. Für Maße, bei denen dieses Messbarkeits-Problem gar nicht auftritt, führt man zunächst eine Terminologie ein.

Definition (**vollständige Maße und Maßräume**). *Man nennt einen Maßraum* $(\Omega, \mathscr{A}, \mu)$ ***vollständig*** *und* μ *ein* ***vollständiges Maß*** *auf* $(\Omega, \mathscr{A})$, *wenn jede Teilmenge einer* μ*-Nullmenge zu* $\mathscr{A}$ *gehört* (*und damit selbst eine* μ*-Nullmenge ist*).

Ist ein Maß μ nicht vollständig, so lässt sich obiges Problem nicht-messbarer Teilmengen von Nullmengen einfach dadurch beheben, dass jede Teilmenge einer μ-Nullmenge per Definition zu einer messbaren Menge erklärt wird. Der folgende Satz sagt, dass dieses naheliegende Vorgehen tatsächlich zu einer Fortsetzung von μ zu einem Maß auf einer größeren σ-Algebra führt:

Satz & Definition (**Vervollständigung** von Maßen und Maßräumen). *Sei* $(\Omega, \mathscr{A}, \mu)$ *ein Maßraum. Dann ist*

$$\mathcal{M}_\mu := \{A \cup T \,:\, T \text{ ist Teilmenge einer } \mu\text{-Nullmenge und } A \in \mathscr{A}.\}$$

eine σ*-Algebra über* Ω, *es gilt* $\mathscr{A} \subset \mathcal{M}_\mu$, *und die Vorschrift*

$$\overline{\mu}(A \cup T) := \mu(A) \qquad \text{für } A \in \mathscr{A} \text{ und jede Teilmenge } T \text{ einer } \mu\text{-Nullmenge}$$

definiert ein vollständiges Maß $\overline{\mu}$ *auf* $(\Omega, \mathcal{M}_\mu)$ *mit* $\overline{\mu}|_{\mathscr{A}} = \mu$. *Man nennt* $\mathcal{M}_\mu$ *die* ***σ-Algebra der μ-messbaren Mengen*** *und bezeichnet* $\overline{\mu}$ *als* ***Vervollständigung*** *von* μ *beziehungsweise* $(\Omega, \mathcal{M}_\mu, \overline{\mu})$ *als* ***Vervollständigung*** *von* $(\Omega, \mathscr{A}, \mu)$.

Bemerkung (zur Minimalität der Vervollständigung). Ist $(\Omega, \widetilde{\mathscr{A}}, \widetilde{\mu})$ irgendein vollständiger Maßraum mit $\mathscr{A} \subset \widetilde{\mathscr{A}}$ und $\widetilde{\mu}|_{\mathscr{A}} = \mu$, so gilt $\mathcal{M}_\mu \subset \widetilde{\mathscr{A}}$ und $\widetilde{\mu}|_{\mathcal{M}_\mu} = \overline{\mu}$. In diesem Sinne ist $\overline{\mu}$ die eindeutig bestimmte Fortsetzung von μ zu einem vollständigen Maß auf dem kleinstmöglichen und natürlichen Definitionsbereich $\mathcal{M}_\mu$.

Beweis des Satzes. Man verifiziert problemlos, dass $\mathscr{A} \subset \mathcal{M}_\mu$ gilt und $\mathcal{M}_\mu$ abgeschlossen unter abzählbarer Vereinigung ist. Um zu sehen, dass $\mathcal{M}_\mu$ abgeschlossen unter Komplementbildung ist, schreibt man das Komplement von $A \cup T \in \mathcal{M}_\mu$ mit $A \in \mathscr{A}$ und $T \subset M$ mit einer μ-Nullmenge M als

$$\Omega \setminus (A \cup T) = \big[\underbrace{\Omega \setminus (A \cup M)}_{\in \mathscr{A}}\big] \cup \big[\underbrace{M \setminus (A \cup T)}_{\subset M}\big] \in \mathcal{M}_\mu\,.$$

Damit ist $\mathcal{M}_\mu$ eine σ-Algebra.

Weiterhin ist zu zeigen, dass $\overline{\mu}$ wohldefiniert ist. Besitze also $M \in \mathcal{M}_\mu$ zwei Darstellungen

$$A \cup T = M = \widetilde{A} \cup \widetilde{T}$$

mit $A, \widetilde{A} \in \mathscr{A}$, $T \subset M$, $\widetilde{T} \subset \widetilde{M}$ und mit μ-Nullmengen M und $\widetilde{M}$. Dann gilt

$$A \cup M \cup \widetilde{M} = \widetilde{A} \cup M \cup \widetilde{M}$$

und mit Grundeigenschaft (1) folgt

$$\mu(A) = \mu(A \cup M \cup \widetilde{M}) = \mu(\widetilde{A} \cup M \cup \widetilde{M}) = \mu(\widetilde{A})\,.$$

Folglich ist die Definition von $\overline{\mu}(M)$ unabhängig davon, ob man die Darstellung $M = A \cup T$ oder die alternative Darstellung $M = \widetilde{A} \cup \widetilde{T}$ betrachtet.

Schließlich erfolgt der Nachweis, dass $\overline{\mu}$ ein Maß ist, relativ problemlos mit den Grundeigenschaften (1) und (2). $\square$

Ausgehend von einem Prämaß η auf einem Halbring $\mathcal{H}$ kann man den Maßfortsetzungssatz aus Abschnitt 2.2 und die Vervollständigung hintereinanderschalten. Man setzt also zuerst η zu einem Maß μ auf $\sigma(\mathcal{H})$ und dann μ zu einem vollständigen Maß $\overline{\mu}$ auf $\mathcal{M}_\mu$ fort. Speziell für das Lebesgue-(Prä-)Maß prägt man folgende Terminologie für den so erhaltenen Maßraum:

Definition (vervollständigtes **Lebesgue-Maß**). *Die Vervollständigung von* $(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \mathscr{L}^N)$ *bezeichnet man mit* $(\mathbb{R}^N, \mathcal{M}^N, \overline{\mathscr{L}^N})$. *Dabei nennt man* $\mathcal{M}^N = \mathcal{M}_{\mathscr{L}^N}$ *die* ***σ-Algebra der Lebesgue-messbaren oder*** $\boldsymbol{\mathscr{L}^N}$***-messbaren Teilmengen*** *des* $\mathbb{R}^N$, *und* $\overline{\mathscr{L}^N}$ *heißt das (vervollständigte)* ***Lebesgue-Maß*** *(auf messbaren Mengen in* $\mathbb{R}^N$*). Oft schreibt man für das Maß* $\overline{\mathscr{L}^N}$ *auf* $(\mathbb{R}^N, \mathcal{M}^N)$ *dennoch kurz* $\mathscr{L}^N$ *und bringt die Notation* $\overline{\mathscr{L}^N}$ *nur dann zum Einsatz, wenn eine präzise Unterscheidung vom Lebesgue-Maß* $\mathscr{L}^N$ *auf* $(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N))$ *erforderlich ist.*

Bemerkung. Die Sätze aus Abschnitt 2.2 zur Translationsinvarianz und zum Skalierungsverhalten von $\mathscr{L}^N$ auf $(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N))$ übertragen sich auf $\overline{\mathscr{L}^N}$ auf $(\mathbb{R}^N, \mathcal{M}^N)$. Mit Hilfe der Minimalität der Vervollständigung kann man außerdem eine Eindeutigkeitsaussage[13] übertragen. In Abschnitt 2.12 wird noch ein „schönerer" Eindeutigkeitssatz für $\overline{\mathscr{L}^N}$ formuliert.

Bemerkungen (zu nicht-Lebesgue-messbaren und nicht-Borelschen Teilmengen von $\mathbb{R}^N$)**.**

(1) In Anbetracht der Translationsinvarianz von $\mathscr{L}^N$ folgt aus dem in der Einleitung des Kapitels angegebenen Satz über die Unlösbarkeit des Maßproblems die *strikte* Inklusion $\mathcal{M}^N \subsetneqq \mathcal{P}(\mathbb{R}^N)$, also die **Existenz nicht-Lebesgue-messbarer Mengen** in $\mathbb{R}^N$, für jedes $N \in \mathbb{N}$. Eine Analyse des früheren Beweises zeigt, dass tatsächlich die dort mit dem Auswahlaxiom konstruierte Menge A nicht in $\mathcal{M}^N$ ist. Solche Beispiele nicht-messbarer Mengen heißen Vitali-Mengen in $\mathbb{R}^N$.

(2) Tatsächlich gilt auch die strikte Inklusion $\mathcal{B}(\mathbb{R}^N) \subsetneqq \mathcal{M}^N$, also insgesamt

$$\boxed{\mathcal{B}(\mathbb{R}^N) \subsetneqq \mathcal{M}^N \subsetneqq \mathcal{P}(\mathbb{R}^N)\,,}$$

für jedes $N \in \mathbb{N}$. Um dies nachzuweisen, kann man für $N \geq 2$ eine Menge in $\mathcal{M}^N \setminus \mathcal{B}(\mathbb{R}^N)$ als $A \times \{0\}^{N-1}$ mit einer beliebigen nicht-Borelschen Teilmenge A von $\mathbb{R}$, beispielsweise einer Vitali-Menge A, konstruieren[14]. In nur einer Dimension gestaltet sich die Konstruktion einer Menge in $\mathcal{M}^1 \setminus \mathcal{B}(\mathbb{R})$ etwas raffinierter, tatsächlich erhält man eine solche aber beispielsweise als Urbild einer nicht-Borelschen Teilmenge von $[0,1]$ unter der sogenannten Cantor-Funktion; man vergleiche mit den Übungen.

(3) Tatsächlich kann man zeigen, dass $\mathcal{M}^N$ gleichmächtig zu $\mathcal{P}(\mathbb{R})$ und $\mathcal{B}(\mathbb{R}^N)$ gleichmächtig zu $\mathbb{R}$ ist. In diesem Sinn enthält $\mathcal{M}^N$ nicht nur mehr, sondern sogar viel mehr Mengen als $\mathcal{B}(\mathbb{R}^N)$. Der Beweis dieser Aussagen beruht aber auf fortgeschrittener Mengenlehre und übersteigt den Rahmen der in der Vorlesung behandelten Maßtheorie.

Zum Abschluss dieses Abschnitts wird festgehalten, dass man ganz zum Lebesgue-Maß auch Lebesgue-Stieltjes-Maße vervollständigen kann:

[13] Es handelt sich um folgenden Sachverhalt: Ist $(\mathbb{R}^N, \mathscr{A}, \mu)$ vollständiger Maßraum und sind $\mathscr{A}$ und μ translationsinvariant mit $\mathcal{B}(\mathbb{R}^N) \subset \mathscr{A}$ und $\mu([0,1)^N) = 1$, so ist $\mathcal{M}^N \subset \mathscr{A}$ und $\mu|_{\mathcal{M}^N} = \overline{\mathscr{L}^N}$.

[14] An dieser Stelle wird auf eine Bemerkung aus Abschnitt 2.2 zurückgegriffen, gemäß der $\mathcal{B}(\mathbb{R}\times\{0\}^{N-1}) = \mathcal{B}(\mathbb{R}^N) \cap \mathcal{P}(\mathbb{R}\times\{0\}^{N-1})$ gilt.

Definition (vervollständigte **Lebesgue-Stieltjes-Maße**)**.** *Seien I ein offenes Intervall in $\mathbb{R}$ und $F\colon I \to \mathbb{R}$ eine nichtfallende Funktion. Man schreibt $(I, \mathcal{M}_F^1, \overline{\mathscr{L}_F^1})$ für die Vervollständigung von $(I, \mathcal{B}(I), \mathscr{L}_F^1)$ und nennt $\overline{\mathscr{L}_F^1}$ das (vervollständigte) Lebesgue-Stieltjes-Maß zu F.*

2.4 Äußere Maße, Carathéodory-Konstruktion, Beweis des Maßfortsetzungssatzes

Der schon früher erwähnte, zweite Zugang zur Maßtheorie arbeitet mit Abbildungen auf der ganzen Potenzmenge und basiert maßgeblich auf dem folgenden Begriff.

Definition (**äußere Maße**)**.** *Eine Abbildung*

$$\alpha\colon \mathcal{P}(\Omega) \to [0, \infty]$$

*heißt ein **äußeres Maß** über Ω, wenn $\alpha(\emptyset) = 0$ gilt und α in folgendem Sinn **monoton und σ-subadditiv** ist:*

- *Für $A, \widetilde{A} \in \mathcal{P}(\Omega)$ mit $A \subset \widetilde{A}$ gilt stets $\alpha(A) \le \alpha(\widetilde{A})$.*
- *Für beliebige $A_1, A_2, A_3, \ldots \in \mathcal{P}(\Omega)$ gilt*

$$\alpha\bigg(\bigcup_{i=1}^{\infty} A_i\bigg) \le \sum_{i=1}^{\infty} \alpha(A_i)\,.$$

Aus einem äußeren Maß lässt sich stets ein Maß gewinnen:

Definition (**Carathéodory-Kriterium** für Messbarkeit)**.** *Sei α ein äußeres Maß über Ω. Eine Menge $A \in \mathcal{P}(\Omega)$ heißt **α-messbar**, wenn sie beliebige Testmengen T additiv zerlegt, wenn also*

$$\alpha(T) = \alpha(T \cap A) + \alpha(T \setminus A)$$

für alle $T \in \mathcal{P}(\Omega)$ gilt. Man bezeichnet das System der α-messbaren Teilmengen von Ω mit $\mathcal{M}_\alpha$.

Satz (von C. Carathéodory, $\sim$1914)**.** *Sei α ein äußeres Maß über Ω. Dann ist $\mathcal{M}_\alpha$ eine σ-Algebra über Ω, und $\alpha|_{\mathcal{M}_\alpha}$ ist ein vollständiges Maß auf $(\Omega, \mathcal{M}_\alpha)$.*

Beweis. Zunächst wird wie folgt gezeigt, dass $\mathcal{M}_\alpha$ abgeschlossen unter endlichen Mengenoperationen ist. Sind $A, \widetilde{A} \in \mathcal{M}_\alpha$ und ist $T \in \mathcal{P}(\Omega)$ eine beliebige Testmenge, so ergibt sich aus der Messbarkeit von A (erster und dritter Schritt) und von $\widetilde{A}$ (zweiter Schritt)

$$\alpha(T) = \alpha(T\cap A)+\alpha(T\setminus A) = \alpha(T\cap A\cap\widetilde{A})+\alpha((T\cap A)\setminus\widetilde{A})+\alpha(T\setminus A) = \alpha(T\cap A\cap\widetilde{A})+\alpha(T\setminus(A\cap\widetilde{A}))\,,$$

wobei der dritte Schritt benutzt, dass $T \setminus (A \cap \widetilde{A})$ durch A in $[T \setminus (A \cap \widetilde{A})] \cap A = (T \cap A) \setminus \widetilde{A}$ und $[T \setminus (A \cap \widetilde{A})] \setminus A = T \setminus A$ zerlegt wird. Damit ist $A \cap \widetilde{A} \in \mathcal{M}_\alpha$ nachgerechnet, und $\mathcal{M}_\alpha$ ist bezüglich endlichem Durchschnitt abgeschlossen. Man sieht leicht, dass $\mathcal{M}_\alpha$ bezüglich Komplementbildung und folglich auch bezüglich endlicher Vereinigung abgeschlossen ist.

Verifiziert man noch, dass $\mathcal{M}_\alpha$ auch unter abzählbarer Vereinigung *disjunkter* Mengen abgeschlossen ist, so folgt, dass $\mathcal{M}_\alpha$ eine σ-Algebra ist (denn für beliebige Mengen $B_i \in \mathcal{M}_\alpha$ ist dann auch $\bigcup_{i=1}^\infty B_i = \bigcup_{i=1}^\infty \big[B_i \cap (\Omega \setminus B_{i-1}) \cap \ldots \cap (\Omega \setminus B_2) \cap (\Omega \setminus B_1)\big] \in \mathcal{M}_\alpha$). Weist man außerdem σ-Additivität von $\alpha|_{\mathcal{M}_\alpha}$ nach, so folgt sofort, dass $\alpha|_{\mathcal{M}_\alpha}$ ein Maß ist.

Konkret bleibt also zu zeigen, dass für disjunkte $A_1, A_2, A_3, \ldots \in \mathcal{M}_\alpha$ stets

$$A := \bigcup_{i=1}^{\infty} A_i \in \mathcal{M}_\alpha \qquad \text{und} \qquad \alpha(A) = \sum_{i=1}^{\infty} \alpha(A_i)$$

gelten. Dazu seien $T \in \mathcal{P}(\Omega)$ und $k \in \mathbb{N}$. Mit der α-Messbarkeit der A_i erhält man

$$\begin{aligned}
\alpha\Big(T \cap \bigcup_{i=1}^{k} A_i\Big) &= \alpha(T \cap A_1) + \alpha\Big(T \cap \bigcup_{i=2}^{k} A_i\Big) \\
&= \alpha(T \cap A_1) + \alpha(T \cap A_2) + \alpha\Big(T \cap \bigcup_{i=3}^{k} A_i\Big) \\
&= \ldots = \sum_{i=1}^{k} \alpha(T \cap A_i)\,.
\end{aligned}$$

Da gemäß dem ersten Beweisschritt $\bigcup_{i=1}^{k} A_i \in \mathcal{M}_\alpha$ gilt, folgt hieraus

$$\alpha(T) = \alpha\Big(T \setminus \bigcup_{i=1}^{k} A_i\Big) + \alpha\Big(T \cap \bigcup_{i=1}^{k} A_i\Big) \geq \alpha(T \setminus A) + \sum_{i=1}^{k} \alpha(T \cap A_i)\,.$$

Durch Grenzübergang $k \to \infty$ ergibt sich

$$\alpha(T) \geq \alpha(T \setminus A) + \sum_{i=1}^{\infty} \alpha(T \cap A_i) \geq \alpha(T \setminus A) + \alpha(T \cap A) \geq \alpha(T)\,,$$

wobei die letzten beiden Ungleichungen Konsequenzen der σ-Subadditivität von α sind. Aus dieser Ungleichungskette liest man einerseits $A \in \mathcal{M}_\alpha$ und mit der speziellen Wahl $T = A$ andererseits $\alpha(A) = \sum_{i=1}^{\infty} \alpha(A_i)$ ab.

Schließlich ergibt sich die Vollständigkeit des Maßraums $(\Omega, \mathcal{M}_\alpha, \alpha|_{\mathcal{M}_\alpha})$ aus den Beobachtungen, dass für Teilmengen A von Nullmengen $\alpha(A) = 0$ gilt, und, dass $A \in \mathcal{P}(\Omega)$ mit $\alpha(A) = 0$ stets in $\mathcal{M}_\alpha$ liegt (denn $\alpha(T \cap A) = 0$ und $\alpha(T) \leq \alpha(T \cap A) + \alpha(T \setminus A) = \alpha(T \setminus A) \leq \alpha(T)$). $\square$

Umgekehrt liefert der nächste Satz ein allgemeines Verfahren zur Konstruktion äußerer Maße durch abzählbares Überdecken, mit dem man jedes Maß zu einem äußeren Maß fortsetzen kann.

Satz (über die **Carathéodory-Konstruktion**). *Sei $\mathscr{S}$ ein beliebige Teilmenge von $\mathcal{P}(\Omega)$ mit $\emptyset \in \mathscr{S}$, und sei $\eta\colon \mathscr{S} \to [0, \infty]$ eine beliebige Abbildung mit $\eta(\emptyset) = 0$. Dann definiert die Festlegung*

$$\eta^*(A) := \inf\Big\{ \sum_{j=1}^{\infty} \eta(S_j) \,:\, S_j \in \mathscr{S} \text{ und } A \subset \bigcup_{j=1}^{\infty} S_j \Big\} \qquad \text{für } A \in \mathcal{P}(\Omega)$$

(mit der Konvention $\inf \emptyset = \infty$) ein äußeres Maß η^* über Ω mit $\eta^*|_{\mathscr{S}} \leq \eta$. Ist η ein Prämaß auf einem Halbring $\mathcal{H}$ über Ω, so gilt sogar $\eta^*|_{\mathcal{H}} = \eta$, und jede Menge aus $\mathcal{H}$ ist η^*-messbar, d.h. $\mathcal{H} \subset \mathcal{M}_{\eta^*}$ und folglich $\sigma(\mathcal{H}) \subset \mathcal{M}_{\eta^*}$.

Der *Beweis* folgt in Kürze.

Bemerkungen (zur Carathéodory-Konstruktion)**.**

(1) Für $\eta\colon \mathscr{S} \to [0,\infty]$ mit $\eta(\emptyset) = 0$ und jedes Mengensystem $\mathscr{A}$ mit $\mathscr{S} \subset \mathscr{A}$ gilt

$$\big(\eta^*|_{\mathscr{A}}\big)^* = \eta^* .$$

In diesem Sinn ändert eine wiederholte Anwendung der Carathéodory-Konstruktion nichts.

Beweis. Die Ungleichung '$\geq$' ergibt sich, wenn man in der Definition von $\big(\eta^*|_{\mathscr{A}}\big)^*$ benutzt, dass η^* ein äußeres Maß ist. Ausgehend von $\eta^*|_{\mathscr{S}} \leq \eta$ erhält man außerdem $\big(\eta^*|_{\mathscr{A}}\big)^* \leq \big(\eta^*|_{\mathscr{S}}\big)^* \leq \eta^*$ und damit die umgekehrte Ungleichung '$\leq$'. □

(2) Ist η ein Prämaß auf einem Halbring $\mathcal{H}$, so kann man sich in der Definition von η^* auf die Betrachtung disjunkter S_j beschränken. Dies sieht man, indem man die disjunkt gemachten Mengen $\widetilde{S}_j := S_j \setminus (S_1 \cup S_2 \cup \ldots \cup S_{j-1})$ wie früher als Vereinigung disjunkter $H_{i,j} \in \mathcal{H}$ schreibt und $\eta(S_j) \geq \sum_{i=1}^{m_j} \eta(H_{i,j})$ abschätzt.

(3) Ist $\eta\colon \mathscr{A} \to [0,\infty]$ mit $\eta(\emptyset) = 0$ monoton und σ-subadditiv auf einer σ-Algebra $\mathscr{A}$, wie es insbesondere im Fall eines Maßes η auf $(\Omega, \mathscr{A})$ eintritt, so lässt sich die Definition von η^* vereinfachen zu

$$\eta^*(A) = \min\{\eta(\widetilde{A}) \,:\, \widetilde{A} \in \mathscr{A}\,,\, A \subset \widetilde{A}\}\,.$$

Für äußere Maße η über Ω (dann $\mathscr{A} = \mathcal{P}(\Omega)$) trivialisiert sich $\eta^* = \eta$ natürlich.

Der *Beweis* dieser Bemerkung ist Thema der Übungen.

Eine technische Besonderheit aller mittels Carathéodory-Konstruktion erhaltenen äußeren Maße kann mit der folgenden Begriffsbildung beschrieben werden:

Definition (**reguläre äußere Maße**)**.** *Sei α ein äußeres Maß über Ω und $\mathscr{R} \subset \mathcal{M}_\alpha$. Dann nennt man α ein **$\mathscr{R}$-reguläres äußeres Maß**, wenn für alle $A \in \mathcal{P}(\Omega)$ gilt:*

$$\alpha(A) = \inf\{\alpha(R) \,:\, R \in \mathscr{R} \text{ und } A \subset R\}\,.$$

*Als **reguläres äußeres Maß** bezeichnet man ein $\mathcal{M}_\alpha$-reguläres äußeres Maß α.*

Bemerkung (zur **Regularität** (bei) der Carathéodory-Konstruktion)**.** Aus obigen Bemerkungen (1) und (3) folgt, dass das äußere Maß η^* des Satzes stets

$$\eta^*(A) = \Big(\eta^*|_{\sigma(\mathscr{S})}\Big)^*(A) = \min\{\eta^*(S) \,:\, S \in \sigma(\mathscr{S}) \text{ und } A \subset S\} \qquad \text{für alle } A \in \mathcal{P}(\Omega)$$

erfüllt. Ist η zumindest ein Prämaß auf einem Halbring $\mathcal{H} = \mathscr{S}$, so ist das **per Carathéodory-Konstruktion erhaltene äußere Maß** η^* wegen $\sigma(\mathcal{H}) \subset \mathcal{M}_{\eta^*}$ also **stets** $\sigma(\mathcal{H})$-regulär und damit insbesondere **regulär**.

Beweis des Satzes über die Carathéodory-Konstruktion. Monotonie von η^* und die Ungleichung $\eta^*|_{\mathscr{S}} \leq \eta$ ergeben sich direkt aus der Definition von η^*.

Als erster echter Beweisschritt wird σ-Subadditivität von η^* mit dem $\frac{\varepsilon}{2^i}$-Trick nachgewiesen. Seien dazu $A_1, A_2, A_3 \ldots \in \mathcal{P}(\Omega)$, wobei ohne Einschränkung $\eta^*(A_i) < \infty$ für alle $i \in \mathbb{N}$ angenommen werden kann. Es gibt dann zu beliebigem $\varepsilon \in \mathbb{R}_{>0}$ stets $S_{i,j} \in \mathscr{S}$ mit $A_i \subset \bigcup_{j=1}^\infty S_{i,j}$ und

$$\sum_{j=1}^\infty \eta(S_{i,j}) \leq \eta^*(A_i) + \frac{\varepsilon}{2^i}\,.$$

Wegen $\bigcup_{i=1}^{\infty} A_i \subset \bigcup_{i,j=1}^{\infty} S_{i,j}$ folgt

$$\eta^*\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i,j=1}^{\infty} \eta(S_{i,j}) \leq \sum_{i=1}^{\infty} \left[\eta^*(A_i) + \frac{\varepsilon}{2^i}\right] = \sum_{i=1}^{\infty} \eta^*(A_i) + \varepsilon\,.$$

Wegen der Beliebigkeit von $\varepsilon \in \mathbb{R}_{>0}$ ist η^* somit σ-subadditiv, also ein äußeres Maß.

Ist η ein Prämaß auf einem Halbring $\mathcal{H}$, so erhält man die η^*-Messbarkeit einer beliebigen Menge $A \in \mathcal{H}$ wie folgt: Seien $T \in \mathcal{P}(\Omega)$ und $S_1, S_2, S_3, \ldots \in \mathcal{H}$ mit $T \subset \bigcup_{j=1}^{\infty} S_j$. Mit der Halbringeigenschaft von $\mathcal{H}$ kann man dann $S_j \setminus A = \bigcup_{i=1}^{m_j} H_{i,j}$ für disjunkte $H_{1,j}, H_{2,j}, \ldots, H_{m_j,j} \in \mathcal{H}$ schreiben. Mit der Additivität von η und der Definition von η^* (angewandt für die Überdeckungen von $T \cap A$ durch die $S_j \cap A$ und von $T \setminus A$ durch die $H_{i,j}$) erhält man erst

$$\sum_{j=1}^{\infty} \eta(S_j) = \sum_{j=1}^{\infty} \eta(S_j \cap A) + \sum_{j=1}^{\infty} \sum_{i=1}^{m_j} \eta(H_{i,j}) \geq \eta^*(T \cap A) + \eta^*(T \setminus A)\,.$$

Da die S_j eine beliebige Überdeckung von T durch Mengen aus $\mathcal{H}$ bilden, folgt per Definition von η^* dann

$$\eta^*(T) \geq \eta^*(T \cap A) + \eta^*(T \setminus A)\,.$$

In Anbetracht der Subadditivität von η^* zeigt dies die η^*-Messbarkeit von A.

Schließlich ist, weiter im Fall eines Prämaßes η auf einem Halbring $\mathcal{H}$, noch $\eta^*(A) \geq \eta(A)$ für $A \in \mathcal{H}$ nachzuweisen. Nach Definition der Carathéodory-Konstruktion und der zugehörigen Bemerkung (2) genügt es dazu, für *disjunkte* $S_1, S_2, S_3, \ldots \in \mathcal{H}$ mit $A \subset \bigcup_{j=1}^{\infty} S_j$ gerade

$$\sum_{j=1}^{\infty} \eta(S_j) \geq \eta(A)$$

zu zeigen. Zum Beweis dieser Ungleichung schließt man zunächst aus der Definition des Halbrings, dass $S_j \cap A \in \mathcal{H}$ sowie $S_j \setminus A = \bigcup_{i=1}^{m_j} H_{i,j}$ für disjunkte $H_{1,j}, H_{2,j}, \ldots, H_{m_j,j} \in \mathcal{H}$ gelten. Mit der σ-Additivität von η ergibt sich dann wie benötigt

$$\sum_{j=1}^{\infty} \eta(S_j) = \sum_{j=1}^{\infty} \left(\eta(S_j \cap A) + \sum_{i=1}^{m_j} \eta(H_{i,j})\right) \geq \sum_{j=1}^{\infty} \eta(S_j \cap A) = \eta(A)\,. \qquad \square$$

Mit den in diesem Abschnitt behandelten Sätzen lässt sich der erste Teil des Maßfortsetzungssatzes aus Abschnitt 2.2 problemlos beweisen:

Beweis des Existenzteils des Maßfortsetzungssatzes. Sei $\eta\colon \mathcal{H} \to [0,\infty]$ ein Prämaß auf einem Halbring $\mathcal{H}$ über Ω. Gemäß dem Satz über die Carathéodory-Konstruktion ist η^* ein äußeres Maß mit $\eta^*|_{\mathcal{H}} = \eta$ und $\sigma(\mathcal{H}) \subset \mathcal{M}_{\eta^*}$. Nach dem ersten Satz dieses Abschnitts ist dann $\eta^*|_{\mathcal{M}_{\eta^*}}$ und erst recht $\eta^*|_{\sigma(\mathcal{H})}$ eine Fortsetzung von η zu einem Maß. $\square$

Um den zweiten Teil des Maßfortsetzungssatzes zu beweisen, wird zusätzlich folgendes Eindeutigkeitslemma für äußere Maße verwendet:

Eindeutigkeitslemma. *Sei $\eta\colon \mathcal{H} \to [0,\infty]$ ein σ-endliches Prämaß auf einem Halbring $\mathcal{H}$ über Ω. Für jede Fortsetzung μ von η zu einem Maß auf $(\Omega, \mathscr{A})$ mit $\mathcal{H} \subset \mathscr{A} \subset \mathcal{M}_{\eta^*}$ gilt $\mu^* = \eta^*$.*

Beweis. Der Satz über die Carathéodory-Konstruktion gibt $\mu^* \leq \eta^*$ und $\mu^*|_{\mathcal{H}} = \eta = \eta^*|_{\mathcal{H}}$.

Als Nächstes zeigt man $\eta^*(A) \leq \mu(A)$ für alle $A \in \mathscr{A}$ mit $\eta^*(A) < \infty$. Für solches A und $\varepsilon \in \mathbb{R}_{>0}$ gibt es $S_1, S_2, S_3, \ldots \in \mathcal{H}$ mit $A \subset \bigcup_{j=1}^{\infty} S_j$ und

$$\sum_{j=1}^{\infty} \eta(S_j) \leq \eta^*(A) + \varepsilon\,.$$

Es wird nun angenommen, dass die S_j disjunkt sind (denn andernfalls kann man sie gemäß Bemerkung (2) zur Carathéodory-Konstruktion disjunkt machen). Mit der σ-Additivität von η^* auf den messbaren Mengen $A, S_j \in \mathcal{M}_{\eta^*}$ erhält man

$$\mu^*\biggl(\bigcup_{j=1}^{\infty} S_j \setminus A\biggr) \leq \eta^*\biggl(\bigcup_{j=1}^{\infty} S_j \setminus A\biggr) = \sum_{j=1}^{\infty} \eta^*(S_j) - \eta^*(A) \leq \varepsilon\,.$$

Aufgrund von $\bigcup_{j=1}^{\infty} S_j \in \sigma(\mathcal{H}) \subset \mathscr{A}$ folgt

$$\eta^*(A) \leq \sum_{j=1}^{\infty} \eta(S_j) = \mu\biggl(\bigcup_{j=1}^{\infty} S_j\biggr) = \mu(A) + \mu\biggl(\bigcup_{j=1}^{\infty} S_j \setminus A\biggr) \leq \mu(A) + \varepsilon\,.$$

Wegen der Beliebigkeit von ε, muss also $\eta^*(A) \leq \mu(A)$ gelten. Unter Verwendung der σ-Endlichkeitsvoraussetzung folgt $\eta^*(A) \leq \mu(A)$ für *alle* $A \in \mathscr{A}$, und mit Bemerkung (1) zur Carathéodory-Konstruktion erhält man schließlich $\eta^* = (\eta^*|_{\mathscr{A}})^* \leq \mu^*$. □

Beweis des Eindeutigkeitsteils des Maßfortsetzungssatzes. Sei $\eta\colon \mathcal{H} \to [0,\infty]$ ein fixiertes σ-endliches Prämaß auf einem Halbring $\mathcal{H}$ über Ω, und sei μ eine Fortsetzung von η zu einem Maß auf $(\Omega, \sigma(\mathcal{H}))$. Gemäß dem Satz über die Carathéodory-Konstruktion gelten $\mu = \mu^*|_{\sigma(\mathcal{H})}$ und $\sigma(\mathcal{H}) \subset \mathcal{M}_{\eta^*}$. Die Anwendung des vorausgehenden Lemmas mit $\mathscr{A} = \sigma(\mathcal{H})$ zeigt $\mu^* = \eta^*$. Insgesamt gilt also $\mu = \eta^*|_{\sigma(\mathcal{H})}$, und μ ist durch η eindeutig bestimmt. □

Für messbare Mengen bezüglich eines Maßes μ und eines äußeren Maßes α wurden sehr ähnliche Bezeichnungen eingeführt: Die σ-Algebra $\mathcal{M}_\mu$ der μ-messbaren Mengen (im vorigen Abschnitt im Rahmen der Vervollständigung definiert) und die σ-Algebra $\mathcal{M}_\alpha$ der α-messbaren Mengen (über das Carathéodory-Kriterium definiert) sind nur dadurch unterscheidbar, dass einmal ein Maß, einmal ein äußeres Maß als Index auftritt. Diese Ähnlichkeit ist beabsichtigt und wird nun dadurch gerechtfertigt, dass die beiden Konzepte messbarer Mengen im Wesentlichen übereinstimmen: Ist einerseits α ein äußeres Maß, so wurde $\mu := \alpha|_{\mathcal{M}_\alpha}$ schon als vollständiges Maß erkannt, und $\mathcal{M}_\mu = \mathcal{M}_\alpha$ gilt trivial. Geht man andererseits von einem σ-endlichen Maß μ aus, so besagt der nächste Satz, dass für $\alpha := \mu^*$ ebenfalls $\mathcal{M}_\alpha = \mathcal{M}_\mu$ gilt.

Satz („μ^*-Messbarkeit gleich μ-Messbarkeit“). *Sei μ ein σ-endliches*[15] *Maß auf $(\Omega, \mathscr{A})$ und μ^* das mittels Carathéodory-Konstruktion gebildete äußere Maß. Dann ist $(\Omega, \mathcal{M}_{\mu^*}, \mu^*|_{\mathcal{M}_{\mu^*}})$ die Vervollständigung von $(\Omega, \mathscr{A}, \mu)$, d.h. es gelten $\mathcal{M}_{\mu^*} = \mathcal{M}_\mu$ und $\mu^*|_{\mathcal{M}_{\mu^*}} = \overline{\mu}$.*

[15] Auf die σ-Endlichkeitsvoraussetzung kann nicht verzichtet werden: Der Maßraum $(\Omega, \mathscr{A}, \xi|_{\mathscr{A}})$ mit der σ-Algebra $\mathscr{A} := \{A \in \mathcal{P}(\Omega) \;:\; A \text{ oder } \Omega \setminus A \text{ ist abzählbar}\}$ und dem Zählmaß ξ ist vollständig, also $\mathcal{M}_{\xi|_{\mathscr{A}}} = \mathscr{A}$. Aber für das äußere Maß $(\xi|_{\mathscr{A}})^* = \xi$ ist $\mathcal{M}_\xi$ ganz $\mathcal{P}(\Omega)$ und damit bei überabzählbarem Ω echt größer als $\mathscr{A}$.

Beweis des Satzes. Gemäß dem ersten Satz dieses Abschnitts ist $(\Omega, \mathcal{M}_{\mu^*}, \mu^*|_{\mathcal{M}_{\mu^*}})$ ein vollständiger Maßraum, und nach dem Satz über die Carathéodory-Konstruktion gelten $\mathscr{A} \subset \mathcal{M}_{\mu^*}$ und $\mu^*|_{\mathscr{A}} = \mu$. Mit der Bemerkung zur Minimalität der Vervollständigung folgen daher $\mathcal{M}_\mu \subset \mathcal{M}_{\mu^*}$ und $\mu^*|_{\mathcal{M}_\mu} = \overline{\mu}$. Um die umgekehrte Inklusion $\mathcal{M}_{\mu^*} \subset \mathcal{M}_\mu$ zu zeigen, betrachtet man ein $A \in \mathcal{M}_{\mu^*}$ und reduziert mit Hilfe der vorausgesetzten σ-Endlichkeit auf den Fall $\mu^*(A) < \infty$. Die frühere Bemerkung (3) liefert erst ein $\widetilde{A} \in \mathscr{A}$ mit $A \subset \widetilde{A}$ und $\mu^*(A) = \mu(\widetilde{A}) = \mu^*(\widetilde{A})$ und in Folge auch ein $M \in \mathscr{A}$ mit $\widetilde{A} \setminus A \subset M$ und $\mu^*(\widetilde{A} \setminus A) = \mu(M) = \mu^*(M)$. Mit dem Carathéodory-Kriterium für Messbarkeit und der Endlichkeit von $\mu^*(A)$ ergibt sich

$$0 = \mu^*(\widetilde{A}) - \mu^*(A) = \mu^*(\widetilde{A} \setminus A) = \mu^*(M)\,.$$

Damit ist $A = (A \setminus M) \cup (A \cap M) \in \mathcal{M}_\mu$, denn $A \setminus M = \widetilde{A} \setminus M$ ist in $\mathscr{A}$, und $A \cap M$ ist Teilmenge der μ-Nullmenge M. □

Insgesamt ergibt sich aus den Resultaten dieses Abschnitts die folgende Korrespondenz zwischen den beiden (schon in der Einleitung angesprochenen) Zugängen zur Maßtheorie:

Korollar (Korrespondenz zwischen Maßen und äußeren Maßen). *Für jede Menge Ω ist die Carathéodory-Konstruktion $(\Omega, \mathscr{A}, \mu) \mapsto \mu^*$ eine Bijektion von den σ-endlichen, vollständigen Maßräumen über Ω auf die σ-endlichen[16], regulären äußeren Maße über Ω. Die Umkehrung dieser Bijektion ist die Einschränkung $\alpha \mapsto \big(\Omega, \mathcal{M}_\alpha, \alpha|_{\mathcal{M}_\alpha}\big)$.*

Beweis. Aus den vorausgehenden Sätzen ergibt sich, dass die angegebenen Abbildungen wohldefiniert und zueinander invers sind. Im Einzelnen ist μ^* nach dem Satz über die Carathéodory-Konstruktion und den folgenden Bemerkungen ein reguläres äußeres Maß mit $\mu^*|_{\mathscr{A}} = \mu$. Damit ist klar, dass μ^* die σ-Endlichkeit von μ erbt, und nach dem letzten Satz gilt $\mathcal{M}_{\mu^*} = \mathcal{M}_\mu = \mathscr{A}$ (weil μ vollständig ist). Umgekehrt ist $(\Omega, \mathcal{M}_\alpha, \alpha|_{\mathcal{M}_\alpha})$ nach dem ersten Satz des Abschnitts ein vollständiger Maßraum, der die σ-Endlichkeit von α erbt, und wegen der Regularität von α und $\big(\alpha|_{\mathcal{M}_\alpha}\big)^*$ gilt $(\alpha|_{\mathcal{M}_\alpha})^* = \alpha$. □

Zum Abschluss dieses Kapitels sei noch festgehalten:

Korollar & Definition (äußeres Lebesgue-Maß). *Die aus dem Lebesgue-Prämaß und dem (vervollständigten) Lebesgue-Maß gebildeten äußeren Maße $\big(\lambda^N\big)^*$, $\big(\mathscr{L}^N\big)^*$ und $\big(\overline{\mathscr{L}^N}\big)^*$ stimmen überein, und die σ-Algebra ihrer messbaren Mengen ist $\mathcal{M}^N$. Man nennt das übereinstimmende äußere Maß das* ***(N-dimensionale) äußere Lebesgue-Maß****.*

Beweis. Aus dem Eindeutigkeitslemma dieses Abschnitts entnimmt man $\big(\mathscr{L}^N\big)^* = \big(\lambda^N\big)^*$, und aus dem letzten Satz folgt $\mathcal{M}_{(\mathscr{L}^N)^*} = \mathcal{M}^N$. Aufgrund der letzten Gleichheit kann das Lemma erneut angewandt werden, um auch $\big(\overline{\mathscr{L}^N}\big)^* = \big(\mathscr{L}^N\big)^*$ einzusehen. □

2.5 Messbare Funktionen

Dieser Abschnitt beschäftigt sich mit messbaren Funktionen, die, ganz grob gesprochen, messbare Mengen in einem Messraum in messbare Mengen in einem anderen Messraum überführen und damit die Messraum-Struktur(en) respektieren. Als Wertebereich werden dabei des Öfteren die erweiterten reellen Zahlen

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

[16] Für äußere Maße wurde σ-Endlichkeit, genau genommen, noch nicht definiert, deshalb hier die zu (Prä-)Maßen analoge Definition: Eine äußeres Maß α über Ω heißt σ-endlich, wenn es $E_1, E_2, E_3, \ldots \in \mathcal{P}(\Omega)$ mit $\alpha(E_i) < \infty$ für alle $i \in \mathbb{N}$ und $\bigcup_{i=1}^\infty E_i = \Omega$ gibt.

auftreten, die als topologischer[17] Raum mit zugehöriger Borel-σ-Algebra $\mathcal{B}(\overline{\mathbb{R}})$ aufgefasst werden und für die ab jetzt immer die Konventionen $0 \cdot (\pm\infty) := 0 =: (\pm\infty) \cdot 0$ unterstellt werden.

Definition (messbare Funktionen). *Seien $(\Omega, \mathscr{A})$ und $(\mathcal{X}, \mathscr{S})$ Messräume. Eine Funktion $f \colon \Omega \to \mathcal{X}$ heißt **$(\mathscr{A}, \mathscr{S})$-messbar**, wenn $f^{-1}(S) \in \mathscr{A}$ für alle $S \in \mathscr{S}$ gilt.*

Bemerkungen (zur Messbarkeit von Funktionen)**.**

- Gelegentlich notiert man für $(\mathscr{A}, \mathscr{S})$-messbares $f \colon \Omega \to \mathcal{X}$ auch $f \colon (\Omega, \mathscr{A}) \overset{\text{mb.}}{\to} (\mathcal{X}, \mathscr{S})$.
- Mit der Notation $f^{-1}(\mathscr{S}) := \{f^{-1}(S) : S \in \mathscr{S}\}$ kann die Bedingung aus der Definition der $(\mathscr{A}, \mathscr{S})$-Messbarkeit zu $f^{-1}(\mathscr{S}) \subset \mathscr{A}$ verkürzt werden.
- Dass man in der Definition mit Urbildern $f^{-1}(S)$ von $S \in \mathscr{S}$, nicht etwa mit Bildern $f(A)$ von $A \in \mathscr{A}$ arbeitet, liegt unter anderem daran, dass Urbilder sich besser mit Mengenoperationen vertragen: Zum Beispiel gilt stets $f^{-1}(S_1 \cap S_2) = f^{-1}(S_1) \cap f^{-1}(S_2)$, aber im Allgemeinen nur $f(A_1 \cap A_2) \subset f(A_1) \cap f(A_2)$. Tatsächlich ist mit $\mathscr{S}$ auch $f^{-1}(\mathscr{S})$ stets eine σ-Algebra, die sogenannte Urbild-σ-Algebra von $\mathscr{S}$ unter f.

Bezeichnungen (für messbare Funktionen)**.** Für Funktionen $f \colon \Omega \to \mathcal{X}$ mit topologischem Zielraum $\mathcal{X}$ benutzt man folgenden Abkürzungen:

(1) Anstelle von $(\mathscr{A}, \mathcal{B}(\mathcal{X}))$-messbar schreibt man kurz **$\mathscr{A}$-messbar**. Ist μ ein Maß auf $(\Omega, \mathscr{A})$, so vereinbart man als Abkürzung für $\mathcal{M}_\mu$-messbar die Bezeichnung **μ-messbar**[18].

(2) Ist Ω ebenfalls ein topologischer Raum, so heißen die $(\mathcal{B}(\Omega), \mathcal{B}(\mathcal{X}))$-messbaren Funktionen auch Borel-messbare Funktionen, Borelsche Funktionen oder kurz **Borel-Funktionen**.

(3) Für Lebesgue-messbares $\Omega \subset \mathbb{R}^N$, mit anderen Worten also für $\Omega \in \mathcal{M}^N$, nennt man $(\mathcal{M}^N|\Omega, \mathcal{B}(\mathcal{X}))$-messbare Funktionen $\Omega \to \mathcal{X}$ auch **$\mathscr{L}^N$-messbar** oder **Lebesgue-messbar**.

Bemerkungen (zu messbaren Funktionen)**.**

(1) Konstante Funktionen $f \colon \Omega \to \mathcal{X}$ sind stets $(\mathscr{A}, \mathscr{S})$-messbar.

(2) Eine charakteristische Funktion $\mathbb{1}_A \colon \Omega \to \overline{\mathbb{R}}$ mit $A \in \mathcal{P}(\Omega)$ ist genau dann $\mathscr{A}$-messbar, wenn $A \in \mathscr{A}$ gilt, wenn also A selbst $\mathscr{A}$-messbar ist.

(3) Den Nachweis der $(\mathscr{A}, \mathscr{S})$-Messbarkeit von f kann man sich oft beträchtlich erleichtern, wenn man einen Erzeuger $\mathscr{E}$ mit $\sigma(\mathscr{E}) = \mathscr{S}$ kennt. Zeigt man dann nämlich nur $f^{-1}(\mathscr{E}) \subset \mathscr{A}$, so folgt schon $(\mathscr{A}, \mathscr{S})$-Messbarkeit von f.

Zum *Beweis* dieser Behauptung zeigt man die allgemeine Gleichheit $f^{-1}(\sigma(\mathscr{E})) = \sigma(f^{-1}(\mathscr{E}))$ (denn dann bleibt $f^{-1}(\mathscr{S})$ als die von $f^{-1}(\mathscr{E})$ erzeugte σ-Algebra in $\mathscr{A}$). Die Inklusion „$\supset$“ der Gleichheit ergibt sich daraus, dass die σ-Algebra $f^{-1}(\sigma(\mathscr{E}))$ mit $f^{-1}(\mathscr{E})$ auch $\sigma(f^{-1}(\mathscr{E}))$ enthält. Die (hier eigentlich benötigte) Inklusion „$\subset$“ ergibt sich, weil $\{S \in \mathcal{P}(\mathcal{X}) : f^{-1}(S) \in \sigma(f^{-1}(\mathscr{E}))\}$ eine σ-Algebra ist, die $\mathscr{E}$ und damit auch $\sigma(\mathscr{E})$ enthält. □

[17] Tatsächlich kann man $\overline{\mathbb{R}}$ sogar als metrischen Raum auffassen, beispielsweise mit der bei Konvention $\arctan(\pm\infty) := \pm\frac{\pi}{2}$ durch $d(x,y) := |\arctan y - \arctan x|$ für $x, y \in \overline{\mathbb{R}}$ definierten Metrik d auf $\overline{\mathbb{R}}$. Die Topologie von $\overline{\mathbb{R}}$ ergibt sich dann als System der bezüglich dieser Metrik offenen Mengen. Dieselbe Topologie erhält man aber auch aus etwas anderen Metriken, weshalb man auf $\overline{\mathbb{R}}$ tatsächlich nur die Topologie, aber keine Metrik als kanonisch betrachtet.

[18] Übrigens ist μ-Messbarkeit per Definition dasselbe wie $\overline{\mu}$-Messbarkeit bezüglich der Vervollständigung $\overline{\mu}$.

(4) Insbesondere lässt sich der vorige Punkt auf Erzeuger der Borel-σ-Algebra anwenden und liefert beispielsweise folgende Aussage: Wenn $f^{-1}((a,\infty]) \in \mathscr{A}$ für alle $a \in \mathbb{Q}$ gilt, so ist $f\colon \Omega \to \overline{\mathbb{R}}$ schon $\mathscr{A}$-messbar.

(5) **Kompositionsregel**: Für $(\mathscr{A},\mathscr{S})$-messbares $f\colon \Omega \to \mathcal{X}$ und $(\mathscr{S},\mathscr{G})$-messbares $g\colon \mathcal{X} \to \mathcal{Y}$ ist die Komposition $g \circ f\colon \Omega \to \mathcal{Y}$ stets $(\mathscr{A},\mathscr{G})$-messbar.

Bemerkungen (zu Borel-Funktionen und Lebesgue-messbaren Funktionen)**.** Für topologische Räume Ω, $\mathcal{X}$ und $\mathcal{Y}$ ergeben sich aus den vorausgehenden Bemerkungen einige Folgerungen:

(1) Ist $f\colon \Omega \to \mathcal{X}$ stetig, so ist für jede offene Menge O in $\mathcal{X}$ das Urbild $f^{-1}(O)$ offen in Ω. Deshalb **ist jede stetige Funktion $\boldsymbol{f\colon \Omega \to \mathcal{X}}$ eine Borel-Funktion**.

(2) Für Borel-Funktionen $f\colon \Omega \to \mathcal{X}$ und $g\colon \mathcal{X} \to \mathcal{Y}$ ist auch $g\circ f\colon \Omega \to \mathcal{Y}$ stets Borel-Funktion.

Ist μ ein Maß über Ω mit $\mathcal{B}(\Omega) \subset \mathcal{M}_\mu$ (zum Beispiel $\mathscr{L}^N$, eines der später eingeführten Radon-Maße oder eine Einschränkung eines solchen auf eine seiner messbaren Mengen), so gilt zudem:

(3) **Jede Borel-Funktion $\boldsymbol{f\colon \Omega \to \mathcal{X}}$ ist $\boldsymbol{\mu}$-messbar**.

(4) Für μ-messbares $f\colon \Omega \to \mathcal{X}$ und Borelsches $g\colon \mathcal{X} \to \mathcal{Y}$ ist $g \circ f\colon \Omega \to \mathcal{Y}$ stets μ-messbar.

(5) Dagegen ist für stetiges $f\colon \mathbb{R}^N \to \mathbb{R}^N$ und $\mathscr{L}^N$-messbares $g\colon \mathbb{R}^N \to \mathbb{R}$ die Komposition $g \circ f$ im Allgemeinen nicht[19] $\mathscr{L}^N$-messbar.

Tatsächlich sind auch stückweise stetige Funktionen sowie alle „vernünftigen“ und praktisch relevanten Funktionen Borel-Funktionen und damit μ-messbar. Außerdem zeigen die Kompositionsregeln, dass **Summen, Linearkombinationen, Produkte, Quotienten, Maximum, Minimum, etc.**[20] von endlich vielen Borelschen beziehungsweise μ-messbaren Funktionen, soweit definiert, **wieder Borelsch beziehungsweise $\boldsymbol{\mu}$-messbar** sind. Dass Messbarkeit von Funktionen zudem bei Grenzübergängen erhalten bleibt, wird in diesem Abschnitt noch gezeigt.

Es folgt eine einfache und zugleich extrem nützliche Begriffsbildung der Maßtheorie:

Definition (fast-überall bestehende Eigenschaften). *Sei $(\Omega,\mathscr{A},\mu)$ ein Maßraum, und für jedes $x \in \Omega$ sei eine von x abhängige Eigenschaft oder Aussage $E(x)$ gegeben. Man spricht davon, dass $E(x)$ für* ***μ-fast-alle*** *$x \in \Omega$ gilt, wenn $\{x \in \Omega : E(x) \text{ gilt nicht.}\}$ eine $\overline{\mu}$-Nullmenge ist. In gleicher Bedeutung sagt man auch, die Eigenschaft E gelte* ***μ-fast-überall*** *auf Ω.*

Bemerkung (zu fast-überall bestehenden Eigenschaften)**.**

- Die Definition wird häufig auf Aussagen des Typs $f(x) \bowtie g(x)$ mit $\bowtie \in \{=,<,>,\le,\ge,\neq\}$ und Funktionen $f,g\colon \Omega \to \overline{\mathbb{R}}$ angewandt. Insbesondere bedeutet eine μ-fast-überall auf Ω gültige Gleichheit $f = g$, dass $\{x \in \Omega : f(x) \neq g(x)\}$ eine $\overline{\mu}$-Nullmenge ist.

[19] Für $N \geq 2$ sieht man dies, wenn man $f(x) := (x_1,0,0,\ldots,0)$ und $g := \mathbb{1}_{A\times\{0\}^{N-1}}$ mit einer Vitali-Menge A in $\mathbb{R}$ wählt, denn dann ist $(g\circ f)^{-1}(\{1\}) = A \times \mathbb{R}^{N-1} \notin \mathcal{M}^N$. Im Fall $N=1$ sieht man es an $f := (\mathrm{id}{+}F)^{-1}$ mit der Cantor-Funktion F und $g := \mathbb{1}_{f(A)}$ mit einer nicht-Lebesgue-messbaren Menge A, so dass $f(A)$ Teilmenge der Cantor-Menge ist.

[20] Es geht bei all diesen Operationen um Kompositionen des Typs $g \circ (f_1,f_2,\ldots,f_M)$, bei denen $f_1,f_2,\ldots,f_M$ mittels einer Borel-Funktion $g\colon \overline{\mathbb{R}}^M \to \overline{\mathbb{R}}$ zusammensetzt werden. Um $\mathscr{A}$-Messbarkeit der Komposition zu folgern, weist man mit Hilfe eines geeigneten Erzeugers von $\mathcal{B}(\overline{\mathbb{R}}^M)$ (wie $\{O_1{\times}O_2{\times}\ldots{\times}O_M : O_i \text{ offen}\}$) nach, dass aus $\mathscr{A}$-Messbarkeit der f_i die $\mathscr{A}$-Messbarkeit von $(f_1,f_2,\ldots,f_M)$ folgt, und wendet dann die Kompositionsregel an.

- Ein weiterer wichtiger Spezialfall ist die **μ-fast-überall Konvergenz** $f_k \underset{k\to\infty}{\longrightarrow} f$ von Funktionen $f_k\colon \Omega \to \mathcal{X}$ gegen eine Grenzfunktion $f\colon \Omega \to \mathcal{X}$.

Als erstes Beispiel-Resultat zu fast-überall bestehenden Eigenschaften wird festgehalten:

Proposition. *Sei $(\Omega, \mathscr{A}, \mu)$ ein Maßraum, und $f, g, f_k\colon \Omega \to \mathcal{X}$ seien Funktionen in einen topologischen Raum $\mathcal{X}$.*

- *Gilt μ-fast-überall $f = g$ auf Ω mit μ-messbarem g, so ist f ebenfalls μ-messbar.*
- *Ist $\mathcal{X}$ sogar metrischer Raum, sind die f_k alle μ-messbar, und tritt μ-fast-überall auf Ω Konvergenz $f_k \underset{k\to\infty}{\longrightarrow} f$ ein, so ist f ebenfalls μ-messbar.*

Beweis. Der Beweis der ersten Aussage ist einfacher als der der zweiten (und bei metrischem $\mathcal{X}$ sogar ein echter Spezialfall). Deshalb wird nur die zweite Aussage behandelt. Für

$$G := \{x \in \Omega \,:\, f_k(x) \text{ konvergiert bei } k \to \infty \text{ gegen } f(x)\}$$

ist nach Voraussetzung $\Omega \setminus G$ eine $\overline{\mu}$-Nullmenge und damit insbesondere $G \in \mathcal{M}_\mu$. Sei nun O offen in $\mathcal{X}$. Für $x \in G$ gilt dann

$$f(x) \in O \iff \exists i, k \in \mathbb{N}\colon \forall \ell \in \mathbb{N}_{\geq k}\colon \operatorname{dist}(f_\ell(x), \mathcal{X} \setminus O) > \tfrac{1}{i}\,,$$

und dies lässt sich zu

$$G \cap f^{-1}(O) = G \cap \bigcup_{i=1}^\infty \bigcup_{k=1}^\infty \bigcap_{\ell=k}^\infty f_\ell^{-1}(\underbrace{\{y \in \mathcal{X} \,:\, \operatorname{dist}(y, \mathcal{X} \setminus O) > \tfrac{1}{i}\}}_{\text{offen in } \mathcal{X}})$$

umschreiben. Wegen der μ-Messbarkeit der f_ℓ liest man $G \cap f^{-1}(O) \in \mathcal{M}_\mu$ ab. Als Teilmenge einer μ-Nullmenge ist auch $f^{-1}(O) \setminus G \in \mathcal{M}_\mu$ messbar, daher ist $f^{-1}(O) \in \mathcal{M}_\mu$ für jede offene Menge O in $\mathcal{X}$ gezeigt. Da die offenen Mengen $\mathcal{B}(\mathcal{X})$ erzeugen, folgt die behauptete $(\mathcal{M}_\mu, \mathcal{B}(\mathcal{X}))$-Messbarkeit von f. □

Bemerkung. Allgemeine Aussagen über μ-messbare Funktionen gelten analog für $\mathscr{A}$-messbare Funktionen bezüglich beliebiger σ-Algebren $\mathscr{A}$, wenn man μ-fast-überall bestehende durch überall bestehende Eigenschaften ersetzt. Formal kann man dies durch Anwendung von „μ-Resultaten“ auf ein Maß μ mit $\mathcal{M}_\mu = \mathscr{A}$ einsehen; konkret kann man dazu $\mu = \xi|_{\mathscr{A}}$ mit dem Zählmaß ξ wählen (bei dem es ja keine-nichttrivialen Nullmengen gibt und deshalb tatsächlich $\mathcal{M}_\mu = \mathscr{A}$ ist). Insbesondere ergibt sich somit aus der Proposition, dass punktweise Limites von Borel-Funktionen (mit Werten in einem metrischen Raum) wieder Borelsch sind.

Aus den früheren Folgerungen und der Proposition ergibt sich, dass alle praktisch relevanten Operationen mit Folgen von Borel-Funktionen beziehungsweise μ-messbaren Funktionen wieder Borel-Funktionen beziehungsweise μ-messbare Funktionen ergeben. Beispielsweise sieht man das für die Operationen sup und lim sup ein, indem man sie folgendermaßen in punktweise Limites umschreibt:

$$\sup_{k\in\mathbb{N}} f_k = \lim_{k\to\infty} \max\{f_1, f_2, \ldots, f_k\} \qquad \text{und} \qquad \limsup_{k\to\infty} f_k = \lim_{m\to\infty} \left(\sup_{k\in\mathbb{N}_{\geq m}} f_k \right).$$

Analog begründet man die Erhaltung von Messbarkeit bei inf und lim inf sowie bei Reihen.

Als Nächstes wird eine Definition messbarer Treppenfunktionen gegeben und die (monotone) Approximation durch solche diskutiert.

Definition (Treppenfunktionen). *Sei $(\Omega, \mathscr{A})$ ein Messraum, und sei $g\colon \Omega \to \overline{\mathbb{R}}$ eine $\mathscr{A}$-messbare Funktion. Man nennt g eine **($\mathscr{A}$-) Treppenfunktion**, wenn g höchstens abzählbar viele Werte annimmt. Nimmt g nur endlich viele Werte an, ist also endliche Linearkombination charakteristischer Funktionen, so heißt g eine ($\mathscr{A}$-) Treppenfunktion mit endlich vielen Stufen.*

Verfahren (Standard-Ausdehnungsprozedur der Integrationstheorie). *Eine sehr häufige Vorgehensweise der Integrationstheorie basiert auf der Verwendung von Treppenfunktionen und verläuft wie folgt: Man beweist eine Aussage über $\mathscr{A}$-messbare Funktionen (oft im Zusammenhang mit Integrierbarkeit und Integralen und bei linearem Auftreten der Funktion), indem man sie schrittweise verifiziert,*

- *erst für charakteristische Funktionen $\mathbb{1}_A$ zu $A \in \mathscr{A}$,*
- *dann für $\mathscr{A}$-Treppenfunktionen $g\colon \Omega \to \mathbb{R}_{\geq 0}$ (mit endlich vielen Stufen), die man als endliche oder unendliche Linearkombinationen charakteristische Funktionen darstellt,*
- *dann für nichtnegative $\mathscr{A}$-messbare Funktionen $f\colon \Omega \to [0, \infty]$ mit dem nächsten Lemma,*
- *schließlich für messbare Funktionen $f\colon \Omega \to \overline{\mathbb{R}}$ durch Zerlegung $f = f_+ - f_-$ in Positiv- und Negativteil,*
- *und eventuell noch für $\mathbb{R}^M$-wertige messbare Funktionen $f\colon \Omega \to \mathbb{R}^M$ durch Betrachtung der Komponentenfunktionen.*

Auch im weiteren Verlauf der Vorlesung wird noch einige Male so vorgegangen. Der Übergang von Treppenfunktionen zu nichtnegativen Funktionen basiert dabei oft auf:

Lemma (zur Approximation nichtnegativer, messbarer Funktionen durch Treppenfunktionen)**.** *Sei $(\Omega, \mathscr{A})$ ein Messraum, und sei $f\colon \Omega \to [0, \infty]$ eine nichtnegative, $\mathscr{A}$-messbare Funktion. Dann gibt es eine Folge von $\mathscr{A}$-Treppenfunktionen $g_k\colon \Omega \to \mathbb{R}_{\geq 0}$ mit endlich vielen Stufen, so dass $g_1 \leq g_2 \leq g_3 \leq \ldots$ gilt und $g_k \underset{k\to\infty}{\longrightarrow} f$ punktweise auf Ω konvergiert.*

Zum *Beweis* wählt man die g_k konstant auf gewissen Zwischenniveaumengen von f. Im Einzelnen gibt es etliche Vorgehensweisen, eine solche ist Thema der Übungen.

Als Beispiel für eine Anwendung der Standard-Ausdehnungsprozedur und des Lemmas wird eine erwähnenswerte Beobachtung über den Zusammenhang zwischen Lebesgue- und Borel-Messbarkeit festgehalten. Tatsächlich unterscheiden sich diese Konzepte — wie ja auch $\mathcal{M}^N$ und $\mathcal{B}(\mathbb{R}^N)$ — nur um Nullmengen:

Satz. *Sei $(\Omega, \mathscr{A}, \mu)$ ein Maßraum. Eine Funktion $f\colon \Omega \to \overline{\mathbb{R}}$ ist genau dann μ-messbar, wenn es ein $\mathscr{A}$-messbares $g\colon \Omega \to \overline{\mathbb{R}}$ gibt, so dass μ-fast-überall $g = f$ auf Ω gilt. Man nennt ein solches g einen $\mathscr{A}$-messbaren Repräsentanten von f.*

Insbesondere ist $f\colon A \to \overline{\mathbb{R}}$ mit $A \in \mathcal{M}^N$ genau dann $\mathscr{L}^N$-messbar, wenn es eine Borel-Funktion $g\colon A \to \overline{\mathbb{R}}$ gibt, so dass $\mathscr{L}^N$-fast-überall $g = f$ auf A gilt. Man nennt ein solches g einen Borel-Repräsentanten von f.

Beweisskizze. Die Implikation „ $\Longleftarrow$ “ folgt direkt aus der letzten Proposition. Die umgekehrte Implikation „ $\Longrightarrow$ “ beruht darauf, dass jede $\mathcal{M}_\mu$-messbare Menge sich nach Definition der Vervollständigung nur um eine $\overline{\mu}$-Nullmenge von einer $\mathscr{A}$-messbaren Menge unterscheidet. Aus dieser Beobachtung ergibt sich Behauptung für charakteristische Funktionen und damit auch für Treppenfunktionen; man vergleiche mit Fußnote 24 im nächsten Abschnitt. Für nichtnegatives f wendet man jetzt das vorige Lemma (mit $\mathcal{M}_\mu$ anstelle von $\mathscr{A}$) an und macht die so erhaltenen $\mathcal{M}_\mu$-Treppenfunktionen g_k durch Modifikation auf $\overline{\mu}$-Nullmengen zu $\mathscr{A}$-messbaren $\widetilde{g}_k$ mit $0 \le \widetilde{g}_1 \le \widetilde{g}_2 \le \widetilde{g}_3 \le \ldots$ auf Ω. Wegen Monotonie konvergieren die $\widetilde{g}_k$ punktweise gegen eine $[0,\infty]$-wertige Funktion g, die μ-fast-überall auf Ω mit f übereinstimmt. Aufgrund der letzten Proposition ist g als Limes der $\widetilde{g}_k$ ebenfalls $\mathscr{A}$-messbar, und damit ist die Behauptung für nichtnegatives f gezeigt. Für allgemeines f folgt sie problemlos durch Zerlegung $f = f_+ - f_-$. □

Wie schon zuvor diskutiert, ist Stetigkeit unter der schwachen Annahme $\mathcal{B}(\Omega) \subset \mathcal{M}_\mu$ eine deutlich stärkere Anforderung an eine Funktion $f \colon \Omega \to \mathcal{X}$ als μ-Messbarkeit, und gleichmäßige Konvergenz ist eine deutlich stärkere Anforderung an eine Funktionenfolge als ihre Konvergenz μ-fast-überall. Nichtsdestotrotz stellen die folgenden Sätze von Lusin und Egoroff verblüffend enge Beziehungen zwischen diesen Begriffsbildungen her: Grob gesprochen, ist gemäß dem Satz von Lusin **jede μ-messbare Funktion *fast* stetig**, und gemäß dem Satz von Egoroff ist **jede μ-fast-überall konvergente Folge *fast* gleichmäßig konvergent**. Entscheidend ist natürlich das Wort „fast“, das hier „bis auf Mengen kleinen Maßes“ bedeutet. Eine präzise Fassung der Sätze folgt:

Satz (**von Lusin**, ∼1912)**.** *Seien Ω eine offene oder abgeschlossene Teilmenge von $\mathbb{R}^N$ und $\mu = \mathscr{L}^N$, oder sei allgemeiner μ ein Radon-Maß (dieser Begriff wird in Abschnitt 2.12 definiert) auf einem für solche zulässigen topologischen Raum Ω. Sei zudem $\mathcal{X}$ ein metrischer Raum, der eine abzählbare dichte Teilmenge enthält, und sei $f \colon \Omega \to \mathcal{X}$ eine μ-messbare Funktion. Dann gibt es zu jedem $\delta \in \mathbb{R}_{>0}$ eine abgeschlossene Teilmenge A_δ von Ω mit*

$$f|_{A_\delta} \text{ stetig} \qquad\qquad \text{und} \qquad\qquad \mu(\Omega \setminus A_\delta) < \delta\,.$$

Im Fall $\mu(\Omega) < \infty$ kann A_δ zudem kompakt gewählt werden.

Bemerkung (zum Satz von Lusin)**.** Die Stetigkeit von $f|_{A_\delta}$ im Satz von Lusin impliziert *nicht*, dass die auf ganz Ω definierte Funktion f an irgendeiner Stelle in Ω oder A_δ stetig sein muss. Dies sieht man schon für $\mu = \mathscr{L}^N$ am Beispiel der Borel-Funktion $f = \mathbb{1}_{\mathbb{Q}}$. Im Fall $\mathcal{X} = \mathbb{R}^M$ kann man aber tatsächlich (was hier nicht bewiesen wird) $f|_{A_\delta}$ stetig auf ganz Ω fortsetzen, d.h. stetige Funktionen $g_\delta \colon \Omega \to \mathbb{R}^M$ mit $g_\delta = f$ auf A_δ konstruieren.

Beweisskizze. Zuerst wird der Fall $\mu(\Omega) < \infty$ behandelt. Wegen der Abzählbarkeitsvoraussetzung lässt sich $\mathcal{X}$ für jedes $i \in \mathbb{N}$ durch abzählbar viele Kugeln vom Radius $\frac{1}{2i}$ überdecken. Aus der Kugelüberdeckung ergibt sich eine Zerlegung von $\mathcal{X}$ in abzählbar viele *disjunkte* Borel-Mengen $Y_{i,1}, Y_{i,2}, Y_{i,3}, \ldots$ vom Durchmesser $\le \frac{1}{i}$. Die Urbilder $B_{i,j} := f^{-1}(Y_{i,j}) \in \mathcal{M}_\mu$ bilden (weiter bei festem i) eine *disjunkte* Zerlegung von Ω, und gemäß einer Regularitätseigenschaft von Radon-Maßen μ (die in den Übungen im Wesentlichen schon bewiesen wurde) findet man abgeschlossene Teilmengen $A_{i,j}$ von Ω mit $\overline{\mu}(B_{i,j} \setminus A_{i,j}) < 2^{-i-j}\delta$. Da Ω als abzählbare Vereinigung kompakter Mengen geschrieben werden kann (für offenes oder abgeschlossenes $\Omega \subset \mathbb{R}^N$ klar und bei Radon-Maßen später leicht einzusehen), lässt sich ohne Einschränkung annehmen, dass die $A_{i,j}$ kompakt sind. Unter Verwendung der Endlichkeitsvoraussetzung $\mu(\Omega) < \infty$ folgt $\lim_{k\to\infty} \mu\big(\Omega \setminus \bigcup_{j=1}^k A_{i,j}\big) = \mu\big(\Omega \setminus \bigcup_{j=1}^\infty A_{i,j}\big) = \sum_{j=1}^\infty \overline{\mu}(B_{i,j} \setminus A_{i,j}) < 2^{-i}\delta$, und daher gibt es zu jedem $i \in \mathbb{N}$ ein $k_i \in \mathbb{N}$ mit $\mu\big(\Omega \setminus \bigcup_{j=1}^{k_i} A_{i,j}\big) < 2^{-i}\delta$. Für

$$A_\delta := \bigcap_{i=1}^\infty \bigcup_{j=1}^{k_i} A_{i,j}$$

folgt $\mu(\Omega\setminus A_\delta) \le \sum_{i=1}^\infty \mu\big(\Omega \setminus \bigcup_{j=1}^{k_i} A_{i,j}\big) < \delta$. Jetzt betrachtet man für jedes $i \in \mathbb{N}$ eine Funktion f_i auf $\bigcup_{j=1}^{k_i} A_{i,j}$, die auf den $A_{i,j} \neq \emptyset$ jeweils konstant einen beliebig gewählten Wert in $Y_{i,j}$ annimmt. Nach Konstruktion sind die f_i stetig und konvergieren gleichmäßig auf der kompakten Menge A_δ gegen f. Nach dem Satz über die Stetigkeit der Grenzfunktion ist $f|_{A_\delta}$ stetig.

Zu Behandlung des Falls $\mu(\Omega) = \infty$ überlegt man sich, dass sich $\Omega = \bigcup_{\ell=1}^\infty \Omega_\ell$ mit μ-endlichen, offenen Mengen Ω_ℓ schreiben lässt, wobei jeder Punkt von Ω in höchstens zwei der Mengen Ω_ℓ enthalten ist (für $\mu = \mathscr{L}^N$ klar; Radon-Maße μ haben diese Eigenschaft allgemein). Nach dem schon Bewiesenen gibt es dann kompakte $A_\ell \subset \Omega_\ell$ mit $\mu(\Omega_\ell \setminus A_\ell) < 2^{-\ell}\delta$, so dass $f|_{A_\ell}$ stetig ist. Für $A_\delta := \bigcup_{\ell=1}^\infty A_\ell$ folgt $\mu(\Omega \setminus A_\delta) < \delta$, und wegen der Bedingung, dass jeder Punkt von Ω zu höchstens zwei Ω_ℓ gehört, ist $A_\delta := \bigcup_{\ell=1}^\infty A_\ell$ abgeschlossen sowie $f|_{A_\delta}$ stetig. □

Satz (von Egoroff/Jegorow[21], ∼1911). *Sei $(\Omega, \mathscr{A}, \mu)$ ein Maßraum mit $\mu(\Omega) < \infty$, sei $\mathcal{X}$ ein metrischer Raum, und die Funktionen $f_k \colon \Omega \to \mathcal{X}$ seien μ-messbar. Liegt μ-fast-überall auf Ω Konvergenz $f_k \underset{k\to\infty}{\longrightarrow} f$ vor, so gibt es zu jedem $\delta > 0$ ein $A_\delta \in \mathscr{A}$ mit*

$$f_k \underset{k\to\infty}{\longrightarrow} f \textbf{ \textit{gleichmäßig} } \text{auf } A_\delta \qquad \text{und} \qquad \mu(\Omega \setminus A_\delta) < \delta\,.$$

Beweis. Gemäß der Proposition dieses Abschnitts überträgt sich die μ-Messbarkeit der f_k auf f. Außerdem wird ohne Einschränkung angenommen, dass $(\Omega, \mathscr{A}, \mu)$ vollständig ist. Für $i, j \in \mathbb{N}$ sei

$$A_{i,j} := \bigcap_{\ell=j}^\infty \big\{x \in \Omega \,:\, \mathrm{d}_\mathcal{X}(f_\ell(x), f(x)) \le \tfrac{1}{i}\big\} \in \mathscr{A}\,.$$

Dann gilt stets $A_{i,j} \subset A_{i,j+1}$, die Menge, auf der $f_k \to f$ punktweise konvergiert, ist genau $\bigcap_{i=1}^\infty \bigcup_{j=1}^\infty A_{i,j}$, und deshalb ist $\Omega \setminus \bigcup_{j=1}^\infty A_{i,j}$ für jedes $i \in \mathbb{N}$ eine μ-Nullmenge. Wegen $\mu(\Omega) < \infty$ folgt $\lim_{j\to\infty} \mu(\Omega \setminus A_{i,j}) = 0$, und es gibt zu jedem $i \in \mathbb{N}$ ein $j_i \in \mathbb{N}$ mit $\mu(\Omega \setminus A_{i,j_i}) < 2^{-i}\delta$. Für

$$A_\delta := \bigcap_{i=1}^\infty A_{i,j_i} \in \mathscr{A}$$

erhält man $\mu(\Omega \setminus A_\delta) < \delta$, und nach Konstruktion ist die Bedingung

$$\mathrm{d}_\mathcal{X}(f_\ell, f) \le \tfrac{1}{i} \qquad \text{auf } A_\delta$$

bei gegebenem $i \in \mathbb{N}$ für alle $\ell \in \mathbb{N}_{\ge j_i}$ erfüllt. Insbesondere liegt gleichmäßige Konvergenz $f_k \to f$ auf $\Omega \setminus A_\delta$ vor. □

2.6 Das Maßintegral

In diesem Abschnitt wird die Integration messbarer Funktionen bezüglich beliebiger Maße eingeführt und mit diesem allgemeinen Integralbegriff das ein zentrales Ziel des aktuellen Vorlesungskapitels erreicht. Im Vordergrund steht dabei stets das Lebesgue-Integral, d.h. die Integration bezüglich des Lebesgue-Maßes $\mathscr{L}^N$.

Wie die Definition des Riemann-Integrals baut auch die Definition des allgemeinen Maßintegrals auf elementaren Integralen von Treppenfunktionen auf — mit dem entscheidenden Unterschied, dass nun viel allgemeinere Treppenfunktionen erlaubt werden (können):

[21] Es handelt sich um zwei verschiedene Transkriptionen desselben russischen Namens. Beide sind gebräuchlich.

Definition (elementares Integral von Treppenfunktionen). *Sei $(\Omega, \mathscr{A}, \mu)$ ein Maßraum. Das* ***(elementare $\boldsymbol{\mu}$-)Integral*** *einer $\mathscr{A}$-Treppenfunktion $g\colon \Omega \to \overline{\mathbb{R}}$ über Ω wird definiert als*

$$\int_\Omega g \,\mathrm{d}\mu := \sum_{t\in\overline{\mathbb{R}}} t\,\mu\big(g^{-1}(t)\big) := \sum_{t\in[0,\infty]} t\,\mu(g^{-1}(t)) + \sum_{t\in[-\infty,0]} t\,\mu(g^{-1}(t))$$

(wobei $g^{-1}(t)$ natürlich für das Urbild $g^{-1}(t) := \{x \in \Omega : g(x) = t\}$ steht).

Bemerkungen (zur Verständnis der Summen in der vorigen Definition)**.**

(1) Da $\mathscr{A}$-Treppenfunktionen höchstens abzählbar viele Werte annehmen, sind höchstens abzählbar viele Summanden der obigen Summen von Null verschieden.

(2) Die beiden Teilsummen $\sum_{t\in[0,\infty]} t\,\mu(g^{-1}(t)) \in [0,\infty]$ und $\sum_{t\in[-\infty,0]} t\,\mu(g^{-1}(t)) \in [-\infty,0]$ (mit eventuell unendlichen Werten) sind stets definiert. Das Integral und die Gesamtsumme $\sum_{t\in\overline{\mathbb{R}}} t\,\mu(g^{-1}(t))$ werden nur dann als definiert betrachtet, wenn bei der Addition der beiden Teilsummen nicht der Ausdruck $\infty-\infty$ auftritt.

(3) Insgesamt hat damit $\int_\Omega g\,\mathrm{d}\mu$ entweder einen Wert in $\overline{\mathbb{R}}$ oder ist undefiniert vom Typ $\infty-\infty$.

Definition (Integrale, Integrierbarkeit, Summierbarkeit). *Sei $(\Omega, \mathscr{A}, \mu)$ ein Maßraum, und sei $f\colon \Omega \to \overline{\mathbb{R}}$ eine beliebige Funktion. Eine $\mathscr{A}$-Treppenfunktion $g\colon \Omega \to \overline{\mathbb{R}}$ heißt* ***($\boldsymbol{\mu}$-) Oberfunktion*** *zu f, wenn $g \ge f$ auf Ω gilt und das elementare μ-Integral $\int_\Omega g\,\mathrm{d}\mu$ definiert ist. Analog heißt g eine* ***($\boldsymbol{\mu}$-) Unterfunktion*** *zu f, wenn $g \le f$ auf Ω gilt und $\int_\Omega g\,\mathrm{d}\mu$ definiert ist. Man nennt*

$$\int^*_\Omega f\,\mathrm{d}\mu := \inf\left\{\int_\Omega g\,\mathrm{d}\mu \,:\, g \text{ ist } \mu\text{-Oberfunktion zu } f\right\} \in \overline{\mathbb{R}}$$

(mit Konvention $\inf\emptyset = \infty$) das ***($\boldsymbol{\mu}$-) Oberintegral*** *von f (über Ω) und*

$$\int_{*\Omega} f\,\mathrm{d}\mu := \sup\left\{\int_\Omega g\,\mathrm{d}\mu \,:\, g \text{ ist } \mu\text{-Unterfunktion zu } f\right\} \in \overline{\mathbb{R}}$$

(mit Konvention $\sup\emptyset = -\infty$) das ***($\boldsymbol{\mu}$-) Unterintegral*** *von f (über Ω). Aufbauend hierauf bezeichnet man f als* ***($\boldsymbol{\mu}$-)integrierbar*** *(über Ω), wenn f zumindest μ-messbar*[22] *ist und ${}^*\!\int_\Omega f\,\mathrm{d}\mu$ mit ${}_*\!\int_\Omega f\,\mathrm{d}\mu$ übereinstimmt, und in diesem Fall heißt der gemeinsame Wert*

$$\int_\Omega f\,\mathrm{d}\mu := \int^*_\Omega f\,\mathrm{d}\mu = \int_{*\Omega} f\,\mathrm{d}\mu \in \overline{\mathbb{R}}$$

das ***($\boldsymbol{\mu}$-)Integral*** *von f (über Ω). Schließlich heißt ein (μ-)integrierbares f mit endlichem Wert des Integrals auch* ***($\boldsymbol{\mu}$-)summierbar***[23]

[22] In vielen Lehrbüchern wird an dieser Stelle $\mathscr{A}$-Messbarkeit von f vorausgesetzt, was für nicht-vollständige Maße wegen $\mathscr{A} \subsetneqq \mathcal{M}_\mu$ eine stärkere Forderung ist und zu einem weniger allgemeinen Integralbegriff führt. Meist ist dieser Unterschied aber unwesentlich, da man zur Vervollständigung oder dem aus Abschnitt 2.5 bekannten $\mathscr{A}$-messbaren Repräsentanten einer μ-messbaren Funktion übergehen kann.

[23] Der Begriff der *Integrierbarkeit* wird in Teilen der Literatur nur für solche Funktionen verwendet, die zu einem *endlichen* Integralwert führen. Hier werden solche Funktionen aber entsprechend der Definition *summierbar* genannt, wohingegen bei *integrierbaren* Funktionen auch unendliche Werte des Integrals zugelassen werden.

Bemerkungen (zu Integrierbarkeit und Integralen).

(1) Das Oberintegral und das Unterintegral existieren für beliebige Funktionen f und erfüllen
$$\int^{*}_{\Omega} f \,\mathrm{d}\mu \geq {}_{*}\!\!\int_{\Omega} f \,\mathrm{d}\mu$$
in $\overline{\mathbb{R}}$. Um Integrierbarkeit eines μ-messbaren f zu zeigen, reicht also der Nachweis der umgekehrten Ungleichung.

(2) Treppenfunktionen (mit definiertem elementaren Integral) sind Ober- und Unterfunktionen zu sich selbst. Deshalb sind sie integrierbar, und ihr Integral stimmt mit ihrem elementaren Integral überein — was die Verwendung identischer Notation für elementare Integrale und Integrale erst rechtfertigt.

(3) Insbesondere sind **charakteristische Funktionen** $\mathbb{1}_A$ mit $A \in \mathscr{A}$ stets μ-integrierbar mit
$$\boxed{\int_{\Omega} \mathbb{1}_A \,\mathrm{d}\mu = \mu(A)\,.}$$
Allgemeiner gelten $\int_{\Omega} \mathbb{1}_A \,\mathrm{d}\mu = \overline{\mu}(A)$ für $A \in \mathcal{M}_\mu$ und $\int^{*}_{\Omega} \mathbb{1}_A \,\mathrm{d}\mu = \mu^*(A)$ für alle $A \in \mathcal{P}(\Omega)$.

(4) μ-Integrierbarkeit und μ-(Ober-/Unter-)Integrale stimmen mit den entsprechenden Konzepten für die Vervollständigung $\overline{\mu}$ überein[24]. Daher kann man sich stets auf die Behandlung von Integralen bezüglich vollständigen Maßen beschränken.

Notationen (für Integrale).

(1) *Ist $X \in \mathscr{A}$ eine messbare Teilmenge von Ω und ist f auf einem Definitionsbereich D mit $X \subset D \subset \Omega$ definiert, so vereinbart man für die **Integration über** die **Teilmenge** X die abkürzende Notation*
$$\int_X f \,\mathrm{d}\mu$$
für das eigentlich mit den Einschränkungen von f und μ zu bildende Integral $\int_X f|_X \,\mathrm{d}\mu|_{\mathscr{A}|X}$. Entsprechend spricht man von μ-Integrier-/Summierbarkeit von f über die Teilmenge X.

(2) *Gleichbedeutend mit $\int_X f \,\mathrm{d}\mu$ werden die Notationen $\int_X f(x)\,\mathrm{d}\mu(x)$ und $\int_X f(x)\,\mathrm{d}\mu\, x$ verwendet. Bei diesen wird der Integrand nicht als Funktion f, sondern durch einen Funktions*term *$f(x)$ angegeben und die (Integrations-) Variable x explizit bezeichnet.*

(3) *Im wichtigsten Fall von Lebesgue-Integralen, also Integralen bezüglich des Lebesgue-Maßes $\mathscr{L}^N$, wird manchmal nur kurz ‚$\mathrm{d}x$' statt ‚$\mathrm{d}\mathscr{L}^N x$' notiert.*

[24] *Beweis.* Sei $f\colon \Omega \to \overline{\mathbb{R}}$ beliebig. Da μ-Oberfunktionen auch $\overline{\mu}$-Oberfunktionen mit gleichem Wert des elementaren Integrals sind, gilt $\int^{*}_{\Omega} f \,\mathrm{d}\overline{\mu} \leq \int^{*}_{\Omega} f \,\mathrm{d}\mu$. Um die umgekehrte Ungleichung herzuleiten, betrachtet man eine $\overline{\mu}$-Oberfunktionen g zu f. Sind $t_1, t_2, t_3, \ldots \in \overline{\mathbb{R}}$ die abzählbar vielen Werte, die g annimmt, und $B_i := g^{-1}(t_i) \in \mathcal{M}_\mu$ die zugehörigen Niveaumengen, so gibt es $\mathscr{A}$-messbare $A_i \subset B_i$ mit $\mu(A_i) = \overline{\mu}(B_i)$, und $M := \Omega \setminus \bigcup_{i=1}^{\infty} A_i \in \mathscr{A}$ ist eine μ-Nullmenge. Folglich ist $\widetilde{g} := \infty \cdot \mathbb{1}_M + \sum_{i=1}^{\infty} t_i \mathbb{1}_{A_i}$ eine μ-Oberfunktion zu f mit
$$\int_{\Omega} \widetilde{g} \,\mathrm{d}\mu = \infty \cdot \mu(M) + \sum_{i=1}^{\infty} t_i \mu(A_i) = \sum_{i=1}^{\infty} t_i \overline{\mu}(B_i) = \int_{\Omega} g \,\mathrm{d}\overline{\mu}\,.$$
Deshalb gilt auch $\int^{*}_{\Omega} f \,\mathrm{d}\mu \leq \int^{*}_{\Omega} f \,\mathrm{d}\overline{\mu}$, und μ-Oberintegrale sind dasselbe wie $\overline{\mu}$-Oberintegrale. Analog oder durch Übergang zu $-f$ verifiziert man Entsprechendes für Unterintegrale. Da μ-Messbarkeit und $\overline{\mu}$-Messbarkeit per Definition gleichbedeutend sind, folgt die Gleichheit der Integrale und der Integrierbarkeitsbegriffe. □

Grundeigenschaften (von Maßintegralen)**.** Seien μ, μ_i Maße auf $(\Omega, \mathscr{A})$, seien $r, s \in \mathbb{R}$ und $\omega_i \in [0, \infty]$ Parameter, und seien $f, g\colon \Omega \to \overline{\mathbb{R}}$ beliebige Funktionen.

(1) **Identität:** Gilt μ-fast-überall $f = g$ auf Ω mit μ-intergrierbarem g, so ist f ebenfalls μ-integrierbar mit

$$\int_\Omega f \,\mathrm{d}\mu = \int_\Omega g \,\mathrm{d}\mu\,.$$

(2) **Monotonie im Integranden:** Sind f und g beide μ-integrierbar und gilt μ-fast-überall $f \leq g$ auf Ω, so folgt

$$\int_\Omega f \,\mathrm{d}\mu \leq \int_\Omega g \,\mathrm{d}\mu\,.$$

Durch Anwendung auf die Funktionen $-|f| \leq f \leq |f|$ folgt aus dieser Regel die **Dreiecksungleichung**

$$\left|\int_\Omega f \,\mathrm{d}\mu\right| \leq \int_\Omega |f| \,\mathrm{d}\mu$$

für μ-integrierbares f (wobei μ-Integrierbarkeit von $|f|$ in Anbetracht des nächsten Satzes automatisch vorliegt und deshalb nicht explizit vorausgesetzt werden muss).

(3) Sind f, g beide μ-*summierbar* und gilt μ-fast-überall $f \leq g$ auf Ω, so hat man die Äquivalenz

$$\int_\Omega f \,\mathrm{d}\mu = \int_\Omega g \,\mathrm{d}\mu \iff f = g \text{ gilt } \mu\text{-fast-überall auf } \Omega\,.$$

(4) **Linearität im Integranden:** Für μ-integrierbare f und g ist[25] $rf{+}sg$ auch μ-integrierbar mit

$$\int_\Omega (rf{+}sg) \,\mathrm{d}\mu = r\int_\Omega f \,\mathrm{d}\mu + s\int_\Omega g \,\mathrm{d}\mu\,,$$

vorausgesetzt die Summe auf der rechten Seite ist nicht vom Typ $\infty{-}\infty$.

(5) **Linearität im Maß:** Ist I eine beliebige Indexmenge und ist f für alle $i \in I$ stets μ_i-integrierbar und zudem $\mathscr{A}$- oder zumindest $\big(\sum_{i\in I} \omega_i\mu_i\big)$-messbar, so ist f auch $\big(\sum_{i\in I} \omega_i\mu_i\big)$-integrierbar mit

$$\int_\Omega f \,\mathrm{d}\Big(\sum_{i\in I} \omega_i\mu_i\Big) = \sum_{i\in I} \omega_i \int_\Omega f \,\mathrm{d}\mu_i\,,$$

vorausgesetzt die Summe auf der rechten Seite ist nicht vom Typ $\infty{-}\infty$. Ist I höchstens abzählbar[26], so ist die zusätzliche Messbarkeitsvoraussetzung an f hierbei automatisch erfüllt und muss nicht explizit gefordert werden.

[25] Da $r\int_\Omega f\,\mathrm{d}\mu + s\int_\Omega g\,\mathrm{d}\mu$ nach Voraussetzung nicht vom Typ $\infty{-}\infty$ ist, wird einer der beiden Werte ∞ und $-\infty$ sowohl von rf als auch von sg nur auf einer $\overline{\mu}$-Nullmenge angenommen, und insbesondere bilden die $x \in \Omega$, für die $rf(x){+}sg(x)$ undefiniert vom Typ $\infty{-}\infty$ ist, eine $\overline{\mu}$-Nullmenge. Die behauptete Aussage ist gültig, wenn man für diese x die Werte $rf(x){+}sg(x)$ in beliebiger Weise festlegt.

[26] Bei überabzählbarem I allerdings kann man auf die Voraussetzung der $\big(\sum_{i\in I} \omega_i\mu_i\big)$-Messbarkeit von f nicht verzichten. Dies zeigt das Beispiel $I = \Omega = \mathbb{R}$, $\mathscr{A} = \sigma(\{\{x\} : x \in \mathbb{R}\})$, $\mu_x = \delta_x|_{\mathscr{A}}$, $\mu = \sum_{x\in\mathbb{R}} \mu_x = \xi|_{\mathscr{A}}$, in dem $\mathcal{M}_{\mu_x} = \mathcal{P}(\mathbb{R})$ und jede Funktion μ_x-messbar, aber wegen $\mathcal{M}_\mu = \mathscr{A}$ jede charakteristische Funktion eines überabzählbaren $A \subset \mathbb{R}$ mit überabzählbarem Komplement nicht μ-messbar ist.

(6) **σ-Additivität im Integrationsbereich:** Sind $X_1, X_2, X_3, \ldots \in \mathscr{A}$ *disjunkte* messbare Mengen und ist f für alle $i \in \mathbb{N}$ stets μ-integrierbar über X_i, so ist f auch μ-integrierbar über $\bigcup_{i=1}^{\infty} X_i$ mit

$$\int_{\bigcup_{i=1}^{\infty} X_i} f \,\mathrm{d}\mu = \sum_{i\in\mathbb{N}} \int_{X_i} f \,\mathrm{d}\mu \,,$$

vorausgesetzt die Summe auf der rechten Seite ist nicht vom Typ $\infty-\infty$.

Ein einfacher, doch oft nützlicher Spezialfall dieser Eigenschaft ist die Regel

$$\int_\Omega \mathbb{1}_A f \,\mathrm{d}\mu = \int_A f \,\mathrm{d}\mu \qquad \text{für } A \in \mathscr{A} \,.$$

Beweise der Grundeigenschaften. Ohne Einschränkung sei $(\Omega, \mathscr{A}, \mu)$ vollständig.

Zuerst wird die Eigenschaft (4) nachgewiesen, wobei auf Beweise einfacher Multiplikationsregeln verzichtet und nur der Fall $r = s = 1$ behandelt wird. Vorerst seien f und g Treppenfunktionen, für die die Summe der elementaren Integrale $\int_\Omega f \,\mathrm{d}\mu + \int_\Omega g \,\mathrm{d}\mu$ definiert ist. Im Fall, dass μ-fast-überall $f \in \mathbb{R}$ und $g \in \mathbb{R}$ gilt, also unendliche Werte nur auf Nullmengen angenommen werden, ergibt sich mit der Definition des elementaren Integrals und der σ-Additivität von μ (wobei stets alle bis auf abzählbar viele Summanden verschwinden)

$$\begin{aligned}
\int_\Omega (f{+}g) \,\mathrm{d}\mu &= \sum_{t\in\mathbb{R}} t\, \mu((f{+}g)^{-1}(t)) \\
&= \sum_{t\in\mathbb{R}} \sum_{\substack{(\alpha,\beta)\in\mathbb{R}^2 \\ \alpha+\beta=t}} (\alpha{+}\beta)\, \mu(f^{-1}(\alpha) \cap g^{-1}(\beta)) \\
&= \sum_{\alpha\in\mathbb{R}} \alpha \sum_{t\in\mathbb{R}} \mu(f^{-1}(\alpha) \cap g^{-1}(t{-}\alpha)) + \sum_{\beta\in\mathbb{R}} \beta \sum_{t\in\mathbb{R}} \mu(f^{-1}(t{-}\beta) \cap g^{-1}(\beta)) \\
&= \sum_{\alpha\in\mathbb{R}} \alpha\mu(f^{-1}(\alpha)) + \sum_{\beta\in\mathbb{R}} \beta\mu(g^{-1}(\beta)) \\
&= \int_\Omega f \,\mathrm{d}\mu + \int_\Omega g \,\mathrm{d}\mu \,.
\end{aligned}$$

Treten dagegen bei f oder g unendliche Werte auf einer Menge positiven Maßes auf, so ist von den beiden Größen $\mu(f^{-1}(\infty)){+}\mu(g^{-1}(\infty))$ und $\mu(f^{-1}(-\infty)){+}\mu(g^{-1}(-\infty))$ eine positiv und eine gleich Null (denn der Fall zweier Nullen wurde gerade behandelt, und im Fall zweier positiver Zahlen wäre $\int_\Omega f \,\mathrm{d}\mu + \int_\Omega g \,\mathrm{d}\mu$ vom Typ $\infty-\infty$). Die zwei noch möglichen Fälle laufen aber darauf hinaus, dass entweder beide Seiten von $\int_\Omega (f{+}g) \,\mathrm{d}\mu = \int_\Omega f \,\mathrm{d}\mu + \int_\Omega g \,\mathrm{d}\mu$ gleich ∞ oder beide Seiten gleich $-\infty$ sind. Somit gilt die letzte Gleichheit elementarer Integrale für alle Treppenfunktionen f und g, für die die rechte Seite definiert ist. Mit der Definition des Oberintegrals folgt hieraus

$$\int_\Omega^* (f{+}g) \,\mathrm{d}\mu \le \int_\Omega^* f \,\mathrm{d}\mu + \int_\Omega^* g \,\mathrm{d}\mu$$

für beliebige Funktionen f und g, solange die rechte Seite definiert ist. In Kombination mit der entsprechenden Ungleichung für Unterintegrale ergibt sich wie behauptet $\int_\Omega (f{+}g) \,\mathrm{d}\mu = \int_\Omega f \,\mathrm{d}\mu + \int_\Omega g \,\mathrm{d}\mu$.

Zum Beweis der Eigenschaft (2) sei $M := \{x \in \Omega : f(x) > g(x)\} \in \mathscr{A}$. Da μ-fast-überall $f \le g$ gilt, ist M eine μ-Nullmenge. Zu einer beliebigen Oberfunktion h zu g bildet man nun die Treppenfunktion $\widetilde{h}$ mit $\widetilde{h} = h$ auf $\Omega \setminus M$ und $\widetilde{h} \equiv \infty$ auf M. Dann ist $\widetilde{h}$ eine Oberfunktion zu f mit $\int_\Omega \widetilde{h}\,\mathrm{d}\mu = \int_\Omega h\,\mathrm{d}\mu$. Insgesamt folgt aus dieser Überlegung ${}^*\!\!\int_\Omega f\,\mathrm{d}\mu \le {}^*\!\!\int_\Omega g\,\mathrm{d}\mu$, und wegen der vorausgesetzten Integrierbarkeit von f und g können die Oberintegrale in der letzten Ungleichung durch Integrale ersetzt werden. Damit ist die Eigenschaft (2) verifiziert.

In der Situation (1) ist μ-Messbarkeit von f bereits bekannt (Proposition aus Abschnitt 2.5). Die Eigenschaft (1) ergibt sich dann aus der beim Nachweis der Eigenschaft (2) verwendeten Ungleichung für Oberintegrale und der analogen Ungleichung für Unterintegrale.

Schließlich wird die Äquivalenz aus (3) gezeigt: Die Rück-Implikation ist gemäß der Eigenschaft (1) klar. Für die Hin-Implikation betrachtet man $M_k := \{x \in \Omega : f(x)+\frac{1}{k} \le g(x)\} \in \mathscr{A}$ für $k \in \mathbb{N}$. Dann gilt μ-fast-überall $g \ge f+\frac{1}{k}\mathbb{1}_{M_k}$ auf Ω, und gemäß den schon bekannten Regeln (2) und (4) folgt

$$\int_\Omega g\,\mathrm{d}\mu \ge \int_\Omega \big(f+\tfrac{1}{k}\mathbb{1}_{M_k}\big)\,\mathrm{d}\mu = \int_\Omega f\,\mathrm{d}\mu + \tfrac{1}{k}\mu(M_k)\,.$$

Im Fall $\int_\Omega f\,\mathrm{d}\mu = \int_\Omega g\,\mathrm{d}\mu$ muss folglich — die Integrale haben wegen der Summierbarkeitsvoraussetzung ja einen endlichen Wert — $\mu(M_k) = 0$ sein. Damit ist $\bigcup_{k=1}^\infty M_k = \{x \in \Omega : f(x) < g(x)\}$ eine μ-Nullmenge, und $f \ge g$ gilt μ-fast-überall auf Ω. Da die umgekehrte Ungleichung per Voraussetzung gilt, stimmen f und g sogar μ-fast-überall auf Ω überein. Also ist die Eigenschaft (3) gezeigt.

Bei der Behandlung der Eigenschaft (5) wird wieder auf den Nachweis einer einfachen Multiplikationsregel verzichtet und nur der Fall $\omega_i = 1$ für alle $i \in I$ behandelt. Sei zunächst f eine $\mathscr{A}$-Treppenfunktion, für die bei $\sum_{i\in I}\int_\Omega f\,\mathrm{d}\mu_i$ alle (elementaren) Integrale existieren und die Summe nicht vom Typ $\infty-\infty$ ist. Dann ergibt sich

$$\int_\Omega f\,\mathrm{d}\Big(\sum_{i\in I}\mu_i\Big) = \sum_{t\in\overline{\mathbb{R}}} t \sum_{i\in I}\mu_i\big(f^{-1}(t)\big) = \sum_{i\in I}\sum_{t\in\overline{\mathbb{R}}} t\,\mu_i\big(f^{-1}(t)\big) = \sum_{i\in I}\int_\Omega f\,\mathrm{d}\mu_i$$

samt Existenz des (elementaren) Integrals auf der linken Seite einfach durch Vertauschung der Summen, die wegen des Ausschlusses von „$\infty-\infty$“ sinnvoll und erlaubt ist. Als Nächstes wird ein allgemeines μ_i- und $\big(\sum_{i\in I}\mu_i\big)$-messbares f, weiterhin mit wohldefinierten Integralen und wohldefinierter Summe in $\sum_{i\in I}\int_\Omega f\,\mathrm{d}\mu_i$, behandelt. Ist die Indexmenge I endlich und ist für $i \in I$ jeweils g_i eine beliebige Oberfunktion zu f mit $\int_\Omega g_i\,\mathrm{d}\mu_i < \infty$, so ist auch $g := \min_{i\in I} g_i$ eine zulässige Oberfunktion zu f mit

$${}^*\!\!\int_\Omega f\,\mathrm{d}\Big(\sum_{i\in I}\mu_i\Big) \le \int_\Omega g\,\mathrm{d}\Big(\sum_{i\in I}\mu_i\Big) = \sum_{i\in I}\int_\Omega g\,\mathrm{d}\mu_i \le \sum_{i\in I}\int_\Omega g_i\,\mathrm{d}\mu_i\,.$$

Infimumbildung bezüglich der endlich vielen g_i ergibt wegen der Existenz von $\int_\Omega f\,\mathrm{d}\mu_i$ dann

$${}^*\!\!\int_\Omega f\,\mathrm{d}\Big(\sum_{i\in I}\mu_i\Big) \le \sum_{i\in I}\int_\Omega f\,\mathrm{d}\mu_i\,.$$

In Kombination mit der entsprechenden Ungleichung für das Unterintegral folgen $\big(\sum_{i\in I}\mu_i\big)$-Integrierbarkeit von f und die behauptete Gleichheit. Bei unendlichem I scheitert die vorausgehende Argumentation allerdings daran, dass $\inf_{i\in I} g_i$ keine Treppenfunktion mehr sein muss. Mit einem alternativen Vorgehen, das im Vorgriff auf den nächsten Satz bereits die Existenz von

$\int_\Omega f_\pm \,\mathrm{d}\big(\sum_{i\in I}\mu_i\big)$ für die nichtnegativen, $\big(\sum_{i\in I}\mu_i\big)$-messbaren Funktionen $f_\pm$ verwendet, kann man aber auch unendliche I einschließen. Hierzu beobachtet man

$$\int_\Omega g\,\mathrm{d}\Big(\sum_{i\in I}\mu_i\Big) = \sum_{i\in I}\int_\Omega g\,\mathrm{d}\mu_i \le \sum_{i\in I}\int_\Omega f_\pm\,\mathrm{d}\mu_i$$

für alle $\big(\sum_{i\in I}\mu_i\big)$-Unterfunktionen g zu $f_\pm$ und erhält daraus durch Supremumbildung bezüglich g die Ungleichungen

$$\int_\Omega f_\pm\,\mathrm{d}\Big(\sum_{i\in I}\mu_i\Big) \le \sum_{i\in I}\int_\Omega f_\pm\,\mathrm{d}\mu_i\,.$$

Durch Argumentation mit Oberfunktionen ergeben sich auch die umgekehrten Ungleichungen, also insgesamt $\int_\Omega f_\pm\,\mathrm{d}\big(\sum_{i\in I}\mu_i\big) = \sum_{i\in I}\int_\Omega f_\pm\,\mathrm{d}\mu_i$. Die Eigenschaft (4) für $f=f_+-f_-$ liefert dann auch allgemein $\big(\sum_{i\in I}\mu_i\big)$-Integrierbarkeit von f mit $\int_\Omega f\,\mathrm{d}\big(\sum_{i\in I}\mu_i\big) = \sum_{i\in I}\int_\Omega f\,\mathrm{d}\mu_i$.

Nun wird noch die ergänzende Behauptung in (5) zu automatischer $\big(\sum_{i\in I}\omega_i\mu_i\big)$-Messbarkeit bei höchstens abzählbarem I hergeleitet. Sei dazu f für alle $i\in I$ jeweils μ_i-messbar, und sei $B\in\mathcal{B}(\overline{\mathbb{R}})$. Per Definition gilt dann $f^{-1}(B)\in\mathcal{M}_{\mu_i}$, also $f^{-1}(B)=A_i\cup T_i$ mit $A_i\in\mathscr{A}$, $T_i\subset M_i\in\mathscr{A}$, $\mu_i(M_i)=0$, für alle $i\in I$. Mit der Abzählbarkeit von I folgt $f^{-1}(B)=\big(\bigcup_{i\in I}A_i\big)\cup\big(\bigcap_{i\in I}T_i\big)$ mit $\bigcup_{i\in I}A_i\in\mathscr{A}$, $\bigcap_{i\in I}T_i\subset\bigcap_{i\in I}M_i\in\mathscr{A}$, $\sum_{i\in I}\omega_i\mu_i\big(\bigcap_{i=1}^\infty M_i\big)=0$, also $f^{-1}(B)\in\mathcal{M}_{\sum_{i\in I}\omega_i\mu_i}$. Somit ist f wie behauptet $\big(\sum_{i\in I}\omega_i\mu_i\big)$-messbar.

Zum Beweis der Eigenschaft (6) sei $X=\bigcup_{i=1}^\infty X_i\in\mathscr{A}$ die Vereinigung der disjunkten $X_i\in\mathscr{A}$. Im Fall einer $\mathscr{A}$-Treppenfunktion f mit wohldefiniertem Term $\sum_{i=1}^\infty\int_{X_i}f\,\mathrm{d}\mu$ ergibt sich mit der Definition des elementaren Integrals und der σ-Additivität von μ direkt

$$\int_X f\,\mathrm{d}\mu = \sum_{t\in\overline{\mathbb{R}}} t\,\mu(X\cap f^{-1}(t)) = \sum_{i=1}^\infty\sum_{t\in\overline{\mathbb{R}}} t\,\mu(X_i\cap f^{-1}(t)) = \sum_{i=1}^\infty\int_{X_i}f\,\mathrm{d}\mu$$

samt der μ-Integrierbarkeit von f über X. Für beliebige Funktionen f wird nun gezeigt, dass bei definierter rechter Seite stets

$$\int_X^* f\,\mathrm{d}\mu \le \sum_{i=1}^\infty \int_{X_i}^* f\,\mathrm{d}\mu$$

gilt. Hierfür kann $\int_{X_i}^* f\,\mathrm{d}\mu<\infty$ für alle $i\in I$ angenommen werden. Zu beliebigem $\varepsilon>0$ gibt es dann für jedes $i\in\mathbb{N}$ eine $\mathscr{A}|X_i$-Treppenfunktion g_i mit $g_i\ge f$ und $\int_{X_i}g_i\,\mathrm{d}\mu\le\int_{X_i}^* f\,\mathrm{d}\mu+2^{-i}\varepsilon$. Die g_i ergeben durch die Festlegung $g(x):=g_i(x)$ für $x\in X_i$ eine $\mathscr{A}|X$-Treppenfunktion g mit $g\ge f$, und mit deren Hilfe folgt

$$\int_X^* f\,\mathrm{d}\mu \le \int_X g\,\mathrm{d}\mu = \sum_{i=1}^\infty\int_{X_i}g\,\mathrm{d}\mu = \sum_{i=1}^\infty\int_{X_i}g_i\,\mathrm{d}\mu \le \sum_{i=1}^\infty\int_{X_i}^* f\,\mathrm{d}\mu+\varepsilon\,.$$

Wegen der Beliebigkeit von $\varepsilon>0$ erhält man die behauptete Ungleichung für Oberintegrale. Eine analoge Ungleichung gilt für Unterintegrale, und die Kombination der beiden liefert die Eigenschaft (6) in der behaupteten Allgemeinheit. Alternativ lässt sich die Eigenschaft (6) auch als Spezialfall der Eigenschaft (5) mit $\omega_i=1$, $\mu_i=\mathbb{1}_{X_i}\cdot\mu$ und $\sum_{i\in I}\mu_i=\mathbb{1}_X\cdot\mu$ lesen, wenn man noch $\mu_i|_{\mathscr{A}|X_i}=\mu|_{\mathscr{A}|X_i}$ und $\int_\Omega f\,\mathrm{d}\mu_i=\int_{X_i}f\,\mathrm{d}\mu_i=\int_{X_i}f\,\mathrm{d}\mu$ sowie Entsprechendes mit X anstelle X_i bemerkt. □

Definition (**Gewichtung**). *Sei $(\Omega, \mathscr{A}, \mu)$ ein Maßraum. Für μ-integrierbares $f\colon \Omega \to \overline{\mathbb{R}}$ erklärt man die **Gewichtung** $f\cdot\mu\colon \mathscr{A} \to \overline{\mathbb{R}}$ von μ mit f durch*

$$(f\cdot\mu)(A) := \int_A f\,\mathrm{d}\mu \qquad \textit{für } A \in \mathscr{A}\,.$$

*In diesem Kontext nennt man μ das **Grundmaß** und f die **Gewichts- oder Dichtefunktion**.*

Um das Grundmaß schon bei der Benennung einer Dichtefunktion explizit anzugeben, bezeichnet man eine solche Funktion gelegentlich als μ-Dichte, speziell für $\mu = \mathscr{L}^N$ als Lebesgue-Dichte und für $\mu = \xi$ als Zähldichte.

Bemerkung (zur Gewichtung von Maßen). Gilt $f \geq 0$ auf Ω, so ist $f\cdot\mu$ gemäß der vorausgehenden Grundeigenschaft (6) ein Maß auf $(\Omega, \mathscr{A})$, man spricht daher auch vom **gewichteten Maß** $f\cdot\mu$. Für allgemeines μ-integrierbares $f\colon \Omega \to \overline{\mathbb{R}}$ kann die σ-additive Abbildung $f\cdot\mu$ auf $\mathscr{A}$ sowohl positive als auch negative Werte annehmen. Man spricht dann von einem (gewichteten) *signierten* Maß.

Die Gewichtung $\mathbb{1}_X\cdot\mu$ mit einer charakteristischen Funktion $\mathbb{1}_X$ zu $X \in \mathscr{A}$ als Dichte stimmt, wie in Abschnitt 2.1 schon angekündigt, mit der dort durch Einschränkung und Fortsetzung beschriebenen Operation überein.

Als Nächstes wird gezeigt, dass nichtnegative (und nichtpositive) μ-messbare Funktionen immer μ-integrierbar sind. Nimmt ein μ-messbares f sowohl positive als auch negative Werte an, so kann man $f = f_+ - f_-$ in seinen Positivteil $f_+ := \max\{f, 0\}$ und seinen Negativteil $f_- := \max\{-f, 0\}$ zerlegen. Nach Grundeigenschaft (4) berechnet sich das μ-Integral von f dann als

$$\int_\Omega f\,\mathrm{d}\mu = \int_\Omega f_+\,\mathrm{d}\mu - \int_\Omega f_-\,\mathrm{d}\mu\,,$$

aber nur solange rechts nicht der Ausdruck $\infty - \infty$ auftritt. Dass der **„Fall $\infty - \infty$“ tatsächlich der *einzige* Nichtexistenzfall für Integrale messbarer Funktionen** ist, garantiert der nächste Satz:

Satz (über **Integrierbarkeitskriterien für messbare Funktionen**). *Sei $(\Omega, \mathscr{A}, \mu)$ ein Maßraum, und $f\colon \Omega \to \overline{\mathbb{R}}$ sei μ-messbar.*

(I) *Gilt $f \geq 0$ (oder $f \leq 0$) auf Ω, so ist f stets μ-integrierbar mit Integralwert $\int_\Omega f\,\mathrm{d}\mu$ in $[0,\infty]$ (oder in $[-\infty, 0]$).*

(II) *Die Funktion f ist genau dann μ-integrierbar, wenn eines der beiden (gemäß (I) stets definierten) Integrale $\int_\Omega f_+\,\mathrm{d}\mu$ und $\int_\Omega f_-\,\mathrm{d}\mu$ endlich ist.*

(III) *Die Funktion f ist genau dann μ-summierbar, wenn das Integral $\int_\Omega |f|\,\mathrm{d}\mu$ endlich ist.*

Beweis. Ohne Einschränkung sei $(\Omega, \mathscr{A}, \mu)$ vollständig. Zum Beweis von (I) betrachtet man im Fall $f \geq 0$ für beliebiges $s \in \mathbb{R}_{>1}$ die nichtnegative $\mathscr{A}$-Treppenfunktion g_s mit

$$g_s(x) := \begin{cases} s^z & \text{falls } s^z \leq f(x) < s^{z+1} \text{ für ein } z \in \mathbb{Z} \\ f(x) & \text{falls } f(x) \in \{0, \infty\} \end{cases}.$$

Dann ist g_s eine Unterfunktion zu f, und sg_s ist eine Oberfunktion zu f. Deshalb folgt

$$\int_\Omega^* f\,\mathrm{d}\mu \leq \int_\Omega sg_s\,\mathrm{d}\mu = s\int_\Omega g_s\,\mathrm{d}\mu \leq s\int_{*\Omega} f\,\mathrm{d}\mu\,.$$

Durch Grenzübergang $s \searrow 1$ erhält man

$$\int_\Omega^* f \,\mathrm{d}\mu \leq {}_*\!\int_\Omega f \,\mathrm{d}\mu \,,$$

also μ-Integrierbarkeit von f. Die Aussage für den Fall $f \leq 0$ folgt durch Übergang zu $-f$.

Zum Nachweis von (II) zeigt man die beiden Implikationen der behaupteten Äquivalenz: Ist einerseits eines der Integrale $\int_\Omega f_+ \,\mathrm{d}\mu$ und $\int_\Omega f_- \,\mathrm{d}\mu$ endlich, so ist f nach der Vorüberlegung mit Grundeigenschaft (4) μ-integrierbar. Liegt andererseits ein μ-integrierbares f vor, so unterscheidet man zwei Fälle: Im Fall $\int_\Omega f \,\mathrm{d}\mu < \infty$ gibt es eine Oberfunktion h zu f mit $\int_\Omega h \,\mathrm{d}\mu < \infty$. Mit der Definition des elementaren Integrals folgt $\int_\Omega h_+ \,\mathrm{d}\mu < \infty$, und da h_+ eine Oberfunktion zu f_+ ist, erhält man $\int_\Omega f_+ \,\mathrm{d}\mu < \infty$. Im Fall $\int_\Omega f \,\mathrm{d}\mu > -\infty$ zeigt ein analoges Argument $\int_\Omega f_- \,\mathrm{d}\mu < \infty$.

Bei der Äquivalenz in Teil (III) des Satzes folgt die Hin-Implikation aus dem für Teil (II) gegebenen Argument und der Beobachtung

$$\int_\Omega |f| \,\mathrm{d}\mu = \int_\Omega f_+ \,\mathrm{d}\mu + \int_\Omega f_- \,\mathrm{d}\mu \,.$$

Die Rück-Implikation ist mit dieser Formel und der Vorüberlegung zum Satz ebenfalls klar. $\square$

Im Folgenden werden Integrale bezüglich speziellen (Klassen von) Maßen, zunächst bezüglich diskreten Maßen, diskutiert. Dieser erste Fall ist in der Integrationstheorie aber nur am Rande von Interesse, denn die Integrale vereinfachen sich tatsächlich zu Summen:

Beispiele (zu Integralen bezüglich diskreter Maße)**.** Für eine beliebige Menge Ω und eine beliebige Funktion $f\colon \Omega \to \overline{\mathbb{R}}$ gelten:

(1) f ist δ_x-integrierbar bezüglich Dirac-Maßen δ_x zu $x \in \Omega$ mit

$$\int_\Omega f \,\mathrm{d}\delta_x = f(x) \,.$$

(2) f ist ξ-integrierbar bezüglich dem Zählmaß ξ auf Ω mit

$$\int_\Omega f \,\mathrm{d}\xi = \sum_{x\in\Omega} f(x) \,,$$

vorausgesetzt die Summe rechts ist nicht vom Typ $\infty-\infty$.

(3) f ist ω-integrierbar bezüglich einem diskreten Maß $\omega := \sum_{x\in\Omega} \omega_x \delta_x$ (zusammengesetzt aus Punktmassen $\omega_x \in [0,\infty]$) mit

$$\int_\Omega f \,\mathrm{d}\omega = \sum_{x\in\Omega} \omega_x f(x) \,,$$

vorausgesetzt die Summe rechts ist nicht vom Typ $\infty-\infty$.

Die wichtigsten Maßintegrale sind bezüglich dem Lebesgue-Maß $\mathscr{L}^N$ gebildete **Lebesgue-Integrale**. Insbesondere ist das Lebesgue-Integral bezüglich $\mathscr{L}^1$ im folgenden Sinn eine echte Verallgemeinerung des Riemann-Integrals:

Bemerkungen (zu **Riemann- und Lebesgue-$\mathscr{L}^1$-Integralen**). Sei $a < b$ in $\mathbb{R}$.

(0) Da einzelne Punkte $\mathscr{L}^1$-Nullmengen sind, gilt für $\mathscr{L}^1$-messbares $f\colon [a,b] \to \overline{\mathbb{R}}$ stets

$$\int_{[a,b]} f \,\mathrm{d}\mathscr{L}^1 = \int_{(a,b)} f \,\mathrm{d}\mathscr{L}^1 = \int_{[a,b)} f \,\mathrm{d}\mathscr{L}^1 = \int_{(a,b]} f \,\mathrm{d}\mathscr{L}^1 \,,$$

und die an Riemann-Integrale erinnernde Notation $\int_a^b f \,\mathrm{d}\mathscr{L}^1 := \int_{(a,b)} f \,\mathrm{d}\mathscr{L}^1$ macht Sinn. Für Lebesgue-Stieltjes-Maße $\mathscr{L}^1_F$ *mit stetigem F* verhält sich dies analog.

(1) **Ist eine beschränkte Funktion** $f\colon [a,b] \to \mathbb{R}$ über $[a,b]$ **Riemann-integrierbar, so ist sie auch $\mathscr{L}^1$-integrierbar** über $[a,b]$ mit gleichem Integralwert

$$\int_a^b f \,\mathrm{d}\mathscr{L}^1 = \text{Riemann-}\!\!\int_a^b f \,.$$

Dies ergibt sich sofort aus folgender Beobachtung: Die Ober- und Unterfunktionen der Riemannschen Theorie sind auch in der Lebesgueschen Theorie zulässig, und deshalb sind das Lebesguesche Ober- und Unterintegral zwischen dem Riemannschen Ober- und Unterintegral eingeschachtelt.

(2) Es gibt viele **$\mathscr{L}^1$-integrierbare Funktionen**, die **nicht Riemann-integrierbar** sind, beispielsweise existieren $\int_0^1 \mathbb{1}_\mathbb{Q} \,\mathrm{d}\mathscr{L}^1 = 0$ und $\int_0^1 \mathbb{1}_{\mathbb{R}\setminus\mathbb{Q}} \,\mathrm{d}\mathscr{L}^1 = 1$ (sogar als *elementare* $\mathscr{L}^1$-Integrale), aber $\mathbb{1}_\mathbb{Q}$ und $\mathbb{1}_{\mathbb{R}\setminus\mathbb{Q}}$ sind nicht Riemann-integrierbar über $[0,1]$.

(3) Tatsächlich lässt sich Riemann-Integrierbarkeit im Rahmen der Lebesgueschen Integrationstheorie sogar besser und ziemlich abschließend verstehen:

Satz (**Charakterisierung der Riemann-Integrierbarkeit**). *Eine beschränkte Funktion $f\colon [a,b] \to \mathbb{R}$ ist genau dann Riemann-integrierbar über $[a,b]$, wenn sie an $\mathscr{L}^1$-fast allen Stellen von $[a,b]$ stetig ist.*

Beweis. Seien $\mathcal{S}_f$ die Menge der Stetigkeitsstellen und $\mathcal{U}_f$ die der Unstetigkeitsstellen von f in $[a,b]$. Für $k \in \mathbb{N}$ sei zudem $\mathcal{S}_{f,k}$ die Menge der $x \in [a,b]$ mit $\max\{f(x), \limsup_{y\to x} f(y)\} < \min\{f(x), \liminf_{y\to x} f(y)\} + \frac{1}{k}$, und es sei $\mathcal{U}_{f,k} := [a,b] \setminus \mathcal{S}_{f,k}$. Dann sind die $\mathcal{U}_{f,k}$ abgeschlossen, und $\mathcal{S}_f = \bigcap_{k=1}^\infty \mathcal{S}_{f,k}$ und $\mathcal{U}_f = \bigcup_{k=1}^\infty \mathcal{U}_{f,k}$ sind Borel-Mengen. Mit diesen Notationen werden nun die beiden Implikationen der behaupteten Äquivalenz einzeln gezeigt.

Für die Hin-Implikation verwendet man ein Widerspruchsargument: Angenommen, es ist $\mathscr{L}^1(\mathcal{U}_f) > 0$. Dann ist auch $\mathscr{L}^1(\mathcal{U}_{f,k}) > 0$ für ein $k \in \mathbb{N}$. Weil f Riemann-integrierbar ist, gibt es eine Riemannsche Oberfunktion h zu f und eine Riemannsche Unterfunktion g zu f mit $\int_a^b h < \int_a^b g + \frac{1}{k}\mathscr{L}^1(\mathcal{U}_{f,k})$. Für $\mathscr{L}^1$-fast-alle $x \in \mathcal{U}_{f,k}$ und genauer für alle $x \in \mathcal{U}_{f,k}$ außer den endlich vielen Unstetigkeitsstellen von g und h gilt aber $h(x) \ge \max\{f(x), \limsup_{y\to x} f(y)\} \ge \min\{f(x), \liminf_{y\to x} f(y)\} + \frac{1}{k} \ge g(x) + \frac{1}{k}$, und deshalb ist $\int_a^b h \ge \int_a^b g + \frac{1}{k}\mathscr{L}^1(\mathcal{U}_{f,k})$ (wobei Übereinstimmung der elementaren Riemann- und Lebesgue-Integrale verwendet wurde). Damit ist ein Widerspruch erreicht. Also muss $\mathscr{L}^1(\mathcal{U}_f) = 0$ sein.

Für die Rück-Implikation sei $M := \sup_{[a,b]} |f| < \infty$, und $k \in \mathbb{N}$ sei beliebig fixiert. Da $\mathcal{U}_f$ und damit auch $\mathcal{U}_{f,k}$ nach Voraussetzung $\mathscr{L}^1$-Nullmengen sind, ist $\mathcal{U}_{f,k}$, wie in den Übungen schon gesehen, in einer offenen Menge von $\mathscr{L}^1$-Maß $<\frac{1}{k}$ enthalten. Die offene Menge kann als Vereinigung offener Intervalle geschrieben werden, und, weil $\mathcal{U}_{f,k}$ abgeschlossen und dann auch kompakt ist, gibt es eine endliche Überdeckung $\mathcal{U}_f \subset \bigcup_{j=1}^{\ell} I_j$ mit offenen Intervallen I_j und $\mathscr{L}^1\big(\bigcup_{j=1}^{\ell} I_j\big) < \frac{1}{k}$. Das Komplement von $\bigcup_{j=1}^{\ell} I_j$ in $[a,b]$ hat die Form $\bigcup_{i=1}^{m} [s_i,t_i]$ mit *disjunkten* $[s_i,t_i]$ in $\mathcal{S}_{f,k}$. Man überlegt sich nun, dass ein $\delta \in \mathbb{R}_{>0}$ gibt, so dass für Intervalle $J \subset \bigcup_{i=1}^{m} [s_i,t_i]$ der Länge $<\delta$ stets $\sup_J f < \inf_J f + \frac{1}{k}$ eintritt. Wäre dies nämlich falsch, so gäbe es Folgen $(x_n)_{n\in\mathbb{N}}$ und $(\widetilde{x}_n)_{n\in\mathbb{N}}$ in $\bigcup_{i=1}^{m} [s_i,t_i]$ mit $\lim_{n\to\infty} |\widetilde{x}_n - x_n| = 0$ und $\liminf_{n\to\infty} f(x_n) \geq \limsup_{n\to\infty} f(\widetilde{x}_n) + \frac{1}{k}$; diese hätten einen gemeinsamen Häufungspunkt in $\bigcup_{i=1}^{m} [s_i,t_i]$, der im Widerspruch zur bisherigen Konstruktion auch in $\mathcal{U}_{f,k}$ liegen müsste. Also gibt es das behauptete δ, und nach Zerlegung von $[s_i,t_i]$ in Intervalle der Länge $<\delta$ kann man eine Riemannsche Oberfunktion h zu f aus den Suprema von f auf den Teilintervallen und eine Riemannsche Unterfunktion g zu f aus den Infima von f auf den Teilintervallen konstruieren, so dass $\int_{s_i}^{t_i} h \leq \int_{s_i}^{t_i} g + \frac{t_i - s_i}{k}$ gilt. Setzt man noch $g \coloneqq -M$ und $h \coloneqq M$ auf $\bigcup_{j=1}^{\ell} I_j$, so ist g auf ganz $[a,b]$ definiert und Riemannsche Oberfunktion zu f, und h ist dort Riemannsche Unterfunktion. Durch Zusammensetzen der Abschätzungen auf $[s_i,t_i]$ und der Abschätzung $h \leq g + 2M$ auf der Menge $\bigcup_{j=1}^{\ell} I_j$ vom $\mathcal{L}^1$-Maß $<\frac{1}{k}$ ergibt sich insgesamt $\int_a^b h \leq \int_a^b g + \frac{b-a}{k} + \frac{2M}{k}$. Da $k \in \mathbb{N}$ beliebig war, ist damit Riemann-Integrierbarkeit von f über $[a,b]$ gezeigt. □

(4) Existieren das "normale" $\mathscr{L}^1$-Integral und das uneigentliche Riemann-Integral (über ein unbeschränktes Intervall oder für einen unbeschränkten Integranden), so stimmen sie ebenfalls überein. Kürzungseffekte können aber dazu führen, dass das uneigentliche Riemann-Integral einen endlichen Wert annimmt, während das $\mathscr{L}^1$-Integral undefiniert vom Typ $\infty - \infty$ ist.

(5) In ähnlicher Weise verallgemeinern $\mathscr{L}_F^1$-Integrale Riemann-Stieltjes-Integrale und das $\mathscr{L}^N$-Integral iterierte Riemann-Integrale.

Nach (1) ist klar, dass konkrete Berechnungen von $\mathscr{L}^1$-Integralen mit den von Riemann-Integralen bekannten Techniken und Rechenregeln durchgeführt werden können. Die Berechnung von $\mathscr{L}^N$-Integralen (und anderen mehrdimensionalen Maßintegralen) kann in vielen Fällen hierauf zurückgeführt werden. Eine erstes Resultat hierzu, das allerdings nur in sehr speziellen Situationen anwendbar ist, sieht wie folgt aus:

Satz (Formel für radiale Integration). *Für $0 \leq r \leq R \leq \infty$ und jede $\mathscr{L}^1$-messbare Funktion $\varphi\colon (r,R) \to \overline{\mathbb{R}}$ gilt (wobei beide Integrale existieren oder keines)*

$$\int_{\mathrm{B}_R(0)\setminus \mathrm{B}_r(0)} \varphi(|x|)\,\mathrm{d}\mathscr{L}^N(x) = N\omega_N \int_r^R \varphi(\varrho)\varrho^{n-1}\,\mathrm{d}\mathscr{L}^1(\varrho)$$

mit N-dimensionalen Kugeln $\mathrm{B}_r(x) = \mathrm{B}_r^N(x) = \{y \in \mathbb{R}^N : |y-x| < r\}$ und dem ***Volumen*** *$\omega_N \coloneqq \mathscr{L}^N(\mathrm{B}_1^N(0))$* ***der N-dimensionalen Einheitskugel****.*

Beweisskizze (nach Standard-Ausdehnungsprozedur). Man verifiziert die Behauptung zuerst für $\varphi = \mathbb{1}_{[a,b)}$ mit $r < a \leq b \leq R$ durch Berechnung der Integrale auf beiden Seiten: Aufgrund von $\varphi(|x|) = \mathbb{1}_{\mathrm{B}_b^N(0)\setminus \mathrm{B}_a^N(0)}(x)$ erhält man links $\mathscr{L}^N\big(\mathrm{B}_b^N(0) \setminus \mathrm{B}_a^N(0)\big) = \omega_N b^N - \omega_N a^N$, rechts ergibt sich $N\omega_N \int_a^b \varrho^{N-1}\,\mathrm{d}\mathscr{L}^1(\varrho) = \omega_N b^N - \omega_N a^N$, und damit ist die Behauptung für $\varphi = \mathbb{1}_{[a,b)}$ gezeigt.

Betrachtet man allgemeiner $\varphi = \mathbb{1}_A$ mit $A \in \mathcal{B}((r,R))$, so kann man beide Seiten der Behauptung als Maße (in A) verstehen, die gemäß dem schon Gezeigten auf dem Erzeuger $\mathcal{I}_1 \cap \mathcal{P}((r,R))$

von $\mathcal{B}((r,R))$ übereinstimmen. Nach dem Maßfortsetzungssatz gilt die Behauptung daher allgemein für $\varphi = \mathbb{1}_A$, und mit Grundeigenschaft (4) folgt sie für alle $\mathcal{B}((r,R))$-Treppenfunktionen φ mit endlich vielen Stufen.

Ist $\varphi\colon (r,R) \to [0,\infty]$ eine $\mathscr{L}^1$-messbare Funktion, so lässt sich φ gemäß Abschnitt 2.5 (Satz über den Borel-Repräsentanten und Approximationslemma) von unten durch eine $\mathscr{L}^1$-fast-überall konvergente Folge von $\mathcal{B}((r,R))$-Treppenfunktionen $g_k\colon (r,R) \to \mathbb{R}_{\ge 0}$ mit endlich vielen Stufen approximieren. Es folgen die Konvergenzen $g_k(\varrho)\varrho^{n-1} \to \varphi(\varrho)\varrho^{n-1}$ für $\mathscr{L}^1$-fast-alle $\varrho \in (r,R)$ und $g_k(|x|) \to \varphi(|x|)$ für $\mathscr{L}^N$-fast-alle[27] $x \in \mathrm{B}^N_R(0) \setminus \mathrm{B}^N_r(0)$. Ein Konvergenzsatz des folgenden Abschnitts 2.7 stellt Konvergenz der zugehörigen Integrale sicher und erlaubt die Übertragung der Behauptung von den g_k auf φ.

Final folgt die behauptete Version des Satzes durch Zerlegung $\varphi = \varphi_+ - \varphi_-$. □

Anwendungen (der Formel für radiale Integration auf **Integrale von Potenzfunktionen**).

(1) Für $s \in \mathbb{R}$ gilt

$$\int_{\mathrm{B}^N_1(0)} |x|^s \,\mathrm{d}\mathscr{L}^N(x) = N\omega_N \int_0^1 \varrho^{s+N-1} \,\mathrm{d}\mathscr{L}^1(\varrho) = \begin{cases} \frac{N\omega_N}{s+N} & \text{für } s > -N \\ \infty & \text{für } s \le -N \end{cases}.$$

(2) Für $s \in \mathbb{R}$ gilt

$$\int_{\mathbb{R}^N \setminus \mathrm{B}^N_1(0)} |x|^s \,\mathrm{d}\mathscr{L}^N(x) = N\omega_N \int_1^\infty \varrho^{s+N-1} \,\mathrm{d}\mathscr{L}^1(\varrho) = \begin{cases} \infty & \text{für } s \ge -N \\ \frac{N\omega_N}{-s-N} & \text{für } s < -N \end{cases}.$$

(3) Aus (1) und (2) folgt

$$\int_{\mathbb{R}^N} |x|^s \,\mathrm{d}\mathscr{L}^N(x) = \infty \qquad \text{für alle } s \in \mathbb{R}\,.$$

(4) Da $\int_{\mathrm{B}^N_1(0)} x_i^2 \,\mathrm{d}\mathscr{L}^N$ für alle $i \in \{1,2,\ldots,N\}$ den gleichen Wert hat (wie sich aus der Invarianz des Lebesgue-Maßes und des Lebesgue-Integrals unter Koordinatenpermutationen, einem einfacher einzusehenden Spezialfall der später in Abschnit 2.11 gezeigten Rotationsinvarianz, ergibt), erhält man außerdem

$$\int_{\mathrm{B}^N_1(0)} x_i^2 \,\mathrm{d}\mathscr{L}^N(x) = \frac{1}{N} \int_{\mathrm{B}^N_1(0)} |x|^2 \,\mathrm{d}\mathscr{L}^N(x) = \frac{\omega_N}{N+2} \qquad \text{für } i \in \{1,2,\ldots,N\}\,.$$

Statt $\overline{\mathbb{R}}$-wertigen Funktionen, wie sie bisher betrachtet wurden, kann man auch komplexwertige und vektorwertige Funktionen integrieren:

Definition (Integrale $\mathbb{C}$-wertiger und vektorwertiger Funktionen). *Sei $(\Omega, \mathscr{A}, \mu)$ ein Maßraum. Das (μ-)Integral einer μ-messbaren $\mathbb{C}$-wertigen Funktion $f = \mathcal{Re}(f) + \mathbf{i}\,\mathcal{Im}(f)\colon \Omega \to \mathbb{C}$ erklärt man als*

$$\int_\Omega f \,\mathrm{d}\mu := \int_\Omega \mathcal{Re}(f) \,\mathrm{d}\mu + \mathbf{i} \int_\Omega \mathcal{Im}(f) \,\mathrm{d}\mu \in \mathbb{C}\,,$$

[27] An dieser Stelle wird implizit benutzt, dass für jede $\mathscr{L}^1$-Nullmenge M die Menge $\{x \in \mathbb{R}^N : |x| \in M\}$ eine $\mathscr{L}^N$-Nullmenge ist. Dies folgt aus dem vorigen Beweisschritt, zunächst für Borelsches M und dann auch allgemein.

und das (μ-)Integral einer μ-messbaren $\mathbb{K}^M$-wertigen Funktion $f = (f_1, f_2, \dots, f_M)\colon \Omega \to \mathbb{K}^M$ definiert man darauf aufbauend als

$$\int_\Omega f \,\mathrm{d}\mu := \left(\int_\Omega f_1 \,\mathrm{d}\mu\,,\, \int_\Omega f_2 \,\mathrm{d}\mu\,,\, \dots\,,\, \int_\Omega f_M \,\mathrm{d}\mu \right) \in \mathbb{K}^M\,.$$

Dabei wird das Integral in beiden Fällen nur dann als definiert betrachtet und f als μ-summierbar bezeichnet, ***wenn alle rechts auftretenden Integrale definiert und endlich sind****.*

Bemerkungen (zur Integration $\mathbb{K}^M$-wertiger Funktionen)**.**

(1) Die Integralbegriffe für $\mathbb{C}^M$- und $\mathbb{R}^{2M}$-wertige Funktionen stimmen bei Identifikation $\mathbb{C} = \mathbb{R}^2$ durch $a+\mathbf{i}b = (a, b)$ überein.

(2) **Praktisch alles bisher Behandelte gilt** für Integrale $\mathbb{C}$- und $\mathbb{K}^M$-wertiger Funktionen **analog**. Zusätzlich seien hier nur zwei Punkte angemerkt: Zum einen gilt die Dreiecksungleichung mit der Euklidischen Norm von $\mathbb{C}$ oder $\mathbb{K}^M$ (und tatsächlich auch mit anderen Normen) ganz analog. Zum anderen hängen Integrale $\mathbb{C}^M$-wertiger Funktionen $\mathbb{C}$-linear vom Integranden ab, d.h. die Grundeigenschaft (4) gilt in diesem Fall sogar für Faktoren $r, s \in \mathbb{C}$.

2.7 Konvergenzsätze für Integrale

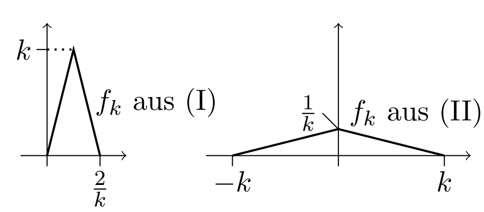
In der Integrationstheorie (und im weiteren Verlauf der Vorlesung) tritt häufig die Frage auf, ob man aus einer Konvergenz $f_k \underset{k\to\infty}{\longrightarrow} f$ von Integranden auf Konvergenz $\int_\Omega f_k \,\mathrm{d}\mu \underset{k\to\infty}{\longrightarrow} \int_\Omega f \,\mathrm{d}\mu$ der zugehörigen Integrale schließen kann. Mit anderen Worten handelt es sich dabei um die Frage, ob Vertauschungen von Limes und Integral gerechtfertigt sind und Maßintegrale stetig vom Integranden abhängen. Diese Fragen lassen sich einfach mit „Ja" beantworten, wenn $(\Omega, \mathscr{A}, \mu)$ ein *endlicher* Maßraum und die Konvergenz $f_k \underset{k\to\infty}{\longrightarrow} f$ mit μ-messbaren $f_k, f\colon \Omega \to \mathbb{R}$ sogar *gleichmäßig* auf Ω ist. Dies zeigt man genau wie in der Riemannschen Integrationstheorie durch direkte Abschätzung $|\int_\Omega f_k \,\mathrm{d}\mu - \int_\Omega f \,\mathrm{d}\mu| \le \int_\Omega |f_k - f| \,\mathrm{d}\mu \le \mu(\Omega) \sup_\Omega |f_k - f| \underset{k\to\infty}{\longrightarrow} 0$. Dass die Gleichmäßigkeits- und die Endlichkeits-Voraussetzung in diesem „trivialen Konvergenzsatz" schon für $\mu = \mathscr{L}^1$ unverzichtbar sind, zeigen folgende Gegenbeispiele mit Zackenfunktionen:

Beispiel. Sei $f_k\colon \mathbb{R} \to [0, \infty]$ definiert durch

(I) $f_k(x) := k(1-|1-kx|)_+$

oder

(II) $f_k(x) := k^{-2}(k-|x|)_+$.



In beiden Fällen konvergiert f_k punktweise auf $\mathbb{R}$ gegen 0 bei $k \to \infty$, aber $\int_\Omega f_k \,\mathrm{d}\mathscr{L}^1 = 1$ konvergiert nicht gegen 0. Bei (I) sind außerdem die Träger der f_k gleichmäßig beschränkt und man kann dieses Beispiel genauso gut auf der $\mathscr{L}^1$-endlichen Menge $[0, 2]$ betrachten, während bei (II) die Konvergenz gleichmäßig ist.

Im Folgenden wird untersucht, was sich trotz solcher Beispiele über stetige Abhängigkeit des Integrals vom Integranden aussagen lässt. Als Erstes sei dazu festgehalten, dass im Fall *nichtnegativer* Integranden zwar nicht Stetigkeit, aber ***Unterhalb*****stetigkeit** generell vorliegt:

Lemma (von Fatou). *Sei* $(\Omega, \mathscr{A}, \mu)$ *ein Maßraum, und die Funktionen* $f_k \colon \Omega \to [0,\infty]$ *seien* μ*-messbar. Dann gilt*

$$\liminf_{k\to\infty} \int_\Omega f_k \,\mathrm{d}\mu \geq \int_\Omega \big(\liminf_{k\to\infty} f_k\big)\,\mathrm{d}\mu\,.$$

Es werden zwei alternative Beweise des Lemmas ausgeführt:

1. Beweis mit dem Satz von Egoroff. Ohne Einschränkung sei μ vollständig, und vorab sei festgehalten, dass $\widetilde{f}_k := \inf_{\ell\in\mathbb{N}_{\geq k}} f_\ell$ gegen die messbare Funktion $f := \liminf_{k\to\infty} f_k$ konvergieren.

Zuerst wird der Fall behandelt, dass $\mu(\Omega) < \infty$ gilt und die f_k sowie f nur endliche Werte in $\mathbb{R}_{\geq 0}$ annehmen. Nach dem Satz von Egoroff gibt es dann Mengen $A_1 \subset A_2 \subset A_3 \subset \ldots$ in $\mathscr{A}$, so dass für jedes $i \in \mathbb{N}$ die Konvergenz $\widetilde{f}_k \to f$ auf A_i gleichmäßig ist und $\mu(\Omega \setminus A_i) < \frac{1}{i}$ gilt. Insbesondere konvergiert $\int_{A_i} f_k \,\mathrm{d}\mu$ für $k \to \infty$ gegen $\int_{A_i} f\,\mathrm{d}\mu$, und $\Omega \setminus \bigcup_{i=1}^\infty A_i$ ist eine μ-Nullmenge. Deshalb ergibt sich mit einer Stetigkeitseigenschaft des gewichteten Maßes $f{\cdot}\mu$ die Ungleichungskette

$$\liminf_{k\to\infty} \int_\Omega f_k\,\mathrm{d}\mu \geq \lim_{k\to\infty} \int_{A_i} \widetilde{f}_k\,\mathrm{d}\mu = \int_{A_i} f\,\mathrm{d}\mu \xrightarrow[i\to\infty]{} \int_{\bigcup_{i=1}^\infty A_i} f\,\mathrm{d}\mu = \int_\Omega f\,\mathrm{d}\mu$$

und somit die Behauptung.

Mit einem Standard-Argument überträgt man die gezeigte Aussage als Nächstes auf den Fall, dass μ nur σ-endlich ist (wobei f_k, f weiterhin nur endliche Werte annehmen).

Schließlich wird der allgemeine Fall betrachtet, wobei ohne Einschränkung angenommen werden kann, dass μ vollständig ist und $\liminf_{k\to\infty} \int_\Omega f_k\,\mathrm{d}\mu < \infty$ gilt. Somit gibt es ein $k_0 \in \mathbb{N}$ mit $\int_\Omega \widetilde{f}_k\,\mathrm{d}\mu < \infty$ für alle $k \in \mathbb{N}_{\geq k_0}$, und aus der Existenz endlicher Obersummen folgt, dass alle $\Omega_k := \{x \in \Omega : \widetilde{f}_k(x) > 0\}$ mit $k \in \mathbb{N}_{\geq k_0}$ und folglich auch $\Omega_* := \bigcup_{k=k_0}^\infty \Omega_k$ jeweils μ-σ-endliche Mengen sind. Auf Ω_* kann der vorige Schritt für die Funktionen $\min\{f_k, M\}$ und $\min\{f, M\}$ mit beliebigem $M \in \mathbb{N}$ angewandt werden und ergibt

$$\liminf_{k\to\infty} \int_\Omega f_k\,\mathrm{d}\mu \geq \liminf_{k\to\infty} \int_{\Omega_*} \min\{f_k, M\}\,\mathrm{d}\mu \geq \int_{\Omega_*} \min\{f, M\}\,\mathrm{d}\mu = \int_\Omega \min\{f, M\}\,\mathrm{d}\mu\,,$$

wobei im letzten Schritt einging, dass $f = \lim_{k\to\infty} \widetilde{f}_k$ auf $\Omega \setminus \Omega_*$ verschwindet. Mit $\{f{\leq}M\} := \{x \in \Omega : f(x) \leq M\}$ und analogen Notationen, den Beobachtungen $\{f \leq M\} \subset \{f \leq M{+}1\}$ und $\bigcup_{M=1}^\infty \{f \leq M\} = \{f < \infty\}$ sowie erneut der Stetigkeitseigenschaft von $f\cdot\mu$, lässt sich die Ungleichungskette fortsetzen durch

$$\ldots \geq \int_{\{f\leq M\}} f\,\mathrm{d}\mu + M\mu\big(\{f=\infty\}\big) \xrightarrow[M\to\infty]{} \int_{\{f<\infty\}} f\,\mathrm{d}\mu + \infty\,\mu\big(\{f=\infty\}\big) = \int_\Omega f\,\mathrm{d}\mu\,.$$

Insgesamt folgt somit die Behauptung des Lemmas im allgemeinen Fall. □

2. Beweis durch direkte Betrachtung von (*Super-*)*Niveaumengen.* Sei μ vollständig. Außerdem sei $f := \liminf_{k\to\infty} f_k \geq 0$, und $h \colon \Omega \to \mathbb{R}_{\geq 0}$ sei eine μ-Unterfunktion zu f. Dann gilt $h = \sum_{i=1}^\infty t_i \mathbb{1}_{A_i}$ für $t_i \in \mathbb{R}_{\geq 0}$ und disjunkte $A_i \in \mathscr{A}$. Für fixiertes $\varepsilon \in \mathbb{R}_{>0}$ setzt man

$$A_{i,k} := \bigcap_{\ell=k}^\infty \{x \in A_i : f_\ell(x) \geq (1{-}\varepsilon)t_i\}$$

und erhält $A_{i,k} \subset A_{i,k+1}$ sowie $\bigcup_{k=1}^{\infty} A_{i,k} = A_i$ gemäß der Wahl von f und h. Insgesamt folgt

$$\liminf_{k\to\infty} \int_\Omega f_k \,\mathrm{d}\mu \geq \liminf_{k\to\infty} \sum_{i=1}^{m} \int_{A_{i,k}} f_k \,\mathrm{d}\mu \geq (1-\varepsilon) \liminf_{k\to\infty} \sum_{i=1}^{m} t_i \mu(A_{i,k}) = (1-\varepsilon) \sum_{i=1}^{m} t_i \mu(A_i)$$

für alle $m \in \mathbb{N}$. Mit den Grenzübergängen $m \to \infty$ und $\varepsilon \searrow 0$ ergibt sich

$$\liminf_{k\to\infty} \int_\Omega f_k \,\mathrm{d}\mu \geq \int_\Omega h \,\mathrm{d}\mu\,.$$

Die letzte Ungleichung überträgt sich von μ-Unterfunktionen h zu f mit Werten in $\mathbb{R}_{\geq 0}$ auf beliebige μ-Unterfunktionen h zu f. Weil das Integral der nichtnegativen messbaren Funktion f stets existiert, ergibt sich durch Supremumsbildung also

$$\liminf_{k\to\infty} \int_\Omega f_k \,\mathrm{d}\mu \geq \int_\Omega f \,\mathrm{d}\mu\,, \qquad \square$$

Wichtiger noch als die allgemeine Unterhalbstetigkeitsaussage des Lemmas ist aber die Erkenntnis, dass Stetigkeit relativ oft sichergestellt werden kann:

Hauptsatz (Konvergenzsätze für Integrale). *Sei $(\Omega, \mathscr{A}, \mu)$ ein Maßraum, die Funktionen $f_k \colon \Omega \to \overline{\mathbb{R}}$ seien μ-messbar, und es liege μ-fast-überall die Konvergenz $f_k \underset{k\to\infty}{\longrightarrow} f$ vor. Dann folgt Konvergenz der Integrale*

$$\lim_{k\to\infty} \int_\Omega f_k \,\mathrm{d}\mu = \int_\Omega f \,\mathrm{d}\mu\,,$$

vorausgesetzt ***eine der folgenden Zusatzbedingungen*** *ist erfüllt:*

(I) ***(Satz von Beppo Levi über monotone Konvergenz)***

$$0 \leq f_1 \leq f_2 \leq f_3 \leq \ldots \qquad \mu\text{-fast-überall auf } \Omega\,,$$

(II) ***(Satz von Lebesgue über dominierte/majorisierte Konvergenz)***

$$|f_k| \leq g \qquad \mu\text{-fast-überall auf } \Omega$$

für alle $k \in \mathbb{N}$ mit μ-summierbarer Majorante g (d.h. g ist μ-messbar mit $\int_\Omega g \,\mathrm{d}\mu < \infty$),

(III) ***(Konvergenzsatz von Vitali)*** *$\mu(\Omega) < \infty$, μ-fast-überall $f_k, f \in \mathbb{R}$ auf Ω und $(f_k)_{k\in\mathbb{N}}$ gleichgradig μ-integrierbar über Ω, wobei letzteres bedeutet, dass zu jedem $\varepsilon \in \mathbb{R}_{>0}$ ein $\delta \in \mathbb{R}_{>0}$ mit der Eigenschaft $\big(B \in \mathscr{A}\,,\ \mu(B) < \delta \implies \sup_{k\in\mathbb{N}} \int_B |f_k| \,\mathrm{d}\mu < \varepsilon\big)$ existiert.*

Bemerkungen (zu Varianten der Konvergenzssätze)**.**

(1) Das Lemma von Fatou und der Satz über monotone Konvergenz bleiben richtig, wenn man statt Nichtnegativität der f_k nur voraussetzt, dass μ-fast-überall $f_k \geq g$ durch ein μ-messbares g mit $\int_\Omega g_- \,\mathrm{d}\mu < \infty$ abgeschätzt werden kann.

(2) Der Satz über dominierte Konvergenz gilt auch dann, wenn nur μ-fast-überall $|f_k| \leq g_k$ mit *von k abhängigen Majoranten* g_k abgeschätzt werden kann, aber μ-fast-überall $g_k \to g$ konvergiert und die Konvergenz $\int_\Omega g_k \,\mathrm{d}\mu \to \int_\Omega g \,\mathrm{d}\mu$ der Integrale der Majoranten bereits sichergestellt ist.

Beweis des Satzes. Unter der Zusatzvoraussetzung (I) ist μ-fast-überall $f_k \le f$ auf Ω, und deshalb folgt $\limsup_{k\to\infty} \int_\Omega f_k \,\mathrm{d}\mu \le \int_\Omega f \,\mathrm{d}\mu$. Andererseits gilt $\liminf_{k\to\infty} \int_\Omega f_k \,\mathrm{d}\mu \ge \int_\Omega f \,\mathrm{d}\mu$ gemäß dem Lemma von Fatou.

In der Situation (II) gelten $|f| \le g$ und $2g - |f_k - f| \ge 2g - |f_k| - |f| \ge 0$ jeweils μ-fast-überall auf Ω. Mit Fatou folgt

$$\int_\Omega 2g \,\mathrm{d}\mu \le \lim_{k\to\infty} \int_\Omega (2g - |f_k - f|) \,\mathrm{d}\mu \le \int_\Omega 2g \,\mathrm{d}\mu \,,$$

und mit der Dreiecksungleichung ergibt sich

$$\left| \int_\Omega f_k \,\mathrm{d}\mu - \int_\Omega f \,\mathrm{d}\mu \right| \le \int_\Omega |f_k - f| \,\mathrm{d}\mu \xrightarrow[k\to\infty]{} 0 \,.$$

Damit ist die behauptete Konvergenz der Integrale auch unter (II) bewiesen.

Zur Behandlung der Voraussetzung (III) sei μ o.E. vollständig. Man überlegt sich zuerst bei festem $k \in \mathbb{N}$, dass $\lim_{\ell\to\infty} \mu(\{x \in \Omega : |f_k(x)| > \ell\}) = 0$ (wegen fast-überall endlicher Werte und Stetigkeitseigenschaft) und $\int_{\{x\in\Omega : |f_k(x)| > \ell_k\}} |f_k| \,\mathrm{d}\mu < 1$ für ein ausreichend großes $\ell_k \in \mathbb{N}$ (gemäß gleichgradiger Integrierbarkeit mit $\varepsilon = 1$) gelten. Damit gilt $\int_\Omega |f_k| \,\mathrm{d}\mu < 1 + \ell_k \mu(\Omega) < \infty$, und $\int_\Omega f_k \,\mathrm{d}\mu \in \mathbb{R}$ ist überhaupt wohldefiniert. Ein ähnliches Argument, bei dem die gleichgradige Integrierbarkeit der f_k mit dem Lemma von Fatou auf f übertragen wird, zeigt Wohldefiniertheit von $\int_\Omega f \,\mathrm{d}\mu \in \mathbb{R}$. Zum Nachweis der eigentlichen Konvergenzaussage sei $\varepsilon \in \mathbb{R}_{>0}$ beliebig. Für das zugehörige $\delta \in \mathbb{R}_{>0}$ mit der vorausgesetzten Eigenschaft liefert der Satz von Egoroff ein $A_\delta \in \mathscr{A}$, so dass $\lim_{k\to\infty} \sup_{A_\delta} |f_k - f| = 0$ und $\mu(\Omega \setminus A_\delta) < \delta$ gelten. Gemäß der Eigenschaft von δ ist damit insbesondere $\int_{\Omega\setminus A_\delta} |f_k| \,\mathrm{d}\mu < \varepsilon$ für alle $k \in \mathbb{N}$, und mit dem Lemma von Fatou folgt $\int_{\Omega\setminus A_\delta} |f| \,\mathrm{d}\mu \le \varepsilon$. Alles in allem ergibt sich aus diesen Beobachtungen

$$\begin{aligned}
\left| \int_\Omega f_k \,\mathrm{d}\mu - \int_\Omega f \,\mathrm{d}\mu \right| &\le \left| \int_{\Omega\setminus A_\delta} f_k \,\mathrm{d}\mu \right| + \left| \int_{A_\delta} (f_k - f) \,\mathrm{d}\mu \right| + \left| \int_{\Omega\setminus A_\delta} f \,\mathrm{d}\mu \right| \\
&\le \int_{\Omega\setminus A_\delta} |f_k| \,\mathrm{d}\mu + \mu(\Omega) \underbrace{\sup_{A_\delta} |f_k - f|}_{\xrightarrow[k\to\infty]{} 0} + \int_{\Omega\setminus A_\delta} |f| \,\mathrm{d}\mu < 3\varepsilon
\end{aligned}$$

für ausreichend großes $k \in \mathbb{N}$. Damit ist die behauptete Konvergenz gezeigt. □

Korollar (über Funktionenreihen). *Sei* $(\Omega, \mathscr{A}, \mu)$ *ein Maßraum. Für nichtnegatives* μ*-messbares* $f_k \colon \Omega \to [0, \infty]$ *gilt stets*
.

$$\sum_{k=1}^{\infty} \int_\Omega f_k \,\mathrm{d}\mu = \int_\Omega \left(\sum_{k=1}^{\infty} f_k \right) \mathrm{d}\mu \,.$$ □

Korollar (**Absolutstetigkeit des Integrals**). *Sei* $(\Omega, \mathscr{A}, \mu)$ *ein Maßraum. Für* μ*-summierbares* $f \colon \Omega \to \overline{\mathbb{R}}$ *gibt es zu jedem* $\varepsilon \in \mathbb{R}_{>0}$ *ein* $\delta \in \mathbb{R}_{>0}$*, so dass für alle* $B \in \mathscr{A}$ *gilt:*

$$\mu(B) < \delta \implies \int_B |f| \,\mathrm{d}\mu < \varepsilon \,.$$

Beweis. Sei $f_k := \min\{|f|, k\}$. Gemäß dem Satz über dominierte Konvergenz mit Majorante $|f|$ gibt es ein $k \in \mathbb{N}$ mit

$$\int_\Omega (|f| - f_k)\,\mathrm{d}\mu < \frac{\varepsilon}{2}\,.$$

Wählt man jetzt $\delta := \frac{\varepsilon}{2k}$, so gilt im Fall $\mu(B) < \delta$ stets

$$\int_B |f|\,\mathrm{d}\mu = \int_B f_k\,\mathrm{d}\mu + \int_B (|f| - f_k)\,\mathrm{d}\mu \leq k\mu(B) + \frac{\varepsilon}{2} < \varepsilon\,. \qquad \square$$

In Anbetracht des vorigen Korollars stellt sich der Satz über dominierte Konvergenz übrigens als Spezialfall des Konvergenzsatzes von Vitali heraus.

Korollar (über **stetige Abhängigkeit von Parametern**). *Seien $(\Omega, \mathscr{A}, \mu)$ ein Maßraum, P ein metrischer (Parameter-)Raum und $f\colon \Omega \times P \to \overline{\mathbb{R}}$ eine Abbildung. Ist dann einerseits $f(x, \cdot)\colon P \to \overline{\mathbb{R}}$ für μ-fast-alle $x \in \Omega$ eine stetige Funktion und andererseits $f(\cdot, p)\colon \Omega \to \overline{\mathbb{R}}$ für alle $p \in \Omega$ eine μ-summierbare Funktion, die μ-fast-überall $|f(\cdot, p)| \leq g$ für eine von p unabhängige μ-summierbare Majorante g erfüllt, so ist*

$$P \to \mathbb{R},\ p \mapsto \int_\Omega f(x, p)\,\mathrm{d}\mu(x)$$

eine stetige Funktion.

Beweis. Nach dem Satz über dominierte Konvergenz gilt für $p_k \to p$ in P stets

$$\lim_{k\to\infty} \int_\Omega f(x, p_k)\,\mathrm{d}\mu(x) = \int_\Omega f(x, p)\,\mathrm{d}\mu(x)\,.$$

Damit hängt das Integral Folgen-stetig, also stetig von $p \in P$ ab. $\square$

Hängen Integranden differenzierbar von einem Parameter ab und existiert eine Parameter-unabhängige und summierbare Majorante für die Ableitung, so erhält man genau wie in der Riemannschen Integrationstheorie auch die **differenzierbare Abhängigkeit des Integrals von Parametern**. Diese Aussage wird hier aber nicht im Detail ausformuliert.

2.8 Produktmaße und der Satz von Fubini

In diesem Abschnitt wird das Produkt zweier Maßräume $(\mathcal{X}, \mathscr{A}, \mu)$ und $(\mathcal{Y}, \mathscr{S}, \nu)$ über dem kartesischen Produkt $\mathcal{X}\times\mathcal{Y}$ als neuem Grundraum erklärt. Dazu gilt es, geeignete Produkt-Operationen mit σ-Algebren und Maßen zu definieren. Vorbereitend werden aber zunächst etwas einfachere Bildungen für Halbringe und Prämaße behandelt:

Definition & Satz (**Produkte von Halbringen und Prämaßen**). *Seien $\mathcal{X}$ und $\mathcal{Y}$ beliebige Mengen, $\mathscr{A} \subset \mathcal{P}(\mathcal{X})$ und $\mathscr{S} \subset \mathcal{P}(\mathcal{Y})$ beliebige Mengensysteme. Dann setzt man*

$$\mathscr{A}\times\mathscr{S} := \{A\times S : A \in \mathscr{A} \text{ und } S \in \mathscr{S}\} \subset \mathcal{P}(\mathcal{X}\times\mathcal{Y})\,.$$

Sind $\mathscr{A}$ und $\mathscr{S}$ Halbringe, so ist auch $\mathscr{A}\times\mathscr{S}$ ein Halbring, und sind außerdem $\mu\colon \mathscr{A} \to [0, \infty]$ und $\nu\colon \mathscr{S} \to [0, \infty]$ Prämaße, so definiert die Festlegung

$$(\mu\times\nu)(A\times S) := \mu(A)\,\nu(S) \qquad \text{für alle } A \in \mathscr{A} \text{ und } S \in \mathscr{S}$$

ein Prämaß $\mu\times\nu\colon \mathscr{A}\times\mathscr{S} \to [0, \infty]$, das als ***Produktprämaß*** *von μ und ν bezeichnet wird.*

Notation (für **Schnitte**). *Für $G \subset \mathcal{X}\times\mathcal{Y}$, $x \in \mathcal{X}$ und $y \in \mathcal{Y}$ vereinbart man die Schreibweisen*

$$\begin{aligned} {}_xG &:= \{y \in \mathcal{Y} \,:\, (x,y) \in G\} \subset \mathcal{Y}\,, \\ G_y &:= \{x \in \mathcal{X} \,:\, (x,y) \in G\} \subset \mathcal{X}\,. \end{aligned}$$

Speziell für Mengen von Produktgestalt $A\times S \in \mathscr{A}\times\mathscr{S} \subset \mathcal{P}(\mathcal{X}\times\mathcal{Y})$ ergeben sich damit die Gleichheiten

$$\nu({}_x(A\times S)) = \mathbb{1}_A(x)\,\nu(S) \qquad \text{und} \qquad \mu((A\times S)_y) = \mu(A)\,\mathbb{1}_S(y)$$

für alle $x \in \mathcal{Y}$ beziehungsweise $y \in \mathcal{Y}$.

Beweis des Satzes. Der Beweis der Halbringeigenschaft von $\mathscr{A}\times\mathscr{S}$ ist eine abstrakte Variante des früheren Nachweises, dass $\mathcal{I}_2$ ein Halbring ist, und wird deshalb nicht explizit ausgeführt.

Für das Produktprämaß $\mu\times\nu$ ist nur σ-Additivität zu verifizieren. Dazu betrachtet man $A, A_1, A_2, A_3, \ldots \in \mathscr{A}$ und $S, S_1, S_2, S_3, \ldots \in \mathscr{S}$, so dass die $A_i\times S_i$ disjunkt sind und

$$A\times S = \bigcup_{i=1}^{\infty}(A_i\times S_i)$$

gilt. Mit der vorausgehenden Notation und dem Korollar zu Funktionenreihen in Abschnitt 2.7 ergibt sich (wobei $\widehat{\mu}$ irgendeine Fortsetzung von μ zu einem Maß bezeichnet und die Existenz einer solchen durch den Maßfortsetzungssatz gesichert ist)

$$\begin{aligned} (\mu\times\nu)(A\times S) = \mu(A)\,\nu(S) &= \int_{\mathcal{X}} \mathbb{1}_A(x)\,\nu(S)\,\mathrm{d}\widehat{\mu}(x) = \int_{\mathcal{X}} \nu({}_x(A\times S))\,\mathrm{d}\widehat{\mu}(x) \\ &= \int_{\mathcal{X}} \nu\bigg(\bigcup_{i=1}^{\infty} {}_x(A_i\times S_i)\bigg)\,\mathrm{d}\widehat{\mu}(x) = \int_{\mathcal{X}} \sum_{i=1}^{\infty} \nu({}_x(A_i\times S_i))\,\mathrm{d}\widehat{\mu}(x) \\ &= \sum_{i=1}^{\infty} \int_{\mathcal{X}} \nu({}_x(A_i\times S_i))\,\mathrm{d}\widehat{\mu}(x) = \sum_{i=1}^{\infty} \int_{\mathcal{X}} \mathbb{1}_{A_i}(x)\,\nu(S_i)\,\mathrm{d}\widehat{\mu}(x) \\ &= \sum_{i=1}^{\infty} \mu(A_i)\,\nu(S_i) = \sum_{i=1}^{\infty} (\mu\times\nu)(A_i\times S_i)\,. \end{aligned}$$

□

Produkte von Halbringen und Prämaßen sind aber nur ein erster Schritt auf dem Weg zu den wichtigeren Produktbildungen bei σ-Algebren und bei Maßen. Tatsächlich ist für zwei σ-Algebren $\mathscr{A}$ und $\mathscr{S}$ das einfache Produkt $\mathscr{A}\times\mathscr{S}$ im Allgemeinen keine σ-Algebra[28], weshalb man zur von $\mathscr{A}\times\mathscr{S}$ erzeugten σ-Algebra übergeht:

Definitionen (Produkt-σ-Algebren, Produktmaße). *Seien $(\mathcal{X},\mathscr{A})$ und $(\mathcal{Y},\mathscr{S})$ Messräume.*

- *Man nennt die σ-Algebra*

$$\mathscr{A}\otimes\mathscr{S} := \sigma(\mathscr{A}\times\mathscr{S}) \subset \mathcal{P}(\mathcal{X}\times\mathcal{Y})$$

  *über $\mathcal{X}\times\mathcal{Y}$ die **Produkt-σ-Algebra** von $\mathscr{A}$ und $\mathscr{S}$.*

[28] Tatsächlich ist $\mathscr{A}\times\mathscr{S}$ sogar *nur* dann eine σ-Algebra, wenn $\mathscr{A} = \{\emptyset,\mathcal{X}\}$ oder $\mathscr{S} = \{\emptyset,\mathcal{Y}\}$ gilt. Andernfalls ist $\mathscr{A}\times\mathscr{S}$ nicht abgeschlossen bezüglich Komplementbildung, da sich Komplemente von nichttrivialen kartesischen Produkten nicht wieder als kartesische Produkte schreiben lassen.

- *Sind μ ein Maß auf $(\mathcal{X},\mathscr{A})$ und ν ein Maß auf $(\mathcal{Y},\mathscr{S})$, so heißt das mit Hilfe der Carathéodory-Konstruktion gebildete Maß*

$$(\mu{\otimes}\nu) := (\mu{\times}\nu)^*|_{\mathscr{A}\otimes\mathscr{S}}$$

*auf $(\mathcal{X}{\times}\mathcal{Y},\mathscr{A}{\otimes}\mathscr{S})$ das **Produktmaß** von μ und ν.*

Bemerkungen (zu Produkt-σ-Algebren)**.** Seien $(\mathcal{X},\mathscr{A})$ und $(\mathcal{Y},\mathscr{S})$ Messräume.

(1) Sind $\mathscr{E}$ ein Erzeuger von $\mathscr{A}$ und $\mathscr{F}$ ein Erzeuger von $\mathscr{S}$, so ist $\mathscr{E}{\times}\mathscr{F}$ ein Erzeuger von $\mathscr{A}{\otimes}\mathscr{S}$, mit anderen Worten gilt also $\sigma(\mathscr{E}{\times}\mathscr{F}) = \sigma(\mathscr{E}){\otimes}\sigma(\mathscr{F})$.

Beweis. Aus $\mathscr{E}{\times}\mathscr{F} \subset \sigma(\mathscr{E}){\times}\sigma(\mathscr{F})$ folgt sofort $\sigma(\mathscr{E}{\times}\mathscr{F}) \subset \sigma(\sigma(\mathscr{E}){\times}\sigma(\mathscr{F})) = \sigma(\mathscr{E}){\otimes}\sigma(\mathscr{F})$. Zum Nachweis der Umkehrinklusion zeigt man erst, dass $\{A \in \mathcal{P}(\mathcal{X}) : A{\times}F \in \sigma(\mathscr{E}{\times}\mathscr{F})\}$ für festes $F \in \mathscr{F}$ eine σ-Algebra über $\mathcal{X}$ ist, die mit $\mathscr{E}$ auch $\sigma(\mathscr{E})$ enthält. Somit gilt $\sigma(\mathscr{E}){\times}\mathscr{F} \subset \sigma(\mathscr{E}{\times}\mathscr{F})$, und zusammen mit der analogen Regel $\sigma(\mathscr{E}){\times}\sigma(\mathscr{F}) \subset \sigma(\sigma(\mathscr{E}){\times}\mathscr{F})$ folgt $\sigma(\mathscr{E}){\otimes}\sigma(\mathscr{F}) = \sigma(\sigma(\mathscr{E}){\times}\sigma(\mathscr{F})) \subset \sigma(\sigma(\mathscr{E}){\times}\mathscr{F}) \subset \sigma(\mathscr{E}{\times}\mathscr{F})$. □

(2) **Schnitte messbarer Mengen sind messbar**, genauer gelten für $G \in \mathscr{A}{\otimes}\mathscr{S}$, $x \in \mathcal{X}$ und $y \in \mathcal{Y}$ stets ${}_xG \in \mathscr{S}$ und $G_y \in \mathscr{A}$.

Beweis. Bei $\{G \in \mathcal{P}(\mathcal{X}{\times}\mathcal{Y}) : {}_xG \in \mathscr{S}$ für alle $x \in \mathcal{X}$, $G_y \in \mathscr{A}$ für alle $y \in \mathcal{Y}\}$ handelt es sich um eine σ-Algebra über $\mathcal{X}{\times}\mathcal{Y}$, die $\mathscr{A}{\times}\mathscr{S}$ und folglich auch $\mathscr{A}{\otimes}\mathscr{S}$ enthält. □

(e) Ist $(\mathcal{Z},\mathcal{G})$ ein weiterer Messraum und $f\colon \mathcal{X}{\times}\mathcal{Y} \to \mathcal{Z}$ eine $(\mathscr{A}{\otimes}\mathscr{S},\mathscr{G})$-messbare Funktion, so ist $f(x,\,\cdot\,)\colon \mathcal{Y} \to \mathcal{Z}$ mit $x \in \mathcal{X}$ stets $(\mathscr{S},\mathscr{G})$-messbar, und $f(\,\cdot\,,y)\colon \mathcal{X} \to \mathcal{Z}$ mit $y \in \mathcal{Y}$ ist $(\mathscr{A},\mathscr{G})$-messbar.

Beweis. Für $G \in \mathscr{G}$ ist $f(x,\cdot)^{-1}(G) = {}_x(f^{-1}(G)) \in \mathscr{S}$ und $f(\cdot,y)^{-1}(G) = (f^{-1}(G))_y \in \mathscr{A}$. □

Bemerkungen (zur Definition des Produktmaßes)**.**

(3) Die aus Abschnitt 2.4 bekannten Eigenschaften der Carathéodory-Konstruktion garantieren

$$(\mu{\otimes}\nu)(A{\times}S) = (\mu{\times}\nu)(A{\times}S) = \mu(A)\,\nu(S) \qquad \text{für } A \in \mathscr{A} \text{ und } S \in \mathscr{S}$$

und $\mathscr{A}{\otimes}\mathscr{S} = \sigma(\mathscr{A}{\times}\mathscr{S}) \subset \mathcal{M}_{(\mu\times\nu)^*}$. Aus dem ersten Satz des Abschnitts 2.4 (zur Einschränkung von äußeren Maßen auf ihre messbaren Mengen) erhält man deshalb, dass $\mu{\otimes}\nu$ tatsächlich ein(e Fortsetzung von $\mu{\times}\nu$ zu einem) Maß auf $(\mathcal{X}{\times}\mathcal{Y},\mathscr{A}{\otimes}\mathscr{S})$ ist.

Für σ-endliche Maße μ, ν ist $\mu{\times}\nu$ ebenfalls σ-endlich, und dann ist die Fortsetzung von $\mu{\times}\nu$ zu einem Maß auf $(\mathcal{X}{\times}\mathcal{Y},\mathscr{A}{\otimes}\mathscr{S})$ gemäß dem Maßfortsetzungssatz auch *eindeutig bestimmt*. Im Allgemeinen ist die Verwendung von Produktmaßen nur in dieser σ-endlichen Situation, in der auch Eindeutigkeit vorliegt, sinnvoll.

Ein Hauptgrund für die Bedeutung von Produktmaßen liegt darin, dass **Lebesgue-Maße von Produktstruktur** sind und genauer das Lebesgue-Maß $\mathscr{L}^N$ mit $N \geq 2$ das Produktmaß nieder-dimensionaler Lebesgue-Maße ist:

Satz (Produkte von Lebesgue-Maßen). *Seien $k, \ell \in \mathbb{N}$.*

(I) *Für Produkte von Borel-σ-Algebren*[29] *und Lebesgue-Maßen auf Borel-Mengen gelten*

$$\mathcal{B}(\mathbb{R}^k)\otimes\mathcal{B}(\mathbb{R}^\ell) = \mathcal{B}(\mathbb{R}^{k+\ell}) \qquad \textit{und} \qquad \mathscr{L}^k\otimes\mathscr{L}^\ell = \mathscr{L}^{k+\ell}.$$

(II) *Bei den σ-Algebren Lebesgue-messbarer Mengen gilt*

$$\mathcal{M}^k\otimes\mathcal{M}^\ell \subsetneqq \mathcal{M}^{k+\ell},$$

und das vervollständige Lebesgue-Maß $\mathscr{L}^{k+\ell}$ auf $\mathcal{M}^{k+\ell}$ (für das, wie früher vereinbart, die gleiche Notation wie für das Lebesgue-Maß auf Borel-Mengen verwendet wird) ist die Vervollständigung

$$\overline{\mathscr{L}^k\otimes\mathscr{L}^\ell} = \mathscr{L}^{k+\ell}$$

des Produktmaßes $\mathscr{L}^k\otimes\mathscr{L}^\ell$ auf $\mathcal{M}^k\otimes\mathcal{M}^\ell$. Insofern unterscheidet sich $\mathcal{M}^{k+\ell}$ von $\mathcal{M}^k\otimes\mathcal{M}^\ell$ nur durch Nullmengen.

Beweis zu Teil (I) des Satzes. Aus der elementaren Beobachtung $\mathcal{I}_k\times\mathcal{I}_\ell = \mathcal{I}_{k+\ell}$ und der vorausgehenden Bemerkung (1) ergibt sich $\mathcal{B}(\mathbb{R}^k)\otimes\mathcal{B}(\mathbb{R}^\ell) = \sigma(\mathcal{I}_k)\otimes\sigma(\mathcal{I}_\ell) = \sigma(\mathcal{I}_k\times\mathcal{I}_\ell) = \sigma(\mathcal{I}_{k+\ell}) = \mathcal{B}(\mathbb{R}^{k+\ell})$. Die Produktformel $\mathscr{L}^k\otimes\mathscr{L}^\ell = \mathscr{L}^{k+\ell}$ folgt mit dem Eindeutigkeitsteil der Maßfortsetzungssatzes aus der Übereinstimmung von $\lambda^k\times\lambda^\ell$ und $\lambda^{k+\ell}$ auf $\mathcal{I}_k\times\mathcal{I}_\ell = \mathcal{I}_{k+\ell}$. □

Beweis zu Teil (II) des Satzes. Man zeigt zuerst, dass $\mathcal{M}^k\times\mathcal{M}^\ell \subset \mathcal{M}^{k+\ell}$ und als Konsequenz $\mathcal{M}^k\otimes\mathcal{M}^\ell \subset \mathcal{M}^{k+\ell}$ gelten: Sind $A \in \mathcal{M}^k$ und $\widetilde{A} \in \mathcal{M}^\ell$ gegeben, so gibt es Borel-Mengen B und $\widetilde{B}$, eine $\mathscr{L}^k$-Nullmenge $T \in \mathcal{M}^k$ und eine $\mathscr{L}^\ell$-Nullmenge $\widetilde{T} \in \mathcal{M}^\ell$ mit $A = B\cup T$ und $\widetilde{A} = \widetilde{B}\cup\widetilde{T}$. Es folgt

$$A\times\widetilde{A} = \underbrace{(B\times\widetilde{B})}_{\text{Borel-Menge}} \cup \underbrace{(B\times\widetilde{T})\cup(T\times\widetilde{B})\cup(T\times\widetilde{T})}_{\mathscr{L}^{k+\ell}\text{-Nullmenge}} \in \mathcal{M}^{k+\ell}$$

(wobei man die angegebenen Mengen mit Teil (I) des Satzes als eine Borel-Menge und eine $\mathscr{L}^{k+\ell}$-Nullmenge identifiziert). Dass $\mathcal{M}^k\times\mathcal{M}^\ell$ sogar eine *echte* Teilmenge von $\mathcal{M}^{k+\ell}$ ist, erkennt man anhand der mit einer Vitali-Menge $A \in \mathcal{P}(\mathbb{R}^k)\setminus\mathcal{M}^k$ gebildeten Beispielmenge $A\times\{0\}^\ell$. Tatsächlich ist dann $A\times\{0\}^\ell$ als $\mathscr{L}^{k+\ell}$-Nullmenge in $\mathcal{M}^{k+\ell}$ enthalten, kann aber nach der obigen Bemerkung (2) nicht zu $\mathcal{M}^k\times\mathcal{M}^\ell$ gehören.

Zum Beweis der Produktformel für Lebesgue-Maße beobachtet man, dass gemäß Teil (I), der Definition der Vervollständigung und den vorausgehenden Argumenten

$$\mathscr{L}^k\otimes\mathscr{L}^\ell = \mathscr{L}^{k+\ell}|_{\mathcal{M}^k\otimes\mathcal{M}^\ell}$$

gilt. Da die beiden Vervollständigungen $\overline{\mathscr{L}^k\otimes\mathscr{L}^\ell}$ und $\mathscr{L}^{k+\ell}$ jeweils vollständige Fortsetzungen dieses Maßes sind, kann man mit der Minimalitätseigenschaft der Vervollständigung auf $\overline{\mathscr{L}^k\otimes\mathscr{L}^\ell} = \mathscr{L}^{k+\ell}$ schließen. □

Es folgt das zentrale Resultat dieses Abschnitts, das den wohl **wichtigsten und nützlichsten Satz über Produktmaße** überhaupt darstellt.

[29] Allgemeiner bleibt $\mathcal{B}(X)\otimes\mathcal{B}(Y) = \mathcal{B}(X\times Y)$ für Teilmengen $X \subset \mathbb{R}^k$ und $Y \subset \mathbb{R}^\ell$ richtig, wenn X, Y und $X\times Y$ mit der Spurtopologie versehen werden. Noch allgemeiner gilt $\mathcal{B}(\mathcal{X})\otimes\mathcal{B}(\mathcal{Y}) = \mathcal{B}(\mathcal{X}\times\mathcal{Y})$ auch für separable metrische Räume $\mathcal{X}$ und $\mathcal{Y}$ mit der Produkttopologie auf $\mathcal{X}\times\mathcal{Y}$. Für beliebige topologische Räume $\mathcal{X}$ und $\mathcal{Y}$ gilt aber im Allgemeinen nur die Inklusion $\mathcal{B}(\mathcal{X})\otimes\mathcal{B}(\mathcal{Y}) \subset \mathcal{B}(\mathcal{X}\times\mathcal{Y})$; siehe [3, Kapitel III.5] für Weiteres.

Hauptsatz (**Satz von Fubini**(-Tonelli), $\sim$1907)**.** *Für Maßräume* $(\mathcal{X},\mathscr{A},\mu)$, $(\mathcal{Y},\mathscr{S},\nu)$ *mit* σ-*endlichen Maßen* μ, ν *und eine* $\mu\otimes\nu$*-integrierbare Funktion* $f\colon \mathcal{X}\times\mathcal{Y}\to\overline{\mathbb{R}}$ *gelten:*

(I) *Für* μ*-fast-alle* $x\in\mathcal{X}$ *ist* $f(x,\,\cdot\,)\colon\mathcal{Y}\to\overline{\mathbb{R}}$ *eine* ν*-integrierbare Funktion, und*

$$\mathcal{X}\to\overline{\mathbb{R}}\,,\ x\mapsto\int_{\mathcal{Y}} f(x,y)\,\mathrm{d}\nu(y)$$

ist μ*-integrierbar (bei beliebiger Festlegung des Wertes* $\int_{\mathcal{Y}} f(x,y)\,\mathrm{d}\nu(y)$ *auf der* $\overline{\mu}$*-Nullmenge der* x*, für die* $f(x,\,\cdot\,)$ *nicht* ν*-integrierbar ist).*

(II) *Für* ν*-fast-alle* $y\in\mathcal{Y}$ *ist* $f(\,\cdot\,,y)\colon\mathcal{X}\to\overline{\mathbb{R}}$ *eine* μ*-integrierbare Funktion, und*

$$\mathcal{Y}\to\overline{\mathbb{R}}\,,\ y\mapsto\int_{\mathcal{X}} f(x,y)\,\mathrm{d}\mu(x)$$

ist ν*-integrierbar (wieder bei beliebiger Festlegung auf der verbleibenden Nullmenge).*

(III) *Es gilt, dies ist die* **Hauptaussage des Satzes***,*

$$\boxed{\int_{\mathcal{X}\times\mathcal{Y}} f\,\mathrm{d}(\mu\otimes\nu)=\int_{\mathcal{X}}\left[\int_{\mathcal{Y}} f(x,y)\,\mathrm{d}\nu(y)\right]\mathrm{d}\mu(x)=\int_{\mathcal{Y}}\left[\int_{\mathcal{X}} f(x,y)\,\mathrm{d}\mu(x)\right]\mathrm{d}\nu(y)}$$

Ein *Beweis* des Satzes von Fubini folgt in Kürze.

Der Satz besagt, dass **Integrale bezüglich Produktmaßen** sich **als „geschachtelte“ Integrale** bezüglich den beiden Faktoren schreiben lassen. Dabei führen bei den geschachtelten Integralen die beiden möglichen Reihenfolgen der Integration zum gleichen Ergebnis, und insofern sind **Vertauschungen der Integrationsreihenfolge** erlaubt. In der Praxis wird der Satz oft angewandt, um durch Berechnung geschachtelter Integrale das Produktmaß messbarer Teilmengen G — die selbst nicht notwendig Produktgestalt haben müssen — von $\mathcal{X}\times\mathcal{Y}$ zu bestimmen oder Integrale über solche Teilmengen bezüglich Produktmaßen auszurechnen. Dabei kommen Schnitte ${}_xG$ beziehungsweise G_y ins Spiel:

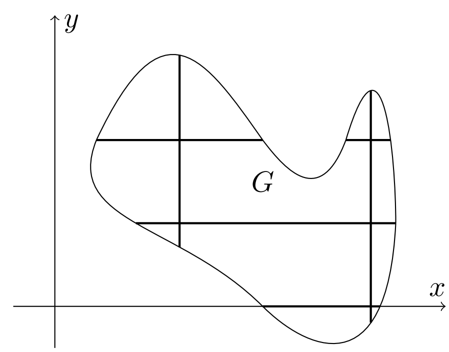


Abb. 4: Zwei vertikale Schnitte ${}_xG$ und drei horizontale Schnitte G_y zu $G\in\mathcal{B}(\mathbb{R}^2)=\mathcal{B}(\mathbb{R})\otimes\mathcal{B}(\mathbb{R})$

Korollar (Satz von Fubini auf Teilmengen)**.** *Unter den Voraussetzungen des Satzes gelten für* $(\mu\otimes\nu)$*-messbare Teilmengen* G *von* $\mathcal{X}\times\mathcal{Y}$ *die Formeln*

$$\boxed{\begin{gathered}\big(\overline{\mu\otimes\nu}\big)(G)=\int_{\mathcal{X}}\overline{\nu}({}_xG)\,\mathrm{d}\mu(x)=\int_{\mathcal{Y}}\overline{\mu}(G_y)\,\mathrm{d}\nu(y)\,,\\ \int_G f\,\mathrm{d}\big(\overline{\mu\otimes\nu}\big)=\int_{\mathcal{X}}\int_{{}_xG} f(x,y)\,\mathrm{d}\overline{\nu}(y)\,\mathrm{d}\mu(x)=\int_{\mathcal{Y}}\int_{G_y} f(x,y)\,\mathrm{d}\overline{\mu}(x)\,\mathrm{d}\nu(y)\,.\end{gathered}}$$

Im Fall $G\in\mathscr{A}\otimes\mathscr{S}$ *können dabei die* $\overline{(\cdot)}$*s weggelassen werden.*

Beweis. Man wendet den Satz mit $\mathbb{1}_G$ beziehungsweise $\mathbb{1}_G f$ anstelle von f an. $\square$

Korollar (Satz von Fubini für Lebesgue-Maße). *Für Lebesgue-Maße auf $\mathbb{R}^{k+\ell}=\mathbb{R}^k\times\mathbb{R}^\ell$ gelten die Formeln*

$$\boxed{\begin{gathered}\mathscr{L}^{k+\ell}(G)=\int_{\mathbb{R}^k}\mathscr{L}^\ell({}_xG)\,\mathrm{d}\mathscr{L}^k(x)=\int_{\mathbb{R}^\ell}\mathscr{L}^k(G_y)\,\mathrm{d}\mathscr{L}^\ell(y)\,,\\ \int_G f\,\mathrm{d}\mathscr{L}^{k+\ell}=\int_{\mathbb{R}^k}\int_{{}_xG}f(x,y)\,\mathrm{d}\mathscr{L}^\ell(y)\,\mathrm{d}\mathscr{L}^k(x)=\int_{\mathbb{R}^\ell}\int_{G_y}f(x,y)\,\mathrm{d}\mathscr{L}^k(x)\,\mathrm{d}\mathscr{L}^\ell(y)\end{gathered}}$$

für $k,\ell\in\mathbb{N}$, $\mathscr{L}^{k+\ell}$-messbare Teilmengen G von $\mathbb{R}^{k+\ell}$ und $\mathscr{L}^{k+\ell}$-integrierbare f.

Beweis. Man kombiniert das vorige Korollar mit dem in diesem Abschnitt schon behandelten Satz über Produkte von Lebesgue-Maßen. □

Unter den Voraussetzungen des Satzes von Fubini kann man die Forderung der σ-Endlichkeit von μ und ν im Wesentlichen als technische Annahme verstehen, die in Anwendungen meist[30] erfüllt ist. Entscheidend ist dagegen die Integrierbarkeitsvoraussetzung an f:

Bemerkungen (zur **Integrierbarkeitsvoraussetzung** bei Fubini)**.** Da (Produkt-)Messbarkeit von f fast immer erfüllt ist (siehe die nächste Serie von Bemerkungen), **schließt die Voraussetzung der $\boldsymbol{\mu\otimes\nu}$-Integrierbarkeit von $\boldsymbol{f}$** im Wesentlichen nur **den Nichtexistenzfall $\boldsymbol{\infty-\infty}$ aus**. Nichtsdestotrotz ist diese Voraussetzung **wesentlich für die Gültigkeit des Satzes von von Fubini** (sowie seiner Korollare), denn das unten folgende Beispiel liefert bereits für $\mu=\nu=\mathscr{L}^1$ einen Beispielintegranden, für den alle iterierten Integrale existieren und einen endlichen Wert ergeben, aber das $\mu\otimes\nu$-Integral undefiniert vom Typ $\infty-\infty$ ist.

Für das praktische Rechnen bedeutet dies, dass nichtnegative (oder nichtpositive) messbare Integranden unproblematisch sind und man den Satz stets auf solche anwenden darf. Nimmt ein Integrand f aber positive *und* negative Werte an, so **ist zuerst sicherzustellen, dass mindestens eines der Integrale $\boldsymbol{\int_{\mathcal{X}\times\mathcal{Y}}f_+\,\mathrm{d}(\overline{\mu\otimes\nu})}$, $\boldsymbol{\int_{\mathcal{X}\times\mathcal{Y}}f_-\,\mathrm{d}(\overline{\mu\otimes\nu})}$, $\boldsymbol{\int_{\mathcal{X}\times\mathcal{Y}}|f|\,\mathrm{d}(\overline{\mu\otimes\nu})}$ endlich ist**. Diese Endlichkeitsbedingungen verifiziert man dabei, indem man den Satz mit den nichtnegativen Integranden f_+, f_- oder $|f|$ verwendet. Erst danach darf man ihn gegebenenfalls auf f selbst anwenden.

Beispiel (zur **Nicht-Anwendbarkeit von Fubini im Fall „$\infty-\infty$“**)**.** Ein **warnendes Beispiel** im Zusammenhang mit dem Satz von Fubini ist das $\mathcal{L}^2$-Integral

$$\int_{[0,1]^2}f(x,y)\,\mathrm{d}\mathcal{L}^2(x,y)\qquad\text{mit Integrand } f(x,y):=\frac{x^2-y^2}{(x^2+y^2)^2}\text{ für } x,y\in\mathbb{R}\,.$$

Tatsächlich folgt aus der Symmetrie $f(y,x)=-f(x,y)$ direkt, dass dieses Integral nur mit Wert 0 existieren oder undefiniert „$\infty-\infty$“ sein kann. Übersieht man dies aber und berechnet die iterierten Integrale, so erhält man

$$\int_0^1\int_0^1 f(x,y)\,\mathrm{d}y\,\mathrm{d}x=\int_0^1\left[\frac{y}{x^2+y^2}\right]_{y=0}^1\mathrm{d}x=\int_0^1\frac{1}{1+x^2}\,\mathrm{d}x=\big[\arctan x\big]_{x=0}^1=\frac{\pi}{4}\,,$$
$$\int_0^1\int_0^1 f(x,y)\,\mathrm{d}x\,\mathrm{d}y=\int_0^1\left[\frac{-x}{x^2+y^2}\right]_{x=0}^1\mathrm{d}y=\int_0^1\frac{-1}{1+y^2}\,\mathrm{d}y=\big[-\arctan y\big]_{y=0}^1=-\frac{\pi}{4}\,.$$

[30] Eine gewisse Vorsicht ist aber beispielsweise im Zusammenhang mit den demnächst eingeführten und nicht generell σ-endlichen Hausdorff-Maßen geboten. Auf diese Maße kann der Satz nur in solchen Fällen angewandt werden, in denen man sich auf σ-endliche Mengen zurückziehen kann.

Nimmt man nur eine dieser beiden Berechnungen vor, so scheinen Rechnung und Ergebnis zunächst unverdächtig. Erst wenn man beide Rechnungen durchführt und die nicht übereinstimmenden Ergebnisse erhält (oder natürlich, wenn man die Integrierbarkeitsvoraussetzung für Fubini genau prüft), wird klar, dass keines der beiden Ergebnisse den Wert des $\mathcal{L}^2$-Integrals gibt. Es muss tatsächlich die Situation vorliegen, in der der **Satz von Fubini nicht anwendbar** ist: Das $\mathcal{L}^2$-Integral ist **undefiniert vom Typ $\infty - \infty$**.

Als Nächstes folgen einige Bemerkungen zur Produktmessbarkeit, die hier allgemein für $\mathscr{A} \otimes \mathscr{S}$-Messbarkeit mit beliebigen σ-Algebren $\mathscr{A}$ und $\mathscr{S}$ über $\mathcal{X}$ und $\mathcal{Y}$ formuliert werden. Genau wie $\mathcal{M}^k \otimes \mathcal{M}^\ell \subset \mathcal{M}^{k+\ell}$ sieht man auch $\mathcal{M}_\mu \otimes \mathcal{M}_\nu \subset \mathcal{M}_{\mu\otimes\nu}$ für allgemeine Maße μ und ν ein, daher lassen sich aus diesen Bemerkungen auch Kriterien gewinnen, die die in der $(\mu\otimes\nu)$-Integrierbarkeitsvoraussetzung bei Fubini enthaltene $\mu\otimes\nu$-Messbarkeit für alle praktisch relevanten Fälle sichern.

Bemerkungen (zu Produkmessbarkeit)**.** Seien $(\mathcal{X}, \mathscr{A})$, $(\mathcal{Y}, \mathscr{S})$, $(\mathcal{W}, \mathscr{G})$, $(\mathcal{Z}, \mathscr{H})$ Messräume.

(1) Ein allgemeines Produkt[31] $f \times g \colon \mathcal{X} \times \mathcal{Y} \to \mathcal{W} \times \mathcal{Z}$ zweier Abbildungen $f \colon \mathcal{X} \to \mathcal{W}$ und $g \colon \mathcal{Y} \to \mathcal{Z}$ ist gegeben durch

$$(f \times g)(x, y) := (f(x), g(y)) \in \mathcal{W} \times \mathcal{Z} \qquad \text{für } x \in \mathcal{X} \text{ und } y \in \mathcal{Y}.$$

Für eine $(\mathscr{A}, \mathscr{G})$-messbare Abbildung $f \colon \mathcal{X} \to \mathcal{W}$ und eine $(\mathscr{S}, \mathscr{H})$-messbare Abbildung $g \colon \mathcal{Y} \to \mathcal{Z}$ ist dieses Produkt $f \times g$ stets $(\mathscr{A} \otimes \mathscr{S}, \mathscr{G} \otimes \mathscr{H})$-messbar.

Beweis der Messbarkeitsaussage. Aus der Definition von $f \times g$ ergibt sich $(f \times g)^{-1}(\mathscr{G} \times \mathscr{H}) \subset \mathscr{A} \times \mathscr{S} \subset \mathscr{A} \otimes \mathscr{S}$. Da $\mathscr{G} \times \mathscr{H}$ per Definition ein Erzeuger von $\mathscr{G} \otimes \mathscr{H}$ ist, folgt $(\mathscr{A} \otimes \mathscr{S}, \mathscr{G} \otimes \mathscr{H})$-Messbarkeit von $f \times g$ gemäß Bemerkung (3) zur Messbarkeitsdefinition in Abschnitt 2.5. □

(2) Ist für $\mathscr{A}$-messbares $g \colon \mathcal{X} \to \overline{\mathbb{R}}$ und $\mathscr{S}$-messbares $h \colon \mathcal{Y} \to \overline{\mathbb{R}}$ beispielsweise durch

$$f(x, y) := g(x) + h(y) \qquad \text{oder} \qquad f(x, y) := \frac{g(x)}{h(y)} \qquad \text{oder} \qquad f(x, y) := \max\{g(x), h(y)\}$$

(mit irgendeiner Konvention bei undefinierten Ausdrücken) oder ganz allgemein für $\mathscr{A}$-messbares $g \colon \mathcal{X} \to \overline{\mathbb{R}}^k$ und $\mathscr{S}$-messbares $h \colon \mathcal{Y} \to \overline{\mathbb{R}}^\ell$ durch

$$f(x, y) := b(g(x), h(y)) \qquad \text{mit einer Borel-Funktion } b \colon \overline{\mathbb{R}}^k \times \overline{\mathbb{R}}^\ell \to \overline{\mathbb{R}}$$

eine Funktion $f \colon \mathcal{X} \times \mathcal{Y} \to \overline{\mathbb{R}}$ gegeben, so ist f stets $\mathscr{A} \otimes \mathscr{S}$-messbar. Diese Regel liefert die **Produktmessbarkeit aller praktisch relevanten Funktionen**.

Beweis. Es liegen $(\mathscr{A} \otimes \mathscr{S}, \mathcal{B}(\overline{\mathbb{R}}^k) \otimes \mathcal{B}(\overline{\mathbb{R}}^\ell))$-Messbarkeit von $g \times h$ (vorige Bemerkung) und $(\mathcal{B}(\overline{\mathbb{R}}^k \times \overline{\mathbb{R}}^\ell), \mathcal{B}(\overline{\mathbb{R}}))$-Messbarkeit von b vor. Mit der Produktstruktur $\mathcal{B}(\overline{\mathbb{R}}^k \times \overline{\mathbb{R}}^\ell) = \mathcal{B}(\overline{\mathbb{R}}^k) \otimes \mathcal{B}(\overline{\mathbb{R}}^\ell)$ der Borel-σ-Algebra und der Kompositionsregel aus Abschnitt 2.5 folgt $(\mathscr{A} \otimes \mathscr{S}, \mathcal{B}(\overline{\mathbb{R}}))$-Messbarkeit von $f = b \circ (g \times h)$. □

(3) Für $\mathscr{A}$-messbare $f_1, f_2, \ldots f_M \colon \Omega \to \overline{\mathbb{R}}$ ist auch $(f_1, f_2, \ldots, f_M) \colon \Omega \to \overline{\mathbb{R}}^M$ stets $\mathscr{A}$-messbar. Dies wurde bereits in Fußnote 20 bemerkt, lässt sich aber jetzt systematisch begründen: Man schreibt $(f_1, f_2, \ldots, f_M) = (f_1 \times f_2 \times \ldots \times f_M) \circ \Delta_\Omega^M$ dazu als Komposition der $(\mathscr{A}, \mathscr{A} \otimes \mathscr{A} \otimes \ldots \otimes \mathscr{A})$-messbaren Diagonalenabbildung $\Delta_\Omega^M \colon \Omega \to \Omega^M$, $x \mapsto (x, x, \ldots, x)$ und der $(\mathscr{A} \otimes \mathscr{A} \otimes \ldots \otimes \mathscr{A}, \mathcal{B}(\overline{\mathbb{R}}) \otimes \mathcal{B}(\overline{\mathbb{R}}) \otimes \ldots \otimes \mathcal{B}(\overline{\mathbb{R}}))$-messbaren Produktabbildung $f_1 \times f_2 \times \ldots \times f_M$. Dann folgt die $\mathscr{A}$-Messbarkeit von $(f_1, f_2, \ldots, f_M)$ direkt aus der Kompositionsregel und $\mathcal{B}(\overline{\mathbb{R}}) \otimes \mathcal{B}(\overline{\mathbb{R}}) \otimes \ldots \otimes \mathcal{B}(\overline{\mathbb{R}}) = \mathcal{B}(\overline{\mathbb{R}}^M)$.

Beweisskizze zum Satz von Fubini. Es wurde bereits gezeigt, dass Schnitte (Produkt-)messbarer Mengen messbar sind, dass also für $G \in \mathscr{A} \otimes \mathscr{S}$ und $x \in \mathcal{X}$ stets ${}_xG \in \mathscr{S}$ gilt und $\nu({}_xG)$ somit definiert ist. Nun wird durch schrittweise Argumentation erst für spezielle, dann für immer allgemeinere Mengen $G \in \mathscr{A} \otimes \mathscr{S}$ nachgewiesen, dass zudem gilt:

$$x \mapsto \nu({}_xG) \text{ ist } \mu\text{-messbar} \qquad \text{mit} \qquad (\mu \otimes \nu)(G) = \int_{\mathcal{X}} \nu({}_xG) \, \mathrm{d}\mu(x). \tag{$*$}$$

[31] Für diese Art der Produktbildung ist neben $\times$ auch das Symbol $\bowtie$ gebräuchlich.

Schritt 1: Für kartesische Produkte $G \in \mathscr{A} \times \mathscr{S}$ erhält man $(*)$ aus den ersten Überlegungen zu Schnitten.

Schritt 2: Sei

$$(\mathscr{A} \times \mathscr{S})_\sigma := \left\{ \bigcup_{i=1}^{\infty} G_i \,:\, G_i \in \mathscr{A} \times \mathscr{S} \text{ disjunkt} \right\} \subset \mathscr{A} \otimes \mathscr{S}\,.$$

Für $G \in (\mathscr{A} \times \mathscr{S})_\sigma$ erhält man $(*)$ aus der Darstellung $\nu\big({}_x(\bigcup_{i=1}^\infty G_i)\big) = \sum_{i=1}^\infty \nu({}_xG_i)$, der Proposition zur Messbarkeitserhaltung aus Abschnitt 2.5 und dem Korollar zu Funktionenreihen aus Abschnitt 2.7.

Schritt 3: Sei

$$(\mathscr{A} \times \mathscr{S})_{\sigma\delta} := \left\{ \bigcap_{i=1}^{\infty} G_i \,:\, G_1 \supset G_2 \supset \ldots \text{ sind } \mu\otimes\nu\text{-endlich in } (\mathscr{A} \times \mathscr{S})_\sigma \right\} \subset \mathscr{A} \otimes \mathscr{S}\,.$$

Für μ-fast-alle $x \in \mathcal{X}$ folgt aus $(\mu\otimes\nu)(G_1) < \infty$ und Schritt 2 schon $\nu({}_xG_1) < \infty$, und daraus erhält man $\nu\big({}_x(\bigcap_{i=1}^\infty G_i)\big) = \lim_{i\to\infty} \nu({}_xG_i)$. Mit der Proposition aus Abschnitt 2.5 und dominierter (oder auch monotoner) Konvergenz schließt man auf die Gültigkeit von $(*)$ für alle $G \in (\mathscr{A} \times \mathscr{S})_{\sigma\delta}$.

Schritt 4: Zu jedem $\mu\otimes\nu$-endlichen $G \in \mathscr{A} \otimes \mathscr{S}$ existiert eine $\mu\otimes\nu$-maßgleiche Obermenge $\widetilde{G} \in (\mathscr{A} \times \mathscr{S})_{\sigma\delta}$ (man vergleiche hierzu die Bemerkungen zur Carathéodory-Konstruktion). Deshalb folgt $(*)$ erst für alle $\mu\otimes\nu$-Nullmengen G und dann für alle $\mu\otimes\nu$-endlichen $G \in \mathscr{A} \otimes \mathscr{S}$ (denn für letztere ist $\widetilde{G} \setminus G$ eine $\mu\otimes\nu$-Nullmenge).

Schritt 5: Unter Verwendung der σ-Endlichkeit von $\mu\otimes\nu$ ergibt sich $(*)$ jetzt für alle $G \in \mathscr{A} \otimes \mathscr{S}$: Dazu schreibt man G als abzählbare disjunkte Vereinigung endlicher Mengen und argumentiert dann wie in Schritt 2. Aus Symmetriegründen gelten analoge Aussagen für die umgekehrte Integrationsreihenfolge, und damit ist der Satz von Fubini für charakteristische Funktionen $f = \mathbb{1}_G$ zu $G \in \mathscr{A} \otimes \mathscr{S}$ gezeigt.

Schritt 6: Die Verallgemeinerung auf beliebige Integranden bewerkstelligt man mit der Standard-Ausdehnungsprozedur und verwendet dabei das Approximationslemma aus Abschnitt 2.5 und wieder einmal den Satz über monotone Konvergenz. □

Durch iterierte Anwendung des Satzes von Fubini gelangt man zur nächsten Folgerung, die die Berechnung von $\mathscr{L}^N$-Maßen und $\mathscr{L}^N$-Integralen auf N geschachtelte $\mathscr{L}^1$-Integrale zurückführt und viele explizite Integralberechnungen erlaubt.

Korollar (iterierter Satz von Fubini). *Die Produktmaß-Operation $\otimes$ ist assoziativ, auf Borel-Mengen gilt*

$$\underbrace{\mathscr{L}^1 \otimes \mathscr{L}^1 \otimes \ldots \otimes \mathscr{L}^1}_{N-mal} = \mathscr{L}^N\,,$$

und für $\mathscr{L}^N$-integrierbare f gilt

.

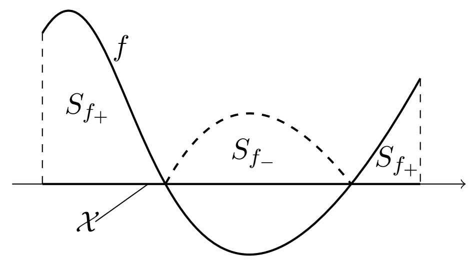
$$\int_{\mathbb{R}^N} f \,\mathrm{d}\mathscr{L}^N = \underbrace{\int_{\mathbb{R}} \ldots \int_{\mathbb{R}} \int_{\mathbb{R}}}_{N \text{ Integrale}} f(x_1, x_2, \ldots, x_N) \,\mathrm{d}\mathscr{L}^1(x_1) \,\mathrm{d}\mathscr{L}^1(x_2) \ldots \mathrm{d}\mathscr{L}^1(x_N)\,. \qquad \square$$

Auf eine ausführlichere Diskussion iterierter Produktmaße und des iterierten Satzes von Fubini wird hier verzichtet. Stattdessen werden noch ein weiteres Korollar und eine konkrete Anwendung behandelt:

Eine bekannte Grundidee der Integrationstheorie besteht darin, dass 1-dimensionale Integrale den Flächeninhalt unter dem Graphen des Integranden f und allgemeiner N-dimensionale Integrale den $(N{+}1)$-dimensionalen Inhalt unter dem Graphen messen. Für allgemeinere Maßräume lässt sich die Vorstellung des Inhalts unter dem Graphen durch ein Produktmaß des Subgraphen S_f verallgemeinern, wobei S_f für nichtnegatives f auf $\mathcal{X}$ als

$$S_f := \{(x,y) \in \mathcal{X}{\times}\mathbb{R} \,:\, 0 \le y < f(x)\} \subset \mathcal{X} \times \mathbb{R}\,,$$

definiert ist. Genauer erhält man:



Korollar (Integral als Maß des Subgraphen). *Sei $(\mathcal{X},\mathscr{A},\mu)$ ein σ-endlicher Maßraum. Für μ-integrierbares $f\colon \mathcal{X} \to \overline{\mathbb{R}}$ sind S_{f_+} und S_{f_-} stets in $\mathcal{M}_\mu{\otimes}\mathcal{B}(\mathbb{R})$, und es gilt*

$$\int_{\mathcal{X}} f \,\mathrm{d}\mu = (\overline{\mu}{\otimes}\mathscr{L}^1)(S_{f_+}) - (\overline{\mu}{\otimes}\mathscr{L}^1)(S_{f_-})\,.$$

Speziell für $\mathscr{L}^N$-integrierbares $f\colon D \to \overline{\mathbb{R}}$ mit $D \in \mathcal{M}^N$ ergibt sich

$$\int_D \,\mathrm{d}\mathscr{L}^N = \mathscr{L}^{N+1}(S_{f_+}) - \mathscr{L}^{N+1}(S_{f_-})\,.$$

Beweis. Ohne Einschränkung sei f nichtnegativ. Aus $\mathbb{1}_{S_f} = \mathbb{1}_{\{(w,y)\in\overline{\mathbb{R}}\times\mathbb{R}\,:\,0\le y<w\}} \circ (f{\times}\mathrm{id}_{\mathbb{R}})$ entnimmt man $S_f \in \mathcal{M}_\mu{\otimes}\mathcal{B}(\mathbb{R})$. Der Satz von Fubini liefert daher

$$\begin{aligned}(\overline{\mu}{\otimes}\mathscr{L}^1)(S_f) &= \int_{\mathcal{X}} \mathscr{L}^1\big({}_x(S_f)\big)\,\mathrm{d}\mu(x)\\ &= \int_{\mathcal{X}} \mathscr{L}^1(\,[0,f(x))\,)\,\mathrm{d}\mu(x) = \int_{\mathcal{X}} f(x)\,\mathrm{d}\mu(x)\,.\end{aligned}$$ □

Schließlich folgt die angekündigte Anwendung:

Anwendung (Berechnung des Einheitskugelvolumens $\boldsymbol{\omega_N}$ mit Fubini). Im Folgenden wird eine für alle $N \in \mathbb{N}$ gültige Formel für das Volumen $\omega_N = \mathscr{L}^N(\mathrm{B}_1^N)$ der N-dimensionalen Einheitskugel $\mathrm{B}_1^N = \mathrm{B}_1^N(0) \subset \mathbb{R}^N$ hergeleitet. Das Volumen einer allgemeinen Kugel erhält man dann gemäß Translationsinvarianz und Skalierungsverhalten von $\mathscr{L}^N$ als $\mathscr{L}^N(\mathrm{B}_r^N(x)) = \omega_N r^N$.

(0) Für $N = 1$ ist $\mathrm{B}_1^1 = (-1,1)$ und daher $\omega_1 = 2$.

(1) Für $N = 2$ liefert der Satz von Fubini

$$\begin{aligned}\omega_2 = \mathscr{L}^2(\mathrm{B}_1^2) &= \int_{\mathbb{R}} \mathscr{L}^1({}_x(\mathrm{B}_1^2))\,\mathrm{d}\mathscr{L}^1(x)\\ &= \int_{-1}^1 \mathscr{L}^1\big(\big(-\sqrt{1-x^2},\sqrt{1-x^2}\big)\big)\,\mathrm{d}\mathscr{L}^1(x) = 2\int_{-1}^1 \sqrt{1-x^2}\,\mathrm{d}x \overset{x=\sin\varphi}{=} \pi\,.\end{aligned}$$

(2) Für $N \geq 3$ erhält man aus $\mathscr{L}^2(\mathrm{B}_r^2) = \pi r^2$, dem Satz von Fubini und der Formel für radiale Integration wie folgt eine Rekursionsformel in Zweier-Schritten:

$$\begin{aligned}\omega_N = \mathscr{L}^N(\mathrm{B}_1^N) &= \int_{\mathbb{R}^{N-2}} \mathscr{L}^2\big({}_x(\mathrm{B}_1^N)\big)\,\mathrm{d}\mathscr{L}^{N-2}(x)\\ &= \int_{\mathrm{B}_1^{N-2}} \mathscr{L}^2\Big(\mathrm{B}^2_{\sqrt{1-|x|^2}}\Big)\,\mathrm{d}\mathscr{L}^{N-2}(x) = \pi \int_{\mathrm{B}_1^{N-2}} (1-|x|^2)\,\mathrm{d}\mathscr{L}^{N-2}(x)\\ &= \pi(N-2)\omega_{N-2}\int_0^1 (1-\varrho^2)\varrho^{N-3}\,\mathrm{d}\varrho = \frac{2\pi}{N}\omega_{N-2}\,.\end{aligned}$$

(3) Auflösung der Rekursion führt auf die gesuchten Formeln

$$\boxed{\omega_N = \frac{1}{k!}\pi^k \quad \text{für } N = 2k\,,\ k \in \mathbb{N}} \qquad \text{und} \qquad \boxed{\omega_N = \frac{2^N k!}{N!}\pi^k \quad \text{für } N = 2k{+}1\,,\ k \in \mathbb{N}_0}$$

Diese lassen sich zusammenfassend schreiben als

$$\omega_N = \frac{\pi^{N/2}}{\Gamma\big(\frac{N}{2}{+}1\big)} \qquad \text{für alle } N \in \mathbb{N}$$

mit der sogenannten Gamma-Funktion Γ zu

$$\Gamma(z) := \int_0^\infty t^{z-1}\mathrm{e}^{-t}\,\mathrm{d}t \qquad \text{für } z \in \mathbb{C} \text{ mit } \mathcal{Re}(z) > 0\,.$$

Diese Funktion erfüllt $\Gamma(k{+}1) = k!$ und kann insofern als eine Verallgemeinerung der Fakultäten natürlicher Zahlen verstanden werden. Des Weiteren sei über Γ nur erwähnt, dass die Funktionalgleichung $\Gamma(z{+}1) = z\Gamma(z)$ für alle $z \in \mathbb{C}$ mit $\mathcal{Re}(z) > 0$ gilt und $\Gamma(\frac{1}{2}) \overset{s=\sqrt{t}}{=} 2\int_0^\infty \mathrm{e}^{-s^2}\,\mathrm{d}s = \sqrt{\pi}$ ist. Diese Eigenschaften genügen, um die angegebene Formel einzusehen. Auf eine weitergehende Diskussion der Gamma-Funktion wird hier verzichtet.

2.9 L^p-Räume und Integralungleichungen

Definition (wesentliche Suprema und Infima). *Sei $(\Omega, \mathscr{A}, \mu)$ ein Maßraum, und $f\colon \Omega \to \overline{\mathbb{R}}$ sei eine beliebige Funktion. Das **(μ-)wesentliche Supremum** $\mu\text{-}\operatorname{ess\,sup}_\Omega f$ und das **(μ-)wesentliche Infimum** $\mu\text{-}\operatorname{ess\,inf}_\Omega f$ von f auf Ω sind definiert als*

$$\begin{aligned}\mu\text{-}\operatorname*{ess\,sup}_\Omega f &:= \min\{L \in \overline{\mathbb{R}} : f \leq L \textit{ gilt } \mu\textit{-fast-überall auf } \Omega\} \in \overline{\mathbb{R}}\,,\\ \mu\text{-}\operatorname*{ess\,inf}_\Omega f &:= \max\{L \in \overline{\mathbb{R}} : f \geq L \textit{ gilt } \mu\textit{-fast-überall auf } \Omega\} \in \overline{\mathbb{R}}\,.\end{aligned}$$

Besteht Klarheit über das Maß μ, so schreibt man kürzer $\operatorname{ess\,sup}_\Omega f$ *beziehungsweise* $\operatorname{ess\,inf}_\Omega f$.

Bemerkung (zur „wesentlichen Variante" der fundamentalen Abschätzung)**.** Für μ-integrierbares $f\colon \Omega \to \overline{\mathbb{R}}$ gilt stets

$$\mu(\Omega)\operatorname*{ess\,inf}_\Omega f \leq \int_\Omega f\,\mathrm{d}\mu \leq \mu(\Omega)\operatorname*{ess\,sup}_\Omega f\,.$$

Definitionen ($\mathbf{L}^p$-Normen, $\mathcal{L}^p$-Räume). *Sei $(\Omega, \mathscr{A}, \mu)$ ein Maßraum, und sei $p \in [1, \infty]$.*

- *Die* ***$\mathbf{L}^p$-Norm*** $\|f\|_{\mathrm{L}^p(\Omega,\mu;\mathbb{K}^M)} \in [0,\infty]$ *einer μ-messbaren Funktion $f\colon \Omega \to \mathbb{K}^M$ ist definiert als*
$$\|f\|_{\mathrm{L}^p(\Omega,\mu;\mathbb{K}^M)} := \begin{cases} \left(\displaystyle\int_\Omega |f|^p \,\mathrm{d}\mu\right)^{\frac{1}{p}} & \text{falls } p < \infty \\ \mu\text{-}\operatorname*{ess\,sup}_\Omega |f| & \text{falls } p = \infty \end{cases}$$
- *Der (Lebesgue-)* ***$\mathcal{L}^p$-Raum*** $\mathcal{L}^p(\Omega, \mathscr{A}, \mu\,; \mathbb{K}^M)$ *besteht aus allen $\mathscr{A}$-messbaren Funktionen $f\colon \Omega \to \mathbb{K}^M$ mit $\|f\|_{\mathrm{L}^p(\Omega,\mu;\mathbb{K}^M)} < \infty$.*

Notation. *Auf das Ausschreiben der Argumente Ω, $\mathscr{A}$, μ, $\mathbb{K}^M$ kann (teilweise) verzichtet werden, wenn der Bezug aus dem Kontext klar ist. Als Alternative zu $\|f\|_{\mathrm{L}^p(\Omega)}$ notiert man auch $\|f\|_{p;\Omega}$ oder nur $\|f\|_p$.*

Bemerkung. Für vollständiges μ enthält $\mathcal{L}^1(\Omega, \mathscr{A}, \mu\,; \mathbb{K}^M)$ genau die (μ-)summierbaren Funktionen, und die Funktionen in $\mathcal{L}^\infty(\Omega, \mathscr{A}, \mu\,; \mathbb{K}^M)$ heißen (μ-)wesentlich beschränkte Funktionen.

Satz (über die **Hölder-Ungleichung**). *Sei $(\Omega, \mathscr{A}, \mu)$ ein Maßraum, und seien $p, q \in [1, \infty]$* ***konjugierte Exponenten****, d.h. solche mit*
$$\frac{1}{p} + \frac{1}{q} = 1\,.$$
Für $f \in \mathcal{L}^p(\Omega, \mathscr{A}, \mu\,; \mathbb{K}^M)$ und $g \in \mathcal{L}^q(\Omega, \mathscr{A}, \mu\,; \mathbb{K}^M)$ ist die als punktweises Euklidisches Skalarprodukt von f und g gebildete Funktion $f \cdot g$ stets μ-summierbar mit
$$\left|\int_\Omega (f \cdot g)\,\mathrm{d}\mu\right| \le \|f\|_p\,\|g\|_q$$
(wobei die Normen bezüglich $(\Omega, \mu\,; \mathbb{K}^M)$ zu bilden sind).

Beweis. Für $p = 1$, $q = \infty$ erhält man mit der Cauchy-Schwarz-Ungleichung des Euklidischen Skalarprodukts
$$\left|\int_\Omega (f \cdot g)\,\mathrm{d}\mu\right| \le \int_\Omega |f \cdot g|\,\mathrm{d}\mu \le \int_\Omega |f|\,|g|\,\mathrm{d}\mu \le \int_\Omega |f|\,\|g\|_\infty\,\mathrm{d}\mu = \|f\|_1\,\|g\|_\infty\,,$$
und der Fall $p = \infty$, $q = 1$ ergibt sich daraus durch Vertauschung von f und g.

Sei nun $p \in (1, \infty)$. Verschwindet die rechte Seite der behaupteten Ungleichung, so ist eine beteiligten Funktionen μ-fast-überall auf Ω gleich Null, und die Behauptung gilt trivial. Andernfalls kann man nach Übergang zu $\frac{f}{\|f\|_p}$ und $\frac{g}{\|g\|_q}$ annehmen, dass $\|f\|_p = \|g\|_q = 1$ gilt und erhält mit der Youngschen Ungleichung[32]
$$\left|\int_\Omega (f \cdot g)\,\mathrm{d}\mu\right| \le \int_\Omega |f|\,|g|\,\mathrm{d}\mu \le \int_\Omega \left(\tfrac{1}{p}|f|^p + \tfrac{1}{q}|g|^q\right)\mathrm{d}\mu = \tfrac{1}{p}\|f\|_p^p + \tfrac{1}{q}\|g\|_q^q = \tfrac{1}{p} + \tfrac{1}{q} = 1\,. \qquad \square$$

[32] Die Youngsche Ungleichung besagt $ab \le \frac{1}{p}a^p + \frac{1}{q}b^q$ für $a, b \in \mathbb{R}_{\ge 0}$ und konjugierte Exponenten $p, q \in [1, \infty]$. Sie kann beispielsweise als Stützfunktionungleichung $h_p(a) \ge h_p(a_0) + h_p'(a_0)(a - a_0)$ der konvexen Funktion h_p mit $h_p(a) := a^p$ an der Stelle $a_0 = b^{\frac{1}{p-1}}$ hergeleitet werden.

Bemerkung. Ist $\mu(\Omega) < \infty$, beispielsweise μ die Einschränkung des Lebesgue-Maßes $\mathscr{L}^N$ auf beschränktes $\Omega \subset \mathbb{R}^N$, so gilt die Inklusion

$$\mathcal{L}^q(\Omega, \mathscr{A}, \mu\,; \mathbb{K}^M) \subset \mathcal{L}^p(\Omega, \mathscr{A}, \mu\,; \mathbb{K}^M) \qquad \text{für } p < q \text{ in } [1, \infty]\,.$$

Dies entnimmt man aus folgender Abschätzung mit der Hölder-Ungleichung zu den konjugierten Exponenten $\frac{q}{p} \in (1, \infty]$ und $s := (1-\frac{p}{q})^{-1} \in [1, \infty)$:

$$\|f\|_p = \left(\int_\Omega \mathbb{1}_\Omega \cdot |f|^p \,\mathrm{d}\mu\right)^{\frac{1}{p}} \leq \left(\|\mathbb{1}_\Omega\|_s \big\||f|^p\big\|_{\frac{q}{p}}\right)^{\frac{1}{p}} = \mu(\Omega)^{\frac{1}{p}-\frac{1}{q}}\|f\|_q\,.$$

Satz (über die **Minkowski-Ungleichung**). *Sei $(\Omega, \mathscr{A}, \mu)$ ein Maßraum, und sei $p \in [1, \infty]$. Für $f, g \in \mathcal{L}^p(\Omega, \mathscr{A}, \mu\,; \mathbb{K}^M)$ gilt dann stets $f+g \in \mathcal{L}^p(\Omega, \mathscr{A}, \mu\,; \mathbb{K}^M)$ mit*

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

(wobei die Normen bezüglich $(\Omega, \mu\,; \mathbb{K}^M)$ zu bilden sind).

Beweis. Die Fälle $p = 1$ und $p = \infty$ sind einfach. Für $p \in (1, \infty)$ schließt man zuerst aus $|f+g|^p \leq (2\max\{|f|, |g|\})^p \leq 2^p(|f|^p + |g|^p)$ auf $\|f+g\|_p \leq 2(\|f\|_p + \|g\|_p) < \infty$ und damit auf $f+g \in \mathcal{L}^p$. Um die Ungleichung in der behaupteten Form ohne den Vorfaktor 2 nachzuweisen, schätzt man mit der Hölderschen Ungleichung zu den konjugierten Exponenten $p \in (1, \infty)$ und $q := \frac{p}{p-1} \in (1, \infty)$ wie folgt ab:

$$\begin{aligned}
\|f+g\|_p^p &= \int_\Omega |f+g|\,|f+g|^{p-1} \,\mathrm{d}\mu \\
&\leq \int_\Omega |f|\,|f+g|^{p-1} \,\mathrm{d}\mu + \int_\Omega |g|\,|f+g|^{p-1} \,\mathrm{d}\mu \\
&\leq \|f\|_p \big\||f+g|^{p-1}\big\|_q + \|g\|_p \big\||f+g|^{p-1}\big\|_q = (\|f\|_p + \|g\|_p)\|f+g\|_p^{p-1}\,.
\end{aligned}$$

Da $\|f+g\|_p < \infty$ gezeigt wurde und sich der Fall $\|f+g\|_p = 0$ trivial erledigt, kann man auf beiden Seiten durch $\|f+g\|_p^{p-1}$ dividieren und erhält die behauptete Ungleichung. □

Zusammen mit der einfachen Regel $\|sf\|_p = |s|\,\|f\|_p$ für $s \in \mathbb{K}$ folgt aus dem vorigen Satz, dass die Räume $\mathcal{L}^p$ lineare Räume und die L^p-Normen absolut homogene Abbildungen mit Dreiecksungleichung sind. Genauer sei festhalten:

Für jedes $p \in [1, \infty]$ ist $\mathcal{L}^p(\Omega, \mathscr{A}, \mu\,; \mathbb{K}^M)$ ein $\mathbb{K}$-Vektorraum mit Seminorm $\|\cdot\|_p$.

Tatsächliche Normen sind die L^p-Normen in den meisten Fällen aber (noch) nicht, denn eine Norm darf nur auf dem Nullelement des Vektorraums den Wert Null annehmen, die L^p-Norm hat aber nicht nur auf der Nullfunktion den Wert Null, sondern auch auf allen Funktionen, die μ-*fast*-überall mit dieser übereinstimmen. Normalerweise (genauer, immer dann, wenn $\mathscr{A}$ eine Nullmenge $\neq\emptyset$ enthält) verschwindet die L^p-Norm damit auf dem Unterraum positiver Dimension

$$\mathcal{N}_\mu := \{f\colon \Omega \to \mathbb{K}^M \,:\, f \text{ ist } \mathscr{A}\text{-messbar und verschwindet } \mu\text{-fast-überall auf } \Omega\}\,.$$

Um die L^p-Normen dennoch als echte Normen auffassen zu können, behilft man sich wie folgt:

Definition ($\mathbf{L}^p$-Räume und Fastfunktionen). *Sei $(\Omega, \mathscr{A}, \mu)$ ein Maßraum, und sei $p \in [1, \infty]$. Der (Lebesgue-)* ***$\mathbf{L}^p$-Raum*** *$\mathrm{L}^p(\Omega, \mu\,; \mathbb{K}^M)$ ist der Faktorraum*

$$\mathrm{L}^p(\Omega, \mu\,; \mathbb{K}^M) := \mathcal{L}^p(\Omega, \mathscr{A}, \mu\,; \mathbb{K}^M)/\mathcal{N}_\mu = \{f{+}\mathcal{N}_\mu \,:\, f \in \mathcal{L}^p(\Omega, \mathscr{A}, \mu\,; \mathbb{K}^M)\}$$

nach dem im Vortext eingeführten (übrigens nicht von p abhängigen) Unterraum $\mathcal{N}_\mu$. Die Restklassen $f{+}\mathcal{N}_\mu \in \mathrm{L}^p(\Omega, \mu\,; \mathbb{K}^M)$ nennt man ***Lebesgue-Klassen*** *oder* ***(μ-)Fastfunktionen****.*

Bemerkungen (zu Fastfunktionen und L^p-Räumen)**.**

(1) Die Faktorbildung nach dem Unterraum $\mathcal{N}_\mu$ kann auch als Faktorbildung bezüglich der Äquivalenzrelation “Gleichheit μ-fast-überall” verstanden werden, also als **Identifikation μ-fast-überall gleicher Funktionen**. Da μ-messbare Funktionen dann stets mit dem $\mathscr{A}$-messbaren Repräsentanten aus Abschnitt 2.5 identifiziert werden, spielt die zugrundegelegte σ-Algebra $\mathscr{A}$ für L^p-Räume praktisch keine Rolle mehr und wird nicht mehr notiert.

(2) Per Definition ist $\mathrm{L}^p(\Omega, \mu\,; \mathbb{K}^M)$ ein $\mathbb{K}$-Vektorraum, und die Summe zweier Fastfunktionen sowie die Multiplikation einer Fastfunktion mit einem Skalar sind erklärt. Auch **etliche weitere Operationen** mit $\mathcal{L}^p$-Funktionen **können auf Fastfunktionen übertragen werden**, da sie nicht von der Wahl von Repräsentanten abhängen. Beispielsweise ist dies bei Rechenoperationen, Limites von Funktionenfolgen, der Bildung abzählbarer Suprema/Infima und auch bei Integralen $\int_\Omega (b \circ f)\,\mathrm{d}\mu$ mit einer festen Borel-Funktion b und $\mathcal{L}^p$-Funktionen f der Fall. Im Folgenden werden alle Notationen für derartige Operationen auch für Fastfunktionen verwendet, und in vielen unkritischen Fällen wird **nicht mehr explizit zwischen Fastfunktionen und Funktionen unterschieden**.

Zu beachten ist aber nichtsdestotrotz, dass **Auswertungen $f(x)$ in einzelnen Punkten** $x \in \Omega$ (zumindest in solchen mit $\mu(\{x\}) = 0$) oder Einschränkungen auf μ-Nullmengen **für Fastfunktionen f nicht sinnvoll** sind.

(3) Für zwei Repräsentanten derselben Fastfunktion, also für $f, g \in \mathcal{L}^p(\Omega, \mathscr{A}, \mu\,; \mathbb{K}^M)$ mit $f{+}\mathcal{N}_\mu = g{+}\mathcal{N}_\mu$, gilt insbesondere $\|f\|_p = \|g\|_p$. Daher ist die Festlegung $\|f{+}\mathcal{N}_\mu\|_p := \|f\|_p$ sinnvoll und **macht $\mathbf{L}^p(\Omega, \mu\,; \mathbb{K}^M)$ zu einem normierten Raum**.

(4) Gemäß der Hölderschen Ungleichung ist $f \cdot g$ für $f, g \in \mathcal{L}^2(\Omega, \mathscr{A}, \mu\,; \mathbb{K}^M)$ stets μ-summierbar, und durch

$$\langle f, g\rangle_{\mathrm{L}^2(\Omega,,\mu;\mathbb{K}^M)} := \int_\Omega f \cdot g \,\mathrm{d}\mu \in \mathbb{K}$$

erhält man eine positiv semidefinite, symmetrische Bilinearform (Fall $\mathbb{K} = \mathbb{R}$) beziehungsweise eine positiv semidefinite, hermitesche Sesquilinearform[33] auf $\mathcal{L}^2(\Omega, \mathscr{A}, \mu\,; \mathbb{K}^M)$. Man schreibt kurz $\langle f, g\rangle_{\mathrm{L}^2}$ statt $\langle f, g\rangle_{\mathrm{L}^2(\Omega,\mu;\mathbb{K}^M)}$.

Natürlich ist $\langle\,\cdot\,,\,\cdot\,\rangle_{\mathrm{L}^2}$ auch auf Fastfunktionen sinnvoll, ist auf $\mathrm{L}^2(\Omega, \mu\,; \mathbb{K}^M)$ sogar positiv definit und dort ein Skalarprodukt, das sogenannte **$\mathbf{L}^2$-Skalarprodukt**.

Es gilt $\|f\|_2 = \sqrt{\langle f, f\rangle_{\mathrm{L}^2}}$, also induziert das L^2-Skalarprodukt die L^2-Norm, und die Hölder-Ungleichung zu $p = q = 2$ ist gerade die Cauchy-Schwarz-Ungleichung des L^2-Skalarprodukts.

[33] Es sei daran erinnert, dass das komplexe Standard-Skalarprodukt auf $\mathbb{C}^N$ die Form $f \cdot g = \sum_{i=1}^N f_i \overline{g_i}$ mit der komplexen Konjugation $\overline{(\cdot)}$ hat.

Bemerkungen (zu Varianten von $\mathcal{L}^p$- und L^p-Räumen).

(1) Analog zu den vorigen Definitionen werden $\mathcal{L}^p$- und L^p-Räume zum Zielraum $\overline{\mathbb{R}}$ definiert. Für $f \in \mathcal{L}^p(\Omega, \mathscr{A}, \mu; \overline{\mathbb{R}})$ gilt dann μ-fast-überall $f \in \mathbb{R}$. Folglich hat jedes $f \in \mathrm{L}^p(\Omega, \mu; \overline{\mathbb{R}})$ einen $\mathbb{R}$-wertigen Repräsentanten, es ist $\mathrm{L}^p(\Omega, \mu; \overline{\mathbb{R}}) = \mathrm{L}^p(\Omega, \mu; \mathbb{R})$ und deshalb bringt $\overline{\mathbb{R}}$-Wertigkeit keinen echten Gewinn an Allgemeinheit.

(2) Ist Ω ein topologischer Raum mit $\mathcal{B}(\Omega) \subset \mathscr{A}$, so definiert man den lokalisierten Raum $\mathcal{L}^p_{\mathrm{lok}}(\Omega, \mathscr{A}, \mu\,; \mathbb{K}^M)$ als die Menge der $\mathscr{A}$-messbaren $f \colon \Omega \to \mathbb{K}^M$, so dass jedes $x \in \Omega$ in einem offenen O mit $f|_O \in \mathcal{L}^p(O, \mathscr{A}|O, \mu|_{\mathscr{A}|O}\,; \mathbb{K}^M)$ liegt. Bei gutartigem Ω (z.B. für einen lokalkompakten Hausdorff-Raum Ω) ist es äquivalent, $f|_K \in \mathcal{L}^p(K, \mathscr{A}|K, \mu|_{\mathscr{A}|K}\,; \mathbb{K}^M)$ für jedes Kompaktum K in Ω zu fordern. Den entsprechenden Raum von Fastfunktionen nennt man $\mathrm{L}^p_{\mathrm{lok}}(\Omega, \mu\,; \mathbb{K}^M)$.

Eine wichtigen Eigenschaften der L^p-Räume besteht darin, dass sie (adäquat) viele konvergente Folgen besitzen. Dies ist ein **entscheidender Vorteil gegenüber** (entsprechenden Räumen in) **der Riemannschen Integrationstheorie** und wird nun präzisiert und bewiesen:

Satz (**von Riesz-Fischer**, ~1907)**.** *Sei* $(\Omega, \mathscr{A}, \mu)$ *ein Maßraum. Für alle* $p \in [1, \infty]$ *ist der normierte Raum* $\mathrm{L}^p(\Omega, \mu\,; \mathbb{K}^M)$ *vollständig, d. h. in* $\mathrm{L}^p(\Omega, \mu\,; \mathbb{K}^M)$ *konvergiert jede Cauchy-Folge.*

Beweis. Im Fall $p = \infty$ folgt der Satz aus der bekannten Tatsache, dass gleichmäßige Cauchy-Folgen gleichmäßig konvergieren.

Sei jetzt $p \in [1, \infty)$, und sei $(f_k)_{k \in \mathbb{N}}$ eine Cauchy-Folge in L^p. Dann gibt es eine Teilfolge f_{k_i} mit $\|f_{k_{i+1}} - f_{k_i}\|_p \le 2^{-i}$. Mit dem Lemma von Fatou und der Minkowski-Ungleichung folgt

$$\left\| \sum_{i=1}^{\infty} |f_{k_{i+1}} - f_{k_i}| \right\|_p \le \sum_{i=1}^{\infty} \|f_{k_{i+1}} - f_{k_i}\|_p \le 1\,,$$

und deshalb liegt μ-fast-überall auf Ω Konvergenz $\sum_{i=1}^{\infty} |f_{k_{i+1}} - f_{k_i}| < \infty$ vor. Gemäß der Vollständigkeit von $\mathbb{K}^M$ existiert μ-fast-überall auf Ω auch

$$f := \lim_{m\to\infty} f_{k_m} = f_{k_1} + \lim_{m\to\infty} \sum_{i=1}^{m-1} (f_{k_{i+1}} - f_{k_i}) = f_{k_1} + \sum_{i=1}^{\infty} (f_{k_{i+1}} - f_{k_i})\,.$$

Für beliebiges $\varepsilon \in \mathbb{R}_{>0}$ erhält man aus dem Lemma von Fatou und der Cauchy-Eigenschaft der f_k nun

$$\int_\Omega |f_k - f|^p \,\mathrm{d}\mu \le \liminf_{i\to\infty} \int_\Omega |f_k - f_{k_i}|^p \,\mathrm{d}\mu \le \varepsilon \qquad \text{für } k \ge k_0(\varepsilon)\,,$$

also liegt Konvergenz $\lim_{k\to\infty} f_k = f$ in L^p vor. □

Bemerkungen (zur Vollständigkeit von L^p)**.**

(1) **$\mathbf{L}^p$ ist also ein Banachraum**, d. h. ein vollständiger normierter Raum, und **$\mathbf{L}^2$ ist sogar ein Hilbertraum**, d. h. ein vollständiger Skalarproduktraum.

(2) Die L^p-Räume sind **fundamentale Funktionenräume der Analysis** und wichtig für alle mathematischen Teilgebiete, für die auch Integration eine Rolle spielt.

Zum Abschluss des Kapitels wird eine weitere wichtige und nützliche Integralungleichung, die Jensensche Ungleichung (für Erwartungswerte), behandelt:

Satz (über die **Jensensche Ungleichung**). *Seien* $(\Omega,\mathscr{A},\mu)$ *ein* ***Wahrscheinlichkeits****raum,* $\Phi\colon \mathbb{K}^M \to \mathbb{R}\cup\{\infty\}$ *eine* ***konvexe***[34] *Borel-Funktion und* $f\colon \Omega\to\mathbb{K}^M$ *eine* μ*-summierbare Funktion. Dann ist* $\int_\Omega (\Phi\circ f)_-\,\mathrm{d}\mu<\infty$*, und es gilt*

$$\Phi\Big(\int_\Omega f\,\mathrm{d}\mu\Big)\le\int_\Omega(\Phi\circ f)\,\mathrm{d}\mu\qquad \textit{in } \mathbb{R}\cup\{\infty\}\,.$$

Bemerkungen (zur Jensenschen Ungleichung)**.**

(1) Für einen beliebigen Maßraum $(\Omega,\mathscr{A},\mu)$ mit $0<\mu(\Omega)<\infty$ ist $\frac{1}{\mu(\Omega)}\mu$ ein Wahrscheinlichkeitsmaß, daher gilt die Jensensche Ungleichung allgemeiner in der Form

$$\Phi\Big(\frac{1}{\mu(\Omega)}\int_\Omega f\,\mathrm{d}\mu\Big)\le\frac{1}{\mu(\Omega)}\int_\Omega(\Phi\circ f)\,\mathrm{d}\mu\,.$$

(2) Es bringt Vorteile, für die konvexe Funktion Φ in der Jensenschen Ungleichung nicht nur reelle Werte, sondern auch den Wert ∞ zuzulassen:

- Ist K eine konvexe Teilmenge von $\mathbb{K}^M$, so kann man die Jensen-Ungleichung mit der konvexen Funktion $\infty\cdot\mathbb{1}_{\mathbb{K}^M\setminus K}$ anwenden und dadurch einsehen, dass **bei** μ-summierbarem f mit (μ-fast-überall) **Werten in** $\boldsymbol{K}$ **auch der Integralmittelwert** $\frac{1}{\mu(\Omega)}\int_\Omega f\,\mathrm{d}\mu$ **in der konvexen Menge** $\boldsymbol{K}$ liegt. Diese Aussage spielt auch bei manchen Beweisen der Jensenschen Ungleichung ein Rolle; dazu vergleiche unten.
- Ist eine konvexe Funktion Φ nicht auf ganz $\mathbb{K}^M$, sondern nur auf einer konvexen Teilmenge $K\subset\mathbb{K}^M$ definiert, so lässt sich die Jensensche Ungleichung auf die konvexe Fortsetzung $\Phi\cdot\mathbb{1}_K+\infty\cdot\mathbb{1}_{\mathbb{K}^M\setminus K}$ anwenden und behält für Funktionen f mit Werten in K (und folglich Mittelwert in K) genau die oben angegebene Form.

(3) Ist Φ reell-wertig, so ist Φ als konvexe Funktion automatisch stetig und damit eine Borel-Funktion. Wird wie im Satz auch der Wert ∞ zugelassen, so muss aber in Randpunkten von $\Phi^{-1}(\infty)$ nicht notwendig Stetigkeit vorliegen und auf die explizite Voraussetzung der Borel-Messbarkeit kann nicht ohne Weiteres verzichtet werden.

(4) Alle Dreiecksungleichungen für Integrale sind Spezialfälle der Jensenschen Ungleichung mit Beträgen oder Normen als konvexe Funktionen. Auch die mit der Inklusion $\mathcal{L}^q\subset\mathcal{L}^p$ verbundene Ungleichung (siehe die Bemerkung nach der Hölder-Ungleichung) stellt sich als Spezialfall zur konvexen Funktion $y\mapsto|y|^{q/p}$ heraus.

Teilbeweis der Jensenschen Ungleichung. Da $\mathbb{C}^M$ mit $\mathbb{R}^{2M}$ identifiziert werden kann, wird ohne Einschränkung $\mathbb{K}=\mathbb{R}$ angenommen. Ein Basis-Argument zum Beweis der Jensenschen Ungleichung funktioniert immer dann, wenn Φ an der Stelle $a:=\int_\Omega f\,\mathrm{d}\mu\in\mathbb{R}^M$ einen endlichen Wert $\Phi(a)<\infty$ annimmt und dort eine affin lineare Stützfunktion in dem Sinn besitzt, dass es ein $w\in\mathbb{R}^M$ mit $\Phi(y)\ge\Phi(a)+w\cdot(y-a)$ für alle $y\in\mathbb{R}^M$ gibt. Man kann dann nämlich mit der Bedingung $\mu(\Omega)=1$ und der Linearität des Integrals die Rechnung

$$\begin{aligned}\Phi\Big(\int_\Omega f\,\mathrm{d}\mu\Big)&=\Phi(a)+w\cdot\Big(\int_\Omega f\,\mathrm{d}\mu-a\Big)\\&=\int_\Omega\Phi(a)\,\mathrm{d}\mu+w\cdot\Big(\int_\Omega f\,\mathrm{d}\mu-\int_\Omega a\,\mathrm{d}\mu\Big)=\int_\Omega\big[\Phi(a)+w\cdot(f-a)\big]\,\mathrm{d}\mu\le\int_\Omega\Phi(f)\,\mathrm{d}\mu\end{aligned}$$

[34]Konvexität von Φ wird auch bei Werten in $\mathbb{R}\cup\{\infty\}$ „ganz normal“ über die Konvexitätsungleichung definiert, bedeutet also $\Phi(\lambda x+(1-\lambda)y)\le\lambda\Phi(x)+(1-\lambda)\Phi(y)$ für alle $x,y\in\mathbb{K}^M$ und $\lambda\in[0,1]$.

durchführen und erhält die Jensen-Ungleichung samt $\int_\Omega (\Phi(f))_- \,\mathrm{d}\mu < \infty$. Die angenommene Existenz einer Stützfunktion ist dabei für konvexe $\Phi \in \mathrm{C}^1(\mathbb{R}^M, \mathbb{R})$ aus der Analysis bekannt (dann einfach $w = \nabla\Phi(a)$ wählen), und mit Hilfe des nächsten Lemmas kann sie allgemeiner für konvexe $\Phi\colon \mathbb{R}^M \to \mathbb{R} \cup \{\infty\}$ an Stellen $a \in \mathbb{R}^M$ mit $\Phi < \infty$ auf einer Umgebung von a gezeigt werden. Mit einer leichten Modifikation des obigen Arguments erhält man die Jensensche Ungleichung zudem auch bei Werten in $\mathbb{R} \cup \{\infty\}$ zumindest für *unterhalbstetige* konvexe $\Phi\colon \mathbb{R}^M \to \mathbb{R}\cup\{\infty\}$, indem man solche als Supremum affiner Funktionen darstellt. Alles in allem ergibt sich hiermit die Behauptung des Satzes unter der schwachen Zusatzvoraussetzung, dass Φ entweder $\mathbb{R}$-wertig oder unterhalbstetig ist. □

Der Beweis der Jensenschen Ungleichung bei Werten in $\mathbb{R}\cup\{\infty\}$ und ohne Unterhalbstetigkeitsvoraussetzung gelingt mit etwas (mehr) konvexer Analysis, ist aber ein Detail, das über den Vorlesungsstoff hinaus geht und hier nur der Vollständigkeit halber ausgeführt wird. Man beginnt mit einem geometrisch plausiblen Lemma:

Lemma (über die Existenz einer Normalenrichtung). *Sei K eine konvexe Menge in $\mathbb{R}^M$. Dann gibt es zu jedem $p \in \mathbb{R}^M \setminus \mathring{K}$ einen Einheitsvektor $v \in \mathbb{R}^M$ mit $(x{-}p)\cdot v \geq 0$ für alle $x \in K$ (und dasselbe sogar mit >0 falls $p \notin \overline{K}$).*

Beweis. Sei $K \neq \emptyset$. Man unterscheidet die Fälle $p \notin \overline{K}$ und $p \in \partial K$.

Ist $p \notin \overline{K}$, so sei p_* der nächste Punkt zu p in $\overline{K}$. Dann gilt

$$0 \leq \frac{1}{\lambda}\Big[|\lambda x{+}(1{-}\lambda)p_*{-}p|^2 - |p_*{-}p|^2\Big] = 2(x{-}p_*)\cdot(p_*{-}p) + \lambda|x{-}p_*|^2$$

für alle $x \in K$ und $\lambda \in (0,1]$. Durch Grenzübergang $\lambda \searrow 0$ erhält man erst $(x{-}p_*)\cdot(p_*{-}p) \geq 0$ und dann mit $(x{-}p)\cdot(p_*{-}p) = (x{-}p_*)\cdot(p_*{-}p) + |p_*{-}p|^2 > 0$ die Behauptung für $v := \frac{p_*-p}{|p_*-p|}$.

Ein Randpunkt $p \in \partial K$ kann wegen der Konvexität von K kein innerer Punkt von $\overline{K}$ sein und kann deshalb als $p = \lim_{k\to\infty} p_k$ mit $p_k \in \mathbb{R}^M \setminus \overline{K}$ geschrieben werden. Aus den gemäß dem vorigen Teil des Arguments zugehörigen Einheitsvektoren $v_k \in \mathbb{R}^M$ lässt sich eine gegen einen Einheitsvektor $v \in \mathbb{R}^M$ konvergente Teilfolge auswählen, und die Behauptung folgt problemlos. □

Lemma (zur Mittelwertbildung über konvexe Mengen). *Sei $(\Omega, \mathscr{A}, \mu)$ ein Wahrscheinlichkeitsraum, K sei eine konvexe Menge in $\mathbb{R}^M$, und $f\colon \Omega \to \mathbb{R}^M$ sei μ-summierbar. Gilt μ-fast-überall $f \in K$ auf Ω, so folgt $\int_\Omega f\,\mathrm{d}\mu \in K$.*

Bemerkung. Insbesondere besagt das Lemma für $\Omega = K \subset V = \mathbb{R}^N$ mit $0 < \mathscr{L}^N(K) < \infty$, dass der **Schwerpunkt** $\frac{1}{\mathscr{L}^N(K)} \int_K x \,\mathrm{d}\mathscr{L}^N(x)$ einer konvexen Menge K stets selbst in K liegt.

Beweis. Man beweist das Lemma durch vollständige Induktion nach $M \in \mathbb{N}$. Für $M = 1$ ist K ein Intervall, und die Behauptung ergibt sich aus in Abschnitt 2.6 behandelten Grundeigenschaften (2) und (3) des Maßintegrals. Somit ist der Induktionsanfang gemacht, es folgt der Induktionsschluss: Sei $M \geq 2$ und das Lemma für konvexe Mengen in $\mathbb{R}^{M-1}$ bewiesen. Ist $p := \int_\Omega f\,\mathrm{d}\mu$ in K, so ist nichts zu zeigen. Andernfalls ist $p \in \mathbb{R}^M \setminus K$, und mit dem Einheitsvektor w des vorigen Lemmas gilt μ-fast-überall $(f{-}p)\cdot w \geq 0$. Außerdem ist

$$\int_\Omega (f{-}p)\cdot w \,\mathrm{d}\mu = \bigg(\int_\Omega f\,\mathrm{d}\mu - p\bigg)\cdot w = 0$$

und deshalb μ-fast-überall schon $(f{-}p)\cdot w = 0$. Somit hat $f{-}p$, eventuell nach Abänderung auf einer μ-Nullmenge, Werte im $(M{-}1)$-dimensionalen orthogonalen Komplement $\{v\}^\perp := \{x \in \mathbb{R}^M : x \cdot v = 0\}$ von v. Da $v^\perp$ als $\mathbb{K}$-Vektorraum isomorph zu $\mathbb{R}^{M-1}$ ist und der zugehörige Isomorphismus wegen Linearität aus dem Integral gezogen werden darf, gilt die Behauptung nach Induktionsannahme für die Funktion $f{-}p$ und die konvexe Teilmenge $(K{-}p)\cap v^\perp$ von $v^\perp$. Es folgt $\int_\Omega (f{-}p)\,\mathrm{d}\mu \in (K{-}p)\cap v^\perp$, und aufgrund von $\mu(\Omega) = 1$ gilt damit auch $\int_\Omega f\,\mathrm{d}\mu \in K$. □

Beweis der Jensenschen Ungleichung. Ohne Einschränkung sei $\mathbb{K} = \mathbb{R}$ und μ vollständig.

Man verifiziert die Jensensche Ungleichung zunächst unter der Annahme, dass $\int_\Omega |\Phi\circ f|\,\mathrm{d}\mu < \infty$ gilt. In diesem Fall kann $\Phi \circ f$, eventuell nach Abänderung von f auf einer Nullmenge als $\mathbb{R}$-wertig angenommen werden. Für die konvexe Menge

$$K_\Phi := \{(x,y) \in \mathbb{R}^M \times \mathbb{R} \,:\, y \geq \Phi(x)\} \subset \mathbb{R}^{M+1}$$

gilt μ-fast-überall $(f, \Phi\circ f) \in K_\Phi$ auf Ω. Mit dem vorigen Lemma folgt deshalb

$$\bigg(\int_\Omega f\,\mathrm{d}\mu, \int_\Omega (\Phi\circ f)\,\mathrm{d}\mu\bigg) \in K_\Phi\,,$$

was offensichtlich eine Umformulierung der Jensenschen Ungleichung ist.

Im allgemeinen Fall ist zunächst $\int_\Omega (\Phi\circ f)_- \,\mathrm{d}\mu < \infty$ nachzuweisen: Dies gilt trivial, wenn $\Omega_- := \{x \in \Omega : \Phi(f(x)) \leq 0\}$ eine μ-Nullmenge ist. Hat Ω_- dagegen positives μ-Maß, so kann das bereits Bewiesene mit $\Phi_k := \max\{\Phi, -k\}$ angewandt werden, da $\Phi_k \circ f$ auf Ω_- beschränkt und somit μ-summierbar über Ω_- ist. Man erhält

$$\frac{1}{\mu(\Omega_-)}\int_{\Omega_-} (\Phi_k \circ f)\,\mathrm{d}\mu \geq \Phi_k\Big(\frac{1}{\mu(\Omega_-)}\int_{\Omega_-} f\,\mathrm{d}\mu\Big) \geq \Phi\Big(\frac{1}{\mu(\Omega_-)}\int_{\Omega_-} f\,\mathrm{d}\mu\Big) > -\infty\,.$$

Damit gilt $\sup_{k\in\mathbb{N}} \int_\Omega (\Phi_k \circ f)_- \,\mathrm{d}\mu < \infty$, und mit Hilfe des Lemmas von Fatou folgt $\int_\Omega (\Phi \circ f)_- \,\mathrm{d}\mu < \infty$. Es muss somit entweder $\int_\Omega |\Phi \circ f|\,\mathrm{d}\mu < \infty$ oder $\int_\Omega (\Phi \circ f)\,\mathrm{d}\mu = \infty$ gelten. Für den ersten dieser Fälle wurde die Jensensche Ungleichung oben bewiesen, im zweiten gilt sie trivial. □

2.10 Fourier-Reihen, Fourier-Transformation

In weitgehender Analogie zum bekannten 1-dimensionalen Fall erklärt man Fourier-Reihen in mehreren Dimensionen wie folgt:

Definition (Fourier-Koeffizienten, Fourier-Reihen). *Zu jedem $k = (k_1, k_2, \ldots, k_N) \in \mathbb{Z}^N$ wird ein trigonometrisches Monom $\mathrm{e}_k \colon \mathbb{R}^N \to \mathbb{C}$ durch*

$$\mathrm{e}_k(x) := \mathrm{e}^{\mathbf{i}k\cdot x} = \exp(\mathbf{i}(k_1x_1{+}k_2x_2{+}\ldots{+}k_Nx_N)) \qquad \text{für } x = (x_1, x_2, \ldots, x_N) \in \mathbb{R}^N$$

definiert. Für eine $\mathscr{L}^N$-summierbare Funktion $f \colon [0, 2\pi)^N \to \mathbb{C}$ (oder auch deren in jeder Variablen 2π-periodische Fortsetzung auf ganz $\mathbb{R}^N$) bezeichnet man dann die Zahlen

$$c_k := \frac{1}{(2\pi)^N}\int_{[0,2\pi)^N} f\,\overline{\mathrm{e}_k}\,\mathrm{d}\mathscr{L}^N \in \mathbb{C} \qquad \text{zu } k \in \mathbb{Z}^N$$

*als **Fourier-Koeffizienten** von f und nennt die mit diesen gebildete Reihe*

$$\sum_{k\in\mathbb{Z}^N} c_k\mathrm{e}_k$$

*die **(formale) Fourier-Reihe** von f.*

Wie im 1-dimensionalen Fall kann man die Fourier-Koeffizienten

$$c_k = f \cdot \mathrm{e}_k$$

als geeignet normierte L^2-Skalarprodukte von f und e_k, d.h. als Skalarprodukte von f und e_k in $\mathcal{L}^2\big([0, 2\pi)^N, \mathcal{M}^N, (2\pi)^{-N}\mathscr{L}^N; \mathbb{C}\big)$ beziehungsweise $\mathrm{L}^2\big([0, 2\pi)^N, (2\pi)^{-N}\mathscr{L}^N; \mathbb{C}\big)$ auffassen (jedenfalls wenn f überhaupt in diesem $\mathcal{L}^2$- beziehungsweise L^2-Raum liegt; auch sonst ist die Verwendung der Schreibweise $f \cdot \mathrm{e}_k$ aber oft noch sinnvoll).

Neben schon bekannte Sätze über das Konvergenzverhalten von Fourier-Reihen tritt nun folgender allgemeingültige (aber auch in einer Dimension noch neue) Sachverhalt:

Satz (klassischer Satz von Riesz-Fischer, ∼1907). *Die Abbildung*

$$\mathrm{L}^2\big([0, 2\pi)^N, (2\pi)^{-N}\mathscr{L}^N; \mathbb{C}\big) \to \ell^2\big(\mathbb{Z}^N; \mathbb{C}\big)\,,\ f \mapsto (f \cdot \mathrm{e}_k)_{k\in\mathbb{Z}^N}$$

*ist ein isometrischer Isomorphismus, insbesondere gelten die **Besselsche Gleichung***

$$\|f\|^2_{\mathrm{L}^2} = \sum_{k\in\mathbb{Z}^N} |f \cdot \mathrm{e}_k|^2 \qquad \text{für } f \in \mathrm{L}^2\big([0, 2\pi)^N, (2\pi)^{-N}\mathscr{L}^N; \mathbb{C}\big)$$

und die allgemeinere ***Parsevalsche Gleichung***

$$f \cdot g = \sum_{k \in \mathbb{Z}^N} (f \cdot \mathrm{e}_k)(\overline{g \cdot \mathrm{e}_k}) \qquad \text{für } f, g \in \mathrm{L}^2\big([0,2\pi)^N, (2\pi)^{-N}\mathscr{L}^N; \mathbb{C}\big)\,,$$

und jedes $f \in \mathrm{L}^2\big([0,2\pi)^N, (2\pi)^{-N}\mathscr{L}^N; \mathbb{C}\big)$ *hat die* ***Darstellung*** $f = \sum_{k \in \mathbb{Z}^N} (f \cdot \mathrm{e}_k)\mathrm{e}_k$ ***als in*** $\mathbf{L^2}\big([0,2\pi)^N, (2\pi)^{-N}\mathscr{L}^N; \mathbb{C}\big)$ ***konvergente Fourier-Reihe****.*

Ausgeschrieben bedeutet die durch den Satz garantierte L^2-Konvergenz von $\sum_{k \in \mathbb{Z}^N} (f \cdot \mathrm{e}_k)\mathrm{e}_k$ gegen f hierbei

$$\lim_{m \to \infty} \bigg\| f - \sum_{i=1}^m (f \cdot \mathrm{e}_{k_i})\mathrm{e}_{k_i} \bigg\|_{\mathrm{L}^2} = 0$$

für eine und dann automatisch[35] jede Abzählung $(k_i)_{i \in \mathbb{N}}$ von $\mathbb{Z}^N$. Diese Konvergenz wird auch als **Konvergenz** der Reihe **im quadratischen Mittel** bezeichnet.

Der Beweis des Satzes beruht maßgeblich auf der in Abschnitt 2.9 nachgewiesenen Vollständigkeit des Raums L^2 und auf folgenden Basis-Aspekten abstrakter Hilbertraum-Theorie:

Definitionen (Orthonormalsysteme, Orthonormalbasen). *Seien H ein Skalarproduktraum, I eine beliebige Indexmenge und $(e_k)_{k \in I}$ eine Familie von Elementen von H.*

(I) *Man nennt $(e_k)_{k \in I}$ ein* ***Orthonormalsystem****, wenn $e_k \cdot e_\ell = \delta_{k\ell}$ für alle $k, \ell \in I$ gilt (wobei $\delta_{k\ell} := 1$ im Fall $k = \ell$ und $\delta_{k\ell} := 0$ im Fall $k \neq \ell$ ist).*

(II) *Ein Orthonormalsystem $(e_k)_{k \in I}$ heißt* ***vollständig****, wenn für $f \in H$ aus $f \cdot e_k = 0$ für alle $k \in I$ schon $f = 0$ folgt.*

(III) *Ein Orthonormalsystem $(e_k)_{k \in I}$ heißt eine* ***Orthonormalbasis****, wenn I abzählbar ist und jedes $f \in H$ die Darstellung $\sum_{k \in I} (f \cdot e_k)e_k = f$ besitzt (wobei die Konvergenz der Reihe, ähnlich wie oben, $\lim_{m \to \infty} \big\| f - \sum_{i=1}^m (f \cdot e_{k_i})e_{k_i} \big\|_H = 0$ für eine und damit jede Abzählung $(k_i)_{i \in \mathbb{N}}$ von I bedeutet).*

Damit kann formuliert werden:

Satz. *Sei H ein Hilbertraum mit abzählbarem Orthonormalsystem $(e_k)_{k \in \mathbb{N}}$. Dann konvergiert für jedes $f \in H$ die Reihe $\sum_{k=1}^\infty (f \cdot e_k)e_k$ in der Norm von H. Ist das Orthonormalsystem $(e_k)_{k \in \mathbb{N}}$ zudem vollständig, so konvergiert die Reihe auch stets gegen f, es handelt sich bei $(e_k)_{k \in \mathbb{N}}$ also dann um eine Orthonormalbasis.*

Beweis. Der „Satz des Pythagoras", $\|f+g\|^2 = \|f\|_H^2 + \|g\|_H^2$ für $f, g \in H$ mit $f \cdot g = 0$, liefert $\|f\|_H^2 = \|f - \sum_{k=1}^n (f \cdot e_k)e_k\|_H^2 + \|\sum_{k=1}^n (f \cdot e_k)e_k\|_H^2 \geq \|\sum_{k=1}^n (f \cdot e_k)e_k\|_H^2 = \sum_{k=1}^n |f \cdot e_k|^2$ für alle $n \in \mathbb{N}$. Damit sind die Partialsummen der Reihe $\sum_{k=1}^\infty |f \cdot e_k|^2$ beschränkt, und Konvergenz dieser Reihe in $\mathbb{R}$ ist nachgewiesen. Aufgrund der analogen Formel $\|\sum_{k=m}^n (f \cdot e_k)e_k\|_H^2 = \sum_{k=m}^n |f \cdot e_k|^2$ für $m < n$ in $\mathbb{N}$ überträgt sich die Cauchy-Eigenschaft auf die Reihe $\sum_{k=1}^\infty (f \cdot e_k)e_k$ in H, wegen der Vollständigkeit von H konvergiert $\sum_{k=1}^\infty (f \cdot e_k)e_k$ also in H. Außerdem gilt

[35] Dieser Automatismus wird auch aus dem Beweis des Satzes klar, dennoch sei der Grund schon festgehalten: Für Abzählungen $(k_i)_{i \in \mathbb{N}}$, $(\ell_j)_{j \in \mathbb{N}}$ von $\mathbb{Z}^N$, $n \in \mathbb{N}$, $m_n := \max\{m \in \mathbb{N} : \{k_1, k_2, \ldots, k_m\} \subset \{\ell_1, \ell_2, \ldots, \ell_n\}\}$ und die Rest-Indexmenge $R_n := \{\ell_1, \ell_2, \ldots, \ell_n\} \setminus \{k_1, k_2, \ldots, k_{m_n}\}$ liefert Orthogonalität der e_k die Abschätzung $\big\| f - \sum_{i=1}^{m_n} (f \cdot \mathrm{e}_{k_i})\mathrm{e}_{k_i} \big\|_{\mathrm{L}^2}^2 = \big\| f - \sum_{j=1}^n (f \cdot \mathrm{e}_{\ell_j})\mathrm{e}_{\ell_j} \big\|_{\mathrm{L}^2}^2 + \big\| \sum_{k \in R_n} (f \cdot \mathrm{e}_k)\mathrm{e}_k \big\|_{\mathrm{L}^2}^2 \geq \big\| f - \sum_{j=1}^n (f \cdot \mathrm{e}_{\ell_j})\mathrm{e}_{\ell_j} \big\|_{\mathrm{L}^2}^2$. Wegen $\lim_{n \to \infty} m_n = \infty$ folgt aus der Konvergenz der Reihe „mit k_i" dann die Konvergenz der Reihe „mit ℓ_j".

$\big(f - \sum_{k=1}^m (f \cdot e_k)e_k\big) \cdot e_\ell = f \cdot e_\ell - \sum_{k=1}^m (f \cdot e_k)\delta_{k\ell} = 0$ für $m \ge \ell$ in $\mathbb{N}$, und mit Stetigkeit des Skalarprodukts (Cauchy-Schwarz-Ungleichung!) folgt $\big(f - \sum_{k=1}^\infty (f \cdot e_k)e_k\big) \cdot e_\ell = 0$ für alle $\ell \in \mathbb{N}$. Bei vollständigem Orthonormalsystem $(e_k)_{k\in\mathbb{N}}$ muss damit $f - \sum_{k=1}^\infty (f \cdot e_k)e_k = 0$ sein. □

Korollare.

(I) *In einem Hilbertraum ist ein **abzählbares vollständiges Orthonormalsystem nichts anderes als** eine **abzählbare Orthonormalbasis** (denn die Hin-Richtung gilt gemäß dem vorausgehenden Satz, die Rück-Richtung gilt trivial).*

(II) *Für einen Hilbertraum H sind folgende Aussagen äquivalent:*

(1) *H besitzt eine abzählbare Orthonormalbasis.*

(2) *Es gibt einen Unterraum U von H, so dass U in H dicht liegt und U eine abzählbare Basis (im Sinn endlicher Linearkombinationen wie in der linearen Algebra) besitzt.*

(3) *H ist separabel, d.h. eine abzählbare Teilmenge A von H liegt dicht in H.*

Beweis. (1) $\Longrightarrow$ (2): Der von der abzählbaren Orthonormalbasis (im Sinn der linearen Algebra) aufgespannte Unterraum liegt dicht in H.

(2) $\Longrightarrow$ (3): Sei $(b_k)_{k\in\mathbb{N}}$ eine abzählbare Basis (oder auch nur ein abzählbares Erzeugendensystem) von U. Mit U liegt dann auch der von $(b_k)_{k\in\mathbb{N}}$ aufgespannte $\mathbb{Q}$-Unterraum $A := \big\{ \sum_{k=1}^m q_k b_k \,:\, m \in \mathbb{N}\,,\, q_1, q_2, \ldots, q_m \in \mathbb{Q}\big\}$ dicht in H, und A ist abzählbar.

(3) $\Longrightarrow$ (1): Sei $A = \{a_k \,:\, k \in \mathbb{N}\}$ eine abzählbare dichte Teilmenge von H. Durch sukzessive Auswahl des nächsten geeigneten a_k findet man $k_1 < k_2 < k_3 < \ldots$ in $\mathbb{N}$, so dass die $a_{k_1}, a_{k_2}, \ldots, a_{k_{m-1}}, a_{k_m}$ für alle $m \in \mathbb{N}$ linear unabhängig sind und denselben Unterraum wie $a_1, a_2, \ldots, a_{k_m-1}, a_{k_m}$ erzeugen. Durch Gram-Schmidt-Orthonormalisierung erzeugt man aus den a_{k_i} ein Orthonormalsystem $(e_i)_{i\in\mathbb{N}}$, so dass auch $e_1, e_2, \ldots, e_{m-1}, e_m$ für alle $m \in \mathbb{N}$ denselben Unterraum wie $a_1, a_2, \ldots, a_{k_m-1}, a_{k_m}$ erzeugen. Insbesondere enthält der vom ganzen Orthonormalsystem $(e_i)_{i\in\mathbb{N}}$ erzeugte Unterraum U die Ausgangsmenge A und liegt somit dicht in H. Jedes $f \in H$ kann daher als $f = \lim_{\ell\to\infty} f_\ell$ mit $f_\ell \in U$ geschrieben werden. Ist $f \cdot e_i = 0$ für alle $i \in \mathbb{N}$, so folgt erst $f \cdot f_\ell = 0$ für alle $\ell \in \mathbb{N}$, dann $\|f\|_H^2 = \lim_{\ell\to\infty}(f \cdot f_\ell) = 0$ und schließlich $f = 0$. Also ist das Orthonormalsystem $(e_i)_{i\in\mathbb{N}}$ vollständig und bildet gemäß (I) eine Orthonormalbasis von H. □

(III) ***Isomorphie-Satz**: **Jeder separable Hilbertraum ist isometrisch isomorph zum Raum** $\ell^2(\mathbb{K})$ **der Quadrat-summierbaren Folgen** — und tatsächlich ist für jede Orthonormalbasis $(e_k)_{k\in\mathbb{N}}$ von H die Abbildung $H \to \ell^2(\mathbb{K})\,,\, f \mapsto (f \cdot e_k)_{k\in\mathbb{N}}$ ein solcher Isomorphismus.*

Beweis. Nach (II) existiert eine Orthonormalbasis $(e_k)_{k\in\mathbb{N}}$ von H. Für eine solche gilt $\sum_{k=1}^m (f \cdot e_k)\mathrm{e}_k \cdot \sum_{\ell=1}^m (g \cdot e_\ell)\mathrm{e}_\ell = \sum_{k,\ell=1}^m (f \cdot e_k)(\overline{g \cdot e_\ell})(e_k \cdot e_\ell) = \sum_{k=1}^m (f \cdot e_k)(\overline{g \cdot e_k})$ für beliebige $f, g \in H$ und $m \in \mathbb{N}$. Durch Grenzübergang $m \to \infty$ mit dem Satz folgt hieraus $f \cdot g = \langle (f \cdot e_k)_{k\in\mathbb{N}}, (g \cdot e_k)_{k\in\mathbb{N}} \rangle_{\ell^2}$, also die Wohldefiniertheit und Isometrie der angegebenen Abbildung. Da Linearität dieser Abbildung offensichtlich ist, bleibt nur ihre Surjektivität nachzuweisen. Dazu sei $(c_k)_{k\in\mathbb{N}} \in \ell^2(\mathbb{K})$. Aufgrund von $\|\sum_{k=m}^n c_k e_k\|_H^2 = \sum_{k=m}^n |c_k|^2$ für $m < n$ in $\mathbb{N}$ überträgt sich die Cauchy-Eigenschaft von den Partialsummen von $\sum_{k=1}^\infty |c_k|^2 < \infty$ auf die von $\sum_{k=1}^\infty c_k e_k$. Folglich existiert $f := \sum_{k=1}^\infty c_k e_k \in H$, und mit der Stetigkeit des Skalarprodukts und der Vollständigkeit von $(e_k)_{k\in\mathbb{N}}$ folgt wie im Beweis des Satzes $c_\ell = f \cdot e_\ell$ für alle $\ell \in \mathbb{N}$. Damit ist Surjektivität nachgewiesen. □

Beweisskizze zum klassischen Satz von Riesz-Fischer. Für die durch $\mathrm{e}_k(x) := \mathrm{e}^{\mathbf{i}k\cdot x}$ definierten Funktionen ergibt sich mit dem Satz von Fubini

$$\begin{aligned}\mathrm{e}_k \cdot \mathrm{e}_\ell = \int_{[0,2\pi)^N} \mathrm{e}_k\overline{\mathrm{e}_\ell}\,\mathrm{d}\frac{\mathscr{L}^N}{(2\pi)^N} &= \frac{1}{(2\pi)^N}\int_{[0,2\pi)^N} \mathrm{e}^{\mathbf{i}(k-\ell)\cdot x}\,\mathrm{d}\mathscr{L}^N(x)\\ &= \prod_{j=1}^N \left(\frac{1}{2\pi}\int_0^{2\pi} \mathrm{e}^{\mathbf{i}(k_j-\ell_j)x_j}\,\mathrm{d}x_j\right) = \prod_{j=1}^N \delta_{k_j\ell_j} = \delta_{k\ell}\end{aligned}$$

für alle $k,\ell \in \mathbb{Z}^N$ (wobei die Berechnung der 1-dimensionalen Integrale mit dem HDI oder mit Symmetrieüberlegungen erfolgen kann). Damit ist $(e_k)_{k\in\mathbb{Z}^N}$ ein Orthonormalsystem in $\mathrm{L}^2 = \mathrm{L}^2\big([0,2\pi)^N,(2\pi)^{-N}\mathscr{L}^N;\mathbb{C}\big)$.

Es wird nun nachgewiesen, dass das Orthonormalsystem $(e_k)_{k\in\mathbb{Z}^N}$ auch vollständig ist. Dazu wird zuerst die Fourier-Darstellung $\sum_{k\in\mathbb{Z}^N}(f\cdot\mathrm{e}_k)\mathrm{e}_k = f$ solcher $f\colon [0,2\pi)^N \to \mathbb{C}$ verifiziert, die die Form $f(x) = \prod_{j=1}^N f_j(x_j)$ für $x \in [0,2\pi)^N$ mit 2π-periodischen C^1-Funktionen $f_i\colon \mathbb{R}\to\mathbb{C}$ haben: In dieser Situation erhält man gemäß dem Satz von Fubini die Fourier-Koeffizienten $f\cdot\mathrm{e}_k$ von f in Produktform

$$f\cdot\mathrm{e}_k = \frac{1}{(2\pi)^N}\int_{[0,2\pi)^N}\prod_{j=1}^N f_j(x_j)\mathrm{e}^{-\mathbf{i}k_jx_j}\,\mathrm{d}\mathscr{L}^N(x) = \prod_{j=1}^N\left(\frac{1}{2\pi}\int_0^{2\pi} f_j(x_j)\mathrm{e}^{-\mathbf{i}k_jx_j}\,\mathrm{d}x_j\right) = \prod_{j=1}^N(f_j\cdot\mathrm{e}_{k_j})$$

aus den Fourier-Koeffizienten $f_j\cdot\mathrm{e}_{k_j}$ der f_j. Für die ganze Fourier-Reihe ergeben sich damit Produktgestalt und (in $x\in[0,2\pi)^N$ gleichmäßige und sogar normale) Konvergenz

$$\sum_{k\in\mathbb{Z}^N}(f\cdot\mathrm{e}_k)\mathrm{e}_k(x) = \sum_{k\in\mathbb{Z}^N}\prod_{j=1}^N(f_j\cdot\mathrm{e}_{k_j})\mathrm{e}_{k_j}(x_j) = \prod_{j=1}^N\sum_{k_j\in\mathbb{Z}}(f_j\cdot\mathrm{e}_{k_j})\mathrm{e}_{k_j}(x_j) = \prod_{j=1}^N f_j(x_j) = f(x)$$

aus der bekannten gleich,äßigen Konvergenz der Fourier-Reihen von C^1-Funktionen in Dimension 1. Insbesondere liegen die f der betrachteten Form im Abschluss $\overline{U}$ des von $(\mathrm{e}_k)_{k\in\mathbb{Z}^N}$ aufgespannten Unterraums U von L^2. Da charakteristische Funktionen $\mathbb{1}_Q$ von Quadern $Q \in \mathcal{I}_N \cap \mathcal{P}\big([0,2\pi)^N\big)$ in der L^2-Norm durch die betrachteten f beliebig gut approximiert werden können, liegen auch solche $\mathbb{1}_Q$ in $\overline{U}$. In weiteren Schritten können erst charakteristische Funktionen $\mathbb{1}_B$ zu $B \in \mathcal{B}\big([0,2\pi)^N\big)$ und dann (Standard-Ausdehnungsprozedur!) beliebige $g\in\mathrm{L}^2$ durch Linearkombinationen von Funktionen $\mathbb{1}_Q$ beliebig gut in der L^2-Norm approximiert werden, beliebige $g\in\mathrm{L}^2$ liegen also ebenfalls in $\overline{U}$. Damit gilt $\overline{U} = \mathrm{L}^2$, und U liegt dicht in L^2. Wie beim Beweis von Teil (II) des vorigen Korollars ist $(\mathrm{e}_k)_{k\in\mathbb{Z}^N}$ somit ein vollständiges Orthonormalsystem in L^2. Mit der in Abschnitt 2.9 nachgewiesenen Vollständigkeit des Skalarproduktraums L^2 (der damit ein Hilbertraum ist) und dem letzten Satz beziehungsweise dem Teil (I) des Korollars lässt sich schließen, dass $(\mathrm{e}_k)_{k\in\mathbb{Z}^N}$ sogar eine Orthonormalbasis von L^2 ist.

Jetzt folgen alle Behauptungen des Satzes: Gemäß Teil (III) des Korollars ist die Abbildung $\mathrm{L}^2 \mapsto \ell^2(\mathbb{Z}^N;\mathbb{C})$, $f\mapsto (f\cdot\mathrm{e}_k)_{k\in\mathbb{Z}^N}$ wohldefiniert und isometrischer Isomorphismus. Schreibt man aus, dass dieser Isomorphismus Norm und Skalarprodukt erhält, so ergeben sich die Besselsche und die Parsevalsche Gleichung. Die Darstellung $f = \sum_{k\in\mathbb{Z}^N}(f\cdot\mathrm{e}_k)\mathrm{e}_k$ von $f\in\mathrm{L}^2$ als Fourier-Reihe, schließlich, entspricht der bereits nachgewiesenen Eigenschaft, dass $(\mathrm{e}_k)_{k\in\mathbb{Z}^N}$ eine Orthonormalbasis von L^2 ist. □

Von großer Bedeutung ist auch die folgende Bildung, die viel mit Fourier-Reihen gemein hat, aber statt diskreter Frequenzen $k \in \mathbb{Z}^N$ nun kontinuierliche Frequenzen $\xi \in \mathbb{R}^N$ verwendet:

Definition (**Fourier-Transformierte, Fourier-Transformation**)**.** *Die **Fourier-Transformierte** $\mathscr{F}f = \widehat{f}\colon \mathbb{R}^N \to \mathbb{C}$ von $f \in \mathrm{L}^1 = \mathrm{L}^1(\mathbb{R}^N, \mathscr{L}^N; \mathbb{C})$ wird definiert durch*

$$\mathscr{F}f(\xi) = \widehat{f}(\xi) := \frac{1}{(2\pi)^{N/2}} \int_{\mathbb{R}^N} f(x) \mathrm{e}^{-\mathbf{i}\xi\cdot x} \,\mathrm{d}\mathscr{L}^N(x) \in \mathbb{C} \qquad \text{für } \xi \in \mathbb{R}^N$$

(*wobei wieder* $\mathrm{e}^{-\mathbf{i}\xi\cdot x} = \exp(-\mathbf{i}(\xi_1 x_1 + \xi_2 x_2 + \ldots \xi_N x_N))$ *zu verstehen ist*)*. Unter der **Fourier-Transformation*** (*auf* L^1) *versteht man die Zuordnungsvorschrift* $\mathscr{F}\colon f \mapsto \widehat{f}$*.*

Bemerkungen (zur Fourier-Transformation)**.** Wie in der Definition sei $f \in \mathrm{L}^1$.

(0) Die Fourier-Transformierte $\widehat{f}$ ist wohldefiniert und beschränkt, genauer ergibt sich aus der Dreiecksungleichung die Abschätzung

$$\big|\widehat{f}(\xi)\big| \le (2\pi)^{-N/2} \|f\|_{\mathrm{L}^1} \qquad \text{für alle } \xi \in \mathbb{R}^N\,.$$

(1) Die **Fourier-Transformation** ist eine **$\mathbb{C}$-lineare Operation**, d.h. für $f, g \in \mathrm{L}^1$ und $r, s \in \mathbb{C}$ gilt stets

$$\widehat{rf{+}sg} = r\widehat{f}{+}s\widehat{g}\,.$$

(2) In Verschärfung von (0) gilt sogar

$$\widehat{f} \in \mathrm{C}^0(\mathbb{R}^N; \mathbb{C}) \qquad\qquad \text{mit Abfallbedingung } \lim_{|\xi|\to\infty} \widehat{f}(\xi) = 0\,,$$

wobei der zweiten Teil dieser Aussage auch als **Riemann-Lebesgue-Lemma** bekannt ist.

Beweis. Stetigkeit von $\widehat{f}$ folgt sofort aus dem Satz über dominierte Konvergenz (mit Majorante $|f|$). Die Abfallbedingung verifiziert man zuerst für charakteristische Funktionen von Quadern $Q = [a_1, b_1){\times}[a_2, b_2){\times} \ldots {\times}[a_N, b_N) \in \mathcal{I}_N$. Für solche liefert nämlich der Satz von Fubini

$$\widehat{\mathbb{1}_Q}(\xi) = \frac{1}{(2\pi)^{N/2}} \int_Q \mathrm{e}^{-\mathbf{i}\xi\cdot x} \,\mathrm{d}\mathscr{L}^N(x) = \frac{1}{(2\pi)^{N/2}} \prod_{j=1}^N \int_{a_j}^{b_j} \mathrm{e}^{-\mathbf{i}\xi_j x_j} \,\mathrm{d}x_j\,,$$

das x_j-Integral auf der rechten Seite hat für $\xi_j \neq 0$ den Wert $\frac{\mathbf{i}}{\xi_j}(\mathrm{e}^{-\mathbf{i}\xi_j b_j} {-} \mathrm{e}^{-\mathbf{i}\xi_j a_j})$ (und für $\xi_j = 0$ den Wert $b_j{-}a_j$), so dass man $\lim_{|\xi|\to\infty} \widehat{\mathbb{1}_Q}(\xi) = 0$ erhält. Mit (1) folgt die Behauptung nun für endliche Linearkombinationen h solcher Funktionen $\mathbb{1}_Q$. Mit solchen Linearkombinationen h kann man dann erst $\mathbb{1}_B$ zu $B \in \mathcal{B}(\mathbb{R}^N)$ mit $\mathscr{L}^N(B) < \infty$ (siehe Übungsblatt 5!) und dann beliebige $f \in \mathrm{L}^1$ (Approximation durch Treppenfunktionen plus Konvergenzsatz!) in der L^1-Norm approximieren. Mit der Abschätzung $|\widehat{f}(\xi){-}\widehat{h}(\xi)| \le (2\pi)^{-N/2} \|f{-}h\|_{\mathrm{L}^1}$ gemäß (0) und (1) erhält man gleichmäßige Konvergenz der Fourier-Transformierten der Approximationen und überträgt die Behauptung dann auf beliebige $f \in \mathrm{L}^1$. □

(3) Folgende Rechenregeln für Fourier-Transformierte werden in den Übungen hergeleitet:

- Translationsregel (bei festem $x_0 \in \mathbb{R}^N$): $f_{x_0}(x) = f(x_0{+}x) \implies \widehat{f_{x_0}}(\xi) = \mathrm{e}^{\mathbf{i}\xi\cdot x_0} \widehat{f}(\xi)$,
- Skalierungsregel (bei festem $r \in \mathbb{R} \setminus \{0\}$): $f_r(x) = f(rx) \implies \widehat{f_r}(\xi) = |r|^{-N} \widehat{f}(r^{-1}\xi)$,

- Konjugationsregel: $\widehat{\overline{f}}(\xi) = \overline{\widehat{f}}(-\xi)$,
- **Ableitungsregeln** (bei festem Index $j \in \{1, 2, \ldots, N\}$ beziehungsweise Multiindex $\alpha \in \mathbb{N}_0^N$ und sofern $\partial_j f$ bzw. $\partial^\alpha f$ als stetige L^1-Funktion auf ganz $\mathbb{R}^N$ existiert):

$$\widehat{\partial_j f}(\xi) = \mathbf{i}\xi_j \widehat{f}(\xi) \qquad \text{und} \qquad \widehat{\partial^\alpha f}(\xi) = \mathbf{i}^{|\alpha|}\xi^\alpha \widehat{f}(\xi)\,.$$

(4) Allgemeiner kann man eine Fourier-Transformierte $\widehat{\mu}\colon \mathbb{R}^N \to \mathbb{C}$ auch für endliche Maße μ auf $(\mathbb{R}^N, \mathscr{A})$ mit $\mathcal{B}(\mathbb{R}^N) \subset \mathcal{M}_\mu$ durch $\widehat{\mu}(\xi) := (2\pi)^{-N/2} \int_{\mathbb{R}^N} \mathrm{e}^{-\mathbf{i}\xi\cdot x}\,\mathrm{d}\mu(x) \in \mathbb{C}$ für $\xi \in \mathbb{R}^N$ erklären. Für nichtnegatives $f \in \mathrm{L}^1$ gilt dann $\widehat{f\cdot\mathscr{L}^N} = \widehat{f}$.

Definition (**Fourier-Rücktransformation, inverse Fourier-Transformation**). *Die* ***Fourier-Rücktransformierte*** *oder* ***Invers-Fourier-Transformierte*** $\mathscr{F}^{-1}f = \check{f}\colon \mathbb{R}^N \to \mathbb{C}$ *von* $f \in \mathrm{L}^1 = \mathrm{L}^1(\mathbb{R}^N, \mathscr{L}^N; \mathbb{C})$ *erklärt man durch* $\mathscr{F}^{-1}f(x) = \check{f}(x) := \widehat{f}(-x) \in \mathbb{C}$ *für* $x \in \mathbb{R}^N$.

Diese Benennung wird begründet durch:

Satz (**Fourier-Umkehrsatz, Fourier-Inversions-Satz, Fourier-Umkehrformel**). *Für* $f \in \mathrm{L}^1 = \mathrm{L}^1(\mathbb{R}^N, \mathscr{L}^N; \mathbb{C})$ *mit* $\widehat{f} \in \mathrm{L}^1$ *gilt*

$$\mathscr{F}^{-1}(\mathscr{F}f) = f \ \text{ bzw. } \check{\widehat{f}} = f \qquad \text{in } \mathrm{L}^1\,, \ \text{ also } \mathscr{L}^N\text{-fast-überall auf } \mathbb{R}^N$$

(*und sogar punktweise auf ganz* $\mathbb{R}^N$, *falls* f *stetig ist*).

Auf einen *Beweis* des Satzes wird hier verzichtet, die wesentlichen Aspekte werden im Rahmen der Übungen behandelt. □

Bemerkung (zur Voraussetzung $\widehat{f} \in \mathrm{L}^1$). Die Voraussetzung $\widehat{f} \in \mathrm{L}^1$ im Fourier-Umkehrsatz kann auch ohne explizite Kenntnis der Fourier-Transformierten überprüft werden: Beispielsweise ergibt sich aus der Ableitungsregel der Bemerkung (3) das hinreichende Kriterium, dass für alle $i \in \{1, 2, \ldots, N\}$ die reine $(N{+}1)$-te Ableitung $\partial_i^{N+1}f$ von f existiert, stetig ist und zu L^1 gehört.

Bemerkung (zur Fourier-Inversion und der **Bedeutung der Fourier-Transformation**). Mit anderen Worten liefert der Fourier-Umkehrsatz die Darstellung

$$f(x) = \frac{1}{(2\pi)^{N/2}} \int_{\mathbb{R}^N} \mathrm{e}^{\mathbf{i}\xi\cdot x} \widehat{f}(\xi)\,\mathrm{d}\mathscr{L}^N(\xi) \qquad \text{für } x \in \mathbb{R}^N$$

von f als **Superposition und damit als eine Art „kontinuierliche Linearkombination“ von Wellen/Schwingungen des Basis-Typs** $x \mapsto \mathrm{e}^{\mathbf{i}\xi\cdot x}$ mit $\xi \in \mathbb{R}^N$.

Es besteht somit eine enge Verwandtschaft zu Fourier-Reihen, die ja abzählbare „diskrete“ Linearkombinationen von $x \mapsto \mathrm{e}^{\mathbf{i}k\cdot x}$ mit $k \in \mathbb{Z}^N$ sind. Als (fürs Erste) wesentlicher Unterschied ist aber zu verzeichnen, dass obige Darstellung anders als die Fourier-Reihen-Darstellung keinerlei Periodizität von f erfordert. Zum Preis, den man hierfür zahlt, gehört, dass man nun beliebige „Frequenzen“ $\xi \in \mathbb{R}^N$ statt nur diskreten $k \in \mathbb{Z}^N$ in Betracht ziehen muss.

Vor allem zeigt die Fourier-Umkehrformel, dass die Fourier-Transformierte $\widehat{f}$ alle Informationen über die Fastfunktion f beinhaltet und diese in gewisse Informationen über zu f gehörige Schwingungen und Frequenzen ξ umsetzt. Somit kann man die **Untersuchung von** $\boldsymbol{\widehat{f}}$ (und in periodischen Fällen auch schon die Untersuchung der Fourier-Koeffizienten) als eine **Analyse von** $\boldsymbol{f}$ **„nach Frequenzen“** betrachten. Dieses Vorgehen „nach Frequenzen“ bildet die **eigentliche Grundidee der Fourier-Analysis**.

Zum Abschluss dieses Abschnitts wird noch das kontinuierliche Analogon des klassischen Satzes von Riesz-Fischer behandelt:

Satz (von Plancherel). *Sei* $\mathrm{L}^1 = \mathrm{L}^1(\mathbb{R}^N, \mathscr{L}^N; \mathbb{C})$ *und* $\mathrm{L}^2 = \mathrm{L}^2(\mathbb{R}^N, \mathscr{L}^N; \mathbb{C})$. *Die Fourier-Transformation* $\mathscr{F}$ *bildet* $\mathrm{L}^1 \cap \mathrm{L}^2$ *in* L^2 *ab, und* $\mathscr{F}|_{\mathrm{L}^1\cap\mathrm{L}^2}$ *besitzt eine eindeutige Fortsetzung zu einem isometrischen Isomorphismus* $\mathscr{F}_{\mathrm{L}^2} \colon \mathrm{L}^2 \to \mathrm{L}^2$.

Beweisskizze. Sei

$$\mathcal{G} := \left\{ f \in \mathrm{L}^1 \cap \mathrm{L}^2 \,:\, \widehat{f} \in \mathrm{L}^1 \cap \mathrm{L}^2 \right\} \subset \mathrm{L}^1 .$$

Um das L^2-Skalarprodukt von $f, g \in \mathcal{G}$ geeignet umschreiben zu können, wird nun $g_* \in \mathcal{G}$ durch $g_*(x) := \overline{g}(-x)$ definiert. Gemäß der Skalierungsregel (mit $r = -1$) und der Konjugationsregel aus Bemerkung (3) gilt dann $\widehat{g}_*(\xi) = \widehat{\overline{g}}(-\xi) = \overline{\widehat{g}}(\xi)$. Mit den Definitionen des L^2-Skalarprodukts und der Fourier-Transformierten und durch Translation der Integrationsvariablen erhält man

$$\begin{aligned}
\langle \widehat{f}, \widehat{g} \rangle_{\mathrm{L}^2} = \int_{\mathbb{R}^N} \widehat{f}(\xi) \widehat{g}_*(\xi) \,\mathrm{d}\xi &= \int_{\mathbb{R}^N} \frac{1}{(2\pi)^{N/2}} \int_{\mathbb{R}^N} f(x) \mathrm{e}^{-\mathbf{i}\xi\cdot x} \,\mathrm{d}x \, \frac{1}{(2\pi)^{N/2}} \int_{\mathbb{R}^N} g_*(y) \mathrm{e}^{-\mathbf{i}\xi\cdot y} \,\mathrm{d}y \,\mathrm{d}\xi \\
&= \frac{1}{(2\pi)^{N/2}} \int_{\mathbb{R}^N} \frac{1}{(2\pi)^{N/2}} \int_{\mathbb{R}^N} \int_{\mathbb{R}^N} f(x) g_*(y) \mathrm{e}^{-\mathbf{i}\xi\cdot(x+y)} \,\mathrm{d}y \,\mathrm{d}x \,\mathrm{d}\xi \\
&= \frac{1}{(2\pi)^{N/2}} \int_{\mathbb{R}^N} \frac{1}{(2\pi)^{N/2}} \int_{\mathbb{R}^N} \int_{\mathbb{R}^N} f(x) g_*(z-x) \mathrm{e}^{-\mathbf{i}\xi\cdot z} \,\mathrm{d}z \,\mathrm{d}x \,\mathrm{d}\xi \,.
\end{aligned}$$

Die z- und die x-Integration können wegen $\int_{\mathbb{R}^N} \int_{\mathbb{R}^N} |f(x) g_*(z-x) \mathrm{e}^{-\mathbf{i}\xi\cdot z}| \,\mathrm{d}z \,\mathrm{d}x = \|f\|_{\mathrm{L}^1} \|g_*\|_{\mathrm{L}^1} < \infty$ mit Fubini vertauscht werden. Deshalb ergibt sich weiter

$$\ldots = \frac{1}{(2\pi)^{N/2}} \int_{\mathbb{R}^N} \frac{1}{(2\pi)^{N/2}} \int_{\mathbb{R}^N} \mathrm{e}^{-\mathbf{i}\xi\cdot z} \int_{\mathbb{R}^N} f(x) g_*(z-x) \,\mathrm{d}x \,\mathrm{d}z \,\mathrm{d}\xi \,.$$

Durch im Wesentlichen die gleichen Umformungen ergibt sich in der betrachteten Situation mit $f, g_* \in \mathrm{L}^1$ zudem, dass durch $h(z) := \int_{\mathbb{R}^N} f(x) g_*(z-x) \,\mathrm{d}x$ eine Hilfsfunktion $h \in \mathrm{L}^1$ mit $\widehat{h} = (2\pi)^{N/2} \widehat{f}\, \widehat{g}_*$ definiert wird, und aus $\widehat{f} \in \mathrm{L}^2$, $\widehat{g}_* \in \mathrm{L}^2$ folgt mit der Hölder-Ungleichung $\widehat{h} \in \mathrm{L}^1$. Damit kann obige Rechnung durch

$$\begin{aligned}
\ldots &= \frac{1}{(2\pi)^{N/2}} \int_{\mathbb{R}^N} \frac{1}{(2\pi)^{N/2}} \int_{\mathbb{R}^N} \mathrm{e}^{-\mathbf{i}\xi\cdot z} h(z) \,\mathrm{d}z \,\mathrm{d}\xi \\
&= \frac{1}{(2\pi)^{N/2}} \int_{\mathbb{R}^N} \widehat{h}(\xi) \,\mathrm{d}\xi = \check{\widehat{h}}(0)
\end{aligned}$$

fortgeführt werden. Tatsächlich stellt sich die Hilfsfunktion h sogar als stetig heraus: Bei stetigem g_* mit $g_* \equiv 0$ außerhalb eines Kompaktums $K \subset \mathbb{R}^N$ ergibt sich dies direkt aus dem Satz über dominierte Konvergenz. Für beliebiges $g_* \in \mathrm{L}^2$ kann man stetige g_k mit $g_k \equiv 0$ außerhalb von Kompakta $K_k \subset \mathbb{R}^N$ und mit $g_k \to g_*$ in der L^2-Norm finden. Da gemäß der Hölder-Ungleichung $\left| \int_{\mathbb{R}^N} f(x) g_k(z-x) \,\mathrm{d}x - \int_{\mathbb{R}^N} f(x) g_*(z-x) \,\mathrm{d}x \right| \le \|f\|_{\mathrm{L}^2} \|g_k - g_*\|_{\mathrm{L}^2}$ *gleichmäßig* in $z \in \mathbb{R}^N$ gegen Null geht, folgt die Stetigkeit von h auch im allgemeinen Fall. Mit $h \in \mathrm{L}^1$, $\widehat{h} \in \mathrm{L}^1$ und der Stetigkeit von h liegen alle Voraussetzungen vor, um den Fourier-Inversions-Satz für h sogar in einzelnen Punkten anwenden zu dürfen, und die begonnene Rechnung kann durch

$$\ldots = h(0) = \int_{\mathbb{R}^N} f(x) g_*(-x) \,\mathrm{d}x = \langle f, g \rangle_{\mathrm{L}^2}$$

komplettiert werden. Insgesamt ist gezeigt, dass $\mathscr{F}|_{\mathcal{G}}$ das L^2-Skalarprodukt und damit auch die L^2-Norm erhält und insbesondere injektiv ist. Aus dem Fourier-Inversions-Satz folgt außerdem, dass jedes $f \in \mathcal{G}$ als $f = \mathscr{F}(\mathscr{F} f^\sharp)$ geschrieben werden kann, wobei man $f^\sharp \in \mathcal{G}$ mit $\widehat{f^\sharp} \in \mathcal{G}$ durch Punktspiegelung $f^\sharp(x) := f(-x)$ erhält. Insgesamt ist $\mathscr{F}|_{\mathcal{G}}$ also isometrischer Isomorphismus von $\mathcal{G}$ auf sich.

Als Nächstes gilt es zu begründen, dass der Unterraum $\mathcal{G}$ in L^2 (mit der L^2-Norm) dicht liegt. Hierzu argumentiert man mit Hilfe von Funktionenräumen wie dem **Schwartz-Raum** $\mathcal{S}(\mathbb{R}^N;\mathbb{C}) := \{f \in \mathrm{C}^\infty(\mathbb{R}^N;\mathbb{C}) : \lim_{|x|\to\infty} |x|^k \partial^\alpha f(x) = 0 \text{ für alle } k \in \mathbb{N}\,,\ \alpha \in \mathbb{N}_0^N\}$ der glatten und schneller als polynomial abfallenden Funktionen oder dem Raum der glatten Testfunktionen $\mathcal{D}(\mathbb{R}^N;\mathbb{C}) := \{f \in \mathrm{C}^\infty(\mathbb{R}^N;\mathbb{C}) : f \equiv 0 \text{ außerhalb eines Kompaktums}\}$, für die sich Dichtheit in L^2 mit einer Standard-Methode der fortgeschrittenen Analysis, der sogenannten Glättung nachweisen lässt. Die Dichtheit von $\mathcal{G}$ folgt dann, weil einerseits $\mathcal{D}(\mathbb{R}^N;\mathbb{C}) \subset \mathcal{S}(\mathbb{R}^N;\mathbb{C}) \subset \mathrm{L}^1 \cap \mathrm{L}^2$ gilt und andererseits für $f \in \mathcal{S}(\mathbb{R}^N;\mathbb{C})$ stets[36] $\widehat{f} \in \mathcal{S}(\mathbb{R}^N;\mathbb{C})$ nachgewiesen werden kann (mit Differentiation unter dem Fourier-Integral und Ableitungsregeln; zumindest ansatzweise wird ein ähnliches Vorgehen im Rahmen der Übungen behandelt).

Aufgrund der Dichtheit von $\mathcal{G}$ im Definitionsbereich L^2 und der Vollständigkeit der Zielraums L^2 kann $\mathscr{F}|_{\mathcal{G}}$ jetzt zu einer isometrischen $\mathbb{C}$-linearen Abbildung $\mathscr{F}_{\mathrm{L}^2} \colon \mathrm{L}^2 \to \mathrm{L}^2$ fortgesetzt werden. Konvergieren $f_k \in \mathcal{G}$ in L^2 gegen f, so folgt $f = \lim_{k\to\infty} f_k = \lim_{k\to\infty} \mathscr{F}(\mathscr{F} f_k^\sharp) = \lim_{k\to\infty} \mathscr{F}_{\mathrm{L}^2}(\mathscr{F}_{\mathrm{L}^2} f_k^\sharp) = \mathscr{F}_{\mathrm{L}^2}(\mathscr{F}_{\mathrm{L}^2}(\lim_{k\to\infty} f_k^\sharp)) = \mathscr{F}_{\mathrm{L}^2}(\mathscr{F}_{\mathrm{L}^2} f^\sharp)$ (wobei $^\sharp$ wie oben für die Punktspiegelung steht und alle Limites in L^2 zu nehmen sind). Deshalb ist $\mathscr{F}_{\mathrm{L}^2} \colon \mathrm{L}^2 \to \mathrm{L}^2$ auch surjektiv und insgesamt isometrischer Isomorphismus.

Zuletzt wird begründet, dass die abstrakt erhaltene L^2-Fourier-Transformation $\mathscr{F}_{\mathrm{L}^2}$ nicht nur auf $\mathcal{G}$ (wo dies nach Konstruktion klar ist), sondern auf ganz $\mathrm{L}^1 \cap \mathrm{L}^2$ mit dem konkreten Fourier-Integral $\mathscr{F}$ übereinstimmt. Dazu benutzt man für $f \in \mathrm{L}^1 \cap \mathrm{L}^2$ Approximationen $f_k \in \mathcal{G}$ (wie sie sich beim Nachweis der schon verwendeten Dichtheitsaussagen mit ergeben), so dass f_k in der L^1-Norm *und* in der L^2-Norm gegen f konvergieren. Dann folgt einerseits $\lim_{k\to\infty} \mathscr{F} f_k = \lim_{k\to\infty} \mathscr{F}_{\mathrm{L}^2} f_k = \mathscr{F}_{\mathrm{L}^2} f$ in der L^2-Norm, und entlang einer Teilfolge konvergiert $\lim_{\ell\to\infty} \mathscr{F} f_{k_\ell}(\xi) = \mathscr{F}_{\mathrm{L}^2} f(\xi)$ für $\mathscr{L}^N$-fast-alle $\xi \in \mathbb{R}^N$. Andererseits zeigt direkte Abschätzung der Fourier-Integrale $\lim_{k\to\infty} \mathscr{F} f_k(\xi) = \mathscr{F} f(\xi)$ sogar für alle $\xi \in \mathbb{R}^N$, so dass $\mathscr{F}_{\mathrm{L}^2} f = \mathscr{F} f$ in L^2 gelten muss. □

Bemerkung (zur Interpretation der L^2-Fourier-Transformation)**.** Für $f \in \mathrm{L}^2 \setminus \mathrm{L}^1$ ist die als Integral definierte Fourier-Transformation $\mathscr{F} f$ nicht erklärt, die L^2-Fourier-Transformation $\mathscr{F}_{\mathrm{L}^2} f$ über (zunächst) abstrakte Fortsetzung aber schon. Da die Abschneidungen $\mathbb{1}_{\mathrm{B}_R} f$ bei $R \to \infty$ in L^2 gegen f konvergieren, gilt tatsächlich $\mathscr{F}_{\mathrm{L}^2} f = \lim_{R\to\infty} \mathscr{F}(\mathbb{1}_{\mathrm{B}_R} f)$ in L^2, und zumindest entlang einer Unendlich-Folge $(R_k)_{k\in\mathbb{N}}$ erhält man auch eine Darstellung

$$\mathscr{F}_{\mathrm{L}^2} f(\xi) = \lim_{k\to\infty} \frac{1}{(2\pi)^{N/2}} \int_{\mathrm{B}_{R_k}} f(x) \mathrm{e}^{-\mathbf{i}\xi\cdot x} \,\mathrm{d}\mathscr{L}^N(x) \qquad \text{für } \mathscr{L}^N\text{-fast-alle } \xi \in \mathbb{R}^N$$

als eine Art uneigentliches Fourier-Integral.

[36] Mit anderen Worten bildet die Fourier-Transformation $\mathcal{S}(\mathbb{R}^N;\mathbb{C})$ in sich ab, und wegen des Fourier-Inversions-Satzes ist $\mathcal{S}(\mathbb{R}^N;\mathbb{C})$ (neben L^2, für den dies ja gerade gezeigt wird) dann ein weiteres (Standard-)Beispiel eines Funktionenraums, den die Fourier-Transformation bijektiv auf sich selbst abbildet. In dieser Eigenschaft liegt die eigentliche Bedeutung des Schwartz-Raums $\mathcal{S}(\mathbb{R}^N;\mathbb{C})$.

2.11 Hausdorff-Maße, Transformations- und Flächenformel

An dieser Stelle wird das **Problem des k-dimensionalen Inhalts von Teilmengen von $\mathbb{R}^N$** mit $k \in \{0,1,2,\ldots,N-1,N\}$ erneut aufgegriffen. In den Extremfällen $k=0$ und $k=N$ kann ein solcher Inhalt mit dem Zählmaß beziehungsweise dem Lebesgue-Maß gemessen und das Problem (auf einer prinzipiellen Ebene) als gelöst angesehen werden. Im Folgenden werden darüber hinaus allgemeine Dimensionen $k \in \{0,1,2,\ldots,N\}$ betrachtet und behandelt. Dazu greift man auf eine Variante der Carathéodory-Konstruktion zurück: Man überdeckt $A \in \mathcal{P}(\mathbb{R}^N)$ mit N-dimensionalen Kugeln, summiert aber nur die **Inhalte der k-dimensionalen Äquatorebenen** und erhofft sich eine sinnvolle Messung des k-dimensionalen Inhalts von $A \in \mathcal{P}(\mathbb{R}^N)$ durch den Ausdruck

$$\inf\left\{\sum_{i=1}^{\infty}\omega_k r_i^k \,:\, A \subset \bigcup_{i=1}^{\infty} \mathrm{B}_{r_i}^N(x_i)\right\} \qquad (*)$$

mit $\mathrm{B}_r^N(x) = \{y \in \mathbb{R}^N \,:\, |y-x| < r\}$, $\omega_0 := 1$, $\omega_k = \mathscr{L}^k(\mathrm{B}_1^k(0))$ für $k \in \mathbb{N}$. Der Ansatz $(*)$ ist allerdings insofern zu naiv, als dass er bei Verwendung großer Kugeln einen zu kleinen Flächeninhalt gekrümmter Kurven oder Flächen liefert (rote Linie im Bild). Erzwingt man jedoch, wie in der folgenden Definition, Kleinheit der Radien, so verschwindet dieses Problem und man erhält sogar in metrischen Räumen Ω und für nicht-ganzzahlige Werte s anstelle von k ein sinnvolles Konzept. Motiviert durch die aus Abschnitt 2.8 bekannte Formel $\omega_k = \frac{\pi^{\frac{k}{2}}}{\Gamma(\frac{k}{2}+1)}$ für $k \in \mathbb{N}_0$ mit der Gamma-Funktion Γ verwendet man dabei auch für beliebiges $s \in \mathbb{R}_{\geq 0}$ den Vorfaktor

$$\omega_s := \frac{\pi^{\frac{s}{2}}}{\Gamma(\frac{s}{2}+1)} \in \mathbb{R}_{>0}\,.$$

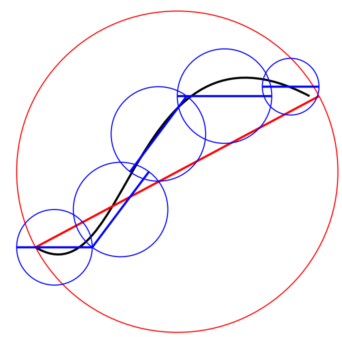


Abb. 5: Gute und schlechte Näherungen an die Kurvenlänge

Definition & Satz (Hausdorff-Maße). *Sei Ω ein metrischer Raum, sei $s \in \mathbb{R}_{\geq 0}$, und sei $\delta \in (0,\infty]$. Für $A \in \mathcal{P}(\Omega)$ setzt man*[37]

$$\mathcal{H}_\delta^s(A) := \inf\left\{\sum_{i=1}^{\infty}\omega_s r_i^s \,:\, A \subset \bigcup_{i=1}^{\infty} \mathrm{B}_{r_i}(x_i) \text{ mit } 0 \leq r_i < \delta\right\}$$

mit offenen Kugeln $\mathrm{B}_r(x) = \{y \in \Omega \,:\, \mathrm{d}_\Omega(x,y) < r\}$ in Ω und

$$\mathcal{H}^s(A) := \lim_{\delta \searrow 0} \mathcal{H}_\delta^s(A)\,.$$

*Dann sind $\mathcal{H}_\delta^s$ und $\mathcal{H}^s$ äußere Maße über Ω, und $\mathcal{H}^s$ heißt das **s-dimensionale** (sphärische) **äußere Hausdorff-Maß** auf Ω. Gemäß Abschnitt 2.4 ist die Einschränkung von $\mathcal{H}^s$ auf die σ-Algebra $\mathcal{M}_{\mathcal{H}^s}$ seiner messbaren Mengen ein vollständiges Maß auf $(\Omega, \mathcal{M}_{\mathcal{H}^s})$. Man bezeichnet diese Einschränkung weiter mit $\mathcal{H}^s$ und nennt sie das **s-dimensionale** (sphärische) **Hausdorff-Maß** auf Ω.*

[37] Durch das Zulassen von $r_i = 0$ mit $\mathrm{B}_{r_i}(x_i) = \emptyset$ werden bei der Infimumbildung endliche Kugelüberdeckungen direkt eingeschlossen. Es ist aber auch üblich und ändert den Wert des Infimums nicht, nur positive r_i zuzulassen.

Bemerkungen (zur Definition und Interpretation von Hausdorff-Maßen)**.**

(1) Bei der Definition von $\mathcal{H}^s_\delta$ handelt es sich um eine Carathéodory-Konstruktion im Sinn von Abschnitt 2.4: Definiert man eine Messvorschrift K^s_δ auf dem System aller Kugeln $\mathrm{B}_r(x)$ in Ω mit $0 \le r < \delta$ durch $K^s_\delta(\mathrm{B}_r(x)) \coloneqq \omega_s r^s$, so gilt $\mathcal{H}^s_\delta = (K^s_\delta)^*$.

(2) Für $\widetilde{\delta} \le \delta$ in $(0, \infty]$ entnimmt man aus der Definition von $\mathcal{H}^s_\delta$ die Monotonie-Eigenschaft $\mathcal{H}^s_\delta(A) \le \mathcal{H}^s_{\widetilde{\delta}}(A)$. Hieraus folgt $\mathcal{H}^s(A) = \sup_{\delta>0} \mathcal{H}^s_\delta(A)$ für beliebiges $A \in \mathcal{P}(\Omega)$.

(3) In der Literatur gibt es etliche Varianten der Definition, und als Hausdorff-Maß ohne den Zusatz „sphärisch" wird meist eine Konstruktion bezeichnet, bei der man mit beliebigen beschränkten Teilmengen von Ω statt Kugeln überdeckt und die Radien r_i durch die halben Durchmesser der überdeckenden Mengen ersetzt. Diese und andere Varianten führen auf „gutartigen" Mengen zum gleichen Wert des Maßes, auf „exotischen" Mengen aber nicht unbedingt.

(4) Man interpretiert $\mathcal{H}^s(A)$ als s-dimensionalen Inhalt einer Teilmenge $A \subset \mathbb{R}^N$. Für ganzzahliges k und gutartiges A wird sich im Lauf dieses Abschnitts zeigen, dass man $\mathcal{H}^k(A)$ durch Integrationen berechnen kann und dabei sinnvolle Werte erhält.

Beweis des Satzes. Es ist nur zu begründen, dass $\mathcal{H}^s_\delta$ und $\mathcal{H}^s$ äußere Maße sind. Für $\mathcal{H}^s_\delta$ folgt dies unmittelbar aus Bemerkung (1) und dem Satz des Abschnitts 2.4 über die Carathéodory-Konstruktion. Um es auch für $\mathcal{H}^s$ nachzuweisen, seien $A_1, A_2, A_3, \ldots \in \mathcal{P}(\Omega)$. Mit Bemerkung (2) und der σ-Subadditivität von $\mathcal{H}^s_\delta$ folgt

$$\mathcal{H}^s\left(\bigcup_{i=1}^\infty A_i\right) = \sup_{\delta>0} \mathcal{H}^s_\delta\left(\bigcup_{i=1}^\infty A_i\right) \le \sup_{\delta>0} \sum_{i=1}^\infty \mathcal{H}^s_\delta(A_i) \le \sum_{i=1}^\infty \mathcal{H}^s(A_i)\,,$$

also σ-Subadditivität von $\mathcal{H}^s$. Der Nachweis der Monotonie von $\mathcal{H}^s$ verläuft analog, und damit ist auch $\mathcal{H}^s$ ein äußeres Maß. □

Um sinnvoll mit den Hausdorff-Maßen $\mathcal{H}^s$ arbeiten zu können, möchte man sicherstellen, dass die σ-Algebren $\mathcal{M}_{\mathcal{H}^s}$ der $\mathcal{H}^s$-messbaren Mengen möglichst groß sind und alle „vernünftigen" Mengen enthalten. Dies gelingt mit dem nächsten Satz:

Satz (über **„viele $\mathcal{H}^s$-messbare Mengen"**)**.** *Seien Ω ein metrischer Raum und $s \in \mathbb{R}_{\ge 0}$. Dann sind alle Borel-Mengen $\mathcal{H}^s$-messbar, d.h. es gilt $\mathcal{B}(\Omega) \subset \mathcal{M}_{h^s}$. Außerdem ist $\mathcal{H}^s$ ein $\mathcal{B}(\Omega)$-reguläres äußeres Maß.*

Ein Beweis des Satzes basiert auf einem Messbarkeits-Kriterium von Carathéodory, das mit dem Begriff des metrischen äußeren Maßes elegant formuliert werden kann.

Definition (**metrische äußere Maße**)**.** *Sei Ω ein metrischer Raum. Ein äußeres Maß α über Ω heißt **metrisches äußeres Maß**, wenn **auf** nicht-leeren **Mengen** $A, B \in \mathcal{P}(\Omega)$ **positiven Abstands** $\operatorname{dist}(A, \widetilde{A}) > 0$ stets die **Additivität** $\alpha(A \cup \widetilde{A}) = \alpha(A) + \alpha(\widetilde{A})$ vorliegt.*

Satz (von Carathéodory)**.** *Sei α ein äußeres Maß über einem metrischen Raum Ω. Genau dann gilt $\mathcal{B}(\Omega) \subset \mathcal{M}_\alpha$, wenn α ein metrisches äußeres Maß ist.*

Beweis des Satzes von Carathéodory. Sei $\mathcal{B}(\Omega) \subset \mathcal{M}_\alpha$, und seien $A, \widetilde{A} \in \mathcal{P}(\Omega)$ nicht-leer mit $\delta \coloneqq \operatorname{dist}(A, \widetilde{A}) > 0$. Dann betrachtet man die offene Menge $U_\delta(A) \coloneqq \{x \in \Omega : \operatorname{dist}(x, A) < \delta\}$, die nach Voraussetzung α-messbar im Sinn des Carathéodory-Kriteriums aus Abschnitt 2.4 ist. Mit $A \subset U_\delta(A)$ und $\widetilde{A} \cap U_\delta(A) = \emptyset$ folgt

$$\alpha(A \cup \widetilde{A}) = \alpha((A \cup \widetilde{A}) \cap U_\delta(A)) + \alpha((A \cup \widetilde{A}) \setminus U_\delta(A)) = \alpha(A) + \alpha(\widetilde{A})\,.$$

Damit ist α ein metrisches äußeres Maß.

Seien umgekehrt α ein metrisches äußeres Maß, A eine abgeschlossene Teilmenge von Ω und T eine beliebige Teilmenge von Ω mit $\alpha(T) < \infty$. Gemäß der Voraussetzung an α folgt dann

$$\alpha(T) \geq \alpha(T \cap A) + \alpha\big(T \setminus U_{1/i}(A)\big)$$

für alle $i \in \mathbb{N}$. Man setzt nun $H_i := \big[T \cap U_{1/i}(A)\big] \setminus U_{1/(i+1)}(A)$ und benutzt, dass H_{2i} und H_{2j} mit $i \neq j$, falls beide nicht-leer sind, stets positiven Abstand haben. Daraus ergibt sich mit der definierenden Eigenschaft des metrischen äußeren Maßes α induktiv $\sum_{i=1}^m \alpha(H_{2i}) = \alpha(\bigcup_{i=1}^m H_{2i}) \leq \alpha(T) < \infty$ und $\sum_{i=1}^m \alpha(H_{2i+1}) = \alpha(\bigcup_{i=1}^m H_{2i+1}) \leq \alpha(T) < \infty$. Man erhält $\sum_{i=1}^\infty \alpha(H_i) < \infty$. Da A abgeschlossen ist, gilt $T \setminus A = (T \setminus U_{1/m}(A)) \cup \bigcup_{i=m}^\infty H_i$, und die σ-Subadditivität von α liefert

$$\alpha(T \setminus A) \leq \alpha(T \setminus U_{1/m}(A)) + \sum_{i=m}^{\infty} \alpha(H_i)\,.$$

Insgesamt ist damit gezeigt, dass

$$\alpha(T) + \sum_{i=m}^{\infty} \alpha(H_i) \geq \alpha(T \cap A) + \alpha(T \setminus A)$$

gilt, wobei der Reihenrest links für $m \to \infty$ verschwindet. Die resultierende Ungleichung bleibt auch im Fall $\alpha(T) = \infty$ richtig, und in Anbetracht der σ-Subadditivität von α folgt

$$\alpha(T) = \alpha(T \cap A) + \alpha(T \setminus A)$$

für alle $T \in \mathcal{P}(\Omega)$. Also liegt jede abgeschlossene Teilmenge A von Ω in $\mathcal{M}_\alpha$, und es folgt $\mathcal{B}(\Omega) \subset \mathcal{M}_\alpha$. □

Beweis des Satzes über „viele $\mathcal{H}^s$-messbare Mengen“. Sind $A, \widetilde{A} \in \mathcal{P}(\Omega)$ nicht-leer mit $\mathrm{dist}(A, \widetilde{A}) > 0$, so trifft bei einer Überdeckung von $A \cup \widetilde{A}$ durch Kugeln vom Radius $< \frac{1}{2}\mathrm{dist}(A, \widetilde{A})$ jede Kugel nur eine der beiden Mengen A und $\widetilde{A}$. Daher lässt sich jede solche Überdeckung auf naheliegende Weise in gleich geartete Überdeckungen von A und $\widetilde{A}$ zerlegen. Durch Verwendung dieser Zerlegung folgt erst $\mathcal{H}^s_\delta(A \cup \widetilde{A}) \geq \mathcal{H}^s_\delta(A) + \mathcal{H}^s_\delta(\widetilde{A})$ für $0 < \delta < \frac{1}{2}\mathrm{dist}(A, \widetilde{A})$ und dann $\mathcal{H}^s(A \cup \widetilde{A}) \geq \mathcal{H}^s(A) + \mathcal{H}^s(\widetilde{A})$. Wegen der Subadditivität von $\mathcal{H}^s$ ist $\mathcal{H}^s$ ein metrisches äußeres Maß, nach dem vorausgehenden Satz gilt also $\mathcal{B}(\Omega) \subset \mathcal{M}_{\mathcal{H}^s}$.

Es bleibt $\mathcal{B}(\Omega)$-Regularität von $\mathcal{H}^s$ zu zeigen. Sei dazu $A \in \mathcal{P}(\Omega)$. Wegen $\mathcal{H}^s_\delta = (K^s_\delta)^* = ((K^s_\delta)^*|_{\mathcal{B}(\Omega)})^* = (\mathcal{H}^s_\delta|_{\mathcal{B}(\Omega)})^*$ gibt es dann zu jedem $\delta \in (0, \infty]$ ein $B_\delta \in \mathcal{B}(\Omega)$ mit $B_\delta \supset A$ und $\mathcal{H}^s_\delta(B_\delta) = \mathcal{H}^s_\delta(A)$. Für $B := \bigcap_{i=1}^\infty B_{1/i} \supset A$ erhält man $\mathcal{H}^s(B) = \mathcal{H}^s(A)$, und $\mathcal{H}^s$ ist $\mathcal{B}(\Omega)$-regulär. □

Bemerkung (zu **Bewegungsinvarianz und Skalierungsverhalten von $\boldsymbol{\mathcal{H}^s}$**). Im Fall $\Omega = \mathbb{R}^N$ liest man aus der Definition die Bewegungsinvarianz und das folgende Skalierungsverhalten von Hausdorff-Maßen ab: Für $A \in \mathcal{P}(\Omega)$, $x \in \mathbb{R}^N$, $T \in \mathcal{O}(\mathbb{R}^N)$, $r \in \mathbb{R}_{>0}$ gelten

$$\mathcal{H}^s(x+TA) = \mathcal{H}^s(A) \qquad \text{und} \qquad \mathcal{H}^s(rA) = r^s\mathcal{H}^s(A)$$

(und insbesondere ist $\mathcal{M}_{\mathcal{H}^s}$ damit bewegungs- und skalierungsinvariant). Der entscheidende Unterschied zur Definition des Lebesgue-Maßes und der Grund dafür, dass die in der Bewegungsinvarianz enthaltene Rotationsinvarianz bei Hausdorff-Maßen klar ist, bestehen darin, dass vom rotationsinvarianten Mengensystem aller Kugeln (statt einem nicht rotationsinvarianten System von achsenparallelen Quadern) ausgegangen wird.

Satz (über **Spezialfälle von $\boldsymbol{\mathcal{H}^s}$**).

(I) *Auf jedem metrischen Raum Ω ist das 0-dimensionale Maß $\mathcal{H}^0$ gleich dem Zählmaß.*

(II) *Auf $\Omega = \mathbb{R}^N$ ist das s-dimensionale Hausdorff-Maß $\mathcal{H}^s$ mit $s > N$ das Nullmaß.*

(III) ***Auf $\boldsymbol{\Omega = \mathbb{R}^N}$ ist das $\boldsymbol{N}$-dimensionale Hausdorff-Maß $\boldsymbol{\mathcal{H}^N}$ gleich dem Lebesgue-Maß $\boldsymbol{\mathscr{L}^N}$***, *genauer gilt $\mathcal{H}^N = \mathscr{L}^N$ auf Mengen der übereinstimmenden σ-Algebren $\mathcal{M}_{\mathcal{H}^N} = \mathcal{M}^N$ und $\mathcal{H}^N = \big(\mathscr{L}^N\big)^*$ auf beliebigen Mengen in $\mathcal{P}(\mathbb{R}^N)$.*

Beweis von Teil (I) *des Satzes.* Ist $A \in \mathcal{P}(\Omega)$ endlich und $d_A := \min\{d(x,y) : x \neq y \text{ in } A\}$, so folgt $\mathcal{H}^0_\delta(A) = \xi(A)$ für $0 < \delta \leq \frac{1}{2}d_A$ und damit $\mathcal{H}^0(A) = \xi(A)$. Mit Monotonie ergibt sich daraus die Behauptung $\mathcal{H}^0 = \xi$ samt $\mathcal{M}_{\mathcal{H}^0} = \mathcal{P}(\Omega)$. □

Beweis von Teil (II) *des Satzes.* Für beliebiges $\ell \in \mathbb{N}$ wird der Würfel $[0,1)^N$ durch die ℓ^N Kugeln $\mathrm{B}^N_{\sqrt{N}/\ell}(z/\ell)$ mit $z \in \{0,1,2,\ldots,\ell{-}1\}^\ell$ überdeckt. Es ergibt sich

$$\mathcal{H}^s\big([0,1)^N\big) = \lim_{\ell\to\infty} \mathcal{H}^s_{2\sqrt{N}/\ell}\big([0,1)^N\big) \leq \lim_{\ell\to\infty} \ell^N \omega_s\big(\sqrt{N}/\ell\big)^s = 0$$

für $s > N$. Wegen der Translationsinvarianz von $\mathcal{H}^s$ folgt hieraus $\mathcal{H}^s \equiv 0$ für $s > N$. □

Beweis von Teil (III) *des Satzes.* Wie in diesem Abschnitt schon gezeigt, kann $\mathcal{H}^N$ zu einem translationsinvarianten Maß auf $(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N))$ eingeschränkt werden, und mit dem soeben für Teil (II) des Satzes verwendeten Argument erhält man $\mathcal{H}^N\big([0,1)^N\big) < \infty$. Gemäß der aus Abschnitt 2.2 bekannten Eindeutigkeitsaussage für $\mathscr{L}^N$ folgt daraus schon $\mathcal{H}^N = \gamma\mathscr{L}^N$ auf Borel-Mengen in $\mathbb{R}^N$ mit einem festen $\gamma \in \mathbb{R}_{\geq 0}$.

Mit der Umformulierung $\mathcal{H}^N_\delta(A) = \inf\{\sum_{i=1}^N \mathscr{L}^N(\mathrm{B}_{r_i}(x_i)) : A \subset \bigcup_{i=1}^\infty \mathrm{B}_{r_i}(x_i)\,,\, 0 \leq r_i < \delta\}$ der Definition von $\mathcal{H}^N_\delta$ und der σ-Subadditivität von $\mathscr{L}^N$ erhält man für beliebiges $\delta > 0$ außerdem $\mathcal{H}^N \geq \mathcal{H}^N_\delta \geq \mathscr{L}^N$ auf Borel-Mengen in $\mathbb{R}^N$. Daher muss für den obigen Vorfaktor $\gamma \geq 1$ gelten.

Die Ungleichung $\gamma \leq 1$ ist etwas schwieriger nachzuweisen, denn man kommt um die Konstruktion einer „guten“ Kugel-Überdeckung (nach Wissensstand des Dozenten) nicht herum. Im Einzelnen überlegt man sich zuerst, dass es zu jeder offenen Menge O in $\mathbb{R}^N$ und jedem $\delta > 0$ eine Überdeckung $\bigcup_{i=1}^\infty W_i = O$ durch abzählbar viele disjunkte halboffene N-dimensionale Würfel $W_i \subset O$ mit Seitenlängen $\ell_i < 2\delta$ (eventuell $\ell_i = 0$) gibt. In jedes W_i kann eine Kugel $B_i \subset O$ mit Radius $r_i = \ell_i/2 < \delta$ (eventuell $r_i = 0$) einbeschrieben werden, und aus $\mathscr{L}^N(B_i) = \frac{\omega_N}{2^N}\mathscr{L}^N(W_i)$ ergibt sich insgesamt $\sum_{i=1}^\infty \mathscr{L}^N(B_i) = \frac{\omega_N}{2^N}\mathscr{L}^N(O)$. Hieraus folgt, dass es zu offenem O und $\delta > 0$ stets ein $m \in \mathbb{N}$ und endlich viele disjunkte Kugeln $B_1, B_2 \ldots, B_m \subset O$ mit Radien $< \delta$ und mit $\mathscr{L}^N(O \setminus \bigcup_{i=1}^m B_i) \leq (1{-}\frac{\omega_N}{2^{N+1}})\mathscr{L}^N(O)$ gibt. Diese Konstruktion wird nun bei fixiertem $\delta > 0$ wie folgt iterativ angewandt. Zuerst wählt man $O = (0,1)^N$ und erhält ein $m_1 \in \mathbb{N}$ und disjunkte Kugeln $K_1, K_2 \ldots, K_{m_1} \subset (0,1)^N$ mit Radien $< \delta$ und $\mathscr{L}^N((0,1)^N \setminus \bigcup_{i=1}^{m_1} K_i) \leq 1{-}\frac{\omega_N}{2^{N+1}}$. Im nächsten Schritt wählt man $O = (0,1)^N \setminus \bigcup_{i=1}^{m_1} \overline{K_i}$ und erhält ein $m_2 \in \mathbb{N}_{>m_1}$ und weitere (auch zu denen des ersten Schritts) disjunkte Kugeln $K_{m_1+1}, K_{m_1+2} \ldots, K_{m_2} \subset (0,1)^N$ mit Radien $< \delta$ und $\mathscr{L}^N((0,1)^N \setminus \bigcup_{i=1}^{m_2} K_i) \leq (1{-}\frac{\omega_N}{2^{N+1}})^2$. Mit $O = (0,1)^N \setminus \bigcup_{i=1}^{m_2} \overline{K_i}$ findet man $m_3 \in \mathbb{N}_{>m_2}$ und weitere disjunkte Kugeln $K_{m_2+1}, K_{m_2+2} \ldots, K_{m_3} \subset (0,1)^N$ mit Radien $< \delta$ und $\mathscr{L}^N((0,1)^N \setminus \bigcup_{i=1}^{m_3} K_i) \leq (1{-}\frac{\omega_N}{2^{N+1}})^3$. Durch Fortsetzung dieser Argumentation erhält man insgesamt abzählbar viele disjunkte Kugeln $K_1, K_2, K_3, \ldots \subset (0,1)^N$ mit Radien $< \delta$, mit $\mathscr{L}^N((0,1)^N \setminus \bigcup_{i=1}^\infty K_i) = 0$ und folglich mit $\sum_{i=1}^\infty \mathscr{L}^N(K_i) = 1$. Verwendet man diese Kugeln in der Definition von $\mathcal{H}^N_\delta$, so folgt $\mathcal{H}^N_\delta\big((0,1)^N\big) \leq 1$, und daraus erhält man $\gamma \leq 1$ für den oben betrachteten Vorfaktor γ.

Insgesamt ist mit der bisherigen Argumentation $\mathcal{H}^N = \mathcal{H}^N_\delta = \mathscr{L}^N$ auf Borel-Mengen in $\mathbb{R}^N$ gezeigt, und die Gleichheit überträgt sich problemlos auf Mengen in $\mathcal{M}^N$. Gemäß Bemerkung (1) zur Carathéodory-Konstruktion folgt zudem $\mathcal{H}^N_\delta = \big(K^N_\delta\big)^* = \big(\big(K^N_\delta\big)^*|_{\mathcal{B}(\mathbb{R}^N)}\big)^* = \big(\mathcal{H}^N_\delta|_{\mathcal{B}(\mathbb{R}^N)}\big)^* = \big(\mathscr{L}^N\big)^*$ auf beliebigen Mengen in $\mathcal{P}\big(\mathbb{R}^N\big)$. Folglich hängt $\mathcal{H}^N_\delta$ nicht von δ ab, und auch $\mathcal{H}^N$ stimmt auf ganz $\mathcal{P}\big(\mathbb{R}^N\big)$ mit $\big(\mathscr{L}^N\big)^*$ überein. Gemäß dem Satz über μ^*- und μ-Messbarkeit aus Abschnitt 2.4 folgt insbesondere $\mathcal{M}_{\mathcal{H}^N} = \mathcal{M}_{(\mathscr{L}^N)^*} = \mathcal{M}^N$. □

Die Gleichheit $\mathscr{L}^N = \mathcal{H}^N$ in Teil (III) des Satzes zeigt, dass eine Konstruktion des Lebesgue-Maßes auch über Kugeln statt über Quader erfolgen kann. Beide Herangehensweisen haben sowohl Vor- als auch Nachteile: Ein wesentlicher Vorteil der allgemein üblichen Quader-Definition liegt darin, dass die halboffenen Quader einen Halbring bilden und der Maßfortsetzungssatz in der früher besprochenen Weise greift. Ist die Verbindung mit Kugeln, entweder per Definition oder per Satz, einmal hergestellt, so besteht ein wichtiger Vorteil darin, dass Rotationsinvarianz direkt und problemlos zur Verfügung steht:

Folgerungen (aus der Übereinstimmung $\mathcal{H}^N = \mathscr{L}^N$ auf $\mathbb{R}^N$)**.**

(1) **Die $\boldsymbol{\sigma}$-Algebra $\mathcal{M}^N$ der Lebesgue-messbaren Mengen und das Lebesgue-Maß $\mathscr{L}^N$** (auf $\mathcal{B}(\mathbb{R}^N)$ oder auf $\mathcal{M}^N$ oder auch als äußeres Lebesgue-Maß auf $\mathcal{P}(\mathbb{R}^N)$) **sind bewegungsinvariant.**

(2) Die **Einschränkung von $\boldsymbol{\mathcal{H}^k}$ auf einen $\boldsymbol{k}$-dimensionalen Unterraum** U von $\mathbb{R}^N$ mit $k \in \{1, 2, \dots, N\}$ **entspricht dem Lebesgue-Maß** $\mathscr{L}^k$: Schreibt man genauer $U = T(\mathbb{R}^k)$ als Bild einer linearen Isometrie $T\colon \mathbb{R}^k \to \mathbb{R}^N$, so gilt $\mathcal{H}^k(TA) = \mathscr{L}^k(A)$ für alle $A \in \mathcal{M}^k$ (und für alle $A \in \mathcal{P}(\mathbb{R}^k)$ außerdem $A \in \mathcal{M}^k \iff TA \in \mathcal{M}_{\mathcal{H}^k}$ sowie $\mathcal{H}^k(TA) = (\mathscr{L}^k)^*(A)$).

Beweis. Die Folgerung (1) ergibt sich aufgrund der Gleichheit $\mathscr{L}^N = \mathcal{H}^N$ aus der Bewegungsinvarianz von $\mathcal{H}^N$ und $\mathcal{M}_{\mathcal{H}^N}$ (Und für diesen Schluss reicht übrigens die Erkenntnis $\gamma\,\mathscr{L}^N = \mathcal{H}^N$ mit $\gamma \in \mathbb{R}_{>0}$, er erfordert also nur die ersten einfachen Argumente des für Teil (III) des vorigen Satzes gegebenen Beweises).

Um die Folgerung (2) zu erhalten, argumentiert man wie folgt: Für $x \in \mathbb{R}^N$ bezeichne x_* den zu x nächstgelegenen Punkt in U. Aus der Inklusion $U \cap \mathrm{B}_r(x) \subset \mathrm{B}_r(x_*)$ sieht man dann, dass man sich beim Ausschreiben der Definition von $\mathcal{H}^k(TA)$ mit beliebigem $A \in \mathcal{P}(\mathbb{R}^k)$ auf Kugeln in $\mathbb{R}^N$ mit Mittelpunkten in U beschränken kann. Man erhält Äquatorebenen solcher Kugeln als T-Bilder von Kugeln mit gleichem Radius in $\mathbb{R}^k$, deshalb treten in der Definition von $\mathcal{H}^k(A)$ die gleichen Näherungssummen auf, und es gilt $\mathcal{H}^k(TA) = \mathcal{H}^k(A)$. Wegen $\mathcal{H}^k = \mathscr{L}^k$ beziehungsweise $\mathcal{H}^k = (\mathscr{L}^k)^*$ auf $\mathbb{R}^k$ folgen hieraus die Behauptungen. □

Proposition ($\boldsymbol{\mathcal{H}^s}$-Schranke bei Lipschitz-Bildern). *Seien Ω und $\mathcal{X}$ metrische Räume, und sei $s \in \mathbb{R}_{\ge 0}$. Für das $\mathcal{H}^s$-Maß des Bildes von $A \in \mathcal{P}(\Omega)$ unter einer Lipschitz-stetigen Abbildung $f\colon \Omega \to \mathcal{X}$ mit Lipschitz-Konstante $\operatorname{Lip} f \in \mathbb{R}_{\ge 0}$ gilt die Abschätzung*

$$\mathcal{H}^s(f(A)) \le (\operatorname{Lip} f)^s \mathcal{H}^s(A)\,.$$

Beweis. Aus einer Überdeckung $A \subset \bigcup_{i=1}^\infty \mathrm{B}_{r_i}(x_i)$ von A durch Kugeln in Ω mit Radien r_i gewinnt man eine Überdeckung $f(A) \subset \bigcup_{i=1}^\infty \mathrm{B}_{(\operatorname{Lip} f) r_i}(f(x_i))$ von $f(A)$ durch Kugeln in $\mathcal{X}$ mit neuen Radien $(\operatorname{Lip} f) r_i$. Aus diesem Zusammenhang ergibt sich gemäß Definition von $\mathcal{H}^s_\delta$ und $\mathcal{H}^s$ erst $\mathcal{H}^s_{(\operatorname{Lip} f)\delta}(f(A)) \le (\operatorname{Lip} f)^s \mathcal{H}^s_\delta(A)$ für $\delta > 0$ und dann $\mathcal{H}^s(f(A)) \le (\operatorname{Lip} f)^s \mathcal{H}^s(A)$. □

Mit Hilfe von Hausdorff-Maßen lässt sich auch ein sehr allgemeiner Dimensionsbegriff bilden. Dieser wird in der Vorlesung nicht vertieft untersucht, zumindest die Definition sei aber kurz festgehalten:

Bemerkungen (zur **Hausdorff-Dimension**)**.**

(1) Sei Ω ein metrischer Raum. Aus der Definition der Hausdorff-Maße lässt sich ableiten, dass es zu jedem $A \in \mathcal{P}(\Omega)$ eine eindeutig bestimmte Zahl $d \in [0, \infty]$ mit

$$\mathcal{H}^s(A) = \infty \quad \text{für alle } s \in [0, d) \qquad\qquad \text{und} \qquad\qquad \mathcal{H}^s(A) = 0 \quad \text{für alle } s \in (d, \infty)$$

gibt. Man nennt d die **Hausdorff-Dimension** von A.

(2) Interessanterweise muss die Hausdorff-Dimension keineswegs ganzzahlig sein. Ein berühmtes Beispiel einer Menge gebrochener Hausdorff-Dimension ist die Cantor-Menge $C \subset [0,1]$, die bekanntlich durch iteratives Entfernen mittlerer Drittel entsteht und Hausdorff-Dimension $\frac{\log 2}{\log 3} \in (0,1)$ hat. Dass $\frac{\log 2}{\log 3}$ zumindest ein Kandidat für die Dimension von C ist, kann man sich anhand der disjunkten Zerlegung $C = \frac{1}{3}C \cup (\frac{2}{3}+\frac{1}{3}C)$ von C in zwei selbstähnliche Teile plausibel machen: Mit Additivität, Translationsinvarianz und Skalierungsverhalten von $\mathcal{H}^s$ folgt aus dieser Zerlegungseigenschaft nämlich $\mathcal{H}^s(C) = 2\big(\frac{1}{3}\big)^s\mathcal{H}^s(C)$. Wenn es also überhaupt ein $s \in \mathbb{R}_{\geq 0}$ mit $0 < \mathcal{H}^s(C) < \infty$ gibt, so muss dieses $2\big(\frac{1}{3}\big)^s = 1$ und folglich $s = \frac{\log 2}{\log 3}$ erfüllen. Um für dieses s das Eintreten von $\mathcal{H}^s(C) = 0$ oder $\mathcal{H}^s(C) = \infty$ auszuschließen und somit vollständig zu beweisen, dass s die Hausdorff-Dimension von C ist, braucht man allerdings noch etwas feinere Abschätzungen für $\mathcal{H}^s_\delta$-Maße. Genaueres findet man etwa in [4, Example 2.7] (allerdings für eine leicht andere Definition von $\mathcal{H}^s$ als in diesen Notizen).

(3) Darüber hinaus hat F. Hausdorff bereits 1919 für jedes $N \in \mathbb{N}$ und *jedes* $d \in [0,N]$ Beispielmengen in $\mathcal{P}\big(\mathbb{R}^N\big)$ mit Hausdorff-Dimension d angegeben.

Als Nächstes wird die zweite wichtige Thematik dieses Abschnitts, nämlich die Transformation von Maßen und Integralen, angegangen. Zuerst wird dabei ein abstraktes Konzept geprägt, danach geht es aber schwerpunktmäßig um konkretere Transformationen von Lebesgue-Maßen und -Integralen sowie allgemeineren Hausdorff-Maßen und -Integralen.

Definition (**Bildmaße**)**.** *Seien ein Maßraum $(\Omega, \mathscr{A}, \mu)$, ein Messraum $(\mathcal{X}, \mathscr{S})$ und eine $(\mathscr{A}, \mathscr{S})$-messbare Abbildung $F\colon \Omega \to \mathcal{X}$ gegeben. Dann wird durch*

$$F_\sharp\mu(S) = \mu^F(S) := \mu(F^{-1}(S)) \qquad \text{für } S \in \mathscr{S}$$

*ein Maß $F_\sharp\mu = \mu^F$ auf $(\mathcal{X}, \mathscr{S})$ definiert. Man nennt dieses Maß das **Bildmaß** von μ unter F.*

Man kann μ- und μ^F-Integrale dann wie folgt ineinander transformieren:

Proposition (**abstrakte Transformationsformel** für Maßintegrale)**.** *Seien ein Maßraum $(\Omega, \mathscr{A}, \mu)$, ein Messraum $(\mathcal{X}, \mathscr{S})$ und eine $(\mathscr{A}, \mathscr{S})$-messbare Abbildung $F\colon \Omega \to \mathcal{X}$ gegeben. Für μ^F-messbares $f\colon \mathcal{X} \to \overline{\mathbb{R}}$ ist $f \circ F$ stets μ-messbar mit*

$$\int_S f \,\mathrm{d}\mu^F = \int_{F^{-1}(S)} (f \circ F)\,\mathrm{d}\mu \qquad \text{für alle } S \in \mathscr{S}\,,$$

wobei beide Integrale existieren oder keines.

Beweis. Für $\widetilde{S} \in \mathscr{S}$ erhält man aus den Definitionen

$$\begin{aligned}\int_S \mathbb{1}_{\widetilde{S}}\,\mathrm{d}\mu^F &= \mu^F(S \cap \widetilde{S}) = \mu(F^{-1}(S \cap \widetilde{S}))\\ &= \mu(F^{-1}(S) \cap F^{-1}(\widetilde{S})) = \int_{F^{-1}(S)} \mathbb{1}_{F^{-1}(\widetilde{S})}\,\mathrm{d}\mu = \int_{F^{-1}(S)} (\mathbb{1}_{\widetilde{S}} \circ F)\,\mathrm{d}\mu\end{aligned}$$

samt $\mathscr{A}$-Messbarkeit von $\mathbb{1}_{\widetilde{S}} \circ F$, also gilt die Behauptung für $f = \mathbb{1}_{\widetilde{S}}$. Der allgemeine Fall lässt sich dann weitgehend problemlos mit der Standard-Ausdehnungsprozedur erledigen. $\square$

Anwendung (der abstrakten Transformationsformel in der Wahrscheinlichkeitstheorie)**.** In der Wahrscheinlichkeitstheorie modelliert man beobachtbare Größen, die von einem Zufallselement abhängen, durch **Zufallsvariablen** X, d.h. durch $\mathscr{A}$-messbare Abbildungen $X\colon \Omega \to \mathbb{R}$ auf einem Wahrscheinlichkeitsraum $(\Omega, \mathscr{A}, P)$. Man bezeichnet das Bildmaß P^X auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ hierbei als (Wahrscheinlichkeits-)**Verteilung von $\boldsymbol{X}$**, denn die Wahrscheinlichkeit, dass man für die betrachtete Größe den Wert $x \in \mathbb{R}$ beziehungsweise einen Wert in $[a,b) \subset \mathbb{R}$ beobachtet, ist gerade $P^X(\{x\}) = P(\{\omega \in \Omega : X(\omega) = x\})$ beziehungsweise $P^X([a,b)) = P(\{\omega \in \Omega : a \le X(\omega) < b\})$.

Für den (wie folgt definierten) Erwartungswert $\mathrm{E}_P[X]$ einer Zufallsvariable $X \in \mathcal{L}^1(\Omega, \mathscr{A}, P)$ erlaubt die vorausgehende Proposition dann die Umschreibung

$$\mathrm{E}_P[X] := \int_{\mathbb{R}} x \,\mathrm{d}P^X(x) = \int_\Omega X(\omega)\,\mathrm{d}P(\omega)\,,$$

wobei beide Integralausdrücke Vorteile mit sich bringen: Am P-Integral erkennt man Linearität von $\mathrm{E}_P[X]$ in X, für die praktische Berechnung von $\mathrm{E}_P[X]$ eignet sich dagegen das P^X-Integral, denn normalerweise ist überhaupt nur die Verteilung P^X und nicht das Hintergrund-Wahrscheinlichkeitsmaß P konkret bekannt.

Im Folgenden geht es aber vor allem um die Transformation von Lebesgue- und Hausdorff-Maßen mit Hilfe gutartiger Transformationen F. Die hierzu gehörigen Integraltransformationen übernehmen bei Integrationen in mehreren Variablen die Rolle der Substitutionsregel und sind für praktische Maß- und Integralberechnungen extrem nützlich:

Hauptsatz (Jacobische **Transformationsformel für $\boldsymbol{\mathscr{L}^N}$**, Substitutionsregel in N Variablen)**.** *Seien D eine offene Teilmenge von $\mathbb{R}^N$ und $F\colon D \to \mathbb{R}^N$ ein C^1-Diffeomorphismus von D auf die offene Menge $F(D)$. Sei außerdem die **Jacobi-Determinante** oder **Jacobische** $\mathrm{J}F\colon D \to \mathbb{R}_{\ge 0}$ durch*

$$\mathrm{J}F(x) := |\det(\mathrm{D}F(x))| \qquad \textit{für } x \in D$$

mit der Jacobi-Matrix $\mathrm{D}F(x) \in \mathbb{R}^{N\times N}$ definiert, und sei A eine $\mathscr{L}^N$-messbare Teilmenge von D. Dann gelten:

(I) *Auch das Bild $F(A)$ ist $\mathscr{L}^N$-messbar mit*

$$\boxed{\mathscr{L}^N(F(A)) = \int_A \mathrm{J}F \,\mathrm{d}\mathscr{L}^N}\,.$$

(II) *Für $\mathscr{L}^N$-messbares $f\colon F(A) \to \overline{\mathbb{R}}$ ist $f \circ F\colon A \to \overline{\mathbb{R}}$ ebenfalls $\mathscr{L}^N$-messbar mit*

$$\boxed{\int_{F(A)} f \,\mathrm{d}\mathscr{L}^N = \int_A (f\circ F)\,\mathrm{J}F\,\mathrm{d}\mathscr{L}^N}\,,$$

wobei beide Integrale existieren oder keines.

Hauptsatz (**Flächenformel**, Transformationsformel für $\mathcal{H}^k$)**.** *Sei $k \le N$ in $\mathbb{N}$, seien D eine offene Teilmenge von $\mathbb{R}^k$ und $F\colon D \to \mathbb{R}^N$ eine injektive C^1-Abbildung auf D. Sei die (allgemeinere) **Jacobi-Determinante** oder **Jacobische** durch*

$$\mathrm{J}F(x) := \sqrt{\det\left(\mathrm{D}F(x)^{\mathrm{T}}\mathrm{D}F(x)\right)} \qquad \textit{für } x \in D$$

mit der Funktionalmatrix $\mathrm{D}F(x) \in \mathbb{R}^{N\times k}$ und der positiv semidefiniten, symmetrischen Matrix $\mathrm{D}F(x)^{\mathrm{T}}\mathrm{D}F(x) \in \mathbb{R}^{k\times k}$ definiert, und sei A eine $\mathscr{L}^k$-messbare Teilmenge von D. Dann gelten:

(I) *Das Bild $F(A)$ ist $\mathcal{H}^k$-messbar mit*

$$\boxed{\mathcal{H}^k(F(A)) = \int_A \mathrm{J}F \,\mathrm{d}\mathscr{L}^k\,.}$$

(II) *Für $\mathcal{H}^k$-messbares $f\colon F(A) \to \overline{\mathbb{R}}$ ist $(f\circ F)\mathrm{J}F\colon A \to \overline{\mathbb{R}}$ eine $\mathscr{L}^k$-messbare Abbildung mit*

$$\boxed{\int_{F(A)} f\,\mathrm{d}\mathcal{H}^k = \int_A (f\circ F)\,\mathrm{J}F\,\mathrm{d}\mathscr{L}^k\,,}$$

wobei beide Integrale existieren oder keines.

Bemerkungen (zur Transformations- und Flächenformel)**.**

(1) Die Transformationsformel und die Flächenformel sind **zentrale Hilfsmittel zur Berechnung von Maßen und Integralen** bezüglich $\mathscr{L}^N$ und $\mathcal{H}^k$. **Beispiele und Anwendungen folgen** nach dem Beweis.

(2) In der Situation der Transformationsformel für $\mathscr{L}^N$ bilden die C^1-Diffeomorphismen F und F^{-1} Borel-Mengen (sowieso) auf Borel-Mengen und Lebesgue-messbare Mengen gemäß der Aussage (I) auf Lebesgue-messbare Mengen ab. Daher kann man die Aussage (I) der $\mathscr{L}^N$-Transformationsformel auch als $(F^{-1})_\sharp\mathscr{L}^N = \mathrm{J}F{\cdot}\mathscr{L}^N$ (mit der $(\mathcal{M}^k|F(D),\mathcal{M}^k|D)$-messbaren Abbildung $F^{-1}\colon F(D) \to D$) oder äquivalent als $F_\sharp(\mathrm{J}F{\cdot}\mathscr{L}^N) = \mathscr{L}^N$ (mit der $(\mathcal{M}^k|D,\mathcal{M}^k|F(D))$-messbaren Abbildung $F\colon D \to F(D)$) ausdrücken. Die Aussage (II) ist nichts anderes als die zur letzten Formulierung gehörige Integraltransformation aus der vorausgehenden Proposition.

(3) In der Situation der Flächenformel ist $F(D)$ eine Borel-Menge[38]. Analog zur vorigen Bemerkung bilden F und F^{-1} dann Borel-Mengen auf Borel-Mengen (trivial für F^{-1}; für F siehe Beweis) und in geeigneter[39] Weise messbare Mengen auf messbare Mengen ab, so dass man die Aussage (I) als $(F^{-1})_\sharp\mathcal{H}^k = \mathrm{J}F{\cdot}\mathscr{L}^k$ oder $F_\sharp(\mathrm{J}F{\cdot}\mathscr{L}^k) = \mathcal{H}^k$ formulieren kann. Bei (II) handelt es sich wieder um die zugehörige Integraltransformation.

(4) Die Flächenformel enthält die Jacobische Transformationsformel als Spezialfall, ist aber selbst für $k = N$ etwas allgemeiner, da F^{-1} in der Allgemeinheit der Flächenformel (in Nullstellen von $\mathrm{J}F$) nicht differenzierbar sein muss.

Der Beweis der Flächenformel basiert auf dem folgenden Lemma, das dem Spezialfall linearer Transformationen F entspricht. Der allgemeine Fall wird dann unten durch lokales Linearisieren von F darauf zurückgeführt.

[38] $F(D)$ ist sogar F_σ, denn D und damit auch $F(D)$ sind abzählbare Vereinigungen von Kompakta.

[39] Tatsächlich ergibt sich aus der angegebenen Aussage zur Borel-Borel-Messbarkeit und Teil (I) der Flächenformel, dass eine Teilmenge $S = F(A) \subset F(D)$ genau dann $\mathcal{H}^k$-messbar ist, wenn die zugehörige Teilmenge $A = F^{-1}(S) \subset D$ für das gewichtete Maß $\mathrm{J}F{\cdot}\mathscr{L}^k$ messbar ist. Damit ist F eine $(\mathcal{M}_{\mathrm{J}F\cdot\mathscr{L}^k}|D,\mathcal{M}_{\mathcal{H}^k}|F(D))$-messbare und F^{-1} eine $(\mathcal{M}_{\mathcal{H}^k}|F(D),\mathcal{M}_{\mathrm{J}F\cdot\mathscr{L}^k}|D)$-messbare Abbildung, so dass die angegebenen Bildmaße Sinn machen. Hat $\mathrm{J}F$ keine Nullstellen, so kann $\mathrm{J}F{\cdot}\mathscr{L}^k$ bei diesen Messbarkeitsaussagen auch durch $\mathscr{L}^k$ ersetzt werden, bei Nullstellen von $\mathrm{J}F$ ist dies nicht mehr möglich, denn dann kann eine $\mathcal{H}^k$-Nullmenge $S = F(A)$ tatsächlich ein nicht-Lebesgue-messbares F-Urbild $A = F^{-1}(S)$ besitzen. Dies ist auch der Grund, warum bei der Aussage (II) im Allgemeinen nur $(f\circ F)\,\mathrm{J}F$ und nicht $(f\circ F)$ selbst $\mathscr{L}^k$-messbar ist. Ein (nicht ganz einfaches) Beispiel für diese Problematik schon im Fall $k = N = 1$ erhält man durch die Konstruktionen einer Cantor-Menge $C \subset [0,1]$ mit $\mathscr{L}^1(C) > 0$, einer streng monotonen C^1-Funktion $F\colon [0,1] \to \mathbb{R}$ mit $F' \equiv 0$ auf C und einer nicht-$\mathscr{L}^1$-messbaren Teilmenge A von C.

Lemma (lineare Transformationsregel). *Sei $k \leq N$ in $\mathbb{N}$, und sei $L \in \mathbb{R}^{N\times k}$. Dann ist für $A \in \mathcal{M}^k$ stets $LA \in \mathcal{M}_{\mathcal{H}^k}$ mit*

$$\mathcal{H}^k(LA) = \sqrt{\det(L^{\mathrm{T}}L)}\,\mathscr{L}^k(A)\,.$$

Beweis des Lemmas. Gemäß der Singulärwertzerlegung[40] lässt sich $L = TDS$ mit einer Diagonalmatrix $D = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_k) \in \mathbb{R}^{k\times k}$, einer orthogonalen Matrix $S \in \mathbb{R}^{k\times k}$, $S^{\mathrm{T}}S = \mathbb{I}_k$ und einer orthogonal ergänzbaren Matrix $T \in \mathbb{R}^{N\times k}$, $T^{\mathrm{T}}T = \mathbb{I}_k$ (d.h. mit anderen Worten einer Isometrie $T \in \mathcal{L}(\mathbb{R}^k, \mathbb{R}^N)$) schreiben. Es gilt dann $\det(L^{\mathrm{T}}L) = \prod_{i=1}^k \lambda_i^2$, und die Behauptung ergibt sich aus der (unten noch erläuterten) Rechnung

$$\begin{aligned}\mathcal{H}^k(LA) &= \mathcal{H}^k(TDSA) = \mathscr{L}^k(DSA)\\ &= \left(\prod_{i=1}^k |\lambda_i|\right)\mathscr{L}^k(SA) = \left(\prod_{i=1}^k |\lambda_i|\right)\mathscr{L}^k(A) = \sqrt{\det(L^{\mathrm{T}}L)}\mathscr{L}^k(A)\,.\end{aligned}$$

Maßgeblich ging hierbei vor allem die frühere Bemerkung (2) zum Lebesgue-artigen Verhalten von $\mathcal{H}^k$ auf dem Unterraum $T(\mathbb{R}^k)$ ein, und beim Übergang zur zweiten Zeile wurde die naheliegende Skalierungsregel für $\mathscr{L}^k$ bei Multiplikation jeder Komponente x_i mit einem anderen Faktor λ_i verwendet (wie man sie problemlos aus Fubini oder anhand der ursprünglichen Konstruktion von $\mathscr{L}^N$ als Fortsetzung von λ^N erhält). Schließlich wurde noch die Rotationsinvarianz von $\mathscr{L}^k$ ausgenutzt. □

Er folgt zunächst ein Beweis der Flächenformel unter einer schwachen Zusatzannahme. Der dadurch erledigte Fall enthält die Transformationsformel für $\mathscr{L}^N$ in ihrer Allgemeinheit. Für den allgemeinen Fall der Flächenformel benötige Ergänzungen werden im Anschluss nachgetragen.

Beweis der Flächenformel für den Fall, dass $\mathrm{J}F$ in D keine Nullstellen besitzt. Aus der Definition $\mathrm{J}F = \sqrt{\det\big((\mathrm{D}F)^{\mathrm{T}}\mathrm{D}F\big)}$ folgt, dass $\mathrm{D}F$ im hier behandelten Fall auf ganz D vollen Rang k hat. Für jedes $x \in D$ ist dann $\mathrm{D}F(x)$ eine injektive lineare Abbildung $\mathbb{R}^k \to \mathbb{R}^N$, und es gibt eine Konstante $m_x \in \mathbb{R}_{>0}$ mit $|\mathrm{D}F(x)v| \geq m_x|v|$ für alle $v \in \mathbb{R}^k$. Es lässt sich nun ohne Einschränkung annehmen, dass D beschränkt und $\mathrm{D}F$ auf D gleichmäßig stetig ist, und dass $m|v| \leq |\mathrm{D}F(x)v| \leq M|v|$ für alle $x \in D$ und $v \in \mathbb{R}^k$ mit festen Konstanten $m, M \in \mathbb{R}_{>0}$ gilt. Sind diese Gleichmäßigkeitsannahmen nämlich nicht erfüllt, so kann man für den Hauptteil der Argumentation von D und A zu den beschränkten Teilmengen $D_\delta^* := \{x \in D : \mathrm{dist}(x, \mathbb{R}^N \setminus D) > \delta\,,\, |x| < \delta^{-1}\} \subset D$ und $A_\delta^* := A \cap D_\delta^*$ mit $\delta \in \mathbb{R}_{>0}$ übergehen, auf D_δ^* alle benötigte Gleichmäßigkeit aus der Kompaktheit von $\overline{D_\delta^*} \subset D$ und Überdeckungsargumenten erhalten und schließlich durch Grenzübergang $\delta \searrow 0$ zu D und A zurückkehren. Seien also die angegebenen Gleichmäßigkeitsannahmen erfüllt, und sei $\varepsilon \in \big(0, \frac{m}{2}\big]$ beliebig fixiert. Dann lässt sich $D = \bigcup_{i=1}^\infty D_i$ in abzählbar viele disjunkte konvexe Mengen $D_i \in \mathcal{B}(\mathbb{R}^k)$

[40]Die hier verwendete Variante der Singulärwertzerlegung kann man wie folgt aus der bekannteren Hauptachsentransformation ableiten: Da $L^{\mathrm{T}}L \in \mathbb{R}^{k\times k}$ symmetrisch und positiv semidefinit ist, liefert die Hauptachsentransformation eine orthogonale Matrix $S \in \mathbb{R}^{k\times k}$, $S^{\mathrm{T}}S = \mathbb{I}_k$, die S gemäß $SL^{\mathrm{T}}LS^{\mathrm{T}} = \mathrm{diag}(\lambda_1^2, \lambda_2^2, \ldots, \lambda_k^2)$ in eine Diagonalmatrix transformiert. Die Diagonaleinträge konnten dabei in der Form λ_i^2 mit $\lambda_i \in \mathbb{R}$ geschrieben werden, da es sich bei Ihnen um nichtnegative Eigenwerte von $L^{\mathrm{T}}L$ handelt. Sind alle λ_i ungleich Null, so kann man jetzt $T := LS^{\mathrm{T}}\mathrm{diag}(\lambda_1^{-1}, \lambda_2^{-1}, \ldots, \lambda_k^{-1})$ wählen und $L = T\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_k)S$ sowie $T^{\mathrm{T}}T = \mathbb{I}_k$ problemlos verifizieren. Verschwinden gewisse λ_i, so trägt man in die zugehörigen, durch das Vorausgehende nicht definierten Spalten von T beliebige Vektoren ein, die mit den anderen Spalten ein Orthonormalsystem bilden. Damit lassen sich die Behauptungen auch in diesem Fall problemlos verifizieren.

(z.B. in geeignete halboffene Quader) mit[41] $F(D_i) \in \mathcal{B}(\mathbb{R}^N)$ und zugehörigen $x_i \in D$ zerlegen. Zudem können aufgrund der angenommenen gleichmäßigen Stetigkeit die Durchmesser der D_i klein genug gewählt werden, dass

$$|(\mathrm{D}F-\mathrm{D}F(x_i))v| \le \varepsilon|v| \qquad \text{und} \qquad |\mathrm{J}F-\mathrm{J}F(x_i)| \le \varepsilon \qquad \text{auf } D_i \tag{$*$}$$

für alle $v \in \mathbb{R}^k$, $i \in \mathbb{N}$ gelten. Für beliebige $x, \widetilde{x} \in D_i$ erhält man durch Integration über die (in D_i enthaltene) Strecke von x nach $\widetilde{x}$ dann die Hilfsabschätzung

$$\begin{aligned}
|F(\widetilde{x})-F(x)| &= \left|\int_0^1 \mathrm{D}F(x+t(\widetilde{x}-x))(\widetilde{x}-x)\,\mathrm{d}t\right| \\
&\ge |\mathrm{D}F(x_i)(\widetilde{x}-x)| - \int_0^1 |(\mathrm{D}F(x+t(\widetilde{x}-x))-\mathrm{D}F(x_i))(\widetilde{x}-x)|\,\mathrm{d}t \\
&\ge m|\widetilde{x}-x| - \varepsilon|\widetilde{x}-x| \ge \frac{m}{2}|\widetilde{x}-x|\,,
\end{aligned}$$

und für $y, \widetilde{y} \in F(D_i)$ folgt (mit der auf $F(D)$ erklärten Umkehrabbildung F^{-1})

$$|F^{-1}(\widetilde{y})-F^{-1}(y)| \le \frac{2}{m}|\widetilde{y}-y|\,. \tag{$**$}$$

Für den Hauptteil der Argumentation sei A eine Borelsche Teilmenge von D, und es sei $A_i := A \cap D_i$. Zunächst garantiert dann $(**)$ Stetigkeit von F^{-1} auf $F(D_i) \in \mathcal{B}(\mathbb{R}^N)$, so dass mit den A_i sowohl die $F(A_i)$ als auch das ganze Bild $F(A) = \bigcup_{i=1}^\infty A_i$ Borel-Mengen und insbesondere $\mathcal{H}^k$-messbar sind. Als weitere Abkürzungen werden jetzt $L_i := \mathrm{D}F(x_i) \in \mathbb{R}^{N\times k}$ und $\mathrm{J}L_i := \sqrt{\det\left(L_i^{\mathrm{T}}L_i\right)} = \mathrm{J}F(x_i) \in \mathbb{R}_{>0}$ verwendet. Mit der Abschätzung $(*)$ ergibt sich dann

$$\left|\mathrm{J}L_i\,\mathscr{L}^k(A_i) - \int_{A_i} \mathrm{J}F\,\mathrm{d}\mathscr{L}^k\right| \le \varepsilon\mathscr{L}^k(A_i)\,. \tag{$*{*}*$}$$

Als Nächstes bemerkt man, dass L_i^{-1} auf $L_i\mathbb{R}^k$ erklärt ist (denn $L_i = \mathrm{D}F(x_i)$ ist wegen $m|v| \le |L_iv|$ für $v \in \mathbb{R}^k$ injektiv), und man weist für $F\circ L_i^{-1}$ auf L_iD_i Lipschitz-Stetigkeit mit Lipschitz-Konstante $1+\frac{\varepsilon}{m}$ nach. Tatsächlich ergibt sich die behauptete Lipschitz-Stetigkeit durch Umschreiben $F\circ L_i^{-1} = \mathrm{id}+(F-L_i)\circ L_i^{-1}$, denn gemäß dem Schrankensatz und $(*)$ ist $F-L_i$ auf D_i Lipschitz-stetig mit Lipschitz-Konstante ε und als Konsequenz aus $m|v| \le |L_iv|$ ist L_i^{-1} auf $L_i\mathbb{R}^k$ Lipschitz-stetig mit Lipschitz-Konstante $\frac{1}{m}$. Mit der so begründeten Lipschitz-Stetigkeit von $F\circ L_i^{-1}$ und der $\mathcal{H}^k$-Schranke für Lipschitz-Bilder folgt

$$\mathcal{H}^k(F(A_i)) = \mathcal{H}^k((F\circ L_i^{-1})L_iA_i) \le \left(1+\frac{\varepsilon}{m}\right)^k \mathcal{H}^k(L_iA_i)\,.$$

Eine ähnliche Argumentation zeigt Lipschitz-Stetigkeit von $L_i\circ F^{-1}$ auf $F(D_i)$ mit Lipschitz-Konstante $1+\frac{2\varepsilon}{m}$. Tatsächlich schreibt man hierzu $L_i\circ F^{-1} = \mathrm{id}+(L_i-F)\circ F^{-1}$ und verwendet auch die durch $(**)$ grantierte Lipschitz-Stetigkeit von F^{-1} mit Konstante $\frac{2}{m}$ (woher dieses Mal der Faktor 2 rührt). Aus der angegebenen Lipschitz-Stetigkeit von $L_i\circ F^{-1}$ ergibt sich dann ganz analog

$$\mathcal{H}^k(L_iA_i) \le \left(1+\frac{2\varepsilon}{m}\right)^k \mathcal{H}^k(F(A_i))\,.$$

[41] Die Bedingung $F(D_i) \in \mathcal{B}(\mathbb{R}^N)$ folgt aus $D_i \in \mathcal{B}(\mathbb{R}^k)$, wenn die D_i als abzählbare Vereinigungen kompakter Mengen geschrieben werden können, wie es bei den üblichen halboffenen Quadern ja der Fall ist.

Die so erhaltenen Abschätzungen der $\mathcal{H}^k$-Maße von $F(A_i)$ und L_iA_i gegeneinander lassen sich in eine Abschätzung für die Differenz dieser Maße umschreiben. Man erhält

$$\begin{aligned} |\mathcal{H}^k(F(A_i)) - \mathcal{H}^k(L_iA_i)| &\le \left[\Big(1+\frac{2\varepsilon}{m}\Big)^k - 1\right] \max\{\mathcal{H}^k(F(A_i)), \mathcal{H}^k(L_iA_i)\} \\ &\le \left[\Big(1+\frac{2\varepsilon}{m}\Big)^k - 1\right] M^k \mathscr{L}^k(A_i)\,, \end{aligned} \tag{$****$}$$

wobei im zweiten Schritt zusätzlich verwendet wurde, dass sowohl F als auch L_i Lipschitz-stetig mit Lipschitz-Konstante M sind, dass die $\mathcal{H}^k$-Schranke entsprechend greift, und dass $\mathcal{H}^k$ auf (Borel-Mengen in) $\mathbb{R}^k$ mit $\mathscr{L}^k$ übereinstimmt. Mit Hilfe der Abschätzungen ($***$) und ($****$) sowie der linearen Transformationsregel $\mathcal{H}^k(L_iA_i) = \mathrm{J}L_i\,\mathscr{L}^k(A_i)$ aus dem vorausgehenden Lemma erhält man insgesamt

$$\begin{aligned} &\left|\mathcal{H}^k(F(A_i)) - \int_{A_i} \mathrm{J}F\,\mathrm{d}\mathscr{L}^k\right| \\ &\quad\le \left|\mathcal{H}^k(F(A_i)) - \mathcal{H}^k(L_iA_i)\right| + \left|\mathcal{H}^k(L_iA_i) - \mathrm{J}L_i\,\mathscr{L}^k(A_i)\right| + \left|\mathrm{J}L_i\,\mathscr{L}^k(A_i) - \int_{A_i} \mathrm{J}F\,\mathrm{d}\mathscr{L}^k\right| \\ &\quad\le \varepsilon\mathscr{L}^k(A_i) + \left[\Big(1+\frac{2\varepsilon}{m}\Big)^k - 1\right] M^k \mathscr{L}^k(A_i)\,. \end{aligned}$$

Diese Abschätzung kann über alle $i \in \mathbb{N}$ aufsummiert werden und nimmt dann die Form

$$\left|\mathcal{H}^k(F(A)) - \int_A \mathrm{J}F\,\mathrm{d}\mathscr{L}^k\right| \le \varepsilon\mathscr{L}^k(A) + \left[\Big(1+\frac{2\varepsilon}{m}\Big)^k - 1\right] M^k \mathscr{L}^k(A)$$

an. Wegen der Beschränktheitsannahme an D ist insbesondere $\mathscr{L}^k(A) < \infty$, daher erhält man durch Grenzübergang $\varepsilon \searrow 0$ die Behauptung (I) des Satzes

$$\mathcal{H}^k(F(A)) = \int_A \mathrm{J}F\,\mathrm{d}\mathscr{L}^k$$

für Borel-Teilmengen A von D.

Ist A nur eine $\mathscr{L}^k$-messbare Teilmenge von D, so gibt es per Definition der Vervollständigung Borel-Mengen $\widetilde{A}$ und M mit $\mathscr{L}^k(M) = 0$ und $\widetilde{A} \subset A \subset \widetilde{A} \cup M \subset D$. Das bereits Gezeigte ist auf $\widetilde{A}$ und M anwendbar. Deshalb folgt aus $F(\widetilde{A}) \subset F(A) \subset F(\widetilde{A}) \cup F(M)$ mit Borel-Mengen $F(\widetilde{A})$ und $F(M)$, $\mathcal{H}^k(F(M)) = 0$ die $\mathcal{H}^k$-Messbarkeit von $F(A)$ samt $\mathcal{H}^k(F(A)) = \mathcal{H}^k(F(\widetilde{A})) = \int_{\widetilde{A}} \mathrm{J}F\,\mathrm{d}\mathscr{L}^k = \int_A \mathrm{J}F\,\mathrm{d}\mathscr{L}^k$. Damit ist die Aussage (I) der Flächenformel im Fall ohne Nullstellen von $\mathrm{J}F$ bewiesen.

Zum Beweis der Aussage (II) betrachtet man eine $\mathcal{H}^k$-messbare Teilmenge S von $F(D)$. Da unter den gemachten Annahmen $\mathcal{H}^k(F(D)) < \infty$ gemäß Aussage (I) gilt, gibt es Borel-Mengen $\widetilde{S}$ und M mit $\mathcal{H}^k(M) = 0$ und $\widetilde{S} \subset S \subset \widetilde{S} \cup M \subset F(D)$. Damit gilt auch $F^{-1}(\widetilde{S}) \subset F^{-1}(S) \subset F^{-1}(\widetilde{S}) \cup F^{-1}(M) \subset D$ mit Borel-Mengen $F^{-1}(\widetilde{S})$ und $F^{-1}(M)$. Gemäß (I) folgt zudem $\int_{F^{-1}(M)} \mathrm{J}F\,\mathrm{d}\mathscr{L}^k = 0$ und im betrachteten Fall ohne Nullstellen von $\mathrm{J}F$ auch $\mathscr{L}^k(F^{-1}(M)) = 0$. Damit ist $F^{-1}(S)$ eine $\mathscr{L}^k$-messbare Menge, und insgesamt ist $(\mathcal{M}^k|D, \mathcal{M}_{\mathcal{H}^k}|F(D))$-Messbarkeit von F gezeigt. Aufgrund dieser Messbarkeitseigenschaft kann die Aussage (I) der Flächenformel, wie schon erwähnt, als $F_\sharp(\mathrm{J}F\cdot\mathscr{L}^k) = \mathcal{H}^k$ geschrieben werden. Die Aussage (II) ergibt sich hieraus gemäß der abstrakten Transformationsformel der früheren Proposition. $\square$

Beweis der Flächenformel im Fall, dass $\mathrm{J}F$ *Nullstellen besitzt.* Wie bei der vorausgehenden Argumentation lässt sich annehmen, dass D beschränkt ist, dass $\mathrm{D}F$ auf D gleichmäßig stetig ist, und dass $|\mathrm{D}F(x)v| \le M|v|$ für alle $x \in D$, $v \in \mathbb{R}^k$ mit festem $M \in \mathbb{R}_{>0}$ gilt. Vor allem ist nun zu zeigen, dass das Bild $F(C)$ der relativ abgeschlossenen Teilmenge $C := \{x \in D : \mathrm{J}F(x) = 0\}$ von D eine $\mathcal{H}^k$-Nullmenge ist, dass also $\mathcal{H}^k(C) = 0$ gilt, denn dann überträgt sich die Behauptung (I) problemlos von Teilmengen $A \setminus C$ der offenen Menge $D \setminus C$ (in der $\mathrm{J}F$ ja keine Nullstellen hat) auf Teilmengen A von D.

Für das Hauptargument sei $\varepsilon \in \mathbb{R}_{>0}$. Ähnlich der vorausgehenden Argumentation kann man $C = \bigcup_{i=1}^\infty C_i$ in abzählbar viele disjunkte Mengen $C_i \in \mathcal{B}(\mathbb{R}^k)$ mit $F(C_i) \in \mathcal{B}(\mathbb{R}^N)$ und zugehörigen $x_i \in C_i$ zerlegen, so dass jedes C_i außerdem in einer konvexen Teilmenge D_i von D liegt, und die Durchmesser der D_i können ausreichend klein gewählt werden, dass

$$|(\mathrm{D}F - \mathrm{D}F(x_i))v| \le \varepsilon|v| \qquad \text{auf } D_i$$

für alle $v \in \mathbb{R}^k$, $i \in \mathbb{N}$ gilt. Für jedes $i \in \mathbb{N}$ bedeutet $x_i \in C_i \subset C$ zudem $\mathrm{J}F(x_i) = 0$, also hat $\mathrm{D}F(x_i)$ Rang echt kleiner als k hat, und es gibt eine Nullrichtung $v_i \in \mathbb{R}^k$, $|v_i| = 1$ mit $\mathrm{D}F(x_i)v_i = 0$. Aus der vorausgehenden Abschätzung folgt dann

$$|\mathrm{D}F(x)v_i| \le \varepsilon \qquad \text{für alle } x \in D_i\,.$$

Man betrachtet jetzt (aus technischen Gründen, die unten klarer werden) die invertierbare Matrix $L_i \in \mathbb{R}^{k\times k}$ mit $L_i v_i = v_i$ und $L_i w = \frac{\varepsilon}{M} w$ für alle zu v_i orthogonalen $w \in \mathbb{R}^k$ und erhält $\mathscr{L}^k(L_i^{-1}C_i) = \big(\frac{M}{\varepsilon}\big)^{k-1}\mathscr{L}^k(C_i)$ aus der linearen Transformationsregel des Lemmas. Für $H_i := F \circ L_i \colon L_i^{-1}D_i \to \mathbb{R}^N$ und $x \in L_i^{-1}D_i$ ergeben sich

$$|\mathrm{D}H_i(x)v_i| = |\mathrm{D}F(L_i x)L_i v_i| = |\mathrm{D}F(L_i x)v_i| \le \varepsilon\,,$$
$$|\mathrm{D}H_i(x)w| = |\mathrm{D}F(L_i x)L_i w| = \tfrac{\varepsilon}{M}|\mathrm{D}F(L_i x)w| \le \varepsilon|w| \quad \text{für alle zu } v_i \text{ orthogonalen } w \in \mathbb{R}^k\,.$$

Gemäß dem Schrankensatz (der auf der konvexen Menge $L_i^{-1}D_i$ angewandt werden kann) folgt hieraus, dass H_i auf $L_i^{-1}D_i$ und insbesondere auf $L_i^{-1}C_i$ Lipschitz-stetig mit Lipschitz-Konstante 2ε ist. Die $\mathcal{H}^k$-Schranke für Lipschitz-Bilder liefert daher

$$\mathcal{H}^k(F(C_i)) = \mathcal{H}^k(H_i(L_i^{-1}C_i)) \le 2^k\varepsilon^k\mathscr{L}^k(L_i^{-1}C_i) = 2^k M^{k-1}\varepsilon\mathscr{L}^k(C_i)\,.$$

Diese Ungleichung kann über $i \in \mathbb{N}$ aufsummiert werden und besagt dann

$$\mathcal{H}^k(F(C)) \le 2^k M^{k-1}\varepsilon\mathscr{L}^k(C)\,.$$

Da D als beschränkt angenommen wurde, ist $\mathscr{L}^k(C) < \infty$, und durch Grenzübergang $\varepsilon \searrow 0$ erhält man $\mathcal{H}^k(F(C)) = 0$. Wie eingangs erklärt, ist die Aussage (I) der Flächenformel damit auch in der allgemeinen Situation bewiesen.

Zur Ableitung der Aussage (II) kann man sich wieder auf die abstrakte Transformationsformel berufen. Der Unterschied zum Fall ohne Nullstellen von $\mathrm{J}F$ besteht lediglich darin, dass man allgemein nur $(\mathcal{M}_{\mathrm{J}F\cdot\mathscr{L}^k}|D, \mathcal{M}_{\mathcal{H}^k}|F(D))$-Messbarkeit und nicht unbedingt $(\mathcal{M}^k|D, \mathcal{M}_{\mathcal{H}^k}|F(D))$-Messbarkeit von F erhält. □

Anwendungen (der Transformationsformel für $\mathscr{L}^N$)**.**

(1) Als **Spezialfälle** sind in der **Transformationsformel für Lebesgue-Maße** (und im Wesentlichen auch schon im Lemma für den linearen Fall)

- die **Bewegungsinvarianz** (für $x \in \mathbb{R}^N$, $T \in \mathcal{O}(\mathbb{R}^N)$)

$$\mathscr{L}^N(x+TA) = \mathscr{L}^N(A)\,, \qquad \int_{x+TA} f\,\mathrm{d}\mathscr{L}^N = \int_A f(x+Ty)\,\mathrm{d}\mathscr{L}^N(y)$$

- sowie das **Skalierungsverhalten** (für $r \in \mathbb{R}_{>0}$)

$$\mathscr{L}^N(rA) = r^N\mathscr{L}^N(A)\,, \qquad \int_{rA} f\,\mathrm{d}\mathscr{L}^N = r^N\int_A f(rx)\,\mathrm{d}\mathscr{L}^N(x)$$

von Lebesgue-Maßen und Lebesgue-Integralen enthalten.

Sehr nützliche Folgerungen aus der Bewegungsinvarianz, die für $\mathscr{L}^N$-messbare $A \subset \mathbb{R}^N$ und $\mathscr{L}^N$-messbare $f\colon A \to \mathbb{R}$ sowie beliebige $T \in \mathcal{O}(\mathbb{R}^N)$ gelten, aber **vor allem** auf **Spiegelungen** T Anwendung finden, sind:

- Regel für **„gerade Symmetrie"**: Ist A (bis auf $\mathscr{L}^N$-Nullmengen) eine disjunkte Vereinigung von $\widetilde{A} \in \mathcal{M}^N$ und $T\widetilde{A}$, und gilt $f \circ T = f$, so folgt $\int_A f \,\mathrm{d}\mathscr{L}^N = 2\int_{\widetilde{A}} f \,\mathrm{d}\mathscr{L}^N$.
- Regel für **„ungerade Symmetrie"**: Ist TA (bis auf $\mathscr{L}^N$-Nullmengen) gleich A, und gilt $f \circ T = -f$, so folgt $\int_A f \,\mathrm{d}\mathscr{L}^N = 0$, *sofern das Integral überhaupt existiert.*

Die Existenz des Integrals stellt man dabei typischerweise durch Angabe einer Majorante sicher (die in den folgenden Beispielen mit beschränkten Integranden auf beschränkten Mengen einfach konstant gewählt werden kann).

Einfache konkrete Beispiele, bei denen man durch „Hinsehen" ungerade Symmetrie erkennt und die man daher ohne jede Rechnung (!) behandeln kann, sind:

- $\int_{[-1,1]\times[-2,2]\times[1,3]}(x_1^3{-}3x_1x_2^2{+}x_1^2x_2)x_3^2 \,\mathrm{d}\mathscr{L}^N(x) = 0$ (da der Integrand ungerade Symmetrie in (x_1,x_2) aufweist und $[-1,1]{\times}[-2,2]{\times}[1,3]$ symmetrisch in (x_1,x_2) ist),
- $\int_{\mathrm{B}_1^N(0)} \frac{\partial}{\partial x_i}|x|^s \,\mathrm{d}\mathscr{L}^N(x) = 0$ für $i \in \{1,2,\ldots,N\}$ (da $|x|^s$ gerade Symmetrie in x_i und $\frac{\partial}{\partial x_i}|x|^s$ deshalb ungerade Symmetrie in x_i aufweist und $\mathrm{B}_1^N(0)$ symmetrisch in x_i ist),
- $\int_{\mathrm{B}_1^N(0)} x_ix_j \,\mathrm{d}\mathscr{L}^N(x) = 0$ für $i \neq j$ in $\{1,2,\ldots,N\}$ (da x_ix_j ungerade Symmetrie in x_i aufweist und $\mathrm{B}_1^N(0)$ symmetrisch in x_i ist).

Das letzte Integral mit $i = j$ wurde übrigens schon am Ende von Abschnitt 2.6 über Invarianz unter Koordinatenpermutationen (die auch als Spezialfall der Transformationsformel angesehen werden kann) berechnet.

(2) Mit der Transformationsformel kann man Maß- und Integralberechnungen in **ebenen Polarkoordinaten** durchführen, wozu man mittels

$$F(\varrho,\varphi) := (\varrho\cos\varphi\,,\ \varrho\sin\varphi) \qquad \text{für } (\varrho,\varphi) \in D := \mathbb{R}_{>0} \times (-\pi,\pi)$$

transformiert. Hier ist $F(D)$ ganz $\mathbb{R}^2$ ohne die $\mathscr{L}^2$-Nullmenge $\mathbb{R}_{\le 0}{\times}\{0\}$, und man berechnet als zugehörige Jacobische

$$\mathrm{J}F(\varrho,\varphi) = \varrho\,.$$

Speziell für Ringsegmente $S_{r,R}^{\alpha,\beta} := F(\,(r,R){\times}(\alpha,\beta)\,)$ mit $r \le R$ in $[0,\infty]$ und $\alpha \le \beta$ in $[-\pi,\pi]$ ergeben sich dann die besonders einfachen Transformationen

$$\mathcal{L}^2\big(S_{r,R}^{\alpha,\beta}\big) = \int_{(r,R)\times(\alpha,\beta)} \varrho \,\mathrm{d}\mathcal{L}^2(\varrho,\varphi) = \tfrac{1}{2}(R^2{-}r^2)(\beta{-}\alpha)\,,$$

$$\int_{S_{r,R}^{\alpha,\beta}} f \,\mathrm{d}\mathcal{L}^2 = \int_{(r,R)\times(\alpha,\beta)} f(\varrho\cos\varphi,\varrho\sin\varphi)\varrho \,\mathrm{d}\mathcal{L}^2(\varrho,\varphi)\,.$$

(3) In ähnlicher Weise kann man mit **Zylinderkoordinaten**

$$F(\varrho,\varphi,h) := (\varrho\cos\varphi\,,\ \varrho\sin\varphi, h) \qquad \text{für } (\varrho,\varphi,h) \in D := \mathbb{R}_{>0} \times (-\pi,\pi) \times \mathbb{R}$$

arbeiten. Für diese ist $F(D)$ ganz $\mathbb{R}^3$ ohne die $\mathscr{L}^3$-Nullmenge $\mathbb{R}_{\le 0}{\times}\{0\}{\times}\mathbb{R}$, und man berechnet als zugehörige Jacobische

$$\mathrm{J}F(\varrho,\varphi,h) = \varrho\,.$$

(4) Ein weiteres wichtiges Koordinatensystem sind **räumliche Polarkoordinaten** oder **Kugelkoordinaten** zu

$$F(\varrho,\varphi,\vartheta) := (\varrho\cos\varphi\sin\vartheta\,,\ \varrho\sin\varphi\sin\vartheta, \varrho\cos\vartheta) \quad \text{für } (\varrho,\varphi,\vartheta) \in D := \mathbb{R}_{>0}\times(-\pi,\pi)\times(0,\pi)\,.$$

Für diese ist $F(D)$ ganz $\mathbb{R}^3$ ohne die $\mathscr{L}^3$-Nullmenge $\mathbb{R}_{\le 0}\times\{0\}\times\mathbb{R}$, und man berechnet als zugehörige Jacobische

$$\mathrm{J}F(\varrho,\varphi,\vartheta) = \varrho^2\sin\vartheta\,.$$

An Bildern zu diesen Koordinatensystemen (wie sie in der Vorlesung gezeigt werden) wird plausibel, dass die Jacobische $\mathrm{J}F$ in der Transformationsformel die beim Übergang zum F-Bild durch (Kombinationen von) Streckungen und Stauchungen zustande kommenden Volumenänderungen berücksichtigt. Man kann sich $\mathrm{J}F\,\mathrm{d}\mathscr{L}^N$ auch als durch F abgebildetes infinitesimales Volumenelement und die **Jacobische** $\mathrm{J}F$ als **infinitesimalen Volumenänderungsfaktor** bei Abbildung von Mengen durch F vorstellen. Dies manifestiert sich in der (unter den Voraussetzungen $\mathscr{L}^N$-Transformationsformel gültigen und mit dieser leicht zu beweisenden) Formel

$$\mathrm{J}F(x) = \lim_{\varepsilon\searrow 0}\frac{\mathscr{L}^N\big(F\big(x+(0,\varepsilon)^N\big)\big)}{\mathscr{L}^N\big(x+(0,\varepsilon)^N\big)} \qquad\qquad \text{für } x\in D\,.$$

Ähnlich wie die Transformationsformel kann auch die Flächenformel zum Wechsel von Koordinatensystemen eingesetzt werden. Im Folgenden sollen aber vor allem Anwendungen in einer etwas anderen Richtung aufgezeigt werden, die nebenbei eine sehr elegante Bestimmung des $(N{-}1)$-dimensionalen Inhalts $\mathcal{H}^{N-1}(\mathrm{S}^{N-1}_\varrho(x_0))$ von Sphären $\mathrm{S}^{N-1}_\varrho(x_0)\subset\mathbb{R}^N$ ermöglichen:

Anwendungen (der Flächenformel)**.** Für Kugeln und Sphären werden stets die Notationen $\mathrm{B}^N_\varrho(x_0) := \{x\in\mathbb{R}^N : |x{-}x_0| < \varrho\}$ und $\mathrm{S}^{N-1}_\varrho(x_0) := \{x\in\mathbb{R}^N : |x{-}x_0| = \varrho\}$ verwendet.

(1) Eine Folgerung aus der Flächenformel ist die **Formel für sphärische Integration** oder sphärische Koflächenformel. Mit dieser Formel kann ein Integral über ein Ringgebiet als Doppelintegral geschrieben werden, bei dem man erst über im Ring enthaltene konzentrische Sphären um den Ring-Mittelpunkt und dann nach den Radien dieser Sphären integriert. Tatsächlich handelt es sich um die Formel

$$\int_{\mathrm{B}^N_R(x_0)\setminus\mathrm{B}^N_r(x_0)} f\,\mathrm{d}\mathscr{L}^N = \int_r^R\int_{\mathrm{S}^{N-1}_\varrho(x_0)} f\,\mathrm{d}\mathcal{H}^{N-1}\,\mathrm{d}\mathscr{L}^1(\varrho)\,.$$

für $x_0\in\mathbb{R}^N$, $r\le R$ in $[0,\infty]$ und $\mathscr{L}^N$-integrierbares $f\colon \mathrm{B}^N_R(x_0)\setminus\mathrm{B}^N_r(x_0)\to\overline{\mathbb{R}}$. Analog zum Satz von Fubini ist dabei $\int_{\mathrm{S}^{N-1}_\varrho(x_0)} f\,\mathrm{d}\mathcal{H}^{N-1}$ für $\mathscr{L}^1$-fast alle $\varrho\in(r,R)$ definiert, und $\varrho\mapsto\int_{\mathrm{S}^{N-1}_\varrho(x_0)} f\,\mathrm{d}\mathcal{H}^{N-1}$ ist $\mathscr{L}^1$-integrierbar über (r,R).

Beweisskizze. Es reicht denn Fall $x_0=0$, $r=0$, $R=\infty$ zu behandeln (denn der allgemeine Fall folgt durch Translation und Einsetzen solcher f, die auf $\mathrm{B}^N_R(x_0)\setminus\mathrm{B}^N_r(x_0)$ verschwinden). Man betrachtet nun „sektorielle“ Mengen

$$S_{a,b} := \{\varrho\omega : \varrho\in(a,b)\,,\ \omega\in S\}$$

mit $a<b$ in $\mathbb{R}_{>0}$ und einer Borel-Menge $S\subset\mathrm{S}^{N-1}_1(0)$, die als Bild $S=\Phi(A)$ einer Borel-Menge $A\subset D$ unter einer injektiven C^1-Abbildung $\Phi\colon D\to\mathrm{S}^{N-1}_1(0)$ auf offenem $D\subset\mathbb{R}^{N-1}$

geschrieben werden kann. Die hier vorausgesetzte Darstellung von S als injektives C^1-Bild ist dabei kaum[42] einschränkend und für jede Borel-Teilmenge S, die ganz in einer offenen Halbkugel von $\mathrm{S}_1^{N-1}(0)$ liegt, sicher erfüllt (denn die „obere" Halbkugel $\mathrm{S}_1^{N-1}(0) \cap (\mathbb{R}^{N-1} \times \mathbb{R}_{>0})$ ist das Bild von $\mathrm{B}_1^{N-1}(0)$ unter der Graphenabbildung $x \mapsto \big(x, \sqrt{1-|x|^2}\big)$, und jede andere offene Halbkugel geht aus dieser durch eine Drehung hervor). Unter dieser Voraussetzung kann jedenfalls $S_{a,b} = F((a,b) \times A)$ durch die injektive C^1-Abbildung $F \colon \mathbb{R}_{>0} \times D \to \mathbb{R}^N$ mit $F(\varrho, x) := \varrho\, \Phi(x)$ und[43] $\mathrm{J}F(\varrho, x) = \varrho^{N-1} \mathrm{J}\Phi(x)$ für $\varrho \in \mathbb{R}_{>0}$ und $x \in D$ parametrisiert werden. Mit der Transformationsformel, dem Satz von Fubini, der Flächenformel und dem Skalierungsverhalten von $\mathcal{H}^{N-1}$ erhält man deshalb

$$\begin{aligned}
\int_{\mathbb{R}^N} \mathbb{1}_{S_{a,b}} \,\mathrm{d}\mathscr{L}^N &= \int_{(a,b)\times A} \mathrm{J}F \,\mathrm{d}\mathscr{L}^N = \int_a^b \varrho^{N-1} \int_A \mathrm{J}\Phi \,\mathrm{d}\mathscr{L}^{N-1} \,\mathrm{d}\mathscr{L}^1(\varrho) \\
&= \int_a^b \varrho^{N-1} \mathcal{H}^{N-1}(S) \,\mathrm{d}\mathscr{L}^1(\varrho) = \int_a^b \mathcal{H}^{N-1}(\varrho S) \,\mathrm{d}\mathscr{L}^1(\varrho) \\
&= \int_a^b \int_{\mathrm{S}_\varrho^{N-1}(0)} \mathbb{1}_{\varrho S} \,\mathrm{d}\mathscr{H}^{N-1} \,\mathrm{d}\mathscr{L}^1(\varrho) \\
&= \int_0^\infty \int_{\mathrm{S}_\varrho^{N-1}(0)} \mathbb{1}_{(a,b)}(\varrho) \mathbb{1}_{\varrho S}(\eta) \,\mathrm{d}\mathscr{H}^{N-1}(\eta) \,\mathrm{d}\mathscr{L}^1(\varrho) \\
&= \int_0^\infty \int_{\mathrm{S}_\varrho^{N-1}(0)} \mathbb{1}_{S_{a,b}}(\eta) \,\mathrm{d}\mathscr{H}^{N-1}(\eta) \,\mathrm{d}\mathscr{L}^1(\varrho)\,.
\end{aligned}$$

Dies ist die Behauptung mit der charakteristischen Funktion $\mathbb{1}_{S_{a,b}}$ anstelle von f. Weil „sektorielle" Mengen $S_{a,b}$ ähnlich wie Quader oder Kugeln geeignete Umgebungssysteme von Punkten in $\mathbb{R}^N \setminus \{0\}$ beinhalten, erzeugen sie die Borel-σ-Algebra von $\mathbb{R}^N$, und die Behauptung folgt für charakteristische Funktionen Borelscher Mengen. Mit der Standard-Ausdehnungsprozedur geht man dann zu beliebigen $\mathscr{L}^N$-integrierbaren f über. □

(2) Für $x_0 \in \mathbb{R}^N$, $R \in \mathbb{R}_{>0}$ und $f \colon \mathbb{R}^N \to \overline{\mathbb{R}}$, das $\mathscr{L}^N$-summierbar auf einer Umgebung von $\overline{\mathrm{B}}_R(x_0)$ sowie stetig nahe $\mathrm{S}_R^{N-1}(x_0)$ ist, ergibt sich aus (1) die Identität

$$\frac{\mathrm{d}}{\mathrm{d}R} \int_{\mathrm{B}_R^N(x_0)} f \,\mathrm{d}\mathscr{L}^N = \int_{\mathrm{S}_R^{N-1}(x_0)} f \,\mathrm{d}\mathcal{H}^{N-1}\,.$$

Die **Ableitung eines Kugelintegrals nach dem Radius** ergibt also das zugehörige Sphären-/Oberflächenintegral.

[42] Tatsächlich kann sogar *jede* Borel-Menge $S \subsetneqq \mathrm{S}_1^{N-1}(0)$ in dieser Form geschrieben werden, denn durch die sogenannte stereographische Projektion kann man $\mathrm{S}_1^{N-1}(0)$ ohne einen beliebig gewählten Fußpunkt glatt parametrisieren. Für den Beweis zu (1) braucht man dies aber nicht zu wissen, sondern es reicht die oben begründete Parametrisierbarkeit von Halbkugeln.

[43] Um den angegebenen Zusammenhang zwischen den Jacobischen von Φ und F zu verifizieren, berechnet man die Funktionalmatrix $\mathrm{D}F(\varrho, \cdot\,) = \big(\,\Phi \,\big|\, \varrho\,\mathrm{D}\Phi\,\big) \in \mathbb{R}^{N\times N}$ und die hieraus gebildete Matrix

$$(\mathrm{D}F(\varrho,\cdot\,))^{\mathrm{T}} \mathrm{D}F(\varrho,\cdot\,) = \left(\begin{array}{c|c} \Phi^{\mathrm{T}}\Phi & \varrho\,\Phi^{\mathrm{T}}\mathrm{D}\Phi \\ \hline \varrho\,(\mathrm{D}\Phi)^{\mathrm{T}}\Phi & \varrho^2(\mathrm{D}\Phi)^{\mathrm{T}}\mathrm{D}\Phi \end{array}\right) = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & \varrho^2(\mathrm{D}\Phi)^{\mathrm{T}}\mathrm{D}\Phi \end{array}\right) \in \mathbb{R}^{N\times N}$$

(wobei aus $|\Phi| \equiv 1$ sowohl $\Phi^{\mathrm{T}}\Phi = |\Phi|^2 \equiv 1$ als auch $(\Phi^{\mathrm{T}}\mathrm{D}\Phi)^{\mathrm{T}} = (\mathrm{D}\Phi)^{\mathrm{T}}\Phi = \frac{1}{2}\nabla\big(|\Phi|^2\big) \equiv 0$ folgt). Insgesamt erhält man $\mathrm{J}F(\varrho,\cdot\,) = \sqrt{\det((\mathrm{D}F(\varrho,\cdot\,))^{\mathrm{T}}\mathrm{D}F(\varrho,\cdot\,))} = \sqrt{\det(\varrho^2(\mathrm{D}\Phi)^{\mathrm{T}}\mathrm{D}\Phi)} = \varrho^{N-1}\mathrm{J}\Phi$.

Beweis. Aus (1) (mit $r=0$) und dem HDI erhält man

$$\frac{\mathrm{d}}{\mathrm{d}R}\int_{\mathrm{B}^N_R(x_0)} f\,\mathrm{d}\mathscr{L}^N = \frac{\mathrm{d}}{\mathrm{d}R}\int_0^R \int_{\mathrm{S}^{N-1}_\varrho(x_0)} f\,\mathrm{d}\mathcal{H}^{N-1}\,\mathrm{d}\mathscr{L}^1(\varrho) = \int_{\mathrm{S}^{N-1}_R(x_0)} f\,\mathrm{d}\mathcal{H}^{N-1}\,,$$

wenn nur $\varrho \mapsto \int_{\mathrm{S}^{N-1}_\varrho(x_0)} f\,\mathrm{d}\mathcal{H}^{N-1}$ nahe R stetig ist.

Um die benötigte Stetigkeitsaussage zu verifizieren, benutzt man am besten die (aus der Definition des Integrals und der Skalierung von $\mathcal{H}^{N-1}$ folgende) Transformation $\int_{\mathrm{S}^{N-1}_\varrho(x_0)} f\,\mathrm{d}\mathcal{H}^{N-1} = \varrho^{N-1}\int_{\mathrm{S}^{N-1}_1(0)} f(x_0{+}\varrho\omega)\,\mathrm{d}\mathcal{H}^{N-1}(\omega)$ auf ein Integral über die Einheitssphäre. Auf der rechten Seite ist der Vorfaktor ϱ^{N-1} stetig in $\varrho \in \mathbb{R}_{>0}$ und das Integral stetig in $\varrho \in (R{-}\delta, R{+}\delta)$ mit ausreichende kleinem $\delta \in (0,R)$. Letzteres folgt dabei durch naheliegende Abschätzung mit Hilfe gleichmäßiger Stetigkeit von f auf $\mathrm{B}^N_{R+\delta}(x_0) \setminus \mathrm{B}^N_{R-\delta}(x_0)$. ☐

(3) Für **Flächeninhalte von Sphären** ergeben sich aus dem Spezialfall $f \equiv 1$ von (2) die Formeln

$$\mathcal{H}^{N-1}(\mathrm{S}^{N-1}_1(x_0)) = N\omega_N \qquad \text{und} \qquad \mathcal{H}^{N-1}(\mathrm{S}^{N-1}_\varrho(x_0)) = N\omega_N\varrho^{N-1}\,.$$

Bemerkung (zur allgemeinen Koflächenformel)**.** Die Formel der Anwendung (1) ist übrigens ein Spezialfall der **(allgemeinen) Koflächenformel**

$$\int_A f\,\mathrm{J}H\,\mathrm{d}\mathscr{L}^N = \int_{\mathbb{R}^k}\int_{A\cap H^{-1}(y)} f\,\mathrm{d}\mathcal{H}^{N-k}\,\mathrm{d}\mathscr{L}^k(y)$$

für $k \le N$ in $\mathbb{N}$, eine C^1-Abbildung $H\colon D \to \mathbb{R}^k$ auf offenem $D \subset \mathbb{R}^N$ sowie $\mathscr{L}^N$-messbare $A \subset D$ und $f\colon A \to \overline{\mathbb{R}}$, für die das Integral auf der linken Seite existiert. Die Jacobische $\mathrm{J}H$ ist dabei als

$$\mathrm{J}H = \sqrt{\det\left(\mathrm{D}H(\mathrm{D}H)^{\mathrm{T}}\right)}$$

definiert, und man kann die Koflächenformel als eine Art „verbogener Satz von Fubini“ verstehen, bei dem die inneren Integrale über $(N{-}k)$-dimensionale gekrümmte Fasern $A \cap H^{-1}(y)$ statt über flache Schnitte laufen. Zum Beweis der Koflächenformel kann man so vorgehen, dass man erst die (Fast-überall-)Existenz geeigneter lokaler Parametrisierungen der Fasern nachweist und dann ähnliche Rechnung wie beim Beweis der Anwendung (1) durchführt. Auf Details hierzu wird verzichtet.

2.12 Radon-Maße und Regularität

Über einem topologischen Raum Ω als Grundraum interessiert man sich besonders für Maße, die in gewisser Weise mit der Topologie von Ω verträglich sind. Die wichtigsten solchen Maße sind Radon-Maße, die sich vor allem durch gutartiges Verhalten auf Kompakta auszeichnen und unter folgenden schwachen Zusatzvoraussetzungen an den Grundraum Ω sinnvoll erklärt werden können:

Definition (**Lokal- und $\boldsymbol{\sigma}$-kompakte Hausdorff-Räume**)**.** *Sei Ω ein topologischer Raum.*

- *Man nennt Ω einen **Hausdorff-Raum**, wenn man je zwei Punkte in Ω durch offene Mengen trennen kann, wenn es also zu $x \ne y$ in Ω stets offene Teilmengen U und V von Ω mit $x \in U$, $y \in V$ und $U \cap V = \emptyset$ gibt.*

- *Man nennt Ω **lokalkompakt**, wenn jedes Element von Ω im Innern einer kompakten Teilmenge von Ω liegt.*
- *Man spricht von **σ-kompaktem** Ω, wenn Ω abzählbare Vereinigung kompakter Mengen ist.*

Bemerkungen (zur Hausdorff-Eigenschaft).

(1) Jeder metrische Raum ist ein Hausdorff-Raum.

(2) Die Bedeutung der Hausdorff-Eigenschaft liegt darin, dass (wie in der Analysis II gezeigt) alle Kompakta in einem Hausdorff-Raum abgeschlossen sind.

Bemerkungen (zu Lokal- und σ-Kompaktheit).

(1) Jede offene und jede abgeschlossene Teilmenge von $\mathbb{R}^N$ ist ein lokal- und σ-kompakter Hausdorff-Raum — ebenso jeder Schnitt einer offenen mit einer abgeschlossenen Teilmenge von $\mathbb{R}^N$. Zudem ist jede lokalkompakte Teilmenge in einem σ-kompakten metrischen Grundraum wie $\mathbb{R}^N$ auch σ-kompakt.

(2) Allerdings sind recht einfache Mengen wie $\mathbb{Q}$ und $(\mathbb{R}_{>0}\times\mathbb{R})\cup\{(0,0)\}$ nicht lokalkompakt, und $\mathbb{R}\setminus\mathbb{Q}$ ist weder lokal- noch σ-kompakt.

(3) Lokal- und σ-Kompaktheit sind nützlich, um die Existenz ausreichend vieler Kompakta sicherzustellen. Insbesondere ist Lokalkompaktheit äquivalent damit, dass jedes Kompaktum im Innern eines weiteren Kompaktums liegt.

Als Nächstes werden innere und äußere Regularität von Maßen (wobei die äußere Regularität dem Regularitätsbegriff des Abschnitts 2.4 für *äußere* Maße entspricht) und dann die Klasse der Radon-Maße eingeführt:

Definition (**reguläre Maße**). *Sei μ ein Maß auf $(\Omega,\mathscr{A})$ und $\mathscr{R}\subset\mathscr{A}$. Dann heißt μ **von außen $\mathscr{R}$-regulär**, wenn*

$$\mu(A)=\inf\{\mu(R)\,:\,R\in\mathscr{R}\text{ und }A\subset R\}\qquad\text{für alle }A\in\mathscr{A}$$

*gilt, und man nennt μ **von innen $\mathscr{R}$-regulär**, wenn*

$$\mu(A)=\sup\{\mu(R)\,:\,R\in\mathscr{R}\text{ und }R\subset A\}\qquad\text{für alle }A\in\mathscr{A}$$

gilt.

Definition (**Radon-Maße**). *Sei $(\Omega,\mathscr{A},\mu)$ ein Maßraum. Man nennt μ ein **Radon-Maß** auf $(\Omega,\mathscr{A})$, wenn Ω ein lokal- und σ-kompakter Hausdorff-Raum ist, alle Borel-Mengen messbar sind (d.h. $\mathcal{B}(\Omega)\subset\mathscr{A}$) und μ **von außen offen-regulär** (d.h. von außen $\mathcal{T}_\Omega$-regulär) sowie **lokal endlich** (d.h. jedes $x\in\Omega$ liegt in einer offenen μ-endlichen Menge) ist.*

Bemerkungen (zu Radon-Maßen).

(1) Die lokale Endlichkeit von μ ist (unter der in der Definition vorausgesetzten lokalen Kompaktheit von Ω) dazu äquivalent, dass $\mu(K)<\infty$ für jede kompakte Teilmenge K von Ω gilt.

(2) Radon-Maße sind (aufgrund der vorausgesetzten σ-Kompaktheit von Ω) stets σ-endlich.

Beispiele (von Radon-Maßen).

(1) Lebesgue-Maße $\mathscr{L}^N$ und Lebesgue-Stieltjes-Maße $\mathscr{L}^1_F$ sind Radon-Maße (und zwar sowohl auf der Borel-σ-Algebra als auch auf der σ-Algebra ihrer messbaren Mengen). Die benötigte Offen-Regularität ergibt sich dabei aus dem letzten Satz dieses Abschnitts.

(2) Dirac-Maße auf einem lokal- und σ-kompakten Hausdorff-Raum wie $\mathbb{R}^N$ sind Radon-Maße.

(3) Zählmaße und Hausdorff-Maße $\mathcal{H}^s$ mit $s < N$ sind *keine* Radon-Maße *auf ganz* $\mathbb{R}^N$, denn sie sind nicht σ-endlich. Ist aber A eine $\mathcal{H}^s$-endliche[44] Teilmenge von $\mathbb{R}^N$, so ist $\mathbb{1}_A \cdot \mathcal{H}^s$ doch ein Radon-Maß; auch dies folgt aus dem letzten Satz dieses Abschnitts.

Die Bedeutung von Radon-Maßen liegt in der folgenden Charakterisierung der Messbarkeit und der Möglichkeit zur gutartigen Approximation messbarer Mengen:

Satz (Charakterisierungen messbarer Mengen). *Sei μ ein Radon-Maß auf $(\Omega, \mathscr{A})$. Für $A \in \mathcal{P}(\Omega)$ sind folgende Aussagen* ***äquivalent***:

(0) *Es gilt $A \in \mathcal{M}_\mu$, d.h. A ist μ-messbar.*

(I) *Zu jedem $\varepsilon \in \mathbb{R}_{>0}$ gibt es eine in Ω offene Menge $O \supset A$ mit $\mu^*(O \setminus A) < \varepsilon$.*

(II) *Zu jedem $\varepsilon \in \mathbb{R}_{>0}$ gibt es eine in Ω abgeschlossene Menge $C \subset A$ mit $\mu^*(A \setminus C) < \varepsilon$.*

(III) *Es gibt eine G_δ-Menge G in Ω und eine $\overline{\mu}$-Nullmenge M mit $A = G \setminus M$.*

(IV) *Es gibt eine F_σ-Menge F in Ω und eine $\overline{\mu}$-Nullmenge M mit $A = F \cup M$.*

Im Fall $\mu^*(A) < \infty$ sind auch folgende Aussagen ***äquivalent*** *zu den vorausgehenden*:

(V) *Zu jedem $\varepsilon \in \mathbb{R}_{>0}$ gibt es eine kompakte Menge $K \subset A$ mit $\mu^*(A \setminus K) < \varepsilon$.*

(VI) *Es gilt*

$$\inf\{\mu(O) : O \text{ offen in } \Omega \text{ mit } A \subset O\} = \sup\{\mu(K) : K \text{ kompakt mit } K \subset A\}.$$

Es sei hier daran erinnert, dass $\mu^*(A) = \overline{\mu}(A)$ für $A \in \mathcal{M}_\mu$ und $\mu^*(A) = \overline{\mu}(A) = \mu(A)$ für $A \in \mathscr{A}$ gelten. Insbesondere können somit für ein vollständiges Radon-Maß μ und eine bekanntermaßen messbare Menge A mit $\mu(A) < \infty$ die Eigenschaften (I)–(VI) des Satzes alle mit μ anstelle von μ^* und $\overline{\mu}$ verwendet werden.

Zudem folgt aus dem Satz, dass Radon-Maße immer auch eine Regularität von innen vorliegt:

Korollar. *Jedes Radon-Maß μ auf $(\Omega, \mathscr{A})$ ist von innen kompakt-regulär (d.h. es ist von innen $\mathcal{K}_\Omega$-regulär mit dem System $\mathcal{K}_\Omega$ aller kompakten Teilmengen von Ω).*

Beweis des Satzes. Da Ω per Definition eines Radon-Maßes σ-kompakt ist, gibt es eine ohne Einschränkung aufsteigende Folge $K_1 \subset K_2 \subset K_3 \subset \ldots$ von Kompakta mit $\bigcup_{i=1}^\infty K_i = \Omega$.

Für „(0) $\Longrightarrow$ (I)“ schreibt man $A = B \setminus M$ mit $B \in \mathscr{A}$ und $\overline{\mu}(M) = 0$. Da $\mu(B \cap K_i) < \infty$ gilt, gibt es zu $\varepsilon \in \mathbb{R}_{>0}$ gemäß der Offen-Regularität von μ offene $O_i \supset B \cap K_i$ mit $\mu(O_i \setminus (B \cap K_i)) < \frac{\varepsilon}{2^i}$. Es folgt, dass $O := \bigcup_{i=1}^\infty O_i$ offen ist mit $A \subset B \subset O$ und

$$\overline{\mu}(O \setminus A) = \mu(O \setminus B) \le \sum_{i=1}^\infty \mu(O_i \setminus (B \cap K_i)) < \varepsilon\,.$$

Die Implikation „(I) $\Longrightarrow$ (III)“ erhält man durch Schneiden der offenen Mengen zu $\varepsilon = 1, \frac{1}{2}, \frac{1}{3}, \ldots$, und „(III) $\Longrightarrow$ (0)“ gilt trivial. Somit ist die Äquivalenz von (0), (I), (III) gezeigt, und die Äquivalenz zu (II) und (IV) folgt daraus durch Übergang zu Komplementen.

Für die restliche Argumentation sei $\mu^*(A) < \infty$.

Für „(0)+(II) $\Longrightarrow$ (V)“ kann man dann von $A \in \mathcal{M}_\mu$ mit $\overline{\mu}(A) < \infty$ ausgehen, und es gibt zu $\varepsilon \in \mathbb{R}_{>0}$ ein abgeschlossenes $C \subset A$ mit $\mu(C) > \overline{\mu}(A) - \varepsilon$. Wegen $\mu(C) = \lim_{i\to\infty} \mu(C \cap K_i)$ gibt es auch ein $i \in \mathbb{N}$ mit $\mu(C \cap K_i) > \mu(A) - \varepsilon$, und da $C \cap K_i$ als abgeschlossene Teilmenge eines Kompaktums wieder kompakt ist, ist (V) nachgewiesen. Da Ω nach Definition eines Radon-Maßes ein Hausdorff-Raum ist und dort, wie schon bemerkt, alle Kompakta abgeschlosssen sind, gilt „(V) $\Longrightarrow$ (II)“ trivial. Also ist die Äquivalenz der Bedingungen (0)–(V) gezeigt.

Um „(VI) $\Longrightarrow$ (I)“ einzusehen, überlegt man sich, dass der übereinstimmende Wert des Infimums und des Supremums in (VI) aus Monotonie-Gründen auch mit $\mu^*(A)$ übereinstimmt. Wegen $\mu^*(A) < \infty$ folgt dann (I). Schließlich ist „(I)+(V) $\Longrightarrow$ (VI)“ einfach einzusehen. Damit ist die Äquivalenz aller Bedingungen (0)–(VI) nachgewiesen. □

[44] Allgemeiner ist $\mathbb{1}_A \cdot \mathcal{H}^s$ immer dann ein Radon-Maß, wenn es lokal endlich ist. Dafür reicht $\mathcal{H}^s$-σ-Endlichkeit von A allerdings noch nicht, wie für $s = 1$ und $N = 2$ das Beispiel $A = \big\{(x,y) \in \mathbb{R}^2 : x \ne 0, \frac{y}{x} \in \mathbb{Q}\big\}$ zeigt.

Beweis des vorausgehenden Korollars. Da $\mathcal{K}_\Omega \subset \mathcal{B}(\Omega) \subset \mathscr{A}$ gilt, ist nur $\mu(A) = \sup\{\mu(K) : K \in \mathcal{K}_\Omega \text{ und } K \subset A\}$ für alle $A \in \mathscr{A}$ zu begründen: Im Fall $\mu(A) < \infty$ liest man dies per Monotonie aus Eigenschaft (VI) des Satzes ab. Im Fall $\mu(A) = \infty$ gilt $\lim_{i\to\infty} \mu(A \cap K_i) = \infty$ für die Kompakta K_i aus dem Beweis des Satzes. Wegen $\mu(A \cap K_i) \le \mu(K_i) < \infty$ liefert Teil (V) des Satzes dann weitere Kompakta $\widetilde{K}_i \subset A \cap K_i$ mit $\lim_{i\to\infty} \mu(\widetilde{K}_i) = \infty$. Somit ist das vorausgehende Supremum unendlich und stimmt mit $\mu(A) = \infty$ überein. □

Korollar (über die **Eindeutigkeit von $\mathscr{L}^N$**). *Das vervollständigte Lebesgue-Maß $\mathscr{L}^N$ ist das einzige translationsinvariante und vollständige Radon-Maß auf $\mathbb{R}^N$ mit $\mathscr{L}^N\big((0,1)^N\big) = 1$.*

Beweis. Nach dem Eindeutigkeitssatz des Abschnitts 2.2 stimmt jedes weitere Maß μ mit den angegebenen Eigenschaften auf $\mathcal{B}(\mathbb{R}^N)$ mit $\mathscr{L}^N$ überein. Per Minimalität der Vervollständigung folgt dann $\mathcal{M}^N \subset \mathcal{M}_\mu$ und $\mu|_{\mathcal{M}^N} = \mathscr{L}^N$. Also bleibt nur $\mathcal{M}_\mu \subset \mathcal{M}^N$ zu zeigen: Gemäß dem vorigen Satz gibt es zu $A \in \mathcal{M}_\mu$ erst ein $G \in \mathcal{B}(\mathbb{R}^N)$ und eine μ-Nullmenge M mit $A = G \setminus M$ und dann auch ein $\widetilde{G} \in \mathcal{B}(\mathbb{R}^N)$ und eine μ-Nullmenge $\widetilde{M}$ mit $M = \widetilde{G} \setminus \widetilde{M}$. Wegen $\mathscr{L}^N(\widetilde{G}) = \mu(\widetilde{G}) \le \mu(M)+\mu(\widetilde{M}) = 0$ folgen $M \in \mathcal{M}^N$ und $A \in \mathcal{M}^N$. □

Korollar („**Radon-Maße auf Intervallen sind Lebesgue-Stieltjes-Maße**“). *Sei I ein offenes Intervall in $\mathbb{R}$. Zu jedem vollständigen Radon-Maß μ auf $(I, \mathcal{M}_\mu)$ gibt es ein nichtfallendes $F\colon I \to \mathbb{R}$ mit $\mu = \mathscr{L}^1_F$ auf $\mathcal{M}_\mu = \mathcal{M}^1_F$.*

Beweis. Für beliebig fixiertes $x_0 \in I$ wird $F\colon I \to \mathbb{R}$ durch $F(x) := \mu([x_0, x))$ für $x \ge x_0$ und $F(x) := -\mu([x, x_0))$ für $x \le x_0$ definiert. Dann ist F nichtfallend, und für $a \le b$ in $I_{\ge x_0}$ gilt

$$\mathscr{L}^1_F([a,b)) = F(b-) - F(a-) = \mu([x_0,b)) - \mu([x_0,a)) = \mu([a,b))\,.$$

Nach analoger Behandlung andere Fälle erhält man, dass μ auf $\mathcal{I}_1 \cap \mathcal{P}(I)$ mit λ^1_F übereinstimmt. Gemäß dem Maßfortsetzungssatz folgt $\mu = \mathscr{L}^1_F$ auf $\mathcal{B}(I)$. Die Übereinstimmung $\mu = \mathscr{L}^1_F$ auf $\mathcal{M}_\mu = \mathcal{M}^1_F$ folgt mit derselben Argumentation wie im Beweis des vorigen Korollars. □

Bemerkungen (zur Klassifikation der Radon-Maße auf Intervallen).

(1) Sei $x_0 \in I$ beliebig fixiert. Der Beweis zeigt dann, dass F stets linksseitig stetig mit $F(x_0) = 0$ gewählt werden kann. Stellt man diese beiden Zusatzbedingungen, so ist F eindeutig durch $\mu = \mathscr{L}^1_F$ (und sogar durch λ^1_F) bestimmt. Folglich ist $F \mapsto \mathscr{L}^1_F$ eine Bijektion zwischen den linksseitig stetigen, nichtfallenden Funktionen $F\colon I \to \mathbb{R}$ mit $F(x_0) = 0$ und den vollständigen Radon-Maßen auf I.

(2) In der Wahrscheinlichkeitstheorie gibt man für $I = \mathbb{R}$ anstelle eines Wertes $F(x_0)$ das Verhalten von F im Unendlichen vor (was aber nur für endliche Maße sinnvoll ist) und betrachtet die Einschränkungen $\mathscr{L}^1_F$ auf die Borelsche σ-Algebra. Man erhält dann eine Bijektion $F \mapsto \mathscr{L}^1_F$ zwischen den linksseitig stetigen, nichtfallenden $F\colon \mathbb{R} \to \mathbb{R}$ mit

$$\lim_{x\to\infty} F(x) = 0\,, \qquad \lim_{x\to\infty} F(x) = 1$$

und den Wahrscheinlichkeitsmaßen auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Ist ein Wahrscheinlichkeitsmaß μ auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ gegeben, so nennt man die derart zugehörige Funktion F die **Verteilungsfunktion** von μ. Sie erfüllt $\mu = \mathscr{L}^1_F$ und ist gegeben durch

$$F(x) = \mu\big((-\infty, x)\big) \qquad \text{für } x \in \mathbb{R}\,.$$

(3) Alles geht analog mit rechtsseitiger Stetigkeit, wenn man linksoffene statt rechtsoffener Intervalle verwendet.

Der Nachweis der Radon-Maß-Eigenschaft wird oft beträchtlich erleichtert durch den folgenden Satz, gemäß dem man in der metrischen Situation statt Offen-Regularität nur Borel-Regularität zu zeigen braucht. Für Lebesgue- und Hausdorff-Maße wurde von dieser Tatsache bereits Gebrauch gemacht.

Satz. *Wird die Topologie auf Ω durch eine Metrik induziert, so ist es äquivalent in der Definition des Radon-Maßes nur äußere $\mathcal{B}(\Omega)$-Regularität statt äußerer Offen-Regularität zu fordern.*

Der Beweis des Satzes ähnelt den bereits auf Übungsblatt 5 behandelten Überlegungen und wird hier nur skizziert:

Beweisskizze. Sei μ ein Maß auf $(\Omega, \mathscr{A})$, das von außen $\mathcal{B}(\Omega)$-regulär ist und außer Offen-Regularität alle Bedingungen aus der Definition des Radon-Maßes erfüllt. Dann prüft man mit dem $\frac{\varepsilon}{2^i}$-Trick nach, dass das System $\mathscr{G}$ aller Mengen $A \in \mathscr{A}$, die die Gleichheiten

$$\begin{aligned}\mu(A) &= \inf\{\mu(O) : O \text{ offen in } \Omega \text{ mit } A \subset O\}\\ &= \sup\{\mu(C) : C \text{ abgeschlossen in } \Omega \text{ mit } C \subset A\}\end{aligned}$$

erfüllen, abgeschlossen unter abzählbarer Vereinigung ist. Aus der Definition ist klar, dass $\mathscr{G}$ unter Komplementbildung abgeschlossen und folglich eine σ-Algebra ist. Zeigen wir nun, dass $\mathscr{G}$ alle abgeschlossenen und somit auch alle offenen Mengen enthält, so folgt $\mathcal{B}(\Omega) \subset \mathscr{G}$ und $\mu(A) = \inf\{\ldots\}$ gilt für alle $A \in \mathcal{B}(\Omega)$. Mit der Borel-Regularität folgt diese Gleichheit auch für alle $A \in \mathscr{A}$, und μ ist von außen offen-regulär.

Zum Beweis des Satzes bleibt also nur zu verifizieren, dass die Gleichheit $\mu(A) = \inf\{\ldots\}$ für jede abgeschlossene Menge A richtig ist (denn $\mu(A) = \sup\{\ldots\}$ gilt für abgeschlossenes A trivial). Dazu argumentiert man wie folgt: Wegen σ-Kompaktheit von Ω gibt es Kompakta $K_1, K_2, K_3, \ldots$ mit $\Omega = \bigcup_{i=1}^\infty K_i$. Mit Hilfe der lokalen Kompaktheit von Ω lässt sich einsehen, dass die Kompakta $A \cap K_i$ im Innern weiterer Kompakta liegen. Unter Rückgriff auf die metrische Struktur folgt dann, dass für $k \gg 1$ die offenen Mengen

$$U_{1/k}(A \cap K_i) := \{x \in \Omega : \operatorname{dist}(x, A \cap K_i) < 1/k\}$$

endliches Maß haben und daher

$$\mu(A \cap K_i) = \lim_{k\to\infty} \mu\big(U_{1/k}(A \cap K_i)\big)$$

gilt. Damit gibt es zu $\varepsilon \in \mathbb{R}_{>0}$ und $i \in \mathbb{N}$ stets eine offene Obermenge O_i von $A \cap K_i$ mit $\mu(O_i \setminus (A \cap K_i)) < \frac{\varepsilon}{2^i}$, und $O := \bigcup_{i=1}^\infty O_i$ ist eine offene Obermenge von A mit $\mu(O \setminus A) < \varepsilon$. Es folgt $\mu(A) = \inf\{\ldots\}$, und nur dies blieb zu zeigen. $\square$

Kapitel 3

Untermannigfaltigkeiten, Differentialformen, Integralsätze

3.1 Untermannigfaltigkeiten des $\mathbb{R}^N$

Das folgende **Konzept $\boldsymbol{k}$-dimensionaler Flächen in $\mathbb{R}^{\boldsymbol{N}}$** ist **von zentraler Bedeutung** für weite Teile der modernen Mathematik:

Definition (**Untermannigfaltigkeiten**). *Seien $k, N \in \mathbb{N}$ mit $k \le N$ und $m \in \mathbb{N} \cup \{\infty, \omega\}$. Man nennt eine Teilmenge M von $\mathbb{R}^N$ eine **$\boldsymbol{k}$-dimensionale $\mathrm{C}^{\boldsymbol{m}}$-Untermannigfaltigkeit mit Rand** des $\mathbb{R}^N$, wenn es zu jedem $a \in M$ offene Umgebungen U von 0 in $\mathbb{R}^k$ und V von a in $\mathbb{R}^N$ sowie eine C^m-Abbildung $F \colon U \to \mathbb{R}^N$ von maximalem Rang[1] mit $F(0) = a$ gibt, so dass F Homöomorphismus entweder von U auf $M \cap V$ (innerer Fall) oder von $U_{\le 0} := \{\xi \in U : \xi_1 \le 0\}$ auf $M \cap V$ (Randfall) ist.*

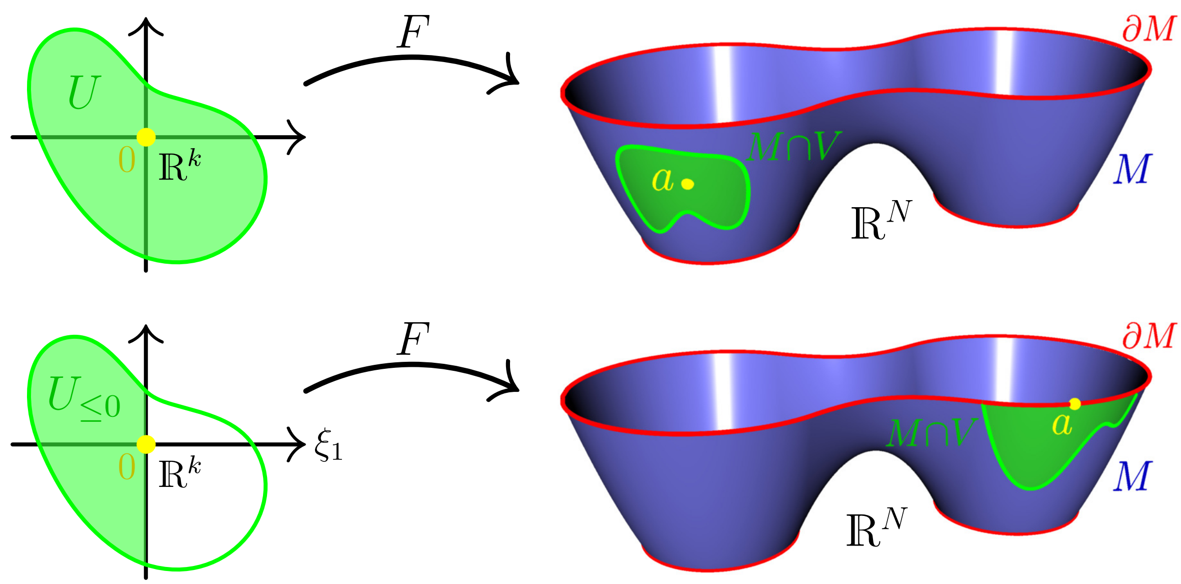


Abb. 6: Definition einer Untermannigfaltigkeit im inneren Fall und im Randfall

[1] Diese Rangbedingung meint $\mathrm{Rang}\, F' \equiv k$ auf U für die totale Ableitung F'. Äquivalent ist $\mathrm{Rang}(\mathrm{D}F) \equiv k$ auf U für die Jacobi-Matrix $\mathrm{D}F$ und auch $\mathrm{J}F > 0$ auf U für die Jacobische $\mathrm{J}F = \sqrt{\det((\mathrm{D}F)^{\mathrm{T}}\mathrm{D}F)}$.

Bemerkungen (zur Definition von Untermannigfaltigkeiten)**.**

(1) Eine Abbildung F mit den Eigenschaften der Definition bezeichnet man als **reguläre $\mathbf{C}^m$-Parametrisierung** des Teils $M \cap V = F(U)$ beziehungsweise $M \cap V = F(U_{\le 0})$ von M. Die Rangbedingung (die durch das Adjektiv „regulär“ zum Ausdruck gebracht wird) stellt dabei sicher, dass das Stück $M \cap V = F(U)$ beziehungsweise $M \cap V = F(U_{\le 0})$ in der Tat einem Stück von $\mathbb{R}^k$ beziehungsweise $(\mathbb{R}^k)_{\le 0}$ „ähnelt“ und in diesem Sinn k-dimensional ist.

(2) Wie erst etwas später anhand des Tangentialkegels präzise begründet wird, können für einen Punkt $a \in M$ — auch für unterschiedliche Wahlen von U, V und F — nie sowohl die Situation des inneren Fall als auch die des Randfalls vorliegen.

Die Zweiteilung in den inneren Fall und den Randfall führt nun zu weiteren Begriffen.

Definitionen (**geometrische Begriffe bei Untermannigfaltigkeiten**)**.** *Sei M eine k-dimensionale C^m-Untermannigfaltigkeit mit Rand des $\mathbb{R}^N$. In Anbetracht der vorausgehenden Bemerkung bezeichnet man die Menge der Punkte $a \in M$, für die der innere Fall der vorigen Definition eintritt, als das* ***geometrisch Innere*** *$M \setminus \partial M$ von M, und die Menge der Punkte $a \in M$, für die der Randfall der vorigen Definition eintritt, als den* ***geometrischen Rand*** *∂M von M. Im Fall $\partial M = \emptyset$ heißt M eine* ***Untermannigfaltigkeit ohne Rand****. Eine kompakte Untermannigfaltigkeit ohne Rand nennt man auch* ***geschlossene Untermannigfaltigkeit****.*

Bemerkungen (zu Untermannigfaltigkeiten)**.**

(0) Gelegentlich ist es sinnvoll, als ergänzende Konventionen festzulegen, dass man unter einer 0-dimensionalen Untermannigfaltigkeit des $\mathbb{R}^n$ eine Teilmenge von $\mathbb{R}^n$ versteht, die nur aus isolierten Punkten besteht, und, dass für solche $\partial M = \emptyset$ sein soll.

(1) Ist M eine Untermannigfaltigkeit mit Rand des $\mathbb{R}^N$, so sind das geometrisch Innere $M \setminus \partial M$ und der geometrische Rand ∂M Untermannigfaltigkeiten *ohne Rand* des $\mathbb{R}^N$. Diese haben (mindestens) dieselbe C^m-Regularität wie M selbst, $M \setminus \partial M$ hat dieselbe Dimension wie M, und ∂M hat die nächst niedrigere Dimension.

Die Behauptungen über $M \setminus \partial M$ folgen dabei direkt aus den Definitionen, die über ∂M begründet man durch Einschränkung von Parametrisierungen auf $U_{=0} := \{\xi \in U : \xi_1 = 0\}$ und wieder mit dem schon erwähnten Sachverhalt, dass sich innerer Fall und Randfall ausschließen.

(2) Es gilt den **geometrischen Rand** — trotz gleicher Notation mit dem Operator ∂ — **sorgfältig vom topologischen Rand** zu **unterscheiden**, denn im Fall $k < N$ ist der geometrische Rand eine deutlich kleinere Menge als der topologische Rand.

Beispiele und Bemerkungen (zu Untermannigfaltigkeiten)**.**

(1) Die folgenden Abbildungen 7 bis 11 illustrieren **typische Beispiele 1-dimensionaler Untermannigfaltigkeiten**.

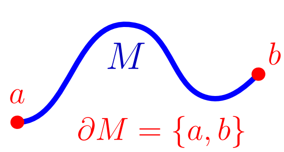


Abb. 7: Eine Kurve M inklusive Endpunkte

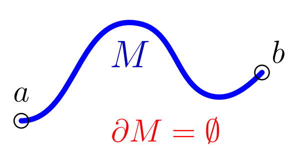


Abb. 8: Eine Kurve M exklusive Endpunkte

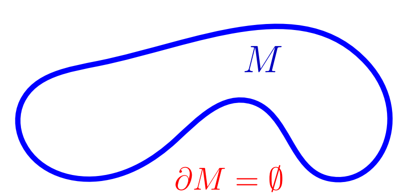


Abb. 9: Eine geschlossene Kurve M

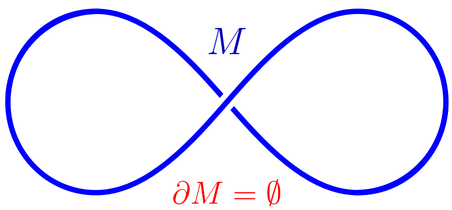


Abb. 10: Eine weitere geschlossene Kurve M (ohne Selbstschnitt in $\mathbb{R}^3$)

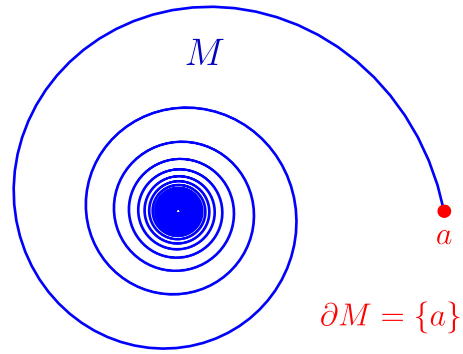


Abb. 11: Eine Spiralkurve (eventuell unendlicher Länge) exklusive Spiralpunkt

Die nun folgenden Abbildungen 12 bis 14 zeigen dagegen Mengen, die ***keine*** **1-dimensionalen Untermannigfaltigkeiten** sind.

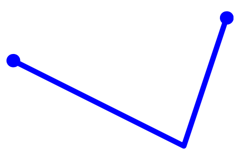


Abb. 12: Eine Kurve mit Knick

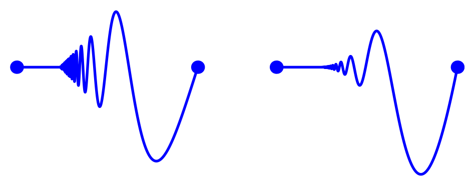


Abb. 13: Graphen von nicht differenzierbaren und Nicht-C^1-Funktionen

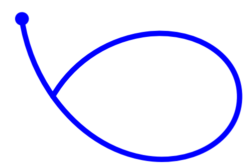


Abb. 14: Eine Schlinge mit „Fast-Selbstschnitt“

Die Kurve der Abbildung 12 kann hierbei mit einer Nullstelle der Ableitung beim Knick C^1-parametrisiert werden, erlaubt aber keine *reguläre* C^1-Parametrisierung ohne solche Nullstellen. Genauso verhält es sich, sofern die Graphenkurve endliche Länge hat, beim Nicht-C^1-Graph der Abbildung 13. Die Schlinge der Abbildung 14 schließlich kann sogar ohne Nullstellen der Ableitung über einem halboffenen Intervall C^1-parametrisiert werden, doch eine solche Parametrisierung kann nie *homöomorph* auf ein Stück der Schlinge (im Sinn einer relativen Umgebung) um den „Abzweigepunkt“ abbilden. Somit wird klar, dass die Homöomorphie-Forderung in der Definition von Untermannigfaltigkeiten gerade zum Ausschluss von Beispielen des Typs der Abbildung 14 dient.

(2) Die folgenden Abbildungen 15 bis 19 zeigen **typische Beispiele 2-dimensionaler Untermannigfaltigkeiten**.

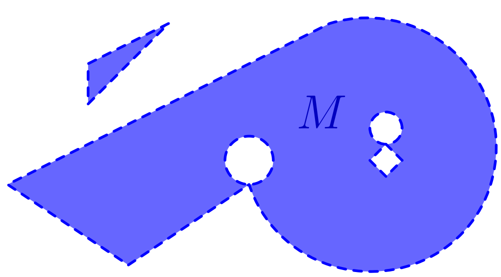


Abb. 15: Ein beliebiger offener Bereich M in der Ebene mit $\partial M = \emptyset$

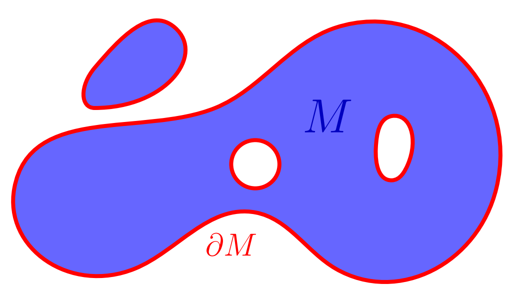


Abb. 16: Ein abgeschlossener Bereich M mit C^m-Rand ∂M in der Ebene

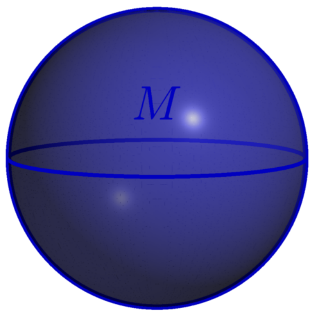


Abb. 17: Die Einheitssphäre $M = \mathrm{S}_1^2 = \partial\mathrm{B}_1^3$ mit $\partial M = \emptyset$

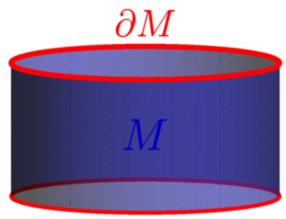


Abb. 18: Der Zylindermantel $M = \mathrm{S}_1^1{\times}[0,1]$ mit $\partial M = \mathrm{S}_1^1{\times}\{0,1\}$

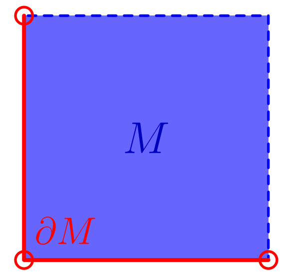


Abb. 19: Das teiloffene Einheitsquadrat $M = [0,1)^2\setminus\{(0,0)\}$ mit $\partial M = ((0,1){\times}\{0\})\,\dot\cup\,(\{0\}{\times}(0,1))$

Hierbei darf in der Situation von Abbildung 15 der nicht in M enthaltene topologische Rand von M auch Knicke und Singularitäten aufweisen, während sich der in einer Untermannigfaltigkeit M enthaltene Rand aus Abbildung 16 gutartig verhalten muss. Weiterhin handelt es sich unter den Beispielen der Abbildungen 15 bis 19 einzig bei der Einheitssphäre der Abbildung 17 um eine geschlossene Untermannigfaltigkeit. Schließlich ist beim teiloffenen Einheitsquadrat der Abbildung 19 das Augenmerk vor allem auf die vier Eckpunkte in $\{0,1\}^2$ zu richten, da bei diesen der demnächst eingeführte Tangentialkegel ein Quadrant ist und M nahe diesen Punkten nicht im Sinn der Definition regulär C^1-parametrisiert werden kann. Nur da alle vier Eckpunkte *nicht* zur Menge M gehören, ist diese tatsächlich eine Untermannigfaltigkeit. Dagegen ist das in Abbildung 20 gezeigte halboffene Einheitsquadrat inklusive des problematischen Eckpunkts $(0,0)$ ***keine*** **2-dimensionale Untermannigfaltigkeit**.

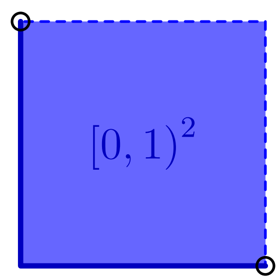


Abb. 20: Das halboffene Einheitsquadrat $[0,1)^2$

(3) Jede offene Teilmenge des $\mathbb{R}^N$ ist eine N-dimensionale C^ω-Untermannigfaltigkeit ohne Rand des $\mathbb{R}^N$ (denn in diesen Fällen maximaler Dimension ergeben sich die regulären C^m-Parametrisierungen trivial als Verschiebungen $\xi \mapsto \xi{+}a$, die für eine in der offenen Menge enthaltene offene Umgebung V von a nur $U = V{-}a$ auf V verschieben).

Der nächste Satz liefert alternative Beschreibungen von Untermannigfaltigkeiten, die auch als zum obigen Vorgehen äquivalente Definitionen angesehen werden können.

Satz (über **Charakterisierungen von inneren Punkten einer Untermannigfaltigkeit**). *Seien $k, N \in \mathbb{N}$ mit $k \le N$ und $m \in \mathbb{N} \cup \{\infty\}$. Dann sind folgende Bedingungen an eine Teilmenge M von $\mathbb{R}^N$ nahe eines fixierten Punktes $a \in M$* ***äquivalent*****:**

(1) ***Parametrische Darstellung*** (*wie in der Definition*): *Es gibt offene Umgebungen* U *von* 0 *in* $\mathbb{R}^k$ *und* V *von* a *in* $\mathbb{R}^N$ *sowie eine* C^m*-Abbildung* $F\colon U \to \mathbb{R}^N$ *von maximalem Rang* (*d.h.* $\operatorname{Rang} F' \equiv k$ *auf* U) *mit* $F(0) = a$, *so dass* F *homöomorph von* U *auf*

$$M \cap V = F(U)$$

abbildet.

(2) ***Explizite Darstellung*** *als* ***gedrehter Graph***: *Es gibt offene Umgebungen* U *von* 0 *in* $\mathbb{R}^k$ *und* V *von* a *in* $\mathbb{R}^N$, *eine lineare Isometrie* $T \in \mathcal{O}(\mathbb{R}^N)$ *sowie eine* C^m*-Abbildung* $f\colon U \to \mathbb{R}^{N-k}$ *mit* $f(0) = 0$ *und*

$$M \cap V = a + T(\operatorname{Graph} f)$$

(*wobei natürlich* $\operatorname{Graph} f := \{(x, f(x)) : x \in U\}$).

(3) ***Implizite Darstellung*** *als gemeinsame Lösungsmenge von* $\boldsymbol{(N-k)}$ ***Gleichungen*** $g_i(x)=0$: *Es gibt eine offene Umgebung* V *von* a *in* $\mathbb{R}^N$ *und eine* C^m*-Abbildung* $g\colon V \to \mathbb{R}^{N-k}$ *von maximalem Rang* (*d.h.* $\operatorname{Rang} g' \equiv N-k$ *auf* V) *mit*

$$M \cap V = g^{-1}(0)\,.$$

(4) ***Darstellung durch Geradebiegen*** (*mittels* Φ^{-1}): *Es gibt offene Umgebungen* Σ *von* 0 *in* $\mathbb{R}^N$ *und* V *von* a *in* $\mathbb{R}^N$ *sowie einen* C^m*-Diffeomorphismus* $\Phi\colon \Sigma \to V$ *von* Σ *auf* V *mit* $\Phi(0) = a$ *und*

$$M \cap V = \Phi((\mathbb{R}^k \times \{0\}) \cap \Sigma)\,.$$

Beweis. Als Erstes wird die Implikation „(1) $\Longrightarrow$ (2)“ behandelt. Wegen $\operatorname{Rang} F'(0) = k$ finden sich in der Jacobi-Matrix $\mathrm{D}F(0) \in \mathbb{R}^{N\times k}$ stets k linear unabhängige Zeilen, die Komponentenfunktionen von F entsprechen. Es kann ohne Einschränkung angenommen werden, dass es sich hierbei um die ersten k Zeilen bzw. Komponentenfunktionen handelt, also $(F_1, F_2, \ldots, F_k)'(0)$ invertierbar ist, und zudem $a = 0$ gilt, denn andernfalls kann man diese Eigenschaften durch Übergang von F zu $T^{-1}(F-a)$ mit einer Koordinatenpermutation $T \in \mathcal{O}(\mathbb{R}^N)$ erreichen. Unter diesen Annahmen gibt es dann nach Umkehrsatz Nullumgebungen $U' \subset U \subset \mathbb{R}^k$ und $U'' \subset \mathbb{R}^k$ sowie eine C^m-Inverse $G\colon U'' \to U'$ zu $(F_1, F_2, \ldots, F_k)\colon U' \to U''$ mit $G(0) = 0$. Da F homöomorph abbildet, ist zudem $F(U')$ relativ offen in $M \cap V$ und enthält 0, daher gibt es eine Nullumgebung $V' \subset V \subset \mathbb{R}^N$ mit $F(U') = M \cap V'$. Insgesamt folgt

$$M \cap V' = F(U') = F(G(U'')) = \{(x, (F_{k+1}, F_{k+2}, \ldots, F_N)(G(x))) : x \in U''\} = \operatorname{Graph} f$$

für die durch $f(x) := (F_{k+1}, F_{k+2}, \ldots, F_N)(G(x))$ definierte C^m-Abbildung $f\colon U'' \to \mathbb{R}^{N-k}$ mit $f(0) = 0$. Somit ist die explizite Darstellung (mit U'', V', $\mathbb{I}_N$ anstelle von U, V, T) hergeleitet.

Zum Nachweis der Implikation „(2) $\Longrightarrow$ (1)“ gibt man die C^m-Abbildung $F\colon U \to \mathbb{R}^N$ mit $F(0) = a$ durch

$$F(\xi) := a + T(\xi, f(\xi)) \qquad \text{für } \xi \in U$$

explizit an. Wegen $\mathrm{D}F = T\binom{\mathbb{I}_k}{\mathrm{D}f}$ mit der $(k\times k)$-Einheitsmatrix $\mathbb{I}_k$ hat F dann maximalen Rang, wegen $|F(\xi) - F(\xi')| = |(\xi - \xi', f(\xi) - f(\xi'))| \geq |\xi - \xi'|$ für $\xi, \xi' \in U$ bildet F homöomorph ab, und mit

$$F(U) = a + T(\operatorname{Graph} f) = M \cap V\,.$$

ist die parametrische Darstellung hergestellt.

Für die Implikation „(2)$\Longrightarrow$(4)“ definiert man die C^m-Abbildung $\Phi\colon U\times\mathbb{R}^{N-k}\to\mathbb{R}^N$ mit $\Phi(0)=a$ in ähnlich expliziter Weise durch

$$\Phi(\xi,\eta):=a+T(\xi,f(\xi)+\eta)\qquad \text{für } \xi\in U\,,\ \eta\in\mathbb{R}^{N-k}\,.$$

Da $(\xi,\eta)\mapsto(\xi,f(\xi)+\eta)$ und seine C^m-Inverse $(\xi,\eta)\mapsto(\xi,-f(\xi)+\eta)$ bijektiv von $U\times\mathbb{R}^{N-k}$ auf sich abbilden, ist Φ tatsächlich C^m-Diffeomorphismus von $U\times\mathbb{R}^{N-k}$ auf $\Omega:=a+T(U\times\mathbb{R}^{N-k})$ und insbesondere von der offenen Nullumgebung $\Sigma:=\Phi^{-1}(V\cap\Omega)\subset U\times\mathbb{R}^{N-k}$ in $\mathbb{R}^N$ auf die offene Umgebung $V\cap\Omega$ von a in $\mathbb{R}^N$. Zudem folgt

$$\Phi((\mathbb{R}^k\times\{0\})\cap\Sigma)=(a+T(\operatorname{Graph}f))\cap V\cap\Omega=M\cap V\cap\Omega\,,$$

also ergibt sich Darstellung durch Geradebiegen (mit $V\cap\Omega$ anstelle von V).

Zum Nachweis von „(4)$\Longrightarrow$(3)“ wählt man die C^m-Abbildung $g\colon V\to\mathbb{R}^{N-k}$ als die hinteren $N{-}k$ Komponentenfunktionen von $\Phi^{-1}\colon V\to\Sigma$. Aufgrund von $\mathrm{D}g=(0\ \mathbb{I}_{N-k})\mathrm{D}(\Phi^{-1})$ mit $\mathrm{D}(\Phi^{-1})\in\mathrm{GL}_N$ hat g maximalen Rang, und mit

$$g^{-1}(0)=\{x\in V:\Phi^{-1}(x)\in\mathbb{R}^k\times\{0\}\}=\Phi((\mathbb{R}^k\times\{0\})\cap\Sigma)=M\cap V$$

erhält man die implizite Darstellung.

Zur abschließenden Verifikation der Implikation „(3)$\Longrightarrow$(2)“ bemerkt man zunächst, dass $\mathrm{D}g(a)\in\mathbb{R}^{(N-k)\times N}$ wegen der Rangbedingung $N{-}k$ linear unabhängige Spalten enthält. Ähnlich wie im ersten Beweisteil kann angenommen werde, dass $a=0$ gilt und es sich um die letzten $N{-}k$ Spalten handelt, dass also $\frac{\partial g}{\partial(x_{k+1},x_{k+2},\ldots,x_N)}(0)$ invertierbar ist (denn andernfalls kann man dies durch Übergang zu $g\circ(a+T)$ mit einer Koordinatenpermutation $T\in\mathcal{O}(\mathbb{R}^N)$ erreichen). In dieser Situation liefert der Satz über implizite Funktionen Nullumgebungen $U\subset\mathbb{R}^k$ und $V'\subset V\subset\mathbb{R}^N$ sowie eine C^m-Funktion $f\colon U\to\mathbb{R}^N$ mit $\operatorname{Graph}f=g^{-1}(0)\cap V'$. Insgesamt erhält man dann mit

$$\operatorname{Graph}f=g^{-1}(0)\cap V'=M\cap V'$$

die explizite Darstellung (mit V', $\mathbb{I}_N$ anstelle von V, T). □

Bemerkungen (zu Darstellungen von Untermannigfaltigkeiten)**.**

(1) Im Kontext des vorigen Satzes ergibt sich auch, dass für zwei reguläre C^m-Parametisierungen $F\colon U\to\mathbb{R}^N$ und $\widetilde{F}\colon\widetilde{U}\to\mathbb{R}^N$ einer Untermannigfaltigkeit $M\subset\mathbb{R}^N$ nahe eines inneren Punkt $a\in M\setminus\partial M$ nach geeigneter Verkleinerung der Nullumgebungen $U,\widetilde{U}\subset\mathbb{R}^k$ der **Koordinatenwechsel** $\widetilde{F}^{-1}\circ F\colon U\to\widetilde{U}$ wohldefiniert und ein C^m-Diffeomorphismus ist.

Beweis. Sei V eine Umgebung von a in $\mathbb{R}^N$, so dass $M\cap V$ eine explizite Darstellung wie in (2) mit $T\in\mathcal{O}(\mathbb{R}^N)$ und C^m-Abbildung $f\colon U'\to\mathbb{R}^{N-k}$ hat. Nach Verkleinerung von U, $\widetilde{U}$, U' lässt sich $F(U)=F(\widetilde{U})=\operatorname{Graph}f$ annehmen. Dann ist die C^m-Abbildung $\Phi\colon U\times\mathbb{R}^{N-k}\to\mathbb{R}^N$ mit $\Phi(\xi,\eta):=F(\xi)+T(0,\eta)$ injektiv (denn aus $\Phi(\xi,\eta)=\Phi(\xi',\eta')$ folgt per Graphendarstellung $\eta=\eta'$, $F(\xi)=F(\xi')$ und dann per Injektivität von F auch $\xi=\xi'$). Zudem folgt aus der Graphendarstellung, dass Φ maximalen Rang hat (denn bei $\mathrm{D}(T^{-1}F)=\binom{Z_1}{Z_2}\in\mathbb{R}^{N\times k}$ hängen die unteren $N{-}k$ Zeilen $Z_2=\mathrm{D}f(\ldots)Z_1\in\mathbb{R}^{(N-k)\times k}$ von den oberen k Zeilen $Z_1\in\mathbb{R}^{k\times k}$ ab, die Rangbedingung für F erfordert daher $\operatorname{Rang}Z_1\equiv k$, und aus $T^{-1}\mathrm{D}\Phi=\mathrm{D}(T^{-1}\Phi)=\begin{pmatrix}Z_1&0\\Z_2&\mathbb{I}_{N-k}\end{pmatrix}\in\mathbb{R}^{N\times N}$ folgt $\operatorname{Rang}(\mathrm{D}\Phi)\equiv N$). Nach dem Umkehrsatz ist somit Φ Diffeomorphismus von $U\times\mathbb{R}^{N-k}$ auf ihr offenes Bild $a+T(U'\times\mathbb{R}^{N-k})$, und analog ergibt die Festlegung $\widetilde{\Phi}(\xi,\eta):=\widetilde{F}(\xi)+T(0,\eta)$ einen Diffeomorphismus $\widetilde{\Phi}$ von $\widetilde{U}\times\mathbb{R}^{N-k}$ auf $a+T(U'\times\mathbb{R}^{N-k})$. Damit ist $\widetilde{\Phi}^{-1}\circ\Phi$ Diffeomorphismus von $U\times\mathbb{R}^{N-k}$ auf $\widetilde{U}\times\mathbb{R}^{N-k}$. Da nach Konstruktion $(\widetilde{\Phi}^{-1}\circ\Phi)(\xi,\eta)=((\widetilde{F}^{-1}\circ F)(\xi),\eta)$ gilt, folgt hieraus, dass $\widetilde{F}^{-1}\circ F$ Diffeomorphismus von U auf $\widetilde{U}$ ist. □

(2) Für **Randpunkte** einer Untermannigfaltigkeit bestehen **größtenteils analoge Charakterisierungen**. Genauer sind für $k \le N$ in $\mathbb{N}$, $m \in \mathbb{N} \cup \{\infty\}$, $M \subset \mathbb{R}^N$, $a \in M$ **äquivalent**:

(1) Es gibt (wie in der Definition) offene Umgebungen U von 0 in $\mathbb{R}^k$ und V von a in $\mathbb{R}^N$ sowie eine C^m-Abbildung $F\colon U \to \mathbb{R}^N$ von maximalem Rang mit $F(0) = a$, so dass F homöomorph von $U_{\le 0}$ auf
$$M \cap V = F(U_{\le 0})$$
abbildet.

(2) Es gibt offene Umgebungen U von 0 in $\mathbb{R}^k$ und V von a in $\mathbb{R}^N$, eine lineare Isometrie $T \in \mathcal{O}(\mathbb{R}^N)$ sowie eine C^m-Abbildung $f\colon U \to \mathbb{R}^{N-k}$ mit $f(0) = 0$ und
$$M \cap V = a + T\big(\mathrm{Graph}\big(f|_{U_0}\big)\big)$$
für eine k-dimensionale C^m-Untermannigfaltigkeit $U_0 \subset U$ mit Rand des $\mathbb{R}^k$ mit $0 \in \partial U_0$.

(3) Es gibt eine offene Umgebung V von a in $\mathbb{R}^N$ und eine C^m-Abbildung $g\colon V \to \mathbb{R}^{1+N-k}$ von maximalem Rang mit
$$M \cap V = g^{-1}(\mathbb{R}_{\le 0} \times \{0\})$$
($M \cap V$ ist gemeinsame Lösungsmenge zu einer Ungleichung und $N{-}k$ Gleichungen).

(4) Es gibt offene Umgebungen Σ von 0 in $\mathbb{R}^N$ und V von a in $\mathbb{R}^N$ sowie einen C^m-Diffeomorphismus $\Phi\colon \Sigma \to V$ von Σ auf V mit $\Phi(0) = a$ und
$$M \cap V = \Phi(((\mathbb{R}^k)_{\le 0} \times \{0\}) \cap \Sigma)\,.$$

Zum Beweis. Über weite Strecken lässt sich analog zum inneren Fall argumentieren. Unterschiede und Modifikationen der Argumentation werden nun kurz angedeutet:

Für „(1) $\Longrightarrow$ (2)“ ist zu beachten, dass G zwar U'' auf U', aber $(U'')_{\le 0}$ nicht unbedingt exakt auf $(U')_{\le 0}$ abbildet. Deshalb ergibt sich die Graphendarstellung im Allgemeinen nicht über $(U'')_{\le 0}$, sondern nur über der k-dimensionalen Untermannigfaltigkeit $U_0 := G^{-1}((U')_{\le 0}) \subset U'' \subset \mathbb{R}^k$.

Bei „(2) $\Longrightarrow$ (1)“ und „(2) $\Longrightarrow$ (4)“ gilt es zusätzlich zu benutzen, dass die k-dimensionale Untermannigfaltigkeit $U_0 \subset \mathbb{R}^k$ nahe $0 \in \partial U_0$ als Bild $F_0((U_0')_{\le 0})$ unter einer regulären Parametrisierung F_0 dargestellt ist.

Bei „(4) $\Longrightarrow$ (3)“ kombiniert man zur Definition von g die erste Komponentenfunktion von Φ^{-1} mit den auch im inneren Fall genutzten hinteren $N{-}k$ Komponentenfunktionen von Φ^{-1}.

Lediglich bei „(3) $\Longrightarrow$ (2)“ kann man sich nicht ohne Weiteres an den inneren Fall anlehnen. Statt den Nachweis dieser Implikation zu führen, kann man den Beweis aber mit folgenden beiden Schritten komplettieren.

Man zeigt „(3) $\Longrightarrow$ (4)“, indem man nach geeigneter Koordinatenpermutation aus g durch Ergänzen trivialer Komponentenfunktionen einen Diffeomorphismus erzeugt, dessen Umkehrabbildung die Eigenschaften von Φ hat. (Das Argument ähnelt einer bekannten Herleitung des Satzes über implizite Funktionen aus dem Umkehrsatz.)

Schließlich gewinnt man die für „(4) $\Longrightarrow$ (1)“ benötigte reguläre Parametrisierung $F\colon U_{\le 0} \to \mathbb{R}^N$ einfach durch $F(\xi) := \Phi(\xi, 0)$, wobei $U = \Sigma_0 \subset \mathbb{R}^k$ ein Schnitt von $\Sigma \subset \mathbb{R}^N$ ist, mit anderen Worten also $U \times \{0\} = (\mathbb{R}^k \times \{0\}) \cap \Sigma$. □

Zentrale Konzepte im Zusammenhang mit (Unter-)Mannigfaltigkeiten sind Tangentialkegel und Tangentialräume. Diese Bildungen beschreiben das Erster-Ordnung-Verhalten nahe eines gegebenen Punkts und werden nun eingeführt und betrachtet.

Satz (über **Tangentialverhalten von Untermannigfaltigkeiten**). *Seien $k, N \in \mathbb{N}$ mit $k \le N$ und M eine k-dimensionale C^1-Untermannigfaltigkeit mit Rand des $\mathbb{R}^N$. Dann ist der Tangentialkegel*
$$\mathrm{Tan}(M,a) := \left\{ \lim_{\ell\to\infty} \frac{x_\ell - a}{r_\ell} \;:\; (x_\ell)_{\ell\in\mathbb{N}} \text{ Folge in } M \text{ mit } \lim_{\ell\to\infty} x_\ell = a\,,\ (r_\ell)_{\ell\in\mathbb{N}} \text{ Folge in } \mathbb{R}_{>0} \right\}$$

für geometrisch innere Punkte $a \in M \setminus \partial M$ *ein* k*-dimensionaler Unterraum von* $\mathbb{R}^N$*, der bei Vorliegen der Darstellungen aus dem vorigen Satz durch die Formeln*

$$\begin{aligned}\mathrm{Tan}(M,a) &= \mathrm{Bild}(F'(0)) = T(\mathrm{Graph}(f'(0))) = \mathrm{Kern}(g'(a)) = \Phi'(0)(\mathbb{R}^k \times \{0\}) \\ &= \{c'(0) : c \textit{ ist } \mathrm{C}^1\textit{-Kurve in } \mathbb{R}^N \textit{ mit } c(0)=a\,,\ c(t) \in M \textit{ für } |t| \ll 1\}\end{aligned}$$

(*mit* $\mathrm{Graph}(f'(0)) = \{(v, \partial_v f(0)) : v \in \mathbb{R}^k\}$) *gegeben ist. Für geometrische Randpunkte* $a \in \partial M$ *ist* $\mathrm{Tan}(M,a)$ *dagegen ein* k*-dimensionaler Halbraum*[2] *in* $\mathbb{R}^N$*, der bei Vorliegen der Darstellungen aus der vorigen Bemerkung* (2) *durch die Formeln*

$$\begin{aligned}\mathrm{Tan}(M,a) &= F'(0)((\mathbb{R}^k)_{\le 0}) = \{T(v, \partial_v f(0)) : v \in \mathrm{Tan}(U_0, 0)\} \\ &= \{w \in \mathbb{R}^N : g'(a)(w) \in \mathbb{R}_{\le 0} \times \{0\}\} = \Phi'(0)((\mathbb{R}^k)_{\le 0} \times \{0\}) \\ &= \{c'(0) : c \textit{ ist } \mathrm{C}^1\textit{-Kurve in } \mathbb{R}^N \textit{ mit } c(0)=a\,,\ c(t) \in M \textit{ für } 0 \le t \ll 1\}\end{aligned}$$

gegeben ist.

Der *Beweis* mit Methoden der mehrdimensionalen Differentialrechnung wird zumindest für den inneren Fall in den Übungen behandelt.

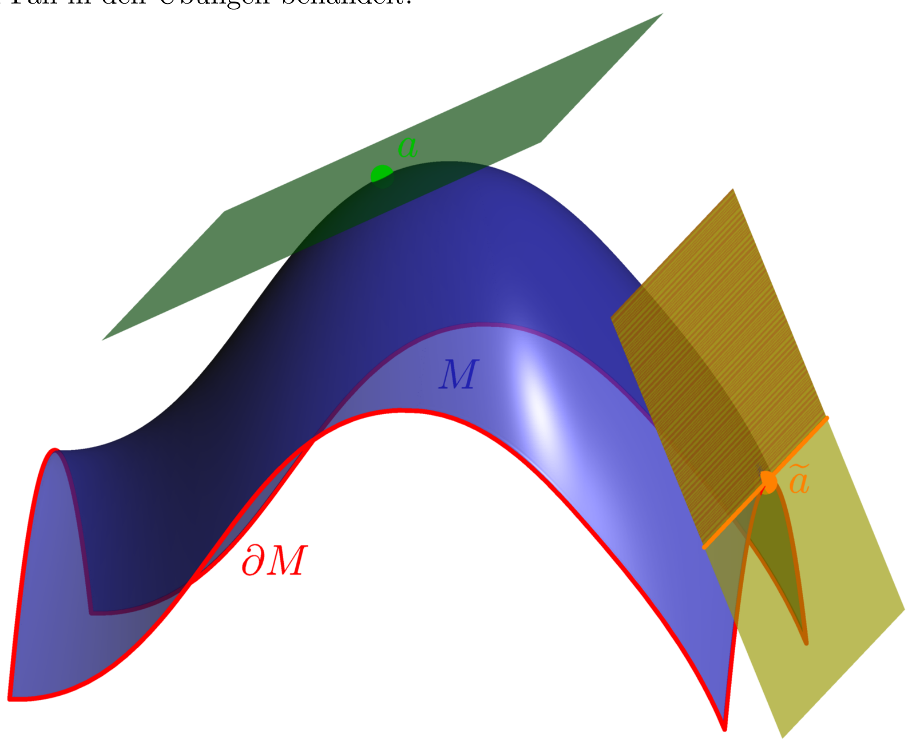


Abb. 21: Ein innerer Punkt $a \in M \setminus \partial M$ mit einem Stück der affinen Tangentialebene $a+\mathrm{T}_a M = a+\mathrm{Tan}(M,a)$ (grünes Ebenenstück) an eine 2-dimensionale Untermannigfaltigkeit M sowie ein Randpunkt $\widetilde{a} \in \partial M$ mit Stücken des affinen Tangentialkegels $\widetilde{a}+\mathrm{Tan}(M,\widetilde{a})$ (orange-gelbes Halbebenenstück), des affinen Tangentialraums $\widetilde{a}+\mathrm{T}_{\widetilde{a}}M$ (gelbes Ebenenstück) und des affinen Tangentialraums $\widetilde{a}+\mathrm{T}_{\widetilde{a}}\partial M$ an den Rand ∂M (oranges Geradenstück)

[2] Unter einem k-dimensionalen Halbraum H in $\mathbb{R}^N$ ist die „Hälfte" eines k-dimensionalen Unterraums von $\mathbb{R}^N$ „auf einer Seite" eines enthaltenen $(k-1)$-dimensionalen Unterraums zu verstehen. Präziser (aber vielleicht weniger eingängig) definiert man die Halbraumeigenschaft von H über die Existenz einer linearen Parametrisierung $H = L((\mathbb{R}^k)_{\le 0})$ mit $L \in \mathbb{R}^{N\times k}$, $\mathrm{Rang}\, L = k$ oder alternativ die Beschreibung als Lösungsmenge eines linearen Ungleichung-Gleichungen-Systems $H = \{x \in \mathbb{R}^N : Ax \in \mathbb{R}_{\le 0} \times \{0\}\}$ mit $A \in \mathbb{R}^{(1+N-k)\times N}$, $\mathrm{Rang}\, A = 1+N-k$.

Definition (**Tangentialraum**). *Für M wie im vorigen Satz und $a \in M$ bezeichnet man den k-dimensionalen Unterraum*[3]

$$\mathrm{T}_a M := \begin{cases} \mathrm{Tan}(M,a) & \text{falls } a \in M \setminus \partial M \\ \mathrm{Span}(\mathrm{Tan}(M,a)) & \text{falls } a \in \partial M \end{cases}$$

*von $\mathbb{R}^N$ als den **Tangentialraum** an M in a und den k-dimensionalen affinen Unterraum $a+\mathrm{T}_a M$ von $\mathbb{R}^N$ als den **affinen Tangentialraum** an M in a.*

Tatsächlich erweist es sich bei Abbildungen zwischen (Unter-)Mannigfaltigkeiten als sehr natürlich, ihre Ableitungen zwischen Tangentialräumen operieren zu lassen. Weiteres hierzu fällt in den Bereich der Differentialgeometrie, hier wird das Konzept nur kurz angerissen:

Ausblick (**Differentiation von Abbildungen zwischen (Unter-)Mannigfaltigkeiten**). Seien M eine k-dimensionale C^1-Untermannigfaltigkeit mit Rand des $\mathbb{R}^N$ und $\widetilde{M}$ eine $\widetilde{k}$-dimensionale C^1-Untermannigfaltigkeit mit Rand des $\mathbb{R}^{\widetilde{N}}$, und sei $F\colon U \to \mathbb{R}^N$ eine reguläre C^1-Parametrisierung von M nahe $a = F(0) \in M$. Dann nennt man f total differenzierbar in a, wenn $f \circ F$ total differenzierbar in 0 ist, und erklärt die **Ableitung** $f'(a) \in \mathcal{L}(\mathrm{T}_a M, \mathrm{T}_{f(a)}\widetilde{M})$ als **lineare Abbildung zwischen Tangentialräumen** durch Postulieren der Kettenregel, genauer durch $f'(a)(F'(0)(v)) := (f \circ F)'(0)(v)$ für $v \in \mathbb{R}^k$ (beachte $F'(0)(v) \in \mathrm{T}_a M$ und $(f \circ F)'(0)(v) \in \mathrm{T}_{f(a)}\widetilde{M}$) oder äquivalent durch $f'(a)(c'(0)) := (f \circ c)'(0)$ für C^1-Kurven wie im vorigen Satz (beachte $c'(0) \in \mathrm{T}_a M$ und $(f \circ c)'(0) \in \mathrm{T}_{f(a)}\widetilde{M}$). Mit der Diffeomorphismus-Eigenschaft der Koordinatenwechsel zeigt man, dass diese Konzepte insofern geometrisch sinnvoll sind, dass sie nicht von der Wahl der Parametrisierung F, sondern nur von M, f und a abhängen.

Schließlich wird zur Vorbereitung auf den späteren Abschnitt 3.4 die Orientierung von Vektorräumen und Untermannigfaltigkeiten eingeführt.

Definitionen & Bemerkungen (**Orientierung von Vektorräumen**). *Seien $k \in \mathbb{N}$ und V ein k-dimensionaler $\mathbb{R}$-Vektorraum.*

(I) *Zwei Basen $b_1, b_2, \ldots, b_k$ und $\widetilde{b}_1, \widetilde{b}_2, \ldots, \widetilde{b}_k$ von V heißen **gleichsinnig orientiert**, wenn $\det(T) > 0$ für den Basiswechsel $T \in \mathrm{GL}(V)$ mit $Tb_i = \widetilde{b}_i$ für $i = 1, 2, \ldots, k$ gilt. Man spricht in dieser Situation auch vom **Orientierung-erhaltenden Basiswechsel** T. Andernfalls, wenn also $\det(T) < 0$ gilt, heißen die Basen **gegensinnig orientiert** und T **Orientierung-umkehrender Basiswechsel**. (Man beachte, dass diese Begriffe nur für geordnete Basen sinnvoll sind, d.h. es kommt nicht nur auf die Menge der Basisvektoren, sondern auch auf die Reihenfolge, in der diese Vektoren aufgezählt werden, an. Deshalb wird unter einer Basis hier und im Folgenden stets eine geordnete Basis mit fixierter Reihenfolge verstanden.)*

(II) *Man prüft recht problemlos, dass gleichsinnige Orientierung eine **Äquivalenzrelation auf der Menge aller Basen** von V ist. Weil die Komposition zweier Orientierung-umkehrender Basiswechsel einen Orientierung-erhaltenden Basiswechsel ergibt, zerfällt die Menge aller Basen von V bezüglich dieser Äquivalenzrelation in **genau zwei Äquivalenzklassen**. Unter einer **Orientierung von V** versteht man die Auswahl einer dieser beiden Äquivalenzklassen, und bezüglich einer fixierten Orientierung bezeichnet man die Basen der gewählten Äquivalenzklasse als **positiv orientierte Basen**, die der anderen*

[3]Äquivalent kann man $\mathrm{T}_a M := \mathrm{Span}(\mathrm{Tan}(M,a))$ für alle $a \in M$ setzen und so obige Fallunterscheidung vermeiden. Oben soll aber insbesondere deutlich werden, dass für $a \in M \setminus \partial M$ einfach $\mathrm{T}_a M = \mathrm{Tan}(M,a)$ ist.

*Äquivalenzklasse als **negativ orientierte Basen**.*

(III) *Die **kanonische Orientierung von** $\mathbb{R}^k$ ist die (Auswahl der) Äquivalenzklasse von Basen von $\mathbb{R}^k$, die die kanonische Basis von $\mathbb{R}^k$ enthält. Mit anderen Worten ist die kanonische Basis von $\mathbb{R}^k$ bezüglich der kanonischen Orientierung von $\mathbb{R}^k$ positiv orientiert.*

Bemerkung (zu **Orientierung in Dimension 0**)**.** Als ergänzende Konvention versteht man unter einer Orientierung des Nullvektorraums die Auswahl einer der Zahlen 1, -1. Dies wird sich etwa für 0-dimensionale Ränder 1-dimensionaler Untermannigfaltigkeiten als sinnvoll erweisen.

Definitionen (**Orientierung von Untermannigfaltigkeiten**)**.** *Seien $k, N \in \mathbb{N}$ mit $k \le N$ und M eine k-dimensionale C^1-Untermannigfaltigkeit mit Rand des $\mathbb{R}^N$.*

(I) *Eine **Orientierung von M** ist eine Auswahl von Orientierungen der Tangentialräume T_aM in allen Punkten $a \in M$, so dass folgende **Kohärenzbedingung** erfüllt ist: Für jedes $a \in M$ gibt es eine reguläre C^1-Parametrisierung $F \colon U \to M$ von M nahe a, für die für jedes $\xi \in U$ die Basis $\partial_1 F(\xi), \partial_2 F(\xi), \ldots, \partial_k F(\xi)$ von $\mathrm{T}_{F(\xi)}M$ positiv orientiert ist. Man spricht in dieser Situation auch von positiv orientierten Parametrisierungen F und bezeichnet M samt einer gewählten Orientierung von M als **orientierte Untermannigfaltigkeit**.*

(II) *Ist M eine orientierte Untermannigfaltigkeit, so wird die **induzierte Orientierung des Randes ∂M** für $k \ge 2$ so festgelegt, dass für alle $a \in \partial M$ und alle Basen $b_1, \ldots, b_{k-1}$ von $\mathrm{T}_a\partial M$ gilt:*

$$\begin{aligned}
&b_1, \ldots, b_{k-1} \text{ positiv orientierte Basis in } \mathrm{T}_a\partial M \\
\iff\ &\nu, b_1, \ldots, b_{k-1} \text{ positiv orientierte Basis in } \mathrm{T}_aM \text{ für ein/jedes } \nu \in \mathrm{T}_aM \setminus \mathrm{Tan}(M,a)
\end{aligned}$$

(Dies ist aufgrund folgender Tatsachen sinnvoll: Wegen $\mathrm{T}_a\partial M \subset \mathrm{T}_aM$ und der Dimensionen wird jede Basis $b_1, \ldots, b_{k-1}$ von $\mathrm{T}_a\partial M$ durch jeden Vektor $\nu \in \mathrm{T}_aM \setminus \mathrm{T}_a\partial M$ zu einer Basis $\nu, b_1, \ldots, b_{k-1}$ von T_aM ergänzt, und die Orientierung von $\nu, b_1, \ldots, b_{k-1}$ hängt für gegebene $b_1, \ldots, b_{k-1}$ nur davon ab, ob $\nu \in \mathrm{T}_aM \setminus \mathrm{Tan}(M,a)$ äußerer Tangentialvektor oder $\nu \in \mathrm{Tan}(M,a) \setminus \mathrm{T}_a\partial M$ innerer Tangentialvektor ist.)

Für $k = 1$ legt man die Orientierung des Randes ergänzend fest durch:

$$\begin{aligned}
&\text{Orientierung } 1 \text{ von } \mathrm{T}_a\partial M \\
&\qquad \iff \text{ein/jedes } \nu \in \mathrm{T}_aM \setminus \mathrm{Tan}(M,a) \text{ positiv orientierte Basis in } \mathrm{T}_aM\,.
\end{aligned}$$

Interpretation (von Orientierungen)**.**

(1) Für **1-dimensionale Untermannigfaltigkeiten** lässt sich Orientierung, wie in Abbildung 22 illustriert, als **Umlaufsinn** interpretieren.

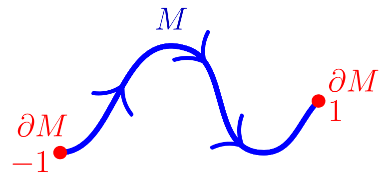


Abb. 22: Eine orientierte 1-dimensionale Untermannigfaltigkeit M in $\mathbb{R}^2$ und ihr 0-dimensionaler Rand ∂M mit der induzierten Orientierung

(2) Für **2-dimensionale Untermannigfaltigkeiten** lässt sich Orientierung, wie in Abbildung 23 illustriert, als **Drehsinn** interpretieren.

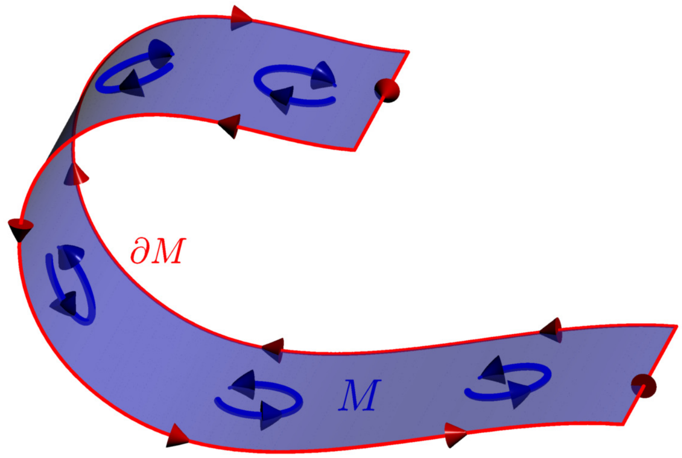


Abb. 23: Eine orientierte 2-dimensionale Untermannigfaltigkeit M in $\mathbb{R}^3$ und ihr 1-dimensionaler Rand ∂M mit der induzierten Orientierung

(3) Für **3-dimensionale Untermannigfaltigkeiten** entspricht Orientierung einem im Tangentialraum an jeden Punkt der Untermannigfaltigkeit festzulegenden **Schraubsinn**. Abbildung 24 illustriert zumindest die beiden möglichen Schraubsinne in $\mathbb{R}^3$. Dabei gehen die beiden linken Bilder durch einen Orientierung-erhaltenden Basiswechsel, nämlich eine Drehung, ineinander über und repräsentieren daher denselben Schraubsinn. Analoges verhält es sich mit den beiden rechten Bildern, die den gegenüber den linken Bildern gegensinnigen Schraubsinn repräsentieren.

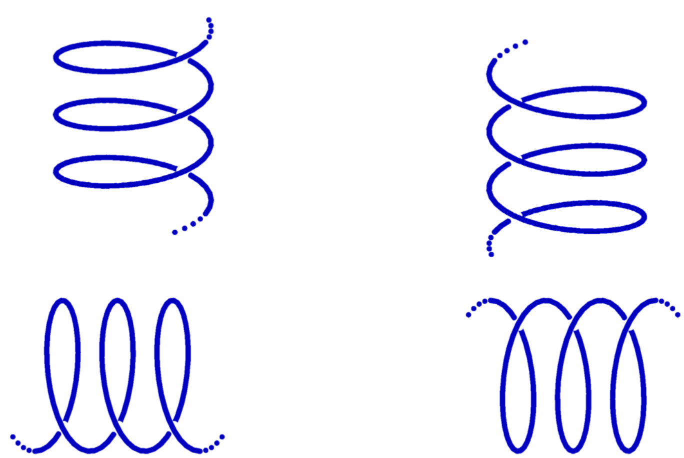


Abb. 24: Schraubsinne in $\mathbb{R}^3$

Bemerkung (zu Nicht-Orientierbarkeit)**.** Es gibt ***nicht*** **orientierbare Untermannigfaltigkeiten** wie das in Abbildung 25 gezeigte **Möbiusband**

$$M := \big\{\big((5+t\cos(\varphi/2))\cos\varphi\,,\ (5+t\cos(\varphi/2))\sin\varphi\,,\ t\sin(\varphi/2)\big) : \varphi \in \mathbb{R}\,,\ t \in [-1,1]\big\} \subset \mathbb{R}^3\,.$$

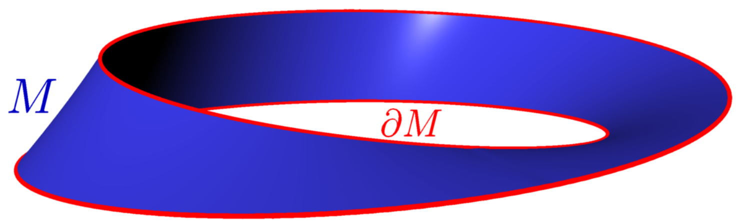


Abb. 25: Das Möbiusband M

Literaturverzeichnis

Die vorliegende Ausarbeitung basiert auf einem Vorlesungsskript von K. Steffen, einem älteren Maßtheorie-Skript des Autors und unter den folgenden Büchern vor allem auf der Quelle [3]:

[1] H. Amann, J. Escher, *Analysis III.* Birkhäuser, 2009.

[2] H. Bauer, *Maß- und Integrationstheorie.* De Gruyter, 1990.

[3] J. Elstrodt, *Maß- und Integrationstheorie.* Springer, 1996.

[4] K. Falconer, *Fractal Geometry. Mathematical Foundations and Applications. Second Edition.* Wiley, 2003.

[5] H. Federer, *Geometric Measure Theory.* Springer, 1969.

[6] O. Forster, *Analysis 3.* Vieweg+Teubner, 2012.

[7] P. Mattila, *Geometry of Sets and Measures in Euclidean Space. Fractals and Rectifiability.* Cambridges University Press, 1995.

[8] W. Rudin, *Real and Complex Analysis.* McGraw-Hill, 1987.