Ein Skript für Ergodentheorie

Chris Preston

2003

Inhaltsverzeichnis

1 Maßtheoretische Vorbereitungen **3**

2 Der Birkhoffsche Ergodensatz **7**

3 Stationäre Folgen **14**

4 Ergodische und mischende Abbildungen **21**

5 Endliche Markov-Ketten **30**

6 Kompakte Gruppen **40**

7 Isomorphie von maßtreuen Abbildungen **50**

8 Entropie **55**

9 Subadditive und multiplikative Ergodensätze **80**

10 Topologische Dynamische Systeme **90**

11 Expansive Abbildungen **100**

12 Markov-Partitionen **105**

13 Die Pseudo-Orbit tracing Eigenschaft **113**

Literatur **119**

Index **120**

1 Maßtheoretische Vorbereitungen

Ein *Messraum* ist ein Paar $(X, \mathcal{F})$ bestehend aus einer nichtleeren Menge X und einer σ-Algebra $\mathcal{F} \subset \mathcal{P}(X)$. Sei $(X, \mathcal{F})$ ein Messraum; ein Maß μ auf $\mathcal{F}$ heißt *Wahrscheinlichkeitsmaß*, wenn $\mu(X) = 1$. Die Menge aller Wahrscheinlichkeitsmaße auf $\mathcal{F}$ wird mit $\mathrm{P}(X, \mathcal{F})$ bezeichnet. Ein *Wahrscheinlichkeitsraum* ist ein Tripel $(X, \mathcal{F}, \mu)$ mit $(X, \mathcal{F})$ ein Messraum und $\mu \in \mathrm{P}(X, \mathcal{F})$.

Sei $(X, \mathcal{F}, \mu)$ ein Wahrscheinlichkeitsraum und sei $p : X \to \mathbb{B} = \{\mathrm{T}, \mathrm{F}\}$ eine Aussage. Man sagt dann, dass p *μ-fast sicher* gilt, wenn es $N \in \mathcal{F}$ mit $\mu(N) = 0$ gibt, so dass $\{x \in X : p(x) = \mathrm{F}\} \subset N$.

Sei X eine Menge; eine Teilmenge $\mathcal{E}$ von $\mathcal{P}(X)$ heißt d-*System*, wenn gilt:

(1) $X \in \mathcal{E}$,

(2) $E_2 \setminus E_1 \in \mathcal{E}$ für alle E_1, $E_2 \in \mathcal{E}$ mit $E_1 \subset E_2$.

Lemma 1.1 *Für jede Teilmenge $\mathcal{S}$ von $\mathcal{P}(X)$ gibt es ein kleinstes* d-*System* $\mathrm{d}(\mathcal{S})$, *das $\mathcal{S}$ enthält. (Mit anderen Worten:* $\mathrm{d}(\mathcal{S})$ *ist ein* d-*System mit $S \subset \mathrm{d}(\mathcal{S})$ und* $\mathrm{d}(\mathcal{S}) \subset \mathcal{B}$ *für jedes* d-*System $\mathcal{B}$ mit $\mathcal{S} \subset \mathcal{B}$.)*

Beweis Der Durchschnitt einer beliebigen Familie von d-Systemen ist wieder ein d-System; ferner ist $\mathcal{P}(X)$ ein d-System, das $\mathcal{S}$ enthält. Sei also $\mathrm{d}(\mathcal{S})$ der Durchschnitt aller d-Systemen, die $\mathcal{S}$ enthalten. Dann ist $\mathrm{d}(\mathcal{S})$ ein d-System mit $\mathcal{S} \subset \mathrm{d}(\mathcal{S})$ und $\mathrm{d}(\mathcal{S}) \subset \mathcal{B}$ für jedes d-System $\mathcal{B}$ mit $\mathcal{S} \subset \mathcal{B}$. $\square$

Eine Teilmenge $\mathcal{S} \subset \mathcal{P}(X)$ heißt *durchschnittsstabil*, wenn $A_1 \cap A_2 \in \mathcal{S}$ für alle A_1, $A_2 \in \mathcal{S}$.

Lemma 1.2 *Für jede durchschnittsstabile Teilmenge $\mathcal{S} \subset \mathcal{P}(X)$ ist* $\mathrm{d}(\mathcal{S})$ *eine Algebra.*

Beweis Da $X \in \mathrm{d}(\mathcal{S})$, ist $\mathrm{d}(\mathcal{S}) \neq \varnothing$, und es ist klar dass $X \setminus A \in \mathrm{d}(\mathcal{S})$ für alle $A \in \mathrm{d}(\mathcal{S})$. Sei $\mathcal{C} = \{A \in \mathrm{d}(\mathcal{S}) : A \cap B \in \mathrm{d}(\mathcal{S})$ für alle $B \in \mathcal{S}\}$; dann ist $\mathcal{S} \subset \mathcal{C} \subset \mathrm{d}(\mathcal{S})$ und man sieht leicht, dass $\mathcal{C}$ ein d-System ist. Damit ist $\mathcal{C} = \mathrm{d}(\mathcal{S})$, d.h., $A \cap B \in \mathrm{d}(\mathcal{S})$ für alle $A \in \mathrm{d}(\mathcal{S})$, $B \in \mathcal{S}$. Setze nun

$$\mathcal{D} = \{A \in \mathrm{d}(\mathcal{S}) : A \cap B \in \mathrm{d}(\mathcal{S}) \text{ für alle } B \in \mathrm{d}(\mathcal{S})\} ;$$

dann ist $\mathcal{S} \subset \mathcal{D} \subset \mathrm{d}(\mathcal{S})$ und wieder ist $\mathcal{D}$ ein d-System. Damit ist $\mathcal{D} = \mathrm{d}(\mathcal{S})$, d.h., $A \cap B \in \mathrm{d}(\mathcal{S})$ für alle A, $B \in \mathrm{d}(\mathcal{S})$. Dies zeigt, dass $\mathrm{d}(\mathcal{S})$ eine Algebra ist. $\square$

Satz 1.1 *Sei* $(X, \mathcal{F})$ *ein Messraum und sei* $\mathcal{S}$ *eine durchschnittsstabile Teilmenge von* $\mathcal{P}(X)$ *mit* $\sigma(\mathcal{S}) = \mathcal{F}$. *Sind* μ_1, $\mu_2 \in \mathrm{P}(X, \mathcal{F})$ *Wahrscheinlichkeitsmaße mit* $\mu_1(A) = \mu_2(A)$ *für alle* A_1, $A_2 \in \mathcal{S}$, *so ist* $\mu_1 = \mu_2$.

Beweis Setze $\mathcal{A} = \{B \in \mathcal{F} : \mu_1(B) = \mu_2(B)\}$; es ist klar, dass $\mathcal{A}$ ein d-System ist, und damit ist $\mathrm{d}(\mathcal{S}) \subset \mathcal{A}$. Nach Lemma 1.2 ist aber $\mathrm{d}(\mathcal{S})$ eine Algebra und daraus ergibt sich nach Lemma 1.6 (Analysis III), dass $\sigma(\mathrm{d}(\mathcal{S})) \subset \mathcal{A}$. Folglich ist $\mathcal{A} = \mathcal{F}$. □

Lemma 1.3 *Sei* $(X, \mathcal{F}, \mu)$ *ein Wahrscheinlichkeitsraum und sei* $f : X \to \mathbb{R}$ *eine* $\mathcal{F}$*-messbare Abbildung mit* $\mu(\{x \in X : f(x) \leq \alpha\})$ *entweder* 0 *oder* 1 *für jedes* $\alpha \in \mathbb{R}$. *Dann gibt es ein* $c \in \mathbb{R}$, *so dass* $f = c$ μ*-fast sicher.*

Beweis Übung. □

Im Folgenden sei $(X, \mathcal{F}, \mu)$ ein Wahrscheinlichkeitsraum. Eine komplexwertige Abbildung $f : X \to \mathbb{C}$ heißt $\mathcal{F}$*-messbar*, wenn die Abbildungen $\operatorname{Re} f : X \to \mathbb{R}$ und $\operatorname{Im} f : X \to \mathbb{R}$ beide $\mathcal{F}$-messbar sind. Eine $\mathcal{F}$-messbare Abbildung $f : X \to \mathbb{C}$ heißt μ*-integrierbar*, wenn $\operatorname{Re} f$ und $\operatorname{Im} f$ beide μ-integrierbar sind und in diesem Fall wird das *Integral* $\int f \, d\mu$ von f bezüglich μ durch

$$\int f \, d\mu = \int \operatorname{Re} f \, d\mu + i \int \operatorname{Im} f \, d\mu$$

definiert. Die Menge aller beschränkten $\mathcal{F}$-messbaren Abbildungen von X nach $\mathbb{R}$ bzw. von X nach $\mathbb{C}$ wird mit $\mathrm{B}(X, \mathcal{F})$ bzw. mit $\mathrm{B}_{\mathbb{C}}(X, \mathcal{F})$ bezeichnet.

Sei $p \geq 1$; dann gilt $\mathcal{L}^p(X, \mathcal{F}, \mu) \subset \mathcal{L}^1(X, \mathcal{F}, \mu)$ und ferner ist $\|f\|_1 \leq \|f\|_p$ für alle $f \in \mathcal{L}^p(X, \mathcal{F}, \mu)$, da nach der Hölderschen Ungleichung

$$\|f\|_1 = \|1f\|_1 \leq \|1\|_q \|f\|_p = \|f\|_p \,.$$

Lemma 1.4 *Sei* $(Y, \mathcal{B})$ *ein Messraum und sei* $f : X \to Y$ *eine* $\mathcal{F}|\mathcal{B}$*-messbare Abbildung. Dann ist die durch* $\nu(B) = \mu(f^{-1}(B))$ *für jedes* $B \in \mathcal{B}$ *definierte Abbildung* $\nu : \mathcal{B} \to \mathbb{R}^+$ *ein Wahrscheinlichkeitsmaß auf* $\mathcal{B}$.

Beweis Es gilt $\nu(\varnothing) = \mu(f^{-1}(\varnothing)) = \mu(\varnothing) = 0$, und ist $\{B_n\}_{n \geq p}$ eine Folge paarweise disjunkter Elemente aus $\mathcal{B}$, so ist $\{f^{-1}(B_n)\}_{n \geq p}$ eine Folge paarweise disjunkter Elemente aus $\mathcal{F}$ und damit ist

$$\begin{aligned}
\nu\left(\bigcup_{n=p}^{\infty} B_n\right) &= \mu\left(f^{-1}\left(\bigcup_{n=p}^{\infty} B_n\right)\right) \\
&= \mu\left(\bigcup_{n=p}^{\infty} f^{-1}(B_n)\right) = \sum_{n=p}^{\infty} \mu(f^{-1}(B_n)) = \sum_{n=p}^{\infty} \nu(B_n) \,.
\end{aligned}$$

Folglich ist ν ein Maß auf $\mathcal{B}$ mit $\nu(Y) = \mu(f^{-1}(Y)) = \mu(X) = 1$. $\square$

Das Maß ν in Lemma 1.4 heißt *das Bildmaß von μ unter f*.

Satz 1.2 *Sei $(Y, \mathcal{B})$ ein Messraum, $f : X \to Y$ eine $\mathcal{F}|\mathcal{B}$-messbare Abbildung, sei ν das Bildmaß von μ unter f und sei $g \in \mathrm{M}(Y, \mathcal{B})$. Dann ist $g \circ f \in \mathcal{L}^1(X, \mathcal{F}, \mu)$ genau, wenn $g \in \mathcal{L}^1(Y, \mathcal{B}, \nu)$ und in diesem Fall gilt*

$$\int g \circ f \, d\mu = \int g \, d\nu \, .$$

Beweis Für jedes $B \in \mathcal{B}$ gilt

$$\int \chi_B \circ f \, d\mu = \int \chi_{f^{-1}(B)} \, d\mu = \mu(f^{-1}(B)) = \nu(B) = \int \chi_B \, d\nu$$

und durch Linearität gilt dann $\int g \circ f \, d\mu = \int g \, d\nu$ für alle $g \in \mathrm{E}(Y, \mathcal{B})$. Durch die Definition des Integrals gilt also $g \circ f \in \mathcal{L}^1_+(X, \mathcal{F}, \mu)$ mit $\int g \circ f \, d\mu = \int g \, d\nu$ genau dann, wenn $g \in \mathcal{L}^1_+(Y, \mathcal{B}, \nu)$. Damit ist $g \circ f = (g^+ \circ f) - (g^- \circ f) \in \mathcal{L}^1(X, \mathcal{F}, \mu)$ genau, wenn $g \in \mathcal{L}^1(Y, \mathcal{B}, \nu)$, und in diesem Fall gilt

$$\begin{aligned}\int g \circ f \, d\mu &= \int ((g^+ \circ f) - (g^- \circ f)) \, d\mu \\ &= \int g^+ \circ f \, d\mu - \int g^- \circ f \, d\mu = \int g^+ \, d\nu - \int g^- \, d\nu = \int g \, d\nu \, . \quad \square\end{aligned}$$

Sei $\mathcal{E}$ eine Unter-σ-Algebra von $\mathcal{F}$, d.h., $\mathcal{E}$ ist eine σ-Algebra mit $\mathcal{E} \subset \mathcal{F}$. Nach Satz 10.6 (Analysis 3) gibt es dann zu jedem $f \in \mathcal{L}^1(X, \mathcal{F}, \mu)$ ein Element $\mathrm{E}_\mu(f|\mathcal{E}) \in \mathrm{M}(X, \mathcal{E}) \cap \mathcal{L}^1(X, \mathcal{F}, \mu)$, so dass

$$\int_E \mathrm{E}_\mu(f|\mathcal{E}) \, d\mu = \int_E f \, d\mu$$

für alle $E \in \mathcal{E}$; ferner ist $\mathrm{E}_\mu(f|\mathcal{E})$ μ-fast sicher eindeutig. $\mathrm{E}_\mu(f|\mathcal{E})$ heißt die *bedingte Erwartung von f bezüglich $\mathcal{E}$*.

Satz 1.3 *(1) Für alle f_1, $f_2 \in \mathcal{L}^1(X, \mathcal{F}, \mu)$ und alle c_1, $c_2 \in \mathbb{R}$ gilt*

$$\mathrm{E}_\mu(c_1 f_1 + c_2 f_2|\mathcal{E}) = c_1 \mathrm{E}_\mu(f_1|\mathcal{E}) + c_2 \mathrm{E}_\mu(f_2|\mathcal{E}) \ \mu\text{-fast sicher} \, .$$

(2) Es gilt $\mathrm{E}_\mu(f|\mathcal{E}) \geq 0$ μ-fast sicher für jedes $f \in \mathcal{L}^1(X, \mathcal{F}, \mu)$ mit $f \geq 0$ μ-fast sicher.

(3) Ist $g \in \mathcal{L}^1(X, \mathcal{F}, \mu)$ $\mathcal{E}$-messbar, so ist $\mathrm{E}_\mu(g|\mathcal{E}) = g$ μ-fast sicher. Insbesondere gilt $\mathrm{E}_\mu(\mathrm{E}_\mu(f|\mathcal{E})|\mathcal{E}) = \mathrm{E}_\mu(f|\mathcal{E})$ μ-fast sicher für jedes $f \in \mathcal{L}^1(X, \mathcal{F}, \mu)$.

(4) Für alle $f \in \mathcal{L}^1(X, \mathcal{F}, \mu)$, $g \in \mathrm{B}(X, \mathcal{E})$ gilt

$$\int \mathrm{E}_\mu(f|\mathcal{E}) g \, d\mu = \int f g \, d\mu \, .$$

Beweis Übung. □

Satz 1.4 *Sei $p \geq 1$; für jedes $f \in \mathcal{L}^p(X, \mathcal{F}, \mu)$ ist $\mathrm{E}_\mu(f|\mathcal{E}) \in \mathcal{L}^p(X, \mathcal{F}, \mu)$ und es gilt $\|\mathrm{E}_\mu(f|\mathcal{E})\|_p \leq \|f\|_p$.*

Beweis Übung. □

Satz 1.5 *Sei X ein kompakter metrischer Raum, sei $\mathcal{B}_X$ die Borelsche σ-Algebra und sei $\mu \in \mathrm{P}(X, \mathcal{B}_X)$. Zu jedem $f \in \mathcal{L}^1(X, \mathcal{B}_X, \mu)$ und jedem $\varepsilon > 0$ gibt es dann eine stetige Abbildung $g : X \to \mathbb{R}$, so dass $\int |f - g| \, d\mu < \varepsilon$.*

Beweis Übungsaufgabe 23, Analysis III. □

2 Der Birkhoffsche Ergodensatz

Sei $(X, \mathcal{F}, \mu)$ ein Wahrscheinlichkeitsraum. Eine Abbildung $T : X \to X$ heißt *messbar*, wenn $T^{-1}(\mathcal{F}) \subset \mathcal{F}$ für alle $F \in \mathcal{F}$ und eine messbare Abbildung T heißt *maßtreu* bezüglich μ, wenn $\mu(T^{-1}(F)) = \mu(F)$ für alle $F \in \mathcal{F}$, d.h., wenn μ selbst das Bildmaß von μ unter T ist. In der maßtheoretischen Ergodentheorie werden die Iterierten $\{T^n\}_{n\geq 0}$ einer maßtreuen Abbildung T studiert, wobei $T^0 = \mathrm{id}_X$, $T^1 = T$ und (für $n \geq 2$) $T^n = T \circ T^{n-1}$.

Im Folgenden sei $(X, \mathcal{F}, \mu)$ ein Wahrscheinlichkeitsraum und $T : X \to X$ eine messbare Abbildung, die maßtreu bezüglich μ ist.

Satz 2.1 (Wiederkehrsatz von Poincaré) *Für jedes $B \in \mathcal{F}$ gilt*

$$\mu(\{x \in B : T^n(x) \in B \textit{ für unendlich viele } n \geq 0\}) = \mu(B)\ .$$

Beweis Sei $B \in \mathcal{F}$ fest und für jedes $n \geq 0$ setze

$$B_n = \{x \in X : T^m(x) \in B \text{ für mindestens ein } m \geq n\}\ .$$

Dann gilt $T^{-1}(B_n) = B_{n+1}$ für jedes $n \geq 0$ und damit ist $\mu(B_n) = \mu(B_0)$ für alle $n \geq 0$, da T maßtreu ist. Nun ist die Folge $\{B_n\}_{n\geq 0}$ monoton fallend und daraus ergibt sich, dass $\mu(B_\infty) = \mu(B_0)$, wobei $B_\infty = \bigcap_{n\geq 0} B_n$. Folglich ist

$$\mu(B \setminus (B \cap B_\infty)) = \mu(B \setminus B_\infty) \leq \mu(B_0 \setminus B_\infty) = 0\ ,$$

da $B \subset B_0$, d.h., $\mu(B \cap B_\infty) = \mu(B)$. Aber

$$B \cap B_\infty = \{x \in B : T^n(x) \in B \text{ für unendlich viele } n \geq 0\}\ . \quad \square$$

Für jedes $B \in \mathcal{F}$ und jedes $x \in X$ gilt

$$\Big(\sum_{n\geq 0} \chi_B \circ T^n\Big)(x) = \sum_{n\geq 0} \chi_B(T^n(x)) = |\{n \geq 0 : T^n(x) \in B\}|$$

und also kann man Satz 2.1 wie folgt umformulieren: Es gilt

$$\sum_{n\geq 0} \chi_B \circ T^n = \infty \quad \mu\text{-fast sicher auf } B$$

für jedes $B \in \mathcal{F}$. In der Tat gilt $\sum_{n\geq 0} \chi_B \circ T^n = \infty \chi_{B^*}$ μ-fast sicher für jedes $B \in \mathcal{F}$, wobei $B^* = \{x \in X : T^n(x) \in B \text{ für ein } n \geq 0\}$. (Der Beweis dafür ist eine Übung.)

Setze nun $\mathcal{I}_T = \{B \in \mathcal{F} : T^{-1}(B) = B\}$; man sieht leicht, dass $\mathcal{I}_T$ eine σ-Algebra ist; sie heißt die *invariante σ-Algebra von T*.

Lemma 2.1 *Eine $\mathcal{F}$-messbare Abbildung $f : X \to \mathbb{R}$ ist $\mathcal{I}_T$-messbar genau dann, wenn $f \circ T = f$.*

Beweis Gilt $f \circ T = f$, so ist $f^{-1}(B) = T^{-1}(f^{-1}(B))$ für jedes $B \in \mathcal{B} = \mathcal{B}(\mathbb{R})$. Damit ist $f^{-1}(B) \in \mathcal{I}_T$ für jedes $B \in \mathcal{B}$, d.h., f ist $\mathcal{I}_T$-messbar. Umgekehrt gilt $\chi_E \circ T = \chi_E$ für jedes $E \in \mathcal{I}_T$, folglich gilt $f \circ T = f$ für jedes $f \in \mathrm{E}(X, \mathcal{I}_T)$ und daraus ergibt sich nach Satz 5.4 (Analysis 3), dass $f \circ T = f$ für jedes $f \in \mathrm{M}(X, \mathcal{I}_T)$. $\square$

Satz 2.2 (Birkhoffscher Ergodensatz) *Für jedes $f \in \mathcal{L}^1(X, \mathcal{F}, \mu)$ gilt*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k = \mathrm{E}_\mu(f|\mathcal{I}_T) \quad \mu\text{-fast sicher.}$$

Beweis Dieser erstreckt sich über die nächsten drei Seiten. $\square$

Der Schlüssel zum Beweis ist der maximale Ergodensatz von Hopf:

Satz 2.3 *Sei $U : \mathcal{L}^1(X, \mathcal{F}, \mu) \to \mathcal{L}^1(X, \mathcal{F}, \mu)$ eine lineare Abbildung, für die gilt:*

($*$) Für jedes $f \in \mathcal{L}^1_+(X, \mathcal{F}, \mu)$ ist $Uf \geq 0$ μ-fast sicher und $\int Uf\, d\mu \leq \int f\, d\mu$.

Für alle $f \in \mathcal{L}^1(X, \mathcal{F}, \mu)$ gilt dann $\int_{E_\infty(f)} f\, d\mu \geq 0$, wobei

$$E_\infty(f) = \left\{ x \in X : \text{es gibt ein } n \geq 0 \text{ mit } \sum_{k=0}^{n} (U^k f)(x) > 0 \right\}.$$

Beweis (von Adriano Garsia) Für $f \in \mathcal{L}^1(X, \mathcal{F}, \mu)$, $n \geq 0$, sei

$$\begin{aligned} E_n(f) &= \left\{ x \in X : \text{es gibt ein } 0 \leq k \leq n \text{ mit } \sum_{j=0}^{k} (U^k f)(x) > 0 \right\} \\ &= \left\{ x \in X : \max_{0\leq k\leq n} \sum_{j=0}^{k} (U^j f)(x) > 0 \right\}. \end{aligned}$$

Dann ist $\{E_n(f)\}_{n\geq 0}$ eine monoton wachsende Folge mit $E_\infty(f) = \bigcup_{n\geq 0} E_n(f)$, und nach dem Satz von der dominierten Konvergenz genügt es daher zu zeigen, dass $\int_{E_n(f)} f\, d\mu \geq 0$ für alle $f \in \mathcal{L}^1(X, \mathcal{F}, \mu)$, $n \geq 0$.

Für jedes $n \geq 1$ definiere $\varphi_n : \mathbb{R}^n \to \mathbb{R}^+$ durch

$$\varphi_n(x_1, \ldots, x_n) = \Big(\max_{1\leq k\leq n} (x_1 + \cdots + x_k) \Big)^+$$

für alle $x_1, \dots, x_n \in \mathbb{R}$, wobei $y^+ = \max\{y, 0\}$; insbesondere ist φ_n stetig. Sind $x_1, \dots, x_{n+1} \in \mathbb{R}$ mit $\varphi_n(x_1, \dots, x_n) > 0$, so gilt

$$\begin{aligned}\varphi_n(x_1, \dots, x_n) &= \max_{1 \le k \le n}(x_1 + \cdots + x_k) = x_1 + \Big(\max_{2 \le k \le n}(x_2 + \cdots + x_k)\Big)^+ \\ &\le x_1 + \Big(\max_{2 \le k \le n+1}(x_2 + \cdots + x_k)\Big)^+ = x_1 + \varphi_n(x_2, \dots, x_{n+1})\ .\end{aligned}$$

Für alle $x_1, \dots, x_{n+1} \in \mathbb{R}$ mit $\varphi_n(x_1, \dots, x_n) > 0$ gilt also

$$x_1 + \varphi_n(x_2, \dots, x_{n+1}) \ge \varphi_n(x_1, \dots, x_n) \tag{1}$$

Für $f_1, \dots, f_n \in \mathcal{L}^1(X, \mathcal{F}, \mu)$ ist $\varphi_n(f_1, \dots, f_n) \in \mathrm{M}(X, \mathcal{F})$, da φ_n stetig ist, und da $\varphi_n(f_1, \dots, f_n) \le |f_1| + \cdots + |f_n|$, ist dann $\varphi_n(f_1, \dots, f_n) \in \mathcal{L}^1(X, \mathcal{F}, \mu)$. Nun ist $E_n(f) = \{x \in X : \varphi_{n+1}(f, \dots, U^n f)(x) > 0\}$ und daraus folgt nach (1), dass

$$f \ge \varphi_{n+1}(f, \dots, U^n f) - \varphi_{n+1}(Uf, \dots, U^{n+1} f) \ \text{ auf } E_n(f) \tag{2}$$

Setze $g_0 = 0$ und für $k = 1, \dots, n$ sei $g_k = \sum_{j=0}^{k-1} U^j f$; dann gilt

$$\begin{aligned}\varphi_{n+1}(Uf, \dots, U^{n+1} f) &= \Big(\max_{1 \le k \le n+1} \sum_{j=1}^{k} U^j f\Big)^+ = \Big(\max_{1 \le k \le n+1} U \sum_{j=0}^{k-1} U^j f\Big)^+ \\ &= \Big(\max_{1 \le k \le n+1} U g_k\Big)^+ = \max_{0 \le k \le n+1} U g_k\ ,\end{aligned}$$

da $Ug_0 = 0$, und genauso gilt $\varphi_{n+1}(f, \dots, U^n f) = \max_{0 \le k \le n+1} g_k$. Ferner ist

$$\max_{0 \le k \le n+1} U g_k \le U\Big(\max_{0 \le k \le n+1} g_k\Big) \ \ \mu\text{-fast sicher}\ .$$

(Sei $h = \max_{0 \le k \le n+1} g_k$; für jedes k ist $Uh - Ug_k = U(h - g_k) \ge 0$ μ-fast sicher, da $h - g_k \ge 0$. Daher ist $\max_{0 \le k \le n+1} Ug_k \le Uh$ μ-fast sicher.) Also gilt

$$\varphi_{n+1}(Uf, \dots, U^{n+1} f) \le U\varphi_{n+1}(f, \dots, U^n f) \ \ \mu\text{-fast sicher} \tag{3}$$

Aus (2) und (3) folgt nun, dass

$$f \ge \varphi_{n+1}(f, \dots, U^n f) - U\varphi_{n+1}(f, \dots, U^n f) \ \ \mu\text{-fast sicher auf } E_n(f) \tag{4}$$

Nun ist $U\varphi_{n+1}(f, \dots, U^n f) \ge 0$ μ-fast sicher, da $\varphi_{n+1}(f, \dots, U^n f) \ge 0$, und $\varphi_{n+1}(f, \dots, U^n f) = 0$ auf $X \setminus E_n(f)$. Damit ist nach (4)

$$\begin{aligned}\int_{E_n(f)} f\, d\mu &\ge \int_{E_n(f)} \varphi_{n+1}(f, \dots, U^n f)\, d\mu - \int_{E_n(f)} U\varphi_{n+1}(f, \dots, U^n f)\, d\mu \\ &= \int \varphi_{n+1}(f, \dots, U^n f)\, d\mu - \int_{E_n(f)} U\varphi_{n+1}(f, \dots, U^n f)\, d\mu \\ &\ge \int \varphi_{n+1}(f, \dots, U^n f)\, d\mu - \int U\varphi_{n+1}(f, \dots, U^n f)\, d\mu \ge 0\ ,\end{aligned}$$

wobei der letzte Schritt aus $(*)$ folgt, da $\varphi_{n+1}(f, \dots, U^n f) \geq 0$. $\square$

Sei nun $U : \mathrm{M}(X, \mathcal{F}) \to \mathrm{M}(X, \mathcal{F})$ die Abbildung, die definiert ist durch $Uf = f \circ T$ für alle $f \in \mathrm{M}(X, \mathcal{F})$. Es ist klar, dass U linear ist.

Lemma 2.2 *Für jede Abbildung $f \in \mathcal{L}^1(X, \mathcal{F}, \mu)$ ist $Uf \in \mathcal{L}^1(X, \mathcal{F}, \mu)$ und es gilt $\int Uf\, d\mu = \int f\, d\mu$.*

Beweis Dies folgt direkt aus Satz 1.2, da μ das Bildmaß von μ unter T ist. $\square$

Die Einschränkung von U auf $\mathcal{L}^1(X, \mathcal{F}, \mu)$ wird auch mit U bezeichnet. Für alle $f \in \mathcal{L}^1_+(X, \mathcal{F}, \mu)$ ist $Uf \in \mathcal{L}^1_+(X, \mathcal{F}, \mu)$ und es gilt $\int Uf\, d\mu = \int f\, d\mu$. Damit erfüllt U die Bedingung $(*)$.

Lemma 2.3 *Für jede Folge $\{a_n\}_{n \geq 0}$ reeller Zahlen gilt*

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} a_k = \limsup_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} a_k \ .$$

Beweis Dies folgt unmittelbar aus der Tatsache, dass

$$\frac{1}{n} \sum_{k=1}^{n} a_k = \frac{n+1}{n} \Big(\frac{1}{m} \sum_{k=0}^{m-1} a_k \Big) - \frac{a_0}{n}$$

für jedes $n \geq 1$, wobei $m = n + 1$. $\square$

Lemma 2.4 *Sei $g \in \mathcal{L}^1(X, \mathcal{F}, \mu)$, $\varepsilon > 0$ und sei*

$$D = \Big\{ x \in X : \limsup_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} (g \circ T^k)(x) > \varepsilon \Big\} \ .$$

Dann ist $D \in \mathcal{I}_U$ und ferner ist $D \subset E_\infty(h)$, wobei $h = (g - \varepsilon)\chi_D$.

Beweis Nach Lemma 2.3 gilt

$$\begin{aligned} T^{-1}(D) &= \Big\{ x \in X : \limsup_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} (g \circ T^k)(T(x)) > \varepsilon \Big\} \\ &= \Big\{ x \in X : \limsup_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} (g \circ T^k)(x) > \varepsilon \Big\} = D \end{aligned}$$

und daraus ergibt sich, dass $D \in \mathcal{I}_T$. Sei nun $x \in D$; dann gibt es ein $N \geq 1$, so dass $\sum_{k=0}^{N-1}(g \circ T^k)(x) > N\varepsilon$, und folglich ist

$$\begin{aligned} 0 &< \sum_{k=0}^{N-1}(g \circ T^k)(x) - N\varepsilon = \sum_{k=0}^{N-1}((g \circ T^k)(x) - \varepsilon) = \sum_{k=0}^{N-1}(g-\varepsilon)(T^k(x)) \\ &= \sum_{k=0}^{N-1}((g-\varepsilon)\chi_D)(T^k(x)) = \sum_{k=0}^{N-1} h(T^k(x)) = \sum_{k=0}^{N-1}(U^k h)(x)\ , \end{aligned}$$

da $x \in D$ und $D \in \mathcal{I}_T$ und damit $T^k(x) \in D$ für alle k, d.h., $\sum_{k=0}^{N-1}(U^k h)(x) > 0$. Also ist $x \in E_\infty(h)$ und dies zeigt, dass $D \subset E_\infty(h)$. □

Lemma 2.5 *Für jedes $g \in \mathcal{L}^1(X, \mathcal{F}, \mu)$ mit $\mathrm{E}_\mu(g|\mathcal{I}_T) = 0$ μ-fast sicher gilt*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} g \circ T^k = 0 \quad \mu\text{-fast sicher.}$$

Beweis Sei $g \in \mathcal{L}^1(X, \mathcal{F}, \mu)$ mit $\mathrm{E}_\mu(g|\mathcal{I}_T) = 0$ μ-fast sicher, sei $\varepsilon > 0$ und setze

$$D_\varepsilon = \Big\{x \in X : \limsup_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1}(g \circ T^k)(x) > \varepsilon\Big\}$$

und $h_\varepsilon = (g-\varepsilon)\chi_{D_\varepsilon}$. Nach Satz 2.3 und Lemma 2.4 ist dann

$$0 \leq \int_{E_\infty(h_\varepsilon)} h_\varepsilon\, d\mu = \int_{E_\infty(h_\varepsilon)} (g-\varepsilon)\chi_{D_\varepsilon}\, d\mu = \int_{D_\varepsilon} (g-\varepsilon)\, d\mu = \int_{D_\varepsilon} g\, d\mu - \varepsilon\mu(D_\varepsilon)$$

und da $D_\varepsilon \in \mathcal{I}_T$ ist $\int_{D_\varepsilon} g\, d\mu = \int_{D_\varepsilon} \mathrm{E}_\mu(g|\mathcal{I}_T)\, d\mu = 0$. Damit ist $\mu(D_\varepsilon) = 0$ und da dies für jedes $\varepsilon > 0$ gilt, ist also

$$\mu\Big(\Big\{x \in X : \limsup_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} g \circ T^k(x) > 0\Big\}\Big) = \mu\Big(\bigcup_{m\geq 1} D_{1/m}\Big) = 0\ .$$

Mit anderen Worten, es gilt $\limsup_{n\to\infty} n^{-1}\sum_{k=0}^{n-1} g \circ T^k \leq 0$ μ-fast sicher.

Nun gilt auch $\mathrm{E}_\mu(-g|\mathcal{I}_T) = 0$ μ-fast sicher, und daraus ergibt sich, dass

$$\liminf_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} g \circ T^k = -\limsup_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1}(-g) \circ T^k \geq 0 \quad \mu\text{-fast sicher.}$$

Dies zeigt dann, dass $\lim_{n\to\infty} n^{-1}\sum_{k=0}^{n-1} g \circ T^k = 0$ μ-fast sicher. □

Beweis für Satz 2.2 Sei $f \in \mathcal{L}^1(X, \mathcal{F}, \mu)$ und setze $g = f - \mathrm{E}_\mu(f|\mathcal{I}_T)$. Nach Lemma 2.1 gilt $\mathrm{E}_\mu(f|\mathcal{I}_T) \circ T = \mathrm{E}_\mu(f|\mathcal{I}_T)$, da $\mathrm{E}_\mu(f|\mathcal{I}_T)$ $\mathcal{I}_T$-messbar ist, also gilt $\mathrm{E}_\mu(f|\mathcal{I}_T) \circ T^k = \mathrm{E}_\mu(f|\mathcal{I}_T)$ für jedes $k \geq 0$ und folglich ist

$$\frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k = \frac{1}{n}\sum_{k=0}^{n-1} g \circ T^k + \mathrm{E}_\mu(f|\mathcal{I}_T)$$

für jedes $n \geq 1$. Aber $\mathrm{E}_\mu(g|\mathcal{I}_T) = 0$ μ-fast sicher, da die Abbildung 0 $\mathcal{I}_T$-messbar ist und $\int_B 0\, d\mu = \int_B (f - \mathrm{E}_\mu(f|\mathcal{I}_T))\, d\mu = \int_B g\, d\mu$ für alle $B \in \mathcal{F}$. Daraus ergibt sich nach Lemma 2.5, dass

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k = \mathrm{E}_\mu(f|\mathcal{I}_T) \quad \mu\text{-fast sicher} . \quad \square$$

Sei $p \geq 1$; dann gilt $\mathcal{L}^p(X, \mathcal{F}, \mu) \subset \mathcal{L}^1(X, \mathcal{F}, \mu)$ und nach Satz 1.4 ist $\mathrm{E}_\mu(f|\mathcal{I}_T)$ ein Element von $\mathcal{L}^p(X, \mathcal{F}, \mu)$ für jedes $f \in \mathcal{L}^p(X, \mathcal{F}, \mu)$.

Satz 2.4 *Sei $p \geq 1$; für jedes $f \in \mathcal{L}^p(X, \mathcal{F}, \mu)$ gilt dann*

$$\lim_{n\to\infty} \left\| \frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k - \mathrm{E}_\mu(f|\mathcal{I}_T) \right\|_p = 0 .$$

Beweis Für jedes $f \in \mathcal{L}^p(X, \mathcal{F}, \mu)$, $n \geq 1$, setze $\sigma_n(f) = 1/n \sum_{k=0}^{n-1} f \circ T^k$. Man beachte, dass $\|\sigma_n(f)\|_p \leq \|f\|_p$, da $\|f \circ T^k\|_p = \|f\|_p$ für alle $k \geq 0$, und nach Satz 2.2 gilt $\lim_{n\to\infty} \sigma_n(f) = \mathrm{E}_\mu(f|\mathcal{I}_T)$ μ-fast sicher.

Betrachte zunächst eine $\mathcal{F}$-messbare beschränkte Abbildung $g : X \to \mathbb{R}$ und sei $\alpha \in \mathbb{R}^+$ eine obere Schranke für $|g|$. Dann gilt $|g \circ T^k| \leq \alpha$ für alle $k \geq 0$ und auch $|\mathrm{E}_\mu(g|\mathcal{I}_S)| \leq \alpha$ μ-fast sicher und damit ist $|\sigma_n(g) - \mathrm{E}_\mu(g|\mathcal{I}_T)|^p \leq (2\alpha)^p$ für alle $k \geq 0$. Nach Satz 2.2 und dem Satz von der dominierten Konvergenz ist daher $\lim_{n\to\infty} \int |\sigma_n(g) - \mathrm{E}_\mu(g|\mathcal{I}_T)|^p\, d\mu = 0$, d.h., $\lim_{n\to\infty} \|\sigma_n(g) - \mathrm{E}_\mu(g|\mathcal{I}_T)\|_p = 0$, und insbesondere ist $\{\sigma_n(g)\}_{n\geq 1}$ eine Cauchy-Folge in $\mathcal{L}^p(X, \mathcal{F}, \mu)$.

Sei nun $f \in \mathcal{L}^p(X, \mathcal{F}, \mu)$ und $\varepsilon > 0$. Dann gibt es eine $\mathcal{F}$-messbare beschränkte Abbildung $g : X \to \mathbb{R}$ mit $\|f - g\|_p < \varepsilon/3$ und da $\{\sigma_n(g)\}_{n\geq 1}$ eine Cauchy-Folge ist, gibt es ein $N \geq 1$, so dass $\|\sigma_n(g) - \sigma_m(g)\|_p < \varepsilon/3$ für alle m, $n \geq N$. Für alle m, $n \geq N$ gilt dann

$$\begin{aligned} \|\sigma_n(f) - \sigma_m(f)\|_p &= \|\sigma_n(f - g) - \sigma_m(f - g) + \sigma_n(g) - \sigma_m(g)\|_p \\ &\leq \|\sigma_n(f - g)\|_p + \|\sigma_m(f - g)\|_p + \|\sigma_n(g) - \sigma_m(g)\|_p \\ &\leq 2\|f - g\|_p + \|\sigma_n(g) - \sigma_m(g)\|_p < \varepsilon , \end{aligned}$$

und dies zeigt, dass $\{\sigma_n(f)\}_{n\geq 1}$ eine Cauchy-Folge in $\mathcal{L}^p(X,\mathcal{F},\mu)$ ist. Folglich gibt es $h \in \mathcal{L}^p(X,\mathcal{F},\mu)$, so dass $\lim_{n\to\infty} \|\sigma_n(f) - h\|_p = 0$, und daraus ergibt sich, dass $\lim_{n\to\infty} \|\sigma_n(f) - \mathrm{E}(f|\mathcal{I}_T)\|_p = 0$, da $\lim_{n\to\infty} \sigma_n(f) = \mathrm{E}_\mu(f|\mathcal{I}_T)$ μ-fast sicher und es eine Teilfolge $\{n_k\}_{k\geq 1}$ gibt, so dass auch $\lim_{k\to\infty} \sigma_{n_k}(f) = h$ μ-fast sicher. $\square$

Eine σ-Algebra $\mathcal{E} \subset \mathcal{F}$ heißt *μ-trivial*, wenn $\mu(B)$ entweder 0 oder 1 ist für jedes $B \in \mathcal{E}$. Ist die σ-Algebra $\mathcal{E}$ μ-trivial, so ist $\mathrm{E}_\mu(f|\mathcal{E}) = \int f\,d\mu$ μ-fast sicher für jedes $f \in \mathcal{L}^1(X,\mathcal{F},\mu)$, da die konstante Abbildung $\int f\,d\mu$ $\mathcal{E}$-messbar ist und für jedes $B \in \mathcal{F}$ mit $\mu(B)$ entweder 0 oder 1 gilt $\int_B (\int f\,d\mu)\,d\mu = \int_B f\,d\mu$.

Die maßtreue Abbildung T heißt *ergodisch bezüglich* μ, wenn die σ-Algebra $\mathcal{I}_T$ μ-trivial ist. Ist T ergodisch, so folgt aus Satz 2.2, dass für jedes $f \in \mathcal{L}^1(X,\mathcal{F},\mu)$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k = \int f\,d\mu \quad \mu\text{-fast sicher}\,.$$

3 Stationäre Folgen

Sei $(E, \mathcal{E})$ ein Messraum und sei I entweder $\mathbb{Z}$ oder $\mathbb{N} = \{0, 1, \ldots\}$. Setze nun $E_\infty = \mathrm{Abb}(I, E)$; wie üblich werden die Elemente von E_∞ dargestellt als Folgen $\{x_n\}_{n \in I}$ mit $x_n \in E$ für jedes $n \in I$. Für jedes $m \in I$ sei $\mathrm{p}_m : E_\infty \to E$ die durch $\mathrm{p}_m(\{x_n\}_{n\in I}) = x_m$ gegebene Projektionsabbildung. Setze $\mathcal{E}_\infty = \sigma\big(\bigcup_{n\in I} \mathrm{p}_m^{-1}(\mathcal{E})\big)$.

Definiere $S : E_\infty \to E_\infty$ durch $S(\{x_n\}_{n\in I}) = \{x'_n\}_{n\in I}$, wobei $x'_n = x_{n+1}$. Dann gilt $\mathrm{p}_{m+1} = \mathrm{p}_m \circ S$ für jedes $m \in I$ und daraus ergibt sich, dass $S^{-1}(\mathcal{E}_\infty) \subset \mathcal{E}_\infty$.

Lemma 3.1 *Die Abbildung S ist maßtreu bezüglich eines Maßes $\nu \in \mathrm{P}(E_\infty, \mathcal{E}_\infty)$ genau dann, wenn für alle $B_k, \ldots, B_\ell \in \mathcal{E}$ und alle $k, \ell \in I$ mit $k \le \ell$*

$$\nu\Big(\bigcap_{n=k}^{\ell} \mathrm{p}_n^{-1}(B_n)\Big) = \nu\Big(\bigcap_{n=k}^{\ell} \mathrm{p}_{n+1}^{-1}(B_n)\Big) \tag{5}$$

Beweis Ist S maßtreu bezüglich ν, dann gilt (1), da

$$\bigcap_{n=k}^{\ell} \mathrm{p}_{n+1}^{-1}(B_n) = S^{-1}\Big(\bigcap_{n=k}^{\ell} \mathrm{p}_n^{-1}(B_n)\Big) .$$

Nehme umgekehrt an, dass (1) gilt für alle $B_k, \ldots, B_\ell \in \mathcal{E}$ und alle $k, \ell \in I$ mit $k \le \ell$. Dann ist $\nu(D) = \nu(S^{-1}(D))$ für alle $D \in \mathcal{D}$, wobei $\mathcal{D}$ die Menge aller Teilmengen von E_∞ ist, die eine Darstellung besitzen der Form $\bigcap_{n=k}^{\ell} \mathrm{p}_n^{-1}(B_n)$ mit $B_k, \ldots, B_\ell \in \mathcal{E}$ und $k, \ell \in I$ mit $k \le \ell$. Aber $\mathcal{D}$ ist durchschnittsstabil und $\mathcal{E} = \sigma(\mathcal{D})$, und folglich ist nach Satz 1.1 $\nu(B) = \nu(S^{-1}(B))$ für alle $B \in \mathcal{E}$, d.h., S ist maßtreu bezüglich ν. $\square$

Bezeichne mit $\mathcal{J}$ die Menge aller Teilmengen von I der Form $\{k, \ldots, \ell\}$ mit $k \le \ell$. Für $J \in \mathcal{J}$ sei $E^J = \mathrm{Abb}(J, E)$ und für $m \in J$ sei $\mathrm{p}_m^J : E^J \to E$ die Projektionsabbildung, die gegebene ist durch $\mathrm{p}_m^J(\{x_n\}_{n\in J}) = x_m$. Setze nun $\mathcal{E}^J = \sigma\big(\bigcup_{m\in J} (\mathrm{p}_m^J)^{-1}(\mathcal{E})\big)$, also ist $\mathcal{E}^J$ die Produkt-σ-Algebra auf E^J.

Für jedes $J \in \mathcal{J}$ sei $\nu_J \in \mathrm{P}(E^J, \mathcal{E}^J)$; die Familie von Maßen $\{\nu_J\}_{J\in\mathcal{J}}$ heißt *verträglich*, wenn für alle $J, K \in \mathcal{J}$ mit $J \subset K$

$$\mu_K\Big(\bigcap_{m\in J} (\mathrm{p}_m^K)^{-1}(B_m)\Big) = \mu_J\Big(\bigcap_{m\in J} (\mathrm{p}_m^J)^{-1}(B_m)\Big)$$

gilt für alle Folgen $\{B_m\}_{m\in J}$ aus $\mathcal{E}$.

Im Folgenden sei (E, ϱ) ein σ-kompakter metrischer Raum (d.h., es gibt also eine Folge $\{F_n\}_{n\ge 0}$ kompakter Teilmengen von E mit $E = \bigcup_{n\ge 0} F_n$) und sei $\mathcal{E}$ die Borel-σ-Algebra auf E. Die wichtigsten Anwendungen sind mit $E = \mathbb{R}^d$ (mit der üblichen Metrik) und mit E einer endlichen Menge (mit der diskreten Metrik). Für jedes $J \in \mathcal{J}$ ist dann auch E^J mit der Produktmetrik ein σ-kompakter metrischer Raum und $\mathcal{E}^J$ ist die Borel-σ-Algebra auf E^J. Insbesondere gibt es zu jedem $B \in \mathcal{E}^J$ und jedem $\varepsilon > 0$ eine kompakte Teilmenge F vom E^J mit $F \subset B$ und $\nu^J(B \setminus F) < \varepsilon$.

Der folgende Satz ist im Wesentlichen der Fortsetzungssatz von Kolmogorov.

Satz 3.1 *Sei $\{\nu_J\}_{J\in\mathcal{J}}$ eine verträgliche Familie von Maßen (mit $\nu_J \in \mathrm{P}(E^J, \mathcal{E}^J)$ für jedes $J \in \mathcal{J}$). Dann gibt es ein eindeutiges Maß $\nu \in \mathrm{P}(E_\infty, \mathcal{E}_\infty)$, so dass*

$$\nu\Big(\bigcap_{m\in J} \mathrm{p}_m^{-1}(B_m)\Big) = \nu_J\Big(\bigcap_{m\in J} (\mathrm{p}_m^J)^{-1}(B_m)\Big)$$

für alle $J \in \mathcal{J}$ und alle Folgen $\{B_m\}_{m\in J}$ aus $\mathcal{E}$.

Beweis Für jedes $J \in \mathcal{J}$ definiere $\mathrm{p}_J : E_\infty \to E^J$ durch $\mathrm{p}_J(\{x_n\}_{n\ge 0}) = \{x_n\}_{n\in J}$ und setze $\mathcal{E}_J = \mathrm{p}_J^{-1}(\mathcal{E}^J)$; es ist klar, dass $\mathcal{E}_J = \sigma\big(\bigcup_{m\in J} \mathrm{p}_m^{-1}(\mathcal{E})\big)$. Man beachte, dass $\mathcal{E}_J \subset \mathcal{E}_K$, falls $J \subset K$. Da die Abbildung p_J surjektiv ist, gibt es ein (eindeutiges) $\nu'_J \in \mathrm{P}(E_\infty, \mathcal{E}_J)$, so dass $\nu'_J(\mathrm{p}_J^{-1}(B)) = \nu_J(B)$ für alle $B \in \mathcal{E}^J$. Insbesondere ist für jede Folge $\{B_m\}_{m\in J}$ aus $\mathcal{E}$

$$\nu'_J\Big(\bigcap_{m\in J} \mathrm{p}_m^{-1}(B_m)\Big) = \nu'_J\Big(\mathrm{p}_J^{-1}\big(\bigcap_{m\in J} (\mathrm{p}_m^J)^{-1}(B_m)\big)\Big) = \nu_J\Big(\bigcap_{m\in J} (\mathrm{p}_m^J)^{-1}(B_m)\Big) ,$$

da $\mathrm{p}_m^J \circ \mathrm{p}_J = \mathrm{p}_m$ für jedes $m \in J$. Seien J, $K \in \mathcal{J}$ mit $J \subset K$; dann gilt

$$\begin{aligned}\nu'_K\Big(\bigcap_{m\in J} \mathrm{p}_m^{-1}(B_m)\Big) &= \nu_K\Big(\bigcap_{m\in J} (\mathrm{p}_m^K)^{-1}(B_m)\Big)\\ &= \nu_J\Big(\bigcap_{m\in J} (\mathrm{p}_m^J)^{-1}(B_m)\Big) = \nu'_J\Big(\bigcap_{m\in J} \mathrm{p}_m^{-1}(B_m)\Big)\end{aligned}$$

für jede Folge $\{B_m\}_{m\in J}$ aus $\mathcal{E}$, d.h., $\nu'_K(B) = \nu'_J(B)$ für alle $B \in \mathcal{D}_J$, wobei

$$\mathcal{D}_J = \Big\{ \bigcap_{m\in J} \mathrm{p}_m^{-1}(B_m) : B_m \in \mathcal{E} \text{ für jedes } m \in J \Big\} .$$

Folglich gilt $\nu'_K(B) = \nu'_J(B)$ für alle $B \in \mathcal{E}_J$, da $\mathcal{D}_J$ durchschnittsstabil ist und $\mathcal{E}_J = \sigma(\mathcal{D}_J)$. Setze nun $\mathcal{A} = \bigcup_{J\in\mathcal{J}} \mathcal{E}_J$; also ist $\mathcal{A}$ eine Algebra mit $\sigma(\mathcal{A}) = \mathcal{E}_\infty$. Da $\nu'_K(B) = \nu'_J(B)$ für alle $B \in \mathcal{E}_J$, $J \subset K$, gibt es eine eindeutige Abbildung $\nu : \mathcal{A} \to \mathbb{R}^+$ mit $\nu(B) = \nu'_J(B)$ für alle $B \in \mathcal{E}_J$, $J \in \mathcal{J}$ und insbesondere ist

$$\nu\Big(\bigcap_{m\in J} \mathrm{p}_m^{-1}(B_m)\Big) = \nu_J\Big(\bigcap_{m\in J} (\mathrm{p}_m^J)^{-1}(B_m)\Big)$$

für alle $J \in \mathcal{J}$ und alle Folgen $\{B_m\}_{m\in J}$ aus $\mathcal{E}$. Es bleibt dann nur zu zeigen, dass ν ein Maß auf $\mathcal{A}$ und damit ein Maß auf $\mathcal{E}_\infty$ ist. (Die Eindeutigkeit ist klar.) Nun sieht man leicht, dass ν additiv auf $\mathcal{A}$ ist, und daher ist ν ein Maß genau dann, wenn $\lim_{n\to\infty} \nu(A_n) = 0$ gilt für jede monoton fallende Folge $\{A_n\}_{n\geq 0}$ aus $\mathcal{A}$ mit $\bigcap_{n\geq 0} A_n = \varnothing$.

Betrachte eine monoton fallende Folge $\{A_n\}_{n\geq 0}$ aus $\mathcal{A}$ und nehme an, es gibt ein $\varepsilon > 0$, so dass $\nu(A_n) \geq \varepsilon$ für alle $n \geq 0$. Da es für alle $J_1, J_2 \in \mathcal{J}$ ein $J \in \mathcal{J}$ mit $J_1 \cup J_2 \subset J$ gibt, gibt es eine monoton wachsende Folge $\{J_n\}_{n\geq 0}$ aus $\mathcal{J}$, so dass $A_n \in \mathcal{E}_{J_n}$ für jedes $n \geq 0$. Sei $A'_n = \mathrm{p}_{J_n}(A_n)$; damit ist $A'_n \in \mathcal{E}^{J_n}$ und $A_n = \mathrm{p}_{J_n}^{-1}(A'_n)$. Es gibt nun eine kompakte Teilmenge F_n von E^J mit $F_n \subset A'_n$ und $\nu_J(A'_n \setminus F_n) < 2^{-n-2}\varepsilon$. Dann ist $C_n = \mathrm{p}_J^{-1}(F_n)$ eine Teilmenge von A_n mit

$$\nu(A_n \setminus C_n) = \nu'_{J_n}(\mathrm{p}_{J_n}^{-1}(A_n \setminus C_n)) = \nu'_{J_n}(A'_n \setminus F_n) < 2^{-n-2}\varepsilon\ .$$

Für jedes $n \geq 0$ sei $B_n = \bigcap_{k=0}^n C_k$; daher ist $\{B_n\}_{n\geq 0}$ eine monoton fallende Folge aus $\mathcal{A}$ mit $B_n \subset A_n$ und $\nu(B_n) > \nu(A_n) - \varepsilon/2 > 0$. Insbesondere ist $B_n \neq \varnothing$; für jedes $n \geq 0$ wähle also $x_n \in B_n$. Aber $\{\mathrm{p}_m(x_n)\}_{n\geq m}$ ist eine Folge aus der kompakten Menge F_m für jedes $m \geq 0$ und folglich gibt es nach dem üblichen Diagonalargument eine Teilfolge $\{n_k\}_{k\geq 0}$, so dass die Folge $\{\mathrm{p}_m(x_{n_k})\}_{k\geq 0}$ für jedes $m \geq 0$ konvergiert. Setze $u_m = \lim_{k\to\infty} \mathrm{p}_m(x_{n_k})$; dann sieht man leicht, dass es ein $x \in E_\infty$ gibt, so dass $\mathrm{p}_{J_m}(x) = u_m$ für jedes $m \geq 0$, und da $u_m \in F_m$, ist $x \in \bigcap_{m\geq 0} C_m$. Insbesondere ist $\bigcap_{m\geq 0} A_m \neq \varnothing$. Dies ziegt, dass $\lim_{n\to\infty} \nu(A_n) = 0$ gilt für jede monoton fallende Folge $\{A_n\}_{n\geq 0}$ aus $\mathcal{A}$ mit $\bigcap_{n\geq 0} A_n = \varnothing$, und damit, dass ν ein Maß auf $\mathcal{A}$ ist. □

Satz 3.2 *Für jedes $n \in I$ sei $\nu_n \in \mathrm{P}(E, \mathcal{E})$. Dann gibt es ein eindeutiges Maß $\nu_\infty \in \mathrm{P}(E_\infty, \mathcal{E}_\infty)$, so dass*

$$\nu_\infty\Big(\bigcap_{n=k}^{\ell} \mathrm{p}_n^{-1}(B_n)\Big) = \prod_{n=k}^{\ell} \nu_n(B_n)$$

für alle $B_k, \ldots, B_\ell \in \mathcal{E}$ und alle $k, \ell \in I$ mit $k \leq \ell$. Ferner ist S maßtreu bezüglich ν_∞, wenn $\nu_n = \nu_m$ für alle $m, n \in I$.

Beweis Für jedes $J \in \mathcal{J}$ sei ν_J das Produktmaß $\prod_{m\in J} \nu_m$ auf $(E^J, \mathcal{E}^J)$. Dann ist es klar, dass die Familie $\{\nu_J\}_{J\in\mathcal{J}}$ verträglich ist, und ν_∞ ist das in Satz 3.1 definierte Maß. Die zweite Aussage folgt umittelbar aus Lemma 3.1. □

Sei nun $(X, \mathcal{F}, \mu)$ ein Wahrscheinlichkeitsraum und sei wieder $(E, \mathcal{E})$ ein beliebiger Messraum. Für jedes $n \in I$ sei $f_n : X \to E$ eine $\mathcal{F}|\mathcal{E}$-messbare Abbildung (d.h., es gilt $f_n^{-1}(\mathcal{E}) \subset \mathcal{F}$); die Folge $\{f_n\}_{n\in I}$ heißt *stationär* bezüglich μ, wenn

$$\mu\Big(\bigcap_{n=k}^{\ell} f_n^{-1}(B_n)\Big) = \mu\Big(\bigcap_{n=k}^{\ell} f_{n+1}^{-1}(B_n)\Big)$$

für alle $B_k, \ldots, B_\ell \in \mathcal{E}$ und alle $k, \ell \in I$ mit $k \leq \ell$.

Satz 3.3 *Es gibt ein eindeutiges $\nu \in \mathrm{P}(E_\infty, \mathcal{E}_\infty)$, so dass*

$$\nu\Big(\bigcap_{n=k}^{\ell} \mathrm{p}_n^{-1}(B_n)\Big) = \mu\Big(\bigcap_{n=k}^{\ell} f_n^{-1}(B_n)\Big)$$

für alle $B_k, \ldots, B_\ell \in \mathcal{E}$ und alle $k, \ell \in I$ mit $k \leq \ell$. Ferner ist S maßtreu bezüglich ν genau dann, wenn die Folge $\{f_n\}_{n\in I}$ stationär bezüglich μ ist.

Beweis Definiere eine Abbildung $\mathbf{f} : X \to E_\infty$ durch $\mathbf{f}(x) = \{f_n(x)\}_{n\in I}$ für jedes $x \in I$. Da $\mathrm{p}_n \circ \mathbf{f} = f_n$, ist $\mathbf{f}^{-1}(\mathrm{p}_n^{-1}(\mathcal{E})) = f_n^{-1}(\mathcal{E}) \subset \mathcal{F}$ für jedes $n \in I$ und damit ist $\mathbf{f}^{-1}(\mathcal{E}) \subset \mathcal{F}$. Sei ν das Bildmaß von μ unter $\mathbf{f}$, d.h., $\nu \in \mathrm{P}(E_\infty, \mathcal{E}_\infty)$ ist das Maß, das gegeben ist durch $\nu(B) = \mu(\mathbf{f}^{-1}(B))$ für jedes $B \in \mathcal{E}$. Dann hat ν die gewünschte Eigenschaft, da $\bigcap_{n=k}^{\ell} f_n^{-1}(B_n) = \mathbf{f}^{-1}\big(\bigcap_{n=k}^{\ell} \mathrm{p}_n^{-1}(B_n)\big)$ für alle $B_k, \ldots, B_\ell \in \mathcal{E}$ und alle $k, \ell \in I$ mit $k \leq \ell$. Sei ν' ein weiteres Maß mit dieser Eigenschaft. Dann gilt $\nu'(B) = \nu(B)$ für alle $B \in \mathcal{D}$, wobei $\mathcal{D}$ wie im Beweis für Lemma 3.1 ist, und daraus ergibt sich nach Satz 1.1, dass $\nu' = \nu$. Die zweite Aussage in Satz 3.1 folgt unmittelbar aus Lemma 3.1. □

Lemma 3.2 *Sei $T : X \to X$ eine Abbildung, die maßtreu bezüglich μ ist, sei $f : X \to E$ eine $\mathcal{F}|\mathcal{E}$-messbare Abbildung und für $n \geq 0$ definiere $f_n : X \to E$ durch $f_n = f \circ T^n$. Dann ist die Folge $\{f_n\}_{n\geq 0}$ stationär bezüglich μ.*

Beweis Für alle $B_k, \ldots, B_\ell \in \mathcal{E}$ und alle $k, \ell \in I$ mit $k \leq \ell$ gilt

$$\begin{aligned}\bigcap_{n=k}^{\ell} f_{n+1}^{-1}(B_n) &= \bigcap_{n=k}^{\ell} (T^{n+1})^{-1}(f^{-1}(B_n)) \\ &= T^{-1}\Big(\bigcap_{n=k}^{\ell} (T^n)^{-1}(f^{-1}(B_n))\Big) = T^{-1}\Big(\bigcap_{n=k}^{\ell} f_n^{-1}(B_n)\Big)\end{aligned}$$

und damit $\mu\big(\bigcap_{n=k}^{\ell} f_{n+1}^{-1}(B_n)\big) = \mu\big(T^{-1}\big(\bigcap_{n=k}^{\ell} f_n^{-1}(B_n)\big)\big) = \mu\big(\bigcap_{n=k}^{\ell} f_n^{-1}(B_n)\big)$. □

Sei nun $\{f_n\}_{n\geq 0}$ eine reellwertige stationäre Folge. Hier ist also $(E, \mathcal{E}) = (\mathbb{R}, \mathcal{B})$ (mit $\mathcal{B} = \mathcal{B}(\mathbb{R})$) und $I = \mathbb{N}$ und insbesondere ist $f_n : X \to \mathbb{R}$ eine $\mathcal{F}$-messbare Abbildung für jedes $n \geq 0$. Da die Folge $\{f_n\}_{n\geq 0}$ stationär ist, hängt $\mu(f_n^{-1}(B))$ nicht von n ab für jedes $B \in \mathcal{B}$ und damit ist $f_0 \in \mathcal{L}^1(X, \mathcal{F}, \mu)$ genau dann, wenn $f_n \in \mathcal{L}^1(X, \mathcal{F}, \mu)$ für alle $n \geq 0$.

Sei $\mathbf{f} : X \to \mathbb{R}_\infty = \mathrm{Abb}(\mathbb{N}, \mathbb{R})$ gegeben durch $\mathbf{f}(x) = \{f_n(x)\}_{n\geq 0}$. und setze

$$\mathcal{T}_\infty = \bigcap_{n\geq 0} \sigma\Big(\bigcup_{m\geq n} f_m^{-1}(\mathcal{B})\Big) ,$$

also ist $\mathcal{T}_\infty \subset \mathcal{F}$ und $\mathcal{T}_\infty$ heißt die σ-Algebra der *terminalen Ereignisse* der Folge $\{f_n\}_{n\geq 0}$.

Satz 3.4 *Ist $f_0 \in \mathcal{L}^1(X, \mathcal{F}, \mu)$, so gilt*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f_k = \mathrm{E}_\mu(f_0|\mathbf{f}^{-1}(\mathcal{I}_S)) \quad \mu\text{-}\textit{fast sicher}\,.$$

Außerdem ist $\mathbf{f}^{-1}(\mathcal{I}_S) \subset \mathcal{T}_\infty$ und insbesondere ist

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f_k = \int f_0 \, d\mu \quad \mu\text{-}\textit{fast sicher}\,,$$

wenn $\mathcal{T}_\infty$ μ-trivial ist.

Beweis Sei $\nu \in \mathrm{P}(\mathbb{R}_\infty, \mathcal{B}_\infty)$ das Bildmaß von μ unter $\mathbf{f}$. Nach Satz 3.1 ist S maßtreu bezüglich ν und da $\mathrm{p}_0 \circ \mathbf{f} = f_0$, ist nach Satz 1.2 $\mathrm{p}_0 \in \mathcal{L}^1(\mathbb{R}_\infty, \mathcal{B}_\infty, \nu)$. Nach Satz 2.2 ist also

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathrm{p}_k = \mathrm{E}_\nu(\mathrm{p}_0|\mathcal{I}_S) \quad \nu\text{-fast sicher}\,,$$

da $\mathrm{p}_m = \mathrm{p}_0 \circ S^m$ für jedes $m \geq 0$. Folglich ist

$$\begin{aligned}
1 &= \nu\Big(\Big\{z \in \mathbb{R}_\infty : \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathrm{p}_k(z) = \mathrm{E}_\nu(\mathrm{p}_0|\mathcal{I}_S)(z)\Big\}\Big) \\
&= \mu\Big(\mathbf{f}^{-1}\Big(\Big\{z \in \mathbb{R}_\infty : \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathrm{p}_k(z) = \mathrm{E}_\nu(\mathrm{p}_0|\mathcal{I}_S)(z)\Big\}\Big)\Big) \\
&= \mu\Big(\Big\{x \in X : \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathrm{p}_k(\mathbf{f}(x)) = \mathrm{E}_\nu(\mathrm{p}_0|\mathcal{I}_S)(\mathbf{f}(x))\Big\}\Big) \\
&= \mu\Big(\Big\{x \in X : \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f_k(x) = \mathrm{E}_\nu(\mathrm{p}_0|\mathcal{I}_S)(\mathbf{f}(x))\Big\}\Big)
\end{aligned}$$

und daraus ergibt sich, dass

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f_k = \mathrm{E}_\nu(\mathrm{p}_0|\mathcal{I}_S) \circ \mathbf{f} \quad \mu\text{-fast sicher}\,.$$

Aber $\mathrm{E}_\nu(\mathrm{p}_0|\mathcal{I}_S) \circ \mathbf{f}$ ist $\mathbf{f}^{-1}(\mathcal{I}_S)$-messbar und für jedes $B \in \mathcal{I}_S$ ist nach Satz 1.2

$$\begin{aligned}
\int_{\mathbf{f}^{-1}(B)} \mathrm{E}_\nu(\mathrm{p}_0|\mathcal{I}_S) \circ \mathbf{f} \, d\mu &= \int_B \mathrm{E}_\nu(\mathrm{p}_0|\mathcal{I}_S) \, d\nu \\
&= \int_B \mathrm{p}_0 \, d\nu = \int_{\mathbf{f}^{-1}(B)} \mathrm{p}_0 \circ \mathbf{f} \, d\nu = \int_{\mathbf{f}^{-1}(B)} f_0 \, d\nu\,.
\end{aligned}$$

Damit ist $\mathrm{E}_\nu(\mathrm{p}_0|\mathcal{I}_S) \circ \mathbf{f} = \mathrm{E}_\mu(f|\mathbf{f}^{-1}(\mathcal{I}_S))$ μ-fast sicher, und dies zeigt, dass

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f_k = \mathrm{E}_\mu(f_0|\mathbf{f}^{-1}(\mathcal{I}_S)) \quad \mu\text{-fast sicher} .$$

Sei nun $B \in \mathcal{I}_S$; dann gilt $S^{-1}(B) = B$ und also ist $(S^n)^{-1}(B) = B$ für jedes $n \geq 0$. Folglich ist $\mathcal{I}_S \subset (S^n)^{-1}(\mathcal{B}_\infty)$ für jedes $n \geq 0$. Aber

$$\begin{aligned}(S^n)^{-1}(\mathcal{B}_\infty) &= (S^n)^{-1}\Big(\sigma\big(\bigcup_{k\geq 0} \mathrm{p}_k^{-1}(\mathcal{B})\big)\Big) \\ &= \sigma\Big((S^n)^{-1}\big(\bigcup_{k\geq 0} \mathrm{p}_k^{-1}(\mathcal{B})\big)\Big) = \sigma\Big(\bigcup_{m\geq n} \mathrm{p}_m^{-1}(\mathcal{B})\Big) ,\end{aligned}$$

da $(S^n)^{-1}(\mathrm{p}_k^{-1}(\mathcal{B})) = \mathrm{p}_{n+k}^{-1}(\mathcal{B})$, und daraus ergibt sich, dass für jedes $n \geq 0$

$$\begin{aligned}\mathbf{f}^{-1}(\mathcal{I}_S) \subset \mathbf{f}^{-1}((S^n)^{-1}(\mathcal{B}_\infty))) &= \mathbf{f}^{-1}\Big(\sigma\big(\bigcup_{m\neq n} \mathrm{p}_m^{-1}(\mathcal{B})\big)\Big) \\ &= \sigma\Big(\mathbf{f}^{-1}\big(\bigcup_{m\neq n} \mathrm{p}_m^{-1}(\mathcal{B})\big)\Big) = \sigma\Big(\bigcup_{m\geq n} f_m^{-1}(\mathcal{B})\Big) .\end{aligned}$$

Daher ist $\mathbf{f}^{-1}(\mathcal{I}_S) \subset \bigcap_{n\geq 0} \sigma\Big(\bigcup_{m\geq n} f_m^{-1}(\mathcal{B})\Big) = \mathcal{T}_\infty$. $\square$

Für jedes $n \geq 0$ sei $f_n : X \to \mathbb{R}$ eine $\mathcal{F}$-messbare Abbildung. Die Folge $\{f_n\}_{n\geq 0}$ heißt *unabhängig* bezüglich μ, wenn

$$\mu\Big(\bigcap_{n=k}^{\ell} f_n^{-1}(B_n)\Big) = \prod_{n=k}^{\ell} \mu(f_n^{-1}(B_n))$$

für alle $B_k, \ldots, B_\ell \in \mathcal{E}$ und alle $k, \ell \in I$ mit $k \leq \ell$. Die Folge heißt *identisch verteilt*, wenn $\mu(f_n^{-1}(B)) = \mu(f_0^{-1}(B))$ für alle $B \in \mathcal{B}$, $n \geq 0$. Es ist klar, dass eine stationäre Folge identisch verteilt ist; umgekehrt ist eine unabhängige identisch verteilte Folge stationär.

Satz 3.5 (Null-Eins-Gesetz von Kolmogorov) *Ist $\{f_n\}_{n\geq 0}$ unabhängig, so ist die σ-Algebra $\mathcal{T}_\infty$ der terminalen Ereignisse μ-trivial.*

Beweis Für jedes $m \geq 0$ setze $\mathcal{T}_m = \sigma\big(\bigcup_{k>m} f_k^{-1}(\mathcal{B})\big)$ und $\mathcal{A}_m = \sigma\big(\bigcup_{k=0}^{m} f_k^{-1}(\mathcal{B})\big)$ und sei $\mathcal{A} = \bigcup_{m\geq 0} \mathcal{A}_m$. Dann ist $\mathcal{T}_\infty = \bigcap_{m\geq 0} \mathcal{T}_m$, und da $f_m^{-1}(\mathcal{B}) \subset \mathcal{A}_m$ und $\mathcal{T}_\infty \subset \sigma\big(\bigcup_{m\geq 0} f_m^{-1}(\mathcal{B})\big)$, ist ferner $\mathcal{T}_\infty \subset \sigma(\mathcal{A})$. Für jedes $m \geq 0$ setze auch

$$\mathcal{D}_m = \Big\{ \bigcap_{k=m+1}^{\ell} f_k^{-1}(B_k) : B_{m+1}, \ldots, B_\ell \in \mathcal{B},\ \ell \geq m+1 \Big\}$$

und $\mathcal{C}_m = \{\bigcap_{k=0}^m f_k^{-1}(B_k) : B_0, \ldots, B_m \in \mathcal{B}\}$. Dann gilt $\mu(C \cap D) = \mu(C)\mu(D)$ für alle $C \in \mathcal{C}_m$, $D \in \mathcal{D}_m$, da die Folge $\{f_n\}_{n\geq 0}$ unabhängig ist, und folglich ist $\mu(A \cap B) = \mu(A)\mu(B)$ für alle $A \in \mathcal{A}_m$, $B \in \mathcal{T}_m$, da $\mathcal{C}_m$ und $\mathcal{D}_m$ durchschnittsstabil sind und $\mathcal{A}_m = \sigma(\mathcal{C}_m)$, $\mathcal{T}_m = \sigma(\mathcal{D}_m)$. Sei nun $B \in \mathcal{T}_\infty$ und $\varepsilon > 0$; da $B \in \sigma(\mathcal{A})$ und $\mathcal{A}$ eine Algebra ist, gibt es ein $A \in \mathcal{A}$ mit $\mu(A \triangle B) < \varepsilon/2$. Aber $A \in \mathcal{A}_n$ für ein $n \geq 0$ und daher ist $\mu(A \cap B) = \mu(A)\mu(B)$, da $B \in \mathcal{T}_n$. Also ist

$$\begin{aligned} 0 &\leq \mu(B) - \mu(B)\mu(B) = \mu(B) - \mu(A \cap B) + \mu(A)\mu(B) - \mu(B)\mu(B) \\ &= \mu(B \setminus A) + (\mu(A) - \mu(B))\mu(B) \leq \mu(A \triangle B) + \mu(A \triangle B)\mu(B) < \varepsilon \,, \end{aligned}$$

und da $\varepsilon > 0$ beliebig ist, ist dann $\mu(B) = \mu(B)\mu(B)$, d.h., $\mu(B)$ ist entweder 0 oder 1. □

Satz 3.6 (Starkes Gesetz der großen Zahlen) *Sei $\{f_n\}_{n\geq 0}$ unabhängig und identisch verteilt mit $f_0 \in \mathcal{L}^1(X, \mathcal{F}, \mu)$. Dann gilt*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f_k = \int f_0 \, d\mu \quad \mu\text{-fast sicher}\,.$$

Beweis Dies folgt unmittelbar aus Satz 3.2 und Satz 3.3. □

4 Ergodische und mischende Abbildungen

Im Folgenden sei $(X, \mathcal{F}, \mu)$ ein Wahrscheinlichkeitsraum und $T : X \to X$ eine messbare Abbildung, die maßtreu bezüglich μ ist. Für $k \geq 0$ wird meistens T^{-k} statt $(T^k)^{-1}$ geschrieben.

Eine σ-Algebra $\mathcal{E} \subset \mathcal{F}$ heißt *μ-trivial*, wenn $\mu(B)$ entweder 0 oder 1 ist für jedes $B \in \mathcal{E}$. Ist die σ-Algebra $\mathcal{E}$ μ-trivial, so ist $\mathrm{E}_\mu(f|\mathcal{E}) = \int f \, d\mu$ μ-fast sicher für jedes $f \in \mathcal{L}^1(X, \mathcal{F}, \mu)$, da die konstante Abbildung $\int f \, d\mu$ $\mathcal{E}$-messbar ist und für jedes $B \in \mathcal{F}$ mit $\mu(B)$ entweder 0 oder 1 gilt $\int_B (\int f \, d\mu) \, d\mu = \int_B f \, d\mu$.

Die maßtreue Abbildung T heißt *ergodisch bezüglich* μ, wenn die σ-Algebra $\mathcal{I}_T$ μ-trivial ist. Ist T ergodisch, so folgt aus Satz 2.2, dass für jedes $f \in \mathcal{L}^1(X, \mathcal{F}, \mu)$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k = \int f \, d\mu \quad \mu\text{-fast sicher} .$$

Satz 4.1 *Die folgenden Aussagen sind äquivalent für T:*

(1) T ist ergodisch bezüglich μ.

(2) Für jedes $B \in \mathcal{F}$ mit $\mu(T^{-1}(B) \triangle B) = 0$ ist $\mu(B)$ entweder 0 oder 1.

(3) Für jedes $B \in \mathcal{F}$ mit $\mu(B) > 0$ gilt $\mu\big(\bigcup_{n\geq 0} T^{-n}(B)\big) = 1$.

(4) Für alle A, $B \in \mathcal{F}$ mit $\mu(A) > 0$ und $\mu(B) > 0$ gibt es ein $n \geq 0$, so dass $\mu(T^{-n}(A) \cap B) > 0$.

(5) Für alle A, $B \in \mathcal{F}$ gilt $\lim_{n\to\infty} n^{-1} \sum_{k=0}^{n-1} \mu(T^{-k}(A) \cap B) = \mu(A)\mu(B)$.

(6) Ist $f \in \mathrm{M}(X, \mathcal{F})$ mit $f \circ T = f$, so gibt es ein $c \in \mathbb{R}$, so dass $f = c$ μ-fast sicher.

(7) Ist $f \in \mathrm{M}(X, \mathcal{F})$ mit $f \circ T = f$ μ-fast sicher, so gibt es ein $c \in \mathbb{R}$, so dass $f = c$ μ-fast sicher.

Beweis (1) $\Rightarrow$ (2): Sei $B \in \mathcal{F}$ mit $\mu(T^{-1}(B) \triangle B) = 0$ und setze

$$A = \{x \in X : T^n(x) \in B \text{ für unendlich viele } n \geq 0\} .$$

Es ist klar, dass $T^{-1}(A) = A$, d.h., $A \in \mathcal{I}_T$. Nun gilt

$$\begin{aligned} A \setminus B &\subset \bigcup_{n\geq 0} \{x \in X : T^{n+1}(x) \in B,\, T^n(x) \notin B\} \\ &= \bigcup_{n\geq 0} T^{-n}(\{x \in X \setminus B : T(x) \in B\}) \subset \bigcup_{n\geq 0} T^{-n}(T^{-1}(B) \triangle B) , \end{aligned}$$

und damit ist $\mu(A \setminus B) = 0$, da $\mu(T^{-n}(T^{-1}(B) \triangle B)) = \mu(T^{-1}(B) \triangle B) = 0$ für jedes $n \geq 0$. Nach Satz 2.1 ist aber $\mu(A \cap B) = \mu(B)$ und daher $\mu(B \setminus A) = 0$. Daraus folgt, dass $\mu(B \triangle A) = \mu(A \setminus B) + \mu(B \setminus A) = 0$. Also ist $\mu(B) = \mu(A)$ und $\mu(A)$ ist entweder 0 oder 1, da $A \in \mathcal{I}_T$.

(2) ⇒ (3): Sei $B \in \mathcal{F}$ mit $\mu(B) > 0$ und setze $A = \bigcup_{n\geq 0} T^{-n}(B)$; insbesondere ist $\mu(A) > 0$, da $B \subset A$. Nun ist $T^{-1}(A) = \bigcup_{n\geq 1} T^{-n}(B) \subset A$ und folglich gilt $\mu(T^{-1}(A) \triangle A) = \mu(A \setminus T^{-1}(A)) = \mu(A) - \mu(T^{-1}(A)) = 0$. Also ist $\mu(A)$ entweder 0 oder 1 und damit ist $\mu(A) = 1$.

(3) ⇒ (4): Seien A, $B \in \mathcal{F}$ mit $\mu(A) > 0$ und $\mu(B) > 0$. Dann ist

$$0 < \mu(B) = \mu\Big(B \cap \bigcup_{n\geq 0} T^{-1}(A)\Big) = \mu\Big(\bigcup_{n\geq 0} (T^{-n}(A) \cap B)\Big) ,$$

da $\mu\big(\bigcup_{n\geq 0} T^{-n}(A)\big) = 1$, und damit gibt es ein $n \geq 0$, so dass $\mu(T^{-n}(A) \cap B) > 0$.

(4) ⇒ (1): Sei $B \in \mathcal{I}_T$; für jedes $n \geq 0$ gilt dann $T^{-n}(B) = B$ und damit $\mu(T^{-n}(B) \cap (X \setminus B)) = 0$. Daraus folgt, dass $\mu(B) = 0$ oder $\mu(X \setminus B) = 0$, d.h., $\mu(B)$ ist entweder 0 oder 1.

(5) ⇒ (1): Sei $B \in \mathcal{I}_T$; dann ist $T^{-k}(B) = B$ für alle $k \geq 0$ und folglich ist

$$\frac{1}{n} \sum_{k=0}^{n-1} \mu(T^{-k}(B) \cap B) = \mu(B)$$

für jedes $n \geq 1$. Daraus ergibt sich, dass $\mu(B) = \mu(B)\mu(B)$, d.h., $\mu(B)$ ist entweder 0 oder 1.

(1) ⇒ (5): Sei $A \in \mathcal{F}$; da T ergodisch ist, ist $\mathrm{E}_\mu(\chi_A|\mathcal{I}_T) = \mu(A)$ μ-fast sicher und folglich gilt nach Satz 2.2, dass $\lim_{n\to\infty} n^{-1} \sum_{k=0}^{n-1} \chi_A \circ T^k = \mu(A)$ μ-fast sicher. Aber $0 \leq n^{-1} \sum_{k=0}^{n-1} \chi_A \circ T^k \leq 1$ für jedes $n \geq 1$ und daraus ergibt sich nach dem Satz von der dominierten Konvergenz, dass

$$\lim_{n\to\infty} \int_B \Big(\frac{1}{n} \sum_{k=0}^{n-1} \chi_A \circ T^k\Big)\, d\mu = \int_B \mu(A)\, d\mu ,$$

d.h., $\lim\limits_{n\to\infty} n^{-1} \sum\limits_{k=0}^{n-1} \mu(T^{-k}(A) \cap B) = \mu(A)\mu(B)$ für jedes $B \in \mathcal{F}$.

(7) ⇒ (6): Dies ist klar.

(6) ⇒ (1): Sei $B \in \mathcal{I}_T$; dann ist $\chi_B \circ T = \chi_B$ und damit gibt es ein $c \in \mathbb{R}$, so dass $\chi_B = c$ μ-fast sicher. Aber dann muss c entweder 0 oder 1 sein und folglich ist $\mu(B) = \int \chi_B\, d\mu = c$ entweder 0 oder 1.

(2) ⇒ (7): Sei $f : X \to \mathbb{R}$ $\mathcal{F}$-messbar mit $f \circ T = f$ μ-fast sicher und setze $A = \{x \in X : (f \circ T)(x) = f(x)\}$, also ist $\mu(A) = 1$. Für jedes $\alpha \in \mathbb{R}$ sei

$B_\alpha = \{x \in X : f(x) \le \alpha\}$; dann gilt $B_\alpha \cap A = T^{-1}(B_\alpha) \cap A$ und damit $\mu(T^{-1}(B_\alpha) \bigtriangleup B_\alpha) = 0$. Für jedes $\alpha \in \mathbb{R}$ ist also $\mu(B_\alpha)$ entweder 0 oder 1, und daher gibt es nach Lemma 1.3 ein $c \in \mathbb{R}$, so dass $f = c$ μ-fast sicher. □

Satz 4.2 *Sei Y ein kompakter metrischer Raum, sei $\mathcal{B}_Y$ die Borelsche σ-Algebra und sei $\nu \in \mathrm{P}(Y, \mathcal{B}_Y)$ mit $\nu(U) > 0$ für jede nichtleere offene Teilmenge U von Y. Sei $T : Y \to Y$ eine stetige Abbildung, die ergodisch bezüglich ν ist. Dann ist*

$$\nu(\{y \in Y : \mathrm{O}_T(y) \text{ ist eine dichte Teilmenge von } Y\}) = 1\ ,$$

wobei $\mathrm{O}_T(y) = \{z \in Y : z = T^n(y) \text{ für ein } n \ge 0\}$.

Beweis Übung. □

Nach Satz 4.1 ist T genau dann ergodisch, wenn

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu(T^{-k}(A) \cap B) = \mu(A)\mu(B)$$

für alle A, $B \in \mathcal{F}$. Nun werden stärkere Eigenschaften dieser Art betrachtet. Zunächst heißt T *schwach-mischend*, wenn

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} |\mu(T^{-k}(A) \cap B) - \mu(A)\mu(B)| = 0$$

für alle A, $B \in \mathcal{F}$, und T heißt *stark-mischend*, wenn

$$\lim_{n\to\infty} \mu(T^{-n}(A) \cap B) = \mu(A)\mu(B)$$

für alle A, $B \in \mathcal{F}$.

Es ist klar, dass jede schwach-mischende Abbildung ergodisch ist; ferner ist jede stark-mischende Abbildung schwach-mischend. (Ist $\{c_n\}_{n\ge1}$ eine Folge aus $\mathbb{R}$ mit $\lim_{n\to\infty} c_n = 0$, so gilt auch $\lim_{n\to\infty} n^{-1} \sum_{k=1}^n |c_k| = 0$.)

Für jede Folge reeller Zahlen $\{a_k\}_{k\ge0}$ setze für jedes $n \ge 1$

$$\sigma_n(\{a_k\}_{k\ge0}) = \frac{1}{n} \sum_{k=0}^{n-1} a_k\ .$$

Für $J \subset \mathbb{N}$ sei $\{\delta_n^J\}_{n\ge0}$ die Folge mit $\delta_n^J = \chi_J(n)$ für jedes $n \in \mathbb{N}$. Die Bedingung in der Definition einer schwach-mischenden Abbildung kann man mit Hilfe des folgenden Lemmas umformulieren:

Lemma 4.1 *Sei $\gamma : \mathbb{R}^+ \to \mathbb{R}^+$ eine stetige Abbildung mit $\gamma^{-1}(\{0\}) = \{0\}$ und sei $\{a_n\}_{n\geq 0}$ eine beschränkte Folge aus $\mathbb{R}^+$. Dann sind äquivalent:*

(1) Es gilt $\lim\limits_{n\to\infty} \sigma_n(\{a_k\}_{k\geq 0}) = 0$.

(2) Es gilt $\lim\limits_{n\to\infty} \sigma_n(\{\gamma(a_k)\}_{k\geq 0}) = 0$.

(3) Es gibt $J \subset \mathbb{N}$ mit $\lim\limits_{n\to\infty} \sigma_n(\{\delta_k^J\}_{k\geq 0}) = 0$ *und* $\lim\limits_{n\to\infty} a_n(1-\delta_n^J) = 0$.

Beweis (3) $\Rightarrow$ (1): Sei $\alpha \in \mathbb{R}^+$ eine obere Schranke der Folge $\{a_n\}_{n\geq 0}$ und sei $J \subset \mathbb{N}$ mit $\lim_{n\to\infty} \sigma_n(\{\delta_k^J\}_{k\geq 0}) = 0$ und $\lim_{n\to\infty} a_n(1-\delta_n^J) = 0$; dann gilt auch $\lim_{n\to\infty} \sigma_n(\{a_k(1-\delta_k^J)\}_{k\geq 0}) = 0$. Sei $\varepsilon > 0$; also gibt es ein $N \geq 1$, so dass $\sigma_n(\{a_k(1-\delta_k^J)\}_{k\geq 0}) < \varepsilon$ und $\sigma_n(\{\delta_k^J\}_{k\geq 0}) < \varepsilon$ für alle $n \geq N$. Folglich gilt

$$\begin{aligned}\sigma_n(\{a_k\}_{k\geq 0}) &= \sigma_n(\{a_k(1-\delta_k^J)\}_{k\geq 0}) + \sigma_n(\{a_k\delta_k^J\}_{k\geq 0}) \\ &\leq \sigma_n(\{a_k(1-\delta_k^J)\}_{k\geq 0}) + \alpha\sigma_n(\{\delta_k^J\}_{k\geq 0}) < (1+\alpha)\varepsilon\end{aligned}$$

für alle $n \geq N$, und daher ist $\lim\limits_{n\to\infty} \sigma_n(\{a_k\}_{k\geq 0}) = 0$.

(1) $\Rightarrow$ (3): Für jedes $m \geq 1$ sei $J_m = \{n \in \mathbb{N} : a_n \geq 1/m\}$; für alle $n \geq 1$ gilt dann $\sigma_n(\{a_k\}_{k\geq 0}) \geq m^{-1}\sigma_n(\{\delta_k^{J_m}\}_{k\geq 0})$ und damit ist $\lim_{n\to\infty} \sigma_n(\{\delta_k^{J_m}\}_{k\geq 0}) = 0$ für jedes $m \geq 1$. Es gibt also eine streng monoton wachsende Folge $\{\ell_m\}_{m\geq 1}$ aus $\mathbb{N}$, so dass $\sigma_n(\{\delta_k^{J_m}\}_{k\geq 0}) < m^{-1}$ für alle $n \geq \ell_m$ und alle $m \geq 1$. Setze

$$J = \bigcup_{m\geq 1} \{\ell_m, \ell_m + 1, \ldots, \ell_{m+1} - 1\} \cap J_m \,.$$

Sei nun $n \in \mathbb{N}$ mit $\ell_m \leq n < \ell_{m+1}$; da die Folge $\{J_n\}_{n\geq 0}$ monoton wachsend ist, ist $J \cap \{0, \ldots, n-1\} \subset J_m \cap \{0, \ldots, n-1\}$ und da $n \geq \ell_m$, ergibt sich daraus, dass $\sigma_n(\{\delta_k^J\}_{k\geq 0}) \leq \sigma_n(\{\delta_k^{J_m}\}_{k\geq 0}) < m^{-1}$. Folglich gilt $\lim_{n\to\infty} \sigma_n(\{\delta_k^J\}_{k\geq 0}) = 0$. Sei andererseits $n \notin J$ mit $\ell_m \leq n < \ell_{m+1}$; dann ist $n \notin J_m$ und damit $a_n < m^{-1}$. Daher ist $\lim_{n\to\infty} a_n(1-\delta_n^J) = 0$.

(2) $\Leftrightarrow$ (3): Nach dem Beweis für (1) $\Leftrightarrow$ (3) angewendet auf die beschränkte Folge $\{\gamma(a_n)\}_{n\geq 0}$ gilt (2) genau dann, wenn es ein $J \subset \mathbb{N}$ mit $\lim_{n\to\infty} \sigma_n(\{\delta_k^J\}_{k\geq 0}) = 0$ und $\lim_{n\to\infty} \gamma(a_n)(1-\delta_n^J) = 0$. Aber $\lim_{n\to\infty} \gamma(a_n)(1-\delta_n^J) = 0$ gilt genau dann, wenn $\lim_{n\to\infty} a_n(1-\delta_n^J) = 0$. $\square$

Satz 4.3 *Sei $p > 0$. Dann sind äquivalent:*

(1) Die Abbildung T ist schwach-mischend.

(2) Für alle $A, B \in \mathcal{F}$ gilt

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} |\mu(T^{-k}(A) \cap B) - \mu(A)\mu(B)|^p = 0 \,.$$

(3) Für jedes A, $B \in \mathcal{F}$ gibt es $J \subset \mathbb{N}$ mit $\lim_{n\to\infty} \sigma_n(\{\delta^J\}_{k\ge 0}) = 0$, so dass

$$\lim_{n\to\infty, n\notin J} \mu(T^{-n}(A) \cap B) = \mu(A)\mu(B) \; .$$

Beweis Dies folgt unmittelbar aus Lemma 4.1. □

Lemma 4.2 *Sei $\mathcal{A}$ eine Algebra mit $\mathcal{F} = \sigma(\mathcal{A})$. Dann gilt:*

(1) T ist ergodisch genau dann, wenn für alle A, $B \in \mathcal{A}$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu(T^{-k}(A) \cap B) = \mu(A)\mu(B) \; .$$

(2) T ist schwach-mischend genau dann, wenn für alle A, $B \in \mathcal{A}$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} |\mu(T^{-k}(A) \cap B) - \mu(A)\mu(B)| = 0 \; .$$

(3) T ist stark-mischend genau dann, wenn für alle A, $B \in \mathcal{A}$

$$\lim_{n\to\infty} \mu(T^{-n}(A) \cap B) = \mu(A)\mu(B) \; .$$

Beweis Seien A, $B \in \mathcal{F}$ und $\varepsilon > 0$; dann gibt es A', $B' \in \mathcal{A}$ mit $\mu(A \triangle A') < \varepsilon/2$ und $\mu(B \triangle B') < \varepsilon/2$. Da aber $(C \cap D) \triangle (C' \cap D') \subset (C \triangle C') \cup (D \triangle D')$ für alle C, C', D, $D' \in \mathcal{F}$, ist für alle $k \ge 0$

$$\begin{aligned}(T^{-k}(A) \cap B) \triangle (T^{-k}(A') \cap B') &\subset (T^{-k}(A) \triangle T^{-k}(A')) \cup (B \triangle B') \\ &= (T^{-k}(A \triangle A')) \cup (B \triangle B')\end{aligned}$$

und daraus ergibt sich, dass $\mu((T^{-k}(A) \cap B) \triangle (T^{-k}(A') \cap B')) < \varepsilon$. Folglich ist $|\mu((T^{-k}(A) \cap B) - \mu(T^{-k}(A') \cap B')| < \varepsilon$ für alle $k \ge 0$. Andererseits ist

$$\begin{aligned}|\mu(A)\mu(B) - \mu(A')\mu(B')| &= |\mu(A)(\mu(B) - \mu(B')) - \mu(B')(\mu(A) - \mu(A'))| \\ &\le |\mu(B) - \mu(B')| + |\mu(A) - \mu(A')| \le \mu(A \triangle A') + \mu(B \triangle B') < \varepsilon \; .\end{aligned}$$

Aus diesen Abschätzungen folgen unmittelbar alle drei Teile des Lemmas. □

Lemma 4.3 *Sei* B *entweder* $\mathrm{B}(X, \mathcal{F})$ *oder* $\mathrm{B}_{\mathbb{C}}(X, \mathcal{F})$.

(1) T ist ergodisch genau dann, wenn für alle f, $g \in \mathrm{B}$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \int (f \circ T^k) g \, d\mu = \int f \, d\mu \int g \, d\mu \; .$$

(2) T ist schwach-mischend genau dann, wenn für alle f, $g \in \mathrm{B}$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \Big| \int (f \circ T^k) g \, d\mu - \int f \, d\mu \int g \, d\mu \Big| = 0 \, .$$

(3) T ist stark-mischend genau dann, wenn für alle f, $g \in \mathrm{B}$

$$\lim_{n\to\infty} \int (f \circ T^n) g \, d\mu = \int f \, d\mu \int g \, d\mu \, .$$

Beweis Übung. □

Lemma 4.4 *Seien $(X_1, \mathcal{F}_1, \mu_1)$ und $(X_2, \mathcal{F}_2, \mu_2)$ Wahrscheinlichkeitsräume und für $i = 1, 2$ sei $T_i : X_i \to T_i$ maßtreu bezüglich μ_i. Dann ist die durch*

$$(T_1 \times T_2)(x_1, x_2) = (T_1(x_1), T_2(x_2))$$

definierte Abbildung $T_1 \times T_2 : X_1 \times X_2 \to X_1 \times X_2$ maßtreu bezüglich $\mu_1 \times \mu_2$.

Beweis Sei $\mathcal{R}$ die Menge aller Teilmengen von $X_1 \times X_2$ der Form $B_1 \times B_2$ mit $B_1 \in \mathcal{F}_1$, $B_2 \in \mathcal{F}_2$; also ist $\mathcal{F}_1 \times \mathcal{F}_2 = \sigma(\mathcal{R})$. Für alle $B_1 \in \mathcal{F}_1$, $B_2 \in \mathcal{F}_2$ gilt $(T_1 \times T_2)^{-1}(B_1 \times B_2) = T_1^{-1}(B_1) \times T_2^{-1}(B_2) \in \mathcal{F}_1 \times \mathcal{F}_2$ und daraus folgt, dass

$$(T_1 \times T_2)^{-1}(\mathcal{F}_1 \times \mathcal{F}_2) = (T_1 \times T_2)^{-1}(\sigma(\mathcal{R})) = \sigma((T_1 \times T_2)^{-1}(\mathcal{R}))) \subset \mathcal{F}_1 \times \mathcal{F}_2 \, .$$

Für alle $B_1 \in \mathcal{F}_1$, $B_2 \in \mathcal{F}_2$ gilt ferner

$$\begin{aligned} (\mu_1 \times \mu_2))((T_1 \times T_2)^{-1}(B_1 \times B_2)) &= (\mu_1 \times \mu_2)(T_1^{-1}(B_1) \times T_2^{-1}(B_2)) \\ &= \mu_1(T_1^{-1}(B_1))\mu_2(T_2^{-1}(B_2)) = \mu_1(B_1)\mu_2(B_2) \\ &= (\mu_1 \times \mu_2)(B_1 \times B_2) \, ; \end{aligned}$$

d.h., $(\mu_1 \times \mu_2))((T_1 \times T_2)^{-1}(R)) = (\mu_1 \times \mu_2)(R)$ für alle $R \in \mathcal{R}$. Daraus ergibt sich, dass $(\mu_1 \times \mu_2))((T_1 \times T_2)^{-1}(F)) = (\mu_1 \times \mu_2)(F)$ für alle $F \in \mathcal{F}_1 \times \mathcal{F}_2$, da $\mathcal{R}$ durchschnittsstabil ist. Also ist $T_1 \times T_2$ maßtreu bezüglich $\mu_1 \times \mu_2$. □

Satz 4.4 *Für die bezüglich μ maßtreue Abbildung $T : X \to X$ sind äquivalent:*

(1) T ist schwach-mischend bezüglich μ.

(2) $T \times T$ ist ergodisch bezüglich $\mu \times \mu$.

(3) $T \times T$ ist schwach-mischend bezüglich $\mu \times \mu$.

Beweis (1) $\Rightarrow$ (3): Für $i = 1, 2$ seien $A_i, B_i \in \mathcal{F}$; nach Satz 4.3 gibt es $J_i \subset \mathbb{N}$ mit $\lim_{n\to\infty} \sigma_n(\{\delta_k^{J_i}\}_{k\geq 0}) = 0$, so dass $\lim_{n\to\infty, n\notin J_i} \mu(T^{-n}(A_i) \cap B_i) = \mu(A_i)\mu(B_i)$. Setze $J = J_1 \cup J_2$; dann ist $\lim_{n\to\infty} \sigma_n(\{\delta_k^{J_1}\}_{k\geq 0}) = 0$ und

$$\begin{aligned}\lim_{n\to\infty, n\notin J} &(\mu\times\mu)((T\times T)^{-n}(A_1\times A_2)\cap(B_1\times B_2))\\ &= \lim_{n\to\infty, n\notin J} \mu(T^{-n}(A_1)\cap B_1)\mu(T^{-n}(A_2)\cap B_2)\\ &= \mu(A_1)\mu(B_1)\mu(A_2)\mu(B_2) = (\mu\times\mu)(A_1\times A_2)(\mu\times\mu)(B_1\times B_2)\,.\end{aligned}$$

Daraus ergibt sich nach Lemma 4.2, dass

$$\lim_{n\to\infty} \sum_{k=0}^{n-1} |(\mu\times\mu)((T\times T)^{-k}(C_1)\cap C_2) - (\mu\times\mu)(C_1)(\mu\times\mu)(C_2)| = 0$$

für alle C_1, $C_2 \in \mathcal{R}$ (mit $\mathcal{R}$ wie im Beweis für Lemma 4.4) und folglich gilt dies auch für alle C_1, $C_2 \in \mathcal{A}$, wobei $\mathcal{A}$ die Menge aller Teilmengen von $X\times X$ ist, die eine Darstellung als endliche disjunkte Vereinigung von Elementen aus $\mathcal{R}$ besitzen. Aber $\mathcal{A}$ ist eine Algebra mit $\sigma(\mathcal{A}) = \mathcal{F}$ und daher ist nach Lemma 4.2 $T\times T$ schwach-mischend bezüglich $\mu\times\mu$.

(3) $\Rightarrow$ (2): Dies ist klar.

(2) $\Rightarrow$ (1): Seien A, $B \in \mathcal{F}$; dann gilt

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\mu(T^{-k}(A)\cap B) &= \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}(\mu\times\mu)((T\times T)^{-k}(A\times X)\cap(X\times B))\\ &= (\mu\times\mu)((A\times X)\cap(X\times B)) = \mu(A)\mu(B)\end{aligned}$$

und genauso gilt auch

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{n}&\sum_{k=0}^{n-1}\big(\mu(T^{-k}(A)\cap B)\big)^2\\ &= \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}(\mu\times\mu)((T\times T)^{-k}(A\times A)\cap(B\times B))\\ &= (\mu\times\mu)((A\times A)\cap(B\times B)) = \big(\mu(A)\mu(B)\big)^2\,.\end{aligned}$$

Daraus ergibt sich, dass

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}|\mu(T^{-k}(A)\cap B) - \mu(A)\mu(B)|^2 &= \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\big(\mu(T^{-k}(A)\cap B)\big)^2\\ &- 2\mu(A)\mu(B)\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\mu(T^{-k}(A)\cap B) + \big(\mu(A)\mu(B)\big)^2 = 0\end{aligned}$$

und folglich ist nach Satz 4.3 T schwach-mischend bezüglich μ. $\square$

Im folgenden Satz sei (E, ϱ) ein σ-kompakter metrischer Raum mit $\mathcal{E}$ der Borel-σ-Algebra auf E und sei I entweder $\mathbb{Z}$ oder $\mathbb{N}$. Sei $\nu \in \mathrm{P}(E, \mathcal{E})$. Nach Satz 3.2 gibt es dann ein eindeutiges Maß $\nu_\infty \in \mathrm{P}(E_\infty, \mathcal{E}_\infty)$, so dass

$$\nu_\infty\Big(\bigcap_{n=k}^{\ell} \mathrm{p}_n^{-1}(B_n)\Big) = \prod_{n=k}^{\ell} \nu(B_n)$$

für alle $B_k, \ldots, B_\ell \in \mathcal{E}$ und alle $k, \ell \in I$ mit $k \leq \ell$. Ferner ist S maßtreu bezüglich ν_∞.

Satz 4.5 *Die Abbildung S ist stark-mischend bezüglich ν_∞.*

Beweis Für $J \in \mathcal{J}$ setze $\mathcal{E}_J = \sigma\big(\bigcup_{m \in J} \mathrm{p}_m^{-1}(\mathcal{E})\big)$ und

$$\mathcal{D}_J = \Big\{ \bigcap_{m \in J} \mathrm{p}_m^{-1}(B_m) : B_m \in \mathcal{E} \text{ für jedes } m \in J \Big\} ;$$

dann ist $\mathcal{D}_J$ durchschnittsstabil und $\mathcal{E}_J = \sigma(\mathcal{D}_J)$. Ist $J = \{k, \ldots, \ell\}$, so ist $S^{-n}(\mathcal{E}_J) \subset \mathcal{E}_{J+n}$ für alle $n \geq 0$ mit $J + n = \{k+n, \ldots, \ell+n\}$. Für $J, K \in \mathcal{J}$ mit $J \cap K = \varnothing$ gilt ferner $\nu_\infty(A \cap B) = \nu_\infty(A)\nu_\infty(B)$ für alle $A \in \mathcal{D}_J$, $B \in \mathcal{D}_K$ und damit auch für alle $A \in \mathcal{E}_J$, $B \in \mathcal{E}_K$. Setze nun $\mathcal{A} = \bigcup_{J \in \mathcal{J}} \mathcal{E}_J$, also ist $\mathcal{A}$ eine Algebra mit $\sigma(\mathcal{A}) = \mathcal{F}$. Sind $A, B \in \mathcal{A}$, so gibt es $J, K \in \mathcal{J}$ mit $A \in \mathcal{E}_J$, $B \in \mathcal{E}_K$ und dann gibt es $N \geq 1$, so dass $\nu_\infty(S^{-n}(A) \cap B) = \nu_\infty(A)\nu_\infty(B)$ für alle $n \geq N$, (da es $N \geq 1$ gibt mit $(J+n) \cap K = \varnothing$ für alle $n \geq N$). Dies zeigt, dass $\lim_{n\to\infty} \nu_\infty(S^{-n}(A) \cap B) = \nu_\infty(A)\nu_\infty(B)$ für alle $A, B \in \mathcal{A}$, und daraus folgt nach Lemma 4.2, dass S stark-mischend bezüglich ν_∞ ist. $\square$

Im Folgenden sei nun $(X, \mathcal{F})$ ein Messraum und sei $T : X \to X$ eine messbare Abbildung. Ein Maß $\mu \in \mathrm{P}(X, \mathcal{F})$ heißt *T-invariant*, wenn T maßtreu bezüglich μ ist, und die Menge aller T-invarianten Elemente in $\mathrm{P}(X, \mathcal{F})$ wird mit $\mathrm{P}_T(X, \mathcal{F})$ bezeichnet. Man beachte, dass $\mathrm{P}_T(X, \mathcal{F})$ leer sein kann.

Lemma 4.5 *Die Menge $\mathrm{P}_T(X, \mathcal{F})$ ist eine konvexe Teilmenge von $\mathrm{P}(X, \mathcal{F})$: Für alle $\mu_1, \mu_2 \in \mathrm{P}_T(X, \mathcal{F})$ und alle $0 \leq t \leq 1$ ist $t\mu_1 + (1-t)\mu_2 \in \mathrm{P}_T(X, \mathcal{F})$.*

Beweis Seien $\mu_1, \mu_2 \in \mathrm{P}_T(X, \mathcal{F})$ und $0 \leq t \leq 1$; setze $\mu = t\mu_1 + (1-t)\mu_2$. Dann ist $\mu \in \mathrm{P}(X, \mathcal{F})$ und für jedes $B \in \mathcal{F}$ gilt

$$\mu(T^{-1}(B)) = t\mu_1(T^{-1}(B)) + (1-t)\mu_2(T^{-1}(B)) = t\mu_1(B) + (1-t)\mu_2(B) = \mu(B)$$

und damit ist μ T-invariant. $\square$

Ist Q eine konvexe Teilmenge von $\mathrm{P}(X, \mathcal{F})$, so heißt $\mu \in Q$ *extremaler Punkt* von Q, wenn aus $\mu = t\mu_1 + (1-t)\mu_2$ mit $\mu_1, \mu_2 \in Q$ und $0 < t < 1$ stets folgt, dass $\mu_1 = \mu_2$ (und damit $\mu = \mu_1 = \mu_2$).

Satz 4.6 *Ein Maß $\mu \in \mathrm{P}_T(X, \mathcal{F})$ ist ein extremaler Punkt von $\mathrm{P}_T(X, \mathcal{F})$ genau dann, wenn T ergodisch bezüglich μ ist.*

Beweis Nehme zunächst an, dass T nicht ergodisch ist. Dann gibt es ein $B \in \mathcal{I}_T$ mit $0 < \mu(B) < 1$. Setze $t = \mu(B)$, also ist $0 < t < 1$, und definiere Maße μ_1, $\mu_2 \in \mathrm{P}(X, \mathcal{F})$ durch $\mu_1(A) = t^{-1}\mu(A \cap B)$ und $\mu_2(A) = (1-t)^{-1}\mu(A \cap (X \setminus B))$ für alle $A \in \mathcal{F}$. Für alle $A \in \mathcal{F}$ gilt dann

$$\begin{aligned}\mu_1(T^{-1}(A)) = t^{-1}\mu(T^{-1}(A) \cap B) &= t^{-1}\mu(T^{-1}(A) \cap T^{-1}(B))\\ &= t^{-1}\mu(T^{-1}(A \cap B)) = t^{-1}\mu(A \cap B) = \mu_1(A)\end{aligned}$$

und damit ist $\mu_1 \in \mathrm{P}_T(X, \mathcal{F})$. Genauso ist $\mu_2 \in \mathrm{P}_T(X, \mathcal{F})$. Aber ist es klar, dass $\mu = t\mu_1 + (1-t)\mu_2$ und $\mu_1 \neq \mu_2$, da $\mu_1(B) = 1 \neq 0 = \mu_2(B)$, und folglich ist μ kein extremaler Punkt von $\mathrm{P}_T(X, \mathcal{F})$.

Nehme nun umgekehrt an, dass μ kein extremaler Punkt von $\mathrm{P}_T(X, \mathcal{F})$ ist. Es gibt also μ_1, $\mu_2 \in \mathrm{P}_T(X, \mathcal{F})$ mit $\mu_1 \neq \mu_2$ und $0 < t < 1$, so dass $\mu = t\mu_1 + (1-t)\mu_2$. Dann ist $\mu_1 \ll \mu$ (μ_1 ist schwach absolut stetig bezüglich μ). Nach dem Satz von Radon-Nikodym gibt es daher ein $f \in \mathcal{L}^1(X, \mathcal{F}, \mu)$, so dass $\mu_1(B) = \int_B f\, d\mu$ für alle $B \in \mathcal{F}$. Für jedes $B \in \mathcal{F}$ gilt dann

$$\begin{aligned}\int_{B \cap T^{-1}(B)} f\, d\mu + \int_{B \setminus T^{-1}(B)} f\, d\mu &= \int_B f\, d\mu = \mu_1(B) = \mu_1(T^{-1}(B))\\ &= \int_{T^{-1}(B)} f\, d\mu = \int_{B \cap T^{-1}(B)} f\, d\mu + \int_{T^{-1}(B) \setminus B} f\, d\mu\end{aligned}$$

und daraus ergibt sich, dass

$$\int_{B \setminus T^{-1}(B)} f\, d\mu = \int_{T^{-1}(B) \setminus B} f\, d\mu \,.$$

Ferner gilt auch $\mu(B \setminus T^{-1}(B)) = \mu(T^{-1}(B) \setminus B)$, da

$$\begin{aligned}\mu(B \cap T^{-1}(B)) + \mu(B \setminus T^{-1}(B)) &= \mu(B)\\ &= \mu(T^{-1}(B)) = \mu(B \cap T^{-1}(B)) + \mu(T^{-1}(B) \setminus B) \,.\end{aligned}$$

Für jedes $\alpha \in \mathbb{R}$ sei nun $B_\alpha = \{x \in X : f(x) \leq \alpha\}$; dann gilt

$$\begin{aligned}\int_{T^{-1}(B_\alpha) \setminus B_\alpha} f\, d\mu &= \int_{B_\alpha \setminus T^{-1}(B_\alpha)} f\, d\mu\\ &\leq \alpha\mu(B_\alpha \setminus T^{-1}(B_\alpha)) = \alpha\mu(T^{-1}(B_\alpha) \setminus B_\alpha)\end{aligned}$$

und $\int_{T^{-1}(B_\alpha) \setminus B_\alpha} f\, d\mu > \alpha\mu(T^{-1}(B_\alpha) \setminus B_\alpha)$, falls $\mu(T^{-1}(B_\alpha) \setminus B_\alpha) > 0$. Folglich ist $\mu(T^{-1}(B_\alpha) \setminus B_\alpha) = \mu(B \setminus T^{-1}(B_\alpha)) = 0$, d.h., $\mu(T^{-1}(B_\alpha) \triangle B_\alpha) = 0$ für jedes $\alpha \in \mathbb{R}$. Aber es gibt $\alpha \in \mathbb{R}$ mit $0 < \mu(B_\alpha) < 1$. (Sonst wäre nach Lemma 1.3 $f = c$ μ-fast sicher für ein $c \in \mathbb{R}$ und damit $\mu = \mu_1 = \mu_2$.) Nach Satz 4.1 ist also T nicht ergodisch bezüglich μ. $\square$

Gibt es genau ein T-invariantes Wahrscheinlichkeitsmaß μ, so ist nach Satz 4.6 T ergodisch bezüglich μ. In diesem Fall heißt T *eindeutig ergodisch*.

5 Endliche Markov-Ketten

Im Folgenden sei E eine nichtleere endliche Menge. Dann heißt eine Abbildung $Q : E \times E \to \mathbb{R}^+$ *stochastische Matrix*, wenn für jedes $u \in E$

$$\sum_{v \in E} Q(u,v) = 1 \; ;$$

eine Abbildung $\alpha : E \to \mathbb{R}^+$ heißt *Wahrscheinlichkeitsvektor*, wenn

$$\sum_{v \in E} \alpha(v) = 1 \; .$$

Für Abbildungen P, $P' : E \times E \to \mathbb{R}$ und $\beta : E \to \mathbb{R}$ werden Abbildungen $PP' : E \times E \to \mathbb{R}$ und $\beta P : E \to \mathbb{R}$ (wie bei Matrizenmultiplikation) definiert durch $(PP')(u,v) = \sum_{w \in E} P(u,w)P'(w,v)$ und $(\beta P)(v) = \sum_{u \in E} \beta(u)P(u,v)$. Ferner sei $\Delta : E \times E \to \mathbb{R}$ die 'Einheitsmatrix' (mit $\Delta(u,u) = 1$ für alle $u \in E$ und $\Delta(u,v) = 0$ für alle $u \neq v$) und setze $Q^0 = \Delta$, $Q^1 = Q$ und $Q^n = QQ^{n-1}$ für alle $n > 1$.

Ein Wahrscheinlichkeitsvektor α heißt *stationär* bezüglich einer stochastischen Matrix Q, wenn $\alpha = \alpha Q$, d.h., wenn für jedes $v \in E$

$$\alpha(v) = \sum_{u \in E} \alpha(u) Q(u,v) \; .$$

Im Folgenden sei $Q : E \times E \to \mathbb{R}^+$ eine stochastische Matrix.

Satz 5.1 *Es gibt einen stationären Wahrscheinlichkeitsvektor für Q.*

Beweis Sei β ein beliebiger Wahrscheinlichkeitsvektor und für jedes $n \geq 1$ sei $\alpha_n = 1/n \sum_{k=0}^{n-1} \beta Q^k$; also ist α_n ebenfalls ein Wahrscheinlichkeitsvektor. Nun ist $\{\alpha_n\}_{n \geq 1}$ eine Folge im folgenkompakten Raum $[0,1]^E$ und damit gibt es eine Teilfolge $\{n_j\}_{j \geq 1}$ und $\alpha \in [0,1]^E$, so dass $\{\alpha_{n_j}\}_{j \geq 1}$ gegen α konvergiert. Insbesondere ist α ein Wahrscheinlichkeitsvektor und es gilt

$$\begin{aligned} \alpha Q = \lim_{j \to \infty} \Big(\frac{1}{n_j} \sum_{k=0}^{n_j - 1} \beta Q^k \Big) Q &= \lim_{j \to \infty} \Big(\frac{1}{n_j} \sum_{k=1}^{n_j} \beta Q^k \Big) \\ &= \lim_{j \to \infty} \Big(\alpha_{n_j} + \frac{1}{n_j} (\beta Q^{n_j} - \beta) \Big) = \alpha \; . \quad \square \end{aligned}$$

Bezeichne nun mit $\mathrm{St}(Q)$ die Menge der stationären Wahrscheinlichkeitsvektoren für Q. Nach Satz 5.1 ist $\mathrm{St}(Q)$ nichtleer und es ist klar, dass $\mathrm{St}(Q)$ eine konvexe Teilmenge von $\mathbb{R}^E$ ist: Es gilt $(1-t)\alpha_1 + t\alpha_2 \in \mathrm{St}(Q)$ für alle α_1, $\alpha_2 \in \mathrm{St}(Q)$, $t \in [0,1]$. Für jedes $\alpha \in \mathrm{St}(Q)$ sei $E_\alpha = \{u \in E : \alpha(u) > 0\}$ und setze

$$E_Q = \{u \in E : \alpha(u) > 0 \text{ für ein } \alpha \in \mathrm{St}(Q)\} \; ,$$

d.h., $E_Q = \bigcup_{\alpha \in \mathrm{St}(Q)} E_\alpha$.

Lemma 5.1 *Es gibt ein $\alpha \in \mathrm{St}(Q)$ mit $E_\alpha = E_Q$.*

Beweis Für jedes $u \in E_Q$ gibt es ein $\alpha_u \in \mathrm{St}(Q)$ mit $\alpha_u(u) > 0$. Wähle einen Wahrscheinlichkeitsvektor $\beta : E_Q \to \mathbb{R}^+$ mit $\beta(u) > 0$ für jedes $u \in E_Q$ und sei $\alpha = \sum_{u \in E_Q} \beta(u)\alpha_u$. Dann ist $\alpha \in \mathrm{St}(Q)$ und $E_\alpha = E_Q$. □

Die stochastische Matrix Q wird *unzerlegbar* genannt, wenn $E_Q = E$. In diesem Fall gibt es nach Lemma 5.1 ein $\alpha \in \mathrm{St}(Q)$ mit $\alpha(u) > 0$ für alle $u \in E$.

Lemma 5.2 *Sei $\alpha \in \mathrm{St}(Q)$; dann gilt $Q(u,v) = 0$ für alle $u \in E_\alpha$, $v \in E \setminus E_\alpha$. Insbesondere ist die Einschränkung $Q_\alpha : E_\alpha \times E_\alpha \to \mathbb{R}^+$ von Q auf $E_\alpha \times E_\alpha$ eine unzerlegbare stochastische Matrix.*

Beweis Sei $u \in E_\alpha$, $v \in E \setminus E_\alpha$; dann gilt

$$0 = \alpha(v) = (\alpha Q)(v) = \sum_{w \in E} \alpha(w)Q(w,v) \geq \alpha(u)Q(u,v)$$

und damit ist $Q(u,v) = 0$, da $\alpha(u) > 0$. □

Satz 5.2 (Ergodensatz für stochastische Matrizen) *Der Limes*

$$R(u,v) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} Q^k(u,v)$$

existiert für alle u, $v \in E$. Ferner ist der Grenzwert R eine stochastische Matrix mit $RQ = QR = R^2 = R$, und für jedes $\beta : E \to \mathbb{R}$ mit $\beta Q = \beta$ gilt $\beta R = \beta$.

Beweis Um die Existenz des Limes zu beweisen, braucht man lediglich Sätze aus Linearer Algebra I und II; siehe, zum Beispiel Fritz, Huppert und Willems [5]. Der Grenzwert R ist dann eine stochastische Matrix, da $1/n \sum_{k=0}^{n-1} Q^k$ dies ist für jedes n. Ferner gilt

$$RQ = \Big(\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} Q^k\Big) Q = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} Q^k = R$$

und genauso ist $QR = R$. Für jedes $k \geq 0$ ist dann auch $Q^k R = R$, also ist

$$R^2 = \Big(\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} Q^k\Big) R = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} Q^k R = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} R = R\,.$$

Ist schließlich $\beta : E \to \mathbb{R}$ mit $\beta Q = \beta$, so ist auch $\beta Q^k = \beta$ für jedes $k \geq 0$ und daraus ergibt sich, dass

$$\beta R = \beta\Big(\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} Q^k\Big) = \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} \beta Q^k = \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} \beta = \beta\ . \ \square$$

Im Folgenden sei R wie in Satz 5.2 und seien $\{\alpha_u\}_{u\in E}$ die Zeilen von R, d.h., $\alpha_u : E \to \mathbb{R}^+$ ist die Abbildung mit $\alpha_u(v) = R(u,v)$ für alle $v \in E$. Da R eine stochastische Matrix ist, ist α_u ein Wahrscheinlichkeitsvektor, und da $RQ = R$, gilt $\alpha_u = \alpha_u Q$, d.h., $\alpha_u \in \mathrm{St}(Q)$ für jedes $u \in E$.

Satz 5.3 *Für jedes $\alpha \in \mathrm{St}(Q)$ gilt $\alpha = \sum_{u\in E} \alpha(u)\alpha_u$, und insbesondere ist $\mathrm{St}(Q)$ die konvexe Hülle der Vektoren $\{\alpha_u\}_{u\in E}$.*

Beweis Da $\alpha Q = \alpha$, ist auch $\alpha R = \alpha$, und für alle $v \in E$ gilt also

$$\alpha(v) = (\alpha R)(v) = \sum_{u\in E} \alpha(u) R(u,v) = \Big(\sum_{u\in E} \alpha(u)\alpha_u\Big)(v)\ .$$

Damit ist $\alpha = \sum\limits_{u\in E} \alpha(u)\alpha_u$. $\square$

Es wird nun stets angenommen, dass Q unzerlegbar ist. (Wenn dies nicht der Fall ist, soll man dann mit Q_α arbeiten, wobei $\alpha \in \mathrm{St}(Q)$ mit $E_\alpha = E_Q$.) Die Matrix Q heißt *irreduzibel*, wenn es zu jedem $u, v \in E$ ein $n \geq 0$ mit $Q^n(u,v) > 0$ gibt.

Satz 5.4 *Die folgenden Aussagen sind äquivalent:*

(1) Die Matrix Q ist irreduzibel.

(2) Es gilt $R(u,v) > 0$ für alle $u, v \in E$.

(3) Es gilt $\alpha_u = \alpha_v$ für alle $u, v \in E$ (d.h., die Zeilen von R sind alle gleich).

(4) Der Eigenraum von Q zum Eigenwert 1 hat die Dimension 1 (d.h., 1 ist ein einfacher Eigenwert von Q).

(5) Es gibt einen eindeutigen stationären Wahrscheinlichkeitsvektor für Q.

Beweis (2) $\Rightarrow$ (1): Dies ist klar, da $\lim\limits_{n\to\infty} 1/n \sum_{k=0}^{n-1} Q^k(u,v) = R(u,v)$ für alle $u, v \in E$.

(1) $\Rightarrow$ (2): Seien $u, v \in E$; da R eine stochastische Matrix ist, gibt es $w \in E$ mit $R(u,w) > 0$ und nach Voraussetzung gibt es ein $n \geq 0$ mit $Q^n(w,v) > 0$. Aber $R(u,v) \geq R(u,w)Q^n(w,v)$, da $R = RQ^n$, und damit ist $R(u,v) > 0$.

(2) $\Rightarrow$ (3): Sei $w \in E$ fest und wähle $v \in E$, so dass $R(v,w) \geq R(u,w)$ für alle $u \in E$. Da $R^2 = R$, ist dann

$$\begin{aligned}\sum_{u\in E} R(v,u)(R(v,w) - R(u,w)) &= \sum_{u\in E} R(v,u)R(v,w) - \sum_{u\in E} R(v,u)R(u,w)\\ &= R(v,w) - R^2(v,w) = 0\end{aligned}$$

und daraus folgt, dass $R(u,w) = R(v,w)$ für alle $u \in E$, da $R(v,w) \geq R(u,w)$ und $R(v,u) > 0$ für alle $u \in E$. Damit ist $R(u,w)$ unabhängig von u für jedes $w \in E$ und dies zeigt, dass $\alpha_u = \alpha_v$ für alle u, $v \in E$.

(3) $\Rightarrow$ (4): Sei α der Wahrscheinlichkeitsvektor mit $\alpha = \alpha_u$ für alle $u \in E$ und betrachte $\beta : E \to \mathbb{R}$ mit $\beta Q = \beta$; d.h., β ist entweder 0 oder ein (linker) Eigenvektor von Q zum Eigenwert 1. Dann gilt auch $\beta R = \beta$ und damit ist

$$\beta(v) = \sum_{u\in E} \beta(u)R(u,v) = \big(\sum_{u\in E} \beta(u)\big)\alpha(v)$$

für alle $v \in E$, d.h., β ist ein Vielfaches von α. Also hat der Eigenraum von Q zum Eigenwert 1 die Dimension 1.

(4) $\Rightarrow$ (5): Seien α, $\alpha' \in \mathrm{St}(Q)$; da α und α' linke Eigenvektoren von Q zum Eigenwert 1 sind, ist α' ein Vielfaches von α und damit ist $\alpha' = \alpha$, da α und α' beide Wahrscheinlichkeitsvektoren sind.

(5) $\Rightarrow$ (2): Sei α das eindeutige Element in $\mathrm{St}(Q)$. Da Q unzerlegbar ist, folgt aus Lemma 5.1, dass $\alpha(v) > 0$ für alle $v \in E$. Aber $\alpha_u \in \mathrm{St}(Q)$ und damit ist $R(u,v) = \alpha_u(v) = \alpha(v) > 0$ für alle u, $v \in E$. $\square$

Satz 5.5 *Sei Q irreduzibel mit $\mathrm{St}(Q) = \{\alpha\}$. Dann sind äquivalent:*

(1) Es gibt ein $n \geq 1$, so dass $Q^n(u,v) > 0$ für alle u, $v \in E$.

(2) Es gilt $\lim_{n\to\infty} Q^n(u,v) = \alpha(v)$ für alle u, $v \in E$.

(3) Für jeden Wahrscheinlichkeitsvektor $\beta : E \to \mathbb{R}^+$ gilt $\lim_{n\to\infty} \beta Q^n = \alpha$.

Beweis (2) $\Rightarrow$ (1): Dies ist klar, da $\alpha(v) > 0$ für alle $v \in E$.

(1) $\Rightarrow$ (3): Sei $m \geq 1$ und nehme an, dass $\lim_{n\to\infty} Q^{mn}(u,v) = \alpha(v)$ für alle u, $v \in E$. Für jedes $k \geq 0$ gilt dann auch

$$\lim_{n\to\infty} Q^{mn+k}(u,v) = \lim_{n\to\infty} \sum_{w\in E} Q^{mn}(u,w)Q^k(w,v) = \sum_{w\in E} \alpha(w)Q^k(w,v) = \alpha(v)$$

und damit $\lim_{n\to\infty} Q^n(u,v) = \alpha(v)$ für alle u, $v \in E$. Ohne Beschränkung der Allgemeinheit kann man also annehmen, dass $Q(u,v) > 0$ für alle u, $v \in E$.

Sei $\varphi : \mathbb{R} \to \mathbb{R}$ zweimal differenzierbar mit $\varphi''(x) > 0$ für alle $x \in \mathbb{R}$ (zum Beispiel $\varphi(x) = x^2$). Für jedes $\gamma : E \to \mathbb{R}$ setze

$$F(\gamma) = \sum_{u \in E} \alpha(u)\varphi(\hat{\gamma}(u))$$

wobei $\hat{\gamma} : E \to \mathbb{R}$ durch $\hat{\gamma}(u) = \gamma(u)/\alpha(u)$ definiert ist. Sei nun $\pi : E \to \mathbb{R}^+$ ein Wahrscheinlichkeitsvektor und setze $\varrho = \pi Q$. Dann gilt

$$\begin{aligned} F(\pi) - F(\pi Q) &= \sum_{u \in E} \alpha(u)\varphi(\hat{\pi}(u)) - \sum_{v \in E} \alpha(v)\varphi(\hat{\varrho}(v)) \\ &= \sum_{u \in E} \alpha(u)\varphi(\hat{\pi}(u)) \sum_{v \in E} Q(u,v) - \sum_{v \in E}\sum_{u \in E} \alpha(u)Q(u,v)\varphi(\hat{\varrho}(v)) \\ &= \sum_{u \in E}\sum_{v \in E} \alpha(u)Q(u,v)(\varphi(\hat{\pi}(u)) - \varphi(\hat{\varrho}(v))) \,. \end{aligned}$$

Für jede Abbildung $b : E \to \mathbb{R}$ gilt ferner

$$\begin{aligned} &\sum_{u \in E}\sum_{v \in E} \alpha(u)Q(u,v)(\hat{\pi}(u) - \hat{\varrho}(v))b(v) \\ &\qquad = \sum_{v \in E}\sum_{u \in E} \Big(\pi(u)Q(u,v) - \frac{\alpha(u)}{\alpha(v)}Q(u,v)\varrho(v)\Big)b(v) \\ &\qquad = \sum_{v \in E} \Big(\varrho(v) - \frac{\alpha(v)}{\alpha(v)}Q(u,v)\varrho(v)\Big)b(v) = 0 \,, \end{aligned}$$

und insbesondere mit $b(v) = \varphi'(\hat{\varrho}(v))$ für jedes $v \in E$ ist dann

$$F(\pi) - F(\pi Q) = \sum_{u,\, v \in E} \alpha(u)Q(u,v)(\varphi(\hat{\pi}(u)) - \varphi(\hat{\varrho}(v)) - (\hat{\pi}(u) - \hat{\varrho}(v))\varphi'(\hat{\varrho}(v))) \,.$$

Nach dem Mittelwertsatz ist aber $\varphi(x) - \varphi(y) - (x-y)\varphi'(y) > 0$ für alle $x \neq y$, da $\varphi'' > 0$. Daraus folgt, dass $F(\pi) \geq F(\pi Q)$ und $F(\pi) = F(\pi Q)$ nur dann, wenn $\hat{\pi}(u) = \hat{\varrho}(v)$ für alle $u,\, v \in E$, d.h., $F(\pi) = F(\pi Q)$ gilt nur dann, wenn $\pi = \pi Q = \alpha$.

Sei $\beta : E \to \mathbb{R}^+$ ein Wahrscheinlichkeitsvektor, und setze $\beta_n = \beta Q^n$ für jedes $n \geq 0$. Da $\beta_{n+1} = \beta_n Q$, ist $F(\beta_{n+1}) \leq F(\beta_n)$, und damit ist die Folge $\{F(\beta_n)\}_{n \geq 0}$ monoton fallend. Da nun $[0,1]^E$ folgenkompakt ist, besitzt die Folge $\{\beta_n\}_{n \geq 0}$ mindestens einen Häufungspunkt. Betrachte also einen solchen Häufungspunkt γ; es gibt dann eine Teilfolge $\{n_j\}_{j \geq 1}$, so dass $\{\beta_{n_j}\}_{j \geq 1}$ gegen γ konvergiert. Insbesondere ist γ ein Wahrscheinlichkeitsvektor und es gilt

$$F(\gamma Q) = F\Big(\lim_{j \to \infty} \beta_{n_j} Q\Big) = F\Big(\lim_{j \to \infty} \beta_{n_j+1}\Big) = \lim_{j \to \infty} F(\beta_{n_j+1})$$

und auch $F(\gamma) = \lim_{j\to\infty} F(\beta_{n_j})$. Da $n_{j+1} \geq n_j + 1$, ist $F(\beta_{n_j+1}) \geq F(\beta_{n_{j+1}})$ für jedes $j \geq 1$ und daher ist $F(\gamma Q) \geq F(\gamma)$. Daraus ergibt sich, dass $\gamma = \gamma Q = \alpha$. Dies zeigt, dass α der einzige Häufungspunkt der Folge $\{\beta_n\}_{n\geq 1}$ ist, und folglich konvergiert $\{\beta_n\}_{n\geq 1}$ gegen α.

(3) $\Rightarrow$ (2): Sei $u \in E$ und definiere $\beta : E \to \mathbb{R}^+$ durch $\beta(u) = 1$ und $\beta(w) = 0$ für alle $w \neq u$. Da $Q^n(u,v) = (\beta Q^n)(v)$, gilt dann für alle $v \in E$, dass

$$\lim_{n\to\infty} Q^n(u,v) = \lim_{n\to\infty} (\beta Q^n)(v) = \alpha(v)\ . \quad \square$$

Bemerkung: Wie im Beweis für Satz 5.6 ((1) $\Rightarrow$ (3)) nehme an, dass $Q(u,v) > 0$ für alle u, $v \in E$, und sei $\beta : E \to \mathbb{R}^+$ ein Wahrscheinlichkeitsvektor; dann gilt $\beta_n(v) = (\beta Q^n)(v) > 0$ für alle $v \in E$, $n \geq 1$. Daraus sieht man leicht, dass es für diesen Beweis genügt, eine zweimal differenzierbare Abbildung $\varphi : (0,+\infty) \to \mathbb{R}$ mit $\varphi''(x) > 0$ für alle $x > 0$ zu haben. Insbesondere kann man $\varphi(x) = x \log x$ nehmen; in diesem Fall ist

$$F(\gamma) = \sum_{u\in E} \gamma(u) \log\Big(\frac{\gamma(u)}{\alpha(u)}\Big)$$

eine Art 'freier Energie'.

Sei nun I entweder $\mathbb{Z}$ oder $\mathbb{N}$. Wie in Kapitel 3 setze $E_\infty = \mathrm{Abb}(I,E)$ und sei $\mathcal{E}_\infty = \sigma\big(\bigcup_{n\in I} \mathrm{p}_m^{-1}(\mathcal{E})\big)$, wobei $\mathcal{E} = \mathcal{P}(E)$ und sei $\mathrm{p}_m : E_\infty \to E$ die durch $\mathrm{p}_m(\{x_n\}_{n\in I}) = x_m$ gegebene Projektionsabbildung ist. Ferner sei $S : E_\infty \to E_\infty$ definiert durch $S(\{x_n\}_{n\in I}) = \{x'_n\}_{n\in I}$, wobei $x'_n = x_{n+1}$.

Satz 5.6 *Sei $\alpha \in \mathrm{St}(Q)$; dann gibt es ein eindeutiges $\nu \in \mathrm{P}(E_\infty, \mathcal{E}_\infty)$, so dass*

$$\nu\Big(\bigcap_{n=k}^{\ell} \mathrm{p}_n^{-1}(B_n)\Big) = \sum_{u_k\in B_k} \cdots \sum_{u_\ell\in B_\ell} \alpha(u_k)Q(u_k,u_{k+1})\cdots Q(u_{\ell-1},u_\ell)$$

für alle k, $\ell \in I$ mit $k \leq \ell$ und alle $B_k, \ldots, B_\ell \subset E$. Ferner ist S maßtreu bezüglich ν.

Beweis Sei $J = \{k,\ldots,\ell\} \in \mathcal{J}$ und definiere $\gamma_J : E^J \to \mathbb{R}^+$ durch

$$\gamma_J(u_k,\ldots,u_\ell) = \alpha(u_k)Q(u_k,u_{k+1})\cdots Q(u_{\ell-1},u_\ell)$$

für alle $(u_k,\ldots,u_\ell) \in E^J$. Man sieht leicht, dass $\sum_{w\in E^J} \gamma_J(w) = 1$ und damit kann ein Maß $\nu_J \in \mathrm{P}(E^J,\mathcal{E}^J)$ definiert werden durch $\nu_J(B) = \sum_{w\in B} \gamma_J(w)$ für jedes $B \in \mathcal{E}^J = \mathcal{P}(E^J)$. Insbesondere ist dann

$$\nu_J\Big(\bigcap_{n=k}^{\ell} (\mathrm{p}_n^J)^{-1}(B_n)\Big) = \sum_{u_k\in B_k} \cdots \sum_{u_\ell\in B_\ell} \alpha(u_k)Q(u_k,u_{k+1})\cdots Q(u_{\ell-1},u_\ell)\ .$$

Sei nun $B_k, \ldots, B_\ell \subset E$ und $L = \{k, \ldots, \ell+1\}$. Dann gilt

$$\nu_L\Big(\bigcap_{m\in J} (\mathrm{p}_m^J)^{-1}(B_m)\Big) = \nu_L\Big(\bigcap_{m\in L} (\mathrm{p}_m^L)^{-1}(B_m)\Big)$$

mit $B_{\ell+1} = E$ und daraus ergibt sich, dass

$$\begin{aligned} &\nu_L\Big(\bigcap_{m\in J} (\mathrm{p}_m^L)^{-1}(B_m)\Big) \\ &\quad = \sum_{u_k\in B_k} \cdots \sum_{u_\ell\in B_\ell} \sum_{v\in E} \alpha(u_k)Q(u_k,u_{k+1})\cdots Q(u_{\ell-1},u_\ell)Q(u_\ell,v) \\ &\quad = \sum_{u_k\in B_k} \cdots \sum_{u_\ell\in B_\ell} \alpha(u_k)Q(u_k,u_{k+1})\cdots Q(u_{\ell-1},u_\ell) = \nu_J\Big(\bigcap_{m\in J} (\mathrm{p}_m^J)^{-1}(B_m)\Big) . \end{aligned}$$

Ist andererseits $M = \{k-1, \ldots, \ell\}$, so ist

$$\nu_M\Big(\bigcap_{m\in J} (\mathrm{p}_m^M)^{-1}(B_m)\Big) = \nu_M\Big(\bigcap_{m\in M} (\mathrm{p}_m^M)^{-1}(B_m)\Big)$$

mit $B_{k-1} = E$ und wie oben ergibt sich daraus, dass

$$\nu_M\Big(\bigcap_{m\in J} (\mathrm{p}_m^M)^{-1}(B_m)\Big) = \nu_J\Big(\bigcap_{m\in J} (\mathrm{p}_m^J)^{-1}(B_m)\Big) ,$$

da $\sum_{v\in E} \alpha(v)Q(v,u_k) = \alpha(u_k)$ für jedes $u_k \in B_k$. Diese zwei speziellen Fälle zeigen, dass die Maße $\{\nu_J\}_{J\in\mathcal{J}}$ verträglich sind: Ist $K = \{k-i, \ldots, \ell+j\}$, so folgt nach i bzw. j Anwendungen des zweiten bzw. des ersten Falls, dass

$$\nu_K\Big(\bigcap_{m\in J} (\mathrm{p}_m^K)^{-1}(B_m)\Big) = \nu_J\Big(\bigcap_{m\in J} (\mathrm{p}_m^J)^{-1}(B_m)\Big) .$$

Nach Satz 3.1 gibt es also ein eindeutiges Maß $\nu \in \mathrm{P}(E_\infty, \mathcal{E}_\infty)$, so dass

$$\nu\Big(\bigcap_{m\in J} \mathrm{p}_m^{-1}(B_m)\Big) = \nu_J\Big(\bigcap_{m\in J} (\mathrm{p}_m^J)^{-1}(B_m)\Big)$$

für alle $J \in \mathcal{J}$. Damit ist ν das eindeutige Wahrscheinlichkeitsmaß mit

$$\nu\Big(\bigcap_{n=k}^{\ell} \mathrm{p}_n^{-1}(B_n)\Big) = \sum_{u_k\in B_k} \cdots \sum_{u_\ell\in B_\ell} \alpha(u_k)Q(u_k,u_{k+1})\cdots Q(u_{\ell-1},u_\ell)$$

für alle $k, \ell \in I$ mit $k \le \ell$ und alle $B_k, \ldots, B_\ell \subset E$. Ferner ist nach Lemma 3.1 S maßtreu bezüglich ν, da

$$\begin{aligned} \nu\Big(\bigcap_{n=k}^{\ell} \mathrm{p}_n^{-1}(B_n)\Big) &= \sum_{u_k\in B_k} \cdots \sum_{u_\ell\in B_\ell} \alpha(u_k)Q(u_k,u_{k+1})\cdots Q(u_{\ell-1},u_\ell) \\ &= \nu\Big(\bigcap_{n=k}^{\ell} \mathrm{p}_{n+1}^{-1}(B_n)\Big) \end{aligned}$$

für alle $B_k, \ldots, B_\ell \subset E$ und alle $k, \ell \in I$ mit $k \leq \ell$. $\square$

Im Folgenden sei $\alpha \in \mathrm{St}(Q)$ mit $\alpha(u) > 0$ für alle $u \in E$ (und nach Lemma 5.2 gibt es ein solches α, da Q unzerlegbar ist) und sei $\nu \in \mathrm{P}(E_\infty, \mathcal{E}_\infty)$ das durch Satz 5.6 bestimmte Maß.

Für jedes $u \in E$ setze $\psi_u = \chi_{\mathrm{p}_0^{-1}(\{u\})}$ (und damit ist $\psi_u(\{z_n\}_{n\in I}) = \Delta(u, z_0)$). Man beachte, das $\int \psi_u \, d\nu = \nu(\mathrm{p}_0^{-1}(\{u\})) = \alpha(u)$.

Lemma 5.3 *Für alle $u, v \in E$ gilt*

$$R(u,v) = \frac{1}{\alpha(u)} \int \psi_u \mathrm{E}_\nu(\psi_v | \mathcal{I}_S) \, d\nu \; .$$

Beweis Nach Satz 2.2 gilt $\lim_{n\to\infty} 1/n \sum_{k=0}^{n-1} \psi_v \circ S^k = \mathrm{E}_\nu(\psi_v|\mathcal{I}_S)$ ν-fast sicher. Da nun $0 \leq 1/n \sum_{k=0}^{n-1} \psi_v \circ S^k \leq 1$ und $0 \leq \psi_u \leq 1$, folgt dann aus dem Satz von der dominierten Konvergenz, dass

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \int \psi_u(\psi_v \circ S^k) \, d\nu = \int \psi_u \mathrm{E}_\nu(\psi_v|\mathcal{I}_S) \, d\nu \; .$$

Aber $\psi_v \circ S^k = \chi_{\mathrm{p}_k^{-1}(\{v\})}$ und damit ist

$$\int \psi_u(\psi_v \circ S^k) \, d\nu = \nu(\mathrm{p}_0^{-1}(\{u\}) \cap \mathrm{p}_k^{-1}(\{v\})) = \alpha(u) Q^k(u,v) \; .$$

Für alle $u, v \in E$ gilt also

$$R(u,v) = \lim_{n\to\infty} 1/n \sum_{k=0}^{n-1} Q^k(u,v) = \frac{1}{\alpha(u)} \int \psi_u \mathrm{E}_\nu(\psi_v|\mathcal{I}_S) \, d\nu \; . \quad \square$$

Bemerkung: Der Beweis für Lemma 5.3 zeigt, dass Satz 5.2 für eine unzerlegbare stochastische Matrix aus dem Birkhoffschen Ergodensatz folgt. (Der allgemeine Fall folgt aber nicht vollständig direkt aus dem unzerlegbaren Fall: Man muss dann immer noch zeigen, dass $\lim_{n\to\infty} 1/n \sum_{k=0}^{n-1} Q^k(u,v) = 0$ für alle $u \in E \setminus E_Q$, $v \in E$.)

Satz 5.7 *Die folgenden Aussagen sind äquivalent:*

(1) Die Abbildung S ist ergodisch bezüglich ν.

(2) Die Matrix Q ist irreduzibel (und dann ist $\mathrm{St}(Q) = \{\alpha\}$*).*

Beweis (1) ⇒ (2): Da $\mathcal{I}_S$ ν-trivial ist, ist $\mathrm{E}_\nu(\psi_v|\mathcal{I}_S) = \int \psi_v \, d\nu$ ν-fast sicher, und damit ist $\alpha_u(v) = R(u,v) = (\alpha(u))^{-1} \int \psi_u \, d\nu \int \psi_v \, d\nu = (\alpha(u))^{-1}\alpha(u)\alpha(v) = \alpha(v)$ für alle u, $v \in E$. Folglich ist nach Satz 5.4 ((3) ⇒ (1)) Q irreduzibel.

(2) ⇒ (1): Seien J, $K \in \mathcal{J}$ mit $J = [a, \dots, b]$ und $K = [c, \dots, d]$, und setze $A = \bigcap_{\ell=a}^{b} \mathrm{p}_\ell^{-1}(\{u_\ell\})$ und $B = \bigcap_{m=c}^{d} \mathrm{p}_m^{-1}(\{v_m\})$, wobei $u_a, \dots, u_b, v_c, \dots, v_d \in E$. Für $k \geq k_0 = \max\{0, d-a\}$ ist dann

$$\begin{aligned}
\nu(S^k(A) \cap B) &= \nu\Big(\bigcap_{m=c}^{d} \mathrm{p}_m^{-1}(\{v_m\}) \cap \bigcap_{\ell=a}^{b} \mathrm{p}_{\ell+k}^{-1}(\{u_\ell\})\Big) \\
&= \alpha(v_c)Q(v_c, v_{c+1}) \cdots Q(v_{d-1}, v_d)Q^{k-d+a}(v_d, u_a)Q(u_a, u_{a+1}) \cdots Q(u_{b-1}, u_b) \\
&= \nu(B)\nu(A)\frac{1}{\alpha(u_a)}Q^{k-d+a}(v_d, u_a)
\end{aligned}$$

und daraus ergibt sich nach Satz 5.4 ((1) ⇒ (3)), dass

$$\begin{aligned}
\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} \nu(S^{-1}(A) \cap B) &= \nu(A)\nu(B)\frac{1}{\alpha(u_a)} \lim_{n\to\infty} \frac{1}{n}\sum_{k=k_0}^{n-1} Q^{k-d+a}(v_d, u_a) \\
&= \nu(A)\nu(B)\frac{1}{\alpha(u_a)}\alpha(u_a) = \nu(A)\nu(B) \; .
\end{aligned}$$

Für jedes $J \in \mathcal{J}$ sei $\mathcal{E}_J = \sigma\big(\bigcup_{m\in J} \mathrm{p}_m^{-1}(\mathcal{E})\big)$; also ist $\mathcal{E}_J$ endlich und jedes Element in $\mathcal{E}_J$ hat eine Darstellung als endliche disjunkte Vereinigung von Mengen der Form $\bigcap_{m\in J} \mathrm{p}_m^{-1}(\{u_m\})$. Damit gilt $\lim_{n\to\infty} 1/n \sum_{k=0}^{n-1} \nu(S^{-1}(A) \cap B) = \nu(A)\nu(B)$ für alle $A \in \mathcal{E}_J$, $B \in \mathcal{E}_K$, J, $K \in \mathcal{J}$ und folglich ist nach Lemma 4.2 S ergodisch bezüglich ν, da $\mathcal{A} = \bigcup_{J\in\mathcal{J}} \mathcal{E}_J$ eine Algebra ist mit $\sigma(\mathcal{A}) = \mathcal{E}$. □

Satz 5.8 *Die folgenden Aussagen sind äquivalent:*

(1) Die Abbildung S ist schwach-mischend bezüglich ν.

(2) Die Abbildung S ist stark-mischend bezüglich ν.

(3) Es gibt ein $n \geq 1$, so dass $Q^n(u,v) > 0$ für alle u, $v \in E$.

Beweis (2) ⇒ (1): Dies ist klar.

(1) ⇒ (3): Für jedes u, $v \in E$ gibt es nach Satz 4.3 eine Teilmenge $J(u,v)$ von $\mathbb{N}$ mit $\lim_{n\to\infty} \sigma_n(\{\delta^{J(u,v)}\}_{k\geq 0}) = 0$, so dass

$$\lim_{n\to\infty, n\notin J(u,v)} \nu(S^{-n}(\mathrm{p}_0^{-1}(\{v\})) \cap \mathrm{p}_0^{-1}(\{u\})) = \nu(\mathrm{p}_0^{-1}(\{u\}))\nu(\mathrm{p}_0^{-1}(\{v\})) = \alpha(v)\alpha(u)$$

und hier gilt ferner

$$\nu(S^{-n}(\mathrm{p}_0^{-1}(\{v\})) \cap \mathrm{p}_0^{-1}(\{u\})) = \nu(\mathrm{p}_n^{-1}(\{v\})) \cap \mathrm{p}_0^{-1}(\{u\})) = \alpha(u)Q^n(u,v) \; .$$

Setze $J = \bigcup_{u,\, v \in E} J(u,v)$; dann gilt $\lim_{n\to\infty} \sigma_n(\{\delta^J\}_{k\geq 0}) = 0$ und

$$\lim_{n\to\infty, n\notin J} Q^n(u,v) = \alpha(v)$$

für alle $u,\, v \in E$. Insbesondere gibt es $n \notin J$ mit $Q^n(u,v) \geq \frac{1}{2}\alpha(v) > 0$ für alle $u,\, v \in E$.

(3) $\Rightarrow$ (2): Dieser ist im Wesentlichen identisch mit dem Beweis für (2) $\Rightarrow$ (1) in Satz 5.7, da nach Satz 5.5 $\lim\limits_{n\to\infty} Q^n(u,v) = \alpha(v)$ für alle $u,\, v \in E$. $\square$

6 Kompakte Gruppen

Sei $\mathrm{K} = \{z \in \mathbb{C} : |z| = 1\}$ und sei $\mathcal{B}_{\mathrm{K}}$ die Borelsche σ-Algebra von K. Es gibt ein eindeutiges Maß $\lambda_{\mathrm{K}} \in \mathrm{P}(\mathrm{K}, \mathcal{B}_{\mathrm{K}})$, so dass

$$\lambda_{\mathrm{K}}(\{\exp(2\pi i\alpha) : u < \alpha \leq v\}) = v - u$$

für alle $0 \leq u < v \leq 1$. Es gilt $\lambda_{\mathrm{K}}(aA) = \lambda_{\mathrm{K}}(A)$ für alle $A \in \mathcal{B}_{\mathrm{K}}$, $a \in \mathrm{K}$, d.h., λ_{K} ist rotationsinvariant, und daraus folgt, dass $\int f(az)\, d\lambda_{\mathrm{K}}(z) = \int f(z)\, d\lambda_{\mathrm{K}}(z)$ für alle $f \in \mathcal{L}^1(\mathrm{K}, \mathcal{B}_{\mathrm{K}}, \lambda_{\mathrm{K}})$, $a \in \mathrm{K}$.

Sei $a \in \mathrm{K}$ fest und definiere $T : \mathrm{K} \to \mathrm{K}$ durch $T(x) = ax$; dann ist T maßtreu bezüglich λ_{K} (da $T^{-1}(A) = a^{-1}A$).

Satz 6.1 *Die Abbildung T ist ergodisch bezüglich λ_{K} genau dann, wenn $a^p \neq 1$ für alle $p \geq 1$.*

Beweis Übung. □

Satz 6.2 *Die Abbildung T ist nicht schwach-mischend bezüglich λ_{K}.*

Beweis Sei $\ell \geq 1$ fest und seien f, $g \in \mathrm{B}_{\mathbb{C}}(\mathrm{K}, \mathcal{B}_{\mathrm{K}})$ die Abbildungen mit $f(z) = z^\ell$ und $g(z) = z^{-\ell}$ für alle $z \in \mathrm{K}$; für alle $b \in \mathrm{K}$ ist dann

$$\int f(z)\, d\lambda_{\mathrm{K}}(z) = \int f(bz)\, d\lambda_{\mathrm{K}}(z) = b^\ell \int f(z)\, d\lambda_{\mathrm{K}}(z)$$

und damit ist $\int f\, d\lambda_{\mathrm{K}} = 0$; genauso ist $\int g\, d\lambda_{\mathrm{K}} = 0$. Nun ist

$$(f \circ T^k)(z) = f(T^k(z)) = f(a^k z) = a^{k\ell} z^\ell$$

und folglich ist $|\int (f \circ T^k) g\, d\lambda_{\mathrm{K}} - \int f\, d\lambda_{\mathrm{K}} \int g\, d\lambda_{\mathrm{K}}| = |a^{k\ell}| = 1$ für alle $k \geq 0$. Daraus ergibt sich, dass

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \Big| \int (f \circ T^k) g\, d\lambda_{\mathrm{K}} - \int f\, d\lambda_{\mathrm{K}} \int g\, d\lambda_{\mathrm{K}} \Big| = 1 ,$$

und also ist nach Lemma 4.3 T nicht schwach-mischend bezüglich λ_{K}. □

Eine *metrische Gruppe* ist ein Paar (G, d) bestehend aus einer Gruppe G und einer Metrik $d : G \times G \to \mathbb{R}^+$ auf G, so dass die Abbildungen $(x, y) \mapsto xy$ (von $G \times G$ nach G) und $x \mapsto x^{-1}$ (von G nach G) beide stetig sind. Meistens schreibt man nur G statt (G, d).

Beispiele: $\mathbb{R}^n$ mit der üblichen Metrik; K (als Untergruppe der multiplikativen Gruppe $\mathbb{C}^\times$) mit der induzierten Metrik als Teilmenge von $\mathbb{C}$.

Sind $G_1, \ldots, G_n$ metrische Gruppen, so ist $G_1 \times \cdots \times G_n$ (mit der Produktmetrik) auch eine metrische Gruppe. Insbesondere ist der *n-dimensionale Torus* K^n eine metrische Gruppe.

Im Folgenden sei G eine kompakte metrische Gruppe und sei $\mathcal{B}_G$ die Borelsche σ-Algebra von G.

Satz 6.3 *Es gibt ein eindeutiges Maß $\mathrm{m}_G \in \mathrm{P}(G, \mathcal{B}_G)$, für das gilt:*

$$\mathrm{m}_G(xB) = \mathrm{m}_G(B) \text{ für alle } B \in \mathcal{B}_G,\ x \in G\ ,$$

wobei $xB = \{xy : y \in B\}$ (und man merke, dass $xB \in \mathcal{B}_G$ für alle $B \in \mathcal{B}_G$, da $xB = f^{-1}(B)$ mit $f(y) = x^{-1}y$ und f stetig ist).

Beweis □

Das Element m_G heißt das *Haarsche Maß* der Gruppe G.

Lemma 6.1 *Für alle $f \in \mathcal{L}^1(G, \mathcal{B}_G, \mathrm{m}_G)$, $x \in G$ ist die Abbildung $y \mapsto f(xy)$ in $\mathcal{L}^1(G, \mathcal{B}_G, \mathrm{m}_G)$ und es gilt*

$$\int f(xy)\, d\mathrm{m}_G(y) = \int f(y)\, d\mathrm{m}_G(y)\ .$$

Beweis Dies folgt unmittelbar aus Satz 1.2. □

Lemma 6.2 *(1) Es gilt $\mathrm{m}_G(Bx) = \mathrm{m}_G(B)$ für alle $B \in \mathcal{B}_G$, $x \in G$.*

(2) Für jede offene nichtleere Teilmenge U von G ist $\mathrm{m}_G(U) > 0$.

Beweis (1) Sei $x \in G$ fest; dann kann ein Element $\mu \in \mathrm{P}(G, \mathcal{B}_G)$ definiert werden durch $\mu(B) = \mathrm{m}_G(Bx)$ für alle $B \in \mathcal{B}_G$. Für alle $y \in G$, $B \in \mathcal{B}_G$ ist nun

$$\mu(yB) = \mathrm{m}_G((yB)x) = \mathrm{m}_G(y(Bx)) = \mathrm{m}_G(Bx) = \mu(B)$$

und nach der Eindeutigkeit von m_G ist also $\mu = \mathrm{m}_G$, d.h., $\mathrm{m}_G(Bx) = \mathrm{m}_G(B)$ für alle $B \in \mathcal{B}_G$.

(2) Sei $U \subset G$ offen mit $U \neq \varnothing$; dann ist $\{xU\}_{x \in G}$ eine offene Überdeckung von G und also gibt es $x_1, \ldots, x_n \in G$ mit $G = \bigcup_{k=1}^n x_k U$. Daraus folgt, dass

$$1 = \mathrm{m}_G(G) \leq \sum_{k=1}^n \mathrm{m}_G(x_k U) = \sum_{k=1}^n \mathrm{m}_G(U) = n\mathrm{m}_G(U)$$

und damit ist $\mathrm{m}_G(U) \geq 1/n > 0$. □

Ist $G = G_1 \times \cdots \times G_n$ das Produkt der metrischen Gruppen $G_1, \ldots, G_n$, so ist $\mathcal{B}_G = \mathcal{B}_{G_1} \times \cdots \times \mathcal{B}_{G_n}$ und aus der Eindeutigkeit von m_G folgt unmittelbar, dass $\mathrm{m}_G = \mathrm{m}_{G_1} \times \cdots \times \mathrm{m}_{G_n}$. Insbesondere ist $\mathrm{m}_{\mathrm{K}^n} = \lambda_{\mathrm{K}}^n$, da $\mathrm{m}_{\mathrm{K}} = \lambda_{\mathrm{K}}$.

Satz 6.4 *Die Abbildung $T : G \to G$ ist maßtreu bezüglich m_G in jedem der folgenden Fälle:*

(1) $T(x) = ax$ für alle $x \in G$ für ein $a \in G$.

(2) $T(x) = f(x)$ für alle $x \in G$, wobei f ein stetiger surjektiver Endomorphismus von G ist.

(2) $T(x) = af(x)$ für alle $x \in G$, wobei $a \in G$ und $f : G \to G$ ein stetiger surjektiver Endomorphismus ist.

Beweis In allen drei Fällen ist T stetig und damit ist $T^{-1}(\mathcal{B}_G) \subset \mathcal{B}_G$.

(1) Hier ist $T^{-1}(B) = a^{-1}B$ und also ist $\mathrm{m}_G(T^{-1}(B)) = \mathrm{m}_G(a^{-1}B) = \mathrm{m}_G(B)$ für alle $B \in \mathcal{B}_G$.

(2) Sei $\mu \in \mathrm{P}(G, \mathcal{B}_G)$ das Bildmaß von m_G unter f, also ist $\mu(B) = \mathrm{m}_G(f^{-1}(B))$ für alle $B \in \mathcal{B}_G$. Seien nun $x \in G$, $B \in \mathcal{B}_G$; da f surjektiv ist, gibt es ein $y \in G$ mit $f(y) = x$, daher gilt

$$f^{-1}(xB) = \{z \in G : f(z) \in f(y)B\} = y\{z \in G : f(z) \in B\} = yf^{-1}(B)$$

und folglich ist $\mu(xB) = \mathrm{m}_G(f^{-1}(xB)) = \mathrm{m}_G(yf^{-1}(B)) = \mathrm{m}_G(f^{-1}(B)) = \mu(B)$. Daraus ergibt sich nach der Eindeutigkeit von m_G, dass $\mu = \mathrm{m}_G$ und damit ist $\mathrm{m}_G(T^{-1}(B)) = \mathrm{m}_G(f^{-1}(B)) = \mathrm{m}_G(B)$ für alle $B \in \mathcal{B}_G$.

(3) Dies folgt unmittelbar aus (1) und (2), da die Zusammensetzung von zwei maßtreuen Abbildungen wieder maßtreu ist. □

Lemma 6.3 *Sei H eine Untergruppe von* K. *Dann gilt entweder $\bar{H} = \mathrm{K}$ oder $H = W_p$ für ein $p \geq 1$, wobei $W_p = \{z \in \mathrm{K} : z^p = 1\}$.*

Beweis Nehme zunächst an, dass H endlich ist. Für jedes $z \in H$ gibt es dann $n \geq 1$, so dass $z^n = 1$ und damit ist H eine Untergruppe von W_q für ein $q \geq 1$. Folglich ist $H = W_p$ mit p einem Teiler von q. Nehme nun an, dass H unendlich ist und sei $\varepsilon > 0$. Dann gibt es $x_1, x_2 \in H$ mit $x_1 \neq x_2$ und $|x_1 - x_2| < \varepsilon$. Setze $x = x_2^{-1}x_1$; also ist $x \in H$ mit $x \neq 1$ und $|x - 1| = |x_2^{-1}||x_1 - x_2| = |x_1 - x_2| < \varepsilon$. Für jedes $y \in \mathrm{K}$ gibt es dann ein $m \geq 1$, so dass $|y - x^m| < \varepsilon$, und dies zeigt, dass $\bar{H} = \mathrm{K}$. □

Für jedes $m \in \mathbb{Z}$ definiere $\theta_m : \mathrm{K} \to \mathrm{K}$ durch $\theta_m(z) = z^m$. Es ist klar, dass θ_m ein stetiger Endomorphismus von K ist. Man merke, dass $z^{-1} = \bar{z}$ für jedes $z \in \mathrm{K}$ und damit ist $\theta_{-m} = \bar{\theta}_m$ für jedes $m \in \mathbb{Z}$.

Lemma 6.4 *Sei $\theta : \mathrm{K} \to \mathrm{K}$ ein stetiger Endomorphismus von* K. *Dann gibt es ein eindeutiges $p \in \mathbb{Z}$, so dass $\theta = \theta_p$.*

Beweis Da K zusammenhängend und kompakt ist und die Abbildung θ stetig ist, ist $\theta(\mathrm{K})$ eine abgeschlossene zusammenhängende Untergruppe von K. Folglich ist nach Lemma 6.3 $\theta(\mathrm{K})$ entweder K oder $W_1 = \{1\}$, da W_p nicht zusammenhängend ist für jedes $p \geq 2$. Da $\theta = \theta_0$, falls $\theta(\mathrm{K}) = \{1\}$, kann angenommen werden, dass $\theta(\mathrm{K}) = \mathrm{K}$. Aber $\mathrm{Kern}\,\theta$ ist auch eine abgeschlossene Untergruppe von K und $\mathrm{Kern}\,\theta \neq \mathrm{K}$, da $\theta(\mathrm{K}) = \mathrm{K}$. Nach Lemma 6.3 gibt es also ein $p \geq 1$, so dass $\mathrm{Kern}\,\theta = W_p$. Es folgt nun daraus (Übungsaufgabe), dass θ entweder θ_p oder θ_{-p} ist. Die Eindeutigkeit ist klar, da $\theta_q \neq \theta_p$, falls $q \neq p$. □

Für $u \in \mathbb{Z}^n$ mit $u = (m_1, \ldots, m_n)$ definiere nun $\theta_u : \mathrm{K}^n \to \mathrm{K}$ durch

$$\theta_u(z_1, \ldots, z_n) = \prod_{k=1}^{n} z_k^{m_k}$$

für alle $(z_1, \ldots, z_n) \in \mathrm{K}^n$; also ist θ_u ein stetiger Homomorphismus.

Lemma 6.5 *Sei $\theta : \mathrm{K}^n \to \mathrm{K}$ ein stetiger Homomorphismus. Dann gibt es ein eindeutiges $u \in \mathbb{Z}^n$, so dass $\theta = \theta_u$.*

Beweis Für $k = 1, \ldots, n$ gibt es den stetigen Homomorphismus $i_k : \mathrm{K} \to \mathrm{K}^n$, der gegeben ist durch $i_k(z) = (1, \ldots, 1, z, 1, \ldots, 1)$ mit z in der k-ten Stelle. Also ist $\theta \circ i_k : \mathrm{K} \to \mathrm{K}$ ein stetiger Endomorphismus und folglich gibt es nach Lemma 6.4 ein eindeutiges $m_k \in \mathbb{Z}$, so dass $\theta \circ i_k = \theta_{m_k}$. Damit ist $\theta = \theta_u$ mit $u = (m_1, \ldots, m_n)$ und u ist eindeutig durch θ bestimmt. □

Bezeichne mit $\mathrm{M}(n \times n, \mathbb{Z})$ die Menge aller $n \times n$ Matrizen über $\mathbb{Z}$ und für jedes $A = (a_{ij}) \in \mathrm{M}(n \times n, \mathbb{Z})$ sei $\theta_A : \mathrm{K}^n \to \mathrm{K}^n$ definiert durch

$$\theta_A(z_1, \ldots, z_n) = \Big(\prod_{k=1}^{n} z_k^{a_{1k}}, \ldots, \prod_{k=1}^{n} z_k^{a_{nk}}\Big)$$

für alle $(z_1, \ldots, z_n) \in \mathrm{K}^n$; also ist θ_A ein stetiger Endomorphismus von K und

$$\theta_A(z) = (\theta_{u_1}(z), \ldots, \theta_{u_n}(z))$$

für alle $z \in \mathrm{K}^n$ mit u_k die k-te Zeile von A. Man sieht leicht, dass

$$\theta_A \circ \theta_B = \theta_{AB}$$

für alle A, $B \in \mathrm{M}(n \times n, \mathbb{Z})$. Definiere ferner $\varrho_n : \mathbb{R}^n \to \mathrm{K}^n$ durch

$$\varrho_n(x_1, \ldots, x_n) = (\exp(2\pi i x_1), \ldots, \exp(2\pi i x_n))$$

für alle $(x_1, \ldots, x_n) \in \mathbb{R}^n$; also ist ϱ_n ein stetiger surjektiver Homomorphismus. Für $A \in \mathrm{M}(n \times n, \mathbb{Z})$ sei $\varphi_A : \mathbb{R}^n \to \mathbb{R}^n$ die entsprechende lineare Abbildung (mit A die Matrix von φ_A bezüglich der kanonischen Basis von $\mathbb{R}^n$); dann ist φ_A auch ein (Gruppen-)Endomorphismus von $\mathbb{R}^n$ und es gilt $\varrho_n \circ \varphi_A = \theta_A \circ \varrho_n$, da

$$\begin{aligned}(\theta_A \circ \varrho_n)(x_1, \ldots, x_n) &= \theta_A(\exp(2\pi i x_1), \ldots, \exp(2\pi i x_n)) \\ &= \Big(\prod_{k=1}^n (\exp(2\pi i x_k))^{a_{1k}}, \ldots, \prod_{k=1}^n (\exp(2\pi i x_k))^{a_{nk}}\Big) \\ &= \Big(\exp\big(2\pi i \sum_{k=1}^n a_{1k} x_k\big), \ldots, \exp\big(2\pi i \sum_{k=1}^n x_k a_{nk}\big)\Big) = (\varrho_n \circ \varphi_A)(x_1, \ldots, x_n)\ .\end{aligned}$$

Satz 6.5 *Sei $\theta : \mathrm{K}^n \to \mathrm{K}^n$ ein stetiger Endmorphismus von K^n. Dann gibt es ein eindeutiges $A \in \mathrm{M}(n \times n, \mathbb{Z})$, so dass $\theta = \theta_A$. Ferner ist θ_A surjektiv genau dann, wenn $\det(A) \neq 0$, und θ_A ist ein Automorphismus genau dann, wenn $\det(A)$ entweder -1 oder 1 ist.*

Beweis Die erste Aussage folgt unmittelbar aus Lemma 6.5, da für jedes k die Projektionsabbildung $\mathrm{p}_k : \mathrm{K}^n \to \mathrm{K}$ ein stetiger Homomorphismus ist.

Ist $\det(A) \neq 0$, so ist φ_A eine Bijektion und damit surjektiv. Folglich ist θ_A surjektiv, da $\theta_A \circ \varrho_n = \varrho_n \circ \varphi_A$ und ϱ_n surjektiv ist. Ist dagegen $\det(A) = 0$, so ist $\dim \varphi_A(\mathbb{R}^n) < n$ und insbesondere ist $\lambda^n(\varphi_A(\mathbb{R}^n)) = 0$ (mit λ^n das Lebesguesche Maß auf $\mathbb{R}^n$). Also ist auch $\lambda^n(C) = 0$, wobei $C = \bigcup_{u \in \mathbb{Z}^n}(\varphi_A(\mathbb{R}^n) + u)$ und insbesondere ist $\mathbb{R}^n \setminus C \neq \varnothing$. Aber für jedes $z \in (\varrho_n \circ \varphi_A)(\mathbb{R}^n)$ ist $\varrho^{-1}(\{u\}) \subset C$ und damit ist $\varrho_n(u) \notin (\varrho_n \circ \varphi_A)(\mathbb{R}^n)$ für jedes $u \in \mathbb{R}^n \setminus C$. Daraus ergibt sich, dass $\theta_A \circ \varrho_n = \varrho_n \circ \varphi_A$ nicht surjektiv ist. Da ϱ_n surjektiv ist, ist dann θ_A auch nicht surjektiv.

Nehme nun an, dass θ_A ein Automorphismus ist; dann ist $(\theta_A)^{-1}$ stetig. (Sind X und Y kompakte metrische Räume und ist $f : X \to X$ eine stetige Bijektion, so ist die Umkehrabbildung $f^{-1} : Y \to X$ automatisch stetig.) Folglich gibt es nach dem ersten Teil ein $B \in \mathrm{M}(n \times n, \mathbb{Z})$, so dass $(\theta_A)^{-1} = \theta_B$; also ist

$$\theta_{E_n} = \mathrm{id}_{\mathrm{K}^n} = \theta_A \circ (\theta_A)^{-1} = \theta_A \circ \theta_B = \theta_{AB}$$

und damit ist $AB = E_n$. Da aber $\det(A)\det(B) = \det(AB) = \det E_n = 1$ und $\det(A), \det(B) \in \mathbb{Z}$, ist dann $\det(A)$ entweder -1 oder 1. Sei umgekehrt $A \in \mathrm{M}(n \times n, \mathbb{Z})$ mit $\det(A)$ entweder -1 oder 1. Nach der expliziten Formel für die Inverse einer Matrix ist $A^{-1} \in \mathrm{M}(n \times n, \mathbb{R})$ tatsächlich ein Element von $\mathrm{M}(n \times n, \mathbb{Z})$. Folglich ist θ_A ein Automorphismus von K^n mit $(\theta_A)^{-1} = \theta_{A^{-1}}$. $\square$

Sei $A \in \mathrm{M}(n \times n, \mathbb{Z})$ mit $\det(A) \neq 0$. Nach Satz 6.4 (2) und Satz 6.5 ist dann θ_A maßtreu bezüglich $\mathrm{m}_{\mathrm{K}^n}$.

Lemma 6.6 *(1) Sei $a \in G$ und definiere $T : G \to G$ durch $T(x) = ax$ für alle $x \in G$. Ist T ergodisch bezüglich m_G, so liegt die Folge $\{a^n\}_{n\geq 0}$ dicht in G.*

(2) Liegt die Folge $\{a^n\}_{n\geq 0}$ dicht in G für ein $a \in G$, so ist G eine abelsche Gruppe.

Beweis (1) Nach Lemma 6.2 (2) ist $\mathrm{m}_G(U) > 0$ für jede nichtleere offene Teilmenge von G und folglich gilt nach Satz 4.2, dass

$$\mathrm{m}_G(\{x \in G : \mathrm{O}_T(x) \text{ ist eine dichte Teilmenge von } G\}) = 1\,,$$

wobei hier $\mathrm{O}_T(x) = \{y \in G : y = a^n x \text{ für ein } n \geq 0\}$. Insbesondere gibt es ein $x \in G$, so dass die Folge $\{a^n x\}_{n\geq 0}$ dicht in G liegt. Aber dann liegt auch die Folge $\{a^n\}_{n\geq 0}$ dicht in G, da die Abbildung $y \mapsto yx^{-1}$ stetig ist.

(2) Seien $x,\, y \in G$; da $\{a^n\}_{n\geq 0}$ dicht in G liegt, gibt es Teilfolgen $\{m_k\}_{k\geq 0}$ und $\{n_k\}_{k\geq 0}$, so dass $x = \lim_{k\to\infty} a^{m_k}$ und $y = \lim_{k\to\infty} a^{n_k}$ und folglich ist dann

$$xy = \lim_{k\to\infty} a^{m_k} a^{n_k} = \lim_{k\to\infty} a^{m_k+n_k} = \lim_{k\to\infty} a^{n_k} a^{m_k} = yx\,. \quad \square$$

Nehme nun zusätzlich an, dass G abelsch ist, also ist G eine kompakte metrische abelsche Gruppe. (Die Gruppenoperation wird aber immer noch multiplikativ geschrieben.)

Ein stetiger Homomorphismus $\gamma : G \to \mathrm{K}$ heißt *Charakter* von G. Sind $\gamma_1,\, \gamma_2$ Charaktere von G, so ist $\gamma_1\gamma_2$ auch ein Charakter (mit $(\gamma_1\gamma_2)(x) = \gamma_1(x)\gamma_2(x)$ für alle $x \in G$), die Abbildung $1 : G \to \mathrm{K}$ (mit $1(x) = 1$ für alle $x \in G$) ist ein Charakter und für jeden Charakter γ ist $\bar{\gamma}$ ein Charakter mit $\gamma\bar{\gamma} = 1$ (wobei $\bar{\gamma}(x) = \overline{\gamma(x)}$ für alle $x \in G$). Die Charaktere von G bilden also eine abelsche Gruppe $\widehat{G}$ unter punktweiser Multiplikation; $\widehat{G}$ heißt die *Charaktergruppe* von G.

Nach Lemma 6.5 ist $\widehat{\mathrm{K}^n} = \{\theta_u : u \in \mathbb{Z}^n\}$ und da $\theta_u\theta_v = \theta_{u+v}$ für alle $u,\, v \in \mathbb{Z}^n$, sind die Gruppen $\widehat{\mathrm{K}^n}$ und $\mathbb{Z}^n$ isomorph.

Satz 6.6 *Zu jedem $x \in G \setminus \{1\}$ gibt es ein $\gamma \in \widehat{G}$ mit $\gamma(x) \neq 1$.*

Beweis $\square$

Für $G = \mathrm{K}^n$ kann man direkt sehen, dass Satz 6.6 richtig ist.

Für metrische Räume X und Y bezeichnet $\mathrm{C}(X, Y)$ die Menge aller stetigen Abbildungen von X nach Y. Sei $A(G)$ die Teilmenge von $\mathrm{C}(G, \mathbb{C})$ bestehend aus allen Abbildungen der Form $\sum_{k=1}^n c_k\gamma_k$ mit $\gamma_1, \ldots, \gamma_n \in \widehat{G}$ und $c_1, \ldots, c_n \in \mathbb{C}$.

Satz 6.7 *Zu jedem $f \in \mathrm{C}(G, \mathbb{C})$ und jedem $\varepsilon > 0$ gibt es ein $g \in A(G)$, so dass $|f(x) - g(x)| < \varepsilon$ für alle $x \in G$.*

Beweis $A(G)$ ist ein Untervektorraum des komplexen Vektorraums $\mathrm{C}(G,\mathbb{C})$ und $fg \in A(G)$ für alle f, $g \in A(G)$; ferner ist $1 \in A(G)$, $\bar{f} \in A(G)$ für alle $f \in A(G)$ und für jedes x, $y \in G$ mit $x \neq y$ gibt es ein $f \in A(G)$ mit $f(x) \neq f(y)$. (Da $x^{-1}y \neq 1$, gibt es nach Satz 6.6 ein $\gamma \in \widehat{G} \subset A(G)$ mit $\gamma(x^{-1}y) \neq 1$ und daher mit $\gamma(y) = \gamma(x^{-1}y)\gamma(x) \neq \gamma(x)$.) Nach dem Satz von Stone-Weierstraß gibt es also zu jedem $f \in \mathrm{C}(G,\mathbb{C})$ und jedem $\varepsilon > 0$ ein $g \in A(G)$, so dass $|f(x) - g(x)| < \varepsilon$ für alle $x \in G$. $\square$

Lemma 6.7 *Für alle* γ_1, $\gamma_2 \in \widehat{G}$ *gilt*

$$\int \gamma_1\bar{\gamma}_2 \, \mathrm{dm}_G = \begin{cases} 1, & \textit{falls } \gamma_1 = \gamma_2 \; , \\ 0, & \textit{falls } \gamma_1 \neq \gamma_2 \; , \end{cases}$$

und insbesondere ist $\int \gamma \, \mathrm{dm}_G = 0$ *für alle* $\gamma \in \widehat{G} \setminus \{1\}$.

Beweis Sei $\gamma \in \widehat{G}$ mit $\gamma \neq 1$. Für jedes $x \in G$ gilt dann nach Lemma 6.1

$$\int \gamma \, \mathrm{dm}_G = \int \gamma(xy) \, \mathrm{dm}_G(y) = \gamma(x) \int \gamma(y) \, \mathrm{dm}_G(y) = \gamma(x) \int \gamma \, \mathrm{dm}_G$$

und da $\gamma(x) \neq 1$ für mindestens ein $x \in G$, ist $\int \gamma \, \mathrm{dm}_G = 0$. Andererseits ist $\int 1 \, d\mathrm{m}_G = 1$. Seien nun γ_1, $\gamma_2 \in \widehat{G}$; dann ist $\gamma_1\bar{\gamma}_2 \in \widehat{G}$ und $\gamma_1 = \gamma_2$ genau dann, wenn $\gamma_1\bar{\gamma}_2 = 1$. Daraus ergibt sich, dass

$$\int \gamma_1\bar{\gamma}_2 \, \mathrm{dm}_G = \begin{cases} 1, & \text{falls } \gamma_1 = \gamma_2 \; , \\ 0, & \text{falls } \gamma_1 \neq \gamma_2 \; . \end{cases} \qquad \square$$

Lemma 6.8 *Sei* $f \in \mathrm{B}(G,\mathcal{B}_G)$.

(1) Gilt $\int f\gamma \, \mathrm{dm}_G = 0$ *für alle* $\gamma \in \widehat{G}$, *so ist* $f = 0$ m_G*-fast sicher.*

(2) Gilt $\int f\gamma \, \mathrm{dm}_G = 0$ *für alle* $\gamma \in \widehat{G} \setminus \{1\}$, *so ist* $f = \int f \, \mathrm{dm}_G$ m_G*-fast sicher.*

Beweis (1) Da $\int f\gamma \, \mathrm{dm}_G = 0$ für alle $\gamma \in \widehat{G}$, gilt dann auch $\int fg \, \mathrm{dm}_G = 0$ für alle $g \in A(G)$. Daraus ergibt sich nach Satz 6.7, dass $\int fg \, \mathrm{dm}_G = 0$ für alle $g \in \mathrm{C}(G,\mathbb{C})$ und insbesondere gilt dann $\int fg \, \mathrm{dm}_G = 0$ für alle $g \in \mathrm{C}(G,\mathbb{R})$. Nach Satz 1.5 gibt es aber zu jedem $\varepsilon > 0$ ein $g \in \mathrm{C}(G,\mathbb{R})$, so dass $\int |f - g| \, \mathrm{dm}_G < \varepsilon$ und folglich ist

$$\int f^2 \, \mathrm{dm}_G = \int f(f-g) \, \mathrm{dm}_G + \int fg \, \mathrm{dm}_G = \int f(f-g) \, \mathrm{dm}_G \leq \varepsilon M \; ,$$

wobei M eine obere Schranke von $|f|$ ist. Damit ist $\int f^2 \, \mathrm{dm}_G = 0$, da $\varepsilon > 0$ beliebig ist, d.h., $f = 0$ m_G-fast sicher.

(2) Setze $g = f - \int f \, \mathrm{dm}_G$. Für alle $\gamma \in \widehat{G} \setminus \{1\}$ gilt nach Lemma 6.7, dass

$$\int g\gamma \, \mathrm{dm}_G = \int f\gamma \, \mathrm{dm}_G - \int f \, \mathrm{dm}_G \int \gamma \, \mathrm{dm}_G = 0 - 0 = 0$$

und $\int g1 \, \mathrm{dm}_G = \int f \, \mathrm{dm}_G - \int f \, \mathrm{dm}_G = 0$, d.h., $\int g\gamma \, \mathrm{dm}_G = 0$ für alle $\gamma \in \widehat{G}$. Nach (1) ist also $g = 0$ m_G-fast sicher, d.h., $f = \int f \, \mathrm{dm}_G$ m_G-fast sicher. $\square$

Satz 6.8 *Sei $a \in G$ und definiere $T : G \to G$ durch $T(x) = ax$ für alle $x \in G$. (Nach Satz 6.4 (1) ist also T maßtreu bezüglich m_G).*

(1) Die Abbildung T ist ergodisch bezüglich m_G genau dann, wenn die Folge $\{a^n\}_{n\geq 0}$ dicht in G liegt.

(2) Ist $G \neq \{1\}$, so ist T nicht schwach-mischend bezüglich m_G.

Beweis (1) Nehme an, dass $\{a^n\}_{n\geq 0}$ dicht in G liegt und sei $f \in \mathrm{B}(G, \mathcal{B}_G)$ mit $f \circ T = f$. Für $\gamma \in \widehat{G}$ setze $c(\gamma) = \int f\gamma \, \mathrm{dm}_G$; dann gilt

$$\begin{aligned} c(\gamma) &= \int f(x)\gamma(x) \, \mathrm{dm}_G(x) = \int f(ax)\gamma(ax) \, \mathrm{dm}_G(x) \\ &= \int (f \circ T)(x)\gamma(a)\gamma(x) \, \mathrm{dm}_G(x) = \gamma(a) \int f(x)\gamma(x) \, \mathrm{dm}_G(x) = \gamma(a)c(\gamma) \, , \end{aligned}$$

und damit ist $c(\gamma) = 0$ oder $\gamma(a) = 1$. Ist aber $\gamma(a) = 1$, so ist $\gamma(a^n) = (\gamma(a))^n = 1$ für jedes $n \geq 0$ und dann ist $\gamma = 1$, da γ stetig ist und $\{a^n\}_{n\geq 0}$ dicht in G liegt. Dies zeigt, dass $\int f\gamma \, \mathrm{dm}_G = 0$ für jedes $\gamma \in \widehat{G} \setminus \{1\}$, und daraus ergibt sich nach Lemma 6.8 (2), dass $f = \int f \, \mathrm{dm}_G$ m_G-fast sicher. Nach Satz 4.1 ist also T ergodisch bezüglich m_G. Die Umkehrung ist Lemma 6.6 (1).

(2) Da $G \neq \{1\}$ gibt es nach Satz 6.6 ein $\gamma \in \widehat{G} \setminus \{1\}$. Dann gilt

$$(\gamma \circ T)(x) = \gamma(a^k x) = \gamma(a^k)\gamma(x)$$

für jedes $x \in G$ und nach Lemma 6.7 ist $\int \gamma \, \mathrm{dm}_G = \int \bar{\gamma} \, \mathrm{dm}_G = 0$. Folglich ist

$$\begin{aligned} \Big| \int (\gamma \circ T^k)\bar{\gamma} \, \mathrm{dm}_G - \int \gamma \, \mathrm{dm}_G \int \bar{\gamma} \, \mathrm{dm}_G \Big| &= \Big| \int (\gamma \circ T^k)\bar{\gamma} \, \mathrm{dm}_G \Big| \\ &= \Big| \gamma(a^k) \int \gamma\bar{\gamma} \, \mathrm{dm}_G \Big| = |\gamma(a^k)| = 1 \end{aligned}$$

für jedes $k \geq 0$ und daraus ergibt sich, dass

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \Big| \int (\gamma \circ T^k)\bar{\gamma} \, \mathrm{dm}_G - \int \gamma \, \mathrm{dm}_G \int \bar{\gamma} \, \mathrm{dm}_G \Big| = 1 \, .$$

Nach Lemma 4.3 (2) ist also T nicht schwach-mischend bezüglich m_G. $\square$

Sei $\theta : G \to G$ ein stetiger Endomorphismus. Für jeden Charakter γ ist dann $\gamma \circ \theta$ auch ein Charakter und folglich ist $\gamma \circ \theta^n \in \widehat{G}$ für alle $\gamma \in \widehat{G}$, $n \geq 0$.

Satz 6.9 *Sei $\theta : G \to G$ ein stetiger surjektiver Endomorphismus (und nach Satz 6.4 (2) ist also θ maßtreu bezüglich m_G). Dann sind äquivalent:*

(1) Für alle $\gamma \in \widehat{G} \setminus \{1\}$ und alle $n \geq 1$ ist $\gamma \circ \theta^n \neq \gamma$.

(2) Die Abbildung θ ist ergodisch bezüglich m_G.

(3) Die Abbildung θ ist stark-mischend bezüglich m_G.

Beweis (2) $\Rightarrow$ (1): Nehme an, dass es $\gamma \in \widehat{G} \setminus \{1\}$ gibt, so dass $\gamma \circ \theta^n = \gamma$ für ein $n \geq 1$. Sei $m = \min\{n \geq 1 : \gamma \circ \theta^n = \gamma\}$ und setze $f = \gamma + \gamma \circ \theta + \cdots + \gamma \circ \theta^{m-1}$. Dann ist $f \circ \theta = \gamma \circ \theta + \cdots + \gamma \circ \theta^m = f$ und da θ ergodisch ist, folgt daraus, dass $f = c\,\mathrm{m}_G$-fast sicher für ein $c \in \mathbb{C}$. (Da $f \circ \theta = f$, gilt $(\operatorname{Re} f) \circ \theta = \operatorname{Re}(f \circ \theta) = \operatorname{Re} f$ und $(\operatorname{Im} f) \circ \theta = \operatorname{Im}(f \circ \theta) = \operatorname{Im} f$; nach Satz 4.1 gibt es also a, $b \in \mathbb{R}$, so dass $\operatorname{Re} f = a$ m_G-fast sicher und $\operatorname{Im} f = b$ m_G-fast sicher. Folglich ist $f = a + ib$ m_G-fast sicher.) Da aber $\gamma \neq \gamma \circ \theta^k$ für $k = 1, \ldots, m-1$, ist nach Lemma 6.7

$$c \int \bar{\gamma}\, d\mathrm{m}_G = \int f\bar{\gamma}\, d\mathrm{m}_G = \int (\gamma + \cdots + \gamma \circ \theta^{m-1})\bar{\gamma}\, d\mathrm{m}_G = \int \gamma\bar{\gamma}\, d\mathrm{m}_G = 1$$

und dies ist ein Widerspruch, da $\int \bar{\gamma}\, d\mathrm{m}_G = 0$.

(3) $\Rightarrow$ (2): Dies ist klar.

(1) $\Rightarrow$ (3): Seien γ, $\gamma' \in \widehat{G}$ mit $\gamma \neq 1$; dann gibt es höchstens ein $n \geq 0$ mit $\gamma \circ \theta^n = \bar{\gamma}'$. (Gilt $\gamma \circ \theta^n = \bar{\gamma}' = \gamma \circ \theta^m$ mit $m > n$, so ist $(\gamma \circ \theta^n) \circ \theta^{m-n} = \gamma \circ \theta^n$ und damit $\gamma \circ \theta^n = 1$. Aber $\gamma \circ \theta^n \neq 1$, da $\gamma \neq 1$ und θ^n surjektiv ist.) Nach Lemma 6.7 gilt also $\lim_{n\to\infty} \int (\gamma \circ \theta^n)\gamma'\, d\mathrm{m}_G = 0 = \int \gamma\, d\mathrm{m}_G \int \gamma'\, d\mathrm{m}_G$ und folglich ist

$$\lim_{n\to\infty} \int (\gamma \circ \theta^n)\gamma'\, d\mathrm{m}_G = \int \gamma\, d\mathrm{m}_G \int \gamma'\, d\mathrm{m}_G$$

für alle γ, $\gamma' \in \widehat{G}$, da dies trivial richtig ist, wenn $\gamma = 1$. Daraus ergibt sich, dass

$$\lim_{n\to\infty} \int (f \circ \theta^n) g\, d\mathrm{m}_G = \int f\, d\mathrm{m}_G \int g\, d\mathrm{m}_G$$

für alle f, $g \in A(G)$, daher nach Satz 6.7 auch für alle f, $g \in \mathrm{C}(G, \mathbb{C})$ und dann (wie im Beweis für Lemma 6.8 (1)) für alle f, $g \in \mathrm{B}(G, \mathcal{B}_G)$. Nach Lemma 4.3 (3) ist also θ stark-mischend bezüglich m_G. $\square$

Der Fall mit $G = \mathrm{K}^n$ wird nun untersucht. Sei $a = (a_1, \ldots, a_n) \in \mathrm{K}^n$ und wähle $b_1, \ldots, b_n \in \mathbb{R}$, so dass $a_k = \exp(2\pi i b_k)$ für jedes $k = 1, \ldots, n$. Dann liegt $\{a^m\}_{m\geq 0}$ dicht in K^n genau, wenn $1, b_1, \ldots, b_n$ linear unabhängig über $\mathbb{Z}$ sind (d.h., wenn aus $m_0 + m_1 b_1 + \cdots + m_n b_n = 0$ mit $m_0, \ldots, m_n \in \mathbb{Z}$ stets $m_k = 0$ für jedes k folgt). Dies ist ein Satz von Kronecker [7], siehe auch Jacobs [6].

Für die Drehung $T : \mathrm{K}^n \to \mathrm{K}^n$, die durch $T(x) = ax$ mit

$$a = (\exp(2\pi i b_1), \ldots, \exp(2\pi i b_n))$$

gegeben ist, ist also T genau dann ergodisch bezüglich $\mathrm{m}_{\mathrm{K}^n}$, wenn 1, b_1, $\ldots$, b_n linear unabhängig über $\mathbb{Z}$ sind; ferner ist T nie schwach-mischend bezüglich $\mathrm{m}_{\mathrm{K}^n}$.

Sei nun $A \in \mathrm{M}(n \times n, \mathbb{Z})$ mit $\det(A) \neq 0$. Ist $u \in \mathbb{Z}^n$, so sieht man leicht, dass $\theta_u \circ \theta^m = \theta_u$ genau dann gilt, wenn $A^m u = u$, und da $A^m \in \mathrm{M}(n \times n, \mathbb{Z})$, gilt $A^m v = v$ mit $v \in \mathbb{Z}^n \setminus \{0\}$ genau dann, wenn 1 ein Eigenwert von A^m ist. Aber 1 ist Eigenwert von A^m genau dann, wenn A einen Eigenwert λ mit $\lambda^m = 1$ besitzt. Folglich ist θ_A ergodisch (und dann gleichzeitig stark-mischend) bezüglich $\mathrm{m}_{\mathrm{K}^n}$ genau dann, wenn A keine Einheitswurzel als Eigenwert besitzt.

7 Isomorphie von maßtreuen Abbildungen

Lemma 7.1 *Sei $(X, \mathcal{F}, \mu)$ ein Wahrschlichkeitsraum und sei $T : X \to X$ eine messbare Abbildung, die maßtreu bezüglich μ ist. Sei ferner $(Y, \mathcal{B})$ ein Messraum, sei $\varphi : X \to Y$ eine $\mathcal{F}|\mathcal{B}$-messbare Abbildung und sei $S : Y \to Y$ eine messbare Abbildung mit $S \circ \varphi = \varphi \circ T$. Dann ist S maßtreu bezüglich des Bildmaßes von μ unter φ.*

Beweis Sei ν das Bildmaß von μ unter φ. Für alle $B \in \mathcal{B}$ gilt dann

$$\begin{aligned}\nu(S^{-1}(B)) &= \mu(\varphi^{-1}(S^{-1}(B))) = \mu((S \circ \varphi)^{-1}(B)) \\ &= \mu((\varphi \circ T)^{-1}(B)) = \mu(T^{-1}(\varphi^{-1}(B))) = \mu(\varphi^{-1}(B)) = \nu(B) \ . \ \square\end{aligned}$$

Sei $(X, \mathcal{F}, \mu)$ ein Wahrscheinlichkeitsraum. Für jedes $Y \in \mathcal{F}$ mit $\mu(Y) = 1$ setze $\mathcal{F}_Y = \{A \subset Y : A \in \mathcal{F}\}$ und definiere $\mu_Y : \mathcal{F}_Y \to \mathbb{R}^+$ durch $\mu_Y(A) = \mu(A)$ für jedes $A \in \mathcal{F}_Y$. Dann ist $\mathcal{F}_Y \subset \mathcal{P}(Y)$ eine σ-Algebra und $\mu_Y \in \mathrm{P}(Y, \mathcal{F}_Y)$, also ist $(Y, \mathcal{F}_Y, \mu_Y)$ ein Wahrscheinlichkeitsraum.

Sei nun $T : X \to X$ eine messbare Abbildung, die maßtreu bezüglich μ ist und mit $T(Y) \subset Y$. Dann kann eine messbare Abbildung $T_Y : Y \to Y$ definiert werden durch $T_Y(y) = T(y)$ für jedes $y \in Y$ und es ist klar, dass T_Y maßtreu bezüglich μ_Y ist.

Lemma 7.2 *Die Abbildung $T : X \to X$ ist ergodisch bzw. schwach-mischend bzw. stark-mischend bezüglich μ genau dann, wenn die Abbildung $T_Y : Y \to Y$ ergodisch bzw. schwach-mischend bzw. stark-mischend bezüglich μ_Y ist.*

Beweis Dies ist klar, da $(T^{-k}(A) \cap B) \cap Y = T_Y^{-k}(A \cap Y) \cap (B \cap Y)$ für alle A, $B \in \mathcal{F}$, $k \geq 0$. $\square$

Seien $(X_1, \mathcal{F}_1)$, $(X_2, \mathcal{F}_2)$ Messräume. Eine bijektive Abbildung $\varphi : X_1 \to X_2$ heißt *bimessbar*, wenn $\varphi(\mathcal{F}_1) = \mathcal{F}_2$, d.h., wenn $\varphi : X_1 \to X_2$ und die Umkehrabbildung $\varphi^{-1} : X_2 \to X_1$ beide messbar sind. In diesem Fall ist die Umkehrabbildung φ^{-1} auch bimessbar.

Im Folgenden seien $(X_1, \mathcal{F}_1, \mu_1)$ und $(X_2, \mathcal{F}_2, \mu_2)$ Wahrscheinlichkeitsräume und für $k = 1, 2$ sei $T_k : X_k \to X_k$ eine messbare Abbildung, die maßtreu bezüglich μ_k ist. Dann heißen T_1 und T_2 *strikt isomorph*, wenn es eine bimessbare Bijektion $\varphi : X_1 \to X_2$ gibt, so dass $T_2 \circ \varphi = \varphi \circ T_1$ und μ_2 das Bildmaß von μ_1 unter φ ist (d.h., $\mu_2(A) = \mu_1(\varphi^{-1}(A))$ für alle $A \in \mathcal{F}_2$).

Diese Definition ist symmetrisch in T_1 und T_2, da $\psi = \varphi^{-1} : X_2 \to X_1$ bimessbar ist mit $T_1 \circ \psi = \psi \circ T_2$ und μ_1 das Bildmaß von μ_2 unter ψ.

Lemma 7.3 *Sind T_1 und T_2 strikt isomorph, so ist die Abbildung $T_1 : X_1 \to X_1$ ergodisch bzw. schwach-mischend bzw. stark-mischend bezüglich μ_1 genau dann, wenn die Abbildung $T_2 : X_2 \to X_2$ ergodisch bzw. schwach-mischend bzw. stark-mischend bezüglich μ_2 ist.*

Beweis Dies ist klar: Ist $\varphi : X_1 \to X_2$ bimessbar mit $T_2 \circ \varphi = \varphi \circ T_1$ und μ_2 das Bildmaß von μ_1 unter φ, so ist

$$\begin{aligned}\mu_2(T_2^{-k}(A) \cap B) &= \mu_1(\varphi^{-1}(T_2^{-k}(A) \cap B)) \\ &= \mu_1(\varphi^{-1}(T_2^{-k}(A)) \cap \varphi^{-1}(B)) = \mu_1(T_1^{-k}(\varphi^{-1}(A)) \cap \varphi^{-1}(B))\end{aligned}$$

und $\mu_2(A)\mu_2(B) = \mu_1(\varphi^{-1}(A))\mu_1(\varphi^{-1}(B))$ für alle A, $B \in \mathcal{F}$, $k \geq 0$. □

Lemma 7.4 *Sei $(X, \mathcal{F}, \mu)$ ein Wahrscheinlichkeitsraum und sei $T : X \to X$ eine messbare Abbildung, die maßtreu bezüglich μ ist. Sei ferner $(Y, \mathcal{B})$ ein Messraum und sei $\varphi : X \to Y$ eine bimessbare Bijektion. Sei ν das Bildmaß von μ unter φ und definiere $S : Y \to Y$ durch $S = \varphi \circ T \circ \varphi^{-1}$. Dann ist S maßtreu bezüglich ν und T und S sind strikt isomorph.*

Beweis Dies folgt unmittelbar aus Lemma 7.1. □

Die Abbildungen T_1 und T_2 heißen *isomorph*, wenn es für $k = 1, 2$ ein $Y_k \in \mathcal{F}_k$ mit $\mu_k(Y_k) = 1$ und $T_k(Y_k) \subset Y_k$ gibt, so dass $(T_1)_{Y_1}$ und $(T_2)_{Y_2}$ strikt isomorph sind. Insbesondere sind strikt isomorphe Abbildungen isomorph.

Lemma 7.5 *Sei $T_3 : X_3 \to X_3$ eine weitere maßtreue Abbildung. Sind T_1 und T_2 sowie T_2 und T_3 isomorph, so sind auch T_1 und T_3 isomorph.*

Beweis Übung. □

Nach Lemma 7.5 definiert Isomorphie eine Äquivalenzrelation auf der Klasse aller maßtreuen Abbildungen, da es klar ist, dass diese Relation reflexiv und symmetrisch ist.

Satz 7.1 *Sind T_1 und T_2 isomorph, so ist die Abbildung $T_1 : X_1 \to X_1$ ergodisch bzw. schwach-mischend bzw. stark-mischend bezüglich μ_1 genau dann, wenn die Abbildung $T_2 : X_2 \to X_2$ ergodisch bzw. schwach-mischend bzw. stark-mischend bezüglich μ_2 ist.*

Beweis Dies folgt unmittelbar aus Lemma 7.2 und Lemma 7.3. □

Es werden nun einige Beispiele betrachtet. Sei $J = [0, 1)$, sei $\mathcal{B}_J$ die σ-Algebra der Borel-Teilmengen von J und sei ϱ die Einschränkung des Lebesgueschen Maßes auf $\mathcal{B}_J$, also ist $\varrho \in \mathrm{P}(J, \mathcal{B}_J)$. Definiere $\varphi : J \to \mathrm{K}$ durch $\varphi(x) = \exp(2\pi ix)$; dann ist φ eine stetige Bijektion und man sieht leicht, dass φ bimessbar ist, (obwohl die Umkehrabbildung $\varphi^{-1} : \mathrm{K} \to J$ nicht stetig ist). Ferner gilt $\varrho(\varphi^{-1}(A)) = \mathrm{m_K}(A)$ für alle $A \in \mathcal{B}_\mathrm{K}$, d.h., $\mathrm{m_K}$ ist das Bildmaß von ϱ unter φ.

Definiere $U : \mathrm{K} \to \mathrm{K}$ durch $U(z) = z^2$ und $T : J \to J$ durch

$$T(x) = \begin{cases} 2x & \text{falls } 0 \le x < \frac{1}{2} \, , \\ 2x - 1 & \text{falls } \frac{1}{2} \le x < 1 \, . \end{cases}$$

Dann gilt $\varphi \circ T = U \circ \varphi$, da $\exp(2\pi i(2x)) = \exp(2\pi i(2x - 1)) = (\exp(2\pi ix))^2$ für alle $x \in J$. Ferner ist U ist ein stetiger surjektiver Endomorphismus von K und damit ist nach Satz 6.4 (2) U maßtreu bezüglich $\mathrm{m_K}$. Setze $\psi = \varphi^{-1}$; dann ist ϱ das Bildmaß von $\mathrm{m_K}$ unter ψ und es gilt $T = \psi \circ U \circ \psi^{-1}$. Nach Lemma 7.4 ist also T maßtreu bezüglich ϱ, und U und T sind strikt isomorph.

Sei nun $E = \{0, 1\}$ und setze $E_\infty = \mathrm{Abb}(\mathbb{N}, E)$, also ist

$$\mathrm{Abb}(\mathbb{N}, E) = \{\{z_n\}_{n\ge0} : z_n \in \{0, 1\} \text{ für alle } n \ge 0\} \, .$$

Für jedes $m \in \mathbb{N}$ sei $\mathrm{p}_m : E_\infty \to E$ gegeben durch $\mathrm{p}_m(\{x_n\}_{n\ge0}) = x_m$ und setze $\mathcal{E}_\infty = \sigma\big(\bigcup_{n\ge0} \mathrm{p}_m^{-1}(\mathcal{P}(E))\big)$; definiere $S : E_\infty \to E_\infty$ durch $S(\{x_n\}_{n\ge0}) = \{x'_n\}_{n\ge0}$, wobei $x'_n = x_{n+1}$. Nach Satz 3.2 gibt es ein eindeutiges $\nu \in \mathrm{P}(E_\infty, \mathcal{E}_\infty)$, so dass

$$\nu\Big(\bigcap_{n=k}^{\ell} \mathrm{p}_n^{-1}(B_n)\Big) = 2^{k-\ell-1} \prod_{n=k}^{\ell} |B_n|$$

für alle $B_k, \ldots, B_\ell \subset E$ und alle $k, \ell \in \mathbb{N}$ mit $k \le \ell$. Ferner ist S maßtreu bezüglich ν. Es wird gezeigt, dass S und T isomorph sind. Für jedes $n \ge 1$ sei $D_n = \{k/2^n : 0 \le k < 2^n\}$ und setze $D = \bigcup_{n\ge1} D_n$ und $Y = J \setminus D$. Da $T^{-1}(D_n) \subset D_{n+1}$ gilt $T(Y) \subset Y$, und da D abzählbar ist, ist $Y \in \mathcal{B}_J$ und $\varrho(Y) = 1$. Nun hat jedes $y \in Y$ eine eindeutige Darstellung $y = \sum_{n\ge0} 2^{-n-1}\epsilon_n(y)$ mit $\epsilon_n(y) \in E$ für jedes n, und $\{n \ge 0 : \epsilon_n(y) = 0\}$ und $\{n \ge 0 : \epsilon_n(y) = 1\}$ sind beide unendliche Mengen, d.h., $\sum_{n\ge0} \epsilon_n(y) = \sum_{n\ge0}(1 - \epsilon_n(y)) = \infty$. Sei

$$Z = \Big\{\{z_n\}_{n\ge0} \in E_\infty : \sum_{n\ge0} z_n = \sum_{n\ge0}(1 - z_n) = \infty\Big\} \, ;$$

eine Abbildung $\varphi : Y \to Z$ kann also definiert werden durch

$$\varphi(y) = \{\epsilon_n(y)\}_{n\ge0}$$

für jedes $y \in Y$ und man sieht leicht, dass φ eine Bijektion ist. Nun gilt $S(Z) \subset Z$ und ferner ist $Z \in \mathcal{E}_\infty$ und $\nu(Z) = 1$, da $E_\infty \setminus Z$ abzählbar ist und $\nu(\{z\}) = 0$ für jedes $z \in E_\infty$. Sei $y \in Y \cap [0, \frac{1}{2})$; dann gilt

$$T(y) = 2\sum_{n\geq 0} 2^{-n-1}\epsilon_n(y) = \sum_{n\geq 0} 2^{-n}\epsilon_n(y) = \sum_{n\geq 0} 2^{-n-1}\epsilon_{n+1}(y)$$

da $\epsilon_0(y) = 0$, und für $y \in Y \cap [\frac{1}{2}, 1)$ gilt andererseits

$$T(y) = 2\sum_{n\geq 0} 2^{-n-1}\epsilon_n(y) - 1 = \sum_{n\geq 0} 2^{-n}\epsilon_n(y) - 1 = \sum_{n\geq 0} 2^{-n-1}\epsilon_{n+1}(y)$$

da $\epsilon_0(y) = 1$. Daraus folgt, dass $\epsilon_n(T(y)) = \epsilon_{n+1}(y)$ für alle $y \in Y$, $n \geq 0$, und damit ist $\varphi(T(y)) = S(\varphi(y))$ für alle $y \in Y$, d.h., $\varphi \circ T_Y = S_Z \circ \varphi$.

Für jedes $n \geq 0$ sei $\mathcal{A}_n$ die Menge aller Teilmengen von J der Form $\bigcup_{k\in\Lambda} I_n(k)$, wobei $I_n(k) = [k2^{-n-1}, (k+1)2^{-n-1})$ und $\Lambda \subset \{0, 1, \ldots, 2^{n+1} - 1\}$. Dann ist $\mathcal{A}_n$ eine (endliche) σ-Algebra, $\mathcal{A}_n \subset \mathcal{A}_{n+1}$ für jedes $n \geq 0$ und $\mathcal{A} = \bigcup_{n\geq 0} \mathcal{A}_n$ ist eine Algebra mit $\sigma(\mathcal{A}) = \mathcal{B}_J$. Für jedes $n \geq 0$ sei ferner $\mathcal{A}'_n$ die Menge aller Teilmengen von E_∞ der Form $\bigcup_{k\in\Lambda'} I'_n(\gamma)$, wobei $I'_n((u_0, \ldots, u_n)) = \bigcap_{k=0}^n \mathrm{p}_k^{-1}(\{u_k\})$ und wobei $\Lambda' \subset \{(u_0, \ldots, u_n) : u_k \in E \text{ für jedes } k\} = E^{n+1}$ Dann ist $\mathcal{A}'_n$ eine (endliche) σ-Algebra, $\mathcal{A}'_n \subset \mathcal{A}'_{n+1}$ für jedes $n \geq 0$ und $\mathcal{A}' = \bigcup_{n\geq 0} \mathcal{A}'_n$ ist eine Algebra mit $\sigma(\mathcal{A}') = \mathcal{E}_\infty$.

Für jedes $n \geq 0$ werden die beiden σ-Algebren $\mathcal{A}_n$ und $\mathcal{A}'_n$ jeweils von 2^{n+1} Atomen erzeugt (nämlich von $I_n(k)$, $0 \leq k < 2^{n+1}$, bzw. von $I'_n(\gamma)$, $\gamma \in E^{n+1}$) und $\varrho(I_n(k) = \nu(I'_n(\gamma)) = 2^{-n-1}$ für jedes $0 \leq k < 2^{n+1}$ und jedes $I'_n(\gamma)$, $\gamma \in E^{n+1}$. Seien $u_0, \ldots, u_n \in E$' dann sieht man leicht, dass

$$\varphi^{-1}(I'_n((u_0, \ldots, u_n)) \cap Z) = I_n\Big(\sum_{k=0}^n u_k 2^{n-k}\Big) \cap Y\ .$$

Folglich ist $\varphi^{-1}(B \cap Z) \in \mathcal{B}_J$ für jedes $B \in \mathcal{A}'_n$, $\varphi(A \cap Y) \in \mathcal{E}_\infty$ für jedes $A \in \mathcal{A}_n$ und $\varrho(\varphi^{-1}(B \cap Z)) = \nu(B \cap Z)$ für jedes $B \in \mathcal{A}'_n$. und damit ist $\varphi^{-1}(B \cap Z) \in \mathcal{B}_J$ für jedes $B \in \mathcal{A}'$, $\varphi(A \cap Y) \in \mathcal{E}_\infty$ für jedes $A \in \mathcal{A}$ und $\varrho(\varphi^{-1}(B \cap Z)) = \nu(B \cap Z)$ für jedes $B \in \mathcal{A}'$. Da aber $\mathcal{E}_\infty = \sigma(\mathcal{A}')$ und $\mathcal{B}_J = \sigma(\mathcal{A})$, ergibt sich daraus, dass $\varphi^{-1}(B \cap Z) \in \mathcal{B}_J$ für jedes $B \in \mathcal{E}_\infty$, $\varphi(A \cap Y) \in \mathcal{E}_\infty$ für jedes $A \in \mathcal{B}_J$ und $\varrho(\varphi^{-1}(B \cap Z)) = \nu(B \cap Z)$ für jedes $B \in \mathcal{E}_\infty$. Dies zeigt, dass S_Z und T_Y strikt isomorph sind und also sind S und T isomorph.

Sei nun $I = [0, 1]$, sei $\mathcal{B}_I$ die σ-Algebra der Borel-Teilmengen von I und sei λ die Einschränkung des Lebesgueschen Maßes auf $\mathcal{B}_I$, also ist $\lambda \in \mathrm{P}(I, \mathcal{B}_I)$. Definiere eine Abbildung $T' : I \to I$ durch

$$T'(x) = \begin{cases} 2x & \text{falls } 0 \leq x \leq \frac{1}{2}\ , \\ 2 - 2x & \text{falls } \frac{1}{2} \leq x \leq 1\ . \end{cases}$$

Dann ist T' maßtreu bezüglich λ (Übungsaufgabe 3), und es wird jetzt gezeigt, dass T' zu einer Abbildung $S' : E_\infty \to E_\infty$ isomorph ist. Definiere $\tau : E_\infty \to E_\infty$ durch $\tau(\{z_n\}_{n\geq 0}) = \{z'_n\}_{n\geq 0}$, wobei $z'_0 = z_0$ und (für $n \geq 1$)

$$z'_n = \begin{cases} z_n & \text{falls } z_0 = 0 \ , \\ 1 - z_n & \text{falls } z_0 = 1 \ . \end{cases}$$

Man sieht leicht, dass τ maßtreu bezüglich ν ist und also ist auch $S' = S \circ \tau$ maßtreu bezüglich ν. Man beachte, dass $S'(\{z_n\}_{n\geq 0}) = \{z''_n\}_{n\geq 0}$, wobei

$$z''_n = \begin{cases} z_{n+1} & \text{falls } z_0 = 0 \ , \\ 1 - z_{n+1} & \text{falls } z_0 = 1 \ . \end{cases}$$

Ferner ist $T'(Y) \subset Y$, $Y \in \mathcal{B}_I$ und $\lambda(Y) = 1$. Ist $y \in Y$, so gilt

$$\epsilon_n(T'(y)) = \begin{cases} \epsilon_{n+1}(y) & \text{falls } \epsilon_0(y) = 0 \ , \\ 1 - \epsilon_{n+1}(y) & \text{falls } \epsilon_0(y) = 1 \ , \end{cases}$$

und damit ist $\varphi \circ T'_Y = S'_Z \circ \varphi$. Daraus ergibt sich (wie im letzten Fall), dass S'_Z und T'_Y strikt isomorph sind, und also sind S' und T' isomorph.

Lemma 7.6 *Die Abbildungen S und S' sind strikt isomorph.*

Beweis Definiere $\psi : E_\infty \to E_\infty$ durch $\psi(\{z_n\}_{n\geq 0}) = \{w_n\}_{n\geq 0}$, wobei

$$w_n = \frac{1}{2}\Big(1 - \prod_{k=0}^{n}(1 - 2z_k)\Big) \ .$$

Dann ist ψ eine bimessbare Bijektion mit $\nu(\psi^{-1}(B)) = \nu(B)$ für alle $B \in \mathcal{E}_\infty$ und es gilt $\psi \circ S = S' \circ \psi$. (Der Beweis dafür ist eine Übung.) $\square$

Da S' und T' sowie S und T isomorph sind, folgt nun aus Lemma 7.6, dass auch T und T' isomorph sind.

8 Entropie

Sei $\Phi : \mathbb{R}^+ \to \mathbb{R}$ die Abbildung, die gegeben ist durch

$$\Phi(t) = \begin{cases} t \log t & \text{falls } t > 0 \ , \\ 0 & \text{falls } t = 0 \ . \end{cases}$$

Dann ist $\Phi(t) < 0$ für alle $t \in (0,1)$, $\Phi(t) > 0$ für alle $t > 1$, $\Phi(t) > -\mathrm{e}^{-1} = \Phi(\mathrm{e}^{-1})$ für alle $t \in \mathbb{R}^+ \setminus \{\mathrm{e}^{-1}\}$ und Φ ist stetig, da $\lim_{t \downarrow 0} t \log t = 0$.

Lemma 8.1 *Die Abbildung Φ ist streng konvex: Es gilt*

$$\Phi(\lambda x + (1-\lambda)y) < \lambda\Phi(x) + (1-\lambda)\Phi(y)$$

für alle $0 \le x < y$ und alle $\lambda \in (0,1)$.

Beweis Man beachte, dass $\Phi''(t) = 1/t > 0$ für alle $t > 0$. Sei $0 \le x < y$ und $\lambda \in (0,1)$ und setze $z = \lambda x + (1-\lambda)y$; also ist $x < z < y$. Nach dem Mittelwertsatz gibt es $u \in (x,z)$ und $v \in (z,y)$, so dass $\Phi(z) - \Phi(x) = \Phi'(u)(z-x)$ und $\Phi(y) - \Phi(z) = \Phi'(v)(y-z)$, und da $u < v$ und $\Phi''(t) > 0$ für alle $t > 0$, ist $\Phi'(u) < \Phi'(v)$. Daraus ergibt sich, dass

$$\begin{aligned} \Phi(z) - \Phi(x) &= \Phi'(u)(z-x) \\ &= \Phi'(u)(y-x)(1-\lambda) = \Phi'(u)(y-z)(1-\lambda)/\lambda \\ &< \Phi'(v)(y-z)(1-\lambda)/\lambda = (\Phi(y) - \Phi(z))(1-\lambda)/\lambda \end{aligned}$$

und folglich gilt $\lambda(\Phi(z) - \Phi(x)) < (1-\lambda)(\Phi(y) - \Phi(z))$, d.h.,

$$\Phi(\lambda x + (1-\lambda)y) = \Phi(z) < \lambda\Phi(x) + (1-\lambda)\Phi(y) \ . \quad \square$$

Aus Lemma 8.1 gilt insbesondere für alle x, $y \in \mathbb{R}^+$ und alle $\lambda \in [0,1]$, dass

$$\Phi(\lambda x + (1-\lambda)y) \le \lambda\Phi(x) + (1-\lambda)\Phi(y) \ .$$

Per Induktion nach n folgt dann, dass für alle $n \ge 2$

$$\Phi\Big(\sum_{k=1}^{n} \lambda_k x_k\Big) \le \sum_{k=1}^{n} \lambda_k \Phi(x_k)$$

für alle $x_1, \ldots, x_n \in \mathbb{R}^+$ und alle $\lambda_1, \ldots, \lambda_n \in [0,1]$ mit $\sum_{k=1}^{n} \lambda_k = 1$.

Sei E eine nichtleere endliche Menge und sei $\alpha : E \to \mathbb{R}^+$ ein Wahrscheinlichkeitsvektor (d.h., $\sum_{u \in E} \alpha(u) = 1$). Setze nun

$$\mathrm{H}(\alpha) = -\sum_{u \in E} \Phi(\alpha(u)) \ ;$$

$\mathrm{H}(\alpha)$ heißt die *Entropie* von α.

Lemma 8.2 *Es gilt* $0 \le \mathrm{H}(\alpha) \le \log|E|$.

Beweis Da $\Phi(t) \le 0$ für alle $t \in [0,1]$, ist $\mathrm{H}(\alpha) \ge 0$, und nach der Konvexität von Φ gilt ferner, dass

$$\begin{aligned}\mathrm{H}(\alpha) &= -|E| \sum_{u \in E} |E|^{-1}\Phi(\alpha(u)) \\ &\le -|E|\Phi\Big(\sum_{u \in E} |E|^{-1}\alpha(u)\Big) = -|E|\Phi(|E|^{-1}) = \log|E| \;. \quad \square\end{aligned}$$

Seien nun E und F nichtleere endliche Mengen und sei $\alpha : E \times F \to \mathbb{R}^+$ ein Wahrscheinlichkeitsvektor. Definiere $\alpha_E : E \to \mathbb{R}^+$ und $\alpha_F : F \to \mathbb{R}^+$ durch $\alpha_E(u) = \sum_{v \in F} \alpha(u,v)$ und $\alpha_F(v) = \sum_{u \in E} \alpha(u,v)$, und also sind α_E und α_F die Bildmaße von α unter den Projektionen $E \times F \to E$ und $E \times F \to F$. Für $u \in E$ und $v \in F$ definiere ferner

$$\alpha(u|v) = \begin{cases} \alpha(u,v)/\alpha_F(v) & \text{falls } \alpha_F(v) > 0 \;, \\ 0 & \text{falls } \alpha_F(v) = 0 \;, \end{cases}$$

$$\alpha(v|u) = \begin{cases} \alpha(u,v)/\alpha_E(u) & \text{falls } \alpha_E(u) > 0 \;, \\ 0 & \text{falls } \alpha_E(u) = 0 \;, \end{cases}$$

wobei $\alpha(u|v)$ und $\alpha(v|u)$ natürlich als bedingte Wahrscheinlichkeiten angesehen werden sollen. Definiere nun

$$\begin{aligned}\mathrm{H}(\alpha|F) &= \sum_{v \in F}\Big(-\sum_{u \in E} \Phi(\alpha(u|v))\Big)\alpha_F(v) \;, \\ \mathrm{H}(\alpha|E) &= \sum_{u \in E}\Big(-\sum_{v \in F} \Phi(\alpha(u|v))\Big)\alpha_E(u) \;;\end{aligned}$$

$\mathrm{H}(\alpha|F)$ bzw. $\mathrm{H}(\alpha|E)$ heißt die *bedingte Entropie* von α bezüglich F bzw. E. Ist $v \in F$ mit $\alpha_F(v) > 0$, so ist $\alpha(\cdot|v) : E \to \mathbb{R}^+$ ein Wahrscheinlichkeitsvektor und daraus ergibt sich, dass

$$\mathrm{H}(\alpha|F) = \sum_{v \in F} \mathrm{H}(\alpha(\cdot|v))\alpha_F(v) \;,$$

und genauso gilt $\mathrm{H}(\alpha|E) = \sum_{u \in E} \mathrm{H}(\alpha(\cdot|u))\alpha_E(u)$.

Satz 8.1 *Seien E und F nichtleere endliche Mengen und sei $\alpha : E \times F \to \mathbb{R}^+$ ein Wahrscheinlichkeitsvektor. Dann gilt:*

(1) $\mathrm{H}(\alpha) = \mathrm{H}(\alpha_E) + \mathrm{H}(\alpha|E)$.

(2) $\mathrm{H}(\alpha|E) \le \mathrm{H}(\alpha_F)$.

(3) $\mathrm{H}(\alpha) \le \mathrm{H}(\alpha_E) + \mathrm{H}(\alpha_F)$.

Beweis (1) Sei $u \in E$ mit $\alpha_E(u) > 0$; dann gilt

$$\begin{aligned}\Phi(\alpha(u,v)) - \Phi(\alpha(v|u))\alpha_E(u) &= \alpha(u,v)\log\alpha(u,v) - \alpha(u,v)\log(\alpha(u,v)/\alpha_E(u)) \\ &= \alpha(u,v)\log(\alpha_E(u))\ ,\end{aligned}$$

und daraus ergibt sich, dass

$$\begin{aligned}\sum_{v\in F}\Phi(\alpha(u,v)) - \sum_{v\in F}\Phi(\alpha(v|u))\alpha_E(u) &= \sum_{v\in F}\alpha(u,v)\log(\alpha_E(u)) \\ &= \alpha_E(u)\log\alpha_E(u) = \Phi(\alpha_E(u))\ .\end{aligned}$$

Für alle $u \in E$ gilt also

$$\sum_{v\in F}\Phi(\alpha(u,v)) = \Phi(\alpha_E(u)) + \sum_{v\in F}\Phi(\alpha(v|u))\alpha_E(u)\ ,$$

da dies trivial richtig ist, wenn $\alpha_E(u) = 0$. Folglich ist

$$\begin{aligned}\mathrm{H}(\alpha) &= -\sum_{u\in E}\sum_{v\in F}\Phi(\alpha(u,v)) \\ &= -\sum_{u\in E}\Phi(\alpha_E(u)) - \sum_{u\in E}\sum_{v\in F}\Phi(\alpha(v|u))\alpha_E(u) = \mathrm{H}(\alpha_E) + \mathrm{H}(\alpha|E)\ .\end{aligned}$$

(2) Man beachte, dass $\alpha(u,v) = \alpha(v|u)\alpha_E(u)$ für alle $u \in E$, $v \in F$. Sei $v \in F$; aus der Konvexität von Φ folgt, dass

$$\Phi(\alpha_F(v)) = \Phi\Big(\sum_{u\in E}\alpha(u,v)\Big) = \Phi\Big(\sum_{u\in E}\alpha_E(u)\alpha(v|u)\Big) \le \sum_{u\in E}\Phi(\alpha(v|u))\alpha_E(u)\ ,$$

da $\sum_{u\in E}\alpha_E(u) = 1$. Daraus ergibt sich, dass

$$\mathrm{H}(\alpha|E) = \sum_{v\in F}\Big(-\sum_{u\in E}\Phi(\alpha(v|u))\alpha_E(u)\Big) \le -\sum_{v\in F}\Phi(\alpha_F(v)) = \mathrm{H}(\alpha_F)\ .$$

(3) Dies folgt unmittelbar aus (1) und (2). $\square$

Im Folgenden sei $(X, \mathcal{F}, \mu)$ ein Wahrscheinlichkeitsraum und sei $T : X \to X$ eine messsbare Abbildung, die maßtreu bezüglich μ ist. Sei E eine nichtleere endliche Menge und sei $f : X \to E$ eine $\mathcal{F}$-messbare Abbildung (d.h., $f^{-1}(\{u\}) \in \mathcal{F}$ für jedes $u \in E$). Für jedes $n \ge 1$ definiere $\alpha_n : E^n \to \mathbb{R}^+$ durch

$$\alpha_n((u_0,\ldots,u_{n-1})) = \mu\Big(\bigcap_{k=0}^{n-1}(f\circ T^k)^{-1}(\{u_k\})\Big)$$

für alle $(u_0,\ldots,u_{n-1}) \in E^n$. Dann ist α_n ein Wahrscheinlichkeitsvektor, und α_n ist das Bildmaß von μ unter der Abbildung $(f, f\circ T, \ldots, f\circ T^{n-1}) : X \to E^n$.

Satz 8.2 *Die Folge* $\{n^{-1}\mathrm{H}(\alpha_n)\}_{n\geq 1}$ *konvergiert, und der Grenzwert*

$$\mathrm{h}(T,f) = \lim_{n\to\infty} \frac{1}{n}\mathrm{H}(\alpha_n)$$

liegt im Intervall $[0, \log|E|]$.

Beweis Nach Lemma 8.2 gilt $0 \leq n^{-1}\mathrm{H}(\alpha^n) \leq n^{-1}\log|E^n| = \log|E|$ für jedes $n \geq 1$. Seien m, $n \geq 1$; dann ist $E^{m+n} = A \times B$, wobei $A = E^m$ und $B = E^n$; setze $\beta = \alpha_{m+n}$. Nun ist es klar, dass $\beta_A = \alpha_m$; andererseits ist

$$\begin{aligned}
\beta_B((v_0,\ldots,v_{n-1})) &= \sum_{(u_0,\ldots,u_{m-1})\in E^m} \alpha_{m+n}((u_0,\ldots,u_{m-1},v_0,\ldots,v_{n-1}))\\
&= \sum_{(u_0,\ldots,u_{m-1})\in E^m} \mu\Big(\bigcap_{k=0}^{m-1}(f\circ T^k)^{-1}(\{u_k\}) \cap \bigcap_{\ell=0}^{n-1}(f\circ T^{m+\ell})^{-1}(\{v_\ell\})\Big)\\
&= \mu\Big(\bigcap_{\ell=0}^{n-1}(f\circ T^{m+\ell})^{-1}(\{v_\ell\})\Big) = \mu\Big(T^{-m}\Big(\bigcap_{\ell=0}^{n-1}(f\circ T^{\ell})^{-1}(\{v_\ell\})\Big)\Big)\\
&= \mu\Big(\bigcap_{\ell=0}^{n-1}(f\circ T^{\ell})^{-1}(\{v_\ell\})\Big) = \alpha_n((v_0,\ldots,v_{n-1}))\ ;
\end{aligned}$$

d.h., $\beta_B = \alpha_n$. Nach Satz 8.1 (3) ist also

$$\mathrm{H}(\alpha_{m+n}) = \mathrm{H}(\beta) \leq \mathrm{H}(\beta_A) + \mathrm{H}(\beta_B) = \mathrm{H}(\alpha_m) + \mathrm{H}(\alpha_n)\ ,$$

und damit ist $\mathrm{H}(\alpha_{m+n}) \leq \mathrm{H}(\alpha_m) + \mathrm{H}(\alpha_n)$ für alle m, $n \geq 1$. Die Existenz des Limes $\lim_{n\to\infty} n^{-1}\mathrm{H}(\alpha_n)$ folgt nun unmittelbar aus dem nächsten Lemma. □

Eine reelle Folge $\{a_n\}_{n\geq 1}$ heißt *subadditiv*, wenn $a_{m+n} \leq a_m + a_n$ für alle m, $n \geq 1$.

Lemma 8.3 *Für jede subadditive reelle Folge* $\{a_n\}_{n\geq 1}$ *gilt*

$$\lim_{n\to\infty} a_n/n = \inf_{n\geq 1} a_n/n\ ,$$

obwohl es möglich ist, dass $\inf_{n\geq 1} a_n/n = -\infty$.

Beweis Man beachte zunächst, dass $a_{km} \leq k a_m$ für alle k, $m \geq 1$. Sei $m \geq 1$ und setze $b_m = \max\{a_1,\ldots,a_m\}$. Nun hat jedes $n \geq 2m$ eine Darstellung der Form $n = km + \ell$ mit $k \geq 1$ und $1 \leq \ell \leq m$ und folglich ist für $n \geq 2m$

$$a_n/n = a_{km+\ell}/n \leq a_{km}/n + a_\ell/n \leq (k a_m)/n + b_m/n \leq a_m/m + b_m/n\ .$$

Damit ist $\limsup_{n\ge 1} a_n/n \le a_m/m$ für jedes $m \ge 1$ und also ist

$$\limsup_{n\ge 1} a_n/n \le \inf_{m\ge 1} a_m/m \le \liminf_{n\ge 1} a_n/n \; .$$

Dies zeigt, dass $\lim\limits_{n\to\infty} a_n/n = \inf\limits_{n\ge 1} a_n/n$. $\square$

Sei wieder $E_\infty = \mathrm{Abb}(\mathbb{N}, E)$ und definiere eine Abbildung $f \triangleleft T : X \to E_\infty$ durch $(f \triangleleft T)(x) = \{(f \circ T^n)(x)\}_{n\ge 0}$. Für $u = \{u_n\}_{n\ge 0} \in E_\infty$, $m \ge 1$, setze

$$C_m(u) = \{\{z_n\}_{n\ge 0} \in E_\infty : z_n = u_n \text{ für } n = 0, \dots, m-1\} \; ;$$

es ist klar, dass $C_m(u) \in \mathcal{E}_\infty$. Sei ν das Bildmaß von μ unter $f \triangleleft T$.

Lemma 8.4 *Es gilt* $\mathrm{H}(\alpha_m) = -\int \log(\nu(C_m(u)))\, d\nu(u)$ *für jedes* $m \ge 1$, *und damit ist insbesondere nach Satz 8.2*

$$\lim_{m\to\infty} -\frac{1}{m} \int \log(\nu(C_m(u)))\, d\nu(u) = \mathrm{h}(T, f) \; .$$

Beweis Für $(u_0, \dots, u_{m-1}) \in E^m$ setze

$$C'_m((u_0, \dots, u_{m-1})) = \{\{z_n\}_{n\ge 0} \in E_\infty : z_n = u_n \text{ für } n = 0, \dots, m-1\} \; ;$$

für jedes $u = \{u_n\}_{n\ge 0} \in E_\infty$ gilt also $C'_m((u_0, \dots, u_{m-1})) = C_m(u)$. Nun ist

$$\begin{aligned}
\mathrm{H}(\alpha_m) &= \sum_{(u_0,\dots,u_{m-1})\in E^m} \Phi(\alpha_n((u_0, \dots, u_{m-1}))) \\
&= \sum_{(u_0,\dots,u_{m-1})\in E^m} \Phi\Big(\mu\Big(\bigcap_{k=0}^{m-1} (f \circ T^k)^{-1}(\{u_k\})\Big)\Big) \\
&= \sum_{(u_0,\dots,u_{m-1})\in E^m} \Phi(\mu((f \triangleleft T)^{-1}(C'_m(u_0, \dots, u_{m-1})))) \\
&= \sum_{(u_0,\dots,u_{m-1})\in E^m} \Phi(\nu(C'_m(u_0, \dots, u_{m-1}))) \\
&= - \sum_{(u_0,\dots,u_{m-1})\in E^m} \log(\nu(C'_m(u_0, \dots, u_{m-1})))\nu(C'_m(u_0, \dots, u_{m-1})) \\
&= - \int \log(\nu(C_m(u)))\, d\nu(u) \; . \quad \square
\end{aligned}$$

Nach Satz 3.3 und Lemma 3.2 ist die Abbildung $S : E_\infty \to E_\infty$ maßtreu bezüglich ν, und ist T ergodisch bezüglich μ, so ist S ergodisch bezüglich ν. (Sei $B \in \mathcal{E}_\infty$ mit $S^{-1}(B) = B$; da $(f \triangleleft T) \circ T = S \circ (f \triangleleft T)$, ist $(f \triangleleft T)^{-1}(B) = T^{-1}((f \triangleleft T)^{-1}(B))$, und damit ist $\nu(B) = \mu((f \lhd T)^{-1}(B))$ entweder 0 oder 1.) Ist S ergodisch, so gilt in der Tat, dass

$$\lim_{m\to\infty} -\frac{1}{m}\log(\nu(C_m(\cdot))) = \mathrm{h}(T,f) \quad \nu\text{-fast sicher} .$$

Dies ist der Satz von Shannon-McMillan-Brieman (der aber hier nicht behandelt wird).

Sei $W(T)$ die Menge aller reellen Zahlen x, für die es eine nichtleere endliche Menge E und eine $\mathcal{F}$-messbare Abbildung $f : X \to E$ gibt, so dass $\mathrm{h}(T,f) = x$. Die *Entropie* $\mathrm{h}(T)$ von T wird nun definiert durch

$$\mathrm{h}(T) = \sup W(T) ;$$

es gilt also $0 \leq \mathrm{h}(T) \leq +\infty$.

Lemma 8.5 *Sei $(Y, \mathcal{B})$ ein Messraum, $\varphi : X \to Y$ eine $\mathcal{F}|\mathcal{B}$-messbare Abbildung und $S : Y \to Y$ eine messbare Abbildung mit $S \circ \varphi = \varphi \circ T$. Sei ν das Bildmaß von μ unter φ (und nach Lemma 7.1 ist also S maßtreu bezüglich ν). Dann gilt $\mathrm{h}(S) \leq \mathrm{h}(T)$.*

Beweis Sei E eine nichtleere endliche Menge und sei $f : Y \to E$ eine $\mathcal{B}$-messbare Abbildung. Dann ist $f \circ \varphi : X \to E$ $\mathcal{F}$-messbar und $(f \circ \varphi) \lhd T = (f \lhd S) \circ \varphi$, da $(f \circ \varphi) \circ T^n = (f \circ S^n) \circ \varphi$ für jedes $n \geq 0$. Damit sind das Bildmaß von μ unter $(f \circ \varphi) \lhd T$ und das Bildmaß von ν unter $f \lhd S$ gleich, und daraus ergibt sich nach Lemma 8.4, dass $\mathrm{h}(S,f) = \mathrm{h}(T, f \circ \varphi)$. Dies zeigt, dass $W(S) \subset W(T)$ und also ist $\mathrm{h}(S) \leq \mathrm{h}(T)$. $\square$

Satz 8.3 *Sind T_1 und T_2 isomorphe maßtreue Abbildungen, so ist $\mathrm{h}(T_1) = \mathrm{h}(T_2)$.*

Beweis Sind T_1 und T_2 strikt isomorph, so folgt unmittelbar aus Lemma 8.5, dass $\mathrm{h}(T_1) = \mathrm{h}(T_2)$. Betrachte also $Y \in \mathcal{F}$ mit $\mu(Y) = 1$ und $T(Y) \subset Y$. Für die Abbildung $i : Y \to X$ mit $i(y) = y$ für alle $y \in Y$ gilt $T \circ i = i \circ T_Y$ und μ ist das Bildmaß von μ_Y unter i; also ist nach Lemma 8.5 $\mathrm{h}(T) \leq \mathrm{h}(T_Y)$. Sei umgekehrt E eine nichtleere endliche Menge und $f : Y \to E$ eine $\mathcal{F}_Y$-messbare Abbildung; wähle $v \in E$ und definiere $f' : X \to E$ durch

$$f'(x) = \begin{cases} f(x) & \text{falls } x \in Y , \\ v & \text{falls } x \in X \setminus Y . \end{cases}$$

Dann sind das Bildmaß von μ_Y unter $f \lhd T_Y$ und das Bildmaß von μ unter $f' \lhd T$ gleich und folglich ist nach Lemma 8.4 $\mathrm{h}(T_Y, f) = \mathrm{h}(T, f')$. Daraus ergibt sich, dass $\mathrm{h}(T_Y) \leq \mathrm{h}(T)$ und daher ist $\mathrm{h}(T_Y) = \mathrm{h}(T)$. $\square$

Für die Berechnung der Entropie ist es nützlich, mit einer Umformulierung der Definition von $\mathrm{h}(T)$ zu arbeiten. Eine endliche Teilmenge ξ von $\mathcal{F} \setminus \{\varnothing\}$ heißt *endliche Partition* vom Messraum $(X, \mathcal{F})$, wenn jedes Element von X in genau einem Element von ξ liegt. (Die Mengen in ξ sind also paarweise disjunkt und ihre Vereinigung ist ganz X.)

Ist E eine nichtleere endliche Menge und $f : X \to E$ eine $\mathcal{F}$-messbare Abbildung, so ist $\{f^{-1}(\{u\}) : u \in f(X)\}$ eine endliche Partition von $(X, \mathcal{F})$, die mit $\pi(f)$ bezeichnet wird. Für jede endliche Partition ξ von $(X, \mathcal{F})$ definiere $\mathsf{f}_\xi : X \to \xi$ durch $\mathsf{f}_\xi(x) = A$ für jedes $x \in A \in \xi$; dann ist f_ξ $\mathcal{F}$-messbar und $\pi(\mathsf{f}_\xi) = \xi$.

Für eine endliche Partition ξ setze $\mathrm{h}(T, \xi) = \mathrm{h}(T, \mathsf{f}_\xi)$.

Satz 8.4 *Es gilt* $\mathrm{h}(T) = \sup\{\mathrm{h}(T, \xi) : \xi$ *eine endliche Partition von* $(X, \mathcal{F})\}$.

Beweis Für jede endliche Partition ξ von $(X, \mathcal{F})$ ist $\mathrm{h}(T, \xi) = \mathrm{h}(T, \mathsf{f}_\xi) \leq \mathrm{h}(T)$. Ist andererseits $f : X \to E$ eine $\mathcal{F}$-messbare Abbildung mit E einer nichtleeren endlichen Menge, so sieht man leicht, dass $\mathrm{h}(T, f) = \mathrm{h}(T, \mathsf{f}_\xi) = \mathrm{h}(T, \xi)$, wobei $\xi = \pi(f)$. (Für $k = 1, 2$ sei E_k eine nichtleere endliche Menge und sei $f_k : X \to E_k$ eine $\mathcal{F}$-messbare Abbildung. Gibt es eine bijektive Abbildung $\tau : E_1 \to E_2$ mit $\tau \circ f_1 = f_2$, so ist $\mathrm{h}(T, f_1) = \mathrm{h}(T, f_2)$.) □

Sind $\xi_m, \ldots, \xi_n$ endliche Partitionen von $(X, \mathcal{F})$, so wird mit $\bigvee_{k=m}^n \xi_k$ (oder mit $\xi_m \vee \cdots \vee \xi_n$) die Menge aller nichtleeren Teilmengen von X bezeichnet, die eine Darstellung der Form $\bigcap_{k=m}^n A_k$ haben mit $A_k \in \xi_k$ für $k = m, \ldots, n$; natürlich ist $\bigvee_{k=m}^n \xi_k$ wieder eine endliche Partition von $(X, \mathcal{F})$.

Lemma 8.6 *Für* $k = m, \ldots, n$ *sei* $f_k : X \to E_k$ *eine* $\mathcal{F}$*-messbare Abbildung mit* E_k *einer nichtleeren endlichen Menge. Sei* $g = (f_m, \ldots, f_n) : X \to E_m \times \cdots \times E_n$ *die Abbildung mit* $g(x) = (f_m(x), \ldots, f_n(x))$ *für alle* $x \in X$. *Dann gilt*

$$\pi(g) = \bigvee_{k=m}^{n} \pi(f_k) \, .$$

Beweis Übung. □

Sei ξ eine endliche Partition von $(X, \mathcal{F})$; für jede messbare Abbildung $S : X \to X$ setze $S^{-1}[\xi] = S^{-1}(\xi) \setminus \{\varnothing\}$, also ist $S^{-1}[\xi]$ die endliche Partition von $(X, \mathcal{F})$, die aus den nichtleeren Elementen der Form $S^{-1}(A)$ mit $A \in \xi$ besteht.

Lemma 8.7 *Ist* E *eine nichtleere endliche Menge,* $f : X \to E$ *eine* $\mathcal{F}$*-messbare Abbildung und* $S : X \to X$ *eine messbare Abbildung, so ist* $\pi(f \circ S) = S^{-1}[\pi(f)]$.

Beweis Dies ist klar. □

Für jede endliche Partition ξ von $(X,\mathcal{F})$ definiere die *Entropie* $\mathrm{H}(\xi)$ von ξ durch

$$\mathrm{H}(\xi) = -\sum_{A\in\xi} \Phi(\mu(A)) \ .$$

Lemma 8.8 *Sei $f : X \to E$ eine $\mathcal{F}$-messbare Abbildung mit E einer nichtleeren endlichen Menge und definiere einen Wahrscheinlichkeitsvektor $\alpha : E \to \mathbb{R}^+$ durch $\alpha(u) = \mu(f^{-1}(\{u\}))$ für jedes $u \in E$ (und also ist α das Bildmaß von μ unter f). Dann gilt* $\mathrm{H}(\alpha) = \mathrm{H}(\pi(f))$.

Beweis Es gilt

$$\begin{aligned}\mathrm{H}(\alpha) = -\sum_{u\in E} \Phi(\alpha(u)) &= -\sum_{u\in E} \Phi(\mu(f^{-1}(\{u\}))) \\ &= -\sum_{u\in f(X)} \Phi(\mu(f^{-1}(\{u\}))) = \mathrm{H}(\pi(f)) \ . \ \square\end{aligned}$$

Satz 8.5 *Für jede endliche Partition ξ von $(X,\mathcal{F})$ gilt*

$$\mathrm{h}(T,\xi) = \lim_{n\to\infty} \frac{1}{n} \mathrm{H}\Big(\bigvee_{k=0}^{n-1} T^{-k}[\xi]\Big) \ .$$

Beweis Nach Satz 8.2 (und nach der Definition von $\mathrm{h}(T,\mathsf{f}_\xi)$) gilt

$$\mathrm{h}(T,\xi) = \lim_{n\to\infty} \frac{1}{n} \mathrm{H}(\alpha_n) \ ,$$

wobei α_n das Bildmaß von μ unter der Abbildung $(\mathsf{f}_\xi, \mathsf{f}_\xi \circ T, \ldots, \mathsf{f}_\xi \circ T^{n-1})$ ist, und nach Lemma 8.8 ist $\mathrm{H}(\alpha_n) = \mathrm{H}(\pi((\mathsf{f}_\xi, \mathsf{f}_\xi \circ T, \ldots, \mathsf{f}_\xi \circ T^{n-1})))$. Aber

$$\pi((\mathsf{f}_\xi, \mathsf{f}_\xi \circ T, \ldots, \mathsf{f}_\xi \circ T^{n-1})) = \bigvee_{k=0}^{n-1} \pi(\mathsf{f}_\xi \circ T^k) = \bigvee_{k=0}^{n-1} T^{-k}[\pi(\mathsf{f}_\xi)] = \bigvee_{k=0}^{n-1} T^{-k}[\xi]$$

(nach Lemma 8.6 und Lemma 8.7), und daraus ergibt sich, das

$$\mathrm{h}(T,\xi) = \lim_{n\to\infty} \frac{1}{n} \mathrm{H}(\alpha_n) = \lim_{n\to\infty} \frac{1}{n} \mathrm{H}\Big(\bigvee_{k=0}^{n-1} T^{-k}[\xi]\Big) \ . \ \square$$

Ist ξ eine endliche Partition von $(X,\mathcal{F})$, so ist $\sigma(\xi) = \{\bigcup_{A\in\eta} A : \eta \subset \xi\}$ und insbesondere ist $\sigma(\xi)$ eine endliche Unter-σ-Algebra von $\mathcal{F}$. Ist umgekehrt $\mathcal{C}$ eine endliche Unter-σ-Algebra von $\mathcal{F}$, so gibt es eine eindeutige endliche Partition $\pi(\mathcal{C})$ von $(X,\mathcal{F})$ mit $\mathcal{C} = \sigma(\pi(\mathcal{C}))$. ($\pi(\mathcal{C})$ besteht aus den minimalen nichtleeren Elementen in $\mathcal{C}$).

Für eine endliche Unter-σ-Algebra $\mathcal{C}$ von $\mathcal{F}$ setze $\mathrm{h}(T,\mathcal{C}) = \mathrm{h}(T,\pi(\mathcal{C}))$.

Satz 8.6 *Es gilt* $\mathrm{h}(T) = \sup\{\mathrm{h}(T,\mathcal{C}) : \mathcal{C}$ *eine endliche Unter-σ-Algebra von* $\mathcal{F}\}$.

Beweis Dies folgt unmittelbar aus Satz 8.4. □

Sind $\mathcal{E}_m, \ldots, \mathcal{E}_n$ (mit $m \le n$) Unter-σ-Algebren von $\mathcal{F}$, so bezeichnet $\bigvee_{k=m}^n \mathcal{E}_k$ (oder $\mathcal{E}_m \vee \cdots \vee \mathcal{E}_n$) die σ-Algebra $\sigma\big(\bigcup_{k=m}^n \mathcal{E}_k\big)$, d.h., $\bigvee_{k=m}^n \mathcal{E}_k$ ist die kleinste σ-Algebra, die $\mathcal{E}_m, \ldots, \mathcal{E}_n$ enthält.

Lemma 8.9 *Für endliche Unter-σ-Algebren* $\mathcal{C}_m, \ldots, \mathcal{C}_n$ *von* $\mathcal{F}$ *gilt*

$$\pi\Big(\bigvee_{k=m}^n \mathcal{C}_k\Big) = \bigvee_{k=m}^n \pi(\mathcal{C}_k) .$$

Beweis Übung. □

Ist $S : X \to X$ eine messbare Abbildung und $\mathcal{C}$ eine endliche Unter-σ-Algebra von $\mathcal{F}$, so ist auch $S^{-1}(\mathcal{C})$ eine endliche Unter-σ-Algebra von $\mathcal{F}$.

Lemma 8.10 *Ist* $\mathcal{C}$ *eine endliche Unter-σ-Algebra von* $\mathcal{F}$ *und* $S : X \to X$ *eine messbare Abbildung, so ist* $S^{-1}[\pi(\mathcal{C})] = \pi(S^{-1}(\mathcal{C}))$.

Beweis Dies ist klar. □

Für eine endliche Unter-σ-Algebra $\mathcal{C}$ von $\mathcal{F}$ setze nun $\mathrm{H}(\mathcal{C}) = \mathrm{H}(\pi(\mathcal{C}))$, d.h.,

$$\mathrm{H}(\mathcal{C}) = -\sum_{A\in\pi(\mathcal{C})} \Phi(\mu(A)) .$$

Satz 8.7 *Für jede endliche Unter-σ-Algebra* $\mathcal{C}$ *von* $\mathcal{F}$ *gilt*

$$\mathrm{h}(T,\mathcal{C}) = \lim_{n\to\infty} \frac{1}{n}\mathrm{H}\Big(\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{C})\Big) .$$

Beweis Nach Lemma 8.9 und Lemma 8.10 gilt

$$\pi\Big(\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{C})\Big) = \bigvee_{k=0}^{n-1} \pi(T^{-k}(\mathcal{C})) = \bigvee_{k=0}^{n-1} T^{-k}[\pi(\mathcal{C})]$$

und daraus ergibt sich nach Satz 8.5, dass

$$\begin{aligned}\mathrm{h}(T,\mathcal{C}) &= \mathrm{h}(T,\pi(\mathcal{C})) = \lim_{n\to\infty} \frac{1}{n}\mathrm{H}\Big(\bigvee_{k=0}^{n-1} T^{-k}[\pi(\mathcal{C})]\Big)\\ &= \lim_{n\to\infty} \frac{1}{n}\mathrm{H}\Big(\pi\Big(\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{C})\Big)\Big) = \lim_{n\to\infty} \frac{1}{n}\mathrm{H}\Big(\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{C})\Big) .\end{aligned}$$ □

Die Berechnung der Entropie $\mathrm{h}(T)$ folgt am Besten über die Sätze 8.6 und 8.7.

Seien $\mathcal{C}$ und $\mathcal{E}$ Unter-σ-Algebren von $\mathcal{F}$ mit $\mathcal{C}$ endlich. Dann gilt $\mathrm{E}_\mu(\chi_A|\mathcal{E}) \geq 0$ μ-fast sicher für jedes $A \in \pi(\mathcal{C})$ und $\sum_{A\in\pi(\mathcal{C})} \mathrm{E}_\mu(\chi_A|\mathcal{E}) = 1$ μ-fast sicher; nach Lemma 8.2 ist also $0 \leq -\sum_{A\in\pi(\mathcal{C})} \Phi(\mathrm{E}_\mu(\chi_A|\mathcal{E})) \leq \log|\pi(\mathcal{C})|$ μ-fast sicher. Die *bedingte Entropie von $\mathcal{C}$ gegeben $\mathcal{E}$* wird definiert durch

$$\mathrm{H}(\mathcal{C}|\mathcal{E}) = -\int \sum_{A\in\pi(\mathcal{C})} \Phi(\mathrm{E}_\mu(\chi_A|\mathcal{E}))\, d\mu\ .$$

Lemma 8.11 *Sei $\mathcal{C}$ eine endliche Unter-σ-Algebra von $\mathcal{F}$.*

(1) Es gilt $\mathrm{H}(\mathcal{C}|\mathcal{F}) = 0$.

(2) Für die triviale σ-Algebra $\mathcal{N} = \{\varnothing, X\}$ ist $\mathrm{H}(\mathcal{C}|\mathcal{N}) = \mathrm{H}(\mathcal{C})$.

Beweis (1) Es gilt $\Phi(0) = \Phi(1) = 0$ und für jedes $A \in \pi(\mathcal{C})$ ist $\mathrm{E}_\mu(\chi_A|\mathcal{F}) = \chi_A$ μ-fast sicher.

(2) Für jedes $A \in \pi(\mathcal{C})$ ist $\mathrm{E}_\mu(\chi_A|\mathcal{N}) = \mu(A)$ μ-fast sicher. □

Lemma 8.12 *Seien $\mathcal{A}$, $\mathcal{C}$ endliche Unter-σ-Algebren von $\mathcal{F}$. Dann gilt*

$$\mathrm{H}(\mathcal{A}|\mathcal{C}) = \sum_{B\in\pi(\mathcal{C})} \Big(-\sum_{A\in\pi(\mathcal{A})} \Phi(\mu(A|B))\Big)\mu(B)\ ,$$

wobei $\mu(A|B) = \mu(A\cap B)/\mu(B)$ *mit* $0/0 = 0$.

Beweis Übung. □

Satz 8.8 *Seien $\mathcal{C}$, $\mathcal{D}$, $\mathcal{E}$ Unter-σ-Algebren von $\mathcal{F}$ mit $\mathcal{C}$ und $\mathcal{D}$ endlich. Dann gilt*

$$\mathrm{H}(\mathcal{C}\vee\mathcal{D}|\mathcal{E}) = \mathrm{H}(\mathcal{C}|\mathcal{E}) + \mathrm{H}(\mathcal{D}|\mathcal{C}\vee\mathcal{E})\ .$$

Beweis In diesem Beweis soll $a/0$ stets als 0 interpretiert werden. Man beachte, dass $\Phi(a) = \Phi(b)(a/b) + \Phi(a/b)b$, falls a, $b \in \mathbb{R}^+$ mit $b \geq a$. Seien A, $B \in \mathcal{F}$; da $\mathrm{E}_\mu(\chi_A|\mathcal{E}) - \mathrm{E}_\mu(\chi_{A\cap B}|\mathcal{E}) = \mathrm{E}_\mu(\chi_{A\setminus B}|\mathcal{E}) \geq 0$ μ-fast sicher, ist

$$\Phi\Big(\frac{\mathrm{E}_\mu(\chi_{A\cap B}|\mathcal{E})}{\mathrm{E}_\mu(\chi_A|\mathcal{E})}\Big)\mathrm{E}_\mu(\chi_A|\mathcal{E}) + \Phi(\mathrm{E}_\mu(\chi_A|\mathcal{E}))\frac{\mathrm{E}_\mu(\chi_{A\cap B}|\mathcal{E})}{\mathrm{E}_\mu(\chi_A|\mathcal{E})} = \Phi(\mathrm{E}_\mu(\chi_{A\cap B}|\mathcal{E}))$$

μ-fast sicher. Da nun $\mathrm{E}_\mu(\chi_{A\cap B}|\mathcal{E}) = 0$ μ-fast sicher, wenn $A\cap B = \varnothing$, ist

$$\mathrm{H}(\mathcal{C}\vee\mathcal{D}|\mathcal{E}) = -\int \sum_{A\in\pi(\mathcal{C})} \sum_{B\in\pi(\mathcal{D})} \Phi(\mathrm{E}_\mu(\chi_{A\cap B}|\mathcal{E}))\, d\mu\ ,$$

und da $\sum_{B\in\pi(\mathcal{D})} \frac{\mathrm{E}_\mu(\chi_{A\cap B}|\mathcal{E})}{\mathrm{E}_\mu(\chi_A|\mathcal{E})} = 1$ μ-fast sicher auf $\{x \in X : \mathrm{E}_\mu(\chi_A|\mathcal{E})(x) > 0\}$, ist

$$\mathrm{H}(\mathcal{C}|\mathcal{E}) = -\int \sum_{A\in\pi(\mathcal{C})} \sum_{B\in\pi(\mathcal{D})} \Phi(\mathrm{E}_\mu(\chi_A|\mathcal{E})) \frac{\mathrm{E}_\mu(\chi_{A\cap B}|\mathcal{E})}{\mathrm{E}_\mu(\chi_A|\mathcal{E})}\, d\mu\, .$$

Für jedes $B \in \mathcal{F}$ setze

$$g_B = \sum_{A\in\pi(\mathcal{C})} \frac{\mathrm{E}_\mu(\chi_{A\cap B}|\mathcal{E})}{\mathrm{E}_\mu(\chi_A|\mathcal{E})}\chi_A\ ;$$

dann ist g_B $\mathcal{C}\vee\mathcal{E}$-messbar, und für jedes $A \in \pi(\mathcal{C})$, $E \in \mathcal{E}$ gilt

$$\begin{aligned}\int_{E\cap A} g_B\, d\mu &= \int_E \frac{\mathrm{E}_\mu(\chi_{A\cap B}|\mathcal{E})}{\mathrm{E}_\mu(\chi_A|\mathcal{E})}\chi_A\, d\mu \\ &= \int_E \frac{\mathrm{E}_\mu(\chi_{A\cap B}|\mathcal{E})}{\mathrm{E}_\mu(\chi_A|\mathcal{E})}\mathrm{E}_\mu(\chi_A|\mathcal{E})\, d\mu = \int_E \mathrm{E}_\mu(\chi_{A\cap B}|\mathcal{E})\, d\mu = \int_{E\cap A} \chi_B\, d\mu\ ,\end{aligned}$$

da $(\mathrm{E}_\mu(\chi_{A\cap B}|\mathcal{E})/\mathrm{E}_\mu(\chi_A|\mathcal{E}))\mathrm{E}_\mu(\chi_A|\mathcal{E})\, d\mu = \mathrm{E}_\mu(\chi_{A\cap B}|\mathcal{E})$ μ-fast sicher. Folglich gilt $\int_F g_B\, d\mu = \int_F \chi_B\, d\mu$ für alle $F \in \mathcal{C}\vee\mathcal{E}$, da jedes Element in $\mathcal{C}\vee\mathcal{E}$ eine Darstellung der Form $\bigcup_{A\in\pi(\mathcal{C})}(E_A \cap A)$ hat mit $E_A \in \mathcal{E}$ für jedes $A \in \pi(\mathcal{C})$. Dies zeigt, dass $g_B = \mathrm{E}_\mu(\chi_B|\mathcal{C}\vee\mathcal{E})$ μ-fast sicher, und folglich ist

$$\begin{aligned}\mathrm{H}(\mathcal{D}|\mathcal{C}\vee\mathcal{E}) &= -\int \sum_{B\in\pi(\mathcal{D})} \Phi(\mathrm{E}_\mu(\chi_B|\mathcal{C}\vee\mathcal{E})\, d\mu = -\int \sum_{B\in\pi(\mathcal{D})} \Phi(g_B)\, d\mu \\ &= -\int \sum_{B\in\pi(\mathcal{D})} \Phi\Big(\sum_{A\in\pi(\mathcal{C})} \frac{\mathrm{E}_\mu(\chi_{A\cap B}|\mathcal{E})}{\mathrm{E}_\mu(\chi_A|\mathcal{E})}\chi_A\Big)\, d\mu \\ &= -\int \sum_{A\in\pi(\mathcal{C})} \sum_{B\in\pi(\mathcal{D})} \Phi\Big(\frac{\mathrm{E}_\mu(\chi_{A\cap B}|\mathcal{E})}{\mathrm{E}_\mu(\chi_A|\mathcal{E})}\Big)\chi_A\, d\mu \\ &= -\int \sum_{A\in\pi(\mathcal{C})} \sum_{B\in\pi(\mathcal{D})} \Phi\Big(\frac{\mathrm{E}_\mu(\chi_{A\cap B}|\mathcal{E})}{\mathrm{E}_\mu(\chi_A|\mathcal{E})}\Big)\mathrm{E}_\mu(\chi_A|\mathcal{E})\, d\mu \\ &= -\int \sum_{A\in\pi(\mathcal{C})} \sum_{B\in\pi(\mathcal{D})} \Big(\Phi(\mathrm{E}_\mu(\chi_{A\cap B}|\mathcal{E})) - \Phi(\mathrm{E}_\mu(\chi_A|\mathcal{E}))\frac{\mathrm{E}_\mu(\chi_{A\cap B}|\mathcal{E})}{\mathrm{E}_\mu(\chi_A|\mathcal{E})}\Big)\, d\mu \\ &= \mathrm{H}(\mathcal{C}\vee\mathcal{D}|\mathcal{E}) - \mathrm{H}(\mathcal{C}|\mathcal{E})\ .\ \square\end{aligned}$$

Im nächsten Satz wird die folgende einfache Version der Jensenschen Ungleichung benötigt:

Lemma 8.13 *Sei $g : X \to \mathbb{R}^+$ beschränkt und $\mathcal{F}$-messbar. Dann gilt*

$$\Phi(\mathrm{E}_\mu(g|\mathcal{D})) \le \mathrm{E}_\mu(\Phi(g)|\mathcal{D}) \quad \mu\text{-fast sicher}$$

für jede Unter-σ-Algebra $\mathcal{D}$ von $\mathcal{F}$.

Beweis Für jedes $t > 0$ definiere $\tau_t : \mathbb{R} \to \mathbb{R}$ durch $\tau_t(s) = \Phi'(t)(s-t) + \Phi(t)$. Da Φ konvex ist, ist $\tau_t(s) \le \Phi(s)$ für alle $s \in \mathbb{R}^+$, und also gilt (μ-fast sicher)

$$\tau_t(\mathrm{E}_\mu(g|\mathcal{D})) = \mathrm{E}_\mu(\tau_t(g)|\mathcal{D}) = \mathrm{E}_\mu(\Phi(g)|\mathcal{D}) - \mathrm{E}_\mu(\Phi(g) - \tau_t(g)|\mathcal{D}) \le \mathrm{E}_\mu(\Phi(g)|\mathcal{D}) \;,$$

da $\Phi(g) - \tau_t(g) \ge 0$. Man sieht aber leicht, dass $\sup\{\tau_t(s) : t \in \mathbb{Q}^+\} = \Phi(s)$ für alle $s \in \mathbb{R}^+$, wobei $\mathbb{Q}^+ = \{t \in \mathbb{Q} : t > 0\}$, und folglich ist

$$\Phi(\mathrm{E}_\mu(g|\mathcal{D})) = \sup\{\tau_t(\mathrm{E}_\mu(g|\mathcal{D})) : t \in \mathbb{Q}^+\} \le \mathrm{E}_\mu(\Phi(g)|\mathcal{D}) \quad \mu\text{-fast sicher} \;. \quad \square$$

Satz 8.9 *Für Unter-σ-Algebren $\mathcal{C}$, $\mathcal{D}$, $\mathcal{E}$ von $\mathcal{F}$ mit $\mathcal{C}$ endlich und $\mathcal{D} \subset \mathcal{E}$ gilt*

$$\mathrm{H}(\mathcal{C}|\mathcal{D}) \ge \mathrm{H}(\mathcal{C}|\mathcal{E}) \;.$$

Beweis Für jedes $A \in \pi(\mathcal{C})$ sei $g_A = \mathrm{E}_\mu(\chi_A|\mathcal{E})$; dann ist $\mathrm{E}_\mu(\chi_A|\mathcal{D}) = \mathrm{E}_\mu(g_A|\mathcal{D})$ μ-fast sicher, da $\mathcal{D} \subset \mathcal{E}$, und damit gilt nach Lemma 8.13, dass

$$\Phi(\mathrm{E}_\mu(\chi_A|\mathcal{D})) = \Phi(\mathrm{E}_\mu(g_A|\mathcal{D})) \le \mathrm{E}_\mu(\Phi(g_A)|\mathcal{D}) = \mathrm{E}_\mu(\Phi(\mathrm{E}_\mu(\chi_A|\mathcal{E}))|\mathcal{D})$$

μ-fast sicher. Folglich ist

$$\int \Phi(\mathrm{E}_\mu(\chi_A|\mathcal{D}))\, d\mu \le \int \mathrm{E}_\mu(\Phi(\mathrm{E}_\mu(\chi_A|\mathcal{E}))|\mathcal{D})\, d\mu = \int \Phi(\mathrm{E}_\mu(\chi_A|\mathcal{E}))\, d\mu$$

und daraus ergibt sich, dass

$$\begin{aligned}\mathrm{H}(\mathcal{C}|\mathcal{D}) = -\int \sum_{A \in \pi(\mathcal{C})} \Phi(\mathrm{E}_\mu(\chi_A|\mathcal{D}))\, d\mu &\ge -\int \sum_{A \in \pi(\mathcal{C})} \Phi(\mathrm{E}_\mu(\chi_A|\mathcal{E}))\, d\mu \\ &= \mathrm{H}(\mathcal{C}|\mathcal{E}) \;. \quad \square\end{aligned}$$

Lemma 8.14 *Sei $\mathcal{E}$ eine Unter-σ-Algebra von $\mathcal{F}$. Dann gilt*

$$\mathrm{E}_\mu(f|\mathcal{E}) \circ T = \mathrm{E}_\mu(f \circ T|T^{-1}(\mathcal{E})) \quad \mu\text{-fast sicher}$$

für jedes $f \in \mathcal{L}^1(X, \mathcal{F}, \mu)$.

Beweis Die Abbildung $\mathrm{E}_\mu(f|\mathcal{E}) \circ T$ ist $T^{-1}(\mathcal{E})$-messbar und für jedes $A \in \mathcal{E}$ gilt nach Satz 1.2, da μ das Bildmaß von μ unter T ist, dass

$$\int_{T^{-1}(A)} \mathrm{E}_\mu(f|\mathcal{E}) \circ T\, d\mu = \int_A \mathrm{E}_\mu(f|\mathcal{E})\, d\mu = \int_A f\, d\mu = \int_{T^{-1}(A)} f \circ T\, d\mu \;.$$

Folglich ist $\mathrm{E}_\mu(f|\mathcal{E}) \circ T = \mathrm{E}_\mu(f \circ T|T^{-1}(\mathcal{E}))$ μ-fast sicher. $\square$

Satz 8.10 *Seien* $\mathcal{C}$, $\mathcal{E}$ *Unter-*σ*-Algebren von* $\mathcal{F}$ *mit* $\mathcal{C}$ *endlich. Dann gilt*

$$\mathrm{H}(T^{-1}(\mathcal{C})|T^{-1}(\mathcal{E})) = \mathrm{H}(\mathcal{C}|\mathcal{E}) \ .$$

Beweis Für jedes $B \in \mathcal{F}$ gilt nach Lemma 8.14 und Satz 1.2, dass

$$\begin{aligned}\int \Phi(\mathrm{E}_\mu(\chi_{T^{-1}(B)}|T^{-1}(\mathcal{E})))\,d\mu &= \int \Phi(\mathrm{E}_\mu(\chi_B|\mathcal{E}) \circ T)\,d\mu \\ &= \int \Phi(\mathrm{E}_\mu(\chi_B|\mathcal{E})) \circ T\,d\mu = \int \Phi(\mathrm{E}_\mu(\chi_B|\mathcal{E}))\,d\mu \ .\end{aligned}$$

Nach Lemma 8.10 ist aber $\pi(T^{-1}(\mathcal{C})) = T^{-1}[\pi(\mathcal{C})] = T^{-1}(\pi(\mathcal{C}))\setminus\{\varnothing\}$, und daraus ergibt sich, dass

$$\begin{aligned}\mathrm{H}(T^{-1}(\mathcal{C})|T^{-1}(\mathcal{E})) &= -\int \sum_{A\in T^{-1}(\pi(\mathcal{C}))\setminus\{\varnothing\}} \Phi(\mathrm{E}_\mu(\chi_A|T^{-1}(\mathcal{E})))\,d\mu \\ &= -\int \sum_{A\in T^{-1}(\pi(\mathcal{C}))} \Phi(\mathrm{E}_\mu(\chi_A|T^{-1}(\mathcal{E})))\,d\mu \\ &= -\int \sum_{B\in\pi(\mathcal{C})} \Phi(\mathrm{E}_\mu(\chi_{T^{-1}(B)}|T^{-1}(\mathcal{E})))\,d\mu \\ &= -\int \sum_{B\in\pi(\mathcal{C})} \Phi(\mathrm{E}_\mu(\chi_B|\mathcal{E}))\,d\mu = \mathrm{H}(\mathcal{C}|\mathcal{E}) \ . \quad \square\end{aligned}$$

Satz 8.11 *Für Unter-*σ*-Algebren* $\mathcal{A}$, $\mathcal{C}$, $\mathcal{D}$, $\mathcal{E}$ *von* $\mathcal{F}$ *mit* $\mathcal{A}$ *und* $\mathcal{C}$ *endlich gilt:*

(1) $\mathrm{H}(\mathcal{A}\vee\mathcal{C}|\mathcal{E}) = \mathrm{H}(\mathcal{A}|\mathcal{E}) + \mathrm{H}(\mathcal{C}|\mathcal{A}\vee\mathcal{E})$.

(2) $\mathrm{H}(\mathcal{C}|\mathcal{E}) \geq \mathrm{H}(\mathcal{C}|\mathcal{D})$, *falls* $\mathcal{E} \subset \mathcal{D}$.

(3) $\mathrm{H}(\mathcal{A}|\mathcal{E}) \leq \mathrm{H}(\mathcal{C}|\mathcal{E})$, *falls* $\mathcal{A} \subset \mathcal{C}$.

(4) $\mathrm{H}(\mathcal{A}\vee\mathcal{C}|\mathcal{E}) \leq \mathrm{H}(\mathcal{A}|\mathcal{E}) + \mathrm{H}(\mathcal{C}|\mathcal{E})$.

(5) $\mathrm{H}(T^{-1}(\mathcal{C})|T^{-1}(\mathcal{E})) = \mathrm{H}(\mathcal{C}|\mathcal{E})$.

Beweis (1) bzw. (2) bzw. (5) ist Satz 8.8 bzw. Satz 8.9 bzw. Satz 8.10.

(3) Da $\mathcal{A} \subset \mathcal{C}$, ist $\mathcal{A}\vee\mathcal{C} = \mathcal{C}$, und damit ist nach (1)

$$\mathrm{H}(\mathcal{C}|\mathcal{E}) = \mathrm{H}(\mathcal{A}\vee\mathcal{C}|\mathcal{E}) = \mathrm{H}(\mathcal{A}|\mathcal{E}) + \mathrm{H}(\mathcal{C}|\mathcal{A}\vee\mathcal{E}) \geq \mathrm{H}(\mathcal{A}|\mathcal{E}) \ .$$

(4) Dies folgt unmittelbar aus (1) und (2). $\square$

Satz 8.12 *Für Unter-σ-Algebren $\mathcal{A}$, $\mathcal{C}$, $\mathcal{D}$ von $\mathcal{F}$ mit $\mathcal{A}$ und $\mathcal{C}$ endlich gilt:*

(1) $\mathrm{H}(\mathcal{A} \vee \mathcal{C}) = \mathrm{H}(\mathcal{A}) + \mathrm{H}(\mathcal{C}|\mathcal{A})$.

(2) $\mathrm{H}(\mathcal{C}) \geq \mathrm{H}(\mathcal{C}|\mathcal{D})$.

(3) $\mathrm{H}(\mathcal{A}) \leq \mathrm{H}(\mathcal{C})$, *falls* $\mathcal{A} \subset \mathcal{C}$.

(4) $\mathrm{H}(\mathcal{A} \vee \mathcal{C}) \leq \mathrm{H}(\mathcal{A}) + \mathrm{H}(\mathcal{C})$.

(5) $\mathrm{H}(T^{-1}(\mathcal{C})) = \mathrm{H}(\mathcal{C})$.

Beweis Nach Lemma 8.11 (2) folgt dies aus Satz 8.11 mit $\mathcal{E} = \mathcal{N} = \{\varnothing, X\}$. □

Lemma 8.15 *Sei $\mathcal{C}$ eine endliche Unter-σ-Algebra von $\mathcal{F}$. Dann gilt*

$$\mathrm{H}\Big(\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{C})\Big) = \mathrm{H}(\mathcal{C}) + \sum_{k=1}^{n-1} \mathrm{H}\Big(\mathcal{C}| \bigvee_{\ell=1}^{k} T^{-\ell}(\mathcal{C})\Big)$$

für jedes $n \geq 1$.

Beweis Die Aussage ist trivial richtig, wenn $n = 1$. Sei $n \geq 1$ und nehme nun an, dass die Aussage für dieses n gilt. Nach Satz 8.11 (1) und (5) gilt dann

$$\begin{aligned}
\mathrm{H}\Big(\bigvee_{k=0}^{n} T^{-k}(\mathcal{C})\Big) &= \mathrm{H}\Big(\Big(\bigvee_{k=1}^{n} T^{-k}(\mathcal{C})\Big) \vee \mathcal{C}\Big) \\
&= \mathrm{H}\Big(\bigvee_{k=1}^{n} T^{-k}(\mathcal{C})\Big) + \mathrm{H}\Big(\mathcal{C}\Big|\bigvee_{k=1}^{n} T^{-k}(\mathcal{C})\Big) \\
&= \mathrm{H}\Big(T^{-1}\Big(\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{C})\Big)\Big) + \mathrm{H}\Big(\mathcal{C}\Big|\bigvee_{k=1}^{n} T^{-k}(\mathcal{C})\Big) \\
&= \mathrm{H}\Big(\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{C})\Big) + \mathrm{H}\Big(\mathcal{C}\Big|\bigvee_{k=1}^{n} T^{-k}(\mathcal{C})\Big) \\
&= \mathrm{H}(\mathcal{C}) + \sum_{k=1}^{n-1} \mathrm{H}\Big(\mathcal{C}\Big|\bigvee_{\ell=1}^{k} T^{-\ell}(\mathcal{C})\Big) + \mathrm{H}\Big(\mathcal{C}\Big|\bigvee_{k=1}^{n} T^{-k}(\mathcal{C})\Big) \\
&= \mathrm{H}(\mathcal{C}) + \sum_{k=1}^{n} \mathrm{H}\Big(\mathcal{C}\Big|\bigvee_{\ell=1}^{k} T^{-\ell}(\mathcal{C})\Big) .
\end{aligned}$$

Durch Induktion nach n ist also die Aussage richtig für jedes $n \geq 1$. □

Satz 8.13 *Für jede endliche Unter-σ-Algebra $\mathcal{C}$ von $\mathcal{F}$ ist die Folge*

$$\Big\{\frac{1}{n}\mathrm{H}\Big(\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{C})\Big)\Big\}_{n\geq 1}$$

monoton fallend.

Beweis Sei $n \geq 1$; nach Satz 8.11 (2) und Lemma 8.15 gilt

$$\mathrm{H}\Big(\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{C})\Big) \geq n\mathrm{H}\Big(\mathcal{C}\Big|\bigvee_{k=1}^{n} T^{-k}(\mathcal{C})\Big)$$

und daraus folgt nach Satz 8.12 (1) und (5), dass

$$\begin{aligned} n\mathrm{H}\Big(\bigvee_{k=0}^{n} T^{-k}(\mathcal{C})\Big) &= n\Big(\mathrm{H}\Big(\bigvee_{k=1}^{n} T^{-k}(\mathcal{C})\Big) + \mathrm{H}\Big(\mathcal{C}\Big|\bigvee_{k=1}^{n} T^{-k}(\mathcal{C})\Big)\Big) \\ &= n\Big(\mathrm{H}\Big(\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{C})\Big) + \mathrm{H}\Big(\mathcal{C}\Big|\bigvee_{k=1}^{n} T^{-k}(\mathcal{C})\Big)\Big) \leq (n+1)\mathrm{H}\Big(\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{C})\Big)\ , \end{aligned}$$

da $\bigvee_{k=1}^{n} T^{-k}(\mathcal{C}) = T^{-1}\big(\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{C})\big)$. $\square$

Die Abbildung T wird *invertierbar* genannt, wenn $T : X \to X$ eine bimessbare Bijektion ist. In diesem Fall ist T^{-1} auch maßtreu bezüglich μ.

Satz 8.14 *Für endliche Unter-σ-Algebren $\mathcal{A}$, $\mathcal{C}$ von $\mathcal{F}$ gilt:*

(1) $\mathrm{h}(T,\mathcal{C}) \leq \mathrm{H}(\mathcal{C})$.

(2) $\mathrm{h}(T,\mathcal{A}\vee\mathcal{C}) \leq \mathrm{h}(T,\mathcal{A}) + \mathrm{h}(T,\mathcal{C})$.

(3) $\mathrm{h}(T,\mathcal{A}) \leq \mathrm{h}(T,\mathcal{C})$, *falls* $\mathcal{A} \subset \mathcal{C}$.

(4) $\mathrm{h}(T,\mathcal{A}) \leq \mathrm{h}(T,\mathcal{C}) + \mathrm{H}(\mathcal{A}|\mathcal{C})$.

(5) $\mathrm{h}(T,T^{-1}(\mathcal{C})) = \mathrm{h}(T,\mathcal{C})$.

(6) $\mathrm{h}(T,\mathcal{C}) = \mathrm{h}(T,\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{C}))$ *für jedes* $n \geq 1$.

(7) $\mathrm{h}(T,\mathcal{C}) = \mathrm{h}(T,\bigvee_{k=-n}^{n} T^{-k}(\mathcal{C}))$ *für jedes* $n \geq 1$, *wenn* T *invertierbar ist.*

Beweis (1) Für jedes $n \geq 1$ gilt nach Satz 8.12 (4) und (5)

$$\frac{1}{n}\mathrm{H}\Big(\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{C})\Big) \leq \frac{1}{n}\sum_{k=0}^{n-1}\mathrm{H}(T^{-k}(\mathcal{C})) = \frac{1}{n}\sum_{k=0}^{n-1}\mathrm{H}(\mathcal{C}) = \mathrm{H}(\mathcal{C})$$

und folglich ist nach Satz 8.7 $\mathrm{h}(T,\mathcal{C}) \leq \mathrm{H}(\mathcal{C})$.

(2) Für jedes $n \geq 1$ gilt nach Satz 8.12 (4)

$$\begin{aligned}
\mathrm{H}\Big(\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{A} \vee \mathcal{C})\Big) &= \mathrm{H}\Big(\bigvee_{k=0}^{n-1} (T^{-k}(\mathcal{A}) \vee T^{-k}(\mathcal{C}))\Big) \\
&= \mathrm{H}\Big(\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{A}) \vee \bigvee_{k=0}^{n-1} T^{-k}(\mathcal{C})\Big) \leq \mathrm{H}\Big(\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{A})\Big) + \mathrm{H}\Big(\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{C})\Big)
\end{aligned}$$

und damit ist nach Satz 8.7 $\mathrm{h}(T, \mathcal{A} \vee \mathcal{C}) \leq \mathrm{h}(T, \mathcal{A}) + \mathrm{h}(T, \mathcal{C})$.

(3) Da $\mathcal{A} \subset \mathcal{C}$, ist $\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{A}) \subset \bigvee_{k=0}^{n-1} T^{-k}(\mathcal{C})$ und damit nach Satz 8.12 (3)

$$\mathrm{H}\Big(\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{A})\Big) \leq \mathrm{H}\Big(\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{C})\Big)$$

für jedes $n \geq 1$. Nach Satz 8.7 ist also $\mathrm{h}(T, \mathcal{A}) \leq \mathrm{h}(T, \mathcal{C})$.

(4) Für jedes $n \geq 1$ ist nach Satz 8.12 (1)

$$\begin{aligned}
\mathrm{H}\Big(\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{A} \vee \mathcal{C})\Big) &= \mathrm{H}\Big(\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{A}) \vee \bigvee_{k=0}^{n-1} T^{-k}(\mathcal{C})\Big) \\
&= \mathrm{H}\Big(\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{A})\Big) + \mathrm{H}\Big(\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{A}) \Big| \bigvee_{k=0}^{n-1} T^{-k}(\mathcal{C})\Big) .
\end{aligned}$$

Aber nach Satz 8.11 (4), (2) und (5) gilt

$$\begin{aligned}
\mathrm{H}\Big(\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{A}) \Big| \bigvee_{k=0}^{n-1} T^{-k}(\mathcal{C})\Big) &\leq \sum_{k=0}^{n-1} \mathrm{H}\Big(T^{-k}(\mathcal{A}) \Big| \bigvee_{\ell=0}^{n-1} T^{-\ell}(\mathcal{C})\Big) \\
&\leq \sum_{k=0}^{n-1} \mathrm{H}(T^{-k}(\mathcal{A}) | T^{-k}(\mathcal{C})) = \sum_{k=0}^{n-1} \mathrm{H}(\mathcal{A}|\mathcal{C}) = n\mathrm{H}(\mathcal{A}|\mathcal{C}) ,
\end{aligned}$$

und daraus ergibt sich, dass für jedes $n \geq 1$

$$\frac{1}{n}\mathrm{H}\Big(\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{A} \vee \mathcal{C})\Big) \leq \frac{1}{n}\mathrm{H}\Big(\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{A})\Big) + \mathrm{H}(\mathcal{A}|\mathcal{C}) .$$

Nach Satz 8.7 ist also $\mathrm{h}(T, \mathcal{A} \vee \mathcal{C}) \leq \mathrm{h}(T, \mathcal{C}) + \mathrm{H}(\mathcal{A}|\mathcal{C})$ und insbesondere ist dann nach (3) $\mathrm{h}(T, \mathcal{A}) \leq \mathrm{h}(T, \mathcal{C}) + \mathrm{H}(\mathcal{A}|\mathcal{C})$.

(5) Nach Satz 8.7 und Satz 8.12 (5) gilt

$$\begin{aligned}
\mathrm{h}(T, T^{-1}(\mathcal{C})) &= \lim_{n\to\infty} \frac{1}{n}\mathrm{H}\Big(\bigvee_{k=0}^{n-1} T^{-k}(T^{-1}(\mathcal{C}))\Big) \\
&= \lim_{n\to\infty} \frac{1}{n}\mathrm{H}\Big(T^{-1}\Big(\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{C})\Big)\Big) = \lim_{n\to\infty} \frac{1}{n}\mathrm{H}\Big(\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{C})\Big) = \mathrm{h}(T, \mathcal{C}) .
\end{aligned}$$

(6) Nach Satz 8.7 gilt

$$\begin{aligned}
&\mathrm{h}\Big(T,\bigvee_{k=0}^{n-1}T^{-k}(\mathcal{C})\Big)\\
&\qquad=\lim_{m\to\infty}\frac{1}{m}\mathrm{H}\Big(\bigvee_{\ell=0}^{m-1}T^{-\ell}\Big(\bigvee_{k=0}^{n-1}T^{-k}(\mathcal{C})\Big)\Big)=\lim_{m\to\infty}\frac{1}{m}\mathrm{H}\Big(\bigvee_{k=0}^{m+n-2}T^{-k}(\mathcal{C})\Big)\\
&\qquad=\lim_{m\to\infty}\frac{m+n-1}{m}\frac{1}{m+n-1}\mathrm{H}\Big(\bigvee_{k=0}^{m+n-2}T^{-k}(\mathcal{C})\Big)=\mathrm{h}(T,\mathcal{C})\ .
\end{aligned}$$

(7) Nach (5) und (6) gilt

$$\begin{aligned}
\mathrm{h}\Big(T,\bigvee_{k=-n}^{n}T^{-k}(\mathcal{C})\Big)&=\mathrm{h}\Big(T,T^{-n}\Big(\bigvee_{k=-n}^{n}T^{-k}(\mathcal{C})\Big)\Big)\\
&=\mathrm{h}\Big(T,\bigvee_{k=0}^{2n}T^{-k}(\mathcal{C})\Big)=\mathrm{h}(T,\mathcal{C})\ . \quad\square
\end{aligned}$$

Satz 8.15 *Es gilt* $\mathrm{h}(\mathrm{id}_X)=0$ *(mit* $\mathrm{id}_X : X\to X$ *der Identitätsabbildung).*

Beweis Ist $\mathcal{C}$ eine endliche Unter-σ-Algebra von $\mathcal{F}$, so ist

$$\bigvee_{k=0}^{n-1}\mathrm{id}_X^{-k}(\mathcal{C})=\bigvee_{k=0}^{n-1}\mathcal{C}=\mathcal{C}$$

für jedes $n\ge 1$ und damit ist nach Satz 8.7 $\mathrm{h}(\mathrm{id}_X,\mathcal{C})=\lim_{n\to\infty}n^{-1}\mathrm{H}(\mathcal{C})=0$. Daraus folgt nach Satz 8.6, dass $\mathrm{h}(\mathrm{id}_X)=0$. $\square$

Satz 8.16 *(1) Es gilt* $\mathrm{h}(T^m)=m\,\mathrm{h}(T)$ *für jedes* $m\in\mathbb{N}$.

(2) Ist T *invertierbar, so gilt* $\mathrm{h}(T^m)=|m|\,\mathrm{h}(T)$ *für jedes* $m\in\mathbb{Z}$.

Beweis Übung. $\square$

Satz 8.17 *Sei* T *periodisch, d.h., es gibt ein* $m\ge 1$*, so dass* $T^m=\mathrm{id}_X$*. Dann ist* $\mathrm{h}(T)=0$.

Beweis Dies folgt unmittelbar aus Satz 8.15 und Satz 8.16. $\square$

Ist $\{\mathcal{E}_\omega\}_{\omega\ge\Omega}$ eine Familie von Unter-σ-Algebren von $\mathcal{F}$, so wird die σ-Algebra $\sigma(\bigcup_{\omega\in\Omega}\mathcal{E}_\omega)$ mit $\bigvee_{\omega\in\Omega}\mathcal{E}_\omega$ bezeichnet.

Lemma 8.16 *Sei $\{\mathcal{E}_n\}_{n\geq 1}$ eine monoton wachsende Folge von Unter-σ-Algebren von $\mathcal{F}$ und setze $\mathcal{E} = \bigvee_{n\geq 1} \mathcal{E}_n$. Für jedes $A \in \mathcal{F}$ gilt dann*

$$\lim_{n\to\infty} \|\mathrm{E}_\mu(\chi_A|\mathcal{E}) - \mathrm{E}_\mu(\chi_A|\mathcal{E}_n)\|_2 = 0 \,.$$

Beweis Übung. □

Satz 8.18 *Sei $\mathcal{C}$ eine endliche Unter-σ-Algebra von $\mathcal{F}$ und $\{\mathcal{E}_n\}_{n\geq 1}$ eine monoton wachsende Folge von Unter-σ-Algebren von $\mathcal{F}$. Dann gilt*

$$\mathrm{H}\Big(\mathcal{C}\Big|\bigvee_{n\geq 1} \mathcal{E}_n\Big) = \lim_{n\to\infty} \mathrm{H}(\mathcal{C}|\mathcal{E}_n) \,.$$

Beweis Sei $\mathcal{E} = \bigvee_{n\geq 1} \mathcal{E}_n$; es genügt zu zeigen, dass für jedes $A \in \pi(\mathcal{C})$

$$\lim_{n\to\infty} \int \Phi(\mathrm{E}_\mu(\chi_A|\mathcal{E}_n))\, d\mu = \int \Phi(\mathrm{E}_\mu(\chi_A|\mathcal{E}))\, d\mu \,.$$

Sei $\varepsilon > 0$; da Φ gleichmäßig stetig auf dem kompakten Intervall $[0,1]$ ist, gibt es ein $\delta > 0$, so dass $|\Phi(s) - \Phi(t)| < \varepsilon/2$ für alle s, $t \in [0,1]$ mit $|s-t| < \delta$, und da $0 \leq \mathrm{E}_\mu(\chi_A|\mathcal{B}) \leq 1$ μ-fast sicher für jede Unter-σ-Algebra $\mathcal{B}$ von $\mathcal{F}$, ist dann

$$\begin{aligned}
&\Big|\int \Phi(\mathrm{E}_\mu(\chi_A|\mathcal{E}_n))\, d\mu - \int \Phi(\mathrm{E}_\mu(\chi_A|\mathcal{E}))\, d\mu\Big| \\
&\qquad \leq \int |\Phi(\mathrm{E}_\mu(\chi_A|\mathcal{E}_n)) - \Phi(\mathrm{E}_\mu(\chi_A|\mathcal{E}))|\, d\mu \\
&\qquad \leq \varepsilon/2 + \mu(\{x \in X : |\Phi(\mathrm{E}_\mu(\chi_A|\mathcal{E}_n))(x) - \Phi(\mathrm{E}_\mu(\chi_A|\mathcal{E}))(x)| \geq \delta\}) \\
&\qquad \leq \varepsilon/2 + \delta^{-2}\|\Phi(\mathrm{E}_\mu(\chi_A|\mathcal{E}_n)) - \Phi(\mathrm{E}_\mu(\chi_A|\mathcal{E}))\|_2^2
\end{aligned}$$

für jedes $n \geq 1$, da $|\Phi(s)| \leq \mathrm{e}^{-1} \leq 1$ für alle $s \in [0,1]$. Nach Lemma 8.16 gibt es nun ein $N \geq 1$, so dass $\|\Phi(\mathrm{E}_\mu(\chi_A|\mathcal{E}_n)) - \Phi(\mathrm{E}_\mu(\chi_A|\mathcal{E}))\|_2 < \delta\sqrt{\varepsilon/2}$ für alle $n \geq N$, und also ist $|\int \Phi(\mathrm{E}_\mu(\chi_A|\mathcal{E}_n))\, d\mu - \int \Phi(\mathrm{E}_\mu(\chi_A|\mathcal{E}))\, d\mu| < \varepsilon$ für alle $n \geq N$. □

Satz 8.19 *Für jede endliche Unter-σ-Algebra $\mathcal{C}$ von $\mathcal{F}$ gilt*

$$\mathrm{h}(T,\mathcal{C}) = \mathrm{H}\Big(\mathcal{C}\Big|\bigvee_{n\geq 1} T^{-n}(\mathcal{C})\Big) \,.$$

Beweis Nach Satz 8.7 und Lemma 8.15 ist

$$\mathrm{h}(T,\mathcal{C}) = \lim_{n\to\infty} \frac{1}{n}\mathrm{H}\Big(\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{C})\Big) = \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n-1} \mathrm{H}\Big(\mathcal{C}|\bigvee_{\ell=1}^{k} T^{-\ell}(\mathcal{C})\Big)$$

und nach Satz 8.17 ist

$$\mathrm{H}\Big(\mathcal{C}\Big|\bigvee_{n\geq 1} T^{-n}(\mathcal{C})\Big) = \lim_{n\to\infty} \mathrm{H}\Big(\mathcal{C}\Big|\bigvee_{k=1}^{n} T^{-k}(\mathcal{C})\Big) = \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n-1} \mathrm{H}\Big(\mathcal{C}|\bigvee_{\ell=1}^{k} T^{-\ell}(\mathcal{C})\Big)\ ,$$

da $\lim\limits_{n\to\infty} n^{-1}\sum_{k=1}^{n-1} a_k = a$ für jede reelle Folge $\{a_n\}_{n\geq 1}$, die gegen a konvergiert. □

Für A, $B \in \mathcal{F}$ schreibt man $A \subset_\mu B$, wenn $\mu(A \setminus B) = 0$, und $A =_\mu B$, wenn $A \subset_\mu B$ und $B \subset_\mu A$, und also gilt $A =_\mu B$ genau dann, wenn $\mu(A \triangle B) = 0$. Sind $\mathcal{D}$ und $\mathcal{E}$ Unter-σ-Algebren von $\mathcal{F}$, so bedeutet $\mathcal{D} \subset_\mu \mathcal{E}$, dass es zu jedem $A \in \mathcal{D}$ ein $B \in \mathcal{E}$ mit $A =_\mu B$ gibt. Ferner bedeutet $\mathcal{D} =_\mu \mathcal{E}$, dass $\mathcal{D} \subset_\mu \mathcal{E}$ und $\mathcal{E} \subset_\mu \mathcal{D}$. Gilt $\mathcal{D} =_\mu \mathcal{E}$, so sieht man leicht, dass auch $T^{-1}(\mathcal{D}) =_\mu T^{-1}(\mathcal{E})$.

Lemma 8.17 *Seien $\mathcal{D}$, $\mathcal{E}$ Unter-σ-Algebren von $\mathcal{F}$ mit $\mathcal{D} =_\mu \mathcal{E}$ und sei $\mathcal{C}$ eine endliche Unter-σ-Algebra von $\mathcal{F}$.*

(1) Dann gilt $\mathrm{H}(\mathcal{C}|\mathcal{D}) = \mathrm{H}(\mathcal{C}|\mathcal{E})$.

(2) Es gilt $\mathrm{H}(\mathcal{C}|\mathcal{E}) = 0$ *genau dann, wenn* $\mathcal{C} \subset_\mu \mathcal{E}$.

Beweis Übung. □

Satz 8.20 *Für jede endliche Unter-σ-Algebra $\mathcal{C}$ von $\mathcal{F}$ gilt* $\mathrm{h}(T, \mathcal{C}) = 0$ *genau dann, wenn* $\mathcal{C} \subset_\mu \bigvee_{n\geq 1} T^{-n}(\mathcal{C})$.

Beweis Dies folgt unmittelbar aus Satz 8.19 und Lemma 8.17 (2). □

Satz 8.21 *Es gilt* $T^{-1}(\mathcal{F}) =_\mu \mathcal{F}$, *wenn* $\mathrm{h}(T) = 0$.

Beweis Da $T^{-1}(\mathcal{F}) \subset \mathcal{F}$, genügt es zu zeigen, dass $\mathcal{F} \subset_\mu T^{-1}(\mathcal{F})$. Sei $B \in \mathcal{F}$; dann ist $\mathcal{C} = \{\varnothing, X, B, X \setminus B\}$ eine endliche Unter-σ-Algebra von $\mathcal{F}$ und es gilt $\mathrm{h}(T, \mathcal{C}) \leq \mathrm{h}(T) = 0$. Daraus ergibt sich nach Satz 8.20, dass

$$\mathcal{C} \subset_\mu \bigvee_{n\geq 1} T^{-n}(\mathcal{C}) = T^{-1}\Big(\bigvee_{n\geq 0} T^{-n}(\mathcal{C})\Big) \subset T^{-1}(\mathcal{F})$$

und damit ist $\mathcal{F} \subset_\mu T^{-1}(\mathcal{F})$. □

Satz 8.22 (Kolmogorov-Sinai) *Sei $\mathcal{C}$ eine endliche Unter-σ-Algebra von $\mathcal{F}$.*

(1) Gilt $\bigvee_{n\geq 0} T^{-n}(\mathcal{C}) =_\mu \mathcal{F}$, *so ist* $\mathrm{h}(T) = \mathrm{h}(T, \mathcal{C})$.

(2) Ist T invertierbar und gilt $\bigvee_{n\in\mathbb{Z}} T^{-n}(\mathcal{C}) =_\mu \mathcal{F}$, *so ist* $\mathrm{h}(T) = \mathrm{h}(T, \mathcal{C})$.

Beweis (1) Sei $\mathcal{A}$ eine endliche Unter-σ-Algebra von $\mathcal{F}$. Nach Satz 8.14 (4) und (6) gilt dann für jedes $n \geq 1$, dass

$$\begin{aligned}\mathrm{h}(T,\mathcal{A}) &\leq \mathrm{h}\Big(T, \bigvee_{k=0}^{n-1} T^{-k}(\mathcal{C})\Big) + \mathrm{H}\Big(\mathcal{A}\Big|\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{C})\Big) \\ &= \mathrm{h}(T,\mathcal{C}) + \mathrm{H}\Big(\mathcal{A}\Big|\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{C})\Big) .\end{aligned}$$

Da aber $\mathcal{A} \subset_\mu \bigvee_{n\geq 0} T^{-n}(\mathcal{C})$, ist nach Satz 8.18 und Lemma 8.17 (2)

$$\lim_{n\to\infty} \mathrm{H}\Big(\mathcal{A}\Big|\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{C})\Big) = \mathrm{H}\Big(\mathcal{A}\Big|\bigvee_{n\geq 0} T^{-n}(\mathcal{C})\Big) = 0 .$$

Dies zeigt, dass $\mathrm{h}(T,\mathcal{A}) \leq \mathrm{h}(T,\mathcal{C})$ für jede endliche Unter-σ-Algebra von $\mathcal{F}$ und damit ist nach Satz 8.6 $\mathrm{h}(T) = \mathrm{h}(T,\mathcal{C})$.

(2) Genauso wie (1), aber unter Verwendung von Satz 8.14 (7) statt Satz 8.14 (6) und mit $\bigvee_{k=-n}^{n} T^{-k}(\mathcal{C})$ statt $\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{C})$. □

Satz 8.23 *Sei T invertierbar und nehme an, dass $\bigvee_{n\geq 0} T^{-n}(\mathcal{C}) =_\mu \mathcal{F}$ für eine endliche Unter-σ-Algebra $\mathcal{C}$ von $\mathcal{F}$. Dann ist $\mathrm{h}(T) = 0$.*

Beweis Da T invertierbar ist, gilt $T^{-1}(\mathcal{F}) = \mathcal{F}$ und daraus ergibt sich nach Satz 8.22 (1), Satz 8.19, Lemma 8.17 (1) und Lemma 8.11 (1), dass

$$\begin{aligned}\mathrm{h}(T) = \mathrm{h}(T,\mathcal{C}) = \mathrm{H}\Big(\mathcal{C}\Big|\bigvee_{n\geq 1} T^{-n}(\mathcal{C})\Big) &= \mathrm{H}\Big(\mathcal{C}\Big|T^{-1}\Big(\bigvee_{n\geq 0} T^{-n}(\mathcal{C})\Big)\Big) \\ &= \mathrm{H}(\mathcal{C}|T^{-1}(\mathcal{F})) = \mathrm{H}(\mathcal{C}|\mathcal{F}) = 0 .\end{aligned}$$ □

Lemma 8.18 *Sei $\mathcal{A}$ eine Unter-σ-Algebra von $\mathcal{F}$ und $0 < \delta < \mathrm{e}^{-1}$. Dann gilt*

$$-\int \Phi(\mathrm{E}_\mu(\chi_B|\mathcal{A}))\,d\mu \leq (2/\delta)\mu(B \triangle A) + \max\{-\Phi(\delta), -\Phi(1-\delta)\}$$

für alle $A \in \mathcal{A}$, $B \in \mathcal{F}$.

Beweis Seien $A \in \mathcal{A}$, $B \in \mathcal{F}$, und setze $\eta = \mu(B \triangle A)$. Sei ferner

$$\begin{aligned}D_\delta^+ &= \{x \in X : \mathrm{E}_\mu(\chi_B - \chi_A|\mathcal{A})(x) > \delta\} , \\ D_\delta^- &= \{x \in X : \mathrm{E}_\mu(\chi_B - \chi_A|\mathcal{A})(x) < -\delta\}\end{aligned}$$

und $D_\delta = D_\delta^+ \cup D_\delta^-$. Dann gilt

$$\delta\mu(D_\delta^+) \leq \int_{D_\delta^+} \mathrm{E}_\mu(\chi_B - \chi_A|\mathcal{A})\, d\mu = \int_{D_\delta^+} (\chi_B - \chi_A)\, d\mu \leq \int |\chi_B - \chi_A|\, d\mu = \eta$$

und genauso gilt $\delta\mu(D_\delta^-) \leq \eta$. Daraus ergibt sich, dass

$$\begin{aligned}
-\int \Phi(\mathrm{E}_\mu(\chi_B|\mathcal{A}))\, d\mu &= -\int \Phi(\mathrm{E}_\mu(\chi_B - \chi_A|\mathcal{A}) + \chi_A)\, d\mu \\
&= -\int_{D_\delta} \Phi(\mathrm{E}_\mu(\chi_B - \chi_A|\mathcal{A}) + \chi_A)\, d\mu - \int_{X\setminus D_\delta} \Phi(\mathrm{E}_\mu(\chi_B - \chi_A|\mathcal{A}) + \chi_A)\, d\mu \\
&\leq \mu(D_\delta) + \max\{-\Phi(\delta), -\Phi(1-\delta)\} \\
&\leq (2/\delta)\mu(B \triangle A) + \max\{-\Phi(\delta), -\Phi(1-\delta)\}\ . \quad \square
\end{aligned}$$

Lemma 8.19 *Sei $\mathcal{G}$ eine Algebra mit $\mathcal{G} \subset \mathcal{F}$ und $\sigma(\mathcal{G}) =_\mu \mathcal{F}$ und sei $\mathcal{C}$ eine endliche Unter-σ-Algebra von $\mathcal{F}$. Zu jedem $\varepsilon > 0$ gibt es dann eine endliche Unter-σ-Algebra $\mathcal{A}$ von $\mathcal{F}$ mit $\mathcal{A} \subset \mathcal{G}$ und* $\mathrm{H}(\mathcal{C}|\mathcal{A}) < \varepsilon$.

Beweis Sei $\varepsilon > 0$ und wähle $0 < \delta < \mathrm{e}^{-1}$ mit $\max\{-\Phi(\delta), -\Phi(1-\delta)\} < \varepsilon/(2m)$, wobei $m = 2^{|\pi(\mathcal{C})|}$. Da $\sigma(\mathcal{G}) =_\mu \mathcal{F}$, gibt es zu jedem $C \in \pi(\mathcal{C})$ ein $A_C \in \mathcal{G}$ mit $\mu(C \triangle A_C) < \delta\varepsilon/(4m)$. Setze $\mathcal{A} = \sigma(\{A_C : C \in \pi(\mathcal{C})\})$; also ist $\mathcal{A}$ eine endliche Unter-σ-Algebra von $\mathcal{F}$ mit $\mathcal{A} \subset \mathcal{G}$ und man sieht leicht, dass $|\pi(\mathcal{A})| \leq m$. Sei $C \in \mathcal{C}$; da $A_C \in \mathcal{A}$, folgt aus Lemma 8.18, dass

$$\begin{aligned}
-\int \Phi(\mathrm{E}_\mu(\chi_C|\mathcal{A}))\, d\mu &\leq (2/\delta)\mu(C \triangle A_C) + \max\{-\Phi(\delta), -\Phi(1-\delta)\} \\
&< (2/\delta)\delta\varepsilon/(4m) + \varepsilon/(2m) = \varepsilon/m\ .
\end{aligned}$$

Daraus ergibt sich, dass

$$\mathrm{H}(\mathcal{C}|\mathcal{A}) = -\int \sum_{C\in\pi(\mathcal{C})} \Phi(\mathrm{E}_\mu(\chi_C)|\mathcal{A})\, d\mu < |\pi(\mathcal{C})|(\varepsilon/m) = \varepsilon\ . \quad \square$$

Satz 8.24 *Sei $\mathcal{G}$ eine Algebra mit $\mathcal{G} \subset \mathcal{F}$ und $\sigma(\mathcal{G}) =_\mu \mathcal{F}$. Dann gilt*

$$\mathrm{h}(T) = \sup\{\mathrm{h}(T,\mathcal{A}) : \mathcal{A} \textit{ eine endliche Unter-}\sigma\textit{-Algebra von } \mathcal{F} \textit{ mit } \mathcal{A} \subset \mathcal{G}\}\ .$$

Beweis Sei $\varepsilon > 0$; nach Satz 8.6 gibt es eine endliche Unter-σ-Algebra $\mathcal{C}$ von $\mathcal{F}$ mit $\mathrm{h}(T,\mathcal{C}) > \mathrm{h}(T) - \varepsilon/2$ und nach Lemma 18.19 gibt es dann eine endliche Unter-σ-Algebra $\mathcal{C}$ von $\mathcal{F}$ mit $\mathcal{A} \subset \mathcal{G}$ und $\mathrm{H}(\mathcal{C}|\mathcal{A}) < \varepsilon/2$. Nach Satz 8.14 (4) ist also $\mathrm{h}(T,\mathcal{A}) \geq \mathrm{h}(T,\mathcal{C}) - \mathrm{H}(\mathcal{C}|\mathcal{A}) > \mathrm{h}(T) - \varepsilon$. $\square$

Satz 8.25 *Für $k = 1, 2$ sei $(X_k, \mathcal{F}_k, \mu_k)$ ein Wahrscheinlichkeitsraum und sei $T_k : X_k \to X_k$ eine messbare Abbildung, die maßtreu bezüglich μ_k ist. Dann gilt*

$$\mathrm{h}(T_1 \times T_2) = \mathrm{h}(T_1) + \mathrm{h}(T_2) .$$

Beweis Sei $\mathcal{G}$ die Menge aller Teilmengen von $X_1 \times X_2$, die eine Darstellung der Form $\bigcup_{k=1}^n A_k^1 \times A_k^2$ mit $A_k^1 \in \mathcal{F}_1$ und $A_k^2 \in \mathcal{F}_2$ haben. Dann ist $\mathcal{G}$ eine Algebra und (per Definition) gilt $\sigma(G) = \mathcal{F}_1 \times \mathcal{F}_2$. Sei $\mathcal{A}$ eine endliche Unter-σ-Algebra von $\mathcal{F}_1 \times \mathcal{F}_2$ mit $\mathcal{A} \subset \mathcal{G}$. Dann sieht man leicht, dass es für $k = 1, 2$ eine endliche Unter-σ-Algebra $\mathcal{C}_k$ von $\mathcal{F}_k$ gibt mit $\mathcal{A} \subset \mathcal{C}_1 \times \mathcal{C}_2$. Aber $\mathcal{C}_1 \times \mathcal{C}_2$ ist eine endliche Unter-σ-Algebra von $\mathcal{F}_1 \times \mathcal{F}_2$ und folglich ist nach Satz 8.14 (3) und Satz 8.24

$$\begin{aligned}\mathrm{h}(T_1 \times T_2) = \sup\{&\mathrm{h}(T_1 \times T_2, \mathcal{C}_1 \times \mathcal{C}_2) : \\ &\mathcal{C}_k \text{ eine endliche Unter-}\sigma\text{-Algebra von } \mathcal{F}_k \text{ für } k = 1, 2\} .\end{aligned}$$

Es genügt also zu zeigen: Ist $\mathcal{C}_k$ eine endliche Unter-σ-Algebra von $\mathcal{F}_k$ für $k = 1, 2$, so gilt $\mathrm{h}(T_1 \times T_2, \mathcal{C}_1 \times \mathcal{C}_2) = \mathrm{h}(T_1, \mathcal{C}_1) + \mathrm{h}(T_2, \mathcal{C}_2)$. Aber für jedes $n \geq 1$ gilt

$$\bigvee_{k=0}^{n-1} (T_1 \times T_2)^{-k}(\mathcal{C}_1 \times \mathcal{C}_2) = \bigvee_{k=0}^{n-1} T_1^{-k}(\mathcal{C}_1) \times \bigvee_{k=0}^{n-1} T_2^{-k}(\mathcal{C}_2) ,$$

da $(T_1 \times T_2)^{-k}(\mathcal{C}_1 \times \mathcal{C}_2) = T_1^{-k}(\mathcal{C}_1) \times T_2^{-k}(\mathcal{C}_2)$ für jedes $k \geq 0$. Nach Satz 8.7 genügt es daher zu zeigen, dass $\mathrm{H}(\mathcal{C}_1 \times \mathcal{C}_2) = \mathrm{H}(\mathcal{C}_1) + \mathrm{H}(\mathcal{C}_2)$. Nun ist

$$\pi(\mathcal{C}_1 \times \mathcal{C}_2) = \{A \times B : A \in \pi(\mathcal{C}_1),\ B \in \pi(\mathcal{C}_2)\}$$

und daraus ergibt sich, da $\Phi(st) = s\Phi(t) + t\Phi(s)$ für alle $s, t \in \mathbb{R}^+$, dass

$$\begin{aligned}\mathrm{H}(\mathcal{C}_1 \times \mathcal{C}_2) &= -\sum_{A\in\pi(\mathcal{C}_1)} \sum_{B\in\pi(\mathcal{C}_2)} \Phi((\mu_1 \times \mu_2)(A \times B)) \\ &= -\sum_{A\in\pi(\mathcal{C}_1)} \sum_{B\in\pi(\mathcal{C}_2)} \Phi(\mu_1(A)\mu_2(B)) \\ &= -\sum_{A\in\pi(\mathcal{C}_1)} \sum_{B\in\pi(\mathcal{C}_2)} (\mu_1(A)\Phi(\mu_2(B)) + \mu_2(B)\Phi(\mu_1(A))) \\ &= -\sum_{A\in\pi(\mathcal{C}_1)} \Phi(\mu_1(A)) - \sum_{B\in\pi(\mathcal{C}_2)} \Phi(\mu_2(B)) = \mathrm{H}(\mathcal{C}_1) + \mathrm{H}(\mathcal{C}_2) . \quad \square\end{aligned}$$

Beispiele:

1. Sei $a \in \mathrm{K} = \{z \in \mathbb{C} : |z| = 1\}$ und definiere $T : \mathrm{K} \to \mathrm{K}$ durch $T(z) = az$ für alle $z \in \mathrm{K}$; T ist maßtreu bezüglich m_{K}. Es gilt $\mathrm{h}(T) = 0$.

Beweis Gilt $a^p = 1$ für ein $p \geq 1$, so ist $T^p = \mathrm{id}_{\mathrm{K}}$ und hier ist nach Satz 8.17 $\mathrm{h}(T) = 0$. Nehme also an, dass $a^p \neq 1$ für alle $p \geq 1$. Nach Übungsaufgabe 4 liegt in diesem Fall die Folge $\{a^n\}_{n\geq 0}$ dicht in K. Sei $\mathcal{C} = \{\varnothing, J, \mathrm{K} \setminus J, \mathrm{K}\}$, wobei $J = \{\exp(2\pi i\theta) : 0 \leq \theta < \frac{1}{2}\}$. Dann ist $a^{-m}J = T^{-m}(J) \in \bigvee_{n\geq 0} T^{-n}(\mathcal{C})$ für jedes $m \geq 0$ und da die Folge $\{a^{-n}\}_{n\geq 0}$ auch dicht in K liegt, folgt daraus, dass $\bigvee_{n\geq 0} T^{-n}(\mathcal{C}) = \mathcal{B}_{\mathrm{K}}$. Nach Satz 18.23 ist $\mathrm{h}(T) = 0$, da T invertierbar ist. □

2. Sei $n \geq 1$ und sei $a \in \mathrm{K}^n$; definiere $T : \mathrm{K}^n \to \mathrm{K}^n$ durch $T(z) = az$ für alle $z \in \mathrm{K}^n$; T ist maßtreu bezüglich $\mathrm{m}_{\mathrm{K}^n}$. Es gilt $\mathrm{h}(T) = 0$.

Beweis Ist $a = (a_1, \ldots, a_n)$, so ist $T = T_1 \times \cdots \times T_n$, wobei $T_k : \mathrm{K} \to \mathrm{K}$ durch $T_k(z) = a_k z$ definiert ist. Nach 1. ist $\mathrm{h}(T_k) = 0$ für jedes $k = 1, \ldots, n$ und nach Satz 8.25 ist $\mathrm{h}(T) = \mathrm{h}(T_1) + \cdots + \mathrm{h}(T_n)$. Damit ist $\mathrm{h}(T) = 0$. □

3. Sei E eine endliche Menge und $Q : E \times E \to \mathbb{R}^+$ eine stochastische Matrix. Sei I entweder $\mathbb{Z}$ oder $\mathbb{N}$, setze $E_\infty = \mathrm{Abb}(I, E)$ und sei $\mathcal{E}_\infty = \sigma\big(\bigcup_{n\in I} \mathrm{p}_m^{-1}(\mathcal{E})\big)$, wobei $\mathcal{E} = \mathcal{P}(E)$ und $\mathrm{p}_m : E_\infty \to E$ die durch $\mathrm{p}_m(\{x_n\}_{n\in I}) = x_m$ gegebene Projektionsabbildung ist. Ferner sei $S : E_\infty \to E_\infty$ definiert durch $S(\{x_n\}_{n\in I}) = \{x'_n\}_{n\in I}$, wobei $x'_n = x_{n+1}$. Sei $\alpha \in \mathrm{St}(Q)$; nach Satz 5.6 gibt es dann ein eindeutiges $\nu \in \mathrm{P}(E_\infty, \mathcal{E}_\infty)$, so dass für alle $k, \ell \in I$ mit $k \leq \ell$ und alle $B_k, \ldots, B_\ell \subset E$

$$\nu\Big(\bigcap_{n=k}^{\ell} \mathrm{p}_n^{-1}(B_n)\Big) = \sum_{u_k\in B_k} \cdots \sum_{u_\ell\in B_\ell} \alpha(u_k)Q(u_k, u_{k+1}) \cdots Q(u_{\ell-1}, u_\ell) ,$$

und S ist maßtreu bezüglich ν. Hier ist

$$\mathrm{h}(S) = -\sum_{u\in E}\sum_{v\in E} \alpha(u)\Phi(Q(u, v)) .$$

Beweis Für jedes $n \in I$ ist $S^{-n}(\mathrm{p}_0^{-1}(\mathcal{E})) = \mathrm{p}_n^{-1}(\mathcal{E})$ und folglich gilt

$$\mathcal{E}_\infty = \sigma\Big(\bigcup_{n\in I} \mathrm{p}_n^{-1}(\mathcal{E})\Big) = \bigvee_{n\in I} S^{-n}(\mathcal{C}) ,$$

wobei $\mathcal{C} = \mathrm{p}_0^{-1}(\mathcal{E})$. Daraus folgt nach Satz 8.22, dass $\mathrm{h}(S) = \mathrm{h}(S, \mathcal{C})$. (Man beachte, dass S invertierbar ist, wenn $I = \mathbb{Z}$.) Nach Übungsaufgabe 32 ist aber

$$\mathrm{h}(S, \mathcal{C}) = -\sum_{u\in E}\sum_{v\in E} \alpha(u)\Phi(Q(u, v)) . \ \square$$

4. Sei E eine endliche Menge und sei $\alpha : E \to \mathbb{R}^+$ ein Wahrscheinlichkeitsvektor. Sei $(E_\infty, \mathcal{E}_\infty)$ und $S : E_\infty \to E_\infty$ wie in Beispiel 3 (mit I entweder $\mathbb{N}$ oder $\mathbb{Z}$).

Dann gibt es ein eindeutiges Produktmaß $\nu \in \mathrm{P}(E_\infty, \mathcal{E}_\infty)$, so dass für alle k, $\ell \in I$ mit $k \leq \ell$ und alle $B_k, \ldots, B_\ell \subset E$

$$\nu\Big(\bigcap_{n=k}^{\ell} \mathrm{p}_n^{-1}(B_n)\Big) = \prod_{n=k}^{\ell} \sum_{u_n \in B_n} \alpha(u_n) ,$$

und S ist maßtreu bezüglich ν. Dies ist der Spezialfall von 3. mit $Q(u, v) = \alpha(v)$ für alle u, $v \in E$ und daraus folgt, dass hier

$$\mathrm{h}(S) = -\sum_{v \in E} \Phi(\alpha(v)) .$$

5. Definiere $U : \mathrm{K} \to \mathrm{K}$ durch $U(z) = z^2$ und $T' : [0,1] \to [0,1]$ durch

$$T'(x) = \begin{cases} 2x & \text{falls } 0 \leq x < \frac{1}{2} , \\ 2 - 2x & \text{falls } \frac{1}{2} \leq x \leq 1 . \end{cases}$$

Dann ist U maßtreu bezüglich m_{K} und T' maßtreu bezüglich der Einschränkung des Lebesgueschen Maßes auf $[0,1]$. Ferner ist in Kapitel 7 gezeigt worden, dass U, T' und S isomorph sind, wobei $S : E_\infty \to E_\infty$ der Spezialfall in 4. mit $E = \{0,1\}$, $\alpha(0) = \alpha(1) = \frac{1}{2}$ und $I = \mathbb{N}$. Nach Satz 8.3 gilt also

$$\mathrm{h}(U) = \mathrm{h}(T') = \mathrm{h}(S) = \log 2 ,$$

da $\mathrm{h}(S) = -(\Phi(\frac{1}{2}) + \Phi(\frac{1}{2})) = \log 2$.

6. Sei $p \in \mathbb{Z} \setminus \{0\}$ und definiere $U : \mathrm{K} \to \mathrm{K}$ durch $U(z) = z^p$ für alle $z \in \mathrm{K}$; dann ist U maßtreu bezüglich m_{K}. Es gilt $\mathrm{h}(U) = \log |p|$.

Beweis Da U invertierbar ist mit $U^{-1}(z) = z^{-p}$ für alle $z \in \mathrm{K}$ und da nach Satz 8.16 (2) $\mathrm{h}(U^{-1}) = \mathrm{h}(U)$, kann man annehmen, dass $p > 0$. Betrachte den Spezialfall in 4. mit $E = \{1, 2, \ldots, p\}$, $\alpha(k) = 1/p$ für jedes k und $I = \mathbb{N}$. Dann ist $\mathrm{h}(S) = \log p$. Aber genauso wie im Fall $p = 2$ kann man zeigen, dass S und U isomorph sind und folglich ist nach Satz 8.3 $\mathrm{h}(U) = \mathrm{h}(S) = \log p$. $\square$

7. Sei $A \in \mathrm{M}(n \times n, \mathbb{Z})$ mit $\det(A) \neq 0$ und sei $\theta_A : \mathrm{K}^n \to \mathrm{K}^n$ der entsprechende surjektive Endomorphismus von K^n; also ist θ_A maßtreu bezüglich $\mathrm{m}_{\mathrm{K}^n}$. Seien $\lambda_1, \ldots, \lambda_n$ die Eigenwerte von A; dann gilt

$$\mathrm{h}(\theta_A) = \sum_{k=1}^{n} \log |\lambda_k| .$$

Beweis Dies wird nur für einen sehr einfachen Fall bewiesen: Nehme an, dass es eine Matrix $P \in \mathrm{M}(n \times n, \mathbb{Z})$ mit $|\det(P)| = 1$ gibt, so dass $D = P^{-1}AP$ eine Diagonalmatrix ist. Dann gilt $\theta_D = (\theta_P)^{-1} \circ \theta_A \circ \theta_P$ und daraus folgt, dass

θ_A und θ_D strikt isomorph sind. Nach Satz 8.3 ist also $\mathrm{h}(\theta_A) = \mathrm{h}(\theta_D)$. Aber $\theta_D = T_1 \times \cdots \times T_n$, wobei $T_k : \mathrm{K} \to \mathrm{K}$ gegeben ist durch $T_k(z) = z^{\eta_k}$ und wobei $\eta_1, \ldots, \eta_n$ die Diagonaleinträge von D sind. Nach Satz 8.23 und 6. gilt daher

$$\mathrm{h}(\theta_A) = \mathrm{h}(\theta_D) = \sum_{k=1}^{n} \log|\eta_k| = \sum_{k=1}^{n} \log|\lambda_k| \ ,$$

da $\eta_1, \ldots, \eta_n$ eine Permutation von $\lambda_1, \ldots, \lambda_n$ ist. $\square$

Man beachte: Im diesem Spezialfall müssen die Eigenwerte von A alle Elemente aus $\mathbb{Z} \setminus \{0\}$ sein.

9 Subadditive und multiplikative Ergodensätze

Im Folgenden sei $(X, \mathcal{F}, \mu)$ ein Wahrscheinlichkeitsraum und $T : X \to X$ eine messbare Abbildung, die maßtreu bezüglich μ ist.

Eine Abbildung $g : X \to [-\infty, \infty)$ heißt $\mathcal{F}$-messbar, wenn $f^{-1}([-\infty, t)) \in \mathcal{F}$ für jedes $t \in \mathbb{R}$. In diesem Fall ist die durch $g^+(x) = \max\{g(x), 0\}$ für jedes $x \in X$ definierte Abbildung $g^+ : X \to \mathbb{R}^+$ auch $\mathcal{F}$-messbar. Ist $g : X \to [-\infty, \infty)$ eine $\mathcal{F}$-messbare Abbildung mit $g^+ \in \mathcal{L}^1(X, \mathcal{F}, \mu)$, so macht das Integral $\int g\, d\mu$ einen Sinn (mit $\int g\, d\mu \in [-\infty, \infty)$).

Satz 9.1 (Subadditiver Ergodensatz von Kingman) *Für jedes $n \geq 1$ sei $f_n : X \to [-\infty, \infty)$ eine $\mathcal{F}$-messbare Abbildung mit $f_n^+ \in \mathcal{L}^1(X, \mathcal{F}, \mu)$ und mit*

$$f_{m+n} \leq f_m + f_n \circ T^m \quad \mu\text{-fast sicher}$$

für alle m, $n \geq 1$. Dann gibt es eine $\mathcal{F}$-messbare Abbildung $f : X \to [-\infty, \infty)$, so dass $\lim_{n\to\infty} n^{-1} f_n = f$ μ-fast sicher, und es gilt ferner $f \circ T = f$ μ-fast sicher, $f^+ \in \mathcal{L}^1(X, \mathcal{F}, \mu)$ und $\int f\, d\mu = \inf\limits_{n\geq 1} n^{-1} \int f_n\, d\mu$.

Bemerkung: Sei $g \in \mathcal{L}^1(X, \mathcal{F}, \mu)$ und für jedes $n \geq 1$ setze $f_n = \sum_{k=0}^{n-1} g \circ T^k$. Dann gilt $f_{m+n} = f_m + f_n \circ T^m$ für alle m, $n \geq 1$ und folglich ist Satz 9.1 eine Verallgemeinerung des Birkhoffschen Ergodensatzes.

Beweis Für jedes $N \geq 1$ setze $f_n^{(N)} = \max\{f_n, -nN\}$. Für a, b, c, $d \in [-\infty, \infty)$ gilt $\max\{a + b, c + d\} \leq \max\{a, c\} + \max\{b, d\}$ und damit ist

$$\begin{aligned} f_{m+n}^{(N)} &= \max\{f_{m+n}, -(m+n)N\} \leq \max\{f_m + f_n \circ T^m, -mN - nN\} \\ &\leq \max\{f_m, -mN\} + \max\{f_n \circ T^m, -nN\} = f_m^{(N)} + f_n^{(N)} \circ T^m \end{aligned}$$

μ-fast sicher für alle m, $n \geq 1$, d.h.,

$$f_{m+n}^{(N)} \leq f_m^{(N)} + f_n^{(N)} \circ T^m \quad \mu\text{-fast sicher}$$

für alle m, $n \geq 1$. Nehme an, dass es für jedes $N \geq 1$ ein $g_N \in \mathcal{L}^1(X, \mathcal{F}, \mu)$ mit $g_N \circ T = g_N$ μ-fast sicher gibt, so dass $\lim_{n\to\infty} n^{-1} f_n^{(N)} = g_N$ μ-fast sicher und $\int g_N\, d\mu = \inf_{n\geq 1} n^{-1} \int f_n^{(N)}\, d\mu$. Dann gilt $g_N = \lim_{n\to\infty} \max\{n^{-1} f_n, -N\}$ μ-fast sicher für jedes $N \geq 1$, und folglich gibt es eine $\mathcal{F}$-messbare Abbildung $f : X \to [-\infty, \infty)$, so dass $g_N = \max\{f, -N\}$ μ-fast sicher für jedes $N \geq 1$. Nun gilt $\lim_{n\to\infty} n^{-1} f_n = f$ μ-fast sicher und es ist klar, dass $f \circ T = f$ μ-fast sicher und $f^+ \in \mathcal{L}^1(X, \mathcal{F}, \mu)$. Ferner gilt

$$\begin{aligned} \int f\, d\mu = \inf_{N\geq 1} \int g_N\, d\mu &= \inf_{N\geq 1} \inf_{n\geq 1} \frac{1}{n} \int f_n^{(N)}\, d\mu \\ &= \inf_{n\geq 1} \inf_{N\geq 1} \frac{1}{n} \int f_n^{(N)}\, d\mu = \inf_{n\geq 1} \frac{1}{n} \int f_n\, d\mu\ . \end{aligned}$$

Im Beweis des Satzes kann man daher annehmen, dass es ein $N \geq 1$ gibt mit $f_n \geq -nN$ für alle $n \geq 1$. Insbesondere ist $f_n \in \mathcal{L}^1(X, \mathcal{F}, \mu)$ für jedes $n \geq 1$.

Sei wieder $\mathcal{I}_T$ die invariante σ-Algebra von T. Für jedes $B \in \mathcal{I}_T$ gilt dann $\chi_B f_{m+n} \leq \chi_B f_m + (\chi_B f_n) \circ T^m$ μ-fast sicher und folglich ist

$$\begin{aligned}\int_B f_{m+n}\, d\mu &\leq \int_B (f_m + f_n \circ T^m)\, d\mu \\ &= \int_B f_m\, d\mu + \int_B f_n \circ T^m\, d\mu = \int_B f_m\, d\mu + \int_B f_n\, d\mu\end{aligned}$$

für alle m, $n \geq 1$. Daraus ergibt sich, dass für jedes $B \in \mathcal{I}_T$

$$\lim_{n\to\infty} \frac{1}{n} \int_B f_n\, d\mu = \inf_{n\geq 1} \frac{1}{n} \int_B f_n\, d\mu \ ,$$

da Lemma 8.3 auch für subadditive Folgen aus $[-\infty, \infty)$ gilt. Insbesondere ist

$$\lim_{n\to\infty} \frac{1}{n} \int f_n\, d\mu = \inf_{n\geq 1} \frac{1}{n} \int f_n\, d\mu \ .$$

Ferner gilt für alle m, $n \geq 1$, dass

$$\begin{aligned}\int_B \mathrm{E}_\mu(f_{m+n}|\mathcal{I}_T)\, d\mu &= \int \chi_B f_{m+n}\, d\mu \leq \int \chi_B f_m\, d\mu + \int \chi_B f_n\, d\mu \\ &= \int_B \mathrm{E}_\mu(f_m|\mathcal{I}_T)\, d\mu + \int_B \mathrm{E}_\mu(f_n|\mathcal{I}_T)\, d\mu\end{aligned}$$

für jedes $B \in \mathcal{I}_T$ und folglich ist

$$\mathrm{E}_\mu(f_{m+n}|\mathcal{I}_T) \leq \mathrm{E}_\mu(f_m|\mathcal{I}_T) + \mathrm{E}_\mu(f_n|\mathcal{I}_T) \ \ \mu\text{-fast sicher}$$

für alle m, $n \geq 1$. Damit gilt nach Lemma 8.3, dass

$$\lim_{n\to\infty} n^{-1}\mathrm{E}_\mu(f_n|\mathcal{I}_T) = \inf_{n\geq 1} n^{-1}\mathrm{E}_\mu(f_n|\mathcal{I}_T) \ \ \mu\text{-fast sicher} \ .$$

Sei $\psi = \lim_{n\to\infty} n^{-1}\mathrm{E}_\mu(f_n|\mathcal{I}_T)$; dann ist ψ $\mathcal{I}_T$-messbar und $\psi \geq -N$ μ-fast sicher, da $\mathrm{E}_\mu(f_n|\mathcal{I}_T) \geq -nN$ μ-fast sicher für jedes n.

Lemma 9.1 *Es gilt $\psi \in \mathcal{L}^1(X, \mathcal{F}, \mu)$ und $\psi \circ T = \psi$ μ-fast sicher, und für jedes $B \in \mathcal{I}_T$ ist $\int_B \psi\, d\mu = \lim\limits_{n\to\infty} n^{-1} \int_B f_n\, d\mu$.*

Beweis Da $-N \leq \psi \leq \mathrm{E}_\mu(f_1|\mathcal{I}_T)$ μ-fast sicher, ist $\psi \in \mathcal{L}^1(X, \mathcal{F}, \mu)$, und nach Lemma 2.1 ist $\psi \circ T = \psi$. Nun ist $\psi = \lim_{n\to\infty} 2^{-n}\mathrm{E}_\mu(f_{2^n}|\mathcal{I}_T)$ und für jedes

$k \geq 1$ ist $(2k)^{-1}\mathrm{E}_\mu(f_{2k}|\mathcal{I}_T) \leq k^{-1}\mathrm{E}_\mu(f_k|\mathcal{I}_T)$ μ-fast sicher; nach dem Satz von der monotonen Konvergenz gilt also

$$\int_B \psi\, d\mu = \lim_{n\to\infty} 2^{-n} \int_B \mathrm{E}_\mu(f_{2^n}|\mathcal{I}_T)\, d\mu = \lim_{n\to\infty} 2^{-n} \int_B f_{2^n}\, d\mu$$
$$= \lim_{n\to\infty} n^{-1} \int_B f_n\, d\mu\ . \quad \square$$

Nach Lemma 9.1 genügt es zu zeigen, dass $\lim\limits_{n\to\infty} n^{-1} f_n = \psi$ μ-fast sicher.

Lemma 9.2 *Sei $\varphi : X \to [-\infty,\infty]$ eine $\mathcal{F}$-messbare Abbildung mit $\varphi \circ T \leq \varphi$ μ-fast sicher. Dann gilt $\varphi \circ T = \varphi$ μ-fast sicher. Genauso gilt $\varphi \circ T = \varphi$ μ-fast sicher, wenn $\varphi \circ T \geq \varphi$ μ-fast sicher.*

Beweis Übung. $\square$

Setze $g = \limsup_{n\to\infty} n^{-1} f_n$ und $h = \liminf_{n\to\infty} n^{-1} f_n$. Dann gilt

$$g \circ T = \limsup_{n\to\infty} n^{-1} f_n \circ T \geq \limsup_{n\to\infty} n^{-1}(f_{n+1} - f_1) = g$$

μ-fast sicher und damit ist nach Lemma 9.2 $g \circ T = g$ μ-fast sicher. Genauso ist natürlich $h \circ T = h$ μ-fast sicher.

Das folgende Lemma ist die subadditive Version des maximalen Ergodensatzes von Hopf:

Lemma 9.3 *Für jedes $n \geq 1$ sei $g_n : X \to [-\infty,\infty)$ eine $\mathcal{F}$-messbare Abbildung mit $g_n^+ \in \mathcal{L}^1(X,\mathcal{F},\mu)$ und mit $g_{m+n} \leq g_m + g_n \circ T^m$ μ-fast sicher für alle m, $n \geq 1$. Dann gilt $\int_{E_\infty} g_1\, d\mu \geq 0$, wobei*

$$E_\infty = \{x \in X : \textit{es gibt ein } n \geq 1 \textit{ mit } g_n(x) > 0\}\ .$$

Beweis Dieser ist fast identisch mit dem Beweis für Satz 2.3. $\square$

Seien $\varepsilon > 0$, $k \geq 1$ fest, setze $D_\varepsilon = \{x \in X : g(x) > \psi(x) + \varepsilon\}$ und für jedes $n \geq 1$ sei $g_n = (k^{-1} f_{kn} - n\psi - n\varepsilon)\chi_{D_\varepsilon}$; dann ist $g_n \in \mathcal{L}^1(X,\mathcal{F},\mu)$ und da $\psi \circ T = \psi$ und $\chi_{D_\varepsilon} \circ T = \chi_{D_\varepsilon}$ μ-fast sicher, gilt $g_{m+n} \leq g_m + g_n \circ T^m$ μ-fast sicher für alle m, $n \geq 1$. Daraus folgt nach Lemma 9.3, dass

$$\int_{D_\varepsilon \cap E_{\varepsilon,k}} (k^{-1} f_k - \psi - \varepsilon)\, d\mu \geq 0\ ,$$

wobei $E_{\varepsilon,k} = \{x \in X : \text{es gibt ein } n \geq 1 \text{ mit } (nk)^{-1} f_{nk}(x) > \psi(x) + \varepsilon\}$. Setze nun

$$D_{\varepsilon,k} = \{x \in X : \limsup_{n\to\infty} (nk)^{-1} f_{nk}(x) > \psi(x) + \varepsilon\}\ ;$$

dann ist $D_{\varepsilon,k} \subset E_{\varepsilon,k}$.

Lemma 9.4 *Es gilt* $\mu(D_\varepsilon \setminus D_{\varepsilon,k}) = 0$.

Beweis Man beachte zunächst: Ist $\{a_n\}_{n\geq 0}$ eine reelle Folge, für die die Folge $\{n^{-1}\sum_{k=0}^{n-1} a_k\}_{n\geq 1}$ konvergiert, so gilt insbesondere $\lim\limits_{n\to\infty} n^{-1}a_n = 0$, da

$$n^{-1}a_n = \frac{n+1}{n}(n+1)^{-1}\sum_{k=0}^{n} a_k - n^{-1}\sum_{k=0}^{n-1} a_k$$

für jedes $n \geq 1$. Daraus ergibt sich nach dem Birkhoffschen Ergodensatz, dass $\lim_{n\to\infty} n^{-1} f_p \circ T^n = 0$ μ-fast sicher für jedes $p \geq 1$. Also gibt es ein $C \in \mathcal{F}$ mit $\mu(C) = 1$, so dass $\lim_{n\to\infty} n^{-1} f_p \circ T^n(x) = 0$ für $p = 1, \ldots, k$ für jedes $x \in C$. Aber $(nk)^{-1}f_{nk} \geq (nk)^{-1}f_{nk+p} - (nk)^{-1}f_p \circ T^{nk}$ μ-fast sicher und dies zeigt, dass $\mu((D_\varepsilon \cap C) \setminus D_{\varepsilon,k}) = 0$. Damit ist auch $\mu(D_\varepsilon \setminus D_{\varepsilon,k}) = 0$. □

Nach Lemma 9.4 gilt für jedes $k \geq 1$, dass

$$\begin{aligned}\int_{D_\varepsilon} (k^{-1}f_k - \psi)\,d\mu &= \int_{D_\varepsilon} (k^{-1}f_k - \psi - \varepsilon)\,d\mu + \varepsilon\mu(D_\varepsilon)\\ &= \int_{D_\varepsilon \cap E_{\varepsilon,k}} (k^{-1}f_k - \psi - \varepsilon)\,d\mu + \varepsilon\mu(D_\varepsilon) \geq \varepsilon\mu(D_\varepsilon)\ ,\end{aligned}$$

da $D_\varepsilon \setminus E_{\varepsilon,k} \subset D_\varepsilon \setminus D_{\varepsilon,k}$. Da aber $\chi_{D_\varepsilon} \circ T = \chi_{D_\varepsilon}$ μ-fast sicher, gibt es nach Übungsaufgabe 6 ein $C_\varepsilon \in \mathcal{I}_T$ mit $\mu(D_\varepsilon \bigtriangleup C_\varepsilon) = 0$ und damit gilt nach der letzten Aussage in Lemma 9.1, dass

$$\begin{aligned}\lim_{k\to\infty}\int_{D_\varepsilon} (k^{-1}f_k - \psi)\,d\mu &= \lim_{k\to\infty}\int_{C_\varepsilon} (k^{-1}f_k - \psi)\,d\mu\\ &= \lim_{k\to\infty} k^{-1}\int_{C_\varepsilon} f_k\,d\mu - \int_{C_\varepsilon} \psi\,d\mu = 0\ .\end{aligned}$$

Dies zeigt, dass $\mu(D_\varepsilon) = 0$ für jedes $\varepsilon > 0$, und also ist $g \leq \psi$ μ-fast sicher. Es bleibt nun zu zeigen, dass $\psi \leq h$ μ-fast sicher und dafür wird der folgende Satz von Derriennic benötigt:

Satz 9.2 *Für jedes* $n \geq 1$ *sei* $h_n : X \to [-\infty, \infty)$ *eine* $\mathcal{F}$*-messbare Abbildung mit* $h_n^+ \in \mathcal{L}^1(X, \mathcal{F}, \mu)$ *und mit* $h_{m+n} \leq h_m + h_n \circ T^m$ μ*-fast sicher für alle* $m, n \geq 1$. *Dann gilt* $\lim\limits_{n\to\infty} n^{-1}\int_G h_n\,d\mu \leq 0$, *wobei*

$$G = \{x \in X : \liminf_{n\to\infty} n^{-1}h_n(x) < 0\}\ .$$

Beweis Siehe Derriennic [4]. □

Sei $\varepsilon > 0$ und setze $F_\varepsilon = \{x \in X : \psi(x) > h(x) + \varepsilon\}$; für jedes $n \geq 1$ sei $h_n = (f_n - n\psi + n\varepsilon)\chi_{F_\varepsilon}$. Dann ist $h_n \in \mathcal{L}^1(X, \mathcal{F}, \mu)$ und für alle $m, n \geq 1$ gilt $h_{m+n} \leq h_m + h_n \circ T^m$ μ-fast sicher, da $\psi \circ T = \psi$ und $\chi_{F_\varepsilon} \circ T = \chi_{F_\varepsilon}$ μ-fast sicher. Außerdem ist $\liminf_{n\to\infty} n^{-1} h_n = (h - \psi + \varepsilon)\chi_{F_\varepsilon}$ und daraus folgt, dass $F_\varepsilon = \{x \in X : \liminf_{n\to\infty} n^{-1} h_n(x) < 0\}$. Nach Satz 9.2 und Lemma 9.1 ist also

$$\mu(F_\varepsilon) = \mu(F_\varepsilon) + \lim_{n\to\infty} \int_{F_\varepsilon} (n^{-1} f_n - \psi)\, d\mu = \lim_{n\to\infty} n^{-1} \int_{F_\varepsilon} (f_n - n\psi + n\varepsilon)\, d\mu \leq 0$$

(wieder unter Verwendung von Übungsaufgabe 6). Folglich ist $\mu(F_\varepsilon) = 0$ für alle $\varepsilon > 0$ und dies zeigt, dass $\psi \leq h$ μ-fast sicher. Daraus ergibt sich, dass

$$\lim_{n\to\infty} n^{-1} f_n = \psi \quad \mu\text{-fast sicher}\ ,$$

und damit ist der subadditive Ergodensatz von Kingman bewiesen. □

Für jedes $0 \leq m < n$ sei $f_{m,n} : X \to [-\infty, \infty)$ eine $\mathcal{F}$-messbare Abbildung. Die Familie $\{f_{m,n}\}_{0 \leq m < n}$ heißt *subadditiv*, wenn

$$f_{\ell,n} \leq f_{\ell,m} + f_{m,n} \quad \mu\text{-fast sicher}$$

für alle $0 \leq \ell < m < n$, und sie heißt *stationär* bezüglich μ, wenn

$$\mu\Big(\bigcap_{k=1}^{\ell} f^{-1}_{m_k,n_k}(B_k)\Big) = \mu\Big(\bigcap_{k=1}^{\ell} f^{-1}_{m_k+1,n_k+1}(B_k)\Big)$$

für alle $B_1, \ldots, B_\ell \in \mathcal{B}_{[-\infty,\infty)}$, $0 \leq m_k < n_k$, $k = 1, \ldots, \ell$ und alle $\ell \geq 1$, wobei $\mathcal{B}_{[-\infty,\infty)} \subset \mathcal{P}([-\infty, \infty))$ die kleinste σ-Algebra ist, die die Intervalle $[-\infty, t)$, $t \in \mathbb{R}$, enthält.

Lemma 9.5 *Für jedes $n \geq 1$ sei $g_n : X \to [-\infty, \infty)$ eine $\mathcal{F}$-messbare Abbildung mit $g_{m+n} \leq g_m + g_n \circ T^m$ μ-fast sicher für alle $m, n \geq 1$, und für $0 \leq m < n$ sei $f_{m,n} = g_{n-m} \circ T^m$. Dann ist die Familie $\{f_{m,n}\}_{0 \leq m < n}$ subadditiv und stationär bezüglich μ.*

Beweis Für alle $0 \leq m < n$ gilt

$$\begin{aligned}
\mu(\{x \in X : f_{\ell,n}(x) \leq f_{\ell,m}(x) + f_{m,n}(x)\}) \\
&= \mu(\{x \in X : (g_{n-\ell} \circ T^\ell)(x) \leq (g_{m-\ell} \circ T^\ell)(x) + (g_{n-m} \circ T^m)(x)\}) \\
&= \mu(T^{-\ell}(\{x \in X : g_{n-\ell}(x) \leq g_{m-\ell}(x) + (g_{n-m} \circ T^{m-\ell})(x)\})) \\
&= \mu(\{x \in X : g_{n-\ell}(x) \leq g_{m-\ell}(x) + (g_{n-m} \circ T^{m-\ell})(x)\}) = 1\ ;
\end{aligned}$$

d.h., $f_{\ell,n} \leq f_{\ell,m} + f_{m,n}$ μ-fast sicher. Ferner ist die Familie $\{f_{m,n}\}_{0 \leq m < n}$ stationär, da T maßtreu bezüglich μ ist. □

Satz 9.3 (Stochastische Version von Satz 9.1) *Sei $\{f_{m,n}\}_{0\le m<n}$ eine subadditive stationäre Familie mit $f_{0,1}^+ \in \mathcal{L}^1(X,\mathcal{F},\mu)$. Dann gibt es eine $\mathcal{F}$-messbare Abbildung $f : X \to [-\infty,\infty)$, so dass $\lim_{n\to\infty} n^{-1} f_{0,n} = f$ μ-fast sicher, und es gilt ferner $f^+ \in \mathcal{L}^1(X,\mathcal{F},\mu)$ und*

$$\lim_{n\to\infty} n^{-1} \int f_{0,n}\, d\mu = \inf_{n\ge 1} n^{-1} \int f_{0,n}\, d\mu = \int f\, d\mu\ .$$

Beweis Sei $Z = \mathrm{Abb}(\mathbb{J}, [-\infty,\infty))$, wobei $\mathbb{J} = \{(m,n) \in \mathbb{N}\times\mathbb{N} : m < n\}$, für $(m,n) \in \mathbb{J}$ definiere $\mathrm{p}_{m,n} : Z \to [-\infty,\infty)$ durch $\mathrm{p}_{m,n}(\{z_{k,\ell}\}_{0\le k<\ell}) = z_{m,n}$ und sei $\mathcal{G} = \sigma\big(\bigcup_{0\le m<n} \mathrm{p}_{m,n}^{-1}(\mathcal{B}_{[-\infty,\infty)})\big)$. Definiere eine Abbildung $T : Z \to Z$ durch $T(\{z_{m,n}\}_{0\le m<n}) = \{z'_{m,n}\}_{0\le m<n}$, wobei $z'_{m,n} = z_{m+1,n+1}$. Dann ist T messbar, da $\mathrm{p}_{m,n} \circ T = \mathrm{p}_{m+1,n+1}$ für alle $0 \le m < n$. Sei $\mathbf{f} : X \to Z$ die Abbildung, die definiert ist durch $\mathbf{f}(x) = \{f_{m,n}(x)\}_{0\le m<n}$, also gilt $\mathbf{f}^{-1}(\mathcal{G}) \subset \mathcal{F}$, und sei ν das Bildmaß von μ unter $\mathbf{f}$. Da $\{f_{m,n}\}_{0\le m<n}$ stationär bezüglich μ ist, ist T maßtreu bezüglich ν. Für jedes $n \ge 1$ sei $g_n = \mathrm{p}_{0,n}$; dann ist $g_n^+ \in \mathcal{L}^1(Z,\mathcal{G},\nu)$, da

$$\begin{aligned}\int g_n^+\, d\nu = \int \mathrm{p}_{0,n}^+\, d\nu = \int \mathrm{p}_{0,n}^+ \circ \mathbf{f}\, d\mu \\ = \int f_{0,n}^+\, d\mu \le \int (f_{0,n}^+ + \cdots + f_{n-1,n}^+)\, d\mu = n \int f_{0,1}^+\, d\mu\ .\end{aligned}$$

Seien nun m, $n \ge 1$ und sei $F = \{x \in X : f_{0,m+n}(x) \le f_{0,m}(x) + f_{m,m+n}(x)\}$; also ist $\mu(F) = 1$. Für jedes $x \in F$ gilt

$$\begin{aligned} g_{m+n}(\mathbf{f}(x)) &= f_{0,m+n}(x) \le f_{0,m}(x) + f_{m,m+n}(x) \\ &= g_m(\mathbf{f}(x)) + \mathrm{p}_{m,m+n}(\mathbf{f}(x)) = g_m(\mathbf{f}(x)) + g_n \circ T^m(\mathbf{f}(x)) \end{aligned}$$

und daraus ergibt sich, dass

$$\begin{aligned} &\nu(\{z \in Z : g_{m+n}(z) \le g_m(z) + g_n \circ T^m(z)\}) \\ &\quad = \mu(\mathbf{f}^{-1}(\{z \in Z : g_{m+n}(z) \le g_m(z) + g_n \circ T^m(z)\})) \\ &\quad = \nu(\{x \in X : g_{m+n}(\mathbf{f}(x)) \le g_m(\mathbf{f}(x)) + g_n \circ T^m(\mathbf{f}(x))\}) \ge \mu(F) = 1\ ; \end{aligned}$$

d.h., $g_{m+n} \le g_m + g_n \circ T^m$ ν-fast sicher für alle m, $n \ge 1$. Nach Satz 9.1 gibt es daher eine $\mathcal{G}$-messbare Abbildung $g : Z \to [-\infty,\infty)$, so dass $\lim_{n\to\infty} n^{-1} g_n = g$ ν-fast sicher, und außerdem ist $g^+ \in \mathcal{L}^1(Z,\mathcal{G},\nu)$ und

$$\lim_{n\to\infty} n^{-1} \int g_n\, d\nu = \inf_{n\ge 1} n^{-1} \int g_n\, d\nu = \int g\, d\nu\ .$$

Aber $f_{0,n} = g_n \circ \mathbf{f}$, und folglich gilt $\lim_{n\to\infty} n^{-1} f_{0,n} = f$ μ-fast sicher, wobei $f = g \circ \mathbf{f}$. Ferner ist nach Satz 1.2 $f^+ \in \mathcal{L}^1(X,\mathcal{F},\mu)$, da $\int f^+\, d\mu = \int g^+\, d\nu$ und

$$\lim_{n\to\infty} n^{-1} \int f_{0,n}\, d\mu = \inf_{n\ge 1} n^{-1} \int f_{0,n}\, d\mu = \int f\, d\mu\ ,$$

da $\int f_{0,n}\, d\mu = \int g_n\, d\nu$ für jedes $n \geq 1$ und $\int f\, d\mu = \int g\, d\nu$. $\square$

Sei $p \geq 1$ und sei $(\mathrm{End}(\mathbb{R}^p), \|\cdot\|_\infty)$ die Banachalgebra aller Endomorphismen von $\mathbb{R}^p$ mit der Norm $\|\cdot\|_\infty$ definiert durch $\|f\|_\infty = \sup\{\|f(x)\| : \|x\| \leq 1\}$.

In den folgenden zwei Sätzen sei $f : X \to \mathrm{End}(\mathbb{R}^p)$ eine $\mathcal{F}$-messbare Abbildung mit $(\log \|f(\cdot)\|_\infty)^+$ ein Element von $\mathcal{L}^1(X, \mathcal{F}, \mu)$.

Satz 9.4 *Es gibt eine $\mathcal{F}$-messbare Abbildung $\chi : X \to [-\infty, \infty)$, so dass*

$$\lim_{n\to\infty} n^{-1} \log \|f(T^{n-1}(x)) \circ \cdots \circ f(T(x)) \circ f(x)\|_\infty = \chi(x) \quad \textit{für } \mu\textit{-fast alle } x \in X\ ,$$

und es gilt $\chi \circ T = \chi$ μ-fast sicher, $\chi^+ \in \mathcal{L}^1(X, \mathcal{F}, \mu)$ und

$$\begin{aligned} &\lim_{n\to\infty} \frac{1}{n} \int \log \|f(T^{n-1}(x)) \circ \cdots \circ f(T(x)) \circ f(x)\|_\infty \, d\mu(x) \\ &\qquad = \inf_{n\geq 1} \frac{1}{n} \int \log \|f(T^{n-1}(x)) \circ \cdots \circ f(T(x)) \circ f(x)\|_\infty \, d\mu(x) = \int \chi \, d\mu\ . \end{aligned}$$

Beweis Für jedes $n \geq 1$ definiere $g_n : X \to [-\infty, \infty)$ durch

$$g_n(x) = \log \|f(T^{n-1}(x)) \circ \cdots \circ f(T(x)) \circ f(x)\|_\infty\ ;$$

also ist g_n $\mathcal{F}$-messbar und es gilt

$$g_n(x) \leq \log \prod_{k=0}^{n-1} \|f(T^k(x))\|_\infty = \sum_{k=0}^{n-1} \log \|f(T^k(x))\|_\infty = \sum_{k=0}^{n-1} (\log \|f(\cdot)\|_\infty \circ T^k)(x)$$

für alle $x \in X$. Folglich gilt auch

$$g_n^+ \leq \sum_{k=0}^{n-1} (\log \|f(\cdot)\|_\infty)^+ \circ T^k$$

und damit ist $g_n^+ \in \mathcal{L}^1(X, \mathcal{F}, \mu)$ für jedes $n \geq 1$. Für alle m, $n \geq 1$ gilt nun

$$\begin{aligned} g_{m+n}(x) &= \log \|f(T^{m+n-1}(x)) \circ \cdots \circ f(T(x)) \circ f(x)\|_\infty \\ &\leq \log \|f(T^{m-1}(x)) \circ \cdots \circ f(T(x)) \circ f(x)\|_\infty \\ &\qquad\qquad + \log \|f(T^{m+n-1}(x)) \circ \cdots \circ f(T^m(x))\|_\infty \\ &= g_m(x) + (g_n \circ T^m)(x) \end{aligned}$$

und also ist das Ergebnis eine direkte Folgerung aus Satz 9.1. $\square$

Bezeichne nun mit $\mathrm{End}^+(\mathbb{R}^p)$ die Teilmenge von $\mathrm{End}(\mathbb{R}^p)$ bestehend aus allen diagonalisierbaren Endomorphismen, für die alle Eigenwerte in $\mathbb{R}^+$ liegen. Ein Element $f \in \mathrm{End}(\mathbb{R}^p)$ liegt in $\mathrm{End}^+(\mathbb{R}^p)$ genau dann, wenn es $0 \leq \lambda_1 < \cdots < \lambda_r$ (mit $r \geq 1$) und eine direkte Summe $\mathbb{R}^p = \bigoplus_{k=1}^r U_k$ gibt, so dass $f(v) = \lambda_k v$ für alle $v \in U_k$, $1 = 1, \ldots, r$.

Das folgende Ergebnis ist der multiplikative Ergodensatz von Oseledec.

Satz 9.5 *Es gibt eine $\mathcal{F}$-messbare Abbildung $\Lambda : X \to \mathrm{End}(\mathbb{R}^p)$ und ein $F \in \mathcal{F}$ mit $T(F) \subset F$ und $\mu(F) = 1$, so dass*

(1) $\Lambda(x) \in \mathrm{End}^+(\mathbb{R}^p)$ für jedes $x \in F$.

(2) Für $x \in F$ seien $0 \le \lambda_1(x) < \cdots < \lambda_{r(x)}(x)$ die Eigenwerte von $\Lambda(x)$ und $U_1(x), \ldots, U_{r(x)}(x)$ die entsprechenden Eigenräume; dann gilt:

(a) Für alle $x \in F$ ist $r(T(x)) = r(x)$.

(b) Für alle $k = 1, \ldots, r(x)$, $x \in F$ ist $\lambda_k(T(x)) = \lambda_k(x)$.

(c) Für alle $k = 1, \ldots, r(x)$, $x \in F$ ist $\dim U_k(T(x)) = \dim U_k(x)$.

(d) Für alle $u \in V_k(x) \setminus V_{k-1}(x)$, $k = 1, \ldots, r(x)$, $x \in F$, ist

$$\lim_{n\to\infty} n^{-1} \log \|(f(T^{n-1}(x)) \circ \cdots \circ f(T(x)) \circ f(x))(u)\| = \log \lambda_k(x) ,$$

wobei $V_0(x) = \{0\}$ und $V_k(x) = U_1(x) + \cdots U_k(x)$ für $k = 1, \ldots, r(x)$.

Beweis Sei $g \in \mathrm{End}^+(\mathbb{R}^p)$, $n \ge 1$; dann gibt es ein eindeutiges $g^{1/n} \in \mathrm{End}^+(\mathbb{R}^p)$ mit $(g^{1/n})^n = g$. (Sind $0 \le \lambda_1 < \cdots < \lambda_r$ die Eigenwerte von g und $U_1, \ldots, U_r$ die entsprechenden Eigenräume, so gilt $g^{1/n}(u) = \lambda_k^{1/n} u$ für jedes $u \in U_k$.) Für jedes $g \in \mathrm{End}(\mathbb{R}^p)$ sei g^* der zu g adjungierte Endomorphismus, also ist g^* das eindeutige Element von $\mathrm{End}(\mathbb{R}^p)$ mit $\langle g(u), v\rangle = \langle u, g^*(v)\rangle$ für alle $u, v \in \mathbb{R}^p$. Man sieht leicht, dass $g^* \circ g \in \mathrm{End}^+(\mathbb{R}^p)$ für alle $g \in \mathrm{End}(\mathbb{R}^p)$. Für jedes $g \in \mathrm{End}(\mathbb{R}^p)$ setze $\mathrm{c}(g) = (g^* \circ g)^{1/2}$. Für jedes $v \in \mathbb{R}^p$ gilt

$$\|g(v)\| = \langle g(v), g(v)\rangle^{1/2} = \langle (g^* \circ g)(v), v\rangle^{1/2} ,$$

und daraus folgt, dass $\|g\|_\infty = \lambda^{1/2}$, wobei λ der größte Eigenwert von $g^* \circ g$ ist. Damit ist $\|g\|_\infty$ der größte Eigenwert von $\mathrm{c}(g)$.

Der Schlüssel zum Beweis von Satz 9.5 ist das folgende Ergebnis von Ruelle:

Satz 9.6 *Sei $\{g_n\}_{n\ge1}$ eine Folge aus $\mathrm{End}(\mathbb{R}^p)$ mit $\limsup_{n\to\infty} n^{-1} \log \|g_n\|_\infty \le 0$. Für $n \ge 1$ sei $h_n = g_n \circ \cdots \circ g_1$ und seien $0 \le t_1^{(n)} \le \cdots \le t_p^{(n)}$ die Eigenwerte von $\mathrm{c}(h_n)$ (d.h., alle p Eigenwerte, nicht nur die verschiedenen). Nehme an, dass für jedes $k = 1, \ldots, p$ der Grenzwert $\lim_{n\to\infty} n^{-1} \log t_k^{(n)}$ existiert. Dann gibt es einen Endomorphismus $\Gamma \in \mathrm{End}^+(\mathbb{R}^p)$, so dass*

$$\lim_{n\to\infty} (\mathrm{c}(h_n))^{1/n} = \Gamma .$$

Seien $0 \le \lambda_1 < \cdots < \lambda_r$ die verschiedenen Eigenwerte von Γ und $U_1, \ldots, U_r$ die entsprechenden Eigenräume. Dann gilt

$$\lim_{n\to\infty} n^{-1} \log \|h_n(u)\| = \log \lambda_k$$

für alle $u \in V_k \setminus V_{k-1}$, wobei $V_0 = \{0\}$ und $V_k = U_1 + \cdots + U_k$ für $k = 1, \ldots, r$.

Beweis Siehe Ruelle [9]. □

Sei $h \in \mathrm{End}(\mathbb{R}^p)$; für jedes $k = 1, \ldots, p$ sei $h^{\wedge k} \in \mathrm{End}(\mathcal{A}^k(\mathbb{R}^p, \mathbb{R}))$ der durch

$$h^{\wedge k}(\alpha)(v_1, \ldots, v_k) = \alpha(h^*(v_1), \ldots, h^*(v_k))$$

für alle $\alpha \in \mathcal{A}^k(\mathbb{R}^p, \mathbb{R})$, $(v_1, \ldots, v_k) \in (\mathbb{R}^p)^k$ definierte Endomorphismus. Für alle $h_1, h_2 \in \mathrm{End}(\mathbb{R}^p)$ gilt $(h_1 \circ h_2)^{\wedge k} = h_1^{\wedge k} \circ h_2^{\wedge k}$.

Es gilt $h^{\wedge k} \in \mathrm{End}^+(\mathcal{A}^k(\mathbb{R}^p, \mathbb{R}))$ für alle $h \in \mathrm{End}^+(\mathbb{R}^p)$: Ist $h \in \mathrm{End}^+(\mathbb{R}^p)$, so ist auch $h^* \in \mathrm{End}^+(\mathbb{R}^p)$ und h^* und h haben die gleichen Eigenwerte. Sei $(v_1, \ldots, v_p)$ eine aus Eigenvektoren von h^* bestehende Basis von $\mathbb{R}^p$ mit $h^*(v_j) = \lambda_j v_j$ für jedes $j = 1, \ldots, p$. Sei $(\varphi_1, \ldots, \varphi_p)$ die zugehörige Dualbasis von $(\mathbb{R}^p)^*$; also bilden die Elemente $\{\varphi_{i_1} \wedge \cdots \wedge \varphi_{i_k} : 1 \le i_1 < \cdots < i_k \le p\}$ eine Basis von $\mathcal{A}^k(\mathbb{R}^p, \mathbb{R})$ und man rechnet leicht nach, dass für alle $1 \le i_1 < \cdots < i_k \le p$

$$h^{\wedge k}(\varphi_{i_1} \wedge \cdots \wedge \varphi_{i_k}) = \lambda_{i_1} \times \cdots \times \lambda_{i_k} \varphi_{i_1} \wedge \cdots \wedge \varphi_{i_k} .$$

Ist $\lambda_1 \le \cdots \le \lambda_p$, so ist insbesondere $\lambda_p \times \cdots \times \lambda_{p-k+1}$ der größte Eigenwert von $h^{\wedge k}$.

Sei $(e_1, \ldots, e_p)$ die kanonische Basis von $\mathbb{R}^p$ und sei $(\tau_1, \ldots, \tau_p)$ die zugehörige Dualbasis von $(\mathbb{R}^p)^*$; dann gibt es ein eindeutiges Skalarprodukt auf $\mathcal{A}^k(\mathbb{R}^p, \mathbb{R})$, so dass die Elemente $\{\tau_{i_1} \wedge \cdots \wedge \tau_{i_k} : 1 \le i_1 < \cdots < i_k \le p\}$ eine orthonormale Basis bilden. Bezüglich dieses Skalarproduktes wird $\mathcal{A}^k(\mathbb{R}^p, \mathbb{R})$ als euklidischer Vektorraum betrachtet. Mit einem gewissen Aufwand rechnet man nach, dass $(h^{\wedge k})^* = (h^*)^{\wedge k}$ für alle $h \in \mathrm{End}(\mathbb{R}^p)$.

Sei $h \in \mathrm{End}(\mathbb{R}^p)$ und seien $0 \le t_1 \le \cdots \le t_p$ die Eigenwerte von $\mathrm{c}(h)$. Da $(h^{\wedge k})^* \circ h^{\wedge k} = (h^* \circ h)^{\wedge k}$, ist $\|h^{\wedge k}\|_\infty = t_p \times \cdots \times t_{p-k+1}$, und insbesondere ist $\|h^{\wedge k}\|_\infty \le \|h\|_\infty^k$, da $\|h\|_\infty = t_p$.

Sei nun $\{h_n\}_{n\ge1}$ eine Folge aus $\mathrm{End}(\mathbb{R}^p)$ und seien $0 \le t_1^{(n)} \le \cdots \le t_p^{(n)}$ die Eigenwerte von $\mathrm{c}(h_n)$. Nehme an, dass für jedes $k = 1, \ldots, p$ der Grenzwert $\lim_{n\to\infty} n^{-1} \log \|h_n^{\wedge k}\|_\infty$ existiert. Dann existiert $\lim_{n\to\infty} n^{-1} \sum_{j=p-k+1}^{p} \log t_j^{(n)}$ für jedes $k = 1, \ldots, p$ und also existiert auch $\lim_{n\to\infty} n^{-1} \log t_k^{(n)}$ für $k = 1, \ldots, p$.

Betrachte jetzt die $\mathcal{F}$-messbare Abbildung $f : X \to \mathrm{End}(\mathbb{R}^p)$ mit $(\log \|f(\cdot)\|_\infty)^+$ ein Element von $\mathcal{L}^1(X, \mathcal{F}, \mu)$. Für $k = 1, \ldots, p$ ist $f^{\wedge k} : X \to \mathrm{End}(\mathcal{A}^k(\mathbb{R}^p, \mathbb{R}))$ $\mathcal{F}$-messbar und $(\log \|f^{\wedge k}(\cdot)\|_\infty)^+ \in \mathcal{L}^1(X, \mathcal{F}, \mu)$, da

$$(\log \|f^{\wedge k}(x)\|_\infty)^+ \le (\log \|f(x)\|_\infty^k)^+ = k(\log \|f(x)\|_\infty)^+ .$$

Für $n \ge 1$ sei $g_n(x) = f(T^{n-1}(x)) \circ \cdots \circ f(x)$; dann gilt

$$g_n^{\wedge k}(x) = (f(T^{n-1}(x)) \circ \cdots \circ f(x))^{\wedge k} = f^{\wedge k}(T^{n-1}(x)) \circ \cdots \circ f^{\wedge k}(x) ,$$

und also gibt es nach Satz 9.4 eine $\mathcal{F}$-messbare Abbildung $\chi_k : X \to [-\infty, \infty)$ und $F_k \in \mathcal{F}$ mit $\mu(F_k) = 1$, so dass

$$\lim_{n\to\infty} n^{-1}\|g_n^{\wedge k}(x)\|_\infty = \chi_k(x) = \chi_k(T(x))$$

für alle $x \in F_k$. Nach dem Birkhoffschen Ergodensatz gilt aber

$$\lim_{n\to\infty} n^{-1}\sum_{j=0}^{n-1}(\log\|f(T^j(\cdot))\|_\infty)^+ = \mathrm{E}_\mu((\log\|f(\cdot)\|_\infty)^+|\mathcal{I}_T) \quad \mu\text{-fast sicher}$$

und damit gibt es ein $F_0 \in \mathcal{F}$ mit $\mu(F_0) = 1$, so dass

$$\limsup_{n\to\infty} n^{-1}\log\|f(T^{n-1}(x))\|_\infty \le 0$$

für alle $x \in F_0$. Setze nun $F = \bigcap_{n\ge 0} T^{-n}\big(\bigcap_{j=0}^p F_j\big)$; dann ist $F \in \mathcal{F}$ mit $T(F) \subset F$ und $\mu(F) = 1$, und für alle $x \in F$ gilt

$$\lim_{n\to\infty} n^{-1}\|g_n^{\wedge k}(x)\|_\infty = \chi_k(x) = \chi_k(T(x))$$

für $k = 1, \ldots, p$ und $\limsup_{n\to\infty} n^{-1}\log\|f(T^{n-1}(x))\|_\infty \le 0$. Nach Satz 9.6 gibt es für jedes $x \in F$ ein $\Lambda(x) \in \mathrm{End}^+(\mathbb{R}^p)$, so dass $\lim_{n\to\infty}(\mathrm{c}(g_n(x)))^{1/n} = \Lambda(x)$, und für $x \in X \setminus F$ setze $\Lambda(x) = 0$; es ist klar, dass $\Lambda : X \to \mathrm{End}(\mathbb{R}^p)$ $\mathcal{F}$-messbar ist. Für $x \in F$ seien $0 \le t_1(x) \le \cdots \le t_p(x)$ die Eigenwerte von $\Lambda(x)$; man sieht leicht, dass $\chi_k(x) = \sum_{j=p-k+1}^p \log t_j(x)$ für jedes $k = 1, \ldots, p$ und daraus ergibt sich, dass $t_k(x) = t_k(T(x))$ für $k = 1, \ldots, p$ und alle $x \in F$.

Wenn aber $0 \le \lambda_1(x) < \cdots < \lambda_{r(x)}(x)$ die verschiedenen Eigenwerte von $\Lambda(x)$ sind und $U_1(x), \ldots, U_{r(x)}(x)$ die entsprechenden Eigenräume, so sind die Zahlen $r(x)$, $\lambda_1(x), \ldots, \lambda_{r(x)}(x)$ und $\dim U_1(x), \ldots, \dim U_{r(x)}(x)$ eindeutig durch die Zahlen $t_1(x), \ldots, t_{r(x)}(x)$ bestimmt. Dies zeigt, dass (2) (a), (b) und (c) richtig sind und (2) (d) folgt unmittelbar aus der letzten Aussage in Satz 9.6. □

10 Topologische Dynamische Systeme

Sei (X, d) ein kompakter metrischer Raum und sei $T : X \to X$ eine stetige Abbildung. Die topologischen Eigenschaften der iterierten Abbildungen werden jetzt studiert.

Typische Beispiele sind:

(1) $X = G$ eine kompakte metrische Gruppe mit $T : G \to G$ definiert entweder durch $T(x) = ax$ oder $T(x) = f(x)$ für alle $x \in G$, wobei $a \in G$ und f ein stetiger Endomorphismus ist.

(2) $X = [0, 1]$ und $T(x) = \mu x(1 - x)$ für alle $x \in [0, 1]$, wobei $\mu \in [0, 4]$.

(3) (*Shift-Räume*) Sei (E, ϱ) ein kompakter metrischer Raum und sei I enweder $\mathbb{Z}$ oder $\mathbb{N} = \{0, 1, \ldots\}$. Setze $E_\infty = \mathrm{Abb}(I, E)$; wie üblich werden die Elemente dargestellt als Folgen E_∞ mit $x_n \in E$ für jedes $n \in I$. Für jedes $m \in I$ sei $\mathrm{p}_m : E_\infty \to E$ die durch $\mathrm{p}_m(\{x_n\}_{n\in I}) = x_m$ gegebene Projektionsabbildung. Definiere $\varrho_\infty : E_\infty \times E_\infty \to \mathbb{R}^+$ durch

$$\varrho_\infty(\{x_n\}_{n\in I}, \{y_n\}_{n\in I}) = \sum_{n\in I} 2^{-|n|} \varrho(x_n, y_n)$$

für alle $\{x_n\}_{n\in I}$, $\{y_n\}_{n\in I} \in E_\infty$. (Man beachte: Da E kompakt ist, gibt es ein $B \in \mathbb{R}^+$, so dass $\varrho(x, y) \le B$ für alle x, $y \in E$ und folglich konvergiert diese Reihe.)

Lemma 10.1 *(1) ϱ_∞ ist eine Metrik auf E_∞ und der metrische Raum $(E_\infty, \varrho_\infty)$ ist kompakt.*

(2) Eine Folge $\{x_m\}_{m\ge p}$ aus E_∞ konvergiert genau dann gegen $x \in E_\infty$, wenn $\lim_{m\to\infty} \varrho(\mathrm{p}_n(x_m), \varrho_n(x))) = 0$ für jedes $n \in I$.

(3) Die Abbildungen $\mathrm{p}_n : E_\infty \to E$, $n \in I$, sind stetig.

(4) Ist Y ein topologischer Raum, so ist eine Abbildung $f : Y \to E_\infty$ stetig genau dann, wenn für jedes $n \in I$ die Abbildung $\mathrm{p}_n \circ f : Y \to E$ stetig ist.

Beweis Übung. □

Definiere $S : E_\infty \to E_\infty$ durch $S(\{x_n\}_{n\in I}) = \{x'_n\}_{n\in I}$, wobei $x'_n = x_{n+1}$. Dann gilt $\mathrm{p}_{n+1} = \mathrm{p}_n \circ S$ für jedes $n \in I$ und daraus ergibt sich nach Lemma 10.1, dass S stetig ist. Ferner ist S ein Homeomorphismus, falls $I = \mathbb{Z}$.

(4) Sei X ein kompakter metrischer Raum und sei $T : X \to X$ eine stetige Abbildung; sei Y eine abgeschlossene Teilmenge von X mit $T(Y) \subset Y$. Dann ist Y mit der induzierten Metrik ein kompakter metrischer Raum und die durch

$T_Y(y) = T(y)$ für alle $y \in Y$ definierte Abbildung $T_Y : Y \to Y$ ist stetig. Ist T ein Homeomorphismus und ist $T(Y) = Y$, so ist auch T_Y ein Homeomorphismus.

Ist $X = E_\infty$ wie in (3) und $T = S$, so heißt (Y, S_Y) ein *Subshift.*

(5) Sei E endlich und I entweder $\mathbb{Z}$ oder $\mathbb{N}$; sei $M : E \times E \to \{0, 1\}$ und setze

$$Y = \{\{x_n\}_{n \in I} \in E_\infty : M(x_n, x_{n+1}) = 1 \text{ für alle } n \in I\} .$$

Nach Lemma 10.1 ist Y eine abgeschlossene Teilmenge von E_∞ und $S(Y) \subset Y$. (Y, S_Y) heißt *Markov-Subshift mit Übergangsmatrix* M.

Im Folgenden sei (X, d) ein kompakter metrischer Raum mit $X \neq \varnothing$ und sei $T : X \to X$ eine stetige Abbildung. Die Abbildung T heißt *minimal*, wenn $\overline{\{T^n(x) : n \geq 0\}} = X$ für jedes $x \in X$.

Satz 10.1 *Äquivalent sind:*

(1) T ist minimal.

(2) Für jede abgeschlossene Teilmenge E von X mit $T(E) \subset E$ gilt entweder $E = X$ oder $E = \varnothing$.

(3) Für jede nichtleere offene Teilmenge U von X gilt $\bigcup_{n \geq 0} T^{-n}(U) = X$.

Beweis (1) $\Rightarrow$ (2): Sei E eine nichtleere abgeschlossene Teilmenge von X mit $T(E) \subset E$ und sei $x \in E$. Da $T(E) \subset E$, ist $\{T^n(x) : n \geq 0\} \subset E$, und damit ist $X = \overline{\{T^n(x) : n \geq 0\}} \subset E$, da E abgeschlossen ist, d.h., $E = X$.

(2) $\Rightarrow$ (3): Sei $U \subset X$ nichtleer und offen und setze $E = X \setminus \bigcup_{n \geq 0} T^{-n}(U)$. Dann ist E abgeschlossen und $T(E) \subset E$; ferner ist $E \neq X$, da $E \cap U = \varnothing$ und folglich ist $E = \varnothing$, d.h., $\bigcup_{n \geq 0} T^{-n}(U) = X$.

(3) $\Rightarrow$ (1): Sei $x \in X$ und sei U eine nichtleere offene Teilmenge von X. Dann gibt es ein $m \geq 0$ mit $x \in T^{-m}(x) \in U$, d.h., mit $T^m(x) \in U$. Folglich gilt $\{T^n(x) : n \geq 0\} \cap U \neq \varnothing$ und damit ist $\overline{\{T^n(x) : n \geq 0\}} = X$. $\square$

Sei Y eine Menge und $g : Y \to Y$ eine Abbildung; für $m \geq 1$ heißt $y \in Y$ *periodischer Punkt* von g mit *Periode* m, wenn $g^m(y) = y$ aber $g^k(y) \neq y$ für $k = 1, \ldots, m-1$. Die Menge der periodischen Punkte mit Periode m wird mit $\mathrm{Per}(m, g)$ bezeichnet. Insbesondere ist $\mathrm{Per}(1, g) = \mathrm{Fix}(g)$, die Menge der *Fixpunkte* von g, und für jedes $m \geq 1$ gilt $\mathrm{Per}(m, g) \subset \mathrm{Fix}(g^m)$.

Sei nun $x \in \mathrm{Per}(m, T)$ und setze $E = \{x, T(x), \ldots, T^{m-1}(x)\}$. Dann ist die Menge E abgeschlossen und $T(E) \subset E$. Ist also X unendlich und T minimal, so folgt aus Satz 10.1, dass $\mathrm{Per}(m, T) = \varnothing$ für jedes $m \geq 1$. Dies deutet an, dass minimale Abbildungen eher "untypisch" sind.

Satz 10.2 *Sei T ein Homeomorphismus. Dann sind äquivalent:*

(1) T ist minimal.

(2) Für jedes $x \in X$ gilt $\overline{\{T^n(x) : n \in \mathbb{Z}\}} = X$.

(3) Für jede abgeschlossene Teilmenge E von X mit $T(E) = E$ gilt entweder $E = X$ oder $E = \varnothing$.

(4) Für jede nichtleere offene Teilmenge U von X gilt $\bigcup_{n\in\mathbb{Z}} T^{-n}(U) = X$.

Beweis (1) $\Rightarrow$ (2): Dies ist klar.

(2) $\Rightarrow$ (3)$\Rightarrow$ (4)$\Rightarrow$ (2): Dies ist fast identisch mit dem Beweis für Satz 10.1.

(4) $\Rightarrow$ (1): Sei U eine nichtleere offene Teilmenge von X. Da $\bigcup_{n\in\mathbb{Z}} T^{-n}(U) = X$ und X kompakt ist, gibt es eine endliche Menge $F \subset \mathbb{Z}$ mit $\bigcup_{n\in F} T^{-n}(U) = X$. Wähle $m \geq 0$, so dass $m + n \geq 0$ für alle $n \in F$. Dann gilt

$$\bigcup_{n\geq 0} T^{-n}(U) \supset \bigcup_{n\in F} T^{-n-m}(U) = T^{-1}(\bigcup_{n\in F} T^{-n}(U)) = T^{-m}(X) = X$$

und damit ist $\bigcup_{n\geq 0} T^{-n}(U) = X$. Nach Satz 10.1 ist also T minimal. □

Eine Teilmenge B eines topologischen Raumes Y heißt *dicht*, wenn $\bar{B} = Y$. Natürlich ist B genau dann dicht, wenn $U \cap B \neq \varnothing$ für jede nichtleere offene Menge $U \subset Y$.

Lemma 10.2 *Sind $U_1, \ldots, U_n$ dichte offene Teilmengen von Y, so ist die offene Menge $\bigcap_{k=1}^n U_k$ auch dicht.*

Beweis Sei U eine nichtleere offene Teilmenge von Y; da U_1 dicht ist, ist $U \cap U_1$ eine nichtleere offene Menge, und da U_2 dicht ist, ist dann $U \cap U_1 \cap U_2$ eine nichtleere offene Menge. Eine $(n-1)$-malige Wiederholung dieses Verfahrens zeigt dann, dass $U \cap \bigcap_{k=1}^n U_k$ nichtleer ist. Also ist $\bigcap_{k=1}^n U_k$ dicht. □

Eine Teilmenge B von Y heißt G_δ*-Menge*, wenn es eine Folge $\{U_n\}_{n\geq 0}$ offener Teilmengen von Y mit $B = \bigcap_{n\geq 0} U_n$ gibt.

Lemma 10.3 (Satz von Baire) *Sei Y ein vollständiger metrischer Raum und sei $\{U_n\}_{n\geq 0}$ eine Folge dichter offener Teilmengen von Y. Dann ist die G_δ-Menge $\bigcap_{n\geq 0} U_n$ auch dicht.*

Beweis Sei U eine nichtleere offene Teilmenge von X. Da U_0 offen und dicht ist, ist $U \cap U_0$ eine nichtleere offene Menge und also gibt es $x_0 \in X$ und $\varepsilon_0 > 0$, so dass $\mathrm{B}(x_0, \varepsilon_0) \subset U \cap U_0$. Da U_1 offen und dicht und $\mathrm{B}(x_0, \varepsilon_0/2)$ eine nichtleere offene Menge ist, ist $\mathrm{B}(x_0, \varepsilon_0/2) \cap U_1$ eine nichtleere offene Menge. Folglich gibt es $x_1 \in X$ und $0 < \varepsilon_1 \le \varepsilon_0/2$, so dass $\mathrm{B}(x_1, \varepsilon_1) \subset \mathrm{B}(x_0, \varepsilon_0/2) \cap U_1$. Auf diese Weise kann man eine Folge $\{x_n\}_{n\ge 0}$ aus X und eine Folge $\{\varepsilon_n\}_{n\ge 0}$ konstruieren, so dass $0 < \varepsilon_n \le \varepsilon_{n-1}/2$ und $\mathrm{B}(x_n, \varepsilon_n) \subset \mathrm{B}(x_{n-1}, \varepsilon_{n-1}/2) \cap U_n$ für jedes $n \ge 1$. Da $x_m \in \mathrm{B}(x_n, \varepsilon_n)$ für alle $m \ge n$ und $\varepsilon_n \le 2^{-n}\varepsilon_0$, ist $\{x_n\}_{n\ge 0}$ eine Cauchy-Folge und damit konvergiert $\{x_n\}_{n\ge 0}$ gegen ein $x \in X$. Da aber $x_m \in \mathrm{B}(x_n, \varepsilon_n/2)$ für alle $m > n$, ist $x \in \overline{\mathrm{B}(x_n, \varepsilon_n/2)} \subset \mathrm{B}(x_n, \varepsilon_n) \subset U_n$ für alle $n \ge 1$ und ferner ist $X \in U \cap U_0$. Daraus ergibt sich, dass $x \in U \cap \bigcap_{n\ge 0} U_n$, d.h., $U \cap \bigcap_{n\ge 0} U_n \ne \varnothing$. Dies zeigt, dass $\bigcap_{n\ge 0} U_n$ dicht ist. $\square$

Erinnerung: Ein kompakter metrischer Raum ist vollständig.

Die stetige Abbildung $T : X \to X$ heißt *einseitig topologisch transitiv*, wenn es ein $x \in X$ gibt, so dass $\overline{\{T^n(x) : n \ge 0\}} = X$. Ist T ein Homeomorphismus, so heißt T *topologisch transitiv*, wenn es ein $x \in X$ gibt, so dass $\overline{\{T^n(x) : n \in \mathbb{Z}\}} = X$. Da $X \ne \varnothing$, ist eine minimale Abbildung stets einseitig topologisch transitiv und ein einseitig topologisch transitiver Homeomorphismus ist stets topologisch transitiv.

Satz 10.3 *Äquivalent sind:*

(1) T ist einseitig topologisch transitiv und $T(X) = X$.

(2) Für jede abgeschlossene Teilmenge E von X mit $T(E) \subset E$ gilt entweder $E = X$ oder $E^\circ = \varnothing$.

(3) Für jede nichtleere offene Teilmenge U von X ist $\bigcup_{n\ge 0} T^{-n}(U)$ eine dichte Teilmenge von X.

(4) $\{x \in X : \overline{\{T^n(x) : n \ge 0\}} = X\}$ ist eine dichte G_δ-Menge.

(5) $\{x \in X : \overline{\{T^n(x) : n \ge 0\}} = X\}$ ist eine dichte Teilmenge von X.

Beweis (1) $\Rightarrow$ (2): Sei $E \subset X$ abgeschlossen mit $T(E) \subset E$ und $E^\circ \ne \varnothing$, also gibt es eine nichtleere offene Teilmenge U von X mit $U \subset E$. Sei $x \in X$ mit $\overline{\{T^n(x) : n \ge 0\}} = X$; dann gibt es ein $m \ge 0$, so dass $T^m(x) \in U \subset E$, und dann gilt $T^n(x) \in E$ für alle $n \ge m$, da $T(E) \subset E$. Folglich ist

$$\begin{aligned} X = T^m(X) = T^m(\overline{\{T^n(x) : n \ge 0\}}) &\subset \overline{T^m(\{T^n(x) : n \ge 0\})} \\ &= \overline{\{T^n(x) : n \ge m\}} \subset \bar{E} = E\,; \end{aligned}$$

d.h., $E = X$. (Man beachte: Sind Y, Z topologische Räume und ist $f : Y \to Z$ stetig, so gilt $f(\bar{B}) \subset \overline{f(B)}$ für jedes $B \subset Y$.)

(2) $\Rightarrow$ (3): Sei $U \subset X$ nichtleer und offen und setze $E = X \setminus \bigcup_{n\ge 0} T^{-n}(U)$. Dann ist E abgeschlossen und $T(E) \subset E$; ferner ist $E \ne X$, da $E \cap U = \varnothing$ und folglich ist $E^\circ = \varnothing$, d.h., $\overline{\bigcup_{n\ge 0} T^{-n}(U)} = X$.

(3) $\Rightarrow$ (4): Sei $\{U_m\}_{m\geq 1}$ eine Basis für die Topologie auf X. Für jedes $x \in X$ gilt $\overline{\{T^n(x) : n \geq 0\}} = X$ genau dann, wenn $\{T^n(x) : n \geq 0\} \cap U_m \neq \varnothing$ für jedes $m \geq 1$, d.h., genau dann, wenn $x \in \bigcup_{n\geq 0} T^{-n}(U_m)$ für alle $m \geq 1$. Damit ist

$$\{x \in X : \overline{\{T^n(x) : n \geq 0\}} = X\} = \bigcap_{m\geq 1} \bigcup_{n\geq 0} T^{-n}(U_m)\,.$$

Aber für jedes $m \geq 1$ ist $\bigcup_{n\geq 0} T^{-n}(U_m)$ eine dichte offene Teilmenge von X und daraus ergibt sich nach dem Satz von Baire, dass $\{x \in X : \overline{\{T^n(x) : n \geq 0\}} = X\}$ eine dichte G_δ-Menge ist.

(4) $\Rightarrow$ (5): Dies ist klar.

(5) $\Rightarrow$ (1): Da $X \neq \varnothing$, ist es klar, dass T einseitig topologisch transitiv ist. Setze $D = \{x \in X : \overline{\{T^n(x) : n \geq 0\}} = X\}$ und sei $x \in D$; dann gilt

$$X = \overline{\{T^n(x) : n \geq 0\}} = \{x\} \cup \overline{\{T^n(x) : n \geq 1\}} \subset \{x\} \cup T(X)$$

und folglich ist $X \setminus T(X) \subset \{x\}$ für jedes $x \in D$. Insbesondere ist $T(X) = X$, falls D aus mehr als einem Element besteht. Ist aber $D = \{z\}$ für ein $z \in X$, so ist auch $X = \{z\}$ und dann ist $T(X) = X$ trivial richtig. $\square$

Sei $x \in X$ mit $\overline{\{T^n(x) : n \geq 0\}} = X$; dann gilt $X \setminus T(X) \subset \{x\}$ und damit ist entweder $T(X) = X$ oder $X \setminus T(X) = \{x\}$. Aber $X \setminus T(X)$ ist offen und daraus ergibt sich, dass $X \setminus T(X) = \{x\}$ nur möglich ist, wenn x ein isolierter Punkt von X ist. Ist also T einseitig topologisch transitiv und enthält X keinen isolierten Punkt, so ist $T(X) = X$ automatisch erfüllt. Betrachte andererseits das Beispiel mit $X = \{0\} \cup \{1/n : n \geq 1\}$ (mit der induzierten Metrik als Teilmenge von $\mathbb{R}$) und mit $T : X \to X$ gegeben durch $T(0) = 0$ und $T(1/n) = 1/(n+1)$ für jedes $n \geq 1$. Hier ist $\overline{\{T^n(1) : n \geq 0\}} = X$ aber $T(X) = X \setminus \{1\}$.

Satz 10.4 *Sei T ein Homeomorphismus. Dann sind äquivalent:*

(1) T ist topologisch transitiv.

(2) Für jede abgeschlossene Teilmenge E von X mit $T(E) = E$ gilt entweder $E = X$ oder $E^\circ = \varnothing$.

(3) Für jede nichtleere offene Teilmenge U von X ist $\bigcup_{n\in\mathbb{Z}} T^{-n}(U)$ eine dichte Teilmenge von X.

(4) $\{x \in X : \overline{\{T^n(x) : n \in \mathbb{Z}\}} = X\}$ ist eine dichte G_δ-Menge.

(5) $\{x \in X : \overline{\{T^n(x) : n \in \mathbb{Z}\}} = X\}$ ist eine dichte Teilmenge von X.

Beweis Dieser ist sehr ähnlich zum Beweis für Satz 10.3. $\square$

Sei $X = \{0\} \cup \{1/n : n \geq 1\} \cup \{2 - 1/n : n \geq 1\} \cup \{2\}$ (mit der induzierten Metrik als Teilmenge von $\mathbb{R}$) und sei $T : X \to X$ gegeben durch $T(0) = 0$, $T(1/n) = 1/(n+1)$ für jedes $n \geq 1$, $T(2 - 1/n) = 2 - 1/(n-1)$ für jedes $n \geq 2$ und $T(2) = 2$. Dann ist T ein Homeomorphismus, $\overline{\{T^n(x) : n \in \mathbb{Z}\}} = X$ für alle $x \in X \setminus \{0, 2\}$ und $\overline{\{T^n(x) : n \geq 0\}} \neq X$ für alle $x \in X$. Also ist T topologisch transitiv aber nicht einseitig topologisch transitiv.

Satz 10.5 *Sei G eine kompakte metrische Gruppe, sei $a \in G$ und sei $U : G \to G$ gegeben durch $U(x) = ax$ für alle $x \in G$. Dann sind äquivalent:*

(1) U ist minimal.

(2) U ist einseitig topologisch transitiv.

(3) U ist topologisch transitiv.

(4) U ist ergodisch bezüglich m_G.

(5) $\{a^n : n \geq 0\}$ ist dicht in G.

(Gilt aber eine von diesen Bedingungen, so muss nach Lemma 6.6 (2) G abelsch sein.)

Beweis (1) $\Rightarrow$ (5) $\Rightarrow$ (2) $\Rightarrow$ (3) sind klar und (4) $\Leftrightarrow$ (5) ist Satz 6.8 (2).

(3) $\Rightarrow$ (1): Es gibt ein $x \in G$, so dass $\overline{\{U^n(x) : n \in \mathbb{Z}\}} = G$ und daraus folgt, dass $\overline{\{U^n(y) : n \in \mathbb{Z}\}} = G$ für alle $y \in G$, da $U^n(y) = a^n y = U^n(x)x^{-1}y$ und die Abbildung $z \mapsto zx^{-1}y$ ein Homeomorphismus ist. Damit ist nach Satz 10.2 U minimal. □

Satz 10.6 *Sei $\mu \in \mathrm{P}(X, \mathcal{B}_X)$ mit $\mu(U) > 0$ für jede nicht-leere offene Teilmenge U von X und mit T maßtreu bezüglich μ. Ist T ergodisch bezüglich μ, so ist T einseitig topologisch transitiv.*

Beweis Dies folgt unmittelbar aus Satz 4.2. □

Sei G eine kompakte abelsche metrische Gruppe und sei $\theta : G \to G$ ein stetiger surjektiver Endomorphismus mit $\gamma \circ \theta^n \neq \gamma$ für alle $n \geq 1$ und alle $\gamma \in \widehat{G} \setminus \{1\}$. Nach Satz 10.6 und Satz 6.9 ist dann θ einseitig topologisch transitiv. (Ist aber $G \neq \{1\}$, so ist θ nicht minimal, da $1 \in \mathrm{Fix}(\theta)$.) Ist insbesondere $A \in \mathrm{M}(n \times n, \mathbb{Z})$ mit $\det(A) \neq 0$ und besitzt A keine Einheitswurzel als Eigenwert, so ist die Abbildung $\theta_A : \mathrm{K}^n \to \mathrm{K}^n$ einseitig topologisch transitiv.

Sei $T : [0, 1] \to [0, 1]$ die Abbildung, die gegeben ist durch

$$T(x) = \begin{cases} 2x & \text{falls } 0 \leq x \leq \frac{1}{2}\,, \\ 2 - 2x & \text{falls } \frac{1}{2} \leq x \leq 1\,. \end{cases}$$

Nach Satz 10.6 und Übungsaufgabe 3 ist T einseitig topologisch transitiv. (Da $0 \in \mathrm{Fix}(T)$, ist T nicht minimal, und in der Tat sieht man leicht, dass hier die Menge $\bigcup_{m\geq 1} \mathrm{Per}(m,T)$ dicht in $[0,1]$ liegt.)

Satz 10.7 *Sei E eine nichtleere Menge und sei (Y, S_Y) das Markov-Subshift mit Übergangsmatrix $M : E \times E \to \{0,1\}$ (und mit I entweder $\mathbb{N}$ oder $\mathbb{Z}$). Dann ist S_Y einseitig topologisch transitiv, falls die Matrix M irreduzibel ist. (Irreduzibel bedeutet, dass es für jedes u, $v \in E$ ein $n \geq 1$ mit $M^n(u,v) > 0$ gibt.)*

Beweis Übung. □

Setze nun

$$\Omega(T) = \{x \in X : \text{zu jeder Umgebung } U \text{ von } x \text{ gibt es ein } n \geq 1 \text{ mit } T^{-n}(U) \cap U \neq \varnothing\} \; ;$$

$\Omega(T)$ heißt die *nicht-wandernde Menge* von T.

Satz 10.8 *Es gilt:*

(1) $\Omega(T)$ ist eine nichtleere abgeschlossene Teilmenge von X.

(2) Für jedes $m \geq 1$ ist $\mathrm{Per}(m,T) \subset \Omega(T)$.

(3) $T(\Omega(T)) \subset \Omega(T)$.

(4) Ist T ein Homeomorphismus, so ist $\Omega(T^{-1}) = \Omega(T)$ und $T(\Omega(T)) = \Omega(T)$.

Beweis (1) Es ist klar, dass $X \setminus \Omega(T)$ offen ist und damit ist $\Omega(T)$ abgeschlossen. Sei $x \in X$; da X kompakt ist, gibt es eine Teilfolge $\{n_k\}_{k\geq 0}$ und $y \in X$, so dass $\lim_{k\to\infty} T^{n_k}(x) = y$. Sei U eine Umgebung von y; dann gibt es $m \geq 0$, so dass $T^{n_k}(x) \in U$ für alle $k \geq m$. Daraus ergibt sich, dass $T^{n_m}(x) \in U$ und $T^{n_{m+1}-n_m}(T^{n_m}(x)) = T^{n_{m+1}}(x) \in U$, und folglich ist $T^{-(n_{m+1}-n_m)}(U) \cap U \neq \varnothing$. Dies zeigt, dass $y \in \Omega(T)$.

(2) Sei $x \in \mathrm{Per}(m,T)$; dann gilt $x \in T^{-1}(U) \cap U$ für jede Umgebung U von x und damit ist $\mathrm{Per}(m,T) \subset \Omega(T)$.

(3) Sei $x \in \Omega(T)$ und sei U eine Umgebung von $T(x)$; dann ist $T^{-1}(U)$ eine Umgebung von x und folglich gibt es ein $n \geq 1$, so dass

$$T^{-1}(T^{-n}(U) \cap U) = T^{-n}(T^{-1}(U)) \cap T^{-1}(U) \neq \varnothing \; .$$

Also ist $T^{-n}(U) \cap U \neq \varnothing$ und dies zeigt, dass $T(x) \in \Omega(T)$.

(4) Hier gilt $T^{-n}(U) \cap U = T^{-n}(T^n(U) \cap U)$ für jedes $U \subset X$ und daraus folgt, dass $\Omega(T^{-1}) = \Omega(T)$. Nach (3) ist $T(\Omega(T)) \subset \Omega(T)$ und auch

$$T^{-1}(\Omega(T)) = T^{-1}(\Omega(T^{-1})) \subset \Omega(T^{-1}) = \Omega(T) \; ,$$

und damit ist $T(\Omega(T)) = \Omega(T)$. □

Lemma 10.4 *Es gilt* $\Omega(T) = \Omega'(T)$, *wobei*

$$\Omega'(T) = \{x \in X : \text{zu jeder Umgebung } U \text{ von } x \text{ und jedem } m \geq 1 \text{ gibt es ein } n \geq m \text{ mit } T^{-n}(U) \cap U \neq \varnothing\} \;;$$

Beweis Es ist klar, dass $\Omega'(T) \subset \Omega(T)$. Sei umgekehrt $x \in \Omega(T)$, sei U eine Umgebung von x und $m \geq 1$. Nehme zunächst an, dass x nicht periodisch ist; dann ist $T^n(x) \neq x$ für alle $n \geq 1$ und folglich gibt es eine Umgebung V von x mit $V \subset U$ und $T^{-k}(V) \cap V = \varnothing$ für $k = 1, \ldots, m-1$. Da $x \in \Omega(T)$, gibt es $n \geq 1$ mit $T^{-n}(V) \cap V \neq \varnothing$ und natürlich muss $n \geq m$ sein, und es gilt $T^{-n}(U) \cap U \neq \varnothing$. Also ist $x \in \Omega'(T)$. Ist dagegen $x \in \mathrm{Per}(p, T)$, so ist $T^{-np}(U) \cap U \neq \varnothing$ für jedes $n \geq 1$, und wieder ist $x \in \Omega'(T)$. □

Satz 10.9 *(1) Es gilt* $\Omega(T) = X$, *wenn* T *surjektiv und einseitig topologisch transitiv ist.*

(2) Ist T *ein Homeomorphismus, so ist* T *einseitig topologisch transitiv genau dann, wenn* T *topologisch transitiv ist und* $\Omega(T) = X$.

Beweis (1) Sei T surjektiv und einseitig topologisch transitiv, und sei $y \in X$. Nun gibt es ein $x \in X$ mit $\overline{\{T^n(x) : n \geq 0\}} = X$, und da $T^m(X) = X$, ist auch

$$\overline{\{T^n(x) : n \geq m\}} = T^m(\overline{\{T^n(x) : n \geq 0\}}) = X$$

für jedes $m \geq 0$. Es gibt also eine Teilfolge $\{n_k\}_{k \geq 1}$, so dass $\lim_{k\to\infty} T^{n_k}(x) = y$ und genauso wie im Beweis für Satz 10.8 (1) ist dann $y \in \Omega(T)$. Dies zeigt, dass $\Omega(T) = X$.

(2) Ist T einseitig topologisch transitiv, dann ist nach (1) $\Omega(T) = X$ und T ist offensichtlich topologisch transitiv. Sei umgekehrt T topologisch transitiv mit $\Omega(T) = X$. Seien U, V nichtleere offene Teilmengen von X. Nach Satz 10.4 ist $\bigcup_{n\in\mathbb{Z}} T^{-n}(U)$ eine dichte Teilmenge von X und folglich gibt es $k \in \mathbb{Z}$ mit $T^{-k}(U) \cap V \neq \varnothing$. Da $\Omega(T) = X$, folgt aus Lemma 10.4, dass es ein $m \geq |k|$ mit $T^{-m}(T^{-k}(U) \cap V) \cap (T^{-k}(U) \cap V) \neq \varnothing$ gibt. Da $m+k \geq 0$, ergibt sich daraus, dass $\bigcup_{n\geq 0} T^{-n}(U) \cap V \neq \varnothing$, Dies zeigt, dass $\bigcup_{n\geq 0} T^{-n}(U)$ eine dichte Teilmenge von X ist für jede nichtleere offene Teilmenge U von X, und ist damit nach Satz 10.3 einseitig topologisch transitiv. □

Für $k = 1, 2$ sei X_k ein kompakter metrischer Raum und sei $T_k : X_k \to X_k$ eine stetige Abbildung. Dann heißen (X_1, T_1) und (X_2, T_2) *konjugiert*, wenn es einen Homeomorphismus $\varphi : X_1 \to X_2$ mit $T_2 \circ \varphi = \varphi \circ T_1$ gibt. Ferner heißt (X_2, T_2) ein *Faktor* von (X_1, T_1), wenn es eine stetige surjektive Abbildung $\varphi : X_1 \to X_2$ gibt, so dass $T_2 \circ \varphi = \varphi \circ T_1$.

Sind (X_1, T_1) und (X_2, T_2) konjugiert, so ist (X_2, T_2) ein Faktor von (X_1, T_1) und (X_1, T_1) ein Faktor von (X_2, T_2).

Satz 10.10 *Sei (X_2, T_2) ein Faktor von (X_1, T_1). Ist T_1 minimal bzw. einseitig topologisch transitiv, so ist auch T_2 minimal bzw. einseitig topologisch transitiv. Sind T_1 und T_2 zusätzlich Homeomorphismen und ist T_1 topologisch transitiv, so ist auch T_2 topologisch transitiv.*

Beweis Sei $\varphi : X_1 \to X_2$ eine stetige surjektive Abbildung mit $T_2 \circ \varphi = \varphi \circ T_1$; für jedes $n \geq 0$ ist also $T_2^n \circ \varphi = \varphi \circ T_1^n$. Sei $x \in X_1$ mit $\overline{\{T_1^n(x) : n \geq 0\}} = X_1$; dann gilt

$$\begin{aligned}\overline{\{T_2^n(\varphi(x)) : n \geq 0\}} = \overline{\{\varphi(T_1^n(x)) : n \geq 0\}} &= \overline{\varphi(\{T_1^n(x) : n \geq 0\})} \\ &= \varphi(\overline{\{T_1^n(x) : n \geq 0\}}) = \varphi(X_1) = X_2\end{aligned}$$

und daraus ergibt sich, dass T_2 minimal bzw. einseitig topologisch transitiv ist, wenn T_1 minimal bzw. einseitig topologisch transitiv ist. Sind T_1 und T_2 Homeomorphismen, so gilt auch $T_2^{-n} \circ \varphi = \varphi \circ T_1^{-n}$ für jedes $n \geq 0$ und folglich ist dann $\overline{\{T_2^n(\varphi(x)) : n \in \mathbb{Z}\}} = X_2$, wenn $\overline{\{T_1^n(x) : n \in \mathbb{Z}\}} = X_1$. Damit ist T_2 topologisch transitiv, wenn T_1 topologisch transitiv ist. □

Die Abbildung $T : X \to X$ heißt *Isometrie* (bezüglich der Metrik d), wenn $d(T(x), T(y)) = d(x, y)$ für alle x, $y \in X$. Eine Isometrie ist automatisch ein Homeomorphismus. (Der Beweis dafür ist eine Übung.)

Die Abbildung $T : X \to X$ heißt *isometrisierbar*, wenn es eine zu d äquivalente Metrik d' auf X gibt, so dass T eine Isometrie bezüglich d' ist.

Satz 10.11 *Sei T ein Homeomorphismus. Dann sind äquivalent:*

(1) T ist topologisch transitiv und isometrisierbar.

(2) T ist minimal und isometrisierbar.

(3) Es gibt eine kompakte abelsche metrische Gruppe G und ein Element $a \in G$ mit $\overline{\{a^n : n \geq 0\}} = G$, so dass (X, T) und (G, U) konjugiert sind, wobei $U(z) = az$ für jedes $z \in G$.

Beweis Übung. □

Sei E eine nichtleere endliche Menge und I entweder $\mathbb{N}$ oder $\mathbb{Z}$; sei $m \geq 2$ und sei B eine Teilmenge von E^m. Definiere $Y_B \subset E_\infty$ durch

$$Y_B = \{\{z_n\}_{n \in I} \in E_\infty : (z_n, \ldots, z_{n+m-1}) \in B \text{ für alle } n \in I\} ;$$

es ist klar, dass $S(Y_B) \subset Y_B$ und nach Lemma 10.1 ist $S(Y_B)$ abgeschlossen. Folglich ist (Y_B, S_{Y_B}) ein Subshift. Sei nun Y eine abgeschlossene Teilmenge von E_∞ mit $S(Y) \subset Y$; dann heißt das Subshift (Y, S_Y) ein Subshift von *endlichem Typ*, wenn $Y = Y_B$ für ein $B \subset E^m$, $m \geq 2$. Jedes Markov-Subshift ist ein Subshift von endlichem Typ. (Ist (Y, S_Y) ein Markov-Subshift mit Übergangsmatrix M, so ist $Y = Y_B$, wobei $B = \{(z_1, z_2) \in E^2 : M(z_1, z_2) = 1\}$.)

Satz 10.12 *Sei $B \subset E^m$ mit $m \geq 2$ und sei (Y, S_Y) das Subshift von endlichem Typ mit $Y = Y_B$. Dann ist (Y, S_Y) zu dem Markov-Subshift mit Übergangsmatrix $M : B \times B \to \{0, 1\}$ konjugiert, wobei*

$$M((z_1, \ldots, z_m), (w_1, \ldots, w_m)) = \begin{cases} 1 & \text{falls } w_k = z_{k+1} \text{ für } k = 1, \ldots, m-1 , \\ 0 & \text{sonst} . \end{cases}$$

Beweis Übung. □

11 Expansive Abbildungen

Im Folgenden sei (X, d) ein kompakter metrischer Raum mit $X \neq \varnothing$ und sei $T : X \to X$ eine stetige Abbildung bzw. ein Homeomorphismus. Sei $I = \mathbb{N}$ bzw. $I = \mathbb{Z}$, wenn T eine stetige Abbildung bzw. ein Homeomorphismus ist. Eine Zahl $\delta > 0$ heißt *expansive Konstante* für T, wenn es für alle $x, y \in X$ mit $x \neq y$ ein $n \in I$ gibt, so dass $d(T^n(x), T^n(y)) > \delta$. Gibt es eine expansive Konstante für T, so heißt T eine *expansive Abbildung* bzw. ein *expansiver Homeomorphismus*.

Beispiele:

(1) Sei E eine nichtleere endliche Menge und betrachte $E_\infty = \mathrm{Abb}(I, E)$ als metrischen Raum mit der Metrik

$$\varrho_\infty(\{x_n\}_{n\in I}, \{y_n\}_{n\in I}) = \sum_{n\in I} 2^{-|n|}\varrho(x_n, y_n)$$

für alle $\{x_n\}_{n\in I}$, $\{y_n\}_{n\in I} \in E_\infty$, wobei $\varrho(u, u) = 0$ für alle $u \in E$ und $\varrho(u, v) = 1$ für alle $u \neq v$. Dann ist $S : E_\infty \to E_\infty$ eine expansive Abbildung ($I = \mathbb{N}$) bzw. ein expansiver Homeomorphismus ($I = \mathbb{Z}$): Seien $\{x_n\}_{n\in I} \neq \{y_n\}_{n\in I}$; dann gibt es $m \in I$ mit $x_m \neq y_m$ und damit ist

$$\varrho_\infty(S^m(\{x_n\}_{n\in I}), S^m(\{y_n\}_{n\in I})) = \sum_{n\in I} 2^{-|n|}\varrho(x_{n+m}, y_{n+m}) \geq 1 \; .$$

(2) Sei Y eine nichtleere abgeschlossene Teilmenge von X mit $T(Y) \subset Y$ bzw. mit $T(Y) = Y$. Ist $T : X \to X$ eine expansive Abbildung bzw. ein expansiver Homeomorphismus, dann ist es auch $T_Y : Y \to Y$. Insbesondere ist jedes Markov-Subshift eine expansive Abbildung bzw. ein expansiver Homeomorphismus.

(3) Sei $A \in \mathrm{M}(n \times n, \mathbb{Z})$ mit $|\det(A)| = 1$. Nach Satz 6.5 ist $\theta_A : \mathrm{K}^n \to \mathrm{K}^n$ ein Automorphismus und θ_A ist ein expansiver Homeomorphismus genau dann, wenn A keinen Eigenwert λ mit $|\lambda| = 1$ besitzt. (Der Beweis dafür ist eine Übung.)

(4) Sei $A \in \mathrm{M}(n \times n, \mathbb{Z})$ mit $\det(A) \neq 0$. Nach Satz 6.5 ist $\theta_A : \mathrm{K}^n \to \mathrm{K}^n$ ein surjektiver Endomorphismus und θ_A ist eine expansive Abbildung genau dann, wenn $|\lambda| > 1$ für jeden Eigenwert λ von A. (Der Beweis dafür ist eine Übung.)

Für jedes $B \subset X$ sei $\mathrm{diam}(B) = \sup\{d(x, y) : x, y \in B\}$.

Eine *endliche Überdeckung* von X ist eine endliche Teilmenge $\mathcal{C}$ von $\mathcal{P}(X)$ mit $X = \bigcup_{C\in\mathcal{C}} C$; in diesem Fall wird $\mathrm{diam}(\mathcal{C}) = \max\{\mathrm{diam}(C) : C \in \mathcal{C}\}$ gesetzt. Eine endliche Überdeckung $\mathcal{C}$ von X heißt *offen* bzw. *abgeschlossen*, wenn jedes Element in $\mathcal{C}$ eine offene bzw. abgeschlossene Teilmenge von X ist.

Lemma 11.1 *Sei $\mathcal{C}$ eine endliche offene Überdeckung von X. Dann gibt es ein $\delta > 0$, so dass es für jedes $B \subset X$ mit $\mathrm{diam}(B) \leq \delta$ ein $U \in \mathcal{C}$ mit $B \subset U$ gibt.*

Beweis Für jedes $x \in X$ gibt es $\delta_x > 0$, so dass $\mathrm{B}(x, \delta_x) \subset U$ für ein $U \in \mathcal{C}$. Dann ist $\{\mathrm{B}(x, \delta_x/2)\}_{x \in X}$ eine offene Überdeckung von X und folglich gibt es eine endliche Menge $\Delta \subset X$ mit $X = \bigcup_{x \in \Delta} \mathrm{B}(x, \delta_x/2)$; setze $\delta = \min\{\delta_x/2 : x \in \Delta\}$. Sei $B \subset X$ mit $\mathrm{diam}(B) \leq \delta$; dann gibt es $x \in \Delta$, so dass $B \cap \mathrm{B}(x, \delta_x/2) \neq \varnothing$, also ist $B \subset \mathrm{B}(x, \delta_x)$ und damit ist $B \subset U$ für ein $U \in \mathcal{C}$. $\square$

Lemma 11.1 ist eine Version des Lebesgueschen Überdeckungslemmas und daher heißt die Zahl δ eine *Lebesgue-Konstante* für die Überdeckung $\mathcal{C}$.

Sei $\mathcal{C}$ eine endliche offene Überdeckung von X. Dann heißt $\mathcal{C}$ *Generator* für T, wenn für jede Folge $\{A_n\}_{n \in I}$ aus $\mathcal{C}$ die Menge $\bigcap_{n \in I} T^{-n}(\bar{A}_n)$ aus höchstens einem Element besteht. Ferner heißt $\mathcal{C}$ *schwacher Generator* für T, wenn für jede Folge $\{A_n\}_{n \in I}$ aus $\mathcal{C}$ die Menge $\bigcap_{n \in I} T^{-n}(A_n)$ aus höchstens einem Element besteht.

Satz 11.1 *Äquivalent sind:*

(1) Es gibt einen schwachen Generator für T.

(2) Es gibt einen Generator für T.

(3) T ist eine expansive Abbildung bzw. ein expansiver Homeomorphismus.

Beweis (2) $\Rightarrow$ (1) ist klar.

(1) $\Rightarrow$ (3): Sei $\mathcal{C}$ ein schwacher Generator für T und sei $\delta > 0$ eine Lebesgue-Konstante für $\mathcal{C}$. Seien nun $x, y \in X$ mit $d(T^n(x), T^n(y)) \leq \delta$ für alle $n \in I$; für jedes $n \in I$ gibt es dann $A_n \in \mathcal{C}$, so dass $T^n(x)$ und $T^n(y)$ beide in A_n liegen. Dann aber liegen x und y beide in $\bigcap_{n \in I} T^{-n}(A_n)$ und damit ist $x = y$, da $\mathcal{C}$ ein schwacher Generator für T ist. Dies zeigt, dass δ eine expansive Konstante für T ist und folglich ist T eine expansive Abbildung bzw. ein expansiver Homeomorphismus.

(3) $\Rightarrow$ (2): Sei $\delta > 0$ eine expansive Konstante für T, und sei $\mathcal{C}$ eine endliche offene Überdeckung von X mit $\mathrm{diam}(A) \leq \delta$ und daher auch mit $\mathrm{diam}(\bar{A}) \leq \delta$ für jedes $A \in \mathcal{C}$. Sei $\{A_n\}_{n \in I}$ eine Folge aus $\mathcal{C}$ und seien $x, y \in \bigcap_{n \in I} T^{-n}(\bar{A}_n)$. Dann liegen $T^n(x)$ und $T^n(y)$ beide in $\bar{A}_n$ und damit ist $d(T^n(x), T^n(y)) \leq \delta$ für jedes $n \in I$. Folglich ist $x = y$, da δ eine expansive Konstante für T ist, und dies zeigt, dass $\mathcal{C}$ ein Generator für T ist. $\square$

Satz 11.1 zeigt insbesondere, dass es nicht von der Metrik sondern nur von der Topologie auf X abhängt, ob T expansiv ist.

Satz 11.2 *Sei $n \in \mathbb{N} \setminus \{0\}$ bzw. $n \in \mathbb{Z} \setminus \{0\}$; dann ist T eine expansive Abbildung bzw. ein expansiver Homeomorphismus genau, wenn T^n es ist.*

Beweis Ist T^n expansiv, dann ist es klar, dass T expansiv ist. Ist ferner T ein Homeomorphismus, so ist es auch klar, dass T genau dann expansiv ist, wenn T^{-1} es ist. Sei also T eine expansive Abbildung bzw. ein expansiver Homeomorphismus und sei $n \ge 1$. Nach Satz 11.1 gibt es einen schwachen Generator $\mathcal{C}$ für T; setze

$$\mathcal{C}^{(n)} = \Big\{ \bigcap_{k=0}^{n-1} T^{-k}(A_k) : A_0, \ldots, A_{n-1} \in \mathcal{C} \Big\} ;$$

dann ist $\mathcal{C}^{(n)}$ eine endliche offene Überdeckung von X. Sei $\{B_m\}_{m\in I}$ eine Folge aus $\mathcal{C}^{(n)}$; für jedes $m \in I$ gibt es $A_0^{(m)}, \ldots, A_{n-1}^{(m)} \in \mathcal{C}$ mit $B_m = \bigcap_{k=0}^{n-1} T^{-k}(A_k^{(m)})$ und folglich besteht die Menge

$$\bigcap_{m\in I} T^{-nm}(B_m) = \bigcap_{m\in I} T^{-nm}\Big(\bigcap_{k=0}^{n-1} T^{-k}(A_k^{(m)})\Big) = \bigcap_{m\in I}\bigcap_{k=0}^{n-1} T^{-nm-k}(A_k^{(m)})$$

aus höchstens einem Element, da $\mathcal{C}$ ein schwacher Generator für T ist. Damit ist $\mathcal{C}^{(n)}$ ein schwacher Generator für T^n und daraus ergibt sich nach Satz 11.1, dass T^n expansiv ist. $\square$

Für eine endliche Überdeckung $\mathcal{D}$ von X und J eine endliche Teilmenge von I sei $\bigvee_{n\in J} T^{-n}(\mathcal{D})$ die endliche Überdeckung von X bestehend aus allen Teilmengen von X der Form $\bigcap_{n\in J} T^{-n}(A_n)$ mit $A_n \in \mathcal{D}$ für jedes $n \in J$. Für jedes $m \ge 1$ setze $I_m = \{0, \ldots, m-1\}$ bzw. $I_m = \{-m, \ldots, m-1\}$.

Lemma 11.2 *Sei T expansiv, sei $\delta > 0$ eine expansive Konstante für T und sei $\mathcal{D}$ eine endliche Überdeckung von X mit* $\mathrm{diam}(\mathcal{D}) \le \delta$. *Dann gibt es zu jedem $\varepsilon > 0$ ein $m \ge 1$, so dass* $\mathrm{diam}(\bigvee_{n\in I_m} T^{-n}(\mathcal{D})) < \varepsilon$.

Beweis Für jedes $m \ge 1$ und jedes $B \in \bigvee_{n\in I_m} T^{-n}(\mathcal{D})$ gilt $d(T^n(u), T^n(v)) < \delta$ für alle u, $v \in B$, $n \in I_m$. Sei nun $\varepsilon > 0$ und nehme an, dass es für jedes $m \ge 1$ ein $W_m \in \bigvee_{n\in I_m} T^{-n}(\mathcal{D})$ mit $\mathrm{diam}(W_n) \ge \varepsilon$ gibt. Für jedes $m \ge 1$ gibt es also x_m, $y_m \in W_m$ mit $d(x_m, y_m) \ge \varepsilon/2$, und da X kompakt ist, gibt es dann x, $y \in X$ und eine Teilfolge $\{m_k\}_{k\ge1}$, so dass $x = \lim_{k\to\infty} x_{m_k}$ und $y = \lim_{k\to\infty} y_{m_k}$, und insbesondere ist $d(x, y) \ge \varepsilon/2$. Aber $d(T^n(x_m), T^n(y_m)) \le \delta$ für alle $n \in I_m$, $m \ge 1$, und daraus ergibt sich, dass $d(T^n(x), T^n(y)) \le \delta$ für alle $n \in I$. Damit ist $x = y$, da δ eine expansive Konstante für T ist. Dies ist ein Widerspruch, und folglich gibt es zu jedem $\varepsilon > 0$ ein $m \ge 1$, so dass $\mathrm{diam}(\bigvee_{n\in I_m} T^{-n}(\mathcal{D})) < \varepsilon$. $\square$

Für $B \subset X$ bezeichnet B^o das Innere der Menge B und ∂B den Rand von B (also ist $\partial B = \bar{B} \cap \overline{X \setminus B} = \bar{B} \setminus B^o$). Für eine endliche Überdeckug $\mathcal{C}$ von X setze $\partial\mathcal{C} = \bigcup_{A\in\mathcal{C}} \partial A$.

Lemma 11.3 *Für abgeschlossene Mengen $F_1, \ldots, F_n \subset X$ gilt* $\Big(\bigcup_{k=1}^{n} \partial F_k\Big)^o = \varnothing$.

Beweis Für jede abgeschlossene Teilmenge F von X ist $\partial F = F \setminus F^o$ und damit ist $(\partial F)^o = \varnothing$, da $(\partial F)^o \subset (F \setminus F^o)^o \subset F^o$; folglich ist $X \setminus \partial F$ eine dichte offene Menge. Daraus ergibt sich nach Lemma 10.2, dass

$$\Big(\bigcup_{k=1}^{n} \partial F_k\Big)^o = X \setminus \overline{\bigcap_{k=1}^{n}(X \setminus \partial F_k)} = X \setminus X = \varnothing\ . \ \square$$

Lemma 11.4 *Für jede endliche abgeschlossene Überdeckug $\mathcal{C}$ von X ist $X \setminus \partial\mathcal{C}$ eine dichte offene Teilmenge von X. Ferner gilt $X = \partial\mathcal{C} \cup \bigcup_{A\in\mathcal{C}} A^o$.*

Beweis Die erste Aussage folgt unmittelbar aus Lemma 11.3. Ferner ist

$$X = \bigcup_{A\in\mathcal{C}} A = \bigcup_{A\in\mathcal{C}}(A \setminus A^o) \cup \bigcup_{A\in\mathcal{C}} A^o = \bigcup_{A\in\mathcal{C}} \partial A \cup \bigcup_{A\in\mathcal{C}} A^o = \partial\mathcal{C} \cup \bigcup_{A\in\mathcal{C}} A^o\ . \ \square$$

Lemma 11.5 *Zu jedem $\delta > 0$ gibt es eine endliche abgeschlossene Überdeckung $\mathcal{C}$ von X, so dass* $\mathrm{diam}(\mathcal{C}) \le \delta$ *und $A_1^o \cap A_2^o = \varnothing$ für alle $A_1,\ A_2 \in \mathcal{C}$ mit $A_1 \ne A_2$.*

Beweis Da der kompakte metrische Raum (X, d) total beschränkt ist, gibt es $x_1, \ldots, x_n \in X$, so dass $X = \bigcup_{k=1}^{n} \mathrm{B}(x_k, \delta/2)$. Setze $D_1 = \overline{\mathrm{B}(x_1, \delta/2)}$ und sei

$$D_k = \overline{\mathrm{B}(x_k, \delta/2)} \setminus \bigcup_{j=1}^{k-1} \mathrm{B}(x_j, \delta/2)$$

für $2 \le k \le n$. Dann ist $\mathcal{C} = \{D_k : 1 \le k \le n\}$ eine endliche abgeschlossene Überdeckung von X mit $\mathrm{diam}(\mathcal{C}) \le \delta$ und man sieht leicht, dass $D_j^o \cap D_k^o = \varnothing$ für alle $1 \le j < k \le n$. $\square$

Satz 11.3 *Sei T ein expansiver Homeomorphismus. Dann gibt es eine nichtleere endliche Menge E und eine abgeschlossene Teilmenge Y von $E_\infty =$* $\mathrm{Abb}(\mathbb{Z}, E)$ *mit $S(Y) = Y$, so dass (X, T) ein Faktor von (Y, S_Y) ist.*

Beweis Sei $\delta > 0$ eine expansive Konstante für T. Nach Lemma 11.5 gibt es eine endliche abgeschlossene Überdeckung E von X, so dass $\mathrm{diam}(\mathsf{E}) \le \delta$ und $C^o \cap D^o = \varnothing$ für alle $C,\ D \in \mathsf{E}$ mit $C \ne D$. Setze $\Delta = \partial\mathsf{E}$; nach Lemma 11.4 ist $X \setminus \Delta$ eine dichte offene Teilmenge von X und für jedes $x \in X \setminus \Delta$ gibt es ein eindeutiges $D \in \mathsf{E}$ mit $x \in D$. Da T ein Homeomorphismus ist, ist $T^{-n}(X \setminus \Delta)$ auch eine dichte offene Teilmenge für jedes $n \in \mathbb{Z}$. Sei $\Delta^\infty = \bigcup_{n\in\mathbb{Z}} T^{-n}(\Delta)$; da

$$X \setminus \Delta^\infty = \bigcap_{n\in\mathbb{Z}}(X \setminus T^{-n}(\Delta)) = \bigcap_{n\in\mathbb{Z}} T^{-n}(X \setminus \Delta)\ ,$$

folgt aus dem Satz von Baire (Lemma 10.3), dass $X \setminus \Delta^\infty$ eine dichte G_δ-Menge ist. Sei $x \in X \setminus \Delta^\infty$; für jedes $n \in \mathbb{Z}$ ist dann $T^n(x) \in X \setminus \Delta$ und folglich gibt es ein eindeutiges $D_n(x) \in \mathsf{E}$ mit $T^n(x) \in D_n(x)$. Definiere eine Abbildung $\psi : X \setminus \Delta^\infty \to \mathsf{E}_\infty = \mathrm{Abb}(\mathbb{Z}, \mathsf{E})$ durch $\psi(x) = \{D_n(x)\}_{n\in I}$.

Seien $x, y \in X \setminus \Delta^\infty$ mit $\psi(x) = \psi(y)$; dann liegen $T^n(x)$ und $T^n(y)$ beide in $D_n(x)$ und damit ist $d(T^n(x), T^n(y)) \le \delta$ für alle $n \in \mathbb{Z}$. Daraus ergibt sich, dass $x = y$, da δ eine expansive Konstante für T ist, und dies zeigt, dass ψ injektiv ist. Setze $Y_0 = \psi(X \setminus \Delta^\infty)$ und sei $\varphi : Y_0 \to X$ die eindeutige Abbildung mit $\varphi(\psi(x)) = x$ für alle $x \in X \setminus \Delta^\infty$; für jedes $z = \{D_n\}_{n\in I} \in Y_0$ ist $\varphi(z)$ das eindeutige Element in $X \setminus \Delta^\infty$ mit $T^n(\varphi(z)) \in D_n$ für alle $n \in I$. Natürlich ist $\varphi(Y_0) = X \setminus \Delta^\infty$ und $S(Y_0) = Y_0$, da $T(X \setminus \Delta^\infty) = X \setminus \Delta^\infty$. Ferner gilt $T(\varphi(z)) = \varphi(S(z))$ für alle $z \in Y_0$.

Es wird nun gezeigt, dass φ gleichmäßig stetig ist. Sei $\varepsilon > 0$; nach Lemma 11.2 gibt es dann ein $m \ge 1$, so dass $\mathrm{diam}(B) < \varepsilon$ für jedes $B \in \bigvee_{n\in I_m} T^{-n}(\mathsf{E})$, wobei $I_m = \{-m, \ldots, m-1\}$. Seien $z = \{D_n\}_{n\in\mathbb{Z}}$, $z' = \{D'_n\}_{n\in\mathbb{Z}} \in Y_0$ mit $\varrho_\infty(z, z') < 2^{-m}$; folglich gilt $D_n = D'_n$ für jedes $n \in I_m$ und daher liegen $\varphi(z)$ und $\varphi(z')$ beide in $\bigcap_{n\in I_m} T^{-n}(D_n)$. Also ist $d(\varphi(z), \varphi(z')) < \varepsilon$.

Da $\varphi : Y_0 \to X$ gleichmäßig stetig ist, läßt sich φ eindeutig zu einer stetigen Abbildung $\varphi : \bar{Y}_0 \to X$ fortsetzen. Setze $Y = \bar{Y}_0$; dann ist Y abgeschlossen und $S(Y) = S(\bar{Y}_0) = \overline{S(Y_0)} = \bar{Y}_0 = Y$. Da $T(\varphi(z)) = \varphi(S(z))$ für alle $z \in Y_0$, gilt auch $T(\varphi(z)) = \varphi(S(z))$ für alle $z \in Y$, d.h., $T \circ \varphi = \varphi \circ S_Y$. Da schließlich Y abgeschlossene Teilmenge des kompakten Raumes E_∞ ist und φ stetig ist, gilt

$$\varphi(Y) = \overline{\varphi(Y)} \supset \overline{\varphi(Y_0)} = \overline{X \setminus \Delta^\infty} = X \ ,$$

d.h., φ ist surjektiv, und damit ist (X, T) ein Faktor von (Y, S_Y). □

Satz 11.4 *Sei T eine expansive Abbildung und nehme an, dass für jede dichte offene Teilmenge U von X die offene Menge $T^{-1}(U)$ auch dicht ist. Dann gibt es eine nichtleere endliche Menge E und eine abgeschlossene Teilmenge Y von $E_\infty = \mathrm{Abb}(\mathbb{N}, E)$ mit $S(Y) \subset Y$, so dass (X, T) ein Faktor von (Y, S_Y) ist.*

Beweis Dieser ist fast identisch mit dem Beweis für Satz 11.3. □

Sei T eine expansive Abbildung bzw. ein expansiver Homeomorphismus und sei $\mathcal{C}$ eine endliche Unter-σ-Algebra von $\mathcal{B}_X$ mit $\mathrm{diam}(\pi(\mathcal{C})) \le \delta$, wobei δ eine expansive Konstante für T ist. Dann gilt $\bigvee_{n\in I} T^{-n}(\mathcal{C}) = \mathcal{B}_X$. (Der Beweis dafür ist eine Übung.) Ist also T maßtreu bezüglich $\mu \in \mathrm{P}(X, \mathcal{B}_X)$, so folgt aus Satz 8.22, dass $\mathrm{h}(T) = \mathrm{h}(T, \mathcal{C})$.

12 Markov-Partitionen

Im Folgenden sei (X, d) ein kompakter metrischer Raum mit $X \neq \varnothing$ und sei $T : X \to X$ eine stetige Abbildung.

Ist T expansiv und ist die Menge $T^{-1}(U)$ dicht für jede dichte offene Teilmenge U von X, so folgt aus Satz 11.4, dass (X, T) ein Faktor eines Subshifts ist: Es gibt eine nichtleere endliche Menge E und eine abgeschlossene Teilmenge Y von $E_\infty = \mathrm{Abb}(\mathbb{N}, E)$ mit $S(Y) \subset Y$, so dass (X, T) ein Faktor von (Y, S_Y) ist. In diesem Kapitel wird untersucht, unter welchen Bedingungen (X, T) ein Faktor eines Markov-Subshifts sein kann.

Es werden nur stetige Abbildungen (und nicht Homeomorphismen) betrachtet. Also ist hier ein Markov-Subshift stets ein *einseitiges* Markov-Subshift mit $I = \mathbb{N}$.

Eine stetige Abbildung $g : X \to X$ heißt *lokaler Homeomorphismus*, wenn es zu jedem $x \in X$ eine offene Umgebung U von x gibt, so dass die Menge $g(U)$ offen und die Einschränkung $g_{|U} : U \to g(U)$ ein Homeomorphismus ist.

Lemma 12.1 *Sei (Y, S_Y) ein Subshift vom endlichen Typ; dann ist $S_Y : Y \to Y$ ein lokaler Homeomorphismus.*

Beweis Nehme zunächst an, dass (Y, S_Y) ein Markov-Subshift ist mit Übergangsmatrix $M : E \times E \to \{0, 1\}$. Für jedes $u \in E$ sei $E^u_\infty = \{\{z_n\}_{n \geq 0} \in E_\infty : z_0 = u\}$; dann ist E^u_∞ eine offene Teilmenge von E_∞ und S bildet E^u_∞ bijektiv auf E_∞ ab. Da E^u_∞ auch eine kompakte Teilmenge von E_∞ ist, ist die Einschränkung S^u von S auf E^u_∞ ein Homeomorphismus. Setze nun $V_u = Y \cap E^u_\infty$; dann ist V_u eine offene Teilmenge von Y und S_Y bildet V_u bijektiv auf die offene Menge $\bigcup_{w \in M_u} V_w$ ab, wobei $M_u = \{w \in E : M(u, w) = 1\}$. Ferner ist die Einschränkung von S_Y auf V_u (als Einschränkung von S^u) ein Homeomorphismus. Dies zeigt, dass S_Y ein lokaler Homeomorphismus ist, da $\bigcup_{u \in E} V_u = Y$. Ist nun (Y, S_Y) ein Subshift vom endlichen Typ, so ist nach Satz 10.12 (Y, S_Y) zu einem Markov-Subshift konjugiert, und damit ist S_Y auch ein lokaler Homeomorphismus. □

Lemma 12.2 *Sei $A \in \mathrm{M}(n \times n, \mathbb{Z})$ mit $\det(A) \neq 0$; dann ist $\theta_A : \mathrm{K}^n \to \mathrm{K}^n$ ein lokaler Homeomorphismus.*

Beweis Übung. □

Eine *Markov-Überdeckung* für T ist eine endliche abgeschlossene Überdeckung $\mathcal{C}$ von X, für die gilt: Für jedes $C \in \mathcal{C}$ gibt es eine Teilmenge $\tau(C) \subset \mathcal{C}$, so dass $T(C) = \bigcup_{D \in \tau(C)} D$.

Eine endliche abgeschlossene Überdeckung $\mathcal{R}$ von X heißt *Markov-Partition* für T, wenn gilt:

(1) Für jedes $R \in \mathcal{R}$ ist $R \neq \varnothing$ und es gilt $R = \overline{R^o}$.

(2) Für alle $R_1,\ R_2 \in \mathcal{R}$ mit $R_1 \neq R_2$ ist $R_1^o \cap R_2^o = \varnothing$.

(3) Für alle $R_1,\ R_2 \in \mathcal{R}$ mit $R_1^o \cap T(R_2^o) \neq \varnothing$ ist $R_1 \subset T(R_2)$.

Satz 12.1 *Sei T ein lokaler Homeomorphismus, der eine expansive Abbildung ist. Dann sind äquivalent:*

(1) (X, T) ist ein Faktor eines Markov-Subshifts.

(2) Zu jedem $\varepsilon > 0$ gibt es eine Markov-Überdeckung $\mathcal{C}$ für T mit $\operatorname{diam}(\mathcal{C}) < \varepsilon$.

(3) Zu jedem $\varepsilon > 0$ gibt es eine Markov-Partition $\mathcal{R}$ für T mit $\operatorname{diam}(\mathcal{R}) < \varepsilon$.

(4) Es gibt ein Markov-Subshift (Y, S_Y) und eine stetige surjektive Abbildung $\varphi : Y \to X$ mit $T \circ \varphi = \varphi \circ S_Y$, so dass φ injektiv über eine dichte G_δ-Menge in X ist und $|\varphi^{-1}(\{x\})| \leq N$ für alle $x \in X$ für ein $N \geq 1$.

Beweis Dies wird aus Satz 12.2, Satz 12.3 und Satz 12.4 folgen. □

Satz 12.2 *Sei (X, T) ein Faktor eines Markov-Subshifts. Dann gibt es zu jedem $\varepsilon > 0$ eine Markov-Überdeckung $\mathcal{C}$ für T mit $\operatorname{diam}(\mathcal{C}) < \varepsilon$.*

Beweis Es gibt ein Markov-Subshift (Y, S_Y) und eine stetige surjektive Abbildung $\varphi : Y \to X$ mit $T \circ \varphi = \varphi \circ S_Y$; sei $M : E \times E \to \{0, 1\}$ die Übergangsmatrix für (Y, S_Y). Sei $m \geq 1$ und für jedes $\eta \in E^m$ sei $C_m(\eta) = \varphi(Y_m(\eta))$, wobei $Y_m(\eta) = \{\{z_n\}_{n \geq 0} \in Y : (z_0, \ldots, z_{m-1}) = \eta\}$. Dann ist $\mathcal{C}_m = \{C_m(\eta) : \eta \in E^m\}$ eine endliche abgeschlossene Überdeckung von X, und in der Tat ist $\mathcal{C}_m$ eine Markov-Überdeckung für T, da

$$\begin{aligned}
&T(C_m((z_0, \ldots, z_{m-1}))) \\
&\qquad = (T \circ \varphi)(Y_m((z_0, \ldots, z_{m-1}))) = (\varphi \circ S_Y)(Y_m((z_0, \ldots, z_{m-1}))) \\
&\qquad = \varphi\Big(\bigcup_{z \in E} Y_m(z_1, \ldots, z_{m-1}, z)\Big) = \bigcup_{z \in E} D_m(z_1, \ldots, z_{m-1}, z)\ .
\end{aligned}$$

Aber φ ist gleichmäßig stetig und $\operatorname{diam}(Y_m(\eta)) \leq 2^{-m}$ für jedes $\eta \in E^m$. Folglich gilt $\lim_{m \to \infty} \operatorname{diam}(\mathcal{C}_m) = 0$, und also gibt es zu jedem $\varepsilon > 0$ ein $m \geq 1$ mit $\operatorname{diam}(\mathcal{C}_m) < \varepsilon$. □

Lemma 12.3 *Sei $g : X \to X$ ein lokaler Homeomorphismus. Für jedes $A \subset X$ mit $\bar{A} = X$ gilt dann auch $\overline{g^{-1}(A)} = X$.*

Beweis Sei $A \subset X$ mit $\overline{g^{-1}(A)} \neq X$; also ist $V = X \setminus \overline{g^{-1}(A)}$ eine nichtleere offene Menge mit $V \cap g^{-1}(A) = \varnothing$, und da g ein lokaler Homeomorphismus ist, gibt es dann eine offene Menge U mit $U \cap V \neq \varnothing$, so dass $g(U)$ offen und die Einschränkung von g auf U ein Homeomorphismus ist. Folglich ist $g(U \cap V)$ eine nichtleere offene Menge mit $g(U \cap V) \cap A = \varnothing$, und damit ist $\bar{A} \neq X$. □

Lemma 12.4 *Sei $g : X \to X$ ein lokaler Homeomorphismus. Dann gibt es ein $\delta > 0$ mit folgender Eigenschaft: Für jedes $B \subset X$ mit $\operatorname{diam}(B) \leq \delta$ gibt es eine offene Menge U mit $B \subset U$, so dass die Menge $g(U)$ offen und die Einschränkung $g_{|U} : U \to g(U)$ ein Homeomorphismus ist.*

Beweis Da X kompakt ist, gibt es eine endliche offene Überdeckung $\mathcal{U}$ von X, so dass für jedes $U \in \mathcal{U}$ die Menge $g(U)$ offen ist mit der Einschränkung von g auf U ein Homeomorphismus. Sei $\delta > 0$ eine Lebesgue-Konstante für $\mathcal{U}$. (Für jedes $B \subset X$ mit $\operatorname{diam}(B) \leq \delta$ gibt es also ein $U \in \mathcal{U}$ mit $B \subset U$.) Dann hat diese Zahl δ die gewünschte Eigenschaft. □

Die Zahl δ in Lemma 12.4 wird *Lebesgue-Konstante* für g genannt.

Lemma 12.5 *Sei $g : X \to X$ ein lokaler Homeomorphismus und sei δ eine Lebesgue-Konstante für g.*

(1) Für jedes $A \subset X$ mit $\operatorname{diam}(A) \leq \delta$ ist $(g(A))^o = g(A^o)$.

(2) Ist $\mathcal{C}$ eine endliche Überdeckung $\mathcal{C}$ von X mit $\operatorname{diam}(\mathcal{C}) \leq \delta/2$, so ist die Einschränkung von g auf $C_1 \cup C_2$ injektiv für alle $C_1, C_2 \in \mathcal{C}$ mit $C_1 \cap C_2 \neq \varnothing$ (und insbesondere ist die Einschränkung von g auf C injektiv für alle $C \in \mathcal{C}$).

Beweis Dies ist klar. □

Eine stetige Abbildung $g : X \to X$ heißt *offen*, wenn $g(U)$ offen ist für jede offene Teilmenge U von X.

Lemma 12.6 *Jeder lokale Homeomorphismus $g : X \to X$ ist offen.*

Beweis Sei U ein offene Teilmenge von X und sei $\delta > 0$ eine Lebesgue-Konstante für g. Zu jedem $x \in U$ gibt es ein $0 < \varepsilon_x < \delta$ mit $\mathrm{B}(x, \varepsilon_x) \subset U$ und dann ist die Einschränkung von g auf $\mathrm{B}(x, \varepsilon_x)$ ein Homeomorphismus. Damit ist $g(\mathrm{B}(x, \varepsilon_x))$ offen für jedes $x \in U$ und folglich ist $g(U) = \bigcup_{x \in U} g(\mathrm{B}(x, \varepsilon_x))$ auch offen. □

Sind $\mathcal{C}$ und $\mathcal{D}$ endliche Überdeckungen von X, so bedeutet $\mathcal{D} \prec \mathcal{C}$, dass es zu jedem $D \in \mathcal{D}$ ein $C \in \mathcal{C}$ mit $D \subset C$ gibt.

Satz 12.3 *Sei T ein lokaler Homeomorphismus und $\mathcal{C}$ eine Markov-Überdeckung für T mit $\mathrm{diam}(\mathcal{C}) \le \delta/2$, wobei $\delta > 0$ eine Lebesgue-Konstante für T ist. Dann gibt es eine Markov-Partition $\mathcal{R}$ für T mit $\mathcal{R} \prec \mathcal{C}$.*

Beweis Sei Y eine Menge und sei $\mathcal{E}$ eine endliche Überdeckung von Y. Die von $\mathcal{E}$ erzeugte Partition wird mit $\pi(\mathcal{E})$ bezeichnet; also ist $\pi(\mathcal{E})$ die endliche Teilmenge von $\mathcal{P}(Y)$ bestehend aus allen nichtleeren Mengen der Form

$$\bigcap_{F\in\mathcal{F}} F \setminus \bigcup_{E\in\mathcal{E}\setminus\mathcal{F}} E$$

mit $\mathcal{F} \subset \mathcal{E}$. Es gilt $X = \bigcup_{A\in\pi(\mathcal{E})} A$, $A_1 \cap A_2 = \varnothing$ für alle A_1, $A_2 \in \pi(\mathcal{E})$ mit $A_1 \ne A_2$ und jedes Element in $\mathcal{E}$ hat eine eindeutige Darstellung der Form $\bigcup_{A\in\mathcal{A}} A$ mit $\mathcal{A} \subset \pi(\mathcal{E})$. Sei $U : Y \to Y$ eine Abbildung; eine endliche Überdeckung $\mathcal{E}$ von Y wird nun *Markov-Überdeckung* für U genannt, wenn es für jedes $E \in \mathcal{E}$ eine Teilmenge $\tau(E) \subset \mathcal{E}$ gibt, so dass $U(E) = \bigcup_{F\in\tau(E)} F$. (Dies ist die gleiche Definition wie vorher aber jetzt ohne die Topologie.) Eine endliche Überdeckung $\mathcal{C}$ von Y wird *U-regulär* genannt, wenn für alle C_1, $C_2 \in \mathcal{C}$ mit $C_1 \cap C_2 \ne \varnothing$ die Einschränkung von U auf $C_1 \cup C_2$ injektiv ist.

Lemma 12.7 *Sei $U : Y \to Y$ eine Abbildung und $\mathcal{E}$ eine U-reguläre Markov-Überdeckung für U. Dann ist $\pi(\mathcal{E})$ auch eine Markov-Überdeckung für U.*

Beweis Sei $A \in \pi(\mathcal{E})$; dann gibt es $\mathcal{F} \subset \mathcal{E}$, so dass

$$A = \bigcap_{F\in\mathcal{F}} F \setminus \bigcup_{E\in\mathcal{E}\setminus\mathcal{F}} E = \bigcap_{F\in\mathcal{F}} F \setminus \bigcup_{E\in\mathcal{G}} E\,,$$

wobei $\mathcal{G} = \{E \in \mathcal{E} \setminus \mathcal{F} : E \cap \bigcap_{F\in\mathcal{F}} F \ne \varnothing\}$. Man beachte: Für alle G_1, $G_2 \subset Y$ gilt $U(G_1) \setminus U(G_2) \subset U(G_1 \setminus G_2)$, wenn aber $U(G_1) \setminus U(G_2) \ne U(G_1 \setminus G_2)$, so gibt es $y \in G_1 \setminus G_2$ und $z \in G_2$ mit $U(y) = U(z)$. Da nun $\mathcal{E}$ U-regulär ist, ist

$$\begin{aligned} U(A) &= U\Big(\bigcap_{F\in\mathcal{F}} F \setminus \bigcup_{E\in\mathcal{G}} E\Big) = U\Big(\bigcap_{F\in\mathcal{F}} F\Big) \setminus U\Big(\bigcup_{E\in\mathcal{G}} E\Big) \\ &= \bigcap_{F\in\mathcal{F}} U(F) \setminus \bigcup_{E\in\mathcal{G}} U(E) = \bigcap_{F\in\mathcal{F}} \bigcup_{C\in\tau(F)} C \setminus \bigcup_{E\in\mathcal{G}} \bigcup_{D\in\tau(E)} D\,. \end{aligned}$$

(Es gilt $U(\bigcup_{E\in\mathcal{G}} E) = \bigcup_{E\in\mathcal{G}} U(E)$ immer, und $U(\bigcap_{F\in\mathcal{F}} F) = \bigcap_{F\in\mathcal{F}} U(F)$ gilt, da $\varnothing \ne A \subset \bigcap_{F\in\mathcal{F}} F$ und $\mathcal{E}$ U-regulär ist.) Also ist $U(A)$ in der Form $\bigcup_{B\in\mathcal{D}} B$ mit $\mathcal{D} \subset \pi(\mathcal{E})$, und dies zeigt, dass $\pi(\mathcal{E})$ eine Markov-Überdeckung für U ist. □

Sei nun $\mathcal{C}$ eine Markov-Überdeckung für T mit $\mathrm{diam}(\mathcal{C}) \le \delta/2$. Dann ist nach Lemma 12.5 (2) $\mathcal{C}$ T-regulär und folglich ist nach Lemma 12.7 $\pi(\mathcal{C})$ eine Markov-Überdeckung von T (aber nur im zweiten Sinn, da im Allgemeinen die Elemente von $\pi(\mathcal{C})$ nicht abgeschlossen sein werden). Für alle A_1, $A_2 \in \pi(\mathcal{C})$ mit $A_1 \ne A_2$ ist $A_1 \cap A_2 = \varnothing$ und damit insbesondere $A_1^o \cap A_2^o = \varnothing$.

Lemma 12.8 *Es gilt* $A \setminus A^o \subset \partial\mathcal{C}$ *und* $\overline{A^o} \subset \partial\mathcal{C} \cup A^o$ *für jedes* $A \in \pi(\mathcal{C})$*. Ferner ist* $X = \partial\mathcal{C} \cup \bigcup_{A\in\pi(\mathcal{C})} A^o$*.*

Beweis Sei $A \in \pi(\mathcal{C})$; dann gibt es $\mathcal{D} \subset \mathcal{C}$, so dass

$$A = \bigcap_{D\in\mathcal{D}} D \setminus \bigcup_{E\in\mathcal{C}\setminus\mathcal{D}} E = \bigcap_{D\in\mathcal{D}} D \cap \bigcap_{E\in\mathcal{C}\setminus\mathcal{D}} (X \setminus E)\,,$$

und daraus ergibt sich, dass

$$A \setminus A^o = \Big(\bigcap_{D\in\mathcal{D}} D \setminus \bigcap_{D\in\mathcal{D}} D^o\Big) \cap \bigcap_{E\in\mathcal{C}\setminus\mathcal{D}} (X \setminus E) \subset \bigcup_{D\in\mathcal{D}} (D \setminus D^o) = \bigcup_{D\in\mathcal{D}} \partial D \subset \partial\mathcal{C}\,.$$

Folglich gilt dann auch

$$X = \bigcup_{A\in\pi(\mathcal{C})} A = \bigcup_{A\in\pi(\mathcal{C})} (A \setminus A^o) \cup \bigcup_{A\in\pi(\mathcal{C})} A^o \subset \partial\mathcal{C} \cup \bigcup_{A\in\pi(\mathcal{C})} A^o\,.$$

Für $B \in \pi(\mathcal{C})$ mit $B \neq A$ ist $A^o \cap B^o = \varnothing$ und damit $\overline{A^o} \cap B^o = \varnothing$. Daher gilt

$$\overline{A^o} = \overline{A^o} \cap \Big(\partial\mathcal{C} \cup \bigcup_{B\in\pi(\mathcal{C})} B^o\Big) = (\overline{A^o} \cap \partial\mathcal{C}) \cup A^o \subset \partial\mathcal{C} \cup A^o\,. \quad \square$$

Sei nun $\mathcal{R} = \{\overline{A^o} : A \in \pi^+(\mathcal{C})\}$, wobei $\pi^+(\mathcal{C}) = \{A \in \pi(\mathcal{C}) : A^o \neq \varnothing\}$. Also ist $\mathcal{R}$ die Teilmenge von $\mathcal{P}(X)$ bestehend aus allen nichtleeren Mengen der Form $\overline{A^o}$ mit $A \in \pi(\mathcal{C})$; Nach Lemma 12.8 und Lemma 11.4 ist

$$\bigcup_{R\in\mathcal{R}} R = \bigcup_{A\in\pi(\mathcal{C})} \overline{A^o} = \overline{\bigcup_{A\in\pi(\mathcal{C})} A^o} \supset \overline{X \setminus \partial\mathcal{C}} = X\,,$$

und damit ist $\mathcal{R}$ eine endliche abgeschlossene Überdeckung von X. Ferner ist es klar, dass $\mathcal{R} \prec \mathcal{C}$.

Es wird gezeigt, dass $\mathcal{R}$ eine Markov-Partition für T ist, und es folgt unmittelbar aus Lemma 12.9, dass $R = \overline{R^o}$ für jedes $R \in \mathcal{R}$.

Lemma 12.9 *Sei* $F = \bar{G}$ *mit* G *offen. Dann gilt* $F = \overline{F^o}$*.*

Beweis Da $F^o \subset F$, ist $\overline{F^o} \subset \bar{F} = F$. Andererseits ist G offen mit $G \subset F$, also ist $G \subset F^o$ und folglich ist $F = \bar{G} \subset \overline{F^o}$. $\square$

Seien $A_1,\ A_2 \in \pi(\mathcal{C})$ mit $A_1 \neq A_2$; nach Lemma 12.8 und Lemma 11.3 gilt dann

$$(\overline{A_1^o})^o \cap (\overline{A_2^o})^o = (\overline{A_1^o} \cap \overline{A_2^o})^o \subset ((\partial\mathcal{C} \cup A_1^o) \cap (\partial\mathcal{C} \cup A_2^o))^o = (\partial\mathcal{C})^o = \varnothing$$

und dies zeigt, dass $R_1^o \cap R_2^o = \varnothing$ für alle R_1, $R_2 \in \mathcal{R}$ mit $R_1 \neq R_2$.

Seien R_1, $R_2 \in \mathcal{R}$ mit $R_1^o \cap T(R_2^o) \neq \varnothing$ und für $k = 1, 2$ sei A_k das Element von $\pi(\mathcal{C})$ mit $R_k = \overline{A_k^o}$. Da nach Lemma 12.6 $T(R_2^o)$ offen ist und $R_1^o \subset R_1 = \overline{A_1^o}$, ist dann $A_1^o \cap T(R_2^o) \neq \varnothing$. Daher ist auch $T^{-1}(A_1^o) \cap R_2^o \neq \varnothing$. Daraus ergibt sich, dass $T^{-1}(A_1^o) \cap A_2^o \neq \varnothing$, da $T^{-1}(A_1^o)$ offen ist und $R_2^o \subset R_2 = \overline{A_2^o}$. Insbesondere ist dann $T^{-1}(A_1) \cap A_2 \neq \varnothing$ und folglich ist $A_1 \cap T(A_2) \neq \varnothing$. Da aber $\pi(\mathcal{C})$ eine Markov-Überdeckung für T (und auch eine richtige Partition von X) ist, gilt in der Tat $A_1 \subset T(A_2)$, und also ist nach Lemma 12.5 (1) (da $\operatorname{diam}(A_2) \leq \delta$)

$$R_1 = \overline{A_1^o} \subset \overline{(T(A_2))^o} = \overline{T(A_2^o)} = T(\overline{A_2^o}) = T(R_2)\ .$$

Dies zeigt, $\mathcal{R}$ ist eine Markov-Partition für T. Damit ist Satz 12.3 bewiesen. □

Eine endliche Überdeckung $\mathcal{C}$ von X wird *schwacher Generator* für T genannt, wenn für jede Folge $\{C_n\}_{n\geq 0}$ aus $\mathcal{C}$ die Menge $\bigcap_{n\geq 0} T^{-n}(C_n)$ aus höchstens einem Element besteht. (Anders als in Kapitel 11 wird hier nicht verlangt, dass ein schwacher Generator aus offenen Mengen besteht.)

Lemma 12.10 *Sei T expansiv und sei $\delta > 0$ eine expansive Konstante für T. Dann ist jede endliche Überdeckung $\mathcal{C}$ von X mit* $\operatorname{diam}(\mathcal{C}) < \delta$ *ein schwacher Generator für T.*

Beweis Dieser ist identisch mit dem Beweis für Satz 11.1 (3) ⇒ (2). □

Satz 12.4 *Sei T ein lokaler Homeomorphismus, der eine expansive Abbildung ist, und sei $\mathcal{R}$ eine Markov-Partition für T mit* $\operatorname{diam}(\mathcal{R}) < \min\{\delta/2, \eta\}$*, wobei δ eine Lebesgue-Konstante und η eine expansive Konstante für T ist. Definiere $M : \mathcal{R} \times \mathcal{R} \to \{0, 1\}$ durch*

$$M(R_1, R_2) = \begin{cases} 1 & \textit{falls } R_2^o \cap T(R_1^o) \neq \varnothing\ , \\ 0 & \textit{sonst}\ , \end{cases}$$

und sei (Y, S_Y) (mit $Y \subset \operatorname{Abb}(\mathbb{N}, \mathcal{R})$) das Markov-Subshift mit Übergangsmatrix M. Dann gibt es eine stetige surjektive Abbildung $\varphi : Y \to X$ mit $T \circ \varphi = \varphi \circ S_Y$, so dass gilt:

(1) Für jedes $z = \{R_n\}_{n\geq 0} \in Y$ ist $T^n(\varphi(z)) \in R_n$ für jedes $n \geq 0$.

(2) Die Abbildung φ ist injektiv über die dichte G_δ-Menge $X \setminus \bigcup_{n\geq 0} T^{-n}(\partial\mathcal{R})$.

(3) Sind $z = \{R_n\}_{n\geq 0}$, $z' = \{R'_n\}_{n\geq 0} \in Y$ mit $R_{m-1} \cap R'_{m-1} \neq \varnothing$ und $R_m = R'_m$ für ein $m \geq 1$, so ist $R_{m-1} = R'_{m-1}$.

(4) Für alle $x \in X$ ist $|\varphi^{-1}(\{x\})| \leq |\mathcal{R}|$.

Beweis Da $R_1^o \cap R_2^o = \varnothing$ für alle R_1, $R_2 \in \mathcal{R}$ mit $R_1 \neq R_2$, gilt

$$X = \bigcup_{R \in \mathcal{R}} R = \bigcup_{R \in \mathcal{R}} (R \setminus R^o) \cup \bigcup_{R \in \mathcal{R}} R^o \subset \partial\mathcal{R} \cup \bigcup_{R \in \mathcal{R}} R^o \; ;$$

d.h., $X = \partial\mathcal{R} \cup \bigcup_{R \in \mathcal{R}} R^o$. Für jedes $x \in X \setminus \partial\mathcal{R}$ gibt es also ein eindeutiges $R \in \mathcal{R}$, so dass $x \in R^o$. Für $z = \{R_n\}_{n \geq 0} \in Y$ und $m \geq 0$ setze

$$F_m(z) = \{x \in X : T^k(x) \in R_k \text{ für } k = 0, \ldots, m\} = \bigcap_{k=0}^{m} T^{-k}(R_k) \; .$$

Dann ist $\{F_m(z)\}_{m \geq 0}$ eine monoton fallende Folge abgeschlossener Teilmengen von X mit $\bigcap_{m \geq 0} F_m(z) = \bigcap_{n \geq 0} T^{-n}(R_n)$ und damit besteht $\bigcap_{m \geq 0} F_m(z)$ aus höchstens einem Element, da nach Lemma 12.10 $\mathcal{R}$ ein schwacher Generator für T ist.

Lemma 12.11 *Für jedes $m \geq 0$ ist $F_m(z) \neq \varnothing$.*

Beweis Sei $k \geq 1$; da $M(R_{k-1}, R_k) = 1$, ist $R_k^o \cap T(R_{k-1}^o) \neq \varnothing$ und folglich ist $R_k \subset T(R_{k-1})$. Nun sei $m \geq 0$ fest, und wähle $x_m \in R_m$; da $R_k \subset T(R_{k-1})$ für $k = m, \ldots, 1$, kann man $x_{m-1}, \ldots, x_0$ finden, so dass $x_{k-1} \in R_{k-1}$ und $x_k = T(x_{k-1})$ für $k = m, \ldots, 1$. Aber dann ist $x_0 \in F_m(z)$. $\square$

Nach Lemma 12.11 besteht $\bigcap_{n \geq 0} T^{-n}(R_n)$ aus genau einem Element für jedes $z = \{R_n\}_{n \geq 0} \in Y$ und damit gibt es eine eindeutige Abbildung $\varphi : Y \to X$, so dass $T^n(\varphi(z)) \in R_n$ für jedes $n \geq 0$ für jedes $z = \{R_n\}_{n \geq 0} \in Y$. Ferner gilt $T \circ \varphi = \varphi \circ S_Y$: Sei $z = \{R_n\}_{n \geq 0} \in Y$; dann liegen $T^n(T(\varphi(z))$ und $T^n(\varphi(S_Y(z))$ beide in R_{n+1} für jedes $n \geq 0$, und daher ist $T(\varphi(z)) = \varphi(S_Y(z))$.

Die Abbildung φ ist stetig: Für jedes $m \geq 1$ sei $z_m = \{R_n^{(m)}\}_{n \geq 0} \in Y$ mit $\lim_{m \to \infty} z_m = z = \{R_n\}_{n \geq 0}$, und sei $N \geq 1$. Dann gibt es ein $m_0 \geq 1$, so dass $R_n^{(m)} = R_n$ für jedes $n = 0, \ldots, N$ für alle $m \geq m_0$, und daraus folgt, dass $\varphi(z_m) \in F_N(z)$ für alle $m \geq m_0$. Aber $\lim_{N \to \infty} \operatorname{diam}(F_N(z)) = 0$, und damit gilt $\lim_{m \to \infty} \varphi(z_m) = \varphi(z)$.

Die Abbildung φ ist surjektiv: Setze $\Delta^\infty = \bigcup_{n \geq 0} T^{-n}(\partial\mathcal{R})$. Nach Lemma 12.3 und Lemma 11.4 ist $T^{-n}(X \setminus \partial\mathcal{R})$ eine dichte offene Menge für jedes $n \geq 0$ und also ist nach dem Satz von Baire (Lemma 10.3) $X \setminus \Delta^\infty = \bigcap_{n \geq 0} T^{-n}(X \setminus \mathcal{R})$ eine dichte G_δ-Menge. Sei $x \in X \setminus \Delta^\infty$; für jedes $n \geq 0$ ist dann $T^n(x) \in X \setminus \partial\mathcal{R}$ und folglich gibt es ein eindeutiges $R_n \in \mathcal{R}$, so dass $T^n(x) \in R_n^o$. Daraus ergibt sich, dass $T^{n+1}(x) \in R_{n+1}^o \cap T(R_n^o)$ und daher ist $M(R_n, R_{n+1}) = 1$ für jedes $n \geq 0$, d.h., $z = \{R_n\}_{n \geq 0} \in Y$ und nach der Definition von φ ist $\varphi(z) = x$. Dies zeigt, dass $\varphi(Y) \supset X \setminus \Delta^\infty$, und da $\varphi(Y)$ abgeschlossen ist, ist dann $\varphi(Y) = X$.

Dieser Beweis für die Surjektivität zeigt natürlich auch, dass φ injektiv über die dichte G_δ-Menge $X \setminus \Delta^\infty$ ist.

Seien $z = \{R_n\}_{n\geq 0}$, $z' = \{R'_n\}_{n\geq 0} \in Y$ mit $R_{m-1} \cap R'_{m-1} \neq \varnothing$ und $R_m = R'_m$ für ein $m \geq 1$. Dann gilt $R_m \subset T(R_{m-1})$ und $R_m = R'_m \subset T(R'_{m-1})$, und da die Einschränkung von T auf $R_{m-1} \cup R'_{m-1}$ injektiv ist, ist $R_m \subset T(R_{m-1} \cap R'_{m-1})$. Da $R^o_m \neq \varnothing$, ist nach Lemma 12.5 (1) $R^o_{m-1} \cap (R'_{m-1})^o = (R_{m-1} \cap R'_{m-1})^o \neq \varnothing$ und daher ist $R_{m-1} = R'_{m-1}$.

Seien $z = \{R_n\}_{n\geq 0}$, $z' = \{R'_n\}_{n\geq 0} \in Y$ mit $z \neq z'$ und $\varphi(z) = \varphi(z') = x$; dann ist $T^n(x) \in R_n \cap R'_n$ und damit $R_n \cap R'_n \neq \varnothing$ für jedes $n \geq 0$. Gilt $R_m = R'_m$ für ein $m \geq 1$, so folgt aus (3), dass auch $R_{m-1} = R'_{m-1}$. Also gibt es ein $p \geq 0$, so dass $R_n \neq R'_n$ für alle $n \geq p$. Sei nun $x \in X$, $m \geq 2$, und für jedes $k = 1, \ldots, m$ sei $z_k = \{R^{(k)}_n\}_{n\geq 0} \in \varphi^{-1}(\{x\})$ mit $z_j \neq z_k$ für alle $1 \leq j < k \leq m$. Für jedes $1 \leq j < k \leq m$ gibt es also ein $p_{j,k} \geq 0$, so dass $R^{(j)}_n \neq R^{(k)}_n$ für alle $n \geq p_{j,k}$. Setze $p = \max\{p_{j,k} : 1 \leq j < k \leq m\}$; dann sind $R^{(1)}_p, \ldots, R^{(m)}_p$ m verschiedene Elemente aus $\mathcal{R}$ und daher muss $m \leq |\mathcal{R}|$ sein. Dies zeigt, dass $|\varphi^{-1}(\{x\})| \leq |\mathcal{R}|$ für jedes $x \in X$.

Damit ist Satz 12.4 bewiesen. $\square$

Beweis für Satz 12.1 (1) $\Rightarrow$ (2) ist Satz 12.2, (2) $\Rightarrow$ (3) ist Satz 12.3, (3) $\Rightarrow$ (4) ist Satz 12.4 und (4) $\Rightarrow$ (1) ist klar. $\square$

13 Die Pseudo-Orbit tracing Eigenschaft

Im Folgenden sei (X,d) ein kompakter metrischer Raum mit $X \neq \varnothing$ und sei $T : X \to X$ eine stetige Abbildung bzw. ein Homeomorphismus. Sei $I = \mathbb{N}$ bzw. $I = \mathbb{Z}$, wenn T eine stetige Abbildung bzw. ein Homeomorphismus ist.

Eine Folge $\{x_n\}_{n\in I}$ aus X heißt *δ-Pseudo-Orbit*, wenn $d(T(x_n), x_{n+1}) < \delta$ für alle $n \in I$. Ein Punkt $x \in X$ heißt *ε-Schatten* einer Folge $\{x_n\}_{n\in I}$ aus X, wenn $d(T^n(x), x_n) < \varepsilon$ für jede $n \in I$. Nun hat T die *Pseudo-Orbit tracing Eigenschaft*, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass jeder δ-Pseudo-Orbit einen ε-Schatten besitzt.

Satz 13.1 *Hat T die Pseudo-Orbit tracing Eigenschaft und ist T expansiv, so ist (X,T) ein Faktor eines Markov-Subshifts (Y, S_Y).*

Beweis Sei $\varepsilon = \kappa/2$, wobei $\kappa > 0$ eine expansive Konstante für T ist. Da T die Pseudo-Orbit tracing Eigenschaft hat, gibt es ein $\delta > 0$, so dass jeder δ-Pseudo-Orbit einen ε-Schatten besitzt, und da T gleichmäßig stetig ist, gibt es dann ein $\gamma > 0$ mit $\gamma \leq \min\{\varepsilon, \delta/2\}$, so dass $d(T(x), T(y)) < \delta/2$ für alle x, $y \in X$ mit $d(x,y) < \gamma$. Sei nun E eine endliche Teilmenge von X mit $\bigcup_{u\in E} \mathrm{B}(u,\gamma) = X$ und definiere $M : E \times E \to \{0,1\}$ durch

$$M(u,v) = \begin{cases} 1 & \text{falls } d(T(u),v) < \delta\ , \\ 0 & \text{sonst}\ . \end{cases}$$

Sei (Y, S_Y) das Markov-Subshift mit Übergangsmatrix M. Eine Folge $\{x_n\}_{n\in I}$ aus X besitzt höchstens einen ε-Schatten: Sind x und x' beide ε-Schatten, so gilt

$$d(T^n(x), T^n(x')) \leq d(T^n(x), x_n) + d(T^n(x'), x_n) < 2\varepsilon = \kappa$$

für alle $n \in I$, und damit ist $x = x'$. Dies bedeutet, dass jeder δ-Pseudo-Orbit genau einen ε-Schatten besitzt.

Sei $z = \{z_n\}_{n\in I} \in Y$; als Folge von Elementen aus X ist $\{z_n\}_{n\in I}$ ein δ-Pseudo-Orbit, und folglich gibt es ein eindeutiges $\varphi(z) \in X$ mit $d(T^n(\varphi(z)), z_n) < \varepsilon)$ für alle $n \in I$.

Die Abbildung $\varphi : Y \to X$ ist stetig: Für jedes $m \geq 1$ sei $z^{(m)} = \{z_n^{(m)}\}_{n\in I} \in Y$ mit $\lim_{m\to\infty} z^{(m)} = z = \{z_n\}_{n\in I}$ und nehme an, dass $\lim_{m\to\infty} \varphi(z^{(m)}) = x$ für ein $x \in X$. Für jedes $n \in I$ gibt es dann $m_n \geq 1$, so dass $z_n^{(m)} = z_n$ für alle $m \geq m_n$ und also gilt

$$\begin{aligned} &d(T^n(\varphi(z)), T^n(\varphi(z^{(m)}))) \\ &\qquad \leq d(T^n(\varphi(z)), z_n) + d(z_n, z_n(m)) + d(T^n(\varphi(z^{(m)})), z_n^{(m)}) \\ &\qquad < 2\varepsilon + d(z_n, z_n^{(m)}) = 2\varepsilon = \kappa \end{aligned}$$

für alle $m \geq m_n$. Daraus ergibt sich, dass $d(T^n(\varphi(z)), T^n(x)) \leq \kappa$ für alle $n \in I$ und daher ist $\varphi(z) = x$. Dies zeigt, dass φ stetig ist. (Aus der Kompaktheit von Y folgt: Ist $g : Y \to X$ eine Abbildung, die an der Stelle $y \in Y$ nicht stetig ist, so gibt es eine Folge $\{y_m\}_{m\geq 1}$ aus Y mit $\lim_{m\to\infty} y_m = y$, für die $\{g(y_m)\}_{m\geq 1}$ konvergiert aber mit $\lim_{m\to\infty} g(y_m) \neq g(y)$.)

Die Abbildung φ ist surjektiv: Sei $x \in X$ und für jedes $n \in I$ wähle $z_n \in E$ mit $d(T^n(x), z_n) < \gamma$. Dann gilt

$$\begin{aligned} d(T(z_n), z_{n+1}) &\leq d(T(z_n), T^{n+1}(x)) + d(T^{n+1}(x), z_{n+1}) \\ &= d(T(T^n(x))), T(z_n)) + d(T^{n+1}(x), z_{n+1}) < \delta/2 + \gamma \leq \delta \end{aligned}$$

für jedes $n \in I$ und also ist $z = \{z_n\}_{n\in I} \in Y$. Aber $d(T^n(x), z_n) < \gamma \leq \varepsilon$ für jedes $n \in I$ und damit ist $\varphi(z) = x$.

Es gilt $\varphi \circ S_Y = T \circ \varphi$: Sei $z = \{z_n\}_{n\in I} \in Y$; da $d(T^n(\varphi(S_Y(z))), z_{n+1}) < \varepsilon$ und $d(T^n(\varphi(z)), z_n) < \varepsilon$ für jedes $n \in I$, ist dann

$$\begin{aligned} &d(T^n(\varphi(S_Y(z))), T^n(T(\varphi(z)))) \\ &\qquad \leq d(T^n(\varphi(S_Y(z))), z_{n+1}) + d(T^{n+1}(\varphi(z)), z_{n+1}) < \varepsilon + \varepsilon = \kappa \end{aligned}$$

für alle $n \in I$, und daraus folgt, dass $\varphi(S_Y(z)) = T(\varphi(z))$. $\square$

Eine stetige Abbildung $g : X \to X$ wird *expandierend* genannt, wenn g surjektiv und offen ist und es $\eta > 0$ und $\lambda > 1$ gibt, so dass

$$d(T(x), T(y)) \geq \lambda d(x, y)$$

für alle x, $y \in X$ mit $d(x, y) \leq \eta$.

Lemma 13.1 *Eine expandierende Abbildung ist expansiv (und dann ist η eine expansive Konstante).*

Beweis Seien x, $y \in X$ mit $x \neq y$. Dann gilt $d(g^n(x), g^n(y)) \geq \lambda^n d(x, y)$, falls $d(g^k(x), g^k(y)) \leq \eta$ für $k = 0, \ldots, n-1$. Folglich ist $d(g^n(x), g^n(y)) > \eta$ für ein $n \leq \min\{k \geq 0 : \lambda^k > \eta/d(x, y)\}$. $\square$

Lemma 13.2 *Sei (Y, S_Y) ein Subshift vom endlichen Typ; dann ist $S_Y : Y \to Y$ expandierend.*

Beweis Wie im Beweis für Lemma 12.1 kann man annehmen, dass (Y, S_Y) ein Markov-Subshift ist, und in diesem Fall sieht man leicht, dass S_Y expandierend ist mit $\lambda = 2$ und $\eta = 1$. $\square$

Lemma 13.3 *Sei* $A \in \mathrm{M}(n \times n, \mathbb{Z})$ *mit* $\det(A) \neq 0$*; dann ist* $\theta_A : \mathrm{K}^n \to \mathrm{K}^n$ *expandierend genau, wenn* $|\lambda| > 1$ *für jeden Eigenwert* λ *von* A*.*

Beweis Übung. □

Satz 13.2 *Ist* T *expandierend, so hat* T *die Pseudo-Orbit tracing Eigenschaft.*

Beweis Eine endliche Folge $\{x_n\}_{n=0}^m$ aus X heißt *endlicher* δ*-Pseudo-Orbit*, wenn $d(T(x_n), x_{n+1}) < \delta$ für $n = 0, \ldots, m-1$. Ein Punkt $x \in X$ heißt ε*-Schatten* einer endlichen Folge $\{x_n\}_{n=0}^m$, wenn $d(T^n(x), x_n) < \varepsilon$ für jede $n = 0, \ldots, m$.

Lemma 13.4 *Gibt es zu jedem* $\varepsilon > 0$ *ein* $\delta > 0$*, so dass jeder endliche* δ*-Pseudo-Orbit einen* ε*-Schatten besitzt, dann hat* T *die Pseudo-Orbit tracing Eigenschaft.*

Beweis Sei $\varepsilon > 0$ und wähle $\delta > 0$, so dass jeder endliche δ-Pseudo-Orbit einen $\varepsilon/2$-Schatten besitzt. Sei nun $\{x_n\}_{n \geq 0}$ ein δ-Pseudo-Orbit; für jedes $m \geq 0$ ist $\{x_n\}_{n=0}^m$ ein endlicher δ-Pseudo-Orbit, der dann einen ε-Schatten z_m besitzt. Da X kompakt ist, gibt es $z \in X$ und eine Teilfolge $\{m_k\}_{k \geq 0}$, so dass $\{z_{m_k}\}_{k \geq 0}$ gegen z konvergiert. Sei $n \geq 0$; für alle $m \geq n$ ist $d(T^n(z_m), x_n) < \varepsilon/2$ und folglich gilt $d(T^n(z), x_n) \leq \varepsilon/2 < \varepsilon$. Damit ist z ein ε-Schatten der Folge $\{x_n\}_{n \geq 0}$. □

Lemma 13.5 *Sei* $g : X \to X$ *eine stetige offene Abbildung und sei* $\beta > 0$*. Dann gibt es ein* $\alpha > 0$*, so dass es für alle* $x, y \in X$ *mit* $d(g(x), y) < \alpha$ *ein* $z \in \mathrm{B}(x, \beta)$ *mit* $g(z) = y$ *gibt.*

Beweis Da X kompakt und damit total beschränkt ist, gibt es eine endliche Menge $\Delta \subset X$, so dass $X = \bigcup_{z \in \Delta} \mathrm{B}(z, \beta/3)$. Für jedes $z \in X$ gilt

$$\overline{g(\mathrm{B}(z, \beta/3))} = g(\overline{\mathrm{B}(z, \beta/3)}) \subset g(\mathrm{B}(z, 2\beta/3)) \; ;$$

also ist $\overline{g(\mathrm{B}(z, \beta/3))}$ eine kompakte Teilmenge der offenen Menge $g(\mathrm{B}(z, 2\beta/3))$. Daher gibt es $\alpha_z > 0$, so dass $\mathrm{B}(y, \alpha_z) \subset g(\mathrm{B}(z, 2\beta/3))$ für alle $y \in g(\mathrm{B}(z, \beta/3))$ und folglich gilt dann $\mathrm{B}(g(x), \alpha_z) \subset g(\mathrm{B}(z, 2\beta/3))$ für alle $x \in \mathrm{B}(z, \beta/3)$. Setze $\alpha = \min\{\alpha_z : z \in \Delta\}$ und seien $x, y \in X$ mit $d(g(x), y) < \alpha$. Wähle $u \in \Delta$ mit $x \in \mathrm{B}(u, \beta/3)$; da $y \in \mathrm{B}(g(x), \alpha) \subset g(\mathrm{B}(u, 2\beta/3))$, gibt es ein $z \in \mathrm{B}(u, 2\beta/3)$ mit $g(z) = y$, und es gilt $d(x, z) \leq d(x, u) + d(z, u) < \beta/3 + 2\beta/3 = \beta$. □

Sei nun $\eta > 0$ und $\lambda > 1$, so dass $d(T(x), T(y)) \geq \lambda d(x, y)$ für alle $x, y \in X$ mit $d(x, y) < \eta$. Sei $\alpha > 0$ wie in Lemma 13.5 mit $g = T$ und $\beta = \eta$.

Lemma 13.6 *Sei $m \geq 0$. Für alle x, $y \in X$ mit $d(T^m(x), y) < \alpha$ gibt es dann ein $z \in X$ mit $T^m(z) = y$ und $d(T^k(x), T^k(z)) \leq \lambda^{k-m} d(T^m(x), y)$ für $k = 0, \ldots, m$.*

Beweis (Durch Induktion nach m.) Für $m = 0$ ist die Aussage trivial richtig. Sei also $m > 1$, und nehme an, die Aussage gilt für $m' = m - 1$. Seien x, $y \in X$ mit $d(T^m(x), y) < \alpha$ und setze $x' = T(x)$; da $d(T^{m-1}(x'), y) = d(T^m(x), y) < \alpha$, gibt es ein $z' \in X$ mit $T^{m-1}(z') = y$ und $d(T^k(x'), T^k(z')) \leq \lambda^{k-m+1} d(T^{m-1}(x'), y)$ für $k = 0, \ldots, m-1$. Insbesondere ist dann

$$d(T(x), z') = d(x', z') \leq \lambda^{-m+1} d(T^{m-1}(x'), y) = \lambda^{-m+1} d(T^m(x), y) < \alpha \, .$$

Nach Lemma 13.5 gibt es nun ein $z \in \mathrm{B}(x, \eta)$ mit $T(z) = z'$ und da $d(x, z) < \eta$, gilt dann $d(T(x), T(z)) \geq \lambda d(x, z)$. Also gilt $T^m(z) = T^{m-1}(z') = y$ und

$$\begin{aligned} d(x, z) \leq \lambda^{-1} d(T(x), T(z)) &= \lambda^{-1} d(x', z') \\ &\leq \lambda^{-m} d(T^{m-1}(x'), y) = \lambda^{-m} d(T^m(x), y) \, , \end{aligned}$$

und für $k = 1, \ldots, m$ gilt auch

$$\begin{aligned} d(T^k(x), T^k(z)) &= d(T^{k-1}(x'), T^{k-1}(z')) \\ &\leq \lambda^{(k-1)-m+1} d(T^{m-1}(x'), y) = \lambda^{k-m} d(T^m(x), y) \, . \quad \square \end{aligned}$$

Nun sei $\varepsilon > 0$ und setze $\delta = \min\{\alpha, (\lambda-1)\varepsilon\}$. Sei $\{x_n\}_{n=0}^m$ ein endlicher δ-Pseudo-Orbit. Da T^m surjektiv ist, gibt es $u_0 \in X$ mit $T^m(u_0) = x_0$. Sei $1 \leq p \leq m$ und nehme an, es gibt $u_0, \ldots, u_{p-1} \in X$, so dass $T^m(u_k) = x_k$ für $k = 0, \ldots, p-1$ und $d(T^j(u_k), T^{j+1}(u_{k-1})) < \lambda^{j-m}\delta$ für alle $j = 0, \ldots, m$ und $k = 1, \ldots, p-1$. Dann gilt $T^m(T(u_{p-1})) = T(x_{p-1})$ und $d(T(x_{p-1}), x_p) < \delta \leq \alpha$; also gibt es nach Lemma 13.6 ein $u_p \in X$ mit $T^m(u_p) = x_p$ und

$$d(T^j(u_p), T^{j+1}(u_{p-1})) = d(T^j(T(u_{p-1})), T^j(u_p)) \leq \lambda^{j-m} d(T(x_{p-1}), x_p) < \lambda^{j-m}\delta$$

für $j = 0, \ldots, m$. Dies zeigt, dass es $u_0, \ldots, u_m \in X$ gibt, so dass $T^m(u_k) = x_k$ für $k = 0, \ldots, m$ und $d(T^j(u_k), T^{j+1}(u_{k-1})) < \lambda^{j-m}\delta$ für alle $j = 0, \ldots, m$ und $k = 1, \ldots, m$. Sei $0 \leq n < m$; da $x_n = T^m(u_n)$, gilt

$$\begin{aligned} d(T^n(u_m), x_n) &\leq \sum_{k=0}^{m-n-1} d(T^{n+k}(u_{m-k}), T^{n+k+1}(u_{m-k-1})) \\ &\leq \delta \sum_{k=0}^{m-n-1} \lambda^{n+k-m} = \delta \sum_{j=1} \lambda^{-j} < \delta(\lambda-1)^{-1} \leq \varepsilon \, , \end{aligned}$$

und $d(T^m(u_m), x_m) = 0$, da $x_m = T^m(u_m)$, Folglich ist u_m ein ε-Schatten der Folge $\{x_n\}_{n=0}^m$. Daraus ergibt sich nach Lemma 13.4, dass T die Pseudo-Orbit tracing Eigenschaft hat. $\square$

Satz 13.3 *Sei T ein lokaler Homeomorphismus, der expandierend ist. Dann gibt es ein Markov-Subshift (Y, S_Y) und eine stetige surjektive Abbildung $\varphi : Y \to X$ mit $T \circ \varphi = \varphi \circ S_Y$, so dass φ injektiv über eine dichte G_δ-Menge in X ist und $|\varphi^{-1}(\{x\})| \le N$ für alle $x \in X$ für ein $N \ge 1$.*

Beweis Dies folgt unmittelbar aus Satz 12.1, Satz 13.1 und Satz 13.2. □

Für Abbildungen g, $g' : X \to X$ sei

$$\mathbf{d}(g, g') = \sup\{d(g(x), g'(x)) : x \in X\} .$$

Seien T, $S : X \to X$ stetige Abbildungen bzw. Homeomorphismen. Dann heißt T ein *ε-Faktor* von S, wenn es eine stetige Abbildung $\varphi : X \to X$ mit $\mathbf{d}(\varphi, \mathrm{id}_X) \le \varepsilon$ gibt, so dass $\varphi \circ S = T \circ \varphi$. (Man beachte: Hier wird nicht verlangt, dass φ surjektiv ist.)

Die stetige Abbildung bzw. der Homeomorphismus $T : X \to X$ heißt *topologisch stabil*, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass T ein ε-Faktor ist von jeder stetigen Abbildung bzw. von jedem Homeomorphismus $S : X \to X$ mit $\mathbf{d}(T, S) < \delta$.

Satz 13.4 *Eine expansive Abbildung bzw. ein expansiver Homeomorphismus T mit der Pseudo-Orbit tracing Eigenschaft ist topologisch stabil.*

Beweis Sei $\kappa > 0$ eine expansive Konstante für T und sei $\varepsilon > 0$ mit $\varepsilon < \kappa/2$. Sei $\delta > 0$, so dass jeder δ-Pseudo-Orbit einen ε-Schatten besitzt. Sei $S : X \to X$ eine stetige Abbildung bzw. ein Homeomorphismus mit $\mathbf{d}(T, S) < \delta$. Für jedes $x \in X$ ist dann $\{S^n(x)\}_{n \in I}$ ein δ-Pseudo-Orbit (bezüglich T), da

$$d(T(S^n(x)), S^{n+1}(x)) = d(T(S^n(x)), S(S^n(x))) \le \mathbf{d}(T, S) < \delta$$

für alle $n \in I$. Folglich besitzt $\{S^n(x)\}_{n \in I}$ einen ε-Schatten, der eindeutig sein muss. (Sind z und z' beide ε-Schatten, so ist

$$d(T^n(z), T^n(z')) \le d(T^n(z), S^n(x)) + d(T^n(z'), S^n(x)) < \varepsilon + \varepsilon < \kappa$$

und damit ist $z' = z$.) Bezeichne diesen eindeutigen ε-Schatten mit $\varphi(x)$; auf diese Weise wird eine Abbildung $\varphi : X \to X$ definiert.

Sei $x \in X$; für alle $n \in I$ gilt dann

$$\begin{aligned} &d(T^n(\varphi(S(x))), T^n(T(\varphi(x)))) \\ &\qquad \le d(T^n(\varphi(S(x))), S^{n+1}(x)) + d(S^{n+1}(x), T^n(T(\varphi(x)))) \\ &\qquad = d(T^n(\varphi(S(x))), S^n(S(x))) + d(T^{n+1}(\varphi(x)), S^{n+1}(x)) < \varepsilon + \varepsilon < \kappa \end{aligned}$$

und daraus folgt, dass $\varphi(S(x)) = T(\varphi(x))$, d.h., $\varphi \circ S = T \circ \varphi$. Da ferner $\varphi(x)$ ein ε-Schatten von $\{S^n(x)\}_{n\in I}$ ist, ist insbesondere

$$d(\varphi(x), \mathrm{id}_X(x)) = d(\varphi(x), x) = d(T^0(\varphi(x)), S^0(x)) < \varepsilon$$

für alle $x \in X$ und daher ist $\mathbf{d}(\varphi, \mathrm{id}_X) \leq \varepsilon$.

Um zu zeigen, dass φ stetig ist, wird das folgende Lemma benötigt:

Lemma 13.7 *Zu jedem $\eta > 0$ gibt es ein $m \geq 1$ mit der folgenden Eigenschaft: Sind x, $y \in X$ mit $d(T^n(x), T^n(y)) \leq \kappa$ für alle $n \in I$ mit $|n| \leq m$, so ist $d(x, y) < \eta$.*

Beweis Nehme an, es gibt ein $\eta > 0$ und für jedes $m \geq 1$ Elemente x_m, $y_m \in X$ mit $d(x_m, y_m) \geq \eta$, so dass $d(T^n(x_m), T^n(y_m)) \leq \kappa$ für alle $n \in I$ mit $|n| \leq m$. Da X kompakt ist, gibt es also eine Teilfolge $\{m_k\}_{k\geq 1}$ und x, $y \in X$, so dass $\lim_{k\to\infty} x_{m_k} = x$ und $\lim_{k\to\infty} y_{m_k} = y$. Dann ist aber $d(T^n(x), T^n(y)) \leq \kappa$ für alle $n \in I$ und $d(x, y) \geq \eta$. Insbesondere ist $x \neq y$, und dies ist ein Widerspruch, da κ eine expansive Konstante für T ist. $\square$

Sei nun $\eta > 0$ und sei $m \geq 1$ wie in Lemma 13.7. Da jedes S^k gleichmäßig stetig ist, gibt es ein $\gamma > 0$, so dass $d(S^n(x), S^n(y)) < \kappa - 2\varepsilon$ für alle $n \in I$ mit $|n| \leq m$ und alle x, $y \in X$ mit $d(x, y) < \gamma$. Seien x, $y \in X$ mit $d(x, y) < \gamma$; dann gilt

$$\begin{aligned} d(T^n(\varphi(x)), T^n(\varphi(y))) &= d(\varphi(S^n(x)), \varphi(S^n(y))) \\ &\leq d(\varphi(S^n(x)), S^n(x)) + d(S^n(x), S^n(y)) + d(\varphi(S^n(y)), S^n(y)) \\ &< \varepsilon + (\kappa - 2\varepsilon) + \varepsilon = \kappa \end{aligned}$$

und daraus ergibt sich, dass $d(\varphi(x), \varphi(y)) < \eta$. Dies zeigt, dass φ stetig ist, und folglich ist T ein ε-Faktor von S. Damit ist Satz 13.4 bewiesen. $\square$

Bemerkung: Seien T, $S : X \to X$ stetige Abbildungen bzw. Homeomorphismen und sei $\varphi : X \to X$ eine stetige Abbildung mit $\varphi \circ S = T \circ \varphi$ und $\mathbf{d}(\varphi, \mathrm{id}_X) \leq \varepsilon$. Ist S expansiv mit expansiver Konstante $\kappa \geq 2\varepsilon$, so ist φ injektiv: Seien x, $y \in X$ mit $\varphi(x) = \varphi(y)$. Für jedes $n \in I$ ist dann

$$\begin{aligned} &d(S^n(x), S^n(y)) \\ &\qquad \leq d(S^n(x), \varphi(S^n(x))) + d(\varphi(S^n(x)), \varphi(S^n(y))) + d(\varphi(S^n(y)), S^n(y)) \\ &\qquad \leq \varepsilon + d(\varphi(S^n(x)), \varphi(S^n(y))) + \varepsilon \\ &\qquad = 2\varepsilon + d(T^n(\varphi(x)), T^n(\varphi(y))) = 2\varepsilon \leq \kappa \end{aligned}$$

und folglich ist $x = y$.

Literatur

[1] Billingsley, P. (1965): Ergodic theory and information. Wiley

[2] Cornfeld, I., Fomin, S. Sinai, Ya. (1982): Ergodic Theory. Springer

[3] Denker, M., Grillenberg, Ch., Sigmund, K. (1976): Ergodic theory on compact spaces. Springer

[4] Derriennic, F. (1975): Sur le théorème ergodique sous-additiv, C. R. Acad. Sc. Paris, 281A, 985-988

[5] Fritz, F., Huppert, B., Willems, W. (1979): Stochastische Matrizen. Springer

[6] Jacobs, K. (1973): Gleichverteilung mod 1, Selecta Mathematica IV

[7] Kronecker, L. (1884): Die Periodensysteme von Funktionen reeller Variablen, Kgl. Preuß. Akad. Wiss. Berlin, 1071-1080

[8] Pollicott, M. (1998): Dynamical systems and ergodic theory. Cambridge University Press

[9] Ruelle, D. (1979): Ergodic theory of differentiable dynamical systems, Publ. I.H.E.S., 50, 275-306

[10] Walters, P. (2000); An Introduction to Ergodic Theory. Springer

Index

G_δ-Menge, 92
δ-Pseudo-Orbit, 113
    endlicher, 115
d-System, 3
σ-Algebra
    der terminalen Ereignisse, 17
    invariante, 7
    triviale, 13, 21
ε-Faktor, 117
ε-Schatten, 113, 115

Abbildung
    bimessbare, 50
    eindeutig ergodische, 29
    einseitig topologisch transitive, 93
    ergodische, 13, 21
    expandierende, 114
    expansive, 100
    integrierbare, 4
    invertierbare, 69
    isometrisierbare, 98
    maßtreue, 7
    messbare, 4, 7
    minimale, 91
    offene, 107
    schwach-mischende, 23
    stark-mischende, 23
    topologisch stabile, 117
    topologisch transitive, 93
Abbildungen
    isomorphe, 51
    strikt isomorphe, 50

bedingte Entropie, 56, 64
bedingte Erwartung, 5
Bildmaß, 5
bimessbare Abbildung, 50
Birkhoffscher Ergodensatz, 8

Charakter, 45
Charaktergruppe, 45

dichte Teilmenge, 92
Durchschnittsstabilität, 3

eindeutig ergodische Abbildung, 29
endliche Überdeckung, 100
    abgeschlossene, 100
    offene, 100
endliche Partition, 61
endlicher δ-Pseudo-Orbit, 115
Entropie, 55, 60, 62
    bedingte, 56, 64
Ergodensatz
    Birkhoffscher, 8
    für stochastische Matrizen, 31
ergodische Abbildung, 13, 21
Erwartung
    bedingte, 5
expandierende Abbildung, 114
expansive Abbildung, 100
expansive Konstante, 100
expansiver Homeomorphismus, 100
extremaler Punkt, 28

Faktor, 97, 117
Familie
    stationäre, 84
    subadditive, 84
fast sicheres Ereignis, 3
Fixpunkt, 91
Folge
    identisch verteilte, 19
    stationäre, 16
    subadditive, 58
    unabhängige, 19

Generator, 101
    schwacher, 101, 110
Gruppe
    metrische, 40

Haarsches Maß, 41

Homeomorphismus
  expansiver, 100
  lokaler, 105

identisch verteilte Folge, 19
Integral, 4
integrierbare Abbildung, 4
invariante σ-Algebra, 7
invariantes Maß, 28
invertierbare Abbildung, 69
irreduzible stochastische Matrix, 32
Isometrie, 98
isometrisierbare Abbildung, 98
isomorphe Abbildungen, 51

konjugierte Räume, 97
Konstante
  expansive, 100
  Lebesgue-, 101, 107

Lebesgue-Konstante, 101, 107
lokaler Homeomorphismus, 105

Markov-Überdeckung, 105, 108
Markov-Partition, 106
Markov-Subshift, 91
Matrix
  stochastische, 30
Maximaler Ergodensatz von Hopf, 8
Maß
  Haarsches, 41
  invariantes, 28
maßtreue Abbildung, 7
messbare Abbildung, 4, 7
Messraum, 3
metrische Gruppe, 40
minimale Abbildung, 91

nicht-wandernde Menge, 96
Null-Eins-Gesetz von Kolmogorov, 19

offene Abbildung, 107

Partition
  endliche, 61
  Markov-, 106
Periode, 91
periodischer Punkt, 91
Pseudo-Orbit, 113
Pseudo-Orbit tracing Eigenschaft, 113

reguläre Überdeckung, 108

Satz von Baire, 92
Satz von Kolmogorov-Sinai, 73
Schatten, 113, 115
schwach-mischende Abbildung, 23
schwacher Generator, 101, 110
Shift-Raum, 90
stark-mischende Abbildung, 23
Starkes Gesetz der großen Zahlen, 20
stationäre Familie, 84
stationäre Folge, 16
stochastische Matrix, 30
  irreduzible, 32
  unzerlegbare, 31
strikt isomorphe Abbildungen, 50
subadditive Familie, 84
subadditive Folge, 58
Subadditiver Ergodensatz
  von Kingman, 80
Subshift, 91
  Markov-, 91
  vom endlichen Typ, 98

topologisch stabile Abbildung, 117
topologisch transitive Abbildung, 93
Torus, 41
triviale σ-Algebra, 13, 21

Überdeckung
  endliche, 100
  Markov-, 105, 108
  reguläre, 108
unabhängige Folge, 19
unzerlegbare stochastische Matrix, 31

verträgliche Familie von Maßen, 14

Wahrscheinlichkeitsmaß, 3

Wahrscheinlichkeitsraum, 3
Wahrscheinlichkeitsvektor, 30
    stationärer, 30
Wiederkehrsatz von Poincaré, 7