TU Wien

Seminararbeit WS 15/16

Stationäre Prozesse und Ergodentheorie

Autor: Daniel Waschmann

Betreuer: Dr. Stefan Gerhold

29. Februar 2016

Inhaltsverzeichnis

1. Einleitung

Bei der Untersuchung einer Größe in laufender Zeit treten besteht in der Regel eine gewisse Abhängigkeit der Werte. Unter der Annahme, die gemeinsame Verteilung der Werte bleibe unter zeitlicher Verschiebung gleich, lässt sich die Größe mithilfe eines stationären Prozesses modellieren. Leider ist das Gesetz der Großen Zahlen nur im speziellen Fall einer *iid* (*i*ndependent, *i*dentical *d*istributed) Folge an Zufallsgrößen anwendbar; sofern eine Abhängigkeit besteht, bekommt man keine Aussage über die Konvergenz der Mittelwerte.

Die Ergodentheorie liefert auch in diesem Fall Konvergenzaussagen und findet daher beispielsweise in der Zeitreihenanalyse, z.B. bei gewissen AR (Auto Regression) und MA (Moving Average) -Prozessen, Anwendung. Im Allgemeinen wird der Grenzwert eine Zufallsgröße sein, genauer ein bedingter Erwartungswert, im ergodischen Fall ist die Sigmaalgebra in der Bedingung trivial und der Grenzwert reduziert sich auf eine Konstante. In dieser Arbeit wird neben dem Ergodensatz von Birkhoff, sowie einigen stetigen und multivarianten Verallgemeinerungen, auch die Möglichkeit der Zerlegung eines Prozesses in ergodische Komponenten (Zufallsmaße) vorgestellt.

Die meisten Sätze sind allgemein für eine Zufallsgröße ξ auf einem Raum $(S, \mathcal{S})$ formuliert. Man stelle sich ξ beispielsweise als diskreten Prozess $(\xi_k)_{k\in\mathbb{N}}$ auf dem Folgenraum $(S, \mathcal{S}) = (\mathbb{R}^\infty, \mathcal{B}^\infty)$ vor.

2. Begriffe und Definitionen

- Eine Transformation auf einem Maßraum $(S, \mathcal{S}, \mu)$ heißt *maßerhaltend*, falls $\mu \circ T^{-1} = \mu$.
- Die unter T invarianten Mengen bilden eine σ-Algebra, weiters mit $\mathcal{I}$ bezeichnet.
- $\theta(x_0, x_1, \dots) := (x_1, x_2, \dots)$ bezeichnet den *Shift-Operator*.
- Eine Folge von Zufallsvariablen $X = (X_n)_{n \in \mathbb{N}}$ heißt *stationär*, falls $\theta \circ X \overset{\mathrm{d}}{=} X$.
- $\pi_0(x_0, x_1, \dots) := x_0$ bezeichnet die Projektion einer Folge auf ihr erstes Element.
- Zur vereinfachten Schreibweise sei $\mathbb{E}^{\mathcal{J}} : \xi \mapsto \mathbb{E}[\xi \mid \mathcal{J}]$ und $\mathcal{J}_\xi := \xi^{-1}\mathcal{J}$ für Zufallsgrößen ξ und σ-Algebren $\mathcal{J}$.

3. Der Ergodensatz von Birkhoff

Definition 3.1. (*Ergodische Transformation*) Eine Transformation T auf einem Maßraum $(S, \mathcal{S}, \mu)$ heißt *ergodisch* bzw. *μ-ergodisch*, falls T maßerhaltend ist und für alle $I \in \mathcal{I}$ gilt: $\mu(I) \in \{0, 1\}$.

Je nach Betrachtungsweise sagt man auch μ ist T-ergodisch oder μ ist ein ergodisches Maß für T. Einen stochastischen Prozess $\xi = (\xi_0, \xi_1, \dots)$ nennt man ergodisch, falls das Bildmaß θ-ergodisch ist. Bezeichnet $(\Omega, \mathcal{F}, \mathbb{P})$ den Wahrscheinlichkeitsraum von ξ, so ist $\mathcal{I}_\xi$ im ergodischen Fall $\mathbb{P}$-trivial.

Lemma 3.2.
Vor.: Sei ξ eine Zufallsgröße und T eine messbare Transformation T auf $(S, \mathcal{S})$.
Beh.: Es gilt $T\xi = \xi$ in Verteilung genau dann, wenn $(T^n\xi)$ stationär ist. In diesem Fall ist auch $(f \circ T^n\xi)$ für beliebiges, messbares f stationär. Umgekehrt besitzt jeder stationäre Prozess so eine Darstellung.

Beweis. Sei $T\xi = \xi$. Es gilt

$$\theta(f \circ T^n\xi) = (f \circ T^{n+1}\xi) = (f \circ T^nT\xi) = (f \circ T^n\xi)$$

in Verteilung. Somit ist $(f \circ T^n\xi)$ stationär. Die Umkehrung ist trivial.
Ist $\eta = (\eta_0, \eta_1, \dots)$ stationär, so gilt $\eta_n = \pi_0(\theta^n \circ \eta)$, wobei $\pi_0(x_0, x_1, \dots) = x_0$. ■

Lemma 3.3.
Vor.: Sei ξ eine Zufallsgröße in S mit Verteilung μ und T eine maßerhaltende Transformation.
Beh.: Dann ist ξ T-ergodisch genau dann, wenn $(T^n\xi)$ θ-ergodisch ist. In diesem Fall ist auch $\eta = (f \circ T^n\xi)$ θ-ergodisch für beliebiges, messbares f.

Beweis. Für messbares $f : S \to S'$ definiere $F = (f \circ T^n\xi)_{n\in\mathbb{N}}$, also $F \circ T = \theta \circ F$. Ist $I \subset (S')^\infty$ θ-invariant, dann gilt $T^{-1}F^{-1}I = F^{-1}\theta^{-1}I = F^{-1}I$, also ist $F^{-1}I$ T-invariant in S.
Setzt man ξ als ergodisch voraus, erhält man $\mathbb{P}(\eta \in I) = \mathbb{P}(\xi \in F^{-1}I) \in \{0, 1\}$. Daher ist η θ-ergodisch.
Sei nun $(T^n\xi)_{n\in\mathbb{N}}$ ergodisch. Definiere $F = (T^n)_{n\in\mathbb{N}}$ und $A = \{s \in S^\infty : s_n \in I\}$ für ein

T-invariantes I in S. Dann ist $I = F^{-1}A$ und A ist θ-invariant. Es folgt $\mathbb{P}(\xi \in I) = \mathbb{P}((T^n\xi) \in A) \in \{0,1\}$, d.h. ξ ist ergodisch. ■

Lemma 3.4. (*Maximaler Ergodensatz*):
Vor.: Sei $\xi = (\xi_k)_{k\in\mathbb{N}}$ eine stationäre Folge von integrierbaren Zufallsgrößen und setze $S_n = \xi_1 + \cdots + \xi_n$.
Beh.: Dann ist

$$\mathbb{E}[\xi_1 \mathbb{1}_{[\sup\limits_n S_n > 0]}] \geq 0 \tag{3.1}$$

Beweis. Setze $M_n = \max(S_1, \ldots, S_n)$. Es gilt

$$S_k = \xi_1 + S_{k-1} \circ \theta \leq \xi_1 + (M_n \circ \theta)_+, \qquad k = 1, \ldots, n$$

Durch Maximumsbildung erhält man $M_n \leq \xi_1 + (M_n \circ \theta)_+$, und die Stationarität von ξ liefert

$$\begin{aligned}\mathbb{E}[\xi_1 \mathbb{1}_{[M_n>0]}] &\geq \mathbb{E}[M_n - (M_n \circ \theta)_+ \mathbb{1}_{[M_n>0]}] \\ &\geq \mathbb{E}[(M_n)_+ - (M_n \circ \theta)_+] = 0.\end{aligned}$$

Da $M_n \uparrow \sup\limits_n S_n$ folgt wegen der Integrierbarkeit von ξ_1 die Behauptung mit monotoner Konvergenz. ■

Satz 3.5. (*Ergodensatz von Birkhoff*):
Vor.: Sei ξ eine Zufallsvariable in S mit Verteilung μ, T eine maßerhaltende Transformation und $\mathcal{I}$ die σ-Algebra der T-invarianten Mengen.
Beh.: Für jede messbare Funktion $f \geq 0$ auf S ist

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f(T^k\xi) = \mathbb{E}[f(\xi)|\mathcal{I}_\xi] \quad f.s. \tag{3.2}$$

Für $f \in L^p$ gilt L^p-Konvergenz.

Beweis. Sei $f \in L^1$, $\epsilon > 0$ und setze $\eta_k = f(T^{k-1}\xi), k \geq 0$. Da $\mathbb{E}[\eta_1|\mathcal{I}_\xi]$ eine T-invariante Funktion von ξ ist, ist die Folge $\zeta_k = \eta_k - \mathbb{E}[\eta_1|\xi^{-1}\mathcal{I}]$ nach Lemma 3.2 stationär. Definiere $S_n = \sum\limits_{k=1}^{n} \zeta_k$, $A_\epsilon = \{\limsup\limits_n \frac{S_n}{n} > \epsilon\}$. Nach Lemma 3.4, und da $A_\epsilon \in \xi^{-1}\mathcal{I}$, gilt

$$\begin{aligned}0 \;&\leq\; \mathbb{E}[(\zeta_1 - \epsilon)\mathbb{1}_{A_\epsilon}\mathbb{1}_{[\sup\limits_n (S_n - n\epsilon)\mathbb{1}_{A_\epsilon} > 0]}] = \mathbb{E}[(\zeta_1 - \epsilon)\mathbb{1}_{A_\epsilon}] \\ &=\; \mathbb{E}[\mathbb{E}[\zeta_1|\mathcal{I}_\xi]\mathbb{1}_{A_\epsilon}] - \epsilon\mathbb{P}(A_\epsilon) = -\epsilon\mathbb{P}(A_\epsilon),\end{aligned}$$

woraus $\mathbb{P}(A_\epsilon) = 0$ folgt. Da ϵ beliebig war, ist $\limsup_n \frac{S_n}{n} \leq 0$. Selbiges angewendet auf $-S_n$ liefert $\liminf_n \frac{S_n}{n} \geq 0$, also $\frac{S_n}{n} \to 0$ f.s.
Sei nun $f \in L^p, p \geq 1$. Für ein $r > 0$ liefert die Jensen-Ungleichung

$$\begin{aligned} \mathbb{E}[\mathbb{1}_A \| \frac{1}{n} \sum_{k=1}^{n} f(T^k \xi) \|^p] &\leq \frac{1}{n} \sum_{k=1}^{n} \mathbb{E}[\| f(T^k \xi) \|^p] \\ &\leq r^p \mathbb{P}(A) + \mathbb{E}[\| f(\xi) \|^p \mathbb{1}_{\| f(\xi) \| > r}], \end{aligned}$$

wobei die rechte Seite für $\mathbb{P}(A) \to 0$ und $r \to \infty$ gegen 0 geht. Daher ist die Funktionenfamilie $\{\| \frac{1}{n} \sum_{k=1}^{n} f(T^k \xi) \|^p, n \in \mathbb{N}\}$ gleichgradig integrierbar, woraus die L^p-Konvergenz folgt.
Zuletzt sei $f \geq 0$ beliebig, setze $\mathbb{E}[f(\xi) | \mathcal{I}_\xi] = \eta^*$. Auf $\{\eta^* < \infty\}$ gilt (2.2) f.s., für $\eta^* = \infty$ und $r > 0$ folgt wegen

$$\begin{aligned} \liminf_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} f(T^k \xi) &\geq \lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} (\min(f(T^k \xi), r) \\ &= \mathbb{E}[\min(f(\xi), r) | \mathcal{I}_\xi] \xrightarrow{r \to \infty} \eta^* \qquad f.s. \end{aligned}$$

die Behauptung. ■

Es sei bemerkt, dass durch Zerlegung von f und Positiv- und Negativteil f^+, f^- die Aussage auch für beliebige, messbare f gilt.
Damit lässt sich nun das Gesetz der Großen Zahlen, sowie einige Verallgemeinerungen beweisen:

Korollar 3.6. (*Kolmogorov's 2. Gesetz der großen Zahlen*):
Vor.: Sei $(X_n)_{n \in \mathbb{N}}$ eine Folge von unabhängig, identisch verteilten Zufallsgrößen auf $(\Omega, \mathcal{F}, \mathbb{P})$, deren Erwartungswert $\mathbb{E}[X_1]$ existiert.
Beh.:

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} X_i = \mathbb{E}[X_1] \qquad \mathbb{P} - f.s. \tag{3.3}$$

Beweis. Es ist $X_i = \pi_0(\theta^k X)$. Eine *iid* Folge an Zufallsgrößen ist auch stationär, daher gilt $\theta \circ X \stackrel{\mathrm{d}}{=} X$ und es folgt

$$\mu[\theta^{-1} A] = \mathbb{P}[X^{-1} \theta^{-1} A] = \mathbb{P}[(\theta X)^{-1} A] = \mathbb{P}[X^{-1} A] = \mu[A].$$

Also ist θ maßerhaltend. Sei nun $k \in \mathbb{N}$ beliebig und $I \in \mathcal{I}$, der σ-Algebra der θ-invarianten Mengen. Wegen $\theta^k X \stackrel{\mathrm{d}}{=} X$ gilt

$$\begin{aligned} & X^{-1} I = (\theta^k X)^{-1} I \in \sigma(X_k, X_{k+1}, ...) \\ & \qquad \Longrightarrow X^{-1} I \in \bigcap_{j=1}^{\infty} \sigma(X_j, X_{j+1}, ...) =: \mathfrak{S}_\infty, \end{aligned}$$

also ist $\mathcal{I}_X \subseteq \mathfrak{S}_\infty$. Nach dem 0–1 - Gesetz von Kolmogorov für iid Folgen an Zufallsgrößen [2, S. 98] ist $\mathfrak{S}_\infty$ $\mathbb{P}$-trivial und damit auch $\mathcal{I}_X$. Da π_0 messbar ist, folgt mit dem Ergodensatz von Birkhoff

$$\frac{1}{n}\sum_{i=1}^{n} X_i = \frac{1}{n}\sum_{i=1}^{n} \pi_0(\theta^k X) \xrightarrow{n\to\infty} \mathbb{E}[\pi_0(X)|\mathcal{I}_X] = \mathbb{E}[X_1].$$

■

Korollar 3.7.
Vor.: Sei ξ eine Zufallsgröße in S mit Verteilung μ, T eine maßerhaltende Transformation, f und $f_{m,k}$ messbare Funktionen auf S.
Beh.:

(i) Falls $f_{m,k} \xrightarrow{m,k\to\infty} f$ fast sicher und $\sup\limits_{m,k} \|f_{m,k}\| \in L^1$, dann gilt

$$\lim_{m,n\to\infty} \frac{1}{n}\sum_{k=1}^{n} f_{m,k}(T^k\xi) = \mathbb{E}[f(\xi)|\mathcal{I}_\xi] \quad f.s. \tag{3.4}$$

(ii) Falls $f_{m,k} \xrightarrow{m,k\to\infty} f$ in L^p, $p \geq 0$, so gilt die Konvergenz aus (i) in L^p.

Beweis. (i) Setze $\overline{f}_{m,k} = f_{m,k} - f$, also $\overline{f}_{m,k} \xrightarrow{m,k\to\infty} 0$, und $g_r = \sup\limits_{m,k>r} \|\overline{f}_{m,k}\|$. Es gilt

$$\limsup_{m,n\to\infty} \|\frac{1}{n}\sum_{k=1}^{n} \overline{f}_{m,k}(T^k\xi)\| \;\leq\; \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} g_r(T^k\xi) = \mathbb{E}[g_r(\xi)|\mathcal{I}_\xi]$$

nach Satz 3.5. Da $g_r \to 0$ und in L^1 ist, folgt mit dominierter Konvergenz $\mathbb{E}[g_r(\xi)|\mathcal{I}_\xi] \to 0$ f.s.
(ii) Mit $\overline{f}_{m,k}$ wie eben, der Minowski-Ungleichung und der Invarianz von μ gilt

$$\|\frac{1}{n}\sum_{k=1}^{n} \overline{f}_{m,k} \circ T^k\|_p \;\leq\; \frac{1}{n}\sum_{k=1}^{n} \|\overline{f}_{m,k}\|_p \;\to\; 0.$$

■

Korollar 3.8. (*Ergodensatz in stetiger Zeit*):

Vor.: Sei ξ eine Zufallsgröße in S mit Verteilung μ, $(T_s)_{s\in\mathbb{R}_0^+}$ eine Familie an maßerhaltenden Transformationen bzgl. μ auf S mit der Eigenschaft, dass $T_{s+t} = T_s T_t$.

Beh.: Für jede messbare Funktion $f \geq 0$ auf S gilt

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t f(T_s\xi)ds = \mathbb{E}[f(\xi)|\mathcal{I}_\xi] \quad f.s. \tag{3.5}$$

Falls $f \in L^p(\mu)$, so gilt die Konvergenz auch in L^p.

Beweis. Sei $X_s = f(T_s\xi)$. Mit der Jensen-Ungleichung und dem Satz von Fubini folgt

$$\mathbb{E}[|\frac{1}{t}\int_0^t X_s ds|^p] \;\leq\; \mathbb{E}[\frac{1}{t}\int_0^t X_s^p ds] \;=\; \frac{1}{t}\int_0^t \mathbb{E}[X_s^p]ds \;=\; \mathbb{E}[X_0^p] < \infty.$$

Weiters gilt

$$\frac{1}{n}\int_0^n f(T_s\xi)ds \;=\; \sum_{k=0}^{n} \int_0^1 f(T_s T^k \xi)ds, \qquad n\in\mathbb{N} \tag{3.6}$$

mit dem diskreten Shift $T = T_1$, und die Konvergenz folgt mit dem Ergodensatz von Birkhoff angewandt auf die Funktion $g(x) = \int_0^1 f(T_s x)ds$.

Um zu zeigen, dass der Grenzwert (3.5) entspricht, sei $f \in L^1$ und wir führen die invariante Version

$$\overline{f}(\xi) \;:=\; \lim_{r\to\infty} \limsup_{n\to\infty} \frac{1}{n}\int_r^{r+n} f(T_s\xi)ds$$

ein, welche auch $\mathcal{I}_\xi$-messbar ist. Wegen der Stationarität von $T_s\xi$ gilt $\mathbb{E}^{\mathcal{I}_\xi} f(T_s\xi) = \mathbb{E}^{\mathcal{I}_\xi} f(\xi)$ $f.s.$ für alle $s \geq 0$. Mit der L^1-Konvergenz und dem Satz von Fubini folgt

$$\mathbb{E}^{\mathcal{I}_\xi} f(\xi) \;=\; \mathbb{E}^{\mathcal{I}_\xi}\frac{1}{t}\int_0^t f(T_s\xi)ds \longrightarrow \mathbb{E}^{\mathcal{I}_\xi}\overline{f}(\xi) \;=\; \overline{f}(\xi).$$

Für $f \in L^p$ verfahre man wie in 3.7. ∎

4. Multivariate Ergodensätze und Zufallsmaße

Proposition 4.1. (*Momenten-Ungleichung*):
Vor.: Sei $\xi = \xi_k$ eine stationäre Folge von Zufallsgrößen, $S_n = \sum_{k=1}^{n}$, $M = \sup_n \frac{S_n}{n}$ und $p > 1$.
Beh.:

(i) $\mathbb{E}[|M|^p] \leq c\mathbb{E}[|\xi_1|^p]]$

(ii) $\mathbb{E}[|M| \log_+^m |M|] \leq d(1 + \mathbb{E}[|\xi_1| \log_+^{m+1} |\xi_1|]$

für Konstanten $c, d > 0$ und $\log_+(x) = \log(\max(1, x))$

Der Beweis benötigt eine Verallgemeinerung von Lemma 2.4.

Lemma 4.2. (*Maximal-Ungleichung*):
Vor.: Sei $\xi = (\xi_k)_{k \in \mathbb{N}}$ stationär in L^1.
Beh.: Es gilt

$$r\mathbb{P}[\{\sup_n (S_n/n) > 2r\}] \;\leq\; \mathbb{E}[\xi_1 \mathbb{1}_{[\xi_1 > r]}], \qquad r > 0. \tag{4.1}$$

Beweis. Sei $r > 0$ fest und setze $\xi_k^r = \xi_k \mathbb{1}_{[\xi_k > r]}$. ξ sei auf dem Folgenraum R^∞ definiert und wir bekommen mit $\xi_k \leq \xi_k^r + r$ und $A_n = S_n/n$

$$A_n - 2r \;=\; A_n \circ (\xi - 2r) \;\leq\; A_n \circ (\xi^r - r),$$

was impliziert, dass $M - 2r \leq M \circ (\xi^r - r)$. 10.4 angewandt auf die Folge $\xi^r - r$ liefert

$$\begin{aligned} r\mathbb{P}[\{M > 2r\}] \;&\leq\; r\mathbb{P}[\{M \circ (\xi^r - r) > 0\}] \\ &\leq\; \mathbb{E}[\xi_1^r \mathbb{1}_{[M \circ (\xi^r - r) > 0]}] \\ &\leq\; \mathbb{E}[\xi_1^r] \;=\; \mathbb{E}[\xi_1 \mathbb{1}_{[\xi_1 > r]}]. \end{aligned}$$

■

Beweis (*Proposition 4.1*): (i) Wir nehmen an, dass $\xi_1 \geq 0$ *f.s.* Mit Lemma 4.2, dem Satz von Fubini sowie einiger Rechnung folgt

$$\begin{aligned}
\mathbb{E}[M^p] &= p\mathbb{E}[\int_0^M r^{p-1}dr] = p\int_0^\infty P[\{M > r\}]r^{p-1}dr \\
&\leq 2p\int_0^\infty \mathbb{E}[\xi_1 \mathbb{1}_{[2\xi_1 > r]}]r^{p-2}dr = 2p\mathbb{E}[\xi_1 \int_0^{2\xi_1} r^{p-2}dr] \\
&= \frac{2p}{p-1}\mathbb{E}[\xi_1(2\xi_1)^{p-1}] = \frac{2^p p}{p-1}\mathbb{E}[\xi_1^p].
\end{aligned}$$

(ii) Für $m = 0$ gilt

$$\begin{aligned}
\mathbb{E}[M] - 1 &\leq \mathbb{E}[(M-1)^+] = \int_1^\infty \mathbb{P}[\{M > r\}]dr \\
&\leq 2\int_1^\infty \mathbb{E}[\xi_1 \mathbb{1}_{[2\xi_1 > r]}]r^{-1}dr \\
&= 2\mathbb{E}[\xi_1 \int_1^{2\xi \vee 1} r^{-1}dr] = 2\mathbb{E}[\xi_1 \log_+ 2\xi_1],
\end{aligned}$$

und für $m > 0$

$$\begin{aligned}
\mathbb{E}[M \log_+^m(M)] &= \int_0^\infty \mathbb{P}[\{M \log_+^m(M) > r\}]dr \\
&= \int_1^\infty \mathbb{P}[\{M > t\}](m \log^{m-1} t + \log^m t)dt \\
&\leq 2\int_1^\infty \mathbb{E}[\xi_1 \mathbb{1}_{[2\xi_1 > t]}](m \log^{m-1} t + \log^m t)dt \\
&= 2\mathbb{E}[\xi_1 \int_0^{\log_+ 2\xi_1} (mx^{m-1} + x^m)dx] \\
&= 2\mathbb{E}[\xi_1(\log_+^m 2\xi_1 + \frac{\log_+^{m+1} 2\xi_1}{m+1})] \\
&\leq 2e + 4\mathbb{E}[\xi_1 \log_+^{m+1} 2\xi_1 \mathbb{1}[2\xi_1 > e]] \\
&\leq d(1 + \mathbb{E}[\log_+^{m+1} \xi_1]).
\end{aligned}$$

■

Im Folgenden sei $L \log^m L(\mu)$ die Menge aller messbaren Funktionen f mit $\int |f| \log_+^m |f| d\mu < \infty$. Da $|f| \log_+^{d-1} |f| \leq |f|^d$, gilt der nächste Satz insbesondere für $f \in L^d(\mu)$.

Theorem 4.3. (*Multivariater Ergodensatz*):

Vor.: Sei ξ eine Zufallsgröße in S mit Verteilung μ, $T_1, \ldots, T_d$ maßerhaltende Transformationen auf S mit zugehörigen invarianten σ-Algebren $\mathcal{I}_1, \ldots, \mathcal{I}_d$ und setze $\mathcal{J}_k = \xi^{-1}\mathcal{I}_k$.

Beh.: Dann gilt für jedes $f \in L \log^{d-1} L(\mu)$

$$\lim_{n_1,\ldots,n_d \to \infty} \frac{1}{n_1, \ldots, n_d} \sum_{k_1=1}^{n_1} \cdots \sum_{k_d=1}^{n_d} f(T_1^{k_1} \ldots T_d^{k_d} \xi) = \mathbb{E}^{\mathcal{J}_d} \ldots \mathbb{E}^{\mathcal{J}_1} f(\xi) \qquad f.s. \quad (4.2)$$

Beweis. Mittels Induktion. Für d=1 folgt die Aussage aus Satz 3.5. (4.2) kann in der Form $\frac{1}{n} \sum_{k=1}^{n} f_m \circ T_{d+1}^k$ geschrieben werden, mit $m = (n_1, \ldots, n_d)$, und $f_m \to \mathbb{E}^{\mathcal{J}_d} \ldots \mathbb{E}^{\mathcal{J}_1} \circ f$. Mit iterierter Anwendung der Momenten-Ungleichung 4.1 folgt $\sup_m \|f_m\| \in L^1$, also nach Korollar 3.7

$$\lim_{m,n \to \infty} \frac{1}{n} \sum_{k=1}^{n} f_m \circ T_{d+1}^k = \mathbb{E}^{\mathcal{J}_{d+1}} \mathbb{E}^{\mathcal{J}_d} \ldots \mathbb{E}^{\mathcal{J}_1} \circ f \tag{4.3}$$

■

Korollar 4.4.

Vor.: Sei $\mathcal{J} = \bigcap_{k=1}^{d} \mathcal{J}_k$ und $L^1(\xi)$ die Menge aller integrierbaren, $\sigma(\xi)$ - messbaren Zufallsgrößen.

Beh.: Falls die Transformationen aus Theorem 2.10 kommutieren, dann gilt

$$\mathbb{E}^{\mathcal{J}_1} \ldots \mathbb{E}^{\mathcal{J}_d} = \mathbb{E}^{\mathcal{J}} \quad \text{auf } L^1(\xi) \tag{4.4}$$

Beweis. Da damit auch $T_1^{k_1}, \ldots, T_d^{k_d}$ für beliebigen $k_1, \ldots, k_d \in \mathbb{Z}^+$ kommutieren, folgt aus dem multivariaten Ergodensatz

$$\mathbb{E}^{\mathcal{J}_1} \cdots \mathbb{E}^{\mathcal{J}_d} f(\xi) \;=\; \mathbb{E}^{\mathcal{J}_{p_1}} \cdots \mathbb{E}^{\mathcal{J}_{p_d}} f(\xi), \qquad f.s. \tag{4.5}$$

für jede messbare Funktion $f \geq 0$ auf S und Permutation $p_1, \ldots, p_d$ von $1, \ldots, d$. Der Ausdruck in (4.5) ist $\mathcal{J}_k$-messbar für jedes k, daher auch $\mathcal{J}$-messbar. Es gilt

$$\mathbb{E}[(\mathbb{E}^{\mathcal{J}_1} \cdots \mathbb{E}^{\mathcal{J}_d} f(\xi)) \mathbb{1}_A] = \mathbb{E}[f(\xi) \mathbb{1}_A], \qquad A \in \mathcal{J},$$

woraus die Behauptung folgt. ■

Theorem 4.5. (*Monotoner, multivariater Ergodensatz*):

Vor.: Sei ξ eine Zufallsgröße in S mit Verteilung μ, $(T_s)_{s\in\mathbb{R}^d}$ eine Familie an messbaren, maßerhaltenden Transformationen mit $T_{s+t} = T_s T_t$, $\mathcal{I}$ die zugehörige σ-Algebra der invarianten Mengen und $B_1 \subset B_2 \subset \dots$ eine Folge an beschränkten, konvexen Mengen in $\mathcal{B}^d$ mit $r(B_n) \to \infty$.

Beh.: Dann gilt für jede messbare Funktion $f \geq 0$ auf S

$$\lim_{n\to\infty} \frac{1}{|B_n|} \int\limits_{B_n} f(T_s\xi)ds = \mathbb{E}[f(\xi)|\mathcal{I}_\xi] \qquad f.s. \tag{4.6}$$

Die Konvergenz gilt auch in L^p, falls $f \in L^p$.

Der Beweis benötigt einige Lemmata. Es bezeichne $\partial_\epsilon B$ die ϵ-Umgebung um den Rand von B.

Lemma 4.6.

Vor.: Sei $B \subset \mathbb{R}^d$ konvex, $\epsilon > 0$.

Beh.:

(i) $|B - B| \leq \binom{2d}{d}|B|$

(ii) $|\partial_\epsilon B| \leq 2((1 + \frac{\epsilon}{r(B)})^d - 1)|B|$

Lemma 4.7.

Vor.: Seien $B_1 \subset B_2 \subset \dots \subset B_m \in \mathcal{B}^d$ beschränkt und konvex mit $|B_1| > 0$, $K \subset \mathcal{B}^d$ beschränkt, $p : K \to \{1, \dots, n\}$.

Beh.: Dann existiert eine endliche Menge $H \subset K$ sodass die Mengen $B_{p(x)} + x$, $x \in H$ disjunkt sind und $|K| \leq \binom{2d}{d} \sum\limits_{x\in H} |B_{p(x)}|$.

Beweis. Setze $C_x = B_{p(x)} + x$ und bestimme $x_1, x_2, \dots$ rekursiv wiefolgt: Für gewählte $x_1, \dots, x_{j-1}$ sei $x_j \in K$ mit größtmöglichen $p(x)$, sodass $C_{x_i} \cap C_{x_j} = \emptyset$ für $i < j$. Die Konstruktion bricht ab, falls kein solches x_j mehr existiert.
Fixiere ein $y \in K$. Nach der Konstruktion von H gilt $C_x \cap C_y \neq \emptyset$ für ein $x \in H$ mit $p(x) \geq p(y)$, weshalb

$$y \in B_{p(x)} - B_{p(y)} + x \subset B_{p(x)} - B_{p(x)} + x. \tag{4.7}$$

Damit ist $K \subset \bigcup\limits_{x\in H} (B_{p(x)} - B_{p(x)} + x)$, und nach Lemma 2.13

$$|K| \leq \sum_{x\in H} |B_{p(x)} - B_{p(x)}| \leq \binom{2d}{d} \sum_{x\in H} |B_{p(x)}| \tag{4.8}$$

■

Außerdem brauchen wir eine multivariate Version von Lemma 4.2. Die Mengenfunktion $\eta B = \int_B f(T_s\xi)ds$ in Theorem 4.5 ist ein *stationäres Zufallsmaß* auf $\mathbb{R}^d$, und die *Intensität* m von η, definiert über die Relation $\mathbb{E}[\eta] = m\lambda_d$, entspricht $\mathbb{E}[f(\xi)]$.

Definition 4.8.
$\xi : \Omega \times \mathbb{R}^d \to \mathbb{R}$ heißt *Zufallsmaß auf* $\mathbb{R}^d$, falls $\xi(\omega, B)$ für fixes $\omega \in \Omega$ ein lokal-endliches Maß in B und eine Zufallsgröße in ω für jedes beschränkte $B \in \mathcal{B}^d$ ist.

Sei $\mathcal{M}(\mathbb{R}^d)$ die Menge der lokal-endlichen Maße μ auf $\mathbb{R}^d$. Dann kann ξ als Zufallsgröße $\Omega \to \mathcal{M}(\mathbb{R}^d)$ gesehen werden, $\mathcal{M}(\mathbb{R}^d)$ versehen mit der σ-Algebra erzeugt von den Abbildungen $\mu \mapsto \mu(B)$, $B \in \mathcal{B}$.

Definition 4.9.
Ein Zufallsmaß ξ heißt *stationär*, falls $\theta_s\xi = \xi$, wobei die Shift-Operatoren θ_s auf $\mathbb{R}^d$ definiert sind als $(\theta_s\mu)(B) = \mu(B + s)$.

$\overline{\xi} = \mathbb{E}[\xi\mathbb{1}_B|\mathcal{I}_\xi]/|B|$, $B \in \mathcal{B}$ heißt *sample intensity* von ξ und ist im Falle von Stationarität unabhängig von der Wahl von B. ($\mathcal{I}$ bezeichnet dabei die Menge aller Shift-invarianten, messbaren Mengen in $\mathcal{M}(\mathbb{R}^d)$)

Lemma 4.10.
Vor.: Sei ξ ein stationäres Zufallsmaß auf $\mathbb{R}^d$ mit Intensität m, $B_1 \subseteq B_2 \subseteq \ldots$ beschränkte , konvexe Mengen in $\mathcal{B}^d$ mit $|B_1| > 0$.
Beh.: Dann ist

$$r\mathbb{P}[\{\sup_k(\xi B_k/|B_k|) > r\}] \ \leq \ m\binom{2d}{d}, \qquad r > 0. \tag{4.9}$$

Beweis. Seien $a, r > 0$ sowie $n \in \mathbb{N}$ fest, definiere einen Prozess ν auf $\mathbb{R}^d$ sowie eine Zufallsmenge K in $S_a = \{x \in \mathbb{R}^d \mid \|x\| \leq a\}$ mit

$$\begin{aligned} \nu(x) &= \inf\{k \in \mathbb{N} \mid \xi(B_k + x) > r|B_k|\}, \qquad x \in \mathbb{R}^d, \\ K &= \{x \in S_a \mid \nu(x) \leq n\}. \end{aligned}$$

Nach Lemma 4.7 existiert eine endliche Menge $H \subseteq K$ sodass $B_{\nu(x)} + x$, $x \in H$, disjunkt sind und $|K| \leq \binom{2d}{d} \sum_{x\in H} |B_{\nu(x)}|$. Mit $b = \sup_{x\in B_n} \|x\|$ gilt

$$\xi S_{a+b} \ \geq \ \sum_{x\in H} \xi(B_{\nu(x)} + x) \ \geq \ r\sum_{x\in H} |B_{\nu(x)}| \ \geq \ r|K|\binom{2d}{d}^{-1}.$$

Durch Erwartungswertbildung, dem Satz von Fubini sowie der Stationarität von ν erhält man

$$\begin{aligned} m\binom{2d}{d}|S_{a+b}| \ &\geq \ r\mathbb{E}[|K|] \ = \ r\int_{S_a} \mathbb{P}[\{\nu(x) \leq n\}]dx \\ &= \ r|S_a|\mathbb{P}[\{\max_{k\leq n}(\xi B_k/|B_k|) > r\}]. \end{aligned}$$

Die Behauptung folgt, indem man durch $|S_a|$ dividiert, $a \to \infty$ und dann $n \to \infty$ schickt. ■

Für das letzte Lemma seien einige Begriffe aus der Funktionalanalysis erwähnt.
Ein *Hilbertraum* H ist ein (reeller oder komplexer) Vektorraum mit einem Skalarprodukt, der vollständig bezüglich der vom Skalarprodukt induzierten Norm ist. Eine *Kontraktion* auf H ist definiert als ein linearer Operator T sodass $\|T\xi\| \le \|\xi\|$ für alle $\xi \in H$. Für einen linearen Unterraum $M \subseteq H$ bezeichne $M^\perp$ den Orthogonalraum und $\overline{M}$ den Abschluss von M. Der *adjungierte Operator* T^* von T ist definiert über die Relation $\langle \xi, T\eta \rangle = \langle T^*\xi, \eta \rangle$, mit $\xi, \eta \in H$.

Lemma 4.11. (*invarianter Unterraum*):
Vor.: Sei $\mathcal{T}$ eine Familie an Kontraktionen auf einem Hilbertraum H, N der $\mathcal{T}$-invariante Unterraum und R der von $\{\xi - T\xi \mid \xi \in H, T \in \mathcal{T}\}$ aufgespannte lineare Unterraum.
Beh.: Dann ist $N^\perp \subseteq \overline{R}$.

Beweis. Für $\xi \perp R$ gilt

$$\langle \xi - T^*\xi, \eta \rangle = \langle \xi, \eta - T\eta \rangle = 0, \qquad T \in \mathcal{T}, \eta \in H,$$

also $T^*\xi = \xi$ für jedes $T \in \mathcal{T}$. Wir haben $\langle T\xi, \xi \rangle = \langle \xi, T^*\xi \rangle = \|\xi\|^2$, und mit der Kontraktionseigenschaft von T folgt

$$\begin{aligned} 0 \;\le\; \|T\xi - \xi\|^2 \;&=\; \|T\xi\|^2 + \|\xi\|^2 - 2\langle T\xi, \xi \rangle \\ \le\; 2\,\|\xi\|^2 - 2\,\|\xi\|^2 \;&=\; 0, \end{aligned}$$

was impliziert, dass $T\xi = \xi$. Damit ist $R^\perp \subseteq N$, also $N^\perp \subseteq (R^\perp)^\perp = \overline{R}$. ■

Beweis (*Theorem 4.5*): Sei $f \in L^1$ und definiere

$$\overline{T}_s f = f \circ T_s, \qquad A_n : g \mapsto |B_n|^{-1} \int_{B_n} \overline{T}_s g \; ds.$$

Sei $\epsilon > 0$ beliebig. Nach Lemma 4.11 existiert eine messbare Zerlegung

$$f = f^\epsilon + \sum_{k \le m} (g_k^\epsilon - \overline{T}_{s_k} g_k^\epsilon) + h^\epsilon,$$

wobei $f^\epsilon \in L^2$ für alle $s \in \mathbb{R}^d$ $\overline{T}_s$-invariant ist, die Funktionen g_i^ϵ beschränkt sind und $\mathbb{E}[|h^\epsilon(\xi)] < \epsilon$. Wegen der Invarianz von f^ϵ ist $A_n f^\epsilon \equiv f^\epsilon$. Mit Lemma 4.6 (ii) folgt für festes $k \le m$

$$\begin{aligned} \left\| A_n(g_k^\epsilon - \overline{T}_{s_k} g_k^\epsilon) \right\| \;&\le\; (|(B_n + s_k) \triangle B_n| / |B_n|)\, \|g_k^\epsilon\| \\ &\le\; 2((1 + \|s_k\|\, r(B_n)^{-1})^d - 1)\, \|g_k^\epsilon\| \xrightarrow{n \to \infty} 0. \end{aligned}$$

Lemma 4.10 liefert

$$r\mathbb{P}[\sup_n A_n \|h^\epsilon(\xi)\| \geq r] \;\leq\; \binom{2d}{d}\mathbb{E}[h^\epsilon(\xi)] \;\leq\; \binom{2d}{d}\epsilon, \qquad r>0,$$

was impliziert, dass $\sup_n A_n \|h^\epsilon(\xi)\| \xrightarrow{P} 0$ für $\epsilon \to 0$. Es folgt $\liminf_n A_n f(\xi) < \infty$ *f.s.*, und damit

$$\begin{aligned}(\limsup_n - \liminf_n) A_n f(\xi) \;&=\; (\limsup_n - \liminf_n) A_n h^\epsilon(\xi) \\ &\leq\; 2\sup_n A_n \|h^\epsilon(\xi)\| \xrightarrow{P} 0.\end{aligned}$$

Also geht die linke Seite *f.s.* gegen 0 und die Konvergenz folgt. Ist $f \in L^p$, so folgt die L^p-Konvergenz aus der gleichmäßigen Integrierbarkeit der Familie $|A_n f(\xi)|^p, n \in \mathbb{N}$. Um zu zeigen, dass der Grenzwert (4.6) entspricht, verfahre man wie im Beweis von Korollar 3.8 und erweitere die Konvergenz für beliebige $f \geq 0$ wie im Beweis von Satz 3.5. ■

Korollar 4.12.

Vor.: Sei ξ ein stationäres Zufallsmaß auf $\mathbb{R}^d$ und $B_1 \subset B_2 \subset \ldots$ eine Folge an konvexen, beschränken Mengen in $\mathcal{B}^d$ mit $r(B_n) \to \infty$.

Beh.: Dann gilt $\xi\mathbb{1}_{B_n}/|B| \to \overline{\xi}$ f.s., wobei $\overline{\xi}\lambda^d = \mathbb{E}[\xi|\mathcal{I}_\xi]$. Falls $\xi\mathbb{1}_{[0,1]^d} \in L^p$, gilt die Konvergenz auch in L^p.

Das nächste Korollar zeigt, dass die L^p-Konvergenz in 4.5 und 4.12 auch unter schwächeren Bedingungen gilt.
Wir nennen Verteilungen μ_n *asymptotisch invariant*, falls $\|\mu_n - \mu_n * \delta_s\| \xrightarrow{n} 0$ für jedes $s \in \mathbb{R}^d$, wobei $\|\cdot\|$ die Totalvariations-Norm bezeichnet. Eine Folge f_n an Dichtefunktionen heißt *asymptotisch invariant*, falls $\lambda_d|f_n - \theta_s f_n| \xrightarrow{n} 0$ für jedes $s \in \mathbb{R}^d$. Es ist zu beachten, dass sich die Aussage in 4.5 auch als $\mu_n X \to \overline{X}$ schreiben lässt, mit $\mu_n = (\mathbb{1}_{B_n} \cdot \lambda_d)/|B_n|$, $X_s = f(T_s\xi)$ und $\overline{X} = \mathbb{E}[f(\xi)|\mathcal{I}_\xi]$.

Korollar 4.13. (*Mittelwert-Ergodensatz*):

Vor.: Sei $p \geq 1$, X ein stationärer, messbarer L^p-wertigen Prozess auf $\mathbb{R}^d$, μ_n asymptotisch invariante Verteilungen, ξ ein stationäres Zufallsmaß auf $\mathbb{R}^d$ mit endlicher *sample intensity* und f_n asymptotisch invariante Dichtefunktionen

Beh.:

(i) $\mu_n X \to \mathbb{E}[X|\mathcal{I}_X]$ in L^p

(ii) $\xi f_n \to \overline{\xi}$ in L^p

Eine Folge $(c_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ heißt *subadditiv*, falls $c_{n+m} \leq c_n + c_m \;\forall\; m,n \in \mathbb{N}$.

Lemma 4.14. (*Subadditivität*):

Vor.: Sei $c_1, c_2, \dots \in \mathbb{R}$ eine subadditive Folge.

Beh.: Dann ist

$$\lim_{n\to\infty} \frac{c_n}{n} = \inf_n \frac{c_n}{n} \in [-\infty, \infty).$$

Setzt man $\xi_{j,k} = \eta_{j+1} \dots + \eta_k$ für eine stationäre, integrierbare Folge an Zufallsvariablen $(\eta_k)_{k\in\mathbb{N}}$, so kovergiert $\xi_{0,n}/n$ f.s. in L^1 nach dem Ergodensatz von Birkhoff . Ein ähnliches Ergebnis erhält man für subadditive, doppelt indizierte Folgen $(\xi_{j,k})$ die stationär bezüglich Shifts in beiden Indizes sind, also $(\xi_{j+1,k+1}) \overset{\mathrm{d}}{=} (\xi_{j,k})$. Die schwächere Forderung

$$(\xi_{k,2k}, \xi_{2k,3k}, \dots) \overset{\mathrm{d}}{=} (\xi_{0,k}, \xi_{k,2k}, \dots), \qquad k \in \mathbb{N}, \tag{4.10}$$

$$(\xi_{k,k+1}, \xi_{k+1,k+2}, \dots) \overset{\mathrm{d}}{=} (\xi_{0,1}, \xi_{1,2}, \dots), \qquad k \in \mathbb{N}, \tag{4.11}$$

reicht hierfür aber aus.

Wir nenne eine doppelt indizierte Folge $\xi_{j,k}$ *subadditiv*, falls

$$\xi_{0,n} \le \xi_{0,m} + \xi_{m,n}, \qquad 0 < m < n.$$

Theorem 4.15. (*Subadditiver Ergodensatz*):

Vor.: Sei $(\xi_{j,k})$ eine subadditive Folge an Zufallsvariablen, welche (2.12) und (2.13) erfüllen, mit $\mathbb{E}[\xi_{0,1}^+] < \infty$.

Beh.: Dann konvergiert $\xi_{0,n}/n$ f.s. gegen eine Zufallsvariable $\bar{\xi}$ in $[-\infty, \infty)$, mit $\mathbb{E}[\bar{\xi}] = \inf_n \mathbb{E}[\xi_{0,n}/n] = c$. Falls $c > -\infty$, so gilt die Konvergenz auch in L^1. Ist die Folge in (2.12) ergodisch, so ist $\bar{\xi}$ f.s. konstant.

Theorem 4.16. (*Zufallsmatrizen*):

Vor.: Sei $X^k = (X_{ij}^k)$ eine stationäre Folge an $d \times d$ Zufallsmatrizen, sodass $X_{ij} > 0$ f.s. und $\mathbb{E}[|\log(X_{ij})|] < \infty$, $\forall i, j$.

Beh.: Dann konvergiert $\frac{1}{n}\log((X^1 \dots X^n)_{ij})$ $\forall i, j$ in L^1 mit $n \to \infty$ und der Grenzwert ist unabhängig von (i, j).

Für Beweise von 4.12 bis 4.16 sei auf [1, S. 190-194] verwiesen.

5. Zerlegung in ergodische Komponenten

Mit den Voraussetzungen von Satz 2.5, wobei S Borel sei, gilt mit $\eta = \mathbb{P}[\xi \in \cdot|\mathcal{I}_\xi]$

$$\mathcal{L}(\xi) = \mathbb{E}[\mathbb{P}[\xi \in \cdot|\mathcal{I}_\xi]] = \mathbb{E}[\eta]. \tag{5.1}$$

Außerdem ist

$$\eta(I) = \mathbb{P}[\xi \in I|\mathcal{I}_\xi] = \mathbb{1}_{\{\xi \in I\}}, \quad I \in \mathcal{I}.$$

Theorem 5.1.
Vor.: Sei ξ eine Zufallsgröße in einem Borel-Raum S mit Verteilung μ und $\mathcal{T} = \{T_s, s \in \mathbb{R}^d\}$ eine messbare Gruppe von maßerhaltenden Transformationen auf S bezüglich μ, mit invarianter σ-Algebra $\mathcal{I}$.
Beh.: Dann ist $\eta = \mathbb{P}[\xi \in \cdot|\mathcal{I}_\xi]$ f.s. invariant und ergodisch unter $\mathcal{T}$.

Für den Beweis fixieren wir eine aufsteigende Folge an konvexen Mengen $B_n \in \mathcal{B}^d$ mit $r(B_n) \to \infty$ und führen auf S die *Wahrscheinlichkeitskerne*

$$\mu_n(x, A) := \frac{1}{|B_n|} \int_{B_n} \mathbb{1}_A(T_s x) ds, \qquad x \in S, A \in \mathcal{S} \tag{5.2}$$

sowie die damit verbundenen *empirischen Verteilungen* $\eta_n := \mu_n(\xi, \cdot)$ ein. Nach Theorem 2.12 gilt $\eta_n \circ f \to \eta \circ f$ für jede beschränkte, messbare Funktion f auf S, wobei $\eta = \mathbb{P}[\xi \in \cdot|\mathcal{I}_\xi]$. Im Folgenden nennen wir eine Klasse $\mathcal{C} \subset \mathcal{S}$ *maßbestimend*, falls jedes Wahrscheinlichkeitsmaß auf S durch seine Werte auf $\mathcal{C}$ eindeutig bestimmt ist.

Lemma 5.2. (*degenerierter Grenzwert*):
Vor.: Sei $A_1, A_2, \ldots \in \mathcal{S}$ eine Folge an maßbestimmenden Mengen, sodass $\eta_n A_k \to \mathbb{P}[\xi \in A_k]$ *f.s.* für jedes $k \in \mathbb{N}$.
Beh.: Dann ist ξ ergodisch.

Beweis. Nach Theorem 2.12 und mit der Voraussetzung gilt $\eta_n A_k \to \eta A_k = P[\xi \in A_k]$ für alle $k \in \mathbb{N}$. Da jedes A_k maßbestimmend ist, folgt $\eta = \mathcal{L}(\xi)$ *f.s.* Für jedes $I \in \mathcal{I}$ gilt daher

$$\mu(I) = \mathbb{P}[\xi \in I] = \eta I = \mathbb{P}[\xi \in I|\mathcal{I}_\xi] = \mathbb{1}_I(\xi) \in \{0, 1\},$$

also ist ξ ergodisch. ■

Beweis (*Theorem 2.22*):
Aufgrund der Stationarität von ξ gilt für jedes $A \in \mathcal{S}$ und $s \in \mathbb{R}^d$

$$(\eta \circ T_s^{-1})A = \mathbb{P}[T_s\xi \in A|\mathcal{I}_\xi] = \mathbb{P}[\xi \in A|\mathcal{I}_\xi] = \eta A \quad f.s. \tag{5.3}$$

Da $\mathcal{S}$ Borel ist, folgt $\eta \circ T_s^{-1} = \eta$ $f.s$ für jedes s. Setze $C = [0,1]^d$ und $\overline{\eta} = \int_C (\eta \circ T_s^{-1}) ds$. Da η invariant unter Shifts in $\mathbb{Z}^d$ ist, ist $\overline{\eta}$ invariant unter beliebigen Shifts. Es folgt

$$\lambda^d(\{s \in C : \eta \circ T_s^{-1} = \eta\}) = 1 \quad f.s.$$

und damit $\overline{\eta} = \eta$ $f.s.$ Also ist η f.s. $\mathcal{T}$-invariant. Sei nun $A_1, A_2, \dots$ eine Folge an maßbestimmenden Mengen in S (welche existiert, da S Borel ist). Mit $\eta_n A_k \to \eta A_k$ für beliebiges k gilt

$$\eta \bigcap_k \{x \in S : \mu_n(x, A_k) \to \eta A_k\} = \mathbb{P}[\bigcap_k \{\eta_n A_k \to \eta A_k\}|\mathcal{I}_\xi] = 1 \quad f.s.$$

η ist ein f.s. $\mathcal{T}$-invariantes Wahrscheinlichkeitsmaß auf S, daher lässt sich Lemma 2.23 für jedes $\omega \in \Omega$ außerhalb einer $\mathbb{P}$-Nullmenge anwenden. Es folgt, dass η ergodisch ist. ■

Das eben bewiesene Theorem zeigt also, dass mit (2.14) eine Zerlegung der Verteilung μ einer Zufallsgröße ξ in ergodische Komponenten vorliegt. Das nächste Ergebnis wird sein, dass diese Zerlegung eindeutig ist und die ergodischen Maße als Extrempunkte in der konvexen Menge an invarianten Maßen charakterisiert sind.
Zur Wiederholung: ein linearer Raum M heißt konvex, falls $tx+(1-t)y \in M$ für $x, y \in M$, $t \in (0,1)$. Wir sagen, $m \in M$ ist *extrem*, falls für beliebige $m_1, m_2 \in M$ und $t \in (0,1)$ aus der Beziehung $m = tm_1 + (1-t)m_2$ folgt, dass $m_1 = m_2 = m$. Zu jeder Menge M an Maßen auf $(S, \mathcal{S})$ bezeichne $\mathcal{M}$ die σ-Algebra erzeugt durch die Auswertungs-Abbildungen $\pi_B : \mu \to \mu(B)$, $\mu \in M$, $B \in \mathcal{S}$.

Theorem 5.3.
Vor.: Sei $\mathcal{T} = \{T_s : s \in \mathbb{R}^d\}$ eine messbare Gruppe an Transformationen auf einem Borel-Raum S.
Beh.: Dann bilden die $\mathcal{T}$-invarianten Verteilungen auf S eine konvexe Menge M, dessen Extrempunkte den ergodischen Maßen in M entsprechen. Außerdem besitzt jedes Maß $\mu \in M$ eine eindeutige Darstellung $\mu = \int m \; \nu(dm)$, mit ν eingeschränkt auf die Menge der ergodischen Maße.

Beweis. M ist offensichtlich konvex, und nach Theorem 2.22 besitzt jedes $\mu \in M$ eine Darstellung $\mu = \int m \; \nu(dm)$, wobei ν ein Wahrscheinlichkeitsmaß auf der Menge der ergodischen Maße in M ist. Um die Eindeutigkeit von ν zu zeigen, sei $\eta = \mu[\cdot|\mathcal{I}]$ eine reguläre, bedingte Wahrscheinlichkeit auf S. Es gilt $\mu_n A \to \eta A$ (μ_n aus (2.15)) nach Theorem 2.12, und für jede Folge $A_1, A_2, \dots$ in $\mathcal{S}$ haben wir

$$m \bigcap_k \{x \in S : \mu_n(x, A_k) \to \eta(x, A_k)\} = 1 \quad \nu - f.s.$$

Dieselbe Beziehung gilt für $\eta(x, A_k)$ ersetzt durch mA_k, da ν auf die ergodischen Maße in M beschränkt ist. Sind die A_k maßbestimmend, so folgt $m\{x : \eta(x, \cdot) = m\} = 1\ \nu - f.s.$, und für jedes $A \in \mathcal{M}$ gilt

$$\mu\{\eta \in A\} = \int m\{\eta \in A\}\nu(dm) = \int \mathbb{1}_A(m)\nu(dm) = \nu A.$$

Also ist $\nu = \mu \circ \eta^{-1}$.
Um die Äquivalenz von Ergodizität und Extremität zu zeigen, sei $\mu \in M$ mit $\mu = \int m\ \nu(dm)$. Angenommen μ ist extrem, aber nicht ergodisch. Dann ist ν nichtdegeneriert, und es existieren $\nu_1 \perp \nu_2$ sowie $c \in (0,1)$ mit $\nu = c\nu_1 + (1-c)\nu_2$. Da μ extrem ist, gilt $\int m\ \nu_1(dm) = \int m\ \nu_2(dm)$, und wegen der Eindeutigkeit der Darstellung $\nu_1 = \nu_2$. Also ist μ ergodisch.
Sei nun μ als ergodisch, aber nicht extrem angenommen. Es gilt $\nu = \delta_\mu$ und es existieren $c \in (0,1), \mu_1, \mu_2$ mit $\mu = c\mu_1 + (1-c)\mu_2$. Wenn $\mu_i = \int m\ \nu_i(dm)$ für $i = 1, 2$, dann ist $\delta_\mu = c\nu_1 + (1-c)\nu_2$ aufgrund der Eindeutigkeit der Darstellung. Das ist nur möglich, wenn $\nu_1 = \nu_2 = \nu$, und damit $\mu_1 = \mu_2 = \mu$. Daher ist μ extrem. ■

Literaturverzeichnis

[1] O. Kallenberg. *Foundations of Modern Probability (Second Edition)*. Springer, 2002.

[2] N. Kusolitsch. *Maß- und Wahrscheinlichkeitstheorie*. Springer Spektrum, 2014.