

Analysis 3

Wintersemester 2023/24

Gandalf Lechner¹

Gegenstand dieser Vorlesung ist eine systematische Entwicklung der Maß- und Integrations-
theorie. Insbesondere werden das Lebesgue-Maß und L^p -Räume zusammen mit Anwendun-
gen in der Fourieranalysis besprochen. Weitere Themen sind Differentialformen und die In-
tegralsätze von Gauß und Stokes.



¹Department Mathematik, FAU Erlangen-Nürnberg, Cauerstr. 11 Erlangen,
gandalf.lechner@fau.de

Inhaltsverzeichnis

1	Mengensysteme und Maße	1
1.1	Motivation und Einführung: Das Inhalts- und Maßproblem	1
1.2	σ -Algebren und andere Mengensysteme	7
1.3	Inhalte, Prämaße, und Maße	12
1.4	Fortsetzungen von Prämaßen zu Maßen	20
1.5	Das Lebesgue-Maß und messbare Funktionen	28
2	Integration messbarer Funktionen	35
2.1	Mehr über messbare Funktionen	35
2.2	Integrale nach einem Maß und integrierbare Funktionen	39
2.3	Nullmengen und Vergleich Lebesgue-/Riemannintegral	45
2.4	Majorisierte Konvergenz und Anwendungen	49
2.5	L^p -Räume	54
2.6	Produktmaße und der Satz von Fubini	61
2.7	Die Transformationsformel	72
3	Einführung in die Fourier-Analysis	80
3.1	Definition der Fouriertransformation und erste Eigenschaften	80
3.2	Umkehrabbildung und Anwendungen	86
3.3	Die Fouriertransformation auf L^2	91
4	Differentialformen und Integration über Untermannigfaltigkeiten	97
4.1	Alternierende Multilinearformen	99
4.2	Differentialformen	104
4.3	Integration von Differentialformen über orientierte Untermannigfaltigkeiten	112
4.4	Der Satz von Stokes	123
4.5	Maßsensoren und Integration von Funktionen über Untermannigfaltigkeiten	133

1 Mengensysteme und Maße

1.1 Motivation und Einführung: Das Inhalts- und Maßproblem

In diesem einführenden Abschnitt werden die Fragen und Begriffe der Maßtheorie motiviert und eingeführt. Um eine erste Vorstellung von der Thematik zu haben, denken wir an *Volumenmessungen*: Für Teilmengen $A \subset \mathbb{R}^d$ möchten wir erklären, was unter einem “ d -dimensionalen Volumen” $\nu(A)$ zu verstehen ist. Wir werden dann auch motivieren, wie diese eher geometrische Fragestellung mit Anwendungen in Integrations-, Wahrscheinlichkeits-, Spektral-, und Quantenfeldtheorie verknüpft ist, was die wichtigsten Anwendungsgebiete der Maßtheorie sind.

Beginnen wir mit Überlegungen zu “ d -dimensionalem Volumen”. Damit meinen wir für $d = 1$ die “Länge” von $A \subset \mathbb{R}$, für $d = 2$ die “Fläche” von $A \subset \mathbb{R}^2$, für $d = 3$ das “Volumen” von $A \subset \mathbb{R}^3$, und für $d > 3$ eine geeignete höherdimensionale Verallgemeinerung.

Im Fall $d = 1$ sagt uns unsere geometrische Anschauung, dass ein Intervall $A = [a, b]$, $a \leq b$, die Länge $\nu(A) = b - a$ haben sollte. Weitere Beispiele, bei denen uns geometrische Anschauung eine Antwort nahelegt, sind

- $A = [0, 1] \cup [2, 4] \Rightarrow \nu(A) = 3$
- $A = [a_1, b_1] \times [a_2, b_2] \subset \mathbb{R}^2 \Rightarrow \nu(A) = (b_1 - a_1) \cdot (b_2 - a_2)$ für $a_1 \leq b_1, a_2 \leq b_2$
- $A = \{x \in \mathbb{R}^3 : \|x\|_2 \leq r\} \Rightarrow \nu(A) = \frac{4}{3}\pi r^3$
- $A = \{(x, y) \in \mathbb{R}^2 : a \leq x \leq b, 0 \leq y \leq x^2 e^x\} \Rightarrow \nu(A) = \int_a^b x^2 e^x dx$
- $A = \emptyset \Rightarrow \nu(A) = 0$
- $A = \{x\} \subset \mathbb{R} \Rightarrow \nu(A) = 0$
- $A = [0, \infty) \subset \mathbb{R} \Rightarrow \nu(A) = +\infty$

Aber wie sieht es mit dem Volumen einer *allgemeinen* Teilmenge $A \subset \mathbb{R}^d$ aus? Es stellen sich verschiedene Fragen: Wie soll der Begriff “Volumen” überhaupt streng definiert werden? Gibt es mehr als einen sinnvollen Volumenbegriff? Wie lässt sich das Volumen einer gegebenen Menge bestimmen? Was ist z.B. das Volumen von $[0, 1] \cap \mathbb{Q} \subset \mathbb{R}$ oder das Volumen der Cantormenge?

Als Antwort auf die erste Frage machen wir nun den Versuch einer Definition. Dabei lassen wir uns von folgenden Überlegungen leiten:

- a) Das Volumen einer Teilmenge $A \subset \mathbb{R}^d$ sollte 0, eine positive reelle Zahl, oder $+\infty$ sein, aber nie negativ oder $-\infty$.
- b) Das Volumen der leeren Menge sollte Null sein, $\nu(\emptyset) = 0$.
- c) *Endliche Additivität*: Sind $A, B \subset \mathbb{R}^d$ disjunkt, so sollte das Volumen der Vereinigung $A \cup B$ gleich der Summe der Volumina von A und B sein, d.h.

$$\nu(A \cup B) = \nu(A) + \nu(B), \quad A \cap B = \emptyset.$$

Disjunkte Vereinigungen werden uns in der Analysis 3 oft begegnen. Manche Autoren schreiben $\dot{\cup}$ oder \sqcup statt \cup , um anzudeuten, dass eine Vereinigung disjunkt ist.

d) *Bewegungsinvarianz*: Ist $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}^d$ eine Bewegung², so sollte für alle $A \subset \mathbb{R}^d$

$$\nu(\varphi(A)) = \nu(A)$$

gelten.

e) *Normierung*: Das Volumen des Einheitswürfels sollte 1 sein, d.h.

$$\nu([0, 1]^d) = 1.$$

Einige Bemerkungen zu diesen sehr natürlich erscheinenden Annahmen:

- Falls $\nu(\emptyset) \in \mathbb{R}$ (also $\nu(\emptyset) < \infty$), folgt b) aus c): $\nu(\emptyset) = \nu(\emptyset \cup \emptyset) = \nu(\emptyset) + \nu(\emptyset) = 2\nu(\emptyset)$.
- Per Induktion lässt sich c) leicht auf endlich viele paarweise disjunkte Teilmengen verallgemeinern: Sind $A_1, \dots, A_n \subset \mathbb{R}^d$ und $A_k \cap A_l = \emptyset$ für $k \neq l$, so gilt

$$\nu(A_1 \cup \dots \cup A_n) = \sum_{k=1}^n \nu(A_k).$$

Dies erklärt die Bezeichnung “endliche Additivität”.

- Da “Volumina” auch unendlich sein können, werden wir es mit Abbildungen in eine um $+\infty$ erweiterte Zahlengerade zu tun haben. Wir benutzen im ganzen Semester die Bezeichnungen

$$\begin{aligned} \overline{\mathbb{R}} &:= \mathbb{R} \cup \{\pm\infty\} = [-\infty, \infty], \\ \overline{\mathbb{R}}_+ &:= \mathbb{R} \cup \{+\infty\} = [0, \infty]. \end{aligned}$$

Die Ordnung von \mathbb{R} wird naheliegenderweise durch $-\infty < x < \infty$ für alle $x \in \mathbb{R}$ auf $\overline{\mathbb{R}}$ fortgesetzt. Bzgl. Addition und Multiplikation verwenden wir die folgenden Konventionen (die z.B. bei der endlichen Additivität relevant sind, wenn einige $\nu(A_k)$ unendlich sind):

$$\begin{aligned} x + (\pm\infty) &= \pm\infty + x = \pm\infty, & x \in \mathbb{R}, \\ x \cdot (\pm\infty) &= (\pm\infty) \cdot x = \pm\infty, & x \in [0, \infty], \\ x \cdot (\pm\infty) &= (\pm\infty) \cdot x = \mp\infty, & x \in [-\infty, 0], \\ 0 \cdot (\pm\infty) &= (\pm\infty) \cdot 0 = 0. \end{aligned}$$

Ausdrücke der Art $\infty + (-\infty)$ sind nicht definiert.

- Wenn Sie sich vorstellen, die Fläche einer Kreisscheibe durch Überdeckung mit vielen paarweise disjunkten Rechtecken zu bestimmen, sehen Sie schnell, dass endlich viele Rechtecke nicht genug sind. Wir werden deshalb häufig eine verschärfte Form von Additivität betrachten, die sogenannte σ -Additivität:

²Siehe Lineare Algebra 2: Eine Bewegung ist eine abstandserhaltene Abbildung, d.h. genau eine Abbildung der Form $\varphi(x) = Ux + a$ mit $a \in \mathbb{R}^d$ und $U \in O(d)$.

Eine Abbildung $\nu : \mathcal{P}(\mathbb{R}^d) \rightarrow [0, \infty]$ heißt σ -*additiv*³, falls für jede Folge $A_k \subset \mathbb{R}^d$, $k \in \mathbb{N}$, von paarweise disjunkten Teilmengen⁴

$$\nu \left(\bigcup_{k=1}^{\infty} A_k \right) = \sum_{k=1}^{\infty} \nu(A_k). \quad (1.1)$$

Forderung d) entspricht der Anschauung, dass Verschiebungen, Drehungen, und Spiegelungen Volumina nicht ändern sollten. Forderung e) erscheint notwendig, um ein eindeutiges Volumen zu erhalten, denn ansonsten genügt mit ν auch jedes positive Vielfache von ν unseren Forderungen. Verzichten wir auf d) und e), so spricht man von einem Inhalt.

Definition 1.1. Ein *Inhalt auf* $\mathcal{P}(\mathbb{R}^d)$ ist eine Abbildung $\nu : \mathcal{P}(\mathbb{R}^d) \rightarrow [0, \infty]$, die $\nu(\emptyset) = 0$ und endliche Additivität erfüllt. Ein *Maß auf* $\mathcal{P}(\mathbb{R}^d)$ ist eine Abbildung $\mu : \mathcal{P}(\mathbb{R}^d) \rightarrow [0, \infty]$, die $\mu(\emptyset) = 0$ und σ -Additivität erfüllt.

In diesem Zusammenhang gibt es folgenden erstaunlichen Satz:

Satz 1.2.

- a) Für $d \geq 3$ gibt es keinen normierten bewegungsinvarianten Inhalt auf $\mathcal{P}(\mathbb{R}^d)$.
- b) Für $d \in \{1, 2\}$ gibt es normierte bewegungsinvariante Inhalte auf $\mathcal{P}(\mathbb{R}^d)$, aber diese Inhalte sind nicht eindeutig.
- c) Für beliebiges $d \in \mathbb{N}$ gibt es kein normiertes bewegungsinvariantes Maß auf $\mathcal{P}(\mathbb{R}^d)$.

Wir werden hier keinen Beweis dieses Satzes geben (Teil c) wird später in der Vorlesung bewiesen). Um die Aussage besser zu illustrieren, erwähnen wir noch den folgenden Satz:

Satz 1.3 (Banach-Tarski Paradoxon). Es gibt eine Zerlegung der Kugel $K := \{x \in \mathbb{R}^3 : \|x\|_2 \leq 1\}$ in fünf disjunkte Teilmengen $K_1, \dots, K_5 \subset K$, die durch Bewegungen des \mathbb{R}^3 zu zwei disjunkten Kopien von K zusammengesetzt werden können, d.h.

$$\bigcup_{i=1}^5 \varphi_i(K_i) = K \cup (K + (3, 0, 0)),$$

mit Bewegungen $\varphi_1, \dots, \varphi_5$.

Dieses überraschende Resultat erscheint wirklich als Paradoxon, denn scheinbar lässt sich durch Zerschneiden, Verschieben und Drehen das Volumen einer Kugel verdoppeln! Es sei angemerkt, dass die Teilmengen K_i sehr kompliziert sind und sich tatsächlich gar nicht explizit angeben lassen. Mit Hilfe des Auswahlaxioms lässt sich allerdings ihre Existenz beweisen.

³Im Laufe Ihres Studiums werden Ihnen einige mathematische Objekte bzw. Eigenschaften begegnen, die ein “ σ ” im Namen tragen (z.B. σ -Algebra). Das σ deutet dabei stets auf eine Kompatibilität mit abzählbarer Unendlichkeit hin, meist abzählbare Vereinigungen. Historisch wurde σ als Abkürzung für “Summe” eingeführt, da die Vereinigung von zwei Mengen früher als ihre Summe bezeichnet wurde.

⁴Würden wir Additivität nicht nur für *abzählbare*, sondern sogar für *beliebige* disjunkte Vereinigungen fordern, so würden wir die meisten interessanten Fälle ausschließen: Wir erwarten zB, dass die Einpunktmengen $\{x\}$, $x \in \mathbb{R}^d$, Volumen $\nu(\{x\}) = 0$ haben sollten. Aber jede Teilmenge $A \subset \mathbb{R}^d$ erfüllt $A = \bigcup_{x \in A} \{x\}$, so dass Additivität für beliebige Vereinigungen auf $\nu(A) = 0$ für alle A führen würde.

Sie sollten sich die K_i als fraktale Gebilde / Punktwolken vorstellen. Die Schlussfolgerung aus diesem Paradoxon (das die Aussagen a) und c) von Satz 1.2 für $d = 3$ impliziert) ist, dass es unmöglich ist, *jeder* Teilmenge von \mathbb{R}^d ein Volumen in einer unserer geometrischen Anschauung entsprechenden Art und Weise zuzuordnen. Wir werden deshalb beim Aufbau der Maßtheorie zunächst einige Zeit damit verbringen, Mengen von nicht zu irregulären als messbar definierten Mengen zu diskutieren, die echte Teilmengen von $\mathcal{P}(\mathbb{R}^n)$ sind. Diese Aufgabe werden im nächsten Abschnitt beginnen.

Bis jetzt haben wir noch kein explizites Beispiel eines Inhaltes oder Maßes gesehen, was wir nun nachholen werden. Die folgenden Abbildungen entsprechen zwar nicht unserer geometrischen Vorstellung von Volumen (sie verletzen die Forderungen d) bzw e)), sind aber trotzdem Inhalte/Maße auf $\mathcal{P}(\mathbb{R}^d)$ im Sinne von Definition 1.1.

Beispiel 1.4.

- a) **(Das Dirac-Maß)** Sei $x \in \mathbb{R}^d$ und betrachte

$$\delta_x : \mathcal{P}(\mathbb{R}^d) \rightarrow [0, \infty], \quad \delta_x(A) := \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}.$$

Dann ist δ_x ein Maß. Es ist normiert, falls $x \in [0, 1]^d$, aber für kein x bewegungs-invariant (Übung).

- b) **(Das Zählmaß)** Ein weiteres einfaches Maß bekommt man durch Zählen der Elemente von A . (Erinnerung an Analysis 1: $|A|$ bezeichnet die Kardinalität einer Menge, also $|A| < \infty \Leftrightarrow (A \text{ ist endlich})$)

$$\zeta : \mathcal{P}(\mathbb{R}^d) \rightarrow [0, \infty], \quad \zeta(A) := \begin{cases} |A| & |A| < \infty \\ +\infty & |A| \not< \infty \end{cases}.$$

Dies ist ein bewegungsinvariantes aber nicht normiertes Maß (Beweis: Übung).

- c) Als Variation des vorigen Beispiels betrachten wir

$$\mu : \mathcal{P}(\mathbb{R}^d) \rightarrow [0, \infty], \quad \mu(A) := \begin{cases} 0 & |A| < \infty \\ \infty & |A| \not< \infty \end{cases}.$$

Diese Abbildung ist ein Inhalt (leichte Übung). Allerdings ist μ kein Maß, da μ nicht σ -additiv ist: Betrachten wir (für $d = 1$) z.B. die Mengen $A_n = \{n\}$, $n \in \mathbb{N}$, so gilt $\mu(A_n) = 0$ für alle n , also auch $\sum_{n=1}^{\infty} \mu(A_n) = 0$. Aber $\bigcup_{n=1}^{\infty} A_n = \mathbb{N}$ hat unendlich viele Elemente, also $\mu(\mathbb{N}) = +\infty$, im Widerspruch zu (1.1).

Insbesondere zeigt dieses Beispiel, dass σ -Additivität nicht automatisch aus endlicher Additivität folgt.

- d) Noch ein langweiliges Beispiel: $\mu(A) = 0$ für alle $A \subset \mathbb{R}^d$ (das **Nullmaß**).

Die oben angegebenen Beispiele belegen, dass der Begriff eines Maßes nicht leer ist. Ein Maß von zentraler Bedeutung in dieser Vorlesung ist das sogenannte *Lebesgue-Maß*, das z.B. $\mu([a, b]) = b - a$, $a \leq b$, erfüllt, also unserer geometrischen Anschauung einer Länge entspricht. Es

ist allerdings wesentlich komplizierter als die obigen Beispiele und kann nicht auf *beliebigen* Teilmengen von \mathbb{R}^d erklärt werden. Wir werden es deshalb erst später einführen.

Zum Abschluss dieses einführenden und motivierenden Abschnittes wollen wir aber auch noch einen Ausblick auf Anwendungen der Maßtheorie geben, die über die geometrische Frage von Volumenmessungen hinausgehen.

Beginnen wir mit der **Integrationstheorie**. Dazu stellen wir uns vor, dass wir bereits eine gute Menge $\mathcal{A} \subset \mathcal{P}(\mathbb{R}^d)$ von nicht zu wilden “messbaren” Mengen $A \subset \mathbb{R}^d$ (bzw. $\mathcal{A} \subset [0, 1]$, falls wir nur an Funktionen auf dem Einheitsintervall interessiert sind, oder gleich $A \subset \Omega$ mit einer allgemeinen Menge Ω) identifiziert haben, zusammen mit einem wie in Definition 1.1 definierten Maß $\mu : \mathcal{A} \rightarrow [0, \infty]$, das jetzt aber nur auf Mengen $A \in \mathcal{A}$ erklärt ist. Die Mengen $A \in \mathcal{A}$ nennen wir dann *meßbar*.

Zu jeder meßbaren Menge $A \subset \Omega$ betrachten wir ihre *charakteristische Funktion*

$$\chi_A(x) := \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}.$$

Das Integral von χ_A sollte dann mit dem Volumen von A übereinstimmen, d.h. wir werden $\int_{\Omega} \chi_A d\mu := \mu(A)$ setzen. Das “ $d\mu$ ” erinnert daran, dass diese Definition von dem gewählten Maß μ , also der Volumenmessvorschrift, abhängen wird. Ähnlich wie beim in Analysis 1 besprochenen Riemann-Integral dehnt sich diese Definition dann leicht auf Linearkombinationen von charakteristischen Funktionen aus und auf allgemeinere Funktionen per Approximation.

Ist $f : \mathbb{R}^d \rightarrow \mathbb{R}$ eine Funktion mit endlichem Bild (dh f nimmt nur endlich viele verschiedene Werte y_1, \dots, y_n an), so betrachten wir die Urbilder $A_k := f^{-1}(\{y_k\})$. Es gilt dann $f = \sum_{k=1}^n y_k \cdot \chi_{A_k}$ und wir setzen $\int f d\mu = \sum_{k=1}^n y_k \mu(A_k)$. Falls das Maß μ auf den Urbildmengen A_k definiert ist (d.h. falls diese Mengen nicht “zu wild” sind; dies ist eine Regularitätsbedingung an f), so gibt dieses Vorgehen ein wohldefiniertes Integral von f . Dieses “Integral nach einem Maß” wird also direkt auf die Idee der Volumenmessung mit einem Maß zurückgeführt und gründet auf der Idee, den Bild- statt den Definitionsbereich von f geeignet zu zerlegen. Da wir statt den beim Riemann-Integral verwendeten charakteristischen Funktionen von Intervallen sehr viel allgemeinere charakteristische Funktionen von “messbaren Mengen” zur Verfügung haben, wird sich zeigen, dass wir auf diese Art und Weise einen wesentlich allgemeineren und besser handhabbaren Integralbegriff als den Begriff des Riemann-Integrals erhalten.

Eine andere Motivation für die Maßtheorie ist die **Wahrscheinlichkeitstheorie**. Wir betrachten ein elementares Beispiel.

Beispiel 1.5. Es wird mit zwei sechseckigen Würfeln gewürfelt und die Augenzahl addiert. Die möglichen Ergebnisse bilden also die Menge

$$\Omega := \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$$

und treten “zufällig”^a auf. Wir sind jetzt daran interessiert, wie “wahrscheinlich”^a es ist, dass das Ergebnis dieses Experiments in einer gegebenen Teilmenge $A \subset \Omega$ liegt, z.B. $A = \{2\}$ oder $A = \{4, 5, 6\}$. Das heißt, wir wollen jeder Teilmenge $A \subset \Omega$ eine

Wahrscheinlichkeit, d.h. eine Zahl $\mu(A) \in [0, 1]$ zuordnen.

Natürlich sollte dabei $\mu(\emptyset) = 0$ (die Wahrscheinlichkeit, dass wir gar kein Ergebnis erhalten, ist Null) und, für disjunkte $A, B \subset \Omega$, auch $\mu(A \cup B) = \mu(A) + \mu(B)$ (die Wahrscheinlichkeit, dass das Ergebnis in A oder der dazu disjunkten Menge B liegt, ist die Summe der beiden Einzelwahrscheinlichkeiten) gelten. Wir sehen also, dass wir es hier automatisch mit einer Struktur zu tun haben, die unseren Ideen zu Volumenmessungen stark ähnelt. Der einzige Unterschied ist, dass hier die zu messenden Teilmengen nicht Teilmengen von \mathbb{R}^d , sondern einer anderen Menge Ω sind.

Zusätzlich erwarten wir hier noch $\mu(\Omega) = 1$ (die Wahrscheinlichkeit, dass wir irgendein Ergebnis erhalten, ist 100%). Solche Maße werden wir Wahrscheinlichkeitsmaße nennen.

Durch Abzählung der Möglichkeiten werden Sie sicher schnell $\mu(\{2\}) = \frac{1}{36}$, $\mu(\{2, 3\}) = \frac{1}{12}$ etc. erkennen und so eine Abbildung $\mu : \mathcal{P}(\Omega) \rightarrow \overline{\mathbb{R}}_+$ erhalten, die die in Def. 1.1 geforderten Eigenschaften hat.

Diese Abzählung beruht jedoch auf der Annahme, dass alle Würfelresultate eines einzelnen Würfels gleich wahrscheinlich sind. Wenn mit gezinkten Würfeln gespielt wird, verschieben sich die Wahrscheinlichkeiten, die dann durch ein anderes Maß modelliert werden. Im Extremfall zeigen beide Würfel bei jedem Wurf eine 6, was durch das Dirac-Maß δ_{12} (auf $\mathcal{P}(\Omega)$) repräsentiert wird.

^aWas das genau bedeutet, wird in der Wahrscheinlichkeitstheorie erklärt.

Die Nichtexistenz eines normierten bewegungsinvarianten Maßes auf $\mathcal{P}(\mathbb{R}^d)$ und das Beispiel mit den Würfeln legen beide aus unterschiedlichen Gründen nahe, dass wir Maße nicht nur auf $\mathcal{P}(\mathbb{R}^d)$, sondern auf allgemeineren Mengensystemen studieren sollten. Im ersten Fall sollten wir uns für eine Teilmenge von $\mathcal{P}(\mathbb{R}^d)$ interessieren, die aus "hinreichend regulären" Mengen besteht, im zweiten Fall ist das fragliche Mengensystem einfach die (endliche) Potenzmenge von $\{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$. Das Studium von geeigneten Mengensystemen wird im nächsten Abschnitt begonnen.

Eine wichtige Anwendung findet die Maßtheorie in der *Spektraltheorie*, wo es um eine Verallgemeinerung des Begriffs eines Eigenwertes für lineare Abbildungen T zwischen gewissen unendlichdimensionalen Vektorräumen V geht. An die Stelle der aus Linearer Algebra bekannten Zerlegung $V = \bigoplus_{\lambda} V_{\lambda}$ in Eigenräume V_{λ} von T zu Eigenwert λ tritt eine Zerlegung, die grob gesprochen sowohl diskrete als auch kontinuierliche Anteile haben kann und durch Integrale und Maße (auf der Menge aller verallgemeinerten Eigenwerte von T) beschrieben wird.

Weitere interessante Maße kommen aus Anwendungen in der Physik. Interessieren wir uns zB für ein sich mit konstanter Geschwindigkeit aber zufälliger Richtung im \mathbb{R}^3 bewegendes Teilchen. Falls das Teilchen zur Zeit $t = 0$ im Ursprung $0 \in \mathbb{R}^3$ startet, wie wahrscheinlich ist es dann, dass es sich bis zu einer Zeit $T > 0$ nicht weiter als höchstens Abstand $R > 0$ vom Ursprung entfernt hat?

Um dies mathematisch zu modellieren, könnten wir die Menge $\Omega := C([0, T], \mathbb{R}^3)$ aller stetigen Wege $[0, T] \rightarrow \mathbb{R}^3$ betrachten, die das Teilchen einschlagen könnte, und ein (nach physikalischen Gesetzmäßigkeiten definiertes) Wahrscheinlichkeitsmaß μ auf Teilmengen von Ω . Die

Teilmenge

$$A_R = \{f \in C([0, T], \mathbb{R}^3) : f(0) = 0, \|f(t)\|_2 \leq R \text{ für alle } t \in [0, T]\}$$

enthält dann genau all die stetigen Pfade, die zur Zeit $t = 0$ bei $f(0) = 0$ beginnen und in dem Zeitintervall $[0, T]$ nicht weiter als höchstens Radius R vom Ursprung entfernen. Die gesuchte Wahrscheinlichkeit wäre also $\mu(A_R)$.

Dieses Beispiel macht wiederum deutlich, dass auch Maße auf anderen Räumen als \mathbb{R}^d von Interesse sind.

Noch abenteuerlichere Maße finden sich in der Quantenfeldtheorie. Ob die zugrundeliegenden physikalischen Ideen wirklich in die rigorose Definition eines Maßes überführt werden können, ist hier häufig noch Gegenstand aktueller Forschung.

1.2 σ -Algebren und andere Mengensysteme

Motiviert durch die Überlegungen im vorigen Abschnitt betrachten wir nun eine beliebige nicht leere Grundmenge Ω und zugehörige Teilmengen der Potenzmenge $\mathcal{P}(\Omega)$ von Ω . Wir werden es im Folgenden viel mit Mengenoperationen zu tun haben – erinnern Sie sich an die grundlegenden Operationen \cup, \cap sowie das *Komplement* einer Teilmenge $A \subset \Omega$ (wie immer in diesem Kurs schließt das Symbol $A \subset \Omega$ die Möglichkeit $A = \Omega$ ein), nämlich

$$A^c := \Omega \setminus A.$$

Die de'Morganschen Regeln besagen für Teilmengen $A_i \subset \Omega$ einer beliebigen Menge Ω

$$\left(\bigcup_{i \in I} A_i\right)^c = \bigcap_{i \in I} A_i^c, \quad \left(\bigcap_{i \in I} A_i\right)^c = \bigcup_{i \in I} A_i^c.$$

Hierbei kann die Indexmenge I beliebig (also endlich, abzählbar unendlich oder überabzählbar) sein.

Weiterhin sei an die *Differenz* von $A, B \subset \Omega$ erinnert,

$$A \setminus B := \{x \in A : x \notin B\} = A \cap B^c.$$

Definition 1.6. Sei Ω eine nicht leere Menge. Eine σ -Algebra (von Ω) ist eine Teilmenge $\mathcal{A} \subset \mathcal{P}(\Omega)$ der Potenzmenge von Ω , die die folgenden drei Eigenschaften hat:

- $\emptyset \in \mathcal{A}$,
- $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$
- Für jede Folge $A_n \in \mathcal{A}$, $n \in \mathbb{N}$, gilt $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.

Bemerkungen: Teilmengen der Potenzmenge $\mathcal{P}(\Omega)$ (also Elemente von $\mathcal{P}(\mathcal{P}(\Omega))$) werden auch oft *Mengensysteme* genannt. Eine σ -Algebra ist also ein spezielles Mengensystem. Die Axiome a) und c) wurden bereits im vorigen Kapitel motiviert. Um b) zu motivieren, denken wir an Wahrscheinlichkeiten: Ist Ω eine Menge von möglichen Ergebnissen eines zufälligen

Ereignisses und $A \subset \Omega$, so bedeutet ein Ergebnis im Komplement A^c genau, dass das Ergebnis *nicht* in A liegt. Kann A also sinnvoll eine Wahrscheinlichkeit $\mu(A)$ zugeordnet werden, so kann auch A^c sinnvoll eine Wahrscheinlichkeit zugeordnet werden, nämlich $1 - \mu(A)$. Auch im Kontext von Volumenmessungen wird sich Axiom b) als eine wichtige Eigenschaft herausstellen.

Aufgrund der Eigenschaften a) und b) enthält jede σ -Algebra von Ω neben \emptyset auch die ganze Menge Ω .

Aufgrund der de'Morganschen Regeln gilt für jede Folge $(A_n)_n \subset \mathcal{A}$ auch

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

σ -Algebren sind also unter dem Bilden von abzählbaren Vereinigungen und abzählbaren Schnitten abgeschlossen. Wählt man einige der A_n als leere Menge, sieht man, dass σ -Algebren insbesondere unter dem Bilden von *endlichen* Vereinigungen und Schnitten abgeschlossen sind.

Beispiel 1.7. Sei Ω eine nicht leere Menge.

- Die kleinste^a σ -Algebra von Ω ist $\{\emptyset, \Omega\}$.
- Die größte σ -Algebra von Ω ist die Potenzmenge $\mathcal{P}(\Omega)$.
- Für $A \subset \Omega$ ist auch $\{\emptyset, A, A^c, \Omega\}$ eine σ -Algebra von Ω .
- $\mathcal{A} := \{A \subset \Omega : A \text{ oder } A^c \text{ ist abzählbar}\}$ ist eine σ -Algebra auf Ω (Übung).

^aHier beziehen sich die Begriffe "kleiner" und "größer" auf die partielle Ordnung durch mengentheoretische Inklusion, d.h. eine σ -Algebra \mathcal{A} ist "größer" als eine σ -Algebra \mathcal{B} , falls $\mathcal{A} \supset \mathcal{B}$.

Die Idee einer σ -Algebra ist, die messbaren Teilmengen von Ω zu identifizieren. Ein Maß wird dann eine Abbildung $\mu : \mathcal{A} \rightarrow [0, \infty]$ auf einer σ -Algebra sein, die \emptyset auf 0 abbildet und σ -additiv ist, genau wie in den Beispielen aus Definition 1.1, in der wir uns noch auf $\Omega = \mathbb{R}^d$ und $\mathcal{A} = \mathcal{P}(\mathbb{R}^d)$ beschränkt hatten. Über Maße werden wir im nächsten Kapitel sprechen; vorerst konzentrieren wir uns auf ihre Definitionsbereiche, dh auf σ -Algebren.

In den meisten Fällen wird die Grundmenge Ω unendlich sein. Hat man es dann nicht mit einem eher künstlichen Beispiel wie in Beispiel 1.7 zu tun, ist es meist schwierig, eine σ -Algebra explizit anzugeben. Noch schwieriger ist es dann, ein Maß auf einer σ -Algebra anzugeben. Die Idee zur Definition eines Maßes hat oft eine gewisse Analogie zur Linearen Algebra, wo man eine lineare Abbildung auf einem Vektorraum durch ihre Wirkung auf eine kleinere Menge, nämlich eine Basis, eindeutig definieren kann. Im Folgenden werden wir uns mit einigen weiteren Mengensystemen (Halbringe, Ringe, ...) beschäftigen, die einfacher als σ -Algebren sind und uns später dabei helfen werden, Maße zu definieren.

Denken Sie z.B. an die Menge aller Intervalle, sagen wir von der rechts halboffenen Form $[a, b)$. Fügen wir noch die leere Menge und ganz \mathbb{R} hinzu, so haben wir

$$\mathcal{I} = \{\emptyset, \mathbb{R}, [a, b) : a < b, a, b \in \mathbb{R}\} \subset \mathcal{P}(\mathbb{R}).$$

Dies ist noch ein recht übersichtliches Mengensystem, und im Sinne von Längenmessungen möchten wir ein "Maß" darauf definieren durch $\mu([a, b)) := b - a$, $\mu(\emptyset) = 0$, $\mu(\mathbb{R}) = +\infty$.

Allerdings ist \mathcal{I} keine σ -Algebra, da \mathcal{I} weder unter Komplementen noch unter abzählbaren Vereinigungen abgeschlossen ist (z.B. $[0, 1) \cup [4, 8) \notin \mathcal{I}$). Um eine σ -Algebra zu erhalten, müssen wir \mathcal{I} also vergrößern, indem wir alle endlichen Vereinigungen und Komplemente hinzufügen. Dies führt uns auf den Begriff eines Erzeugendensystems.

Lemma 1.8. Sei Ω eine nicht leere Menge und $\mathcal{E} \subset \mathcal{P}(\Omega)$ ein Mengensystem. Dann gibt es eine kleinste σ -Algebra, die \mathcal{E} enthält. Diese σ -Algebra wird als die von \mathcal{E} erzeugte σ -Algebra $\sigma(\mathcal{E})$ bezeichnet, und \mathcal{E} heißt ein Erzeuger von $\sigma(\mathcal{E})$.

Beweis. Da $\mathcal{P}(\Omega)$ eine σ -Algebra von Ω ist, gibt es \mathcal{E} enthaltende σ -Algebren. Also ist

$$\sigma(\mathcal{E}) := \bigcap_{\substack{\mathcal{A} \supset \mathcal{E} \\ \mathcal{A} \text{ } \sigma\text{-Algebra}}} \mathcal{A}$$

nicht leer und per Definition in jeder \mathcal{E} enthaltenden σ -Algebra enthalten. Aus Def. 1.6 folgt direkt, dass $\sigma(\mathcal{E})$ eine σ -Algebra ist. \square

Für eine beliebige nicht leere Menge Ω betrachten wir wie in Beispiel 1.7 iv) die σ -Algebra $\mathcal{A} = \{A \subset \Omega : A \text{ oder } A^c \text{ ist höchstens abzählbar}\}$. Zeigen Sie: $\mathcal{E} = \{\{x\} : x \in \Omega\}$ ist ein Erzeuger von \mathcal{A} . Wann gilt $\mathcal{A} \neq \mathcal{P}(\Omega)$?

Wir können z.B. die von allen offenen Mengen eines metrischen Raumes erzeugte σ -Algebra betrachten:

Definition 1.9. Sei (X, d) ein metrischer Raum. Die von den offenen Teilmengen von X erzeugte σ -Algebra von X heißt die *Borel- σ -Algebra* von (X, d) und wird mit $\mathcal{B}(X)$ bezeichnet. Die Elemente von $\mathcal{B}(X)$ heißen *Borelmengen*.

Die Borel- σ -Algebra von X ist per Definition von allen offenen Teilmengen von X erzeugt. Da sie aber unter Komplementbildung abgeschlossen ist, wird sie auch von allen abgeschlossenen Teilmengen von X erzeugt.

Die für uns wichtigste σ -Algebra ist die Borel- σ -Algebra von \mathbb{R}^d . Es spielt keine Rolle, welche Norm auf \mathbb{R}^d wir zur Definition der Metrik (und damit der offenen Mengen) zugrundelegen, da all diese Normen äquivalent sind (siehe Analysis 2). Die Menge $\mathcal{B}(\mathbb{R}^d)$ ist anschaulich sehr groß, so leicht fällt einem keine Teilmenge von \mathbb{R}^d ein, die nicht in $\mathcal{B}(\mathbb{R}^d)$ enthalten ist: Per Definition ist jede offene Teilmenge von \mathbb{R}^d eine Borelmenge, und wegen $(A \in \mathcal{B}(\mathbb{R}^d) \Rightarrow A^c \in \mathcal{B}(\mathbb{R}^d))$ ist auch jede abgeschlossene Teilmenge von \mathbb{R}^d Borel. Die rationalen Zahlen $\mathbb{Q} \subset \mathbb{R}$ (weder offen noch abgeschlossen) sind auch eine Borelmenge, denn \mathbb{Q} ist abzählbar und die Einpunktmengen $\{x\}$ sind abgeschlossen, also Borel. Vielleicht denken Sie jetzt an die Cantormenge $C \subset \mathbb{R}$ (Aufgabe H4.4 in Analysis 2), die überabzählbar ist. Aber C ist abgeschlossen, also auch Borel. Wir werden später aber sehen, dass $\mathcal{B}(\mathbb{R}^d)$ eine echte Teilmenge der Potenzmenge von \mathbb{R}^d ist.

Wir wollen die Borel- σ -Algebra von \mathbb{R}^d nun etwas näher verstehen, indem wir ein handhabbareres Erzeugendensystem als das Mengensystem aller offenen Mengen angeben. Wir definieren dazu ähnlich wie in Analysis 2 einen *Quader* als eine Teilmenge von \mathbb{R}^d der Gestalt

$$Q := [a_1, b_1) \times [a_2, b_2) \times \dots \times [a_d, b_d)$$

mit $a_j \leq b_j, j = 1, \dots, d$, und \mathcal{Q}^d als das Mengensystem all solcher Quader. Wir haben die Quader hier als rechts halboffen gewählt, da dies gut zu Zerlegungen passt (z.B. $[0, 1) \cup [1, 2) = [0, 2)$). Wir nennen Q einen *Würfel*, falls alle Kantenlängen gleich sind, dh falls $b_1 - a_1 = \dots = b_d - a_d$.

Da es recht viel Schreibarbeit ist, einen Quader zu notieren, vereinbaren wir noch folgende abkürzende Schreibweisen: Für $a, b \in \mathbb{R}^d$ definieren wir⁵

$$a < b \Leftrightarrow a_k < b_k, k = 1, \dots, d,$$

$$a \leq b \Leftrightarrow a_k \leq b_k, k = 1, \dots, d,$$

und die "mehrdimensionalen Intervalle"

$$[a, b] := \{x \in \mathbb{R}^d : a \leq x \leq b\}, \quad [a, b) := \{x \in \mathbb{R}^d : a \leq x < b\},$$

und analog $(a, b], (a, b)$. Dann ist $[a, b)$ genau der oben ausführlich geschriebene Quader, und

$$\mathcal{Q}^d = \{[a, b) : a, b \in \mathbb{R}^d, a \leq b\}$$

das Mengensystem aller Quader in \mathbb{R}^d .

Um die von \mathcal{Q}^d erzeugte σ -Algebra zu verstehen, beweisen wir das folgende Lemma.

Lemma 1.10. *Jede offene Menge $U \subset \mathbb{R}^d$ ist eine disjunkte Vereinigung von höchstens abzählbar vielen Würfeln mit rationalen Eckpunkten.*

Beweis. Für $n \in \mathbb{N}_0$ betrachten wir die Menge von Würfeln

$$\mathcal{W}_n := \left\{ \left[\frac{a_1}{2^n}, \frac{a_1+1}{2^n} \right) \times \dots \times \left[\frac{a_d}{2^n}, \frac{a_d+1}{2^n} \right) : a_1, \dots, a_d \in \mathbb{Z} \right\}$$

mit Kantenlänge 2^{-n} und Eckpunkten in $2^{-n}\mathbb{Z}^d$. Diese Menge ist abzählbar unendlich und bildet eine disjunkte Überdeckung von \mathbb{R}^d (je zwei unterschiedliche Elemente aus \mathcal{W}_n sind disjunkt, und $\mathbb{R}^d = \bigcup_{W \in \mathcal{W}_n} W$). Weiterhin sind zwei Würfel $W \in \mathcal{W}_n, \tilde{W} \in \mathcal{W}_m$ mit $n, m \in \mathbb{N}_0, n \leq m$, entweder disjunkt oder ineinander enthalten, d.h. $\tilde{W} \subset W$.

Wir betrachten nun die offene Menge $U \subset \mathbb{R}^d$ und definieren \mathcal{W}_0^U als die Menge aller Würfel $W \in \mathcal{W}_0$, die ganz in U enthalten sind. Induktiv definieren wir $\mathcal{W}_n^U, n \geq 1$, als die Menge aller Würfel $W \in \mathcal{W}_n$, die ganz in U enthalten sind, aber in keinem der Würfel aus $\bigcup_{k < n} \mathcal{W}_k^U$ enthalten sind.

Die Vereinigung all dieser Mengen, also $\mathcal{W}^U := \bigcup_{n=0}^{\infty} \mathcal{W}_n^U$, ist die gesuchte Überdeckung von U . Als abzählbare Vereinigung abzählbarer Mengen ist \mathcal{W}^U abzählbar. Per Konstruktion sind je zwei unterschiedliche Würfel aus \mathcal{W}^U disjunkt. Um noch zu zeigen, dass $U = \bigcup_{W \in \mathcal{W}^U} W$ gilt, betrachten wir einen beliebigen Punkt $x \in U$. Dann gibt es für jedes $n \in \mathbb{N}$ einen eindeutigen Würfel $W_n(x)$ mit den zwei Eigenschaften i) $x \in W_n(x)$ und ii) $W_n(x) \in \mathcal{W}_n$. Dieser Würfel ist in der ε -Umgebung in der Norm $\|\cdot\|_{\infty}$ mit $\varepsilon = 2^{-n}$, also $U_{2^{-n}}^{\infty}(x) = \{y \in \mathbb{R}^d : \|x - y\|_{\infty} < 2^{-n}\}$, enthalten. Falls $W_n(x) \not\subset U$ für alle $n \in \mathbb{N}$, liegt x also nicht im Inneren U° von U . Aber U ist nach Voraussetzung offen, dh $U = U^{\circ}$. Dies zeigt, dass für jedes $x \in U$ der Würfel $W_n(x)$ für groß genügendes n ganz in U liegt, d.h. $U = \bigcup_{W \in \mathcal{W}^U} W$. \square

⁵Dies ist eine *partielle Ordnung* auf \mathbb{R}^d , d.h. eine transitive, reflexive und antisymmetrische Relation. Für $d > 1$ ist es aber keine Ordnung, da für zwei $a, b \in \mathbb{R}^d$ nicht notwendigerweise eine der Aussagen $a < b, a = b, a > b$ gelten muss.

Satz 1.11. Die Borel- σ -Algebra $\mathcal{B}(\mathbb{R}^d)$ wird erzeugt von der Menge $\mathcal{Q}_{\mathbb{Q}}^d$ aller Quader mit rationalen Eckpunkten, d.h. $\sigma(\mathcal{Q}_{\mathbb{Q}}^d) = \sigma(\mathcal{Q}^d) = \mathcal{B}(\mathbb{R}^d)$.

Beweis. Wir bemerken zunächst, dass jeder Quader $Q \in \mathcal{Q}^d$ eine Borelmenge ist: Sei $e_n := (\frac{1}{n}, \dots, \frac{1}{n}) \in \mathbb{R}^d$. Dann ist der Quader $[a, b)$ (mit $a, b \in \mathbb{R}^d$, dh in mehrdimensionaler Intervallnotation) gleich dem Durchschnitt all der offenen Quader $(a - e_n, b) \in \mathcal{B}(\mathbb{R})$, $n \in \mathbb{N}$, und nach den Eigenschaften einer σ -Algebra ist $\mathcal{B}(\mathbb{R})$ unter abzählbaren Durchschnitten abgeschlossen. Also $[a, b) \in \mathcal{B}(\mathbb{R})$.

Wir haben also $\mathcal{Q}^d \subset \sigma(\mathcal{Q}^d) \subset \mathcal{B}(\mathbb{R}^d)$. Bezeichnet \mathcal{O} die Menge aller offenen Teilmengen von \mathbb{R}^d , so gilt nach dem vorausgehenden Lemma $\mathcal{O} \subset \sigma(\mathcal{Q}_{\mathbb{Q}}^d)$, also auch $\mathcal{B}(\mathbb{R}^d) = \sigma(\mathcal{O}) \subset \sigma(\mathcal{Q}_{\mathbb{Q}}^d)$. Insgesamt haben wir

$$\mathcal{B}(\mathbb{R}^d) \subset \sigma(\mathcal{Q}_{\mathbb{Q}}^d) \subset \sigma(\mathcal{Q}^d) \subset \mathcal{B}(\mathbb{R}^d),$$

was die Aussage des Satzes impliziert. □

Dieser Satz sagt uns, dass das übersichtlich kleine Mengensystem \mathcal{Q}^d die unübersichtlich große σ -Algebra $\mathcal{B}(\mathbb{R}^d)$ erzeugt. Während wir leicht Volumenabbildungen auf \mathcal{Q}^d angeben können, ist das auf $\mathcal{B}(\mathbb{R}^d)$ weniger klar. Wir befassen uns deshalb zunächst etwas näher mit Mengensystemen, die Eigenschaften analog zu \mathcal{Q}^d haben. Dies ist keine σ -Algebra (z.B. ist \mathcal{Q}^d nicht abgeschlossen unter Vereinigungen), aber wenigstens ein sogenannter Halbring.

Definition 1.12. Ein Mengensystem $\mathcal{A} \subset \mathcal{P}(\Omega)$ auf einer nicht leeren Menge Ω heißt *Halbring*, falls die folgenden drei Eigenschaften gelten:

- a) $\emptyset \in \mathcal{A}$,
- b) \mathcal{A} ist abgeschlossen unter Schnitten^a, d.h. $A, B \in \mathcal{A} \Rightarrow A \cap B \in \mathcal{A}$,
- c) für alle $A, B \in \mathcal{A}$ existieren endlich viele paarweise disjunkte Mengen $C_1, \dots, C_n \in \mathcal{A}$, so dass

$$A \setminus B = \bigcup_{k=1}^n C_k.$$

Falls \mathcal{A} auch unter endlichen Vereinigungen abgeschlossen ist, so heißt \mathcal{A} *Ring*.

^aMan sagt auch: \mathcal{A} ist *schnittstabil*.

Zeigen Sie: Ein \emptyset enthaltendes Mengensystem $\mathcal{R} \subset \mathcal{P}(\Omega)$ ist ein Ring genau dann, wenn er unter Vereinigungen und Differenzen abgeschlossen ist.

Lemma 1.13. Das Mengensystem \mathcal{Q}^d aller Quader in \mathbb{R}^d ist ein Halbring.

Beweis. Eigenschaft i) ist klar: $\emptyset = [a, a)$ für jedes $a \in \mathbb{R}^d$. Für ii) fassen wir max und min komponentenweise auf, dh $\max\{a, b\} = (\max\{a_1, b_1\}, \dots, \max\{a_d, b_d\})$ und analog für min. Dann gilt $[a, b) \cap [a', b') = [\max\{a, a'\}, \min\{b, b'\}) \in \mathcal{Q}^d$.

Für den Beweis von Eigenschaft iii) betrachten wir zunächst $d = 1$. Dann gilt (Skizze)

$$[a, b) \setminus [a', b') = [a, \min\{b, a'\}) \sqcup [\max\{a, b'\}, b),$$

eine disjunkte Vereinigung von höchstens zwei Elementen von \mathcal{Q}^1 .

Für $d > 1$ ist $[a, b) \setminus [a', b')$ im Allgemeinen eine disjunkte Vereinigung von mehreren Quadern in \mathcal{Q}^d , aber Eigenschaft iii) gilt immer noch. Ein Beweis kann per Induktion über d gegeben werden (Übung). \square

Ein Halbring \mathcal{H} ist im Allgemeinen kein Ring, denn \mathcal{H} muss nicht unter endlichen Vereinigungen abgeschlossen sein. Dieses Defizit lässt sich aber leicht beheben, indem wir den von \mathcal{H} erzeugten Ring betrachten. Dazu gehen wir ganz analog zu der erzeugten σ -Algebra eines Mengensystems vor (Lemma 1.8). Ist $\mathcal{H} \subset \mathcal{P}(\Omega)$ ein Halbring, so betrachten wir

$$\mathcal{R} := \bigcap_{\substack{R \supset \mathcal{H} \\ R \text{ Ring auf } \Omega}} R.$$

Da Schnitte von Ringen Ringe sind, ist \mathcal{R} ein Ring, der \mathcal{H} enthält, und zwar der kleinste Ring mit dieser Eigenschaft. Wir nennen \mathcal{R} den von \mathcal{H} erzeugten Ring.

Lemma 1.14. Sei $\mathcal{H} \subset \mathcal{P}(\Omega)$ ein Halbring. Dann stimmt die aus endlichen disjunkten Vereinigungen von Elementen von \mathcal{H} bestehende Menge,

$$\mathcal{R} := \left\{ \bigcup_{k=1}^n A_k : n \in \mathbb{N}, A_1, \dots, A_n \in \mathcal{H} \text{ paarweise disjunkt} \right\},$$

mit dem von \mathcal{H} erzeugten Ring überein. Insbesondere ist der von \mathcal{Q}^d erzeugte Ring der Ring $\mathcal{R}(\mathbb{R}^d)$ aller Quadersummen (Menge aller Mengen, die endliche disjunkte Vereinigungen von Quadern sind).

Beweis. Der Beweis wird in den Hausaufgaben geführt (Blatt 2). \square

1.3 Inhalte, Prämaße, und Maße

Nachdem wir uns etwas mit Mengensystemen beschäftigt haben, können wir nun die grundlegenden Definitionen der Maßtheorie angeben.

Definition 1.15.

- Ein *Messraum* ist ein Paar (Ω, \mathcal{A}) bestehend aus einer Menge Ω und einer σ -Algebra \mathcal{A} von Ω . Die Elemente von \mathcal{A} werden *messbare Mengen* genannt.
- Sei (Ω, \mathcal{A}) ein Messraum. Ein *Maß* ist eine Abbildung $\mu : \mathcal{A} \rightarrow [0, \infty]$, die $\mu(\emptyset) = 0$ und σ -Additivität erfüllt.
- Ein *Maßraum* ist ein Tripel $(\Omega, \mathcal{A}, \mu)$ bestehend aus einer Menge Ω , einer σ -Algebra \mathcal{A} von Ω , und einem Maß auf \mathcal{A} .
- Ein Maß μ heißt *endlich*, falls $\mu(\Omega) < \infty$. Ein Maß heißt *Wahrscheinlichkeitsmaß*,

falls $\mu(\Omega) = 1$.

- Einige Beispiele auf der σ -Algebra $\mathcal{P}(\mathbb{R}^d)$ haben wir bereits in Beispiel 1.4 kennengelernt. Das Dirac-Maß lässt sich auch auf einem allgemeinen Messraum (Ω, \mathcal{A}) definieren: Für $\omega \in \Omega$ setzen wir,

$$\delta_\omega(A) := \begin{cases} 1 & \omega \in A, \\ 0 & \omega \notin A \end{cases}, \quad A \subset \Omega,$$

und sehen leicht, dass δ_ω ein Wahrscheinlichkeitsmaß auf Ω ist.

- Ist (Ω, \mathcal{A}) ein Messraum und $\mu_n, n \in \mathbb{N}$, Maße auf \mathcal{A} , so ist für beliebige positive Koeffizienten $c_n \geq 0$ auch $\mu(A) := \sum_{n=1}^{\infty} c_n \mu_n(A)$ ein Maß auf \mathcal{A} (Übung).

Für abzählbare Mengen Ω können wir sogar alle Maße auf $\mathcal{P}(\Omega)$ angeben.

Beispiel 1.16. Sei Ω eine nicht leere höchstens abzählbar unendliche Menge und $m : \Omega \rightarrow [0, \infty]$ eine beliebige Abbildung. Dann ist

$$\mu : \mathcal{P}(\Omega) \rightarrow [0, \infty], \quad \mu(A) := \sum_{\omega \in A} m(\omega)$$

ein Maß auf $\mathcal{A} = \mathcal{P}(\Omega)$, und jedes Maß auf $\mathcal{P}(\Omega)$ ist von dieser Form. Die Notation $\sum_{\omega \in A} m(\omega)$ bereitet keine Schwierigkeiten (die Reihenfolge der Summation ist nicht fixiert), da $m(\omega) \geq 0$ und die Reihe damit unabhängig von Umordnungen gegen $+\infty$ divergiert oder absolut konvergiert (Analysis 1).

Der Beweis dieser Aussage wird in den Hausaufgaben geführt (Blatt 1).

Lemma 1.17 (Grundlegende Eigenschaften von Maßen). Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum.

a) Für $A, B \in \mathcal{A}$ mit $A \subset B$ gilt

$$\mu(B) = \mu(A) + \mu(B \setminus A),$$

insbesondere also $\mu(A) \leq \mu(B)$ (Monotonie).

b) Für alle $A, B \in \mathcal{A}$ gilt

$$\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B).$$

c) Für eine aufsteigende Folge $(A_n)_{n \in \mathbb{N}} \subset \mathcal{A}$, $A_n \subset A_{n+1}$ für alle $n \in \mathbb{N}$, gilt

$$\mu \left(\bigcup_{n=1}^{\infty} A_n \right) = \lim_{n \rightarrow \infty} \mu(A_n) = \sup_{n \in \mathbb{N}} \mu(A_n).$$

d) Für eine absteigende Folge $(A_n)_{n \in \mathbb{N}} \subset \mathcal{A}$, $A_n \supset A_{n+1}$ für alle $n \in \mathbb{N}$, mit $\mu(A_1) < \infty$ gilt

$$\mu \left(\bigcap_{n=1}^{\infty} A_n \right) = \lim_{n \rightarrow \infty} \mu(A_n) = \inf_{n \in \mathbb{N}} \mu(A_n).$$

e) Für eine beliebige Folge $(A_n)_n \subset \mathcal{A}$ gilt (σ -Subadditivität)

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

Beweis. a) $B \setminus A = B \cap A^c$ liegt in \mathcal{A} nach den Eigenschaften einer σ -Algebra. Da $B = A \sqcup (B \setminus A)$ (disjunkte Vereinigung), folgt die erste Aussage aus Additivität von μ . Wegen $\mu(B \setminus A) \geq 0$ folgt auch die zweite Aussage.

b) Aus $A \cup B = A \sqcup (B \setminus A)$ und $B = (B \setminus A) \sqcup (A \cap B)$ (disjunkte Vereinigungen) folgt per Additivität von μ

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

c) Wir definieren $B_1 := A_1$, $B_{n+1} := A_{n+1} \setminus A_n$, $n \geq 1$. Dann sind die B_n , $n \in \mathbb{N}$, paarweise disjunkt und erfüllen $A_n = B_1 \sqcup \dots \sqcup B_n$, also $A := \bigcup_{n \in \mathbb{N}} A_n = \bigsqcup_{n \in \mathbb{N}} B_n$. Die σ -Additivität von μ liefert nun

$$\mu(A) = \sum_{n=1}^{\infty} \mu(B_n) = \lim_{n \rightarrow \infty} \sum_{k=1}^n \mu(B_k) = \lim_{n \rightarrow \infty} \mu(A_n).$$

Aufgrund der Monotonie von μ und $A_n \subset A_{n+1}$ ist $\mu(A_n)$ eine monoton wachsende Folge. Also $\lim_{n \rightarrow \infty} \mu(A_n) = \sup_{n \in \mathbb{N}} \mu(A_n)$.

d) Wir definieren $B_n := A_1 \setminus A_n$, $n \in \mathbb{N}$ und $A := \bigcap_{n \in \mathbb{N}} A_n$. Dann ist (B_n) eine aufsteigende Folge, d.h. $B_n \subset B_{n+1}$ für alle $n \in \mathbb{N}$, und

$$\bigcup_{n \in \mathbb{N}} B_n = \bigcup_{n \in \mathbb{N}} A_1 \setminus A_n = A_1 \cap \bigcup_{n \in \mathbb{N}} A_n^c = A_1 \cap \left(\bigcap_{n \in \mathbb{N}} A_n \right)^c = A_1 \setminus A.$$

Da $(B_n)_n$ aufsteigend ist, folgt aus c) $\sup_n \mu(B_n) = \mu(A_1 \setminus A) = \mu(A_1) - \mu(A)$ (im letzten Schritt haben wir a) auf die Inklusion $A \subset A_1$ angewendet und $\mu(A_1) < \infty$ benutzt, so dass $\mu(A_1) - \mu(A)$ wohldefiniert ist). Andererseits gilt $\sup_n \mu(B_n) = \sup_n (\mu(A_1) - \mu(A_n)) = \mu(A_1) - \inf_n \mu(A_n)$. Dies impliziert $\mu(A) = \inf_n \mu(A_n)$. Da die Folge (A_n) absteigend ist, ist $(\mu(A_n))_n$ monoton fallend, also $\inf_n \mu(A_n) = \lim_n \mu(A_n)$.

e) Wir setzen ähnlich wie in c) $A := \bigcup_{n \in \mathbb{N}} A_n$ sowie $B_1 := A_1$, $B_{n+1} := A_{n+1} \setminus \bigcup_{k=1}^n A_k$, $n \geq 1$, und haben $A = \bigsqcup_{n \in \mathbb{N}} B_n$. Die σ -Additivität von μ ergibt $\mu(A) = \sum_{n=1}^{\infty} \mu(B_n)$. Wegen $B_{n+1} \subset A_{n+1}$ und Monotonie folgt die Behauptung. \square

Einige Bemerkungen zu diesem Lemma:

- Wir dürfen im Allgemeinen b) nicht als $\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$ umformulieren, da sowohl $\mu(A)$ als auch $\mu(A \cap B)$ unendlich sein könnten, was auf den nicht definierten Ausdruck " $\infty - \infty$ " führen würde. Bei Minuszeichen ist immer Vorsicht geboten!
- In Teil d) kann die Voraussetzung $\mu(A_1) < \infty$ nicht weggelassen werden. Beispiel: Sei μ das Zählmaß auf $\mathcal{P}(\mathbb{N})$ und $A_n := \{k \in \mathbb{N} : k \geq n\}$. Dann gilt $A_{n+1} \subset A_n$ und $\mu(A_n) = \infty$ für alle n . Aber $\bigcap_n A_n = \emptyset$ und $\mu(\emptyset) = 0 \neq \lim_{n \rightarrow \infty} \mu(A_n)$.

- In der Situation von c), also im Falle einer aufsteigenden Folge $A_n \subset A_{n+1}$ mit $A := \bigcup_n A_n$, schreiben wir auch kurz $A_n \nearrow A$ und nennen $(A_n)_n$ eine *Ausschöpfung von A*. Analog schreiben wir in der Situation von d) $A_n \searrow A$, d.h. falls $A_{n+1} \subset A_n$ und $A = \bigcap_n A_n$.

Definition 1.18. Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum. Eine Menge $M \subset \Omega$ heißt μ -Nullmenge (oder kurz *Nullmenge* wenn das Maß aus dem Kontext klar ist), falls es $A \in \mathcal{A}$ gibt mit $M \subset A$ und $\mu(A) = 0$.

Der Maßraum heißt *vollständig*, falls jede μ -Nullmenge in \mathcal{A} enthalten ist.

Lemma 1.19. Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum.

- Eine meßbare Menge $A \in \mathcal{A}$ ist genau dann eine Nullmenge, wenn $\mu(A) = 0$.
- Teilmengen von Nullmengen sind Nullmengen.
- Abzählbare Vereinigungen von Nullmengen sind Nullmengen.

Beweis. a) Sei $A \in \mathcal{A}$ eine Nullmenge. Per Definition gibt es $B \in \mathcal{A}$ mit $A \subset B$ und $\mu(B) = 0$. Dann folgt per Monotonie $0 \leq \mu(A) \leq \mu(B) = 0$, also $\mu(A) = 0$. Gilt hingegen $\mu(A) = 0$, so zeigt $A \subset A$, dass A eine Nullmenge ist.

b) ist klar.

c) Seien $A_n \subset \Omega$, $n \in \mathbb{N}$, μ -Nullmengen. Dann gibt es $B_n \in \mathcal{A}$ mit $A_n \subset B_n$ und $\mu(B_n) = 0$. Es gilt $\bigcup_n A_n \subset \bigcup_n B_n \in \mathcal{A}$, und per σ -Subadditivität $\mu(\bigcup_n B_n) \leq \sum_n \mu(B_n) = \sum_n 0 = 0$. \square

Beispiel 1.20. Sei $\mu : \mathcal{B}(\mathbb{R}) \rightarrow [0, \infty]$ ein Maß auf der Borel- σ -Algebra von \mathbb{R} (kurz: ein Borelmaß), dass $\mu((a, b)) = b - a$ für alle $a \leq b$ erfüllt. (Wir werden ein solches Maß, das sogenannte Lebesgue-Borel-Maß, später konstruieren). Dann sind alle abzählbaren Teilmengen von \mathbb{R} μ -Nullmengen.

Beweis: Wir betrachten zunächst die Einpunktmengen $\{x\} \in \mathcal{B}(\mathbb{R})$ ($\{x\}$ ist abgeschl.), und $\{x\} = \bigcap_{n \in \mathbb{N}} (x - \frac{1}{n}, x + \frac{1}{n})$. Die Folge $A_n := (x - \frac{1}{n}, x + \frac{1}{n})$ ist absteigend, und $\mu(A_1) = \mu(x - 1, x + 1) = 2 < \infty$. Also können wir Teil d) von Lemma 1.17 verwenden und schließen

$$\mu(\{x\}) = \inf_n \mu(x - \frac{1}{n}, x + \frac{1}{n}) = \inf_n \frac{2}{n} = 0.$$

Nach Lemma 1.19 sind damit auch alle abzählbaren Mengen, die ja Vereinigungen ihrer abzählbar vielen Elemente sind, Nullmengen. \square

Nebenbei bemerkt impliziert dies auch

$$\mu([a, b]) = \mu([a, b)) = \mu((a, b]) = \mu((a, b)) = b - a$$

für alle $a \leq b$. Denn zum Beispiel $\mu((a, b]) = \mu((a, b) \sqcup \{b\}) = \mu((a, b)) + \mu(\{b\}) = b - a + 0 = b - a$.

Beispiel 1.21. Wir betrachten eine Menge Ω , ein Element $\omega \in \Omega$, und das Dirac-Maß δ_ω auf einer σ -Algebra \mathcal{A} auf Ω mit $\{\omega\} \in \mathcal{A}$. Dann ist $M \subset \Omega$ eine δ_ω -Nullmenge genau dann, wenn $\omega \notin M$.

Beweis: Angenommen, $M \subset \Omega$ ist eine Nullmenge, dh es gibt $A \in \mathcal{A}$ mit $M \subset A$ und $\delta_\omega(A) = 0 \Leftrightarrow \omega \notin A$. Da $M \subset A$, haben wir auch $\omega \notin M$.

Angenommen, $\omega \notin M$, d.h. $M \subset \{\omega\}^c$. Aber $\{\omega\}^c \in \mathcal{A}$, da \mathcal{A} unter Komplementen abgeschlossen ist, und offenbar ist $\{\omega\}^c$ eine δ_ω -Nullmenge. Also ist M eine δ_ω -Nullmenge.

Bemerkung: In dieser Situation ist δ_ω vollständig genau dann, wenn $\mathcal{A} = \mathcal{P}(\Omega)$.

Ein unvollständiger Maßraum $(\Omega, \mathcal{A}, \mu)$ lässt sich immer vervollständigen, indem alle nicht messbaren Nullmengen zu \mathcal{A} hinzugefügt werden:

Lemma 1.22. Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und $\mathcal{N} \subset \mathcal{P}(\Omega)$ die Menge aller μ -Nullmengen. Dann ist

$$\tilde{\mathcal{A}} := \{A \cup N : A \in \mathcal{A}, N \in \mathcal{N}\}$$

eine σ -Algebra, und

$$\tilde{\mu} : \tilde{\mathcal{A}} \rightarrow [0, \infty], \quad \tilde{\mu}(A \cup N) := \mu(A)$$

ein vollständiges Maß. Wir nennen $\tilde{\mu}$ die Vervollständigung von μ .

Der Beweis wird in den Übungen geführt.

Die Lemmata 1.17, 1.19 und 1.22 sind Beispiele dafür, wie Sie mit Maßen in einem axiomatischen Zugang arbeiten können. Es beantwortet aber nicht die Frage, wie wir uns interessante Beispiele von Maßen verschaffen können; nur in den seltensten Fällen wird das so explizit wie in Beispiel 1.16 möglich sein.

Da die Konstruktion von Maßen oft über Halbringe und Ringe geschehen wird, geben wir jetzt zwei Abschwächungen des Maßbegriffs.

Definition 1.23. Sei $\mathcal{H} \subset \mathcal{P}(\Omega)$ ein Halbring. Ein *Inhalt* ist eine Abbildung $\nu : \mathcal{H} \rightarrow [0, \infty]$, die $\nu(\emptyset) = 0$ und endliche Additivität, d.h. für paarweise disjunkte $A_1, \dots, A_N \in \mathcal{H}$,

$$\nu \left(\bigsqcup_{n=1}^N A_n \right) = \sum_{n=1}^N \nu(A_n), \quad \text{falls } \bigsqcup_{n=1}^N A_n \in \mathcal{H},$$

erfüllt. Ein Inhalt, der sogar σ -additiv ist, d.h. der für paarweise disjunkte $A_n \in \mathcal{H}$,

$n \in \mathbb{N}$,

$$\nu \left(\bigsqcup_{n \in \mathbb{N}} A_n \right) = \sum_{n=1}^{\infty} \nu(A_n), \quad \text{falls } \bigsqcup_{n \in \mathbb{N}} A_n \in \mathcal{H},$$

erfüllt, heißt *Prämaß*.

Beachten Sie, dass die Bedingung $\bigcup_{n=1}^N A_n \in \mathcal{H}$ bzw. $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{H}$ gestellt werden muss, da Halbringe im Allgemeinen nicht unter (endlichen bzw. abzählbaren) Vereinigungen abgeschlossen sind. Dies ist die Verallgemeinerung von Def. 1.1 auf Halbringe. Natürlich können wir insbesondere Inhalte bzw. Prämaße auf Ringen betrachten, denn Ringe sind ja spezielle Halbringe.

Sehr viele Beispiele von Inhalten auf dem Halbring \mathcal{Q}^1 (tatsächlich alle Inhalte ν auf \mathcal{Q}^1 , die endlich sind in dem Sinne, dass $\nu(I) < \infty$ für alle $I \in \mathcal{Q}^1$) lernen Sie in einer Hausaufgabe auf Blatt 2 kennen.

Jeder Inhalt auf einem Halbring setzt sich automatisch zu einem Inhalt auf dem erzeugten Ring (siehe Lemma 1.14) fort:

Satz 1.24. *Sei $\mathcal{H} \subset \mathcal{P}(\Omega)$ ein Halbring und $\nu : \mathcal{H} \rightarrow [0, \infty]$ ein Inhalt. Dann gibt es genau einen Inhalt $\bar{\nu}$ auf dem von \mathcal{H} erzeugten Ring \mathcal{R} mit $\bar{\nu}|_{\mathcal{H}} = \nu$, nämlich*

$$\bar{\nu} \left(\bigsqcup_{k=1}^n A_k \right) = \sum_{k=1}^n \nu(A_k)$$

für paarweise disjunkte $A_k \in \mathcal{H}$.

Beweis. Sind $A_1, \dots, A_n \in \mathcal{H}$ paarweise disjunkt, so liegt $\bigsqcup_{k=1}^n A_k$ in \mathcal{R} (Lemma 1.14). Ein Inhalt $\bar{\nu}$ auf \mathcal{R} , der sich zu ν auf \mathcal{H} einschränkt, muss wegen der endlichen Additivität von Inhalten $\bar{\nu}(\bigsqcup_{k=1}^n A_k) = \sum_{k=1}^n \nu(A_k)$ erfüllen. Falls eine Fortsetzung $\bar{\nu}$ von ν existiert, ist sie also eindeutig.

Um die Existenz der Fortsetzung $\bar{\nu}$ zu zeigen, wollen wir $\bar{\nu}$ durch $\bar{\nu}(\bigsqcup_{k=1}^n A_k) = \sum_{k=1}^n \nu(A_k)$ definieren. Es stellt sich die Frage, ob das eine wohldefinierte Definition ergibt, denn eine Menge $A = \bigsqcup_{k=1}^n A_k$ könnte ja auch in anderer Weise $A = \bigsqcup_{l=1}^m B_l$ als endliche disjunkte Vereinigung von Mengen $B_l \in \mathcal{H}$ dargestellt werden. Wir müssen also zeigen, dass $\sum_{k=1}^n \nu(A_k) = \sum_{l=1}^m \nu(B_l)$ gilt. Dazu setzen wir $C_{kl} := A_k \cap B_l \in \mathcal{H}$ und bemerken, dass die C_{kl} paarweise disjunkt sind. Da ν ein Inhalt ist, gilt

$$\begin{aligned} \sum_{k=1}^n \nu(A_k) &= \sum_{k=1}^n \nu(A_k \cap A) = \sum_{k=1}^n \nu \left(A_k \cap \bigcup_{l=1}^m B_l \right) = \sum_{k=1}^n \nu \left(\bigcup_{l=1}^m C_{kl} \right) \\ &= \sum_{k=1}^n \sum_{l=1}^m \nu(C_{kl}) = \sum_{l=1}^m \nu(B_l). \end{aligned}$$

Also ist $\bar{\nu} : \mathcal{R} \rightarrow [0, \infty]$ wohldefiniert. Offensichtlich gilt $\bar{\nu}(\emptyset) = 0$. Sind $A, B \in \mathcal{R}$ disjunkt, so finden wir nach Lemma 1.14 endlich viele paarweise disjunkte A_k, B_l mit $A =$

$\bigcup_k A_k, B = \bigcup_l B_l$. Also haben wir

$$\bar{\nu}(A \cup B) = \bar{\nu}\left(\bigcup_k A_k \cup \bigcup_l B_l\right) = \sum_k \nu(A_k) + \sum_l \nu(B_l) = \bar{\nu}(A) + \bar{\nu}(B),$$

d.h. $\bar{\nu}$ ist auch ein Inhalt. □

Zeigen Sie, dass ein Inhalt ν auf einem Halbring \mathcal{H} monoton und subadditiv ist:

- $A, B \in \mathcal{H}, A \subset B \Rightarrow \nu(A) \leq \nu(B)$.
- Für $A_1, \dots, A_n \in \mathcal{H}$ mit $A := \bigcup_{k=1}^n A_k \in \mathcal{H}$ gilt $\nu(A) \leq \sum_{k=1}^n \nu(A_k)$.

Es ist mitunter nützlich, alternative Charakterisierungen der wichtigen Eigenschaft der σ -Additivität eines Inhalts zu kennen. Der nächste Satz liefert solche Charakterisierungen.

Satz 1.25. Sei $\nu : \mathcal{R} \rightarrow [0, \infty]$ ein Inhalt auf einem Ring \mathcal{R} . Dann sind die folgenden drei Aussagen äquivalent:

- a) ν ist ein Prämaß, d.h. σ -additiv.
- b) ν ist σ -subadditiv (siehe Lemma 1.17 e)).
- c) ν erhält Suprema aufsteigender Folgen, d.h. $\nu(A) = \sup_n \nu(A_n)$ für jede Ausschöpfung $A_n \nearrow A$ (mit $A_n, A \in \mathcal{R}$).

Beweis. b) \Rightarrow a): Sei $(A_n)_n$ eine Folge paarweise disjunkter Mengen in \mathcal{R} mit $A := \bigsqcup_{n=1}^{\infty} A_n \in \mathcal{R}$. Da ν als Inhalt endlich additiv und monoton (siehe Übung vor diesem Satz) ist, gilt für jedes $n \in \mathbb{N}$

$$\sum_{k=1}^n \nu(A_k) = \nu\left(\bigsqcup_{k=1}^n A_k\right) \leq \nu(A).$$

Also ist $\nu(A)$ eine obere Schranke für die monoton wachsende Folge $n \mapsto \sum_{k=1}^n \nu(A_k)$ und damit auch für ihren Grenzwert $\sum_{k=1}^{\infty} \nu(A_k)$. Die umgekehrte Ungleichung $\nu(A) \leq \sum_{k=1}^{\infty} \nu(A_k)$ gilt per Annahme der σ -Subadditivität b).

a) \Rightarrow c) Dieser Schritt ist identisch zu dem Beweis von Lemma 1.17 c) – Sie müssen nur prüfen, dass die Mengen $B_1 := A_1, B_{n+1} := A_{n+1} \setminus A_n$ in \mathcal{R} liegen (Eigenschaft eines Rings).

c) \Rightarrow b) Sei $(A_n)_n$ eine beliebige Folge in \mathcal{R} mit der Eigenschaft $A := \bigcup_n A_n \in \mathcal{R}$. Setze $B_n := \bigcup_{k=1}^n A_k \in \mathcal{R}$, dies ist eine Ausschöpfung von A . Also haben wir

$$\nu(A) = \nu\left(\bigcup_{n=1}^{\infty} B_n\right) = \lim_{n \rightarrow \infty} \nu(B_n) = \lim_{n \rightarrow \infty} \nu\left(\bigcup_{k=1}^n A_k\right),$$

und aufgrund der endlichen Subadditivität von ν (siehe Übung vor diesem Satz) gilt

$$\nu\left(\bigcup_{k=1}^n A_k\right) \leq \sum_{k=1}^n \nu(A_k) \xrightarrow{n \rightarrow \infty} \sum_{k=1}^{\infty} \nu(A_k).$$

□

Unser wichtigstes Beispiel eines Halbringes, der kein Ring ist, ist der Halbring \mathcal{Q}^d aller Quader in \mathbb{R}^d . Unser wichtigstes Beispiel eines Ringes, der keine σ -Algebra ist, ist der von \mathcal{Q}^d erzeugte Ring $\Omega(\mathbb{R}^d)$ der Quadersummen (endliche disjunkte Vereinigungen von Quadern, siehe Lemma 1.14). Und unser wichtigstes Beispiel eines Inhalts ist das Elementarvolumen, das wir jetzt diskutieren.

Satz 1.26. *Wir betrachten den Halbring \mathcal{Q}^d aller Quader in \mathbb{R}^d . Das Elementarvolumen, dh die Abbildung^a*

$$\text{Vol}_d : \mathcal{Q}^d \rightarrow [0, \infty), \quad \text{Vol}_d([a, b]) := \prod_{j=1}^d (b_j - a_j),$$

ist ein Inhalt.

^aWenn die Dimension $d \in \mathbb{N}$ aus dem Kontext klar ist, schreiben wir Vol statt Vol_d .

Beweis. Per Definition gilt $\text{Vol}(\emptyset) = 0$. Wir müssen Additivität zeigen: Sind $Q_1, \dots, Q_n \in \mathcal{Q}^d$ paarweise disjunkte rechtsoffene Quader, deren Vereinigung $Q = \bigsqcup_{k=1}^n Q_k$ auch ein rechtsoffener Quader ist, so gilt $\text{Vol}(Q) = \sum_{k=1}^n \text{Vol}(Q_k)$.

Wir zeigen diese Gleichung per Induktion über die Dimension d . Für $d = 1$ sind die $Q_k = [a_k, b_k]$ Intervalle. Da es sowohl bei Vereinigungen als auch bei endlichen Summen nicht auf die Reihenfolge ankommt, dürfen wir die Quader so sortieren, dass $b_1 < b_2 < \dots < b_n$. Dann gilt

$$a_1 < b_1 = a_2 < b_2 = a_3 < \dots < b_{n-1} = a_n < b_n,$$

und $Q = [a_1, b_n]$. Per Definition von Vol_1 gilt

$$\sum_{k=1}^n \text{Vol}_1(Q_k) = \sum_{k=1}^n (b_k - a_k) = b_n - a_1 = \text{Vol}_1(Q),$$

und der Beweis des Induktionsanfangs ist abgeschlossen.

Für den Induktionsschritt sei $d \geq 2$ und wir nehmen an, dass $\text{Vol}_{d-1} : \mathcal{Q}^{d-1} \rightarrow [0, \infty]$ ein Inhalt ist. Seien $Q_k \in \mathcal{Q}^d$ paarweise disjunkte Quader, wobei der Index k über eine endliche Menge K läuft, und sei $Q := \bigcup_{k \in K} Q_k$ ebenfalls ein Quader in \mathcal{Q}^d .

Jeder d -dimensionale Quader ist ein Kreuzprodukt von einem eindimensionalen Quader (Intervall $[a, b]$) und einem $(d - 1)$ -dimensionalen Quader, d.h.

$$Q_k = I_k \times Q'_k, \quad Q = I \times Q',$$

mit $I, I_k \in \mathcal{Q}^1$, $Q_k, Q'_k \in \mathcal{Q}^{d-1}$. Per Definition des Volumens (Produkt der Seitenlängen) gilt

$$\text{Vol}_d(Q_k) = \text{Vol}_1(I_k) \cdot \text{Vol}_{d-1}(Q'_k), \quad \text{Vol}_d(Q) = \text{Vol}_1(I) \cdot \text{Vol}_{d-1}(Q').$$

Die Intervalle I_k, I beschreiben die x_1 -Koordinaten der Punkte in Q_k, Q und erfüllen deshalb $I = \bigcup_k I_k$. Sie müssen aber nicht paarweise disjunkt sein (Skizze für $d = 2$ hilft hier). Wir können aber zu einer feineren Zerlegung von I übergehen, dh es gibt eine endliche Menge L

und paarweise disjunkte Intervalle $J_l \in \mathcal{Q}^1$, $l \in L$, so dass $I_k = \bigsqcup_{l \in L_k} J_l$ für eine Teilmenge $L_k \subset L$.

Mit dieser Unterteilung haben wir $Q_k = \bigsqcup_{l \in L_k} J_l \times Q'_k$ und wegen $\text{Vol}_1(I_k) = \sum_{l \in L_k} \text{Vol}_1(J_l)$

$$\text{Vol}_d(Q_k) = \text{Vol}_1(I_k) \cdot \text{Vol}_{d-1}(Q'_k) = \sum_{l \in L_k} \text{Vol}_1(J_l) \cdot \text{Vol}_{d-1}(Q'_k).$$

Für die Vereinigung Q der Quader gilt $Q = \bigsqcup_{k \in K} \bigsqcup_{l \in L_k} J_l \times Q'_k$. Sei $K_l := \{k \in K : l \in L_k\}$. Dann gilt $Q = \bigsqcup_{l \in L} \bigsqcup_{k \in K_l} J_l \times Q'_k$, und $Q' = \bigsqcup_{k \in K_l} Q'_k$ für jedes $l \in L$. Dies ergibt mit der Induktionsvoraussetzung $\text{Vol}_{d-1}(Q') = \sum_{k \in K_l} \text{Vol}_{d-1}(Q'_k)$.

So erhalten wir

$$\begin{aligned} \text{Vol}_d(Q) &= \text{Vol}_1(I) \text{Vol}_{d-1}(Q') = \sum_{l \in L} \text{Vol}_1(J_l) \text{Vol}_{d-1}(Q') \\ &= \sum_{l \in L} \sum_{k \in K_l} \text{Vol}_1(J_l) \text{Vol}_{d-1}(Q'_k) = \sum_{k \in K} \sum_{l \in L_k} \text{Vol}_1(J_l) \text{Vol}_{d-1}(Q'_k) = \sum_{k \in K} \text{Vol}_d(Q_k). \end{aligned}$$

□

Gemäß Satz 1.24 dehnt sich das Elementarvolumen automatisch auf den Ring aller Quadersummen aus. Wir werden etwas später sehen, dass dieser Inhalt sogar ein Prämaß ist, das sogenannte Lebesgue-Prämaß.

1.4 Fortsetzungen von Prämaßen zu Maßen

Was konkrete Volumenmessungen angeht, haben wir im letzten Abschnitt den elementargeometrischen Inhalt $\text{Vol}_d : \mathfrak{Q}(\mathbb{R}^d) \rightarrow [0, \infty)$ auf dem Ring aller Quadersummen konstruiert. Wir wollen diesen Inhalt jetzt zu einem Maß auf einer geeigneten σ -Algebra erweitern, um das Lebesgue-Maß zu erhalten.

Abstrakt gesehen stehen wir also vor der folgenden Frage: Gegeben einen Inhalt $\nu : \mathcal{H} \rightarrow [0, \infty]$ auf einem Halbring \mathcal{H} (über einer nicht leeren Menge Ω), wie können wir eine natürliche \mathcal{H} enthaltende σ -Algebra \mathcal{A} und eine Fortsetzung von ν zu einem Maß μ auf \mathcal{A} (dh $\mu|_{\mathcal{R}} = \nu$) konstruieren? Betrachten wir zB die von \mathcal{H} erzeugte σ -Algebra $\sigma(\mathcal{H})$, ist es nicht klar, dass eine solche Fortsetzung existiert und eindeutig ist. Wir können zwar immer auf den erzeugten Ring fortsetzen, aber der Schritt zur erzeugten σ -Algebra ist nicht offensichtlich.

Der Weg, den wir in diesem Abschnitt einschlagen werden, lässt sich folgendermaßen beschreiben: Zuerst konstruieren wir aus ν ein sogenanntes *äußeres Maß* $\mu^* : \mathcal{P}(\Omega) \rightarrow [0, \infty]$ auf der ganzen Potenzmenge von Ω (das ist im Allgemeinen kein Maß). Mittels des Satzes von Carathéodory zeigen wir dann, dass es eine natürliche σ -Algebra $\mathcal{A}(\mu^*) \subset \mathcal{P}(\Omega)$ gibt, so dass die Einschränkung $\mu^*|_{\mathcal{A}(\mu^*)}$ ein Maß (sogar ein vollständiges Maß) ist. Der Satz von Hahn wird uns dann zeigen, dass im Falle eines σ -additiven Inhalts ν (also eines Prämaßes) die Inklusion $\sigma(\mathcal{H}) \subset \mathcal{A}(\mu^*)$ besteht, und wie gewünscht $\mu^*|_{\mathcal{A}(\mu^*)}$ eine Fortsetzung von ν ist.

In Anwendung auf das Elementarvolumen werden wir folgende wichtige Aussage erhalten:

Satz 1.27 (Satz vom Lebesgue-Maß). *Es gibt genau ein Maß λ_d auf der Borel- σ -Algebra $\mathcal{B}(\mathbb{R}^d)$, dass $\lambda_d(Q) = \text{Vol}_d(Q)$ für jeden Quader $Q \in \mathcal{Q}^d$ erfüllt. Dieses Maß wird das (d -dimensionale) Lebesgue-Maß genannt.*

Nachdem wir diesen Satz haben, werden wir die Eigenschaften des Lebesgue-Maßes studieren und es zur Integration verwenden. Die Details der Konstruktion werden dann keine große Rolle mehr spielen.

Sie können sich die oben skizzierte Idee der Konstruktion eines Maßes so vorstellen, dass wir eine auf ν angepasste σ -Algebra konstruieren werden, die echt größer sein kann als $\sigma(\mathcal{R})$. Betrachten Sie zB das Dirac-Maß δ_0 auf dem Ring $\mathcal{Q}(\mathbb{R}^d)$ aller Quadersummen im \mathbb{R}^d , so ist die erzeugte σ -Algebra "nur" die Borel- σ -Algebra $\mathcal{B}(\mathbb{R}^d)$, obwohl δ_0 auf der ganzen Potenzmenge $\mathcal{P}(\mathbb{R}^d)$ ein Maß definiert.

Definition 1.28. Sei Ω eine nicht leere Menge. Ein *äußeres Maß* ist eine Abbildung $\mu^* : \mathcal{P}(\Omega) \rightarrow [0, \infty]$ mit den drei Eigenschaften:

- a) Striktheit: $\mu^*(\emptyset) = 0$.
- b) Monotonie: $A \subset B \subset \Omega \Rightarrow \mu^*(A) \leq \mu^*(B)$.
- c) σ -Subadditivität: Für jede Folge $(A_n)_{n \in \mathbb{N}}$ in $\mathcal{P}(\Omega)$ gilt

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

Wir haben im vorigen Abschnitt gesehen, dass jedes Maß auf der ganzen Potenzmenge insbesondere ein äußeres Maß ist, die Umkehrung gilt im Allgemeinen aber nicht. Das Sternchen in μ^* ist Bestandteil des Namens des äußeren Maßes und soll daran erinnern, dass μ^* kein Maß zu sein braucht (das wir üblicherweise mit μ bezeichnen).

Beispiel 1.29. Die Abbildung

$$\mu^* : \mathcal{P}(\mathbb{R}) \rightarrow [0, \infty], \quad \mu^*(A) := \begin{cases} 0 & A = \emptyset \\ 1 & A \neq \emptyset \end{cases}$$

ist ein äußeres Maß (leichter Check aller drei Eigenschaften), aber kein Maß, denn $\mu^*({1, 2}) = 1 \neq 1 + 1 = \mu^*({1}) + \mu^*({2})$ (nicht additiv).

Obwohl äußere Maße meist keine Maße sind, werden sie sich als ein nützliches Konzept zur Konstruktion von Maßen herausstellen. Wir zeigen zunächst, dass jeder Inhalt auf einem Halbring ein äußeres Maß definiert.

Lemma 1.30. *Sei $\nu : \mathcal{H} \rightarrow [0, \infty]$ ein Inhalt auf einem Halbring $\mathcal{H} \subset \mathcal{P}(\Omega)$. Dann ist*

$$\mu^* : \mathcal{P}(\Omega) \rightarrow [0, \infty]$$

$$\mu^*(A) := \inf \left\{ \sum_{n=1}^{\infty} \nu(B_n) : (B_n)_n \text{ Folge in } \mathcal{H} \text{ mit } A \subset \bigcup_{n=1}^{\infty} B_n \right\}$$

Die Definition von μ^* erklärt den Begriff “äußeres Maß”: Man überdeckt die fragliche Menge A mit einem abzählbaren System von Mengen $B_n \in \mathcal{H}$ und sucht dann nach dem Infimum von $\sum_n \nu(B_n)$, was einer Approximation von A von außen entspricht. Frühere Zugänge zur Maßtheorie verwendeten zusätzlich ein “inneres Maß” μ_* , was in moderneren Zugängen aber nicht mehr erforderlich ist und bei uns deshalb nicht auftreten wird.

Beachten Sie, dass das Infimum in $[0, \infty]$ per Definition $\inf \emptyset = \infty$ erfüllt.

Beweis. Die Striktheit und Monotonie von μ^* sind direkt aus der Definition ersichtlich. Für die σ -Additivität betrachten wir eine beliebige Folge $(A_n)_n$ von Teilmengen von Ω und müssen $\mu^*(\bigcup_n A_n) \leq \sum_n \mu^*(A_n)$ zeigen. Dies ist offensichtlich richtig, wenn die rechte Seite gleich $+\infty$ ist, dh wir dürfen $\sum_n \mu^*(A_n) < \infty$ annehmen. Insbesondere $\mu^*(A_n) < \infty$ für alle n , dh jedes A_n lässt sich durch eine Folge in \mathcal{H} abdecken, $A_n \subset \bigcup_{k \in \mathbb{N}} B_{n,k}$ mit $B_{n,k} \in \mathcal{H}$ (siehe Def. von μ^*). Nach der Definition des Infimums bedeutet das, dass wir für jedes $\varepsilon > 0$ die Mengen $B_{n,k} \in \mathcal{H}$ so wählen können, dass

$$\sum_{k=1}^{\infty} \nu(B_{n,k}) \leq \mu^*(A_n) + \frac{\varepsilon}{2^n}, \quad n \in \mathbb{N}.$$

Wir haben $\bigcup_n A_n \subset \bigcup_{n,k} B_{n,k}$ mit $B_{n,k} \in \mathcal{H}$. Benutzen wir die Monotonie von μ^* und dann $\mu^*(\bigcup_{n,k} B_{n,k}) \leq \sum_{k,n} \nu(B_{n,k})$ (Def. von μ^*), so erhalten wir

$$\begin{aligned} \mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) &\leq \mu^*\left(\bigcup_{n=1}^{\infty} \bigcup_{k=1}^{\infty} B_{n,k}\right) \leq \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \nu(B_{n,k}) \\ &\leq \sum_{n=1}^{\infty} \left(\mu^*(A_n) + \frac{\varepsilon}{2^n}\right) = \sum_{n=1}^{\infty} \mu^*(A_n) + \varepsilon. \end{aligned}$$

Im Limes $\varepsilon \rightarrow 0$ ergibt sich die σ -Subadditivität von μ^* . □

Beispiel 1.31. Das **äußere Lebesgue-Maß** ist das aus dem Elementarvolumen $\text{Vol}_d : \mathcal{Q}^d \rightarrow [0, \infty]$ konstruierte äußere Maß auf \mathbb{R}^d . Es wird meist mit λ^* bezeichnet. Das äußere Lebesgue-Maß ist nicht additiv, also kein Maß (ohne Beweis).

Satz 1.32. Sei Q ein (offener, abgeschlossener, oder halboffener) Quader und λ^* das äußere Lebesgue-Maß. Dann gilt $\lambda^*(Q) = \text{Vol}(Q)$.

Beweis. Sei $Q \in \mathcal{Q}^d$ und $\varepsilon > 0$. Dann gibt es Quader $A, B \in \mathcal{Q}^d$, so dass⁶ $A \subset \bar{A} \subset Q \subset B^\circ \subset B$ und

$$\text{Vol}(A) + \varepsilon \geq \text{Vol}(Q) \geq \text{Vol}(B) - \varepsilon \tag{*}$$

⁶Hier bezeichnet \bar{A} den Abschluss von A und B° das Innere von B , siehe Analysis 2. Für Quader gilt zB $\overline{[a, b]} = [a, b]$ und $[a, b]^\circ = (a, b)$.

(Beweis als leichte Übung). Wir betrachten die Folge $(Q_n)_n \subset \mathcal{Q}^d$, die durch $Q_1 := B$, $Q_n := \emptyset$ für $n > 1$, definiert ist. Dann gilt $Q \subset \bigcup_n Q_n$, und nach Definition von λ^* folgt $\lambda^*(Q) \leq \sum_n \text{Vol}_d(Q_n) = \text{Vol}(B) \leq \text{Vol}(Q) + \varepsilon$. Dies zeigt $\lambda^*(Q) \leq \text{Vol}(Q)$, da $\varepsilon > 0$ beliebig war.

Für die umgekehrte Ungleichung wählen wir wieder $\varepsilon > 0$. Nach Definition von λ^* gibt es eine Folge $(Q_n)_n \subset \mathcal{Q}^d$, so dass $Q \subset \bigcup_n Q_n$ und $\sum_n \text{Vol}(Q_n) \leq \lambda^*(Q) + \varepsilon$. Anwendung von (\star) auf jedes Q_n liefert uns Quader $A, B_n \in \mathcal{Q}^d$ mit

$$\begin{aligned} \bar{A} &\subset Q, & \text{Vol}(Q) &\leq \text{Vol}(A) + \varepsilon, \\ Q_n &\subset B_n^\circ, & \text{Vol}(B_n) &\leq \text{Vol}(Q_n) + \frac{\varepsilon}{2^n}. \end{aligned}$$

Insbesondere haben wir $\bar{A} \subset \bigcup_n B_n^\circ$, dh die (B_n°) bilden eine offene Überdeckung der kompakten Menge \bar{A} . Da \bar{A} kompakt ist, gibt es eine endliche Teilüberdeckung, dh $\bar{A} \subset \bigcup_{n=1}^N B_n$.

Nun benutzen wir noch, dass Vol als Inhalt auf einem Halbring monoton und subadditiv ist:

$$\begin{aligned} \text{Vol}(Q) &\leq \text{Vol}(A) + \varepsilon \leq \text{Vol}\left(\bigcup_{n=1}^N B_n\right) + \varepsilon \leq \sum_{n=1}^N \text{Vol}(B_n) + \varepsilon \\ &\leq \sum_{n=1}^N \left(\text{Vol}(Q_n) + \frac{\varepsilon}{2^n}\right) + \varepsilon \leq \sum_{n=1}^N \text{Vol}(Q_n) + 2\varepsilon \leq \lambda^*(Q) + 3\varepsilon. \end{aligned}$$

Im Grenzübergang $\varepsilon \rightarrow 0$ folgt $\text{Vol}(Q) \leq \lambda^*(Q)$. □

Korollar 1.33. Das Elementarvolumen $\text{Vol}_d : \mathcal{Q}^d \rightarrow [0, \infty)$ ist ein Prämaß für jede Dimension $d \in \mathbb{N}$.

Beweis. Nach Satz 1.26 ist Vol_d ein Inhalt. Da Vol_d auf \mathcal{Q}^d mit λ^* übereinstimmt, ist Vol_d auch σ -subadditiv. Also ist Vol_d nach Satz 1.25 ein Prämaß. □

Zu diesem Zeitpunkt wissen wir, dass jeder Inhalt ein äußeres Maß definiert. Ein wesentlicher Schritt besteht jetzt darin, von diesem äußeren Maß zu einem echten Maß auf einer σ -Algebra überzugehen. Dazu müssen wir eine Menge $\mathcal{A}(\mu^*) \subset \mathcal{P}(\Omega)$ von messbaren Mengen in Ω identifizieren.

Definition 1.34. Sei $\mu^* : \mathcal{P}(\Omega) \rightarrow [0, \infty]$ ein äußeres Maß. Dann heißt $A \subset \Omega$ μ^* -messbar, falls

$$\mu^*(X) = \mu^*(X \cap A) + \mu^*(X \cap A^c)$$

für alle $X \in \mathcal{P}(\Omega)$ gilt. Die Menge aller μ^* -messbaren Mengen wird mit $\mathcal{A}(\mu^*)$ bezeichnet.

Die Messbarkeitsbedingung von Carathéodory besagt, dass eine messbare Menge A die folgende Eigenschaft haben sollte: Wird eine beliebige Menge X durch A in die zwei disjunkten Teile $X \cap A$ und $X \cap A^c$ geteilt, so sollte das äußere Maß bzgl. dieser Teilung additiv sein.

Dies erinnert stark an Additivität (wir müssen $\mathcal{A}(\mu^*)$ so definieren, dass μ^* darauf σ -additiv ist), und die Bedingung ist offensichtlich symmetrisch in $A \leftrightarrow A^c$, und von \emptyset erfüllt. Auch wenn sich die Bedingung im Detail wegen der beliebigen Mengen X nur schwer anschaulich motivieren lässt, belegt der folgende Satz, dass sie genau die richtigen Konsequenzen hat.

Satz 1.35 (Maßerweiterungssatz von Carathéodory). *Sei μ^* ein äußeres Maß auf $\mathcal{P}(\Omega)$. Dann ist $\mathcal{A}(\mu^*)$ eine σ -Algebra und $\mu^* : \mathcal{A}(\mu^*) \rightarrow [0, \infty]$ ein vollständiges Maß.*

Beweis. Wir haben bereits oben erwähnt, dass $\mathcal{A}(\mu^*)$ die leere Menge enthält und abgeschlossen unter Komplementbildung $A \mapsto A^c$ ist.

Wir möchten nun zeigen, dass $\mathcal{A}(\mu^*)$ abgeschlossen unter Vereinigungen ist. Sei also $A, B \in \mathcal{A}(\mu^*)$ und $X \subset \Omega$. Dann gilt

$$\begin{aligned} \mu^*(X \cap (A \cup B)) + \mu^*(X \cap (A \cup B)^c) &= \mu^*((X \cap A) \cup (X \cap B)) + \mu^*(X \cap A^c \cap B^c) \\ &= \mu^*((X \cap A) \cup (X \cap A^c \cap B)) + \mu^*(X \cap A^c \cap B^c) \\ &\leq \mu^*(X \cap A) + \mu^*(X \cap A^c \cap B) + \mu^*(X \cap A^c \cap B^c), \end{aligned}$$

wobei wir die Subadditivität von μ^* benutzt haben. Da $A, B \in \mathcal{A}(\mu^*)$, haben wir

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

Eingesetzt in die erste Gleichung erhalten wir $\mu^*(X \cap (A \cup B)) + \mu^*(X \cap (A \cup B)^c) \leq \mu^*(X)$. Da die Umkehrung dieser Ungleichung aufgrund der Subadditivität von μ^* sowieso gilt, haben wir $A \cup B \in \mathcal{A}(\mu^*)$ gezeigt.

Für eine σ -Algebra brauchen wir aber abzählbare Vereinigungen. Als Schritt dorthin betrachten wir zunächst abzählbare disjunkte Vereinigungen, dh eine Folge $(A_n)_n$ paarweise disjunkter Mengen in $\mathcal{A}(\mu^*)$.

Wir setzen

$$B_n := \bigcup_{k=1}^n A_k, \quad B := \bigcup_{k=1}^{\infty} A_k = \bigcup_{n=1}^{\infty} B_n.$$

Aufgrund der schon gezeigten Stabilität von $\mathcal{A}(\mu^*)$ unter Vereinigungen folgt $B_n \in \mathcal{A}(\mu^*)$ für alle n . Wir müssen $B \in \mathcal{A}(\mu^*)$ zeigen.

Sei $X \subset \Omega$ beliebig. Wir behaupten für alle $n \in \mathbb{N}$

$$\mu^*(X \cap B_n) = \sum_{k=1}^n \mu^*(X \cap A_k),$$

und beweisen diese Behauptung per Induktion in n . Für $n = 1$ ist das trivial. Für den Induktionsschritt haben wir

$$\begin{aligned} \mu^*(X \cap B_{n+1}) &= \mu^*(X \cap B_{n+1} \cap B_n) + \mu^*(X \cap B_{n+1} \cap B_n^c) \quad (\text{da } B_n \in \mathcal{A}(\mu^*)) \\ &= \mu^*(X \cap B_n) + \mu^*(X \cap A_{n+1}) \\ &= \sum_{k=1}^{n+1} \mu^*(X \cap A_k). \end{aligned}$$

Das schließt den Induktionsbeweis ab.

Also haben wir für alle $n \in \mathbb{N}$ per Monotonie μ^*

$$\begin{aligned}\mu^*(X) &= \mu^*(X \cap B_n) + \mu^*(X \cap B_n^c) \\ &\geq \mu^*(X \cap B_n) + \mu^*(X \cap B^c) = \sum_{k=1}^n \mu^*(X \cap A_k) + \mu^*(X \cap B^c).\end{aligned}$$

Da diese Ungleichung für alle n gilt, erhalten wir im Grenzwert und per σ -Subadditivität von μ^*

$$\begin{aligned}\mu^*(X) &\geq \sum_{k=1}^{\infty} \mu^*(X \cap A_k) + \mu^*(X \cap B^c) & (\star) \\ &\geq \mu^*\left(\bigcup_{k=1}^{\infty} (X \cap A_k)\right) + \mu^*(X \cap B^c) \\ &= \mu^*(X \cap B) + \mu^*(X \cap B^c).\end{aligned}$$

Also ist $\mathcal{A}(\mu^*)$ unter abzählbaren disjunkten Vereinigungen abgeschlossen. Weiterhin ist $\mathcal{A}(\mu^*)$ wegen $A \cap B = (A^c \cup B^c)^c$ auch unter Schnitten abgeschlossen. Dies impliziert auch, dass $\mathcal{A}(\mu^*)$ unter Differenzen $A \cap B = A \cap B^c$ abgeschlossen ist.

Ist nun $(A_n)_n$ eine nicht notwendigerweise paarweise disjunkte Folge von Elementen in $\mathcal{A}(\mu^*)$, so setzen wir $B_n := \bigcup_{k=1}^n A_k$ (dies sind Elemente von $\mathcal{A}(\mu^*)$), und beachten $A_n = B_n \setminus B_{n-1}$ (mit $B_0 = \emptyset$). Dann ist $\bigcup_n A_n = \bigsqcup_n (B_n \setminus B_{n-1})$ eine abzählbare disjunkte Vereinigung von Elementen in $\mathcal{A}(\mu^*)$, also in $\mathcal{A}(\mu^*)$ enthalten.

Dies schließt den Beweis ab, dass $\mathcal{A}(\mu^*)$ eine σ -Algebra ist.

Wir betrachten noch einmal die Ungleichung (\star) , und zwar mit $X = B = \bigcup_{k=1}^{\infty} A_k$. Dies ergibt $\mu^*(\bigcup_{k=1}^{\infty} A_k) \geq \sum_{k=1}^{\infty} \mu^*(A_k)$. Da die umgekehrte Ungleichung aufgrund der σ -Subadditivität von μ^* sowieso gilt, haben wir hier sogar Gleichheit, dh μ^* ist auf $\mathcal{A}(\mu^*)$ σ -additiv und damit ein Maß.

Es bleibt die Vollständigkeit von $\mu^*|_{\mathcal{A}(\mu^*)}$ zu zeigen. Sei also $N \in \mathcal{A}(\mu^*)$ eine Nullmenge und $A \subset N$. Für beliebiges $X \subset \Omega$ gilt dann $X \cap A \subset N$, per Monotonie also $\mu^*(X \cap A) \leq \mu^*(N) = 0$ und deshalb

$$\mu^*(X) \geq \mu^*(X \cap A^c) = \mu^*(X \cap A) + \mu^*(X \cap A^c).$$

Da $\mu^*(X \cap A) + \mu^*(X \cap A^c) \leq \mu^*(X)$ ohnehin gilt, zeigt dies, dass A in $\mathcal{A}(\mu^*)$ liegt. Also ist $\mu^*|_{\mathcal{A}(\mu^*)}$ vollständig. \square

Der Satz von Carathéodory ist ein wesentlicher Schritt in unserer Konstruktion von Maßen – ausgehend von einem Inhalt auf einem Halbring haben wir ein Maß auf einer σ -Algebra konstruiert. Aber wie hängt dieses Maß mit dem anfänglich betrachteten Inhalt zusammen? Ist es überhaupt eine Erweiterung? Diese Fragen klärt der Satz von Hahn.

Satz 1.36 (Fortsetzungssatz von Hahn). Sei $\nu : \mathcal{H} \rightarrow [0, \infty]$ ein Inhalt auf einem Halbring \mathcal{H} , und $\mu^* : \mathcal{A}(\mu^*) \rightarrow [0, \infty]$ die Einschränkung des zugehörigen äußeren Maßes auf die σ -Algebra $\mathcal{A}(\mu^*)$. Dann gilt

- a) $\sigma(\mathcal{H}) \subset \mathcal{A}(\mu^*)$ und,
- b) $\mu^*|_{\mathcal{H}} = \nu$ genau dann, wenn ν ein Prämaß ist.

Beweis. a) Wir zeigen zunächst $\sigma(\mathcal{H}) \subset \mathcal{A}(\mu^*)$. Da \mathcal{H} und der von \mathcal{H} erzeugte Ring \mathcal{R} dieselbe σ -Algebra erzeugen, dürfen wir dabei annehmen, dass \mathcal{H} ein Ring ist.

Gegeben $A \in \mathcal{H}$ müssen wir die Carathéodory-Messbarkeitsbedingung zeigen, d.h.

$$\mu^*(X) \geq \mu^*(X \cap A) + \mu^*(X \cap A^c)$$

für alle $X \subset \Omega$. Wir dürfen annehmen, dass $\mu^*(X) < \infty$, denn sonst ist die Behauptung trivial. Für jedes $\varepsilon > 0$ gibt es laut Definition von μ^* eine Folge $(B_n)_n \subset \mathcal{H}$ mit

$$X \subset \bigcup_{n=1}^{\infty} B_n, \quad \sum_{n=1}^{\infty} \nu(B_n) < \mu^*(X) + \varepsilon,$$

also

$$X \cap A \subset \bigcup_{n=1}^{\infty} (B_n \cap A), \quad X \cap A^c \subset \bigcup_{n=1}^{\infty} (B_n \setminus A).$$

Da $B_n \cap A$ und $B_n \setminus A$ in \mathcal{H} liegen (hier benutzen wir, dass \mathcal{H} bereits ein Ring ist), sind dies Überdeckungen mit Elementen aus \mathcal{H} , wie in der Definition von μ^* . Dies impliziert per endlicher Additivität von ν

$$\begin{aligned} \mu^*(X \cap A) + \mu^*(X \cap A^c) &\leq \sum_{n=1}^{\infty} \nu(B_n \cap A) + \sum_{k=1}^{\infty} \nu(B_k \setminus A) \\ &= \sum_{n=1}^{\infty} (\nu(B_n \cap A) + \nu(B_n \setminus A)) = \sum_{n=1}^{\infty} \nu(B_n) < \mu^*(X) + \varepsilon. \end{aligned}$$

Da $\varepsilon > 0$ beliebig war, folgt die behauptete Ungleichung und damit $\sigma(\mathcal{H}) \subset \mathcal{A}(\mu^*)$.

b) Falls $\mu^*|_{\mathcal{H}} = \nu$, dann ist ν σ -subadditiv, da μ^* als äußeres Maß σ -subadditiv ist. Nach Satz 1.25 ist ν dann ein Prämaß.

Für die andere Richtung nehmen wir an, dass ν ein Prämaß ist und müssen für jedes $A \in \mathcal{H}$ die Identität $\mu^*(A) = \nu(A)$ zeigen. Die Ungleichung $\mu^*(A) \leq \nu(A)$ folgt automatisch (Def. von μ^*). Für die umgekehrte Ungleichung können wir $\mu^*(A) < \infty$ annehmen. Dann gibt es zu jedem $\varepsilon > 0$ eine Überdeckung $A \subset \bigcup_n B_n$, $B_n \in \mathcal{H}$, mit $\sum_n \nu(B_n) < \mu^*(A) + \varepsilon$.

Da $B_n \cap A \in \mathcal{H}$, folgt aus der σ -Subadditivität und Monotonie von ν

$$\nu(A) = \nu\left(\bigcup_{n=1}^{\infty} (B_n \cap A)\right) \leq \sum_{n=1}^{\infty} \nu(B_n \cap A) \leq \sum_{n=1}^{\infty} \nu(B_n) < \mu^*(A) + \varepsilon,$$

was für $\varepsilon \rightarrow 0$ die Behauptung liefert. □

Der Beweis des Satzes von Hahn sagt insbesondere, dass es für einen nicht σ -additiven Inhalt ν auf einem Halbring \mathcal{H} eine Menge $A \in \mathcal{H}$ gibt, die $\mu^*(A) < \nu(A)$ erfüllt.

Definition 1.37. Sei $\nu : \mathcal{H} \rightarrow [0, \infty]$ ein Prämaß auf einem Halbring \mathcal{H} . Dann liefern die Sätze von Carathéodory und Hahn eine \mathcal{H} enthaltende σ -Algebra $\mathcal{A}(\mu^*)$ und ein Maß $\mu = \mu^*|_{\mathcal{A}(\mu^*)} \rightarrow [0, \infty]$ mit $\mu|_{\mathcal{H}} = \nu$. Dieses Maß heißt die *Carathéodory-Hahn-Fortsetzung von ν* .

Für das Prämaß $\text{Vol}_d : \mathcal{Q}^d \rightarrow [0, \infty)$ bezeichnen wir das zugehörige äußere Maß mit λ_d^* und nennen $\mathcal{L}^d := \mathcal{A}(\lambda_d^*) \subset \mathcal{P}(\mathbb{R}^d)$ die *Lebesgue'sche σ -Algebra*. Die Teilmengen $A \in \mathcal{L}^d$ heißen *Lebesgue-messbar*, und $\lambda_d := \lambda_d^*|_{\mathcal{L}^d}$ heißt das *Lebesgue-Maß*.

Einige Bemerkungen zum Lebesgue-Maß:

- Das Lebesgue-Maß ist das wichtigste Maß in dieser Vorlesung.
- Wenn die Dimension d aus dem Kontext klar ist, schreiben wir auch kürzer $\mathcal{L}, \lambda, \lambda^*$.
- Da \mathcal{Q}^d die Borel- σ -Algebra $\mathcal{B}(\mathbb{R}^d)$ erzeugt, gilt $\mathcal{B}(\mathbb{R}^d) \subset \mathcal{L}^d$: Alle Borelmengen sind messbar. Insbesondere ist nun für alle offenen und alle abgeschlossenen Teilmengen von \mathbb{R}^d ein Volumen definiert.
- Wie groß sind diese σ -Algebren? Es lässt sich zeigen, dass $\mathcal{B}(\mathbb{R}^d)$ die gleiche Kardinalität wie \mathbb{R} hat. Die Lebesgue-Algebra \mathcal{L}^d hat hingegen die Kardinalität von $\mathcal{P}(\mathbb{R}^d)$, ist in diesem Sinne also unvergleichlich viel größer. Insbesondere $\mathcal{B}(\mathbb{R}^d) \neq \mathcal{L}^d$.
- Die Einschränkung $\lambda|_{\mathcal{B}(\mathbb{R}^d)}$ des Lebesgue-Maßes auf Borelmengen wird *Lebesgue-Borel-Maß* genannt. Das Lebesgue-Maß ist die Vervollständigung des Lebesgue-Borel-Maßes ist – dies folgt aus einem allgemeinen Satz zur Carathéodory-Hahn-Konstruktion, den wir hier aus Zeitgründen aber nicht besprechen wollen. Siehe aber Satz 1.39 für den Beweis, dass jede Lebesgue-messbare Menge A von der Form $A = B \cup N$ mit einer Borelmenge B und einer Lebesgue-Nullmenge N ist.
- Es gibt nicht Lebesgue-messbare Mengen, $\mathcal{L}^d \neq \mathcal{P}(\mathbb{R}^d)$. Der Beweis davon kommt später.

Aus Zeitgründen beweisen wir hier die Eindeutigkeitsaussage aus Satz 1.27 nicht. Sie gründet sich auf folgender Tatsache:

Satz 1.38 (Maßeindeutigkeitsatz). Sei (Ω, \mathcal{A}) ein Maßraum, $\mathcal{E} \subset \mathcal{A}$ ein schnittstabiler Erzeuger, und $\mu_1, \mu_2 : \mathcal{A} \rightarrow [0, \infty]$ Maße mit $\mu_1|_{\mathcal{E}} = \mu_2|_{\mathcal{E}}$. Falls $\mu_1|_{\mathcal{E}}$ die Eigenschaft hat:

- Es existieren $(A_n)_{n \in \mathbb{N}} \in \mathcal{E}$ mit $\Omega = \bigcup_{n=1}^{\infty} A_n$ und $\mu_1(A_n) < \infty$ für alle $n \in \mathbb{N}$ (σ -Endlichkeit),

so gilt $\mu_1 = \mu_2$.

Da \mathcal{Q}^d als Halbring schnittstabil und das Elementarvolumen Vol_d offenbar σ -endlich ist (betrachte zB die Quader $[-n, n]^d$), impliziert dieser Satz die Eindeutigkeitsaussage in Satz 1.27.

1.5 Das Lebesgue-Maß und messbare Funktionen

Nachdem wir das Lebesgue-Maß λ konstruiert haben, wollen wir einige seiner Eigenschaften ableiten. Wir beginnen mit einer Regularitätseigenschaft, die besagt, dass das Lebesgue-Maß $\lambda(A)$ von außen durch offene und von innen durch abgeschlossene Mengen approximiert werden kann.

Satz 1.39. Sei $A \subset \mathbb{R}^d$. Die folgenden Aussagen sind äquivalent:

- a) A ist Lebesgue-messbar, dh. $A \in \mathcal{L}$.
- b) Für jedes $\varepsilon > 0$ existiert eine offene Menge $U \subset \mathbb{R}^d$ mit $A \subset U$ und $\lambda^*(U \setminus A) < \varepsilon$ (Man sagt: λ ist von außen regulär).
- c) Für jedes $\varepsilon > 0$ existieren eine offene Menge $U \subset \mathbb{R}^d$ und eine abgeschlossene Menge $V \subset \mathbb{R}^d$, so dass $V \subset A \subset U$ und $\lambda(U \setminus V) < \varepsilon$ (Man sagt: λ ist von innen regulär).

Beweis. In den Hausaufgaben werden die Implikationen $a) \Rightarrow b)$ und $a) \Rightarrow c)$ gezeigt (für $d = 1$, aber der Beweis verallgemeinert sich leicht auf $d \in \mathbb{N}$).

$c) \Rightarrow b)$ Nach Voraussetzung gibt es offenes U und abgeschlossenes V mit $V \subset A \subset U$ und $\lambda^*(U \setminus V) = \lambda(U \setminus V) < \varepsilon$. Aufgrund der Monotonie des äußeren Maßes λ^* gilt $\lambda^*(U \setminus A) \leq \lambda^*(U \setminus V) < \varepsilon$.

$c) \Rightarrow a)$ Wir wählen $\varepsilon = \frac{1}{n}$, $n \in \mathbb{N}$, und finden nach Voraussetzung offene Mengen U_n und abgeschlossene Mengen V_n mit $V_n \subset A \subset U_n$ und $\lambda(U_n \setminus V_n) < \frac{1}{n}$. Sei $V := \bigcup_n V_n$ und $U := \bigcap_n U_n$. Sowohl U als auch V sind Borelmengen, und es gilt $V \subset A \subset U$ sowie $\lambda(U \setminus V) \leq \lambda(U_n \setminus V_n) < \frac{1}{n}$, also $\lambda(U \setminus V) = 0$. Deshalb ist $A = V \cup (A \setminus V)$ Vereinigung der Borelmenge V und der Teilmenge $A \setminus V$ der Nullmenge $U \setminus V$. Also ist $A \setminus V$ eine Nullmenge. Da das Lebesgue-Maß vollständig ist (Satz 1.35), ist $A \setminus V$ und damit auch $A = V \cup (A \setminus V)$ messbar. \square

In diesem Beweis haben wir insbesondere gesehen, dass jede Lebesguemenge Vereinigung einer Borelmenge mit einer Lebesguenullmenge ist. Wie bereits vorher erwähnt, ist das Lebesgue-Maß die Vervollständigung des Lebesgue-Borel-Maßes.

Beispiel 1.40. Wir illustrieren die Approximation messbarer Mengen von innen und außen an dem einfachen Beispiel eines Intervalls $A = [a, b)$, $a < b$. Gegeben $\varepsilon > 0$, so erfüllen die offene Menge $U = (a - \frac{\varepsilon}{4}, b)$ und die abgeschlossene Menge $V = [a, b - \frac{\varepsilon}{4}]$ offenbar $V \subset A \subset U$ und $\lambda(U \setminus A) = \lambda((a - \frac{\varepsilon}{4}, a)) = \frac{\varepsilon}{4} < \varepsilon$ und $\lambda(U \setminus V) = \lambda((a - \frac{\varepsilon}{4}, a) \sqcup (b - \frac{\varepsilon}{4}, b)) = \frac{\varepsilon}{2} < \varepsilon$.

Ein interessanteres Beispiel ist $A = \mathbb{Q} \subset \mathbb{R}$. Dies wird in den Übungen besprochen. Um zu verdeutlichen, dass es wichtig ist, dass wir *offene* Mengen für die Approximation von *außen*, und *abgeschlossene* Mengen für die Approximation von *innen* verwenden, versuchen wir es andersherum: Sei $U \subset \mathbb{Q}$ offen und $\mathbb{Q} \subset V$ abgeschlossen. Da jedes nicht leere offene Intervall sowohl rationale als auch irrationale Zahlen enthält, folgt $U = \emptyset$. Da jede reelle Zahl Grenzwert einer Folge von rationalen Zahlen ist, gilt $\overline{\mathbb{Q}} =$

\mathbb{R} und damit $V = \mathbb{R}$. Die Approximationseigenschaften in Teil b) und c) des Satzes scheitern also.

Bei unseren anfänglichen Überlegungen in Abschnitt 1.1 hatten wir *Bewegungsinvarianz* als eine relevante Eigenschaft eines unserer geometrischen Anschauung entsprechenden Volumens identifiziert. In den späteren allgemeinen Maßkonstruktionen ist Bewegungsinvarianz aber bisher nicht aufgetreten. Wir wollen diesen Punkt nun angehen, und gleich etwas allgemeiner das Verhalten des Lebesgue-Maßes unter geeigneten Abbildungen betrachten. Dies führt uns auf den für die Maßtheorie zentralen Begriff einer messbaren Funktion.

Definition 1.41 (Messbare Funktionen). Seien (Ω, \mathcal{A}) und $(\tilde{\Omega}, \tilde{\mathcal{A}})$ zwei Messräume. Eine Abbildung $f : \Omega \rightarrow \tilde{\Omega}$ heißt *messbar*, wenn die Urbilder messbarer Mengen messbar sind, dh wenn $f^{-1}(\tilde{A}) \in \mathcal{A}$ für alle $\tilde{A} \in \tilde{\mathcal{A}}$.

Wenn die σ -Algebren betont werden müssen, sprechen wir genauer von $(\mathcal{A}, \tilde{\mathcal{A}})$ -Messbarkeit.

Beachten Sie, dass für den Begriff einer messbaren Funktion, genau wie für den Begriff einer messbaren Menge, kein Maß benötigt wird. So wie uns stetige Funktionen viel in Analysis 1 und 2 beschäftigt haben, werden uns messbare Funktionen viel in Analysis 3 beschäftigen; insbesondere im Zusammenhang mit Integration.

Beispiel 1.42.

- a) Konstante Funktionen sind messbar: Sei $f : \Omega \rightarrow \tilde{\Omega}$, $f(x) = c$ für alle $x \in \Omega$. Denn ist $\tilde{A} \in \tilde{\mathcal{A}}$ messbar, so ist

$$f^{-1}(\tilde{A}) = \{x \in \Omega : f(x) \in \tilde{A}\} = \{x \in \Omega : c \in \tilde{A}\}$$

entweder \emptyset (messbar) oder Ω (messbar).

- b) Charakteristische Funktionen messbarer Mengen sind messbar: Sei $M \subset \Omega$, und sei $\chi_M : \Omega \rightarrow \mathbb{R}$ die zugehörige charakteristische Funktion (Wir betrachten hier \mathbb{R} als Messraum $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ mit der Borel- σ -Algebra). Dann ist $\chi_M^{-1}(A) \in \mathcal{A}$ für alle $A \in \mathcal{B}(\mathbb{R})$ und wir sehen, dass χ_M genau dann als Funktion messbar ist, wenn $M \subset \Omega$ als Teilmenge messbar ist.

Eine Funktion $f : X \rightarrow Y$ zwischen zwei metrischen Räumen X, Y (z.B. $f : \mathbb{R} \rightarrow \mathbb{R}$ oder $f : [0, 1] \rightarrow \mathbb{R}^3$) nennen wir messbar (oder genauer: *Borel-messbar*), wenn auf Definitionsbereich und Bildbereich die jeweilige Borel- σ -Algebra verwendet wird.

Das Beispiel der charakteristischen Funktionen hat bereits gezeigt, dass messbare Funktionen nicht stetig sein müssen. Das beinhaltet auch "sehr unstetige" Funktionen wie zB die messbare *Dirichlet-Funktion*

$$\chi_{\mathbb{Q}} : \mathbb{R} \rightarrow \mathbb{R}, \quad \chi_{\mathbb{Q}}(x) = \begin{cases} 1 & x \text{ rational} \\ 0 & x \text{ irrational} \end{cases}$$

Andererseits sind stetige Funktionen stets messbar. Um diesen Punkt zu verstehen, arbeiten wir nicht mit der ε - δ -Charakterisierung von Stetigkeit, sondern viel einfacher mit der Charakterisierung von Stetigkeit durch offene Mengen (siehe Analysis 1, Satz 4.15 b)):

Satz 1.43. Eine Funktion $f : X \rightarrow Y$ zwischen zwei metrischen Räumen X, Y ist genau dann stetig, wenn für jede offene Menge $U \subset Y$ das Urbild $f^{-1}(U) \subset X$ offen ist.

Aufgrund der Äquivalenzaussage in diesem Satz können wir Stetigkeit von Funktionen also auf eine ganz ähnliche Art und Weise wie Messbarkeit von Funktionen durch Urbilder zu definieren. Da im Folgenden oft Urbilder vorkommen, einige Kommentare zur Notation: Gegeben zwei Mengen $\Omega, \tilde{\Omega}$ und eine Abbildung $f : \Omega \rightarrow \tilde{\Omega}$, so schreiben wir $f(A), f^{-1}(\tilde{A})$ für $A \subset \Omega$ und $\tilde{A} \subset \tilde{\Omega}$ wie gewohnt für Bilder und Urbilder. Analog schreiben wir für Mengensysteme $\mathcal{M} \subset \mathcal{P}(\Omega)$ und $\tilde{\mathcal{M}} \subset \mathcal{P}(\tilde{\Omega})$ auch

$$f(\mathcal{M}) := \{f(M) : M \in \mathcal{M}\} \subset \mathcal{P}(\tilde{\Omega}),$$

$$f^{-1}(\tilde{\mathcal{M}}) := \{f^{-1}(\tilde{M}) : \tilde{M} \in \tilde{\mathcal{M}}\} \subset \mathcal{P}(\Omega).$$

Mit dieser Notation schreibt sich $(\mathcal{A}, \tilde{\mathcal{A}})$ -Messbarkeit einer Funktion $f : \Omega \rightarrow \tilde{\Omega}$ einfach als $f^{-1}(\tilde{\mathcal{A}}) \subset \mathcal{A}$.

Seien $\Omega, \tilde{\Omega}$ nicht leere Mengen und $f : \Omega \rightarrow \tilde{\Omega}$ eine Abbildung. Zeigen Sie:

- Für jede Indexmenge I und beliebige $A, A_i \subset \Omega, i \in I$ gilt $f^{-1}(\bigcup_{i \in I} A_i) = \bigcup_{i \in I} f^{-1}(A_i)$ und $f^{-1}(\bigcap_{i \in I} A_i) = \bigcap_{i \in I} f^{-1}(A_i)$ sowie $f^{-1}(A^c) = f^{-1}(A)^c$ (Erinnerung an das 1. Semester). Was geht schief, wenn Sie f statt f^{-1} verwenden?
- $f^{-1}(\sigma(\tilde{\mathcal{M}})) = \sigma(f^{-1}(\tilde{\mathcal{M}}))$ für jedes Mengensystem $\tilde{\mathcal{M}} \subset \mathcal{P}(\tilde{\Omega})$.

Satz 1.44. Seien X, Y metrische Räume und $f : X \rightarrow Y$ stetig. Dann ist f Borel-messbar.

Beweis. Sei \mathcal{O}_Y die Menge aller offenen Teilmengen von Y . Dann haben wir nach Definition der Borel- σ -Algebra $\mathcal{O}_Y \subset \sigma(\mathcal{O}_Y) = \mathcal{B}(Y)$. Da f stetig ist, gilt für jedes $O \in \mathcal{O}_Y$ auch $f^{-1}(O) \in \mathcal{O}_X \subset \mathcal{B}(X)$, dh Urbilder offener Mengen sind messbar, und $f^{-1}(\mathcal{O}_Y) \subset \mathcal{O}_X$. Mit der obenstehenden Übung folgt

$$f^{-1}(\mathcal{B}(Y)) = f^{-1}(\sigma(\mathcal{O}_Y)) = \sigma(f^{-1}(\mathcal{O}_Y)) \subset \sigma(\mathcal{O}_X) = \mathcal{B}(X).$$

□

Dieser Satz gibt uns eine große Menge von messbaren Funktionen, was vor allem für die Integrationstheorie von Interesse sein wird. Zunächst verwenden wir messbare Funktionen dazu, ein Maß von einer σ -Algebra zu einer anderen zu transportieren.

Definition und Satz 1.45. Seien $(\Omega, \mathcal{A}), (\tilde{\Omega}, \tilde{\mathcal{A}})$ zwei Messräume, $\mu : \mathcal{A} \rightarrow [0, \infty]$ ein Maß, und $f : \Omega \rightarrow \tilde{\Omega}$ eine messbare Funktion. Dann ist

$$f_*\mu : \tilde{\mathcal{A}} \rightarrow [0, \infty], \quad (f_*\mu)(\tilde{A}) := \mu(f^{-1}(\tilde{A})),$$

ein Maß auf $\tilde{\mathcal{A}}$. Das Maß $f_*\mu$ heißt das *Bildmaß* von μ unter f .

Die Notation f_* tritt an verschiedenen Stellen in der Mathematik auf und bezeichnet einen “pushforward” (meist nicht auf Deutsch übersetzt), mit dem eine Struktur (hier ein Maß) vom Definitions- auf den Bildbereich einer Abbildung $f : \Omega \rightarrow \tilde{\Omega}$ transportiert wird. Das dazu duale Konzept ist der “pullback” (Rücktransport) f^* , wir werden etwas später darauf eingehen.

Beweis. Die Abbildung $f_*\mu$ ist wohldefiniert, denn aufgrund der Messbarkeit von f gilt $f^{-1}(\tilde{A}) \in \mathcal{A}$ für alle $\tilde{A} \in \tilde{\mathcal{A}}$. Wir müssen zeigen, dass $f_*\mu$ tatsächlich ein Maß ist.

$(f_*\mu)(\emptyset) = \mu(f^{-1}(\emptyset)) = \mu(\emptyset) = 0$, also ist $f_*\mu$ strikt. Für die σ -Additivität betrachten wir paarweise disjunkte messbare Mengen $\tilde{A}_n \in \tilde{\mathcal{A}}$, $n \in \mathbb{N}$. Dann sind auch die Urbilder $f^{-1}(\tilde{A}_n)$ paarweise disjunkt und wir erhalten

$$(f_*\mu) \left(\bigsqcup_{n=1}^{\infty} \tilde{A}_n \right) = \mu \left(f^{-1} \left(\bigsqcup_{n=1}^{\infty} \tilde{A}_n \right) \right) = \mu \left(\bigsqcup_{n=1}^{\infty} f^{-1}(\tilde{A}_n) \right) = \sum_{n=1}^{\infty} \mu \left(f^{-1}(\tilde{A}_n) \right),$$

was mit $\sum_{n=1}^{\infty} (f_*\mu)(\tilde{A}_n)$ übereinstimmt. □

Genau wie Kompositionen stetiger Funktionen stetig sind, sind auch Kompositionen messbarer Funktionen messbar. Es gilt:

Lemma 1.46. *Seien $(\Omega_1, \mathcal{A}_1)$, $(\Omega_2, \mathcal{A}_2)$, $(\Omega_3, \mathcal{A}_3)$ Messräume sowie $f : \Omega_1 \rightarrow \Omega_2$ und $g : \Omega_2 \rightarrow \Omega_3$ messbar. Dann ist auch $g \circ f : \Omega_1 \rightarrow \Omega_3$ messbar. Ist μ ein Maß auf \mathcal{A}_1 , so gilt $(g \circ f)_*\mu = g_*(f_*\mu)$.*

Der Beweis reduziert sich auf ein bloßes Einsetzen der Definitionen und wird deshalb als Übung überlassen.

Wir haben nun alle Werkzeuge parat, um die Bewegungsinvarianz des Lebesgue-Maßes zu beweisen. Wir beginnen mit dem einfachen Fall der *Translationen*: Für $y \in \mathbb{R}^d$ sei

$$\tau_y : \mathbb{R}^d \rightarrow \mathbb{R}^d, \quad \tau_y(x) := x + y,$$

die Translation um y . Offenbar ist τ_y stetig (also insbesondere Borel-messbar) und bijektiv mit Umkehrabbildung $\tau_y^{-1} = \tau_{-y}$, und es gilt $\tau_y\tau_z = \tau_{y+z}$ sowie $\tau_0 = \text{id}_{\mathbb{R}^d}$ (die Translationen bilden eine eine zu $(\mathbb{R}, +)$ isomorphe Gruppe).

Bilder von Teilmengen $M \subset \mathbb{R}^d$ unter Translation schreiben wir kürzer als $M + y := \tau_y(M) = \{m + y : m \in M\}$.

Satz 1.47.

- a) *Das Lebesgue-Borel-Maß ist translationsinvariant: $(\tau_y)_*\lambda = \lambda$ für alle $y \in \mathbb{R}^d$. Auch das Lebesgue-Maß ist translationsinvariant.*
- b) *Sei μ ein translationsinvariantes Maß auf $\mathcal{B}(\mathbb{R}^d)$ mit $\mu([0, 1]^d) = m < \infty$. Dann gilt $\mu = m \cdot \lambda$.*

Beweis. a) Sei $Q := [a, b) \in \mathcal{Q}^d$ ein Quader, und $y \in \mathbb{R}^d$. Dann haben wir

$$((\tau_y)_*\lambda)(Q) = \lambda(Q - y) = \lambda([a - y, b - y)) = \prod_{k=1}^d (b_k - a_k) = \lambda(Q).$$

Nach dem Maßeindeutigkeitssatz 1.38 folgt nun $\lambda = (\tau_y)_*\lambda$ auf $\mathcal{B}(\mathbb{R}^d)$. Für den Beweis der Translationsinvarianz der Vervollständigung, also des Lebesgue-Maßes, verweisen wir auf das Buch von Elstrodt (Korollar II.6.5).

b) Wir betrachten $n_1, \dots, n_d \in \mathbb{N}$ und die disjunkte Zerlegung

$$[0, 1)^d = \bigsqcup_{k_1=1}^{n_1} \dots \bigsqcup_{k_d=1}^{n_d} \left[\frac{k_1-1}{n_1}, \frac{k_1}{n_1} \right) \times \dots \times \left[\frac{k_d-1}{n_d}, \frac{k_d}{n_d} \right) .$$

Aufgrund von Translationsinvarianz und Additivität von μ folgt

$$\mu \left(\left[0, \frac{1}{n_1} \right) \times \dots \times \left[0, \frac{1}{n_d} \right) \right) = \frac{m}{n_1 \cdot \dots \cdot n_d} .$$

Jeder Quader Q mit rationalen Eckpunkten ist eine disjunkte Vereinigung von Quadern dieser Form, so dass $\mu(Q) = m \cdot \lambda(Q)$ folgt. Die Behauptung folgt nun aus dem Maßeindeutigkeitssatz. \square

Unsere Definition des Lebesgue-Maßes basiert auf dem Volumen eines Quaders. Diese Quader wurden stets als achsenparallel, dh von der Produktform $[a_1, b_1) \times \dots \times [a_d, b_d)$, angenommen. Wie sieht das Volumen aus, wenn wir die Quader (oder beliebige messbare Mengen) drehen, spiegeln oder strecken? Um dies zu untersuchen, betrachten wir das Verhalten von λ unter linearen Abbildungen $\mathbb{R}^d \ni x \mapsto Tx$ mit einer $(d \times d)$ -Matrix $T \in M_d(\mathbb{R})$, die wir zunächst als invertierbar annehmen möchten. Da \mathbb{R}^d ein endlichdimensionaler Vektorraum ist, ist T stetig und insbesondere Borel-messbar, so dass wir das Bildmaß $T_*\lambda$ betrachten können.

Satz 1.48. Sei $T \in GL(d)$ eine invertierbare $(d \times d)$ -Matrix. Dann gilt für jede Borelmenge $B \subset \mathbb{R}^d$

$$(T_*\lambda)(B) = \lambda(T^{-1}B) = |\det T|^{-1} \lambda(B).$$

Beweis. Zuerst bemerken wir, dass $T_*\lambda$ ein translationsinvariantes Maß auf $\mathcal{B}(\mathbb{R}^d)$ ist: Für $x \in \mathbb{R}^d$ und $B \in \mathcal{B}(\mathbb{R}^d)$ gilt

$$(T_*\lambda)(B - x) = \lambda(T^{-1}B - T^{-1}(x)) = \lambda(T^{-1}B) = (T_*\lambda)(B).$$

Da der Einheitswürfel $[0, 1]^d$ kompakt und T^{-1} als lineare Abbildung stetig ist, ist auch $T^{-1}([0, 1]^d)$ kompakt (Satz 9.15 aus Analysis 2). Also gibt es einen beschränkten Quader $Q \supset T^{-1}([0, 1]^d) \supset T^{-1}([0, 1)^d)$, und es folgt

$$m_T := (T_*\lambda)([0, 1)^d) = \lambda(T^{-1}([0, 1)^d)) \leq \lambda(Q) < \infty.$$

Nach Satz 1.47 gilt dann $T_*\lambda = m_T \cdot \lambda$. Wir müssen $m_T = |\det T|^{-1}$ zeigen.

Wir betrachten zunächst den Spezialfall einer Diagonalmatrix $T = \text{diag}(t_1, \dots, t_d)$ mit reellen Zahlen $t_k \neq 0$. Dann streckt T^{-1} die x_k -Koordinate um $\frac{1}{t_k}$, und es ergibt sich $m_T = \frac{1}{|t_1| \cdots |t_d|} = |\det T|^{-1}$, wie behauptet.

Als nächsten Spezialfall betrachten wir eine orthogonale Matrix T , d.h. $\|Tx\|_2 = \|x\|_2$ für alle $x \in \mathbb{R}^d$ (äquivalent: $T^{-1} = T^t$). Dann gilt offenbar $T^{-1}(B_1) = B_1$ für die Einheitskugel $B_1 = \{x \in \mathbb{R}^d : \|x\|_2 \leq 1\}$, und deshalb $m_T = \frac{\lambda(T^{-1}(B_1))}{\lambda(B_1)} = 1 = |\det T|^{-1}$ (hier haben wir die aus Lineare Algebra bekannte Tatsache benutzt, dass $|\det T| = 1$ für orthogonale Matrizen T gilt).

Um diese Spezialfälle zu erweitern, bemerken wir als nächstes die Multiplikationsformel $m_{TS} = m_T m_S$ für $T, S \in \text{GL}(d)$ (die ja auch für die Determinante gilt). Denn es gilt für jede Borelmenge B mit Lemma 1.46

$$\begin{aligned} m_{TS} \cdot \lambda(B) &= (TS)_* \lambda(B) = (T_*(S_* \lambda))(B) = (T_*(m_S \cdot \lambda))(B) \\ &= m_S (T_* \lambda)(B) = m_S m_T \cdot \lambda(B). \end{aligned}$$

Eine beliebige Matrix $T \in \text{GL}(d)$ kann immer in der Form $T = U_1 D U_2$ mit orthogonalen Matrizen U_1, U_2 und einer diagonalen Matrix D geschrieben werden (siehe Lineare Algebra 2).

Also gilt

$$m_T = m_{U_1 D U_2} = m_{U_1} m_D m_{U_2} = m_D = |\det D|^{-1} = |\det(U_1 D U_2)|^{-1} = |\det T|^{-1},$$

und der Beweis ist abgeschlossen. \square

Bemerkung: Ersetzen Sie in dem Satz T durch T^{-1} , so erhalten Sie auch die Formel

$$\lambda(T(B)) = |\det T| \cdot \lambda(B)$$

für beliebige Borelmengen $B \in \mathcal{B}(\mathbb{R}^d)$ und beliebige invertierbare Matrizen $T \in \text{GL}(d)$.

Beispiel 1.49.

- a) (**Zentralstreckungen**). Sei $c > 0$ und $S_c : \mathbb{R}^d \rightarrow \mathbb{R}^d, x \mapsto c \cdot x$, die Streckung um c . Dann gilt für jede Borelmenge $B \subset \mathbb{R}^d$

$$\lambda(S_c(B)) = \lambda(cB) = c^d \cdot \lambda(B),$$

denn S_c wird durch die Matrix $c \cdot 1$ dargestellt, die Determinante c^d hat.

- b) (**Volumen eines Parallelotops**) Seien x_1, \dots, x_d linear unabhängige Vektoren in \mathbb{R}^d . Das von diesen Vektoren aufgespannte Parallelotop ist per Definition die Menge

$$P = \left\{ \sum_{k=1}^d t_k x_k : 0 \leq t_k \leq 1 \right\}.$$

Wir sehen, dass P das Bild des Einheitswürfels $[0, 1]^d$ unter der Matrix A ist, die die Standardbasisvektoren e_1, \dots, e_d auf die x_k abbildet, $Ae_k = x_k$. Also ist P

abgeschlossen, damit messbar, und A invertierbar (linear unabhängige Spalten). Das Volumen des Parallelotops ist $\lambda(P) = |\det A|$.

Satz 1.48 stimmt für nicht invertierbares T nicht, da dann $\det T = 0$ gilt. In diesem Fall ist $T\mathbb{R}^d \subset \mathbb{R}^d$ ein Unterraum mit $\dim(T\mathbb{R}^d) < d$, und für Borelmengen $B \subset (T\mathbb{R}^d)^c$ gilt $(T_*\lambda)(B) = \lambda(T^{-1}(B)) = \lambda(\emptyset) = 0$. Nichtsdestotrotz gilt $\lambda(T(B)) = 0 = |\det T|\lambda(B)$ für alle $B \in \mathcal{L}$, da $T(\mathbb{R}^d)$ als Unterraum mit Dimension $< d$ eine Nullmenge ist (Übung).

Beispiel 1.50 (Beispiel einer nicht messbaren Menge). Wir geben ein Beispiel einer nicht Lebesgue-messbaren Teilmenge $M \subset \mathbb{R}$. (Die Konstruktion lässt sich auf \mathbb{R}^d verallgemeinern.)

Dazu betrachten wir auf dem Einheitsintervall $[0, 1]$ die Äquivalenzrelation $x \sim y : \Leftrightarrow x - y \in \mathbb{Q}$ (es lässt sich leicht prüfen, dass dies wirklich eine Äquivalenzrelation ist). Sei $\mathcal{K} := [0, 1] / \sim$ die Menge der Äquivalenzklassen. Wir wählen aus jeder Äquivalenzklasse $K \in \mathcal{K}$ genau einen Repräsentanten x_K aus, und definieren die *Vitali-Menge*

$$V := \{x_K : K \in \mathcal{K}\} \subset [0, 1].$$

Diese Konstruktion benutzt das *Auswahlaxiom*^a.

Wir zeigen jetzt, dass V nicht Lebesgue-messbar ist. Dazu bemerken wir zunächst, dass für rationale Zahlen $r, r' \in \mathbb{Q}$, $r \neq r'$,

$$(r + V) \cap (r' + V) = \emptyset$$

gilt: Denn jedes Element $x \in (r + V) \cap (r' + V)$ ist von der Form $x_K + r = x_{K'} + r' \Rightarrow x_K \sim x_{K'} \Rightarrow x_K = x_{K'} \Rightarrow r = r'$, ein Widerspruch.

Nun betrachten wir die abzählbare Menge $R := [-1, 1] \cap \mathbb{Q}$. Es gilt

$$[0, 1] \subset \bigsqcup_{r \in R} (r + V) \subset [-1, 2].$$

Beweis: Zu jedem $x \in [0, 1]$ gibt es ein rationales r und einen Repräsentanten x_K , so dass $x = r + x_K$, dh $|r| = |x - x_K| \leq 1$. Wegen $[0, 1] \supset V$ ergibt dies die Behauptung.

Angenommen, V wäre Lebesgue-messbar. Dann gilt aufgrund der σ -Additivität, Translationsinvarianz, und Additivität von λ

$$1 = \lambda([0, 1]) \leq \sum_{r \in R} \lambda(V) \leq \lambda([-1, 2]) = 3.$$

Da es eine solche Zahl $\lambda(V)$ nicht gibt, haben wir einen Widerspruch und schließen, dass V nicht Lebesgue-messbar ist. Insbesondere ist V keine Borelmenge.

^aDieses Axiom der Mengenlehre besagt: Zu jeder Menge \mathcal{A} von nicht leeren Mengen $A \in \mathcal{A}$ gibt es eine Auswahlfunktion, dh eine Abbildung $f : \mathcal{A} \rightarrow \bigcup_{A \in \mathcal{A}} A$ mit $f(A) \in A$ für alle $A \in \mathcal{A}$. Das Auswahlaxiom wird von den allermeisten MathematikerInnen akzeptiert, ist aber logisch unabhängig von den üblichen Axiomen der Zermelo-Fraenkel-Mengenlehre.

2 Integration messbarer Funktionen

Gegeben einen Maßraum $(\Omega, \mathcal{A}, \mu)$, werden wir im Folgenden ein Integral $\int_{\Omega} f d\mu$ für messbare Funktionen $f : \Omega \rightarrow \mathbb{R}$ (oder $f : \Omega \rightarrow \overline{\mathbb{R}}$) definieren, wobei \mathbb{R} mit der Borel- σ -Algebra ausgestattet ist. Im speziellen aber wichtigen Fall des Maßraums $(\mathbb{R}^d, \mathcal{L}(\mathbb{R}^d), \lambda_d)$ erhalten wir das sogenannte d -dimensionale Lebesgue-Integral.

Die wesentlichen Schritte der Konstruktion sind:

- Für charakteristische Funktionen messbarer Mengen definieren wir $\int_{\Omega} \chi_A d\mu := \mu(A)$. Das Integral hängt also stark davon ab, welches Maß zugrundeliegt - zB $\int_{\mathbb{R}} \chi_{[1,3]} d\lambda = 2$ für das Lebesgue-Maß λ auf \mathbb{R} , und $\int_{\mathbb{R}} \chi_{[1,3]} d\mu_0 = 0$ für das Dirac-Maß μ_0 auf \mathbb{R} .
- Für einfache Funktionen (endliche Linearkombinationen charakteristischer Funktionen messbarer Mengen) definieren wir $\int_{\Omega} \sum_{k=1}^n c_k \chi_{A_k} d\mu := \sum_{k=1}^n c_k \mu(A_k)$.
- Für allgemeine messbare Funktionen werden wir eine Zerlegung in positive und negative Anteile und eine Approximation durch einfache Funktionen verwenden.

Wir möchten dann die wesentlichen Eigenschaften des Integrals untersuchen, zB Linearität, Monotonie ($\int_{\Omega} f d\mu \leq \int_{\Omega} g d\mu$ für $f \leq g$), das Verhalten unter Limiten ($\lim_n \int_{\Omega} f_n d\mu \stackrel{?}{=} \int_{\Omega} \lim_n f_n d\mu$), und mit dem Riemann-Integral vergleichen.

Aus diesem Plan wird klar, dass wir zuerst etwas mehr Wissen über messbaren Funktionen erarbeiten sollten, bevor wir das Integral definieren.

2.1 Mehr über messbare Funktionen

In der Maßtheorie kommt es häufig vor, dass gewisse Größen wie z.B. Maße $\mu(A)$ unendlich sein können. Wir werden deshalb auch zulassen, dass die zu integrierenden Funktionen den Wert $f(x) = \infty$ annehmen können. Dies kommt insbesondere bei Grenzwerten von Funktionenfolgen leicht vor, zB

$$f_n(x) = n\chi_{[0, \frac{1}{n}]}(x) \xrightarrow{n \rightarrow \infty} \begin{cases} \infty & x = 0 \\ 0 & x \neq 0 \end{cases}.$$

Wir betrachten deshalb Funktionen der Form

$$f : \Omega \rightarrow \overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}.$$

Diese Funktionen werden *numerische Funktionen* genannt, da sie - bis auf die möglichen Werte $\pm\infty$ - zahlenwertig sind und nicht Werte in einer abstrakten Menge $\tilde{\Omega}$ annehmen. Auf der erweiterten reellen Achse $\overline{\mathbb{R}}$ haben wir die σ -Algebra $\mathcal{B}(\overline{\mathbb{R}})$, die aus allen Mengen der Form $B, B \cup \{\infty\}, B \cup \{-\infty\}, B \cup \{\pm\infty\}$ mit $B \in \mathcal{B}(\mathbb{R})$ besteht.

Eine oft hilfreiche Charakterisierung von messbaren Funktionen $f : \Omega \rightarrow \overline{\mathbb{R}}$ verwendet *Superniveaumengen* und *Subniveaumengen*, nämlich die Urbilder (mit $a \in \mathbb{R}$)

$$\{x \in \Omega : f(x) > a\} = f^{-1}((a, \infty]) =: \{f > a\}.$$

Ganz analog werden die Kurzschreibweisen $\{f < a\}, \{f \geq a\}, \{f \leq a\}$ definiert. Beachten Sie, dass dies alle Urbilder messbarer Mengen in $\mathcal{B}(\overline{\mathbb{R}})$ unter f sind. Für messbares $f : (\Omega, \mathcal{A}) \rightarrow (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ liegen $\{f > a\}, \{f < a\}, \{f \geq a\}, \{f \leq a\}$ allesamt in \mathcal{A} . Noch genauer können wir sagen:

Lemma 2.1. Sei (Ω, \mathcal{A}) ein Messraum. Eine numerische Funktion $f : \Omega \rightarrow \overline{\mathbb{R}}$ ist genau dann messbar, wenn eine der folgenden äquivalenten Bedingungen erfüllt ist:

- a) $\{f < a\} \in \mathcal{A}$ für alle $a \in \mathbb{R}$,
- b) $\{f > a\} \in \mathcal{A}$ für alle $a \in \mathbb{R}$,
- c) $\{f \leq a\} \in \mathcal{A}$ für alle $a \in \mathbb{R}$,
- d) $\{f \geq a\} \in \mathcal{A}$ für alle $a \in \mathbb{R}$.

Beweis. Wie bereits erläutert, impliziert die Messbarkeit von f a)–d). Für die Äquivalenz der Bedingungen a)–d) untereinander beachten wir $\{f \geq a\} = \bigcap_{n=1}^{\infty} \{f > a - \frac{1}{n}\}$ (also b) \Rightarrow d)), $\{f > a\} = \bigcup_{n=1}^{\infty} \{f \geq a + \frac{1}{n}\}$ (also d) \Rightarrow b)), $\{f < a\} = \{f \geq a\}^c$ (also d) \Leftrightarrow a)) und $\{f \leq a\} = \{f > a\}^c$ (also b) \Leftrightarrow c)).

Wir zeigen noch, dass d) die Messbarkeit von f impliziert. Da $f^{-1}([a, \infty])$ nach Voraussetzung messbar ist, und $f^{-1}(\sigma(\mathcal{E})) = \sigma(f^{-1}(\mathcal{E}))$ für jedes $\mathcal{E} \subset \mathcal{P}(\overline{\mathbb{R}})$, genügt es zu zeigen, dass die von den Intervallen $[a, \infty]$, $a \in \mathbb{R}$, erzeugte σ -Algebra Σ mit $\mathcal{B}(\overline{\mathbb{R}})$ übereinstimmt. Wegen $\{\infty\} = \bigcap_{n \in \mathbb{N}} [n, \infty]$ und $\{-\infty\} = \bigcap_{n \in \mathbb{N}} (\overline{\mathbb{R}} \setminus [-n, \infty])$ liegen $\{\pm\infty\}$ in Σ . Außerdem enthält Σ alle halboffenen Intervalle $[a, b) = [a, \infty] \setminus [b, \infty]$. Da diese $\mathcal{B}(\overline{\mathbb{R}})$ erzeugen, folgt $\Sigma = \mathcal{B}(\overline{\mathbb{R}})$. \square

Beispiel 2.2. Sei $p \in \mathbb{R}$. Für $p \geq 0$ ist die Potenzfunktion $x \mapsto |x|^p$ stetig, also messbar. Für $p < 0$ betrachten wir die Funktion

$$f : \mathbb{R} \rightarrow \overline{\mathbb{R}}, \quad f(x) := \begin{cases} |x|^p & x \neq 0 \\ \infty & x = 0 \end{cases},$$

und behaupten, dass f messbar (aber nicht stetig) ist. Dazu prüfen wir die Bedingung aus Lemma 2.1 a): Für beliebiges $a \in \mathbb{R}$ gilt

$$\{f < a\} = \{x \in \mathbb{R} : |x|^p < a\} = \begin{cases} \emptyset & a \leq 0 \\ (-\infty, -a^{1/p}) \cup (a^{1/p}, \infty) & a > 0 \end{cases},$$

und diese Mengen sind messbar (sogar offen).

Der Wert von f bei $x = 0$ spielt hier keine Rolle, genauso gut hätten wir $f(0) = y$ mit einem beliebigen $y \in \overline{\mathbb{R}}$ setzen können.

Ist $f : (\Omega, \mathcal{A}) \rightarrow \overline{\mathbb{R}}$ messbar, so sind für jedes $a \in \mathbb{R}$ die a -Niveaumenge $\{f = a\} := f^{-1}(\{a\}) = \{f \geq a\} \cap \{f \leq a\}$ messbar. Aber aus der Messbarkeit aller $\{f = a\}$, $a \in \mathbb{R}$, folgt nicht die Messbarkeit von f (Knobel-Übung, finden Sie ein Gegenbeispiel).

Wir betrachten nun einige Operationen, die messbare Funktionen in messbare Funktionen überführen.

Satz 2.3. Sei (Ω, \mathcal{A}) ein Messraum und $f, g : \Omega \rightarrow \overline{\mathbb{R}}$ messbare numerische Funktionen. Dann sind auch die folgenden Funktionen messbar:

- a) $c \cdot f$ für alle $c \in \mathbb{R}$,
- b) $\sup\{f, g\} : x \mapsto \max\{f(x), g(x)\}$ und $\inf\{f, g\} : x \mapsto \min\{f(x), g(x)\}$,
- c) der positive Teil f_+ von f , definiert als $f_+ := \sup\{f, 0\}$, und der negative Teil f_- von f , definiert als $f_- := \sup\{-f, 0\}$,
- d) $|f|^p$ für alle $p > 0$,
- e) $f + g$, falls überall definiert (d.h. es gibt kein $x \in \Omega$ mit $f(x) = \pm\infty$ und $g(x) = \mp\infty$),
- f) $f \cdot g$.

Beweis. a) und d) folgen aus der Messbarkeit von $x \mapsto c \cdot x$ und $x \mapsto |x|^p$. b) und e) wurden schon in den Übungen gezeigt, und c) ist ein Spezialfall von b). Für f) bemerken wir wie in Übung P4.4, dass $h : \Omega \rightarrow \mathbb{R}^2$, $h(x) := (f(x), g(x))$ messbar ist. Da auch $\mathbb{R}^2 \rightarrow \mathbb{R}$, $m : (x, y) \mapsto xy$ als stetige Funktion messbar ist, ist $fg = m \circ h$ als Komposition messbarer Funktionen messbar. \square

Positiv- und Negativteil einer Funktion sind Ihnen bereits aus Analysis 1 bekannt (siehe Analysis 1, Lemma 6.10). Erinnern Sie: f_{\pm} sind beide positiv, und es gilt $f = f_+ - f_-$ und $|f| = f_+ + f_-$. Insbesondere sehen wir als Korollar von Satz 2.3, dass eine numerische Funktion $f : \Omega \rightarrow \overline{\mathbb{R}}$ genau dann messbar ist, wenn f_+ und f_- messbar sind.

Das Integral nach einem Maß wird durch einen Approximationsprozess definiert. Deshalb sehen wir uns auch die Eigenschaften messbarer Funktionen unter Limiten, Suprema, etc an: Gegeben eine Folge von Funktionen $f_n : \Omega \rightarrow \overline{\mathbb{R}}$, $n \in \mathbb{N}$, so setzen wir

$$\begin{aligned} \sup_{n \in \mathbb{N}} f_n : x \mapsto \sup_{n \in \mathbb{N}} f_n(x), & \quad \inf_{n \in \mathbb{N}} f_n : x \mapsto \inf_{n \in \mathbb{N}} f_n(x), \\ \limsup_{n \in \mathbb{N}} f_n : x \mapsto \inf_{n \in \mathbb{N}} \sup_{k \geq n} f_k(x), & \quad \liminf_{n \in \mathbb{N}} f_n : x \mapsto \sup_{n \in \mathbb{N}} \inf_{k \geq n} f_k(x). \end{aligned}$$

Diese Funktionen existieren als Funktionen $\mathbb{R} \rightarrow \overline{\mathbb{R}}$, können also die Werte $\pm\infty$ annehmen, selbst wenn $f_n(x) \in \mathbb{R}$ für alle x, n . Dies zeigt bereits, dass es nützlich ist, mit Funktionen nach $\overline{\mathbb{R}}$ zu arbeiten. Der punktweise Limes $\lim_n f_n : x \mapsto \lim_n f_n(x)$ muss nicht existieren. Wenn er existiert, stimmt er mit $\limsup_n f_n$ und $\liminf_n f_n$ überein.

Lemma 2.4. Sei (Ω, \mathcal{A}) ein Messraum und $f_n : \Omega \rightarrow \overline{\mathbb{R}}$, $n \in \mathbb{N}$, messbar. Dann sind auch die Funktionen

$$\sup_n f_n, \quad \inf_n f_n, \quad \limsup_n f_n, \quad \liminf_n f_n,$$

messbar.

Beweis. Aufgrund der Definition von \liminf und \limsup genügt es, die Aussage für \sup und \inf zu zeigen. Dies folgt direkt aus $\{\sup_n f_n \leq a\} = \bigcap_n \{f_n \leq a\}$ und $\{\inf_n f_n \geq a\} = \bigcap_n \{f_n \geq a\}$. \square

Messbarkeit ist also stabil unter Suprema unter Infima.

Wie bereits am Beginn dieses Kapitels angekündigt, wird das Integral über charakteristische Funktionen und ihre Linearkombinationen aufgebaut werden. Wir definieren diese wie folgt.

Definition 2.5. Sei (Ω, \mathcal{A}) ein Messraum. Eine *einfache Funktion* ist eine messbare numerische Funktion $f : \Omega \rightarrow \overline{\mathbb{R}}$, die nur endlich viele Werte annimmt, d.h. $|f(\Omega)| < \infty$.

Sei $f : \Omega \rightarrow \overline{\mathbb{R}}$ einfach, mit Bild $f(\Omega) = \{y_1, \dots, y_n\}$. Dann sind die Urbilder $A_j := f^{-1}(\{y_j\}) = \{f = y_j\}$ messbare Teilmengen von Ω , und es gilt

$$f = \sum_{j=1}^n y_j \cdot \chi_{A_j}.$$

Diese Darstellung ist allerdings nicht eindeutig, zB $\chi_{[0,1]} = \chi_{[0, \frac{1}{2}]} + \chi_{(\frac{1}{2}, 1]}$.

Es ist klar, dass Linearkombinationen von einfachen Funktionen einfach sind. Da für Produkte von charakteristischen Funktionen

$$\chi_A \cdot \chi_B = \chi_{A \cap B}$$

gilt, sind auch Produkte von einfachen Funktionen einfach. Für $\Omega = \mathbb{R}$ sieht dies alles sehr ähnlich wie beim Riemann-Integral aus, aber wir haben nun sehr viel mehr charakteristische Funktionen als nur solche von Intervallen zur Verfügung. Außerdem können wir sehr viel allgemeinere Definitionsbereiche (zB $\Omega = \mathbb{R}^d$ oder Ω ein Funktionenraum mit einer geeigneten σ -Algebra) betrachten.

Die folgende Approximierbarkeit messbarer Funktionen durch einfache Funktionen wird für uns wesentlich sein.

Satz 2.6. Sei (Ω, \mathcal{A}) ein Messraum und $f : \Omega \rightarrow \overline{\mathbb{R}}_+$ eine messbare nicht negative Funktion. Dann gibt es eine aufsteigende Folge^a nicht negativer einfacher Funktionen $\varphi_n : \Omega \rightarrow \mathbb{R}_+$ mit $\lim_n \varphi_n(x) = \sup_n \varphi_n(x) = f(x)$ für alle $x \in \Omega$.

^ad.h. $0 \leq \varphi_n(x) \leq \varphi_{n+1}(x)$ für alle $x \in \Omega, n \in \mathbb{N}$

Beweis. Wir zerlegen den Bildbereich $[0, \infty]$ von f in die $n2^n + 1$ disjunkten Intervalle

$$\left[0, \frac{1}{2^n}\right), \quad \left[\frac{1}{2^n}, \frac{2}{2^n}\right), \quad \dots \quad \left[\frac{n2^n-1}{2^n}, \frac{n2^n}{2^n}\right), \quad [n, \infty].$$

Durch Urbildbildung erhalten wir eine Zerlegung von Ω in paarweise disjunkte Mengen,

$$\Omega = \left\{0 \leq f < \frac{1}{2^n}\right\} \sqcup \dots \sqcup \left\{\frac{n2^n-1}{2^n} \leq f < \frac{n2^n}{2^n}\right\} \sqcup \{n \leq f\}.$$

Wir definieren nun die einfache Funktion

$$\varphi_n := 0 \cdot \chi_{\{0 \leq f < \frac{1}{2^n}\}} + \frac{1}{2^n} \cdot \chi_{\{\frac{1}{2^n} \leq f < \frac{2}{2^n}\}} + \dots + \frac{n2^n - 1}{2^n} \cdot \chi_{\{\frac{n2^n-1}{2^n} \leq f < \frac{n2^n}{2^n}\}} + n \cdot \chi_{\{n \leq f\}}.$$

Per Konstruktion gilt $0 \leq \varphi_n(x) \leq f(x)$ für alle $x \in \Omega$ und $f(x) < \varphi_n(x) + \frac{1}{2^n}$ für alle $x \in \{f \leq n\}$. Also haben wir den punktweisen Grenzwert $\varphi_n(x) \rightarrow f(x)$ für alle $x \in \Omega$. Durch Vergleich der die charakteristischen Funktionen definierenden Mengen sieht man auch $\varphi_n \leq \varphi_{n+1}$ ein: Wir setzen $A_{n,k} := \{\frac{k}{2^n} \leq f < \frac{k+1}{2^n}\}$. Dann gilt $A_{n,k} = A_{n+1,2k} \sqcup A_{n+1,2k+1}$ und Vergleich der Funktionswerte ($\frac{k}{2^n}$ für φ_n vs. $\frac{2k}{2^{n+1}}$ bzw. $\frac{2k+1}{2^{n+1}}$ für φ_{n+1}) liefert die Behauptung. \square

2.2 Integrale nach einem Maß und integrierbare Funktionen

Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum. Wir beginnen nun die Konstruktion des Integrals $\int_{\Omega} f d\mu$ für Borel-messbare numerische Funktionen $f : \Omega \rightarrow \overline{\mathbb{R}}$. Dabei gehen wir in mehreren Schritten vor.

Für eine nicht-negative einfache Funktion $f : \Omega \rightarrow \overline{\mathbb{R}}_+$ mit Bild $f(\Omega) = \{y_1, \dots, y_n\}$ und Urbildern $A_j := f^{-1}(\{y_j\})$ setzen wir

$$\int_{\Omega} f d\mu := \sum_{j=1}^n y_j \cdot \mu(A_j) \in \overline{\mathbb{R}}_+.$$

Man kann leicht zeigen, dass diese Definition nicht von der Darstellung f abhängt: Falls $f = \sum_{k=1}^m z_k \chi_{B_k}$, so gilt $\sum_{j=1}^n y_j \cdot \mu(A_j) = \sum_{k=1}^m z_k \cdot \mu(B_k)$ (Übung, betrachten Sie die Mengen $B_k \cap A_j$).

Wir können auch nur über eine messbare Teilmenge von Ω integrieren: Für $A \in \mathcal{A}$ setzen wir

$$\int_A f d\mu := \int_{\Omega} f \chi_A d\mu = \sum_{j=1}^n y_j \cdot \mu(A_j \cap A)$$

und nennen dieses Integral das μ -Integral von f über A .

Lemma 2.7. Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und $\varphi : \Omega \rightarrow [0, \infty]$ eine einfache Funktion. Dann ist

$$\nu : \mathcal{A} \rightarrow [0, \infty], \quad \nu(A) := \int_A \varphi d\mu$$

ein Maß.

Beweis. Offenbar gilt $\nu(\emptyset) = 0$. Für die σ -Additivität betrachten wir eine Folge (B_l) paarweise disjunkter Mengen in \mathcal{A} . Mit $B := \bigsqcup_l B_l$ gilt

$$\begin{aligned} \nu(B) &= \int_B \varphi d\mu = \sum_{j=1}^n y_j \mu(A_j \cap B) \\ &= \sum_{j=1}^n y_j \mu \left(A_j \cap \bigsqcup_{l=1}^{\infty} B_l \right) = \sum_{j=1}^n y_j \sum_{l=1}^{\infty} \mu(A_j \cap B_l) \\ &= \sum_{l=1}^{\infty} \sum_{j=1}^n y_j \mu(A_j \cap B_l) = \sum_{l=1}^{\infty} \nu(B_l). \end{aligned}$$

□

Wir notieren einige Eigenschaften dieses Integrals:

Lemma 2.8. Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum. Für nicht-negative einfache Funktionen $\varphi, \psi : \Omega \rightarrow [0, \infty]$, messbare Mengen $A, B \in \mathcal{A}$, und $c \in [0, \infty]$ gilt:

$$\begin{aligned} \int_A c\varphi d\mu &= c \int_A \varphi d\mu, \\ \int_A (\varphi + \psi) d\mu &= \int_A \varphi d\mu + \int_A \psi d\mu, \\ \varphi \leq \psi &\Rightarrow \int_A \varphi d\mu \leq \int_A \psi d\mu, \\ \mu(A) = 0 &\Rightarrow \int_A \varphi d\mu = 0, \\ A \subset B &\Rightarrow \int_A \varphi d\mu \leq \int_B \varphi d\mu. \end{aligned}$$

Beweis. Die ersten beiden und die vierte Eigenschaft sind direkte Konsequenzen der Definition des Integrals. Die dritte Eigenschaft (Monotonie) folgt aus der zweiten (Additivität), weil $\psi - \varphi$ ebenfalls eine nicht-negative einfache Funktion ist. Die letzte Eigenschaft folgt, da $A \subset B$ die Ungleichung $\chi_A \leq \chi_B$ impliziert. \square

Das Integral nach μ über eine nicht-negative messbare Funktion wird nun durch Approximation von unten definiert.

Definition 2.9. Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und $f : \Omega \rightarrow [0, \infty]$ eine nicht negative messbare Funktion. Dann definieren wir das *Integral von f über Ω bzgl. μ* (oder: das μ -Integral von f) als

$$\int_{\Omega} f d\mu := \sup \left\{ \int_{\Omega} \varphi d\mu : 0 \leq \varphi \leq f, \varphi \text{ ist einfach} \right\}.$$

Für einfache Funktionen stimmt diese Definition mit der bereits gegebenen Definition von $\int_{\Omega} f d\mu$ überein: Denn einerseits können wir, falls f einfach ist, unter dem Supremum $\varphi = f$ wählen, was zeigt, dass die neue Definition $\geq \int_{\Omega} f d\mu$ ist. Und andererseits gilt unter dem Supremum aufgrund der Monotonie aus Lemma 2.8 stets $\int_{\Omega} \varphi d\mu \leq \int_{\Omega} f d\mu$, was zeigt, dass die neue Definition $\leq \int_{\Omega} f d\mu$ ist.

Was die Notation des Integrals angeht, so schreiben wir $\int_{\Omega} f d\mu$ oder $\int_{\Omega} f(x) d\mu(x)$. Wenn Ω und/oder μ aus dem Zusammenhang klar sind, schreiben wir auch knapper $\int f d\mu$, $\int_{\Omega} f$ oder $\int f$.

Das Supremum existiert immer in $[0, \infty]$ (könnte also unendlich sein). Wir werden das Supremum gleich noch genauer beschreiben. Dazu benötigen wir die Monotonie des Integrals.

Lemma 2.10. Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und $f, g : \Omega \rightarrow \overline{\mathbb{R}}_+$ messbare nicht-negative Funktionen mit $f \leq g$. Dann gilt $\int_{\Omega} f d\mu \leq \int_{\Omega} g d\mu$.

Beweis. Zur Abkürzung schreiben wir E_h für die Menge aller einfachen Funktionen φ , die $0 \leq \varphi \leq h$ erfüllen. Dann schreibt sich das Integral als $\int_{\Omega} h d\mu = \sup_{\varphi \in E_h} \int_{\Omega} \varphi d\mu$.

Aus $f \leq g$ folgt $E_f \subset E_g$, und damit $\int f = \sup_{\varphi \in E_f} \int \varphi \leq \sup_{\varphi \in E_g} \int \varphi = \int g$. \square

Der folgende Satz von der monotonen Konvergenz ist der zweitwichtigste Konvergenzsatz der Lebesgue'schen Integrationstheorie

Satz 2.11 (Satz von der monotonen Konvergenz; Satz von Beppo Levi). Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und $(f_n)_n$ eine monoton steigende Folge messbarer nicht-negativer numerischer Funktionen $f_n : \Omega \rightarrow [0, \infty]$. Dann ist $f := \lim_n f_n$ messbar (siehe Lemma 2.4), und es gilt

$$\lim_{n \rightarrow \infty} \int_{\Omega} f_n d\mu = \int_{\Omega} f d\mu.$$

Beweis. Da die Funktionenfolge $(f_n)_n$ monoton wächst, wächst nach dem vorausgehenden Lemma auch die Folge der Integrale $(\int_{\Omega} f_n d\mu)_n$. Also existiert $I := \lim_n \int_{\Omega} f_n d\mu \in [0, \infty]$. Aufgrund der Monotonie der Folge gilt $f_n \leq f$ und deshalb $\int_{\Omega} f_n d\mu \leq \int_{\Omega} f d\mu$ für alle $n \in \mathbb{N}$, also $I \leq \int_{\Omega} f d\mu$.

Für die umgekehrte Ungleichung zeigen wir, dass jede einfache Funktion φ mit $0 \leq \varphi \leq f$

$$\int_{\Omega} \varphi d\mu \leq I$$

erfüllt, das ergibt dann die Behauptung. Sei dazu eine solche einfache Funktion φ gegeben. Wir betrachten $0 \leq c < 1$ und betrachten $A_n := \{f_n \geq c\varphi\}$. Es gilt $A_n \subset A_{n+1}$ wegen der Monotonie der Folge $(f_n)_n$, und $\bigcup_{n \in \mathbb{N}} A_n = \Omega$. Zum Beweis der zweiten Aussage sei $x \in \Omega$. Falls $\varphi(x) > 0$, so gilt $c\varphi(x) < \varphi(x) \leq f(x) \leftarrow f_n(x)$, und deshalb gibt es $n \in \mathbb{N}$, so dass $c\varphi(x) \leq f_n(x)$, dh $x \in A_n$. Für $\varphi(x) = 0$ gilt sogar $x \in A_n$ für alle n . Nun erhalten wir

$$I \geq \int_{\Omega} f_n d\mu \geq \int_{A_n} f_n d\mu \geq \int_{A_n} c\varphi d\mu = c \int_{A_n} \varphi d\mu.$$

Da $A \mapsto \int_A \varphi d\mu$ ein Maß ist, können wir Satz 1.25 c) anwenden und sehen, dass

$$\lim_{n \rightarrow \infty} \int_{A_n} \varphi d\mu = \int_{\Omega} \varphi d\mu$$

gilt. Also haben wir $I \geq c \int_{\Omega} \varphi d\mu$, und im Limes $c \nearrow 1$ erhalten wir die gewünschte Ungleichung. \square

Der Satz von der monotonen Konvergenz gilt nicht für das Riemann-Integral: Es gibt monoton wachsende Folgen Riemann-integrierbarer Funktionen, deren Grenzwert nicht Riemann-integrierbar ist.

Beispiel 2.12. Für das Riemann-Integral liefert uns der Hauptsatz der Differential- und Integralrechnung eine effiziente Methode, um Integrale zu berechnen, die uns momentan für das Lebesgue- und allgemeinere μ -Integrale noch fehlt. Wir illustrieren den Satz von Beppo Levi deshalb nur an zwei sehr einfachen Beispielen.

a) Wir betrachten die Funktionen

$$f_n : \mathbb{R} \rightarrow \mathbb{R}, \quad f_n(x) := x^{1/n} \chi_{[0,1]}(x).$$

Die Folge (f_n) ist monoton wachsend ($x^{1/n} = e^{-\frac{1}{n}|\log x|}$ für $0 \leq x \leq 1$), und der punktweise Grenzwert ist

$$f(x) = \lim_{n \rightarrow \infty} f_n(x) = \begin{cases} 1 & 0 < x \leq 1 \\ 0 & \text{sonst} \end{cases},$$

also $f = \chi_{(0,1]}$. Damit folgt für das Lebesgue-Integral über f

$$\lim_n \int_{\mathbb{R}} f_n(x) d\lambda(x) = \int_{\mathbb{R}} \chi_{(0,1]}(x) d\lambda(x) = \lambda((0, 1]) = 1.$$

- b) Ohne die Monotonie-Bedingung gilt der Satz von Beppo Levi nicht: Die Funktionen $f_n : \mathbb{R} \rightarrow \mathbb{R}$, $f_n := \frac{1}{n} \chi_{[0,n]}$ sind messbare einfache nicht-negative Funktionen mit $f_n \rightarrow 0$ für $n \rightarrow \infty$, aber $\int_{\mathbb{R}} f_n d\lambda = 1$ für alle $n \in \mathbb{N}$.

Als Korollar des Satzes 2.11 von Beppo Levi und der Approximierbarkeit messbarer positiver Funktionen (Satz 2.11) erhalten wir eine Charakterisierung des μ -Integrals als Grenzwert.

Korollar 2.13. Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und $f : \Omega \rightarrow [0, \infty]$ eine nicht negative messbare Funktion. Sei weiterhin $(\varphi_n)_n$ eine monoton wachsende Folge nicht negativer einfacher Funktionen mit $\varphi_n \nearrow f$ (die es nach Satz 2.6 gibt). Dann ist das μ -Integral von f

$$\int_{\Omega} f = \lim_{n \rightarrow \infty} \int_{\Omega} \varphi_n = \sup_{n \in \mathbb{N}} \int_{\Omega} \varphi_n.$$

Beweis. Die erste Gleichung folgt direkt aus dem Satz von der monotonen Konvergenz, und die zweite Gleichung aus der Monotonie des Integrals ($n \mapsto \int_{\Omega} \varphi_n$ ist monoton wachsend). \square

Es ist auch möglich, das μ -Integral von f durch $\int_{\Omega} f := \lim_{n \rightarrow \infty} \int_{\Omega} \varphi_n$ mit einer approximierenden Folge (φ_n) von nicht negativen einfachen Funktionen zu definieren. Dann muss man aber zeigen, dass diese Definition nicht von der Wahl der Folge abhängt, was in unserer (äquivalenten) Definition 2.9 offensichtlich ist.

Ein erster Vergleich mit der Definition des Riemann-Integrals (Analysis 1, Kapitel 6) bietet sich hier an: Das Riemann-Integral einer Funktion $f : [a, b] \rightarrow \mathbb{R}$ ist definiert, falls Ober- und Unterintegral übereinstimmen. Das Unterintegral ist durch Approximation von f durch Treppenfunktionen von unten definiert. Dabei sind Treppenfunktionen für Riemannintegrale Linearkombinationen von charakteristischen Funktionen χ_I von Intervallen $I \subset [a, b]$. Da Intervalle messbar sind, sind solche Funktionen insbesondere einfache Funktionen für den Maßraum $([a, b], \mathcal{B}([a, b]), \lambda)$. Bei der Konstruktion des λ -Integrals haben wir also sehr viel mehr einfache Funktionen zur Verfügung als beim Riemann-Integral. Weiterhin basiert das Riemann-Integral auf einer Zerlegung des Definitionsbereichs $[a, b]$, während das Lebesgue-Integral eine Zerlegung des Bildbereichs involviert, vgl. Satz 2.6.

Wir werden die beiden Integralbegriff etwas später noch genauer vergleichen. Zuerst dehnen wir μ -Integrale auf nicht notwendigerweise nicht negative Funktionen aus. Dazu notieren wir folgendes elementare Lemma.

Lemma 2.14. Für messbare Funktionen $f, g : \Omega \rightarrow \overline{\mathbb{R}}_+$ und $c \in \mathbb{R}_+$ gilt

$$\int_{\Omega} (cf) d\mu = c \int_{\Omega} f d\mu,$$

$$\int_{\Omega} (f + g) d\mu = \int_{\Omega} f d\mu + \int_{\Omega} g d\mu.$$

Beweis. Wir wählen Folgen (φ_n) und (ψ_n) monoton wachsender nicht negativer einfacher Funktionen mit $\varphi_n \nearrow f$ und $\psi_n \nearrow g$. Dann gilt $c\varphi_n \nearrow cf$ und $\varphi_n + \psi_n \nearrow f + g$, und deshalb

$$\int_{\Omega} cf d\mu = \lim_{n \rightarrow \infty} \int_{\Omega} c\varphi_n d\mu = c \lim_{n \rightarrow \infty} \int_{\Omega} \varphi_n d\mu = c \int_{\Omega} f d\mu,$$

$$\int_{\Omega} (f + g) d\mu = \lim_{n \rightarrow \infty} \int_{\Omega} (\varphi_n + \psi_n) d\mu = \lim_{n \rightarrow \infty} \left(\int_{\Omega} \varphi_n d\mu + \int_{\Omega} \psi_n d\mu \right)$$

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

□

Definition 2.15. Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum. Eine Funktion $f : \Omega \rightarrow \overline{\mathbb{R}}$ heißt μ -integrierbar, wenn sie messbar ist und für die beiden (messbaren) Funktionen $f_{\pm} : \Omega \rightarrow \overline{\mathbb{R}}_+$

$$\int_{\Omega} f_{\pm} d\mu < \infty$$

gilt. In diesem Fall definieren wir das μ -Integral von f als

$$\int_{\Omega} f d\mu := \int_{\Omega} f_+ d\mu - \int_{\Omega} f_- d\mu.$$

Die Menge aller μ -integrierbaren Funktionen $f : \Omega \rightarrow \mathbb{R}$ wird mit $\mathcal{L}^1(\Omega, \mathcal{A}, \mu)$ (oder kürzer $\mathcal{L}^1(\Omega, \mu)$ bzw $\mathcal{L}^1(\Omega)$) bezeichnet.

Für den Fall $\Omega = \mathbb{R}^d$, $\mathcal{A} = \mathcal{B}(\mathbb{R}^d)$, $\mu = \lambda^d$ (Lebesgue-Maß) sprechen wir von *Lebesgue-Integrierbarkeit* und dem *Lebesgue-Integral*.

Im Fall $f \geq 0$ gilt $f = f_+$ und $f_- = 0$, so dass die obige Definition des μ -Integrals mit der bisherigen Definition für nicht negative Funktionen übereinstimmt. Eine nicht negative messbare Funktion heißt aber nur dann integrierbar, wenn ihr μ -Integral endlich ist.

Bemerken Sie, dass eine messbare Funktion $f : \Omega \rightarrow \overline{\mathbb{R}}$ genau dann μ -integrierbar ist, wenn

$$\int_{\Omega} |f| d\mu < \infty$$

gilt. Denn $|f| = f_+ + f_-$, also mit Lemma 2.14 $\int_{\Omega} |f| d\mu = \int_{\Omega} f_+ d\mu + \int_{\Omega} f_- d\mu$.

Wenn das Maß μ aus dem Zusammenhang klar ist, sprechen wir auch kürzer von integrierbaren (statt: μ -integrierbaren) Funktionen. Wir sammeln einige Eigenschaften integrierbarer Funktionen:

Satz 2.16. Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und $f, g : \Omega \rightarrow \overline{\mathbb{R}}$ integrierbar.

a) Für jedes $c \in \mathbb{R}$ ist cf integrierbar, und

$$\int_{\Omega} cf \, d\mu = c \int_{\Omega} f \, d\mu.$$

b) Ist $f + g$ auf ganz Ω definiert, so ist auch $f + g$ integrierbar, und es gilt

$$\int_{\Omega} (f + g) \, d\mu = \int_{\Omega} f \, d\mu + \int_{\Omega} g \, d\mu.$$

c) Falls $f \leq g$, so gilt

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

d) Die Dreiecksungleichung für Integrale gilt:

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

Insbesondere ist $\mathcal{L}^1(\Omega, \mathcal{A}, \mu)$ ein Vektorraum, und $f \mapsto \int_{\Omega} f \, d\mu$ ein lineares Funktional.

Beweis. a) Für $c \geq 0$ gilt $(cf)_{\pm} = cf_{\pm}$ und deshalb $\int (cf)_{\pm} = \int cf_{\pm} = c \int f_{\pm} < \infty$, und $\int (cf) = \int cf_+ - \int (cf_-) = c \int f_+ - c \int f_- = c \int f$. Für $c = -1$ gilt $(-f)_{\pm} = f_{\mp}$, also $\int (-f) = \int f_- - \int f_+ = -\int f$.

b) Es gilt

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

Wegen $(f + g)_{\pm} \leq f_{\pm} + g_{\pm}$ sind alle Integrale endlich. Das impliziert die Behauptung.

c) $f \leq g$ impliziert $f_+ \leq g_+$ und $f_- \geq g_-$, also

$$\int f = \int f_+ - \int f_- \leq \int g_+ - \int g_- = \int g.$$

d) folgt aus $-|f| \leq f \leq |f|$. □

Damit ist das Integral nach einem Maß vollständig definiert. Wir bemerken noch, dass wie für beim Integral über einfache Funktionen für messbare Mengen $A \in \mathcal{A}$

$$\int_A f \, d\mu := \int_{\Omega} \chi_A f \, d\mu$$

gesetzt wird. Dies erklärt das Integral über (messbare) Teilmengen des Definitionsbereichs (siehe Übung).

Analog zu Lemma 2.7 sieht man, dass für eine nicht negative messbare Funktion $f : \Omega \rightarrow \overline{\mathbb{R}}$

$$\nu : \mathcal{A} \rightarrow \overline{\mathbb{R}}_+, \quad \nu(A) := \int_A f d\mu$$

ein Maß definiert (Übung Blatt 5). Dieses Maß heißt das *Maß mit Dichte f bzgl. μ* und wird auch kurz als $f \cdot \mu$ bezeichnet.

Es ist nicht schwierig zu zeigen, dass $\mu(A) = 0 \Rightarrow (f \cdot \mu)(A) = 0$ gilt, dh Nullmengen für μ sind auch Nullmengen für $\nu := f \cdot \mu$ (die Umkehrung gilt im Allgemeinen nicht). In dieser Situation sagt man, dass ν absolut stetig bzgl. μ ist, und schreibt $\nu \ll \mu$.

Der *Satz von Radon-Nikodým* besagt, dass unter gewissen Bedingungen auch die Umkehrung gilt: Falls μ, ν zwei σ -endliche Maße auf einem Messraum (Ω, \mathcal{A}) sind, so dass $\nu \ll \mu$, dann besitzt ν eine Dichte bzgl μ : Es gibt eine messbare nicht negative Funktion $f : \Omega \rightarrow \mathbb{R}$, so dass $\nu = f \cdot \mu$. Die Funktion f heißt dann Radon-Nikodým-Ableitung von ν bzgl. μ und als $f = \frac{d\nu}{d\mu}$ geschrieben.

Dieser Satz hat viele Anwendungen, insbesondere in der Wahrscheinlichkeitstheorie. Nicht kommutative Varianten spielen auch in der Theorie der von Neumann Algebren eine wichtige Rolle, führen aber über den Rahmen dieser Vorlesung hinaus.

2.3 Nullmengen und Vergleich Lebesgue-/Riemannintegral

In der Konstruktion von Maßen aus Prämaßen haben wir den Begriff einer Nullmenge kennengelernt (Def. 1.18). Weiterhin haben wir gesehen, dass für das vollständige Lebesguemaß λ alle Nullmengen messbar sind, d.h. genau die Mengen $A \subset \mathbb{R}^d$ mit $\lambda(A) = 0$ sind.

Auch bei Integralen treten Nullmengen auf sehr natürliche Art und Weise auf.

Satz 2.17. Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum.

a) Für eine messbare numerische Funktion $f : \Omega \rightarrow \overline{\mathbb{R}}$ gilt

$$\int_{\Omega} |f| d\mu = 0 \iff \{x \in \Omega : f(x) \neq 0\} \text{ ist eine Nullmenge.}$$

b) Ist $f : \Omega \rightarrow \overline{\mathbb{R}}$ integrierbar, so ist $\{x \in \Omega : f(x) = \pm\infty\}$ eine Nullmenge.

c) Seien $f, g : \Omega \rightarrow \overline{\mathbb{R}}$ messbar, so dass $\{x \in \Omega : f(x) \neq g(x)\}$ eine Nullmenge ist. Dann ist f genau dann integrierbar, wenn g integrierbar ist, und in diesem Fall gilt

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

Beweis. a) Die Menge $S := \{x \in \Omega : f(x) \neq 0\} = f^{-1}(\overline{\mathbb{R}} \setminus \{0\})$ ist messbar, da f messbar ist. Sei $\int_{\Omega} |f| d\mu = 0$, und betrachte $\varphi_n := \inf\{n|f|, \chi_S\}$, $n \in \mathbb{N}$. Dann sind die φ_n messbar und bilden eine monoton wachsende Folge mit $\varphi_n \nearrow \chi_S$. Weiterhin gilt $\int \varphi_n \leq$

$\int n|f| = n \int |f| = 0$, also $\int \varphi_n = 0$. Mit dem Satz über monotone Konvergenz folgt also $\mu(S) = \int \chi_S = \lim_n \int \varphi_n = 0$, d.h. S ist eine Nullmenge.

Für die andere Beweisrichtung nehmen wir an, dass S eine Nullmenge ist, und definieren $\psi_n := \inf\{n\chi_S, |f|\}$. Dann sind die ψ_n eine monoton wachsende Folge mit $\psi_n \nearrow |f|$. Da $0 \leq \int \psi_n \leq n \int \chi_S = n\mu(S) = 0$, gilt $\int \psi_n = 0$ für alle n , und damit $\int |f| = \lim_n \int \psi_n = 0$ (nach Satz über monotone Konvergenz).

b) Sei $M := \{x \in \Omega : f(x) = \pm\infty\} = \{x \in \Omega : |f(x)| = \infty\}$; dies ist eine messbare Menge. Dann gilt $n\chi_M \leq |f|$ für alle $n \in \mathbb{N}$, und deshalb

$$n\mu(M) = \int_{\Omega} n\chi_M d\mu \leq \int_{\Omega} |f| d\mu < \infty,$$

da f integrierbar ist. Das impliziert $\mu(M) = 0$.

c) Die Menge $M := \{f \neq g\} = \{x \in \Omega : f(x) \neq g(x)\}$ ist messbar, da f und g messbar sind. Die Funktionen $f \cdot \chi_{M^c} = g \cdot \chi_{M^c}$ und $f \cdot \chi_M$ sind als Produkte messbarer Funktionen messbar. Da M nach Voraussetzung eine Nullmenge ist, gilt nach a) $\int |f\chi_M| = 0$. Da $f = f \cdot \chi_{M^c} + f \cdot \chi_M$, haben wir

$$\int |f| = \int |f\chi_M + f\chi_{M^c}| \leq \int |f\chi_{M^c}| \leq \int |f|.$$

Also ist f genau dann integrierbar, wenn $f\chi_{M^c}$ integrierbar ist, und $\int |f| = \int |f\chi_{M^c}|$. Genauso können wir für g argumentieren, und wegen $f\chi_{M^c} = g\chi_{M^c}$ folgt die Behauptung. \square

Dieser Satz lehrt uns, dass es meist irrelevant ist, ob wir Funktionen $f : \Omega \rightarrow \mathbb{R}$ oder $f : \Omega \rightarrow \overline{\mathbb{R}}$ betrachten: Falls f integrierbar ist, ist $\{|f| = \infty\}$ eine Nullmenge, und wenn wir f auf dieser Nullmenge ändern (z.B. $= 0$ setzen), hat das keinen Einfluss auf das Integral.

Als knappe Sprechweise vereinbaren wir das Folgende: Ist $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und P eine Aussage über Punkte $x \in \Omega$, so sagen wir P **gilt μ -fast überall**, falls die Menge der Punkte $x \in \Omega$, für die $P(x)$ falsch ist, eine μ -Nullmenge ist. Im Falle des Lebesgue-Maßes sagen wir auch **Lebesgue-fast überall**.

Beispiel 2.18. Die Aussagen von Satz 2.17 können wir mit dieser Sprechweise auch so formulieren:

- a) $\int_{\Omega} |f| d\mu = 0 \iff f = 0$ μ -fast überall.
- b) f integrierbar $\Rightarrow |f| < \infty$ μ -fast überall.
- c) $f = g$ μ -fast überall $\Rightarrow \int f d\mu = \int g d\mu$.

Weiteres Beispiel: Eine Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$, die abzählbare viele Nullstellen hat, erfüllt $f \neq 0$ Lebesgue-fast überall.

Wir vergleichen nun das Riemann- und das Lebesgueintegral. Dazu betrachten wir ein kompaktes Intervall $[a, b] \subset \mathbb{R}$ und beschränkte Funktionen $f : [a, b] \rightarrow \mathbb{R}$. Für diese Funktionen haben wir den Begriff der Riemann-Integrierbarkeit. Erinnerung dazu: f ist Riemann-integrierbar, falls es für jedes $\varepsilon > 0$ Treppenfunktionen $\varphi, \psi : [a, b] \rightarrow \mathbb{R}$ mit $\varphi \leq f \leq \psi$

und $\int_a^b \psi(x)dx - \int_a^b \varphi(x)dx < \varepsilon$ gibt – siehe Analysis 1, Lemma 6.7. Hierbei ist eine Treppenfunktion eine Linearkombination von charakteristischen Funktionen von Intervallen, insbesondere also eine Lebesgue-einfache Funktion. Das Integral von Treppenfunktionen φ war so definiert, dass es mit $\int_{[a,b]} \varphi(x)d\lambda(x)$ übereinstimmt.

Angenommen, $f : [a, b] \rightarrow \mathbb{R}$ ist Riemann-integrierbar. Dann können wir zu jedem $n \in \mathbb{N}$ Treppenfunktionen $\varphi_n, \psi_n : [a, b] \rightarrow \mathbb{R}$ wählen, die $\varphi_n \leq f \leq \psi_n$ und

$$\int_a^b \psi_n(x)dx - \int_a^b \varphi_n(x)dx < \frac{1}{n}$$

erfüllen. Wir dürfen annehmen, dass (φ_n) monoton wachsend und (ψ_n) monoton fallend ist, denn ansonsten ersetzen wir φ_n durch $\tilde{\varphi}_n := \sup\{\varphi_1, \dots, \varphi_n\}$ und ψ_n durch $\tilde{\psi}_n := \inf\{\psi_1, \dots, \psi_n\}$.

Nach Definition des Riemann-Integrals gilt

$$\int_a^b f(x)dx = \lim_{n \rightarrow \infty} \int_a^b \psi_n(x)dx = \lim_{n \rightarrow \infty} \int_a^b \varphi_n(x)dx.$$

Wir betrachten $\hat{\varphi}_n := \varphi_n - \varphi_1$ und $\hat{\psi}_n := \psi_1 - \psi_n$. Dann gilt $\hat{\varphi}_n, \hat{\psi}_n \geq 0$ und sowohl $(\hat{\varphi}_n)$ als auch $(\hat{\psi}_n)$ sind monoton wachsend. Die Funktionen $\varphi := \varphi_1 + \lim_n \hat{\varphi}_n = \lim_n \varphi_n$ und $\psi := \psi_1 - \lim_n \hat{\psi}_n = \lim_n \psi_n$ sind messbar, und nach dem Satz von Beppo Levi gilt

$$\begin{aligned} \lim_{n \rightarrow \infty} \int_{[a,b]} \varphi_n(x)d\lambda(x) &= \lim_{n \rightarrow \infty} \int_{[a,b]} \hat{\varphi}_n(x)d\lambda(x) + \int_{[a,b]} \varphi_1(x)d\lambda(x) = \int_{[a,b]} \varphi(x)d\lambda(x), \\ \lim_{n \rightarrow \infty} \int_{[a,b]} \psi_n(x)d\lambda(x) &= \int_{[a,b]} \psi_1(x)d\lambda(x) - \lim_{n \rightarrow \infty} \int_{[a,b]} \hat{\psi}_n(x)d\lambda(x) = \int_{[a,b]} \psi(x)d\lambda(x). \end{aligned}$$

Da beide Limiten gleich dem Riemann-Integral von f sind, haben wir

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

gezeigt. Weiterhin gilt $\varphi \leq f \leq \psi$, also $\psi - \varphi \geq 0$ und $\int_{[a,b]} |\psi - \varphi|d\lambda = 0$. Satz 2.17 a) impliziert nun, dass $\psi = \varphi$ λ -fast überall gilt, also auch $f = \varphi$ λ -fast überall. Ersetzen wir also f durch φ (was f nur auf einer Nullmenge ändert), so folgt $\int_a^b f(x)dx = \int_{[a,b]} f d\lambda$. Zusammengefasst:

Satz 2.19. Sei $[a, b] \subset \mathbb{R}$ ein kompaktes Intervall und $f : [a, b] \rightarrow \mathbb{R}$ eine Riemann-integrierbare Funktion. Dann ist f , evtl. nach einer Abänderung auf einer Lebesgue-Nullmenge, Borel-messbar und das Riemannsche und Lebesguesche Integral stimmen überein,

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

Bemerkungen:

- Es gibt beschränkte Funktionen auf kompakten Intervallen, die Lebesgue-integrierbar, aber nicht Riemann-integrierbar sind. Ein typisches Beispiel ist $f = \chi_{\mathbb{Q} \cap [0,1]}$ – siehe Übungen.

- Aufgrund des obigen Satzes schreibt man für Lebesgue-Integrale auch oft $\int_a^b f(x)dx$ oder $\int_{[a,b]} f(x)dx$ anstelle von $\int_{[a,b]} f(x)d\lambda(x)$ oder $\int_{[a,b]} f d\lambda$. Diese Notation ist aber nur für das Lebesgueintegral reserviert und wird nicht für Integrale bzgl anderen Maßen verwendet.
- Als noch genaueren Vergleich von Riemann- und Lebesgueintegral lässt sich zeigen: Eine beschränkte Funktion $f : [a, b] \rightarrow \mathbb{R}$ ist genau dann Riemann-integrierbar, wenn die Menge der Unstetigkeitsstellen von f eine Lebesgue-Nullmenge ist. Für den Beweis verweisen wir auf das Buch von Elstrodt, Satz 6.1.

In Analysis 2, Kapitel 7, haben wir auch *uneigentliche* Riemann-Integrale kennengelernt. Erinnerung: Für ein Intervall $I \subset \mathbb{R}$ (offen, halboffen, mit endlichen oder unendlichen Grenzen, also zB (a, b) , (a, ∞) , $(-\infty, a]$) heißt $f : I \rightarrow \mathbb{R}$ (uneigentlich) Riemann-integrierbar, wenn für jedes kompakte Intervall $K \subset I$ die Einschränkung $f|_K$ Riemann-integrierbar ist, und $\lim_{K \nearrow I} \int_K f =: \int_I f$ existiert.

Satz 2.20. Sei $I \subset \mathbb{R}$ ein Intervall und $f : I \rightarrow \mathbb{R}$ eine Funktion, so dass $f|_K$ für jedes kompakte Intervall $K \subset I$ Riemann-integrierbar ist. Dann ist f (evtl nach Abänderung auf einer Nullmenge) genau dann Lebesgue-integrierbar über I , wenn $|f|$ uneigentlich über I Riemann-integrierbar ist, und das uneigentliche Riemann-Integral $\int_I f$ stimmt mit dem Lebesgue-Integral $\int_I f d\lambda$ überein.

Beweis. Wir nehmen an, dass $I = (a, b)$ offen ist (mit $-\infty \leq a < b \leq +\infty$), im Falle eines halboffenen Intervalls ist der Beweis ähnlich.

Wir wählen monotone Folgen $a_n \searrow a$ und $b_n \nearrow b$ mit $a < a_n < b_n < b$. Nach Satz 2.19 ist $f \cdot \chi_{[a_n, b_n]}$ Lebesgue-integrierbar (evtl nach Abänderung auf einer Nullmenge), und $f = \lim_n f \chi_{[a_n, b_n]}$ als Grenzwert einer Folge messbarer Funktionen messbar. Mit Hilfe des Satzes über monotone Konvergenz sehen wir, dass das uneigentliche Riemann-Integral über $|f|$

$$\int_a^b |f(x)|dx = \lim_{n \rightarrow \infty} \int_{a_n}^{b_n} |f(x)|dx = \lim_{n \rightarrow \infty} \int_I |f| \chi_{[a_n, b_n]} d\lambda = \int_I |f| d\lambda$$

ist. Das zeigt, dass f genau dann Lebesgue-integrierbar über I ist, wenn $|f|$ uneigentlich Riemann-integrierbar über I ist.

Um die Gleichheit des uneigentlichen Riemann-Integrals von f (nicht $|f|$) mit dem Lebesgueintegral $\int_I f d\lambda$ zu zeigen, würden wir gerne so argumentieren:

$$\int_a^b f(x)dx = \lim_{n \rightarrow \infty} \int_{a_n}^{b_n} f(x) dx = \lim_{n \rightarrow \infty} \int_I f \chi_{[a_n, b_n]} d\lambda = \int_I f d\lambda.$$

Alle Schritte bis auf den letzten (Vertauschung Limes und Integral) sind gerechtfertigt. Im letzten Schritt können wir aber nicht den Satz von Beppo Levi verwenden, da die Konvergenz im Allgemeinen nicht monoton ist. Wir brauchen also einen anderen Approximationssatz für Lebesgueintegrale, der uns diesen Schritt erlaubt. Wir werden diesen Satz im nächsten Abschnitt kennenlernen (Satz 2.23), der dann auch diese Lücke im Beweis schließt. \square

Zum Abschluss unseres Vergleichs von Riemann- und Lebesgueintegralen eine Warnung: Es gibt uneigentliche Riemann-Integrale, die *nicht* als Lebesgueintegrale existieren. Der wesentliche Punkt ist, dass Lebesgue-Integrierbarkeit äquivalent zu $\int |f| d\lambda < \infty$ ist. Uneigentliche

Riemann-Integrale können aber existieren, indem sich positive und negative Beiträge zum Integral geeignet wegheben. Das folgende Beispiel illustriert das.

Beispiel 2.21. Wir betrachten das Fresnel'sche Integral $\int_0^\infty \sin(x^2) dx$ (siehe Beispiel 7.11 in Analysis 2). Dieses Integral existiert als ein uneigentliches Riemann-Integral, aber $\int_0^\infty |\sin(x^2)| dx$ existiert *nicht*. (Man zeigt wie in Beispiel 7.11 in Analysis 2, dass dieses Integral mit $\frac{1}{2} \int_0^\infty \frac{|\sin t|}{\sqrt{t}} dt$ übereinstimmt, aber letzteres Integral existiert nicht, da $t \mapsto t^{-1/2}$ zu langsam abfällt.)

Aber die Existenz von $\int_0^\infty |\sin(x^2)| dx$ als uneigentliches Riemann-Integral ist nach Satz 2.20 äquivalent zur Existenz (Endlichkeit) des Lebesgueintegrals $\int_0^\infty \sin(x^2) d\lambda(x)$, also existiert dieses Integral nicht.

2.4 Majorisierte Konvergenz und Anwendungen

Wir kommen jetzt zu dem wichtigsten Approximationssatz der Lebesgueschen Integrations-theorie, dem Satz über majorisierte Konvergenz (auf Englisch “dominated convergence”) von Lebesgue.

Zuerst betrachten wir das Lemma von Fatou.

Satz 2.22 (Lemma von Fatou). Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und $f_n : \Omega \rightarrow \overline{\mathbb{R}}_+$ eine Folge nicht negativer messbarer Funktionen. Dann gilt

$$\int_{\Omega} \liminf_{n \rightarrow \infty} f_n d\mu \leq \liminf_{n \rightarrow \infty} \int_{\Omega} f_n d\mu.$$

Einige Bemerkungen zum Lemma von Fatou:

- Dieser Satz hat nur sehr wenige Voraussetzungen: Es wird einzig und allein gefordert, dass die f_n nicht negativ und messbar sind. Die Existenz von $\liminf_n f_n$ ist automatisch (es ist ein Grenzwert einer monoton wachsenden Folge). Allerdings ist auch die Aussage des Lemmas schwächer als zB beim Satz von Beppo Levi, da hier nur eine Ungleichung vorliegt.
- Der Beweis des Lemmas und Beispiele, für die $\int_{\Omega} \liminf_{n \rightarrow \infty} f_n d\mu < \liminf_{n \rightarrow \infty} \int_{\Omega} f_n d\mu$ gilt, werden in Hausaufgabe H5.2 besprochen.
- Um sich die Richtung der Ungleichung zu merken, sollten Sie sich klarmachen, dass in $\liminf_n f_n$ der Limes Inferior für jeden Punkt x einzeln gebildet wird, während in $\liminf_n \int f_n$ nur der Limes Inferior über die zahlenwertige Folge der Integrale gebildet wird – es ist also zu erwarten, dass die linke Seite kleiner ausfallen kann als die rechte.

Satz 2.23 (Satz über majorisierte Konvergenz). Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und $f_n : \Omega \rightarrow \mathbb{R}$, $n \in \mathbb{N}$, eine Folge messbarer Funktionen, die μ -fast überall punktweise gegen eine Funktion $f : \Omega \rightarrow \overline{\mathbb{R}}$ konvergiert. Falls es eine integrierbare Majorante gibt, also eine integrierbare Funktion $M : \Omega \rightarrow \mathbb{R}_+$ mit $|f_n| \leq M$ μ -fast überall für alle $n \in \mathbb{N}$, so

folgt: f_n und f sind integrierbar, und

$$\lim_{n \rightarrow \infty} \int_{\Omega} f_n d\mu = \int_{\Omega} f d\mu, \quad \lim_{n \rightarrow \infty} \int_{\Omega} |f_n - f| d\mu = 0.$$

Beweis. Alle behaupteten Aussagen über f_n und f ändern sich nicht, wenn diese Funktionen auf Nullmengen abgeändert werden. Da es eine Nullmenge $N \in \mathcal{A}$ gibt, so dass $f_n(x) \rightarrow f(x)$ für $x \in N^c$, können wir nach den Ersetzungen $f_n \mapsto \chi_{N^c} f_n$ und $f \mapsto f \chi_{N^c}$ annehmen, dass $f_n(x) \rightarrow f(x)$ für alle $x \in \Omega$ gilt. Analog dürfen wir auch $|f_n(x)| \leq M(x)$ für alle $x \in \Omega$ annehmen.

Die Annahme über die Majorante impliziert sofort, dass die f_n integrierbar sind: $\int |f_n| \leq \int M < \infty$. Da $M \geq |f_n| \rightarrow |f|$, folgt auch die Integrierbarkeit von f .

Wir haben die Ungleichung

$$0 \leq |f_n - f| \leq |f_n| + |f| \leq 2M$$

und wenden das Lemma von Fatou auf die Folge $g_n := 2M - |f_n - f| \geq 0$ an: Wegen $\liminf_n g_n = \lim_n g_n = 2M$ gilt

$$\begin{aligned} \int_{\Omega} 2M d\mu &= \int_{\Omega} \liminf_{n \rightarrow \infty} (2M - |f_n - f|) d\mu \\ &\leq \liminf_{n \rightarrow \infty} \int_{\Omega} (2M - |f_n - f|) d\mu \\ &= \int_{\Omega} 2M d\mu - \limsup_{n \rightarrow \infty} \int_{\Omega} |f_n - f| d\mu, \end{aligned}$$

also $\limsup_n \int_{\Omega} |f_n - f| d\mu = 0$. Der größte Häufungspunkt der nicht negativen Folge $I_n := \int_{\Omega} |f_n - f| d\mu$ ist also Null. Das impliziert $\lim_n \int_{\Omega} |f_n - f| d\mu = 0$, wie in der zweiten behaupteten Gleichung angegeben.

Um auch die erste behauptete Gleichung einzusehen, genügt nun die Dreiecksungleichung:

$$\left| \int_{\Omega} f_n d\mu - \int_{\Omega} f d\mu \right| = \left| \int_{\Omega} (f_n - f) d\mu \right| \leq \int_{\Omega} |f_n - f| d\mu \rightarrow 0.$$

□

Beispiel 2.24.

a) Wir möchten den Grenzwert

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}} \frac{n \sin \frac{x}{n}}{x(1+x^2)} dx$$

bestimmen. Bei $x = 0$ kann der Integrand $f_n : \mathbb{R} \rightarrow \mathbb{R}$, $f_n(x) = \frac{n \sin \frac{x}{n}}{x(1+x^2)}$, $n \in \mathbb{N}$, durch $f_n(0) = 1$ stetig fortgesetzt werden (Erinnerung: $\frac{\sin \varepsilon}{\varepsilon} \rightarrow 1$ für $\varepsilon \rightarrow 0$). Die f_n sind also alle stetig (sogar glatt), also auf jeden Fall messbar.

Es gilt $f_n(x) \rightarrow (1+x^2)^{-1} =: f(x)$ für alle $x \in \mathbb{R}$ für $n \rightarrow \infty$, die Grenzfunktion f ist auch messbar und sogar integrierbar (siehe Kapitel 7 in Analysis 2

– beachten Sie, dass $f(x)$ schneller als $\frac{1}{x^{1+\varepsilon}}$, $\varepsilon > 0$, abfällt).

Weiterhin gilt $|\sin t| \leq |t|$ für alle $t \in \mathbb{R}$. Also ist $M(x) := \frac{1}{1+x^2} = f(x)$ eine integrierbare Majorante, und es folgt

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}} \frac{n \sin \frac{x}{n}}{x(1+x^2)} dx = \int_{\mathbb{R}} \frac{1}{1+x^2} dx = [\arctan x]_{-\infty}^{\infty} = \pi.$$

- b) Das nächste Beispiel zeigt, dass die Annahme über die integrierbare Majorante nicht fallengelassen werden kann: Sei $f_n := \chi_{[n, n+1]}$. Dann gilt $\int f_n(x) dx = 1$ für alle $n \in \mathbb{N}$, und $\lim_{n \rightarrow \infty} f_n(x) = 0$ für alle $x \in \mathbb{R}$, also $\int_{\mathbb{R}} \lim_n f_n(x) dx = 0$.

In diesem Fall gibt es keine integrierbare Majorante. Denn falls $M : \mathbb{R} \rightarrow \overline{\mathbb{R}}_+$ die Ungleichung $M(x) \geq |f_n(x)|$ für alle n und x erfüllt, so gilt $\int_{\mathbb{R}} M(x) dx \geq \int_1^{\infty} M(x) dx \geq \int_1^{\infty} 1 dx = \infty$. Also ist M nicht integrierbar.

Es gibt extrem viele Anwendungen des Satzes über majorisierte Konvergenz. Wir illustrieren drei typische Situationen.

Lemma 2.25. Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und (A_n) eine Ausschöpfung von Ω (eine Folge messbarer Mengen mit $A_n \subset A_{n+1}$ für alle n , und $\Omega = \bigcup_{n \in \mathbb{N}} A_n$). Falls eine messbare Funktion $f : \Omega \rightarrow \mathbb{R}$ über jedes A_n integrierbar ist (d.h. $\int_{A_n} |f| d\mu < \infty$), und

$$\lim_{n \rightarrow \infty} \int_{A_n} |f| d\mu < \infty,$$

so ist f über ganz Ω integrierbar, mit

$$\int_{\Omega} f d\mu = \lim_{n \rightarrow \infty} \int_{A_n} f d\mu.$$

Beweis. Die Folge $f_n := f \cdot \chi_{A_n}$ konvergiert gegen f , und die Folge $|f_n| \chi_{A_n}$ wächst monoton gegen $M := |f|$. Nach dem Satz über monotone Konvergenz ist f integrierbar. Da $M \geq |f_n|$, haben wir eine integrable Majorante. Die Behauptung folgt nun aus dem Satz über majorisierte Konvergenz. \square

Dieser Satz schließt die Lücke im Beweis von Satz 2.20.

Wir geben ein Beispiel mit Lebesgue-Maß in Dimension $d = 3$.

Beispiel 2.26. Die Funktion $f : \mathbb{R}^3 \rightarrow \mathbb{R}$, $f(x) := e^{-\|x\|_2}$, ist Lebesgue-integrierbar.

Beweis: Als stetige Funktion ist f messbar. Wir haben derzeit noch keine Methode, das mehrdimensionale Integral $\int_{\mathbb{R}^3} e^{-\|x\|_2} d\lambda(x)$ zu berechnen, können aber immerhin argumentieren, dass es endlich ist.

Dazu betrachten wir die 3-dimensionalen Kugeln $B_n := \{x \in \mathbb{R}^3 : \|x\|_2 \leq n\}$ mit Radius $n \in \mathbb{N}$, und die "Kugelschalen" $B_{n+1} \setminus B_n$. Da $\mathbb{R}_+ \ni r \mapsto e^{-r}$ monoton fallend

ist, gilt

$$e^{-\|x\|_2} \leq e^{-n}, \quad x \in B_{n+1} \setminus B_n.$$

Somit erhalten wir

$$\int_{B_{n+1} \setminus B_n} e^{-\|x\|_2} dx \leq e^{-n} \int_{B_{n+1} \setminus B_n} dx = e^{-n} \lambda(B_{n+1} \setminus B_n) = e^{-n} (\lambda(B_{n+1}) - \lambda(B_n)).$$

Wie aus der Schule bekannt, gilt $\lambda(B_n) = \frac{4}{3}\pi n^3$. Wir können das aber aktuell noch nicht beweisen, haben aber wenigstens Beispiel 1.49 a) zur Verfügung:

$$\lambda(B_n) = \lambda(n \cdot B_1) = n^3 \cdot \lambda(B_1).$$

Da B_1 kompakt ist, gilt $\lambda(B_1) < \infty$. Somit erhalten wir

$$\int_{B_{n+1} \setminus B_n} e^{-\|x\|_2} dx \leq \lambda(B_1) \cdot e^{-n} ((n+1)^3 - n^3) = \lambda(B_1) \cdot e^{-n} (3n^2 + 3n + 1) < \infty.$$

Wir betrachten nun $B_{n+1} = B_1 \sqcup (B_2 \setminus B_1) \sqcup \dots \sqcup (B_{n+1} \setminus B_n)$ erhalten per Additivität

$$\begin{aligned} \int_{B_{n+1}} e^{-\|x\|_2} dx &= \int_{B_1} e^{-\|x\|_2} dx + \sum_{k=1}^n \int_{B_{k+1} \setminus B_k} e^{-\|x\|_2} dx \\ &\leq \lambda(B_1) + \sum_{k=1}^n \lambda(B_1) \cdot e^{-k} (3k^2 + 3k + 1). \end{aligned}$$

Da die Reihe $\sum_{k=1}^{\infty} e^{-k} (3k^2 + 3k + 1)$ konvergiert (siehe Analysis 1), können wir Lemma 2.25 anwenden und erhalten $\int_{\mathbb{R}^3} e^{-\|x\|_2} dx < \infty$.

Als zwei weitere Anwendungen betrachten wir parameterabhängige Integrale. Darunter verstehen wir Integrale der Form $\int f(x, y) dx$, dh der Integrand hängt von zwei Variablen ab, von denen nur eine integriert wird (x), und wir interessieren uns für die Abhängigkeit des Integrals von der anderen Variable (y).

Satz 2.27 (Parameterabhängige Integrale – Stetigkeit). Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum, $U \subset \mathbb{R}^d$ eine offene Teilmenge, und $y_0 \in U$. Weiter sei

$$f : \Omega \times U \rightarrow \mathbb{R}, \quad (x, y) \mapsto f(x, y)$$

eine Funktion mit den folgenden drei Eigenschaften:

- Für jedes $y \in U$ ist $x \mapsto f(x, y)$ μ -integrierbar auf Ω .
- Für jedes $x \in \Omega$ ist die Funktion $y \mapsto f(x, y)$ stetig in y_0 .
- Es gibt eine integrierbare Funktion $M : \Omega \rightarrow \overline{\mathbb{R}}_+$ mit

$$|f(x, y)| \leq M(x) \quad \text{für alle } (x, y) \in \Omega \times U.$$

Dann ist

$$g : U \rightarrow \mathbb{R}, \quad g(y) := \int_{\Omega} f(x, y) d\mu(x)$$

stetig in y_0 .

Beweis. Wir betrachten eine Folge $(y_n) \in U$ mit $y_n \rightarrow y_0$ und müssen $g(y_n) \rightarrow g(y_0)$ zeigen.

Aufgrund der Stetigkeit der Funktionen $f(x, \cdot) : y \mapsto f(x, y)$ gilt $f(\cdot, y_n) \rightarrow f(\cdot, y_0)$. Die Annahmen a) und c) garantieren, dass der Satz über majorisierte Konvergenz angewendet werden darf, und es folgt

$$\begin{aligned} \lim_{n \rightarrow \infty} g(y_n) &= \lim_{n \rightarrow \infty} \int_{\Omega} f(x, y_n) d\mu(x) = \int_{\Omega} \lim_{n \rightarrow \infty} f(x, y_n) d\mu(x) \\ &= \int_{\Omega} f(x, y_0) d\mu(x) = g(y_0). \end{aligned}$$

□

Beispiel 2.28. Die Gamma-Funktion

$$\Gamma : (0, \infty) \rightarrow \mathbb{R}, \quad \Gamma(s) := \int_0^{\infty} e^{-x} x^{s-1} dx$$

ist stetig (Übung).

Fragen wir uns nicht nur nach Stetigkeit, sondern nach (partieller) Differenzierbarkeit von g , so gibt uns die folgende Variante Auskunft:

Satz 2.29 (Parameterabhängige Integrale – Differenzierbarkeit). Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und $I \subset \mathbb{R}$ ein offenes nicht leeres Intervall. Weiter sei

$$f : \Omega \times I \rightarrow \mathbb{R}, \quad (x, t) \mapsto f(x, t)$$

eine Funktion mit den folgenden drei Eigenschaften:

- Für jedes $t \in I$ ist $x \mapsto f(x, t)$ μ -integrierbar auf Ω .
- Für jedes $x \in \Omega$ ist die Funktion $t \mapsto f(x, t)$ differenzierbar in I .
- Es gibt eine integrierbare Funktion $M : \Omega \rightarrow \overline{\mathbb{R}}_+$ mit

$$\left| \frac{\partial f}{\partial t}(x, t) \right| \leq M(x) \quad \text{für alle } (x, t) \in \Omega \times I.$$

Dann ist

$$g : I \rightarrow \mathbb{R}, \quad g(t) := \int_{\Omega} f(x, t) d\mu(x)$$

differenzierbar. Für jedes $t \in I$ ist $x \mapsto \frac{\partial f}{\partial t}(x, t)$ μ -integrierbar über Ω , und es gilt

$$g'(t) = \int_{\Omega} \frac{\partial f}{\partial t}(x, t) d\mu(x).$$

Beweis. Seit $t \in I$ fest und $h \neq 0$ so, dass $t+h \in I$. Wir betrachten die Differentialquotienten

$$f_h(x) := \frac{f(x, t+h) - f(x, t)}{h},$$

die aufgrund unserer Differenzierbarkeitsannahme $f_h(x) \rightarrow \frac{\partial f}{\partial t}(x, t)$ für $h \rightarrow 0$ erfüllen. Nach dem Mittelwertsatz der Differentialrechnung (Analysis 1, Satz 5.16 b)) existiert $\theta \in [0, 1]$ so, dass $f_h(x) = \frac{\partial f}{\partial t}(x, t + \theta h)$. Das zeigt, dass die Majorante M aus Annahme c) $|f_h(x)| \leq M(x)$ für alle $x \in \Omega$ erfüllt. Wir können als majorisierte Konvergenz anwenden, und erhalten

$$g'(t) = \lim_{h \rightarrow 0} \frac{g(t+h) - g(t)}{h} = \lim_{h \rightarrow 0} \int_{\Omega} f_h(x) d\mu(x) = \int_{\Omega} \frac{\partial f}{\partial t}(x, t) d\mu(x),$$

wie behauptet. □

Beispiel 2.30. Die Gamma-Funktion

$$\Gamma : (0, \infty) \rightarrow \mathbb{R}, \quad \Gamma(s) := \int_0^{\infty} e^{-x} x^{s-1} dx$$

ist beliebig oft differenzierbar, mit Ableitungen

$$\Gamma^{(n)}(s) = \int_0^{\infty} e^{-x} x^{s-1} (\log x)^n dx, \quad n \in \mathbb{N}.$$

2.5 L^p -Räume

Wir möchten nun den Raum der integrierbaren Funktionen auf einem Maßraum $(\Omega, \mathcal{A}, \mu)$ genauer untersuchen. Zuerst erinnern wir uns daran, dass $\mathcal{L}^1(\Omega) := \mathcal{L}^1(\Omega, \mathcal{A}, \mu)$ die Menge aller endlichen numerischen Funktionen $f : \Omega \rightarrow \mathbb{R}$ ist, die integrierbar sind, d.h. die $\int_{\Omega} |f| d\mu < \infty$ erfüllen. Wir beschränken uns hier auf die endlichen Funktionen (anstelle von $f : \Omega \rightarrow \overline{\mathbb{R}}$) aus zwei Gründen: Einerseits stellen wir so sicher, dass Linearkombinationen $c_1 f_1 + c_2 f_2$ für $c_1, c_2 \in \mathbb{R}$ und $f_1, f_2 \in \mathcal{L}^1(\Omega)$ wohldefiniert sind ($\infty - \infty$ tritt nicht auf), und $\mathcal{L}^1(\Omega)$ ist ein Vektorraum über \mathbb{R} , und andererseits sagt uns Satz 2.17, dass eine integrierbare Funktion $f : \Omega \rightarrow \overline{\mathbb{R}}$ nach einer Abänderung auf einer Nullmenge zu einer Funktion $g : \Omega \rightarrow \mathbb{R}$ wird, deren Integral mit dem von f übereinstimmt.

Sowohl in der Definition der Integrierbarkeit als auch in der Konvergenzaussage des Satzes über majorisierte Konvergenz treffen wir auf die Größe $\int_{\Omega} |f| d\mu$. Wir verabreden deshalb die Abkürzung

$$\|f\|_1 := \int_{\Omega} |f| d\mu \in [0, \infty]$$

für eine beliebige messbare Funktion $f : \Omega \rightarrow \mathbb{R}$. Dann ist

$$\mathcal{L}^1(\Omega) = \{f : \Omega \rightarrow \mathbb{R} \text{ messbar} : \|f\|_1 < \infty\},$$

und $\|\cdot\|_1$ erinnert uns sehr an eine Norm (Def. 8.1 in Analysis 2): Sicherlich nimmt $\|\cdot\|_1$ auf $\mathcal{L}^1(\Omega)$ nur nicht negative Werte an, und außerdem haben wir

- $\|cf\|_1 = |c|\|f\|_1$ für $c \in \mathbb{R}$, $f \in \mathcal{L}^1(\mathbb{R})$ (Satz 2.16)
- $\|f + g\|_1 \leq \|f\|_1 + \|g\|_1$ für $f, g \in \mathcal{L}^1(\mathbb{R})$

Aber ein Problem gibt es: Es gibt Funktionen $f \neq 0$ in $\mathcal{L}^1(\Omega)$ mit $\|f\|_1 = 0$. Deshalb ist $\|\cdot\|_1$ keine Norm, sondern nur eine sogenannte Halbnorm. Immerhin wissen wir aus Satz 2.17 für $f \in \mathcal{L}^1(\Omega)$

$$\|f\|_1 = 0 \iff f = 0 \text{ } \mu\text{-fast überall.}$$

Wir wollen diesen Misstand jetzt beheben, indem wir Funktionen miteinander identifizieren, die sich nur auf Nullmengen unterscheiden.

Dazu bemerken wir zuerst, dass aufgrund der Halbnormeigenschaften von $\|\cdot\|_1$

$$\mathcal{N} := \{f \in \mathcal{L}^1(\Omega) : \|f\|_1 = 0\}$$

ein Untervektorraum von $\mathcal{L}^1(\Omega)$ ist. Wir wollen alle Funktionen in \mathcal{N} mit der Nullfunktion $f = 0$ identifizieren. Dazu betrachten wir den Quotientenvektorraum⁷

$$L^1(\Omega) = L^1(\Omega, \mathcal{A}, \mu) = \mathcal{L}^1(\Omega)/\mathcal{N}.$$

Klassen $[f] \in L^1(\Omega)$ können wir eine Norm zuordnen durch

$$\|[f]\|_1 := \|f\|_1.$$

Dies ist wohldefiniert, also unabhängig von der Wahl eines Repräsentanten der Klasse $[f]$, da für jedes $g \in [f]$ gilt: $g = f + n$, $n \in \mathcal{N}$, und deshalb $\|g\|_1 = \|f + n\|_1 \leq \|f\|_1 + 0$ sowie $\|f\|_1 = \|g - n\|_1 \leq \|g\|_1$, also $\|f\|_1 = \|g\|_1$.

Weiterhin gilt

$$\|[f]\|_1 = 0 \Rightarrow f = 0 \text{ fast überall} \Rightarrow [f] = [0].$$

Wir halten unsere Überlegungen in folgendem Satz fest.

Satz 2.31. Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum. Dann ist $L^1(\Omega, \mathcal{A}, \mu)$ ein normierter Vektorraum.

Die Elemente von $L^1(\Omega)$ sind also Äquivalenzklassen von Funktionen $f : \Omega \rightarrow \mathbb{R}$, die sich voneinander nur auf Nullmengen unterscheiden. Es ist üblich, die Äquivalenzklassen einfach mit f statt $[f]$ zu bezeichnen, und Elemente von $L^1(\Omega)$ als "Funktionen" zu bezeichnen,

⁷Erinnerung an Lineare Algebra: Ist V ein Vektorraum und $U \subset V$ ein Untervektorraum, so ist der Quotientenvektorraum V/U die Menge aller Äquivalenzklassen $[v]$ bzgl. der Äquivalenzrelation $v_1 \sim v_2 := \Leftrightarrow v_1 - v_2 \in U$. Dies ist ein Vektorraum mit den Operationen $[v] + [w] := [v + w]$ und $c[v] := [cv]$ für $v, w \in V$, $c \in \mathbb{R}$.

obwohl es ja Äquivalenzklassen von Funktionen sind. Das ist zwar eine kleine mathematische Schlamperei, aber für Zwecke der Integration gerechtfertigt, da sich $\int_{\Omega} f d\mu$ nicht ändert, wenn f auf einer Nullmenge geändert wird (Satz 2.17).

Allerdings hat eine "Funktion" (Äquivalenzklasse) in $L^1(\Omega)$ zumeist keinen wohldefinierten Wert an einem Punkt $x \in \Omega$. Betrachten wir z.B. den Lebesgue-Borel Maßraum $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ und $f \in L^1(\mathbb{R})$, so macht $f(x)$ (für ein $x \in \mathbb{R}$) keinen Sinn, da f durch Änderung auf der Nullmenge $\{x\}$ zu einem anderen Wert von $f(x)$ führt, dh $f(x)$ hängt ab von der Wahl eines Repräsentanten in $[f]$.

Nachdem wir $L^1(\Omega)$ definiert haben, gehen wir einen Schritt weiter und definieren $L^p(\Omega)$, $p \in [1, \infty]$, analog zu unserem Vorgehen in Analysis 2 und den dort diskutierten p -Normen. Wir betrachten zunächst $1 \leq p < \infty$ und definieren

$$\|f\|_p := \|f^p\|_1^{1/p} = \left(\int_{\Omega} |f|^p d\mu \right)^{1/p} \in [0, \infty].$$

Genau wie für $\|\cdot\|_1$ ist klar, dass $\|\cdot\|_p$ auf dem Quotientenvektorraum

$$L^p(\Omega) := \mathcal{L}^p(\Omega)/\mathcal{N}_p, \quad \mathcal{L}^p(\Omega) := \{f : \Omega \rightarrow \mathbb{R} \text{ messbar} : \|f\|_p < \infty\}, \\ \mathcal{N}_p := \{f \in \mathcal{L}^p(\Omega) : f = 0 \text{ fast überall}\}$$

wohldefiniert ist. Um zu zeigen, dass $\|\cdot\|_p$ eine Norm ist, müssen wir die Dreiecksungleichung nachprüfen (alle anderen Eigenschaften sind klar). Das geht in enger Analogie zu unserem Vorgehen für die Normen $\|x\|_p = (|x_1|^p + \dots + |x_d|^p)^{1/p}$ für $x \in \mathbb{R}^d$ in Analysis 2 (siehe Satz 8.8 und 8.10), so dass wir auf den Beweis des folgenden Satzes verzichten können.

Satz 2.32. Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und $f, g : \Omega \rightarrow \overline{\mathbb{R}}$ messbar.

a) Gegeben zwei reelle Zahlen $p, q > 1$ mit $\frac{1}{p} + \frac{1}{q} = 1$, so gilt die Höldersche Ungleichung

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

b) Für jedes $p \geq 1$ gilt die Dreiecksungleichung

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

c) Für jedes $p \geq 1$ ist $L^p(\Omega)$ ein normierter Vektorraum.

Beispiel 2.33.

a) Die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = e^{-|x|}$, liegt in $L^p(\mathbb{R})$ für alle $p \geq 1$. Beweis: f

ist stetig, also messbar, und

$$\begin{aligned}
 \|f\|_p^p &= \int_{\mathbb{R}} e^{-p|x|} d\lambda(x) \\
 &= 2 \lim_{R \rightarrow \infty} \int_0^R e^{-px} dx \\
 &= \frac{2}{p} \lim_{R \rightarrow \infty} \int_0^{pR} e^{-y} dy \\
 &= -\frac{2}{p} \lim_{R \rightarrow \infty} [e^{-y}]_0^{pR} \\
 &= \frac{2}{p} < \infty.
 \end{aligned}$$

b) Sei $\alpha > 0$. Die Funktion $g_\alpha : \mathbb{R} \rightarrow \mathbb{R}$, $g_\alpha(x) = x^{-\alpha} \chi_{[1, \infty)}(x)$, liegt in $L^p(\mathbb{R})$ genau dann, wenn $p\alpha > 1$.

Da g_α messbar ist, müssen wir nur prüfen, ob $\|g_\alpha\|_p$ endlich ist.

$$\begin{aligned}
 \|g_\alpha\|_p^p &= \int_{\mathbb{R}} (x^{-\alpha} \chi_{[0, \infty)}(x))^p d\lambda(x) = \int_1^\infty x^{-\alpha p} dx \\
 &= \lim_{R \rightarrow \infty} \int_1^R x^{-\alpha p} dx = \begin{cases} \lim_{R \rightarrow \infty} \frac{1}{1-\alpha p} [x^{1-\alpha p}]_1^R & \alpha p \neq 1 \\ \lim_{R \rightarrow \infty} [\log(x)]_1^R & \alpha p = 1 \end{cases} \\
 &= \begin{cases} \frac{1}{\alpha p - 1} < \infty & \alpha p > 1 \\ \infty & \alpha p < 1 \\ \infty & \alpha p = 1 \end{cases}.
 \end{aligned}$$

c) Sei $\alpha > 0$. Die Funktion $h_\alpha : \mathbb{R} \rightarrow \mathbb{R}$, $h_\alpha(x) = x^{-\alpha} \chi_{[0, 1]}(x)$ liegt in $L^p(\mathbb{R})$ genau dann, wenn $\alpha p < 1$.

Wir gehen vor wie in b) und berechnen

$$\begin{aligned}
 \|h_\alpha\|_p^p &= \int_0^1 x^{-\alpha p} dx = \lim_{\varepsilon \searrow 0} \int_\varepsilon^1 x^{-\alpha p} dx \\
 &= \begin{cases} \lim_{\varepsilon \searrow 0} \frac{1}{1-\alpha p} [x^{1-\alpha p}]_\varepsilon^1 & \alpha p \neq 1 \\ \lim_{\varepsilon \searrow 0} [\log(x)]_\varepsilon^1 & \alpha p = 1 \end{cases} \\
 &= \begin{cases} \frac{1}{\alpha p - 1} < \infty & \alpha p < 1 \\ \infty & \alpha p > 1 \\ \infty & \alpha p = 1 \end{cases}.
 \end{aligned}$$

Da $L^p(\Omega)$ ein normierter Vektorraum ist, haben wir nun eine natürliche Art und Weise, Konvergenz von Folgen $(f_n) \subset L^p(\Omega)$ zu betrachten. Wir sagen, dass $f_n \rightarrow f$ (für $f \in$

$L^p(\Omega)$ in L^p -Norm konvergiert, falls

$$0 = \lim_{n \rightarrow \infty} \|f_n - f\|_p = \lim_{n \rightarrow \infty} \left(\int_{\Omega} |f_n - f|^p d\mu \right)^{1/p}.$$

Es ist wichtig, diesen Konvergenzbegriff sauber vom Begriff der punktweisen Konvergenz zu unterscheiden. Konvergenz in L^p -Norm bezieht sich auf eine durch das Integral gemittelte Konvergenz, die weder punktweise Konvergenz impliziert noch von punktweiser Konvergenz impliziert wird. Die folgenden Beispiele illustrieren diese Behauptung.

Beispiel 2.34.

- a) Die Folge $f_n : \mathbb{R} \rightarrow \mathbb{R}$, $f_n := n\chi_{[1/n, 2/n]}$ liegt in $L^p(\mathbb{R})$ für alle $p \geq 1$, und konvergiert offenbar punktweise gegen die Nullfunktion, $\lim_{n \rightarrow \infty} f_n(x) = 0$ für alle $x \in \mathbb{R}$. Aber

$$\|f_n - 0\|_p^p = \int_{1/n}^{2/n} n^p dx = n^{p-1} \geq 1,$$

also konvergiert $\|f_n - 0\|_p^p$ nicht gegen Null.

- b) Die Folge $f_n : \mathbb{R} \rightarrow \mathbb{R}$, $f_n(x) = \frac{1}{1+n^2x^2}$ liegt in $L^1(\mathbb{R})$ und konvergiert gegen $f = 0$ in L^1 -Norm:

$$\|f_n\|_1 = \int_{\mathbb{R}} \frac{1}{1+n^2x^2} dx = \frac{1}{n} \int_{\mathbb{R}} \frac{1}{1+y^2} dy = \frac{\pi}{n} \rightarrow 0$$

Punktweise gilt hingegen:

$$\lim_{n \rightarrow \infty} f_n(x) = \lim_{n \rightarrow \infty} \frac{1}{1+n^2x^2} = \begin{cases} 0 & x \neq 0 \\ 1 & x = 0 \end{cases} = \chi_{\{0\}}(x).$$

Also konvergiert f_n nicht punktweise gegen 0. Allerdings gilt $f_n \rightarrow 0$ fast überall.

Es gibt aber auch Beispiele von Funktionenfolgen (f_n) , die in L^p -Norm konvergieren, aber für kein x konvergiert $f_n(x)$.

Für einen normierten Raum $(V, \|\cdot\|)$ können wir Konvergenz in Norm und Cauchyfolgen in Norm betrachten. Erinnerungen an Analysis 2:

- Eine Folge $(v_n) \subset V$ heißt *normkonvergent*, falls es $v \in V$ gibt (den Grenzwert der Folge), so dass $\|v_n - v\| \rightarrow 0$ für $n \rightarrow \infty$.
- Eine Folge $(v_n) \subset V$ heißt *Cauchyfolge in Norm*, falls $\|v_n - v_m\| \rightarrow 0$ für $n, m \rightarrow \infty$.
- Jede normkonvergente Folge ist eine Cauchyfolge in Norm.
- Falls auch jede Cauchyfolge in Norm normkonvergent ist, so nennen wir $(V, \|\cdot\|)$ *Banachraum* oder sagen, dass $(V, \|\cdot\|)$ *vollständig in Norm* ist.

In einem Banachraum gibt es also zu jeder Cauchyfolge (v_n) in Norm einen Vektor $v \in V$, so dass $\|v_n - v\| \rightarrow 0$ für $n \rightarrow \infty$.

Satz 2.35. Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und $p \geq 1$. Dann ist $(L^p(\Omega), \|\cdot\|_p)$ ein Banachraum, d.h. zu jeder Cauchyfolge $(f_n) \subset L^p(\Omega)$ in Norm gibt es $f \in L^p(\Omega)$, so dass $\|f_n - f\|_p \rightarrow 0$ für $n \rightarrow \infty$. Weiterhin existiert eine Teilfolge $(f_{n_k})_k$ von $(f_n)_n$, die punktweise fast überall gegen f konvergiert.

Beweis. Wir führen den Beweis der Einfachheit halber nur für $p = 1$. Den allgemeinen Fall können Sie zB im Buch von Fischer (Seite 138) nachlesen.

Da (f_n) eine Cauchyfolge in Norm ist, können wir eine Teilfolge $(f_{n_k})_k$ so finden, dass

$$\|f_n - f_{n_k}\|_1 < \frac{1}{2^k}, \quad n \geq n_k.$$

Wir behaupten, dass die Teilfolge $(f_{n_k})_k$ punktweise fast überall konvergiert. Dazu setzen wir $g_k := f_{n_{k+1}} - f_{n_k}$. Nach Konstruktion gilt dann $\|g_k\|_1 < \frac{1}{2^k}$. Da $l \mapsto \sum_{k=1}^l |g_k|$ eine monoton wachsende Folge nicht-negativer messbarer Funktion ist, können wir den Satz über monotone Konvergenz anwenden, um zu schließen:

$$\int_{\Omega} \sum_{k=1}^{\infty} |g_k| d\mu = \sum_{k=1}^{\infty} \int_{\Omega} |g_k| d\mu = \sum_{k=1}^{\infty} \|g_k\|_1 < \sum_{k=1}^{\infty} \frac{1}{2^k} = 1.$$

Die Funktion $h := \sum_{k=1}^{\infty} |g_k|$ ist also integrierbar. Nach Satz 2.17 b) ist $N := \{h = \infty\}$ eine Nullmenge.

Wir geben nun den Kandidaten für den Grenzwert der Folge $(f_{n_k})_k$ an, nämlich

$$f(x) := \begin{cases} f_{n_1}(x) + \sum_{k=1}^{\infty} (f_{n_{k+1}}(x) - f_{n_k}(x)) & x \notin N, \\ 0 & x \in N. \end{cases}$$

Für $x \notin N$ gilt dann

$$f(x) = \lim_{K \rightarrow \infty} (f_{n_1}(x) + \sum_{k=1}^K (f_{n_{k+1}}(x) - f_{n_k}(x))) = \lim_{K \rightarrow \infty} f_{n_K}(x),$$

wie behauptet.

Außerdem haben wir für $x \notin N$ und beliebiges $k \in \mathbb{N}$ die Abschätzung

$$\begin{aligned} |f_{n_k}(x)| &= |f_{n_1}(x) + f_{n_2}(x) - f_{n_1}(x) + f_{n_3}(x) - f_{n_2}(x) + \dots + f_{n_k}(x) - f_{n_{k-1}}(x)| \\ &\leq |f_{n_1}(x)| + \sum_{l=2}^k |f_{n_l}(x) - f_{n_{l-1}}(x)| \\ &\leq |f_{n_1}(x)| + \sum_{l=1}^{\infty} |f_{n_{l+1}}(x) - f_{n_l}(x)| =: M(x). \end{aligned}$$

Wir haben bereits oben gezeigt, dass $M = |f_{n_1}| + \sum_{l=1}^{\infty} |g_l|$ integrierbar ist. Da auch $f_{n_k} \rightarrow f$ fast überall und $|f_{n_k}| \leq M$ fast überall, können wir den Satz über majorisierte Konvergenz anwenden und schließen

$$0 = \lim_{k \rightarrow \infty} \int_{\Omega} |f_{n_k} - f| d\mu = \|f_{n_k} - f\|_1,$$

wie behauptet. □

Bisher haben wir nur $p \in [1, \infty)$ betrachtet. Den Fall $p = \infty$ definieren wir nun.

Definition 2.36. Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und $f : \Omega \rightarrow \mathbb{R}$ messbar.

- Eine *wesentliche Schranke* von f ist eine Zahl $C > 0$, so dass $\{|f| > C\}$ eine Nullmenge ist.
- f heißt *wesentlich beschränkt*, falls sie eine wesentliche Schranke hat.
- Das *wesentliche Supremum* von f ist definiert als

$$\|f\|_\infty := \inf\{C \in \mathbb{R} : C \text{ ist wesentliche Schranke von } f\},$$

mit der Konvention $\inf \emptyset = \infty$.

- Der Raum der wesentlich beschränkten Funktionen ist

$$\mathcal{L}^\infty(\Omega) := \{f : \Omega \rightarrow \mathbb{R} \text{ messbar} : \|f\|_\infty < \infty\}.$$

Beispiel 2.37. Die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = 2e^{-x^2} + \chi_{\mathbb{Q}}(x)e^{+x^2}$, ist unbeschränkt aber wesentlich beschränkt. Das wesentliche Supremum ist 2.

Bei der Notation $\|f\|_\infty$ ist Vorsicht angebracht, da das gleiche Symbol mitunter für die Supremumsnorm $\sup_x |f(x)|$ verwendet wird. Manchmal wird deshalb auch $\|f\|_{L^\infty}$ für das wesentliche Supremum von f geschrieben.

Genau wie für $p < \infty$ definieren wir nun den Quotientenvektorraum

$$L^\infty(\Omega) := \mathcal{L}^\infty(\Omega) / \{f \in \mathcal{L}^\infty(\Omega) : \|f\|_\infty = 0\}.$$

Die oben bewiesenen Aussagen für $p < \infty$ übertragen sich auf den Fall $p = \infty$:

Zeigen Sie:

- Für messbare Funktionen $f, g : \Omega \rightarrow \mathbb{R}$ gilt die Höldersche Ungleichung

$$\|fg\|_1 \leq \|f\|_\infty \|g\|_1.$$

- $\|\cdot\|_\infty$ ist eine Norm auf $L^\infty(\Omega)$.
- $(L^\infty(\Omega), \|\cdot\|_\infty)$ ist ein Banachraum.

Die L^p -Räume gehören zu den wichtigsten Räumen der Analysis überhaupt. Unter ihnen sticht L^2 als Spezialfall hervor, da in diesem Fall die Norm von einem Skalarprodukt kommt. Wie in Übung P7.1 gezeigt wird, ist

$$\begin{aligned} \langle \cdot, \cdot \rangle &: L^2(\mathbb{R}) \times L^2(\mathbb{R}) \rightarrow \mathbb{R}, \\ \langle f, g \rangle &:= \int_{\Omega} fg \, d\mu \end{aligned}$$

ein wohldefiniertes Skalarprodukt, dass die Norm $\|\cdot\|_2$ via

$$\|f\|_2 = \sqrt{\langle f, f \rangle} = \left(\int_{\Omega} |f|^2 d\mu \right)^{1/2}$$

festlegt. Ein Banachraum, dessen Norm auf diese Art und Weise durch ein Skalarprodukt gegeben ist, heißt *Hilbertraum*. $L^2(\Omega)$ ist also ein Beispiel eines Hilbertraums.

Hilberträume treten insbesondere in der Funktionalanalysis und ihren Anwendungen auf. Analog zu \mathbb{R}^d mit Euklidischem Skalarprodukt können wir in Hilberträumen von Orthogonalität und Orthonormalbasen sprechen. So stehen zB die beiden Funktionen

$$f, g : \mathbb{R} \rightarrow \mathbb{R}, \quad f(x) = xe^{-x^2}, \quad g(x) = e^{-x^2}$$

bzgl des L^2 -Skalarproduktes aufeinander senkrecht, denn

$$\langle f, g \rangle = \int_{\mathbb{R}} x e^{-2x^2} dx = 0$$

aufgrund der Antisymmetrie des Integranden.

Eine *Orthonormalbasis* in einem Hilbertraum \mathcal{H} mit Skalarprodukt $\langle \cdot, \cdot \rangle$ ist eine Menge $(e_n)_{n \in N} \in \mathcal{H}$ von Vektoren indiziert durch eine Indexmenge N (oft $N = \mathbb{N}$), so dass a) $\langle e_n, e_m \rangle = \delta_{n,m}$ (die Basisvektoren stehen paarweise senkrecht aufeinander) und b) der einzige Vektor $\psi \in \mathcal{H}$, der $\langle \psi, e_n \rangle = 0$ für alle $n \in N$ erfüllt, ist $\psi = 0$ (die Basisvektoren "spannen den ganzen Raum auf").

Eine Orthonormalbasis erlaubt es, beliebige Vektoren $\xi \in \mathcal{H}$ in die Basis zu entwickeln, und es gilt $\xi = \sum_{n \in N} \langle e_n, \xi \rangle e_n$. Die Konvergenz dieser Reihe ist im Sinne der durch das Skalarprodukt gegebenen Norm zu verstehen.

Analog zur Linearen Algebra, in der sie sich mit endlichdimensionalen Hilberträumen befasst haben, kann man auch im unendlichdimensionalen Fall fragen: Gegeben eine lineare Abbildung $A : \mathcal{H} \rightarrow \mathcal{H}$, gibt es eine Orthonormalbasis aus Eigenvektoren? Diese Fragen sind im unendlichdimensionalen Fall sehr viel reichhaltiger und werden in der Spektraltheorie studiert.

2.6 Produktmaße und der Satz von Fubini

Bisher haben wir eine gut funktionierende Theorie der Integration entwickelt, die allerdings noch einen Schönheitsfehler hat: Bis auf Integrale, die wir auf (uneigentliche) Riemann-Integrale über Teilmengen von \mathbb{R} zurückführen können, haben wir noch keine Methoden entwickelt, um Integrale explizit zu bestimmen. Um dies zu beheben, möchten wir nun Integrale über Teilmengen von \mathbb{R}^d betrachten, also "in mehreren Variablen integrieren", was in engem Zusammenhang mit der Berechnung von Volumina von Teilmengen von \mathbb{R}^d steht.

Wir denken an Funktionen $\mathbb{R}^d \ni (x_1, \dots, x_d) \mapsto f(x_1, \dots, x_d) \in \mathbb{R}$ und, für messbares $U \subset \mathbb{R}^d$, an Integrale der Art

$$\int_U f(x_1, \dots, x_d) d\lambda_d(x_1, \dots, x_d).$$

Bei der Entwicklung der Differentialrechnung in mehreren Variablen (Analysis 2) konnten wir Differentiale im \mathbb{R}^d in vielerlei Hinsicht auf Ableitungen in einer Variable (partielle Ableitungen) zurückführen. Für die Integralrechnung stellt sich analog die Frage, wie sich ein

Integral über eine Teilmenge von \mathbb{R}^d , $d > 1$, (was wir ja bereits definiert haben) auf mehrere Integrale über Teilmengen von \mathbb{R} zurückführen lässt. Andersherum stellt sich allgemein auch die Frage, wie eine Funktion $(x, y) \mapsto f(x, y)$ von zwei (oder mehreren) Variablen integriert werden kann, wenn x aus einem Maßraum $(\Omega, \mathcal{A}, \mu)$ und y aus einem anderen Maßraum $(\Omega', \mathcal{A}', \mu')$ kommt, dh wie wir ein geeignetes Maß auf $\Omega \times \Omega'$ definieren können. In diesem Abschnitt werden all diese Fragen geklärt werden.

Wir beginnen als Vorbereitung mit einer Betrachtung von Lebesgue-Maßen von Teilmengen von \mathbb{R}^d . Dazu zerlegen wir $\mathbb{R}^d = \mathbb{R}^k \times \mathbb{R}^l$ mit $k + l = d$, d.h. wir betrachten Punkte $(x, y) \in \mathbb{R}^d$ mit $x \in \mathbb{R}^k$ und $y \in \mathbb{R}^l$. Gegeben eine Menge $A \subset \mathbb{R}^k \times \mathbb{R}^l$, betrachten wir die Schnitte

$$A^y = \{x \in \mathbb{R}^k : (x, y) \in A\}, \quad y \in \mathbb{R}^l.$$

Falls $A = B \times C$ eine Produktmenge ist, ist $A^y = B$ für $y \in C$ und $A^y = \emptyset$ für $y \notin C$. Im Allgemeinen wird A^y aber auf kompliziertere Weise von y abhängen.

Der folgende Satz verbindet das Volumen von A (bzgl. dem Lebesguemaß λ_d in $d = k + l$ Dimensionen) mit dem Volumen der A^y (bzgl. dem Lebesguemaß λ_k in k Dimensionen). Dieser Satz und Varianten davon sind als *Cavalierische Prinzip* bekannt.

Satz 2.38. Sei $A \subset \mathbb{R}^k \times \mathbb{R}^l$ eine Borelmenge.

a) Für jedes $y \in \mathbb{R}^l$ ist A^y eine Borelmenge (in \mathbb{R}^k).

b) Die Funktion

$$F_A : \mathbb{R}^l \rightarrow \overline{\mathbb{R}}_+, \quad F_A(y) := \lambda_k(A^y)$$

ist Borel-messbar.

c) Es gilt mit $d = k + l$

$$\lambda_d(A) = \int_{\mathbb{R}^l} \lambda_k(A^y) d\lambda_l(y).$$

Wir verzichten an dieser Stelle auf einen Beweis, da wir bald ein allgemeineres Ergebnis beweisen werden, das diesen Satz als Spezialfall enthält.

Beispiel 2.39. Wir verwenden das Cavalierische Prinzip, um das Volumen der Kugel $K_d(r) := \{x \in \mathbb{R}^d : \|x\|_2 \leq r\}$ mit Radius $r \geq 0$ im d -dimensionalen Raum zu bestimmen. Wegen $K_d(r) = r \cdot K_d(1)$ wissen wir bereits $\lambda_d(K_d(r)) = r^d \cdot \lambda_d(K_d(1))$. Wir interessieren uns also für die Zahlen

$$V_d := \lambda_d(K_d(1)).$$

Für $d = 1$ ist $K_1(1) = [-1, 1]$, also $V_1 = 2$. Für $d = 2$ und $d = 3$ wissen wir aus der Schule $V_2 = \pi$ und $V_3 = \frac{4\pi}{3}$, haben aber noch keinen Beweis. Das holen wir jetzt nach.

Zur Bestimmung der Schnittmengen von $K_d(1)$ betrachten wir $x \in \mathbb{R}^{d-1}$, $y \in \mathbb{R}$, und

haben

$$(x, y) \in K_d(1) \Leftrightarrow \|x\|_2^2 + y^2 \leq 1 \Leftrightarrow K_d(1)^y = \begin{cases} K_{d-1}(\sqrt{1-y^2}) & |y| \leq 1 \\ \emptyset & |y| > 1 \end{cases}.$$

Das Cavalierische Prinzip ergibt also

$$V_d = \int_{-1}^1 \lambda_{d-1}(K_{d-1}(\sqrt{1-y^2})) dy = V_{d-1} \cdot \int_{-1}^1 (1-y^2)^{\frac{d-1}{2}} dy.$$

Eine schöne Gelegenheit für die Substitutionsregel! (Siehe Analysis 1, Beispiel 6.22). Die Substitution $x = \sin(y)$ führt auf

$$V_d = V_{d-1} \cdot I_d, \quad I_d := 2 \int_0^{\pi/2} \cos^d(x) dx.$$

Dieses Integral können Sie auf verschiedene Weisen berechnen, zum Beispiel per partieller Integration oder indem Sie $\cos(x) = \frac{1}{2}(e^{ix} + e^{-ix})$ schreiben, ausmultiplizieren und die Exponentialfunktionen integrieren. Das Ergebnis ist (Übung)

$$I_{2d} = \pi \prod_{m=1}^d \frac{2m-1}{2m}, \quad I_{2d+1} = 2 \prod_{m=1}^d \frac{2m}{2m+1}.$$

Zum Beispiel erhalten wir $I_2 = \frac{\pi}{2}$ und damit $V_2 = V_1 \cdot I_2 = \pi$ als Fläche der zweidimensionalen Einheitskreisscheibe, oder $I_3 = \frac{4}{3}$ und damit $V_3 = \frac{4\pi}{3}$ als Volumen der dreidimensionalen Einheitskugel.

Für allgemeine Dimension d erhält man aus den obigen Formeln für die I_d

$$V_{2d} = \frac{\pi^d}{d!}, \quad V_{2d+1} = \frac{2^{d+1}}{1 \cdot 3 \cdot \dots \cdot (2d+1)} \pi^d.$$

Bemerkenswerterweise gilt $\lim_{d \rightarrow \infty} V_d = 0$. Das Volumen des $K_d(1)$ enthaltenen Würfels $[-1, 1]^d$ ist hingegen 2^d , dh insbesondere geht der Anteil des Volumens, das in der Kugel liegt, für große Dimension gegen Null.

Beispiel 2.40. Als weitere Anwendung betrachten wir ein kompaktes Intervall $I = [a, b] \subset \mathbb{R}$, eine stetige Funktion $f : I \rightarrow \mathbb{R}_+$, und die Menge

$$A := \{(x, y) \in \mathbb{R}^2 : x \in I, 0 \leq y \leq f(x)\}.$$

Dies ist die Fläche zwischen dem Intervall I und dem Graphen von f , nach Schulwissen sollte ihr Flächeninhalt $\lambda_2(A)$ gleich dem Riemannintegral $\int_a^b f(x) dx$ sein.

Das können wir mit Cavalieri jetzt auch beweisen. Dazu betrachten wir die Schnitte von A (hier entlang der y - statt entlang der x -Achse, was natürlich genauso geht), nämlich

$A_x := \{y \in \mathbb{R} : (x, y) \in A\} = [0, f(x)]$, und erhalten

$$\lambda_2(A) = \int_a^b \lambda_1([0, f(x)]) dx = \int_a^b f(x) dx.$$

Nach diesen Anwendungen (weitere Anwendungen gibt es in den Übungen) werfen wir nun einen allgemeineren Blick auf das Prinzip von Cavalieri. Die Cavalieri-Formel

$$\lambda_d(A) = \int_{\mathbb{R}^l} \lambda_k(A^y) d\lambda_l(y) = \int_{\mathbb{R}^k} \lambda_l(A_x) d\lambda_k(x)$$

stellt einen Zusammenhang zwischen den beiden Maßen λ_k auf \mathbb{R}^k und λ_l auf \mathbb{R}^l mit dem Maß λ_d auf $\mathbb{R}^k \times \mathbb{R}^l$ her. Auf Produktmengen $A = B \times C$ gilt $\lambda_d(B \times C) = \lambda_k(B) \cdot \lambda_l(C)$, so dass wir λ_d als das ‘‘Produktmaß’’ von λ_k und λ_l bezeichnen mögen. Wir interessieren uns nun für zwei Fragen:

- Wie definiert man allgemeine Produktmaße als ‘‘Produkt’’ von zwei Maßen μ_1, μ_2 ?
- Wie bestimmt man konkret Integrale bzgl. eines Produktmaßes?

Wir beginnen mit der Definition von Produktmaßen auf Grundlage von zwei gegebenen Maßen.

Definition 2.41. Seien $(\Omega_1, \mathcal{A}_1)$ und $(\Omega_2, \mathcal{A}_2)$ Messräume. Die *Produkt- σ -Algebra* von \mathcal{A}_1 und \mathcal{A}_2 ist

$$\begin{aligned} \mathcal{A}_1 \otimes \mathcal{A}_2 &:= \sigma(\mathcal{E}(\mathcal{A}_1, \mathcal{A}_2)) \subset \mathcal{P}(\Omega_1 \times \Omega_2), \\ \mathcal{E}(\mathcal{A}_1, \mathcal{A}_2) &:= \{A_1 \times A_2 : A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2\}. \end{aligned}$$

Messbarkeit bzgl. der Produkt- σ -Algebra wird auch als *Produkt-Messbarkeit* bezeichnet.

$\mathcal{A}_1 \otimes \mathcal{A}_2$ ist also die kleinste σ -Algebra auf $\Omega_1 \times \Omega_2$, die alle Produkte $A_1 \times A_2$ von messbaren Mengen enthält.

Wir betrachten nun Produktmessbarkeit von Funktionen.

Lemma 2.42. Seien $(\Omega_1, \mathcal{A}_1)$, $(\Omega_2, \mathcal{A}_2)$, (Ω, \mathcal{A}) drei Messräume. Eine Funktion $f : \Omega \rightarrow \Omega_1 \times \Omega_2$, $x \mapsto (f_1(x), f_2(x))$ ist genau dann produkt-messbar, wenn $f_1 : \Omega \rightarrow \Omega_1$ und $f_2 : \Omega \rightarrow \Omega_2$ messbar sind.

Beweis. Für $A_1 \in \mathcal{A}_1$ und $A_2 \in \mathcal{A}_2$ gilt $f^{-1}(A_1 \times A_2) = f_1^{-1}(A_1) \cap f_2^{-1}(A_2)$. Falls f_1, f_2 messbar sind, ist dies eine messbare Menge, dh $f^{-1}(A_1 \times A_2) \in \mathcal{A}$. Da die Produktmengen $A_1 \times A_2$ einen Erzeuger der Produkt- σ -Algebra bilden, folgt, dass f messbar ist (siehe Übung auf Seite 30).

Ist hingegen f als produkt-messbar angenommen, so betrachten wir die Mengen $A_1 \times \Omega_2$ und $\Omega_1 \times A_2$ (mit $A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2$) und erhalten, dass $f^{-1}(A_1 \times \Omega_2) = f_1^{-1}(A_1)$ und $f^{-1}(\Omega_1 \times A_2) = f_2^{-1}(A_2)$ messbar sind. Also sind f_1, f_2 messbar. \square

Insbesondere sind die kanonischen Projektionen $\pi_1 : \Omega_1 \times \Omega_2 \rightarrow \Omega_1$, $\pi_1(x_1, x_2) = x_1$, und $\pi_2 : \Omega_1 \times \Omega_2 \rightarrow \Omega_2$, $\pi_2(x_1, x_2) = x_2$, messbar, da sie die Komponenten der (messbaren) Identitätsfunktion $\Omega_1 \times \Omega_2 \rightarrow \Omega_1 \times \Omega_2$ sind.

Für das wichtige Beispiel der Borel- σ -Algebren auf \mathbb{R}^d sieht das so aus:

Lemma 2.43. Seien $k, l \in \mathbb{N}$ und $d := k + l$. Dann ist die Borel- σ -Algebra von $\mathbb{R}^d = \mathbb{R}^k \times \mathbb{R}^l$

$$\mathcal{B}^d = \mathcal{B}^k \otimes \mathcal{B}^l.$$

Beweis. Die Projektionen $\pi_1 : \mathbb{R}^d \rightarrow \mathbb{R}^k$, $\pi_1(x, y) = x$, und $\pi_2 : \mathbb{R}^d \rightarrow \mathbb{R}^l$, $\pi_2(x, y) = y$, sind stetig, also messbar. Für Borelmengen $B \subset \mathbb{R}^k$ und $B' \subset \mathbb{R}^l$ gilt also

$$B \times B' = (B \times \mathbb{R}^l) \cap (\mathbb{R}^k \times B') = \pi_1^{-1}(B) \cap \pi_2^{-1}(B') \in \mathcal{B}(\mathbb{R}^d),$$

d.h. nach Übergang zu den erzeugten σ -Algebren $\mathcal{B}^k \otimes \mathcal{B}^l \subset \mathcal{B}^d$.

Andererseits sind halboffene Quader Produktmengen, also Elemente von $\mathcal{B}^k \otimes \mathcal{B}^l$. Da die halboffenen Quader die Borel- σ -Algebra erzeugen, impliziert das $\mathcal{B}^d \subset \mathcal{B}^k \otimes \mathcal{B}^l$. \square

Dieses Ergebnis zeigt insbesondere, dass die Produkt- σ -Algebra im Allgemeinen viel mehr Mengen als nur Produkte $B \times B'$ von messbaren Mengen enthält.

Wir betrachten nun das Cavalierische Prinzip in diesem allgemeinen maßtheoretischen Setting. In folgendem Lemma zeigen wir in a) eine Verallgemeinerung der Aussage a) des Cavalierischen Prinzips. Aussage b) ist eine Vorbereitung für die Integration bzgl. Produktmaßen.

Lemma 2.44. Seien $(\Omega_1, \mathcal{A}_1)$, $(\Omega_2, \mathcal{A}_2)$, $(\tilde{\Omega}, \tilde{\mathcal{A}})$ Messräume und $(x_1^0, x_2^0) \in \Omega_1 \times \Omega_2$.

a) Für $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$ gilt

$$A_{x_1^0} := \{x_2 \in \Omega_2 : (x_1^0, x_2) \in A\} \in \mathcal{A}_2,$$

$$A^{x_2^0} := \{x_1 \in \Omega_1 : (x_1, x_2^0) \in A\} \in \mathcal{A}_1.$$

b) Ist $f : \Omega_1 \times \Omega_2 \rightarrow \tilde{\Omega}$ messbar, so sind auch $f_{x_1^0} : \Omega_2 \rightarrow \tilde{\Omega}$, $f_{x_1^0}(x_2) = f(x_1^0, x_2)$ und $f^{x_2^0} : \Omega_1 \rightarrow \tilde{\Omega}$, $f^{x_2^0}(x_1) = f(x_1, x_2^0)$ messbar.

Beweis. Wir zeigen den Beweis für $A_{x_1^0}$ und $f_{x_1^0}$, die Argumente für $A^{x_2^0}$ und $f^{x_2^0}$ verlaufen analog. Wir fixieren also $x_1^0 \in \Omega_1$ und betrachten die Abbildung

$$p : \Omega_2 \rightarrow \Omega_1 \times \Omega_2, \quad p(x_2) := (x_1^0, x_2).$$

Die Komponentenfunktionen von p sind die konstante Funktion $p_1 : x_2 \mapsto x_1^0$ und die Identität $p_2 : x_2 \mapsto x_2$, die beide messbar sind. Also ist p produkt-messbar (Lemma 2.42). Damit folgt, dass $A_{x_1^0} = p^{-1}(A)$ eine messbare Menge und $f_{x_1^0} = f \circ p$ eine messbare Funktion sind. \square

Mit diesem Lemma sehen wir bereits, dass für eine messbare Funktion $f : \Omega_1 \times \Omega_2 \rightarrow \mathbb{R}_+$ das nur bzgl einer Variablen gebildete Integral $x_2 \mapsto \int_{\Omega_1} f(x_1, x_2) d\mu_1(x_1)$ existiert. Wir werden bald analysieren, wann diese Funktion wiederum in x_2 messbar ist, und was das Integral $\int_{\Omega_2} \left(\int_{\Omega_1} f(x_1, x_2) d\mu_1(x_1) \right) d\mu_2(x_2)$ mit einem “Integral in zwei Variablen” (bzgl eines Produktmaßes) zu tun hat.

Wir kommen nun zu der zweiten Aussage des Cavalierischen Prinzips im allgemeinen maßtheoretischen Kontext. Hier müssen wir verlangen, dass wir es mit σ -endlichen Maßen (siehe Satz 1.38) zu tun haben. Diese Annahme ist für das Lebesguemaß erfüllt.

Lemma 2.45. *Seien $(\Omega_1, \mathcal{A}_1, \mu_1)$ und $(\Omega_2, \mathcal{A}_2, \mu_2)$ zwei σ -endliche Maßräume. Dann sind für jede produkt-messbare Menge $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$ die beiden Funktionen*

$$\begin{aligned} \mu_1^A : \Omega_1 &\rightarrow [0, \infty], & \mu_1^A(x_1) &:= \mu_2(A_{x_1}), \\ \mu_2^A : \Omega_2 &\rightarrow [0, \infty], & \mu_2^A(x_2) &:= \mu_1(A^{x_2}), \end{aligned}$$

messbar.

Beweis. Wir führen den Beweis für μ_1^A , das Argument für μ_2^A ist analog. Sei

$$\Sigma := \{A \in \mathcal{A}_1 \otimes \mathcal{A}_2 : x_1 \mapsto \mu_1^A(x_1) = \mu_2(A_{x_1}) \text{ ist messbar}\}.$$

Wir müssen zeigen, dass $\Sigma = \mathcal{A}_1 \otimes \mathcal{A}_2$ gilt. Insbesondere müssen wir sehen, dass Σ eine σ -Algebra ist.

Wie schon beim Cavalierischen Prinzip argumentiert, sehen wir, dass Σ alle Produktmengen $A_1 \times A_2$ mit $A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2$, enthält. Insbesondere gilt $\emptyset \in \Sigma$.

Für Komplemente betrachten wir $A \in \Sigma$. Wegen $(A^c)_{x_1} = (A_{x_1})^c$ gilt dann

$$\mu_1^{A^c}(x_1) = \mu_2(\Omega_2 \setminus A_{x_1}) = \mu_2(\Omega_2) - \mu_2(A_{x_1}) = \mu_2(\Omega_2) - \mu_1^A(x_1).$$

Dies ist allerdings nur wohldefiniert, wenn nicht $\infty - \infty$ auftritt. Wir nehmen deshalb zunächst an, dass μ_2 ein endliches Maß ist. In diesem Fall sehen wir, dass auch $\mu_1^{A^c}$ messbar ist, d.h. wir haben $A \in \Sigma \Rightarrow A^c \in \Sigma$.

Ist $A = \bigsqcup_n A_n$ eine abzählbare disjunkte Vereinigung von Mengen in Σ , so ist auch $A_{x_1} = \bigsqcup_n (A_n)_{x_1}$ eine abzählbare disjunkte Vereinigung von Mengen in \mathcal{A}_2 . Wie im Satz von Carathéodory (Satz 1.35) kann man dann argumentieren, dass Σ eine σ -Algebra ist. Da Σ den Erzeuger $\mathcal{E}(\mathcal{A}_1, \mathcal{A}_2)$ enthält, folgt $\Sigma = \mathcal{A}_1 \otimes \mathcal{A}_2$.

Wir müssen noch die Annahme der Endlichkeit von μ_2 zu σ -Endlichkeit abschwächen. Dazu wählen wir eine Folge $C_n \nearrow \Omega_2$ in \mathcal{A}_2 mit $\mu_2(C_n) < \infty$ für alle n und sehen, dass $\mu_1^A(x_1) = \lim_{n \rightarrow \infty} \mu_2(A_{x_1} \cap C_n)$ als Limes messbarer Funktionen messbar ist. \square

Nach diesen diversen Vorbereitungen kommen wir nun zum Satz über das Produktmaß.

Satz 2.46. *Seien $(\Omega_1, \mathcal{A}_1, \mu_1)$ und $(\Omega_2, \mathcal{A}_2, \mu_2)$ zwei σ -endliche Maßräume. Dann gibt es genau ein Maß ν auf $\mathcal{A}_1 \otimes \mathcal{A}_2$, das*

$$\nu(A_1 \times A_2) = \mu_1(A_1) \cdot \mu_2(A_2), \quad A_i \in \mathcal{A}_i,$$

erfüllt. Für $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$ gilt die Cavalieri-Formel

$$\nu(A) = \int_{\Omega_1} \mu_2(A_{x_1}) d\mu_1(x_1) = \int_{\Omega_2} \mu_1(A^{x_2}) d\mu_2(x_2).$$

Dieses Maß wird das Produktmaß von μ_1 und μ_2 genannt, und mit $\nu = \mu_1 \otimes \mu_2$ bezeichnet.

Wir sehen sofort, dass das Produktmaß der Lebesgue-Borel-Maße λ_k auf \mathcal{B}^k und λ_l auf \mathcal{B}^l das Lebesgue-Borel-Maß λ_d auf \mathcal{B}^d , $d = k + l$, ist. Somit impliziert der obige Satz das Cavalieri-Prinzip im \mathbb{R}^d . Als Knobelaufgabe können Sie sich aber überlegen:

Sei \mathcal{L}_d die Lebesgue- σ -Algebra auf \mathbb{R}^d . Zeigen Sie $\mathcal{L}_2 \neq \mathcal{L}_1 \otimes \mathcal{L}_1$.

Beweis. (des Satzes vom Produktmaß) Die Idee des Beweises ist, zu zeigen, dass die Definition $\nu(A) = \int_{\Omega_1} \mu_2(A_{x_1}) d\mu_1(x_1)$ ein Maß auf der Produkt- σ -Algebra definiert, dass $\nu(A_1 \times A_2) = \mu_1(A_1) \cdot \mu_2(A_2)$ erfüllt. Die Eindeutigkeitsaussage folgt dann aus dem Maßeindeutigkeitssatz Satz 1.38.

Zuerst bemerken wir, dass $\nu(A) \in [0, \infty]$ für jedes $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$ wegen Lemma 2.45 wohldefiniert ist.

Für Produktmengen $A = A_1 \times A_2$ (mit $A_i \in \mathcal{A}_i$) gilt aufgrund von $\mu_2((A_1 \times A_2)_{x_1}) = \chi_{A_1}(x_1) \mu_2(A_2)$

$$\nu(A_1 \times A_2) = \int_{\Omega_1} \chi_{A_1}(x_1) \mu_2(A_2) d\mu_1(x_1) = \mu_1(A_1) \cdot \mu_2(A_2),$$

wie behauptet.

Nun zeigen wir, dass ν ein Maß ist: Offenbar gilt $\nu(\emptyset) = \nu(\emptyset \times \emptyset) = \mu_1(\emptyset) \cdot \mu_2(\emptyset) = 0 \cdot 0 = 0$.

Für die σ -Additivität betrachten wir eine Folge paarweise disjunkter Mengen $A_n \in \mathcal{A}_1 \otimes \mathcal{A}_2$. Dann ist $(\bigsqcup_n A_n)_{x_1} = \bigsqcup_n ((A_n)_{x_1})$, und die $(A_n)_{x_1}$ liegen in \mathcal{A}_2 (Lemma 2.44 a)). Da μ_2 ein Maß ist (also σ -additiv ist), erhalten wir

$$\begin{aligned} \nu\left(\bigsqcup_n A_n\right) &= \int_{\Omega_1} \mu_2\left(\bigsqcup_n (A_n)_{x_1}\right) d\mu_1(x_1) \\ &= \int_{\Omega_1} \sum_{n=1}^{\infty} \mu_2((A_n)_{x_1}) d\mu_1(x_1) \\ &= \sum_{n=1}^{\infty} \int_{\Omega_1} \mu_2((A_n)_{x_1}) d\mu_1(x_1) \\ &= \sum_{n=1}^{\infty} \nu(A_n), \end{aligned}$$

wobei wir im vorletzten Schritt den Satz über monotone Konvergenz auf die Folge $N \mapsto f_N(x_1) := \sum_{n=1}^N \mu_2((A_n)_{x_1})$ angewendet haben.

Das Maß ν existiert also auf $\mathcal{A}_1 \otimes \mathcal{A}_2$ und erfüllt die erste behauptete Integralformel. Genau-
sogut könnten wir ein Maß ν' mit Hilfe der zweiten behaupteten Integralformel definieren.

Da diese Maße σ -endlich sind und auf dem schnittstabilen Erzeuger aller Produktmengen übereinstimmen, sind sie gleich. Das zeigt auch die Eindeutigkeitsaussage. \square

Nach diesem wesentlichen Schritt fragen wir uns nun, wie bzgl. eines Produktmaßes integriert wird. Wir werden gleich sehen, dass dies mittels iterierter Integrale ausgedrückt werden kann, d.h. dass eine Formel der Art

$$\begin{aligned} \int_{\Omega_1 \times \Omega_2} f(x_1, x_2) d(\mu_1 \otimes \mu_2)(x_1, x_2) &= \int_{\Omega_1} \left(\int_{\Omega_2} f(x_1, x_2) d\mu_2(x_2) \right) d\mu_1(x_1) \\ &= \int_{\Omega_2} \left(\int_{\Omega_1} f(x_1, x_2) d\mu_1(x_1) \right) d\mu_2(x_2) \end{aligned} \quad (\star)$$

gilt – dh die zu integrierende Funktion $(x_1, x_2) \mapsto f(x_1, x_2)$ von zwei Variablen x_1, x_2 wird zunächst als Funktion der einen Variablen, sagen wir x_1 , angesehen (bei festem $x_2 \in \Omega_2$) und bzgl μ_1 integriert. Das Integral über x_1 ist dann eine Funktion von x_2 , und kann bzgl μ_2 integriert werden. Die Behauptung ist, dass dieses Doppelintegral existiert, nicht von der Reihenfolge der Integrationen (erst in x_1 oder erst in x_2) abhängt, und gleich dem Integral von f bzgl. des Produktmaßes ist.

Um (\star) zu zeigen und die genauen Voraussetzungen an f zu klären, betrachten wir zunächst den Fall, dass die messbare Funktion $f = \chi_A$ eine charakteristische Funktion einer messbaren Menge $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$ ist. Dann gilt wegen $x_2 \in A_{x_1} \Leftrightarrow (x_1, x_2) \in A$

$$f(x_1, x_2) = \begin{cases} 1 & (x_1, x_2) \in A \\ 0 & (x_1, x_2) \notin A \end{cases} = \begin{cases} 1 & x_2 \in A_{x_1} \\ 0 & x_2 \notin A_{x_1} \end{cases} = \chi_{A_{x_1}}(x_2),$$

also $f(x_1, \cdot) = \chi_{A_{x_1}}$. Dies ist nach Lemma 2.45 eine messbare Funktion, und die erste Formel in (\star) stimmt mit der ersten Integralformel in Satz 2.46 überein: Die linke Seite ist $\int_{\Omega_1 \times \Omega_2} \chi_A d(\mu_1 \otimes \mu_2) = (\mu_1 \otimes \mu_2)(A)$, die rechte Seite ist

$$\int_{\Omega_1} \left(\int_{\Omega_2} \chi_{A_{x_1}}(x_2) d\mu_2(x_2) \right) d\mu_1(x_1) = \int_{\Omega_1} \mu_2(A_{x_1}) d\mu_1(x_1).$$

Für die zweite Formel lässt sich genauso argumentieren. Also haben wir (\star) bereits für charakteristische Funktionen gezeigt.

Da die Integrale auf beiden Seiten von (\star) linear sind, folgt (\star) sofort auch für nicht negative einfache Funktionen.

Gehen wir noch einen Schritt weiter und betrachten eine nicht negative messbare Funktion $f : \Omega_1 \times \Omega_2 \rightarrow \overline{\mathbb{R}}_+$. Dann gibt es nach Satz 2.6 eine aufsteigende Folge nicht negativer einfacher Funktionen f_n mit $f_n \nearrow f$. Dann gilt aber auch $f_n(x_1, \cdot) \nearrow f(x_1, \cdot)$ und per monotoner Konvergenz $\int_{\Omega_2} f_n(x_1, \cdot) d\mu_2 \nearrow \int_{\Omega_2} f(x_1, \cdot) d\mu_2$.

Also ist auch $x_1 \mapsto \int_{\Omega_2} f_n(x_1, \cdot) d\mu_2$ eine monoton wachsende Folge messbarer Funktionen, und nochmalige Anwendungen des Satzes über monotone Konvergenz ergeben

$$\begin{aligned} \int_{\Omega_1} \int_{\Omega_2} f_n(x_1, x_2) d\mu_2(x_2) d\mu_1(x_1) &\rightarrow \int_{\Omega_1} \int_{\Omega_2} f(x_1, x_2) d\mu_2(x_2) d\mu_1(x_1), \\ \int_{\Omega_1 \times \Omega_2} f_n(x_1, x_2) d(\mu_1 \otimes \mu_2)(x_1, x_2) &\rightarrow \int_{\Omega_1 \times \Omega_2} f d(\mu_1 \otimes \mu_2). \end{aligned}$$

Da die Integrale auf den linken Seiten übereinstimmen (die f_n sind einfache Funktionen), stimmen auch die Integrale auf den rechten Seiten überein.

Diese Argumentationsstrategie ist oft nützlich: Wenn Sie zeigen sollen, dass zwei Integralausdrücke über allgemeine messbare/integrierbare Funktionen übereinstimmen, so prüfen Sie es zuerst für charakteristische Funktionen und denen dann per Linearität auf einfache, und per Approximation auf allgemeine Funktionen aus.

Wir fassen unsere Überlegungen in folgendem Satz zusammen:

Satz 2.47 (Satz von Tonelli). Seien $(\Omega_i, \mathcal{A}_i, \mu_i)$, $i = 1, 2$, zwei σ -endliche Maßräume und $f : \Omega_1 \times \Omega_2 \rightarrow [0, \infty]$ produkt-messbar. Dann sind die Funktionen

$$x_1 \mapsto \int_{\Omega_2} f(x_1, x_2) d\mu_2(x_2), \quad x_2 \mapsto \int_{\Omega_1} f(x_1, x_2) d\mu_1(x_1)$$

messbar, und es gilt

$$\begin{aligned} \int_{\Omega_1 \times \Omega_2} f d(\mu_1 \otimes \mu_2) &= \int_{\Omega_1} \int_{\Omega_2} f(x_1, x_2) d\mu_2(x_2) d\mu_1(x_1) \\ &= \int_{\Omega_2} \int_{\Omega_1} f(x_1, x_2) d\mu_1(x_1) d\mu_2(x_2). \end{aligned}$$

Beispiel 2.48. Für integrierbare Funktionen $f, g : \mathbb{R} \rightarrow \mathbb{R}$ ist auch $F : \mathbb{R}^2 \rightarrow \mathbb{R}$, $F(x, y) := f(x)g(y)$, integrierbar.

Zuerst überlegen wir uns, dass F Borel-messbar ist: Die Funktionen $(x, y) \mapsto f(x)$ und $(x, y) \mapsto g(y)$ sind messbar (durch Betrachtung Urbilder messbarer Mengen), und die Produktabbildung $p : \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x, y) \mapsto x \cdot y$ ist stetig, also Borel-messbar. Also ist F als Komposition messbarer Funktionen messbar.

Für die Integrierbarkeit von F müssen wir zeigen, dass das Integral $\int_{\mathbb{R}^2} |F(x, y)| d\lambda_2(x, y)$ endlich ist:

$$\begin{aligned} \int_{\mathbb{R}^2} |F(x, y)| d\lambda_2(x, y) &= \int_{\mathbb{R}} \int_{\mathbb{R}} |f(x)| |g(y)| dx dy \\ &= \int_{\mathbb{R}} |g(y)| \int_{\mathbb{R}} |f(x)| dx dy = \int_{\mathbb{R}} |g(y)| dy \cdot \int_{\mathbb{R}} |f(x)| dx < \infty. \end{aligned}$$

Im Satz von Tonelli würden wir gerne noch die Annahme, dass f nicht negativ ist, fallenlassen, und Messbarkeit von f durch Integrierbarkeit ersetzen. Das wird allerdings nur bis auf Nullmengen gehen, wie wir an einem Beispiel verdeutlichen:

Sei $b : \mathbb{R} \rightarrow \mathbb{R}$ eine beliebige messbare Funktion und

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x_1, x_2) = \begin{cases} b(x_1) & x_2 = 0 \\ 0 & x_2 \neq 0 \end{cases}.$$

Dann ist f messbar, und es gilt $f = 0$ fast überall bzgl. des Produktmaßes $\lambda_2 = \lambda_1 \otimes \lambda_1$. Aber die Einschränkung $f(\cdot, 0) = b$ war eine beliebige messbare Funktion, muss also nicht integrierbar sein.

Satz 2.49 (Satz von Fubini). Seien $(\Omega_i, \mathcal{A}_i, \mu_i)$, $i = 1, 2$, zwei σ -endliche Maßräume und $f : \Omega_1 \times \Omega_2 \rightarrow \overline{\mathbb{R}}$ eine bzgl. des Produktmaßes $\mu_1 \otimes \mu_2$ integrierbare Funktion.

a) Es gibt eine μ_1 -Nullmenge N_1 , so dass

$$f_{x_1} : \Omega_2 \rightarrow \overline{\mathbb{R}}, \quad f_{x_1}(x_2) = f(x_1, x_2),$$

für alle $x_1 \in \Omega_1 \setminus N_1$ μ_2 -integrierbar ist, und es gibt eine μ_2 -Nullmenge N_2 , so dass

$$f^{x_2} : \Omega_1 \rightarrow \overline{\mathbb{R}}, \quad f^{x_2}(x_1) := f(x_1, x_2),$$

für alle $x_2 \in \Omega_2 \setminus N_2$ μ_1 -integrierbar ist.

b) Die Funktionen

$$\begin{aligned} \Omega_1 \ni x_1 &\mapsto \int_{\Omega_2} \chi_{\Omega_1 \setminus N_1}(x_1) f_{x_1}(x_2) d\mu_2(x_2), \\ \Omega_2 \ni x_2 &\mapsto \int_{\Omega_1} \chi_{\Omega_2 \setminus N_2}(x_2) f^{x_2}(x_1) d\mu_1(x_1), \end{aligned}$$

sind integrierbar, und es gilt

$$\begin{aligned} \int_{\Omega_1 \times \Omega_2} f d(\mu_1 \otimes \mu_2) &= \int_{\Omega_1} \left(\int_{\Omega_2} f(x_1, x_2) d\mu_2(x_2) \right) d\mu_1(x_1) \\ &= \int_{\Omega_2} \left(\int_{\Omega_1} f(x_1, x_2) d\mu_1(x_1) \right) d\mu_2(x_2). \end{aligned}$$

Beweis. a) Lemma 2.44 liefert die Messbarkeit der Funktionen f_{x_1} und f^{x_2} , und der Satz von Tonelli ergibt $\int_{\Omega_1} \left(\int_{\Omega_2} |f(x_1, x_2)| d\mu_2(x_2) \right) d\mu_1(x_1) = \int_{\Omega_1 \times \Omega_2} |f| d(\mu_1 \otimes \mu_2) < \infty$

Nach Satz 2.17 gibt es eine Nullmenge $N_1 \subset \Omega_1$, so dass der innere Integrand $\int_{\Omega_2} |f(x_1, x_2)| d\mu_2(x_2)$ für $x_1 \notin N_1$ endlich ist, dh f_{x_1} ist integrierbar für $x_1 \notin N_1$. Analog argumentiert man für f^{x_2} .

b) Man zerlegt $f = f_+ - f_-$ und wendet den Satz von Tonelli auf den Positiv- und Negativteil an. □

Der Satz von Fubini ist ein zentrales Ergebnis der Maßtheorie, das auch enormen praktischen Nutzen hat. Wir machen einige Bemerkungen zu diesem Satz und illustrieren ihn an einigen Beispielen.

- Per vollständiger Induktion dehnt sich der Satz von Fubini auf Funktionen in $d \in \mathbb{N}$ Variablen aus. Insbesondere können wir das Integral über eine bzgl. des d -dimensionalen Lebesguemaßes integrierbare Funktion $f : \mathbb{R}^d \rightarrow \mathbb{R}$ als ein d -faches Integral über Funktionen $\mathbb{R} \rightarrow \mathbb{R}$ schreiben:

$$\int_{\mathbb{R}^d} f(x_1, \dots, x_d) d\lambda_d(x_1, \dots, x_d) = \int_{\mathbb{R}} \int_{\mathbb{R}} \dots \int_{\mathbb{R}} f(x_1, \dots, x_d) dx_1 \dots dx_d.$$

Die Reihenfolge der Integrationen ist hierbei irrelevant.

- Wird über einen Produktbereich wie zB den Quader $Q = [a_1, b_1] \times \dots \times [a_d, b_d]$ integriert, so erhalten wir

$$\int_Q f d\lambda_d = \int_{\mathbb{R}^d} \chi_Q f d\lambda_d = \int_{a_1}^{b_1} \int_{a_2}^{b_2} \dots \int_{a_d}^{b_d} f(x_1, \dots, x_d) dx_d dx_{d-1} \dots dx_1.$$

Hier sind die Integrale und zugehörigen dx_i also von innen nach außen auszuführen, dh zuerst Integration über x_d über das Intervall $[a_d, b_d]$, dann Integration über x_{d-1} über das Intervall $[a_{d-1}, b_{d-1}]$, etc. Vielerorts (vor allem in der Physik) findet sich für solch ein Integral auch die Notation

$$\int_{a_1}^{b_1} dx_1 \int_{a_2}^{b_2} dx_2 \dots \int_{a_d}^{b_d} dx_d f(x_1, \dots, x_d),$$

die andeutet, welches Integral zu welcher Variable gehört. Wir werden uns in diesem Skript aber an die oben erwähnte Mathematik-Schreibweise halten.

Beispiel 2.50.

- a) Integral über einen dreidimensionalen Würfel: Unter Benutzung von $\sin(a + \frac{\pi}{2}) - \sin(a - \frac{\pi}{2}) = 2 \cos a$ erhalten wir

$$\begin{aligned} & \int_{[-\frac{\pi}{2}, \frac{\pi}{2}]^3} \cos(x_1 + x_2 + x_3) d\lambda_3(x_1, x_2, x_3) \\ &= \int_{-\pi/2}^{\pi/2} \int_{-\pi/2}^{\pi/2} \int_{-\pi/2}^{\pi/2} \cos(x_1 + x_2 + x_3) dx_1 dx_2 dx_3 \\ &= \int_{-\pi/2}^{\pi/2} \int_{-\pi/2}^{\pi/2} \left(\sin(x_2 + x_3 + \frac{\pi}{2}) - \sin(x_2 + x_3 - \frac{\pi}{2}) \right) dx_2 dx_3 \\ &= 2 \int_{-\pi/2}^{\pi/2} \int_{-\pi/2}^{\pi/2} \cos(x_2 + x_3) dx_2 dx_3 \\ &= 2 \int_{-\pi/2}^{\pi/2} \left(\sin(x_3 + \frac{\pi}{2}) - \sin(x_3 - \frac{\pi}{2}) \right) dx_3 \\ &= 4 \int_{-\pi/2}^{\pi/2} \cos(x_3) dx_3 \\ &= 4 \left(\sin \frac{\pi}{2} - \sin \left(-\frac{\pi}{2} \right) \right) = 8. \end{aligned}$$

Dieses Integral ließ sich mit Fubini schön vereinfachen, da der Integrationsbereich W ein Produktbereich war, $W = [-\frac{\pi}{2}, \frac{\pi}{2}] \times [-\frac{\pi}{2}, \frac{\pi}{2}] \times [-\frac{\pi}{2}, \frac{\pi}{2}]$. Für allgemeinere Bereiche brauchen wir noch weitere Methoden.

- b) Das folgende Beispiel ist eine Warnung, dass im Satz von Fubini die Integrierbarkeitsannahme an f nicht fallengelassen werden kann. Wir betrachten

$$f : [0, 1] \times [0, 1] \rightarrow \mathbb{R}, \quad f(x, y) := \frac{x^2 - y^2}{(x^2 + y^2)^2}.$$

Wegen $\arctan'(t) = \frac{1}{1+t^2}$ gilt

$$\begin{aligned}\frac{\partial}{\partial y} \arctan \frac{x}{y} &= \frac{-\frac{x}{y^2}}{1 + \frac{x^2}{y^2}} = -\frac{x}{x^2 + y^2}, \\ \frac{\partial^2}{\partial x \partial y} \arctan \frac{x}{y} &= \frac{x^2 - y^2}{(x^2 + y^2)^2} = f(x, y).\end{aligned}$$

Damit erhalten wir

$$\begin{aligned}\int_0^1 \left(\int_0^1 f(x, y) dx \right) dy &= \int_0^1 \left[-\frac{x}{x^2 + y^2} \right]_{x=0}^{x=1} dy = -\int_0^1 \frac{1}{1 + y^2} dy \\ &= -[\arctan y]_0^1 = -\frac{\pi}{4}.\end{aligned}$$

Allerdings gilt auch $\int_0^1 \left(\int_0^1 f(x, y) dy \right) dx = -\frac{\pi}{4}$, da der Integrand bei Vertauschung von x und y sein Vorzeichen wechselt.

Hier sind die beiden iterierten Integrale $\int_0^1 \left(\int_0^1 f(x, y) dy \right) dx$ und $\int_0^1 \left(\int_0^1 f(x, y) dx \right) dy$ also unterschiedlich! Das liegt daran, dass f nicht integrierbar ist (Überprüfen Sie $\int_0^1 \int_0^1 |f(x, y)| dx dy = \infty$).

- c) Wir betrachten die Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) = x \cdot y$ und die Viertelkreisscheibe $V = \{(x, y) \in \mathbb{R}^2 : \|x\|_2 \leq 1, x, y \geq 0\}$. Dann ist f über V integrierbar (d.h. $f \chi_V \in L^1(\mathbb{R}^2)$), denn f ist stetig und V kompakt. Mit Fubini erhalten wir

$$\begin{aligned}\int_V xy d\lambda_2(x, y) &= \int_{\mathbb{R}} \int_{\mathbb{R}} \chi_V(x, y) xy dx dy = \int_0^1 y \int_0^{\sqrt{1-y^2}} x dx dy \\ &= \frac{1}{2} \int_0^1 y(1 - y^2) dy = \frac{1}{8}.\end{aligned}$$

Merkregel: Man merkt sich die Voraussetzungen und Ergebnisse des Satzes von Fubini am besten so: Gegeben eine messbare Funktion $f : \Omega_1 \times \Omega_2 \rightarrow \overline{\mathbb{R}}$, so prüft man, ob eines der drei Integrale

$$\int_{\Omega_1 \times \Omega_2} |f| d(\mu_1 \otimes \mu_2), \quad \int_{\Omega_1} \int_{\Omega_2} |f| d\mu_2 d\mu_1, \quad \int_{\Omega_2} \int_{\Omega_1} |f| d\mu_1 d\mu_2$$

endlich ist. Dann folgt, dass alle drei Integrale gleich sind, und dass die entsprechenden drei Integrale auch gleich sind, wenn überall $|f|$ durch f ersetzt wird.

Beginnen Sie mit $\int_{\Omega_1 \times \Omega_2} |f| d(\mu_1 \otimes \mu_2) < \infty$, so ist dies genau die Aussage des Satzes von Fubini. Beginnen Sie stattdessen mit $\int_{\Omega_1} \int_{\Omega_2} |f| d\mu_2 d\mu_1 < \infty$, so liefert der Satz von Tonelli $\int_{\Omega_1 \times \Omega_2} |f| d(\mu_1 \otimes \mu_2) < \infty$, und die Behauptung folgt wieder mit dem Satz von Fubini.

2.7 Die Transformationsformel

Mit dem Satz von Fubini haben wir ein gutes Werkzeug zum Bestimmen von Integralen in der Hand. Allerdings lässt sich dieser Satz in vielen Situationen nicht direkt oder nur mit Mühe anwenden, wie das folgende Beispiel verdeutlicht.

Betrachten wir Polarkoordinaten, dh beschreiben Punkte $(x, y) \in \mathbb{R}^2$ durch ihren Radius $r := \sqrt{x^2 + y^2}$ und Winkel φ , nämlich $x = r \cos \varphi$, $y = r \sin \varphi$ mit $0 \leq \varphi < 2\pi$. Sei nun $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ eine integrierbare Funktion, und $K \subset \mathbb{R}^2$ der "Kreisstreifen" all der Punkte $(x, y) \in \mathbb{R}^2$, die $R_1 < r < R_2$ und $\alpha_1 \leq \varphi \leq \alpha_2$ erfüllen.

Wollen wir nun das Integral $\int_K f(x, y) dx dy$ bestimmen, so sehen die Integrale über x bzw. y recht kompliziert aus, da wir ausdrücken müssen, für welches $x \in \mathbb{R}$ welche Punkte (x, y) in K liegen. Dabei sieht die Geometrie ziemlich einfach aus, wenn wir in Radius und Winkel rechnen. Wir sollten also statt in x und y besser in r und φ integrieren und müssen das Integral auf diese neuen Variablen "umrechnen". Das geschieht mit der sogenannten Transformationsformel, der mehrdimensionalen Verallgemeinerung der Substitutionsregel aus Analysis 1, die wir in diesem Abschnitt unter Zuhilfenahme von Methoden der Analysis 2 besprechen. Dieses Ergebnis gilt nur für das Lebesgueintegral, wir werden hier also keine allgemeinen Maßräume betrachten.

Wir betrachten zuerst zwei Spezialfälle.

- a) Ist $T \in GL(d)$ eine invertierbare reelle $(d \times d)$ -Matrix und $a \in \mathbb{R}^d$, so betrachten wir die affine invertierbare Abbildung

$$\varphi : \mathbb{R}^d \rightarrow \mathbb{R}^d, \quad \varphi(x) := Tx + a.$$

Aufgrund der Translationsinvarianz des Lebesgue-Borel-Maßes wissen wir aus Satz 1.48, dass das Bildmaß unter φ durch $(\varphi^* \lambda)(V) = \lambda(\varphi^{-1}(V)) = |\det T|^{-1} \lambda(V)$, $V \in \mathcal{B}^d$ gegeben ist. In Hausaufgabe 7.3 haben Sie gezeigt, dass dies für jede Borelmenge $V \subset \mathbb{R}^d$ und jede integrierbare Funktion $f : V \rightarrow \overline{\mathbb{R}}$ die Formel

$$\int_V f(y) dy = \int_{\varphi^{-1}(V)} f(\varphi(x)) |\det T| dx$$

impliziert. (Hier und im Folgenden schreiben wir wieder dx für $d\lambda_d(x)$.)

- b) In der Analysis 1 haben wir die Kettenregel und die zugehörige Substitutionsregel kennengelernt. Zur Erinnerung betrachten wir ein kompaktes Intervall $V = [a, b] \subset \mathbb{R}$, eine stetig differenzierbare streng monotone Funktion $\varphi : V \rightarrow \mathbb{R}$ (also $\varphi'(x) \neq 0$), und eine stetige Funktion $f : [a, b] \rightarrow \mathbb{R}$. Dann gilt

$$\int_V f(y) dy = \int_{\varphi^{-1}(V)} f(\varphi(x)) |\varphi'(x)| dx.$$

Um das einzusehen (in Analysis 1 hatten wir es etwas anders geschrieben) unterscheiden wir die beiden Fälle, dass φ streng monoton wächst oder fällt. Falls φ wächst, so gilt $\varphi'(x) > 0$, dh die Betragsstriche können entfallen, und $\varphi^{-1}([a, b]) = [\varphi^{-1}(a), \varphi^{-1}(b)]$, dh das Integral auf der rechten Seite ist $\int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)}$. Falls φ fällt, so gilt $\varphi'(x) < 0$, und die Betragsstriche geben einen Faktor (-1) . Dieser Faktor wird dadurch kompensiert, dass in diesem Fall $\varphi^{-1}([a, b]) = [\varphi^{-1}(b), \varphi^{-1}(a)]$ und wir für Riemannintegrale $\int_c^d = -\int_d^c$ definiert haben (Im Unterschied zu maßtheoretisch definierten Integralen).

Die allgemeine Transformationsformel enthält die sich formal sehr ähnlich sehenden Spezialfälle a) und b). Um die Verallgemeinerung zu formulieren, betrachten wir einen (C^1) -Diffeomorphismus (Analysis 2, Def. 13.6) $\varphi : U \rightarrow V$ zwischen offenen Teilmengen $U, V \subset \mathbb{R}^d$

(Erinnerung: Das heißt, dass φ bijektiv und stetig differenzierbar ist, und die Umkehrabbildung $\varphi^{-1} : V \rightarrow U$ ebenfalls stetig differenzierbar ist). Dann ist die das Differential $(D\varphi)(x)$ (die Jacobimatrix) invertierbar, d.h. $\det(D\varphi)(x) \neq 0$ für alle $x \in U$ (Analysis 2, Lemma 13.7).

Satz 2.51 (Transformationsformel). Seien $U, V \subset \mathbb{R}^d$ offen, $\varphi : U \rightarrow V$ ein C^1 -Diffeomorphismus, und $f : V \rightarrow \overline{\mathbb{R}}$ messbar. Dann ist f über V integrierbar genau dann, wenn $|\det(D\varphi)| \cdot (f \circ \varphi)$ über $U = \varphi^{-1}(V)$ integrierbar ist. In diesem Fall gilt

$$\int_V f(y) dy = \int_{\varphi^{-1}(V)} f(\varphi(x)) |\det(D\varphi)(x)| dx.$$

Da wir den Beweis etwas später machen, führen wir hier zunächst nur eine kleine Plausibilitätsüberlegung durch und denken uns dazu $\varphi^{-1}(V)$ in viele kleine Quader Q zerlegt. Da φ stetig differenzierbar ist, ist es auf Q affin approximierbar (Analysis 2, Lemma 12.8), dh $\varphi(x) \approx (D\varphi)(x_0)(x - x_0) + \varphi(x_0)$, wobei $x_0 \in Q$ ein Referenzpunkt ist. Auf jedem Quader können wir also φ durch eine affine Transformation wie in Spezialfall a) ersetzen. Durch Zusammensetzen aller Quader sollte sich dann die behauptete Transformationsformel ergeben.

Zunächst geben wir einige Anwendungsbeispiele der Transformationsformel.

Beispiel 2.52 (Polarkoordinaten). Jeder Punkt $(x, y) \in \mathbb{R}^2$ lässt sich auch durch Radius (Abstand vom Ursprung in Euklidischer Norm) und Winkel $\theta \in (-\pi, \pi]$ beschreiben:

$$x = r \cos \theta, \quad y = r \sin \theta.$$

Für $x = y = 0$ (also $r = 0$) ist der Winkel nicht eindeutig festgelegt. Weiterhin ist der Winkelbereich $(-\pi, \pi]$ (den wir so gewählt haben, damit jeder Punkt mit $r > 0$ einen eindeutigen Winkel hat) nicht offen. Für die Konstruktion eines Diffeomorphismus müssen wir diese störenden Punkte entfernen, dh wir betrachten $r > 0$ und $\theta \in (-\pi, \pi)$, nämlich

$$U := \mathbb{R}_+ \times (-\pi, \pi), \quad M := \{(x, 0) : x \leq 0\}, \\ \varphi : U \rightarrow V := \mathbb{R}^2 \setminus M, \quad \varphi(r, \theta) := (r \cos \theta, r \sin \theta).$$

Dann sind U und V offen, und φ ist bijektiv und stetig differenzierbar. Ihre Jacobimatrix ist

$$(D\varphi)(r, \theta) = \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix} \Rightarrow \det(D\varphi)(r, \theta) = r > 0.$$

Dies zeigt, dass $(D\varphi)(r, \theta)$ für alle $(r, \theta) \in U$ invertierbar ist. Nach dem Satz über die Umkehrabbildung (Analysis 2, Korollar 13.14) folgt, dass φ ein Diffeomorphismus ist.

Da M eine Borel-Nullmenge ist, besagt die Transformationsformel für jede integrierbare

Funktion $f : \mathbb{R}^2 \rightarrow \overline{\mathbb{R}}$

$$\int_{\mathbb{R}^2} f(x, y) dx dy = \int_{-\pi}^{\pi} \int_0^{\infty} f(r \cos \theta, r \sin \theta) r dr d\theta.$$

Als konkretes Anwendungsbeispiel berechnen wir das Integral $I := \int_{\mathbb{R}} e^{-x^2} dx$ mit einem Trick:

$$\begin{aligned} I^2 &= \int_{\mathbb{R}} e^{-x^2} dx \cdot \int_{\mathbb{R}} e^{-y^2} dy = \int_{\mathbb{R}} \int_{\mathbb{R}} e^{-x^2-y^2} dx dy = \int_{-\pi}^{\pi} \int_0^{\infty} e^{-r^2} r dr d\theta \\ &= -\frac{1}{2} \int_{-\pi}^{\pi} \int_0^{\infty} \frac{de^{-r^2}}{dr} dr d\theta = \frac{1}{2} \int_{-\pi}^{\pi} d\theta = \pi, \end{aligned}$$

also $I = \sqrt{\pi}$. Beachten Sie, dass $x \mapsto e^{-x^2}$ keine elementare Stammfunktion hat, das Quadrieren des Integrals per Fubini und Polarkoordinaten aber auf die Funktion $r \mapsto re^{-r^2}$ mit einer elementaren Stammfunktion geführt hat.

An anderes noch einfacheres Beispiel ist, die Fläche der Kreisscheibe $K_R = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < R^2\}$ mit Radius $R > 0$ zu berechnen. Wegen $\varphi^{-1}(K_R \setminus \{(x, 0) : x < 0\}) = (0, R) \times (-\pi, \pi)$ gilt

$$\lambda(K_R) = \int_{\mathbb{R}^2} \chi_{K_R}(x, y) dx dy = \int_{-\pi}^{\pi} \int_0^R r dr d\theta = \pi R^2.$$

Beispiel 2.53 (Kugelkoordinaten). Als nächstes Beispiel betrachten wir Kugelkoordinaten in drei Dimensionen. Hier beschreiben wir Punkte $(x, y, z) \in \mathbb{R}^3$ durch ihren Abstand vom Ursprung $r = \sqrt{x^2 + y^2 + z^2}$, den Winkel $\alpha \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, den (x, y, z) mit der $\{z = 0\}$ -Ebene einschließt (Breitengrad), und dem Winkel $\beta \in (-\pi, \pi)$, den $(x, y, 0)$ mit der positiven x -Achse einschließt (Längengrad). Nach elementargeometrischen Überlegungen gilt

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r \cos \alpha \cos \beta \\ r \cos \alpha \sin \beta \\ r \sin \alpha \end{pmatrix} =: \varphi(r, \alpha, \beta)$$

Dann ist $\varphi : \mathbb{R}_+ \times (-\frac{\pi}{2}, \frac{\pi}{2}) \times (-\pi, \pi) \rightarrow \mathbb{R}^3 \setminus N$ (mit einer Nullmenge N) bijektiv und stetig differenzierbar. Die Jacobimatrix ist

$$\begin{aligned} (D\varphi)(r, \alpha, \beta) &= \begin{pmatrix} \cos \alpha \cos \beta & -r \sin \alpha \cos \beta & -r \cos \alpha \sin \beta \\ \cos \alpha \sin \beta & -r \sin \alpha \sin \beta & r \cos \alpha \cos \beta \\ \sin \alpha & r \cos \alpha & 0 \end{pmatrix} \\ \Rightarrow \det(D\varphi)(r, \alpha, \beta) &= r^2 \cos \alpha > 0. \end{aligned}$$

Wie im vorigen Beispiel schließen wir, dass φ ein Diffeomorphismus ist. Für integrier-

bare Funktionen $f : \mathbb{R}^3 \rightarrow \overline{\mathbb{R}}$ gilt also

$$\begin{aligned} \int_{\mathbb{R}^3} f \, d\lambda_3 &= \int_{\mathbb{R}^3} f(x, y, z) \, dx \, dy \, dz \\ &= \int_{-\pi}^{\pi} \int_{-\pi/2}^{\pi/2} \int_0^{\infty} f(r \cos \alpha \cos \beta, r \cos \alpha \sin \beta, r \sin \alpha) r^2 \cos \alpha \, dr \, d\alpha \, d\beta. \end{aligned}$$

Als konkretes Anwendungsbeispiel berechnen wir das Gravitationspotential einer rotations-symmetrischen Masseverteilung. Wir betrachten also die Kugel $K_R = \{\mathbf{x} \in \mathbb{R}^3 : \|\mathbf{x}\| \leq R\}$ und eine integrierbare positive Funktion $m : [0, R] \rightarrow \mathbb{R}_+$. Die Gesamtmasse M der Masseverteilung ist das Integral über die Dichte, also

$$M := \int_{K_R} m(\|\mathbf{x}\|) \, d\mathbf{x} = \int_{-\pi}^{\pi} \int_{-\pi/2}^{\pi/2} \int_0^R m(r) r^2 \cos \alpha \, dr \, d\alpha \, d\beta = 4\pi \int_0^R m(r) r^2 \, dr.$$

Das Gravitationspotential dieser Masseverteilung ist (bis auf Naturkonstanten) im Punkt $\mathbf{y} \in \mathbb{R}^3 \setminus K_R$

$$V(\mathbf{y}) = \int_{K_R} \frac{m(\|\mathbf{x}\|)}{\|\mathbf{x} - \mathbf{y}\|} \, d\mathbf{x}.$$

Um dieses Integral zu berechnen, behaupten wir zunächst, dass $V(\mathbf{y})$ nur von $\|\mathbf{y}\|$ abhängt. Sei dazu $T \in \text{SO}(3)$ eine Rotationsmatrix. Dann gilt wegen $\|T\mathbf{x}\| = \|\mathbf{x}\|$

$$V(T\mathbf{y}) = \int_{K_R} \frac{m(\|\mathbf{x}\|)}{\|\mathbf{x} - T\mathbf{y}\|} \, d\mathbf{x} = \int_{K_R} \frac{m(\|T^{-1}\mathbf{x}\|)}{\|T^{-1}\mathbf{x} - \mathbf{y}\|} \, d\mathbf{x} = \int_{K_R} \frac{m(\|\mathbf{x}\|)}{\|\mathbf{x} - \mathbf{y}\|} \, d\mathbf{x} = V(\mathbf{y}).$$

Im vorletzten Schritt haben wir die Transformationsformel für den linearen Diffeomorphismus $\varphi(\mathbf{x}) = T^{-1}\mathbf{x}$ (mit $\det T = 1$) und $T(K_R) = K_R$ (eine Kugel ist rotations-symmetrisch) benutzt.

Zum Berechnen von $V(\mathbf{y})$ können wir uns also auf $\mathbf{y} = (0, 0, z)$, $z = \|\mathbf{y}\| > R$ beschränken. Dazu berechnen wir zunächst die im Integranden auftretende Norm

$$\|\mathbf{x} - (0, 0, z)\|^2 = \left\| \begin{pmatrix} r \cos \alpha \cos \beta \\ r \cos \alpha \sin \beta \\ r \sin \alpha - z \end{pmatrix} \right\|^2 = r^2 + z^2 - 2rz \sin \alpha.$$

Unser Integral ist dann in Kugelkoordinaten (mit der Substitution $t = \sin \alpha$)

$$\begin{aligned}
 V(\mathbf{y}) &= \int_{-\pi}^{\pi} \int_{-\pi/2}^{\pi/2} \int_0^R \frac{m(r) r^2 \cos \alpha}{\sqrt{r^2 + z^2 - 2rz \sin \alpha}} dr d\alpha d\beta \\
 &= 2\pi \int_{-1}^1 \int_0^R \frac{m(r) r^2}{\sqrt{r^2 + z^2 - 2rz t}} dr dt \\
 &= 2\pi \int_0^R m(r) r^2 \left[\frac{1}{-rz} \sqrt{r^2 + z^2 - 2rz t} \right]_{t=-1}^{t=1} dr \\
 &= 2\pi \int_0^R m(r) \frac{r}{z} \left[\sqrt{(r+z)^2} - \sqrt{(r-z)^2} \right] dr \\
 &= \frac{4\pi}{z} \int_0^R m(r) r^2 dr \\
 &= \frac{M}{\|\mathbf{y}\|}.
 \end{aligned}$$

Beim Auflösen der Wurzeln haben wir benutzt, dass \mathbf{y} nicht in K_R liegt, also $z = \|\mathbf{y}\| > R \geq r$ für alle r im Integrationsbereich.

Die Interpretation dieses Ergebnisses ist, dass das Potential der Kugel mit dem Potential einer in $\mathbf{x} = 0$ konzentrierten Punktmasse M übereinstimmt. Diese Tatsache ist auch als *Newton's Theorem* bekannt.

Nachdem wir uns von der Nützlichkeit der Transformationsformel überzeugt haben, machen wir uns an den Beweis.

Beweis der Transformationsformel. Wir bemerken zunächst, dass aufgrund der Stetigkeit und Invertierbarkeit von φ und wegen $|\det(D\varphi)(x)| \neq 0$ die Messbarkeit von f äquivalent ist zu der Messbarkeit von $(f \circ \varphi) \cdot |\det D\varphi|$. Wir zerlegen dann in Positiv- und Negativteil $f = f_+ - f_-$ bzw. $(f \circ \varphi) \cdot |\det D\varphi| = (f_+ \circ \varphi) \cdot |\det D\varphi| - (f_- \circ \varphi) \cdot |\det D\varphi|$, dh wir müssen die behauptete Formel nur für $f \geq 0$ zeigen.

Angenommen, wir können die Ungleichung

$$\int_V f d\lambda \leq \int_U (f \circ \varphi) |\det D\varphi| d\lambda \quad (\star)$$

zeigen. Dann gilt mit $g := (f \circ \varphi) |\det D\varphi|$ anstelle von f und $\psi := \varphi^{-1}$ anstelle von φ und

U anstelle von V

$$\begin{aligned} \int_U (f \circ \varphi) |\det D\varphi| d\lambda &= \int_U g d\lambda \\ &\leq \int_V (f \circ \varphi \circ \varphi^{-1})(x) |\det(D\varphi)(\varphi^{-1}(x))| |\det(D\varphi^{-1})(x)| d\lambda(x) \\ &= \int_V f d\lambda, \end{aligned}$$

also die umgekehrte Ungleichung. Es genügt also, (\star) zu zeigen. (Wir haben hier die Formel für das Differential der Umkehrabbildung verwendet, siehe Analysis 2, Lemma 13.7.)

Mit den mittlerweile gewohnten Approximationsargumenten genügt es, (\star) für charakteristische Funktionen $f = \chi_B$ von Borelmengen $B \subset V$ zu zeigen, dh mit $A := \varphi^{-1}(B) \subset U$

$$\lambda(\varphi(A)) \leq \int_A |\det D\varphi| d\lambda. \quad (\star\star)$$

Nun müssen wir aber wirklich etwas beweisen und zeigen zunächst, dass $(\star\star)$ für kompakte (nicht leere) Würfel A gilt.

Dazu bezeichnen wir mit $l > 0$ die Seitenlänge des Würfels A und wählen $c \geq 0$ so, dass

$$\lambda(\varphi(A)) = c \int_A |\det D\varphi| d\lambda$$

gilt. Das ist möglich, da das Integral ungleich Null ist. Ziel des Beweises ist also $c \leq 1$.

Zerlegen wir nun A in 2^d kompakte bis auf Nullmengen disjunkte Würfel, indem wir alle Seiten von A halbieren, so muss für mindestens einen von den kleineren Würfeln, sagen wir A_1 , die Ungleichung $\lambda(\varphi(A_1)) \geq c \int_{A_1} |\det D\varphi| d\lambda$ gelten, den anderenfalls gibt es einen Widerspruch zu $\lambda(\varphi(A)) = c \int_A |\det D\varphi| d\lambda$ wegen der Additivität beider Seiten der Gleichung unter disjunkter Zerlegung.

Wir iterieren dieses Verfahren und erhalten so eine absteigende Folge (A_n) kompakter Würfel mit Seitenlänge $2^{-n}l$ und $\lambda(\varphi(A_n)) \geq c \int_{A_n} |\det D\varphi| d\lambda$. Die Mittelpunkte der Würfel bilden eine Cauchyfolge mit Grenzwert a , der $a \in \bigcap_n A_n$ erfüllt. Indem wir eine affine Transformation ausführen, für die wir die Transformationsformel bereits kennen, dürfen wir $a = 0 = \varphi(a)$ annehmen.

Nun benutzen wir die Differenzierbarkeit von φ in $a = 0$: Mit $T := (D\varphi)(0)$ gilt

$$\varphi(x) = Tx + r(x), \quad \lim_{x \rightarrow 0} \frac{r(x)}{\|x\|_\infty} = 0.$$

Wir arbeiten hier mit der (an Würfel angepassten) Maximumsnorm $\|\cdot\|_\infty$ auf \mathbb{R}^d . Zu jedem $\varepsilon > 0$ gibt es $N_\varepsilon \in \mathbb{N}$ mit $\|\varphi(x) - Tx\|_\infty < \varepsilon l_n$ für alle $x \in A_n$. Also liegt $\varphi(A_n)$ in $TA_n + l_n B_\varepsilon$ mit $B_\varepsilon = \{y \in \mathbb{R}^d : \|y\|_\infty < \varepsilon\}$. Der skalierte Würfel A_n/l_n hat Seitenlänge 1, ist also von der Form $A_n/l_n = [0, 1]^d + b_n$ mit $b_n \in [-1, 0]^d$. Nach dem Satz von Stone-Weierstraß hat (b_n) eine konvergente Teilfolge; wir dürfen also die Existenz von $b := \lim_n b_n$ annehmen. Die Stetigkeit von T impliziert dann für hinreichend großes n

$$TA_n/l_n + B_\varepsilon = T([0, 1]^d + b_n) + B_\varepsilon \subset T[0, 1]^d + Tb + B_{2\varepsilon}.$$

Nun erhalten wir (beachten Sie $l_n^d = \lambda(A_n)$)

$$\begin{aligned} c \int_{A_n} |\det D\varphi| d\lambda &\leq \lambda(\varphi(A_n)) \leq \lambda(TA_n + l_n B_\varepsilon) \leq l_n \lambda(T[0, 1]^d + B_{2\varepsilon}) \\ &\Rightarrow c \frac{\int_{A_n} |\det D\varphi| d\lambda}{\lambda(A_n)} \leq \lambda(T[0, 1]^d + B_{2\varepsilon}). \end{aligned}$$

Im Limes $n \rightarrow \infty$ konvergiert die linke Seite aufgrund der Stetigkeit von $|\det D\varphi|$ gegen $c|\det(D\varphi)(0)|$, dh $c|\det(D\varphi)(0)| \leq \lambda(T[0, 1]^d + B_{2\varepsilon})$. Im Limes $\varepsilon \rightarrow 0$ konvergiert nun die rechte Seite gegen $\lambda(T[0, 1]^d) = |\det T| = |\det(D\varphi)(0)|$. Dies ergibt $c \leq 1$ wie behauptet.

Im letzten Schritt des Beweises müssen wir die Ungleichung $(\star\star)$ von kompakten Würfeln auf messbare Mengen verallgemeinern. Zuerst bemerken wir, dass $(\star\star)$ auch für halboffene statt kompakte Würfel gilt, da ein halboffener Würfel $[a, b) \subset V$ von innen durch kompakte Würfel approximiert werden kann, $[a, b_n] \nearrow [a, b)$, und beide Seiten von $(\star\star)$ unter dieser Approximation monoton wachsend konvergieren.

Da jede offene Menge eine disjunkte Vereinigung von abzählbar vielen halboffenen Würfeln ist (Lemma 1.10), folgt die Gültigkeit von $(\star\star)$ für alle offenen Mengen $A \subset V$.

Sei nun $A \subset U$ eine Borelmenge. Dann ist $U_n := \{|\det D\varphi| < n\} \cap U$ offen, und $U_n \nearrow U$ für $n \rightarrow \infty$. Also $A \cap U_n \nearrow A$. Es genügt also, den Beweis für die Mengen A_n zu führen, dh wir dürfen annehmen, dass $|\det D\varphi|$ beschränkt ist.

Sei nun $\varepsilon > 0$. Dann gibt es eine offene Menge $O \supset A$ mit $\lambda(O \setminus A) \leq \varepsilon$. Betrachten wir $U \cap O$ anstelle von O , dürfen wir $O \subset U$ annehmen. Dann erhalten wir mit dem bereits gezeigten Ergebnis für offene Mengen

$$\begin{aligned} \lambda(\varphi(A)) &\leq \lambda(\varphi(O)) \leq \int_O |\det D\varphi| d\lambda \\ &= \int_A |\det D\varphi| d\lambda + \int_{O \setminus A} |\det D\varphi| d\lambda \leq \int_A |\det D\varphi| d\lambda + \varepsilon \sup_{x \in U} |\det(D\varphi)(x)| \end{aligned}$$

Im Limes $\varepsilon \rightarrow 0$ ergibt sich dann die Behauptung. \square

An dieser Stelle haben wir die wesentlichen Züge der Maß- und Integrationstheorie kennengelernt. Im nächsten Kapitel werden wir Anwendungen in der Fourieranalysis besprechen.

3 Einführung in die Fourier-Analyse

Als Motivation für die folgenden Überlegungen denken wir an ein akustisches Signal

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad t \mapsto f(t),$$

in Abhängigkeit einer Zeitvariablen t , das wir auf seine Frequenzen hin untersuchen wollen. Falls $f(t) = \sin(kt)$ mit $k \in \mathbb{R}$, so ist f periodisch mit Periodenlänge $\frac{2\pi}{k}$, dh Frequenz k . Im Allgemeinen wird ein Signal viele Frequenzen enthalten, zB enthält $f(t) = \sum_{n=1}^N c_n \sin(k_n t)$ die Frequenzen k_1, \dots, k_N mit Amplituden c_1, \dots, c_N . Im Fall von kontinuierlich variierenden Frequenzen denken wir an Funktionen der Form $f(t) = \int_{\mathbb{R}} c(k) \sin(kt) dk$. Welche Funktionen ergeben sich auf diese Art und Weise? Gegeben ein solches f , wie kann man c (Frequenzinformation) zurückgewinnen?

Diese Fragen stehen am historischen Beginn der Fourier-Analyse. Heutzutage ist Fourier-Analyse ein weites Feld, das sowohl in reiner Mathematik (harmonische Analysis) als auch den Anwendungen (Physik, Datenkompression) eine wichtige Rolle spielt.

Aus der Perspektive von Kapitel 2 dieses Skripts wird uns die Fourier-Analyse die Möglichkeit geben, alle besprochenen Ergebnisse der Integrationstheorie anzuwenden. Alle Integrale werden bzgl. des Lebesguemaßes sein.

3.1 Definition der Fouriertransformation und erste Eigenschaften

Bisher haben wir uns mit (messbaren bzw integrierbaren) Funktionen $f : \Omega \rightarrow \mathbb{R}$ beschäftigt. Im Folgenden wird es sich als günstig erweisen, statt der trigonometrischen Funktionen die komplexe Exponentialfunktion $e^{ikx} = \cos(kx) + i \sin(kx)$ zu verwenden. Wir werden deshalb komplexwertige Funktionen $f : \Omega \rightarrow \mathbb{C}$ betrachten. Das Integral über eine komplexwertige Funktion ist als

$$\int_{\Omega} f d\mu = \int_{\Omega} \operatorname{Re}(f) d\mu + i \int_{\Omega} \operatorname{Im}(f) d\mu$$

definiert und damit komplex linear. Wir werden ab nun mit $\mathcal{L}^p(\Omega)$ die Menge aller komplexwertigen Funktionen $f : \Omega \rightarrow \mathbb{C}$ meinen, die $\int_{\Omega} |f|^p d\mu < \infty$ erfüllen, und analog für die $L^p(\Omega)$.

Definition 3.1. Für $f \in \mathcal{L}^1(\mathbb{R})$ definieren wir die *Fourier-Transformierte von f* als die Funktion

$$\tilde{f} : \mathbb{R} \rightarrow \mathbb{C}, \quad \tilde{f}(k) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x) e^{-ikx} dx.$$

Allgemeiner definieren wir für $f \in \mathcal{L}^1(\mathbb{R}^d)$, $d \in \mathbb{N}$, die *Fourier-Transformierte von f* als die Funktion

$$\tilde{f} : \mathbb{R}^d \rightarrow \mathbb{C}, \quad \tilde{f}(k) := \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} f(x) e^{-i\langle k, x \rangle} dx,$$

wobei $\langle k, x \rangle = k_1 x_1 + \dots + k_d x_d$ das Euklidische Skalarprodukt auf \mathbb{R}^d ist.

- Der Vorfaktor $(2\pi)^{-d/2}$ ist Konvention. Wir werden später sehen, wieso dieser Vorfaktor eine gute Wahl ist.
- Da das Integral $\int_{\mathbb{R}} f(x)e^{-ikx} dx$ für jedes $k \in \mathbb{R}$ definiert ist, können wir es als Funktion von k betrachten – dies ist genau die Definition der Fouriertransformierten von f . Weiterhin sehen wir, dass eine Änderung von f auf Nullmengen zu der gleichen Fouriertransformierten \tilde{f} führt, dh wir können auch (Äquivalenzklassen) $f \in L^1(\mathbb{R}^d)$ eine Fouriertransformierte zuordnen.
- Der Faktor $e^{-ikx} = \cos(kx) - i \sin(kx)$ begründet sich in der Motivation als “Frequenzanalyse”. In Dimension $d > 1$ verwenden wir den Faktor

$$e^{-i\langle k, x \rangle} = e^{-i(k_1 x_1 + \dots + k_d x_d)} = e^{-ik_1 x_1} \dots e^{-ik_d x_d}.$$

Ist $f \in L^1(\mathbb{R}^d)$ von Produktform, nämlich $f(x_1, \dots, x_d) = f_1(x_1) \cdots f_d(x_d)$ mit $f_1, \dots, f_d \in L^1(\mathbb{R})$, so folgt mit dem Satz von Fubini

$$\tilde{f}(k) = \tilde{f}_1(k_1) \cdots \tilde{f}_d(k_d),$$

was die Wahl von $e^{-i\langle k, x \rangle}$ in der Definition der Fouriertransformation begründet.

Zur Illustration berechnen wir einige Fouriertransformierte in einer Dimension.

Beispiel 3.2. Sei $I = [a, b] \subset \mathbb{R}$ ein kompaktes Intervall. Dann ist $f := \chi_I \in L^1(\mathbb{R})$, und

$$\tilde{f}(k) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-ikx} dx = \frac{1}{\sqrt{2\pi}} \frac{i}{k} (e^{-ikb} - e^{-ika}).$$

Für $\alpha > 0$ ist $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = e^{-\alpha|x|}$, in $L^1(\mathbb{R})$, und

$$\begin{aligned} \tilde{f}(k) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-\alpha|x|} e^{-ikx} dx = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} e^{-(\alpha+ik)x} dx + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 e^{(\alpha-ik)x} dx \\ &= \frac{1}{\sqrt{2\pi}} \left(\frac{1}{\alpha+ik} + \frac{1}{\alpha-ik} \right) \\ &= \sqrt{\frac{2}{\pi}} \frac{\alpha}{\alpha^2 + k^2}. \end{aligned}$$

Definition 3.1 legt zu jeder Funktion $f \in L^1(\mathbb{R}^d)$ eine neue Funktion \tilde{f} fest. Im Folgenden wollen wir verstehen, wie Eigenschaften von f mit Eigenschaften von \tilde{f} zusammenhängen. Wenn wir die *Fouriertransformation* betrachten, also die auf $L^1(\mathbb{R}^d)$ definierte Abbildung $\mathcal{F} : f \mapsto \tilde{f}$, so klären wir zuerst die Frage, was der Bildbereich von \mathcal{F} ist.

Lemma 3.3. Für $f \in L^1(\mathbb{R}^d)$ ist \tilde{f} stetig und beschränkt.

Beweis. Für die Stetigkeit verwenden wir Satz 2.27: Wegen $|f(x)e^{-i\langle k, x \rangle}| = |f(x)|$ haben wir die integrierbare Majorante $|f|$, so dass die Stetigkeit von \tilde{f} folgt. Die Beschränktheit erhalten wir per Dreiecksungleichung,

$$|\tilde{f}(k)| \leq \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}} |f(x)e^{-i\langle k, x \rangle}| dx = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}} |f(x)| dx = \frac{\|f\|_1}{(2\pi)^{d/2}} < \infty. \quad \square$$

Wir können die Fouriertransformation also als eine Abbildung

$$\mathcal{F} : L^1(\mathbb{R}^d) \rightarrow C_b(\mathbb{R}^d), \quad f \mapsto \tilde{f}.$$

auffassen, wobei $C_b(\mathbb{R}^d)$ den Vektorraum aller stetigen beschränkten Funktionen $\mathbb{R}^d \rightarrow \mathbb{C}$ bezeichnet. Diese Abbildung ist aufgrund der Linearität des Integrals linear. Natürliche Fragen sind: Ist \mathcal{F} injektiv? surjektiv? bijektiv? Kann man ggf. die inverse Abbildung \mathcal{F}^{-1} bestimmen?

Die beiden Beispiele oben zeigen, dass die Fouriertransformierte \tilde{f} üblicherweise völlig anders aussieht als die Funktion f selbst. Der Zusammenhang zwischen f und \tilde{f} ist also trickreich.

Wir betrachten zunächst den Zusammenhang zwischen Differentiation und Fouriertransformation, was im Wesentlichen auf der einfachen Beziehung

$$\frac{\partial e^{-i\langle k, x \rangle}}{\partial k_\mu} = -ix_\mu e^{-i\langle k, x \rangle}, \quad \mu = 1, \dots, d,$$

beruht. Wir werden es also mit Multiplikation mit Komponenten x_μ von $x \in \mathbb{R}^d$ zu tun haben, und definieren deshalb

$$Q_\mu(f) := x_\mu \cdot f(x),$$

wobei $f : \mathbb{R}^d \rightarrow \mathbb{C}$ eine beliebige Funktion ist. In Dimension $d = 1$ gibt es nur eine Komponente und wir schreiben einfach Q für die Abbildung $(Qf)(x) := xf(x)$. Wir können die Q_μ auch mehrfach anwenden und schreiben dann $Q_\mu^2(f) := Q_\mu(Q_\mu(f))$ für die Funktion $x \mapsto x_\mu^2 f(x)$, und analog für höhere Potenzen Q_μ^n , $n \in \mathbb{N}$.

Satz 3.4. Für beliebiges $d \in \mathbb{N}$ und $\mu \in \{1, \dots, d\}$ gilt

a) Sei $f \in L^1(\mathbb{R}^d)$ so, dass $Q_\mu(f) \in L^1(\mathbb{R}^d)$. Dann ist \tilde{f} partiell nach seiner μ -ten Variablen differenzierbar, und es gilt

$$\widetilde{Q_\mu(f)} = i\partial_\mu \tilde{f}.$$

b) Sei $f \in L^1(\mathbb{R}^d) \cap C^1(\mathbb{R}^d)$ mit $\partial_\mu f \in L^1(\mathbb{R}^d)$ und $f(x) \rightarrow 0$ für $\|x\| \rightarrow \infty$. Dann gilt

$$\widetilde{\partial_\mu f} = iQ_\mu(\tilde{f}).$$

Bei diesen Formeln muss man genau auf die Reihenfolge der Operationen achten: In a) wird auf der linken Seite erst mit der μ -ten Komponente der Variablen multipliziert, und dann Fourier-transformiert. Auf der rechten Seite der Gleichung wird hingegen zuerst Fourier-transformiert und dann nach der μ -ten Variable differenziert:

$$\begin{aligned} \widetilde{Q_\mu(f)}(k) &= (2\pi)^{-d/2} \int_{\mathbb{R}^d} (-ix_\mu f(x) e^{-i\langle k, x \rangle}) dx \\ &= (2\pi)^{-d/2} \frac{\partial}{\partial k_\mu} \int_{\mathbb{R}^d} f(x) e^{-i\langle k, x \rangle} dx = \frac{\partial \tilde{f}(k)}{\partial k_\mu}. \end{aligned}$$

Analoges gilt für b).

Beweis. a) Wir benutzen Satz 2.29 (Differenzierbarkeit parameterabhängiger Integrale). Für jedes feste $x \in \mathbb{R}^d$ ist $k \mapsto f(x)e^{-i\langle k, x \rangle}$ beliebig oft differenzierbar; insbesondere existieren die partiellen Ableitungen $\frac{\partial \tilde{f}(k)}{\partial k_\mu}$, $\mu = 1, \dots, d$. Die partielle Ableitung des Integranden ist

$$\frac{\partial}{\partial k_\mu} (f(x)e^{-i\langle k, x \rangle}) = -ix_\mu \cdot f(x) e^{-i\langle k, x \rangle}.$$

Um Satz 2.29 anwenden zu können, brauchen wir eine integrierbare Majorante M dieser partiellen Ableitung, dh $M \in L^1(\mathbb{R}^d)$ mit $|-ix_\mu \cdot f(x) e^{-i\langle k, x \rangle}| = |x_\mu f(x)| \leq M(x)$. Da $x \mapsto x_\mu f(x)$ per Annahme integrierbar ist, haben wir unsere Majorante gefunden.

b) Nach Annahme sind sowohl f als auch $\partial_\mu f$ integrierbar. Wir schreiben die linke Seite der behaupteten Gleichung mit dem Satz von Fubini als

$$\widetilde{\partial_\mu f}(k) = (2\pi)^{-d/2} \int_{\mathbb{R}} \dots \int_{\mathbb{R}} e^{-i\langle k, x \rangle} \frac{\partial f(x)}{\partial x_\mu} dx_1 \dots dx_d.$$

Da wir uns die Reihenfolge der Integrationen aussuchen dürfen, führen wir zuerst die Integration über x_μ aus. Das x_μ -Integral ist somit (partielle Integration)

$$\begin{aligned} \lim_{\substack{a \rightarrow -\infty \\ b \rightarrow \infty}} \int_a^b e^{-i\langle k, x \rangle} \frac{\partial f(x)}{\partial x_\mu} dx_\mu &= - \lim_{\substack{a \rightarrow -\infty \\ b \rightarrow \infty}} \int_a^b f(x) \frac{\partial e^{-i\langle k, x \rangle}}{\partial x_\mu} dx_\mu + \lim_{\substack{a \rightarrow -\infty \\ b \rightarrow \infty}} [e^{-i\langle k, x \rangle} f(x)]_{x_\mu=a}^{x_\mu=b} \\ &= ik_\mu \int_{\mathbb{R}} f(x) e^{-i\langle k, x \rangle} dx_\mu + 0. \end{aligned}$$

Führen wir nun auch die anderen Integrationen aus, so erhalten wir das behauptete Ergebnis

$$\widetilde{\partial_\mu f}(k) = ik_\mu \tilde{f}(k).$$

□

Dieser Satz sagt, dass Fouriertransformation (partielle) Ableitungen in Multiplikation mit Komponenten des Arguments umwandelt. Diese Eigenschaft ist grundlegend in vielen Anwendungen der Fouriertransformation.

Beispiel 3.5.

- a) Wir hatten bereits geprüft, dass die Fouriertransformierte von $f(x) = e^{-|x|}$ die Funktion $\tilde{f}(k) = \sqrt{\frac{2}{\pi}} \frac{1}{1+k^2}$ ist. Was ist die Fouriertransformation von $g(x) := xe^{-|x|}$? Nach Teil a) des Satzes gilt

$$\tilde{g}(k) = i \frac{d}{dk} \sqrt{\frac{2}{\pi}} \frac{1}{1+k^2} = \sqrt{\frac{2}{\pi}} \frac{-2ik}{(1+k^2)^2}.$$

- b) Was ist die Fouriertransformierte von $f(x) := e^{-\frac{1}{2}x^2}$? Wir berechnen Sie mit einem Trick. Dazu bemerken wir zuerst, dass

$$f'(x) = -xf(x)$$

gilt. Da $f \in L^1(\mathbb{R})$ und die Voraussetzungen von Satz 3.4 b) erfüllt sind, können

wir beide Seiten dieser Gleichung Fourier-transformieren, und erhalten

$$ik\tilde{f}(k) = -i\frac{d\tilde{f}(k)}{dk} \Rightarrow \frac{d\tilde{f}(k)}{dk} = -k\tilde{f}(k).$$

Die Fouriertransformierte \tilde{f} erfüllt also die selbe (Differential-)Gleichung wie f . Ist nun h eine beliebige differenzierbare Funktion $\mathbb{R} \rightarrow \mathbb{C}$, die $h'(x) = -xh(x)$ für alle $x \in \mathbb{R}$ erfüllt, so ist $H(x) := h(x)e^{\frac{1}{2}x^2}$ konstant (die Ableitung ist Null, wie man leicht nachrechnet). Daraus schließen wir, dass $h(x) = ce^{-\frac{1}{2}x^2}$ mit einer Konstante $c \in \mathbb{C}$. Wir haben also schon gelernt:

$$\tilde{f}(k) = ce^{-\frac{1}{2}k^2}.$$

Um auch noch die Konstante c zu bestimmen, bemerken wir

$$c = \tilde{f}(0) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-\frac{1}{2}x^2} dx = \frac{1}{\sqrt{\pi}} \int_{\mathbb{R}} e^{-y^2} dy = 1,$$

wobei wir Beispiel 2.52 verwendet haben. Die Gaussfunktion f stimmt also mit ihrer Fouriertransformierten überein.

Die Fouriertransformation übersetzt lokale Eigenschaften von f wie Stetigkeit, Differenzierbarkeit, in globale Eigenschaften von \tilde{f} , wie Abfallgeschwindigkeit bei unendlich. In Beispiel 3.2 hatten wir eine unstetige Funktion f betrachtet, ihre Fouriertransformierte \tilde{f} fiel nur langsam ab, nämlich wie $|k|^{-1}$. Im zweiten Teils des Beispiels hatten wir eine stetige Funktion betrachtet, deren Fouriertransformierte fiel etwas schneller ab (wie $|k|^{-2}$). Im Allgemeinen haben wir den folgenden Zusammenhang. Der Einfachheit halber beschränken wir uns auf eine Dimension, $d = 1$.

Korollar 3.6. Falls $f \in L^1(\mathbb{R})$ n -mal stetig differenzierbar mit Ableitungen $f^{(l)} \in L^1(\mathbb{R})$ ist, und $f^{(l)}(x) \rightarrow 0$ für $|x| \rightarrow \infty$ für alle $l = 0, 1, \dots, n-1$ gilt, so gibt es eine Konstante c_n mit

$$|\tilde{f}(k)| \leq c_n \cdot (1 + |k|)^{-n}, \quad x \in \mathbb{R}.$$

Beweis. Die Voraussetzungen an f erlauben es uns, Satz 3.4 b) iteriert anzuwenden und

$$\widetilde{f^{(l)}}(k) = i^l k^l \tilde{f}(k), \quad l = 0, 1, \dots, n,$$

zu erhalten. Also gibt es Konstanten C_l , so dass $|k^l \tilde{f}(k)| \leq C_l$ für alle $k \in \mathbb{R}$. Damit erhalten wir

$$(1 + |k|)^n |\tilde{f}(k)| = \sum_{l=0}^n \binom{n}{l} |k|^l |\tilde{f}(k)| \leq \sum_{l=0}^n \binom{n}{l} C_l =: c_n,$$

wie behauptet. □

Analog können Sie beweisen: Falls $f \in L^1(\mathbb{R})$

$$|f(x)| \leq c(1 + |x|)^{-n}$$

für ein $n \in \mathbb{N}$ und eine Konstante $c > 0$ erfüllt, so ist die Fouriertransformierte \tilde{f} $(n-2)$ -mal differenzierbar. Denn die Abschätzung impliziert $|x^{n-2}f(x)| \leq c \frac{x^{n-2}}{(1+|x|)^n}$, was integrierbar ist. Also $Q^{n-2}(f) \in L^1(\mathbb{R})$, und wiederholte Anwendung von Satz 3.4 gibt die Behauptung.

An dieser Stelle zwei Warnungen zu häufigen Irrtümern:

- Eine Funktion in $L^1(\mathbb{R})$ muss nicht beschränkt sein – auch nicht, wenn sie stetig ist.
- Ist eine Funktion differenzierbar und beschränkt, so muss ihre Ableitung nicht beschränkt sein.

Diese Tatsachen machen die etwas kompliziert erscheinenden Annahmen zB in Korollar 3.6 nötig.

Wir betrachten nun einen Funktionenraum, der besonders gut auf die Fouriertransformation angepasst ist.

Definition 3.7. Eine Funktion $f : \mathbb{R}^d \rightarrow \mathbb{C}$ heißt *Schwartz-Funktion*, falls

- f glatt ist: $f \in C^\infty(\mathbb{R}^d)$, und
- f , und alle Ableitungen von f , schneller als polynomiell für $\|x\| \rightarrow \infty$ gegen Null gehen: Für jedes Paar von Multiindizes ${}^a \mu, \nu \in \mathbb{N}_0^d$ gibt es eine Konstante $c_{\mu, \nu}$, so dass

$$|x^\mu (\partial^\nu f)(x)| \leq c_{\mu, \nu}, \quad x \in \mathbb{R}^d.$$

Die Menge aller Schwartzfunktionen heißt der *Schwartzraum* (über \mathbb{R}^d) und wird mit $\mathcal{S}(\mathbb{R}^d)$ bezeichnet.

^aErinnerung an Analysis 2 (Def. 12.33): Für einen Multiindex $\mu = (\mu_1, \dots, \mu_d)$ schreiben wir $(\partial^\mu f)(x) = (\partial_1^{\mu_1} \cdots \partial_d^{\mu_d} f)(x)$ $\mu, \mu' \in \mathbb{N}_0^d$ und $|\mu| := \mu_1 + \dots + \mu_d$ für die Ordnung von μ . Wir schreiben auch $(Q^\mu f)(x) := x^\mu f(x) := x_1^{\mu_1} \cdots x_d^{\mu_d} f(x)$.

Für $d = 1$ vereinfacht sich die Definition, da es nur eine Ableitung statt mehrerer partieller Ableitungen und nur ein Argument x statt mehrerer Komponenten x_μ gibt. Es gilt

$$\mathcal{S}(\mathbb{R}) = \{f \in C^\infty(\mathbb{R}) : \sup_{x \in \mathbb{R}} |x^m f^{(n)}(x)| < \infty \forall n, m \in \mathbb{N}_0\}.$$

Beispiel 3.8. Die Funktion $f(x) = e^{-x^2}$ liegt im Schwartzraum: Sie ist glatt, und ihre n -te Ableitung ist von der Form $f^{(n)}(x) = p(x)e^{-x^2}$ mit einem Polynom p . Also gilt für jedes $m \in \mathbb{N}$

$$|x^m f^{(n)}(x)| = x^m p(x) e^{-x^2},$$

was für $|x| \rightarrow \infty$ gegen Null geht und insbesondere beschränkt ist.

Lemma 3.9. Sei $f \in \mathcal{S}(\mathbb{R}^d)$.

- a) Für jeden Multiindex μ gilt $\partial^\mu f \in L^1(\mathbb{R}^d)$.
- b) $\tilde{f} \in \mathcal{S}(\mathbb{R}^d)$.

Beweis. a) Aufgrund der Schwatz-Eigenschaft von f finden wir eine Konstante c , so dass

$$|(\partial^\mu f)(x)| \leq \frac{c}{(1+x_1^2) \cdots (1+x_d^2)}$$

gilt. Wir können also das Integral über $|\partial^\mu f|$ wie folgt abschätzen (Fubini):

$$\int_{\mathbb{R}^d} |(\partial^\mu f)(x)| dx \leq \int_{\mathbb{R}^d} \frac{c}{(1+x_1^2) \cdots (1+x_d^2)} dx = c \left(\int_{\mathbb{R}} \frac{dx}{1+x^2} \right)^d = c\pi^d < \infty.$$

b) Wir wenden Satz 3.4 a) wiederholt an und sehen, dass \tilde{f} glatt ist (partielle Ableitungen beliebiger Ordnung existieren). Da $\partial^\mu f \in L^1(\mathbb{R}^d)$, gilt nach Teil b) dieses Satzes

$$Q^\mu \partial^\nu \tilde{f} = (-i)^{|\nu|} Q^\mu \widetilde{Q^\nu f} = (-i)^{|\alpha|+|\nu|} \widetilde{\partial^\mu Q^\nu f}.$$

Per Produktregel lassen sich die partiellen Ableitungen auf der rechten Seite in eine Linearkombination endlich vieler Terme der Form $Q^{\nu'} \partial^{\mu'} f$ umformen. Dies sind nach a) L^1 -Funktionen, also ist ihre Fouriertransformierte beschränkt (Lemma 3.3). \square

Diese Ergebnisse illustrieren die angenehmen Eigenschaften von Schwartzfunktionen. Insbesondere können wir die Fouriertransformation also zu einer Abbildung

$$\mathcal{F} : \mathcal{S}(\mathbb{R}^d) \rightarrow \mathcal{S}(\mathbb{R}^d)$$

einschränken.

3.2 Umkehrabbildung und Anwendungen

Wir studieren jetzt die Frage, auf welchen Räumen die Fouriertransformation invertierbar ist, und wie die Umkehrabbildung aussieht. Wir beginnen mit einigen vorbereitenden Überlegungen.

Lemma 3.10. Seien $f, g \in L^1(\mathbb{R}^d)$, $a \in \mathbb{R}^d$, und $r > 0$. Dann gilt

- a) Die Fouriertransformierte von $f_a(x) := f(x-a)$ ist $\tilde{f}_a(k) = e^{-i\langle k, a \rangle} \tilde{f}(k)$.
- b) Die Fouriertransformierte von $g_r(x) := g(\frac{x}{r})$ ist $\tilde{g}_r(k) = r^d \tilde{g}(r \cdot k)$.
- c) Das Produkt der Fouriertransformierten ist

$$\tilde{f} \cdot \tilde{g} = (2\pi)^{-d/2} \widetilde{f * g},$$

wobei $(f * g)(x) = \int_{\mathbb{R}^d} f(y)g(x-y)dy$ die Faltung von f und g ist.

Der Beweis dieses Lemmas erfolgt in den Übungen.

Lemma 3.11. Für $f, g \in L^1(\mathbb{R}^d)$ sind $\tilde{f} \cdot g$ und $f \cdot \tilde{g}$ integrierbar, und es gilt

$$\int_{\mathbb{R}^d} \tilde{f}(x)g(x) dx = \int_{\mathbb{R}^d} f(k)\tilde{g}(k) dk.$$

Beweis. Wir wissen bereits, dass \tilde{f}, \tilde{g} stetig und beschränkt sind. Also sind die Funktionen $\tilde{f} \cdot g$ und $f \cdot \tilde{g}$ messbar. Nach der Hölderschen Ungleichung gilt $\|\tilde{f}g\|_1 \leq \|\tilde{f}\|_\infty \|g\|_1 < \infty$ und $\|f\tilde{g}\|_1 \leq \|f\|_1 \|\tilde{g}\|_\infty < \infty$. Hier haben wir $f, g \in L^1(\mathbb{R}^d)$ und $\tilde{f}, \tilde{g} \in C_b(\mathbb{R}^d)$ benutzt – letztere Funktionen sind beschränkt, also insbesondere wesentlich beschränkt. Wir schließen, dass $\tilde{f}g$ und $f\tilde{g}$ integrierbar sind.

Mit dem Satz von Fubini erhalten wir nun

$$\begin{aligned} \int_{\mathbb{R}^d} \tilde{f}(x)g(x) dx &= (2\pi)^{-d/2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(k)e^{-i\langle k,x \rangle} dk g(x) dx \\ &= (2\pi)^{-d/2} \int_{\mathbb{R}^d} f(k) \int_{\mathbb{R}^d} g(x) e^{-i\langle k,x \rangle} dx dk \\ &= \int_{\mathbb{R}^d} f(k) \tilde{g}(k) dk. \end{aligned}$$

□

Als nächsten Punkt besprechen wir eine sehr oft nützliche Approximationstechnik. Wir beginnen mit einem Beispiel in einer vereinfachten Situation.

Beispiel 3.12. Wir betrachten in Dimension $d = 1$ die Funktionen $g_r = \frac{1}{2r}\chi_{[-r,r]}$. Im Limes $r \rightarrow 0$ haben wir

$$g_r(x) \rightarrow \begin{cases} 0 & x \neq 0 \\ \infty & x = 0 \end{cases},$$

die Grenzfunktion ist also ziemlich singulär und fast überall gleich Null. Sei nun $f : \mathbb{R} \rightarrow \mathbb{C}$ stetig. Dann ist $f g_r \in L^1(\mathbb{R})$, und es gilt mit einer Stammfunktion F von f

$$\begin{aligned} \int_{\mathbb{R}} f(x)g_r(x)dx &= \frac{1}{2r} \int_{-r}^r f(x)dx = \frac{F(r) - F(-r)}{2r} \\ &= \frac{F(r) - F(0)}{2r} + \frac{F(0) - F(-r)}{2r} \\ &\xrightarrow{r \rightarrow 0} F'(0) = f(0). \end{aligned}$$

Der Grenzwert des Integrals ist also nicht singulär. Aufgrund der wachsenden Konzentration der g_r bei $x = 0$ konvergiert das Integral gegen den Wert von f an der Stelle $x = 0$.

Wir können das auch aus der Sicht von Maßen betrachten: Die Maße $g_r \cdot \lambda$ (also die Maße mit der Dichte g_r bzgl. dem Lebesguemaß λ) erfüllen für alle stetigen $f : \mathbb{R} \rightarrow \mathbb{C}$

$$\int_{\mathbb{R}} f d(g_r \lambda) \xrightarrow{r \rightarrow 0} \int_{\mathbb{R}} f d\delta_0 = f(0).$$

Das Ergebnis $f(0)$ sagt insbesondere, dass dieses Argument nicht für $f \in L^1(\mathbb{R})$ funktionieren kann, da $f(0)$ für die Äquivalenzklassen $f \in L^1(\mathbb{R})$ nicht definiert ist.

Ein allgemeineres Approximationsargument sieht folgendermaßen aus.

Lemma 3.13. Sei $\rho \in \mathcal{L}^1(\mathbb{R}^d)$ mit

$$\int_{\mathbb{R}^d} \rho(x) dx = 1,$$

und sei $\rho_r(x) := r^{-d} \rho(\frac{x}{r})$, $r > 0$. Dann gilt für alle $f \in C_b(\mathbb{R}^d)$

$$\lim_{r \rightarrow 0} (\rho_r * f)(x) = f(x).$$

Beweis. Wir haben

$$(\rho_r * f)(x) = \int_{\mathbb{R}^d} f(x - y) \frac{\rho(\frac{y}{r})}{r^d} dy = \int_{\mathbb{R}^d} f(x - r \cdot z) \rho(z) dz,$$

wobei wir $z := y/r$ substituiert haben (Der Integrand liegt in $\mathcal{L}^1(\mathbb{R}^d)$ wegen $\rho \in \mathcal{L}^1(\mathbb{R}^d)$ und $f \in C_b(\mathbb{R}^d)$). Der Integrand hat die integrierbare Majorante $M(z) := \|f\|_\infty |\rho(z)|$, die $|f(x - rz)\rho(z)| \leq M(z)$ für alle $r > 0$ erfüllt. Da der Limes $r \rightarrow 0$ unter dem Integral $\int_{\mathbb{R}^d} f(x)\rho(z)dz = f(x)$ ergibt (hier haben wir die Stetigkeit von f benutzt), folgt die Behauptung. \square

Beispiel 3.14. Eine oft verwendete Wahl von ρ ist eine Gaußfunktion, nämlich

$$\rho(x) = (2\pi)^{-d/2} e^{-\frac{\|x\|^2}{2}}.$$

Dann gilt $\int_{\mathbb{R}^d} \rho(x) dx = 1$, wie man aus Beispiel 3.5 ersieht. (Da die Gaußfunktion $\rho(x) = \prod_{i=1}^d \left((2\pi)^{-1/2} \exp(-\frac{x_i^2}{2}) \right)$ eine Produktfunktion ist, zerfällt das Integral in ein Produkt von Integralen über \mathbb{R} .)

Es gibt viele weitere Varianten von dieser Technik. Zum Beispiel ist es richtig, dass für $f \in L^1(\mathbb{R}^d)$

$$\lim_{r \rightarrow 0} \|f - \rho_r * f\|_1 = 0$$

gilt.

Wir kommen nun zur Umkehrformel der Fouriertransformation. Da für $f \in L^1(\mathbb{R}^d)$ die Fouriertransformierte \tilde{f} in $C_b(\mathbb{R}^d)$ liegt, betrachten wir die symmetrische Funktion, dass sowohl f als auch \tilde{f} in $L^1(\mathbb{R}^d) \cap C_b(\mathbb{R}^d)$ liegen. Das ist sicherlich für alle $f \in \mathcal{S}(\mathbb{R}^d)$ der Fall.

Satz 3.15 (Umkehrabbildung der Fouriertransformation).

a) Sei $f \in L^1(\mathbb{R}^d) \cap C_b(\mathbb{R}^d)$ so, dass auch $\tilde{f} \in L^1(\mathbb{R}^d)$. Dann gilt für alle $x \in \mathbb{R}^d$

$$f(x) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} \tilde{f}(k) e^{i\langle k, x \rangle} dk.$$

b) Die Fouriertransformation ist eine bijektive lineare Abbildung $\mathcal{F} : \mathcal{S}(\mathbb{R}^d) \rightarrow \mathcal{S}(\mathbb{R}^d)$, mit Umkehrabbildung

$$(\mathcal{F}^{-1}f)(x) = (\mathcal{F}f)(-x) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} f(k) e^{i\langle k, x \rangle} dk.$$

Beweis. a) Sei ρ die Gaußfunktion aus Beispiel 3.14. Dann gilt: i) $\rho \in \mathcal{L}^1(\mathbb{R}^d)$, ii) $\rho(0) = (2\pi)^{-d/2}$, iii) $\tilde{\rho} = \rho$, insbesondere $\tilde{\rho}(-y) = \tilde{\rho}(y)$ für alle $y \in \mathbb{R}^d$ (siehe Beispiel 3.5 b)), und iv) $|\rho(p)| \leq (2\pi)^{-d/2}$ für alle $p \in \mathbb{R}^d$.

Mit Lemma 3.10 b) sehen wir, dass $\tilde{\rho}_r : y \mapsto \frac{\tilde{\rho}(\frac{-y}{r})}{r^d}$ die Fouriertransformierte von $k \mapsto \rho(r \cdot k)$ ist. Mit der Abkürzung $f_{-x}(y) := f(x + y)$ gilt dann aufgrund von Lemma 3.10 a) und Lemma 3.11 für jedes $x \in \mathbb{R}^d$

$$\begin{aligned} (\rho_r * f)(x) &= \int_{\mathbb{R}^d} f(x - y) \frac{\tilde{\rho}(\frac{-y}{r})}{r^d} dy \\ &= \int_{\mathbb{R}^d} f_{-x}(y) \tilde{\rho}_r(y) dy \\ &= \int_{\mathbb{R}^d} \tilde{f}(k) e^{i\langle x, k \rangle} \rho(r \cdot k) dk. \end{aligned}$$

Da $f \in C_b(\mathbb{R}^d)$, geht die linke Seite für $r \rightarrow 0$ gegen $(\rho_r * f)(x) \rightarrow f(x)$ (Lemma 3.13).

Für die rechte Seite haben wir $|\tilde{f}(k) e^{i\langle x, k \rangle} \rho(r \cdot k)| \leq (2\pi)^{-d/2} |\tilde{f}(k)| =: M(k)$ für alle $r > 0$. Da $\tilde{f} \in L^1(\mathbb{R}^d)$, ist dies eine r -unabhängige integrierbare Majorante. Wegen $\rho(rk) \rightarrow \rho(0) = (2\pi)^{-d/2}$ für $r \rightarrow 0$ erhalten wir so die behauptete Umkehrformel

$$f(x) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} \tilde{f}(k) e^{i\langle k, x \rangle} dk.$$

b) Da Schwartzfunktionen die Voraussetzungen von a) erfüllen, sehen wir anhand der Formel, dass $\mathcal{F} : \mathcal{S}(\mathbb{R}^d) \rightarrow \mathcal{S}(\mathbb{R}^d)$ injektiv ist (falls $\tilde{f}(k) = 0$ für alle $k \in \mathbb{R}^d$, so auch $f = 0$).

Nun definieren wir $(\hat{\mathcal{F}}f)(k) := \tilde{f}(-k)$. Aufgrund der großen Ähnlichkeit zu \mathcal{F} sehen wir direkt, dass $\hat{\mathcal{F}} : \mathcal{S}(\mathbb{R}^d) \rightarrow \mathcal{S}(\mathbb{R}^d)$ linear und injektiv ist. Angesichts der Umkehrformel gilt

$$\hat{\mathcal{F}} \circ \mathcal{F} = \text{id}_{\mathcal{S}(\mathbb{R}^d)}, \quad \mathcal{F} \circ \hat{\mathcal{F}} = \text{id}_{\mathcal{S}(\mathbb{R}^d)}.$$

Dies impliziert, dass \mathcal{F} und $\hat{\mathcal{F}}$ bijektiv und Umkehrabbildungen voneinander sind. \square

Teil a) dieses Satzes gilt auch unter der schwächeren Voraussetzung $f, \tilde{f} \in L^1(\mathbb{R}^d)$; in diesem Fall gilt die Umkehrformel fast überall.

Das Überraschende an der Umkehrabbildung der Fouriertransformation ist, dass sie bis auf das Vorzeichen im Exponenten der ebenen Welle mit \mathcal{F} übereinstimmt. Insbesondere gilt $\mathcal{F}^2(f)(x) = f(-x)$ für alle $f \in \mathcal{S}(\mathbb{R}^d)$.

Die Umkehrabbildung der Fouriertransformation ist eine wesentliche Eigenschaft von \mathcal{F} . Sie sagt uns, dass jede Funktion (zumindest jede Schwartzfunktion) als Überlagerung ebener Wellen geschrieben werden kann, und zwar auf eindeutige Art und Weise. Die Fouriertransformierte Funktion \tilde{f} enthält also alle Information über f .

Beispiel 3.16. Was ist die Fouriertransformierte von $h : \mathbb{R} \rightarrow \mathbb{R}$, $h(x) := \frac{1}{1+x^2}$? Wir wissen aus Beispiel 3.2, dass $f(x) := e^{-|x|}$ die Fouriertransformierte $\tilde{f}(k) = \sqrt{\frac{2}{\pi}} h(k)$ hat. Also gilt

$$\mathcal{F}(h)(k) = \sqrt{\frac{\pi}{2}} \mathcal{F}(\mathcal{F}(f))(k) = \sqrt{\frac{\pi}{2}} \mathcal{F}^{-1}(\mathcal{F}(f))(-k) = \sqrt{\frac{\pi}{2}} f(-k) = \sqrt{\frac{\pi}{2}} e^{-|k|}.$$

Eine Anwendung der Umkehrformel auf Faltungen:

Beispiel 3.17. Für $f, g \in \mathcal{S}(\mathbb{R})$ gilt

$$\widetilde{fg} = (2\pi)^{-d/2} f * g.$$

Wir wissen bereits $\tilde{f} \cdot \tilde{g} = (2\pi)^{-d/2} \widetilde{f * g}$ (Lemma 3.10), nach Anwendung von \mathcal{F}^{-1} also

$$(2\pi)^{-d/2} f * g = \mathcal{F}^{-1}(\mathcal{F}(f) \cdot \mathcal{F}(g)) \Rightarrow (2\pi)^{-d/2} \mathcal{F}^{-1}(f) * \mathcal{F}^{-1}(g) = \mathcal{F}^{-1}(f \cdot g).$$

Da diese Gleichung für alle $f, g \in \mathcal{S}(\mathbb{R}^d)$ gilt und \mathcal{F}^{-1} surjektiv ist, folgt die Behauptung.

Wir geben ein Anwendungsbeispiel, in dem viele der bisher etablierten Eigenschaften der Fouriertransformation zusammenspielen.

Beispiel 3.18. Wir betrachten einen Metalldraht, der zur Zeit $t \in \mathbb{R}$ an der Stelle $x \in \mathbb{R}$ die Temperatur $T(t, x) \in \mathbb{R}$ hat. Die allmähliche Verteilung der Wärme wird durch die *Wärmeleitungsgleichung* beschrieben, nämlich

$$\frac{\partial T(t, x)}{\partial t} - a \frac{\partial^2 T(t, x)}{\partial x^2} = 0,$$

wobei $a > 0$ ein Parameter ist, der die Temperaturleitfähigkeit des Materials wiedergibt. Wir nehmen an, dass die Temperaturverteilung zur Zeit $t = 0$ eine bekannte Funktion $T_0 \in \mathcal{S}(\mathbb{R})$ ist.

Eine typische Fragestellung ist dann das *Anfangswertproblem*: Finde eine (die?) Lösung T der Wärmeleitungsgleichung, die

$$T(0, x) = T_0(x)$$

erfüllt. Wir machen außerdem eine Regularitätsannahme an die Lösung: Zu jedem $t > 0$ möge es $\varepsilon > 0$ und $c > 0$ geben, so dass

$$\sup_{t-\varepsilon < t' < t+\varepsilon} |\partial_{t'} T(t', x)| \leq \frac{c}{1+x^2}$$

für alle $x \in \mathbb{R}$ gilt.

Wir beschreiben hier, wie man dieses Problem per Fouriertransformation lösen kann. Es werden nicht alle Details ausgeführt, die volle Lösung soll in den Hausaufgaben geführt werden. Insbesondere die mit (\star) bezeichneten Aussagen sollen da detailliert begründet werden.

Die Hauptidee ist, die Lösung zu fester Zeit zu betrachten, dh $x \mapsto T_t(x) = T(t, x)$, und diese Funktionen zu Fouriertransformieren. Mit Satz 3.4 erhalten wir

$$\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \partial_t T_t(x) e^{-ikx} dx = a \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \partial_x^2 T_t(x) e^{-ikx} dx = -ak^2 \tilde{T}_t(k),$$

und die linke Seite ist $\partial_t \tilde{T}_t(k)$ (\star) . Wir haben es also mit der viel einfacheren Gleichung

$$\partial_t \tilde{T}_t(k) = -ak^2 \tilde{T}_t(k)$$

zu tun, die die eindeutige Lösung $\tilde{T}_t(k) = \tilde{T}_0(k) \cdot e^{-atk^2}$ hat (\star) . Um $T(t, x)$ zu erhalten, führen wir eine inverse Fouriertransformation aus und bekommen

$$T(t, x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \tilde{T}_0(k) \cdot e^{-atk^2} e^{ikx} dk.$$

Nach einer kleinen Rechnung ergibt sich (\star)

$$T(t, x) = \frac{1}{\sqrt{4\pi ta}} \int_{\mathbb{R}} T(0, y) e^{-\frac{(x-y)^2}{4ta}} dy, \quad t > 0.$$

Diese Formel beschreibt die Lösung zur Zeit $t > 0$.

Noch eine Bemerkung zum Schluss: Angenommen, $T : \mathbb{R}^2 \rightarrow \mathbb{R}$, $(t, x) \mapsto T(t, x)$, ist eine Lösung der Wärmeleitungsgleichung, die in $\mathcal{S}(\mathbb{R}^2)$ liegt. Dann können wir in t und x Fouriertransformieren und erhalten für die Fouriertransformierte $\tilde{T}(\omega, k)$

$$(i\omega + ak^2)\tilde{T}(\omega, k) = 0$$

durch Fouriertransformation der Wärmeleitungsgleichung. (Die Annahme $T \in \mathcal{S}(\mathbb{R}^2)$ ist unnötig stark.) Da $i\omega + ak^2$ nur für $\omega = 0$ und $k = 0$ verschwindet, sehen wir $\tilde{T} = 0$ (denn \tilde{T} ist stetig). Also $T = 0$. Es gibt also keine nichtverschwindenden Lösungen der Wärmeleitungsgleichung, die in *beiden* Variablen Schwartz sind.

3.3 Die Fouriertransformation auf L^2

Zum Abschluss unserer kleinen Tour durch die Fourier-Analyse wollen wir die Fouriertransformation auf $L^2(\mathbb{R}^d)$ besprechen. Da L^2 -Funktionen nicht in L^1 liegen müssen, wir \mathcal{F} bisher aber nur auf L^1 definiert hatten, werden wir unsere Definition erweitern müssen.

Erinnern Sie sich, dass wir unter $L^2(\mathbb{R}^d)$ nun alle *komplexwertigen* (Äquivalenzklassen von) messbaren Funktionen $f : \mathbb{R}^d \rightarrow \mathbb{C}$ verstehen, für die

$$\|f\|_2^2 = \int_{\mathbb{R}^d} |f(x)|^2 dx < \infty$$

gilt. Analog zu \mathbb{C}^n definieren wir ein sesquilineares Skalarprodukt via

$$\langle f, g \rangle := \int_{\mathbb{R}^d} \overline{f(x)} g(x) dx.$$

Die komplexe Konjugation auf dem linken Faktor bewirkt, dass $\langle f, f \rangle = \|f\|_2^2$ die Norm reproduziert.

Lemma 3.19. *Seien $f, g \in \mathcal{S}(\mathbb{R}^d)$. Dann gilt die Plancherel-Formel*

$$\langle f, g \rangle = \langle \tilde{f}, \tilde{g} \rangle.$$

Die Fouriertransformation $\mathcal{F} : \mathcal{S}(\mathbb{R}^d) \rightarrow \mathcal{S}(\mathbb{R}^d)$ erhält also die L^2 -Norm,

$$\|\mathcal{F}(f)\|_2 = \|f\|_2, \quad f \in \mathcal{S}(\mathbb{R}^d).$$

Beweis. Aufgrund der Formeln für \mathcal{F} und \mathcal{F}^{-1} gilt $\overline{\mathcal{F}(f)} = \mathcal{F}^{-1}(\overline{f})$. Somit erhalten wir aus Lemma 3.11

$$\begin{aligned} \langle \tilde{f}, \tilde{g} \rangle &= \int_{\mathbb{R}^d} \overline{\mathcal{F}(f)(k)} \mathcal{F}(g)(k) dk \\ &= \int_{\mathbb{R}^d} \mathcal{F}^{-1}(\overline{f})(k) \mathcal{F}(g)(k) dk = \int_{\mathbb{R}^d} \overline{f(x)} g(x) dx = \langle f, g \rangle. \end{aligned}$$

Die zweite Aussage ergibt sich, indem man $g = f$ setzt. □

Die Fouriertransformation ist also eine isometrische und damit insbesondere eine stetige Abbildung $\mathcal{S}(\mathbb{R}^d) \rightarrow \mathcal{S}(\mathbb{R}^d)$, wenn wir den Schwartzraum mit der L^2 -Norm $\|\cdot\|_2$ ausstatten. Allerdings ist $\mathcal{S}(\mathbb{R}^d)$ nicht vollständig in dieser Norm, wie wir in folgendem Beispiel illustrieren.

Beispiel 3.20. Sei für $n \in \mathbb{N}$

$$f_n : \mathbb{R} \rightarrow \mathbb{R}, \quad f_n(x) := e^{-x^2} \tanh(nx).$$

Diese Funktionen liegen in $\mathcal{S}(\mathbb{R})$, denn sie sind glatt und die Ableitungen sind wegen $\tanh'(x) = \frac{1}{\cosh^2(x)}$ alle schnell abfallend für $|x| \rightarrow \infty$ (Übung).

Für $n \rightarrow \infty$ haben wir punktweise

$$\lim_{n \rightarrow \infty} f_n(x) = \begin{cases} e^{-x^2} & x > 0 \\ 0 & x = 0 \\ -e^{-x^2} & x < 0 \end{cases} =: \text{sign}(x) e^{-x^2}$$

mit der Vorzeichenfunktion sign . Die Grenzfunktion $g(x) := \text{sign}(x) \cdot e^{-x^2}$ ist also unstetig und liegt damit nicht in $\mathcal{S}(\mathbb{R})$.

Betrachten wir den Abstand von f_n zu g in L^2 -Norm,

$$\|f_n - g\|_2^2 = \int_{\mathbb{R}} e^{-2x^2} |\tanh(nx) - \text{sign}(x)|^2 dx,$$

so können wir wegen $e^{-2x^2} |\tanh(nx) - \text{sign}(x)|^2 \leq 4e^{-2x^2}$ (wegen $|\tanh(y)| \leq 1$ für alle $y \in \mathbb{R}$) majorisierte Konvergenz verwenden und sehen $\|f_n - g\|_2 \rightarrow 0$ für $n \rightarrow \infty$.

Insbesondere ist (f_n) eine in $\mathcal{S}(\mathbb{R})$ liegende Folge, die bzgl L^2 -Norm eine Cauchyfolge ist. Der Grenzwert liegt aber nicht in $\mathcal{S}(\mathbb{R})$, also ist $(\mathcal{S}(\mathbb{R}), \|\cdot\|_2)$ nicht vollständig.

Wir diskutieren nun, wie L^2 -Funktionen durch Schwartzfunktionen approximiert werden können. Dazu erinnern wir uns zunächst an Analysis 1: Eine Teilmenge M eines metrischen Raumes X heißt *dicht*, falls ihr Abschluss \bar{M} gleich dem ganzen Raum X ist. In diesem Fall gibt es also zu jedem $x \in X$ eine Folge $(m_n) \subset M$ mit $m_n \rightarrow x$. Wir möchten zeigen, dass $\mathcal{S}(\mathbb{R}^d) \subset L^2(\mathbb{R}^d)$ dicht in der von der L^2 -Norm induzierten Metrik ist.

Dazu erwähnen wir zunächst folgenden Satz (zuerst eine Definition).

Definition 3.21.

- a) Sei X ein metrischer Raum, V ein Vektorraum, und $f : X \rightarrow V$ eine Funktion. Der *Träger* (Englisch: support) von f ist

$$\text{supp}(f) := \overline{\{x \in X : f(x) \neq 0\}}.$$

- b) $C_c(\mathbb{R}^d)$ ist der Vektorraum aller stetiger Funktionen $f : \mathbb{R}^d \rightarrow \mathbb{C}$ mit kompaktem Träger.

Eine Funktion $f : \mathbb{R}^d \rightarrow \mathbb{C}$ hat kompakten Träger genau dann, wenn es $R > 0$ gibt, so dass f außerhalb der Kugel $K_R = \{x \in \mathbb{R}^d : \|x\| \leq R\}$ verschwindet (damit ist der Träger beschränkt, abgeschlossen ist er per Definition). Damit ist klar, dass $C_c(\mathbb{R}^d)$ ein Vektorraum ist. Analog schreiben wir $C_c^\infty(\mathbb{R}^d)$ für den Raum aller glatten Funktionen $f : \mathbb{R}^d \rightarrow \mathbb{C}$, die kompakten Träger haben. Es gilt $C_c^\infty(\mathbb{R}^d) \subset \mathcal{S}(\mathbb{R}^d)$.

Satz 3.22. Sei $d \in \mathbb{N}$ und $p \in [1, \infty)$. Der Raum $C_c(\mathbb{R}^d)$ aller stetigen Funktionen mit kompaktem Träger liegt dicht in $L^p(\mathbb{R}^d)$ in der von der L^p -Norm induzierten Metrik.

Aus Zeitgründen geben wir den Beweis nicht an, siehe das Buch von Forster Satz 6 in §12. Wir illustrieren das Argument an einem einfachen Beispiel.

Beispiel 3.23. Sei $I := [a, b] \subset \mathbb{R}$ ein kompaktes Intervall und $\chi_I \in L^p(\mathbb{R})$ die zugehörige charakteristische Funktion. Definiere für $n \in \mathbb{N}$

$$f_n : \mathbb{R} \rightarrow \mathbb{R}, \quad f_n(x) := \begin{cases} n(x - a + \frac{1}{n}) & a - \frac{1}{n} \leq x < a \\ 1 & a \leq x \leq b \\ -n(x - b - \frac{1}{n}) & b < x \leq b + \frac{1}{n} \\ 0 & \text{sonst} \end{cases}.$$

Dann gilt $f_n \in C_c(\mathbb{R})$, nämlich $\text{supp}(f_n) \subset [a - \frac{1}{n}, b + \frac{1}{n}]$, und $\|f_n\|_\infty = 1$. Da f_n mit χ_I auf I übereinstimmt, erhalten wir

$$\|\chi_I - f_n\|_p^p = \int_{a-\frac{1}{n}}^a |f_n(x)|^p dx + \int_b^{b+\frac{1}{n}} |f_n(x)|^p dx \leq \frac{1}{2n} \xrightarrow{n \rightarrow \infty} 0.$$

Also kann χ_I in L^p -Norm von Funktionen in $C_c(\mathbb{R})$ approximiert werden, dh $\chi_I \in \overline{C_c(\mathbb{R})}$. Ähnlich kann man für charakteristische Funktionen von Quadern in höherer Dimension argumentieren.

Satz 3.24.

a) Sei $\rho \in L^1(\mathbb{R}^d)$ mit $\int_{\mathbb{R}^d} \rho(x) dx = 1$. Dann gilt für alle $f \in L^p(\mathbb{R}^d)$ (mit $p \in [1, \infty]$)

$$\lim_{r \rightarrow 0} \|\rho_r * f - f\|_p = 0,$$

wobei $\rho_r(x) = r^{-d} \rho(x/r)$.

b) Sei $1 \leq p < \infty$. Der Raum $C_c^\infty(\mathbb{R}^d)$ aller glatten Funktionen mit kompaktem Träger liegt dicht in $L^p(\mathbb{R}^d)$ in der von der L^p -Norm induzierten Metrik. Insbesondere ist $\mathcal{S}(\mathbb{R}^d) \subset L^p(\mathbb{R}^d)$ dicht.

Beweis. a) Dies ist die L^p -Version von Lemma 3.13. Aus Zeitgründen geben wir den Beweis für diesen Teil nicht an. Siehe zB. Folland "Real Analysis" Theorem 8.14.

b) Es genügt zu zeigen, dass für eine Funktion $f \in C_c(\mathbb{R}^d)$ eine Folge $f_n \in C_c^\infty(\mathbb{R}^d)$ mit $\|f - f_n\|_p \rightarrow 0$ existiert. Wir wählen dazu $\rho \in C_c^\infty(\mathbb{R}^d)$ mit $\int_{\mathbb{R}^d} \rho d\lambda = 1$ wie in Teil a), und setzen $f_n := \rho_{1/n} * f$. Dann gilt nach Lemma 3.13 $f_n \rightarrow f$ punktweise, und nach Teil a) gilt dieser Limes auch in L^p -Norm. Nach Hausaufgabe H8.1 e) ist die Faltung $\rho_{1/n} * f$ glatt. Nach Hausaufgabe H8.1c) hat $\rho_{1/n} * f$ kompakten Träger, dh $f_n \in C_c^\infty(\mathbb{R}^d)$.

Wegen $C_c^\infty(\mathbb{R}^d) \subset \mathcal{S}(\mathbb{R}^d)$ folgt insbesondere die Dichtheit von $\mathcal{S}(\mathbb{R}^d)$. □

Nun können wir die Fouriertransformation auf $L^2(\mathbb{R}^d)$ ausdehnen.

Satz 3.25 (Satz von Plancherel). *Es gibt eine eindeutige lineare bijektive Abbildung*

$$\hat{\mathcal{F}} : L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d)$$

mit folgenden Eigenschaften:

a) $\hat{\mathcal{F}}(f) = \mathcal{F}(f)$ und $\hat{\mathcal{F}}^{-1}(f) = \mathcal{F}^{-1}(f)$ für $f \in \mathcal{S}(\mathbb{R}^d)$,

b) $\|\hat{\mathcal{F}}(f)\|_2 = \|f\|_2$ für $f \in L^2(\mathbb{R}^d)$.

Die Abbildung $\hat{\mathcal{F}}$ erweitert also die Fouriertransformation \mathcal{F} auf $L^2(\mathbb{R}^d)$. Üblicherweise wird auch $\hat{\mathcal{F}}$ einfach mit \mathcal{F} bezeichnet.

Beweis. Eindeutigkeit: Falls $\hat{\mathcal{F}}$ und $\check{\mathcal{F}}$ zwei Abbildungen mit den angegebenen Eigenschaften sind, so gilt $G(f) := \hat{\mathcal{F}}(f) - \check{\mathcal{F}}(f) = 0$ für alle $f \in \mathcal{S}(\mathbb{R}^d)$. Nach b) ist $G : L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d)$ aber stetig. Ist also $f \in L^2(\mathbb{R}^d)$ beliebig und $f_n \in \mathcal{S}(\mathbb{R}^d)$ eine Folge mit $\|f_n - f\|_2 \rightarrow 0$, so gilt $G(f) = \lim_n G(f_n) = 0$. Also $\hat{\mathcal{F}} = \check{\mathcal{F}}$.

Existenz: Sei $f \in L^2(\mathbb{R}^d)$ beliebig. Nach Satz 3.24 existiert eine Folge $(f_n) \subset \mathcal{S}(\mathbb{R}^d)$ mit $\|f_n - f\|_2 \rightarrow 0$ für $n \rightarrow \infty$ (dh (f_n) konvergiert gegen f in L^2 -Norm). Mittels Lemma 3.19 sehen wir $\|\mathcal{F}(f_n) - \mathcal{F}(f_m)\|_2 = \|f_n - f_m\|_2 \rightarrow 0$ für $n, m \rightarrow \infty$. Also ist $(\mathcal{F}(f_n))_n$ eine Cauchyfolge bzgl L^2 -Norm. Da $L^2(\mathbb{R}^d)$ vollständig ist (Satz 2.35), gibt es einen Grenzwert in $L^2(\mathbb{R}^d)$. Wir definieren

$$\hat{\mathcal{F}}(f) := \lim_{n \rightarrow \infty} \mathcal{F}(f_n).$$

Diese Definition ist unabhängig von der ausgesuchten Folge (f_n) , die in L^2 -Norm gegen f konvergiert: Ist (f'_n) eine weitere solche Folge, so gilt $f_n - f'_n \rightarrow 0$ in L^2 -Norm, und damit $\mathcal{F}(f_n) - \mathcal{F}(f'_n) \rightarrow 0$ in L^2 -Norm. Also $\lim_n \mathcal{F}(f_n) = \lim_n \mathcal{F}(f'_n)$.

Wählen wir für $f, g \in L^2(\mathbb{R}^d)$ approximierende Folgen $(f_n), (g_n)$, so gilt $\lambda f_n + g_n \rightarrow \lambda f + g$. Die Linearität von \mathcal{F} impliziert dann, dass $\hat{\mathcal{F}}$ linear ist.

Per umgekehrter Dreiecksungleichung sehen wir

$$\left| \|\hat{\mathcal{F}}(f)\|_2 - \|\mathcal{F}(f)\|_2 \right| \leq \|\hat{\mathcal{F}}(f) - \mathcal{F}(f_n)\|_2 \rightarrow 0,$$

also $\|\hat{\mathcal{F}}(f)\|_2 = \lim_n \|\mathcal{F}(f_n)\|_2 = \lim_n \|f_n\|_2 = \|f\|_2$. Also gilt Eigenschaft b) ($\hat{\mathcal{F}}$ ist isometrisch), insbesondere ist $\hat{\mathcal{F}}$ injektiv. Für den Beweis der Surjektivität sei $g \in L^2(\mathbb{R}^d)$. Wir wählen eine Folge $(g_n) \subset \mathcal{S}(\mathbb{R}^d)$ mit $g_n \rightarrow g$ in L^2 -Norm, und betrachten $f := \lim_n \mathcal{F}^{-1}(g_n)$ (dieser Grenzwert existiert nach demselben Argument wie oben). Dann gilt $\hat{\mathcal{F}}(f) = \lim_n \mathcal{F}(\mathcal{F}^{-1}(g_n)) = \lim_n g_n = g$. Also ist $\hat{\mathcal{F}}$ surjektiv, damit bijektiv.

Die erste Eigenschaft in a) gilt per Definition, und die zweite folgt analog für $\hat{\mathcal{F}}^{-1}$ anstelle von $\hat{\mathcal{F}}$. \square

In diesem Beweis haben wir eigentlich nur benutzt, dass $\mathcal{S}(\mathbb{R}^d) \subset L^2(\mathbb{R}^d)$ ein dichter Unterraum und $\mathcal{F} : \mathcal{S}(\mathbb{R}^d) \rightarrow \mathcal{S}(\mathbb{R}^d)$ eine bijektive isometrische Abbildung ist. Methoden dieser Art werden viel in der Funktionalanalysis verwendet.

Wir haben die Fouriertransformation jetzt auf $L^2(\mathbb{R}^d)$ ausgedehnt und bezeichnen sie im Folgenden einfach mit \mathcal{F} bzw $\tilde{f} = \mathcal{F}(f)$, $f \in L^2(\mathbb{R}^d)$. Die Situation ist ähnlich symmetrisch wie beim Schwartzraum; die Fouriertransformation ist eine Bijektion $L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d)$. Aber wie bestimmt man \tilde{f} für eine Funktion $f \in L^2(\mathbb{R}^d)$, die nicht in $L^1(\mathbb{R}^d)$ liegt? Das Fourier-Integral $\int f(x)e^{-i\langle k, x \rangle} dx$ muss dann nicht existieren.

Die Lösung zu dieser Frage lautet wie folgt: Da $\mathcal{S} \subset L^1 \cap L^2$, liegt insbesondere $L^1 \cap L^2$ dicht in L^2 . Auf L^1 funktioniert unsere ursprüngliche Definition von \mathcal{F} . Man kann sich leicht überzeugen, dass $\hat{\mathcal{F}}(f) = \lim_n \mathcal{F}(f_n)$ für eine Folge $f_n \in L^1 \cap L^2$ mit $\|f_n - f\|_2 \rightarrow 0$. Gegeben $f \in L^2(\mathbb{R}^d)$, ist ein Beispiel einer solchen Folge

$$f_n := f \cdot \chi_{K_{R_n}}, \quad K_{R_n} := \{x \in \mathbb{R}^d : \|x\|_2 \leq R_n\},$$

wobei (R_n) eine Folge von Radien mit $R_n \rightarrow \infty$ ist. Dann gilt also

$$\tilde{f}_n(k) = (2\pi)^{-d/2} \int_{K_{R_n}} f(x) e^{-i\langle k, x \rangle} dx,$$

und \tilde{f} ist der Limes von \tilde{f}_n in L^2 -Norm. Da Limiten in L^2 -Norm auch Limiten im quadratischen Mittel heißen, kürzt man dies ab als

$$\tilde{f}(k) = (2\pi)^{-d/2} \text{l.i.m.}_{R \rightarrow \infty} \int_{K_R} f(x) e^{-i\langle k, x \rangle} dx,$$

wobei l.i.m. für "limit in mean" steht.

Wir können auch direkt sehen, dass dieser Limes existiert:

$$\|\tilde{f} - \tilde{f}_n\|_2^2 = \|f - f_n\|_2^2 = \int_{\mathbb{R}^d \setminus K_{R_n}} |f(x)|^2 dx \xrightarrow{n \rightarrow \infty} 0.$$

(Für die Auswertung des Limes wurde majorisierte Konvergenz verwendet.)

Beispiel 3.26. In Beispiel 3.2 hatten wir gesehen, dass die charakteristische Funktion $\chi_{[-1,1]}$ die Fouriertransformierte $\tilde{\chi}_{[-1,1]}(k) = \sqrt{\frac{2}{\pi}} \frac{\sin k}{k}$ hat. Diese Funktion liegt in L^2 , aber nicht in L^1 . Ihre inverse L^2 -Fouriertransformation ist

$$f(x) := \text{l.i.m.}_{R \rightarrow \infty} \frac{1}{\pi} \int_{-R}^R \frac{\sin k}{k} e^{ikx} dk = \text{l.i.m.}_{R \rightarrow \infty} \frac{1}{\pi} \int_{-R}^R \frac{\sin k \cos(kx)}{k} dk.$$

Bei der zweiten Gleichung haben wir verwendet, dass $\frac{\sin k}{k} i \sin(kx)$ ungerade in k ist und deshalb nicht zum Integral beiträgt.

Mit etwas Mühe könnten wir den obigen Grenzwert berechnen. Wir wissen aber bereits, dass er (im Sinne von L^2) mit $\chi_{[-1,1]}(x)$ übereinstimmen muss, da es sich um die inverse Fouriertransformation von $\tilde{\chi}_{[-1,1]}$ handelt, und die L^2 -Fouriertransformation die L^1 -Fouriertransformation ausdehnt und bijektiv ist.

4 Differentialformen und Integration über Untermannigfaltigkeiten

Unsere bisherigen Überlegungen zur Integration haben uns auf eine Verallgemeinerung des Riemannintegrals durch maßtheoretisch definierte Integrale geführt. Im Vergleich zum Riemannintegral gibt es aber zwei wesentliche Unterschiede: Erstens ist das Riemannintegral $\int_a^b f(x)dx$ antisymmetrisch in den Grenzen a, b des Integrationsintervalls $[a, b]$, stimmt mit dem Lebesgueintegral $\int_{[a,b]} f(x)d\lambda_1(x)$ also nur für $a \leq b$ überein. Beim Riemannintegral haben wir es mit einem orientierten Intervall zu tun (intuitiv gesprochen integrieren wir von links nach rechts und andersherum mit einem Minuszeichen), beim Lebesgueintegral aber mit einem nicht orientierten Intervall $[a, b]$.

Zweitens haben wir bisher kein Analogon des Hauptsatzes der Differential- und Integralrechnung. Da die Formel $\int_a^b f(x)dx = F(b) - F(a)$ (mit einer Stammfunktion F von f) antisymmetrisch in a, b ist, können wir eine solche Verallgemeinerung auch nur für Integrale erwarten, die eine Orientierung berücksichtigen.

Im Folgenden werden wir eine Integrationstheorie diskutieren, die geeignete Objekte (sogenannte Differentialformen) über orientierte Untermannigfaltigkeiten des \mathbb{R}^d integriert. Das können d -dimensionale offene Teilmengen $U \subset \mathbb{R}^d$ sein (hier wird es einen engen Zusammenhang mit dem Lebesgueintegral geben), aber auch Untermannigfaltigkeiten niedriger Dimensionen, wie zB Kurven und Flächen im dreidimensionalen Raum. Diese Theorie wird insbesondere natürliche Integrationsfragen wie “Was ist der Flächeninhalt der Kugeloberfläche $S^2 = \{x \in \mathbb{R}^3 : \|x\|_2 = 1\} \subset \mathbb{R}^3$ ” beantworten.⁸

Um die folgenden Entwicklungen zu motivieren, betrachten wir ein Intervall $[a, b] \subset \mathbb{R}$ und eine stetige Funktion $f : [a, b] \rightarrow \mathbb{R}$. Wir unterteilen das Intervall an Zwischenpunkten $a = t_0, t_1, \dots, t_N = b$ und betrachten die Summen $\sum_{n=1}^N f(t_n) \cdot \Delta t_n$, $\Delta t_n := t_{n+1} - t_n$, die im Limes immer feinerer Unterteilungen gegen $\int_a^b f(t)dt$ konvergieren. Die Summanden $f(t_n)\Delta t_n$ hängen also linear von den Intervalllängen Δt_n ab, mit einer von f und dem Punkt t_n abhängigen Proportionalitätskonstante $f(t_n)$. Falls F eine Stammfunktion von f ist, so gilt

$$\sum_{n=1}^N f(t_n) \cdot \Delta t_n = \sum_{n=1}^N F'(t_n) \cdot \Delta t_n \approx \sum_{n=1}^N (F(t_{n+1}) - F(t_n)) = F(b) - F(a),$$

und im Grenzwert beliebig feiner Unterteilungen wird die Approximation \approx exakt.

Betrachten wir nun eine Kurve $\gamma : I := [0, 1] \rightarrow \mathbb{R}^3$ und ein noch zu bestimmendes Objekt ω , dass wir “über γ integrieren” möchten, so dass $\int_\gamma \omega$ eine analog zum Riemann-Integral definierte reelle Zahl ist. Wir unterteilen dazu I wie vorher und die Kurve an zugehörigen Zwischenpunkten $v_0 := \gamma(t_0), \dots, v_N := \gamma(t_N)$. Im Limes immer feinerer Unterteilungen konvergieren die Differenzvektoren $v_n - v_{n-1}$ gegen Vektoren, die tangential zur Kurve liegen (genügend Regularität von γ vorausgesetzt). In Analogie zu $\int_a^b f(t)dt$ sollte ω Tangentialvektoren an die Kurve linear in Zahlen abbilden, mit einer von ω und dem Punkt $x = \gamma(t)$ abhängigen Proportionalitätskonstanten. Das heißt, für jedes $x \in \gamma(I)$ sollte $\omega(x)$ ein lineares Funktional $T_x\gamma(I) \rightarrow \mathbb{R}$ auf dem Tangentialraum $T_x\gamma(I)$ sein. Solche Objekte werden

⁸Hier meinen wir nicht, dass S^2 bzgl des Lebesguemaßes λ_3 Volumen $\lambda_3(S^2) = 0$ hat, sondern den aus der Schule bekannten Flächeninhalt 4π .

wir als Differentialformen definieren (in diesem Fall 1-Formen). Das Integral $\int_{\gamma} \omega$ sollte sich dann als Grenzwert immer feinerer Unterteilungen von $\sum_n \omega(x)(\Delta x_n)$ (mit den Vektoren $x_n := \gamma(t_n)$ und $\Delta x_n = x_{n+1} - x_n$) ergeben.

Gegeben eine differenzierbare Funktion $f : \mathbb{R}^3 \rightarrow \mathbb{R}$, so hängen die Richtungsableitungen $(\partial_v f)(x) = \langle \text{grad} f(x), v \rangle$ linear von der Richtung v ab, geben uns also lineare Funktionale $df(x) : \mathbb{R}^3 \rightarrow \mathbb{R}$, $v \mapsto (\partial_v f)(x)$, die wir insbesondere zu linearen Funktionalen auf Tangentialvektoren an die Kurve γ einschränken können. Da $df(x)(v) \approx f(x+v) - f(x)$ für kleine v gilt, erwarten wir auch hier, dass $\sum_{n=1}^N (df)(x_n)(\Delta x_n) \approx f(b) - f(a)$ auf $\int_{\gamma} df = f(\gamma(1)) - f(\gamma(0))$ als Analogie zum Hauptsatz der Differential- und Integralrechnung führt.

Als weiteres Beispiel in dieser informellen Diskussion betrachten wir eine orientierte zweidimensionale Fläche $M \subset \mathbb{R}^3$. Eine genaue Definition von Orientierung erfolgt später; stellen Sie sich vor, dass eine Seite der Fläche M als "Außenseite" deklariert ist.

Der Einfachheit halber nehmen wir an, dass $M = \phi([0, 1]^2)$ mit einer genügend regulären Funktion $\phi : [0, 1]^2 \rightarrow \mathbb{R}^3$, die die Rolle von γ aus der Diskussion von Kurven übernimmt. Dann unterteilen wir $[0, 1]^2$ in viele kleine Rechtecke, mit den Ecken (t_n, t_m) , $(t_n + \Delta t, t_m)$, $(t_n, t_m + \Delta t)$, $(t_n + \Delta t, t_m + \Delta t)$ und nähern M mit den Vierecken $V_{n,m}$ mit den Ecken $\phi(t_n, t_m)$, $\phi(t_n + \Delta t, t_m)$, $\phi(t_n, t_m + \Delta t)$, $\phi(t_n + \Delta t, t_m + \Delta t)$ an. Im Limes immer feinerer Unterteilungen konvergieren die Seiten dieser Vierecke gegen Vektoren, die tangential an M liegen. Im Unterschied zu der Kurve ist der Tangentialraum $T_x M$ jetzt zwei- statt eindimensional, und jedes der Vierecke $V_{n,m}$ ist von zwei statt einem Tangentialvektor aufgespannt. Wir betrachten deshalb eine Abbildung ω , die jedem $x \in M$ ein bilineares Funktional $\omega(x) : T_x M \times T_x M \rightarrow \mathbb{R}$ zuweist (der lokale Anteil am Integral sollte proportional zu den beiden Längen der beiden Kantenvektoren $\xi_{n,m}$, $\eta_{n,m}$ des Vierecks $V_{n,m}$ sein). Da wir für parallele Vektoren $\xi_{n,m}$, $\eta_{n,m}$ ein Viereck mit Flächeninhalt 0 haben, fordern wir unserer geometrischen Anschauung entsprechend weiterhin $\omega(x)(v, v) = 0$. Aufgrund der Bilinearität von $\omega(x)$ ist dies zu der Antisymmetrie $\omega(x)(v, w) = -\omega(x)(w, v)$ äquivalent. Wir haben also zu jedem Punkt $x \in M$ eine antisymmetrische Bilinearform $\omega(x) : (T_x M)^2 \rightarrow \mathbb{R}$. Dies wird die Definition einer Differentialform (in diesem Fall einer 2-Form) sein. Das Integral $\int_M \omega$ sollte sich dann als Limes immer feinerer Zerlegungen von $\sum_{n,m} \omega(x_{n,m})(\xi_{n,m}, \eta_{n,m})$ ergeben. Die Antisymmetrie von ω drückt die Abhängigkeit von der Orientierung aus: Wenn die Argumente von $\omega(x)$ vertauscht werden (dies kann als Vertauschung von "Außen- und Innenseite" von M angesehen werden), ändert sich das Integral um ein Vorzeichen.

Allgemeiner werden wir n -Formen definieren und über orientierte n -dimensionale Untermannigfaltigkeiten des \mathbb{R}^d integrieren. Dazu werden wir zunächst die algebraische Theorie der alternierenden Multilinearformen (siehe Lineare Algebra 2) diskutieren und dann zu den Differentialformen kommen. In diesem Zusammenhang werden wir die sogenannte äußere Ableitung kennenlernen, die erstmals dem Symbol " dx " eine unabhängige mathematische Bedeutung zuweist. Ein Hauptziel dieses Kapitels ist dann der Satz von Stokes, der eine Verallgemeinerung des Hauptsatzes der Differential- und Integralrechnung auf höhere Dimensionen und Untermannigfaltigkeiten darstellt.

Als Wiederholungsthemen zu diesem Kapitel empfehlen sich: Untermannigfaltigkeiten (Analysis 2 Kapitel 13.4), Dualräume (Lineare Algebra 2, Kapitel 15) und Multilinearformen (Lineare Algebra 2, Kapitel 23).

4.1 Alternierende Multilinearformen

Wir beginnen mit der rein algebraischen Theorie der alternierenden Multilinearformen. In diesem Abschnitt gibt es noch keinen direkten Zusammenhang zu Differentiation und Integration.

Definition 4.1. Sei V ein reeller Vektorraum.

- a) Eine *Multilinearform* der Ordnung $n \in \mathbb{N}$ ist eine Abbildung $\varphi : V^n \rightarrow \mathbb{R}$, die in jedem ihrer Argumente linear ist: Gegeben $k \in \{1, \dots, n\}$ und Vektoren $v_1, \dots, v_{k-1}, v_{k+1}, \dots, v_n \in V$, so ist

$$V \ni x \mapsto \varphi(v_1, \dots, v_{k-1}, x, v_{k+1}, \dots, v_n) \in \mathbb{R}$$

linear.

- b) Eine *alternierende (oder: antisymmetrische) Multilinearform* der Ordnung $n \in \mathbb{N}$ ist eine Multilinearform $\varphi : V^n \rightarrow \mathbb{R}$ mit der Eigenschaft, dass $\varphi(v_1, \dots, v_n) = 0$, falls zwei Argumente v_i übereinstimmen (d.h. es $i, j \in \{1, \dots, n\}$, so dass $i \neq j$ und $v_i = v_j$).
- c) Für festes $n \in \mathbb{N}$ nennt man die alternierenden Multilinearformen $V^n \rightarrow \mathbb{R}$ auch *alternierende n -Formen* oder *äußere n -Formen*. Die Menge aller alternierenden n -Formen wird mit $\bigwedge^n(V^*)$ bezeichnet. Per Konvention setzen wir $\bigwedge^0(V^*) := \mathbb{R}$.

Es lässt sich leicht zeigen, dass $\bigwedge^n(V^*)$ ein \mathbb{R} -Vektorraum ist bzgl. der punktweisen Operationen ($\varphi, \psi \in \bigwedge^n(V^*)$, $\lambda \in \mathbb{R}$)

$$(\varphi + \lambda \cdot \psi)(v_1, \dots, v_n) := \varphi(v_1, \dots, v_n) + \lambda \cdot \psi(v_1, \dots, v_n).$$

Zeigen Sie: Eine Multilinearform $\varphi : V^n \rightarrow \mathbb{R}$ ist genau dann alternierend, wenn sie unter der Vertauschung von zwei Argumenten ihr Vorzeichen wechselt, dh

$$\varphi(v_1, \dots, v_i, \dots, v_j, \dots, v_n) = -\varphi(v_1, \dots, v_j, \dots, v_i, \dots, v_n)$$

gilt. Zeigen Sie auch: Ist $\pi \in S_n$ eine Permutation von $(1, \dots, n)$ und $\varphi \in \bigwedge^n(V^*)$, so gilt

$$\varphi(v_{\pi(1)}, \dots, v_{\pi(n)}) = \text{sign}(\pi) \cdot \varphi(v_1, \dots, v_n).$$

Beispiel 4.2.

- a) Für $n = 1$ sind die Multilinearformen $\varphi : V \rightarrow \mathbb{R}$ nichts anderes als lineare Funktionale (also Linearformen; hier gibt es nur ein Argument), d.h. die 1-Formen stimmen mit dem Dualraum V^* von V überein.

Erinnerung an Lineare Algebra (Dualräume): Der Dualraum V^* eines Vektorraums V ist der Raum aller linearen Abbildungen $V \rightarrow \mathbb{R}$. Ist (v_1, \dots, v_d) eine Basis des (endlichdimensionalen) Vektorraums V , so bilden die durch $v_i^*(v_j) := \delta_{ij}$ definierten Funktionale v_i^* eine Basis (v_1^*, \dots, v_d^*) von V^* . Insbesondere gilt $V^* \cong V$ (Isomorphismus von Vektorräumen).

Ist V mit einem Skalarprodukt $\langle \cdot, \cdot \rangle$ ausgestattet, so dass (v_1, \dots, v_d) eine Orthonormalbasis ist (dh $\langle v_i, v_j \rangle = \delta_{ij}$), so kann der Isomorphismus $V \rightarrow V^*$, $v \mapsto v^*$, konkret als $v^*(w) = \langle v, w \rangle$ (mit $v, w \in V$) angegeben werden.

Da für $n = 1$ die Alternierungseigenschaft leer ist, gilt $\bigwedge^1(V^*) = V^*$.

b) Für jede $(d \times d)$ -Matrix A ist

$$\mathbb{R}^d \times \mathbb{R}^d \ni (v, w) \mapsto \langle v, Aw \rangle \in \mathbb{R}$$

eine Bilinearform (Multilinearform mit $n = 2$). Diese Form ist genau dann alternierend, wenn A antisymmetrisch ist, dh $A^T = -A$.

c) Ein Beispiel einer alternierenden d -Form über $V = \mathbb{R}^d$ ist die Determinante, wenn wir die Argumente als Spaltenvektoren einer Matrix lesen:

$$\det : (\mathbb{R}^d)^n \rightarrow \mathbb{R}, \quad \det(v_1, \dots, v_d) := \det \begin{pmatrix} (v_1)_1 & \dots & (v_d)_1 \\ (v_1)_2 & \dots & (v_d)_2 \\ \vdots & \ddots & \vdots \\ (v_1)_d & \dots & (v_d)_d \end{pmatrix}.$$

In Linearer Algebra wurde gezeigt, dass die Determinante bis auf Vielfache sogar die eindeutige alternierende d -Form ist, dh.

$$\bigwedge^d((\mathbb{R}^d)^*) = \mathbb{R} \det = \{c \cdot \det : c \in \mathbb{R}\}.$$

Aufgrund der Bedeutung der Determinante für Volumenmessungen deutet sich hier eine erste Verbindung zur Maß- und Integrationstheorie an.

Ein Beispiel einer Linearform auf einem unendlichdimensionalen Vektorraum ist das Integral

$$\mathcal{L}_{\mathbb{R}}^1(\mathbb{R}^d) \ni f \mapsto \int_{\mathbb{R}^d} f(x) dx \in \mathbb{R}.$$

Dieses Funktional ist nicht von der Form $f \mapsto \langle g, f \rangle$ für ein $g \in \mathcal{L}_{\mathbb{R}}^1(\mathbb{R}^d)$. Es ist noch nicht einmal klar, welches Skalarprodukt gemeint ist, und wenn wir zB das L^2 -Skalarprodukt wählen, so ist die Aussage falsch. Im Folgenden werden wir es nur mit Multilinearformen über endlichdimensionalen Vektorräumen zu tun haben.

Wir definieren nun ein "Produkt" von Linearformen.

Definition 4.3. Das *äußere Produkt* (oder *Dachprodukt*) von Linearformen $\varphi_1, \dots, \varphi_n \in V^*$ ist die Abbildung

$$\varphi_1 \wedge \dots \wedge \varphi_n : V^n \rightarrow \mathbb{R}, \quad (v_1, \dots, v_n) \mapsto \det \begin{pmatrix} \varphi_1(v_1) & \dots & \varphi_1(v_n) \\ \varphi_2(v_1) & \dots & \varphi_2(v_n) \\ \vdots & \ddots & \vdots \\ \varphi_n(v_1) & \dots & \varphi_n(v_n) \end{pmatrix}.$$

Bemerkungen:

- Das äußere Produkt von n Linearformen $\varphi_1, \dots, \varphi_n$ ist eine äußere n -Form: Offensichtlich ist $\varphi_1 \wedge \dots \wedge \varphi_n$ multilinear (die Determinante ist in den Zeilen linear), und die Alternierungseigenschaft der Determinante impliziert, dass $\varphi_1 \wedge \dots \wedge \varphi_n$ alternierend ist.
- Die Abbildung $(V^*)^n \ni (\varphi_1, \dots, \varphi_n) \mapsto \varphi_1 \wedge \dots \wedge \varphi_n \in \bigwedge^n(V^*)$ ist in jedem seiner Argumente linear, und alternierend, und es gilt

$$\varphi_{\pi(1)} \wedge \dots \wedge \varphi_{\pi(n)} = \text{sign}(\pi) \cdot \varphi_1 \wedge \dots \wedge \varphi_n$$

für jede Permutation $\pi \in S_n$.

- Für $\varphi_1, \dots, \varphi_n \in V^*$ gilt (Übung)

$$\varphi_1 \wedge \dots \wedge \varphi_n = 0 \Leftrightarrow \{\varphi_1, \dots, \varphi_n\} \text{ ist linear abhängig.}$$

Diese Tatsache erklärt die Bezeichnung "äußeres" Produkt: $\varphi_1 \wedge \dots \wedge \varphi_n$ ist nur dann von 0 verschieden, wenn φ_n nicht in $\text{span}\{\varphi_1, \dots, \varphi_{n-1}\}$, sondern im "Äußeren" des Untervektorraums $\text{span}\{\varphi_1, \dots, \varphi_{n-1}\}$ liegt.

Ist (v_1, \dots, v_d) eine Basis von V , so ist die duale Basis (v_1^*, \dots, v_d^*) eine Basis des Dualraums $V^* = \bigwedge^1(V^*)$. Wir beschreiben nun Basen für $\bigwedge^n(V^*)$ für allgemeines $n \in \mathbb{N}$.

Lemma 4.4. Sei V ein endlichdimensionaler Vektorraum und $\{\varphi_i : 1 \leq i \leq d\}$ eine Basis von V^* . Dann ist

$$\{\varphi_{i_1} \wedge \dots \wedge \varphi_{i_n} : 1 \leq i_1 < i_2 < \dots < i_n \leq d\}$$

eine Basis von $\bigwedge^n(V^*)$. Insbesondere gilt

$$\dim \bigwedge^n(V^*) = \binom{\dim V}{n}$$

und $\bigwedge^n(V^*) = \{0\}$ für $n > d = \dim V$.

Beweis. Wir wählen eine Basis (v_1, \dots, v_d) von V , so dass $\varphi_i = v_i^*$ gilt, dh $\varphi_i(v_j) = \delta_{ij}$. Nach Definition des äußeren Produktes gilt dann $(\varphi_{i_1} \wedge \dots \wedge \varphi_{i_n})(v_{j_1}, \dots, v_{j_n}) = 1$, falls $(i_1, \dots, i_n) = (j_1, \dots, j_n)$ (Determinante der Einheitsmatrix). Falls es einen Index j_k gibt, so dass $j_k \notin \{i_1, \dots, i_n\}$, so gilt $(\varphi_{i_1} \wedge \dots \wedge \varphi_{i_n})(v_{j_1}, \dots, v_{j_n}) = 0$ (Determinante einer Matrix mit einer Null-Spalte). Das impliziert, dass für zwei geordnete n -Tupel $i_1 < \dots < i_n$ und $j_1 < \dots < j_n$

$$(\varphi_{i_1} \wedge \dots \wedge \varphi_{i_n})(v_{j_1}, \dots, v_{j_n}) = \begin{cases} 1 & (i_1, \dots, i_n) = (j_1, \dots, j_n) \\ 0 & \text{sonst} \end{cases}$$

gilt. Sei nun $\omega \in \bigwedge^n(V^*)$. Dann gilt mit der Wahl $\omega_{i_1, \dots, i_n} := \omega(v_{i_1}, \dots, v_{i_n})$

$$\omega = \sum_{i_1 < \dots < i_n} \omega_{i_1, \dots, i_n} \cdot \varphi_{i_1} \wedge \dots \wedge \varphi_{i_n}, \quad (*)$$

denn beide Seiten sind alternierende n -Formen und stimmen auf allen Argumenten der Form $(v_{j_1}, \dots, v_{j_n})$ mit $j_1 < \dots < j_n$ überein. Aufgrund der Multilinearität und Alternierungseigenschaft stimmen sie dann auf beliebigen (w_1, \dots, w_n) , $w_k \in V$, überein. Weiterhin ist diese Darstellung von ω eindeutig, denn Anwendung auf $(v_{j_1}, \dots, v_{j_n})$ liefert $\omega(v_{j_1}, \dots, v_{j_n}) = \omega_{i_1, \dots, i_n}$. Das zeigt, dass die behauptete Menge eine Basis von $\bigwedge^n(V^*)$ ist.

Beachten Sie, dass die Summe in (\star) nur über alle *geordneten* n -Tupel läuft, dh über alle $(i_1, \dots, i_n) \in \{1, \dots, d\}^n$ mit $i_1 < \dots < i_n$. Solche n -Tupel gibt es für $n > d$ nicht, also $\bigwedge^n(V^*) = \{0\}$ für $n > d$. Für $1 \leq n \leq d$ entspricht die Wahl eines geordneten n -Tupels der Wahl einer n -elementigen Teilmenge von $\{1, \dots, d\}$, was $\dim \bigwedge^n(V^*) = \binom{d}{n}$ zeigt. \square

Die Zahlen ω_{i_1, \dots, i_n} werden die *Koeffizienten* von ω genannt. Sie hängen natürlich von der Basis $(\varphi_1, \dots, \varphi_d)$ ab, was in der Notation aber unterdrückt wird.

Die im Beweis benutzte Formel

$$(v_{i_1}^* \wedge \dots \wedge v_{i_n}^*)(v_{j_1}, \dots, v_{j_n}) = \begin{cases} 1 & (i_1, \dots, i_n) = (j_1, \dots, j_n) \\ 0 & \text{sonst} \end{cases}$$

für geordnete Indizes $i_1 < \dots < i_n$, $j_1 < \dots < j_n$, ist häufig nützlich. Weiterhin sollte man sich merken, dass $\varphi_{i_1} \wedge \dots \wedge \varphi_{i_n} = 0$, falls zwei i_j übereinstimmen, d.h. $i_a = i_b$ mit $a \neq b$.

Beispiel 4.5. Sei (e_1, \dots, e_d) die Standardbasis von \mathbb{R}^d . Dann gilt für $x = (x_1, \dots, x_d)$, $y = (y_1, \dots, y_d) \in \mathbb{R}^d$, $d \geq 2$,

$$(e_1^* \wedge e_2^*)(x, y) = \det \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix} = x_1 y_2 - x_2 y_1.$$

Anstelle der Determinante kann man auch mit den Eigenschaften von \wedge rechnen:

$$\begin{aligned} (e_1^* \wedge e_2^*)(x, y) &= x_1 y_1 (e_1^* \wedge e_2^*)(e_1, e_1) + x_1 y_2 (e_1^* \wedge e_2^*)(e_1, e_2) \\ &\quad + x_2 y_1 (e_1^* \wedge e_2^*)(e_2, e_1) + x_2 y_2 (e_1^* \wedge e_2^*)(e_2, e_2) \\ &= 0 + x_1 y_2 \cdot 1 + x_2 y_1 \cdot (-1) + 0 \\ &= x_1 y_2 - x_2 y_1. \end{aligned}$$

Wir können nun die Definition des äußeren Produktes auf alternierende Multilinearformen mit unterschiedlichen Anzahlen von Argumenten ausdehnen.

Definition und Satz 4.6. Sei V ein endlichdimensionaler Vektorraum und $k, l \in \mathbb{N}$. Es gibt genau eine bilineare Abbildung

$$\bigwedge^k V^* \times \bigwedge^l V^* \rightarrow \bigwedge^{k+l} V^*, \quad (\omega, \eta) \mapsto \omega \wedge \eta,$$

so dass für $\psi_1, \dots, \psi_k, \xi_1, \dots, \xi_l \in V^*$

$$(\psi_1 \wedge \dots \wedge \psi_k) \wedge (\xi_1 \wedge \dots \wedge \xi_l) = \psi_1 \wedge \dots \wedge \psi_k \wedge \xi_1 \wedge \dots \wedge \xi_l$$

gilt. Diese Abbildung heißt auch das *äußere Produkt*.

Beweis. Wir wählen eine Basis $(\varphi_1, \dots, \varphi_d)$ von V^* . Die Formen $\omega \in \bigwedge^k V^*$, $\eta \in \bigwedge^l V^*$ können wir dann in die Basis aus Lemma 4.4 entwickeln,

$$\omega = \sum_{i_1 < \dots < i_k} \omega_{i_1, \dots, i_k} \varphi_{i_1} \wedge \dots \wedge \varphi_{i_k}, \quad \eta = \sum_{j_1 < \dots < j_l} \eta_{j_1, \dots, j_l} \varphi_{j_1} \wedge \dots \wedge \varphi_{j_l}.$$

Das äußere Produkt von ω und η ist dann als

$$\omega \wedge \psi := \sum_{i_1 < \dots < i_k} \sum_{j_1 < \dots < j_l} \omega_{i_1, \dots, i_k} \psi_{j_1, \dots, j_l} \cdot \varphi_{i_1} \wedge \dots \wedge \varphi_{i_k} \wedge \varphi_{j_1} \wedge \dots \wedge \varphi_{j_l}$$

definiert. Dies ist offensichtlich bilinear und hat die gewünschte Eigenschaft $(\psi_1 \wedge \dots \wedge \psi_k) \wedge (\xi_1 \wedge \dots \wedge \xi_l) = \psi_1 \wedge \dots \wedge \psi_k \wedge \xi_1 \wedge \dots \wedge \xi_l$ für Linearformen ψ_i, ξ_j . Dies zeigt die Existenz des Produktes. Da jede k -Form gemäß Lemma 4.4 eine Linearkombination von Dachprodukten von Linearformen ist, folgt auch die Eindeutigkeit. Insbesondere hängt die Definition von \wedge nicht von der Wahl einer Basis ab. \square

Es ist oft praktisch, das äußere Produkt auch für äußere 0-Formen (Zahlen) zu definieren: Für $a \in \bigwedge^0 V^*$, $\omega \in \bigwedge^k V^*$ mit $k \in \mathbb{N}_0$

$$a \wedge \omega := \omega \wedge a := a \omega.$$

Da sich beim äußeren Produkt $\omega \wedge \eta$ die Ordnungen von ω und η addieren, bietet es sich an, alle alternierenden Multilinearformen (von beliebiger Ordnung) in einem Vektorraum zusammenzufassen.

Wir betrachten nun die Gesamtheit aller alternierenden n -Formen für beliebige n .

Definition und Satz 4.7. Sei V ein reeller Vektorraum. Die *äußere Algebra* über V ist

$$\bigwedge V^* := \bigoplus_{n=0}^{\infty} \bigwedge^n V^* \cong \bigoplus_{n=0}^{\dim V} \bigwedge^n V^*.$$

Mit der komponentenweisen Addition und dem äußeren Produkt ist $\bigwedge V^*$ eine assoziative gradiert kommutative Algebra.

Das heißt: Für $\omega \in \bigwedge^k V^*$, $\eta \in \bigwedge^l V^*$, $\psi \in \bigwedge^n V^*$ gilt

$$\begin{aligned} (\omega \wedge \eta) \wedge \psi &= \omega \wedge (\eta \wedge \psi), \\ \omega \wedge \eta &= (-1)^{kl} \eta \wedge \omega. \end{aligned}$$

Beweis. Alle Eigenschaften bis auf die gradierte Kommutativität $\omega \wedge \eta = (-1)^{kl} \eta \wedge \omega$ sind aus der Definition klar. Um auch diese Eigenschaft zu prüfen, bemerken wir, dass die Permutation

$$(i_1, \dots, i_k, j_1, \dots, j_l) \mapsto (j_1, \dots, j_l, i_1, \dots, i_k)$$

das Signum $(-1)^{kl}$ hat. Damit folgt die Behauptung. \square

Beispiel 4.8. Die *kanonische symplektische Form auf dem \mathbb{R}^{2d}* ist die 2-Form

$$\omega := \sum_{i=1}^d \varphi_i \wedge \varphi_{d+i},$$

wobei $\varphi_i = e_i^*$ die Dualbasis zur kanonischen Basis auf \mathbb{R}^{2d} ist. Konkret für $d = 2$ ergibt sich $\omega = \varphi_1 \wedge \varphi_3 + \varphi_2 \wedge \varphi_4$. Da $\omega \wedge \omega$ eine 4-Form auf \mathbb{R}^4 ist, muss $\omega \wedge \omega = c \cdot \det$ mit einer Konstanten c gelten.

Als Übung zum Umgang mit äußeren Produkten zeigen wir $c = -2$:

$$\begin{aligned} \omega \wedge \omega &= (\varphi_1 \wedge \varphi_3 + \varphi_2 \wedge \varphi_4) \wedge (\varphi_1 \wedge \varphi_3 + \varphi_2 \wedge \varphi_4) \\ &= \varphi_1 \wedge \varphi_3 \wedge \varphi_1 \wedge \varphi_3 + \varphi_1 \wedge \varphi_3 \wedge \varphi_2 \wedge \varphi_4 + \varphi_2 \wedge \varphi_4 \wedge \varphi_1 \wedge \varphi_3 \\ &\quad + \varphi_2 \wedge \varphi_4 \wedge \varphi_2 \wedge \varphi_4 \\ &= 0 + (-1)^1 \varphi_1 \wedge \varphi_2 \wedge \varphi_3 \wedge \varphi_4 + (-1)^3 \varphi_1 \wedge \varphi_2 \wedge \varphi_3 \wedge \varphi_4 + 0 \\ &= -2 \det. \end{aligned}$$

Die symplektische Form spielt eine zentrale Rolle in der klassischen Mechanik, wobei e_1, \dots, e_d den Orts- und e_{d+1}, \dots, e_{2d} den Impulskoordinaten entsprechen.

Ein Vektor $v \in \mathbb{R}^d$ definiert die $(d-1)$ -Form $\omega_v \in \wedge^{d-1}(\mathbb{R}^d)$,

$$\omega_v(w_1, \dots, w_{d-1}) := \det(v, w_1, \dots, w_{d-1}).$$

Zeigen Sie: Speziell für $d = 3$ ist dies $\omega_v = v_1 e_2^* \wedge e_3^* + v_2 e_3^* \wedge e_1^* + v_3 e_1^* \wedge e_2^*$, und für beliebige $v, w \in \mathbb{R}^3$ gilt

$$v^* \wedge w^* = \omega_{v \times w},$$

wobei \times in der letzten Zeile das *Kreuzprodukt* des \mathbb{R}^3 bezeichnet.

4.2 Differentialformen

Nach unseren Vorarbeiten im letzten Abschnitt führen wir nun Differentialformen als von Punkt zu Punkt in einer offenen Menge $U \subset \mathbb{R}^d$ variierende alternierende Multilinearformen ein. Die Idee ist, dass wir zu jedem $p \in U$ eine Form $\omega(p) \in \wedge^n(T_p^*(U))$ haben, wobei $T_p^*(U)$ der Dualraum des Tangentialraums⁹ $T_p(U)$ an U im Punkt $p \in U$ ist. Diese Sichtweise wird uns dann auch zu Differentialformen auf Untermannigfaltigkeiten des \mathbb{R}^d führen.

Definition 4.9. Sei $U \subset \mathbb{R}^d$ offen. Eine *Differentialform der Ordnung $n \in \mathbb{N}_0$* (oder *n -Form*) auf U ist eine Abbildung

$$\omega : U \rightarrow \bigcup_{p \in U} \wedge^n T_p^*(U)$$

⁹Die Elemente von $T_p^*(U)$ werden Kotangentialvektoren genannt.

mit $\omega(p) \in \bigwedge^n T_p^*(U)$ für alle $p \in U$.

Eine 0-Form ist nichts anderes als eine Abbildung $U \rightarrow \mathbb{R}$, da $\bigwedge^0 T_p^*(U) = \mathbb{R}$. Beispiele von 1-Formen erhalten wir durch Differentiation: Ist $f : U \rightarrow \mathbb{R}$ differenzierbar, so kennen wir aus Analysis 2 den Gradienten $\text{grad}(f)(p) = ((\partial_1 f)(p), \dots, (\partial_d f)(p)) \in \mathbb{R}^d$. Aufgefasst als Kovektor haben wir (am Punkt $p \in U$ angewendet auf den Tangentialvektor $v \in T_p(U)$)

$$df(p)(v) := \langle \text{grad}(f)(p), v \rangle,$$

was eine 1-Form definiert. Dies ist genau die Richtungsableitung $(\partial_v f)(p)$ von f am Punkt p in Richtung v (Analysis 2, Lemma 12.13).

Insbesondere können wir die Koordinatenfunktionen $f_i(x_1, \dots, x_d) := x_i$ betrachten, deren zugehörige 1-Formen wir kurz als dx_i bezeichnen¹⁰, also

$$dx_i(p) = e_i^*.$$

Da die alternierenden Multilinearformen $dx_{i_1}(p) \wedge \dots \wedge dx_{i_n}(p)$, $1 \leq i_1 < \dots < i_n \leq d$, eine Basis von $\bigwedge^n T_p^*(U)$ bilden, kann jede Differentialform ω der Ordnung n auf U als

$$\omega = \sum_{i_1 < \dots < i_n} \omega_{i_1, \dots, i_n} dx_{i_1} \wedge \dots \wedge dx_{i_n}$$

geschrieben werden mit eindeutig bestimmten Koeffizientenfunktionen $\omega_{i_1, \dots, i_n} : U \rightarrow \mathbb{R}$. Zum Beispiel gilt für eine Funktion (0-Form) $f : U \rightarrow \mathbb{R}$

$$df(p) = \sum_{i=1}^d \frac{\partial f}{\partial x_i}(p) dx_i,$$

wie man durch Vergleich mit $df(p)(v) = \langle \text{grad}(f)(p), v \rangle = \sum_{i=1}^d (\partial_i f)(p) e_i^*(v)$ sieht.

Wir nennen eine Differentialform ω stetig (bzw. k -mal stetig differenzierbar), wenn alle ihre Koeffizientenfunktionen stetig (bzw. k -mal stetig differenzierbar) sind. Wir bezeichnen den Raum aller glatten Differentialformen der Ordnung n auf U mit $\Omega^n(U)$, und mit $\Omega(U) := \bigoplus_{n=0}^{\infty} \Omega^n(U)$ den Raum aller glatten Differentialformen beliebiger Ordnung.

Da Differentialformen ausgewertet an einem Punkt p nichts anderes als alternierende Multilinearformen sind, können wir sie punktweise addieren, mit dem äußeren Produkt multiplizieren, und mit Funktionen multiplizieren:

$$\begin{aligned} (\omega + \varphi)(p) &:= \omega(p) + \varphi(p), & \omega, \varphi &\in \Omega^n(U), \\ (f\omega)(p) &:= f(p)\omega(p), & f &\in C^\infty(U) = \Omega^0(U), \omega \in \Omega^n(U), \\ (\omega \wedge \varphi)(p) &:= \omega(p) \wedge \varphi(p), & \omega &\in \Omega^n(U), \varphi \in \Omega^k(U). \end{aligned}$$

Diese Operationen lassen sich natürlich auch völlig analog für Differentialformen mit niedrigerem Regularitätsgrad (zB stetig statt glatt) definieren.

¹⁰Dies ist eine etwas unschöne Notation, eigentlich wäre etwas wie df_i besser. Ähnlich wie bei Ausdrücken der Form $\frac{df}{dx}$ ist diese Notation aber so weit verbreitet, dass wir sie hier auch benutzen werden.

Definition 4.10. Sei $U \subset \mathbb{R}^d$ offen. Die *äußere Ableitung* (oder *das Differential*) einer stetig differenzierbaren n -Form $\omega = \sum_{i_1 < \dots < i_n} \omega_{i_1, \dots, i_n} dx_{i_1} \wedge \dots \wedge dx_{i_n}$ auf U ist die durch

$$d\omega := \sum_{i_1 < \dots < i_n} d\omega_{i_1, \dots, i_n} \wedge dx_{i_1} \wedge \dots \wedge dx_{i_n}$$

definierte $(n + 1)$ -Form.

Beispiel 4.11.

a) Sei $f \in C^\infty(U) = \Omega^0(U)$ eine glatte Funktion. Dann ist ihre äußere Ableitung

$$df = \sum_{i=1}^d \frac{\partial f}{\partial x_i} dx_i.$$

b) Für eine 1-Form $\omega \in \Omega^1(\mathbb{R}^3)$ ist $\omega = \sum_{i=1}^3 \omega_i dx_i$ und

$$\begin{aligned} d\omega &= (d\omega_1) \wedge dx_1 + (d\omega_2) \wedge dx_2 + (d\omega_3) \wedge dx_3 \\ &= \left(\frac{\partial \omega_1}{\partial x_1} dx_1 + \frac{\partial \omega_1}{\partial x_2} dx_2 + \frac{\partial \omega_1}{\partial x_3} dx_3 \right) \wedge dx_1 \\ &\quad + \left(\frac{\partial \omega_2}{\partial x_1} dx_1 + \frac{\partial \omega_2}{\partial x_2} dx_2 + \frac{\partial \omega_2}{\partial x_3} dx_3 \right) \wedge dx_2 \\ &\quad + \left(\frac{\partial \omega_3}{\partial x_1} dx_1 + \frac{\partial \omega_3}{\partial x_2} dx_2 + \frac{\partial \omega_3}{\partial x_3} dx_3 \right) \wedge dx_3 \\ &= \left(\frac{\partial \omega_2}{\partial x_1} - \frac{\partial \omega_1}{\partial x_2} \right) dx_1 \wedge dx_2 + \left(\frac{\partial \omega_3}{\partial x_2} - \frac{\partial \omega_2}{\partial x_3} \right) dx_2 \wedge dx_3 + \left(\frac{\partial \omega_3}{\partial x_1} - \frac{\partial \omega_1}{\partial x_3} \right) dx_1 \wedge dx_3. \end{aligned}$$

c) Für eine 2-Form $\omega = \omega_{12} dx_1 \wedge dx_2 + \omega_{23} dx_2 \wedge dx_3 + \omega_{31} dx_3 \wedge dx_1 \in \Omega^2(\mathbb{R}^3)$ ist die äußere Ableitung

$$d\omega = \left(\frac{\partial \omega_{12}}{\partial x_3} + \frac{\partial \omega_{23}}{\partial x_1} + \frac{\partial \omega_{31}}{\partial x_2} \right) dx_1 \wedge dx_2 \wedge dx_3.$$

d) Für $\omega \in \Omega^d(\mathbb{R}^d)$ gilt $d\omega = 0$, da $d\omega \in \Omega^{d+1}(\mathbb{R}^d) = \{0\}$.

e) Für jedes $p \in \mathbb{R}^d$ hat der Vektorraum $\bigwedge^{d-1} T_p^* U$ der $(d-1)$ -Formen die Dimension $\binom{d}{d-1} = d$. Als Basis können wir die Formen

$$(-1)^{l-1} (dx_1 \wedge \dots \wedge \widehat{dx_l} \wedge \dots \wedge dx_d), \quad l \in \{1, \dots, d\},$$

wählen, wobei $\widehat{dx_l}$ andeutet, dass dieser Faktor ausgelassen wird. Eine $(d-1)$ -Form $\omega \in \Omega^{d-1}(U)$ lässt sich also stets in der Form

$$\omega = \sum_{l=1}^d (-1)^{l-1} \omega_l dx_1 \wedge \dots \wedge \widehat{dx_l} \wedge \dots \wedge dx_d$$

schreiben. Ist ω stetig differenzierbar, so erhalten wir als äußere Ableitung

$$\begin{aligned} d\omega &= \sum_{l=1}^d (-1)^{l-1} \sum_{i=1}^d \frac{\partial \omega_l}{\partial x_i} dx_i \wedge dx_1 \wedge \dots \wedge \widehat{dx_l} \wedge \dots \wedge dx_d \\ &= \sum_{l=1}^d \frac{\partial \omega_l}{\partial x_l} dx_1 \wedge \dots \wedge dx_d. \end{aligned}$$

Satz 4.12 (Eigenschaften der äußeren Ableitung). Sei $U \subset \mathbb{R}^d$ offen.

a) Die äußere Ableitung ist linear: Sind ω_1, ω_2 stetig differenzierbare n -Formen auf U und $\lambda_1, \lambda_2 \in \mathbb{R}$, so gilt

$$d(\lambda_1 \omega_1 + \lambda_2 \omega_2) = \lambda_1 d\omega_1 + \lambda_2 d\omega_2.$$

b) Die äußere Ableitung ist eine gradierte Derivation: Für eine stetig differenzierbare n -Form ω und eine stetig differenzierbare k -Form φ auf U gilt

$$d(\omega \wedge \varphi) = (d\omega) \wedge \varphi + (-1)^n \omega \wedge (d\varphi).$$

c) Für jede zweimal stetig differenzierbare n -Form ω auf U gilt

$$d(d\omega) = 0.$$

Kurz gesagt: $d^2 = 0$.

Beweis. a) ergibt sich direkt aus der Definition von d .

b) Diese Eigenschaft beruht auf der Produktregel der Differentialrechnung. Wir betrachten zunächst $n = k = 0$, dh $\omega, \varphi : U \rightarrow \mathbb{R}$ sind stetig differenzierbare Funktionen. Die Produktregel für die partiellen Ableitungen, $\partial_i(\omega\varphi) = (\partial_i\omega)\varphi + \omega(\partial_i\varphi)$ impliziert

$$d(\omega\varphi) = \sum_{i=1}^d ((\partial_i\omega)\varphi + \omega(\partial_i\varphi)) dx_i = d\omega \wedge \varphi + (-1)^0 \omega \wedge d\varphi,$$

wie behauptet. Für den allgemeinen Fall dürfen wir uns auf $\omega = \omega_{i_1, \dots, i_n} dx_{i_1} \wedge \dots \wedge dx_{i_n}$ und $\varphi = \varphi_{j_1, \dots, j_k} dx_{j_1} \wedge \dots \wedge dx_{j_k}$ mit sortierten Multiindizes beschränken. Wir schreiben zur Abkürzung $\omega_I := \omega_{i_1, \dots, i_n}$, $\varphi_J := \varphi_{j_1, \dots, j_k}$ und $dx_I := dx_{i_1} \wedge \dots \wedge dx_{i_n}$, $dx_J := dx_{j_1} \wedge \dots \wedge dx_{j_k}$ und erhalten

$$\begin{aligned} d(\omega \wedge \varphi) &= d(\omega_I \varphi_J) \wedge dx_I \wedge dx_J \\ &= ((d\omega_I)\varphi_J + \omega_I(d\varphi_J)) \wedge dx_I \wedge dx_J \\ &= d\omega_I \wedge dx_I \wedge \varphi_J dx_J + (-1)^n \omega_I dx_I \wedge d\varphi_J \wedge dx_J \\ &= (d\omega) \wedge \varphi + (-1)^n \omega \wedge (d\varphi). \end{aligned}$$

c) Diese Eigenschaft beruht auf dem Satz von Schwarz (Analysis 2, Satz 12.25: Für C^2 -Funktionen gilt $\partial_i \partial_j f = \partial_j \partial_i f$). Wir behandeln wieder zuerst den Fall einer Funktion (0-

Form). Per Definition von d gilt

$$d(df) = d\left(\sum_{i=1}^d \frac{\partial f}{\partial x_i} dx_i\right) = \sum_{i=1}^d d\left(\frac{\partial f}{\partial x_i}\right) \wedge dx_i = \sum_{i=1}^d \sum_{j=1}^d \frac{\partial^2 f}{\partial x_j \partial x_i} dx_j \wedge dx_i.$$

Wegen $f \in C^2(U)$ gilt $\frac{\partial^2 f}{\partial x_j \partial x_i} dx_j \wedge dx_i = \frac{\partial^2 f}{\partial x_i \partial x_j} dx_j \wedge dx_i = -\frac{\partial^2 f}{\partial x_i \partial x_j} dx_i \wedge dx_j$. In der Summe heben sich die Terme (i, j) und (j, i) also weg, dh $d^2 f = 0$. Insbesondere gilt $d^2 x_i = 0$.

Für eine n -Form $\sum_{i_1 < \dots < i_n} f_{i_1, \dots, i_n} dx_{i_1} \wedge \dots \wedge dx_{i_n}$ gilt mit der Produktregel aus Teil b)

$$\begin{aligned} d^2 \omega &= d \sum_{i_1 < \dots < i_n} (df_{i_1, \dots, i_n}) \wedge dx_{i_1} \wedge \dots \wedge dx_{i_n} \\ &= \sum_{i_1 < \dots < i_n} (d^2 f_{i_1, \dots, i_n}) \wedge dx_{i_1} \wedge \dots \wedge dx_{i_n} \\ &\quad + \sum_{l=1}^n (-1)^l \sum_{i_1 < \dots < i_n} (df_{i_1, \dots, i_n}) \wedge dx_{i_1} \wedge \dots \wedge d^2 x_{i_l} \wedge \dots \wedge dx_{i_n} = 0. \end{aligned}$$

□

Die äußere Ableitung ist auch in der klassischen Vektoranalysis (grad, div, rot) auf \mathbb{R}^3 nützlich. Um diesen Zusammenhang zu sehen, betrachten wir eine offene Menge $U \subset \mathbb{R}^3$ mit einer skalaren stetig differenzierbaren Funktion $f : U \rightarrow \mathbb{R}$ und einem stetig differenzierbaren Vektorfeld $v : U \rightarrow \mathbb{R}^3$. Zu diesen Abbildungen haben wir die aus Analysis 2 bekannten Differentialausdrücke

$$\begin{aligned} \text{grad } f &:= (\partial_1 f, \partial_2 f, \partial_3 f), \\ \text{div } v &:= \frac{\partial v_1}{\partial x_1} + \frac{\partial v_2}{\partial x_2} + \frac{\partial v_3}{\partial x_3}, \\ \text{rot } v &:= \left(\frac{\partial v_3}{\partial x_2} - \frac{\partial v_2}{\partial x_3}, \frac{\partial v_1}{\partial x_3} - \frac{\partial v_3}{\partial x_1}, \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2} \right). \end{aligned}$$

Divergenz und Gradient können auch in beliebiger Dimension $d \in \mathbb{N}$ definiert werden.

Weiterhin betrachten wir das “vektorielle Streckenelement”

$$d\vec{s} := (dx_1, dx_2, dx_3),$$

das “vektorielle Flächenelement”

$$d\vec{S} := (dx_2 \wedge dx_3, dx_3 \wedge dx_1, dx_1 \wedge dx_2),$$

und das “Volumenelement”

$$dV := dx_1 \wedge dx_2 \wedge dx_3.$$

Auf U definierte Differentialformen können so wie folgt geschrieben werden (mit dem als Punkt notierten Euklidischen Skalarprodukt):

$$\begin{aligned} \text{Ordnung 0:} & \quad f \\ \text{Ordnung 1:} & \quad v \cdot d\vec{s} := v_1 dx_1 + v_2 dx_2 + v_3 dx_3 \\ \text{Ordnung 2:} & \quad v \cdot d\vec{S} := v_1(dx_2 \wedge dx_3) + v_2(dx_3 \wedge dx_1) + v_3(dx_1 \wedge dx_2), \\ \text{Ordnung 3:} & \quad f dV. \end{aligned}$$

Zeigen Sie:

- Beliebige stetig differenzierbare Differentialformen der Ordnung 0, 1, 2, 3 lassen sich in der oben angegebenen Art und Weise durch geeignetes f bzw. v schreiben.
- Für die Differentiale gilt

$$\begin{aligned}df &= \operatorname{grad} f \cdot d\vec{s}, \\d(v \cdot d\vec{s}) &= \operatorname{rot}(v) \cdot d\vec{S}, \\d(v \cdot d\vec{S}) &= \operatorname{div}(v) dV.\end{aligned}$$

- Zeigen Sie per äußerer Ableitung für zweimal stetig differenzierbares f und v

$$\operatorname{rot}(\operatorname{grad} f) = 0, \quad \operatorname{div}(\operatorname{rot} v) = 0.$$

Definition 4.13. Sei $U \subset \mathbb{R}^d$ offen.

- Eine stetig differenzierbare n -Form ω in U heißt *geschlossen*, wenn

$$d\omega = 0.$$

- Für $n \geq 1$ heißt eine stetige n -Form ω in U *exakt*, falls es eine stetig differenzierbare $(n-1)$ -Form ψ in U gibt mit

$$\omega = d\psi.$$

Wegen $d^2 = 0$ ist jede exakte Form geschlossen. Im Allgemeinen müssen geschlossene Formen nicht exakt sein. Das *Poincarésche Lemma* (das wir hier nicht im Detail besprechen) besagt, dass für sternförmiges U jede geschlossene Form exakt ist. Für 1-Formen lässt sich aber auf Grundlage von Satz 12.31 aus Analysis 2 leicht zeigen, dass Geschlossenheit Exaktheit impliziert: Mit einer 1-Form $\omega = \sum_i v_i dx_i$ können wir das Vektorfeld $v : U \rightarrow \mathbb{R}^d$, $x \mapsto (v_1(x), \dots, v_d(x))$ assoziieren, und $d\omega = 0$ ist äquivalent zu $\partial_i v_j = \partial_j v_i$ für alle i, j (Übung). Für sternförmiges U wissen wir, dass ein solches Vektorfeld ein Gradientenfeld ist, d.h. $v = \operatorname{grad} f$ für ein $f \in C^1(U)$. Die Funktion f ist eine 0-Form auf U , und $v = \operatorname{grad} f$ ist äquivalent zu $\omega = df$ (Übung).

Als letzten Aspekt von Differentialformen, den wir für die Integration benötigen, wenden wir uns nun dem Rücktransport von Differentialformen zu.

Definition 4.14. Seien $U \subset \mathbb{R}^d$ und $V \subset \mathbb{R}^k$ offen, $\varphi = (\varphi_1, \dots, \varphi_k) : U \rightarrow V$ stetig differenzierbar, und

$$\omega = \sum_{i_1 < \dots < i_n} \omega_{i_1, \dots, i_n} dx_{i_1} \wedge \dots \wedge dx_{i_n} \in \Omega^n(V)$$

eine n -Form auf V . Der Rücktransport ("pullback") $\varphi^*\omega$ ist die Form

$$\varphi^*\omega := \sum_{i_1 < \dots < i_n} (\omega_{i_1, \dots, i_n} \circ \varphi) d\varphi_{i_1} \wedge \dots \wedge d\varphi_{i_n} \in \Omega^n(U).$$

Im Gegensatz zum beim Bildmaß verwendeten pushforward ist hier der pullback die richtige Operation, da Differenzierbarkeit erhalten werden soll, φ aber nicht als invertiert oder Diffeomorphismus vorausgesetzt ist. Wir werden den pullback später dazu verwenden, Differentialformen auf Untermannigfaltigkeiten zu integrieren.

Beispiel 4.15.

- a) Falls ω eine 1-Form ist, also $\omega = \sum_{i=1}^k f_i dx_i$, so können wir $\varphi^*\omega$ wie folgt beschreiben. Seien y_1, \dots, y_d die kanonischen Koordinatenfunktionen von $\mathbb{R}^d \supset U$, also $d\varphi_i = \sum_{j=1}^d \frac{\partial \varphi_i}{\partial y_j} dy_j$. Dann ist

$$\varphi^*\omega = \sum_{i=1}^k \sum_{j=1}^d (f_i \circ \varphi) \frac{\partial \varphi_i}{\partial y_j} dy_j = \sum_{j=1}^d g_j dy_j, \quad g_j := \sum_{i=1}^k (f_i \circ \varphi) \frac{\partial \varphi_i}{\partial y_j}.$$

Mit der Jacobimatrix $D\varphi$ von φ lassen sich die Funktionen $g = (g_1, \dots, g_d)$ (als Zeilenvektor) als

$$g = (f \circ \varphi) D\varphi$$

schreiben.

- b) Wir betrachten die Polarkoordinaten

$$\varphi : \mathbb{R}_+ \times (-\pi, \pi) \rightarrow \mathbb{R}^2, \quad \varphi(r, \theta) = (r \cos \theta, r \sin \theta)$$

und die 2-Form $\omega = f dx_1 \wedge dx_2$. Mit der in Polarkoordinaten geschriebenen Funktion $\hat{f} := f \circ \varphi$ ergibt sich

$$\begin{aligned} d\varphi_1(r, \theta) &= \cos \theta dr - r \sin \theta d\theta, & d\varphi_2(r, \theta) &= \sin \theta dr + r \cos \theta d\theta, \\ (\varphi^*\omega)(r, \theta) &= (\hat{f} d\varphi_1 \wedge d\varphi_2)(r, \theta) \\ &= f(r \cos \theta, r \sin \theta) \cdot (r \cos^2 \theta dr \wedge d\theta - r \sin^2 \theta d\theta \wedge dr) \\ &= r f(r \cos \theta, r \sin \theta) dr \wedge d\theta. \end{aligned}$$

Zeigen Sie: Im Falle einer d -Form $\omega = f dx_1 \wedge \dots \wedge dx_d$ ist der pullback

$$\varphi^*\omega = (f \circ \varphi) \cdot \det(D\varphi) dx_1 \wedge \dots \wedge dx_d.$$

Satz 4.16 (Eigenschaften des Rücktransports). Seien $U \subset \mathbb{R}^d$ und $V \subset \mathbb{R}^k$ offen, $\varphi : U \rightarrow V$ stetig differenzierbar, und $\omega, \omega_1, \omega_2 \in \Omega^n(V)$ Differentialformen der Ordnung n in V sowie $\eta \in \Omega^l(V)$ eine l -Form in V , und $\lambda, \mu \in \mathbb{R}$. Dann gilt

- a) $\varphi^*(\lambda\omega_1 + \mu\omega_2) = \lambda\varphi^*\omega_1 + \mu\varphi^*\omega_2$ (Linearität des Rücktransports),

$$b) \varphi^*(\omega \wedge \eta) = (\varphi^*\omega) \wedge (\varphi^*\eta),$$

c) Ist ω stetig differenzierbar und φ zweimal stetig differenzierbar, so gilt

$$d(\varphi^*\omega) = \varphi^*(d\omega).$$

d) Ist $\psi : W \subset \mathbb{R}^p \rightarrow U$ stetig differenzierbar, so gilt

$$\psi^*(\varphi^*\omega) = (\varphi \circ \psi)^*\omega.$$

Beweis. a) und b) ergeben sich direkt aus der Definition des Rücktransports, siehe auch Hausaufgabe H11.1.

Aufgrund der Linearität genügt es, in c) und d) $\omega = f dx_{i_1} \wedge \dots \wedge dx_{i_n}$ zu betrachten. Mit der Kettenregel gilt (die kanonischen Koordinaten in \mathbb{R}^d nennen wir wieder y_1, \dots, y_d)

$$d(f \circ \varphi) = \sum_{i=1}^d \frac{\partial(f \circ \varphi)}{\partial y_i} dy_i = \sum_{i=1}^d \sum_{j=1}^k \left(\frac{\partial f}{\partial x_j} \circ \varphi \right) \frac{\partial \varphi_j}{\partial y_i} dy_i = \sum_{j=1}^k \left(\frac{\partial f}{\partial x_j} \circ \varphi \right) d\varphi_j$$

Damit haben wir

$$\begin{aligned} d(\varphi^*\omega) &= d((f \circ \varphi) d\varphi_{i_1} \wedge \dots \wedge d\varphi_{i_n}) \\ &= d(f \circ \varphi) \wedge d\varphi_{i_1} \wedge \dots \wedge d\varphi_{i_n} \\ &= \sum_{j=1}^k \left(\frac{\partial f}{\partial x_j} \circ \varphi \right) d\varphi_j \wedge d\varphi_{i_1} \wedge \dots \wedge d\varphi_{i_n} \\ &= \varphi^* \left(\sum_{j=1}^k \frac{\partial f}{\partial x_j} dx_j \wedge dx_{i_1} \wedge \dots \wedge dx_{i_n} \right) = \varphi^*(d\omega). \end{aligned}$$

d) Die Komponenten der verketteten Abbildung $\varphi \circ \psi$ sind $(\varphi \circ \psi)_i = \varphi_i \circ \psi$, also gilt nach c) $\psi^*(d\varphi_i) = d(\psi^*\varphi_i) = d(\varphi_i \circ \psi) = d(\varphi \circ \psi)_i$. Damit erhalten wir mit b)

$$\begin{aligned} \psi^*(\varphi^*\omega) &= \psi^*((f \circ \varphi) d\varphi_{i_1} \wedge \dots \wedge d\varphi_{i_n}) \\ &= (f \circ \varphi \circ \psi) \cdot \psi^*(d\varphi_{i_1}) \wedge \dots \wedge \psi^*(d\varphi_{i_n}) \\ &= (f \circ \varphi \circ \psi) \cdot d(\varphi \circ \psi)_{i_1} \wedge \dots \wedge d(\varphi \circ \psi)_{i_n} = (\varphi \circ \psi)^*\omega. \end{aligned}$$

□

Unsere Definition der äußeren Ableitung $d\omega$ ist direkt in einer Darstellung der zu differenzierenden Form ω gegeben, es ist also eine Definition "in Koordinaten". Es gibt aber auch eine "koordinatenfreie Definition": Es gibt genau eine Sequenz linear Abbildungen

$$0 \xrightarrow{d} \Omega^0 M \xrightarrow{d} \Omega^1 M \xrightarrow{d} \Omega^2 M \xrightarrow{d} \dots$$

(für eine Mannigfaltigkeit M , insbesondere für eine offene Teilmenge $U \subset \mathbb{R}^d$), so dass die folgenden drei Bedingungen erfüllt sind:

a) Für $f \in \Omega^0 M$ ist df das Differential von f .

b) $d \circ d = 0$.

$$c) d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^{\deg(\omega)} \omega \wedge d\eta.$$

Als Ausblick auf die Kohomologietheorie erwähnen wir, dass $d \circ d = 0$ bedeutet, dass das Bild von $d : \Omega^{n-1}M \rightarrow \Omega^n M$ im Kern von $d : \Omega^n M \rightarrow \Omega^{n+1}M$ enthalten ist. Wir können also den Quotientenvektorraum

$$H^n M := \ker(d : \Omega^n M \rightarrow \Omega^{n+1}M) / \text{Im}(d : \Omega^{n-1}M \rightarrow \Omega^n M)$$

betrachten, der *n-te de Rham Kohomologiegruppe von M* heißt. In dieser Formulierung besagt das Poincarésche Lemma also $H^n M = \{0\}$ für $n \geq 1$, falls M sternförmig oder *einfach zusammenhängend* ist. Das bedeutet anschaulich, dass jede geschlossene Kurve in M auf einen Punkt zusammengezogen werden kann, M also “keine Löcher” hat. Diese Objekte werden in der algebraischen Topologie studiert; man kann zeigen, dass $H^n M$ nur von topologischen Eigenschaften von M abhängt.

4.3 Integration von Differentialformen über orientierte Untermannigfaltigkeiten

Wir kommen nun zur Integration von Differentialformen und beginnen mit Formen maximaler Ordnung.

Definition 4.17. Sei $U \subset \mathbb{R}^d$ offen und $A \subset U$ eine Borelmenge. Eine d -Form $\omega = f dx_1 \wedge \dots \wedge dx_d$ in U heißt über A integrierbar, wenn $f|_A$ Lebesgue-integrierbar ist. In diesem Fall definieren wir

$$\int_A \omega := \int_A f(x) d\lambda_d(x) = \int_A f(x) dx.$$

Auf der rechten Seite bezeichnet “ dx ” Lebesgue-Integration bzgl der Variable x , keine Differentialform. Diese Definition übersetzt also gewissermaßen die 1-Formen dx_i in ihre bisherige Bedeutung bzgl Integralen,

$$\int_A f dx_1 \wedge \dots \wedge dx_n = \int_A f(x) dx_1 \dots dx_n.$$

Sie können sich diese Definition so merken, dass $dx_1 \wedge \dots \wedge dx_d = \det$ die *Standardvolumenform* auf \mathbb{R}^d ist (eine d -Form), und f angibt, wie das Volumen lokal skaliert wird.

Unter der obigen Übersetzung wird der pullback zur Transformationsformel. Dazu zuerst eine Vorbemerkung: Seien $U, V \subset \mathbb{R}^d$ offen und $\varphi : U \rightarrow V$ ein Diffeomorphismus (wie bei der Transformationsformel, Satz 2.51). Dann gilt $\det(D\varphi)(x) \neq 0$ für alle $x \in U$ (Analysis 2, Lemma 13.7). Falls U *zusammenhängend* ist (d.h. zu jedem Paar $x, y \in U$ gibt es einen stetigen Weg von x nach y , also eine stetige Abbildung $\gamma : [0, 1] \rightarrow U$ mit $\gamma(0) = x$ und $\gamma(1) = y$), so muss das Vorzeichen von $\det(D\varphi)(x)$ konstant sein, da $x \mapsto \det(D\varphi)(x)$ stetig und nicht verschwindend ist (Zwischenwertsatz). Ist U nicht zusammenhängend, so muss das nicht so sein.

Satz 4.18. Seien $U, V \subset \mathbb{R}^d$ offen und $\varphi : U \rightarrow V$ ein Diffeomorphismus mit konstantem Vorzeichen ε von $\det(D\varphi)(x)$, $x \in U$, und $A \subset V$ eine Borelmenge. Dann ist für jede über A integrierbare d -Form ω auf V die Form $\varphi^* \omega$ über die Borelmenge $\varphi^{-1}(A) \subset U$

integrierbar, und es gilt

$$\int_{\varphi^{-1}(A)} \varphi^* \omega = \varepsilon \int_A \omega.$$

Beweis. Wir können ω in der Form $\omega = f dx_1 \wedge \dots \wedge dx_d$ schreiben. Dann ist der pullback

$$\varphi^* \omega = (f \circ \varphi) \cdot \det(D\varphi) dx_1 \wedge \dots \wedge dx_d$$

wie in einer Übung gezeigt wurde (Präsenzübungen Blatt 12). Also gilt per Transformationsformel

$$\begin{aligned} \int_{\varphi^{-1}(A)} \varphi^* \omega &= \int_{\varphi^{-1}(A)} (f \circ \varphi) \det(D\varphi) dx_1 \wedge \dots \wedge dx_d \\ &= \int_{\varphi^{-1}(A)} f(\varphi(x)) \det(D\varphi)(x) dx_1 \dots dx_d \\ &= \varepsilon \int_{\varphi^{-1}(A)} f(\varphi(x)) |\det(D\varphi)(x)| dx \\ &= \varepsilon \int_A f(y) dy \\ &= \varepsilon \int_A \omega. \end{aligned}$$

□

Das in dieser Formel auftretende Vorzeichen ε deutet bereits an, dass wir auf dem richtigen Weg zur Verallgemeinerung des Hauptsatzes der Differential- und Integralrechnung sind.

Wir betrachten nun die noch sehr viel interessantere Frage der Integration einer n -Form über eine n -dimensionale Untermannigfaltigkeit $M \subset \mathbb{R}^d$, also $n \in \{0, \dots, d\}$. Zuerst erinnern wir uns an die Definition (Analysis 2, Def. 13.21).

Definition 4.19. Es seien $n, d \in \mathbb{N}$ mit $n \leq d$. Eine n -dimensionale (differenzierbare) Untermannigfaltigkeit des \mathbb{R}^d ist eine nicht leere Teilmenge $M \subset \mathbb{R}^d$ mit der folgenden Eigenschaft: Für jeden Punkt $p \in M$ gibt es eine offene Umgebung $U \subset \mathbb{R}^d$ von p und eine offene Teilmenge $V \subset \mathbb{R}^d$ und einen Diffeomorphismus $\varphi : U \rightarrow V$ mit

$$\varphi(M \cap U) = V \cap (\mathbb{R}^n \times \{0\}).$$

Hier ist $\{0\}$ als Teilmenge von \mathbb{R}^{d-n} aufzufassen.

Wir werden Untermannigfaltigkeiten häufig lokal durch Koordinaten (Parameter) beschreiben, also durch Abbildungen von Teilmengen von \mathbb{R}^k nach M . Dazu ist folgender Satz relevant.

Satz 4.20 (Karten und Kartenwechsel).

- a) Eine Teilmenge $M \subset \mathbb{R}^d$ ist genau dann eine n -dimensionale Untermannigfaltigkeit des \mathbb{R}^d , wenn es zu jedem Punkt $p \in M$ eine in M offene Umgebung $p \in U \subset M$ gibt^a, eine offene Menge $V \subset \mathbb{R}^n$, und eine stetig differenzierbare Abbildung $h :$

$V \rightarrow \mathbb{R}^d$, die V homöomorph^b auf U abbildet und

$$\text{rang}(Dh)(x) = n, \quad x \in V,$$

erfüllt.

- b) Sei $M \subset \mathbb{R}^d$ eine n -dimensionale Untermannigfaltigkeit und seien $h_i : V_i \rightarrow U_i \subset M$, $i = 1, 2$, zwei Karten mit $U := U_1 \cap U_2 \neq \emptyset$. Dann sind $W_i := h_i^{-1}(U) \subset V_i$ offen und

$$h_2^{-1} \circ h_1 : W_1 \rightarrow W_2$$

ein Diffeomorphismus.

^a $U \subset M \subset \mathbb{R}^d$ heißt in M offen, wenn es zu jedem $x \in U$ eine offene Umgebung $p \in O \subset \mathbb{R}^d$ gibt, so dass $O \cap M \subset U$

^bErinnerung: Homöomorph heißt bijektiv und stetig mit stetiger Umkehrabbildung.

Wir werden diesen Satz hier nicht beweisen (siehe Forster 3, Satz 4&5, Seite 163f), um nicht zu weit von unserem eigentlichen Integrationsthema abzuschweifen.

Sie sollten sich h als eine Kartenabbildung vorstellen, die ein Stück der n -dimensionalen Mannigfaltigkeit durch die n Koordinaten $(x_1, \dots, x_n) \in V$ parametrisiert. Eine Untermannigfaltigkeit kommt nicht mit eindeutig bestimmten Kartenabbildungen, aber man kann glatt zwischen Karten mit überlappenden Kartengebieten hin- und herwechseln. Um die Analogie zur Kartographie noch etwas weiter zu treiben, definieren wir einen *Atlas* einer Untermannigfaltigkeit als eine Menge $\mathcal{A} = \{h_i : V_i \rightarrow U_i : i \in \mathcal{I}\}$ von Karten, die die ganze Mannigfaltigkeit abdecken, dh

$$\mathcal{A} = \bigcup_{i \in \mathcal{I}} U_i.$$

Die Indexmenge kann hierbei beliebig sein. Für kompaktes M genügt aber immer eine endliche Indexmenge.

Beispiel 4.21.

- a) Wir betrachten einen zweidimensionalen Zylinder

$$M := \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = r^2\},$$

wobei $r > 0$ ein vorgegebener Radius ist. Mit dem Satz vom regulären Wert (angewandt auf die Funktion $(x, y, z) \mapsto x^2 + y^2 - r^2$) sieht man leicht, dass M eine zweidimensionale Untermannigfaltigkeit des \mathbb{R}^3 ist.

Als Karten betrachten wir die beiden Abbildungen

$$h_1 : V_1 := (-\pi, \pi) \times \mathbb{R} \rightarrow M, \quad h_1(\alpha, z) := \begin{pmatrix} r \cos \alpha \\ r \sin \alpha \\ z \end{pmatrix},$$

$$h_2 : V_2 := (-\pi, \pi) \times \mathbb{R}_+ \rightarrow M, \quad h_2(\alpha, z) := \begin{pmatrix} r \cos \alpha \\ r \sin \alpha \\ z^2 \end{pmatrix}.$$

Für beide Abbildungen prüft man leicht nach, dass sie Karten sind. Das Bild $U_1 = h_1(V_1)$ der ersten Karte besteht aus dem ganzen Zylinder mit Ausnahme aller Punkte $(x, y, z) \in M$ mit $x < 0$ und $y = 0$. Das Bild $U_2 = h_2(V_2)$ der zweiten Karte besteht aus den Punkten $(x, y, z) \in M$ mit $z > 0$ außer denen, die $x < 0, y = 0$ erfüllen.

Für einen Kartenwechsel betrachten wir $U := U_1 \cap U_2 = U_2$ sowie $W_1 = h_1^{-1}(U) = (-\pi, \pi) \times \mathbb{R}_+$ und $W_2 = V_2$ sowie

$$h_2^{-1} \circ h_1 : (-\pi, \pi) \times \mathbb{R}_+ \rightarrow (-\pi, \pi) \times \mathbb{R}_+, \quad (\alpha, z) \mapsto (\alpha, \sqrt{z}),$$

was ein Diffeomorphismus ist.

- b) Als ein weiteres Beispiel betrachten wir eine Kreiskurve mit Radius 1, nämlich $M := \{(x, y) : x^2 + y^2 = 1\}$. Dies ist eine eindimensionale Untermannigfaltigkeit des \mathbb{R}^2 .

Zwei Karten von M sind:

$$h_1 : (-\pi, \pi) \rightarrow M, \quad h_1(\alpha) = \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix},$$

$$h_2 : (-1, 1) \rightarrow M, \quad h_2(x) = \begin{pmatrix} x \\ \sqrt{1-x^2} \end{pmatrix}.$$

Während das Bild von h_1 der ganze Kreis mit Ausnahme von $(-1, 0)$ ist, ist das Bild von h_2 nur der obere offene Halbkreis. Der Kartenwechsel ist hier

$$h_2^{-1} \circ h_1 : (0, \pi) \rightarrow (-1, 1), \quad \alpha \mapsto \cos \alpha,$$

ein Diffeomorphismus ($\cos' \alpha = -\sin \alpha \neq 0$ für $0 < \alpha < \pi$).

Eine weitere Karte wäre

$$h_3 : (-1, 1) \rightarrow M, \quad h_3(y) := \begin{pmatrix} -\sqrt{1-y^2} \\ -y \end{pmatrix}.$$

Diese Karte deckt den offenen linken Halbkreis ab. Die Mengen $\{h_1, h_3\}$ oder $\{h_1, h_2, h_3\}$ sind also Atlanten von M .

Während Karten für das konkrete Rechnen mit einer Untermannigfaltigkeit M sehr wichtig sind, sind sie nicht kanonisch durch M gegeben. Für "koordinatenunabhängige" Definitionen, die mit Hilfe von Karten gegeben werden, werden wir also stets prüfen müssen, dass sie

nicht von der Wahl der Karten abhängen.

Wir kommen nun zur Integration zurück und betrachten zunächst den übersichtlichen Fall, dass $A \subset M$ eine kompakte Teilmenge einer n -dimensionalen Untermannigfaltigkeit M des \mathbb{R}^d ist, die in ein Kartengebiet passt, dh es gibt eine Kartenabbildung $h : V \subset \mathbb{R}^n \rightarrow U \subset M$ mit $A \subset U$. Weiterhin sei ω eine stetige n -Form, die auf einer offenen Umgebung O (mit $U \subset O \subset \mathbb{R}^d$) definiert ist. In diesem Kontext wollen wir ein Integral $\int_A \omega$ definieren. Geleitet von unseren bisherigen Ergebnissen zu d -Formen setzen wir

$$\int_A \omega := \int_{h^{-1}(A)} h^* \omega.$$

Damit diese Definition sinnvoll ist, sollte sie aber nicht von der Wahl der Karte h abhängen. Ist $h_2 : V_2 \subset \mathbb{R}^n \rightarrow U_2$ eine weitere Karte mit $A \subset U_2$, so dürfen wir $U_1 = U_2 =: U$ annehmen. Es gibt dann also einen Diffeomorphismus

$$\psi : h^{-1}(U) \rightarrow h_2^{-1}(U), \quad \psi = h_2^{-1} \circ h.$$

Wir nehmen an, dass $h^{-1}(U)$ zusammenhängend ist. Dann hat $\det(D\psi)(x)$ konstantes Vorzeichen $\varepsilon = \pm 1$ für alle $x \in h^{-1}(U)$. Mit Hilfe von Satz 4.16 d) und Satz 4.18 erhalten wir

$$\int_{h^{-1}(A)} h^* \omega = \int_{(h_2 \circ \psi)^{-1}(A)} (h_2 \circ \psi)^* \omega = \int_{\psi^{-1}(h_2^{-1}(A))} \psi^*(h_2^* \omega) = \varepsilon \int_{h_2^{-1}(A)} h_2^* \omega.$$

Dieses Ergebnis sagt uns, dass unsere Definition des Integrals $\int_A \omega$ zwar nicht sehr stark, aber doch etwas von der Wahl der Karte h bzw h_2 abhängt: Das Integral kann sich im Vorzeichen ändern, je nachdem, welche Karte verwendet wird.

Diese Beobachtung motiviert die folgende Definition.

Definition 4.22.

- a) Seien $U, V \subset \mathbb{R}^n$ offen. Eine invertierbare stetig differenzierbare Abbildung $\varphi : U \rightarrow V$ heißt *orientierungstreu*, falls $\det(D\varphi)(x) > 0$ für alle $x \in U$, und *orientierungsumkehrend*, falls $\det(D\varphi)(x) < 0$ für alle $x \in U$.
- b) Sei M eine n -dimensionale Untermannigfaltigkeit von \mathbb{R}^d . Zwei Karten $h_i : V_i \rightarrow U_i, i = 1, 2$, mit überlappenden Kartengebieten, $U_1 \cap U_2 \neq \emptyset$, heißen *gleich orientiert*, falls ihr Kartenwechsel (siehe Satz 4.20 b)) orientierungstreu ist.
- c) Sei M eine n -dimensionale Untermannigfaltigkeit von \mathbb{R}^d . Ein Atlas von M heißt *orientiert*, falls alle Kartenwechsel gleich orientiert sind. M heißt *orientierbar*, falls M einen orientierten Atlas hat. Eine *orientierte* Untermannigfaltigkeit ist eine Untermannigfaltigkeit mit einem orientierten Atlas. Eine *Orientierung* ist eine Äquivalenzklasse von orientierten Atlanten von M bzgl der Äquivalenzrelation $\mathcal{A}_1 \sim \mathcal{A}_2 \Leftrightarrow$ jede Karte aus \mathcal{A}_1 ist zu jeder Karte aus \mathcal{A}_2 gleich orientiert.
- d) Auf einer orientierten Untermannigfaltigkeit heißt eine Karte *orientierungserhaltend*, wenn sie zu allen Karten eines die Orientierung definierenden Atlas orientierungstreu ist.

Beispiel 4.23.

- a) Jede Untermannigfaltigkeit, die von einer einzigen Karte abgedeckt werden kann, ist orientierbar, da sie einen aus nur einer Karte bestehenden Atlas besitzt.
- b) Für eine offene Teilmenge $U \subset \mathbb{R}^d$ (also einer d -dimensionalen Untermannigfaltigkeit von \mathbb{R}^d) ist $\text{id}_U : U \rightarrow U$ eine Karte, die den Atlas $\{\text{id}_U\}$ bildet. Die *kanonische Orientierung von U* ist die durch diesen Atlas gegebene Orientierung.
- c) Für einen n -dimensionalen Untervektorraum M von \mathbb{R}^d (dies ist eine n -dimensionale Untermannigfaltigkeit) mit geordneter Basis (v_1, \dots, v_n) ist die durch $h : \mathbb{R}^n \rightarrow M$, $e_i \mapsto v_i$, $i = 1, \dots, n$ definierte lineare Abbildung eine Karte. Zwei solche Karten sind orientierungstreu genau dann, wenn ihre Determinanten $\det(v_1, \dots, v_n)$ das gleiche Vorzeichen haben.
- d) Der Kreis $S^1 \subset \mathbb{R}^2$ ist orientierbar, da der in Beispiel 4.21 angegebene Atlas $\{h_1, h_3\}$ orientiert ist. (Der Kartenwechsel ist $-\sin : (-\pi, -\frac{\pi}{2}) \cup (\frac{\pi}{2}, \pi) \rightarrow (-1, 1)$, mit positiver Ableitung $-\cos$).

Die beiden Karten h_1, h_2 sind hingegen nicht gleich orientiert: Hier ist der Kartenwechsel der Diffeomorphismus $\cos : (0, \pi) \rightarrow (-1, 1)$ mit Ableitung $-\sin$, die auf $(0, \pi)$ *negative* Werte annimmt. Das liegt daran, dass die Karte h_1 den Kreis im Gegenuhrzeigersinn und die Karte h_2 den Kreis im Uhrzeigersinn durchläuft. Für eindimensionale Mannigfaltigkeiten entspricht die Wahl einer Orientierung der Wahl einer Durchlaufrichtung. Wir bemerken ohne Beweis, dass alle eindimensionalen Mannigfaltigkeiten orientierbar sind.

- e) Nicht jede Untermannigfaltigkeit ist orientierbar, ein Gegenbeispiel ist das Möbiusband (siehe Hausaufgaben).

Wir vereinbaren noch, dass eine Orientierung eines Punktes $\{x\}$ (also einer 0-dimensionalen Untermannigfaltigkeit des \mathbb{R}^d) die Wahl eines Vorzeichens ± 1 ist.

Der Genauigkeit halber sollte man eine orientierte Untermannigfaltigkeit statt mit M mit einem Symbol (M, \mathcal{A}) oder $(M, [\mathcal{A}])$ (mit \mathcal{A} einem orientierten Atlas die Orientierung $[\mathcal{A}]$ definierenden Atlas) bezeichnen. Üblicherweise wird die Orientierung in der Notation aber unterdrückt. Wir gehen später noch näher auf den Begriff der Orientierung ein.

Wir kommen nun zur Definition des Integrals von n -Formen über Teilmengen von n -dimensionalen Untermannigfaltigkeiten $M \subset \mathbb{R}^d$. Zuerst einige Vorbemerkungen.

Definition 4.24. Sei $M \subset \mathbb{R}^d$ eine n -dimensionale Untermannigfaltigkeit. Eine Teilmenge $A \subset M$ heißt *messbar* (bzw. *Nullmenge*), wenn für eine Überdeckung von A durch Kartengebiete $h_i : V_i \rightarrow U_i$ von M die Mengen $h_i^{-1}(U_i \cap A) \subset \mathbb{R}^n$ Lebesgue-messbar bzw. Lebesgue-Nullmengen sind.

Zeigen Sie, dass diese Definition nicht von der Wahl von Karten abhängt.

Der Einfachheit halber beschränken wir uns im Folgenden auf Untermannigfaltigkeiten, die in endlich viele paarweise disjunkte meßbare Teilmengen A_i zerlegt werden können, die je-

weils in nur einem Kartengebiet U_i (einer Karte $h_i : V_i \rightarrow U_i \subset M$) enthalten sind, und nennen solche Untermannigfaltigkeiten "einfach".

Definition und Satz 4.25. Sei $M \subset \mathbb{R}^d$ eine orientierte n -dimensionale einfache Untermannigfaltigkeit. Eine auf einer offenen Umgebung von M definierte n -Form ω heißt *integrierbar*, wenn für eine Zerlegung $M = \bigcup_{j=1}^J A_j$ in messbare in jeweils nur einem Kartengebiet enthaltene Mengen A_j (s.o.) und passende orientierungserhaltene Karten $h_j : V_j \rightarrow U_j \subset M$ die Form $h_j^* \omega$ über $h_j^{-1}(A_j) \subset \mathbb{R}^n$ integrierbar ist, $j = 1, \dots, J$.

In diesem Fall ist das *Integral von ω über A*

$$\int_A \omega := \sum_{j=1}^J \int_{h_j^{-1}(A_j \cap A)} h_j^* \omega.$$

Diese Konstruktion hängt nicht von der Wahl der Zerlegung und der Karten ab.

Beweis. Es seien $A_j, j = 1, \dots, J$ und $\tilde{A}_j, j = 1, \dots, \tilde{J}$, Zerlegungen von M in messbare jeweils in einem Kartengebiet enthaltene Mengen, und $h_j : V_j \rightarrow U_j$ bzw. $\tilde{h}_j : \tilde{V}_j \rightarrow \tilde{U}_j$ zugehörige orientierungserhaltene Karten. Wir nehmen an, dass die n -Form ω die Bedingungen bzgl. der A_j und h_j erfüllt. Dann ist

$$\int_A \omega = \sum_{j=1}^J \int_{h_j^{-1}(A_j \cap A)} h_j^* \omega$$

eine Summe von wohldefinierten Lebesgue-Integralen über messbare Teilmengen von \mathbb{R}^n . Wir zerlegen nun $A_j \cap A = \bigsqcup_{k=1}^{\tilde{J}} A_j \cap \tilde{A}_k \cap A$ in paarweise disjunkte messbare Mengen und erhalten so mit den Eigenschaften des Lebesgueintegrals

$$\int_A \omega = \sum_{j=1}^J \sum_{k=1}^{\tilde{J}} \int_{h_j^{-1}(A_j \cap \tilde{A}_k \cap A)} h_j^* \omega.$$

Da die Karten h_j und \tilde{h}_k alle orientierungserhaltend sind, also orientierungstreue Kartenwechsel mit allen Karten des ausgewählten orientierten Atlas' von M haben, sind auch die Kartenwechsel zwischen den h_j und \tilde{h}_k orientierungstreu (Kettenregel und Multiplikativität der Determinante).

Wir können also den Kartenwechsel $h_j^{-1}(A_j \cap \tilde{A}_k \cap A) \rightarrow \tilde{h}_k^{-1}(A_j \cap \tilde{A}_k \cap A)$ betrachten, und erhalten mit der Rechnung vor Def. 4.22

$$\begin{aligned} \int_A \omega &= \sum_{j=1}^J \sum_{k=1}^{\tilde{J}} \int_{h_j^{-1}(A_j \cap \tilde{A}_k \cap A)} h_j^* \omega = \sum_{j=1}^J \sum_{k=1}^{\tilde{J}} (+1) \int_{\tilde{h}_k^{-1}(A_j \cap \tilde{A}_k \cap A)} \tilde{h}_k^* \omega \\ &= \sum_{k=1}^{\tilde{J}} \int_{\tilde{h}_k^{-1}(\tilde{A}_k \cap A)} \tilde{h}_k^* \omega. \end{aligned}$$

Dies zeigt, dass $\int_A \omega$ nicht von den gemachten Wahlen abhängt. □

Ist (M, \mathcal{A}) eine orientierte n -dimensionale Untermannigfaltigkeit (hierbei bezeichnet \mathcal{A} einen orientierten Atlas von M , der die Orientierung $[\mathcal{A}]$ definiert), so können wir M auch mit der *umgekehrten Orientierung* ausstatten. Dazu betrachten wir

$$J : \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad J(x_1, \dots, x_n) := (-x_1, x_2, \dots, x_n)$$

und für eine Karte $h : V \rightarrow U \subset M$

$$\tilde{V} := J(V), \quad \tilde{h} : \tilde{V} \rightarrow U, \quad \tilde{h}(x) := h(J(x)).$$

Dies ist ebenfalls eine Karte. Betrachten wir den die Orientierung von M definierenden orientierten Atlas $\mathcal{A} := \{h_i : i \in \mathcal{I}\}$, so ist $\tilde{\mathcal{A}} := \{\tilde{h}_i : i \in \mathcal{I}\}$ ebenfalls ein orientierter Atlas. Wegen dem Minuszeichen in J ist jede Karte von $\tilde{\mathcal{A}}$ bzgl. jeder Karte von \mathcal{A} orientierungsumkehrend. Man nennt die durch $\tilde{\mathcal{A}}$ definierte Orientierung $[\tilde{\mathcal{A}}]$ von M die zu $[\mathcal{A}]$ umgekehrte Orientierung.

Beachten Sie, dass die Definition des Integrals $\int_M \omega$ von der Orientierung abhängt: Es gilt

$$\int_{(M, [\mathcal{A}])} \omega = - \int_{(M, [\tilde{\mathcal{A}}])} \omega.$$

Meistens wird die Orientierung aber nicht explizit in der Notation aufgeführt.

Beispiel 4.26 (Integrale von Einsformen / Vektorfeldern entlang Kurven). Wir betrachten eine glatte Kurve $\gamma : (a, b) \rightarrow \mathbb{R}^d$, deren Bild eine orientierbare eindimensionale Untermannigfaltigkeit $M = \gamma((a, b))$ mit dem Atlas $\{\gamma\}$ ist. Auf \mathbb{R}^d betrachten wir ein stetiges Vektorfeld $v : \mathbb{R}^d \rightarrow \mathbb{R}^d$ und die assoziierte 1-Form $\omega = v_1 dx_1 + \dots + v_d dx_d$. Dann ist der Rücktransport von ω bzgl. γ

$$(\gamma^* \omega)(t) = \sum_{i=1}^d v_i(\gamma(t)) d\gamma_i(t) = \sum_{i=1}^d v_i(\gamma(t)) \gamma'_i(t) dt = \langle v(\gamma(t)), \gamma'(t) \rangle dt.$$

Das Integral von ω entlang der Kurve ist

$$\int_M \omega = \int_a^b \langle v(\gamma(t)), \gamma'(t) \rangle dt.$$

Unser allgemeiner Integralbegriff gibt uns also insbesondere eine Möglichkeit, ein Vektorfeld entlang einer Kurve zu integrieren.

- Betrachten wir den Kreis $rS^1 \subset \mathbb{R}^2$ und eine 1-Form $\omega = v_1 dx_1 + v_2 dx_2$. Wir wählen die Orientierung von rS^1 , in der die Karte $h : (-\pi, \pi) \rightarrow S^1$, $\alpha \mapsto (r \cos \alpha, r \sin \alpha)$ orientierungserhaltend ist. Das Bild dieser Karte deckt rS^1 bis auf die Nullmenge $\{(-r, 0)\}$ ab, so dass wir uns für Integrale auf eine Karte beschränken dürfen. So ergibt sich

$$\int_{rS^1} \omega = r \int_{-\pi}^{\pi} (-\sin \alpha \cdot v_1(r \cos \alpha, r \sin \alpha) + \cos \alpha \cdot v_2(r \cos \alpha, r \sin \alpha)) d\alpha.$$

- Betrachten wir den speziellen Fall, dass ω exakt ist, also $\omega = df$ für eine Nullform (C^1 -Funktion $f : \mathbb{R}^d \rightarrow \mathbb{R}$). Dies ist äquivalent dazu, dass v ein Gradientenfeld

ist, $v = \text{grad} f$. In diesem Fall bemühen wir die Kettenregel aus Analysis 2, um $\langle \text{grad}(f)(\gamma(t)), \gamma'(t) \rangle = (f \circ \gamma)'(t)$ zu sehen. Damit erhalten wir

$$\int_M df = \int_a^b (f \circ \gamma)'(t) dt = f(\gamma(b)) - f(\gamma(a)),$$

was wir bereits als eine Verallgemeinerung des Hauptsatzes der Differential- und Integralrechnung ansehen können.

Insbesondere sehen wir, dass $\int_M df$ nur vom Anfangs- und Endpunkt der Kurve, nicht aber ihrem genauen Verlauf abhängt. Weiterhin sehen wir, dass $\int_M df = 0$ für jede *geschlossene* Kurve ($\gamma(a) = \gamma(b)$) gilt.

Beispiel 4.27. Wir betrachten die Einsform $\omega(x, y) = x dy$ auf \mathbb{R}^2 . Integriert über den mittels der Karte $h : (-\pi, \pi) \rightarrow rS^1, \alpha \mapsto (r \cos \alpha, r \sin \alpha)$ orientierten Kreises ergibt sich

$$\int_{rS^1} \omega = \int_{(-\pi, \pi)} h^* \omega = \int_{-\pi}^{\pi} r \cos \alpha d(r \sin \alpha) = r^2 \int_{-\pi}^{\pi} \cos^2 \alpha d\alpha = \pi r^2.$$

Wir machen folgende interessante Beobachtung: Die Zweiform $d\omega = dx \wedge dy$ kann über die offene Kreisscheibe $K := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < r^2\}$ integriert werden, dies ergibt $\int_K d\omega = \pi r^2$, den gleichen Wert wie $\int_{rS^1} \omega$.

Dies ist kein Zufall, sondern ein Ausdruck des Satzes von Stokes: Der Rand von K ist $\partial K = rS^1$, und es gilt $\int_K d\omega = \int_{\partial K} \omega$ (mehr dazu später).

Beispiel 4.28. Wir betrachten die zweidimensionale Untermannigfaltigkeit $H := S^2 \cap \{(x, y, z) \in \mathbb{R}^3 : z > 0\} \subset \mathbb{R}^3$ (offene obere Halbsphäre) mit der durch die Karte

$$h : D := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\} \rightarrow \mathbb{R}^3, \\ h(x, y) := (x, y, \sqrt{1 - x^2 - y^2})$$

gegebenen Orientierung. Mit einer stetigen Funktion $g : \mathbb{R}^3 \rightarrow \mathbb{R}$ haben wir die 2-Form

$$\omega := g dy \wedge dz,$$

deren Rücktransport bzgl h

$$(h^* \omega)(x, y) = \frac{x \cdot g(x, y, \sqrt{1 - x^2 - y^2})}{\sqrt{1 - x^2 - y^2}} dx \wedge dy$$

ist. Das Integral von ω über H ist also

$$\int_H \omega = \int_D \frac{x \cdot g(x, y, \sqrt{1 - x^2 - y^2})}{\sqrt{1 - x^2 - y^2}} dx dy.$$

Um den Begriff der Orientierung einer n -dimensionalen Untermannigfaltigkeit $M \subset \mathbb{R}^d$ und des Integrals von n -Formen besser zu verstehen, betrachten wir das Zusammenspiel

von Orientierung und Tangentialräumen. Aus Satz 4.20 a) wissen wir: Ist $p \in M$ und $h : V \subset \mathbb{R}^n \rightarrow U \subset M$ eine Karte mit $p = h(x) \in U$, so bilden die Vektoren $((\partial_1 h)(x), \dots, (\partial_n h)(x))$ eine geordnete Basis von $T_p M$.

Ist M orientiert, so nennen wir eine Basis der Form $((\partial_1 h)(x), \dots, (\partial_n h)(x))$ von $T_p M$ *positiv orientiert*, wenn h eine orientierungserhaltende Karte ist. Auf diese Weise definiert eine Orientierung von M (und die kanonische Orientierung von \mathbb{R}^n) eine Orientierung ihrer Tangentialräume $T_p M$.

Wir spezialisieren uns nun auf den Fall einer *Hyperfläche*, also einer $(d - 1)$ -dimensionalen Untermannigfaltigkeit von \mathbb{R}^d (z.B. die Kugeloberfläche S^2 im dreidimensionalen Raum \mathbb{R}^3).

Definition 4.29. Ein stetiges *Einheits-Normalenfeld* ν einer Hyperfläche M (eine $(d-1)$ -dim. Untermannigfaltigkeit von \mathbb{R}^d) ist eine stetige Abbildung

$$\nu : M \rightarrow \mathbb{R}^d,$$

so dass a) $\nu(p) \perp T_p M$ für alle $p \in M$ und b) $\|\nu(p)\| = 1$ für alle $p \in M$.

Ist M orientiert, so heißt ein Einheitsnormalenfeld ν von M *positiv orientiert*, wenn für jede positiv orientierte Basis (v_1, \dots, v_{d-1}) von $T_p M$ die Basis $(\nu(p), v_1, \dots, v_{d-1})$ eine positiv orientierte Basis von \mathbb{R}^d ist.

Ein Einheitsnormalenfeld gibt also an jedem Punkt der Mannigfaltigkeit einen Einheitsvektor, der senkrecht auf dem Tangentialraum steht. Da der Tangentialraum einer Hyperfläche Kodimension 1 hat, gibt es genau zwei Einheitsvektoren, die auf $T_p M$ senkrecht stehen. Die Orientierungsbedingung wählt einen von ihnen aus (dh eine "Seite" der Mannigfaltigkeit wird als Außenseite deklariert). Wenn die Vektoren $\nu(p)$ so punktweise ausgewählt werden, sieht man leicht, dass es immer ein positiv orientiertes Einheits-Normalenfeld gibt. Allerdings muss dieses Feld nicht stetig sein, wie das Möbiusband-Beispiel zeigt.

Satz 4.30. Sei $M \subset \mathbb{R}^d$ eine Hyperfläche ($d \geq 2$).

- a) Falls M eine Orientierung besitzt, so gibt es genau ein stetiges positiv orientiertes Einheits-Normalenfeld von M .
- b) Falls M ein stetiges Einheitsnormalenfeld ν hat, so gibt es genau eine Orientierung von M , bzgl der ν positiv orientiert ist.

Beweis. a) Wie oben erläutert, ist die Eindeutigkeit klar: An jedem Punkt $p \in M$ wählen wir eine positiv orientierte Basis (v_1, \dots, v_{d-1}) von $T_p M$ und fixieren $\nu(p)$ eindeutig durch $\|\nu(p)\| = 1$ und die positive Orientierung von $(\nu(p), v_1, \dots, v_{d-1})$. Um die Existenzbehauptung zu zeigen, müssen wir zeigen, dass ν stetig ist. Dies können wir lokal, also in der Nähe eines ausgewählten Punktes $p \in M$, prüfen.

Dazu behaupten wir zunächst, dass M in einer Umgebung von p durch $f(q) = 0$ für eine C^1 -Funktion $f : G \subset \mathbb{R}^d \rightarrow \mathbb{R}$ gegeben ist (also 0 ein regulärer Wert von f ist; Beweis dieser Behauptung als Übung mit dem Satz über die Umkehrfunktion). Dann wissen wir (Analysis 2, Lemma 13.29), dass $\text{grad} f(p) \neq 0$ senkrecht auf $T_p M$ steht. Wir betrachten das

Einheitsnormalenfeld

$$\tilde{\nu}(q) := \frac{\text{grad}f(q)}{\|\text{grad}f(q)\|},$$

das $\tilde{\nu}(p) = \pm\nu(p)$ am Punkt $q = p$ erfüllt. Sei nun h eine positiv orientierte Karte bei p . Dann ist per Konstruktion die Funktion

$$\Delta : x \mapsto \det(\pm\tilde{\nu}(h(x)), \partial_1 h(x), \dots, \partial_{d-1} h(x))$$

bei $x = h^{-1}(p)$ positiv. Da Δ stetig ist, ist sie auch in einer offenen Umgebung von x positiv. Also stimmen ν und $\tilde{\nu}$ in einer Umgebung von p überein, dh ν ist stetig.

b) Sei $\nu : M \rightarrow \mathbb{R}^n$ ein stetiges Einheitsnormalenfeld und $p \in M$. Wir betrachten eine Karte $h : V \rightarrow U \ni p$, evtl nach Verkleinerung von V dürfen wir annehmen, dass V zusammenhängend ist. Die schon in a) betrachtete stetige Funktion

$$\Delta : V \rightarrow \mathbb{R} \setminus \{0\}, \quad \Delta : x \mapsto \det(\nu(h(x)), \partial_1 h(x), \dots, \partial_{d-1} h(x)),$$

hat keine Nullstellen, nimmt also auf ganz V das gleiche Vorzeichen an. Sollte das Vorzeichen negativ sein, schalten wir $(x_1, \dots, x_n) \mapsto (x_1, \dots, x_{n-1}, -x_n)$ vor die Karte h . Wir dürfen also $\Delta(x) > 0$ für alle $x \in V$ annehmen.

Da p beliebig war, erhalten wir auf diese Weise einen Atlas und behaupten, dass dieser Atlas orientiert ist. Seien also $h : V \rightarrow U$, $\tilde{h} : \tilde{V} \rightarrow \tilde{U}$ zwei solche Karten, mit überlappenden Kartengebieten und $\psi : h^{-1}(U \cap \tilde{U}) \rightarrow \tilde{h}^{-1}(U \cap \tilde{U})$, $\psi = \tilde{h}^{-1} \circ h$.

Mit $x = h^{-1}(p)$, $\tilde{x} = \tilde{h}^{-1}(p)$ betrachten wir im Tangentialraum $T_p M$ die Basen $v_j := (\partial_j h)(x)$ und $\tilde{v}_j = (\partial_j \tilde{h})(\tilde{x})$. Nach der Kettenregel gilt

$$v_j = (\partial_j(\tilde{h} \circ \psi))(x) = \sum_{i=1}^{d-1} (\partial_i \tilde{h})(\psi(x)) (\partial_j \psi_i)(x) = \sum_{i=1}^{d-1} \tilde{v}_i (D\psi)(x)_{ij}.$$

Die Matrix $(D\psi)(x)$ bildet also die eine auf die andere Basis ab. Da nach Voraussetzung

$$\det(\nu(p), v_1, \dots, v_{d-1}) > 0, \quad \det(\nu(p), \tilde{v}_1, \dots, \tilde{v}_{d-1}) > 0,$$

folgt auch $\det(D\psi)(x) > 0$. Also sind die beiden Karten h, \tilde{h} gleich orientiert. \square

Beispiel 4.31. Wir betrachten einige Beispiele von Hyperflächen mit stetigen Einheitsnormalenfeldern.

- $M = \mathbb{R}^2 \times \{0\} \subset \mathbb{R}^3$ hat $\nu(x, y) = (0, 0, 1)$ für alle $x, y \in \mathbb{R}$.
- $M = S^d \subset \mathbb{R}^{d+1}$ hat $\nu(x) = x$ für alle $x \in S^d$.
- $M = \{(t, t^2) : t \in \mathbb{R}\} \subset \mathbb{R}^2$ (Parabelkurve) hat im Punkt (t, t^2) Tangentenvektor $(1, 2t)$, also Normalenvektor $\pm(-2t, 1)$. Wir wählen das positive Vorzeichen und normieren, das ergibt $\nu((t, t^2)) = \frac{1}{\sqrt{1+4t^2}}(-2t, 1)$.
- Etwas allgemeiner als im vorigen Beispiel betrachten wir eine offene Menge $U \subset \mathbb{R}^n$ und eine C^1 -Funktion $f : U \rightarrow \mathbb{R}$. Dann ist der Graph

$$M := \{(x, f(x)) : x \in U\} \subset \mathbb{R}^{n+1}$$

eine n -dimensionale Untermannigfaltigkeit (Analysis 2, Beispiel 13.22). Wir können M durch eine einzelne Karte beschreiben, nämlich $h : U \rightarrow \mathbb{R}^{n+1}$, $h(x) := (x, f(x))$. Diese Karte gibt uns als Basis des Tangentialraums $T_{(x,f(x))}M$

$$\begin{aligned}(\partial_1 h)(x) &= (1, 0, \dots, 0, (\partial_1 f)(x)), \\(\partial_2 h)(x) &= (0, 1, \dots, 0, (\partial_2 f)(x)), \dots, \\(\partial_n h)(x) &= (0, 0, \dots, 1, (\partial_n f)(x)).\end{aligned}$$

Daran erkennen wir, dass die beiden Vektoren $\pm(-\text{grad}f(x), 1)$ senkrecht auf $T_{(x,f(x))}M$ stehen. Nach Normierung erhalten wir also die beiden stetigen Einheitsnormalenfelder

$$\nu((x, f(x))) = \pm \frac{1}{\sqrt{1 + \|\text{grad}f(x)\|^2}} \begin{pmatrix} -\text{grad}f(x) \\ 1 \end{pmatrix},$$

die die beiden möglichen Orientierungen auf M induzieren.

Anschaulich ist eine Orientierung einer eindimensionalen Untermannigfaltigkeit eine Durchlaufrichtung, eine Orientierung einer zweidimensionalen Untermannigfaltigkeit eine konsistente Festlegung eines Drehsinns in jedem Tangentialraum, und eine Orientierung einer dreidimensionalen Untermannigfaltigkeit eine Festlegung eines Schraubensinns (Rechte-Hand-Regel) in jedem Tangentialraum.

4.4 Der Satz von Stokes

Wir kommen jetzt zu dem zentralen Satz der Integrationstheorie von Differentialformen, dem Satz von Stokes. Dieser Satz ist die Verallgemeinerung des Hauptsatzes der Differential- und Integralrechnung auf höhere Dimensionen und Untermannigfaltigkeiten. Der Hauptsatz besagt für eine C^1 -Funktion f auf \mathbb{R}

$$\int_a^b f'(x) dx = f(b) - f(a).$$

Im Rahmen von Differentialformen haben wir schon gelernt, dass i) die adäquate Verallgemeinerung des Intervalls (a, b) eine orientierte n -dimensionale Untermannigfaltigkeit $M \subset \mathbb{R}^d$ (ein Beispiel ist $(a, b) \subset \mathbb{R}$ mit der kanonischen Orientierung), ii) die adäquate Verallgemeinerung des Integranden eine n -Form, und iii) die adäquate Verallgemeinerung der Ableitung f' die äußere Ableitung $d\omega$ ist (ein Beispiel ist die 0-Form f und die 1-Form $df = f' dx$). Die linke Seite des Hauptsatzes verallgemeinert sich also zu $\int_M d\omega$. Da M eine n -Mannigfaltigkeit ist, sollte ω eine $(n - 1)$ -Form sein.

Die rechte Seite des Hauptsatzes betrifft nur f , in der Verallgemeinerung auf Formen erwarten wir hier also die $(n - 1)$ -Form ω . Die Punkte a, b bilden den *Rand* des Intervalls $[a, b]$ (das wir hier als abgeschlossen ansehen). Wir werden im Folgenden definieren, was eine berandete Untermannigfaltigkeit $M \subset \mathbb{R}^d$ ist, dies wird eine Untermannigfaltigkeit ∂M mit $\dim(\partial M) = \dim M - 1$ sein. Wir werden weiterhin definieren, wie eine Orientierung von M eine Orientierung des Randes ∂M festlegt. Dies wird so gemacht sein, dass im Beispiel des kanonisch orientierten Intervalls $[a, b]$ der Rand $\partial[a, b] = \{a\} \cup \{b\}$ die Orientierung -1 bei a und $+1$ bei b trägt, so dass die rechte Seite des Hauptsatzes mit $\int_{[a,b]} df = \int_{\{a\} \cup \{b\}} f = f(b) - f(a)$ übereinstimmt.

Der Satz von Stokes wird dann durch

$$\int_M d\omega = \int_{\partial M} \omega$$

wiedergegeben (die genauen Voraussetzungen werden später beschrieben).

Wir beginnen mit der Diskussion von Rändern und warnen schon gleich zu Anfang davor, dass der im Folgenden entwickelte Begriff "Rand" im Allgemeinen von dem aus Analysis 2 bzw. Topologie bekannten Begriff des Randes ∂A einer Teilmenge A eines metrischen Raums verschieden ist, aber trotzdem mit demselben Symbol ∂A bezeichnet wird.

Zunächst betrachten wir den berandeten Halbraum, $d \geq 2$,

$$\mathbb{R}_-^d := \{(x_1, \dots, x_d) \in \mathbb{R}^d : x_1 \leq 0\} \subset \mathbb{R}^d.$$

Genau wie der \mathbb{R}^d das lokale Modell einer Untermannigfaltigkeit ist, ist der berandete Halbraum \mathbb{R}_-^d das lokale Modell einer berandeten Untermannigfaltigkeit. Beachten Sie, dass \mathbb{R}_-^d keine Untermannigfaltigkeit ist: Die Randpunkte $(0, x_2, \dots, x_d)$ liegen nicht in Bildern von auf offenen Gebieten $V \subset \mathbb{R}^d$ definierten Karten.

Das Innere von $M := \mathbb{R}_-^d$ von \mathbb{R}_-^d ist hingegen eine offene Teilmenge von \mathbb{R}^d , also eine d -dimensionale Untermannigfaltigkeit von \mathbb{R}^d , der wir ihre kanonische Orientierung geben. Der Rand ∂M von M wird hier genau wie in Analysis 2 als die berandende Ebene $\partial M := \{x_1 = 0\} \subset \mathbb{R}^d$ definiert (später kommt eine allgemeinere Definition). Dann ist ∂M eine $(d-1)$ -dimensionale Untermannigfaltigkeit des \mathbb{R}^d , die wir durch die globale Karte

$$h : \mathbb{R}^{d-1} \rightarrow \mathbb{R}^d, \quad h(x_1, \dots, x_{d-1}) := (0, x_1, \dots, x_{d-1})$$

orientieren. Die Tangentialräume $T_p \partial M$ erhalten dadurch die Basis $(\partial_1 h(x), \dots, \partial_{d-1} h(x)) = (e_2, \dots, e_d)$. Das eindeutige stetige positiv orientierte Einheitsnormalenfeld von ∂M , das diese Basis zu einer positiv orientierten Basis von $T_p M = \mathbb{R}^d$ erweitert, ist also das nach außen zeigende Vektorfeld $\nu(x) = e_1$.

Lemma 4.32. *Sei ω eine stetig differenzierbare $(d-1)$ -Form im \mathbb{R}^d mit kompaktem Träger. Dann gilt mit den oben eingeführten Bezeichnungen und Orientierungen*

$$\int_{\mathbb{R}_-^d} d\omega = \int_{\partial \mathbb{R}_-^d} \omega.$$

Beweis. Wir können die Form als $\omega = \sum_{j=1}^d (-1)^{j-1} f_j dx_1 \wedge \dots \wedge \widehat{dx_j} \wedge \dots \wedge dx_d$ mit C^1 -Funktionen f_1, \dots, f_d schreiben. Für die oben definierte Karte h gilt $dh_1 = 0$ und $dh_j = dx_{j-1}$ für $j > 1$. Damit erhalten wir

$$\begin{aligned} h^* \omega &= \sum_{j=1}^d (f_j \circ h) (-1)^{j-1} dh_1 \wedge \dots \wedge \widehat{dh_j} \wedge \dots \wedge dh_d \\ &= (f_1 \circ h) dx_1 \wedge \dots \wedge dx_d, \end{aligned}$$

also

$$\int_{\partial \mathbb{R}_-^d} \omega = \int_{\mathbb{R}^{d-1}} f_1(0, x_1, \dots, x_{d-1}) dx_1 \dots dx_{d-1}.$$

Dieses Integral ist wohldefiniert und endlich, da ω (also insbesondere f_1) stetig ist und kompakten Träger hat.

Für das Integral von $d\omega = \sum_{j=1}^d dx_1 \wedge \dots \wedge df_j \wedge \dots \wedge dx_d = \sum_{j=1}^d \frac{\partial f_j}{\partial x_j} dx_1 \wedge \dots \wedge dx_d$ über $\mathbb{R}_-^d = \mathbb{R}_- \times \mathbb{R}^{d-1}$ betrachten wir zunächst den Term mit $j = 1$ und berechnen

$$\int_{-\infty}^0 \frac{\partial f_1}{\partial x_1}(x_1, \dots, x_d) dx_1 = f_1(0, x_2, \dots, x_d) - 0$$

(weil f_1 kompakten Träger hat). Damit erhalten wir

$$\int_{\mathbb{R}_-^d} \frac{\partial f_1}{\partial x_1}(x_1, \dots, x_d) dx_1 \cdots dx_d = \int_{\mathbb{R}^{d-1}} f_1(0, x_2, \dots, x_d) dx_2 \cdots dx_d,$$

was bereits mit $\int_{\partial \mathbb{R}_-^d} \omega$ übereinstimmt. Die Terme mit $j > 1$ müssen nach Integration also verschwinden. Da f_j kompakten Träger hat, haben wir $\int_{\mathbb{R}} \frac{\partial f_j}{\partial x_j} dx_j = 0$ (dieses Integral ist gleich der Auswertung von $f_j(x_1, \dots, x_d)$ an zwei Punkten $x_j = -R$ und $x_j = R$, die für R groß genug verschwindet). Also ergibt sich für $j > 1$

$$\int_{\mathbb{R}_-^d} \frac{\partial f_j}{\partial x_j} dx_1 \cdots dx_d = \int_{\mathbb{R}_- \times \mathbb{R}^{d-2}} \left(\int_{\mathbb{R}} \frac{\partial f_j}{\partial x_j} dx_j \right) dx_1 \cdots \widehat{dx_j} \cdots dx_d = 0,$$

und der Beweis ist abgeschlossen. □

Nach dieser vorbereitenden Überlegung kommen wir zu allgemeinen berandeten Untermannigfaltigkeiten. Die folgende Diskussion bis zum Satz von Stokes wird aus Zeitgründen nicht in allen Details ausgeführt.

Definition 4.33. Eine *berandete n -dimensionale Untermannigfaltigkeit* von \mathbb{R}^d ist eine Teilmenge $M \subset \mathbb{R}^d$, so dass es zu jedem Punkt $p \in M$ eine Karte $h : V \subset \mathbb{R}_-^n \rightarrow U \subset M$, $p \in U$, gibt (mit $V \subset \mathbb{R}_-^n$ und $U \subset M$ relativ offen). Der Punkt $p = h(x)$ heißt *Randpunkt von M* wenn p Bild eines Randpunktes $x \in \partial \mathbb{R}_-^n$ ist. Die Menge aller Randpunkte von M wird mit ∂M bezeichnet.

Wie schon zuvor heißt hier $V \subset \mathbb{R}_-^n$ relativ offen, wenn es für jedes $x \in V$ eine offene Teilmenge $O \subset \mathbb{R}^n$ gibt, so dass $x \in O \cap \mathbb{R}_-^n$. Zum Beispiel ist \mathbb{R}_-^n relativ offen in \mathbb{R}_-^n , obwohl \mathbb{R}_-^n nicht offen ist in \mathbb{R}^n .

Mit ‘Karte’ ist in obiger Definition eine Abbildung genau wie in Satz 4.20 gemeint. Für die Differenzierbarkeitseigenschaften von h auf der nicht notwendigerweise offenen Menge V ist gemeint, dass es eine C^1 -Ausdehnung auf eine offene Umgebung von V gibt.

Bei einer berandeten Untermannigfaltigkeit kann die Umgebung V nicht nur in \mathbb{R}_-^n , sondern sogar in \mathbb{R}^n offen sein, dh nicht den Rand des Halbraums \mathbb{R}_-^n treffen. Dies sind die inneren Punkte von M . Insbesondere kann es passieren, dass M nur innere Punkte, also leeren Rand hat – dies sind die Untermannigfaltigkeiten ohne Rand, also genau die bisher betrachteten Untermannigfaltigkeiten.

Berandete Untermannigfaltigkeiten werden oft auch ‘Untermannigfaltigkeiten mit Rand’ genannt – beachten Sie aber, dass eine berandete Untermannigfaltigkeit im Allgemeinen keine Untermannigfaltigkeit ist.

Beispiel 4.34.

- a) Die abgeschlossene Kreisscheibe $K = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\} \subset \mathbb{R}^2$ ist eine berandete Untermannigfaltigkeit des \mathbb{R}^2 , mit Rand $\partial K = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$: Für die Punkte im Inneren $K \setminus \partial K$ finden wir Karten wie gewohnt. Für den Randpunkt $(1, 0)$ betrachten wir die Karte

$$h : K - (1, 0) \subset \mathbb{R}_-^2 \rightarrow \mathbb{R}^2, \quad h(x, y) := (x + 1, y),$$

die $(0, 0) \in \partial \mathbb{R}_-^2$ auf $(1, 0) \in \partial K$ abbildet. Für einen allgemeinen Randpunkt $(x, y) = (\cos \alpha, \sin \alpha) \in K$ kann man analog argumentieren, indem man noch eine Rotation um α mit der Karte komponiert.

In diesem Fall stimmt der topologische Rand mit dem differentialgeometrischen Rand von K überein.

- b) Die abgeschlossene Halbsphäre

$$M = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1, x \geq 0\} \subset \mathbb{R}^3$$

ist eine berandete Untermannigfaltigkeit des \mathbb{R}^3 , mit Rand

$$\partial M = \{(x, y, 0) \in \mathbb{R}^3 : y^2 + z^2 = 1\}.$$

- c) Der Zylinder

$$Z = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = r^2, -L \leq z \leq L\}$$

ist eine zweidimensionale berandete Untermannigfaltigkeit des \mathbb{R}^3 . In diesem Fall ist der Rand nicht zusammenhängend, sondern besteht aus zwei disjunkten Kreisen,

$$\partial Z = \{(x, y, z) : x^2 + y^2 = r^2, z = L\} \cup \{(x, y, z) : x^2 + y^2 = r^2, z = -L\}.$$

In diesen Beispielen ist der Rand ∂M jeweils eine Untermannigfaltigkeit (ohne Rand) der Dimension $\dim(\partial M) = \dim M - 1$. Dies ist auch ganz allgemein so.

Satz 4.35. Sei $M \subset \mathbb{R}^d$ eine n -dimensionale berandete Untermannigfaltigkeit.

- a) $\partial M \subset \mathbb{R}^d$ ist eine $(n - 1)$ -dimensionale Untermannigfaltigkeit.
b) Ist M orientierbar, so ist auch ∂M orientierbar.

Beweis. (Skizze) a) Wir betrachten einen Atlas \mathcal{A} von M und einen Randpunkt $p \in \partial M$. Dann gibt es eine Karte $h : V \subset \mathbb{R}_-^n \rightarrow U \subset M$ in \mathcal{A} . Wir betrachten $b : \mathbb{R}_-^{n-1} \rightarrow \partial \mathbb{R}_-^n$, $b(x_1, \dots, x_{n-1}) = (0, x_1, \dots, x_{n-1})$ und setzen $\tilde{V} := b^{-1}(\partial \mathbb{R}_-^n \cap V)$. Unsere Karte für ∂M ist dann

$$\tilde{h} := h \circ b : \tilde{V} \rightarrow h(\partial \mathbb{R}_-^n \cap V), \quad \tilde{h}(x_1, \dots, x_{n-1}) = h(0, x_1, \dots, x_{n-1}).$$

Man prüft nach, dass dies eine Karte im Sinne der üblichen Definition (ohne Rand) ist.

b) Wir nehmen nun an, dass \mathcal{A} orientiert ist und betrachten zwei Karten mit am Rand überlappenden Kartengebieten, $h_i : V_i \rightarrow U_i$, $i = 1, 2$, mit $U_1 \cap U_2 \cap \partial M \neq \emptyset$. Wir behaupten, dass die beiden daraus gemäß a) konstruierten Karten $\tilde{h}_i = h_i \circ b$ gleich orientiert sind. Wir wissen, dass der Kartenwechsel $\psi = h_2^{-1} \circ h_1$ orientierungstreu ist. Dann ist der Kartenwechsel der Randkarten

$$\tilde{\psi} = \tilde{h}_2^{-1} \circ \tilde{h}_1^{-1}, \quad \tilde{\psi}(x) = (\psi_2(0, x), \dots, \psi_n(0, x)).$$

Es lässt sich zeigen, dass die Jacobimatrix von $\tilde{\psi}$ positive Determinante hat, weil die Jacobimatrix von ψ positive Determinante hat. \square

In Teil b) dieser Beweisskizze haben wir nicht nur gezeigt, dass mit M auch ∂M orientierbar ist, sondern auch gleich eine Konvention für die Wahl der Orientierung von ∂M bei gegebener Orientierung von M getroffen. Wir nennen dies die von M auf ∂M *induzierte Orientierung*.

Diese Konvention kann wie folgt ausgedrückt werden: Ist $p \in \partial M$ ein Randpunkt und $h : V \rightarrow U$, $h(x) = p$ eine Karte, so sind die partiellen Ableitungen $(\partial_1 h(x), \dots, \partial_n h(x))$ eine Basis von $T_p M$ wie üblich. Ein Vektor $v \in T_p M$ heißt *nach außen (innen) zeigend*, wenn in dieser Karte seine erste Komponente positiv (negativ) ist. (Dieser Konvention liegt unsere Abbildung nach \mathbb{R}_-^n zugrunde, für die e_1 das nach außen zeigende Einheitsnormalenfeld ist.)

Die Orientierungskonvention ist dann, dass eine Basis (w_1, \dots, w_{n-1}) von $T_p \partial M$ genau dann positiv orientiert ist, wenn für einen (dann jeden) nach außen zeigenden Vektor $v \in T_p M$ die Basis (v, w_1, \dots, w_{n-1}) von $T_p M$ positiv orientiert ist.

Beispiel 4.36.

- a) Sei $M = \{(x, y) : x^2 + y^2 \leq 1\} \subset \mathbb{R}^2$ mit $\partial M = \{(x, y) : x^2 + y^2 = 1\}$. Orientieren wir ∂M im mathematisch positiven Umlaufssinn. Dies ist die von der kanonischen Orientierung von M induzierte Orientierung von ∂M : Der positiv orientierte Tangentialvektor an ∂M bei Winkel α ist $w = (-\sin \alpha, \cos \alpha)$, und ein nach außen zeigender Vektor ist der nach außen zeigende Normalenvektor $v = (\cos \alpha, \sin \alpha)$. Wegen $\det(v, w) = 1$ folgt die Behauptung.
- b) Sei $Z = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = 1, -L \leq z \leq L\} \subset \mathbb{R}^3$ mit $\partial Z = R_- \cup R_+$, $R_{\pm} := S^1 \times \{\pm L\}$. Wir orientieren Z durch das nach außen zeigende Einheitsnormalenfeld

$$\nu(x, y, z) = (x, y, 0), \quad (x, y, z) \in Z.$$

Dann ist am Punkt $p = (x, y, \pm L) \in \partial Z$ eine positiv orientierte Basis von $T_p Z$ durch

$$(v, e_3), \quad v := \begin{pmatrix} -y \\ x \\ 0 \end{pmatrix}$$

gegeben, und $\{v\}$ bildet eine Basis von $T_p \partial Z$. Wir behaupten, dass $\{-v\}$ positiv orientiert für $T_p R_+$ und $\{+v\}$ positiv orientiert für $T_p R_-$ ist. Dazu beachten wir, dass $\pm e_3$ am Rand R_{\pm} nach außen zeigt. Also ist $(\pm e_3, \mp v)$ positiv orientiert (ein

Vorzeichen von \pm und \mp , und ein Vorzeichen vom Umdrehen der Reihenfolge der Basisvektoren).

Also induziert die Orientierung von Z die durch $\mp v$ gegebene Orientierung auf R_{\pm} , dh bei Betrachtung von Z aus der positiven z -Richtung im mathematisch positiven (für R_+) bzw. negativen (für R_-) Umlaufsinn.

Man kann sich merken, dass die Orientierungskonvention für eine dreidimensionale berandete Untermannigfaltigkeit M von \mathbb{R}^3 besagt, dass die Oberfläche ∂M bei Draufsicht im mathematisch positiven Umlaufsinn orientiert ist.

Wir haben nun alle Konzepte eingeführt, die wir für den Satz von Stokes brauchen.

Satz 4.37 (Satz von Stokes). Sei $O \subset \mathbb{R}^d$ offen, $M \subset O \subset \mathbb{R}^d$ eine kompakte n -dimensionale berandete orientierte Untermannigfaltigkeit, und ω eine auf O definierte stetig differenzierbare $(n-1)$ -Form. Dann gilt

$$\int_M d\omega = \int_{\partial M} \omega,$$

wenn ∂M die von M induzierte Orientierung trägt.

Die Annahme über die Kompaktheit von M verdient einige Kommentare.

- Beginnen wir mit einem einfachen Beispiel: Das Intervall $[a, b] \subset \mathbb{R}$ ist eine kompakte kanonisch orientierte 1-dimensionale berandete Untermannigfaltigkeit von \mathbb{R} , mit Rand $\partial[a, b] = \{a, b\}$. Die induzierte Orientierung von $\{a, b\}$ ist $+1$ bei b und -1 bei a . Eine 0-Form auf einer offenen Umgebung $O = (a - \varepsilon, b + \varepsilon)$ von $[a, b]$ (mit $\varepsilon > 0$) ist eine Funktion $f : (a - \varepsilon, b + \varepsilon) \rightarrow \mathbb{R}$, mit äußerer Ableitung $df = f' dx$. In diesem Fall sagt der Satz von Stokes also

$$\int_{[a,b]} f'(x) dx = \int_{\partial[a,b]} f(x) dx = f(b) - f(a),$$

und reproduziert damit den Hauptsatz der Differential- und Integralrechnung.

- Wenn wir das obige Beispiel etwas modifizieren und $M' := [a, b)$ oder $M'' := (a, b)$ anstelle von $[a, b]$ betrachten, so sind M' und M'' nicht mehr kompakt, aber trotzdem noch kanonisch orientierte 1-dimensionale berandete Untermannigfaltigkeiten von \mathbb{R} . Ihre Ränder sind $\partial M' = \{a\}$ und $\partial M'' = \emptyset$. Da $\{b\}$ und $\{a, b\}$ Lebesgue-Nullmengen sind, haben wir in diesen Fällen

$$\begin{aligned} \int_{[a,b)} f'(x) dx &= \int_{[a,b]} f'(x) dx = f(b) - f(a) \neq -f(a) = \int_{\partial[a,b)} f(x) dx, \\ \int_{(a,b)} f'(x) dx &= \int_{[a,b]} f'(x) dx = f(b) - f(a) \neq 0 = \int_{\partial(a,b)} f(x) dx. \end{aligned}$$

Der Satz von Stokes gilt hier also nicht, was die Annahme über die Kompaktheit von M erklärt. Nichtsdestotrotz kann mit dem Satz von Stokes argumentiert werden: Im

Falle von $M = (a, b)$ betrachten wir $N := [a + \varepsilon, b - \varepsilon]$, eine kompakte Untermannigfaltigkeit mit Rand $\{a + \varepsilon, b - \varepsilon\}$, und erhalten

$$\int_{[a+\varepsilon, b-\varepsilon]} f'(x) dx = f(b - \varepsilon) - f(a + \varepsilon),$$

was im Limes $\varepsilon \rightarrow 0$ das erwartete Ergebnis liefert.

- Wir können auch Lemma 4.32 als einen Spezialfall des Satzes von Stokes ansehen, obwohl die dort betrachtete berandete Mannigfaltigkeit $M := \mathbb{R}_-^d$ nicht kompakt (da unbeschränkt) ist. In dem Lemma wurde ja ω als mit kompaktem Träger $\text{supp } \omega$ vorausgesetzt, sagen wir $\text{supp } \omega \subset [-R, R]^d$ für geeignetes $R > 0$. Dann können wir die kompakte berandete Untermannigfaltigkeit $M' := \mathbb{R}_-^d \cap [-R - 1, R + 1]^d$ betrachten und den Satz von Stokes anwenden. Wegen $\text{supp } \omega \subset [-R, R]^d$ verschwindet ω aber auf allen Punkten des Randes, die nicht im Rand von \mathbb{R}_-^d liegen. Dies resultiert in der Aussage von Lemma 4.32.

Beweis. (Skizze) Wir betrachten den Fall, dass wir nur ein Kartengebiet benötigen (globale Karte), dh $h : V \subset \mathbb{R}_-^n \rightarrow M$. Dann erhalten wir

$$\int_M d\omega = \int_V h^* d\omega = \int_V d(h^* \omega).$$

Da ω und damit auch $h^* \omega$ kompakten Träger hat, können wir $h^* \omega$ durch Null zu einer auf ganz \mathbb{R}_-^n stetig differenzierbaren n -Form $\hat{\omega}$ fortsetzen, so dass das obige Integral als

$$\int_M d\omega = \int_{\mathbb{R}_-^n} d\hat{\omega} = \int_{\partial \mathbb{R}_-^n} \hat{\omega}$$

geschrieben werden kann, wobei im zweiten Schritt Lemma 4.32 verwendet wurde.

Andererseits erhalten wir mit der von h induzierten Karte $\tilde{h} = h \circ b : \tilde{V} \subset \mathbb{R}_-^{n-1} \rightarrow \partial M$

$$\int_{\partial M} \omega = \int_{\tilde{V}} b^* h^* \omega = \int_{\mathbb{R}_-^{n-1}} b^* \hat{\omega} = \int_{\partial \mathbb{R}_-^n} \hat{\omega}.$$

Für den allgemeinen Fall muss M mit mehreren Kartengebieten abgedeckt werden. Man kann M nicht einfach in disjunkte Teile zerlegen wie bei der Integration, da dies zu Problemen mit der Differenzierbarkeit führen würde. Stattdessen bedient man sich einer anderen Technik (glatte Zerlegung der Eins), auf die wir hier aber aus Zeitgründen nicht eingehen. \square

Korollar 4.38. Sei $M \subset \mathbb{R}^d$ eine n -dimensionale kompakte Untermannigfaltigkeit und ω eine auf einer offenen Umgebung von M definierte stetig differenzierbare $(n - 1)$ -Form. Dann gilt

$$\int_M d\omega = 0.$$

Beweis. Da M leeren Rand hat, folgt mit dem Satz von Stokes $\int_M d\omega = \int_{\emptyset} \omega = 0$. \square

Für Anwendungen des Stokesschen Integralsatzes verweisen wir auf die Hausaufgaben.

Der Satz von Stokes besticht durch seine Allgemeinheit und elegante Formulierung. In Anwendungen sind aber meist konkretere Spezialfälle dieses Satzes wichtiger, die wir nun diskutieren. Wir konzentrieren uns auf Untermannigfaltigkeiten des Anschauungsraums \mathbb{R}^3 und Untermannigfaltigkeiten der Dimensionen $n = 1, 2, 3$. Wir erinnern uns an die Beschreibung von n -Formen und ihren äußeren Ableitungen durch die vektoriellen Strecken- und Flächenelemente und das Volumenelement (Seite 108)

$$d\vec{s} = (dx_1, dx_2, dx_3), \quad d\vec{S} = (dx_2 \wedge dx_3, dx_3 \wedge dx_1, dx_1 \wedge dx_2), \quad dV = dx_1 \wedge dx_2 \wedge dx_3.$$

Satz 4.39. Sei $U \subset \mathbb{R}^3$ offen und $f : U \rightarrow \mathbb{R}$ stetig differenzierbar. Dann gilt

$$\int_M \text{grad } f \cdot d\vec{s} = f(q) - f(p)$$

für jede orientierte kompakte eindimensionale Untermannigfaltigkeit $M \subset U$ von p nach q .

Beweis. Die Funktion f ist eine Nullform, und ihre äußere Ableitung ist $df = \text{grad } f \cdot d\vec{s}$. Der Rand von M ist $\{p, q\}$ mit Orientierung $+1$ bei p und -1 bei q . Also gilt nach dem Satz von Stokes

$$\int_M \text{grad } f \cdot d\vec{s} = \int_M df = \int_{\partial M} f = (+1) \cdot f(q) + (-1) \cdot f(p).$$

□

Satz 4.40 (Stoke'scher Integralsatz). Sei $U \subset \mathbb{R}^3$ offen und $v : U \rightarrow \mathbb{R}^3$ ein stetig differenzierbares Vektorfeld. Dann gilt

$$\int_M \text{rot } v \cdot d\vec{S} = \int_{\partial M} v \cdot d\vec{s}$$

für jede orientierte kompakte zweidimensionale Untermannigfaltigkeit $M \subset U$.

Beweis. Das Vektorfeld v definiert die Einsform $\omega = v \cdot d\vec{s}$, mit äußerer Ableitung $d\omega = \text{rot } v \cdot d\vec{S}$. Also gilt nach dem Satz von Stokes

$$\int_M \text{rot } v \cdot d\vec{S} = \int_M d(v \cdot d\vec{s}) = \int_{\partial M} v \cdot d\vec{s}.$$

□

Satz 4.41 (Gauß'scher Integralsatz). Sei $U \subset \mathbb{R}^3$ offen und $v : U \rightarrow \mathbb{R}^3$ ein stetig differenzierbares Vektorfeld. Dann gilt

$$\int_M \text{div } v \, dV = \int_{\partial M} v \cdot d\vec{S}$$

für jede orientierte kompakte dreidimensionale Untermannigfaltigkeit $M \subset U$.

Beweis. Das Vektorfeld v definiert die Zweiform $\omega = v \cdot d\vec{S}$ mit dem vektoriellen Flächenelement $d\vec{S} = (dx_2 \wedge dx_3, dx_3 \wedge dx_1, dx_1 \wedge dx_2)$, mit äußerer Ableitung $d\omega = \operatorname{div} v \, dV$ mit dem Volumenelement $dV := dx_1 \wedge dx_2 \wedge dx_3$. Also gilt nach dem Satz von Stokes

$$\int_M \operatorname{div} v \, dV = \int_M d(v \cdot d\vec{S}) = \int_{\partial M} v \cdot d\vec{S}.$$

□

Wir geben zwei Beispiele zum Gauß'schen Integralsatz.

Beispiel 4.42. Wir betrachten ein Vektorfeld $v : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, das wir uns als das Geschwindigkeitsfeld einer strömenden Flüssigkeit vorstellen (das heißt $v(x) \in \mathbb{R}^3$ gibt die Richtung und Geschwindigkeit $\|v(x)\|$ der Flüssigkeit an der Stelle x an). Gegeben eine Fläche F (d.h. eine orientierte zweidimensionale Untermannigfaltigkeit $F \subset \mathbb{R}^3$), so gibt $\int_F v \cdot d\vec{S}$ an, wie viel Flüssigkeit durch die Fläche F (pro Zeiteinheit) strömt. Das auf der Orientierung von F beruhende Vorzeichen von $\int_F v \cdot d\vec{S}$ gibt dabei an, ob die Strömungsbilanz als positiv oder negativ gerechnet wird, also gegen oder mit dem positiv orientierten Einheitsnormalenfeld von F strömt.

Ist zum Beispiel konkret

$$v(x, y, z) = \begin{pmatrix} x^2 - z \\ xy \\ -3zx \end{pmatrix}$$

und $F = \{(0, y, z) : y^2 + z^2 < 1\} = \{0\} \times D$ eine Einheitskreisscheibe in der (y, z) -Ebene, so haben wir die globale Karte $h : D \rightarrow \mathbb{R}^3$, $h(y, z) = (0, y, z)$, und damit (wegen $dh_1 = 0$, $dh_2 = dy$, $dh_3 = dz$)

$$\begin{aligned} & \int_F (v_1 \, dy \wedge dz + v_2 \, dz \wedge dx + v_3 \, dx \wedge dy) \\ &= \int_F ((x^2 - z) \, dy \wedge dz + xy \, dz \wedge dx - 3zx \, dx \wedge dy) \\ &= \int_D (x^2 - z) \, dy \, dz \\ &= \int_0^1 dr \int_0^{2\pi} d\alpha r (r^2 \cos^2 \alpha - r \sin \alpha) \\ &= \frac{1}{8}. \end{aligned}$$

Das positive Vorzeichen informiert uns darüber, dass die Strömung mehr in Richtung der positiv orientierten Normalen e_1 von F als entgegengesetzt verläuft.

Wie sieht es aus, wenn wir eine komplizierte Fläche betrachten, zB $F = S^2$ (sagen wir mit dem nach außen zeigenden Einheitsnormalenfeld) oder ein Torus $F = T$? Was erstmal nach etwas Rechnung aussieht, wird mit dem Satz von Gauß ganz leicht: $S^2 = \partial K$ ist

der Rand der orientierten berandeten kompakten dreidimensionalen Untermannigfaltigkeit $K = \{(x, y, z) : x^2 + y^2 + z^2 \leq 1\}$, dh $\int_{S^2} v \cdot d\vec{S} = \int_K \operatorname{div} v \, dV$. Wir müssen also nur die Divergenz über das dreidimensionale Volumen K integrieren. Wegen

$$\operatorname{div} v(x, y, z) = 2x + x - 3x = 0$$

ist dieses Integral Null. Anschaulich bedeutet das, dass aus K genauso viel Flüssigkeit pro Zeit hinein- wie hinausströmt. Das ist intuitiv klar, wenn v keine "Quellen" (oder "Senken") hat, was genau durch $\operatorname{div} v = 0$ ausgedrückt wird.

Beispiel 4.43. Als weiteres Beispiel betrachten wir das *Gauß'sche Gesetz*, eine der Maxwell-Gleichungen aus der Elektrodynamik. Es besagt, dass Ladungen die Quellen des elektrischen Feldes sind: Das elektrische Feld $E : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ und die Ladungsdichte $\rho : \mathbb{R}^3 \rightarrow \mathbb{R}$ sind durch die Gleichung

$$\operatorname{div} E = \rho$$

verknüpft (wir nehmen an, dass E stetig differenzierbar und damit ρ stetig ist). Eine häufige Fragestellung ist, E bei gegebenem ρ zu bestimmen. Wir wollen das für den Fall tun, dass sowohl ρ als auch E kugelsymmetrisch sind, dh nur von $\|x\|$ (mit $x = (x_1, x_2, x_3) \in \mathbb{R}^3$) abhängen, und $E(\|x\|) = f(\|x\|) \frac{x}{\|x\|}$ mit einer Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$. Weiterhin nehmen wir an, dass ρ kompakten Träger hat (endlich ausgedehnte Ladungskonfiguration), also $\rho(x) = 0$ für alle $\|x\| \geq R$ für ein groß genugenden Radius $R > 0$.

Mit $K_R := \{x \in \mathbb{R}^3 : \|x\| \leq R\}$ ist die Gesamtladung der Ladungskonfiguration (per Definition)

$$Q := \int_{\mathbb{R}^3} \rho(x) \, d\lambda_3(x) = \int_{K_R} \rho(x) \, d\lambda_3(x) = \int_{K_R} \operatorname{div} E(x) \, dV(x)$$

Wir betrachten K_R als kanonisch orientierte dreidimensionale kompakte Untermannigfaltigkeit von \mathbb{R}^3 . Nach dem Satz von Gauß haben wir dann

$$Q = \int_{\partial K_R} E \cdot d\vec{S}.$$

Dieses Integral bestimmen wir in Kugelkoordinaten. Dazu verwenden wir die in Beispiel 2.53 beschriebene Karte von ∂K_R , die bis auf Nullmengen ganz ∂K_R abdeckt,

$$h : \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \times (-\pi, \pi) \rightarrow \partial K_R,$$

$$h(\alpha, \beta) = R \begin{pmatrix} \cos \alpha \cos \beta \\ \cos \alpha \sin \beta \\ \sin \alpha \end{pmatrix}.$$

Eine Rechnung zeigt

$$Q = \int_{\partial K_R} E \cdot d\vec{S} = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} d\alpha \int_{-\pi}^{\pi} d\beta R^2 f(R) \cos \alpha d\alpha d\beta = 4\pi R^2 f(R),$$

also $f(R) = \frac{Q}{4\pi R^2}$.

4.5 Maßensoren und Integration von Funktionen über Untermannigfaltigkeiten

Unsere Überlegungen zum Satz von Stokes waren von einer Verallgemeinerung des Hauptsatzes der Differential- und Integralrechnung geleitet, der das Studium der alternierenden Multilinearformen und Differentialformen motiviert hat.

Aus der Perspektive der Maßtheorie können wir uns aber auch nach Maßen auf Untermannigfaltigkeiten fragen, und zwar insbesondere nach Maßen, die unserer geometrischen Anschauung von Länge einer Kurve (1-dimensionale Untermannigfaltigkeit), Flächeninhalt einer Fläche (2-dimensionale Untermannigfaltigkeit), oder allgemeiner nach einem “ n -dimensionalen Volumen” einer n -dimensionalen Untermannigfaltigkeit $M \subset \mathbb{R}^d$ entsprechen. Wir hatten in Def. 4.24 bereits festgelegt, was eine messbare Teilmenge $A \subset M$ einer n -dimensionalen Untermannigfaltigkeit $M \subset \mathbb{R}^d$ ist. Als Teilmenge von \mathbb{R}^d hat A aber bzgl. des d -dimensionalen Lebesguemaßes λ_d stets Maß $\lambda_d(A) = 0$, falls $n < d$. Wir müssen also ein anderes Maß konstruieren. Diese Überlegung wollen wir in diesem Abschnitt verfolgen.

Wir beginnen die Untersuchung des n -dimensionalen Volumens mit der linearen Situation. Dazu betrachten wir den Vektorraum \mathbb{R}^n , Vektoren $b_1, \dots, b_n \in \mathbb{R}^n$, und das von ihnen aufgespannte Parallelotop

$$P_B := \left\{ \sum_{k=1}^n t_k b_k : 0 \leq t_k \leq 1 \right\}.$$

ist (auch Spat oder Parallelepipid genannt). Wie in Beispiel 1.49 gezeigt wurde, ist das Lebesguemaß dieser (messbaren) Menge

$$\lambda_n(P_B) = |\det(B)| = \sqrt{\det(B^T B)},$$

wobei B die $(n \times n)$ -Matrix mit den Spaltenvektoren b_1, \dots, b_n ist. Die zweite Gleichung $|\det(B)| = \sqrt{\det(B^T B)}$ gilt für jede $(n \times n)$ -Matrix und wird sich noch als nützlich herausstellen.

Nun betrachten wir eine injektive lineare Abbildung $A : \mathbb{R}^n \rightarrow \mathbb{R}^d$, mit $n \leq d$. Wie sollten wir das n -dimensionale Volumen von $A(P_B)$ als Teilmenge des n -dimensionalen Untervektorraums $A\mathbb{R}^n \subset \mathbb{R}^d$ definieren? In Anbetracht der Rotationsinvarianz des Lebesguemaßes sollte diese Definition so getroffen werden, dass sie sich nicht ändert, wenn wir die Menge $A(P_B)$ im \mathbb{R}^d drehen, also zum Bild unter $S \in \text{SO}(d)$ übergehen. Wählen wir $S \in \text{SO}(d)$ so, dass $SA(P_B) \subset \mathbb{R}^n \times \{0, \dots, 0\}$ (mit $d - n$ vielen Nullen). Dann können wir auf $\mathbb{R}^n \subset \mathbb{R}^d$ projizieren und das Lebesguemaß im \mathbb{R}^n verwenden. Sei also $\pi : \mathbb{R}^d \rightarrow \mathbb{R}^n$ die orthogonale

Projektion in den kanonischen Basen (e_1, \dots, e_d) von \mathbb{R}^d und (f_1, \dots, f_n) von \mathbb{R}^n ist dies

$$\pi(e_i) = \begin{cases} f_i & 1 \leq i \leq n \\ 0 & i > n \end{cases}, \quad \pi = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & 0 & \dots & 0 \end{pmatrix}$$

Dann ist unsere Definition des n -dimensionalen Volumens von $A(P_B)$

$$\text{Vol}_n(A(P_B)) := \lambda_n(\pi S A(P_B)).$$

Wir erwarten, dass diese Definition nicht von S abhängt. Die Menge $\pi S A(P_B) \subset \mathbb{R}^n$ ist das von den Vektoren $\pi S A b_1, \dots, \pi S A b_n$ aufgespannte Parallelotop. Um zu einer handlicheren Formel für $\text{Vol}_n(A(P_B))$ zu gelangen, schreiben wir

$$\text{Vol}_n(A(P_B)) = |\det(\pi S A B)| = \sqrt{\det(B^T A^T S^T \pi^T \pi S A B)}.$$

Aufgrund der Definition von π und S gilt $\pi^T \pi v = v$ für alle v im Bild von S . Außerdem gilt $S^T S = 1$ wegen $S \in \text{SO}(d)$. Somit erhalten wir (beachten Sie $B, B^T, A^T A \in M_n$)

$$\begin{aligned} \text{Vol}_n(A(P_B)) &= \sqrt{\det(B^T A^T A B)} = \sqrt{\det(B^T) \det(A^T A) \det(B)} \\ &= \sqrt{\det(A^T A)} \cdot \lambda_n(P_B). \end{aligned}$$

Diese interessante Formel erklärt, wie sich das n -dimensionale Volumen des Parallelotops P_B unter der linearen Einbettung in den höherdimensionalen Raum \mathbb{R}^d verändert.

Wir bemerken noch, dass die Spalten a_1, \dots, a_n der Matrix A die Bilder $a_i = A f_i$ der Basisvektoren f_1, \dots, f_n von \mathbb{R}^n sind. Damit gilt, $i, j \in \{1, \dots, n\}$,

$$(A^T A)_{ij} = \sum_{k=1}^d A_{ik} A_{jk} = \langle a_i, a_j \rangle \quad (\text{Euklidisches Skalarprodukt im } \mathbb{R}^d).$$

Die Einträge der Matrix $A^T A$ informieren uns also über die Längen der Bilder der Basisvektoren ($\|a_i\| = \sqrt{(A^T A)_{ii}}$) und ihre relativen Winkel. Insbesondere ist $A^T A$ stets positiv definit: Für $v \in \mathbb{R}^n \setminus \{0\}$ gilt

$$\langle v, A^T A v \rangle = \|A v\|^2 > 0,$$

wobei die Ungleichung durch die Injektivität von A folgt.

Wir übertragen diese linearen Vorüberlegungen nun auf n -dimensionale Untermannigfaltigkeiten $M \subset \mathbb{R}^d$, wo wir sie als lokales Modell für nichtlineare Einbettungen (Karten) verwenden. Gegeben eine Karte $h : V \rightarrow U \subset M$ und einen Punkt $p \in U$, so erhalten wir eine Basis des Tangentialraums als $((\partial_1 h)(x), \dots, (\partial_n h)(x))$, dh die Jacobimatrix $(Dh)(x)$ (mit $x = h^{-1}(p)$) übernimmt die Rolle von A .

Definition 4.44. Sei $M \subset \mathbb{R}^d$ eine n -dimensionale Untermannigfaltigkeit und $h : V \subset \mathbb{R}^n \rightarrow U \subset M$ eine Karte. Bezüglich dieser Karte definieren wir den *Maßstensor* (oder *metrischen Tensor*) als die $(n \times n)$ -Matrix-wertige Funktion

$$G : V \rightarrow M_n(\mathbb{R}), \quad G(x) := (Dh)(x)^T (Dh)(x).$$

Die *Gramsche Determinante*^a von M bzgl. der Karte ist

$$g : V \rightarrow \mathbb{R}_+, \quad g(x) := \det G(x).$$

^aAuch die Bezeichnung $|g|$ für $\det G$ ist weit verbreitet.

Aus unseren Vorüberlegungen erhalten wir sofort:

Lemma 4.45. *Sei $h : V \rightarrow U \subset M$ eine Karte und $x \in V$. Dann ist der durch h definierte Maßstensor $G(x)$ positiv definit. Es gilt $\det G(x) > 0$ und $G_{ij}(x) = \langle (\partial_i h)(x), (\partial_j h)(x) \rangle$. Die Gramsche Determinante $g : V \rightarrow \mathbb{R}_+$ ist stetig.*

(Die Stetigkeit folgt, da die Karten C^1 sind).

Da der Maßstensor mit Hilfe einer willkürlich gewählten Karte definiert wurde, sollten wir prüfen, wie er sich unter Kartenwechseln verhält.

Lemma 4.46. *Seien $h : V \rightarrow U \subset M$ und $\tilde{h} : \tilde{V} \rightarrow \tilde{U} \subset M$ Karten mit zugehörigen Maßstensenoren G, \tilde{G} und Gram'schen Determinanten g, \tilde{g} . Sei weiterhin $p \in U \cap \tilde{U}$ und $x = h^{-1}(p)$ und $\psi : h^{-1}(U \cap \tilde{U}) \rightarrow \tilde{h}^{-1}(U \cap \tilde{U})$, $\psi = \tilde{h}^{-1} \circ h$ der Kartenwechsel. Dann gilt*

$$G(x) = (D\psi)(x)^T \tilde{G}(\psi(x)) (D\psi)(x), \quad g(x) = |\det(D\psi)(x)|^2 \cdot \tilde{g}(\psi(x)).$$

Beweis. Die Kettenregel besagt $(Dh)(x) = (D(\tilde{h} \circ \psi))(x) = (D\tilde{h})(\psi(x))(D\psi)(x)$ und somit

$$\begin{aligned} G(x) &= (Dh)(x)^T (Dh)(x) \\ &= (D\psi)(x)^T (D\tilde{h})(\psi(x))^T (D\tilde{h})(\psi(x)) (D\psi)(x) \\ &= (D\psi)(x)^T \tilde{G}(\psi(x)) (D\psi)(x). \end{aligned}$$

Determinantenbildung ergibt $g(x) = \det G(x) = (\det(D\psi)(x))^2 \cdot \tilde{g}(\psi(x))$. \square

Wir betrachten nun eine messbare Teilmenge $A \subset M$, die vorerst der Einfachheit halber in ein einziges Kartengebiet passen soll, dh es gebe eine Karte $h : U \rightarrow V \supset A$. Dann möchten wir das n -dimensionale Volumen von A als

$$\text{Vol}_n(A) := \int_V \chi_A(h(x)) \sqrt{g(x)} dx = \int_{h^{-1}(A)} \sqrt{g(x)} dx$$

definieren, wobei g die Gram'sche Determinante der Karte h ist. Dieses Lebesgueintegral (keine Differentialformen) über die Teilmenge $h^{-1}(A)$ des \mathbb{R}^n ist wohldefiniert, da $h^{-1}(A)$ messbar ist (dies ist die Definition der Messbarkeit von A) und $x \mapsto \sqrt{g(x)}$ stetig (da g stetig und positiv), also messbar ist. Weiterhin gilt $\text{Vol}_n(A) \geq 0$.

Außerdem hängt diese Definition nicht von der Wahl der Karte ab: Ist $\tilde{h} : \tilde{U} \rightarrow \tilde{V} \supset A$ eine weitere Karte, die A in ihrem Bildbereich enthält, so ist aufgrund von Lemma 4.46 und der Transformationsformel (Satz 2.51)

$$\int_{h^{-1}(A)} \sqrt{g(x)} dx = \int_{\psi^{-1}(\tilde{h}^{-1}(A))} |\det(D\psi)(x)| \sqrt{\tilde{g}(\psi(x))} dx = \int_{\tilde{h}^{-1}(A)} \sqrt{\tilde{g}(x)} dx$$

mit der Gramschen Determinante \tilde{g} von \tilde{h} .

Bis auf die Einschränkung, dass A in ein Kartengebiet passen muss, handelt es sich also um eine vernünftige Begriffsbildung. Wir besprechen zunächst zwei Beispiele.

Beispiel 4.47.

- a) Ist $\gamma : (a, b) \rightarrow \mathbb{R}^d$ eine glatte Kurve, so können wir sie als 1-Mannigfaltigkeit C mit der Karte γ ansehen. In diesem Fall ist die Gramsche Determinante gleich der (1×1) -Matrix

$$G(t) = \langle \gamma'(t), \gamma'(t) \rangle = \|\gamma'(t)\|^2,$$

also $g(t) = \|\gamma'(t)\|$. Damit wird das 1-dimensionale Volumen der Kurve gleich

$$\text{Vol}_1(C) = \int_a^b \|\gamma'(t)\| dt,$$

was wir bereits in Analysis 2 als die Länge einer Kurve kennengelernt haben (Ana. 2, Definition 11.5).

- b) Testen wir unsere Definition an einer Kugeloberfläche und betrachten die Halbkugel $H = \{(x, y, z) : x^2 + y^2 + z^2 = r^2, z > 0\}$ mit der Karte $h : (0, \frac{\pi}{2}) \times (-\pi, \pi) \rightarrow \mathbb{R}^3$, $h(\alpha, \beta) = (r \cos \alpha \cos \beta, r \cos \alpha \sin \beta, r \sin \alpha)$. Dann ist

$$(\partial_\alpha h)(\alpha, \beta)^T = (-r \sin \alpha \cos \beta, -r \sin \alpha \sin \beta, r \cos \alpha),$$

$$(\partial_\beta h)(\alpha, \beta)^T = (-r \cos \alpha \sin \beta, r \cos \alpha \cos \beta, 0),$$

und damit der Maßtensor und die Gramsche Determinante

$$G(\alpha, \beta) = \begin{pmatrix} r^2 & 0 \\ 0 & r^2 \cos^2 \alpha \end{pmatrix}, \quad g(\alpha, \beta) = r^4 \cos^2 \alpha.$$

Unsere Definition des zweidimensionalen Volumens der Halbkugeloberfläche ergibt somit

$$\text{Vol}_2(H) = \int_0^{\frac{\pi}{2}} d\alpha \int_{-\pi}^{\pi} d\beta r^2 \cos \alpha = 2\pi r^2 \int_0^{\pi/2} \cos \alpha d\alpha = 2\pi r^2.$$

Dies ist genau die Hälfte des aus der Schule bekannten Flächeninhalts $4\pi r^2$ der Oberfläche einer Kugel mit Radius r .

Für eine beliebige messbare Teilmenge $A \subset M$ zerlegen wir wieder in jeweils in ein Kartengebiet passende Teile. Wie bei Satz 4.25 beschränken wir uns der Einfachheit halber wieder auf einfache Untermannigfaltigkeiten, die in endlich viele paarweise disjunkte meßbare Teilmengen A_i zerlegt werden können, die jeweils in nur einem Kartengebiet U_i (einer Karte $h_i : V_i \rightarrow U_i \subset M$) enthalten sind.

Definition und Satz 4.48. Sei $M \subset \mathbb{R}^d$ eine n -dimensionale einfache Untermannigfaltigkeit und $A \subset M$ messbar. Das n -dimensionale Volumen von A ist wie folgt definiert: Sei $A = \bigsqcup_i A_i$ eine Zerlegung in paarweise disjunkte meßbare Teilmengen A_i und $h_i : V_i \rightarrow U_i \supset A_i$ Karten. Dann setzen wir

$$\text{Vol}_n(A) := \mu_M(A) := \sum_i \int_{h_i^{-1}(A \cap A_i)} \sqrt{g_i}$$

mit der Gram'schen Determinante g_i von h_i .

- Diese Definition hängt nicht von der gewählten Zerlegung $M = \bigsqcup_i A_i$ und den gewählten Karten ab.
- Die Menge aller messbarer Mengen auf M bildet eine σ -Algebra Σ_M , und $\mu_M : \Sigma_M \rightarrow [0, \infty]$ ist ein Maß, das *Oberflächenmaß von M* .

Beweis. a) Seien $M = \bigsqcup_i A_i = \bigsqcup_j \tilde{A}_j$ zwei Zerlegungen und $h_i : V_i \rightarrow U_i \supset A_i$, $\tilde{h}_j : \tilde{V}_j \rightarrow \tilde{U}_j \supset \tilde{A}_j$ passende Karten. Dann gilt mit den Kartenwechseln ψ_{ij} von h_i zu \tilde{h}_j

$$\begin{aligned} \sum_i \int_{h_i^{-1}(A \cap A_i)} \sqrt{g_i} &= \sum_{i,j} \int_{h_i^{-1}(A \cap A_i \cap \tilde{A}_j)} \sqrt{g_i} \\ &= \sum_{i,j} \int_{\tilde{h}_j^{-1}(A \cap A_i \cap \tilde{A}_j)} \sqrt{\tilde{g}_j} \\ &= \sum_j \int_{\tilde{h}_j^{-1}(A \cap \tilde{A}_j)} \sqrt{\tilde{g}_j}. \end{aligned}$$

b) Offenbar enthält Σ_M die leere Menge. Sei $(U_i)_{i \in I}$ eine Überdeckung von M durch Kartengebiete und $A \in \Sigma_M$, dh $h_i^{-1}(U_i \cap A)$ ist eine Lebesguemenge für alle $i \in I$. Dann ist auch $h_i^{-1}(U_i \cap A^c) = V_i \setminus h_i^{-1}(U_i \cap A)$ eine Lebesguemenge, da die V_i offen und damit messbar sind. Ist $(A_n)_{n \in \mathbb{N}}$ eine abzählbare Menge von messbaren Mengen, so ist auch $\bigcup_n A_n$ messbar, denn $h_i^{-1}(U_i \cap \bigcup_n A_n) = \bigcup_n h_i^{-1}(U_i \cap A_n)$. Also ist Σ_M eine σ -Algebra.

Offenbar bildet Vol_n messbare Mengen nach $[0, \infty]$ ab und erfüllt $\text{Vol}_n(\emptyset) = 0$. Auch die σ -Additivität erhalten wir vom Lebesguemaß: Sind (A_n) paarweise disjunkte messbare Mengen auf M und $M = \bigsqcup_i B_i$ eine disjunkte messbare Zerlegung, so haben wir

$$\begin{aligned} \mu_M\left(\bigsqcup_n A_n\right) &= \sum_i \int_{h_i^{-1}(B_i \cap \bigsqcup_n A_n)} \sqrt{g_i} \\ &= \sum_i \sum_n \int_{h_i^{-1}(B_i \cap A_n)} \sqrt{g_i} \\ &= \sum_n \mu_M(A_n). \end{aligned}$$

Also ist μ_M ein Maß. □

Das Maß μ_M ist so konstruiert, dass für eine Bewegung φ von \mathbb{R}^d und messbares $A \subset M$ stets $\mu_M(A) = \mu_{\varphi(M)}(\varphi(A))$ gilt (Übung). Weiterhin betrachten wir, analog zum Lebesguemaß, das Verhalten des n -dimensionalen Volumens unter Skalierungen.

Lemma 4.49. Sei $M \subset \mathbb{R}^d$ eine n -dimensionale Untermannigfaltigkeit und $r > 0$. Dann ist auch $rM := \{r \cdot x : x \in M\}$ eine n -dimensionale Untermannigfaltigkeit von \mathbb{R}^d . Für jede messbare Menge $A \subset M$ ist auch $rA \subset rM$ messbar, und es gilt

$$\mu_{rM}(rA) = r^n \mu_M(A).$$

Beweis. Ist $h : V \rightarrow U \subset M$ eine Karte von M , so ist $h_r : V \rightarrow rU \subset rM$, $h_r(x) := r \cdot h(x)$, eine Karte von rM . Man zeigt dann leicht, dass rM eine n -dimensionale Untermannigfaltigkeit ist. Die Maßstensoren bzgl h und h_r sind wegen $(Dh_r)(x) = r(Dh)(x)$ durch

$$G_r(x) = (Dh_r)(x)^T (Dh_r)(x) = r^2 (Dh)(x)^T (Dh)(x) = r^2 G(x)$$

verbunden. Für die Gram'schen Determinanten folgt

$$g_r(x) = \sqrt{\det(r^2 G(x))} = \sqrt{r^{2n} \det(G(x))} = r^n g(x).$$

Dies impliziert die Behauptung. □

Damit hat $\text{Vol}_n = \mu_M$ alle Eigenschaften, die wir von einem n -dimensionalen Volumen erwarten.

Das Maß Vol_n auf M wird oft auch S (für "surface") genannt, wobei die Abhängigkeit von M unterdrückt wird. Da (M, Σ_M, μ_M) ein Maßraum ist, haben wir gemäß Abschnitt 2.2 auch einen Integralbegriff $\int_M f d\mu_M$ für Borel-messbare numerische Funktionen $f : M \rightarrow \overline{\mathbb{R}}$. Dieser Integralbegriff kommt ohne Orientierungen und Differentialformen aus, liefert allerdings auch keinen Hauptsatz der Differential- und Integralrechnung.

Explizit können wir dieses Integral wie folgt beschreiben.

Satz 4.50. Sei $f : M \rightarrow \overline{\mathbb{R}}$ eine numerische Funktion und $M = \bigsqcup_i A_i$ eine disjunkte Zerlegung von M mit zugehörigen Karten $h_i : V_i \rightarrow U_i \supset A_i$. Dann ist f genau dann bzgl des Maßes μ_M integrierbar, falls für jedes i die Funktion

$$h_i^{-1}(A_i) \rightarrow \overline{\mathbb{R}}, \quad x \mapsto f(h_i(x)) \sqrt{g_i(x)}$$

(mit der Gram'schen Determinante g_i von h_i) Lebesgue-integrierbar ist. In diesem Fall gilt

$$\int_M f d\mu_M = \sum_i \int_{h_i^{-1}(A_i)} f(h_i(x)) \sqrt{g_i(x)} dx.$$

Beweis. Es genügt, die Aussage für nicht-negative Funktionen mit Träger in einem A_i zu beweisen (sonst Zerlegung). Für die charakteristische Funktion $f = \chi_A$ einer messbaren Menge ist die Aussage per Definition von μ_M richtig, was sich sogleich auf Treppenfunktionen überträgt. Der allgemeine Fall folgt per Approximation (monotone Konvergenz). □

Beispiel 4.51. Wir betrachten den ellipsenförmigen Zylinder

$$Z = \{(x, y, z) \in \mathbb{R}^3 : a^2 x^2 + y^2 = r^2\}$$

mit $0 < a < 1$ und $r > 0$, und die Funktion $f : Z \rightarrow \mathbb{R}$, $f(x, y, z) = \frac{x^2}{1+z^2}$. Um $\int_Z f d\mu_Z$ zu bestimmen, betrachten wir die Karte

$$h : (-\pi, \pi) \times \mathbb{R} \rightarrow Z, \quad h(\alpha, z) := \begin{pmatrix} \frac{r}{a} \cos \alpha \\ r \sin \alpha \\ z \end{pmatrix}.$$

Das Bild von h unterscheidet sich von Z nur um eine Nullmenge, und wir haben $g(\alpha, z) = r^2(a^{-2} \sin^2 \alpha + \cos^2 \alpha)$ für alle α und z . Also

$$\begin{aligned} \int_Z f d\mu_Z &= \int_{-\pi}^{\pi} d\alpha \int_{\mathbb{R}} dz \frac{\frac{r^2}{a^2} \cos^2 \alpha}{1+z^2} r \sqrt{a^{-2} \sin^2 \alpha + \cos^2 \alpha} \\ &= \frac{2\pi r^3}{a^2} \int_{-\pi}^{\pi} d\alpha \cos^2 \alpha \sqrt{a^{-2} \sin^2 \alpha + \cos^2 \alpha} \end{aligned}$$

Wir haben nun zwei Integralbegriffe für Untermannigfaltigkeiten: Zum einen den Integralbegriff für Differentialformen über orientierte Untermannigfaltigkeiten, und zum anderen den Integralbegriff für Funktionen über nicht notwendigerweise orientierte Untermannigfaltigkeiten. Wir wollen diese Begriffe nun vergleichen.

Dazu betrachten wir eine Hyperfläche, also eine $(d-1)$ -dimensionale Untermannigfaltigkeit $M \subset \mathbb{R}^d$.

Satz 4.52. Sei $M \subset \mathbb{R}^d$ eine orientierte Hyperfläche und $\nu : M \rightarrow \mathbb{R}^d$ das zugehörige stetige positiv orientierte Einheits-Normalenfeld. Sei $O \subset \mathbb{R}^d$ eine M enthaltende offene Menge und $v : O \rightarrow \mathbb{R}^d$ ein stetiges Vektorfeld mit kompaktem Träger. Dann gilt

$$\int_M v(x) \cdot d\vec{S}(x) = \int_M \langle v(x), \nu(x) \rangle dS(x).$$

Hierbei ist auf der linken Seite $d\vec{S} = (dS_1, \dots, dS_d)$, $dS_k := (-1)^{k-1} dx_1 \wedge \dots \wedge \widehat{dx_k} \wedge \dots \wedge dx_d$ und $v(x) \cdot d\vec{S}(x) = \sum_{k=1}^d v_k(x) dS_k$ eine $(d-1)$ -Form auf O , und das Integral \int_M ist im Sinne von Differentialformen über die orientierte Untermannigfaltigkeit M zu lesen.

Auf der rechten Seite ist das Integral als Integral der numerischen Funktion $M \ni x \mapsto \langle v(x), \nu(x) \rangle \in \mathbb{R}$ bzgl. des Oberflächenmaß $S = \mu_M$ von M zu lesen.

Formal merkt man sich diesen Satz am besten als “ $d\vec{S} = \nu dS$ ”, wobei $d\vec{S}$ das vektorielle Flächenelement und S das Oberflächenmaß sind.

In dem Beweis des Satzes werden wir auf zwei Determinanten stoßen, für die wir ein Lemma vorbereiten.

Lemma 4.53. Sei $\mathbf{1}$ die $(n \times n)$ -Einheitsmatrix und $a = (a_1, \dots, a_n)^T, b = (b_1, \dots, b_n)^T \in \mathbb{R}^n$ zwei Vektoren (also $(n \times 1)$ -Matrizen). Dann gilt

$$\det(\mathbf{1} + ab^T) = \det \begin{pmatrix} \mathbf{1} & a \\ -b^T & 1 \end{pmatrix} = 1 + \langle a, b \rangle.$$

Beweis. Durch Matrixmultiplikation überprüft man

$$\begin{pmatrix} \mathbf{1} + ab^T & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{1} & 0 \\ -b^T & 1 \end{pmatrix} = \begin{pmatrix} \mathbf{1} & a \\ -b^T & 1 \end{pmatrix},$$

$$\begin{pmatrix} \mathbf{1} & 0 \\ b^T & 1 \end{pmatrix} \begin{pmatrix} \mathbf{1} + ab^T & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{1} & 0 \\ -b^T & 1 \end{pmatrix} = \begin{pmatrix} \mathbf{1} & a \\ 0 & 1 + \langle a, b \rangle \end{pmatrix}.$$

Beachten Sie, dass diese Matrizen auf Vektoren der Form $(x, t)^T$, mit $x \in \mathbb{R}^n$ und $t \in \mathbb{R}$, wie beispielsweise in

$$\begin{pmatrix} \mathbf{1} & 0 \\ b^T & 1 \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix} = \begin{pmatrix} x \\ \langle b, x \rangle + t \end{pmatrix}$$

wirken. So zeigt man die oben angeschriebenen Matrixgleichungen.

Bildet man in der ersten (zweiten) Gleichung Determinanten, folgt die erste (zweite) behauptete Gleichung. \square

Beweis. Da beide Integrale additiv bzgl. einer Zerlegung von M in messbare Teilmengen sind, dürfen wir annehmen, dass M Graph einer Funktion ist: Wir dürfen also voraussetzen, dass es offenes $V \subset \mathbb{R}^{d-1}$ und eine stetig differenzierbare Funktion $m : V \rightarrow \mathbb{R}$ gibt, so dass

$$M = \{(\mathbf{x}, x_d) : \mathbf{x} \in V, x_d = m(\mathbf{x})\}.$$

Wir haben dann die globale Karte

$$h : V \rightarrow M, \quad h(\mathbf{x}) = (\mathbf{x}, m(\mathbf{x}))$$

und betrachten die Orientierung von M , in der h positiv orientiert ist.

Um das Integral auf der linken Seite über die $(d-1)$ -Form $\omega := v \cdot d\vec{S}$ zu berechnen, bestimmen wir für $k < d$

$$\begin{aligned} h^*(dS_k) &= (-1)^{k-1} dx_1 \wedge \dots \wedge \widehat{dx_k} \wedge \dots \wedge dm(\mathbf{x}) \\ &= (-1)^{k-1} dx_1 \wedge \dots \wedge \widehat{dx_k} \wedge \dots \wedge \frac{\partial m}{\partial x_k} dx_k \\ &= (-1)^d \frac{\partial m}{\partial x_k} dx_1 \wedge \dots \wedge dx_{d-1}, \end{aligned}$$

und für $k = d$

$$h^*(dS_d) = (-1)^{d-1} dx_1 \wedge \dots \wedge dx_{d-1}.$$

Damit ist der Pullback von ω

$$h^*\omega = (-1)^{d-1} \left(- \sum_{k=1}^d (v_k \circ h) \partial_k m + v_d \circ h \right) dx_1 \wedge \dots \wedge dx_{d-1}.$$

Also ist das Integral auf der linken Seite

$$\begin{aligned} \int_M \omega &= \int_V h^*\omega = (-1)^{d-1} \int_V \left(- \sum_{k=1}^d v_k(h(\mathbf{x})) \partial_k m(\mathbf{x}) + v_d(h(\mathbf{x})) \right) d\mathbf{x} \\ &= (-1)^{d-1} \int_V \left\langle v(h(\mathbf{x})), \begin{pmatrix} -\text{grad}m(\mathbf{x}) \\ 1 \end{pmatrix} \right\rangle d\mathbf{x}. \end{aligned} \quad (\star)$$

Nun betrachten wir das Integral auf der rechten Seite der behaupteten Gleichung. Wie in Beispiel 4.31 gezeigt, hat M genau die beiden stetigen Einheitsnormalenfelder

$$\nu(\mathbf{x}, m(\mathbf{x})) = \pm \frac{1}{\sqrt{1 + \|\text{grad}m(\mathbf{x})\|^2}} \begin{pmatrix} -\text{grad}m(\mathbf{x}) \\ 1 \end{pmatrix}.$$

Das Vorzeichen muss so gewählt werden, dass ν positiv orientiert ist; zu diesem Punkt kommen wir bald. Der metrische Tensor von h hat wegen $\partial_i h(\mathbf{x}) = e_i + \partial_i m(\mathbf{x}) \cdot e_d$ die Einträge

$$G_{ij}(\mathbf{x}) = \langle e_i + \partial_i m(\mathbf{x}) \cdot e_d, e_j + \partial_j m(\mathbf{x}) \cdot e_d \rangle = \delta_{ij} + \partial_i m(\mathbf{x}) \partial_j m(\mathbf{x}),$$

was mit dem (i, j) -Eintrag der Matrix $\mathbf{1} + (\text{grad}m(\mathbf{x})) \cdot (\text{grad}m(\mathbf{x}))^T$ übereinstimmt. Also können wir Lemma 4.53 verwenden und erhalten die Gram'sche Determinante

$$g(\mathbf{x}) = \det G(\mathbf{x}) = 1 + \|\text{grad}m(\mathbf{x})\|^2.$$

Also ist das Integral auf der rechten Seite im Satz gleich

$$\begin{aligned} \int_V \langle \nu(\mathbf{x}, m(\mathbf{x})), \nu(\mathbf{x}, m(\mathbf{x})) \rangle \sqrt{1 + \|\text{grad}m(\mathbf{x})\|^2} d\mathbf{x} \\ = \pm \int_V \left\langle v(h(\mathbf{x})), \begin{pmatrix} -\text{grad}m(\mathbf{x}) \\ 1 \end{pmatrix} \right\rangle d\mathbf{x}. \end{aligned}$$

Dies stimmt mit dem linken Integral (\star) überein, falls das Vorzeichen \pm als $(-1)^{d-1}$ gewählt wird.

Wir müssen noch prüfen, ob diese Vorzeichenwahl mit der durch die Orientierung fixierte Vorzeichenwahl übereinstimmt. Dazu müssen wir gemäß Definition 4.29 bestimmen, für welches Vorzeichen die Basis $(\nu(\mathbf{x}, m(\mathbf{x})), \partial_1 h(\mathbf{x}), \dots, \partial_{d-1} h(\mathbf{x}))$ positiv orientiert ist. Wir berechnen dazu

$$\partial_1 h(\mathbf{x}) = (1, 0, \dots, 0, \partial_1 m(\mathbf{x})), \quad \partial_2 h(\mathbf{x}) = (0, 1, 0, \dots, \partial_2 m(\mathbf{x})),$$

etc, und erhalten mit $m_k := \partial_k m(\mathbf{x})$

$$\begin{aligned} \det(\nu(\mathbf{x}, m(\mathbf{x})), \partial_1 h(\mathbf{x}), \dots, \partial_{d-1} h(\mathbf{x})) \\ = \pm \frac{1}{\sqrt{1 + \|\text{grad}m(\mathbf{x})\|^2}} \det \begin{pmatrix} -m_1 & 1 & & & \\ -m_2 & & 1 & & \\ \vdots & & & \ddots & \\ -m_{d-1} & & & & 1 \\ 1 & m_1 & \dots & & m_{d-1} \end{pmatrix}. \end{aligned}$$

Die Determinante der Matrix in dieser Formel ist

$$(-1)^{d-1} \det \begin{pmatrix} \mathbf{1} & -\text{grad}m(\mathbf{x}) \\ (\text{grad}m(\mathbf{x}))^T & 1 \end{pmatrix} = (-1)^{d-1} (1 + \|\text{grad}m(\mathbf{x})\|^2),$$

wie man durch Sortieren der ersten Spalte nach ganz hinten und Anwendung von Lemma 4.53 verifiziert.

Damit haben wir

$$\det(\nu(\mathbf{x}, m(\mathbf{x})), \partial_1 h(\mathbf{x}), \dots, \partial_{d-1} h(\mathbf{x})) = \pm (-1)^{d+1} \sqrt{1 + \|\text{grad}m(\mathbf{x})\|^2}.$$

Wir müssen das Vorzeichen von ν also als $(-1)^{d+1}$ wählen, und somit stimmt die behauptete Gleichung auch im Vorzeichen. \square

Mit diesem Ergebnis betrachten wir noch einmal die Integralsätze von Gauß und Stokes, die wir nun auch in ihrer "klassischen Version", dh ohne Differentialformen, angeben können.

Satz 4.54 (Gauß'scher Integralsatz – klassische Version in allgemeiner Dimension).

Sei $O \subset \mathbb{R}^d$ offen, $M \subset O$ eine kompakte d -dimensionale Untermannigfaltigkeit mit ihrer kanonischen Orientierung, und ν das stetige nach außen zeigende Einheitsnormalenfeld des Randes ∂M . Dann gilt für jedes stetig differenzierbare Vektorfeld $v : O \rightarrow \mathbb{R}^d$

$$\int_M \text{div} v \, d\lambda_d = \int_{\partial M} \langle v, \nu \rangle \, d\mu_{\partial M}.$$

Beweis. In Satz 4.41 hatten wir den Gauß'schen Integralsatz in seiner Differentialformenversion der Einfachheit halber nur für $d = 3$ angegeben. Da die $(d-1)$ -Form $\omega := v \cdot d\vec{S}$, mit $dS_k = (-1)^{k-1} dx_1 \wedge \dots \wedge \widehat{dx_k} \wedge \dots \wedge dx_d$, äußere Ableitung

$$\begin{aligned} d\omega &= \sum_{k=1}^d (-1)^{k-1} dv_k \wedge dx_1 \wedge \dots \wedge \widehat{dx_k} \wedge \dots \wedge dx_d \\ &= \sum_{k=1}^d (-1)^{k-1} \frac{\partial v_k}{\partial x_k} dx_k \wedge dx_1 \wedge \dots \wedge \widehat{dx_k} \wedge \dots \wedge dx_d \\ &= \sum_{k=1}^d \frac{\partial v_k}{\partial x_k} dx_1 \wedge \dots \wedge dx_d \\ &= \text{div} v \, dx_1 \wedge \dots \wedge dx_d \end{aligned}$$

hat, gilt nach dem Satz von Stokes

$$\int_M \text{div} v \, dx_1 \wedge \dots \wedge dx_d = \int_{\partial M} v \cdot d\vec{S},$$

wenn ∂M die von der kanonischen Orientierung von M induzierte Orientierung trägt. Dies ist die Orientierung, die von dem nach außen zeigenden Einheitsnormalenfeld $\nu : \partial M \rightarrow \mathbb{R}^d$ gegeben ist. Nach Satz 4.52 folgt also $\int_{\partial M} v \cdot d\vec{S} = \int_{\partial M} \langle v, \nu \rangle d\mu_{\partial M}$. Das Integral auf der linken Seite ist gleich $\int_M \text{div} v \, d\lambda_d$ per Definition des Integrals über Differentialformen. \square

Für Integration von 1-Formen über eindimensionale orientierte Untermannigfaltigkeiten (Kurven) $C \subset \mathbb{R}^d$ hatten wir bereits eine differentialformenfreie Beschreibung kennengelernt (Beispiel 4.26). Die Orientierung von C entspricht einem Durchlaufsinne der Kurve. Ist $\gamma : (a, b) \rightarrow C$ eine globale Karte, so liegt $\gamma'(t)$ tangential on C im Punkt $\gamma(t)$. Falls $\gamma'(t)$ in die Richtung der Orientierung von C zeigt (also positiv orientiert ist), gilt für ein auf einer offenen Umgebung von C definiertes stetiges Vektorfeld v

$$\int_C v \cdot d\vec{s} = \int_a^b \langle v(\gamma(t)), \gamma'(t) \rangle dt.$$

Die Gram'sche Determinante ist hier $g(t) = \|\gamma'(t)\|$ (siehe Beispiel 4.47). Definieren wir ein normiertes positiv orientiertes Tangentenfeld $\tau := \gamma'/\|\gamma'\|$, so erhalten wir

$$\int_C v \cdot d\vec{s} = \int_C \langle v, \tau \rangle d\mu_C.$$

Satz 4.55 (Stoke'scher Integralsatz – klassische Version).

Sei $O \subset \mathbb{R}^3$ offen, $M \subset O$ eine orientierte kompakte zweidimensionale Untermannigfaltigkeit mit positiv orientiertem stetigen Einheitsnormalenfeld ν , und $\tau : \partial M \rightarrow \mathbb{R}^3$ das positiv orientierte normierte Tangentenfeld. Dann gilt für jedes stetig differenzierbare Vektorfeld $v : O \rightarrow \mathbb{R}^3$

$$\int_M \langle \text{rot } v, \nu \rangle d\mu_M = \int_{\partial M} \langle v, \tau \rangle d\mu_{\partial M}.$$

Beweis. Nach Satz 4.52 gilt $\int_M \langle \text{rot } v, \nu \rangle d\mu_M = \int_M \text{rot } v \cdot d\vec{S}$, was nach Satz 4.40 mit $\int_{\partial M} v \cdot d\vec{s}$ übereinstimmt, wenn ∂M die von M induzierte Orientierung trägt. Da das Tangenteinheitsfeld τ positiv orientiert ist, stimmt dieses Integral wiederum mit $\int_{\partial M} \langle v, \tau \rangle d\mu_{\partial M}$ überein, was den Beweis abschließt. \square

In dieser Form werden die Integralsätze oft verwendet. Wir geben ein Beispiel.

Beispiel 4.56. In der Magnetostatik interessiert man sich für die Frage, was für ein Magnetfeld (stetig differenzierbares Vektorfeld $B : \mathbb{R}^3 \rightarrow \mathbb{R}^3$) von einer gegebenen Stromdichte (stetiges Vektorfeld $j : \mathbb{R}^3 \rightarrow \mathbb{R}^3$) erzeugt wird. Es ist bekannt, dass das Magnetfeld den Gleichungen

$$\begin{aligned} \text{div } B &= 0, \\ \text{rot } B &= j \end{aligned}$$

genügt. Wir betrachten als Beispiel einen unendlich langen stromdurchflossenen Draht, also

$$j(x_1, x_2, x_3) = f((x_1^2 + x_2^2)^{1/2}) \cdot e_3$$

mit einer stetigen Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ mit Träger $\text{supp } f \subset [0, r]$; hierbei ist $r > 0$ der Durchmesser des Drahtes.

Um das Magnetfeld B außerhalb des Drahtes zu bestimmen, bemerken wir, dass für $B = \operatorname{rot} A$ mit einem C^2 -Vektorfeld $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ automatisch $\operatorname{div} B = \operatorname{div}(\operatorname{rot} A) = 0$ folgt. In Differentialformensprache betrachten wir $\omega := B \cdot d\vec{S}$, dann gilt $d\omega = \operatorname{div}(B)dV = 0$. Also ist ω geschlossen. Da wir einen sternförmigen Definitionsbereich haben, ist nach dem Poincaré-Lemma ω exakt, dh $\omega = d\psi$ für eine 1-Form ψ . In drei Dimensionen sind 1-Formen von der Form $\psi = A \cdot d\vec{s}$, also $d\psi = \operatorname{rot} A \cdot d\vec{S}$. Deshalb ist das Magnetfeld notwendigerweise von der Form $B = \operatorname{rot} A$ (man nennt A das Vektorpotential).

Die Symmetrie der Stromverteilung j legt nahe, dass A nur von $\sqrt{x_1^2 + x_2^2}$ abhängen und nur in x_3 -Richtung zeigen sollte. Ohne dies genau zu begründen, suchen wir deshalb B in der Form $B = \operatorname{rot} A$ mit $A(x_1, x_2, x_3) = a(\sqrt{x_1^2 + x_2^2})e_3$. Damit erhalten wir

$$B(x_1, x_2, x_3) = -\frac{a'((x_1^2 + x_2^2)^{1/2})}{(x_1^2 + x_2^2)^{1/2}} \begin{pmatrix} -x_2 \\ x_1 \\ 0 \end{pmatrix}.$$

Um die Funktion a zu bestimmen, betrachten wir die Scheibe $D_R := \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1^2 + x_2^2 \leq R^2, x_3 = 0\}$ mit Radius $R > r$. Dies ist eine berandete zweidimensionale Untermannigfaltigkeit von \mathbb{R}^3 , die wir durch das konstante Einheitsnormalenfeld $\nu = e_3$ orientieren. Ihr Rand ist dann durch die Parametrisierung $\gamma : (0, 2\pi) \rightarrow \partial D_R$, $t \mapsto (R \cos t, R \sin t)$ positiv orientiert. Wir wenden den Satz von Stokes an und erhalten

$$\begin{aligned} \int_{\partial D_R} \langle B, \gamma' \rangle d\mu_{\partial D_R} &= - \int_0^{2\pi} \left\langle \frac{a'(R)}{R} \begin{pmatrix} -R \sin t \\ R \cos t \\ 0 \end{pmatrix}, \begin{pmatrix} -R \sin t \\ -R \cos t \\ 0 \end{pmatrix} \right\rangle dt \\ &= -2\pi R a'(R), \end{aligned}$$

sowie

$$\begin{aligned} \int_{D_R} \langle \operatorname{rot} B, \nu \rangle d\mu_{D_R} &= \int_{D_R} \langle j, \nu \rangle d\mu_{D_R} \\ &= \int_{D_R} f((x_1^2 + x_2^2)^{1/2}) dx_1 dx_2 \\ &= \int_0^{2\pi} d\alpha \int_0^r d\rho \rho f(\rho) \\ &=: 2\pi I. \end{aligned}$$

Nach dem Satz von Stokes ergibt sich $a'(R) = -\frac{I}{R}$, also

$$B(x_1, x_2, x_3) = \frac{I}{(x_1^2 + x_2^2)^{3/2}} \begin{pmatrix} -x_2 \\ x_1 \\ 0 \end{pmatrix}, \quad x_1^2 + x_2^2 > r^2.$$

Dieser Sachverhalt ist als Biot-Savart Gesetz bekannt.

Eine Riemannsche Mannigfaltigkeit ist eine Mannigfaltigkeit M , für die an jedem Punkt $p \in M$ eine positiv definite symmetrische Bilinearform $g_p : T_p M \times T_p M \rightarrow \mathbb{R}$ erklärt ist, die differenzierbar von p abhängt. Dies ist genau das, was im Konkreten von dem metrischen Tensor G geleistet wird.

Eine Riemannsche Mannigfaltigkeit (M, g) ist insbesondere ein metrischer Raum: Der metrische Tensor erlaubt es, über Längen von in M verlaufenden Kurven zu sprechen. Gegeben zwei Punkte $p, q \in M$ definiert man $d(p, q)$ dann als Infimum über die Längen aller stückweise differenzierbarer Wege in M von p nach q . Insbesondere kann man dann sinnvoll nach Wegen kürzester Länge von p nach q in M fragen, den sogenannten Geodäten. Diese lassen sich als Lösungen von gewissen Differentialgleichungen bestimmen.

Eine wichtige Verallgemeinerung von Riemannschen Mannigfaltigkeiten erhält man, wenn die Bilinearform g_p nicht positiv definit, sondern eine andere Signatur hat. Dies beinhaltet insbesondere vierdimensionale Mannigfaltigkeiten mit Signatur $(-+++)$, die als Modelle von Raumzeiten in der Allgemeinen Relativitätstheorie von zentraler Bedeutung sind.