

SKRIPTUM ZUR VORLESUNG

---

---

# FRAKTALE

---

---

VON

MARKUS PASSENBRUNNER

*Institut für Analysis*

*Johannes Kepler Universität Linz*

WINTERSEMESTER 2015/16



JOHANNES KEPLER  
UNIVERSITÄT LINZ | JKU

# Inhaltsverzeichnis

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>Hausdorff-Metrik und iterierte Funktionensysteme</b>                    | <b>3</b>  |
| 1.1      | Ein erstes Beispiel – Die Cantormenge . . . . .                            | 3         |
| 1.2      | Hausdorff-Metrik . . . . .   | 5         |
| 1.3      | Iterierte Funktionensysteme . . . . .                                      | 8         |
| 1.4      | Beispiele . . . . .  | 10        |
| 1.5      | Iterierte Funktionensysteme mit Kondensation . . . . .                     | 12        |
| <b>2</b> | <b>Hausdorffmaß und -dimension</b>   | <b>13</b> |
| 2.1      | Erinnerung an die Maßtheorie . . . . .                                     | 13        |
| 2.2      | Hausdorffmaß . . . . .   | 18        |
| 2.3      | Hausdorffdimension . . . . .   | 24        |
| 2.4      | Einige Hilfsmittel zur Berechnung der Hausdorffdimension . . . . .         | 25        |
| 2.5      | Maße auf Fraktalen . . . . .   | 28        |
| 2.6      | Dimension von selbstähnlichen Mengen . . . . .                             | 29        |
| <b>3</b> | <b>Markovketten und Ergodensätze</b>                                       | <b>32</b> |
| 3.1      | Bedingte Erwartung . . . . .   | 32        |
| 3.2      | Markovketten . . . . .   | 37        |
| 3.3      | Ein Ergodensatz . . . . .  | 40        |
| 3.4      | Anwendungen . . . . .  | 45        |
| <b>4</b> | <b>Weitere Fraktale</b>  | <b>49</b> |
| 4.1      | Spezielle Teilmengen von $[0, 1)$ . . . . .                                | 49        |
| 4.2      | Brownsche Bewegung . . . . .   | 51        |
| <b>5</b> | <b>Julia-Mengen</b>  | <b>55</b> |
| 5.1      | Definition und einfache Eigenschaften von ausgefüllten Juliamengen . . . . | 55        |
| 5.2      | Die Mandelbrot-Menge . . . . .   | 61        |
| 5.3      | Fixpunkte . . . . .  | 68        |
| 5.4      | Periodische Punkte und Zyklen . . . . .                                    | 69        |
| 5.5      | Julia-Mengen $J(f)$ . . . . .  | 71        |
| 5.6      | Weitere Eigenschaften von Julia-Mengen und der Mandelbrot-Menge . . .      | 76        |

# 1 Hausdorff-Metrik und iterierte Funktionensysteme

## 1.1 Ein erstes Beispiel – Die Cantormenge

Jedes  $x \in [0, 1]$  hat eine 3-adische Entwicklung, es gibt also für alle  $x \in [0, 1]$  eine Folge  $a \in \{0, 1, 2\}^{\mathbb{N}}$  mit

$$x = \sum_{j=1}^{\infty} a_j 3^{-j} = 0, a_1 a_2 \dots_3.$$

Diese Entwicklung ist eindeutig, falls  $x$  nicht von der Form  $x = p3^{-k}$  für  $p, k \in \mathbb{N}$  ist. Falls  $x = p3^{-k}$  mit  $3 \nmid p$ , so gibt es zwei Entwicklungen von  $x$ . Für eine ist  $a_k = 1$  und für die andere ist  $a_k = 0$  oder  $a_k = 2$ . Wir verstehen im folgenden unter einer Entwicklung von  $x$  immer die letztere mit  $a_k \neq 1$ . Beispielsweise gilt

$$\begin{aligned} a_1 = 1 &\iff \frac{1}{3} < x < \frac{2}{3}, \\ a_1 \neq 1, a_2 = 1 &\iff \frac{1}{9} < x < \frac{2}{9} \text{ oder } \frac{7}{9} < x < \frac{8}{9}. \end{aligned}$$

**Definition 1.1.** Die Cantormenge  $C$  ist die Menge aller  $x \in [0, 1]$  mit Basis-3 Entwicklung  $x = \sum_{j=1}^{\infty} a_j 3^{-j}$ , sodass für alle  $j \in \mathbb{N}$  gilt  $a_j \neq 1$ .

Die obige, etwas künstliche Wahl der jeweiligen Entwicklung von  $x$  dient dazu, dass  $C$  abgeschlossen ist.  $C$  erhält man also, indem man beginnend mit  $[0, 1]$  rekursiv die offenen mittleren Drittel der jeweils verbliebenen Intervalle entfernt. Wir beweisen einige Eigenschaften von  $C$ .

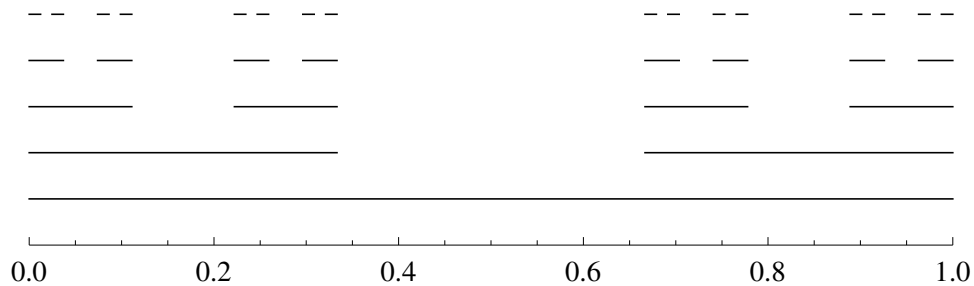


Abbildung 1: Die ersten 5 Iterationen der Konstruktion der Cantormenge

**Proposition 1.2.** Sei  $C$  die Cantormenge. Dann gilt

- (i)  $C$  ist kompakt,
- (ii)  $\lambda(C) = 0$  (dabei ist  $\lambda$  das Lebesguemaß),
- (iii)  $\text{card } C = \mathfrak{c}$ .

Unter der letzten Aussage verstehen wir, dass  $C$  die Kardinalität des Kontinuums besitzt, dass also eine Bijektion von  $C$  nach  $[0, 1]$  existiert.  $C$  ist also eine Menge, die im mengentheoretischen Sinne nicht von  $[0, 1]$  unterscheidbar ist, aber im maßtheoretischen Sinne eine Menge mit Maß 0.

*Beweis.* (i) Die Eigenschaft der Kompaktheit von  $C$  ist in  $\mathbb{R}$  äquivalent zur Abgeschlossenheit und Beschränktheit von  $C$ . Da  $C \subset [0, 1]$ , ist  $C$  beschränkt. Weiters ist  $C$  der Durchschnitt von abgeschlossenen Mengen und somit selbst abgeschlossen, also als Teilmenge von  $[0, 1]$  kompakt.

$$(ii) \lambda(C) = 1 - \frac{1}{3} - \frac{2}{9} - \frac{4}{27} - \dots = 1 - \sum_{j=0}^{\infty} \frac{2^j}{3^{j+1}} = 1 - \frac{1}{3} \frac{1}{1-2/3} = 0.$$

(iii) Sei  $x \in C$  mit  $x = \sum_j a_j 3^{-j}$ , wobei  $a_j = 0$  oder  $a_j = 2$  für alle  $j$ . Wir definieren  $f : C \rightarrow [0, 1]$  durch

$$f(x) = \sum_j b_j 2^{-j} \quad \text{mit } b_j = \frac{a_j}{2}.$$

Diese Reihe ist die Basis-2 Entwicklung einer Zahl in  $[0, 1]$ . Klarerweise erhält man auf diese Weise jede Zahl in  $[0, 1]$ .  $f$  ist also eine Surjektion. Da die identische Abbildung  $\text{id} : C \rightarrow [0, 1]$ ,  $x \mapsto x$  eine injektive Abbildung von  $C$  nach  $[0, 1]$ , sind  $C$  und  $[0, 1]$  insgesamt gleichmächtig. (vgl. Satz von Cantor-Schröder-Bernstein)  $\square$

Weiters ist leicht zu sehen, dass  $C$  keine isolierten Punkte besitzt, nirgends dicht (das Innere von  $C$  ist leer) und total unzusammenhängend ( $C$  besitzt außer den einpunktigen Teilmengen keine zusammenhängenden Mengen) ist. Dabei ist ein Punkt  $a \in C$  ein isolierter Punkt, falls eine Umgebung  $U \subset \mathbb{R}$  von  $a$  existiert, sodass  $U$  keine weiteren (außer  $a$ ) Elemente aus  $C$  enthält.

**Bemerkung 1.3.** Die Abbildung  $f$  des vorigen Beweises ist monoton, genauer gilt

$$x < y \Rightarrow f(x) < f(y),$$

falls  $x$  und  $y$  nicht die beiden Endpunkte eines in der Konstruktion von  $C$  entfernten Intervalls sind. Wir können also  $f$  zu einer monotonen Abbildung

$$f : [0, 1] \rightarrow [0, 1]$$

erweitern, indem wir sie konstant fortsetzen auf jedem Intervall das in  $C$  fehlt. Da  $f$  zusätzlich surjektiv ist, ist  $f$  auch stetig. Diese Funktion nennt man die *Cantor-Funktion* oder auch „*devils staircase*“. Aus Proposition 1.2,2. folgt, dass  $f$  fast überall bzgl. des Lebesguemaßes konstant ist, also  $f' \equiv 0$  fast überall und wir erhalten

$$1 = f(1) - f(0) \neq \int_0^1 f'(t) dt = 0.$$

Vergleiche hierzu folgende Definition und den anschließenden Satz:

**Definition 1.4.** Eine Funktion  $f : \mathbb{R} \rightarrow \mathbb{C}$  ist absolut stetig, falls für alle  $\varepsilon > 0$  ein  $\delta > 0$  existiert, sodass für alle endlichen Mengen von disjunkten Intervallen  $(a_1, b_1), \dots, (a_n, b_n)$  gilt

$$\sum_{j=1}^n (b_j - a_j) < \delta \Rightarrow \sum_{j=1}^n |f(b_j) - f(a_j)| < \varepsilon.$$

Weiters ist  $f$  absolut stetig auf  $[a, b]$  falls diese Bedingung für alle disjunkten Intervalle in  $[a, b]$  erfüllt ist.

**Satz 1.5** (Hauptsatz der Differential- und Integralrechnung für Lebesgue-Integrale). Für  $a, b \in \mathbb{R}$  mit  $a < b$  sind äquivalent

- (i)  $f$  ist absolut stetig auf  $[a, b]$ ,
- (ii)  $f(x) - f(a) = \int_a^x g(t) dt$  für ein  $g \in L^1[a, b] := \{f : [a, b] \rightarrow \mathbb{R} \text{ mb.}, \int_a^b |f(x)| dx < \infty\}$ ,
- (iii)  $f$  ist fast überall differenzierbar auf  $[a, b]$ ,  $f' \in L^1[a, b]$  und  $f(x) - f(a) = \int_a^x f'(t) dt$ .

Vergleiche hierzu den Satz von Radon-Nikodým (Satz 2.15), der die Äquivalenz der ersten zwei Punkte dieses Satzes auf allgemeinere Maßräume ausdehnt.

Es fällt auf, dass die Cantormenge mit einem rekursiven Prozess konstruiert wurde. Weiters ist  $C$  kompakt. Wir werden sehen, dass allgemeinere fraktale Mengen diese Eigenschaften auch besitzen. Dabei verstehen wir unter einem Fraktal eine Teilmenge eines euklidischen Raumes die keine ganzzahlige Dimension hat. Wir werden den Begriff der Dimension später noch präzisieren.

## 1.2 Hausdorff-Metrik

Sei  $(X, d)$  ein metrischer Raum. Das heißt,  $X$  ist eine Menge und  $d$  ist eine Abbildung von  $X \times X$  in die Menge der nichtnegativen reellen Zahlen mit den drei Eigenschaften

- (i)  $d(x, y) = 0$  für  $x, y \in X$  genau dann wenn  $x = y$ ,
- (ii)  $d(x, y) = d(y, x)$ ,  $x, y \in X$ ,
- (iii)  $d(x, y) \leq d(x, z) + d(z, y)$ ,  $x, y, z \in X$ .

Sei  $A \subset X$  eine Teilmenge von  $X$ . Dann ist für  $x \in X$  der Abstand  $d(x, A)$  von  $x$  zu  $A$  definiert durch

$$d(x, A) := \inf\{d(x, y) : y \in A\}.$$

Analog ist für eine weitere Menge  $B \subset X$

$$d(B, A) := \inf\{d(x, A) : x \in B\}$$

der Abstand von  $B$  zu  $A$ . Eine Teilmenge  $A$  eines metrischen Raumes  $X$  ist kompakt, falls jede Folge in  $A$  eine in  $A$  konvergente Teilfolge besitzt.

**Definition 1.6.** Wir bezeichnen mit

$$\mathcal{K}(X) := \{K \subset X : \emptyset \neq K \text{ kompakt}\}$$

die Menge der nichtleeren, kompakten Teilmengen des metrischen Raums  $X$ .

Sei ab jetzt  $(X, d)$  ein vollständiger metrischer Raum. Das heißt, dass jede Cauchyfolge in  $X$  einen Grenzwert in  $X$  besitzt. Wir werden dann für  $\mathcal{K}(X)$  eine Metrik einführen (die sogenannte *Hausdorff-Metrik*) und zeigen, dass  $\mathcal{K}(X)$  mit dieser Metrik ebenfalls ein vollständiger Raum ist.

**Definition 1.7** (Hausdorff-Metrik). Seien  $A \in \mathcal{K}(X)$  und  $\delta > 0$ . Dann definieren wir

$$A_\delta := \{x \in X : d(x, A) \leq \delta\},$$

die  $\delta$ -Parallelmenge von  $A$ . Die Hausdorff-Metrik  $h$  ist dann definiert durch

$$h(A, B) := \inf\{\delta > 0 : A \subset B_\delta, B \subset A_\delta\} \quad \text{für } A, B \in \mathcal{K}(X).$$

Einfach zu zeigen ist folgendes

**Lemma 1.8.**  $h$  ist eine Metrik auf  $\mathcal{K}(X)$ .

*Beweis.* Übung. □

**Satz 1.9.** Sei  $(X, d)$  ein vollständiger metrischer Raum. Dann ist  $(\mathcal{K}(X), h)$  ebenfalls ein vollständiger metrischer Raum und es gilt für eine Cauchyfolge  $A_n$  in  $(\mathcal{K}(X), h)$ , dass

$$\lim_{n \rightarrow \infty} A_n = \{a \in X : \text{es existiert eine Folge } (a_n) \text{ mit } a_n \in A_n \text{ und } \lim a_n = a\} =: A$$

und es gilt  $A \in \mathcal{K}(X)$ .

Für den Beweis dieses Satzes benötigen wir das folgende

**Lemma 1.10** (Erweiterung von Cauchy-Folgen). Sei  $(A_n)$  eine Cauchyfolge in  $(\mathcal{K}(X), h)$  und sei  $(n_j)_{j=1}^\infty$  eine streng wachsende Folge natürlicher Zahlen. Sei weiters  $(a_{n_j})_{j=1}^\infty$  eine Cauchyfolge von Punkten  $a_{n_j} \in A_{n_j}$ . Dann existiert eine Cauchyfolge  $(\tilde{a}_n)$  mit  $\tilde{a}_n \in A_n$  und  $\tilde{a}_{n_j} = a_{n_j}$  für alle  $j \in \mathbb{N}$ .

*Beweis.* Wir definieren die Folge  $\tilde{a}_n$  induktiv. Für  $j = 1$  und  $1 \leq n \leq n_1$  wählen wir  $\tilde{a}_n \in A_n$  so, dass  $d(a_{n_1}, A_n) = d(a_{n_1}, \tilde{a}_n)$ . Klarerweise gilt dann  $\tilde{a}_{n_1} = a_{n_1}$ . Aufgrund der Kompaktheit von  $A_n$  ist diese Wahl möglich. Für  $n_j + 1 \leq n \leq n_{j+1}$  wählen wir  $\tilde{a}_n \in A_n$  so, dass  $d(a_{n_{j+1}}, A_n) = d(a_{n_{j+1}}, \tilde{a}_n)$ . Somit gilt für alle  $n$ , dass  $\tilde{a}_n \in A_n$  und für alle  $j$ , dass  $\tilde{a}_{n_j} = a_{n_j}$ . Der Beweis, dass die so definierte Folge eine Cauchyfolge ist, involviert ein  $\varepsilon/3$ -Argument und das Ausnutzen, dass sowohl  $(a_{n_j})_j$  als auch  $(A_n)_n$  eine Cauchyfolge sind. Wir wollen das dem Leser als Übungsaufgabe überlassen. □

Wir benötigen auch noch ein paar klassische Definitionen: Eine Teilmenge  $S \subset X$  eines metrischen Raums  $(X, d)$  heißt *beschränkt*, falls es  $a \in X$  und  $R > 0$  gibt, sodass für alle  $x \in S$  gilt

$$d(a, x) \leq R.$$

Eine Teilmenge  $S \subset X$  eines metrischen Raums  $(X, d)$  heißt *präkompakt*, falls für alle  $\varepsilon > 0$  eine endliche Menge  $\{y_1, \dots, y_n\} \subset S$  existiert, sodass es für alle  $x \in S$  ein  $1 \leq i \leq n$  gibt mit

$$d(y_i, x) \leq \varepsilon.$$

Anders gesagt ist  $S$  präkompakt genau dann wenn es zu jedem  $\varepsilon > 0$  ein endliches  $\varepsilon$ -Netz für  $S$  gibt. Bekanntermaßen gilt dann folgender

**Satz 1.11.** Sei  $(X, d)$  ein vollständiger, metrischer Raum. Eine Teilmenge  $S \subset X$  ist genau dann kompakt, wenn  $S$  abgeschlossen und präkompakt ist.

Wir beweisen nun die Vollständigkeit von  $(\mathcal{K}(X), h)$ .

*Beweis von Satz 1.9.* Sei  $(A_n)$  eine Cauchyfolge in  $(\mathcal{K}(X), h)$ . Wir zerlegen den Beweis in mehrere Schritte

**Schritt 1;  $A \neq \emptyset$ :** Wir zeigen die Existenz einer Cauchyfolge  $(a_i)$  mit  $a_i \in A_i$  für alle  $i$ . Da  $(A_n)$  eine Cauchyfolge ist, existiert eine Folge positiver Zahlen  $N_1 < N_2 < \dots$  mit

$$h(A_m, A_n) \leq 2^{-i} \quad \text{für alle } m, n \geq N_i.$$

Wir wählen nun  $x_{N_1} \in A_{N_1}$  beliebig. Dann gibt es wegen  $h(A_{N_1}, A_{N_2}) < 1/2$  ein  $x_{N_2} \in A_{N_2}$  mit  $d(x_{N_1}, x_{N_2}) < 1/2$ . Induktiv erhalten wir eine Folge  $(x_{N_j})_j$  mit  $d(x_{N_{j+1}}, x_{N_j}) < 2^{-j}$ . Dies ist eine Cauchyfolge in  $X$  (Beweis!). Nach Lemma 1.10 existiert eine Cauchyfolge  $(a_j)$  mit  $a_{N_i} = x_{N_i}$  und  $a_j \in A_j$  für alle  $i, j$ . Diese Folge hat einen Limes ( $X$  ist vollständig!), der definitionsgemäß in  $A$  liegt, also ist  $A \neq \emptyset$ .

**Schritt 2;  $A$  ist abgeschlossen und somit vollständig:** Sei  $a = \lim_{i \rightarrow \infty} a_i$  mit  $a_i \in A$ . Wir zeigen  $a \in A$ . Zu jedem  $a_i$  existiert nach Definition von  $A$  eine Folge  $(x_{in})_n$  mit  $x_{in} \in A_n$  und  $a_i = \lim_{n \rightarrow \infty} x_{in}$ . Weiters existieren eine wachsende Folge natürlicher Zahlen  $(N_i)_i$  mit  $d(a_{N_i}, a) < 1/i$  und zu jedem  $N_i$  ein  $m_i$  mit  $d(x_{N_i m_i}, a_{N_i}) \leq 1/i$ . Setzen wir  $y_{m_i} := x_{N_i m_i}$ , so ist  $y_{m_i} \in A_{m_i}$  und  $d(y_{m_i}, a) \leq 2/i$ . Somit gilt also  $a = \lim_{i \rightarrow \infty} y_{m_i}$ . Nach Lemma 1.10 kann  $(y_{m_i})_i$  zu einer Cauchyfolge  $(z_i)_i$  mit  $z_i \in A_i$  und  $a = \lim_{i \rightarrow \infty} z_i$  erweitert werden. Es gilt also  $a \in A$ .

**Schritt 3; Für  $\varepsilon > 0$  existiert ein  $N$ , sodass für  $n \geq N$  gilt  $A \subset (A_n)_\varepsilon$ :** Fixiere  $\varepsilon > 0$ . Dann existiert ein  $N$ , sodass für  $m, n \geq N$  gilt  $h(A_m, A_n) \leq \varepsilon$ . Wählen wir nun  $n \geq N$ , dann gilt für alle  $m \geq n$ , dass  $A_m \subset (A_n)_\varepsilon$ . Für  $a \in A$  existiert eine Folge  $(a_m)$  mit  $a_m \in A_m$  und  $a = \lim_{m \rightarrow \infty} a_m$ . Nun ist  $a_m \in A_m \subset (A_n)_\varepsilon$ . Da  $(A_n)_\varepsilon$  abgeschlossen ist (warum?) gilt auch  $\lim a_m = a \in (A_n)_\varepsilon$ .

**Schritt 4;  $A$  ist kompakt:** Wir zeigen, dass  $A$  präkompakt ist. Nehmen wir an, es wäre nicht so, dann gibt es ein  $\varepsilon > 0$ , sodass kein endliches  $\varepsilon$ -Netz existiert. Dann existiert eine Folge  $(x_i)_i$ , sodass für alle  $i \neq j$  gilt:  $d(x_i, x_j) \geq \varepsilon$ . Nach Schritt 3 gilt aber für hinreichend große  $n$ , dass  $A \subset (A_n)_{\varepsilon/3}$ . Zu jedem  $x_i$  existiert also ein  $y_i \in A_n$  mit  $d(x_i, y_i) \leq \varepsilon/3$ . Da  $A_n$  kompakt ist, konvergiert eine Teilfolge  $(y_{n_i})$  von  $(y_i)$ . Es ist also

$$d(x_{n_i}, x_{n_j}) \leq d(x_{n_i}, y_{n_i}) + d(y_{n_i}, y_{n_j}) + d(y_{n_j}, x_{n_j}) < \varepsilon$$

für  $i, j$  hinreichend groß im Widerspruch zur Annahme, somit ist  $A$  präkompakt. Da  $A$  aber nach Schritt 2 auch vollständig ist, folgt aus Satz 1.11 die Kompaktheit von  $A$ .

**Schritt 5;  $\lim_{n \rightarrow \infty} A_n = A$ :** Nach Schritt 1 und Schritt 4 gilt  $A \in \mathcal{K}(X)$ . Wegen Schritt 3 reicht es zu zeigen

$$\forall \varepsilon > 0 \exists N \forall n \geq N : A_n \subset A_\varepsilon.$$

Sei  $\varepsilon > 0$  beliebig. Da  $(A_n)$  eine Cauchyfolge ist, erhalten wir eine streng monotone Folge  $(N_j)_{j=0}^\infty$  natürlicher Zahlen mit

$$A_n \subset (A_{N_1})_{\varepsilon/2} \quad \text{für alle } n \geq N_0 \quad \text{und} \quad A_{N_j} \subset (A_{N_{j+1}})_{2^{-(j+1)}\varepsilon} \quad \text{für alle } j \geq 1$$

Wir zeigen nun, dass für  $n \geq N_0$  gilt:  $A_n \subset A_\varepsilon$ . Sei also  $y \in A_n$ . Dann existiert ein  $x_{N_1} \in A_{N_1}$  mit  $d(y, x_{N_1}) \leq \varepsilon/2$  und induktiv existieren  $x_{N_{j+1}} \in A_{N_{j+1}}$  sodass  $d(x_{N_{j+1}}, x_{N_j}) \leq 2^{-(j+1)}\varepsilon$ . Aus der Dreiecksungleichung folgt nun  $d(y, x_{N_j}) < \varepsilon$  für alle  $j$ .  $(x_{N_j})_j$  ist eine Cauchyfolge, die nach Lemma 1.10 einen Grenzwert  $x \in A$  besitzt, es gilt somit auch  $d(y, x) \leq \varepsilon$ , also  $y \in A_\varepsilon$ .  $\square$

### 1.3 Iterierte Funktionensysteme

Sei  $(X, d)$  ein vollständiger metrischer Raum.

**Definition 1.12.** Sei  $f : X \rightarrow X$  eine Abbildung. Die Lipschitzkonstante von  $f$  ist definiert als

$$\text{Lip } f := \sup_{x \neq y} \frac{d(f(x), f(y))}{d(x, y)}.$$

Wir nennen  $f$  eine Kontraktion, falls  $\text{Lip } f < 1$ . Falls  $f$  eine Kontraktion ist, nennen wir  $\text{Lip } f$  auch Kontraktionsfaktor. Weiters sei  $f^0 := \text{Id}_X$  und  $f^n := f \circ f^{n-1}$  für  $n \geq 1$  die  $n$ -fache Hintereinanderausführung von  $f$ .

Ist  $f : X \rightarrow X$  eine Abbildung, so induziert  $f$  auch eine Abbildung auf der Potenzmenge  $\mathcal{P}(X)$ , die wir auch mit  $f$  bezeichnen und definiert ist als

$$f(A) = \{f(a) : a \in A\}, \quad A \subset X.$$

Ist weiters  $f$  eine stetige Abbildung, so gilt für  $A \in \mathcal{K}(X)$ , dass  $f(A) \in \mathcal{K}(X)$  (Übung). Zusätzlich überträgt sich der Kontraktionsfaktor von  $f$  auf die Mengenabbildung  $f$ :

**Lemma 1.13.** Sei  $w : X \rightarrow X$  eine Kontraktion mit Kontraktionsfaktor  $s$ . Dann ist  $w : \mathcal{K}(X) \rightarrow \mathcal{K}(X)$  eine Kontraktion mit Kontraktionsfaktor  $s$ .

*Beweis.* Seien  $A, B \in \mathcal{K}(X)$  und  $A \subset B_\delta$ . Dann gilt für alle  $a \in A$ , dass  $d(a, B) \leq \delta$ . Für fixes  $a \in A$  existiert also ein  $b \in B$  mit  $d(a, b) \leq \delta$  (warum?). Da  $w : X \rightarrow X$  eine Kontraktion mit Faktor  $s$  ist, folgt für alle  $a \in A$  die Existenz von  $b \in B$  mit  $d(w(a), w(b)) \leq s\delta$ . Das bedeutet aber  $w(A) \subset w(B)_{s\delta}$ . Analog folgt aus  $B \subset A_\delta$ , dass  $w(B) \subset w(A)_{s\delta}$ . Nach Übergang zum Infimum folgt somit

$$h(w(A), w(B)) \leq sh(A, B),$$

also die Behauptung.  $\square$

**Lemma 1.14.** Seien  $A, B, C, D \in \mathcal{K}(X)$ . Dann gilt

$$h(A \cup B, C \cup D) \leq h(A, C) \vee h(B, D).$$

Dabei ist für zwei reelle Zahlen  $s, t$  der Ausdruck  $s \vee t$  definiert als  $\max\{s, t\}$ .

*Beweis.* Es folgt aus  $B \subset D_\delta, A \subset C_\delta, D \subset B_\delta, C \subset A_\delta$ , dass

$$A \cup B \subset C_\delta \cup D_\delta \quad \text{und} \quad C \cup D \subset A_\delta \cup B_\delta.$$

Da  $(A \cup B)_\delta = A_\delta \cup B_\delta$  (Beweis?), folgt nach Übergang zum Infimum die Behauptung.  $\square$



**Lemma 1.15.** Sei  $\{w_n : 1 \leq n \leq N\}$  eine endliche Menge von Kontraktionen mit Kontraktionsfaktoren  $\{s_n : 1 \leq n \leq N\}$ . Für  $B \in \mathcal{K}(X)$  definieren wir

$$W(B) := \bigcup_{n=1}^N w_n(B).$$

Dann ist  $W : \mathcal{K}(X) \rightarrow \mathcal{K}(X)$  eine Kontraktion mit Kontraktionsfaktor  $\max_{1 \leq n \leq N} s_n$ .

*Beweis.* Induktion nach  $N$ . Der Fall  $N = 1$  folgt aus Lemma 1.13. Zusätzlich erhalten wir aus Lemma 1.14 und der Induktionsvoraussetzung

$$\begin{aligned} h(W(B), W(C)) &= h\left(\bigcup_{n=1}^N w_n(B), \bigcup_{n=1}^N w_n(C)\right) \\ &\leq h\left(\bigcup_{n=1}^{N-1} w_n(B), \bigcup_{n=1}^{N-1} w_n(C)\right) \vee h(w_N(B), w_N(C)) \\ &\leq \left(\max_{1 \leq n \leq N-1} s_n \vee s_N\right) h(B, C) = \max_{1 \leq n \leq N} s_n \cdot h(B, C). \quad \square \end{aligned}$$

Wir sind nun in der Lage die Definition von iterierten Funktionensystemem zu geben.

**Definition 1.16.** Ein iteriertes Funktionensystem (IFS) besteht aus einem vollständigen metrischen Raum  $(X, d)$  zusammen mit einer endlichen Menge von Abbildungen  $w_n : X \rightarrow X$ ,  $1 \leq n \leq N$ . Sind die  $w_n$  Kontraktionen, so spricht man von einem hyperbolischen IFS. Sein Kontraktionsfaktor ist  $\max_{1 \leq j \leq N} s_j$ , wobei  $s_j$  der Kontraktionsfaktor von  $w_j$  ist. Für ein IFS schreiben wir auch kurz  $\{X; w_n\}$ .

Zur Erinnerung geben wir noch eine Formulierung des Fixpunktsatzes von Banach:

**Satz 1.17** (Fixpunktsatz von Banach). Sei  $f : X \rightarrow X$  eine Kontraktion auf dem vollständigen metrischen Raum  $(X, d)$ . Dann hat  $f$  einen eindeutig bestimmten Fixpunkt  $x_f$ , das heißt  $f(x_f) = x_f$  und es gilt für alle  $x \in X$

$$\lim_{n \rightarrow \infty} f^n(x) = x_f.$$

Sektion 1.2 und obige Lemmata lassen sich nun wie folgt zusammenfassen:

**Satz 1.18.** Sei  $\{X; w_n\}$  ein IFS mit Kontraktionsfaktor  $s$ . Dann ist  $W : \mathcal{K}(X) \rightarrow \mathcal{K}(X)$ , definiert durch

$$W(B) := \bigcup_{n=1}^N w_n(B),$$

eine Kontraktion auf  $(\mathcal{K}(X), h)$ . Weiters existiert ein eindeutiger Fixpunkt  $A \in \mathcal{K}(X)$  von  $W$  und es gilt für alle  $B \in \mathcal{K}(X)$

$$A = \lim_{n \rightarrow \infty} W^n(B).$$

*Beweis.* Dies folgt sofort aus obigen Resultaten und dem Fixpunktsatz von Banach (Satz 1.17).  $\square$

**Definition 1.19.** Der eindeutige Fixpunkt  $A$  in obigem Satz heißt der Attraktor des IFS  $\{X; w_n\}$ .

## 1.4 Beispiele

**Beispiel 1.20** (Cantormenge). Seien  $X = [0, 1]$  und

$$w_1(x) = \frac{x}{3}, \quad w_2(x) = \frac{x}{3} + \frac{2}{3}.$$

Dann ist der eindeutige Attraktor des IFS  $\{X; w_n\}$

$$C = \lim_{n \rightarrow \infty} W^n(X)$$

die Cantormenge (vgl. Abbildung 1, abgebildet sind die Mengen  $X, W^1(X), \dots, W^4(X)$ ). Allgemeiner erhält man eine vergleichbare Menge mit  $\beta \in [0, 1/2]$  und

$$w_1(x) = \beta x, \quad w_2(x) = \beta x + 1 - \beta$$

als Attraktor.

**Beispiel 1.21** (Sierpiński-Dreieck). Das *Sierpiński-Dreieck* erhält man, indem man ein Dreieck mittels Halbierung der Seiten in vier gleiche Teildreiecke aufteilt, das mittlere Dreieck entfernt und iteriert (vgl. Abbildung 2). Dabei ist  $X$  das Dreieck mit den Eckpunkten  $(0, 0), (1, 0), (1/2, 1)$  und

$$w_1(x) = \frac{x}{2}, \quad w_2(x) = \frac{x}{2} + (1/2, 0), \quad w_3(x) = \frac{x}{2} + (1/4, 1/2).$$

Der eindeutige Attraktor von  $\{X; w_n\}$

$$\mathcal{S} := \lim_{n \rightarrow \infty} W^n(X)$$

ist dann das Sierpiński-Dreieck.

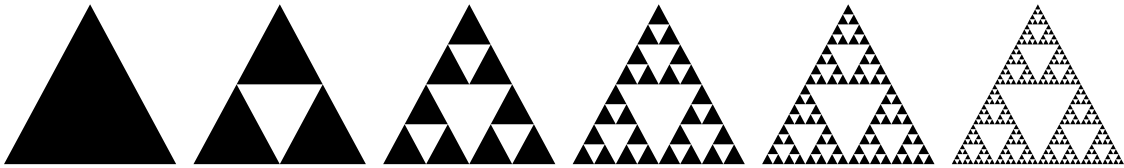


Abbildung 2: Sierpiński-Dreieck: Abgebildet sind  $X, W^1(X), \dots, W^5(X)$  mit  $X$  und  $W$  aus Beispiel 1.21.

**Beispiel 1.22** (Sierpiński-Teppich). Analog zum Sierpiński-Dreieck wird beim Sierpiński-Teppich ein Quadrat in 9 gleich große Unterquadrate aufgeteilt, das mittlere entfernt und wieder iteriert (vgl. Abbildung 3). Dabei ist  $X = [0, 1] \times [0, 1]$  und

$$w_j = \frac{x}{3} + b_j, \quad 1 \leq j \leq 8,$$

wobei der Vektor  $b$  eine Aufzählung der Menge  $\{0, 1/3, 2/3\}^2 \setminus \{(1/3, 1/3)\}$  ist. Der Sierpiński-Teppich ist der eindeutig bestimmte Attraktor dieses IFS.

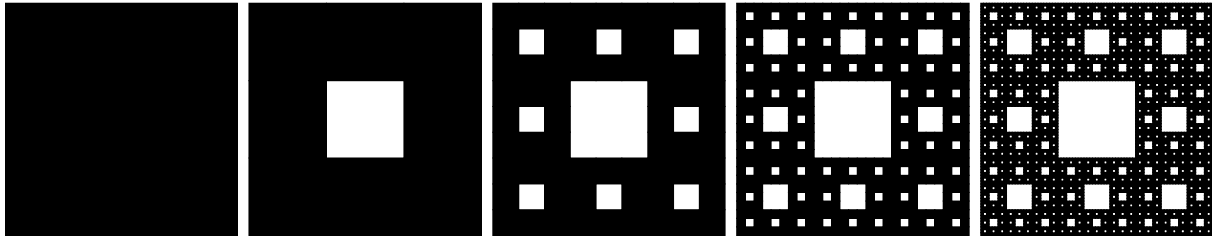


Abbildung 3: Sierpiński-Teppich: Abgebildet sind  $X, W^1(X), \dots, W^4(X)$  mit  $X$  und  $W$  aus Beispiel 1.22.

**Beispiel 1.23** (Koch-Schneeflocke). Die Kochsche Schneeflocke erhält man dadurch, dass man sich eine Einheitsstrecke vorgibt, das mittlere Drittel  $M$  dieser Strecke durch die anderen zwei Seiten eines gleichseitigen Dreiecks mit Basis  $M$  ersetzt und iteriert. Dabei gibt es zwei solche Dreiecke; man muss eines auswählen. Seien  $X = \mathbb{R}^2$  und  $L$  die Strecke von  $(0, 0)$  nach  $(1, 0)$ . Wir setzen zusätzlich

$$\begin{aligned} w_1(x) &= \frac{x}{3}, \\ w_2(x) &= \frac{R_{\pi/3}x}{3} + (1/3, 0), \\ w_3(x) &= \frac{R_{-\pi/3}x}{3} + (1/2, \sqrt{3}/6), \\ w_4(x) &= \frac{x}{3} + (2/3, 0) \end{aligned}$$

mit der Rotation  $R_\theta$  (gegen den Uhrzeigersinn) um den Winkel  $\theta$ . Die Koch-Schneeflocke ist der eindeutig bestimmte Attraktor dieses IFS. Damit wir wirklich eine „Schneeflocke“ erhalten, müssen wir diesen Prozess auf jede Seite eines gleichseitigen Dreiecks anwenden (Abbildung 5).

**Beispiel 1.24** (Pythagoras-Baum). Bei der Konstruktion des Pythagoras-Baums gibt man sich ein Einheitsquadrat  $D = [0, 1]^2$  vor und teilt die obere Kante  $[0, 1] \times \{1\}$  in einem Verhältnis  $t : (1 - t)$  für beliebiges  $0 < t < 1$ . Man zeichne dann das eindeutige rechtwinkelige Dreieck mit den Eckpunkten  $(0, 1), (1, 1)$  und dem Punkt mit  $x$ -Koordinate  $t$  und  $y$ -Koordinate  $> 1$  und nehme die Seiten dieses Dreiecks als Basis für Quadrate. Dann entstehen, wie in Abbildung 6 gesehen werden kann, zwei zusätzliche Quadrate. Dieser Prozess wird nun mit den jeweils neu entstandenen Quadraten iteriert. Formal sei  $X = \mathbb{R}^2$  und  $D = [0, 1]^2$  das Einheitsquadrat. Dann definieren wir die Abbildungen

$$\begin{aligned} w_0(A) &= D \quad \text{für alle } A \in \mathcal{K}(\mathbb{R}^2), \\ w_1(x) &= \sqrt{t}R_{\arctan \sqrt{\frac{1-t}{t}}}x + (0, 1), \\ w_2(x) &= \sqrt{1-t}R_{-\arctan \sqrt{\frac{t}{1-t}}}x + (t, 1 + \sqrt{t(1-t)}), \end{aligned}$$

wobei wieder  $R_\theta$  die Drehung um den Winkel  $\theta$  gegen den Uhrzeigersinn bezeichnet. Der Pythagoras Baum ist dann der Attraktor dieses IFS.

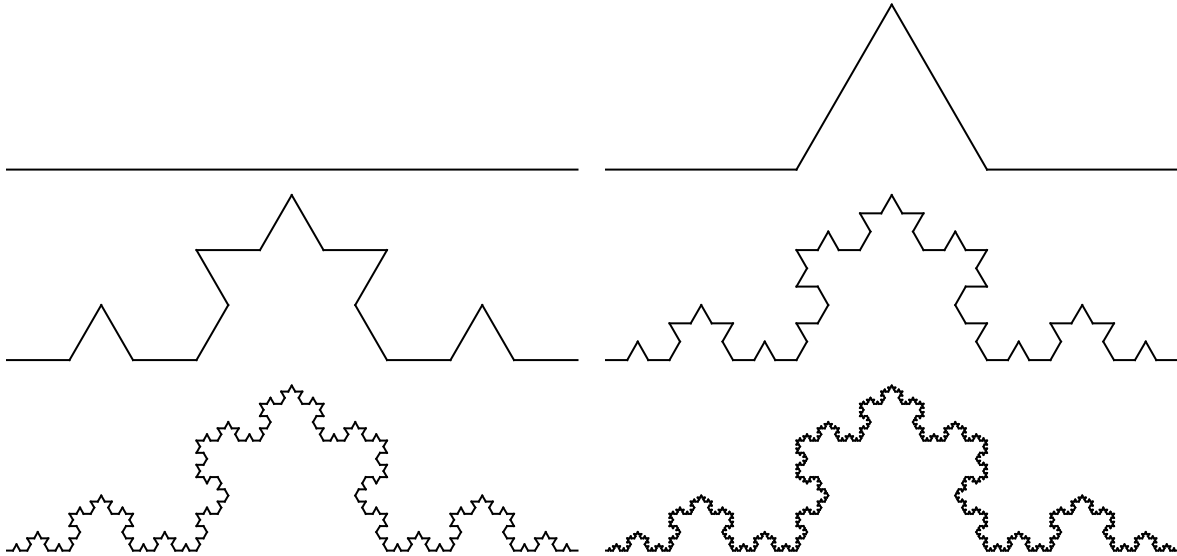


Abbildung 4: Koch-Schneeflocke: Abgebildet sind  $L, W^1(L), \dots, W^5(L)$  mit  $L$  und  $W$  aus Beispiel 1.23.

## 1.5 Iterierte Funktionensysteme mit Kondensation

Manche Fraktale, wie zB der Pythagoras-Baum, besitzen ein dominantes Strukturelement, das in immer kleineren Versionen vorkommt. Deshalb kann es hilfreich sein, sich jede Iteration zu „merken“, bevor man sie mit den Abbildungen  $w_1, \dots, w_N$  „kopiert“. Wie schon beim Pythagoras-Baum  $w_0(A) = D$  gesetzt wurde für alle  $A$ , kann dieses Vorgehen verallgemeinert werden.

**Definition 1.25.** Sei  $(X, d)$  ein metrischer Raum und sei  $C \in \mathcal{K}(X)$ . Definiere die Transformation  $w_0 : \mathcal{K}(X) \rightarrow \mathcal{K}(X)$  durch

$$w_0(B) = C$$

für alle  $B \in \mathcal{K}(X)$ . Dann heißt  $w_0$  Kondensationsabbildung und  $C$  ist die dazugehörige Kondensationsmenge. Ist  $\{X; w_1, \dots, w_N\}$  ein IFS und ist  $w_0$  eine Kondensationsabbildung, so heißt  $\{X; w_0, w_1, \dots, w_N\}$  ein IFS mit Kondensation.

Sei nun  $\{X; w_0, \dots, w_N\}$  ein IFS mit Kondensation und Kondensationsmenge  $C$ . Da die Abbildung  $w_0$  die Lipschitzkonstante 0 besitzt, ist klar, dass ein IFS mit Kondensation die gleiche Lipschitzkonstante wie das gleiche IFS ohne die Abbildung  $w_0$ . Analog zu Satz 1.18 besitzt ein IFS mit Kondensation dann einen eindeutigen Fixpunkt, den man sogar noch genauer angeben kann: Für  $B \in \mathcal{K}(X)$  seien  $W_0(B) = \bigcup_{j=0}^N w_j(B)$  und  $W(B) = \bigcup_{j=1}^N w_j(B)$ . Definiere die Folge  $C_n := W_0^n(C)$ . Dann ist  $C_n$  eine Cauchyfolge in  $\mathcal{K}(X)$ , die gegen den Attraktor  $A_0$  von  $\{X; w_0, \dots, w_N\}$  konvergiert. Beachte weiters, dass (warum?)

$$C_n = C \cup W(C) \cup \dots \cup W^n(C)$$

und es folgt (Übung!), dass

$$A_0 = \overline{\bigcup_{j=0}^{\infty} W^j(C)}.$$

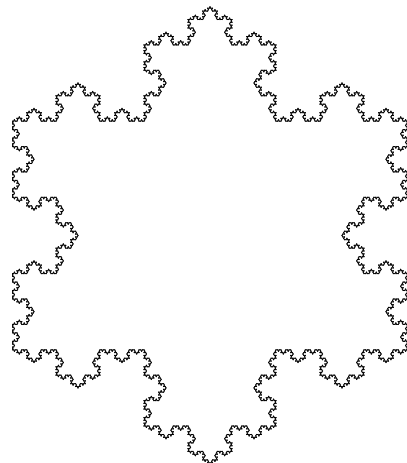


Abbildung 5: Koch-Schneeflocke:  $W^5(L)$  angewandt auf die drei Seiten eines gleichseitigen Dreiecks.

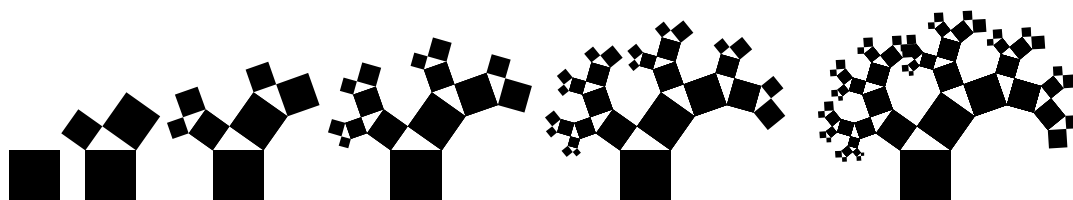


Abbildung 6: Die ersten 5 Iterationen der Konstruktion des Pythagoras-Baums mit Parameter  $t = 1/3$

**Beispiel 1.26** (Ein fraktaler Baum). Seien  $0 < \theta < \pi$ ,  $0 < r < 1$  und  $L$  die Strecke von  $(0, 0)$  nach  $(0, 1)$ . Dann definieren wir

$$w_0(B) = L, \quad w_1(x) = rR_\theta x + (0, 1), \quad w_2(x) = rR_{-\theta} x + (0, 1).$$

Siehe [Abbildung 7](#). Der eindeutig bestimmte Attraktor dieses IFS mit Kondensation ist dann  $\bigcup_{j=0}^{\infty} W^j(L)$ , wobei wie oben  $W(L) = \bigcup_{j=1}^N w_j(L)$ .

## 2 Hausdorffmaß und -dimension

In diesem Abschnitt wollen wir präzisieren, was wir mit dem Begriff der Dimension einer Teilmenge eines euklidischen Raums meinen und führen die sogenannte Hausdorffdimension von Mengen ein. Ein dazu verwandter Begriff ist das Hausdorffmaß einer Menge. Bevor wir damit beginnen, diese beiden Konzepte einzuführen, benötigen wir noch ein paar Resultate aus der Maßtheorie.

### 2.1 Erinnerung an die Maßtheorie

#### Äußere Maße

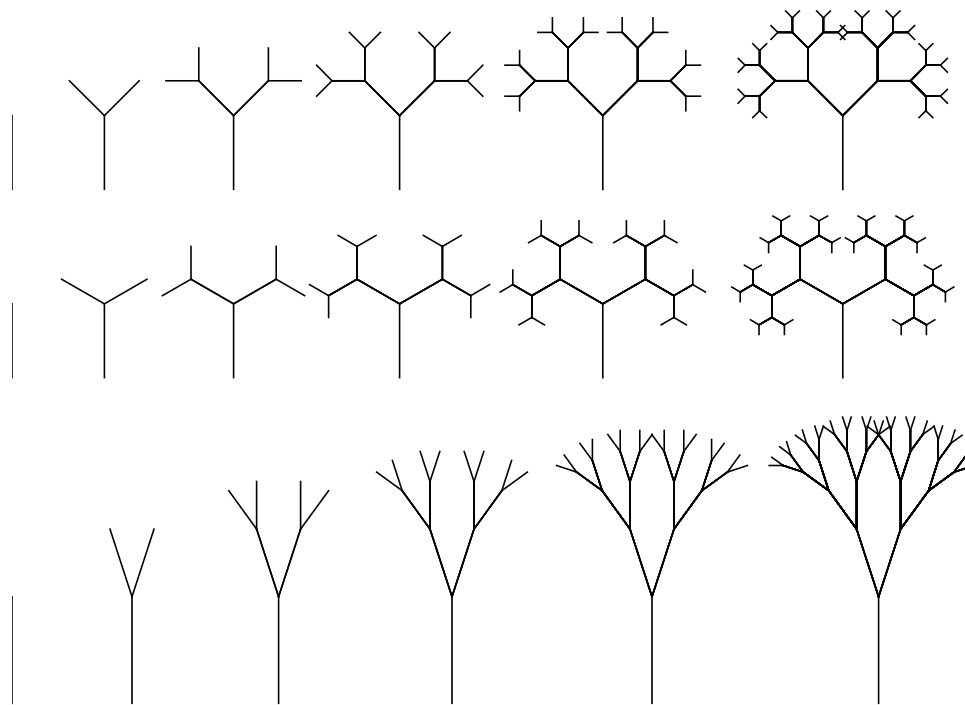


Abbildung 7: Die ersten 5 Iterationen der Konstruktion des fraktalen Baumes von Beispiel 1.26 mit Parametern  $r = 2/3$  und a)  $\theta = \pi/4$ ; b)  $\theta = \pi/3$ ; c)  $\theta = \pi/10$ .

**Definition 2.1.** Sei  $X$  eine Menge und  $\mu^* : \mathcal{P}(X) \rightarrow [0, \infty]$  eine Mengenfunktion mit

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

(ii)  $\mu^*(A) \leq \mu^*(B)$  für  $A \subset B$  (Monotonie),

(iii)  $\mu^*(\bigcup_{j=1}^{\infty} A_j) \leq \sum_{j=1}^{\infty} \mu^*(A_j)$ ,  $A_j \subset X$  (Subadditivität).

Dann heißt  $\mu^*$  ein äußeres Maß. Ist  $X$  zusätzlich ein metrischer Raum mit Metrik  $d$ , so heißt  $\mu^*$  ein metrisches äußeres Maß, falls  $\mu^*$  ein äußeres Maß ist und die Bedingung

$$\mu^*(A \cup B) = \mu^*(A) + \mu^*(B) \quad \text{für } d(A, B) > 0$$

erfüllt.

**Definition 2.2.** Sei  $X$  eine Menge und  $\mu^*$  ein äußeres Maß auf  $X$ . Eine Teilmenge  $A \subset X$  heißt  $\mu^*$ -messbar, falls

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c) \quad \text{für alle } E \subset X$$

gilt.

**Bemerkung 2.3.** Natürlich folgt aus der Subadditivität von  $\mu^*$  die Ungleichung  $\mu^*(E) \leq \mu^*(E \cap A) + \mu^*(E \cap A^c)$ . Um zu zeigen, dass eine Menge  $A$  messbar ist, reicht es also die umgekehrte Ungleichung zu zeigen. Für  $\mu^*(E) = \infty$  ist diese trivial, es gilt also, dass  $A \subset X$   $\mu^*$ -messbar ist, genau dann wenn

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

für alle  $E \subset X$  mit  $\mu^*(E) < \infty$ .

**Definition 2.4.** Sei  $(X, \mathcal{A}, \mu)$  ein Maßraum. Ist für alle  $A \in \mathcal{A}$  mit  $\mu(A) = 0$  auch jede Teilmenge von  $A$  in  $\mathcal{A}$ , so nennt man das Maß  $\mu$  vollständig.

**Definition 2.5.** Ist  $(X, d)$  ein metrischer Raum, so ist die Borelsche  $\sigma$ -Algebra  $\mathcal{B}_X$  definiert als die von den offenen Mengen in  $(X, d)$  erzeugte  $\sigma$ -Algebra. Eine Menge  $A \in \mathcal{B}_X$  heißt Borelmenge.

Die Borelsche  $\sigma$ -Algebra wird natürlich auch von den abgeschlossenen Teilmengen von  $X$  erzeugt.

**Satz 2.6** (Carathéodory). Sei  $\mu^*$  ein äußeres Maß auf einer Menge  $X$ . Dann ist die Familie  $\mathcal{A}$  der  $\mu^*$ -messbaren Mengen eine  $\sigma$ -Algebra und die Restriktion von  $\mu^*$  auf  $\mathcal{A}$  ist ein vollständiges Maß.

**Proposition 2.7.** Ist  $\mu^*$  ein metrisches äußeres Maß auf einem metrischen Raum  $(X, d)$ , dann ist jede Borelmenge  $\mu^*$ -messbar.

*Beweis.* Da die abgeschlossenen Mengen die Borelsche  $\sigma$ -Algebra erzeugen, reicht es zu zeigen, dass jede abgeschlossene Menge  $F \subset X$   $\mu^*$ -messbar ist. Wir müssen also für gegebenes  $A \subset X$  mit  $\mu^*(A) < \infty$  zeigen, dass

$$\mu^*(A) \geq \mu^*(A \cap F) + \mu^*(A \setminus F).$$

Sei  $B_n = \{x \in A \setminus F : d(x, F) \geq 1/n\}$ . Dann ist  $B_n$  eine steigende Folge von Mengen mit Vereinigung  $A \setminus F$  (da  $F$  abgeschlossen ist). Zusätzlich gilt  $d(B_n, F) \geq 1/n$ . Also folgt aus der Monotonie von  $\mu^*$  und der Definition eines metrischen äußeren Maßes

$$\mu^*(A) \geq \mu^*((A \cap F) \cup B_n) = \mu^*(A \cap F) + \mu^*(B_n).$$

Es bleibt also zu zeigen, dass  $\mu^*(A \setminus F) = \lim_{n \rightarrow \infty} \mu^*(B_n)$ . Sei nun  $C_n = B_{n+1} \setminus B_n$ . Für  $x \in C_{n+1}$  und  $y \in B_n$  folgt aus der Dreiecksungleichung

$$\frac{1}{n} \leq d(y, F) \leq d(x, y) + d(x, F) < d(x, y) + \frac{1}{n+1},$$

also erhalten wir nach Bilden des Infimums über alle  $x, y$ , dass  $d(C_{n+1}, B_n) \geq \frac{1}{n(n+1)}$ . Mit Induktion schließen wir weiter

$$\begin{aligned} \mu^*(B_{2k+1}) &\geq \mu^*(C_{2k} \cup B_{2k-1}) = \mu^*(C_{2k}) + \mu^*(B_{2k-1}) \\ &\geq \mu^*(C_{2k}) + \mu^*(C_{2k-2} \cup B_{2k-3}) \geq \dots \geq \sum_{j=1}^k \mu^*(C_{2j}), \end{aligned}$$

und analog  $\mu^*(B_{2k}) \geq \sum_{j=1}^k \mu^*(C_{2j-1})$ . Da  $\mu^*(B_n) \leq \mu^*(A) < \infty$ , folgt, dass die beiden Reihen  $\sum_{j=1}^{\infty} \mu^*(C_{2j})$  und  $\sum_{j=1}^{\infty} \mu^*(C_{2j-1})$  konvergieren. Aus der Subadditivität von  $\mu^*$  folgt nun

$$\mu^*(A \setminus F) \leq \mu^*(B_n) + \sum_{j=n+1}^{\infty} \mu^*(C_j).$$

Für  $n \rightarrow \infty$  verschwindet die letzte Summe und wir erhalten

$$\mu^*(A \setminus F) \leq \liminf \mu^*(B_n) \leq \limsup \mu^*(B_n) \leq \mu^*(A \setminus F),$$

also  $\mu^*(A \setminus F) = \lim \mu^*(B_n)$  wie gewünscht.  $\square$

**Bildmaße und der Transformationssatz** Seien  $(\Omega, \mathcal{F}), (\Omega', \mathcal{F}')$  zwei Maßräume,  $T : (\Omega, \mathcal{F}) \rightarrow (\Omega', \mathcal{F}')$  eine messbare Abbildung und  $\mu$  ein Maß auf  $(\Omega, \mathcal{F})$ . Dann definiert die Abbildung

$$A' \mapsto \mu(T^{-1}(A'))$$

ein Maß auf  $(\Omega', \mathcal{F}')$ . Das so entstandene Maß heißt das *Bildmaß* oder *push-forward* von  $\mu$  unter der Abbildung  $T$  und wird mit  $\mu_T$  bezeichnet.

**Beispiel 2.8.** Ist  $X : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$  eine reelle Zufallsvariable auf dem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, \mathbb{P})$ , so ist  $\mathbb{P}_X(A') = \mathbb{P}([X \in A'])$  für  $A \in \mathcal{B}_{\mathbb{R}}$  die Wahrscheinlichkeit, dass die Zufallsvariable  $X$  in der Menge  $A'$  liegt.  $\mathbb{P}_X$  heißt die Verteilung der Zufallsvariable  $X$ . Dabei ist  $[X \in A'] := X^{-1}(A')$ .

Für Bildmaße gilt der Transformationssatz:

**Satz 2.9.** Sei  $f : (\Omega', \mathcal{F}') \rightarrow (\overline{\mathbb{R}}, \overline{\mathcal{B}}_{\mathbb{R}})$  eine messbare Abbildung ( $\mathcal{B}_{\mathbb{R}}$  ist die Borelsche  $\sigma$ -Algebra auf den reellen Zahlen),  $T : (\Omega, \mathcal{F}) \rightarrow (\Omega', \mathcal{F}')$  eine messbare Abbildung und  $\mu$  ein Maß auf  $(\Omega, \mathcal{F})$ . Dann ist  $f$  genau dann bezüglich  $\mu_T$  integrierbar, falls  $f \circ T$  bezüglich  $\mu$  integrierbar ist und es gilt

$$\int f d\mu_T = \int f \circ T d\mu.$$

**Bemerkung 2.10.**  $\overline{\mathbb{R}}$  ist definiert als  $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$  und eine Teilmenge  $B \subset \overline{\mathbb{R}}$  ist in eine Borelmenge in  $\overline{\mathbb{R}}$  (also in  $\overline{\mathcal{B}}_{\mathbb{R}}$ ), falls  $B \cap \mathbb{R}$  eine Borelmenge in  $\mathbb{R}$  ist.

## Rieszscher Darstellungssatz

**Definition 2.11.** Ein lineares Funktional  $\Lambda : C(X) \rightarrow \mathbb{R}$  heißt positiv, falls für alle  $f \in C(X)$  mit  $f \geq 0$  gilt, dass auch  $\Lambda(f) \geq 0$ .

**Satz 2.12** (Rieszscher Darstellungssatz). Sei  $X$  ein kompakter metrischer Raum und  $\Lambda$  ein positives lineares Funktional auf  $C(X)$  mit  $\Lambda(1) = 1$ . Dann existiert ein eindeutiges Wahrscheinlichkeitsmaß  $\mu$  auf den Borelmengen  $\mathcal{B}_X$  von  $X$ , sodass

$$\Lambda(f) = \int_X f d\mu \quad \text{für alle } f \in C(X).$$

## Der Satz von Radon-Nikodým

**Definition 2.13.** Seien  $\mu, \nu$  zwei Maße auf einem Maßraum  $(\Omega, \mathcal{F})$ . Dann heißt  $\mu$  absolut stetig bzgl.  $\nu$  (i.Z.  $\mu \ll \nu$ ), falls für alle  $A \in \mathcal{F}$  gilt

$$\nu(A) = 0 \Rightarrow \mu(A) = 0.$$

**Definition 2.14.** Ein Maß  $\mu$  auf einem Maßraum  $(\Omega, \mathcal{F})$  heißt  $\sigma$ -endlich, falls es Mengen  $A_n \in \mathcal{F}$  gibt mit  $\mu(A_n) < \infty$  und  $\Omega = \bigcup_{n=1}^{\infty} A_n$ .

**Satz 2.15** (Radon-Nikodým). Seien  $\mu, \nu$  zwei Maße auf einem Maßraum  $(\Omega, \mathcal{F})$ . Ist  $\nu$   $\sigma$ -endlich, so sind äquivalent:



- $\mu$  besitzt eine Dichte  $f$  bezüglich  $\nu$  (das heißt, es gibt  $f \geq 0$  messbar auf  $\Omega$  mit  $\mu(A) = \int_A f d\nu$  für alle  $A \in \mathcal{F}$ ).
- $\mu \ll \nu$ .

Weiters ist  $f$   $\nu$ -fast überall eindeutig bestimmt.

Man vergleiche den Satz von Radon-Nikodým mit Satz 1.5.

### Der Erweiterungssatz von Kolmogorov

**Definition 2.16.** Sei  $(X_n, \mathcal{F}_n)$  eine Familie von Mengen  $X_n$  ausgestattet mit einer  $\sigma$ -Algebra  $\mathcal{F}_n$ . Für jede Indexmenge  $I \subset \mathbb{N}$  schreiben wir  $(X^I, \mathcal{F}^I) := \prod_{n \in I} (X_n, \mathcal{F}_n)$ . Für alle  $x \in X^{I_1}$  und alle  $I_2 \subset I_1 \subset \mathbb{N}$  schreiben wir  $\pi_{I_1 I_2}(x)$  für die Restriktion von  $x$  auf  $I_2$ . Eine Familie  $\{\mu_E : E \subset \mathbb{N} \text{ endlich}\}$  von Maßen heißt konsistent, falls

(i)  $\mu_E$  auf  $\mathcal{F}^E$  definiert ist für alle endlichen  $E \subset \mathbb{N}$  und

(ii)  $\mu_{E_2}(A) = \mu_{E_1}(\pi_{E_1 E_2}^{-1}(A))$  für alle  $A \in \mathcal{F}^{E_2}$  und alle endlichen  $E_2 \subset E_1 \subset \mathbb{N}$  gilt.

**Satz 2.17.** Sei  $(X_n, \mathcal{F}_n)$  eine Familie von separablen, vollständigen metrischen Räumen  $X_n$  ausgestattet mit einer  $\sigma$ -Algebra  $\mathcal{F}_n$ . Ist  $\{\mu_E : E \subset \mathbb{N} \text{ endlich}\}$  eine konsistente Familie von Maßen, dann existiert ein Maß  $\mu$  auf  $\mathcal{F}^{\mathbb{N}}$ , sodass  $\mu_E(A) = \mu(\pi_{\mathbb{N} E}^{-1}(A))$  für alle  $A \in \mathcal{F}^E$  für alle endlichen Mengen  $E \subset \mathbb{N}$ .

**Approximation von Mengen in einer  $\sigma$ -Algebra mit Mengen in einer erzeugenden Algebra** Im nächsten Lemma behandeln wir die maßtheoretische Approximation von Mengen einer  $\sigma$ -Algebra durch Mengen der erzeugenden Algebra. Im Unterschied zu einer  $\sigma$ -Algebra muss in einer Algebra  $\mathcal{F}_0$  nur die *endliche* Vereinigung von Mengen in  $\mathcal{F}_0$  wieder in  $\mathcal{F}_0$  enthalten sein. Dazu benötigen wir die symmetrische Differenz  $A \Delta B$  zweier Mengen  $A, B \subseteq \Omega$ , die definiert ist durch  $A \Delta B := (A \setminus B) \cup (B \setminus A)$ .

**Lemma 2.18.** Sei  $\mathcal{F}_0$  eine Algebra, die die  $\sigma$ -Algebra  $\mathcal{F}$  erzeugt, also für die gilt  $\mathcal{F} = \sigma(\mathcal{F}_0)$ . Sei weiters  $\mu$  ein Wahrscheinlichkeitsmaß auf  $\mathcal{F}$ . Dann gibt es für alle Mengen  $A \in \mathcal{F}$  und alle  $\varepsilon > 0$  eine Menge  $A_\varepsilon \in \mathcal{F}_0$  mit

$$\mu(A \Delta A_\varepsilon) < \varepsilon.$$

*Beweis.* Wir definieren

$$\mathcal{G} := \{A \in \mathcal{F} : \forall \varepsilon > 0 \exists A_\varepsilon \in \mathcal{F}_0 : \mu(A \Delta A_\varepsilon) < \varepsilon\}$$

und zeigen, dass  $\mathcal{G}$  eine  $\sigma$ -Algebra ist. Es gilt klarerweise  $\Omega \in \mathcal{G}$ , da  $\Omega \in \mathcal{F}_0$  ist und aus  $A \in \mathcal{G}$  folgt  $A^c \in \mathcal{G}$ , da

$$A \Delta A_\varepsilon = A^c \Delta A_\varepsilon^c.$$

Sei nun eine Folge von Mengen  $A_n$  in  $\mathcal{G}$  gegeben. Dann ist auch  $A = \bigcup_{k=1}^{\infty} A_k$  in  $\mathcal{G}$ . Um das zu zeigen, geben wir uns  $\varepsilon > 0$  vor und wählen  $n_* \in \mathbb{N}$  so, dass

$$\mu\left(A \Delta \bigcup_{k=1}^{n_*} A_k\right) = \mu\left(A \setminus \bigcup_{k=1}^{n_*} A_k\right) < \varepsilon/2.$$

Danach wählen wir für  $1 \leq k \leq n_*$  Mengen  $A_{k,\varepsilon} \in \mathcal{F}_0$  mit

$$\mu(A_k \Delta A_{k,\varepsilon}) < \frac{\varepsilon}{2n_*}.$$

Dies ist möglich, da  $A_k \in \mathcal{G}$  ist. Da

$$\left( \bigcup_{k=1}^{n_*} A_k \right) \Delta \left( \bigcup_{k=1}^{n_*} A_{k,\varepsilon} \right) \subset \bigcup_{k=1}^{n_*} A_k \Delta A_{k,\varepsilon},$$

erhalten wir

$$\mu \left( \left( \bigcup_{k=1}^{n_*} A_k \right) \Delta \left( \bigcup_{k=1}^{n_*} A_{k,\varepsilon} \right) \right) \leq \sum_{k=1}^{n_*} \mu(A_k \Delta A_{k,\varepsilon}) < \frac{\varepsilon}{2}.$$

Deshalb folgt aus der Relation  $C \Delta E \subset (C \Delta D) \cup (D \Delta E)$  für beliebige Mengen  $C, D, E$  die Abschätzung

$$\mu \left( A \Delta \left( \bigcup_{k=1}^{n_*} A_{k,\varepsilon} \right) \right) \leq \mu \left( A \Delta \left( \bigcup_{k=1}^{n_*} A_k \right) \right) + \mu \left( \left( \bigcup_{k=1}^{n_*} A_k \right) \Delta \left( \bigcup_{k=1}^{n_*} A_{k,\varepsilon} \right) \right) < \varepsilon,$$

und somit nach Definition  $A \in \mathcal{G}$ . Zusammenfassend ist  $\mathcal{G}$  also eine  $\sigma$ -Algebra, die (trivialerweise)  $\mathcal{F}_0$  enthält, daher gilt  $\mathcal{G} = \mathcal{F}$ .  $\square$

**Beispiel 2.19.** Sei  $\Omega = [0, 1]$  und  $\lambda$  das Lebesgue-Maß auf der Borelschen  $\sigma$ -Algebra  $\mathcal{F} = \mathcal{B}_\Omega$ . Dann wird  $\mathcal{F}$  von der Algebra der endlichen Vereinigungen von dyadischen Intervallen  $(m2^{-k}, (m+1)2^{-k})$  mit  $k \in \mathbb{N}, 0 \leq m \leq 2^k - 1$  erzeugt.

## 2.2 Hausdorffmaß

Sei  $U$  eine offene Teilmenge von  $\mathbb{R}^n$ . Der Durchmesser  $|U|$  von  $U$  ist definiert als

$$|U| := \text{diam } U := \sup\{|x - y| : x, y \in U\}.$$

Falls  $\{U_i\}$  eine abzählbare (oder endliche) Familie von Mengen mit Durchmesser  $\leq \delta$  ist mit  $F \subset \bigcup_{i=1}^{\infty} U_i$ , dann heißt  $\{U_i\}$  eine  $\delta$ -Überdeckung von  $F$ .

**Definition 2.20.** Seien  $F \subset \mathbb{R}^n$  und  $s \geq 0$ . Für  $\delta > 0$  definieren wir

$$\mathcal{H}_\delta^s(F) := \inf \left\{ \sum_{i=1}^{\infty} |U_i|^s : \{U_i\} \text{ ist eine } \delta\text{-Überdeckung von } F \right\}.$$

Falls  $\delta$  abnimmt, wird die Klasse der Überdeckungen von  $F$  kleiner, also steigt  $\mathcal{H}_\delta^s(F)$  für  $\delta \searrow 0$ .

**Definition 2.21.** Seien  $F \subset \mathbb{R}^n$  und  $s \geq 0$ . Das  $s$ -dimensionale Hausdorffmaß  $\mathcal{H}^s$  von  $F$  ist definiert als

$$\mathcal{H}^s(F) := \sup_{\delta > 0} \mathcal{H}_\delta^s(F) = \lim_{\delta \searrow 0} \mathcal{H}_\delta^s(F).$$

Dieser Limes existiert für alle Teilmengen  $F$  von  $\mathbb{R}^n$ , kann jedoch auch den Wert  $\infty$  annehmen.

**Bemerkung 2.22.** Die Intuition hinter dieser Definition von  $\mathcal{H}^s$  ist folgende: Für  $s \in \mathbb{N}$  und eine „ $s$ -dimensionale“ Teilmenge von  $\mathbb{R}^n$  (wie etwa eine relativ offene Teilmenge eines  $s$ -dimensionalen Unterraums von  $\mathbb{R}^n$ ) ist der Anteil von  $A$ , der in einer Region mit Durchmesser  $r$  enthalten ist, proportional zu  $r^s$ .

Wir werden nun sehen, dass das oben definierte  $s$ -dimensionale Hausdorffmaß seinem Namen gerecht wird, das heißt es gilt

**Proposition 2.23.**  $\mathcal{H}^s$  ist ein metrisches äußeres Maß.

*Beweis.* Es ist eine gute Übungsaufgabe, zu zeigen, dass  $\mathcal{H}_\delta^s$  ein äußeres Maß ist für alle  $\delta > 0$ . Daraus folgt dann, dass auch  $\mathcal{H}^s$  ein äußeres Maß ist. Wir wollen nun zeigen, dass für  $d(A, B) > 0$  gilt

$$\mathcal{H}^s(A \cup B) = \mathcal{H}^s(A) + \mathcal{H}^s(B).$$

Sei dazu  $\{C_j\}$  eine Überdeckung von  $A \cup B$  mit  $|C_j| < \delta < d(A, B)$  für alle  $j$ , dann schneidet kein  $C_j$   $A$  und  $B$ , es folgt also

$$\sum_{j=1}^{\infty} |C_j|^s \geq \mathcal{H}_\delta^s(A) + \mathcal{H}_\delta^s(B),$$

nach Übergang zum Infimum über alle Überdeckungen von  $A \cup B$  somit

$$\mathcal{H}_\delta^s(A \cup B) \geq \mathcal{H}_\delta^s(A) + \mathcal{H}_\delta^s(B).$$

Für  $\delta \rightarrow 0$  erhalten wir die gewünschte Ungleichung

$$\mathcal{H}^s(A \cup B) \geq \mathcal{H}^s(A) + \mathcal{H}^s(B),$$

da die umgekehrte Ungleichung aus der Subadditivität von  $\mathcal{H}^s$  folgt.  $\square$

Es folgt also aus dieser Proposition, Satz 2.6 und Proposition 2.7 der für uns wichtige

**Satz 2.24.**  $\mathcal{H}^s$ , eingeschränkt auf die Borelmengen  $\mathcal{B}_{\mathbb{R}^n}$  von  $\mathbb{R}^n$  ist ein Maß (das  $s$ -dimensionale Hausdorffmaß).

**Bemerkung 2.25.** Das Volumen der Einheitskugel im  $\mathbb{R}^n$  ist

$$\omega_n := \frac{\pi^{n/2}}{\Gamma(n/2 + 1)},$$

wobei  $\Gamma(x) := \int_0^\infty e^{-t} t^{x-1} dt$  die Gammafunktion bezeichnet. Man kann zeigen, dass auf  $\mathbb{R}^n$  das Maß  $\omega_n 2^{-n} \mathcal{H}^n$  gleich dem  $n$ -dimensionalen Lebesguemaß  $\lambda^n$  ist. Beachte, dass  $\omega_n 2^{-n}$  das Volumen der Kugel im  $\mathbb{R}^n$  mit Durchmesser 1 ist.

Das Lebesguemaß  $\lambda^n$  auf  $\mathbb{R}^n$  ist homogen vom Grad  $n$ , das heißt für  $A$  messbar und  $t > 0$  gilt

$$\lambda^n(tA) = t^n \lambda^n(A).$$

Man wird ein ähnliches Verhalten auch von  $\mathcal{H}^s$  erwarten, tatsächlich gilt

**Proposition 2.26.** Seien  $F \subset \mathbb{R}^n$  und  $f, g : F \rightarrow \mathbb{R}^n$  zwei Abbildungen mit

$$|f(x) - f(y)| \leq c|g(x) - g(y)|^\alpha \quad \text{für alle } x, y \in F.$$

Dann gilt

$$\mathcal{H}^{s/\alpha}(f(F)) \leq c^{s/\alpha} \mathcal{H}^s(g(F)). \quad (2.1)$$

*Beweis.* Sei  $\{B_j\}$  eine  $\delta$ -Überdeckung von  $g(F)$ . Dann ist  $\{U_j\} = \{f(g^{-1}(B_j))\}$  eine  $c\delta^\alpha$ -Überdeckung von  $f(F)$ , denn für  $x, y \in g^{-1}(B_j)$  gilt

$$|f(x) - f(y)| \leq c|g(x) - g(y)|^\alpha \leq c\delta^\alpha.$$

Es folgt also

$$\mathcal{H}_{c\delta^\alpha}^{s/\alpha}(f(F)) \leq \sum_j |U_j|^{s/\alpha} \leq \sum_j (c|B_j|^\alpha)^{s/\alpha} = c^{s/\alpha} \sum_j |B_j|^s.$$

Nach Übergang zum Infimum über alle Überdeckungen von  $g(F)$  folgt somit  $\mathcal{H}_{c\delta^\alpha}^{s/\alpha}(f(F)) \leq c^{s/\alpha} \mathcal{H}_\delta^s(g(F))$ . Lassen wir  $\delta$  gegen 0 gehen, so folgt die gewünschte Ungleichung

$$\mathcal{H}^{s/\alpha}(f(F)) \leq c^{s/\alpha} \mathcal{H}^s(g(F)). \quad \square$$

**Bemerkung 2.27.** Es ist instruktiv, sich ein paar Spezialfälle dieser Formel anzusehen. Ist  $f(x) = tx$  bzw.  $f(x) = x/t$  und  $g(x) = x$ , so folgt insgesamt aus (2.1)

$$\mathcal{H}^s(tF) = t^s \mathcal{H}^s(F), \quad t > 0,$$

also ist  $\mathcal{H}^s$  homogen vom Grad  $s$ . Ist  $f$  eine Lipschitz-Abbildung mit Lipschitz-Konstante  $c$  und  $g$  die Identität, so wird (2.1) zu

$$\mathcal{H}^s(f(F)) \leq c^s \mathcal{H}^s(F).$$

Ist  $f$  eine Isometrie (d.h.  $|f(x) - f(y)| = |x - y|$  für alle  $x, y \in F$ ), dann gilt sogar

$$\mathcal{H}^s(f(F)) = \mathcal{H}^s(F).$$

Insbesondere ist also das Hausdorffmaß invariant unter Rotationen und Translationen.

Unser weiteres Ziel in diesem Abschnitt ist der Beweis einer Verallgemeinerung der Transformationsformel für Integrale in  $\mathbb{R}^n$  auf Untermannigfaltigkeiten von  $\mathbb{R}^n$ . Dazu definieren wir

**Definition 2.28.** Sei  $1 \leq k \leq n$ . Eine  $k$ -dimensionale  $C^1$ -Untermannigfaltigkeit ist eine Menge  $M \subset \mathbb{R}^n$  mit der Eigenschaft: Für alle  $x \in M$  existiert eine Umgebung  $U$  von  $x$  in  $\mathbb{R}^n$ , eine offene Menge  $V \subset \mathbb{R}^k$  und eine injektive  $C^1$ -Abbildung  $f : V \rightarrow U$  mit  $f(V) = M \cap U$  und injektiver Jacobimatrix überall in  $V$ .  $f$  heißt dann Parametrisierung von  $M$ .

Dabei bedeutet die letzte Forderung an  $f$ , dass für alle  $x \in V$  die lineare Abbildung, induziert durch die Jacobimatrix  $D_x f$ , injektiv ist. Jede Untermannigfaltigkeit  $M$  von  $\mathbb{R}^n$  kann mit abzählbar vielen Parametrisierungen überdeckt werden. Wir nehmen bis auf weiteres an, dass  $M = M \cap U$  eine globale Parametrisierung besitzt.

**Beispiel 2.29.** Sei  $k = 1$  und  $n = 2$  und sei durch die injektive Funktion  $f$  eine Kurve parametrisiert. Die Forderung, dass die Jacobimatrix von  $f$  (in diesem Fall ein zweidimensionaler Vektor) injektiv ist bedeutet, dass  $f'(x)$  für alle  $x \in V$  als lineare Abbildung von  $\mathbb{R}$  nach  $\mathbb{R}^2$  injektiv ist und das ist genau dann der Fall, wenn  $|f'(x)| > 0$  für alle  $x \in V$ .

Sei nun  $T$  eine lineare Abbildung von  $\mathbb{R}^k$  nach  $\mathbb{R}^n$ ;  $T^* : \mathbb{R}^n \rightarrow \mathbb{R}^k$  ist ihre Transponierte. Dann ist klarerweise  $T^*T : \mathbb{R}^k \rightarrow \mathbb{R}^k$  positiv semidefinit, man kann also definieren

$$J(T) := \sqrt{\det(T^*T)}.$$

**Proposition 2.30.** Für  $k \leq n$ ,  $A \in \mathcal{B}_{\mathbb{R}^k}$  und eine lineare Abbildung  $T : \mathbb{R}^k \rightarrow \mathbb{R}^n$  gilt

$$\mathcal{H}^k(T(A)) = J(T)\mathcal{H}^k(A). \quad (2.2)$$

*Beweis.* Für  $k = n$  ist  $J(T) = |\det T|$  und die Aussage reduziert sich auf den bekannten Fall für das Lebesguemaß:

$$\lambda^n(T(A)) = |\det T| \cdot \lambda^n(A).$$

Für  $k < n$  sei  $R$  eine Rotation, die das Bild von  $T$  auf den Unterraum  $\mathbb{R}^k \times \{0\}$  in  $\mathbb{R}^n$  abbildet und sei  $S := RT$ . Dann gilt  $S^*S = T^*R^*RT = T^*T$  und somit  $J(S) = J(T)$ . Da  $\mathcal{H}^k$  rotationsinvariant ist (siehe Bemerkung 2.27), gilt  $\mathcal{H}^k(S(A)) = \mathcal{H}^k(T(A))$ . Sei weiters  $P$  die Projektion von  $\mathbb{R}^k \times \{0\}$  auf  $\mathbb{R}^k$ . Dann ist  $P$  eine Isometrie und  $\mathcal{H}^k$  ist invariant bezüglich  $P$ . Es folgt also, wenn wir (2.2) wieder für quadratische Matrizen anwenden:

$$\begin{aligned} \mathcal{H}^k(T(A)) &= \mathcal{H}^k(S(A)) = \mathcal{H}^k(PS(A)) = J(PS)\mathcal{H}^k(A) \\ &= \sqrt{\det(S^*S)}\mathcal{H}^k(A) = \sqrt{\det(T^*T)}\mathcal{H}^k(A). \end{aligned} \quad \square$$

Um diese Formel auf allgemeine Funktionen  $T$  erweitern zu können, benötigen wir noch ein Lemma.

**Lemma 2.31.** Sei  $M$  eine  $k$ -dimensionale  $C^1$ -Untermannigfaltigkeit von  $\mathbb{R}^n$  parametrisiert mit  $f : V \rightarrow \mathbb{R}^n$ . Für alle  $\alpha > 1$  existiert eine Folge  $\{B_j\}$  von disjunkten Borel-Teilmengen von  $V$  sodass  $V = \bigcup_j B_j$  und eine Folge von linearen Abbildungen  $T_j : \mathbb{R}^k \rightarrow \mathbb{R}^n$ , sodass

$$\alpha^{-1}|T_j z| \leq |(D_x f)z| \leq \alpha|T_j z|, \quad x \in B_j, z \in \mathbb{R}^k \quad (2.3)$$

und

$$\alpha^{-1}|T_j x - T_j y| \leq |f(x) - f(y)| \leq \alpha|T_j x - T_j y|, \quad x, y \in B_j. \quad (2.4)$$

*Beweis.* Seien  $\varepsilon > 0$  und  $\beta > 1$  mit

$$\alpha^{-1} + \varepsilon < \beta^{-1} < 1 < \beta < \alpha - \varepsilon$$

und sei  $\mathcal{T}$  eine abzählbare dichte Teilmenge der Menge der linearen Abbildungen von  $\mathbb{R}^k$  nach  $\mathbb{R}^n$  (zB. die Menge der  $n \times k$  Matrizen mit rationalen Einträgen). Für  $T \in \mathcal{T}$  und  $m \in \mathbb{N}$  sei  $E(T, m)$  die Menge aller  $x \in V$  sodass

$$\beta^{-1}|Tz| \leq |(D_x f)z| \leq \beta|Tz| \quad \text{für alle } z \in \mathbb{R}^k \quad \text{und} \quad (2.5)$$

$$\alpha^{-1}|Tx - Ty| \leq |f(x) - f(y)| \leq \alpha|Tx - Ty| \quad \text{für alle } y \in V \text{ mit } |y - x| < m^{-1}. \quad (2.6)$$

Diese Definition von  $E(T, m)$  bleibt unverändert, falls  $y, z$  in abzählbaren dichten Teilmengen von  $V$  bzw.  $\mathbb{R}^k$  liegen. Da dann  $E(T, m)$  definiert ist über abzählbar viele Ungleichungen mit stetigen Funktionen, ist  $E(T, m)$  eine Borelmenge. Es reicht also nun zu zeigen, dass die Mengen  $E(T, m)$   $V$  überdecken, da jedes  $E(T, m)$  eine abzählbare Vereinigung von Mengen  $E_i(T, m)$  mit Durchmesser  $< m^{-1}$  ist und eine Disjunktifizierung der abzählbaren Familie

$$\{E_i(T, m) : T \in \mathcal{T}, i, m \in \mathbb{N}\}$$

die gewünschten Folgen  $T_j, B_j$  liefert.

Sei also  $x \in V$ . Wir müssen zeigen, dass es eine lineare Abbildung  $T$  und eine natürliche Zahl  $m$  gibt mit  $x \in E(T, m)$ . Dazu definieren wir  $\delta_0 := \inf\{|(D_x f)z| : |z| = 1\}$ . Da  $D_x f$  injektiv ist (nach Definition einer Parametrisierung) und  $S^1$  kompakt, ist  $\delta_0 > 0$ . Wähle  $\delta > 0$  mit  $\delta \leq (\beta - 1)\delta_0$  und  $\delta \leq (1 - \beta^{-1})\delta_0$ . Wähle weiters  $T \in \mathcal{T}$  mit

$$\|T - D_x f\| < \delta.$$

Dann gilt

$$|Tz| \leq |(D_x f)z| + |(D_x f)z - Tz| \leq |(D_x f)z| + \delta|z| \leq \beta|(D_x f)z|.$$

In analoger Weise erhält man  $|Tz| \geq \beta^{-1}|(D_x f)z|$ , also die erste Ungleichung in (2.5). Zusätzlich zeigt diese Ungleichung, dass  $T$  injektiv ist. Daher gilt

$$\eta := \inf\{|Tz| : |z| = 1\} > 0.$$

Nun nutzen wir die Differenzierbarkeit von  $f$  bei  $x$  aus und erhalten ein  $m \in \mathbb{N}$  mit

$$|f(y) - f(x) - (D_x f)(y - x)| \leq \varepsilon\eta|y - x| \leq \varepsilon|T(y - x)| \quad \text{für alle } |x - y| \leq m^{-1}.$$

Dann gilt aber

$$\begin{aligned} |f(y) - f(x)| &\leq |f(y) - f(x) - (D_x f)(y - x)| + |D_x f(y - x)| \\ &\leq \varepsilon|Ty - Tx| + \beta|Ty - Tx| < \alpha|Ty - Tx| \end{aligned}$$

für alle  $|x - y| \leq m^{-1}$ . In ähnlicher Weise folgt nun auch die Ungleichung in die andere Richtung  $|f(y) - f(x)| > \alpha^{-1}|Ty - Tx|$ . Das bedeutet aber  $x \in E(T, m)$ , der Beweis ist also abgeschlossen.  $\square$

**Satz 2.32.** *Sei  $M$  eine  $k$ -dimensionale  $C^1$ -Untermannigfaltigkeit von  $\mathbb{R}^n$  parametrisiert durch  $f : V \rightarrow \mathbb{R}^n$ . Für eine Borelmenge  $A \subset V$  ist  $f(A)$  eine Borelmenge in  $\mathbb{R}^n$  und*

$$\mathcal{H}^k(f(A)) = \int_A J(D_x f) d\mathcal{H}^k(x). \quad (2.7)$$

Weiters gilt für eine Borel-messbare Funktion  $\phi$  auf  $M$ , die entweder nichtnegativ oder in  $L^1(M, \mathcal{H}^k)$  ist:

$$\int_M \phi(y) d\mathcal{H}^k(y) = \int_V \phi(f(x)) J(D_x f) d\mathcal{H}^k(x). \quad (2.8)$$

*Beweis.* Wir zeigen zuerst, dass  $f(A)$  messbar ist: Da  $V$  offen ist in  $\mathbb{R}^k$ , ist  $V$  eine abzählbare Vereinigung von kompakten Teilmengen. Eine abgeschlossene Teilmenge  $A$  von  $V$  ist somit auch eine abzählbare Vereinigung von kompakten Teilmengen. Da  $f$  stetig ist, ist auch  $f(A)$  eine abzählbare Vereinigung von kompakten Mengen.

Da  $f$  injektiv ist, ist die Familie  $\mathcal{F}$  der Mengen  $A \subset V$ , sodass  $f(A)$  messbar ist, eine  $\sigma$ -Algebra, die die abgeschlossenen Mengen enthält (wie gerade gezeigt wurde), also besteht  $\mathcal{F}$  aus allen Borelteilmengen von  $V$ .

Wir zeigen nur (2.7), da (2.8) aus (2.7) mit den üblichen Linearitäts- und Approximationssargumenten folgt. Sei nun  $\alpha > 1$ , Weiters wähle Folgen  $\{B_j\}, \{T_j\}$  wie im letzten Lemma. Definiere dann  $A_j := A \cap B_j$ . Dann folgt aus (2.3) und Proposition 2.26

$$\alpha^{-k} \mathcal{H}^k(T_j(E)) \leq \mathcal{H}^k((D_x f)(E)) \leq \alpha^k \mathcal{H}^k(T_j(E)),$$

für alle  $x \in A_j$  und alle Borelmengen  $E \subset \mathbb{R}^k$ . Daraus erhalten wir mit Proposition 2.30

$$\alpha^{-k} J(T_j) \leq J(D_x f) \leq \alpha^k J(T_j) \quad \text{für alle } x \in A_j.$$

Es folgt aber auch aus (2.4) und Proposition 2.26:

$$\alpha^{-k} \mathcal{H}^k(T_j(A_j)) \leq \mathcal{H}^k(f(A_j)) \leq \alpha^k \mathcal{H}^k(T_j(A_j)).$$

Nun erhalten wir aber aus diesen Ungleichungen und Proposition 2.30

$$\begin{aligned} \alpha^{-2k} \mathcal{H}^k(f(A_j)) &\leq \alpha^{-k} J(T_j) \mathcal{H}^k(A_j) \leq \int_{A_j} J(D_x f) d\mathcal{H}^k(x) \\ &\leq \alpha^k J(T_j) \mathcal{H}^k(A_j) \leq \alpha^{2k} \mathcal{H}^k(f(A_j)). \end{aligned}$$

Nun summieren wir über alle  $j$  und bekommen schließlich

$$\alpha^{-2k} \mathcal{H}^k(f(A)) \leq \int_A J(D_x f) d\mathcal{H}^k(x) \leq \alpha^{2k} \mathcal{H}^k(f(A))$$

für alle  $\alpha > 1$ . Die Behauptung folgt nun mit  $\alpha \rightarrow 1$ . □

**Beispiel 2.33** (Hausdorffmaß einer Fläche in  $\mathbb{R}^3$ ). Seien  $k = 2$ ,  $n = 3$  und  $f : V \rightarrow \mathbb{R}^3$  die Parametrisierung einer Fläche. Dann folgt aus Satz 2.32

$$\mathcal{H}^2(f(A)) = c_n \int_A |f_{x_1} \times f_{x_2}| dx_1 dx_2$$

mit einer Konstante  $c_n$ , die nur von  $n$  abhängt (vgl. Bemerkung 2.25), denn

$$\begin{aligned} J(D_x f)^2 &= \det((D_x f)^T (D_x f)) = \det \begin{pmatrix} |f_{x_1}|^2 & \langle f_{x_1}, f_{x_2} \rangle \\ \langle f_{x_1}, f_{x_2} \rangle & |f_{x_2}|^2 \end{pmatrix} \\ &= |f_{x_1}|^2 |f_{x_2}|^2 - \langle f_{x_1}, f_{x_2} \rangle^2 = |f_{x_1} \times f_{x_2}|^2. \end{aligned}$$

### 2.3 Hausdorffdimension

Aus Definition 2.20 folgt, dass für eine beliebige Menge  $A$  und für alle  $\delta < 1$  die Funktion  $s \mapsto \mathcal{H}_\delta^s(A)$  monoton fallend ist, also ist auch  $s \mapsto \mathcal{H}^s(A)$  monoton fallend. Es gilt sogar mehr:

**Proposition 2.34.** *Für alle  $s \geq 0$  und  $q > s$  gelten*

$$\begin{aligned}\mathcal{H}^s(A) < \infty &\Rightarrow \mathcal{H}^q(A) = 0, \\ \mathcal{H}^q(A) > 0 &\Rightarrow \mathcal{H}^s(A) = \infty.\end{aligned}$$

*Beweis.* Da die zweite Folgerung die Kontraposition der ersten ist, zeigen wir nur die erste. Sei also  $\mathcal{H}^s(A) < \infty$  und wähle für  $\delta > 0$  eine  $\delta$ -Überdeckung  $\{B_j\}$  von  $A$ , sodass

$$\sum_{j=1}^{\infty} |B_j|^s \leq \mathcal{H}^s(A) + 1.$$

Dann gilt für  $q > s$

$$\sum_{j=1}^{\infty} |B_j|^q = \sum_{j=1}^{\infty} |B_j|^s |B_j|^{q-s} \leq \delta^{q-s} (\mathcal{H}^s(A) + 1).$$

Es gilt also  $\mathcal{H}_\delta^q(A) \leq \delta^{q-s} (\mathcal{H}^s(A) + 1)$  für alle  $\delta > 0$ . Lassen wir  $\delta$  gegen 0 gehen und beachten die Endlichkeit von  $\mathcal{H}^s(A)$ , erhalten wir die Behauptung.  $\square$

**Definition 2.35** (Hausdorffdimension). *Sei  $A \subset \mathbb{R}^n$ . Die Hausdorffdimension  $\dim A$  von  $A$  ist definiert als*

$$\dim A := \inf\{s \geq 0 : \mathcal{H}^s(A) = 0\} = \sup\{s \geq 0 : \mathcal{H}^s(A) = \infty\}.$$

Nach Proposition 2.34 sind beide Ausdrücke gleich. Es gilt also

$$\mathcal{H}^s(A) = \begin{cases} \infty & s < \dim A, \\ 0 & s > \dim A. \end{cases}$$

Für  $s = \dim A$  kann  $\mathcal{H}^s(A) = 0$  oder  $\mathcal{H}^s(A) = \infty$  gelten. Falls aber

$$0 < \mathcal{H}^s(A) < \infty,$$

so heißt  $A$  eine  $s$ -Menge.

**Bemerkung 2.36.** Manchmal ist es wünschenswert, eine schärfere Unterteilung der Dimension oben zu haben als die Hausdorffdimension. Um das zu erreichen, geben wir uns eine stetige, steigende Funktion  $h : [0, \infty] \rightarrow [0, \infty)$  vor, die wir *Dimensionsfunktion* nennen. Analog zu Definition 2.20 definieren wir

$$\mathcal{H}_\delta^h(F) := \inf \left\{ \sum_{i=1}^{\infty} h(|U_i|) : \{U_i\} \text{ ist eine } \delta\text{-Überdeckung von } F \right\}$$

für eine Teilmenge  $F \subset \mathbb{R}^n$ . Dies führt zu einem Maß  $\mathcal{H}^h(F) = \lim_{\delta \rightarrow 0} \mathcal{H}_\delta^h(F)$ . Dem klassischen Hausdorffmaß  $\mathcal{H}^s$  entspricht die Wahl  $h(t) = t^s$ .



## 2.4 Einige Hilfsmittel zur Berechnung der Hausdorffdimension

Wir beginnen mit einer einfachen unteren Schranke der Hausdorffdimension einer Menge  $F \subset \mathbb{R}^n$ .

**Proposition 2.37.** *Sei  $\mu$  ein äußeres Maß auf  $F$  mit  $\mu(F) = 1$ . Weiters seien  $s, c, \delta$  so, dass*

$$\mu(U) \leq c|U|^s \quad \text{für alle } U \subset F \text{ mit } |U| \leq \delta.$$

Dann gelten

$$\mathcal{H}^s(F) \geq c^{-1} \quad \text{und} \quad \dim F \geq s.$$

*Beweis.* Sei  $\{U_i\}$  eine  $\delta$ -Überdeckung von  $F$ . Dann gilt

$$0 < 1 = \mu(F) \leq \mu\left(\bigcup_i U_i\right) \leq \sum_i \mu(U_i) \leq c \sum_i |U_i|^s.$$

Es folgt also nach Übergang zum Infimum  $\mathcal{H}_\eta^s(F) \geq c^{-1}$  für  $\eta \leq \delta$ . Somit folgt die Behauptung  $\mathcal{H}^s(F) \geq c^{-1}$ .  $\square$

Für die nächste Abschätzung der Hausdorffdimension benötigen wir ein Überdeckungslemma, das unabhängig von der folgenden Anwendung auch von eigenständigem Interesse ist.

**Lemma 2.38.** *Sei  $\mathcal{U}$  eine Familie von Kugeln  $B$  in  $\mathbb{R}^n$ , die in einer beschränkten Teilmenge von  $\mathbb{R}^n$  enthalten sind. Dann gibt es eine (endliche oder abzählbare) disjunkte Teilfamilie  $\{B_j\}$  von  $\mathcal{U}$ , sodass*

$$\bigcup_{B \in \mathcal{U}} B \subset \bigcup_{j=1}^{\infty} 5B_j,$$

wobei  $5B_j$  die Kugel mit gleichem Mittelpunkt und 5fachem Radius wie  $B_j$  ist.

*Beweis. Schritt 1, Wählen der disjunkten Kugeln:* Sei  $r_1$  der Radius einer Kugel  $B_1 \in \mathcal{U}$  mit

$$r_1 > \frac{1}{2} \sup_{B \in \mathcal{U}} r(B), \quad r(B) := \text{Radius von } B.$$

Nun gehen wir induktiv vor und setzen voraus, dass die Kugeln  $B_1, \dots, B_k$  mit zugehörigen Radien  $r_1, \dots, r_k$  bereits gewählt sind. Dann wählen wir  $B_{k+1}$  mit Radius  $r_{k+1}$  so, dass  $B_{k+1}$  disjunkt ist zu allen Kugeln  $B_j$  ( $1 \leq j \leq k$ ) und

$$r_{k+1} > \frac{1}{2} \sup_{B \in \mathcal{U}} \{r(B) : B \text{ disjunkt zu } B_1, \dots, B_k\}.$$

Diese Folge kann abbrechen. Ist dies der Fall, so hat diese endliche Folge die gewünschten Eigenschaften (Übung). Wir betrachten im folgenden nur mehr den abzählbar unendlichen Fall.

**Schritt 2, diese Folge besitzt die gewünschten Eigenschaften:** Da alle Kugeln in  $\mathcal{U}$  in einer beschränkten Teilmenge von  $\mathbb{R}^n$  enthalten sind und die Kugeln  $B_j$  disjunkt sind, konvergiert die Folge  $r_j$  der Radien von  $B_j$  gegen 0. Wir müssen nun zeigen, dass für alle  $B \in \mathcal{U}$  gilt, dass  $B \subset \bigcup_{j=1}^{\infty} 5B_j$ . Sei also  $B \in \mathcal{U}$  mit Radius  $r$  gegeben.

Sei  $k$  die kleinste Zahl mit  $r_{k+1} < r/2$ .  $B$  schneidet nun eine der Kugeln  $B_1, \dots, B_k$ , da sonst im Iterationsprozess  $B$  statt  $B_{k+1}$  genommen hätte werden sollen. Sei also  $1 \leq k_0 \leq k$  so, dass  $B \cap B_{k_0} \neq \emptyset$ . Nach Konstruktion der  $B_j$  gilt  $r_{k_0} \geq r/2$ . Es ist nun eine leichte Folgerung der Dreiecksungleichung (Übung), dass  $B \subset 5B_{k_0}$ .  $\square$

**Satz 2.39.** *Sei  $\mu$  ein Wahrscheinlichkeitsmaß auf  $\mathbb{R}^n$  und  $F \in \mathcal{B}_{\mathbb{R}^n}$  eine Borelmenge in  $\mathbb{R}^n$ . Sei weiters  $0 < c < \infty$  konstant. Dann gelten:*

(i) Falls  $\limsup_{r \rightarrow 0} \frac{\mu(B(x, r))}{r^s} < c$  für alle  $x \in F$ , dann folgt  $\mathcal{H}^s(F) \geq \mu(F)/c$ .

(ii) Falls  $\limsup_{r \rightarrow 0} \frac{\mu(B(x, r))}{r^s} > c$  für alle  $x \in F$ , dann folgt  $\mathcal{H}^s(F) \leq 10^s \mu(\mathbb{R}^n)/c$ .

*Beweis.* (i) Sei  $\delta > 0$  und definiere

$$F_\delta := \{x \in F : \mu(B(x, r)) < (c - \varepsilon)r^s \text{ für alle } 0 < r \leq \delta \text{ für ein } \varepsilon > 0\}.$$

Sei dann  $\{U_i\}$  eine  $\delta$ -Überdeckung von  $F$  und daher eine von  $F_\delta$ . Für alle  $U_i$ , die einen Punkt  $x$  von  $F_\delta$  enthalten, gilt klarerweise, dass die Kugel  $B(x, |U_i|)$  die Menge  $U_i$  enthält. Aus der Definition von  $F_\delta$  folgt nun für solche  $U_i$

$$\mu(U_i) \leq \mu(B(x, |U_i|)) < c|U_i|^s,$$

sodass wir also erhalten

$$\mu(F_\delta) \leq \sum_{U_i \cap F_\delta \neq \emptyset} \mu(U_i) \leq c \sum_i |U_i|^s.$$

Nach Übergang zum Infimum über alle  $\delta$ -Überdeckungen wird das zu  $\mu(F_\delta) \leq c\mathcal{H}_\delta^s(F) \leq c\mathcal{H}^s(F)$ . Da  $F_\delta \uparrow F$  für  $\delta \downarrow 0$ , ist das die Behauptung.

(ii) Sei  $F$  beschränkt. Für fixes  $\delta > 0$  geben wir uns die Menge von Kugeln

$$\mathcal{U} = \{B(x, r) : x \in F, 0 < r \leq \delta, \mu(B(x, r)) > cr^s\}$$

vor. Die Voraussetzung liefert, dass  $\mathcal{U}$  eine  $2\delta$ -Überdeckung von  $F$  ist. Mit dem Überdeckungslemma 2.38 bekommen wir eine Folge  $\{B_j\}_j$  von disjunkten Kugeln, sodass

$$\bigcup_{B \in \mathcal{U}} B \subset \bigcup_{j=1}^{\infty} 5B_j.$$

Das heißt,  $\{5B_j\}_j$  ist eine  $10\delta$ -Überdeckung von  $F$ . Wir folgern weiter aus der Definition von  $\mathcal{U}$

$$\mathcal{H}_{10\delta}^s(F) \leq \sum_{j=1}^{\infty} |5B_j|^s = 5^s \sum_{j=1}^{\infty} |B_j|^s \leq 10^s \sum_{j=1}^{\infty} \frac{\mu(B_j)}{c} \leq 10^s c^{-1} \mu(\mathbb{R}^n),$$

da die Kugeln  $B_j$  disjunkt sind. Da diese Ungleichung für alle  $\delta > 0$  und alle dazugehörigen Überdeckungen  $\mathcal{U}$  gilt, erhalten wir die Behauptung des Satzes  $\mathcal{H}^s(F) \leq 10^s c^{-1} \mu(\mathbb{R}^n)$  für beschränkte  $F$ . Ist  $F$  unbeschränkt, so folgt die Behauptung durch Grenzübergang.  $\square$

**Bemerkung 2.40.** Durch eine bessere Version des Überdeckungslemmas (den Überdeckungssatz von Vitali) erhält man im zweiten Punkt des vorherigen Satzes den besseren Vorfaktor  $2^s$  anstatt  $10^s$ .

Wir schließen nun diesen Abschnitt mit einer potentialtheoretischen Methode, die Hausdorffdimension zu berechnen, ab. Dazu definieren wir für  $s \geq 0$  das  $s$ -Potential des endlichen Maßes  $\mu$  mit  $\mu(\mathbb{R}^n) > 0$  in einem Punkt  $x \in \mathbb{R}^n$  durch

$$\phi_s(x) = \int \frac{\mu(dy)}{|x-y|^s}.$$

Für  $n = 3$  und  $s = 1$  ist das im wesentlichen das Gravitationspotential der Massenverteilung  $\mu$ . Weiters definieren wir die  $s$ -Energie von  $\mu$  durch

$$I_s(\mu) = \int \phi_s(x)\mu(dx) = \int \int \frac{\mu(dx)\mu(dy)}{|x-y|^s}.$$

Der folgende Satz verbindet Hausdorffdimension mit potentialtheoretischen Methoden.

**Satz 2.41.** *Sei  $F$  eine Teilmenge von  $\mathbb{R}^n$ . Dann gelten*

(i) *Wenn es ein endliches (äußeres) Maß  $\mu$  auf  $F$  mit  $I_s(\mu) < \infty$  gibt, dann gilt  $\mathcal{H}^s(F) = \infty$  und  $\dim F \geq s$ .*

(ii) *Ist  $F$  eine Borelmenge mit  $\mathcal{H}^s(F) > 0$ , dann gibt es ein endliches Maß auf  $F$  mit  $\mu(F) > 0$ , sodass  $I_t(\mu) < \infty$  für alle  $t < s$ .*

*Beweis.* Wir benötigen in weiterer Folge nur den ersten Teil und beweisen auch nur diesen. Sei also  $I_s(\mu) < \infty$ . Wir definieren

$$F_1 := \left\{ x \in F : \limsup_{r \rightarrow 0} \frac{\mu(B(x, r))}{r^s} > 0 \right\}.$$

Für  $x \in F_1$  existiert also eine Folge von Radien  $(r_i)$  mit  $r_i \downarrow 0$  und ein  $\varepsilon > 0$ , sodass

$$\mu(B(x, r_i)) \geq \varepsilon r_i^s.$$

Weiters folgt aus  $I_s(\mu) = \int \int \frac{\mu(dx)\mu(dy)}{|x-y|^s} < \infty$ , dass  $\mu(\{x\}) = 0$  für alle  $x \in \mathbb{R}^n$ . Es folgt also aus der Stetigkeit von  $\mu$  für  $q_i$  hinreichend klein, dass

$$\mu(A_i) \geq \frac{1}{4} \varepsilon r_i^s \quad \text{für } A_i = B(x, r_i) \setminus B(x, q_i).$$

Nach eventuellem Übergang zu einer Teilfolge, können wir weiters annehmen, dass  $r_{i+1} < q_i$  für alle  $i$ . Daher sind die  $A_i$  disjunkte Kreisringe mit Mittelpunkt  $x$ . Für  $x \in F_1$  schließen wir nun aus  $|x-y|^{-s} \geq r_i^{-s}$  auf  $A_i$

$$\phi_s(x) = \int \frac{\mu(dy)}{|x-y|^s} \geq \sum_{i=1}^{\infty} \int_{A_i} \frac{\mu(dy)}{|x-y|^s} \geq \sum_{i=1}^{\infty} \frac{1}{4} \varepsilon r_i^s r_i^{-s} = \infty.$$

Aus der Voraussetzung  $I_s(\mu) = \int \phi_s(x)\mu(dx) < \infty$  folgt aber  $\phi_s(x) < \infty$  für  $\mu$ -fast alle  $x$ . Das bedeutet  $\mu(F_1) = 0$ . Da nach Definition  $\limsup_{r \rightarrow 0} \frac{\mu(B(x, r))}{r^s} = 0$  für  $x \in F \setminus F_1$ , erhalten wir aus dem ersten Teil von Satz 2.39, dass für alle  $c > 0$  gilt

$$\mathcal{H}^s(F) \geq \mathcal{H}^s(F \setminus F_1) \geq \mu(F \setminus F_1)/c \geq (\mu(F) - \mu(F_1))/c = \mu(F)/c.$$

Daraus folgt aber  $\mathcal{H}^s(F) = \infty$ , also die Behauptung.  $\square$

## 2.5 Maße auf Fraktalen

Sei  $(X, d)$  ein vollständiger metrischer Raum.

**Notation 2.42.** Sei  $(w_1, \dots, w_N)$  eine Familie von Abbildungen von  $X$  nach  $X$ . Für  $x \in X, E \in \mathcal{B}_X, \mu$  ein Wahrscheinlichkeitsmaß auf  $X$  und Indizes  $i_1, \dots, i_k \in \{1, \dots, N\}$  schreiben wir

$$\begin{aligned} x_{i_1 \dots i_k} &= w_{i_1} \circ \dots \circ w_{i_k}(x), & E_{i_1 \dots i_k} &= w_{i_1} \circ \dots \circ w_{i_k}(E), \\ \mu_{i_1 \dots i_k}(E) &= \mu((w_{i_1} \circ \dots \circ w_{i_k})^{-1}(E)). \end{aligned}$$

**Satz 2.43.** Sei  $\{X; w_1, \dots, w_N\}$  ein IFS mit Lipschitzkonstanten  $0 < c_1, \dots, c_N < 1$ . Sei weiters  $A$  der Attraktor dieses IFS und  $0 < p_1, \dots, p_N < 1$  mit  $\sum_{j=1}^N p_j = 1$ . Dann existiert ein eindeutiges Wahrscheinlichkeitsmaß  $\mu$  auf  $X$  mit  $\text{supp } \mu = A$  und

$$\mu = \sum_{j=1}^N p_j \mu_j.$$

$\mu$  heißt dann das invariante Maß bezüglich des IFS  $\{X; w_1, \dots, w_N\}$  mit den Wahrscheinlichkeiten  $0 < p_1, \dots, p_N < 1$ .

*Beweis. Schritt 1, Definition einer Folge  $\mu^k$  von Maßen:* Wir wählen  $x \in A$  beliebig, sei weiters  $\delta_x$  die Punktmasse in  $x$ . Für  $k \in \mathbb{N}$  definieren wir

$$\mu^k = \sum_{i_1, \dots, i_k=1}^N p_{i_1} \cdots p_{i_k} [\delta_x]_{i_1 \dots i_k},$$

das heißt, für  $f \in C(A)$  gilt

$$\int f d\mu^k = \sum_{i_1, \dots, i_k=1}^N p_{i_1} \cdots p_{i_k} f(x_{i_1 \dots i_k}).$$

**Schritt 2,  $\mu^k(f)$  ist eine Cauchyfolge:** Seien  $f \in C(A)$  und  $\varepsilon > 0$  beliebig. Weiters definiere  $c := \max_{1 \leq j \leq N} c_j$ . Dann existiert ein  $K > 0$ , sodass für alle  $x, y$  mit  $|x - y| < c^K |A|$  gilt, dass  $|f(x) - f(y)| < \varepsilon$  (aufgrund der gleichmäßigen Stetigkeit von  $f$ ). Sei nun  $l > k \geq K$ . Da  $x_{i_1 \dots i_l} \in A_{i_1 \dots i_k}$  ( $A$  ist ein Fixpunkt von  $W$ !) und  $|A_{i_1 \dots i_k}| \leq c_{i_1} \cdots c_{i_k} |A| \leq c^k |A|$ , erhalten wir

$$|f(x_{i_1 \dots i_k}) - f(x_{i_1 \dots i_k i_{k+1} \dots i_l})| < \varepsilon.$$

Summieren über  $i_{k+1}, \dots, i_l$  liefert

$$\left| f(x_{i_1 \dots i_k}) - \sum_{i_{k+1}, \dots, i_l=1}^N p_{i_{k+1}} \cdots p_{i_l} f(x_{i_1 \dots i_k i_{k+1} \dots i_l}) \right| < \varepsilon,$$

da  $1 = (\sum_{j=1}^N p_j)^m = \sum_{j_1, \dots, j_m=1}^N p_{j_1} \cdots p_{j_m}$ . Aus dem gleichen Grund erhalten wir nun nach Summation über  $i_1, \dots, i_k$ , dass  $|\int f d\mu^k - \int f d\mu^l| < \varepsilon$ . Das heißt, die Folge  $\{\int f d\mu^k\}$  konvergiert für alle  $f \in C(A)$ .

**Schritt 3, der Grenzwert ist das gesuchte Maß:** Der Limes

$$\Lambda(f) := \lim_{k \rightarrow \infty} \int f d\mu^k$$

definiert ein positives lineares Funktional auf  $C(A)$  mit  $\Lambda(1) = \lim \mu^k(A) = 1$ . Nach dem Rieszschen Darstellungssatz 2.12 existiert genau ein Wahrscheinlichkeitsmaß  $\mu$  das dieses Funktional darstellt. Da  $x_{i_1 \dots i_k} \in A_{i_1 \dots i_k}$ ,  $|A_{i_1 \dots i_k}| \leq c^k |A|$  und  $A = \bigcup A_{i_1 \dots i_k}$ , sind die Punkte  $\{x_{i_1 \dots i_k} : k \in \mathbb{N}, 1 \leq i_1, \dots, i_k \leq N\}$  dicht in  $A$ , also  $\text{supp } \mu = A$ . Schließlich ergibt sich

$$\begin{aligned} \int f d\mu &= \lim_{k \rightarrow \infty} \sum_{i_1, \dots, i_k=1}^N p_{i_1} \cdots p_{i_k} f(x_{i_1 \dots i_k}) = \sum_{j=1}^N p_j \lim_{k \rightarrow \infty} \sum_{i_2, \dots, i_k=1}^N p_{i_2} \cdots p_{i_k} f(x_{j i_2 \dots i_k}) \\ &= \sum_{j=1}^N p_j \lim_{k \rightarrow \infty} \int f \circ w_j d\mu^{k-1} = \sum_{j=1}^N p_j \int f \circ w_j d\mu = \sum_{j=1}^N p_j \int f d\mu_j \end{aligned}$$

und somit die Behauptung.

**Schritt 4, Eindeutigkeit von  $\mu$ :** Übung! □.

## 2.6 Dimension von selbstähnlichen Mengen

Wir werden in diesem Abschnitt die Hausdorffdimension der in Sektion 1.4 angegebenen Beispiele berechnen. Wir werden hierzu eine allgemeine Formel für die Dimension von (noch zu definierenden) selbstähnlichen Mengen angeben.

**Definition 2.44.** Eine Abbildung  $S : \mathbb{R}^n \rightarrow \mathbb{R}^n$  heißt Ähnlichkeit, falls ein  $0 < r < \infty$  existiert, sodass für alle  $x, y \in \mathbb{R}^n$  gilt

$$|S(x) - S(y)| = r|x - y|.$$

Seien  $h_r, t_b : \mathbb{R}^n \rightarrow \mathbb{R}^n$  definiert durch

$$h_r(x) := rx, \quad t_b(x) := x + b.$$

$h_r$  heißt Homothetie und  $t_b$  ist eine Translation.

Ähnlichkeiten auf  $\mathbb{R}^n$  lassen sich charakterisieren:

**Satz 2.45.**  $S : \mathbb{R}^n \rightarrow \mathbb{R}^n$  ist eine Ähnlichkeit genau dann, wenn  $S$  die Komposition einer orthogonalen Transformation  $U$ , einer Homothetie  $h_r$  und einer Translation  $t_b$  ist.

*Beweis.* Ist  $S$  die Zusammensetzung solcher Abbildungen, so ist  $S$  klarerweise eine Ähnlichkeit. Ist andererseits  $S$  eine Ähnlichkeit, so definieren wir für  $r := \text{Lip } S$  die Funktion  $U(x) := \frac{1}{r}(S(x) - S(0))$ . Dann ist  $U$  eine Isometrie mit Fixpunkt 0. Weiters erhält  $U$  innere Produkte, was man aus der folgenden Rechnung sieht:

$$\begin{aligned} \langle x, y \rangle &= \frac{1}{2}(|x|^2 + |y|^2 - |x - y|^2) \\ &= \frac{1}{2}(|U(x)|^2 + |U(y)|^2 - |U(x) - U(y)|^2) = \langle U(x), U(y) \rangle. \end{aligned}$$

Sei nun  $\{e_i : 1 \leq i \leq n\}$  eine Orthonormalbasis von  $\mathbb{R}^n$ . Dann ist demnach auch  $\{U(e_i) : 1 \leq i \leq n\}$  eine Orthonormalbasis von  $\mathbb{R}^n$  und daher gilt

$$U(x) = \sum_{i=1}^n \langle U(x), U(e_i) \rangle U(e_i) = \sum_{i=1}^n \langle x, e_i \rangle U(e_i).$$

Daher ist  $U$  linear und somit orthogonal. Es gilt also  $S(x) = rUx + S(0) = t_{S(0)} \circ h_r \circ U(x)$ , also die Behauptung, dass  $S$  Komposition einer orthogonalen Abbildung, einer Homothetie und einer Translation ist.  $\square$

**Definition 2.46.** Ein Attraktor  $A$  eines IFS  $\{X; w_j\}$  heißt selbstähnlich, falls alle Abbildungen  $w_j$  Ähnlichkeiten mit Lipschitzkonstante  $0 < c_j < 1$  sind.

**Bemerkung 2.47.** Ist  $A$  eine selbstähnliche Menge eines IFS  $\{X; w_j : 1 \leq j \leq N\}$  mit Lipschitzkonstanten  $\{c_j : 1 \leq j \leq N\}$ , so berechnen wir unter der Annahme, dass die  $w_j(A)$  „annähernd disjunkt“ sind

$$\mathcal{H}^s(A) = \sum_{j=1}^N \mathcal{H}^s(w_j(A)) = \sum_{j=1}^N c_j^s \mathcal{H}^s(A).$$

Ist weiters gesichert, dass  $A$  eine  $s$ -Menge ist, das heißt  $0 < \mathcal{H}^s(A) < \infty$ , so erhalten wir, dass die Hausdorffdimension  $s$  der Menge  $A$  die folgende Gleichung erfüllen muss:

$$\sum_{j=1}^N c_j^s = 1. \quad (2.9)$$

Wir präzisieren im folgenden die Voraussetzungen an das IFS, die benötigt werden, dass die Hausdorffdimension  $s$  von  $A$  durch diese Gleichung gegeben ist.

**Definition 2.48.** Ein IFS  $\{\mathbb{R}^n; w_j\}$  mit Ähnlichkeiten  $w_j$  erfüllt die open set condition, falls eine nichtleere, offene und beschränkte Menge  $V$  existiert, sodass

$$W(V) \subset V, \quad \text{und} \quad w_i(V) \cap w_j(V) = \emptyset \text{ für } i \neq j.$$

**Beispiel 2.49.** Ist  $([0, 1]; w_1, w_2)$  das IFS der Cantormenge, so erfüllt dieses die open set condition mit  $V = (0, 1)$

Wie angekündigt erhalten wir folgendes Resultat:

**Satz 2.50.** Sei  $\{\mathbb{R}^n; w_1, \dots, w_N\}$  ein IFS mit Ähnlichkeiten  $w_j$ , Lipschitzkonstanten  $0 < c_j < 1$  und Attraktor  $A$ . Falls dieses IFS die open set condition erfüllt, so hat  $A$  positives, endliches  $s$ -Hausdorffmaß  $\mathcal{H}^s(A)$ , wobei  $s$  die Gleichung

$$\sum_{j=1}^N c_j^s = 1$$

erfüllt. Weiters gilt  $\mathcal{H}^s(w_i(A) \cap w_j(A)) = 0$  für alle  $i \neq j$ .

Um diesen Satz zu beweisen, benötigen wir das folgende

**Lemma 2.51.** *Seien  $C_1, C_2, \delta$  positive Zahlen. Weiters sei  $\{U_\alpha\}_{\alpha \in A}$  eine Familie von disjunkten offenen Mengen in  $\mathbb{R}^n$  sodass es für alle  $\alpha \in A$  Kugeln  $B_1, B_2$  mit Radius  $C_1\delta$  bzw.  $C_2\delta$  gibt mit*

$$B_1 \subset U_\alpha \subset B_2.$$

*Dann schneidet keine Kugel mit Radius  $\delta$  mehr als  $(1 + 2C_2)^n C_1^{-n}$  der Mengen  $\overline{U_\alpha}$ .*

*Beweis.* Sei  $B$  eine Kugel mit Radius  $\delta$  und  $B \cap \overline{U_\alpha} \neq \emptyset$ . Dann ist  $\overline{U_\alpha}$  enthalten in der Kugel mit Radius  $(1 + 2C_2)\delta$  und konzentrisch mit  $B$ . (Übung!). Also, falls  $N$  Mengen der  $U_\alpha$ 's  $B$  schneiden, dann gibt es  $N$  disjunkte Kugeln mit Radius  $C_1\delta$ , die enthalten sind in einer Kugel mit Radius  $(1 + 2C_2)\delta$ . Wenn man ihr Volumen addiert, erhält man

$$N(C_1\delta)^n \leq ((1 + 2C_2)\delta)^n,$$

also  $N \leq (1 + 2C_2)^n C_1^{-n}$ . □

*Beweis von Satz 2.50.* Für  $k \in \mathbb{N}$  gilt  $A = W^k(A) = \bigcup A_{i_1 \dots i_k}$ ,  $|A_{i_1 \dots i_k}| = c_{i_1} \cdots c_{i_k} |A|$ . Sei weiters  $c := \max_{1 \leq j \leq N} c_j$  die Lipschitzkonstante des IFS. Zusätzlich sei  $\mu$  das invariante Wahrscheinlichkeitsmaß auf  $A$  von Satz 2.43 mit  $p_j = c_j^s$ .

**Schritt 1,  $\mathcal{H}^s(A) < \infty$ :** Sei  $\delta_k := c^k |A|$ . Dann gilt

$$\mathcal{H}_{\delta_k}^s(A) \leq \sum_{i_1, \dots, i_k=1}^N |A_{i_1 \dots i_k}|^s = \sum_{i_1, \dots, i_k} (c_{i_1} \cdots c_{i_k})^s |A|^s = |A|^s.$$

Da  $c^k \rightarrow 0$  mit  $k \rightarrow \infty$ , erhalten wir nach Grenzübergang für  $\delta \rightarrow 0$ , dass  $\mathcal{H}^s(A) \leq |A|^s < \infty$ .

**Schritt 2, Definition von  $N_K$ :** Sei  $V$  die Menge aus der open set condition und seien die Konstanten  $C_1, C_2$  so, dass  $V$  eine Kugel mit Radius  $C_1$  enthält und in einer Kugel mit Radius  $C_2$  enthalten ist. Sei dann  $N_K := \frac{(1+2C_2)^n}{C_1^n (\min_i c_i)^n}$ .

**Schritt 3:** Für alle Kugeln  $B$  mit Radius  $\delta < 1$  gilt  $\mu(B) \leq N_K \delta^s$ : Für eine Folge  $\sigma \in \{1, \dots, N\}^{\mathbb{N}}$ ,  $\sigma = (i_1, i_2, \dots)$ , sei  $k \in \mathbb{N}$  die kleinste natürliche Zahl, sodass

$$(\min_i c_i) \delta \leq c_{i_1} \cdots c_{i_k} \leq \delta.$$

Sei nun  $Q$  die Menge aller (endlichen) Folgen  $(i_1, \dots, i_k)$ , die auf diese Weise entstehen. Da es für jede Folge  $\sigma$  genau ein  $k$  mit dieser Eigenschaft gibt, ist  $\{V_{i_1 \dots i_k} : (i_1, \dots, i_k) \in Q\}$  eine Familie disjunkter Mengen. Da  $A \subset \overline{V}$  (Übung!), folgt weiter

$$A \subset \bigcup_{(i_1, \dots, i_k) \in Q} A_{i_1 \dots i_k} \subset \bigcup_{(i_1, \dots, i_k) \in Q} \overline{V_{i_1 \dots i_k}}$$

Nach Schritt 2 enthält  $V_{i_1 \dots i_k}$  eine Kugel vom Radius  $C_1 c_{i_1} \cdots c_{i_k} \geq C_1 (\min_i c_i) \delta$  und ist enthalten in einer Kugel vom Radius  $C_2 c_{i_1} \cdots c_{i_k} \leq C_2 \delta$ . Sei nun  $Q_1$  die Menge

aller Folgen  $(i_1, \dots, i_k)$  in  $Q$ , sodass  $B \overline{V_{i_1 \dots i_k}}$  schneidet. Nach Lemma 2.51 sind das maximal  $N_K = \frac{(1+2C_2)^n}{(C_1 \min_i c_i)^n}$  Folgen. Aus der Identität

$$\mu = \sum_{(i_1, \dots, i_k) \in Q} (c_{i_1} \cdots c_{i_k})^s \mu_{i_1 \dots i_k}$$

folgt nun

$$\mu(B) = \mu(A \cap B) = \sum_{(i_1, \dots, i_k) \in Q} (c_{i_1} \cdots c_{i_k})^s \mu_{i_1 \dots i_k}(A_{i_1 \dots i_k} \cap B) \leq \delta^s N_K.$$

**Schritt 4:** Für jede Überdeckung  $\{E_j\}_j$  von  $A$  mit Mengen vom Durchmesser  $< 1$  gilt:

$$\sum |E_j|^s \geq N_K^{-1}.$$

Sei nun  $\{E_j\}_j$  eine Überdeckung mit  $A$  mit Mengen mit Durchmesser  $< 1$ . Da jede Menge  $E_j$  in einer abgeschlossenen Kugel mit gleichem Durchmesser enthalten ist, reicht es, für eine Überdeckung von  $A$  mit Kugeln  $\{B_j\}_j$  mit Durchmesser  $\delta_j < 1$  die Ungleichung

$$\sum \delta_j^s \geq N_K^{-1}.$$

zu zeigen. Diese folgt allerdings aus Schritt 3, denn

$$1 = \mu(A) \leq \sum \mu(B_j) \leq N_K \sum \delta_j^s.$$

Nun folgt aus diesem Schritt die untere Schranke  $\mathcal{H}^s(A) \geq N_K^{-1}$ .

**Schritt 5,**  $\mathcal{H}^s(w_i(A) \cap w_j(A)) = 0$ : Da die  $w_j$  Ähnlichkeiten sind, folgt  $\mathcal{H}^s(w_j(A)) = c_j^s \mathcal{H}^s(A)$ , also  $\mathcal{H}^s(A) = \sum_{j=1}^N \mathcal{H}^s(w_j(A))$  und somit, da  $A = \bigcup_{j=1}^N w_j(A)$

$$\mathcal{H}^s(w_i(A) \cap w_j(A)) = 0. \quad \square$$

**Korollar 2.52.** *Wir erhalten folgende Werte für die Hausdorffdimension verschiedener Fraktale:*

| $A$                   | $\dim A$   | $\approx$ |
|-----------------------|------------|-----------|
| Cantormenge           | $\log_3 2$ | 0.63      |
| Sierpiński-Dreieck    | $\log_2 3$ | 1.58      |
| Sierpiński-Teppich    | $\log_3 8$ | 1.89      |
| Kochsche Schneeflocke | $\log_3 4$ | 1.26      |

*Beweis.* Übung! □

## 3 Markovketten und Ergodensätze

### 3.1 Bedingte Erwartung

Seien im folgenden  $(\Omega, \mathcal{F}, \mathbb{P})$  ein Wahrscheinlichkeitsraum und  $\mathcal{F}' \subset \mathcal{F}$  eine Unter- $\sigma$ -Algebra von  $\mathcal{F}$ . Weiters sei  $(X, \mathcal{E})$  eine Menge  $X$  ausgestattet mit einer  $\sigma$ -Algebra  $\mathcal{E}$ .



**Definition 3.1.** Sei  $\xi : \Omega \rightarrow \mathbb{R}$  eine messbare Abbildung mit  $\mathbb{E}|\xi| < \infty$ . Eine  $\mathcal{F}'$ -messbare Abbildung  $\eta$  heißt bedingte Erwartung von  $\xi$  unter  $\mathcal{F}'$  (i.Z.  $\mathbb{E}(\xi|\mathcal{F}')$ ), falls

$$\int_A \xi d\mathbb{P} = \int_A \eta d\mathbb{P} \quad \text{für alle } A \in \mathcal{F}' \text{ gilt.}$$

**Lemma 3.2.** Die bedingte Erwartung existiert und ist  $\mathbb{P}$ -fast sicher eindeutig bestimmt.

*Beweis.* Sei  $\xi = \xi_+ - \xi_-$  die kanonische Zerlegung in die Differenz zweier nichtnegativer Funktionen  $\xi_+, \xi_-$ . Definiere dann zwei Maße  $\mu_{\pm}$  auf  $(\Omega, \mathcal{F}')$  durch

$$\mu_{\pm}(A) := \int_A \xi_{\pm} d\mathbb{P}, \quad \text{für } A \in \mathcal{F}'.$$

Dann sind  $\mu_{\pm}$  absolut stetig bezüglich  $\mathbb{P}$ , es gibt also nach dem Satz von Radon-Nikodým Dichten  $\eta_{\pm}$  für  $\mu_{\pm}$  bezüglich  $\mathbb{P}$ . Die  $\mathcal{F}'$  messbare Abbildung  $\eta_+ - \eta_-$  ist dann die ( $\mathbb{P}$ -fast sicher eindeutig bestimmte) bedingte Erwartung  $\mathbb{E}(\xi|\mathcal{F}')$ .  $\square$

**Beispiel 3.3.** (i) Ist  $\mathcal{F}' = \{\emptyset, \Omega\}$  die triviale  $\sigma$ -Algebra, so gilt  $\mathbb{E}(\xi|\mathcal{F}') = \mathbb{E}\xi$   $\mathbb{P}$ -fast sicher.

(ii) Ist  $\xi$   $\mathcal{F}'$ -messbar, so gilt  $\mathbb{E}(\xi|\mathcal{F}') = \xi$   $\mathbb{P}$ -fast sicher.

(iii) Ist  $\xi$  unabhängig von  $\mathcal{F}'$ , das heißt es gilt für alle  $B \in \mathcal{B}_{\mathbb{R}}$  und  $A \in \mathcal{F}'$ , dass

$$\mathbb{P}([\xi \in B] \cap A) = \mathbb{P}(\xi \in B)\mathbb{P}(A),$$

dann gilt  $\mathbb{E}(\xi|\mathcal{F}') = \mathbb{E}\xi$   $\mathbb{P}$ -fast sicher. Es gilt nämlich für  $A \in \mathcal{F}'$

$$\int_A \xi d\mathbb{P} = \int 1_A \xi d\mathbb{P} = \mathbb{E}(1_A \xi).$$

Aufgrund der Unabhängigkeit von  $1_A$  und  $\xi$  erhalten wir weiters

$$\mathbb{E}(1_A \xi) = \mathbb{E}1_A \mathbb{E}\xi = \mathbb{P}(A)\mathbb{E}\xi = \int_A \mathbb{E}\xi d\mathbb{P},$$

also gilt nach Definition der bedingten Erwartung die Behauptung.

(iv) Sei  $\Omega = \bigcup_{j=1}^{\infty} \Omega_j$  wobei die  $\Omega_j$  paarweise disjunkt seien.  $\{\Omega_j\}_j$  ist dann eine Partition von  $\Omega$ . Wir nehmen weiters an, dass  $\mathbb{P}(\Omega_j) > 0$  für alle  $j$  gilt. Sei dann  $\mathcal{F}' := \sigma(\Omega_1, \Omega_2, \dots)$  die von dieser Partition erzeugte  $\sigma$ -Algebra. Dann gilt

$$\mathbb{E}(\xi|\mathcal{F}')(\omega) = \sum_{j=1}^{\infty} \frac{\mathbb{E}(\xi; \Omega_j)}{\mathbb{P}(\Omega_j)} 1_{\Omega_j}(\omega). \quad (3.1)$$

Es gilt für ein Atom  $\Omega_k$  aus  $\mathcal{F}'$ , dass

$$\int_{\Omega_k} \xi d\mathbb{P} = \mathbb{E}(\xi; \Omega_k) = \int_{\Omega_k} \frac{\mathbb{E}(\xi; \Omega_k)}{\mathbb{P}(\Omega_k)} d\mathbb{P},$$

nach Definition der bedingten Erwartung, also gilt 3.1.

**Definition 3.4.** Wir definieren die bedingte Wahrscheinlichkeit eines Ereignisses  $A \in \mathcal{F}$  unter  $\mathcal{F}'$  durch

$$\mathbb{P}(A|\mathcal{F}') := \mathbb{E}(1_A|\mathcal{F}').$$

**Bemerkung 3.5.** Nach Definition ist

$$\int_B \mathbb{P}(A|\mathcal{F}') d\mathbb{P} = \int_B \mathbb{E}(1_A|\mathcal{F}') d\mathbb{P} = \int_B 1_A d\mathbb{P} = \mathbb{P}(A \cap B) \quad \text{für alle } B \in \mathcal{F}'$$

die definierende Eigenschaft der bedingten Wahrscheinlichkeit.

Im Falle der bedingten Wahrscheinlichkeit, bzw. wenn  $\xi = 1_A$  eine Indikatorfunktion ist, wird (3.1) zu

$$\mathbb{P}(A|\mathcal{F}')(\omega) = \sum_{j=1}^{\infty} \mathbb{P}(A|\Omega_j) 1_{\Omega_j}(\omega). \quad (3.2)$$

Dabei ist die elementare bedingte Wahrscheinlichkeit  $\mathbb{P}(A|\Omega_j)$  definiert als  $\mathbb{P}(A \cap \Omega_j)/\mathbb{P}(\Omega_j)$ .

**Beispiel 3.6** (Bayes-Formel). Seien  $A \in \mathcal{F}$ ,  $G \in \mathcal{F}'$ . Dann gilt

$$\mathbb{P}(G|A) = \frac{\int_G \mathbb{P}(A|\mathcal{F}') d\mathbb{P}}{\int_{\Omega} \mathbb{P}(A|\mathcal{F}') d\mathbb{P}}.$$

Die Behauptung folgt direkt aus der obigen Bemerkung 3.5. Ist nun  $\mathcal{F}'$  die  $\sigma$ -Algebra, die von der Partition  $(\Omega_j)_j$  von  $\Omega$  erzeugt wird, so wird diese Formel mittels (3.2) zu

$$\mathbb{P}(\Omega_k|A) = \frac{\mathbb{P}(A|\Omega_k)\mathbb{P}(\Omega_k)}{\sum_{j=1}^{\infty} \mathbb{P}(A|\Omega_j)\mathbb{P}(\Omega_j)},$$

also der klassischen Bayes-Formel. Nun noch ein kurzes Beispiel zur Bayes-Formel: Sei  $X$  gleichmäßig verteilt auf  $\{1, \dots, 6\}$  und  $Y$  die Anzahl der „Adler“, wenn  $X$  Münzen geworfen werden. Was ist nun die Wahrscheinlichkeit, dass  $X = i$  war, wenn nur bekannt ist, dass  $Y = 0$ . Wir erhalten aus der Bayes-Formel

$$\mathbb{P}(X = i|Y = 0) = \frac{\mathbb{P}(Y = 0|X = i)\mathbb{P}(X = i)}{\sum_{j=1}^6 \mathbb{P}(Y = 0|X = j)\mathbb{P}(X = j)} = \frac{2^{-i}6^{-1}}{\sum_{j=1}^6 2^{-j}6^{-1}} = \frac{2^{-i}}{1 - 2^{-6}}.$$

**Definition 3.7.** Seien  $\xi$  eine Zufallsvariable auf  $(\Omega, \mathcal{F}, \mathbb{P})$  und  $\eta : \Omega \rightarrow X$  messbar. Dann definieren wir die bedingte Erwartung von  $\xi$  unter  $\eta$  als

$$\mathbb{E}(\xi|\eta) := \mathbb{E}(\xi|\sigma(\eta)),$$

wobei  $\sigma(\eta)$  die kleinste  $\sigma$ -Algebra auf  $\Omega$  bezeichnet, bezüglich der  $\eta$  messbar ist.

**Beispiel 3.8.** Seien  $X_1, X_2$  unabhängig und poissonverteilt mit Intensität  $\lambda$  und weiters  $Y := X_1 + X_2$ . Es gilt also

$$\mathbb{P}(X_j = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k \geq 0, j = 1, 2.$$

Mit  $\Omega_k := [Y = k]$  ist  $(\Omega_k)_{k=0}^\infty$  eine Partition von  $\Omega$  und  $\Omega_k$  sind Atome von  $\sigma(Y)$ . Es folgt also aus (3.2), dass

$$\mathbb{P}(X_1 = i|Y)(\omega) = \sum_{k=0}^{\infty} \mathbb{P}(X_1 = i|Y = k)1_{[Y=k]}$$

mit

$$\mathbb{P}(X_1 = i|Y = k) = \frac{\mathbb{P}(X_1 = i, Y = k)}{\mathbb{P}(Y = k)} = \frac{\mathbb{P}(X_1 = i, X_2 = k - i)}{\mathbb{P}(Y = k)} = \binom{k}{i} 2^{-k}.$$

**Beispiel 3.9.** Seien  $X$  und  $Y$  integrierbare Zufallsvariablen mit gemeinsamer Dichte  $f_{X,Y}$  bezüglich des Lebesguemaßes (das heißt  $\mathbb{P}(X \in A, Y \in B) = \int_A \int_B f_{X,Y}(x, y) dy dx$ ). Dann gilt

$$\mathbb{E}(X|Y)(\omega) = \frac{\int_{\mathbb{R}} x f_{X,Y}(x, Y(\omega)) dx}{\int_{\mathbb{R}} f_{X,Y}(w, Y(\omega)) dw}.$$

Es ist nämlich für  $B = [Y \in C]$

$$\begin{aligned} \int_B X d\mathbb{P} &= \int_C \int_{\mathbb{R}} x f_{X,Y}(x, y) dx dy = \int_C \frac{\int_{\mathbb{R}} f_{X,Y}(z, y) dz}{\int_{\mathbb{R}} f_{X,Y}(w, y) dw} \int_{\mathbb{R}} x f_{X,Y}(x, y) dx dy \\ &= \int_{\mathbb{R}} \int_C \frac{\int_{\mathbb{R}} x f_{X,Y}(x, y) dx}{\int_{\mathbb{R}} f_{X,Y}(w, y) dw} f_{X,Y}(z, y) dy dz = \int_B \frac{\int_{\mathbb{R}} x f_{X,Y}(x, Y(\omega)) dx}{\int_{\mathbb{R}} f_{X,Y}(w, Y(\omega)) dw} d\mathbb{P}(\omega). \end{aligned}$$

**Satz 3.10** (Eigenschaften der bedingten Erwartung). 1. (Linearität) Seien  $X, Y$  Zufallsvariablen und  $a, b \in \mathbb{C}$ . Dann gilt

$$\mathbb{E}(aX + bY|\mathcal{F}') = a\mathbb{E}(X|\mathcal{F}') + b\mathbb{E}(Y|\mathcal{F}').$$

2. (Monotonie) Ist  $X \leq Y$   $\mathbb{P}$ -fast sicher, so folgt  $\mathbb{E}(X|\mathcal{F}') \leq \mathbb{E}(Y|\mathcal{F}')$ .

3. (Monotone Konvergenz) Sei  $X_n \geq 0$ ,  $X_n \uparrow X$ ,  $\mathbb{E}X < \infty$ . Dann folgt  $\mathbb{E}(X_n|\mathcal{F}') \uparrow \mathbb{E}(X|\mathcal{F}')$ .

4. (Jensen Ungleichung) Ist  $\varphi$  konvex und gelten  $\mathbb{E}|X|, \mathbb{E}|\varphi(X)| < \infty$ . Dann gilt

$$\varphi(\mathbb{E}(X|\mathcal{F}')) \leq \mathbb{E}(\varphi(X)|\mathcal{F}').$$

5. (Tower property) Seien  $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \mathcal{F}$ . Dann gilt

$$\mathbb{E}(\mathbb{E}(X|\mathcal{F}_1)|\mathcal{F}_2) = \mathbb{E}(X|\mathcal{F}_1) = \mathbb{E}(\mathbb{E}(X|\mathcal{F}_2)|\mathcal{F}_1).$$

6. Ist  $X$   $\mathcal{F}'$ -messbar und gelten  $\mathbb{E}|Y|, \mathbb{E}|XY| < \infty$ , so folgt

$$\mathbb{E}(XY|\mathcal{F}') = X\mathbb{E}(Y|\mathcal{F}').$$

7. Sind  $X$  und  $Y$  unabhängig, so gilt

$$\mathbb{E}(Y|X) = \mathbb{E}Y.$$

8. Sei  $\mathbb{E}X^2 < \infty$ .  $L^2(\Omega, \mathcal{F})$  ist ein Hilbertraum und  $L^2(\Omega, \mathcal{F}')$  ist ein abgeschlossener Unterraum. Weiters ist  $\mathbb{E}(X|\mathcal{F}')$  die Projektion von  $X$  auf  $L^2(\mathcal{F}')$ .

*Beweis.* 1. Sei  $A \in \mathcal{F}'$  Dann gilt nach der Linearität des Integrals und der Definition der bedingten Erwartung

$$\begin{aligned} \int_A a\mathbb{E}(X|\mathcal{F}') + b\mathbb{E}(Y|\mathcal{F}')d\mathbb{P} &= a \int_A \mathbb{E}(X|\mathcal{F}')d\mathbb{P} + b \int_A \mathbb{E}(Y|\mathcal{F}')d\mathbb{P} \\ &= a \int_A Xd\mathbb{P} + b \int_A Yd\mathbb{P} = \int_A aX + bYd\mathbb{P}, \end{aligned}$$

also die Behauptung.

2. Es gilt aufgrund der Definition der bedingten Erwartung und der Monotonie des Integrals für  $A \in \mathcal{F}'$

$$\int_A \mathbb{E}(X|\mathcal{F}')d\mathbb{P} = \int_A Xd\mathbb{P} \leq \int_A Yd\mathbb{P} = \int_A \mathbb{E}(Y|\mathcal{F}')d\mathbb{P}.$$

Daher gilt  $\mathbb{E}(X|\mathcal{F}') \leq \mathbb{E}(Y|\mathcal{F}')$   $\mathbb{P}$ -fast sicher.

3. Definiere  $Y_n := X - X_n$ , dann ist zu zeigen, dass  $\mathbb{E}(Y_n|\mathcal{F}') \downarrow 0$ . Da  $Y_n \downarrow$ , folgt aus Punkt 2., dass  $Z_n := \mathbb{E}(Y_n|\mathcal{F}') \downarrow Z_\infty$  für ein  $Z_\infty$ . Ist nun  $A \in \mathcal{F}'$ , so gilt nach der Definition des bedingten Erwartungswerts

$$\int_A Z_n d\mathbb{P} = \int_A Y_n d\mathbb{P}.$$

Für  $n \rightarrow \infty$  und dem Satz von der dominierten Konvergenz folgt  $\int_A Z_\infty d\mathbb{P} = 0$  für alle  $A \in \mathcal{F}'$ , also folgt  $Z_\infty = 0$  fast sicher.

4. Sei ohne Beschränkung der Allgemeinheit  $\varphi$  nicht linear. Definieren wir

$$S := \{(a, b) : a, b \in \mathbb{Q}, ax + b \leq \varphi(x) \text{ für alle } x\}.$$

Dann ist  $\varphi(x) = \sup\{ax + b : (a, b) \in S\}$ . Ist nun  $\varphi(x) \geq ax + b$ , dann folgt aus 1. und 2.

$$\mathbb{E}(\varphi(X)|\mathcal{F}') \geq a\mathbb{E}(X|\mathcal{F}') + b \quad \mathbb{P} - \text{fast sicher.}$$

Nach Übergang zum Supremum folgt die Jensen-Ungleichung.

5.  $\mathbb{E}(X|\mathcal{F}_1)$  ist  $\mathcal{F}_2$ -messbar, daraus folgt die erste Gleichung. Sei nun  $A \in \mathcal{F}_1 \subset \mathcal{F}_2$ . Dann gilt nach Definition der bedingten Erwartung

$$\int_A \mathbb{E}(X|\mathcal{F}_1)d\mathbb{P} = \int_A Xd\mathbb{P} = \int_A \mathbb{E}(X|\mathcal{F}_2)d\mathbb{P},$$

also folgt wieder nach Definition der bedingten Erwartung

$$\mathbb{E}(\mathbb{E}(X|\mathcal{F}_2)|\mathcal{F}_1) = \mathbb{E}(\mathbb{E}(X|\mathcal{F}_1)|\mathcal{F}_1) = \mathbb{E}(X|\mathcal{F}_1),$$

also die zweite Gleichung.

6. Wir zeigen das zunächst für eine Indikatorfunktion  $X = 1_B$  mit  $B \in \mathcal{F}'$ . Für allgemeine Funktionen  $X$  folgt das durch übliche Linearitäts- und Approximationsargumente aus den obigen Eigenschaften. Sei zusätzlich  $A \in \mathcal{F}'$ . Dann gilt nach Definition der bedingten Erwartung

$$\int_A 1_B \mathbb{E}(Y|\mathcal{F}')d\mathbb{P} = \int_{A \cap B} \mathbb{E}(Y|\mathcal{F}')d\mathbb{P} = \int_{A \cap B} Yd\mathbb{P} = \int_A 1_B Yd\mathbb{P},$$

also  $\mathbb{E}(1_B Y | \mathcal{F}') = 1_B \mathbb{E}(Y | \mathcal{F}')$  und somit die Behauptung für Indikatorfunktionen  $X$ .

7. siehe Beispiel 3.3, Punkt 3.

8. Sei  $Z \in L^2(\Omega, \mathcal{F}')$ . Dann gilt nach Punkt 6., dass  $Z \mathbb{E}(X | \mathcal{F}') = \mathbb{E}(ZX | \mathcal{F}')$  und somit

$$\mathbb{E}(Z \mathbb{E}(X | \mathcal{F}')) = \mathbb{E} \mathbb{E}(ZX | \mathcal{F}') = \mathbb{E}(ZX)$$

nach Punkt 5. Also ist  $\mathbb{E}(Z(X - \mathbb{E}(X | \mathcal{F}')))) = 0$  für alle  $Z \in L^2(\Omega, \mathcal{F}')$ . Sei nun  $Y \in L^2(\Omega, \mathcal{F}')$  und  $Z = Y - \mathbb{E}(X | \mathcal{F}')$ , dann folgt daraus insgesamt

$$\mathbb{E}(X - Y)^2 = \mathbb{E}(X - \mathbb{E}(X | \mathcal{F}') - Z)^2 = \mathbb{E}(X - \mathbb{E}(X | \mathcal{F}'))^2 + \mathbb{E}Z^2.$$

Das heißt also schließlich, dass  $\mathbb{E}(X - Y)^2$  minimal ist für  $Z = 0$  also für  $Y = \mathbb{E}(X | \mathcal{F}')$ , also ist  $\mathbb{E}(X | \mathcal{F}')$  die Projektion auf  $L^2(\Omega, \mathcal{F}')$ .  $\square$

**Definition 3.11.** Seien  $A \in \mathcal{F}$  und  $x \in X$ . Dann heißt eine Zufallsvariable  $g : X \rightarrow \mathbb{R}$  mit

$$\mathbb{P}(A \cap [\eta \in B]) = \int_B g(x) \mathbb{P}_\eta(dx), \quad \text{für alle } B \in \mathcal{E}$$

die bedingte Wahrscheinlichkeit von  $A$  unter  $\eta = x$ , i.Z.  $\mathbb{P}(A | \eta = x)$ .

*Beweis.* Der Beweis der Existenz von  $g$  und der  $\mathbb{P}_\eta$ -fast sicheren Eindeutigkeit erfolgt analog zum Beweis der Existenz und Eindeutigkeit der bedingten Erwartung und wird als Übung überlassen.  $\square$

**Bemerkung 3.12.** Man kann zeigen, dass für eine  $\sigma(\eta)$ -messbare reelle Zufallsvariable  $Z : \Omega \rightarrow \mathbb{R}$  eine messbare Funktion  $g : (X, \mathcal{E}) \rightarrow (\mathbb{R}, \mathcal{B}_\mathbb{R})$  existiert mit

$$Z = g \circ \eta.$$

Sei nun  $g : (X, \mathcal{E}) \rightarrow (\mathbb{R}, \mathcal{B}_\mathbb{R})$  die nach diesem Schema zu  $\mathbb{E}(\xi | \eta)$  zugeordnete Funktion. Dann gilt

$$\mathbb{E}(\xi | \eta) = g \circ \eta$$

und  $g$  ist charakterisiert durch

$$\int_{[\eta \in B]} \xi d\mathbb{P} = \int_{[\eta \in B]} \mathbb{E}(\xi | \eta) d\mathbb{P} = \int_{[\eta \in B]} g \circ \eta d\mathbb{P} = \int_B g d\mathbb{P}_\eta.$$

Man kann also analog zur obigen Definition auch

$$\mathbb{E}(\xi | \eta = x) := g(x)$$

setzen und es folgt dann  $\mathbb{E}(\xi | \eta)(\omega) = \mathbb{E}(\xi | \eta = \eta(\omega))$   $\mathbb{P}$ -fast sicher. Der Spezialfall  $\xi = 1_A$  liefert wieder Definition 3.11.

## 3.2 Markovketten

**Definition 3.13.** Sei  $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \dots \subset \mathcal{F}$  eine Folge von  $\sigma$ -Algebren und  $\xi_n : \Omega \rightarrow X$  ( $\mathcal{F}_n, \mathcal{E}$ )-messbare Abbildungen. Die Folge  $(\xi_n)$  heißt Markov-Kette, falls für alle  $n \geq 0$  und alle  $B \in \mathcal{E}$  gilt

$$\mathbb{P}(\xi_{n+1} \in B | \mathcal{F}_n) = \mathbb{P}(\xi_{n+1} \in B | \xi_n).$$

Die Markovkette heißt homogen, falls  $x \mapsto \mathbb{P}(\xi_{n+1} \in B | \xi_n = x)$  für alle  $n$  gleich ist.

**Beispiel 3.14.** Seien  $\eta_1, \eta_2, \dots : \Omega \rightarrow \mathbb{R}$  unabhängige Zufallsvariablen und sei  $A \in \mathcal{B}_{\mathbb{R}}$ . Seien  $S_0 = 0$  und  $S_{n+1} = S_n + \eta_{n+1}$ . Dann ist  $(S_n)$  eine Markovkette. Falls die  $\eta_i$  identisch verteilt sind, so ist diese Markovkette homogen. Wir berechnen zuerst  $\mathbb{P}(S_{n+1} \in A | S_n)$ . Es gilt für eine  $S_n$ -messbare Menge  $B = [S_n \in C]$

$$\begin{aligned} \int_B 1_A(S_{n+1}) d\mathbb{P} &= \int_B 1_A(S_n + \eta_{n+1}) d\mathbb{P} \\ &= \int_C \int_{\mathbb{R}} 1_A(x + y) d\mathbb{P}_{\eta_{n+1}}(y) d\mathbb{P}_{S_n}(x) = \int_C \mathbb{P}_{\eta_{n+1}}(A - x) d\mathbb{P}_{S_n}(x) \\ &= \int_B \mathbb{P}_{\eta_{n+1}}(A - S_n(\omega)) d\mathbb{P}(\omega), \end{aligned}$$

also folgt nach Definition der bedingten Wahrscheinlichkeit  $\mathbb{P}(S_{n+1} \in A | S_n)(\omega) = \mathbb{P}_{\eta_{n+1}}(A - S_n(\omega))$   $\mathbb{P}$ -fast sicher.

Nun berechnen wir  $\mathbb{P}(S_{n+1} \in A | S_n, \dots, S_0)$ . Es reicht, Mengen der Art  $B = [\eta_0 \in B_0, \dots, \eta_n \in B_n]$  in  $\sigma(S_0, \dots, S_n)$  zu betrachten. Mit so einer Menge erhalten wir

$$\begin{aligned} \int_B 1_A(S_{n+1}) d\mathbb{P} &= \int_{B_0} \cdots \int_{B_n} \int_{\mathbb{R}} 1_A(x_0 + \cdots + x_n + y) d\mathbb{P}_{\eta_{n+1}}(y) d\mathbb{P}_{\eta_n}(x_n) \cdots d\mathbb{P}_{\eta_0}(x_0) \\ &= \int_{B_0} \cdots \int_{B_n} \mathbb{P}_{\eta_{n+1}}(A - x_0 - \cdots - x_n) d\mathbb{P}_{\eta_n}(x_n) \cdots d\mathbb{P}_{\eta_0}(x_0) \\ &= \int_B \mathbb{P}_{\eta_{n+1}}(A - S_n(\omega)) d\mathbb{P}(\omega), \end{aligned}$$

also insgesamt die Markoveigenschaft  $\mathbb{P}(S_{n+1} \in A | S_n) = \mathbb{P}(S_{n+1} \in A | S_n, \dots, S_0)$ .

Wir werden in Zukunft immer die Folge der  $\sigma$ -Algebren  $\mathcal{F}_n := \sigma(\xi_0, \dots, \xi_n)$  betrachten. Definitionsgemäß ist also in einer Markovkette der Wert der  $(n+1)$ -ten Zufallsvariable nur vom Wert der  $n$ -ten Zufallsvariable abhängig (und nicht von der gesamten Vergangenheit). Die bedingte Wahrscheinlichkeit  $\mathbb{P}(\xi_{n+1} \in B | \xi_n = x)$  wird als  $(n+1)$ -te Übergangswahrscheinlichkeitsdichte von  $x$  nach  $B$  interpretiert. Der folgende Satz garantiert die Existenz einer Markovkette mit bestimmten Übergangswahrscheinlichkeiten.

Dazu setzen wir ab jetzt voraus, dass  $X$  ein polnischer Raum (separabel, vollständig, metrisch) und  $\mathcal{E}$  die Borelsche  $\sigma$ -Algebra  $\mathcal{B}_X$  sind.

**Satz 3.15.** Sei  $p : X \times \mathcal{B}_X \rightarrow [0, 1]$  eine Abbildung mit den Eigenschaften

- (i) Für alle  $A \in \mathcal{B}_X$  ist  $p(\cdot, A) : X \rightarrow [0, 1]$   $(\mathcal{B}_X, \mathcal{B}_{\mathbb{R}})$ -messbar.
- (ii) Für alle  $x \in X$  ist  $p(x, \cdot) : \mathcal{B}_X \rightarrow [0, 1]$  ein Wahrscheinlichkeitsmaß.

Sei weiters  $\nu$  ein Wahrscheinlichkeitsmaß auf  $\mathcal{B}_X$ . Dann existieren ein Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, \mathbb{P}^\nu)$  und eine homogene Markovkette  $(\xi_n)$  mit der Anfangsverteilung  $\nu$  (das heißt die Verteilung von  $\xi_0$  ist  $\nu$ ) und den Übergangswahrscheinlichkeiten

$$\mathbb{P}^\nu(\xi_{n+1} \in B | \xi_n = x) = p(x, B).$$

*Beweis.* Sei  $\Omega = X^{\mathbb{N}}$ ,  $\mathcal{F} = \mathcal{B}_\Omega$ . Dann ist  $\mathcal{F}$  die kleinste  $\sigma$ -Algebra, sodass die Projektionen  $\xi_n : \Omega \rightarrow X$ , definiert durch

$$\omega = (x_0, x_1, \dots) \mapsto \xi_n(\omega) = x_n,$$

$(\mathcal{F}, \mathcal{B}_X)$  messbar sind. Dies folgt aus der Tatsache, dass  $X$  separabel ist (Übung!). Somit sind auch die Abbildungen

$$\pi_n : \Omega \rightarrow X^{n+1}, \quad (x_0, x_1, \dots) \mapsto (x_0, \dots, x_n)$$

messbar. Sei  $A \in \mathcal{B}_{X^{n+1}}$ . Dann definieren wir

$$\mathbb{P}(\pi_n \in A) := \mathbb{P}^\nu(\pi_n \in A) := \int_X \nu(dx_0) \int_X p(x_0, dx_1) \cdots \int_X 1_A(x_0, \dots, x_n) p(x_{n-1}, dx_n).$$

Nach dem Erweiterungssatz von Kolmogorov (Satz 2.17) existiert somit ein eindeutiges Maß  $\mathbb{P}$  auf  $\mathcal{F}$ , das diese endlichdimensionalen Verteilungen besitzt. Die gewünschte Markovkette sind nun die oben definierten Projektionen  $\xi_n$ . Dazu zeigen wir

**Schritt 1, die Anfangsverteilung:** Es gilt nach Definition von  $\mathbb{P}$ , dass  $\mathbb{P}(\xi_0 \in A) = \int_X 1_A(x_0) \nu(dx_0) = \nu(A)$ , also  $\mathbb{P}_{\xi_0} = \nu$ .

**Schritt 2, Berechnung von  $\mathbb{P}(\xi_{n+1} \in B | \xi_0, \dots, \xi_n)$ :** Seien  $C \in \mathcal{B}_{X^{n+1}}$  und  $B \in \mathcal{B}_X$ . Dann gilt

$$\begin{aligned} & \mathbb{P}(\xi_{n+1} \in B, (\xi_0, \dots, \xi_n) \in C) \\ &= \int_{\Omega} 1_C((\xi_0, \dots, \xi_n)(\omega)) 1_B(\xi_{n+1}(\omega)) \mathbb{P}(d\omega) \\ &= \int_X \nu(dx_0) \int_X p(x_0, dx_1) \cdots \int_X 1_C(x_0, \dots, x_n) 1_B(x_{n+1}) p(x_n, dx_{n+1}) \\ &= \int_X \nu(dx_0) \int_X p(x_0, dx_1) \cdots \int_X p(x_n, B) 1_C(x_0, \dots, x_n) p(x_{n-1}, dx_n) \\ &= \int_{[(\xi_0, \dots, \xi_n) \in C]} p(\xi_n(\omega), B) \mathbb{P}(d\omega), \end{aligned}$$

also nach Definition  $\mathbb{P}(\xi_{n+1} \in B | \xi_0, \dots, \xi_n)(\omega) = p(\xi_n(\omega), B)$   $\mathbb{P}$ -fast sicher.

**Schritt 3, Berechnung von  $\mathbb{P}(\xi_{n+1} \in B | \xi_n)$ :** Seien  $A, B \in \mathcal{B}_X$ . Dann gilt

$$\begin{aligned} \mathbb{P}(\xi_{n+1} \in B, \xi_n \in A) &= \int 1_B(\xi_{n+1}(\omega)) 1_A(\xi_n(\omega)) \mathbb{P}(d\omega) \\ &= \int_X \nu(dx_0) \int_X p(x_0, dx_1) \cdots \int_X 1_A(x_n) 1_B(x_{n+1}) p(x_n, dx_{n+1}) \\ &= \int_X \nu(dx_0) \int_X p(x_0, dx_1) \cdots \int_X 1_A(x_n) p(x_n, B) p(x_{n-1}, dx_n) \\ &= \int 1_A(\xi_n(\omega)) p(\xi_n(\omega), B) \mathbb{P}(d\omega) \\ &= \int_{[\xi_n \in A]} p(\xi_n(\omega), B) \mathbb{P}(d\omega), \end{aligned}$$

also  $\mathbb{P}(\xi_{n+1} \in B | \xi_n)(\omega) = p(\xi_n(\omega), B)$   $\mathbb{P}$ -fast sicher.

Die Kombination dieser Schritte zeigt also, dass  $(\xi_n)$  eine homogene Markovkette mit Anfangsverteilung  $\nu$  ist.  $\square$

**Beispiel 3.16.** Unser Interesse an Markovketten liegt primär an folgendem Beispiel: Sei  $(X; w_1, \dots, w_N)$  ein IFS mit Wahrscheinlichkeiten  $p_1, \dots, p_N$ , wobei  $X$  ein kompakter, metrischer Raum (beachte (Übung), dass ein kompakter metrischer Raum separabel und vollständig ist) sei. Wir definieren dann die Übergangswahrscheinlichkeit  $p(x, B)$  als

$$p(x, B) := \sum_{i=1}^N p_i 1_B(w_i(x)) = \sum_{i=1}^N p_i \delta_{w_i(x)}(B),$$

also ist  $p(x, \{w_i(x)\}) = p_i$ . Das bedeutet, wir starten an einem Punkt  $x \in X$ . Dann wählen wir gemäß den Wahrscheinlichkeiten  $p_i$  ein  $w_i$  aus und wenden dieses auf  $x$  an. Dieser Prozess wird iteriert. Diese Setzung von  $p$  erfüllt die Eigenschaften von Satz 3.15, wir erhalten also zu beliebig vorgegebener Startverteilung  $\nu$  von  $x$  eine Markovkette  $(\xi_n)$  auf dem Wahrscheinlichkeitsraum  $(X^{\mathbb{N}}, \mathcal{B}_{X^{\mathbb{N}}}, \mathbb{P}^\nu)$  mit den vorgegebenen Übergangswahrscheinlichkeiten  $p(x, B)$ .

**Definition 3.17.** Sei  $p(x, B)$  eine Übergangswahrscheinlichkeit (die Abbildung  $p$  erfülle also die beiden Voraussetzungen aus Satz 3.15). Sei weiters  $\mathcal{M}^1(X)$  der Raum der Wahrscheinlichkeitsmaße auf  $X$ . Der Markov-Operator  $M : \mathcal{M}^1(X) \rightarrow \mathcal{M}^1(X)$  ist definiert als

$$M(\mu)(B) := \int_X p(x, B) \mu(dx) \quad \text{für } \mu \in \mathcal{M}^1(X), B \in \mathcal{B}_X.$$

$\mu \in \mathcal{M}^1(X)$  heißt ein bezüglich  $p(x, B)$  invariantes Maß, falls

$$M(\mu) = \mu.$$

In Bezug auf eine Markovkette heißt  $\mu$  dann eine invariante Anfangsverteilung.

**Beispiel 3.18.** Sei  $p(x, B)$  die Übergangswahrscheinlichkeit aus Beispiel 3.16. Dann ist das Maß  $\mu$  aus Satz 2.43 die eindeutige invariante Anfangsverteilung der zu  $p(x, B)$  gehörigen homogenen Markovkette. Es gilt nämlich

$$M(\mu)(B) = \int_X \sum_{i=1}^N p_i \delta_{w_i(x)}(B) \mu(dx) = \sum_{i=1}^N p_i \mu(w_i \in B) = \sum_{i=1}^N p_i \mu_i(B)$$

in der Notation von Satz 2.43.

### 3.3 Ein Ergodensatz

**Definition 3.19.** Eine Folge  $(\xi_n)$  von Zufallsvariablen  $\xi_n : \Omega \rightarrow X$  heißt stationär, falls für alle  $k$   $(\xi_1, \xi_2, \dots)$  und  $(\xi_k, \xi_{k+1}, \dots)$  dieselbe Verteilung besitzen, falls also für alle  $k \in \mathbb{N}$  und für alle  $B \in \mathcal{B}_{X^{\mathbb{N}}}$  gilt

$$\mathbb{P}((\xi_1, \xi_2, \dots) \in B) = \mathbb{P}((\xi_k, \xi_{k+1}, \dots) \in B).$$

**Beispiel 3.20.** Jede unabhängige identisch verteilte Folge von Zufallsvariablen ist stationär. Für eine solche Folge gilt das Gesetz der großen Zahlen: Falls  $\mathbb{E}|\xi_1| < \infty$  und  $\mathbb{E}\xi_1 = m$ , dann gilt

$$\lim_{n \rightarrow \infty} \frac{\xi_1 + \dots + \xi_n}{n} = m \quad \mathbb{P}\text{-fast sicher.}$$



Wir wollen in diesem Abschnitt eine Verallgemeinerung des Gesetzes der großen Zahlen auf stationäre Folgen (den Birkhoffschen Ergodensatz) kennenlernen.

**Bemerkung 3.21.** (Übung!)

(i) Die Folge  $(\xi_n)$  ist genau dann stationär, falls  $(\xi_1, \xi_2, \dots)$  und  $(\xi_2, \xi_3, \dots)$  dieselbe Verteilung besitzen.

(ii) Alle Zufallsvariablen  $\xi_i$  haben bei einer stationären Folge dieselbe Verteilung.

**Definition 3.22.** Sei  $T : \Omega \rightarrow \Omega$  eine  $(\mathcal{F}, \mathcal{F})$ -messbare Abbildung.  $T$  heißt maßtreu, falls

$$\mathbb{P}(T^{-1}(A)) = \mathbb{P}(A) \quad \text{für alle } A \in \mathcal{F}$$

gilt.

**Proposition 3.23.** Sei der Maßraum  $(X^{\mathbb{N}}, \mathcal{B}_{X^{\mathbb{N}}}, \mathbb{P})$  gegeben mit den Projektionsabbildungen  $\xi_n((x_0, x_1, \dots)) = x_n$  und der Shifttransformation  $T((x_0, x_1, \dots)) := (x_1, x_2, \dots)$ . Dann ist die Folge  $(\xi_n)$  genau dann stationär, falls  $T$  maßtreu ist.

*Beweis.* Die Aussage ergibt sich aus den Tatsachen, dass für  $A \in \mathcal{B}_{X^{\mathbb{N}}}$  gilt

$$A = [(\xi_0, \xi_1, \dots) \in A], \quad T^{-1}(A) = [(\xi_1, \xi_2, \dots) \in A].$$

Setzen wir nämlich die Stationarität von  $(\xi_n)$  voraus, so gilt auch

$$\mathbb{P}(T^{-1}(A)) = \mathbb{P}[(\xi_1, \xi_2, \dots) \in A] = \mathbb{P}[(\xi_0, \xi_1, \dots) \in A] = \mathbb{P}(A),$$

also, dass  $T$  maßtreu ist. Ist andererseits  $T$  maßtreu, so erhalten wir

$$\mathbb{P}[(\xi_k, \xi_{k+1}, \dots) \in A] = \mathbb{P}(T^{-k}(A)) = \mathbb{P}(A) = \mathbb{P}[(\xi_0, \xi_1, \dots) \in A],$$

also die Stationarität von  $(\xi_n)$ . □

**Definition 3.24.** Sei  $(\Omega, \mathcal{F}, \mathbb{P})$  ein Wahrscheinlichkeitsraum. Ein Ereignis  $A \in \mathcal{F}$  heißt invariant (bzgl.  $T$ ), falls  $T^{-1}(A) = A$ . Eine maßtreue Abbildung  $T : \Omega \rightarrow \Omega$  heißt ergodisch (bzgl.  $\mathbb{P}$ ), falls für jedes invariante Ereignis  $A$  gilt, dass  $\mathbb{P}(A) \in \{0, 1\}$ .

Gleichmaßen definieren wir diese Begriffe für stationäre Folgen. Sei also die Folge von Zufallsvariablen  $\xi_i : \Omega \rightarrow X$  stationär. Eine Menge  $A = [(\xi_1, \xi_2, \dots) \in C]$  für  $C \in \mathcal{B}_{X^{\mathbb{N}}}$  heißt invariant, falls für alle  $k \geq 1$  gilt

$$A = [(\xi_k, \xi_{k+1}, \dots) \in C].$$

Weiters heißt diese stationäre Folge ergodisch, falls für jede invariante Menge  $A$  gilt, dass  $\mathbb{P}(A) \in \{0, 1\}$ .

**Bemerkung 3.25.** (Übung!) Seien der Wahrscheinlichkeitsraum, die Folge  $(\xi_n)$  und die Shiftabbildung  $T$  definiert wie in Proposition 3.23. Dann ist  $T$  genau dann ergodisch, falls  $(\xi_n)$  ergodisch ist.

**Satz 3.26.** Eine homogene Markovkette  $(\xi_n)$  ist genau dann stationär, falls  $\xi_0$  und  $\xi_1$  dieselbe Verteilung besitzen.

*Beweis.* Besitzen  $\xi_0$  und  $\xi_1$  dieselbe Verteilung, so gilt

$$\mathbb{P}_{\xi_0}(dx_1) = \mathbb{P}_{\xi_1}(dx_1) = \int_X p(x_0, dx_1) \mathbb{P}_{\xi_0}(dx_0). \quad (3.3)$$

Um zu zeigen, dass  $(\xi_0, \xi_1, \dots)$  und  $(\xi_1, \xi_2, \dots)$  dieselbe Verteilung besitzen, reicht es, dies auf den Zylindermengen  $B_0 \times B_1 \times \dots \times B_n \times X \times X \times \dots$  nachzuweisen. Einerseits erhalten wir aus der Definition

$$\begin{aligned} & \mathbb{P}(\xi_0 \in B_0, \dots, \xi_n \in B_n) \\ &= \int_X \int_X \dots \int_X 1_{B_0 \times \dots \times B_n}(x_0, \dots, x_n) p(x_{n-1}, dx_n) \dots p(x_0, dx_1) \mathbb{P}_{\xi_0}(dx_0), \end{aligned} \quad (3.4)$$

andererseits gilt mit (3.3)

$$\begin{aligned} & \mathbb{P}(\xi_1 \in B_0, \dots, \xi_{n+1} \in B_n) \\ &= \mathbb{P}(\xi_0 \in X, \xi_1 \in B_0, \dots, \xi_{n+1} \in B_n) \\ &= \int_X \int_X \dots \int_X 1_{X \times B_0 \times \dots \times B_n}(x_0, \dots, x_{n+1}) p(x_n, dx_{n+1}) \dots p(x_0, dx_1) \mathbb{P}_{\xi_0}(dx_0) \\ &= \int_X \int_X \dots \int_X 1_{B_0 \times \dots \times B_n}(x_1, \dots, x_{n+1}) p(x_n, dx_{n+1}) \dots p(x_1, dx_2) \mathbb{P}_{\xi_0}(dx_1) \end{aligned}$$

Nach Umbenennung der Variablen ist klar, dass dies identisch ist mit dem Ausdruck in (3.4), es folgt also die Behauptung des Satzes.  $\square$

Aus (3.3) und der Definition der Invarianz eines Maßes folgt

**Korollar 3.27.** *Eine homogene Markovkette  $(\xi_n)$  ist genau dann stationär, falls die Anfangsverteilung  $\mu$  ein invariantes Maß bezüglich den zur Markovkette gehörigen Übergangswahrscheinlichkeiten  $p(x, B)$  ist.*

Ohne Beweis erwähnen wir den später benötigten Satz

**Satz 3.28.** *Sei  $\mu$  die einzige invariante Anfangsverteilung einer homogenen Markovkette  $(\xi_n)$ . Dann ist dieser Prozess mit  $\mu$  als Anfangsverteilung ergodisch.*

Sei nun  $(\Omega, \mathcal{F}, \mathbb{P})$  ein Wahrscheinlichkeitsraum und  $T : \Omega \rightarrow \Omega$  eine messbare Abbildung. Sei dann  $\mathcal{J}$  die Familie der invarianten Mengen.  $\mathcal{J}$  ist dann eine  $\sigma$ -Algebra

**Satz 3.29** (Birkhoffscher Ergodensatz). *Sei  $(\Omega, \mathcal{F}, \mathbb{P})$  ein Wahrscheinlichkeitsraum und  $Z : \Omega \rightarrow \mathbb{R}$  eine Zufallsvariable mit  $\mathbb{E}|Z| < \infty$ . Ist  $T$  eine maßtreue Abbildung, so existiert der Limes*

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=0}^{k-1} Z(T^j \omega) \quad \mathbb{P}\text{-fast sicher.}$$

Falls zusätzlich  $T$  ergodisch ist, so gilt sogar

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=0}^{k-1} Z(T^j \omega) = \mathbb{E}Z \quad \mathbb{P}\text{-fast sicher.}$$

*Beweis. Schritt 1, Beweisidee:* Wir zeigen, dass mit

$$\overline{Z}(\omega) := \limsup_{k \rightarrow \infty} \frac{1}{k} \sum_{j=0}^{k-1} Z(T^j \omega) \quad \text{bzw.} \quad \underline{Z}(\omega) := \liminf_{k \rightarrow \infty} \frac{1}{k} \sum_{j=0}^{k-1} Z(T^j \omega)$$

für alle invarianten Mengen  $A \subset \Omega$  (d.h.  $T^{-1}(A) = A$ ) gilt

$$\int_A \overline{Z} d\mathbb{P} \leq \int_A Z d\mathbb{P} \leq \int_A \underline{Z} d\mathbb{P}.$$

Denn daraus folgt dann  $\underline{Z} = \overline{Z}$  fast sicher, der besagte Limes existiert also und er ist gleichzeitig (nach Definition) eine Version der bedingten Erwartung von  $Z$  unter der  $\sigma$ -Algebra  $\mathcal{I}$  der invarianten Mengen. Da für ergodische Abbildungen  $T$  jede invariante Menge entweder Maß 0 oder 1 hat, ist diese bedingte Erwartung fast sicher gleich dem Erwartungswert  $\mathbb{E}Z$ ; wir erhalten also in diesem Fall die zweite Behauptung

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=0}^{k-1} Z(T^j \omega) = \mathbb{E}Z \quad \mathbb{P}\text{-fast sicher.}$$

Wir zeigen im folgenden nur die erste Ungleichung (für invariante Mengen  $A$ )

$$\int_A \overline{Z} d\mathbb{P} \leq \int_A Z d\mathbb{P}$$

für positive Zufallsvariablen  $Z \geq 0$ . Die zweite Ungleichung mit  $\underline{Z}$  erhält man mit einem symmetrischen Argument und dann folgt der allgemeine Fall für  $Z$  integrierbar nach einer Aufspaltung von  $Z = Z_+ - Z_-$  in die Differenz zweier nichtnegativer Zufallsvariablen.

**Schritt 2,**  $\int_A \overline{Z} d\mathbb{P} \leq \int_A Z d\mathbb{P}$ :  $\overline{Z}$  ist (fast sicher) invariant (unter  $T$ ), da

$$\limsup_{k \rightarrow \infty} \frac{1}{k} \sum_{j=0}^{k-1} Z(T^{j+1} \omega) = \limsup_{k \rightarrow \infty} \frac{1}{k} \sum_{j=0}^{k-1} Z(T^j \omega) \quad \text{auf } [Z < \infty] \quad (3.5)$$

und  $\mathbb{P}([Z = \infty]) = 0$  aufgrund der Integrierbarkeit von  $Z$ . Fixieren wir nun eine Konstante  $M$  und ein  $\varepsilon > 0$ , so definieren wir

$$\overline{Z}_M := \overline{Z} \wedge M, \quad \tau(\omega) := \min\{k \geq 1 : \overline{Z}_M \leq \frac{1}{k} \sum_{j=0}^{k-1} Z(T^j \omega) + \varepsilon\}.$$

Da  $\overline{Z}$  invariant ist, ist es auch  $\overline{Z}_M$ , wir erhalten also aus (3.5)

$$\sum_{j=0}^{\tau(\omega)-1} \overline{Z}_M(T^j \omega) \leq \sum_{j=0}^{\tau(\omega)-1} Z(T^j \omega) + \tau(\omega)\varepsilon \quad \text{auf } [Z < \infty]. \quad (3.6)$$

Nach Definition ist  $\tau$  überall endlich, somit existiert ein  $N$ , sodass  $\mathbb{P}([\tau > N]) \leq \varepsilon/M$ . Basierend auf  $Z$  definieren wir nun

$$\tilde{Z}(\omega) := \begin{cases} Z(\omega) & \tau(\omega) \leq N, \\ Z(\omega) \vee M & \tau(\omega) > N. \end{cases}$$

Analog zu  $\tau$  setzen wir

$$\tilde{\tau}(\omega) := \min\{k \geq 1 : \overline{Z_M}(\omega) \leq \frac{1}{k} \sum_{j=0}^{k-1} \tilde{Z}(T^j \omega) + \varepsilon\}.$$

Der Vorteil dieser Hilfsgrößen besteht nun darin, dass  $\tilde{\tau}$  beschränkt ist durch  $N$  (weil  $\tilde{\tau} = \tau$  auf  $[\tau \leq N]$  und  $\tilde{\tau} = 1$  auf  $[\tau > N]$ ) und es gilt für  $\tilde{Z}$  und jede invariante Menge  $A$  die Abschätzung

$$\int_A \tilde{Z}(\omega) \mathbb{P}(d\omega) \leq \int_A Z(\omega) \mathbb{P}(d\omega) + \int_{A \cap [\tau > n]} M \mathbb{P}(d\omega) \leq \int_A Z(\omega) \mathbb{P}(d\omega) + \varepsilon. \quad (3.7)$$

Analog zu (3.6) gilt nun

$$\sum_{j=0}^{\tilde{\tau}(\omega)-1} \overline{Z_M}(T^j \omega) \leq \sum_{j=0}^{\tilde{\tau}(\omega)-1} \tilde{Z}(T^j \omega) + \tilde{\tau}(\omega) \varepsilon \quad \text{fast sicher.} \quad (3.8)$$

Um den Beweis abzuschließen, teilen wir die Summation in geeignete Blöcke: Sei  $L$  so groß, dass  $NM/L < \varepsilon$  und wir definieren induktiv

$$\tau_0(\omega) := 0, \quad \tau_k(\omega) := \tau_{k-1}(\omega) + \tilde{\tau}(T^{\tau_{k-1}(\omega)} \omega) \quad \text{für } k \geq 1.$$

Dann gilt

$$\sum_{j=0}^{L-1} \overline{Z_M}(T^j \omega) = \sum_{k=1}^{k(\omega)} \sum_{j=\tau_{k-1}(\omega)}^{\tau_k(\omega)-1} \overline{Z_M}(T^j \omega) + \sum_{j=\tau_k(\omega)}^{L-1} \overline{Z_M}(T^j \omega),$$

wobei  $k(\omega)$  die maximale natürliche Zahl ist, für die  $\tau_k(\omega) \leq L - 1$ . Aus dieser Aufspaltung, der Invarianz von  $\overline{Z_M}$ , (3.8) und der Positivität von  $\tilde{Z}$  erhalten wir die Abschätzung

$$\sum_{j=0}^{L-1} \overline{Z_M}(T^j \omega) \leq \sum_{j=0}^{L-1} \tilde{Z}(T^j \omega) + L\varepsilon + NM.$$

Dabei ist zu beachten, dass  $L - \tau_k(\omega) \leq N$ . Da  $A$  und  $\overline{Z_M}$  invariant sind, folgt daraus durch Integration und Division durch  $L$

$$\int_A \overline{Z_M} d\mathbb{P} \leq \int_A \tilde{Z} d\mathbb{P} + \varepsilon + NM/L \leq \int_A \tilde{Z} d\mathbb{P} + 2\varepsilon \leq \int_A Z d\mathbb{P} + 3\varepsilon.$$

Da diese Ungleichung für alle  $\varepsilon > 0$  gilt, erhalten wir nun aus dem Satz über die monotone Konvergenz die Behauptung

$$\int \overline{Z} d\mathbb{P} \leq \int Z d\mathbb{P}. \quad \square$$

### 3.4 Anwendungen

Für ein IFS  $\{X; w_1, \dots, w_N\}$  mit Wahrscheinlichkeiten  $0 < p_1, \dots, p_N < 1$  definieren wir ein Wahrscheinlichkeitsmaß  $\tilde{\mathbb{P}}$  auf  $D^{\mathbb{N}}$  mit  $D = \{1, \dots, N\}$  durch

$$\tilde{\mathbb{P}}(\{i_1\} \times \dots \times \{i_k\} \times D \times D \times \dots) := p_{i_1} \cdots p_{i_k}.$$

Dann gilt für die Übergangswahrscheinlichkeit von Beispiel 3.16 und die zugehörige Markovkette  $(\xi_n)$  mit beliebiger Anfangsverteilung  $\nu$ , dass

$$\mathbb{P}^\nu((\xi_0, \xi_1, \dots) \in B) = \int_X \mathbb{P}^{\delta_x}((\xi_0, \xi_1, \dots) \in B) \nu(dx) \quad (3.9)$$

und

$$\mathbb{P}^{\delta_x}((\xi_0, \xi_1, \dots) \in B) = \tilde{\mathbb{P}}((i_1, i_2, \dots) : (x, w_{i_1}(x), w_{i_2} \circ w_{i_1}(x), \dots) \in B) \quad (3.10)$$

**Satz 3.30** (Elton). *Seien  $(X, d)$  ein kompakter metrischer Raum,  $\{X; w_1, \dots, w_N\}$  ein IFS mit Wahrscheinlichkeiten  $0 < p_1, \dots, p_N < 1$  und  $\mu$  das (einzige) invariante Maß dieses IFS. Dann gibt es ein  $S \subset \{1, \dots, N\}^{\mathbb{N}}$  mit  $\tilde{\mathbb{P}}(S) = 1$ , sodass für alle Folgen  $(i_1, i_2, \dots) \in S$ , alle  $x_0 \in X$  und alle  $f \in C(X)$  gilt*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} f(w_{i_{k-1}} \circ \dots \circ w_{i_1}(x_0)) = \int_X f d\mu.$$

*Beweis.* Wir betrachten die Übergangswahrscheinlichkeiten  $p(x, B)$  von Beispiel 3.16 und die stationäre (siehe Satz 3.26), homogene Markovkette  $(\xi_n)$  mit der (eindeutigen, bzgl.  $p(x, B)$ ) invarianten Anfangsverteilung  $\mu$  auf dem Wahrscheinlichkeitsraum  $(X^{\mathbb{N}}, \mathcal{B}_{X^{\mathbb{N}}}, \mathbb{P}^\mu)$ . Die Shiftabbildung  $T : X^{\mathbb{N}} \rightarrow X^{\mathbb{N}}$ ,  $(x_0, x_1, \dots) \mapsto (x_1, x_2, \dots)$  ist maßtreu (dies folgt aus Proposition 3.23) und ergodisch (dies ist eine Konsequenz von Satz 3.28 und Bemerkung 3.25).

Sei nun  $f \in C(X)$  beliebig. Wir wenden den Ergodensatz 3.29 mit der maßtreuen und ergodischen Abbildung  $T$  und der Zufallsvariable  $Z = f \circ \xi_0$  an. Dann erhalten wir

$$\mathbb{P}^\mu \left( \left\{ (x_0, x_1, \dots) \in X^{\mathbb{N}} : \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} f(x_k) = \int_X f(x) \mu(dx) \right\} \right) = 1.$$

Setzen wir nun  $G(x_0, f) := \{(i_1, i_2, \dots) : \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} f(w_{i_{k-1}} \circ \dots \circ w_{i_1}(x_0)) = \int_X f d\mu\}$  auf die von  $x_0$  und  $f$  abhängige Menge, für die die Aussage des Satzes gilt, so erhalten wir aus (3.9) und (3.10), dass

$$\tilde{\mathbb{P}}(G(x_0, f)) = 1$$

für  $\mu$ -fast alle  $x_0 \in X$ . Nutzt man nun die Tatsachen aus, dass die  $w_i$  Kontraktionen sind und dass  $f$  stetig ist, so erhält man  $\tilde{\mathbb{P}}(G(x_0, f)) = 1$  für alle  $x_0 \in X$ . Die Separabilität von  $C(X)$  liefert uns schließlich die Behauptung, dass wir die Menge  $S \subset \{1, \dots, N\}^{\mathbb{N}}$  unabhängig von  $f$  wählen können.  $\square$

**Bemerkung 3.31.** Dieser Satz lässt sich folgendermaßen interpretieren: Für jeden Startpunkt  $x_0 \in X$  konvergiert fast jede Trajektorie  $(x_0, x_1, \dots)$  bei der Anwendung des Zufallsiterationsalgorithmus  $x_1 = w_i(x_0)$  mit Wahrscheinlichkeit  $p_i$ ,  $x_2 = w_j(w_i(x_0))$  mit Wahrscheinlichkeit  $p_i p_j$ , usw.

$$\frac{1}{n} \sum_{k=0}^{n-1} \int_X f d\delta_{x_k} \rightarrow \int_X f d\mu, \quad f \in C(X).$$

Es konvergiert also das Maß  $\frac{1}{n} \sum_{k=0}^{n-1} \delta_{x_k}$  *vag* gegen das invariante Maß  $\mu$ . Dabei heißt eine Folge von Wahrscheinlichkeitsmaßen  $\mu_n$  auf einem kompakten metrischen Raum  $X$  *vag konvergent* gegen  $\mu$ , falls für alle  $f \in C(X)$  gilt:  $\int_X f d\mu_n \rightarrow \int_X f d\mu$ . Man kann zeigen, dass folgende Aussagen äquivalent sind:

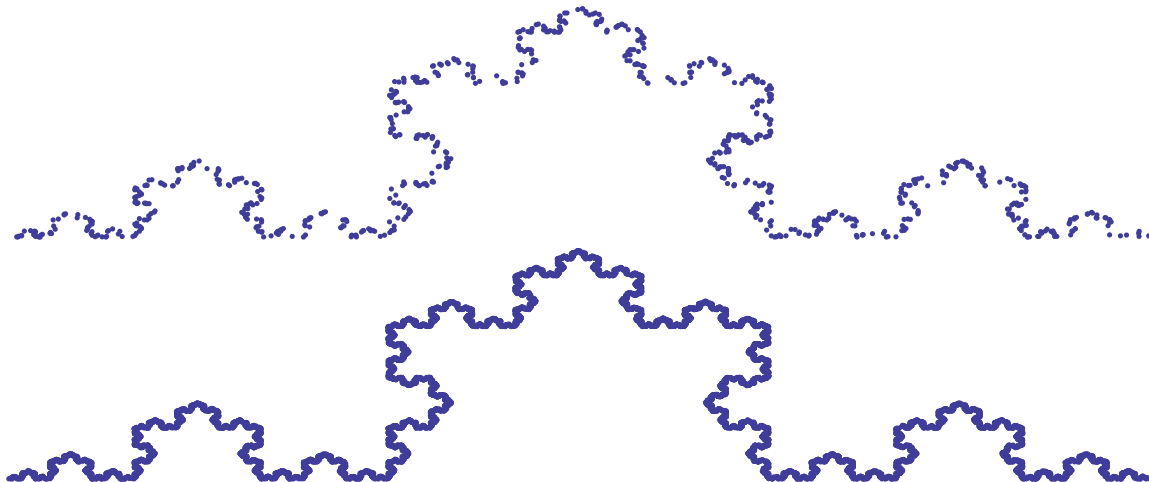
- (i)  $\mu_n \rightarrow \mu$  *vag*
- (ii)  $\limsup \mu_n(C) \leq \mu(C)$  für alle abgeschlossenen Teilmengen  $C$  von  $X$
- (iii)  $\liminf \mu_n(U) \geq \mu(U)$  für alle offenen Teilmengen  $U$  von  $X$
- (iv)  $\lim \mu_n(A) = \mu(A)$  für alle Borelmengen  $A$  mit  $\mu(\partial A) = 0$ .

Diese Bemerkung liefert nun ein Korollar zu Satz 3.30:

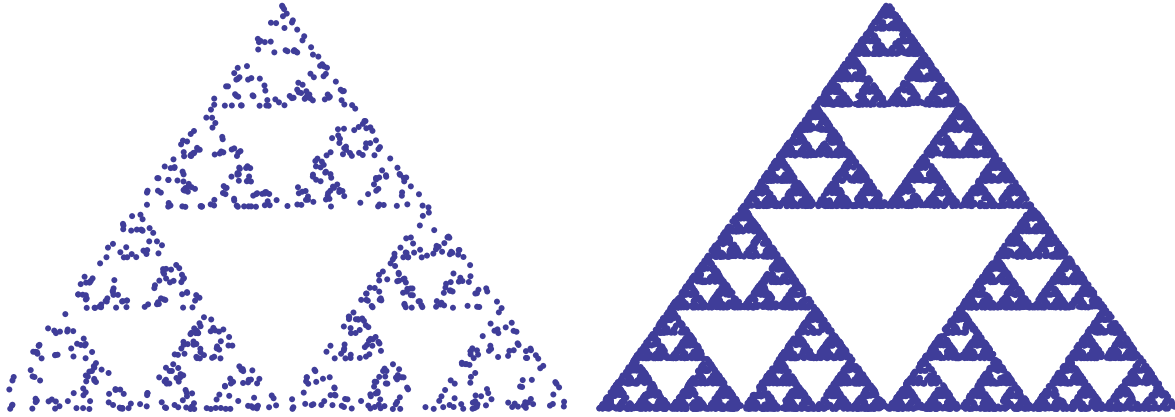
**Korollar 3.32.** Sei  $B$  eine Borelmenge in  $X$  mit  $\mu(\partial B) = 0$ . Dann gilt für  $\tilde{\mathbb{P}}$ -fast alle Folgen  $(i_1, i_2, \dots)$  in  $\{1, \dots, N\}^{\mathbb{N}}$ , dass die mittlere Verweilzeit der entsprechenden Trajektorie  $x_k = w_{i_k} \circ \dots \circ w_{i_1}(x_0)$  bei beliebigem Startpunkt  $x_0$  gegen  $\mu(B)$  konvergiert, es gilt also

$$\lim_{n \rightarrow \infty} \frac{|\{0 \leq i \leq n-1 : x_i \in B\}|}{n} = \mu(B).$$

**Beispiel 3.33** (Näherung der Koch-Schneeflocke). Betrachten wir das IFS von Beispiel 1.23 mit den Wahrscheinlichkeiten  $p_i = 1/4$  für  $1 \leq i \leq 4$  so erhalten wir mit 1000 bzw. 10000 Iterationspunkten und Entfernen der ersten 100 Iterationen die Abbildungen



**Beispiel 3.34** (Näherung des Sierpiński-Dreiecks). Betrachten wir das IFS von Beispiel 1.21 mit den Wahrscheinlichkeiten  $p_i = 1/3$  für  $1 \leq i \leq 3$  so erhalten wir mit 1000 bzw. 10000 Iterationspunkten und Entfernen der ersten 100 Iterationen die Abbildungen



**Das starke Gesetz der großen Zahlen mittels des Ergodensatzes** Um zu zeigen, dass das starke Gesetz der großen Zahlen aus dem Ergodensatz folgt, benötigen wir ein Resultat aus der Wahrscheinlichkeitsrechnung. Dazu folgende Definition

**Definition 3.35.** Sei  $(\Omega, \mathcal{F}, \mathbb{P})$  ein Wahrscheinlichkeitsraum und  $(\xi_n)$  eine Folge von Zufallsvariablen. Die  $\sigma$ -Algebra

$$\mathcal{F}_t := \bigcap_{n=1}^{\infty} \sigma(\xi_n, \xi_{n+1}, \dots)$$

heißt die  $\sigma$ -Algebra der terminalen Ereignisse.

**Satz 3.36** (Null-Eins-Gesetz von Kolmogorov). Sei  $(\xi_n)$  eine Folge von unabhängigen Zufallsvariablen und  $A$  ein bezüglich dieser Folge terminales Ereignis (dh.  $A \in \mathcal{F}_t$ ). Dann gilt  $\mathbb{P}(A) \in \{0, 1\}$ .

*Beweis.* Die Idee ist zu zeigen, dass  $A \in \mathcal{F}_t$  von sich selbst unabhängig ist, denn dann folgt aus  $\mathbb{P}(A)^2 = \mathbb{P}(A)$ , dass  $\mathbb{P}(A) \in \{0, 1\}$ . Sei  $A \in \mathcal{F}_t$ . Dann gilt auch

$$A \in \Sigma := \sigma(\xi_1, \xi_2, \dots) = \sigma\left(\bigcup_{n=1}^{\infty} \sigma(\xi_1, \xi_2, \dots, \xi_n)\right).$$

$\bigcup_{n=1}^{\infty} \sigma(\xi_1, \xi_2, \dots, \xi_n)$  ist eine Algebra und somit gilt nach Lemma 2.18: für alle  $\varepsilon > 0$  und alle  $A \in \Sigma$  existieren  $n \in \mathbb{N}$  und  $A_n \in \sigma(\xi_1, \dots, \xi_n)$  mit

$$\mathbb{P}(A \Delta A_n) < \varepsilon.$$

Daher gibt es eine Folge  $A_n \in \sigma(\xi_1, \dots, \xi_n)$  mit  $\mathbb{P}(A \Delta A_n) \rightarrow 0$ . Daraus erhalten wir

$$\begin{aligned} \mathbb{P}(A \setminus A_n) &= \mathbb{P}(A) - \mathbb{P}(A \cap A_n) \rightarrow 0, \\ \mathbb{P}(A) - \mathbb{P}(A_n) &= \mathbb{P}(A \setminus A_n) - \mathbb{P}(A_n \setminus A) \rightarrow 0. \end{aligned}$$

Da  $A$  und  $A_n$  unabhängig sind folgt also

$$\mathbb{P}(A) = \lim_{n \rightarrow \infty} \mathbb{P}(A \cap A_n) = \mathbb{P}(A) \lim_{n \rightarrow \infty} \mathbb{P}(A_n) = \mathbb{P}(A)^2,$$

und somit nach obigen Bemerkungen  $\mathbb{P}(A) \in \{0, 1\}$ . □

Sei  $\nu$  ein Wahrscheinlichkeitsmaß auf  $\mathbb{R}$ . Setzen wir in der Notation von Satz 3.15 für alle  $x \in \mathbb{R}$  und  $B \in \mathcal{B}_{\mathbb{R}}$  die Übergangswahrscheinlichkeit auf  $p(x, B) := \nu(B)$ , so erhalten wir eine unabhängige, identisch verteilte Folge von Zufallsvariablen  $(\xi_n)$  auf  $\mathbb{R}^{\mathbb{N}}$ , wobei  $\xi_n((x_1, x_2, \dots)) = x_n$  die Projektion auf die  $n$ -te Koordinate ist. Wie schon in Beispiel 3.20 bemerkt, ist  $(\xi_n)$  eine stationäre Folge. Nach Definition ist eine invariante Menge  $A$  bezüglich der Folge  $(\xi_n)$  in der terminalen  $\sigma$ -Algebra  $\mathcal{F}_t$  enthalten und hat somit Wahrscheinlichkeit 0 oder 1. Daraus folgt aber die Ergodizität der Folge  $(\xi_n)$ . Aus Bemerkung 3.25 und dem Ergodensatz erhalten wir nun mit der Shiftabbildung  $T$  und der Funktion  $f : \mathbb{R}^{\mathbb{N}} \rightarrow \mathbb{R}$ ,  $f((x_1, x_2, \dots)) := x_1$  das Korollar

**Korollar 3.37** (Starkes Gesetz der großen Zahlen). *Sei  $(\xi_n)$  eine Folge von unabhängigen identisch verteilten reellen Zufallsvariablen auf dem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, \mathbb{P})$  mit  $\mathbb{E}|\xi_1| < \infty$ . Dann gilt  $\mathbb{P}$ -fast sicher*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \xi_i(\omega) = \mathbb{E}\xi_1.$$

**Anwendungen in der Zahlentheorie** Wir setzen  $\Omega = [0, 1)$ ,  $\mathcal{F} = \mathcal{B}_{\Omega}$  und  $\lambda$  gleich dem Lebesguemaß auf  $\Omega$ . Wir definieren die dyadische Transformation  $T : \Omega \rightarrow \Omega$

$$T(x) := 2x \pmod{1}.$$

$T$  ist maßtreu und zusätzlich gilt

**Lemma 3.38.** *Die dyadische Transformation  $T$  ist ergodisch.*

*Beweis.* Sei  $A \subset \Omega$  gegeben mit  $A = T^{-1}(A)$ . Dann ist zu zeigen, dass  $\lambda(A) \in \{0, 1\}$ . Zunächst folgt aus  $A = T^{-1}(A)$ , dass  $x \in A \Leftrightarrow Tx \in A$ . Ist  $x' = x + 1/2 \pmod{1}$ , so gilt auch, dass  $Tx \in A \Leftrightarrow Tx' \in A$ , da nur die erste Stelle der binären Entwicklung von  $x = .x_1x_2x_3\cdots$  verschieden ist zur binären Entwicklung von  $x' = .x'_1x_2x_3\cdots$  und somit  $Tx = Tx' = .x_2x_3\cdots$ . Daraus folgt nun  $\lambda(A \cap E^c) = \lambda(A \cap E)$  für  $E = [0, 1/2)$  und wir erhalten

$$\lambda(A) = 2 \cdot \lambda(A \cap E) = \frac{\lambda(A \cap E)}{\lambda(E)},$$

also die Unabhängigkeit von  $A$  und  $E$ . Diese Unabhängigkeit gilt nun auch für endliche Vereinigungen  $E$  von dyadischen Intervallen (Übung!). Sei nun  $\varepsilon > 0$  beliebig und  $E$  eine solche endliche Vereinigung von dyadischen Intervallen mit der Eigenschaft

$$\lambda(A \Delta E) < \varepsilon. \quad (\text{Die Existenz eines solchen } E \text{ folgt aus Lemma 2.18})$$

Dann gelten insbesondere  $|\lambda(A) - \lambda(E)| < \varepsilon$  und  $|\lambda(A) - \lambda(A)\lambda(E)| = |\lambda(A) - \lambda(A \cap E)| < \varepsilon$ . Daraus schließen wir

$$|\lambda(A) - \lambda(A)^2| \leq |\lambda(A) - \lambda(A)\lambda(E)| + |\lambda(A)\lambda(E) - \lambda(A)^2| < 2\varepsilon.$$

Da diese Ungleichung für alle  $\varepsilon > 0$  gilt, folgt also  $\lambda(A) = \lambda(A)^2$  und daraus schließlich  $\lambda(A) \in \{0, 1\}$ .  $\square$



**Definition 3.39.** Eine Zahl  $x \in [0, 1)$  heißt normal zur Basis  $b$ , falls die relative Häufigkeit von  $j$ , ( $0 \leq j \leq b-1$ ) in den ersten  $k$  Stellen der  $b$ -adischen Entwicklung von  $x$  gegen  $1/b$  konvergiert für  $k \rightarrow \infty$ . Also formal genau dann, falls

$$\lim_{k \rightarrow \infty} \frac{n_j(x|_k)}{k} = \frac{1}{b}, \quad j \in \{0, \dots, b-1\}$$

wobei  $n_j(x|_k)$  die Anzahl der Vorkommnisse der Ziffer  $j$  in der Entwicklung von  $x$  zur Basis  $b$  in den ersten  $k$  Stellen bezeichnet.

**Beispiel 3.40.** Ein Beispiel einer normalen Zahl zur Basis 10 ist die *Champernowne-Zahl*

$$.123456789101112131415 \dots,$$

die durch Aneinanderreihung aller natürlichen Zahlen gebildet wird. Ein weiteres Beispiel stellt die *Copeland-Erdős-Zahl*

$$.235791113171923293137 \dots$$

dar, die durch Aneinanderreihung aller Primzahlen entsteht.

Aus dem Ergodensatz folgt nun, dass fast alle Zahlen normal (zu jeder Basis) sind.

**Korollar 3.41.** Fast alle Zahlen bezüglich des Lebesgue-Maßes in  $[0, 1]$  sind normal zur Basis 2.

*Beweis.* Wir wenden den Ergodensatz mit der dyadischen Transformation  $Tx = 2x \pmod{1}$  (diese erfüllt die Voraussetzungen; siehe Lemma 3.38) und  $f = 1_{[0,1/2)}$  an. Dann erhalten wir mit

$$\lim_{k \rightarrow \infty} \frac{n_0(x|_k)}{k} = \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=0}^{k-1} f(T^j x) = \int_0^1 f d\lambda = \frac{1}{2} \quad \lambda\text{-fast sicher}$$

die Behauptung. □

Dieses Argument lässt sich nun leicht auf alle Basen  $b \geq 2$  ausdehnen, indem man statt  $T$  die Abbildung  $x \mapsto bx \pmod{1}$  betrachtet. Wir erhalten somit für alle Basen  $b$  eine Menge  $G_b \subset [0, 1)$  mit  $\lambda(G_b) = 1$ , sodass alle  $x \in G_b$  normal zur Basis  $b$  sind. Schneiden wir nun über alle  $b$ , so ist  $G = \bigcap_{b \in \mathbb{N}, b \geq 2} G_b$  eine Menge mit Lebesguemaß 1, sodass alle  $x \in G$  normal zu jeder Basis sind.

## 4 Weitere Fraktale

### 4.1 Spezielle Teilmengen von $[0, 1)$

Wir diskutierten im vorigen Abschnitt, dass die Menge der normalen Zahlen im Einheitsintervall  $[0, 1]$  Länge 1 besitzen. Alternativ dazu kann man sich auch für eine fixe Basis  $b \in \mathbb{N}$ ,  $b \geq 2$  die Mengen

$$F(p_0, \dots, p_{b-1}) := \left\{ x \in [0, 1) : \lim_{k \rightarrow \infty} \frac{n_j(x|_k)}{k} = p_j \text{ für alle } 0 \leq j \leq b-1 \right\}$$

ansetzen. Dabei sind  $0 \leq p_j \leq 1$  mit  $\sum_{j=0}^{b-1} p_j = 1$ . Die anschließende Proposition enthält eine Formel für die Hausdorffdimension dieser Mengen.

**Proposition 4.1.** Sei  $F = F(p_0, \dots, p_{b-1})$ . Dann gilt

$$\dim F = -\frac{1}{\log b} \sum_{j=0}^{b-1} p_j \log p_j.$$

*Beweis.* Sei  $I_{i_1 \dots i_k}$  das Intervall der Länge  $b^{-k}$  der Zahlen  $x \in [0, 1)$ , deren  $b$ -adische Entwicklung mit  $i_1 \dots i_k$  beginnt. Dabei gilt natürlich  $0 \leq i_j \leq b-1$ . Wir definieren ein Wahrscheinlichkeitsmaß auf  $([0, 1), \mathcal{B}_{[0,1)})$  durch

$$\mathbb{P}(I_{i_1 \dots i_k}) := p_{i_1} \cdots p_{i_k}. \quad (4.1)$$

Sei  $\xi_k(x) : [0, 1) \rightarrow \{0, \dots, b-1\}$  die  $k$ -te Ziffer der  $b$ -adischen Entwicklung von  $x$ . Dann ist  $(\xi_k)_k$  eine Folge von unabhängigen Zufallsvariablen und für fixes  $j \in \{0, \dots, b-1\}$  sei  $Z_k(x) := 1_{[\xi_k=j]}(x)$ . Dann sind die  $Z_k$  ebenfalls unabhängig und identisch verteilt und aus dem starken Gesetz der großen Zahlen (Korollar 3.37) folgt

$$\lim_{k \rightarrow \infty} \frac{n_j(x|_k)}{k} = \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{k=1}^n Z_k = \mathbb{E}Z_1 = p_j \quad \mathbb{P}\text{-fast sicher,}$$

also insbesondere  $\mathbb{P}(F) = 1$ . Sei nun  $I_k(x)$  das Intervall  $I_{i_1 \dots i_k}$  der Länge  $b^{-k}$ , sodass  $x \in I_{i_1 \dots i_k}$ . Für fixes  $y \in [0, 1)$  folgt aus (4.1) nach Logarithmieren

$$\log \mathbb{P}(I_k(y)) = \sum_{j=0}^{b-1} n_j(y|_k) \log p_j.$$

Für  $y \in F$  gilt außerdem nach Definition  $\lim_{k \rightarrow \infty} n_j(y|_k)/k = p_j$  für alle  $j$ , es ergibt sich also für  $y \in F$

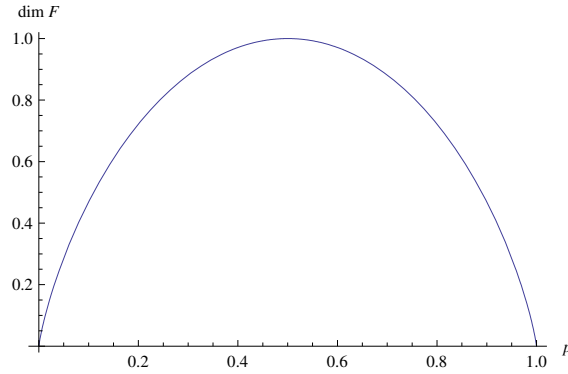
$$\frac{1}{k} \log \frac{\mathbb{P}(I_k(y))}{|I_k(y)|^s} = \frac{1}{k} \log \mathbb{P}(I_k(y)) - \frac{1}{k} \log b^{-ks} \rightarrow \sum_{j=0}^{b-1} p_j \log p_j + s \log b$$

falls  $k \rightarrow \infty$ . Somit ist mit  $\theta = -\frac{1}{\log b} \sum_{j=0}^{b-1} p_j \log p_j$

$$\lim_{k \rightarrow \infty} \frac{\mathbb{P}(I_k(y))}{|I_k(y)|^s} = \begin{cases} 0, & \text{für } s < \theta, \\ \infty, & \text{für } s > \theta. \end{cases}$$

Wir sind nun in der Situation von Satz 2.39. Der Beweis überträgt sich wörtlich, falls wir  $\frac{\mu(B(x,r))}{r^s}$  durch  $\frac{\mathbb{P}(I_k(y))}{|I_k(y)|^s}$  ersetzen. Es ergibt sich also  $\mathcal{H}^s(F) = \infty$  für  $s < \theta$  und  $\mathcal{H}^s(F) = 0$  für  $s > \theta$  und somit  $\dim F = \theta$ .  $\square$

**Bemerkung 4.2.** Wir betrachten den Spezialfall  $b = 2$  und setzen  $p = p_0$ ,  $1 - p = p_1$ . Wenn wir die Hausdorffdimension  $\dim F = -\frac{p \log p + (1-p) \log(1-p)}{\log 2}$  von  $F$  in Abhängigkeit von  $p$  plotten, erhalten wir



Diese Funktion ist klarerweise symmetrisch um  $1/2$  und je weiter sich  $p$  von  $1/2$  entfernt, desto kleiner wird die Hausdorffdimension von  $F$ .

## 4.2 Brownsche Bewegung

Sei  $(\Omega, \mathcal{F}, \mathbb{P})$  ein Wahrscheinlichkeitsraum. Eine Funktion  $W : \Omega \times [0, \infty) \rightarrow \mathbb{R}^d$  ist eine ( $\mathbb{R}^d$ -wertige) *Brownsche Bewegung* oder *Wiener Prozess*, falls  $W$  die folgenden drei Eigenschaften besitzt:

- (i) Auf einer Menge mit Wahrscheinlichkeit 1 gelten  $W(\omega, 0) = 0$  und  $t \mapsto W(\omega, t)$  ist eine stetige Funktion.
- (ii) Für alle  $t \geq 0$  und alle  $h > 0$  ist der Zuwachs  $W(\cdot, t+h) - W(\cdot, t)$  normalverteilt mit Mittelwert 0 und Varianz  $h$ , es gilt also

$$\mathbb{P}(W(\cdot, t+h) - W(\cdot, t) \in A) = (2\pi h)^{-d/2} \int_A \exp\left(-\frac{|u|^2}{2h}\right) du.$$

- (iii) Für  $0 \leq t_1 \leq t_2 \leq \dots \leq t_{2m}$  sind die Zuwächse  $W(\cdot, t_1), W(\cdot, t_2) - W(\cdot, t_1), \dots, W(\cdot, t_{2m}) - W(\cdot, t_{2m-1})$  unabhängig.

Nach Eigenschaft 2. sind die Zuwächse unabhängig von  $t$  verteilt, sie sind also *stationär*. Es ist eine nichttriviale Tatsache, dass eine Brownsche Bewegung existiert. Der natürliche Wahrscheinlichkeitsraum auf dem eine Brownsche Bewegung lebt ist der Raum stetigen Funktionen von  $[0, \infty)$  nach  $\mathbb{R}^d$ ,  $\Omega = C([0, \infty), \mathbb{R}^d)$  (der sogenannte Wiener-Raum) und einem Maß  $\mathbb{P}$  (dem Wiener-Maß), sodass obige Eigenschaften erfüllt sind. Dies ist der Fall, da die Brownsche Bewegung fast sicher stetige Pfade besitzt. In diesem Wahrscheinlichkeitsraum ist dann für eine stetige Funktion  $\omega$ :  $W(\omega, t) = \omega(t)$ . Der Einfachheit halber werden wir in Zukunft den Parameter  $\omega$  auslassen und von den Zufallsvariablen  $W(t)$  sprechen. Wir bemerken noch, dass Eigenschaft 2. insbesondere nach Übergang zu Polarkoordinaten die Gleichung

$$\mathbb{P}(|W(t+h) - W(t)| \leq \rho) = c h^{-d/2} \int_0^\rho r^{n-1} \exp\left(-\frac{r^2}{2h}\right) dr$$

mit einer Konstanten  $c$ , die nur von  $n$  abhängt, liefert.

Eine grundlegende Eigenschaft der Brownschen Bewegung ist deren lokale Hölder-Stetigkeit für alle Exponenten  $\lambda < 1/2$ .

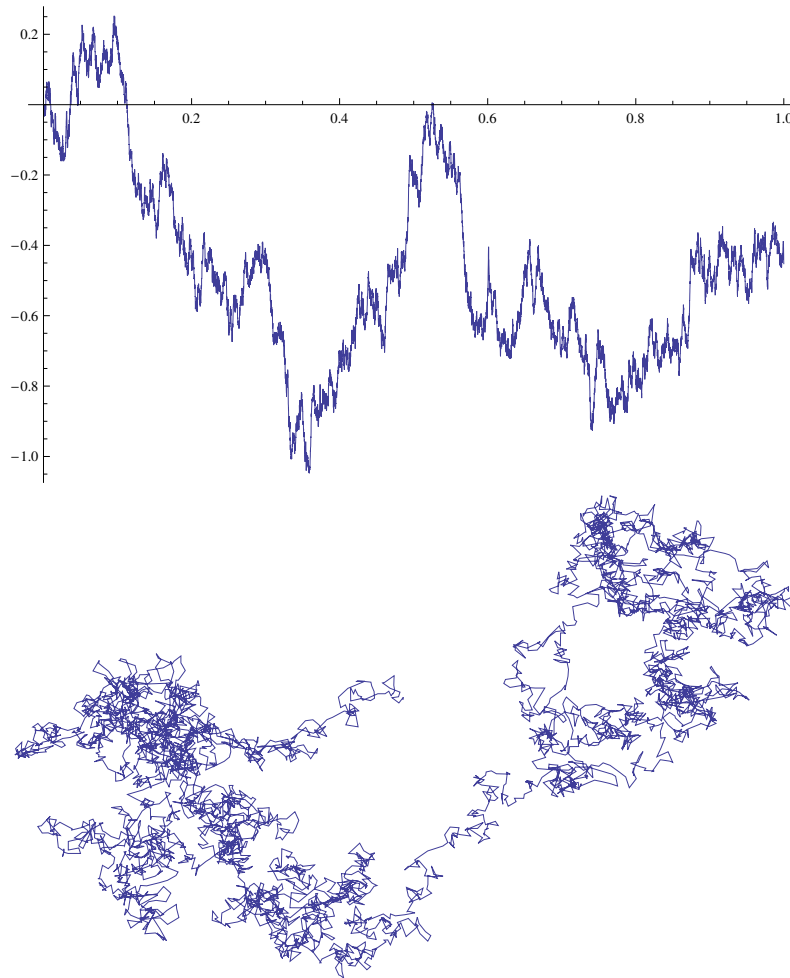


Abbildung 8: Simulation eines Graphs einer Brownschen Bewegung und einer 2-dimensionalen Brownschen Bewegung.

**Proposition 4.3.** Sei  $0 < \lambda < 1/2$ . Mit Wahrscheinlichkeit 1 erfüllt die Brownsche Bewegung die Hölderbedingung

$$|W(t+h) - W(t)| \leq b|h|^\lambda \quad \text{für } |h| < H_0$$

für ein  $H_0 > 0$ , wobei  $b$  nur von  $\lambda$  abhängt.

*Beweis.* Übung in mehreren Schritten:

**Schritt 1:** Benutze Polarkoordinaten, um für alle natürlichen Zahlen  $i$  zu zeigen, dass es eine Konstante  $c_i$ , die nur von  $i$  abhängt gibt mit

$$\mathbb{E}|W(t+h) - W(t)|^{2i} = c_i h^i \quad \text{für alle } h \geq 0.$$

**Schritt 2:** Zeige das Borel-Cantelli Lemma: Für eine Folge von Ereignissen  $(A_n)$  gilt

$$\sum_{n=0}^{\infty} \mathbb{P}(A_n) < \infty \implies \mathbb{P}(\limsup A_n) = 0,$$

wobei  $\limsup A_n = \bigcap_{n=0}^{\infty} \bigcup_{k \geq n} A_k$  (beachte, dass  $\omega \in \limsup A_n$  genau dann wenn  $\omega \in A_n$  für unendlich viele  $n$ ).

**Schritt 3:** Definiere  $A_n = \bigcup_{m=1}^{2^n} [|W((m-1)2^{-n}) - W(m2^{-n})| > 2^{-n\lambda}]$ . Benutze die Chebyshev Ungleichung und Schritt 1, um zu zeigen, dass  $\mathbb{P}(A_n) \leq 2^{-\gamma n}$  für ein geeignetes  $\gamma > 0$  und folgere  $\sum_{n=0}^{\infty} \mathbb{P}(A_n) < \infty$

**Schritt 4:** Mit dem Borel Cantelli Lemma folgt nun die Existenz einer Konstante  $K = K(\omega)$ , sodass mit Wahrscheinlichkeit 1 gilt

$$|W((m-1)2^{-n}) - W(m2^{-n})| \leq K2^{-n\lambda} \quad \text{für alle } n \geq 0.$$

**Schritt 5:** Nun muss man noch die Stetigkeit der Brownschen Bewegung und eine Zerlegung von  $[t, t+h]$  in geeignete dyadische Intervalle ausnutzen, um diese Ungleichung für alle Parameterwerte beweisen zu können.  $\square$

Wir berechnen nun die Hausdorffdimension einer Brownschen Bewegung und eines Graphen der Brownschen Bewegung. Es gilt

**Satz 4.4.** *Mit Wahrscheinlichkeit 1 hat ein Pfad der Brownschen Bewegung in  $\mathbb{R}^d (d \geq 2)$  Hausdorffdimension 2.*

*Beweis.* Für alle  $\lambda < 1/2$  erfüllt die Brownsche Bewegung nach der letzten Proposition eine Hölderbedingung mit Wahrscheinlichkeit 1, es folgt also aus Proposition 2.26 die obere Schranke  $\dim W([0, 1]) \leq \lambda^{-1} \dim[0, 1] = \lambda^{-1}$ . Also hat die Brownsche Bewegung höchstens Hausdorffdimension 2.

Für die untere Schranke benutzen wir die potentialtheoretische Methode von Satz 2.41. Sei  $1 < s < 2$ . Analog zu Schritt 1 des Beweises der vorigen Proposition gilt

$$\mathbb{E}|W(t+h) - W(t)|^{-s} = c_s h^{-s/2},$$

wobei  $c_s$  eine Konstante ist, die nur von  $s$  abhängt. Da  $s < 2$  folgt dann die Endlichkeit von

$$\begin{aligned} \mathbb{E} \int_0^1 \int_0^1 |W(t) - W(u)|^{-s} dt du &= \int_0^1 \int_0^1 \mathbb{E}|W(t) - W(u)|^{-s} dt du \\ &= \int_0^1 \int_0^1 c_s |t - u|^{-s/2} dt du. \end{aligned} \quad (4.2)$$

Wir definieren nun für einen fixen Pfad  $W(\omega, \cdot)$  ein Wahrscheinlichkeitsmaß darauf:

$$\mu^\omega(A) := \lambda(\{0 \leq t \leq 1 : \omega(t) \in A\}) = \lambda_\omega(A),$$

wobei  $\lambda$  das Lebesguemaß bezeichnet. Dann gilt  $\int g(x) d\mu^\omega(x) = \int_0^1 g(\omega(t)) dt$  für alle Funktionen  $g$ . Somit wird die Endlichkeit von (4.2) zur Endlichkeit von

$$\int_{\Omega} \int \int |x - y|^{-s} d\mu^\omega(x) d\mu^\omega(y) d\mathbb{P}(\omega).$$

Also gilt  $\int \int |x - y|^{-s} d\mu^\omega(x) d\mu^\omega(y) < \infty$   $\mathbb{P}$ -fast sicher, wobei  $\mu^\omega$  ein Wahrscheinlichkeitsmaß auf  $\omega([0, 1])$  ist. Es folgt also aus dem ersten Teil von Satz 2.41 die untere Abschätzung  $\dim W([0, 1]) \geq 2$   $\mathbb{P}$ -fast sicher.  $\square$

**Bemerkung 4.5.** Man kann zeigen, dass die Pfade der Brownschen Bewegung in  $\mathbb{R}^d$  für  $d \geq 2$  2-dimensionales Hausdorffmaß 0 besitzen. Genauer haben diese Pfade positives endliches Maß bezüglich der Funktion  $h(t) = t^2 \log(1/t) \log \log \log(1/t)$  für  $d = 2$  und bezüglich  $h(t) = t^2 \log \log(1/t)$  für  $d \geq 3$  (vergleiche Bemerkung 2.36). In diesem Sinne hat also die Brownsche Bewegung eine Dimension, die „logarithmisch kleiner“ ist als 2.

Um die Dimension eines Graphs der Brownschen Bewegung zu bestimmen, benötigen wir noch ein kleines Lemma. Vorher erinnern wir an die Definition eines Graphen einer Funktion: Ist  $A$  eine Menge und  $f : A \rightarrow B$  eine Funktion, so ist die Menge

$$\text{gr } f := \{(x, f(x)) : x \in A\} \subset A \times B$$

der *Graph* der Funktion  $f$ .

**Lemma 4.6.** Sei  $f : [0, 1] \rightarrow \mathbb{R}$  eine stetige Funktion. Es gelte die Abschätzung

$$|f(t) - f(u)| \leq c|t - u|^{2-s}, \quad t, u \in [0, 1],$$

für eine Konstante  $c > 0$  und ein  $1 \leq s \leq 2$ . Dann ist  $\mathcal{H}^s(\text{gr } f) < \infty$  und folglich  $\dim \text{gr } f \leq s$ .

*Beweis.* Übung: Fixiere  $\delta > 0$  und überdecke den Graph mit Würfeln der Seitenlänge  $\delta$ . Wieviel werden dafür benötigt?  $\square$

**Satz 4.7.** Mit Wahrscheinlichkeit 1 hat der Graph einer Brownschen Bewegung  $W(\omega, \cdot) : [0, 1] \rightarrow \mathbb{R}$  Hausdorffdimension  $1\frac{1}{2}$ .

*Beweis.* Nach dem vorigen Lemma und der Hölderbedingung der Brownschen Bewegung hat der Graph höchstens Hausdorffdimension  $2 - \lambda$  für alle  $\lambda < 1/2$ , also höchstens Dimension  $1\frac{1}{2}$ .

Für die untere Abschätzung führen wir eine ähnliche Rechnung wie im letzten Satz durch:

$$\begin{aligned} \mathbb{E}((|W(t+h) - W(t)|^2 + h^2)^{-s/2}) &= ch^{-1/2} \int_0^\infty (r^2 + h^2)^{-s/2} \exp\left(-\frac{r^2}{2h}\right) dr \\ &= \frac{1}{2}c \int_0^\infty (uh + h^2)^{-s/2} u^{-1/2} \exp\left(-\frac{u}{2}\right) du. \end{aligned}$$

Wenn wir dieses Integral aufspalten in die Integration von 0 bis  $h$  und in die von  $h$  bis  $\infty$ , so können wir das nach oben abschätzen mit

$$\frac{1}{2}c \int_0^h (h^2)^{-s/2} u^{-1/2} du + \frac{1}{2}c \int_h^\infty (uh)^{-s/2} u^{-1/2} \leq c_1 h^{1/2-s} \quad (4.3)$$

mit einer Konstanten  $c_1$  für alle  $h < 1$ . Wir benutzen nun wieder potentialtheoretische Methoden für die untere Abschätzung der Hausdorffdimension und definieren ein Wahrscheinlichkeitsmaß auf einem Graphen der Brownschen Bewegung

$$\mu^\omega(A) = \lambda(\{0 \leq t \leq 1 : (t, \omega(t)) \in A\}),$$