

Analysis 3

Dirk Lorenz

Inhaltsverzeichnis

Vorbemerkung	3
1 Einleitung und Motivation - Das Maßproblem	5
2 Mengenalgebren und -ringe	10
3 σ -Algebren	14
4 Inhalte, Prämaße und Maße	19
5 Das Lebesguesche Prämaß	23
6 Fortsetzung von Prämaßen und äußere Maße	27
7 Von äußeren Maßen zu Maßen	32
8 Messbare und Maß-Räume	36
9 Messbare Abbildungen	40
10 Das Integral bezüglich eines Maßes	45
11 Integrierbarkeit	50
12 Konvergenzsätze	55
13 Anwendungen und Folgerungen der Konvergenzsätze	60
14 Der Raum L^1	65
15 Die Invarianz der Lebesgue-Maße	70
16 Iterierte Integrale und die Sätze von Fubini und Tonelli	75
17 Der Transformationssatz	80
18 Anwendungen der Transformationsformel	85
19 Darstellung von k -dimensionalen Flächen	90
20 Untermannigfaltigkeiten des \mathbb{R}^d	95

21	Tangential- und Normalraum	100
22	Extrema unter Nebenbedingungen, Lagrange-Multiplikatoren	105
23	Karten und der Maßtensor	110
24	Das Integral über eine Untermannigfaltigkeit	115
25	Integrieren entlang des Radius	119
26	Kompakta mit glattem Rand	124
27	Die wichtigsten Differentialoperatoren	129
28	Der Divergenzsatz von Gauß	134

Vorbemerkung

Dies ist das Handout zur Vorlesung „Analysis 3“, gehalten an der TU Braunschweig im Wintersemester 2020/21. Es keinerlei Anspruch, auch ohne Besuch der Vorlesung verständlich zu sein. Diese Version enthält mit an Sicherheit grenzender Wahrscheinlichkeit zahlreiche Fehler und das, obwohl viele Leserinnen und Leser schon zahlreiche Fehler gefunden haben; seien Sie also aufmerksam beim Lesen.

Die Aufzeichnungen der Vorlesungen zu diesem Skript finden Sie entweder im Stud.IP-Eintrag zur Vorlesung oder auf dem YouTube Kanal

<https://www.youtube.com/channel/UCn9kJUL71fuWXXxX-o6090g>.

Alle weiteren Materialien finden Sie in Stud.IP.

Braunschweig, den 12. Februar 2021

Dirk Lorenz
d.lorenz@tu-braunschweig.de

1 Einleitung und Motivation - Das Maßproblem

Wir wollen nun im mehrdimensionalen integrieren. Genauer: Ist $\Omega \subset \mathbb{R}^d$, und $f : \Omega \rightarrow \mathbb{R}$, so wollen wir das Integral von f über Ω definieren. Für $d = 2$ soll das Integral dabei das "signierte Volumen zwischen Graph und x_1 - x_2 -Ebene" sein. Insgesamt werden wir hier anders vorgehen als in einer Dimension. Das liegt an zwei Dingen:

1. In einer Dimension hatten wir ein Intervall in Teil-Intervalle unterteilt und Funktionen betrachtet, die auf diesen Teilintervallen konstant sind. In mehreren Dimensionen gibt es keine naheliegende Unterteilung einer beliebigen Menge $\Omega \subset \mathbb{R}^d$.
2. Das Riemann-Integral führt zu Problemen bei Grenzübergängen: Sind Funktionen f_n alle Riemann-integrierbar über einem Intervall $[a, b]$, und konvergiert $f_n \rightarrow f$ punktweise, so folgt nicht, dass f Riemann-integrierbar ist.

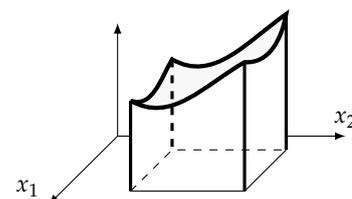
Wir gehen also anders vor:

1. Wir klären zuerst, wie wir für möglichst beliebige Mengen $A \subset \mathbb{R}^d$ die "Größe" der Menge bestimmen können (für $d = 1$ also die "Länge" für $d = 2$ die "Fläche", usw.). Dies führt auf den Begriff des *Maßes*.
2. Anschließend gehen wir zu Integralen über: Dazu werden wir nicht wieder den *Definitionsbereich* Ω einer Funktion $f : \Omega \rightarrow \mathbb{R}$ unterteilen, sondern den *Wertebereich*, d.h. wir betrachten Mengen der Form

$$\{x \in \Omega \mid \lambda_1 \leq f(x) \leq \lambda_2\}.$$

Dieser Zugang zu Integralen wurde 1902 von Henri Léon Lebesgue entwickelt.

Kommen wir zuerst zum Begriff des Maßes. In der Antike wurden Flächeninhalte (und Volumina) dadurch bestimmt, dass man die Fläche zerlegte und die Teile durch geometrische Konstruktionen auf einfache Flächen zurückführte, von denen man den Flächeninhalt einfach postulierte (z.B. für Rechtecke durch Länge mal Breite). Für beliebige Teilmengen des \mathbb{R}^2 (und allgemeiner im \mathbb{R}^d) ist nicht mehr klar, was ein sinnvoller Begriff eines Flächeninhaltes sein soll. Im ersten Semester hatten wir schon gesehen, dass wir Inhalte von Flächen, die wir mit Hilfe von Riemann-integrierbaren Funktionen beschreiben können, durch ein Integral berechnen können. Diese Konstruktion ist allerdings nicht sehr flexibel, da sie auf einer Zerlegung des Definitionsbereiches in Intervalle beruht (was insbesondere auch der Grund für das schlechte Verhalten bezüglich Grenzwerten ist).



Hier ein einfaches Beispiel: Es sei $I = [0, 1]$ und $Q = I \cap \mathbb{Q}$ die Menge der rationalen Zahlen in I . Wir können Q abzählen als $Q = \{q_k \mid k \in \mathbb{N}\}$ und definieren

$$f_n(x) = \begin{cases} 1, & \text{falls } x \in \{q_1, \dots, q_n\} \\ 0, & \text{sonst.} \end{cases}$$

Für

$$f(x) = \begin{cases} 1, & \text{falls } x \in Q \\ 0, & \text{sonst} \end{cases}$$

gilt $f_n \rightarrow f$ punktweise. Die Funktionen f_n sind stückweise stetig mit endlich vielen Unstetigkeitsstellen und daher Riemann integrierbar (mit $\int_0^1 f_n(x) dx = 0$). Die Grenzfunktion f hingegen ist nicht Riemann integrierbar.

Das Problem, allgemeinen Teilmengen des \mathbb{R}^d ein Maß (was die Begriffe Länge, Fläche und Volumen in $d = 1, 2$ bzw. 3 verallgemeinert) zuzuweisen, nennt man das *Maßproblem*. Wir wählen einen axiomatischen Zugang und stellen zuerst eine Wunschliste von Eigenschaften zusammen. Von einem Maß hätten wir gerne:

(M1) **Definitionsgebiet:** Ein Maß μ ist eine Abbildung, die jeder Teilmenge $M \subset \mathbb{R}^d$ eine nicht-negative Zahl oder ggf. auch “ $+\infty$ ” zuweist.

Jede Menge soll ein Maß haben, negative Maße scheinen nicht sinnvoll.

(M2) **Additivität:** Zerlegt man eine Menge in disjunkte Teile, dann soll das Maß der Menge gleich der Summe der Maße der Teile sein.

(M3) **Translationsinvarianz:** Für $M \subset \mathbb{R}^d$ und $a \in \mathbb{R}^d$ soll für die verschobene Menge $a + M = \{a + x \mid x \in M\}$ gelten

$$\mu(a + M) = \mu(M)$$

Verschieben soll das Maß gleich lassen.

(M4) **Normierung:** Für den Einheitswürfel $W^d := \{x \in \mathbb{R}^d \mid 0 \leq x_i \leq 1\}$ gilt

$$\mu(W^d) = 1$$

Dies normiert das Maß. Im Prinzip ist der Wert 1 willkürlich; aber sinnvoll.

Diese Axiome erscheinen nicht zu ambitioniert; jede Forderung für sich klingt einleuchtend und naheliegend. Allerdings tauchen schnell Probleme auf.

1. Problem: Beliebige Additivität funktioniert nicht. Das erste Problem betrifft (M2), (M3) und (M4): Wir können jede Menge, insbesondere den Einheitswürfel W^d , als disjunkte Vereinigung ihrer Punkte darstellen:

$$W^d = \bigcup_{x \in W^d} \{x\}.$$

Das Maß $\mu(W^d)$ soll nach (M4) gleich eins sein. Mit (M2) sollte auch eins herauskommen, wenn wir die Maße der Mengen $\{x\}$ zusammenzählen. Das Problem ist: setzen wir $\mu(\{x\}) = 0$, also das Maß eine einpunktigen Menge auf Null, ergibt sich bei Summieren immer noch Null. Setzen wir $\mu(\{x\}) = \epsilon > 0$, so muss dies wegen (M3) für alle x die gleiche Zahl ϵ sein, aber dann ist die Reihe über die Maße aller Punkte in W^d divergent und wieder nicht gleich eins.

Wir können also nicht (M₁)-(M₄) gleichzeitig haben. Wir wollen aber an der Forderung, dass einzelne Punkte das Maß Null haben, festhalten, d.h. $\mu(\{x\}) = 0$ für alle x .

Daraus folgt mit der Additivität sofort, dass es bei einem Intervall für das Maß nicht auf die Randpunkte ankommt, d.h. es gilt $\mu([a, b]) = \mu([a, b[) = \mu(]a, b]) = \mu(]a, b[)$. Außerdem folgt aus der Additivität sofort, dass die leere Menge das Maß Null hat.

Versuchen wir, die geforderte Additivität abzuschwächen:

(M₂)' **abzählbare Additivität (auch σ -Additivität):** Hat man abzählbar viele disjunkte Mengen $M_1, M_2, \dots \subset \mathbb{R}^d$, dann soll gelten

$$\mu\left(\bigcup_{j \in \mathbb{N}} M_j\right) = \sum_{j=1}^{\infty} \mu(M_j)$$

Hieraus folgt insbesondere auch:

(M₅) [Monotonie:] Ist $A \subset B$, so ist $B = A \cup (B \setminus A)$ und die Vereinigung ist disjunkt. Daher gilt

$$\begin{aligned} \mu(B) &= \mu(A \cup (B \setminus A)) = \mu(A) + \underbrace{\mu(B \setminus A)}_{\geq 0} \\ &\geq \mu(A) \end{aligned}$$

immer wenn $B \supset A$ gilt.

Dies verbietet die Zerlegung in überabzählbar viele Teile und das Problem, dass bei Zerlegung in einzelne Punkte auftritt wäre damit behoben. Es bleibt aber immer noch ein Problem über; dies ist etwas diffiziler.

2. Problem: Abzählbare Additivität funktioniert nicht. Dies Problem wurde schon 1904 von Giuseppe Vitali entdeckt und beruht auf der Konstruktion einer sehr "pathologischen" Teilmenge des Einheitsintervalls, der sogenannten Vitali-Menge:

Wir betrachten das halboffene Einheitsintervall $[0, 1[$. Wir definieren die folgende Äquivalenzrelation:

$$x \sim^{\mathbb{Q}} y \iff x - y \in \mathbb{Q}.$$

Dann zerfällt $[0, 1[$ in disjunkte Äquivalenzklassen und es sei $N \subset [0, 1[$ eine Menge, die aus jeder der Äquivalenzklassen genau ein Element enthält. Diese Menge N wird *Vitali-Menge* genannt. Zu diesem N definieren wir rationale Verschiebungen, d.h. zu $r \in \mathbb{Q}$ sei

$$N_r = \{x + r \mid x \in N\}.$$

Es gilt:

i) Ist $r \neq s$ für $r, s \in \mathbb{Q} \cap [0, 1[$, so gilt $N_r \cap N_s = \emptyset$.

Eine gute Anschauung für diese Menge gibt es leider nicht. Das liegt daran, dass ihre Konstruktion benötigt, dass man aus unendlich vielen Mengen gleichzeitig jeweils ein Element auswählen muss. Dies ist kein explizites Konstruktionsprinzip und lässt sich nicht systematisch durchführen. Das Prinzip, aus einer beliebigen Menge von Mengen eine neue Menge zu konstruieren, indem man jeweils ein Element aus jeder der Mengen nimmt, ist eines der Axiome der Mengenlehre und wird *Auswahlaxiom* genannt. Dieses Axiom haben wir schon benutzt ohne es zu erwähnen. Es ermöglicht viele einfache abstrakte Konstruktionen, führt aber auch zu seltsamen Ergebnissen, wie z.B. der Existenz der Vitali-Menge.

- ii) Ist $\{r_n\}_{n \in \mathbb{N}}$ ein Abzählung der Menge $T = [-1, 1] \cap \mathbb{Q}$, so gilt

$$[0, 1[\subset \bigcup_{n=1}^{\infty} N_{r_n}.$$

- iii) Es gilt

$$\bigcup_{n=1}^{\infty} N_{r_n} \subset [-1, 2]$$

- i) Angenommen, es wäre $N_r \cap N_s \neq \emptyset$. Dann gäbe es $x, y \in N$, so dass $x + r = y + s$. Dann wäre aber $x - y = s - r \in \mathbb{Q}$ und damit wäre $x \sim^{\mathbb{Q}} y$ und x, y wären in der gleichen Äquivalenzklasse und könnten nicht beide in N sein. Ein Widerspruch.
- ii) Es sei $a \in [0, 1[$, d.h. $0 \leq a < 1$. Dieses a liegt in einer der Äquivalenzklassen der Relation $\sim^{\mathbb{Q}}$ und daher können wir es als $a = r + x$ mit $x \in N$ und $r \in \mathbb{Q}$ schreiben. Da $0 \leq a, x < 1$ gilt, muss $r = a - x \in [-1, 1]$ gelten und daher folgt die behauptete Inklusion.
- iii) Für $|r| \leq 1$ ist $N_r \subset [-1, 2]$.

Und hier haben wir das Problem:

- Wegen der Normiertheit gilt

$$\mu([0, 1]) = 1$$

und wegen

$$[-1, 2] = [-1, 0[\cup [0, 1[\cup [1, 2]$$

und der Verschiebungsinvarianz gilt

$$\mu([-1, 2]) = 3.$$

- Es folgt also mit Monotonie (M₅)

$$1 = \mu([0, 1]) \stackrel{\text{ii)}}{\leq} \mu\left(\bigcup_{n=1}^{\infty} N_{r_n}\right) \stackrel{\text{iii)}}{\leq} \mu([-1, 2]) = 3.$$

- Wegen der abzählbaren Additivität folgt, da die Mengen N_r paarweise disjunkt sind,

$$1 \leq \sum_{n=1}^{\infty} \mu(N_{r_n}) \leq 3.$$

- Wegen der Verschiebungsinvarianz haben die Mengen N_{r_n} alle das gleiche Maß wie die Menge N .

Wir sehen: Egal welches Maß wir der Menge N zuweisen, wir bekommen einen Widerspruch in der Ungleichung $1 \leq \sum_{n=1}^{\infty} \mu(N) \leq 3$.

Hier könnte Abhilfe darin bestehen, nur noch zu fordern

(M2) **endliche Additivität:** Für endlich viele disjunkte Mengen $M_1, \dots, M_r \subset \mathbb{R}^d$ soll gelten

$$\mu\left(\bigcup_{j=1}^r M_j\right) = \sum_{j=1}^r \mu(M_j)$$

Auch das führt immer noch zu einem Problem

3. Problem: Auch endliche Additivität funktioniert nicht! In drei Dimensionen gibt es das *Banach-Tarski Paradoxon*. Dies besagt, dass man die Einheitskugel in \mathbb{R}^3 in fünf disjunkte Teile zerlegen kann, so dass diese durch Drehungen und Verschiebungen zu zwei neuen Einheitskugeln zusammengesetzt werden können. Für diese fünf Mengen lässt sich also kein endlich-additives Maß definieren, da sich einerseits die fünf Maße zu eins aufaddieren müssen, aber andererseits für jeweils zwei und drei davon das gleiche gelten muss.

Man kann dies Paradoxon, grob gesprochen, dadurch erklären, dass die Zerlegung dermaßen komplizierte Mengen benutzt, dass der Begriff des "Volumens" für sie keinen Sinn mehr ergibt.

Wir waren also doch zu ambitioniert und müssen unsere Axiome abändern. Wir fragen uns, welche unserer Forderungen (M1) bis (M4) zu ambitioniert war und was weggelassen werden soll. Zwei Möglichkeiten bieten sich an:

- Wir fordern nicht mehr, dass *alle* Mengen ein Maß haben sollen, sondern nur noch "genügend viele" Mengen. Es gibt ja überabzählbar viele Elemente in \mathbb{R}^d und die Menge aller Teilmengen von \mathbb{R}^d noch einmal größer (sogar in einem präzisen Sinn, Stichwort "Cantors Diagonalargument").
- Wir verzichten auf die "Additivität" in jeder Form und fordern hier weniger.

Wir nehmen den ersten Weg, d.h. wir schwächen (M1) ab (indem wir einen kleineren Definitionsbereich nehmen) und benutzen weiterhin (M2)'. Daher werden wir in den nächsten Abschnitten zuerst allgemeine Systeme von Teilmengen einer Menge untersuchen.

2 Mengenalgebren und -ringe

Es sei Ω eine Menge. Mit $\mathfrak{P}(\Omega)$ bezeichnen wir die *Potenzmenge*, d.h. die Menge aller Teilmengen von Ω . Für zwei Mengen $A, B \in \mathfrak{P}(\Omega)$ gibt es

$$\begin{aligned} A \cup B & \quad (\text{Vereinigung}) \\ A \cap B & \quad (\text{Schnitt}) \\ A \setminus B & \quad (\text{Mengendifferenz}) \\ A^c := \Omega \setminus A & \quad (\text{Komplement}) \\ A \Delta B := (A \setminus B) \cup (B \setminus A) & \quad (\text{symmetrische Differenz}) \end{aligned}$$

Für Familien von Mengen, also für $A_i \in \mathfrak{P}(\Omega), i \in I$ ist

$$\begin{aligned} \bigcup_{i \in I} A_i &= \{x \in \Omega \mid \exists j \in I : x \in A_j\}, \\ \bigcap_{i \in I} A_i &= \{x \in \Omega \mid \forall j \in I : x \in A_j\}, \end{aligned}$$

wobei I eine beliebige (endliche, abzählbar unendliche oder auch überabzählbar unendliche Menge) ist.

Satz 2.1. Für $A, B, C, X, Y, A_i, X_i \in \mathfrak{P}(\Omega)$ mit $i \in I$ gilt

- i) $A \Delta \emptyset = A, A \Delta A = \emptyset$
- ii) $A \Delta B = B \Delta A$
- iii) $(A \Delta B) \Delta C = A \Delta (B \Delta C)$
- iv) $(A \Delta B) \cap C = (A \cap C) \Delta (B \cap C)$
- v) Wenn $A \cap B = \emptyset$, dann $A \Delta B = A \cup B$
- vi) Wenn $B \subset A$, dann $A \Delta B = A \setminus B$
- vii) Wenn $X \cap Y = \emptyset$, dann $A \cap B \subset (X \Delta A) \cup (Y \Delta B)$
- viii) $\left(\bigcup_{i \in I} X_i\right) \Delta \left(\bigcup_{i \in I} A_i\right) \subset \bigcup_{i \in I} (X_i \Delta A_i)$

Beweis.

Die Regeln i) bis vi) macht man sich schnell mit einem Bild klar.

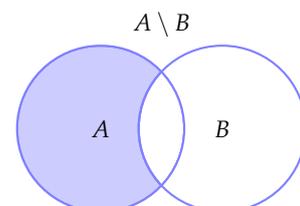
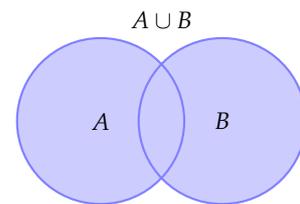
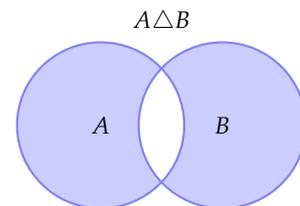
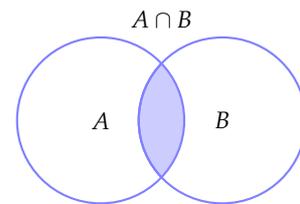
Zu vii): Sei $x \in A \cap B$. Wir unterscheiden die beiden Fälle $x \in X$ und $x \notin X$:

1. Ist $x \in X$, so ist (wegen $X \cap Y = \emptyset$) $x \notin Y$. Dann ist aber $x \in Y \Delta B$ und damit auch in $(X \Delta A) \cup (Y \Delta B)$.
2. Ist $x \notin X$, so ist $x \in A \setminus X$ und damit ebenfalls in $(X \Delta A) \cup (Y \Delta B)$ was die Behauptung zeigt.

Zu viii): Sei $x \in \left(\bigcup_{i \in I} X_i\right) \Delta \left(\bigcup_{i \in I} A_i\right)$. Auch hier unterscheiden wir zwei Fälle, nämlich $x \in \bigcup_{i \in I} X_i$ und $x \notin \bigcup_{i \in I} X_i$.

Manchmal wird die Potenzmenge auch mit 2^Ω bezeichnet. Dies ist dadurch motiviert, dass jede Teilmenge $A \subset \Omega$ durch eine Funktion f_A von Ω in die zwei-elementige Menge $\{0,1\}$ beschrieben werden kann, nämlich durch $f_A(x) = 1$ für $x \in A$ und $= 0$ sonst.

Für leere Vereinigungen und Schnitte gibt es die Konventionen $\bigcup_{i \in \emptyset} A_i = \emptyset$ und $\bigcap_{i \in \emptyset} A_i = \Omega$.



1. Ist $x \in \bigcup_{i \in I} X_i$, so ist $x \notin \bigcup_{i \in I} A_i$. Das bedeutet, dass x in mindestens einem X_{i_0} liegt und in keinem A_i . Dann ist aber $x \in X_{i_0} \Delta A_{i_0}$ und also auch in $\bigcup_{i \in I} (X_i \Delta A_i)$.
2. Ist $x \notin \bigcup_{i \in I} X_i$, so gilt $x \in \bigcup_{i \in I} A_i$ und das Argument ist analog.

□

Definition 2.2. Ist Ω eine Menge und $\mathfrak{A} \subset \mathfrak{P}(\Omega)$ eine Menge von Teilmengen. Dann heißt \mathfrak{A} eine *Mengenalgebra* auf Ω , wenn gilt

- i) $\emptyset \in \mathfrak{A}$,
- ii) $A \in \mathfrak{A} \implies A^c \in \mathfrak{A}$,
- iii) $A, B \in \mathfrak{A} \implies A \cup B \in \mathfrak{A}$.

Beispiel. Für jede Menge Ω sind

$$\mathfrak{A} = \{\emptyset, \Omega\}, \quad \text{und} \quad \mathfrak{A} = \mathfrak{P}(\Omega)$$

Mengenalgebren.

Definition 2.3. Ein Mengensystem $\mathfrak{R} \subset \mathfrak{P}(\Omega)$ heißt *Mengenring*, wenn gilt

- i) $\emptyset \in \mathfrak{R}$,
- ii) $A, B \in \mathfrak{R} \implies A \setminus B \in \mathfrak{R}$,
- iii) $A, B \in \mathfrak{R} \implies A \cup B \in \mathfrak{R}$.

Jede Mengenalgebra ist ein Mengenring und ein Mengenring \mathfrak{R} ist eine Mengenalgebra, wenn $\Omega \in \mathfrak{R}$ gilt.

Wir könnten ii) und iii) auch durch ii') " $A, B \in \mathfrak{R} \implies A \Delta B \in \mathfrak{R}$ " und iii') " $A, B \in \mathfrak{R} \implies A \cap B \in \mathfrak{R}$ " ersetzen.

Dass ii)' und iii)' aus ii) und iii) folgen, sieht man schnell, und für die Rückrichtung zeigt man zuerst die Identitäten $A \setminus B = (A \cap B) \Delta A$ und $A \cup B = (A \Delta B) \Delta (A \cap B)$.

Beispiel. a) Ist Ω eine Menge, so ist $\mathfrak{R} = \{A \subset \Omega \mid A \text{ endlich}\}$ ein Mengenring und eine Mengenalgebra genau dann, wenn Ω endlich ist.

b) Ist Ω eine Menge, so ist $\mathfrak{A} = \{A \subset \Omega \mid A \text{ oder } A^c \text{ endlich}\}$ eine Mengenalgebra.

c) Für $\Omega = \mathbb{R}$ sei Ω die Menge aller Teilmengen $A \subset \mathbb{R}$ die sich schreiben lassen als

$$A = \bigcup_{i=1}^m [a_i, b_i[, \quad a_i < b_i.$$

(D.h. Ω besteht aus endlichen Vereinigungen von links abgeschlossenen und rechts-offenen Intervallen.) Dieses Ω ist ein Mengenring.

Aus i) und ii) folgt direkt $\Omega = \emptyset^c \in \mathfrak{A}$ und aus ii) und iii) und den Regeln $(A \cap B)^c = A^c \cup B^c$ und $(A \cup B)^c = A^c \cap B^c$ folgt, dass man iii) auch durch iii') " $A, B \in \mathfrak{A} \implies A \cap B \in \mathfrak{A}$ " ersetzen könnte.

Wegen $A \cap B = A \setminus (A \setminus B)$ folgt auch hier, dass $A \cap B$ in \mathfrak{R} liegt, wenn A und B dies tun.

Die leere Menge ist als leere Vereinigung in Ω . Das Vereinigungen von Mengen in Ω wieder in Ω sind, ist per Definition klar.

Bleibt zu zeigen, dass für $A, B \in \Omega$ auch $A \setminus B$ in Ω liegt. Betrachten wir dazu zwei halboffene Intervalle I und J . Dann ist $I \setminus J$ entweder leer, ein halboffenes Intervall oder die Vereinigung von zwei solchen. Für $A = \bigcup_{k=1}^m I_k$ folgt dann

$$A \setminus J = \left(\bigcup_{k=1}^m I_k \right) \setminus J = \bigcup_{k=1}^m (I_k \setminus J)$$

woraus $A \setminus J \in \Omega$ folgt. Ist $B = \bigcup_{l=1}^n J_l$, setzen wir $B_s = \bigcup_{l=1}^s J_l$ und bekommen induktiv

$$A \setminus B_s = (A \setminus B_{s-1}) \setminus J_s \in \Omega$$

woraus die Behauptung folgt.

Definition 2.4. Sind Ω_1 und Ω_2 Mengen so nennen wir eine Teilmenge X in $\Omega_1 \times \Omega_2$ *Rechteck*, wenn sie von der Form $X = A \times B$ mit $A \subset \Omega_1$ und $B \subset \Omega_2$ ist.

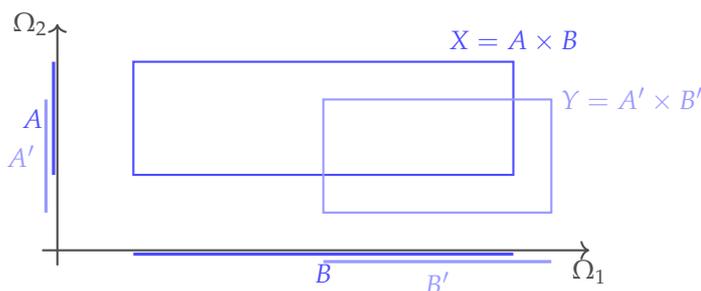
Sind $\mathfrak{A} \subset \mathfrak{P}(\Omega_1)$ und $\mathfrak{B} \subset \mathfrak{P}(\Omega_2)$ Mengenringe, so definieren wir das *Produkt* der Mengenringe $\mathfrak{R} \subset \mathfrak{P}(\Omega_1 \times \Omega_2)$ durch

$$\mathfrak{R} = \mathfrak{A} \boxtimes \mathfrak{B} = \left\{ \bigcup_{i=1}^m A_i \times B_i \mid A_i \in \mathfrak{A}, B_i \in \mathfrak{B} \right\}$$

Dies ist in der Tat ein Mengenring: Wie im Beispiel oben sind $\emptyset \in \mathfrak{A} \boxtimes \mathfrak{B}$ und " $X, Y \in \mathfrak{A} \boxtimes \mathfrak{B} \implies X \cup Y \in \mathfrak{A} \boxtimes \mathfrak{B}$ " klar. Bleibt zu zeigen, dass auch $X \setminus Y \in \mathfrak{A} \boxtimes \mathfrak{B}$ gilt. Wir betrachten zuerst

$$X = A \times B, Y = A' \times B', \quad A, A' \in \mathfrak{A}, B, B' \in \mathfrak{B}.$$

Für $X \setminus Y$ sieht die Situation so aus:



und es ist klar, dass sich $(A \times B) \setminus (A' \times B')$ als Vereinigung von höchstens drei Rechtecken schreiben lässt die alle in $\mathfrak{A} \boxtimes \mathfrak{B}$ liegen (genauer sind das die Mengen $(A \setminus A') \times (B \setminus B')$, $(A \cap A') \times (B \setminus B')$ und $(A \setminus A') \times (B \cap B')$ - machen Sie sich klar, dass das immer stimmt, auch im Fall dass $Y \subset X$!). Also ist $X \setminus Y \in \mathfrak{A} \boxtimes \mathfrak{B}$. Wie im vorigen Beispiel zeigt man weiter, dass auch für $X = \bigcup_{i=1}^m (A_i \times B_i)$, $Y = \bigcup_{j=1}^n (A'_j \times B'_j)$ wieder $X \setminus Y \in \mathfrak{A} \boxtimes \mathfrak{B}$ gilt.

Beispiel. Wir betrachten nun $\Omega = \mathbb{R}^d$ und dort die halboffenen Quader

$$Q = [a_1, b_1[\times \cdots \times [a_d, b_d[$$

(welche wir im Folgenden nur "Quader" nennen).

Mit $\mathfrak{Q}(\mathbb{R}^d)$ bezeichnen wir die Menge aller endlichen Vereinigungen von Quadern. Für $d = 1$ ist das exakt die Menge \mathfrak{Q} aus dem vorigen Beispiel. Für $d > 1$ können wir jeden Quader Q in \mathbb{R}^d schreiben als

$$Q = I \times Q'$$

mit einem halboffenen Intervall I und einem Quader Q' in \mathbb{R}^{d-1} . Es folgt

$$\mathfrak{Q}(\mathbb{R}^d) = \mathfrak{Q}(\mathbb{R}) \boxtimes \mathfrak{Q}(\mathbb{R}^{d-1})$$

und insbesondere ist $\mathfrak{Q}(\mathbb{R}^d)$ ein Mengerring, den wir den *Mengerring der endlichen Quadersummen* nennen.

Es gilt sogar, dass wir jedes Element in $\mathfrak{Q}(\mathbb{R}^d)$ als *disjunkte* endliche Vereinigung von Quadern schreiben können (und hierbei ist es hilfreich, dass wir halboffene Quader gewählt haben).

Man beachte, dass $\mathfrak{Q}(\mathbb{R}^d)$ keine Mengenalgebra ist, da der \mathbb{R}^d selbst nicht in $\mathfrak{Q}(\mathbb{R}^d)$ liegt.

3 σ -Algebren

Führen wir nun die wichtigsten Mengensysteme in der Maßtheorie ein:

Definition 3.1. Ist Ω eine Menge, so nennen wir $\mathfrak{A} \subset \mathfrak{P}(\Omega)$ eine σ -Algebra, wenn folgende Eigenschaften gelten:

- i) $\emptyset \in \mathfrak{A}$,
- ii) $A \in \mathfrak{A} \implies A^c \in \mathfrak{A}$,
- iii) Für $A_k \in \mathfrak{A}, k = 1, 2, \dots$ gilt $\bigcup_{k=1}^{\infty} A_k \in \mathfrak{A}$.

Wir wollen nun häufiger Folgen von Mengen studieren und führen daher ein:

Definition 3.2. Wir schreiben $X_k \nearrow X$ (oder $X_k \nearrow_{k=1}^{\infty} X$), falls $X_1 \subset X_2 \subset \dots$ und $X = \bigcup_{k=1}^{\infty} X_k$ gelten. Analog schreiben wir $Y_k \searrow Y$ (bzw. $Y_k \searrow_{k=1}^{\infty} Y$), falls $Y_1 \supset Y_2 \supset \dots$ und $Y = \bigcap_{k=1}^{\infty} Y_k$ gelten.

Satz 3.3. Eine Mengenalgebra $\mathfrak{A} \subset \mathfrak{P}(\Omega)$ ist genau dann eine σ -Algebra, wenn für jede aufsteigende Folge $A_k \in \mathfrak{A}$ mit $A_k \nearrow A$ gilt, dass $A \in \mathfrak{A}$.

Beweis.

Die Hinrichtung ist klar nach Definition und für die Rückrichtung argumentieren wir wie folgt: Sei \mathfrak{A} eine Mengenalgebra in der alle Vereinigungen von aufsteigenden Folgen wieder enthalten sind. Sind dann $A_1, A_2, \dots \in \mathfrak{A}$, so ist $\tilde{A}_k = \bigcup_{i=1}^k A_i$ eine aufsteigende Folge von Mengen. Dann gilt $A := \bigcup_{k=1}^{\infty} \tilde{A}_k = \bigcup_{k=1}^{\infty} A_k$ und nach Voraussetzung ist $A \in \mathfrak{A}$. Also ist \mathfrak{A} eine σ -Algebra. \square

Beispiel. a) Jede Mengenalgebra mit nur endlich vielen Mengen ist auch eine σ -Algebra, also z.B. $\mathfrak{A} = \{\emptyset, \Omega\}$. Außerdem ist $\mathfrak{A} = \mathfrak{P}(\Omega)$ immer eine σ -Algebra.

b) Ist $\mathfrak{A} \subset \mathfrak{P}(\Omega)$ eine σ -Algebra und $\Omega_1 \subset \Omega$, so definieren wir

$$\mathfrak{A} \cap \Omega_1 := \{A \cap \Omega_1 \mid A \in \mathfrak{A}\}$$

und bekommen dadurch wieder eine σ -Algebra (vgl. Übung). Wir nennen sie die von \mathfrak{A} induzierte σ -Algebra oder auch die Spur von \mathfrak{A} auf Ω_1 .

c) Etwas allgemeiner: Sind Ω und Ω_1 Mengen, $\varphi : \Omega_1 \rightarrow \Omega$ eine Abbildung und $\mathfrak{A} \subset \mathfrak{P}(\Omega)$ eine σ -Algebra, so ist

$$\mathfrak{A}_1 = \varphi^{-1}(\mathfrak{A}) := \{\varphi^{-1}(A) \mid A \in \mathfrak{A}\}$$

σ -Algebren sind also auch bezüglich abzählbarer Vereinigungen abgeschlossen und nicht nur bezüglich endlicher, wie die Mengenalgebren.

Da jedes A_k auch leer sein kann, ist jede σ -Algebra eine Mengenalgebra. Wir könnten in iii) auch die Vereinigung auch durch einen Schnitt ersetzen.

Im Fall von $X_k \nearrow X$ (bzw. $X_k \searrow X$) nennen wir X den *aufsteigenden* (bzw. *absteigenden*) Grenzwert. Beachte, dass dies nur erklärt ist, wenn die Mengenfolge tatsächlich aufsteigend (bzw. absteigend) ist. Im allgemeinen sind Grenzwerte von Mengen nicht definiert und ein delikates Thema.

(also die Menge der Urbilder von Elementen der σ -Algebra \mathfrak{A}) selbst eine σ -Algebra auf Ω_1 . Das folgt direkt aus den Rechenregeln für Urbilder

$$\varphi^{-1}(\Omega \setminus A) = \Omega_1 \setminus \varphi^{-1}(A), \quad \varphi^{-1}\left(\bigcup_{i \in I} A_i\right) = \bigcup_{i \in I} \varphi^{-1}(A_i).$$

Dies umfasst das vorige Beispiel, in dem man $\Omega_1 \subset \Omega$ und die Einbettungsabbildung $\iota: \Omega_1 \rightarrow \Omega, \iota(x) = x$ nimmt, denn dann ist $\mathfrak{A} \cap \Omega_1 = \iota^{-1}(\mathfrak{A})$.

Hat man mehrere σ -Algebren $\mathfrak{A}_i, i \in I$ auf der gleichen Menge Ω , so kann man deren Durchschnitt

$$\mathfrak{A} = \bigcap_{i \in I} \mathfrak{A}_i$$

betrachten und erhält wieder eine σ -Algebra.

Jedes Element $A \in \mathfrak{A}$ ist also in allen \mathfrak{A}_i enthalten, damit folgen sofort die drei Eigenschaften einer σ -Algebra für \mathfrak{A} .

Damit können wir definieren:

Definition 3.4. Ist $\mathfrak{E} \subset \mathfrak{P}(\Omega)$ irgend ein System von Teilmengen, so ist

$$\sigma(\mathfrak{E}) := \bigcap_{\substack{\mathfrak{A} \subset \mathfrak{P}(\Omega) \\ \mathfrak{A} \text{ } \sigma\text{-Algebra}} \mathfrak{A}$$

der Schnitt aller der σ -Algebren, die alle Mengen in \mathfrak{E} enthalten und damit in gewissem Sinne die kleinste σ -Algebra die alle Mengen in \mathfrak{E} enthält. Wie nennen $\sigma(\mathfrak{E})$ auch die von \mathfrak{E} erzeugte σ -Algebra (vgl. Übung).

Beispiel. Im Folgenden besonders wichtig ist die *Borelsche σ -Algebra* in einem topologischen Raum: Ist (Ω, τ) ein topologischer Raum, so betrachten wir die σ -Algebra, die von den offenen Mengen erzeugt wird, also

$$\mathcal{B}(\Omega) := \sigma(\tau).$$

Wir nennen $\mathcal{B}(\Omega)$ auch die *Borel-Algebra* von Ω und die Elemente in $\mathcal{B}(\Omega)$ heißen *Borel-Mengen*.

Ein knappe Charakterisierung der Mengen, die in der Borelschen σ -Algebra liegen, ist nicht möglich. Um sie besser zu verstehen, untersuchen wir $\mathcal{B}(\Omega)$ für ein $\Omega \subset \mathbb{R}^d$ (mit der Standard-Topologie):

Beispiel. Für $\Omega \subset \mathbb{R}^d$ gilt:

- $\mathcal{B}(\Omega)$ enthält neben allen offenen Mengen auch alle abgeschlossenen Mengen, da σ -Algebren abgeschlossen bezüglich Komplementbildung sind.

Einbettungsabbildungen schreiben wir auch abkürzend einfach als $\iota: \Omega_1 \hookrightarrow \Omega$.

- b) Da einzelne Punkte abgeschlossen sind, sind auch alle einelementigen Mengen in $\mathcal{B}(\Omega)$.
- c) Da eine σ -Algebra abgeschlossen bezüglich abzählbarer Vereinigungen ist, ist auch $\mathbb{Q}^d \cap \Omega$ in $\mathcal{B}(\Omega)$.

Im \mathbb{R}^d können wir die Borel-Algebra sogar durch Quader erzeugen! Dazu zeigen wir zuerst ein Lemma:

Lemma 3.5. *Jede offene Menge $U \subset \mathbb{R}^d$ ist die disjunkte Vereinigung abzählbar vieler halboffener Würfel mit rationalen Eckpunkten.*

Beweis.

Zu $k \geq 0$ bezeichnen wir mit \mathfrak{W}_k die (abzählbare) Menge aller Würfel der Form

$$W = \left[\frac{m_1}{2^k}, \frac{m_1+1}{2^k} \right[\times \dots \times \left[\frac{m_d}{2^k}, \frac{m_d+1}{2^k} \right[, \quad m_1, \dots, m_d \in \mathbb{Z}$$

(dies sind die Würfel mit *dyadischen* Eckpunkten). Die Würfel aus \mathfrak{W}_k sind eine disjunkte Überdeckung des \mathbb{R}^d und für $k' \geq k$ und $W \in \mathfrak{W}_k$ und $W' \in \mathfrak{W}_{k'}$ gilt: Entweder sind W und W' disjunkt oder es ist $W' \subset W$.

Wir konstruieren nun die benötigte Menge von Würfel induktiv: Zu Beginn sei M_0 die Menge aller Würfel $W \in \mathfrak{W}_0$ (also diejenigen mit ganzzahligen Eckpunkten), welche ganz in U enthalten sind. Für $k \geq 1$ sei dann M_k die Menge aller Würfel in \mathfrak{W}_k , welche ganz in $U \setminus \bigcup_{i=0}^{k-1} M_i$ liegen (also in U , aber in keinem der größeren Würfel, die wir schon hinzugenommen haben).

Wir setzen schließlich

$$M = \bigcup_{k=0}^{\infty} M_k.$$

Nach Konstruktion sind alle Würfel in M disjunkt und es bleibt zu zeigen, dass $M = U$ gilt (wobei $M \subset U$ nach Konstruktion schon erfüllt ist): Sei dazu $a \in U$. Da U offen ist, gibt es $\epsilon > 0$, so dass $B_\epsilon(a) \subset U$ gilt. Die Würfel in \mathfrak{W}_k haben den Durchmesser $2^{-k}\sqrt{d}$. Daher gibt es ein k und einen Würfel $W \in \mathfrak{W}_k$ mit $a \in W \subset B_\epsilon(a) \subset U$. Dieser Würfel W ist entweder im konstruierten M_k , oder in einem größeren $W' \in M_{k'}$ enthalten. Es folgt $a \in M$ und also $U \subset M$. \square

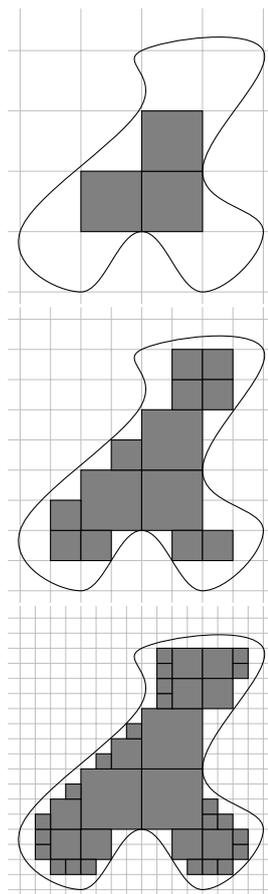
Satz 3.6. *Ist $\mathfrak{Q}(\mathbb{R}^d)$ der Mengenring der endlichen Quadersummen aus Abschnitt 2, so gilt*

$$\mathcal{B}(\mathbb{R}^d) = \sigma(\mathfrak{Q}(\mathbb{R}^d)).$$

Bezeichnen wir mit $\mathfrak{Q}_Q(\mathbb{R}^d)$ die Menge aller Quader mit rationalen Endpunkten, so gilt ebenfalls

$$\mathcal{B}(\mathbb{R}^d) = \sigma(\mathfrak{Q}_Q(\mathbb{R}^d)).$$

Die Konstruktion der Würfel im Beweis sieht wie folgt aus:



Beweis.

Jeder halboffene Quader $Q = [a_1, b_1[\times \cdots \times [a_d, b_d[$ lässt sich als Durchschnitt von abzählbar vielen offenen Quadern darstellen, nämlich

$$Q = \bigcap_{n=1}^{\infty} \left(]a_1 - 1/n, b_1[\times \cdots \times]a_d - 1/n, b_d[\right).$$

Ist τ die Topologie auf dem \mathbb{R}^d (d.h. das System aller offenen Mengen im \mathbb{R}^d), so folgt

$$\Omega(\mathbb{R}^d) \subset \sigma(\tau) = \mathcal{B}(\mathbb{R}^d)$$

und damit auch $\sigma(\Omega(\mathbb{R}^d)) \subset \mathcal{B}(\mathbb{R}^d)$. Aus Lemma 3.5 folgt

$$\tau \subset \sigma(\Omega_Q(\mathbb{R}^d))$$

und damit $\mathcal{B}(\mathbb{R}^d) = \sigma(\tau) \subset \sigma(\Omega_Q(\mathbb{R}^d))$. Damit haben wir

$$\mathcal{B}(\mathbb{R}^d) \subset \sigma(\Omega_Q(\mathbb{R}^d)) \subset \sigma(\Omega(\mathbb{R}^d)) \subset \mathcal{B}(\mathbb{R}^d)$$

und es muss also überall Gleichheit herrschen. \square

Definition 3.7 (Produkt von σ -Algebren). Sind Ω_1 und Ω_2 Mengen und $\mathfrak{A} \subset \mathfrak{P}(\Omega_1)$ und $\mathfrak{B} \subset \mathfrak{P}(\Omega_2)$ zwei σ -Algebren, so definieren wir das *Produkt der σ -Algebren* als

$$\mathfrak{A} \otimes \mathfrak{B} := \sigma(\mathfrak{A} \boxtimes \mathfrak{B}).$$

Es gilt

$$\mathcal{B}(\mathbb{R}^{m+n}) = \mathcal{B}(\mathbb{R}^m) \otimes \mathcal{B}(\mathbb{R}^n).$$

Definition 3.8 (Monotone Klassen). Eine nicht-leere Mengenfamilie $\mathfrak{M} \subset \mathfrak{P}(\Omega)$ heißt *monotone Klasse*, wenn gilt

- i) Für $X_k \in \mathfrak{M}$ und $X_k \nearrow X$ gilt $X \in \mathfrak{M}$, und
- ii) Für $Y_k \in \mathfrak{M}$ und $Y_k \searrow Y$ gilt $Y \in \mathfrak{M}$.

- Jede σ -Algebra ist nach Satz 3.3 eine monotone Klasse und jede Mengenalgebra, die eine monotone Klasse ist, ist eine σ -Algebra (vgl. Übung).
- Der Schnitt von monotonen Klassen ist wieder eine monotone Klasse, daher können wir für jede Familie \mathfrak{E} den Schnitt über alle monotonen Klassen bilden, die \mathfrak{E} enthalten und bekommen so die kleinste monotone Klasse, die \mathfrak{E} enthält. Wir nennen diese, die von \mathfrak{E} erzeugte *monotone Klasse*.

Satz 3.9. Ist $\mathfrak{A} \subset \mathfrak{P}(\Omega)$ ein Mengerring und \mathfrak{M} die von \mathfrak{A} erzeugte monotone Klasse, so ist \mathfrak{M} wieder ein Mengerring.

Bemerke, dass $\mathfrak{A} \boxtimes \mathfrak{B}$ im Allgemeinen keine σ -Algebra ist. Nach dieser Konstruktion ist $\mathfrak{A} \otimes \mathfrak{B}$ die kleinste σ -Algebra, welche alle Rechtecke $A \times B$ mit $A \in \mathfrak{A}$ und $B \in \mathfrak{B}$ enthält.

Mit anderen Worten: \mathfrak{M} ist abgeschlossen bezüglich aufsteigenden und absteigenden Grenzwerten.

Aber nicht jede monotone Klasse ist eine σ -Algebra: Die Familie bestehend aus den Intervallen $]0, a[$ und $[0, a]$ ist zwar eine monotone Klasse, aber nicht einmal ein Mengerring.

Beweis.

Zu $X \subset \Omega$ definieren wir

$$\kappa(\mathfrak{M}, X) = \{Y \in \mathfrak{P}(\Omega) \mid X \setminus Y \in \mathfrak{M}, Y \setminus X \in \mathfrak{M}, X \cup Y \in \mathfrak{M}\}.$$

Offensichtlich gilt $Y \in \kappa(\mathfrak{M}, X)$ genau dann, wenn $X \in \kappa(\mathfrak{M}, Y)$.
Es gilt: \mathfrak{M} ist ein Mengerring, genau dann, wenn für jedes $X \in \mathfrak{M}$ gilt

$$\mathfrak{M} \subset \kappa(\mathfrak{M}, X).$$

Zeigen wir, dass diese Inklusion gilt: Zuerst bemerken wir, dass $\kappa(\mathfrak{M}, X)$ eine monotone Klasse ist: Gilt nämlich $Y_1 \subset Y_2 \subset \dots$ und $Y_k \nearrow Y$, so gilt ebenfalls

$$(X \setminus Y_k) \searrow (X \setminus Y), (Y_k \setminus X) \nearrow (Y \setminus X), (X \cup Y_k) \nearrow (X \cup Y)$$

(und analog für absteigende Grenzwerte).

Sei nun $A \in \mathfrak{A}$. Da \mathfrak{A} ein Mengerring ist, folgt

$$\mathfrak{A} \subset \kappa(\mathfrak{M}, A)$$

(denn ist auch $B \in \mathfrak{A}$, so sind auch $A \setminus B, B \setminus A$ und $A \cup B$ in $\mathfrak{A} \subset \mathfrak{M}$ und daher ist $B \in \kappa(\mathfrak{M}, A)$). Da $\kappa(\mathfrak{M}, A)$ eine monotone Klasse ist, die \mathfrak{A} enthält, und \mathfrak{M} die kleinste solche monotone Klasse ist, gilt

$$\mathfrak{M} \subset \kappa(\mathfrak{M}, A).$$

Es folgt also: Für alle $X \in \mathfrak{M}$ gilt $X \in \kappa(\mathfrak{M}, A)$ und also auch $A \in \kappa(\mathfrak{M}, X)$. Da dies für alle $A \in \mathfrak{A}$ gilt, folgt

$$\mathfrak{A} \subset \kappa(\mathfrak{M}, X)$$

und daher wie eben auch die gewünschte Bedingung $\mathfrak{M} \subset \kappa(\mathfrak{M}, X)$.

□

Korollar 3.10. Ist $\mathfrak{A} \subset \mathfrak{P}(\Omega)$ eine Mengenalgebra und \mathfrak{M} eine monotone Klasse mit $\mathfrak{A} \subset \mathfrak{M}$, dann gilt

$$\sigma(\mathfrak{A}) \subset \mathfrak{M}.$$

4 Inhalte, Prämaße und Maße

Wir wollen "Maße" für möglichst viele Mengen definieren. Daher ist es sinnvoll, auch "unendlich große" Mengen mit einzuschließen, und ihnen das "Maß" ∞ zuzuordnen. Da wir mit Maßen von Mengen auch rechnen wollen, erweitern wie die Zahlengerade wie folgt:

Definition 4.1. Die erweiterte Zahlengerade ist $\bar{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$. Wir setzen die Ordnung von \mathbb{R} auf $\bar{\mathbb{R}}$ fort durch

$$a \in \mathbb{R} \implies -\infty < a < \infty.$$

Für Addition und Multiplikation erklären wir:

$$\begin{aligned} a + (\pm\infty) &= \pm\infty + a = \pm\infty && \text{falls } a \in \mathbb{R} \\ a \cdot (\pm\infty) &= (\pm\infty) \cdot a = \pm\infty && \text{falls } 0 < a \leq \infty \\ a \cdot (\pm\infty) &= (\pm\infty) \cdot a = \mp\infty && \text{falls } -\infty \leq a < 0 \\ 0 \cdot (\pm\infty) &= (\pm\infty) \cdot 0 = 0. \end{aligned}$$

Oft arbeiten wir mit

$$[0, \infty] = \{a \in \bar{\mathbb{R}} \mid a \geq 0\} = [0, \infty[\cup \{\infty\}.$$

Ausdrücklich *nicht* definiert die Ausdrücke $\infty + (-\infty)$ und $-\infty + \infty$.

Definition 4.2 (Inhalte). Es sei $\mathfrak{A} \subset \mathfrak{P}(\Omega)$ ein Mengenring. Ein *Inhalt* auf \mathfrak{A} ist eine Funktion

$$\mu : \mathfrak{A} \rightarrow [0, \infty]$$

mit den Eigenschaften

- i) $\mu(\emptyset) = 0$,
- ii) sind $A, B \in \mathfrak{A}$ disjunkt, so gilt $\mu(A \cup B) = \mu(A) + \mu(B)$.
- Inhalte sind immer monoton, in dem Sinn, dass für $A \subset B$ immer $\mu(A) \leq \mu(B)$ gilt.

Wir können disjunkt zerlegen $B = A \cup (B \setminus A)$ und aus i) und ii) folgt $\mu(B) = \mu(A) + \mu(B \setminus A) \geq \mu(A)$.

- Da wir disjunkt zerlegen können als $A \cup B = A \cup (B \setminus A)$ und $B = (A \cap B) \cup (B \setminus A)$, folgt

$$\mu(A \cup B) = \mu(A) + \mu(B \setminus A) \quad \text{und} \quad \mu(B) = \mu(A \cap B) + \mu(B \setminus A).$$

Ist dann $\mu(A \cap B) < \infty$, so folgt

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

Mit diesen Konventionen gilt für Folgen (c_k) die uneigentlich divergieren $a + \lim c_k = \lim(a + c_k)$ und $a \lim c_k = \lim(ac_k)$.

Beachte außerdem, dass nicht alle Rechenregeln weiterhin gelten: Aus $a + \infty = b + \infty$ kann z.B. nicht geschlossen werden, dass a und b gleich sind!

Per Induktion folgt, dass für paarweise disjunkte Mengen $A_1, \dots, A_m \in \mathfrak{A}$ gilt, das $\mu(\bigcup_{i=1}^m A_i) = \sum_{i=1}^m \mu(A_i)$ gilt, d.h. ein Inhalt ist (*endlich*) *additiv*.

Definition 4.3 (σ -Additivität, Prämaße und Maße, endliche und σ -endliche Inhalte und Maße). Ein Inhalt μ auf einem Mengenring \mathfrak{A} heißt σ -additiv (oder abzählbar additiv), wenn für jede Folge (A_k) von paarweise disjunkten Mengen mit $A = \bigcup_{i=1}^{\infty} A_i$ gilt

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

Ein σ -additiver Inhalt heißt *Prä-Maß*. Ein σ -additiver Inhalt auf einer σ -Algebra heißt *Maß*.

Ein Inhalt (bzw. Maß) μ heißt *endlich*, wenn für alle $A \in \mathfrak{A}$ gilt $\mu(A) < \infty$ und er heißt σ -endlich, wenn es Folge $A_k \in \mathfrak{A}$ gibt mit

$$\bigcup_{k=1}^{\infty} A_k = \Omega \quad \text{und} \quad \mu(A_k) < \infty.$$

σ -endlich heißt also, dass die ganze Menge mit Mengen von endlichem Inhalt ausgeschöpft werden kann.

Satz 4.4. Es sei μ ein Inhalt auf einem Mengenring $\mathfrak{A} \subset \mathfrak{P}(\Omega)$. Wir betrachten folgende Aussagen:

- a) μ ist σ -additiv.
- b) μ ist σ -subadditiv, d.h. für jede Folge $A_k \in \mathfrak{A}$ mit $\bigcup_{k=1}^{\infty} A_k = A \in \mathfrak{A}$ gilt $\mu(A) \leq \sum_{k=1}^{\infty} \mu(A_k)$.
- c) μ ist stetig von unten, d.h. für aufsteigende $A_k \in \mathfrak{A}$ mit $A_k \nearrow A \in \mathfrak{A}$ gilt $\mu(A_k) \nearrow \mu(A)$.
- d) μ ist stetig von oben, d.h. für absteigende $B_k \in \mathfrak{A}$ mit $B_k \searrow B \in \mathfrak{A}$ und $\mu(B_1) < \infty$ gilt $\mu(B_k) \searrow \mu(B)$.

Dann gilt

$$a) \iff b) \iff c) \implies d).$$

Ist μ endlich, so gilt auch noch $d) \implies c)$.

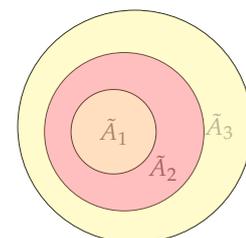
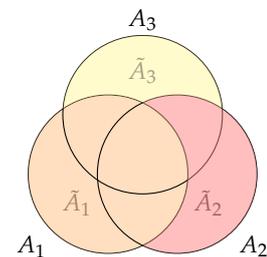
Beweis.

b) \implies a): Es sei $A_k \in \mathfrak{A}$ und $\bigcup_{k=1}^{\infty} A_k = A$. Ohne Einschränkung können wir annehmen, dass die A_k paarweise disjunkt sind (ansonsten betrachten wir $\tilde{A}_1 = A_1, \tilde{A}_2 = A_2 \setminus A_1, \tilde{A}_3 = A_3 \setminus (A_1 \cup A_2), \dots$). Dann gilt wegen der endlichen Additivität und Monotonie von μ

$$\sum_{k=1}^m \mu(A_k) = \mu\left(\bigcup_{k=1}^m A_k\right) \leq \mu(A).$$

Es folgt also $\sum_{k=1}^{\infty} \mu(A_k) \leq \mu(A)$ und wegen der σ -Subadditivität gilt sowieso $\mu(A) \leq \sum_{k=1}^{\infty} \mu(A_k)$ und damit Gleichheit.

a) \implies c): Ist $A_k \in \mathfrak{A}$ aufsteigend mit $A_k \nearrow A \in \mathfrak{A}$, so sind die Mengen $\tilde{A}_1 = A_1, \tilde{A}_k = A_k \setminus A_{k-1}$ ($k \geq 2$) paarweise



disjunkt mit $\bigcup_{k=1}^m \tilde{A}_k = A_m$ und $\bigcup_{k=1}^{\infty} \tilde{A}_k = A$. Aus der endlichen Additivität folgt

$$\sum_{k=1}^m \mu(\tilde{A}_k) = \mu(A_m)$$

und wegen a) gilt

$$\mu(A) = \sum_{k=1}^{\infty} \mu(\tilde{A}_k) = \lim_{m \rightarrow \infty} \sum_{k=1}^m \mu(\tilde{A}_k) = \lim_{m \rightarrow \infty} \mu(A_m).$$

c) \implies b): Ist $A_k \in \mathfrak{A}$ und $\bigcup_{k=1}^{\infty} A_k = A \in \mathfrak{A}$, so definieren wir $\tilde{A}_m = \bigcup_{k=1}^m A_k$. Es gilt dann immer

$$\mu(\tilde{A}_m) \leq \sum_{k=1}^m \mu(A_k). \quad (*)$$

Da nun $\tilde{A}_m \nearrow A$, folgt aus der Stetigkeit von unten $\mu(\tilde{A}_m) \nearrow \mu(A)$ und es folgt per Grenzübergang in (*)

$$\mu(A) \leq \sum_{k=1}^{\infty} \mu(A_k).$$

Damit haben wir die Äquivalenz von a), b) und c) gezeigt.

c) \implies d): Sei nun μ stetig von unten und $B_k \in \mathfrak{A}$ absteigend mit $B_k \searrow B \in \mathfrak{A}$. Wir definieren damit die aufsteigende Folge $A_k = B_1 \setminus B_k \in \mathfrak{A}$ und für diese gilt $A_k \nearrow B_1 \setminus B$. Aus der Stetigkeit von unten folgt $\mu(A_k) \nearrow \mu(B_1 \setminus B) = \mu(B_1) - \mu(B)$ und da $\mu(A_k) = \mu(B_1) - \mu(B_k)$ gilt, folgt $\mu(B_k) \searrow \mu(B)$.

Unter der Voraussetzung, dass μ endlich ist zeigen wir die letzte Implikation:

d) \implies c): Sei $\mu(A) < \infty$ für alle $A \in \mathfrak{A}$, und $A_k \in \mathfrak{A}$ aufsteigend mit $A_k \nearrow A \in \mathfrak{A}$. Dann sind die Mengen $B_k = A \setminus A_k$ absteigend mit $B_k \searrow \emptyset$. Wegen der Stetigkeit von oben gilt $\mu(B_k) \searrow 0$ und wegen $\mu(B_k) = \mu(A) - \mu(A_k)$ folgt $\mu(A_k) \nearrow \mu(A)$.

□

Beispiel. a) Es sei Ω eine Menge und \mathfrak{A} der Mengerring aller endlichen oder abzählbar unendlichen Teilmengen von Ω . Auf diesem Mengerring definieren wir das *Zählmaß*

$$\# : \mathfrak{A} \rightarrow [0, \infty], \quad \#(A) = \begin{cases} \text{Anzahl der Elemente in } A, & \text{falls } A \text{ endlich} \\ \infty, & \text{falls } A \text{ abzählbar unendlich.} \end{cases}$$

Das Zählmaß ist ein σ -additiver Inhalt auf \mathfrak{A} . Es ist σ -endlich genau dann, wenn Ω höchstens abzählbar unendlich ist.

- b) Ist Ω eine Menge, $a \in \Omega$ und $\mathfrak{A} \subset \mathfrak{P}(\Omega)$ eine σ -Algebra, so definieren wir

$$\delta_a(A) = \begin{cases} 1, & \text{falls } a \in A \\ 0, & \text{falls } a \notin A. \end{cases}$$

Das gibt ein Maß auf \mathfrak{A} und man nennt es das *Punktmaß in a* oder auch das *Dirac-Maß in a* .

- c) Ist \mathfrak{A} eine σ -Algebra und sind μ_1, \dots, μ_m Maße auf \mathfrak{A} und $c_1, \dots, c_m \geq 0$, so ist auch die Linearkombination

$$\mu := \sum_{k=1}^m c_k \mu_k, \quad \mu(A) = \sum_{k=1}^m c_k \mu_k(A)$$

ein Maß. Gleiches gilt für abzählbare Linearkombinationen mit nicht-negativen Koeffizienten.

- d) Ist Ω eine abzählbar unendliche Menge und

$$\mathfrak{A} = \{A \subset \Omega \mid A \text{ endlich oder } A^c \text{ endlich}\}$$

die Mengenalgebra der endlichen und co-endlichen Mengen, so definieren wir

$$\mu(A) = \begin{cases} 0, & \text{falls } A \text{ endlich} \\ 1, & \text{falls } A^c \text{ endlich.} \end{cases}$$

Das ist ein Inhalt, der nicht σ -additiv ist: Ist A abzählbar (mit endlichem Komplement), so ist $A = \{a_1, a_2, \dots\}$ und es gilt $A = \bigcup_{k=1}^{\infty} \{a_k\}$ (die Menge A ist die disjunkte abzählbare Vereinigung aller ihrer Punkte). Allerdings ist $\mu(A) = 1 \neq 0 = \sum_{k=1}^{\infty} \mu(\{a_k\})$

5 Das Lebesguesche Prämaß

Wir wollen nun anfangen, ein Maß zu konstruieren, welches dem natürlichen Inhaltsbegriff für Längen, Fläche und Volumina entspricht. Wir benutzen dafür den entsprechenden Begriff für Quader und benutzen also zuerst den Mengenring $\mathfrak{Q}(\mathbb{R}^d)$ der endlichen Quadersummen (siehe Beispiel in Abschnitt 2): Wir definieren den Inhalt

$$\lambda^d : \mathfrak{Q}(\mathbb{R}^d) \rightarrow [0, \infty[$$

wie folgt: Für

$$Q = [a_1, b_1[\times \cdots \times [a_d, b_d[, \quad a_i < b_i$$

setzen wir

$$\text{Vol}_d(Q) := \prod_{i=1}^d (b_i - a_i).$$

Ist $A \in \mathfrak{Q}(\mathbb{R}^d)$ eine endliche disjunkte Vereinigung von Quadern Q_1, \dots, Q_m , so setzen wir

$$\lambda^d(A) := \sum_{k=1}^m \lambda^d(Q_k) = \sum_{k=1}^m \text{Vol}_d(Q_k).$$

Die Additivität von λ^d auf $\mathfrak{Q}(\mathbb{R}^d)$ ist nach Definition klar, jedoch haben wir hier ein kleines technisches Problem: Es ist nicht klar, dass λ^d wohldefiniert ist, d.h. es ist nicht sofort klar, dass verschiedene Darstellungen einer Menge A als disjunkte Vereinigung von Quadern den gleichen Wert für $\lambda^d(A)$ ergeben. Um dies zu zeigen, zeigen wir zuerst ein Hilfsresultat: Wenn ein Quader auf verschiedene Weise als disjunkte Vereinigung von Quadern geschrieben wird, liefern beide Darstellungen gleichen Inhalt ergibt.

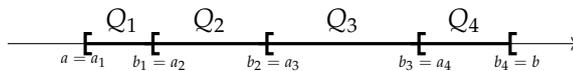
Lemma 5.1. *Es sei Q ein Quader in \mathbb{R}^d und $Q = \bigcup_{k=1}^m Q_k$ mit disjunkten Quadern Q_1, \dots, Q_m . Dann gilt*

$$\text{Vol}_d(Q) = \sum_{k=1}^m \text{Vol}_d(Q_k).$$

Beweis.

Wir führen Induktion nach d durch:

Anfang $d = 1$: Hier sind die Quader Intervalle $Q = [a, b[$ und $Q_k = [a_k, b_k[$ und wir sortieren die Quader Q_k von links nach rechts.



d.h. beide Darstellungen liefern den gleichen Wert für $\lambda^d(A)$.

Damit haben wir gezeigt:

$$\lambda^d : \mathfrak{Q}(\mathbb{R}^d) \rightarrow [0, \infty[\text{ ist ein Inhalt.}$$

Als Hilfresultat für das Folgende beweisen wir:

Lemma 5.2. *Ist $Q \subset \mathbb{R}^d$ ein Quader und $\epsilon > 0$, so gibt es weitere Quader Q', Q''*

$$\begin{aligned} \overline{Q'} \subset Q \subset (Q'')^\circ, \quad \text{Vol}_d(Q') &\geq (1 - \epsilon) \text{Vol}_d(Q) \\ \text{Vol}_d(Q'') &\leq (1 + \epsilon) \text{Vol}_d(Q). \end{aligned}$$

Beweis.

Ist $Q = [a_1, b_1[\times \cdots \times [a_d, b_d[$, so wählt man einfach

$$Q' = [a_1, b_1 - \delta[\times \cdots \times [a_d, b_d - \delta[, \quad Q'' = [a_1 - \delta, b_1[\times \cdots \times [a_d - \delta, b_d[,$$

denn für diese gilt

$$\begin{aligned} \text{Vol}_d(Q') &= \prod_{i=1}^d (b_i - \delta - a_i) = \text{Vol}_d(Q) - \delta \sum_{i=1}^d \prod_{j \neq i} (b_j - a_j) + \mathcal{O}(\delta^2) \\ \text{Vol}_d(Q'') &= \prod_{i=1}^d (b_i - (a_i - \delta)) = \text{Vol}_d(Q) + \delta \sum_{i=1}^d \prod_{j \neq i} (b_j - a_j) + \mathcal{O}(\delta^2). \end{aligned}$$

Da $\sum_{i=1}^d \prod_{j \neq i} (b_i - a_i) > 0$ ist, gibt es ein $\delta > 0$ (klein genug), so dass die Behauptung gilt. mit $\delta > 0$ klein genug. \square

Satz 5.3. *Der Inhalt $\lambda^d : \mathfrak{Q}(\mathbb{R}^d) \rightarrow [0, \infty[$ ist ein σ -endliches Prämaß und wir nennen λ^d das Lebesguesche Prämaß.*

Beweis.

Wir müssen zeigen, dass λ^d σ -additiv und σ -endlich ist.

- Nach Satz 3.3 reicht es, zu zeigen, dass λ^d σ -subadditiv ist. Es seien dazu $A, A_k \in \mathfrak{Q}(\mathbb{R}^d)$, $k \geq 1$ mit $A = \bigcup_{k=1}^{\infty} A_k$ und $\epsilon > 0$. Nach Lemma 5.2 gibt es $A', A_k'' \in \mathfrak{Q}(\mathbb{R}^d)$, so dass

$$\begin{aligned} \overline{A'} \subset A, \quad A_k \subset (A_k'')^\circ \\ \lambda^d(A') \geq (1 - \epsilon) \lambda^d(A), \quad \lambda^d(A_k'') \leq (1 + \epsilon) \lambda^d(A_k). \end{aligned}$$

Da $\overline{A'}$ kompakt ist und von den $(A_k'')^\circ$ überdeckt wird, reichen endlich viele $(A_k'')^\circ$ zur Überdeckung, d.h. es gilt insbesondere

$$A' \subset \bigcup_{k=1}^m A_k''.$$

Es folgt

$$(1 - \epsilon) \lambda^d(A) \leq \lambda^d(A') \leq \sum_{k=1}^m \lambda^d(A_k'') \leq (1 + \epsilon) \sum_{k=1}^{\infty} \lambda^d(A_k).$$

Per Grenzübergang $\epsilon \rightarrow 0$ folgt also

$$\lambda^d(A) \leq \sum_{k=1}^{\infty} \lambda^d(A_k).$$

Da.h. λ^d ist σ -subadditiv.

- Für $\Omega_m = [-m, m[\times \cdots \times [-m, m[$ gilt $\lambda^d(\Omega_m) = (2m)^d$ und

$$\mathbb{R}^d = \bigcup_{m=1}^{\infty} \Omega_m.$$

Daher ist λ^d σ -endlich.

□

Unser weiteres Vorgehen ist es nun, das Lebesguesche Prämaß zu einem Maß auf einer möglichst großen σ -Algebra zu erweitern. Hierfür gibt es eine allgemeine Konstruktion zur Fortsetzung von Prämaßen, die wir in den nächsten beiden Abschnitten durchführen.

6 Fortsetzung von Prämaßen und äußere Maße

In diesem Abschnitt gehen wir davon aus, dass $\mathfrak{A} \subset \mathfrak{P}(\Omega)$ ein Mengenring ist und wir ein σ -endliches Prämaß $\mu : \mathfrak{A} \rightarrow [0, \infty]$ gegeben haben. Unser Ziel ist es, dieses Prämaß zu einem Maß

$$\tilde{\mu} : \sigma(\mathfrak{A}) \rightarrow [0, \infty]$$

fortzusetzen. Dies verläuft in mehreren Schritten.

Als erstes setzen wir μ auf die Menge \mathfrak{A}^\uparrow fort, welche aus allem Mengen besteht, die wir als aufsteigenden Grenzwert bekommen können, d.h.

$$\mathfrak{A}^\uparrow := \{A \subset \Omega \mid \exists A_k \in \mathfrak{A} : A_k \nearrow A\}.$$

Da μ σ -endlich ist, gilt $\Omega \in \mathfrak{A}^\uparrow$. Außerdem gilt

$$\mathfrak{A} \subset \mathfrak{A}^\uparrow \subset \sigma(\mathfrak{A}).$$

Der naheliegende Versuch für eine Fortsetzung ist

$$\mu^\uparrow(A) := \lim_{k \rightarrow \infty} \mu(A_k) \quad \text{falls } A_k \nearrow A, A_k \in \mathfrak{A}.$$

Insbesondere ist für $A \in \mathfrak{A}$ immer $\mu^\uparrow(A) = \mu(A)$ (da wir $A_k = A$ wählen können) und μ^\uparrow ist damit tatsächlich eine Fortsetzung.

Wir müssen hier allerdings wieder zeigen, dass μ^\uparrow wohldefiniert ist, d.h. unabhängig von der speziellen Folge A_k :

Beweis.

(Beweis der Wohldefiniertheit von μ^\uparrow .) Gelte $A_k \nearrow A, B_k \nearrow A, A_k, B_k \in \mathfrak{A}$. Fixieren wir k , so gilt (da $A_k \subset A$)

$$B_j \cap A_k \nearrow_{j=1}^{\infty} A_k.$$

Da μ σ -additiv ist, ist es nach Satz 3.3 stetig von unten und es folgt

$$\lim_{j \rightarrow \infty} \mu(B_j \cap A_k) = \mu(A_k)$$

Es gilt aber auch $\mu(B_j \cap A_k) \leq \mu(B_j)$ und daher ist $\lim_{j \rightarrow \infty} \mu(B_j) \geq \lim_{j \rightarrow \infty} \mu(B_j \cap A_k) = \mu(A_k)$ für alle k . Daher können wir auch rechts zum Grenzwert $k \rightarrow \infty$ übergehen und bekommen

$$\lim_{j \rightarrow \infty} \mu(B_j) \geq \lim_{k \rightarrow \infty} \mu(A_k).$$

Führen wir die gleiche Argumentation mit vertauschten Rolle von B_j und A_k durch, erhalten wir die umgekehrte Ungleichung und damit Gleichheit. \square

Untersuchen wir das Mengensystem \mathfrak{A}^\uparrow :

Satz 6.1. Für \mathfrak{A}^\uparrow gilt

i) Für $A_1, \dots, A_m \in \mathfrak{A}^\uparrow$ gilt $A_1 \cap \dots \cap A_m \in \mathfrak{A}^\uparrow$.

Es gilt für beliebige $B_k \in \mathfrak{A}$ immer $\bigcup_{k=1}^m B_k \nearrow \bigcup_{k=1}^{\infty} B_k$. Also ist \mathfrak{A}^\uparrow die Menge aller $A \subset \Omega$, welche sich als Vereinigung von abzählbar vielen Mengen in \mathfrak{A} schreiben lassen.

Der Grenzwert auf der rechten Seite existiert immer in $[0, \infty]$, da eine wachsende Folge vorliegt.

ii) Für $A_k \in \mathfrak{A}^\uparrow, k \geq 1$ gilt $\bigcup_{k=1}^\infty A_k \in \mathfrak{A}^\uparrow$.

iii) Ist $A \in \mathfrak{A}^\uparrow$ und $B \in \mathfrak{A}$, so ist $A \setminus B \in \mathfrak{A}^\uparrow$,

Beweis.

i) Gilt $A_k \nearrow A, B_k \nearrow B$, so ist $A_k \cap B_k$ aufsteigend und es gilt $A_k \cap B_k \nearrow A \cap B$ und es folgt $A \cap B \in \mathfrak{A}^\uparrow$.

ii) Für jedes k gilt $A_k = \bigcup_{i=1}^\infty A_{k,i}$ mit $A_{k,i} \in \mathfrak{A}$. Also ist $A = \bigcup_{k=1}^\infty \bigcup_{i=1}^\infty A_{k,i}$, d.h. A ist eine abzählbare Vereinigung von Mengen aus \mathfrak{A} und damit in \mathfrak{A}^\uparrow .

iii) Gilt $A_k \nearrow A, A_k \in \mathfrak{A}$, so ist $A_k \setminus B \in \mathfrak{A}$ und $A_k \setminus B \nearrow A \setminus B$. Es folgt also $A \setminus B \in \mathfrak{A}^\uparrow$.

□

Es folgt nicht, dass \mathfrak{A}^\uparrow ein Mengerring ist, das aus $A, B \in \mathfrak{A}^\uparrow$ im Allgemeinen nicht folgt, dass $A \setminus B$ in \mathfrak{A}^\uparrow liegt.

Satz 6.2. Für die Fortsetzung $\mu^\uparrow : \mathfrak{A}^\uparrow \rightarrow [0, \infty]$ von $\mu : \mathfrak{A} \rightarrow [0, \infty]$ gilt

a) μ^\uparrow ist monoton (vgl. Bemerkung nach Definition 4.2).

b) μ^\uparrow ist endlich additiv (vgl. Definition 4.2).

c) μ^\uparrow ist stetig von unten (vgl. Satz 4.4).

d) μ^\uparrow ist σ -subadditiv (vgl. Satz 4.4).

Beweis.

a) Ist $A \subset B, A_k \nearrow A$ und $B_k \nearrow B$ (mit $A_k, B_k \in \mathfrak{A}$), so ist $A_k \cup B_k \nearrow B$. Es folgt $\mu^\uparrow(A) = \lim_{k \rightarrow \infty} \mu(A_k) \leq \lim_{k \rightarrow \infty} \mu(A_k \cup B_k) = \mu^\uparrow(B)$.

b) Sind $A, B \in \mathfrak{A}^\uparrow$ disjunkt und $A_k \nearrow A$ und $B_k \nearrow B$ (mit $A_k, B_k \in \mathfrak{A}$), so sind auch A_k und B_k jeweils disjunkt. Also gilt $\mu(A_k \cup B_k) = \mu(A_k) + \mu(B_k)$ und da $A_k \cup B_k \nearrow A \cup B$, folgt die Behauptung.

c) Es sei $A = \bigcup_{k=1}^\infty A_k$ mit $A_k \in \mathfrak{A}^\uparrow$. Für jedes k gibt es $A_{ki} \in \mathfrak{A}, i \geq 1$ mit $A_{ki} \nearrow_{i=1}^\infty A_k$. Wir definieren

$$\tilde{A}_m = \bigcup_{k,i \leq m} A_{ki}$$

und bemerken $\tilde{A}_m \in \mathfrak{A}, \tilde{A}_m \nearrow A$. Also ist nach Definition der Fortsetzung $\mu^\uparrow(A) = \lim_{m \rightarrow \infty} \mu(\tilde{A}_m)$. Also gibt es zu jedem $M <^\uparrow \mu(A)$ ein $m \geq 1$, so dass $M < \mu(\tilde{A}_m)$. Wegen $\tilde{A}_m \subset \bigcup_{k=1}^m A_k \subset A$ folgt

$$M < \mu(\tilde{A}_m) \leq \mu\left(\bigcup_{k=1}^m A_k\right) = \mu^\uparrow\left(\bigcup_{k=1}^m A_k\right) \leq \mu^\uparrow(A).$$

Da dies für alle $M < \mu^\uparrow(A)$ gilt, folgt $\lim_{m \rightarrow \infty} \mu\left(\bigcup_{k=1}^m A_k\right) = \mu^\uparrow(A)$.

d) Wie in b) zeigt man $\mu(A_1 \cup A_2) \leq \mu(A_1) + \mu(A_2)$, woraus durch Induktion folgt $\mu(\bigcup_{k=1}^m A_k) \leq \sum_{k=1}^m \mu(A_k)$. Durch Grenzwertbildung $m \rightarrow \infty$ folgt mit c) die Behauptung.

□

Wir haben nun μ von \mathfrak{A} auf \mathfrak{A}^\uparrow fortgesetzt. Leider ist \mathfrak{A}^\uparrow noch keine σ -Algebra, nämlich noch nicht einmal ein Mengerring. Wir setzen nun μ^\uparrow weiter fort, sogar auf die ganze Potenzmenge $\mathfrak{P}(\Omega)$. Da wir im Folgenden nicht mehr mit μ und μ^\uparrow gleichzeitig arbeiten, sondern für alle $A \in \mathfrak{A}^\uparrow$ die Fortsetzung verwenden, lassen wir den Pfeil an μ ab hier wieder weg, d.h. wir schreiben μ statt μ^\uparrow auch auf \mathfrak{A}^\uparrow .

Definition 6.3. Für ein wie eben erklärt fortgesetztes Prämaß $\mu : \mathfrak{A} \rightarrow [0, \infty]$ auf \mathfrak{A}^\uparrow erklären wir das *äußere Maß* $\mu^* : \mathfrak{P}(\Omega) \rightarrow [0, \infty]$ durch

$$\mu^*(X) = \inf\{\mu(A) \mid A \supset X, A \in \mathfrak{A}^\uparrow\}.$$

Aus Satz 6.2 a) folgt, dass für $A \in \mathfrak{A}^\uparrow$ gilt $\mu(A) = \mu^*(A)$.

Satz 6.4. Für μ^* gilt

- a) $\mu^*(\emptyset) = 0$
- b) μ^* ist monoton.
- c) μ^* ist σ -subadditiv.

Beweis.

- a) Klar, da $\emptyset \in \mathfrak{A}^\uparrow$ und $\mu(\emptyset) = 0$.
- b) Klar, da für $X \subset Y$ die Menge der A mit $A \supset X$ größer ist, als die Menge der A mit $A \supset Y$.
- c) Zu zeigen ist: $\mu^*(\bigcup_{k=1}^\infty X_k) \leq \sum_{k=1}^\infty \mu^*(X_k)$. Ist $\mu^*(X_k) = \infty$, für mindestens ein k , so ist die Ungleichung immer erfüllt. Sei also $\mu^*(X_k)$ für alle k endlich und $\epsilon > 0$. Nach Definition von μ^* als Infimum gibt es $B_k \in \mathfrak{A}^\uparrow$ mit $X_k \subset B_k$ und

$$\mu^*(X_k) \leq \mu(B_k) \leq \mu^*(X_k) + \epsilon/2^k.$$

Dann ist $\bigcup_{k=1}^\infty X_k \subset \bigcup_{k=1}^\infty B_k$ zusammen mit Monotonie von μ^* und Satz 6.2 d) folgt

$$\begin{aligned} \mu^*\left(\bigcup_{k=1}^\infty X_k\right) &\leq \mu^*\left(\bigcup_{k=1}^\infty B_k\right) = \mu\left(\bigcup_{k=1}^\infty B_k\right) \leq \sum_{k=1}^\infty \mu(B_k) \\ &\leq \sum_{k=1}^\infty \left[\mu^*(X_k) + \frac{\epsilon}{2^k}\right] = \sum_{k=1}^\infty \mu^*(X_k) + \epsilon, \end{aligned}$$

woraus die Behauptung folgt.

□

Der für uns wichtigste Fall ist der folgende:

Beispiel. Wir betrachten Lebesgueschen Prämaß $\lambda^d : \mathfrak{Q}(\mathbb{R}^d) \rightarrow [0, \infty[$ aus Satz 5.3. Hier ist $\mathfrak{Q}(\mathbb{R}^d)^\uparrow$ die Menge aller abzählbaren Vereinigungen von Quadern. Das entsprechende äußere Maß ist

$$(\lambda^d)^*(X) = \inf\{\lambda^d(A) \mid A \in \mathfrak{Q}(\mathbb{R}^d)^\uparrow, A \supset X\}.$$

Da $\mathfrak{Q}(\mathbb{R}^d)^\uparrow$ nach Lemma 3.5 alle offenen Mengen umfasst erscheint es plausibel, dass

$$(\lambda^d)^*(X) = \inf\{\lambda^d(U) \mid U \text{ offen und } U \supset X\}. \quad (*)$$

Zeigen wir, dass (*) gilt:

Die Ungleichung " \leq " ist klar (da in $(\lambda^d)^*$ über mehr Mengen ein Infimum gebildet wird). Aus Lemma 5.2 folgt allerdings, dass zu jeder abzählbaren Quadersumme A und für jedes $\epsilon > 0$ ein offenes U mit $A \subset U$ und $\lambda^d(U) \leq (1 + \epsilon)\lambda^d(A)$ existiert. Daher lässt sich für jedes X die rechte Seite in (*) nach oben durch $(1 + \epsilon)(\lambda^d)^*(X)$ abschätzen, was die Behauptung zeigt.

Wir nennen $(\lambda^d)^*$ das *Lebesguesche äußere Maß*.

Definition 6.5. Im Allgemeinen nennt man jede Abbildung $\mu^* : \mathfrak{P}(\Omega) \rightarrow [0, \infty]$ mit $\mu^*(\emptyset) = 0$ ein *äußeres Maß*, wenn μ^* monoton und σ -subadditiv ist.

Satz 6.6. Ist $\mu^* : \mathfrak{P}(\Omega) \rightarrow [0, \infty]$ ein äußeres Maß, so gilt:

i) Für alle $X, Y, Z \in \mathfrak{P}(\Omega)$ gilt

$$\mu^*(X \Delta Z) \leq \mu^*(X \Delta Y) + \mu^*(Y \Delta Z).$$

ii) Sind $X, Y \in \mathfrak{P}(\Omega)$ mit $\mu^*(X), \mu^*(Y) < \infty$, so gilt

$$|\mu^*(X) - \mu^*(Y)| \leq \mu^*(X \Delta Y).$$

Beweis. i) Es gilt immer $X \Delta Z = (X \Delta Y) \Delta (Y \Delta Z) \subset (X \Delta Y) \cup (Y \Delta Z)$. Also folgt die Behauptung aus Monotonie und Subadditivität von μ^* .

ii) Es gilt $X \cup Y = (X \cap Y) \cup (X \Delta Y)$ und daher $\mu^*(X \cup Y) \leq \mu^*(X \cap Y) + \mu^*(X \Delta Y)$. Es folgt also

$$0 \leq \mu^*(X \cup Y) - \mu^*(X \cap Y) \leq \mu^*(X \Delta Y).$$

Weiterhin ist $X \cap Y \subset X \subset X \cup Y$ (und ebenso für Y) und daher ist gilt

$$\mu^*(X \cap Y) \leq \mu^*(X), \mu^*(Y) \leq \mu^*(X \cup Y).$$

Es folgt (zusammen mit der ersten Ungleichung)

$$\mu^*(X) - \mu^*(Y) \leq \mu^*(X \cup Y) - \mu^*(X \cap Y) \leq \mu^*(X \Delta Y).$$

Vertauschen von X und Y zeigt $\mu^*(Y) - \mu^*(X) \leq \mu^*(X \Delta Y)$ und beides zusammen liefert die Behauptung. \square

Direkt aus Teil a) des Satzes erhält man:

Korollar 6.7. Ist μ^* ein äußeres Maß auf $\mathfrak{P}(\Omega)$, so ist

$$d(X, Y) := \mu^*(X \Delta Y)$$

eine Halbmetrik (auch Pseudometrik) auf $\mathfrak{P}(\Omega)$, d.h. es gilt

1. $X = Y \implies d(X, Y) = 0$,
2. $d(X, Y) = d(Y, X)$,
3. $d(X, Z) \leq d(X, Y) + d(Y, Z)$.

Wie jede Metrik, liefert auch diese Metrik einen Konvergenzbegriff, nämlich

$$Y_k \xrightarrow{\mu^*} X, \quad \text{falls } d(Y_k, X) \rightarrow 0.$$

Aus Teil b) von Satz 6.6 folgt, dass μ^* stetig (sogar Lipschitz-stetig mit Konstante 1) bezüglich dieser Metrik ist, d.h.

$$Y_k \xrightarrow{\mu^*} X \implies \mu^*(Y_k) \rightarrow \mu^*(X).$$

Der einzige Unterschied von Halbmetrik und Metrik ist, dass bei einer Halbmetrik aus $d(X, Y) = 0$ nicht $X = Y$ folgen muss.

7 Von äußeren Maßen zu Maßen

Definition 7.1. Es sei $\mathfrak{A} \subset \mathfrak{P}(\Omega)$ ein Mengerring, $\mu : \mathfrak{A} \rightarrow [0, \infty[$ ein Prämaß, und $\mu^* : \mathfrak{P}(\Omega) \rightarrow [0, \infty]$ das zugehörige äußere Maß (nach Definition 6.3), so nennen wir $X \subset \Omega$ \mathfrak{A} -*approximierbar* bezüglich μ^* , falls es zu jedem $\epsilon > 0$ ein $A \in \mathfrak{A}$ gibt, sodass

$$\mu^*(X \Delta A) < \epsilon.$$

Für endliche Prämaße haben wir vorerst:

Satz 7.2. Es sei $\mathfrak{A} \subset \mathfrak{P}(\Omega)$ eine Mengenalgebra, $\mu : \mathfrak{A} \rightarrow [0, \infty[$ ein endliches Prämaß, und $\mu^* : \mathfrak{P}(\Omega) \rightarrow [0, \infty]$ das zugehörige äußere Maß. Dann ist das Mengensystem

$$\tilde{\mathfrak{A}} := \{X \in \mathfrak{P}(\Omega) \mid X \text{ ist } \mathfrak{A}\text{-approximierbar bezüglich } \mu^*\}$$

eine σ -Algebra.

Beweis.

1. Abgeschlossenheit bezüglich Komplementen: Es gilt $X \Delta A = X^c \Delta A^c$. Wird also X durch $A_k \in \mathfrak{A}$ approximiert, so wird X^c durch $A_k^c \in \mathfrak{A}$ approximiert. Also gilt: Ist $X \in \tilde{\mathfrak{A}}$, so ist $X^c \in \tilde{\mathfrak{A}}$.
2. Abgeschlossenheit bezüglich abzählbarer Vereinigungen: Es sei $X_1, X_2, \dots \in \tilde{\mathfrak{A}}$. Zu zeigen ist $X := \bigcup_{k=1}^{\infty} X_k \in \tilde{\mathfrak{A}}$. Da jedes X_k \mathfrak{A} -approximierbar bezüglich μ^* ist, so gibt es zu jedem $\epsilon > 0$ Mengen $A_k \in \mathfrak{A}$, so dass $\mu^*(X_k \Delta A_k) < \epsilon/2^k$. Wir setzen $\tilde{A}_m := \bigcup_{k=1}^m A_k$ und $A := \bigcup_{k=1}^{\infty} A_k \in \mathfrak{A}^\uparrow$. Dann gilt $\tilde{A}_m \nearrow A$ und (nach Satz 6.2 c)) auch $\mu(\tilde{A}_m) \nearrow \mu(A) < \infty$.

Wir wählen nun m_0 so groß, dass $\mu(A) - \mu(\tilde{A}_{m_0}) < \epsilon$ gilt. Wir setzen dann

$$A' = A \setminus \tilde{A}_{m_0} = \bigcup_{k=m_0+1}^{\infty} A_k \in \mathfrak{A}^\uparrow.$$

Da $A = A' \cup \tilde{A}_{m_0}$ eine diskunkte Vereinigung ist, gilt $\mu(A) = \mu(A') + \mu(\tilde{A}_{m_0})$ und also ist $\mu(A') = \mu(A) - \mu(\tilde{A}_{m_0}) < \epsilon$.

Es gilt

$$X \Delta \tilde{A}_{m_0} = X \Delta (A \Delta A') = (X \Delta A) \Delta A' \subset (X \Delta A) \cup A'.$$

Nach Satz 2.1 viii) gilt

$$X \Delta A = \left(\bigcup_{k=1}^{\infty} X_k \right) \Delta \left(\bigcup_{k=1}^{\infty} A_k \right) \subset \bigcup_{k=1}^{\infty} (X_k \Delta A_k)$$

und mit der vorigen Gleichung folgt

$$X \Delta \tilde{A}_{m_0} \subset \bigcup_{k=1}^{\infty} (X_k \Delta A_k) \cup A'.$$

Anders gesagt: $X \in \mathfrak{P}(\Omega)$ heißt \mathfrak{A} -approximierbar bezüglich μ^* , wenn es eine Folge $A_k \in \mathfrak{A}$ gibt, mit $A_k \xrightarrow{\mu^*} X$, d.h. $\mu^*(X \Delta A_k) \rightarrow 0$, woraus nach Satz 6.6 ii) auch $\mu(A_k) \rightarrow \mu^*(X)$ folgt.

Da jedes $X \in \mathfrak{A}$ selbst \mathfrak{A} -approximierbar bezüglich μ^* ist, gilt $\mathfrak{A} \subset \tilde{\mathfrak{A}}$. Aus Satz 7.2 folgt dann $\mathfrak{A} \subset \mathfrak{A}^\uparrow \subset \sigma(\mathfrak{A}) \subset \tilde{\mathfrak{A}}$.

Aus der Subadditivität von μ^* folgt also

$$\mu^*(X \triangle \tilde{A}_{m_0}) \leq \sum_{k=1}^{\infty} \mu^*(X_k \triangle A_k) + \mu^*(A') < \sum_{k=1}^{\infty} \epsilon/2^k + \epsilon = 2\epsilon.$$

Es folgt $X \in \tilde{\mathfrak{A}}$ wie gefordert. □

Satz 7.3. Ist $\mathfrak{A} \subset \mathfrak{P}(\Omega)$ eine Mengenalgebra, $\mu : \mathfrak{A} \rightarrow [0, \infty[$ ein endliches Prämaß und $\tilde{\mathfrak{A}}$ die σ -Algebra der \mathfrak{A} -approximierbaren Mengen. Dann lässt sich μ eindeutig zu einem Maß $\tilde{\mu} : \tilde{\mathfrak{A}} \rightarrow [0, \infty]$ fortsetzen, und es gilt für alle $A \in \tilde{\mathfrak{A}}$

$$\tilde{\mu}(A) = \mu^*(A)$$

mit dem μ zugeordneten äußeren Maß μ^* .

Beweis.

1. Eindeutigkeit der Fortsetzung: Es seien $\mu_1, \mu_2 : \tilde{\mathfrak{A}} \rightarrow]0, \infty[$ zwei Fortsetzung von μ . Dann gilt für alle $B \in \mathfrak{A}^\uparrow$

$$\mu_1(B) = \mu_2(B) = \mu(B).$$

Sei nun $X \in \tilde{\mathfrak{A}}$. Da X \mathfrak{A} -approximierbar ist, gibt es $A \in \mathfrak{A}$ und $B \in \mathfrak{A}^\uparrow$, so dass

$$X \triangle A \subset B \quad \text{und} \quad \mu(B) < \epsilon.$$

Es folgt $A \setminus B \subset X \subset A \cup B$ und daher für $i = 1, 2$

$$\mu_i(A \setminus B) \leq \mu_i(X) \leq \mu_i(A \cup B).$$

Es gilt aber

$$\mu_i(A \setminus B) \geq \mu_i(A) - \mu_i(B) = \mu(A) - \mu(B) > \mu(A) - \epsilon$$

und

$$\mu_i(A \cup B) \leq \mu_i(A) + \mu_i(B) = \mu(A) + \mu(B) < \mu(A) + \epsilon.$$

Zusammen folgt

$$\mu(A) - \epsilon < \mu_i(X) < \mu(A) + \epsilon, \quad i = 1, 2$$

und damit

$$|\mu_1(X) - \mu_2(X)| < 2\epsilon.$$

Da dies für alle $\epsilon > 0$ gilt, folgt $\mu_1(X) = \mu_2(X)$.

2. Zeige (endliche) Additivität von μ^* in $\tilde{\mathfrak{A}}$: Es seien $X, Y \in \tilde{\mathfrak{A}}$ disjunkt und wir wollen zeigen, dass $\mu^*(X \cup Y) = \mu^*(X) + \mu^*(Y)$. Wir wählen dazu $A_k, B_k \in \mathfrak{A}$ mit $\mu^*(A_k \triangle X), \mu^*(B_k \triangle Y) \rightarrow 0$

0. Aus Satz 2.1 viii) folgt $(X \cup Y) \Delta (A_k \cup B_k) \subset (X \Delta A_k) \cup (Y \Delta B_k)$ und daher

$$\mu^*((X \cup Y) \Delta (A_k \cup B_k)) \leq \mu^*(X \Delta A_k) + \mu^*(Y \Delta B_k) \rightarrow 0.$$

Es folgt $\mu^*(A_k \cup B_k) \rightarrow \mu^*(X \cup Y)$.

Außerdem gilt nach Satz 2.1 vii) $A_k \cap B_k \subset (X \Delta A_k) \cup (Y \Delta B_k)$ und daher

$$\mu(A_k \cap B_k) \leq \mu^*(X \Delta A_k) + \mu^*(Y \Delta B_k) \rightarrow 0.$$

Es folgt

$$\lim_{k \rightarrow \infty} \mu(A_k \cup B_k) = \lim_{k \rightarrow \infty} (\mu(A_k) + \mu(B_k) - \mu(A_k \cap B_k)) = \lim_{k \rightarrow \infty} \mu(A_k) + \lim_{k \rightarrow \infty} \mu(B_k)$$

und daher gilt $\mu^*(X \cup Y) = \mu^*(X) + \mu^*(Y)$.

3. Da $\mu^*|_{\tilde{\mathfrak{A}}}$ nach obigem (endlich) additiv ist und insgesamt (als äußeres Maß) σ -subadditiv, folgt aus Satz 3.3, dass μ^* auf $\tilde{\mathfrak{A}}$ σ -additiv ist. Damit ist gezeigt, dass μ^* auf $\tilde{\mathfrak{A}}$ ein Maß ist.

□

Bleibt noch, die Konstruktion auf den Fall zu erweitern, dass das Prämaß nur σ -endlich ist und nicht endlich (denn der Fall ist für das Lebesguesche Prämaß relevant):

Satz 7.4. Sei $\mathfrak{A} \subset \mathfrak{P}(\Omega)$ ein Mengenring auf Ω und $\mu : \mathfrak{A} \rightarrow [0, \infty]$ ein σ -endliches Prämaß. Dann ist $\tilde{\mu} = \mu^*|_{\mathfrak{B}}$ mit $\mathfrak{B} = \sigma(\mathfrak{A})$ die eindeutige Fortsetzung von μ zu einem Maß auf der σ -Algebra \mathfrak{B} .

Beweis.

Wir setzen das Ergebnis aus dem vorigen Satz auf den nicht-endlichen (aber σ -endlichen Fall) fort: Da μ σ -endlich ist, gibt es eine aufsteigende Folge $\Omega_m \nearrow \Omega$ mit $\Omega_m \in \mathfrak{A}$ und $\mu(\Omega_m) < \infty$. Damit definieren wir

$$\mathfrak{A}_m := \mathfrak{A} \cap \Omega_m, \quad \mathfrak{B}_m := \mathfrak{B} \cap \Omega_m.$$

Es gilt: \mathfrak{A}_m ist sogar eine Mengenalgebra (da $\Omega_m \in \mathfrak{A}$ ist) und sogar $\mathfrak{B}_m = \sigma(\mathfrak{A}_m)$. Außerdem ist $\mu|_{\mathfrak{A}_m}$ ein endliches Prämaß. Dieses setzen wir mit Satz 7.3 zu $\tilde{\mu}_m : \mathfrak{B}_m \rightarrow]0, \infty[$ fort und da diese Fortsetzung eindeutig ist, folgt $\tilde{\mu}_m|_{\mathfrak{B}_{m-1}} = \tilde{\mu}_{m-1}$.

Für jedes $X \in \mathfrak{B} = \sigma(\mathfrak{A})$ gilt $X \cap \Omega_m \nearrow X$ und also muss eine Fortsetzung $\tilde{\mu} : \mathfrak{B} \rightarrow [0, \infty]$ die Forderung

$$\tilde{\mu}(X) = \sup_m \tilde{\mu}_m(X \cap \Omega_m)$$

erfüllen (was schon die Eindeutigkeit zeigt). Daher definieren wir $\tilde{\mu}$ auf diese Weise und es bleibt zu zeigen, dass $\tilde{\mu}$ ein Maß ist.

Zeige $\tilde{\mu}(X) = \mu^*(X)$: Wir setzen

$$X_m := X \cap \Omega_m, \quad Y_1 := X_1, \quad Y_m := X_m \setminus X_{m-1}, \quad m \geq 2.$$

Die Fortsetzung eines Prämaßes auf einem Mengenring zu einem Maß auf einer σ -Algebra geht also wie folgt: Setze das Prämaß zu einem äußeren Maß (auf der ganzen Potenzmenge) fort, und schränke dieses auf die von \mathfrak{A} erzeugte σ -Algebra ein.

Nach Satz 7.3 ist $\mu^*(X_m) = \tilde{\mu}_m(X_m)$ und $\mu^*(Y_m) = \tilde{\mu}_m(Y_m)$. Außerdem ist $X_m = \bigcup_{k=1}^m Y_k$ und $X = \bigcup_{k=1}^{\infty} Y_k$ und die Vereinigungen sind beide disjunkt. Die σ -Subadditivität von μ^* liefert

$$\mu^*(X) \leq \sum_{m=1}^{\infty} \mu^*(Y_m) = \sup_m \sum_{k=1}^m \mu^*(Y_k) = \sup_m \sum_{k=1}^m \tilde{\mu}_m(Y_k) = \sup_m \tilde{\mu}_m(X_m) = \tilde{\mu}(X).$$

Andererseits ist

$$\tilde{\mu}(X) = \sup_m \tilde{\mu}_m(X_m) = \sup_m \mu^*(X_m) \leq \mu^*(X).$$

Zusammen folgt $\mu^*(X) = \tilde{\mu}(X)$ wie gewünscht.

Zeige: $\tilde{\mu}$ ist additiv. Seien dazu X und Y disjunkt. Wir setzen

$$X_m := X \cap \Omega_m, \quad Y_m := Y \cap \Omega_m.$$

Dann gilt

$$\tilde{\mu}_m(X_m) \nearrow \tilde{\mu}(X), \quad \tilde{\mu}_m(Y_m) \nearrow \tilde{\mu}(Y), \quad \tilde{\mu}_m(X_m \cup Y_m) \nearrow \tilde{\mu}(X \cup Y).$$

Da $\tilde{\mu}_m$ ein Maß ist, gilt $\tilde{\mu}_m(X_m \cup Y_m) = \tilde{\mu}_m(X_m) + \tilde{\mu}_m(Y_m)$ und es folgt

$$\tilde{\mu}(X \cup Y) = \tilde{\mu}(X) + \tilde{\mu}(Y).$$

Schließlich können wir wieder Satz 3.3 anwenden: Aus dem vorigen folgt, dass $\tilde{\mu}$ ein Inhalt ist und da μ^* σ -subadditiv ist und $\tilde{\mu} = \mu^*|_{\mathfrak{B}}$ gilt, muss $\tilde{\mu}$ schon σ -additiv sein. \square

Korollar 7.5. Jedes $X \in \sigma(\mathfrak{A})$ mit $\tilde{\mu}(X) < \infty$ ist \mathfrak{A} -approximierbar.

Wir müssen zeigen, dass zu jedem $\epsilon > 0$ ein $A \in \mathfrak{A}$ existiert mit $\mu^*(X \Delta A) < \epsilon$. Falls $X \subset \Omega_m$ für ein m gilt, ist das klar. Falls nicht, bilden wir $X_m := X \cap \Omega_m$. Dann gilt $\tilde{\mu}(X_m) \nearrow \tilde{\mu}(X)$ und es gibt ein m , so dass

$$\tilde{\mu}(X) - \tilde{\mu}(X_m) < \epsilon/2.$$

Da in diesem Fall $X \setminus X_m = X \Delta X_m$ gilt, folgt $\mu^*(X \Delta X_m) < \epsilon/2$. Da $X_m \subset \Omega_m$ gilt, existiert $A \in \mathfrak{A}$, mit $\mu^*(X_m \Delta A) < \epsilon/2$ und zusammen bekommen wir

$$\mu^*(X \Delta A) \leq \mu^*(X \Delta X_m) + \mu^*(X_m \Delta A) < \epsilon$$

wie gewünscht.

Wir wenden den Prozess der Maßfortsetzung nun auf das Lebesguesche Prämaß $\lambda^d : \Omega(\mathbb{R}^d) \rightarrow [0, \infty[$ an.

Definition 7.6. Die durch Satz 7.4 garantierte Fortsetzung auf die Borelsche σ -Algebra $\mathcal{B}(\mathbb{R}^d) = \sigma(\Omega(\mathbb{R}^d))$ bezeichnen wir wieder mit $\lambda^d : \mathcal{B}(\mathbb{R}^d) \rightarrow [0, \infty]$ und nennen es *Lebesgue-Borelsches Maß*.

8 Messbare und Maß-Räume

Definition 8.1. Ein *messbarer Raum* ist ein Paar (Ω, \mathfrak{A}) bestehend aus einer Menge Ω versehen mit einer σ -Algebra $\mathfrak{A} \subset \mathfrak{P}(\Omega)$. Die Mengen $A \in \mathfrak{A}$ nennen wir *\mathfrak{A} -messbare Mengen* (oft nur *messbare Mengen*).

Ein *Maßraum* ist eine Tripel $(\Omega, \mathfrak{A}, \mu)$, wobei (Ω, \mathfrak{A}) ein messbarer Raum und $\mu : \mathfrak{A} \rightarrow [0, \infty]$ ein Maß ist. Auf einem Maßraum nennen wir die \mathfrak{A} -messbaren Mengen auch *μ -messbare Mengen*.

Beispiel. • Ist X ein topologischer Raum, so können wir den *Borelschen messbaren Raum* $(X, \mathcal{B}(X))$ (mit der von den offenen Mengen erzeugten Borelschen σ -Algebra $\mathcal{B}(X)$) definieren.

- Zu jedem $d \geq 1$ gibt es den *Lebesgue-Borelschen Maßraum* $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \lambda^d)$.
- Ein Maßraum $(\Omega, \mathfrak{A}, P)$ wobei das Maß $P(\Omega) = 1$ erfüllt, nennt man *Wahrscheinlichkeitsraum* (und diese werden in der Stochastik studiert).

Definition 8.2. Es sei $(\Omega, \mathfrak{A}, \mu)$ ein Maßraum und $\mu^* : \mathfrak{P}(\Omega) \rightarrow [0, \infty]$ das zugehörige äußere Maß. Eine Menge $N \subset \Omega$ heißt *μ -Nullmenge*, wenn $\mu^*(N) = 0$ gilt.

Nullmengen bezüglich des Lebesgues-Borelschen Maßes nennen wir *Lebesgue-Nullmengen*.

Aus der σ -Subadditivität von äußeren Maßen (Satz 6.4 c)) folgt, dass die Vereinigung von abzählbar vielen Nullmengen wieder eine Nullmenge ist.

Beispiel. 1. Die leere Menge ist per Definition immer eine Nullmenge.

2. Wir betrachten $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda^1)$, also den Lebesgue-Borelschen Maßraum auf der reellen Achse.

Für jede einpunktige Menge $\{a\}$ und jedes $\epsilon > 0$ gilt

$$\lambda^1(\{a\}) \leq \lambda^1([a - \epsilon, a + \epsilon]) = 2\epsilon$$

und es folgt

$$\lambda^1(\{a\}) = 0.$$

Da sowohl \mathbb{N} also \mathbb{Q} abzählbare Vereinigungen von einpunktigen Mengen sind, gilt

$$\lambda^1(\mathbb{N}) = \lambda^1(\mathbb{Q}) = 0.$$

Beachte das zwar \mathbb{Q} dicht in \mathbb{R} liegt, trotzdem aber eine Lebesgue-Nullmenge ist. Das mag paradox klingen, bedeutet aber nur, dass "topologische Größe" nicht das gleiche ist, wie "Größe im Sinne eines Maßes".

3. Betrachten wir nun den Maßraum $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}), \lambda^d)$ und dort *kompakte* Quader

$$R = [a_1, b_1] \times \cdots \times [a_d, b_d].$$

Wir zeigen $\lambda^d(R) = \prod_{i=1}^d (b_i - a_i)$ (auch für den Fall, dass $a_i = b_i$ gilt). Für halboffene Quader ist das per Definition klar. Für kompakte Quader folgt das mit Hilfe der Quader

$$Q_k = [a_1, b_1 + \frac{1}{k}] \times \cdots \times [a_d, b_d + \frac{1}{k}].$$

Es gilt nämlich $R = \bigcap_{k=1}^{\infty} Q_k$ und also gilt

$$\lambda^d(R) = \lim_{k \rightarrow \infty} \lambda^d(Q_k) = \prod_{i=1}^d (b_i - a_i).$$

Insbesondere sind kompakte Quader bei denen eine Seitenlänge 0 ist Nullmengen.

Es folgt: Achsenparallele Hyperebenen

$$H_i(c) := \{x \in \mathbb{R}^d \mid x_i = c\}, \quad (i \in \{1, \dots, d\}, c \in \mathbb{R})$$

sind Nullmengen (als aufsteigende Vereinigung von kompakten Quadern mit einer Seitenlänge 0).

Satz 8.3 (Charakterisierung von Lebesgue-Nullmengen). *Eine Menge $A \subset \mathbb{R}^d$ ist genau dann eine Lebesgue-Nullmenge, wenn es zu jedem $\epsilon > 0$ eine Folge W_k von kompakten Würfeln in \mathbb{R}^d gibt mit*

$$A \subset \bigcup_{k=1}^{\infty} W_k, \quad \sum_{k=1}^{\infty} \lambda^d(W_k) < \epsilon.$$

Beweis.

\Leftarrow : Ist per Definition klar.

\Rightarrow : Sei A eine Lebesgue-Nullmenge. Aus unserer Untersuchung des Lebesgueschen äußeren Maßes aus Abschnitt 5 wissen wir, dass es zu einer Menge A mit $(\lambda^d)^*(A) = 0$ ein offenes $U \subset \mathbb{R}^d$ mit $A \subset U$ und $\lambda^d(U) < \epsilon$ gibt. Nach Lemma 3.5 können wir U als disjunkte Vereinigung von abzählbar vielen (halboffenen) Würfeln Q_k ausdrücken und also gilt $\lambda^d(U) = \sum_k \lambda^d(Q_k) < \epsilon$. Wir setzen also $W_k = \overline{Q_k}$ und da die nun kompakten Würfel W_k das gleiche Lebesgue-Maß wie die Q_k haben, haben wir die Implikation gezeigt.

□

Satz 8.4. *Ist $U \subset \mathbb{R}^d$ offen und $F : U \rightarrow \mathbb{R}^d$ Lipschitz-stetig, so gilt: Ist $A \subset U$ eine Lebesgue-Nullmenge, so ist auch $F(A)$ eine Lebesgue-Nullmenge.*

Beweis.

Da sich jedes offene U als abzählbare Vereinigung von kompakten Quadern schreiben lässt, genügt es, zu zeigen: Für jeden kompakten Würfel $K \subset U$ ist $F(A \cap K)$ eine Nullmenge.

Ist C eine Lipschitz-Konstante von F , so gilt für $x, y \in K$

$$\|F(x) - F(y)\| \leq C\|x - y\|.$$

Ist K ein Würfel mit Seitenlänge a , so ist $F(A \cap K)$ in einem Würfel der Seitenlänge $\sqrt{d}Ca$ enthalten.

Der Durchmesser von W ist $\sqrt{a^2 + \dots + a^2} = \sqrt{da^2} = \sqrt{d}a$ und wird bei Abbildung mit F um höchstens C gestreckt.

Also gilt

$$\lambda^d(F(A \cap K)) \leq (\sqrt{d}Ca)^d = d^{d/2}C^d a^d = d^{d/2}C^d \lambda^d(K).$$

Da A in einer abzählbaren Vereinigung von Würfeln mit beliebig kleinem Gesamtvolumen enthalten ist, folgt die Behauptung. \square

Da sich beliebige Hyperebenen als Drehung von achsenparallelen Hyperebenen ergeben und Drehungen Lipschitz-stetig (mit Konstante 1) sind, folgt, dass beliebige Hyperebenen ebenfalls Lebesgue-Nullmengen sind.

Aus unserer Definition von Nullmengen folgt nicht automatisch, dass Nullmengen selbst messbar sind, aber es gilt folgendes:

Lemma 8.5. *Ist $(\Omega, \mathfrak{A}, \mu)$ ein Maß-Raum und S eine μ -Nullmenge, so gibt es eine μ -messbare μ -Nullmenge A mit $S \subset A$.*

Beweis.

Ist S eine μ -Nullmenge, so gilt $\mu^*(S) = 0$. Also gibt es eine Folge A_k von μ -messbaren Mengen mit $S \subset A_k$ und $\mu(A_k) \searrow \mu^*(S) = 0$. Dann ist $A = \bigcap_k A_k$ wieder μ -messbar mit $S \subset A$ und $\mu(A) = \lim_{k \rightarrow \infty} \mu(A_k) = 0$. \square

Definition 8.6. Ein Maß-Raum $(\Omega, \mathfrak{A}, \mu)$ heißt *vollständig*, wenn jede μ -Nullmenge auch μ -messbar ist.

Nicht-vollständige Maß-Räume kann man vervollständigen, indem man die σ -Algebra wie folgt vergrößert: Wir definieren $\overline{\mathfrak{A}}^\mu$ durch

$$\overline{\mathfrak{A}}^\mu := \{A \cup S \mid A \in \mathfrak{A}, S \subset \Omega, \mu^*(S) = 0\}$$

Dies ergibt tatsächlich eine σ -Algebra und μ lässt sich durch

$$\bar{\mu}(A \cup S) := \mu(A) \text{ für } A \in \mathfrak{A}, \mu^*(S) = 0$$

darauffortsetzen. Dies $\bar{\mu}$ ist ein Maß und der Maß-Raum $(\Omega, \overline{\mathfrak{A}}^\mu, \bar{\mu})$ ist die *Vervollständigung* von $(\Omega, \mathfrak{A}, \mu)$.

Manchmal wird der Begriff “ μ -Nullmenge” auch anders definiert, nämlich indem man nur μ -messbare Mengen N μ -Nullmengen nennt (und zwar dann, wenn $\mu(N) = 0$ gilt). In diesem Fall nennt man ein Maß vollständig, wenn jede Teilmenge einer μ -Nullmenge μ -messbar ist.

D.h., für die Vervollständigung nehmen wir alle μ -messbaren Mengen und vereinigen sie mit allen μ -Nullmengen.

Beispiel. Die Vervollständigung des Lebesgue-Borelschen Maßraumes $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \lambda^d)$ heißt *Lebesguescher Maßraum* $(\mathbb{R}^d, \overline{\mathcal{B}(\mathbb{R}^d)}^{\lambda^d}, \overline{\lambda^d})$. Die Fortsetzung $\overline{\lambda^d}$ bezeichnen wir wieder mit λ^d und nennen es *Lebesgue-Maß*. Die σ -Algebra $\mathcal{L}(\mathbb{R}^d) = \overline{\mathcal{B}(\mathbb{R}^d)}^{\lambda^d}$ heißt *Lebesgue-Algebra* und die Mengen dort heißen *Lebesgue-messbar*.

Die Mengen in $\mathcal{B}(\mathbb{R}^d)$ heißen auch *Borel-messbare Mengen*.

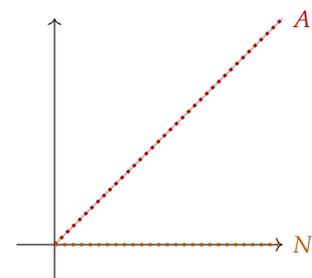
Der Unterschied zwischen dem Lebesgue-Maß und dem Lebesgue-Borel-Maß ist subtil. Da das Lebesgue-Maß vollständig ist, ist es oft einfacher zu benutzen. Allerdings kommt es auch nicht immer zum Einsatz. Wir greifen kurz ein wenig vor:

1. Sind (Ω, \mathfrak{A}) und (Ω', \mathfrak{A}') zwei messbare Räume, so nennt man eine Funktion $f : \Omega \rightarrow \Omega'$ $(\mathfrak{A}, \mathfrak{A}')$ -messbar, wenn für jedes $B \in \mathfrak{A}'$ gilt, dass das Urbild $f^{-1}(B)$ in \mathfrak{A} liegt. Nehmen wir jetzt speziell $\Omega' = \overline{\mathbb{R}}$ und $\Omega = \mathbb{R}^d$. Dann nennt man eine Funktion $f : \mathbb{R}^d \rightarrow \overline{\mathbb{R}}$ *Borel-messbar*, wenn sie $(\mathcal{B}(\overline{\mathbb{R}}), \mathcal{B}(\mathbb{R}^d))$ -messbar ist, also, wenn Urbilder von Borel-Mengen in $\overline{\mathbb{R}}$ immer Borel-messbar in \mathbb{R}^d sind. Der Begriff *Lebesgue-messbar* meint allerdings $(\mathcal{L}(\mathbb{R}^d), \mathcal{B}(\mathbb{R}^d))$ -messbar, das heißt man betrachtet nur Urbilder von Borel-Mengen! Das liegt zum Beispiel daran, dass eine stetig Funktion im Allgemeinen nicht $(\mathcal{L}(\mathbb{R}^d), \mathcal{L}(\mathbb{R}^d))$ -messbar sein muss (aber immer $(\mathcal{L}(\mathbb{R}^d), \mathcal{B}(\mathbb{R}^d))$ -messbar ist).
2. Man könnte glauben, dass für Lebesgue-messbare Mengen $A \subset \mathbb{R}^2$ immer gelten müsste, dass auch deren Projektion $P_1 A = \{x_1 \mid (x_1, x_2) \in A\} \subset \mathbb{R}$ immer Lebesgue-messbar wäre (und auch Henri Lebesgue war dieser Ansicht). Das stimmt allerdings nicht - hier ein einfaches Beispiel: Es sei $N \subset \mathbb{R}$ eine Vitali-Menge, wie sie in Abschnitt 1 konstruiert wurde. Wir betrachten dann in \mathbb{R}^2

$$A = \{(x, x) \mid x \in N\}.$$

Diese Menge ist eine Borel-Nullmenge, da sie in einer Hyperebene enthalten ist, also ist sie Lebesgue-messbar. Die Projektion auf die erste Koordinate ist aber $P_1 A = N$ eine Vitali-Menge und nicht Lebesgue-messbar.

Was richtig ist: Die Projektionen von Borel-messbaren Mengen sind Lebesgue-messbar (was wir aber nicht hier zeigen werden).



9 Messbare Abbildungen

Wir wollen nun Abbildungen auf messbaren Räumen betrachten. Da diese sich zum Integrieren eignen sollen, brauchen wir eine gewisse Verträglichkeit mit den messbaren Strukturen.

Definition 9.1. Es seien (Ω, \mathfrak{A}) und (Ω', \mathfrak{A}') messbare Räume. Eine Abbildung

$$f : \Omega \rightarrow \Omega'$$

heißt $(\mathfrak{A}, \mathfrak{A}')$ -messbar, wenn für jedes $B \in \mathfrak{A}'$ immer $f^{-1}(B) \in \mathfrak{A}$ gilt.

Um zu prüfen, ob eine Abbildung messbar ist, reicht es, Urbilder von Elementen eines Erzeugers zu betrachten:

Lemma 9.2. Ist $\mathfrak{A}' = \sigma(\mathfrak{E}')$, so gilt: f ist $(\mathfrak{A}, \mathfrak{A}')$ -messbar, wenn für $E' \in \mathfrak{E}'$ immer $f^{-1}(E') \in \mathfrak{A}$ folgt.

Beweis.

Es gelte $f^{-1}(\mathfrak{E}') \subset \mathfrak{A}$. Zu zeigen ist $f^{-1}(\mathfrak{A}') \subset \mathfrak{A}$. Wir setzen also

$$\mathfrak{G} := \{A \subset \Omega' \mid f^{-1}(A) \in \sigma(f^{-1}(\mathfrak{E}'))\}.$$

Dieses Mengensystem erfüllt das Prinzip der guten Mengen:

1. Für $E' \in \mathfrak{E}'$ gilt $f^{-1}(E') \in \sigma(f^{-1}(\mathfrak{E}'))$, d.h. $E' \in \mathfrak{G}$.
2. \mathfrak{G} ist eine σ -Algebra, da sich Vereinigungen, Schnitte und Komplemente durch das Urbild ziehen, d.h. für $A, A_i \in \mathfrak{G}$ gilt immer

$$f^{-1}\left(\bigcup A_i\right) = \bigcup f^{-1}(A_i) \in \mathfrak{A},$$

$$f^{-1}\left(\bigcap A_i\right) = \bigcap f^{-1}(A_i) \in \mathfrak{A},$$

$$f^{-1}(A^c) = f^{-1}(A)^c \in \mathfrak{A}.$$

Nach dem Prinzip der guten Mengen folgt $\mathfrak{G} = \mathfrak{A}' = \sigma(\mathfrak{E}')$, d.h. für $A \in \mathfrak{A}'$ gilt $f^{-1}(A) \in \sigma(f^{-1}(\mathfrak{E}')) \subset \mathfrak{A}$, wie behauptet. □

Beispiel. Sind Ω, Ω' topologische Räume, ist jede stetige Abbildung $f : \Omega \rightarrow \Omega'$ Borel-Borel-messbar, d.h. für jede Borel-Menge A in Ω' ist $f^{-1}(A)$ wieder eine Borel-Menge (aber in Ω). Das folgt aus Lemma 9.2, da die offenen Mengen die Borel- σ -Algebra erzeugen.

Um eine reellwertige Funktion zu integrieren, wird es praktisch sein, dabei $\pm\infty$ als Werte zuzulassen. Auch für solche Funktionen wollen wir den Begriff der Messbarkeit definieren.

Definition 9.3. Eine Funktion $f : \Omega \rightarrow \overline{\mathbb{R}}$ nennen wir *numerische Funktion* auf Ω .

Kurz gesagt: Eine Abbildung ist messbar, wenn Urbilder messbarer Mengen wieder messbar sind.

Wenn man mit $f^{-1}(\mathfrak{E}')$ die Menge aller Urbilder von Mengen in \mathfrak{E}' unter f^{-1} bezeichnet, so können wir Messbarkeit einer Funktion auch als $f^{-1}(\mathfrak{A}') \subset \mathfrak{A}$ schreiben.

In obiger Notation heißt " $f^{-1}(E') \in \mathfrak{A}$ für alle $E' \in \mathfrak{E}'$ " also $f^{-1}(\mathfrak{E}') \subset \mathfrak{A}$.

Dieses Lemma beweist man elegant mit Hilfe des *Prinzips der guten Mengen*, was wir zuerst formulieren:

Lemma (Prinzip der guten Mengen). Es sei \mathfrak{A} eine σ -Algebra, die von \mathfrak{E} erzeugt wird und P eine Eigenschaft von Mengen in \mathfrak{A} . Weiterhin sei $\mathfrak{G} \subset \mathfrak{A}$ eine Menge, für die alle Elemente $G \in \mathfrak{G}$ die Eigenschaft P haben. Gilt dann

(a) $\mathfrak{E} \subset \mathfrak{G}$ und

(b) \mathfrak{G} ist eine σ -Algebra

so folgt, dass alle Elemente in \mathfrak{A} ebenfalls die Eigenschaft P haben.

Beweis. Dies gilt, da wegen $\mathfrak{E} \subset \mathfrak{G}$ und $\sigma(\mathfrak{G}) = \mathfrak{G}$ folgt

$$\mathfrak{A} = \sigma(\mathfrak{E}) \subset \sigma(\mathfrak{G}) = \mathfrak{G}. \quad \square$$

Das Mengensystem \mathfrak{G} nennt man dann die "guten Mengen". Das Prinzip in Worten: Enthalten die guten Mengen einen Erzeuger der σ -Algebra und bilden sie selbst eine, so sind sie schon die ganze σ -Algebra.

Wird sowohl im Definitions- als auch im Wertebereich die Borel- σ -Algebra genommen, sagt man auch einfach nur "Borel-messbar".

Auf $\overline{\mathbb{R}}$ definieren wir die Borel- σ -Algebra $\mathcal{B}(\overline{\mathbb{R}})$, als die σ -Algebra, die von den offenen Mengen in \mathbb{R} und von den Mengen $\{\infty\}$ und $\{-\infty\}$ erzeugt wird.

Eine Abbildung $f : \Omega \rightarrow \overline{\mathbb{R}}$ eines messbaren Raumes (Ω, \mathfrak{A}) heißt (\mathfrak{A} -)messbare (numerische) Funktion, wenn sie $(\mathfrak{A}, \mathcal{B}(\overline{\mathbb{R}}))$ -messbar ist.

$\mathcal{B}(\overline{\mathbb{R}})$ enthält also alle Mengen der Form B , $B \cup \{\infty\}$ und $B \cup \{-\infty\}$, $B \cup \{\infty, -\infty\}$ mit $B \in \mathcal{B}(\mathbb{R})$.

Definition 9.4. Für eine numerische Funktion $f : \Omega \rightarrow \overline{\mathbb{R}}$ definieren wir

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

und analog die Mengen $\{f \geq c\}$, $\{f < c\}$, $\{f \leq c\}$, $\{f = c\}$, $\{f \neq c\}$ usw.

Beispiel. Ist $S \subset \Omega$, so ist die charakteristische Funktion von S definiert durch

$$\chi_S(x) := \begin{cases} 1, & \text{falls } x \in S, \\ 0, & \text{sonst.} \end{cases}$$

Es gilt $\{\chi_S \geq 1\} = S = \{\chi_S = 1\}$. Natürlich ist χ_S genau dann messbar, wenn S messbar ist.

Satz 9.5. Es sei (Ω, \mathfrak{A}) ein messbarer Raum und $f : \Omega \rightarrow \overline{\mathbb{R}}$ eine numerische Funktion. Dann ist f genau dann \mathfrak{A} -messbar, wenn eine der folgenden äquivalenten Bedingungen gilt:

- i) $\forall c \in \mathbb{R} : \{f > c\} \in \mathfrak{A}$
- ii) $\forall c \in \mathbb{R} : \{f \geq c\} \in \mathfrak{A}$
- iii) $\forall c \in \mathbb{R} : \{f < c\} \in \mathfrak{A}$
- iv) $\forall c \in \mathbb{R} : \{f \leq c\} \in \mathfrak{A}$

Beweis.

Die Mengensysteme $]c, \infty]$, $[c, \infty]$, $[-\infty, c[$, $[-\infty, c]$ erzeugen alle die gleiche σ -Algebra (vgl. Übungsblatt 2, Aufgabe 4). Und da diese σ -Algebra insbesondere die halboffenen Intervalle $[a, b[$ enthält, handelt es sich dabei um die Borel- σ -Algebra auf $\overline{\mathbb{R}}$. \square

Satz 9.6. Sind $f, g : \Omega \rightarrow \overline{\mathbb{R}}$ \mathfrak{A} -messbare Funktionen auf einem messbaren Raum (Ω, \mathfrak{A}) , so sind die Mengen $\{f > g\}$, $\{f \geq g\}$, $\{f = g\}$ und $\{f \neq g\}$ \mathfrak{A} -messbar.

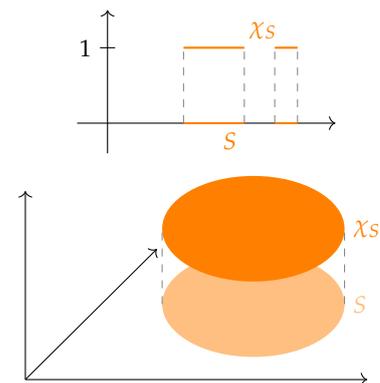
Beweis.

Es gilt

$$f(x) > g(x) \iff \left[\exists r \in \mathbb{Q} : f(x) > r > g(x) \right].$$

Also ist

$$\{f > g\} = \bigcup_{r \in \mathbb{Q}} \left(\underbrace{\{f > r\}}_{\in \mathfrak{A}} \cap \underbrace{\{r > g\}}_{\in \mathfrak{A}} \right) \in \mathfrak{A}.$$



Die restlichen Aussagen folgen aus

$$\{f \geq g\} = \{f < g\}^c, \quad \{f = g\} = \{f \geq g\} \setminus \{f > g\}, \quad \{f \neq g\} = \{f = g\}^c.$$

□

Wir werden in Zukunft statt \mathfrak{A} -messbar nur noch messbar schreiben, wenn die σ -Algebra aus dem Kontext klar ist.

Satz 9.7. Ist (Ω, \mathfrak{A}) ein messbarer Raum und $f, g : \Omega \rightarrow \bar{\mathbb{R}}$ messbar, dann sind auch die folgenden Funktionen messbar:

i) cf für $c \in \mathbb{R}$,

ii) $|f|^p$ für $p > 0$,

iii) $f + g$ (falls überall auf Ω definiert),

iv) fg .

D.h. es gibt kein x in dem f und g beide $\pm\infty$ mit entgegengesetztem Vorzeichen sind.

Beweis.

Die Punkte i) und ii) folgen daraus, dass die Funktionen $x \mapsto cx$ und $x \mapsto |x|^p$ als stetige Funktionen $(\mathcal{B}(\bar{\mathbb{R}}), \mathcal{B}(\bar{\mathbb{R}}))$ -messbar sind.

Zu iii): Ist g messbar, so sind auch $-g$ und $c - g$ messbar, und die Behauptung folgt aus Satz 9.6, da

$$\{f + g > c\} = \{f > c - g\}.$$

Zu iv): Die Menge $\Omega_\infty = \{|f| = \infty\} \cup \{|g| = \infty\}$ ist messbar und $fg|_{\Omega_\infty}$ ist eine messbare Funktion. Wir müssen also nur noch die Messbarkeit von fg auf $\Omega_0 = \Omega \setminus \Omega_\infty$ zeigen. Auf Ω_0 haben aber f und g Werte in \mathbb{R} und es gilt dort

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

Die Behauptung folgt nun mit den Punkten i) bis iii). □

Definition 9.8. Zu numerischen Funktionen $f, g, f_k : \Omega \rightarrow \bar{\mathbb{R}}$, $k = 1, 2, \dots$, definieren wir

$$\begin{aligned} \max(f, g)(x) &:= \max(f(x), g(x)), & \min(f, g)(x) &:= \min(f(x), g(x)), \\ (\sup_k f_k)(x) &:= \sup_k \{f_k(x)\}, & (\inf_k f_k)(x) &:= \inf_k \{f_k(x)\}, \\ (\limsup_k f_k)(x) &:= \limsup_k f_k(x), & (\liminf_k f_k)(x) &:= \liminf_k f_k(x), \\ f_+(x) &:= \max(f(x), 0), & f_-(x) &:= \max(-f(x), 0). \end{aligned}$$

Die Funktionen f_+ und f_- heißen *Positivteil* bzw. *Negativteil* von f und es gilt $f = f_+ - f_-$.

Satz 9.9. Ist (Ω, \mathfrak{A}) ein messbarer Raum und sind $f, g, f_k : \Omega \rightarrow \bar{\mathbb{R}}$ messbar, dann sind auch $\max(f, g)$, $\min(f, g)$, $\sup_k f_k$, $\inf_k f_k$, $\limsup_k f_k$, $\liminf_k f_k$, f_+ und f_- messbar.

Beweis.

Die Fälle $\sup_k f_k$ und $\inf_k f_k$ folgen aus

$$\{\sup_k f_k \leq c\} = \{x \mid \forall k : f_k(x) \leq c\} = \bigcap_k \{f_k \leq c\}$$

$$\{\inf_k f_k \geq c\} = \{x \mid \forall k : f_k(x) \geq c\} = \bigcap_k \{f_k \geq c\}$$

Die weiteren Fälle folgen daraus (die Fälle mit \limsup und \liminf folgen z.B. wegen $\limsup_k f_k = \inf_m (\sup_{k \geq m} f_k)$ und $\liminf_k f_k = \sup_m (\inf_{k \geq m} f_k)$).

□

Definition 9.10. Ist (Ω, \mathfrak{A}) ein messbarer Raum, so heißt messbare Funktion $f : \Omega \rightarrow \mathbb{R}$ *einfache Funktion*, wenn f nur endlich viele Werte annimmt.

Nennen wir die endlich vielen Werte c_1, \dots, c_k , so bilden die Mengen $A_i = f^{-1}(c_i)$, $i = 1, \dots, k$ eine (messbare) Partition von Ω und es gilt

$$f = \sum_{i=1}^k c_i \chi_{A_i}.$$

Umgekehrt gilt: jede Linearkombination von charakteristischen Funktionen von messbaren Mengen ist eine einfache Funktion, d.h. jede einfache Funktion ist von obiger Form.

Bemerkenswerterweise kann man *jede* nicht-negative messbare numerische Funktion punktweise durch eine aufsteigende Folge von einfachen Funktionen approximieren:

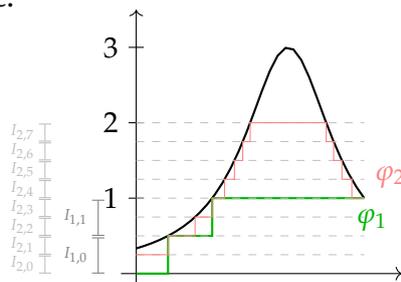
Satz 9.11. Es sei (Ω, \mathfrak{A}) ein messbarer Raum und $f : \Omega \rightarrow [0, \infty]$ messbar. Dann gibt es eine Folge $0 \leq \varphi_1 \leq \varphi_2 \leq \dots$ von einfachen Funktionen $\varphi_k : \Omega \rightarrow \mathbb{R}$ mit $\varphi_k \xrightarrow{k \rightarrow \infty} f$ p.w.

Beweis.

Zu $k = 1, 2, \dots$ und $m = 0, 1, 2, \dots, k2^k - 1$ definieren wir die Intervalle $I_{m,k} =]\frac{m}{2^k}, \frac{m+1}{2^k}]$ und damit

$$\varphi_k(x) = \begin{cases} \frac{m}{2^k}, & \text{falls } f(x) \in I_{m,k} \\ k, & \text{falls } f(x) > k. \end{cases}$$

Hier eine Skizze:



Erinnere: $\varphi_k \xrightarrow{k \rightarrow \infty} f$ p.w. bedeutet, dass für jeden Punkt x gilt $\varphi_k(x) \xrightarrow{k \rightarrow \infty} f(x)$.

Aus der Konstruktion sieht man $\varphi_k \leq \varphi_{k+1}$ und

$$\begin{aligned} f(x) < k &\implies 0 < f(x) - \varphi_k(x) \leq 2^{-k} \\ f(x) > k &\implies \varphi_k(x) = k. \end{aligned}$$

Daraus folgt $\varphi_k(x) \xrightarrow{k \rightarrow \infty} f(x)$ für alle x .

□

Als Fazit dieser Überlegungen kann man festhalten:

- Für numerische Funktionen auf einem topologischen Raum gilt: Alle Funktionen, die man als Summen, Produkte, punktweise Suprema, Infima oder Grenzwerte aus stetigen Funktionen erhält, sind Borel-messbar.
- Für numerische Funktionen auf beliebigen messbaren Räumen gilt: Alle Funktionen, die man als Summen, Produkte, punktweisen Suprema, Infima oder punktweise Grenzwerten aus einfachen Funktionen erhält, sind messbar.

Damit scheint es ziemlich schwierig, überhaupt einmal eine nicht messbare Funktion zu konstruieren (und in der Tat geht das im Allgemeinen auch nicht ohne die Hilfe des Auswahlaxioms, also z.B. mit Hilfe einer charakteristischen Funktion einer nicht messbaren Menge).

10 Das Integral bezüglich eines Maßes

In diesem Abschnitt sei $(\Omega, \mathfrak{A}, \mu)$ immer ein Maßraum.

Wir wollen nun ein Integral einer Funktion $f : \Omega \rightarrow \mathbb{R}$ (oder auch $\rightarrow \overline{\mathbb{R}}$) definieren, welches für das Lebesgue-Maß (also für $\Omega = \mathbb{R}^d, \mu = \lambda^d$) die Fläche, bzw. Volumen etc. zwischen Graph und Definitionsbereich wiedergibt. Das Integral einer Funktion f bezüglich eines Maßes μ bezeichnen wir mit

$$\int f d\mu.$$

Im Allgemeinen hätten wir gerne folgende Eigenschaften:

1. Für jede messbare Menge A soll gelten $\int \chi_A d\mu = \mu(A)$.
2. Das Integral ist linear, d.h. $\int (af + g) d\mu = a \int f d\mu + \int g d\mu$.
3. Das Integral ist monoton, d.h. ist $f \leq g$, so ist $\int f d\mu \leq \int g d\mu$.
4. Das Integral verträgt sich mit (geeigneten) Grenzwerten, d.h. für $f_k \rightarrow f$ (in einem noch zu spezifizierenden Sinn) gilt $\int f_k d\mu \rightarrow \int f d\mu$.

Wie schon beim Riemann-Integral, definieren wir das Integral nun in mehreren Schritten für immer allgemeinere Fälle.

Definition 10.1 (Integral von nicht-negativen, einfachen Funktionen). Ist $\varphi : \Omega \rightarrow [0, \infty[$, eine einfache Funktion, d.h. es gilt $\varphi = \sum_{i=1}^m c_i \chi_{A_i}$ mit $c_i \geq 0$ und messbaren A_i , so definieren wir

$$\int \varphi d\mu := \sum_{i=1}^m c_i \mu(A_i).$$

Man beachte, dass das Integral auch den Wert $+\infty$ annehmen kann, und zwar, wenn eine der Menge A_i unendliches Maß hat.

Dies Integral hat zwar nur einen kleinen Definitionsbereich, erfüllt aber immerhin unsere Wunschliste:

Satz 10.2. Es seien $\varphi, \psi, \varphi_k : \Omega \rightarrow [0, \infty[$ nicht-negative einfache Funktionen und sei $c \in [0, \infty[$. Dann gilt:

- i) $\int (c\varphi + \psi) d\mu = c \int \varphi d\mu + \int \psi d\mu$
- ii) $\varphi \leq \psi \implies \int \varphi d\mu \leq \int \psi d\mu$
- iii) Gilt $\varphi_k \nearrow \psi$ p.w., so gilt $\lim_{k \rightarrow \infty} \int \varphi_k d\mu = \int \psi d\mu$.

Beweis.

i) ist einfach die Linearität von Summen und ii) folgt aus i), denn ist $\varphi \leq \psi$, so ist $\psi - \varphi \geq 0$ und daher ist $\int (\psi - \varphi) d\mu \geq 0$.

Für iii) betrachten wir zuerst den Fall $\int \psi d\mu = \infty$: Da ψ selbst eine einfache Funktion ist, muss also ein Summand schon $+\infty$

Soll der Integrationsbereich mit betont werden, schreibt man auch $\int_{\Omega} f d\mu$. Wird die Integrationsvariable benötigt, schreibt man auch $\int f(x) d\mu(x)$.

Das entspricht der Forderung "Volumen ist Grundfläche mal Höhe".

Die Darstellung einer einfachen Funktion in der Form $\varphi = \sum_{i=1}^m c_i \chi_{A_i}$ ist zwar nicht eindeutig, das Integral allerdings unabhängig von der Darstellung.

Dabei bedeutet $\varphi_k \nearrow \psi$ p.w., dass $\varphi_k(x)$ für jedes x wachsend in k ist und $\varphi_k(x) \xrightarrow{k \rightarrow \infty} \psi(x)$.

sein, und daher gibt es $c > 0$ und $A \in \mathfrak{A}$, so dass

$$\psi \geq c\chi_A, \quad \mu(A) = \infty.$$

Da nur punktweise Konvergenz vorliegt, können wir nicht schließen, dass φ auf ganz A irgendwann nahe an c liegt. Wir gehen also anders vor, und definieren

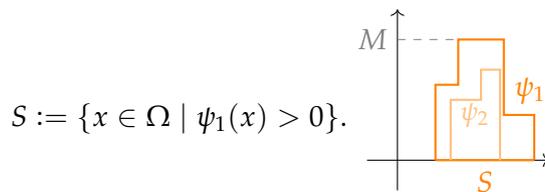
$$A_k := \{x \in A \mid \varphi_k(x) \geq c/2\}$$

(insbesondere ist $A_k \in \mathfrak{A}$, da φ_k einfach ist). Wegen $\varphi_k \nearrow \psi$ folgt $A_k \nearrow A$ und da μ als Maß stetig von unten ist, folgt $\mu(A_k) \rightarrow \infty$. Schließlich gilt nach Definition von A_k auch $\varphi_k \geq \frac{c}{2}\chi_{A_k}$ und es folgt

$$\int \varphi_k d\mu \geq \frac{c}{2}\mu(A_k) \rightarrow \infty.$$

Behandeln wir jetzt den Fall $\int \psi d\mu < \infty$. Wir betrachten $\psi_k = \psi - \varphi_k$. Dann gilt $\psi_k \searrow 0$ und es ist also zu zeigen, dass $\int \psi_k d\mu \xrightarrow{k \rightarrow \infty} 0$.

Dazu führen wir die Größe $M := \sup\{\psi_1(x) \mid x \in \Omega\}$ ein und betrachten die Menge



Da $\int \psi_1 d\mu < \infty$, folgt $\mu(S) < \infty$. Zu $\epsilon > 0$ definieren wir

$$S_k := \{x \in S \mid \psi_k(x) \leq \epsilon\} \subset S$$

und aus $\psi_k \searrow 0$ folgt $\bigcup_k S_k = \{x \in S \mid \exists k : \psi_k(x) \leq \epsilon\} = S$ und das heißt $S_k \nearrow S$. Da μ von unten stetig ist, folgt $\mu(S_k) \rightarrow \mu(S)$ bzw. $\mu(S \setminus S_k) \rightarrow 0$. Außerdem haben wir nach Definition von M und S_k :

$$\psi_k(x) \leq \epsilon\chi_{S_k}(x) + M\chi_{S \setminus S_k}(x).$$

Daher gilt für das Integral

$$\int \psi_k d\mu \leq \epsilon\mu(S_k) + M\mu(S \setminus S_k)$$

und also $\lim_{k \rightarrow \infty} \int \psi_k d\mu \leq \epsilon\mu(S)$. Da wir dies für alle $\epsilon > 0$ zeigen können, folgt $\lim_{k \rightarrow \infty} \int \psi_k d\mu = 0$. \square

Da wir in Satz 9.11 gezeigt haben, dass wir jede nicht-negative messbare Funktion wachsend mit einfachen Funktionen approximieren können und eben gesehen haben, dass sich das Integral

für einfache Funktionen aus Definition 10.1 gut mit wachsenden Grenzwerten verträglich, definieren wir:

Definition 10.3 (Integral für nicht-negative messbare Funktionen). Sei $(\Omega, \mathfrak{A}, \mu)$ ein Maß-Raum und $f : \Omega \rightarrow [0, \infty]$ messbar. Dann gibt es eine aufsteigende Folge $\varphi_k : \Omega \rightarrow [0, \infty[$ von einfachen Funktionen mit $\varphi_k \nearrow f$ p.w. und wir definieren das Integral von f durch

$$\int f d\mu := \lim_{k \rightarrow \infty} \int \varphi_k d\mu.$$

Da die Folge auf der rechten Seite wachsend ist, existiert der Grenzwert auf jeden Fall (ggf. uneigentlich in \mathbb{R}). Es ist aber noch zu zeigen, dass der Grenzwert unabhängig von der Folge φ_k ist:

Sei $\psi_k : \Omega \rightarrow [0, \infty[$ eine weitere Folge einfacher Funktionen mit $\psi_k \nearrow f$. Es reicht zu zeigen, dass dann gelten muss

$$\lim_{k \rightarrow \infty} \int \varphi_k d\mu \geq \lim_{m \rightarrow \infty} \int \psi_m d\mu$$

(denn, wenn wir das zeigen können, können wir durch Vertauschen von φ_k und ψ_k auch die umgekehrte Ungleichung und damit die Gleichheit zeigen). Für $m, k \geq 1$ definieren wir $\varphi_{m,k} := \min(\varphi_k, \psi_m)$. Dann gilt auf jeden Fall $\varphi_{m,k} \leq \varphi_k$. Halten wir m fest, so gilt $\varphi_{m,k} \nearrow_{k=1}^{\infty} \psi_m$ und nach Satz 10.2 ii) und iii) folgt

$$\lim_{k \rightarrow \infty} \int \varphi_k d\mu \geq \lim_{k \rightarrow \infty} \int \varphi_{m,k} d\mu = \int \psi_m d\mu.$$

Gehen wir zum Grenzwert $m \rightarrow \infty$ über, folgt die Behauptung.

Die guten Eigenschaften des Integral aus Satz 10.2 bleiben alle erhalten:

Satz 10.4. Es sei $(\Omega, \mathfrak{A}, \mu)$ ein Maßraum. Dann gilt für messbare Funktionen $f, g, f_k : \Omega \rightarrow [0, \infty]$ und $c \geq 0$:

$$i) \int (cf + g) d\mu = c \int f d\mu + \int g d\mu$$

$$ii) f \leq g \implies \int f d\mu \leq \int g d\mu$$

iii) Ist f_k eine punktweise wachsende Funktionenfolge, so gilt $\lim_{k \rightarrow \infty} \int f_k d\mu = \int \lim_{k \rightarrow \infty} f_k d\mu$.

Beweis.

Wir wählen Folgen von einfachen Funktionen φ_k, ψ_k mit $\varphi_k \nearrow f$ und $\psi_k \nearrow g$ und bekommen mit Hilfe von Satz 10.2:

i) Es gilt $c\varphi_k + \psi_k \nearrow cf + g$ und daher

$$\begin{aligned} \int (cf + g) d\mu &= \lim_{k \rightarrow \infty} \int (c\varphi_k + \psi_k) d\mu = c \lim_{k \rightarrow \infty} \int \varphi_k d\mu + \lim_{k \rightarrow \infty} \int \psi_k d\mu \\ &= c \int f d\mu + \int g d\mu. \end{aligned}$$

ii) Wegen i) reicht es, zu zeigen, dass für $f \geq 0$ auch $\int f d\mu \geq 0$ gilt. Wir können in diesem Fall $\varphi_k \geq 0$ wählen (bzw. $(\varphi_k)_+$ statt φ_k betrachten) und bekommen

$$\int f d\mu = \lim_{k \rightarrow \infty} \int \varphi_k d\mu \geq 0.$$

iii) Für jedes k gibt es eine aufsteigende Folge $\varphi_{k,j}$ von einfachen Funktionen mit $\varphi_{k,j} \nearrow_{j=1}^{\infty} f_k$. Wir setzen $\psi_m = \sup\{\varphi_{k,j} \mid k, j \leq m\}$. Dann gilt $\psi_m \nearrow h := \lim_{k \rightarrow \infty} f_k$ und daher

$$\lim_{m \rightarrow \infty} \int \psi_m d\mu = \int h d\mu.$$

Allerdings gilt ebenfalls $\psi_m \leq f_m \leq h$ und daher

$$\int \psi_m d\mu \leq \int f_m d\mu \leq \int h d\mu$$

und daraus folgt mit dem Sandwich-Lemma $\lim_{m \rightarrow \infty} \int f_m d\mu = \int h d\mu$.

□

Den Fall des Lebesgue-Maßes λ^d werden wir später ausführlicher untersuchen. Hier ein einfacheres Beispiel.

Beispiel. Wir betrachten $\Omega = \mathbb{N} = \{0, 1, 2, \dots\}$ und darauf die σ -Algebra $\mathfrak{A} = \mathfrak{P}(\mathbb{N})$ und das Zählmaß

$$\#A := \begin{cases} \text{Anzahl der Elemente in } A, & \text{falls } A \text{ endlich} \\ \infty, & \text{sonst.} \end{cases}$$

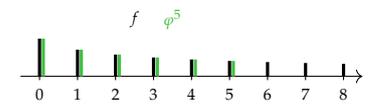
Die Funktionen $f : \Omega \rightarrow \mathbb{R}$ können wir ebenso als reelle Folgen $f_k := f(k)$ auffassen. Da $\mathfrak{A} = \mathfrak{P}(\mathbb{N})$, sind alle Funktionen (bzw. Folgen) $\#$ -messbar. Die einfachen Funktionen sind genau die, bei denen f_k nur endlich viele Werte annehmen kann. In diesem Fall gibt es eine besonders einfache Konstruktion für eine aufsteigende Approximation einer Folge (f_k) mit nicht-negativen Folgengliedern durch einfache Funktionen. Zu einer Folge (f_k) definieren wir eine Folge $(\varphi_k^n)_{k \in \mathbb{N}}$ von Folgen durch

$$\varphi_k^n = \begin{cases} f_k, & \text{für } k \leq n \\ 0, & \text{sonst,} \end{cases}$$

d.h. (φ_k^n) schneidet die Folge (f_k) ab dem n -ten Folgenglied ab und setzt alles danach auf 0. Wir schreiben φ^n als Summe von charakteristischen Funktionen

$$\varphi_k^n = \sum_{i=0}^n \varphi_i^n \chi_{\{i\}}(k)$$

wobei $\chi_{\{i\}}(k)$ genau für $k = i$ den Wert 1 und sonst den Wert 0 annimmt (wir könnten also auch $\chi_{\{i\}}(k) = \delta_{k,i}$ schreiben). Damit



In Aussage iii) hätten wir auch auf die Voraussetzung, dass f messbar ist verzichten können, das dies nach Satz 9.9 aus der Messbarkeit der f_k folgt.

ist das Integral von φ^n nach Definition 10.1

$$\int \varphi^n d\# = \sum_{i=0}^n \underbrace{\varphi_i^n}_{=f_i} \underbrace{\#(\{i\})}_{=1} = \sum_{i=0}^n f_i.$$

Damit bekommen wir das Integral von f bezüglich $\#$ als

$$\int f d\# = \lim_{n \rightarrow \infty} \int \varphi^n d\# = \lim_{n \rightarrow \infty} \sum_{i=0}^n f_i = \sum_{i=0}^{\infty} f_i.$$

Wir halten fest: Das Integral einer (nicht-negativen) Folge bezüglich des Zählmaßes ist genau die Reihe über die Folgenglieder.

Damit können wir Reihen als Spezialfälle von Integralen betrachten.

11 Integrierbarkeit

Wir haben nun geklärt, wie das Integral für nicht-negative messbare Funktionen definiert ist. Untersuchen wir nun den Fall, wenn die Funktion sowohl positive als auch negative Werte annehmen kann. Hier nimmt man die einfachste Idee: Wir teilen f auf in

$$f = f_+ - f_-$$

und definieren

$$\int f d\mu = \int f_+ d\mu - \int f_- d\mu.$$

Doch Achtung! Das geht nicht immer, denn wir haben ja explizit den Wert $+\infty$ für Integrale zugelassen und daher kann hier die Situation $\infty - \infty$ entstehen. Dieser Fall wird einfach per Definition ausgeschlossen.

Definition 11.1. Ist $(\Omega, \mathfrak{A}, \mu)$ ein Maßraum, so nennen wir $f : \Omega \rightarrow \mathbb{R}$ μ -integrierbar, wenn f μ -messbar ist und es gilt $\int f_+ d\mu < \infty$ und $\int f_- d\mu < \infty$ und dann ist

$$\int f d\mu := \int f_+ d\mu - \int f_- d\mu.$$

Satz 11.2. Sei $(\Omega, \mathfrak{A}, \mu)$ ein Maßraum und $f : \Omega \rightarrow \mathbb{R}$ μ -messbar.

- i) f ist integrierbar genau dann, wenn $\int |f| d\mu < \infty$.
- ii) Gibt es zwei nicht-negative μ -messbare Funktionen $f_1, f_2 : \Omega \rightarrow [0, \infty]$ mit $f = f_1 - f_2$ und $\int f_1 d\mu, \int f_2 d\mu < \infty$, so ist f integrierbar mit

$$\int f d\mu = \int f_1 d\mu - \int f_2 d\mu.$$

Beweis.

- i) Es gilt $|f| = f_+ + f_-$ und $f_+, f_- \leq |f|$. Also: Ist f integrierbar, so sind f_+ und f_- integrierbar und da $f_+, f_- \geq 0$ gilt, folgt aus Satz 10.4 i) $\int |f| d\mu = \int f_+ d\mu + \int f_- d\mu < \infty$. Ist umgekehrt $\int |f| d\mu < \infty$, so sind nach Monotonie für nicht negative Funktionen (Satz 10.4 ii)) auch f_+ und f_- integrierbar was gerade die Definition der Integrierbarkeit von f ist.
- ii) In diesem Fall gilt $|f| \leq f_1 + f_2$ und daraus folgt $\int |f| d\mu < \infty$. Nach Teil i) folgt die Integrierbarkeit von f . Außerdem ist $f_1 \geq f_+$ und daher ist wegen $f = f_1 - f_2 = f_+ - f_-$ die Funktion $g := f_1 - f_+ = f_2 - f_-$ nicht-negativ mit $\int g d\mu < \infty$.

Wir haben in Abschnitt 10 das Integral bezüglich μ für alle nicht-negativen, μ -messbaren Funktionen erklärt, nennen f aber nur dann μ -integrierbar, wenn $\int f d\mu < \infty$ gilt. Ist das Maß aus dem Kontext klar, sagen wir auch einfach "integrierbar" statt " μ -integrierbar".

Der Punkt ii) ist ein technisches Hilfsmittel, was wir im nächsten Satz einsetzen werden. Hier muss übrigens gelten, dass f_1 und f_2 nicht gleichzeitig den Wert $+\infty$ annehmen, da sonst f nicht definiert ist.

∞. Mit Satz 10.4 ii) folgt

$$\begin{aligned}\int f_1 d\mu - \int f_2 d\mu &= \int (f_+ + g) d\mu - \int (f_- + g) d\mu \\ &= \int f_+ d\mu - \int f_- d\mu = \int f d\mu.\end{aligned}$$

□

Satz 11.3 (und Definition des Lebesgue-Raumes $\mathcal{L}^1(\Omega, \mu)$). i) Die Menge

$$\mathcal{L}^1(\Omega, \mu) := \{f : \Omega \rightarrow \bar{\mathbb{R}} \mid f \text{ } \mu\text{-integrierbar}\}$$

ist ein Vektorraum..

ii) Für $f, g \in \mathcal{L}^1(\Omega, \mu)$ gilt:

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

iii) Für $f \in \mathcal{L}^1(\Omega, \mu)$ gilt

$$\left| \int f d\mu \right| \leq \int |f| d\mu.$$

Bisher hatten wir Linearität und Monotonie nur für nicht-negative Funktionen gezeigt.

Manchmal schreibt man auch $\mathcal{L}^1(\Omega, \mathfrak{A}, \mu)$ oder $\mathcal{L}^1(\mu)$.

Beweis.

i) Ist $c = 0$ so gilt immer $0 = \int c f d\mu = c \int f d\mu$. Für $c > 0$ gilt $(cf)_+ = c(f_+)$ und $(cf)_- = c(f_-)$ und für $c < 0$ gilt $(cf)_+ = -c(f_-)$ und $(cf)_- = -c(f_+)$. Also gilt für $c > 0$ mit Satz 10.4 i)

$$\begin{aligned}\int (cf) d\mu &= \int (cf)_+ d\mu + \int (cf)_- d\mu = \int c(f_+) d\mu + \int c(f_-) d\mu \\ &= c \left(\int f_+ d\mu + \int f_- d\mu \right) = c \int f d\mu\end{aligned}$$

und für $c < 0$ analog.

Ist $F = f + g$, so ist $F = (f_+ + g_+) - (f_- + g_-)$ (allerding muss nicht $F_+ = f_+ + g_+$ gelten) und nach Satz 11.2 ii) ist F damit integrierbar und es gilt

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

ii) Ist $f \leq g$, so ist $f_+ \leq g_+$ und $f_- \geq g_-$. Es folgt

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

iii) Aus $f \leq |f|$ und $-f \leq |f|$ folgen

$$\int f d\mu \leq \int |f| d\mu \quad \text{und} \quad -\int f d\mu \leq \int |f| d\mu,$$

was die Behauptung zeigt. □

Satz 11.3 i)+ii) sagen, dass die Abbildung

$$I : \mathcal{L}^1(\Omega, \mu) \rightarrow \mathbb{R}, \quad f \mapsto \int f d\mu$$

eine lineare und monotone Abbildung ist.

Komplexwertige Funktionen $f : \Omega \rightarrow \mathbb{C}$ zerlegt man in Real- und Imaginärteil $f = f_1 + if_2$ und wir nennen f μ -integrierbar, wenn f_1 und f_2 integrierbar sind und definieren

$$\int f d\mu := \int f_1 d\mu + i \int f_2 d\mu.$$

Ist $\Omega \subset \mathbb{R}^d$, versehen mit dem Lebesgue-Borel-Maß, so schreiben wir nur $\mathcal{L}^1(\Omega)$.

Wir bemerken, dass die Abbildung $f \mapsto \int |f| d\mu$ positiv homogen ist, und die Dreiecksungleichung erfüllt. Sie ist aber keine Norm auf $\mathcal{L}^1(\Omega, \mu)$, da sie nicht positiv definit ist: Ist eine Funktion f nur auf einer Nullmenge ungleich Null, so ist ihr Integral trotzdem Null, was wir im folgenden Satz präziser machen.

Satz 11.4. Sei $(\Omega, \mathfrak{A}, \mu)$ ein Maßraum.

i) Ist $f : \Omega \rightarrow \overline{\mathbb{R}}$ μ -messbar, so gilt

$$\int |f| d\mu = 0 \iff \{x \in \Omega \mid f(x) \neq 0\} \text{ ist } \mu\text{-Nullmenge.}$$

ii) Ist $f : \Omega \rightarrow \overline{\mathbb{R}}$ μ -integrierbar, so ist

$$\{x \in \Omega \mid f(x) = \pm\infty\}$$

eine μ -Nullmenge.

iii) Es seien $f, g : \Omega \rightarrow \overline{\mathbb{R}}$ μ -messbar und sei $\{x \in \Omega \mid f(x) \neq g(x)\}$ eine μ -Nullmenge. So gilt: f ist genau dann μ -integrierbar, wenn g μ -integrierbar ist und es gilt

$$\int f d\mu = \int g d\mu.$$

Beweis.

i) Die μ -Messbarkeit der Menge $S = \{x \in \Omega \mid f(x) \neq 0\}$ folgt aus der μ -Messbarkeit von f .

Sei zuerst $\int |f| d\mu = 0$. Zu $k = 1, 2, \dots$ setzen wir

$$\varphi_k := \min(k|f|, \chi_S).$$

Dann gilt $\varphi_k \nearrow \chi_S$ und $\varphi_k \leq k|f|$. Also ist $\int \varphi_k d\mu \leq k \int |f| d\mu = 0$ und es folgt nach Satz 10.4 iv) $0 = \lim_{k \rightarrow \infty} \int \varphi_k d\mu = \int \chi_S d\mu = \mu(S)$.

Sei umgekehrt S eine μ -Nullmenge. Wir definieren

$$\psi_k = \min(k\chi_S, |f|)$$

und bemerken $\psi_k \nearrow |f|$. Da $\psi_k \leq k\chi_S$ gilt, folgt $\int \psi_k = 0$ und wieder mit Satz 10.4 iv) folgt $\int |f| d\mu = 0$.

- ii) Ist f integrierbar, so gilt $M := \int |f| d\mu < \infty$. Wir setzen $A = \{x \in \Omega \mid f(x) = \pm\infty\}$ und bemerken, dass für alle k gilt $k\chi_A \leq |f|$. Also ist $\int \chi_A d\mu \leq \frac{1}{k} \int |f| d\mu = M/k$. Da dies für alle k gilt, folgt $\mu(A) = \int \chi_A d\mu = 0$.
- iii) Wir setzen $S = \{x \in \Omega \mid f(x) \neq g(x)\}$ und definieren $h = f\chi_{S^c}$ (beachte, dass $h = g\chi_{S^c}$). Mit $f_1 = f\chi_S$ folgt aus i) $\int |f_1| d\mu = 0$. Da $f = h + f_1$ gilt, ist f genau dann μ -integrierbar, wenn h μ -integrierbar ist. Da diese Argumentation ebenso mit g statt f geht, folgt: g ist μ -integrierbar genau dann, wenn h μ -integrierbar ist.

□

Kommen wir schließlich noch zur sehr praktischen sprechweise “ μ -fast überall”:

Definition 11.5. Ist $(\Omega, \mathfrak{A}, \mu)$ ein Maßraum und P eine Aussage über die Punkte $x \in \Omega$, d.h. wir können P als Abbildung $P : \Omega \rightarrow \{\text{wahr, falsch}\}$ auffassen. Dann sagen wir:

Die Aussage P gilt μ -fast überall, falls die Menge $\{x \in \Omega \mid P(x) \text{ ist falsch}\}$ eine μ -Nullmenge ist.

Ist μ das Lebesgue-Maß λ^d sagen wir *Lebesgue-fast überall*.

Beispiel. Damit können wir die Aussagen von Satz 11.4 wie folgt ausdrücken:

- Es gilt $\int |f| d\mu = 0$ genau dann, wenn $f = 0$ μ -fast überall.
- Ist f μ -integrierbar, so ist $|f| < \infty$ μ -fast überall.
- Gilt $f = g$ μ -fast überall, so gilt $\int f d\mu = \int g d\mu$.

Definition 11.6. Ist $(\Omega, \mathfrak{A}, \mu)$ ein Maßraum und $Z \subset \Omega$ μ -messbar, so heißt $f : \Omega \rightarrow \overline{\mathbb{R}}$ über Z μ -integrierbar, wenn $\chi_Z f$ μ -integrierbar ist und wir definieren

$$\int_Z f d\mu = \int_{\Omega} \chi_Z f d\mu.$$

Ein wichtige Folge von Satz 11.4: Wir können ohne Beschränkung der Allgemeinheit annehmen, dass eine μ -integrierbare Funktion f μ -fast überall endlich ist, denn da $N = \{f = \pm\infty\}$ eine μ -Nullmenge ist, können wir statt f auch (zum Beispiel) $\tilde{f}(x) = \chi_{N^c}(x)f(x)$ betrachten (d.h. wir setzen \tilde{f} auf N einfach gleich Null). Dann ist $\tilde{f} = f$ μ -fast überall (und hat insbesondere das gleiche Integral).

Haben wir einen Maßraum $(\Omega, \mathfrak{A}, \mu)$ und eine nicht-negative μ -messbare Funktion $\varphi : \Omega \rightarrow [0, \infty]$, so können wir für μ -messbare Mengen A definieren

$$\nu(A) := \int_A \varphi d\mu.$$

Satz 11.7 (und Definition von Maßen mit Dichten). *Die eben definierte Abbildung $\nu : \mathfrak{A} \rightarrow [0, \infty]$ ist ein Maß auf \mathfrak{A} . Man nennt ν das Maß mit Dichte φ zum Maß μ und schreibt $\nu = \varphi \cdot \mu$.*

Beweis.

Zuerst sehen wir $\nu(\emptyset) = \int_{\emptyset} \varphi d\mu = 0$.

Sind nun A und B μ -messbar und disjunkt, so gilt $\chi_{A \cup B} = \chi_A + \chi_B$ und daher

$$\begin{aligned} \nu(A \cup B) &= \int_{A \cup B} \varphi d\mu = \int \varphi \chi_{A \cup B} d\mu = \int (\varphi \chi_A + \varphi \chi_B) d\mu \\ &= \int_A \varphi d\mu + \int_B \varphi d\mu = \nu(A) + \nu(B). \end{aligned}$$

Um nun die σ -Additivität zu zeigen, reicht es nach Satz 4.4, zu zeigen, dass ν von unten stetig ist. Sei dazu $A_k \in \mathfrak{A}$ mit $A_k \nearrow A$. Dann folgt für die charakteristischen Funktionen $\chi_{A_k} \nearrow \chi_A$, also auch $\chi_{A_k} \varphi \nearrow \chi_A \varphi$, und nach Satz 10.4 folgt

$$\begin{aligned} \lim_{k \rightarrow \infty} \nu(A_k) &= \lim_{k \rightarrow \infty} \int_{A_k} \varphi d\mu = \lim_{k \rightarrow \infty} \int_{\Omega} \chi_{A_k} \varphi d\mu \\ &= \int_{\Omega} \chi_A \varphi d\mu = \int_A \varphi d\mu = \nu(A). \end{aligned}$$

□

Wir hätten hier auch äquivalent mit der von Z induzierten σ -Algebra $(Z, \mathfrak{A} \cap Z, \mu|_{\mathfrak{A} \cap Z})$ arbeiten können.

12 Konvergenzsätze

Das Vertauschen von Integral und Grenzwert bei punktweiser Konvergenz der Funktionenfolge liefert nicht immer das gleiche Ergebnis.

Beispiel. 1. Wir betrachten $\Omega = \mathbb{R}$ mit dem Lebesgue-Maß λ und die Folge

$$f_k(x) = \chi_{[k,k+1]}(x) = \begin{cases} 1, & \text{falls } k \leq x \leq k+1 \\ 0, & \text{sonst.} \end{cases}$$

Für die Folge gilt $f_k \rightarrow 0$ punktweise denn für einen festen Punkt $x \in \mathbb{R}$ gilt immer $f_k(x) = 0$, sobald $k > x$. Es gilt

$$\int_{\mathbb{R}} f_k d\lambda = \lambda([k, k+1]) = 1.$$

Da die Grenzfunktion konstant Null ist, gilt aber

$$\lim_{k \rightarrow \infty} \int_{\mathbb{R}} f_k d\lambda = 1 \neq 0 = \int_{\mathbb{R}} \lim_{k \rightarrow \infty} f_k d\lambda.$$

Anschaulich gesprochen: Die Masse verschwindet horizontal im Unendlichen.

2. Wir betrachten wieder $\Omega = \mathbb{R}$ mit dem Lebesgue-Maß und

$$f_k(x) = k\chi_{]0,1/k[}(x) = \begin{cases} k, & \text{falls } 0 < x < 1/k \\ 0, & \text{sonst.} \end{cases}$$

Hier gilt

$$\int_{\mathbb{R}} f_k d\lambda = k \frac{1}{k} = 1$$

aber wieder ist $\lim_{k \rightarrow \infty} f_k(x) = 0$ für alle x und wir haben wieder

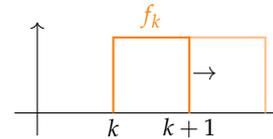
$$\lim_{k \rightarrow \infty} \int_{\mathbb{R}} f_k d\lambda = 1 \neq 0 = \int_{\mathbb{R}} \lim_{k \rightarrow \infty} f_k d\lambda.$$

Anschaulich gesprochen: Die Masse verschwindet vertikal im Unendlichen.

Zeigen wir jetzt, dass sich unser neuer Integralbegriff in einigen Fällen mit punktweise Grenzwerten verträgt. Die erste wichtige Erkenntnis ist, dass im Fall von (punktweise) monotonen Folgen von Funktionen die Vertauschbarkeit von Grenzwert und Integral funktioniert.

Satz 12.1 (Satz von der monotonen Konvergenz). *Ist $(\Omega, \mathfrak{A}, \mu)$ ein Maßraum und sind $f_k : \Omega \rightarrow \mathbb{R}$ μ -integrierbar mit $f_1 \leq f_2 \leq \dots$, so dass es ein M gibt, so dass für alle k gilt*

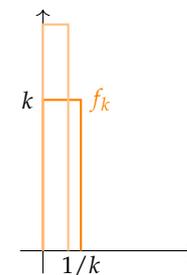
$$\int f_k d\mu \leq M.$$



Das gleiche passiert im Fall von $\Omega = \mathbb{N}$ mit dem Zählmaß $\#$ und der Folgenfolge $a^n = (a_k^n)_{k \in \mathbb{N}}$ mit

$$a_k^n = \delta_{kn} = \begin{cases} 1, & \text{falls } k = n \\ 0, & \text{sonst,} \end{cases}$$

d.h. $a^1 = (1, 0, 0, \dots)$, $a^2 = (0, 1, 0, \dots)$, $a^3 = (0, 0, 1, 0, \dots)$, ...
Es gilt $\int_{\mathbb{N}} a^n d\# = \sum_{k \in \mathbb{N}} a_k^n = 1$ aber es gilt $a_k^n \xrightarrow{n \rightarrow \infty} 0 = a_k$. Daher ist $\lim_{n \rightarrow \infty} \sum_{k \in \mathbb{N}} a_k^n = 1 \neq 0 = \sum_{k \in \mathbb{N}} \lim_{n \rightarrow \infty} a_k^n$.



Dann ist $f := \lim_{k \rightarrow \infty} f_k$ (punktweise erklärt) ebenfalls μ -integrierbar und es gilt

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

Beweis.

Wir definieren $g_k = f_k - f_1$. Dann ist g_k wieder monoton wachsend aber die g_k sind nicht negative Funktionen mit $g_k \rightarrow f - f_1$ und

$$\int g_k d\mu = \int f_k d\mu - \int f_1 d\mu \leq M - \int f_1 d\mu =: M_1.$$

Nach Satz 9.9 folgt, dass $g = \lim_{k \rightarrow \infty} g_k = f - f_1$ (wiederum punktweise erklärt) μ -messbar ist und nach Satz 10.4 iv) folgt, da $g, g_k \geq 0$ gilt,

$$\int g d\mu = \lim_{k \rightarrow \infty} \int g_k d\mu \leq M_1.$$

Also ist g integrierbar und addieren wir $\int f_1 d\mu$ auf beiden Seiten, folgt die Behauptung. \square

Korollar 12.2 (Lemma von Fatou). Für eine Funktionenfolge $f_k : \Omega \rightarrow [0, \infty]$ gilt immer

$$\int \liminf_{k \rightarrow \infty} f_k d\mu \leq \liminf_{k \rightarrow \infty} \int f_k d\mu.$$

Wir betrachten $g_k = \inf_{l \geq k} f_l$, denn diese Folge ist wachsend und konvergiert punktweise gegen $\liminf_{k \rightarrow \infty} f_k$. Die Monotonie des Integrals gibt $\int g_k d\mu = \int (\inf_{l \geq k} f_l) d\mu \leq \int f_k d\mu$ und nach Satz 12.1 (monotone Konvergenz) folgt

$$\int \liminf_{k \rightarrow \infty} f_k d\mu = \int \lim_{k \rightarrow \infty} g_k d\mu = \lim_{k \rightarrow \infty} \int g_k d\mu \leq \liminf_{k \rightarrow \infty} \int f_k d\mu.$$

Vergleichen wir nun das Riemann-Integral mit dem Lebesgue-Integral, d.h. wir betrachten speziell die Menge $\Omega = [a, b]$ und den Maßraum $([a, b], \mathcal{B}([a, b]), \lambda^1)$, also die σ -Algebra der Borel-Mengen im Intervall $[a, b]$ versehen mit dem eindimensionalen Lebesgue-Maß.

Beispiel. Es gibt Lebesgue-integrierbare Funktionen, die nicht Riemann-integrierbar sind: Das einfachste Beispiel ist die Dirichlet'sche Sprungfunktion

$$g(x) = \chi_{\mathbb{Q} \cap [a, b]}(x) = \begin{cases} 1, & \text{falls } x \text{ rational} \\ 0, & \text{sonst.} \end{cases}$$

Dass die Funktion nicht Riemann-integrierbar ist, haben wir schon im ersten Semester gesehen, das Lebesgue-Integral hingegen existiert: Die Menge \mathbb{Q} ist abzählbar, daher eine abzählbare Vereinigung von einpunktigen Mengen, also insbesondere Borel-messbar

Analoges gilt für monoton fallende Funktionenfolgen.

Unsere Beispiele zeigen, dass die Ungleichung tatsächlich strikt sein kann.

mit Lebesgue-Maß 0. Daher ist

$$\int g d\lambda = \int \chi_{\mathbb{Q} \cap [a,b]} d\lambda = \lambda^1(\mathbb{Q} \cap [a, b]) = 0.$$

Andersherum gilt: Jede Riemann-integrierbare Funktion auf einem kompakten Intervall $[a, b]$ ist auch Lebesgue-integrierbar (mit gleichem Wert des Integrals):

Satz 12.3. *Ist $f : [a, b] \rightarrow \mathbb{R}$ Riemann integrierbar, so ist f , ggf. nach Abänderung auf einer Lebesgue-Nullmenge, Borel-messbar und es gilt*

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

Beweis.

Um dies zu zeigen, erinnern wir uns an folgende Charakterisierung der Riemann-Integrierbarkeit: $f : [a, b] \rightarrow \mathbb{R}$ ist genau dann Riemann-integrierbar, wenn es zu jedem $\epsilon > 0$ zwei Treppenfunktionen φ, ψ gibt mit $\varphi \leq f \leq \psi$ und

$$\int_a^b \psi(x) dx - \int_a^b \varphi(x) dx < \epsilon.$$

Treppenfunktionen sind (als endliche Summe von Indikator-Funktionen von Intervallen) Borel-messbar und es gilt

$$\int_a^b \varphi(x) dx = \int_{[a,b]} \varphi d\lambda^1,$$

d.h., das Lebesgue-Integral stimmt mit dem Riemann-Integral überein.

Es gibt also zu jedem k zwei Treppenfunktionen φ_k, ψ_k mit $\varphi_k \leq f \leq \psi_k$ und

$$\int_a^b \psi_k(x) dx - \int_a^b \varphi_k(x) dx < \frac{1}{k}.$$

Ohne Einschränkung können wir annehmen, dass die Folge φ_k wachsend und die Folge ψ_k fallend ist (ansonsten betrachten wir die Folgen $\tilde{\varphi}_k = \max(\varphi_1, \dots, \varphi_k)$ bzw. $\tilde{\psi}_k = \min(\psi_1, \dots, \psi_k)$). Die Definition des Riemann-Integrals liefert

$$\int_a^b f(x) dx = \lim_{k \rightarrow \infty} \int_a^b \varphi_k(x) dx = \lim_{k \rightarrow \infty} \int_a^b \psi_k(x) dx.$$

Um das Lebesgue-Integral von f zu berechnen, benutzen wir den Satz von der monotonen Konvergenz (Satz 12.1): Dieser zeigt, dass φ_k punktweise gegen eine Borel-messbare Funktion φ konvergiert und für diese gilt

$$\int_{[a,b]} \varphi d\lambda = \lim_{k \rightarrow \infty} \int_{[a,b]} \varphi_k d\lambda = \lim_{k \rightarrow \infty} \int_a^b \varphi_k(x) dx.$$

Eine wichtige Konsequenz: Wollen wir Lebesgue-Integrale von Funktionen auf kompakten Intervallen berechnen, so können wir die Techniken benutzen, die wir für Riemann-Integrale kennen.

Ebenso folgt aus dem Satz von der monotonen Konvergenz, dass ψ_k punktweise gegen eine Borel-messbare Funktion ψ konvergiert für die gilt

$$\int_{[a,b]} \psi d\lambda = \lim_{k \rightarrow \infty} \int_{[a,b]} \psi_k d\lambda = \lim_{k \rightarrow \infty} \int_a^b \psi_k(x) dx.$$

Es folgt

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

und da $\varphi \leq f \leq \psi$ gilt, folgt $\psi - \varphi \geq 0$. Wegen

$$\int_{[a,b]} (\psi - \varphi) d\lambda = 0$$

ist $\varphi = \psi$ Lebesgue-fast überall und also auch $f = \varphi$ Lebesgue-fast überall. □

Satz 12.4 (Satz von der dominierten Konvergenz). Sei $(\Omega, \mathfrak{A}, \mu)$ ein Maßraum, $f_k : \Omega \rightarrow \mathbb{R}, k \geq 1$ eine Folge μ -integrierbarer Funktionen mit $f_k \rightarrow f$ punktweise μ -fast überall. Da gilt: Gibt es eine μ -integrierbare Funktion $F : \Omega \rightarrow [0, \infty]$ mit $|f_k| \leq F$ für alle k , so ist f (ggf. nach Abänderung auf einer μ -Nullmenge) μ -integrierbar und es gilt sowohl

$$\int f d\mu = \lim_{k \rightarrow \infty} \int f_k d\mu$$

als auch

$$\lim_{k \rightarrow \infty} \int |f - f_k| d\mu \rightarrow 0.$$

Beweis.

Die Menge $N = \{x \in \Omega \mid f_k(x) \not\rightarrow f(x)\}$ ist eine μ -Nullmenge und wenn wir alle Funktionen mit χ_N multiplizieren, können wir auch $f_k(x) \rightarrow f(x)$ für alle $x \in \Omega$ annehmen.

Wir definieren $g_k(x) = \sup\{f_i(x) \mid i \geq k\}$ und $g_{k,n}(x) = \sup\{f_i(x) \mid k \leq i \leq n\}$. Damit gilt

$$f(x) = \limsup_{k \rightarrow \infty} f_k(x) = \inf_k g_k(x),$$

und außerdem ist $g_{k,n} \nearrow_{n=0}^{\infty} g_k$ und $\int g_{k,n} d\mu \leq \int F d\mu =: M < \infty$ und mit dem Satz von der monotonen Konvergenz (Satz 12.1) folgt die Integrierbarkeit von g_k . Die Folge g_k ist fallend und wegen $\int g_k d\mu \geq -\int F d\mu > -\infty$ folgt (wieder mit monotoner Konvergenz), dass f integrierbar ist mit

$$\int f d\mu = \lim_{k \rightarrow \infty} \int g_k d\mu.$$

Mit $h_k(x) = \inf\{f_i(x) \mid i \geq k\}$ gilt $f(x) = \liminf_k f_k(x) =$

Der Satz heißt auch *Konvergenzsatz von Lebesgue* oder *Satz von der majorisierten Konvergenz*.

Die Funktion F ist die "dominierende" oder "majorisierende Funktion". Man kann immer $F := \sup_k |f_k|$ nehmen und prüfen, ob diese Funktion integrierbar ist.

Die Konvergenz " $f_k \rightarrow f$ punktweise μ -fast überall" bedeutet, dass für μ -fast alle $x \in \Omega$ gilt $f_k(x) \rightarrow f(x)$, anders gesagt: die Menge $\{x \in \Omega \mid f_k(x) \not\rightarrow f(x)\}$ ist eine μ -Nullmenge. Wir schreiben abkürzen p.w. μ -f.ü. und im Fall des Lebesguemaßes auch nur p.w.f.ü.

$\sup_k h_k$ und analog zu vorher schließt man

$$\int f d\mu = \lim_{k \rightarrow \infty} \int h_k d\mu.$$

Wegen $h_k \leq f_k \leq g_k$ folgt mit dem Sandwich-Lemma $\int f d\mu = \lim_{k \rightarrow \infty} \int f_k d\mu$.

Weiterhin ist $|f - f_k| \leq g_k - h_k$ und daher

$$\lim_{k \rightarrow \infty} \int |f - f_k| d\mu \leq \lim_{k \rightarrow \infty} (\int g_k d\mu - \int h_k d\mu) = 0.$$

□

13 Anwendungen und Folgerungen der Konvergenzsätze

Neben der Anwendung im Fall von Lebesgue-Integralen, folgt auch eine Aussage für Reihen:

Korollar 13.1. *Es sei $(a_k^n)_{k \in \mathbb{N}}$ eine Folge von reellen Folgen. Gibt es dann eine summierbare Majorante, d.h. eine Folge (A_k) mit $|a_k^n| \leq A_k$ für alle k und n und $\sum_{k \in \mathbb{N}} A_k < \infty$ so gilt: Ist $a_k = \lim_{n \rightarrow \infty} a_k^n$, so ist auch die Folge (a_k) summierbar und es gilt*

$$\lim_{n \rightarrow \infty} \sum_{k=1}^{\infty} a_k^n = \sum_{k=1}^{\infty} a_k.$$

Das ist genau Satz 12.4 im Fall von $\Omega = \{1, 2, \dots\}$ mit dem Zählmaß #.

Entsprechende Analoga für monotone Konvergenz und das Lemma von Fatou gelten ebenfalls.

Beispiel. Wir betrachten noch einmal die Folge $a_n = (1 + \frac{1}{n})^n$ (von der wir seit dem ersten Semester wissen, dass sie gegen e konvergiert). Wir benutzen den binomischen Lehrsatz und bekommen

$$(1 + \frac{1}{n})^n = \sum_{k=0}^n \binom{n}{k} \frac{1}{n^k}.$$

Wir würden nun gerne rechts einfach $n \rightarrow \infty$ schicken, müssen aber überlegen, in wie fern das gerechtfertigt ist. Wir schreiben

$$a_k^n = \begin{cases} \binom{n}{k} \frac{1}{n^k}, & \text{falls } k \leq n \\ 0, & \text{sonst.} \end{cases}$$

Es gilt für alle k und n

$$\begin{aligned} a_k^n &= \binom{n}{k} \frac{1}{n^k} = \frac{n!}{k!(n-k)!} \frac{1}{n^k} = \frac{1}{k!} \frac{n(n-1) \cdots (n-k+1)}{n^k} \\ &\leq \frac{1}{k!} \frac{n^k}{n^k} = \frac{1}{k!} =: A_k. \end{aligned}$$

Also ist $A_k = \frac{1}{k!}$ eine summierbare Majorante. Für fixiertes k gilt außerdem

$$\binom{n}{k} \frac{1}{n^k} = \frac{1}{k!} \underbrace{\frac{n}{n}}_{=1} \underbrace{\frac{n-1}{n}}_{\rightarrow 1} \cdots \underbrace{\frac{n-k+1}{n}}_{\rightarrow 1} \rightarrow \frac{1}{k!} \text{ für } n \rightarrow \infty$$

und es folgt

$$\begin{aligned} \lim_{n \rightarrow \infty} (1 + \frac{1}{n})^n &= \lim_{n \rightarrow \infty} \sum_{k=0}^n \binom{n}{k} \frac{1}{n^k} = \lim_{n \rightarrow \infty} \sum_{k=0}^{\infty} a_k^n \\ &= \sum_{k=0}^{\infty} \lim_{n \rightarrow \infty} a_k^n = \sum_{k=0}^{\infty} \frac{1}{k!} = e. \end{aligned}$$

Der Satz von der dominierten Konvergenz hat weitere praktische Konsequenzen:

Satz 13.2 (Stetige Abhängigkeit von Parametern). *Es sei $(\Omega, \mathfrak{A}, \mu)$ ein Maßraum und $f : \Omega \times \mathbb{R} \rightarrow \mathbb{R}$ eine Funktion, so dass die Abbildung $f_t(x) := f(x, t)$ für alle $t \in \mathbb{R}$ μ -messbar ist und die Abbildung $f_x(t) = f(x, t)$ für μ -fast alle x stetig ist. Dann gilt: Gibt es ein μ -integrierbares $g : \Omega \rightarrow \mathbb{R}$ mit $|f(x, t)| \leq g(x)$ für alle t und μ -fast alle x so ist f_t für alle t μ -integrierbar und die Funktion*

$$F(t) := \int_{\Omega} f(x, t) d\mu(x)$$

ist stetig.

Beweis.

Um Stetigkeit zu zeigen, betrachten wir eine Folge $t_n \rightarrow t$. Die Funktionen $f_n(x) := f(x, t_n)$ erfüllen die Voraussetzungen des Satzes von der dominierten Konvergenz, insbesondere gilt $f_n(x) = f(x, t_n) \rightarrow f(x, t)$ μ -fast überall und daher gilt

$$\lim_{n \rightarrow \infty} F(t_n) = \lim_{n \rightarrow \infty} \int_{\Omega} f_n d\mu = \int_{\Omega} f(x, t) d\mu(x) = F(t).$$

□

Satz 13.3 (Ableiten unter dem Integral). *Es sei $f : \Omega \times \mathbb{R} \rightarrow \mathbb{R}$ so, dass die Abbildung $x \mapsto f(x, t)$ für alle $t \in \mathbb{R}$ μ -messbar ist, für ein $t_0 \in \mathbb{R}$ μ -integrierbar, und außerdem die partielle Ableitung $\frac{\partial f}{\partial t}(x, t)$ für alle (x, t) existiert. Dann gilt: Gibt es ein integrierbares $g : \Omega \rightarrow \mathbb{R}$ mit*

$$\left| \frac{\partial f}{\partial t}(x, t) \right| \leq g(x),$$

so ist die Abbildung $x \mapsto f(x, t)$ für alle t integrierbar und die Funktion

$$F(t) = \int_{\Omega} f(x, t) d\mu(x)$$

ist differenzierbar mit

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

Beweis.

Nach dem Mittelwertsatz gibt für $t, t_0 \in \mathbb{R}$ ein τ zwischen t und t_0 , so dass

$$f(x, t) - f(x, t_0) = \frac{\partial f}{\partial t}(x, \tau)(t - t_0).$$

Damit folgt

$$|f(x, t) - f(x, t_0)| \leq g(x)|t - t_0|$$

und

$$|f(x, t)| \leq |f(x, t_0)| + g(x)|t - t_0|.$$

Damit folgt $\int_{\Omega} |f(x, t)| d\mu(x) \leq \int_{\Omega} |f(x, t_0)| d\mu(x) + |t - t_0| \int_{\Omega} g(x) d\mu(x)$ und damit die Integrierbarkeit der Funktionen $x \mapsto f(x, t)$ für alle t .

Für die Differenzierbarkeit von F betrachten wir $t_n \rightarrow t$ und den Differenzenquotienten

$$\frac{F(t_n) - F(t)}{t_n - t} = \int_{\Omega} \frac{f(x, t_n) - f(x, t)}{t_n - t} d\mu(x).$$

Nun gilt:

i) Der Integrand konvergiert (punktweise) gegen die partielle Ableitung $\frac{\partial f}{\partial t}(x, t)$.

ii) Nach dem Mittelwertsatz gilt

$$\left| \frac{f(x, t_n) - f(x, t)}{t_n - t} \right| \leq g(x) |t - t_n|$$

und da $t - t_n$ beschränkt ist, ist die rechte Seite eine integrierbare Majorante.

Die Behauptung folgt nun aus dem Satz von der dominierten Konvergenz. \square

Korollar 13.4. Ist in der Situation von Satz 13.3 die Funktion f partiell differenzierbar in $t \in D \subset \mathbb{R}^p$, dann gilt für $i = 1, \dots, p$

$$\frac{\partial F}{\partial t_i}(t) = \int_M \frac{\partial f}{\partial t_i}(x, t) d\mu(x).$$

Beispiel (Γ -Funktion). Zur Berechnung von $\int_0^{\infty} x \exp(-tx) dx$ gehen wir wie folgt vor: Wir betrachten $x \geq 0$ und $t \geq \epsilon$ für ein $\epsilon > 0$. Es gilt $x \exp(-tx) = -\frac{\partial}{\partial t} [\exp(-tx)]$ und da $t > \epsilon$, ist $x \exp(-\epsilon x)$ eine Lebesgue-integrierbare Majorante. Es gilt also

$$\int_0^{\infty} x \exp(-tx) dx = \int_0^{\infty} -\frac{\partial}{\partial t} [\exp(-tx)] dx = -\frac{\partial}{\partial t} \int_0^{\infty} \exp(-tx) dx.$$

Mit der Substitution $y = tx$ folgt $\int_0^{\infty} \exp(-tx) dx = \frac{1}{t} \int_0^{\infty} \exp(-y) dy = \frac{1}{t} [-\exp(-y)]_0^{\infty} = \frac{1}{t}$.

Wir schließen

$$\int_0^{\infty} x \exp(-tx) dx = -\frac{\partial}{\partial t} \left[\frac{1}{t} \right] = \frac{1}{t^2}.$$

Wiederholung obiger Argumente erlaubt wiederholtes Ableiten dieser Gleichung nach t (und Differenzieren unter dem Integral). Es folgt

$$\int_0^{\infty} x^2 \exp(-tx) dx = \frac{2}{t^3} \quad \int_0^{\infty} x^3 \exp(-tx) dx = \frac{6}{t^4}$$

und rekursiv ergibt sich

$$\int_0^{\infty} x^n \exp(-tx) dx = \frac{n!}{t^{n+1}}.$$

Für $t = 1$ bekommen wir für alle $n \in \mathbb{N}$

$$\int_0^{\infty} x^n \exp(-x) dx = n!.$$

Beispiel (Gaußsches Integral). Wir betrachten das sogenannte Gaußsche Integral $\int_0^{\infty} \exp(-x^2/2) dx$. Um dies zu berechnen, betrachten wir zuerst

$$g(t) = \left(\int_0^t \exp(-x^2/2) dx \right)^2.$$

und leiten diese nach t ab. Wir bekommen nach der Kettenregel und dem Fundamentalsatz

$$\begin{aligned} g'(t) &= 2 \left(\int_0^t \exp(-x^2/2) dx \right) \exp(-t^2/2) \\ &= 2 \int_0^t \exp(-(t^2 + x^2)/2) dx. \end{aligned}$$

Mit der Substitution $u = x/t$ wird dies zu

$$g'(t) = 2 \int_0^1 t \exp(-(1+u^2)t^2/2) du.$$

Wir bemerken, dass für den Integranden gilt

$$t \exp(-(1+u^2)t^2/2) = -\frac{\partial}{\partial t} \frac{\exp(-(1+u^2)t^2/2)}{1+u^2}$$

und da dies eine integrierbare Funktion ist, gilt

$$\begin{aligned} g'(t) &= 2 \int_0^1 -\frac{\partial}{\partial t} \frac{\exp(-(1+u^2)t^2/2)}{1+u^2} du \\ &= -2 \frac{\partial}{\partial t} \int_0^1 \frac{\exp(-(1+u^2)t^2/2)}{1+u^2} du \end{aligned}$$

Bilden wir auf beiden Seiten die Stammfunktion, folgt

$$g(t) = -2 \int_0^1 \frac{\exp(-(1+u^2)t^2/2)}{1+u^2} du + C.$$

Für $t \rightarrow 0$, folgt einerseits aus der Definition von g und obiger Formel

$$\begin{aligned} 0 = g(0) &= -2 \int_0^1 \frac{1}{1+u^2} du + C \\ &= -2 \arctan(u) \Big|_{u=0}^1 + C = -\frac{\pi}{2} + C. \end{aligned}$$

Also ist $C = \frac{\pi}{2}$ und für $t \rightarrow \infty$ bekommen wir mit Hilfe des Satzes von der dominierten Konvergenz

$$\begin{aligned} \left(\int_0^{\infty} \exp(-t^2/2) dt \right)^2 &= \lim_{t \rightarrow \infty} g(t) \\ &= \lim_{t \rightarrow \infty} \left(-2 \int_0^1 \frac{\exp(-(1+u^2)t^2/2)}{1+u^2} du + \frac{\pi}{2} \right) \\ &= -2 \int_0^1 \underbrace{\lim_{t \rightarrow \infty} \frac{\exp(-(1+u^2)t^2/2)}{1+u^2}}_{=0} du + \frac{\pi}{2} \\ &= \frac{\pi}{2}. \end{aligned}$$

Dies motiviert die Definition der *Eulerschen Γ -Funktion* als Interpolation der Fakultät: Für $s > 0$ definiert man

$$\Gamma(s) := \int_0^{\infty} x^{s-1} \exp(-x) dx$$

und es gilt $\Gamma(n+1) = n!$.

Es gilt also

$$\int_0^{\infty} \exp(-t^2/2) dt = \sqrt{\frac{\pi}{2}}.$$

14 Der Raum L^1

Es sei $(\Omega, \mathfrak{A}, \mu)$ ein Maßraum. In Satz 11.3 hatten wir schon den zugehörigen Vektorraum $\mathcal{L}^1(\Omega, \mu)$ aller integrierbaren Funktionen kennengelernt. Wir definieren nun für $f : \Omega \rightarrow \mathbb{R}$

$$\|f\|_{L^1} := \int_{\Omega} |f| d\mu \in [0, \infty]$$

Es gilt: $f \in \mathcal{L}^1(\Omega, \mu)$ genau dann, wenn $\|f\|_{L^1} < \infty$.

Tatsächlich ist dies keine Norm auf $\mathcal{L}^1(\Omega, \mu)$, aber es gilt:

- i) $\|f\|_{L^1} = 0 \iff f = 0$ μ -fast überall
- ii) $\|cf\|_{L^1} = |c| \|f\|_{L^1}$ für alle $c \in \mathbb{R}$
- iii) $\|f + g\|_{L^1} \leq \|f\|_{L^1} + \|g\|_{L^1}$, falls $f + g$ μ -fast überall definiert ist.

Weiterhin gilt:

$$\text{iv) } \|f_k - f\|_{L^1} \xrightarrow{k \rightarrow \infty} 0 \implies \int_{\Omega} f_k d\mu \rightarrow \int_{\Omega} f d\mu.$$

Da die positive Definitheit nicht gegeben ist, nennt man $\|\cdot\|_{L^1}$ nur *Seminorm*. Man kann sich aber folgendermaßen abhelfen: Wir definieren

$$\mathcal{N} := \{f \in \mathcal{L}^1(\Omega, \mu) \mid \|f\|_{L^1} = 0\}$$

und bemerken, dass \mathcal{N} ein Unterraum von $\mathcal{L}^1(\Omega, \mu)$ ist.

Wir bilden nun den Quotientenvektorraum

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

Was bedeutet das konkret in diesem Fall: Wir betrachten zwei Funktionen $f, g \in \mathcal{L}^1(\Omega, \mu)$ als äquivalent, wenn $f - g \in \mathcal{N}$ gilt, d.h.

$$f \sim g \iff f = g \text{ } \mu\text{-fast überall.}$$

Die Elemente von $L^1(\Omega, \mu)$ sind Äquivalenzklassen bezüglich dieser Äquivalenzrelation und alle Operationen (wie Summen, skalare Multiplikation, Betrag, Produkt...) sind mit Hilfe von Repräsentanten definiert. Insbesondere gilt: Ist $f \sim g$, so ist auch $\int f d\mu = \int g d\mu$.

Wir werden im Folgenden immer die Äquivalenzklassen unter den Tisch fallen lassen und wieder von Funktionen $f \in L^1(\Omega, \mu)$ sprechen. Das ist nicht ganz ungefährlich, denn eine solche "Funktion" ist nur bis auf Abänderung auf einer beliebigen Nullmenge bestimmt. Insbesondere kann man für eine Funktion $f \in L^1(\Omega, \mu)$ dann nicht mehr vom "Wert in einem Punkt" sprechen, d.h. der Ausdruck $f(x)$ ist nicht mehr wohldefiniert. Hingegen sind Integrale $\int_A f d\mu$ für alle Teilmengen $A \subset \Omega$ immer erklärt.

Beachte die verschiedenen L : L^1 an der Norm \mathcal{L}^1 für den Raum.

Um den Unterschied zwischen “Funktion” (im klassischen Sinn) und Funktion “in $L^1(\Omega, \mu)$ ” zu verdeutlichen, nennen wir Funktionen im zweiten Fall auch L^1 -Funktionen.

Wir untersuchen den Raum $L^1(\Omega, \mu)$ weiter:

Für einen σ -endlichen Maßraum $(\Omega, \mathfrak{A}, \mu)$ betrachten wir den Vektorraum der einfachen Funktionen

$$E(\Omega) = \left\{ \sum_{k=1}^m c_k \chi_{A_k} \mid c_k \in \mathbb{R}, A_k \in \mathfrak{A}, \mu(A_k) < \infty \right\}.$$

Ist \mathfrak{A} von einem Mengerring \mathfrak{A}_0 erzeugt, so betrachten wir außerdem den Vektorraum der *ganz einfachen Funktionen*

$$E_0(\Omega) = \left\{ \sum_{k=1}^m c_k \chi_{A_k} \mid c_k \in \mathbb{R}, A_k \in \mathfrak{A}_0, \mu(A_k) < \infty \right\}.$$

Es gilt $E_0(\Omega) \subset E(\Omega) \subset L^1(\Omega, \mu)$ und es handelt sich jeweils um Unterräume.

Ein wichtiges Beispiel ist $\Omega = \mathbb{R}^d$, $\mathfrak{A} = \mathcal{B}(\mathbb{R}^d)$ und $\mathfrak{A}_0 = \mathcal{Q}(\mathbb{R}^d)$.

Satz 14.1. Die Vektorräume $E_0(\Omega)$ und $E(\Omega)$ liegen dicht in $L^1(\Omega, \mu)$, d.h. zu jedem $f \in L^1(\Omega, \mu)$ und jedem $\epsilon > 0$ gibt es ein $\varphi \in E_0(\Omega)$ mit $\|f - \varphi\|_{L^1} < \epsilon$.

Beweis.

1. Jedes integrierbare f ist per $f = f_+ - f_-$ die Differenz zweier nicht-negativer Funktionen und daher beschränken wir uns auf den Fall $f \geq 0$. Nach Definition der Integrierbarkeit gibt es also $\varphi_k \in E(\Omega)$ mit $\varphi_k \nearrow f$ und $\int \varphi_k d\mu \nearrow \int f d\mu$. Es ist also $f - \varphi_k = |f - \varphi_k| \geq 0$ und es folgt für k groß genug

$$\|f - \varphi_k\|_{L^1} = \int |f - \varphi_k| d\mu = \int f - \varphi_k d\mu < \epsilon.$$

Damit ist gezeigt, dass $E(\Omega)$ in $L^1(\Omega)$ dicht liegt.

2. Um zu zeigen, dass auch $E_0(\Omega)$ dicht liegt, reicht es, zu zeigen, dass sich jede einfache Funktion durch ganz einfacher Funktionen approximieren lässt und dafür reicht es, zu zeigen: Ist $A \in \mathfrak{A}$, mit $\mu(A) < \infty$ und $\epsilon > 0$, so gibt es $S \in \mathfrak{A}_0$ mit $\|\chi_A - \chi_S\|_{L^1} < \epsilon$. Diese Norm ist aber

$$\|\chi_A - \chi_S\|_{L^1} = \int \chi_{A \Delta S} d\mu = \mu(A \Delta S)$$

und nach Korollar 6.7 kann man die rechte Seite beliebig klein machen. □

Im konkreten Fall des Lebesgue-Maßes folgt also: Die Menge der ganz einfachen Funktionen (also der endlichen Linearkombinationen von charakteristischen Funktionen von Quadern) liegt dicht in $L^1(\mathbb{R}^d, \lambda^d)$.

Wir können also i.A. für eine L^1 -Funktion nicht mehr vom Wert $f(x)$ sprechen. Ist $\Omega = \mathbb{R}^d$, und $\mu = \lambda^d$ so können aber immer noch Integrale $\int_{B_r(x)} f d\lambda^d$ betrachten. Wenn wir diese mit dem Lebesgue-Maß des Balles $B_r(x)$ normieren, bekommen wir den Mittelwert von f über diesen Ball: $\frac{1}{\lambda^1(B_r(x))} \int_{B_r(x)} f d\lambda^d$. Die Frage, ob, wann und wogegen diese Mittelwerte für $r \rightarrow 0$ konvergieren ist gar nicht einfach. Man kann (mit einigem Aufwand) zeigen, dass diese Grenzwerte für λ^d -fast alle x existieren und man daher einen wohldefinierten λ^d -fast überall definierten Repräsentanten von f hat (und die Ausnahmemenge ist ebenfalls fixiert). Dies findet man unter dem Stichwort *Lebesguescher Differentialsatz*.

Die Approximation von L^1 -Funktionen mit einfachen Funktionen ist oft nützlich. Manchmal möchte man aber auch mit glatteren Funktionen approximieren. Das geht auch und wir zeigen den Fall von stetigen Funktionen. Wir brauchen dazu einen topologischen Raum X und bezeichnen mit $C(X)$ den Vektorraum der stetigen Funktionen von X nach \mathbb{R} .

Definition 14.2. Zu $f : X \rightarrow \mathbb{R}$ bezeichnet

$$\text{supp}(f) := \overline{\{x \in X \mid f(x) \neq 0\}}$$

den Träger von f .

Konkret betrachten wir $X = U \subset \mathbb{R}^d$ offen. Wir bezeichnen mit

$$C_c(U) := \{f \in C(U) \mid \text{supp}(f) \subset U \text{ kompakt}\}.$$

Jedes $f \in C_c(U)$ lässt sich per

$$\tilde{f}(x) := \begin{cases} f(x), & x \in U \\ 0, & \text{sonst} \end{cases}$$

zu einem $\tilde{f} \in C(\mathbb{R}^d)$ fortsetzen (insbesondere ist \tilde{f} stetig). Daher ist $C_c(U)$ in natürlicher Weise ein Unterraum von $C_c(\mathbb{R}^d)$. Wir wollen zeigen, dass $C_c(U)$ in $L^1(U)$ dicht liegt und zeigen dafür zuerst ein Lemma:

Lemma 14.3. *Es sei $U \subset \mathbb{R}^d$ offen und nicht leer. Dann gibt es eine Folge $h_k \in C_c(U)$ mit $0 \leq h_k \leq 1$ und $h_k(x) \xrightarrow{k \rightarrow \infty} 1$ für alle $x \in U$.*

Beweis.

1. Im ersten Schritt konstruieren wir zu beliebigem kompaktem $K \subset U$ ein $h \in C_c(U)$ welches schon $0 \leq h \leq 1$ und $h|_K \equiv 1$ erfüllt: Dazu definieren wir $r = \text{dist}(K, \mathbb{R}^d \setminus U)$ (und aus Analysis 2 wissen wir, dass $r > 0$ gilt) und setzen

$$h(x) = \max(0, 1 - \frac{2}{r} \text{dist}(x, K)).$$

Man kann hierfür die geforderten Eigenschaften einfach nachrechnen.

2. Im zweiten Schritt nutzen wir, dass U von Innen mit kompakten Mengen ausgeschöpft werden kann, d.h. es gibt eine kompakte Mengen K_m mit $K_m \nearrow U$ und zu jedem gibt es ein h_m wie aus Schritt 1. Gehen wir noch von h_m zu $\max(h_m, h_{m-1})$ über, so ist die Folge der h_m auch wachsend und die Behauptung ist gezeigt.

□

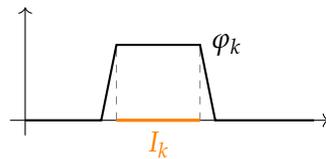
Satz 14.4. *Sei $U \subset \mathbb{R}^d$ offen und $L^1(U) = L^1(U, \mathcal{B}(U), \lambda^d)$. Dann liegt der Raum $C_c(U)$ dicht in $L^1(U)$.*

Beweis.

1. Im ersten Schritt behandeln wir den Fall $U = \mathbb{R}^d$. Da die ganz einfachen Funktionen schon dicht liegen, reicht es, sich auf den Fall $f = \chi_Q$ mit einem Quader Q zu beschränken.

Dazu: Ein $f \in L^1(U)$ approximieren wir zuerst durch ein $g = \sum_{k=1}^m c_k \chi_{Q_k}$, d.h. wir wählen c_k und Q_k so, dass $\|f - g\|_{L^1} < \epsilon/2$ und dann approximieren wir jedes χ_{Q_k} mit einem stetigen $h_k \in C_c(\mathbb{R}^d)$, so dass $\|c_k \chi_{Q_k} - h_k\| \leq \epsilon/(2m)$ denn dann ist $h = \sum_k h_k \in C_c(\mathbb{R}^d)$ und es gilt $\|f - h\|_{L^1} \leq \|f - g\|_{L^1} + \|g - h\|_{L^1} \leq \epsilon/2 + \sum_{k=1}^m \|c_k \chi_{Q_k} - h_k\|_{L^1} \leq \epsilon$.

Dies geht "per Hand", indem wir für $Q = I_1 \times \dots \times I_d$ (also $\chi_Q(x_1, \dots, x_d) = \chi_{I_1}(x_1) \dots \chi_{I_d}(x_d)$) ein stetige φ_k wie in der Skizze definieren



und dann $\varphi(x) = \varphi_1(x_1) \dots \varphi_d(x_d)$ setzen.

2. Im zweiten Schritt sei nun $U \subset \mathbb{R}^d$ offen und $f \in L^1(U)$ wir setzen f durch Null zu $\tilde{f} \in L^1(\mathbb{R}^d)$ fort und nach Teil 1. gibt es $G \in C_c(\mathbb{R}^d)$ mit $\|\tilde{f} - G\|_{L^1(\mathbb{R}^d)} < \epsilon/2$. Schränken wir G zu $g := G|_U$ ein, so ist zwar $g \in C(U)$, hat aber ggf. keinen kompakten Träger mehr. Die können wir aber durch Multiplikation mit einem geeigneten h_k aus Lemma 14.3 erreichen. d.h. wir nehmen $h_k g$ statt g . Mit dem Satz von der dominierten Konvergenz kann man nun die Behauptung zeigen.

□

Schließlich zeigen wir, dass der Raum $L^1(\Omega, \mu)$ vollständig ist. Dazu zeigen wir zuerst ein Ergebnis über den Zusammenhang von Konvergenz bezüglich der L^1 -Norm und punktweiser fast überall Konvergenz.

Zuerst ein Lemma:

Lemma 14.5. Ist $g_k \in L^1(\Omega, \mu)$, mit $\sum_{k=1}^{\infty} \|g_k\|_{L^1} =: M < \infty$, dann konvergiert die Folge $h_m = \sum_{k=1}^m g_k$ der Partialsummen μ -fast überall gegen eine Funktion $g \in L^1(\Omega, \mu)$ und es gilt $\|h_m - g\|_{L^1} \rightarrow 0$.

15 Die Invarianz der Lebesgue-Maßes

In diesem Abschnitt wollen wir untersuchen, wie sich das Lebesgue-Maß bei Verformung mit linearen Abbildungen verhält. D.h. wir wollen untersuchen, wie das Lebesgue-Maß von Mengen der Form

$$TA = \{Tx \mid x \in A\}$$

für (affin) lineare Abbildungen $T : \mathbb{R}^d \rightarrow \mathbb{R}^d$ und Lebesgue-messbare Mengen $A \subset \mathbb{R}^d$ ist.

Der Fall von Verschiebungen ist einfach: Ist $T_a(x) = a + x$ für $a \in \mathbb{R}^d$ und $T_a(A) = a + A$ die verschobene Menge, so ist gilt für das Lebesgue-Maß $\lambda^d(A) = \lambda^d(T_a(A))$. Man sagt auch: Das Lebesgue-Maß ist *translationsinvariant*.

Tatsächlich ist das Lebesgue-Borel-Maß das einzige translationsinvariante Maß auf dem Borelschen Maßraum was dem Einheitswürfel das Maß 1 gibt:

Satz 15.1. *Es sei $\mu : \mathcal{B}(\mathbb{R}^d) \rightarrow [0, \infty]$ ein translationsinvariantes Maß, für das es eine beschränkte offene Borelmenge B_0 mit $0 < \mu(B_0) < \infty$ gibt, so gibt es ein $c > 0$ mit $\mu = c\lambda^d$, d.h. μ ist ein Vielfaches des Lebesgue-Borel-Maßes.*

Beweis.

Es sei

$$W_q = \{x \in \mathbb{R}^d \mid 0 \leq x_i \leq 1/q\}$$

ein Würfel mit Kantenlänge q . Da B_0 offen ist, gibt es ein $q \in \mathbb{N}$ (groß genug) und eine Verschiebung um a , so dass $a + W_q \subset B_0$ liegt. Daher ist $\mu(W_q) \leq \mu(B_0) < \infty$. Da B_0 beschränkt ist, lässt es sich durch endlich viele Mengen der Form $a + W_q$ überdecken, woraus $\mu(W_q) > 0$ folgt. Da der Einheitswürfel W_1 eine disjunkte Vereinigung von q^n Würfeln W_q ist, folgt

$$0 < \underbrace{\mu(W_1)}_{=:c} = q^n \mu(W_q)$$

und es folgt $\mu(W_q) = cq^{-n}$ und auch $\mu(W_1) = c\lambda^d(W_1)$.

Daraus bekommen wir für alle Quader Q ebenfalls $\mu(Q) = c\lambda^d(Q)$. Also stimmen die beiden Maße λ^d und μ/c auf $\mathcal{Q}(\mathbb{R}^d)$ über ein, und damit (wegen der Eindeutigkeit der Fortsetzung) auch auf $\mathcal{B}(\mathbb{R}^d)$. \square

Lemma 15.2. *Es sei $T \in \mathbb{R}^{d \times d}$ invertierbar. Dann ist das Maß $\mu(B) = \lambda^d(T(B))$ ein translationsinvariantes Maß.*

Beweis.

Sind die Mengen B_k paarweise disjunkt, so sind auch die $T(B_k)$

Eine affin lineare Abbildung von \mathbb{R}^d nach \mathbb{R}^d ist von der Form $T(x) = Mx + b$ mit $M \in \mathbb{R}^{d \times d}$, $b \in \mathbb{R}^d$.

Verschobene Quader haben per Definition das gleiche Lebesgue-Maß, und zum Ausschöpfen von $a + A$ kann man die gleichen (nur verschobenen) Quader wie für A nehmen.

paarweise disjunkt, und es gilt

$$\begin{aligned}\mu\left(\bigcup_k B_k\right) &= \lambda^d\left(T\left(\bigcup_k B_k\right)\right) = \lambda^d\left(\bigcup_k T(B_k)\right) \\ &= \sum_k \lambda^d(T(B_k)) = \sum_k \mu(B_k).\end{aligned}$$

Außerdem gilt $T(a + B) = Ta + TB$ und also ist $\mu(a + B) = \lambda^d(Ta + TB) = \lambda^d(TB) = \mu(B)$, d.h. μ ist translationsinvariant.

□

Zusammen mit dem Satz davor folgt:

Korollar 15.3. Zu invertierbarem $T \in \mathbb{R}^{d \times d}$ sei $\mu(B) = \lambda^d(T(B))$. Dann existiert ein $c > 0$, so dass $\mu(B) = c\lambda^d(B)$.

Wir müssen nur zeigen, dass es eine offene Borelmenge B_0 gibt, für die $\mu(B_0) > 0$ gilt. Da T invertierbar ist, existiert die Umkehrabbildung T^{-1} und ist als lineare Abbildung stetig. Daher ist TB_0 für offenes B_0 wieder offen und es folgt $\mu(B_0) = \lambda^d(TB_0) > 0$, da offene Mengen positives Lebesguemaß haben.

Satz 15.4 (Bewegungsinvarianz des Lebesguemaßes). Es sei $S \in \mathbb{R}^{d \times d}$ eine orthonormale Matrix (d.h. es gilt $S^T = S^{-1}$) und $a \in \mathbb{R}^d$. Dann gilt für alle Borel-Mengen B

$$\lambda^d(a + SB) = \lambda^d(B).$$

Beweis.

Da wir schon die Translationsinvarianz des Lebesgue-Maßes gezeigt haben, können wir $a = 0$ annehmen. Wir definieren $\mu(B) := \lambda(S(B))$. Nach Korollar 15.3 gibt es ein c , so dass $\mu(B) = c\lambda(B)$. Für die Einheitskugel $K = \{x \in \mathbb{R}^d \mid \|x\|_2 \leq 1\}$ gilt $S(K) = K$ und daher ist $\mu(K) = \lambda(S(K)) = \lambda(K)$ und daher ist $c = 1$, was die Behauptung zeigt. □

Die Abbildung $x \mapsto a + Sx$ ist eine Drehung/Spiegelung, gefolgt von einer Verschiebung. Solche Abbildungen nennt man auch *Bewegungen* und daher sagt dieser Satz: Das Lebesgue-Maß ist invariant unter Bewegungen.

Hier ein Lemma aus der linearen Algebra:

Lemma 15.5. Zu jedem invertierbarem $A \in \mathbb{R}^{d \times d}$ existieren zwei orthonormale Matrizen $S_1, S_2 \in \mathbb{R}^{d \times d}$ und eine Diagonalmatrix D mit positiver Diagonale, so dass $A = S_1 D S_2$.

Beweis.

Nach dem Spektralsatz für symmetrische Matrizen gibt es ein orthonormales S und ein $D_1 = \text{diag}(\lambda_1, \dots, \lambda_d)$ mit $\lambda_k > 0$, so dass

$$S^T A^T A S = D_1.$$

Wir definieren $D = \text{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_d})$ und $S_1 = A S D^{-1}$ und bekommen

$$S_1^T S_1 = D^{-1} S^T A^T A S D^{-1} = D^{-1} D_1 D^{-1} = E_d$$

Die ist tatsächlich ein Spezialfall der Singulärwertzerlegung.

und daher ist S_1 orthonormal. Setzen wir noch $S_2 = S^{-1}$ (was wieder orthonormal ist), so bekommen wir aus $S_1 = ASD^{-1}$ die behauptete Darstellung $A = S_1DS^{-1} = S_1DS_2$. \square

Satz 15.6 (Transformation des Lebesgue-Maßes unter linearen Abbildungen). *Ist $T \in \mathbb{R}^{d \times d}$ invertierbar, so gilt für alle Borel-Mengen B*

$$\lambda^d(T(B)) = |\det(T)|\lambda^d(B).$$

Beweis.

1. Zeigen wir zuerst den Fall $T = D = \text{diag}(\sigma_1, \dots, \sigma_d)$ einer Diagonalmatrix mit $\sigma_k > 0$. Für die Einheitswürfel W_1 gilt dann $TW_1 = [0, \sigma_1] \times \dots \times [0, \sigma_d]$, und daher gilt

$$\lambda^d(TW_1) = \prod_{i=1}^d \sigma_i = \det(D)\lambda^d(W_1)$$

und die Anwendung von Korollar 15.3 auf das Maß $\mu(B) = \lambda^d(TB)$ zeigt die Behauptung.

2. Für den Allgemeinen Fall zerlegen wir $T = S_1DS_2$ nach Lemma 15.5. Dann gilt $|\det(T)| = \det(D)$ und mit dem ersten Punkt sehen wir

$$\lambda^d(T(B)) = \lambda^d(S_1DS_2B) = \lambda^d(DS_2B) = \det(D)\lambda^d(S_2B) = \det(D)\lambda^d(B),$$

woraus die Behauptung folgt. \square

Im Fall von nicht-invertierbarem T ist die Behauptung auch war: Dann ist nämlich $\det(T) = 0$ und wegen $\text{rang}(T) < d$ und $T(B)$ immer in eine Nullmenge (da $T(B)$ in einer Hyperebene enthalten ist, vgl. Abschnitt 8).

Beispiel. 1. Für eine einfache Skalierung $x \mapsto rx$, die von der Matrix $T = rE_d$ beschrieben wird, gilt also

$$\lambda^d(rB) = |\det(rE_d)|\lambda^d(B) = r^d\lambda^d(B).$$

2. Ein Parallelepipied P lässt sich schreiben als $P = TW_1$ mit einem invertierbarem $T \in \mathbb{R}^{d \times d}$ und dem Einheitswürfel W_1 . Daher folgt

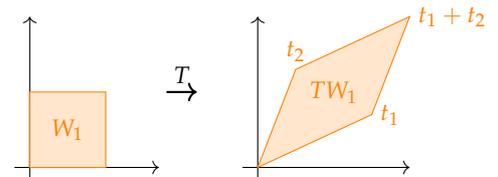
$$\lambda^d(P) = \lambda^d(TW_1) = |\det(T)|.$$

Definition 15.7 (Bildmaß). Es sei $(\Omega, \mathfrak{A}, \mu)$ ein Maßraum, (Ω', \mathfrak{A}') ein messbarer Raum und $T : \Omega \rightarrow \Omega'$ eine messbare Abbildung. Dann ist ν definiert durch

$$\nu(A') := \mu(T^{-1}(A')), \quad A' \in \mathfrak{A}'$$

ein Maß, genannt das Bildmaß von μ unter T und wird auch mit $\nu = T\#\mu$ (oder auch $T_*\mu$) geschrieben.

Für $T = [t_1 t_2] \in \mathbb{R}^{2 \times 2}$ sieht die Situation so aus:



Satz 15.8 (Integration bezüglich des Bildmaßes). *Es sei $(\Omega, \mathfrak{A}, \mu)$ ein Maßraum, (Ω', \mathfrak{A}') ein messbarer Raum und $T : \Omega \rightarrow \Omega'$ eine messbare Abbildung. Dann gilt: Die Funktion $f : \Omega' \rightarrow \mathbb{R}$ ist genau dann integrierbar bezüglich des Bildmaßes $T_{\#}\mu$, wenn $f \circ T : \Omega \rightarrow \mathbb{R}$ bezüglich μ integrierbar ist und es gilt*

$$\int_{\Omega} (f \circ T) d\mu = \int_{\Omega'} f d(T_{\#}\mu).$$

Beweis.

Wir zeigen die Behauptung erst für charakteristische Funktionen, dann für einfache, dann für nicht-negative und schließlich für integrierbare Funktionen.

Das Beweisprinzip hier nennt man auch (vielleicht irreführend) "algebraische Induktion" oder auch "maßtheoretische Induktion".

1. Für messbare charakteristische Funktionen $f = \chi_B$ gilt

$$\begin{aligned} (\chi_B \circ T)(x) = 1 &\iff T(x) \in B \\ &\iff x \in T^{-1}(B) \end{aligned}$$

d.h. es gilt $\chi_B \circ T = \chi_{T^{-1}(B)}$. Also gilt

$$\begin{aligned} \int \chi_B \circ T d\mu &= \int \chi_{T^{-1}(B)} d\mu = \mu(T^{-1}(B)) \\ &= T_{\#}\mu(B) = \int \chi_B d(T_{\#}\mu). \end{aligned}$$

2. Ist f einfach, d.h. $f = \sum_k c_k \chi_{B_k}$ (mit endlicher Summe) so nutzen wir Linearität des Integrals und bekommen aus dem ersten Teil

$$\begin{aligned} \int f \circ T d\mu &= \int \sum_k c_k (\chi_{B_k} \circ T) d\mu = \sum_k c_k \int (\chi_{B_k} \circ T) d\mu \stackrel{1.}{=} \sum_k c_k \int \chi_{B_k} d(T_{\#}\mu) \\ &= \int \sum_k c_k \chi_{B_k} d(T_{\#}\mu) = \int f d(T_{\#}\mu). \end{aligned}$$

3. Sei nun f nicht-negativ und messbar. Dann approximieren wir f von unten durch einfache Funktionen $\varphi_k \nearrow f$ und diese erfüllen dann ebenfalls $\varphi_k \circ T \nearrow f \circ T$. Mit dem zweiten Schritt und monotoner Konvergenz folgt

$$\begin{aligned} \int f \circ T d\mu &= \lim_k \int \varphi_k \circ T d\mu \stackrel{2.}{=} \lim_k \int \varphi_k d(T_{\#}\mu) \\ &= \int f d(T_{\#}\mu) \end{aligned}$$

und insbesondere ist die rechte Seite genau dann endlich (und damit integrierbar) genau dann, wenn die linke Seite endlich (und damit integrierbar) ist.

4. Ist jetzt f integrierbar, so zerlegen wir $f = f_+ - f_-$ und wenden Schritt 3. auf f_+ und f_- an.

□

Zwei direkte Folgerungen im Fall des Lebesgue-Maßes:

Korollar 15.9. Ist $T \in \mathbb{R}^{d \times d}$ invertierbar und $a \in \mathbb{R}^d$, so gilt:

1. $f : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$ genau dann Lebesgue-integrierbar, wenn die Funktion $x \mapsto f(Tx + a)$ Lebesgue-integrierbar ist und es gilt

$$\int_{\mathbb{R}^d} f(Tx + a) d\lambda^d(x) = \frac{1}{|\det(T)|} \int_{\mathbb{R}^d} f(x) d\lambda^d(x).$$

2. Im Fall einer Permutationsmatrix T bekommen wir: Für jede Permutation $\pi : \{1, \dots, d\} \rightarrow \{1, \dots, d\}$ gilt

$$\int_{\mathbb{R}^d} f(x_{\pi(1)}, \dots, x_{\pi(d)}) d\lambda^d(x) = \int_{\mathbb{R}^d} f(x_1, \dots, x_d) d\lambda^d(x).$$

16 Iterierte Integrale und die Sätze von Fubini und Tonelli

In diesem Abschnitt behandeln wir ausschließlich das Lebesgue-Maß. Wir werden also nur messbar statt Lebesgue-messbar usw. schreiben.

Bisher können wir Lebesgue-Integrale eigentlich nur auf zwei Arten ausrechnen: Entweder direkt über die Definition indem wir von unten mit einfachen Funktionen approximieren oder, und das geht nur im eindimensionalen Fall, als Riemann-Integral. Für Lebesgue-Integrale in höheren Dimensionen haben wir bisher nur die erste Methode. Das soll sich nun ändern: Zuerst behandeln wir den Satz von Fubini, der zeigt, wie man Integrale über Rechtecke berechnen kann und später kommen wir zum Transformationsatz, der dann für krummlinig berandete Gebiete funktioniert.

Eine praktische Technik für die Berechnung von Integralen in höheren Dimensionen, ist es, sukzessive über die einzelnen Dimensionen zu integrieren, da wir dann auf die Techniken aus dem ersten Semester zurückgreifen können. Konkret würde dies für eine Funktion $f : [a, b] \times [c, d] \rightarrow \mathbb{R}$ bedeuten, dass wir

$$\int_{[a,b] \times [c,d]} f d\lambda$$

berechnen als

$$\int_c^d \int_a^b f(x_1, x_2) dx_1 dx_2$$

bzw. schrittweise: Erst berechnet man $g(x_2) = \int_a^b f(x_1, x_2) dx_1$ und schließlich $\int_c^d g(x_2) dx_2$ (und dabei benutzt man jeweils z.B. den Hauptsatz).

Es stellt sich allerdings die Frage, ob dies tatsächlich das richtige Ergebnis liefert, denn immerhin handelt es sich um eine Vertauschung von Grenzwerten.

Wir nähern uns der Problematik im Rahmen des Lebesgue-Maßes und fragen uns erst einmal folgendes: Ist $M \subset \mathbb{R}^d$ Lebesgue-messbar, wie hängt dann das Lebesgue-Maß $\lambda^d(M)$ mit den Lebesgue-Maßen der Schnitte von M zusammen, also mit den Mengen

$$M(z) = \{(x_1, \dots, x_{d-1}) \mid (x_1, \dots, x_{d-1}, z) \in M\}$$

und den Größen $\lambda^{d-1}(M(z))$?

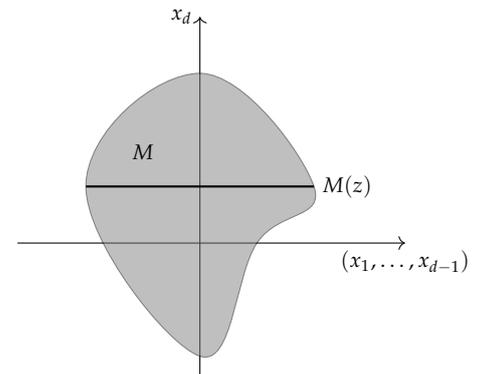
Satz 16.1 (Cavalierisches Prinzip, Schichtkuchen-Formel). *Es sei $M \subset \mathbb{R}^d$ messbar. Dann ist die Menge $M(z) \subset \mathbb{R}^{d-1}$ für fast alle $z \in \mathbb{R}$ ebenfalls messbar und die Funktion*

$$f(z) = \lambda^{d-1}(M(z))$$

(welche fast überall erklärt ist) ist messbar und es gilt

$$\lambda^d(A) = \int_{\mathbb{R}} f(z) d\lambda^1(z) = \int_{\mathbb{R}} \lambda^{d-1}(M(z)) d\lambda^1(z).$$

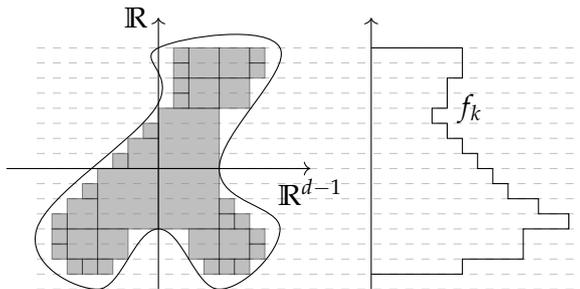
Die Sätze dieses Abschnittes gelten allerdings auch mit wenigen Zusatzannahmen in allgemeineren Fällen.



Ist $\lambda(M) < \infty$, so ist $f \in L^1(\mathbb{R})$.

Beweis.

Sei vorerst M offen. Dann können wir M nach Lemma 3.5 als abzählbare Vereinigung von Gitter-Würfeln darstellen. Bezeichne M_k die Vereinigung der Würfel in M auf dem Gitter der Feinheit 2^{-k} . Also besteht $M(z)$ aus abzählbar vielen $(d-1)$ -dimensionalen Würfeln $M_k(z)$. Wir setzen $f_k(z) := \lambda^{d-1}(M_k(z))$ und bemerken, dass f_k ein Grenzwert von nicht-negativen einfachen Funktionen ist und daher messbar.



Nach Konstruktion ist die Folge f_k punktweise wachsend mit $\int_{\mathbb{R}} f_k d\lambda^1 \leq \lambda^d(M)$. Der Satz 12.1 von der monotonen Konvergenz zeigt, dass für fast alle z der Grenzwert $f(z) := \lim_k f_k(z)$ existiert, f messbar ist und es gilt

$$\lim_k \int_{\mathbb{R}} f_k d\lambda^1 = \int_{\mathbb{R}} f d\lambda^1.$$

Wegen $\lim_k \lambda^d(M_k) = \lambda^d(M)$ folgt $\lambda^d(M) = \int_{\mathbb{R}} f d\lambda^1$ (schließlich ist $M(z) \subset \mathbb{R}^{d-1}$ offen und gleich der Vereinigung aller $M_k(z)$). Es folgt

$$f(z) = \lim_k \lambda^{d-1}(M_k(z)) = \lambda^{d-1}(M(z))$$

und die Behauptung ist für offenes M gezeigt.

Für messbares M folgt die Behauptung wieder mit einem Argument nach dem Prinzip der guten Mengen, was wir hier nicht ausführen. \square

Beispiel. Es sei $B \subset \mathbb{R}^{d-1}$ kompakt. Der Zylinder mit Basis B und Höhe $h > 0$ ist die Menge $Z := B \times [0, h] \subset \mathbb{R}^d$. Die Schnittmengen sind

$$M(z) = \begin{cases} B, & \text{falls } 0 \leq z \leq h \\ \emptyset, & \text{sonst} \end{cases}$$

und daher ist

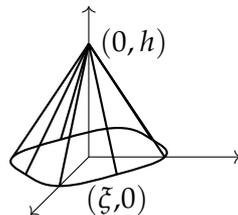
$$\lambda^d(Z) = \int_0^h \lambda^{d-1}(B) d\lambda^1 = h \lambda^{d-1}(B).$$

Wir können auch den *Kegel* mit Basis $B \subset \mathbb{R}^{d-1}$ und Höhe $h > 0$ betrachten:

$$C := \{((1-\alpha)\xi, \alpha h) \in \mathbb{R}^{d-1} \times \mathbb{R} \mid \xi \in B, 0 \leq \alpha \leq 1\}$$

Etwas allgemeiner gilt für $d = k + l$: Zu $y \in \mathbb{R}^l$ sei $A(y) = \{x \in \mathbb{R}^k \mid (x, y) \in A\}$, so ist $y \mapsto \lambda^k(A(y))$ messbar und es gilt $\lambda^d(A) = \int_{\mathbb{R}^l} \lambda^k(A(y)) d\lambda^l(y)$.

(d.h. die Menge besteht aus allen Verbindungsstrecken der Punkt $(\xi, 0)$ mit $\xi \in B$ mit dem Punkt $(0, h)$).



Die Schnittmengen $C(z) = \{\xi' \in \mathbb{R}^{d-1} \mid (\xi', z) \in C\}$ sind für $z < 0$ und $z > h$ leer, und sonst gilt

$$C(z) = \left(1 - \frac{z}{h}\right)B$$

und daher ist

$$\lambda^{d-1}(C(z)) = \left(1 - \frac{z}{h}\right)^{d-1} \lambda^{d-1}(B).$$

Es folgt also

$$\lambda^d(C) = \int_0^h \left(1 - \frac{z}{h}\right)^{d-1} \lambda^{d-1}(B) d\lambda^1(z).$$

Das eindimensionale Integral schreiben wir als Riemann-Integral und berechnen mit der Substitution $z/h = u$

$$\int_0^h \left(1 - \frac{z}{h}\right)^{d-1} dz = \int_0^1 (1-u)^{d-1} h du = \left[-\frac{1}{d}(1-u)^d \right]_{u=0}^1 h = \frac{h}{d}$$

und daher ist das Volumen des Kegels

$$\lambda^d(C) = \frac{h \lambda^{d-1}(B)}{d}.$$

Kommen wir nun zu iterierten Integralen. Hierzu zerlegen wir $\mathbb{R}^d = \mathbb{R}^k \times \mathbb{R}^l$ und wir schreiben immer $(x, y) \in \mathbb{R}^d$ mit $x \in \mathbb{R}^k$ und $y \in \mathbb{R}^l$.

Der Satz von Tonelli behandelt nicht-negative Funktionen.

Satz 16.2 (Satz von Tonelli). *Es sei $f : \mathbb{R}^k \times \mathbb{R}^l \rightarrow [0, \infty]$ eine messbare Funktion. Dann ist die Funktion $x \mapsto f(x, y)$ für fast alle $y \in \mathbb{R}^l$ eine messbare Funktion auf \mathbb{R}^k und insbesondere ist das Integral*

$$\int_{\mathbb{R}^k} f(x, y) d\lambda^k(x)$$

erklärt (mit Wert in $[0, \infty]$). Die Funktion

$$F : \mathbb{R}^l \rightarrow [0, \infty], \quad y \mapsto F(y) := \int_{\mathbb{R}^k} f(x, y) d\lambda^k(x)$$

ist messbar und es gilt

$$\int_{\mathbb{R}^{k+l}} f(x, y) d\lambda^{k+l}(x, y) = \int_{\mathbb{R}^l} F(y) d\lambda^l(y) = \int_{\mathbb{R}^l} \left(\int_{\mathbb{R}^k} f(x, y) d\lambda^k(x) \right) d\lambda^l(y).$$

Beweis.

Der Beweis benutzt wieder algebraische Induktion:

1. Für charakteristische Funktionen ist dies die Aussage von Satz 16.1 (genauer: Der Randbemerkung zum Cavalierischen Prinzip).
2. Für nicht-negative einfache Funktionen folgt die Aussage aus dem ersten Punkt mit der Linearität des Integrals.
3. Für allgemeine messbare nicht-negative Funktionen folgt die Behauptung mit Hilfe des Satzes von der monotonen Konvergenz.

□

Beispiel. Ein separable Funktionen $F : \mathbb{R}^{k+l} \rightarrow \bar{\mathbb{R}}$

$$F(x, y) = f(x)g(y)$$

ist integrierbar, wenn die Funktionen $f : \mathbb{R}^k \rightarrow \bar{\mathbb{R}}$ und $g : \mathbb{R}^l \rightarrow \bar{\mathbb{R}}$ integrierbar sind, denn dann ist

$$\begin{aligned} \int_{\mathbb{R}^{k+l}} |F(x, y)| d\lambda^{k+l}(x, y) &= \int_{\mathbb{R}^l} |g(y)| \left(\int_{\mathbb{R}^k} |f(x)| d\lambda^k(x) \right) d\lambda^l(y) \\ &= \int_{\mathbb{R}^k} |f(x)| d\lambda^k(x) \int_{\mathbb{R}^l} |g(y)| d\lambda^l(y) < \infty. \end{aligned}$$

Lässt man die Nicht-Negativität weg, so braucht man eine Integrierbarkeitsvoraussetzung (damit sich nicht "unendlich große positive und negative Volumina auf verschiedene Arten wegheben können"). Dann ist die gewünschte Aussage jedoch eine direkte Konsequenz aus dem Satz von Tonelli.

Satz 16.3 (Satz von Fubini). Sei $d = k + l$ und $f : \mathbb{R}^d \rightarrow \bar{\mathbb{R}}$ integrierbar. Dann ist die Abbildung $x \mapsto f(x, y)$ für fast alle $y \in \mathbb{R}^l$ integrierbar und $F(y) := \int_{\mathbb{R}^k} f(x, y) d\lambda^k(x)$ ist fast-überall definiert und ebenfalls integrierbar mit

$$\int_{\mathbb{R}^{k+l}} f(x, y) d\lambda^{k+l}(x, y) = \int_{\mathbb{R}^l} F(y) d\lambda^l(y) = \int_{\mathbb{R}^l} \left(\int_{\mathbb{R}^k} f(x, y) d\lambda^k(x) \right) d\lambda^l(y).$$

Beweis.

Da eine Funktion integrierbar ist, wenn Positiv- und Negativteil ein endliches Integral haben, reicht es, wenn wir den Satz für nicht-negative Funktionen $f \geq 0$ zeigen (und dann dies Ergebnis auf Positiv- und Negativteil anwenden). Aus dem Satz von Tonelli folgt schon die Messbarkeit von F und ebenso

$$\int_{\mathbb{R}^l} F(y) d\lambda^l(y) = \int_{\mathbb{R}^{k+l}} f(x, y) d\lambda^{k+l}(x, y) < \infty.$$

Also ist F aus \mathbb{R}^l integrierbar und die Behauptung ist schon gezeigt.

□

Per Symmetrie gilt natürlich auch

$$\begin{aligned}\int_{\mathbb{R}^{k+l}} f(x, y) d\lambda^{k+l}(x, y) &= \int_{\mathbb{R}^k} \left(\int_{\mathbb{R}^l} f(x, y) d\lambda^l(y) \right) d\lambda^k(x) \\ &= \int_{\mathbb{R}^l} \left(\int_{\mathbb{R}^k} f(x, y) d\lambda^k(x) \right) d\lambda^l(y).\end{aligned}$$

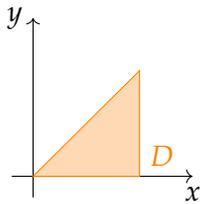
Rekursiv bekommt man

$$\int_{\mathbb{R}^d} f(x_1, \dots, x_d) d\lambda^d(x) = \int_{\mathbb{R}} \left(\dots \left(\int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x_1, \dots, x_d) d\lambda^1(x_1) \right) d\lambda^1(x_2) \right) \dots \right) d\lambda^1(x_d)$$

d.h. wir können Integrale über den \mathbb{R}^d (oder Rechtecke im \mathbb{R}^d) berechnen, indem wir jeweils über die einzelnen Variablen integrieren (wobei uns alle Techniken für das Riemann-Integral zur Verfügung steht). Dabei kommt es nicht auf die Reihenfolge an. Einzige Voraussetzung: Die Funktion muss im Lebesgue-Sinne integrierbar sein, d.h. Positiv- und Negativteil müssen ein endliches Integral haben.

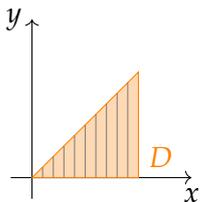
Bei der praktischen Anwendung des Satzes von Fubini muss man oft die Integrationsgrenzen sorgfältig ermitteln:

Beispiel. Wir wollen eine Funktion f über das Dreieck $D = \{(x, y) \in \mathbb{R}^2 \mid 0 \leq x \leq 1, 0 \leq y \leq x\}$ berechnen.

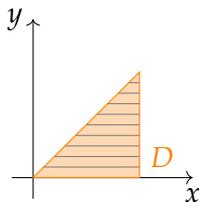


Um dies als iteriertes Integral zu schreiben, haben wir zwei Möglichkeiten, nämlich indem wir zuerst über x und dann über y integrieren, oder anders herum. Für die iterierten Integrale müssen wir untersuchen, über welche y wir für festes x integrieren müssen (bzw. anders herum)

$$\begin{aligned}\int_D f(x, y) d\lambda^2(x, y) &= \int_0^1 \int_0^x f(x, y) dy dx \\ &= \int_0^1 \int_y^1 f(x, y) dx dy.\end{aligned}$$



erst über y , dann über x



erst über x , dann über y

17 Der Transformationssatz

Der Transformationssatz ist die Verallgemeinerung der Substitutionsregel. Er erlaubt es, Integrale von Funktionen, die auf krummlinig berandeten Mengen definiert sind, in Integrale auf einfacheren Gebieten (z.B. Rechtecken) zurückzuführen.

Sind $U, V \subset \mathbb{R}^d$ offen, so heißt ein bijektives $\Phi : U \rightarrow V$ C^1 -invertierbar, wenn sowohl Φ , als auch $\Phi^{-1} : V \rightarrow U$ stetig differenzierbar sind. Die Jacobimatrix von Φ in einem Punkt $x \in U$ ist

$$D\Phi(x) = \begin{pmatrix} \frac{\partial \Phi_1}{\partial x_1} & \dots & \frac{\partial \Phi_1}{\partial x_d} \\ \vdots & & \vdots \\ \frac{\partial \Phi_d}{\partial x_1} & \dots & \frac{\partial \Phi_d}{\partial x_d} \end{pmatrix}$$

und die Jacobimatrix der Umkehrfunktion erfüllt nach dem Umkehrsatz

$$[D\Phi(a)]^{-1} = D\Phi^{-1}(\Phi(a)).$$

Die Bälle bezüglich der ∞ -Norm sind Würfel und wir benutzen die Bezeichnung

$$W_\epsilon(a) := \{x \mid \|x - a\|_\infty \leq \epsilon\}$$

für die abgeschlossenen Würfel.

Um die Transformationsformel zu zeigen, zeigen wir zuerst zwei technische Lemmas.

Zuerst ein Lemma, was angibt, wie sehr sich Würfel beim Abbilden verzerren können, wenn die Ableitung der Abbildung nahe an der Einheitsmatrix liegt:

Lemma 17.1. *Es seien $U, V \subset \mathbb{R}^d$ offen, $F : U \rightarrow V$ C^1 -invertierbar mit Umkehrabbildung $G = F^{-1} : V \rightarrow U$. Weiterhin sei $x_0 \in U$, $y_0 = F(x_0)$ und $r > 0$, so dass $W_r(x_0) \subset U$ und $W_r(y_0) \subset V$ und es gebe $\epsilon \in]0, 1[$, mit*

$$\sup_{x \in W_r(x_0)} \|DF(x) - E_d\| \leq \epsilon, \quad \text{und} \quad \sup_{y \in W_r(y_0)} \|DG(y) - E_d\| \leq \epsilon.$$

Sind dann $\delta \leq r/2$ und $a \in U$ so, dass $x_0 \in W_\delta(a)$ gilt, so gilt für $b = F(a) \in V$:

$$W_{(1-\epsilon)\delta}(b) \subset F(W_\delta(a)) \subset W_{(1+\epsilon)\delta}(b).$$

Beweis.

Ohne Einschränkung sei $x_0 = y_0 = 0$. Da $0 \in W_\delta(a)$ liegt, gilt $\|a\|_\infty \leq \delta \leq r/2$ und also $W_\delta(a) \subset W_r(0)$. Ist jetzt $y \in F(W_\delta(a))$, dann gibt es $x \in W_\delta(a)$ mit $F(x) = y$. Mit dem Mittelwertsatz folgt

$$\|y - b\|_\infty = \|F(x) - F(a)\|_\infty \leq \|DF(\xi)\| \|x - a\|_\infty$$

Hier ein paar Erinnerungen an Begriffe, die wir in diesem Abschnitt benutzen werden: Für $x \in \mathbb{R}^d$ ist $\|x\|_\infty = \max(|x_1|, \dots, |x_d|)$ und es gilt $\frac{1}{\sqrt{d}}\|x\|_2 \leq \|x\|_\infty \leq \|x\|_2$. Die ∞ -Norm induziert die Zeilensummennorm für Matrizen $A = (a_{ij}) \in \mathbb{R}^{d \times d}$

$$\|A\| := \|A\|_{\infty \rightarrow \infty} = \max_i \sum_j |a_{ij}|$$

und es gilt

$$\|A + B\| \leq \|A\| + \|B\|,$$

$$\|AB\| \leq \|A\| \|B\|,$$

$$\|Ax\|_\infty \leq \|A\| \|x\|_\infty.$$

Der mehrdimensionale Mittelwertsatz besagt

$$\begin{aligned} & \Phi(x) - \Phi(a) \\ &= \left(\int_0^1 D\Phi(a + t(x-a)) dt \right) \cdot (x-a) \end{aligned}$$

wenn immer die Verbindungsstrecke von a nach x ganz in U liegt.

Es folgt in diesem Fall $\|\Phi(x) - \Phi(a)\|_\infty \leq C \|x - a\|_\infty$ mit $C = \sup_{0 \leq t \leq 1} \|D\Phi(a + t(x-a))\|$.

mit einem ξ zwischen x und a . Dieses ξ ist also auch in $W_r(0)$ und daher gilt $\|DF(\xi)\| \leq \|DF(\xi) - E_d\| + \|E_d\| \leq 1 + \epsilon$ und es folgt

$$\|y - b\|_\infty \leq (1 + \epsilon)\|x - a\| \leq (1 + \epsilon)\delta.$$

Das heißt aber $y \in W_{(1+\epsilon)\delta}(b)$, also

$$F(W_\delta(a)) \subset W_{(1+\epsilon)\delta}(b).$$

Ebenso sehen wir

$$\|b\|_\infty = \|F(a)\|_\infty \leq (1 + \epsilon)\|a\|_\infty \leq (1 + \epsilon)r/2.$$

Da $\|b\|_\infty \leq (1 + \epsilon)r/2$ gilt, folgt $W_{(1-\epsilon)\delta}(b) \subset W_r(0)$ und d.h., dass für $y \in W_{(1-\epsilon)\delta}(b)$ gilt $\|DG(y)\| \leq 1 + \epsilon$. Es folgt

$$\|G(y) - G(b)\|_\infty \leq (1 + \epsilon)\|y - b\|_\infty \leq (1 + \epsilon)(1 - \epsilon)\delta \leq \delta,$$

d.h. $G(W_{(1-\epsilon)\delta}(b)) \subset W_\delta(a)$. Wenden wir F darauf an, folgt

$$W_{(1-\epsilon)\delta}(b) \subset F(W_\delta(a)).$$

□

Lemma 17.2. *Es seien $U, V \subset \mathbb{R}^d$ offen und $\Phi : U \rightarrow V$ eine C^1 -invertierbare Abbildung. Weiter sei $x_0 \in U$ und $T := D\Phi(x_0)$. Dann gilt: Zu jedem $\epsilon > 0$ existiert ein $\delta > 0$, so dass für jeden achsenparallelen Würfel W mit Seitenlänge δ und $x_0 \in W \subset U$ gilt*

$$\left| \frac{\lambda^d(\Phi(W))}{\lambda^d(W)} - |\det(T)| \right| \leq \epsilon.$$

Beweis.

Ohne Einschränkung nehmen wir $x_0 = 0$ und $\Phi(x_0) = 0$ an. Dann ist $D\Phi(0) = T$ und für die Umkehrabbildung $\Psi = \Phi^{-1}$ gilt $D\Psi = T^{-1}$.

Wir betrachten die Abbildung $F := T^{-1} \circ \Phi : U \rightarrow V_1$ mit $V_1 = T^{-1}V$. Für dieses F gilt dann

$$DF(x) = T^{-1}D\Phi(x) \text{ und } DF(0) = E_d.$$

Ebenso gilt für die Umkehrabbildung $G := F^{-1} = \Psi \circ T : V_1 \rightarrow U$

$$DG(y) = D\Psi(Ty)T \text{ und } DG(0) = E_d.$$

Sei $\epsilon > 0$. Wir setzen $\epsilon_1 = \epsilon/|\det(T)|$ und wählen $\epsilon_2 > 0$ so klein, dass

$$|(1 \pm \epsilon_2)^d - 1| \leq \epsilon_1$$

gilt. Da DF und DG stetig sind, gibt es $\delta > 0$, so dass $W_\delta(0) \subset U$, $W_\delta(0) \subset V_1$ und

$$\sup_{\|x\|_\infty \leq \delta} \|DF(x) - E_d\| \leq \epsilon_2, \quad \sup_{\|y\|_\infty \leq \delta} \|DG(y) - E_d\| \leq \epsilon_2.$$

Nach dem vorigen Lemma folgt daraus für jeden Würfel $W \subset U$ mit Seitenlänge δ und $0 \in W$:

$$(1 - \epsilon_2)^d \leq \frac{\lambda^d(F(W))}{\lambda^d(W)} \leq (1 + \epsilon_2)^d,$$

also

$$\left| \frac{\lambda^d(F(W))}{\lambda^d(W)} - 1 \right| \leq \epsilon_1.$$

Da $F(W) = T^{-1}\Phi(W)$, gilt nach Satz 15.6 $\lambda^d(F(W)) = |\det(T)|^{-1}\lambda^d(\Phi(W))$ und es folgt

$$\left| \frac{\lambda^d(\Phi(W))}{\lambda^d(W)} - 1 |\det(T)| \right| \leq \epsilon_1 |\det(T)| = \epsilon.$$

□

Damit können wir die Transformationsformel für stetige Funktionen zeigen:

Satz 17.3. *Es seien $U, V \subset \mathbb{R}^d$ offen und $\Phi : U \rightarrow V$ C^1 -invertierbar. Dann gilt für jedes stetige $f : V \rightarrow \mathbb{R}$ mit kompaktem Träger*

$$\int_V f(y) d\lambda^d(y) = \int_U f(\Phi(x)) |\det(D\Phi(x))| d\lambda^d(x).$$

Beweis.

Da der Träger von $f \circ \Phi$ in einer endlichen Vereinigung von kompakten Würfeln enthalten ist, reicht es, die Behauptung im Fall von $U = W, V = \Phi(W)$ mit kompaktem Würfel W zu zeigen.

Dazu definieren wir für jeden kompakten Würfel $Q \subset W$

$$\Delta(Q) := \int_{\Phi(Q)} F(y) d\lambda^d(y) - \int_Q F(\Phi(x)) |\det(D\Phi(x))| d\lambda^d(x)$$

und wollen zeigen, dass $\Delta(W) = 0$ gilt.

Nehmen wir an, dass dies nicht gilt. Dann gibt es ein $\alpha > 0$, so dass $|\Delta(W)| = \alpha \lambda^d(W)$. Hat W die Seitenlänge r , so unterteilen wir W durch Halbierung aller Seiten in 2^d kompakte Würfel $W^{(j)}$ mit Seitenlänge $r/2$ (und diese sind, bis auf die Ränder, disjunkt). Da

$$\Delta(W) = \sum_{j=1}^{2^d} \Delta(W^{(j)})$$

muss dann für mindestens einen Teilwürfel $W_1 = W^{(j)}$ gelten

$$|\Delta(W_1)| \geq \alpha \lambda^d(W_1).$$

Wir wenden diesen Schritt rekursiv an und bekommen so eine absteigende Folge $W_1 \supset W_2 \cdots$ von Würfeln, wobei W_k die Kantenlänge $2^{-k}r$ hat und für alle k gilt

$$|\Delta(W_k)| \geq \alpha \lambda^d(W_k). \quad (*)$$

Der Schnitt dieser W_k enthält genau einen Punkt x_0 und wir setzen

$$T = D\Phi(x_0), \quad y_0 = \Phi(x_0), \quad c = F(y_0)$$

und definieren $g : W \rightarrow \mathbb{R}$ und $h : \Phi(W) \rightarrow \mathbb{R}$ durch

$$\begin{aligned} g(x) &:= F(\Phi(x))|\det(D\Phi(x))| - F(\Phi(x_0))|\det(D\Phi(x_0))| \\ h(y) &:= F(y) - F(y_0). \end{aligned}$$

Wir bekommen

$$\begin{aligned} \int_{\Phi(W_k)} F(y) d\lambda^d(y) &= \int_{\Phi(W_k)} (h(y) + F(y_0)) d\lambda^d \\ &= \int_{\Phi(W_k)} h(y) d\lambda^d + c\lambda^d(\Phi(W_k)) \end{aligned}$$

und entsprechend

$$\begin{aligned} \int_{W_k} F(\Phi(x))|\det(D\Phi(x))| d\lambda^d(x) &= \int_{W_k} (g(x) + F(\Phi(x_0))|\det(T)|) d\lambda^d(x) \\ &= \int_{W_k} g(x) d\lambda^d(x) + c|\det(T)|\lambda^d(W_k). \end{aligned}$$

Daraus folgt nach Definition der Funktion Δ

$$\Delta(W_k) = \int_{\Phi(W_k)} h(y) d\lambda^d(y) - \int_{W_k} g(x) d\lambda^d(x) + c\left(\lambda^d(\Phi(W_k)) - |\det(T)|\lambda^d(W_k)\right)$$

und dies ergibt

$$\frac{\Delta(W_k)}{\lambda^d(W_k)} = \frac{1}{\lambda^d(W_k)} \int_{\Phi(W_k)} h(y) d\lambda^d(y) - \frac{1}{\lambda^d(W_k)} \int_{W_k} g(x) d\lambda^d(x) + c\left(\frac{\lambda^d(\Phi(W_k))}{\lambda^d(W_k)} - |\det(T)|\right)$$

Für $k \rightarrow \infty$ gilt für $x \in W_k$ und $y \in \Phi(W_k)$ immer $y \rightarrow y_0$ und $x \rightarrow x_0$ und per Definition gilt $\lim_{y \rightarrow y_0} h(y) = 0$ und $\lim_{x \rightarrow x_0} g(x) = 0$, also gehen die ersten beiden Terme auf der rechten Seite gegen Null. Die letzte Term geht wegen Lemma 17.2 gegen Null für $k \rightarrow \infty$. Das ist aber Widerspruch zu (*), was den Satz beweist. \square

Wir erweitern das Ergebnis aus dem vorigen Satz von stetigen Funktionen auf integrierbare Funktionen:

Satz 17.4. *Es seien $U, V \subset \mathbb{R}^d$ offen und $\Phi : U \rightarrow V$ C^1 -invertierbar. Dann gilt: $f : V \rightarrow \mathbb{R}$ ist genau dann integrierbar, wenn die Funktion $(f \circ \Phi)|\det(D\Phi)| : U \rightarrow \mathbb{R}$ integrierbar ist, und es gilt*

$$\int_U f(\Phi(x))|\det(D\Phi(x))| d\lambda^d(x) = \int_V f(y) d\lambda^d(y).$$

Es reicht, sich zu merken, dass man für $y = \Phi(x)$ im Integral $d\lambda^d(y) = |\det(D\Phi(x))| d\lambda^d(x)$ ersetzen muss (und natürlich den Integrationsbereich entsprechend anpassen).

Beweis.

1. Sei zuerst $f : V \rightarrow \mathbb{R}$ integrierbar, dann ist f messbar und ebenso ist $g := (f \circ \Phi)|\det(D\Phi)| : U \rightarrow \mathbb{R}$ messbar. Nach Satz 14.3 können wir f mit stetigen Funktionen mit kompaktem Träger approximieren, d.h. es existiert eine Folge $f_k \in C_c(V)$, so dass $\|f - f_k\|_{L^1} \rightarrow 0$ gilt. Nach Satz 14.6 gibt es dann eine Teilfolge von f_k (wiederum mit f_k bezeichnet), welche fast überall gegen f konvergiert. Entsprechend konvergiert $g_k := (f_k \circ \Phi)|\det(D\Phi)|$ fast überall gegen g . Satz 17.3 zeigt

$$\int_U g_k(x) d\lambda^d(x) = \int_V f_k(y) d\lambda^d(y) \quad (*)$$

und wegen $\|g_k - g_l\|_{L^1} = \|f_k - f_l\|_{L^1}$ ist g_k eine Cauchy-Folge in L^1 und konvergiert also gegen ein $G \in L^1(U)$. Da aber eine Teilfolge von (g_k) punktweise fast überall gegen g konvergiert, muss $G = g$ f.ü. gelten. Damit folgt die Behauptung durch Grenzübergang in (*).

2. Für die Rückrichtung, vertauscht man die Rollen von f und g , d.h. man betrachtet die Umkehrabbildung $\Psi = \Phi^{-1}$ und wegen $(D\Phi) \circ \Psi = (D\Psi)^{-1}$ folgt

$$(g \circ \Psi)|\det(D\Psi)| = (f \circ \Phi \circ \Psi)|\det(D\Phi \circ \Psi)||\det(D\Psi)| = f$$

und wir können den ersten Teil anwenden.

□

18 Anwendungen der Transformationsformel

Zuerst ein Beispiel mit einer linearen Transformation:

Beispiel. Zu einer symmetrischen, positiv definiten Matrix $C \in \mathbb{R}^{n \times n}$ betrachten wir das Integral

$$\int_{\mathbb{R}^n} \exp(-x^T C x) d\lambda(x).$$

Nach dem Spektralsatz für symmetrische Matrix schreiben wir C als $C = V^T D V$ mit orthonormalem V und diagonalem D welches nur positive Werte auf der Diagonalen hat. Bezeichnen wir mit $D^{1/2}$ die Matrix, die die Wurzeln der Diagonaleinträge von D auf der Diagonalen hat, können wir schreiben $C = (D^{1/2} V)^T (D^{1/2} V)$. Setzen wir $U = V^T D^{-1/2}$, so gibt das $U^T C U = I$.

Betrachten wir die lineare Abbildung $U : \mathbb{R}^n \rightarrow \mathbb{R}^n$, so gilt nach dem Transformationssatz (wir "substituieren $x = Uy, d\lambda(x) = U d\lambda(y)$ ")

$$\begin{aligned} \int_{\mathbb{R}^n} \exp(-x^T C x) d\lambda(x) &= \int_{\mathbb{R}^n} \exp(-(Uy)^T C Uy) |\det(U)| d\lambda(y) \\ &= |\det(U)| \int_{\mathbb{R}^n} \exp(-\|y\|_2^2) d\lambda(y). \end{aligned}$$

Es gilt $1 = \det(U^T C U) = \det(U^T) \det(C) \det(U)$, woraus folgt $\det(U) = 1/\sqrt{\det(C)}$. Mit dem Satz von Fubini und dem Wissen $\int_{-\mathbb{R}} \exp(-x^2) dx = \sqrt{\pi}$ folgt

$$\begin{aligned} \int_{\mathbb{R}^n} \exp(-x^T C x) d\lambda(x) &= \frac{1}{\sqrt{\det(C)}} \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \exp\left(-\sum_{i=1}^n y_i^2\right) dy_1 \cdots dy_n \\ &= \frac{1}{\sqrt{\det(C)}} \left(\int_{\mathbb{R}} \exp(-t^2) dt \right)^n = \frac{\sqrt{\pi}^n}{\sqrt{\det(C)}}. \end{aligned}$$

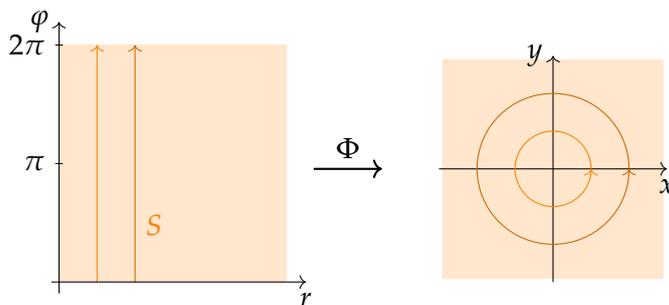
Betrachten wir jetzt Polarkoordinaten in der Ebene: Einen Punkt $(x, y) \in \mathbb{R}^2$ ($x, y \neq 0$) können wir mit Hilfe von Abstand $r = \sqrt{x^2 + y^2}$ vom Nullpunkt und Winkel φ darstellen als

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

und der Winkel ist bis auf ein ganzzahliges Vielfaches von 2π eindeutig bestimmt. Polarkoordinaten werden durch die Abbildung

$$\Phi : [0, \infty[\times [0, 2\pi] \rightarrow \mathbb{R}^2, \quad (r, \varphi) \mapsto (r \cos(\varphi), r \sin(\varphi))$$

beschrieben.



Die Jacobi-Matrix von Φ ist

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

und es folgt

$$\det(D\Phi(r, \varphi)) = r.$$

Aus der Transformationsformel bekommen wir also:

Satz 18.1. Eine Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ ist genau dann integrierbar, wenn die Funktion

$$g : [0, \infty[\times [0, 2\pi] \rightarrow \mathbb{R}, g(r, \varphi) = rf(\Phi(r, \varphi))$$

integrierbar ist, und es gilt

$$\int_{\mathbb{R}^2} f(x, y) d\lambda^2(x, y) = \int_{[0, 2\pi]} \int_{[0, \infty[} f(r \cos(\varphi), r \sin(\varphi)) r d\lambda^1(r) d\lambda^1(\varphi).$$

Beispiel. Mit Hilfe von Polarkoordinaten kann man das Gaußsche Integral sehr einfach berechnen: Wir setzen $I = \int_{-\infty}^{\infty} \exp(-x^2/2) dx$ und betrachten wieder das Quadrat, was wir als ein Integral über den \mathbb{R}^2 betrachten

$$\begin{aligned} I^2 &= \int_{\mathbb{R}} \exp(-x^2/2) dx \int_{\mathbb{R}} \exp(-y^2/2) dy \\ &= \int_{\mathbb{R}^2} \exp(-x^2/2) \exp(-y^2/2) dx dy \\ &= \int_{\mathbb{R}^2} \exp(-(x^2 + y^2)/2) dx dy. \end{aligned}$$

Dies Integral berechnen wir nun mit Hilfe von Polarkoordinaten: Mit $r = \sqrt{x^2 + y^2}$ ist nämlich $\exp(-(x^2 + y^2)/2) = \exp(-r^2/2)$ und daher ist

$$I^2 = \int_0^{2\pi} \int_0^{\infty} \exp(-r^2/2) r dr d\varphi.$$

Eine Stammfunktion von $r \exp(-r^2/2)$ ist schnell gefunden, denn es gilt $(\exp(-r^2/2))' = -r \exp(-r^2/2)$. Es folgt

$$\begin{aligned} I^2 &= \int_0^{2\pi} [-\exp(-r^2/2)]_{r=0}^{\infty} d\varphi \\ &= \int_0^{2\pi} 1 d\varphi = 2\pi \end{aligned}$$

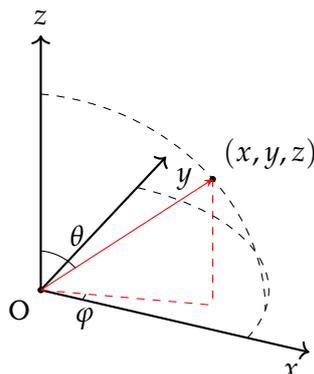
und wir bekommen

$$\int_{\mathbb{R}} \exp(-x^2/2) dx = I = \sqrt{2\pi}.$$

Beispiel (Polar-Koordinaten im \mathbb{R}^3). Wir betrachten die Abbildung

$$f :]0, \infty[\times]0, 2\pi[\times]0, \pi[\rightarrow \mathbb{R}^3, \quad f(r, \varphi, \theta) = \begin{bmatrix} r \cos(\varphi) \sin(\theta) \\ r \sin(\varphi) \sin(\theta) \\ r \cos(\theta) \end{bmatrix}.$$

Dabei ist φ der Winkel in der x, y -Ebene und θ der Winkel von der z -Achse. (Man könnte auch sagen: φ ist der Längengrad und θ ist der Breitengrad, wobei $\theta = 0$ dem Nordpol entspricht, $\theta = \pi/2$ (also 90°) dem Äquator entspricht und $\theta = \pi$ dem Südpol entspricht.)



Diese Abbildung bildet das Rechteck $]0, \infty[\times]0, 2\pi[\times]0, \pi[\subset \mathbb{R}^3$ bijektiv auf den $\mathbb{R}^3 \setminus \{x \in \mathbb{R}^3 \mid x_1 = x_2 = 0\}$ ab. Diese Abbildung ist differenzierbar, und die Jacobi-Matrix ist

$$Df(r, \varphi, \theta) = \begin{bmatrix} \cos(\varphi) \sin(\theta) & -r \sin(\varphi) \sin(\theta) & r \cos(\varphi) \cos(\theta) \\ \sin(\varphi) \sin(\theta) & r \cos(\varphi) \sin(\theta) & r \sin(\varphi) \cos(\theta) \\ \cos(\theta) & 0 & -r \sin(\theta) \end{bmatrix}$$

und ihre Determinante ist

$$\begin{aligned} \det(Df(r, \varphi, \theta)) &= r^2 \left(\cos(\theta) \det \begin{bmatrix} -\sin(\varphi) \sin(\theta) & \cos(\varphi) \cos(\theta) \\ \cos(\varphi) \sin(\theta) & \sin(\varphi) \cos(\theta) \end{bmatrix} \right. \\ &\quad \left. - \sin(\theta) \det \begin{bmatrix} \cos(\varphi) \sin(\theta) & -\sin(\varphi) \sin(\theta) \\ \sin(\varphi) \sin(\theta) & \cos(\varphi) \sin(\theta) \end{bmatrix} \right) \\ &= r^2 \left((\cos(\theta) (-\sin(\varphi)^2 \sin(\theta) \cos(\theta) - \cos(\varphi)^2 \sin(\theta) \cos(\theta)) \right. \\ &\quad \left. - \sin(\theta) (\cos(\varphi)^2 \sin(\theta)^2 + \sin(\varphi)^2 \sin(\theta)^2) \right) \\ &= r^2 \left(-\cos(\theta)^2 \sin(\theta) - \sin(\theta)^2 \sin(\theta) \right) \\ &= -r^2 \sin(\theta). \end{aligned}$$

Da $\sin(\theta) > 0$ für $\theta \in]0, \pi[$ folgt

$$|\det(Df(r, \varphi, \theta))| = r^2 \sin(\theta).$$

Damit können wir das Volumen einer Kugel von Radius $a > 0$ wie folgt berechnen: Es gilt

$$f : \begin{array}{ccc} M & \longrightarrow & B_a(0) \\]0, a[\times]0, 2\pi[\times]0, \pi[& & \{x \mid \|x\|_2 < a\} \end{array}$$

Das Minuszeichen scheint vielleicht seltsam. Es verschwindet z.B., wenn man die Reihenfolge der Variablen so ändert, dass die letzten beiden Spalten der Jacobi-Matrix vertauscht werden.

und die Abbildung C^1 -invertierbar. Mit dem Satz von Fubini und Satz 17.4 folgt

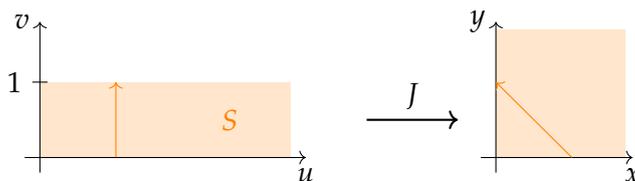
$$\begin{aligned}\lambda^3(B_a(0)) &= \lambda^3(f(M)) = \int_{f(M)} 1 d\lambda^3 = \int_M |\det(Df)| d\lambda^3 \\ &= \int_0^a \int_0^{2\pi} \int_0^\pi r^2 \sin(\theta) d\theta d\varphi dr \\ &= \underbrace{\int_0^a r^2 dr}_{=a^3/3} \underbrace{\int_0^{2\pi} d\varphi}_{=2\pi} \underbrace{\int_0^\pi \sin(\theta) d\theta}_{=-\cos(\theta)|_0^\pi=2} = \frac{4\pi}{3} a^3.\end{aligned}$$

Hier eine weitere Transformation, die manchmal hilfreich ist:

Beispiel. Wir betrachten die Abbildung

$$J :]0, \infty[\times]0, 1[\rightarrow \mathbb{R}^2, \quad (u, v) \mapsto (u(1-v), uv).$$

Die Abbildung bildet die Linie $\{(u, v) \mid 0 < v < 1\}$ welche die Punkte $(u, 0)$ und $(u, 1)$ verbindet auf die Linie ab, welche $(u, 0)$ und $(0, u)$ verbindet. Außerdem wird die Menge $S =]0, \infty[\times]0, 1[$ bijektiv auf $]0, \infty[^2$ abgebildet:



Ist $J(u, v) = (x, y)$, so gilt $u(1-v) = x$ und $uv = y$ und es folgt $x + y = u$. Die Umkehrfunktion von J ist

$$J^{-1}(x, y) = (x + y, y/(x + y)).$$

Die Ableitung von J und ihre Determinante sind

$$DJ(u, v) = \begin{pmatrix} 1-v & -u \\ v & u \end{pmatrix}, \quad \det(DJ(u, v)) = (1-v)u + uv = u.$$

Um eine Anwendung zu zeigen, erinnern wir die Eulersche Γ -Funktion (vgl. Abschnitt 12): Für $s > 0$ ist

$$\Gamma(s) = \int_0^\infty x^{s-1} \exp(-x) dx.$$

Wir wollen nun zeigen, dass für $p, q > 0$ gilt

$$\int_0^1 (1-t)^{p-1} t^{q-1} dt = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}.$$

Man nennt das Integral auch das *Eulersche Beta-Integral*.

Dazu rechnen wir unter Benutzung des Satzes von Fubini und der obigen Substitution

$$\begin{aligned}\Gamma(p)\Gamma(q) &= \int_0^\infty x^{p-1} \exp(-x) dx \int_0^\infty y^{q-1} \exp(-y) dy \\ &= \int_0^\infty \int_0^\infty x^{p-1} y^{q-1} \exp(-(x+y)) dx dy \\ &= \int_0^1 \int_0^\infty u^{p-1} (1-v)^{p-1} u^{q-1} v^{q-1} \exp(-u) u du dv \\ &= \int_0^1 (1-v)^{p-1} v^{q-1} dv \int_0^\infty u^{p+q-1} \exp(-u) du \\ &= \int_0^1 (1-v)^{p-1} v^{q-1} dv \Gamma(p+q)\end{aligned}$$

was die Behauptung zeigt.

19 Darstellung von k -dimensionalen Flächen

Eine Kugeloberfläche im \mathbb{R}^3 lässt sich auf mindestens vier verschiedene Arten mit Hilfe von glatten Funktionen darstellen:

1. **Als Nullstellengebilde einer Funktion $f : \mathbb{R}^3 \rightarrow \mathbb{R}$:** Es sei $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ gegeben durch

$$f(x) = \|x\|_2^2 = x_1^2 + x_2^2 + x_3^2.$$

Dann ist die Oberfläche der Einheitskugel

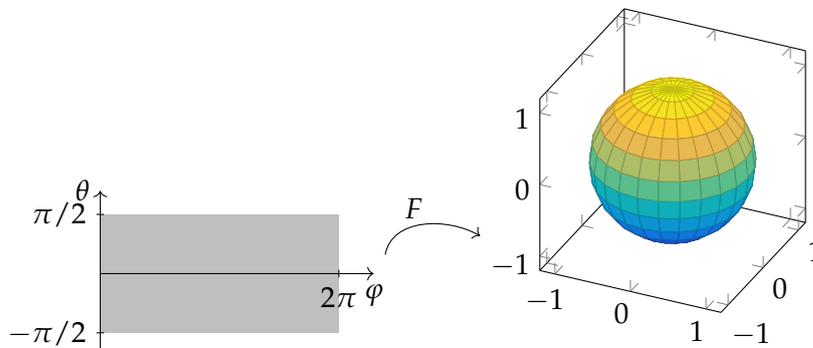
$$K = \{x \in \mathbb{R}^3 \mid f(x) = 1\}.$$

2. **Durch eine Parametrisierung " $F : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ ":** Ein Punkt auf der Oberfläche der Einheitskugel lässt sich durch Angabe von Längen- und Breitengrad beschreiben: Es sei $F : [0, 2\pi[\times]-\pi/2, \pi/2[\rightarrow \mathbb{R}^3$ gegeben durch

$$F(\varphi, \theta) = \begin{bmatrix} \cos(\varphi) \cos(\theta) \\ \sin(\varphi) \cos(\theta) \\ \sin(\theta) \end{bmatrix}.$$

Dann ist die Kugeloberfläche (mit Ausnahme von Nord- und Südpol)

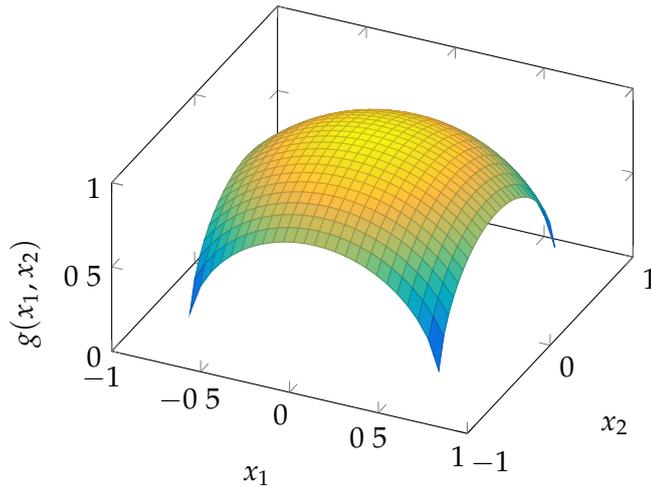
$$K = F([0, 2\pi[\times]-\pi/2, \pi/2[).$$



3. **Lokal als Graphen einer Funktion " $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ ":** Die Nordhalbkugel ist der Graph der Funktion $g : \{x \in \mathbb{R}^2 \mid \|x\|_2^2 \leq 1\} \rightarrow \mathbb{R}$ gegeben durch

$$g(x) = \sqrt{1 - \|x\|_2^2}.$$

Für andere Teile der Kugeloberfläche kann man entsprechende Funktionen finden (und muss ggf. die Koordinaten umnummerieren, z.B. wenn man "die westliche Hemisphäre" beschreiben will). Hier der Graph auf dem Definitionsgebiet $[-1/\sqrt{2}, 1/\sqrt{2}]^2$:



4. **Durch lokales “flachmachen”, also durch Verformen auf eine Ebene:** Eine Umgebung eines Punktes kann so verformt werden, dass die Kugeloberfläche der Umgebung dabei auf eine Koordinatenebene abgebildet wird. Dies geht bei der Kugel und dem Nordpol $(0,0,1)$ z.B. wie folgt: Ist U eine geeignete Umgebung des Nordpols, so definieren wir

$$\Phi : U \rightarrow \mathbb{R}^3, \quad \Phi(x_1, x_2, x_3) = \begin{bmatrix} x_1 \\ x_2 \\ \sqrt{1 - x_1^2 - x_2^2 - x_3^2} \end{bmatrix}.$$

Dann ist das Bild von U in der Ebene mit $x_3 = 0$ enthalten.

Wir wollen nun diese Möglichkeiten zur Darstellung von Oberflächen mathematisch präzise definieren und zeigen, dass diese bei richtiger Formulierung den gleichen Begriff für eine k -dimensionale Oberfläche ergeben. Wir beginnen mit einem vorbereitenden Begriff:

Definition 19.1. Es sei $T \subset \mathbb{R}^k$ offen. Eine C^α Funktion $\varphi : T \rightarrow \mathbb{R}^n$, $(t_1, \dots, t_k) \mapsto \varphi(t_1, \dots, t_k)$ heißt C^α -Immersion, falls die Jacobi-Matrix

$$D\varphi = \begin{bmatrix} \frac{\partial \varphi_1}{\partial t_1} & \cdots & \frac{\partial \varphi_1}{\partial t_k} \\ \vdots & & \vdots \\ \frac{\partial \varphi_n}{\partial t_1} & \cdots & \frac{\partial \varphi_n}{\partial t_k} \end{bmatrix} \in \mathbb{R}^{n \times k}$$

in jedem Punkt $t \in T$ den Rang k hat.

Bemerkung. i) In diesem Zusammenhang nennt man $D\varphi$ auch *Funktionalmatrix* und schreibt auch

$$\frac{\partial(\varphi_1, \dots, \varphi_n)}{\partial(t_1, \dots, t_k)} := D\varphi.$$

- ii) Da $\frac{\partial(\varphi_1, \dots, \varphi_n)}{\partial(t_1, \dots, t_k)} \in \mathbb{R}^{n \times k}$, muss $n \geq k$ gelten, wenn $D\varphi$ den Rang k haben soll.

iii) Die Bedingung $\text{rang}(D\varphi(t)) = k$ heißt nichts anderes, als dass die Vektoren

$$\frac{\partial \varphi}{\partial t_1}(t), \dots, \frac{\partial \varphi}{\partial t_k}(t) \in \mathbb{R}^n$$

linear unabhängig sind.

iv) Die Bedingung $\text{rang}(D\varphi)(t) = k$ lässt sich mit Hilfe von Determinanten wie folgt ausdrücken: Es gilt $\text{rang}(D\varphi)(t) = k$, falls es Indizes $1 \leq i_1 < i_2 < \dots < i_k \leq n$ gibt, so dass

$$\det \frac{\partial(\varphi_{i_1}, \dots, \varphi_{i_k})}{\partial(t_1, \dots, t_k)} \neq 0.$$

Da die Determinante stetig ist, folgt daraus: Hat $D\varphi(t^*)$ Vollrang, so gilt dies auch in einer Umgebung von t^* .

- v) Im Spezialfall $T \subset \mathbb{R}$ (also $k = 1$) ist eine Immersion nichts anderes als eine Kurve im \mathbb{R}^n die nirgends mit Geschwindigkeit Null durchlaufen wird.
- vi) Im anderen Extremfall $k = n$ ist φ nach dem Umkehrsatz lokal invertierbar und stellt also lokal nur eine "Umparametrisierung" des \mathbb{R}^n dar. ■

Wir erinnern kurz an den Begriff des Diffeomorphismus:

Definition 19.2. Sind $U, V \in \mathbb{R}^n$ offen, so heißt $\Phi : U \rightarrow V$ ein C^k -Diffeomorphismus, wenn Φ bijektiv ist und sowohl Φ , als auch Φ^{-1} C^k -Abbildungen sind.

Nun übertragen wir die topologischen Begriffe, die wir im \mathbb{R}^n definiert haben, auf Teilmengen des \mathbb{R}^n , indem wir die von der Standard-Metrik induzierte Metrik und damit auch die entsprechende Teilraumtopologie betrachten:

- Da (d, \mathbb{R}^n) ein metrischer Raum mit der euklidischen Metrik $d(x, y) = \|x - y\|_2$ ist, können wir auf M die induzierte Metrik betrachten.
- Die offenen Bälle in M um $a \in M$ sind dann genau von der Form $B_\epsilon(a) \cap M$.
- Wir sagen $V \subset M$ ist *offen relativ M* , wenn es zu jedem $a \in V$ ein $\epsilon > 0$ gibt, so dass $B_\epsilon(a) \cap M \subset V$.
- Teilmengen von M , die relativ M offen sind, sind im Allgemeinen nicht in \mathbb{R}^n offen! Ist jedoch $U \subset \mathbb{R}^n$ offen, dann ist $U \cap M$ offen relativ M .
- Man kann zeigen, dass sogar alle relativ M offenen Mengen von der Form $U \cap M$ mit offenen $U \subset \mathbb{R}^n$ sind. Man folgert hieraus, dass $K \subset M \subset \mathbb{R}^n$ kompakt relativ M ist, genau dann, wenn K kompakt in \mathbb{R}^n ist.

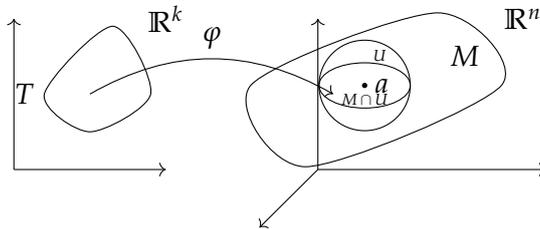
Definition 19.3. Eine Abbildung $\varphi : X \rightarrow Y$ zwischen metrischen Räumen heißt *Homöomorphismus*, wenn φ stetig und umkehrbar ist, und φ^{-1} ebenfalls stetig ist.

Kommen wir nun zum zentralen Begriff, der für uns als Definition von “ k -dimensionaler Fläche” fungieren wird:

Definition 19.4. Eine Menge $M \subset \mathbb{R}^n$ heißt *k -dimensionale Untermannigfaltigkeit des \mathbb{R}^n* der Klasse C^α , wenn es zu jedem $a \in M$ offene Umgebung $U \subset \mathbb{R}^n$ gibt, sowie eine offene Teilmenge $T \subset \mathbb{R}^k$ und eine C^α -Immersion

$$\varphi : T \rightarrow \mathbb{R}^n,$$

so dass T durch φ auf $\varphi(T) = M \cap U$ homöomorph abgebildet wird.



Die Zahl $n - k$ heißt *Kodimension* der Untermannigfaltigkeit; Untermannigfaltigkeiten der Kodimension 1 nennt man auch *Hyperflächen*.

Definition 19.5. Es sei $M \subset \mathbb{R}^n$ eine k -dimensionale Untermannigfaltigkeit und $V \subset M$ offen relativ M . Eine Immersion $\varphi : T \rightarrow V$ (mit offenem $T \subset \mathbb{R}^k$) welche T homöomorph auf V abbildet heißt dann *lokale Parameterdarstellung* von M oder auch *lokale Karte* von M .

Hat jeder Punkt von M eine Umgebung, so dass für die Umgebung ein Karte der Klasse C^α existiert, so nennen wir M eine *Untermannigfaltigkeit der Klasse C^α* (oder auch C^α -Untermannigfaltigkeit).

Wenn man zwei Karten $\varphi_i : T_i \rightarrow V_i, i = 1,2$ gegeben hat, und V_1 und V_2 sich schneiden, so hat man in $V_1 \cap V_2$ zwei verschiedene Parameterdarstellungen. Man kann beide ineinander umrechnen:

Satz 19.6 (und Definition). *Es sei $M \subset \mathbb{R}^n$ eine k -dimensionale Untermannigfaltigkeit der Klasse C^α und*

$$\varphi_i : T_i \rightarrow V_i \quad (i = 1,2)$$

zwei Karten welche C^α -Diffeomorphismen sind und sei $V := V_1 \cap V_2 \neq \emptyset$. Dann sind die Mengen $W_i = \varphi_i^{-1}(V) \subset \mathbb{R}^k$ offene Teilmengen der T_i und die Abbildungen

$$\tau_{ij} := \varphi_j^{-1} \circ \varphi_i : W_i \rightarrow W_j$$

sind C^α -Diffeomorphismen. Wir nennen τ_{ij} auch Kartenwechsel oder Parameter-Transformationen.

Man beachte, dass nicht automatisch das Bild einer Immersion immer eine Untermannigfaltigkeit ist: Ist zum Beispiel T ein offenes Intervall und $\varphi(T) \subset \mathbb{R}^2$ eine Kurve, die sich selbst schneidet, so ist dieses Bild *keine* Untermannigfaltigkeit. Die Immersion φ selbst ist dann ja nicht einmal bijektiv auf das Bild und man kann zeigen, dass keine offene Umgebung des Schnittpunktes das homöomorphe Bild eines offenen Intervalls ist.

Der Begriff der Karte kommt von der Analogie zu Landkarten, die die gekrümmte Oberfläche der Erde lokal in rechtwinkligen Koordinaten beschreiben. Tatsächlich nennt man eine Menge von Karten deren Bilder die gesamte Untermannigfaltigkeit überdecken auch einen *Atlas* der Untermannigfaltigkeit.

An anderen Stellen wird die Abbildung $\varphi^{-1} : V \rightarrow T$ (also von der Untermannigfaltigkeit in den Parameterbereich) als Karte bezeichnet .

Der Beweis ist nicht schwierig, aber auch nicht sehr erhellend, siehe z.B. Analysis 3, O. Forster, §14, Satz 5.

Satz 19.7. *Es sei $T \subset \mathbb{R}^k$ offen und $\varphi : T \rightarrow \mathbb{R}^n$ eine Immersion. Dann gibt es zu jedem $t \in T$ eine offene Umgebung $V \subset T$, so dass $\varphi(V)$ eine k -dimensionale Untermannigfaltigkeit ist und*

$$\varphi|_V : V \rightarrow \varphi(V) \subset \mathbb{R}^n$$

ein Homöomorphismus von V auf $\varphi(V)$ ist.

Beweis.

Für $k = n$ ist dies der Umkehrsatz. Betrachten wir also $k < n$ und ein $t \in T$. Dann existieren Indizes $1 \leq i_1 < \dots < i_k \leq n$, so dass

$$\frac{\partial(\varphi_{i_1}, \dots, \varphi_{i_k})}{\partial(t_1, \dots, t_k)}(t)$$

invertierbar ist und ohne Einschränkung nehmen wir $(i_1, \dots, i_k) = (1, \dots, k)$. Für die Abbildung

$$\tilde{\varphi} := (\varphi_1, \dots, \varphi_k) : T \rightarrow \mathbb{R}^k$$

gelten also die Voraussetzungen des Umkehrsatzes und folglich gibt es eine offene Umgebung $V \subset T$ von t und eine offene Umgebung $U \subset \mathbb{R}^k$ von $\tilde{\varphi}(t)$, so dass $\tilde{\varphi} : V \rightarrow U$ ein Homöomorphismus ist.

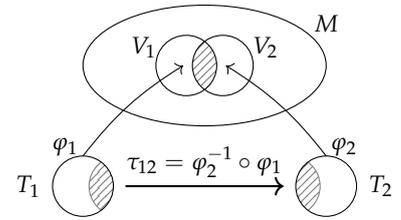
Damit schreiben wir

$$\varphi|_V = (\tilde{\varphi}, \hat{\varphi}) : V \rightarrow U \times \mathbb{R}^{n-k} \subset \mathbb{R}^n$$

mit $\hat{\varphi} = (\varphi_{k+1}, \dots, \varphi_n)$. Da $\tilde{\varphi} : V \rightarrow U$ bijektiv ist, ist auch $\varphi|_V : V \rightarrow \varphi(V) \subset U \times \mathbb{R}^{n-k}$ bijektiv. Dieses $\varphi|_V$ ist sogar ein Homöomorphismus, denn die Umkehrabbildung

$$\psi = (\varphi|_V)^{-1} : \varphi(V) \rightarrow V, \quad \psi(x_1, \dots, x_k, x_{k+1}, \dots, x_n) = (\tilde{\varphi})^{-1}(x_1, \dots, x_k)$$

ist stetig (da $(\tilde{\varphi})^{-1}$ stetig ist). □



20 Untermannigfaltigkeiten des \mathbb{R}^d

Beispiel. 1. Die Kugeloberfläche ist eine 2-dimensionale Untermannigfaltigkeit des \mathbb{R}^3 , denn für die Abbildung

$$F(\varphi, \theta) = \begin{bmatrix} \cos(\varphi) \cos(\theta) \\ \sin(\varphi) \cos(\theta) \\ \sin(\theta) \end{bmatrix}$$

ist

$$DF(\varphi, \theta) = \begin{bmatrix} -\sin(\varphi) \cos(\theta) & -\cos(\varphi) \sin(\theta) \\ \cos(\varphi) \cos(\theta) & -\sin(\varphi) \sin(\theta) \\ 0 & \cos(\theta) \end{bmatrix}$$

immer vom Rang 2.

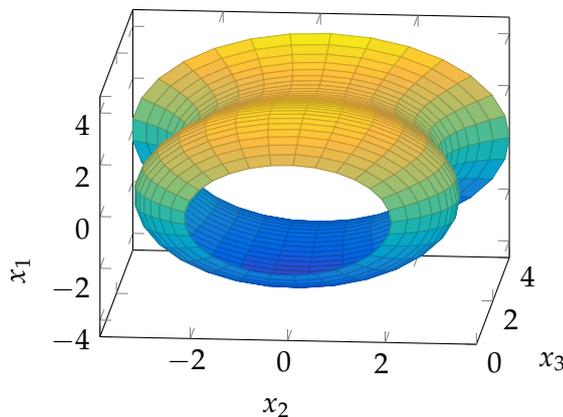
Die Matrix bestehend aus den letzten beiden Zeilen hat immer Vollrang, da $-\pi/2 < \theta < \pi/2$.

2. Ist I ein offenes Intervall, $\alpha : I \rightarrow \mathbb{R}^2$ eine C^1 Kurve, die wir uns als Kurve in der x_1, x_3 -Ebene in \mathbb{R}^3 vorstellen, d.h. $x_1(t) = \alpha_1(t)$, $x_2(t) = 0$, $x_3(t) = \alpha_2(t)$. Außerdem nehmen wir an, dass $\alpha'(t) \neq 0$ für alle t . Rotieren wir diese Kurve um die x_3 -Achse, erhalten wir eine Fläche mit der Parameterdarstellung

$$F : I \times]0, 2\pi[\rightarrow \mathbb{R}^3, \quad (t, \varphi) = \begin{bmatrix} \alpha_1(t) \cos(\varphi) \\ \alpha_1(t) \sin(\varphi) \\ \alpha_2(t) \end{bmatrix}.$$

Wieder ist einfach zu sehen, dass $DF(t, \varphi)$ Rang 2 hat, wenn (z.B.) $\alpha_1(t) \neq 0$ und $\alpha_2'(t) \neq 0$.

Hier das Bild für die Kurve $\alpha(t) = \begin{bmatrix} t \\ (t - \frac{1}{2})(t - 2)(t - 3\frac{1}{2})/4 + 3 \end{bmatrix}$:



Die Kugeloberfläche erhalten wir z.B. als Rotation eines Halbkreises.

Der Begriff *Untermannigfaltigkeit* kommt daher, dass wir a-priori Teilmengen des \mathbb{R}^n betrachten. Unter einer k -dimensionalen (topologischen) Mannigfaltigkeit versteht man einen topologischen Raum M (also einen Raum, auf dem der Begriff der offenen Menge Sinn ergibt), so dass jeder Punkt in M eine Umgebung hat, die homöomorph zu einer offenen Teilmenge des \mathbb{R}^k ist. Hierbei wird auf den umgebenden Raum \mathbb{R}^n verzichtet. Eine erste (gar nicht so einfache) Frage hier ist: Lässt sich jede k -dimensionale Mannigfaltigkeit in den \mathbb{R}^n "einbetten", d.h., ist sie homöomorph zu einer Teilmenge eines \mathbb{R}^n ? Wenn ja, für welche n ? Es sei soviel verraten: Auch für beliebig glatte k -dimensionale Mannigfaltigkeiten reicht nicht immer ein \mathbb{R}^{k+1} , aber immerhin \mathbb{R}^{2k} (siehe dazu die Begriffe "Kleinsche Flasche" und "Whitneyscher Einbettungssatz").

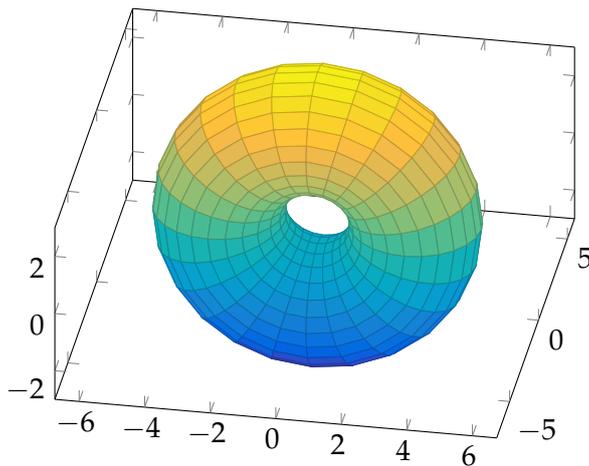
Beispiel. Rotieren wir die Kreislinie

$$t \mapsto \begin{bmatrix} R + r \cos(t) \\ r \sin(t) \end{bmatrix}$$

(für $0 < r < R$) um die x_3 -Achse, bekommen wir

$$F(t, s) = \begin{bmatrix} (R + r \cos(t)) \cos(s) \\ (R + r \cos(t)) \sin(s) \\ r \sin(t) \end{bmatrix}$$

und erhalten also Fläche den sogenannten *Torus*.



Beispiel. Keine Untermannigfaltigkeiten sind zum Beispiel:

i) Der Würfel $W = \{x \in \mathbb{R}^n \mid \|x\|_\infty = 1\}$.

Keine Umgebung der Ecken lässt sich als Bild einer Immersion schreiben (diese muss C^1 sein!).

ii) Das Koordinatenkreuz $\{x \in \mathbb{R}^n \mid \prod_{i=1}^n x_i = 0\}$.

Keine Umgebung des Nullpunktes ist homöomorph zu einer offenen Menge eines \mathbb{R}^k .

Jetzt zeigen wir, dass vier Beschreibungen von Untermannigfaltigkeiten (durch lokale Karten, lokal als Graphen, lokal als Nullstellengebilde und durch lokalen Flachmachen) alle äquivalent sind.

Wir halten dazu noch einmal etwas formeller fest: Eine Menge $M \subset \mathbb{R}^n$ ist

i) *lokal ein Nullstellengebilde*, wenn es zu jedem $a \in M$ eine offene Umgebung $U \subset \mathbb{R}^n$ und $n - k$ Funktionen

$$f_j : U \rightarrow \mathbb{R}, \quad j = 1, \dots, n - k$$

der Klasse C^1 gibt, so dass

$$M \cap U = \{x \in U \mid 0 = f_1(x) = \dots = f_{n-k}(x)\}$$

und für alle $x \in M \cap U$ gilt

$$\text{rang} \frac{\partial(f_1, \dots, f_{n-k})}{\partial(x_1, \dots, x_n)}(x) = n - k.$$

ii) *lokal ein Graph*, wenn es zu jedem $a \in M$ (nach eventueller Umnummerierung der Koordinaten) offene Umgebungen $U' \subset \mathbb{R}^k$ von $a' := (a_1, \dots, a_k)$ und $U'' \subset \mathbb{R}^{n-k}$ von $a'' := (a_{k+1}, \dots, a_n)$ und eine C^1 Abbildung

$$g : U' \rightarrow U''$$

gibt, so dass

$$M \cap (U' \times U'') = \{(x', x'') \in U' \times U'' \mid x'' = g(x')\}.$$

iii) *lokal eine Transformation auf die k -dimensionalen Ebene*, wenn es zu jedem $a \in M$ eine offene Umgebung $U \subset \mathbb{R}^n$ von a , eine offene Menge $V \subset \mathbb{R}^n$ und einen Diffeomorphismus $\Phi : U \rightarrow V$ gibt, so dass für die k -dimensionale Ebene

$$E_k = \{x = (x_1, \dots, x_n) \in \mathbb{R}^n \mid x_{k+1} = \dots = x_n = 0\}$$

gilt, dass

$$\Phi(M \cap U) = E_k \cap V.$$

Satz 20.1. Eine Menge $M \subset \mathbb{R}^n$ ist genau dann eine k -dimensionale Untermannigfaltigkeit, wenn sie lokal ein Nullstellengebilde, lokal ein Graph oder lokal eine Transformation auf eine Ebene ist.

Beweis.

Wir zeigen:

$$\begin{aligned} M \text{ } k\text{-dim. Untermannigfaltigkeit} &\stackrel{1.}{\implies} M \text{ lokal ein Graph} \\ &\stackrel{2.}{\implies} M \text{ lokal Nullstellengebilde} \\ &\stackrel{3.}{\implies} M \text{ lokal Transformation auf eine Ebene} \\ &\stackrel{4.}{\implies} M \text{ } k\text{-dim. Untermannigfaltigkeit} \end{aligned}$$

1. Sei M eine k -dimensionale Untermannigfaltigkeit, $a \in M$ mit entsprechender Umgebung U , $T \subset \mathbb{R}^k$ und $t^* \in T$, $\varphi : T \rightarrow M \cap U$ Immersion und $a = \varphi(t^*)$. Dann gilt ohne Einschränkung (ggf. Umnummerierung der Koordinaten)

$$\det \frac{\partial(\varphi_1, \dots, \varphi_k)}{\partial(t_1, \dots, t_k)} \neq 0.$$

Nach dem Umkehrsatz ist also $\tilde{\varphi} = (\varphi_1, \dots, \varphi_k)$ ein Diffeomorphismus von einer Umgebung $T_1 \subset T$ von t^* auf eine offene Menge $U_1 \subset \mathbb{R}^k$. Es sei $\psi : U_1 \rightarrow T_1$ die Umkehrabbildung von $\tilde{\varphi}$. Dann hat $G := \varphi \circ \psi : U_1 \rightarrow \mathbb{R}^n$ die

Gestalt

$$G(x_1, \dots, x_k) = (x_1, \dots, x_k, g_{k+1}(x'), \dots, g_n(x')), \quad x' = (x_1, \dots, x_k).$$

Für ein geeignetes $U' \subset U_1$ leistet die Abbildung $g = (g_{k+1}, \dots, g_n) : U' \rightarrow \mathbb{R}^{n-k}$ das Gewünschte, denn es gilt mit $x'' = (x_{k+1}, \dots, x_n)$ und $U'' = g(U')$, dass $(x', x'') \in M \cap (U' \times U'')$ wenn

$$g(x') = x''.$$

2. Sei nun M lokal ein Graph und $a \in M$, d.h. M ist in einer Umgebung von a der Graph einer Funktion $g = (g_{k+1}, \dots, g_n)$, d.h. es gilt $a = (x_1, \dots, x_n) \in M$ genau dann, wenn $g_{k+1}(x_1, \dots, x_k) = x_{k+1}, \dots, g_n(x_1, \dots, x_k) = x_n$. Dann ist M dort auch das Nullstellengebilde der $n - k$ Gleichungen

$$f_j(x_1, \dots, x_n) = x_{k+j} - g_{k+j}(x_1, \dots, x_k), \quad j = 1, \dots, n - k.$$

Die Rangbedingung an die Funktionalmatrix der Funktionen f_j ist erfüllt, denn es gilt

$$\frac{\partial(f_1, \dots, f_{n-k})}{\partial(x_1, \dots, x_n)} = \begin{bmatrix} 1 & & 0 \\ * & \ddots & \\ 0 & & 1 \end{bmatrix},$$

(d.h. die hinteren $n - k$ Spalten bilden die Einheitsmatrix).

3. Jetzt sei M lokal um $a \in M$ Nullstellengebilde der Funktionen f_1, \dots, f_{n-k} und ohne Einschränkung nehmen wir an, dass

$$\det \frac{\partial(f_1, \dots, f_{n-k})}{\partial(x_{k+1}, \dots, x_n)}(a) \neq 0.$$

Wir bilden die Funktion

$$\Phi(x_1, \dots, x_n) = \begin{bmatrix} x_1 \\ \vdots \\ x_k \\ f_1(x_1, \dots, x_n) \\ \vdots \\ f_{n-k}(x_1, \dots, x_n) \end{bmatrix}.$$

Dann ist die Jacobi-Matrix von der Form

$$D\Phi = \begin{bmatrix} I_k & 0_{k \times (n-k)} \\ \frac{\partial(f_1, \dots, f_{n-k})}{\partial(x_1, \dots, x_n)} \end{bmatrix}$$

und aus dem Umkehrsatz folgt, dass Φ in einer Umgebung von a ein Diffeomorphismus ist. Das Nullstellengebilde von f_1, \dots, f_{n-k} wird durch Φ auf die Ebene $E_k = \{(x_1, \dots, x_n) \mid x_{k+1} = \dots = x_n = 0\}$ abgebildet.

4. Sei schließlich M in einer Umgebung U von $a \in M$ auf eine k -dimensionale Ebene transformierbar und zwar durch $\Phi : U \rightarrow V$. Dann ist $\Psi = \Phi^{-1}$ differenzierbar und die Abbildung

$$(t_1, \dots, t_k) \mapsto \varphi(t_1, \dots, t_k) := \Psi(t_1, \dots, t_k, 0, \dots, 0)$$

liefert die geforderte Parameterdarstellung von M in U (Nachrechnen!).

□

21 Tangential- und Normalraum

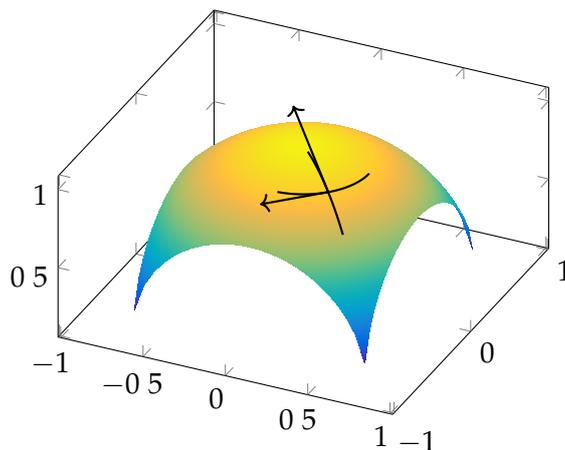
Um mit einer k -dimensionalen Fläche zu arbeiten, ist es hilfreich, ihre Geometrie zu verstehen und benutzen zu können. Diese ist nicht immer direkt aus einer der verschiedenen Darstellungsformen zu erkennen. Hilfreich für den Umgang mit k -dimensionalen Flächen sind die Begriffe “Tangentialvektor” und “Normalenvektor”. Ersteres sind Vektoren, die “tangential an der Fläche anliegen” und zweiteres Vektoren, die senkrecht auf der Fläche stehen. Wir gehen wie folgt vor: Wir definieren einen Tangentialvektor an die Fläche als einen Tangentialvektor einer Kurve, die auf der Fläche verläuft und dann einen Normalenvektor als einen Vektor der senkrecht auf allen Tangentialvektoren steht.

Definition 21.1. Es sei $M \subset \mathbb{R}^n$ eine Untermannigfaltigkeit und $a \in M$. Ein Vektor $v \in \mathbb{R}^n$ heißt *Tangentialvektor an M in a* , wenn es ein $\epsilon > 0$ und eine C^1 -Kurve

$$\alpha :]-\epsilon, \epsilon[\rightarrow M$$

gibt, so dass $\alpha(0) = a$ und $\alpha'(0) = v$.

Hier eine zweidimensionale Untermannigfaltigkeit und zwei Tangentialvektoren:



Definition 21.2 (Tangentialraum, Normalenvektoren, Normalraum). Die Menge aller Tangentialvektoren an M in $a \in M$ nennen wir *Tangentialraum* und wir bezeichnen ihn mit $T_a(M)$.

Ein Vektor $w \in \mathbb{R}^n$, der auf allen $v \in T_a(M)$ senkrecht steht, heißt *Normalenvektor* und die Menge aller Normalenvektoren an M in a nennen wir *Normalraum* und wir bezeichnen ihn mit $N_a(M)$.

Die Begriffe “Tangential-” und “Normalraum” sind gerechtfertigt, da es sich jeweils tatsächlich um Vektorräume handelt, wie der nächste Satz zeigt. Der Satz gibt eine Möglichkeit an, sich Basen dieser Vektorräume zu beschaffen.

Satz 21.3. Es sei $M \subset \mathbb{R}^n$ eine k -dimensionale Untermannigfaltigkeit und $a \in M$. Dann gilt

i) Der Tangentialraum $T_a(M)$ ist ein k -dimensionaler Unterraum des \mathbb{R}^n .

Eine Basis von $T_a(M)$ lässt sich zum Beispiel wie folgt gewinnen: Sei $\varphi : V \rightarrow M$ eine lokale Parametrisierung von M in einer Umgebung von a und sei $t^* \in V$ mit $\varphi(t^*) = a$. Dann sind die Vektoren

$$\frac{\partial \varphi}{\partial t_1}(t^*), \dots, \frac{\partial \varphi}{\partial t_k}(t^*)$$

eine Basis von $T_a(M)$.

ii) Der Normalraum $N_a(M)$ ist ein $n - k$ -dimensionaler Unterraum von \mathbb{R}^n .

Eine Basis von $N_a(M)$ bekommt man zum Beispiel so: Es sei M lokal bei a das Nullstellengebilde zu $f_1, \dots, f_{n-k} : U \rightarrow \mathbb{R}$, d.h. die f_j sind C^1 -Funktionen auf einer offenen Umgebung U von a und es gilt

$$M \cap U = \{x \in U \mid f_1(x) = \dots = f_{n-k}(x) = 0\}$$

und die Funktionalmatrix $\frac{\partial(f_1, \dots, f_{n-k})}{\partial(x_1, \dots, x_n)}$ hat Vollrang. Dann bilden die Vektoren

$$\nabla f_1(a), \dots, \nabla f_{n-k}(a)$$

eine Basis von $N_a(M)$.

Beweis.

Wir bezeichnen den von $\frac{\partial \varphi}{\partial t_1}(t^*), \dots, \frac{\partial \varphi}{\partial t_k}(t^*)$ aufgespannten Vektorraum mit V , den von $\nabla f_1(a), \dots, \nabla f_{n-k}(a)$ aufgespannten Raum mit W und zeigen

$$a) V \subset T_a(M), \quad b) W \subset N_a(M).$$

Damit ist dann auch gezeigt, dass die Dimension von $T_a(M)$ gleich k ist und die Dimension von $N_a(M)$ gleich $n - k$ ist, denn $N_a(M)$ steht senkrecht auf $T_a(M)$ und V ist k -dimensional und W ist $n - k$ -dimensional.

a) Es sei $v \in V$, d.h. v lässt sich schreiben als

$$v = \sum_{i=1}^k c_i \frac{\partial \varphi}{\partial t_i}(t^*)$$

mit $c_i \in \mathbb{R}$. Wir definieren die Kurve

$$\alpha(s) = \varphi(t_* + \begin{bmatrix} c_1 \\ \vdots \\ c_k \end{bmatrix} s)$$

welche für $|s| \leq \epsilon$ und ϵ klein genug wohl definiert ist. Außerdem gilt $\alpha(s) \in M$, da φ eine lokale Parametrisierung ist.

Die Kettenregel liefert

$$\alpha'(s) = D\varphi(t_* + \begin{bmatrix} c_1 \\ \vdots \\ c_k \end{bmatrix} s) \begin{bmatrix} c_1 \\ \vdots \\ c_k \end{bmatrix}, \quad \text{also} \quad \alpha'(0) = \sum_{i=1}^k c_i \frac{\partial \varphi}{\partial t_i}(t^*) = v$$

und folglich ist v der Tangentialvektor der Kurve α und dies zeigt $v \in T_a(M)$.

b) Hier müssen wir nur zeigen, dass $\nabla f_j(a) \in N_a(M)$, und dafür ist zu zeigen, dass $\nabla f_j(a)$ auf jedem $v \in T_a(M)$ senkrecht steht. Ist $v \in T_a(M)$, so ist v von der Form $v = \alpha'(0)$ mit entsprechender Kurve α in M . Da die Kurve α in M verläuft, und alle f_j auf M Null sind, folgt (für die s , für die α definiert ist)

$$f_j(\alpha(s)) = 0.$$

Leiten wir diese Gleichung mit Hilfe der Kettenregel ab, ergibt sich im Punkt $s = 0$ (bedenke $\alpha(0) = a$)

$$0 = Df_j(\alpha(0))\alpha'(0) = \langle \nabla f_j(a), \alpha'(0) \rangle = \langle \nabla f_j(a), v \rangle$$

was die Orthogonalität von $\nabla f_j(a)$ und v zeigt.

□

Wir halten also fest:

- Für eine Basis von $T_a(M)$ leitet man eine Parameterdarstellung ab.
- Für eine Basis von $N_a(M)$ leitet man eine Darstellung als Nullstellengebilde ab.
- Die Dimension des Tangentialraumes $T_a(M)$ ist die Dimension der Untermannigfaltigkeit und die Dimension des Normalraumes $N_a(M)$ die Kodimension der Untermannigfaltigkeit.
- Verschiedene Parameterdarstellungen bzw. verschiedene Darstellungen als Nullstellengebilde geben verschiedene Basen von $T_a(M)$ bzw. $N_a(M)$.
- Es gilt

$$T_a(M) \perp N_a(M) \quad \text{und} \quad T_a(M) \oplus N_a(M) = \mathbb{R}^n.$$

Wir führen noch einen weiteren Begriff ein:

Definition 21.4. Sei M eine Untermannigfaltigkeit mit lokaler Parameterdarstellung $\varphi : T \rightarrow \mathbb{R}^n$. Dann nennen wir die Bilder der achsenparallelen Linien $\{t + se_j\}$ für $t \in T$ die *Parameterlinien*.

Die Parameterlinien werden von vielen Programmen zur Visualisierung gleich mit geplottet und geben oft einen guten Eindruck, wie die Fläche parametrisiert ist. Satz 21.3 sagt also, dass die

Tangentialvektoren entlang der Parameterlinien eine Basis des Tangentialraumes sind.

Beispiel. Wir schauen uns einmal das ganz einfache Beispiel eines affinen linearen Unterraumes des \mathbb{R}^n an. Einen solchen können wir auf verschiedene Art beschreiben:

Durch erzeugende Vektoren: Sind $v_0, v_1, \dots, v_k \in \mathbb{R}^n$ und sind v_1, \dots, v_k linear unabhängig so ist

$$M = \{v_0 + c_1 v_1 + \dots + c_k v_k \mid c_1, \dots, c_k \in \mathbb{R}\} = v_0 + \langle \{v_1, \dots, v_k\} \rangle$$

ein affiner Unterraum und eine Untermannigfaltigkeit der Dimension k . In jedem Punkt $x \in M$ können wir die Wege $\alpha_i(s) = x + s v_i$, $i = 1, \dots, k$ definieren, die ganz in M verlaufen. Da $\alpha_i'(s) = v_i$ gilt, sind alle v_i Tangentialvektoren und bilden zusammen eine Basis des Tangentialraumes.

Fassen wir die Vektoren v_1, \dots, v_k zusammen zu

$$V = [v_1 \ \dots \ v_k] \in \mathbb{R}^{n \times k},$$

so gilt $M = v_0 + \text{Bild}(V)$ und

$$T_x(M) = \text{Bild}(V).$$

Wegen der bekannten Beziehung $\text{Kern}(V) = \text{Bild}(V^T)^\perp$ folgt

$$N_x(M) = \text{Kern}(V^T).$$

Als Lösungsraum: Ist $A \in \mathbb{R}^{m \times n}$ und $b \in \mathbb{R}^m$, so ist die Lösungsmenge

$$M = \{x \in \mathbb{R}^n \mid Ax = b\}$$

ein affiner Unterraum uns insbesondere auch eine Untermannigfaltigkeit der Dimension $n - m$ (falls A den Rang m hat). Um in diesem Fall Tangential- und Normalenraum zu bestimmen, schreiben wir

$$A = \begin{bmatrix} a_1^T \\ \vdots \\ a_m^T \end{bmatrix}, \quad \text{mit } a_1, \dots, a_m \in \mathbb{R}^n.$$

Damit können wir M als Nullstellengebilde zu den Funktionen $g_i(x) = a_i^T x - b_i = \langle a_i, x \rangle - b_i$ schreiben (wobei b_i der i -te Eintrag von b ist). Wir bekommen also Normenvektoren durch Ableiten der g_i , also

$$\nabla g_i(x) = a_i$$

und sehen, dass die Zeilenvektoren von A eine Basis des Normalenraumes $N_x(M)$ bilden (und insbesondere ist dieser unabhängig von x). Mit Begriffen der Linearen Algebra:

$$N_x(M) = \text{Bild}(A^T).$$

Wegen der bekannten Beziehung $\text{Bild}(A^T) = \text{Kern}(A)^\perp$ bekommen wir als Tangentialraum

$$T_x(M) = \text{Kern}(A).$$

Beispiel. Als weiteres einfaches Beispiel betrachten wir die Einheitskugel

$$M = \{x \in \mathbb{R}^n \mid \|x\|_2^2 = 1\}.$$

Sie ist das Nullstellengebilde zu $g(x) = \|x\|_2^2 - 1 = 0$ und daher ist ein Normalenvektor in $a \in \mathbb{R}^n$

$$\nabla g(a) = 2a,$$

d.h. der Normalenraum ist

$$N_a(M) = \langle a \rangle$$

und hat die Dimension 1 und der Tangentialraum ist entsprechend das orthogonale Komplement, also $T_a(M) = a^\perp$, und hat die Dimension $n - 1$.

22 Extrema unter Nebenbedingungen, Lagrange-Multiplikatoren

Der Begriff der Untermannigfaltigkeit erlaubt es, Analysis auf gekrümmten Flächen zu betreiben. So erlauben es lokale Parametrisierungen zum Beispiel, Funktionen, die auf einer Untermannigfaltigkeit definiert sind "entlang der Untermannigfaltigkeit" abzuleiten. In diesem Abschnitt werden wir untersuchen, wie sich Extrema von Funktionen auf Untermannigfaltigkeiten ermitteln lassen. Da sich Untermannigfaltigkeiten auch als Nullstellengebilde $\{x \mid f_1(x) = \dots = f_{n-k}(x) = 0\}$ schreiben lassen, und jede Gleichung $f_j(x) = 0$ sich als eine zusätzliche Bedingung an x auffassen lässt, spricht man auch von "Extrema unter Nebenbedingungen".

Beispiel. Die Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x) = a_1x_1 + a_2x_2$ ist linear und hat auf \mathbb{R}^2 kein Extremum (wenn $[a_1 \ a_2]^T \neq 0$). Wir wollen Sie nun auf dem Einheitskreis betrachten, also auf der Untermannigfaltigkeit $M = \{x \in \mathbb{R}^2 \mid x_1^2 + x_2^2 = 1\}$, und fragen uns: Für welche $x^* \in M$ ist

$$f(x^*) = \max_{x \in M} f(x)?$$

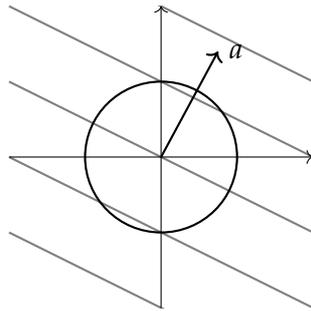
Den Einheitskreis kann man (mit Ausnahme nur eines Punktes) global parametrisieren, z.B. durch $\varphi :]0, 2\pi[\rightarrow \mathbb{R}^2$, $\varphi(t) = [\sin(t) \ \cos(t)]^T$. Ein Maximum von f auf M ist also ein Maximum von $f \circ \varphi$ auf $]0, 2\pi[$ (wobei wir noch den Punkt $[0 \ 1]^T$ extra prüfen müssen). Führt man diese Idee aus, stellt sich heraus, dass dies zwar möglich, aber umständlich ist.

Die Funktion lautet $(f \circ \varphi)(t) = a_1 \sin(t) + a_2 \cos(t)$ mit Ableitung $(f \circ \varphi)'(t) = a_1 \cos(t) - a_2 \sin(t)$. Die Ableitung hat also Nullstellen, wenn $a_1 \cos(t) = a_2 \sin(t)$. Ist $a_2 \neq 0$, so führt dies auf $\frac{a_1}{a_2} = \tan(t)$, was man nach t auflösen kann (allerdings nicht eindeutig).

Wir betrachten die Situation einmal geometrisch: Wir setzen $a = [a_1 \ a_2]^T$ und bemerken, dass die Funktion f entlang aller Linien senkrecht zu a konstant ist.

Es gilt ja $f(x) = \langle a, x \rangle$, also ist für $\langle a, b \rangle = 0$ auch $f(x + b) = \langle a, x + b \rangle = \langle a, x \rangle + \langle a, b \rangle = \langle a, x \rangle = f(x)$.

Zeichnen wie die Niveau-Linien von f und die Menge M in ein Bild, bekommen wir



Außerdem ist f in Richtung von a wachsend (der Gradient ist ja gerade $\nabla f(x) = a$, unabhängig von x). Wir sehen also ohne weitere Rechnung, dass f sein Maximum auf M im Punkt $x^* = a/\|a\|_2$ annimmt und sein Minimum in $-a/\|a\|_2$.

Was steckt hinter der Vorgehensweise im vorigen Beispiel? Wichtig für ein Extremum ist nicht, dass der Gradient verschwindet, sondern, dass er *senkrecht auf allen Tangentialvektoren steht!* (Grob gesagt: Lokale Variationen gibt es entlang von Tangentialvektoren, und daher sind auch nur Variationen in diese Richtung von Bedeutung.) Die Richtungen senkrecht zu Tangentialvektoren sind aber gerade die Normalvektoren. Noch einmal anders ausgedrückt: Ist ein Punkt a^* lokales Extremum, so ist der Gradient $\nabla f(a^*)$ im Normalraum $N_{a^*}(M)$. Dies lässt sich tatsächlich beweisen, und es stellt sich als nützlich notwendiges Kriterium heraus:

Satz 22.1 (Notwendiges Kriterium für Extrema unter Nebenbedingungen). *Es sei $U \subset \mathbb{R}^n$ offen und $M \subset U$ eine Untermannigfaltigkeit von Kodimension r . Diese sei dargestellt als Nullstellengebilde von $g_1, \dots, g_r : U \rightarrow \mathbb{R}$, d.h. $a \in M$ genau dann, wenn $g_1(a) = \dots, g_r(a) = 0$ mit C^1 -Funktionen g_1, \dots, g_r und es gilt*

$$\text{rang} \frac{\partial(g_1, \dots, g_r)}{\partial(x_1, \dots, x_n)} = r.$$

Weiterhin sei $F : U \rightarrow \mathbb{R}$ eine C^1 Funktion und $a^* \in M$ sei lokales Maximum von F in M , d.h. es gibt eine Umgebung V von a^* , so dass für alle $x \in M \cap V$ gilt

$$F(x) \leq F(a^*).$$

Dann gilt: $\nabla F(a^*)$ ist ein Normalenvektor auf M in a^* , d.h. es gibt $\lambda = [\lambda_1 \ \dots \ \lambda_r]^T$, so dass

$$\nabla F(a^*) = [\nabla g_1(a^*) \ \dots \ \nabla g_r(a^*)] \lambda = \sum_{j=1}^r \lambda_j \nabla g_j(a^*).$$

Beweis.

Es sei $\alpha :]-\epsilon, \epsilon[\rightarrow M$ eine C^1 -Kurve mit $\alpha(0) = a^*$. Da F bei a^* ein lokales Maximum bzgl. M hat, hat die Funktion $F \circ \alpha :$

$]-\epsilon, \epsilon[\rightarrow \mathbb{R}$ ein lokales Maximum bei $t = 0$. Es gilt also

$$0 = (F \circ \alpha)'(0) = DF(\alpha(0))\alpha'(0) = \langle \nabla F(a^*), \alpha'(0) \rangle.$$

Dies zeigt, dass $\nabla F(a^*)$ senkrecht auf jedem Tangentialvektor steht (denn diese sind von der Form $\alpha'(0)$). \square

Schreiben wir $g = [g_1 \ \cdots \ g_r]^T : \mathbb{R}^n \rightarrow \mathbb{R}^r$ so schreibt sich die letzte Bedingung auch

$$DF(a^*) = \lambda^T Dg(a^*),^1$$

denn durch transponieren dieser Gleichung folgt $\nabla F(a^*) = DF(a^*)^T = Dg(a^*)^T \lambda$.

Die Zahlen $\lambda_1, \dots, \lambda_r$ heißen *Lagrange-Multiplikatoren*, der Vektor $\lambda \in \mathbb{R}^r$ auch *Lagrange-Multiplikator*.

Zum Rechnen gibt es folgenden Formalismus, der sich ebenfalls gut eignet um sich das Vorgehen einzuprägen: Sucht man lokale Extrema unter Nebenbedingungen, bildet man die *Lagrange-Funktion*

$$\mathcal{L}(x, \lambda) = F(x) - \langle \lambda, g(x) \rangle$$

und sucht x^* und λ^* als Lösung von

$$\begin{aligned} \nabla_x \mathcal{L}(x^*, \lambda^*) &= 0 \\ \nabla_\lambda \mathcal{L}(x^*, \lambda^*) &= 0 \end{aligned}$$

wobei wir mir $\nabla_x \mathcal{L}$ den Gradienten von \mathcal{L} bezüglich der Variable x bezeichnen (und $\nabla_\lambda \mathcal{L}$ entsprechend). Die erste Gleichung ist nämlich äquivalent zu

$$\nabla F(x^*) = Dg(x^*)^T \lambda^*$$

und die zweite ist die Nebenbedingung

$$g(x^*) = 0.$$

Alternativ gibt es auch $\mathcal{L}(x, \lambda) = F(x) + \langle \lambda, g(x) \rangle$ als Definition der Lagrange-Funktion. Dies ist äquivalent (außer dass der Lagrange-Multiplikator umgekehrtes Vorzeichen hat).

Beispiel (Minimum-Norm Lösungen). Es sei $A \in \mathbb{R}^{n \times m}$ mit $n \leq m$ und $\text{rang } A = n$ und $b \in \mathbb{R}^n$. Wir suchen eine Lösung zu

$$\min \|x\|_2 \quad \text{unter der Bedingung} \quad Ax = b.$$

Unter diesen Bedingungen gibt es einen $m - n$ -dimensionalen Kern der Matrix A und die Menge $M = \{x \in \mathbb{R}^m \mid Ax = b\}$ haben es schon im vorigen Abschnitt untersucht. Aktuell ist die Funktion, die wir minimieren wollen nicht differenzierbar (zumindest im Nullpunkt nicht); in diesem Fall können wir sie aber

¹Erinnere: Der Gradient ist die transponierte Jacobi-Matrix ($\nabla F = DF^T$) und die Gradienten der Komponentenfunktionen liegen in den Zeilen der Jacobi-Matrix.

durch Quadrieren differenzierbar machen. Wir betrachten also das äquivalente Problem

$$\min F(x) \quad \text{mit} \quad g(x) = 0$$

mit $F(x) = \|x\|_2^2$, $g(x) = Ax - b$. Satz 22.1 sagt uns, dass wir $\lambda \in \mathbb{R}^n$ und $a^* \in M$ finden sollten, so dass $\nabla F(a^*) = Dg(a^*)^T \lambda$. Wir berechnen

$$\nabla F(a^*) = 2a^*, \quad \text{und} \quad Dg(a^*) = A \text{ (letzteres unabhängig von } a^* \text{)}.$$

Wir bekommen die Gleichung

$$2a^* = A^T \lambda. \quad (*)$$

Da wir auch noch wissen, dass $a^* \in M$ gilt, haben wir auch noch die Gleichung $Aa^* = b$. Setzen wir ineinander ein, bekommen wir für den Lagrange-Multiplikator die Gleichung $AA^T \lambda = 2b$. Es gilt $AA^T \in \mathbb{R}^{n \times n}$ und $\text{rang } AA^T = n$, daher ist AA^T invertierbar und es folgt $\lambda = 2(AA^T)^{-1}b$. Aus (*) folgt nun

$$a^* = A^T(AA^T)^{-1}b.$$

Die Lösung eines unterbestimmten Gleichungssystems $Ax = b$ welche die kleinste Norm hat, ist also $A^T(AA^T)^{-1}b$.

Der Formalismus würde hier wie folgt aussehen: Die Lagrange-Funktion ist

$$\mathcal{L}(x, \lambda) = \|x\|^2 - \langle \lambda, Ax - b \rangle$$

und das Optimalitätssystem lautet

$$\begin{aligned} \nabla_x \mathcal{L}(x^*, \lambda^*) &= 2x^* - A^T \lambda^* = 0 \\ \nabla_\lambda \mathcal{L}(x^*, \lambda^*) &= Ax^* - b = 0. \end{aligned}$$

Als weiteres Beispiel betrachten wir die Operator-Norm von $A \in \mathbb{R}^{n \times m}$.¹ Insbesondere betrachten wir die sogenannte *Spektral-Norm*, welche man erhält, wenn man auf \mathbb{R}^n und \mathbb{R}^m jeweils die 2-Norm wählt. Die Nebenbedingung $x \neq 0$ definiert allerdings keine praktische Untermannigfaltigkeit. Brauchbarer ist die Formulierung

$$\|A\|^2 = \max\{\|Ax\|_2^2 \mid \|x\|_2 = 1\}.$$

Wir werden mit Hilfe von Lagrange-Multiplikatoren zeigen, dass die Spektralnorm die Wurzel aus dem größten Eigenwert von $A^T A$ ist; manchmal geschrieben als

$$\|A\| = \sqrt{\lambda_{\max}(A^T A)}.$$

¹ $\|A\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}$ (bezüglich gewählter Normen auf \mathbb{R}^n und \mathbb{R}^m)

Beispiel (Spektral-Norm). Wir suchen also ein Extremum von $F(x) = \|Ax\|_2^2$ unter der Bedingung $g(x) = \|x\|_2^2 - 1 = 0$.

Es gilt

$$\nabla g(x) = 2x \neq 0$$

was zeigt, dass $M = \{x \mid g(x) = 0\}$ ein Untermannigfaltigkeit ist (was wir aber schon wussten). Leiten wir nun F ab. Berechnen der partiellen Ableitungen ist mühsam, einfacher geht es so: Man beobachtet

$$F(x+h) = \|A(x+h)\|_2^2 = \|Ax\|_2^2 + 2\langle Ax, Ah \rangle + \|Ah\|_2^2$$

und schließt

$$F(x+h) = F(x) + \langle 2A^T Ax, h \rangle + \underbrace{\|Ah\|_2^2}_{=\varphi(h)}.$$

Da $\varphi(h)/\|h\|_2 \rightarrow 0$ für $h \rightarrow 0$ gilt, folgt $DF(x)(h) = \langle 2A^T Ax, h \rangle$, also $DF(x) = 2x^T A^T A$, bzw. $\nabla F(x) = 2A^T Ax$. Für das Lagrange-Funktional $\mathcal{L}(x, \lambda) = \|Ax\|_2^2 - \lambda(\|x\|_2^2 - 1)$ gilt also

$$\nabla_x \mathcal{L}(a, \lambda) = 2A^T Aa - 2\lambda a \quad \nabla_\lambda \mathcal{L}(a, \lambda) = \|a\|_2^2 - 1$$

Setzen wir die Ableitung $\nabla_x \mathcal{L}$ gleich Null, bekommen wir

$$A^T Aa = \lambda a$$

und da auch die Ableitung $\nabla_\lambda \mathcal{L}$ gleich Null sein muss, sehen wir: a ist ein (normierter) Eigenvektor von $A^T A$.

Als Kandidaten für ein globales Maximum haben wir also nur noch die normierten Eigenvektoren von $A^T A$ zu betrachten. Da $A^T A$ symmetrisch und positiv semidefinit ist, gibt es m orthogonale Eigenvektoren v_1, \dots, v_m und zugehörige Eigenwerte $\lambda_1 \geq \dots \geq \lambda_m \geq 0$. Wir testen F für alle v_i :

$$F(v_i) = \|Av_i\|_2^2 = \langle Av_i, Av_i \rangle = \langle A^T Av_i, v_i \rangle = \lambda_i.$$

Offensichtlich ist das Maximum bei λ_1 und wir sehen

$$\|A\| = \sqrt{\lambda_1}.$$

23 Karten und der Maßtensor

Wir wollen uns nun dem Integral über k -dimensionale Untermannigfaltigkeiten des \mathbb{R}^n nähern. Den Spezialfall $k = 1$ haben wir schon einmal gesehen, denn hier handelt es sich um das Integral über eine Kurve $\gamma :]a, b[\rightarrow D \subset \mathbb{R}^n$ mit $\gamma'(t) \neq 0$. Hier ist $M = \gamma(]a, b[)$ eine eindimensionale Untermannigfaltigkeit, und γ eine (globale) Parametrisierung davon. Ist $f : D \rightarrow \mathbb{R}$ eine stetige Funktion, so ist das Kurvenintegral von f über M erklärt als

$$\int_M f = \int_\gamma f = \int_a^b f(\gamma(t)) \|\gamma'(t)\|_2 dt.$$

Das letzte Integral kann man so interpretieren: Wir integrieren über den Parameterbereich und berücksichtigen dabei mit dem Faktor $\|\gamma'(t)\|_2$, wie viel dieser beim Abbilden verzerrt wird. Für das Integral über höherdimensionale Untermannigfaltigkeiten können wir ähnlich vorgehen, müssen aber zwei Dinge klären:

- Wir haben nur noch lokale Parametrisierungen zur Verfügung und müssen diese geeignet zusammenfügen.
- Wir müssen klären, was der Ersatz für den "Verzerrungsfaktor" sein wird.

Stellen wir uns eine Untermannigfaltigkeit mit einer *globalen Karte* vor. D.h. wir betrachten eine Immersion $\varphi : T \rightarrow \mathbb{R}^n$ welche $T \subset \mathbb{R}^k$ homöomorph auf $M = \varphi(T)$ abbildet, so dass M eine Untermannigfaltigkeit ist. Haben wir nun eine Funktion $f : M \rightarrow \mathbb{R}$, so werden wir sehen, dass das Integral von f über M definiert wird als

$$\int_M f dS := \int_T f(\varphi(t)) \sqrt{\det(D\varphi(t)^T D\varphi(t))} d\lambda^k(t). \quad (*)$$

Das Integral über M wird also als ein Integral über den Parameterbereich T ausgedrückt. Wir wollen nun auf verschiedene Weise motivieren, warum der Term $\sqrt{\det(D\varphi(t)^T D\varphi(t))}$ im Integral auftaucht. Dieser Term ist dafür zuständig, die Verzerrung durch die Parametrisierung abzubilden.

1. Beobachtung: Im Fall $n = k$ ist die Formel (*) einfach die Transformationsformel. In diesem Fall ist nämlich $D\varphi(t) \in \mathbb{R}^{n \times n}$ und es gilt

$$\begin{aligned} \sqrt{\det(D\varphi(t)^T D\varphi(t))} &= \sqrt{\det(D\varphi(t)^T) \det(D\varphi(t))} \\ &= |\det(D\varphi(t))|. \end{aligned}$$

2. Beobachtung: Betrachten wir nun die Beschreibung einer Untermannigfaltigkeit durch "Transformation auf eine Ebene": Es sei $M \subset \mathbb{R}^n$ eine k -dimensionale Untermannigfaltigkeit und U eine offene Umgebung der Untermannigfaltigkeit M und $V \subset \mathbb{R}^n$

Es gilt $G = (D\varphi)^T D\varphi$, und daher ist G symmetrisch und positiv semi-definit.

$$\begin{aligned} \langle Gv, v \rangle &= \\ \langle (D\varphi)^T D\varphi v, v \rangle &= \\ \|D\varphi v\|_2^2 &\geq 0 \end{aligned}$$

Es folgt insbesondere, dass $\det(G)$ immer nicht negativ ist (da alle Eigenwerte dies sind).

eine weitere offene Menge. Schließlich sei $\Phi : U \rightarrow V$ ein Diffeomorphismus, der M auf die k -dimensionale Ebene E_k der ersten k Koordinaten abbildet, also $\Phi(M \cap U) = E_k \cap V$. Haben wir nun eine Funktion $f : U \rightarrow \mathbb{R}$ und wollen diese über M integrieren, so versuchen wir, dieses in den "flachen Koordinaten" in $E_k \cap V$ zu tun. Betrachten wir vorerst das Integral von f über die offene Menge $U \supset M$. Dieses können wir mit Hilfe der Transformationsformel auch als Integral über V ausdrücken: Setzen wir $\Psi = \Phi^{-1}$, so gilt $U = \Psi(V)$ und nach der Transformationsformel (Satz 17.4) gilt

$$\int_U f d\lambda = \int_{\Psi(V)} f d\lambda = \int_V f(\Psi(t)) |\det(D\Psi(t))| d\lambda(t).$$

Für die Integration über V müssen wir die Determinante von $D\Psi$ berücksichtigen, die angibt, wie sich das Volumen bei Abbildung mit Ψ lokal vergrößert bzw. verkleinert.

Wir schreiben diese Formel ein wenig um: Wir definieren die Matrix $G(t) = D\Psi(t)^T D\Psi(t)$ mit den Einträgen

$$G_{ij}(t) = \left\langle \frac{\partial \Psi}{\partial t_i}(t), \frac{\partial \Psi}{\partial t_j}(t) \right\rangle.$$

Setzen wir nun $g(t) = \det(G(t))$, so folgt aus der wie eben $\sqrt{g} = |\det(D\Psi)|$ und wir können das Integral über U ausdrücken durch

$$\int_U f d\lambda = \int_{\Psi(V)} f d\lambda = \int_V f(\Psi(t)) \sqrt{g(t)} d\lambda(t).$$

Wollen wir nun nicht über eine Umgebung U von M integrieren, sondern über M selbst, sollten wir auch nicht über V , sondern über $E_k \cap V$ integrieren (was theoretisch kein Problem darstellt, da es sich um eine offene Teilmenge des \mathbb{R}^k handelt, wo wir die Lebesguesche Integrationstheorie entwickelt haben). Dafür sollten die "Verzerrungen durch Ψ in die letzten $n - k$ Richtungen" keine Rolle spielen, d.h., die Ableitungen $\partial \Psi / \partial t_j$ sollen für $k + 1 \leq j \leq n$ nicht vorkommen. Wir können uns die Sache so vorstellen, dass die Ableitungen der Funktion Ψ nach den Variablen $j = k + 1, \dots, n$ alle senkrecht auf dem Tangentialraum der Untermannigfaltigkeit stehen können, d.h. wir haben $\langle \frac{\partial \Psi}{\partial t_i}(t), \frac{\partial \Psi}{\partial t_j}(t) \rangle = 0$, falls $i \leq k$ und $j \geq k + 1$. Nehmen wir auch noch an, dass die Ableitungen $\frac{\partial \Psi}{\partial t_j}(t)$ für $j \geq k + 1$ normiert sind, bekommen wir

$$D\Psi(t)^T D\Psi(t) = \begin{bmatrix} \frac{\partial \Psi}{\partial t_1}(t)^T \\ \vdots \\ \frac{\partial \Psi}{\partial t_n}(t)^T \end{bmatrix} \begin{bmatrix} \frac{\partial \Psi}{\partial t_1}(t) & \cdots \\ \frac{\partial \Psi}{\partial t_n}(t) \end{bmatrix} = \begin{bmatrix} (\langle \frac{\partial \Psi}{\partial t_i}(t), \frac{\partial \Psi}{\partial t_j}(t) \rangle)_{i,j=1,\dots,k} & 0 \\ 0 & E_{n-k} \end{bmatrix}$$

Wir sehen also in diesem Fall

$$|\det(D\Psi(t))| = \sqrt{\det(D\Psi(t)^T D\Psi(t))} = \sqrt{\det((\langle \frac{\partial \Psi}{\partial t_i}(t), \frac{\partial \Psi}{\partial t_j}(t) \rangle)_{i,j=1,\dots,k})}$$

3. Beobachtung: In der großen Übung 8, Aufgabe 3 wurde das Lebesgue-Maß für allgemeine endlichdimensionale reelle Vektorräume erklärt. Ein Ergebnis war das Folgende: Ist ein Unterraum U von den linear unabhängigen Vektoren v_1, \dots, v_k erzeugt, so ist das Lebesgue-Maß des von diesen Vektoren aufgespannten Parallelotops gegeben durch

$$\sqrt{\det((\langle v_i, v_j \rangle)_{i,j=1,\dots,k})}.$$

Haben wir nun eine Karte φ einer Untermannigfaltigkeit, so sind die Vektoren $v_1 = \frac{\partial \varphi}{\partial t_1}(t), \dots, v_k = \frac{\partial \varphi}{\partial t_k}(t)$ eine Basis des Tangentialraumes $T_{\varphi(t)}(M)$ und es gilt

$$\sqrt{\det(D\varphi(t)^T D\varphi(t))} = \sqrt{\det((\langle v_i, v_j \rangle)_{i,j=1,\dots,k})}.$$

All diese Beobachtungen motivieren:

Definition 23.1 (Integral bei globaler Karte). Es sei $M \subset \mathbb{R}^n$ eine k -dimensionale Untermannigfaltigkeit und $\varphi : T \rightarrow M$ eine globale Karte, $G_{ij} : T \rightarrow \mathbb{R}$, $1 \leq i, j \leq k$ definiert durch

$$G_{ij}(t) = \left\langle \frac{\partial \varphi}{\partial t_i}(t), \frac{\partial \varphi}{\partial t_j}(t) \right\rangle.$$

Die $k \times k$ Matrix $G = (G_{ij})$ heißt *Maßtensor* und ihre Determinante

$$g(t) = \det(G(t))$$

heißt *Gramsche Determinante* von M bezüglich der Karte $\varphi : T \rightarrow M$. Damit definieren wir das Integral einer Funktion $f : M \rightarrow \mathbb{R}$ über M durch

$$\int_M f dS := \int_T f(\varphi(t)) \sqrt{g(t)} d\lambda^k(t),$$

falls das Integral auf der rechten Seite existiert.

Zeigen wir noch, dass das Integral wohldefiniert ist, d.h. dass der Wert unabhängig von der gewählten globalen Karte ist. Dazu untersuchen wir, wie sich Maßtensor und Gramsche Determinante unter Kartenwechsel verhalten.

Lemma 23.2. Es sei $M \subset \mathbb{R}^n$ eine k -dimensionale Untermannigfaltigkeit mit zwei globalen Karten $\varphi : T \rightarrow M$, $\tilde{\varphi} : \tilde{T} \rightarrow \tilde{M}$. Dann gilt mit dem Kartenwechsel

$$\tau = \varphi^{-1} \circ \tilde{\varphi} : \tilde{T} \rightarrow T$$

für die Gramsche Determinante \tilde{g} bezüglich $\tilde{\varphi}$

$$\tilde{g}(s) = \det(D\tau(s))^2 g(\tau(s)) \quad \text{für } s \in \tilde{T}.$$

Beweis.

Wie bei Kurvenintegralen stehen wir vor dem Problem, eine aussagekräftige Bezeichnung für das linke Integral über M zu finden. Es gibt hier zahlreiche Konventionen und wir benutzen $\int_M f(x) dS(x)$ wobei S für "Surface" steht. Symbolisch schreibt man " $dS(x) = \sqrt{g(\tilde{t})} d\lambda^k(\tilde{t})$ " und nennt $dS(x)$ das *k-dimensionale Flächenelement*. Weitere gebräuchliche Schreibweisen sind zum Beispiel $\int_M f d^k F$ oder $\int_M f dS_M$.

Wir bezeichnen die Koordinaten in T mit $t = (t_1, \dots, t_k)$ und die in \tilde{T} mit $s = (s_1, \dots, s_k)$ (d.h. $t = \tau(s)$). Nach Definition von τ ist $\tilde{\varphi} = \varphi \circ \tau$ und also gilt nach der Kettenregel

$$D\tilde{\varphi}(s) = D\varphi(\tau(s))D\tau(s)$$

und daher

$$\tilde{G} = (D\tilde{\varphi})^T D\tilde{\varphi} = (D\varphi D\tau)^T D\varphi D\tau = (D\tau)^T \underbrace{(D\varphi)^T D\varphi}_{=G} D\tau.$$

Die Multiplikativität der Determinante zeigt nun

$$\tilde{g}(s) = \det(\tilde{G}(s)) = \det(D\tau(s))^2 g(\tau(s))$$

□

Es lässt sich nun folgern, dass das Integral über Untermannigfaltigkeiten mit globaler Karte unabhängig von der Karte ist:

Satz 23.3. *Es sei $M \subset \mathbb{R}^n$ eine k -dimensionale Untermannigfaltigkeit und $\varphi : T \rightarrow M$ und $\tilde{\varphi} : \tilde{T} \rightarrow M$ zwei globale Karten für M . Dann gilt*

$$\int_M f dS = \int_T f(\varphi(t)) \sqrt{g(t)} d\lambda^k(t) = \int_{\tilde{T}} f(\tilde{\varphi}(s)) \sqrt{\tilde{g}(s)} d\lambda^k(s).$$

Beweis.

Es sei $\tau = \varphi^{-1} \circ \tilde{\varphi} : \tilde{T} \rightarrow T$ der Kartenwechsel. Nach Satz 23.2 gilt

$$\sqrt{\tilde{g}(s)} = |\det(D\tau(s))| \sqrt{g(\tau(s))}$$

und aus der Transformationsformel (Satz 17.4) folgt wegen $\tau(\tilde{T}) = T$

$$\begin{aligned} \int_{\tilde{T}} f(\tilde{\varphi}(s)) \sqrt{\tilde{g}(s)} d\lambda^k(s) &= \int_{\tilde{T}} f(\varphi(\tau(s))) |\det(D\tau(s))| \sqrt{g(\tau(s))} d\lambda^k(s) \\ &= \int_T f(\varphi(t)) \sqrt{g(t)} d\lambda^k(t). \end{aligned}$$

□

Beispiel (Fläche eines Graphen). Es sei $T \subset \mathbb{R}^{n-1}$ und $f : T \rightarrow \mathbb{R}$. Dann ist der Graph $M = \{x \in \mathbb{R}^n \mid x_n = f(x_1, \dots, x_{n-1})\}$ von f eine $n - 1$ -dimensionale Untermannigfaltigkeit. Mit Hilfe der obigen Definition können wir den Flächeninhalt des Graphen wie folgt berechnen. Eine Parameterdarstellung von M ist gegeben durch $\varphi : T \rightarrow \mathbb{R}^n$,

$$\varphi(t_1, \dots, t_{n-1}) = \begin{bmatrix} t_1 \\ \vdots \\ t_{n-1} \\ f(t_1, \dots, t_{n-1}) \end{bmatrix}$$

Es gilt für $1 \leq i \leq n-1$

$$\frac{\partial \varphi}{\partial t_i}(t) = e_i + \frac{\partial f}{\partial t_i}(t)e_n.$$

Daher ist der Maßtensor

$$G_{ij}(t) = \langle e_i + \frac{\partial f}{\partial t_i}(t)e_n, e_j + \frac{\partial f}{\partial t_j}(t)e_n \rangle = \delta_{ij} + \frac{\partial f}{\partial t_i}(t) \frac{\partial f}{\partial t_j}(t)$$

($i, j = 1, \dots, n-1$). In Matrix-Schreibweise ist dies

$$G(t) = I_{n-1} + \nabla f(t) \nabla f(t)^T.$$

Für Matrizen dieser Form gilt

$$g(t) = \det(G(t)) = 1 + \|\nabla f(t)\|_2^2$$

. Folglich ist die Fläche des Graphen von f

$$\int_M 1 dS = \int_T \sqrt{1 + \|\nabla f(t)\|_2^2} d\lambda^{n-1}(t).$$

Für die Parabel $f(x) = \|x\|_2^2$ auf $T = B_1(0) \subset \mathbb{R}^2$, also für die Menge

$$M = \{(x_1, x_2, x_3) \mid x_3 = x_1^2 + x_2^2, \quad x_1^2 + x_2^2 \leq 1\},$$

ergibt sich z.B. (gerechnet in 2-D Polar-Koordinaten)

$$\begin{aligned} \int_M 1 dS &= \int_{B_1(0)} \sqrt{1 + 4\|t\|^2} d\lambda^2(t) = \int_0^{2\pi} \int_0^1 \sqrt{1 + 4r^2} r dr d\theta \\ &= 2\pi \left[\frac{1}{12} (1 + 4r^2)^{\frac{3}{2}} \right]_0^1 = \frac{\pi}{6} (5^{\frac{3}{2}} - 1) \approx 5,33. \end{aligned}$$

Ist $a \in \mathbb{R}^n$ und $A = I + aa^T$, dann ist A symmetrisch und hat als Eigenwerte $1 + \|a\|^2$ (zum Eigenvektor a) und 1 ($n-1$ -fach zu jedem Vektor senkrecht zu a).

24 Das Integral über eine Untermannigfaltigkeit

Wollen wir nun über Untermannigfaltigkeiten integrieren die nicht global parametrisiert sind, sondern nur durch lokale Karten beschrieben werden, so müssen wir aufpassen, wie wir die Bereiche behandeln, an denen sich Karten überlappen.

Wir behandeln den Fall, dass es endlich viele Karten $\varphi_j : T_j \rightarrow V_j, j = 1, \dots, m$ gibt, die die Untermannigfaltigkeit überdecken, also $\bigcup_j V_j = M$. Um das Integral über ganz M aus Integralen über die Bilder der Karten V_j zusammenzufügen, benötigen wir ein wichtiges technisches Hilfsmittel:

Definition 24.1. Es sei M eine Menge, überdeckt von Mengen $V_j, j = 1, \dots, m$, d.h. $M = \bigcup_j V_j$, wobei die V_j nicht disjunkt zu sein brauchen. Eine (der Überdeckung (V_j) untergeordnete) Zerlegung der Eins besteht aus Funktionen $\alpha_j : M \rightarrow \mathbb{R}, j = 1, \dots, m$ mit folgenden Eigenschaften:

- i) Es gilt $0 \leq \alpha_j \leq 1$,
- ii) $\alpha_j(x) = 0$, falls $x \notin V_j$ und
- iii) für jedes $x \in M$ gilt $\sum_{j=1}^m \alpha_j(x) = 1$.

Man spricht von einer C^k -Zerlegung der Eins, wenn die α_j alle C^k -Funktionen sind oder einer integrierbaren Zerlegung der Eins, wenn die α_j alle integrierbar sind (usw.).

In diesem Fall brauchen wir eine lokal integrierbare Zerlegung der Eins, das soll heißen, dass die Funktionen $\alpha_j \circ \varphi_j : T_j \rightarrow \mathbb{R}$ alle jeweils auf jeder kompakten Teilmenge von T_j integrierbar sind.

Eine solche integrierbare Zerlegung der Eins lässt sich z.B. auf einfache Weise wie folgt erhalten: Wir definieren $W_1 = V_1$ und für $j = 2, \dots, m$ dann $W_j = V_j \setminus (\bigcup_{i < j} V_i)$. Dann setzen wir $\alpha_j = \mathbb{1}_{W_j}$.

Wir definieren:

Definition 24.2. Es sei M eine Untermannigfaltigkeit, die mit endlich vielen Karten $\varphi_j : T_j \rightarrow V_j$ dargestellt werden kann. Für eine der Überdeckung (V_j) untergeordnete lokal integrierbare Zerlegung der Eins (α_j) und eine Funktion $f : M \rightarrow \mathbb{R}$ definieren wir

$$\int_M f dS = \sum_{j=1}^m \int_{V_j} \alpha_j(x) f(x) dS(x).$$

Existiert das Integral auf der rechten Seite, so nennen wir $f : M \rightarrow \mathbb{R}$ über M integrierbar.

Es muss noch Wohldefiniertheit gezeigt werden, d.h. dass das Integral unabhängig von der gewählten Zerlegung der Eins ist. Seien (α_j) und (β_l) zwei solche Zerlegungen zu Überdeckungen (V_j) und Karten $\varphi_j : T_j \rightarrow V_j$, bzw. (W_l) und $\psi_l : S_l \rightarrow W_l$.

Hier sind die Integrale auf der rechten Seite durch Definition 23.1 definiert, da sie jeweils φ_j als globale Karte haben.

Da beides Zerlegungen der Eins sind, gilt

$$\sum_l \alpha_j \beta_l = \alpha_j, \quad \sum_j \alpha_j \beta_l = \beta_l,$$

also

$$\sum_{l=1}^p \int_{W_l} \alpha_j(x) \beta_l(x) f(x) dS(x) = \int_M \alpha_j(x) f(x) dS(x)$$

und

$$\sum_{j=1}^m \int_{V_j} \alpha_j(x) \beta_l(x) f(x) dS(x) = \int_M \beta_l(x) f(x) dS(x)$$

Summieren wir über j bzw. l so ergibt sich auf den linken Seite jeweils das gleiche und es folgt die Behauptung.

Mit Hilfe des Integrals über Untermannigfaltigkeiten können wir auch ein k -dimensionales Volumen für Teilmengen des \mathbb{R}^n definieren. Beachte, dass für eine k -dimensionale Untermannigfaltigkeit $M \subset \mathbb{R}^n$ mit $k < n$ immer gilt

$$\lambda^n(M) = 0,$$

d.h. M ist eine Lebesgue Nullmenge.

Definition 24.3. Es sei $M \subset \mathbb{R}^n$ eine k -dimensionale Untermannigfaltigkeit. Eine Menge $A \subset M$ heißt *integrierbare Teilmenge* von M , wenn die Funktion $\mathbb{1}_A$ über M integrierbar ist. Wir definieren das *k -dimensionale Volumen* von A als

$$\text{vol}_k(A) = \int_M \mathbb{1}_A(x) dS(x).$$

Wir sagen, dass $N \subset M$ eine k -dimensionale Nullmenge ist, falls $\text{vol}_k(N) = 0$.

Beispiel (Integrale über die S^2). Wir betrachten die Oberfläche S^2 der Einheitskugel im \mathbb{R}^3 und eine lokale Parametrisierung davon

$$\Phi :]0, 2\pi[\times]-\pi/2, \pi/2[\rightarrow \mathbb{R}^3, \quad (\varphi, \theta) \mapsto \begin{bmatrix} \cos(\varphi) \cos(\theta) \\ \sin(\varphi) \cos(\theta) \\ \sin(\theta) \end{bmatrix}$$

Die Ableitungen der Parametrisierung sind

$$\frac{\partial \Phi}{\partial \varphi}(\varphi, \theta) = \begin{bmatrix} -\sin(\varphi) \cos(\theta) \\ \cos(\varphi) \cos(\theta) \\ 0 \end{bmatrix}, \quad \frac{\partial \Phi}{\partial \theta}(\varphi, \theta) = \begin{bmatrix} -\cos(\varphi) \sin(\theta) \\ -\sin(\varphi) \sin(\theta) \\ \cos(\theta) \end{bmatrix}$$

Also ist der Maßtensor

$$G(\varphi, \theta) = \begin{bmatrix} \|\frac{\partial \Phi}{\partial \varphi}(\varphi, \theta)\|_2^2 & \langle \frac{\partial \Phi}{\partial \varphi}(\varphi, \theta), \frac{\partial \Phi}{\partial \theta}(\varphi, \theta) \rangle \\ \langle \frac{\partial \Phi}{\partial \varphi}(\varphi, \theta), \frac{\partial \Phi}{\partial \theta}(\varphi, \theta) \rangle & \|\frac{\partial \Phi}{\partial \theta}(\varphi, \theta)\|_2^2 \end{bmatrix} = \begin{bmatrix} \cos(\theta)^2 & 0 \\ 0 & 1 \end{bmatrix}.$$

Obige Definition ist für Teilmengen von Untermannigfaltigkeiten befriedigend und sie ermöglicht es, ein k -dimensionales Volumen für glatte Flächen zu erklären. Für nicht glatte Flächen stößt man aber wieder auf Schwierigkeiten (z.B. wenn diese sich nicht in endlich viele glatte Flächen disjunkt zerlegen lassen). Die Maßtheorie bietet hier einen eleganten Weg mit Hilfe der *k -dimensionalen Hausdorff-Maße* für Teilmengen des \mathbb{R}^n . Diese sind Null für Mengen der Dimension $< k$, haben den Wert $+\infty$ für Mengen der Dimension $> k$ und liefern für glatte Flächen A der Dimension k den Wert $\text{vol}_k(A)$, sind aber für eine viel allgemeinere Klasse von Mengen definiert.

Dies gibt die Gramsche Determinante $g(\varphi, \theta) = \cos(\theta)^2$ und da $\cos(\theta) > 0$ für $-\pi/2 < \theta < \pi/2$ folgt

$$\sqrt{g(\varphi, \theta)} = \cos(\theta).$$

Da die Parametrisierung die ganze Kugelfläche mit Ausnahme des Null-Längengrades erreicht, und dieser 2-dimensionales Volumen Null hat, gilt für eine integrierbare Funktion $f : S^2 \rightarrow \mathbb{R}$:

$$\int_{S^2} f dS = \int_0^{2\pi} \int_{-\pi/2}^{\pi/2} f(\Phi(\varphi, \theta)) \cos(\theta) d\theta d\varphi.$$

Zum Beispiel gilt für die Fläche des S^2

$$\text{vol}_2(S^2) = \int_0^{2\pi} \int_{-\pi/2}^{\pi/2} \cos(\theta) d\theta d\varphi = \int_0^{2\pi} 2 d\varphi = 4\pi.$$

Beispiel (Rotationsflächen). Wir betrachten für ein offenes Intervall $I \subset \mathbb{R}$ und stetig differenzierbares $f : I \rightarrow [0, \infty[$ die Menge

$$M = \{x \in \mathbb{R}^3 \mid x_3 \in I, \quad x_1^2 + x_2^2 = f(x_3)^2\}$$

(also die Rotationsfläche des Graphen von f). Dies ist eine zweidimensionale Untermannigfaltigkeit im \mathbb{R}^3 und (bis auf eine Nullmenge) parametrisiert durch

$$\Phi : I \times]0, 2\pi[\rightarrow \mathbb{R}^3, \quad \Phi(t, \varphi) = \begin{bmatrix} f(t) \cos(\varphi) \\ f(t) \sin(\varphi) \\ t \end{bmatrix}.$$

Die Ableitung der Parametrisierung ist

$$\frac{\partial \Phi}{\partial t}(t, \varphi) = \begin{bmatrix} f'(t) \cos(\varphi) \\ f'(t) \sin(\varphi) \\ 1 \end{bmatrix}, \quad \frac{\partial \Phi}{\partial \varphi}(t, \varphi) = \begin{bmatrix} -f(t) \sin(\varphi) \\ f(t) \cos(\varphi) \\ 0 \end{bmatrix}.$$

Der Maßtensor ist also

$$G(t, \varphi) = \begin{bmatrix} 1 + f'(t)^2 & 0 \\ 0 & f(t)^2 \end{bmatrix}$$

und daher ist

$$\sqrt{g(t, \varphi)} = f(t) \sqrt{1 + f'(t)^2}.$$

Die Rotationsfläche des Graphen ist also

$$\text{vol}_2(M) = \int_0^{2\pi} \int_I f(t) \sqrt{1 + f'(t)^2} dt d\varphi = 2\pi \int_I f(t) \sqrt{1 + f'(t)^2} dt \quad (*)$$

(falls das Integral existiert).

Beispiel (Noch einmal die Fläche eines Paraboloids). Im Beispiel im letzten Abschnitt hatten wir die Fläche des Paraboloids

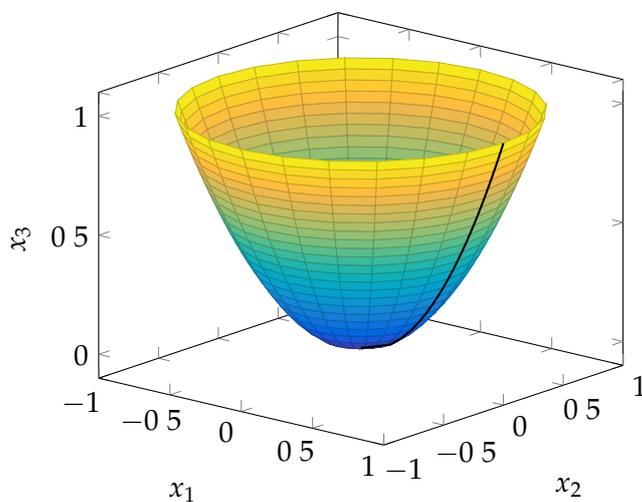
$$M = \{x \in \mathbb{R}^3 \mid x_3 = x_1^2 + x_2^2, \quad x_1^2 + x_2^2 < 1\}$$

berechnet (mit $\text{vol}_2(M) = \frac{\pi}{6}(5^{\frac{3}{2}} - 1)$), indem wir M als Graph von $f : B_1(0) \rightarrow \mathbb{R}$, $f(x_1, x_2) = x_1^2 + x_2^2$ aufgefasst hatten.

Ebenso gut können wir M (bis auf eine Menge von zweidimensionalen Volumen Null) als Rotationsfläche schreiben, nämlich mit $g :]0,1[\rightarrow [0, \infty[$, $g(t) = \sqrt{t}$ betrachten wir

$$M' = \{x \in \mathbb{R}^3 \mid x_3 \in]0,1[, \quad x_1^2 + x_2^2 = (\sqrt{x_3})^2\}$$

(dies entspricht, bis auf den Punkt 0, der Menge M). Hier die Menge M und (etwas fett gedruckt) der Graph $x_1 = \sqrt{x_3}$:



Für die Fläche von M bekommen wir also nach (*) wegen $g'(t) = \frac{1}{2\sqrt{t}}$

$$\begin{aligned} \text{vol}_2(M') &= 2\pi \int_0^1 \sqrt{t} \sqrt{1 + \frac{1}{4t}} dt = 2\pi \int_0^1 \sqrt{t + \frac{1}{4}} dt \\ &= 2\pi \left[\frac{2}{3} \left(t + \frac{1}{4}\right)^{\frac{3}{2}} \right]_{t=0}^1 = \frac{\pi}{6} (5^{\frac{3}{2}} - 1) \end{aligned}$$

wie vorher.

25 Integrieren entlang des Radius

Das n -dimensionale Volumen (bzw. Lebesgue-Maß) verhält sich bei Streckung bzw. Stauchung mit einem Faktor $r \geq 0$ wie r^n , also $\lambda^n(rM) = r^n \lambda^n(M)$. Das k -dimensionale Volumen von Mengen im \mathbb{R}^n skaliert hingegen wie r^k . Dies folgt aus diesem Satz:

Satz 25.1. *Ist $M \subset \mathbb{R}^n$ eine k -dimensionale Untermannigfaltigkeit und $r > 0$, so ist rM ebenfalls eine k -dimensionale Untermannigfaltigkeit. Eine Funktion $f : rM \rightarrow \mathbb{R}$ ist genau dann über rM integrierbar, wenn die Funktion $g : M \rightarrow \mathbb{R}$, $g(y) = f(ry)$ über M integrierbar ist und es gilt*

$$\int_{rM} f(x) dS(x) = \int_M f(ry) r^k dS(y).$$

Insbesondere ist für jede integrierbare Teilmenge $A \subset M$ auch $rA \subset rM$ integrierbar und es gilt

$$\text{vol}_k(rA) = r^k \text{vol}_k(A).$$

Beweis.

Ist $\varphi : T \rightarrow V \subset M$ eine Karte für M , so ist $\tilde{\varphi}(x) := r\varphi(x)$, $\tilde{\varphi} : T \rightarrow rV \subset rM$ eine Karte für rM . Es gilt $D\tilde{\varphi} = rD\varphi$ und damit $\tilde{G} = r^2G$. Es folgt (da $\tilde{G} \in \mathbb{R}^{k \times k}$) für die Gramsche Determinante

$$\tilde{g}(t) = r^{2k} g(t), \quad \text{also} \quad \sqrt{\tilde{g}(t)} = r^k g(t),$$

woraus die Behauptung folgt. \square

Mit Hilfe dieses Satzes beweisen wir folgenden Satz, der ähnlich zur Integration in Kugel-Koordinaten ist:

Satz 25.2. *Es sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ integrierbar. Dann ist für Lebesgue-fast alle $r > 0$ die Funktion f über die Sphäre $M_r = \{x \in \mathbb{R}^n \mid \|x\|_2 = r\}$ integrierbar und es gilt*

$$\int_{\mathbb{R}^n} f d\lambda^n = \int_0^\infty \int_{\|x\|_2=r} f(x) dS(x) d\lambda^1(r).$$

Beweis.

Wir bezeichnen den oberen, bzw. unteren, Halbraum mit

$$H_\pm = \{x \in \mathbb{R}^n \mid \pm x_n > 0\}.$$

Es genügt, die Aussage für $f \mathbb{1}_{H_+}$ zu zeigen.

Es gilt Lebesgue-fast überall $f = f \mathbb{1}_{H_+} + f \mathbb{1}_{H_-}$ und außerdem hat $M_r \cap \{x_n = 0\}$ $(n-1)$ -dimensionales Volumen 0. Schließlich ist der Fall für $f \mathbb{1}_{H_-}$ analog.

Wir betrachten also nur den Fall, dass f nur auf H_+ nicht Null ist.

Wir bezeichnen die offene Einheitskugel im \mathbb{R}^{n-1} mit $U =$

$\{\xi \in \mathbb{R}^{n-1} \mid \|\xi\|_2 < 1\}$ und definieren

$$\Phi : U \times]0, \infty[\rightarrow H_+, \quad \Phi(\xi_1, \dots, \xi_{n-1}, r) = \begin{bmatrix} r\xi_1 \\ \vdots \\ r\xi_{n-1} \\ r\sqrt{1 - \|\xi\|_2^2} \end{bmatrix}.$$

Dieses Φ ist bijektiv und sogar ein C^1 -Diffeomorphismus. Man kann nachrechnen, dass

$$\det(D\Phi(\xi, r)) = \frac{r^{n-1}}{\sqrt{1 - \|\xi\|_2^2}}.$$

Es ist

$$D\Phi = \begin{bmatrix} r & 0 & \dots & 0 & \xi_1 \\ 0 & \ddots & \ddots & \vdots & \xi_2 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \dots & 0 & r & \xi_{n-1} \\ -\frac{r\xi_1}{\sqrt{1 - \|\xi\|_2^2}} & -\frac{r\xi_2}{\sqrt{1 - \|\xi\|_2^2}} & \dots & -\frac{r\xi_{n-1}}{\sqrt{1 - \|\xi\|_2^2}} & \sqrt{1 - \|\xi\|_2^2} \end{bmatrix}.$$

Durch Entwickeln der Determinante nach der letzten Spalte und geschicktes Berechnen der sich ergebenden Unterdeterminanten erhält man die Behauptung.

Nach der Transformationsformel (Satz 17.4) und dem Satz von Fubini gilt

$$\begin{aligned} \int_{H_+} f(x) d\lambda^n(x) &= \int_{U \times]0, \infty[} f(\Phi(\xi, r)) \det(D\Phi(\xi, r)) d\lambda^n(\xi, r) \\ &= \int_0^\infty \int_U f(r\xi, r\sqrt{1 - \|\xi\|_2^2}) \frac{r^{n-1}}{\sqrt{1 - \|\xi\|_2^2}} d\lambda^{n-1}(\xi) d\lambda^1(r). \end{aligned}$$

Es bleibt also zu zeigen, dass

$$\int_U f(r\xi, r\sqrt{1 - \|\xi\|_2^2}) \frac{r^{n-1}}{\sqrt{1 - \|\xi\|_2^2}} d\lambda^{n-1}(\xi) = \int_{\|x\|_2=r} f(x) dS(x).$$

Um dies zu sehen, stellen wir die $n - 1$ -dimensionale Untermannigfaltigkeit $M = \{x \in \mathbb{R}^n \mid \|x\|_2 = r, x_n > 0\}$ als Graph der Funktion

$$F : T \rightarrow \mathbb{R}, \quad T = \{t \in \mathbb{R}^{n-1} \mid \|t\|_2 < r\}, \quad F(t) = \sqrt{r^2 - \|t\|_2^2}$$

dar. Nach dem Beispiel aus Abschnitt 23 ist die Gramsche Determinante

$$g(t) = 1 + \|\nabla F(t)\|_2^2$$

und wegen $\frac{\partial F}{\partial t_i}(t) = \frac{-t_i}{\sqrt{r^2 - \|t\|_2^2}}$ folgt

$$g(t) = 1 + \frac{\|t\|_2^2}{r^2 - \|t\|_2^2} = \frac{r^2}{r^2 - \|t\|_2^2}.$$

Es folgt

$$\begin{aligned} \int_{\|x\|_2=r} f(x) dS(x) &= \int_{\|t\|_2 < r} f(t, F(t)) \frac{r}{\sqrt{r^2 - \|t\|_2^2}} d\lambda^{n-1}(t) \\ &= \int_{\|t\|_2 < r} f(t, \sqrt{r^2 - \|t\|_2^2}) \frac{r}{\sqrt{r^2 - \|t\|_2^2}} d\lambda^{n-1}(t) \\ &= \int_{\|\xi\|_2 < 1} f(r\xi, r\sqrt{1 - \|\xi\|_2^2}) \frac{r^{n-1}}{\sqrt{1 - \|\xi\|_2^2}} d\lambda^{n-1}(\xi) \end{aligned}$$

wobei in der letzten Zeile noch einmal der Transformationssatz benutzt wurde. \square

Kombination von Satz 25.1 und Satz 25.2 liefert:

Korollar 25.3. Es sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ integrierbar. Dann gilt

$$\int_{\mathbb{R}^n} f d\lambda^n = \int_0^\infty \int_{\|y\|_2=1} f(ry) dS(y) r^{n-1} d\lambda^1(r).$$

Korollar 25.4. Es sei $f : [0, \infty[\rightarrow \mathbb{R}$ und

$$\omega_n = \text{vol}_{n-1}(\{y \in \mathbb{R}^n \mid \|y\|_2 = 1\})$$

die $n - 1$ -dimensionale Fläche der n -Sphäre. Dann gilt

$$\int_{\mathbb{R}^n} f(\|x\|_2) d\lambda^n(x) = \omega_n \int_0^\infty f(r) r^{n-1} d\lambda^1(r)$$

falls die rechte Seite existiert (d.h., falls die Abbildung $x \mapsto f(\|x\|_2)$ integrierbar ist).

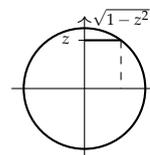
Vorigs Korollar zeigt

$$\begin{aligned} \int_{\mathbb{R}^n} f(\|x\|_2) d\lambda^n &= \int_0^\infty \int_{\|y\|_2=1} f(r) dS(y) r^{n-1} d\lambda^1(r) \\ &= \text{vol}_{n-1}(\{y \in \mathbb{R}^n \mid \|y\|_2 = 1\}) \int_0^\infty f(r) r^{n-1} d\lambda^1(r). \end{aligned}$$

Um noch ω_n zu berechnen, berechnen wir zuerst das Lebesgue-Maß der n -dimensionalen Einheitskugel

$$B_1^n(0) = \{x \in \mathbb{R}^n \mid \|x\|_2 \leq 1\}.$$

Wir bezeichnen $\beta_n = \lambda^n(B_1^n(0))$. Dann ist $\lambda^n(B_1^n(r)) = r^n \beta_n$. Wir nutzen das Cavalierische Prinzip (Satz 16.1): Die Querschnitte auf Höhe z mit $z^2 \leq 1$ sind $(n - 1)$ -dimensionale Kugeln mit Radius $\sqrt{1 - z^2}$. Wir bekommen also die Rekursionsformel



$$\beta_n = \beta_{n-1} \int_{-1}^1 (\sqrt{1-z^2})^{n-1} dz.$$

Mit der Substitution $z = \sin(x)$ folgt $\beta_n = \beta_{n-1} \int_{-\pi/2}^{\pi/2} \cos(x)^n dx$
Wir bezeichnen berechnen mit partieller Integration

$$\begin{aligned} I_n &= \int_{-\pi/2}^{\pi/2} \cos(x)^n dx = \cos(x)^{n-1} \sin(x) \Big|_{-\pi/2}^{\pi/2} + \int_{-\pi/2}^{\pi/2} (n-1) \cos(x)^{n-2} \underbrace{\sin(x) \cos(x)}_{=\sin(x)^2=1-\cos(x)^2} dx \\ &= (n-1) \int_{-\pi/2}^{\pi/2} \cos(x)^{n-2} (1-\cos(x)^2) dx \\ &= (n-1) I_{n-2} - (n-1) I_n \end{aligned}$$

woraus wir die Rekursionsformel $I_n = \frac{n-1}{n} I_{n-2}$ gewinnen. Die ersten beiden Integrale sind

$$I_0 = \int_{-\pi/2}^{\pi/2} 1 dx = \pi \quad I_1 = \int_{-\pi/2}^{\pi/2} \cos(x) dx = 2$$

und es folgt rekursiv

$$I_{2m} = \pi \prod_{k=1}^m \frac{2k-1}{2k}, \quad I_{2m+1} = 2 \prod_{k=1}^m \frac{2k}{2k+1}.$$

Damit bekommen wir $I_n I_{n-1} = 2\pi/n$ und es folgt $\beta_n = \beta_{n-1} I_n = \beta_{n-2} I_n I_{n-1}$, also

$$\beta_{2m} = \frac{\pi}{m} \beta_{2m-2}, \quad \beta_{2m+1} = \frac{2\pi}{2m+1} \beta_{2m-1}.$$

Bekanntermaßen ist $\beta_1 = \lambda^1([-1,1]) = 2$ und $\beta_2 = \pi$ und wir bekommen

$$\beta_{2m} = \frac{\pi^m}{m!}, \quad \beta_{2m+1} = \frac{2^{m+1} \pi^m}{1 \cdot 3 \cdots (2m+1)}.$$

Es ergibt sich folgende Tabelle

n	1	2	3	4	5	6
β_n	2	π	$\frac{4}{3}\pi$	$\frac{1}{2}\pi^2$	$\frac{8}{15}\pi^2$	$\frac{1}{6}\pi^3$
\approx	2	3,14	4,19	4,93	5,26	5,17

Inbesondere ist β_n für $n = 5$ maximal, für $n \geq 5$ fallend und es gilt $\beta_n \rightarrow 0$ für $n \rightarrow \infty$.

Um nun die Fläche der Einheitskugel zu berechnen argumentieren wir wie folgt: Das Volumen der Einheitskugel mit Radius r ist $\beta_n(r) := r^n \beta_n(1) = r^n \beta_n$. Im Sinne des vorigen Satzes bekommen wir die Oberfläche der Einheitskugel als Ableitung des Volumens nach dem Radius, also ist

$$\omega_n(r) = nr^{n-1} \beta_n \quad \text{also} \quad \omega_n = \omega_n(1) = n\beta_n$$

also

$$\omega_{2m} = \frac{2\pi^m}{(m-1)!}, \quad \beta_{2m+1} = \frac{2^{m-1}\pi^m}{1 \cdot 3 \cdots (2m-1)},$$

z.B.

$$\omega_1 = 2, \quad \omega_2 = 2\pi, \quad \omega_3 = 4\pi, \quad \omega_4 = 2\pi^2, \quad \omega_5 = \frac{8}{3}\pi^2.$$

26 Kompakta mit glattem Rand

Der Fundamentalsatz der Integral- und Differentialrechnung

$$\int_a^b f'(t) dt = f(b) - f(a)$$

setzt das Integral über die Ableitung einer Funktion über ein Intervall mit den Werten der Funktion an den Rändern des Intervalls in Beziehung. Eine ähnliche Situation lässt sich in höheren Dimensionen wie folgt beschreiben: Ist $A \subset \mathbb{R}^n$ derart, dass ∂A (also der Rand von A) eine $n - 1$ -dimensionale Untermannigfaltigkeit ist, lässt sich ein Integral über eine Funktion auf ∂A mit einem Integral über eine Ableitung der Funktion über A in Beziehung setzen?

Die Antwort liefert der Gaußsche Integralsatz: Es stellt sich heraus, dass hierbei nicht Funktionen über ∂A integriert werden, sondern *Vektorfelder*. Der Gaußsche Integralsatz lautet:

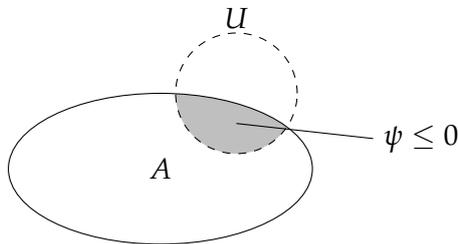
$$\int_A \operatorname{div} F(x) d\lambda^n(x) = \int_{\partial A} \langle F(a), \nu(a) \rangle dS(a).$$

Hierbei ist $F : A \rightarrow \mathbb{R}^n$ ein Vektorfeld, $\operatorname{div} F$ ein (noch zu erklärender) Differentialoperator und $\nu(a)$ ein Einheitsvektor, der außen und senkrecht auf ∂A steht. Im Falle von $n = 1$ erhält man genau den Fundamentalsatz zurück: Denn dann ist $A = [a, b]$, $\operatorname{div} F = F'$, $\partial A = \{a, b\}$, "Randintegral" entspricht der Auswertung von F am Rand und der Vektor ν gibt das Vorzeichen.

Als erstes legen wir die Klasse von Mengen A fest, über deren Rand wir integrieren wollen:

Definition 26.1. Wir nennen $A \subset \mathbb{R}^n$ ein *Kompaktum mit glattem Rand*, wenn A kompakt ist und es zu jedem Randpunkt $a \in \partial A$ eine offene Umgebung $U \subset \mathbb{R}^n$ und eine C^1 Funktion $\psi : U \rightarrow \mathbb{R}$ gibt, so dass gilt

- i) $A \cap U = \{x \in U \mid \psi(x) \leq 0\}$
- ii) $\nabla \psi(x) \neq 0$ für alle $x \in U$.



Satz 26.2. Ist A ein Kompaktum mit glattem Rand und U, ψ wie oben, so gilt

$$\partial A \cap U = \{x \in U \mid \psi(x) = 0\}.$$

Folglich ist ∂A lokal ein Nullstellengebilde und daher eine $n - 1$ -dimensionale Untermannigfaltigkeit.

Insbesondere gibt es $\epsilon > 0$, so dass $a + t \nabla \psi(a) \notin A$ für $0 < t < \epsilon$.

Beweis.

Zeigen wir erst $\partial A \cap U = \{x \in U \mid \psi(x) = 0\}$.

1. “ \subset ”: Es ist zu zeigen, dass für $x \in U$ mit $\psi(x) < 0$ folgt, dass $x \notin \partial A$. Da aber ψ stetig ist, folgt aus $\psi(x) < 0$ auch, dass $\psi(y) < 0$ für y aus einer Umgebung $V \subset U$ von x . Also ist $V \subset A$ und damit x ein innerer Punkt von A , also kein Randpunkt.
2. “ \supset ”: Sei nun $a \in U$ und $\psi(a) = 0$. Wir wollen zeigen, dass $a \in \partial A$.

Wir setzen $v = \nabla\psi(a)$ (und wissen, dass $v \neq 0$). Die Taylor-Entwicklung von ψ in a lautet

$$\psi(a+h) = \psi(a) + \langle \nabla\psi(a), h \rangle + R(\|h\|_2) = \langle v, h \rangle + R(\|h\|_2)$$

mit einem Rest R , für den gilt

$$\frac{R(\|h\|_2)}{\|h\|_2} \rightarrow 0, \quad \text{für } h \rightarrow 0.$$

Speziell für $h = tv$ lautet dies

$$\psi(a+tv) = t\|v\|_2^2 + \tilde{R}(t)$$

mit $\tilde{R}(t)/t \rightarrow 0$ für $t \rightarrow 0$. Also gibt es $\epsilon > 0$, so dass gilt

$$\begin{aligned} \psi(a+tv) &> 0 & \text{für } 0 < t < \epsilon \\ \psi(a+tv) &< 0 & \text{für } -\epsilon < t < 0. \end{aligned}$$

Also ist $a+tv \notin A$, wenn $0 < t < \epsilon$ und $\in A$, wenn $-\epsilon < t < 0$. Dies zeigt, dass jede Umgebung von a Punkte innerhalb und außerhalb von A enthält, was zeigt, dass a ein Randpunkt von A ist.

Im zweiten Schritt haben wir den Rest der Behauptung schon gezeigt. \square

Satz 26.3 (und Definition des äußeren Normalen-Einheitsvektors).
Ist A ein Kompaktum mit glattem Rand, so existiert zu jedem $a \in \partial A$ genau ein Vektor $v(a) \in \mathbb{R}^n$ mit den Eigenschaften

- i) $v(a) \in N_a(\partial A)$ (d.h. $v(a)$ ein **Normalenvektor** für ∂A in a),
- ii) $\|v(a)\|_2 = 1$ (d.h. $v(a)$ ist ein **Einheitsvektor**), und
- iii) es gibt $\epsilon > 0$, so dass für $0 < t < \epsilon$ gilt

$$a + tv(a) \notin A$$

(d.h. $v(a)$ zeigt nach **außen**).

Der Vektor $v(a)$ heißt **äußerer Normalen-Einheitsvektor** von A im Punkt a .

Beweis.

1. Existenz: Es seien U und ψ wie in der Definition eines Kompaktums mit glattem Rand. Dann hat der Vektor

$$v(a) = \frac{\nabla\psi(a)}{\|\nabla\psi(a)\|_2}$$

die geforderten Eigenschaften i) bis iii).

i) folgt, da ∂A nach Satz 26.2 lokal das Nullstellengebilde von ψ ist und nach Satz 21.3 ii) dann $\nabla\psi(a)$ ein Normalenvektor ist. ii) ist nach Konstruktion klar und iii) haben wir eben in Satz 26.2 gezeigt.

2. Eindeutigkeit: Da ∂A eine $n - 1$ -dimensionale Untermannigfaltigkeit ist, ist $N_a(\partial A)$ eindimensional und $\nabla\psi(a)$ ein Basis. Es folgt also $v(a) = \lambda \nabla\psi(a)$. Um normiert zu sein, muss $|\lambda| = \|\nabla\psi(a)\|_2^{-1}$ gelten und damit $v(a)$ nach außen zeigt, muss $\lambda > 0$ gelten. Dies legt $v(a)$ eindeutig fest.

□

Korollar 26.4. Auf einem Kompaktum A mit glattem Rand gibt es ein stetiges Normalen-Einheitsfeld

$$v : \partial A \rightarrow \mathbb{R}^n.$$

Die Stetigkeit von $v(a) = \nabla\psi(a) / \|\nabla\psi(a)\|_2$ folgt daraus, dass ψ stetig differenzierbar ist.

Beispiel (Äußerer Normalen-Einheitsvektor auf einer Kugel). Ist $A = \bar{B}_r(0) = \{x \in \mathbb{R}^n \mid \|x\|_2 \leq r\}$, so ist $\partial A = \{x \mid \psi(x) = \|x\|_2^2 - r^2 = 0\}$, also ist für $\|a\|_2 = r$

$$v(a) = \frac{\nabla\psi(a)}{\|\nabla\psi(a)\|_2} = \frac{2a}{\|2a\|_2} = \frac{a}{r}.$$

Der äußere Normalen-Einheitsvektor entspricht in diesem Fall also dem skalierten Ortsvektor.

Als Vorbereitung auf den nächsten Abschnitt zeigen wir zwei weitere Ergebnisse:

Lemma 26.5. Es sei $\Omega = \Omega' \times I$ mit $\Omega' \in \mathbb{R}^{n-1}$ offen, $I = [a, b]$ ein Intervall, $f : \Omega \rightarrow \mathbb{R}$, $g : \Omega' \rightarrow \mathbb{R}$ differenzierbar, $g(x') \in]a, b[$ für $x' \in \Omega'$. Gibt es dann für $i \in \{1, \dots, n-1\}$ eine integrierbare Funktion $G_i : I \rightarrow \mathbb{R}$ mit $|\frac{\partial}{\partial x_i} f(x', x_n)| \leq G_i(x_n)$, so ist

$$F(x') := \int_a^{g(x')} f(x', x_n) dx_n$$

partiell differenzierbar nach x_i differenzierbar mit

$$\frac{\partial}{\partial x_i} F(x') = \int_a^{g(x')} \frac{\partial}{\partial x_i} f(x', x_n) dx_n + f(x', g(x')) \frac{\partial}{\partial x_i} g(x').$$

Beweis.

Wir setzen

$$h(x', z) = \int_a^z f(x', x_n) dx_n$$

denn dann ist $F(x') = h(x', g(x'))$. Nach der Kettelregel ist

$$\frac{\partial}{\partial x_i} F(x') = \frac{\partial}{\partial x_i} h(x', g(x')) + \frac{\partial}{\partial z} h(x', g(x')) \frac{\partial}{\partial x_i} g(x').$$

Berechnen wir noch die partiellen Ableitungen von h : Mit Satz 13.3 folgt

$$\frac{\partial}{\partial x_i} h(x', z) = \int_a^z \frac{\partial}{\partial x_i} f(x', x_n) dx_n, \quad \frac{\partial}{\partial z} h(x', z) = f(x', z)$$

und das zeigt Behauptung. \square

Hat von $u, v \in C^1([a, b])$ eine Funktion kompakten Trager, so gilt in der Regel der partiellen Integration

$$\int_a^b u'(x)v(x)dx = - \int_a^b u(x)v'(x)dx.$$

Dieses Ergebnis wollen wir nun fur Funktionen auf Teilmengen des \mathbb{R}^n verallgemeinern:

Satz 26.6 (Partielle Integration). *Es sei $U \subset \mathbb{R}^n$ offen. Dann gilt fur $i = 1, \dots, n$:*

i) Fur $\varphi \in C_c^1(U)$ gilt

$$\int_U \frac{\partial \varphi(x)}{\partial x_i} d\lambda^n(x) = 0.$$

ii) Fur $u \in C^1(U)$ und $v \in C_c^1(U)$ gilt

$$\int_U \frac{\partial u(x)}{\partial x_i} v(x) d\lambda^n(x) = - \int_U u(x) \frac{\partial v(x)}{\partial x_i} d\lambda^n(x).$$

Beweis.

Teil ii) folgt aus i) indem man i) auf uv anwendet.

Es ist dann $uv \in C_c^1(U)$ und nach der Produktregel gilt

$$\frac{\partial (uv)(x)}{\partial x_i} (x) = \frac{\partial u(x)}{\partial x_i} v(x) + u(x) \frac{\partial v(x)}{\partial x_i}.$$

Weiterhin konnen wir ohne Einschrankung $U = \mathbb{R}^n$ setzen, da wir φ durch 0 auf ganz \mathbb{R}^n fortsetzen konnen. Auerdem nehmen wir ohne Einschrankung $i = 1$.

Ist $\text{supp } \varphi$ kompakt, so existiert $R > 0$, so dass $\text{supp } \varphi \subset]-R, R[^n$. Es folgt mit dem Satz von Fubini und dem Fundamentalsatz

der Differential- und Integralrechnung

$$\begin{aligned}\int_{\mathbb{R}^n} \frac{\partial \varphi(x)}{\partial x_1} d\lambda^n(x) &= \int_{-R}^R \cdots \int_{-R}^R \frac{\partial \varphi(x)}{\partial x_1} dx_1 \cdots dx_n \\ &= \int_{-R}^R \cdots \int_{-R}^R \underbrace{\left[\varphi(x) \right]_{x_1=-R}^R}_{=0} dx_2 \cdots dx_n \\ &= 0.\end{aligned}$$

□

27 Die wichtigsten Differentialoperatoren

Beispiel (und Definition von Divergenz und Laplace Operator). Ist $U \subset \mathbb{R}^n$, $f \in C_c^1(U)$ und $g \in C^1(U, \mathbb{R}^n)$, dann ist $\nabla f \in C_c(U, \mathbb{R}^n)$ und es gilt nach dem Satz 26.6 über partielle Integration

$$\begin{aligned} \int_U \langle \nabla f(x), g(x) \rangle d\lambda^n(x) &= \int_U \sum_{i=1}^n \frac{\partial f(x)}{\partial x_i} g_i(x) d\lambda^n(x) \\ &= \sum_{i=1}^n \int_U \frac{\partial f(x)}{\partial x_i} g_i(x) d\lambda^n(x) \\ &= - \sum_{i=1}^n \int_U f(x) \frac{\partial g_i(x)}{\partial x_i} d\lambda^n(x) \\ &= - \int_U f(x) \sum_{i=1}^n \frac{\partial g_i(x)}{\partial x_i} d\lambda^n(x) \end{aligned}$$

(und ebenso für $f \in C^1(U)$ und $g \in C_c^1(U, \mathbb{R}^n)$). Für Vektorfelder $g \in C^1(U, \mathbb{R}^n)$ nennen wir den Differentialoperator, der auf der rechten Seite auftaucht *Divergenz*, und schreiben

$$\operatorname{div} g(x) = \sum_{i=1}^n \frac{\partial g_i(x)}{\partial x_i} = \frac{\partial g_1(x)}{\partial x_1} + \dots + \frac{\partial g_n(x)}{\partial x_n}.$$

Anders ausgedrückt $\operatorname{div} g = \operatorname{spur}(Dg)$ (die Spur der Jacobi-Matrix).
Damit können wir schreiben

$$\int_U \langle \nabla f(x), g(x) \rangle d\lambda^n(x) = - \int_U f(x) \operatorname{div} g(x) d\lambda^n(x). \quad (+)$$

Für $f \in C^2(U)$ definieren wir

$$\Delta f = \operatorname{div}(\nabla f) = \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2}$$

und nennen Δ den *Laplace Operator*.

Mit anderen Worten: $\Delta f = \operatorname{spur} H_f$ (die Spur der Hesse-Matrix).

Wesentlich für die Identität

$$\int_U \langle \nabla f(x), g(x) \rangle d\lambda^n(x) = - \int_U f(x) \operatorname{div} g(x) d\lambda^n(x)$$

ist, dass eine der beiden Funktionen kompakten Träger hat. Ist dies nicht der Fall, so treten "Randterme" dazu. Diese Situation erfordert weitere Untersuchung und führt auf den Divergenzsatz von Gauß. Als Einstieg untersuchen wir eine besonders einfache geometrische Situation: Wir betrachten eine Menge folgender Form:

Führen wir für $f, g \in C_c^1(U)$ bzw. $F, G \in C_c^1(U, \mathbb{R}^n)$ die Skalarprodukte

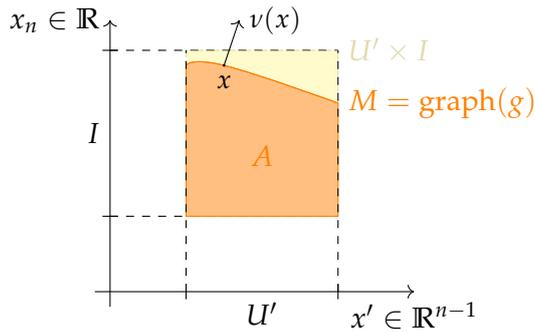
$$\langle f, g \rangle = \int_U f(x)g(x) d\lambda(x),$$

$$\langle F, G \rangle = \int_U \langle F(x), G(x) \rangle_{\mathbb{R}^n} d\lambda(x)$$

ein, so liest sich Gleichung (+) formal als

$$\langle \nabla f, G \rangle = - \langle f, \operatorname{div} G \rangle = \langle f, - \operatorname{div} G \rangle.$$

In diesem Sinne sagt man auch "– div ist der zu ∇ adjungierte Differentialoperator".



Satz 27.1 (Vorabversion des Divergenzsatzes von Gauß). Es sei $U' \subset \mathbb{R}^{n-1}$ offen, $I =]\alpha, \beta[$, $g : U' \rightarrow I$ eine C^1 -Funktion und

$$A := \{(x', x_n) \in U' \times I \mid x_n \leq g(x')\}$$

$$M := \{(x', x_n) \in U' \times I \mid x_n = g(x')\}.$$

Ferner sei für $x \in M$ der Vektor $v(x)$ der äußere Normalen-Einheitsvektor auf M , also

$$v(x) = \frac{\begin{bmatrix} -\nabla g(x') \\ 1 \end{bmatrix}}{\sqrt{1 + \|\nabla g(x')\|_2^2}}.$$

Dann gilt für jedes $f \in C_c^1(U' \times I)$ und $i = 1, \dots, n$

$$\int_A \frac{\partial f}{\partial x_i}(x) d\lambda^n(x) = \int_M f(x) v_i(x) dS(x).$$

Beweis.

Nach Beispiel aus Abschnitt 23 ist die Gramsche Determinante für M bezüglich der Parameterdarstellung $x' \mapsto (x', g(x'))$ gerade $1 + \|\nabla g(x')\|_2^2$. Wir bearbeiten die Fälle $i = n$ und $i = 1, \dots, n-1$ getrennt:

1. Fall: $i = n$. Dies ist der einfachere Fall: Da die Abbildung $x_n \mapsto f(x', x_n)$ für jedes x' kompakten Träger in I hat, folgt mit dem Fundamentalsatz der Integral- und Differentialrechnung

$$\int_{\alpha}^{g(x')} \frac{\partial f}{\partial x_n}(x', x_n) dx_n = f(x', g(x')).$$

Mit dem Satz von Fubini folgt

$$\begin{aligned} \int_A \frac{\partial f}{\partial x_n}(x', x_n) d\lambda^n(x) &= \int_{U'} \left(\int_{\alpha}^{g(x')} \frac{\partial f}{\partial x_n}(x', x_n) dx_n \right) d\lambda^{n-1}(x') \\ &= \int_{U'} f(x', g(x')) d\lambda^{n-1}(x') \\ &= \int_{U'} f(x', g(x')) \frac{\sqrt{1 + \|\nabla g(x')\|_2^2}}{\sqrt{1 + \|\nabla g(x')\|_2^2}} d\lambda^{n-1}(x') \\ &= \int_M f(x) v_n(x) dS(x) \end{aligned}$$

wobei wir am Ende $\nu_n(x) = 1/\sqrt{1 + \|\nabla g(x')\|_2^2}$ und “ $dS(x) = \sqrt{1 + \|\nabla g(x')\|_2^2} d\lambda^{n-1}(x')$ ” benutzt haben.

2. Fall: $1 \leq i \leq n-1$. Da Differenzenquotienten von stetig differenzierbaren Funktionen sogar gleichmäßig auf kompaktem Mengen konvergieren, können wir die Aussage von Lemma 26.5 anwenden und bekommen

$$\begin{aligned} & \frac{\partial}{\partial x_i} \int_{\alpha}^{g(x')} f(x', x_n) dx_n \\ &= \int_{\alpha}^{g(x')} \frac{\partial f}{\partial x_i}(x', x_n) dx_n + f(x', g(x')) \frac{\partial}{\partial x_i} g(x'). \quad (*) \end{aligned}$$

Da die Funktion $\varphi : U' \rightarrow \mathbb{R}$, $\varphi(x') = \int_{\alpha}^{g(x')} f(x', x_n) dx_n$ kompakten Träger hat (da f dies hat), folgt mit Satz 26.6 i):

$$\int_{U'} \frac{\partial}{\partial x_i} \left(\int_{\alpha}^{g(x')} f(x', x_n) dx_n \right) d\lambda^{n-1}(x') = 0 \quad (**)$$

Mit dem Satz von Fubini folgt mit (*) und (**)

$$\begin{aligned} \int_A \frac{\partial f}{\partial x_i}(x) d\lambda^n(x) &= \int_{U'} \int_{\alpha}^{g(x')} \frac{\partial f}{\partial x_i}(x) dx_n d\lambda^{n-1}(x') \\ &\stackrel{(*)}{=} \int_{U'} \frac{\partial}{\partial x_i} \left(\int_{\alpha}^{g(x')} f(x', x_n) dx_n \right) d\lambda^{n-1}(x') \\ &\quad - \int_{U'} f(x', g(x')) \frac{\partial g(x')}{\partial x_i} d\lambda^{n-1}(x') \\ &\stackrel{(**)}{=} - \int_{U'} f(x', g(x')) \frac{\partial g(x')}{\partial x_i} d\lambda^{n-1}(x'). \end{aligned}$$

Im letzten Integral erweitern wir mit $\sqrt{1 + \|\nabla g(x')\|_2^2}$, und erhalten

$$\begin{aligned} & - \int_{U'} f(x', g(x')) \frac{\partial g(x')}{\partial x_i} d\lambda^{n-1}(x') \\ &= \int_{U'} f(x', g(x')) \frac{-\frac{\partial g(x')}{\partial x_i}}{\sqrt{1 + \|\nabla g(x')\|_2^2}} \sqrt{1 + \|\nabla g(x')\|_2^2} d\lambda^{n-1}(x') \\ &= \int_M f(x) \nu_i(x) dS(x) \end{aligned}$$

wobei wir im letzten Schritt wieder die Form von ν_i und die des Flächenelements benutzt haben.

□

Bemerkung. Wenden wir diesen Satz auf Funktionen f_i an und

summieren die Gleichungen über i , so ergibt sich

$$\begin{aligned}\int_A \operatorname{div} f \, d\lambda &= \int_A \sum_{i=1}^n \frac{\partial f_i}{\partial x_i} \, d\lambda \\ &= \int_M \sum_{i=1}^n f_i(x) v_i(x) \, dS(x) \\ &= \int_M \langle f(x), v(x) \rangle \, dS(x).\end{aligned}$$

Mit anderen Worten: Das Integral über die Divergenz eines Vektorfeldes lässt sich als Randintegral ausdrücken. ■

Um den Divergenzatz von Gauß für allgemeine Kompakta A mit glattem Rand zu beweisen, zerlegen wir A in Teile, die die einfache Form aus Satz 27.1 haben, dann wenden wir auf diesen Teilen den Satz an, und schließlich fügen wir alles zusammen. Wir werden sehen, dass wir zwei Dinge dabei berücksichtigen müssen, damit im Beweis alles gut geht:

1. Wir müssen das Problem auf stetig differenzierbare Art auf die Teile aufteilen, und
2. die Zerlegung darf nicht aus "zu kleinen Mengen" bestehen.

Für den ersten Punkt benötigen wir eine C^1 -Zerlegung der Eins und für den zweiten Punkt ein weiteres geometrisches Hilfsmittel für Überdeckungen von Kompakta.

Kommen wir erst zu glatten Zerlegungen der Eins. Wir benötigen diese nur für Zerlegungen, die aus äquidistant positionierten Würfeln bestehen. In diesem Fall gibt es eine ziemlich explizite Konstruktion:

Lemma 27.2 (C^∞ Zerlegung der Eins für Würfel). *Zu $\epsilon > 0$ und $p \in \mathbb{Z}^n$ setzen wir*

$$W_\epsilon(p\epsilon) = \{x \in \mathbb{R}^n \mid \|x - p\epsilon\|_\infty \leq \epsilon\}.$$

Dann gibt es Funktionen $\alpha_p \in C^\infty(\mathbb{R}^n)$, die eine der Überdeckung $(W_\epsilon(p\epsilon))_{p \in \mathbb{Z}^n}$ untergeordnete Zerlegung der Eins in \mathbb{R}^n bilden.

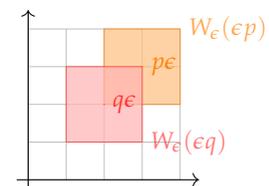
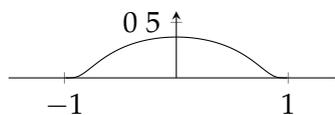
Beweis.

Wir müssen also zu jedem $p \in \mathbb{Z}^n$ ein $\alpha_p \in C^\infty(\mathbb{R}^n)$ konstruieren, so dass gilt:

1. $0 \leq \alpha_p \leq 1$,
2. $\operatorname{supp} \alpha_p \subset W_\epsilon(p\epsilon)$, und
3. für alle $x \in \mathbb{R}^n$ gilt $\sum_{p \in \mathbb{Z}^n} \alpha_p(x) = 1$.

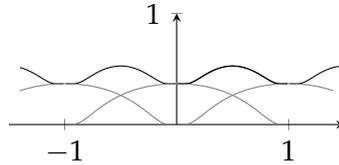
Wir beginnen mit $n = 1$ und definieren $g : \mathbb{R} \rightarrow \mathbb{R}$ durch

$$g(t) = \begin{cases} \exp(-\frac{1}{1-t^2}), & |t| < 1 \\ 0, & |t| \geq 1 \end{cases}.$$



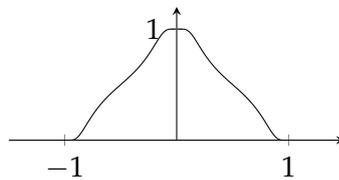
Bekannterweise ist $g \in C_c^\infty(\mathbb{R})$ mit $\text{supp } g = [-1, 1]$. Nun wollen wir durch ganzzahlige Verschiebung hieraus eine Zerlegung der Eins für die Intervalle $[k-1, k+1]$ machen. Hierfür müssen wir noch dafür sorgen, dass sich die Funktionen punktweise zu Eins aufsummieren. Dazu setzen wir

$$G(t) = \sum_{k \in \mathbb{Z}} g(t-k).$$



Es gilt $G \in C^\infty(\mathbb{R})$ und $G(t-k) = G(t)$ für $k \in \mathbb{Z}$ (beachte, dass die Reihe in jedem Punkt konvergiert, da jeweils nur zwei Summanden nicht Null sind). Da $G(t) \neq 0$ für alle t , setzen wir

$$h(t) = \frac{g(t)}{G(t)}.$$



Hierfür gilt $h \in C_c^\infty(\mathbb{R})$ mit $\text{supp } h = [-1, 1]$, und außerdem ist für alle t

$$\sum_{k \in \mathbb{Z}} h(t-k) = \sum_{k \in \mathbb{Z}} \frac{g(t-k)}{G(t-k)} = \sum_{k \in \mathbb{Z}} \frac{g(t-k)}{G(t)} = \frac{1}{G(t)} \sum_{k \in \mathbb{Z}} g(t-k) = 1.$$

Anders gesagt: Die Funktionen $h_k(t) = h(t-k)$ bilden eine der Überdeckung $([k-1, k+1])_{k \in \mathbb{Z}}$ untergeordnete C^∞ -Zerlegung der Eins.

Wir definieren schließlich zu $p \in \mathbb{Z}^n$ die Funktionen $\alpha_p : \mathbb{R}^n \rightarrow \mathbb{R}$ durch

$$\alpha_p(x) = \prod_{i=1}^n h\left(\frac{x_i}{\epsilon} - p_i\right).$$

Man überzeugt sich davon, dass $\alpha_p \in C_c^\infty(\mathbb{R}^n)$, $\text{supp } \alpha_p \subset W_\epsilon(p\epsilon)$ und

$$\sum_{p \in \mathbb{Z}^n} \alpha_p(x) = 1,$$

wie gefordert. □

28 Der Divergenzsatz von Gauß

Das letzte Hilfsmittel, das wir für den Beweis des Satzes von Gauß brauchen, ist die sogenannte *Lebesgue-Zahl* einer offenen Überdeckung. Wir benötigen dazu den Begriff des *Durchmessers* (engl.: diameter) einer Menge. Dieser ist für $A \subset \mathbb{R}^n$

$$\text{diam}(A) = \sup\{\|x - y\|_2 \mid x, y \in A\}$$

(ggf. $+\infty$).

Lemma 28.1 (Lebesgue-Zahl). *Es sei $A \subset \mathbb{R}^n$ kompakt und $U_i \subset \mathbb{R}^n$, $i \in I$ eine offene Überdeckung von A .*

Dann existiert ein $\lambda > 0$ (genannt Lebesgue-Zahl zu A und (U_i)), so dass für jedes kompakte $K \subset \mathbb{R}^n$ mit $\text{diam}(K) \leq \lambda$ welches A schneidet ein $i \in I$ existiert, mit $K \subset U_i$.

Beweis.

Zu jedem $a \in A$ existiert ein $i \in I$ und ein $r_a > 0$, so dass

$$B_{r_a}(a) \subset U_i.$$

Die Familie $(B_{r_a/2}(a))_{a \in A}$ überdeckt A , und da A kompakt ist, gibt es a_1, \dots, a_m , so dass

$$A \subset \bigcup_{k=1}^m B_{r_{a_k}/2}(a_k).$$

Dann ist $\lambda = \min_{1 \leq k \leq m} \left(\frac{r_{a_k}}{2}\right)$ eine Lebesgue-Zahl, denn: Sei $K \subset \mathbb{R}^n$ kompakt mit $\text{diam}(K) \leq \lambda$ und $K \cap A \neq \emptyset$. Zu $a \in K \cap A$ gibt es $k \in \{1, \dots, m\}$ und $i \in I$, so dass

$$a \in B_{r_{a_k}/2}(a_k) \subset B_{r_{a_k}}(a_k) \subset U_i.$$

Da $\text{diam}(K) \leq r_{a_k}/2$, folgt

$$K \subset B_{r_{a_k}}(a_k) \subset U_i.$$

Für $a' \in K$ gilt $\|a - a'\|_2 \leq \text{diam}(K) \leq r_{a_k}/2$ und da auch $\|a - a_k\| \leq r_{a_k}/2$, folgt für $a' \in K$ auch $\|a' - a_k\| \leq r_{a_k}$.

□

Nun können wir den Divergenzsatz von Gauß beweisen:

Satz 28.2 (Divergenzsatz von Gauss). *Es sei $A \subset \mathbb{R}^n$ ein Kompaktum mit glattem Rand, $\nu : \partial A \rightarrow \mathbb{R}^n$ das äußere Einheits-Normalenfeld und $U \subset \mathbb{R}^n$ offen mit $A \subset U$. Dann gilt für jedes $F \in C^1(U, \mathbb{R}^n)$*

$$\int_A \text{div} F(x) d\lambda^n(x) = \int_{\partial A} \langle F(x), \nu(x) \rangle dS(x).$$

Beweis.

Da nach Satz 26.2 der Rand von A lokal Nullstellengebilde ist, gibt es offene Mengen $U_i, i \in I$, so dass $\partial A \subset \bigcup_{i \in I} U_i$, so dass wir (nach Umnummerierung der Koordinaten) schreiben können $U_i = U'_i \times]\alpha_i, \beta_i[$ und C^1 -Funktionen $g_i : U'_i \rightarrow]\alpha_i, \beta_i[$ gibt, so dass

$$U_i \cap A = \{(x', x_n) \in U'_i \times]\alpha_i, \beta_i[\mid x_n \leq g_i(x') \text{ (bzw. } x_n \geq g_i(x'))\}.$$

Um aus den U_i eine Überdeckung von ganz A zu machen, reicht es, noch eine weitere Menge \tilde{U} zu den U_i hinzuzufügen, von der wir annehmen können, dass $\tilde{U} \subset A^\circ$. Für die resultierenden offene Überdeckung von A sei $\lambda > 0$ die Lebesgue-Zahl aus Lemma 28.1.

Wir setzen $\epsilon = \frac{\lambda}{2\sqrt{n}}$ und betrachte die in Lemma 27.2 konstruierte Zerlegung der Eins $(\alpha_p)_{p \in \mathbb{Z}}$. Es gilt $\text{supp } \alpha_p \subset W_\epsilon(p\epsilon)$ und daher hat $\text{supp } \alpha_p$ einen Durchmesser kleiner als die Diagonale eines Würfels mit Kantenlänge 2ϵ , also kleiner als $2\epsilon\sqrt{n} = \lambda$. Es sei

$$P = \{p \in \mathbb{Z}^n \mid \text{supp } \alpha_p \cap A \neq \emptyset\}$$

(beachte: P ist endlich, da A beschränkt). Dann ist für $x \in A$ auch $F(x) = \sum_{p \in P} \alpha_p(x)F(x)$ und da die Divergenz linear ist, folgt

$$\int_A \text{div } F d\lambda^n = \sum_{p \in P} \int_A \text{div}(\alpha_p F) d\lambda^n.$$

Außerdem gilt

$$\int_{\partial A} \langle F(x), \nu(x) \rangle dS(x) = \sum_{p \in P} \int_{\partial A} \langle \alpha_p(x)F(x), \nu(x) \rangle dS(x).$$

Wir müssen die Behauptung also für die Funktionen $\alpha_p F$ zeigen.

Nach unserer Konstruktion ist $\text{supp}(\alpha_p F)$ in einem U_i enthalten. Nun haben wir noch zwei Fälle:

1. Fall: $\text{supp}(\alpha_p F) \subset \tilde{U}$, d.h. $\alpha_p F$ hat kompakten Träger in A° .
Nach der Regel der partiellen Integration (Satz 26.6) gilt

$$\int_A \text{div}(\alpha_p F) d\lambda^n = 0 = \int_{\partial A} \langle \alpha_p(x)F(x), \nu(x) \rangle dS(x)$$

(die linke Gleichung ist die partielle Integration und die rechte Gleichung ist klar, da der Integrand im Integrationsgebiet konstant Null ist).

2. Fall: Das zugehörige U_i schneidet den Rand ∂A . Hier folgt die Behauptung indem wir Satz 27.1 auf die Komponenten der vektorwertigen Funktion $\alpha_p F$ anwenden und aufsummieren.

□

Beispiel. Wir betrachten $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $F(x) = x$. Die Divergenz ist einfach

$$\operatorname{div} F(x) = \sum_{i=1}^n \frac{\partial x_i}{\partial x_i} = n.$$

Daher gilt für jedes Kompaktum mit glattem Rand A

$$\int_{\partial A} \langle x, \nu(x) \rangle dS(x) = n \int_A d\lambda^n = n\lambda^n(A)$$

bzw.

$$\lambda^n(A) = \frac{1}{n} \int_{\partial A} \langle x, \nu(x) \rangle dS(x).$$

Das heißt, wir können das n -dimensionale Volumen einer Menge bestimmen, indem wir die Größe $\langle x, \nu(x) \rangle$ über den Rand der Menge integrieren.

Im Falle der Einheitskugel $K_n = \bar{B}_1(0) \subset \mathbb{R}^n$ ist $\partial K_n = S^{n-1}$ und $\nu(x) = x$ (vgl. Beispiel aus Abschnitt 26) und es folgt

$$\operatorname{vol}_n(K_n) = \frac{1}{n} \int_{S^{n-1}} \underbrace{\langle x, x \rangle}_{=\|x\|_2^2=1} dS(x) = \frac{1}{n} \operatorname{vol}_{n-1}(S^{n-1}).$$

Bemerkung (Bedeutung in der Physik). Für ein Vektorfeld F und das Einheits-Normalenfeld ν auf ein Kompaktum A mit glattem Rand bezeichnen wir für $x \in \partial A$ mit $\alpha(x)$ den Winkel zwischen $F(x)$ und $\nu(x)$, d.h.

$$\langle F(x), \nu(x) \rangle = \|F(x)\| \cos(\alpha(x)).$$

Beschreibt das Vektorfeld $F : U \rightarrow \mathbb{R}^n$ zum Beispiel den Fluss einer Quantität, also z.B. die Geschwindigkeit einer Wasserströmung, so ist die Größe

$$\langle F(x), \nu(x) \rangle$$

der Anteil der Strömung im Punkt x in Richtung $\nu(x)$, also aus A heraus. (Anteile der Strömung tangential an ∂A lassen nichts aus A heraus fließen.) Daher ist das Integral

$$\int_{\partial A} \langle F(x), \nu(x) \rangle dS(x)$$

der *Gesamtfluss durch die Oberfläche ∂A von A* . Der Divergenzsatz sagt, dass dieser Gesamtfluss gleich dem Integral über die Divergenz des Flussfeldes über ganz A ist. Daher nennt man $\operatorname{div} F(x)$ auch die *Quellstärke von F* : Ist z.B. $\operatorname{div} F(x) = 0$, auf ganz A , so folgt, dass genau so viel in A hinein, wie wieder heraus fließt.

Felder F , für die $\operatorname{div} F = 0$ gilt, nennt man auch *quellenfrei* oder *divergenzfrei*.

Strömungen von inkompressiblen Flüssigkeiten (wie zum Beispiel Wasser) müssen zum Beispiel divergenzfrei sein, wenn im Inneren von A keine Wasserquellen oder -senken sind. ■

Wendet man den Divergenzsatz auf gF mit $F \in C^1(U, \mathbb{R}^n)$ und $g \in C^1(U)$, so erhält man wegen

$$\operatorname{div}(gF) = \sum_{i=1}^n \frac{\partial(gF_i)}{\partial x_i} = \sum_{i=1}^n \left[\frac{\partial g}{\partial x_i} F_i + g \frac{\partial F_i}{\partial x_i} \right] = \langle \nabla g, F \rangle + g \operatorname{div} F$$

die folgende Variante der partiellen Integration mit Randtermen:

Korollar 28.3. *In der Situation von Satz 28.2 gilt für $F \in C^1(U, \mathbb{R}^n)$ und $g \in C^1(U, \mathbb{R})$*

$$\int_A g(x) \operatorname{div} F(x) d\lambda^n = - \int_A \langle \nabla g(x), F(x) \rangle d\lambda^n + \int_{\partial A} \langle g(x) F(x), \nu(x) \rangle dS(x).$$

Auf gleiche Weise erhält man eine nützliche Formel, wenn man für $g \in C^1(U)$ und $f \in C^2(U)$ den Divergenzsatz auf das Vektorfeld $F = g\nabla f$ anwendet. Hier ist (setze $F = \nabla f$ in obiger Rechnung)

$$\operatorname{div}(g\nabla f) = \langle \nabla g, \nabla f \rangle + g\Delta f.$$

Wir führen noch folgende Notation ein: Für den Normalen-Einheitsvektor $\nu(x)$ für $x \in \partial A$ schreiben wir

$$\frac{\partial f}{\partial \nu}(x) = \langle \nabla f(x), \nu(x) \rangle$$

für die Ableitung von f in Normalenrichtung.

Wir haben gezeigt:

Korollar 28.4 (1. Greensche Formel). *Für obige f und g gilt in der Situation des Divergenzsatzes von Gauß*

$$\int_A \langle \nabla g, \nabla f \rangle + g\Delta f d\lambda^n = \int_{\partial A} g \frac{\partial f}{\partial \nu} dS.$$

Vertauschen wir f und g und ziehen die beiden Formeln voneinander ab, bekommen wir:

Korollar 28.5 (2. Greensche Formel). *Für obige $f, g \in C^2(U)$ gilt in der Situation des Divergenzsatzes von Gauß*

$$\int_A f\Delta g - g\Delta f d\lambda^n = \int_{\partial A} f \frac{\partial g}{\partial \nu} - g \frac{\partial f}{\partial \nu} dS.$$