

# FH Aachen

## University of Applied Sciences

### Campus Jülich

Fachbereich „Medizintechnik und Technomathematik“

Studiengang: Angewandte Mathematik und Informatik

## **Anwendung Johnson-Lindenstrauss Lemma zur verzerrungsarmen Einbettung von Punkten aus einem hochdimensionalen in einen niedrig-dimensionalen euklidischen Raum**

Seminararbeit von Morteza Montahaee

1. Prüfer: Prof. Dr. rer. Nat. Martin Reißel

Aachen, Mai 2023

# Inhaltsverzeichnis

---

<b>1 Einleitung.....</b>	<b>2</b>
1.1 Motivation.....	2
<b>2 Grundlagen der Maßtheorie.....</b>	<b>3</b>
2.1 Offene, Abgeschlossene, und Kompakte Mengen.....	4
2.2 Rechtecken, Quadern.....	5
2.3 Einleitende Bemerkungen.....	5
2.4 Äußeres Maß und ihre Eigenschaften.....	9
2.5 Messbare Mengen und Lebesgue-Maß.....	12
2.6 Messbare Funktionen.....	16
2.7 Das Lebesgue Integral.....	19
2.7.1 Einfache Funktion.....	19
2.7.2 Beschränkte Funktion.....	20
2.7.3 Nicht-Negative Funktionen.....	21
2.7.4 Allgemeines Fall.....	22
2.8 Die Brunn-Minkowski Ungleichung.....	27
2.9 Rektifizierbare Kurve.....	30
2.10 Die Isoperimetrische Ungleichung.....	34
<b>3 Johnson-Lindenstrauss Lemma.....</b>	<b>37</b>
3.1 Grundlage der Wahrscheinlichkeitstheorie.....	37
3.2 Affine Geometrie.....	39
3.3 Konzentration des Maßes.....	40
3.4 Das Johnson-Lindenstrauss Lemma.....	44
<b>4 Anhang.....</b>	<b>50</b>
4.1 Abbildungsverzeichnis.....	50
4.2 Bibliographie.....	50

## 1 Einleitung

---

Daten können in mehreren Verfahren zur weiteren Verarbeitung modelliert werden. Manchmal können sie als  $m$  Datensätze mit  $n$  Merkmalen in einer relationalen Datenbank beschreiben werden. Sie können aber auch im Lauf eines Data-Mining oder einer numerischen Analyse bzw. Numerik geschöpft werden. Dabei wirken die Dimensionalität der Daten Ressourcenverbrauch (Zeitlauf und Speicherbedarf) während der Bearbeitung der Rechnung stark aus. Verwenden einer Datenabbildung kann zur Reduktion der Dimensionalität ohne Verlust der wesentlichen Datenmerkmale und somit zur Minderung der Ressourcenbelastungen während der Bearbeitung der Daten beitragen.

### 1.1 Motivation

Ein wichtiger Konzeptbestandteil der Dimensionalitätsreduktion kommt beim Johnson-Lindenstrauss Lemma vor. Es besagt dass beliebig  $n$  Datenpunkte im höherdimensionalen euklidischen Raum auf eine niedriger Dimension  $k$  abgebildet werden können, ohne Verzerrung des euklidischen Abstands zwischen zwei beliebigen Punkten ein Faktor  $1 \pm \epsilon$  überschritten zu werden (Abbildung 1<sup>\*</sup> [L. JP009]).

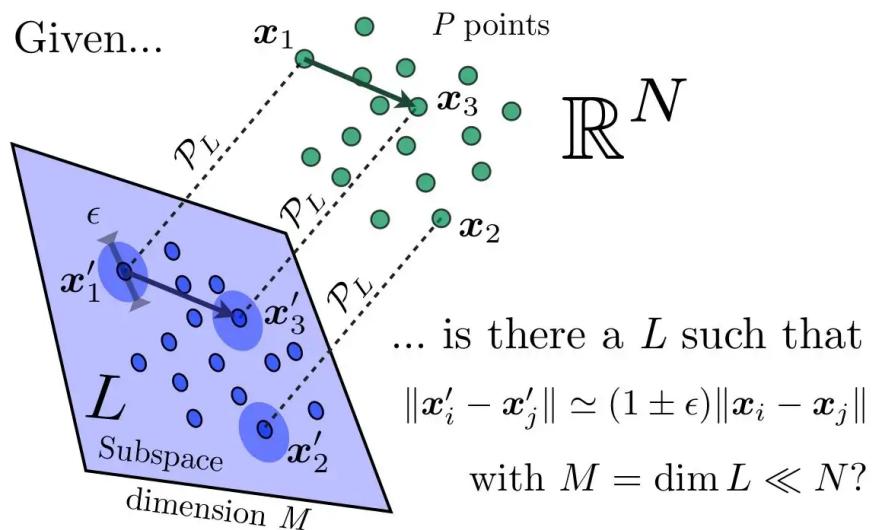


Abbildung 1<sup>\*</sup>: Eine lineare Dimensionalitätsreduktion.

\* Bildquelle von Laurent Jacques

## 2 Grundlagen der Maßtheorie

---

*The sets whose measure we can define by virtue of the preceding ideas we will call measurable sets; we do this without intending to imply that it is not possible to assign a measure to other sets.*

*E. Borel, 1898*

Wir fangen mit der Erörterung einiger elementaren Grundlagen, die Basis für die nachfolgende Theorie nämlich Maßtheorie sind. Die Hauptidee bei der Berechnung des Volumens bzw. des Maßes einer Teilmenge  $\mathbb{R}^d (d \geq 1)$  besteht darin, diese Menge durch Vereinigung andere Menge anzunähern, deren topologische Geometrie einfacher zu verstehen bzw. deren Volumen einfacher zu berechnen sind. Das Wort „Volumen“ bezieht sich angemessen hier um die Maßgröße in  $\mathbb{R}^d$ . Zum Beispiel entspricht diese der Länge für  $d=1$  oder der Fläche für  $d=2$ . Nach dieser Herangehensweise werden wir die Rechtecke und Würfels als die wesentlichen Bauelement der Maßtheorie verwenden. Mit anderen Worten wird in  $\mathbb{R}$  Intervalle und in  $\mathbb{R}^d (d \geq 2)$  Produkten von Intervallen betrachtet und somit kann die Wohldefiniertheit<sup>1</sup> eines Volumens in jeder Dimension als Ergebnis der Produktion aller Intervalllängen unter kartesianische Produkt garantiert werden.

Als Standardnotation [RAR0051] nehmen wir folgenden an:

1. Ein Punkt  $x \in \mathbb{R}^d$  besteht aus einem Tupel von den reellen Zahlen:

$$x = (x_1, x_2, \dots, x_d), \quad x_i \in \mathbb{R}, \forall i \in \{1, \dots, d\}$$

2. Die Addition zwischen zwei Punkten sowie Multiplikation eines Punkts mit einer reellen Zahl sind komponentenweise definiert:

$$x + y = (x_1 + y_1, x_2 + y_2, \dots, x_d + y_d); \lambda x = (\lambda x_1, \lambda x_2, \dots, \lambda x_d), \quad \forall x, y \in \mathbb{R}^d, \lambda \in \mathbb{R}$$

3. Die Norm von  $x$  wird mit  $|x|$  als Standard die Euklidische Norm wie folgt gezeichnet:

$$|x| = (x_1^2 + x_2^2 + \dots + x_d^2)^{\frac{1}{2}}$$

4. Die Menge  $\mathbb{R}^d$  von  $d$ -Tupels mit den oben Eigenschaften (1-3) bilden einen **Vektorraum** über Reellen zusammen ab.

5. Aus den II und III folgt sich die Abstand zwischen zwei Punkten einfach unter dem  $|x - y|$ .

6. Das absolute Komplement einer Menge  $E \subset \mathbb{R}^d$  ist definiert durch

$$E^c = \{x \in \mathbb{R}^d : x \notin E\}$$

7. Für Mengen  $E, F \subset \mathbb{R}^d$  ist folglich das relative Komplement (Differenz) definiert durch

$$E \setminus F = \{x \in \mathbb{R}^d : x \in E, x \notin F\} = E \cap F^c$$

8. Der Abstand zwischen zwei Mengen  $E, F \subset \mathbb{R}^d$  ist definiert durch

$$d(E, F) = \inf |x - y| = \max_{d \in \mathbb{R}} \{d \leq |x - y|\}$$

---

<sup>1</sup> Wohldefiniertheit bezeichnet in der Mathematik die Eigenschaft eines Objekts, eindeutig definiert zu sein.

## 2.1 Offene, Abgeschlossene, und Kompakte Mengen

### Definitionen 2.1.1

1. Der **offene Ball** von  $\mathbb{R}^d$  ist mit dem Mittelpunkt  $x$  und dem Radius  $r$  definiert durch

$$B_r(x) = \{x \in \mathbb{R}^d : |y - x| < r\}$$

2. Eine Menge  $E \subset \mathbb{R}^d$  ist offen, falls Folgendes gilt:

$$\forall x \in E \exists r > 0 : B_r(x) \subset E$$

3. Eine Menge  $E \subset \mathbb{R}^d$  ist abgeschlossen, falls ihr Komplement offen ist.

4. Eine Menge  $E \subset \mathbb{R}^d$  ist abzählbar, falls es finit ist oder existiert eine bijektive<sup>1</sup> Funktion  $f: \mathbb{N} \rightarrow E$ , wobei  $\mathbb{N}$  die Menge der natürlichen Zahlen ist.

5. Eine Menge  $E \subset \mathbb{R}^d$  ist beschränkt, falls sie in einem Ball mit einem endlichem Radius liegt.

6. Eine beschränkte Menge  $E \subset \mathbb{R}^d$  ist kompakt, falls sie auch abgeschlossen ist. Mit anderen Worten  $E \subset \mathbb{R}^d$  ist Kompakt, falls folgende Aussage wahr ist:

Immer wenn  $E$  in der Vereinigung von offenen Mengen liegt, liegt dann  $E$  in der Vereinigung endlich von vielen dieser Mengen. Dies nennt man auch Überdeckungssatz.<sup>2</sup>

$$E \subset \bigcup_j G_j \Rightarrow E \subset \bigcup_{j=1}^N G_{\alpha_j}$$

7. Ein Punkt  $x \in \mathbb{R}^d$  ist **Limit Punkt/ Grenzwert** von  $E \subset \mathbb{R}^d$ , falls Folgendes gilt:

$$\forall r > 0 \exists y \neq x : y \in E \cap B_r(x)$$

8. Ein Punkt  $x \in E \subset \mathbb{R}^d$  ist ein isolierter Punkt, falls Folgendes gilt:

$$\exists r > 0 : \{x\} = B_r(x) \cap E$$

9. Ein Punkt  $x \in E \subset \mathbb{R}^d$  ist ein **innerer** Punkt, falls Folgendes gilt:

$$\exists r > 0 : B_r(x) \subset E$$

10. Eine Menge  $\overset{\circ}{E} \subseteq E \subset \mathbb{R}^d$  heißt die **innere** Menge der  $E$ , falls Folgendes gilt:

$$\overset{\circ}{E} = \{x \in E : B_r(x) \subset E ; r > 0\}$$

11. Eine **abgeschlossene Hülle**  $\bar{E}$  (eng. Closure [FJIR001]) der Menge  $E \subset \mathbb{R}^d$  definiert durch

$$\bar{E} = \{x \in \mathbb{R}^d : x \in E \vee x \text{ ist limit Punkt der } E\}$$

12. Die abgeschlossene Hülle einer Menge ist damit abgeschlossen und eine abgeschlossene Menge enthält alle ihre Limit Punkten.

<sup>1</sup> Bijektivität bedeutet, dass für alle  $e \in E$  genau ein  $n \in \mathbb{N}$  mit  $f(n) = e$  existiert. Also die Elemente von  $E$  und  $\mathbb{N}$  sind gepaart.

<sup>2</sup> Das nennt man auch Heine-Borel's Satz.

13. **Rand** (eng. Boundary) einer Menge  $E \subset \mathbb{R}^d$ ,  $\partial E$  ist die Menge von Punkten, die in der  $\bar{E}$  liegen, aber gleichzeitig nicht zu der inneren Menge/Raum der  $E$ ,  $\mathring{E}$  gehören und wird.

$$\partial E = \bar{E} \setminus \mathring{E} = [\mathring{E} \cup (\mathring{E}^c)]^c$$

14. Eine Menge  $E \subset \mathbb{R}^d$  heißt **perfekt**, falls die Menge  $E$  keinen isolierter Punkt hat.

15. Eine Menge  $E \subset \mathbb{R}^d$  heißt **konvex**, falls  $E = \{x + t(y - x) \mid x, y \in E \wedge t \in [0, 1]\}$ .

### Lemma 2.1.2

Seien Mengen  $F, K \subset \mathbb{R}^d$  abgeschlossen bzw. kompakt und Beiden disjunkt, ist dann  $d(F, K) > 0$ .

*Beweis:* Da  $F$  abgeschlossen ist, existiert für jeden Punkt  $x \in K$ ,  $\delta_x > 0$  mit  $d(x, F) > 3\delta_x^{(*)}$ . Da  $K$  kompakt ist, und  $K$  durch  $\bigcup_{x \in K} B_{2\delta_x}(x)$  überdeckt ist, können wir einen Teilüberdeckung

$\bigcup_{j=1}^N B_{2\delta_j}(x_j)^{(**)}$ . Ist nun  $\delta = \min(\delta_1, \dots, \delta_N)$ , ist dann  $d(K, F) > \delta > 0$ . In diesem Zusammenhang gelten dann für jeden  $x \in K$ ,  $y \in F$  ein  $j$   $|x_j - x| \leq 2\delta_j^{(**)}$  und bzw.  $|y - x_j| \geq 3\delta_j^{(*)}$ . Deshalb

$$|y - x| \geq |y - x_j| - |x_j - x| \geq 3\delta_j - 2\delta_j \geq \delta. \blacksquare$$

## 2.2 Rechtecken, Quadern<sup>1</sup>

Gegeben sei ein abgeschlossenes Rechteck  $R \subset \mathbb{R}^d$  definiert durch das Produkt von  $d$  eindimensionalen abgeschlossene und beschränkte Intervallen

$$R = \prod_{j=1}^n [a_j, b_j] = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_d, b_d],$$

wobei  $a_j \leq b_j \in \mathbb{R}$ ; für alle  $j = 1, 2, \dots, d$ . Dieses kann auch wie gefolgt umformuliert werden:

$$R = \{(x_1, x_2, \dots, x_d) \in \mathbb{R}^d : a_j \leq x_j \leq b_j; j = 1, 2, \dots, d\}.$$

## 2.3 Einleitende Bemerkungen

1. Ein Rechteck ist hier abgeschlossen und ihr Seiten stehen parallel zur Koordinatenachsen.

Präzise sind Rechtecke in  $\mathbb{R}$  (Abbildung 2) die abgeschlossen beschränkten Intervalle,

während sie in  $\mathbb{R}^2$  (Abbildung 3) die übliche vierseitig Rechtecke und in  $\mathbb{R}^3$  (Abbildung 4)

die abgeschlossene Parallelepipede sind.

2. Wir nehmen dabei an, dass die Seitenlängen der Rechteck  $R$ ,  $b_1 - a_1, \dots, b_d - a_d$  sind.

3. Das Volumen des Rechteckes  $R$ ,  $|R|$  ist definiert [RAR0051] durch

$$|R| = (b_1 - a_1) \times (b_2 - a_2) \times \cdots \times (b_d - a_d).$$

Dabei entspricht der Begriff „Volumen“ für  $d = 1$ , der Länge und für  $d = 2$  der Fläche.

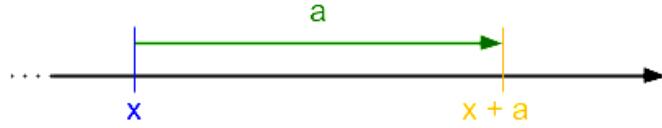
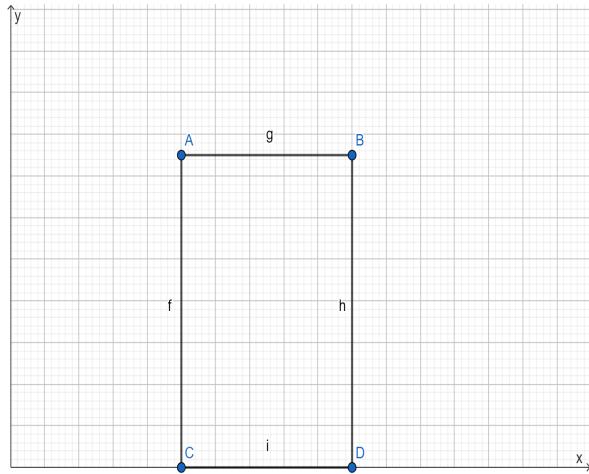
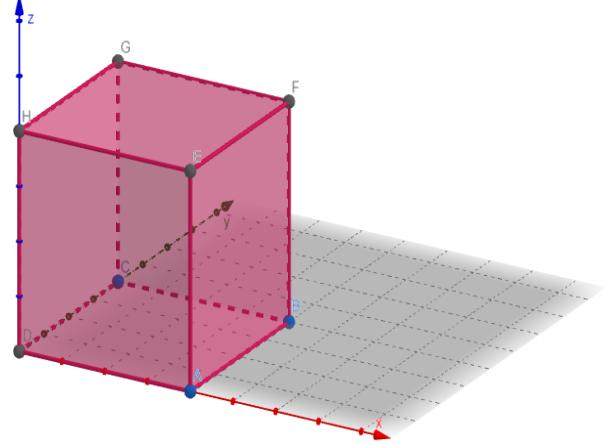
4. Ein offenes Rechteck ist Produkt offener Intervalle und damit Innere Menge/Raum des Rechteckes  $R$ ,  $\mathring{R}$  ist definiert durch

---

1 Eine Verallgemeinerung des Rechtecks und des Quaders auf beliebig viele Dimensionen wird in der Geometrie mit Hyperquader bezeichnet.

$$\overset{\circ}{R} = (a_1, b_1) \times (a_2, b_2) \times \cdots \times (a_d, b_d).$$

5. Ein Kubus/Würfel  $Q$  wird auch als ein Rechteck mit  $b_1 - a_1 = b_2 - a_2 = \cdots = b_d - a_d = l$  betrachtet und damit  $|Q| = l^d$ .

Abbildung 2: Ein Rechteck in  $\mathfrak{R}$ Abbildung 3: Ein Rechteck in  $\mathbb{R}^2$ Abbildung 4: Ein Rechteck in  $\mathbb{R}^3$ 

### Lemma 2.3.1

Ist ein Rechteck  $R$  die Vereinigung von endlich vielen Rechtecken, gilt dann folgende Gleichung:

$$R = \bigcup_{k=1}^N R_k \rightarrow |R| = \sum_{k=1}^N |R_k|$$

**Beweis:** Wir betrachten das Gitter, welche durch eine Verlängerung der Kanten bzw. Seiten aller Rechtecke  $R_1, \dots, R_N$  erstellt wird (Abbildung 5). Aus dieser Konstruktion entstehen endlich viele Rechtecke  $\tilde{R}_1, \tilde{R}_2, \dots, \tilde{R}_M$  und eine Partition  $J_1, \dots, J_N$  von positive ganze Zahlen zwischen eins und  $M$ , sodass die beiden folgenden Vereinigungen unter fast disjunkte Eigenschaft bleiben (Abbildung 6)

$$R = \bigcup_{j=1}^M \tilde{R}_j \wedge R_k = \bigcup_{j \in J_k} \tilde{R}_j.$$

Beispielweise ist für das Rechteck  $R$  ist  $|R| = \sum_{j=1}^M |\tilde{R}_j|$ , weil das Gitter in der Tat die Seiten des  $R$  zerlegt und jedes  $\tilde{R}_j$  aus den Produkten der Intervalle in der Zerlegungen. Wenn wir also die Volumen des  $\tilde{R}_j$  summieren, summieren wir dabei die entsprechende Produkte der Länge der entstandenen Intervallen. Da es ebenso für die anderen Rechtecke  $R_1, \dots, R_N$  gilt, schließen wir das folgende:

$$|R| = \sum_{j=1}^M |\tilde{R}_j| = \sum_{k=1}^N \sum_{j \in J_k} |\tilde{R}_j| = \sum_{k=1}^N |R_k|.$$

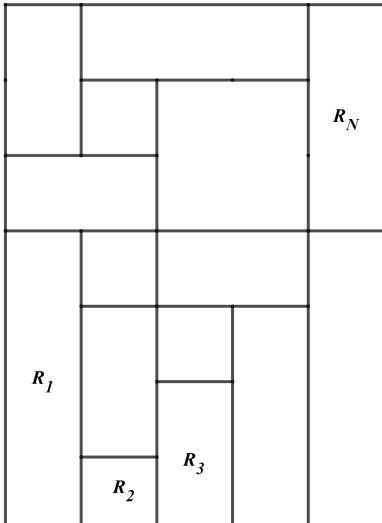


Abbildung 5: Erstellung des Gitters durch die Rechtecke

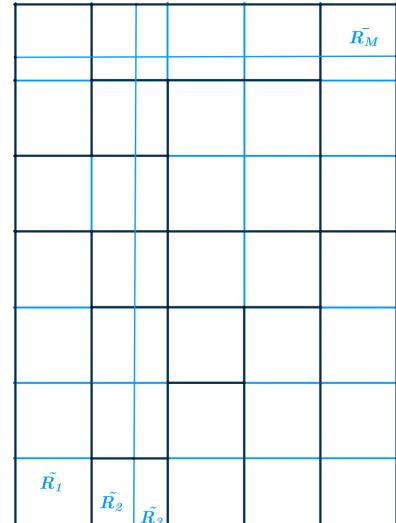


Abbildung 6: Ein konstruiertes Gitter

**Lemma 2.3.2**

Seien  $R, R_1, \dots, R_N$  wie oben definierte Rechtecke mit  $R \subset \bigcup_{k=1}^N R_k$ , es gilt dann:  $|R| \leq \sum_{k=1}^N |R_k|$ .

*Beweis:* Die Hauptidee besteht hier darin, die Rekonstruktion des Gitters diesmal durch Verlängerung aller Seiten der Rechtecke  $R_1, \dots, R_N$  aber auch des Rechtecks  $R$  zu betrachten. Damit müssen die zugehörige  $J_k$  Mengen (siehe den Beweis für das letzte Lemmas) nicht mehr zueinander disjunkt sein. Aus diesem Grund Ansonsten wird alles einfach wie unten gefolgt:

$$0 \leq \sum_{\substack{j=1 \\ j \neq l}}^M |\tilde{R}_j \cap \tilde{R}_l| \leq |\hat{R}| - |R| = \sum_{k=1}^N |R_k| - |R| \Rightarrow |R| \leq \sum_{k=1}^N |R_k|$$

Im folgenden Schnitt wird konkret der Struktur von offenen Mengen in Bezug auf Kubus/Würfel beschrieben.

**Satz 2.3.3**

Jede offene Teilmenge  $G$  von  $\mathfrak{R}^1$  ist eindeutig als abzählbare Vereinigung von disjunkten offenen Intervallen darstellbar.

*Beweis:* Für jedes  $x \in G$  bezeichnen wir  $I_x$  als das größte Intervall, das  $x$  enthält und gleichzeitig in  $G$  liegt. Da  $G$  offen ist, liegt  $x$  in einem kleineren (nicht unbedingt triviale) Intervall. Aus diesem Grund gilt  $a_x < x < b_x$  wobei  $a_x$  und  $b_x$  mit womöglich unbeschränkt Werte definiert durch:

$$a_x = \inf\{a < x : (a, x) \subset G\} \quad \wedge \quad b_x = \sup\{b > x : (x, b) \subset G\}$$

Wenn wir nun  $I_x = (a_x, b_x)$  setzen, haben wir dann aufgrund der Konstruktion  $x \in I_x \wedge I_x \subset G$ .

Daraus folgt:

$$G = \bigcup_{x \in G} I_x.$$

Wir nehmen nun an, dass es mindestens zwei Intervallen  $I_x, I_y$  gäbe, die sich schneiden, somit ihrer Verengung, was folglich auch ein offenes Intervall ist, liegt in  $G$  und enthält  $x$ . Weil  $I_x$  maximal ist muss  $(I_x \cup I_y) \subset I_x$ . Das muss auch für  $I_y$  gelten d.h.  $(I_x \cup I_y) \subset I_y$ . Dies kann nur geschehen, wenn

1 Die reelle Menge  $\mathfrak{R}$  ist nicht abzählbar. Ein elegant Beweis dafür wurde bereits von Herrn Georg Cantor geschrieben.

$I_x = I_y$ . Deshalb müssen zwei beliebig unterschiedliche Intervalle in der Mengenfamilie  $\Lambda = \{I_x\}_{x \in G}$  disjunkt sein. Der Beweis wird vollständig sein, indem wir zeigen, dass genannten Mengenfamilie nur aus abzählbar vielen distinkten Intervallen bestehen. Dies ist jedoch leicht erkennbar, denn jedes offenes Intervall  $I_x$  eine rationale Zahl<sup>2</sup> enthält. Weil verschiedene Intervalle disjunkt sind, müssen sie verschiedene rationale Zahlen enthalten und daher ist  $\Lambda$  abzählbar. Damit ist der Beweis fertig .■

### Korollar 2.3.4

Sollte  $G = \bigcup_{j=1}^{\infty} I_j$  offen, wobei  $I_j$ s disjunkte offene Intervalle sind, sollte Maß von  $G$  gleich  $\sum_{j=1}^{\infty} |I_j|$ .

Da diese Repräsentation eindeutig ist, können wir dies als Definition des Maßes auffassen.

Wir würden dann feststellen, dass immer dann wenn  $G_1$  und  $G_2$  offen und disjunkt sind, ist das Maß ihrer Vereinigung die Summe ihrer Maße ist.

### Satz 2.3.5

jede offene Teilmenge  $G$  von  $\mathbb{R}^d$ ,  $d \geq 1$ , ist als eine abzählbare Vereinigung von fast disjunkt abgeschlossener Kuben bzw. Würfel darstellbar.

*Beweis:* Wir müssen eine abzählbare Mengenfamilie  $\kappa$  von abgeschlossenen Würfeln konstruieren, deren Innenräume disjunkt sind und zwar  $G = \bigcup_{\chi \in \kappa} \chi$ .

Als ersten Schritt betrachten wir das Gitter in  $\mathbb{R}^d$ , welche aus durch eine Verlängerung der Seiten bzw. Kanten mit Länge 1 entsteht, wobei deren Knotenpunkten ganzzahlig sind. Mit anderen Worten betrachten wir das einfache Gitter, welche parallel zum  $\mathbb{Z}^d$ -Axis steht, genauer zu sagen, das Gitter, welche durch planarer Graph bzw. Gittergraph  $\mathbb{Z}^d$  erzeugt wird. Darüber hinaus werden wir auch die Gitter verwenden, die aus Würfeln mit Kantenlängen  $2^{-N}$  erzeugt werden. Diese sind wiederum durch sukzessive Bisektion des ursprünglichen Gitters in Endstand gekommen. Anhand der folgenden Regel werden die Würfel als Teil von  $\kappa$  im ursprünglichen Gittergenerierungsprozess entweder angenommen oder verworfen (Abbildung 7). Die Regel ist beschreiben wie folgt:

- Wenn  $\chi$  vollständig in  $G$  liegt, wird es angenommen.
- Wenn  $\chi$  sowohl  $G$  als auch  $G^C$  schneidet, wird es mit Vorbehalt angenommen.
- Wenn  $\chi$  vollständig in  $G^C$  liegt, wird es verworfen.

In einem zweiten Schritt (Abbildung 8) teilen wir die mit Vorbehalt angenommen Würfel in  $2^d$  Würfel mit der Kantenlänge  $\frac{1}{2}$ ; Dann wiederholen wir unser Verfahren, indem wir die kleineren Würfel annehmen, wenn sie vollständig in  $G$  liegen, indem wir sie mit Vorbehalt annehmen, falls sie sowohl  $G$  als auch  $G^C$  schneidet, oder indem wir sie verworfen, falls sie nur in  $G^C$  liegen.

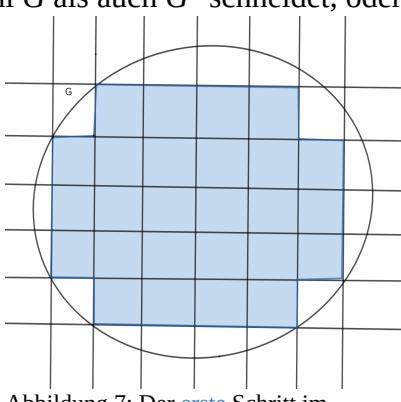


Abbildung 7: Der erste Schritt im Gittergenerierungsprozess für einen offenen Raum in  $\mathbb{R}^2$ .

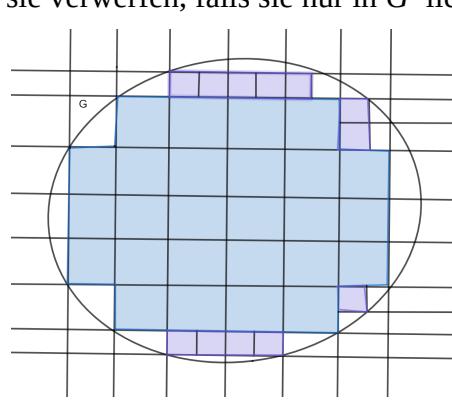


Abbildung 8: Der zweite Schritt im Gittergenerierungsprozess für den offenen Raum in  $\mathbb{R}^2$ .

Diese Vorgehen wird konstruktionsbedingt beliebig wiederholt und somit ist resultierende Mengenfamilie  $\kappa$  aus allen zulässigen Würfeln abzählbar und gleichzeitig besteht aus fast disjunkten Würfeln. Zur Bestätigung dessen, dass die Vereinigung der Würfeln immer noch in  $G$  liegt, stellen wir fest, dass es für  $x \in G$  einen Würfel mit einer Kantenlänge  $2^{-N}$  (gewonnen aus sukzessiven Halbierung. Bzw. Bisektionen des ursprünglichen Gitters) gibt, der  $x$  enthält und der vollständig im  $G$  liegt. Dieser Würfel ist entweder zugelassen oder liegt in einem Würfel, der breites zugelassen worden war. Daraus ergibt sich, dass die Vereinigung aller Würfel in  $\kappa$ ,  $G$  überdeckt. Wenn also  $G = \bigcup_{j=1}^{\infty} R_j$ , wobei die Rechtecke  $R_j$  fast disjunkt sind. Aufgrund vorliegende

Erkenntnisse ist, das folgende Maßgröße für  $G$  auszuweisen:  $G = \sum_{j=1}^{\infty} |R_j|$ . Dies ist verständlich, da

das Grenzvolumen eines Rechtecks gleich null sein sollte, und Überlappung bzw. Überschneidung der Rechtecke nicht zum Volumen  $G$  hinzugefügt wird. Wir stellen dabei fest, dass die oben beschrieben Zerlegung in Würfel nicht eindeutig ist. Darüber hinaus ist dabei eine Unabhängigkeit zwischen die Summe von der Zerlegung nicht sofort zu erschließen. In Sinnes des Wortes sind hierbei Volumen und Fläche für  $\mathbb{R}^d$  mit  $d \geq 2$  sogar wenn es sich um offene Räume handelt. ■

**Vorwort:** Die im nächsten Abschnitt entwickelte allgemeine Theorie(Maßtheorie) enthält in der Tat eine Begriff von Volumen, der den Zerlegungen von offenen Räumen der beiden letzten Sätze entspricht, und für alle Dimensionen (d.h. für alle Offene Räume in  $\mathbb{R}^d$  mit  $d \geq 1$ ) gilt.

## 2.4 Äußeres Maß und ihre Eigenschaften

Der Begriff äußeres Maß ist das erste von zwei wichtigsten Konzepten, die man braucht, um einen Maßtheorie zu entwickeln. Wir beginnen mit der Definition und grundlegenden Eigenschaften für ein äußeres Maß.

Das äußere Maß [RAR0051], wie seine Name andeutet, soll das Volumen einer Menge bzw. eines Raumes  $E \subseteq \mathbb{R}^d$  durch äußere Annäherung beschreiben. Dabei wird der Raum von Quaders bzw. Würfeln überdeckt und wenn die Überdeckung feiner wird, sollte es nicht zur mehr Überlappung der Würfeln führen. Damit sollte das Volumen von  $E$  zur Summe der Volumen der Würfel mehr annähert. Die genaue Definition des äußeren Maßes lautet wie folgt:

### Definition 2.4.1 Äußeres Maß

Für einen Unterraum  $E$  in  $\mathbb{R}^d$  ist ein **äußeres Maß „ $m_*$ “** definiert durch

$$(1) \quad m_*(E) = \inf \sum_{j=1}^{\infty} |Q_j|$$

Dabei ist der Infimum über alle abzählbaren Überdeckungen  $E \subset \bigcup_{j=1}^{\infty} Q_j$  mittels abgeschlossene Quaders berechnet.

### Bemerkung 2.4.2

Das äußere Maß ist immer nicht negative kann aber infinite sein. ( $0 \leq m_*(E) \leq \infty$ ).

### Eigenschaften 2.4.3

1. Das äußere Maß eines Punkts  $p \in \mathbb{R}^d$  ist null. Dies wird sehr deutlich, wenn wir einen Punkt als ein Würfel mit dem Volumen null betrachten, welch sich natürlich überdeckt.
2. Das äußere Maß der leeren Menge/Raum ist ebenfalls null.
3. Das äußere Maß eines abgeschlossenes Würfels ist gleich seinem Volumen.

**Beweis:** Wir nehmen an, dass  $Q$  ein abgeschlossener Würfel im  $\mathbb{R}^d$ . Da  $Q$  selbst überdeckt, wir müssen das Verhältnis  $0 \leq m_*(Q) \leq |Q|$  haben. Deshalb ist genug, die inverse Ungleichung nämlich  $m_*(Q) \geq |Q|$  zu zeigen. Wir betrachten ein beliebige Überdeckung mittels Würfeln  $Q$

$|Q| \subset \bigcup_{j=1}^{\infty} Q_j$  und merken, dass es nur nötig ist, die folgende Ungleichung zu beweisen.

$$(2) \quad |Q| \leq \sum_{j=1}^{\infty} |Q_j|$$

Für ein festes  $\varepsilon > 0$  wählen wir für jeden  $j$  einen offene Würfel  $S_j$ , welche  $Q_j$  beinhaltetet, und  $|S_j| \leq (1 + \varepsilon)|Q_j|$ . Aus der offenen Überdeckung  $\bigcup_{j=1}^{\infty} S_j$  des kompakten Raum  $Q$  können wir eine finite Teilüberdeckung bestimmen, welche nach mögliche Umnummerierung der Rechtecke als  $|Q| \subset \bigcup_{j=1}^N S_j$  beschrieben werden kann. Nimmt man Abschluss bzw. abgeschlossene Hülle der

Würfel  $S_j$ , kann man durch Lemma 2.2 schließen, dass  $|Q| \leq \sum_{j=1}^N |S_j|$ . Daraus ergibt es sich

$$|Q| \leq (1 + \varepsilon) \sum_{j=1}^N |Q_j| \leq (1 + \varepsilon) \sum_{j=1}^{\infty} |Q_j|.$$

Da  $\varepsilon$  beliebig ist, ergibt es sich, dass die Ungleichung (2) erfüllt ist und damit  $m_*(Q) \geq |Q|$ . ■

4. Es sei  $E_1 \subset E_2 \subset \mathbb{R}^d$ , ist dann  $m_*(E_1) \leq m_*(E_2)$  (dies nennt man **Monotonie** für äußeres Maß).

*Beweis:* Diese Ungleichung folgt aus der Beobachtung bzw. Tatsache, dass jede Überdeckung von  $E_2$  durch eine abzählbare Kollektion von Würfeln gleich Überdeckung von  $E_1$  bedeutet. Insbesondere ergibt sich aus der Monotonie, dass jede beschränkte Unterraum (Teilraum) von  $\mathbb{R}^d$  ein finites äußeres Maß hat.

5. Ist  $Q$  ein offenes Würfel, gilt immer noch  $m_*(Q) = |Q|$ .

*Beweis:* Da  $Q$  durch seine eigene abgeschlossene Hülle  $\bar{Q}$  überdeckt ist, und  $|\bar{Q}| = |Q|$ , ergibt es sich direkt, dass  $0 \leq m_*(Q) \leq |Q|$ . Zum Beweis inverser Ungleichung stellen wir folgendes fest:

Ist  $Q_0$  eine abgeschlossene Würfel, welches in  $Q$  liegt, ist dann  $m_*(Q_0) \leq m_*(Q)$  (Monotonie). Diese Ungleichung ist darauf zurückzuziehen, dass jede Überdeckung des Raums  $Q$  durch abzählbare abgeschlossene Würfel gleichzeitig auch eine Überdeckung von  $Q_0$  bedeutet. Daher ist  $|Q_0| \leq m_*(Q)$  und weil die Auswahl  $Q_0$  mit einem Volumen so nah an  $|Q|$  liegen kann, wie das gewünscht ist, müssen wir dann  $m_*(Q) \geq |Q|$  haben. ■

6. Das äußere Maß eines Rechteckes  $R$  ist gleich seinem Volumen.

Wie es bei der Eigenschaft 3 argumentiert, es ist zu sehen  $m_*(R) \geq |R|$ . Zum Beweis der reverse Ungleichung, betrachten wir nun ein Gitter in  $\mathbb{R}^d$ , welches aus Würfeln mit der Kantenlänge  $\frac{1}{k}$  entsteht. Wenn dann  $\mathcal{Q}$  aus der (finiten) Kollektion aller Würfel besteht, die in  $R$  liegen, und  $\mathcal{Q}'$  aus der (finiten) Kollektion aller Würfel besteht, die das Komplement von  $R$ ,  $R^c$  überschneiden, stellen wir zuerst fest,  $R \subset \bigcup_{Q \in \mathcal{Q} \cup \mathcal{Q}'} Q$ . Ein einfache Argumentation ergibt es aus  $\sum_{Q \in \mathcal{Q}} |Q| \leq |R|$ . Darüber hinaus gibt es Würfel<sup>1</sup> in  $\mathcal{Q}'$  mit einem Volumen  $k^{-d}$ , sodass

$$\sum_{Q \in \mathcal{Q}'} |Q| = O\left(\frac{1}{k}\right).$$

Daraus folgt  $\sum_{Q \in (\mathcal{Q} \cup \mathcal{Q}')} |Q| \leq |R| + O\left(\frac{1}{k}\right)$ . Nun als  $k \rightarrow \infty$ , kriegen wir  $m_*(R) \leq |R|$ . ■

---

<sup>1</sup> Die Notation  $f(x) = O(g(x))$  bedeutet, dass für eine Konstante  $C$  und alle  $x$  in einem bestimmten Wertebereich  $|f(x)| \geq C |g(x)|$ . In unserem Fall existiert es hier weniger  $Ck^{d-1}$  Würfel, als  $k \rightarrow \infty$ .

7. Das äußere Maß von  $\mathbb{R}^d$  ist infinite(unendlich).

*Beweis:* Dies folgt aus der Tatsache, dass jede abgeschlossene Hülle von  $\mathbb{R}^d$  gleich eine abgeschlossene Hülle eines beliebigen Würfels  $Q \subset \mathbb{R}^d$  und daher  $m_*(\mathbb{R}^d) \geq |Q|$ . Da  $Q$  kann ein beliebig großes Volumen haben kann, muss  $m_*(\mathbb{R}^d) = \infty$  sein .■

8. Es sei  $E = \bigcup_{j=1}^{\infty} E_j$ , ist dann  $m_*(E) \leq \sum_{j=1}^{\infty} m_*(E_j)$  (Abzählbare Subadditivität)

*Beweis:* Zunächst nehmen wir an, dass  $m_*(E_j) < \infty$ , andernfalls ist die Ungleichung eindeutig erfüllt. Für jede  $\varepsilon > 0$  ergibt es sich aus der Definition des äußeren Maßes für jeden  $j$  eine Überdeckung  $E_j \subset \bigcup_{k=1}^{\infty} Q_{k,j}$ , welche durch abgeschlossene Würfel mit  $\sum_{k=1}^{\infty} |Q_{k,j}| \leq m_*(E_j) + \frac{\varepsilon}{2^j}$  entsteht. Daraus folgt, dass  $E \subset \bigcup_{j,k=1}^{\infty} Q_{k,j}$  eine Überdeckung für den Raum  $E$  durch abgeschlossene Würfel bildet. Deshalb

$$m_*(E) \leq \sum_{j,k=1}^{\infty} |Q_{k,j}| = \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} |Q_{k,j}| \leq \sum_{j=1}^{\infty} (m_*(E_j) + \frac{\varepsilon}{2^j}) = \sum_{j=1}^{\infty} m_*(E_j) + \varepsilon.$$

Da dies für jede  $\varepsilon > 0$  gilt, ist somit der Beweis fertig .■

9. Es sei  $E \subset \mathbb{R}^d$ , ist dann  $m_*(E) = \inf m_*(G)$ , wobei das Infimum alle offene Menge  $G$  erfasst, welche  $E$  beinhalteten.

*Beweis:* Die Monotonie Eigenschaft liefert direkt die Ungleichung  $m_*(E) \leq \inf m_*(G)$ . Zum Zeigen der reversen Ungleichung nehmen wir  $\varepsilon > 0$  und wählen Würfel  $Q_j$  aus, welche  $E \subset \bigcup_{j=1}^{\infty} Q_j$  mit  $\sum_{k=1}^{\infty} |Q_k| \leq m_*(E_j) + \frac{\varepsilon}{2^j}$  erfüllen. Es sei  $\hat{Q}_j$  ein offenes Würfel mit  $|\hat{Q}_j| \leq |Q_j| + \frac{\varepsilon}{2^{j+1}}$ , welches  $Q_j$  beinhaltet. Daher ist  $G = \bigcup_{j=1}^{\infty} \hat{Q}_j$  offen. und aus der Abzählbare Subadditivität(8te)Eigenschaft folgt

$$m_*(G) \leq \sum_{j=1}^{\infty} |m_*(\hat{Q}_j)| = \sum_{j=1}^{\infty} |\hat{Q}_j| \leq \sum_{j=1}^{\infty} (|Q_j| + \frac{\varepsilon}{2^{j+1}}) \leq \sum_{j=1}^{\infty} |Q_j| + \frac{\varepsilon}{2} \leq m_*(E) + \varepsilon.$$

Folglich ist  $\inf m_*(G) \leq m_*(E)$  und damit ist der Beweis fertig .■

10. Es sei  $E = E_1 \cup E_2$  und  $d(E_1, E_2) > 0$ , ist dann  $m_*(E) = m_*(E_1) + m_*(E_2)$ .

*Beweis:* Von der Abzählbare Subadditivität(8te)Eigenschaft wissen wir bereits  $m_*(E) \leq m_*(E_1) + m_*(E_2)$  und somit bleibt die reverse Ungleichung zu zeigen. Dazu wählen wir zuerst  $\delta > 0$  mit  $\delta < d(E_1, E_2)$  aus. Als Nächstes bestimmen wir eine Überdeckung  $E \subset \bigcup_{j=1}^{\infty} Q_j$ , welche durch abgeschlossene Würfel mit  $\sum_{k=1}^{\infty} |Q_k| \leq m_*(E_j) + \varepsilon$  entsteht. Wir können nach der Unterteilung der Würfel  $Q_j$  davon ausgehen, dass jeder  $Q_j$  einen Durchmesser kleiner als  $\delta$  hat. In diesem Fall kann jeder  $Q_j$  höchstens eine der beiden Mengen bzw. Räumen  $E_1$  oder  $E_2$  überschneiden. Bezeichnet man mit  $J_1$  und  $J_2$  die Menge der Indizes  $j$ , für die  $Q_j$   $E_1$  und  $E_2$  überschneidet, ist dann die Schnittmenge  $J_1 \cap J_2$  leer und damit haben wir

$$E_1 \subset \bigcup_{j \in J_1} Q_j \quad \wedge \quad E_2 \subset \bigcup_{j \in J_2} Q_j.$$

Daraus folgt

$$m_*(E_1) + m_*(E_2) \leq \sum_{j \in J_1} |Q_j| + \sum_{j \in J_2} |Q_j| \leq \sum_{j=1}^{\infty} |Q_j| \leq m_*(E) + \varepsilon.$$

Da dies für jede  $\varepsilon > 0$  gilt, ist somit der Beweis fertig .■

11. Es sei die Menge  $E$  aus die abzählbare Vereinigung von fast disjunkte Würfel  $E = \bigcup_{j=1}^{\infty} Q_j$ , ist

$$m_*(E) = \sum_{j=1}^{\infty} |Q_j| .$$

Beweis: Es sei  $\tilde{Q}_j$  ein streng beinhaltet von  $Q_j$  Würfel, für das  $|Q_j| \leq |\tilde{Q}_j| + \frac{\varepsilon}{2^j}$  mit einem festen  $\varepsilon > 0$  gilt. Daraus folgt, dass für jede  $N$  Würfel  $\tilde{Q}_1, \tilde{Q}_2, \dots, \tilde{Q}_N$  disjunkt sind. Mit Betracht des Abstands eines von anderem Würfel und mehrfache Anwendung der Eigenschaft 10 ergibt es sich damit zusammen

$$m_*\left(\bigcup_{j=1}^N Q_j\right) = \sum_{j=1}^N |\tilde{Q}_j| \geq \sum_{j=1}^N \left(|Q_j| - \frac{\varepsilon}{2^j}\right).$$

Da  $\bigcup_{j=1}^N \tilde{Q}_j \subset E$ , wir schließen für jedes positive Integer  $N$ , dass  $m_*(E) \geq \sum_{j=1}^N |\tilde{Q}_j| - \varepsilon$ .

Als  $N \rightarrow \infty$ , bildet für jede  $\varepsilon > 0$  die Ungleichung  $\sum_{j=1}^{\infty} |Q_j| \leq m_*(E) + \varepsilon$  und damit  $\sum_{j=1}^{\infty} |Q_j| \leq m_*(E)$ .

Mit der Kombination der abzählbare Subadditivität Eigenschaft wird also die gewünschte Gleichung erreicht.

#### Bemerkung 2.4.4

Die letzte Eigenschaft (11te Eigenschaft) zeigt, dass wenn eine Menge zu fast disjunkten Würfel zerlegbar ist, ist seine äußeres Maß als Summe des Volumens der Würfel zu berechnen. Insbesondere ergibt es sich aus dem Satz 2.3.5, dass das äußere Maß einer offener Menge gleich der Summe der Volumen der Würfel in einer Zerlegung ist. Außerdem liefert dies auch gleichzeitig einen Beweis für die Unabhängigkeit der Summe von der Zerlegung.

#### Bemerkung 2.4.5

Trotz den letzten zwei Eigenschaften(10 und 11te Eigenschaft) kann man nicht allgemeine schließen, dass es für disjunkte Vereinigung von zwei Teilmengen bzw. Unterräume von  $\mathbb{R}^d$  folgende Aussage gilt:

$$m_*(E_1 \cup E_2) = m_*(E_1) + m_*(E_2).$$

## 2.5 Messbare Mengen und Lebesgue-Maß

Der Begriff Messbarkeit beschränkt sich auf eine Kollektion von Teilmengen in  $\mathbb{R}^d$ , für die das äußere Maß alle gewünschten Eigenschaften inklusive der Additivität(abzählbare Additivität) für disjunkte Vereinigungen von Mengen erfüllt.

#### Definition 2.5.1

Eine Teilmenge  $E$  von  $\mathbb{R}^d$  heißt Lebesgue-messbar oder einfach messbar, wenn für jeden  $\varepsilon > 0$  eine offene Menge  $G$  mit folgenden Eigenschaften existiert:

$$E \subset G \quad \wedge \quad m_*(G \setminus E) \leq \varepsilon.$$

Dies ist vergleichbar mit der neunten Eigenschaft im 2.4.3 , die für alle Mengen  $E$  gilt.

Falls  $E$  messbar ist, definieren wir sein Lebesgue-Maß (oder Maß)  $m(E)$  durch  $m(E) = m_*(E)$ . Damit beerbt ein Lebesgue-Maß all Eigenschaften des äußeren Maß im 2.4.3.

#### Eigenschaften 2.5.2

1. Jede offenen Menge  $G \subset \mathbb{R}^d$  ist messbar.

*Beweis:* Es folgt direkt aus der Definition 2.5.1.

2. Es sei  $m_*(E) = 0$ , ist dann  $E$  messbar. Insbesondere, ist  $F$  eine Teilmenge einer Menge mit äußeres Maß 0, ist  $F$  folglich messbar.

**Beweis:** Gemäß der Eigenschaft 9 in 2.4.3 wissen wir, dass für jede  $\varepsilon > 0$  eine offene Menge  $G$  mit  $E \subset G \wedge m_*(G) \leq m_*(E) + \varepsilon \leq \varepsilon$  existiert. Da  $(G \setminus E) \subset G$ , liefert die Monotonie Eigenschaft (4. Eigenschaft in 2.4.3)  $m_*(G \setminus E) < \varepsilon$ . ■

3. Eine abzählbare Vereinigung von messbaren Mengen ist wiederum messbar.

**Beweis:** Es sei  $E = \bigcup_{j=1}^{\infty} E_j$ , wobei jede  $E_j$  messbar ist. Gegeben sei  $\varepsilon > 0$ . Wir können dann für jedes Index  $j$  eine offene Menge  $O_j$  mit  $E_j \subset O_j \wedge m_*(O_j \setminus E_j) < \frac{\varepsilon}{2^j}$ . Daher wird die Vereinigung  $G = \bigcup_{j=1}^{\infty} O_j$  offen,  $E \subset G$ , und  $(G \setminus E) \subset \bigcup_{j=1}^{\infty} (O_j \setminus E_j)$ . Nun aus Monotonie und Subadditivität Eigenschaften des äußeren Maßes (Eigenschaften 4 und 8 in 2.4.3) folgt

$$m_*(G \setminus E) \leq \sum_{j=1}^{\infty} m_*(O_j \setminus E_j) \leq \varepsilon. ■$$

4. Abgeschlossene Mengen sind messbar.

**Beweis:** Wir zeigen stattdessen, dass kompakte Mengen messbar sind, denn jede abgeschlossene Menge  $F$  kann als die Vereinigung von kompakte Menge beschreiben werden. Also  $F = \bigcup_{k=1}^{\infty} (F \cap B_k)$ . Dabei bezeichnen wir  $B_k$  als ein abgeschlossenes Ball mit dem Radius  $k$  und dem Mittelpunkt Ursprung. Mit der Anwendung der letzten Eigenschaft (Eigenschaft 3) wird dann Beweis fertig. Es sei also  $F$  kompakt (sodass vor allem  $m_*(F) < \varepsilon$  erfüllt). Gemäß der 9. Eigenschaft des äußeren Maßes können wir für jede  $\varepsilon > 0$  eine offene Menge  $G$  mit  $F \subset G \wedge m_*(G) \leq m_*(F) + \varepsilon$  wählen. Da  $F$  abgeschlossen ist, ist der Differenz  $G \setminus F$  offen. Nach dem Satzes 2.3.5 können wir diesen Differenz als eine abzählbare Vereinigung von fast disjunkte Würfel erfassen ( $G \setminus F = \bigcup_{j=1}^{\infty} Q_j$ ).

Für ein feste  $N$  ist die finite Vereinigung  $K = \bigcup_{j=1}^N Q_j$  kompakt. Aus der Lemma 2.1.2 folgt  $d(K, F) > 0$ . Da  $(K \cup F) \subset G$ , ergibt sich mit den Eigenschaften 4, 10, und 11 in 2.4.3 ergibt sich

$$m_*(G) \geq m_*(F) + m_*(K) = m_*(F) + \sum_{j=1}^N m_*(Q_j).$$

Folglich  $\sum_{j=1}^N m_*(Q_j) \leq m_*(G) - m_*(F) \leq \varepsilon$ . Die letzte Ungleichung gilt auch für  $N \rightarrow \infty$ .

Unter Anwendung der Eigenschaft Subadditivität des äußeren Maßes findet man schließlich

$$m_*(G \setminus F) \sum_{j=1}^N m_*(Q_j) \leq \varepsilon. ■$$

5. Das Komplement einer messbaren Menge ist wiederum messbar.

**Beweis:** Es sei  $E$  messbar, können wir dann für jede positive ganze Zahl  $n$  eine offene Menge  $G_n$  mit  $E \subset G_n \wedge m_*(G_n \setminus E) \leq \frac{1}{n}$  wählen. Das Komplement  $G_n^c$  ist abgeschlossen und aus der letzten(4.) Eigenschaft folgt daher, dass es messbar ist. Aus der 3. Eigenschaft ergibt sich es, dass die abzählbare Vereinigung  $S = \bigcup_{n=1}^{\infty} G_n^c$  auch messbar ist. Wir stellen mit (\*) fest, dass  $S \subset E^c$

und  $(E^c \setminus S) \subset (G_n \setminus E)$  mit  $m_*(E^c \setminus S) \leq \frac{1}{n}$  für alle  $n$  existiert. Deshalb ist mit  $n \rightarrow \infty$

$m_*(E^c \setminus S) = 0$  und laut der 2. Eigenschaft ist  $E^c \setminus S$  daher messbar. Da  $S$  und  $(E^c \setminus S)$  beiden messbar, ist damit nach der 3. Eigenschaft ihre Vereinigung nämlich  $E^c$  wiederum messbar. ■

6. Eine abzählbar Schnittmenge von messbare Menge ist wiederum messbar.

**Beweis:** Aus der Eigenschaften 3, und 5 und De-morgansche Gesetze für messbare  $E_j$  folgt

$$\bigcap_{j=1}^{\infty} E_j = \left( \bigcup_{j=1}^{\infty} E_j^c \right)^c. \blacksquare$$

### Satz 2.5.3

Seien  $E_1, E_2, \dots$ , disjunkt messbare Mengen und  $E = \bigcup_{j=1}^{\infty} E_j$ , ist dann  $m(E) = \sum_{j=1}^{\infty} m(E_j)$ .

**Beweis:** Zuerst nehmen wir weiter an, dass jede  $E_j$  beschränkt ist. Folglich kann für jedes  $j$  mittels der Definition der Messbarkeit zum  $E_j^c$  eine abgeschlossene Teilmenge  $F_j$  von  $E_j$  ( $F_j \subseteq E_j$ ) mit  $m_*(E_j \setminus F_j) \leq \frac{\varepsilon}{2^j}$ . Für jede feste  $N$  sind die Mengen  $F_1, \dots, F_N$  kompakt (wegen der Beschränktheit der  $E_j$ ) und disjunkt, sodass  $m\left(\bigcup_{j=1}^N F_j\right) = \sum_{j=1}^N m(F_j)$ . Aus  $\bigcup_{j=1}^N F_j \subset E$  ergibt sich

$$m(E) \geq \sum_{j=1}^N m(F_j) \geq \sum_{j=1}^N m(E_j) - \varepsilon.$$

Damit ist  $m(E) \geq \sum_{j=1}^{\infty} m(E_j)$ , als  $N \rightarrow \infty$  (\*). Nach der Subadditivität Eigenschaft(8. Eigenschaft in

2.4. 3) gilt reverse Ungleichung  $m(E) \leq \sum_{j=1}^{\infty} m(E_j)$ , was zur Gleichung  $m(E) = \sum_{j=1}^{\infty} m(E_j)$  führt.

Wir wählen im Allgemein eine beliebige Folge von Würfel  $\{Q_k\}_{k=1}^{\infty}$ , die zu  $\mathbb{R}^d$  in dem Sinn wächst, dass  $\forall k \geq 1$ ,  $Q_k \subset Q_{k+1}$  und  $\bigcup_{k=1}^{\infty} Q_k = \mathbb{R}^d$ . Wir setzen dann  $S_k = Q_k \setminus Q_{k-1}$  mit  $S_1 = Q_1$  für  $k \geq 2$ .

Wir setzen nun  $E_{j,k} = E_j \cap S_k$  als messbare Mengen, damit ergibt sich  $E = \bigcup_{j,k} E_{j,k}$ . Wir bemerken, dass jede  $E_{j,k}$  beschränkt und die letzte Vereinigung disjunkt ist. Darüber hinaus ist jede Vereinigung  $E_j = \bigcup_{k=1}^{\infty} E_{j,k}$  auch disjunkt. Mit dieser Eigenschaften und durch Erinnerung (\*) ergibt sich

$$m(E) = \sum_{j,k} m(E_{j,k}) = \sum_j \sum_k m(E_{j,k}) = \sum_{j=1}^{\infty} m(E_j). \blacksquare$$

### Bemerkung 2.5.4

Mit dem Satz 2.5.3 ist die abzählbare Additivität des Lebesgue-Maßes auf messbare Mengen festgelegt. Dieses Ergebnis stellt uns die notwendige Zusammenhang zwischen dem Folgenden vor:

- Unserem primären Notation von Volumen ist durch äußeres Maß gegeben,
- Dem genauen Begriff der messbaren Mengen,
- Den abzählbar infiniten Operationen, die auf diesen Mengen erlaubt sind.

### Definition 2.5.5

- Seien  $E_1, E_2, \dots$  eine zu  $E$  wachsende und abzählbare Kollektion(d.h.  $E = \bigcup_{k=1}^{\infty} E_k$  mit  $E_k \subset E_{k+1}$  für alle  $k \geq 1$ , zeichnen wir dies ab jetzt mit  $E_k \nearrow E$
- Seien  $E_1, E_2, \dots$  eine zu  $E$  fallende und abzählbare Kollektion(d.h.  $E = \bigcap_{k=1}^{\infty} E_k$  mit  $E_k \supset E_{k+1}$  für alle  $k \geq 1$ , zeichnen wir dies ab jetzt mit  $E_k \searrow E$ .

### Korollar 2.5.6

(a) Es sei  $E_k \nearrow E$ , ist dann  $m(E) = \lim_{N \rightarrow \infty} m(E_N)$ .

(b) Es sei  $E_k \searrow E$  und  $\exists k \in \mathbb{N}$ :  $m(E_k) < \infty$ , ist dann  $m(E) = \lim_{N \rightarrow \infty} m(E_N)$ .

(a) *Beweis:* Wir setzen  $G_k = E_k \setminus E_{k-1}$  mit  $G_1 = E_1$  für  $k \geq 2$ . Wir stellen bei ihrer Konstruktion fest, dass die Mengen  $G_k$  messbar, disjunkt sind, und dass  $E = \bigcup_{k=1}^{\infty} G_k$ . Daraus

$$\text{folgt } m(E) = \sum_{k=1}^{\infty} m(G_k) = \lim_{N \rightarrow \infty} \sum_k^N m(G_k) = \lim_{N \rightarrow \infty} m\left(\bigcup_{k=1}^N G_k\right) = \lim_{N \rightarrow \infty} m(E_N). \blacksquare$$

(b) *Beweis:* Wir können selbstverständlich annehmen, dass  $m(E_1) < \infty$ . Wir setzen  $G_k = E_k \setminus E_{k+1}$  mit  $E_1 = E \cup \{\bigcup_{k=1}^{\infty} G_k\}$  als disjunkte Vereinigung von messbaren Mengen. Folglich finden wir, dass

$$m(E_1) = m(E) + \lim_{N \rightarrow \infty} \sum_k^N (m(E_k) - m(E_{k+1})) = m(E) + m(E_1) - \lim_{N \rightarrow \infty} m(E_N).$$

$$\text{Weil } m(E_1) < \infty, \text{ sehen wir deutlich } m(E) = \lim_{N \rightarrow \infty} m(E_N). \blacksquare$$

### Bemerkung 2.5.7

Die zweite Teil (b) des Korollars wird ohne die Voraussetzung  $\exists k \in \mathbb{N}: m(E_k) < \infty$  ausbleiben. Dies kann durch Beispiel  $E_n = (n, \infty) \subset \mathfrak{R}$  (für alle  $n \in \mathbb{N}$ ) bestätigt werden.

### Satz 2.5.8

Es sei  $E \subseteq \mathfrak{R}^d$  messbar, gilt die Folgende für alle  $\varepsilon > 0$ :

- i. Es existiert eine offene Menge  $G$  mit  $E \subset G$  und  $m(G \setminus E) \leq \varepsilon$ .
- ii. Es existiert ein abgeschlossene Menge  $F$  mit  $F \subset E$  und  $m(E \setminus F) \leq \varepsilon$ .
- iii. Es existiert eine kompakte Menge  $K$  mit  $K \subset E$  und  $m(E \setminus K) \leq \varepsilon$ , falls  $m(E) < \varepsilon$ .
- iv. Es existiert ein finite Vereinigung  $F = \bigcup_{j=1}^N Q_j$  von abgeschlossenen Würfeln mit  $m(E \Delta F) \leq \varepsilon$ , falls  $m(E) < \varepsilon$ .

Die Notation  $E \Delta F$  steht für die symmetrische Differenz zwischen den Mengen  $E$  und  $F$  wie folgt

$$E \Delta F = (E \setminus F) \cup (F \setminus E)$$

(i) *Beweis:* Dies folgt direkt aus der Definition der Messbarkeit.  $\blacksquare$

(ii) *Beweis:* Wir wissen, dass  $E^c$  messbar ist. Also existiert eine offene Menge  $G$  mit

$$E^c \subset G \wedge m_*(G \setminus E^c) \leq \varepsilon. \text{ Es sei } F = G^c, \text{ ist dann } F \text{ als ein abgeschlossene Menge mit } F = G^c \subset E \wedge E \setminus F = E \cap G = G \setminus E^c. \text{ Damit ist } m(E \setminus F) \leq \varepsilon. \blacksquare$$

(iii) *Beweis:* Zunächst wählen wir eine abgeschlossene Menge  $F \subset E$  mit  $m(E \setminus F) \leq \frac{\varepsilon}{2}$ . Es sei

für jedes positive Integer  $n$ ,  $B_n$  das Ball mit dem Mittelpunkt in der Ursprung. Wir definieren kompakte Mengen  $K_n = (F \cap B_n)$ . Damit ist  $E \setminus K_n$  ein Folge von messbaren Mengen, die fallend auf  $E \setminus F$  ist. Da  $m(E) < \infty$ , finden wir für alle groß genüge  $n$   $m(E \setminus K_n) \leq \varepsilon$ .  $\blacksquare$

(iv) *Beweis:* Es sei  $\{Q_k\}_{k=1}^{\infty}$  eine Mengenfamilie von abgeschlossenen Würfeln mit

$$E \subset \bigcup_{j=1}^{\infty} Q_j \wedge \sum_{j=1}^{\infty} |Q_j| \leq m(E) + \frac{\varepsilon}{2}. \text{ Da } m(E) < \infty, \text{ konvergiert die Summenfolge und daher}$$

$$\text{existiert ein } N > 0 \text{ mit } \sum_{j=N+1}^{\infty} |Q_j| \leq \frac{\varepsilon}{2}. \text{ Falls } F = \bigcup_{j=1}^N Q_j, \text{ ist dann}$$

$$m(E \Delta F) = m(E \setminus F) + m(F \setminus E) \leq m\left(\bigcup_{j=N+1}^{\infty} Q_j\right) + m\left(\bigcup_{j=1}^N Q_j \setminus E\right) \leq \sum_{j=N+1}^{\infty} |Q_j| + \sum_{j=1}^{\infty} |Q_j| \leq \varepsilon. \blacksquare$$

### Invarianzeigenschaften 2.5.9

(a) *Translationsinvarianz:*

Es sei  $E \subset \mathbb{R}^d$  messbar und  $h \in \mathbb{R}^d$ , ist dann auch die Menge  $E_h = E + h = \{x + h \mid x \in E\}$  messbar und  $m(E_h) = m(E + h) = m(E)$ .

Dies ist bei der Erkenntnis aus der Beobachtung einer Würfel einfacher zu sehen. Nach den Eigenschaften des äußeren Maßes (Eigenschaften 1, 3, 5, 10 in 2.4.3) haben wir  $m_*(E_h) = m_*(E)$ .

Um Messbarkeit  $E_h$  zu zeigen, betrachten wir unter Annahme, die die Menge  $E$  messbar ist, eine offene Menge  $G \supset E$  mit  $m_*(G \setminus E) \leq \varepsilon$ . Daher ist Menge  $G_h \supset E_h$  mit  $m_*(G_h \setminus E_h) \leq \varepsilon$  offen.

(b) *Relative Dilatationsinvarianz:*

Es sei  $\delta > 0$  und  $\delta E = \{\delta x \mid x \in E\}$ , können wir mit gleichem Weg wie oben feststellen, dass  $\delta E$  messbar ist, immer wenn  $E$  messbar ist. Und es gilt dabei  $m(\delta E) = \delta^d m(E)$

(c) *Spiegelungsinvarianz:*

Es sei  $-E = \{-x \mid x \in E\}$ , ist dann  $m(-E) = m(E)$ , immer wenn  $E$  messbar ist.

## 2.6 Messbare Funktionen

Die Vorstellung einer Charakteristische Funktion ist der gute Ausgangspunkt für eine messbare Funktion. Diese ist definiert [RAR0051] durch

$$\chi_E(x) := \begin{cases} 1 & \text{falls } x \in E \\ 0 & \text{falls } x \notin E \end{cases}$$

Der nächste Schritt ist ein Übergang zu den Funktionen, die die Grundlage für die Integrations-theorie sind. Für das Riemannsche Integral<sup>1</sup> ist es tatsächlich die Klasse von Treppenfunktionen, jeweils gegeben als eine finite Summe

$$f = \sum_{k=1}^N a_k \chi_{R_k},$$

Wobei die  $R_k$  Rechtecke und  $a_k$  konstant sind.

Die Lebesgue Integralen entsprechen noch allgemeiner Begriff, der durch die sogenannte Einfache Funktion darstellbar ist. Einfache Funktion ist eine finite Summe

$$f = \sum_{k=1}^N a_k \chi_{E_k},$$

wobei jede  $E_k$  eine messbare Menge mit finitem Maß und  $a_k$  konstant sind.

Wir betrachten hier nur reellwertige Funktionen<sup>1</sup> auf  $\mathbb{R}^d$  also  $f: \mathbb{R}^d \rightarrow \mathbb{R}$ . Dabei bezeichnen wir die Funktion  $f$  als endlichwertig (eng. finite-valued), falls  $-\infty < f(x) < \infty$  für alle  $x \in \mathbb{R}^d$ .

### Definition 2.6.1

Ein Funktion  $f$ , welche auf eine Teilmenge  $E \subset \mathbb{R}^d$  definiert ist, heißt **messbar**, falls für jede  $a \in \mathbb{R}$  die Menge  $f^{-1}([-\infty, a)) = \{x \in E \mid f(x) < a\} := \{f(x) < a\}$  messbar ist. Alternative wird durch Messbarkeit  $f^{-1}([-\infty, a]) = \{x \in E \mid f(x) \leq a\} := \{f(x) \leq a\}$  definiert. Äquivalent kann man

- 1 Ein reellwertige Funktion  $f$ , definiert auf ein Intervall  $[0, L]$ , ist *Riemann-Integrierbar*, falls es beschränkt ist, und für jede  $\varepsilon > 0$ , eine Unterteilung  $0 = x_0 < x_1 < \dots < x_{N-1} < x_N = L$  des Intervalls  $[0, L]$  existiert, sodass w   

$$U = \sum_{k=1}^N \left[ \sup_{x_{j-1} \leq x \leq x_j} f(x) \right] (x_j - x_{j-1}) \text{ und } L = \sum_{k=1}^N \left[ \inf_{x_{j-1} \leq x \leq x_j} f(x) \right] (x_j - x_{j-1})$$
 als obere bzw. untere Summe   

$$U - L < \varepsilon$$
 erfüllen.
- 1 Allgemein kann eine messbare Funktion durch zwei gegebenen **Messräumen** abgebildet werden.

stattdessen die Messbarkeit einer Funktion mit Messbarkeit der oberen Mengen verknüpfen:

$f$  ist messbar genau dann, wenn die Menge  $\{x \in E \mid f(x) < a\} := \{f(x) \leq a\}$  für alle  $a$  messbar ist.

Man hat tatsächlich  $\{f(x) \leq a\} = \bigcap_{k=1}^{\infty} \left\{ f(x) < a + \frac{1}{k} \right\}$  und erinnert sich daran, dass eine abzählbare Schnittmenge von messbaren Mengen wiederum messbar ist (6. Eigenschaft 2.5.2). Offenbar ist dann eine endlichwertige Funktion genau dann messbar, wenn die Menge  $\{b < f(x) \leq a\}$  messbar ist.

Wir betrachten, dass es auch analog für andere Richtung  $\{f(x) < a\} = \bigcup_{k=1}^{\infty} \left\{ f(x) \leq a - \frac{1}{k} \right\}$  gilt.

### Eigenschaften 2.6.2

1.  $-f(x)$  ist messbar, wenn auch immer  $f(x)$  messbar ist.

*Beweis:* Dies folgt direkt aus der oberen Definition und der fünften Eigenschaft (2.5.2), wenn man  $f(x)$  durch  $-f(x)$  ersetzt.  $\{-f(x) \leq a\} = \{f(x) \geq -a\} = \{f(x) < -a\}^C$ . ■

2. Die endlichwertige Funktion  $f(x)$  ist genau dann messbar, wenn  $f^{-1}(G)$  für jede offene Menge  $G$  messbar ist, und wenn  $f^{-1}(F)$  für jede abgeschlossene Menge  $F$  messbar ist.

*Beweis:* Diese folgen auch direkt aus der oberen Definition und kann sogar mit der Annahme der Messbarkeit  $f^{-1}(\infty)$  und  $f^{-1}(-\infty)$  erweitert werden. ■

3. i) Es sei  $f(x)$  auf  $\mathfrak{R}^d$  stetig<sup>2</sup>, dann ist  $f(x)$  messbar.

ii) Ferner sei  $f(x)$  endlichwertige und  $g(x)$  stetige Funktion auf  $\mathfrak{R}$ , ist dann  $gof(x)$  auch messbar.

*Beweis:* i) Da die Menge  $G = (-\infty, a) \subset \mathfrak{R}$  ist offen, müssen wir zuerst zeigen, dass  $f^{-1}(G)$  auch offen ist. Es sei  $p \in \mathfrak{R}^d$  und  $f(p) \in G$ . Da  $G$  offen ist, existiert es dann  $\varepsilon > 0$ , sodass  $y = f(x) \in G$ , falls  $d_{\mathfrak{R}}(f(x), f(p)) = |f(x) - f(p)| < \varepsilon$ . Da anderseits  $f$  an der Stelle  $p$  stetig ist, existiert es dann eine  $\delta > 0$ , sodass  $d_{\mathfrak{R}^d}(f(x), f(p)) < \varepsilon$ , falls  $d_{\mathfrak{R}^d}(x, p) = |x - x_0| < \delta$ . Daraus folgt, dass  $x \in f^{-1}(G)$ , sobald  $d_{\mathfrak{R}^d}(x, p) < \delta$ . Somit ist  $f^{-1}(G) := \Phi$  offen.

Mit der ersten Eigenschaft 2.5.2 wird der Beweis für Messbarkeit der  $f(x)$  erreicht. Für das zweite Teil ii) ist analog zum ersten Teil. Dabei sollte beachtet werden, dass Endlichwertigkeit eine notwendige Bedingung für die zweite Aussage ist. ■

4. Es sei  $\{f_n\}_{n=1}^{\infty}$  eine Folge von messbaren Funktionen, sind dann

$$\sup_n f_n(x), \quad \inf_n f_n(x), \quad \limsup_{n \rightarrow \infty} f_n(x), \quad \text{und} \quad \liminf_{n \rightarrow \infty} f_n(x)$$

messbar.

*Beweis:* Für  $\sup_n f_n(x)$  folgt der Beweis direkt aus  $\{\sup_n f_n > a\} = \bigcup_n \{f_n > a\}$ . Dieser kann auch für  $\inf_n f_n(x)$  gelten, da  $\inf_n f_n(x) = -\sup_n (-f_n(x))$ . Für  $\limsup$  und  $\liminf$  folgen damit

$$\limsup_{n \rightarrow \infty} f_n(x) = \inf_k \{ \sup_{n \geq k} f_n \} \quad \text{und} \quad \liminf_{n \rightarrow \infty} f_n(x) = \sup_k \{ \inf_{n \geq k} f_n \}. ■$$

5. Es sei  $\{f_n\}_{n=1}^{\infty}$  eine Kollektion von messbaren Funktionen und  $\lim_{n \rightarrow \infty} f_n(x) = f(x)$ , ist dann  $f$  messbar.

*Beweis:* Der Beweis folgt unmittelbar aus der letzten Eigenschaft, denn

$$\limsup_{n \rightarrow \infty} f_n(x) = f(x) = \liminf_{n \rightarrow \infty} f_n(x). ■$$

2 Eine Funktion  $f(x)$ , definiert auf  $\Omega \subset \mathfrak{R}^d$ , heißt stetig an der Stelle  $x_0 \in \Omega$ , falls es für jede  $\varepsilon > 0$  eine  $\delta > 0$  existiert, sodass  $|f(x) - f(x_0)| < \varepsilon$ , wenn immer  $|x - x_0| < \delta$  für  $x \in \Omega$ .

6. Seien  $f$  und  $g$  messbare Funktionen, dann

- (i) die  $f_k$  sind auch messbar für  $k \geq 1$ .
- (ii)  $f + g$  und  $fg$  sind messbar, falls die beiden Funktionen  $f$  und  $g$  endlichwertige sind.

*Beweis:* Für (i) stellen wir fest, dass

- wenn das Exponent nicht gerade ist, ist dann  $\{f^k > a\} = \{f > a^{1/k}\}$ .
- wenn das Exponent gerade ist, ist dann  $\{f^k > a\} = \{f > a^{1/k}\} \cup \{f < -a^{1/k}\}$ .  
Für (ii) beachten wir, dass  $\{f + g > a\} = \bigcup_{r \in \mathbb{Q}} \{f > a - r\} \cap \{g > r\}$ , wobei hier  $\mathbb{Q}$

als die Menge Rationale Zahlen bezeichnet ist. Aus diesem Grund folgt, dass

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

### Definition 2.6.3

Wir bezeichnen zwei Funktionen  $f$  und  $g$ , definiert auf  $E \subset \mathbb{R}^d$  fast überall gleich(eng. Almost everywhere equal) mit  $f(x) = g(x)$  a.e.  $x \in E$ , falls die Menge  $\{x \in E \mid f(x) \neq g(x)\}$  ein Maß null hat. Also gilt im Allgemein eine Eigenschaft oder Aussage „fast überall(a.e.)“, falls sie trotz der Ausnahme einer Menge mit dem Maß Null wahr ist.

### Bemerkung 2.6.4

- Mit der oberen Definition ist einfach zu sehen, wenn eine Funktion  $f$  messbar und  $f = g$  a.e., ist dann  $g$  auch messbar. Dies folgt direkt aus der Tatsache, dass  $\{f < a\}$  und  $\{g < a\}$  von einer Menge mit dem Maß Null unterscheiden. Somit können auch alle oberen Eigenschaften (1-6) durch geltende a.e. gelockert werden.
- Falls  $f$  und  $g$  fast überall auf eine messbare Teilmenge  $E \subset \mathbb{R}^d$  definiert sind, können dann die Funktionen  $f + g$  und  $fg$  nur auf die Schnittmenge der Definitionsreiche  $f$  und  $g$ . Da die Vereinigung von zwei Mengen mit dem Maß Null wiederum ein Maß null hat, ist  $f + g$  fast überall auf  $E$  definiert.

### Satz 2.6.5

Es sei  $f$  eine nicht-negativ messbare Funktion auf  $\mathbb{R}^d$ , existiert es dann eine wachsende Folge von nicht-negative Einfache-Funktionen  $\{\varphi_k\}_{k=1}^\infty$ , welche Punktweise zur  $f$  konvergiert. Das heißt:

$$\varphi_k(x) \leq \varphi_{k+1}(x) \quad \text{und} \quad \lim_{k \rightarrow \infty} \varphi_k = f(x); \quad \forall x \in \mathbb{R}^d.$$

*Beweis:* Es sei  $Q_N$  eine Würfel unter eine Trunkierung um eine  $N \geq 1$  an der Ursprung. Wir

definieren  $F_N(x) := \begin{cases} f(x) & \text{falls } x \in Q_N \text{ und } f(x) \leq N \\ N & \text{falls } x \in Q_N \text{ und } f(x) > N \\ 0 & \text{sonst} \end{cases}$

Damit ist die  $F_N(x) \rightarrow f(x)$  für alle  $x \in E \subset \mathbb{R}^d$  als  $N \rightarrow \infty$ . Nun teilen wir den Wertebereich der  $F_N$ , nämlich  $[0, N]$  wie unten definiert:

Für feste  $N, M \geq 1$ ,  $E_{l,M} = \left\{ x \in Q_N \mid \frac{l}{M} < F_N(x) \leq \frac{l+1}{M}; 0 \leq l < NM \right\}$ . Somit können

wir eine Einfache-Funktion  $F_{N,M}(x) = \sum_l \frac{l}{M} \chi_{E_{l,M}}(x)$ , welche  $0 \leq F_N(x) - F_{N,M}(x) \leq \frac{1}{M}$  für alle  $x$  in  $\mathbb{R}^d$  befunden sind. Bei der Wahl  $N = M = 2^k$  (mit  $k > 1; k \in \mathbb{N}$ ) und Setzen  $\varphi_k = F_{2^k, 2^k}$  ergibt

sich für alle  $x \in \mathbb{R}^d$ ,  $0 \leq F_N(x) - \varphi_k(x) \leq \frac{1}{2^k}$ . In der Tat  $\varphi_{k-1} = F_{2^{k-1}, 2^{k-1}} \leq F_{2^k, 2^{k-1}} \leq F_{2^k, 2^k} = \varphi_k$ .

Damit ist wegen des Wachstums der Folge  $\varphi_k$  alle gewünschte Eigenschaften im Satz erfüllt. ■

### Bemerkung 2.6.6

Wir betrachten, dass das Ergebnis des letzten Satz für nicht-negative Funktionen gilt, welche als eine erweiterte Funktion\* im Fall eines Limits  $+\infty$  wohl definierbar sind. Dies kann analog für eine nicht-negative Funktion mit

zulässig erweiterte Limit  $-\infty$  gelten.

### Satz 2.6.7

Es sei  $f$  eine messbare Funktion auf  $\mathbb{R}^d$ , existiert es dann eine Folge von Einfache-Funktionen  $\{\varphi_k\}_{k=1}^\infty$ , welche  $|\varphi_k(x)| \leq \varphi_{k+1}(x)$  und  $\lim_{k \rightarrow \infty} \varphi_k = f(x); \forall x \in \mathbb{R}^d$  erfüllt.

Insbesondere gilt  $|\varphi_k| \leq |f(x)|$  (Schränke der  $f$ ) für alle  $x$  und  $k$ .

Beweis: Es sei  $f := f(x) = f^+(x) - f^-(x)$  als eine Zerlegung mit  $f^+(x) = \max\{f(x), 0\}$  und  $f^-(x) = \max\{-f(x), 0\}$ . Dank des letzten Satz wissen wir, dass es wachsende Folgen von nicht-negative Einfache-Funktionen  $\{\varphi_k^{(1)}\}_{k=1}^\infty$  und  $\{\varphi_k^{(2)}\}_{k=1}^\infty$  existiert, welche zur nicht-negativen Funktionen  $f^+$  bzw.  $f^-$  punktweise konvergieren. Aus dem Setzen  $\varphi_k(x) = \varphi_k^{(1)}(x) - \varphi_k^{(2)}(x)$  folgt

$$\lim_{k \rightarrow \infty} \varphi_k = \lim_{k \rightarrow \infty} \varphi_k^{(1)}(x) - \lim_{k \rightarrow \infty} \varphi_k^{(2)}(x) = f^+ - f^- = f; \quad \forall x \in \mathbb{R}^d.$$

Zum Schluss beachten wir, dass  $|f| = |f^+ - f^-| = f; \forall x \in \mathbb{R}^d$  und für jede  $x$  entweder  $\varphi_k^{(1)}(x) \cdot \varphi_k^{(2)}(x) = 0$ . Die letzte Gleichung folgt direkt aus der Tatsache, dass maximal eine der beiden  $f^+(x), f^-(x)$  ungleich null ist,  $0 \leq \varphi_k^{(1)} \leq f^+(x)$ , und  $0 \leq \varphi_k^{(2)} \leq f^-(x)$ . Damit ergibt sich  $|\varphi_k(x)| = \varphi_k^{(1)}(x) + \varphi_k^{(2)}(x)$  offensichtlich als wachsende Folge .■

## 2.7 Das Lebesgue Integral

Wir gehen es in vier Stufen schrittweise Integration vor.

### 2.7.1 Einfache Funktion

Wir erinnern uns vom Kapitel 2.5 an die sogenannten Einfach Funktion  $\varphi(x) = \sum_{k=1}^N a_k \chi_{E_k}(x)$ , wobei

$\varphi(x) = \sum_{k=1}^N a_k \chi_{E_k}(x)$ , wobei die  $E_k$  messbare Menge mit finiten Maß und  $a_k$  konstant sind.

#### Definition 2.7.1.1

Es sei  $\varphi$  eine einfache Funktion mit kanonische Form  $\varphi(x) = \sum_{k=1}^M c_k \chi_{F_k}(x)$ . ist das **Lebesgue**

**Integral** von  $\varphi$  durch  $\int_{\mathbb{R}^d} \varphi(x) dx = \sum_{k=1}^M c_k m(F_k)$  definiert, wobei  $F_k = \{x \in \mathbb{R}^d \mid \varphi(x) = c_k\}$  die

gewünschte kanonische Form repräsentiert. Ferner sei die Menge  $E \subset \mathbb{R}^d$  messbar mit einer finiten Maßgröße, ist  $\varphi(x) \chi_E(x)$  dann auch eine Einfache Funktion mit folgende Definition:

$$\int_E \varphi(x) dx = \int_{\mathbb{R}^d} \varphi(x) \chi_E(x) dx.$$

---

\* Eine erweiterte Funktion ist eine Funktion, deren Wertebereich bis unendlich erweitert wird.

Zur Verdeutlichung der Wahl des Lebesgue Maßes wird manchmal in der Definition das  $dx$  mit einem  $dm(x)$  getaucht. D.h.  $\int_{\mathbb{R}^d} \varphi(x) dm(x)$ .

Im Folgenden werden wir Aussagen(eng. Proposition) zur Eigenschaften von Einfache-Funktionen ohne Beweise zu benenn.

### Proposition 2.7.1.2

Das Integral von oben definierten Einfache-Funktionen erfüllen folgende Eigenschaften:

- i. (Darstellungsunabhängigkeit) Es sei  $\varphi(x) = \sum_{k=1}^N a_k \chi_{E_k}(x)$  ein beliebige Repräsentation von  $\varphi$ , ist dann  $\int \varphi = \sum_{k=1}^N a_k m(E_k)$ .
- ii. (Linearität) Es sei  $\varphi$  und  $\psi$  Einfache-Funktionen und  $a, b \in \mathbb{R}$ , ist dann
$$\int (a\varphi + b\psi) = a \int \varphi + b \int \psi$$
- iii. (Additivität) Es sei  $E$  und  $F$  disjunkte Teilmengen in der  $\mathbb{R}^d$  mit finiter Maßgröße, ist dann
$$\int_{E \cup F} \varphi = \int_E \varphi + \int_F \varphi.$$
- iv. (Monotonie) Es sei  $\varphi < \psi$  Einfache-Funktionen, gilt es dann:
$$\int \varphi \leq \int \psi$$
- v. (Dreieckungleichung) Es sei  $\varphi$  eine Einfache-Funktion, gilt es dann auch für  $|\varphi|$  und
$$|\int \varphi| \leq \int |\varphi|.$$

## 2.7.2 Beschränkte Funktion

### Definition 2.7.2.1

Eine **Stütze** (engl. Support) von messbare Funktion  $f$  ist definiert durch

$$\text{supp}(f) = \{x \mid f(x) \neq 0\}.$$

Ferner ist die Funktion  $f$  auf eine Menge  $E \subset \mathbb{R}^d$  **gestützt**(engl. supported), falls  $f(x) = 0$  immer dann wenn  $x \notin E$ .

### Bemerkung 2.7.2.2

Eine wichtiges Ergebnis vom letzten Kapitel ([Satz 2.6.7](#)) sagt aus, dass wenn die Funktion  $f$  durch  $M$  beschränkt und auf die Menge  $E$  gestützt ist, existiert dann eine Folge  $\{\varphi_n\}_{n=1}^\infty$  von Einfache-Funktionen, welche durch  $M$  beschränkt und auf  $E$  gestützt sind, sodass für alle  $x \in \mathbb{R}^d$ ,  $\varphi_N(x) \rightarrow f(x)$ .

### Lemma 2.7.2.3

Es sei  $f$  ein beschränkte Funktion und gestützt auf eine Menge  $E \subset \mathbb{R}^d$  mit finiter Maßgröße. Ferner Es sei  $\{\varphi_n\}_{n=1}^\infty$  eine Folge von einfache Funktionen beschränkt durch  $M$ , gestützt auf die Menge  $E$ , und  $\varphi_n(x) \rightarrow f(x)$  für a.e.  $x$  gelten dann folgende.

- i. Das  $\lim_{n \rightarrow \infty} \int \varphi_n$  existiert.
- ii. Es sei  $f = 0$  a.e., ist dann  $\lim_{n \rightarrow \infty} \int \varphi_n = 0$ .

**Beweis:** Auf den Beweis<sup>¬</sup> wird wegen Vorkenntnisse wie z.B. über Egorov's Satz und über über Littlewood's Prinzipien verzichtet.

### Definition 2.7.2.3

Für eine beschränkte Funktion, welche auf einer Menge  $E \subset \mathbb{R}^d$  mit finiter Maßgröße gestützt ist, ist ihr **Lebesgue Integral** definiert durch:

$$\int f(x) dx = \lim_{n \rightarrow \infty} \int \varphi_n(x) dx.$$

Dabei ist  $\{\varphi_n\}_{n=1}^{\infty}$  eine Folge von einfache Funktionen mit folgende Eigenschaften

- $|\varphi_n(x)| \leq M$ .
- $\varphi_n$  ist gestützt auf der  $\text{supp}(f)$ .
- $\varphi_N(x) \rightarrow f(x)$  für a.e.  $x$ .

Im Folgenden werden wir Aussagen(eng. Proposition) zur Eigenschaften von Einfache-Funktionen ohne Beweise zu benenn.

### Proposition 2.7.2.2

Seien  $f$  und  $g$  beschränkte Funktionen und gestützt auf die Menge mit finiter Maßgröße, gelten dann folgende Eigenschaften.

- i. (Linearität) Es sei  $a, b \in \mathbb{R}$ , ist dann

$$\int (a f + b g) = a \int f + b \int g.$$

- ii. (Additivität) Es sei  $E$  und  $F$  disjunkte Teilmengen in der  $\mathbb{R}^d$  mit finiter Maßgröße, ist dann

$$\int_{E \cup F} f = \int_E f + \int_F f.$$

- iii. (Monotonie) Es sei  $f < g$ , gilt es dann:

$$\int f \leq \int g.$$

- iv. (Dreieckungleichung) Es sei  $|f|$  beschränkt und gestützt auf eine Menge mit finiter Maßgröße gilt es dann

$$|\int f| \leq \int |f|.$$

### Satz 2.7.2.3 (Beschränkte Konvergenz)

Es sei  $\{f_n\}$  eine Folge von messbaren Funktionen, welche beschränkt durch  $M$  und auf einer Menge  $E \subset \mathbb{R}^d$  mit finiter Maßgröße gestützt sind. Ferner sei  $f_n(x) \rightarrow f(x)$  für a.e.  $x$  als  $n \rightarrow \infty$ . Dann ist  $f$  eine messbare Funktion, welche beschränkt, auf die Menge  $E$  gestützt, und

$$\int |f_n - f| \rightarrow 0 \text{ bzw. } \int |f_n| \rightarrow f \text{ als } n \rightarrow \infty.$$

### Bemerkung 2.7.2.3

Durch den letzten Satz kann man zeigen, dass Riemann-Integrierbare<sup>1</sup> Funktionen auch Lebesgue integrierbar sind.

## 2.7.3 Nicht-Negative Funktionen

Wir setzen nun das Integral von Funktionen fort, die messbar, nicht-negative aber nicht unbedingt beschränkt sind (Erweiterte Funktion\*) [RAR0051].

### Definition 2.7.3.1

Es sei  $f$  eine Funktion, wie die oben beschrieben ist. Sein (erweiterte) **Lebesgue Integral** ist definiert

---

¬ Leser kann aus eigener Interesse im Buch „Princeton Lectures in Analysis“ von Herrschaften M. Stein und Shakarchi die Vorkenntnisse sowie den Beweis finden.

durch

$$\int f(x) dx = \sup_g \int g(x) dx,$$

wobei die Funktion  $g$  ( $0 \leq g \leq f$ ) beschränkt und auf eine Menge mit finiter Maßgröße gestützt ist. Damit ist das Integral der Funktion  $f$  ist endlich oder unendlich. Im Fall  $\int f(x) dx < \infty$  heißt die Funktion  $f$  **Lebesgue integrierbar** oder einfach **integrierbar**.

Analog zur vorherigen Entwicklungsschritten für die Definition des Lebesgue Integrals werden die Eigenschaften für die nicht-negativen Funktionen zum Teil wiederholen.

### Proposition 2.7.3.2

Das Integral von nicht-negative messbare Funktionen besitzen folgende Eigenschaften.

- i. (Linearität) Es sei  $a, b \in \mathbb{R}$ , ist dann

$$\int (a f + b g) dx = a \int f dx + b \int g dx.$$

- ii. (Additivität) Es sei  $E$  und  $F$  disjunkte Teilmengen in der  $\mathcal{R}^d$  mit finiter Maßgröße, ist dann

$$\int_{E \cup F} f dx = \int_E f dx + \int_F f dx.$$

- iii. (Monotonie) Es sei  $f < g$ , gilt es dann:

$$\int f dx \leq \int g dx.$$

- iv. Es sei  $g$  integrierbar und  $0 \leq g \leq f$ , gilt es dann auch für  $f$ .

- v. Es sei  $f$  integrierbar, ist dann  $f(x) < \infty$  a.e.  $x$ .

- vi. Es sei  $\int f dx = 0$ , dann  $f(x) = 0$  fast für jede(a.e.)  $x$ .

### Lemma 2.7.3.3 (Fatou)

Es sei  $\{f_n\}$  eine Folge von messbaren Funktionen mit  $f_n \geq 0$ . Wenn  $\lim_{n \rightarrow \infty} f_n(x) = f(x)$  für a.e.  $x$ , gilt

$$\int f dx \leq \liminf_{n \rightarrow \infty} \int f_n dx.$$

### Korollare 2.7.3.4

Es sei  $f$  nicht-negative messbare Funktion, und  $\{f_n\}$  eine Folge von nicht-negative messbare Funktionen mit  $f_n(x) \leq f(x)$  und  $\lim_{n \rightarrow \infty} f_n(x) = f(x)$  für a.e.  $x$ , gilt dann

$$\lim_{n \rightarrow \infty} \int f_n dx = \int f dx.$$

### Korollare 2.7.3.5 (Monotone Konvergenz Satz)

Es sei  $\{f_n\}$  eine Folge von messbaren Funktionen mit  $f_n \nearrow f$  (2.5). Dann gilt

$$\lim_{n \rightarrow \infty} \int f_n dx = \int f dx.$$

### Korollare 2.7.3.6

Betrachte  $\sum_{k=1}^{\infty} a_k(x)$ , wobei  $a_k(x) \geq 0$  messbar für jede  $k \geq 1$  ist. Dann gilt

$$\int \sum_{k=1}^{\infty} a_k(x) dx = \sum_{k=1}^{\infty} \int a_k(x) dx$$

Ferner sei  $\sum_{k=1}^{\infty} \int a_k(x) dx$  endlich. Dann konvergiert die Summe  $\sum_{k=1}^{\infty} a_k(x)$  für a.e.  $x$ .

## 2.7.4 Allgemeines Fall

### Definition 2.7.4.1

Im Sinnes letzten Schnitt (2.7.3) wir bezeichnen eine reellwertige Funktion  $f$  als **Lebesgue**

**integrierbar**(oder einfach **integrierbar**), falls die nicht-negative Funktion  $|f|$  integrierbar ist.

#### Bemerkung 2.7.4.2

In der Tat wird zur Realisierung der obere Definition die Funktion  $f$  wie unten beschreiben zerlegt:

$$f := f(x) = f^+(x) - f^-(x) \text{ mit } f^+(x) = \max\{f(x), 0\} \text{ und } f^-(x) = \max\{-f(x), 0\}.$$

Nun da die beiden nicht-negativen Funktionen  $f^+, f^- \leq |f|$ , sind die Beiden integrierbar immer wenn die Funktion  $f$  integrierbar ist und damit ist das **Lebesgue Integral** von Funktion  $f$  wohl-definiert durch folgende.

$$\int f = \int f^+ - \int f^-.$$

#### Proposition 2.7.4.3

Das Integral von Lebesgue integrierbare Funktionen ist linear, additive, monoton und erfüllt die Dreieckungleichung.

#### Proposition 2.7.4.4

Es sei  $f$  eine integrierbare Funktion auf  $\mathbb{R}^d$ . Es gelte dann für jede  $\varepsilon > 0$  folgenden:

- i. Es existiert eine Menge endlicher Maßgröße  $B$  (z.B. ein **Ball**), sodass

$$\int_{B^c} |f| < \varepsilon.$$

- ii. Es gibt eine  $\delta > 0$ , sodass

$$\int_{B^c} |f| < \varepsilon \text{ immer, wenn } m(E) < \delta.$$

#### Satz 2.7.4.5

Es sei  $\{f_n\}$  eine Folge von messbaren Funktionen mit  $\lim_{n \rightarrow \infty} f_n(x) = f(x)$  für a.e.  $x$ . Ferner sei  $|f_n(x)| < g(x)$ . es gilt dann folgende

$$\int |f_n - f| \rightarrow 0 \text{ bzw. } \int f_n \rightarrow f \text{ als } n \rightarrow \infty.$$

#### Bemerkung 2.7.7

Generell können wir  $\mathbb{R}^d$  als Produkt  $\mathbb{R}^d = \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ , wobei  $d = d_1 + d_2$ , und  $d_1, d_2 \geq 1$ . Damit kann ein Punkt in der  $\mathbb{R}^d$  in der Form  $(x, y)$  mit  $x \in \mathbb{R}^{d_1}$ ,  $y \in \mathbb{R}^{d_2}$  betrachtet werden. Durch solcher Zerlegung der  $\mathbb{R}^d$  wird der allgemeine Begriff Slice ([Abbildung 9](#)), welche durch Fixierung einer Variable entsteht, verständlich vorstellbar. Für eine Funktion  $f \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$  ist das Slice von  $f$  bzgl. der Variable  $y \in \mathbb{R}^{d_2}$  die Funktion  $f^y$  der  $x \in \mathbb{R}^{d_1}$  definiert durch  $f^y = f(x, y)$ . Analog dazu ist das Slice von  $f$  für ein feste  $x \in \mathbb{R}^{d_1}$  definiert durch  $f_x = f(x, y)$ .

Im Fall einer Menge  $E \subset \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$  sind ihre Slices definiert durch  $E^y = \{x \in \mathbb{R}^{d_1} \mid (x, y) \in E\}$  und  $E_x = \{y \in \mathbb{R}^{d_2} \mid (x, y) \in E\}$

#### Satz 2.7.8<sup>†</sup>

Es sei  $f(x, y)$  integrierbar auf  $\mathbb{R}^{d_2} \times \mathbb{R}^{d_2}$ , ist dann für a.e.  $y \in \mathbb{R}^{d_2}$

- i. Das Slice  $f^y$  integrierbar auf  $\mathbb{R}^{d_1}$ .
- ii. Die Funktion  $F^y = \int_{\mathbb{R}^{d_1}} f^y(x) dx$  integrierbar auf  $\mathbb{R}^{d_2}$ .
- iii.  $\int_{\mathbb{R}^{d_2}} (\int_{\mathbb{R}^{d_1}} f(x, y) dx) dy = \int_{\mathbb{R}^{d_2}} f \cdot F^y$

Es ist klar, dass der Satz in Bezug auf  $x$  und  $y$  symmetrisch ist, sodass wir analog zum Slice  $f^y$  auf

<sup>†</sup> Der Satz ist entwickelt von Herrn Fubini und bekannt genau auch mit diesem Namen.

$\Re^{d_1}$  gleiche für das Slice  $f_x$  auf  $\Re^{d_2}$  für a.e.  $x \in \Re^{d_1}$  aussagen können(d.h. Integrierbarkeit  $f_x$  auf  $\Re^{d_2}$  für a.e.  $x \in \Re^{d_1}$ , also Integrierbarkeit von  $\int_{\Re^{d_2}} f_x(y) dy$  auf  $\Re^{d_1}$  und  $\int_{\Re^{d_1}} (\int_{\Re^{d_2}} f(x, y) dy) dx = \int_{\Re^d} f$

Insbesondere sagt der Satz aus, dass das Integral  $f$  auf  $\Re^d$  durch eine Iteration kleines dimensionalen Integrals und zwar in beliebige Iterationsreihenfolge berechenbar ist. D.h.

$$\int_{\Re^{d_2}} (\int_{\Re^{d_1}} f(x, y) dx) dy = \int_{\Re^{d_1}} (\int_{\Re^{d_2}} f(x, y) dy) dx = \int_{\Re^d} f$$

**Beweis:** Wir werden wegen notwendige Grundkenntnisse aus der Variationsrechnung auf der Beweis verzichten. Einen guten Beweis findet allerdings der Leser mit der Grundkenntnisse im Buch „Princeton Lectures in Analysis“ von Herrschaften M.Stein und Shakarchi [RAR0051].

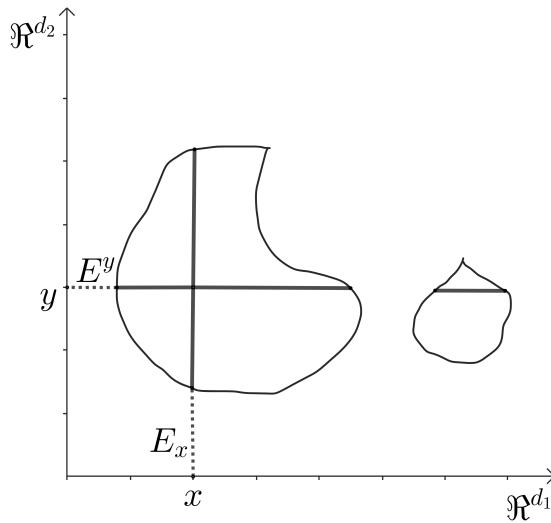


Abbildung 9: Slices  $E^y$  und  $E_x$  (für feste  $x$  und  $y$ ) in einer Menge  $E$ .

### Satz 2.7.9

Es sei  $f(x, y)$  ein nicht negative messbare Funktion auf  $\Re^{d_2} \times \Re^{d_2}$ , gelten dann für a.e.  $y \in \Re^{d_2}$  folgenden:

- i. das Slice  $f^y$  ist messbar auf  $\Re^{d_1}$ .
- ii. Die Funktion  $F^y = \int_{\Re^{d_1}} f^y(x) dx$  ist messbar auf  $\Re^{d_2}$ .
- iii.  $\int_{\Re^{d_2}} (\int_{\Re^{d_1}} f(x, y) dx) dy = \int_{\Re^d} f(x, y) dx dy$  im Sinne einer Extension.

In der Praxis wird dieser Satz<sup>†</sup> oft in Verbindung mit dem Fubini'S Satz verwendet. In der Tat ist für eine gegebene messbare Funktion  $f$  auf  $\Re^d$  Rechnung  $\int_{\Re^d} f$  gefragt. Zur Begründung der iterierten

Integration wenden wir den Satz zuerst für  $|f|$ . Damit können wir ungehindert die iterierte Integralen der nicht negative Funktionen berechnen. Wenn Diese finite sind, garantiert der Satz die Funktion  $f$  ist integrierbar nämlich  $\int |f| < \infty$ .

<sup>†</sup> Der Satz wurde erstmals von Tonelli formuliert.

**Beweis:** Wir betrachten folgende Trunkierung:

$$f_k(x, y) := \begin{cases} f(x, y) & \text{falls } |(x, y)|, f(x) < k \\ 0 & \text{sonst} \end{cases}$$

Jede  $f_k$  ist integrierbar, und durch ersten Teil (i) des Fubini's Satz existiert es ein Menge  $E \subset \mathbb{R}^{d_2}$  mit dem Maß 0, sodass das Slice  $f_k(x)$  für alle  $y \in E_k^c$  messbar. Beim Einsatz  $E = \bigcup_{k=1}^{\infty} E_k$  finden wir nun, dass  $f^y(x)$  für alle  $y \in E^c$  und alle  $k$  messbar ist. Ferner  $m(E) = 0$ . Da  $f_k^y \nearrow f^y$ , folgt aus dem Monotone Konvergenz Satz für jede  $y \notin E$ :

$$F_k^y = \int_{\mathbb{R}^{d_1}} f_k^y(x) dx \nearrow \int_{\mathbb{R}^{d_1}} f^y(x) dx = F^y \text{ als } k \rightarrow \infty.$$

Durch den Fubini's Satz ist  $F_y^k$  messbar für alle  $y \in E^c$  und daher es ist gleich  $F_y$ . Mit der anderen Anwendung von Monotone Konvergenz ergibt sich

$$(1) \quad \int_{\mathbb{R}^{d_2}} \left( \int_{\mathbb{R}^{d_1}} f_k(x, y) dx \right) dy \rightarrow \int_{\mathbb{R}^{d_2}} \left( \int_{\mathbb{R}^{d_1}} f(x, y) dx \right) dy.$$

Aus dem letzten Teil des Fubini's Satz (iii) wissen wir schon

$$(2) \quad \int_{\mathbb{R}^{d_2}} \left( \int_{\mathbb{R}^{d_1}} f_k(x, y) dx \right) dy = \int_{\mathbb{R}^d} f_k.$$

Mit dem Monotone Konvergenz Satz ergibt sich

$$(3) \quad \int_{\mathbb{R}^d} f_k \rightarrow \int_{\mathbb{R}^d} f.$$

Mit der Kombination (1), (2), und (3) wird der Beweis vollständig. ■

### Korollar 2.7.10

Es sei  $E$  eine messbare Menge in  $\mathbb{R}^{d_2} \times \mathbb{R}^{d_2}$ , dann ist für a.e.  $y \in \mathbb{R}^{d_2}$  das Slice

$$E^y = \left\{ x \in \mathbb{R}^{d_1} \mid (x, y) \in E \right\}$$

ein messbare Teilmenge in der  $\mathbb{R}^{d_1}$ . Weiteres ist  $m(E^y)$  ein messbare Funktion von  $y$  und

$$m(E) = \int_{\mathbb{R}^{d_2}} m(E^y) dy.$$

**Beweis:** Dies folgt direkt beim Setzen der Funktion  $f = \chi_E$  im ersten Teil des letzten Satzes.

### Proposition 2.7.11

Es sei  $E = E_1 \times E_2$  (als **Produktion Menge**) mit  $E_j \subset \mathbb{R}^{d_j}$  und  $m_*(E_2) > 0$  messbar. Dann ist die  $E_1$  messbar.

**Beweis:** Mit dem Korollar 2.7.10 ergibt es sich für a.e.  $y \in \mathbb{R}^2$ , dass die Slice-Funktion

$$f^y := (\chi_{E_1 \times E_2})^y(x) = \chi_{E_1}(x) \chi_{E_2}(y)$$

als eine Funktion von  $x$  messbar ist. Das Zeigen des Existenz solcher  $y$  wird mit  $m_*(E_2) > 0$  begründet. In der Tat zeigt man für die Teilmenge  $F \subset \mathbb{R}^{d_2}$ , wo das Slice  $E^y$  messbar ist. Damit ist  $m(F^c) = 0$ . Aus der Zerlegung  $E_2$  in der zwei getrennte Teilmengen  $(E_2 \cap F)$  und  $(E_2 \cap F^c)$  folgt

$$0 < m_*(E_2) \leq m_*(E_2 \cap F) + m_*(E_2 \cap F^c) = m_*(E_2 \cap F). \blacksquare$$

**Lemma 2.7.12**

Es Seien Mengen  $E_1 \subset \mathbb{R}^{d_1}$  und  $E_2 \subset \mathbb{R}^{d_2}$ . Es gilt dann

$$m_*(E_1 \times E_2) \leq m_*(E_1)m_*(E_2),$$

mit dem Verständnis, dass  $m_*(E_1 \times E_2) = 0$ , falls äußeres Maß einer der  $E_j$  gleich null ist.

**Proposition 2.7.13**

Es seien Mengen  $E_1 \subset \mathbb{R}^{d_1}$  und  $E_2 \subset \mathbb{R}^{d_2}$  messbar. Dies gilt dann auch für  $E = E_1 \times E_2$ . Mit dem Verständnis, dass  $m_*(E_1 \times E_2) = 0$ , falls äußeres Maß einer der  $E_j$  gleich null ist, gilt es ferner

$$m(E) = m(E_1)m(E_2).$$

*Beweis:* Im Fall der Messbarkeit der  $E$  folgt die Behauptung über die  $m(E)$  direkt aus dem Korollar 2.7.10. Da jede Menge  $E_j$  messbar ist, existiert es Mengen  $G_j \subset \mathbb{R}^{d_j}$  als eine abzählbar Schnittmenge von offene Mengen\* mit  $G_j \subset E_j$  und  $m_*(G_j \setminus E_j) = 0$  für jede  $j = 1, 2$ . Die letzte Gleichung folgt aus der Messbarkeit der  $E_j$  und daher ist die  $G = (G_1 \times G_2) \in (\mathbb{R}^{d_1} \times \mathbb{R}^{d_2})$  und somit

$$G = (G_1 \times G_2) \setminus (E_1 \times E_2) \subset ((G_1 \setminus E_1) \times G_2) \cup (G_1 \times (G_2 \setminus E_2)).$$

Durch letzten Lemma stellen wir fest, dass  $m_*(G \setminus E) = 0$  und damit ist die  $E$  messbar.  $\blacksquare$

**Korollar 2.7.14**

Es sei  $f$  eine messbare Funktion auf  $\mathbb{R}^{d_1}$ . Die Funktion  $\tilde{f}$ , definiert durch  $\tilde{f}(x, y) = f(x)$  ist messbar in  $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ .

*Beweis:* Wir nehmen an, dass es sich um eine reellwertige Funktion  $f$  handelt und stellen gleichzeitig fest, dass die  $E_1 = \{x \in \mathbb{R}^{d_1} \mid f(x) < a, a \in \mathbb{R}\}$  messbar ist. Da

$$\{(x, y) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \mid \tilde{f}(x, y) < a\} = E_1 \times \mathbb{R}^{d_2},$$

zeigt die letzte Proposition, dass  $\forall a \in \mathbb{R}$  die  $E_1 \times \mathbb{R}^{d_2}$  messbar ist. Damit ist die  $\tilde{f}(x, y)$  eine messbare Funktion auf  $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ .  $\blacksquare$

**Korollar 2.7.15**

Es sei  $f(x)$  eine nicht-negative Funktion auf  $\mathbb{R}^d$  und  $A = \{(x, y) \in \mathbb{R}^d \times \mathbb{R} \mid 0 \leq y \leq f(x)\}$ . Es gelten dann folgende Aussage:

- i.  $f$  ist nur dann messbar auf  $\mathbb{R}^d$ , wenn  $A$  messbar auf  $\mathbb{R}^{d+1}$ .
- ii. Im Fall der Erfüllung der Bedingung in der (i) stellt man fest

$$\int_{\mathbb{R}^d} f(x) dx = m(A).$$

*Beweis:* Wenn  $f$  messbar auf  $\mathbb{R}^d$  ist, garantiert die letzte Proposition, dass die Funktion  $F(x, y) = y - f(x)$

auf  $\mathbb{R}^{d+1}$  messbar ist. Damit wird direkt klar, dass die Menge  $A = \{y \geq 0\} \cap \{F(x) \leq 0\}$  auch messbar ist.

Umgekehrt sei es  $A$  messbare Menge. Wir stellen fest, dass für jede  $x \in \mathbb{R}^d$  das Slice  $A_x = \{y \in \mathbb{R} \mid (x, y) \in A\}$  eine abgeschlossene Segmentierung, nämlich  $A_x = [0, f(x)]$  ist. Im Folge des Korollars 2.7.10 ergibt sich mit dem Tauschen der Rolle  $x$  und  $y$  die Messbarkeit von  $m(A_x) = f(x)$ . Aus diesem Grund folgt

---

\* Solche Menge( $G_\delta$ ) gehören zur einfachste sogenannte Borelmenge.

$$m(E) = \int \chi_A(x, y) dx dy = \int_{\mathbb{R}^d} m(A) dx = \int_{\mathbb{R}^d} f(x) dx. \blacksquare$$

**Bemerkung 2.7.16**

Generell kann man leicht zeigen, dass für eine messbare Funktion  $f$  auf  $\mathbb{R}^d$  die Funktion  $\tilde{f}(x, y) = f(x - y)$  auf  $\mathbb{R}^d \times \mathbb{R}^d$  messbar ist.

## 2.8 Die Brunn-Minkowski Ungleichung

**Definition 2.8.1**

Seien  $A, B \subset \mathbb{R}^d$  messbar, ihr Summe<sup>2</sup> (Abbildung 10<sup>3</sup>) definiert durch

$$A + B = \{x \in \mathbb{R}^d \mid x = x' + x''; x' \in A \wedge x'' \in B\}.$$

**Bemerkung 2.8.2**

Seien die  $A, B \subset \mathbb{R}^d$ , und  $A + B$  messbar. Wir können durch einfache Beispiel von Cantor-Mengen wie  $A = B = C = \left\{ \sum_{j=1}^{\infty} a_k 3^{-k} \mid a_k \in \{0, 2\}, \forall k \right\}$  sehen, dass  $m(A) = m(B) = 0$  aber  $m(A + B) \neq 0$ , denn  $A + B = C + C = \left\{ \sum_{j=1}^{\infty} (a_k + b_k) 3^{-k} \mid a_k, b_k \in \{0, 2\}, \forall k \right\} \supseteq [0, 2]$ . Damit ist leicht zu erkennen, dass es nicht möglich ist, eine obere Schranke für  $m(A + B)$  in Form von  $m(A)$  und  $m(B)$  zu bestimmen.

Im Folgenden werden wir zeigen, dass im Gegenteil eine untere Schranke für  $m(A + B)$  in Form von  $m(A)$  und  $m(B)$  berechnet werden kann.

**Bemerkung 2.8.3**

Seien  $A, B \subset \mathbb{R}^d$  messbar, folgt daraus nicht notwendigerweise, dass  $A + B$  auch messbar ist. Als konkretes Beispiel wählen wir in  $\mathbb{R}^2$ ,  $A = \{0\} \times [0, 1] \wedge B = \mathbb{N} \times \{0\}$ , wobei  $\mathbb{N}$  ein nicht messbare Menge in  $\mathbb{R}$ , damit ist  $A + B = \mathbb{N} \times \mathbb{N}$  nicht messbar.

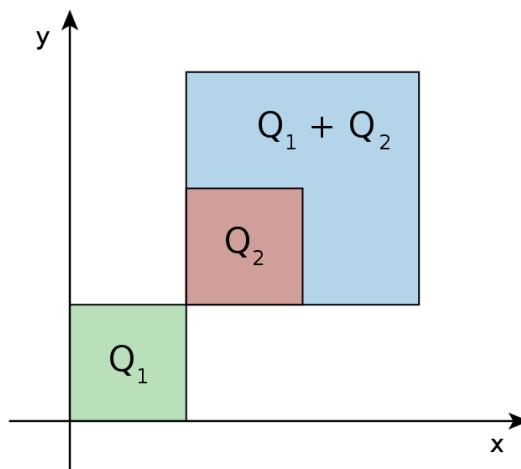


Abbildung 10<sup>3</sup>: Minkowski Addition  $Q_1 + Q_2 = [0, 1]^2 + [1, 2]^2 = [1, 3]^2$ .

**Lemma 2.8.5**

Es sei  $x_j$  nicht negative Zahlen ( $x_j \geq 0$ ), gilt dann  $\frac{1}{d} \sum_{j=1}^d x_j \geq \left( \prod_{j=1}^d x_j \right)^{\frac{1}{d}}$  für jedes  $j$ .

*Beweis:* Dies ist durch ein sukzessives Ersetzen von Elementen leicht zu zeigen. Wir schreiben

2 Auch wird Minkowski Addition genannt.

3 Bildquelle von Wikipedia

schreiben die Ungleichung  $\alpha = \frac{x_1 + x_2 + \dots + x_d}{d} \geq \sqrt[d]{x_1 x_2 \cdots x_d} = \beta$  um und dabei betrachten wir, dass die Gleichung nur für  $x_1 = x_2 = \dots = x_d$  gelten sollte. Falls alle Zahlen nicht gleich sind, existiert dann  $x_i, x_j$  mit  $x_i < \alpha < x_j$ . Es bleibt beim Ersetzen  $x_i$  durch  $\alpha$  und  $x_j$  durch  $(x_i + x_j - \alpha)$  den arithmetischen Mittelwert im Gegenteil des geometrischen Mittelwerts unverändert. In der Tat wird damit der geometrische Mittelwert erhöht, denn  $\alpha(x_i + x_j - \alpha) - x_i x_j = (\alpha - x_i)(x_j - \alpha) > 0$ . Falls immer noch alle Zahlen nicht gleich sind, wiederholen wir den oberen Ersatz weiter. Nach höchstens  $(d - 1)$  solchen Schritten werden die allen Zahlen durch  $\alpha$  ersetzt, während sich der geometrische Mittelwert weiterhin erhöht. Damit ergibt sich nach dem letzten Schritt  $\sqrt[d]{x_1 x_2 \cdots x_d} = \sqrt[d]{\alpha \alpha \cdots \alpha} = \alpha$ .

### Satz 2.8.6

Seien  $A, B \subset \mathbb{R}^d$  und ihr Summe  $A + B$  messbar, es gilt dann

$$(3) \quad (m(A + B))^{\frac{1}{d}} \geq (m(A))^{\frac{1}{d}} + (m(B))^{\frac{1}{d}}.$$

*Beweis:* Zuerst nehmen wir an, dass  $A$  und  $B$  Rechtecke mit den Kantenlängen  $\{a_j\}_{j=1}^d$  und  $\{b_j\}_{j=1}^d$  sind, dann ist nur

$$(2) \quad \left( \prod_{j=1}^d (a_j + b_j) \right)^{\frac{1}{d}} \geq \left( \prod_{j=1}^d a_j \right)^{\frac{1}{d}} + \left( \prod_{j=1}^d b_j \right)^{\frac{1}{d}}$$

zu zeigen. Durch Homogenität der Rechtecken kann die letzte Ungleichung zum speziellen Fall, nämlich  $(a_j + b_j) = 1$  für jedes  $j$ , reduziert werden, indem wir  $a_j, b_j$  durch  $\lambda_j a_j, \lambda_j b_j$  mit  $\lambda_j > 0$  ersetzen. Damit brauchen wir nur die  $\lambda_j = (a_j + b_j)^{-1} = \frac{1}{a_j + b_j}$ . Dies ist allerdings durch die

Division  $\left( \prod_{j=1}^d (a_j + b_j) \right)^{\frac{1}{d}} \neq 0$  von beiden Seiten der Ungleichung (2) mehr anschaulicher. Das heißt  $1 \geq \left( \prod_{j=1}^d \frac{a_j}{a_j + b_j} \right)^{\frac{1}{d}} + \left( \prod_{j=1}^d \frac{b_j}{a_j + b_j} \right)^{\frac{1}{d}}$

Mit dieser Reduktion wird die Ungleichung (2) als direkte Folge der

arithmetische-geometrische Ungleichung  $\frac{1}{d} \sum_{j=1}^d x_j \geq \left( \prod_{j=1}^d a_j \right)^{\frac{1}{d}}, x_j \geq 0$  erreicht, indem wir mit dem

Rücksicht der oberen Reduktion,  $x_j = a_j$  und  $x_j = b_j$  setzen. Also

$$\left( \prod_{j=1}^d \frac{a_j}{a_j + b_j} \right)^{\frac{1}{d}} \leq \frac{1}{d} \sum_{j=1}^d \frac{a_j}{a_j + b_j}, \text{ und } \left( \prod_{j=1}^d \frac{b_j}{a_j + b_j} \right)^{\frac{1}{d}} \leq \frac{1}{d} \sum_{j=1}^d \frac{b_j}{a_j + b_j},$$

und daher durch Lemma 2.8.5

$$\left( \prod_{j=1}^d \frac{a_j}{a_j + b_j} \right)^{\frac{1}{d}} + \left( \prod_{j=1}^d \frac{b_j}{a_j + b_j} \right)^{\frac{1}{d}} \leq \frac{1}{d} \sum_{j=1}^d \frac{a_j}{a_j + b_j} + \frac{1}{d} \sum_{j=1}^d \frac{b_j}{a_j + b_j} = \frac{1}{d} \sum_{j=1}^d \frac{a_j + b_j}{a_j + b_j} = 1.$$

Als Nächstes betrachten wir den Fall, bei dem die Mengen  $A$  und  $B$  die Vereinigung von endlich vielen Rechtecken mit den disjunkten Inneren. Wir werden diesen Fall durch Induktion auf die Anzahl der Rechtecke in der  $A$  und der  $B$  nämlich  $n$  beweisen. Dabei ist wichtig zu beachten, dass bei der unabhängigen Translation (Verschiebung) der Mengen  $A$  und  $B$  die Ungleichung (1) unverändert bleiben wird. In der Tat führt der Ersatz  $A$  und  $B$  durch  $A_h = A + h$  bzw.  $B_h = B + h'$

zum Ersatz  $A + B$  durch  $A + B + h + h'$  und somit ändern sich auch die entsprechenden Maße nicht (Translationsinvarianz). Wir wählen nun zur Zusammenstellung der  $A$  als Kollektion ein disjunkte Rechteckpaare  $R_1$  und  $R_2$  aus und stellen fest, dass  $R_1$  und  $R_2$  durch einer Koordinatendarstellung einer Hyperebene voneinander getrennt werden können. Aus diesem Grund können wir für manche  $j$  Voraussetzen, dass  $R_1$  und  $R_2$  durch geeignetes  $h$  in  $A_- = A \cap \{x_j \leq 0\}$  bzw. in  $A_+ = A \cap \{x_j \geq 0\}$  liegen. Daraus folgt, daß  $A_-$  und  $A_+$  höchstens  $(n - 1)$  Rechtecke enthalten können. Danach stellen wir die Translation der Menge  $B$  so ein, dass  $B_- = B \cap \{x_j \leq 0\}$

und  $B_+ = B \cap \{x_j \geq 0\}$  die Gleichung  $\frac{m(B_\pm)}{m(B)} = \frac{m(A_\pm)}{m(A)}$  erfüllen.

Es gilt allerdings, dass  $A + B \supset (A_+ + B_+) \cup (A_- + B_-)$ , und die Vereinigung in der rechten Seite unbedingt disjunkt sind (die Beiden Teil der Vereinigung liegen in der unterschiedlichen Halbräumen). Daraus folgt durch Induktionshypothese

$$\begin{aligned} m(A_+ + B_+) &\geq (m(A_+)^{\frac{1}{d}} + m(B_+)^{\frac{1}{d}})^d, \text{ und } m(A_- + B_-) \geq (m(A_-)^{\frac{1}{d}} + m(B_-)^{\frac{1}{d}})^d, \text{ dass} \\ m(A + B) &\geq m(A_+ + B_+) + m(A_- + B_-) \geq \left[ m(A_+)^{\frac{1}{d}} + m(B_+)^{\frac{1}{d}} \right]^d + \left[ m(A_-)^{\frac{1}{d}} + m(B_-)^{\frac{1}{d}} \right]^d \\ &= m(A_+) \left[ 1 + \left( \frac{m(B_+)}{m(A_+)} \right)^{\frac{1}{d}} \right]^d + m(A_-) \left[ 1 + \left( \frac{m(B_-)}{m(A_-)} \right)^{\frac{1}{d}} \right]^d = (m(A_+) + m(A_-)) \left[ 1 + \left( \frac{m(B)}{m(A)} \right)^{\frac{1}{d}} \right]^d \\ &= \left[ m(A)^{\frac{1}{d}} + m(B)^{\frac{1}{d}} \right]^d. \end{aligned}$$

Damit ergibt sich die Ungleichung (1), falls  $A$  und  $B$  beiden finite Vereinigung von Rechtecken mit disjunkten Inneren sind. Dies führt anschließend sofort zu dem Ergebnis, als  $A$  und  $B$  offene Mengen mit finitem Maß sind. Tatsächlich können wir nach dem Satzes 2.3.5 für jedes  $\varepsilon > 0$  eine Vereinigungen von fast disjunkten Rechtecken  $A_\varepsilon$  und  $B_\varepsilon$  finden, sodass  $A_\varepsilon \subset A$ ,  $B_\varepsilon \subset B$ , mit  $m(A) \leq m(A_\varepsilon) + \varepsilon$ , und  $m(B) \leq m(B_\varepsilon) + \varepsilon$ . Da  $A + B \supset A_\varepsilon + B_\varepsilon$ , ergibt sich die Ungleichung (1) aus einem Übergang zum einem Grenzwert für die  $A_\varepsilon$  und  $B_\varepsilon$  als  $\varepsilon \rightarrow 0$ :

$$\begin{aligned} m(A + B) &\geq m(A_\varepsilon + B_\varepsilon) \geq \left[ m(A_\varepsilon)^{\frac{1}{d}} + m(B_\varepsilon)^{\frac{1}{d}} \right]^d \geq \left[ (m(A) - \varepsilon)^{\frac{1}{d}} + (m(B) - \varepsilon)^{\frac{1}{d}} \right]^d \\ &= \left[ m(A)^{\frac{1}{d}} + m(B)^{\frac{1}{d}} \right]^d, \text{ als } \varepsilon \rightarrow 0. \end{aligned}$$

Ausgehend davon können wir zum Fall hinführen, wo  $A$  und  $B$  beliebig kompakte Mengen sind. Wir stellen fest, dass  $A + B$  dann auch kompakt ist. Ferner ist durch die Definition  $A^\varepsilon = \{x \in \mathbb{R}^d \mid d(x, A) < \varepsilon\}$  eine offene Menge mit  $A^\varepsilon \setminus A$ , als  $\varepsilon \rightarrow 0$ .

Damit ist aus dem Lemma 2.5.6(b)  $m(A^\varepsilon) \rightarrow m(A)$ , als  $\varepsilon \rightarrow 0$ . Analog dazu werden

$B^\varepsilon$  und  $(A + B)^\varepsilon$  definiert. Wir beachten dazu, dass  $A + B \subset A^\varepsilon + B^\varepsilon \subset (A + B)^{2\varepsilon}$ . Angenommen  $m(A + B) < \infty$  existiert für jedes  $\rho > 0$  ein  $\varepsilon > 0$ , sodass

$$\rho + m(A + B) \geq m((A + B)^{2\varepsilon}) \geq m(A^\varepsilon + B^\varepsilon) \geq \left[ m(A^\varepsilon)^{\frac{1}{d}} + m(B^\varepsilon)^{\frac{1}{d}} \right]^d \geq \left[ m(A)^{\frac{1}{d}} + m(B)^{\frac{1}{d}} \right]^d,$$

und somit wird wiederum die Ungleichung (1) erhalten, als  $\rho \rightarrow 0$ .

Seien nun  $A$ ,  $B$ , und  $A + B$  mit  $m(A + B) < \infty$  (also sind folglich  $m(A)$ ,  $m(B) < \infty$ ) messbar, existiert durch passende Approximation nach innen  $A$  und  $B$ , nämlich  $K_1 \subset A$ ,  $K_2 \subset B$  gemäß zum Teil (iii) des Satzes 2.5.8 also mit  $m(A \setminus K_1)$ ,  $m(B \setminus K_2) < \varepsilon$ . Dann ist  $K_1 + K_2 \subset A + B$ , und

$$m(A + B) \geq m(K_1 + K_2) \geq \left[ m(K_1)^{\frac{1}{d}} + m(K_2)^{\frac{1}{d}} \right]^d \geq \left[ (m(A) - \varepsilon)^{\frac{1}{d}} + (m(B) - \varepsilon)^{\frac{1}{d}} \right]^d$$

$$= \left[ m(A)^{\frac{1}{d}} + m(B)^{\frac{1}{d}} \right]^d, \text{ als } \varepsilon \rightarrow 0. \blacksquare$$

**Bemerkung 2.8.7**

Für eine konvexe Menge  $A$  (2.1.15) wird mit einer  $\lambda > 0$  unter Berufung der Definition (oben) der Menge  $\lambda A = \{\lambda x \mid x \in A\}$  beschrieben. Aus der Maßgröße  $A + \lambda A = (1 + \lambda)A$  nämlich  $m(A + \lambda A) = (1 + \lambda)^d m(A)$  wird Minkowski Ungleichung unter  $(1 + \lambda)^d \geq 1 + \lambda^d$  und  $\lambda > 0$  wiederum bestätigt.

**Satz 2.8.8**

Es sei  $E \in \mathfrak{R}^d$ . Es gilt für die Slices (2.7.7)  $E_a(E_{x=a}), E_b$ , und  $E_c$  mit  $a < b < c$  dann

$$m(E_b) = \min \{m(E_a), m(E_c)\}.$$

*Beweis:* Für den Fall  $m(E_a) = 0$  oder  $m(E_c) = 0$  ist die Aussage trivial. Ansonsten sei es  $t = \frac{b-a}{c-a} \rightarrow t \in (0, 1)$ . Mit der Monotonie und Subadditivität Eigenschaften des Maßes  $E_b$  (2.4), der Konvexität der  $E$ , mit der oben definiert  $b = a + t(c - a)$  und schließlich mit der Minkowski Ungleichung ergibt sich

$$\begin{aligned} (m(E_b))^{\frac{1}{n}} &\geq m((1-t)E_a + tE_c)^{\frac{1}{n}} \geq m((1-t)E_a)^{\frac{1}{n}} + m(tE_c)^{\frac{1}{n}} \\ &= [(1-t)^n m(E_a)]^{\frac{1}{n}} + [t^n m(E_c)]^{\frac{1}{n}} \\ &= (1-t)(m(E_a))^{\frac{1}{n}} + t(m(E_c))^{\frac{1}{n}} \\ &\geq (1-t)\left(\min \{m(E_a), m(E_c)\}\right)^{\frac{1}{n}} + t\left(\min \{m(E_a), m(E_c)\}\right)^{\frac{1}{n}} \\ &= \left(\min \{m(E_a), m(E_c)\}\right)^{\frac{1}{n}}. \blacksquare \end{aligned}$$

**Korollar 2.8.9**

Seien  $A, B \subset \mathfrak{R}^d$  und ihr Summe  $A + B$  messbar, es gilt dann

$$m\left(\frac{(A+B)}{2}\right) \geq \sqrt{m(A)m(B)}.$$

*Beweis:* Es folgt direkt aus dem Satz 2.8.6 und der arithmetisch-geometrischen Ungleichung.

$$\begin{aligned} m\left(\frac{(A+B)}{2}\right)^{\frac{1}{n}} &\geq m\left(\frac{A}{2} + \frac{B}{2}\right)^{\frac{1}{n}} = m\left(\frac{(A)}{2}\right)^{\frac{1}{n}} + m\left(\frac{(B)}{2}\right)^{\frac{1}{n}} = \frac{(m(A))^{\frac{1}{n}} + (m(B))^{\frac{1}{n}}}{2} \\ &\geq \sqrt{(m(A))^{\frac{1}{n}} (m(B))^{\frac{1}{n}}}. \blacksquare \end{aligned}$$

## 2.9 Rektifizierbare Kurve

**Definition 2.9.1**

Wir betrachten **stetige Kurve**  $\Gamma = \{z(t) = (x(t), y(t)) \mid t \in [a, b]\}$  mit stetigen reellwertigen Funktionen  $x$  und  $y$  von  $t$  (in der  $XOY$  Ebene). Die Länge der  $\Gamma$  wird wie oft als das Supremum der Längen von allen polygonalen Geraden approximiert (Abbildung 11), die nacheinander durch endlich viele Punkten der  $\Gamma$  verbunden sind

$$L(\Gamma) = \sup_{\{t_j\}_{j=1}^N} \sum_{j=1}^N |z(t_j) - z(t_{j-1})|.$$

Falls die Summe aller Längen beschränkt ist, nämlich gibt es eine  $M < \infty$  für je Partition, sodass

$$\sum_{j=1}^N |z(t_j) - z(t_{j-1})| \leq M,$$

ist die Kurve  $\Gamma$  **rektifizierbar** [RAR0051].

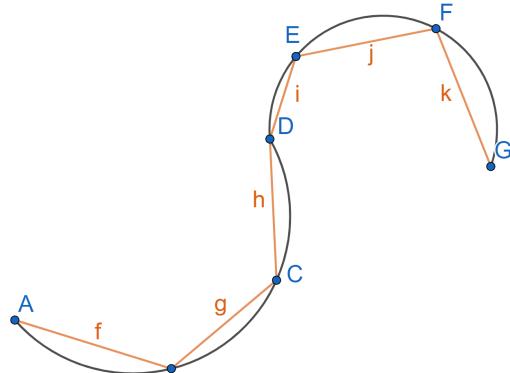


Abbildung 11: Approximation einer rektifizierbaren Kurve durch die Geraden  $f, g, h, i, j$  und  $k$ .

### Bemerkung 2.9.2

Wenn die Funktion  $z(t)$  bzw.  $x(t)$  und  $y(t)$  aber auch stetig differenzierbar sind, wird die Länge der Kurve wie folgt definiert

$$L = \int_a^b [(x'(t))^2 + (y'(t))^2]^{\frac{1}{2}} dt^1.$$

### Definition 2.9.3

Es sei  $F(t)$  eine komplexwertige Funktion definiert auf  $[a, b]$  mit  $a = t_0 < t_1 < \dots < t_N = b$  als eine Partition. Die **Variation** auf diese Partition ist definiert durch

$$\sum_{j=1}^N |F(t_j) - F(t_{j-1})|.$$

Ferner heißt die Funktion  $F$  **beschränkte Variation**, falls die Variationen der Funktion über alle Partitionen beschränkt ist. Mit anderen Worten existiert es eine  $M < \infty$  für je Partition, sodass

$$\sum_{j=1}^N |F(t_j) - F(t_{j-1})| \leq M.$$

### Bemerkung 2.9.4

Die Voraussetzung der Stetigkeit der Funktion  $F(t)$  kommt erst dann vor, wenn sie als eine Kurve nämlich  $F(t) = z(t) = x(t) + iy(t)$  bezeichnet wird. Außerdem gegeben sei es eine Partition  $\tilde{\wp}$  durch  $a = \tilde{t}_0 < \tilde{t}_1 < \dots < \tilde{t}_N = b$  und als eine Verfeinerung einer anderen Partition  $\wp$  durch

$a = t_0 < t_1 < \dots < t_N = b$ . Damit ist die Variation von  $F$  auf  $\tilde{\wp}$  größer oder gleich die Variation von  $F$  auf  $\wp$ .

### Satz 2.9.5

Eine durch  $(x(t), y(t))$  parametrisierte Kurve ist für  $t \in [a, b]$  nur dann rektifizierbar, wenn beide

- 1 Die Bedingung stetige Differenzierbarkeit für  $x(t)$  und  $y(t)$  kann durch absolute Stetigkeit gelockert werden, da  $z(t)$  damit für a.e.  $t \in [a, b]$  differenzierbar ist.
- 2 Eine Partition  $\tilde{\wp}$  von  $[a, b]$  ist eine Verfeinerung einer Partition  $\wp$  von  $[a, b]$ , falls jeder Punkt in der  $\wp$  auch zu  $\tilde{\wp}$  gehört.

Funktionen  $x(t)$  und  $y(t)$  beschränkte Variation sind.

**Beweis:** Es folgt direkt von der Definition der  $F(t) = x(t) + iy(t)$  bzw. die Teleskopsumme Regel für  $l_j = F(t_j) - F(t_{j-1}) = (x(t_j) - x(t_{j-1})) + i(y(t_j) - y(t_{j-1}))$ , falls  $a$  und  $b$  reell sind.  
 $|a + ib| \leq |a| + |b| \leq 2|a + ib|$ . ■

### Definition 2.9.6

- Die **totale Variation** einer Funktion  $F$  auf  $[a, x]$ , wobei  $a \leq x \leq b$ , ist definiert durch

$$T_F(a, x) = \sup_{\{t_j\}} \sum_{j=1}^N |F(t_j) - F(t_{j-1})|,$$

wobei es sich um das Supremum über alle Partitionen von  $[a, x]$  handelt.

- Die **positive Variation** einer Funktion  $F$  auf  $[a, x]$  ist definiert durch

$$P_F(a, x) = \sup_{\{t_j\}^+} \sum_{j=1}^N F(t_j) - F(t_{j-1}),$$

wobei es sich um das Supremum über alle Partitionen von  $[a, x]$  mit  $F(t_j) \geq F(t_{j-1})$  handelt.

- Die **negative Variation** einer Funktion  $F$  auf  $[a, x]$  ist definiert durch

$$N_F(a, x) = \sup_{\{t_j\}^-} \sum_{j=1}^N [F(t_j) - F(t_{j-1})],$$

wobei es sich um das Supremum über alle Partitionen von  $[a, x]$  mit  $F(t_j) \leq F(t_{j-1})$  handelt.

### Lemma 2.9.7

Es sei  $F$  eine reellwertige Funktion und einer von beschränkter Variation auf  $[a, b]$ , dann gelten für alle  $x \in [a, b]$

- $F(x) - F(a) = P_F(a, x) - N_F(a, x)$
- $T_F(a, x) = P_F(a, x) - N_F(a, x)$

### Definition 2.9.8

Eine durch  $(x(t), y(t))$  parametrisierte Kurve  $z(t)$  heißt für  $t \in [a, b]$  **einfach**, falls die Abbildung  $t \mapsto z(t)$  injektiv<sup>1</sup> für  $t \in [a, b]$ . Falls es für  $t \in [a, b]$  mit  $z(a) = z(b)$  gilt, heißt  $z(t)$  eine **abgeschlossene einfache** Kurve. Generell heißt eine Kurve **quasi-einfach** (Abbildung 12), falls die Abbildung  $z(t)$  für  $t$  in der Komplement<sup>2</sup> von endlich vielen Punkten des Intervalls  $[a, b]$  injektiv ist.

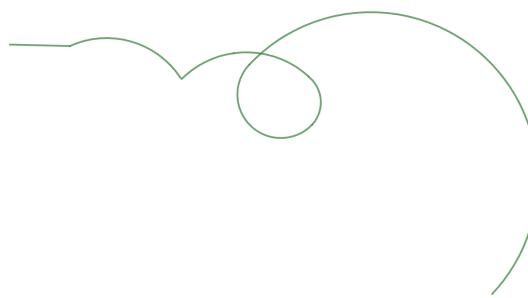


Abbildung 12: Eine quasie-einfache Kurve.

### Definition 2.9.9

Es sei  $K \in \mathbb{R}^2$  eine kompakte Menge (wie z.B  $\Gamma$  oben). Wir zeichnen eine offene Menge  $K^\delta$  als

1 Eine Abbildung  $f: X \rightarrow Y$  heißt injektive, falls es zu jedem Element  $y$  der Zielmenge höchstens ein Element  $x$  der Ausgangsmenge gibt. Damit ergibt sich  $x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)$

eine Menge, die alle Punkten der Menge  $K$  mit einem Abstand kleiner als  $\delta$  enthält (Abbildung 13)

$$K^\delta = \{x \in \mathbb{R}^2 \mid d(x, K) < \delta\}.$$

Wir zeichnen **Minkowski Inhalte**<sup>†</sup> der Menge  $K$  mit

$$\mu(K) = \lim_{\delta \rightarrow 0} \frac{m(K^\delta)}{2\delta},$$

falls der Grenzwert existiert.

### Bemerkung 2.9.10

Für eine kompakte Menge  $K$  ist eine Äquivalenz zur Minkowski Inhalte mit Hilfe der Infimum und Supremum der Grenzwert darstellbar und zwar für die beiden Werten

$$\mu^*(K) = \limsup_{\delta \rightarrow 0} \frac{m(K^\delta)}{2\delta} \text{ und } \mu_*(K) = \liminf_{\delta \rightarrow 0} \frac{m(K^\delta)}{2\delta}$$

muss  $\mu_*(K) = \mu^*(K) < \infty$  gelten.

### Satz 2.9.11

Es sei  $\Gamma = \{z(t) = (x(t), y(t)) \mid t \in [a, b]\}$  eine quasi-einfache Kurve. Die Minkowski Inhalte der  $\Gamma$  existiert nur dann, wenn die  $\Gamma$  rektifizierbar ist. Wenn es der Fall ist, ist dann  $\mu(\Gamma) = L$  (2.9.1).

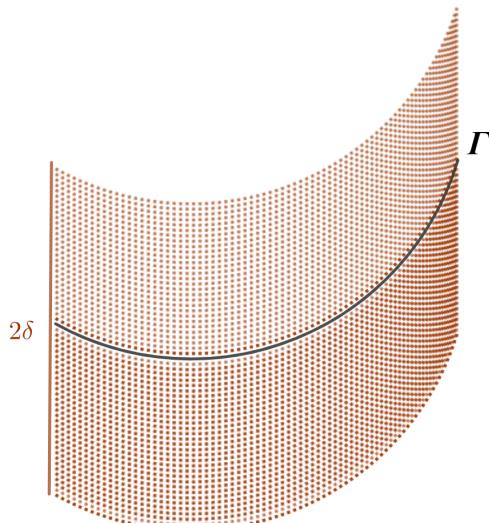


Abbildung 13: Die Kurve  $\Gamma$  und die Menge  $\Gamma^\delta$ .

### Proposition 2.9.12

Es sei  $\Gamma = \{z(t) = (x(t), y(t)) \mid t \in [a, b]\}$  eine quasi-einfache Kurve. Wenn  $\mu_*(\Gamma) < \infty$ , ist dann die Kurve rektifizierbar und ihre Länge  $L \leq \mu_*(\Gamma)$ .

### Lemma 2.9.13<sup>‡</sup>

Es sei  $\Gamma = \{z(t) = (x(t), y(t)) \mid t \in [a, b]\}$  irgendeine Kurve, und  $\Delta = |z(b) - z(a)|$  der Abstand zwischen ihre Endpunkten. Es gilt dann

$$m(\Gamma^\delta) \geq 2\delta\Delta.$$

### Proposition 2.9.14

Es sei  $\Gamma = \{z(t) = (x(t), y(t)) \mid t \in [a, b]\}$  eine rektifizierbare Kurve mit der Länge  $L$ . Es gilt dann

$$\mu^*(\Gamma) \leq L.$$

<sup>†</sup> Diese ist eindimensionale Minkowski Inhalte.

<sup>‡</sup> Zum besseren Verstehen des Hintergrunds dieser Aussage eine Vorkenntnisse über die Segmentierung unter totale Variation hilfreich, worauf wegen des Kernthema unseres Schreibens verzichtet ist.

**Bemerkungen 2.9.15**

- i. Es gibt auch die sogenannte Gammafunktion, die für  $s > 0$  wie folgt definiert ist.

$$\Gamma(s) = \int_0^\infty e^{-x} x^{s-1} dx.$$

Das Integral kann man mit der sogenannte Stirling's Formel wie folgt approximieren.

$$\Gamma(s) = e^{s \log(s)} e^{-s} \frac{\sqrt{2\pi}}{\sqrt{s}} \left( 1 + \sum_{j=1}^N a_j s^{-j} + O(s^{-N}) \right).$$

Dabei sind  $a_i$  (z.B.  $a_1 = \frac{1}{2}$ ) reelle Konstanten für jede  $N > 0$ , wenn  $s \in S_\delta = \{s : |\arg s| \leq \pi - \delta\}$ .

- ii. In Bezug auf die Länge einer Kurve (in der Bemerkung 2.9.2) kann eine Parametrisierung um Variable  $t \in [a, b]$  und damit um die Menge  $s = s(t) = L(a, t)$  durch das Paar  $\tilde{z}(s) = (\tilde{x}(s) + i \tilde{y}(s))$  eingeführt werden, indem  $\tilde{z}(s) = z(t)$  für  $s = s(t)$ . Dabei ist es klar, dass  $|\tilde{z}(s_1) - \tilde{z}(s_2)| \leq |s_1 - s_2|$  (Abstand zwischen zwei Punkten  $\leq$  Länge der Portion) der Kurve. Diese Parametrisierung wird als **Bogenlänge Parametrisierung** bezeichnet.

**Satz 2.9.16**

Es sei  $z(t) = (x(t), y(t))$ ,  $\forall t \in [a, b]$ , eine rektifizierbare Kurve mit Länge  $L$ . Betrachte die oben beschriebene Bogenlänge Parametrisierung  $\tilde{z}(s) = (\tilde{x}(s), \tilde{y}(s))$ , sind dann  $\tilde{x}(t)$  und  $\tilde{y}(t)$  absolut stetig,  $|\tilde{z}'(s)| = 1$  a.e  $s \in [0, L]$ , und

$$L = \int_0^L [(\tilde{x}'(s))^2 + (\tilde{y}'(s))^2]^{1/2} ds.$$

*Beweis:* Wir stellen fest, dass  $|\tilde{z}(s_1) - \tilde{z}(s_2)| \leq |s_1 - s_2|$ . Damit ergibt sich unmittelbar die absolute Stetigkeit der  $\tilde{z}(s)$  und daher ist die Funktion für a.e.  $s \in [0, L]$  differenzierbar. Außerdem führt durch Division dieser Ungleichung auf  $|s_1 - s_2|$  zur  $|\tilde{z}'(s)| \leq 1$  a.e  $s \in [0, L]$ . Nun durch Definition der **totale Variation** von  $\tilde{z}(s)$  gleich mit  $L$ , und die Tatsache

$$L = T_{\tilde{z}}(a, x) = \int_0^L |\tilde{z}'(s)| ds = \int_0^L [(\tilde{x}'(s))^2 + (\tilde{y}'(s))^2]^{1/2} ds.$$

Die zweit Gleichung ist nur dann möglich, wenn fast überall in  $[0, L]$   $|\tilde{z}'(s)| = 1$ . ■

**2.10 Die Isoperimetrische Ungleichung**

Mit Hilfe der Kenntnisse vom letzten Kapitel fangen wir an, die Bedeutung der Ungleichung zuerst in einem 2 dimensionalen Gebiet zu verstehen und zu beweisen.

**Bemerkung 2.10.1**

Die Isoperimetrische Ungleichung sagt aus, dass zwischen alle Kurve einer bestimmten Länge der Kreis das Maximum der Fälsche enthält [RAR0051].

Zur Vorbereitung der Beweis der Ungleichung nehmen wir an, dass  $\Omega \subset \mathbb{R}^2$  ein beschränkte offene Menge, derer Rand  $\overline{\Omega} \setminus \Omega$  eine rektifizierbare Kurve mit einer Länge  $\ell(\Gamma)$  ist. Ob  $\Gamma$  eine abgeschlossene einfache Kurve ist, spielt dabei keine Rolle.

**Satz 2.10.2 ( Isoperimetrische Ungleichung in  $\mathbb{R}^2$  )**

Die Isoperimetrische  $4\pi m(\Omega) < (\ell(\Gamma))^2$ .

*Beweis:* Für jede  $\delta > 0$  betrachten wir äußere Menge

- 
- Für jede nicht-negative komplexe Zahl  $z = r e^{i\theta} = r(\cos \theta + i \sin \theta)$  heißt  $\theta \in \mathbb{R}$  Argument von  $z$  und es wird oft mit  $\arg z$  bezeichnet.

$$\Omega_+(\delta) = \{x \in \mathbb{R}^2 \mid d(x, \overline{\Omega}) < \delta\}$$

und innere Menge

$$\Omega_-(\delta) = \{x \in \mathbb{R}^2 \mid d(x, \Omega^C) \geq \delta\}.$$

Damit ergibt sich  $\Omega_-(\delta) \subset \Omega \subset \Omega_+(\delta)$  (Abbildung 14). Wir beachten nun, dass für

$$\Gamma^\delta = \{x \in \mathbb{R}^2 \mid d(x, \Gamma) < \delta\}$$

$$(1) \quad \Omega_+(\delta) = \Omega_-(\delta) \cup \Gamma^\delta$$

als eine disjunkte Vereinigung. Ferner sei  $D(\delta) = \{x \in \mathbb{R}^2 \mid |x| < \delta\}$  des offenen Balls (Disk mit dem Ursprung als Mittelpunkt und dem Radius  $\delta$ ). Daraus folgt

$$\begin{cases} \Omega_+(\delta) \supset \Omega + D(\delta), \\ \Omega \supset \Omega_-(\delta) + D(\delta) \end{cases}$$

Mit der Brunn-Minkowski Ungleichung ergibt sich für den ersten Schluss

$$m(\Omega_+(\delta)) \geq \left[ (m(\Omega))^{\frac{1}{2}} + m(D(\delta))^{\frac{1}{2}} \right]^2.$$

Da  $m(D(\delta)) = \pi \delta^2$  und  $(A + B)^2 \geq A^2 + 2AB$  (immer wenn  $AB \geq 0$ ), finden wir, dass

$$(2) \quad m(\Omega_+(\delta)) \geq m(\Omega) + 2\pi^{\frac{1}{2}} \delta m(\Omega)^{\frac{1}{2}}.$$

Analog zum ersten Schluss ergibt sich mit dem Zweiten

$$m(\Omega) \geq m(\Omega_-(\delta)) + 2\pi^{\frac{1}{2}} \delta m(\Omega_-)^{\frac{1}{2}}$$

oder äquivalent:

$$(3) \quad -m(\Omega_-(\delta)) \geq 2\pi^{\frac{1}{2}} \delta m(\Omega_-)^{\frac{1}{2}} - m(\Omega).$$

Nun folgt aus der (1)

$$m(\Omega_+(\delta)) - m(\Omega_-(\delta)) = m(\Gamma^\delta),$$

und mit der oberen Ungleichungen (2) und (3) kriegt man

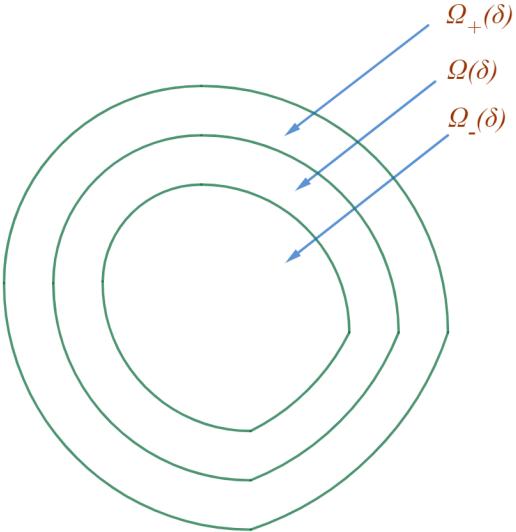
$$m(\Gamma^\delta) \geq 2\pi^{\frac{1}{2}} \delta \left( m(\Omega)^{\frac{1}{2}} + m(\Omega_-)^{\frac{1}{2}} \right).$$

Nun erhält man durch die Division  $2\delta$  auf der beiden Seiten der letzte Ungleichung und gleichzeitig das Bemerkens dessen, dass  $\Omega_-(\delta) \nearrow \Omega$ , als  $\delta \rightarrow 0$

$$\mu^*(\Gamma^\delta) \geq \pi^{\frac{1}{2}} \left( 2m(\Omega)^{\frac{1}{2}} \right).$$

Mit der Proposition 2.9.14 ( $\ell(\Gamma) \geq \mu^*(\Gamma)$ ) ergibt sich

$$\ell(\Gamma) \geq 2(\pi m(\Omega))^{\frac{1}{2}}. \blacksquare$$

Abbildung 14: Die Mengen  $\Omega_-$ ,  $\Omega$ , und  $\Omega_+ \in \mathbb{R}^2$ **Bemerkungen 2.10.3**

- i. Das gleiche Resultat wird sogar ohne Annahme für den Rand, also als eine rektifizierbare Kurve, gelten. In der Tat zeigt der Beweis, dass für jede beschränkte offene Menge mit einem Rand  $\Gamma$  gilt

$$4\pi m(\Omega) < (\mu^*(\Gamma^\delta))^2.$$

- ii. Generell ist das Maßgröße eines beliebigen Balls  $B \in \mathfrak{R}^d$  durch **Invarianzeigenschaften** (Translation und Dilatation) einer messbaren Menge darstellbar. Also  $m(B) = v_d r^d$ , wobei die Konstante  $v_d = m(B_1)$  und  $B_1 = \{x \in \mathfrak{R}^d \mid |x| < 1\}$ . Außerdem kann mit dem **Korollar 2.7.15** gezeigt werden, dass

$$v_d = \int_{-1}^1 (1 - x^2)^{\frac{1}{2}} dx = \pi.$$

Beim gleichen Verfahren findet man die Konstante für das Ball B wie folgt:

$$v_d = 2 \int_0^1 (1 - x^2)^{\frac{(d-1)}{2}} dx = \frac{\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2} + 1)},$$

Wobei  $\Gamma(z)$  die **Gammafunktion** <sup>\*</sup> ist.

**Korollar 2.10.5**

Es sei eine Kurve in  $\mathfrak{R}^d$  als eine stetige Abbildung  $t \mapsto z(t)$  von einem Intervall  $[a, b] \rightarrow \mathfrak{R}^d$ . Analog zur Bedingungen für Rektifizierbarkeit von Kurven, gegebene im **Satz 2.9.5** und **Satz 2.9.16**, gilt sie auch für diese Kurve. Darüber hinaus gilt gemäß der **Definition 2.9.9** die Minkowski Inhalte  $\mu(K)$

von einer Kompaktmenge  $K \in \mathfrak{R}^d$  im Falle Existenz des Grenzwerts mit  $\mu(K) = \lim_{\delta \rightarrow 0} \frac{m(K^\delta)}{m_{d-1}(B(\delta))}$

wobei  $m_{d-1}(B(\delta))$  die Maßgröße (in der  $\mathfrak{R}^{d-1}$ ) vom Ball mit  $B_\delta = \{x \in \mathfrak{R}^{d-1} \mid |x| < \delta\}$ .

Abschließend gelten entsprechende Aussagen gemäß der **Propositionen 2.9.12** und **2.9.14** für die oben beschriebene Kurve.

\* Äquivalent für gerade Dimension kann  $\Gamma(\frac{d}{2} + 1)$  durch die Definition der Funktion als  $\prod_{j=0}^{(d/2)-1} (\frac{d}{2} - j)$  und für ungerade Dimension werden, mit **Wahrnehmung reeller Zahl**  $\frac{d}{2}$ ,  $\Gamma(\frac{d}{2} + 1) = \frac{\sqrt{\pi} (d!!)}{2^{\frac{(d+1)}{2}}}$  mit  $d!! = \prod_{j=0}^{[d/2]} (d - 2j)$ .

### 3 Johnson-Lindenstrauss Lemma

---

Wir fangen zuerst mit notwendige Grundlage der Wahrscheinlichkeitstheorie an, bevor wir zur Beschreibung des Hauptlemmas bzw. die Anwendung des Lemmas in der Stochastik kommen.

#### 3.1 Grundlage der Wahrscheinlichkeitstheorie

##### Definitionen 3.1.1

- a) Der **Ergebnisraum**  $\Omega$  (oft wird auch die **Omega Menge** gezeichnet) ist die Menge aller möglichen Ergebnisse, welche als zufällige Ausgang eines Experiments oder Versuches (Zufallsexperiment) bezeichnet wird.
- b) Eine Menge  $\mathcal{F}$  (oft wird auch mit  $\Sigma$  bezeichnet) als Kollektion von Teilmengen der Menge  $\Omega$  ist (erfüllt die Bedingungen für)  **$\sigma$ -Algebra**<sup>1</sup>, falls
  - (i)  $\mathcal{F}$  nichtleere Menge ist.
  - (ii) Für jede  $X \in \mathcal{F}$  die Menge  $\bar{X} = (\Omega \setminus X) \in \mathcal{F}$ .
  - (iii) Für jede  $X, Y \in \mathcal{F}$  ist die Vereinigung  $(X \cup Y) \in \mathcal{F}$ .
- c) Eine **Wahrscheinlichkeitsmaß** ist eine Abbildung  $\mathbb{P}: \mathcal{F} \rightarrow [0, 1]$  (Zuweisen/ Zuordnung der Wahrscheinlichkeit) zu Ereignissen und hat folgenden Eigenschaften [LVSM006]<sup>2</sup>:
  - (i)  $\mathbb{P}[\Omega] = 1$ .
  - (ii) (Additivität) Für zwei disjunkte Mengen  $X, Y \in \mathcal{F}$  gilt es  $\mathbb{P}[X \cup Y] = \mathbb{P}[X] + \mathbb{P}[Y]$ .
- d) Ein **Wahrscheinlichkeitsraum** ist ein Tripel  $(\Omega, \mathcal{F}, P)$ , wobei  $\Omega$  ein Stichprobenraum,  $\mathcal{F}$  eine  $\sigma$ -Algebra, definiert über  $\Omega$ , und  $P$  ein Wahrscheinlichkeitsmaß sind.
- e) Eine **Zufallsvariable**  $f$  ist eine Abbildung von  $\Omega$  in einer Menge  $G$ . Wir stellen fest, dass die Wahrscheinlichkeit, dass die Zufallsvariable einen Wert in einer gegebenen Teilmenge von Werten annimmt, wohldefiniert ist. Formal ausgedrückt bedeutet dies, dass für jede Teilmenge  $U \subseteq G$  gilt:  $f^{-1}(U) \in \mathcal{F}$ . Das heißt,  $P[f \in U] = P[f^{-1}(U)]$  ist definiert. Z.B sei eine Zufallsvariable definiert durch  $X: \Omega \rightarrow \mathbb{R}$  auf einem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, P)$ . Wir können Fragen stellen, wie wahrscheinlich ist es, dass der Wert von  $X$  gleich zwei ist? Dies entspricht der Wahrscheinlichkeit des Ereignisses  $\{\omega \mid X(\omega) = 2\}$  welche oft als  $P(X = 2)$  bezeichnet.
- f) Wir sagen, dass  $X$  eine **stetige** Zufallsvariable ist, wenn es eine nichtnegative Funktion  $f(x)$  gibt, die für alle reellen  $x \in (-\infty, \infty)$  definiert ist und die Eigenschaft hat, dass für jede Menge  $B$  von reellen Zahlen bezüglich des Lebesgue-Maßes  $\lambda$  folgendes gilt [SMR019]:
 
$$P\{X \in B\} = \mathbb{P}[B] = \int_B f(x) d\lambda(x).$$

Dabei werden Funktion  $f(x)$  als **Wahrscheinlichkeitsdichtefunktion** der Zufallsvariable  $X$  und  $\mathbb{P}[B]$  als ihre **Verteilungsfunktion** (Abbildung 15) bezeichnet.

- g) Falls eine reelle Zufallsvariable  $X$  eine Wahrscheinlichkeitsdichtefunktion  $f(x)$  besitzt und der **Erwartungswert** existiert, kann er wie folgt berechnet werden:

$$\mathbb{E}[X] = \int_B x f(x) d\lambda(x).$$

Im Fall einer **diskreten** Zufallsvariable wird er  $\mathbb{E}[X] = \sum_{x: p(x) > 0} x p(x)$  berechnet.

- h) Sei  $(\Omega, \Sigma, P)^\sigma$  ein Wahrscheinlichkeitsraum und  $X$  eine Zufallsvariable auf diesem Raum. Die **Varianz** ist definiert im Folgenden durch  $\mathbb{E}[X^2] - \mathbb{E}[X]^2$ :

1 Auch Borelmenge, erwähnt in der letzten Kapitel (Seite 26), ist ein  $\sigma$ -Algebra.

2 Die Wahrscheinlichkeitsmaß ist eine spezieller Fall vom **Lebesgue-Maß** und daher hat er diese Eigenschaften.

$$\text{Var}(X) := \mathbb{E}[(X - \mu)^2].$$

- i) Für einen Vektorspalte  $X = (X_1, X_2, \dots, X_n)^T$ , deren Einträgen Zufallsvariablen sind, ist eine **Kovarianzmatrix** wie folgt definiert [SMR019]:

$$\mathfrak{R}^{n \times n} \ni K_{XX} = \begin{bmatrix} K_{X_i X_j} \end{bmatrix}_{1 \leq i, j \leq n} \text{ mit } K_{X_i X_j} = \text{cov}[X_i, X_j] = \mathbb{E}[(X_i - \mathbb{E}[X_i])(X_j - \mathbb{E}[X_j])].$$

- j) Eine stetige Zufallsvariable  $X$  hat eine **Normalverteilung** mit Erwartungswert  $\mathbb{E}[X] = \mu$  und Varianz  $\text{Var}(X) = \sigma^2$ , oft geschrieben als  $X \sim \mathcal{N}(\mu, \sigma^2)$ , falls  $X$  folgende Wahrscheinlichkeitsdichtefunktion hat [ARGB007]:

$$\forall x \in (-\infty, \infty), f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right).$$

Analog zur eindimensionalen Normalverteilung wird die **mehrdimensionale** Normalverteilung definiert wie folgt (für reelle Zufallsvariable  $X \sim \mathcal{N}_d(\mu, \Sigma)$ ):

$$f_X(x) = \frac{1}{\sqrt{(2\pi)^d \det(\Sigma)}} \exp\left(-\frac{1}{2}(x - \mu)^T \Sigma^{-1} (x - \mu)\right).$$

Dabei ist  $\Sigma$  eine reguläre **Kovarianzmatrix**, welche symmetrisch positive definiert<sup>s</sup> ist.

Die Symmetrie bedeutet hier für jede zwei Punkten  $x, y \in \mathfrak{R}^d$  nicht nur  $|x| = |y|$  sondern auch die Wahrscheinlichkeitsdichtefunktion an ihre Stelle gleich sind (d.h.  $f(x) = f(y)$ ).

Die **Projektion** der Normalverteilung in jede Richtung ist eine eindimensionale Normalverteilung. Hält man  $d$  Variablen  $X_1, X_2, \dots, X_d$  durch eindimensionale Normalverteilung, ergibt sich damit ein Punkt  $(X_1, X_2, \dots, X_d)^T$ , der eine mehrdimensionale Normalverteilung hat.

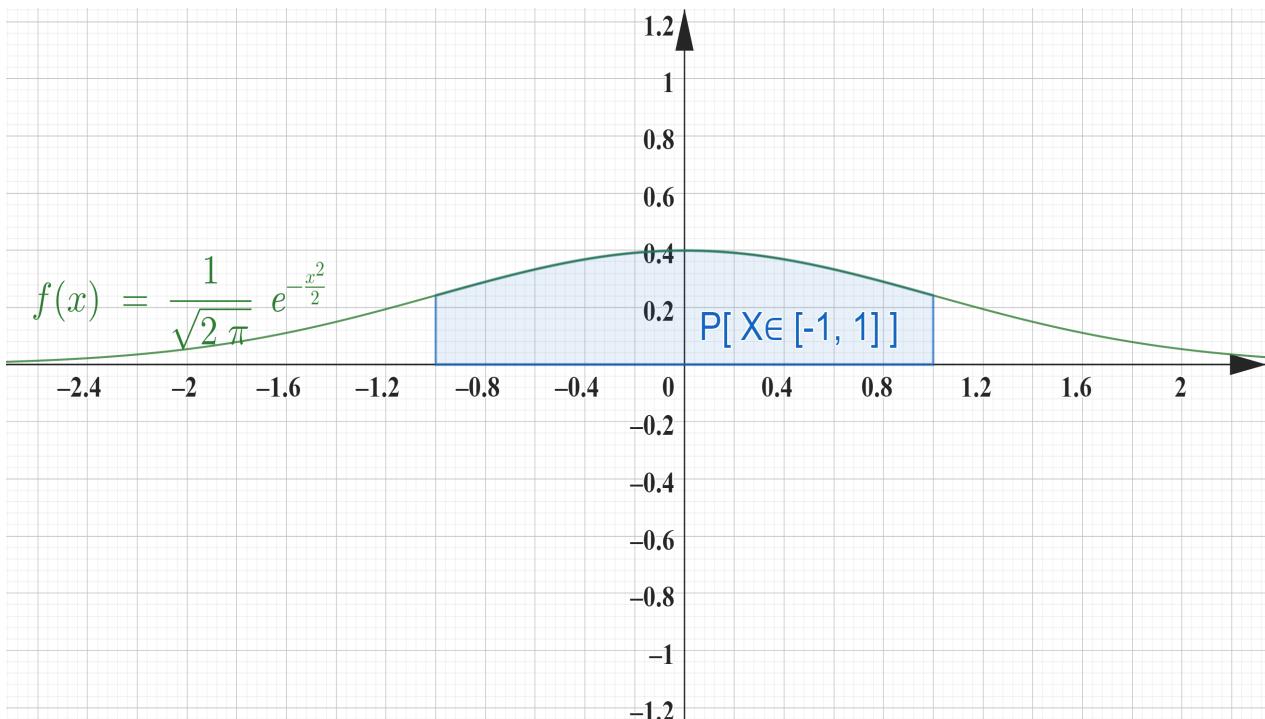


Abbildung 15: Die Standardnormalverteilungsfunktion für eine 2-dimensionale Zufallsvariable.

$\sigma$  Das  $\sum$  ist ein  $\sigma$ -Algebra über der Grundmenge  $\Omega$ .

$\S$  Eine Matrix  $M \in \mathfrak{R}^{n \times n}$  ist symmetrisch positive definite  $\Leftrightarrow \forall x \in \mathfrak{R}^n \setminus \{0\}, x^T M x > 0$ .

### 3.2 Affine Geometrie

Während Konvexe Mengen (2.1.15) die Strecke (eng. Line Segment) zwischen zwei beliebigen Punkten enthalten, enthalten Affine Mengen die gesamte Linie zwischen zwei beliebigen Punkten [OGUE010]. Daher ist die natürliche Einstellung für konvexe Mengen eine affine Teilmenge eines Vektorraums (Grundlagen der Maßtheorie 4).

#### Definition 3.2.1

1. Eine nichtleere Teilmenge  $A$  eines Vektorraums  $E$  wird als affine Menge bezeichnet, wenn für Punkte  $x$  und  $y$  in  $A$  die durch  $x$  und  $y$  verlaufende Linie in  $A$  enthalten werden.

$$x, y \in A \Rightarrow \ell := \{x + t(x - y) = (1 - t)x + y : t \in \mathbb{R}\} \subseteq A.$$

2. Sei  $\mathbf{A}$  und  $\mathbf{B}$  Affine Mengen in Vektorräumen  $\mathbf{E}$  und bzw.  $\mathbf{D}$ . Eine Abbildung  $F: \mathbf{A} \rightarrow \mathbf{B}$  wird als affine Abbildung bezeichnet, falls für alle  $t \in \mathbb{R}$   $F((1-t)x + t y) = (1-t)F(x) + t F(y)$  gilt.

3. Sei  $\{x_k\}_{k=1}^d$  eine endliche Menge von Punkten in dem Vektorraum  $E$ . Eine affine Kombination von  $\{x_k\}_{k=1}^d$  ist jeder Punkt  $y = \sum_{k=1}^d \lambda_k x_k$  mit  $\sum_{k=1}^d \lambda_k = 1$ .
4. Sei  $B$  eine nichtleere Menge im Vektorraum  $E$ . Die affine Hülle (engl. affine hull od. span) der  $B$  ist die Menge aller affinen Kombinationen von Punkten aus  $B$ , das heißt:

$$\left\{ \sum_{k=1}^d \lambda_k b_k \mid b_k \in B, \sum_{k=1}^d \lambda_k = 1, d = 1, 2, \dots \right\}.$$

5. Sei  $A \subseteq E$  eine affine Teilmenge vom Vektorraum  $E$  und  $a \in A$  ein beliebiger Punkt, ist dann  $L := A - a = \{y - a \mid y \in A\}$  ein linearer Subraum des Vektorraums  $E$ , welches unabhängig vom Punkt  $a \in A$ . Als Konsequenz gelten dann  $A = a + L$  und  $L := A - a = \{y - z \mid y, z \in A\}^\dagger$ .

#### Bemerkung 3.2.2

Ein affiner Raum ist auch durch eine Menge  $A$  zusammen mit einem Vektorraum  $V$  unter einer transitiven und freien Gruppenoperation der additiven Gruppe von  $V$  auf die Menge  $A$  wohldefiniert. Explizit bedeutet die obige Definition, dass die Gruppenoperation eine Abbildung ist, die generell als Addition mit den folgenden Eigenschaften bezeichnet wird [BM984]:

$$A \times V \rightarrow A$$

$$(a, v) \mapsto a + v,$$

1. Neutrales Element:

$$\forall a \in A, a + 0 = a, \text{ wobei } 0 \text{ das Nullvektor in } V \text{ ist.}$$

2. Assoziativität:

$$\forall v, w \in V, \forall a \in A, (a + v) + w = a + (v + w) \text{ (hier ist die letzte Addition, eine Addition in } V\text{)}$$

3. Freie und Transitive Operationen:

$$\forall a \in A, \text{ ist die Abbildung } V \rightarrow A: v \mapsto a + v \text{ eine Bijektion (Seite 4).}$$

4. Existenz von injektiven Translationen (Parallelverschiebungen):

$$\forall v \in V, \text{ ist die Abbildung } A \rightarrow A: a \mapsto a + v \text{ eine Bijektion.}$$

5. Subtraktion:

---

† Leser kann aus eigener Interesse einen Beweis für diese Aussage im Buch „Foundation of Optimization“ von Herrn Osman Güler unter Abschnitt 4.1 Affine Geometry finden.

$\forall a, b \in A$ , existiert ein eindeutige  $v \in V$ , bezeichnet als  $b - a$ , sodass  $b = a + v$ .

Damit ist  $\mathbb{R}^n$  auch ein affiner Raum (als ein Vektorraum über selbst) und sein affiner Subräume haben der Form  $E = p + V$ , wobei  $V \subset \mathbb{R}^n$  ein Vektorunterraum ist.

**Beispiel:**

$$\text{Sei } p = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \wedge V = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \mid \begin{pmatrix} x \\ y \end{pmatrix} = t \begin{pmatrix} 2 \\ 3 \end{pmatrix}, t \in \mathbb{R} \right\} \subset \mathbb{R}^2, \text{ ist dann der affiner Unterraum :}$$

$$E = p + V = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \mid \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1+2t \\ -1+3t \end{pmatrix}, t \in \mathbb{R} \right\} \subset \mathbb{R}^2.$$

### Lemma 3.2.3

Für affinen Räumen  $D$  und  $E$  ist eine affine Abbildung in der Form  $F: D \rightarrow E$  mit  $F(x) = Ax + b$  darstellbar, wobei das Matrix-Vektor-Produkt „ $Ax$ “ die Rolle einer linearen Transformation<sup>△</sup> und der Vektor „ $b$ “ die Rolle die Translation (oder Parallelverschiebung) spielen.

**Beweis:** Sei  $t \in \mathbb{R}$  und  $x, y \in D$ , dann  $F((1-t)x + ty) = M((1-t)x + ty) + b$

$$\Delta = (1-t)(Mx + b) + t(My + b) + (t - 1)b - tb + b = (1-t)F(x) + tF(y). \blacksquare$$

## 3.3 Konzentration des Maßes

Die Konzentration des Maßes ist ein anderes Aspekt von der Maßtheorie, welche überwiegend ein mathematisches Phänomen beschreibt, das häufig in der Stochastik vorkommt.

### Definitionen 3.3.1

(a) Ein **metrischer Raum** ist ein zueinander geordnetes Paar  $(M, d)$ , wobei  $M$  eine Menge und  $d$  eine **Metrik** auf  $M$  ist, d.h. eine Funktion  $d: M \times M \rightarrow \mathbb{R}$ , die folgende Voraussetzungen für alle Punkten  $x, y, z \in M$  erfüllt [RAR0052]:

- Der Abstand von einem Punkt zu sich selbst ist Null:  $d(x, x) = 0$ .
- (Positivität) Der Abstand zwischen zwei verschiedenen Punkten ist immer positiv: Wenn  $x \neq y$ , dann  $d(x, y) > 0$ .
- (Symmetrie) Der Abstand von  $x$  zu  $y$  ist immer gleich dem Abstand von  $y$  zu  $x$ :  $d(x, y) = d(y, x)$ .
- (Dreiecksungleichung) Die Dreiecksungleichung gilt:

$$d(x, z) \leq d(x, y) + d(y, z).$$

(b) Gegeben zwei metrische Räume  $(X, d_X)$  und  $(Y, d_Y)$ , wobei  $d_X$  die Metrik auf der Menge  $X$  und  $d_Y$  die Metrik auf der Menge  $Y$  bezeichnet, heißt eine Funktion  $f: X \rightarrow Y$  **Lipschitzstetig**, wenn es eine reelle Konstante  $K \geq 0$  gibt, so dass für alle  $x_1$  und  $x_2$  in  $X$  gilt:

$d_Y(f(x_1), f(x_2)) \leq K d_X(x_1, x_2)$ . Jede solche  $K$  wird als **Lipschitz-Konstante** für die Funktion  $f$  bezeichnet und  $f$  wird oft auch als  $K$ -Lipschitz genannt.

(c) Sei  $(X, d)$  ein metrischer Raum mit einem Maß  $\mu$  auf den Borelmengen (oben) mit  $\mu(X) = 1$ . Sei  $\alpha(\varepsilon) = \sup \{ m(X \setminus B_\varepsilon) \mid B \text{ ist eine Borelmenge und } \mu(B) \geq 1/2 \}$ , wobei die Funktion  $\alpha(\cdot)$  wird als **Konzentrationsrate** des Raumes  $X$  bezeichnet.

(d) In einem metrischen Raum  $M = (X, d_X)$  ist eine Menge  $V$  eine Umgebung eines Punktes  $p$ ,

△ Seien  $V$  und  $W$  Vektorräume über einem gemeinsamen Grundkörper (Menge mit der Wohldefiniertheit für die Operationen Addition und Multiplikation)  $K$ . Eine Abbildung  $f: V \rightarrow W$  heißt lineare Abbildung, falls für alle  $u, v \in V$  und  $\lambda \in K$  die folgende Bedingung gelten:  $f(\lambda u + v) = \lambda f(u) + f(v)$ .

falls es einen offenen Ball (2.1.1) mit Mittelpunkt  $x$  und Radius  $r > 0$  gibt, so dass es in der  $V$  enthalten ist (Abbildung 16<sup>‡</sup>).

(e) Ein **sphärisches** Wahrscheinlichkeitsmaß ist ein Wahrscheinlichkeitsmaß auf der Sphäre, also einer Kugeloberfläche. Es beschreibt die Verteilung von zufälligen Punkten[e] auf der Sphäre. Dabei ist folgenden sind zu betrachten:

- i. Für den Wahrscheinlichkeitsmaß ist Wahrscheinlichkeitsraum  $(\Omega, |\cdot|, P)$ , wobei  $|\cdot|$  die Euklidische Norm (Seite 3) ist.
- ii. Dies ist eine rotationsinvariant auf die Kugeloberfläche  $S^{(n-1)} := \{x \in \mathbb{R}^n \mid |x| = 1\}$  normiert. D.h. für ein  $A \subseteq S^{n-1}$  und eine Rotation Matrix  $T$  gilt:  $m(\hat{A}) = m(TA) = m(A)$ , wobei  $TX = \{\alpha x \mid x \in X, \alpha \in [0, 1]\}$  einfache Darstellung  $X$  nach der Rotation ist (Wir erinnern uns, dass Elemente einer Rotationsmatrix immer zwischen 0 und 1 liegen).
- iii. Wir betrachten eine Funktion  $f: S^{(n-1)} \rightarrow \mathbb{R}$  und stellen wir uns vor, dass eine Wahrscheinlichkeitsdichtefunktion über den Kugel definiert ist. Sei ferner

$$\mathbb{P}[f \leq t] = \mathbb{P}\left[\{x \in S^{(n-1)} \mid f(x) \leq t\}\right].$$

Der **Median** der Funktion  $f$  ist gezeichnet als **med(f)** und definiert durch  $\sup_t \mathbb{P}[f \leq t]$ .

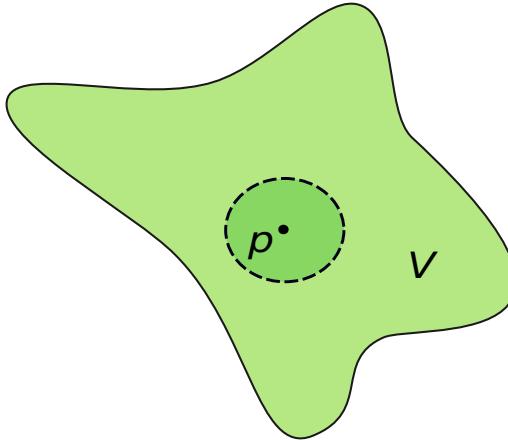


Abbildung 16<sup>‡</sup>: Menge  $V$  ist eine Umgebung des Punktes  $p$

### Lemma 3.3.2

Sei  $F$  ein  $k$ -dimensionaler affiner Unterraum und  $P_F: \mathbb{R}^d \rightarrow F$  eine Projektion, die jeden Punkt  $x \in \mathbb{R}^d$  auf seinen nächste Umgebung zum  $F$  bildet, dann wird  $P$  als die Projektion eine Kontraktion (d.h. 1 Lipschitz-stetig). Mit anderen Worten gilt es folgende Aussage:

$$|P_F(p) - P_F(q)| \leq |p - q|.$$

**Beweis:** Offenbar ist die Projektion  $P_F$  eine affine Abbildung zwischen  $\mathbb{R}^d$  (Subtraktion:) und  $F$ . Nach der Lemma 3.2.3 im letzten Abschnitt existieren eine Matrix  $M$  und einen Vektor  $b$  für  $P_F(p) = o$  wie folgt:  $P_F(p) = Mp + b$ .

Mit der Definition der Projektion ergibt sich:

$$|(M - I)p + b| = |Mp + b - p| = |P_F(p) - p| = |o - p| < \varepsilon^{(1)}.$$

Aus der Dreieckungleichung ist uns bekannt:  $|(M - I)p + b| \leq |(M - I)p| + |b|^{(2)}$ .

Aus (1) und (2) zusammen folgt:  $|(M - I)p| \leq |(M - I)p| + |b| \leq \varepsilon \equiv |(M - I)p| \leq 0$ , als  $\varepsilon \rightarrow 0$  (od.  $|Mp| \leq |p|$ , als  $\varepsilon \rightarrow 0$ ). Nun aus der letzte Ungleichung folgt:

‡ Bildquelle aus dem Wikipedia.

$$|P_F(p) - P_F(q)| \leq |Mp + b - Mq - b| \leq |M(p - q)| \leq |p - q|. \blacksquare$$

**Lemma 3.3.3**

Für jede  $\hat{a} \in T(A) = \widehat{A}$  (od.  $TA$  oben) und  $\hat{b} \in T(B) = \widehat{B}$  gilt:  $\left| \frac{\hat{a} + \hat{b}}{2} \right| \leq 1 - \frac{t^2}{8}$ .

*Beweis:* Seien  $\hat{a} = \alpha a$  und  $\hat{b} = \beta b$  für  $a \in A$  und  $b \in B$  (Abbildung 17) und zum Vereinfachen zeichnen wir  $|Om_1| = |m_1|$ ,  $|Oa| = |a|$ , und  $|Ob| = |b|$ , ferner sei einer positiven Parameter  $t$  mit  $t \leq |a - b|$  daraus folgt :

$$|m_1| = \left| \frac{a + b}{2} \right| = \sqrt{1^2 - \left| \frac{a - b}{2} \right|^2} = \sqrt{1^2 - \left( \frac{t}{2} \right)^2} \leq \sqrt{1^2 - \frac{t^2}{8}} \leq 1 - \frac{t^2}{8} \text{ (1).}$$

In Bezug  $\hat{a}$  und  $\hat{b}$  nehmen wir an, dass  $\alpha \leq \beta$  gilt. Es ist zu beobachten, dass die Größe  $|\hat{a} + \hat{b}|$  maximiert wird, wenn  $\beta = 1$  ist. Demzufolge ergibt sich durch Anwendung der Dreiecksungleichung:

$$\left| \frac{\hat{a} + \hat{b}}{2} \right| = \left| \frac{\alpha a + \beta b}{2} \right| = \left| \frac{\alpha a + b}{2} \right| \leq \left| \frac{\alpha(a + b)}{2} \right| + \left| \frac{(1 - \alpha)b}{2} \right| \leq \alpha \left( 1 - \frac{t^2}{8} \right) + (1 - \alpha) \frac{1}{2} = \tau.$$

Offenbar ist  $\tau$  eine konvexe Kombination von  $\frac{1}{2}$  und  $\left( 1 - \frac{t^2}{8} \right)$ . Insbesondere schließen wir daraus,

dass  $\tau \leq \max \left\{ \left( 1 - \frac{t^2}{8} \right), \frac{1}{2} \right\} \leq \left( 1 - \frac{t^2}{8} \right)$ , da  $t \leq 2$ .  $\blacksquare$

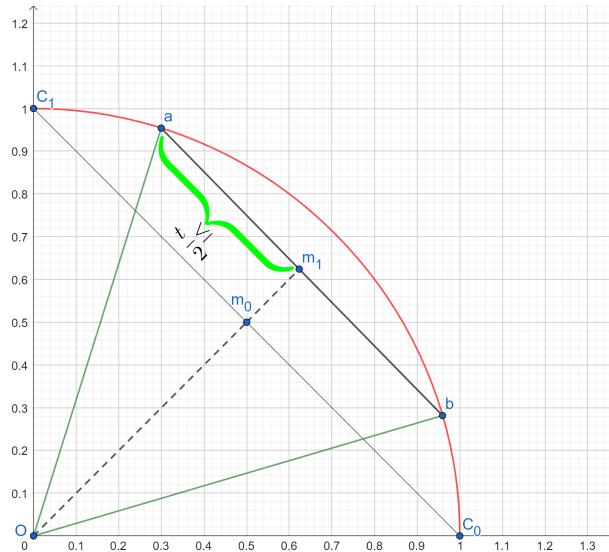


Abbildung 17: Teil der Oberfläche eines Kugelsegments

**Satz 3.3.4 (Maßkonzentration auf der Kugel).**

Sei  $A \subseteq S^{(n-1)}$  eine messbare Menge mit  $\mathbb{P}[A] \geq \frac{1}{2}$  und sei  $A_t$  die Menge der Punkte von  $S^{(n-1)}$  in einem Abstand von höchstens  $t$  von  $A$ , wobei  $t \leq 2$  ist. Dann gilt  $1 - \mathbb{P}[A_t] \leq 2 \exp\left(\frac{-nt^2}{2}\right)$  [SHP018].

*Beweis:* Wir stellen eine leichte Abweichung von der oberen Behauptung bzw. die Abschätzung zum Beweis, indem wir der Wurzel der Euler Funktion auf  $\sqrt{\exp\left(\frac{-nt^2}{2}\right)} = \exp\left(\frac{-nt^2}{4}\right)$  umstellen.

Wir wählen zuerst einen zufälligen Punkt  $p$  gleichmäßig innerhalb der Einheitskugel  $B_1$ . Sei  $\psi$  die

Wahrscheinlichkeit, dass  $p$  in  $\widehat{A}$  liegt. Offensichtlich ist  $\text{vol}(\widehat{A}) = \psi \text{vol}(B_1)$ . Wir betrachten also den normalisierten Punkt  $q = \frac{p}{|p|}$ . Nach der Definition von  $\widehat{A}$  offensichtlich liegt  $p$  in  $\widehat{A}$  genau

dann, wenn  $q$  in  $A$  liegt. Somit ist  $m(\widehat{A}) = \frac{\text{vol}(\widehat{A})}{\text{vol}(B_1)} = \psi = \mathbb{P}[p \in \widehat{A}] = \mathbb{P}[q \in A] = \mathbb{P}[A]$ , weil  $q$

nach Annahme eine gleichmäßige Verteilung auf der Hydrosphäre hat. Sei  $B = S^{(n-1)} \setminus A_t$  und  $\widehat{B} = T B$ , wobei  $A_t = \{x \in \mathbb{R}^n \mid d(x, A) < t\}$  als  $t$ -Verfetten (engl.  $t$ -Inflation) (eine Innere Menge) ist (siehe Abbildung 18<sup>†</sup>). Für alle  $a \in A$  und  $b \in B$  haben wir dann  $\|a - b\| \geq t$ . Durch die Anwendung der letzten Lemma wissen wir, dass  $\left(\frac{\widehat{A} + \widehat{B}}{2}\right)$  im Ball  $rB_1$  mit  $r = \left(1 - \frac{t^2}{8}\right)$  beinhaltet ist. Wir betrachten, dass  $m(\widehat{A}) = m(rB_1) = \frac{\text{vol}(rB_1)}{\text{vol}(B_1)} = r^n = \left(1 - \frac{t^2}{8}\right)^n$ . Mit der Anwendung vom

speziellen Fall der Brunn-Minkowski Ungleichung ergibt es sich

$$\left(1 - \frac{t^2}{8}\right)^n = m(rB_1) = m\left(\frac{\widehat{A} + \widehat{B}}{2}\right) \geq \sqrt{m(\widehat{A})m(\widehat{B})} = \sqrt{\mathbb{P}[A]\mathbb{P}[B]} \geq \sqrt{\frac{\mathbb{P}[B]}{2}}.$$

Damit ist wegen  $1 - x \leq \frac{1}{e^x} \leq e^{-x}$  (für  $x \geq 0$ )  $\mathbb{P}[B] \leq 2\left(1 - \frac{t^2}{8}\right)^n \leq 2 \exp\left(-\frac{2nt^2}{8}\right) = 2 \exp\left(-\frac{t^2}{4}\right)$ .

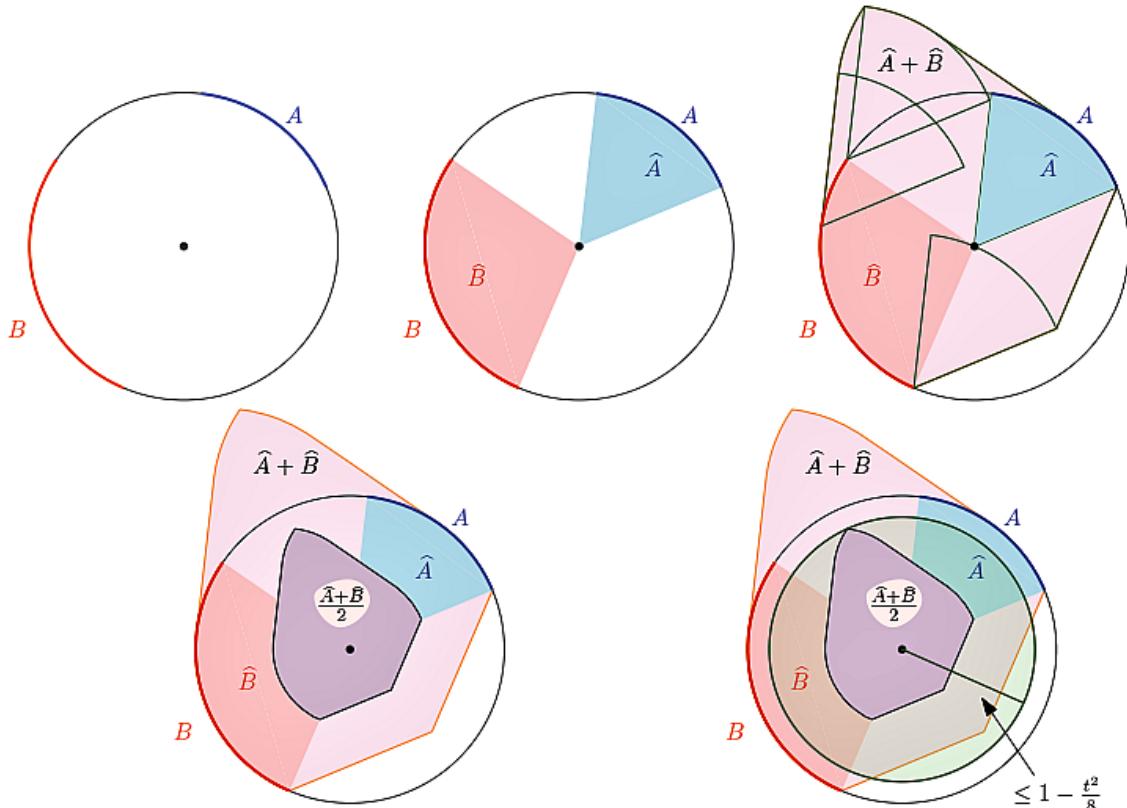


Abbildung 18<sup>†</sup>: Demonstration A und B, ihre Rotationen  $\widehat{A}, \widehat{B}$  auf der Kugeloberfläche  $S^{(n-1)}$  im Satz 3.3.4.

<sup>†</sup> Bildquelle aus dem ursprünglichen Artikel vom Sariel Har-Peled.

**Lemma 3.3.5**

Für die oben beschriebenen Wahrscheinlichkeitsmaß gilt es die folgende:

$$\begin{aligned} \mathbb{P}[f \leq \text{med}(f)] &\leq \frac{1}{2} \text{ und } \mathbb{P}[f > \text{med}(f)] \leq \frac{1}{2}. \\ \text{Beweis:} \quad \text{Weil (2.6.1)} \{f(x) < \text{med}(f)\} &= \bigcup_{k=1}^{\infty} \left\{ f(x) \leq \text{med}(f) - \frac{1}{k} \right\} \\ &\rightsquigarrow = \bigcup_{k=1}^{\infty} \left( -\infty, \text{med}(f) - \frac{1}{k} \right] = (-\infty, \text{med}(f)) \\ &= \bigcup_{k=1}^{\infty} \left( -\infty, \text{med}(f) - \frac{1}{k} \right] = (-\infty, \text{med}(f)), \text{ haben wir} \\ \mathbb{P}[f \leq \text{med}(f)] &= \sup_{k \geq 1} \mathbb{P}[f \leq \text{med}(f) - \frac{1}{k}] \leq \sup_{k \geq 1} \frac{1}{2} = \frac{1}{2}. \end{aligned}$$

Die zweite Behauptung folgt analog wie die Erste durch das folgende symmetrische Argumentation:

$$\{f(x) > a\} = \{f(x) \leq a\}^c; \forall a \in \mathfrak{R}. \blacksquare$$

Die folgende Satz wird oft auch als „Lévy's isometrische Ungleichung“ genannt und befasst sich um die Funktionen, die erste Lipschitz-Konstante erfüllen, und liefert dabei eine obere Abschätzung unter Einbeziehung von Medianen.

**Satz 3.25** (Lévy's isometrische Ungleichung)

Sei  $f: S^{(n-1)} \rightarrow \mathfrak{R}$  eine 1-Lipschitz Funktion. Dann gilt für alle  $t \in [0, 1]$ :

$$\mathbb{P}[f > \text{med}(f) + t] \leq 2 \exp\left(\frac{-nt^2}{2}\right) \text{ und } \mathbb{P}[f < \text{med}(f) - t] \leq 2 \exp\left(\frac{-nt^2}{2}\right)$$

**Beweis:** Sei  $A = \{x \in S^{(n-1)} \mid f(x) \leq \text{med}(f)\}$ . Nach dem letzten Lemma (3.2.4) haben wir  $P[A] \geq \frac{1}{2}$ . Wir betrachten nun einen Punkt  $x \in A_t$ , wobei  $A_t$  eine **t-Inflation** über  $S^{(n-1)}$  ist. Sei  $\eta(x)$  der nächste Punkt in  $A$  zu  $x$ . Wir haben dann durch Definition der  $A_t$   $\|x - \eta(x)\| \leq t$ . Andererseits folgt aus der 1-Lipschitzstetigkeit der Funktion  $f$  und  $\eta(x) \in A$ :

$$f(x) - f(\eta(x)) \leq \|x - \eta(x)\| \leq t \rightarrow f(x) \leq t + f(\eta(x)) \leq t + \text{med}(f).$$

Mit dem **Maßkonzentration auf der Kugel** Satz ergibt es sich:

$$\mathbb{P}[f > \text{med}(f) + t] = 1 - \mathbb{P}[f - \text{med}(f) \leq t] = 1 - \mathbb{P}[A_t] \leq 2 \exp\left(\frac{-nt^2}{2}\right).$$

Die zweite Behauptung folgt analog wie die Erste durch das symmetrische Argumentation.  $\blacksquare$

**3.4 Das Johnson-Lindenstrauss Lemma****Lemma 3.4.1**

Für einen Einheitsvektor  $x \in S^{(n-1)}$  sei  $\xi(x) = (x_1^2, x_2^2, \dots, x_k^2)^{\left(\frac{1}{2}\right)}$  die Länge der Projektion von  $x$  im von den ersten  $k$  Koordinaten gebildeten Unterraum. Ferner sei  $x$  ein zufällige Vektor generiert aus Gleichverteilung vom  $S^{(n-1)}$ . Dann wird die genannte Norm als scharf konzentriert mit der folgenden Eigenschaft bezeichnet:

$$\mathbb{P}[\xi(x) \geq m + t] \leq 2 \exp\left(\frac{-nt^2}{2}\right) \text{ und } \mathbb{P}[\xi(x) \leq m - t] \leq 2 \exp\left(\frac{-nt^2}{2}\right),$$

wobei die Variable  $m$  abhängig von der Anzahl der Dimension vom Original- und Unterraum ( $m = m(n, k)$ ) ist und  $t \in [0, 1]$ . Das Weiteren gilt für  $k \geq 10 \ln(n)$ , dass die Variable  $m \geq \frac{1}{2} \sqrt{\frac{k}{n}}$ .

**Beweis:** Die orthogonale Projektion  $P: \mathbb{R}^n \rightarrow \mathbb{R}^k$  gegeben durch  $p(x_1, \dots, x_n) = (x_1, \dots, x_k)$  ist nach der [Lemma 3.3.2](#) 1-Lipschitz (d.h. durch Projektionen können nur Abstände kleiner werden können). Als solche ist  $\xi(x) = |P(x)|$  1-Lipschitz, da für jedes  $x, y$  gilt

$$|\xi(x) - \xi(y)| = ||P(x)| - |P(y)|| \leq |P(x) - P(y)| \leq |x - y|.$$

[Lévy's isometrische Ungleichung](#) liefert die erforderliche Überschätzung mit  $m = \text{med}(f)$ . Somit müssen wir nur noch die untere Schranke für  $m$  beweisen.

Für ein zufälliges  $x = (x_1, \dots, x_n) \in S^{(n-1)}$  haben wir  $\mathbb{E}[|X|^2] = \mathbb{E}[1] = 1$ . Durch Linearität der Erwartungsfunktion und Symmetrie gilt für jede bestimmte  $1 \leq j \leq n$ .

$$1 = \mathbb{E}[|X|^2] = \mathbb{E}\left[\sum_{i=1}^n x_i^2\right] = \sum_{i=1}^n \mathbb{E}[x_i^2] = n \mathbb{E}[x_j^2].$$

Für  $j = 1, \dots, n$ , ist daher  $\mathbb{E}[x_j^2] = \frac{1}{n}$  und somit ergibt sich wiederum durch die Linearität:

$$\mathbb{E}[(\xi(x))^2] = \mathbb{E}\left[\sum_{i=1}^k x_i^2\right] = \sum_{i=1}^k \mathbb{E}[x_i^2] = \sum_{i=1}^k \frac{1}{n} = \frac{k}{n}.$$

Als nächstes verwenden wir, dass  $\xi(x)$  konzentriert ist, um zu zeigen, dass  $\xi^2(x)$  auch relativ konzentriert ist. Da  $\forall x \in S^{(n-1)}, \xi(x) \leq 1$ , gilt für jedes  $t \geq 0$  dann folgendes

$$\frac{k}{n} = \mathbb{E}[\xi^2] = \mathbb{P}[\xi \leq m + t](m + t)^2 + \mathbb{P}[\xi \geq m + t](1) \leq (1)(m + t)^2 + 2 \exp\left(\frac{-nt^2}{2}\right)^{(1)}.$$

Setze man nun  $t = \sqrt{\frac{k}{5n}}$ , ergibt sich mit  $k \geq 10 \ln(n)$  dass,

$$2 \exp\left(\frac{-nt^2}{2}\right) = 2 \exp\left(\frac{-k}{10}\right) \leq \frac{2}{\exp(\ln(n))} = \frac{2}{n}^{(2)}.$$

Bei (1) und (2) erhalten wir, dass  $\frac{k}{n} \leq \left(m + \sqrt{\frac{k}{5n}}\right)^2 + \frac{2}{n}$ .

Dies bedeutet, dass  $\sqrt{\frac{k-2}{n}} \leq \left(m + \sqrt{\frac{k}{5n}}\right)$  und damit ergibt sich

$$m \geq \sqrt{\frac{k-2}{n}} - \sqrt{\frac{k}{5n}} = \sqrt{\frac{1}{n}} \left( \frac{0.8k - 2}{\sqrt{k-2} + \sqrt{0.2k}} \right)^{(3)}.$$

Nun wissen wir, dass  $\forall n \in \mathbb{N}, k \geq 2 \wedge k \geq 10 \ln(n)$ . Gauner gesagt sollte dann in der letzten Aussage die  $k$  eine Wert größer-gleich  $10 \ln(2)$  haben. Dennoch schon für eine  $k \geq 4.5 < 10 \ln(2)$  gilt folgende Ungleichung:

$$k' = k - 2; \sqrt{k-2} + \sqrt{0.2k} \leq 2\sqrt{0.8(k-2.5)} \equiv \sqrt{k'+0.5} + \sqrt{0.2k'+0.5} \leq 2\sqrt{0.8k''}$$

Ersetze man diese (letzte) Ungleichung in (3) folgt direkt die gewünschte Ungleichung, nämlich:

$$m \geq \frac{1}{2} \sqrt{\frac{k}{n}}. \blacksquare$$

### Bemerkung 3.4.2

#### ➤ Projektion eines Vektors

Als nächstes möchten wir argumentieren, dass die Projektion eines festen Vektors in einen zufälligen  $k$ -dimensionalen Unterraum zu einem zufälligen Vektor führt, dessen Länge stark konzentriert ist. Dies würde bedeuten, dass wir eine Dimensionsreduktion durchführen und trotzdem Abstände zwischen Punkten erhalten können, die uns wichtig sind. Zu diesem Zweck möchten wir die letzte [Lemma](#) umkehren. Anstatt einen Punkt  $x$  zufällig auszuwählen und ihn auf den ersten  $k$ -dimensionalen Raum zu projizieren, möchten wir  $x$  festlegen und den  $k$ -dimensionalen Unterraum zufällig auswählen, in den wir projizieren. Wir müssen diesen zufälligen  $k$ -dimensionalen Raum jedoch sorgfältig auswählen. Tatsächlich, wenn wir diesen zufälligen Unterraum durch eine Transformation  $T$  rotieren, so dass er die ersten  $k$  Dimensionen einnimmt, muss der Punkt  $T(x)$  gleichmäßig auf der Sphäre verteilt sein, um die Lemma (3.4.1) verwendet werden zu können. Daher möchten wir eine zufällige Rotation von  $\mathbb{R}^n$  auswählen. Dies ordnet die Standard-Orthonormalbasis in einen zufällig rotierten Orthonormalsystem zu. Die Wahl des von den ersten  $k$  Vektoren der rotierten Basis aufgespannten (3.2.4) Unterraums ergibt einen  $k$ -dimensionalen zufälligen Unterraum. Eine solche Rotation ist eine orthonormale Matrix mit Determinante 1. Wir können eine solche Matrix erzeugen, indem wir einen Vektor  $e_1 \in S^{(n-1)v}$  zufällig auswählen. Als nächstes setzen wir  $e_1$  als erste Spalte unserer Rotationsmatrix und erzeugen die anderen  $n - 1$  zugehörigen Basisvektoren, indem wir rekursiv  $n - 1$  orthonormale Vektoren im Raum orthogonal zu  $e_1$  erzeugen [SHP018].

#### ➤ Erzeugung eines zufälligen Punktes auf der Sphäre

An dieser Stelle kann sich selbstverständlich der Leser fragen, wie man einen Punkt gleichmäßig von der Einheitssphäre  $S^{(n-1)}$  auswählt. Die Idee ist, einen Punkt aus der mehrdimensionalen Normalverteilung  $X \sim \mathcal{N}_n(\mu=0, \sigma^2=1)$  auszuwählen und ihn so zu normalisieren, dass er die Länge 1 hat. Da die mehrdimensionale Normalverteilung die Dichtefunktion  $2\pi^{\frac{-n}{2}} \exp\left(\frac{-|X|^2}{2}\right)$  hat, welche symmetrisch ist (d.h. alle Punkte haben in

Entfernung  $r$  vom Ursprung dieselbe Verteilung), folgt daraus, dass dies tatsächlich einen Punkt zufällig und gleichmäßig auf  $S^{(n-1)}$  erzeugt. Das Erzeugen eines Vektors mit mehrdimensionaler Normalverteilung ist nicht mehr als das [Auswählen](#) jeder Koordinate entsprechend der Normalverteilung. Gegeben eine Quelle von Zufallszahlen entsprechend der Gleichverteilung, kann dies mit  $O(1)$  Berechnungen pro Koordinate durchgeführt werden, indem man die Box-Muller-Transformation [BM58] verwendet. Insgesamt kann jeder Zufallsvektor in  $O(n)$  Zeit erzeugt werden. Da das Projizieren einer  $n$ -dimensionalen Normalverteilung in den niedriger dimensionierten Raum eine Normalverteilung ergibt, folgt daraus, dass das Erzeugen einer zufälligen Projektion nicht mehr ist als das zufällige Auswählen von  $n$  Vektoren entsprechend der mehrdimensionalen Normalverteilung  $v_1, \dots, v_n$ .

Dann orthonormieren wir sie mit *Graham-Schmidt*, was nicht anders als  $\hat{v}_1 = \frac{v_1}{|v_1|}$  mit

---

<sup>v</sup> Es wird als orthogonal Basisvektor bzw. erste Eigenvektor der orthogonale Matrix versehen.

normalisierte Vektor von  $v_i - w_i$  ist, wobei  $w_i$  die Projektion von  $v_i$  auf den von  $v_1, \dots, v_{i-1}$  aufgespannten (3.2.4) Raum ist. Indem wir diese Vektoren als Spalten einer Matrix nehmen, erzeugen wir eine Matrix  $A$  mit Determinante entweder 1 oder -1. Wir multiplizieren einen der Vektoren mit -1, wenn die Determinante -1 ist. Die resultierende Matrix ist eine zufällige Rotationsmatrix. Wir können nun die letzte [Lemma](#) bzgl. des genannten Szenario neu formulieren, in dem der Vektor festgelegt ist und die Projektion in einen zufälligen Unterraum erfolgt [SHP018].

### Lemma 3.4.3

Sei  $x \in S^{(n-1)}$  ein beliebiger Einheitsvektor. Betrachten wir nun einen zufälligen  $k$ -dimensionalen Unterraum  $F$  und sei dabei  $\xi(x) = (x_1^2, x_2^2, \dots, x_k^2)^{\frac{1}{2}}$  die Länge der Projektion von  $x$  in  $F$ . Dann gibt es  $m = m(n, k)$ , so dass

$$\forall t \in [0, 1], \mathbb{P}[\xi(x) \geq m + t] \leq 2 \exp\left(\frac{-nt^2}{2}\right) \text{ und } \mathbb{P}[\xi(x) \leq m - t] \leq 2 \exp\left(\frac{-nt^2}{2}\right),$$

Außerdem gilt für  $k \geq 10 \ln(n)$ , dass die Variable  $m \geq \frac{1}{2} \sqrt{\frac{k}{n}}$  [SHP018].

*Beweis:* Sei  $v_i$  der i-te orthonormale Vektor mit 1 an der i-ten Koordinate. Sei  $M$  eine zufällige Translation des Raums, die wie oben beschrieben erzeugt wurde. Offensichtlich ist für einen beliebigen festen Einheitsvektor  $x$  der Vektor  $Mx$  gleichmäßig auf der Sphäre verteilt. Nun ist die i-te Spalte der Matrix  $M$  der Zufallsvektor  $e_i$  und  $M^T v_i = e_i$ . Somit haben wir

$$\langle Mx, v_i \rangle = (Mx)^T v_i = x^T (M)^T v_i = x^T e_i = \langle x, e_i \rangle.$$

Insbesondere, ist im Fall  $Mx$  als Zufallsvektor behandelt und ihn auf die ersten  $k$  Koordinaten projiziert, haben wir

$$\xi(x) = \sqrt{\sum_{i=1}^k \langle Mx, v_i \rangle^2} = \sqrt{\sum_{i=1}^k \langle Mx, e_i \rangle^2}.$$

Abere<sub>1, ..., e<sub>k</sub></sub> ist nur eine orthonormale Basis eines zufälligen  $k$ -dimensionalen Unterraums. Somit ist der rechte Ausdruck in der letzten Gleichung die Länge der Projektion von  $x$  in einen  $k$ -dimensionalen zufälligen Unterraum. Daher hat die Länge der Projektion von  $x$  in einen zufälligen  $k$ -dimensionalen Unterraum genau dieselbe Verteilung wie die Länge der Projektion eines Zufallsvektors in die ersten  $k$  Koordinaten. Die Behauptung folgt nun aus [Lemma 3.4.1](#).▪

### Definition 3.4.4

Die Abbildung  $f : \mathfrak{N}^n \rightarrow \mathfrak{N}^k$  ist für  $X \subseteq \mathfrak{N}^n$  **Bilipschitz-Äquivalenz**, falls folgende Bedingung erfüllt wird :

$$\forall p, q \in X, \exists c > 0 \text{ s.t. } c K^{-1} |p - q| \leq |f(p) - f(q)| \leq c |p - q|.$$

Die kleinste Konstante  $K^*$ , die obere Bedingung erfüllt (d.h.  $K^* = \min K = \frac{|f(p) - f(q)|}{c |p - q|}$ ) heißt

**Verzerrung** (eng. Distortion) der Funktion  $f$  und wird oft mit  $dist(f)$  bezeichnet. Die Funktion  $f$  wird hier als eine **K-Einbettung** der  $X$  gezeichnet. Wir betone, dass es grundsätzlich analog zur Definition der Lipschitz-Konstante (3.3.(b)) die eine Bilipschitz-Äquivalenz für  $f: X \rightarrow Y$  existiert, wenn es eine reelle Konstante  $K \geq 1$  gibt, so dass für alle  $x_1$  und  $x_2$  in  $X$  in der metrischen Räumen

$d_X, d_Y$  gilt:

$$\frac{1}{K} d_X(x_1, x_2) \leq d_Y(f(x_1), f(x_2)) \leq K d_X(x_1, x_2).$$

Mit anderem Wort ist die **erste Bedingung** (definiert Sein der Euklidische Norm) nur dann erfüllt, wenn  $K = c^2$  [NW018].

### Bemerkung 3.4.5

Sei  $X \subseteq \mathbb{R}^m$  eine Menge mit  $n$  Punkten, wobei  $m$  möglicherweise viel größer als  $n$  sein kann ( $m \gg n$ ). In diesem Fall kann aufgrund der Beobachtung der Abstände zwischen den Punkten in  $X$ ,  $X$  als eine Menge von Punkten betrachtet werden, die in dem von den Punkten der Menge  $X$  aufgespannten affinen Unterraum  $\mathcal{F}$  liegen. Dieser Unterraum hat die Dimension  $n - 1$ . Daher kann jeder Punkt von  $X$  als  $(n - 1)$ -dimensionaler Punkt in  $\mathcal{F}$  interpretiert werden<sup>¥</sup>. Mit anderen Worten kann angenommen werden, dass die Menge der  $n$  Punkte, die beobachtet werden, in  $\mathbb{R}^n$  (genauer gesagt in  $\mathbb{R}^{n-1}$ ) liegen. Sollte  $m < n$  sein, kann die Koordinaten der Punkte in  $X$  durch Nullen ersetzt werden, sodass sie in  $\mathbb{R}^n$  liegen.

### Satz 3.4.6 (Johnson-Lindenstrauss Lemma)

Gegeben sei eine Menge  $X$  mit  $n$  Punkten in einem euklidischen Raum und ein  $\varepsilon \in (0, 1]$ . Dann existiert eine  $(1 + \varepsilon)$ -**Einbettung** von  $X$  in  $\mathbb{R}^k$  mit  $k = O(\varepsilon^{-2} \log n)$ .

**Beweis:** Nach der letzten Bemerkung ist voraussetzbar, dass  $X \subseteq \mathbb{R}^n$  liegt. Es sei  $k = 200\varepsilon^{-2} \ln(n)$ . Unter der Annahme, dass  $k < n$  ist, seien  $\mathcal{F}$  ein zufälliger  $k$ -dimensionaler **linearer** Unterraum von  $\mathbb{R}^n$  und  $P_{\mathcal{F}}: \mathbb{R}^n \rightarrow \mathcal{F}$  der orthogonale Projektionsoperator von  $\mathbb{R}^n$  in der  $\mathcal{F}$ . Ferner sei  $m$  die Zahl, um die für  $x \in S^{(n-1)}$ ,  $|P_{\mathcal{F}}(x)|$  konzentriert ist, wie in [Lemma 3.4.3](#).

Für zwei Punkte  $x, y \in \mathbb{R}^n$  kann folgende Aussage mit einer Wahrscheinlichkeit größer-gleich

$1 - \left(\frac{1}{n}\right)^2$  bewiesen werden:

$$\left(1 - \frac{\varepsilon}{3}\right)m|x - y| \leq |P_{\mathcal{F}}(x) - P_{\mathcal{F}}(y)| \leq \left(1 + \frac{\varepsilon}{3}\right)m|x - y| (\diamond).$$

Da es insgesamt  $\binom{n}{2}$  Möglichkeiten zum Wählen eines Paares von Punkten in  $X$  gibt, und  $\left(\frac{1}{n}\right)^2$  die Wahrscheinlichkeit eines festen Paares, für das die obere Aussage nicht gilt (d.h.  $\left(\frac{1}{n}\right)\left(\frac{1}{n}\right)$ ) ist, ist die Wahrscheinlichkeit für mindestens ein solches Ereignis (Paare, die nicht obere Ungleichungen (Aussage) erfüllen) damit  $\binom{n}{2}\left(\frac{1}{n}\right)^2 = \frac{n(n-1)}{2}\left(\frac{1}{n^2}\right) = \frac{n-1}{2n}$ . Damit die Wahrscheinlichkeiten für

Gegenereignis nämlich die obere Aussage für alle Paare gelten sollte, ist  $1 - \frac{n-1}{2n} = \frac{n+1}{2n} > \frac{1}{2}$ .

In diesem Fall ist die Abbildung  $P$  eine **D-Einbettung** von  $X$  in  $\mathbb{R}^k$  vergleichbar mit der Definition oben  $\frac{c}{D} \geq \left(1 - \frac{\varepsilon}{3}\right)m$  mit  $c = 1 + \frac{\varepsilon}{3}$ . Ferner ist mit einer Konstante wie unten die Einbettung garantiert:

---

<sup>¥</sup> Ein affiner Raum ist in der Tat ein Unterraum eines projektiven Raums, der wiederum durch eine Äquivalenzrelation aus einem Vektorraum hervorgeht (nicht durch einen linearen Unterraum).

$$D \leq \frac{1 + \frac{\varepsilon}{3}}{(1 - \frac{\varepsilon}{3})} \leq 1 + \frac{\frac{2\varepsilon}{3}}{(1 - \frac{\varepsilon}{3})} \geq \frac{2}{3}$$

Es sei  $u = x - y$ , dann ist  $P_F(u) = P_F(x) - P_F(y)$ , da  $P_F(\cdot)$  ein linearer Operator ist. Somit wird die Bedingung zu  $(1 - \frac{\varepsilon}{3})m|u| \leq |P_F(u)| \leq (1 + \frac{\varepsilon}{3})m|u|$ .

Da die Projektion ein linearer Operator ist, ist die Bedingung für jedes  $\alpha > 0$  äquivalent zu

$$(1 - \frac{\varepsilon}{3})m|\alpha u| \leq |P_F(\alpha u)| \leq (1 + \frac{\varepsilon}{3})m|\alpha u|.$$

Daher kann angenommen werden, dass  $|u| = 1$  mit dem Wählen  $\alpha = \frac{1}{|u|}$  Es muss also gezeigt werden, dass

$$||P_F(u)| - m| \leq \frac{m\varepsilon}{3}.$$

Es sei  $f(u) = |P_F(u)|$ . Da  $m \geq \frac{1}{2}\sqrt{\frac{k}{n}}$  und  $k = 200\varepsilon^{-2} \ln(n)$ , ist für  $t = \frac{\varepsilon m}{3}$  nach der [Lemma 3.4.1](#)

(wobei der Zufallsraum mit dem Zufallsvektor vertauscht wird), die Wahrscheinlichkeit, dass dies nicht erfüllt ist, durch

$$\begin{aligned} \mathbb{P}[|f(u) - m| \leq t] &= \mathbb{P}[f(u) \leq m + t] + \mathbb{P}[f(u) \geq m - t] \leq 4 \exp\left(\frac{-nt^2}{2}\right) = 4 \exp\left(\frac{-nm^2\varepsilon^2}{18}\right) \\ &\leq 4 \exp\left(\frac{-k\varepsilon^2}{72}\right) = 4 \left(e^{-\ln(n)}\right)^{\left(\frac{200}{72}\right)} = \frac{4}{n^{\frac{25}{9}}} < \frac{1}{n^2} \end{aligned}$$

begrenzt.▪

### Bemerkung 3.4.7

Wir betrachten, dass die letzte Ungleichung im oberen Beweis nur für  $n \geq 6$  gelten. Durch Änderung, z.B Lockerung  $(1 - \frac{\varepsilon}{3}), (1 + \frac{\varepsilon}{3})$  in  $(*)$   $(1 - \frac{\varepsilon}{2})$  bzw.  $(1 + \frac{\varepsilon}{2})$  wird die genannte Ungleichung für eine  $n \geq 2$  gelten.

## 4 Anhang

---

### 4.1 Abbildungsverzeichnis

Abbildung 1*: Eine lineare Dimensionalitätsreduktion.....	2
Abbildung 2: Ein Rechteck in.....	6
Abbildung 3: Ein Rechteck in $\mathbb{R}^2$ .....	6
Abbildung 5: Erstellung des Gitters durch die Rechtecke.....	7
Abbildung 6: Ein konstruiertes Gitter.....	7
Abbildung 7: Der erste Schritt im Gittergenerierungsprozess für einen offenen Raum in $\mathbb{R}^2$ .....	8
Abbildung 8: Der zweite Schritt im Gittergenerierungsprozess für den offenen Raum in $\mathbb{R}^2$ .....	8
Abbildung 9: Slices $E_y$ und $E_x$ (für feste $x$ und $y$ ) in einer Menge $E$ .....	24
Abbildung 103: Minkowski Addition $Q_1 + Q_2 = [0, 1]^2 + [1, 2]^2 = [1, 3]^2$ .....	27
Abbildung 11: Approximation einer rektifizierbar-en Kurve durch die Geraden $f, g, h, i, j$ und $k$ ....	31
Abbildung 12: Eine quasie-einfache Kurve.....	32
Abbildung 13: Die Kurve $\Gamma$ und die Menge $\Gamma\delta$ .....	33
Abbildung 14: Die Mengen $\Omega_{-}, \Omega$ , und $\Omega_{+} \in \mathbb{R}^2$ .....	36
Abbildung 15: Die Standardnormalverteilungsfunktion für eine 2-dimensionale Zufallsvariable....	38
Abbildung 16†: Menge $V$ ist eine Umgebung des Punktes $p$ .....	41
Abbildung 17: Teil der Oberfläche eines Kugelsegments.....	42
Abbildung 18†: Demonstration A und B, ihre Rotationen auf der Kugeloberfläche im Satz 3.3.4....	43

### 4.2 Bibliographie

- [L. JP009]: L. Jacques, Presentation slides of Laurent Jacques, 2009
- [NW018]: Nik Weaver, Lipschitz Algebras, 2018
- [SHP018]: Sariel Har-Peled, The Johnson-Lindenstrauss Lemma, 2018
- [RAR0052]: Elias M. Stein, Rami Shakarchi, Real Analysis, 2005
- [BM984]: Marcel Berger, Problems in Geometry, 1984
- [OGUE010]: Osman Güler, Foundations of Optimization, 2010
- [ARGB007]: Alvin C. Rencher and G. Bruce Schaalje, Linear Model In Statistics, 2007
- [SMR019]: Sheldon M. Ross, Introduction to Probability Models, 2019
- [LVSM006]: LaValle, Steven M, Planning Algorithms, 2006
- [FJIR001]: Frank Jones, Lebesgue Integration on Euclidean Space, 2001
- [RAR0051]: Elias M. Stein, Rami Shakarchi, Real Analysis, 2005