Stochastische Dynamik

Vorlesung von
Prof. Peter Imkeller

Humboldt Universität zu Berlin

Sommersemester 2004

Mitschrift von Wolfgang Siegert

Inhaltsverzeichnis

1 Konstruktion und elementare Eigenschaften von Markov-Ketten **1**

2 Invariante Maße und asymptotisches Verhalten **8**

3 Stationäre Prozesse **16**

4 Der Birkhoffsche Ergodensatz **22**

5 Der Subadditive Ergodensatz von Kingman **25**

6 Der Satz von Furstenberg-Kesten **32**

7 Der Multiplikative Ergodensatz von Oseledets **41**

Notationen **52**

Literaturverzeichnis **52**

Index **53**

1. Markov-Ketten: Konstruktion und elementare Eigenschaften

Definition 1.1. Sei $(S, \mathscr{S})$ ein meßbarer Raum. Eine Funktion

$$p : S \times \mathscr{S} \to [0, 1]$$

heißt *Übergangswahrscheinlichkeit*, falls gilt:

(a) für jedes $x \in S$ ist $p(x, \,.\,)$ Wahrscheinlichkeitsmaß auf $(S, \mathscr{S})$;
(b) für jedes $A \in \mathscr{S}$ ist $p(\,.\,, A)$ $\mathscr{S}$-meßbar.

Bemerkung 1.2. Sei p Übergangswahrscheinlichkeit auf einem meßbaren Raum $(S, \mathscr{S})$.

i) Ist $f : S \to \mathbb{R}$ $\mathscr{S}$-$\mathscr{B}^1$-meßbar und beschränkt, so auch $g := \int_S f(x)\, p(\,.\,, dx)$;

ii) Ist μ ein Wahrscheinlichkeitsmaß auf $(S, \mathscr{S})$, so auch $\nu := \int_S p(x, \,.\,)\, \mu(dx)$.

Beweis. i) Wegen 1.1(b) gilt die Aussage für Indikatorfunktionen $f = 1_A$ meßbarer Mengen, also auch (Linearität des Integrals) für Treppenfunktionen.
Ist $f \geq 0$, so existieren approximierende Treppenfunktionen $0 \leq f_n \nearrow f$; hier ist $g_n := \int f_n(x)\, p(\,.\,, dx)$ meßbar (da f_n Treppenfunktion) und (durch die Schranke von f) beschränkt; andererseits gilt (Satz über monotone Konvergenz) $g_n \nearrow \int f(x)\, p(\,.\,, dx) \equiv g$, sodaß g als punktweiser Limes meßbarer Funktionen selbst meßbar ist; g ist ebenfalls durch die Schranke von f beschränkt.
Im allgemeinen Fall ist $f = f^+ - f^-$ mit $f^+, f^- \geq 0$; aufgrund des bisherigen sind $g^\pm := \int f^\pm(x)\, p(\,.\,, dx)$ meßbar und beschränkt, also auch $g = g^+ - g^-$.

ii) Für eine Folge $(A_n)_{n \in \mathbb{N}}$ paarweise disjunkter $A_n \in \mathscr{S}$ gilt:

$$\nu\left(\dot{\bigcup_n} A_n\right) \equiv \int_S p\left(x, \dot{\bigcup} A_n\right) \mu(dx) \overset{1.1(a)}{=} \int_S \sum_n p(x, A_n)\, \mu(dx)$$
$$\overset{\text{mon.Kvgz.}}{=} \sum_n \int_S p(x, A_n)\, \mu(dx) \equiv \sum_n \nu(A_n) \;;$$

ferner ist $\nu(S) \equiv \int_S p(x, S)\, \mu(dx) \overset{1.1(a)}{=} \int_S \mu(dx) = 1$ (μ W.Maß). □

Definition 1.3. Sei S ein polnischer Raum und hierauf $(p_n)_{n \in \mathbb{N}}$ eine Folge von Übergangswahrscheinlichkeiten sowie μ ein Wahrscheinlichkeitsmaß. Dann sei $P_0 := \mu$ und

$$P_n(B_0 \times \cdots \times B_n) := \int\limits_{B_0 \times \cdots \times B_n} p_n(x_{n-1}, dx_n) p_{n-1}(x_{n-2}, dx_{n-1}) \cdots p_1(x_0, dx_1)\, \mu(dx_0)$$

zu $n \in \mathbb{N}$ und $B_i \in \mathscr{S} \equiv \mathscr{B}(S)$.

Mit Bemerkung 1.2 folgt rekursiv, daß P_n wohldefiniert ist auf dem Semiring[1]

$$\mathscr{R}_n := \{B_0 \times \cdots \times B_n : B_i \in \mathscr{S}\}\ .$$

[1] Ein Mengensystem $\mathscr{P}$ heißt *Semiring*, falls gilt (cf. Halmos [HM 74, S.22]):

- zu $E \in \mathscr{P}$ und $F \in \mathscr{P}$ ist auch $E \cap F \in \mathscr{P}$, und
- zu $E \in \mathscr{P}$ und $F \in \mathscr{P}$ mit $E \subset F$ existieren endlich viele $C_0, C_1, \ldots, C_n \in \mathscr{P}$, sodaß

$$E = C_0 \subset C_1 \subset \cdots \subset C_n = F \qquad \text{und} \qquad C_i \setminus C_{i-1} \in \mathscr{P} \quad (i = 1, \ldots, n)\ .$$

Ferner setzt sich P_n fort zu einem Maß auf der von dem Ring $r(\mathscr{R}_n)$ erzeugten σ-Algebra

$$\sigma\left(r(\mathscr{R}_n)\right) = \underbrace{\mathscr{S} \otimes \cdots \otimes \mathscr{S}}_{(n+1)\text{-mal}} \equiv \mathscr{S}^{n+1} \quad \left(\overset{S \text{ poln.}}{=} \mathscr{B}(S^{n+1})\right) .$$

Beweis. P_n induziert einen (endlichen) Inhalt auf dem Ring $r(\mathscr{R}_n)$. Nach Caratheodory muß zur Fortsetzung auf $\sigma\left(r(\mathscr{R}_n)\right)$ gezeigt werden, daß P_n σ-additiv auf dem Ring ist. Da P_n endlicher Inhalt ist, ist die σ-Additivität äquivalent zur „Stetigkeit von oben“, die im folgenden gezeigt wird; wegen Rekursion und Bemerkung 1.2 genügt es dabei, den Fall $n = 1$ zu betrachten:
Sei $(A_k)_{k\in\mathbb{N}}$ eine Folge in $r(\mathscr{R}_1)$ mit $A_k \searrow \emptyset$, so ist zu zeigen: $P_1(A_k) \xrightarrow{k\to\infty} 0$.
Bezeichnet man den Schnitt durch ein $A \in r(\mathscr{R}_1)$ bei $x \in S$ mit

$$A_x := \{y \in S : (x, y) \in A\} ,$$

so folgt wegen $A_k \searrow \emptyset$ für alle $x \in S$:

$$(A_k)_x \searrow \emptyset \quad (k \to \infty) .$$

Aufgrund der „Stetigkeit von oben“ des Maßes $p_1(x, \,.\,)$ folgt daraus

$$p_1(x, (A_k)_x) \xrightarrow{k\to\infty} 0 \qquad (x \in S)$$

und somit wegen majorierter Konvergenz:

$$P_1(A_k) = \int_S p_1(x, (A_k)_x)\ \mu(dx) \xrightarrow{k\to\infty} 0 .$$

□

Das nächste Ziel ist nun, eine Markovkette auf $S^{\mathbb{N}_0}$ mit Übergangswahrscheinlichkeiten $(p_n)_{n\in\mathbb{N}}$ und Startwahrscheinlichkeit μ zu konstruieren; hierbei sei S weiterhin polnisch, versehen mit der Borel-σ-Algebra $\mathscr{B}(S) =: \mathscr{S}$. Dazu wird die Konsistenzbedingung von Kolmogorov für $(P_n)_{n\in\mathbb{N}_0}$ nachgewiesen. Hierfür dienen folgende Definitionen:
Zu $F, G \subset \mathbb{N}_0$ mit $F \subset G$ sei

$$\begin{aligned} \pi_{G,F} \ : \ S^G &\longrightarrow S^F \\ (x_i)_{i\in G} &\longmapsto (x_i)_{i\in F} \end{aligned}$$

die Projektion auf die die kleinere Indexmenge und hiermit $\pi_F := \pi_{\mathbb{N}_0,F}$; entsprechend setze zu $m, n \in \mathbb{N}_0$ mit $m \leq n$

$$\begin{aligned} \pi_{n,m} \ : \ S^{n+1} &\longrightarrow S^{m+1} \\ (x_0, \ldots, x_n) &\longmapsto (x_0, \ldots, x_m) \end{aligned}$$

und zu $m \in \mathbb{N}_0$

$$\begin{aligned} \pi_m \ : \ S^{\mathbb{N}_0} &\longrightarrow S^{m+1} \\ (x_i)_{i\in\mathbb{N}_0} &\longmapsto (x_0, \ldots, x_m) . \end{aligned}$$

Diese Projektionen sind bezüglich den jeweiligen Produkt-σ-Algebren meßbar. Der Maßeindeutigkeitssatz (angewandt auf $\cap$-stabile Erzeuger der σ-Algebren bestehend aus Zylindermengen) ergibt für $m, n \in \mathbb{N}_0$ mit $m \leq n$ die Gleichheit

$$P_n \circ \pi_{n,m}^{-1} \;=\; P_m$$

von Maßen auf $\mathscr{S}^{m+1}$; entsprechend gilt für endliche $F, G \subset \mathbb{N}_0$ mit $F \subset G$ auch

$$P_G \circ \pi_{G,F}^{-1} \;=\; P_{\max F} \circ \pi^{-1}_{\{0,\dots,\max F\},F} \;=:\; P_F\,;$$

diese Konsistenzeigenschaft besagt, daß $(P_F)_{F \subset \mathbb{N}_0 \text{ endl.}}$ ein Promaß auf $(S^{\mathbb{N}_0}, \mathscr{B}(S)^{\mathbb{N}_0})$ definiert. Nach dem Konsistenzsatz von Kolmogorov ist es sogar σ-additiv. Es existiert daher ein eindeutiges Wahrscheinlichkeitsmaß P_μ auf $(S^{\mathbb{N}_0}, \mathscr{B}(S)^{\mathbb{N}_0})$ mit

$$P_\mu \circ \pi_n^{-1} \;=\; P_n \qquad (n \in \mathbb{N}_0)\,. \tag{1}$$

Satz 1.4 (kanonische Markovkette). *Sei S ein polnischer Raum mit Übergangswahrscheinlichkeiten $(p_n)_{n\in\mathbb{N}}$ und Wahrscheinlichkeitsmaß μ; P_μ sei das hiervon induzierte Wahrscheinlichkeitsmaß auf $S^{\mathbb{N}_0}$. Dann ist*

$$X_n \;:=\; \pi_{\{n\}} \;\equiv\; \pi_{\mathbb{N}_0,\{n\}} \qquad (n \in \mathbb{N}_0)$$

eine Markovkette *auf*

$$(\Omega, \mathscr{F}, \mathbb{P}, (\mathscr{F}_n)_{n\in\mathbb{N}_0}) \;:=\; \Big(S^{\mathbb{N}_0}, \mathscr{S}^{\mathbb{N}_0}, P_\mu, (\sigma(\pi_n))_{n\in\mathbb{N}_0}\Big)\,,$$

d.h. es gilt:

i) X_n ist $\mathscr{F}_n$-meßbar, und

ii) für alle $n \in \mathbb{N}_0$ und $B \in \mathscr{S}$ gilt:

$$\mathbb{P}\,(X_{n+1} \in B \,|\, \mathscr{F}_n) \;=\; \mathbb{P}\,(X_{n+1} \in B \,|\, X_n) \;=\; p_{n+1}(X_n, B)\,.$$

Beweis. i) X_n ist meßbar bezüglich $\sigma(X_n) \subset \sigma(X_0, \dots, X_n) \equiv \sigma(\pi_n) \equiv \mathscr{F}_n$.
ii) Es ist zu zeigen:

$$\int_A 1_{\{X_{n+1}\in B\}}\, dP_\mu \;=\; \int_A p_{n+1}(X_n, B)\, dP_\mu \qquad (A \in \mathscr{F}_n)$$

(dann auch $\mathbb{P}(X_{n+1} \in B | X_n) = p_{n+1}(X_n, B)$, weil meßbar bzgl. $\sigma(X_n) \subset \mathscr{F}_n$). Da $\pi_n^{-1}(\mathscr{R}_n)$ ein $\cap$-stabiler Erzeuger von $\mathscr{F}_n$ ist, genügt es, obige Gleichung für

$$A \;=\; \pi_n^{-1}\,(B_0 \times \cdots \times B_n) \;\equiv\; \{X_0 \in B_0, \dots, X_n \in B_n\}$$

mit $B_0, \dots, B_n \in \mathscr{S}$ nachzuweisen, nämlich:

$$\begin{aligned}
\int_A 1_{\{X_{n+1}\in B\}}\, dP_\mu \quad &= \quad P_\mu\{X_0 \in B_0, \dots, X_n \in B_n, X_{n+1} \in B\} \\
&\overset{(1)}{=} \quad P_{n+1}(B_0 \times \cdots \times B_n \times B) \\
&\overset{1.3}{=} \quad \int_{B_0\times\cdots\times B_n} p_{n+1}(x_n, B)\, P_n(dx_0, \dots, dx_n) \\
&\overset{\text{Trafo.satz}}{=} \quad \int_A p_{n+1}(X_n, B)\, dP_\mu\,.
\end{aligned}$$

□

Definition 1.5. Sei $(S, \mathscr{S})$ ein meßbarer Raum. Dann ist auf dem Pfadraum $\Omega \equiv S^{\mathbb{N}_0}$ die Familie $\theta \equiv (\theta_n)_{n\in\mathbb{N}_0}$ der *(kanonischen) Shifts* $\theta_n : \Omega \longrightarrow \Omega$ $(n \in \mathbb{N}_0)$ definiert durch

$$\theta_n(\omega) := \big(m \mapsto \omega(m+n)\big) .$$

Jedes θ_n ist meßbar bezüglich $\mathscr{F} \equiv \mathscr{S}^{\mathbb{N}_0}$.

Als nächstes wird die Markov-Eigenschaft (mit festen Zeiten) und hiermit die Starke Markov-Eigenschaft (mit Stoppzeiten) gezeigt. Hierbei bezeichnen $\mathbb{E}_\mu$ bzw. $\mathbb{E}_x \equiv \mathbb{E}_{\delta_x}$ die bezüglich P_μ bzw. P_{δ_x} auf Ω gebildeten Erwartungswerte bei zugrundegelegten Übergangswahrscheinlichkeiten $(p_n)_{n\in\mathbb{N}}$. Als Vereinfachung wird die Markov-Kette als zeitlich homogen vorausgesetzt:

Definition 1.6. In der Situation von Satz 1.4 heißt die Markov-Kette X *zeitlich-homogen*, falls für alle $n \in \mathbb{N}$ gilt: $p_n = p_1 (=: p)$.

Theorem 1.7 (Markov-Eigenschaft). *In der Situation aus 1.4 sei die Markov-Kette X zeitlich-homogen; Y sei eine beschränkte, $\mathscr{F}$-meßbare Zufallsvariable auf Ω. Dann gilt:*

$$\mathbb{E}_\mu(Y \circ \theta_n \mid \mathscr{F}_n) \;=\; \mathbb{E}_{X_n}(Y) \;\equiv\; \mathbb{E}_x(Y)\Big|_{x=X_n} \qquad (n \in \mathbb{N}_0) .$$

Beweis. Zunächst ist zu bemerken, daß $\mathbb{E}_{X_n}(Y)$ tatsächlich meßbar bzgl. $\mathscr{F}_n$ ist; dies folgt aus der Adaptiertheit von X und der Meßbarkeit von $x \mapsto \mathbb{E}_x(Y)$ [letztere ist nach Definition und rekursiver Anwendung von 1.2 i) klar für Indikatorfunktionen $Y = 1_{\pi_n^{-1}[B_0\times\cdots\times B_n]}$ zu $B_i \in \mathscr{S}$; die allgemeine Aussage ergibt sich aus dem Monotone-Klassen-Theorem, da wegen des Satzes über monotone Konvergenz $\{Y : x \mapsto \mathbb{E}_x(Y) \text{ meßbar}\}$ abgeschlossen bzgl. monotonen Operationen ist]. Es bleibt also, die behauptete Gleichheit nachzuweisen. Aufgrund des Monotone-Klassen-Theorems genügt es, dies für den Fall zu zeigen, daß Y von der Form $\prod_{k=0}^{m} g_k(X_k)$ ist mit beschränkten, $\mathscr{S}$-meßbaren ZVn $g_0, \ldots, g_m$.

1) Zunächst betrachten wir die Mengen aus $\mathscr{F}_n$ der Gestalt $A := \pi_n^{-1}[A_0 \times \cdots \times A_n]$ mit $A_0, \ldots, A_n \in \mathscr{S}$; hiermit gilt:

$$\begin{aligned}
\mathbb{E}_\mu(Y \circ \theta_n \cdot 1_A) \quad &\equiv \quad \mathbb{E}_\mu\left(\prod_{k=0}^{m} g_k(X_{n+k}) \cdot 1_A\right) \\
&\overset{(1),1.3}{=} \quad \int_{A_0} \mu(dx_0) \int_{A_1} p(x_0, dx_1) \cdots \int_{A_n} p(x_{n-1}, dx_n) \;\times \\
&\qquad \times \int_S g_0(x_{n+1})\, p(x_n, dx_{n+1}) \cdots \int_S g_m(x_{n+m})\, p(x_{n+m-1}, dx_{n+m}) \\
&\overset{\text{Trafo.satz}}{=} \quad \mathbb{E}_\mu\left(\mathbb{E}_{X_n}\left(\prod_{k=0}^{m} g_k(X_k)\right) \cdot 1_A\right) \\
&\equiv \quad \mathbb{E}_\mu\big(\mathbb{E}_{X_n}(Y) \cdot 1_A\big) ,
\end{aligned}$$

also die Behauptung für alle $A \in \mathscr{F}_n$, die von der speziellen, obigen Gestalt sind.

2) Sei nun $\mathscr{L} := \big\{A \in \mathscr{F}_n : \text{Aussage aus 1) gilt für } A\big\}$.Gemäß 1) ist $\pi_n^{-1}(\mathscr{R}_n) \subset \mathscr{L}$; da $\pi_n^{-1}(\mathscr{R}_n)$ $\cap$-stabil ist, folgt aufgrund des Dynkin-Lemmas $\mathscr{F}_n = \sigma(\pi_n^{-1}(\mathscr{R}_n)) \subset \mathscr{L}$. $\square$

Nächstes Ziel ist, die Markov-Eigenschaft auf Stoppzeiten auszudehnen.

Definition 1.8. Sei $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\in\mathbb{N}_0})$ ein filtrierter Meßraum; $N : \Omega \to \mathbb{N}_0 \cup \{\infty\}$ heißt $(\mathscr{F}_n)_{n\in\mathbb{N}_0}$-*Stoppzeit*, falls $\{N \le n\} \in \mathscr{F}_n$ für alle $n \in \mathbb{N}_0$ ist. Hierzu äquivalent ist, daß $\{N = n\} \in \mathscr{F}_n$ für alle $n \in \mathbb{N}_0$ gilt.

Zu einer $(\mathscr{F}_n)_n$-Stoppzeit N hat man die σ-Algebra

$$\mathscr{F}_N \ := \ \left\{ A \in \mathscr{F} \ : \ A \cap \left\{ N \overset{(=)}{\le} n \right\} \in \mathscr{F}_n \text{ für alle } n \in \mathbb{N}_0 \right\} ;$$

sie heißt N-*Vergangenheit* oder σ-*Algebra der Ereignisse vor* N.

In der Situation aus 1.4 und 1.5 erweitert man nun formal Ω um $\Delta \notin \Omega$, nimmt $\{\Delta\}$ zu $\mathscr{F}$ hinzu und setzt für eine $(\mathscr{F}_n)_{n\in\mathbb{N}_0}$-Stoppzeit N

$$\theta_N(\omega) \ := \ \begin{cases} \theta_{N(\omega)}(\omega) & , N(\omega) < \infty \\ \Delta & , N(\omega) = \infty \, . \end{cases}$$

Für eine Zufallsvariable Y auf Ω sei $Y(\Delta) := 0$.

Theorem 1.9 (Starke Markov-Eigenschaft). *In der Situation aus 1.4 sei die Markov-Kette X zeitlich-homogen; θ sei der Shift aus 1.5 und N eine $(\mathscr{F}_n)_n$-Stoppzeit. Ist dann eine Familie $(Y_n)_{n\in\mathbb{N}_0}$ $\mathscr{F}$-meßbarer und (gleichmäßig in (n,ω)) beschränkter Zufallsvariable gegeben, so gilt:*

$$\mathbb{E}_\mu(Y_N \circ \theta_N \mid \mathscr{F}_N) \ = \ \mathbb{E}_{X_N}(Y_N) \qquad \textit{auf } \{N < \infty\} \, ;$$

Speziell gilt für eine $\mathscr{F}$-meßbare beschränkte Zufallsvariable Y :

$$\mathbb{E}_\mu(Y \circ \theta_N \mid \mathscr{F}_N) \ = \ \mathbb{E}_{X_N}(Y) \qquad \textit{auf } \{N < \infty\} \, .$$

Beweis. Zunächst ist zu bemerken, daß $\omega \mapsto \mathbb{E}_{X_{N(\omega)}(\omega)}(Y_{N(\omega)})$ tatsächlich $\mathscr{F}_N$-meßbar ist, da sie die Komposition der meßbaren Abbildungen $\omega \mapsto (\omega, N(\omega))$, $(\omega, n) \mapsto (X_n(\omega), n)$ und $(x, n) \mapsto \mathbb{E}_x(Y_n)$ ist.
Mit $A \in \mathscr{F}_N$ gilt dann:

$$\begin{aligned} \mathbb{E}_\mu\left(Y_N \circ \theta_N \cdot 1_{A\cap\{N<\infty\}}\right) &\overset{\text{maj.Kvgz.}}{=} \sum_{n=0}^{\infty} \mathbb{E}_\mu\left(Y_n \circ \theta_n \cdot 1_{A\cap\{N=n\}}\right) \\ &\overset{\text{ME 1.7}}{=} \sum_{n=0}^{\infty} \mathbb{E}_\mu\left(\mathbb{E}_{X_n}(Y_n) \cdot 1_{A\cap\{N=n\}}\right) \\ &\overset{\text{maj.Kvgz.}}{=} \mathbb{E}_\mu\left(\mathbb{E}_{X_N}(Y_N) \cdot 1_{A\cap\{N<\infty\}}\right) \, . \end{aligned}$$

□

Als nächstes sollen invariante Maße einer Markov-Kette betrachtet werden. Dies ist stark korreliert mit deren Wiederkehreigenschaften. Als zweite Vereinfachung wird im folgenden angenommen, daß S abzählbar ist; für die Darstellung mit allgemeinem polnischen Zustandsraum S siehe Meyn & Tweedie [M-T 93].

Im folgenden sei

$$T_y \;:=\; \inf\{\, n \in \mathbb{N} \,:\, X_n = y \,\} \qquad (y \in S)\,,$$

die erste Treffzeit von y und hiermit

$$\rho_{xy} \;:=\; P_x(T_y < \infty) \qquad (x, y \in S)\,.$$

$y \in S$ heißt $\left\{ \begin{array}{c} \textit{rekurrent} \\ \textit{transient} \end{array} \right\}$, falls $\left\{ \begin{array}{c} \rho_{yy} = 1 \\ \rho_{yy} < 1 \end{array} \right\}$ ist. Ist jeder Zustand rekurrent, so heißt auch die Markov-Kette *rekurrent*. Die Anzahl der Besuche in y,

$$H_y \;:=\; \sum_{n=1}^{\infty} 1_{\{X_n = y\}}$$

charakterisiert Rekurrenz und Transienz von y folgendermaßen:

Theorem 1.10 (Transienz und Rekurrenz). *Dei Markov-Kette X aus 1.4 sei zeitlich-homogen mit abzählbarem Zustandsraum S. Dann gilt für $y \in S$:*

$$\begin{array}{lcl} y \textit{ transient} & \Longrightarrow & \mathbb{E}_x(H_y) \;=\; \dfrac{\rho_{xy}}{1-\rho_{yy}} \;<\; \infty \quad (\forall\, x \in S)\,, \\ y \textit{ rekurrent} & \Longleftrightarrow & \mathbb{E}_y(H_y) \;=\; \infty\,. \end{array}$$

Beweis. Zu $k \in \mathbb{N}$ sei T_y^k die Zeit des k-ten Besuches in y. Hiermit gilt:

$$P_x(T_y^k < \infty) \;=\; \rho_{xy} \cdot \rho_{yy}^{k-1} \qquad (x \in S, k \in \mathbb{N})\,; \tag{$\star$}$$

für $k = 1$ ist dies gerade die Definition von ρ_{xy}; für $k > 1$ folgt dies induktiv:

$$\begin{aligned} P_x(T_y^k < \infty) &= P_x\left(T_y^{k-1} < \infty\,,\, T_y \circ \theta_{T_y^{k-1}} < \infty\right) \\ &= \mathbb{E}_x\left(1_{\{T_y^{k-1}<\infty\}}\; \underbrace{\mathbb{E}_x\left(1_{\left\{T_y \circ \theta_{T_y^{k-1}} < \infty\right\}} \,\middle|\, \mathscr{F}_{T_y^{k-1}} \right)}_{\overset{\text{st.ME 1.9}}{\equiv}\; \mathbb{E}_y\left(1_{\{T_y<\infty\}}\right) \;\equiv\; \rho_{yy}} \right) \\ &= \rho_{yy} \;\cdot\; P_x(T_y^{k-1} < \infty) \\ &\overset{\text{I.V.}}{=} \rho_{xy}\rho_{yy}^{k-1}\,. \end{aligned}$$

Hiermit folgt:

$$\mathbb{E}_x(H_y) \;=\; \sum_{n=1}^{\infty} P_x \underbrace{\{H_y \geq n\}}_{\{T_y^n<\infty\}} \;\overset{(\star)}{=}\; \rho_{xy} \sum_{n=1}^{\infty} \rho_{yy}^{n-1} \;=\; \frac{\rho_{xy}}{1-\rho_{yy}}\;;$$

die geometrische Reihe konvergiert für $\rho_{yy} < 1$, und divergiert bei $\rho_{yy} = 1$. $\square$

Als nächstes wird gezeigt, daß Rekurrenz „ansteckend“ ist:

Theorem 1.11. *Die Markov-Kette X aus 1.4 sei zeitlich-homogen mit abzählbarem S. Ist $x \in S$ rekurrent und $\rho_{xy} > 0$ mit einem $y \in S$, so ist auch y rekurrent und es gilt $\rho_{yx} = 1$.*

Beweis. Aufgrund der Rekurrenz von x gilt:

$$\begin{aligned} 0 \;=\; P_x(T_x = \infty) \;&\geq\; P_x\big(T_y < \infty\,,\, T_x \circ \theta_{T_y} = \infty\big) \\ &=\; \mathbb{E}_x\Bigg(1_{\{T_y<\infty\}} \quad \underbrace{\mathbb{E}_x\Big(1_{\{T_x\circ\theta_{T_y}=\infty\}}\,\Big|\,\mathscr{F}_{T_y}\Big)}_{\overset{\text{st.ME 1.9}}{=}\ \mathbb{E}_y\left(1_{\{T_x=\infty\}}\right)\ =\ (1-\rho_{yx})}\Bigg) \\ &=\; \rho_{xy}\,(1-\rho_{yx}) \ ; \end{aligned}$$

Da $\rho_{xy} > 0$ vorausgesetzt war, folgt: $\rho_{yx} = 1$.
Hiermit ergibt sich noch die Rekurrenz von y: Wegen $\rho_{xy} > 0$ und $\rho_{yx} = 1$ existieren $k_1, k_2 \in \mathbb{N}$ mit

$$P_x(X_{k_1} = y) > 0 \quad \text{und} \quad P_y(X_{k_2} = x) > 0\,.$$

Aufgrund der Chapman-Kolmogorov-Gleichung hat man für $n \in \mathbb{N}$:

$$P_y(X_{n+k_1+k_2} = y) \;\geq\; P_y(X_{k_2} = x)\, P_x(X_n = x)\, P_x(X_{k_1} = y)\,,$$

also

$$\mathbb{E}_y(H_y) \;=\; \sum_{n=1}^{\infty} P_y(X_n = y) \;\geq\; \underbrace{P_y(X_{k_2} = x)}_{>0}\; \underbrace{\mathbb{E}_x(H_x)}_{\overset{1.10}{=}\infty}\; \underbrace{P_x(X_{k_1} = y)}_{>0}\,.$$

Also ist auch $\mathbb{E}_y(H_y) = \infty$ und y rekurrent gemäß 1.10. □

Demnach ist die Menge der rekurrenten Zustände in Klassen eingeteilt: Für $x, y \in S$ sei

$$x \sim y \;\;:\Longleftrightarrow\;\; \big(x = y \;\text{ oder }\; (\rho_{xy} > 0 \text{ und } \rho_{yx} > 0)\big)\,.$$

Theorem 1.12. *Die Markov-Kette X aus 1.4 sei zeitlich-homogen mit abzählbarem S. Dann zerfällt die Menge der rekurrenten Punkte $R := \{x \in S : \rho_{xx} = 1\}$ in eine Familie $(R_i)_{i\in I}$ paarweise disjunkter rekurrenter Klassen, die Äquivalenzklassen von $\sim$.*

Beweis. Es ist zu zeigen, daß $\sim$ eine Äquivalenzrelation ist: Die Reflexivität und Symmetrie dieser Relation folgen unmittelbar aus obiger Definition, sodaß nur noch die Transienz von $\sim$ nachzuweisen ist:
Sind also $x, y, z \in R$ fixiert, so ist zu zeigen, daß mit $x \sim y$ und $y \sim z$ auch $x \sim z$ gilt. Dabei sei oE $x \neq y$ und $x \neq z$; nach obiger Definition von $\sim$ gilt also $\rho_{xy} > 0$ und $\rho_{yz} > 0$. Wendet man wie in den Beweisen von 1.10 und 1.11 die starke Markov-Eigenschaft 1.9 an, so folgt hiermit:

$$\rho_{xz} \;\equiv\; P_x(T_z < \infty) \;\geq\; P_x(T_y < \infty\,,\, T_z \circ \theta_{T_y} < \infty) \;=\; \rho_{xy}\,\rho_{yz} \;>\; 0\,,$$

woraus sich mit 1.11 ergibt $(x \in R)$: $\rho_{zx} = 1 > 0$, insgesamt also $x \sim z$. □

2. Invariante Maße und asymptotisches Verhalten

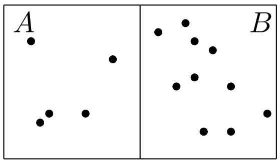
Es sei weiterhin folgende Situation zugrunde gelegt: Der abzählbare Raum S sei der Zustandsraum der kanonischen, zeitlich-homogenen Markov-Kette $(X_n)_{n\in\mathbb{N}_0}$ mit Pfadraum $(\Omega, \mathscr{F}) := (S^{\mathbb{N}_0}, \mathscr{S}^{\mathbb{N}_0})$ und Übergangsmatrix p.

Definition 2.1. Ein Maß μ auf $\mathscr{S}$ heißt *stationär*, wenn für alle $y \in S$ gilt:

$$\mu(y) \;=\; \big(\mu p\big)(y) \;\equiv\; \sum_{x\in S} \mu(x)\, p(x,y) \;<\; \infty\,.$$

Ein Maß μ auf $\mathscr{S}$ heißt *invariant*, wenn es ein stationäres Wahrscheinlichkeitsmaß ist.

Beispiel 2.2 (Ehrenfest-Modell von Diffusionen). In einem System, das aus den Behältern A und B besteht, befinden sich insgesamt r Moleküle. X_n sei die Anzahl der Moleküle in A zum Zeitpunkt $n \in \mathbb{N}_0$. Diese Größe nimmt also Werte in $S := \{0, 1, \ldots, r\}$ an. Durch



$$p(k,m) := \begin{cases} \frac{r-k}{r} & , \; m = k+1 \\ \frac{k}{r} & , \; m = k-1 \\ 0 & , \; \text{sonst} \end{cases}$$

wird eine Übergangswahrscheinlichkeit auf S definiert, die proportional zur Anzahl der Moleküle im Behälter A ist. Zu dieser Übergangsmatrix ist die Binomialverteilung auf S,

$$\mu(k) := \binom{r}{k} 2^{-r} \qquad (k \in S \equiv \{0, 1, \ldots, r\})\,,$$

ein invariantes Maß.

Beweis. Da μ ein Wahrscheinlichkeitsmaß ist, ist nur nachzuweisen, daß $\mu(k) = \sum_{m=0}^{r} p(m,k)\,\mu(m)$ für $k = 0, 1, \ldots, r$ gilt. Bei $k = 1, \ldots, r-1$ ist

$$\begin{aligned} \sum\nolimits_{m=0}^{r} p(m,k)\,\mu(m) &= p(k+1,k)\,\mu(k+1) \;+\; p(k-1,k)\,\mu(k-1) \\ &\equiv 2^{-r}\left[\binom{r}{k+1}\frac{k+1}{r} \;+\; \binom{r}{k-1}\frac{r-(k-1)}{r}\right] \\ &= 2^{-r}\left[\frac{(r-1)!}{k!\,(r-(k+1))!} \;+\; \frac{(r-1)!}{(k-1)!\,(r-k)!}\right] \\ &= 2^{-r}\frac{(r-1)!}{(k-1)!\,(r-k-1))!}\left[\frac{1}{k} \;+\; \frac{1}{r-k}\right] \\ &= 2^{-r}\frac{r!}{k!\,(r-k)!} \\ &\equiv \mu(k)\;. \end{aligned}$$

Bei den Fällen $k = 0$ und $k = r$ ist nur ein Summand ungleich null. $\square$

Nun wird gezeigt, wie jeder Klasse rekurrenter Zustände ein stationäres Maß zugeordnet ist; die Markov-Kette entkoppelt also auf diesen Klassen. Dabei wird laufend benutzt:

$$P_x(X_n = y) \;=\; p^n(x,y) \qquad (x, y \in S\,;\, n \in \mathbb{N})\,,$$

wobei $p^n(x,y)$ das n-fache Matrixprodukt bezeichnet.

Theorem 2.3. *Sei x rekurrent und $T \equiv T_x := \inf\{n \in \mathbb{N} : X_n = x\}$ seine erste Treffzeit.*

$$\mu(y) := \mathbb{E}_x\left(\sum_{n=0}^{T-1} 1_{\{X_n=y\}}\right) = \sum_{n=0}^{\infty} P_x(X_n = y\,,\, T > n) \qquad (y \in S)$$

definiert dann ein stationäres Maß.

Beweis. Zuerst wird die Gleichheit $\mu p = \mu$ gezeigt; anschließend wird hiermit nachgewiesen, daß $\mu(y) < \infty$ für alle $y \in S$ gilt. Beachte, daß $\mu(x) = 1$.

(a) $\sum_{y\in S} \mu(y)\, p(y,z) = \mu(z)$ für alle $z \in S$:

1) Falls $z \neq x$ ist, so folgt mit der Markov-Eigenschaft 1.7:

$$\begin{aligned}
\sum_{y\in S} \mu(y)\, p(y,z) &\overset{\text{Fubini}}{=} \sum_{n=0}^{\infty}\sum_{y\in S} P_x(X_n = y\,,\, T > n)\ \cdot\ P_y(X_1 = z) \\
&\overset{\text{ME}}{=} \sum_{n=0}^{\infty}\sum_{y\in S} P_x(X_n = y\,,\, T > n\,,\, X_{n+1} = z) \\
&= \sum_{n=0}^{\infty} P_x(T > n\,,\, X_{n+1} = z) \\
&\overset{z\neq x}{=} \sum_{n=0}^{\infty} P_x(T > n+1\,,\, X_{n+1} = z) \\
&= \sum_{n=1}^{\infty} P_x(T > n\,,\, X_n = z) \\
&\overset{z\neq x}{=} \sum_{n=0}^{\infty} P_x(T > n\,,\, X_n = z) \qquad \equiv \qquad \mu(z)\,.
\end{aligned}$$

2) Falls $z = x$ ist, so folgt — wiederum mit der Markov-Eigenschaft 1.7:

$$\begin{aligned}
\sum_{y\in S} \mu(y)\, p(y,x) &\overset{\text{ME}}{=} \sum_{n=0}^{\infty}\sum_{y\in S} P_x(X_n = y\,,\, T > n\,,\, X_{n+1} = x) \\
&= \sum_{n=0}^{\infty} P_x(T = n+1) \quad = \quad \rho_{xx} \overset{x \text{ rek.}}{=} 1 \;=\; \mu(x)\,.
\end{aligned}$$

(b) $\mu(y) < \infty$ für alle $y \in S$:

1) Falls $\rho_{xy} > 0$: Iteriert man (a), so folgt: $\mu = \mu\, p^n$ für $n \in \mathbb{N}$ also

$$1 \;=\; \mu(x) \overset{(a)}{=} \big(\mu\, p^n\big)(x) \;=\; \sum\nolimits_{y\in S} \mu(y)\, p^n(y,x) \qquad (n \in \mathbb{N})\,.$$

Folglich muß notwendigerweise $\mu(y) < \infty$ sein, falls $p^n(y,x) > 0$ mit einem $n \in \mathbb{N}$ gilt; da $p^n(y,x) = P_y(X_n = x)$ ist, wird letzteres impliziert durch $\rho_{yx} \equiv P_y(T_x < \infty) > 0$, was aber im betrachteten Fall $\rho_{xy} > 0$ aufgrund der Rekurrenz von x aus 1.11 folgt (also $x \sim y$).

2) Ist $\rho_{xy} = 0$, so folgt aus der Definition von μ: $\mu(y) = 0\,(< \infty)$. □

Theorem 2.4 (Eindeutigkeit stationärer Maße). *$(X_n)_{n\in\mathbb{N}_0}$ sei* irreduzibel*, d.h. S bestehe aus einer einzigen Äquivalenzklasse rekurrenter Zustände. Dann ist das stationäre Maß μ aus Theorem 2.3 bis auf Multiplikation mit Konstanten eindeutig.*

Beweis. Sei $a \in S$ ein rekurrenter Zustand und μ das zu a gemäß 2.3 gebildete stationäre Maß. Bezeichnet ν ein weiteres stationäres Maß, so ist zu zeigen:

$$\nu(z) = \mu(z) \cdot \nu(a) \qquad (z \in S)\,.$$

Aus der Stationarität von ν folgt iterativ für $z \in S$:

$$\begin{aligned}
\nu(z) &= \sum_{y\in S} \nu(y)\, p(y,z) \\
&= \nu(a)\, p(a,z) + \sum_{y\neq a} \nu(y)\, p(y,z) \\
&= \nu(a)\, p(a,z) + \sum_{y\neq a} \left(\sum_{x\in S} \nu(x)\, p(x,y) \right) p(y,z) \\
&= \nu(a)\, p(a,z) + \sum_{y\neq a} \nu(a)\, p(a,y)\, p(y,z) + \sum_{y\neq a}\sum_{x\neq a} \nu(x)\, p(x,y)\, p(y,z) \\
&= \nu(a)\, P_a(X_1 = z) + \sum_{y\neq a} \nu(a)\, P_a(X_1 \neq a\,,\, X_2 = z) \\
&\qquad\qquad + P_\nu(X_0 \neq a\,,\, X_1 \neq a\,,\, X_2 = z) \\
&= \cdots = \\
&= \nu(a) \sum_{m=1}^{n} P_a(X_k \neq a \text{ für } 1 \leq k < m\,,\, X_m = z) \\
&\qquad\qquad + P_\nu(X_0 \neq a\,,\, X_1 \neq a\,,\, \ldots, X_{n-1} \neq a\,,\, X_n = z) \\
&\geq \nu(a) \cdot \mu(z)
\end{aligned}$$

$(n \to \infty)$ nach der Definition von μ; daher folgt für $n \in \mathbb{N}$:

$$\nu(a) = \sum_{z\in S} \nu(z)\, p^n(z,a) \geq \nu(a) \sum_{z\in S} \mu(z)\, p^n(z,a) = \nu(a)\, \mu(a) = \nu(a)\,.$$

In der davor erhaltenen Abschätzung $\nu(z) \geq \nu(a)\, \mu(z)$ kann „>“ also nur gelten, wenn $p^n(z,a) = 0$ für jedes $n \in \mathbb{N}$ ist. Aufgrund der Irreduzibilität existiert aber zu jedem z ein $n \in \mathbb{N}$ mit $p^n(z,a) > 0$. Daher ist $\nu(z) = \nu(a)\, \mu(z)$. □

Nun wird eine notwendige Bedingung für die Normierbarkeit stationärer Maße gegeben:

Satz 2.5. *Existiert ein invariantes Maß μ, so sind alle Zustände y mit $\mu(y) > 0$ rekurrent.*

Beweis. Für $n \in \mathbb{N}$ gilt wegen der Stationarität $\mu = \mu p^n$, also mit Fubini

$$\sum_{n=1}^{\infty} \mu(y) = \sum_{x\in S} \mu(x) \sum_{n=1}^{\infty} p^n(x,y) \overset{1.10}{=} \sum_{x\in S} \mu(x)\, \frac{\rho_{xy}}{1-\rho_{yy}} \leq \frac{\mu(S)}{1-\rho_{yy}}\,.$$

Nach Voraussetzung sind $\sum_{n=1}^{\infty} \mu(y) = \infty$ und $\mu(S) = 1 < \infty$, also $\rho_{yy} = 1$. □

Theorem 2.6. $(X_n)_{n\in\mathbb{N}_0}$ *sei irreduzibel und* μ *ein invariantes Maß. Dann gilt:*

$$\mu(x) \;=\; \frac{1}{\mathbb{E}_x(T_x)} \qquad (x \in S)\,.$$

Beweis. Zunächst ist zu bemerken, daß alle Elemente von S rekurrent sind: Denn jedes Element mit positiver μ-Masse ist gemäß 2.5 rekurrent; da aber X irreduzibel ist, überträgt sich diese Rekurrenz auch auf alle anderen Elemente. Folglich ist zu jedem fixierten $x \in S$ gemäß 2.3 ein stationäres Maß μ_0 gegeben:

$$\mu_0(z) \;\equiv\; \sum\nolimits_{n\in\mathbb{N}_0} P_x(X_n = z\,,\, T_x > n) \qquad \text{und} \qquad \mu_0(x) = 1\,.$$

Hieraus folgt mit Fubini:

$$\sum_{z\in S}\mu_0(z) \;=\; \sum_{n=0}^{\infty}\sum_{z\in S} P_x(X_n = z\,,\, T_x > n) \;=\; \sum_{n=0}^{\infty} P_x(T_x > n) \;=\; \mathbb{E}_x(T_x)\,.$$

Mit der Eindeutigkeitsaussage aus 2.4 heißt das für das normierte Maß μ:

$$\mu(y) \;=\; \frac{\mu_0(y)}{\sum_{z\in S}\mu_0(z)} \;=\; \frac{\mu_0(y)}{\mathbb{E}_x(T_x)} \qquad (y \in S)\,,$$

woraus bei $y = x$ die Behauptung folgt, da $\mu_0(x) = 1$ ist. □

$x \in S$ heißt *positiv rekurrent*, falls $\mathbb{E}_x(T_x) < \infty$ ist; andernfalls heißt x *null-rekurrent*.

„Postitiv rekurrent“ ist stärker als „rekurrent“. Positive und Null-Rekurrenz sind beide Klasseneigenschaften. Im Ehrenfest-Modell 2.2 ist jeder Zustand positiv rekurrent.

Korollar 2.7. $(X_n)_{n\in\mathbb{N}_0}$ *sei irreduzibel. Dann sind folgende Aussagen äquivalent:*

i) Es existiert ein invariantes Maß;

ii) Es existiert ein positiv rekurrenter Zustand;

iii) Alle Zustände sind positiv rekurrent.

Beweis. iii) $\Rightarrow$ ii) trivial.
ii) $\Rightarrow$ i) Sei x der positiv rekurrente Zustand. Gemäß 2.3 existiert ein stationäres Maß μ_0 mit Gesamtmasse $\mu_0(S) = \sum_{z\in S}\mu_0(z) = \mathbb{E}_x(T_x)$ (Beweis von 2.6), die wegen der postiven Rekurrenz endlich ist. Die Normierung μ ist also invariant:

$$\mu(y) \;:=\; \frac{\mu_0(y)}{\mathbb{E}_x(T_x)} \;\equiv\; \frac{1}{\mathbb{E}_x(T_x)}\sum\nolimits_{n\in\mathbb{N}_0} P_x(X_n = y\,,\, T_x > n) \qquad (y \in S)\,.$$

i) $\Rightarrow$ iii) Sei μ das invariante Maß. Wegen der Irreduzibilität ist $\mu(x) > 0$ für alle $x \in S$ (denn jeder Zustand x ist rekurrent, sodaß $\mu_0(x) = 1$ für das gemäß 2.3 zugeordnete stationäre Maß μ_0 gilt; wegen 2.4 muß daher $\mu(x) > 0$ sein). Aus 2.6 folgt insbesondere: $\mathbb{E}_x(T_x) = \frac{1}{\mu(x)} < \infty$ für jedes $x \in S$. □

Als nächstes wird diskutiert, wann p^n gegen das invariante Maß konvergiert.

Beispiel 2.8. Auf $S := \{1,2\}$ definiert $p := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ eine Übergangsmatrix. Dabei gilt:

$$p^{2n} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{und} \quad p^{2n+1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \equiv p \qquad (n \in \mathbb{N})\,.$$

In diesem Fall liegt keine Konvergenz von $p^n(x,y)$ vor.

Periodizität verhindert also Konvergenz gegen das invariantes Maß.

Definition 2.9. Zu einem rekurrenten $x \in S$ sei[2]

$$I_x \;:=\; \{n \in \mathbb{N}_0 \,:\, p^n(x,x) > 0\}\;.$$

Hiermit heißt $d_x := \operatorname{ggT}(I_x)$ die *Periode* von x.

Wegen der Chapman-Kolmogorov-Gleichung ist I_x eine Halbgruppe.
In obigem Beispiel 2.8 ist $I_1 = I_2 = \{\,\text{gerade Zahlen}\}$ und $d_1 = d_2 = 2$.

Lemma 2.10. *Es seien $x, y \in S$ rekurrent mit $x \sim y$. Dann ist $d_x = d_y$.*

Beweis. Es wird gezeigt[3]: $d_y \mid d_x$. Da die folgende Argumentation symmetrisch in x und y ist, folgt hieraus schon die Behauptung, denn nach Vertauschen der Rollen ist damit auch $d_x \mid d_y$ gezeigt.
Ohne Einschränkung gelte $x \neq y$. Aufgrund der Äquivalenz $x \sim y$ ist daher $\rho_{xy} > 0$ und $\rho_{yx} > 0$; insbesondere existieren $m, n \in \mathbb{N}$ mit $p^m(x,y) > 0$ und $p^n(y,x) > 0$. Aus den Chapman-Kolmogorov-Gleichungen folgt hieraus:

$$p^{n+m}(y,y) \;\geq\; p^n(y,x)\,p^m(x,y) \;>\; 0\,.$$

Aus obiger Definition ergibt sich daher $d_y \mid n+m\,.$
Sei nun ein beliebiges $k \in I_x$ fixiert; wegen des eben gezeigten Zwischenschrittes $d_y \mid n+m$ ist noch einzusehen, daß auch $d_y \mid n+m+k$ gilt, da aus diesen beiden Aussagen $d_y \mid k$ und damit die Behauptung folgt. Mit Chapman-Kolmogorov und $k \in I_x$ erhält man aber:

$$p^{n+k+m}(y,y) \;\geq\; p^n(y,x)\,p^k(x,x)\,p^m(x,y) \;>\; 0\,,$$

und damit $d_y \mid n+k+m\,.$ $\square$

Definition 2.11. (a) Eine Zustand $x \in S$ heißt *aperiodisch*, falls $d_x = 1$ gilt.
(b) Eine irreduzible, rekurrente Markovkette heißt *aperiodisch*, falls jeder Zustand aperiodisch ist.

Wie in obigem Beispiel angedeutet wird sich herausstellen, daß Aperiodizität ein Kriterium für die Konvergenz der Übergangswahrscheinlichkeiten gegen das invariante Maß ist. Der Beweis dieses Satzes wird vorbereitet durch folgendes Lemma:

[2] Erinnerung: $p^n(x,y) \equiv P_x(X_n = y)$ für $x, y \in S$ und $n \in \mathbb{N}_0$.
[3] Wie üblich ist „|" die Abkürzung für „teilt".

Lemma 2.12. *Ist x aperiodisch, so existiert $m_0 \in \mathbb{N}$ mit $p^m(x,x) > 0$ für alle $m \geq m_0$.*

Beweis. Zunächst wird gezeigt, daß es ein $N \in \mathbb{N}$ gibt mit $N, N+1 \in I_x$. Hierzu seien $n_0, n_0 + k \in I_x$ fixiert. Ist $k = 1$, so ist man fertig. Ist $k \geq 2$, so wähle man $n_1 \in I_x$ mit $k \nmid n_1$ (da $d_x = 1$). Hierfür hat man (Division mit Rest)

$$n_1 \;=\; m\,k + r_1 \qquad (m \in \mathbb{N}_0\,,\, 0 < r_1 < k)$$

und aufgrund der Halbgruppeneigenschaft von I_x

$$(m+1)(n_0+k) \in I_x \qquad \text{und} \qquad (m+1)n_0 + n_1 \in I_x\,.$$

Für diese beiden Elemente gilt:

$$\begin{aligned} \Big|(m+1)(n_0+k) \;-\; \big((m+1)n_0+n_1\big)\Big| \;&=\; |(m+1)k \;-\; n_1| \\ &\equiv\; |(m+1)k \;-\; (m\,k+r_1)| \;=\; k - r_1 \;<\; k\,. \end{aligned}$$

Ist $k - r_1 = 1$, so gilt die Zwischenbehauptung mit $N := (m+1)n_0 + n_1$. Ist $k - r_1 > 1$, so wiederhole man die Rekursion mit $\tilde{n}_0 := (m+1)n_0 + n_1$ und $\tilde{k} := k - r_1$, um in endlich vielen Schritten $N \in \mathbb{N}$ mit $N, N+1 \in I_x$ zu erhalten. Hieraus folgt nun das Lemma mit $m_0 := N^2$, denn für $m \geq m_0$ hat man

$$m - N^2 \;=\; k\,N + r \qquad (k \in \mathbb{N}_0\,,\, 0 \leq r < N)$$

(Division mit Rest), sodaß

$$m \;=\; N^2 + k\,N + r \;=\; (N - r + k)N + r(1+N) \;\in\; I_x$$

wegen der Halbgruppen-Eigenschaft von I_x gilt. □

Theorem 2.13 (Invariantes Maß ist Limes der Übergangswahrscheinlichkeiten). *Die Markov-Kette $(X_n)_{n\in\mathbb{N}_0}$ sei aperiodisch und besitze das invariante Maß μ . Dann gilt:*

$$p^n(x,y) \xrightarrow{n\to\infty} \mu(y) \;=\; \frac{1}{\mathbb{E}_y(T_y)} \qquad (x, y \in S)\,.$$

Beweis(Kopplung von Prozessen, W. Döblin). Auf $S^2 \equiv S \times S$ definiert

$$q\big((x_1,y_1)\,,\,(x_2,y_2)\big) \;:=\; p(x_1,x_2)\,p(y_1,y_2) \qquad (x_1,x_2,y_1,y_2 \in S)$$

eine Übergangswahrscheinlichkeit. Es sei $(X_n, Y_n)_{n\in\mathbb{N}_0}$ die zu q gehörige kanonische Markov-Kette, also die Markov-Kette in S^2 auf

$$(\Omega, \mathscr{F}, \mathbb{P}) \;:=\; \Big((S^2)^{\mathbb{N}_0},\, (\mathscr{S}^2)^{\mathbb{N}_0},\, P_\varrho\Big)\;,$$

wobei P_ϱ das zu q und einer Anfangsverteilung ϱ (auf $\mathscr{S}^2 \equiv \mathscr{S} \otimes \mathscr{S}$) nach Kolmogorov gebildete Wahrscheinlichkeitsmaß ist.
Mit 2.12 wird nun die Irreduzibilität von $(X_n, Y_n)_{n\in\mathbb{N}_0}$ gezeigt; hieraus ergibt sich, daß dieser gekoppelte Prozeß die Diagonale von S^2 in endlicher Zeit trifft, womit dann die Konvergenz hergeleitet wird:

1) $(X_n, Y_n)_{n\in\mathbb{N}_0}$ ist irreduzibel: Sind $x_1, x_2, y_1, y_2 \in S$ fixiert, so gibt es wegen der Irreduzibilität von X Zeitpunkte $k, l \in \mathbb{N}$ mit

$$p^k(x_1, x_2) \;>\; 0 \qquad \text{und} \qquad p^l(y_1, y_2) \;>\; 0\,.$$

Die Aperiodizität liefert gemäß 2.12 auch ein $m_0 \in \mathbb{N}$, sodaß für $m \geq m_0$ gilt

$$p^{m+l}(x_2, x_2) \;>\; 0 \qquad \text{und} \qquad p^{m+k}(y_2, y_2) \;>\; 0\,.$$

Also ist mit Chapman-Kolmogorov auch

$$\begin{aligned} q^{k+l+m}&\big((x_1, y_1)\,,\,(x_2, y_2)\big) \\ &\equiv\;\; p^{k+l+m}(x_1, x_2)\, p^{k+l+m}(y_1, y_2) \\ &\geq\;\; p^k(x_1, x_2)\, p^{m+l}(x_2, x_2)\;\; p^l(y_1, y_2)\, p^{m+k}(y_2, y_2) \;\;>\;\; 0\,. \end{aligned}$$

Daher besteht S^2 aus einer einzigen Äquivalenzklasse. Für die Irreduzibilität ist noch zu zeigen, daß alle Zustände in S^2 rekurrent sind. Gemäß 2.5 ist genügt hierfür ein q-invariantes Maß ν mit $\nu(x, y) > 0$ für alle $(x, y) \in S^2$. Nun ist aber

$$\nu(x, y) \;:=\; \mu(x)\,\mu(y) \qquad\qquad (x, y \in S)$$

ein q-invariantes Maß auf S^2 wegen der p-Invarianz von μ:

$$\begin{aligned} \sum_{(x_1,x_2)\in S^2} \nu(x_1, x_2)\, q((x_1, x_2), (y_1, y_2)) \;&\equiv\; \sum_{(x_1,x_2)} \mu(x_1)\,\mu(x_2)\, p(x_1, y_1)\, p(x_2, y_2) \\ =\;\; \sum_{x_1} \mu(x_1)\, p(x_1, y_1) \sum_{x_2} \mu(x_2)\, p(x_2, y_2) \;&=\; \mu(y_1)\,\mu(y_2) \;\equiv\; \nu(y_1, y_2) \end{aligned}$$

für $(y_1, y_2) \in S^2$; ferner ist $\nu(y_1, y_2) \;\equiv\; \mu(y_1)\,\mu(y_2) \overset{2.6}{=} \frac{1}{\mathbb{E}_{y_1}(T_{y_1})}\,\frac{1}{\mathbb{E}_{y_2}(T_{y_2})} \overset{2.7\text{ iii)}}{>} 0\,.$

2) Bezeichnet T die erste Treffzeit mit der Diagonalen $D := \{(x, x) : x \in S\}$,

$$T \;:=\; \inf\{\, n \in \mathbb{N} : (X_n, Y_n) \in D\,\}\,,$$

$T_{(x,x)}$ die Erstbesuchszeit in $(x, x) \in D$, so gilt zum einen $T \leq T_{(x,x)}$. Ist ϱ eine beliebige Anfangsverteilung auf S^2, so ist andererseits wegen der in 1) gezeigten Rekurrenz jedes $T_{(x,x)} < \infty$ P_ϱ - f.s.; insbesondere ist auch $T < \infty$ P_ϱ – f.s. .

X_n und Y_n haben auf $\{T \leq n\}$ dieselbe Verteilung ($n \in \mathbb{N}$), da für $y \in S$ gilt:

$$\begin{aligned} P_\varrho\,(X_n = y\,,\, T \leq n) \;&=\; \sum_{m=1}^{n} P_\varrho\,(T = m\,,\, X_n = y) \\ &=\; \sum_{m=1}^{n} \sum_{x\in S} P_\varrho\,(T = m\,,\, X_m = x\,,\, X_n = y) \;\;= \end{aligned}$$

$$
\begin{aligned}
&= \sum_{m=1}^{n} \sum_{x \in S} P_\varrho\left(X_n = y \,\middle|\, T = m\,,\, X_m = x\right)\; P_\varrho\left(T = m\,,\, X_m = x\right) \\
&\overset{\text{ME}}{=} \sum_{m=1}^{n} \sum_{x \in S} P_\varrho\left(X_n = y \,\middle|\, X_m = x\right)\; P_\varrho\left(T = m\,,\, X_m = x\right) \\
&= \sum_{m=1}^{n} \sum_{x \in S} P_\varrho\left(Y_n = y \,\middle|\, Y_m = x\right)\; P_\varrho\left(T = m\,,\, Y_m = x\right) \\
&= \;\ldots\;\ldots\;\ldots \overset{\text{ebenso}}{=} \; P_\varrho\left(Y_n = y\,,\, T \le n\right)\;,
\end{aligned}
$$

wobei einging, daß X und Y dieselbe Übergangswahrscheinlichkeit p besitzen.

3) Nun wird die Behauptung des Satzes nachgewiesen; hiezu zeigen wir folgende (stärkere) Konvergenz:

$$
\sum_{y \in S} |\, p^n(x,y) \;-\; \mu(y)\, | \xrightarrow{n \to \infty} 0
$$

für alle $x \in S$; die Gleichheit $\mu(y) \;=\; 1/\,\mathbb{E}_y(T_y)$ ist ja bereits wegen 2.6 klar.

Ist also ein beliebiges $x \in S$ gegeben, so fixieren wir hierzu das Anfangsmaß

$$
\varrho \;:=\; \delta_x \otimes \mu
$$

auf S^2 für den gekoppelten Prozeß. Hiermit gilt für alle $y \in S$

$$
\begin{aligned}
p^n(x,y) \;&=\; P_\varrho(X_n = y) \\
&=\; P_\varrho\left(X_n = y\,,\, T \le n\right) \;+\; P_\varrho\left(X_n = y\,,\, T > n\right) \\
&\overset{2)}{=}\; P_\varrho\left(Y_n = y\,,\, T \le n\right) \;+\; P_\varrho\left(X_n = y\,,\, T > n\right)
\end{aligned}
$$

wegen der in 2) gezeigten Gleichheit der Verteilungen, und

$$
\mu(y) \;=\; P_\varrho(Y_n = y) \;\equiv\; P_\varrho\left(Y_n = y\,,\, T \le n\right) \;+\; P_\varrho\left(Y_n = y\,,\, T > n\right)
$$

wegen der p-Invarianz von μ; also insgesamt

$$
\begin{aligned}
\sum_{y \in S} |\, p^n(x,y) \;-\; \mu(y)\, | \;&=\; \sum_{y \in S} |\, P_\varrho(X_n = y) \;-\; P_\varrho(Y_n = y)\, | \\
&=\; \sum_{y \in S} |\, P_\varrho\left(X_n = y\,,\, T > n\right) \;-\; P_\varrho\left(Y_n = y\,,\, T > n\right) | \\
&\le\; \sum_{y \in S} \left[\, P_\varrho\left(X_n = y\,,\, T > n\right) \;+\; P_\varrho\left(Y_n = y\,,\, T > n\right) \right] \\
&=\; 2\, P_\varrho\left(T > n\right) \quad \xrightarrow{n \to \infty} \quad 0\,,
\end{aligned}
$$

da T P_ϱ-f.s. endlich ist, wie in 2) gesehen. $\square$

3. Stationäre Prozesse

Im folgenden betrachten wir stochastische Prozesse $X = (X_n)_{n\in\mathbb{N}_0}$ auf einem fixierten Wahrscheinlichkeitsraum $(\Omega, \mathscr{F}, \mathbb{P})$ mit Werten in einem polnischen Raum S (versehen mit der Borel-σ-Algebra $\mathscr{S} := \mathscr{B}(S)$). Diese Familie von $\mathscr{F}$-$\mathscr{S}$-meßbaren Abbildungen faßt man auch auf als Zufallsfolge

$$X : \Omega \longrightarrow S^{\mathbb{N}_0}\,, \qquad \omega \mapsto (X_n(\omega))_{n\in\mathbb{N}_0}\,,$$

die $\mathscr{F}$-$\mathscr{S}^{\mathbb{N}_0}$-meßbar ist, mit der Produkt-σ-Algebra

$$\mathscr{S}^{\mathbb{N}_0} \;:=\; \sigma\Big(\bigcup\nolimits_{n\in\mathbb{N}_0} \pi_{\{n\}}^{-1}[B_n] \;:\; B_n \in \mathscr{S}\Big) \;=\; \sigma\Big(\bigcup\nolimits_{n\in\mathbb{N}_0} \pi_n^{-1}[B] \;:\; B \in \mathscr{S}^{n+1}\Big)\;;$$

dabei ist das zweite Erzeugendensystem $\cap$-stabil, das erste hingegen nicht. Das durch

$$P_X \;\equiv\; P_{(X_n)_{n\in\mathbb{N}_0}} \;:=\; \mathbb{P}\circ X^{-1}$$

definierte Maß auf $\mathscr{S}^{\mathbb{N}_0}$ ist die *Verteilung von X*.
Sind nur Verteilungseigenschaften relevant, so kann statt X ohne Einschränkung auch sein *kanonischer Repräsentant* $(Y)_n := \big(\pi_{\{n\}}\big)_n$ auf $(S^{\mathbb{N}_0}, \mathscr{S}^{\mathbb{N}_0}, P_X)$ betrachtet werden.

Definition 3.1. Ein stochastischer Prozeß $X = (X_n)_{n\in\mathbb{N}_0}$ heißt *stationär*, falls gilt:

$$P_{(X_n)_{n\in\mathbb{N}_0}} = P_{(X_{n+k})_{n\in\mathbb{N}_0}} \qquad (\forall\, k\in\mathbb{N})\,.$$

Ein stationärer Prozeß tritt also hinsichtlich seiner Verteilung „auf der Stelle“; dies wird im folgenden Lemma nochmals formuliert:

Lemma 3.2. *$X = (X_n)_{n\in\mathbb{N}_0}$ ist genau dann stationär, falls gilt:*

$$P_{(X_0,\dots,X_n)} = P_{(X_k,\dots,X_{k+n})} \qquad (k\in\mathbb{N}\,,\; n\in\mathbb{N}_0)\,.$$

Beweis. „$\Rightarrow$“ Für alle $k\in\mathbb{N}$, $n\in\mathbb{N}_0$ und $B\in\mathscr{S}^{n+1}$ gilt:

$$\begin{aligned} P_{(X_0,\dots,X_n)}(B) &\equiv \mathbb{P}\{\,(X_0,\dots,X_n)\in B\,\}\\ &= \mathbb{P}\{\,(X_m)_{m\in\mathbb{N}_0}\in \pi_n^{-1}(B)\,\}\\ &\overset{\text{stat}}{=} \mathbb{P}\{\,(X_{m+k})_{m\in\mathbb{N}_0}\in \pi_n^{-1}(B)\,\}\\ &= \mathbb{P}\{\,(X_k,\dots,X_{n+k})\in B\,\} \quad\equiv\quad P_{(X_k,\dots,X_{k+n})}(B)\,. \end{aligned}$$

„$\Leftarrow$“ Nach Voraussetzung gilt gerade für alle $k\in\mathbb{N}$, $n\in\mathbb{N}_0$ und $B\in\mathscr{S}^{n+1}$:

$$P_{(X_m)_{m\in\mathbb{N}_0}}\big(\pi_n^{-1}(B)\big) \;=\; P_{(X_{m+k})_{m\in\mathbb{N}_0}}\big(\pi_n^{-1}(B)\big)$$

(vgl. obige Rechnung). Da aber $\{\,\bigcup_{n\in\mathbb{N}_0}\pi_n^{-1}(B) \;:\; B\in\mathscr{S}^{n+1}\,\}$ ein $\cap$-stabiler Erzeuger von $\mathscr{S}^{\mathbb{N}_0}$ ist, folgt hieraus $P_{(X_m)_{m\in\mathbb{N}_0}} = P_{(X_{m+k})_{m\in\mathbb{N}_0}}$ mit dem Maßeindeutigkeitssatz. $\square$

Beispiel 3.3 (Markov-Kette mit Übergangswahrscheinlichkeit p). Sei $(X_n)_{n\in\mathbb{N}_0}$ eine Markovkette auf einem abzählberen Raum S (versehen mit $\mathscr{S} := \mathscr{B}(S) \equiv \mathfrak{P}(S)$) mit Übergangswahrscheinlichkeit p und invariantem Maß μ. Dann ist $(X_n)_{n\in\mathbb{N}_0}$ stationär auf $(\Omega, \mathscr{F}, \mathbb{P}) := (S^{\mathbb{N}_0}, \mathscr{S}^{\mathbb{N}_0}, P_\mu)$.

Beweis. Zunächst gilt für alle $B := B_0 \times B_1 \times \cdots \times B_n \in \mathscr{S}^{n+1}$:

$$\begin{aligned}
(P_\mu)_{(X_1,\dots,X_{n+1})}(B) \;&\equiv\; P_\mu\{\, X_1 \in B_0 \,,\, X_2 \in B_1 \,,\, \dots \,,\, X_{n+1} \in B_n \,\} \\
&= P_\mu\{\, X_0 \in S \,,\, X_1 \in B_0 \,,\, X_2 \in B_1 \,,\, \dots \,,\, X_{n+1} \in B_n \,\} \\
&= \sum_{z\in S} \mu(z) \sum_{x_0\in B_0} p(z,x_0) \sum_{x_1\in B_1} p(x_0,x_1) \cdots \sum_{x_n\in B_n} p(x_{n-1},x_n) \\
&= \sum_{x_0\in B_0} \sum_{z\in S} \mu(z)\, p(z,x_0) \sum_{x_1\in B_1} p(x_0,x_1) \cdots \sum_{x_n\in B_n} p(x_{n-1},x_n) \\
&\overset{\text{inv}}{=} \sum_{x_0\in B_0} \mu(x_0) \sum_{x_1\in B_1} p(x_0,x_1) \cdots \sum_{x_n\in B_n} p(x_{n-1},x_n) \\
&= P_\mu\{\, X_0 \in B_0 \,,\, X_1 \in B_1 \,,\, \dots \,,\, X_n \in B_n \,\} \\
&= (P_\mu)_{(X_0,X_1,\dots,X_n)}(B) \,.
\end{aligned}$$

Iteriert man dieses Argument k-mal, so erhält man das Kriterium für Stationarität aus 3.2 . □

Beispiel 3.4 (Rotation des Kreises). Sei $(\Omega, \mathscr{F}, \mathbb{P}) := ([0,1), \mathscr{B}[0,1), \lambda\big|_{\mathscr{F}})$, wobei λ das Lebesguemaß bezeichnet. Dann ist für jedes fixierte $\theta \in [0,1)$ der Prozeß $(X_n)_{n\in\mathbb{N}_0}$,

$$X_n \,:\, \Omega \longrightarrow S := \Omega \,, \qquad X_n(\omega) \;:=\; \omega + n\cdot\theta \pmod 1 \,, \qquad n \in \mathbb{N}_0 \,,$$

eine stationäre Markov-Kette auf $(S^{\mathbb{N}_0}, \mathscr{S}^{\mathbb{N}_0}, P_\lambda)$ bzgl. der Übergangswahrscheinlichkeit

$$p \,:\, S \times \mathscr{S} \longrightarrow [0,1] \,, \qquad p(x,y) \;:=\; \begin{cases} 1\,, & \text{falls } y = x + \theta \pmod 1 \\ 0\,, & \text{sonst.} \end{cases}$$

Beweis. Wegen der Translationsinvarianz des Lebesguemaßes ist λ p-invariant,

$$\lambda(dy) \;=\; \int_0^1 \lambda(dz)\, p(z,dy) \,.$$

Also folgt wie in 3.3 für alle $B := B_0 \times B_1 \times \cdots \times B_n \in \mathscr{S}^{n+1}$:

$$\begin{aligned}
(P_\lambda)_{(X_1,\dots,X_{n+1})}(B) \;&=\; P_\lambda\{\, X_0 \in S \,,\, X_1 \in B_0 \,,\, X_2 \in B_1 \,,\, \dots \,,\, X_{n+1} \in B_n \,\} \\
&= \int_\Omega \lambda(dz) \int_{B_0} p(z,dx_0) \int_{B_1} p(x_0,dx_1) \cdots \int_{B_n} p(x_{n-1},dx_n) \\
&= \int_{B_0} \int_\Omega \lambda(dz)\, p(z,dx_0) \int_{B_1} p(x_0,dx_1) \cdots \int_{B_n} p(x_{n-1},dx_n) \\
&\overset{\text{inv}}{=} \int_{B_0} \lambda(dx_0) \int_{B_1} p(x_0,dx_1) \cdots \int_{B_n} p(x_{n-1},dx_n) \\
&= (P_\lambda)_{(X_0,X_1,\dots,X_n)}(B) \,,
\end{aligned}$$

und damit die Stationarität wiederum aus 3.2 . □

Theorem 3.5. *Der Prozeß $(X_n)_{n\in\mathbb{N}_0}$ mit polnischem Zustandsraum $(S,\mathscr{S})$ sei stationär und $g : S^{\mathbb{N}_0} \longrightarrow S'$ sei $\mathscr{S}^{\mathbb{N}_0}$-$\mathscr{S}'$-meßbar, wobei $(S',\mathscr{S}')$ ebenfalls polnisch ist. Dann ist*

$$Y_k \;:=\; g\,(X_k\,,\,X_{k+1}\,,\,\dots) \qquad (k\in\mathbb{N}_0)$$

stationär (in S').

Beweis. Wegen der Meßbarkeit von g ist für jedes $k\in\mathbb{N}_0$ auch

$$g_k \,:\, S^{\mathbb{N}_0} \longrightarrow S'\,, \qquad x\mapsto g\circ\theta_k(x)$$

meßbar, wobei $\theta\equiv(\theta_k)_{k\in\mathbb{N}_0}$ wieder (siehe 1.5) den meßbaren Shift

$$\theta_k \,:\, S^{\mathbb{N}_0} \longrightarrow S^{\mathbb{N}_0}\,, \qquad (x_n)_n\mapsto(x_{n+k})_n$$

bezeichnet. Sei nun $B\in(\mathscr{S}')^{\mathbb{N}_0}$ fixiert; aufgrund der Meßbarkeit aller g_k ist $A:=(g_0,g_1,\dots)^{-1}(B)$ meßbar und wegen $Y_k=g_k\big((X_n)_n\big)$ folgt für $m\in\mathbb{N}$:

$$\begin{aligned} P_{(Y_k)_{k\in\mathbb{N}_0}}(B) &\equiv \mathbb{P}\big((Y_k)_{k\in\mathbb{N}_0}\in B\big) &&= \mathbb{P}\big((X_n)_{n\in\mathbb{N}_0}\in A\big)\\ &\overset{X\text{ stat}}{=} \mathbb{P}\big((X_{n+m})_{n\in\mathbb{N}_0}\in A\big) &&= \mathbb{P}\big((Y_{k+m})_{k\in\mathbb{N}_0}\in B\big)\\ &\equiv P_{(Y_{k+m})_{k\in\mathbb{N}_0}}(B)\,, \end{aligned}$$

also gerade die Stationarität von Y. □

Beispiel 3.6 (Bernoulli-Shift). Auf $(\Omega,\mathscr{F},\mathbb{P}):=([0,1),\mathscr{B}[0,1),\lambda\big|_{\mathscr{F}})$ ist $(Y_n)_{n\in\mathbb{N}_0}$,

$$Y_n \,:\, \Omega\longrightarrow\Omega\,, \qquad Y_n \;:=\; \begin{cases} \mathrm{id}_\Omega\,, & n=0\\ 2Y_{n-1}\ (\mathrm{mod}\ 1)\,, & n\in\mathbb{N}\,,\end{cases}$$

stationär.

Beweis. Sei $(X_n)_{n\in\mathbb{N}_0}$ eine Bernoulli-Folge zur Rate $\frac{1}{2}$, realisiert als Produktmaß $\tilde{\mathbb{P}}$ auf $\tilde\Omega:=\{0,1\}^{\mathbb{N}_0}$; $(X_n)_n$ sei also eine Folge von iid-ZVen in $S:=\{0,1\}$ mit $\tilde{\mathbb{P}}\{X_n=0\}=\tilde{\mathbb{P}}\{X_n=1\}=\frac{1}{2}$. Dann ist $(X_n)_n$ stationär. Ferner ist

$$g \,:\, \tilde\Omega\equiv\{0,1\}^{\mathbb{N}_0}\longrightarrow\Omega\equiv[0,1)\,, \qquad (x_n)_n\mapsto\sum\nolimits_{n=0}^\infty x_n\,2^{-n-1}\ (\mathrm{mod}\ 1)$$

meßbar und $\tilde{\mathbb{P}}\circ g^{-1}=\mathbb{P}$ (dyadische Intervalle lassen sich als Mengen der Bauart $\{X_0=i_0\,,\,\dots\,,\,X_k=i_k\}$ mit $i_0,\dots,i_k\in\{0,1\}$ schreiben). Wegen 3.5 ist nun

$$Z_k \;:=\; g(X_k\,,\,X_{k+1}\,,\,\dots) \qquad (k\in\mathbb{N}_0)$$

stationär; andererseits gilt:

$$\begin{aligned} 2\,Z_0 \;\equiv\; 2\,g(X_0\,,\,X_1\,,\,\dots) &= X_0+\sum\nolimits_{n=1}^\infty X_n\,2^{-n} \quad (\mathrm{mod}\ 1)\\ &= X_0+\sum\nolimits_{n=0}^\infty X_{n+1}\,2^{-(n+1)} \quad (\mathrm{mod}\ 1)\\ &= g(X_1\,,\,X_2\,,\,\dots)\;\equiv\;Z_1\,; \end{aligned}$$

iterativ erhält man: $2\,Z_{n-1}=Z_n\,(n\in\mathbb{N})$, sodaß mit Z auch Y stationär ist. □

Definition 3.7 (maßtreue Abbildung). Sei $(\Omega, \mathscr{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum. Eine $\mathscr{F}$-$\mathscr{F}$-meßbare Abbildung $\varphi : \Omega \to \Omega$ heißt *maßtreu*, falls gilt: $\mathbb{P} \circ \varphi^{-1} = \mathbb{P}$.

Bemerkung 3.8. Sei φ eine maßtreu auf $(\Omega, \mathscr{F}, \mathbb{P})$ und $X : \Omega \to S$ eine $\mathscr{F}$-$\mathscr{S}$-meßbare Abbildung mit Werten in einem polnischen Raum $(S, \mathscr{S})$. Dann ist $(X_n)_{n\in\mathbb{N}_0}$ mit

$$X_n := \begin{cases} X\,, & n = 0 \\ X \circ \varphi^n\,, & n \in \mathbb{N} \end{cases}$$

stationär.

Beweis. Für $B \in \mathscr{S}^{n+1}$ gilt:

$$\begin{aligned} P_{(X_0,\dots,X_n)}(B) &\equiv \mathbb{P}\big((X_0,\dots,X_n) \in B\big) \\ &\overset{\varphi \text{ m.t.}}{=} \mathbb{P}\big((X_0,\dots,X_n) \circ \varphi^k \in B\big) = P_{(X_k,\dots,X_{k+n})}(B)\,, \end{aligned}$$

sodaß die Stationarität aus 3.2 folgt. □

Die Situation in der vorhergehenden Bemerkung gibt nicht nur ein Beispiel für eine stationäre Folge, sondern schon den allgemeinen Fall:

Satz 3.9 (Standardmodell für stationäre Folgen).
Sei $(Y_n)_{n\in\mathbb{N}_0}$ stationär auf $(\Omega, \mathscr{F}, \mathbb{P})$ mit Werten in einem polnischen Raum $(S, \mathscr{S})$. Dann existiert ein Wahrscheinlichkeitsraum $(\Omega', \mathscr{F}', \mathbb{P}')$ mit einer maßtreuen Abbildung $\varphi : \Omega' \to \Omega'$ und einer ZV $X_0 : \Omega' \to S$ derart, daß mit $X_n := X_0 \circ \varphi^n$ $(n \in \mathbb{N})$ gilt:

$$P'_{(X_n)_{n\in\mathbb{N}_0}} = P_{(Y_n)_{n\in\mathbb{N}_0}}\,.$$

Beweis. Es seien $(\Omega', \mathscr{F}', \mathbb{P}') := (S^{\mathbb{N}_0}, \mathscr{S}^{\mathbb{N}_0}, P_{(Y_n)_{n\in\mathbb{N}_0}})$ und $X_0 := \pi_{\{0\}}$ (Projektion zur Zeit 0) sowie $\varphi := \theta_1$ (Shift). Wegen der Stationarität von Y ist φ maßtreu, denn für $A' \in \mathscr{F}'$ gilt:

$$\begin{aligned} P'\left(\varphi^{-1}(A')\right) &= \mathbb{P}\big((Y_n)_n \in \varphi^{-1}(A')\big) \\ &= \mathbb{P}\big((Y_{n+1})_n \in A'\big) \\ &\overset{Y \text{ stat}}{=} \mathbb{P}\big((Y_n)_n \in A'\big) = P'(A')\,. \end{aligned}$$

Die behauptete Gleichheit der Verteilungen gilt nach Definition von $\mathbb{P}'$. □

Definition 3.10 (invariant, ergodisch). Sei φ maßtreue Abbildung auf $(\Omega, \mathscr{F}, \mathbb{P})$.
$A \in \mathscr{F}$ heißt $\left\{ \begin{array}{l} \textit{invariant} \\ \textit{invariant im strengeren Sinn} \end{array} \right\}$, falls $\left\{ \begin{array}{l} \varphi^{-1}(A) = A \ \ \mathbb{P}\text{-f.s.} \\ \varphi^{-1}(A) = A \end{array} \right\}$ ist.

φ heißt *ergodisch*, falls für alle $A \in \mathscr{I} := \{$ invariante Mengen $\}$ gilt: $\mathbb{P}(A) \in \{0, 1\}$.

Bemerkung 3.11. i) $\mathscr{I}$ ist eine σ-Algebra (Unter-σ-Algebra von $\mathscr{F}$);

ii) Zu $A \in \mathscr{I}$ existiert eine streng invariante Menge $B \in \mathscr{F}$ mit $B = A$ $\mathbb{P}$-f.s.; (z.B. $B := \liminf_{n\to\infty} \varphi^{-n}(A)$)

iii) Zu $A \in \mathscr{I}$ existiert ein $B \in \mathscr{T} := \bigcap_{n=1}^{\infty} \sigma(X_n, X_{n+1}, \dots)$ mit[4] $B = A$ $\mathbb{P}$-f.s.; (z.B. wieder $B := \liminf_{n\to\infty} \varphi^{-n}(A)$, da $B = \varphi^{-k}(B) \in \sigma(X_k, X_{k+1}, \dots)$)

[4] $\mathscr{T}$ ist die σ-Algebra der *terminalen Ereignisse*;

Beispiel 3.12. $(X_n)_{n\in\mathbb{N}_0}$ seien unabhängige Zufallselemente in einem polnischen Raum S (oE auf dem Folgenraum definiert), d.h.:

$$\mathbb{P} \equiv P_X \;=\; \bigotimes_{n\in\mathbb{N}_0} P_{X_n}\,.$$

Dann gilt $\mathbb{P}(A) \in \{0,1\}$ für $A \in \mathscr{T}$; d.h. der Shift $\varphi := \theta_1$ ist ergodisch.

Beispiel 3.13 (Rotation des Kreises). Wie in 3.4 betrachten wir die Transformation

$$\varphi \,:\, \Omega \longrightarrow \Omega\,, \qquad \varphi(\omega) \;:=\; \omega + \theta \;\; (\mathrm{mod}\ 1)\,,$$

auf dem Wahrscheinlichkeitsraum $(\Omega, \mathscr{F}, \mathbb{P}) := ([0,1), \mathscr{B}[0,1), \lambda\big|_{\mathscr{F}})$, wobei λ das Lebesguemaß bezeichnet. Dann ist φ genau dann ergodisch, wenn θ irrational ist.

Beweis. „$\Rightarrow$" Sei θ rational, also $\theta = \frac{m}{n}$ mit natürlichen Zahlen $n \geq m \geq 1$. Ferner sei $B \in \mathscr{F} \equiv \mathscr{B}[0,1)$ mit $0 < \lambda(B) < \frac{1}{n}$. Dann ist $A := \bigcup_{k=1}^{m-1}(B + \frac{k}{n})$ invariant, aber $0 < \lambda(A) < 1$.
„$\Leftarrow$" Dies kann man mit einem Fourierreihen-Argument zeigen; siehe z.B. Shiryaev [Sh 95, p.408] oder auch Kallenberg [KB 97, p.174/9]. $\square$

Beispiel 3.14. Sei $(X_n)_{n\in\mathbb{N}_0}$ die kanonische Markovkette auf $S := \{1,2,3,4\}$ mit Übergangswahrscheinlichkeit

$$p \;:=\; \begin{pmatrix} \frac13 & \frac23 & 0 & 0 \\ \frac23 & \frac13 & 0 & 0 \\ 0 & 0 & \frac12 & \frac12 \\ 0 & 0 & \frac14 & \frac34 \end{pmatrix}$$

(p ist eine stochastische Matrix, da die Zeilensummen gleich 1 sind). Ein Maß μ auf S ist invariant, falls gilt:

$$\mu(j) \;=\; \sum_{i=1}^{4} p(i,j)\mu(i) \qquad (j = 1,2,3,4)\,.$$

Dies wird z.B. erfüllt durch die beiden Maße

$$\mu_0(1) \;=\; \mu_0(2) \;:=\; \frac{1}{2}\,,\;\; \mu_0(3) \;=\; \mu_0(4) \;:=\; 0$$

und

$$\mu_1(1) \;=\; \mu_1(2) \;:=\; 0\,,\;\; \mu_1(3) \;:=\; \frac{1}{3}\,,\;\; \mu_1(4) \;:=\; \frac{2}{3}\,.$$

Dann ist aber auch jedes

$$\mu_\beta \;:=\; (1-\beta)\mu_0 \;+\; \beta\mu_1 \qquad (0 \leq \beta \leq 1)$$

invariant. Bezüglich des kanonischen Shifts $\varphi := \theta_1$ gilt nun:

$$A \;:=\; \{X_n \in \{1,2\}, n \in \mathbb{N}_0\} \;\in\; \mathscr{I} \qquad \text{und} \qquad B \;:=\; \{X_n \in \{3,4\}, n \in \mathbb{N}_0\} \;\in\; \mathscr{I}\,.$$

Hiermit gilt weiter: $P_{\mu_\beta}(A) = 1-\beta$ und $P_{\mu_\beta}(B) = \beta$. Folglich ist φ genau dann ergodisch, wenn $\beta \in \{0,1\}$ ist.

Theorem 3.15. $(X_n)_{n\in\mathbb{N}_0}$ *mit polnischem Zustandsraum* $(S,\mathscr{S})$ *sei ergodisch und* $g : S^{\mathbb{N}_0} \longrightarrow S'$ *sei* $\mathscr{S}^{\mathbb{N}_0}$-$\mathscr{S}'$*-meßbar, wobei* $(S',\mathscr{S}')$ *ebenfalls polnisch ist. Dann ist*

$$Y_k := g\,(X_k\,,\,X_{k+1}\,,\,\ldots) \qquad (k\in\mathbb{N}_0)$$

ergodisch (in S'*).*

Beweis. Ohne Einschränkung sei wieder $(\Omega,\mathscr{F},\mathbb{P}) = (S^{\mathbb{N}_0},\mathscr{S}^{\mathbb{N}_0},P_{(X_n)_{n\in\mathbb{N}_0}})$ und $X_n = \pi_{\{n\}}$ (Projektion zur Zeit n) sowie $\varphi = \theta_1$ (Shift). Ebenso sei $(\Omega',\mathscr{F}',\mathbb{P}') = ((S')^{\mathbb{N}_0},(\mathscr{S}')^{\mathbb{N}_0},P_{(Y_n)_{n\in\mathbb{N}_0}})$ und $\varphi' = \theta_1$. Ferner bezeichnen $\mathscr{I}$ bzw. $\mathscr{I}'$ die zu φ bzw. φ' gehörigen Systeme invarianter Mengen.
Sei nun $A\in\mathscr{I}'$ fixiert; für $B := (g_0,g_1,\ldots)^{-1}(A)$ gilt dann:

$$\begin{aligned}\varphi^{-1}(B) &= (g_1,g_2,\ldots)^{-1}(A)\\ &= (g_0,g_1,\ldots)^{-1}\left((\varphi')^{-1}(A)\right)\\ &= (g_0,g_1,\ldots)^{-1}(A) \quad\equiv\quad B\,,\end{aligned}$$

also $B\in\mathscr{I}$. Wegen der Ergodizität von φ folgt also: $\mathbb{P}'(A)\equiv\mathbb{P}(B)\in\{0,1\}$. □

Beispiel 3.16 (Bernoulli-Shift). Wie in 3.6 betrachten wir iid-ZVn $(X_n)_n$ in $S := \{0,1\}$ mit $\mathbb{P}\{X_n=0\} = \mathbb{P}\{X_n=1\} = \frac{1}{2}$. Ferner sei

$$g\;:\;\{0,1\}^{\mathbb{N}_0} \longrightarrow [0,1)\,, \qquad (x_n)_n \mapsto \sum\nolimits_{n=0}^{\infty} x_n\,2^{-n-1}\ (\mathrm{mod}\ 1)\,.$$

Wegen 3.12 ist nun X ergodisch, also gemäß 3.15 auch

$$Y_k := g\,(X_k\,,\,X_{k+1}\,,\,\ldots) \qquad (k\in\mathbb{N}_0)\,.$$

4. Der Birkhoffsche Ergodensatz

Sei $(\Omega, \mathscr{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum mit einer maßtreuen Abbildung $\varphi : \Omega \to \Omega$ und einer ZV $X : \Omega \to \mathbb{R}$. Wir untersuchen nun das asymptotische Verhalten des durch

$$X_k \;:=\; X \circ \varphi^k \qquad (k \in \mathbb{N}_0)$$

definierten stochastischen Prozesses.

Theorem 4.1 (Ergodensatz, Birkhoff). *Sei $X \in L^1(\mathbb{P})$. Dann gilt $\mathbb{P}$-f.s. und in $L^1(\mathbb{P})$:*

$$\frac{1}{n} \sum_{k=0}^{n-1} X \circ \varphi^k \xrightarrow{n\to\infty} \mathbb{E}(X|\mathscr{I})\,.$$

Der Beweis stützt sich auch folgende Abschätzung:

Lemma 4.2 (Maximal-ergodisches Lemma, Hopf). *In der Situation aus 4.1 seien*

$$\begin{aligned} S_n \;&:=\; X_0 + \cdots + X_{n-1} \;\equiv\; \sum_{k=0}^{n-1} X \circ \varphi^k && (n \in \mathbb{N}) \;\; \textit{und} \\ M_n \;&:=\; \max\{0, S_1, \ldots, S_n\} && (n \in \mathbb{N}_0)\,. \end{aligned}$$

Dann gilt:

$$\mathbb{E}\left(X\, 1_{\{M_n>0\}} \right) \;\geq\; 0 \qquad (n \in \mathbb{N}_0)\,.$$

Beweis. Im Fall $n = 0$ ist nicht zu beweisen. Zuerst zeigen wir folgendes:

$$X\, 1_{\{M_n>0\}} \;\geq\; 1_{\{M_n>0\}}\, (M_n - M_n \circ \varphi) \qquad (n \in \mathbb{N})\,;$$

nach obigen Definitionen gilt $S_k - M_n \leq 0$ für alle $k \in \{1, \ldots, n\}$, also auch

$$X \;\geq\; X + (S_k - M_n) \circ \varphi \;=\; (X + S_k \circ \varphi) - M_n \circ \varphi \;\equiv\; S_{k+1} - M_n \circ \varphi$$

und damit

$$X \;\geq\; \max\{S_1, \ldots, S_n\} - M_n \circ \varphi\,;$$

insbesondere ist damit gezeigt:

$$\begin{aligned} X\, 1_{\{M_n>0\}} \;&\geq\; 1_{\{M_n>0\}} \max\{S_1, \ldots, S_n\} - 1_{\{M_n>0\}}\, M_n \circ \varphi \\ &=\; 1_{\{M_n>0\}}\, (M_n - M_n \circ \varphi) \qquad (n \in \mathbb{N})\,, \end{aligned}$$

also gerade obige Zwischenbehauptung. Hiermit ergibt sich aber gerade:

$$\begin{aligned} \mathbb{E}\left(X\, 1_{\{M_n>0\}} \right) \;&\geq\; \int_{\{M_n>0\}} (M_n - M_n \circ \varphi)\, d\mathbb{P} \\ &=\; \int (M_n - M_n \circ \varphi)\, d\mathbb{P} \;=\; 0\,, \end{aligned}$$

wobei zuletzt noch benutzt wurde, daß φ maßtreu ist. □

Beweis des Birkhoffschen Ergodensatzes 4.1 Ohne Einschränkung sei $\mathbb{E}(X|\mathscr{I}) = 0$; andernfalls betrachten wir $\tilde{X} := X - \mathbb{E}(X|\mathscr{I})$, was wegen der Invarianz $\mathbb{E}(X|\mathscr{I}) \circ \varphi = \mathbb{E}(X|\mathscr{I})$ ($\mathbb{P}$-f.s.) möglich ist.

$\mathbb{P}$-fast sichere Konvergenz: Hierfür werden wir mit

$$\bar{X} \;:=\; \limsup_{n\to\infty} \frac{S_n}{n} \;\equiv\; \limsup_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} X \circ \varphi^k$$

und mit

$$D \;:=\; \{\, \bar{X} > \varepsilon \,\} \;\in\; \mathscr{I} \qquad\qquad \text{(zu beliebigem } \varepsilon > 0)$$

zeigen:

$$\mathbb{P}(D) \;=\; 0\,;$$

analog zeigt man $\liminf \frac{S_n}{n} \geq 0$, indem man $-X$ anstatt X betrachtet.
Um also $\mathbb{P}(D) = 0$ zu zeigen drücken wir D anders aus: Mit

$$\begin{aligned}
X^* \;&:=\; (X - \varepsilon)\, 1_D \\
S_n^* \;&:=\; X^* + X^* \circ \varphi + \cdots + X^* \circ \varphi^{n-1} \\
M_n^* \;&:=\; \max\{0, S_1^*, \ldots, S_n^*\} \\
F_n \;&:=\; \{\, M_n^* > 0 \,\}
\end{aligned}$$

hat man

$$D \;=\; \left\{ \sup_{n\in\mathbb{N}} \frac{S_n^*}{n} \;>\; 0 \right\} \;=\; \bigcup_{n\in\mathbb{N}} F_n\,.$$

Wendet man das maximal-ergodische Lemma 4.2 an auf X^*, so folgt hiermit:

$$\begin{aligned}
0 \quad &\leq \quad \mathbb{E}\,(\, X^*\, 1_{F_n}\,) && \text{(Lemma 4.2)} \\
&\xrightarrow{n\to\infty} \quad \mathbb{E}\Big(\, X^*\, 1_{\bigcup_n F_n} \Big) && \text{(maj. Kvgz., da } X \in L^1) \\
&= \quad \mathbb{E}\,(\, X^*\, 1_D\,) && \text{(vorherige Char. von } D) \\
&\equiv \quad \mathbb{E}\,(\, X\, 1_D\,) - \varepsilon\, \mathbb{P}(D) && \text{(Definition von } X^*) \\
&= \quad -\varepsilon\, \mathbb{P}(D) && (\,\mathbb{E}(X|\mathscr{I}) = 0 \text{ und } D \in \mathscr{I}) \\
&\leq \quad 0\,,
\end{aligned}$$

insgesamt also gerade: $\mathbb{P}(D) = 0$.

L^1-Konvergenz: Hierzu wird X „abgeschnitten“; mit einem festen $K > 0$ sei

$$X' \;:=\; X\, 1_{\{|X|\leq K\}} \qquad \text{und} \qquad X'' \;:=\; X - X'\,.$$

Die oben gezeigte $\mathbb{P}$-fast sichere Konvergenz gilt insbesondere auch für X'; da diese Konvergenz hier aber durch K majoriert ist, folgt insgesamt für X':

$$\frac{1}{n} \sum_{k=0}^{n-1} X' \circ \varphi^k \;\xrightarrow{n\to\infty}\; \mathbb{E}(X'|\mathscr{I}) \quad \text{in } L^1(\mathbb{P})\,.$$

Außerdem hat man

$$\mathbb{E}\left(\left|\frac{1}{n}\sum_{k=0}^{n-1} X''\circ\varphi^k\right|\right) \;\le\; \frac{1}{n}\sum_{k=0}^{n-1}\mathbb{E}\left(|X''|\circ\varphi^k\right) \;=\; \mathbb{E}(|X''|)\,,$$

wobei benutzt wurde, daß φ maßtreu ist; ferner gilt nach Jensen ($|\,.\,|$ ist konvex):

$$\mathbb{E}\left(\,\left|\mathbb{E}(X''|\mathscr{I})\right|\,\right) \;\le\; \mathbb{E}\left(\mathbb{E}(\,|X''|\,|\mathscr{I})\,\right) \;=\; \mathbb{E}(|X''|)\,;$$

faßt man die beiden letzten Ungleichungen zusammen, so ergibt sich:

$$\mathbb{E}\left(\left|\frac{1}{n}\sum_{k=0}^{n-1} X''\circ\varphi^k \;-\; \mathbb{E}(X''|\mathscr{I})\right|\right) \;\le\; 2\,\mathbb{E}(|X''|)\,.$$

Sei nun ein beliebiges $\varepsilon > 0$ fixiert; dann kann man $K > 0$ so groß wählen, daß

$$2\,\mathbb{E}(|X''|) \;<\; \frac{\varepsilon}{2}$$

ist (majorierte Konvergenz, Definition von X''). Mit diesen Parametern ε und K kann man wegen obiger L^1-Konvergenz bei X' ein $n_0 \in \mathbb{N}$ wählen, sodaß gilt:

$$\mathbb{E}\left(\left|\frac{1}{n}\sum_{k=0}^{n-1} X'\circ\varphi^k \;-\; \mathbb{E}(X'|\mathscr{I})\right|\right) \;<\; \frac{\varepsilon}{2} \qquad (n \ge n_0)\,.$$

Da nun $X \equiv X' + X''$ ist, ergeben die vorangehenden drei Abschätzungen:

$$\begin{aligned}&\mathbb{E}\left(\left|\frac{1}{n}\sum_{k=0}^{n-1} X\circ\varphi^k \;-\; \mathbb{E}(X|\mathscr{I})\right|\right)\\ &\le\; \mathbb{E}\left(\left|\frac{1}{n}\sum_{k=0}^{n-1} X'\circ\varphi^k - \mathbb{E}(X'|\mathscr{I})\right|\right) + \mathbb{E}\left(\left|\frac{1}{n}\sum_{k=0}^{n-1} X''\circ\varphi^k - \mathbb{E}(X''|\mathscr{I})\right|\right) \;<\; \varepsilon\end{aligned}$$

$\square$

Beispiel 4.3 (Starkes Gesetz der großen Zahlen). $(X_n)_{n\in\mathbb{N}_0}$ seien iid ZV, oE auf dem Folgenraum $\Omega := \mathbb{R}^{\mathbb{N}_0}$ definiert, mit $\mathbb{P} \equiv P_X = P_{X_0}\otimes P_{X_0}\otimes\cdots$ und ergodischem Shift $\varphi = \theta_1$ (siehe 3.12). Ist dann $X_0 \in L^1(\mathbb{P})$, so folgt aus 4.1 mit 3.9:

$$\frac{1}{n}\sum_{k=0}^{n-1} X_k \;=\; \frac{1}{n}\sum_{k=0}^{n-1} X_0\circ\varphi^k \;\xrightarrow{\mathbb{P}\text{-fs},\,L^1(\mathbb{P})}\; \mathbb{E}(X_0|\mathscr{I}) \;=\; \mathbb{E}(X_0)\,.$$

Beispiel 4.4 (Rotation des Kreises, Weylscher Gleichverteilungssatz). Es sei

$$\varphi\,:\,\Omega \longrightarrow \Omega\,, \qquad \varphi(\omega) \;:=\; \omega + \theta \pmod 1\,,$$

auf $(\Omega,\mathscr{F},\mathbb{P}) := ([0,1),\mathscr{B}[0,1),\lambda\big|_{\mathscr{F}})$ wie in 3.4 und 3.13, wobei λ das Lebesguemaß bezeichnet. Ferner sei $\theta \in \mathbb{Q}^c$. Dann folgt aus 4.1 mit 3.13 für $A \in \mathscr{B}[0,1)$:

$$\frac{1}{n}\sum_{k=0}^{n-1} 1_A\circ\varphi^k \;\xrightarrow{\lambda\text{-fs},\,L^1(\lambda)}\; \lambda(A)\,.$$

5. Der Subadditive Ergodensatz von Kingman

Sei $(\Omega, \mathscr{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum mit einer maßtreuen Transformation $\varphi : \Omega \to \Omega$. Im vorangehenden Abschnitt haben wir das asymptotische Verhalten von $\frac{S_n}{n}$ untersucht, wobei S_n die Gestalt $\sum_{k=0}^{n-1} X \circ \varphi^k$ hat, also insbesondere der *additiven Kozykel-Eigenschaft*

$$S_{n+m} \;=\; S_n \;+\; S_m \circ \varphi^n \qquad\qquad (n, m \in \mathbb{N}_0)$$

genügt. Nun interessieren wir uns für folgende Verallgemeinerung:

Definition 5.1 (Subadditive Folge von Zufallsvariablen). Eine Folge $(Y_n)_n$ von Zufallsvariablen ($n \in \mathbb{N}_0$ oder $\mathbb{N}$; Zustandsraum $\mathbb{R} \cup \{-\infty\}$) heißt *subadditiv*, wenn gilt:

$$Y_{n+m} \;\leq\; Y_n \;+\; Y_m \circ \varphi^n \qquad\qquad (n, m \in \mathbb{N}_0)\,.$$

Eine Folge $(Y_n)_{n\in\mathbb{N}_0}$ heißt *superadditiv*, wenn $(-Y_n)_{n\in\mathbb{N}_0}$ subadditiv ist, und sie heißt *additiv*, wenn sie sowohl sub- als auch superadditiv ist.

Beispiel 5.2. Sei $(X_n)_{n\in\mathbb{N}_0}$ eine Folge von iid ZV, die oE als $X_n = \pi_{\{n\}}$ auf dem Folgenraum $(\Omega, \mathscr{F}, \mathbb{P}, \varphi) = (S^{\mathbb{N}_0}, \mathscr{S}^{\mathbb{N}_0}, P_{(X_n)_{n\in\mathbb{N}_0}}, \theta_1)$ definiert ist. Hierzu sei

$$S_n \;:=\; \sum_{k=0}^{n-1} X_k\,.$$

Dann ist $(S_n)_{n\in\mathbb{N}_0}$ additiv und $(|S_n|)_{n\in\mathbb{N}_0}$ subadditiv.

Beweis. Die Additivität von $(S_n)_n$ folgt direkt, da $\varphi \equiv \theta_1$ ist. Ferner gilt:

$$\begin{aligned}
|S_{n+m}| \;\equiv\; \left|\sum_{k=0}^{n+m-1} X_k\right| \;&\leq\; \left|\sum_{k=0}^{n-1} X_k\right| + \left|\sum_{k=n}^{n+m-1} X_k\right| \\
&=\; |S_n| \;+\; \left|\sum_{k=0}^{m-1} X_k \circ \varphi^n\right| \;=\; |S_n| \;+\; |S_m| \circ \varphi^n\,.
\end{aligned}$$

□

Beispiel 5.3. Auf einem Wahrscheinlichkeitsraum $(\Omega, \mathscr{F}, \mathbb{P})$ mit maßtreuem φ sei eine zufällige Matrix, also eine meßbare Abbildung $A : \Omega \to \mathbb{R}^{d\times d}$ gegeben. Ferner seien

$$\begin{array}{rcll}
A_n & := & (A \circ \varphi^{n-1})(A \circ \varphi^{n-2}) \cdots A & \qquad \text{und hiermit} \\
Y_n & := & \log \| A_n \| & \qquad (n \in \mathbb{N})\,,
\end{array}$$

wobei $\| \,.\, \|$ eine Matrixnorm bezeichnet. Dann ist $(Y_n)_n$ subadditiv.

Beweis.

$$\begin{aligned}
Y_{n+m} \;&=\; \log \| (A \circ \varphi^{m-1} \circ \varphi^n) \cdots (A \circ \varphi^0 \circ \varphi^n)(A \circ \varphi^{n-1}) \cdots A \| \\
&=\; \log \| (A_m \circ \varphi^n)\, A_n \| \\
&\overset{\text{Norm}}{\leq}\; \log [\, (\| A_m \| \circ \varphi^n)\; \| A_n \| \,] \\
&=\; \log\, (\| A_m \| \circ \varphi^n) \;+\; \log \| A_n \| \;\;\equiv\;\; Y_m \circ \varphi^n \;+\; Y_n\,.
\end{aligned}$$

□

Ziel ist es nun, bei subadditivem $(Y_n)_n$ eine Konvergenzaussage für $\frac{Y_n}{n}$ zu erhalten. Dies wird im subadditiven Ergodensatz 5.7 von Kingman geschehen. Hierzu dienen die folgenden drei Lemmata.

Lemma 5.4 (Riesz). *Seien $u_1, \ldots, u_n \in \mathbb{R}$ $(n \in \mathbb{N})$. Mit*

$$s_j \;:=\; \begin{cases} 0\,, & j = 0 \\ u_1 + \cdots + u_j\,, & j \in \{1, \ldots, n\}\,, \end{cases}$$

definiere

$$v_j \;\equiv\; v_{jn} \;:=\; \max_{k \in \{j, \ldots, n\}} (\, s_k - s_j \,) \;\equiv\; \max\big\{\, 0\,,\, u_{j+1}\,,\, u_{j+1} + u_{j+2}\,,\, u_{j+1} + \cdots + u_n \,\big\}$$

für $j = 0, 1, \ldots, n$. Dann gilt:

$$\sum_{j=0}^{n-1} u_{j+1}\, 1_{\{v_{jn} > 0\}} \;\geq\; 0\,.$$

Beweis. 1) Zunächst gilt für alle $j \in \{0, 1, \ldots, n\}$:

$$v_j \;=\; \max\{\, 0\,,\, u_{j+1} + v_{j+1} \,\} \;\equiv\; (u_{j+1} + v_{j+1})^+\,.$$

Dies folgt direkt, da

$$\begin{aligned} v_j \;&=\; \max\{\, 0\,,\, u_{j+1}\,,\, u_{j+1} + u_{j+2}\,,\, u_{j+1} + \cdots + u_n \,\} \qquad \text{und} \\ v_{j+1} \;&=\; \max\{\, 0\,,\, u_{j+2}\,,\, u_{j+2} + u_{j+3}\,,\, u_{j+2} + \cdots + u_n \,\}\,. \end{aligned}$$

2) Wegen 1) gilt:

$$v_j \;\leq\; v_{j+1} + u_{j+1} 1_{\{v_j > 0\}} \qquad (j \in \{0, 1, \ldots, n\})\,.$$

Denn falls $v_j = 0$ ist, ist dies trivial, und im Falle $v_j > 0$ gilt:

$$0 \;<\; v_j \;\overset{1)}{=}\; (u_{j+1} + v_{j+1})^+ \;\overset{v_j > 0}{=}\; v_{j+1} + u_{j+1}\,.$$

3) Aus 2) folgt nun die Behauptung des Lemmas, denn:

$$0 \;\leq\; v_0 \;=\; v_0 - v_n \;=\; \sum_{j=0}^{n-1} (v_j - v_{j+1}) \;\overset{2)}{\leq}\; \sum_{j=0}^{n-1} u_{j+1}\, 1_{\{v_j > 0\}}\,.$$

□

Im Beweis des subadditiven Ergodensatzes von Kingman werden wir subadditive Folgen $(Y_n)_n$ vergleichen mit additiven Folgen $X_n = \sum_{i=0}^{n-1} X_0 \circ \varphi^i$. Dazu dient folgende Hilfsüberlegung, für die das vorangehende Lemma von Riesz benötigt wird:

Lemma 5.5 (Maximalungleichung). $(Y_n)_{n\in\mathbb{N}_0}$ *sei superadditiv auf* $(\Omega, \mathscr{F}, \mathbb{P}, \varphi)$ *und es gelte* $Y_n \geq 0$ *für alle* n. *Ferner sei* $X \geq 0$ *eine integrierbare ZV; hierzu setze*

$$V \;:=\; \sup_{n\in\mathbb{N}_0} (\, Y_n - X_n \,) - Y_0 \,, \qquad \textit{wobei} \qquad X_n \;:=\; \sum_{i=0}^{n-1} X \circ \varphi^i \,.$$

Dann gilt:

$$\mathbb{E}\left(X\, 1_{\{V>0\}} \,|\, \mathscr{I} \right) \;\leq\; \sup_{n\in\mathbb{N}} \frac{\mathbb{E}(\, Y_n \,|\, \mathscr{I}\,)}{n} \,.$$

Beweis. Es sei $v_{jn} := \max\limits_{k\in\{j,\ldots,n\}} \Big(Y_k - Y_j - \sum\limits_{i=j}^{k-1} X \circ \varphi^i \Big)$ für $j = 0, 1, \ldots, n$.

1) Zunächst gilt:

$$Y_n \;\geq\; \sum_{j=0}^{n-1} X \circ \varphi^j \;\; 1_{\{v_{jn}>0\}} \qquad\qquad (n \in \mathbb{N}) \,;$$

denn mit $Y_{j+1} \geq Y_j$ (wegen Superadditivität und $Y_n \geq 0$) erhält man:

$$\begin{aligned} Y_n \;\;\geq\;\; Y_n - Y_0 \;\;&=\;\; \sum_{j=0}^{n-1} (\, Y_{j+1} - Y_j \,) \\ &\geq\;\; \sum_{j=0}^{n-1} (\, Y_{j+1} - Y_j \,)\; 1_{\{v_{jn}>0\}} \;\overset{5.4}{\geq}\; \sum_{j=0}^{n-1} X \circ \varphi^j \;\; 1_{\{v_{jn}>0\}} \,, \end{aligned}$$

wobei der letzte Schritt aus 5.4 mit $u_j := Y_j - Y_{j-1} - X \circ \varphi^{j-1}$ folgt.

2) Aus 1) folgt nun:

$$\mathbb{E}(\, Y_n \,|\, \mathscr{I}\,) \;\geq\; \sum_{k=1}^{n} \mathbb{E}\left(X\, 1_{\{v_{0k}>0\}} \,|\, \mathscr{I} \right) \qquad\qquad (n \in \mathbb{N});$$

denn für $k \geq j$ folgt aus der Superadditivität $Y_k - Y_j \;\geq\; Y_{k-j} \circ \varphi^j$ und daher $v_{jn} \;\geq\; v_{0(n-j)} \circ \varphi^j$, also insgesamt (mit der Maßtreue von φ):

$$\begin{aligned} \mathbb{E}(\, Y_n \,|\, \mathscr{I}\,) \;\;&\overset{1)}{\geq}\;\; \sum_{j=0}^{n-1} \mathbb{E}\left(X \circ \varphi^j \;\; 1_{\{v_{jn}>0\}} \,\Big|\, \mathscr{I} \right) \\ &\geq\;\; \sum_{j=0}^{n-1} \mathbb{E}\left(\Big[X \;\; 1_{\{v_{0(n-j)}>0\}} \Big] \circ \varphi^j \,\Big|\, \mathscr{I} \right) \;=\; \sum_{k=1}^{n} \mathbb{E}\left(X\, 1_{\{v_{0k}>0\}} \,|\, \mathscr{I} \right) . \end{aligned}$$

3) Mit Fatou und $\{v_{0k} > 0\} \nearrow \{V > 0\}$ $(k \to \infty)$ ergibt sich daraus:

$$\sup_{n\in\mathbb{N}} \frac{\mathbb{E}(\, Y_n \,|\, \mathscr{I}\,)}{n} \;\overset{2)}{\geq}\; \liminf_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \mathbb{E}\left(X\, 1_{\{v_{0k}>0\}} \,|\, \mathscr{I} \right) \;\geq\; \mathbb{E}\left(X\, 1_{\{V>0\}} \,|\, \mathscr{I} \right) \,.$$

□

Lemma 5.6. *Ist Z eine meßbare Funktion auf $(\Omega, \mathscr{F}, \mathbb{P}, \varphi)$ mit $Z \geq Z \circ \varphi$, so gilt $Z = Z \circ \varphi$. Ist insbesondere $(Y_n)_{n \in \mathbb{N}_0}$ superadditiv und*

$$\overline{Y} := \limsup_{n\to\infty} \frac{Y_n}{n}, \qquad \textit{bzw.} \qquad \underline{Y} := \liminf_{n\to\infty} \frac{Y_n}{n},$$

so gilt:

$$\overline{Y} = \overline{Y} \circ \varphi, \qquad \textit{bzw.} \qquad \underline{Y} = \underline{Y} \circ \varphi.$$

Beweis. Wir zeigen zunächst die Aussage über Z und nehmen hierzu $Z > Z \circ \varphi$ auf einer Menge positiver Masse an, also

$$\mathbb{P}(Z > q > Z \circ \varphi) > 0$$

für ein $q \in \mathbb{Q}$. Dann folgt aber:

$$\begin{aligned}
\mathbb{P}(Z < q) &\overset{\varphi \text{ m.t.}}{=} \mathbb{P}(Z \circ \varphi < q) \\
&= \mathbb{P}(Z \circ \varphi < q \leq Z) + \mathbb{P}(Z \circ \varphi < q, Z < q) \\
&\overset{Z \geq Z\circ\varphi}{=} \underbrace{\mathbb{P}(Z \circ \varphi < q \leq Z)}_{>0} + \mathbb{P}(Z < q) \\
&> \mathbb{P}(Z < q),
\end{aligned}$$

ein Widerspruch.
Wegen der ebengezeigten Behauptung ist nun nur noch zu sehen:

$$\overline{Y} \geq \overline{Y} \circ \varphi, \qquad \text{bzw.} \qquad \underline{Y} \geq \underline{Y} \circ \varphi;$$

aufgrund der Superadditivität von $(Y_n)_n$ hat man nun:

$$\begin{aligned}
\frac{Y_{n+1}}{n+1} &\geq \frac{Y_1}{n+1} + \frac{Y_n \circ \varphi}{n+1} \\
&= \frac{Y_1}{n+1} + \frac{n}{n+1} \frac{Y_n}{n} \circ \varphi.
\end{aligned}$$

□

Theorem 5.7 (Subadditiver Ergodensatz, Kingman).
Auf $(\Omega, \mathscr{F}, \mathbb{P}, \varphi)$ sei $(Y_n)_{n \in \mathbb{N}}$ eine superadditive Folge integrierbarer ZV. Dann gilt:

$$\frac{Y_n}{n} \xrightarrow[n\to\infty]{\mathbb{P}\text{-}f.s.} \sup_{n \in \mathbb{N}} \frac{1}{n} \mathbb{E}(Y_n \mid \mathscr{I}) =: \gamma \leq \infty.$$

Dabei ist γ genau dann integrierbar, wenn $\sup_{n\in\mathbb{N}} \frac{1}{n} \mathbb{E}(Y_n) < \infty$ ist. In diesem Fall gilt auch

$$\frac{Y_n}{n} \xrightarrow[n\to\infty]{L^1(\mathbb{P})} \gamma.$$

Ferner existiert eine Menge $\tilde{\Omega} \in \mathscr{I}$ mit $\tilde{\Omega} \subset \varphi^{-1} \tilde{\Omega}$ und $\mathbb{P}(\tilde{\Omega}) = 1$, sodaß auch gilt:

$$\frac{Y_n}{n} \xrightarrow[n\to\infty]{} \gamma \quad \textit{auf} \quad \tilde{\Omega}.$$

Beweis. Um die Notation zu vereinfachen setzen wir $Y_0 := 0$; dann ist $(Y_n)_{n\in\mathbb{N}_0}$ weiterhin superadditiv.

1) Wir zeigen zunächst, daß man ohne Einschränkung

$$Y_n \ \geq\ 0 \qquad\qquad (n \in \mathbb{N})$$

annehmen kann. Hierzu sei

$$J_n \ :=\ H_n \ -\ F_n \qquad \text{mit} \qquad H_n \ :=\ \sum_{i=0}^{n-1} Y_1^+ \circ \varphi^i \quad \text{und} \quad F_n \ :=\ \sum_{i=0}^{n-1} Y_1 \circ \varphi^i\,;$$

setzt man noch $G_n := Y_n - F_n$, so schreibt sich Y_n als

$$Y_n \ \equiv\ Y_n \ -\ F_n \ +\ H_n \ -\ J_n \ \equiv\ G_n \ +\ H_n \ -\ J_n\,.$$

Dabei sind die Folgen $(H_n)_n$ und $(J_n)_n$ additiv, so daß die vorausgesetzte Integrabilität mit Birkhoffs Ergodensatz 4.1 ZV γ_H und γ_J liefert, sodaß

$$\frac{H_n}{n} \longrightarrow \gamma_H \qquad\qquad \text{und} \qquad\qquad \frac{J_n}{n} \longrightarrow \gamma_J$$

$\mathbb{P}$-f.s. und in $L^1(\mathbb{P})$ gilt. Die Behauptungen über Y folgen also, wenn auch

$$\frac{G_n}{n} \longrightarrow \gamma_G \qquad\qquad (\mathbb{P}\text{-f.s. und in } L^1(\mathbb{P}))$$

gezeigt ist, da dann insbesondere

$$\frac{Y_n}{n} \longrightarrow \gamma_G + \gamma_H - \gamma_J \qquad\qquad (\mathbb{P}\text{-f.s. und in } L^1(\mathbb{P}))$$

folgt. Nun sieht man aber mit induktiver Anwendung der Superadditivität,

$$\begin{aligned}
G_n \ &=\ -Y_1 \ -\ Y_1 \circ \varphi \ -\ \cdots \ -\ Y_1 \circ \varphi^{n-2} \ -\ Y_1 \circ \varphi^{n-1} \ +\ Y_n \\
&\geq\ -Y_1 \ -\ Y_1 \circ \varphi \ -\ \cdots \ -\ Y_1 \circ \varphi^{n-2} \ +\ Y_{n-1} \\
&\geq\ -Y_1 \ -\ Y_1 \circ \varphi \ -\ \cdots \ +\ Y_{n-2} \\
&\geq\ \cdots\ \cdots \\
&\geq\ -Y_1 \ -\ Y_1 \circ \varphi \ +\ Y_2 \quad \geq \quad 0\,;
\end{aligned}$$

andererseits überträgt sich die Superadditivität von $(Y_n)_n$ auf $(G_n)_n$. Daher sind die Behauptungen bzgl $(Y_n)_n$darauf zurückgeführt, die entsprechenden Konvergenzen für den positiven Prozeß $(G_n)_n$ zu zeigen.

2) Wir zeigen weiter, daß man ohne Einschränkung

$$Y_n \ \geq\ n \qquad\qquad (n \in \mathbb{N})$$

annehmen kann: Wegen $Y_{n+m} + n + m \ \geq\ Y_n + n + (Y_m + m) \circ \varphi^n$ ist mit $(Y_n)_n$ auch $(Y_n + n)_n$ superadditiv. Konvergiert nun $\frac{Y_n + n}{n} \longrightarrow \gamma'$, so auch $\frac{Y_n}{n} \longrightarrow \gamma := \gamma' - 1\,.$

3) $\frac{Y_n}{n} \longrightarrow \gamma$ $\mathbb{P}$-f.s.: Hierzu zeigen wir $\overline{Y} \leq \gamma$ und $\underline{Y} \geq \gamma$, wobei wieder

$$\overline{Y} \;:=\; \limsup_{n\to\infty} \frac{Y_n}{n} \qquad \text{bzw.} \qquad \underline{Y} \;:=\; \liminf_{n\to\infty} \frac{Y_n}{n} \,.$$

$\overline{Y} \leq \gamma$ $\mathbb{P}$-f.s.: Für $r \in \mathbb{N}_{>2}$ definieren wir

$$X^r \;:=\; \min\{\, r \,,\, \overline{Y} - \frac{1}{r} \,\} \;>\; 0\,;$$

hierbei folgt die Ungleichung, da wegen 2) $\overline{Y} \geq 1$ ist. Nach 5.6 gilt ferner $X^r = X^r \circ \varphi$, also $X_n^r := \sum_{i=0}^{n-1} X^r \circ \varphi^i = nX^r$; hiermit hat man

$$V \;:=\; \sup_{n\in\mathbb{N}_0} (\, Y_n - X_n^r \,) - Y_0 \;=\; \sup_{n\in\mathbb{N}_0} (\, Y_n - nX^r \,) \;>\; 0\,;$$

dabei ergibt sich die zuletzt notierte Ungleichung aus der Definition von X^r: wäre nämlich $Y_n \leq nX^r$ für alle $n \in \mathbb{N}_0$, so erhielte man den Widerspruch $\overline{Y} \equiv \limsup \frac{Y_n}{n} \leq \limsup \frac{nX^r}{n} = X^r < \overline{Y}$. Also folgt mit 5.5:

$$X^r \;=\; \mathbb{E}\,(\, X^r \,|\, \mathscr{I}\,) \;\leq\; \sup_{n\in\mathbb{N}} \frac{\mathbb{E}(\, Y_n \,|\, \mathscr{I})}{n} \;\equiv\; \gamma\,.$$

Hieraus folgt durch $r \to \infty$ aber: $\overline{Y} \leq \gamma$ $\mathbb{P}$-f.s..

$\underline{Y} \geq \gamma$ $\mathbb{P}$-f.s.: Zunächst ist $(Y_n)_n$ wegen der Superadditivität und der Positivität monoton wachsend; daraus schließt man:

$$k\, Y_{n+k-1} \;\geq\; \sum_{j=0}^{n-1} (Y_{j+k} - Y_j) \qquad\qquad (k, n \in \mathbb{N})\;;$$

hiermit erhält man für jedes $k \in \mathbb{N}$:

$$\begin{aligned}
\underline{Y} \;&=\; \liminf_{n\to\infty} \frac{Y_{n+k-1}}{n+k-1} \\
&=\; \liminf_{n\to\infty} \frac{Y_{n+k-1}}{n} \\
&=\; \frac{1}{k}\, \liminf_{n\to\infty} \frac{k\, Y_{n+k-1}}{n} \\
&\geq\; \frac{1}{k}\, \liminf_{n\to\infty} \sum_{j=0}^{n-1} \frac{Y_{j+k} - Y_j}{n} && \text{(vorangehende Bem.)} \\
&\geq\; \frac{1}{k}\, \liminf_{n\to\infty} \sum_{j=0}^{n-1} \frac{Y_k \circ \varphi^j}{n} && \text{(Superadditivität)} \\
&=\; \frac{1}{k}\, \mathbb{E}(\, Y_k \,|\, \mathscr{J}\,) && \text{(Birkhoff 4.1)}
\end{aligned}$$

also auch $\underline{Y} \geq \gamma$.

4) γ integrierbar $\Leftrightarrow \sup\limits_{n\in\mathbb{N}} \frac{1}{n}\,\mathbb{E}(Y_n) < \infty$: Hierzu sei $Z_n := \frac{Y_n}{n}$. Aus dem Bisherigen folgt: $Z_n \to \gamma$ $\mathbb{P}$-f.s. und $\mathbb{E}\gamma \geq \mathbb{E}Z_n$; daher ist „$\Rightarrow$“ gezeigt. „$\Leftarrow$“ folgt mit monotoner Konvergenz.

5) Ist γ integrierbar, so gilt: $Z_n \equiv \frac{Y_n}{n} \to \gamma$ in $L^1(\mathbb{P})$. Wegen $0 \leq (\gamma - Z_n)^+ \leq \gamma$ gilt zum einen

$$\mathbb{E}\left((\gamma - Z_n)^+ \right) \;\to\; 0\,;$$

andererseits gilt auch

$$0 \;\leq\; \mathbb{E}(\,\gamma - Z_n\,) \;\to\; 0$$

wegen Fatou, sodaß auch folgt:

$$\mathbb{E}\left((\gamma - Z_n)^- \right) \;=\; -\mathbb{E}(\,\gamma - Z_n\,) + \mathbb{E}\left((\gamma - Z_n)^+ \right) \;\to\; 0\,,$$

und somit insgesamt

$$\mathbb{E}\,(\,|\,\gamma - Z_n\,|\,) \;\to\; 0\,.$$

6) Existenz von $\tilde{\Omega} \in \mathscr{I}$ mit $\tilde{\Omega} \subset \varphi^{-1}\,\tilde{\Omega}$, $\mathbb{P}(\tilde{\Omega}) = 1$ und $\frac{Y_n}{n} \to \gamma$ auf $\tilde{\Omega}$: Wegen 5.6 sind $\overline{Y}$ und $\underline{Y}$ invariant. Deshalb ist auch

$$\tilde{\Omega} \;:=\; \{\,\overline{Y} \;=\; \underline{Y}\,\}$$

invariant; die restlichen Eigenschaften folgen aus dem bisher Gezeigten. $\square$

6. Der Satz von Furstenberg-Kesten

Sei $(\Omega, \mathscr{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum mit einer maßtreuen Abbildung $\varphi : \Omega \to \Omega$ und $A : \Omega \to \mathbb{R}^{d\times d}$ eine zufällige Matrix. Wir untersuchen nun die Asymptotik von

$$A_n(\omega) \;:=\; \big(A\circ\varphi^{n-1}(\omega)\big)\,\big(A\circ\varphi^{n-2}(\omega)\big)\cdots(A\circ\varphi(\omega))\,\big(A(\omega)\big) \qquad (\omega\in\Omega)\,. \tag{2}$$

Beispiel 6.1 (deterministische, symmetrische Matrix). Sei $A \in \mathbb{R}^{d\times d}$ symmetrisch. Dann existiert eine Diagonalisierung von A mit reellen (da A symmetrisch) Eigenwerten $\delta_1 \geq \cdots \geq \delta_d$; es existiert also eine orthogonale Matrix O, sodaß

$$A \;=\; O^*\,DO \qquad \text{mit} \qquad D \;:=\; \begin{pmatrix} \delta_1 & & 0 \\ & \ddots & \\ 0 & & \delta_d \end{pmatrix}$$

gilt. Hierbei gelte $\delta_1 > \cdots > \delta_d$, d.h. die zu δ_i gehörigen Eigenräume E_i seien eindimensional. Ferner sei x_i ein Einheitsvektor in E_i und

$$V_j \;=\; \begin{cases} E_j \oplus E_{j+1} \oplus \cdots \oplus E_d\,, & j = 1,\ldots,d \\ \{0\}\,, & j = d+1\,. \end{cases}$$

Sei hiermit $x \in V_j \setminus V_{j+1}$. Dann schreibt sich x als

$$x \;=\; \sum_{k=j}^{d} \alpha_k\, x_k \qquad \text{mit} \qquad \alpha_j \neq 0\,.$$

Somit gilt wegen der Linearität von A, da die x_k Eigenvektoren sind:

$$A^n x \;=\; \sum_{k=j}^{d} \alpha_k\, A^n x_k \;=\; \sum_{k=j}^{d} \alpha_k\, \delta_k^n\, x_k\,,$$

also

$$\begin{aligned} \frac{1}{n}\log|A^n x| \;&=\; \frac{1}{n}\log\left|\sum_{k=j}^{d}\alpha_k\,\delta_k^n\,x_k\right| \\ &=\; \frac{1}{n}\left[\log\delta_j^n \;+\; \log\left\|\sum_{k=j}^{d}\alpha_k\left(\frac{\delta_k}{\delta_j}\right)^n x_k\right\|\right] \;\xrightarrow{n\to\infty}\; \log\delta_j\,. \end{aligned}$$

Hiervon gilt auch die Umkehrung, also insgesamt:

$$x \in V_j \setminus V_{j+1} \quad\Longleftrightarrow\quad \lim_{n\to\infty}\frac{1}{n}\log\|A^n x\| \;=\; \log\delta_j \qquad (j = 1,\ldots,d)\,.$$

Wir verfolgen nun das Ziel, diese Aussage für $(A^n)_n$ analog auch für die Folge $(A_n)_n$ aus (2) zu zeigen.

Definition-Bemerkung 6.2 (Singulärwertzerlegung). Jedes $A \in \mathbb{R}^{d\times d}$ besitzt eine *Singulärwertzerlegung*, d.h. es gibt orthogonale Matrizen U, V und eine Diagonalmatrix

$$D = \begin{pmatrix} \delta_1 & & 0 \\ & \ddots & \\ 0 & & \delta_d \end{pmatrix}$$

mit $\delta_1 \geq \cdots \geq \delta_d$, sodaß gilt

$$A = VDU .$$

Dabei sind $\delta_1, \ldots, \delta_d$ die Eigenwerte von $(A^*A)^{1/2}$ und für die Operatornorm gilt: $\|A\| = \delta_1$.

Beweis. A hat zunächst eine *polare Zerlegung*, d.h.:

$$A = W(A^*A)^{1/2} \qquad \text{mit einer orthogonalen Matrix } W .$$

(Im Fall, daß A nicht-singulär ist, folgt dies mit $W := A(A^*A)^{-1/2}$). Sei nun $D := \operatorname{diag}(\delta_1, \ldots, \delta_d)$ die Diagonalmatrix mit den Eigenwerten $\delta_1 \geq \cdots \geq \delta_d$ von $(A^*A)^{1/2}$, so schreibt sich die positiv-semidefinite Martix $(A^*A)^{1/2}$ als

$$(A^*A)^{1/2} = U^* DU$$

mit einer orthogonalen Matrix U. Insgesamt gilt also mit $V := WU^*$ gerade

$$A = W(A^*A)^{1/2} = WU^* DU = VDU .$$

$\square$

Bemerkung: Sind $e_1, \ldots, e_d$ die kanonischen Einheitsvektoren des $\mathbb{R}^d$, so ist Ue_i der Vektor in Richtung der i-ten Hauptachse des Ellipsoids $(A^*A)^{1/2}(S^{d-1})$ und δ_i gibt die Streckung in dieser Richtung an.

Zur Konstruktion der Analoga von $\delta_1, \ldots, \delta_d$ in Beispiel 6.1 für die in (2) definierte Folge A_n benötigen wir Information darüber, wie A_n lineare Unterräume des $\mathbb{R}^d$ transformiert:

Definition 6.3 (Äußeres Produkt). Zu einem d-dimensionalen linearen Raum E sei

$$\mathscr{L}^k(E^*) := \{\, k\text{-linearen Multilinearformen auf } (E^*)^k \,\} \qquad (k = 1, \ldots, d) .$$

Hiermit definieren wir $\wedge^k E$, das *k-fache äußere Produkt von E*, als

$$\wedge^k E := \{\, f \in \mathscr{L}^k(E^*) \;:\; f \text{ alternierend} \,\} ,$$

also als die Gesamtheit aller k-linearen, alternierenden Multilinearformen auf $(E^*)^k$.
Ein Element $f \in \wedge^k E$ ist also eine k-lineare Abbildung

$$f : \underbrace{E^* \times \cdots \times E^*}_{k\text{-mal}} \longrightarrow \mathbb{R} ,$$

die alternierend ist, d.h.:

$$f(\ldots, x_i, \ldots, x_j, \ldots) = -f(\ldots, x_j, \ldots, x_i, \ldots) \qquad (i \neq j) .$$

Lemma 6.4 (Alternierende Abbildungen). *Für ein $f \in \mathscr{L}^k(E^*)$ sind äquivalent:*

i) $f \in \wedge^k E$

ii) $f(x_1\,,\,\dots\,,x_k) \;=\; 0$ *, falls* $(x_1\,,\,\dots\,,x_k)$ *nicht paarweise verschieden*

iii) $f(x_1\,,\,\dots\,,x_k) \;=\; 0$ *, falls* $(x_1\,,\,\dots\,,x_k)$ *nicht paarweise linear unabhängig*

iv) $f(x_{\pi(1)}\,,\,\dots\,,x_{\pi(k)}) \;=\; sgn(\pi)\, f(x_1\,,\,\dots\,,x_k)$ *für alle* $\pi \in \mathfrak{S}_k$ *.*

Beweis. i)⇔iv) folgt durch die Darstellung $\pi = \tau_1 \circ \cdots \circ \tau_k$ mit Zweierpermutationen τ_i , also $\operatorname{sgn}(\tau_i) = -1$.

i)⇒ii) Sind $(x_1\,,\,\dots\,,x_k)$ nicht paarweise verschieden, so folgt aus der Definition einer alternierenden Abbildung durch Vertauschen der gleichen Elemente: $f(x_1\,,\,\dots\,,x_k) \;=\; -f(x_1\,,\,\dots\,,x_k)$.

ii)⇒iii) Es sei ohne Einschränkung $x_k = \sum_{i=1}^{k-1} \alpha_i\, x_i$. Dann folgt mit Linearität und ii): $f(x_1\,,\,\dots\,,x_k) \;=\; \sum_{i=1}^{k-1} \alpha_i\, f(x_1\,,\,\dots\,,x_{k-1}\,,x_i) \;=\; 0$.

iii)⇒ii) ist trivial.

ii)⇒i) Es sei ohne Einschränkung $k = 2$. Dann folgt für $x_1, x_2 \in E^*$:

$$\begin{aligned} 0 &\overset{\text{ii)}}{=} f(x_1 + x_2\,,\, x_1 + x_2) \\ &= f(x_1,x_1) \,+\, f(x_1,x_2) \,+\, f(x_2,x_1) \,+\, f(x_2,x_2) \\ &\overset{\text{ii)}}{=} f(x_1,x_2) \,+\, f(x_2,x_1)\,, \end{aligned}$$

also $f(x_1,x_2) \;=\; -f(x_2,x_1)$. □

Definition-Bemerkung 6.5. Es seien $f \in \wedge^k E$ und $g \in \wedge^l E$, wobei E wieder einen d-dimensionalen linearen Raum bezeichnet und $k, l \in \mathbb{N}_0$ sind. Dann heißt

$$f \wedge g\,(\,x_1\,,\,\dots\,,x_{k+l}\,) \;:=\; \frac{1}{k!\,l!} \sum_{\pi \in \mathfrak{S}_{k+l}} \operatorname{sgn}(\pi)\, f\left(x_{\pi(1)}\,,\,\dots\,,x_{\pi(k)}\right)\, g\left(x_{\pi(k+1)}\,,\,\dots\,,x_{\pi(k+l)}\right)$$

das *äußere Produkt von* f *und* g und es gilt: $f \wedge g \in \wedge^{k+l} E$.

Beweis. Es ist $f \wedge g \in \mathscr{L}^m(E^*)$ mit $m := k + l$. Um zu sehen, daß $f \wedge g$ auch alternierend ist, wird 6.4 iv) angewendet; zu beliebigen $x_1, \dots, x_m \in E^*$ sei

$$a(\pi) \;:=\; \frac{1}{k!\,l!}\, f\left(x_{\pi(1)}\,,\,\dots\,,x_{\pi(k)}\right)\, g\left(x_{\pi(k+1)}\,,\,\dots\,,x_{\pi(m)}\right) \qquad (\pi \in \mathfrak{S}_m)\,.$$

Hiermit folgt:

$$\begin{aligned} f \wedge g\,(\,x_{\pi(1)}\,,\,\dots\,,x_{\pi(m)}\,) &= \sum_{\pi' \in \mathfrak{S}_m} \operatorname{sgn}(\pi')\, a(\pi' \circ \pi) \\ &= \operatorname{sgn}(\pi) \sum_{\pi' \in \mathfrak{S}_m} \operatorname{sgn}(\pi' \circ \pi)\, a(\pi' \circ \pi) \\ &= \operatorname{sgn}(\pi) \sum_{\sigma \in \mathfrak{S}_m} \operatorname{sgn}(\sigma)\, a(\sigma) \\ &\equiv \operatorname{sgn}(\pi)\; f \wedge g\,(\,x_1\,,\,\dots\,,x_m\,)\,. \end{aligned}$$

□

Lemma 6.6 (Assoziativität des äußeren Produktes).
Es seien $f \in \wedge^k E$, $g \in \wedge^l E$ *und* $h \in \wedge^m E$ *mit* $k, l, m \in \mathbb{N}_0$. *Dann gilt:*

$$(f \wedge g) \wedge h \;=\; f \wedge (g \wedge h) \;.$$

Beweis. Es sei $n := k + l + m$ und $\mathfrak{T} := \{ \tau \in \mathfrak{S}_n \;:\; \tau(i) = i \text{ für } i > k + l \}$; ferner sei zu beliebigen $x_1, \ldots, x_m \in E^*$ und $\pi \in \mathfrak{S}_m$

$$a(\pi) := f\left(x_{\pi(1)}, \ldots, x_{\pi(k)}\right) g\left(x_{\pi(k+1)}, \ldots, x_{\pi(k+l)}\right) h\left(x_{\pi(k+l+1)}, \ldots, x_{\pi(n)}\right).$$

Hiermit folgt durch zweimaliges Anwenden von 6.5:

$$\begin{aligned}
\big((f \wedge g) \wedge h\big)(x_1, \ldots, x_n) &= \\
&= \frac{1}{(k+l)!\, m!} \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma) \frac{1}{k!\, l!} \sum_{\tau \in \mathfrak{T}} \operatorname{sgn}(\tau)\, a(\sigma \circ \tau) \\
&= \frac{1}{k!\, l!\, m!} \frac{1}{(k+l)!} \sum_{\tau \in \mathfrak{T}} \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma \circ \tau)\, a(\sigma \circ \tau) \\
&= \frac{1}{k!\, l!\, m!} \frac{\operatorname{card}(\mathfrak{T})}{(k+l)!} \sum_{\gamma \in \mathfrak{S}_n} \operatorname{sgn}(\gamma)\, a(\gamma) \\
&= \frac{1}{k!\, l!\, m!} \sum_{\gamma \in \mathfrak{S}_n} \operatorname{sgn}(\gamma)\, a(\gamma) \,.
\end{aligned}$$

Da man dieses Ergebnis aber auch erhält, wenn man (mit denselben Schritten) $(f \wedge (g \wedge h))(x_1, \ldots, x_n)$ berechnet, ist die Behauptung gezeigt. □

Damit ist klar, daß Ausdrücke wie

$$f_1 \wedge \ldots \wedge f_m \qquad \text{mit } f_l \in \wedge^{k_l} E$$

eindeutig bestimmt sind. Hierfür gilt:

Lemma 6.7. *Es seien* $f_l \in \wedge^{k_l} E$ *für* $l \in \{1, \ldots, m\}$. *Dann gilt mit* $n := k_1 + \cdots + k_m$

$$f_1 \wedge \ldots \wedge f_m \;=\; \prod_{1 \le l \le m} \frac{1}{k_l\,!} \cdot \sum_{\pi \in \mathfrak{S}_n} sgn(\pi)\, f_\pi \,,$$

wobei f_π *mit* $i_l := k_1 + \cdots + k_{l-1}$ *definiert ist als*

$$f_\pi(x_1, \ldots, x_n) := f_1\left(x_{\pi(1)}, \ldots, x_{\pi(i_1)}\right) f_2\left(x_{\pi(i_1+1)}, \ldots, x_{\pi(i_2)}\right) \cdots f_m\left(x_{\pi(i_m+1)}, \ldots, x_{\pi(n)}\right).$$

Beweis. Dies folgt mit vollständiger Induktion nach m. Der Fall $l = 2$ ist dabei gerade die Definition in 6.5; der Fall $l = 3$ ist im Beweis von 6.6 gezeigt. □

Lemma 6.8. *Es sei $e_1, \dots, e_d$ eine Basis von $E^{**} \cong E$ und $b_1, \dots, b_d$ sei hierzu die duale Basis von E^*. Dann gilt für alle $f \in \wedge^k E$:*

$$f = \sum_{i_1 < \cdots < i_k} a_{i_1 \dots i_k}\, e_{i_1} \wedge \dots \wedge e_{i_k} \quad \Longleftrightarrow \quad a_{i_1 \dots i_k} = f(b_{i_1}, \dots, b_{i_k}) \text{ für alle } i_1 < \cdots < i_k \,.$$

Beweis. Zunächst folgt für $i_1 < \cdots < i_k$ und $j_1 < \cdots < j_k$ aus 6.7:

$$\begin{aligned} e_{i_1} \wedge \dots \wedge e_{i_k} (b_{j_1}, \dots, b_{j_k}) &= \sum_{\pi \in \mathfrak{S}_k} \operatorname{sgn}(\pi)\, e_{i_1}\Big(b_{j_{\pi(1)}}\Big) \cdots e_{i_k}\Big(b_{j_{\pi(k)}}\Big) \\ &= \sum_{\pi \in \mathfrak{S}_k} \operatorname{sgn}(\pi)\ \delta_{i_1, j_{\pi(1)}} \cdots \delta_{i_k, j_{\pi(k)}} \\ &= \delta_{i_1, j_1} \cdots \delta_{i_k, j_k} \qquad\qquad (\text{da } i_1 < \cdots < i_k) \\ &= \begin{cases} 1\,, & \text{falls } i_1 = j_1\,, \dots, i_k = j_k \\ 0\,, & \text{sonst.} \end{cases} \end{aligned}$$

„$\Rightarrow$“ folgt direkt aus dieser Vorbemerkung;
„$\Leftarrow$“ Hierzu sei $g := \sum_{i_1 < \cdots < i_k} f(b_{i_1}, \dots, b_{i_k})\, e_{i_1} \wedge \dots \wedge e_{i_k} \in \wedge^k E$. Wegen obiger Vorbemerkung im Beweis gilt $f(b_{i_1}, \dots, b_{i_k}) = g(b_{i_1}, \dots, b_{i_k})$ für alle $i_1 < \cdots < i_k$; wegen Linearität sind daher f und g auf ganz $(E^*)^k$ gleich. □

Lemma 6.9. *Es sei $e_1, \dots, e_d$ eine Basis von $E^{**} \cong E$ und $k \in \{1, \dots, d\}$. Dann ist*

$$\{\, e_{i_1} \wedge \dots \wedge e_{i_k} \ : \ 1 \le i_1 < \dots < i_k \le d \,\}$$

eine Basis von $\wedge^k E$. Insbesondere gilt:

$$\dim \wedge^k E = \binom{d}{k}\,.$$

Beweis. Hierzu wählt man eine zu $e_1, \dots, e_d$ duale Basis $b_1, \dots, b_d$ von E^* und wendet 6.8 an. □

Definition 6.10 (Skalarprodukt). Sei $b_1, \dots, b_d$ eine Basis von E^*. Dann definiert

$$\langle f, g \rangle := \sum_{i_1 < \cdots < i_k} f(b_{i_1}, \dots, b_{i_k})\, g(b_{i_1}, \dots, b_{i_k}) \qquad (f, g \in \wedge^k E)$$

ein Skalarprodukt auf $\wedge^k E$.

Lemma 6.11. *Für das Skalarprodukt aus 6.10 gilt:*

$$\langle u_1 \wedge \dots \wedge u_k\,,\, v_1 \wedge \dots \wedge v_k \rangle = \det\big(\langle u_i, v_j \rangle\big)_{1 \le i,j \le k} \qquad (u_i, v_i \in E)\,.$$

Insbesondere gilt für die zugehörige Norm:

$$|\, u_1 \wedge \dots \wedge u_k \,| = \sqrt{\det\big(\langle u_i, v_j \rangle\big)_{1 \le i,j \le k}} \qquad (u_i, v_i \in E)\ .$$

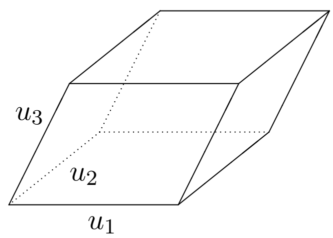
Beweis. Die rechte Seite $h(u_1,\ldots,u_k;v_1,\ldots,v_k) := \det\big(\langle u_i,v_j\rangle\big)_{1\le i,j\le k}$ ist bei festen $v_1,\ldots,v_k$ eine alternierende Multilinearform in $u_1,\ldots,u_k$ und umgekehrt, d.h. $h(\,.\,;v_1,\ldots,v_k)\in\wedge^k E^*$ und $h(u_1,\ldots,u_k;\,.\,)\in\wedge^k E^*$. Bezeichnet $e_1,\ldots,e_d$ die zu $b_1,\ldots,b_d$ duale Basis von E und wendet man 6.8 zweimal an, so folgt also:

$$\begin{aligned}
&h(u_1,\ldots,u_k;v_1,\ldots,v_k)\\
&\quad= \sum_{i_1<\cdots<i_k} h(e_{i_1},\ldots,e_{i_k};v_1,\ldots,v_k)\, b_{i_1}\wedge\ldots\wedge b_{i_k}(u_1,\ldots,u_k)\\
&\quad= \sum_{i_1<\cdots<i_k}\sum_{j_1<\cdots<j_k} h(e_{i_1},\ldots,e_{i_k};e_{j_1},\ldots,e_{j_k})\, b_{i_1}\wedge\ldots\wedge b_{i_k}(u_1,\ldots,u_k)\\
&\qquad\qquad\cdot b_{j_1}\wedge\ldots\wedge b_{j_k}(v_1,\ldots,v_k)\\
&\quad= \sum_{i_1<\cdots<i_k} b_{i_1}\wedge\ldots\wedge b_{i_k}(u_1,\ldots,u_k)\;\; b_{i_1}\wedge\ldots\wedge b_{i_k}(v_1,\ldots,v_k)\\
&\quad= \sum_{i_1<\cdots<i_k} u_1\wedge\ldots\wedge u_k(b_{i_1},\ldots,b_{i_k})\;\; v_1\wedge\ldots\wedge v_k(b_{i_1},\ldots,b_{i_k})\\
&\quad\equiv \langle u_1\wedge\ldots\wedge u_k\,,\, v_1\wedge\ldots\wedge v_k\rangle\,.
\end{aligned}$$

□

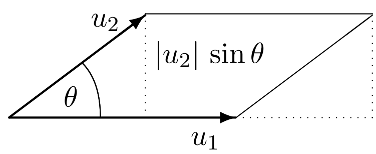
Bemerkung: Insbesondere gilt das Vorangehende für $E := \mathbb{R}^d = E^*$ (mit der kanonischen Basis). Die Norm

$$|\,u_1\wedge\ldots\wedge u_k\,| \;=\; \sqrt{\det\big(\langle u_i,v_j\rangle\big)_{1\le i,j\le k}}$$



gibt dann für $E=\mathbb{R}^d$ das k-dimensionale Volumen des Parallelflachs an, der von $u_1,\ldots,u_k$ aufgespannt wird. So gilt z.B. im Fall $k=2$:

$$\begin{aligned}
|\,u_1\wedge u_2\,| &= \det\begin{pmatrix}\langle u_1,u_1\rangle & \langle u_1,u_2\rangle\\ \langle u_2,u_1\rangle & \langle u_2,u_2\rangle\end{pmatrix}^{1/2}\\
&= \det\begin{pmatrix}|u_1|^2 & |u_1|\,|u_2|\,\cos\theta\\ |u_1|\,|u_2|\,\cos\theta & |u_2|^2\end{pmatrix}^{1/2}\\
&= \big(\,|u_1|^2\,|u_2|^2\,(1-\cos^2\theta)\,\big)^{1/2}\\
&= |u_1|\,|u_2|\,\sin\theta
\end{aligned}$$



□

Da das Ziel ist, die Wirkung auf k-dimensionale Objekte im $\mathbb{R}^d$ zu studieren, betrachten wir nun das k-fache äußere Produkt einer Matrix:

Definition-Bemerkung 6.12. Sei $A \in \mathbb{R}^{d\times d}$. Dann wird wegen 6.9 durch

$$\wedge^k A\,(u_1 \wedge \ldots \wedge u_k) \;:=\; Au_1 \wedge \ldots \wedge Au_k \qquad\qquad (u_i \in \mathbb{R}^d)$$

ein linearer Operator $\wedge^k A : \wedge^k\mathbb{R}^d \to \wedge^k\mathbb{R}^d$ definiert, das *k-fache äussere Produkt der Matrix A*. Hierfür gilt:

i) $\wedge^1 A = A$,

ii) $\wedge^d A = \det A$ (wegen 6.7),

iii) $\wedge^k(AB) = (\wedge^k A)(\wedge^k B)$,

iv) $(\wedge^k A)^{-1} = \wedge^k A^{-1}$ falls A invertierbar,

v) $\wedge^k(cA) = c^k \wedge^k A$ für $c \in \mathbb{R}$,

vi) $\wedge^k U$ orthogonal, falls U orthogonal und in diesem Fall gilt $(\wedge^k U)^* = \wedge^k U^*$.

Lemma 6.13 (Äußeres Produkt einer Matrix und Eigenwerte). *Seien $\lambda_1,\ldots,\lambda_d$ die Eigenwerte von $A \in \mathbb{R}^{d\times d}$. Dann hat $\wedge^k A$ die Eigenwerte*

$$\{\, \lambda_{i_1} \cdots \lambda_{i_k} \,:\, 1 \le i_1 < \cdots < i_k \le d \,\}\,.$$

Beweis. Sind $u_1,\ldots,u_d$ Eigenvektoren zu $\lambda_1,\ldots,\lambda_d$ und fixiert man Indizes $1 \le i_1 < \cdots < i_k \le d$, so folgt:

$$\begin{aligned} \wedge^k A\,(u_{i_1} \wedge \ldots \wedge u_{i_k}) &\equiv Au_{i_1} \wedge \ldots \wedge Au_{i_k} \\ &= \lambda_{i_1} u_{i_1} \wedge \ldots \wedge \lambda_{i_k} u_{i_k} \\ &= (\lambda_{i_1} \cdots \lambda_{i_k})\,(u_{i_1} \wedge \ldots \wedge u_{i_k})\,, \end{aligned}$$

sodaß $\lambda_{i_1} \cdots \lambda_{i_k}$ ein Eigenwert zum Eigenvektor $u_{i_1} \wedge \ldots \wedge u_{i_k}$ ist. Aus Dimensionsgründen müssen dies alle Eigenvektoren und damit alle Eigenwerte sein. □

Lemma 6.14 (Äußeres Produkt einer Matrix und Singulärzerlegung).
Zu $A \in \mathbb{R}^{d\times d}$ seien $\delta_1 \ge \ldots \ge \delta_d \ge 0$ die Singulärwerte und

$$A = VDU$$

eine Singulärwertzerlegung, wobei $D \equiv \operatorname{diag}(\delta_1,\ldots,\delta_d)$ ist. Dann gilt für $k = 1,\ldots,d$:

i) $\wedge^k A = (\wedge^k V)(\wedge^k D)(\wedge^k U)$ *ist Singulärzerlegung von $\wedge^k A$;*

ii) $\wedge^k D = \operatorname{diag}(\,\delta_{i_1} \cdots \delta_{i_k} \,:\, 1 \le i_1 < \cdots < i_k \le d\,)$.
Also ist $\delta_1 \cdots \delta_k$ der größte bzw. $\delta_{d-k+1} \cdots \delta_d$ der kleinste Singulärwert von $\wedge^k A$.

iii) Für die Operatornorm gilt:
$\|\wedge^k A\| = \delta_1 \cdots \delta_k\,, \quad |\det A| = \|\wedge^d A\| = \delta_1 \cdots \delta_d$ *und* $\|\wedge^k A\| \le \|A\|^k$.

Beweis. i) und ii) folgen aus 6.12 und 6.13; iii) ergibt sich aus ii) und der Definition der Operatornorm $\|\cdot\|$. □

Theorem 6.15 (Furstenberg-Kesten). *Sei $(\Omega, \mathscr{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und hierauf $A : \Omega \to \mathbb{R}^{d\times d}$ eine zufällige Matrix, für die gilt*

$$\log^+ \| A(\,.\,) \| \in L^1(\mathbb{P}) \,. \tag{3}$$

Ferner sei wie in (2)

$$A_n \;:=\; \left(A \circ \varphi^{n-1}\right) \left(A \circ \varphi^{n-2}\right) \cdots (A \circ \varphi)\; A$$

mit einer ($\mathbb{P}$-)maßtreuen Abbildung $\varphi : \Omega \to \Omega$.

Dann existiert eine Menge $\tilde{\Omega} \in \mathscr{F}$ mit $\mathbb{P}(\tilde{\Omega}) = 1$ sowie $\tilde{\Omega} \subset \varphi^{-1}(\tilde{\Omega})$, und es existieren meßbare Funktionen

$$\gamma^{(k)} : \Omega \longrightarrow \mathbb{R} \cup \{-\infty\} \qquad\qquad (k = 1, \ldots, d)$$

mit $\gamma^{(k)+} \in L^1(\mathbb{P})$, sodaß für alle $\omega \in \tilde{\Omega}$ und $k, m \in \{1, \ldots, d\}$ gilt:

$$\begin{aligned}
\gamma^{(k)}(\omega) &= \lim_{n\to\infty} \frac{1}{n} \log \| \wedge^k A_n(\omega) \| \,, \\
\gamma^{(k)}\big(\varphi(\omega)\big) &= \gamma^{(k)}(\omega) \,, \\
\gamma^{(k+m)}(\omega) &\le \gamma^{(k)}(\omega) + \gamma^{(m)}(\omega) \,.
\end{aligned}$$

Definiert man rekursiv Zufallsvariable

$$\Lambda_k : \Omega \longrightarrow \mathbb{R} \cup \{-\infty\} \qquad\qquad (k = 1, \ldots, d)$$

durch

$$\Lambda_1 + \ldots + \Lambda_k \;=\; \gamma^{(k)}$$

mit

$$\Lambda_k \;:=\; -\infty \quad \textit{auf} \quad \{\, \gamma^{(k)} = -\infty \,\} \,,$$

so gilt für alle $\omega \in \tilde{\Omega}$ und $k \in \{1, \ldots, d\}$:

$$\begin{aligned}
\Lambda_k(\omega) &= \lim_{n\to\infty} \frac{1}{n} \log \delta_k\big(A_n(\omega)\big) \,, \\
\Lambda_k\big(\varphi(\omega)\big) &= \Lambda_k(\omega) \,, \\
\Lambda_1(\omega) &\ge \Lambda_2(\omega) \;\ge\; \ldots \;\ge\; \Lambda_d(\omega) \;\; (\,\ge\, -\infty\,) \,.
\end{aligned}$$

Ist $\mathbb{P}$ ergodisch, so sind $\gamma^{(k)}$ und Λ_k wegen obiger Invarianz konstant (auf $\tilde{\Omega}$), also $\gamma^{(k)} = \mathbb{E}(\gamma^{(k)})$ und $\Lambda_k = \mathbb{E}(\Lambda_k)$.

Beweis. 1) Sei

$$Y_n^k \;:=\; \log \| \wedge^k A_n \| \qquad (n \in \mathbb{N}, k = 1, \ldots, d) \,;$$

dann ist $(Y_n^k)_n$ für jedes $k = 1, \ldots, d$ subadditiv: Im Falle $k = 1$ wurde dies in 5.3 gezeigt; für $k > 1$ überträgt sich die dortige Rechnung sofort, da für alle Matrizen B, C gilt: $\wedge^k(BC) = (\wedge^k B)(\wedge^k C)$. Daher ist mit A auch jedes $\wedge^k A$ ein *Kozykel*, d.h. es gilt:

$$\wedge^k A_{n+m} \;=\; \wedge^k A_n \circ \varphi^m \,\cdot\, \wedge^k A_m \,.$$

Somit folgt also die Subadditivität von $(Y_n^k)_n$.

2) Die Existenz von $\tilde{\Omega}$ und $\gamma^{(k)}$ mit den behaupteten Eigenschaften folgt aus Satz 5.7, angewandt auf $(-Y_n^k)_n$; es bleibt lediglich zu zeigen:

$$\gamma^{(k+m)} \leq \gamma^{(k)} + \gamma^{(m)} ;$$

dies folgt aber direkt aus der charakterisierenden Eigenschaft der $\gamma^{(k)}$ und der Normungleichung

$$\| \wedge^{k+m} A_n\| \leq \| \wedge^k A_n\| \cdot \| \wedge^m A_n\| .$$

3) Wir zeigen nun die Behauptungen bzgl. Λ_k : Nach 6.14 gilt für $k = 1, \ldots, d$:

$$\frac{1}{n} \log \| \wedge^k A_n\| = \frac{1}{n} \sum_{i=1}^{k} \log \delta_i(A_n).$$

Es ist $\Lambda_1 \equiv \gamma^{(1)}$ und für $\omega \in \tilde{\Omega}$ erhält man daraus sukzessiv:

$$\Lambda_{k+1}(\omega) \equiv \gamma^{k+1}(\omega) - \gamma^k(\omega) = \lim_{n\to\infty} \frac{1}{n} \log \delta_{k+1}(A_n) ,$$

falls $\gamma^k(\omega) > \infty$ ist; bricht dieses Verfahren ab, d.h. ist $\gamma^{k_0}(\omega) = -\infty$, so ist auch $\gamma^k(\omega) = -\infty$ für alle $k \geq k_0$ und damit auch $\Lambda_k = -\infty$ für alle $k \geq k_0$. Die restlichen Aussagen gelten wegen

$$\delta_1(A_n) \geq \delta_2(A_n) \geq \ldots \geq \delta_d(A_n)$$

und die jeweiligen Erwartungswerte existieren nach Voraussetzung. □

7. Der Multiplikative Ergodensatz von Oseledets

Sei $(\Omega, \mathscr{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum mit einer maßtreuen Abbildung $\varphi : \Omega \to \Omega$ und $A : \Omega \to \mathbb{R}^{d\times d}$ eine zufällige Matrix. In Übereinstimmung mit (2) definieren wir

$$A_n \;:=\; \begin{cases} \left(A\circ\varphi^{n-1}\right)\left(A\circ\varphi^{n-2}\right)\cdots\left(A\circ\varphi\right)A\,, & n\in\mathbb{N}\,,\\ I\,, & n=0\,, \end{cases}$$

den *von A erzeugten Kozykel*; A_n ist also *Kozykel* über φ, d.h.es gilt:

$$A_{n+m} \;=\; \left(A_n\circ\varphi^m\right)\cdot A_m \qquad\qquad (m,n\in\mathbb{N}_0)\,,$$

was ja bereits im Beweis des Satzes von Furstenberg-Kesten benutzt worden war.

Wir interessieren uns nun für die Asymptotik von $|A_n x|$ zu $x\in\mathbb{R}^d$ bei $n\to\infty$. Dies führen wir auf den Satz von Furstenberg-Kesten zurück mit Hilfe des folgenden (deterministischen) Satzes 7.3. Um die darin enthaltenen Konvergenzaussagen beweisen zu können zeigen wir zunächst zwei Lemmata:

Lemma 7.1. *Sei $\Phi\in\mathbb{R}^{d\times d}$ symmetrisch mit spektraler Zerlegung*

$$\Phi \;=\; \sum_{i=1}^{r}\lambda_i P_i\,,$$

wobei $r\le d$ ist und λ_i die Eigenwerte sowie P_i die zugehörigen orthogonalen Projektoren auf die Eigenräume bezeichnen. Seien

$$\Phi_n \;=\; \sum_{i=1}^{r_n}\lambda_i^n P_i^n$$

ebenfalls symmetrische $d\times d$-Matrizen, sodaß gilt:

i) $\lambda_k^n \xrightarrow{n\to\infty} \lambda_i$ für alle $k\in\Sigma_i$, wobei $\Sigma_i\neq\emptyset$ Mengen von Indizes sind (i=1,..., r);

ii) $\bar{P}_i^n := \sum_{k\in\Sigma_i} P_k^n \xrightarrow{n\to\infty} P_i$ für alle $i=1,\dots,r$.

Dann folgt: $\Phi_n \xrightarrow{n\to\infty} \Phi$.

Beweis. Mit den Konvergenzvoraussetzungen ergibt sich:

$$\begin{aligned}\Phi_n - \Phi \;&=\; \sum_{i=1}^{r}\sum_{k\in\Sigma_i}\lambda_k^n P_k^n \;-\; \sum_{i=1}^{r}\lambda_i P_i\\ &=\; \sum_{i=1}^{r}\Bigg[\underbrace{\sum_{k\in\Sigma_i}\left(\lambda_k^n-\lambda_i\right)P_k^n}_{\to 0} + \lambda_i\underbrace{\Big(\sum_{k\in\Sigma_i}P_k^n - P_i\Big)}_{\to 0}\Bigg] \;\xrightarrow{n\to\infty}\; 0\,.\end{aligned}$$

$\square$

Lemma 7.2. *Seien P, Q orthogonale Projektoren in $\mathbb{R}^2$, sodaß gilt:*

$$\dim U \;=\; \dim V \;=\; 1\,, \qquad \textit{wobei} \quad U := \operatorname{Im} P \;\textit{und}\; V := \operatorname{Im} Q\,.$$

Dann folgt:

$$\delta(U,V) \;:=\; \|P-Q\| \;=\; |x \wedge y| \;=\; |\sin\alpha| \qquad (x \in U, y \in V \;\textit{mit}\; |x| = |y| = 1)\,,$$

wobei α den Winkel zwischen x und y bezeichnet. Folglich ist δ eine vollständige Metrik auf $\mathbb{P}^1$, dem proketiven Raum aller eindimensionalen Teilräume des $\mathbb{R}^2$.

Beweis. Die zweite Gleichung wurde schon auf S. 37 gezeigt.

$\|P-Q\| \;=\; |x \wedge y|$: Wie auf S. 37 folgt weiter:

$$\begin{aligned}
|x \wedge y| \;&=\; \det \begin{pmatrix} \langle x,x\rangle & \langle x,y\rangle \\ \langle y,x\rangle & \langle y,y\rangle \end{pmatrix}^{1/2} \\
&=\; \sqrt{1 - \langle x,y\rangle^2} \\
&=\; \sqrt{\langle x,y\rangle^2 + \langle x,y^\perp\rangle^2 - \langle x,y\rangle^2} \\
&=\; |\,\langle x,y^\perp\rangle\,| \\
&=\; \|\,(I-Q)\,P\,\| \\
&=\; \|\,(P-Q)\,P\,\| \;\;\le\;\; \|\,P-Q\,\|\,,
\end{aligned}$$

wobei noch die Idempotenz orthogonaler Projektoren benutzt wurde sowie die Tatsache $\|AB\| = \|BA\|$ für orthogonale Projektoren A, B.
Andererseits folgt für $w \in \mathbb{R}^2$:

$$\begin{aligned}
|\,(P-Q)w\,|^2 \;&=\; |\,(P-QP)w - (Q-QP)w\,|^2 \\
&=\; |\,(I-Q)Pw - Q(I-P)w\,|^2 \\
&=\; |\,(I-Q)Pw\,|^2 \;+\; |\,Q(I-P)w\,|^2 \\
&\le\; \|\,(I-Q)P\,\|^2\;|\,Pw\,|^2 \;+\; \underbrace{\|\,Q(I-P)\,\|}_{\|\,(I-Q)P\,\|}{}^2\;\;|\,(I-P)w\,|^2 \\
&=\; \|\,(I-Q)P\,\|^2\,,
\end{aligned}$$

also

$$\|\,P-Q\,\| \;\le\; \|\,(I-Q)P\,\|\,.$$

Insgesamt ist also gezeigt:

$$\|\,P-Q\,\| \;=\; \|\,(I-Q)P\,\| \;=\; |x \wedge y|\,.$$

□

Wie bereits angekündigt dient der folgende deterministische Satz dazu, den Satz von Furstenberg-Kesten anwenden zu können.

Satz 7.3 (Goldsheid-Margulis). *Sei $(A_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{R}^{d\times d}$ mit den Eigenschaften:*

$$\limsup_{n\to\infty} \frac{1}{n} \log \|A_n\| \;\leq\; 0 \tag{4}$$

und $\Phi_n := A_n \cdots A_1$ *erfülle*

$$\lim_{n\to\infty} \frac{1}{n} \log \| \wedge^i \Phi_n\| \;=:\; \gamma^{(i)} \;\in\; \mathbb{R} \cup \{-\infty\} \tag{5}$$

für jedes $i = 1, \ldots, d$ *. Dann gilt:*

i) Es existiert (in der Topologie der Operatornorm) der Limes

$$\Psi \;:=\; \lim_{n\to\infty} (\, \Phi_n^* \, \Phi_n \,)^{1/2n} \;\geq\; 0\,.$$

Definiert man nun sukzessiv Λ_i *für* $i = 1, \ldots, d$ *durch* $\Lambda_1 + \cdots + \Lambda_i \;=\; \gamma^{(i)}$ *(falls* $\gamma^{(i)} = -\infty$ *ist, so setze* $\Lambda_i = -\infty$*), dann sind die Eigenwerte von* Ψ *gerade*

$$e^{\Lambda_1}, \ldots, e^{\Lambda_d}$$

und es gilt

$$\Lambda_i = \lim_{n\to\infty} \frac{1}{n} \log \delta_i(\Phi_n) \qquad\quad (i = 1, \ldots, d)\,.$$

ii) Seien

$$e^{\lambda_p} < \cdots < e^{\lambda_1}$$

die verschiedenen (!) Eigenwerte von Ψ *(wobei* $\lambda_p = -\infty$ *sein kann),* $U_p, \ldots, U_1$ *seien die zugehörigen Eigenräume mit* $d_i := \dim U_i$ *und es sei*

$$V_i \;:=\; \begin{cases} \{0\}\,, & i = p+1 \\ U_p \oplus \cdots \oplus U_i\,, & i = 1, \ldots, p\,. \end{cases}$$

Dann gilt:

$$V_{p+1} \subset V_p \subset V_{p-1} \subset \cdots \subset V_1 = \mathbb{R}^d$$

und für jedes $x \in \mathbb{R}^d \setminus \{0\}$ *existiert der* Lyapunov-Exponent

$$\lambda(x) \;:=\; \lim_{n\to\infty} \frac{1}{n} \log |\Phi_n x|\;;$$

es gilt für alle $i = 1, \ldots, p$ *:*

$$x \in V_i \setminus V_{i+1} \iff \lambda(x) = \lambda_i$$

bzw. äquivalent hierzu:

$$V_i \;=\; \{\, x \in \mathbb{R}^d \,:\, \lambda(x) \leq \lambda_i \,\}\,.$$

Beweis. Im Fall $d = 1$ ist nichts zu zeigen, da dann $\Phi_n \in \mathbb{R}$ ist und die Aussagen direkt aus den Voraussetzungen folgen.
Zur Vereinfachung wollen wir uns nun auf den Fall $d = 2$ beschränken; der allgemeine Fall läßt sich ähnlich beweisen, erfordert allerdings mehr Arbeit (siehe hierzu Arnold [AR 98] S.S.144-152).

$\boxed{\Lambda_i = \lim_{n\to\infty} \frac{1}{n} \log \delta_i(\Phi_n) \ \text{ für } i = 1,2}$ Dies folgt aus (5) mit 6.14 iii):

$$\Lambda_1 \equiv \gamma^1 \overset{(5)}{=} \lim_{n\to\infty} \frac{1}{n} \log \|\Phi_n\| \overset{6.14}{=} \lim_{n\to\infty} \frac{1}{n} \log \delta_1(\Phi_n)\ ;$$

falls nun $\Lambda_1 = -\infty$ ist, also $\gamma^1 = -\infty$, so ist wegen (5) auch $\gamma^2 = -\infty = \Lambda_2$; andererseits hat man in diesem Fall;

$$\frac{1}{n} \log \delta_2(\Phi_n) \leq \frac{1}{n} \log \delta_1(\Phi_n) \longrightarrow -\infty.$$

Ist $\Lambda_1 > -\infty$, so folgt:

$$\begin{aligned} \Lambda_2 \quad \equiv \quad \gamma^2 - \Lambda_1 \quad &= \quad \lim_{n\to\infty} \frac{1}{n} \log \underbrace{\| \wedge^2 \Phi_n\|}_{\delta_1(\Phi_n)\delta_2(\Phi_n)} - \lim_{n\to\infty} \frac{1}{n} \log \delta_1(\Phi_n) \\ &= \quad \lim_{n\to\infty} \frac{1}{n} \log \delta_2(\Phi_n)\,. \end{aligned}$$

$\boxed{\text{Konvergenz der Operatoren und Lyapunov-Exponenten}}$ Sei nun

$$\Phi_n = V_n D_n O_n$$

die Singulärzerlegung von Φ_n, mit

$$D_n \ = \ \begin{pmatrix} \delta_1(\Phi_n) & 0 \\ 0 & \delta_2(\Phi_n) \end{pmatrix}.$$

hiermit ergibt sich:

$$(\Phi_n^* \, \Phi_n)^{1/2n} \ = \ (O_n^* D_n^2 O_n)^{1/2n} \ = \ O_n^* D_n^{1/n} O_n\ ;$$

diese Matrix hat als Eigenwerte $\delta_1(\Phi_n)^{1/n}$ und $\delta_2(\Phi_n)^{1/n}$, die gemäß dem oben gezeigten gegen e^{Λ_1} und e^{Λ_2} konvergieren; man hat also folgende Konvergenz:

$$D_n^{1/n} \equiv \begin{pmatrix} \delta_1^{1/n}(\Phi_n) & 0 \\ 0 & \delta_2^{1/n}(\Phi_n) \end{pmatrix} \xrightarrow{n\to\infty} \begin{pmatrix} e^{\Lambda_1} & 0 \\ 0 & e^{\Lambda_2} \end{pmatrix}.$$

Die Schwierigkeit besteht nun darin, daß die Konvergenz von $O_n$im Allgemeinen nicht gewährleistet ist; es genügt jedoch, daß die jeweiligen Eigenräume konvergieren, wozu 7.1 gezeigt worden ist.

1. FALL: $\Lambda_1 = \Lambda_2 =: \lambda_1$: Wie eben gesehen gilt also $D_n^{1/n} \to e^{\lambda_1} I$ und 7.1 liefert

$$(\Phi_n^* \Phi_n)^{1/2n} \xrightarrow{n\to\infty} e^{\lambda_1} I$$

mit $\bar{P}_1^n := P_1^n + P_2^n$. Ferner gilt dann sofort: $V_1 \equiv U_1 = \mathbb{R}^2$, $p = 1$ und $d_1 = 2$. Daher ist nur noch zu zeigen, daß für alle $x \in \mathbb{R}^2 \setminus \{0\}$ gilt:

$$\lambda(x) \;\equiv\; \lim_{n\to\infty} \frac{1}{n} \log \|\Phi_n x\| \;=\; \lambda_1 \, .$$

Hierfür sei zunächst $\lambda_1 > -\infty$; dann folgt aus der bereits bewiesenen Charakterisierung von Λ_1, daß für jedes $\epsilon > 0$ ein $c_\epsilon \in (0,\infty)$ existiert, mit

$$\frac{1}{c_\epsilon} e^{n(\lambda_1 - \epsilon)} \;\le\; \delta_i(\Phi_n) \;\le\; c_\epsilon e^{n(\lambda_1+\epsilon)}, \qquad\qquad i = 1,2.$$

Setzt man noch $x_n := O_n x$, so folgt

$$|\Phi x| \;=\; |V_n D_n O_n x| \;=\; |D_n x_n| = \left(\delta_1(\Phi_n)^2 (x_n^1)^2 + \delta_2(\Phi_n)^2 (x_n^2)^2\right)^{1/2}$$

mit x_n^i den beiden Komponenten von x_n; insgesamt ist also

$$\frac{|x|}{c_\epsilon}\, e^{n(\lambda_1-\epsilon)} \;\le\; |\Phi_n x| \;\le\; |x| c_\epsilon e^{n(\lambda_1+\epsilon)} \, ,$$

weshalb folgt, daß $\lambda(x) = \lambda_1$ ist.
Ist $\lambda = -\infty$, so kann man ebenso für jedes $r < 0$ ein $c_r \in (0,\infty)$ finden mit

$$0 \;\le\; \delta_i(\Phi_n) \;\le\; c_r e^{nr} \, , \qquad\qquad i = 1,2 \, .$$

Wie oben folgt dann:

$$0 \;\le\; |\Phi_n x| \;\le\; |x| c_r e^{nr},$$

woraus wie oben folgt: $\lambda(x) = \lambda_1$. Somit ist der Satz im Fall $\Lambda_1 = \Lambda_2$ bewiesen.

2. FALL: $\lambda_1 \equiv \Lambda_1 > \Lambda_2 \equiv \lambda_2$: Hier gilt also

$$D_n^{1/n} \equiv \begin{pmatrix} \delta_1^{1/n}(\Phi_n) & 0 \\ 0 & \delta_2^{1/n}(\Phi_n) \end{pmatrix} \xrightarrow{n\to\infty} \begin{pmatrix} e^{\lambda_1} & 0 \\ 0 & e^{\lambda_2} \end{pmatrix} .$$

Um hier die Existenz von Ψ zu beweisen, müssen wir zeigen, daß die Orthoprojektionen P_1^n, P_2^n auf die Eigenräume U_1^n, U_2^n von $(\Phi_n^*\Phi_n)^{1/2n}$ gegen Orthoprojektionen P_1, P_2 konvergieren, denn dann folgt wegen 7.1 gerade:

$$(\Phi_n^* \, \Phi_n)^{1/2n} \;\;\xrightarrow{n\to\infty}\;\; e^{\lambda_1} P_1 + e^{\lambda_2} P_2 \;\;=:\;\; \Psi \, .$$

Dies wird mittels eines Cauchy-Argumentes im folgenden Lemma gezeigt. Hierfür bemerken wir, daß die Eigenvektoren von $(\Phi_n^*\Phi_n)^{1/2n} \;=\; O_n^* D_n^{1/n} O_n$ gerade gegeben sind durch $u_i^n := O_n^* e_i$ $(i = 1,2)$, wobei (e_1, e_2) die Standardbasis des $\mathbb{R}^2$ bezeichnet. Insbesondere ist $U_i^n = \text{span}\,(u_i^n)\, ,\; i = 1,2.$

Lemma 7.4. *In obiger Situation („2. Fall" im Beweis von Satz 7.3) gilt:*

$$\limsup_{n\to\infty} \frac{1}{n} \log \delta\big(U_i^n, U_i^{n+1}\big) \;\le\; \lambda_2 - \lambda_1 \;<\; 0 \qquad (i = 1,2).$$

Insbesondere ist $(U_i^n)_{n\in\mathbb{N}}$ *($i = 1,2$) eine Cauchy-Folge im projektiven Raum* $\mathbb{P}^1$*, konvergiert also gegen ein* $U_i \in P^1$*. Hierüber behaupten wir ferner, daß diese Konvergenz mit „exponentieller Geschwindigkeit" stattfindet:*

$$\limsup_{n\to\infty} \frac{1}{n} \log \delta\big(U_i^n, U_i\big) \;\le\; \lambda_2 - \lambda_1 \qquad (i = 1,2).$$

Beweis von 7.4. Ohne Einschränkung sei hierbei $i = 2$, da $U_1(n)$ orthogonal zu $U_2(n)$ ist, aber die Metrik δ auf $\mathbb{P}^1$ invariant gegen orthogonale Transformationen ist.
Wegen der Orthogonalität aller (u_1^{n+1}, u_2^{n+1}) kann man u_2^n darstellen als

$$u_2^n \;=\; \alpha_n u_1^{n+1} + \beta_n u_2^{n+1} \qquad (n \in \mathbb{N})\,.$$

1) $\delta\big(U_2^n, U_2^{n+1}\big) = |\alpha_n|$, denn:

$$\begin{aligned}\delta\big(U_2^n, U_2^{n+1}\big) \;\overset{7.2}{=}\; |u_2^n \wedge u_2^{n+1}| \;&\equiv\; |(\alpha_n u_1^{n+1} + \beta_n u_2^{n+1}) \wedge u_2^{n+1}| \\ &=\; |\alpha_n|\, |u_1^{n+1} \wedge u_2^{n+1}| \\ &=\; |\alpha_n|\,,\end{aligned}$$

wobei die Orthonormalität von u_1^n und u_2^n benutzt wurde.

2) $\delta\big(U_2^n, U_2^{n+1}\big) \le \|A_{n+1}\| \frac{\delta_2(\Phi_n)}{\delta_1(\Phi_{n+1})}$, denn: Zunächst ist

$$\begin{aligned}|\Phi_{n+1} u_2^n| \;&\equiv\; |\alpha_n \Phi_{n+1} u_1^{n+1} + \beta_n \Phi_{n+1} u_2^{n+1}| \\ &\equiv\; |\alpha_n\, V_{n+1} D_{n+1} O_{n+1}\, O_{n+1}^* e_1 \;+\; \beta_n\, V_{n+1} D_{n+1} O_{n+1}\, O_{n+1}^* e_2| \\ &=\; |\alpha_n\, \delta_1(\Phi_{n+1})\, V_{n+1} e_1 \;+\; \beta_n\, \delta_2(\Phi_{n+1})\, V_{n+1} e_2| \\ &\overset{\text{orth.}}{\ge}\; |\alpha_n\, \delta_1(\Phi_{n+1})\, V_{n+1} e_1| \\ &=\; |\alpha_n|\, \delta_1(\Phi_{n+1})\,;\end{aligned}$$

andererseits ist

$$|\Phi_{n+1} u_2^n| \;\equiv\; |A_{n+1} \Phi_n u_2^n| \;\le\; \|A_{n+1}\|\, |\Phi_n u_2^n| \;=\; \|A_{n+1}\|\, \delta_2(\Phi_n)\,,$$

insgesamt also

$$\delta\big(U_2^n, U_2^{n+1}\big) \;\overset{1)}{=}\; |\alpha_n| \;\le\; \frac{|\Phi_{n+1} u_2^n|}{\delta_1(\Phi_{n+1})} \;\le\; \|A_{n+1}\|\, \frac{\delta_2(\Phi_n)}{\delta_1(\Phi_{n+1})}\,.$$

3) Erste Behauptung des Lemmas: Mit dem eben gezeigten folgt:

$$\begin{aligned}
\limsup_{n\to\infty} \frac{1}{n} \log \delta\big(U_2^n, U_2^{n+1}\big) &\leq \limsup_{n\to\infty} \frac{1}{n} \log \|A_{n+1}\| \\
&\quad + \limsup_{n\to\infty} \frac{1}{n} \log \delta_2(\Phi_n) \\
&\quad - \liminf_{n\to\infty} \frac{1}{n} \log \delta_1(\Phi_{n+1}) \\
&\leq 0 + \lambda_2 - \lambda_1 \,,
\end{aligned}$$

wobei die erste Voraussetzung von Satz 7.3 und die bereits bewiesene Konvergenzaussage benutzt wurden.

4) $(U_2^n)_n$ konvergiert in P^1 gegen ein U_2: Da δ eine vollständige Metrik ist, ist zu zeigen, daß $(U_2^n)_n$ eine δ-Cauchy-Folge ist. Sei hierzu $\varepsilon < \lambda_1 - \lambda_2$; gemäß dem eben gezeigten kann man ein $n_0 \in \mathbb{N}$ wählen, daß

$$\frac{1}{n} \log \delta\big(U_2^n, U_2^{n+1}\big) \;<\; \lambda_2 - \lambda_1 + \varepsilon \quad (< 0) \qquad\qquad (\forall\, n \geq n_0)$$

ist. Dann folgt aber für $n_0 \leq m \leq n$:

$$\begin{aligned}
\delta\big(U_2^n, U_2^{n+1}\big) &\leq \sum_{k=m}^{n-1} \delta\big(U_2^k, U_2^{k+1}\big) \\
&\leq \sum_{k=m}^{n-1} \mathrm{e}^{k(\lambda_2-\lambda_1+\varepsilon)} \\
&\leq \sum_{k=m}^{\infty} \mathrm{e}^{k(\lambda_2-\lambda_1+\varepsilon)} \\
&= \frac{\mathrm{e}^{m(\lambda_2-\lambda_1+\varepsilon)}}{1-\mathrm{e}^{\lambda_2-\lambda_1+\varepsilon}} \quad \xrightarrow{m\to\infty} \quad 0\,,
\end{aligned}$$

wobei die Summenformel für geometrische Reihen einging.

5) Zweite Behauptung des Lemmas: Mit der eben benutzten Argumentation folgt auch:

$$\delta\big(U_2^n, U_2\big) \;\leq\; \mathrm{e}^{n(\lambda_2-\lambda_1+\varepsilon)} \, \frac{1}{1-\mathrm{e}^{\lambda_2-\lambda_1+\varepsilon}}$$

und daher

$$\limsup_{n\to\infty} \frac{1}{n} \log \delta\big(U_2^n, U_2\big) \;\leq\; \lambda_2 - \lambda_1 + \varepsilon\,;$$

nun folgt die Behauptung mit $\varepsilon \to 0$. $\square$

Lem. 7.4

Fortsetzung des Beweises von 7.3. Als Orthoprojektionen P_1, P_2 wählen wir nun natürlich die Projektionen auf die gemäß 7.4 existierenden Räume U_1, U_2. Wegen 7.2 und 7.4 folgt die Konvergenz

$$P_i^n \xrightarrow{n\to\infty} P_i \qquad (i = 1, 2)\,.$$

Insgesamt folgt also

$$(\Phi_n^* \, \Phi_n)^{1/2n} \xrightarrow{n\to\infty} e^{\lambda_1} P_1 + e^{\lambda_2} P_2 \ =: \ \Psi$$

Es ist also nur noch die Behauptung über die *Lyapunov-Exponenten* nachzuweisen; hierbei ist $V_2 = U_2 \subset \mathbb{R}^2 = V_1$, sodaß nun noch zu zeigen ist:

$$\begin{aligned} x \in V_2 \setminus \{0\} \quad &\Longrightarrow \quad \lim_{n\to\infty} \frac{1}{n} \log \|\Phi_n x\| = \lambda_2 \qquad \text{und} \\ x \in \mathbb{R}^2 \setminus V_2 \quad &\Longrightarrow \quad \lim_{n\to\infty} \frac{1}{n} \log \|\Phi_n x\| = \lambda_1 \,; \end{aligned}$$

wobei jeweils oE $|x| = 1$ angenommen werden kann.

$\underline{x \in V_2 \setminus \{0\} \Rightarrow \lim \frac{1}{n} \log |\Phi_n x| = \lambda_2}$: Wir stellen x dar als

$$x \ = \ \alpha_n u_1^n + \beta_n u_2^n \ ,$$

also wiederum

$$\Phi_n x \ = \ \alpha_n \Phi_n u_1^n + \beta_n \Phi_n u_2^n \ = \ \alpha_n \, \delta_1(\Phi_n) \, V_n e_1 + \beta_n \, \delta_2(\Phi_n) \, V_n e_2 \ ,$$

und daher

$$|\beta_n| \, \delta_2(\Phi_n) \ \le \ \left[\alpha_n^2 \, \delta_1(\Phi_n)^2 + \beta_n^2 \, \delta_2(\Phi_n)^2 \right]^{1/2} \ = \ |\Phi_n x| \ ;$$

wie im Beweisteil 1) von 7.4 folgt aus 7.2: $\delta\big(U_2^n, U_2^{n+1}\big) = |\alpha_n|$, da $x \in V_2 = U_2$ ist; also folgt wegen 7.4 auch:

$$\limsup_{n\to\infty} \frac{1}{n} \log |\alpha_n| \ = \ \limsup_{n\to\infty} \frac{1}{n} \log \delta\big(U_2^n, U_2\big) \ \le \ \lambda_2 - \lambda_1 \ < \ 0 \,;$$

daher folgt:

$$\beta_n^2 \ = \ 1 - \alpha_n^2 \xrightarrow{n\to\infty} 1 \,.$$

und somit insgesamt:

$$\begin{aligned} \lambda_2 \ &= \ \lim_{n\to\infty} \frac{1}{n} \log \left(\, |\beta_n| \, \delta_2(\Phi_n) \, \right) \\ &\le \ \liminf_{n\to\infty} \frac{1}{n} \log |\Phi_n x| \ \le \ \limsup_{n\to\infty} \frac{1}{n} \log |\Phi_n x| \\ &= \ \frac{1}{2} \limsup_{n\to\infty} \frac{1}{n} \log \left[\, \alpha_n^2 \, \delta_1(\Phi_n)^2 + \beta_n^2 \, \delta_2(\Phi_n)^2 \right] \\ &\le \ \frac{1}{2} \max \left\{ \limsup_{n\to\infty} \frac{1}{n} \log \alpha_n^2 \, \delta_1(\Phi_n)^2 \ , \ \limsup_{n\to\infty} \frac{1}{n} \log \beta_n^2 \, \delta_2(\Phi_n)^2 \right\} \\ &\le \ \max \{ \, (\lambda_2 - \lambda_1) + \lambda_1 \ , \ 0 + \lambda_2 \, \} \\ &= \ \lambda_2 \,. \end{aligned}$$

$\underline{x \in \mathbb{R}^2 \setminus V_2 \Rightarrow \lim \frac{1}{n} \log |\Phi_n x| = \lambda_1}$: Hier stellen wir x dar als

$$x = \alpha u + \beta v$$

mit Einheitsvektoren $u \in U_1$ und $v \in U_2 = V_2$; diese schreiben wir als

$$v = \alpha_n u_1^n + \beta_n u_2^n \qquad \text{bzw.} \qquad u = \gamma_n u_1^n + \delta_n u_2^n \,.$$

Auch hier folgt aus 7.4 notwendig: $\alpha_n \to 0$, $\delta_n \to 0$ und somit $|\beta_n| \to 1$, $|\gamma_n| \to 1$ (im projektiven Raum gilt ja wegen 7.4: $u_1^n \to u$ und $u_2^n \to v$).
Damit gilt wie oben:

$$\begin{aligned} |\alpha|\,|\gamma_n|\,\delta_1(\Phi_n) &\leq \left[(\alpha\gamma_n + \beta\alpha_n)^2\,\delta_1(\Phi_n)^2 + (\alpha\delta_n + \beta\beta_n)^2\,\delta_2(\Phi_n)^2 \right]^{1/2} \\ &= |\Phi_n x| \,; \end{aligned}$$

beachtet man noch, daß aufgrund der Lage von x immer $\alpha = \langle x, u \rangle \neq 0$ ist, so folgt insgesamt wiederum:

$$\begin{aligned} \lambda_1 &= \lim_{n\to\infty} \frac{1}{n} \log\left(|\alpha|\,|\gamma_n|\,\delta_1(\Phi_n) \right) \\ &\leq \liminf_{n\to\infty} \frac{1}{n} \log |\Phi_n x| \;\leq\; \limsup_{n\to\infty} \frac{1}{n} \log |\Phi_n x| \\ &= \frac{1}{2} \limsup_{n\to\infty} \frac{1}{n} \log \left[(\alpha\gamma_n + \beta\alpha_n)^2\,\delta_1(\Phi_n)^2 + (\alpha\delta_n + \beta\beta_n)^2\,\delta_2(\Phi_n)^2 \right] \\ &\leq \lambda_1 \,. \end{aligned}$$

Somit sind alle Aussagen von 7.3 bewiesen. □

Um den Satz von Goldsheid-Margulis anwenden zu können, bleibt noch, die erste Voraussetzung im speziellen Fall stationärer zufälliger Matrizen nachzuprüfen:

Lemma 7.5. *Sei* $X : \Omega \to \mathbb{R} \cup \{-\infty\}$ *eine ZV mit* $X^+ \in L^1(\Omega, \mathscr{F}, \mathbb{P})$. *Dann ist*

$$\Omega_1 := \left\{ \limsup_{n\to\infty} \frac{1}{n} X \circ \varphi^{n-1} \leq 0 \right\}$$

invariant und es ist $\mathbb{P}(\Omega_1) = 1$.

Beweis. Die Invarianz folgt aus der Definition von Ω_1. Ferner trägt Ω_1 volles Maß hat, denn:

$$\begin{aligned} \sum_{n=1}^{\infty} \mathbb{P}\left\{ \frac{1}{n} X \circ \varphi^{n-1} > \varepsilon \right\} &\overset{\varphi \text{ m.t.}}{=} \sum_{n=1}^{\infty} \mathbb{P}\{X > \varepsilon n\} = \sum_{n=1}^{\infty} \mathbb{P}\left\{X^+ > \varepsilon n\right\} \\ &\leq \frac{1}{\varepsilon}\, \mathbb{E}(X^+) < \infty \,, \end{aligned}$$

also nach Borel-Cantelli: $\mathbb{P}(\Omega_1) = 1$. □

Um den Hauptsatz zu erhalten wird 7.5 angewandt auf $X := \log \|A\|$. Somit folgt:

Theorem 7.6 (Multiplikativer Ergodensatz, Oseledets). *Sei* $A : \Omega \to \mathbb{R}^{d\times d}$ *eine zufällige Matrix auf* $(\Omega, \mathscr{F}, \mathbb{P}, \varphi)$ *und*

$$A_n := \begin{cases} \left(A\circ\varphi^{n-1}\right)\left(A\circ\varphi^{n-2}\right)\cdots(A\circ\varphi)\,A\,, & n\in\mathbb{N}\,,\\ I\,, & n=0\,,\end{cases}$$

der hiervon erzeugte Kozykel auf $\mathbb{R}^d$. *Es gelte*

$$\log^+ \|A\| \in L^1(\Omega, \mathscr{F}, \mathbb{P})\,.$$

Dann existiert $\tilde\Omega \in \mathscr{F}$ *mit* $\tilde\Omega \subset \varphi^{-1}(\tilde\Omega)$ *und* $\mathbb{P}(\tilde\Omega) = 1$, *sodaß für jedes* $\omega\in\tilde\Omega$ *gilt:*

i) Es existiert

$$\Psi(\omega) := \lim_{n\to\infty}\left(A_n^*(\omega)\,A_n(\omega)\right)^{1/2n} \ge 0$$

ii) Sind

$$e^{\lambda_{p(\omega)}(\omega)} < \cdots < e^{\lambda_1(\omega)}$$

die verschiedenen Eigenwerte von $\Psi(\omega)$ *(wobei* $\lambda_{p(\omega)}(\omega) = -\infty$ *sein kann), und* $U_{p(\omega)}(\omega), \dots, U_1(\omega)$ *die zugehörigen Eigenräume mit* $d_i(\omega) := \dim U_i(\omega)$, *so gilt:*

$$\big(\lambda_i\circ\varphi\big)(\omega) = \lambda_i(\omega)\,,\quad \big(d_i\circ\varphi\big)(\omega) = d_i(\omega)\,,\quad \text{und}\quad 1\le i\le p_i(\omega) = \big(p_i\circ\varphi\big)(\omega)\,.$$

iii) Definiert man

$$V_i(\omega) := \begin{cases}\{0\}\,, & i = p(\omega)+1\\ U_{p(\omega)}(\omega)\oplus\cdots\oplus U_i(\omega)\,, & i = 1,\dots,p(\omega)\,.\end{cases}$$

Dann gilt also

$$V_{p(\omega)+1}(\omega) \subset V_{p(\omega)}(\omega) \subset V_{p(\omega)-1}(\omega) \subset \cdots \subset V_1(\omega) = \mathbb{R}^d$$

und für jedes $x\in\mathbb{R}^d\setminus\{0\}$ *existiert*

$$\lambda(\omega, x) := \lim_{n\to\infty}\frac{1}{n}\log|A_n(\omega)x|\,;$$

es gilt für alle $i = 1,\dots,p(\omega)$:

$$x\in V_i(\omega)\setminus V_{i+1}(\omega) \iff \lambda(\omega,x) = \lambda_i(\omega)$$

bzw. äquivalent hierzu:

$$V_i(\omega) = \{\, x\in\mathbb{R}^d : \lambda(\omega,x)\le\lambda_i(\omega)\,\}\,.$$

iv) Ist φ *ergodisch, so sind* p, λ_i *und* d_i *auf* $\tilde\Omega$ *konstant* $\mathbb{P}$*-f.s..*

Beweis. Aufgrund der Integrabilitätsvoraussetzung ist 7.5 anwendbar mit $X := \log\|A\|$ und liefert die invariante Menge

$$\tilde{\Omega}_1 \;:=\; \left\{ \omega \in \Omega \,:\, \limsup_{n\to\infty} \frac{1}{n} \log\|A(\varphi^{n-1}\omega)\| \leq 0 \right\}$$

vollen Masses. Wir wenden nun das deterministische MET 7.3 an auf

$$A_n^\omega \;:=\; A(\varphi^{n-1}\omega) \qquad \text{und} \qquad \Phi_n^\omega \;\equiv\; A_n^\omega \cdots A_1^\omega \overset{\text{Kozykel}}{=} A_n(\omega)\,,$$

wobei (4) nach Definition auf $\tilde{\Omega}_1$ erfüllt ist und (5) wegen des Satzes von Furstenberg-Kesten 6.15 auf einer vorwärts invarianten Menge $\tilde{\Omega}_2$ vollen Maßes gilt; folglich ist 7.3 anwendbar für jedes $\omega \in \tilde{\Omega}_1 \cap \tilde{\Omega}_2 =: \tilde{\Omega}$, einer vorwärts invarianten Menge vollen Maßes, und liefert mit 6.15 die obigen Behauptungen □

Definition 7.7. Die Funktionen λ_i aus dem Satz von Oseledets heißen die *Lyapunov-Exponenten* des linearen Kozykels $(A_n)_{n\in\mathbb{N}_0}$.

Die Lyapunov-Exponenten sind also die Analoga zu den Eigenwerte einer Matrix, siehe 6.1. Die Räume V_i (für $i = 1,\ldots,p$) sind allerdings nicht die Analoga der Eigenräume im Deterministischen. Dazu muß man die Theorie erweitern auf Zeitskala $\mathbb{Z}$, siehe Arnold [AR 98] Theorem 3.4.11. .

Notationen

$\mathbb{R}_+$	$\{t \in \mathbb{R} : t \geq 0\}$
$\mathbb{N}_0$	$\mathbb{N} \cup \{0\}$
$s^\pm$	$(\pm s) \vee 0$; Positiv- bzw. Negativteil einer reellen Zahl oder Funktion s
$\equiv$	Gleichheit nach Definition
$\vert\ \vert$	Norm
$\Vert\ \Vert$	Operatornorm
$M \,\dot\cup\, N$	disjunkte Vereinigung von M und N
$\mathscr{B}(X)$	Borel-σ-Algebra auf dem topologischen Raum X
$\mathscr{B}^n$	$\mathscr{B}(\mathbb{R}^n)$
$\delta_1(A) \geq \cdots \geq \delta_d(A)$	Singulärwerte von $A \in \mathbb{R}^{d\times d}$
$\mathbb{E}(f)$	$\int f\, d\mathbb{P}$; Erwartungswert einer Funktion f nach dem Wahrscheinlichkeitsmaß $\mathbb{P}$
$\mathbb{E}(f\vert\mathscr{F})$	bedingte Erwartung der ZV f unter $\mathscr{F}$
$\mathscr{I}$	σ-Algebra der meßbaren, invarianten Mengen
$\sigma(\mathscr{M})$	von einer Familie $\mathscr{M}$ von Mengen bzw. Funktionen erzeugte σ-Algebra
ZV	Zufallsvariable
oE	ohne Einschränkung

Literaturverzeichnis

[AR 98] L. Arnold. *Random Dynamical Systems.* Springer-Verlag, Berlin Heidelberg, 1998.

[G-M 89] I.Y. Goldsheid & G.A. Margulis. Lyapunov Indices of a product of random matrices. *Russian Mathematical Surveys*, 44:11-71, 1989.

[HM 74] P. Halmos. *Measure Theory.* Springer-Verlag, New York, 1974.

[KB 97] O. Kallenberg. *Foundations of Modern Probability.* Springer-Verlag, New York, 1997.

[KI 68] J.F.C. Kingman. The ergodic theory of subadditive stochastic processes. *J. Royal Statist. Soc. Ser. B*, 30:499-510, 1968.

[M-T 93] S. Meyn & R. Tweedie. *Markov chains and stochastic stability* . Springer, London, 1993.

[OS 68] V.I. Oseledets. A multiplicative ergodic theorem. Lyapunov characteristic numbers for dynamical systems. *Trans. Moscow Math. Soc.*, 19:197-231, 1968.

[Sh 95] A.N. Shiryaev. *Probability.* (second edition) GTM 95 Springer-Verlag, Berlin Heidelberg NewYork, 1995.

Index

additive Kozykel-Eigenschaft 25
äußeres Produkt 33

Bernoulli-Shift 18, 21

Ehrenfest-Modell von Diffusionen 8
Ergodensatz (Birkhoff)
22, 29, 30
ergodisch 19

Furstenberg-Kesten, Satz von
39, 51

invariant
Maß 8
Menge 19

Kingman, Satz von
siehe: subadditiver Ergodensatz
Konsistenzsatz von Kolmogorov 3
Kozykel 39, 41
additiver 25
erzeugter 41
Lyapunov-Exponent 51

Lyapunov-Exponent 51

Markovkette 3
Markoveigenschaft 4
starke 5
maßtreue Abbildung 19
Maximal-ergodisches Lemma (Hopf)
22, 23
Maximalungleichung 27, 30
multiplikativer Ergodensatz (Oseledets)
50
deterministischer 43, 51

rekurrent 6
Riesz-Lemma 26, 27
Rotation des Kreises 17, 20, 24

Semiring 1
Shift auf dem Pfadraum 4
Singulärwertzerlegung 33
starkes Gesetz der großen Zahlen 24
stationär
Maß 8
Prozeß 16
Stoppzeit 5
subadditive Folge von Zufallsvariablen 25
subadditiver Ergodensatz (Kingman)
28, 40

transient 6

Übergangswahrscheinlichkeit 1

Weylscher Gleichverteilungssatz 24

zufälliges dynamisches System:
siehe: Kozykel