3 Stationäre Prozesse (Version November 2017)

3.1 Maßerhaltende Transformationen

In diesem Kapitel führen wir zunächst den Begriff der *maßerhaltenden Transformation* auf einem Wahrscheinlichkeitsraum ein und definieren, wann eine solche Transformation *ergodisch* bzw. *mischend* heißt. Parallel dazu entwickeln wir den Begriff eines *stationären Prozesses* und erarbeiten die Beziehung zwischen diesen Begriffen. Höhepunkt des Kapitels ist der *Ergodensatz*, Satz 3.12. Im ganzen Kapitel ist $(\Omega, \mathcal{F}, \mathbb{P})$ ein fester Wahrscheinlichkeitsraum.

Definition 3.1. Sei $(E, \mathcal{E})$ ein Messraum. Ein $(E, \mathcal{E})$-wertiger Prozess $\mathbb{X} = (X_n)_{n\in\mathbb{N}}$ heißt *stationär*, wenn gilt

$$\mathcal{L}((X_n)_{n\in\mathbb{N}}) = \mathcal{L}((X_{n+1})_{n\in\mathbb{N}}).$$

Bemerkung: Induktiv folgt dann für jedes $k \in \mathbb{N}$

$$\mathcal{L}((X_n)_{n\in\mathbb{N}}) = \mathcal{L}((X_{n+k})_{n\in\mathbb{N}}).$$

Bemerkung: Ein $(E, \mathcal{E})$-wertiger Prozess $\mathbb{X}$ ist stationär, genau dann wenn für jedes $k \in \mathbb{N}$

$$\mathcal{L}((X_1,\, X_2,\, ..., X_k)) = \mathcal{L}((X_2,\, X_3,\, ..., X_{k+1}))$$

gilt.

Bemerkung: Statt $\mathbb{N}$ wird oft auch $\mathbb{N}_0$ als Indexmenge eines stationären Prozesses verwendet.

Warnung: Man beachte, dass es in obiger Definition keineswegs genügt, dass $\mathcal{L}(X_n) = \mathcal{L}(X_{n+1})$ für alle $n \in \mathbb{N}$ gilt: Sind zum Beispiel $Y_1, Y_2, ...$ unabhängig $\mathcal{N}(0,1)$-verteilt und $X_1 = X_2 = Y_2$ und $X_k := Y_k$ für $k \geq 3$, dann gilt $\mathcal{L}(X_n) = \mathcal{L}(X_{n+1}) = \mathcal{N}(0,1)$ für alle $n \in \mathbb{N}$, aber der Prozess $(X_n)_{n\in\mathbb{N}}$ ist *nicht* stationär, da zum Beispiel $\mathcal{L}(X_1, X_2) \neq \mathcal{L}(X_2, X_3)$!

Beispiel: Jede Folge von u.i.v. Zufallsgrößen (mit Werten in einem beliebigen Messraum $(E, \mathcal{E})$) ist ein stationärer Prozess.

Definition 3.2. Eine Abbildung $T : \Omega \to \Omega$ heißt *maßerhaltende Transformation* auf $(\Omega, \mathcal{F}, \mathbb{P})$, wenn T $\mathcal{F}$-$\mathcal{F}$-messbar ist und $\mathbb{P}T^{-1} = \mathbb{P}$ gilt.

Bemerkung: Die messbare Abbildung $T : \Omega \to \Omega$ ist maßerhaltend genau dann, wenn $\mathbb{P}T^{-1}(A) = \mathbb{P}(A)$ für alle A aus einem $\cap$-stabilen Erzeuger von $\mathcal{F}$ gilt.

Lemma 3.3. *Sei T eine maßerhaltende Transformation auf $(\Omega, \mathcal{F}, \mathbb{P})$ und Y eine $(E, \mathcal{E})$-wertige Zufallsgröße auf $(\Omega, \mathcal{F}, \mathbb{P})$. Dann ist $X_1 := Y$, $X_2 := X_1 \circ T$, $X_3 := X_2 \circ T$, ... stationär.*

Beweis. Für $n \in \mathbb{N}$ und $A_1, ..., A_n \in \mathcal{E}$ gilt

$$\begin{aligned}\mathbb{P}\{X_1 \in A_1, ... X_n \in A_n\} &= \mathbb{P}T^{-1}\{X_1 \in A_1, ... X_n \in A_n\}\\ &= \mathbb{P}\{X_1 \circ T \in A_1, ..., X_n \circ T \in A_n\} = \mathbb{P}\{X_2 \in A_1, ..., X_{n+1} \in A_n\}.\end{aligned}$$

Da $\{A_1 \times ... \times A_n \times E \times ...; n \in \mathbb{N}, A_1, ..., A_n \in \mathcal{E}\}$ ein $\cap$-stabiler Erzeuger von $\mathcal{E}^{\mathbb{N}}$ ist, folgt die Behauptung. $\square$

Es stellt sich die Frage nach einer Umkehrung von Lemma 3.3. Nun kann man sehr leicht sehen, dass es zu einem gegebenen stationären Prozess $\mathbb{X} = (X_n)_{n\in\mathbb{N}}$ auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ im allgemeinen keine maßerhaltende Transformation T auf $(\Omega, \mathcal{F}, \mathbb{P})$ und Zufallsgröße Y gibt, so dass

$$X_1 = Y,\ X_2 := X_1 \circ T,\ X_3 := X_2 \circ T,\ ...$$

gilt (siehe das Beispiel weiter unten). Es gilt aber folgendes:

$\mathbb{X} : (\Omega, \mathcal{F}, \mathbb{P}) \to (E^{\mathbb{N}}, \mathcal{E}^{\mathbb{N}})$ ist messbar. Sei $\hat{\mathbb{P}} = \mathbb{P}\mathbb{X}^{-1}$ die Verteilung von $\mathbb{X}$ und $T : E^{\mathbb{N}} \to E^{\mathbb{N}}$ der ***Linksshift*** definiert als $T((x_1, x_2, ...)) = (x_2, x_3, ...)$. Dann ist T maßerhaltend auf $(E^{\mathbb{N}}, \mathcal{E}^{\mathbb{N}}, \hat{\mathbb{P}})$ (wie man sehr leicht überprüft). Daher ist der *kanonische Prozess* $\hat{\mathbb{X}}$ auf $(E^{\mathbb{N}}, \mathcal{E}^{\mathbb{N}}, \hat{\mathbb{P}})$, definiert durch $\hat{X}_n(x_1, x_2, ...) := x_n$ stationär. Es gilt $\mathcal{L}(\hat{\mathbb{X}}) = \mathcal{L}(\mathbb{X}) = \hat{\mathbb{P}}$ und für $Y(x_1, x_2, ...) := x_1$ gilt

$$\hat{X}_1 = Y,\ \hat{X}_2 = Y \circ T,\ \hat{X}_3 = Y \circ T^2,\ ...$$

Beispiel: Sei $\Omega = \{a, b, c\},\ \mathcal{F} = \mathcal{P}(\Omega)$ und $\mathbb{P}(\{a\}) = 1/2,\ \mathbb{P}(\{b\}) = \mathbb{P}(\{c\}) = 1/4$. Weiter gelte: $X_n(a) = 1, X_n(b) = X_n(c) = 0$, falls n gerade ist und $X_n(a) = 0, X_n(b) = X_n(c) = 1$, falls n ungerade ist, also

$$\begin{aligned} \mathbb{X}(a) &= (0, 1, 0, 1, 0, 1, ...) \\ \mathbb{X}(b) = \mathbb{X}(c) &= (1, 0, 1, 0, 1, 0...). \end{aligned}$$

$\mathbb{X}$ ist stationär. Es existiert aber keine maßerhaltende Transformation T auf $(\Omega, \mathcal{F}, \mathbb{P})$, so dass $X_2 = X_1 \circ T$. Angenommen doch, dann gilt

$$1 = X_2(a) = X_1(T(a)),$$

dh. $T(a) \in \{b, c\}$. Angenommen es gilt $T(a) = b$, dann gilt $1/4 = \mathbb{P}(\{b\}) = \mathbb{P}T^{-1}(\{b\}) \geq \mathbb{P}(\{a\}) \geq 1/2$, was nicht möglich ist. Dasselbe gilt im Fall $T(a) = c$.

Definition 3.4. Sei T eine maßerhaltende Transformation auf $(\Omega, \mathcal{F}, \mathbb{P})$.

a) $A \in \mathcal{F}$ heißt T-*invariant*, wenn $T^{-1}(A) = A$ gilt.

b) T heißt *ergodisch*, wenn für jede T-invariante Menge A gilt: $\mathbb{P}(A) \in \{0, 1\}$.

Bezeichnung und Bemerkung: Ist T eine maßerhaltende Transformation auf $(\Omega, \mathcal{F}, \mathbb{P})$, so wird die Familie aller T-invarianten Mengen mit $\mathcal{J}$ bezeichnet. Man überprüft unschwer, dass $\mathcal{J}$ eine Teil-σ-Algebra von $\mathcal{F}$ ist. Man nennt $\mathcal{J}$ die *invariante σ-Algebra* (bezüglich T).

Definition 3.5. Eine maßerhaltende Transformation T auf $(\Omega, \mathcal{F}, \mathbb{P})$ heißt *mischend*, wenn für alle $A, B \in \mathcal{F}$ gilt:

$$\lim_{n\to\infty} \mathbb{P}(A \cap T^{-n}B) = \mathbb{P}(A)\,\mathbb{P}(B).$$

Proposition 3.6. *Jede mischende Transformation ist ergodisch.*

Beweis. Ist T mischend und $A \in \mathcal{F}$ T-invariant, so gilt

$$\mathbb{P}(A) = \mathbb{P}(A \cap A) = \mathbb{P}(A \cap T^{-n}A) \to \mathbb{P}(A) \cdot \mathbb{P}(A),$$

das heißt $\mathbb{P}(A) = (\mathbb{P}(A))^2$, also $\mathbb{P}(A) \in \{0, 1\}$. Somit ist T ergodisch. □

Lemma 3.7. *Sei T eine maßerhaltende Transformation auf $(\Omega, \mathcal{F}, \mathbb{P})$ und $\mathcal{D}$ ein $\cap$-stabiler Erzeuger von $\mathcal{F}$. Wenn*

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

für alle $A, B \in \mathcal{D}$ gilt, dann ist T mischend.

Beweis.

1. Schritt: Für $B \in \mathcal{D}$ sei

$$\mathcal{D}_B := \{A \in \mathcal{F} : \ \mathbb{P}(A \cap T^{-n}B) \to \mathbb{P}(A)\,\mathbb{P}(B)\}.$$

Dann überprüft man sehr leicht, dass $\mathcal{D}_B$ ein Dynkinsystem ist. Für $B \in \mathcal{D}$ gilt also $\mathcal{D} \subseteq \mathcal{D}_B \subseteq \mathcal{F}$ und mit dem Hauptsatz über Dynkinsysteme folgt $\mathcal{D}_B = \mathcal{F}$.

2. Schritt: Für $A \in \mathcal{F}$ sei

$$\mathcal{E}_A := \{B \in \mathcal{F} : \ \mathbb{P}(A \cap T^{-n}B) \to \mathbb{P}(A)\,\mathbb{P}(B)\}.$$

Dann überprüft man ebenso leicht, dass $\mathcal{E}_A$ ein Dynkinsystem ist. Nach dem ersten Schritt folgt $\mathcal{E}_A \supseteq \mathcal{D}$ und daher mit dem Hauptsatz über Dynkinsysteme $\mathcal{E}_A = \mathcal{F}$. □

3.2 Beispiele

Beispiel 3.8. Sei $D := \{z \in \mathbb{C} : |z| = 1\}$ der Rand des komplexen Einheitskreises, $\mathcal{B}_D$ die Borel-σ- Algebra auf D und λ_D das normierte Lebesguemaß auf $(D, \mathcal{B}_D)$. Weiter sei für $c \in D$ die Abbildung $T_c : D \to D$ definiert durch $T_c(\omega) := c\,\omega$. Dann ist T_c maßerhaltend und es gilt:

a) T_c ist ergodisch genau dann wenn c keine Einheitswurzel ist.

b) T_c ist für kein c mischend.

Beweis. Es ist klar, dass T_c für jedes $c \in D$ maßerhaltend ist.
a) Sei $c \in D$ eine Einheitswurzel. Dann existiert ein $n \in \mathbb{N}$ mit $c^n = 1$. Sei $A \in \mathcal{B}_D$ eine beliebige Menge mit der Eigenschaft $0 < \lambda_D(A) < \frac{1}{n}$ und $B := A \cup T_c^{-1}(A) \cup ... \cup T_c^{-(n-1)}(A)$.

Dann ist $B \in \mathcal{B}_D$, $T_c^{-1}(B) = B$ (da $T_c^{-n}(A) = A$) und $0 < \lambda_D(B) < 1$. Also ist $\mathcal{J}$ nicht trivial und damit T_c nicht ergodisch.

Sei nun c keine Einheitswurzel. Für die folgende Überlegung ist es angenehmer, D mit $[0, 1)$ zu identifizieren und T_c mit der Abbildung $T_c(x) = x + c \bmod 1$ (wobei nun $c \in [0, 1)$ irrational ist). Statt λ_D schreiben wir nur λ für das Lebesguemaß auf $[0, 1)$. Sei also $A \in \mathcal{B}_{[0,1)}$ invariant unter T_c. Da $\{\, \mathrm{e}^{2\pi ikt}, k \in \mathbb{Z}\}$ eine ONB des komplexen $L^2([0, 1), \lambda)$ ist und $\mathbf{1}_A \in L^2([0, 1), \lambda)$, existiert (in L^2) eine Darstellung der Form

$$\mathbf{1}_A(t) = \sum_{k \in \mathbb{Z}} c_k \, \mathrm{e}^{2\pi ikt}, \qquad \text{mit } \sum_k |c_k|^2 < \infty.$$

Da A invariant ist, folgt

$$\mathbf{1}_A(t) = \mathbf{1}_A(T_c(t)) = \sum_{k \in \mathbb{Z}} c_k \, \mathrm{e}^{2\pi ik(t+c)}.$$

Da die Darstellung durch eine ONB eindeutig ist, folgt für jedes $k \in \mathbb{Z}$

$$c_k = c_k \, \mathrm{e}^{2\pi ikc},$$

also entweder $c_k = 0$ oder $kc \in \mathbb{Z}$. Da c nach Voraussetzung irrational ist, kann kc allenfalls für $k = 0$ ganzzahlig sein und daher gilt $c_k = 0$ für alle $k \neq 0$. Es folgt also, dass $\mathbf{1}_A$ fast sicher gleich der Konstanten c_0 ist (welche notwendigerweise 0 oder 1 ist). Also folgt $\lambda(A) \in \{0, 1\}$ und damit die Trivialität von $\mathcal{J}$, d.h. T_c ist ergodisch.

b) Es genügt, die Behauptung in dem Fall zu zeigen, in dem c keine Einheitswurzel ist. Dazu beachte man zuerst, dass dann die Menge $\mathcal{C} := \{c^n | n \in \mathbb{N}_0\}$ dicht in D liegt, denn die Folge hat (da D kompakt ist) mindestens einen Häufungspunkt ω_0 in D. Sei $\varepsilon > 0$ und $m > n \geq 0$ so dass $|c^n - \omega_0| < \frac{\varepsilon}{2}$ und $|c^m - \omega_0| < \frac{\varepsilon}{2}$. Dann gilt mit der Dreiecksungleichung $0 < |c^{m-n} - 1| < \varepsilon$ und damit existiert für jedes $\omega \in D$ ein $k \in \mathbb{N}_0$, so dass $|\omega - c^{(m-n)k}| < \varepsilon$, d.h. $\mathcal{C}$ ist dicht in D.

Um zu zeigen, dass im Fall, in dem c keine Einheitswurzel ist, T_c nicht mischend ist, wählen wir die Mengen $A = B = \{\, \mathrm{e}^{2\pi ix} : 0 \leq x \leq \frac{1}{4}\}$. Da $\mathcal{C}$ dicht in D ist, existiert eine Folge $n_k \to \infty$ so dass $A \cap T^{-n_k}(A) = \emptyset$ und daher $\mathbb{P}(A \cap T^{-n_k}(A)) = 0 \neq \frac{1}{16} = \mathbb{P}(A)^2$. $\square$

Beispiel 3.9. Sei $(E, \mathcal{E}, \mu)$ ein Wahrscheinlichkeitsraum. Dann ist der Linksshift $T : (E^{\mathbb{N}}, \mathcal{E}^{\mathbb{N}}, \mu^{\otimes\mathbb{N}}) \to (E^{\mathbb{N}}, \mathcal{E}^{\mathbb{N}}, \mu^{\otimes\mathbb{N}})$ maßerhaltend und mischend.

Beweis. Es ist klar, dass T maßerhaltend ist. Sei $A = A_1 \times ... \times A_k \times E \times ...$ und $B = B_1 \times ... \times B_m \times E \times ...$ mit $A_1, ..., B_m \in \mathcal{E}$. Dann gilt für $n \geq k$

$$\mu^{\otimes\mathbb{N}}(A \cap T^{-n}(B)) = \mu(A_1)...\mu(A_k)\mu(B_1)...\mu(B_m) = \mu^{\otimes\mathbb{N}}(A)\mu^{\otimes\mathbb{N}}(B).$$

Da Mengen A und B der obigen Form einen $\cap$-stabilen Erzeuger von $\mathcal{E}^{\mathbb{N}}$ bilden, folgt aus Lemma 3.7, dass T mischend ist. $\square$

3.3 Birkhoffscher Ergodensatz

Nun wollen wir den Birkhoffschen Ergodensatz formulieren und beweisen. Dafür bedürfen wir des folgenden Lemmas. Im ganzen Abschnitt sei T eine maßerhaltende Transformation auf $(\Omega, \mathcal{F}, \mathbb{P})$ und $\mathcal{J}$ die invariante σ-Algebra.

Lemma 3.10. *Ist* $Y \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbb{P})$ *und* $\mathbb{E}(Y|\mathcal{J}) < 0 \quad \mathbb{P}$*-f.s., so gilt:*

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} Y \circ T^j \leq 0 \quad \mathbb{P}\text{-f.s.}$$

Beweis. Erklärt man

$$M_n := \max\{0, Y, Y + Y\circ T, Y + Y \circ T + Y \circ T^2, \ldots, \sum_{j=0}^{n-1} Y \circ T^j\},$$

so gilt: $0 \leq M_n(\omega) \leq M_{n+1}(\omega)$. Insbesondere existiert also $M(\omega) := \lim_{n\to\infty} M_n(\omega) \leq \infty$. Wir zeigen, dass $M < \infty$ fast sicher gilt. Nun ist

$$Y + M_n \circ T = \max\{Y, Y + Y\circ T, Y + Y \circ T + Y \circ T^2, \ldots, \sum_{j=0}^{n} Y \circ T^j\}$$

und $(Y + M_n \circ T)^+ = \max\{0, Y + M_n \circ T\} = M_{n+1}$, bzw. für $n \to \infty : (Y + M \circ T)^+ = M$. Wegen $|Y(\omega)| < \infty$, folgt hieraus:

$$M(\omega) = \infty \iff (M \circ T)(\omega) = M(T(\omega)) = \infty,$$

d.h. für $A := \{M = \infty\}$ gilt:

$$T^{-1}(A) = \{\omega | T(\omega) \in A\} = \{M \circ T = \infty\} = \{M = \infty\} = A \in \mathcal{J}.$$

Allgemein gilt für reelle Zahlen $a, b : (a + b)^+ = b - \min(b, -a)$. Angewandt auf M_n ergibt dies:

$$M_{n+1} = (Y + M_n \circ T)^+ = M_n \circ T - \min\{M_n \circ T, -Y\} \geq M_n,$$

d.h.

$$\mathbb{E}(\mathbf{1}_A \;\; M_n \circ T) - \mathbb{E}(\mathbf{1}_A \;\; \min\{M_n \circ T, -Y\}) \geq \mathbb{E}(\mathbf{1}_A \;\; M_n).$$

Weil A invariant und T maßerhaltend ist, gilt:

$$\mathbb{E}(\mathbf{1}_A \;\; M_n \circ T) = \mathbb{E}(\mathbf{1}_{T^{-1}(A)} \;\; M_n \circ T) = \mathbb{E}(\mathbf{1}_A \circ T \;\; M_n \circ T) = \mathbb{E}(\mathbf{1}_A \;\; M_n),$$

bzw. – da der letzte Ausdruck endlich ist – $\mathbb{E}(\mathbf{1}_A \;\; \min(M_n \circ T, -Y)) \leq 0$. Hieraus folgt schließlich mit dem Satz von der beschränkten Konvergenz:

$$\mathbb{E}(\underbrace{\mathbf{1}_A \;\; \min\{M \circ T, -Y\}}_{=\mathbf{1}_A \, (-Y)}) \leq 0.$$

Mithin gilt $\mathbb{E}(\mathbf{1}_A\ Y) \geq 0$, woraus sich wegen $\mathbb{E}(Y|\mathcal{J}) < 0 \quad \mathbb{P}$ - f.s. sofort $\mathbb{P}(A) = 0$ ergibt, d.h. $M < \infty \quad \mathbb{P}$ -f.s. Aus

$$\frac{1}{n}\sum_{j=0}^{n-1} Y \circ T^j \leq \frac{1}{n} M_n \leq \frac{1}{n} M$$

folgt dann die Behauptung. □

Lemma 3.11. *Ist Z eine reelle $\mathcal{J}$ messbare Zufallsgröße, so gilt: $Z \circ T = Z$.*

Beweis. Für jede reelle Zahl a ist $Z^{-1}(\{a\}) \in \mathcal{J}$, d.h. $Z^{-1}(\{a\}) = T^{-1}(Z^{-1}(\{a\})) = (Z \circ T)^{-1}(\{a\})$, also gilt: $Z(\omega) = a \iff (Z \circ T)(\omega) = a$, bzw. $Z = Z \circ T$. □

Der Birkhoffsche Ergodensatz lautet nun wie folgt.

Satz 3.12. *Falls $X \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbb{P})$ ist, gilt:*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} X \circ T^j = \mathbb{E}(X|\mathcal{J}) \quad \mathbb{P}\text{-}f.s..$$

Beweis. Für beliebiges $\varepsilon > 0$ und $Y := X - \mathbb{E}(X|\mathcal{J}) - \varepsilon$ gilt $Y \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbb{P})$ und $\mathbb{E}(Y|\mathcal{J}) = -\varepsilon < 0$. Nach Lemma 3.11 ist $\mathbb{E}(X|\mathcal{J}) \circ T^j = \mathbb{E}(X|\mathcal{J})$ für alle j, also:

$$\frac{1}{n}\sum_{j=0}^{n-1} Y \circ T^j = \frac{1}{n}\sum_{j=0}^{n-1} X \circ T^j - \mathbb{E}(X|\mathcal{J}) - \varepsilon$$

bzw.

$$\frac{1}{n}\sum_{j=0}^{n-1} X \circ T^j = \mathbb{E}(X|\mathcal{J}) + \varepsilon + \frac{1}{n}\sum_{j=0}^{n-1} Y \circ T^j.$$

Mit Lemma 3.10 folgt hieraus:

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} X \circ T^j \leq \mathbb{E}(X|\mathcal{J}) + \varepsilon.$$

Die gleiche Überlegung mit $-X \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbb{P})$ liefert:

$$\limsup_{n\to\infty} \Big(-\frac{1}{n}\sum_{j=0}^{n-1} X \circ T^j \Big) \leq -\mathbb{E}(X|\mathcal{J}) + \varepsilon$$

Insgesamt folgt also:

$$-\varepsilon + \mathbb{E}(X|\mathcal{J}) \leq \liminf_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} X \circ T^j \leq \limsup_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} X \circ T^j \leq \mathbb{E}(X|\mathcal{J}) + \varepsilon,$$

also

$$\lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} X \circ T^j = \mathbb{E}(X|\mathcal{J}) \quad \mathbb{P}\text{-f.s..}$$

□

Bemerkung 3.13. Aus Satz 1.50 und Bemerkung 1.51 folgt, dass die Konvergenz im Birkhoffschen Ergodensatz auch in L^1 gilt.

Nun formulieren wir als Folgerung aus dem Birkhoffschen Ergodensatz den Ergodensatz für stationäre Prozesse.

Satz 3.14. *Sei* $\mathbb{X} = (X_1, X_2, ...)$ *ein reellwertiger, stationärer Prozess mit* $\mathbb{E}|X_1| < \infty$ *definiert auf einem Raum* $(\Omega, \mathcal{F}, \mathbb{P})$*. Dann gilt*

$$\frac{1}{n}\sum_{j=1}^{n} X_j \to \mathbb{E}(X_1|\mathbb{X}^{-1}(\mathcal{J}))\, \textit{f.s.},$$

wobei $\mathcal{J}$ *die invariante* σ*-Algebra des Linksshifts* T *auf* $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}^{\mathbb{N}}, \hat{\mathbb{P}})$ *ist, wobei* $\hat{\mathbb{P}} = \mathbb{P}\mathbb{X}^{-1}$ *ist.*

Beweis. Nach Satz 3.12 gilt für $\pi_1 : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$, $\pi_1(x) := x_1$

$$\lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} \pi_1 \circ T^j(x) = \mathbb{E}(\pi_1|\mathcal{J})(x) \qquad \hat{\mathbb{P}}\text{-f.s.},$$

da $\mathbb{E}_{\hat{\mathbb{P}}}|\pi_1(x)| = \mathbb{E}|X_1| < \infty$. Ersetzt man x durch $\mathbb{X}(\omega)$, so folgt

$$\lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} X_{j+1}(\omega) = \mathbb{E}(\pi_1|\mathcal{J})(\mathbb{X}(\omega)) \qquad \mathbb{P}\text{-f.s..}$$

Wir müssen also noch zeigen, dass $\mathbb{E}(\pi_1|\mathcal{J})(\mathbb{X}(\omega)) = \mathbb{E}(X_1|\mathbb{X}^{-1}(\mathcal{J}))(\omega)$ $\mathbb{P}$-fast sicher gilt, also

(i) $\mathbb{E}(\pi_1|\mathcal{J})(\mathbb{X}(\omega))$ ist $\mathbb{X}^{-1}(\mathcal{J})$-messbar und

(ii) $\int_A \mathbb{E}(\pi_1|\mathcal{J})(\mathbb{X}(\omega))\,\mathrm{d}\mathbb{P}(\omega) = \int_A X_1\,\mathrm{d}\mathbb{P}$ für alle $A \in \mathbb{X}^{-1}(\mathcal{J})$.

Zu (i): Dies folgt, da Kompositionen messbarer Abbildungen messbar sind.
Zu (ii): Stelle $A \in \mathbb{X}^{-1}(\mathcal{J})$ in der Form $A = \mathbb{X}^{-1}(\tilde{A})$ mit $\tilde{A} \in \mathcal{J}$ dar. Dann gilt

$$\begin{aligned}\int_{\mathbb{X}^{-1}(\tilde{A})} \mathbb{E}(\pi_1|\mathcal{J})(\mathbb{X}(\omega))\,\mathrm{d}\mathbb{P}(\omega) &= \int_{\tilde{A}} \mathbb{E}(\pi_1|\mathcal{J})(x)\,\mathrm{d}\hat{\mathbb{P}}(x)\\ &= \int_{\tilde{A}} \pi_1(x)\,\mathrm{d}\hat{\mathbb{P}}(x) = \int_{\mathbb{X}^{-1}(\tilde{A})} \pi_1(\mathbb{X}(\omega))\,\mathrm{d}\mathbb{P}(\omega) = \int_{\mathbb{X}^{-1}(\tilde{A})} X_1(\omega)\,\mathrm{d}\mathbb{P}(\omega),\end{aligned}$$

womit die Behauptung gezeigt ist. □

Als Korollar erhalten wir das starke Gesetz der großen Zahlen für u.i.v. Zufallsgrößen.

Korollar 3.15. *Seien* $X_1, X_2, ...$ *u.i.v. in* $L^1(\Omega, \mathcal{F}, \mathbb{P})$*. Dann folgt* $\lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} X_j = \mathbb{E}X_1$ *f.s.*

Beweis. Sei $\hat{\mathbb{P}}$ die Verteilung von $\mathbb{X} = (X_1, X_2, ...)$, d.h. $\hat{\mathbb{P}} = \mathcal{L}(X_1)^{\otimes\mathbb{N}}$. Nach Beispiel 3.9 ist der Linksshift $T : (\mathbb{R}^{\mathbb{N}}, \mathcal{B}^{\mathbb{N}}, \hat{\mathbb{P}}) \to (\mathbb{R}^{\mathbb{N}}, \mathcal{B}^{\mathbb{N}}, \hat{\mathbb{P}})$ ergodisch. Daher ist die zugehörige invariante σ-Algebra $\mathcal{J}$ trivial, also auch $\mathbb{X}^{-1}(\mathcal{J})$, d.h. $\mathbb{E}(X_1|\mathbb{X}^{-1}(\mathcal{J})) = \mathbb{E}X_1$ f.s. und die Behauptung folgt aus Satz 3.14. □

3.4 Anwendungen des Ergodensatzes

Eine Anwendung des Ergodensatzes ist die folgende Proposition.

Proposition 3.16. *Eine maßerhaltende Transformation T auf $(\Omega, \mathcal{F}, \mathbb{P})$ ist ergodisch genau dann, wenn für alle A, $B \in \mathcal{F}$ gilt*

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

Beweis. (nach [1], Seite 426)

"$\Rightarrow$" Sei T ergodisch und $A, B \in \mathcal{F}$. Dann folgt aus dem Ergodensatz

$$\frac{1}{n} \sum_{k=0}^{n-1} \mathbf{1}_{A\cap T^{-k}(B)}(\omega) = \frac{\mathbf{1}_A(\omega)}{n} \sum_{k=0}^{n-1} \mathbf{1}_B(T^k\omega) \to \mathbf{1}_A(\omega)\mathbb{E}\mathbf{1}_B \text{ f.s.},$$

und mit beschränkter Konvergenz folgt die Konvergenz der Erwartungswerte, d.h. die Behauptung.

"$\Leftarrow$" Sei $A \in \mathcal{J}$ (und $B := A$). Dann gilt

$$\mathbb{P}(A) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{P}(A \cap T^{-k}(A)) = \mathbb{P}(A)^2,$$

also folgt $\mathbb{P}(A) \in \{0, 1\}$. Also ist T ergodisch.

□

3.5 Markovketten

Sei $X_0, X_1, X_2, ...$ eine irreduzible, stationäre Markovkette definiert auf $(\Omega, \mathcal{F}, \mathbb{P})$ mit Werten in einer abzählbaren Menge E. Auf E wählen wir die Potenzmenge $\mathcal{E}$ als σ-Algebra. Weiter sei $\hat{\mathbb{P}}$ die Verteilung der (stationären) Markovkette auf dem Raum $(E^{\mathbb{N}_0}, \mathcal{E}^{\mathbb{N}_0})$ und $T : E^{\mathbb{N}_0} \to E^{\mathbb{N}_0}$ der Linksshift. Dann ist T maßerhaltend. Weiter gilt der folgende Satz.

Satz 3.17. *Es gilt*

a) T ist ergodisch.

b) T ist mischend genau dann wenn die Markovkette aperiodisch ist.

Beweis. a) Sei $A \in E^{\mathbb{N}_0}$ invariant, d.h. $T^{-1}(A) = A$. Zu zeigen ist $\hat{\mathbb{P}}(A) \in \{0, 1\}$. Sei $\pi_n : E^{\mathbb{N}_0} \to E$ die Projektion auf die n-te Komponente und $\mathcal{F}_n := \sigma(\pi_m, m \leq n) \subseteq \mathcal{F}_\infty = \mathcal{F} = \mathcal{E}^{\mathbb{N}_0}$, $n \in \mathbb{N}_0$. Für $x \in E$ sei

$$h(x) := \mathbb{E}_x(\mathbf{1}_A)(= \mathbb{E}(\mathbf{1}_A|\pi_0 = x) = \hat{\mathbb{P}}(A|\pi_0 = x)),$$

wobei $\mathbb{E}$ der Erwartungswert bezüglich $\hat{\mathbb{P}}$ ist.

Nun ist $\mathbb{E}(\mathbf{1}_A|\mathcal{F}_n), n \in \mathbb{N}_0$ ein $\mathbb{F}$-Martingal, welches regulär ist und daher für $n \to \infty$ fast sicher und in L^1 gegen $\mathbf{1}_A$ konvergiert.

Weiter gilt für $\omega \in E^{\mathbb{N}_0}$ und $n \in \mathbb{N}_0$

$$\begin{aligned}\mathbb{E}(\mathbf{1}_A|\mathcal{F}_n)(\omega) &= \mathbb{E}(\mathbf{1}_A \circ T^n|\mathcal{F}_n)(\omega) = \mathbb{E}(\mathbf{1}_A \circ T^n|\pi_n)(\omega) \\ &= \hat{\mathbb{P}}(T^{-n}(A)|\pi_n = \omega_n) = \hat{\mathbb{P}}((\pi_n, \pi_{n+1} \ldots) \in A|\pi_n = \omega_n), \\ &= \hat{\mathbb{P}}(A|\pi_0 = \omega_n) = h(\pi_n(\omega)),\end{aligned}$$

wobei wir beim zweiten und vorletzten Gleichheitszeichen die Markoveigenschaft benutzten.

Da $(\pi_n)_{n\in\mathbb{N}_0}$ rekurrent ist und $h(\pi_n)$ fast sicher gegen $\mathbf{1}_A$ konvergiert, folgt entweder $h \equiv 0$ oder $h \equiv 1$, also $\hat{\mathbb{P}}(A) \in \{0, 1\}$.

b) Wir nehmen zuerst an, dass T aperiodisch ist und bezeichnen die Übergangswahrscheinlichkeiten der Markovkette mit $p_{i,j}$, $i, j \in E$ und die Verteilung von X_0 mit μ (μ ist die eindeutige invariante Verteilung der Markovkette). Weiter sei

$$\mathcal{D} := \Big\{\big\{x \in E^{\mathbb{N}_0} : x_0 = v_0, ..., x_m = v_m\big\} : m \in \mathbb{N}_0,\ v_0, ..., v_m \in E\Big\} \cup \emptyset.$$

$\mathcal{D}$ ist ein $\cap$-stabiler Erzeuger von $\mathcal{E}^{\mathbb{N}_0}$. Nun sei $A := \{x \in E^{\mathbb{N}_0} : x_0 = v_0, ..., x_m = v_m\}$ und $B := \{x \in E^{\mathbb{N}_0} : x_0 = w_0, ..., x_k = w_k\}$, wobei $v_0, ..., v_m, w_0, ..., w_k \in E$. Dann gilt für $n \geq m$ mit der Markoveigenschaft

$$\begin{aligned}\hat{\mathbb{P}}(A \cap T^{-n}(B)) &= \mathbb{P}(\{X_0 = v_0, ..., X_m = v_m\} \cap \{X_n = w_0, ..., X_{n+k} = w_k\}) \\ &= \mathbb{P}(\{X_n = w_0, ..., X_{n+k} = w_k\}|X_m = v_m)\mathbb{P}(X_0 = v_0, ..., X_m = v_m).\end{aligned}$$

Der zweite Faktor ist gleich $\mathbb{P}(A)$ und der erste ist gleich $p^{(n-m)}_{v_m,w_0} p_{w_0,w_1} \ldots p_{w_{k-1},w_k}$. Da die Kette aperiodisch ist, konvergiert $p^{(n-m)}_{v_m,w_0}$ nach dem Hauptsatz für Markovketten für $n \to \infty$ gegen μ_{w_0} und damit konvergiert $\mathbb{P}(\{X_n = w_0, ..., X_{n+k} = w_k\}|X_m = v_m)$ gegen $\mathbb{P}(B)$. Die Behauptung folgt nun aus Lemma 3.7.

Wir müssen noch die Umkehrung zeigen, d.h. zeigen, dass eine irreduzible und stationäre Markovkette, die periodisch ist, nicht mischend sein kann. Da dies sehr einfach und gleichzeitig eine gute Übungsaufgabe ist (deren Lösung in [1], Seite 428 steht), verzichten wir auf den Beweis. $\square$

3.6 Ergänzungen

Oft ist der folgende Begriff nützlich.

Definition 3.18. Eine maßerhaltende Transformation T auf $(\Omega, \mathcal{F}, \mathbb{P})$ heißt *schwach mischend*, wenn für alle $A, B \in \mathcal{F}$ gilt:

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

Offensichtlich ist jede mischende Transformation schwach mischend. Aus Proposition 3.16 folgt sofort, dass jede schwach mischende Transformation ergodisch ist. Das folgende Beispiel zeigt, dass aus ergodisch nicht schwach mischend folgt. Es gibt Beispiele von schwach mischenden Transformationen, die nicht mischend sind.

Beispiel 3.19. Sei $\Omega = \{0, 1\}$ mit der Potenzmenge als σ-Algebra und der Gleichverteilung $\mathbb{P}$ als Wahrscheinlichkeitsmaß. Die Abbildung T definiert durch $T(0) = 1$, $T(1) = 0$ ist offensichtlich maßerhaltend und ergodisch aber nicht schwach mischend, wie man sieht, wenn man $A = \{0\}$ und $B = \{1\}$ wählt.

Ein Hauptgrund für die Einführung des Begriffes *schwach mischend* ist die folgende Aussage, deren Beweis sich in [2] findet.

Satz 3.20. *Sei T eine maßerhaltende Transformation auf $(\Omega, \mathcal{F}, \mathbb{P})$ und $T \times T : (\Omega, \mathcal{F}, \mathbb{P}) \otimes (\Omega, \mathcal{F}, \mathbb{P})$ definiert durch $T \times T(\omega_1, \omega_2) := (T(\omega_1), T(\omega_2))$. Dann sind äquivalent*

i) T ist schwach mischend

ii) $T \times T$ ist ergodisch

iii) $T \times T$ ist schwach mischend.

Man beachte, dass daher aus der Ergodizität von T *nicht* die Ergodizität von $T \times T$ folgt. Sie gilt sogar *nie*, wenn T nur ergodisch aber nicht schwach mischend ist. Man überzeuge sich davon, dass in obigem Beispiel $T \times T$ in der Tat nicht ergodisch ist!

References

[1] Klenke, A. (2006). *Wahrscheinlichkeitstheorie.* Springer, Berlin.

[2] Walters, P. (1975). *Ergodic Theory - Introductory Lectures.* Springer LNM 458, Berlin.