Schwach ergodische Prozesse

Von der Fakultät für Naturwissenschaften
der
Universität Duisburg-Essen
(Standort Duisburg)
zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften
genehmigte Dissertation
von

Thorsten Bahne
aus Oberhausen

Referent:	Prof. Dr. L. Rogge
Korreferent:	Prof. Dr. U. Herkenrath
Tag der mündlichen Prüfung:	14. Februar 2003

Inhaltsverzeichnis

§ 1

Einleitung

In der mathematischen Disziplin der Wahrscheinlichkeitstheorie existieren mehrere zentrale Probleme, deren Lösungen von verschiedensten Seiten beleuchtet werden. Eine dieser Fragestellungen ist das starke Gesetz der großen Zahlen (SLLN, „Strong law of large numbers“). Ein stochastischer Prozess $(X_n)_{n\in\mathbb{N}}$ erfüllt das starke Gesetz der großen Zahlen, wenn gilt:

$$P\left(\lim_{n\to\infty}\left(\frac{1}{n}\cdot\sum_{k=1}^{n}X_k \;-\; \frac{1}{n}\cdot\sum_{k=1}^{n}P(X_k)\right) = 0\right) = 1,$$

mit anderen Worten, ein stochastischer Prozess erfüllt das SLLN, falls die Folge $\left(\frac{1}{n}\cdot\sum_{k=1}^{n}X_k \;-\; \frac{1}{n}\cdot\sum_{k=1}^{n}P(X_k)\right)_{n\in\mathbb{N}}$ P-f.s. gegen 0 konvergiert. Ein starkes Gesetz der großen Zahlen für den fairen Münzwurf (fair coin-tossing) wurde im Jahr 1909 von Borel veröffentlicht. Eine der zur Zeit bekanntesten Varianten des SLLN ist das starke Gesetz der großen Zahlen von Kolmogorov aus dem Jahre 1930:

> Es sei $(X_n)_{n\in\mathbb{N}}$ eine unabhängige Folge reeller, integrierbarer Zufallsvariablen. Gilt dann
> $$\sum_{n=1}^{\infty}\frac{V(X_n)}{n^2} \;<\; \infty,$$
> so erfüllt $(X_n)_{n\in\mathbb{N}}$ das starke Gesetz der großen Zahlen.

Eine Variante dieses SLLN wurde 1931 von Birkhoff entdeckt [6]. Als Voraussetzungen in diesem sogenannten *Birkhoff'schen Ergodensatz* werden die Stationarität und die Ergodizität des betreffenden stochastischen Prozesses $(X_n)_{n\in\mathbb{N}}$ benötigt. Die mathematischen Aussagen im Umfeld der bedingten Wahrscheinlichkeiten ermöglichen jedoch noch schwächere Voraussetzungen für die Gültigkeit des SLLN als die Ergodizität. Da die Ergodizität über die 0-1-Eigenschaft der invarianten Mengen des entsprechenden Prozesses definiert wird, folgt auch die Unabhängigkeit der einzelnen Zufallsvariablen des Prozesses von eben diesen Mengen. Diese Unabhängigkeit reicht letztlich als Bedingung für die Gültigkeit des Ergodensatzes von Birkhoff aus.

Im Jahre 1983 wurde von Etemadi [18] eine weitere Variante des starken Gesetzes der großen Zahlen publiziert. Der Nachweis der Gültigkeit des SLLN wurde über die paarweise Unabhängigkeit und die identische Verteilung der einzelnen Zufallsvariablen (ZV) des Prozesses geführt.

Eine Kopplung dieses bedeutenden Ergebnisses der Stochastik und der Überlegungen im Umfeld der Ergodentheorie und der damit verbundenen Aussage von Birkhoff wurde im Jahre 1998 durch Landers und Rogge erreicht [27]. In ihrem Artikel, der im Jahre 2000 unter dem Titel „Weak ergodicity of stationary pairwise independent processes“ in den Proceedings der AMS publiziert wurde, entwickelten sie den Begriff eines *schwach ergodischen Prozesses*. Die schwache Ergodizität umfaßt neben der Stationarität auch die Unabhängigkeit der einzelnen ZV des Prozesses von den invarianten Mengen desselben Prozesses, also genau die oben angesprochenen hinreichenden Voraussetzungen für den Ergodensatz von Birkhoff. Die Verbindung zum Ergebnis von Etemadi kann darin gesehen werden, dass in dem angegebenen Artikel der Nachweis der schwachen Ergodizität eines paarweise unabhängigen, stationären Prozesses mit Werten in einem separablen, metrischen Raum geführt wird. Es wird weiter nachgewiesen, dass es schwach ergodische Prozesse gibt, die nicht ergodisch sind. Somit wurde gezeigt, dass die schwache Ergodizität eine echte Erweiterung der bisherigen Begrifflichkeiten ist. Diese Tatsache sagt aber noch nichts über die Möglichkeiten aus, in den mathematischen Sätzen und Folgerungen die Voraussetzung der Ergodizität durch die schwache Ergodizität zu ersetzen.

Es soll daher untersucht werden, welche Eigenschaften ergodischer stationärer Prozesse auch für schwach ergodische stationäre Prozesse gelten und welche Sätze und Folgerungen für ergodische stationäre Prozesse bei einer Verallgemeinerung der Voraussetzungen ihre Gültigkeit beibehalten. Des Weiteren sollen spezielle stochastische Prozesse betrachtet werden, bei denen die Ergodizität gleichbedeutend mit der schwachen Ergodizität ist. Zum Abschluss werden Aussagen zur Gültigkeit des zentralen Grenzwertsatzes (CLT, „Central Limit Theorem“) gemacht.

Im Rahmen der vorliegenden Arbeit werden zu Beginn die grundlegenden Definitionen und Eigenschaften schwach ergodischer Prozesse betrachtet und dargelegt. Hierzu gehört insbesondere der Ergodensatz für schwach ergodische Prozesse 3.6, der als Pendant zum SLLN angesehen werden kann, wie aus der obigen Darlegung hervorgeht.

Des Weiteren werden Betrachtungen die Klasse der schwach ergodischen Prozesse betreffend durchgeführt. In diesem Umfeld tritt die Frage nach der Abgeschlossenheit der Klasse bezüglich spezieller Konvergenzen auf, aber auch die Vererbung der schwachen Ergodizität auf Kombinationen mit messbaren Funktionen. Neben diesen sind aber auch Fragen von Interesse, wie zwei Maße zueinander liegen, falls ein stochastischer Prozess bezüglich beider schwach ergodisch ist.

In der vorliegenden Arbeit wird nachgewiesen, dass jeder stationäre Prozess bedingt identisch verteilt bezüglich seiner invarianten Mengen $\mathscr{I}(X_n, n \in \mathbb{N})$ ist, d. h. für jedes $n \in \mathbb{N}$ und jedes $A \in \mathscr{A}_1$ gilt (vgl. A.1):

$$P\left(X_n \in A \mid \mathscr{I}(X_n, n \in \mathbb{N})\right) \;=\; P\left(X_1 \in A \mid \mathscr{I}(X_n, n \in \mathbb{N})\right).$$

Angesichts dieser Tatsache und der Aussage von Satz 4.3, dem Satz von de Finetti, wonach jeder vertauschbare Prozess $(X_n)_{n\in\mathbb{N}}$ bedingt identisch verteilt und bedingt unabhängig bezüglich seiner invarianten Mengen ist, werden auch Betrachtungen im Umfeld von vertauschbaren Prozessen durchgeführt. Als besonderes Phänomen in diesem Bereich ist die Tatsache zu werten, dass im Falle der vertauschbaren Prozesse die beiden Begriffe schwach ergodisch und ergodisch zusammenfallen, d. h. ein vertauschbarer Prozess ist genau dann ergodisch, wenn er schwach ergodisch ist.

Zu den Betrachtungen der schwach ergodischen Prozesse gehören auch Informationen, welche verschiedenen Gruppen von stochastischen Prozessen zu dieser neuen Klasse gehören. Aus diesem Grund wird der Beweisgang aus dem Artikel von Landers und Rogge auf eine abstrakte Ebene gehoben und auf notwendige Voraussetzungen hin untersucht. Angesichts der abstrakten Darstellung kann die Aussage von Landes und Rogge dann auch auf andere σ-Algebren übertragen werden. Die hier gefundenen Bedingungen werden anschließend durch verschiedene Beispiele ergänzt. Hierbei wird deutlich, dass die Bedingungen von Landers und Rogge, also die Stationarität und die paarweise Unabhängigkeit des Prozesses, als Spezialfall einer umfassenderen Aussage anzusehen sind.

Betrachtungen der schwachen Ergodizität im Umfeld spezieller Klassen von stochastischen Prozessen schließen sich an die Beispiele an. Entsprechend der Phänomene bei vertauschbaren Prozessen kann auch für den Fall der Markov-Prozesse nachgewiesen werden, dass genau die ergodischen stationären Prozesse schwach ergodisch sind. Dieses liegt an der speziellen Darstellung der invarianten Mengen eines Markov-Prozesses.

Auch der Bereich des Decouplings wird im Zusammenhang mit schwach ergodischen stationären Prozessen betrachtet. Aufgrund der bedingten Unabhängigkeit eines entkoppelt tangierenden Prozesses und der Aussagen im Zusammenhang mit dem schon angesprochenen Satz von de Finetti werden Überlegungen angestellt, ob und wann es einen vertauschbaren entkoppelten tangierenden Prozess gibt und welche Eigenschaften sich von einem schwach ergodischen Prozess auf den ihm zugeordneten entkoppelt tangierenden Prozess übertragen.

Als Abschluss der vorliegenden Bearbeitung werden Martingale resp. Martingal-Differenzenfolgen betrachtet. Im Falle der Ergodizität sind Aussagen über die quadratische Integrierbarkeit gemacht worden, die sich auch auf den schwach ergodischen Fall übertragen lassen. Die interessante Aussage zur Gültigkeit des zentralen Grenzwertsatzes (CLT) von Billingsley [4] kann ebenfalls in vollem Umfang übernommen werden und verallgemeinert somit bereits bekannte Aussagen. Hier sei z. B. das Ergebnis von Dug Hun Hong [25] aus dem Jahre 1994 erwähnt, der die Gültigkeit des CLT für paarweise unabhängige, identisch verteilte und vorzeichen-invariante Zufallsvariablen nachgewiesen hat.

Insgesamt werden also verschiedene Klassen stochastischer Prozesse angesprochen. Einerseits werden Beispiele angegeben, bei denen es zu Verallgemeinerungen der ergodischen Aussagen kommt, andererseits werden Eigenschaften aufgezeigt, die weitere Voraussetzungen benötigen, um ihre Gültigkeit im schwach ergodischen Umfeld beizubehalten. Des Weiteren finden sich Klassen von stochastischen Prozessen, bei denen die schwach ergodischen Repräsentanten gerade die ergodischen sind.

§ 2

Grundlegende Definitionen und Zusammenhänge

Zu Beginn dieser Bearbeitung sollen die grundlegenden Definitionen und Sachverhalte im Umfeld von stationären Prozessen dargelegt werden. Da einzelne Begriffe in der Literatur zum Teil mit leicht variierenden Definitionen gefunden werden können, werden alle in dieser Arbeit verwendeten Begriffe zuvor definiert.

Die schon bekannten Definitionen und Zusammenhänge sind zumeist dem Buch von Bauer [3] oder dem Buch von Breiman [9] entnommen.

Allgemein soll für die gesamte Arbeit gelten, dass ein stochastischer Prozess aus Zufallsvariablen (ZV) besteht, die von einem Wahrscheinlichkeitsraum $(\Omega, \mathscr{A}, P)$ in einen Messraum $(\Omega_1, \mathscr{A}_1)$ abbilden. Für eine σ-Algebra $\mathscr{D} \subset \mathscr{A}$ sei weiter $\mathscr{D}^{\mathbb{N}}$ die entsprechende Produkt-σ-Algebra auf $\Omega^{\mathbb{N}}$, entsprechendes gilt auch für ${\mathscr{D}_1}^{\mathbb{N}}$ auf ${\Omega_1}^{\mathbb{N}}$ mit einer σ-Algebra $\mathscr{D}_1 \subset \mathscr{A}_1$.

Zunächst sollen die Begriffe der *Stationarität* und der *Vertauschbarkeit* eines stochastischen Prozesses definiert werden.

2.1 Definition

(i) Ein stochastischer Prozess $(X_n)_{n\in\mathbb{N}}$ heißt *stationär*, wenn für jedes $k \in \mathbb{N}$ die Prozesse $(X_{k+n})_{n\in\mathbb{N}}$ und $(X_n)_{n\in\mathbb{N}}$ identisch verteilt sind, d.h. wenn für jedes $B \in {\mathscr{A}_1}^{\mathbb{N}}$ gilt:

$$P\left((X_1, X_2, \dots) \in B\right) \;=\; P\left((X_{k+1}, X_{k+2}, \dots) \in B\right).$$

(ii) Ein stochastischer Prozess $(X_n)_{n\in\mathbb{N}}$ heißt *vertauschbar*, wenn für jede endliche Permutation τ von $\mathbb{N}$ die Prozesse $\left(X_{\tau(n)}\right)_{n\in\mathbb{N}}$ und $(X_n)_{n\in\mathbb{N}}$ identisch verteilt sind, d.h. wenn für jedes $B \in \mathscr{A}_1{}^{\mathbb{N}}$ gilt:

$$P\left((X_1, X_2, \dots) \in B\right) \;=\; P\Big((X_{\tau(1)}, X_{\tau(2)}, \dots) \in B\Big)\,.$$

Hierbei ist eine *endliche Permutation* τ von $\mathbb{N}$ eine bijektive Abbildung von $\mathbb{N}$ auf $\mathbb{N}$, für die ein $k_\tau \in \mathbb{N}$ von der Art existiert, dass für alle $j \in \mathbb{N}_{\geq k_\tau}$ stets $\tau(j) = j$ gilt. Die Menge aller endlichen Permutationen von $\mathbb{N}$ wird mit $\tau(\mathbb{N})$ bezeichnet.

2.2 Bemerkung

(i) Ist $(X_n)_{n\in\mathbb{N}}$ ein stationärer Prozess, so sind die ZV $X_n, n \in \mathbb{N}$, identisch verteilt.

(ii) Ist $(X_n)_{n\in\mathbb{N}}$ ein vertauschbarer Prozess, so ist $(X_n)_{n\in\mathbb{N}}$ auch ein stationärer Prozess.

Beweis:

Zu (i) Die Bemerkung ergibt sich aus Definition 2.1, wenn man die folgenden Mengen $B \in \mathscr{A}_1{}^{\mathbb{N}}$ mit $B = A \times \Omega_1{}^{\mathbb{N}}$ und $A \in \mathscr{A}_1$ betrachtet.

Zu (ii) Für den Nachweis der Stationarität des Prozesses $(X_n)_{n\in\mathbb{N}}$ genügt es nach dem Buch von Breiman [9, Chapter 6] für jedes $n \in \mathbb{N}$ und für jedes $A \in \mathscr{A}_1{}^n$ zu zeigen:

$$P\left((X_1, X_2, \dots, X_n) \in A\right) \;=\; P\left((X_2, X_3, \dots, X_n, X_{n+1}) \in A\right).$$

Seien $n \in \mathbb{N}$ und $A \in \mathscr{A}_1{}^n$ beliebig, aber fest. Weiter ist die folgende Abbildung $\tau : \mathbb{N} \to \mathbb{N}$ mit

$$\tau(k) \;:=\; \begin{cases} k+1 & \text{, falls } k \in \{1, 2, \dots, n\}, \\ 1 & \text{, falls } k = n+1, \\ k & \text{ sonst,} \end{cases}$$

eine endliche Permutation. Somit folgt aufgrund der Vertauschbarkeit des Prozesses $(X_n)_{n\in\mathbb{N}}$:

$$\begin{aligned} P\left((X_1, X_2, \dots, X_n) \in A\right) \;&=\; P\left((X_1, X_2, \dots, X_n, X_{n+1}) \in A \times \Omega_1\right) \\ &=\; P\Big((X_{\tau(1)}, X_{\tau(2)}, \dots, X_{\tau(n)}, X_{\tau(n+1)}) \in A \times \Omega_1\Big) \\ &=\; P\left((X_2, X_3, \dots, X_{n+1}, X_1) \in A \times \Omega_1\right) \\ &=\; P\left((X_2, X_3, \dots, X_{n+1}) \in A\right) \end{aligned}$$

Somit ist $(X_n)_{n\in\mathbb{N}}$ stationär. □

Für die Definition der *Ergodizität* resp. der *schwachen Ergodizität* eines stationären Prozesses wird der Begriff der *invarianten Menge* benötigt. In diesem Zusammenhang werden auch zwei verwandte Arten von Mengen definiert, die *terminalen* und die *symmetrischen* Mengen, die für eine Erweiterung der verschiedenen Arten der Ergodizität verwendet werden können.

2.3 Definition

(i) Eine Menge $A \in \mathscr{A}$ heißt *invariant* (bezüglich des Prozesses $(X_n)_{n\in\mathbb{N}}$), wenn eine der beiden äquivalenten Aussagen gilt:

a) Es existiert ein $B \in {\mathscr{A}_1}^{\mathbb{N}}$, so dass für alle $n \in \mathbb{N}$ gilt:

$$A \;=\; \{(X_n, X_{n+1}, \dots) \in B\}\,.$$

b) Es existiert eine invariante Menge $C \in {\mathscr{A}_1}^{\mathbb{N}}$, d. h. eine Menge $C \in {\mathscr{A}_1}^{\mathbb{N}}$, mit $(x_1, x_2, x_3, \dots) \in C \;\Leftrightarrow\; (x_2, x_3, \dots) \in C$, für die gilt:

$$A \;=\; \{(X_1, X_2, X_3, \dots) \in C\}\,.$$

(ii) Eine Menge $A \in \mathscr{A}$ heißt *terminal* (bezüglich des Prozesses $(X_n)_{n\in\mathbb{N}}$), wenn eine der beiden äquivalenten Aussagen gilt:

a) Es gilt:

$$A \;\in\; \bigcap_{n=1}^{\infty} \sigma\,(X_n, X_{n+1}, X_{n+2}, \cdots)\,.$$

b) Es existiert eine Menge C aus der *universellen terminalen σ-Algebra* $\mathcal{T} \subset {\mathscr{A}_1}^{\mathbb{N}}$ mit

$$A \;=\; \{(X_1, X_2, X_3, \dots) \in C\}\,,$$

wobei gilt:

$$\mathcal{T} := \bigcap_{n=1}^{\infty} \mathcal{T}_n, \text{ mit } \mathcal{T}_n := \left\{{\Omega_1}^n \times B \mid B \in {\mathscr{A}_1}^{\mathbb{N}}\right\},\ n \in \mathbb{N}.$$

(iii) Eine Menge $A \in \mathscr{A}$ heißt *symmetrisch* oder *permutierbar* (bezüglich des Prozesses $(X_n)_{n\in\mathbb{N}}$), wenn eine der äquivalenten Bedingungen gilt:

a) Es existiert ein $B \in {\mathscr{A}_1}^{\mathbb{N}}$, so dass für jede endliche Permutation τ von $\mathbb{N}$ gilt:

$$A \;=\; \left\{(X_{\tau(1)}, X_{\tau(2)}, \dots) \in B\right\}.$$

b) Es existiert eine symmetrische Menge $C \in \mathscr{A}_1^{\mathbb{N}}$, d. h. eine Menge $C \in \mathscr{A}_1^{\mathbb{N}}$, mit $(x_1, x_2, x_3, \dots) \in C \Leftrightarrow (x_{\tau(1)}, x_{\tau(2)}, \dots) \in C$ für jede endliche Permutation τ von $\mathbb{N}$, für die gilt:

$$A = \{(X_1, X_2, X_3, \dots) \in C\}.$$

Zum Beweis der Gültigkeit der jeweiligen Äquivalenzen der letzten Definition wird der sogenannte *Shift-Operator* benötigt.

2.4 Definition

Auf einem Messraum $(\Omega^{\mathbb{N}}, \mathscr{A}^{\mathbb{N}})$ sei der *Shift-Operator* $S : \Omega^{\mathbb{N}} \to \Omega^{\mathbb{N}}$ definiert durch

$$S(\omega_1, \omega_2, \omega_3, \dots) = (\omega_2, \omega_3, \dots).$$

Der Shift-Operator ist messbar (vgl. Breiman [9, Proposition 6.10 und 6.11]).

Beweis der Äquivalenzen aus Definition 2.3:

Zum Beweis der Äquivalenz der jeweiligen Definitionen sei hier die Aussage bezüglich der invarinaten Mengen gezeigt:

„$\Rightarrow$ “ Es sei A invariant, d. h. es existiert ein $B \in \mathscr{A}_1^{\mathbb{N}}$, so dass für alle $n \in \mathbb{N}$ gilt:

$$A = \{(X_n, X_{n+1}, \dots) \in B\}.$$

Es ist zu zeigen:
Es existiert ein $C \in \mathscr{A}_1^{\mathbb{N}}$ mit

$$(x_1, x_2, x_3, \dots) \in C \Leftrightarrow (x_2, x_3, \dots) \in C \quad \text{und} \tag{2.1}$$

$$A = \{(X_1, X_2, X_3, \dots) \in C\}. \tag{2.2}$$

Mit Definition 2.4 lässt sich die Voraussetzung umschreiben zu

$$(\forall n \in \mathbb{N}_0) \quad A = \big\{(X_1, X_2, X_3, \dots) \in S^{-n}(B)\big\}. \tag{2.3}$$

Man betrachte nun die folgende Menge C:

$$C := \bigcup_{n=0}^{\infty} \bigcap_{j=n}^{\infty} S^{-j}(B) = \limsup_{j \in \mathbb{N}} S^{-j}(B).$$

Trivialerweise gilt $C \in \mathscr{A}_1^{\mathbb{N}}$.

Zu (2.1) Es gilt die folgende Äquivalenz:

$$
\begin{aligned}
& x := (x_1, x_2, \ldots) \in C \\
\Leftrightarrow\ & x \in S^{-j}(B) \text{ für alle bis auf endlich viele } j \in \mathbb{N}. \\
\Leftrightarrow\ & S(x) \in S\left(S^{-j}(B)\right) = S^{-(j-1)}(B) \text{ für alle bis auf endlich viele } j \in \mathbb{N}. \\
\Leftrightarrow\ & S(x) = (x_2, x_3, \ldots) \in C
\end{aligned}
$$

Also ist C eine invariante Teilmenge von ${\Omega_1}^{\mathbb{N}}$.

Zu (2.2) Sei $\omega \in A$, dann gilt mit (2.3)

$$(\forall n \in \mathbb{N}_0) \quad (X_1, X_2, X_3, \ldots)(\omega) \in S^{-n}(B),$$

also auch

$$(X_1, X_2, \ldots)(\omega) \ \in \ \bigcap_{j=0}^{\infty} S^{-j}(B) \ \subset \ \bigcup_{n=0}^{\infty} \bigcap_{j=n}^{\infty} S^{-j}(B) \ = \ C.$$

Somit ist $\omega \in \{(X_1, X_2, \ldots) \in C\}$.

Sei andererseits $\omega \in \{(X_1, X_2, \ldots) \in C\}$, dann existiert ein $n_0 \in \mathbb{N}_0$ mit

$$(X_1, X_2, \ldots)(\omega) \ \in \ \bigcap_{j=n_0}^{\infty} S^{-j}(B).$$

Also gilt insbesondere

$$(X_1, X_2, \ldots)(\omega) \ \in \ S^{-n_0}(B),$$

daher folgt

$$\omega \ \in \ \left\{(X_1, X_2, \ldots) \in S^{-n_0}(B)\right\} \underset{(2.3)}{=} A.$$

Somit ist insgesamt $A \ = \ \{(X_1, X_2, X_3, \ldots) \in C\}$.

„$\Leftarrow$" Es sei $A \ = \ \{(X_1, X_2, \ldots) \in B\}$, mit einer invarianten Menge $B \in {\mathscr{A}_1}^{\mathbb{N}}$.

Für $\omega \in \Omega$ gilt aufgrund der Invarianz von B

$$(X_1(\omega), X_2(\omega), \ldots) \in B \ \Leftrightarrow \ (X_2(\omega), X_3(\omega), \ldots) \in B.$$

Also ist A invariant. □

Die Systeme aller bezüglich eines Prozesses $(X_n)_{n\in\mathbb{N}}$ jeweils invarianten, terminalen oder symmetrischen Mengen weisen besondere Strukturen auf. Hierzu wird in der folgenden Definition eine Aussage gemacht.

2.5 Definition

a) Das System aller bezüglich eines Prozesses $(X_n)_{n\in\mathbb{N}}$ invarianten Mengen ist eine σ-Algebra und wird mit $\mathscr{I}(X_n, n \in \mathbb{N})$ bezeichnet.

b) Das System aller bezüglich eines Prozesses $(X_n)_{n\in\mathbb{N}}$ terminalen Mengen ist eine σ-Algebra und wird mit $\mathscr{T}(X_n, n \in \mathbb{N})$ bezeichnet.

c) Das System aller bezüglich eines Prozesses $(X_n)_{n\in\mathbb{N}}$ symmetrischen Mengen ist eine σ-Algebra und wird mit $\mathscr{S}(X_n, n \in \mathbb{N})$ bezeichnet.

2.6 Bemerkung

Wie man den Definitionen der jeweiligen Mengen in 2.3 ansehen kann, gelten die folgenden Inklusionen bei festem stochastischen Prozess $(X_n)_{n\in\mathbb{N}}$:

$$\mathscr{I}(X_n, n \in \mathbb{N}) \ \subset \ \mathscr{T}(X_n, n \in \mathbb{N}) \ \subset \ \mathscr{S}(X_n, n \in \mathbb{N}).$$

Zwischen den invarianten Mengen des Urbild- und des Bildraumes eines stochastischen Prozesses bestehen gewisse Zusammenhänge, wie die folgende Bemerkung zeigt. Diese lassen sich wiederum auch für terminale und symmetrische Mengen nachweisen.

2.7 Bemerkung

Bezeichnet man die in Definiton 2.3 eingeführten invarianten Mengen aus ${\mathscr{A}_1}^{\mathbb{N}}$ mit $\mathscr{I}\left({\mathscr{A}_1}^{\mathbb{N}}\right)$, die terminalen Mengen aus ${\mathscr{A}_1}^{\mathbb{N}}$ mit $\mathscr{T}\left({\mathscr{A}_1}^{\mathbb{N}}\right)$ und die symmetrischen Mengen aus ${\mathscr{A}_1}^{\mathbb{N}}$ mit $\mathscr{S}\left({\mathscr{A}_1}^{\mathbb{N}}\right)$, so ergibt sich folgender Zusammenhang zwischen diesen σ-Algebren und den entsprechenden σ-Algebren bezüglich eines stochastischen Prozesses $(X_n)_{n\in\mathbb{N}}$. Berücksichtigt man des Weiteren die Projektionsabbildungen $(\pi_n)_{n\in\mathbb{N}}$, die vom unendlichen Produktraum ${\Omega_1}^{\mathbb{N}}$ in die jeweilige n-te Koordinate $(n \in \mathbb{N})$ abbilden, so ergibt sich weiter unter Berücksichtigung der Tatsache, dass $\mathscr{K}\left({\mathscr{A}_1}^{\mathbb{N}}\right) = \mathscr{K}(\pi_n, n \in \mathbb{N})$ für $\mathscr{K} \in \{\mathscr{I}, \mathscr{T}, \mathscr{S}\}$ gilt :

$$\text{(i)}\ \ \mathscr{I}(X_n, n \in \mathbb{N}) \ = \ {(X_n)_{n\in\mathbb{N}}}^{-1}\left(\mathscr{I}\left({\Omega_1}^{\mathbb{N}}\right)\right) \ = \ {(X_n)_{n\in\mathbb{N}}}^{-1}\left(\mathscr{I}(\pi_n, n \in \mathbb{N})\right),$$

$$\text{(ii)}\ \ \mathscr{T}(X_n, n \in \mathbb{N}) \ = \ {(X_n)_{n\in\mathbb{N}}}^{-1}\left(\mathscr{T}\left({\Omega_1}^{\mathbb{N}}\right)\right) \ = \ {(X_n)_{n\in\mathbb{N}}}^{-1}\left(\mathscr{T}(\pi_n, n \in \mathbb{N})\right) \ \text{ und}$$

$$\text{(iii)}\ \ \mathscr{S}(X_n, n \in \mathbb{N}) \ = \ {(X_n)_{n\in\mathbb{N}}}^{-1}\left(\mathscr{S}\left({\Omega_1}^{\mathbb{N}}\right)\right) \ = \ {(X_n)_{n\in\mathbb{N}}}^{-1}\left(\mathscr{S}(\pi_n, n \in \mathbb{N})\right).$$

Nachdem nun alle relevanten Mengensysteme vorgestellt wurden, soll der Begriff der *Ergodizität* definiert werden, der ein zentraler Begriff dieser Arbeit ist.

2.8 Definition
Ein stationärer Prozess $(X_n)_{n\in\mathbb{N}}$ heißt *ergodisch*, wenn für alle $A \in \mathscr{I}(X_n, n \in \mathbb{N})$ gilt:

$$P(A) \ \in \ \{0, 1\},$$

d. h. wenn die σ-Algebra der bezüglich des Prozesses $(X_n)_{n\in\mathbb{N}}$ invarianten Mengen *trivial* (bzw. *degeneriert*) ist.

Nach der Definition der Ergodizität, die schon aus der einschlägigen Literatur bekannt ist, soll nun die neue Eigenschaft eines stationären Prozesses, die *schwache Ergodizität*, definiert und vorgestellt werden. Sie wurde der mathematischen Öffentlichkeit erstmals im Jahre 2000 von Landers und Rogge in ihrem Artikel „Weak ergodicity of stationary pairwise independent processes“ [27] präsentiert.

2.9 Definition
Ein stationärer Prozess $(X_n)_{n\in\mathbb{N}}$ heißt *schwach ergodisch*, wenn für jedes $n \in \mathbb{N}$ die Zufallsvariable X_n unabhängig von $\mathscr{I}(X_n, n \in \mathbb{N})$ ist, d. h. falls für jedes $A \in \mathscr{A}_1$, $I \in \mathscr{I}(X_n, n \in \mathbb{N})$ und $n \in \mathbb{N}$ gilt:

$$P\left([X_n \in A] \cap I\right) \ = \ P(X_n \in A) \cdot P(I)$$

bzw. falls für jedes $n \in \mathbb{N}$ und jedes $A \in \sigma(X_n)$ gilt:

$$P\left(A \mid \mathscr{I}(X_n, n \in \mathbb{N})\right) \ = \ P(A).$$

2.10 Folgerung
Aus den Definitionen 2.8 und 2.9 erkennt man sofort, dass jeder ergodische stationäre Prozess auch schwach ergodisch ist.

Aufgrund von Folgerung 2.10 ist der von Landers und Rogge neu definierte Begriff des *schwach ergodischen Prozesses* gerechtfertigt. Die definierenden Eigenschaften machen dies recht einfach deutlich. Es handelt sich dabei offensichtlich um eine Abschwächung der 0-1-Eigenschaft der bezüglich des Prozesses invarianten Mengen.

Der Nachweis, dass es sich bei der neuen Eigenschaft um eine echte Verallgemeinerung der bekannten Sachverhalte handelt, wird ebenfalls im Artikel von Landers und Rogge [27] geführt. Hierzu wird das „schöne und trickreiche“ Beispiel von Cuesta und Matrán [14] angeführt, welches unter anderem zeigt, dass es Prozesse gibt, die nach der Begriffsbildung von Landers und Rogge zwar schwach ergodisch sind, jedoch nicht ergodisch.

Neben den bezüglich eines stochastischen Prozesses $(X_n)_{n\in\mathbb{N}}$ invarianten (symmetrischen) Mengen, gibt es auch die entsprechenden invarianten (symmetrischen) Zufallsvariablen.

2.11 Definition

(i) Eine Zufallsvariable Y, die von $(\Omega, \mathscr{A}, P)$ in einen Messraum $(\Omega_2, \mathscr{A}_2)$ abbildet, heißt *invariant* (bezüglich des Prozesses $(X_n)_{n\in\mathbb{N}}$), wenn es eine messbare Funktion $\varphi : ({\Omega_1}^{\mathbb{N}}, {\mathscr{A}_1}^{\mathbb{N}}) \to (\Omega_2, \mathscr{A}_2)$ gibt, so dass für alle $n \in \mathbb{N}$ gilt:

$$Y \;=\; \varphi \circ (X_n, X_{n+1}, \dots).$$

(ii) Eine Zufallsvariable Y, die von $(\Omega, \mathscr{A}, P)$ in einem Messraum $(\Omega_2, \mathscr{A}_2)$ abbildet, heißt *symmetrisch* (bezüglich des Prozesses $(X_n)_{n\in\mathbb{N}}$), wenn es eine messbare Funktion $\varphi : ({\Omega_1}^{\mathbb{N}}, {\mathscr{A}_1}^{\mathbb{N}}) \to (\Omega_2, \mathscr{A}_2)$ gibt, so dass für jede endliche Permutation τ von $\mathbb{N}$ gilt:

$$Y \;=\; \varphi \circ (X_{\tau(1)}, X_{\tau(2)}, \dots).$$

Für reelle Zufallsvariablen kann man einen Zusammenhang zwischen den invarianten (symmetrischen) Mengen bezüglich eines Prozesses und der Invarianz (Symmetrie) der ZV bezüglich desselben Prozesses feststellen.

2.12 Lemma

Eine reelle Zufallsvariable Y auf $(\Omega, \mathscr{A}, P)$ ist genau dann invariant (symmetrisch) bezüglich des Prozesses $(X_n)_{n\in\mathbb{N}}$, wenn sie $\mathscr{I}(X_n, n \in \mathbb{N})$-messbar ($\mathscr{S}(X_n, n \in \mathbb{N})$-messbar) ist.

Beweis:

„$\Rightarrow$“ Sei Y eine invariante Zufallsvariable, dann gilt für $A \in \mathscr{B}$ und für alle $n \in \mathbb{N}$

$$\begin{aligned}\{Y \in A\} \;&=\; \{\varphi(X_n, X_{n+1}, \dots) \in A\}\\ &=\; \{(X_n, X_{n+1}, \dots) \in \varphi^{-1}(A)\}.\end{aligned}$$

Somit ist Y $\mathscr{I}(X_n, n \in \mathbb{N})$-messbar.

„$\Leftarrow$“ Diese Richtung wird unter Verwendung des Prinzips der algebraischen Induktion gezeigt.

Sei $Z(\omega) := 1_I(\omega)$ für ein $I \in \mathscr{I}(X_n, n \in \mathbb{N})$. Dann gilt mit passendem $B \in \mathscr{A}_1{}^{\mathbb{N}}$ für alle $n \in \mathbb{N}$:

$$\begin{aligned} Z(\omega) &= 1_I(\omega) = 1_{\{(X_n, X_{n+1}, \ldots) \in B\}}(\omega) \\ &= 1_B\left((X_n, X_{n+1}, \ldots)(\omega)\right). \end{aligned}$$

Somit ist Z eine invariante Zufallsvariable.

Des Weiteren sei nun $\mathscr{Z}$ die Klasse aller invarianten Zufallsvariablen auf $(\Omega, \mathscr{I})$. $\mathscr{Z}$ ist abgeschlossen unter linearen Kombinationen.

Seien $Z_n \in \mathscr{Z},\ n \in \mathbb{N},\ Z_n(\omega) \uparrow Z_0(\omega)$.
Dann gilt mit $Z_n = \varphi_n \circ (X_j, X_{j+1}, \ldots), j \in \mathbb{N}$:

$$\begin{aligned} Z_0(\omega) &= \lim_{n\to\infty} \left(\varphi_n \circ (X_j(\omega), X_{j+1}(\omega), \ldots)\right), \text{ für alle } j \in \mathbb{N} \\ &= \left(\left(\limsup_{n\to\infty} \varphi_n\right) \cdot 1_{\left[\left|\limsup\limits_{n\to\infty} \varphi_n\right| < \infty\right]}\right) \circ (X_j(\omega), X_{j+1}(\omega), \ldots), \text{ für alle } j \in \mathbb{N} \\ &=: \varphi_0 \circ (X_j(\omega), X_{j+1}(\omega), \ldots), \text{ für alle } j \in \mathbb{N}. \end{aligned}$$

Also ist auch Z_0 eine invariante Zufallsvariable. Somit folgt nach dem Buch von Breiman [9, Satz 2.38], dass $\mathscr{Z}$ jede nicht-negative Zufallsvariable auf $(\Omega, \mathscr{I})$ enthält. Daher ist jede Zufallsvariable auf $(\Omega, \mathscr{I})$, d. h. jede $\mathscr{I}(X_n, n \in \mathbb{N})$-messbare Zufallsvariable auf $(\Omega, \mathscr{A})$, invariant.

Entsprechend kann auch die Aussage für die symmetrischen Zufallsvariablen bewiesen werden. □

Mit Hilfe der bisher eingeführten Begriffe kann nun ein notwendiges und hinreichendes Kriterium für die Ergodizität eines stationären Prozesses angeben werden.

2.13 Lemma
Sei $(X_n)_{n\in\mathbb{N}}$ ein stationärer Prozeß auf $(\Omega, \mathscr{A}, P)$.
$(X_n)_{n\in\mathbb{N}}$ ist genau dann ergodisch, wenn jede invariante Zufallsvariable P-f.s. konstant ist.

Diese Aussage kann offensichtlich noch dahingehend abgeschwächt werden, dass gilt:

2.14 Lemma
Sei $(X_n)_{n\in\mathbb{N}}$ ein stationärer Prozeß auf $(\Omega, \mathscr{A}, P)$.
$(X_n)_{n\in\mathbb{N}}$ ist genau dann ergodisch, wenn jede beschränkte, invariante Zufallsvariable P-f.s. konstant ist.

Diese beiden Lemmata sind ebenfalls dem Buch von Breiman [9, Proposition 6.18 und 6.19] entnommen. Die Beweise können dort nachvollzogen werden.

Stationäre Prozesse werden oft über *maßerhaltende* bzw. *maßtreue Transformationen* eingeführt und definiert. Diese Vorgehensweise ist nicht nur als Spezialfall anzusehen, da bekannt ist, dass sich jeder stationäre Prozess $(X_n)_{n\in\mathbb{N}}$ im Rahmen seiner Verteilung durch einen Prozess $(X \circ T^{n-1})_{n\in\mathbb{N}}$ darstellen lässt, der mit Hilfe einer maßerhaltenden Transformation T und einer Zufallsvariablen X erzeugt wird. Dieses wird im Anschluss erläutert.

Zuvor sollen jedoch die hierfür grundlegende Definition einer maßerhaltenden Tranformation angegeben werden.

2.15 Definition

Eine messbare Funktion $T : (\Omega, \mathscr{A}, P) \to (\Omega, \mathscr{A}, P)$ heißt *maßerhaltende Transformation*, wenn für alle $A \in \mathscr{A}$ gilt:

$$P(T^{-1}A) \;=\; P(A).$$

2.16 Lemma

Sei T eine maßerhaltende Transformation auf $(\Omega, \mathscr{A}, P)$ und X eine Zufallsvariable auf $(\Omega, \mathscr{A})$. Dann ist der Prozess $(X_n)_{n\in\mathbb{N}}$, mit

$$X_n(\omega) \;:=\; X\left(T^{n-1}(\omega)\right) \quad , n \in \mathbb{N},$$

ein stationärer Prozess.

Der Beweis dieses Lemmas ist im Buch von Breiman [9, Proposition 6.9] gegeben.

In Anlehnung an die obige Aussage über den Zusammenhang der Verteilungen von stationären Prozessen und maßerhaltenden Transformationen kann die Behauptung konkret nachgewiesen werden. Man betrachte den Shift-Operator S, der in 2.4 definiert worden ist, den Koordinatenprozess π_1, für den $\pi_1(x_1, x_2, \ldots) = x_1$ gilt und die Zufallsvariable $X := (X_n)_{n\in\mathbb{N}}$ auf dem unendlichen Produktraum. Als Maß auf dem unendlichen Produktraum betrachte man das Verteilungsmaß P_X und den Prozess $(\pi_1 \circ S^{n-1})_{n\in\mathbb{N}}$ und erkennt, dass die Verteilung dieses Prozesses mit der Verteilung des Prozesses $(X_n)_{n\in\mathbb{N}}$ übereinstimmt.

Weiter erkennt man, dass der Shift-Operator S genau dann eine maßerhaltende Transformation ist, wenn der vorliegende Prozess stationär ist. Dies ist ebenfalls im Buch von Breiman [9] gezeigt.

Im Anschluss sollen die zuvor in diesem Paragraphen gegebenen Definitionen auf den Fall der maßerhaltenden Transformationen umformuliert werden.

2.17 Definition
Eine Menge $A \in \mathscr{A}$ ist *invariant* (bezüglich der Transformation T), wenn $T^{-1}A = A$ ist.

Das System aller bezüglich T invarianten Mengen ist eine σ-Algebra und wird mit $\mathscr{I}(T)$ bezeichnet.

2.18 Definition
Sei T eine maßerhaltende Transformation auf $(\Omega, \mathscr{A}, P)$. T heißt *ergodisch*, wenn die σ-Algebra der bezüglich T invarianten Mengen, also $\mathscr{I}(T)$, trivial ist.

Zum Abschluss der umformulierten Definitionen im Zusammenhang mit maßerhaltenden Transformationen sei noch die Definition der invarianten Zufallsvariable für den Fall der Darstellung mittels einer maßerhaltenden Transformation angegeben.

2.19 Definition
Seien X eine Zufallsvariable und T eine maßerhaltende Transformation auf $(\Omega, \mathscr{A}, P)$. X ist eine *invariante Zufallsvariable* (bezüglich T), wenn

$$X(\omega) \;=\; X(T\omega) \quad (\forall \omega \in \Omega).$$

2.20 Lemma
Sei $(X_n)_{n\in\mathbb{N}}$ ein stationärer Prozess auf $(\Omega, \mathscr{A}, P)$, der sich mit Hilfe der maßerhaltenden Transformation T darstellen lässt, d. h. es gilt

$$X_i \;=\; X_1 \circ T^{i-1} \quad (\forall i \in \mathbb{N})\,.$$

Dann gilt:

$$\mathscr{I}(X_n, n \in \mathbb{N}) \;\subset\; \mathscr{I}(T).$$

Beweis:
Sei $A \in \mathscr{I}(X_n, n \in \mathbb{N})$, dann existiert ein $B \in \mathscr{A}_1^{\mathbb{N}}$ mit

$$A \;=\; \{(X_n, X_{n+1}, \ldots) \in B\} \quad (\forall n \in \mathbb{N}). \tag{2.4}$$

Es ist zu zeigen:

$$T^{-1}A = A.$$

Nun gilt:

$$\begin{aligned}
T^{-1}(A) &= \{\omega \mid T(\omega) \in A\} \\
&= \{\omega \mid (X_1, X_2, X_3 \ldots)(T\omega) \in B\} \\
&= \big\{\omega \mid (X_1, X_1 \circ T, X_1 \circ T^2, \ldots)(T\omega) \in B\big\} \\
&= \big\{\omega \mid (X_1 \circ T, X_1 \circ T^2, X_1 \circ T^3, \ldots)(\omega) \in B\big\} \\
&= \{(X_2, X_3, \ldots) \in B\} \underset{(2.4)}{=} A.
\end{aligned}$$

□

2.21 Bemerkung

(i) Im Allgemeinen gilt:

$$\mathscr{I}(X_n, n \in \mathbb{N}) \neq \mathscr{I}(T).$$

Hierzu betrachte man einen Wahrscheinlichkeitsraum $(\Omega, \mathscr{A}, P)$ mit nichttrivialer σ-Algebra $\mathscr{A}$. T sei die Identität auf $(\Omega, \mathscr{A}, P)$, die offensichtlich maßerhaltend ist. Daher folgt also $\mathscr{I}(T) = \mathscr{A}$ und für konstantes X_1 gilt trivialerweise $\mathscr{I}(X_n, n \in \mathbb{N}) = \{\emptyset, \Omega\}$. Daher folgt $\mathscr{I}(X_n, n \in \mathbb{N}) \neq \mathscr{I}(T)$.

(ii) Es gilt für eine maßerhaltende Transformation T:

(a) Ist T ergodisch, so ist der stationäre Prozess $(X \circ T^{n-1})_{n\in\mathbb{N}}$ mit einer beliebigen Zufallsvariablen X ergodisch.

(b) $(X \circ T^{n-1})_{n\in\mathbb{N}}$ ist für jede Zufallsvariable X, die unabhängig von $\mathscr{I}(T)$ ist, schwach ergodisch, wie aus dem Zusammenspiel von Lemma 2.20 und Lemma 2.22 folgt.

Unter Berücksichtigung der Darstellung eines schwach ergodischen stationären Prozesses gilt das folgende Lemma.

2.22 Lemma

Es sei $(X_n)_{n\in\mathbb{N}}$ ein stationärer Prozess.

Der Prozess $(X_n)_{n\in\mathbb{N}}$ ist genau dann schwach ergodisch, wenn X_1 unabhängig von $\mathscr{I}(X_n,\ n \in \mathbb{N})$ ist.

Beweis:

„$\Leftarrow$“ Es seien X_1 unabhängig von $\mathscr{I}(X_n, n \in \mathbb{N})$ und $I \in \mathscr{I}(X_n, n \in \mathbb{N})$.

Nach Definition 2.3 gilt mit passendem $B \in \mathscr{A}_1^{\mathbb{N}}$ für alle $n \in \mathbb{N}$:

$$I = \{(X_1, X_2, \ldots) \in B\} = \{(X_n, X_{n+1}, \ldots) \in B\} \tag{2.5}$$

Für die schwache Ergodizität ist für beliebiges $n \in \mathbb{N}$ und $B_1 \in \mathscr{A}_1$ zu zeigen:

$$P\left((X_n \in B_1) \cap I\right) \;=\; P\left(X_n \in B_1\right) \cdot P(I). \tag{2.6}$$

Es gilt:

$$\begin{aligned}
P\left((X_n \in B_1) \cap I\right) &\underset{(2.5)}{=} P\left((X_n \in B_1) \cap ((X_n, X_{n+1}, \ldots) \in B)\right) \\
&= P\left((X_n, X_{n+1}, \ldots) \in \left(B_1 \times {\Omega_1}^{\mathbb{N}}\right) \cap B\right) \\
&\underset{\text{Stationarität}}{=} P\left((X_1, X_2, \ldots) \in \left(B_1 \times {\Omega_1}^{\mathbb{N}}\right) \cap B\right) \\
&= P\left((X_1 \in B_1) \cap I\right) \\
&= P\left(X_1 \in B_1\right) \cdot P(I) \\
&\underset{\text{Bemerkung 2.2}}{=} P\left(X_n \in B_1\right) \cdot P(I).
\end{aligned}$$

Also gilt (2.6) und damit ist $(X_n)_{n\in\mathbb{N}}$ schwach ergodisch.

„$\Rightarrow$ “ Diese Richtung des Beweises ist mit der Definition der schwachen Ergodizität (Definition 2.9) offensichtlich. □

2.23 Bemerkung

Lemma 2.22 kann aufgrund weiterer Eigenschaften stationärer Prozesse noch einfacher bewiesen werden. Dies wird in Lemma 4.6 geschehen, wenn die für einen kurzen Beweis notwendigen Begriffe und Zusammenhänge definiert und gezeigt sind.

Somit wurden alle grundlegenden Definitionen und Zusammenhänge, die in der vorliegenden Arbeit benötigt werden, aufgezeigt. Mit Hilfe dieser Begriffsbasis können nun die Forschungen im Umfeld der schwach ergodischen Prozesse beginnen.

§ 3

Eigenschaften schwach ergodischer Prozesse

Inhalt des vorigen Paragraphens war die Definition schwach ergodischer Prozesse. Bevor die Darstellung dieser Prozesse im weiteren Verlauf der Arbeit behandelt wird, sollen im Folgenden deren grundlegende Eigenschaften herausgestellt werden. Insbesondere werden in diesem Paragraphen die folgenden Thematiken behandelt:

- $\mathscr{I}(\varphi(X_n, X_{n+1}, \ldots), n \in \mathbb{N}) \ \subset \ \mathscr{I}(X_n, n \in \mathbb{N})$.
- Für einen schwach ergodischen stationären Prozess $(X_n)_{n \in \mathbb{N}}$ ist auch der Prozess $(\varphi \circ X_n)_{n \in \mathbb{N}}$ schwach ergodisch.
- Die Voraussetzung der Ergodizität im Ergodensatz von Birkhoff [6] wird durch die allgemeinere Voraussetzung der schwachen Ergodizität ersetzt.
- Die Beziehung zweier Wahrscheinlichkeitsmaße zueinander wird für den Fall untersucht, dass der stochastische Prozess bezüglich beider Maße schwach ergodisch ist.
- Zum Abschluss wird die Abgeschlossenheitseigenschaft der Menge aller schwach ergodischen stationären Prozesse diskutiert.

Für die weiteren Bearbeitungen soll zunächst einmal die Lage der σ-Algebren der invarianten Mengen verschiedener stochastischer Prozesse betrachtet werden.

3.1 Lemma
Es seien M, M_1 separable metrische Räume und $X_n, \ n \in \mathbb{N}$, M-wertige Zufallsvariablen.

a) Ist $\varphi : M^{\mathbb{N}} \to M_1$ eine Borel-messbare Abbildung, dann gilt:

$$\mathscr{I}\left(\varphi(X_n, X_{n+1}, \dots), n \in \mathbb{N}\right) \subset \mathscr{I}(X_n, n \in \mathbb{N}).$$

b) Ist M zusätzlich auch vollständig und $\varphi : M \to \mathbb{R}$ eine injektive und Borel-messbare Abbildung, dann gilt:

$$\mathscr{I}(\varphi \circ X_n, n \in \mathbb{N}) = \mathscr{I}(X_n, n \in \mathbb{N}).$$

Beweis:

a) Man setze für $(x_1, x_2, \dots) \in M^{\mathbb{N}}$:

$$\Phi(x_1, x_2, \dots) := (\varphi(x_1, x_2, \dots), \varphi(x_2, x_3, \dots), \dots).$$

Dann ist Φ messbar.
Sei $I \in \mathscr{I}\left(\varphi(X_n, X_{n+1}, \dots), n \in \mathbb{N}\right)$. Dann existiert nach Definition 2.3 eine invariante Menge $B \subset {M_1}^{\mathbb{N}}$ mit $I = {(X_n)_{n \in \mathbb{N}}}^{-1}\left(\Phi^{-1}(B)\right)$.

Es genügt zu zeigen:

$$\Phi^{-1}(B) \in \mathscr{I}\left(M^{\mathbb{N}}\right), \tag{3.1}$$

denn dann folgt mit Definition 2.3 und Bemerkung 2.7 sofort $I \in \mathscr{I}(X_n, n \in \mathbb{N})$.
Es gilt:

$$\begin{aligned}
(y_1, y_2, y_3, \dots) \in \Phi^{-1}(B) \quad &\Leftrightarrow \quad \Phi(y_1, y_2, y_3, \dots) \in B \\
&\Leftrightarrow \quad (\varphi(y_1, y_2, \dots), \varphi(y_2, y_3, \dots), \dots) \in B \\
&\underset{B \text{ invariant}}{\Leftrightarrow} \quad (\varphi(y_2, y_3, \dots), \varphi(y_3, y_4, \dots), \dots) \in B \\
&\Leftrightarrow \quad \Phi(y_2, y_3, \dots) \in B \\
&\Leftrightarrow \quad (y_2, y_3, \dots) \in \Phi^{-1}(B)
\end{aligned}$$

Also folgt (3.1).

b) In Anlehnung an a) genügt es hier zu zeigen:

$$\mathscr{I}(\varphi \circ X_n, n \in \mathbb{N}) \supset \mathscr{I}(X_n, n \in \mathbb{N}).$$

Diese Inklusion folgt aus a) angewendet auf $(\varphi \circ X_n)$ anstelle von X_n und $\varphi^{-1} : \varphi(M) \to M$ anstelle von φ. Die für a) notwendige Borel-Messbarkeit von φ^{-1} ist nach dem Buch von Cohn [13, Satz 8.3.7 und Satz 8.3.5] durch die spezielle Wahl des Raumes M gegeben. □

3.2 Bemerkung

Die Aussagen des obigen Lemmas lassen sich auch auf die umfassenderen σ-Algebren der terminalen und der symmetrischen Mengen bezüglich eines stochastischen Prozesses $(X_n)_{n\in\mathbb{N}}$ übertragen. Die terminale Menge $T \in \mathscr{T}\,(\varphi(X_n, X_{n+1}, \dots), n \in \mathbb{N})$ bzw. die symmetrische Menge $S \in \mathscr{S}\,(\varphi(X_n, X_{n+1}, \dots), n \in \mathbb{N})$ lässt sich im Beweisgang des obigen Lemmas nach Definition 2.3 dann mit Hilfe einer terminalen bzw. symmetrischen Menge $B \subset {M_1}^{\mathbb{N}}$ darstellen. Aufgrund der definierenden Bedingungen für diese Mengen folgt daher die Behauptung.

Zunächst soll untersucht werden, ob sich die schwache Ergodizität eines stationären Prozesses bei der Verknüpfung desselben mit einer messbaren Funktion auf den neuentstandenen Prozess vererbt. Die Fragen nach der Stationarität und der Ergodiziät werden im Buch von Breiman [9, Proposition 6.6 und Proposition 6.31] bereits besprochen, hiernach gilt der folgende Satz:

3.3 Satz

Seien $(X_n)_{n\in\mathbb{N}}$ ein stochastischer Prozess von $(\Omega, \mathscr{A}, P)$ in $(\Omega_1, \mathscr{A}_1)$, φ eine messbare Abbildung von $({\Omega_1}^{\mathbb{N}}, {\mathscr{A}_1}^{\mathbb{N}})$ in $(\Omega_2, \mathscr{A}_2)$ und für alle $n \in \mathbb{N}$ sei $Y_n := \varphi(X_n,\ X_{n+1},\ \dots)$. Dann gilt:

(i) Ist $(X_n)_{n\in\mathbb{N}}$ stationär, so auch $(Y_n)_{n\in\mathbb{N}}$.

(ii) Ist $(X_n)_{n\in\mathbb{N}}$ ergodisch, so auch $(Y_n)_{n\in\mathbb{N}}$

Aus Lemma 3.1 wird der obige Sachverhalt für ergodische Prozesse deutlich, da nach der dortigen Aussage gilt:

$$\mathscr{I}(X_n, n \in \mathbb{N}) \supset \mathscr{I}(\varphi(X_k,\ X_{k+1},\ \dots), k \in \mathbb{N}).$$

Im Falle der schwachen Ergodizität wird jedoch nur die Unabhängigkeit jeder einzelnen ZV des stationären Prozesses von den invarianten Mengen desselben Prozesses gefordert, eine allgemeine Unabhängigkeit unter den ZV wird nach Definition nicht verlangt. Man beachte daher, dass nach den Ausssagen der Unabhängigkeitsuntersuchungen zwar z. B. $\sigma(X_1)$ und $\sigma(X_2)$ jeweils unabhängig von $\mathscr{I} := \mathscr{I}(X_n, n \in \mathbb{N})$ sind, jedoch folgt hieraus nur die Unabhängigkeit des Dynkin-Systems $D(X_1, X_2)$ von $\mathscr{I}$. Um jedoch eine Aussage über die Unabhängigkeit von $Y = \varphi(X_1, X_2)$ und $\mathscr{I}$ machen zu können muss zusätzlich die Durchschnittsstabilität von $\sigma(X_1) \cup \sigma(X_2)$ gewährleistet sein (vgl. A.6).

Für schwach ergodische stationäre Prozesse lässt sich der folgende Zusammenhang festhalten.

3.4 Satz

Seien $(X_n)_{n\in\mathbb{N}}$ ein stationärer Prozess von $(\Omega, \mathscr{A}, P)$ nach $(\Omega_1, \mathscr{A}_1)$ und φ eine messbare Abbildung von $(\Omega_1, \mathscr{A}_1)$ nach $(\Omega_2, \mathscr{A}_2)$.
Ist der Prozess $(X_n)_{n\in\mathbb{N}}$ schwach ergodisch, so auch der Prozess $(\varphi(X_n))_{n\in\mathbb{N}}$.

Beweis:

Nach Definition der schwachen Ergodizität ist $\sigma(X_n) = {X_n}^{-1}(\mathscr{A}_1)$ unabhängig von $\mathscr{I}(X_n, n \in \mathbb{N})$ für jedes $n \in \mathbb{N}$. Des Weiteren gilt nach Lemma 3.1:

$$\mathscr{I}\left(\varphi \circ X_n, n \in \mathbb{N}\right) \subset \mathscr{I}(X_n, n \in \mathbb{N}) \tag{3.2}$$

und

$$\left(\varphi \circ X_n\right)^{-1}\left(\mathscr{A}_2\right) = {X_n}^{-1}\left(\varphi^{-1}\left(\mathscr{A}_2\right)\right) \subset {X_n}^{-1}(\mathscr{A}_1). \tag{3.3}$$

Aus (3.2) und (3.3) folgt nun die Behauptung aufgrund der vorausgesetzen schwachen Ergodizität des Prozesses $(X_n)_{n\in\mathbb{N}}$. □

Das nun folgende Beispiel soll aufzeigen, dass die Aussage von Satz 3.3 (ii) nicht allgemein auf schwach ergodische Prozesse $(X_n)_{n\in\mathbb{N}}$ übertragbar ist.

3.5 Beispiel

Es sei $p \in \mathbb{N}$ eine Primzahl mit $p \geq 3$.
Weiter seien Y_0, U und $\{Z_{n\cdot p}, n \in \mathbb{N}_0\}$ unabhängige Zufallsvariablen, die auf der Menge $\{0, 1, 2, \ldots, p-1\}$ gleichverteilt sind.
Man definiere

$$\begin{aligned} Z_{0+k} &:= Z_0 \oplus k \cdot Y_0, && k = 0, 1, \ldots, p-1 \text{ und} \\ Z_{n\cdot p+k} &:= Z_{n\cdot p} \oplus k \cdot Y_0, && k = 0, 1, \ldots, p-1, n \in \mathbb{N}, \end{aligned}$$

wobei $\oplus$ die Summe modulo p meint.
Als Prozess definiere man für $k \in \mathbb{N}$:

$$W_k := Z_{k+U}. \tag{3.4}$$

Weiter setze man für $n \in \mathbb{N}$:

$$X_n := \frac{2}{p-1} \cdot W_n - 1 \tag{3.5}$$

Nach den Angaben aus dem Artikel von Cuesta und Matrán [14] handelt es sich bei dem Prozess $(X_n)_{n\in\mathbb{N}}$ um einen paarweise unabhängigen und stationären Prozess. Der vorliegende Prozess ist somit nach dem Ergebnis von Landers und Rogge, das in Satz 6.1 noch explizit notiert wird, schwach ergodisch, jedoch nicht ergodisch.

Der in diesem Beispiel erwünschte Nachweis wird mit Hilfe einer gewichteten Addition geführt. Man betrachte dazu den Prozess $(B_n)_{n\in\mathbb{N}}$, der durch

$$B_n \;:=\; 2\cdot X_n \,+\, X_{n+1}, \quad n\in\mathbb{N} \tag{3.6}$$

definiert wird. Mit dieser Festlegung gilt dann

$$P(B_1=-3) \underset{(3.6)}{=} P(X_1=-1;X_2=-1) \underset{\text{paarw. unabh.}}{=} \frac{1}{p}\cdot\frac{1}{p} = \frac{1}{p^2}. \tag{3.7}$$

Nach dem Artikel von Cuesta und Matrán folgt, dass

(i) auf $[Y_0=0]$ gilt:

$$\limsup_{n\to\infty}\frac{S_n}{\sigma\cdot\sqrt{2\cdot n\cdot\log(\log(n))}} = +\sqrt{p}$$

und

$$\liminf_{n\to\infty}\frac{S_n}{\sigma\cdot\sqrt{2\cdot n\cdot\log(\log(n))}} = -\sqrt{p}$$

(ii) auf $[Y_0\neq 0]$ gilt:

$$\begin{aligned}\limsup_{n\to\infty}\frac{S_n}{\sigma\cdot\sqrt{2\cdot n\cdot\log(\log(n))}} &= \liminf_{n\to\infty}\frac{S_n}{\sigma\cdot\sqrt{2\cdot n\cdot\log(\log(n))}}\\ &= \lim_{n\to\infty}\frac{S_n}{\sigma\cdot\sqrt{2\cdot n\cdot\log(\log(n))}}\\ &= 0\end{aligned}$$

wobei σ die Standard-Abweichung von X_1 und $(S_n)_{n\in\mathbb{N}}$ die Folge der Partialsummen des Prozesses $(X_n)_{n\in\mathbb{N}}$ ist.
Die obigen Grenzwerte sind in der Literatur unter dem Begriff „Law of the iterated logarithm (LIL)“ bekannt.

Mit diesen Ergebnissen im Umfeld des LIL und (3.6) kann man erkennen, dass für den Prozess $(B_n)_{n\in\mathbb{N}}$ mit den entsprechenden Partialsummen $({}^{B}S_n)_{n\in\mathbb{N}}$ gilt:

$$\begin{aligned} I_B \;&:=\; \left[\limsup_{n\to\infty}\frac{{}^{B}S_n}{\sigma\cdot\sqrt{2\cdot n\cdot\log(\log(n))}} \;>\; 0\right]\\ &= \left[\limsup_{n\to\infty}\frac{S_n}{\sigma\cdot\sqrt{2\cdot n\cdot\log(\log(n))}} \;>\; 0\right]\\ &= [Y_0=0]\\ &\in \mathscr{I}(B_n, n\in\mathbb{N}) \;=:\; \mathscr{I}_B\end{aligned}$$

Nach Definition der Zufallsvariablen Y_0 folgt nun weiter:

$$P(I_B) = P(Y_0 = 0) = \frac{1}{p}. \tag{3.8}$$

Somit gilt im Rahmen der Unabhängigkeitsuntersuchung von B_1 und $\mathscr{I}_B$ nach (3.7) und (3.8):

$$\begin{aligned} P(B_1 = -3) \cdot P(I_B) &= \frac{1}{p^2} \cdot \frac{1}{p} \\ &= \frac{1}{p^3}. \end{aligned} \tag{3.9}$$

Ebenso folgt:

$$\begin{aligned} & P([B_1 = -3] \cap I_B) \\ &= P(X_1 = -1;\ X_2 = -1;\ Y_0 = 0) \\ &= P(W_1 = 0;\ W_2 = 0;\ Y_0 = 0) \\ &= P\left(\bigcup_{k=0}^{p-1} [W_1 = 0;\ W_2 = 0;\ Y_0 = 0;\ U = k]\right) \\ &= \sum_{k=0}^{p-1} P(W_1 = 0;\ W_2 = 0;\ Y_0 = 0;\ U = k) \\ &= \sum_{k=0}^{p-1} P(Z_{1+U} = 0;\ Z_{2+U} = 0;\ Y_0 = 0;\ U = k) \\ &= \sum_{k=0}^{p-3} P(Z_0 \oplus (1+k) \cdot Y_0 = 0;\ Z_0 \oplus (2+k) \cdot Y_0 = 0;\ Y_0 = 0;\ U = k) \\ &\quad + P(Z_0 \oplus (p-1) \cdot Y_0 = 0;\ Z_p = 0;\ Y_0 = 0;\ U = p-2) \\ &\quad + P(Z_1 = 0;\ Z_p \oplus Y_0 = 0;\ Y_0 = 0;\ U = p-1) \\ &= \sum_{k=0}^{p-3} P(Z_0 = 0;\ Y_0 = 0;\ U = k) \\ &\quad + P(Z_0 = 0;\ Z_p = 0;\ Y_0 = 0;\ U = p-2) \\ &\quad + P(Z_p = 0;\ Y_0 = 0;\ U = p-1) \\ &= (p-2) \cdot \frac{1}{p^3} + \frac{1}{p^4} + \frac{1}{p^3} \\ &= (p-1) \cdot \frac{1}{p^3} + \frac{1}{p^4}. \end{aligned} \tag{3.10}$$

Aufgrund der Aussagen aus (3.9) und (3.10) folgt sofort, dass B_1 und I_B nicht unabhängig sind. Somit ist der Prozess $(B_n)_{n \in \mathbb{N}}$ nicht schwach ergodisch.

Wie schon der Name vermuten lässt, ist eine der Haupteigenschaften eines ergodischen stationären Prozesses $(X_n)_{n \in \mathbb{N}}$ die Gültigkeit des SLLN unter Verwendung des

Ergodensatzes von Birkhoff [6] aus dem Jahre 1931, der die P-f.s.-Konvergenz von $\frac{1}{n} \cdot \sum_{k=1}^{n} X_k \underset{n\to\infty}{\longrightarrow} P\left(X_1 \mid \mathscr{I}(X_n, n \in \mathbb{N})\right)$ für stationäre Prozesse sicherstellt. Als adäquate Folgerung aus diesem Ergodensatz wird nun der nächste Satz angegeben, der die Gültigkeit des SLLN für schwach ergodische stationäre Prozesse liefert.

3.6 Satz (Ergodensatz bei schwacher Ergodizität)

Es sei $(X_n)_{n\in\mathbb{N}}$ ein schwach ergodischer stationärer Prozess auf $(\Omega, \mathscr{A}, P)$, und X_1 sei integrierbar, dann gilt:

$$\frac{1}{n} \cdot \sum_{k=1}^{n} X_k \underset{n\to\infty}{\longrightarrow} P(X_1) \quad P\text{-f.s.}$$

und

$$P\left(\left|\frac{1}{n} \cdot \sum_{k=1}^{n} X_k \;-\; P(X_1)\right|\right) \underset{n\to\infty}{\longrightarrow} 0,$$

d. h. es gilt die Konvergenz im ersten Mittel.

Beweis:

Nach Definition eines schwach ergodischen Prozesses (siehe 2.9) ist $(X_n)_{n\in\mathbb{N}}$ ein stationärer Prozess. Somit gilt nach dem Birkhoff'schen Ergodensatz für allgemeine stationäre Prozesse [9, Proposition 6.28]:

$$\frac{1}{n} \cdot \sum_{k=1}^{n} X_k \underset{n\to\infty}{\longrightarrow} P\left(X_1 \mid \mathscr{I}(X_n, n \in \mathbb{N})\right) \quad P\text{-f.s.}. \tag{3.11}$$

Berücksichtigt man nun die zweite definierende Eigenschaft der schwachen Ergodizität in Zusammenspiel mit der Bemerkung A.2 (d), so folgt:

$$P\left(X_1 \mid \mathscr{I}(X_n, n \in \mathbb{N})\right) \;=\; P(X_1) \quad P\text{-f.s.}. \tag{3.12}$$

Also folgt aus (3.11) und (3.12) die erste Behauptung.

Die Behauptung über die Konvergenz im ersten Mittel ergibt sich aus dem Zusammenwirken des allgemeinen Beweises dieser Aussage aus dem Buch von Breiman [9, Corollary 6.25] und Bemerkung A.2 (d). □

3.7 Bemerkung

Im Falle eines nicht-negativen Prozesses $(Y_n)_{n\in\mathbb{N}}$ wird neben der Stationarität keine weitere Eigenschaft für die Gültigkeit des schwachen Ergodensatzes benötigt. Es kann gezeigt werden (vgl. [9, 6.6]), dass für einen nicht-negativen stationären Prozess $(Y_n)_{n\in\mathbb{N}}$ mit $P(Y_1) = \infty$ gilt:

$$\frac{1}{n} \cdot \sum_{k=1}^{n} Y_k \underset{n\to\infty}{\longrightarrow} \infty \quad P\text{-f.s.}.$$

Die schwache Ergodizität eines Prozesses $(X_n)_{n\in\mathbb{N}}$ hängt angesichts der definierenden Eigenschaften dieser Prozesse entscheidend von dem auf einem Messraum $(\Omega, \mathscr{A})$ definierten Wahrscheinlichkeitsmaß P ab. Somit stellt sich die Frage, in welcher Beziehung zwei Maße P und Q auf $(\Omega, \mathscr{A})$ zueinander stehen, wenn ein Prozess $(X_n)_{n\in\mathbb{N}}$ bezüglich beider Maße schwach ergodisch ist.

3.8 Korollar

Sei $(\Omega, \mathscr{A})$ ein Messraum. Des Weiteren seien P und Q zwei Wahrscheinlichkeitsmaße auf $(\Omega, \mathscr{A})$, unter denen $(X_n)_{n\in\mathbb{N}}$, mit $X_n : (\Omega, \mathscr{A}) \to (\Omega_1, \mathscr{A}_1),\ n \in \mathbb{N}$, ein stationärer und schwach ergodischer Prozess ist.

Dann gilt entweder $P_{X_1} = Q_{X_1}$ oder P und Q sind orthogonal, genauer gesagt existiert eine Menge $C \in \mathscr{I}(X_n, n \in \mathbb{N})$ mit $P(C) = 0$ und $Q(C) = 1$.

Beweis:

Sei $P_{X_1} \neq Q_{X_1}$, dann existiert ein $B \in \mathscr{A}_1$ mit

$$P(X_1 \in B) \;=\; P_{X_1}(B) \;\neq\; Q_{X_1}(B) \;=\; Q(X_1 \in B).$$

Man setze

$$Y_n \;:=\; 1_B \circ X_n,\ n \in \mathbb{N}.$$

Dann ist $(Y_n)_{n\in\mathbb{N}}$ nach Satz 3.3 ein stationärer Prozess und nach Satz 3.4 ist jedes $Y_n,\ n \in \mathbb{N}$, unabhängig von den invarianten Mengen $\mathscr{I}(Y_n, n \in \mathbb{N})$. Diese Aussagen gelten bezüglich beider Wahrscheinlichkeitsmaße P und Q. Also gilt unter Zuhilfenahme des *Ergodensatzes bei schwacher Ergodizität* 3.6:

$$\begin{aligned} \frac{1}{n} \cdot \sum_{k=1}^{n} Y_k \underset{n\to\infty}{\longrightarrow} Q(Y_1) \quad & Q\text{-f.s.} \\ = Q(1_B \circ X_1) \quad & Q\text{-f.s.} \\ = Q_{X_1}(B) \quad & Q\text{-f.s..} \end{aligned} \tag{3.13}$$

Sei C die Menge aller $\omega \in \Omega$, für die die Q-f.s.-Konvergenz in (3.13) gilt, dann folgt

$$Q(C) = 1. \tag{3.14}$$

Da $Q_{X_1}(B) \neq P_{X_1}(B)$ vorausgesetzt wurde, ist die Menge aller $\omega \in \Omega$, für die die Konvergenz

$$\frac{1}{n} \cdot \sum_{k=1}^{n} Y_k \underset{n\to\infty}{\longrightarrow} P_{X_1}(B) \tag{3.15}$$

gilt, in $\Omega \setminus C$ enthalten. Die Konvergenz in (3.15) ist aufgrund der Stationarität und der schwachen Ergodizität bezüglich P f.-s. gegeben, daher gilt:

$$1 \;=\; P(\{\omega \mid (3.15) \text{ gilt}\}) \;\leq\; P(\Omega \setminus C) \;\leq\; 1.$$

Also ist $P(\Omega \setminus C) = 1$ oder anders gesagt $P(C) = 0$.
Die Invarianz der Menge C ist aufgrund ihrer Definition klar.
Somit sind P und Q orthogonal. □

3.9 Bemerkung

Die klassische Darstellung dieses Korollares im Umfeld ergodischer, maßerhaltener Transformationen findet man im Buch von Breiman [9, Corollary 6.24]. Um diesen Spezialfall aus dem obigen Ergebnis zu erhalten, sei auf die Darstellung eines stationären Prozesses im Hinblick auf seine Verteilung nach Lemma 2.16 verwiesen. Berücksichtigt man in der dortigen Darstellung die Festlegung $X := id$, so erhält man das Ergebnis aus dem Buch von Breiman. Dort wird zwar vom Maß P und nicht von der Verteilung von X_1 bezüglich P gesprochen, jedoch entsprechen sich in diesem Spezialfall nach den obigen Erläuterungen P_{X_1} und P, da $X_1 = id \circ T^{1-1} = id$ ist.

Eine wichtige mathematische Frage ist die nach der Abgeschlossenheit eines Systems von stochastischen Prozessen bezüglich einer geeigneten Konvergenz. Zum Zwecke dieser Überlegungen wird das folgende Lemma benötigt.

3.10 Lemma

Es seien f eine messbare Funktion von $(\Omega^\circ, \mathscr{A}^\circ, P^\circ)$ in $(\Omega^{\circ\circ}, \mathscr{A}^{\circ\circ})$ und $\mathscr{G} \subset \mathscr{A}^\circ$ eine σ-Algebra.
Seien dann weiter $(P_n)_{n \in \mathbb{N}_0}$ eine Folge von Maßen auf $(\Omega^\circ, \mathscr{A}^\circ)$ und $A \in \mathscr{A}^{\circ\circ}$. Gilt dann:

$$\lim_{n \to \infty} P_n(B) \;=\; P_0(B) \quad (\forall B \in \mathscr{A}^\circ) \tag{3.16}$$

und

$$P_n(1_A \circ f \mid \mathscr{G}) \;=\; P_n(1_A \circ f) \;\; P_n\text{-f.s.} \quad (\forall n \in \mathbb{N})\,, \tag{3.17}$$

so gilt auch

$$P_0(1_A \circ f \mid \mathscr{G}) = P_0(1_A \circ f) \;\; P_0\text{-f.s.}.$$

Beweis:

Sei $A \in \mathscr{A}^{\circ\circ}$ mit der oben geforderten Eigenschaft.

Es ist zu zeigen:

$$P_0(1_A \circ f \mid \mathscr{G}) \;=\; P_0(1_A \circ f) \quad P_0\text{-f.s.}.$$

Die $\mathscr{G}$-Messbarkeit der Konstanten $P_0(1_A \circ f)$ ist klar.

Sei nun $G \in \mathscr{G}$, dann gilt

$$\begin{aligned}
\int_G P_0(1_A \circ f)\, \mathrm{d}P_0 &= P_0(1_A \circ f) \cdot P_0(G) \\
&\underset{(3.16)}{=} \lim_{n\to\infty} \left(P_n(1_A \circ f) \cdot P_n(G)\right) \\
&= \lim_{n\to\infty} \int_G P_n(1_A \circ f)\, \mathrm{d}P_n \\
&\underset{(3.17)}{=} \lim_{n\to\infty} \int_G P_n(1_A \circ f \mid \mathscr{G})\, \mathrm{d}P_n \\
&= \lim_{n\to\infty} \int_G 1_A \circ f\, \mathrm{d}P_n \\
&= \int_G 1_A \circ f\, \mathrm{d}P_0.
\end{aligned}$$

Also folgt insgesamt

$$P_0(1_A \circ f \mid \mathscr{G}) \;=\; P_0(1_A \circ f) \quad P_0\text{-f.s.}.$$

□

Ziel dieser Überlegungen ist eine Aussage über die Abgeschlossenheit der Menge der schwach ergodischen stationären Prozesse $(X_n)_{n\in\mathbb{N}}$ auf einem Wahrscheinlichkeitsraum $(\Omega, \mathscr{A}, P)$ bezüglich einer geeigneten Konvergenz. Diese ist unter Verwendung der soeben dokumentierten Ergebnisse möglich. Hierzu werden die folgenden Festlegungen unter Verwendung der Begrifflichkeiten aus Lemma 3.10 vorgenommen:

(a) $(\Omega^\circ, \mathscr{A}^\circ, P^\circ) \;:=\; ({\Omega_1}^{\mathbb{N}}, {\mathscr{A}_1}^{\mathbb{N}}, {P_1}^{\mathbb{N}})$.

(b) $P_i \;:=\; P_{(X_{i,1}, X_{i,2}, X_{i,3}, \dots)}$ für $i \in \mathbb{N}$, wobei $((X_{i,n})_{n\in\mathbb{N}})_{i\in\mathbb{N}}$ gerade die Folge von schwach ergodischen Prozessen ist, die in einer Weise gegen den Prozess $(X_n)_{n\in\mathbb{N}}$ konvergiert, welche im folgenden Satz noch verdeutlicht wird.

(c) f sei in den folgenden n verschiedenen Fällen ($n \in \mathbb{N}$) durch die Koordinatenabbildungen $\pi_n,\ n \in \mathbb{N}$, gegeben.

(d) $\mathscr{G} \;:=\; \mathscr{I}(\pi_n, n \in \mathbb{N})$.

(e) Nach den obigen Festlegungen ergibt sich $(\Omega^{\circ\circ}, \mathscr{A}^{\circ\circ}, P^{\circ\circ}) \;=\; (\Omega_1, \mathscr{A}_1, P_1)$.

Setzt man nun diese konkreten Werte in die Voraussetzungen des Lemmas 3.10 ein und fordert die Eigenschaft (3.17) für jedes $A \in \mathscr{A}^{\infty} = \mathscr{A}_1$, so ergibt sich unter Berücksichtigung der Tatsache, dass diese verallgemeinerte Anforderung nach Bemerkung 2.7 gerade die definierende Eigenschaft der schwachen Ergodizität ist, der folgende Satz.

3.11 Satz

Seien $(X_{i,n})_{n\in\mathbb{N}}$ für jedes $i \in \mathbb{N}$ schwach ergodische stationäre Prozesse und die Verteilungen der Prozesse $X_i := (X_{i,n})_{n\in\mathbb{N}}$ konvergieren derart gegen die Verteilung des Prozesses $X_0 := (X_{0,n})_{n\in\mathbb{N}}$, dass gilt

$$\lim_{i\to\infty} P_{(X_{i,n})_{n\in\mathbb{N}}}(A) \;=\; P_{(X_{0,n})_{n\in\mathbb{N}}}(A) \quad \left(\forall A \in \mathscr{A}_1{}^{\mathbb{N}}\right),$$

dann ist auch der Prozess $(X_{0,n})_{n\in\mathbb{N}}$ schwach ergodisch.

Bevor ein Beispiel dafür angegeben werden kann, dass man die Konvergenz der Verteilung des gesamten Prozesses benötigt, und man nicht mit der entsprechenden Konvergenz einer beliebigen endlichen Randverteilung zum selben Ergebnis kommt, soll ein weiterer Zusammenhang (bzw. eine Übereinstimmung der Begriffe) zwischen schwacher Ergodizität und Ergodizität aufgezeigt werden. Dieses Lemma wird für das angesprochene Beispiel benötigt.

3.12 Lemma

Es sei $(X_n)_{n\in\mathbb{N}}$ ein stationärer Prozess, für den ein $n_0 \in \mathbb{N}$ existiert mit $X_{n_0} \equiv$id. Der Prozess $(X_n)_{n\in\mathbb{N}}$ ist ergodisch, wenn er schwach ergodisch ist.

Beweis:

Wegen $X_{n_0} \equiv$id gilt

$$\mathscr{A} \;=\; X_{n_0}{}^{-1}(\mathscr{A}),$$

also gilt nach der definierenden Eigenschaft der schwach ergodischen Prozesse:

$$\mathscr{A} \text{ ist unabhängig von } \mathscr{I}(X_n, n \in \mathbb{N}) \;\subset\; \mathscr{A}.$$

Somit folgt die Trivialität von $\mathscr{I}(X_n, n \in \mathbb{N})$. □

Mit diesem Lemma kann nun das im Anschluss an Satz 3.11 angekündigte Gegenbeispiel angegeben werden.

3.13 Beispiel

Es seien $\mu_n \in \mathbb{R} \setminus \mathbb{Q}$, $n \in \mathbb{N}$ mit $\mu_n \downarrow 0$.

Dann gilt auf $\left([0,1[, \mathscr{B}_{[0,1[}, \lambda_{[0,1[}\right)$ mit der Modulo-Abbildung $[.]$, also der Division mit Rest,

$$T(\omega) \;:=\; \omega \text{ und } T_n(\omega) \;:=\; (\omega + \mu_n)\,[1] \quad (\forall \omega \in [0,1[, n \in \mathbb{N}), \text{ dass}$$

$$Y_i := T^{i-1} = id \text{ und } X_{i,n} := {T_n}^{i-1} \quad (\forall i, n \in \mathbb{N})$$

und

$$Y := (Y_n)_{n\in\mathbb{N}} \text{ und } X_i := (X_{i,n})_{n\in\mathbb{N}} \quad (\forall i \in \mathbb{N})$$

stationäre Prozesse sind. X_n ist nach dem Buch von Petersen [31, Example (2), Seite 49] für jedes $n \in \mathbb{N}$ ergodisch, Y aber nicht. Unter Zuhilfenahme des obigen Lemmas 3.12 ist Y auch nicht schwach ergodisch.

Aus der Definition der einzelnen Zufallsvariablen $X_{i,n},\ i, n \in \mathbb{N}$ wird sofort deutlich, dass $X_{i,n}$ für $n \to \infty$ gegen Y_i konvergiert. Somit folgt aber auch die Konvergenz jeder endlichen Auswahl von ZV $(X_{1,n}, X_{2,n}, \ldots, X_{m,n})$ ($m \in \mathbb{N}$ beliebig) für $n \to \infty$ gegen $(Y_1, Y_2, \ldots, Y_m)$. Daher ist die Vorraussetzung der Konvergenz jeder beliebigen endlichen Randverteilung erfüllt. Jedoch ist Y, wie schon vorher gesagt, nicht schwach ergodisch. Also genügt die Konvergenz aller endlichen Randverteilungen einer Folge von schwach ergodischen stationären Prozessen nicht aus, um die schwache Ergodizität des Grenzprozesses zu induzieren. Hierfür muss die Konvergenz für die Verteilungen der gesamten Prozesse erfüllt sein.

Dieses Beispiel macht aber noch etwas anderes deutlich:
Angesichts der Begrifflichkeiten wäre es sicherlich erstrebenswert, dass die schwache Konvergenz einer Folge von schwach ergodischen stationären Prozessen auf $\Omega^{\mathbb{N}}$ für die schwache Ergodizität des Grenzprozesses ausreicht. Nach obiger Bearbeitung gilt für beliebiges $n \in \mathbb{N}$:

$$\left(\forall A \in \mathscr{A}^{\mathbb{N}}\right) \quad P_{(X_{i,k})_{1\le k\le n}}(A) \underset{i\to\infty}{\longrightarrow} P_{(X_{0,k})_{1\le k\le n}}(A). \tag{3.18}$$

Des Weiteren gilt aber auch für $j, n \in \mathbb{N}_0$:

$$P_{(X_{j,k})_{1\le k\le n}}(A) \;=\; P_{(X_{j,k})_{k\in\mathbb{N}}}\left((\pi_1, \pi_2, \ldots, \pi_n)^{-1}(A)\right). \tag{3.19}$$

Unter Beachtung der Definition der schwachen Konvergenz auf dem Folgenraum $\Omega^{\mathbb{N}}$, wie sie im Buch von Billingsley [5, Chapter 1.3] dargestellt wird, stellt man fest, dass der oben dargestellte Umstand ((3.18) und (3.19)) hinreichend für die schwache Konvergenz ist. Somit wird durch das obige Beispiel auch ein Gegenbeispiel dafür gegeben, dass die schwache Konvergenz einer Folge von schwach ergodischen Prozessen für die schwache Ergodizität des Grenzprozesses nicht ausreicht.

Somit kommt man zu folgenden Bemerkungen.

3.14 Bemerkung

(i) Es sei $(X_i)_{i\in\mathbb{N}} := ((X_{i,n})_{n\in\mathbb{N}})_{i\in\mathbb{N}}$ punktweise konvergent gegen Y für $i \in \mathbb{N}$. X_i ist für jedes $i \in \mathbb{N}$ schwach ergodisch $\not\Rightarrow$ Y schwach ergodisch.

(ii) Es sei $(X_i)_{i\in\mathbb{N}} := ((X_{i,n})_{n\in\mathbb{N}})_{i\in\mathbb{N}}$ eine Folge stochastischer Prozesse und $(P_{X_i})_{i\in\mathbb{N}}$ konvergiere schwach gegen P_Y. X_i ist für jedes $i \in \mathbb{N}$ schwach ergodisch $\not\Rightarrow$ Y schwach ergodisch.

In diesem Paragraphen wurde deutlich, dass die Voraussetzung der Ergodizität im Birkhoff'schen Ergodensatz durch die schwache Ergodizität ersetzt werden kann. Angesichts dieser Erkenntnis ist es möglich, mathematische Sachverhalte für ergodische stationäre Prozesse, die die Gültigkeit des Birkhoff'schen Ergodensatzes benötigen, auch auf schwach ergodische stationäre Prozesse zu verallgemeinern. Hierbei muss jedoch beachtet werden, dass sich die schwache Ergodizität des stationären Prozesses $(X_n)_{n\in\mathbb{N}}$ nur auf den stationären Prozess $(\varphi(X_n))_{n\in\mathbb{N}}$ überträgt und nicht unbedingt auf den Prozess $(\psi(X_n, X_{n+1}, \ldots))_{n\in\mathbb{N}}$ (vgl. Satz 3.4 und Satz 3.3). Dieses wäre aber z. B. bei der Verallgemeinerung des folgenden Sachverhaltes nötig [9, Theorem 6.40]:

Es sei $(X_n)_{n\in\mathbb{N}}$ ein ergodischer stationärer Prozess und $(T_n)_{n\in\mathbb{N}}$ sei der stochastische Prozess, der aus den Rekurrenzzeiten für das Eintreffen des Prozesses $(X_n)_{n\in\mathbb{N}}$ in der Menge $A \in \mathscr{A}_1$ gebildet wird, d. h.

$$T_1 := \min\{n \in \mathbb{N} \mid X_n \in A\} \text{ und } T_k := R_k - R_{k-1},\ k \in \mathbb{N}_{\geq 2}$$

mit

$$R_1 := T_1 \text{ und } R_k := \min\{n \in \mathbb{N} \mid X_n \in A, n > R_{k-1}\},\ k \in \mathbb{N}_{\geq 2}.$$

Dann ist $(T_n)_{n\in\mathbb{N}}$ auf $[X_1 \in A]$ ergodisch bezüglich des Maßes $P(\cdot \mid X_1 \in A)$.

§ 4

Bedingte Verteilungen

Der folgende Paragraph beschäftigt sich mit der schon im vorherigen Paragraphen angesprochenen *bedingten Verteilung* von Zufallsvariablen. Die nun vorgestellten Eigenschaften und Zusammenhänge dienen der Motivation der Darstellung schwach ergodischer Prozesse, die im nächsten Paragraphen besprochen werden soll. Aufgrund der Aussagen des Satzes von de Finetti für vertauschbare Prozesse werden äquivalente Aussagen im Falle der stationären Prozesse gesucht. Des Weiteren werden spezielle Inklusionseigenschaften spezieller σ-Algebren für vertauschbare Prozesse angegeben. Zu den konkreten Ergebnissen und Aussagen dieses Paragraphen gehören

- der Satz von de Finetti,
- der Zusammenhang zwischen der Stationarität eines Prozesses und der bedingt identischen Verteilung bezüglich dessen invarianter Mengen,
- die Beziehung der symmetrischen und der f.-s.-invarianten Mengen eines vertauschbaren Prozesses zueinander,
- die Erweiterung der schwachen Ergodizität auf $\mathscr{G} \supset \mathscr{I}(X_n, n \in \mathbb{N})$ und
- die Äquivalenz der schwachen Ergodizität und der Trivialität der symmetrischen Mengen eines vertauschbaren Prozesses.

Zu Beginn dieser Betrachtungen sollen die schon mehrfach angesprochenen Begriffe der *bedingt identischen Verteilung* und der *bedingten Unabhängigkeit* eingeführt werden.

4.1 Definition

Es seien $(X_n)_{n\in\mathbb{N}}$ ein stochastischer Prozess mit $X_n : (\Omega, \mathscr{A}, P) \to (\Omega_1, \mathscr{A}_1),\ n \in \mathbb{N}$ und $\mathscr{G}$ eine Unter-σ-Algebra von $\mathscr{A}$.

(a) X_1 und X_2 heißen *bedingt identisch verteilt bezüglich* $\mathscr{G}$, wenn für jedes $B \in \mathscr{A}_1$ gilt:

$$P(X_1 \in B \mid \mathscr{G}) \ = \ P(X_2 \in B \mid \mathscr{G}) \quad P\text{-f.s.}.$$

(b) $X_1, X_2, \ldots, X_n,\ n \in \mathbb{N}$, heißen *bedingt unabhängig bezüglich* $\mathscr{G}$, wenn für alle $A_i \in \sigma(X_i),\ 1 \le i \le n$, gilt:

$$P(A_1 \cap A_2 \cap \ldots \cap A_n \mid \mathscr{G}) \ = \ \prod_{i=1}^{n} P(A_i \mid \mathscr{G}) \quad P\text{-f.s.}.$$

(c) Der Prozess $(X_n)_{n\in\mathbb{N}}$ heißt *bedingt unabhängig bezüglich* $\mathscr{G}$, wenn für jedes $n \in \mathbb{N}$ die Zufallsvariablen $X_1, X_2, \ldots, X_n$ bedingt unabhängig bezüglich $\mathscr{G}$ sind.

4.2 Bemerkung

(a) Falls $\mathscr{G} = \{\emptyset, \Omega\}$ ist, so ist die bedingte Unabhängigkeit bezüglich $\mathscr{G}$ gerade die allgemeine (unbedingte) Unabhängigkeit.

(b) Falls $\mathscr{G} = \mathscr{A}$ betrachtet wird, so ist jeder stochastische Prozess bedingt unabhängig bezüglich $\mathscr{G}$.

(c) Aus der Unabhängigkeit von Zufallsvariablen folgt i.A. nicht deren bedingte Unabhängigkeit bezüglich einer Unter-σ-Algebra $\mathscr{G} \subset \mathscr{A}$.

(d) Abhängige Zufallsvariablen können bedingt unabhängig bezüglich einer Unter-σ-Algebra $\mathscr{G} \subset \mathscr{A}$ sein.

(e) Für Beispiele betrachte man Seite 221 des Buches von Chow und Teicher [12].

Einer der grundlegenden Sätze im Zusammenhang mit der bedingten Unabhängigkeit und der bedingt identischen Verteilung ist der Satz von de Finetti [12, Theorem 2, Seite 224]. Seine zentrale Aussage ist die Definition der vertauschbaren Prozesse durch die bedingte Unabhängigkeit und die bedingt identische Verteilung bezüglich spezieller σ-Algebren.

4.3 Satz (Satz von de Finetti)

Ein stochastischer Prozess $(X_n)_{n\in\mathbb{N}}$ mit $X_n : (\Omega, \mathscr{A}, P) \to (\Omega_1, \mathscr{A}_1)$, $n \in \mathbb{N}$, ist genau dann vertauschbar, wenn er bedingt identisch verteilt und bedingt unabhängig bezüglich einer Unter-σ-Algebra $\mathscr{G} \subset \mathscr{A}$ ist.

Hierbei kann $\mathscr{G}$ insbesondere als die terminale σ-Algebra $\mathscr{T}(X_n, n \in \mathbb{N}) =: \mathscr{T}$ oder als die symmetrische σ-Algebra $\mathscr{S}(X_n, n \in \mathbb{N}) =: \mathscr{S}$ gewählt werden, und es gilt in diesem Fall für $A \in \mathscr{A}$:

$$P(X_1 \in A \mid \mathscr{S}) \;=\; P(X_1 \in A \mid \mathscr{T}), \quad P\text{-f.s..}$$

Dieser Satz wird in der Dissertation von Heß [24, Nummer 7.3.2] in einer etwas abgewandelten Version vorgestellt. Den anschließenden Bemerkungen zum dortigen siebten Kapitel kann man die Tatsache entnehmen, dass die Gleichheit der bedingten Verteilungen im Satz von de Finetti nicht nur auf $\mathscr{S}$ und $\mathscr{T}$ anwendbar ist, sondern auch auf $\mathscr{I}(X_n, n \in \mathbb{N}) =: \mathscr{I}$ erweitert werden kann. Dieses ist durch den Birkhoff'schen Ergodensatz (vgl. [6] bzw. [9, Proposition 6.28]) begründet. Die notwendigen Betrachtungen im Zusammenhang mit den invarianten Mengen werden im Artikel von Olshen [29] durchgeführt.

Es ergibt sich somit der folgende Satz:

4.4 Satz

Für einen stochastischen Prozess $(X_n)_{n\in\mathbb{N}}$ sind die folgenden Aussagen äquivalent:

(i) $(X_n)_{n\in\mathbb{N}}$ ist ein vertauschbarer Prozess.

(ii) Der Prozess $(X_n)_{n\in\mathbb{N}}$ ist bedingt identisch verteilt und bedingt unabhängig bezüglich einer Unter-σ-Algebra $\mathscr{G} \subset \mathscr{A}$.

(iii) Der Prozess $(X_n)_{n\in\mathbb{N}}$ ist bedingt identisch verteilt und bedingt unabhängig bezüglich $\mathscr{S}(X_n, n \in \mathbb{N})$.

(iv) Der Prozess $(X_n)_{n\in\mathbb{N}}$ ist bedingt identisch verteilt und bedingt unabhängig bezüglich $\mathscr{T}(X_n, n \in \mathbb{N})$.

(v) Der Prozess $(X_n)_{n\in\mathbb{N}}$ ist bedingt identisch verteilt und bedingt unabhängig bezüglich $\mathscr{I}(X_n, n \in \mathbb{N})$.

Es stellt sich die Frage, ob diese Ergebnisse in gewissem Sinne auf stationäre Prozesse übertragbar sind, d. h. ob für stationäre Prozesse eine spezielle bedingte Unabhängigkeit oder bedingt identische Verteilung bezüglich einer speziellen Unter-σ-Algebra $\mathscr{G} \subset \mathscr{A}$ vorliegt. Erste Überlegungen führen zu folgendem Ergebnis.

4.5 Satz
Es sei $(X_n)_{n\in\mathbb{N}}$ ein stationärer Prozess.
Dann ist $(X_n)_{n\in\mathbb{N}}$ bedingt identisch verteilt bezüglich $\mathscr{I}(X_n, n \in \mathbb{N})$.

Beweis:
Nach Definition der bedingten identischen Verteilung in 4.1 ist für beliebiges $m \in \mathbb{N}$ und $B \in \mathscr{A}_1$ zu zeigen:

$$P\left(1_B \circ X_1 \mid \mathscr{I}(X_n, n \in \mathbb{N})\right) \;=\; P\left(1_B \circ X_m \mid \mathscr{I}(X_n, n \in \mathbb{N})\right) \;\; P\text{-f.s.}. \tag{4.1}$$

Unter Berücksichtigung der Definition der bedingten Erwartung in A.1 und der dort gemachten Aussage über die Übereinstimmung zweier verschiedener bedingter Erwartungen bezüglich derselben σ-Algebra, genügt es für (4.1) mit beliebigem $m \in \mathbb{N}, I \in \mathscr{I}(X_n, n \in \mathbb{N})$ und $B \in \mathscr{A}_1$ zu zeigen:

$$\int_I P(1_B \circ X_1 \mid \mathscr{I}(X_n, n \in \mathbb{N}))\, \mathrm{d}P \;=\; \int_I 1_B \circ X_m \, \mathrm{d}P. \tag{4.2}$$

Seien nun $m \in \mathbb{N}$ und $I \in \mathscr{I}(X_n, n \in \mathbb{N})$, dann existiert nach Definition der invarianten Mengen 2.3 eine Menge $C \in \mathscr{A}_1{}^{\mathbb{N}}$ mit

$$I \;=\; (X_n)_{n\in\mathbb{N}}{}^{-1}C \;=\; (X_{n+m})_{n\in\mathbb{N}}{}^{-1}C. \tag{4.3}$$

Des Weiteren sei $B \in \mathscr{A}_1$, dann folgt:

$$\begin{aligned}
\int_I P(1_B \circ X_1 \mid \mathscr{I}(X_n, n \in \mathbb{N}))\, \mathrm{d}P \quad &= \quad \int_I 1_B \circ X_1 \, \mathrm{d}P \\
&= \quad \int (1_B \circ \pi_1 \cdot 1_C) \circ (X_1, X_2, \dots)\; \mathrm{d}P \\
&\underset{2.1}{=} \quad \int (1_B \circ \pi_1 \cdot 1_C) \circ (X_m, X_{m+1}, \dots)\; \mathrm{d}P \\
&= \quad \int 1_B \circ X_m \cdot 1_{(X_m, X_{m+1}, \dots)^{-1}C}\, \mathrm{d}P \\
&\underset{(4.3)}{=} \quad \int_I 1_B \circ X_m \, \mathrm{d}P.
\end{aligned}$$

Somit ist (4.1) gezeigt, und es folgt daher die bedingt identische Verteilung von $(X_n)_{n\in\mathbb{N}}$ bezüglich $\mathscr{I}(X_n, n \in \mathbb{N})$ für jeden stationären Prozess $(X_n)_{n\in\mathbb{N}}$. □

Mit Hilfe dieses Satzes kann nun der in Bemerkung 2.23 angekündigte kurze Beweis für Lemma 2.22 gegeben werden, der die bedingt identische Veteilung eines stationären Prozesses bezüglich seiner invarianten Mengen berücksichtigt.

4.6 Lemma
Sei $(X_n)_{n\in\mathbb{N}}$ ein stationärer Prozess. Dann ist $(X_n)_{n\in\mathbb{N}}$ genau dann schwach ergodisch, wenn X_1 unabhängig von $\mathscr{I}(X_n, n \in \mathbb{N})$ ist.

Beweis:

„$\Rightarrow$“ Diese Richtung folgt aus der Definition der schwachen Ergodizität 2.9.

„$\Leftarrow$“ Sei X_1 unabhängig von $\mathscr{I} := \mathscr{I}(X_n, n \in \mathbb{N})$ und $n \in \mathbb{N}$, dann genügt es nach Bemerkung A.2 (e) für ein beliebiges $B \in \mathscr{A}_1$ zu zeigen:

$$P\left(X_n \in B \mid \mathscr{I}\right) \quad = \quad P\left(X_n \in B\right) \quad P\text{-f.s..} \tag{4.4}$$

Nach Satz 4.5 folgt:

$$P\left(X_1 \in B \mid \mathscr{I}\right) \; = \; P\left(X_n \in B \mid \mathscr{I}\right) \quad P\text{-f.s..} \tag{4.5}$$

Somit folgt weiter:

$$\begin{aligned} P\left(X_n \in B \mid \mathscr{I}\right) \quad & \underset{(4.5)}{=} \quad P\left(X_1 \in B \mid \mathscr{I}\right) \quad P\text{-f.s.} \\ & = \quad P\left(X_1 \in B\right) \quad P\text{-f.s.} \\ & = \quad P\left(X_n \in B\right) \quad P\text{-f.s.} \end{aligned}$$

Also gilt (4.4). □

Bei einem schwach ergodischen Prozess ist nicht nur die Unabhängigkeit von X_1 und den invarianten Mengen gegeben, sondern selbstverständlich auch die Unabhängigkeit von X_1 und allen Mengen, die sich nur um Nullmengen von den invarianten Mengen unterscheiden. Diese Überlegung führt zur folgenden Definition.

4.7 Definition
Man definiere die Vervollständigung $\mathscr{G}^*$ einer Unter-σ-Algebra $\mathscr{G} \subset \mathscr{A}$ auf $(\Omega, \mathscr{A}, P)$ durch

$$\mathscr{G}^* \; := \; \{A \in \mathscr{A} \mid (\exists G \in \mathscr{G}) \; P(A \triangle G) = 0\},$$

wobei die *symmetrische Differenz* $\triangle$ für $B, C \in \mathscr{A}$ folgendermaßen definiert ist:

$$B \triangle C \quad := \quad (B \setminus C) \; \cup \; (C \setminus B).$$

Die Vervollständigung $\mathscr{G}^*$ ist ebenfalls eine Unter-σ-Algebra von $\mathscr{A}$.

4.8 Bemerkung

(i) Entsprechend der Definitionen im Paragraphen 2 lassen sich die definierenden Eigenschaften der schwachen Ergodizität, der Ergodizität und die Trivialitätseigenschaften für σ-Algebren allgemein auf die Vervollständigungen der entsprechenden σ-Algebren übertragen.

(ii) Die Äquivalenzen im verallgemeinerten Satz von de Finetti 4.4 lassen sich auch auf die Vervollständigungen $\mathscr{I}^*(X_n, n \in \mathbb{N})$, $\mathscr{T}^*(X_n, n \in \mathbb{N})$, $\mathscr{S}^*(X_n, n \in \mathbb{N})$ und $\mathscr{G}^*$ ausweiten. Hierbei gilt im Spezialfall ebenfalls die Übereinstimmung der bedingten Verteilungen des Prozesses $(X_n)_{n\in\mathbb{N}}$ bezüglich der hier erwähnten σ-Algebren.

An dieser Stelle sollen erste Eigenschaften angegeben werden, die sich aus dem Zusammenspiel der bedingten Unabhängigkeit resp. Verteilung und den vervollständigten σ-Algebren ergeben.

4.9 Lemma

Sind die Zufallsvariablen X und Y bedingt identisch verteilt bezüglich einer σ-Algebra $\mathscr{G} \subset \mathscr{A}$, so auch bezüglich jeder Unter-σ-Algebra $\mathscr{F} \subset \mathscr{G}$, also sind X und Y insbesondere identisch verteilt.

Beweis:

Für ein beliebiges $A \in \mathscr{A}_1$ gilt:

$$P\,(X \in A \mid \mathscr{G}) \quad = \quad P\,(Y \in A \mid \mathscr{G})$$

Daher folgt mit Lemma A.3 (b) und Bemerkung A.2 (c):

$$\begin{aligned} P\,(X \in A \mid \mathscr{F}) &= P\,(P\,(X \in A \mid \mathscr{G}) \mid \mathscr{F}) \\ &= P\,(P\,(Y \in A \mid \mathscr{G}) \mid \mathscr{F}) \\ &= P\,(Y \in A \mid \mathscr{F}) \end{aligned}$$

□

4.10 Lemma

Für die Vervollständigung $\mathscr{G}^*$ einer σ-Algebra $\mathscr{G}$ gilt:

Sind die Zufallsvariablen X und Y bedingt identisch verteilt bezüglich $\mathscr{G}$, so auch bezüglich $\mathscr{G}^*$.

Beweis:
Es sei $(X_n)_{n\in\mathbb{N}}$ bedingt identisch verteilt bezüglich $\mathscr{G}$, dann ist zu zeigen:

$$P\left(X_n \in A \mid \mathscr{G}^*\right) \;=\; P\left(X_1 \in A \mid \mathscr{G}^*\right) \;\; P\text{-f.s.} \quad (\forall A \in \mathscr{A})\,(\forall n \in \mathbb{N})\,.$$

Seien nun $A \in \mathscr{A}, n \in \mathbb{N}$ und $G^* \in \mathscr{G}^*$. Nach Definition der Vervollständigung einer σ-Algebra existiert ein $G \in \mathscr{G}$ mit $P\left(G \triangle G^*\right) = 0$.
Somit gilt wegen $G^* \uplus (G \setminus G^*) \;=\; G \uplus (G^* \setminus G)$:

$$\begin{aligned}
\int_{G^*} P\left(1_A \circ X_n \mid \mathscr{G}^*\right)\, \mathrm{d}P &= \int_{G^*} 1_A \circ X_n \,\mathrm{d}P \\
&= \int_{G} 1_A \circ X_n \,\mathrm{d}P + \int_{G^*\setminus G} 1_A \circ X_n \,\mathrm{d}P - \int_{G\setminus G^*} 1_A \circ X_n \,\mathrm{d}P \\
&= \int_{G} 1_A \circ X_n \,\mathrm{d}P \\
&= \int_{G} P\left(1_A \circ X_n \mid \mathscr{G}\right)\, \mathrm{d}P \\
&= \int_{G} P\left(1_A \circ X_1 \mid \mathscr{G}\right)\, \mathrm{d}P \\
&= \int_{G} 1_A \circ X_1 \,\mathrm{d}P \\
&= \int_{G} 1_A \circ X_1 \,\mathrm{d}P + \int_{G^*\setminus G} 1_A \circ X_1 \,\mathrm{d}P - \int_{G\setminus G^*} 1_A \circ X_1 \,\mathrm{d}P \\
&= \int_{G^*} 1_A \circ X_1 \,\mathrm{d}P \\
&= \int_{G^*} P\left(1_A \circ X_1 \mid \mathscr{G}^*\right)\, \mathrm{d}P.
\end{aligned}$$

Somit ist $P\left(X_n \in A \mid \mathscr{G}^*\right)$ eine Version von $P\left(X_1 \in A \mid \mathscr{G}^*\right)$ und daher gilt nun auch $P\left(X_n \in A \mid \mathscr{G}^*\right) \;=\; P\left(X_1 \in A \mid \mathscr{G}^*\right)$ P-f.s.. □

4.11 Folgerung
Nach den obigen Lemmata 4.9 und 4.10 und Satz 4.5 folgt:

(a) Die Zufallsvariablen X und Y sind genau dann bedingt identisch verteilt bezüglich der σ-Algebra $\mathscr{G}$, wenn sie bedingt identisch verteilt bezüglich der Vervollständigung $\mathscr{G}^*$ sind und es gilt offensichtlich für $A \in \mathscr{A}_1$:

$$P(X \in A \mid \mathscr{G}) \;=\; P(X \in A \mid \mathscr{G}^*) \quad P\text{-f.s.}.$$

(b) Jeder stationäre Prozess $(X_n)_{n\in\mathbb{N}}$ ist bedingt identisch verteilt bezüglich der Vervollständigung $\mathscr{I}^*(X_n, n \in \mathbb{N})$ der invarianten Mengen $\mathscr{I}(X_n, n \in \mathbb{N})$.
$\mathscr{I}^*(X_n, n \in \mathbb{N})$ wird die σ-Algebra der f.-s.-invarianten Mengen genannt.

Somit lässt sich also für stationäre Prozesse die bedingte identische Verteilung bezüglich der σ-Algebra der invarianten Mengen resp. der f.-s.-invarianten Mengen zeigen.

Im Satz von de Finetti (Satz 4.4) ist jedoch die Rede von einer Äquivalenzaussage für vertauschbare Prozesse. Dadurch stellt sich nun auch die Frage, ob jeder stochastische Prozess $(X_n)_{n\in\mathbb{N}}$, der bedingt identisch verteilt bezüglich $\mathscr{I}(X_n, n \in \mathbb{N})$ ist, auch stationär ist. Mit Hilfe des folgenden Lemmas kann diese Vermutung durch ein Gegenbeispiel widerlegt werden.

4.12 Lemma
Ist $(X_n)_{n\in\mathbb{N}}$ ein stochastischer Prozess und X_0 eine messbare Abbildung, so gilt für den stochastischen Prozess $(X_j)_{j\in\mathbb{N}_0}$:

$$\mathscr{I}(X_j, j \in \mathbb{N}_0) \;\subset\; \mathscr{I}(X_n, n \in \mathbb{N}).$$

Beweis:
Sei $A \in \mathscr{I}(X_j, j \in \mathbb{N}_0)$, dann gilt nach Definition 2.3:

$$(\exists B \in \mathscr{A}_1{}^{\mathbb{N}})(\forall j \in \mathbb{N}_0)\; A = \{(X_j, X_{j+1}, \ldots \in B)\}.$$

Also gilt auch:

$$(\exists B \in \mathscr{A}_1{}^{\mathbb{N}})(\forall n \in \mathbb{N})\; A = \{(X_n, X_{n+1}, \ldots \in B)\}.$$

Daher ist A wiederum nach Definition 2.3 in $\mathscr{I}(X_n, n \in \mathbb{N})$ enthalten. $\square$

4.13 Beispiel
Sei $\Omega := [0, 1[$, $\mathscr{A} := \mathscr{B}_{[0,1[}$ und $P := \lambda_{[0,1[}$.
Des Weiteren betrachte man die folgendermaßen definierte Transformation T:

$$T(\omega) \;:=\; (2 \cdot \omega)\,[1], \text{ für alle } \omega \in \Omega$$

und den folgendermaßen definierten stochastischen Prozess $(X_n)_{n\in\mathbb{N}}$:

$$X_n \;:=\; T^{n-1}, \text{ für alle } n \in \mathbb{N}.$$

Als weitere messbare Abbildung betrachte man

$$X_0(\omega) \;:=\; 1 - \omega, \text{ für alle } \omega \in \Omega.$$

Nach dem Buch von Breiman [9, Problem 10, Seite 117] ist die Transformation T ergodisch und somit ist $\mathscr{I}(T)$ trivial, also nach Lemma 2.20 auch $\mathscr{I}(X_n, n \in \mathbb{N})$.

(Man erkennt dies auch daran, dass die Zufallsvariablen $X_n,\ n \in \mathbb{N}$, unabhängig und identisch verteilt sind.)

Daher folgt nach Lemma 4.12 auch die Trivialität von $\mathscr{I} := \mathscr{I}(X_j, j \in \mathbb{N}_0)$. Somit gilt nun auch für jedes $A \in \mathscr{A}$ und jedes $j \in \mathbb{N}_0$:

$$P(X_j \in A \mid \mathscr{I}) \ = \ P(X_j \in A).$$

Daher folgt die bedingt identische Verteilung des Prozesses bezüglich $\mathscr{I}$ aus der ursprünglichen identischen Verteilung der $(X_j)_{j \in \mathbb{N}_0}$.

Sei nun $B := \left[0, \frac{1}{2}\right] \times \left[0, \frac{1}{2}\right] \times \mathbb{R}^{\mathbb{N}} \in \mathscr{B}^{\mathbb{N}}$. Dann gilt:

$$P\left((X_0, X_1, \ldots) \in B\right) \ = \ P\left(X_0 < \frac{1}{2},\ X_1 < \frac{1}{2}\right) \ = \ 0$$

und

$$P\left((X_1, X_2, \ldots) \in B\right) \ = \ P\left(X_1 < \frac{1}{2},\ X_2 < \frac{1}{2}\right) \ = \ \frac{1}{4}.$$

Demnach ist der stochastische Prozess $(X_j)_{j \in \mathbb{N}_0}$ nicht stationär, jedoch bedingt identisch verteilt bezüglich $\mathscr{I}(X_j, j \in \mathbb{N}_0)$.

Somit bleibt nach den bisherigen Erörterungen neben anderen die Frage nach einer bedingten Unabhängigkeit von stationären Prozessen offen. Besinnt man sich der Aussagen des Satzes von de Finetti 4.3 bzw. des Satzes 4.4 und des Satzes 4.5, so kann man direkt feststellen, dass eine allgemeine bedingte Unabhängigkeit der Zufallsvariablen bezüglich der invarianten Mengen nicht möglich sein kann, da ansonsten die Begriffe *vertauschbarer* und *stationärer* Prozess zusammenfallen würden. Dieses ist aber bekanntlich nicht der Fall.

Daher liegen höchstens die Vermutungen nahe, dass die paarweise bedingte Unabhängigkeit bezüglich der invarianten Mengen oder die bedingte Unkorreliertheit, ebenfalls bezüglich der invarianten Mengen, vorliegen könnte. Jedoch auch für diese Vermutungen kann man Gegenbeispiele finden, wie im Folgenden gezeigt wird.

4.14 Beispiel

Es sei $(X_n)_{n \in \mathbb{N}}$ ein ergodischer stationärer Prozess auf $(\Omega, \mathscr{A}, P)$, bei dem die Zufallsvariablen $X_n, n \in \mathbb{N}$, nicht paarweise unabhängig sind und $\mathscr{I} := \mathscr{I}(X_n, n \in \mathbb{N})$.

Aufgrund der Ergodizität des Prozesses folgt nun für alle $A,\ B \in \mathscr{A}_1$ und alle $n,\ k \in \mathbb{N}$:

$$P(X_n \in A \mid \mathscr{I}) \ = \ P(X_n \in A)$$

und

$$P(X_n \in A,\ X_k \in B \mid \mathscr{I}) \ = \ P(X_n \in A, X_k \in B).$$

Des Weiteren gilt wegen der Ergodizität und der dadurch induzierten Stationarität nach Satz 4.5 die bedingt identische Verteilung bezüglich $\mathscr{I}$. Da die $X_n, n \in \mathbb{N}$, nicht paarweise unabhängig sind, existieren $k, m \in \mathbb{N}$ und $A, B \in \mathscr{A}_1$ mit

$$P(X_k \in A, X_m \in B) \;\neq\; P(X_k \in A) \cdot P(X_m \in B). \tag{4.6}$$

Somit folgt nun

$$\begin{aligned} P\left(X_k \in A,\ X_m \in B \mid \mathscr{I}\right) &= P\left(X_k \in A,\ X_m \in B\right) \\ &\underset{(4.6)}{\neq} P(X_k \in A) \cdot P(X_m \in B) \\ &= P(X_k \in A \mid \mathscr{I}) \cdot P(X_m \in B \mid \mathscr{I}). \end{aligned}$$

Daher folgt, dass der stochastische Prozess $(X_n)_{n \in \mathbb{N}}$ nicht bedingt paarweise unabhängig bezüglich $\mathscr{I}(X_n, n \in \mathbb{N})$ ist.

Die bedingte Unkorreliertheit kann auch an einem speziellen Beispiel der obigen Form widerlegt werden. Hierzu seien $\Omega := [0, 1[$, $\mathscr{A} := \mathscr{B}_{[0,1[}$ und $P := \lambda_{[0,1[}$.
Des Weiteren betrachte man die folgendermaßen definierte Transformation T:

$$T(\omega) \;:=\; (\omega + \mu)\,[1], \text{ für alle } \omega \in \Omega,$$

mit $\mu = 1 - \varepsilon$ für ein $\varepsilon \in]0, 1/2[\backslash\mathbb{Q}$ und den stochastischen Prozess $(X_n)_{n \in \mathbb{N}}$, der durch

$$X_n \;:=\; T^{n-1}, \text{ für alle } n \in \mathbb{N},$$

definiert wird. Dieser Prozess wird in den verschiedenen Lehrbüchern als ein Prozess von der im vorhergehenden Beispiel geforderten Form nachgewiesen (vgl. z. B. [9, Seite 113]). Die Voraussetzung (4.6) erfüllt dieser Prozess z. B. für $k = 1, m = 2$ und $A = B = [0, 1/2]$. Als Gegenbeispiel für die Frage der Unkorreliertheit betrachte man den Prozess $(Y_n)_{n \in \mathbb{N}}$ mit $Y_n := 1_A \circ X_n, n \in \mathbb{N}$.

Die Tatsache, dass alle Implikationen im verallgemeinerten Satz von de Finetti 4.5 gelten, lässt sich mit Hilfe des folgenden Sachverhaltes leicht erklären. Im Jahre 1971 wurde diese Aussage für vertauschbare Prozesse von Olshen [29] veröffentlicht.

4.15 Satz
Sei $(X_n)_{n \in \mathbb{N}}$ ein vertauschbarer Prozess, dann gilt

$$\mathscr{S}(X_n, n \in \mathbb{N}) \;\subset\; \mathscr{I}^*(X_n, n \in \mathbb{N}).$$

Beweis:
Sei $E \in \mathscr{S}$, d. h. $\left(\exists B \in \mathscr{A}_1{}^{\mathbb{N}}\right)(\forall \tau \in \tau(\mathbb{N}))\, E = \{(X_{\tau(1)}, X_{\tau(2)}, \ldots) \in B\}$. Unter Verwendung des Shift-Operators und der Definitionen der invarianten resp. f.-s.-invarianten Mengen bei maßerhaltenden Transformationen (vgl. [32, Definition 3.5.3 und Lemma 3.5.1]) ist zu zeigen:

$$P(E \triangle S^{-1}E) = 0.$$

Sei $\varepsilon > 0$ beliebig, aber fest. Dann existiert nach einem Standardergebnis der Maßtheorie (vgl. [9, Proposition 2.32]) ein $m \in \mathbb{N}$ und ein $W \in \sigma(X_1, \ldots X_m)$, d. h. $(\exists C \in \mathscr{A}_1{}^m)$ mit $W = \{(X_1, X_2, \ldots, X_m) \in C\} = \{(X_1, X_2, \ldots) \in C \times \mathscr{A}_1{}^{\mathbb{N}}\}$, mit

$$P(W \triangle E) < \varepsilon. \tag{4.7}$$

Auf ${(X_n)_{n\in\mathbb{N}}}^{-1}\left(\mathscr{A}_1{}^{\mathbb{N}}\right)$ definiere man die folgende Abbildung a_τ :

$$a_\tau\left(\{(X_1, X_2, \ldots) \in B\}\right) := \{(X_{m+1}, X_1, X_2, X_3, \ldots, X_m, X_{m+2}, \ldots) \in B\},$$

mit $B \in \mathscr{A}_1{}^{\mathbb{N}}$.
Nach Definition der Symmetrie der Menge E (vgl. Definition 2.3) gilt:

$$a_\tau(E) = E. \tag{4.8}$$

Hiermit folgt:

$$a_\tau W \triangle E = a_\tau W \triangle a_\tau E = a_\tau(W \triangle E) \tag{4.9}$$

und daher:

$$P(a_\tau W \triangle E) \underset{(4.9)}{=} P\left(a_\tau(W \triangle E)\right) = P(W \triangle E). \tag{4.10}$$

Unter Verwendung der Tatsache, dass $a_\tau W = S^{-1}W$ ist, folgt mit (4.7) und (4.10) weiter:

$$P(S^{-1}W \triangle W) = P(a_\tau W \triangle W) \leq P(a_\tau W \triangle E) + P(E \triangle W) \leq 2 \cdot \varepsilon. \tag{4.11}$$

Des Weiteren folgt aufgrund der Maßtreue von S sofort

$$P(S^{-1}W \triangle S^{-1}E) = P(S^{-1}(W \triangle E)) = P(W \triangle E). \tag{4.12}$$

Mit (4.7), (4.11) und (4.12) ergibt sich insgesamt:

$$P\left(E \triangle S^{-1}E\right) \leq P\left(E \triangle W\right) + P\left(W \triangle S^{-1}W\right) + P\left(S^{-1}W \triangle S^{-1}E\right) < 4 \cdot \varepsilon.$$

Also gilt die Behauptung. $\square$

Aus diesem Satz folgen einerseits Äquivalenzen bezüglich der Trivialität einer Unter-σ-Algebra $\mathscr{G} \subset \mathscr{A}$ und andererseits Äquivalenzen bezüglich der Unabhängigkeit der einzelnen Zufallsvariablen $X_n,\ n \in \mathbb{N}$, eines vertauschbaren Prozesses $(X_n)_{n\in\mathbb{N}}$ von einer Unter-σ-Algebra $\mathscr{G} \subset \mathscr{A}$.

Des Weiteren ergibt sich hieraus ebenfalls für vertauschbare Prozesse $(X_n)_{n\in\mathbb{N}}$ die Übereinstimmung der bedingten Wahrscheinlichkeit bezüglich $\mathscr{I}(X_n, n \in \mathbb{N})$ und der bedingten Wahrscheinlichkeit bezüglich einer σ-Algebra $\mathscr{G}$ mit $\mathscr{I} \subset \mathscr{G} \subset \mathscr{I}^*$.

4.16 Folgerung
Es sei $(X_n)_{n\in\mathbb{N}}$ ein vertauschbarer Prozess, dann gilt:

(a) Die folgenden Aussagen sind äquivalent:

- (i) $(X_n)_{n\in\mathbb{N}}$ ist ergodisch.
- (ii) $\mathscr{S}(X_n, n \in \mathbb{N})$ ist trivial.

(b) Die folgenden Aussagen sind äquivalent:

- (i) $(X_n)_{n\in\mathbb{N}}$ ist schwach ergodisch.
- (ii) Für jedes $n \in \mathbb{N}$ ist die Zufallsvariable X_n unabhängig von $\mathscr{S}(X_n, n \in \mathbb{N})$.

(c) Für jedes $B \in \mathscr{A}_1$ und jedes $n \in \mathbb{N}$ gilt:

$$P(X_n \in B \mid \mathscr{I}) = P(X_n \in B \mid \mathscr{T}) = P(X_n \in B \mid \mathscr{S}) \quad P\text{-f.s..}$$

Hieraus ergeben sich zwei neue Fragestellungen:

(I) Können die beiden Äquivalenzen aus 4.16 (a) und 4.16 (b) zu einer Äquivalenzaussage zusammengefasst werden?

(II) Gilt die Gleichheit der bedingten Verteilungen auch für stationäre Prozesse?

Während Frage (I) im Anschluss bewiesen wird, kann Frage (II) durch einn Gegenbeispiel widerlegt werden.

4.17 Lemma
Es sei $(X_n)_{n\in\mathbb{N}}$ ein stochastischer Prozess, der bedingt unabhängig bezüglich einer Unter-σ-Algebra $\mathscr{G} \subset \mathscr{A}$ ist und dessen Zufallsvariablen $X_n,\ n \in \mathbb{N}$, jeweils unabhängig von $\mathscr{G}$ sind. Dann sind die $X_n,\ n \in \mathbb{N}$, unabhängig.

Beweis:
Nach der Definition der Unabhängigkeit von Zufallsvariablen ist für jedes $n \in \mathbb{N}$ und jede Auswahl $B_i \in \mathscr{A}_1,\ 1 \leq i \leq n$, zu zeigen:

$$P\left(\bigcap_{i=1}^{n} \{X_i \in B_i\}\right) = \prod_{i=1}^{n} P\left(X_i \in B_i\right). \tag{4.13}$$

Mit obigen Mengen gilt nun für $n \in \mathbb{N}$:

$$\begin{aligned}
P\left(\bigcap_{i=1}^{n} \{X_i \in B_i\}\right) &= P\left(\bigcap_{i=1}^{n} \{X_i \in B_i\} \middle| \{\emptyset, \Omega\}\right) \\
&\underset{A.3}{=} P\left(P\left(\bigcap_{i=1}^{n} \{X_i \in B_i\} \middle| \mathscr{G}\right) \middle| \{\emptyset, \Omega\}\right) \\
&\underset{\text{bed. Unabhängigkeit}}{=} P\left(\prod_{i=1}^{n} P\left(X_i \in B_i \mid \mathscr{G}\right) \middle| \{\emptyset, \Omega\}\right) \\
&\underset{\text{Unabhängigkeit von } \mathscr{G}}{=} P\left(\prod_{i=1}^{n} P\left(X_i \in B_i\right) \middle| \{\emptyset, \Omega\}\right) \\
&= P\left(\prod_{i=1}^{n} P\left(X_i \in B_i\right)\right) \\
&= \prod_{i=1}^{n} P\left(X_i \in B_i\right).
\end{aligned}$$

Also gilt (4.13). □

Somit ergibt sich für vertauschbare Prozesse der folgende Sachverhalt.

4.18 Satz
Ist $(X_n)_{n\in\mathbb{N}}$ ein vertauschbarer Prozess, dann sind äquivalent:

(i) $(X_n)_{n\in\mathbb{N}}$ ist schwach ergodisch.

(ii) $(X_n)_{n\in\mathbb{N}}$ ist unabhängig und identisch verteilt.

Beweis:
Nach dem Satz von de Finetti (Satz 4.3 resp. Satz 4.4) ist der vertauschbare Prozess $(X_n)_{n\in\mathbb{N}}$ bedingt unabhängig bezüglich $\mathscr{I} := \mathscr{I}(X_n, n \in \mathbb{N})$. Des Weiteren sind die Zufallsvariablen $X_n,\ n \in \mathbb{N}$, jeweils unabhängig von $\mathscr{I}$ nach Definition der schwachen Ergodizität. Also sind die Voraussetzungen des Lemmas 4.17 erfüllt und somit ist $(X_n)_{n\in\mathbb{N}}$ ein Prozess aus unabhängigen Zufallsvariablen. Dieser ist aufgrund der Vertauschbarkeit auch identisch verteilt. □

Daher folgt insbesondere im Zusammenhang mit ergodischen Prozessen das folgende Korollar.

4.19 Korollar
Für einen vertauschbaren Prozess $(X_n)_{n\in\mathbb{N}}$ gilt die folgende Äquivalenz:

(i) $(X_n)_{n\in\mathbb{N}}$ ist schwach ergodisch.

(ii) $(X_n)_{n\in\mathbb{N}}$ ist ergodisch (und somit ist auch $\mathscr{S}(X_n, n \in \mathbb{N})$ trivial).

Beweis:
Nach Lemma 4.17 folgt sofort, dass $(X_n)_{n\in\mathbb{N}}$ unabhängig und identisch verteilt ist. Somit folgt nach dem Satz von Hewitt und Savage [3, Theorem 11.6] die Trivialität der symmetrischen Mengen $\mathscr{S}(X_n, n \in \mathbb{N}) \supset \mathscr{I}(X_n, n \in \mathbb{N})$, also ist $(X_n)_{n\in\mathbb{N}}$ ergodisch.
Die Rückrichtung folgt trivialerweise. □

Nun soll das bereits angekündigte Gegenbeispiel für die zweite Fragestellung, die durch Folgerung 4.16 (c) aufgeworfen wurde, angegeben werden.

4.20 Beispiel
Man betrachte den folgenden Raum:

$$\Omega \ := \ \{-1, 1\}^{\mathbb{N}} \text{ und } \mathscr{A} \ := \ \mathfrak{P}(\Omega).$$

Auf $(\Omega, \mathscr{A})$ definiere man das Wahrscheinlichkeitsmaß P durch

$$P(1, -1, 1, -1, 1, \dots) \ = \ P(-1, 1, -1, 1, -1, \dots) \ = \ 1/2$$

und betrachte die kanonischen Koordinatenabbildungen $(\pi_n)_{n\in\mathbb{N}}$ als stochastischen Prozess.

Der Prozess $(\pi_n)_{n\in\mathbb{N}}$ ist auf dem Wahrscheinlichkeitsraum $(\Omega, \mathscr{A}, P)$ offensichtlich stationär und $\mathscr{I}(\pi_n, n \in \mathbb{N})$ ist trivial, denn es gilt für jede invariante Menge $I \in \mathscr{I}(\pi_n, n \in \mathbb{N})$

$$(1, -1, 1, -1, \dots) \in I \ \Leftrightarrow \ (-1, 1, -1, 1, \dots) \in I.$$

Also folgt $P(I) \in \{0, 1\}$ für jedes $I \in \mathscr{I}(\pi_n, n \in \mathbb{N})$, somit ist $(\pi_n)_{n\in\mathbb{N}}$ ergodisch. Nach Bemerkung A.2 c) folgt aus der Ergodizität des hier betrachteten Prozesses sofort:

$$P\left(\pi_k \mid \mathscr{I}(\pi_n, n \in \mathbb{N})\right) \ = \ P(\pi_k) \quad (\forall k \in \mathbb{N}). \tag{4.14}$$

Seien

$$A_1 \;:=\; \left\{\omega \in \Omega \,\middle|\, \lim_{n\to\infty} X_{2n} = 1\right\} \quad \text{und} \quad A_{-1} \;:=\; \left\{\omega \in \Omega \,\middle|\, \lim_{n\to\infty} X_{2n} = -1 \right\},$$

also $(-1, 1, -1, 1, -1, 1, \ldots) \in A_1$ und $(1, -1, 1, -1, 1, -1, \ldots) \in A_{-1}$, so gilt:

$$A_1, A_{-1}, \Omega \;\in\; \mathscr{T}(\pi_n, n \in \mathbb{N}). \tag{4.15}$$

Für $a \in \{-1, 1\}$ und $k \in \mathbb{N}$ gilt weiter:

$$P\left({\pi_k}^{-1}(a) \triangle A_1\right) = 0 \quad \vee \quad P\left({\pi_k}^{-1}(a) \triangle A_{-1}\right) = 0$$

Also folgt nach (4.15), der Definition der Vervollständigung einer σ-Algebra 4.7 und der Messbarkeit einer Abbildung für $k \in \mathbb{N}$:

$$\pi_k \text{ ist } \mathscr{T}^*(\pi_n, n \in \mathbb{N})\text{- messbar}$$

und daher nach Bemerkung A.2 auch:

$$P\left(\pi_k \mid \mathscr{T}^*(\pi_n, n \in \mathbb{N})\right) \;=\; \pi_k \quad P\text{-f.s.}. \tag{4.16}$$

Aus (4.14) und (4.16) folgt nun:

$$P\left(\pi_k \mid \mathscr{T}^*(\pi_n, n \in \mathbb{N})\right) \;=\; \pi_k \;\neq P(\pi_k) \;=\; P\left(\pi_k \mid \mathscr{I}(\pi_n, n \in \mathbb{N})\right) \quad P\text{-f.s.};$$

also stimmen für den allgemeinen stationären Fall die bedingten Erwartungen bezüglich $\mathscr{I}(X_n, n \in \mathbb{N})$ und bezüglich $\mathscr{T}^*(X_n, n \in \mathbb{N})$ nicht notwendigerweise überein. Daher können auch die bedingten Verteilungen nicht übereinstimmen und es ist ein Gegenbeispiel gezeigt.

Dieses Beispiel zeigt auch, dass ein stationärer Prozess nicht bedingt identisch verteilt bezüglich der terminalen Mengen eines stationären Prozesses sein muss:

Aus der Definition des Maßes P folgt:

$$\pi_1 \;\neq \pi_2 \quad P\text{-f.s.}.$$

Kombiniert man diese Tatsache mit (4.16), so folgt P-f.s.:

$$P\left(\pi_1 \mid \mathscr{T}^*(\pi_n, n \in \mathbb{N})\right) \;=\; \pi_1 \;\neq\; \pi_2 \;=\; P\left(\pi_2 \mid \mathscr{T}^*(\pi_n, n \in \mathbb{N})\right). \tag{4.17}$$

Somit stimmen die bedingten Erwartungen bezüglich der fast-sicher terminalen Mengen des Prozesses $(\pi_n)_{n\in\mathbb{N}}$ nicht überein, also ist $(\pi_n)_{n\in\mathbb{N}}$ nicht bedingt identisch verteilt bezüglich $\mathscr{T}^*(\pi_n, n \in \mathbb{N})$. Nach Lemma 4.9 ist $(\pi_n)_{n\in\mathbb{N}}$ auch nicht bedingt identisch verteilt bezüglich $\mathscr{T}(\pi_n, n \in \mathbb{N})$.

Somit ist auch diese Behauptung durch ein Gegenbeispiel widerlegt.

Die Übereinstimmung der bedingten Verteilungen eines Prozesses bezüglich verschiedener σ-Algebren kann auch im Falle stationärer Prozesse vorliegen. Hierzu wird jedoch noch eine weitere Eigenschaft benötigt. Mit Hilfe des daraus resultierenden Ergebnisses kann dann ein erstes Ergebnis für die Erweiterung der schwachen Ergodizität auf die Unabhängigkeit von $\mathscr{I}$-umfassenderen σ-Algebren gegeben werden.

4.21 Satz

Sei $(X_n)_{n\in\mathbb{N}}$ ein reellwertiger, stationärer Prozess, der bedingt identisch verteilt bezüglich einer σ-Algebra

$$\mathscr{G} \ \supset \ \mathscr{I}(X_n, n \in \mathbb{N}) =: \mathscr{I} \tag{4.18}$$

ist. Dann stimmen die bedingten Verteilungen für jede ZV X_n, $n \in \mathbb{N}$, bezüglich $\mathscr{G}$ und bezüglich $\mathscr{I}$ überein, d. h. es gilt für alle $B \in \mathscr{B}$:

$$P(X_n \in B \mid \mathscr{G}) \ = \ P(X_n \in B \mid \mathscr{I}) \ \ P\text{-f.s.}. \tag{4.19}$$

Beweis:

Bedingt durch die Voraussetzungen dieses Satzes reicht es, (4.19) für $n = 1$ zu zeigen.

Angesichts der bedingt identischen Verteilung des Prozesses $(X_n)_{n\in\mathbb{N}}$ bezüglich $\mathscr{G}$ folgt für jede messbare Funktion $\psi : \mathbb{R} \to \mathbb{R}$:

$$P\left(\psi \circ X_n \mid \mathscr{G}\right) \ = \ P\left(\psi \circ X_1 \mid \mathscr{G}\right) \quad (\forall n \in \mathbb{N}). \tag{4.20}$$

Sei nun $\varphi : \mathbb{R} \to \mathbb{R}$ eine injektive, messbare Funktion, so gilt nach Lemma 3.1:

$$\mathscr{I}(X_n, n \in \mathbb{N}) \ = \ \mathscr{I}(\varphi \circ X_n, n \in \mathbb{N}). \tag{4.21}$$

Nach dem Birkhoff'schen Ergodensatz [6] folgt für beschränktes φ wegen (4.21):

$$\frac{1}{n} \cdot \sum_{k=1}^{n} \varphi \circ X_k \ \underset{\substack{k\to\infty \\ \text{im ersten Mittel}}}{\longrightarrow} \ P\left(\varphi \circ X_1 \mid \mathscr{I}(X_n, n \in \mathbb{N})\right).$$

Da $P\left(\varphi \circ X_1 \mid \mathscr{I}(X_n, n \in \mathbb{N})\right)$ nach (4.18) auch $\mathscr{G}$-messbar ist, erhält man unter Beachtung von Lemma A.3 und Bemerkung A.2:

$$\begin{aligned}
P\left(\varphi \circ X_1 \mid \mathscr{G}\right) \ &\underset{(4.20)}{=} \ P\left(\frac{1}{n} \cdot \sum_{k=1}^{n} \varphi \circ X_k \,\middle|\, \mathscr{G}\right) \quad P\text{-f.s.} \\
&\underset{\text{Lemma A.3 (a)}}{\longrightarrow} \ P\left(P\left(\varphi \circ X_1 \mid \mathscr{I}(X_n, n \in \mathbb{N})\right) \mid \mathscr{G}\right) \quad P\text{-f.s.} \\
&\underset{\text{Bemerkung A.2}}{=} \ P\left(\varphi \circ X_1 \mid \mathscr{I}(X_n, n \in \mathbb{N})\right) \quad P\text{-f.s.}.
\end{aligned}$$

Daher folgt:

$$P(\varphi \circ X_1 \mid \mathscr{G}) \;=\; P(\varphi \circ X_1 \mid \mathscr{I}(X_n, n \in \mathbb{N})) \quad P\text{-f.s.}. \tag{4.22}$$

Sei nun $x \in \mathbb{R}$ beliebig, aber fest gewählt. Dann existiert eine Folge $(\varphi_n)_{n\in\mathbb{N}}$ von injektiven, messbaren und beschränkten Funktionen, die punktweise gegen die Funktion $\varphi_0 := 1_{]-\infty,x]}$ konvergiert. Unter Beachtung des Satzes über dominierende Konvergenz für bedingte Wahrscheinlichkeiten folgt aus (4.22) die Behauptung (4.19).

□

In Anlehnung an den verallgemeinerten Satz von de Finetti 4.4 ergibt sich aus dem letzten Satz als Anwendung auf die terminale und die symmetrischen Mengen der folgende Satz.

4.22 Satz
Sei $(X_n)_{n\in\mathbb{N}}$ ein reellwertiger, stationärer Prozess, der bedingt identisch verteilt ist bezüglich $\mathscr{S}(X_n, n \in \mathbb{N})$ (resp. $\mathscr{T}(X_n, n \in \mathbb{N})$), dann stimmen die bedingten Wahrscheinlichkeiten von X_1 bezüglich $\mathscr{I}(X_n, n \in \mathbb{N})$ und bezüglich $\mathscr{S}(X_n, n \in \mathbb{N})$ (resp. $\mathscr{T}(X_n, n \in \mathbb{N})$) überein.

Aus diesem Satz folgt für reellwertige, schwach ergodische Prozesse insbesondere die nächste Aussage:

4.23 Folgerung
Sei $(X_n)_{n\in\mathbb{N}}$ ein reellwertiger, schwach ergodischer Prozess.
Ist der Prozess$(X_n)_{n\in\mathbb{N}}$ bedingt identisch verteilt bezüglich der symmetrischen Mengen $\mathscr{S}(X_n, n \in \mathbb{N})$ (resp. bezüglich der terminalen Mengen $\mathscr{T}(X_n, n \in \mathbb{N})$), dann ist jedes $X_n, n \in \mathbb{N}$, unabhängig von $\mathscr{S}(X_n, n \in \mathbb{N})$ (resp. $\mathscr{T}(X_n, n \in \mathbb{N})$).

§ 5

Darstellung schwach ergodischer Prozesse

Nachdem in den bisherigen Paragraphen dieser Arbeit Eigenschaften und Erweiterungen des Begriffs der schwachen Ergodizität stationäre Prozesse untersucht wurden, soll nun konkret eine Darstellung der Klasse der schwach ergodischen Prozesse betrachtet werden. Hierbei werden insbesondere die Vorgehensweise und die Beweistechniken aus dem Artikel von Landers und Rogge [27] untersucht und ausgewertet.

Im weiteren Verlauf wird von stochastischen Prozessen $(X_n)_{n\in\mathbb{N}}$ mit Werten in einem separablem metrischen Raum M die Rede sein. Dieser separable metrische Raum kann als Unterraum eines überabzählbaren, vollständigen und separablen metrischen Raumes angesehen werden. Daher ist es o.B.d.A. möglich, M als überabzählbaren, vollständigen, separablen metrischen Raum anzusehen. Nach dem Buch von Cohn zur Maßtheorie [13, Satz 8.3.6] existiert ein Borel-Isomorphismus $\psi_M : M \to \mathbb{R}$, d. h. eine bijektive Abbildung von M nach $\mathbb{R}$, für die ψ_M und ${\psi_M}^{-1}$ messbar sind. Wird also im Folgenden von einem Borel-Isomorphismus ψ_M gesprochen, so ist stets dieser eben angesprochene Isomorphismus gemeint.

Um auch Aussagen über die Unabhängigkeit der einzelnen Zufallsvariablen eines stochastischen Prozesses von z. B. der symmetrischen σ-Algebra des entsprechenden Prozesses machen zu können, sollen die nun vorgestellten Betrachtungen auf allgemeiner Basis durchgeführt werden und können anschließend in die Spezialfälle überführt werden. Aus diesem Grund werden jetzt erneut einige Begriffe definiert, die in den vorherigen Paragraphen schon für Spezialfälle angegeben worden sind.

Für die weiteren Betrachtungen in diesem Paragraphen seien $(X_n)_{n\in\mathbb{N}}$ ein stochastischer Prozess mit $X_n : (\Omega, \mathscr{A}, P) \to (M, \mathscr{B})$, $n \in \mathbb{N}$, $\mathcal{F} \subset \Omega^\Omega$ eine Menge von messbaren Funktionen und es sei die σ-Algebra $\mathscr{F}(X_n,\ n \in \mathbb{N}) \subset \mathscr{A}$ wie folgt mit $\mathcal{F}' := \mathcal{F} \cup \{id_\Omega\}$ definiert:

$$\begin{aligned}
&\mathscr{F}(X_n, n \in \mathbb{N}) \\
&\quad := \left\{A \in \mathscr{A} \ :\ \left(\exists B \in \mathscr{B}^{\mathbb{N}}\right) (\forall f \in \mathcal{F}')\, A \ = \ \{(X_1 \circ f, X_2 \circ f, \ldots) \in B\}\right\} \qquad (5.1)
\end{aligned}$$

Für den Fall, dass klar ist welcher Prozess $(X_n)_{n\in\mathbb{N}}$ gemeint ist, wird nur $\mathscr{F}$ anstelle von $\mathscr{F}(X_n, n \in \mathbb{N})$ geschrieben.

Mit dieser Definition ergibt sich die folgende Aussage, die als Äquivalenz zu Lemma 3.1 angesehen werden kann.

5.1 Lemma
Seien $M,\ M_1$ separable metrische Räume und $(X_n)_{n\in\mathbb{N}}$ ein stochastischer Prozess mit Werten in M.

(i) Ist $\psi : M \to M_1$ Borel-messbar, dann gilt:

$$\mathscr{F}(\psi \circ X_n,\ n \in \mathbb{N}) \ \subset \ \mathscr{F}(X_n,\ n \in \mathbb{N}).$$

(ii) Ist $\varphi : M \to \mathbb{R}$ injektiv und Borel-messbar und ist M vollständig, so gilt:

$$\mathscr{F}(\varphi \circ X_n,\ n \in \mathbb{N}) \ = \ \mathscr{F}(X_n,\ n \in \mathbb{N}).$$

Beweis:

Zu (i): Man setze auf $M^{\mathbb{N}}$:

$$\Psi \circ (x_1, x_2, \ldots) \ := \ (\psi \circ x_1, \psi \circ x_2, \ldots),$$

dann ist Ψ messbar. Somit gilt für $F \in \mathscr{F}(\psi \circ X_n, n \in \mathbb{N})$ nach (5.1) für entsprechendes $B \in \mathscr{B}^{\mathbb{N}}$ und für alle $f \in \mathcal{F}'$:

$$\begin{aligned}
F \ &= \ \{((\psi \circ X_1) \circ f, (\psi \circ X_2) \circ f, \ldots) \in B\} \\
&= \ \{\Psi \circ (X_1 \circ f, X_2 \circ f, \ldots) \in B\} \\
&= \ \{(X_1 \circ f, X_2 \circ f, \ldots) \in \Psi^{-1}B\}.
\end{aligned}$$

Also gilt $F \in \mathscr{F}(X_n, n \in \mathbb{N})$.

Zu (ii): Entsprechend des Beweises zu Lemma 3.1 verläuft auch hier der Beweis durch zweifache Anwendung von (i), was durch die spezielle Wahl der Voraussetzungen möglich ist. $\square$

5.2 Definition

Eine Zufallsvariable Y heißt *$\mathcal{F}$-Zufallsvariable* (bezüglich $(X_n)_{n\in\mathbb{N}}$), wenn es eine messbare Funktion φ gibt, so dass für jedes $f \in \mathcal{F}$ gilt:

$$Y \;=\; \varphi(X_1 \circ f, X_2 \circ f, \ldots).$$

5.3 Lemma

Eine reellwertige Zufallsvariable Y ist genau dann eine $\mathcal{F}$-Zufallsvariable, wenn sie $\mathscr{F}(X_n, n \in \mathbb{N})$-messbar ist.

5.4 Definition

Ein stochastischer Prozess $(Y_n)_{n\in\mathbb{N}}$ heißt *$\mathcal{F}$-Prozess*, wenn für jedes $B \in \mathscr{B}^{\mathbb{N}}$ und jedes $f \in \mathcal{F}$ gilt:

$$P\left((Y_1, Y_2, \ldots) \in B\right) \;=\; P\left((Y_1 \circ f, Y_2 \circ f, \ldots) \in B\right).$$

5.5 Bemerkung

i) Sei $(Y_n)_{n\in\mathbb{N}}$ ein $\mathcal{F}$-Prozess, dann sind alle $f \in \mathcal{F}$ bezüglich $P\,|_{\mathscr{F}(X_n, n\in\mathbb{N})}$ maßerhaltende Funktionen.

ii) Ist $(Y_n)_{n\in\mathbb{N}}$ ein $\mathcal{F}$-Prozess, dann ist $(\phi \circ Y_n)_{n\in\mathbb{N}}$ für jede messbare Funktion $\phi : (M, \mathscr{B}) \to (\Omega_2, \mathscr{A}_2)$ ein $\mathcal{F}$-Prozess.

Der folgende Satz ist dem Buch von Ash und Gardner [2, Satz 3.2.8] entnommen.

5.6 Satz

Sei T eine maßerhaltende Transformation, $g : (\Omega, \mathscr{A}) \to (\mathbb{R}, \mathscr{B})$ P-integrierbar, dann gilt für jedes $A \in \mathscr{A}$:

$$\int_{T^{-1}A} g(T\omega)\,\mathrm{d}P(\omega) \;=\; \int_A g(\omega)\,\mathrm{d}P(\omega).$$

Unter Beachtung von Bemerkung 5.5 erhält man folgenden Sachverhalt.

5.7 Lemma

Sei $(Y_n)_{n\in\mathbb{N}}$ ein $\mathcal{F}$-Prozess, dann gilt für jedes $B \in \mathscr{F}(Y_n, n \in \mathbb{N})$ und jedes $f \in \mathcal{F}$

$$\int_B Y_1(\omega)\; dP \;=\; \int_B (Y_1 \circ f)\;(\omega)\; dP.$$

Der nun folgende Satz ist eine Verallgemeinerung des bekannten Ergodensatzes von Birkhoff (vgl. [6] bzw. [2, Satz 3.5.6]).

5.8 Satz

Sei $(X_n)_{n\in\mathbb{N}}$ ein $\mathcal{F}$-Prozess mit $P(|X_1|) < \infty$, für den

(i) $P\left(X_1 \mid \mathscr{F}(X_n, n \in \mathbb{N})\right) \;=\; P\left(X_n \mid \mathscr{F}(X_n, n \in \mathbb{N})\right) \quad (\forall n \in \mathbb{N})$

(ii) $\frac{1}{n} \cdot \sum_{k=1}^{n} X_k$ konvergiert in $\mathscr{L}^1$ und P-f.s. und

(iii) $\limsup_{n\to\infty} \left(\frac{1}{n} \cdot \sum_{k=1}^{n} X_k\right)$ ist eine $\mathcal{F}$-Zufallsvariable

gilt, dann folgt

$$\lim_{n\to\infty} \frac{1}{n} \cdot \sum_{k=1}^{n} X_k \;=\; P(X_1 \mid \mathscr{F}(X_n, n \in \mathbb{N})) \quad P\text{-f.s.}.$$

Beweis:

Nach Voraussetzung (ii) konvergiert $\frac{1}{n} \cdot \sum_{k=1}^{n} X_k$ o.B.d.A. P-f.s. gegen

$$Y := \limsup_{n\to\infty} \left(\frac{1}{n} \cdot \sum_{k=1}^{n} X_k\right).$$

Es ist daher zu zeigen:

$$Y \;=\; P\left(X_1 \mid \mathscr{F}(X_n, n \in \mathbb{N})\right) \; P\text{-f.s.}. \tag{5.2}$$

Hierzu genügt es nach der Definition der bedingten Erwartung A.1, den Nachweis der folgenden Eigenschaften zu führen:

(a) Y ist $\mathscr{F}(X_n, n \in \mathbb{N})$-messbar und

(b) $P\left(X_1 \mid \mathscr{F}(X_n, n \in \mathbb{N})\right) = Y \quad P$-f.s..

Zu (a): Nach Voraussetzung (iii) ist Y eine $\mathcal{F}$-Zufallsvariable. Also ist Y nach Lemma 5.3 $\mathscr{F}(X_n, n \in \mathbb{N})$-messbar.

Zu (b): Aufgrund der nach (ii) und (5.2) gültigen $\mathscr{L}^1$-Konvergenz von $\frac{1}{n} \cdot \sum_{k=1}^{n} X_k$ gegen Y folgt:

$$\begin{aligned}
P\left(X_1 \mid \mathscr{F}(X_n, n \in \mathbb{N})\right) \quad &\underset{(i)}{=} \quad P\left(\frac{1}{n} \cdot \sum_{k=1}^{n} X_k \,\middle|\, \mathscr{F}(X_n, n \in \mathbb{N})\right) \\
&\underset{\substack{\|\cdot\|_1 \\ n\to\infty}}{\longrightarrow} \quad P\left(Y \mid \mathscr{F}(X_n, n \in \mathbb{N})\right) \\
&\underset{(a)}{=} \quad Y.
\end{aligned}$$

Also folgt (b).
Somit gilt insgesamt die Behauptung. □

Bevor ein weiteres Lemma angegeben wird, sei zur Erläuterung eine kurze Bemerkung gemacht:

5.9 Bemerkung
Sei $x \in \mathbb{R}$ gewählt, dann gibt es eine Folge $(\varphi_n^x)_{n\in\mathbb{N}}$ von $]0,1[$-wertigen, messbaren und injektiven Funktionen, die punktweise gegen $1_{]-\infty,x]}$ konvergiert.
Hier sei ein konkretes Beispiel angegeben. Man wähle ein beliebiges $\varepsilon \in \mathbb{R}_+$ und setze

$$\varphi_n^x(t) \;:=\; \begin{cases} \frac{1}{2} + \frac{1}{\pi} \cdot \arctan\left(n\left(x - t + \varepsilon\right)\right) & \text{, falls } t \in\,]-\infty, x], \\ \frac{1}{2} + \frac{1}{\pi} \cdot \arctan\left(n\left(x - t\right)\right) & \text{, falls } t \in\,]x, \infty[. \end{cases}$$

5.10 Lemma
Es sei $(X_n)_{n\in\mathbb{N}}$ ein reellwertiger, stochastischer Prozess auf $(\Omega, \mathscr{A}, P)$. Es seien weiter für beliebiges $x \in \mathbb{R}$ $\varphi^x := 1_{]-\infty,x]}$ und $(\varphi_m^x)_{m\in\mathbb{N}}$ eine Folge von $]0,1[$-wertigen, messbaren injektiven Funktionen, die punktweise gegen φ^x konvergiert. Gilt dann für alle $m \in \mathbb{N}$ mit einer σ-Algebra $\mathscr{G} \subset \mathscr{A}$:

$$P\left(\varphi_m^x \circ X_n\right) \;=\; P\left(\varphi_m^x \circ X_n \mid \mathscr{G}\right) \quad (\forall n \in \mathbb{N}),$$

dann ist jedes $X_n, n \in \mathbb{N}$, unabhängig von $\mathscr{G}$.

Beweis:
Der Beweis beruht auf der Anwendung des Satzes über die majorisierende Konvergenz für Wahrscheinlichkeiten und dem entsprechenden Satz für bedingte Wahscheinlichkeiten. Nach deren Anwendung folgt sofort:

$$P\left(\varphi^x \circ X_n\right) \;=\; P\left(\varphi^x \circ X_n \mid \mathscr{G}\right) \quad (\forall n \in \mathbb{N}),$$

also nach A.2 (e) die Unabhängigkeit von X_n und $\mathscr{G}$. □

Hiermit kommt man zu der folgenden Aussage.

5.11 Satz
Sei $(X_n)_{n\in\mathbb{N}}$ ein $\mathcal{F}$-Prozess mit Werten in einem separablen metrischen Raum M. Des Weiteren sei $\psi_M : M \to \mathbb{R}$ ein Borel-Isomorphismus.
Gilt dann für jedes $x \in \mathbb{R}$, $m \in \mathbb{N}$ und jeweils eine injektive, messbare und $]0,1[$-wertige Funktionenfolge $(\varphi_n^x)_{n\in\mathbb{N}}$, die punktweise gegen $1_{]-\infty,x]}$ konvergiert:

i) $\frac{1}{n} \cdot \sum_{k=1}^{n} (\varphi_m^x \circ \psi_M \circ X_k) \underset{n \to \infty}{\longrightarrow} P(\varphi_m^x \circ \psi_M \circ X_1 \mid \mathscr{F})$ P-f.s.,

ii) $P\left(\varphi_m^x \circ \psi_M \circ X_1 \mid \mathscr{F}\right) \; = \; P\left(\varphi_m^x \circ \psi_M \circ X_n \mid \mathscr{F}\right) \quad (\forall n \in \mathbb{N}),$

iii) $P_{X_1} \; = \; P_{X_n} \quad (\forall n \in \mathbb{N}),$

iv) die Prozesse $(\varphi_m^x \circ \psi_M \circ X_n)_{n \in \mathbb{N}}$ erfüllen das starke Gesetz der großen Zahlen (SLLN).

Dann ist jedes X_n, $n \in \mathbb{N}$, unabhängig von $\mathscr{F}((X_n),\ n \in \mathbb{N})$.

Beweis:
Dieser Beweis ist eine Verallgemeinerung des Beweises von Landers und Rogge [27], in dem die gleiche Methodik verwendet wurde.

Wie schon zu Beginn dieses Paragraphen erwähnt wurde, handelt es sich bei dem in den Voraussetzungen erwähnten Borel-Isomorphismus ψ_M um die von M abhängige messbare Abbildung von $M \to \mathbb{R}$, für die auch ${\psi_M}^{-1}$ messbar ist. Unter Beachtung dieser Tatsache folgt, dass es sich bei $(Y_n)_{n \in \mathbb{N}}$ mit $Y_n := \psi_M \circ X_n,\ n \in \mathbb{N}$, wieder um einen $\mathcal{F}$-Prozess handelt, der den Voraussetzungen dieses Satzes genügt.

Nach Lemma 5.1 gilt

$$\mathscr{F}(Y_n,\ n \in \mathbb{N}) \; = \; \mathscr{F}(X_n,\ n \in \mathbb{N}).$$

Somit reicht es, die Unabhängigkeit jedes $Y_n,\ n \in \mathbb{N}$ und $\mathscr{F}(Y_n,\ n \in \mathbb{N})$ zu zeigen.

Seien nun $n_0 \in \mathbb{N}$ und $x \in \mathbb{R}$ beliebig, aber fest gewählt. Für die gewünschte Unabhängigkeit genügt es zu zeigen, dass

$$\{Y_{n_0} < x\} \text{ und } \mathscr{F}(Y_n,\ n \in \mathbb{N}) \tag{5.3}$$

unabhängig sind.

Die in diesem Satz geforderten Voraussetzungen erfüllen aber gerade die Voraussetzungen von Lemma 5.10. Somit folgt die Behauptung dieses Satzes aus der Anwendung des Lemmas 5.10 auf jedes $Y_n,\ n \in \mathbb{N}$. □

Nachdem nun ein Ergebnis auf einer allgemeiner Basis erreicht wurde, soll die Überührung in die Spezialfälle ermöglicht werden. Zu diesem Zweck sollen die Mengen $\mathcal{F}$ für verschiedene Fälle angegeben werden.

Betrachtet man die Unabhängigkeit der einzelnen Zufallsvariablen $X_n, n \in \mathbb{N}$, von den invarianten Mengen $\mathscr{I}$ des stationären Prozesses $(X_n)_{n \in \mathbb{N}}$, so kann für den Fall, dass sich der Prozess aus einer ZV X und einer maßerhaltenden Transformation T generieren lässt, die Menge $\mathcal{F}$ in folgender Weise dargestellt werden:

$$\mathcal{F}_{\mathscr{I}}^{T} \; = \; \{T^i \mid i \in \mathbb{N}\}.$$

Im allgemeinen Fall eines stationären Prozesses muss auf eine andere Konstruktion zurückgegriffen werden, die auch für die Unabhängigkeit von den symmetrischen Mengen von Bedeutung ist. Hierzu sei der stochastische Prozess $(Y_n)_{n\in\mathbb{N}}$ mit $Y_n : (\Omega_1, \mathscr{A}_1, P_1) \to (M, \mathscr{B})$, $n \in \mathbb{N}$ und $Y := (Y_1, Y_2, Y_3, \ldots)$ gegeben. Dann folgt mit $(\Omega, \mathscr{A}, P) := ({\Omega_1}^{\mathbb{N}}, {\mathscr{A}_1}^{\mathbb{N}}, P_Y)$ und dem stochastischen Prozess $(X_n)_{n\in\mathbb{N}} := (\pi_n)_{n\in\mathbb{N}}$ bestehend aus den Koordinatenabbildungen die folgende Darstellung für die Menge $\mathcal{F}$:

$$\mathcal{F}_{\mathscr{I}} \;=\; \{S^i \mid i \in \mathbb{N}\},$$

wobei S der Shift-Operator aus Definition 2.4 ist.

Für die Unabhängigkeit von den symmetrischen Mengen betrachte man ebenfalls obigen Ansatz mit folgender Menge $\mathcal{F}$:

$$\mathcal{F}_{\mathscr{S}} \;=\; \{f_\tau \mid \tau \in \tau(\mathbb{N})\},$$

mit $f_\tau(\omega_1, \omega_2, \ldots) = (\omega_{\tau(1)}, \omega_{\tau(2)}, \ldots)$.

Der folgende Paragraph beschäftigt sich mit der konkreten Angabe von stochastischen Prozessen, deren einzelnen Zufallsvariablen von speziellen σ-Algebren unabhängig sind. In diesem Zusammenhang wird dann auch der schon oft zitierte und für diese Arbeit grundlegende Satz aus dem Artikel von Landers und Rogge [27] notiert.

Wie die bisherigen Betrachtungen dieser Ausarbeitung aufzeigen, werden vorrangig die Unabhängigkeit von den invarianten und den symmetrischen Mengen eines stochastischen Prozesses behandelt, aber es werden auch explizite Beispiele für die Unabhängigkeit von den terminalen Mengen bzw. der Trivialität der terminalen resp. symmetrischen Mengen gegeben.

§ 6

Beispiele schwach ergodischer Prozesse

Der zweite Teil dieser Arbeit befasst sich mit Beispielen für schwach ergodische Prozesse bzw. mit Anwendungen der bisherigen Ergebnisse über schwach ergodischer Prozesse.

Der vorliegende Paragraph beschäftigt sich insbesondere mit

- dem Ergebnis von Landers und Rogge, dessen Aussage die schwache Ergodizität eines paarweise unabhängigen, stationären Prozesses ist.
- der Erweiterung der Aussage von Landers und Rogge auf $\mathscr{I}$-umfassende σ-Algebren, insbesondere auf die σ-Algebra der symmetrischen Mengen.
- Betrachtungen im Umfeld von 2-vertauschbaren Prozessen, die ebenfalls als Kandidaten für schwach ergodische Prozesse in Frage kommen könnten. Diese Vermutung wird sich jedoch als schon bekannt erweisen.
- der Frage nach der schwachen Ergodizität von nicht-negativen stochastischen Prozessen. Für diese existiert ein starkes Gesetz der großen Zahlen, das von Etemadi im Jahre 1983 [19] veröffentlicht wurde.
- *-mischenden Prozessen, die unter bestimmten Voraussetzungen als schwach ergodisch nachgewiesen werden können. Ein besonderer Vertreter dieser Klasse ist die Menge der *m-abhängigen Prozesse*, bei denen neben der Trivialität der terminalen Mengen auch die Trivialität der symmetrischen Mengen unter Hinzunahme der Vertauschbarkeit des stochastischen Prozesses nachgewiesen werden kann.

Zu Beginn der weiteren Betrachtungen wird nun der Satz von Landers und Rogge [27] angegeben, der die eigentliche Grundlage dieser Arbeit bildet.

6.1 Satz
Sei $(X_n)_{n\in\mathbb{N}}$ ein stationärer Prozess von paarweise unabhängigen Zufallsvariablen mit Werten in einem separablen metrischen Raum. Dann ist $(X_n)_{n\in\mathbb{N}}$ schwach ergodisch, d. h. jedes $X_n,\ n \in \mathbb{N}$, ist unabhängig von $\mathscr{I}(X_n,\ n \in \mathbb{N})$.

Im entsprechenden Beweis wird die Stationarität des Prozesses $(X_n)_{n\in\mathbb{N}}$ für die Gültigkeit des Birkhoff'schen Ergodensatzes benutzt. Ferner wird für jede messbare, injektive Funktion $\varphi : \mathbb{R} \to]0,1[$ das starke Gesetz der großen Zahlen (SLLN) angewandt:

$$\frac{1}{n} \cdot \sum_{i=1}^{n} \varphi(X_i) \ \to\ P(\varphi(X_1)) \quad P\text{-f.s.}. \tag{6.1}$$

Ist der Prozess also stationär, so ist aufgrund von Satz 3.3 für die schwache Ergodizität gerade die „letztere“ Eigenschaft (6.1) nachzuweisen.

Verschärft man in Satz 6.1 die Stationarität zur Vertauschbarkeit, so erhält man zwar eine Aussage über die Unabhängigkeit der einzelnen Zufallsvariablen $X_n,\ n \in \mathbb{N}$, von den symmetrischen Mengen $\mathscr{S}(X_n, n \in \mathbb{N})$, jedoch ist dieses Ergebnis in der Mathematik bereits seit langem bekannt.

Eines der Argumente für die soeben gemachte Aussage ist das folgende Ergebnis von Hu [26]:

6.2 Satz
Sei $(X_n)_{n\in\mathbb{N}}$ ein vertauschbarer Prozess von paarweise unabhängigen Zufallsvariablen. Dann sind die $X_n,\ n \in \mathbb{N}$, unabhängig.

Somit handelt es sich bei der umgeschriebenen Folgerung aus dem Satz von Landers und Rogge um eine andere Notation des wohlbekannten 0-1-Satzes für symmetrische Mengen von Hewitt und Savage [3, Theorem 11.6].

Bei kritischer Betrachtung der Voraussetzungen von Satz 6.1 in Verbindung mit der Definition der schwach ergodischen Prozesse in 2.9 erkennt man, dass die Stationarität für die schwache Ergodizität eines Prozesses zwingend notwendig ist. Des Weiteren erkennt man, dass die paarweise Unabhängigkeit der Zufallsvariablen $X_n,\ n \in \mathbb{N}$, im vorliegenden Fall für die Gültigkeit der SLLN benutzt wurde. Für die Gültigkeit dieses Gesetzes existieren in der mathematischen Literatur jedoch noch weitere hinreichende Bedingungen, so dass es an dieser Stelle zu Variationen in den Voraussetzungen kommen kann. Bei den nun folgenden weiteren Beispielen für schwach ergodische Prozesse muss dem Umstand Rechnung getragen werden, dass das SLLN nicht für die Zufallsvariablen $X_n,\ n \in \mathbb{N}$, gelten muss, sondern insbesondere für $\varphi \circ X_n,\ n \in \mathbb{N}$, mit einer injektiven und beschränkten Funktion φ, wie es bereits im Satz 5.11 dargelegt wurde.

Die folgende Definition beschreibt eine weitere Klasse stochstischer Prozesse, die die Klasse der paarweise unabhängigen Prozesse umfasst. Für diese veröffentlichte Etemadi im Jahre 1996 [20] einen Artikel über die Gültigkeit des SLLN.

6.3 Definition

i) Ein stochastischer Prozess $(X_n)_{n\in\mathbb{N}}$ heißt *2-vertauschbar*, wenn alle geordneten Paare $(X_i, X_j)\,,\ i,j \in \mathbb{N}, i \neq j$ identisch verteilt sind.

ii) Ein stochastischer Prozess $(X_n)_{n\in\mathbb{N}}$ heißt *2-vertauschbar im zweiten Moment*, wenn für $i,j \in \mathbb{N}$ gilt

$$P\left(X_i \cdot X_j\right) \;=\; \begin{cases} P\left(X_1 \cdot X_2\right), \text{ falls } i \neq j, \\ P\left({X_1}^2\right), \text{ falls } i = j. \end{cases}$$

Im Hinblick auf eine Erweiterung des Satzes 6.1 auf die Klasse der 2-vertauschbaren Prozesse ist folgendes Lemma von Bedeutung:

6.4 Lemma

Es sei $(X_n)_{n\in\mathbb{N}}$ ein stationärer 2-vertauschbarer Prozess und φ eine beschränkte messbare Funktion, dann ist der Prozess $(\varphi \circ X_n)_{n\in\mathbb{N}}$ 2-vertauschbar im zweiten Moment.

Beweis:

Seien $i,j \in \mathbb{N}$. Die definierende Bedingung für $i = j$ ergibt sich aus der Stationarität und der damit verbundenen identischen Verteilung des Prozesses $(\varphi \circ X_n)_{n\in\mathbb{N}}$.

Für $i \neq j \in \mathbb{N}$ ist nach obiger Definition zu zeigen:

$$P\left((\varphi \circ X_i)(\varphi \circ X_j)\right) \;=\; P\left((\varphi \circ X_1)(\varphi \circ X_2)\right). \tag{6.2}$$

Nun gilt mit der zweidimendionalen, messbaren Funktion

$$\Phi(x,y) := \varphi(x) \cdot \varphi(y):$$

$$\begin{aligned}
P\left((\varphi\circ X_i)(\varphi\circ X_j)\right) &= \int (\varphi\circ X_i)(\varphi\circ X_j)\ \mathrm{d}P &&= \int \Phi\circ(X_i,X_j)\ \mathrm{d}P \\
&= \int \Phi\ \mathrm{d}P_{(X_i,X_j)} &&= \int \Phi\ \mathrm{d}P_{(X_1,X_2)} \\
&= \int \Phi\circ(X_1,X_2)\ \mathrm{d}P &&= \int (\varphi\circ X_1)(\varphi\circ X_2)\ \mathrm{d}P \\
&= P\left((\varphi\circ X_1)(\varphi\circ X_2)\right).
\end{aligned}$$

Somit gilt (6.2). □

Für die gesuchte Aussage wird nun noch das schon angesprochene Ergebnis von Etemadi [20] benötigt, welches wie folgt lautet.

6.5 Satz

Sei $(X_n)_{n\in\mathbb{N}}$ ein stochastischer Prozess, der 2-vertauschbar im zweiten Moment ist und für den $P(X_1) = P(X_n)$ für $n \in \mathbb{N}$ gilt. Dann folgt:

$$\lim_{n\to\infty} \frac{1}{n}\cdot\sum_{i=1}^{n} X_i \ = \ X \quad P\text{-f.s. und in } L^2,$$

wobei X eine Zufallsvariable mit $P(X) = P(X_1)$ und $\mathrm{Var}(X) = \mathrm{Cov}(X_1, X_2)$ ist.

Für unkorrelierte Zufallsvariablen $X_n,\ n \in \mathbb{N}$, folgt weiter $X = P(X_1)$ P-f.s..

Dem ersten Anschein nach kann mit diesem Ergebnis ein weiteres, verallgemeinertes Beispiel für schwach ergodische Prozesse gegeben werden. Angesichts der gewünschten Anwendung im Beweisgang von Landers und Rogge ist jedoch die kleine Zusatzbemerkung im letzten Satz von ausschlaggebender Bedeutung. Die Unkorreliertheit der Zufallsvariablen sichert die für den Beweisgang notwendige Konvergenz gegen den Erwartungswert. Hierbei muss jedoch beachtet werden, dass nicht der Prozess $(X_n)_{n\in\mathbb{N}}$ den Bedingungen des obigen Satzes von Etemadi genügen muss, sondern die Prozesse $(\varphi_k^x\circ X_n)_{n\in\mathbb{N}}$, $k \in \mathbb{N}, x \in \mathbb{R}$, wobei die Folgen $(\varphi_k^x)_{k\in\mathbb{N}}$ jeweils Folgen von $]0,1[$-wertigen, messbaren, injektiven Funktionen sind, die punktweise gegen $\varphi^x = 1_{]-\infty,x]}$ konvergieren. Die Unkorreliertheit dieser Prozesse, die, wie schon erwähnt, für die Anwendung im Beweis benötigt würde, ergäbe jedoch angesichts der Grenzwertbetrachtungen die paarweise Unabhängigkeit der Zufallsvariablen $X_n,\ n \in \mathbb{N}$. Somit liefert die 2-Vertauschbarkeit des Prozesses $(X_n)_{n\in\mathbb{N}}$ kein neues Ergebnis, sondern nur eine Variation des Ergebnisses von Landers und Rogge.

Aufgrund der besonderen Voraussetzungen und der geforderten Eigenschaften der Borel-messbaren Funktionen $(\varphi_n^x)_{n\in\mathbb{N}},\ x \in \mathbb{R}$, nämlich der $]0,1[$-Wertigkeit, kann jedoch auch ein anderes Ergebnis von Etemadi [19] im Umfeld des SLLN zu Rate gezogen werden, das mit schwächeren Voraussetzungen als der allgemeinen Unkorreliertheit auskommt und vollständig auf die 2-Vertauschbarkeit verzichtet.

6.6 Satz

Sei $(X_n)_{n\in\mathbb{N}}$ ein stochastischer Prozess von nicht-negativen Zufallsvariablen mit endlichen zweiten Momenten, für die mit $S_n = \sum_{i=1}^n X_i,\ n \in \mathbb{N}$, gilt:

(i) $\sup\{P(X_n) \mid n \in \mathbb{N}\} \ < \ \infty$,

(ii) $P(X_i \cdot X_j) \ \leq P(X_i) \cdot P(X_j)$, für $i < j$ und

(iii) $\sum_{i=1}^\infty \frac{\mathrm{Var} X_i}{i^2} \ < \ \infty$.

Dann gilt

$$\lim_{n\to\infty} \left(\frac{S_n}{n} \ - \ \frac{P(S_n)}{n} \right) \ = \ 0 \qquad P\text{-f.s.}.$$

Für die Anwendung dieses Ergebnisses im Beweisgang von Satz 6.1 sind also die Bedingungen (i) bis (iii) für die Prozesse $(\varphi_k^x \circ X_n)_{n\in\mathbb{N}}\,,\ k \in \mathbb{N}, x \in \mathbb{R}$, zu zeigen. Hierbei werden die Punkte (i) und (iii) jedoch schon von diesen speziellen Prozessen erfüllt, denn es gilt für (i):

$$\sup\left\{P\left(\varphi_k^x \circ X_n\right) \mid n \in \mathbb{N}\right\} \ \leq \ 1 \ < \ \infty,$$

da φ_k^x für jedes $x \in \mathbb{R}$ und jedes $k \in \mathbb{N}$ eine $]0,1[$-wertige Funktion ist.

Für die Bedingung (iii) kommen vergleichbare Überlegungen in Betracht, denn auch hier gilt aufgrund der $]0,1[$-Wertigkeit der Funktionen φ_k^x die folgende Ungleichung für $n \in \mathbb{N}$:

$$\mathrm{Var}\left(\varphi_k^x \circ X_n\right) \ = \ P\left(\left(\varphi_k^x \circ X_n\right)^2\right) - P\left(\varphi_k^x \circ X_n\right)^2 \ \leq \ 2.$$

Somit ergibt sich:

$$\sum_{n=1}^{\infty} \frac{\mathrm{Var}\left(\varphi_k^x \circ X_n\right)}{n^2} \ \leq \ 2 \cdot \sum_{n=1}^{\infty} \frac{1}{n^2} \ < \ \infty.$$

Also werden diese beiden Voraussetzungen von den speziellen Prozessen erfüllt.

Angesichts der notwendigen Stationarität des Prozesses $(X_n)_{n\in\mathbb{N}}$ und der Aussage von Satz 3.3 folgt die jeweilige identische Verteilung der Zufallsvariablen des Prozesses $(\varphi_k^x \circ X_n)_{n\in\mathbb{N}}$. Somit gilt in der Aussage von Satz 6.6 also die P-f.s.-Konvergenz der Folge $\left(\frac{1}{n} \cdot S_n\right)_{n\in\mathbb{N}}$ gegen den gemeinsamen Erwartungswert, da ja die Folge $\left(\frac{1}{n} \cdot P(S_n)\right)_{n\in\mathbb{N}}$ im vorliegenden Fall konstant ist.

Nach den obigen Ausführungen bleibt nur die Erfüllung der Bedingung (ii) für die Prozesse $(\varphi_k^x \circ X_n)_{n\in\mathbb{N}}$, $k \in \mathbb{N}$, $x \in \mathbb{R}$, nachzuweisen, um die schwache Ergodizität eines stationären Prozesses zu beweisen. Daher gilt der folgende Sachverhalt.

6.7 Satz

Sei $(X_n)_{n\in\mathbb{N}}$ ein stationärer Prozess mit Werten in einem separablen metrischen Raum, so dass für die Prozesse $(\varphi_k^x \circ \psi_M \circ X_n)_{n\in\mathbb{N}}$ für $x \in \mathbb{R}, k \in \mathbb{N}$, wobei die $(\varphi_k^x)_{k\in\mathbb{N}}$ jeweils Folgen von $]0,1[$-wertigen, messbaren, injektiven Funktionen sind, die punktweise gegen $\varphi^x = 1_{]-\infty,x]}$ konvergieren und ψ_M der im letzten Paragraphen angegebene Isomorphismus ist, gilt:

$$P\left((\varphi_k^x \circ \psi_M \circ X_i) \cdot (\varphi_k^x \circ \psi_M \circ X_j)\right) \;\le\; P\left(\varphi_k^x \circ \psi_M \circ X_1\right)^2 \text{ für } i < j.$$

Dann ist $(X_n)_{n\in\mathbb{N}}$ ein schwach ergodischer Prozess.

Angesichts der Tatsache, dass die im obigen Satz geforderten Voraussetzungen nur mit recht hohem Aufwand nachprüfbar sind, soll hier eine Verallgemeinerung angegeben werden. Diese kann jedoch nicht so umfassend sein wie das Original, dafür jedoch leichter handhabbar.

6.8 Korollar

Sei $(X_n)_{n\in\mathbb{N}}$ ein reellwertiger, stationärer Prozess, und es gilt für $(i,j) \in \mathbb{N}^2$, $i \neq j$ die Aussage:

$$P\left(\left(1_{]-\infty,x]} \circ X_i\right) \cdot \left(1_{]-\infty,x]} \circ X_j\right)\right) \;<\; P\left(1_{]-\infty,x]} \circ X_1\right)^2 \quad (\forall x \in \mathbb{R}).$$

Dann ist $(X_n)_{n\in\mathbb{N}}$ schwach ergodisch.

Im obigen Korollar kann die Bedingung nicht auf die Gleichheit der beiden Ausdrücke erweitert werden, denn wie schon zuvor erwähnt, muss die $\le$-Relation für die konvergente Folge $(\varphi_k^x \circ X_n)_{k\in\mathbb{N}}$ erfüllt sein. Ist die Gleichheit in der obigen Bedingung gegeben, können keine Rückschlüsse auf die Relationen zwischen den Folgegliedern der beiden konvergenten Folgen gezogen werden. Dieses ist nur bei der $<$-Relation möglich, daher die hier verwendete Einschränkung.

Es existieren jedoch noch weitere Bedingungen für die Gültigkeit des SLLN im Umfeld von Unabhängigkeit und 2-Vertauschbarkeit. Hier sei auf die verschiedenen Definitionen von *Mischbarkeits-Koeffizienten* verwiesen [15].

Im Bereich der *Mischbarkeit* existiert ein weiterer Satz, der sich mit hinreichenden Bedingungen für die Gültigkeit des starken Gesetzes der großen Zahlen beschäftigt. Bevor dieser jedoch angegeben werden kann, benötigt man die folgende Definition, die aus dem Buch von Stout [32, Seite 138] stammt.

6.9 Definition
Sei $(X_n)_{n\in\mathbb{N}}$ ein reellwertiger, stochastischer Prozess und

$$\mathscr{K}_{i,j} \ := \ \sigma(X_k, i \le k \le j)$$

für alle $1 \le i \le j < \infty$.
$(X_n)_{n\in\mathbb{N}}$ heißt **-mischend*, wenn es eine Funktion $\psi : \mathbb{N} \to \mathbb{R}$ mit

$$\psi(m) \to 0 \text{ für } m \to \infty$$

und ein $M \in \mathbb{N}$ gibt, so dass für alle $A \in \mathscr{K}_{1n},\ B \in \mathscr{K}_{m+n,m+n} = \sigma(X_{m+n})$ mit $m \ge M$ und $n \in \mathbb{N}$ folgt

$$|P(A \cap B) - P(A)P(B)| \ \le \ \psi(m) \cdot P(A)P(B).$$

Der schon angesprochene Satz über das SLLN stammt aus dem Artikel von Blum, Hanson und Koopmans [7].

6.10 Satz
Sei $(X_n)_{n\in\mathbb{N}}$ ein *-mischender Prozess mit $P(X_n) = 0$ und $P(|X_n|) \le K < \infty$ für jedes $n \in \mathbb{N}$ und

$$\sum_{i=1}^{\infty} \frac{P({X_i}^2)}{i^2} \ < \ \infty. \tag{6.3}$$

Dann erfüllt $(X_n)_{n\in\mathbb{N}}$ das starke Gesetz der großen Zahlen.

Ersetzt man nun das Ergebnis von Etemadi im Satz 6.1 durch dieses Ergebnis von Blum, Hanson und Koopmans, so erhält man den folgenden Sachverhalt.

6.11 Satz
Sei $(X_n)_{n\in\mathbb{N}}$ ein reellwertiger, *-mischender, stationärer Prozess, dann ist $(X_n)_{n\in\mathbb{N}}$ schwach ergodisch.

Der Beweis dieses Satzes erfolgt wiederum nach der bekannten Vorgehensweise, die aus dem grundlegenden Artikel von Landers und Rogge stammt. Wie schon erwähnt, muss dazu gezeigt werden, dass das SLLN für den Prozess $(\varphi \circ X_n)_{n\in\mathbb{N}}$ mit einer messbaren und injektiven Funktion $\varphi : \mathbb{R} \to]0,1[$ gilt. Man wende hierfür den Satz 6.10 auf $Y_n := \varphi \circ X_n - P(\varphi \circ X_n),\ n \in \mathbb{N}$ an. Der Erwartungswert 0 ist trivialerweise erfüllt, die Beschränktheit von $P(|Y_n|), n \in \mathbb{N}$, ist durch die

Beschränktheit von φ durch 1 gesichert, welche auch die Konvergenz der Reihe aus (6.3) bedingt, da $P\left({Y_n}^2\right) = P\left((\varphi \circ X_n)^2\right) - P\left(\varphi \circ X_n\right)^2 \leq 2, n \in \mathbb{N}$, ist. Angesichts der Tatsache, dass für $i \leq j \in \mathbb{N}$ gilt:

$$\sigma(\varphi \circ X_k, i \leq k \leq j) \ \subset \ \sigma(X_k, i \leq k \leq j),$$

folgt aus der *-Mischbarkeit von $(X_n)_{n\in\mathbb{N}}$ die *-Mischbarkeit von $(\varphi \circ X_n)_{n\in\mathbb{N}}$.

Ein Vertreter der *-mischenden Prozesse ist die Klasse der sogenannten *m-abhängigen Prozesse*. Die Verbindung dieser speziellen stochastischen Prozesse zu den unabhängigen Prozessen wird im Folgenden deutlich, denn für m-abhängige Prozesse $(X_n)_{n\in\mathbb{N}}$ sind die terminalen Mengen $\mathscr{T}(X_n, n \in \mathbb{N})$ trivial und für m-abhängige, vertauschbare Prozesse $(Y_n)_{n\in\mathbb{N}}$ sind die symmetrischen Mengen $\mathscr{S}(X_n, n \in \mathbb{N})$ trivial. Doch zunächst zur Definition dieser Prozesse.

6.12 Definition

Ein stochastischer Prozess $(X_n)_{n\in\mathbb{N}}$ heißt *m-abhängig* ($m \in \mathbb{N}$), wenn

$$\{X_1, X_2, \ldots, X_n\} \text{ und } \{X_{n+m+1}, X_{n+m+2}, \ldots\}$$

für alle $n \in \mathbb{N}$ unabhängige Klassen von Zufallsvariablen sind.

Beispiele für derartige Prozesse finden sich z. B. im Buch von Stout [32].

Die durch die Definition deutlich gewordene recht enge Beziehung zu den unabhängigen Zufallsvariablen und die Analyse der Beweise des 0-1-Gesetzes von Kolmogorov und des 0-1-Gesetzes von Hewitt und Savage deuten auf die Möglichkeit der Trivialität der entsprechenden σ-Algebren hin. Diese soll nun konkret bewiesen werden.

6.13 Satz

Es sei $(X_n)_{n\in\mathbb{N}}$ ein m-abhängiger Prozess, dann ist $\mathscr{T}(X_n, n \in \mathbb{N})$ trivial.

Beweis:

Nach Definition der m-Abhängigkeit existiert ein $m_0 \in \mathbb{N}$, derart dass für alle $n \in \mathbb{N}$

$$\{X_1, X_2, \ldots, X_n\} \text{ unabhängig ist von } \{X_{n+m_0+1}, X_{n+m_0+2}, \ldots\}. \tag{6.4}$$

Es sei nun $A \in \bigcap_{n=1}^{\infty} \sigma(X_n, X_{n+1}, X_{n+2}, \ldots) = \mathscr{T}(X_n, n \in \mathbb{N})$ und man setze

$$\mathscr{D} \ := \ \{D \in \mathscr{A} \mid D \text{ ist unabhängig von } A\}.$$

Somit ist zu zeigen:

$$A \in \mathscr{D} \tag{6.5}$$

Nach Definition A.4 ist $\mathscr{D}$ offensichtlich ein Dynkin-System. Aufgrund der Definition der terminalen Mengen und (6.4) folgt nun:

$$(\forall n \in \mathbb{N}) \quad \sigma(X_1, X_2, \ldots, X_n) \in \mathscr{D}. \tag{6.6}$$

Nun ist $(\sigma(X_1, X_2, \ldots, X_n))_{n\in\mathbb{N}}$ eine aufsteigende Folge von σ-Algebren, also ist $\bigcup_{n=1}^{\infty} \sigma(X_1, X_2, \ldots, X_n)$ offenbar $\cap$-stabil. Daher gilt nach Satz A.6 (b) und (6.6):

$$\sigma\left(\bigcup_{n=1}^{\infty} \sigma(X_1, X_2, \ldots, X_n)\right) = D\left(\bigcup_{n=1}^{\infty} \sigma(X_1, X_2, \ldots, X_n)\right) \subset \mathscr{D}. \tag{6.7}$$

Weiter folgt:

$$\begin{aligned}
\sigma(X_k) \quad &\subset \quad \bigcup_{n=1}^{\infty} \sigma(X_1, X_2, \ldots, X_n) \quad (\forall k \in \mathbb{N}) \\
&\Rightarrow \quad \bigcup_{\nu=k}^{\infty} \sigma(X_\nu) \subset \bigcup_{n=1}^{\infty} \sigma(X_1, X_2, \ldots, X_n) \quad (\forall k \in \mathbb{N}) \\
&\Rightarrow \quad \sigma(X_k, X_{k+1}, \ldots) \subset \sigma\left(\bigcup_{n=1}^{\infty} \sigma(X_1, X_2, \ldots, X_n)\right) \underset{(6.7)}{\subset} \mathscr{D} \quad (\forall k \in \mathbb{N}) \\
&\Rightarrow \quad \bigcap_{k=1}^{\infty} \sigma(X_k, X_{k+1}, \ldots) \subset \mathscr{D},
\end{aligned}$$

d. h. $A \in \mathscr{D}$ und somit ist der Beweis für die Aussage dieses Satzes erbracht, d. h. es gilt (6.5). □

Aus diesem Satz und Satz 4.15 ergibt sich nun das folgende Korollar für vertauschbare, m-abhängige Prozesse.

6.14 Korollar

Es sei $(X_n)_{n\in\mathbb{N}}$ ein vertauschbarer, m-abhängiger Prozess, dann ist $\mathscr{S}(X_n,\ n \in \mathbb{N})$ trivial.

Beweis:

Nach Satz 6.13 ist $\mathscr{T}(X_n, n \in \mathbb{N})$ aufgrund der m-Abhängigkeit trivial, folglich ist $\mathscr{I}(X_n, n \in \mathbb{N})$ als Unter-σ-Algebra von $\mathscr{T}(X_n, n \in \mathbb{N})$ degeneriert, also folgt auch die Trivialität von $\mathscr{I}^*(X_n, n \in \mathbb{N})$. Da nach Satz 4.15 gerade die Inklusion $\mathscr{S}(X_n, n \in \mathbb{N}) \subset \mathscr{I}^*(X_n, n \in \mathbb{N})$ gilt, folgt die Trivialität von $\mathscr{S}(X_n, n \in \mathbb{N})$. □

Zum Abschluss dieses Paragraphen soll noch eine kurze Zusammenfassung einiger Beispiele gegeben werden, die als Kurzübersicht anzusehen ist.

6.15 Bemerkung

Für einen stochastischen Prozess $(X_n)_{n\in\mathbb{N}}$ gilt:

$(X_n)_{n\in\mathbb{N}}$ unabhängig und identisch verteilt	$\Rightarrow$	Trivialität von $\mathscr{S}(X_n, n \in \mathbb{N})$
$(X_n)_{n\in\mathbb{N}}$ m-abhängig und vertauschbar	$\Rightarrow$	Trivialität von $\mathscr{S}(X_n, n \in \mathbb{N})$
$(X_n)_{n\in\mathbb{N}}$ m-abhängig	$\Rightarrow$	Trivialität von $\mathscr{T}(X_n, n \in \mathbb{N})$
$(X_n)_{n\in\mathbb{N}}$ paarweise unabhängig und stationär	$\Rightarrow$	schwache Ergodizität

§ 7

Markov-Prozesse und die schwache Ergodizität

Der folgende Paragraph soll Aufschluss darüber geben, unter welchen Bedingungen ein stationärer Markov-Prozess schwach ergodisch ist. Es wird deutlich werden, dass es für diese spezielle Klasse stochastischer Prozesse keinen Unterschied zwischen der schwachen Ergodizität und der Ergodizität gibt. Dieses liegt in der Darstellung der invarianten bzw. der P-f.s.-invarianten Mengen eines Markov-Prozesses begründet. Ein vergleichbares Ergebnis ist bereits bei den vertauschbaren Prozessen in Paragraph 4 zu verzeichnen gewesen.

Doch zunächst die Definition der Markov-Prozesse.

7.1 Definition

Ein stochastischer Prozess $(X_n)_{n\in\mathbb{N}}$ auf einem Wahrscheinlichkeitsraum $(\Omega, \mathscr{A}, P)$ heißt *Markov-Prozess*, wenn gilt:

$$(\forall n \in \mathbb{N})\,(\forall A \in \sigma(X_{n+1}))\quad P(A \mid X_1, X_2, \ldots, X_n) \;=\; P(A \mid X_n) \qquad P\text{-f.s..}$$

7.2 Bemerkung

Die in Definition 7.1 angeführte Bedingung für einen Markov-Prozess ist nach einer Aussage aus dem Buch von Breiman [9, Proposition 7.5] gleichwertig zu folgender Bedingung:

$$(\forall n \in \mathbb{N})\,(\forall B \in \sigma(X_{n+1}, X_{n+2}, \ldots))\; P(B \mid X_1, X_2, \ldots, X_n) \;=\; P(B \mid X_n)\;\; P\text{-f.s..}$$

Aufgrund der Definition von $\mathscr{I}^*$ in Definition 4.7 ist offensichtlich, dass es zu jeder P-f.s.-invarianten Menge I eine terminale Menge T gibt mit $P(I \triangle T) = 0$, denn zu jeder P-f.s.-invarianten Menge I existiert eine invariante Menge I_0 mit $P(I \triangle I_0) = 0$ und nach Bemerkung 2.6 ist diese Menge I_0 auch eine terminale Menge.

Vor diesem Hintergrund kann nun eine Darstellung der P-f.s.-invarianten Mengen eines stationären Markov-Prozesses angegeben werden. Dieser Satz ist im Buch von Ash und Gardner [2, Theorem 3.5.4] oder in einer Abwandlung mit leicht variierenden Voraussetzungen auch im Buch von Stout [32, Lemma 3.6.4] zu finden.

Zuvor werden jedoch noch zwei Eigenschaften angegeben, die für den Beweis des Satzes benötigt werden (vgl. [9, Theorem 5.21] bzw. [2, Lemma 3.5.2]).

7.3 Lemma
Es sei $(X_n)_{n\in\mathbb{N}}$ ein stochastischer Prozess.

(i) Für jede integrierbare ZV Z gilt:

$$P\left(Z \mid X_1, \ldots, X_n\right) \underset{n\to\infty}{\longrightarrow} P\left(Z \mid X_1, X_2, \ldots\right) \quad P\text{-f.s. und in } L_1.$$

(ii) Ist $(X_n)_{n\in\mathbb{N}}$ ein Markov-Prozess und T eine maßerhaltende Transformation mit $X_n = X_1 \circ T^{n-1}$ P-f.s., $n \in \mathbb{N}$, so gilt für jedes $A \in \mathscr{A}$ und $n \in \mathbb{N}$:

$$P\left(A \mid X_n\right)(T\omega) \quad = \quad P\left(T^{-1}A \mid X_{n+1}\right)(\omega) \quad P\text{-f.s..}$$

Hiermit kommt man nun zu folgendem Satz:

7.4 Satz
Sei $(X_n)_{n\in\mathbb{N}}$ ein stationärer Markov-Prozess mit Werten in $(\Omega_1, \mathscr{A}_1)$. Dann lässt sich jede P-f.s.-invariante Menge $A \in \mathscr{I}^*(X_n, n \in \mathbb{N})$ P-f.s. darstellen als:

$$A \;=\; [X_1 \in C] \;\; P\text{-f.s.} \qquad \text{mit } C \in \mathscr{A}_1,$$

also $P\left(A \triangle [X_1 \in C]\right) = 0$.

Beweis:
Für den Beweis des vorliegenden Satzes betrachtet man wieder einen stationären Prozess, der aus einer maßerhaltenden Transformation T und einer Zufallsvariable X besteht. Nach den Bemerkungen im Umfeld von Definition 2.15 und Bemerkung 2.7 folgt, dass der auf diese Art geführte Beweis auch für allgemeine stationäre Markov-Prozesse gültig ist.

Sei nun $A \in \mathscr{I}^*(X_n, n \in \mathbb{N})$. Nach einer Bemerkung am Anfang dieses Paragraphen existiert eine Menge $B \in \mathscr{T}(X_n, n \in \mathbb{N})$ mit $P(A \triangle B) = 0$. Da nun B eine terminale Menge ist, folgt nach Definition von $\mathscr{T}(X_n, n \in \mathbb{N})$:

$$(\forall n \in \mathbb{N}) \quad B \in \sigma\left(X_{n+1}, X_{n+2}, X_{n+3}, \ldots\right).$$

Somit folgt mit der Markov-Eigenschaft des Prozesses $(X_n)_{n\in\mathbb{N}}$ resp. nach Bemerkung 7.2 sofort für alle $n \in \mathbb{N}$: $P(B \mid X_1, \ldots X_n) = P(B \mid X_n) P$-f.s.. Folglich gilt auch:

$$(\forall n \in \mathbb{N}) \quad P(A \mid X_1, \ldots X_n) = P(A \mid X_n) \;\; P\text{-f.s..}$$

Nach 7.3 (i) gilt:

$$P(A \mid X_1, \ldots, X_n) \;\longrightarrow\; P(A \mid X_1, X_2, \ldots) \;\; P\text{-f.s.,} \tag{7.1}$$

ebenso gilt jedoch auch:

$$P(A \mid X_1, X_2, \ldots) \;=\; P(1_A \mid X_1, X_2, \ldots) \;=\; 1_A \;\; P\text{-f.s.,} \tag{7.2}$$

da $A \in \mathscr{I}^*(X_n, n \in \mathbb{N}) \subset \sigma(X_1, X_2, \ldots)$. Aus (7.1) und (7.2) folgt:

$$P(A \mid X_n) \;\longrightarrow\; 1_A \;\; P\text{-f.s..} \tag{7.3}$$

Sei nun $\varepsilon > 0$ und für $n \in \mathbb{N}$ definiere man

$$H_n \;:=\; \{\omega : |P(A \mid X_n)(\omega) \;-\; 1_A(\omega)| \geq \varepsilon\}.$$

Hiermit folgt aufgrund der P-f.s.-Konvergenz aus (7.3) und der daraus resultierenden stochastischen Konvergenz sofort:

$$P(H_n) \underset{n\to\infty}{\longrightarrow} 0. \tag{7.4}$$

Weiter folgt nun für $n \in \mathbb{N}$:

$$\begin{aligned} T^{-1}(H_n) \;&=\; \{\omega : |P(A \mid X_n)(T\omega) \;-\; 1_A(T\omega)| \geq \varepsilon\} \\ &=\; \left\{\omega : \left|P(T^{-1}A \mid X_{n+1})(\omega) \;-\; 1_{T^{-1}A}(\omega)\right| \geq \varepsilon\right\} \\ &=\; \{\omega : |P(A \mid X_{n+1})(\omega) \;-\; 1_A(\omega)| \geq \varepsilon\} \\ &=\; H_{n+1} \;\; P\text{-f.s.} \end{aligned}$$

aufgrund von Lemma 7.3 (ii) und der P-f.s.-Invarianz von A. Da T als maßerhaltend vorausgesetzt war folgt nun für $n \in \mathbb{N}$

$$P(H_n) \;=\; P\left(T^{-1}(H_n)\right) \;=\; P(H_{n+1}).$$

Angesichts der Konvergenz aus (7.4) folgt $P(H_n) \equiv 0$ für $n \in \mathbb{N}$, d. h.

$$(\forall n \in \mathbb{N}) \;\; P(A \mid X_n) \;=\; 1_A \quad P\text{-f.s.},$$

insbesondere also auch $P(A \mid X_1) \;=\; 1_A \;\; P$-f.s..

Nun ist $P(A \mid X_1)$ $\sigma(X_1)$-messbar, kann also nach einer Aussage aus dem Buch von Ash [1, 6.4.2 (c)] durch eine Borel-messbare Funktion $f : \Omega_1 \to \Omega_1$ als $f(X_1)$ dargestellt werden. Somit folgt nun P-f.s.:

$$\begin{aligned} \omega \in A &\Leftrightarrow 1_A(\omega) = 1 \\ &\Leftrightarrow f\left(X_1(\omega)\right) = 1 \quad P\text{-f.s.} \\ &\Leftrightarrow X_1(\omega) \in C \quad P\text{-f.s., mit } C = f^{-1}(1). \end{aligned}$$

Insgesamt folgt also die Aussage des Satzes. □

Für eine Äquivalenzaussage bezüglich der Gestalt aller P-f.s.-invarianten Mengen benötigt man nun noch das folgende Lemma (vgl. [2, Lemma 3.2.4]).

7.5 Lemma

Für $A \in \mathscr{A}$ gilt:

$$A \in \mathscr{I}^*(X_n, n \in \mathbb{N}) \;\Leftrightarrow\; P(T^{-1}A \setminus A) = 0 \text{ oder } P(A \setminus T^{-1}A) = 0,$$

d. h. falls P-f.s. aus $T\omega \in A$ auch $\omega \in A$ bzw. P-f.s. aus $\omega \in A$ auch $T\omega \in A$ folgt.

Beweis:

Es gilt nun:

$$\begin{aligned} P(A \setminus T^{-1}A) &= P(A) - P(A \cap T^{-1}A) \\ &= P(T^{-1}A) - P(A \cap T^{-1}A) \\ &= P(T^{-1}A \setminus A). \end{aligned}$$

Somit folgt weiter:

$$P(A \triangle T^{-1}A) \;=\; 2 \cdot P(A \setminus T^{-1}A) \;=\; 2 \cdot P(T^{-1}A \setminus A).$$

Also folgt die Behauptung. □

Somit gilt insgesamt das folgende Korollar.

7.6 Korollar

Sei $(X_n)_{n\in\mathbb{N}}$ ein stationärer Markov-Prozess. Dann gilt für $A \in \mathscr{A}$:
$A \in \mathscr{I}^* := \mathscr{I}^*(X_n, n \in \mathbb{N})$ genau dann, wenn

$$A = [X_n \in C \text{ für alle } n \in \mathbb{N}] \;\; P\text{-f.s.}$$

mit $C \in \mathscr{A}_1$ ist.

Beweis:
Auch dieser Beweis wird mittels der Darstellung eines stationären Prozesses mit einer maßerhaltenden Transformation geführt.

„$\Leftarrow$" Sei $A \in \mathscr{I}^*$. Nach Satz 7.4 gilt:

$$A = [X_1 \in C] \text{ für ein } C \in \mathscr{A}_1$$

und ebenfalls:

$$A \;=\; T^{-1}A \;\; P\text{-f.s..}$$

Somit folgt aber:

$$[X_1 \in C] \;=\; [X_2 \in C] \;\; P\text{-f.s..}$$

Durch Induktion folgt die rechte Seite.

„$\Rightarrow$" Sei A von obiger Form, dann gilt

$$A \;\subset\; T^{-1}A.$$

Nach Lemma 7.5 folgt die linke Seite. □

Aufgrund der in Korollar 7.6 gegebenen Darstellung einer fast-sicher-invarianten Menge eines stationären Prozesses gilt der folgende Sachverhalt (vgl.[2, Satz 3.5.6]).

7.7 Folgerung
Es sei $(X_n)_{n\in\mathbb{N}}$ ein stationärer Markov-Prozess mit Werten in $(\Omega_1, \mathscr{A}_1)$, dann ist $(X_n)_{n\in\mathbb{N}}$ genau dann nicht ergodisch, wenn es eine Menge $A \in \mathscr{A}_1$ gibt mit

$$0 \;<\; P\left([X_n \in A \quad (\forall n \in \mathbb{N})]\right) \;<\; 1.$$

Hiermit gelangt man zu folgender Aussage.

7.8 Satz
Ein stationärer Markov-Prozess $(X_n)_{n\in\mathbb{N}}$ ist genau dann schwach ergodisch, wenn er ergodisch ist.

Beweis:
Zu zeigen ist nur „$\Rightarrow$".
Sei $(X_n)_{n\in\mathbb{N}}$ nicht ergodisch. Dann existiert nach Folgerung 7.7 eine Menge $C \in \mathscr{A}_1$ mit

$$0 \;<\; P\left([X_n \in C \text{ für alle } n \in \mathbb{N}]\right) \;<\; 1. \tag{7.5}$$

Annahme: $(X_n)_{n\in\mathbb{N}}$ ist schwach ergodisch.

Aufgrund der Annahme gilt für jedes $A \in \mathscr{I}^*(X_n, n \in \mathbb{N})$, $B \in \mathscr{A}_1$ und $n \in \mathbb{N}$ nach Bemerkung 4.8:

$$P(A \cap [X_n \in B]) \;=\; P(A) \cdot P(X_n \in B). \tag{7.6}$$

Nach Satz 7.6 folgt auch für $C \in \mathscr{A}_1$:

$$[X_n \in C \text{ für alle } n \in \mathbb{N}] \;\in\; \mathscr{I}^*.$$

Also folgt hieraus unter Zuhilfenahme von (7.6) für jedes $m \in \mathbb{N}$:

$$\begin{aligned} P(X_m \in C) \cdot P\left([X_n \in C \text{ für alle } n \in \mathbb{N}]\right) \\ &= P\left([X_m \in C\} \cap \{X_n \in C \text{ für alle } n \in \mathbb{N}]\right) \\ &= P\left([X_n \in C \text{ für alle } n \in \mathbb{N}]\right). \end{aligned} \tag{7.7}$$

Da nach (7.5) nun $P\left([X_n \in C \text{ für alle } n \in \mathbb{N}]\right) \neq 0$ gilt, muss aufgrund der Identität in (7.7) $P(X_m \in C) = 1$ gelten. Des Weiteren war m beliebig gewählt, also folgt weiter:

$$P\left(\bigcap_{n=1}^{\infty}\{X_n \in C\}\right) \;=\; 1,$$

was wiederum zur Folge hat, dass

$$P\left([X_n \in C \text{ für alle } n \in \mathbb{N}]\right) \;=\; 1$$

gilt. Diese steht aber im Widerspruch zu (7.5).

Somit ist $(X_n)_{n\in\mathbb{N}}$ nicht schwach ergodisch. Also folgt die Behauptung durch Negation dieser Aussage. □

Im vorliegenden Fall kann also eine direkte Verbindung zwischen der Unabhängigkeit der einzelnen Zufallsvariablen von den invarianten Mengen, also der schwachen Ergodizität und der Trivialität derselben Mengen, also der Ergodizität, hergestellt werden.

Die Aussagen und Äquivalenzen dieses Paragraphen beruhen alle auf der besonderen Darstellung der P-f.s.-invarianten Mengen eines Markov-Prozesses. Angesichts der Tatsache, dass es für umfassendere σ-Algebren, also insbesondere $\mathscr{T}(X_n, n \in \mathbb{N})$ und $\mathscr{S}(X_n, n \in \mathbb{N})$, keine derartige Darstellung gibt, wird deutlich, dass die Aussagen dieses Satzes nicht auf andere Fälle übertragbar sind.

§ 8

Decoupling

Die dargelegten Grundlagen der vorliegenden Dissertation in Paragraph 2 machen deutlich, dass die schwache Ergodizität per Defininition (vgl. Definition 2.9) die Stationarität des entsprechenden Prozesses voraussetzt. In den bisherigen Betrachtungen, insbesondere im Paragraphen 6, wurde deutlich, dass neben der Stationarität des Prozesses auch eine gewisse Art der Unabhängigkeit bzw. *Entkopplung* der einzelnen Zufallsvariablen $X_n, n \in \mathbb{N}$, des stochastischen Prozesses für dessen schwache Ergodizität vorliegen muss.

Das noch recht junge mathematische Fachgebiet des *Decouplings* beschäftigt sich mit eben dieser Fragestellung. Bei der Durchsicht der dortigen Ergebnisse fallen angesichts der bedingt identischen Verteilung gewisse Parallelen zum Satz von de Finetti (Satz 4.3) auf, also der Äquivalenz zur Vertauschbarkeit von stochastischen Prozessen. Im folgenden Paragraphen soll der Frage nachgegangen werden, ob zu gewissen Prozessen ein „entkoppelter" vertauschbarer Prozess existiert. Aufgrund des ursprünglichen Interesses der vorliegenden Arbeit wird danach untersucht, welche Eigenschaften eines schwach ergodischen Prozesses auf den entkoppelten Prozess übertragen werden können. Des Weiteren wird kurz zu der Frage Stellung genommen, ob es auch schwach ergodische Prozesse gibt, die die Vertauschbarkeit des entkoppelten Prozesses bedingen.

Für weitere intensivere Beschäftigungen mit dem Thema *Decoupling* sei das Buch von de la Peña und Giné [30] empfohlen, das auch die Grundlage der Ausführungen dieses Paragraphens darstellt.

Zu Beginn der Ausführungen seien die Grunddefinitionen des Decouplings angegeben.

8.1 Definition

Es seien $(X_n)_{n\in\mathbb{N}}$ und $(Y_n)_{n\in\mathbb{N}}$ zwei stochastische Prozesse auf dem Wahrscheinlichkeitsraum $(\Omega, \mathscr{A}, P)$ und $(\mathscr{F}_n)_{n\in\mathbb{N}}$ eine nicht-absteigende Folge von σ-Algebren, wobei X_n und Y_n $\mathscr{F}_n$-messbar sind für $n \in \mathbb{N}$ $(\mathscr{F}_n \subset \mathscr{A})$.

Die Prozesse $(X_n)_{n\in\mathbb{N}}$ und $(Y_n)_{n\in\mathbb{N}}$ heißen dann *tangierende Prozesse* (bezüglich $(\mathscr{F}_n)_{n\in\mathbb{N}}$), wenn für alle $n \in \mathbb{N}$ X_n und Y_n bedingt identisch verteilt bezüglich $\mathscr{F}_{n-1}$ sind $(\mathscr{F}_0 := \{\emptyset, \Omega\})$.

Mit diesem Grundbegriff können bereits die ersten Ergebnisse vorgestellt werden.

8.2 Satz

Es sei $(X_n)_{n\in\mathbb{N}}$ ein stochastischer Prozess und $(\mathscr{F}_n)_{n\in\mathbb{N}}$ eine aufsteigende Folge von σ-Algebren mit $\sigma(X_i) \subset \mathscr{F}_i,\ i \in \mathbb{N}$. Des Weiteren sei $(Y_n)_{n\in\mathbb{N}}$ ein tangierender Prozess bezüglich $(\mathscr{F}_n)_{n\in\mathbb{N}}$ von $(X_n)_{n\in\mathbb{N}}$. Sei nun $\mathscr{H}$ eine Unter-σ-Algebra von $\mathscr{F}_1$, dann gilt:

(i) $(\forall n \in \mathbb{N}_{\geq 2})(\forall A \in \mathscr{A}_1) \quad P(1_A \circ X_n \mid \mathscr{H}) = P(1_A \circ Y_n \mid \mathscr{H})$.

(ii) $(X_n)_{n\in\mathbb{N}}$ ist genau dann bedingt identisch verteilt bezüglich $\mathscr{H}$, wenn $(Y_n)_{n\in\mathbb{N}}$ bedingt identisch verteilt bezüglich $\mathscr{H}$ ist.

Beweis:

(i) Es gilt für $n \in \mathbb{N}_{\geq 2}$ und $A \in \mathscr{A}_1$:

$$\begin{aligned} P(1_A \circ X_n \mid \mathscr{H}) &= P\left(P\left(1_A \circ X_n \mid \mathscr{F}_{n-1}\right) \mid \mathscr{H}\right) \\ &= P\left(P\left(1_A \circ Y_n \mid \mathscr{F}_{n-1}\right) \mid \mathscr{H}\right) \\ &= P(1_A \circ Y_n \mid \mathscr{H}). \end{aligned}$$

(ii) Für $n \in \mathbb{N}_{\geq 2}$ und $A \in \mathscr{A}_1$ gilt:

$$\begin{aligned} P\left(1_A \circ Y_n \mid \mathscr{H}\right) &\underset{A.3(b)}{=} P\left(P\left(1_A \circ Y_n \mid \mathscr{F}_{n-1}\right) \mid \mathscr{H}\right) \\ &= P\left(P\left(1_A \circ X_n \mid \mathscr{F}_{n-1}\right) \mid \mathscr{H}\right) \\ &= P\left(P\left(1_A \circ X_n \mid \mathscr{H}\right) \mid \mathscr{F}_{n-1}\right) \\ &= P\left(P\left(1_A \circ X_1 \mid \mathscr{H}\right) \mid \mathscr{F}_{n-1}\right) \\ &= P\left(P\left(1_A \circ X_1 \mid \mathscr{F}_{n-1}\right) \mid \mathscr{H}\right) \\ &= P\left(P\left(1_A \circ Y_1 \mid \mathscr{F}_{n-1}\right) \mid \mathscr{H}\right) \\ &= P\left(1_A \circ Y_1 \mid \mathscr{H}\right). \end{aligned}$$

Die Gegenrichtung ergibt sich aus dem gegenseitigen Austausch der Zufallsvariablen X_n und $Y_n, n \in \mathbb{N}$. $\square$

8.3 Folgerung

Es seien $(X_n)_{n\in\mathbb{N}}$ und $(Y_n)_{n\in\mathbb{N}}$ tangierende Prozesse, dann gilt:

(i) Für jedes $n \in \mathbb{N}$ sind X_n und Y_n identisch verteilt.

(ii) $(X_n)_{n\in\mathbb{N}}$ ist genau dann identisch verteilt, also bedingt identisch verteilt bezüglich $\{\emptyset, \Omega\}$, wenn $(Y_n)_{n\in\mathbb{N}}$ identisch verteilt ist, und es gilt für jedes $n \in \mathbb{N}$ und $A \in \mathscr{A}_1$:

$$P(1_A \circ X_n) \;=\; P(1_A \circ X_1) \;=\; P(1_A \circ Y_1) \;=\; P(1_A \circ Y_n).$$

(iii) Ist $(X_n)_{n\in\mathbb{N}}$ stationär und gilt $\mathscr{I}(X_n, n \in \mathbb{N}) \subset \mathscr{F}_1$, dann ist $(Y_n)_{n\in\mathbb{N}}$ bedingt identisch verteilt bezüglich $\mathscr{I}(X_n, n \in \mathbb{N})$.

(iv) Aufgrund der Übereinstimmung der bedingten Wahrscheinlichkeiten von X_n und Y_n bezüglich $\mathscr{F}_{n-1}$, $n \in \mathbb{N}$, folgt für jedes $k \in \mathbb{N}$:
X_k ist genau dann unabhängig von $\mathscr{F}_{k-1}$, wenn Y_k unabhängig von $\mathscr{F}_{k-1}$ ist.

8.4 Definition

Es seien $(Y_n)_{n\in\mathbb{N}}$ ein stochastischer Prozess und $(\mathscr{F}_n)_{n\in\mathbb{N}}$ eine aufsteigende Folge von σ-Algebren, wobei Y_n $\mathscr{F}_n$-messbar ist für $n \in \mathbb{N}$.
Der stochastische Prozess $(Y_n)_{n\in\mathbb{N}}$ erfüllt die *bedingte Unabhängigkeitsbedingung* (BUB) (bezüglich $(\mathscr{F}_n)_{n\in\mathbb{N}}$), wenn eine σ-Algebra $\mathscr{G} \subset \mathscr{A}$ existiert, so dass $(Y_n)_{n\in\mathbb{N}}$ bedingt unabhängig bezüglich $\mathscr{G}$ ist und die bedingte Verteilung der Komponente $Y_n, n \in \mathbb{N}$, bezüglich $\mathscr{G}$ mit der bedingten Verteilung von Y_n bezüglich $\mathscr{F}_{n-1}$ übereinstimmt.

8.5 Definition

Ein stochastischer Prozess $(Y_n)_{n\in\mathbb{N}}$ heißt *entkoppelt tangierender Prozess* bezüglich $((X_n)_{n\in\mathbb{N}}; (\mathscr{F}_n)_{n\in\mathbb{N}})$, wenn $(Y_n)_{n\in\mathbb{N}}$ die bedingte Unabhängigkeitsbedingung bezüglich $(\mathscr{F}_n)_{n\in\mathbb{N}}$ erfüllt und $(X_n)_{n\in\mathbb{N}}$ und $(Y_n)_{n\in\mathbb{N}}$ tangierende Prozesse bezüglich $(\mathscr{F}_n)_{n\in\mathbb{N}}$ sind.

Mit diesem Grundgerüst an Begrifflichkeiten kann nun einer der Hauptsätze der mathematischen Entkopplungstheorie (des Decouplings) formuliert werden.

8.6 Satz

Zu jedem stochastischen Prozess $(X_n)_{n\in\mathbb{N}}$, dessen Komponenten $X_n, n \in \mathbb{N}$, bezüglich einer σ-Algebra $\mathscr{F}_n$ einer aufsteigenden Folge von σ-Algebren messbar sind, existiert ein stochastischer Prozess $(Y_n)_{n\in\mathbb{N}}$, der entkoppelt tangierend bezüglich $((X_n)_{n\in\mathbb{N}}; (\mathscr{F}_n)_{n\in\mathbb{N}})$ ist.

8.7 Bemerkung

(i) Der nach Satz 8.6 existierende entkoppelt tangierende Prozess $(Y_n)_{n\in\mathbb{N}}$ ist möglicherweise auf einem „erweiterten" Wahrscheinlichkeitsraum definiert. Zum Verständnis wird im Anschluss an diese Bemerkung ein entsprechendes Beispiel für die Konstruktion eines entkoppelt tangierenden Prozesses angegeben.

(ii) Einer der Hauptfälle in der Decouplingtheorie beschäftigt sich mit der Eigenschaft $\mathscr{G} = \sigma(X_1, X_2, X_3, \ldots)$. Somit kann in diesem Falle dann stets gefolgert werden:

$$\mathscr{I}(X_n, n\in\mathbb{N}) \subset \mathscr{G}.$$

Man beachte, dass die BUB für den entkoppelt tangierenden Prozess $(Y_n)_{n\in\mathbb{N}}$ gefordert wird und es keine direkte Querverbindung zwischen dem Prozess $(X_n)_{n\in\mathbb{N}}$ und der σ-Algebra $\mathscr{G}$ gibt.

Im Anschluss werden zwei Beispiele für die Konstruktion der entkoppelt tangierenden Prozesse angegeben. Das erste beschäftigt sich mit einem Prozess, der in einen vollständig separablen, metrischen Raum abbildet. Hierbei wird auch die Definition des „erweiterten Wahrscheinlichkeitsraums" gezeigt. Das zweite Beispiel beschäftigt sich mit dem diskreten Modell („Urnenmodell"). Diese Beispiele stammen ebenfalls aus dem Buch von de la Peña und Giné [30].

8.8 Beispiel

(a) Es sei $(M, \mathscr{M})$ ein messbarer Raum, bestehend aus einem vollständig separablen metrischen Raum M und seiner Borel-σ-Algebra $\mathscr{M}$. Es sei weiter $(X_n)_{n\in\mathbb{N}}$ ein M-wertiger Prozess und $(\mathscr{F}_n)_{n\in\mathbb{N}}$ eine aufsteigende Folge von σ-Algebren ($\subset \mathscr{M}$), so dass jedes X_n $\mathscr{F}_n$-messbar ist, $n\in\mathbb{N}$. Sind dann die ZV X_n, $n\in\mathbb{N}$, auf dem Wahrscheinlichkeitsraum $(\Omega, \mathscr{A}, P)$ definiert, so setzt man den „erweiterten" Wahrscheinlichkeitsraum $\left(\widetilde{\Omega}, \widetilde{\mathscr{A}}\right)$ folgendermaßen fest:

$$\left(\widetilde{\Omega}, \widetilde{\mathscr{A}}\right) := \left(M^{\mathbb{N}} \times \Omega, \mathscr{M}^{\mathbb{N}} \otimes \mathscr{A}\right).$$

Man beachte, dass für jedes Wahrscheinlichkeitsmaß $\widetilde{P}$ auf $\left(\widetilde{\Omega}, \widetilde{\mathscr{A}}\right)$ gelten muss, dass es bei der Projektion auf den ursprünglichen Wahrscheinlichkeitsraum $(\Omega, \mathscr{A}, P)$ $\left(\widetilde{\Omega} \to \Omega, [(m_n)_{n\in\mathbb{N}}, \omega] \mapsto \omega\right)$ mit dem Wahrscheinlichkeitsmaß

P übereinstimmt. Mit dieser Erweiterung folgt nun, dass der stochastische Prozess $\left(\widetilde{X}_n\right)_{n\in\mathbb{N}}$ mit $\widetilde{X}_n\left([(m_n)_{n\in\mathbb{N}},\omega]\right) := X_n(\omega)$, $n \in \mathbb{N}$, $[(m_n)_{n\in\mathbb{N}},\omega] \in \widetilde{\Omega}$, *stochastisch identisch* zum ursprünglichen Prozess $(X_n)_{n\in\mathbb{N}}$ ist, d. h. die Verteilungen von $(X_{i_1}, X_{i_2}, \ldots, X_{i_k})$ und $\left(\widetilde{X}_{i_1}, \widetilde{X}_{i_2}, \ldots, \widetilde{X}_{i_k}\right)$ stimmen für jede endliche Menge $\{i_1, i_2, \ldots, i_k\} \subset \mathbb{N}$, $k \in \mathbb{N}$, überein.

Da es sich bei M um einen Polnischen Raum handelt, existiert die bedingte Verteilung von X_n bezüglich $\mathscr{F}_{n-1}$, $n \in \mathbb{N}$ (in Zeichen: $\mathscr{L}(X_n \mid \mathscr{F}_{n-1})$). Zur Erläuterung dieses Sachverhaltes betrachte man das Buch von Dudley [16]. Somit folgt, dass für $\omega \in \Omega$ die bedingte Verteilung des Prozesses $(Y_n)_{n\in\mathbb{N}}$ durch $\otimes_{i=1}^{\infty}\left(\mathscr{L}(X_i \mid \mathscr{F}_{i-1})(\omega)\right)$ gegeben sein muss. Da die Zufallsvariablen $Y_n, n \in \mathbb{N}$, die Koordinatenabbildungen von $M^{\mathbb{N}}$ werden sollen, folgt hieraus und aus der obigen Eigenschaft der Projektion zwingend für das Wahrscheinlichkeitsmaß $\widetilde{P}$:

$$\widetilde{P}\{A \times B\} \;=\; \int_B \left(\otimes_{i=1}^{\infty}\mathscr{L}(X_i \mid \mathscr{F}_{i-1})(\omega)\right)(A)\; \mathrm{d}P(\omega).$$

Formal bedeutet dies, dass die Prozesse $\left(\widetilde{X}_n\right)_{n\in\mathbb{N}}$ und $(Y_n)_{n\in\mathbb{N}}$, mit

$$\widetilde{X}_i\left([(m_n)_{n\in\mathbb{N}},\omega]\right) := X_i(\omega) \text{ und } Y_i\left([(m_n)_{n\in\mathbb{N}},\omega]\right) = m_i, i \in \mathbb{N},$$

auf dem Wahrscheinlichkeitsraum $(\widetilde{\Omega}, \widetilde{\mathscr{A}}, \widetilde{P})$ definiert sind und die folgenden Eigenschaften erfüllen:
Unter den Voraussetzungen $\widetilde{\mathscr{F}}_n := \mathscr{M}^n \otimes \mathscr{F}_n, n \in \mathbb{N}$ und $\widetilde{\mathscr{G}} := \{\emptyset, M^{\mathbb{N}}\} \otimes \mathscr{A}$ gilt nun:

- $(X_n)_{n\in\mathbb{N}}$ und $\left(\widetilde{X}_n\right)_{n\in\mathbb{N}}$ sind stochastisch identisch.
- $\widetilde{X}_n$ und Y_n sind $\widetilde{\mathscr{F}}_{n-1}$-messbar.
- $(Y_n)_{n\in\mathbb{N}}$ erfüllt die BUB und ist bedingt unabhängig bezüglich $\widetilde{\mathscr{G}}$.

Da die Prozesse $(X_n)_{n\in\mathbb{N}}$ und $(\widetilde{X}_n)_{n\in\mathbb{N}}$ stochastisch nicht zu unterscheiden sind, handelt es sich hierbei um ein Beispiel für die Konstruktion eines entkoppelt tangierenden Prozesses, der auf einem „erweiterten Wahrscheilichkeitsraum“ definiert ist. Dieses ist möglich, da die Zufallsvariablen des Prozesses $(X_n)_{n\in\mathbb{N}}$ in einen vollständig separablen metrischen Raum abbilden.

(b) Als **diskretes Beispiel** betrachte man das Ziehen einer Stichprobe vom Umfang n aus einer Urne mit N Bällen $\{B_1, B_2, \ldots, B_N\}$, $n, N \in \mathbb{N}, n \le N < \infty$. Die Folge $(X_i)_{1\le i\le n}$ soll eine Stichprobe ohne Zurücklegen darstellen und die

Folge $(Y_i)_{1\leq i\leq n}$ eine bedingt unabhängige Stichprobe. Um eine solche bedingt unabhängige Folge zu erhalten, verfahre man wie folgt:
An der k-ten Stelle der Stichprobe ohne Zurücklegen ($k \in \{1,\ldots,n\}$) erhält man die Realisationen X_k und Y_k durch gleichmäßiges Ziehen aus der Menge $\{B_1, B_2, \ldots, B_N\} \setminus \{X_1, X_2, \ldots, X_{k-1}\}$. Hierbei geht man nun folgendermaßen vor: Man bestimme zuerst die Realisation von Y_k durch Ziehen eines Balles, lege diesen zurück in die Urne, Danach bestimme man die Realisation von X_k und lege den gezogenen Ball beiseite. Nun erkennt man, dass mit der hier angegebenen Prozedur Realisationen $(X_i)_{1\leq i\leq n}$ und $(Y_i)_{1\leq i\leq n}$ der Stichproben entstehen, die tangierend bezüglich $(\mathscr{F}_n)_{n\in\mathbb{N}}$ sind, wobei $\mathscr{F}_n$ für $n \in \mathbb{N}$ folgendermaßen definiert ist:

$$\mathscr{F}_n := \sigma(X_1, X_2, \ldots, X_n, Y_1, Y_2, \ldots, Y_n).$$

Des Weiteren ist $(Y_i)_{1\leq i\leq n}$ bedingt unabhängig bezüglich der σ-Algebra $\mathscr{G} := \sigma(X_1, X_2, \ldots, X_N)$.

Vor dem Hintergrund der beiden obigen Beispiele wird deutlich, dass keine Aussage getroffen werden kann, wann ein stochastischer Prozess als entkoppelt tangierender Prozess bezüglich eines anderen Prozesses und einer bestimmten Folge von σ-Algebren identifiziert werden kann. Selbst die Annahme, ein stochastischer Prozess $(Y_n)_{n\in\mathbb{N}}$, der bedingt unabhängig bezüglich einer σ-Algebra $\mathscr{G} \subset \mathscr{A}$ ist, wäre ein entkoppelt tangierender Prozess bezüglich $((Y_n)_{n\in\mathbb{N}}; (\mathscr{F}_n)_{n\in\mathbb{N}})$ ist nur dann richtig, wenn für alle $n \in \mathbb{N}$ die bedingte Verteilung von Y_n bezüglich $\mathscr{G}$ mit der bedingten Verteilung von Y_n bezüglich $\mathscr{F}_{n-1}$ übereinstimmt, also z. B. falls der Prozess $(Y_n)_{n\in\mathbb{N}}$ unabhängig ist, sowie $\mathscr{F}_n = \sigma(Y_1, \ldots, Y_n)$ und $\mathscr{G} = \{\emptyset, \Omega\}$ gilt.

Angesichts der Existenz eines entkoppelt tangierenden Prozesses $(Y_n)_{n\in\mathbb{N}}$, der bedingt unabhängig bezüglich einer σ-Algebra $\mathscr{G}$ ist, stellt sich vor dem Hintergrund des Satzes von de Finetti 4.3 die Frage, ob dieser Prozess $(Y_n)_{n\in\mathbb{N}}$ unter gewissen Voraussetzungen auch vertauschbar ist. Hierzu benötigt man letztlich noch die bedingt identische Verteilung bezüglich $\mathscr{G}$.

Berücksichtigt man die Erfüllung der bedingten Unabhängigkeitsbedingung von Seiten des Prozesses $(Y_n)_{n\in\mathbb{N}}$ und die Tangiertheit der Prozesse $(X_n)_{n\in\mathbb{N}}$ und $(Y_n)_{n\in\mathbb{N}}$, so ergibt sich folgender Sachverhalt für die Vertauschbarkeit des Prozesses $(Y_n)_{n\in\mathbb{N}}$ (im Hinblick auf die bedingt identische Verteilung und die bedingte Unabhängigkeit bezüglich $\mathscr{G}$).

8.9 Satz

Es sei $(X_n)_{n\in\mathbb{N}}$ ein stochastischer Prozess, dessen Komponenten $X_n, n \in \mathbb{N}$, bezüglich einer σ-Algebra $\mathscr{F}_n$ einer aufsteigenden Folge von σ-Algebren messbar sind. Existiert dann ein $k \in \mathbb{N}$ mit

$$(\forall A \in \mathscr{A}_1)\,(\forall n, m \in \mathbb{N}_{\geq k}) \quad P\left(1_A \circ X_m \mid \mathscr{F}_{m-1}\right) \;=\; P\left(1_A \circ X_n \mid \mathscr{F}_{n-1}\right), \tag{8.1}$$

so ist der nach Satz 8.6 existierende entkoppelt tangierende Prozess $(Y_n)_{n\in\mathbb{N}_{\geq k}}$ vertauschbar.

Beweis:

Nach den vorherigen Ausführungen genügt es, die bedingt identische Verteilung des Prozesses $(Y_n)_{n\in\mathbb{N}_{\geq k}}$ bezüglich $\mathscr{G}$ zu zeigen.

Es gilt für $A \in \mathscr{A}_1$, $n, m \in \mathbb{N}_{\geq k}$:

$$\begin{aligned} P\left(1_A \circ Y_n \mid \mathscr{G}\right) &= P\left(1_A \circ Y_n \mid \mathscr{F}_{n-1}\right) \\ &= P\left(1_A \circ X_n \mid \mathscr{F}_{n-1}\right) \\ &\underset{(8.1)}{=} P\left(1_A \circ X_m \mid \mathscr{F}_{m-1}\right) \\ &= P\left(1_A \circ Y_m \mid \mathscr{F}_{m-1}\right) \\ &= P\left(1_A \circ Y_m \mid \mathscr{G}\right). \end{aligned}$$

Somit folgt die bedingt identische Verteilung des Prozesses $(Y_n)_{n\in\mathbb{N}_{\geq k}}$ bezüglich der σ-Algebra $\mathscr{G}$ und damit die Aussage des Satzes. □

Im Folgenden soll nun untersucht werden, welche Eigenschaften sich von einem schwach ergodischen Prozess $(X_n)_{n\in\mathbb{N}}$ auf einen gewissen bezüglich $(X_n)_{n\in\mathbb{N}}$ entkoppelten Prozesse $(Y_n)_{n\in\mathbb{N}}$ übertragen lassen. Hierfür wird erneut eine spezielle Forderung an die Folge der σ-Algebren $(\mathscr{F}_n)_{n\in\mathbb{N}}$ gestellt. Bisher war hieran nur die Eigenschaft geknüpft, dass $X_n,\ n \in \mathbb{N}$, $\mathscr{F}_n$-messbar sein sollte. Aus Gründen der Bearbeitung mit „geschachtelten" bedingten Wahrscheinlichkeiten wird im restlichen Verlauf dieses Paragraphens stets

$$\mathscr{I}(X_n, n \in \mathbb{N}) \subset \mathscr{F}_k,\ k \in \mathbb{N}$$

gefordert. Hiermit gilt:

8.10 Satz

Es sei $(X_n)_{n\in\mathbb{N}}$ ein schwach ergodischer Prozess und $(Y_n)_{n\in\mathbb{N}}$ ein tangierender Prozess bezüglich $(X_n)_{n\in\mathbb{N}}$ mit $\mathscr{I} := \mathscr{I}\left(X_n, n \in \mathbb{N}\right) \subset \mathscr{F}_1$. Dann gilt:

Jedes $Y_n,\ n \in \mathbb{N}$, ist unabhängig von $\mathscr{I}$; $(Y_n)_{n\in\mathbb{N}}$ ist bedingt identisch verteilt bezüglich $\mathscr{I}$ und folglich auch identisch verteilt.

Beweis:
Aufgrund der Stationarität von $(X_n)_{n\in\mathbb{N}}$ folgt nach Satz 8.2 sofort die identische Verteilung des Prozesses $(Y_n)_{n\in\mathbb{N}}$ und nach Folgerung 8.3 auch die bedingt identische Verteilung des Prozesses $(Y_n)_{n\in\mathbb{N}}$ bezüglich $\mathscr{I}(X_n, n \in \mathbb{N})$. Für $n \in \mathbb{N}$ und $A \in \mathscr{A}_1$ gilt nun:

$$\begin{aligned}
P\left(1_A \circ Y_n \mid \mathscr{I}\right) &\underset{A.3(b)}{=} P\left(P\left(1_A \circ Y_n \mid \mathscr{F}_{n-1}\right) \mid \mathscr{I}\right) \\
&= P\left(P\left(1_A \circ X_n \mid \mathscr{F}_{n-1}\right) \mid \mathscr{I}\right) \\
&= P\left(P\left(1_A \circ X_n \mid \mathscr{I}\right) \mid \mathscr{F}_{n-1}\right) \\
&= P\left(P\left(1_A \circ X_n\right) \mid \mathscr{F}_{n-1}\right) \\
&= P\left(1_A \circ X_n\right) \\
&\underset{8.3(i)}{=} P\left(1_A \circ Y_n\right).
\end{aligned}$$

Also ist jedes $Y_n,\ n \in \mathbb{N}$, unabhängig von $\mathscr{I}(X_n, n \in \mathbb{N})$. □

Im Falle eines entkoppelt tangierenden Prozesses bezüglich eines schwach ergodischen Prozesses benötigt man für den Nachweis der Vertauschbarkeit des Prozesses $(Y_n)_{n\in\mathbb{N}}$ noch die bedingte Unabhängigkeit desselben Prozesses bezüglich $\mathscr{I}(X_n, n \in \mathbb{N})$. Dies ist erforderlich, da die bedingt identische Verteilung bezüglich eben dieser σ-Algebra im vorherigen Satz gezeigt wurde. Hierbei ist die folgende Äquivalenz von Bedeutung, die aus dem Buch von Bauer [3, Aufgabe 5, Seite 128] stammt.

8.11 Satz
Sei Y eine integrierbare, nicht-negative reelle Zufallsvariable auf einem Wahrscheinlichkeitsraum $(\Omega, \mathscr{A}, P)$. Des Weiteren seien $\mathscr{C}_1, \mathscr{C}_2$ zwei Unter-σ-Algebren von $\mathscr{A}$ und $\mathscr{C}_3 := \sigma\{\mathscr{C}_1, \mathscr{C}_2\}$. Dann sind äquivalent:

(i) $P(X \mid \mathscr{C}_3) \;=\; P(X \mid \mathscr{C}_2)$ P-f.s..

(ii) $P(X{\cdot}Y \mid \mathscr{C}_2) \;=\; P(X \mid \mathscr{C}_2) \cdot P(Y \mid \mathscr{C}_2)$ P-f.s.
für jede $\mathscr{C}_1$-messbare reelle Zufallsvariable $Y \geq 0$.

Eine direkte Anwendung dieses Satzes auf die Fragestellung nach der bedingten Unabhängigkeit des Prozesses $(Y_n)_{n\in\mathbb{N}}$ bezüglich $\mathscr{I}$ liefert nur Forderungen an eben diesen Prozess, der aber im Vorhinein nicht bekannt ist. Vielmehr sind Bedingungen für den schwach ergodischen Prozess $(X_n)_{n\in\mathbb{N}}$ von Interesse.

Hierzu wende man den eben vorgestellten Satz aus dem Buch von Bauer auf den Fall $\mathscr{C}_1 := \mathscr{F}_{m-1}, \mathscr{C}_2 := \mathscr{H} \subset \mathscr{F}_1$ und folglich $\mathscr{C}_3 = \mathscr{F}_{m-1}$ $(m \in \mathbb{N}_{\geq 2})$ an. Dabei beachte man, dass die Aussage für $n, m \in \mathbb{N}$ mit $n < m$ gezeigt wird. Somit gilt also die $\mathscr{F}_{m-1}$-Messbarkeit der ZV Y_n.

Hiermit kommt man nun zu folgendem Ergebnis.

8.12 Satz

Es sei $(X_n)_{n\in\mathbb{N}}$ ein stochastischer Prozess, der bedingt identisch verteilt bezüglich der σ-Algebra $\mathscr{H} \subset \mathscr{A}$ ist. Des Weiteren sei $(Y_n)_{n\in\mathbb{N}}$ ein tangierender Prozess von $(X_n)_{n\in\mathbb{N}}$ bezüglich $(\mathscr{F}_n)_{n\in\mathbb{N}}$ mit $\mathscr{H} \subset \mathscr{F}_1$.

Existiert dann ein $k \in \mathbb{N}$, sodass für jedes $A \in \mathscr{A}_1$ und für jedes $m \in \mathbb{N}_{\geq k}$ gilt:

$$P(1_A \circ X_m \mid \mathscr{F}_{m-1}) \;=\; P(1_A \circ X_m \mid \mathscr{H}),$$

dann ist $(Y_n)_{n\in\mathbb{N}_{\geq k}}$ vertauschbar, d. h. bedingt identisch verteilt und bedingt unabhängig bezüglich $\mathscr{H}$.

Beweis:

Nach der Verallgemeinerung des Beweises von Satz 8.10 genügt es, die bedingte Unabhängigkeit des Prozesses $(Y_n)_{n\in\mathbb{N}}$ bezüglich $\mathscr{H}$ zu zeigen.

Es gilt nach den Voraussetzungen und der Äquivalenz aus Satz 8.11 für $A, B \in \mathscr{A}_1$ und $n, m \in \mathbb{N}_{\geq k}$ mit $n < m$:

$$\begin{aligned}
&P\left(1_A\circ Y_n \cdot 1_B\circ Y_m \mid \mathscr{H}\right)\\
&\quad= P\left(P\left(1_A\circ Y_n \cdot 1_B\circ Y_m \mid \mathscr{F}_{m-1}\right) \mid \mathscr{H}\right)\\
&\quad= P\left(1_A\circ Y_n \cdot P\left(1_B \circ Y_m \mid \mathscr{F}_{m-1}\right) \mid \mathscr{H}\right)\\
&\quad= P\left(P\left(1_A \circ Y_n \mid \mathscr{F}_{m-1}\right) \cdot P\left(1_B \circ Y_m \mid \mathscr{F}_{m-1}\right) \mid \mathscr{H}\right)\\
&\quad= P\left(P\left(1_A \circ X_n \mid \mathscr{F}_{m-1}\right) \cdot P\left(1_B \circ X_m \mid \mathscr{F}_{m-1}\right) \mid \mathscr{H}\right)\\
&\quad= P\left(P\left(1_A \circ X_n \mid \mathscr{F}_{m-1}\right) \mid \mathscr{H}\right) \cdot P\left(P\left(1_B \circ X_m \mid \mathscr{F}_{m-1}\right) \mid \mathscr{H}\right)\\
&\quad= P\left(P\left(1_A \circ Y_n \mid \mathscr{F}_{m-1}\right) \mid \mathscr{H}\right) \cdot P\left(P\left(1_B \circ Y_m \mid \mathscr{F}_{m-1}\right) \mid \mathscr{H}\right)\\
&\quad= P\left(1_A \circ Y_n \mid \mathscr{H}\right) \cdot P\left(1_B \circ Y_m \mid \mathscr{H}\right).
\end{aligned}$$

Somit folgt die bedingte Unabhängigkeit von $(Y_n)_{n\in\mathbb{N}}$ bezüglich $\mathscr{H}$, also insgesamt die Vertauschbarkeit des Prozesses $(Y_n)_{n\in\mathbb{N}_{\geq k}}$. □

8.13 Bemerkung

(i) Gilt Satz 8.12 für $k = 1$, so folgt die Unabhängigkeit von X_1 und $\mathscr{H}$.

(ii) Bei zusätzlicher identischer Verteilung des Prozesses $(X_n)_{n\in\mathbb{N}}$ folgt die Unabhängigkeit von X_n und $\mathscr{H}$ für jedes $n \in \mathbb{N}$.

Spezifiziert man Satz 8.12 für den Fall der schwach ergodischen Prozesse, d. h. $\mathscr{H} := \mathscr{I}(X_n, n \in \mathbb{N})$, so ergibt sich das folgende Korollar.

8.14 Korollar
Ist $(X_n)_{n\in\mathbb{N}}$ ein schwach ergodischer Prozess und $(Y_n)_{n\in\mathbb{N}}$ ein entkoppelt tangierender Prozess von $(X_n)_{n\in\mathbb{N}}$ bezüglich $(\mathscr{F}_n)_{n\in\mathbb{N}}$, mit $\mathscr{I} := \mathscr{I}(X_n, n \in \mathbb{N}) \subset \mathscr{F}_1$.
Existiert dann ein $k \in \mathbb{N}$, sodass für jedes $A \in \mathscr{A}_1$ und jedes $m \in \mathbb{N}_{\geq k}$ gilt

$$P(1_A \circ X_m \mid \mathscr{F}_{m-1}) \;=\; P(1_A \circ X_m \mid \mathscr{I}) \;(=\; P(1_A \circ X_m)),$$

dann ist $(Y_n)_{n\in\mathbb{N}_{\geq k}}$ vertauschbar und $(X_n)_{n\in\mathbb{N}}$ ergodisch.

Beweis:
Nach Satz 8.12 folgt mit $\mathscr{H} = \mathscr{I}$ die Vertauschbarkeit des Prozesses $(Y_n)_{n\in\mathbb{N}_{\geq k}}$.
Da aber für diesen speziellen Fall auch die Unabhängigkeit von X_m und $\mathscr{F}_{m-1}$ für jedes $m \in \mathbb{N}_{\geq k}$ gilt, folgt sofort die Unabhängigkeit des Prozesses $(X_{k+n})_{n\in\mathbb{N}}$. Somit ist $(X_{k+n})_{n\in\mathbb{N}}$ ein ergodischer Prozess und es gilt

$$\mathscr{I}(X_n, n \in \mathbb{N}) \;\subset\; \mathscr{I}(X_{k+n}, n \in \mathbb{N}),$$

also ist auch $\mathscr{I}(X_n, n \in \mathbb{N})$ trivial und damit $(X_n)_{n\in\mathbb{N}}$ ergodisch. □

Angesichts der in diesem Paragraphen vorgestellten Ergebnisse kann geschlossen werden, dass es zu jedem schwach ergodischen stationären Prozess $(X_n)_{n\in\mathbb{N}}$ einen identisch verteilten, entkoppelt tangierenden Prozess $(Y_n)_{n\in\mathbb{N}}$ gibt, der bedingt identisch verteilt bezüglich $\mathscr{I}(X_n, n \in \mathbb{N})$ und bedingt unabhängig bezüglich einer σ-Algebra $\mathscr{G} \subset \mathscr{A}$ ist. Des Weiteren folgt die jeweilige Unabhängigkeit der Zufallsvariablen $Y_n, n \in \mathbb{N}$, von $\mathscr{I}(X_n, n \in \mathbb{N})$, welche aber nicht mit der schwachen Ergodizität des Prozesses $(Y_n)_{n\in\mathbb{N}}$ zu verwechseln ist.

§ 9

Martingale

Zum Abschluss der vorliegenden Arbeit soll der Bereich der Martingale und deren Besonderheiten im Falle der schwachen Ergodizität angesprochen werden. Diese spezielle Klasse der stochastischen Prozesse nimmt seit dem Artikel von Billingsley [4] eine besondere, um nicht zu sagen zentrale Rolle im Umfeld der ergodischen Prozesse ein. In dem angesprochenen Artikel aus dem Jahre 1961 wird die Gültigkeit des zentralen Grenzwertsatzes (CLT) für quadratisch integrierbare und ergodische Martingal-Differenzenfolgen durch den Rückgriff auf Arbeiten von Lindeberg und Lévy im selben Umfeld nachgewiesen. Des Weiteren wurde im Jahre 1985 von Esseen und Janson [17, Theorem 1] eine Möglichkeit vorgestellt, aus gewissen Eigenschaften einer ergodischen Martingal-Differenzenfolge die quadratische Integrierbarkeit zu folgern.

Angesichts dieser allgemeinen Aussagen im Falle der Ergodizität soll in diesem Paragraphen der Versuch unternommen werden, folgende Sachverhalte auf die schwächere Voraussetzung der schwachen Ergodizität zu erweitern:

- Hinreichende Bedingungen für die quadratische Integrierbarkeit.
- Der zentrale Grenzwertsatz (CLT).

Es wird sich zeigen, dass die hinreichenden Bedingungen sowohl für die quadratische Integrierbarkeit, als auch für den zentralen Grenzwertsatze ihre Gültigkeit im Falle der schwachen Ergodizität beibehalten.

Die Gültigkeit des CLT ist eines der zentralen Themen der Stochastik, wie schon der Name verrät. Aufgrund des Artikels von Bradley [8] ist bekannt, dass die Stationarität und paarweise Unabhängigkeit eines stochastischen Prozesses für die Gültigkeit des zentralen Grenzwertsatzes nicht ausreichen. Jedoch waren diese gerade die hinreichenden Bedingungen für die schwache Ergodizität eines stochastischen Prozesses aus Satz 6.1 von Landers und Rogge. Daher ist also die Berücksichtigung zusätzlicher Voraussetzungen für die Gültigkeit des CLT nötig. Der schon erwähnte Satz von Billingsley erscheint in diesem Zusammenhang als hilfreich, denn wie schon andere Sachverhalte in dieser Arbeit belegt haben, lassen sich einige Aussagen über ergodische Prozesse auf schwach ergodische Prozesse ausdehnen. Dies kann auch im vorliegenden Fall geschehen, wie im Weiteren genauer erläutert wird.

In diesem Zusammenhang wird anschließend ein Beispiel für eine schwach ergodische Martingal-Differenzenfolge angegeben, die nicht ergodisch ist. Hierdurch wird die Existenz der vorliegenden Bearbeitung gerechtfertigt.

Bei der Beschäftigung mit Martingal-Differenzenfolgen stößt man auf ein weiteres starkes Gesetz der großen Zahlen (SLLN). Theoretisch hätte man auch dieses in die Überlegungen zu den „konkreten“ Beispielen für schwach ergodische stationäre Prozesse im Paragraphen 6 miteinbeziehen können, jedoch ergeben sich bei dem hier vorliegenden Fall Probleme, wie später in diesem Paragraphen gezeigt wird. Dieses soll dann auch die Betrachtungen in diesem Paragraphen beenden.

Doch zunächst zu den Definitionen eines Martingals.

9.1 Definition
Ein stochastischer Prozess $(X_i)_{i\in\mathbb{I}}$ mit $\mathbb{I} = \{1, 2, \dots, n\}, n \in \mathbb{N}$ oder $\mathbb{I} = \mathbb{N}$ auf einem Wahrscheinlichkeitsraum $(\Omega, \mathscr{A}, P)$ heißt *Martingal*, wenn gilt:

(i) $P(|X_i|) < \infty$, für alle $i \in \mathbb{I}$ und

(ii) $P(X_i \mid X_j, \dots, X_1) \;=\; X_j$ P-f.s., für alle $j < i,\; j, i \in \mathbb{I}$.

9.2 Definition
Ein integrierbarer Prozess $(X_i)_{i\in\mathbb{I}}$ mit $\mathbb{I} = \{1, 2, \dots, n\}, n \in \mathbb{N}$ oder $\mathbb{I} = \mathbb{N}$ auf einem Wahrscheinlichkeitsraum $(\Omega, \mathscr{A}, P)$ heißt *Martingal-Differenzenfolge* (MDF) , wenn gilt:

(i) $P\left(X_1\right) \;=\; 0$ und

(ii) $P\left(X_i \mid X_{i-1}, \dots, X_1\right) \;=\; 0 \quad \left(\forall i \in \mathbb{I} \setminus \{1\}\right).$

9.3 Bemerkung

(a) Sei $(Y_n)_{n\in\mathbb{N}}$ ein integrierbarer stochastischer Prozess, dann ist der Prozess $(X_n)_{n\in\mathbb{N}}$ mit

$$\begin{aligned} X_1 &:= Y_1 \text{ und} \\ X_n &:= X_{n-1} + Y_n - P\left(Y_n \mid X_{n-1}, \ldots, X_1\right),\ n \in \mathbb{N}_{\geq 2}, \end{aligned}$$

ein Martingal.

(b) Sei $(Y_n)_{n\in\mathbb{N}}$ ein Martingal, dann ist der Prozess $(X_n)_{n\in\mathbb{N}}$ mit

$$X_n := Y_{n+1} - Y_n,\ n \in \mathbb{N},$$

eine Martingal-Differenzenfolge.

(c) Ist $(X_n)_{n\in\mathbb{N}}$ eine Martingal-Differenzenfolge, so ist der Prozess $(S_n)_{n\in\mathbb{N}}$ mit

$$S_n := \sum_{i=1}^{n} X_i,\ n \in \mathbb{N},$$

ein Martingal mit $P(S_n) = 0$ für $n \in \mathbb{N}$.

(d) Sei $(X_n)_{n\in\mathbb{N}}$ eine Martingal-Differenzenfolge und gilt $(a_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$. Dann ist der Prozess $(a_n \cdot X_i)_{1\leq i\leq n}$ für jedes $n \in \mathbb{N}$ ebenfalls eine Martingal-Differenzenfolge.

In den folgenden Sätzen dieses Paragraphen, die sich mit der Problematik des CLT beschäftigen, wird stets die quadratische Integrierbarkeit der Martingal-Differenzen gefordert. Im Falle von ergodischen Martingal-Differenzen wurde im Jahre 1985 von Esseen und Janson [17, Theorem 1] eine Möglichkeit vorgestellt, aus gewissen Eigenschaften der Martingal-Differenzenfolge deren quadratische Integrierbarkeit zu folgern.

Die Voraussetzung der Ergodizität lässt sich auf die allgemeinere Forderung der schwachen Ergodizität abschwächen. Somit ergibt sich der folgende Satz:

9.4 Satz

Sei $(X_n)_{n\in\mathbb{N}}$ eine schwach ergodische Martingal-Differenzenfolge. Falls

$$\liminf_{n\to\infty} P\left(\left|\frac{\sum_{k=1}^{n} X_k}{\sqrt{n}}\right|\right) < \infty$$

gilt, so ist jedes X_n, $n \in \mathbb{N}$, quadratisch integrierbar.

Zum Beweis des obigen Satzes benötigt man das folgende Lemma, das z. B. für $C{=}2$ von Burkholder [10] und für $C{=}6$ von Garsia [22] bewiesen wurde.

9.5 Lemma

Es sei $(X_n)_{n\in\mathbb{N}}$ eine Martingal-Differenzenfolge und $S_n := \sum_{k=1}^n X_k$, $n \in \mathbb{N}$. Dann gilt:

$$P\left(\left(\sum_{k=1}^{n}\left({X_k}^2\right)\right)^{\frac{1}{2}} > \lambda\right) \leq \frac{C}{\lambda}\cdot P\left(|S_n|\right),$$

mit einer Konstanten $C \in \mathbb{R}$.

Beweis des Satzes 9.4:

O.B.d.A. nehme man für alle $n \in \mathbb{N}$ $P\left(|S_n/\sqrt{n}|\right) \leq A < \infty$ an, ansonsten betrachte man eine Teilfolge. Mit Lemma 9.5 ergibt sich nun für $n \in \mathbb{N}$ unter Berücksichtigung von $(\lambda \cdot \sqrt{n})$ anstelle von λ:

$$\begin{aligned}
P\left(\left(\frac{1}{n}\cdot\sum_{k=1}^{n}\left({X_k}^2\right)\right)^{\frac{1}{2}} > \lambda\right) &= P\left(\left(\sum_{k=1}^{n}\left({X_k}^2\right)\right)^{\frac{1}{2}} > \lambda\cdot\sqrt{n}\right) \\
&\leq \frac{C}{\lambda\cdot\sqrt{n}}\cdot P\left(|S_n|\right) \\
&\leq \frac{C}{\lambda}\cdot P\left(\left|\frac{S_n}{\sqrt{n}}\right|\right) \\
&\leq \frac{C\cdot A}{\lambda} \\
&< 1, \qquad (9.1)
\end{aligned}$$

für λ groß genug.

Nach dem Ergodensatz für schwach ergodische Prozesse 3.6 und den anschließenden Bemerkungen folgt:

$$\frac{1}{n}\cdot\sum_{k=1}^{n}{X_k}^2 \underset{n\to\infty}{\longrightarrow} P\left({X_1}^2\right) \quad P\text{-f.s.}.$$

Wäre nun $P\left({X_1}^2\right) = \infty$, so folgte auch die Konvergenz von $\left(\frac{1}{n}\cdot\sum_{k=1}^{n}\left({X_k}^2\right)\right)_{n\in\mathbb{N}}$ gegen ∞ und somit würde auch die linke Seite der Gleichung (9.1) gegen 1 konvergieren, welches der Aussage von Gleichung (9.1) widerspräche.

Demnach folgt also $P\left({X_1}^2\right) < \infty$, d. h. es folgt aufgrund der identischen Verteilung der Zufallsvariablen $X_n, n \in \mathbb{N}$, insgesamt die quadratische Integrierbarkeit der X_n, $n \in \mathbb{N}$. □

Nachdem hinreichende Bedingungen für die quadratische Integrierbarkeit einer Martingal-Differenzenfolge gefunden wurden, wird die Frage nach der Gültigkeit des zentralen Grenzwertsatzes für Martingal-Differenzenfolgen behandelt.

Der Versuch den Beweis des zentralen Grenzwertsatzes für ergodische Martingal-Differenzenfolgen $(Y_n)_{n\in\mathbb{N}}$ von Billingsley [4], der auf den Grundzügen der Arbeiten von Lindeberg und Lévy basiert, für eine schwach ergodische Martingal-Differenzenfolge $(X_n)_{n\in\mathbb{N}}$ zu adaptieren, erscheint bei der Nacharbeitung des entsprechenden Beweises als erfolgversprechend, zumal die Voraussetzung der Ergodizität nur an zwei Stellen benötigt wird. Es wird jedoch deutlich, dass letztlich nicht direkt die Ergodizität der Martingal-Differenzenfolge $(X_n)_{n\in\mathbb{N}}$ benutzt wird, sondern es wird auf die besonderen Eigenschaften einer ergodischen maßerhaltenden Transformation zurückgegriffen, die in Bemerkung 2.21 dargelegt wurden.

Die Komplexität der sich theoretisch hierbei ergebenden Bedingung schränken die Nutzbarkeit der Aussage entscheidend ein. Auch die abgewandelte Darstellung des entsprechenden Beweises im Buch von Gänssler und Stute [21] bringt hier nicht den gewünschten Erfolg. Aus diesem Grund wird der zentrale Grenzwertsatz für Doppelfolgen von Hall und Heyde [23, Theorem 3.2] als Grundlage für weitere Betrachtungen herangezogen. Eine ausführlichere Darlegung dieser Zusammenhänge findet man auch im Buch von Liptser und Shiryayev [28, §5].

9.6 Satz

Es sei $(X_{n,i})_{1\leq i\leq n}$ für jedes $n \in \mathbb{N}$ eine Martingal-Differenzenfolge, die jeweils aus quadratisch integrierbaren ZV besteht. Gilt dann für $n \to \infty$:

(i) $\max_{1\leq i\leq n} |X_{n,i}| \xrightarrow{p} 0$,

(ii) $\sum_{i=1}^n {X_{n,i}}^2 \xrightarrow{p} c \in \mathbb{R}_+$ und

(iii) $P\left(\max_{1\leq i\leq n} {X_{n,i}}^2\right)$ ist bezüglich n beschränkt,

so konvergiert $S_n = \frac{1}{\sqrt{c}} \cdot \sum_{i=1}^n X_{n,i}$ für $n \to \infty$ schwach gegen eine $N(0,1)$-verteilte Zufallsvariable.

9.7 Bemerkung

(i) Bei der Bedingung (i) im obigen Satz 9.6 handelt es sich um eine *asymptotische Vernachlässigbarkeitsannahme* („assumption of asymptotic negligibility"). Dem Buch von Hall und Heyde zufolge [23, Abschnitt 3.2] ist die hier angesetzte Forderung äquivalent zur *schwachen Lindeberg-Bedingung*:

$$(\forall \varepsilon > 0) \quad \sum_{i=1}^n {X_{n,i}}^2 \cdot 1_{[|X_{n,i}|>\varepsilon]} \quad \xrightarrow{p} \quad 0 \qquad \text{für } n \to \infty.$$

(ii) Ist $(X_n)_{n\in\mathbb{N}}$ ein integrierbarer und identisch verteilter stochastischer Prozess, so ist die Doppelfolge $(X_{n,i})_{1\leq i\leq n, n\in\mathbb{N}}$, mit $X_{n,i} := \frac{1}{\sqrt{n}} \cdot X_i, 1 \leq i \leq n, n \in \mathbb{N}$, *gleichmäßig asymptotisch vernachlässigbar* (vgl. das Buch von Hall und Heyde [23, 3.2 (i), Seite 53], denn es gilt für jedes $\varepsilon > 0$:

$$\begin{aligned}
\max_{1\le i\le n} P\left(|X_{n,i}| > \varepsilon\right) &= \max_{1\le i\le n} P\left(\left|\frac{1}{\sqrt{n}}\cdot X_i\right| > \varepsilon\right)\\
&= \max_{1\le i\le n} P\left(\left|\frac{1}{\sqrt{n}}\cdot X_1\right| > \varepsilon\right)\\
&= P\left(\left|\frac{1}{\sqrt{n}}\cdot X_1\right| > \varepsilon\right)\\
&\underset{n\to\infty}{\longrightarrow}\ ,\text{ da } X_1 \text{ integrierbar ist.}
\end{aligned}$$

Jedoch ist diese Bedingung i. A. schwächer als die schwache Lindeberg-Bedingung.

Es ergibt sich somit der folgende Satz.

9.8 Satz

Es sei $(X_n)_{n\in\mathbb{N}}$ eine quadratisch integrierbare Martingal-Differenzenfolge und man definiere $\sigma := \sqrt{P\left({X_1}^2\right)}$.

Ist $(X_n)_{n\in\mathbb{N}}$ schwach ergodisch, so konvergiert $\left(\frac{1}{\sigma\cdot\sqrt{n}}\cdot\sum_{k=1}^n X_k\right)_{n\in\mathbb{N}}$ schwach gegen eine $N(0,1)$-verteilte Zufallsvariable.

Beweis:

Nach Satz 9.6 ist für $X_{n,i} := \frac{1}{\sqrt{n}}\cdot X_i$, $1\le i\le n, n\in\mathbb{N}$ zu zeigen:

(a) $\frac{1}{\sqrt{n}}\cdot\max_{1\le i\le n}|X_i| \stackrel{p}{\longrightarrow} 0$,

(b) $\sum_{i=1}^n\left(\frac{1}{\sqrt{n}}\cdot X_i\right)^2 \stackrel{p}{\longrightarrow} c\in\mathbb{R}_+$ und

(c) $P\left(\max_{1\le i\le n}\left(\frac{1}{\sqrt{n}}\cdot X_i\right)^2\right)$ ist bezüglich n beschränkt.

Zu (a) Der Nachweis dieser Eigenschaft erfolgt über den Nachweis der nach Bemerkung 9.7 (i) äquivalenten schwachen Lindeberg-Bedingung.

Es seien nun $\varepsilon > 0$ und $N\in\mathbb{N}$, dann gilt trivialerweise für jedes $n\in\mathbb{N}_{\ge N}$ und $i\in\mathbb{N}$:

$$ {X_i}^2\cdot 1_{\left[|X_i|>\varepsilon\cdot\sqrt{n}\right]} \le {X_i}^2\cdot 1_{\left[|X_i|>\varepsilon\cdot\sqrt{N}\right]} \tag{9.2}$$

Nach Satz 3.4 und Satz 3.6 gilt nun für jedes $\varepsilon > 0$ und jedes $N\in\mathbb{N}$:

$$\begin{aligned}
0 &\leq \liminf_{n\to\infty} \sum_{k=1}^{n} \left(\frac{1}{\sqrt{n}} \cdot X_k\right)^2 \cdot 1_{\left[\left|\frac{1}{\sqrt{n}}\cdot X_k\right|>\varepsilon\right]} \\
&= \liminf_{n\to\infty} \frac{1}{n} \cdot \sum_{k=1}^{n} {X_k}^2 \cdot 1_{\left[|X_k|>\varepsilon\cdot\sqrt{n}\right]} \\
&\leq \limsup_{n\to\infty} \frac{1}{n} \cdot \sum_{k=1}^{n} {X_k}^2 \cdot 1_{\left[|X_k|>\varepsilon\cdot\sqrt{n}\right]} \\
&\underset{(9.2)}{\leq} \limsup_{n\to\infty} \frac{1}{n} \cdot \sum_{k=1}^{n} {X_k}^2 \cdot 1_{\left[|X_k|>\varepsilon\cdot\sqrt{N}\right]} \\
&= \lim_{n\to\infty} \frac{1}{n} \cdot \sum_{k=1}^{n} {X_k}^2 \cdot 1_{\left[|X_k|>\varepsilon\cdot\sqrt{N}\right]} \\
&= P\left({X_1}^2 \cdot 1_{\left[|X_1|>\varepsilon\cdot\sqrt{N}\right]}\right) \quad P\text{-f.s..}
\end{aligned}$$

Somit folgt insgesamt:

$$\begin{aligned}
0 &\leq \liminf_{n\to\infty} \frac{1}{n} \cdot \sum_{k=1}^{n} {X_k}^2 \cdot 1_{\left[\left|\frac{1}{\sqrt{n}}\cdot X_k\right|>\varepsilon\right]} \\
&\leq \limsup_{n\to\infty} \frac{1}{n} \cdot \sum_{k=1}^{n} {X_k}^2 \cdot 1_{\left[\left|\frac{1}{\sqrt{n}}\cdot X_k\right|>\varepsilon\right]} \\
&\leq \inf_{N\in\mathbb{N}} P\left({X_1}^2 \cdot 1_{\left[|X_1|>\varepsilon\cdot\sqrt{N}\right]}\right) \quad P\text{-f.s.} \\
&= 0.
\end{aligned}$$

Also konvergiert die Folge $\left(\frac{1}{n} \cdot \sum_{k=1}^{n} {X_k}^2 \cdot 1_{\left[\left|\frac{1}{\sqrt{n}}\cdot X_k\right|>\varepsilon\right]}\right)_{n\in\mathbb{N}}$ für $n \to \infty$ P-f.s. gegen 0, also folgt auch die gesuchte stochastische Konvergenz dieses Ausdrucks gegen 0.

Zu (b) Es gilt:

$$\begin{aligned}
\sum_{i=1}^{n} \left(\frac{1}{\sqrt{n}} \cdot X_i\right)^2 &= \frac{1}{n} \cdot \sum_{i=1}^{n} {X_i}^2 \\
&\underset{\text{Satz 3.6}}{\longrightarrow} P\left({X_1}^2\right) \quad (< \infty) \quad P\text{-f.s..}
\end{aligned}$$

Also gilt die P-f.s.-Konvergenz von $\sum_{i=1}^{n} \left(\frac{1}{\sqrt{n}} \cdot X_i\right)^2$ gegen $c = P\left({X_1}^2\right)$, somit also auch die entsprechende stochastische Konvergenz.

Zu (c)

$$\begin{aligned}
P\left(\max_{1\leq i\leq n}\left(\frac{1}{\sqrt{n}}\cdot X_i\right)^2\right) &= P\left(\frac{1}{n}\cdot\max_{1\leq i\leq n}{X_i}^2\right)\\
&\leq P\left(\frac{1}{n}\cdot\sum_{i=1}^{n}{X_i}^2\right)\\
&= \frac{1}{n}\cdot P\left(\sum_{i=1}^{n}{X_i}^2\right)\\
&= \frac{1}{n}\cdot\sum_{i=1}^{n}P\left({X_i}^2\right)\\
&= \frac{1}{n}\cdot\sum_{i=1}^{n}P\left({X_1}^2\right)\\
&= P\left({X_1}^2\right)\quad(<\infty).
\end{aligned}$$

Die gesuchte schwache Konvergenz von $\left(\frac{1}{\sigma\cdot\sqrt{n}}\cdot\sum_{k=1}^{n}X_k\right)_{n\in\mathbb{N}}$ gegen eine $N(0,1)$-verteilte Zufallsvariable folgt somit nach Satz 9.6 □

Zur Rechtfertigung der Untersuchungen von schwach ergodischen Prozessen im Umfeld von Martingal-Differenzenfolgen bedarf es eines entsprechenden Beispiels. Es soll eine quadratisch-integrierbare, schwach ergodische Martingal-Differenzenfolge angegeben werden, die nicht ergodisch ist. Hierzu wird das folgende Lemma benötigt.

9.9 Lemma

Es sei $(X_n)_{n\in\mathbb{N}}$ ein stochastischer Prozess auf $(\Omega,\mathscr{A})$, dann ist die Menge aller Wahrscheinlichkeitsmaße auf $(\Omega,\mathscr{A})$, bezüglich derer $(X_n)_{n\in\mathbb{N}}$ eine Martingal-Differenzenfolge ist, konvex. Ist also $(X_n)_{n\in\mathbb{N}}$ bezüglich P_1 und bezüglich P_2 eine MDF, so auch bezüglich jeden Maßes Q mit $Q = a\cdot P_1 + (1-a)\cdot P_2$ für jedes $a\in[0,1]$.

Beweis:

Sei $(X_n)_{n\in\mathbb{N}}$ eine MDF bezüglich P_1 und bezüglich P_2. Weiter sei $a\in[0,1]$ und $Q := a\cdot P_1 + (1-a)\cdot P_2$.

Für den Beweis reicht es nach Definition 9.2 und Definition A.1 zu zeigen:

- Jedes $X_n, n\in\mathbb{N}$ ist bezüglich Q integrierbar.

- Für jedes $n\in\mathbb{N}$ und $A\in\sigma_n := \sigma(X_1,\dots X_{n-1})$ $(\sigma_1 = \{\emptyset,\Omega\})$ gilt:

$$\int_A X_n \,\mathrm{d}Q \;=\; 0.$$

Die Q-Integrierbarkeit ist aufgrund der P_1- und der P_2-Integrierbarkeit offensichtlich, also muss nur der zweite Punkt nachgewiesen werden.

Nun gilt für $n \in \mathbb{N}$ und $A \in \sigma_n$:

$$\begin{aligned}
\left|\int_A X_n \,\mathrm{d}Q\right| &\le \left|a \cdot \int_A X_n \,\mathrm{d}P_1 + (1-a) \cdot \int_A X_n \,\mathrm{d}P_2\right| \\
&\le a \cdot \left|\int_A X_n \,\mathrm{d}P_1\right| + (1-a) \cdot \left|\int_A X_n \,\mathrm{d}P_2\right| \\
&= a \cdot \left|\int_A P_1(X_n \mid \sigma_n) \,\mathrm{d}P_1\right| + (1-a) \cdot \left|\int_A P_2(X_n \mid \sigma_n) \,\mathrm{d}P_2\right| \\
&= a \cdot \left|\int_A 0 \,\mathrm{d}P_1\right| + (1-a) \cdot \left|\int_A 0 \,\mathrm{d}P_2\right| \\
&= 0.
\end{aligned}$$

Somit folgt also für jedes $A \in \sigma_n$:

$$\int_A X_n \,\mathrm{d}Q = 0.$$

Daher gilt insgesamt die Behauptung des Lemmas. □

Mit diesem Lemma kommt man zu folgendem Beispiel.

9.10 Beispiel

Auf $(\mathbb{R}^{\mathbb{N}}, \mathscr{B}^{\mathbb{N}})$ betrachte man die Projektionen $(\pi_n)_{n\in\mathbb{N}}$ als stochastischen Prozess. Weiter wähle man ein Maß R auf $(\mathbb{R}^{\mathbb{N}}, \mathscr{B}^{\mathbb{N}})$ bezüglich dessen $(\pi_n)_{n\in\mathbb{N}}$ eine quadratisch-integrierbare, ergodische Martingal-Differenzenfolge bildet, deren Komponenten π_n, $n \in \mathbb{N}$, bezüglich R nicht unabhängig sind. Die Existenz eines solchen Maßes ist nach Literatur gesichert. Hierzu vergleiche man Ausführungen aus dem Buch von Hall und Heyde [23] zur Gültigkeit des CLT für quadratisch-integrierbare, ergodische Martingal-Differenzenfolgen . Weiter sei angemerkt, dass sich die entsprechende Fachliteratur über quadratisch-integrierbare, ergodische Martingal-Differenzenfolgen für den Fall, dass ein entsprechendes Maß R nicht existieren würde, stets mit unabhängigen und identisch verteilten Zufallsvariablen beschäftigen würde.

Weiter wähle man ein Maß Q, sodass $(\pi_n)_{n\in\mathbb{N}}$ bezüglich Q quadratisch-integrierbar, identisch verteilt und unabhängig ist, also ebenfalls eine ergodische MDF bildet. Zusätzlich soll $R_{\pi_1} = Q_{\pi_1}$ gelten. Die Existenz dieses zweiten Maßes ist ebenfalls laut Literatur gesichert, man denke hier z. B. an das Produktmaß.

Man definiere sich nun weiter das folgende Maß P auf $(\mathbb{R}^{\mathbb{N}}, \mathscr{B}^{\mathbb{N}})$:

$$P := \frac{1}{2} \cdot R + \frac{1}{2} \cdot Q.$$

Es ist nun nachzuweisen, dass der Prozess $(\pi_n)_{n\in\mathbb{N}}$ bezüglich P eine quadratisch-integrierbare, schwach ergodische MDF bildet, die nicht ergodisch ist.

Quadratische Integrierbarkeit: Die quadratische Integrierbarkeit von $(\pi_n)_{n\in\mathbb{N}}$ bezüglich P folgt offensichtlich aus der quadratischen Integrierbarkeit von $(\pi_n)_{n\in\mathbb{N}}$ bezüglich R und bezüglich Q.

Stationarität: Seien hierzu $B \in \mathscr{B}_1{}^{\mathbb{N}}$ und $n \in \mathbb{N}$, dann folgt mit der Stationarität von $(\pi_n)_{n\in\mathbb{N}}$ bezüglich R und bezüglich Q:

$$\begin{aligned} P\left((\pi_1, \pi_2, \dots) \in B\right) &= \left(\frac{1}{2} \cdot R + \frac{1}{2} \cdot Q\right)\left((\pi_1, \pi_2, \dots) \in B\right) \\ &= \frac{1}{2} \cdot R\left((\pi_1, \pi_2, \dots) \in B\right) \;+\; \frac{1}{2} \cdot Q\left((\pi_1, \pi_2, \dots) \in B\right) \\ &= \frac{1}{2} \cdot R\left((\pi_n, \pi_{n+1}, \dots) \in B\right) \;+\; \frac{1}{2} \cdot Q\left((\pi_n, \pi_{n+1}, \dots) \in B\right) \\ &= P\left((\pi_n, \pi_{n+1}, \dots) \in B\right). \end{aligned}$$

Schwache Ergodizität: Aufgrund der Voraussetzungen ist der Prozess $(\pi_n)_{n\in\mathbb{N}}$ schwach ergodisch und nach Lemma 2.22 reicht es, hierfür die Unabhängigkeit von π_1 und $\mathscr{I}(\pi_n,\ n \in \mathbb{N})$ zu zeigen. Seien also $B \in \mathscr{B}$ und $I \in \mathscr{I}(X_n,\ n \in \mathbb{N})$. Wegen $R_{\pi_1} = Q_{\pi_1}$ folgt insbesondere auch:

$$P(\pi_1 \in B) \;=\; R(\pi_1 \in B) \;=\; Q(\pi_1 \in B). \tag{9.3}$$

Also folgt:

$$\begin{aligned} P([\pi_1 \in B] \cap I) &= \left(\frac{1}{2} \cdot R + \frac{1}{2} \cdot Q\right)([\pi_1 \in B] \cap I) \\ &= \frac{1}{2} \cdot R([\pi_1 \in B] \cap I) \;+\; \frac{1}{2} \cdot Q([\pi_1 \in B] \cap I) \\ &= \frac{1}{2} \cdot R([\pi_1 \in B]) \cdot R(I) \;+\; \frac{1}{2} \cdot Q([\pi_1 \in B]) \cdot Q(I) \\ &\underset{(9.3)}{=} \frac{1}{2} \cdot P([\pi_1 \in B]) \cdot R(I) \;+\; \frac{1}{2} \cdot P([\pi_1 \in B]) \cdot Q(I) \\ &= P([\pi_1 \in B]) \cdot \frac{1}{2} \cdot (R(I) + Q(I)) \\ &= P([\pi_1 \in B]) \cdot P(I). \end{aligned}$$

Daher gilt die Unabhängigkeit von π_1 und $\mathscr{I}(\pi_n,\ n \in \mathbb{N})$ und somit die schwache Ergodizität.

Martingal-Differenzenfolge: Da es sich bei $(\pi_n)_{n\in\mathbb{N}}$ um eine Martingal-Differenzenfolge sowohl bezüglich R als auch bezüglich Q handelt, folgt mit Lemma 9.9 und $a = \frac{1}{2}$ sofort, dass es sich auch um eine MDF bezüglich P handelt.

Es ist noch der Nachweis zu führen, dass es sich nicht um einen ergodischen Prozess handelt.

Nicht-Ergodizität: Nach den Voraussetzungen handelt es sich bei $(\pi_n)_{n\in\mathbb{N}}$ um einen stochastichen Prozess, der bezüglich R und bezüglich Q ergodisch ist, jedoch stimmen R und Q nicht überein. Somit existiert nach einem Korollar aus dem Buch von Breiman [9, Corollary 6.24] bzw. Korollar 3.8 und der anschließenden Bemerkung, eine invariante Menge $I_0 \in \mathscr{I}(\pi_n,\ n \in \mathbb{N})$, sodass gilt:

$$R(I_0) \ = \ 1 \quad \text{und} \quad Q(I_0) = 0.$$

Somit folgt aber auch:

$$P(I_0) \ = \ \frac{1}{2} \cdot R(I_0) \ + \ \frac{1}{2} \cdot Q(I_0) \ = \ \frac{1}{2}.$$

$\mathscr{I}(\pi_n,\ n \in \mathbb{N})$ ist daher bezüglich P nicht trivial, also ist der Prozess $(\pi_n)_{n\in\mathbb{N}}$ nicht ergodisch.

Insgesamt konnte also nachgewiesen werden, dass es quadratisch-integrierbare, schwach ergodische MDF gibt, die nicht ergodisch sind.

Vor dem Hintergrund der Frage nach einer schwach ergodischen MDF, die nicht ergodisch ist, der Darstellung von schwach ergodischen Prozessen in Paragraph 5 mit Hilfe des starken Gesetzes der großen Zahlen und der Existenz eines SLLN für MDF stellt sich die Frage, ob auch gewisse MDF aufgrund dieser Zusammenhänge als schwach ergodische Prozesse erkannt werden können.

Hierzu sei zuerst das starke Gesetz der großen Zahlen für Martingale [11] angegeben, das von Chow im Jahre 1965 veröffentlicht wurde.

9.11 Satz
Sei $(X_n)_{n\in\mathbb{N}}$ eine Martingal-Differenzenfolge.
Gilt dann:

$$\sum_{i=1}^{\infty} \frac{P\left(|X_i|^2 \mid X_{i-1}, X_{i-2}, \ldots, X_1\right)}{i^2} \ < \ \infty, \tag{9.4}$$

so folgt:

$$\frac{1}{n} \cdot \sum_{i=1}^{n} X_i \ \longrightarrow \ 0 \ \ (= P(X_1)) \ P\text{-f.s.}$$

Bei der Einarbeitung dieses Sachverhaltes in den Beweisgang von Landers und Rogge von Satz 6.1 benötigt man jedoch nicht direkt die Eigenschaft, dass es sich beim Prozess $(X_n)_{n\in\mathbb{N}}$ um eine Martingal-Differenzenfolge handeln muss, sondern man benötigt die Voraussetzung, dass die Prozesse $\left(\varphi_k^x \circ X_n\right)_{n\in\mathbb{N}}, k \in \mathbb{N}, x \in \mathbb{R}$, für die die Funktionenfolgen $(\varphi_k^x)_{k\in\mathbb{N}}, x \in \mathbb{R}$ jeweils Folgen von $]0,1[$-wertigen, messbaren, injektiven Funktionen sind, die punktweise gegen $\varphi^x = 1_{]-\infty,x]}$ konvergieren, Martingal-Differenzenfolgen sind (vgl. Paragraph 5). Somit folgt nach Definition sofort, dass der jeweilige Erwartungswert 0 ist. Dies bedeutet aber doch, dass für jedes $k \in \mathbb{N}$ und jedes $x \in \mathbb{R}$ $\varphi_k^x \circ X_1 = 0$ sein muss.

Somit wird deutlich, dass eine Nutzung dieser Version des SLLN für die Darstellung spezieller Beispiele von schwach ergodischen Prozessen nach Vorgabe der Beweisidee von Landers und Rogge nicht möglich ist.

Mit diesen Ausführungen sollen die Untersuchungen und Betrachtungen im Umfeld von schwach ergodischen stationären Prozessen beendet werden.

§ 10

Zusammenfassung

Die vorliegende Betrachtung der schwach ergodischen Prozesse, ihrer Darstellung und Eigenschaften und die Untersuchung einzelner Spezialfälle wird zum Abschluss noch einmal kurz zusammengefasst.

Der Begriff der schwachen Ergodizität erweitert den bisherigen Begriffsapparat im Umfeld der stationären Prozesse. Wie schon im Artikel von Landers und Rogge [27] mit dem Beispiel von Cuesta und Matrán [14] gezeigt wurde, handelt es sich hierbei um eine echte Erweiterung der bekannten Begriffe, da die definierende Bedingung eine Abschwächung der bisherigen Voraussetzungen darstellt.

Angesichts der Darstellung schwach ergodischer Prozesse in Paragraph 5 und den anschließenden Beispielen für schwach ergodische Prozessse in Paragraph 6 können neue starke Gesetze der großen Zahlen aufgestellt werden. Wie schon in der Einleitung geschildert, handelt es sich beim Ergodensatz von Birkhoff im (schwach) ergodischen Fall (vgl. Satz 3.6) um das *Strong law of large numbers*. Die Klassen von stochastischen Prozessen, die aufgrund der vorliegenden Bearbeitung als schwach ergodisch identifiziert wurden, erfüllen somit auch das SLLN.

In den ersten Paragraphen dieser Arbeit wurde schon deutlich, dass die Verallgemeinerung der Voraussetzung „ergodischer stationärer Prozess“ hin zur Voraussetzung „schwach ergodischer stationärer Prozess“ nicht bei allen Anwendungen möglich ist. Dieses wird insbesondere in Satz 3.4 und dem anschließenden Beispiel deutlich. Diese Einschränkungen machen es daher unmöglich, einige interessante und bedeutende Ergebnisse aus der allgemeinen Ergodentheorie auf die neue Klasse von stationären Prozessen zu erweitern. Als Beispiel sei hier, wie bereits am Ende von Paragraph 3, nur angegeben, dass bei einem ergodischen Prozess auch die entsprechenden Rekkurenzzeiten wieder ergodisch sind (vgl. [9]). Ebenso zeigt sich, dass gewisse Aussagen aus der Ergodentheorie auch nur auf einer abstrakteren Ebene übernommen werden können. Dies zeigt sich insbesondere bei der schwachen Ergodizität bezüglich zweier Maße in Satz 3.8.

Gewisse Klassen stochastischer Prozesse weisen eine weitere Besonderheit auf. Die neue Eigenschaft der schwachen Ergodizität erweist sich für spezielle stationäre Prozesse als äquivalent zur bekannten Ergodizität, dies wurde insbesondere für vertauschbare Prozesse in Paragraph 4 und für Markov-Prozesse in Paragraph 7 nachgewiesen. Im ersten Fall folgte diese Tatsache aufgrund der dort vorliegenden Unabhängigkeit und identischen Verteilung der einzelnen Zufallsvariablen und im zweiten Fall aufgrund der speziellen Darstellung der invarianten resp. der P-f.s.-invarianten Mengen des entsprechenden Prozesses.

Abschließend wurden noch Bedingungen für die Gültigkeit des CLT für schwach ergodische Prozesse aufgestellt. Auch in diesem Beispiel, das im Rahmen der Martingal-Differenzenfolgen auftauchte, wurden einige Aussagen, deren Gültigkeit bereits für ergodische stationäre Prozesse nachgewiesen worden war, auf schwach ergodische stationäre Prozesse erweitert. Somit zeigte sich, dass die bisherigen Bedingungen nur hinreichend, aber nicht notwendig waren. Gerade an diesem Beispiel zeigte sich, dass über die (schwache) Ergodentheorie weitere Ergebnisse in anderen Forschungszweigen erreicht werden konnten. Es konnte hier eine Alternative des CLT für paarweise unabhängige Zufallsvariablen angegeben werden, die in der Tradition der Untersuchungen zu paarweise unabhängigen ZV und zu vorzeichen-invarianten ZV steht, wie sie z. B. im Artikel von Hong [25] publiziert wurden.

Insgesamt kann keine allgemein gültige Aussage über die Anwendbarkeit und Nützlichkeit des neuen Begriffes gemacht werden. Wie die obigen Schilderungen und die Ausführungen der vorliegenden Arbeit deutlich macht, ist hierfür der jeweilige Anwendungszweck von entscheidender Bedeutung. In zahlreichen Fällen sind die schwächeren Bedingungen von schwach ergodischen Prozesse bereits hinreichend für das gesuchte Ergebnis, jedoch kann diese Aussage resp. Wertung nicht verallgemeinert werden.

Anhang A

Mathematische Ergänzungen

Bedingte Erwartung

An dieser Stelle soll die Definition eines bedingten Erwartungswertes und der bedingten Wahrscheinlichkeit bezüglich einer σ-Algebra $\mathscr{G}$ gegeben werden.

A.1 Definition
Es sei X eine integrierbare Zufallsvariable auf $(\Omega, \mathscr{A}, P)$ und $\mathscr{G}$ eine Unter-σ-Algebra von $\mathscr{A}$.

(a) Eine $\mathscr{G}$-messbare Zufallsvariable Y heißt die *bedingte Erwartung von X bezüglich $\mathscr{G}$*, wenn für jedes $G \in \mathscr{G}$ gilt:

$$\int_G Y \, \mathrm{d}P \; = \; \int_G X \, \mathrm{d}P.$$

Die Zufallsvariable Y wird dann mit $P(X \mid \mathscr{G})$ bezeichnet.
Je zwei verschiedene bedingte Erwartungen bezüglich derselben σ-Algebra stimmen P-f.s. überein.

(b) Für ein Ereignis $A \in \mathscr{A}$ heißt

$$P(A \mid \mathscr{G}) \; := \; P(1_A \mid \mathscr{G})$$

die *bedingte Wahrscheinlichkeit von A bezüglich $\mathscr{G}$*.

Im Folgenden sind erste Eigenschaften dieser bedingten Erwartung notiert. Gerade im Umfeld von schwach ergodischen Prozessen sind die bedingten Wahrscheinlichkeiten von Interesse, da sie als Indikator für die Unabhängigkeit von zwei σ-Algebren dienen können.

A.2 Bemerkung

Für Zufallsvariablen X, Y, die entweder beide nicht negativ oder beide integrierbar sind und eine Unter-σ-Algebra $\mathscr{G} \subset \mathscr{A}$, gelten die folgenden Aussagen:

(a) Für $\alpha, \beta \in \mathbb{R}$ gilt:

$$P(\alpha \cdot X + \beta \cdot Y \mid \mathscr{G}) \;=\; \alpha \cdot P(X \mid \mathscr{G}) + \beta \cdot P(Y \mid \mathscr{G}) \quad P\text{-f.s.}.$$

(b) Ist die Zufallsvariable X $\mathscr{G}$-messbar, so folgt:

$$P(X \mid \mathscr{G}) \;=\; X \quad P\text{-f.s.}.$$

(c) Ist $X = Y$ P-f.s., so gilt auch:

$$P(X \mid \mathscr{G}) \;=\; P(Y \mid \mathscr{G}) \quad P\text{-f.s.}.$$

(d) Ist die Zufallsvariable X unabhängig von $\mathscr{G}$, so gilt:

$$P(X \mid \mathscr{G}) \;=\; P(X) \quad P\text{-f.s.}.$$

(e) Zwei Unter-σ-Algebren $\mathscr{G}$ und $\mathscr{F}$ von $\mathscr{A}$ sind genau dann voneinander unabhängig, wenn für jedes $G \in \mathscr{G}$ gilt:

$$P(G \mid \mathscr{F}) \;=\; P(G) \quad P\text{-f.s.}$$

(vgl. Bauer [3, Seite 128]).

A.3 Lemma

(a) Konvergiert der stochastische Prozess $(X_n)_{n\in\mathbb{N}}$ P-f.s. gegen die ZV X und existiert eine integrierbare ZV Y mit $|X_n| \leq Y$ für alle $n \in \mathbb{N}$, dann gilt:

$$\lim_{n\to\infty} P(X_n \mid \mathscr{G}) \;=\; P(X \mid \mathscr{G}) \quad P\text{-f.s.},$$

für eine Unter-σ-Algebra $\mathscr{G} \subset \mathscr{A}$.

(b) Gilt $\mathscr{G}_1 \subset \mathscr{G}_2$ für zwei Unter-σ-Algebren $\mathscr{G}_1, \mathscr{G}_2$ von $\mathscr{A}$, so folgt für eine ZV X:

$$P\left(P\left(X \mid \mathscr{G}_1\right) \mid \mathscr{G}_2\right) \;=\; P\left(P\left(X \mid \mathscr{G}_2\right) \mid \mathscr{G}_1\right) \;=\; P\left(X \mid \mathscr{G}_1\right) \quad P\text{-f.s.}.$$

Dynkin-Systeme

A.4 Definition
Es sei $\Omega \neq \emptyset$, $\emptyset \neq \mathscr{D} \subset \mathfrak{P}(\Omega)$.
$\mathscr{D}$ heißt *Dynkin-System*, wenn gilt:

(1) $\Omega \in \mathscr{D}$,

(2) $D_1, D_2 \in \mathscr{D},\ D_1 \subset D_2 \ \Rightarrow\ D_2 \setminus D_1 \in \mathscr{D}$ und

(3) $D_n \in \mathscr{D}, D_n \subset D_{n+1}$ für alle $n \in \mathbb{N} \ \Rightarrow\ \bigcup_{n=1}^{\infty} D_n \in \mathscr{D}$.

A.5 Definition
Ein System $\mathscr{E} \subset \mathfrak{P}(\Omega)$ heißt *durchschnittsstabil* ($\cap$-stabil), wenn gilt:

$$E_1, E_2 \in \mathscr{E} \ \Rightarrow\ E_1 \cap E_2 \in \mathscr{E}.$$

A.6 Satz
Es sei $\Omega \neq \emptyset$, $\emptyset \neq \mathscr{E} \subset \mathfrak{P}(\Omega)$. Dann gilt:

(a) $\mathscr{E}$ $\cap$-stabiles Dynkin-System $\Leftrightarrow$ $\mathscr{E}$ σ-Algebra.

(b) $\mathscr{E}$ $\cap$-stabil $\Rightarrow$ $D(\mathscr{E}) = \sigma(\mathscr{E})$.

Anhang B

Symbolik

Die folgende Liste enthält Symbole und Notationen, die im Verlauf dieser Arbeit von Bedeutung sind.
Hierbei seien $a, b, c \in \mathbb{R}$, $(\Omega, \mathscr{A}, P)$ ein Wahrscheinlichkeitsraum, $(\Omega_1, \mathscr{A}_1)$ ein Messraum und $f : (\Omega, \mathscr{A}) \to (\Omega_1, \mathscr{A}_1)$ eine messbare Abbildung. Des Weiteren seien $A, B \in \mathscr{A}$, $A_1 \in \mathscr{A}_1$, $(X_n)_{n\in\mathbb{N}}$ sei ein stochastischer Prozess auf $(\Omega, \mathscr{A}, P)$ und $I \subset \mathbb{N}$ eine beliebige Indexmenge.

Symbol	**Bedeutung**
$\mathbb{N}$	die natürlichen Zahlen
$\mathbb{N}_{\geq y}$	$\{x \in \mathbb{N} \mid x \geq y\}$
$\mathbb{N}_0$	$\mathbb{N} \cup \{0\}$
$\mathbb{R}$	die reellen Zahlen
$\mathbb{Q}$	die rationalen Zahlen
$[c, d]$	abgeschlossene Intervallgrenzen
$]c, d[$	offene Intervalgrenzen
$a[b] = c$	Modulo-Abbildung, d.h. $a[b] = c \quad \Leftrightarrow \quad (\exists d \in \mathbb{R}) \;\; a = d \cdot b + c$

Fortsetzung auf der nächsten Seite

Fortsetzung der vorherigen Seite	
$[a]$	Gauss-Klammern, d.h. $[a] = \max\{h \in \mathbb{N} \mid h \leq a\}$
1_A	die charakteristische Funktion von A, d.h. $1_A(\omega) = \begin{cases} 1 & \text{, für } \omega \in A \\ 0 & \text{, für } \omega \notin A \end{cases}$
$\mathfrak{P}(\Omega)$	Potenzmenge von Ω
$\sigma(X_i, i \in I)$	kleinste σ-Algebra, die die von $X_i, i \in I$, erzeugten σ-Algebren enthält
$D(X_i, i \in I)$	kleinstes Dynkin-System, das die von $X_i, i \in I$, erzeugten Dynkin-Systeme umfasst
P_{X_i}	Verteilung von X_i bezüglich P
$\underset{\|\cdot\|_1}{\longrightarrow}$	Konvergenz im ersten Mittel, d. h. Konvergenz in der 1-Norm.
$\xrightarrow{p}$	stochastische Konvergenz.
$\xrightarrow{w}$	schwache Konvergenz.
$P - \lim$	Grenzwert bezüglich der stochstischen Konvergenz
$A \triangle B$	$(A \setminus B) \cup (B \setminus A)$

Literaturverzeichnis

[1] Ash, R.B.: *Real Analysis and Probability.* First Edition. Academic Press, New York, San Francisco, London, 1972

[2] Ash, R.B. ; Gardner, M.F.: *Topics in stochastic processes.* First Edition. New York, San Francisco, London : Academic Press, 1975

[3] Bauer, H.: *Probability Theory.* First Edition. Berlin, New York : de Gruyter Studies in Mathematics, 1996

[4] Billingsley, P.: The Lindeberg-Lévy Theorem for Martingales. In: *Americal Mathematical Society Proceedings* 12 (1961), S. 788–792

[5] Billingsley, P.: *Convergence of Probability Measures.* New York, Chichester, Brisbane, Toronto : John Wiley & Sons, 1968

[6] Birkhoff, G.D.: Proof of the Ergodic Theorem. In: *Proc. nat. Acad. Sci. U.S.A.* 17 (1931), S. 656–660

[7] Blum, J.R. ; Hanson, D.L. ; Koopmans, L.H.: On the strong law of large numbers for a class of stochastic processes. In: *Z. Wahrsch. Verw. Geb.* 2 (1963), S. 1–11

[8] Bradley, R.C.: A stationary, pairwise independent, absolutely regular sequence for which the central limit theorem fails. In: *Probability Theory and Rel. F.* 81 (1989), S. 1–10

[9] Breiman, L.: *Probability.* First Edition. Reading, Massachusetts : Addison Wesley Publishing Company, 1968

[10] Burkholder, D.L.: A sharp inequality for martingale transforms. In: *Ann. Probability* 7 (1979), S. 858–863

[11] Chow, Y.S.: Local Convergence of Martingales and the Law of Large Numbers. In: *Ann. Math. Statist.* 36 (1965), S. 552–558

[12] Chow, Y.S. ; Teicher, H.: *Probability Theory.* Second Edition. New York : Springer Verlag, 1988

[13] Cohn, D.L.: *Measure Theory.* Boston und Stuttgart : Birkhäuser, 1980

[14] Cuesta, J.A. ; Matrán, C.: On the asymptotic behavior of sums of pairwise independent random variables. In: *Statistics and Prob. Lett.* 11 (1991), S. 201–210

[15] Doukhan, P.: *Mixing.* New York, Berlin, Heidelberg : Springer Verlag, 1994

[16] Dudley, R. M.: *Real Analysis and Probability.* Pacific Grove, California : Wadsworth & Brooks, 1989

[17] Esseen, C. G. ; Janson, S.: On moment conditions for normed sums of independent variables and martingale differences. In: *Stoch. proc. and their appl.* 19 (1985), S. 173–182

[18] Etemadi, N.: An elementary proof of the strong law of large numbers. In: *Z. Wahrsch. Verw. Geb.* 55 (1983), S. 187–193

[19] Etemadi, N.: On the laws of large numbers for nonnegative random variables. In: *Journal of Multivariate Analysis* 13 (1983), S. 187–193

[20] Etemadi, N. ; Kaminski, M.: Strong Law of Large Numbers for 2-Exchangeable Random Variables. In: *Statistics and Probability Letters* 28 (1996), S. 245–250

[21] Gänssler, P. ; Stute, W.: *Wahrscheinlichkeitstheorie.* Berlin : Springer-Verlag, 1977

[22] Garsia, A.M.: *Martingales Inequalities.* Reading, MA : Benjamin Inc., 1973

[23] Hall, P. ; Heyde, C.C.: *Martingale limit theory and its applications.* First Edition. New York, London, Toronto, Sydney, San Francisco : Academic Press, 1980

[24] Hess, K.-T.: *Bedingt unabhängige und identisch verteilte Folgen von Zufallsvariablen*, Technische Universität Dresden, Dissertation, 1997

[25] Hong, Dug H.: A Remark on the CLT for Sums of Pairwise I.I.D. Random Variables. In: *Math. Japonica* 42 (1995), S. 87–89

[26] Hu, T.-C.: On pairwise independent and independent exchangeable random variables. In: *Stochastic Analysis and Applications* 15 (1) (1997), S. 51–57

[27] Landers, D. ; Rogge, L.: Weak ergodicity of stationary pairwise independent processes. In: *Proc. Am. Math. Soc.* 128, No. 4 (2000), S. 1203–1206

[28] Liptser, R. S. ; Shiryayev, A.N.: *Theory of Martingales*. Dordrecht, Boston, London : Kluwer Academic Publishers, 1989

[29] Olshen, R.A.: The coincidence of measure algebras under an exchangeable probability. In: *Z. Wahrsch. verw. Geb.* 18 (1971), S. 153–158

[30] Peña, V. de l. ; Giné, E.: *Decoupling*. New York, Berlin, Heidelberg : Springer-Verlag, 1999

[31] Petersen, K.E.: *Ergodic Theory*. Cambridge : Cambridge University Press, 1983

[32] Stout, W.F.: *Almost sure convergence*. New York, San Francisco, London : Academic Press, 1974

Schlagwortverzeichnis

Danksagung

An dieser Stelle möchte ich die Gelegenheit nutzen und mich von ganzem Herzen bei meinen Eltern, Günter und Maria-Theresia Bahne, bedanken. Ohne Ihren bedingungslosen Beistand und moralische Unterstützung wäre mir ein Hochschulstudium und anschliessendes Promotionsstudium nicht möglich gewesen. Hierfür und für Alles, was Sie bisher für mich getan haben, möchte ich Ihnen meinen aufrichtigen und herzlichen Dank sagen.

In fachlicher Hinsicht sei Herrn Prof. Dr. Lothar Rogge mein Dank für seine ausdauernde und intensive Betreuung dieser Dissertation ausgesprochen. Die von ihm aufgeworfenen Fragen und die hieraus resultierenden Diskussionen führten zu weiteren Forschungsansätzen und waren stets eine wertvolle Hilfe.

Weiter möchte ich mich auch bei Herrn Prof. Dr. Günter Törner bedanken, der mir trotz aufwendiger Projektarbeiten stets die Möglichkeit eingeräumt hat, die Bearbeitung dieser Dissertation fortzusetzen.

Last but not least sei auch Frau Christiane Boßmann und Herrn Dipl.-Ing. Marcus Schmidt ein besonderer Dank für Ihre Unterstützung bei der Korrektur der Rechtschreibung dieser Arbeit ausgesprochen.

Bei ihnen und allen, die mich in meinem Vorhaben bestärkt und hierbei unterstützt haben, möchte ich mich bedanken.

Erklärung

Hiermit erkläre ich, Thorsten Bahne, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht.

Die Arbeit wurde weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Oberhausen, den 11. November 2002

(Thorsten Bahne)