Ingo Bürk

Skript vom Wintersemester 2011/2012

Stochastische Prozesse

Mathematics, rightly viewed, possesses not only truth, but supreme beauty.
Bertrand Russell

Universität Stuttgart
2011

Dieses Skript entstand im Rahmen der Vorlesung „Stochastische Prozesse“ bei Prof. Ingo Steinwart als Vorlesungsmitschrieb.

Es kann nicht garantiert werden, dass dieses Dokument fehlerfrei ist und der Autor übernimmt für möglicherweise entstandene Schäden jeglicher Art keine Haftung. Dieser Mitschrieb ist kein offizielles Dokument der Universität Stuttgart, Mitarbeiter eben dieser tragen daher ebenfalls keine Verantwortung.

Mit freundlichen Grüßen
Ingo Bürk

Inhaltsverzeichnis

Vorwort

Mittels eines *stochastischen Prozesses* beschreibt man in der Mathematik zeitlich geordnete, zufällige Vorgänge. Die Theorie dieser Prozesse erweitert die Wahrscheinlichkeitstheorie und öffnet das Tor zur stochastischen Analysis.

1
Übersicht und Einführung

In diesem Kapitel wollen wir die grundlegenden Begriffe für die stochastischen Prozesse definieren, eine kurze Einführung in die Thematik geben und einfache Eigenschaften herleiten.

1.1. Stochastische Prozesse

Im ersten Schritt wollen wir uns den namensgebenden Begriff anschauen und festlegen, was wir unter einem stochastischen Prozess verstehen:

Definition 1.1.1 **Stochastischer Prozess**
Sei $(\Omega, \mathscr{A}, P)$ ein Wahrscheinlichkeitsraum, $(\mathscr{X}, \mathscr{B})$ ein Messraum und $T \neq \emptyset$. Dann heißt eine Familie $X = (X_t)_{t \in T}$ mit für alle $t \in T$ messbaren Funktionen $X_t : \Omega \to \mathscr{X}$ ein stochastischer Prozess.

Da der Begriff die Grundlage für die ganze Thematik darstellt, werden wir desweiteren einige Eigenschaften kennenlernen, um verschiedene Arten von stochastischen Prozessen zu unterscheiden.

Klassifikation 1.1.2
Wir legen folgende klassifizierenden Eigenschaften für stochastische Prozesse fest:

- Ist $\mathscr{X} = \mathbf{R}$, so sprechen wir von reellwertigen stochastischen Prozessen.
- Ein stochastischer Prozess mit endlichem/abzählbarem Zustandsraum $\mathscr{X}$ heißt selbst endlich/abzählbar.
- Ist $\mathscr{X} = \mathbf{R}^d$, so sprechen wir von einem Punktprozess.
- Ist $T \in \{\mathbf{N}, \mathbf{N}_0, \mathbf{Z}\}$, so nennen wir den Prozess einen zeitdiskreten stochastischen Prozess.

- Falls $T \subset \mathbf{R}$ ein Intervall ist, so sprechen wir von einem zeitkontinuierlichen stochastischen Prozess.

Wir werden uns im Wesentlichen jedoch auf zeitdiskrete/kontinuierliche stochastische Prozesse mit Zustandsraum $\mathbf{R}$ oder höchstens abzählbaren Mengen beschränken, es gibt in der Theorie jedoch noch viel mehr Fälle wie z. B. Mengen von Funktionen als Zustandsraum.

Beispiel *Bereits bekannte stochastische Prozesse*
Stochastische Prozesse sind zu diesem Zeitpunkt keine völlig neuartigen Objekte, wir kennen bereits folgende Vertreter:

- $(X_n)_{n\in\mathbf{N}}$ mit i. i. d. Zufallsvariablen X_n.
- $(\overline{X_n})_{n\in\mathbf{N}}$ für eine Folge von i. i. d. Zufallsvariablen $(X_n)_{n\in\mathbf{N}}$ mit $\overline{X_n} = \frac{1}{n}\sum_{i=1}^n X_i$.
- $(X_n^*)_{n\in\mathbf{N}}$ für i. i. d. Zufallsvariablen X_n mit $X_n \in \mathscr{L}_2$ und $\sigma^2 := \operatorname{Var} X_1 > 0$, sowie

$$X_n^* := \frac{1}{\sqrt{n\sigma^2}} \sum_{i=1}^n (X_i - \mathbf{E}X_i).$$

Seien $(X_i)_{i\in\mathbf{N}}$ i. i. d. mit $X_1 \in \mathscr{L}_2$ und $\operatorname{Var} X_1 > 0$, dann ist $(\overline{X_n})_{n\in\mathbf{N}}$ weder unabhängig noch identisch verteilt und $(X_n^*)_{n\in\mathbf{N}}$ nicht unabhängig, aber eventuell identisch verteilt. //

Definition 1.1.3 **Pfad/Trajektorie**
Sei $X = (X_t)_{t\in T}$ ein stochastischer Prozess, dann heißt für $\omega \in \Omega$ die Abbildung

$$\begin{aligned} X(\omega)\colon T &\to \mathscr{X}, \\ t &\mapsto X_t(\omega) \end{aligned}$$

Pfad oder Trajektorie von X bezüglich ω.

Stochastische Prozesse erzeugen also zufällige Abbildungen von T in den Zustandsraum $\mathscr{X}$. Abhängig von der Klasse des Prozesses (zeitdiskret, zeitkontinuerlich, ...) ist die erzeugte Funktion mitunter eine Folge oder eine „normale“ reellwertige Funktion.

Lemma 1.1.4
Wir betrachten die Menge $\mathscr{X}^T := \times_{t\in T} \mathscr{X}$ der Abbildungen $T \to \mathscr{X}$, ausgestattet mit der Produkt-σ-Algebra $\mathscr{B}^T := \bigotimes_{t\in T} \mathscr{B}$. Ferner sei $X = (X_t)_{t\in T}$ ein $\mathscr{X}$-wertiger stochastischer Prozess über $(\Omega, \mathscr{A}, P)$. Dann ist

$$\begin{aligned} X\colon \Omega &\to \mathscr{X}^T, \\ \omega &\mapsto X(\omega) = (t \mapsto X_t(\omega)) \end{aligned}$$

eine messbare Abbildung, das heißt X ist eine $\mathscr{X}^T$-wertige Zufallsvariable.

Beweis: Die Aussage des Lemma folgt unmittelbar aus [WTSkript11, Lemma I.9.16]. □

Wir haben nun zwar einen neuen Begriff, allerdings noch kein Ziel, das wir anstreben. Daher wollen wir nun auf einige typische Fragestellungen eingehen. Häufig wird untersucht, ob die Trajektorien beschränkt sind, wie ihr Wachstumsverhalten aussieht, ob sie für $t \to \infty$ konvergieren, ob sie stetig sind, wann ihre Austrittszeiten sind (d. h. wann verlässt der Pfad ein gewisses Intervall) et cetera.

Stochastische Prozesse finden in vielen Gebieten breite Anwendung. Beispiele für typische Anwendungsfelder für stochastische Prozesse sind:

i) *Finanzmathematik* – Bewertung von Finanzprodukten, Risikomangement, Investitionsstrategien

ii) *Physik* – Diffusionssysteme, stochastische Thermodynamik, Quantenphysik

iii) *Biologie* – Populationsmodelle, „Genomics“

iv) *Ingenieurswissenschaften* – Warteschlangenprobleme, Steuerungsprobleme

1.2. Einfache Eigenschaften

Wir wollen zunächst einige einfache Eigenschaften stochastischer Prozesse untersuchen. Die erste Eigenschaft bezieht sich auf die Gleichheit stochastischer Prozesse, dazu seien $X = (X_t)_{t\in T}$ und $Y = (Y_t)_{t\in T}$ zwei $\mathscr{X}$-wertige stochastische Prozesse über dem gleichen Wahrscheinlichkeitsraum $(\Omega, \mathscr{A}, P)$. Als Abbildungen $\Omega \times T \to \mathscr{X}$ würden wir Gleichheit durch $X_t(\omega) = Y_t(\omega)$ für alle $t \in T$ und alle $\omega \in \Omega$ definieren. Da wir mit stochastischen Prozessen nach Lemma 1.1.4 jedoch auch Zufallsvariablen vorliegen haben, wollen wir auch andere Begriffe für die Unterscheidbarkeit einführen.

Definition 1.2.1 **Unterscheidbarkeit**
Die wie eben definierten stochastischen Prozesse X und Y heißen nicht unterscheidbar genau dann, wenn für P-fast alle $\omega \in \Omega$ und für alle $t \in T$ die Gleichheit $X_t(\omega) = Y_t(\omega)$ gilt. Dies ist genau dann der Fall, wenn $P(\{\omega : X(\omega) = Y(\omega)\}) = 1$ gilt.
Die Prozesse sind also nicht unterscheidbar, wenn fast alle Pfade gleich sind.

Die Aussage ist also, dass Prozesse nicht unterscheidbar sind, wenn es eine von t unabhängige Nullmenge gibt, außerhalb welcher die Pfade übereinstimmen. Wir wollen jetzt gewissermaßen die Reihenfolge der Quantoren vertauschen und zulassen, dass diese Nullmenge von t abhängt.

Definition 1.2.2 **Version**
Seien X und Y wieder wie eben definiert. Dann heißt X Version von Y (und Y Version von X) genau dann, wenn

$$P(\{\omega : X_t(\omega) = Y_t(\omega)\}) = 1 \qquad \text{für alle } t \in T$$

gilt, das heißt wir fordern Gleichheit nur noch punktweise für $t \in T$.

Lemma 1.2.3
Seien X und Y die zwei obigen stochastischen Prozesse. Dann gilt:

i) X und Y sind nicht unterscheidbar impliziert, dass X eine Version von Y ist.

ii) Ist X eine Version von Y und T abzählbar, so sind X und Y nicht unterscheidbar.

Beweis: Der Beweis wird in den Übungen geführt. □

Wir wollen uns nun noch davon überzeugen, dass diese beiden Begriffe im Allgemeinen wirklich verschieden sind. Die dahinterstehende Idee wird sein, dass eine überabzählbare Vereinigung von Nullmengen im Allgemeinen keine Nullmenge mehr ist.

Beispiel 1.2.4
Im Allgemeinen gilt die Umkehrung von i) aus Lemma 1.2.3 nicht. Wir setzen $T := [0,1]$, $\Omega := [0,1]$ und P sei die Gleichverteilung. Ferner sei $X_t(\omega) := 0$ für alle $t \in T$ und $\omega \in \Omega$, sowie $Y_t(\omega) = \begin{cases} 0 & \text{für } t \neq \omega \\ 1 & \text{für } t = \omega \end{cases}$. Dann gilt für $t \in [0,1]$

$$P(\{\omega : X_t(\omega) \neq Y_t(\omega)\}) = P(\{t\}) = 0,$$

also ist X eine Version von Y. Außerdem gilt für jedes $\omega \in \Omega$

$$(t \mapsto X_t(\omega)) = X(\omega) \neq Y(\omega) = (t \mapsto Y_t(\omega)),$$

also sind X und Y unterscheidbar. //

Definition 1.2.5 **Stetigkeit**
Sei $X = (X_t)_{t\in T}$ ein zeitkontinuierlicher stochastischer Prozess. Dann heißt X links-/rechts- bzw. stetig genau dann, wenn P-fast alle Pfade (links-/rechts-)stetig sind, das heißt für P-fast alle $\omega \in \Omega$ ist die Abbildung $X(\omega)\colon T \to \mathbf{R}$ (links/rechts-)stetig.

Ist X (links-/rechts-)stetig, so folgt, dass ein nicht unterscheidbarer stochastischer Prozess Y existiert, für welchen *alle* Pfade (links-/rechts-)stetig sind. Dazu übernimmt man in Y einfach die stetigen Pfade und ersetzt nicht-stetige $X(\omega)$ durch die Nullfunktion. Nach Voraussetzung geschieht dies nur auf einer Nullmenge, womit X und Y nicht unterscheidbar sind.

Stetige stochastische Prozesse sind Zufallsvariablen, deren Bild fast sicher im Raum $C(T)$ der stetigen Funktionen über T liegt. Wie eben gesehen, finden wir damit einen nicht unterscheidbaren Prozess, dessen Bild komplett in diesem Raum liegt.

Satz 1.2.6
Seien X, Y (links-/rechts-)stetige stochastische Prozesse und Y eine Version von X. Dann sind X und Y nicht unterscheidbar.

Jeder stochastische Prozess hat also bis auf Ununterscheidbarkeit höchstens eine stetige Version.

Beweis: Ohne Einschränkung seien X und Y stetige Prozesse und $T = \mathbf{R}$. Für stetige Funktionen $f, g\colon \mathbf{R} \to \mathbf{R}$ gilt die Äquivalenz

$$f = g \;\Leftrightarrow\; f(x) = g(x) \forall_{x\in\mathbf{Q}},$$

es genügt also $P\left(\{\omega : X_t(\omega) = Y_t(\omega) \forall_{t\in\mathbf{Q}}\}\right) = 1$ zu zeigen. Dies folgt jedoch bereits aus Lemma 1.2.3 ii) für $T' = \mathbf{Q}$. □

Satz 1.2.7
Sei $X = (X_t)_{t \in T}$ ein $\mathbf{N}_0$-wertiger und zeitkontinuierlicher stochastischer Prozess. Dann sind folgende Aussagen äquivalent:

i) X ist rechtsstetig.

ii) Für P-fast alle $\omega \in \Omega$ und alle $t \in T$ existiert ein $\varepsilon > 0$ mit $X_s(\omega) = X_t(\omega)$ für alle $s \in [t, t+\varepsilon) \cap T$.

Ändert sich also für einen rechtsstetigen Prozess die Trajektorie, so verbleibt sie danach für eine gewisse Zeit in diesem Zustand.

Beweis: Für die Richtung i) $\Rightarrow$ ii) sei $\omega \in \Omega$ mit $X(\omega)$ rechtsstetig und $t \in T$. Wegen der rechtsseitigen Stetigkeit existiert ein $\varepsilon > 0$ mit $|X_t(\omega) - X_s(\omega)| < \frac{1}{2}$ für alle $s \in [t, t+\varepsilon)$ und wegen $X_t(\omega) \in \mathbf{N}_0$ und $X_s(\omega) \in \mathbf{N}_0$ folgt daher $X_t(\omega) = X_s(\omega)$. Die andere Richtung des Satzes ist trivial. $\square$

Wir geben an dieser Stelle eine Klassifikation der Pfade für obige ($\mathbf{N}_0$-wertige) stochastische Prozesse:

i) Der Pfad hat nur endlich viele Sprungstellen.

ii) Der Pfad hat unendlich viele Sprungstellen, aber nur endlich viele in jedem beschränkten Zeitintervall.

iii) Der Pfad hat unendlich viele Sprungstellen in (mindestens) einem beschränkten Zeitintervall.

1.3. Endlichdimensionale Randverteilungen

In diesem Abschnitt wollen wir stochastische Prozesse konstruieren und Gleichheit für unterschiedliche Wahrscheinlichkeitsräume diskutieren. Hierfür benötigen wir jedoch die namensgebenden endlichen Randverteilungen.

Definition 1.3.1 **Randverteilung**
Sei $X = (X_t)_{t\in T}$ ein $\mathbf{R}$-wertiger stochastischer Prozess über dem Wahrscheinlichkeitsraum $(\Omega, \mathscr{A}, P)$. Ferner seien $n \geq 1$ und $t_1, \ldots, t_n \in T$. Dann nennen wir $P_{X_{t_1},\ldots,X_{t_n}}$, die durch

$$P_{X_{t_1},\ldots,X_{t_n}}(A) := P\left(\left(X_{t_1}, \ldots, X_{t_n}\right) \in A\right) \qquad \text{für alle } A \in \mathscr{B}^n$$

definiert war, die Randverteilung von X bezüglich $t_1, \ldots, t_n$. Wir werden hierfür die kürzere Schreibweise $P_{t_1,\ldots,t_n}$ verwenden.

Satz 1.3.2
Sei X ein $\mathbf{R}$-wertiger stochastischer Prozess über $(\Omega, \mathscr{A}, P)$. Dann erfüllen die Randverteilungen die folgenden Eigenschaften:

i) Permutationsinvarianz, das heißt für jedes $n \geq 1$, jede Permutation $\pi \colon \{1, \ldots, n\} \to \{1, \ldots, n\}$ und alle $t_1, \ldots, t_n \in T$ gilt

$$P_{t_{\pi(1)},\ldots,t_{\pi(n)}}(A) = P_{t_1,\ldots,t_n}\left(\pi^{-1}(A)\right) \qquad \text{für alle } A \in \mathscr{B}^n,$$

wobei $\pi^{-1}(A) := \{(x_1, \ldots, x_n) : (x_{\pi(1)}, \ldots, x_{\pi(n)}) \in A\}$ ist.

ii) Konsistenz, das heißt für alle $n \geq m \geq 1$ und alle $t_1, \ldots, t_n \in T$ gilt

$$P_{t_1,\ldots,t_m}(A) = P_{t_1,\ldots,t_n}\left(A \times \mathbf{R}^{n-m}\right) \qquad \text{für alle } A \in \mathscr{B}^m.$$

Die erste Eigenschaft besagt also, dass das gleichzeitige Permutieren der t_i und A keinen Effekt hat, während die zweite Eigenschaft ausdrückt, dass niederdimensionale Randverteilungen durch höherdimensionale Randverteilungen ausgedrückt werden können.

Beweis: Wir beweisen beide Eigenschaften getrennt:

i) Es genügt, Mengen der Form $A = (-\infty, a]$ für $a = (a_1, \ldots, a_n) \in \mathbf{R}^n$ zu betrachten, da diese $\cap$-stabile Erzeugendensysteme sind (vgl. [WTSkript11, Korollar I.5.3]). Es gilt

$$\pi^{-1}(A) = \left\{(x_1, \ldots, x_n) : x_{\pi(i)} \leq a_i \quad \forall_{i=1,\ldots,n}\right\}.$$

Daraus folgt dann

$$\begin{aligned} P_{t_{\pi(1)},\ldots,t_{\pi(n)}}(A) &= P\left(X_{t_{\pi(i)}} \le a_i \quad \forall_{i=1,\ldots,n}\right) \\ &= P_{t_1,\ldots,t_n}\left(\pi^{-1}(A)\right). \end{aligned}$$

ii) Diese Aussage folgt wegen

$$\begin{aligned} P_{t_1,\ldots,t_n}(A \times \mathbf{R} \times \ldots \times \mathbf{R}) &= P\left(X_{t_i} \in A \quad \forall_{i=1,\ldots,m} \quad \text{und} \quad X_{t_i} \in \mathbf{R} \quad \forall_{i=m+1,\ldots,n}\right) \\ &= P_{t_1,\ldots,t_m}(A). \end{aligned}$$

□

Satz 1.3.3 **Existenzsatz von Kolmogorov**
Sei $T \neq \emptyset$ und $\mathscr{P} = \left(P_{t_1,\ldots,t_n}\right)_{t_1,\ldots,t_n \in T}$ eine Familie von Wahrscheinlichkeitsmaßen, so dass $P_{t_1,\ldots,t_n}$ ein Wahrscheinlichkeitsmaß auf $\mathbf{R}^n$ ist. Ferner sei $\mathscr{P}$ permutationsinvariant und konsistent. Dann gibt es einen $\mathbf{R}$-wertigen stochastischen Prozess $X = (X_t)_{t \in T}$, für welchen die Familie $\mathscr{P}$ die Familie der endlichdimensionalen Randverteilungen ist, das heißt

$$P_{X_{t_1},\ldots,X_{t_n}} = P_{t_1,\ldots,t_n} \quad \text{für alle } n \ge 1 \text{ und } t_1,\ldots,t_n \in T.$$

Permutationsinvariante, konsistente Familien definieren also einen stochastischen Prozess. Wir werden den Satz hier nicht beweisen, da der Beweis zu sehr in andere Fachgebiete reicht. Er findet sich in [Meintrup04, Korollar A.20].

In der Literatur wird häufig $\mathscr{P} = (P_I)_{I \subset T}$ mit endlichen I betrachtet. Die Permutationsinvarianz ist hier gewissermaßen bereits eingebaut und wird daher nicht mehr extra erwähnt.

Definition 1.3.4 **Gleiche Randverteilung**
Seien $X = (X_t)_{t \in T}$ und $Y = (Y_t)_{t \in T}$ zwei $\mathbf{R}$-wertige stochastische Prozesse über $(\Omega, \mathscr{A}, P)$ bzw. $(\Omega', \mathscr{A}', P')$. Dann haben X und Y die gleichen Randverteilungen genau dann, wenn

$$P_{X_{t_1},\ldots,X_{t_n}} = P_{Y_{t_1},\ldots,Y_{t_n}} \quad \text{für alle } n \ge 1 \text{ und } t_1,\ldots,t_n \in T$$

gilt. Wir schreiben hierfür $X \stackrel{d}{=} Y$.

Wenn X eine Version von Y ist, so folgt insbesondere $X \stackrel{d}{=} Y$. Wir überlassen den Beweis hierfür jedoch dem Leser. Der stochastische Prozess aus Satz 1.3.3 ist bis auf „$\stackrel{d}{=}$“ eindeutig.

Definition 1.3.5 **Stationärer Prozess**
Ein **R**-wertiger stochastischer Prozess $(X_t)_{t\in T}$ heißt stationär genau dann, wenn T bezüglich + abgeschlossen ist und für alle $n \geq 1$ und alle $t_1, \dots, t_n, s \in T$ gilt:

$$P_{t_1+s,\dots,t_n+s} = P_{t_1,\dots,t_n}$$

Randverteilungen ändern sich bei stationären stochastischen Prozessen also nicht mit der Zeit.

Definition 1.3.6 **Zuwächse**
Ein **R**-wertiger stochastischer Prozess $(X_t)_{t\in T}$ hat

i) unabhängige Zuwächse genau dann, wenn für alle $n \geq 1$ und $t_0 < t_1 < \ldots < t_n$ die Zufallsvariablen $X_{t_1} - X_{t_0}, \dots, X_{t_n} - X_{t_{n-1}}$ unabhängig sind.

ii) stationäre Zuwächse genau dann, wenn T bezüglich + abgeschlossen ist und für alle $r, s, t \in T$ gilt:

$$X_{r+s+t} - X_{r+t} \stackrel{\mathrm{d}}{=} X_{r+s} - X_r$$

Das heißt der Zuwachs innerhalb der Zeit s ändert sich nicht durch zeitliche Verschiebung.

1.4. Einfache Beispiele

Beispiel 1.4.1 *Unabhängig und identisch verteilt*
Sei $X = (X_n)_{n\geq 1}$ i. i. d. Dann gilt:

i) Die Randverteilungen sind gerade die Produktmaße (siehe auch „Kanonisches Modell“ in [WTSkript11, Satz II.2.5]).

ii) X ist stationär.

iii) X hat stationäre Zuwächse.

iv) X hat im Allgemeinen keine unabhängigen Zuwächse.

Die Aussagen ii) und iii) folgen aus i), iv) wird dem Leser zum Beweis überlassen. //

Beispiel 1.4.2 *Irrfahrt*
Es sei $(Y_n)_{n\geq 0}$ ein i. i. d. stochastischer Prozess mit $Y_n \sim \mathrm{B}\left(1, \frac{1}{2}\right)$. Dann heißt $X = (X_n)_{n\geq 0}$ mit $X_n := \sum_{i=0}^{n} Y_i$ für $n \geq 0$ symmetrische Irrfahrt. Ist $Y_n \sim \mathrm{B}(1,p)$ für $p \in [0,1]$, so sprechen wir von einer asymmetrischen Irrfahrt. Der Prozess X hat stationäre und unabhängige Zuwächse, ist jedoch nicht stationär.

Zunächst gilt für $m < n$, dass

$$X_n - X_m = \sum_{i=m+1}^{n} Y_i \tag{*}$$

ist. Da die Y_i unabhängig sind, folgt die Unabhängigkeit der Zuwächse aus [WTSkript11, Satz II.2.7, Satz II.2.9] und der Tatsache, dass kein Index i doppelt vorkommt. Die Stationarität der Zuwächse folgt ebenfalls aus (*). Da $X_n \sim \mathrm{B}(n,p)$ gilt sind die eindimensionalen Randverteilungen nicht gleich, d. h. $P_n \neq P_{n+1}$.
In der Regel wird die symmetrische Irrfahrt jedoch anders definiert, da man in beide Richtungen gehen können möchte. Man setzt hierfür $Z_n := \sum X_n - 1$. Dies entspricht der Summe von Zufallsvariablen mit Werten in $\{\pm 1\}$.

Die Irrfahrt wird u. a. zur Untersuchung von einfachen Wettspielen verwendet. Beispielsweise gebe es zwei Spieler und eine Münze, die wiederholt geworfen wird. Falls sie Kopf zeigt, so verliert Spieler 1 und zahlt seinem Kontrahenten einen Euro, entsprechend umgekehrt für den Fall „Zahl“. Man kann sich nun fragen, wie lange es dauert, bis einer der Spieler pleite ist (ein gewisses Startkapital sei gegeben) oder wie groß die Wahrscheinlichkeit dafür ist, dass Spieler 1 verliert et cetera. //

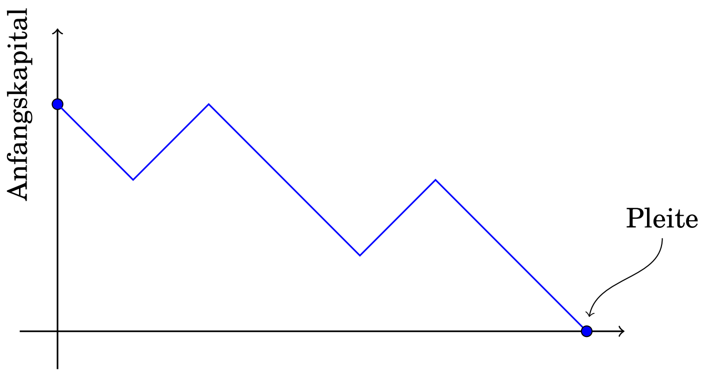

Abbildung 1.1.: Darstellung einer Wettspiel-Irrfahrt für einen Spieler.

Beispiel 1.4.3 *Gleitendes Mittel*
Es sei $(X_n)_{n\in\mathbf{Z}}$ ein $\mathbf{R}$-wertiger stochastischer Prozess, $k \in \mathbf{N}$ und $c_0,\ldots,c_k \in [0,1]$ mit $\sum_{i=0}^k c_i = 1$. Wir definieren nun $Y = (Y_n)_{n\in\mathbf{Z}}$ vermöge $Y_n := \sum_{i=0}^k c_i X_{n-i}$. Dann heißt Y gleitendes Mittel von X bezüglich der Gewichtung $c_0,\ldots,c_k$. Gleitende Mittel glätten den Prozess X gewissermaßen und werden unter anderem bei der Zeitreihenanalyse eingesetzt (z. B. Aktienkurse). Der Prozess Y ist stationär. Gilt zudem, dass $(X_n)_{n\in\mathbf{Z}}$ i. i. d. mit $X_i \in \mathscr{L}_2$ ist, so ist $\operatorname{Var} Y_n \leq \operatorname{Var} X_n$, wobei die echte Ungleichung in der Mehrzahl der Fälle gilt. Der Beweis für diese Eigenschaften wird dem Leser überlassen. //

Beispiel 1.4.4 *Untypisches Beispiel*
Es sei $P = \bigotimes_{i=1}^\infty \lambda_{[0,1]}$ auf $[0,1]^{\mathbf{N}}$, ausgestattet mit der Produkt-σ-Algebra. Wir definieren für $\omega \in [0,1]$, also eine Folge $\omega = (\omega_i)_{i\geq 1}$, und $t \in [0,\infty)$ die Zufallsvariablen

$$X_t(\omega) := \sum_{i=1}^\infty \frac{\omega_i}{i!} t^i.$$

Die Reihe konvergiert absolut und gleichmäßig auf allen Kompakta. Ferner bildet $X = (X_t)_{t\geq 0}$ einen stochastischen Prozess und jede Trajektorie $X(\omega) = (t \mapsto X_t(\omega))$ ist eine analytische Funktion. Ist also $X_t(\omega)$ für alle $t \in (t_1,t_2)$ bekannt, so ist bereits die gesamte Trajektorie bekannt. //

Solche Prozesse wie in Beispiel 1.4.4 interessieren uns hier allerdings nicht, weshalb wir sie nicht mehr betrachten werden. Stattdessen haben die Prozesse, die wir betrachten werden, eine wesentlich lockerere Beziehung zwischen Vergangenheit und Zukunft. Dafür müssen wir jedoch erst mehr Theorie entwickeln.

1.5. Bedingte Erwartungen

Dieses Kapitel ist von *zentraler* Wichtigkeit für alles, was wir machen werden. Wir verweisen an dieser Stelle daher explizit auf [Meintrup04, Kapitel 8] und [Klenke06].

Bisher hatten wir bedingte Wahrscheinlichkeiten durch $P(B \mid A) := \frac{P(A \cap B)}{P(A)}$ definiert, sofern $P(A) > 0$ gilt. Die Interpretation war, dass $P(B \mid A)$ die Wahrscheinlichkeit für B beschreibt, falls A eintritt. Dies wollen wir auf zwei Weisen verallgemeinern:

i) Es soll $P(A) = 0$ erlaubt sein.

ii) Es soll mehr als ein Ereignis A erlaubt sein.

Der Hintergrund hierfür ist folgender: Sind X und Y zwei $\mathbf{R}$-wertige Zufallsvariablen, so wollen wir Fragen beantworten wie zum Beispiel:

i) Wie groß ist die Wahrscheinlichkeit von $Y \in B$, falls $X(\omega)$ eintritt? Die Wahrscheinlichkeiten solcher Elementarereignisse verschwinden sehr oft, weshalb wir die erste Verallgemeinerung benötigen.

ii) Wie groß ist die Wahrscheinlichkeit von $Y \in B$, falls für $\omega \in \Omega$ und jedes $A \in \mathscr{B} = \sigma(X)$ bekannt ist, ob $X(\omega) \in A$ gilt?

Wir wählen die „heuristische“ Herangehensweise. Es sei $(\Omega, \mathscr{A}, P)$ ein Wahrscheinlichkeitsraum, $Y \in \mathscr{L}_1(P)$, $A \in \mathscr{A}$ mit $P(A) > 0$ und $X : \Omega \to \mathbf{R}$ messbar. Wir setzen nun

$$\mathbf{E}(Y \mid A) := \frac{\mathbf{E}(Y \cdot \mathbf{1}_A)}{P(A)} \tag{*}$$

und verallgemeinern dadurch $P(B \mid A)$, denn es gilt

$$\begin{aligned}\mathbf{E}(\mathbf{1}_B \mid A) &= \frac{\mathbf{E}(\mathbf{1}_B \cdot \mathbf{1}_A)}{P(A)} = \frac{\mathbf{E}(\mathbf{1}_{A \cap B})}{P(A)} = \frac{P(A \cap B)}{P(A)} \\ &= P(B \mid A).\end{aligned}$$

Nun betrachten wir für $x \in \mathbf{R}$ die Menge $A_x := \{\omega : X(\omega) = x\}$. Dafür nehmen wir an, dass X diskret verteilt ist, das heißt, es existieren höchstens abzählbar viele $x \in X$ mit $P(A_x) > 0$ und $\sum_{x \in X} P(A_x) = 1$. Nun definieren wir $g : \mathbf{R} \to \mathbf{R}$ vermöge

$$g(x) := \begin{cases} \mathbf{E}(Y \mid A_x) & \text{für } P(A_x) > 0 \\ 0 & \text{sonst} \end{cases}$$

und darauf aufbauend $Z := g \circ X$, wir erhalten also das Diagramm in 1.2. Damit folgt nun, dass Z $\sigma(X)$-messbar ist und überdies gilt für alle $B \in \sigma(X)$:

$$\mathbf{E}(Z \cdot \mathbf{1}_B) = \int_B Z \, \mathrm{d}P = \int_B Y \, \mathrm{d}P = \mathbf{E}(Y \cdot \mathbf{1}_B).$$

$$\begin{tikzcd}
\Omega \arrow[rr, "Z"] \arrow[dr, "X"'] & & \mathbf{R} \\
& \mathbf{R} \arrow[ur, "g"'] &
\end{tikzcd}$$

Abbildung 1.2.: Darstellung der Beziehung zwischen den Räumen und Abbildungen.

Also ist Z eine $\sigma(X)$-messbare Zufallsvariable, die sich bezüglich Integration von Y über $\sigma(X)$ nicht unterscheidet.

Beweis: Wir müssen zunächst zeigen, dass Z tatsächlich $\sigma(X)$-messbar ist. Dazu betrachten wir

$$Z^{-1}(\mathscr{B}) = (g \circ X)^{-1}(\mathscr{B}) = X^{-1}(g^{-1}(\mathscr{B})) \subset X^{-1}(\mathscr{B}) = \sigma(X),$$

wobei die Inklusion gilt, da g messbar ist. Für die Eigenschaft bezüglich der Integration sei zunächst $B \in \sigma(X)$, dann existiert ein $A \in \mathscr{B}$ mit $B = X^{-1}(A)$. Nun erhalten wir mit Hilfe des Transformationssatzes

$$\int_B Z \,\mathrm{d}P = \int g \circ X \cdot \mathbf{1}_A \circ X \,\mathrm{d}P = \int g \cdot \mathbf{1}_A \,\mathrm{d}P_X = \sum_{x \in A} g(x) P(A_x)$$

Wegen $E(Y \mid A_x) \cdot P(A_x) = \mathbf{E}(Y \cdot \mathbf{1}_{A_x})$ erhalten wir

$$= \sum_{x \in A, P(A_x) > 0} \mathbf{E}(Y \cdot \mathbf{1}_{A_x}) = \mathbf{E}\left(Y \sum_{x \in A, P(A_x) > 0} \mathbf{1}_{A_x}\right)$$

Da $\bigcup_{x \in A} \{\omega : X(\omega) = x\} = B$ gilt, folgt

$$= \mathbf{E}(Y \cdot \mathbf{1}_B).$$

□

Im Weiteren werden wir Z mit dem eben Gesehenem definieren bzw. konstruieren und so die zweite Frage beantworten, während uns die Faktorisierung $Z = g \circ X$ die erste Frage beantworten wird.

Definition 1.5.1 Bedingte Erwartung
Es sei $(\Omega, \mathscr{A}, P)$ ein Wahrscheinlichkeitsraum, $Y \in \mathscr{L}_1(P)$ und $\mathscr{B} \subset \mathscr{A}$ eine σ-Algebra. Dann heißt eine Zufallsvariable $Z : \Omega \to \mathbf{R}$ bedingte Erwartung von Y unter $\mathscr{B}$, falls gilt:

i) Z ist $\mathscr{B}$-messbar.

ii) Für alle $B \in \mathscr{B}$ gilt

$$\mathbf{E}(Y \cdot \mathbf{1}_B) = \mathbf{E}(Z \cdot \mathbf{1}_B).$$

Ist $\mathscr{B} = \sigma(X)$ für eine Zufallsvariable $X : \Omega \to \mathbf{R}$, so sind die beiden Eigenschaften der heuristischen Herangehensweise erfüllt. Zunächst müssen wir uns jedoch fragen, ob es solche Objekte überhaupt gibt und ob sie eindeutig sind.

Satz 1.5.2 **Existenz und Eindeutigkeit**
Es sei $(\Omega, \mathscr{A}, P)$ ein Wahrscheinlichkeitsraum, $Y \in \mathscr{L}_1(P)$ und $\mathscr{B} \subset \mathscr{A}$ eine σ-Algebra. Dann gilt:

i) Es existiert eine bedingte Erwartung Z von Y unter $\mathscr{B}$.

ii) Sind Z und Z' zwei bedingte Erwartungen von Y unter $\mathscr{B}$, so gilt P-fast sicher $Z = Z'$.

Bevor wir dies beweisen, wollen wir die Notation $\mathbf{E}(Y \mid \mathscr{B}) := Z$ einführen. Wir nennen Z Version der bedingten Erwartung. Falls $\mathscr{B} = \sigma(X)$ für eine Zufallsvariable $X : \Omega \to \mathbf{R}$ gilt, so schreiben wir $\mathbf{E}(Y \mid X) := \mathbf{E}(Y \mid \sigma(X))$. Es gilt zu beachten, dass $\mathbf{E}(\mathbf{E}(Y \mid \mathscr{B}) \cdot \mathbf{1}_B) = \mathbf{E}(Y \cdot \mathbf{1}_B)$ für alle $B \in \mathscr{B}$ gilt.

Beweis: Wir beweisen zunächst die Existenz und betrachten den Fall, dass $Y \geq 0$ gilt. Dann definieren wir $\nu : \mathscr{B} \to [0, \infty)$ durch $\nu(B) := \int_B Y \, \mathrm{d}P$ für $B \in \mathscr{B}$. Dies ist ein endliches Maß. Ferner ist $P|_{\mathscr{B}} : \mathscr{B} \to [0,1]$ ein Wahrscheinlichkeitsmaß auf $\mathscr{B}$ und es gilt: aus $P|_{\mathscr{B}}(B) = 0$ folgt $\nu(B) = 0$. Also ist $\nu \ll P|_{\mathscr{B}}$. Der Satz von Radon-Nikodym (vgl. [WTSkript11, Satz I.12.1]) gibt uns dann eine Dichte $Z : \Omega \to \mathbf{R}$ von ν bezüglich $P|_{\mathscr{B}}$. Diese ist nach Konstruktion sogar $\mathscr{B}$-messbar und für $B \in \mathscr{B}$ gilt:

$$\int_B Z \, \mathrm{d}P = \int_B \underbrace{Z \, \mathrm{d}P|_{\mathscr{B}}}_{=\nu} = \nu(B) = \int_B Y \, \mathrm{d}P$$

Im zweiten Fall sei $Y = Y^+ - Y^-$ mit $Y^+ := \max\{0, Y\} \geq 0$ und $Y^- := \max\{0, -Y\} \geq 0$. Mit Hilfe des ersten Falls erhalten wir $Z^+, Z^- \geq 0$ und setzen nun $Z := Z^+ - Z^-$.

Wir kommen nun zur Eindeutigkeit. Es seien Z und Z' zwei bedingte Erwartungen von Y unter $\mathscr{B}$. Dann gilt $\int_B Z \, \mathrm{d}P = \int_B Z' \, \mathrm{d}P$ für alle $B \in \mathscr{B}$. Wir betrachten im Speziellen die Menge $B := \{Z > Z'\}$, die offenbar messbar ist. Für B gilt nun

$$0 = \int_B \underbrace{Z - Z'}_{\geq 0} \, \mathrm{d}P,$$

woraus wir $P(B) = 0$ und damit P-fast sicher $Z \leq Z'$ erhalten. Eine symmetrische Betrachtung durch Vertauschen der Rollen von Z und Z' liefert schließlich P-fast sicher $Z = Z'$. □

Satz 1.5.3 **Elementare Eigenschaften**
Es sei $(\Omega, \mathscr{A}, P)$ ein Wahrscheinlichkeitsraum, $X, Y \in \mathscr{L}_1(P)$ und $\mathscr{B} \subset \mathscr{A}$ eine σ-Algebra. Dann gelten die folgenden Eigenschaften:

i) $\mathbf{E}(\mathbf{E}(X \mid \mathscr{B})) = \mathbf{E}X$.

ii) Ist X überdies $\mathscr{B}$-messbar, so gilt P-fast sicher $\mathbf{E}(X \mid \mathscr{B}) = X$.

iii) Für $\alpha, \beta \in \mathbf{R}$ gilt P-fast sicher

$$\mathbf{E}(\alpha X + \beta Y \mid \mathscr{B}) = \alpha \mathbf{E}(X \mid \mathscr{B}) + \beta \mathbf{E}(Y \mid \mathscr{B}).$$

iv) Gilt P-fast sicher $X \leq Y$, so gilt auch P-fast sicher $\mathbf{E}(X \mid \mathscr{B}) \leq \mathbf{E}(Y \mid \mathscr{B})$.

Insbesondere gilt unter den Voraussetzungen im Satz also auch $|\mathbf{E}(X \mid \mathscr{B})| \leq \mathbf{E}(|X| \mid \mathscr{B})$.

Beweis: Wir werden einige der Eigenschaften beweisen und die anderen dem Leser überlassen:

i) Für $B := \Omega \in \mathscr{B}$ und $Y' := \mathbf{E}(X \mid \mathscr{B})$ folgt P-fast sicher $\mathbf{E}Y' = \mathbf{E}(Y' \cdot \mathbf{1}_B) = \mathbf{E}(X \cdot \mathbf{1}_B) = \mathbf{E}X$.

ii) Dies folgt aus der Definition der bedingten Erwartung (Definition 1.5.1) und der Eindeutigkeit (Satz 1.5.2), denn dann ist $\mathbf{E}(X \mid \mathscr{B}) = \mathbf{E}(\mathbf{E}(X \mid \mathscr{B}) \cdot \mathbf{1}_B) = \mathbf{E}(X \cdot \mathbf{1}_B)$ für alle $B \in \mathscr{B}$.

iii) Dieser Beweis wird dem Leser überlassen.

iv) Es genügt wegen iii) zu zeigen, dass aus $X \geq 0$ folgt, dass $\mathbf{E}(X \mid \mathscr{B}) \geq 0$ gilt. Dies wurde jedoch bereits im Beweis von Satz 1.5.2 im ersten Fall des Existenzbeweises gezeigt. □

Wir wollen nun noch zeigen, dass die aus der Wahrscheinlichkeitstheorie bekannten Konvergenzsätze im Wesentlichen auch für bedingte Erwartungen gelten.

Satz 1.5.4 **Konvergenzsätze**
Es sei $(\Omega, \mathscr{A}, P)$ ein Wahrscheinlichkeitsraum, $X, X_n \in \mathscr{L}_1(P)$ für alle $n \geq 1$ und $\mathscr{B} \subset \mathscr{A}$ eine σ-Algebra. Dann gelten die folgenden Sätze:

i) Satz von der monotonen Konvergenz: Ist P-fast sicher $0 \leq X_n \nearrow X$, so gilt auch $\mathbf{E}(X_n \mid \mathscr{B}) \nearrow \mathbf{E}(X \mid \mathscr{B})$.

ii) Satz von der dominierten Konvergenz: Ist $Y \in \mathscr{L}_1(P)$ mit $|X_n| \leq Y$ für alle $n \geq 1$ und $X_n \to X$ P-fast sicher, so folgt P-fast sicher auch $\mathbf{E}(X_n \mid \mathscr{B}) \to \mathbf{E}(X \mid \mathscr{B})$.

Beweis: Der Beweis beider Sätze erfolgt getrennt:

i) Der „normale" Satz von Beppo Levi (vgl. Satz A.2 im Anhang) liefert uns $\mathbf{E}(X - X_n) \to 0$. Aus Satz 1.5.3 folgt dann

$$\mathbf{E}|\mathbf{E}(X \mid \mathscr{B}) - \mathbf{E}(X_n \mid \mathscr{B})| \overset{\text{iv)}}{=} \mathbf{E}(\mathbf{E}(X \mid \mathscr{B}) - \mathbf{E}(X_n \mid \mathscr{B})) \overset{\text{i), iii)}}{=} \mathbf{E}(X - X_n) \to 0.$$

Damit konvergiert $\mathbf{E}(X_n \mid \mathscr{B}) - \mathbf{E}(X \mid \mathscr{B})$ in $\mathscr{L}_1(P)$ gegen 0. Aus der Wahrscheinlichkeitstheorie wissen wir mit [WTSkript11, Satz II.7.4], dass dann $\mathbf{E}(X_n \mid \mathscr{B}) - \mathbf{E}(X \mid \mathscr{B}) \to 0$

stochastisch gilt. Aus der Wahrscheinlichkeitstheorie wissen wir, dass es dann eine Teilfolge mit $\mathbf{E}(X_{n_k} \mid \mathbf{B}) - \mathbf{E}(X \mid \mathscr{B}) \to 0$ P-fast sicher gibt. Da die Folge $\mathbf{E}(X_{n_k} \mid \mathscr{B})$ monoton steigend ist, folgt aber auch die P-fast sichere Konvergenz der Gesamtfolge.

ii) Wir definieren zunächst $Z_n := \sup_{k \geq n} |X_k - X|$, dann gilt P-fast sicher $Z_n \searrow 0$ und $|X_n - X| \leq Z_n$. Ferner erhalten wir

$$|\mathbf{E}(X_n \mid \mathscr{B}) - \mathbf{E}(X \mid \mathscr{B})| = |\mathbf{E}(X_n - X \mid \mathscr{B})| \leq \mathbf{E}(|X_n - X| \mid \mathscr{B}) \leq \mathbf{E}(Z_n \mid \mathscr{B}).$$

Daher werden wir nun zeigen, dass $\mathbf{E}(Z_n \mid \mathscr{B}) \to 0$ gilt. Dazu erinnern wir uns, dass P-fast sicher $Z_n \searrow 0$ gilt. Daraus folgt, dass auch $\mathbf{E}(Z_n \mid \mathscr{B}) \searrow$ gilt und wir setzen $U := \lim \mathbf{E}(Z_n \mid \mathscr{B})$. Klar ist, dass $U \geq 0$ gilt. Um $U = 0$ zu zeigen genügt es daher, $\mathbf{E}U = 0$ zu beweisen. Da $|X_n| \leq Y$ gilt, folgt auch $|X| \leq Y$ und damit $|X_n - X| \leq 2Y$. Nun erhalten wir $0 \leq Z_n \leq 2Y$ und der (normale) Satz von der dominierten Konvergenz[1] liefert uns dann $\mathbf{E}Z_n \to 0$. Damit gilt $0 \leq \mathbf{E}U \leq \mathbf{E}(\mathbf{E}(Z_n \mid \mathscr{B})) = \mathbf{E}Z_n \to 0$.

□

Satz 1.5.5 **Ungleichung von Jensen**
Es sei $(\Omega, \mathscr{A}, P)$ ein Wahrscheinlichkeitsraum, $I \subset \mathbf{R}$ ein Intervall, $X : \Omega \to I$ eine P-integrierbare Abbildung, $\phi : I \to \mathbf{R}$ konvex und $\mathscr{B} \subset \mathscr{A}$ eine σ-Algebra. Dann ist P-fast sicher $\mathbf{E}(X \mid \mathscr{B}) \in I$ und gilt zusätzlich $\phi \circ X \in \mathscr{L}_1(P)$, so folgt

$$\phi(\mathbf{E}(X \mid \mathscr{B})) \leq \mathbf{E}(\phi \circ X \mid \mathscr{B}).$$

Insbesondere gilt $\phi(\mathbf{E}X) \leq \mathbf{E}(\phi \circ X)$, falls $\phi \circ X \in \mathscr{L}_1(P)$ ist.

Die Ungleichung von Jensen ist also gewissermaßen eine Verallgemeinerung des Konvexitätsbegriffes auf nicht-abzählbare Mengen bzw. Wahrscheinlichkeitsmaße.

Beweis: Wir setzen zunächst $U := \{v : I \to \mathbf{R} \mid v \text{ ist affin linear und } v \leq \phi\}$. Dann gilt $\phi(x) = \sup_{v \in U} v(x)$ für alle $x \in I$. Ferner folgt P-fast sicher $\mathbf{E}(X \mid \mathscr{B}) \in I$ aus $x \in I$ und der Monotonie. Außerdem ist nun $\phi(X) = \sup_{v \in U} v(X)$. Für $v_0 \in U$ gilt dann

$$v_0(\mathbf{E}(X \mid \mathscr{B})) = \mathbf{E}(v_0 \circ X \mid \mathscr{B}) \leq \mathbf{E}\left(\sup_{v \in U} v(X) \mid \mathscr{B}\right) = \mathbf{E}(\phi \circ X \mid \mathbf{B})$$

und wir erhalten schließlich

$$\phi(\mathbf{E}(X \mid \mathscr{B})) = \sup_{v \in U} v(\mathbf{E}(X \mid \mathscr{B})) \leq \mathbf{E}(\phi \circ X \mid \mathscr{B}).$$ □

1 Dieser findet sich im Anhang als Satz A.4.

Korollar 1.5.6 **$\mathscr{L}_p$-Zugehörigkeit**
Es sei $(\Omega, \mathscr{A}, P)$ ein Wahrscheinlichkeitsraum, $p \geq 1$ und $X \in \mathscr{L}_p(P)$, sowie $\mathscr{B} \subset \mathscr{A}$ eine σ-Algebra. Dann gilt $\mathbf{E}(X \mid \mathscr{B}) \in \mathscr{L}_p(P)$ und

$$\|\mathbf{E}(X \mid \mathscr{B})\|_{\mathscr{L}_p(P)} \leq \|X\|_{\mathscr{L}_p(P)}.$$

Der Beweis ist eher einfach und wird daher zur Übung überlassen.

Satz 1.5.7 **Einfluss von $\mathscr{B}$ auf die bedingte Erwartung**
Es sei $(\Omega, \mathscr{A}, P)$ ein Wahrscheinlichkeitsraum und $X \in \mathscr{L}_1(P)$, dann gelten die folgenden Eigenschaften:

i) Für $\mathscr{B} = \{\emptyset, \Omega\}$ gilt $\mathbf{E}(X \mid \mathscr{B}) = \mathbf{E}X$.

ii) Sind $\mathscr{C} \subset \mathscr{B} \subset \mathscr{A}$ jeweils σ-Algebren, so gilt $\mathbf{E}(\mathbf{E}(X \mid \mathscr{B}) \mid \mathscr{C}) = \mathbf{E}(X \mid \mathscr{C})$. Insbesondere gilt $\mathbf{E}(\mathbf{E}(X \mid \mathscr{B})) = \mathbf{E}(X \mid \mathscr{B})$.

iii) Sei $\mathscr{B} \subset \mathscr{A}$ eine σ-Algebra und $\mathscr{B}$ und $\sigma(X)$ seien unabhängig, dann gilt $\mathbf{E}(X \mid \mathscr{B}) = \mathbf{E}X$.

Auch dieser Beweis wird zur Übung überlassen.

Satz 1.5.8 **Produkte**
Es sei $(\Omega, \mathscr{A}, P)$ ein Wahrscheinlichkeitsraum, $\mathscr{B} \subset \mathscr{A}$ eine σ-Algebra, $Y \in \mathscr{L}_1(P)$ und $X : \Omega \to \mathbf{R}$ eine $\mathscr{B}$-messbare Abbildung mit $X \circ Y \in \mathscr{L}_1(P)$. Dann gilt

$$\mathbf{E}(XY \mid \mathscr{B}) = X \cdot \mathbf{E}(Y \mid \mathscr{B}).$$

Da X messbar ist, gilt $\mathbf{E}(X \mid \mathscr{B}) = X$ und die Gleichung lässt sich nun auch als $\mathbf{E}(XY \mid \mathscr{B}) = \mathbf{E}(X \mid \mathbf{B})\mathbf{E}(Y \mid \mathscr{B})$ schreiben.

Beweis: Der Beweis gliedert sich in die selben vier Schritte, die wir auch bei der Konstruktion für Integrale in [WTSkript11] durchlaufen haben.

i) Es sei $X = \mathbf{1}_A$ für $A \in \mathscr{B}$. Dann ist $X \cdot \mathbf{E}(Y \mid \mathscr{B})$ $\mathscr{B}$-messbar und P-integrierbar. Für $B \in \mathscr{B}$ folgt dann

$$\mathbf{E}(X \cdot \mathbf{E}(Y \mid \mathscr{B}) \cdot \mathbf{1}_B) = \mathbf{E}(\mathbf{E}(X \mid \mathscr{B}) \cdot \mathbf{1}_{A \cap B}) = \mathbf{E}(Y \cdot \mathbf{1}_{A \cap B}) = \mathbf{E}(XY \cdot \mathbf{1}_B).$$

ii) Nun sei X eine Treppenfunktion, dann folgt die Behauptung aus der Linearität der bedingten Erwartung (vgl. Satz 1.5.3).

iii) Es sei $X \geq 0$ messbar, dann existieren Treppenfunktionen $X_n \nearrow X$. Wir zerlegen $Y = Y^+ - Y^-$ wie gewohnt und erhalten $X_n Y^+ \nearrow XY^+$ und $X_n Y^- \nearrow XY^-$. Mit der monotonen Konvergenz aus Satz 1.5.4 und der Linearität erhalten wir dann die Behauptung.

iv) Nun sei X messbar und wir zerlegen $X = X^+ - X^-$ wie gewohnt. Mit der Linearität und den vorherigen Schritten folgt die Behauptung. □

Satz 1.5.9 **Bestapproximation**
Es sei $(\Omega, \mathscr{A}, P)$ ein Wahrscheinlichkeitsraum, $B \subset \mathscr{A}$ eine σ-Algebra und $X \in \mathscr{L}_2(P)$. Dann nimmt die Abbildung $\mathscr{L}_2(P|_{\mathscr{B}}) \to \mathbf{R}$ vermöge $Y \mapsto \|X - Y\|^2_{\mathscr{L}_2(P|_{\mathscr{B}})} = \mathbf{E}(X - Y)^2$ ihr einziges Minimum in $Y^* = \mathbf{E}(X \mid \mathscr{B})$ an.

Beweis: Es sei $Y \in \mathscr{L}_2(P|_{\mathscr{B}})$, nach Satz 1.5.8 gilt dann $\mathbf{E}(XY \mid \mathscr{B}) = Y\mathbf{E}(X \mid \mathscr{B}) = Y \cdot Y^*$. Daraus folgt $\mathbf{E}(XY) = \mathbf{E}(\mathbf{E}(XY \mid \mathscr{B})) = \mathbf{E}(Y \cdot Y^*)$ und für $Y = Y^* \in \mathscr{L}_2(P|_{\mathscr{B}})$ gilt dann $\mathbf{E}(XY) = \mathbf{E}(Y^*)^2$. Dies ergibt

$$\begin{aligned}\mathbf{E}(X-Y)^2 - \mathbf{E}(X-Y^*)^2 &= -2\mathbf{E}(XY) + \mathbf{E}Y^2 + \underbrace{2\mathbf{E}(XY^*) - \mathbf{E}(Y^*)^2}_{=\mathbf{E}(Y^*)^2} = \mathbf{E}Y^2 - 2\mathbf{E}YY^* + \mathbf{E}(Y^*)^2 \\ &= \mathbf{E}(Y - Y^*)^2 \geq 0.\end{aligned}$$

Daraus können wir nun $\mathbf{E}(X - Y^*)^2 \leq \mathbf{E}(X - Y)^2$ folgern, wobei Gleichheit genau für $\mathbf{E}(Y - Y^*)^2 = 0$ gilt, was wiederum genau für $Y = Y^*$ P-fast sicher der Fall ist. □

Korollar 1.5.10 **$\mathscr{L}_2$-Projektionseigenschaften**
Es sei $(\Omega, \mathscr{A}, P)$ ein Wahrscheinlichkeitsraum und $\mathscr{B} \subset \mathscr{A}$ eine σ-Algebra. Dann ist $\mathbf{E}(\cdot \mid \mathscr{B})\colon \mathscr{L}_2(P) \to \mathscr{L}_2(P)$ eine orthogonale Projektion auf $\mathscr{L}_2(P|_{\mathscr{B}})$, das heißt es gilt

i) $\mathbf{E}(\mathbf{E}(\cdot \mid \mathbf{B}) \mid \mathscr{B}) = \mathbf{E}(\cdot \mid \mathscr{B})$.

ii) $\ker \mathbf{E}(\cdot \mid \mathscr{B}) = \mathscr{L}_2^{\perp}(P|_{\mathscr{B}}) = (\operatorname{im} \mathbf{E}(\cdot \mid \mathscr{B}))^{\perp}$, wobei $\mathscr{L}_2^{\perp}(P|_{\mathscr{B}}) = \{Y \in \mathscr{L}_2(P) : \mathbf{E}(YZ) = 0 \quad \forall Z \in \mathscr{L}_2(P|_{\mathscr{B}})\}$ gilt.

Eigentlich sollte dies mit L_2 statt $\mathscr{L}_2$ formuliert werden, dafür haben wir bisher jedoch zu wenig Theorie. Dabei ist $\mathscr{L}_2(P) = \{f : X \to \mathbf{R}$ messbar mit $\|f\|_2^2 := \int |f|^2 \, \mathrm{d}P < \infty\}$, allerdings ist $\|\cdot\|_2^2$ keine Norm, sondern nur eine Seminorm, da es $f : X \to \mathbf{R}$ messbar mit $f \neq 0$ gibt, so dass $\|f\|_2^2 = 0$ ist. Daher definiert man $f \sim g :\Leftrightarrow f = g$ P-fast sicher und faktorisiert solche Funktionen dann mittels $L_2(P) := \mathscr{L}_2(P)/\sim$ heraus, dieser Raum besteht also eigentlich aus Funktionenklassen. Dieser Unterschied wird in der Praxis jedoch häufig ignoriert.

Beweis: Die Aussage kann aus Satz 1.5.9 gefolgert werden, wir werden dies hier jedoch elementarer beweisen. Die erste Aussage folgt aus Satz 1.5.7, wir widmen uns nun also dem zweiten Teil. Nach Satz 1.5.3 gilt $\operatorname{im}\mathbf{E}(\cdot \mid \mathscr{B}) = \mathscr{L}_2(P\big|_{\mathscr{B}})$. Wir zeigen nun die Inklusion „$\subset$": Sei $Y \in \ker \mathbf{E}(\cdot \mid \mathscr{B})$, das heißt $\mathbf{E}(Y \mid \mathscr{B}) = 0$ und $Z \in \mathscr{L}_2(P\big|_{\mathscr{B}})$, dann gilt

$$\mathbf{E}(YZ) = \mathbf{E}(\mathbf{E}(YZ \mid \mathscr{B})) \overset{1.5.8}{=} \mathbf{E}(Z \cdot \mathbf{E}(Y \mid \mathscr{B})) = 0.$$

Für die andere Inklusion „$\supset$" sei $Y \in \mathscr{L}_2^{\perp}(P\big|_{\mathscr{B}})$, das heißt $Y \in \mathscr{L}_2(P)$ und $\mathbf{E}(YZ) = 0$ für alle $Z \in \mathscr{L}_2(P\big|_{\mathscr{B}})$. Wir wollen zeigen, dass $\mathbf{E}(Y \mid \mathscr{B}) = 0$ gilt. Dazu setzen wir $Z := \mathbf{E}(Y \mid \mathscr{B}) \in \mathscr{L}_2(P\big|_{\mathscr{B}})$ und es folgt

$$0 = \mathbf{E}(YZ) = \mathbf{E}(\mathbf{E}(YZ \mid \mathscr{B})) \overset{1.5.8}{=} \mathbf{E}(Z \cdot \mathbf{E}(Y \mid \mathscr{B})) = \mathbf{E}Z^2,$$

also gilt P-fast sicher $Z = 0$. $\square$

Wir wollen nun einige verschiedene Interpretationen anführen und diskutieren:

i) Die $\mathscr{B}$-Messbarkeit von Y bedeutet, dass nur Informationen, die durch $\mathscr{B}$ ausdrückbar sind, zur Konstruktion von Y verwendet werden können. Präzisiert wird dies durch das Approximationslemma aus der Wahrscheinlichkeitstheorie (vgl. [WTSkript11]).

ii) Die bedingte Erwartung $\mathbf{E}(X \mid \mathscr{B})$ vereinfacht X in dem Sinne, dass alle nicht durch $\mathscr{B}$ ausdrückbaren Informationen weggelassen werden. Präzise bedeutet dies, dass $\ker \mathbf{E}(\cdot \mid \mathscr{B}) = \mathscr{L}_2^{\perp}(P\big|_{\mathscr{B}})$ weggelassen wird, wobei dieser Raum orthogonal zu den erlaubten Funktionen steht.

iii) Die bedingte Erwartung $\mathbf{E}(X \mid \mathscr{B})$ ist die beste Approximation für X innerhalb aller Funktionen, die nur $\mathscr{B}$ als Information zur Verfügung haben. Haben wir also $\mathscr{B}$, so ist $\mathbf{E}(X \mid \mathscr{B})$ also die beste Prognose von X. Präziser haben wir dies in Satz 1.5.9 diskutiert.

Im letzten Punkt haben wir davon gesprochen, was passiert, wenn wir $\mathscr{B}$ „haben", sind jedoch nicht darauf eingegangen, was dies genau bedeutet. Es seien $A, B \in \Omega$ messbar und $\omega \in \Omega$ eine zufällige Beobachtung, dann ist $P(A)$ die Wahrscheinlichkeit für $\omega \in A$ und $P(A \mid B)$ die Wahrscheinlichkeit für $\omega \in A$, falls wir zusätzlich wissen, dass $\omega \in B$ gilt. Analog gibt uns $P(A \mid B^c)$ die Wahrscheinlichkeit für $\omega \in A$, falls wir $\omega \notin B$ wissen. Damit ist

$$\omega \mapsto \begin{cases} P(A \mid B) & \text{falls } \omega \in B \\ P(A \mid B^c) & \text{falls } \omega \notin B \end{cases}$$

die Wahrscheinlichkeit für $\omega \in A$, falls wir über $\omega \in B$ Bescheid wissen. Dies kann auf σ-Algebren $\mathscr{A}$ mit $P(A) > 0$ für alle $\emptyset \neq A \in \mathscr{A}$ verallgemeinert werden (vgl. die Konstruktion zu Beginn des Kapitels). Dies führte zu Messbarkeit und einer Gleichung, beides zusammen machte die Konstruktion dann eindeutig. Für die Verallgemeinerung haben wir die Messbarkeit und die Gleichung zur Definition erhoben. In diesem Sinne ist die bedingte Wahrscheinlichkeit $P(A \mid \mathscr{B})(\omega) = \mathbf{E}(\mathbf{1}_A \mid \mathscr{B})(\omega)$ die Wahrscheinlichkeit von $\omega \in A$, falls wir $\omega \in B$ für alle $B \in \mathscr{B}$ entscheiden können.

An dieser Stelle wollen wir betonen, dass $P(\cdot \mid \mathscr{B})(\omega)$ im Allgemeinen *kein* Wahrscheinlichkeitsmaß ist, denn für jedes $A \in \mathscr{A}$ arbeiten wir mit Versionen und es im Allgemeinen zuviele A gibt um diese zusammenzuführen.

Lemma 1.5.11 **Faktorisierungslemma**
Es sei $(\Omega, \mathscr{A}, P)$ ein Wahrscheinlichkeitsraum, $(\Omega', \mathscr{A}')$ ein Messraum, $X : \Omega \to \Omega'$ messbar und $Y : \Omega \to \mathbf{R}$. Dann sind folgende Aussagen äquivalent:

i) Y ist $\sigma(X)$-messbar.

ii) Es existiert eine messbare Abbildung $g : \Omega' \to \mathbf{R}$, so dass das folgende Diagramm kommutiert:

$$\begin{tikzcd} \Omega \arrow[rr, "Y"] \arrow[dr, "X"'] & & \mathbf{R} \\ & (\Omega', \mathscr{A}') \arrow[ur, "g"'] & \end{tikzcd}$$

Ist X surjektiv, so folgt ferner, dass g eindeutig ist.

Beweis: Die Richtung von ii) nach i) ist trivial. Für die andere Richtung führen wir wieder aufeinanderfolgende Schritte aus. Im ersten Schritt sei $Y = \mathbf{1}_A$, dann ist $A \in \sigma(X)$, da Y nach Voraussetzung $\sigma(X)$-messbar ist. Dann existiert ein $A' \in \mathscr{A}'$ mit $A = X^{-1}(A')$ und wir setzen $g := \mathbf{1}_{A'}$. Für den zweiten Schritt sei $Y = \sum_{i=1}^n \alpha_i \mathbf{1}_{A_i}$ und wir setzen $g := \sum_{i=1}^n \alpha_i \mathbf{1}_{A_i'}$, wobei wie im ersten Schritt $X^{-1}(A_i') = A_i$ setzen.

Im dritten Schritt sei $Y \geq 0$, dann existiert eine Folge $Y_n \nearrow Y$ von Treppenfunktionen und mit Hilfe des zweiten Schritts existiert entsprechend eine Folge g_n mit $Y_n = g_n \circ X$. Da die Konstruktion im zweiten Schritt monoton ist, folgt $g_n \nearrow g := \lim g_n$. Im letzten Schritt zerlegen wir schließlich wie gewohnt $Y = Y^+ - Y^-$ und erhalten so den Rest der Aussage.

Zu zeigen bleibt nun noch die Eindeutigkeit, wenn X surjektiv ist. Sind $\omega_1, \omega_2 \in \Omega$ mit $X(\omega_1) = X(\omega_2)$, so gilt auch $Y(\omega_1) = g(X(\omega_1)) = g(X(\omega_2)) = Y(\omega_2)$. Sei nun $x \in \Omega'$, dann existiert ein $\omega \in \Omega$ mit $X(\omega) = x$. Dann erhalten wir $g(x) = g(X(\omega)) = Y(\omega)$ und g ist eindeutig, wobei $Y(\omega)$ nach unseren Vorüberlegungen von der speziellen Wahl des ω unabhängig ist. $\square$

Wir wollen diese Ergebnisse nun wieder interpretieren. Dazu sei $A, B \in \mathscr{A}$ und $X := \mathbf{1}_B$, sowie $\mathscr{B} := \sigma(X)$. Dann ist $\mathscr{B} = \{\emptyset, B, B^c, \Omega\}$ und nach Lemma 1.5.11 gilt $P(A \mid \mathbf{1}_B) = P(A \mid \mathscr{B}) = \mathbf{E}(\mathbf{1}_A \mid \mathscr{B}) = g \circ \mathbf{1}_B$. Für $\omega \in B$ folgt dann

$$\begin{aligned} P(B)P(A \mid \mathbf{1}_B)(\omega) = P(B)g(1) = \int g(1)\mathbf{1}_B \,\mathrm{d}P = \int g \circ \mathbf{1}_B \,\mathrm{d}P = \int_B \mathbf{E}(\mathbf{1}_A \mid B) \,\mathrm{d}P = \int_B \mathbf{1}_A \,\mathrm{d}P \\ = P(A \cap B). \end{aligned}$$

Ist nun $P(B) > 0$, so folgt $P(A \mid \mathbf{1}_B)(\omega) = P(A \mid B)$ für alle $\omega \in B$, aber wir wissen bereits, dass die linke Seite auch für $P(B) = 0$ definiert ist.

Definition 1.5.12 Faktorisierter bedingter Erwartungswert
Es sei $(\Omega, \mathscr{A}, P)$ ein Wahrscheinlichkeitsraum, $(\Omega', \mathscr{A}')$ ein Messraum, $Y \in \mathscr{L}_1(P)$ und $X : \Omega \to \Omega'$ messbar und surjektiv. Dann ist $\mathbf{E}(Y \mid X)$ $\sigma(X)$-messbar und mit Lemma 1.5.11 existiert genau ein $g : \Omega' \to \mathbf{R}$ mit $\mathbf{E}(Y \mid X) = g \circ X$. Für $x \in \Omega'$ schreiben wir dann $\mathbf{E}(Y \mid X = x) := g(x)$ und nennen $\mathbf{E}(Y \mid X = \cdot)$ den faktorisierten bedingten Erwartungswert. Analog ist $P(A \mid X = x) = \mathbf{E}(\mathbf{1}_A \mid X = x)$.

Für $A \in \mathscr{A}$ ist $P(A \mid X = \cdot)$ nur bis auf P_X-Nullmengen eindeutig, da $\mathbf{E}(Y \mid X)$ nur bis auf P-Nullmengen bestimmt ist. Für eine weitere Ausführung verweisen wir auf [Klenke06, Kap. 8.3, S. 182].

2

Zeitdiskrete Martingale

Martingale waren ursprünglich eine Klasse von Wettstrategien im 18. Jahrhundert Frankreichs, aber auch ein Begriff für bestimmte Pferdezügel. In der Theorie der stochastischen Prozesse sind Martingale bestimmte Klassen von Prozessen, welche die stochastischen Prozesse gewissermaßen einschnüren.

2.1. Definition und einfache Beispiele

Im Folgenden sei $(\Omega, \mathscr{A}, P)$ immer ein Wahrscheinlichkeitsraum, sowie $T \subset [0,\infty)$ eine Indexmenge. Da wir vor allem zeitdiskrete Martingale betrachten werden, wird häufig $T = \mathbf{N}$ oder $T = \mathbf{N}_0$ sein.

Definition 2.1.1 **Filtration**
Eine Familie $\mathscr{F} = (\mathscr{F}_t)_{t\in T}$ von σ-Unteralgebren von $\mathscr{A}$ heißt Filtration genau dann, wenn $\mathscr{F}_s \subset \mathscr{F}_t$ für alle $s, t \in T$ mit $s \leq t$ gilt.

Wird $\mathscr{F}_t$ als unser Kenntnisstand zum Zeitpunkt t betrachtet, so sichert die Filtrationseigenschaft, dass wir höchstens dazulernen, unser Kenntnisstand mit der Zeit also größer wird.

Definition 2.1.2 **Adaptierter stochastischer Prozess**
Sei $X = (X_t)_{t\in T}$ ein $\mathscr{X}$-wertiger stochastischer Prozess und $\mathscr{F} = (\mathscr{F}_t)_{t\in T}$ eine Filtration, dann heißt X an $\mathscr{F}$ adaptiert genau dann, wenn X_t für alle $t \in T$ $\mathscr{F}_t$-messbar ist.

Dies bedeutet also, dass X_t mit unserem Kenntnisstand $\mathscr{F}_t$ beschreibbar ist.

Definition 2.1.3 Vorhersagbarer stochastischer Prozess
Sei $X = (X_n)_{n\geq 0}$ ein $\mathscr{X}$-wertiger stochastischer Prozess und $\mathscr{F} = (\mathscr{F}_n)_{n\geq 0}$ eine Filtration, dann heißt X vorhersagbar oder previsibel bezüglich $\mathscr{F}$ genau dann, wenn X_n für alle $n \geq 1$ $\mathscr{F}_{n-1}$-messbar und X_0 konstant ist.

In diesem Fall können wir den Prozess also bereits aus dem Kenntnisstand vom Zeitpunkt zuvor beschreiben und daher gewissermaßen vorhersagen. Ist X previsibel, so ist X auch adaptiert. Eine häufig verwendete Filtration ist die natürliche Filtration:
Ist $(X_t)_{t\in T}$ ein stochastischer Prozess und $\mathscr{F}_t := \sigma(X_s : T \ni s \leq t)$, dann ist X an $\mathscr{F} := (\mathscr{F}_t)_{t\in T}$ adaptiert. Dies ist die kleinste Filtration bezüglich welcher X adaptiert ist. Oft werden wir hierfür $\mathscr{F}_t^X := \mathscr{F}_t$ schreiben. Wir kommen nun zum eigentlichen Objekt unserer Begierde:

Definition 2.1.4 Martingal
Sei $M = (M_n)_{n\geq 0}$ ein $\mathbf{R}$-wertiger stochastischer Prozess mit $M_n \in \mathscr{L}_1(P)$ für alle $n \geq 0$, sowie $\mathscr{F} = (\mathscr{F}_n)_{n\geq 0}$ eine Filtration, an welcher M adaptiert ist. Dann heißt M

i) (zeitdiskretes) Martingal bezüglich $\mathscr{F}$ genau dann, wenn $\mathbf{E}(M_n \mid \mathscr{F}_{n-1}) = M_{n-1}$ P-fast sicher für alle $n \geq 1$ gilt.

ii) Sub-Martingal genau dann, wenn $\mathbf{E}(M_n \mid \mathscr{F}_{n-1}) \geq M_{n-1}$ P-fast sicher für alle $n \geq 1$ gilt.

iii) Super-Martingal genau dann, wenn $\mathbf{E}(M_n \mid \mathscr{F}_{n-1}) \leq M_{n-1}$ P-fast sicher für alle $n \geq 1$ gilt.

Wir wollen an dieser Stelle einige einfache Eigenschaften festhalten:

i) Ist M ein Martingal, so ist es natürlich auch ein Sub- und ein Super-Martingal.

ii) M ist ein Sub-Martingal genau dann, wenn $-M$ ein Super-Martingal ist.

iii) Ist M ein Martingal, so ist auch $(M_n - M_0)_{n\geq 0}$ ein Martingal.

iv) Ist M ein Martingal, so gilt $\mathbf{E}M_n = \mathbf{E}(\mathbf{E}(M_n \mid \mathscr{F}_{n-1})) = \mathbf{E}M_{n-1}$ für alle $n \geq 1$, die Erwartungswerte von Martingalen ändern sich also nicht über die Zeit. Für Sub-Martingale (Super-Martingale) gilt mit analoger Begründung, dass die Erwartungswerte monoton steigen (fallen).

Die beste Prognose von M_n zum Zeitpunkt $n-1$ ist, wie wir im letzten Kapitel gesehen haben, $\mathbf{E}(M_n \mid \mathscr{F}_{n-1})$. Die Martingaleigenschaft besagt nun gerade, dass die Prognose gleich M_{n-1} ist.

Alternativ wollen wir Martingale nun als ein Spiel interpretieren: Es sei $M = (M_n)_{n\geq 0}$ ein $\mathbf{R}$-wertiger stochastischer Prozess, der an $(\mathscr{F}_n)_{n\geq 0}$ adaptiert ist. Dabei soll M_n unser Vermögen zum Zeitpunkt n beschreiben, also ist $M_n - M_{n-1}$ unser Gewinn in der n-ten Spielrunde. Ist M ein Martingal, so gilt mit der Martingal- und Adaptionseigenschaft

$$\mathbf{E}(M_n - M_{n-1} \mid \mathscr{F}_{n-1}) = \mathbf{E}(M_n \mid \mathscr{F}_{n-1}) - \mathbf{E}(M_{n-1} \mid \mathscr{F}_{n-1}) = M_{n-1} - M_{n-1} = 0.$$

In der n-ten Spielrunde ist das Spiel also fair, sofern unser Kenntnisstand zum Zeitpunkt $n-1$ zugrundegelegt wird. Ist M ein Super-Martingal, so folgt analog, dass $\mathbf{E}(M_{n-1} - M_n \mid \mathscr{F}_{n-1}) \leq 0$ gilt, also ist das Spiel auf unsere Kosten unfair.

Beispiel 2.1.5 *Summen unabhängiger Zufallsvariablen*
Ist $(X_n)_{n\geq 0}$ eine Folge unabhängiger Zufallsvariablen mit $X_n \in \mathscr{L}_1(P)$ für alle $n \geq 0$ und $X_0 = 0$. Dann ist

$$M_n := \sum_{i=0}^{n} (X_i - \mathbf{E}X_i) \qquad \text{für } n \geq 0$$

ein Martingal bezüglich der natürlichen Filtration von M_n. Die Adaptionseigenschaft und Integrierbarkeit sind offensichtlich, zu zeigen bleibt die Martingaleigenschaft. Es ist

$$\mathbf{E}(M_{n+1} - M_n \mid \mathscr{F}_n) = \mathbf{E}(X_{n+1} - \mathbf{E}X_{n+1} \mid \mathscr{F}_n) = \mathbf{E}(X_{n+1} \mid \mathscr{F}_n) - \mathbf{E}X_{n+1} = \mathbf{E}X_{n+1} - \mathbf{E}X_{n+1} = 0,$$

wobei wir ausgenutzt haben, dass X_{n+1} und $\mathscr{F}_n$ unabhängig sind. //

Beispiel 2.1.6 *Martingal durch Filtration*
Es sei $X \in \mathscr{L}_1(P)$ und $(\mathscr{F}_n)_{n\geq 0}$ eine Filtration. Dann definiert

$$M_n := \mathbf{E}(X \mid \mathscr{F}_n) \qquad \text{für } n \geq 0$$

ein Martingal. Auch hier sind die Adaptionseigenschaft und Integrierbarkeit klar, zu beweisen bleibt also wieder die Martingaleigenschaft. Wegen $\mathscr{F}_{n-1} \subset \mathscr{F}_n$ gilt

$$\mathbf{E}(M_n \mid \mathscr{F}_{n-1}) = \mathbf{E}(\mathbf{E}(X \mid \mathscr{F}_n) \mid \mathscr{F}_{n-1}) = \mathbf{E}(X \mid \mathscr{F}_{n-1}) = M_{n-1}. \quad //$$

Beispiel 2.1.7 *Produkte unabhängiger Zufallsvariablen*
Es sei $(X_n)_{n\geq 0}$ eine Folge unabhängiger Zufallsvariablen mit $\mathscr{L}_1(P) \ni X_n \geq 0$ und $\mathbf{E}X_n = 1$ für alle $n \geq 0$. Wir setzen $M_0 := 0$, $\mathscr{F}_0 := \{\emptyset, \Omega\}$ und dann $M_n := X_1 \cdot \ldots \cdot X_n$ und $\mathscr{F}_n := \sigma(X_1, \ldots, X_n)$. Dann ist M_n ein Martingal. Die Adaptionseigenschaft ist nach Konstruktion klar und die Integrierbarkeit folgt aus $X_1 \cdot \ldots \cdot X_n \in \mathscr{L}_1(P)$. Wir zeigen nun noch die Martingaleigenschaft. Es gilt

$$\mathbf{E}(M_n \mid \mathscr{F}_{n-1}) = \mathbf{E}(M_{n-1} \cdot X_n \mid \mathscr{F}_{n-1}) \overset{1.5.8}{=} M_{n-1} \cdot \mathbf{E}(X_n \mid \mathscr{F}_{n-1}) = M_{n-1} \cdot \mathbf{E}X_n = M_{n-1} \cdot 1. \quad //$$

Nach diesen etwas einfacheren Beispielen wollen wir nun noch ein komplizierteres Beispiel diskutieren:

Beispiel 2.1.8 *Diskretes stochastisches Integral*
Es sei $(X_n)_{n\geq 0}$ ein $\mathscr{F} = (\mathscr{F}_n)_{n\geq 0}$-adaptierter stochastischer Prozess, sowie $(H_n)_{n\geq 1}$ ein $\mathscr{F}$-vorhersagbarer stochastischer Prozess. Dann heißt der Prozess $H.X$, der durch

$$(H.X)_n := X_0 + \sum_{m=1}^{n} H_m(X_m - X_{m-1}) \qquad \text{für } n \geq 0$$

definiert ist, diskretes stochastisches Integral. Ist X ein Martingal, so heißt $H.X$ auch Martingaltransformierte von X.

Wir wollen nun erläutern, woher diese Bezeichnung kommt. Dazu benötigen wir das Stieltjes-Integral, das durch

$$\int_0^t h \,\mathrm{d}g(x) := \lim \sum_{k=1}^{n} h(t_{k-1})(g(t_k) - g(t_{k-1}))$$

definiert ist, wobei der Grenzwert des Rangs der Zerlegung $0 = t_0 < t_1 < \ldots < t_n = t$ gegen Null betrachtet wird, das heißt die Zerlegung immer feiner gemacht wird. Dieser Grenzwert existiert, wenn g hinreichend regulär ist. Bei $H.X$ wird die Grenzwertbildung weggelassen und die Zerlegung von $[0,n]$ ist sehr speziell gewählt, nämlich $t_m = m$. Dies motiviert die Bezeichnung „diskretes Integral“ für $H.X$. Ferner wählen wir $g(t_k) = X_k(\omega)$, das heißt wir integrieren bezüglich einer zufälligen Funktion, was die Bezeichnung „stochastisches (Integral)“ motiviert. Man kann später den Grenzwert sogar durchführen, dazu benötigt man jedoch die Regularität der Trajektorien von X. //

Lemma 2.1.9
Es sei $(X_n)_{n\geq 0}$ an eine Filtration $\mathscr{F}$ adaptiert mit $X_i \in \mathscr{L}_1(P)$ für alle $i \geq 0$. Ferner sei $(H_n)_{n\geq 1}$ vorhersagbar bezüglich $\mathscr{F}$ mit $H_n(X_n - X_{n-1}) \in \mathscr{L}_1$ für alle $n \geq 1$. Dann gilt:

i) Ist X ein Martingal, so ist auch $H.X$ ein Martingal.

ii) Ist X ein Sub-/Super-Martingal und $H \geq 0$, so ist auch $H.X$ ein Sub-/Super-Martingal.

Beweis: Wir beweisen beide Eigenschaften gleichzeitig. Es gilt

$$\begin{aligned}\mathbf{E}((H.X)_n - (H.X)_{n-1} \mid \mathscr{F}_{n-1}) = \mathbf{E}(H_n(X_n - X_{n-1}) \mid \mathscr{F}_{n-1}) &\overset{1.5.8}{=} H_n \cdot \mathbf{E}(X_n - X_{n-1} \mid \mathscr{F}_{n-1}) \\ &= \begin{cases} 0 & \text{falls } X \text{ Martingal} \\ \geq 0 & \text{falls } X \text{ Sub-Martingal} \\ \leq 0 & \text{falls } X \text{ Super-Martingal} \end{cases}.\end{aligned}$$

Damit ist alles gezeigt. □

2.2. Grundlegende Eigenschaften

In diesem Abschnitt wollen wir zwei Ziele erreichen. Zunächst wollen wir einige Möglichkeiten kennenlernen, aus Martingalen neue Martingale zu konstruieren. Danach wollen wir zeigen, dass stochastische Prozesse im Wesentlichen aus Martingalen und vorhersagbaren stochastischen Prozessen bestehen.

Satz 2.2.1
Es seien $X = (X_n)_{n\geq 0}$ und $Y = (Y_n)_{n\geq 0}$ reellwertige, $\mathscr{F}$-adaptierte stochastische Prozesse, wobei $\mathscr{F}$ eine Filtration ist. Dann gilt:

i) Sind X und Y Martingale und $\alpha, \beta \in \mathbf{R}$, so ist auch $\alpha X + \beta Y$ ein Martingal.

ii) Sind X und Y beide Super- oder Sub-Martingale und $\alpha, \beta \geq 0$, so ist auch $\alpha X + \beta Y$ ein Super- bzw. Sub-Martingal.

iii) Sind X und Y Super-Martingale, so ist $X \wedge Y := \min\{X, Y\}$ ebenfalls ein Super-Martingal. Sind X und Y Sub-Martingale, so ist analog $\max\{X, Y\}$ ein Sub-Martingal.

iv) Ist X ein Super-Martingal, $(T_n)_{n\geq 1} \subset \mathbf{N}$ mit $T_n \to \infty$ und $\mathbf{E}X_{T_n} \geq \mathbf{E}X_0$ für alle $n \geq 1$, so ist X ein Martingal.

Beweis: Die ersten beiden Eigenschaften lassen sich durch simples Nachrechnen unter Verwendung der Linearität des bedingten Erwartungswertes beweisen. Für iii) setzen wir $Z_n := X_n \wedge Y_n$. Wegen $|Z_n| \leq |X_n| + |Y_n|$ folgt dann $Z_n \in \mathscr{L}_1$ und die Tatsache, dass Z_n adaptiert ist. Dann erhalten wir

$$\mathbf{E}(Z_n \mid \mathscr{F}_{n-1}) \leq \mathbf{E}(X_n \mid \mathscr{F}_{n-1}) \leq \mathbf{E}(X_{n-1} \mid \mathscr{F}_{n-1}) \leq \mathbf{E}(Y_{n-1} \mid \mathscr{F}_{n-1}) \leq \mathbf{E}(Z_{n-1} \mid \mathscr{F}_{n-1}).$$

Nun wollen wir iv) beweisen. Dazu sei $m \in \mathbf{N}$, dann existiert ein $n \in \mathbf{N}$ mit $m < T_n$ und wir setzen $Y_i := \mathbf{E}(X_{T_n} \mid \mathscr{F}_i)$ für $i < T_n$. Zuerst zeigen wir, dass P-fast sicher $X_i = Y_i$ für alle $i < T_n$ gilt. Dazu betrachte

$$\begin{aligned} Y_i = \mathbf{E}(X_{T_n} \mid \mathscr{F}_i) &= \mathbf{E}(\mathbf{E}(X_{T_n} \mid \mathscr{F}_{T_n - 1}) \mid \mathscr{F}_i) \\ &\leq \mathbf{E}(X_{T_n - 1} \mid \mathscr{F}_i) \\ &\;\;\vdots \\ &\leq \mathbf{E}(X_i \mid \mathscr{F}_i) = X_i. \end{aligned}$$

Mit dieser Rechnung gilt nun

$$\begin{aligned}\mathbf{E}X_0 \le \mathbf{E}X_{T_n} &= \mathbf{E}(\mathbf{E}(X_{T_n} \mid \mathscr{F}_i)) = \mathbf{E}Y_i \\ &\le \mathbf{E}X_i = \mathbf{E}(\mathbf{E}(X_i \mid \mathscr{F}_{i-1})) \\ &\le \mathbf{E}X_{i-1} \\ &\quad\vdots \\ &\le \mathbf{E}X_0.\end{aligned}$$

Damit müssen all diese Ungleichungen also tatsächlich Gleichungen sein und wir erhalten insbesondere $\mathbf{E}X_i = \mathbf{E}Y_i$ für alle $i < T_n$. Fassen wir dies zusammen, so erhalten wir $X_i = Y_i$ P-fast sicher für alle $i < T_n$. Nun gilt

$$\mathbf{E}(X_m \mid \mathscr{F}_{m-1}) = \mathbf{E}(\underbrace{\mathbf{E}(X_{T_n} \mid \mathscr{F}_m)}_{=Y_m} \mid \mathscr{F}_{m-1}) = \mathbf{E}(X_{T_n} \mid \mathscr{F}_{m-1}) = Y_{m-1} = X_{m-1}.$$ □

Satz 2.2.2
Es sei $X = (X_n)_{n\ge 0}$ ein Martingal und $\phi\colon \mathbf{R} \to \mathbf{R}$ eine konvexe Abbildung. Dann gilt:

i) Falls für den positiven Anteil der Komposition $\mathbf{E}(\phi \circ X)^+ < \infty$ für alle $n \ge 0$ gilt, so folgt, dass $\phi \circ X$ ein Sub-Martingal ist.

ii) Ist $X_n \in \mathscr{L}_p$ für ein $p \in [1,\infty)$ und alle $n \ge 0$, so ist $|X|^p$ ein Sub-Martingal.

Beweis: Die Aussage ii) folgt aus i) für $\phi(t) := |t|^p$. Beweisen wir also i). Dazu sei $v(x) = ax + b$ affin linear für $x \in \mathbf{R}$ mit $v \le \phi$ (vgl. Beweis von Satz 1.5.5). Dann gilt $\phi^- = \max\{0, -\phi\} \le \max\{0, -v\} = v^-$ und wir erhalten

$$\mathbf{E}(\phi \circ X_n)^- \le \mathbf{E}(v \circ X)^- \le |a|\mathbf{E}|X_n| + |b| < \infty.$$

Dann ist $\phi \circ X_n \in \mathscr{L}_1$. Nun folgt mit Satz 1.5.5

$$\mathbf{E}(\phi \circ X_n \mid \mathscr{F}_{n-1}) \ge \phi \mathbf{E}(X_n \mid \mathscr{F}_{n-1}) = \phi \circ X_{n-1}.$$ □

Lemma 2.2.3
Es sei $M = (M_n)_{n\ge 0}$ ein $\mathscr{F}$-adaptiertes, vorhersagbares Martingal. Dann gilt P-fast sicher $M_n = M_0$ für alle $n \ge 0$.

Beweis: Es genügt, induktiv zu zeigen, dass P-fast sicher $M_n = M_{n-1}$ für alle $n \ge 1$ gilt. Dies gilt aber wegen $M_n = \mathbf{E}(M_n \mid \mathscr{F}_{n-1}) = M_{n-1}$. □

Lemma 2.2.4
Es sei $M = (M_n)_{n\geq 0}$ ein Martingal, dann folgt $\mathbf{E}(M_n \cdot M_m) = \mathbf{E}M_m^2$ für alle $0 \leq m \leq n$.

Beweis: Wir können ohne Einschränkung $n > m$ annehmen, da der Fall $n = m$ trivial ist. Dann gilt mit iterierter Anwendung der Martingaleigenschaft

$$\mathbf{E}(M_n \cdot M_m) = \mathbf{E}(\mathbf{E}(M_n \cdot M_m \mid \mathscr{F}_m)) = \mathbf{E}(M_m(\underbrace{\mathbf{E}(M_n \mid \mathscr{F}_m)}_{=M_m})) = \mathbf{E}(M_m \cdot M_m). \qquad \square$$

Satz 2.2.5 **Doob-Zerlegung**
Es sei $X = (X_n)_{n\geq 0}$ ein an $\mathscr{F} = (\mathscr{F}_n)_{n\geq 0}$ adaptierter stochastischer Prozess mit $X_n \in \mathscr{L}_1(P)$ für alle $n \geq 0$. Dann existiert – bis auf Ununterscheidbarkeit – genau ein Martingal M und ein $\mathscr{F}$-vorhersagbarer stochastischer Prozess A mit $A_0 = 0$, so dass $X = M + A$ gilt. Ferner ist X ein Sub-Martingal genau dann, wenn A monoton wachsend ist.

Mit anderen Worten bestehen stochastische Prozesse also aus Martingalen und vorhersagbaren Prozessen.

Beweis: Wir zeigen zunächst die Existenz. Für $n \geq 0$ sei dazu

$$M_n := X_0 + \sum_{k=1}^{n} (X_k - \mathbf{E}(X_k \mid \mathscr{F}_{k-1})),$$

sowie

$$A_n := -\sum_{k=1}^{n} (X_{k-1} - \mathbf{E}(X_k \mid \mathscr{F}_{k-1})).$$

Dann gilt, dass A_n nach Konstruktion $\mathscr{F}_{n-1}$-messbar ist, da es als Summe über $\mathscr{F}_{k-1}$-messbare Terme entsteht und $\mathscr{F}_{k-1} \subset \mathscr{F}_{n-1}$ für $k \in \{1, \ldots, n\}$ gilt. Dann folgt, dass A $\mathscr{F}$-vorhersagbar ist; dass $A_0 = 0$ gilt ist klar. Ferner ist M_n nach Konstruktion mit analoger Argumentation $\mathscr{F}_n$-messbar und integrierbar, wir müssen also noch die Martingaleigenschaft zeigen. Dazu betrachten wir

$$\mathbf{E}(M_n - M_{n-1} \mid \mathscr{F}_{n-1}) = \mathbf{E}(X_n - \mathbf{E}(X_n \mid \mathscr{F}_{n-1}) \mid \mathscr{F}_{n-1}) = \mathbf{E}(X_n \mid \mathscr{F}_{n-1}) - \mathbf{E}(X_n \mid \mathscr{F}_{n-1}) = 0.$$

Nun bleibt noch zu zeigen, dass M und A tatsächlich X zerlegen. Dies gilt wegen

$$\begin{aligned} M_n + A_n &= X_0 + \sum_{k=1}^{n} (X_k - \mathbf{E}(X_k \mid \mathscr{F}_{k-1})) - \sum_{k=1}^{n} X_{k-1} + \sum_{k=1}^{n} \mathbf{E}(X_k \mid \mathscr{F}_{k-1}) \\ &= X_n. \end{aligned}$$

Damit ist die Existenz gezeigt und wir zeigen die Eindeutigkeit. Dazu sei $X = M + A = M' + A'$, dann gilt $M - M' = A' - A$. Nach Satz 2.2.1 ist $M - M'$ ein Martingal und $A' - A$ ist vorhersagbar, also auch $M - M'$. Mit Lemma 2.2.3 folgt dann $M_n - M_n' = M_0 - M_0' = A_0 - A_0' = 0 - 0 = 0$, also folgt P-fast sicher $M = M'$.

Zum Schluss beweisen wir noch, dass X ein Sub-Martingal ist genau dann, wenn A wachsend ist. Dazu überlegen wir uns zunächst

$$\mathbf{E}(X_n \mid \mathscr{F}_{n-1}) = \mathbf{E}(M_n \mid \mathscr{F}_{n-1}) + \mathbf{E}(A_n \mid \mathscr{F}_{n-1}) = M_{n-1} + A_n.$$

Also ist X ein Sub-Martingal. Daraus folgt $M_{n-1} + A_n = \mathbf{E}(X_n \mid \mathscr{F}_{n-1}) \geq X_{n-1} = M_{n-1} + A_{n-1}$ und wir erhalten $A_n \geq A_{n-1}$. Die andere Richtung folgt analog. □

Korollar 2.2.6
Es sei $X = (X_n)_{n\geq 0}$ ein quadrat-integrierbares $\mathscr{F}$-Martingal, d. h. es gilt $X_n \in \mathscr{L}_2$ für alle $n \geq 0$. Dann gilt:

i) X^2 ist ein Sub-Martingal.

ii) Es gibt genau einen $\mathscr{F}$-vorhersagbaren stochastischen Prozess A mit $A_0 = 0$, so dass $X^2 - A$ ein Martingal ist. Dieser Prozess A heißt quadratischer Variationsprozess von X und wird mit $\langle X \rangle := A$ bezeichnet.

iii) Der Prozess $\langle X \rangle$ ist monoton wachsend.

iv) Es gilt $\langle X \rangle_n = \sum_{i=1}^{n} \mathbf{E}((X_i - X_{i-1})^2 \mid \mathscr{F}_{i-1})$ und $\mathbf{E}\langle X \rangle_n = \operatorname{Var}(X_n - X_0)$.

Beweis: Aussage i) folgt unmittelbar aus Satz 2.2.2. Aussage ii) folgt aus Satz 2.2.5, denn für X^2 mit $X^2 = M + A$ erhalten wir $X^2 - A = M$. Aussage iii) folgt ebenfalls aus Satz 2.2.5. Wir zeigen also noch iv). Dazu betrachten wir

$$\sum_{i=1}^{n} \mathbf{E}((X_i - X_{i-1})^2 \mid \mathscr{F}_{i-1}) = \sum_{i=1}^{n} \mathbf{E}(X_i^2 - 2X_iX_{i-1} + X_{i-1}^2 \mid \mathscr{F}_{i-1}),$$

durch Auseinanderziehen und Anwenden der Definitionen der bedingten Erwartung und Martingalen erhalten wir dann

$$\begin{aligned} &= \sum_{i=1}^{n} \left(\mathbf{E}(X_i^2 \mid \mathscr{F}_{i-1}) - X_{i-1}^2\right) \overset{2.2.5}{=} A_n \\ &= \langle X \rangle_n. \end{aligned}$$

Nun wollen wir noch die Formel für den Erwartungswert beweisen. Es gilt

$$\mathbf{E}A_n = \sum_{i=1}^{n} \mathbf{E}(\mathbf{E}(X_i - X_{i-1})^2 \mid \mathscr{F}_{i-1}) = \sum_{i=1}^{n} \mathbf{E}\left(X_i^2 - 2X_iX_{i-1} + X_{i-1}^2\right),$$

und wieder durch Auseinanderziehen und Lemma 2.2.4 erhalten wir

$$\begin{aligned} &= \sum_{i=1}^{n} \left(\mathbf{E}X_i^2 - \mathbf{E}X_{i-1}^2\right) = \mathbf{E}X_n^2 - \mathbf{E}X_0^2 = \mathbf{E}X_n^2 - 2\mathbf{E}X_nX_0 + \mathbf{E}X_0^2 = \mathbf{E}(X_n - X_0)^2 \\ &= \operatorname{Var}(X_n - X_0), \end{aligned}$$

da $\mathbf{E}X_n = \mathbf{E}X_0$ gilt, wie wir bereits gesehen haben. □

Beispiel 2.2.7
Es seien $(Y_i)_{i\geq 1}$ unabhängig mit $Y_i \in \mathscr{L}_2(P)$ und $\mathbf{E}Y_i = 0$ für alle $i \geq 0$, sowie $X_n := \sum_{i=1}^n Y_i$ für $n \geq 0$. Dann gilt

$$\langle X \rangle_n = \sum_{i=1}^n \mathbf{E}Y_i^2 = \sum_{i=1}^n \operatorname{Var} Y_i. \quad /\!\!/$$

Beispiel 2.2.8
Es seien $(Y_i)_{i\geq 0}$ unabhängig mit $Y_i \in \mathscr{L}_2(P)$ und $\mathbf{E}Y_i = 1$ für alle $i \geq 0$, sowie $X_n := \prod_{i=1}^n Y_i$ für $n \geq 0$. Dann gilt

$$\langle X \rangle_n = \sum_{i=1}^n X_{i-1}^2 \operatorname{Var} Y_i. \quad /\!\!/$$

2.3. Stoppzeiten

Wir würden stochastische Prozesse gerne in Abhängigkeit ihres Verhaltens bis zum Eintritt eines bestimmten Ereignisses untersuchen beziehungsweise modifizieren. Dazu sei $(\Omega, \mathscr{A}, P)$ wie üblich ein Wahrscheinlichkeitsraum.

Definition 2.3.1 **Stoppzeit**
Es sei $\mathscr{F} = (\mathscr{F}_n)_{n \geq 0}$ eine Filtration. Eine Abbildung $\tau : \Omega \to \overline{\mathbf{N}_0} := \mathbf{N}_0 \cup \{\infty\}$ heißt Stoppzeit genau dann, wenn

$$\{\tau \leq n\} \in \mathscr{F}_n \qquad \text{für alle } n \geq 0.$$

Falls wir einen stochastischen Prozess bei der Zeit $\tau(\omega)$ in Abhängigkeit des Verhaltens bis zur Zeit $\tau(\omega)$ anhalten wollen – wir werden darauf gleich noch näher eingehen –, so darf $\tau(\omega)$ nur von unserem Kenntnisstand bis zur Zeit $\tau(\omega)$ abhängen. Die Menge $\{\tau \leq n\}$ enthält alle Zeiten vor dem Zeitpunkt n, die nur von der Kenntnis bis zu diesem Zeitpunkt abhängen.

Für zeitdiskrete stochastische Prozesse – auf die wir uns hier ja im Wesentlichen beschränken –, kann man statt der in Definition 2.3.1 verwendeten Menge auch $\{\tau = n\}$ verwenden.

Beispiel 2.3.2 *Eintrittszeit*
Es sei $(X_n)_{n \geq 0}$ ein $\mathscr{F}$-adaptierter Prozess und $A \subset \mathbf{R}$ messbar. Dann ist die Eintrittszeit $\tau_A : \Omega \to \overline{\mathbf{N}_0}$ mit $\omega \mapsto \inf\{n \in \mathbf{N}_0 : X_n(\omega) \in A\}$ eine Stoppzeit, d. h. $\tau_A(\omega)$ ist der erste Zeitpunkt, zu welchem die Trajektorie $X(\omega)$ in A liegt. Betrachte hierzu

$$\{\tau_A \leq n\} = \{\omega : \exists_{k \in \{0,\dots,n\}} : X_k(\omega) \in A\} = \bigcup_{k=0}^{n} \{X_k \in A\} \in \mathscr{F}_n. \quad /\!/$$

Beispiel 2.3.3 *Konstante Stoppzeit*
Es sei $\mathscr{F}$ eine Filtration und $n_0 \in \mathbf{N}_0$. Dann ist $\tau : \Omega \to \overline{\mathbf{N}_0}$ mit $\omega \mapsto n_0$ eine (konstante) Stoppzeit. Betrachte

$$\{\tau \leq n\} = \begin{cases} \emptyset & \text{falls } n < n_0 \\ \Omega & \text{falls } n \geq n_0 \end{cases} \in \mathscr{F}_n. \quad /\!/$$

Lemma 2.3.4
Seien σ und τ Stoppzeiten, dann gilt

i) $\sigma \wedge \tau := \min\{\sigma, \tau\}$ und $\sigma \vee \tau := \max\{\sigma, \tau\}$ sind Stoppzeiten.

ii) $\sigma + \tau$ ist eine Stoppzeit.

iii) $\tau + s$ ist für alle $s \geq 0$ eine Stoppzeit.

Insbesondere gilt zu beachten, dass die Bedingung $s \geq 0$ in Aussage iii) des Lemmas essentiell ist und die Aussage für $s < 0$ im Allgemeinen nicht gilt.

Beweis: Für i) betrachten wir $\{\tau \vee \sigma \leq n\} = \{\tau \leq n\} \cap \{\sigma \leq n\}$ und $\{\tau \wedge \sigma \leq n\} = \{\tau \leq n\} \cup \{\sigma \leq n\}$, diese Mengen sind offenbar messbar.

Für ii) sei $n \in \mathbf{N}_0$, dann folgt aus i), dass $\tau \vee n$ und $\tau \wedge n$ Stoppzeiten sind. Für $m < n$ gilt nun $\{\tau \wedge n \leq m\} \in \mathscr{F}_m \subset \mathscr{F}_n$, für $m \geq n$ gilt $\{\tau \wedge n \leq m\} = \Omega \in \mathscr{F}_n$. Damit ist $\tau' := (\tau \wedge n) + \mathbf{1}_{\{\tau > n\}}$ eine $\mathscr{F}_n$-messbare Zufallsvariable. Analog erhalten wir die $\mathscr{F}_n$-Messbarkeit von $\sigma' := (\sigma \wedge n) + \mathbf{1}_{\{\sigma > n\}}$. Damit folgt $\tau' + \sigma' \in \mathscr{F}_n$. Schließlich gilt für $\tau \leq n$ und $\sigma \leq n$, dass $\tau' + \sigma' = \tau + \sigma$ gilt, ist aber zum Beispiel $\tau > n$, so ist $\tau' = n + 1 > n$. Damit folgt $\{\sigma + \tau \leq n\} = \{\sigma' + \tau' \leq n\} \in \mathscr{F}_n$.

Aussage iii) folgt aus Aussage ii) und Beispiel 2.3.3. Wir wollen noch zeigen, warum die Aussage für $s < 0$ nicht stimmt. Es ist $\{\tau + s \leq n\} = \{\tau \leq n - s\} \in \mathscr{F}_{n-s}$, aber im Allgemeinen gilt $\mathscr{F}_{n-s} \not\subset \mathscr{F}_n$. $\square$

Definition 2.3.5 Gestoppter stochastischer Prozess
Sei τ eine endliche Stoppzeit (d. h. $\tau < \infty$ P-fast sicher) und $X = (X_n)_{n \geq 0}$ ein $\mathscr{F}$-adaptierter stochastischer Prozess. Wir definieren:

i) Es sei $X_\tau : \Omega \to \mathbf{R}$ eine Abbildung vermöge $\omega \mapsto X_{\tau(\omega)}(\omega)$. Dies ist der Wert der Trajektorie $X(\omega)$ zum Stoppzeitpunkt $\tau(\omega)$.

ii) Gestoppter Prozess: $X^\tau = (X_n^\tau)_{n \geq 0}$ sei definiert durch

$$X_n^\tau := X_{\tau \wedge n} = \begin{cases} X_n & \text{für } n \leq \tau \\ X_\tau & \text{für } n \geq \tau \end{cases}.$$

Die für die stochastischen Prozesse typische Frage ist nun, in welchem Sinne X_τ und X^τ messbar sind, von welchen Kenntnisständen diese Variablen also abhängen. Dies wollen wir nun diskutieren.

Definition 2.3.6 τ-Vergangenheit
Es sei $(\Omega, \mathscr{A}, P)$ ein Wahrscheinlichkeitsraum, $\mathscr{F} = (\mathscr{F}_n)_{n \geq 0}$ eine Filtration und τ eine $\mathscr{F}$-Stoppzeit. Dann ist

$$\mathscr{F}_\tau := \{A \in \mathscr{A} : A \cap \{\tau \leq n\} \in \mathscr{F}_n \text{ für alle } n \geq 0\}$$

eine σ-Algebra. Diese wird die σ-Algebra der τ-Vergangenheit genannt.

Lemma 2.3.7
Es seien σ und τ Stoppzeiten. Dann gelten die folgenden Aussagen:

i) Gilt $\sigma \le \tau$, so folgt $\mathscr{F}_\sigma \subset \mathscr{F}_\tau$.

ii) Für $n \ge 0$ gilt $\mathscr{F}_{\tau \wedge n} \subset \mathscr{F}_n$.

Beweis: Aussage ii) folgt aus i), da $\sigma := \tau \wedge n$ eine Stoppzeit mit $\sigma \le \tau$ ist. Wir beweisen nun also i). Sei $A \in \mathscr{F}_\sigma$ und $n \ge 0$, dann gilt wegen $\{\tau \le n\} \subset \{\sigma \le n\}$

$$A \cap \{\tau \le n\} = \underbrace{A \cap \{\sigma \le n\}}_{\in \mathscr{F}_n} \cap \underbrace{\{\tau \le n\}}_{\in \mathscr{F}_n} \in \mathscr{F}_n,$$

also ist $A \in \mathscr{F}_\tau$. □

Satz 2.3.8 **Messbarkeit von X_τ und X^τ**
Sei $X = (X_n)_{n \ge 0}$ ein $\mathscr{F}$-adaptierter stochastischer Prozess und τ eine endliche Stoppzeit. Dann gilt:

i) X_τ ist $\mathscr{F}_\tau$-messbar.

ii) X^τ ist $\mathscr{F}$-adaptiert und auch $\mathscr{F}^\tau$-adaptiert, wobei $\mathscr{F}^\tau_n := \mathscr{F}_{\tau \wedge n}$ ist.

Beweis: Wir beweisen zunächst i). Sei $B \subset \mathbf{R}$ messbar. Dann müssen wir $\{X_\tau \in B\} \in \mathscr{F}_\tau$, also $\{X_\tau \in B\} \cap \{\tau \le n\} \in \mathscr{F}_n$ für alle $n \ge 0$ zeigen. Dies geschieht über den folgenden, typischen Ansatz:

$$\begin{aligned}\{X_\tau \in B\} \cap \{\tau \le n\} = \{X_\tau \in B\} \cap \bigcup_{k=0}^n \{\tau = k\} &= \bigcup_{k=0}^n (\{X_\tau \in B\} \cap \{\tau = k\}) \\ &= \bigcup_{k=0}^n \{X_k \in B\} \in \mathscr{F}_n.\end{aligned}$$

Kommen wir zu ii). Da $\tau \wedge n$ eine Stoppzeit ist, ist $X^\tau_n = X_{\tau \wedge n}$ nach i) $\mathscr{F}_{\tau \wedge n}$-messbar und nach Lemma 2.3.7 damit auch $\mathscr{F}_n$-messbar. □

2.4. Optional Stopping Theorem

Sei X ein Martingal und τ eine Stoppzeit. Die Kernaussage wird dann im Wesentlichen sein, dass X^τ ein Martingal ist.

Satz 2.4.1 **Optional Stopping Theorem (Doob)**
Sei $X = (X_n)_{n\geq 0}$ ein $\mathscr{F}$-adaptierter stochastischer Prozess und τ eine endliche $\mathscr{F}$-Stoppzeit. Dann gilt:

i) Ist X ein (Super-/Sub-)Martingal, so ist auch X^τ ein (Super-/Sub-)Martingal.

ii) Ist X ein Martingal, so gilt $\mathbf{E}X_{\tau\wedge n} = \mathbf{E}X_0$ für alle $n \geq 0$. Ist τ P-fast sicher beschränkt, so gilt ferner $\mathbf{E}X_\tau = \mathbf{E}X_0$.

iii) Aussage ii) gilt auch analog für Super-/Sub-Martingale.

Beweis: Wir beweisen zunächst i). Dazu definieren wir $H := (H_n)_{n\geq 1}$ durch $H_n := \mathbf{1}_{\{\tau\geq n\}} = 1 - \mathbf{1}_{\{\tau<n\}} \in \mathscr{F}_{n-1}$, also ist H vorhersagbar. Außerdem gilt $0 \leq H_n \leq 1$ für alle $n \geq 1$, also ist $H_n \cdot (X_n - X_{n-1}) \in \mathscr{L}_1$. Mit Lemma 2.1.9 folgt dann, dass $H.X$ ein Martingal ist. Wir müssen noch $H.X = X^\tau$ zeigen. Dazu wählen wir $n \geq 0$, dann gilt

$$
\begin{aligned}
(H.X)_n &= X_0 + \sum_{k=1}^{n} H_k(X_k - X_{k-1}) = X_0 + \sum_{k=1}^{n} \mathbf{1}_{\{\tau\geq k\}} \cdot (X_k - X_{k-1}) \\
&= X_0 + \sum_{k=1}^{\tau\wedge n} (X_k - X_{k-1}) = X_{\tau\wedge n} = X_n^\tau.
\end{aligned}
$$

Für Super-/Sub-Martingale geht dies analog. Wir kommen nun zu ii). Da X^τ ein Martingal ist, folgt $\mathbf{E}X_{\tau\wedge n} = \mathbf{E}X_n^\tau = \mathbf{E}X_0^\tau = \mathbf{E}X_{\tau\wedge 0} = \mathbf{E}X_0$. Für den zweiten Teil der Aussage sei τ nun P-fast sicher beschränkt, dann gilt $\tau \wedge n \to \tau$ P-fast sicher für $n \to \infty$. Sei nun $N \in \mathbf{N}_0$ mit $\tau \leq N$, dann ist $h = \max_{k=0,\ldots,N} |X_k| \in \mathscr{L}_1$ und $|X_{\tau\wedge n}| \leq h$ für alle $n \geq 0$. Mit dem Satz zur majorisierten Konvergenz folgt dann $\mathbf{E}X_{\tau\wedge n} = \mathbf{E}X_0 \to \mathbf{E}X_\tau$, also gilt $\mathbf{E}X_\tau = \mathbf{E}X_0$. All dies lässt sich analog für Super-/Sub-Martingale zeigen. □

Beispiel 2.4.2 *Symmetrische Irrfahrt*
Seien $(Y_i)_{i\geq 1}$ i. i. d. Zufallsvariablen mit $P(Y_i = 1) = P(Y_i = -1) = \frac{1}{2}$ und $X_n := \sum_{i=1}^n Y_i$ für $n \geq 0$. Ferner seien $a < 0$ und $b > 0$. Dann betrachten wir die Stoppzeiten

$$
\begin{aligned}
\tau_a &:= \inf\{n \geq 0 : X_n = a\}, \\
\tau_b &:= \inf\{n \geq 0 : X_n = b\}, \\
\tau_{a,b} &:= \tau_a \wedge \tau_b = \tau_{(-\infty,a]\cup[b,\infty)}.
\end{aligned}
$$

Zunächst möchten wir dies als ein faires Wettspiel interpretieren. Spieler 1 hat das Kapital $-a > 0$, Spieler 2 das Kapital $b > 0$. Dann gibt $Y_i = 1$ an, dass Spieler 1 gewinnt und eine Geldeinheit von Spieler 2 erhält. Die Stoppzeit $\tau_{a,b}$ gibt dann die Spieldauer bis zum Bankrott einer der Spieler an. Die Menge $A = \{\tau_{a,b} = \tau_a\}$ beschreibt also das Ereignis, dass Spieler 1 verliert, dass der Prozess X also zuerst in a (und nicht in b) ankommt. Wir wollen nun $P(A)$ und die mittlere Spieldauer $\mathbf{E}\tau_{a,b}$ bestimmen.

Unser erstes Ziel ist es, $P(\tau < \infty) = 1$ für $\tau := \tau_{a,b}$ zu zeigen. Dies bedeutet, dass das Spiel P-fast sicher irgendwann endet. Dazu setzen wir $c := b - a$ und $A_k := \bigcap_{i=(k-1)c+1}^{kc} \{Y_i = 1\}$. Tritt A_k ein, so ist das Spiel spätestens zum Zeitpunkt kc beendet, da Spieler 2 pleite ist. Es gilt $A_k \subset \{\tau \le kc\}$ und daraus folgt $\bigcup_{k=1}^n A_k \subset \{\tau \le nc\}$. Ferner ist $(A_k)_k$ eine unabhängige Familie und es gilt $P(A_k) = \left(\frac{1}{2}\right)^c$. Damit erhalten wir

$$\begin{aligned} P(\tau = \infty) &= \lim_{n\to\infty} P(\tau > nc) \le \lim_{n\to\infty} P\left(\bigcap_{k=1}^n A_k^C\right) = \lim_{n\to\infty} \prod_{k=1}^n P(A_k^C) = \lim_{n\to\infty} \left(1 - \left(\frac{1}{2}\right)^c\right)^n \\ &= 0. \end{aligned}$$

Als nächstes wollen wir nun $P(A)$ bestimmen. Klar ist, dass $\tau_{a,b} \wedge n \to \tau_{a,b} =: \tau$ P-fast sicher für $n \to \infty$, $X_n^\tau = X_{\tau\wedge n} \to X_\tau$ P-fast sicher und $\left\|X_n^\tau\right\|_\infty = \|X_{\tau\wedge n}\|_\infty \le \max\{|a|, |b|\}$ gilt. Mit dem Satz zur majorisierten Konvergenz folgt dann $\mathbf{E}X_{\tau\wedge n} \to \mathbf{E}X_\tau$. Satz 2.4.1 sagt nun, dass $\mathbf{E}X_n^\tau = \mathbf{E}X_0^\tau = 0$ gilt und mit obiger Überlegung erhalten wir $\mathbf{E}X_\tau = 0$. Damit folgt

$$0 = \mathbf{E}X_\tau = aP(\tau_{a,b} = \tau_a) + bP(\tau_{a,b} = \tau_b) = aP(A) + b(1 - P(A)).$$

Durch Auflösen der Gleichung erhalten wir $P(A) = \frac{b}{b-a}$.

Das dritte Ziel ist die Berechnung der mittleren Spieldauer $\mathbf{E}\tau_{a,b}$. Da $X_n \in \mathscr{L}_2$ für alle $n \ge 0$ gilt, existiert der quadratische Variationsprozess[1] $\langle X\rangle$ und es gilt $\langle X\rangle_n = \sum_{i=1}^n \mathbf{E}Y_i^2 = n$ für alle $n \ge 0$. Nach Korollar 2.2.6 ist $Z = X^2 - \langle X\rangle$ ein Martingal mit $Z_n = X_n^2 - n$. Mit Satz 2.4.1 folgt nun

$$0 = \mathbf{E}Z_0 = \mathbf{E}Z_n^\tau = \mathbf{E}(\underbrace{X_{\tau\wedge n}^2}_{\to X_\tau^2} - \underbrace{(\tau \wedge n)}_{\to\tau}) \to \mathbf{E}X_\tau^2 - \mathbf{E}\tau,$$

woraus wir $0 = \mathbf{E}X_\tau^2 - \mathbf{E}\tau$ erhalten. Dann ist

$$\mathbf{E}\tau = \mathbf{E}X_\tau^2 = a^2P(A) + b^2(1 - P(A)) = -ab. \quad /\!\!/$$

Satz 2.4.3 **Optional Sampling Theorem I**
Sei $X = (X_n)_{n\ge0}$ ein $\mathscr{F}$-Martingal und $\sigma \le \tau$ seien $\mathscr{F}$-Stoppzeiten. Ist τ beschränkt, so folgt

$$\mathbf{E}(X_\tau \mid \mathscr{F}_\sigma) = X_\sigma.$$

1 Diesen haben wir in Korollar 2.2.6 kennengelernt.

Ist $(\tau_n)_{n\geq 0}$ eine aufsteigende Folge beschränkter Stoppzeiten, so ist $(X_{\tau_n})_{n\geq 0}$ ein $(\mathscr{F}_{\tau_n})_{n\geq 0}$-Martingal. Dabei ist $(\mathscr{F}_{\tau_n})_{n\geq 0}$ nach Lemma 2.3.7 eine Filtration und X_{τ_n} ist nach Satz 2.3.8 $\mathscr{F}_{\tau_n}$-messbar. Die Martingaleigenschaft liefert dann Satz 2.4.3.

Beweis: Es gelte P-fast sicher $\tau \leq N$. Dann ist X_τ integrierbar, da $|X_\tau| \leq \max_{n\leq N} |X_n|$ und $X_n \in \mathscr{L}_1$ gilt. Zunächst wollen wir $\mathbf{E}(X_N \mid \mathscr{F}_\tau) = X_\tau$ P-fast sicher zeigen. Dazu sei $A \in \mathscr{F}_\tau$, d. h. $A \cap \{\tau = k\} \in \mathscr{F}_k$ für alle $k \geq 0$. Dann ist

$$\begin{aligned}
\mathbf{E}(X_N \cdot \mathbf{1}_A) &= \sum_{k=0}^{N} \mathbf{E} X_N \mathbf{1}_A \mathbf{1}_{\{\tau=k\}} = \sum_{k=0}^{N} \mathbf{E} X_N \mathbf{1}_{A\cap\{\tau=k\}} = \sum_{k=0}^{N} \mathbf{E}(\mathbf{1}_{A\cap\{\tau=k\}} \mathbf{E}(X_N \mid \mathscr{F}_k)) \\
&= \sum_{k=0}^{N} \mathbf{E}(\mathbf{1}_{A\cap\{\tau=k\}} X_k) = \sum_{k=0}^{N} \mathbf{E} X_\tau \mathbf{1}_A \mathbf{1}_{\{\tau=k\}} \\
&= \mathbf{E} X_\tau \mathbf{1}_A .
\end{aligned}$$

Dadurch erhalten wir gerade die Gleichung, die wir beweisen wollten. Nun nutzen wir dies für den Beweis des Satzes. Es gilt $\sigma \leq \tau \leq N$ und mit Satz 2.3.7 erhalten wir damit $\mathscr{F}_\sigma \subset \mathscr{F}_\tau$. Dann folgt

$$\mathbf{E}(X_\tau \mid \mathscr{F}_\sigma) = \mathbf{E}(\mathbf{E}(X_N \mid \mathscr{F}_\tau) \mid \mathscr{F}_\sigma) = \mathbf{E}(X_N \mid \mathscr{F}_\sigma) = X_\sigma . \qquad \square$$

2.5. Konvergenzsätze für Martingale I

In diesem Abschnitt wollen wir untersuchen, ob ein X existiert, so dass $X_n \to X$ gilt, falls X_n ein (Super-/Sub-)Martingal ist.

Wir wollen für das kommende Lemma 2.5.1 einen Ansatz diskutieren. Dazu sei $(x_n) \subset \mathbf{R}$ eine Folge, die weder eigentlich noch uneigentlich konvergiert. Dann ist $\liminf x_n < \limsup x_n$ und es seien $a < b$ mit $\liminf x_n < a < b < \limsup x_n$. Das Intervall $[a,b]$ wird von der Folge dann unendlich oft überquert. Um Konvergenz zu zeigen, genügt es daher, die Anzahl der Überquerungen abzuschätzen. In unserem Fall haben wir es natürlich mit Prozessen und nicht mit einfachen Zahlenfolgen zu tun.

Sei nun X ein $\mathscr{F}$-adaptierter stochastischer Prozess und $a < b$. Dann definieren wir $\sigma_0 := 0$, $\tau_k = \inf\{n \geq \sigma_{k-1} : X_n \leq a\}$ und $\sigma_k = \inf\{n \geq \tau_k : X_n \geq b\}$. Der Nachweis, dass dies tatsächlich Stoppzeiten definiert, wird zur Übung überlassen. Ferner gilt $\tau_k \leq \sigma_k \leq \tau_{k+1}$. Nun gibt τ_1 an, wann das erste Mal $X_n \leq a$ gilt, während σ_1 den ersten Zeitpunkt angibt, zu welchem $X_n \geq b$ gilt, nachdem $X_n \leq a$ war. Dann gibt τ_2 den zweiten Zeitpunkt mit $X_n \leq a$ an, nachdem $X_n \geq b$ eingetreten ist et cetera. In Abbildung 2.1 wird dies veranschaulicht.

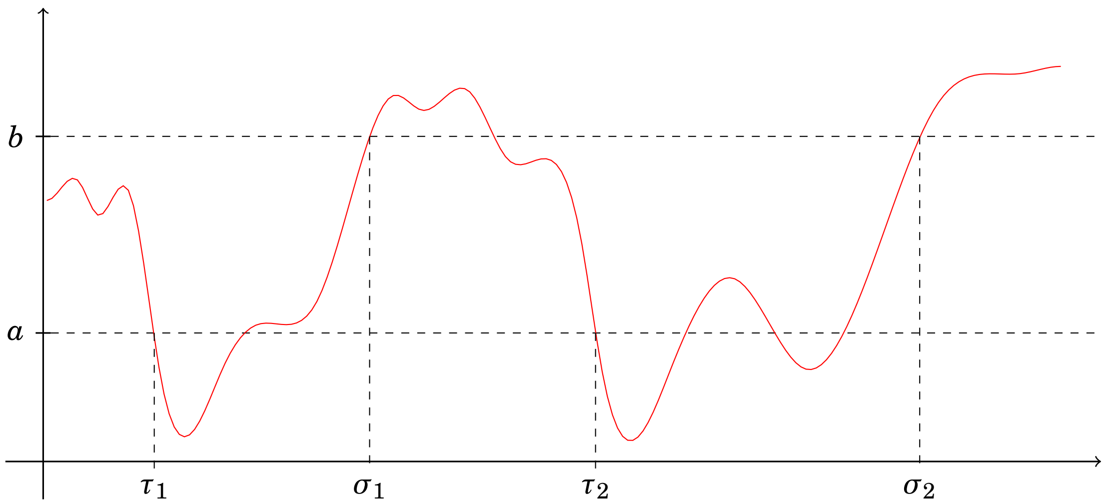

Abbildung 2.1.: Darstellung der Überquerungsstoppzeiten.

Für $N \in \mathbf{N}_0$ betrachten wir ferner $U_{a,b}^N(\omega) := \sup\{k \in \mathbf{N}_0 : \sigma_k(\omega) \leq N\}$, was die Anzahl der Aufwärtsüberquerungen, die wir auch Upcrossings nennen, darstellt. Nach obigem Plan wollen wir nun $U_{a,b}^N$ abschätzen.

Lemma 2.5.1
Sei $X = (X_n)_{n\geq 0}$ ein Sub-Martingal, $N \in \mathbf{N}_0$ und $a < b$. Dann gilt

$$\mathbf{E}U_{a,b}^N \leq \frac{\mathbf{E}(X_N - a)^+}{b-a}.$$

Beweis: Wir betrachten $Y_n := (X_n - a)^+$, dann ist $(Y_n)_n$ ein Sub-Martingal nach Satz 2.2.1 und es gilt $Y_n \geq 0$. Ohne Einschränkung können wir daher $X_n \geq 0$ und $a = 0$ annehmen. Nun setzen wir

$$H_n := \sum_{k=1}^{\infty} \mathbf{1}_{\underbrace{\{\tau_k < n \leq \sigma_k\}}_{\in \mathscr{F}_{n-1}}} \qquad \text{für } n \geq 1,$$

dann ist $(H_n)_n$ ein vorhersagbares Martingal mit $H_n \geq 0$ für alle $n \geq 1$. Ferner gilt $H_n \leq 1$, denn für $\omega \in \{\tau_k < n \leq \sigma_k\}$ gilt wegen $\sigma_k \leq \tau_{k+1}$, dass $\tau_{k+1}(\omega) \geq n$ ist, also ist $\omega \notin \{\tau_{k+1} < n \leq \sigma_{k+1}\}$. Mit Lemma 2.1.9 folgt nun, dass $(1-H).X$ ein Sub-Martingal ist und wir erhalten $\mathbf{E}((1-H).X)_N \geq \mathbf{E}((1-H).X)_0 = \mathbf{E}X_0$. Ferner gilt

$$\begin{aligned}((1-H).X)_N &= X_0 + \sum_{m=1}^{N}(1-H_m)(X_m - X_{m-1}) = X_N - \sum_{m=1}^{N} H_m(X_m - X_{m-1}) \\ &= X_N + X_0 - (H.X)_N.\end{aligned}$$

Daraus folgt $\mathbf{E}(H.X)_N \leq \mathbf{E}X_N$. Weiter gilt nun

$$H_k = \begin{cases} 1 & \text{falls } k \in \{\tau_i + 1, \dots, \sigma_i\} \text{ für ein } i \in \mathbf{N} \\ 0 & \text{sonst} \end{cases}.$$

Für $m \geq 1$ und $\sigma_m < \infty$ gilt dann unter Verwendung dieser Eigenschaft

$$\begin{aligned}(H.X)_{\sigma_m} &= X_0 + \sum_{k=1}^{\sigma_m} H_k(X_k - X_{k-1}) = X_0 + \sum_{i=1}^{m} \sum_{k=\tau_i+1}^{\sigma_i} (X_k - X_{k-1}) \\ &= X_0 + \sum_{i=1}^{m}(X_{\sigma_i} - X_{\tau_i}) \\ &\geq bm,\end{aligned}$$

wobei die letzte Abschötzung wegen $X_{\sigma_i} \geq b$ und $X_{\tau_i} \leq a = 0$ gilt. Sei nun $N \in \mathbf{N}_0$, dann betrachten wir zwei Fälle. Im ersten Fall ist $N \in \{\sigma_m, \dots, \tau_{m+1}\}$ für ein $m \in \mathbf{N}$. Dann gilt $U_{a,b} = m$ und damit folgt

$$(H.X)_N = X_0 + \sum_{k=1}^{N} H_k(X_k - X_{k-1}) = X_0 + \sum_{k=1}^{\sigma_m} H_k(X_k - X_{k-1}) \geq bm = bU_{a,b}^N.$$

Im zweiten Fall sei $N \in \{\tau_m + 1, \dots, \sigma_m\}$ für ein $m \in \mathbf{N}$. Dann ist $U_{a,b}^N = m - 1$ und wir erhalten

$$\begin{aligned}(H.X)_N &= X_0 + \sum_{k=1}^{N} H_k(X_k - X_{k-1}) = X_0 + \sum_{k=1}^{\tau_m} H_k(X_k - X_{k-1}) + \sum_{k=\tau_m+1}^{N} H_k(X_k - X_{k-1}) \\ &= X_0 + \sum_{k=1}^{\sigma_{m-1}} H_k(X_k - X_{k-1}) + X_N - X_{\tau_m} = (H.X)_{\sigma_{m-1}} + X_N - X_{\tau_m} \\ &\geq (H.X)_{\sigma_{m-1}} = b(m-1) = bU_{a,b}^N.\end{aligned}$$

Nehmen wir beide Fälle zusammen, so erhalten wir also $(H.X)_N \geq bU_{a,b}^N$ und damit

$$b\mathbf{E}U_{a,b}^N \leq \mathbf{E}(H.X)_N \leq \mathbf{E}X_N.$$

□

Satz 2.5.2 **Martingalkonvergenz I**
Es sei $(X_n)_{n\geq 0}$ ein Sub-Martingal mit $\sup_{n\geq 0}\mathbf{E}X_n^+ < \infty$. Wir setzen $\mathscr{F}_\infty := \sigma\left(\bigcup_{n=0}^\infty \mathscr{F}_n\right)$. Dann existiert eine $\mathscr{F}_\infty$-messbare Zufallsvariable $X_\infty \in \mathscr{L}_1$ mit $X_n \to X_\infty$ P-fast sicher.

Beweis: Sei $a < b$. Dann gilt $\mathbf{E}(X_n - a)^+ \leq |a| + \mathbf{E}X_n^+ < c < \infty$ für ein $c > 0$ und alle $n \geq 0$. Mit Lemma 2.5.1 folgt nun

$$\mathbf{E}U_{a,b}^n \leq \frac{|a| + \mathbf{E}X_n^+}{b-a} < \frac{c}{b-a} < \infty.$$

Für $n \to \infty$ gilt ferner $U_{a,b}^n \nearrow U_{a,b} := \lim U_{a,b}^n$ P-fast sicher. Mit dem Satz von Beppo Levi[2] folgt dann

$$\mathbf{E}U_{a,b} = \lim_{n\to\infty} \mathbf{E}U_{a,b}^n \leq \frac{c}{b-a} < \infty.$$

Damit erhalten wir $P(U_{a,b} < \infty) = 1$. Sei nun $C^{a,b} := \{\liminf X_n < a\} \cap \{\limsup X_n > b\}$. Es ist klar, dass $X_n(\omega)$ für $\omega \in C^{a,b}$ nicht konvergiert. Ferner gilt offenbar $C^{a,b} \subset \{U_{a,b} = \infty\}$, was aber eine Nullmenge ist. Für $C := \bigcup_{a<b\in\mathbf{Q}} C^{a,b}$ gilt also $P(C) = 0$. Nach Konstruktion ist jedoch X_n auf $\Omega \setminus C$ konvergent. Daher existiert X_∞, so dass P-fast sicher $X_n \to X_\infty$ gilt.

Da X_n wegen $\mathscr{F}_n \subset \mathscr{F}_\infty$ auch $\mathscr{F}_\infty$-messbar ist, folgt, dass X_∞ auch $\mathscr{F}_\infty$-messbar ist. Da X ein Sub-Martingal ist, folgt zudem $\mathbf{E}X_0 \leq \mathbf{E}X_n = \mathbf{E}X_n^+ - \mathbf{E}X_n^-$. Dann ist $\mathbf{E}|X_n| = \mathbf{E}X_n^+ + \mathbf{E}X_n^- \leq \mathbf{E}X_n^+ + \mathbf{E}X_n^+ - \mathbf{E}X_0 < c'$ für ein $c' < \infty$. Mit dem Lemma von Fatou[3] gilt dann $\mathbf{E}|X|_\infty \leq \liminf \mathbf{E}|X_n| \leq c' < \infty$. □

Korollar 2.5.3
Ist X ein nicht-negatives Super-Martingal, so existiert eine $\mathscr{F}_\infty$-messbare Zufallsvariable $0 \leq X_\infty \in \mathscr{L}_1$ mit $\mathbf{E}X_\infty \leq \mathbf{E}X_0$ und $X_n \to X_\infty$ P-fast sicher.

Beweis: Offenbar ist $-X \leq 0$ ein Sub-Martingal. Dann ist $\mathbf{E}(-X_n)^+ \leq 0$ für alle $n \geq 1$. Mit Satz 2.5.2 existiert dann ein $0 \geq -X_\infty \in \mathscr{L}_1$, das $\mathscr{F}_\infty$-messbar ist, mit $-X_n \to -X_\infty$ P-fast sicher. Mit dem Lemma von Fatou[4] gilt zudem $\mathbf{E}X_\infty \leq \liminf \mathbf{E}X_n \leq \mathbf{E}X_0$. □

Aus der Wahrscheinlichkeitstheorie kennen wir verschiedene Konvergenzbegriffe. Da wir nun immer integrierbare Zufallsvariablen vorliegen haben, wollen wir untersuchen, wann $X_n \to X_\infty$ in $\mathscr{L}_1$ gilt. Dazu müssen wir im nächsten Abschnitt jedoch einige Vorbereitungen treffen.

2 Dieser findet sich im Anhang als Satz A.2.
3 Dieses findet sich im Anhang als Lemma A.3.
4 Dieses findet sich im Anhang als Lemma A.3.

2.6. Gleichgradige Integrierbarkeit

Definition 2.6.1 **Gleichgradige Integrierbarkeit**
Eine Familie $(X_i)_{i\in I}$ von Zufallsvariablen heißt gleichgradig integrierbar oder gleichmäßig integrierbar genau dann, wenn

$$\sup_{i\in I} \int_{|X_i|\geq c} |X_i| \, \mathrm{d}P \to 0 \qquad \text{für } c \to \infty.$$

Analog wird die gleichgradige Integrierbarkeit für Mengen definiert. Falls $c \nearrow \infty$ ist, so ist die Konvergenz in Definition 2.6.1 ebenfalls monoton.

Zwischen gleichgradiger und gleichmäßiger Stetigkeit besteht ein Unterschied, während bei der Integrierbarkeit beide Begriffe für den selben Sachverhalt verwendet werden.

Beispiel 2.6.2
Ist $X \in \mathscr{L}_1$, so ist $\{X\}$ gleichgradig integrierbar. //

Lemma 2.6.3
Sind $K_1, \ldots, K_m$ gleichmäßig integrierbar, so ist auch $K_1 \cup \ldots \cup K_m$ gleichmäßig integrierbar. Mit Beispiel 2.6.2 folgt dann insbesondere, dass endliche Mengen integrierbarer Zufallsvariablen gleichmäßig integrierbar sind.

Beweis: Dieser Beweis wird zur Übung überlassen. □

Satz 2.6.4 **Projektionsfamilien**
Sei $Y \in \mathscr{L}_1$ und $\mathscr{F}_i \subset \mathscr{A}$ seien Sub-σ-Algebren für $i \in I \neq \emptyset$. Wir setzen $X_i := \mathbf{E}(Y \mid \mathscr{F}_i)$. Dann ist $(X_i)_{i\in I}$ gleichmäßig integrierbar.

Ist $I = \mathbf{N}_0$ und $(\mathscr{F}_n)_{n\geq 0}$ eine Filtration, so ist insbesondere $(X_n)_{n\geq 0}$ ein gleichgradig integrierbares Martingal. Im nächsten Abschnitt werden wir sehen, dass gleichgradig integrierbare Martingale stets von dieser Form sind.

Beweis: Wir definiren ein Maß $\mu \colon \mathscr{A} \to [0,\infty)$ durch $\mu(A) := \int_A |Y| \, \mathrm{d}P$ für $A \in \mathscr{A}$. Klar ist, dass μ ein endliches Maß mit $\mu \ll P$ ist. Sei nun $\varepsilon > 0$. Mit [WTSkript11, Korollar I.12.3][5] folgt,

5 Auch zu finden in [Meintrup04, Korollar 2.40].

dass es ein $\delta > 0$ gibt, so dass $P(A) < \delta$ impliziert, dass $\mu(A) < \varepsilon$ gilt. Mit der Ungleichung von Markov folgt zudem

$$\begin{aligned}\sup_{i\in I} P(|X_i| \geq c) &\leq \sup_{i\in I} \frac{\mathbf{E}|X_i|}{c} = \sup_{i\in I} \frac{\mathbf{E}|\mathbf{E}(Y \mid \mathscr{F}_i)|}{c} \\ &\leq \sup_{i\in I} \frac{\mathbf{E}|Y|}{c} \to 0\end{aligned}$$

für $c \to \infty$. Daher existiert ein $c > 0$ mit $\sup_{i\in I} P(|X_i| \geq c) < \delta$. Damit folgt

$$\begin{aligned}\sup_{i\in I} \mathbf{E}|X_i|\mathbf{1}_{|X_i|\geq c} &\leq \sup_{i\in I} \mathbf{E}(\mathbf{E}(|Y| \mid \mathscr{F}_i)\mathbf{1}_{|X_i|\geq c}) = \sup_{i\in I} \mathbf{E}(|Y|\mathbf{1}_{|X_i|\geq c}) = \sup_{i\in I} \mu(|X_i| \geq c) \\ &\leq \varepsilon.\end{aligned}$$

□

Lemma 2.6.5 **Gleichmäßige Beschränktheit**
Ist K eine gleichmäßig integrierbare Menge, so ist K auch gleichmäßig beschränkt in $\mathscr{L}_1$, das heißt

$$\sup_{X\in K} \|X\|_1 < \infty.$$

Die Umkehrung gilt im Allgemeinen nicht.

Beweis: Der Beweis des Lemmas wird zur Übung überlassen. □

An dieser Stelle wollen wir noch ein Beispiel dafür geben, dass die Umkehrung von Lemma 2.6.5 im Allgemeinen falsch ist. Betrachte dazu die Gleichverteilung auf $[0,1]$ und die Zufallsvariablen $X_n := n \cdot \mathbf{1}_{[0,\frac{1}{n}]}$. Dann ist $\|X_n\|_1 = 1$ für alle $n \geq 1$, aber für $c \leq n$ gilt auch $\mathbf{E}|X_n|\mathbf{1}_{|X_n|>c} = 1$. Dann ist auch das Supremum davon 1, was jedoch einen Widerspruch darstellt.

Man kann auf $\mathscr{L}_1$ die kleinste Topologie betrachten, bezüglich welcher alle linearen Funktionale die Form $f \mapsto \int fh \, \mathrm{d}P$ für festes $h \in \mathscr{L}_\infty$ haben. Dies ist die sogenannte schwache oder w-Topologie. Dann gibt es den Satz von Dunford-Pettis, der besagt, dass K genau dann gleichmäßig integrierbar ist, wenn $\overline{K}^w$ w-kompakt ist.

Satz 2.6.6 **Charakterisierung I**
Es sei $(X_i)_{i\in I}$ eine Familie von Zufallsvariablen. Dann sind die folgenden Aussagen äquivalent:

i) $(X_i)_{i\in I}$ ist gleichmäßig integrierbar.

ii) $(X_i)_{i\in I}$ ist gleichmäßig beschränkt in $\mathscr{L}_1$ und für alle $\varepsilon > 0$ existiert ein $\delta > 0$, so dass für alle $A \in \mathscr{A}$ gilt:

$$P(A) < \delta \qquad \Rightarrow \qquad \sup_{i\in I} \int_A |X_i| \, \mathrm{d}P < \varepsilon. \tag{N266A2}$$

Beweis: Wir betrachten zunächst die Richtung von i) nach ii). Die gleichmäßige Beschränktheit folgt dann aus Lemma 2.6.5. Sei nun $\varepsilon > 0$, dann existiert aufgrund der gleichmäßigen Integrierbarkeit ein $c > 0$ mit $\int_{|X_i|\geq c} |X_i| \, \mathrm{d}P < \varepsilon$ für alle $i \in I$. Nun setzen wir $\delta := \frac{\varepsilon}{c}$, für $A \in \mathscr{A}$ mit $P(A) < \delta$ gilt dann

$$\begin{aligned}\int_A |X_i| \, \mathrm{d}P &= \int_{A\cap\{|X_i|<c\}} |X_i| \, \mathrm{d}P + \int_{A\cap\{|X_i|\geq c\}} |X_i| \, \mathrm{d}P \\ &\leq cP(A) + \varepsilon \\ &\leq 2\varepsilon.\end{aligned}$$

Nun beweisen wir die Richtung von ii) nach i). Sei $\varepsilon > 0$ und $\delta > 0$ so, dass (N266A2) erfüllt ist. Wir setzen $c := \frac{1}{\delta} \sup_{i\in I} \int |X_i| \, \mathrm{d}P < \infty$. Mit der Ungleichung von Markov folgt dann

$$P(|X_i| \geq c) \leq \frac{\mathbf{E}|X_i|}{c} \leq \delta.$$

Wenden wir (N266A2) nun auf $A := \{|X_i| \geq c\}$ an, so erhalten wir für alle $i \in I$

$$\int_{|X_i|\geq c} |X_i| \, \mathrm{d}P < \varepsilon. \qquad \square$$

Satz 2.6.7 **Charakterisierung II**
Es sei $(X_n)_{n\geq 1} \subset \mathscr{L}_1$ und X_∞ eine Zufallsvariable. Dann sind die folgenden Aussagen äquivalent:

i) Die Familie $(X_n)_{n\geq 1}$ ist gleichmäßig integrierbar und $X_n \to X_\infty$ stochastisch.

ii) Es gilt $X_\infty \in \mathscr{L}_1$ und $X_n \to X_\infty$ in $\mathscr{L}_1$.

Beweis: Wir beweisen zunächst die Richtung von i) nach ii). In der Wahrscheinlichkeitstheorie haben wir gelernt, dass eine Teilfolge X_{n_k} mit $X_{n_k} \to X_\infty$ P-fast sicher existiert. Mit dem Lemma von Fatou[6] folgt dann

$$\mathbf{E}|X_\infty| \leq \liminf \mathbf{E}|X_{n_k}| < \infty,$$

wobei der letzte Schritt wegen der gleichmäßigen Beschränktheit gilt. Dann ist also $X_\infty \in \mathscr{L}_1$. Ferner betrachten wir nun $\{X_n : n \in \overline{\mathbf{N}}\}$. Diese Menge ist gleichmäßig integrierbar nach Beispiel

6 Dieses findet sich im Anhang als Lemma A.3.

2.6.2 und Lemma 2.6.3. Sei nun $\varepsilon > 0$. Mit Satz 2.6.6 folgt dann, dass es ein $\delta > 0$ gibt, so dass (N266A2) erfüllt ist für unsere Menge. Da $X_n \to X_\infty$ stochastisch, existiert ein n_0, so dass für alle $n \geq n_0$ $P(|X_n - X_\infty| > \varepsilon) < \delta$ gilt. Für $A := \{|X_n - X_\infty| > \varepsilon\}$ gilt mit (N266A2) dann $\int_A |X_n| \, \mathrm{d}P < \varepsilon$ für alle $n \geq n_0$ oder $n = \infty$. Dann ist

$$\begin{aligned} \int |X_n - X_\infty| \, \mathrm{d}P &\leq \int_{|X_n - X_\infty| < \varepsilon} |X_n - X_\infty| \, \mathrm{d}P + \int_{|X_n - X_\infty| > \varepsilon} |X_n - X_\infty| \, \mathrm{d}P \\ &\leq \varepsilon + \int_{|X_n - X_\infty| > \varepsilon} |X_n| + |X_\infty| \, \mathrm{d}P \\ &\leq 3\varepsilon. \end{aligned}$$

Kommen wir zur Richtung von ii) nach i). Nach unsereren Kenntnissen aus der Wahrscheinlichkeitstheorie genügt es, gleichmäßige Integrierbarkeit zu zeigen. Da (X_n) in $\mathscr{L}_1$ konvergiert, ist (X_n) gleichmäßig beschränkt. Daher zeigen wir noch (N266A2) aus Satz 2.6.6. Sei also $\varepsilon > 0$. Da X_∞ gleichmäßig integrierbar ist, folgt die Existenz eines $\delta > 0$, so dass für alle $A \in \mathscr{A}$ mit $P(A) < \delta$ gilt, dass $\int_A |X_\infty| \, \mathrm{d}P < \varepsilon$. Sei nun $n_0 \geq 1$ mit $\|X_n - X_\infty\|_1 < \varepsilon$ für alle $n \geq n_0$. Für $A \in \mathscr{A}$ mit $P(A) < \delta$ folgt dann

$$\begin{aligned} \int_A |X_n| \, \mathrm{d}P &\leq \int_A |X_n - X_\infty| \, \mathrm{d}P + \int_A |X_\infty| \, \mathrm{d}P \\ &\leq 2\varepsilon. \end{aligned}$$

Da ferner $\{X_1, \dots, X_{n_0 - 1}\}$ gleichmäßig integrierbar ist, existiert mit Satz 2.6.6 ein $\delta' > 0$ mit

$$P(A) < \delta' \qquad \Rightarrow \qquad \sup_{n \leq n_0 - 1} \int_A |X_n| \, \mathrm{d}P < \varepsilon.$$

Setzen wir nun $\delta'' := \min\{\delta, \delta'\}$, so sind wir fertig. □

Satz 2.6.8 **Majorante**
Es sei $(X_i)_{i \in I}$ eine Familie von Zufallsvariablen und $Y \in \mathscr{L}_1$ mit $|X_i| \leq Y$ für alle $i \in I$. Dann ist $(X_i)_{i \in I}$ gleichmäßig integrierbar.

Beweis: Der Beweis wird zur Übung überlassen. □

Korollar 2.6.9 **Majorisierte Konvergenz**
Es sei $X_n \to X$ stochastisch und $Y \in \mathscr{L}_1$ mit $|X_n| \leq Y$ für alle $n \geq 1$. Dann gilt $X_n \to X$ in $\mathscr{L}_1$ und $\mathbf{E}X_n \to \mathbf{E}X$.

Beweis: Satz 2.6.8 zeigt, dass $(X_n)_{n \geq 1}$ gleichmäßig integrierbar ist. Aus Satz 2.6.7 folgt dann, dass $X \in \mathscr{L}_1$ und $X_n \to X$ in $\mathscr{L}_1$ gilt. Damit ist

$$|\mathbf{E}X_n - \mathbf{E}X| \leq \mathbf{E}|X_n - X| \to 0.$$ □

Satz 2.6.10 **Beschränktheit**
Sei $(X_i)_{i\in I}$ eine Familie von Zufallsvariablen, $\phi\colon \mathbf{R} \to [0,\infty)$ messbar und $\lim_{t\to\infty} \frac{\phi(t)}{t} = \infty$. Gilt $K := \sup_{i\in I} \mathbf{E}\phi \circ |X_i| < \infty$, so ist $(X_i)_{i\in I}$ gleichmäßig beschränkt.

Beweis: Sei $\varepsilon > 0$ und setze $M := \frac{K}{\varepsilon}$. Dann existiert $c > 0$ mit $\frac{\phi(t)}{t} \geq M$ für alle $t \geq c$ und wir erhalten

$$\int_{|X_i|\geq c} |X_i| \,\mathrm{d}P \leq \frac{1}{M} \sup_{|X_i|\geq c} \phi(|X_i|) \,\mathrm{d}P \leq \frac{1}{M} K = \varepsilon.$$ □

Ist ϕ monoton wachsend und konvex und gilt $\frac{\phi(t)}{t} \to \infty$, so sind die Bedingungen für Satz 2.6.10 erfüllt. Ferner gibt es den Satz de la Vallée-Poussin: Ist (X_i) gleichmäßig integrierbar, so gilt $\sup_{i\in I} \mathbf{E}\phi \circ |X_i| < \infty$ für alle konvexen und wachsenden ϕ, die zudem $\frac{\phi(t)}{t} \to \infty$ erfüllen.

Korollar 2.6.11 **$\mathscr{L}_p$-Beschränktheit**
Ist $(X_i)_{i\in I} \subset \mathscr{L}_p$ für ein $p > 1$ gleichmäßig beschränkt, so ist $(X_i)_{i\in I}$ gleichmäßig integrierbar.

Satz 2.6.11 gilt auch für $p = \infty$, nicht aber für $p = 1$.

Beweis: Wir setzen $\phi(t) := t^p$ und wenden Satz 2.6.10 an, da $\mathbf{E}\phi \circ |X_i| = \|X_i\|_p^p$. □

2.7. Konvergenzsätze für Martingale II

Satz 2.7.1 **$\mathscr{L}_1$-Martingalkonvergenz**
Es sei $(X_n)_{n\geq 0}$ ein gleichmäßig integrierbares Sub-Martingal, dann existiert ein $X_\infty \in \mathscr{L}_1$ mit $X_n \to X_\infty$ P-fast sicher und in $\mathscr{L}_1$.

Beweis: Da die Familie $(X_n)_{n\geq 0}$ gleichmäßig integrierbar ist, ist sie auch gleichmäßig beschränkt in $\mathscr{L}_1$ und es gilt $\sup_{n\geq 0} \mathbf{E}|X_n| < \infty$. Da $\mathbf{E}X_n^+ \leq \mathbf{E}X_n^+ + \mathbf{E}X_n^- = \mathbf{E}|X_n|$ gilt, folgt $\mathbf{E}X_n^+ < \infty$. Mit Satz 2.5.2 folgt dann die Existenz einer Zufallsvariable $X_\infty \in \mathscr{L}_1$ mit $X_n \to X_\infty$ P-fast sicher. Mit Satz 2.6.7 folgt dann, dass diese Konvergenz auch in $\mathscr{L}_1$ gilt. □

Es gilt zu beachten, dass X_∞ auch bezüglich $\mathscr{F}_\infty := \sigma(\bigcup_{n\geq 0} \mathscr{F}_n)$ messbar ist.

Satz 2.7.2 **Charakterisierung gleichmäßig integrierbarer Martingale**
Sei $(X_n)_{n\geq 0}$ ein Martingal. Dann sind die folgenden Aussagen äquivalent:

i) $(X_n)_{n\geq 0}$ ist gleichmäßig integrierbar.

ii) Es existiert eine $\mathscr{F}_\infty$-messbare Zufallsvariable $X_\infty \in \mathscr{L}_1$ mit $X_n = \mathbf{E}(X_\infty \mid \mathscr{F}_n)$ für $n \geq 0$.

iii) $(X_n)_{n\geq 0}$ konvergiert in $\mathscr{L}_1$ gegen eine Zufallsvariable X_∞.

Beweis: Mit Satz 2.7.1 folgt i) ⇒ iii). Für iii) ⇒ ii) gilt wegen $X_n \to X_\infty$ in $\mathscr{L}_1$, dass $X_\infty \in \mathscr{L}_1$ gilt und diese Zufallsvariable auch $\mathscr{F}_\infty$-messbar ist. Zu zeigen ist also noch die angegebene Identität. Sei dazu $A \in \mathscr{F}_n$, für $m \geq n$ folgt wegen $\mathbf{E}(X_m \mid \mathscr{F}_n) = X_n$ dann

$$\mathbf{E}X_n \mathbf{1}_A = \mathbf{E}(\mathbf{E}(X_m \mid \mathscr{F}_n)\mathbf{1}_A) = \mathbf{E}\mathbf{E}(X_m \mathbf{1}_A \mid \mathscr{F}_n) = \mathbf{E}X_m \mathbf{1}_A.$$

Damit folgt

$$|\mathbf{E}X_n \mathbf{1}_A - \mathbf{E}X_\infty \mathbf{1}_A| = |\mathbf{E}X_m \mathbf{1}_A - \mathbf{E}X_\infty \mathbf{1}_A| \leq \mathbf{E}|X_m - X_\infty| \to 0.$$

Nun wollen wir noch ii) ⇒ i), diese Richtung haben wir jedoch bereits in Satz 2.6.4 bewiesen. □

Satz 2.7.3 **$\mathscr{L}_1$-Konvergenz ⇒ P-fast sichere Konvergenz**
Sei $(X_n)_{n\geq 0}$ ein Sub-Martingal und $X_\infty \in \mathscr{L}_1$ mit $X_n \to X_\infty$ in $\mathscr{L}_1$. Dann gilt $X_n \to X_\infty$ P-fast sicher und $X_n \leq \mathbf{E}(X_\infty \mid \mathscr{F}_n)$ für alle $n \geq 0$. Wird insbesondere ein Martingal vorausgesetzt, so gilt nach Satz 2.7.2 sogar Gleichheit.

Beweis: Aus der $\mathscr{L}_1$-Konvergenz folgt $\sup_{n\geq 0} \mathbf{E}|X_n| < \infty$ und damit auch $\sup_{n\geq 0} \mathbf{E}X_n^+ < \infty$. Mit Satz 2.5.2 folgt dann die Existenz von $X \in \mathscr{L}_1$ mit $X_n \to X$ P-fast sicher. Dass tatsächlich $X = X_\infty$ gilt, ist klar. Für den Beweis der Abschätzung sei nun $\delta > 0$ und $m \geq n \geq 0$. Mit dem Satz von Markov gilt dann

$$\begin{aligned} P(\mathbf{E}(X_\infty \mid \mathscr{F}_n) - X_n < -\delta) &\leq P(\mathbf{E}(X_\infty \mid \mathscr{F}_n) - \mathbf{E}(X_m \mid \mathscr{F}_n) < -\delta) \\ &\leq \frac{\mathbf{E}|\mathbf{E}(X_\infty - X_m \mid \mathscr{F}_n)|}{\delta} \\ &\leq \frac{\|X_\infty - X_m\|}{\delta} \xrightarrow{m\to\infty} 0. \end{aligned}$$

Für $\delta \to 0$ erhalten wir dann die behauptete Abschätzung. □

Satz 2.7.4 **Projektionsmartingale**
Sei $X \in \mathscr{L}_1$ und $X_n := \mathbf{E}(X \mid \mathscr{F}_n)$ für $n \geq 0$. Dann gilt $\lim\limits_{n\to\infty} X_n = \mathbf{E}(X \mid \mathscr{F}_\infty)$. In unser üblichen Notation bedeutet dies also $X_\infty = \mathbf{E}(X \mid \mathscr{F}_\infty)$.

Beweis: Wir wissen aus Satz 2.6.4, dass $(X_n)_{n\geq 0}$ gleichmäßig integrierbar ist. Dann existiert nach Satz 2.7.1 eine $\mathscr{F}_\infty$-messbare Zufallsvariable $X_\infty \in \mathscr{L}_1$ mit $X_n \to X_\infty$ in $\mathscr{L}_1$. Aus Satz 2.7.3 wissen wir dann, dass diese Konvergenz auch P-fast sicher gilt. Betrachte nun die Maße $\mu(A) := \int_A X \, \mathrm{d}P$ und $\mu_\infty(A) := \int_A X_\infty \, \mathrm{d}P$ auf $\mathscr{F}_\infty$. Wir wollen nun $\mu = \mu_\infty$ zeigen, denn daraus folgt unmittelbar die gewünschte Identität. Für $A \in \mathscr{F}_n$ gilt mit Satz 2.7.2 nun $\mathbf{E}(X_\infty \mid \mathscr{F}_n) = X_n = \mathbf{E}(X \mid \mathscr{F}_n)$. Daraus folgt $\mu = \mu_\infty$ auf $\mathscr{F}_n$ für alle $n \geq 0$. Da $\bigcup_{n\geq 0} \mathscr{F}_N$ ein $\cap$-stabiles Erzeugendensystem von $\mathscr{F}_\infty = \sigma(\bigcup_{n\geq 0} \mathscr{F}_n)$ bildet, folgt damit $\mu = \mu_\infty$ auf $\mathscr{F}_\infty$. □

Als nächstes betrachten wir ein in $\mathscr{L}_p$ gleichmäßig beschränktes Martingal $(X_n)_{n\geq 0}$. Dann gilt $X_n \to X_\infty$ in $\mathscr{L}_1$ und P-fast sicher. Wir wollen uns mit der Frage beschäftigen, ob auch Konvergenz in $\mathscr{L}_p$ vorliegt.

Sei $(X_n)_{n\geq 0}$ ein stochstischer Prozess, dann schreiben wir

$$X_n^* := \sup_{k\leq n} X_k$$

und

$$|X|_n^* := \sup_{k\leq n} |X_k|.$$

Wir formulieren nun folgendes Lemma, das im Wesentlichen eine Verschärfung der Ungleichung von Markov darstellt und überlassen den Beweis zur Übung:

Lemma 2.7.5
Ist $(X_n)_{n\geq 0}$ ein Sub-Martingal und $\lambda > 0$, so folgt

$$\lambda P(X_n^* \geq \lambda) \leq \mathbf{E}X_n \mathbf{1}_{\{X_n^* \geq \lambda\}}.$$

Lemma 2.7.6
Seien $X, Y \geq 0$ Zufallsvariablen mit

$$\lambda P(X \geq Y) \leq \mathbf{E} Y \mathbf{1}_{\{X \geq \lambda\}} \tag{*}$$

für alle $\lambda > 0$. Dann gilt für $p > 1$ und $p' := \frac{p}{p-1}$ die Abschätzung

$$\|X\|_p \leq p' \|Y\|_p .$$

Beweis: Ohne Einschränkung sei $Y \in \mathscr{L}_p$. Im ersten Fall sei auch $X \in \mathscr{L}_p$. Dann gilt

$$\|X\|_p^p = \mathbf{E} X^p = \mathbf{E}\left(\int_0^X p t^{p-1} \, \mathrm{d}t\right) = \mathbf{E}\left(\int_0^\infty p t^{p-1} \mathbf{1}_{\{X \geq t\}} \, \mathrm{d}t\right)$$

und mit dem Satz von Fubini-Tonelli erhalten wir

$$\begin{aligned} &= \int_0^\infty p t^{p-1} \mathbf{E} \mathbf{1}_{\{X \geq t\}} \, \mathrm{d}t \\ &\leq \int_0^\infty p t^{p-2} \mathbf{E} Y \mathbf{1}_{\{X \geq t\}} \, \mathrm{d}t \end{aligned}$$

und durch erneute Anwendung des Satzes von Fubini-Tonelli

$$= \mathbf{E} Y \int_0^\infty p t^{p-2} \mathbf{1}_{\{X \geq t\}} \, \mathrm{d}t = \mathbf{E} Y \int_0^X p t^{p-2} \, \mathrm{d}t = p' \mathbf{E} Y X^{p-1}.$$

Mittels der Ungleichung von Hölder gilt nun

$$\leq p' \|Y\|_p \left\|X^{p-1}\right\|_{p'} = p' \|Y\|_p \|X\|_p^{p-1} .$$

Teilen wir durch $\|X\|_p^{p-1}$, so erhalten wir die gewünschte Aussage.

Im zweiten Fall sei nun $X \geq 0$ allgemein. Dann betrachten wir $X_n := X \wedge n \in \mathscr{L}_p$ für $n \geq 1$. Wir wollen zeigen, dass X_n die Abschätzung (*) erfüllt. Für $\lambda \leq n$ gilt $X_n \geq \lambda \Leftrightarrow X \wedge n \geq \lambda \Leftrightarrow X \geq \lambda$. Daher folgt $\{X_n \geq \lambda\} = \{X \geq \lambda\}$ und (*) gilt für X_n. Für $\lambda > n$ gilt $\{X_n \geq \lambda\} = \emptyset$ und daher gilt auch hier (*) für X_N. Damit folgt aus dem ersten Fall

$$\|X_n\|_p \leq p' \|Y\|_p .$$

Da $X_n \nearrow X$ gilt, folgt die gewünschte Abschätzung mit dem Satz von Beppo Levi[7]. □

7 Dieser findet sich im Anhang als Satz A.2.

Satz 2.7.7 **Doobsche Ungleichung**
Sei $(X_n)_{n\geq 0}$ ein Martingal oder ein nicht-negatives Sub-Martingal. Dann gilt:

i) Für $p \geq 1$, $\lambda > 0$ und $n \geq 0$ gilt
$$\lambda^p P(|X|_n^* \geq \lambda) \leq \mathbf{E}|X_n|^p.$$

ii) Für $p > 1$ und $n \geq 0$ gilt
$$\|X_n\|_p \leq \left\| |X|_n^* \right\|_p \leq p' \|X_n\|_p .$$

Wir wollen zunächst eine kurze Fassung des Satzes angeben:

i) $X_n \in \mathscr{L}_p \Longrightarrow |X|_n^* \in \mathscr{L}_{p,\infty}$.
ii) $|X|_n^* \in \mathscr{L}_p \Longleftrightarrow X_n \in \mathscr{L}_p$.

Beweis: Betrachten wir zunächst den Fall, dass X ein Martingal ist. Mit Satz 2.2.2 folgt dann, dass $(|X_n|^p)_{n\geq 0}$ ein Sub-Martingal ist. Der Beweis dieses Satzes funktioniert völlig analog aber auch für nicht-negative Sub-Martingale. Mit Lemma 2.7.5 folgt dann die Behauptung.

Im zweiten Fall ist die erste Ungleichung trivial, die zweite Ungleichung folgt aus Lemma 2.7.5 und 2.7.6. □

Satz 2.7.8 **$\mathscr{L}_p$-Konvergenz für Martingale**
Sei $(X_n)_{n\geq 0}$ ein Martingal, das für $p > 1$ gleichmäßig $\mathscr{L}_p$-beschränkt ist. Dann existiert $X_\infty \in \mathscr{L}_p$ mit $X_n \to X_\infty$ P-fast sicher und in $\mathscr{L}_p$.

Beweis: Mit Korollar 2.6.11 wissen wir, dass (X_n) gleichmäßig integrierbar ist. Satz 2.7.2 und Satz 2.7.3 liefern dann die Existenz von $X_\infty \in \mathscr{L}_1$ mit $X_n \to X_\infty$ P-fast sicher und in $\mathscr{L}_1$. Satz 2.7.7 zeigt dann
$$\left\| |X|_n^* \right\|_p \leq p' \|X_n\|_p < p'k < \infty$$
für alle $n \geq 0$ und geeignetes k. Für $X^* := \sup_{n\geq 0} |X_n|$ gilt $|X|_n^* \nearrow X^*$ und mit dem Satz von Beppo Levi[8] folgt dann $\|X^*\|_p \leq p'k$. Dies bedeutet aber gerade $(X^*)^p \in \mathscr{L}_1$. Dann gilt $|X_n - X_\infty| \leq 2X^* \in \mathscr{L}_p$, also $X_\infty \in \mathscr{L}_p$. Ferner gilt $|X_n - X_\infty|^p \leq 2^p (X^*)^p \in \mathscr{L}_1$ und daher $|X_n - X_\infty|^p \to 0$ P-fast sicher. Mit dem Satz zur majorisierten Konvergenz folgt dann $\mathbf{E}|X_n - X_\infty|^p \to 0$. □

8 Dieser findet sich im Anhang als Satz A.2.

2.8. Optional Sampling

Das Ziel dieses Abschnittes ist es, Optional Sampling für beliebige Stoppzeiten zu betreiben. Dies wird jedoch nicht für allgemeine, sondern für gleichmäßig integrierbare Martingale gelten.

Sei dazu $(X_n)_{n\geq 0}$ ein gleichmäßig integrierbares Martingal. Dann existiert $X_\infty \in \mathscr{L}_1$ mit $X_n \to X_\infty$ P-fast sicher und in $\mathscr{L}_1$. Ferner ist $X_n = \mathbf{E}(X_\infty \mid \mathscr{F}_n)$. Betrachte nun eine beliebige Stoppzeit $\tau\colon \Omega \to \overline{\mathbf{N}_0}$. Für $\omega \in \Omega$ mit $\tau(\omega) = \infty$ setzen wir $X_\tau(\omega) := X_\infty(\omega)$, ansonsten wie gewohnt als $X_\tau(\omega) := X_{\tau(\omega)}(\omega)$. Bis jetzt haben wir in Satz 2.4.3 gesehen, dass wenn $\sigma \leq \tau$ Stoppzeiten sind und τ beschränkt ist, folgt, dass $\mathbf{E}(X_\tau \mid \mathscr{F}_\sigma) = X_\sigma$ ist. Auf die Voraussetzung der Beschränktheit wollen wir nun verzichten:

Satz 2.8.1 **Optional Sampling Theorem II**
Sei $(X_n)_{n\geq 0}$ ein gleichmäßig integrierbares Martingal und $\sigma \leq \tau$ Stoppzeiten. Dann gilt $X_\sigma, X_\tau \in \mathscr{L}_1$ und

$$\mathbf{E}(X_\tau \mid \mathscr{F}_\sigma) = X_\sigma.$$

Beweis: Zuerst wollen wir $X_\tau \in \mathscr{L}_1$ zeigen. Da $(X_n)_{n\geq 0}$ gleichmäßig integrierbar ist, folgt mit Satz 2.7.2, dass $X_n = \mathbf{E}(X_\infty \mid \mathscr{F}_n)$ für alle $n \geq 0$ gilt. Daraus folgt $|X_n| \leq \mathbf{E}(|X_\infty| \mid \mathscr{F}_n)$. Für $n \geq 0$ erhalten wir so

$$\mathbf{E}|X_n|\mathbf{1}_{\{\tau=n\}} \leq \mathbf{E}\mathbf{E}(|X_\infty| \mid \mathscr{F}_n)\mathbf{1}_{\{\tau=n\}} = \mathbf{E}|X_\infty|\mathbf{1}_{\{\tau=n\}}. \tag{*}$$

Ferner gilt

$$|X_\tau|\mathbf{1}_{\{\tau\leq n\}} \nearrow |X_\tau|\mathbf{1}_{\{\tau<\infty\}} \tag{**}$$

und

$$|X_\tau|\mathbf{1}_{\{\tau\leq n\}} = \left|\sum_{k=0}^{n} X_\tau \mathbf{1}_{\{\tau=k\}}\right| \leq \sum_{k=0}^{n} |X_k|\mathbf{1}_{\{\tau=k\}}. \tag{***}$$

Mit dem Satz von Beppo Levi[9] folgt nun

$$\begin{aligned}
\mathbf{E}|X_\tau|\mathbf{1}_{\{\tau<\infty\}} &\overset{(**)}{=} \lim_{n\to\infty} \mathbf{E}|X_\tau|\mathbf{1}_{\{\tau\leq n\}} \\
&\overset{(***)}{\leq} \sum_{k=0}^{\infty} \mathbf{E}|X_\tau|\mathbf{1}_{\{\tau=k\}} \\
&\overset{(*)}{\leq} \sum_{k=0}^{\infty} \mathbf{E}|X_\infty|\mathbf{1}_{\{\tau=k\}} \\
&= \mathbf{E}|X_\infty|\mathbf{1}_{\{\tau<\infty\}} \\
&\leq \mathbf{E}|X_\infty| < \infty.
\end{aligned}$$

9 Dieser findet sich im Anhang als Satz A.2.

Außerdem ist $\mathbf{E}|X_\tau|\mathbf{1}_{\{\tau=\infty\}} \le \mathbf{E}|X_\infty|$ und insgesamt erhalten wir $\mathbf{E}|X_\tau| \le 2\mathbf{E}|X_\infty| < \infty$, also gilt $X_\tau \in \mathscr{L}_1$. Analog gilt dies natürlich auch für $X_\sigma \in \mathscr{L}_1$. Im zweiten Schritt zeigen wir nun $\mathbf{E}(X_\tau \mid \mathscr{F}_\sigma) = X_\tau$. Dazu verwenden wir Satz 2.3.8, der zeigte, dass X_τ auch $\mathscr{F}_\sigma$-messbar ist. Der Satz zeigte dies nur für endliche Stoppzeiten, der Beweis ist nahezu wortwörtlich aber auch auf unsere Situation hier übertragbar. Wir wollen nun zunächst $\mathbf{E}(X_\infty \mathbf{1}_A) = \mathbf{E}(X_\tau \mathbf{1}_A)$ für alle $A \in \mathscr{F}_\tau$ zeigen. Sei dazu $n \ge 0$, dann sind $n \ge \tau \wedge n$ endliche Stoppzeiten und mit Satz 2.4.3 erhalten wir $\mathbf{E}(X_n \mid \mathscr{F}_{\tau\wedge n}) = X_{\tau\wedge n}$. Daraus folgt

$$\mathbf{E}(X_\infty \mid \mathscr{F}_{\tau\wedge n}) = \mathbf{E}(\mathbf{E}(X_\infty \mid \mathscr{F}_n) \mid \mathscr{F}_{\tau\wedge n}) = X_{\tau\wedge n}. \tag{$\heartsuit$}$$

Ferner gilt für $A \in \mathscr{F}_\tau$ nun $A \cap \{\tau \le n\} = A \cap \{\tau \le \tau \wedge n\} \in \mathscr{F}_{\tau\wedge n}$ und es folgt

$$\begin{aligned} \mathbf{E}X_\infty \mathbf{1}_{A\cap\{\tau\le n\}} &= \mathbf{E}(\mathbf{E}X_\infty \mathbf{1}_{A\cap\{\tau\le n\}} \mid \mathscr{F}_{\tau\wedge n}) \\ &= \mathbf{E}(\mathbf{E}(X_\infty \mid \mathscr{F}_{\tau\wedge n})\mathbf{1}_{A\cap\{\tau\le n\}}) \\ &\overset{\heartsuit}{=} \mathbf{E}X_{\tau\wedge n}\mathbf{1}_{A\cap\{\tau\le n\}}. \end{aligned} \tag{$\heartsuit\heartsuit$}$$

Ohne Einschränkung sei $X_\infty \ge 0$, sonst zerlegen wir wie üblich $X_\infty = X_\infty^+ - X_\infty^-$ und führen das folgende Argument zweimal durch. Nun folgt $X_n = \mathbf{E}(X_\infty \mid \mathscr{F}_n) \ge 0$ und $X_\tau \ge 0$. Mit dem Satz von Beppo Levi[10] und dem Satz von Lebesgue[11] folgt wegen

$$X_\infty \mathbf{1}_{A\cap\{\tau\le n\}} \longrightarrow X_\infty \mathbf{1}_{A\cap\{\tau<\infty\}}$$

und

$$X_{\tau\wedge n}\mathbf{1}_{A\cap\{\tau\le n\}} \longrightarrow X_\tau \mathbf{1}_{A\cap\{\tau<\infty\}}$$

daher mit Hilfe von ($\heartsuit\heartsuit$) die folgende Identität:

$$\underbrace{\mathbf{E}X_\infty \mathbf{1}_{A\cap\{\tau\le n\}}}_{\to \mathbf{E}X_\infty \mathbf{1}_{A\cap\{\tau<\infty\}}} = \underbrace{\mathbf{E}X_{\tau\wedge n}\mathbf{1}_{A\cap\{\tau\le n\}}}_{\to \mathbf{E}X_\tau \mathbf{1}_{A\cap\{\tau<\infty\}}}.$$

Nach Definition gilt außerdem $X_\tau \mathbf{1}_{\{\tau=\infty\}} = X_\infty = X_\infty \mathbf{1}_{\{\tau=\infty\}}$, wir erhalten damit also nun $X_\tau \mathbf{1}_{A\cap\{\tau=\infty\}} = X_\infty \mathbf{1}_{A\cap\{\tau=\infty\}}$ und insgesamt $\mathbf{E}X_\infty \mathbf{1}_A = \mathbf{E}X_\tau \mathbf{1}_A$. Dann folgt $\mathbf{E}(X_\infty \mid \mathscr{F}_\tau) = X_\tau$ und damit

$$\mathbf{E}(X_\tau \mid \mathscr{F}_\sigma) = \mathbf{E}(\mathbf{E}(X_\infty \mid \mathscr{F}_\tau) \mid \mathscr{F}_\sigma) = \mathbf{E}(X_\infty \mid \mathscr{F}_\sigma) = X_\sigma. \qquad \square$$

10 Dieser findet sich im Anhang als Satz A.2.
11 Dieser findet sich im Anhang als Satz A.4.

2.9. Anwendungen

Martingale können auf vielen verschiedenen Gebieten eingesetzt werden, zum Beispiel:

i) Beweis der Gesetze der großen Zahlen.

ii) Beweis des Satzes von Radon-Nikodym (vgl. [Klenke06, Satz 11.17]).

iii) Viele Anwendungen aus der Praxis, die nicht theoretischer Natur sind.

iv) Andere Fragen auf dem Gebiet der stochastischen Prozesse.

In diesem Abschnitt wollen wir nun ein paar Beispielanwendungen der Martingaltheorie diskutieren.

Verzweigungsprozesse

Unser Ziel ist die Modellierung der Größe einer Population ohne Sex. Dazu bezeichne $X_{n,i}$ die Anzahl der Nachkommen des i-ten Individuums in der n-ten Generation. Ferner setzen wir $Z_0 := 1$ und $Z_{n+1} := \sum_{i=1}^{Z_n} X_{n,i}$, diese Größe beschreibt dann die Anzahl der Individuen in der $(n+1)$-ten Generation. Man nennt dies einen Galton-Watson-Prozess, falls gilt:

i) Die $(X_{n,i})$ sind i. i. d. – das heißt also, dass Fruchtbarkeit nicht vererbt wird.

ii) Es gilt $P(X_{1,1} = k) = p_k$ mit $\sum_{k=0}^{\infty} p_k = 1$.

iii) Es gilt $M := \mathbf{E}X_{1,1} < \infty$ und $\sigma^2 := \operatorname{Var} X_{1,1} < \infty$.

iv) Es ist $\mathscr{F}_n := \sigma\{X_{k,i} : k < n, i \in \mathbf{N}\}$, dann ist Z_n $\mathscr{F}_n$-messbar und $(Z_n)_{n\geq 0}$ ist $(\mathscr{F}_n)_{n\geq 0}$-adaptiert.

Wir setzen nun $W_n := M^{-n} Z_n$. Die Idee ist es hierbei, dass bei konstanter Nachkommenzahl M exakt $Z_n = M^n$ gelten würde und der Faktor M^{-n} den Prozess daher normiert. Wir fragen uns nun, wie sich Z_n in Abhängigkeit von M verhält. Ist $Z_n(\omega) = 0$ für ein $n \in \mathbf{N}$, so gilt $Z_m(\omega) = 0$ für alle $m \geq n$, da eine einmal ausgestorbene Population nicht mehr wachsen kann. Außerdem ist $Z_n(\omega) = 0$ genau dann, wenn $W_n(\omega) = 0$ ist.

Lemma 2.9.1
W ist ein Martingal und es gilt $\mathbf{E}Z_n = M^n$.

Beweis: Für $n \geq 0$ gilt

$$\begin{aligned}\mathbf{E}(W_{n+1} \mid \mathscr{F}_n) &= M^{-(n+1)}\mathbf{E}\left(\sum_{i=1}^{Z_n} X_{n,i} \mid \mathscr{F}_n\right) = M^{-(n+1)}\mathbf{E}\left(\sum_{k=1}^{\infty} \mathbf{1}_{\{Z_n=k\}} \sum_{i=1}^{Z_n} X_{n,i} \mid \mathscr{F}_n\right)\\ &= M^{-(n+1)} \sum_{k=1}^{\infty} \mathbf{E}\left(\mathbf{1}_{\{Z_n=k\}} \sum_{i=1}^{k} X_{n,i} \mid \mathscr{F}_n\right) = M^{-(n+1)} \sum_{k=1}^{\infty} \mathbf{1}_{\{Z_n=k\}} \mathbf{E}\left(\sum_{i=1}^{k} X_{n,i} \mid \mathscr{F}_n\right)\\ &= M^{-(n+1)} \sum_{k=1}^{\infty} \mathbf{1}_{\{Z_n=k\}} \cdot k \cdot M\\ &= M^{-n} Z_n = W_n.\end{aligned}$$

Dabei wurde verwendet, dass $X_{n,i}$ von $\mathbf{F}_n$ unabhängig ist. Nun folgt insbesondere

$$M^{-n}\mathbf{E}Z_n = \mathbf{E}W_n = \mathbf{E}W_0 = \mathbf{E}Z_0 = 1. \qquad \square$$

Satz 2.9.2
Es existiert $W_\infty := \lim_{n\to\infty} W_n$ P-fast sicher. Falls $\operatorname{Var} X_{1,1} \in (0,\infty)$ gilt (also keine konstante Nachkommenschaft vorliegt), folgt aus $M \leq 1$, dass $\mathbf{E}W_\infty = 0$ ist und für $M > 1$ folgt, dass $\mathbf{E}W_\infty = 1$ ist.

Mit anderen Worten sagt der zweite Teil des Satzes aus, dass die Population genau dann nicht ausstirbt, wenn $M > 1$ ist.

Beweis: Da $W_n \geq 0$ für alle $n \geq 0$ gilt, folgt mit Korollar 2.5.3, dass $W_\infty = \lim_{n\to\infty} W_n$ existiert. Wir betrachten nun zunächst $M < 1$, dann ist $\mathbf{E}Z_n = M^n \to 0$ und wegen $Z_n \geq 0$ folgt $Z_n \to 0$ in $\mathscr{L}_1$. Mit der Markov-Ungleichung erhalten wir dann $P(Z_n > 0) = P\left(Z_n > \frac{1}{2}\right) \to 0$. Damit folgt wegen $\{Z_n = 0\} \subset \{Z_m = 0\}$ für $m \geq n$

$$P\left(\bigcup_{n=1}^{\infty} \{Z_n = 0\}\right) = \lim_{n\to\infty} P(Z_n = 0) = 1.$$

Sei nun $\omega \in \bigcup_{n=1}^{\infty} \{Z_n = 0\}$. Dann existiert ein n, so dass für alle $m \geq n$ mit $0 = Z_m(\omega)$ folgt, dass $0 = W_m(\omega)$ ist. Also ist $W_\infty(\omega) = 0$ P-fast sicher und damit $W_\infty = 0$ P-fast sicher.

Wir betrachten nun den Fall $M = 1$, dann ist $Z = W$ und damit $W_n(\omega) \in \mathbf{N}$ für alle ω und n. Diese Eigenschaft gilt dann P-fast sicher auch für W_∞. Für $k \geq 0$ setzen wir $A_k := \{W_\infty = k\}$. Wir wollen zeigen, dass $P(A_k) = 0$ für alle $k \geq 1$ gilt, da dann $P(A_0) = 1$ folgt. Dazu führen wir einen Widerspruchsbeweis und nehmen an, dass ein $k \geq 1$ mit $P(A_k) > 0$ existiert. Ferner setzen wir $B_n := \bigcap_{i=1}^k \{X_{n,i} = 0\}$. Da $\mathbf{E}X_{1,1} = 1$ und $\operatorname{Var} X_{1,1} > 0$ gilt, folgt $P(X_{1,1} = 0) =: \nu > 0$. Da die X_i unabhängig sind, folgt nun $P(B_n) = \nu^k > 0$. Wir erhalten $\sum_{n=1}^{\infty} P(B_n) = \infty$ und mit dem Satz von Borel-Cantelli[12] damit, dass $P(\limsup B_n) = 1$ gilt. Dann folgt $P(A_k \cap \limsup B_n) > 0$ und ferner ist $Z_n = W_n \to W_\infty$ P-fast sicher, daher folgt

$$P(\{\lim Z_n = W_\infty\} \cap A_k \cap \limsup B_n) > 0.$$

Dann existiert also $\omega \in \Omega$ mit $Z_n(\omega) \to W_\infty(\omega)$, $\omega \in A_k$ und $\omega \in \limsup B_n$. Nun folgt aber, dass es ein $n_0 \in \mathbf{N}$ gibt, so dass $Z_n(\omega) = k$ für alle $n \geq n_0$ gilt. Aus der Definition der Z_n wissen wir dann, dass $k = \sum_{i=1}^k X_{n,i}(\omega)$ für alle $n \geq n_0$ und

$$\omega \in \limsup B_n = \bigcap_{l\geq 1} \bigcup_{m \geq l} \bigcap_{i=1}^k \{X_{m,i} = 0\} \subset \bigcup_{m \geq n_0} \bigcap_{i=1}^k \{X_{m,i} = 0\}$$

gilt, das heißt es existiert $m \geq n_0$ mit $X_{m,i}(\omega) = 0$ für alle $m \geq n_0$ und $i = 1, \ldots, k$. Dann folgt aber

$$k = \sum_{i=1}^k X_{m,i}(\omega) = 0,$$

was im Widerspruch zu unserer Voraussetzung steht.

Wir kommen nun zum letzten Fall und betrachten $M > 1$. Es gilt $\mathbf{E}Z_{n-1} = M^{n-1}$ und die Familie $Z_{n-1}, X_{n-1,1}, X_{n-1,2}, \ldots$ ist unabhängig. Wie man leicht zeigen kann, gilt

$$\begin{aligned}\operatorname{Var} Z_n = \operatorname{Var} \sum_{i=1}^{Z_{n-1}} X_{n-1,i} &= (\mathbf{E}X_{1,1})^2 \operatorname{Var} Z_{n-1} + \mathbf{E}Z_{n-1} \operatorname{Var} X_{1,1} \\ &= M^2 \operatorname{Var} Z_{n-1} + M^{n-1}\sigma^2.\end{aligned}$$

Ferner kann man durch Nachrechnen zeigen, dass

$$\operatorname{Var} W_n = M^{-2n} \operatorname{Var} Z_n = \operatorname{Var} W_{n-1} + M^{-(n+1)}\sigma^2$$

gilt. Außerdem ist $\operatorname{Var} W_0 = \operatorname{Var} Z_0 = 0$ und mittels vollständiger Induktion kann man

$$\operatorname{Var} W_n = \sigma^2 \sum_{k=2}^{n+1} M^{-k} \leq \sigma^2 \frac{M}{M-1} < \infty$$

für alle $n \geq 1$ zeigen. Insgesamt erhalten wir damit, dass $(W_n)_{n\geq 0}$ gleichmäßig $\mathscr{L}_2$-beschränkt ist. Dann folgt, dass $W_n \to W_\infty$ in $\mathscr{L}_2$ und damit auch in $\mathscr{L}_1$ gilt. Wir erhalten nun $1 = \mathbf{E}Z_0 = \mathbf{E}W_0 = \mathbf{E}(\mathbf{E}(W_\infty \mid \mathscr{F}_0)) = \mathbf{E}W_\infty$. $\square$

12 Dieser findet sich im Anhang als Satz A.5.

3
Markovketten

Markovketten sind stochastische Prozesse, deren Zukunft nur von der Gegenwart und nicht von der Vergangenheit abhängt. Erstmals wurden solche Prozesse 1906–1908 von Andrej Markov bei der Untersuchung von Verallgemeinerungen des starken Gesetzes der großen Zahlen betrachtet. Markovketten spielen in der Praxis eine große Rolle.

3.1. Definitionen, Beispiele und einfache Eigenschaften

Im Folgenden sei $(X_n)_{n\geq 0}$ ein stochastischer Prozess mit Werten in einer höchsten abzählbaren Menge S. Auf S betrachten wir die σ-Algebra $\mathscr{P}(S)$. Ferner sei $\mathscr{F}$ die von $(X_n)_{n\geq 0}$ induzierte, natürliche Filtration. Wir erinnern uns daran, dass gilt:

$$P(X \in A \mid \mathscr{B}) = \mathbf{E}(\mathbf{1}_A \circ X \mid \mathscr{B})$$

Definition 3.1.1 **Markovkette**
Ein stochastischer Prozess $(X_n)_{n\geq 0}$ heißt Markovkette genau dann, wenn für alle $A \subset S$ und alle $n \geq 0$ die Identität

$$P(X_{n+1} \in A \mid X_0, \ldots, X_n) = P(X_{n+1} \in A \mid X_n)$$

gilt. Das Bildmaß P_{X_0} heißt dann Startverteilung.

Hierbei verwenden wir die übliche Schreibweise $P(A \mid X_n) := P(A \mid \sigma(X_n))$.

Die unmittelbare Zukunft von Markovketten hängt also lediglich von ihrem Zustand in der Gegenwart ab.

Lemma 3.1.2
Die folgenden Aussagen sind äquivalent:

i) Der Prozess $(X_n)_{n\geq 0}$ ist eine Markovkette.

ii) Für alle $A \subset S$, $n \geq 0$ und $i_0, \ldots, i_n \in S$ gilt

$$P(X_{n+1} \in A \mid X_0 = i_0, \ldots, X_n = i_n) = P(X_{n+1} \in A \mid X_n = i_n).$$

An dieser Stelle wollen wir die Konvention festlegen, dass wir immer, wenn wir von „$P(\quad \mid B)$ für alle B mit der Eigenschaft …“ sprechen, implizit meinen, dass dies für alle B mit der Eigenschaft aus Lemma 3.1.2 und $P(B) > 0$ gelten soll. Insbesondere gilt dies auch, wenn wir von „für alle $i_0, \ldots, i_n \in S$ gilt …“ sprechen.

Beweis: Wir setzen $B_j := \{X_j = i_j\}$ und $B := \bigcap_{j=0}^n B_j$. Nehmen wir $P(B) > 0$ an, so folgt $P(B_j) > 0$. Im Abschnitt 1.5 haben wir bei der heuristischen Herangehensweise für bedingte Erwartungen gezeigt, dass P-fast sicher

$$P(X_{n+1} \in A \mid \mathscr{F}_n)(\omega) = P(X_{n+1} \in A \mid B) \qquad \text{für } \omega \in B$$

gilt. Analog gilt

$$P(X_{n+1} \in A \mid X_n)(\omega) = P(X_{n+1} \in A \mid B_n) \qquad \text{für } \omega \in B_n.$$

Für i) $\Rightarrow$ ii) gilt wegen $P(B) > 0$, dass es ein $\omega \in B$ mit den obigen Formeln geben muss. Für ii) $\Rightarrow$ i) gilt, dass $(\{X_0 = i_0, \ldots, X_n = i_n\})_{i_0, \ldots, i_n \in S}$ eine abzählbare Partition von Ω ist. Auf jeder Partitionszelle B mit $P(B) > 0$ gilt nach obigen Überlegungen und ii)

$$P(X_{n+1} \in A \mid \mathscr{F}_n) \overset{P\text{-f.s.}}{=} P(X_{n+1} \in A \mid B) \overset{\text{ii)}}{=} P(X_{n+1} \in A \mid B_n) \overset{B \subset B_n}{=} P(X_{n+1} \in A \mid X_n).$$

Vereinigen wir diese Partitionszellen, so erhalten wir die Behauptung. □

Korollar 3.1.3
Es sei $(X_n)_{n\geq 0}$ eine Markovkette, dann gilt für alle $n \geq 0$ und $i_0, \ldots, i_n \in S$

$$P(X_0 = i_0, \ldots, X_n = i_n) = P(X_0 = i_0) \cdot P(X_1 = i_1 \mid X_0 = i_0) \cdot \ldots \cdot P(X_n = i_n \mid X_{n-1} = i_{n-1}).$$

Beweis: Mit Lemma 3.1.2 gilt

$$\begin{aligned} P(X_0 = i_0, \ldots, X_n = i_n) &= P(X_n = i_n \mid X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) \cdot P(X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) \\ &= P(X_n = i_n \mid X_{n-1} = i_{n-1}) \cdot P(X_{n-1} = i_{n-1}, \ldots, X_0 = i_0). \end{aligned}$$

Führt man dieses Argument induktiv fort, so erhalten wir die Aussage des Korollars. □

Die Wahrscheinlichkeit des Verhaltens des Pfades bis zur Zeit n ist also gleich dem Produkt der Startverteilung und den Übergangswahrscheinlichkeiten $P(X_{j+1} = i_{j+1} \mid X_j = i_j)$.

Satz 3.1.4 **Chapman-Kolmogorov**
Sei $(X_n)_{n\geq 0}$ eine Markovkette und $k < m < n$, dann gilt für $i, j \in S$

$$P(X_n = j \mid X_k = i) = \sum_{l\in S} P(X_n = j \mid X_m = l) \cdot P(X_m = l \mid X_k = i).$$

Beweis: Es ist

$$\begin{aligned} P(X_n = j, X_k = i) &= \sum_{l\in S} P(X_n = j, X_m = l, X_k = i) \\ &= \sum_{l\in S} P(X_k = i, X_m = l) \cdot P(X_n = j \mid X_k = i, X_m = l) \\ &= \sum_{l\in S} P(X_k = i, X_m = l) \cdot P(X_n = j \mid X_m = l). \end{aligned}$$

Teilt man durch $P(X_k = i)$, so erhält man die gewünschte Identität. □

Beispiel 3.1.5 *Summen unabhängiger Zufallsvariablen*
Seien $X_0, Y_1, Y_2, \dots$ unabhängige Zufallsvariablen mit Werten in $\mathbf{Z}^d$. Wir setzen

$$X_n = X_0 + \sum_{i=1}^{n} Y_i$$

und erhalten dadurch die Markovkette $(X_n)_{n\geq 0}$. //

Beispiel 3.1.6 *Simples stochastisches Wettermodell*
Wir nehmen an, dass es drei Wetterzustände gibt:

i) Ein verregnetes Wetter kodieren wir mit 1.

ii) Ist es bewölkt, so kodieren wir dies mit 2.

iii) Einen sonnigen Tag kodieren wir mit 3.

Ferner sei eine Matrix $(p_{i,j})_{i,j=1}^{3}$ von Übergangswahrscheinlichkeiten „$p_{i,j} = P(\text{Morgen } j \mid \text{Heute } i)$“ gegeben. Beispielsweise könnte diese Matrix in Los Alamos und Bremen wie in Abbildung 3.1 gegeben sein. Dieses Beispiel wird in den Übungen ausführlich behandelt. //

	1	2	3
1	0.001	0.0499	0.95
2	0.001	0.001	0.998
3	0.001	0.002	0.997

	1	2	3
1	0.95	0.04	0.01
2	0.8	0.19	0.01
3	0.8	0.18	0.02

Abbildung 3.1.: Exemplarische Übergangsmatrizen des Wettermodells für Los Alamos (links) und Bremen (rechts).

Beispiel 3.1.7 *Einfaches Warteschlangenmodell*
Seien $n = 0, 1, \dots$ Zeitpunkte, zu denen ein Skilift einen Skifahrer mitnehmen kann. Zwischen den Zeitpunkten n und $n+1$ kommen Y_n neue Skifahrer an den Lift. Die Länge der Warteschlange am Lift sei $(X_n)_{n\geq 0}$, dann gilt $X_0 = 0$ und $X_n = \max\{0, X_{n-1} - 1\} + Y_{n-1}$. Sind die Y_i unabhängig, so ist $(X_n)_{n\geq 0}$ eine Markovkette. //

Beispiel 3.1.8 *Stepping Stone Model*
Es seien $k \geq 2$ Farben und eine $(m \times m)$-Matrix $(a_{i,j})_{i,j=1}^m$ mit $a_{i,j} \in \{1, \dots, k\}$ für alle i und j gegeben. Ist S die Menge aller solchen Matrizen, so ist $|S| = k^{m^2}$, die Größe wächst also sehr schnell. Nun zum Übergang von n nach $n+1$: Für $a_{i,j}^{(n)}$ wählen wir zufällig einen der acht Nachbarn, wobei das Quadrat als Torusoberfläche gedeutet wird. Diesen Nachbarn wollen wir $a_{\tilde{i},\tilde{j}}^{(n)}$ nennen und setzen $a_{i,j}^{(n+1)} := a_{\tilde{i},\tilde{j}}^{(n)}$. Dies führen wir für alle i, j durch und erhalten so eine Markovkette. Jetzt kann man untersuchen, was für $n \to \infty$ passiert. //

Dabei ist in den Beispielen 3.1.6 und 3.1.8 eigentlich unklar, ob es zu gegebenen Übergangswahrscheinlichkeiten wirklich eine Markovkette gibt. Dies wollen wir im nächsten Abschnitt aber zeigen.

3.2. Homogene Markovketten

Bisher kann die Übergangswahrscheinlichkeit $P(X_{n+1} \in A \mid X_n = i)$ noch von n abhängen. Diese Abhängigkeit wollen wir bei homogenen Markovketten verbieten.

Definition 3.2.1 **Homogene Markovkette**
Eine Markovkette $(X_n)_{n\geq 0}$ heißt homogen genau dann, wenn $P(X_{n+1} = j \mid X_n = i) = P(X_1 = j \mid X_0 = i)$ für alle $n \geq 0$ und alle $i, j \in S$ gilt. In diesem Fall setzen wir die Übergangswahrscheinlichkeit $p_{i,j} := P(X_1 = j \mid X_0 = i)$ und $\alpha_i := P(X_0 = i)$. Dann ist $\mathbb{P} := (p_{i,j})_{i,j\in S}$ die Übergangsmatrix und $\alpha := (\alpha_i)_{i\in S}$ die Startverteilung.

Homogene Markovketten lassen sich gut mit Graphen veranschaulichen. Ein Beispiel wird in Abbildung 3.2 gegeben. Pfeile bedeuten hierbei echt positive Übergangswahrscheinlichkeiten.

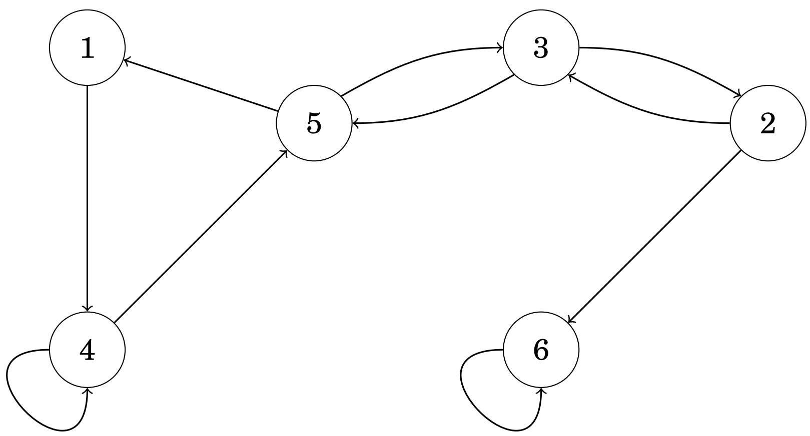

Abbildung 3.2.: Darstellung einer homogenen Markovkette. Das Vorhandensein eines Pfeiles symbolisiert eine positive Übergangswahrscheinlichkeit.

Lemma 3.2.2 **Stochastische Matrix**
Ist $(X_n)_{n\geq 0}$ eine homogene Markovkette, so ist $\mathbb{P} = (p_{i,j})_{i,j\in S}$ eine stochastische Matrix, das heißt es gilt

i) $p_{i,j} \geq 0$ für alle $i, j \in S$.

ii) $\sum_{j\in S} p_{i,j} = 1$ für alle $i \in S$.

Lemma 3.2.3 **Chapman-Kolmogorov II**
Ist $(X_n)_{n\geq 0}$ eine homogene Markovkette mit der Übergangsmatrix $\mathbb{P}$ und der Startverteilung α, so gilt

$$P(X_{n+m} = j \mid X_n = i) = p_{i,j}^{(m)},$$

wobei $\mathbb{P}^m = (p_{i,j}^{(m)})_{i,j\in S}$ die m-te Potenz von $\mathbb{P}$ ist. Insbesondere ist $p_{i,j}^{(1)} = p_{i,j}$ und

$$p_{i,j}^{(m+1)} = \sum_{l\in S} p_{i,l}^{(m)} p_{l,j}^{(1)}.$$

Ferner gilt

$$P(X_m = j) = (\mathbb{P}^m \alpha)_j = \sum_{i\in S} \alpha_i p_{i,j}^{(m)}.$$

Beweis: Wir führen eine Induktion über m für $n = 0$ durch. Dies genügt, da die Markovkette homogen ist. Für $m = 1$ ist die Aussage offensichtlich, dann ist

$$\begin{aligned} P(X_{m+1} = j \mid X_0 = i) &= \sum_{l\in S} P(X_{m+1} = j \mid X_m = l) P(X_m = l \mid X_0 = i) \\ &= \sum_{l\in S} p_{l,j} p_{i,l}^{(m)}. \end{aligned}$$

Ferner gilt mit der Formel von der totalen Wahrscheinlichkeit[1]

$$P(X_m = j) = \sum_{i\in S} P(X_m = j \mid X_0 = i) P(X_0 = i) = \sum_{i\in S} p_{i,j}^{(m)} \alpha_i. \qquad \square$$

Wahrscheinlichkeiten von homogenen Markovketten lassen sich also durch Matrixmultiplikationen berechnen. Für nicht-homogene Markovketten wird diese Potenz zu einem Produkt verschiedener Matrizen. Bis jetzt haben wir also homogene Markovketten vorgegeben und gewinnen daraus $\mathbb{P}$ und α. Wir wollen uns nun fragen, ob wir umgekehrt aus gegebenen $\mathbb{P}$ und α auch eine Markovketten erhalten. Dies bejahen wir in folgendem Satz:

Satz 3.2.4
Ist $\mathbb{P}$ eine stochastische Matrix über S und $\alpha = (\alpha_i)_{i\in S}$ mit $\alpha_i \geq 0$ für alle $i \in S$ und $\sum_{i\in S} \alpha_i = 1$, so existiert eine S-wertige Markovkette $(X_n)_{n\geq 0}$ mit $P(X_1 = j \mid X_0 = i) = p_{i,j}$ für alle $i, j \in S$ und $P(X_0 = i) = \alpha_i$ für alle $i \in S$.

Beweis: Der Beweis wird hier nicht geführt, findet sich aber in [Meintrup04, Satz A.11]. $\square$

1 Diese findet sich im Anhang unter Satz A.1.

Insbesondere existieren also die Markovketten aus Beispiel 3.1.6 und Beispiel 3.1.8.

Beispiel 3.2.5 *Eindimensionale Irrfahrt*
Es sei $(Y_n)_{n\geq 0}$ mit $P(Y_n = 1) =: p > 0$, $P(Y_n = -1) = 1 - p =: q > 0$ und $(Y_n)_{n\geq 0}$ sei unabhängig. Dann setzen wir $X_n := \sum_{i=0}^{n} Y_i$. In Beispiel 3.1.5 haben wir bereits gesehen, dass dies eine Markovkette ist. Außerdem kann man leicht sehen, dass sie homogen ist. Für $i \in \mathbf{Z}$ gilt

$$p_{i,i+1} = P(X_{n+1} = i+1 \mid X_n = i) = p$$

und

$$p_{i,i-1} = P(X_{n+1} = i-1 \mid X_n = i) = 1 - p.$$

Damit ist also $p_{i,j} = 0$ für $i \in S$ und $j \notin \{i \pm 1\}$. Außerdem ist $\alpha_1 = p$, $\alpha_{-1} = 1 - p$ und $\alpha_j = 0$ für alle anderen j.

Eine Variante ist die Irrfahrt mit absorbierendem Rand, dann ist $S = \{-1, \ldots, N\}$. Es gilt $p_{i,i+1} = p$ und $p_{i,i-1} = 1 - p$, falls $-1 < i < N$ ist, sowie $p_{-1,-1} = 1$ und $p_{N,N} = 1$. //

Beispiel 3.2.6 *Galton-Watson-Prozess*
Der im Abschnitt 2.9 betrachtete Galton-Watson-Prozess ist ebenfalls eine homogene Markovkette. //

Definition 3.2.7 **Absorbierender Zustand**
Ein Zustand $i \in S$ heißt absorbierend genau dann, wenn $p_{i,i} = 1$ gilt.

3.3. Die starke Markoveigenschaft

In diesem Abschnitt wollen wir weitere Untersuchungen der Zukunft im Vergleich zur Gegenwart oder Vergangenheit anstellen.

Lemma 3.3.1
Es seien $A,B,C_i \subset \Omega$ messbar für $i \in I$, wobei I eine abzählbare Indexmenge ist. Ferner seien die $(C_i)_{i\in I}$ paarweise disjunkt und es existiere ein $p \in (0,1]$ mit $P(A \mid B \cap C_i) = p$ für alle $i \in I$. Dann gilt für $C := \bigcup_{i\in I} C_i$

$$P(A \mid B \cap C) = p.$$

Beweis: Es gilt

$$\begin{aligned} pP(B\cap C) &= \sum_{i\in I} pP(B\cap C_i) = \sum_{i\in I} P(A \mid B\cap C_i)P(B\cap C_i) = \sum_{i\in I} P(A\cap B\cap C_i) = P(A\cap B\cap C) \\ &= P(A \mid B\cap C)P(B\cap C). \end{aligned}$$

□

Satz 3.3.2 **Markoveigenschaft**
Sei $(X_n)_{n\geq 0}$ eine (homogene) Markovkette, dann gilt für $0 < n < m$, $i_n \in S$, $V \subset S^n$ und $Z \subset S^{m-n}$ die Identität

$$P((X_{n+1},\dots,X_m) \in Z \mid X_n = i_n, (X_0,\dots,X_{n-1}) \in V) = P((X_{n+1},\dots,X_m) \in Z \mid X_n = i_n).$$

Eine beliebige Zukunft Z hängt also nur von der Gegenwart i_n ab, nicht jedoch von der Vergangenheit V. Es ist jedoch essentiell, dass die Gegenwart nur durch einen einzigen Zustand beschrieben wird – „$X_n = i_n$“ darf also nicht durch „$X_n \in \tilde{G}$“ ersetzt werden.

Beweis: Wegen der σ-Additivität genügt es, $Z = \{(i_{n+1},\dots,i_n)\}$ zu betrachten. Seien nun $i_0,\dots,i_{n-1} \in S$, dann gilt

$$\begin{aligned} P((X_{n+1},\dots,X_m) \in Z \mid X_n = i_n,\dots,X_0 = i_0) &= \frac{P(X_0 = i_0,\dots,X_m = i_m)}{P(X_0 = i_0,\dots,X_n = i_n)} \overset{3.1.3}{=} \frac{\alpha_{i_0}\prod_{j=0}^{m-1} p_{i_j i_{j+1}}}{\alpha_{i_0}\prod_{j=0}^{n-1} p_{i_j i_{j+1}}} \\ &= \prod_{j=n}^{m-1} p_{i_j i_{j+1}}, \end{aligned}$$

also ist dieser Ausdruck von $i_0,\dots,i_{n-1}$ unabhängig. Da sich $\{(X_0,\dots,X_{n-1}) \in V\}$ disjunkt in $\{X_0 = i_0,\dots,X_{n-1} = i_{n-1}\}$ und analog auch $\{(X_0,\dots,X_{n-1}) \in S^n\}$ zerlegen lassen, folgt die Behauptung mit Lemma 3.3.1.

□

Korollar 3.3.3
Sei $(X_n)_{n\geq 0}$ eine homogene Markovkette und $Z \subset \bigotimes_{n=0}^{\infty} \mathscr{P}(S)$. Für $i \in S$ und $V \in S^{n+1}$ gilt dann

$$P((X_{n+1}, X_{n+2}, \ldots) \in Z \mid X_n = i_n, (X_0, \ldots, X_n) \in V) = P((X_1, X_2, \ldots) \in Z \mid X_0 = i_0).$$

Beweis: Der Beweis verwendet Homogenität und wird wie im Beweis von Satz 3.3.2 auf einem $\cap$-stabilen Erzeugendensystem durchgeführt. □

Die bedingte Zukunft einer homogenen Markovkette unterscheidet sich also nicht von der bedingten Zukunft einer neu gestarteten homogenen Markovkette.

Satz 3.3.4 **Starke Markoveigenschaft**
Es sei $(X_n)_{n\geq 0}$ eine homogene Markovkette und $\tau : \Omega \to \overline{\mathbf{N}_0}$ eine Stoppzeit mit $P(\tau < \infty) = 1$. Dann gilt für jedes $V \in \mathscr{F}_\tau$, $Z \in \bigotimes_{n=0}^{\infty} \mathscr{P}(S)$ und $i \in S$ die Eigenschaft

$$P((X_{\tau+1}, X_{\tau+2}, \ldots) \in Z \mid X_\tau = i, V) = P((X_1, X_2, \ldots) \in Z \mid X_0 = i).$$

Satz 3.3.4 sagt also, dass die Zukunft $X_\tau = i$ das selbe wie das Verhalten nach der Zeit 0 ist. Man spricht an dieser Stelle auch von einer „Wiedergeburt“.

Beweis: Es genügt, eine Zukunft der Form $Z = \{i_1\} \times \ldots \times \{i_n\} \times S \times \ldots$ zu betrachten, da diese ein $\cap$-stabiles Erzeugendensystem von $\bigotimes_{n=0}^{\infty} \mathscr{P}(S)$ bilden. Ferner überlegen wir uns für disjunkte C_i

$$\begin{aligned} P(A \cap C \mid B) &= \sum_i P(A \cap C_i \mid B) = \sum_i \frac{P(A \cap C_i \cap B)}{P(B)} = \sum_i \frac{P(C_i \cap B)}{P(B)} \frac{P(A \cap C_i \cap B)}{P(C_i \cap B)} \\ &= \sum_i P(C_i \mid B) P(A \mid B \cap C_i). \end{aligned}$$

Mit $C = \bigcup C_i$ für $C_i = \{\tau = i\}$ gilt dann $C = \Omega$ und wir erhalten

$$\begin{aligned} P(\underbrace{X_{\tau+1} = i_1, \ldots, X_{\tau+m} = i_m}_{=:A\cap C} \mid \underbrace{X_\tau = i, V}_{=:B}) \\ &= \sum_{l=0}^{\infty} P(\tau = l \mid X_\tau = i, V) P(X_{\tau+1} = i_1, \ldots, X_{\tau+m} = i_m \mid X_\tau = i, V, \tau = l) \\ &= \sum_{l=0}^{\infty} P(\tau = l \mid X_\tau = i, V) P(X_{l+1} = i_1, \ldots, X_{l+m} = i_m \mid X_l = i, V, \tau = l) \\ &\overset{3.3.3}{=} \underbrace{\sum_{l=0}^{\infty} P(\tau = l \mid X_\tau = i, V)}_{=P(\Omega \mid X_\tau = i, V) = 1} \underbrace{P(X_1 = i_1, \ldots, X_m = i_m \mid X_0 = i)}_{\text{von } l \text{ unabhängig}}. \end{aligned}$$

Dies ist aber gerade die Behauptung. □

3.4. Klassifikation von Zuständen

Wir wollen nun ein Beispiel diskutieren, um verschiedene Arten von Zuständen kennenzulernen. Dazu betrachten wir die homogene Markovkette, die in Abbildung 3.3 dargestellt wird. Wir sehen, dass wir von Zustand 3 aus in jeden anderen Zustand gelangen können, dieser die Kette aber in zwei Gebiete aufteilt, zwischen denen man nie wieder wechseln kann. Zwischen Zustand 1 und Zustand 2 können wir beliebig oft wechseln und zwischen den Zustanden 4, 5 und 6 gibt es eine Art „Kreisverkehr“.

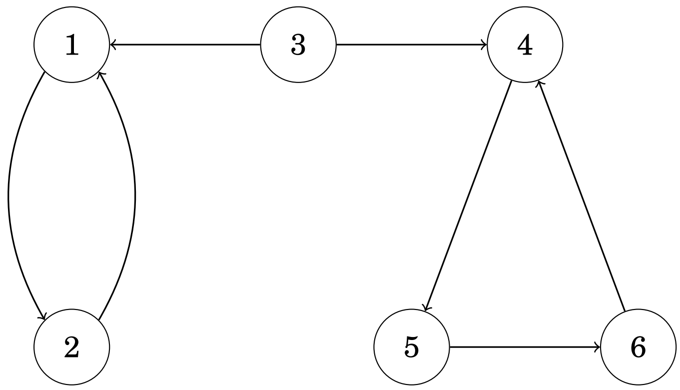

Abbildung 3.3.: Beispiel einer Markovkette mit verschiedenen Typen von Zuständen.

Definition 3.4.1 Erreichbarkeit, Kommunizierend
Sei $(X_n)_{n\geq 0}$ eine homogene Markovkette mit Übergangsmatrix $\mathbb{P}$ und Startverteilung α. Für $i, j \in S$ sagen wir,

i) dass j von i aus erreichbar ist und schreiben $i \to j$ genau dann, wenn ein $m \geq 0$ mit $p_{i,j}^{(m)} > 0$ existiert.

ii) dass i und j kommunizieren und schreiben $i \leftrightarrow j$ genau dann, wenn $i \to j$ und $j \to i$ gilt.

Wir erinnern uns daran, dass wir $p_{ij}^{(0)} = \delta_{ij}$ definiert haben und daher $i \leftrightarrow j$ für alle $i \in S$ gilt. Ferner ist $i \to j$ genau dann, wenn es einen Pfad von i nach j gibt. Dies ist im Wesentlichen die Aussage von Chapman-Kolmogorov (vergleiche Satz 3.1.4). Die Äquivalenzrelation „$\leftrightarrow$“ beschreibt die sog. *starken Zusammenhangskomponenten* des Graphen, wenn man die Kanten (i,i) hinzufügt. Die zugehörigen Äquivalenzklassen heißen Kommunikationsklassen. Im Beispiel aus Abbildung 3.3 wären dies $\{1,2\}$, $\{3\}$ und $\{4,5,6\}$.

Ist C eine Kommunikationsklasse und sind $i, j \in C$, so existiert ein Pfad in C, der von i nach j führt. Für den Beweis nehmen wir an, dass es $i, j \in C$ gibt, so dass jeder Pfad von i nach j durch ein $l \in S \setminus C$ verläuft. Solch einen Pfad mit $l \notin C$ wollen wir nun fixieren, dann gilt $i \to l$ und $l \to j \to i$. Damit gilt aber $i \leftrightarrow l$ und daher $l \in C$, was einen Widerspruch darstellt.

Unser nächstes Ziel ist die Beschreibung von Zustandsmengen, aus denen wir nicht wieder herauskommen können. Dazu führen wir den folgenden Begriff ein:

Definition 3.4.2 Abgeschlossene Menge
Es sei $(X_n)_{n\geq 0}$ eine homogene Markovkette mit Übergangsmatrix $\mathbb{P}$ und Startverteilung α. Eine Menge $A \subset S$ heißt abgeschlossen genau dann, wenn gilt

$$\sum_{j\in A} p_{ij} = 1 \qquad \text{für alle } i \in A.$$

Mit anderen Worten existiert also kein $j \notin A$ mit $p_{ij} > 0$. Im Beispiel aus Abbildung 3.3 wären $\{1,2\}$, $\{4,5,6\}$ oder auch $\{1,2,4,5,6\}$ und S selbst abgeschlossene Mengen.

Definition 3.4.3 Irreduzibilität
Es sei $(X_n)_{n\geq 0}$ eine homogene Markovkette. Eine Menge $A \subset S$ heißt irreduzibel genau dann, wenn $i \leftrightarrow j$ für alle $i, j \in A$ gilt. Ferner nennen wir die Markovkette (X_n) selbst irreduzibel genau dann, wenn S irreduzibel ist.

Aus der Irreduzibilität folgt im Allgemeinen nicht die Abgeschlossenheit. Dazu kann man Zustand 3 in Abbildung 3.3 betrachten. Ist S irreduzibel, so ist eine Kommunikationsklasse namensgebend. Die größte irreduzible Menge, die ein $i \in S$ enthält, ist dessen Kommunikationsklasse. Im Allgemeinen sind Kommunikationsklassen aber nicht abgeschlossen.

Wir sind nun an der Wahrscheinlichkeit dafür interessiert, dass man in endlicher Zeit von i nach j gelangen kann. Die Wahrscheinlichkeit, nach genau m Schritten von i nach j zu gelangen, haben wir mit $p_{ij}^{(m)}$ bezeichnet.

Für $j \in S$ definieren wir die Rückkehrzeit

$$\tau_j := \inf\{n \geq 1 : X_n = j\}.$$

Für diese Stoppzeit gilt $\tau_j \geq 1$, insbesondere gilt also $X_0(\omega) = j \not\Rightarrow \tau_j(\omega) = 0$. Für $i, j \in S$ setzen wir nun

$$\rho_{i,j} := P(\tau_j < \infty \mid X_0 = i).$$

Daher beschreibt $\rho_{i,j}$ gerade die Wahrscheinlichkeit dafür, in endlicher Zeit zu j zu gelangen, wenn man in i startet. Insbesondere ist $\rho_{i,i}$ gerade die Rückkehrwahrscheinlichkeit. Gelangen wir sicher in endlicher Zeit von i nach i, ist also $\rho_{i,i} = 1$, so liegt aufgrund der starken Markoveigenschaft die Vermutung nahe, dass i unendlich oft besucht wird. Wir werden sehen, dass i nur endlich oft besucht wird, wenn $\rho_{i,i} < 1$ ist.

Definition 3.4.4 Transienz, Rekurrenz
Es sei $(X_n)_{n\geq 0}$ eine homogene Markovkette. Ein Zustand $i \in S$ heißt dann

i) transient genau dann, wenn $\rho_{i,i} < 1$ ist.

ii) rekurrent genau dann, wenn $\rho_{i,i} = 1$ ist.

iii) positiv rekurrent genau dann, wenn der Zustand rekurrent ist und zusätzlich $\mathbf{E}(\tau_i \mid X_0 = i) < \infty$ gilt.

iv) nullrekurrent genau dann, wenn der Zustand rekurrent ist und $\mathbf{E}(\tau_i \mid X_0 = i) = \infty$ gilt.

Jeder Zustand $i \in S$ ist offenbar entweder transient oder rekurrent. Außerdem ist jeder rekurrente Zustand entweder positiv rekurrent oder nullrekurrent. Für $i \in S$ setzen wir nun

$$N_i := \sum_{n=0}^{\infty} \mathbf{1}_{\{X_n = i\}}$$

für die Anzahl der Besuche in i. Dann gilt $\mathbf{E}(N_i \mid X_0 = i) = \sum_{n=0}^{\infty} P(X_n = i \mid X_0 = i) = \sum_{n=0}^{\infty} p_{ii}^{(n)}$ für die erwarteten Besuche in i. Ferner betrachten wir die Wahrscheinlichkeit dafür, dass $X_n = i$ unendlich oft eintritt. Diese ist gegeben durch

$$P(N_i = \infty \mid X_0 = i) = P(\limsup\{X_n = i\} \mid X_0 = i).$$

Wir erwarten nun, dass diese Größen für transiente und (positive bzw. null-)rekurrente Zustände unterschiedlich sind.

Satz 3.4.5
Sei $(X_n)_{n \geq 0}$ eine homogene Markovkette und $i \in S$.

i) Ist i rekurrent, so gilt $P(N_i = \infty \mid X_0 = i) = 1$ und $\mathbf{E}(N_i \mid X_0 = i) = \sum_{n=0}^{\infty} p_{ii}^{(n)} = \infty$.

ii) Ist i transient, so gilt $P(N_i = \infty \mid X_0 = i) = 0$ und $\mathbf{E}(N_i \mid X_0 = i) = \sum_{n=0}^{\infty} p_{ii}^{(n)} = \frac{1}{1-\rho_{i,i}} < \infty$.

Beweis: Wir betrachten die letzte Besuchszeit, die gegeben ist durch

$$\begin{aligned} L_i &: \Omega \to \overline{\mathbf{N}_0}, \\ L_i(\omega) &:= \sup\{n \geq 0 : X_n(\omega) = i\}. \end{aligned}$$

Es gilt zu beachten, dass L_i keine Stoppzeit definiert, da die Zukunft $X_m(\omega) \neq i$ für $m > n$ beschrieben wird. Ist $X_0(\omega) = i$, so folgt $\{n \geq 0 : X_n(\omega) = i\} \neq \emptyset$ und $L_i(\omega) \geq 0$. Als Vorbemerkung führen wir folgende Rechnung an:

$$P(A \cap C \mid B) = \frac{P(A \cap C \cap B)}{P(B)} = \frac{P(A \cap C \cap B)}{P(B \cap C)} \cdot \frac{P(B \cap C)}{P(B)} = P(A \mid B \cap C) \cdot P(C \mid B).$$

Damit erhalten wir für $n \geq 0$

$$\begin{aligned} P(L_i = n \mid X_0 = i) &= P(X_n = i \text{ und } X_m \neq i\, \forall_{m>n} \mid X_0 = i) \\ &= P(X_m \neq i\, \forall_{m>n} \mid X_n = i \text{ und } X_0 = i) \cdot P(X_n = i \mid X_0 = i) \\ &\overset{3.3.3}{=} P(X_m \neq i\, \forall_{m>0} \mid X_0 = i) \cdot P(X_n = i \mid X_0 = i) \\ &= (1 - \rho_{i,i}) p_{ii}^{(n)}. \end{aligned} \tag{*}$$

Summieren wir über diesen Ausdruck, so erhalten wir

$$\begin{aligned} P(L_i < \infty \mid X_0 = i) &= (1-\rho_{i,i}) \sum_{n=0}^{\infty} p_{ii}^{(n)} \\ &= (1-\rho_{i,i})\mathbf{E}(N_i \mid X_0 = i). \end{aligned}$$

Ferner besucht (X_n) den Zustand i genau dann unendlich oft, wenn $L_i = \infty$ ist. Damit folgt

$$P(L_i < \infty \mid X_0 = i) = 1 - P(N_i = \infty \mid X_0 = i). \tag{**}$$

Setzen wir dies zusammen, so erhalten wir

$$(1-\rho_{i,i})\mathbf{E}(N_i \mid X_0 = i) = 1 - P(N_i = \infty \mid X_0 = i). \tag{***}$$

Ist i rekursiv, so gilt per Definition $\rho_{i,i} = 1$. Wegen (*) folgt damit $P(L_i = n \mid X_0 = i) = 0$ für alle $n \geq 0$ und damit $P(L_i < \infty \mid X_0 = i) = 0$. Aus (**) folgt dann, dass $P(N_i = \infty \mid X_0 = i) = 1$ gilt. Mit dem Lemma von Borel-Cantelli[2] folgt dann

$$\mathbf{E}(N_i \mid X_0 = i) = \sum_{n=1}^{\infty} P(X_n = i \mid X_0 = i) = \infty.$$

Ist i hingegen transient, so gilt $\rho_{i,i} < 1$ und mit (***) folgt daher

$$\mathbf{E}(N_i \mid X_0 = i) = \frac{1 - P(N_i = \infty \mid X_0 = i)}{1-\rho_{i,i}} < \infty.$$

Mit dem Lemma von Borel-Cantelli folgt dann wiederum $P(N_i = \infty \mid X_0 = i) = 0$ und mit (***) schließlich

$$\mathbf{E}(N_i \mid X_0 = i) = \frac{1}{1-\rho_{i,i}}.$$

□

Satz 3.4.6 **Abhängigkeit von Kommunikationsklassen**
Es sei $(X_n)_{n\geq 0}$ eine homogene Markovkette und $i, j \in S$ mit $i \leftrightarrow j$. Dann gilt:

i) i ist rekurrent genau dann, wenn j rekurrent ist.

ii) i ist transient genau dann, wenn j transient ist.

Beweis: Da jeder Zustand entweder transient oder rekurrent ist, müssen wir lediglich die erste Aussage beweisen. Da diese zudem symmetrisch in i und j ist, genügt es, nur eine Implikation zu beweisen. Sei daher i transient. Dann folgt wegen der kommunizierenden Eigenschaft, dass es $k, m \geq 0$ mit $p_{ij}^{(k)} > 0$ und $p_{ji}^{(m)} > 0$ gibt. Dann gilt mit dem Satz von Chapman-Kolmogorov aber $p_{ii}^{(k+l+m)} \geq p_{ij}^{(k)} p_{jj}^{(l)} p_{ji}^{(m)}$ für alle $l \geq 0$. Summieren wir über l, so erhalten wir

$$\sum_{l=0}^{\infty} p_{jj}^{(l)} \leq \frac{1}{p_{ij}^{(k)} p_{ji}^{(m)}} \sum_{l=0}^{\infty} p_{ii}^{(k+l+m)} < \infty,$$

da i nach Voraussetzung transient war. Damit ist j nicht rekurrent und daher transient. □

2 Dieses findet sich im Anhang unter Lemma A.5.

Satz 3.4.6 garantiert die Wohldefiniertheit der folgenden Sprechweise: Wir sagen, dass eine Kommunikationsklasse A rekurrent bzw. transient heißt, wenn es einen rekurrenten bzw. transienten Zustand $i \in A$ gibt, da dies dann auch für alle anderen Zustände in A gilt.

Satz 3.4.7 **Abgeschlossenheit von Kommunikationsklassen**
Es sei $(X_n)_{n\geq 0}$ eine homogene Markovkette und $R \subset S$ eine rekurrente Kommunikationsklasse. Dann ist R abgeschlossen im Sinne von Definition 3.4.2.

Beweis: Wir nehmen an, dass R nicht abgeschlossen ist. Dann existieren $i \in R$, $j \notin R$ und $m \geq 0$ mit $P(X_m = j \mid X_0 = i) > 0$. Da R eine Kommunikationsklasse ist, in der j nicht enthalten ist, gilt außerdem $P(X_n = i \text{ und } X_m = j \mid X_0 = i) = 0$ für alle $n > m$. Dann folgt auch $P(X_n = i \text{ unendlich oft und } X_m = j \mid X_0 = i) = 0$. Wir erhalten

$$\begin{aligned} P(X_n = i \ \infty\text{-oft} \mid X_0 = i) &\leq P(X_n = i \ \infty\text{-oft und } X_m = j \mid X_0 = i) + P(X_m \neq j \mid X_0 = i) \\ &\leq 0 + (1 - P(X_m = j)) \\ &< 1. \end{aligned}$$

Mit Satz 3.4.5 folgt dann, dass i nicht rekurrent ist. Nach Satz 3.4.6 stellt dies jedoch einen Widerspruch dar. □

Satz 3.4.8 **Zerlegung des Zustandsraumes**
Es sei $(X_n)_{n\geq 0}$ eine homogene Markovkette. Dann lässt sich der Zustandsraum zerlegen in

$$S = T \cup \bigcup_{l\in L} R_l,$$

wobei T die Menge der transienten Zustände, L eine höchstens abzählbare Indexmenge und R_l für jedes $l \in L$ eine irreduzibele, abgeschlossene und rekurrente Kommunikationsklasse ist.

Beweis: Seien $(C_k)_{k\in K}$ die Kommunikationsklassen, dann ist $S = \bigcup_{k\in K} C_k$ eine paarweise disjunkte Vereinigung, wobei K höchstens abzählbar ist, da auch S höchstens abzählbar ist. Wir setzen nun T als die Vereinigung über die transienten C_k. Dann gilt

$$S = T \quad \cup \quad \bigcup_{\substack{k\in K \\ C_k \text{ rekurrent}}} C_k.$$

Mit Satz 3.4.7 ist jede rekurrente Kommunikationsklasse aber auch abgeschlossen. Da Kommunikationsklassen zudem nach Definition irreduzibel sind, folgt die Behauptung. □

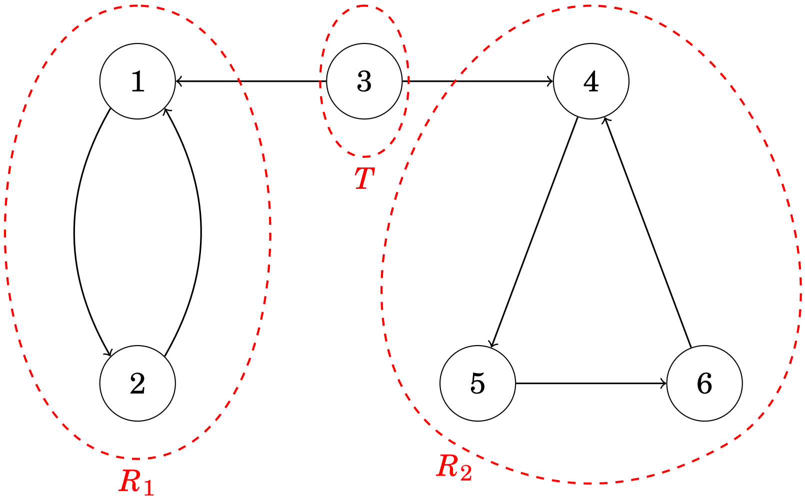

Abbildung 3.4.: Zerlegung des Beispiels aus Abbildung 3.3 gemäß Satz 3.4.8.

Eine Menge $A \subset S$ ist abgeschlossen, wenn kein Pfad A verlässt und irreduzibel, wenn es für jedes $i \in A$ einen Pfad zu jedem $j \in A$ gibt. Ist A endlich, abgeschlossen und irreduzibel, so würden wir erwarten, dass jeder Zustand häufig besucht wird.

Satz 3.4.9 Rekurrenz endlicher, irreduzibler Markovketten
Sei $(X_n)_{n\geq 0}$ eine homogene Markovkette und $A \subset S$ endlich, abgeschlossen und irreduzibel. Dann ist A positiv rekurrent.

Insbesondere sind also auch endliche, irreduzible und abgeschlossene Markovketten bereits positiv rekurrent.

Beweis: Es sei $i \in A$, dann gilt $i \leftrightarrow j$ für alle $j \in A$ und daher existiert für alle $j \in A$ ein k_j mit $P(\tau_i \leq k_j \mid X_0 = j) > 0$. Da A endlich ist, können wir das Maximum $k := \max\{k_j : j \in A\} < \infty$ wählen. Nun sei

$$\varepsilon := P(\tau_i \leq k \mid X_0 = j) > 0. \tag{*}$$

Wir wollen zeigen, dass $P(\tau_i > nk \mid X_0 = j) = (1-\varepsilon)^n$ ist und führen hierfür eine vollständige Induktion durch. Der Induktionsanfang ist bereits in (*) gegeben. Für den Induktionsschritt beachten wir nun, dass $\tau_i \leq k$ genau dann gilt, wenn $X_{\tau_i \wedge k} = i$ ist. Dann gilt

$$\begin{aligned}
P(\tau_i > (n+1)k \mid X_0 = j) &= P(X_{\tau_i \wedge (n+1)k} \neq i \mid X_0 = j) \\
&= \sum_{i \neq l \in A} P(X_{\tau_i \wedge (n+1)k} \neq i, X_{\tau_i \wedge nk} = l \mid X_0 = j) \\
&= \sum_{i \neq l \in A} P(X_{\tau_i \wedge (n+1)k} \neq i \mid X_{\tau_i \wedge nk} = l, X_0 = j) P(X_{\tau_i \wedge nk} = l \mid X_0 = j).
\end{aligned}$$

Wir wenden auf den ersten Faktor die starke Markoveigenschaft und (*) an und erhalten

$$
\begin{aligned}
&= (1-\varepsilon) \sum_{i \neq l \in A} P(X_{\tau_i \wedge nk} = l \mid X_0 = j) \\
&= (1-\varepsilon)(1-\varepsilon)^n \\
&= (1-\varepsilon)^{n+1}.
\end{aligned}
$$

Mit einem Satz aus der Wahrscheinlichkeitstheorie und dem Integral-Vergleichskriterium folgt dann

$$
\begin{aligned}
\mathbf{E}(\tau_i \mid X_0 = i) = \int_0^\infty P(\tau_i \geq t \mid X_0 = i)\, \mathrm{d}t &= \sum_{n=0}^\infty P(\tau_i > n \mid X_0 = i) \\
&\leq k \sum_{n=0}^\infty P(\tau_i > nk \mid X_0 = i) \\
&< \infty.
\end{aligned}
$$

□

Korollar 3.4.10
Sei $(X_n)_{n\geq 0}$ eine homogene Markovkette und $A \subset S$ eine endliche Kommunikationsklasse. Dann sind die folgenden Aussagen äquivalent:

i) A ist abgeschlossen.

ii) A ist rekurrent.

iii) A ist positiv rekurrent.

Beweis: Für i) ⇒ iii) folgt, nach Satz 3.4.9, dass A irreduzibel ist, da A nach Voraussetzung eine Kommunikationsklasse ist. Der Schritt iii) ⇒ ii) ist offensichtlich und ii) ⇒ i) folgt schließlich aus Satz 3.4.7. □

Wir können zusammenfassend also sagen, dass wir endliche Kommunikationsklassen bezüglich der Rekurrenz gut verstehen. Die nächste Frage wäre also, wie es mit unendlichen Kommunikationsklassen aussieht. Hier lassen sich solche Aussagen nicht treffen, wie wir im folgenden Beispiel sehen werden.

Beispiel 3.4.11 *Irrfahrt auf* $\mathbf{Z}$
Sei $(Y_n)_{n\geq 0}$ i. i. d. mit $P(Y_n = 1) =: p \in (0,1)$ und $P(Y_n = -1) = 1-p$. Dann setzen wir $X_n := \sum_{i=0}^n Y_n$ und in Beispiel 3.2.5 haben wir bereits gesehen, dass dies eine homogene Markovkette mit $p_{i,i+1} = p$ und $p_{i,i-1} = 1-p$ für $i \in \mathbf{Z}$ und $p_{i,j} = 0$ sonst ist. Dass (X_n) irreduzibel ist, ist klar, da $i \leftrightarrow i+1$ für alle Zustände $i \in \mathbf{Z}$ gilt. Außerdem gilt $X_n = 0$ P-fast sicher und es genügt, die Rekurrenz in $i = 0$ zu untersuchen. Man kann sich leicht überlegen, dass $p_{0,0}^{(2n+1)} = 0$ für alle $n \geq 0$ gilt.

Wie sieht es für eine gerade Anzahl von Schritten aus? Gelangen wir von 0 nach 0 in $2n$ Schritten, so müssen wir genau n Schritte in eine und n Schritte in die andere Richtung machen. Die Wahrscheinlichkeit von einer solchen Folge von Schritten beträgt $p^n(1-p)^n$ und es gibt $\binom{2n}{n}$ verschiedene solcher Folgen. Damit erhalten wir

$$p_{0,0}^{(2n)} = \binom{2n}{n} p^n (1-p)^n.$$

Mit Satz 3.4.5 reicht es, die Reihe $\sum_{n=0}^{\infty} p_{0,0}^{(2n)}$ zu untersuchen. Dazu verwenden wir die Stirling-Formel[3], die im Wesentlichen $n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$ besagt. Damit erhalten wir

$$\binom{2n}{n} = \frac{(2n)!}{(n!)^2} \sim \frac{4^n}{\sqrt{\pi n}}.$$

Im ersten Fall sei $p = 1-p$, dann ist $4^n p^n (1-p)^n = 1$ und wir erhalten

$$\sum_{n=0}^{\infty} p_{0,0}^{(2n)} \geq c \sum_{n=0}^{\infty} \frac{1}{\sqrt{\pi n}} = \infty,$$

also ist die Irrfahrt rekurrent. Im zweiten Fall sei $p \neq 1-p$, dann gilt $4^n(1-p)^n = a^n$ für ein $a \in (0,1)$. Dann ist die Reihe konvergent und die Irrfahrt daher transient. //

Man kann zeigen, dass die symmetrische Irrfahrt in $\mathbf{Z}^d$ genau dann rekurrent ist, wenn $d \leq 2$ ist. Anschaulich gesprochen bewegt man sich stets um die 0 herum. Für Details verweisen wir auf [Meintrup04, S. 248].

Satz 3.4.6 hatte für einen rekurrenten Zustand i mit $i \leftrightarrow j$ gezeigt, dass dann auch j rekurrent ist. Dies wollen wir nun verallgemeinern.

Satz 3.4.12 **Allgemeine Erreichbarkeitswahrscheinlichkeiten**
Sei $(X_n)_{n\geq 0}$ eine homogene Markovkette und $i \in S$ ein rekurrenter Zustand. Für $j \in S$ betrachten wir

$$\rho_{i,j} := P(\tau_j < \infty \mid X_0 = i).$$

Gilt $i \to j$, so ist auch j rekurrent und es gilt $\rho_{i,j} = \rho_{j,i} = 1$.

Beweis: Es sei also $i \to j$. Dann existieren paarweise verschiedene $i_0, \ldots, i_k$ mit $i_0 = i$, $i_k = j$ und $P(X_1 = i_1, \ldots, X_k = i_k \mid X_0 = i) > 0$, insbesondere gilt also $p_{i,j}^{(k)} > 0$. Nun folgt

$$\begin{aligned} 0 = 1 - \rho_{i,i} &= P(\tau_i = \infty \mid X_0 = i) \\ &\geq P(X_1 = i_1, \ldots, X_k = i_k, \tau_i = \infty \mid X_0 = i) \\ &= P(X_1 = i_1, \ldots, X_k = i_k \mid X_0 = i) P(\tau_i = \infty \mid X_l = i_l \text{ für } l = 0, \ldots, k) \end{aligned}$$

3 Siehe [WTSkript11, Satz 1.8.6].

und mit der Markoveigenschaft erhalten wir

$$\begin{aligned}
&= P(X_1 = i_1, \dots, X_k = i_k \mid X_0 = i) P(\tau_i = \infty \mid X_0 = i_k) \\
&= \underbrace{P(X_1 = i_1, \dots, X_k = i_k \mid X_0 = i)}_{>0}(1 - \rho_{j,i}).
\end{aligned}$$

Damit folgt $1 - \rho_{j,i} = 0$ und daher $\rho_{j,i} = 1$. Daher existiert ein $l \geq 1$ mit $p_{j,i}^{(l)} > 0$ und für $n \geq 0$ folgt mit der Chapman-Kolmogorov-Gleichung $p_{j,j}^{(l+n+k)} \geq p_{j,i}^{(l)} p_{i,i}^{(n)} p_{i,j}^{(k)}$. Dann gilt

$$\mathbf{E}(N_j \mid X_0 = j) \overset{3.4.5}{=} \sum_{n=0}^{\infty} p_{j,j}^{(n)} \geq \sum_{n=0}^{\infty} p_{j,i}^{(l)} p_{i,i}^{(n)} p_{i,j}^{(k)} = p_{j,i}^{(l)} p_{i,j}^{(k)} \sum_{n=0}^{\infty} p_{i,i}^{(n)} = \infty,$$

da i rekurrent ist. Wieder mit Satz 3.4.5 folgt also, dass auch j rekurrent ist. Vertauschen wir die Rollen von i und j, so folgt der Rest. $\square$

3.5. Stationarität

Wir wollen untersuchen, ob es zu einer gegebenen Übergangsmatrix $\mathbb{P}$ eine Startverteilung α gibt, so dass die homogene Markovkette stationär ist. Dazu beobachten wir, dass zu einem gegebenen $\mathbb{P}$ jede Startverteilung eine homogene Markovkette ergibt, deren Irreduzibilität und Rekurrenz von α unabhängig ist. Im Folgenden schreiben wir P_α statt P, wenn die Wahrscheinlichkeiten der Markovkette mit der Startverteilung α betrachtet werden.

Definition 3.5.1 **Stationarität**
Sei $\mathbb{P}$ eine stochastische Matrix über S. Eine Verteilung π auf S heißt stationär bezüglich $\mathbb{P}$ genau dann, wenn für alle $j \in S$ gilt:

$$\sum_{i \in S} \pi(i) p_{i,j} = \pi(j).$$

Anschaulich ausgedrückt ist die Wahrscheinlichkeit, im ersten Schritt nach j zu gelangen, wenn der Start i-verteilt ist, also $\pi(j)$. In der Matrixschreibweise gilt also $\pi\mathbb{P} = \pi$ und π ist ein Eigenvektor von $\mathbb{P}$ zum Eigenwert 1.

Satz 3.5.2 **Stationäre homogene Markovketten**
Sei $\mathbb{P}$ eine stochastische Matrix und π eine stationäre Verteilung von $\mathbb{P}$, sowie α irgendeine Startverteilung und $(X_n)_{n\geq 0}$ die homogene Markovkette bezüglich α und $\mathbb{P}$. Für $B \in \bigotimes_{\mathbf{N}} \mathscr{P}(S)$ gilt dann

$$P_\pi((X_n, X_{n+1}, \ldots) \in B) = P_\pi((X_0, X_1, \ldots) \in B).$$

Satz 3.5.2 besagt also, dass (π, P) eine stationäre, homogene Markovkette definiert.

Beweis: Wir wissen, dass $\pi\mathbb{P} = \pi$ gilt. Induktiv folgt dann auch $\pi\mathbb{P}^n = \pi$ für $n \geq 1$. Mit Lemma 3.2.3 folgt dann

$$P_\pi(X_n = j) = \sum_{i \in S} \pi(j) p_{i,j}^{(n)} = \pi(j).$$

Damit erhalten wir

$$\begin{aligned} P_\pi((X_n, X_{n+1}, \ldots) \in B) &= \sum_{i \in S} P_\pi(X_n = i) P_\pi((X_n, \ldots) \in B \mid X_n = i) \\ &= \sum_{i \in S} \pi(i) P_\pi((X_0, \ldots) \in B \mid X_0 = i) \\ &= P_\pi((X_0, \ldots) \in B). \end{aligned}$$

□

Wir wollen nun Bedingungen für die Existenz stationärer Verteilungen studieren.

Lemma 3.5.3 **Positivität stationärer Verteilungen**
Sei π eine stationäre Verteilung bezüglich $\mathbb{P}$ und $A \subset S$ abgeschlossen und irreduzibel. Gibt es ein $i_0 \in A$ mit $\pi(i_0) > 0$, so gilt $\pi(i) > 0$ für alle $i \in A$.

Ist $(X_n)_{n\geq 0}$ eine irreduzible, homogene Markovkette und π eine stationäre Verteilung dieser Markovkette, so folgt also insbesondere $\pi(i) > 0$ für alle $i \in S$.

Beweis: Wir setzen $S_+ := \{i \in A : \pi(i) > 0\}$ und $S_0 := \{i \in A : \pi(i) = 0\}$. Offenbar sind S_+ und S_0 disjunkt und es gilt $S_+ \neq \emptyset$ nach Voraussetzung. Zu zeigen ist nun $S_0 = \emptyset$. Wir nehmen an, dass $S_0 \neq \emptyset$ gilt, dann existiert also ein $j_0 \in S_0$ mit $\pi(j_0) = 0$. Für $j \in S_0$ folgt

$$\sum_{i \in S} \pi(i) p_{i,j} = \pi(j) = 0.$$

Da $\pi(i) > 0$ für alle $i \in S_+$ gilt, folgt $p_{i,j} = 0$ für alle $i \in S_+$ und $j \in S_0$. Es gibt also keine Verbindung von S_+ nach S_0, also gibt es keinen Pfad in A, der von i_0 nach j_0 führt. Dies steht jedoch im Widerspruch zur Irreduzibilität von A. □

Satz 3.5.4
Sei $(X_n)_{n\geq 0}$ eine homogene Markovkette mit stationärer Verteilung π, $A \subset S$ abgeschlossen und irreduzibel und es existiere ein $i_0 \in A$ mit $\pi(i_0) > 0$. Dann gilt

$$\mathbf{E}(\tau_i \mid X_0 = i) = \frac{P_\pi(\tau_i < \infty)}{\pi(i)}$$

für alle $i \in A$. Insbesondere sind alle $i \in A$ positiv rekurrent und ist $(X_n)_{n\geq 0}$ irreduzibel, so gilt für alle $i \in S$

$$\pi(i) = \frac{1}{\mathbf{E}(\tau_i \mid X_0 = i)},$$

das heißt in diesem Fall ist π eindeutig.

Beweis: Für $i \in S$ zeigen wir zunächst

$$\pi(i)\mathbf{E}(\tau_i \mid X_0 = i) = P_\pi(\tau_i < \infty). \tag{*}$$

Dazu betrachten wir die disjunkte Zerlegung

$$\{\tau_i \leq n\} = \bigcup_{l=1}^{n} \{X_l = i, X_{l+1} \neq i, \ldots, X_n \neq i\} = \bigcup_{l=0}^{n-1} \{X_{n-l} = i, X_{n-l+1} \neq i, \ldots, X_n \neq i\},$$

dann folgt mit Satz 3.5.2

$$\begin{aligned}
P_\pi(\tau_i<\infty) = \lim_{n\to\infty} P_\pi(\tau_i \le n) &= \lim_{n\to\infty} \sum_{l=0}^{n-1} P_\pi(X_{n-l}=i, X_{n-l+1}\neq i,\dots,X_n\neq i)\\
&= \lim_{n\to\infty} \sum_{l=0}^{n-1} P_\pi(X_0=i, X_1\neq i,\dots,X_l\neq i)\\
&= \sum_{l=0}^{\infty} P(X_0=i,\tau_i>l),
\end{aligned}$$

da $X_j \neq i$ genau dann gilt, wenn $\tau_i > l$ ist. Weiter folgt dann

$$\begin{aligned}
&= \sum_{l=0}^{\infty} P_\pi(\tau_i>l \mid X_0=i) P_\pi(X_0=i)\\
&= \sum_{l=0}^{\infty} P(\tau_i>l \mid X_0=i)\pi(i)\\
&= \pi(i)\mathbf{E}(\tau_i \mid X_0=i).
\end{aligned}$$

Dies beweist (*). Sei nun $i \in A$, dann folgt mit Lemma 3.5.3 wegen $\pi(i_0)>0$ auch $\pi(i)>0$. Wir können (*) daher durch $\pi(i)$ teilen und erhalten die gewünschte Gleichung.

Ferner gilt mit Satz 3.4.12 auch

$$\begin{aligned}
P_\pi(\tau_i<\infty) &= \sum_{j\in S} P(\tau_i<\infty \mid X_0=j) P_\pi(X_0=j) = \sum_{j\in S} 1\cdot\pi(j)\\
&= 1.
\end{aligned}$$

□

Satz 3.5.5
Ist $(X_n)_{n\geq 0}$ eine homogene Markovkette und $i \in S$ ein positiv rekurrenter Zustand, so existiert eine stationäre Verteilung.

Beweis: Für $j \in S$ sei

$$c(j) := \sum_{n=0}^{\infty} P(X_n=j,\tau_i>n \mid X_0=i).$$

Dann wollen wir zeigen, dass $\pi(j) := \frac{c_j}{\mathbf{E}(\tau_i|X_0=i)}$ eine stationäre Verteilung definiert. Wir zeigen zunächst, dass π eine Verteilung ist. Dafür betrachten wir

$$\sum_{j\in S} c(j) = \sum_{n=0}^{\infty} P(\tau_i>n \mid X_0=i) = \mathbf{E}(\tau_i \mid X_0=i) < \infty.$$

Für die Stationarität sei $c_n(j) := P(X_n = j, \tau_i > n \mid X_0 = i)$ für $n \geq 0$ und $j \in S$. Dann wollen wir zeigen, dass für alle $k \in S$

$$\sum_{j \in S} c(j) p_{jk} = \sum_{n=0}^{\infty} \sum_{j \in S} c_n(j) p_{jk} \stackrel{!}{=} c(k)$$

gilt. Im ersten Fall sei $i \neq k$. Da $\{\tau_i > n\} \in \sigma(X_0, \ldots, X_n)$ gilt, folgt

$$\begin{aligned}
\sum_{n=0}^{\infty} \sum_{j \in S} c_n(j) p_{jk} &= \sum_{n=0}^{\infty} \sum_{j \in S} P(X_n = j, \tau_i > n \mid X_0 = i) P(X_{n+1} = k \mid X_n = j) \\
&= \sum_{n=0}^{\infty} \sum_{j \in S} P(X_n = j, \tau_i > n \mid X_0 = i) P(X_{n+1} = k \mid X_n = j, \tau_i > n).
\end{aligned}$$

Mit der uns bereits bekannten Formel $P(A \cap C \mid B) = P(A \mid C \cap B) P(C \mid B)$ gilt dann

$$\begin{aligned}
&= \sum_{n=0}^{\infty} \sum_{j \in S} P(X_n = j, X_{n+1} = k, \tau_i > n \mid X_0 = i) \\
&= \sum_{n=0}^{\infty} P(X_{n+1} = k, \tau_i > n \mid X_0 = i).
\end{aligned}$$

Wegen $X_{n+1} = k \neq i$ folgt $\tau_i \neq n+1$ und daher

$$\begin{aligned}
&= \sum_{n=0}^{\infty} P(X_{n+1} = k, \tau_i > n+1 \mid X_0 = i) \\
&= \sum_{n=0}^{\infty} c_{n+1}(k).
\end{aligned}$$

Wegen $c_0(k) = P(X_0 = k, \tau_i > n \mid X_0 = i) = 0$ fällt der erste Summand weg und wir erhalten

$$\begin{aligned}
&= \sum_{n=0}^{\infty} c_n(k) \\
&= c(k).
\end{aligned}$$

Im zweiten Fall sei $i = k$, dann erhalten wir zunächst wie eben

$$\sum_{n=0}^{\infty} \sum_{j \in S} c_n(j) p_{ij} = \sum_{n=0}^{\infty} P(X_{n+1} = i, \tau_i > n \mid X_0 = i)$$

und nun mit $\tau_i > 0$ und $\tau_i < \infty$ P-fast sicher

$$\begin{aligned}
&= \sum_{n=0}^{\infty} P(\tau_i = n+1 \mid X_0 = i) \\
&= 1 \\
&= P(X_0 = i, \tau_i > 0 \mid X_0 = i) \\
&= \sum_{n=0}^{\infty} P(X_n = i, \tau_i > n \mid X_0 = i).
\end{aligned}$$

Das Ereignis in der Summe ist für $n > 0$ aber nie erfüllt und wir erhalten schließlich

$$= c(i). \qquad \square$$

Korollar 3.5.6
Es sei $(X_n)_{n\geq 0}$ eine irreduzible, homogene Markovkette. Dann sind die folgenden Aussagen äquivalent:

i) Es existiert ein positiv rekurrenter Zustand $i \in S$.

ii) Alle Zustände $i \in S$ sind positiv rekurrent.

iii) Es existiert eine stationäre Verteilung.

iv) Es existiert genau eine stationäre Verteilung.

Ist eine dieser Aussagen – und damit alle – erfüllt, so gilt für $i \in S$

$$\pi(i) = \frac{1}{\mathbf{E}(\tau_i \mid X_0 = i)}.$$

Beweis: Für i) ⇒ iii) verwenden wir Satz 3.5.5, für iii) ⇒ iv) und iii) ⇒ ii) verwenden wir Satz 3.5.4 und ii) ⇒ i) und iv) ⇒ iii) sind trivial. □

3.6. Grenzverhalten

Wenn wir wissen, dass $i \in S$ transient ist, so folgt mit Satz 3.4.5, dass $p_{ii}^{(n)} \to 0$ gilt. Was gilt jedoch im Allgemeinen? Dazu betrachten wir

$$\mathbb{P} := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

dann gilt $\mathbb{P}^2 = E_2$, wobei E_n die Einheitsmatrix ist. Wir erhalten also

$$\mathbb{P}^n = \begin{cases} \mathbb{P} & \text{falls } n \text{ ungerade} \\ E_2 & \text{sonst} \end{cases}.$$

Dann konvergiert $p_{ii}^{(n)}$ jedoch nicht. Ein solches periodisches Verhalten wollen wir also ausschließen.

Definition 3.6.1 Periodizität
Sei $(X_n)_{n\geq 0}$ eine homogene Markovkette und $i \in S$.

i) Die Periode von i ist der größte gemeinsame Teiler von

$$J_i := \left\{ n \geq 1 : p_{ii}^{(n)} > 0 \right\}.$$

ii) Der Zustand i heißt aperiodisch, falls er die Periode 1 hat.

Betrachten wir nochmal das obige Beispiel. Da $J_1 = J_2 = 2\mathbf{N}$ gilt, haben $i = 1$ und $i = 2$ jeweils die Periode 2.

Beachte, dass $p_{ii}^{(n)} > 0$ genau dann gilt, wenn man in *genau* n Schritten von i nach i gelangen kann. In Abbildung 3.5 sehen wir einen gängigen Denkfehler für die Periode eines Zustandes.

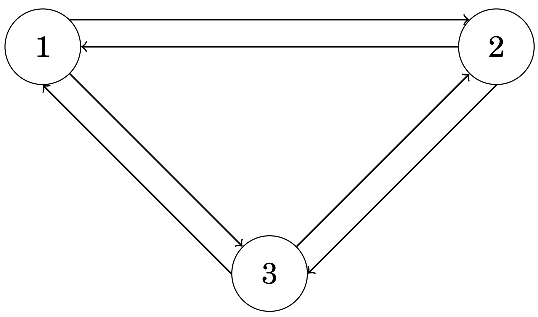

Abbildung 3.5.: Von jedem Zustand i gelangt man in sowohl zwei als auch drei Schritten von i nach i, die Periode ist daher 1.

Lemma 3.6.2
Sei $A \subset \mathbf{N}$ mit $\mathrm{ggT}(A) = 1$ und A sei abgeschlossen bezüglich der Addition. Dann existiert ein n_0, so dass $n \in A$ für alle $n \geq n_0$ gilt.

Beweis: Da die Aussage rein zahlentheoretischer Natur ist, wollen wir sie hier nicht beweisen und verweisen stattdessen auf [Meintrup04, Lemma 9.41]. □

Lemma 3.6.3
Sei $(X_n)_{n\geq 0}$ eine homogene Markovkette und $i \in S$. Dann sind folgende Aussagen äquivalent:

i) Der Zustand i ist aperiodisch.

ii) Es existiert ein n_0, so dass $p_{ii}^{(n)} > 0$ für alle $n \geq n_0$ gilt.

Beweis: Für i) ⇒ ii) setzen wir $A := \{n \geq 0 : p_{ii}^{(n)} > 0\}$. Da i aperiodisch ist, gilt $\mathrm{ggT}(A) = 1$. Für $n, m \in A$ gilt ferner $p_{ii}^{(n+m)} \geq p_{ii}^{(n)} p_{ii}^{(m)} > 0$ und daher $n + m \in A$. Mit Lemma 3.6.2 folgt dann die Aussage.

Für ii) ⇒ i) sei $d = \mathrm{ggT}(A)$. Nach Voraussetzung existiert ein n_0 mit der entsprechenden Eigenschaft. Also gilt $d \mid n_0$ und $d \mid n_0 + 1$. Dann teilt d aber auch $(n_0 + 1) - n_0 = 1$ und wir erhalten $d = 1$. □

Satz 3.6.4
Ist $(X_n)_{n\geq 0}$ eine homogene Markovkette, $i \in S$ aperiodisch und $A \subset S$ die Kommunikationsklasse von i, so existiert ein n_0, so dass $p_{jk}^{(n)} > 0$ für alle $j, k \in A$ und alle $n \geq n_0$ gilt.

Setzen wir $j = k$, so folgt insbesondere, dass j aperiodisch ist.

Beweis: Es seien $j, k \in A$. Dann existieren $l, m \geq 0$ mit $p_{ji}^{(l)} > 0$ und $p_{ij}^{(k)} > 0$. Mit Lemma 3.6.3 existiert dann ein n_0, so dass $p_{ii}^{(n)} > 0$ für alle $n \geq n_0$ gilt. Damit erhalten wir

$$p_{jk}^{(l+n+m)} \geq p_{ji}^{(l)} p_{ii}^{(n)} p_{ik}^{(m)} > 0.$$ □

Definition 3.6.5 **Kopplung**
Es seien $(X_n)_{n\geq 0}$ und $(Y_n)_{n\geq 0}$ stochastische Prozesse über $(\Omega, \mathscr{A}, P)$ mit Zustandsraum $\mathscr{X}$.

Dann sind (X_n) und (Y_n) ein Kopplungspaar, falls eine P-fast sicher endliche Stoppzeit τ existiert, so dass für alle $n \geq 0$ und $\omega \in \Omega$ gilt:

$$n \geq \tau(\omega) \Longrightarrow X_n(\omega) = Y_n(\omega).$$

Wir nennen τ in diesem Fall Kopplungszeit.

Aneinander gekoppelte stochastische Prozesse sind nach der Zeit τ also gleich.

Satz 3.6.6 **Konvergenz von Kopplungen**
Es sei τ die Kopplungszeit zum Kopplungspaar (X_n) und (Y_n). Für eine messbare Menge $A \subset \mathscr{X}$ gilt dann

$$\lim_{n\to\infty} (P(X_n \in A) - P(Y_n \in A)) = 0.$$

Beachte: Der Grenzwert $\lim P(X_n \in A)$ selbst existiert im Allgemeinen nicht.

Beweis: Wir betrachten

$$\begin{aligned}
|P(X_n \in A) - P(Y_n \in A)| &\leq |P(X_n \in A, \tau \leq n) - P(Y_n \in A, \tau \leq n)| \\
&\qquad + |P(X_n \in A, \tau > n) - P(Y_n \in A, \tau > n)| \\
&\leq 2P(\tau > n) \\
&\to 0.
\end{aligned}$$

□

Konstruktion von Produkt-Markovketten

Es seien $(X_n)_{n\geq 0}$ und $(Y_n)_{n\geq 0}$ unabhängige, homogene Markovketten bezüglich $(\alpha, \mathbb{P})$ und $(\beta, \mathbb{P})$, sie unterscheiden sich also nur in ihrer Startverteilung. Es sei $Z_n := (X_n, Y_n) \in S \times S$, dann kann man leicht zeigen, dass $(Z_n)_{n\geq 0}$ eine homogene Markovkete auf $S \times S$ mit Startverteilung $\alpha \otimes \beta$ und der Übergangsmatrix $\overline{\mathbb{P}}$ vermöge

$$\overline{p}_{(i_0,j_0),(i_1,j_1)} = p_{i_0,i_1} p_{j_0,j_1}$$

bildet.

Satz 3.6.7
Es seien $(X_n)_{n\geq 0}$, $(Y_n)_{n\geq 0}$ und $(Z_n)_{n\geq 0}$ wie eben definiert. Ist $(Z_n)_{n\geq 0}$ irreduzibel und rekurrent, so gelten die folgenden Aussagen:

i) Die Stoppzeit $\tau_{i_0} := \inf\{n \geq 1 : X_n = Y_n = i_0\}$ für $i_0 \in S$ ist P-fast sicher endlich.

ii) Es sei

$$W_n := \begin{cases} X_n & \text{falls } n \le \tau_{i_0} \\ Y_n & \text{falls } n > \tau_{i_0} \end{cases}.$$

Dann ist W_n eine homogene Markovkette mit Startverteilung α bezüglich $\mathbb{P}$.

iii) Für alle $A \subset S$ gilt

$$P(X_n \in A) - P(Y_n \in A) \to 0.$$

Beweis: Für den Beweis von i) betrachten wir die Rückkehrzeit $\tau = \tau_{i_0}$ von (Z_n) zum Zustand (i_0, i_0). Da (Z_n) rekurrent ist, gilt P-fast sicher $\tau < \infty$.

Die Aussage ii) werden wir hier nicht beweisen. Wir verweisen stattdessen auf [Meintrup04, S. 258].

Für iii) beachten wir, dass (Y_n) und (W_n) für τ_{i_0} ein Kopplungspaar sind. Nach ii) sind (W_n) und (X_n) identisch verteilt, da beide Markovketten bezüglich $(\alpha, \mathbb{P})$ sind. Dann wenden wir Satz 3.6.6 an. □

Satz 3.6.8 **Konvergenz homogener Markovketten**
Es sei $(X_n)_{n\ge 0}$ eine irreduzible, homogene Markovkette mit stationärer Verteilung π. Dann gilt

$$P(X_n = j) \overset{n\to\infty}{\longrightarrow} \pi(j)$$

für alle $j \in S$.

Ist $\alpha = \delta_{\{i\}}$, so gilt insbesondere $p_{ij}^{(n)} \to \pi(j)$ für alle $i, j \in S$.

Beweis: Es sei $(\alpha, \mathbb{P})$ die Konfiguration von $(X_n)_{n\ge0}$ und $(Y_n)_{n\ge0}$ eine von $(X_n)_{n\ge0}$ unabhängige, homogene Markovkette mit der Konfiguration $(\pi, \mathbb{P})$. Wir wollen zeigen, dass $Z_n := (X_n, Y_n)$ irreduzibel und rekurrent ist, um Satz 3.6.7 anzuwenden. Für die Irreduzibilität seien $(i_0, j_0), (i_1, j_1) \in S \times S$, dann folgt mit Lemma 3.6.3, dass ein $r \ge 0$ mit $p_{i_1,i_1}^{(r)} > 0$ und $p_{j_1,j_1}^{(r)} > 0$ existiert. Mit Lemma 3.6.3 folgt dann $p_{i_0,i_1}^{(n)} p_{j_0,j_1}^{(n)} > 0$ für alle $n \ge 0$. Damit erhalten wir

$$\overline{p}_{(i_0,j_0),(i_1,j_1)}^{(n+r)} = p_{i_0,i_1}^{(r)} p_{j_0,j_1}^{(r)} > 0,$$

also ist Z_n irreduzibel. Für die Rekurrenz kann man leicht zeigen, dass $\pi \otimes \pi$ eine stationäre Verteilung von (X_n, Y_n) ist. Nach Korollar 3.5.6 ist (Z_n) also rekurrent. Mit Satz 3.6.7 erhalten wir

$$P(X_n = j) - \pi(j) = P(X_n = j) - P(Y_n = j) \longrightarrow 0.$$ □

Für endliche Zustandsräume S lässt sich der Beweis auch elementarer führen. Hierfür verweisen wir auf die Literatur.

Es sei $(X_n)_{n\geq 0}$ irreduzibel und aperiodisch. Dann wollen wir folgende Eigenschaften festhalten:

i) Ist (X_n) zusätzlich transient, so gilt

$$\sum_n p_{ij}^{(n)} \leq \frac{1}{p_{ji}^{(r)}} \sum_n p_{ij}^{(n)} p_{ji}^{(r)} \leq \frac{1}{p_{ji}^{(r)}} \sum_n p_{ii}^{(n+r)} < \infty.$$

Also gilt $p_{ij}^{(n)} \to 0$.

ii) Ist $(X_n)_{n\geq 0}$ positiv rekurrent, so existiert nach Korollar 3.5.6 eine stationäre Verteilung π mit $\pi(j) > 0$ für alle $j \in S$. Nach Satz 3.6.8 gilt dann $p_{ij}^{(n)} \to \pi(j)$ und $\sum_n p_{ij}^{(n)} = \infty$.

iii) Den dritten Fall halten wir in Satz 3.6.9 fest.

Satz 3.6.9
Sei $(X_n)_{n\geq 0}$ eine irreduzible, aperiodische und rekurrente, aber nicht positiv rekurrente, homogene Markovkette. Für alle $i, j \in S$ gilt dann

$$p_{ij}^{(n)} \longrightarrow 0$$

und

$$\sum_n p_{ij}^{(n)} = \infty.$$

Beweis: Der Beweis findet sich in [Meintrup04, Satz 9.48] und wird hier nicht geführt. □

4
Poissonprozesse

Poissonprozesse sind spezielle Zählprozesse mit poissonverteilten Zuwächsen, die zum Beispiel das wiederholte Eintreten eines Ereignisses modellieren.

4.1. Definition und Eigenschaften

Im Folgenden betrachten wir einen $\mathbf{N}_0$-wertigen stochastischen Prozess $(X_t)_{t\geq 0}$ mit kontinuierlicher Zeit. Satz 1.2.7 besagte, dass X genau dann rechtsstetig ist, wenn für alle $\omega \in \Omega$ und alle $t > 0$ ein $\varepsilon > 0$ mit $X_s(\omega) = X_t(\omega)$ für alle $s \in [t, t+\varepsilon)$ existiert. Der Prozess ist rechts von Stetigkeitsstellen also in einer gewissen Umgebung konstant. Insbesondere gibt es für jede Trajektorie genau drei Möglichkeiten:

i) Die Trajektorie besitzt endlich viele Sprünge.

ii) Die Trajektorie hat unendlich viele Sprünge, aber nur endlich viele in jedem beschränkten Intervall.

iii) Es gibt ein Zeitintervall, in welchem unendlich viele Sprünge auftreten. Man spricht in diesem Fall von *Explosionen*.

Wir werden sehen, dass bei Poissonprozessen P-fast sicher nur der zweite Fall eintritt. Daher wollen wir diese Eigenschaft näher untersuchen. In der Abbildung 4.1 sehen wir einen solchen Prozess und die Bezeichnungen für die entsprechenden Zeiten, die wir in diesem Kapitel verwenden wollen. Wir sehen bereits, dass eine solche Nummerierung der T_i im dritten Fall nicht ohne Weiteres möglich wäre und dass wir im ersten Fall nur endlich viele T_i haben.

Im zweiten Fall lässt sich $(X_t)_{t\geq 0}$ vollständig durch die Angabe der Werte $(S_n)_{n\geq 1}$ und der Sprungzeiten $(T_n)_{n\geq 0}$ bzw. durch Angabe der Werte $(S_n)_{n\geq 0}$ und der Wartezeiten $(W_n)_{n\geq 1}$ beschreiben. Die Rechtsstetigkeit und der diskrete Zustandsraum erlaubt es uns also, einen zeitkontinuierlichen Prozess durch zwei zeitdiskrete Prozesse zu beschreiben. Dies wollen wir nun formalisieren.

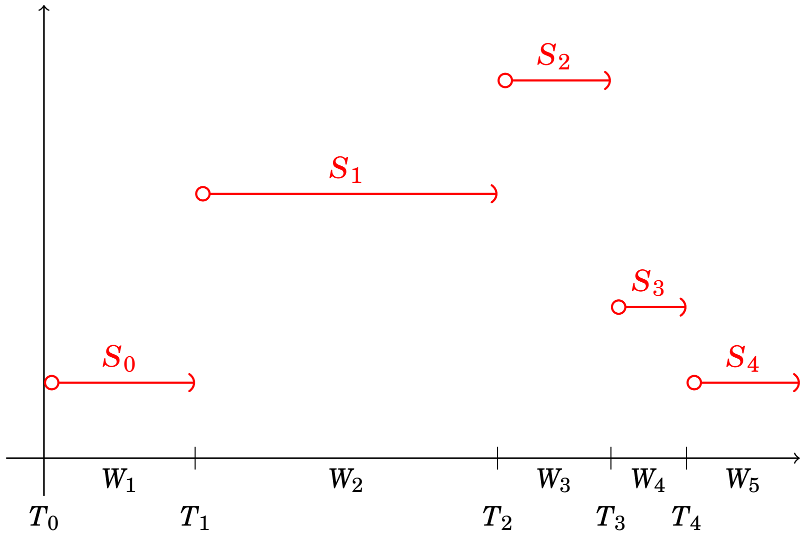

Abbildung 4.1.: Die Abbildung zeigt eine typische Trajektorie eines rechtsstetigen, zeitkontinuierlichen stochastischen Prozesses, sowie die Bezeichnungen für dieses Kapitel.

Definition 4.1.1 Sprungprozesse
Sei $(X_t)_{t\geq 0}$ ein zeitkontinuierliches, rechtsstetiger und $\mathbf{N}_0$-wertiger stochastischer Prozess.

i) Die Sprungzeiten $(T_n)_{n\geq 0}$ sind definiert durch $T_0 := 0$ und $T_{n+1} := \inf\{t \geq T_n : X_t \neq X_{T_n}\}$ für $n > 0$ mit $\inf\emptyset := \infty$.

ii) Der Prozess (X_t) heißt explosionsfrei, wenn P-fast sicher $\lim_{n\to\infty} T_n = \infty$ gilt.

iii) Die Wartezeiten $(W_n)_{n\geq 1}$ sind definiert durch

$$W_n := \begin{cases} T_n - T_{n-1} & \text{falls } T_{n-1} < \infty \\ \infty & \text{sonst} \end{cases}.$$

iv) Der Sprungprozess $(S_n)_{n\geq 0}$ ist definiert durch

$$S_n := \begin{cases} X_{T_n} & \text{falls } T_n < \infty \\ X_{T_m} & \text{sonst} \end{cases},$$

wobei $m := \max\{r \in \mathbf{N}_0 : T_r < \infty\}$ die letzte endliche Sprungzeit beschreibt.

Besitzt ein Pfad nur endlich viele Sprünge, so sind sowohl Warte- als auch Sprungzeit nach dem letzten Sprung unendlich. Ist $T_{n-1} < \infty$, so besagt Satz 1.2.7, dass $T_{n-1} < T_n$ gilt. Der Satz besagt außerdem $W_n > 0$.

Ist (X_t) explosionsfrei, so bestimmen (S_n) und (W_n) bzw. (S_n) und (T_n) den Prozess $(X_t)_{t\geq 0}$ durch $X_n = S_n$ für $T_n \leq t < T_{n+1}$.

Satz 4.1.2
Sei $(W_n)_{n\geq 1}$ i. i. d. mit $W_i \in \mathscr{L}_1$ und $P(W_n \in (0,\infty)) = 1$. Für $n \geq 0$ und $t \geq 0$ setzen wir $T_n := \sum_{i=1}^n W_i$ und $X_t := \sum_{i=1}^\infty \mathbf{1}_{(0,t]}(T_i)$ als die Anzahl der Sprünge bis zur Zeit t. Dann ist $(X_t)_{t\geq 0}$ ein $\mathbf{N}_0$-wertiger, rechtsstetiger, monoton wachsender und explosionsfreier stochastischer Prozess, dessen Wartezeiten die $(W_n)_{n\geq 1}$ sind und dessen Sprungprozess (S_n) durch $S_n = n$ für $n \geq 0$ gegeben ist. Für alle $n \geq 0$ gilt ferner P-fast sicher $T_n < \infty$ und $T_n < T_{n+1}$.

Beweis: Es gilt $P(T_{n+1} - T_n > 0) = P(W_{n+1} > 0) = 1$. Für $n \geq 0$ und $t \geq 0$ folgt damit $\{X_t = n\} = \{T_n \leq t < T_{n+1}\}$. Wir wollen zeigen, dass P-fast sicher $X_t \in \mathbf{N}_0$ für alle $t \geq 0$ gilt. Dazu betrachten wir

$$\{X_t = \infty\} = \{\lim T_n \leq t\} = \left\{\sum_{i=1}^\infty W_i \leq t\right\} = \left\{\sum_{i=1}^n W_i \leq t \quad \forall_{n\geq 0}\right\} = \left\{0 \leq \frac{1}{n}\sum_{i=1}^n W_i \leq \frac{t}{n} \quad \forall_{n\geq 0}\right\}$$
$$\subset \left\{\lim \frac{1}{n}\sum_{i=1}^n W_i = 0\right\}.$$

Das starke Gesetz der großen Zahlen impliziert nun P-fast sicher $\frac{1}{n}\sum_{i=1}^n W_i \to \mathbf{E}W_1 > 0$ und wir erhalten $P(X_t = \infty) = 0$.

Wir wollen nun zeigen, dass (X_t) explosionsfrei ist. Da (X_t) offenbar monoton wachsend und $\mathbf{N}_0$-wertig ist, ist keine Explosion möglich, denn wäre bei t_0 eine Explosion, so hätten wir $\lim_{t\nearrow t_0} X_t = \infty$ und damit gäbe es $t < t_0$ mit $X_t > X_{t_0} \in \mathbf{N}_0$, was im Widerspruch zur Monotonie steht.

Wir zeigen nun, dass (X_t) rechtsstetig ist. Dazu sei $t \geq 0$, $\omega \in \Omega$ und $n := X_t(\omega)$. Wegen $T_n < T_{n+1}$ existiert dann $\varepsilon > 0$ mit $T_n(\omega) \leq t + \varepsilon < T_{n+1}(\omega)$ P-fast sicher. Dann gilt $X_s(\omega) = n$ für alle $s \in [t, t+\varepsilon)$ und mit Satz 1.2.7 folgt dann die Rechtsstetigkeit.

Ferner gilt wegen der Monotonie der T_n

$$X_{T_n} = \sum_{i=0}^\infty \mathbf{1}_{(0,T_n]}(T_i) = \sum_{i=1}^n \mathbf{1}_{(0,T_n]}(T_i) = n. \tag{*}$$

Nun wollen wir zeigen, dass die T_n die Sprungzeiten von (X_t) sind. Dazu seien $(\tilde{T}_n)$ die Sprungzeiten von (X_t) und wir zeigen mit vollständiger Induktion, dass $T_n = \tilde{T}_n$ gilt. Für den Induktionsanfang betrachte einfach $T_0 = 0 = \tilde{T}_0$. Für den Induktionsschritt gilt

$$\tilde{T}_{n+1} = \inf\{t \geq \tilde{T}_n : X_t \neq X_{\tilde{T}_n}\} = \inf\{t \geq T_n : X_t \neq X_{T_n}\} \overset{(*)}{=} \inf\left\{t \geq T_n : \sum_{i=0}^\infty \mathbf{1}_{(0,t]}(T_i) \neq n\right\}.$$

Wir können nun die Summe in $\sum_{i=0}^{n} \mathbf{1}_{(0,t]}(T_i) + \sum_{i=n+1}^{\infty} \mathbf{1}_{(0,t]}(T_i)$ aufspalten, wobei die erste dieser Teilsummen gerade n ergibt. Wir erhalten daher aus der Nicht-Negativität

$$
\begin{aligned}
&= \inf\left\{ t \geq T_n : \sum_{i=n+1}^{\infty} \mathbf{1}_{(0,t]}(T_i) > 0 \right\} = \inf\{t \geq T_n : t \geq T_{n+1}\} \\
&= T_{n+1}.
\end{aligned}
$$

Wir zeigen nun noch, dass P-fast sicher $T_n < \infty$ gilt. Es ist $T_n = \sum_{i=1}^{n} W_i$ und wegen $P(W_i < \infty) = 1$ ist dann auch $P(T_n < \infty) = 1$. Damit sind die W_n die Wartezeiten von (X_t). Schließlich gilt mit (**) noch $S_n = X_{T_n} = n$. □

Sind (W_n), (T_n) und (X_t) wie im Satz 4.1.2 gegeben, so heißt $(X_t)_{t\geq 0}$ Zählprozess, da dieser Prozess zählt, wie viele der konsekutiven Ereignisse, die zu den Zeiten T_i eintreten, bereits zur Zeit t eingetreten sind. Satz 4.1.2 ermöglicht recht allgemein die Konstruktion solcher Zählprozesse.

Wir werden als nächstes Poissonprozesse einführen, die Zählprozesse mit exponentialverteilten Wartezeiten sind. Daher wollen wir zunächst die Exponentialverteilung $\operatorname{Exp}(\lambda)$ für $\lambda > 0$ in Erinnerung rufen. Dies ist eine Lebesgue-absolut stetige Verteilung auf $\mathbf{R}$, welche durch die Dichte

$$
h(t) := \begin{cases} \lambda \exp(-\lambda t) & \text{falls } t \geq 0 \\ 0 & \text{sonst} \end{cases}
$$

gegeben ist. Ist $Y \sim \operatorname{Exp}(\lambda)$, so gelten die folgenden Eigenschaften:

i) Es ist $P(Y \in (0,\infty)) = 1$.

ii) Es gilt $\mathbf{E}Y = \lambda^{-1}$ und $\operatorname{Var} Y = \lambda^{-2}$.

iii) Die Verteilungsfunktion von Y ist gegeben durch

$$
F_Y(t) := \begin{cases} 1 - \exp(-\lambda t) & \text{falls } t \geq 0 \\ 0 & \text{sonst} \end{cases}.
$$

iv) Die Zufallsvariable Y ist gedächtnislos, für alle $t, s \geq 0$ gilt also

$$
P(Y > t + s \mid Y > s) = P(Y > t).
$$

v) Sind $Y_1, \ldots, Y_n \sim \operatorname{Exp}(\lambda)$ unabhängig, so gilt $\sum_{i=1}^{n} Y_i \sim \Gamma(n, \lambda)$, wobei dies die Gammaverteilung ist, welche durch die Lebesgue-Dichte

$$
h_n(t) := \begin{cases} \lambda \frac{(\lambda t)^{n-1}}{(n-1)!} \exp(-\lambda t) & \text{falls } t > 0 \\ 0 & \text{sonst} \end{cases}
$$

gegeben ist.

Mit diesen elementaren Eigenschaften können wir nun zur Definition des Poissonprozesses, der ein Spezialfall von Satz 4.1.2 darstellt. Dieser sichert die Existenz, die Explosionsfreiheit und das Vorhandensein unendlich vieler Sprünge, insbesondere also auch P-fast sicher $T_n < \infty$ und $W_n < \infty$.

Definition 4.1.3 **Poissonprozess**
Ein rechtsstetiger und $\mathbf{N}_0$-wertiger stochastischer Prozess $(N_t)_{t\geq 0}$ heißt homogener Poissonprozess mit Rate $\lambda > 0$, wenn die folgenden Eigenschaften erfüllt sind:

i) Es gilt P-fast sicher $N_0 = 0$.

ii) Die Wartezeiten $(W_n)_{n\geq 1}$ sind i. i. d. mit $W_n \sim \mathrm{Exp}(\lambda)$.

iii) Der Sprungprozess (S_n) erfüllt $S_n = n$ für alle $n \geq 0$.

Poissonprozesse sind durch den Parameter λ bereits eindeutig beschrieben, da Wartezeiten und Sprungprozess den stochastischen Prozess bereits festlegen.

4.2. Äquivalente Beschreibung

Wir wollen uns nun der Frage widmen, wieso Poissonprozesse ihren Namen tragen und dafür eine äquivalente Beschreibung dieser Prozesse angeben.

Satz 4.2.1 **Charakterisierung**
Ein $\mathbf{N}_0$-wertiger, rechtsstetiger stochastischer Prozess $(N_t)_{t\geq 0}$ ist genau dann ein homogener Poissonprozess mit Rate $\lambda > 0$, wenn die folgenden Eigenschaften erfüllt sind:

i) Es gilt P-fast sicher $N_0 = 0$.

ii) Der Prozess (N_t) hat unabhängige Zuwächse.

iii) Die Zuwächse sind poissonverteilt, für $s, t \geq 0$ gilt also $N_{s+t} - N_s \sim \mathrm{Pois}(\lambda t)$.

Bevor wir diesen Satz beweisen, erinnern wir zunächst daran, dass $\mathrm{Pois}(\lambda)$ durch die Dichte $\frac{\lambda^k}{k!}e^{-\lambda}$ bezüglich des Zählmaßes auf $\mathbf{N}_0$ gegeben ist. Ferner weisen wir darauf hin, dass iii) impliziert, dass die Zuwächse stationär sind und für $s = 0$ folgt mit i) auch $N_t \sim \mathrm{Pois}(\lambda t)$. Insbesondere gilt $\mathbf{E}N_t = \lambda t$, das heißt im Intervall $[0, t]$ treten im Mittel λt Ereignisse ein. Pro Zeiteinheit treten im Mittel also λ Ereignisse ein, es gilt also $\frac{\mathbf{E}N_t}{t} = \lambda$, wie wir eben schon gesehen haben. Diese Tatsache erklärt den Begriff *Rate* für den Parameter λ.

Um Satz 4.2.1 zu beweisen, benötigen wir zunächst zwei Lemmata, die wir zunächst behandeln.

Lemma 4.2.2
Sei $\alpha_n : \mathbf{R}^n \to \mathbf{R}$ gegeben durch $x \mapsto \sum_{i=1}^n x_i$. Für $t \geq 0$ gilt dann

$$\int_0^\infty \cdots \int_0^\infty \mathbf{1}_{\{\alpha_n(x)\leq t\}} \,\mathrm{d}x_1 \cdots \mathrm{d}x_n = \frac{t^n}{n!}.$$

Aufgrund der Nicht-Negativität des Integranden können wir das Integral mit dem Satz von Tonelli auch als Integral über das Produktmaß schreiben oder die Integrationsreihenfolge beliebig verändern.

Beweis: Für den Beweis führen wir eine vollständige Induktion durch. Für $n = 1$ betrachten wir

$$\int_0^\infty \mathbf{1}_{\{\alpha_1(x)\leq t\}} \,\mathrm{d}x = \int_0^t 1 \,\mathrm{d}x = t = \frac{t^1}{1!}.$$

Wir kommen nun zum Induktionsschritt. Es gilt

$$\begin{aligned}\int_0^\infty \cdots \int_0^\infty \mathbf{1}_{\{\alpha_{n+1}(x)\le t\}}\,\mathrm{d}x_1\cdots\mathrm{d}x_n &= \int_0^\infty \cdots \int_0^\infty \mathbf{1}_{\{\alpha_n(x)\le t-x_{n+1}\}}\mathbf{1}_{\{x_{n+1}\le t\}}\,\mathrm{d}x_1\cdots\mathrm{d}x_n \\ &= \int_0^t \frac{(t-x_{n+1})^n}{n!}\,\mathrm{d}x_{n+1} \\ &= \frac{t^{n+1}}{(n+1)!}.\end{aligned}$$

□

Lemma 4.2.3
Sei $(N_t)_{t\ge0}$ ein homogener Poissonprozess mit Rate $\lambda>0$. Dann gilt für alle $s,t\ge0$ für den Zuwachs $N_{s+t}-N_s\sim\mathrm{Pois}(\lambda t)$ und dieser Zuwachs ist von $\sigma(N_r : r\le s)$ unabhängig.

Beweis: Wir wollen für $r\le s$, $k,l\in\mathbf{N}_0$ zunächst zeigen, dass

$$P(N_r=k, N_{s+t}-N_s=l)=\exp(-\lambda r)\frac{(\lambda r)^k}{k!}\exp(-\lambda t)\frac{(\lambda t)^l}{l!} \tag{*}$$

gilt. Wir betrachten zunächst den Spezialfall $r=s$ und die Wartezeiten $(W_n)_{n\ge1}$. Wir wissen, dass die Wartezeiten unabhängig und identisch verteilt sind mit $W_n\sim\mathrm{Exp}(\lambda)$. Die Dichte h von $(W_n)_{n=1}^{k+l+1}$ ist daher das Temsorprodukt der einzelnen Dichten, das heißt für $x\ge0$ ist $h\colon\mathbf{R}^{k+l+1}\to\mathbf{R}$ gegeben durch

$$x\mapsto\prod_{i=1}^{k+l+1}\lambda e^{-\lambda x_i}=\lambda^{k+l+1}\exp(-\lambda\alpha_{k+l+1}(x)),$$

wobei α_n wie in Lemma 4.2.2 gegeben ist. Da $N_r=n$ äquivalent zu $T_n\le r<T_{n+1}$ ist, erhalten wir

$$\begin{aligned}P(N_s=k,N_{s+t}-N_s=l)&=P(T_k\le s<T_{k+1},T_{k+l}\le s+t<T_{k+l+1})\\ &=P\left(\sum_{i=1}^k W_i\le s<\sum_{i=1}^{k+1}W_i,\sum_{i=1}^{k+l}W_i\le s+t<\sum_{i=1}^{k+l+1}W_i\right)\end{aligned}$$

und mit dem Transformationssatz und der Dichte h erhalten wir

$$=\int_0^\infty\cdots\int_0^\infty \mathbf{1}_{\{\alpha_k(x)\le s<\alpha_{k+1}(x),\alpha_{k+l}(x)\le s+t<\alpha_{k+l+1}(x)\}}h(x)\,\mathrm{d}x_1\cdots\mathrm{d}x_{k+l+1}.$$

Im ersten Schritt berechnen wir das Integral bezüglich x_{k+l+1}. Dazu beachten wir, dass $\mathbf{1}_{\{\alpha_k(x)\le s<\alpha_{k+1}(x)\}}$ nur von $x_1,\ldots,x_{k+1}$ abhängt und damit wegen $\mathbf{1}_{A\cap B}=\mathbf{1}_A\mathbf{1}_B$ aus dem Integral herausgezogen und daher in diesem Schritt ignoriert werden kann. Mit $y = \alpha_{k+l+1}(x) = \alpha_{k+l}(x) + x_{k+l+1}$ ist nun

$$\begin{aligned}\int_0^\infty \lambda^{k+l+1} \exp(-\lambda\alpha_{k+l+1}(x))\mathbf{1}_{\{\alpha_{k+l}(x)\le s+t<\alpha_{k+l+1}(x)\}}\, \mathrm{d}x_{k+l+1}&\\
&= \int_{\alpha_{k+l}(x)}^\infty x^{k+l+1} \exp(-\lambda)\mathbf{1}_{\{\alpha_{k+l}(x)\le s+t<y\}}\, \mathrm{d}y\\
&= \int_{s+t}^\infty \lambda^{k+l+1} \exp(-\lambda y)\mathbf{1}_{\{\alpha_{k+l}(x)\le s+t\}}\, \mathrm{d}y\\
&= \lambda^{k+l} \exp(-\lambda(s+t))\mathbf{1}_{\{\alpha_{k+l}(x)\le s+t\}}.\end{aligned}$$

Im zweiten Schritt integrieren wir nun bezüglich $x_{k+1},\dots,x_{k+l+1}$. Es ist

$$\begin{aligned}\int_0^\infty \cdots \int_0^\infty \mathbf{1}_{\{\alpha_k(x)\le s<\alpha_{k+1}(x),\alpha_{k+l}(x)\le s+t<\alpha_{k+l+1}(x)\}} h(x)\, \mathrm{d}x_{k+1}\cdots \mathrm{d}x_{k+l+1}&\\
= \int_0^\infty \cdots \int_0^\infty \lambda^{k+l} \exp(-\lambda(s+t)) \int_{\{\alpha_k(x)\le s<\alpha_{k+1}(x),\alpha_{k+l}(x)\le s+t\}}\, \mathrm{d}x_{k+1}\cdots \mathrm{d}x_{k+l}&\end{aligned}$$

und mit $y_1 = \alpha_{k+1}(x) - s - \alpha_k(x) - x_{k+1} - s$ und $y_2 = x_{k+2},\dots,y_l = x_{k+l}$ ist $s < \alpha_k(x) + x_{k+1}$ genau dann, wenn $y > 0$ gilt. Außerdem ist $y_1 + \ldots + y_l = \alpha_{k+l}(x) - s$ und wir erhalten

$$= \lambda^{k+l} \exp(-\lambda(s+t)) \int_0^\infty \cdots \int_0^\infty \mathbf{1}_{\{\alpha_k(x)\le s\}}\mathbf{1}_{\{\alpha_l(y)\le t\}}\, \mathrm{d}y_1\cdots \mathrm{d}y_l$$

Die erste Indikatorfunktion im Integranden ist von y unabhängig und wir erhalten mit Lemma 4.2.2 daher

$$= \lambda^{k+l} \exp(-\lambda(s+t))\mathbf{1}_{\{\alpha_k(x)\le s\}}\frac{t^l}{l!}.$$

Im dritten Schritt berechnen wir nun das gesamte Integral. Es ist mit Lemma 4.2.2

$$\begin{aligned}\int_0^\infty \cdots \int_0^\infty \mathbf{1}_{\{\alpha_k(x)\le s<\alpha_{k+1}(x),\alpha_{k+l}(x)\le s+t<\alpha_{k+l+1}(x)\}} h(x)\, \mathrm{d}x_1\cdots \mathrm{d}x_{k+l+1}&\\
&= \int_0^\infty \cdots \int_0^\infty \lambda^{k+l} \exp(-\lambda(s+t))\frac{t^l}{l!}\mathbf{1}_{\{\alpha_k(x)\le s\}}\, \mathrm{d}x_1\cdots \mathrm{d}x_k\\
&= \lambda^{k+l} \exp(-\lambda(s+t))\frac{t^l}{l!}\frac{s^k}{k!}\\
&= \exp(-\lambda s)\frac{(\lambda s)^k}{k!} \exp(-\lambda t)\frac{(\lambda t)^l}{l!}.\end{aligned}$$

Nun müssen wir noch den Fall $r < s$ betrachten. Mit einer analogen Rechnung zur bisherigen erhalten wir

$$\begin{aligned}P(N_r = k, N_{s+t} - N_s = l) &= \sum_{m=0}^\infty P(N_r = k, N_s - N_r = m, N_{s+t} - N_s = l)\\
&= \sum_{m=0}^\infty \exp(-\lambda r)\frac{(\lambda r)^k}{k!} \exp(-\lambda(s-r))\frac{(\lambda(s-r))^m}{m!} \exp(-\lambda t)\frac{(\lambda t)^l}{l!}\\
&= \exp(-\lambda r)\frac{(\lambda r)^k}{k!} \exp(-\lambda t)\frac{(\lambda t)^l}{l!}.\end{aligned}$$

Damit haben wir (*) bewiesen. Nun folgt für $r \le s$

$$\begin{aligned} P(N_{s+t} - N_s = l) &= \sum_{k=0}^{\infty} P(N_r = k, N_{s+t} - N_s = l) \overset{(*)}{=} \sum_{k=0}^{\infty} \exp(-\lambda r) \frac{(\lambda r)^k}{k!} \exp(-\lambda t) \frac{(\lambda t)^l}{l!} \\ &= \exp(-\lambda t) \frac{(\lambda t)^l}{l!}. \end{aligned}$$

Also gilt $N_{s+t} - N_s \sim \mathrm{Pois}(\lambda t)$. Für die Unabhängigkeit merken wir an, dass $\sigma(N_r : r \le s)$ von $\{N_r = k\}$ für $r \le s$ und $k \in \mathbf{N}_0$ erzeugt wird. Da $N_r = N_{0+r} - N_0 \sim \mathrm{Pois}(\lambda r)$ ist, erhalten wir

$$P(N_r = k, N_{s+t} - N_s = l) \overset{(*)}{=} \exp(-\lambda r) \frac{(\lambda r)^k}{k!} \exp(-\lambda t) \frac{(\lambda t)^l}{l!} = P(N_r = k) P(N_{s+t} - N_s = l).$$

Damit sind $\{N_r = k\}$ und $\{N_{s+t} - N_s = l\}$ unabhängig. Außerdem bilden die Mengen $\{N_{s+t} - N_s = l\}$ ein $\cap$-stabiles Erzeugendensystem von $\sigma(N_{s+t} - N_s)$, allerdings bilden $\{N_r = k\}$ kein solches $\cap$-stabiles Erzeugendensystem für $\sigma(N_r : r \le s)$. Ein $\cap$-stabiles Erzeugendensystem hiervon wäre zum Beispiel $\{N_{i_1} = k_{i_1}\} \cap \ldots \cap \{N_{i_s} = k_{i_s}\}$, womit man die gesamte Rechnung für (*) nochmals durchführen müsste. Dies werden wir an dieser Stelle nicht ausführen und verweisen daher auf die Literatur. □

Beweis von Satz 4.2.1: Wir beweisen zunächst „⇒“. Dann ist i) per Definition erfüllt und ii) und iii) folgen aus Lemma 4.2.3 mit der Unabhängigkeit von $\sigma(N_r : r \le s)$.

Für „⇐“ sei (N_t) ein stochastischer Prozess, so dass i), ii) und iii) gelten. Ferner sei $(\tilde{N}_t)_{t \ge 0}$ ein homogener Poissonprozess mit Rate $\lambda > 0$. Wir betrachten

$$\begin{aligned} P(N_{t_1} = k_1, \ldots, N_{t_n} = k_n) &= P(N_{t_1} = k_1, N_{t_2} - N_{t_1} = k_2 - k_1, \ldots, N_{t_n} - N_{t_{n-1}} = k_n - k_{n-1}) \\ &= P(N_{t_1} = k_1) \cdot \underbrace{P(N_{t_2} - N_{t_1} = k_2 - k_1)}_{=\mathrm{Pois}((t_2 - t_1)\lambda)(\{k_2 - k_1\})} \cdot \ldots \cdot P(N_{t_n} - N_{t_{n-1}} = k_n - k_{n-1}) \\ &= \tilde{P}(\tilde{N}_{t_1} = k_1) \cdot \tilde{P}(\tilde{N}_{t_2} - \tilde{N}_{t_1} = k_2 - k_1) \cdot \ldots \cdot \tilde{P}(\tilde{N}_{t_n} - \tilde{N}_{t_{n-1}} = k_n - k_{n-1}) \\ &= \tilde{P}(\tilde{N}_{t_1} = k_1, \ldots, \tilde{N}_{t_n} = k_n). \end{aligned}$$

Die Prozesse (N_t) und $(\tilde{N}_t)$ besitzen also die selben endlichdimensionalen Randverteilungen, insbesondere gilt dies dann auch für (W_n) und $(\tilde{W}_n)$, sowie für (S_n) und $(\tilde{S}_n)$. Daher ist (N_t) ein homogener Poissonprozess mit Rate λ. □

Korollar 4.2.4 **Starke Markoveigenschaft**
Sei $(N_t)_{t \ge 0}$ ein homogener Poissonprozess mit Rate $\lambda > 0$. Dann ist

$$\hat{N}_t := N_{s+t} - N_s$$

mit $t \ge 0$ ein homogener Poissonprozess mit Rate λ, der von $\sigma(N_r : r \le s)$ unabhängig ist.

Ein homogener Poissonprozess kann zum Zeitpunkt s also unabhängig neugestartet werden.

Beweis: Wir sehen, dass $\hat{N}_0 = N_s - N_s = 0$ gilt. Für die dritte Eigenschaft gemäß Satz 4.2.1 sehen wir

$$\hat{N}_{r+t} - \hat{N}_r = N_{s+r+t} - N_s - N_{s+r} + N_s = N_{s+r+t} - N_{s+r} \sim \text{Pois}(\lambda t).$$

Für die zweite Eigenschaft gilt $\hat{N}_{r+t} - \hat{N}_r = N_{s+r+t} - N_{s+r}$, die unabhängigen Zuwächse von (N_t) garantieren dann die unabhängigen Zuwächse von $(\hat{N}_t)$. Mit Satz 4.2.1 ist $(\hat{N}_t)$ also ein homogener Poissonprozess mit Rate λ. Die Unabhängigkeit von $\sigma(N_r : r \leq s)$ folgt aus Lemma 4.2.3. $\square$

4.3. Weitere Eigenschaften

In diesem Abschnitt wollen wir weitere Eigenschaften von Poissonprozessen untersuchen, die für praktische Anwendungen von Interesse sind.

Satz 4.3.1 **Summenbildung**
Seien $(N_t)_{t\geq 0}$ und $(M_t)_{t\geq 0}$ unabhängige, homogene Markovprozesse mit der Rate λ bzw. μ. Dann ist $(N_t + M_t)_{t\geq 0}$ ein homogener Poissonprozess mit Rate $\lambda + \mu$.

Beweis: Wir verwenden für den Beweis die Charakterisierung aus Satz 4.2.1. Klar ist, dass i) erfüllt ist. Für ii) gilt, dass (N_t) und (M_t) unabhängige Zuwächse haben und voneinander unabhängig sind. Daher besitzt auch die Summe unabhängige Zuwächse.

Wir wollen nun noch iii) zeigen, dass die Zuwächse $N_{s+t} + M_{s+t} - N_s - M_s$ also poissonverteilt mit Rate $\lambda + \mu$ sind. Dazu betrachten wir

$$
\begin{aligned}
P(N_{s+t} + M_{s+t} - N_s - M_s = k) &= \sum_{m=0}^{k} P(N_{s+t} - N_s = m \text{ und } M_{s+t} - M_s = k - m) \\
&= \sum_{m=0}^{k} P(N_{s+t} - N_s = m) P(M_{s+t} - M_s = k - m) \\
&= \sum_{m=0}^{k} e^{-\lambda t} \frac{(\lambda t)^m}{m!} e^{-\mu t} \frac{(\mu t)^{k-m}}{(k-m)!}.
\end{aligned}
$$

Wir können die Exponentialfunktionen aus der Summe ziehen und müssen nur noch die verbleibende Summe betrachten. Wir erhalten mit dem binomischen Lehrsatz

$$
\begin{aligned}
\sum_{m=0}^{k} \frac{(\lambda t)^m}{m!} \frac{(\mu t)^{k-m}}{(k-m)!} &= \frac{t^k}{k!} \sum_{m=0}^{k} \frac{k!}{m!(k-m)!} \lambda^m \mu^{k-m} \\
&= \frac{(\lambda + \mu)^k t^k}{k!} \sum_{m=0}^{k} \binom{k}{m} \left(\frac{\lambda}{\lambda + \mu} \right)^m \left(\frac{\mu}{\lambda + \mu} \right)^{k-m} \\
&= \frac{(\lambda + \mu)^k t^k}{k!}.
\end{aligned}
$$

Dann folgt also

$$
P(N_{s+t} + M_{s+t} - N_s - M_s = k) = e^{-(\lambda+\mu)t} \frac{((\lambda + \mu)t)^k}{k!}. \qquad \square
$$

Es seien nun $(N_t^{(1)}), \ldots, (N_t^{(m)})$ unabhängige, homogene Poissonprozesse, jeweils mit den Raten $\lambda_1, \ldots, \lambda_m$ und den Wartezeiten $(W_n^{(1)}), \ldots, (W_n^{(m)})$. Wir betrachten dann $N_t := N_t^{(1)} + \ldots + N_t^{(m)}$ und wollen untersuchen, was für die Wartezeiten von (N_t) gilt.

Zunächst ist klar, dass $W_1 := \min\{W_1^{(1)}, \dots, W_1^{(m)}\}$ die erste Wartezeit von (N_t) ist. Es sei $J := \operatorname{argmin}_{j=1,\dots,m} W_1^{(j)}$ der Index des Prozesses, der das erste Ereignis von N_t verursacht. Ferner gilt $W_1 \sim \operatorname{Exp}(\lambda)$ und $W_1^{(J)} = W_1$, wobei $\lambda := \lambda_1 + \ldots + \lambda_m$ ist.

Satz 4.3.2
Mit den obigen Bezeichnungen gelten die folgenden Aussagen:

i) J und W_1 sind unabhängig.

ii) Es gilt $P(J = i) = \dfrac{\lambda_j}{\lambda}$ für alle $i = 1, \dots, m$.

Für den Beweis des Satzes benötigen wir ein Lemma, das wir dem Beweis voranstellen.

Lemma 4.3.3
Es seien X_1 und X_2 unabhängige Zufallsvariablen, deren Verteilungen die Lebesgue-Dichten f_1 und f_2 besitzen. Ferner sei $g\colon \mathbf{R}^2 \to \mathbf{R}$ mit $g(X_1, X_2) \in \mathscr{L}_1$. Dann gelten die folgenden Aussagen:

i) Es ist $\mathbf{E}g(X_1, X_2) = \displaystyle\int_{-\infty}^{\infty} \mathbf{E}g(x, X_2) f_1(x)\,\mathrm{d}x$.

ii) Es gilt $P(t \le X_1 \le X_2) = \displaystyle\int_t^{\infty} P(X_2 \ge x) f_1(x)\,\mathrm{d}x$ für alle $t \in \mathbf{R}$.

Beweis: Für die erste Eigenschaft betrachten wir

$$\begin{aligned}
\mathbf{E}g(X_1, X_2) = \mathbf{E}_{P_{X_1,X_2}} g = \mathbf{E}_{P_{X_1} \otimes P_{X_2}} g &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x_1, x_2) f_1(x_1) f_2(x_2)\,\mathrm{d}x_2\,\mathrm{d}x_1 \\
&= \int_{-\infty}^{\infty} f_1(x_1) \int_{-\infty}^{\infty} g(x_1, x_2) f_2(x_2)\,\mathrm{d}x_2\,\mathrm{d}x_1 \\
&= \int_{-\infty}^{\infty} \mathbf{E}g(x_1, X_2) f_1(x_1)\,\mathrm{d}x_1.
\end{aligned}$$

Für den Beweis der zweiten Eigenschaft betrachten wir $g(X_1, X_2) := \mathbf{1}_{[t,\infty)}(x_1)\mathbf{1}_{[x_1,\infty)}(x_2)$. Dann gilt $\mathbf{E}g(X_1, X_2) = P(t \le X_1 \le X_2)$ und $\mathbf{E}g(x, X_2) = \mathbf{1}_{[t,\infty)}(x) P(X_2 \ge x)$. Dann folgt die Aussage mit der ersten Eigenschaft. □

Beweis von Satz 4.3.2: Für $j = 1, \dots, m$ und $t \ge 0$ zeigen wir zunächst

$$P(J = i, W_1 \ge t) = \frac{\lambda_j}{\lambda} e^{-\lambda t}. \tag{*}$$

Dazu sei o. B. d. A. $j = 1$. Dann ist $V := \min\{W_1^{(2)}, \dots, W_1^{(m)}\}$ die erste Wartezeit von $(N_t^{(2)} + \dots + N_t^{(m)})$ und es gilt $V \sim \mathrm{Exp}(\lambda_2 + \dots + \lambda_m)$. Wir erhalten

$$\begin{aligned} P(J = 1, W_1 \geq t) = P(t \leq W_1^{(1)} \leq V) &\overset{4.3.3}{=} \int_t^\infty P(V \geq x)\lambda_1 e^{-\lambda_1 x}\,\mathrm{d}x \\ &= \int_t^\infty \exp\left(-x\sum_{i=2}^m \lambda_i\right)\lambda_1 e^{-\lambda_1 x}\,\mathrm{d}x \\ &= \lambda_1 \int_t^\infty e^{-\lambda x}\,\mathrm{d}x \\ &= \frac{\lambda_1}{\lambda} e^{-\lambda t}. \end{aligned}$$

Damit ist (*) gezeigt. Für $t \searrow 0$ folgt nun $P(J = j, W_1 \geq t) \to P(J = j)$, aber auch $P(J = j, W_1 \geq t) \to \frac{\lambda_j}{\lambda}$, es gilt also $P(J = j) = \frac{\lambda_j}{\lambda}$. Dies beweist ii). Für i) gilt

$$P(J = j, W_1 \geq t) \overset{(*)}{=} \frac{\lambda_j}{\lambda} e^{-\lambda t} = P(J = j)P(W_1 \geq t),$$

also erhalten wir die Unabhängigkeit auf einem ∩-stabilem Erzeugendensystem und sind fertig. □

Beispiel 4.3.4 *Bank*
Wir betrachten eine Bank mit drei Eingängen. Wir nehmen an, dass die Kunden durch diese Eingänge poissonverteilt mit den Raten $\lambda_1, \lambda_2, \lambda_3$ kommen. Mit den Bezeichnungen des Satzes 4.3.2 folgt, dass die Wahrscheinlichkeit dafür, dass der erste Kunde durch den Eingang j kommt, gerade $\frac{\lambda_j}{\lambda_1+\lambda_2+\lambda_3}$ beträgt. Für die durchschnittliche Wartezeit auf den ersten Kunden gilt ferner

$$\mathbf{E}W_1 = \frac{1}{\lambda_1 + \lambda_2 + \lambda_3}.$$

Die starke Markoveigenschaft aus Korollar 4.2.4 sichert hierbei, dass diese Formeln auch für alle weiteren Kunden gelten. //

Wir wollen uns auch noch mit der Frage beschäftigen, was passiert, wenn die Kunden in Beispiel 4.3.4 auch Ein- oder Auszahlungen vornehmen.

Definition 4.3.5 **Compound Process**
Sei (N_t) ein homogener Poissonprozess mit Rate λ. Ferner sei (Y_n) eine Folge von i. i. d. Zufallsvariablen auf $\mathbf{R}$, die von (N_t) unabhängig ist. Dann heißt der Prozess $(C_t)_{t\geq 0}$, der durch

$$C_t := \sum_{i=1}^{N_t} Y_i$$

definiert ist, der Compound Process von (N_t) und (Y_n).

Jedes von (N_t) gezählte Ereignis löst eine „Aktion X_i" aus C_t aus. Nun beschreibt C_t die Summe dieser Aktionen bis zur Zeit t. Im Beispiel 4.3.4 kann man Y_i als Einzahlung des i-ten Kunden betrachten, wobei $Y_i(\omega) < 0$ einer Auszahlung entspricht. Dann beschreibt C_t die Bilanz dieser Vorgänge bis zur Zeit t aus Sicht der Bank.

Satz 4.3.6
Sei N eine $\mathbf{N}_0$-wertige Zufallsvariable und (Y_n) eine Folge von i. i. d. Zufallsvariablen, die von N unabhängig ist. Wir setzen

$$C := \sum_{i=1}^{N} Y_i.$$

Dann gelten die folgenden Aussagen:

i) Waldsche Identität: Sind $N, Y_1 \in \mathscr{L}_1$, so gilt $C \in \mathscr{L}_1$ und $\mathbf{E}C = \mathbf{E}N\mathbf{E}Y_1$.

ii) Blackwell-Girshick: Sind $N, Y_1 \in \mathscr{L}_2$, so gilt $C \in \mathscr{L}_2$ und

$$\operatorname{Var} C = \mathbf{E}N \operatorname{Var} Y_1 + \operatorname{Var} N (\mathbf{E}Y_1)^2.$$

iii) Ist $N \sim \operatorname{Pois}(\lambda)$ und $N, Y_1 \in \mathscr{L}_2$, so ist $C \in \mathscr{L}_2$ und es gilt $\mathbf{E}C = \lambda \mathbf{E}Y_1$ und $\operatorname{Var} C = \lambda \mathbf{E}Y_1^2$.

Beweis: Der Beweis wird zur Übung überlassen. □

Beispiel *Fortsetzung von Beispiel 4.3.4*
Wir hatten uns bereits gefragt, was passiert, wenn man in Beispiel 4.3.4 auch Ein- und Auszahlungen erlaubt. Wir nehmen daher an, dass Kunde i die Einzahlung Y_i durchführt. Wir fragen uns nun, wie die Bilanz der Bank zur Zeit t aussieht.

Dazu sei $N_t \sim \operatorname{Pois}(\lambda t)$ mit $\lambda = \lambda_1 + \lambda_2 + \lambda_3$, die Y_i seien i. i. d. und es sei $C_t := \sum_{i=1}^{N_t} Y_i$. Aus Satz 4.3.3 folgt dann $\mathbf{E}C_t = \lambda t \mathbf{E}Y_1$ und $\operatorname{Var} C_t = \lambda t \mathbf{E}Y_1^2$, sowohl Erwartungswert als auch Varianz verändern sich also linear mit der Zeit. Für $a \geq 0$ gilt dann mit der Markov-Ungleichung

$$P(|C_t| \geq a) \leq \frac{\mathbf{E}C_t^2}{a^2} = \frac{\lambda t \mathbf{E}Y_1^2 + \lambda^2 t^2 (\mathbf{E}Y_1)^2}{a^2}.$$

Diese Abschätzung ist jedoch nicht besonders gut und kann stark verbessert werden. Für die Verteilung von C_t lässt sich auch explizit eine Formel angeben, die jedoch kompliziert und häufig schwierig auszuwerten ist. //

Wir nehmen nun an, dass k Ereignisse beobachtet wurden. Wie sieht dann die Verteilung der T_i aus? Das folgende Lemma 4.3.7 besagt, dass die bedingte Verteilung für T_1 die Gleichverteilung auf $[0, t]$ ist.

Lemma 4.3.7 **Bedingte Verteilung von T_1**
Sei (N_t) ein homogener Poissonprozess mit Rate λ. Dann gilt für $0 \leq s \leq t$

$$P(T_1 \leq s \mid N_t = 1) = \frac{s}{t}.$$

Beweis: Da $T_1 \leq s$ äquivalent zu $N_s = 1$ ist, gilt mit $N_0 = 0$, der Unabhängigkeit der Zuwächse und mit der Dichte der Poissonverteilung

$$\begin{aligned} P(T_1 \leq s \mid N_t = 1) &= \frac{P(T_1 \leq s, N_t = 1)}{P(N_t = 1)} = \frac{P(N_s - N_0 = 1, N_t - N_s = 0)}{P(N_t = 1)} \\ &= \frac{P(N_s = 1)P(N_t - N_s = 0)}{P(N_t = 1)} = \frac{\frac{\lambda s e^{-\lambda s}}{1!} \frac{(\lambda(t-s))^0 e^{-\lambda(t-s)}}{0!}}{\frac{\lambda t e^{-\lambda t}}{1!}} \\ &= \frac{s}{t}. \end{aligned}$$

□

Satz 4.3.8 **Bedingte Verteilung von $T_1, \dots, T_k$**
Sei (N_t) ein homogener Poissonprozess mit Rate λ. Dann hat die Verteilung der $T_1, \dots, T_k$ unter der Bedingung $N_t = k$, das heißt $P((T_1, \dots, T_k) \in \bullet \mid N_t = k)$, auf $\mathbf{R}^k$ die Lebesgue-Dichte

$$\begin{aligned} h : \mathbf{R}^k &\to \mathbf{R} \\ (t_1, \dots, t_k) &\mapsto \frac{k!}{t^k} \mathbf{1}_{\{0 \leq t_1 \leq \ldots \leq t_k \leq t\}}. \end{aligned}$$

Dies ist die Dichte der Ordnungsstatistik von k unabhängigen und auf $[0, t]$ gleichverteilten Zufallsvariablen.

Beweis: Wir werden den Beweis hier nicht führen, da er konzeptionell einfach ist, aber eine längere Rechnung erfordert. Nachzulesen ist er in [Meintrup04, Satz 10.27]. □

4.4. Inhomogene Poissonprozesse

Bisher war die Rate λ konstant, also von der Zeit t unabhängig. Wir wollen nun untersuchen, was passiert, wenn sie sich mit der Zeit ändert.

Definition 4.4.1 Inhomogener Poissonprozess
Sei $(N_t)_{t\geq 0}$ ein rechtsstetiger, $\mathbf{N}_0$-wertiger stochastischer Prozess und $\lambda\colon [0,\infty) \to (0,\infty)$ Lebesgue-integrierbar auf allen kompakten Intervallen. Dann heißt $(N_t)_{t\geq 0}$ ein nicht-homogener oder inhomogener Poissonprozess genau dann, wenn folgende Eigenschaften gelten:

i) Es gilt $N_0 = 0$.

ii) Der Prozess (N_t) hat unabhängige Zuwächse.

iii) Für $s,t \geq 0$ ist der Zuwachs $N_{s+t} - N_s$ poissonverteilt mit der Rate $\displaystyle\int_s^{s+t} \lambda(u)\,\mathrm{d}u$.

Wir haben also die Charakterisierung aus Satz 4.2.1 verwendet, um von den homogenen auf die inhomogenen Poissonprozesse zu verallgemeinern. Jeder homogene Poissonprozess ist insbesondere ein inhomogener Poissonprozess mit konstanter Ratenfunktion $\lambda(t) := \lambda_0$. Den Nachweis zur Existenz und Eindeutigkeit echt inhomogener Poissonprozesse werden wir hier nicht führen und verweisen daher auf die Literatur.

Definition 4.4.2 Mittelwertfunktion
Ist (N_t) ein inhomogener Poissonprozess mit Ratenfunktion λ, so heißt $m\colon [0,\infty) \to (0,\infty)$ mit $t \mapsto \mathbf{E}N_t$ Mittelwertfunktion.

Satz 4.4.3 Eigenschaften der Mittelwertfunktion
Sei (N_t) ein inhomogener Poissonprozess mit Ratenfunktion λ und Mittelwertfunktion m. Dann gelten die folgenden Aussagen:

i) Die Mittelwertfunktion m ist eine Stammfunktion von λ, also gegeben durch

$$m(t) = \int_0^t \lambda(u)\,\mathrm{d}u.$$

ii) Die Mittelwertfunktion m ist streng monoton wachsend und stetig.

iii) Ist m^* streng monoton wachsend und stetig, so existiert λ^* mit $m^*(t) = \int_0^t \lambda^*(u)\,\mathrm{d}u$ für $t \geq 0$, also definiert m^* einen inhomogenen Poissonprozess mit Ratenfunktion λ^*. Ferner ist m^* Lebesgue-fast überall differenzierbar und es gilt $(m^*)' = \lambda^*$ fast sicher, also ist λ^* fast sicher eindeutig.

Im Fall homogener Poissonprozesse gilt also $m(t) = \mathbf{E}N_t = \lambda t$ und damit $m' = \lambda$.

Beweis: Für i) beachten wir, dass $N_t = N_t - N_0$ poissonverteilt mit Rate $\int_0^t \lambda(u)\,\mathrm{d}u$ ist. Dann gilt $m(t) = \mathbf{E}N_t = \int_0^t \lambda(u)\,\mathrm{d}u$, da der Erwartungswert α-poissonverteilter Zufallsvariablen gerade α ist.

Die strenge Monotonie in ii) ist mit i) klar, da $\lambda(t) > 0$ gilt. Für die Stetigkeit betrachten wir für $0 \le s \le t \le b$ das endliche Maß Q_b auf $\mathbf{R}$, welches die Lebesguedichte

$$h_b(u) = \begin{cases} \lambda(u) & \text{falls } u \in [0,b] \\ 0 & \text{sonst} \end{cases}$$

besitzt. Dann gilt

$$m(t) - m(s) = \int_s^t \lambda(u)\,\mathrm{d}u = Q_b([s,t])$$

und wir erhalten die Stetigkeit von m aus der Stetigkeit von Q_b.

Ist m^* streng monoton wachsend, so spricht man davon, dass m^* von beschränkter Variation ist. Es folgt, dass m^* fast überall differenzierbar mit den Eigenschaften aus iii) ist. Wir wollen dies hier nicht ausführen und verweisen daher auf [Kestelman60] oder [Graves56]. □

Korollar 4.4.4 **Compound Process**
Sei (N_t) ein inhomogener Poissonprozess mit Mittelwertfunktion m. Dann gelten für den durch

$$C_t := \sum_{i=1}^{N_t} Y_i$$

definierten Compound Process von N_t und einer von (N_t) unabhängigen Folge mit i. i. d. Zufallsvariablen Y_i die folgenden Eigenschaften:

i) Ist $Y_1 \in \mathscr{L}_1$, so folgt $C_t \in \mathscr{L}_1$ und es gilt

$$\mathbf{E}C_t = m(t)\mathbf{E}Y_1.$$

ii) Ist $Y_1 \in \mathscr{L}_2$, so folgt $C_t \in \mathscr{L}_2$ und es gilt

$$\operatorname{Var} C_t = m(t)\mathbf{E}Y_1^2.$$

Insbesondere sind $t \mapsto \mathbf{E}C_t$ und $t \mapsto \operatorname{Var} C_t$ streng monoton (oder konstant Null) und stetig.

Beweis: Für den Erwartungswert erhält man mit Satz 4.3.6

$$\mathbf{E}C_t = \mathbf{E}N_t \mathbf{E}Y_1 = m(t)\mathbf{E}Y_1.$$

Für die Varianz erhält man wegen $\mathbf{E}X = \alpha = \operatorname{Var} X$ für $X \sim \operatorname{Pois}(\alpha)$ wiederum mit Satz 4.3.6

$$\operatorname{Var} C_t = \mathbf{E}N_t \operatorname{Var} Y_1 + \operatorname{Var} N_t (\mathbf{E}Y_1)^2 = m(t)\mathbf{E}Y_1^2.$$ □

Korollar 4.4.5 **Summen**
Sind $(N_t^{(1)})$ und $(N_t^{(2)})$ unabhängige, inhomogene Poissonprozesse mit Mittelwert- und Ratenfunktionen m_1 und λ_1 bzw. m_2 und λ_2, so ist $(N_t^{(1)} + N_t^{(2)})$ ein inhomogener Poissonprozess mit Mittelwertfunktion $m_1 + m_2$ und Ratenfunktion $\lambda_1 + \lambda_2$.

Beweis: Für den Beweis, dass die Summe einen inhomogenen Poissonprozess definiert, müssen wir die drei Eigenschaften aus Definition 4.4.1 nachweisen. Die erste Eigenschaft ist jedoch klar und die zweite Eigenschaft können wir wie in Satz 4.3.1 zeigen. Die dritte und letzte Eigenschaft ist eine ähnliche Rechnung wie im Beweis von Satz 4.3.1 und wird hier nicht bewiesen. □

Beispiel *Fortsetzung von Beispiel 4.3.4 – Teil 2*
Wieder betrachten wir eine Bank, der erste Eingang sei nun jedoch gegenüber einer Bushaltestelle, der zweite Eingang münde in eine Fußgängerzone und der dritte Eingang ist ein Zugang zu einem Bürohochhaus, dessen Angestellte zwischen 12 Uhr und 13 Uhr Mittagspause haben. Die Rechnungen verlaufen dann analog zu denen bei homogenen Poissonprozessen. //

5 Wiener-Prozess

Im Jahr 1827 beobachtete Robert Brown die Bewegungen von Pollen in Wasser und erkannte eine unregelmäßige Bewegung. Der Durchbruch in der Beschreibung dieses Verhaltens gelang 1905 von Albert Einstein, unabhängig von vorherigen wichtigen Arbeiten zu diesem Thema, indem er die Wärmebewegung der Moleküle im Wasser betrachtete. Dazu musste er jedoch die Existenz eines stochastischen Prozesses postulieren, die erstmals 1923 vom US-Amerikaner Norbert Wiener (*26.11.1894 – †18.03.1964) bewiesen wurde.

Wiener-Prozesse finden besonders in der Physik, in den Ingenieurs-Wissenschaften und in der Finanzmathematik breite Anwendung. Im Gegensatz zu den Prozessen, die wir bisher kennengelernt haben, sind Wiener Prozesse sowohl in der Zeit als auch im Zustandsraum stetig.

5.1. Definition und einfache Eigenschaften

Definition 5.1.1 Wiener-Prozess
Ein $\mathbf{R}$-wertiger stochastischer Prozess $(W_t)_{t\geq 0}$ heißt Wiener-Prozess oder Brownsche Bewegung, wenn die folgenden Eigenschaften erfüllt sind:

i) Es gilt $W_0 = 0$ P-fast sicher.

ii) Der Prozess (W_t) hat unabhängige Zuwächse.

iii) Die Zuwächse $W_{s+t} - W_s$ sind $\mathcal{N}(0,t)$-verteilt für alle $s \geq 0$ und $t > 0$.

iv) Der Prozess (W_t) ist stetig, P-fast alle Pfade sind also stetig.

Die Existenz eines stochastischen Prozesses, der i) – iii) erfüllt, ist mit dem, was wir bisher kennengelernt haben, nicht besonders schwierig. Die Stetigkeit als zusätzliche Eigenschaft stellt sich schwieriger beim Beweis der Existenz heraus.

Man nennt einen stochastischen Prozess Lévy-Prozess, wenn die Zuwächse unabhängig und stationär sind. Homogene Poissonprozesse und Wiener-Prozesse sind Beispiele hierfür.

Die drei ersten Eigenschaften in Definition 5.1.1 legen die endlich-dimensionalen Randverteilungen bereits fest und sichern damit die Eindeutigkeit. Wir werden dies später noch genauer untersuchen.

Multivariate Normalverteilung

Wir wollen an dieser Stelle nochmals an multivariate Normalverteilungen erinnern und unseren Kenntnisstand aus der Wahrscheinlichkeitstheorie ergänzen. Für mehr Details verweisen wir auf [Meintrup04, Kap. 7].

Sei $\mu \in \mathbf{R}^d$ und Σ eine strikt positiv definite und symmetrische $(d \times d)$-Matrix. Dann hat die multivariate Normalverteilung $\mathcal{N}(\mu, \Sigma)$ die Lebesgue-Dichte

$$h(x) = \frac{1}{(2\pi)^{\frac{d}{2}} (\det \Sigma)^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu) \right).$$

Für $X \sim \mathcal{N}(\mu, \Sigma)$ mit $X = (X_1, \ldots, X_d)$ gilt $\mathbf{E}X_i = \mu_i$ und $\operatorname{Cov}(X_i, X_j) := \mathbf{E}X_i X_j - \mathbf{E}X_i \mathbf{E}X_j = \Sigma_{ij}$ für alle $i, j = 1, \ldots, d$. Daher nennen wir Σ auch Kovarianzmatrix.

Lemma 5.1.2
Es sei $X \sim \mathcal{N}(\mu, \Sigma)$, dann sind die Komponenten $X_1, \ldots, X_d$ unabhängig genau dann, wenn Σ eine Diagonalmatrix ist, wenn also $\Sigma_{ij} = 0$ für alle $i \neq j$ gilt.

Beweis: Der Beweis findet sich in [Meintrup04, Satz 7.33]. □

Satz 5.1.3
Eine Zufallsvariable X ist genau dann multivariat normalverteilt, wenn für alle $t \in \mathbf{R}^d \setminus \{0\}$ gilt, dass $\sum_{i=1}^d t_i X_i$ normalverteilt ist.

Beweis: Der Beweis findet sich in [Meintrup04, Satz 7.35]. □

Definition 5.1.4 **Gauß-Prozess**
Ein stochastischer Prozess heißt Gauß-Prozess, wenn jede endlich-dimensionale Randverteilung multivariat normalverteilt ist.

Wir werden sehen, dass Wiener-Prozesse auch Gauß-Prozesse sind. Dies wollen wir im Folgenden näher diskutieren.

Lemma 5.1.5
Sei $\mu\colon [0,\infty) \to \mathbf{R}$ und $\Gamma\colon [0,\infty) \times [0,\infty) \to \mathbf{R}$ symmetrisch und strikt positiv definit, dann gibt es einen Gauß-Prozess $(X_t)_{t\geq 0}$ mit $\mathbf{E}X_t = \mu(t)$ für $t \geq 0$ und $\operatorname{Cov}(X_t, X_s) = \Gamma(t,s)$ für $t,s \geq 0$.

Ist (Y_t) ein weiterer stochastischer Prozess dieser Art, so haben (X_t) und (Y_t) die gleichen endlich-dimensionalen Randverteilungen.

Beweis: Wir werden den Beweis hier nicht führen und überlassen ihn zur Übung. □

Satz 5.1.6 **Wiener-Prozess als Gauß-Prozess**
Sei $(W_t)_{t\geq 0}$ ein stochastischer Prozess. Dann sind folgende Aussagen äquivalent:

i) (W_t) ist ein Wiener-Prozess.

ii) (W_t) ist ein stetiger, zentrierter Gauß-Prozess (es gilt also $\mathbf{E}W_t = 0$ für alle $t \geq 0$) mit der Kovarianzfunktion $\Gamma(s,t) = \min\{s,t\}$ für alle $s,t \geq 0$.

Beweis: Wir betrachten zunächst i) ⇒ ii). Die Stetigkeit ist klar und für die Zentriertheit betrachten wir $\mathbf{E}W_t = \mathbf{E}(W_t - W_0) = 0$. Für beliebige $t_1 < t_2 < \ldots < t_n$ sind $W_{t_0}, W_{t_1} - W_{t_0}, \ldots, W_{t_n} - W_{t_{n-1}}$ unabhängig und normalverteilt. Die gemeinsame Verteilung ist dann multivariat normal. Nach Satz 5.1.3 ist dann jede nicht-triviale Linearkombination der Zuwächse normalverteilt. Dann ist auch jede nicht-triviale Linearkombination von $W_{t_1}, \ldots, W_{t_n}$ normalverteilt und wieder mit Satz 5.1.3 folgt, dass $(W_{t_1}, \ldots, W_{t_n})$ multivariat normalverteilt ist. Damit ist (W_t) ein Gauß-Prozess. Für $s \leq t$ gilt

$$\begin{aligned}
\operatorname{Cov}(W_t, W_s) &= \mathbf{E}W_tW_s - \mathbf{E}W_t\mathbf{E}W_s = \mathbf{E}W_tW_s = \mathbf{E}(W_t - W_s + W_s)W_s = \mathbf{E}(W_t - W_s)W_s + \mathbf{E}W_s^2 \\
&= \mathbf{E}(W_t - W_s)(W_s - W_0) + s = \mathbf{E}(W_t - W_s)\mathbf{E}(W_s - W_0) + s \\
&= s.
\end{aligned}$$

Wegen $s \leq t$ ist $\Gamma(s,t) = \min\{s,t\} = s$ und wir haben daher auch diese Eigenschaft gezeigt.

Für ii) ⇒ i) ist die Stetigkeit klar, da wir diese voraussetzen. Ferner gilt $\mathbf{E}W_0 = 0$ und $\operatorname{Var} W_0 = \operatorname{Cov}(W_0, W_0) = 0$ und wir erhalten $W_0 = 0$ P-fast sicher. Nach Satz 5.1.3 ist jeder Zuwachs $W_{s+t} - W_s$ normalverteilt und es gilt $\mathbf{E}(W_{s+t} - W_s) = 0$ und

$$\begin{aligned}
\operatorname{Var}(W_{s+t} - W_s) &= \mathbf{E}(W_{s+t} - W_s)^2 = \mathbf{E}W_{s+t}^2 - 2\mathbf{E}W_{s+t}W_s + \mathbf{E}W_s^2 \\
&= \Gamma(s+t, s+t) - 2\Gamma(s+t, s) + \Gamma(s,s) = s + t - 2s + s \\
&= t,
\end{aligned}$$

also gilt $W_{s+t} - W_s \sim \mathcal{N}(0,t)$. Sei nun $0 \leq t_0 < \ldots < t_n$. Da wir einen Gauß-Prozess haben, ist $(W_{t_1}, \ldots, W_{t_n})$ multivariat normalverteilt und mit Satz 5.1.3 folgt, dass jede nicht-triviale

Linearkombination der Komponenten normalverteilt ist, damit auch jede nicht-triviale Linearkombination von $W_{t_1}, W_{t_2} - W_{t_1}, \ldots, W_{t_n} - W_{t_{n-1}}$ normalverteilt und wieder mit Satz 5.1.3 ist daher $(W_{t_1}, W_{t_2} - W_{t_1}, \ldots, W_{t_n} - W_{t_{n-1}})$ multivariat normalverteilt. Für $i < j$ gilt ferner

$$
\begin{aligned}
\operatorname{Cov}(W_{t_i} - W_{t_{i-1}}, W_{t_j} - W_{t_{j-1}}) \\
&= \mathbf{E}(W_{t_i} - W_{t_{i-1}})(W_{t_j} - W_{t_{j-1}}) \\
&= \mathbf{E}W_{t_i}W_{t_j} - \mathbf{E}W_{t_i}W_{t_{j-1}} - \mathbf{E}W_{t_{i-1}}W_{t_j} + \mathbf{E}W_{t_{i-1}}W_{t_{j-1}} \\
&= \operatorname{Cov}(W_{t_i}, W_{t_j}) - \operatorname{Cov}(W_{t_i}, W_{t_{j-1}}) - \operatorname{Cov}(W_{t_{i-1}}, W_{t_j}) + \operatorname{Cov}(W_{t_{i-1}}, W_{t_{j-1}}) \\
&= \Gamma(t_i, t_j) - \Gamma(t_i, t_{j-1}) - \Gamma(t_{i-1}, t_j) + \Gamma(t_{i-1}, t_{j-1}) \\
&= t_i - t_i - t_i + t_i \\
&= 0.
\end{aligned}
$$

Damit sind die Zuwächse nach Lemma 5.1.2 unabhängig. □

Korollar 5.1.7
Ist $(W_t)_{t\geq 0}$ ein Wiener-Prozess, so sind auch die folgenden Prozesse Wiener-Prozesse:

i) $(-W_t)_{t\geq 0}$.

ii) $(W_{s+t} - W_s)_{t\geq 0}$ für festes $s \geq 0$.

iii) $\left(\frac{1}{c} W_{c^2 t}\right)_{t\geq 0}$ für festes $c > 0$.

iv) $(W_{t_0} - W_{t_0 - t})_{t\in[0,t_0]}$ für festes $t_0 > 0$.

Beweis: Der Beweis wird zur Übung überlassen. □

Korollar 5.1.7 besagt also unter anderem, dass der Neustart zur Zeit s wieder einen Wiener-Prozess ergibt, dass ein Prozess in der Zeit umskaliert werden kann und, bis auf einen Faktor, die Trajektorien gleich aussehen.

Lemma 5.1.8
Es sei (X_t) ein stochastischer Prozess mit $X_0 = 0$. Dann sind die folgenden Aussagen äquivalent:

i) Der Prozess (X_t) hat unabhängige Zuwächse.

ii) Für alle $t \geq s \geq 0$ ist $X_t - X_s$ von $\sigma(X_r : r \leq s)$ unabhängig.

Beweis: Wir betrachten i) $\Rightarrow$ ii) und fixieren $0 \leq s \leq t$. Für $n \geq 1$ und $0 \leq s_0 < s_1 < \ldots < s_n = s$ gilt dann

$$\sigma(X_{s_0}, X_{s_1}, \ldots, X_{s_n}) = \sigma(X_{s_0}, X_{s_1} - X_{s_0}, \ldots, X_{s_n} - X_{s_{n-1}}).$$

Diese σ-Algebra ist nach Voraussetzung von $X_t - X_s$ unabhängig. Damit erhalten wir, dass

$$\bigcup_{n=1}^{\infty} \left\{(X_{s_0}, X_{s_1}, \ldots, X_{s_n}) : 0 \leq s_0 < s_1 < \ldots < s_n = s\right\}$$

von $X_t - X_s$ unabhängig ist und dies ist außerdem ein $\cap$-stabiles Erzeugendensystem von $\sigma(X_r : r \leq s)$.

Die Richtung ii) $\Rightarrow$ i) ist klar. □

Korollar 5.1.9 **Markov-Eigenschaft**
Sei (W_t) ein Wiener-Prozess und $s \geq 0$. Dann ist der Wiener-Prozess $(W_{s+t} - W_s)_{t \geq 0}$ von $\sigma(X_r : r \leq s)$ unabhängig.

Beweis: Dass tatsächlich ein Wiener-Prozess vorliegt folgt aus Satz 5.1.7, die Unabhängigkeit folgt mit Lemma 5.1.8. □

5.2. Existenz des Wiener-Prozesses

In diesem Abschnitt wollen wir zeigen, dass Wiener-Prozesse tatsächlich existieren. Dazu werden wir zeigen, dass gewisse stochastische Prozesse eine stetige Version besitzen und zentrierte Gauß-Prozesse mit der Kovarianzfunktion $\Gamma(s,t) = \min\{s,t\}$ gehören hier dazu. Der Wiener-Prozess ist dann die stetige Version dieses Gauß-Prozesses.

Definition 5.2.1 **Hölder-Stetigkeit**
Sei $I \subset \mathbf{R}$ ein Intervall $f : I \to \mathbf{R}$. Dann heißt f lokal Hölder-stetig mit Ordnung $\gamma > 0$ genau dann, wenn für alle $x \in I$ ein $c > 0$ und $\delta > 0$ existieren, so dass

$$|f(y) - f(z)| \le c|y - z|^\gamma$$

für alle $y, z \in I$ mit $|x - y| < \delta$ und $|x - z| < \delta$ gilt.

Ferner heißt f Hölder-stetig mit Ordnung γ, wenn ein $c > 0$ existiert, so dass

$$|f(y) - f(z)| \le c|y - z|^\gamma$$

für alle $y, z \in I$ gilt.

Für $\gamma = 1$ ist dies Lipschitz-Stetigkeit, Hölder-Stetigkeit mit der Ordnung $\gamma > 1$ impliziert, dass f konstant ist. Es gibt Hölder-stetige Funktionen, die nicht Lipschitz-stetig sind, zum Beispiel $\sqrt{\cdot} : [0,\infty) \to [0,\infty)$. Lokale Hölder-Stetigkeit impliziert ferner Stetigkeit.

Lemma 5.2.2
Sei $f : I \to \mathbf{R}$ lokal Hölder-Stetigkeit mit Ordnung γ. Dann gelten die folgenden Aussagen:

i) Die Abbildung f ist Hölder-stetig zur Ordnung γ' für alle $\gamma' \in (0, \gamma]$.

ii) Ist I kompakt, so ist f Hölder-stetig mit Ordnung γ.

iii) Ist I beschränkt mit Länge T und gibt es $\delta > 0$ und $c > 0$ mit

$$|f(y) - f(z)| \le c|y - z|^\gamma$$

für alle $y, z \in I$ mit $|y - z| < \delta$, so ist f Hölder-stetig der Ordnung γ mit $c' := c\left(\frac{T}{\delta}\right)^{1-\gamma}$.

Beweis: Die erste Eigenschaft ist klar.

Für ii) geben wir ein typisches Kompaktheitsargument an: Schreibe c_x und δ_x für die Konstanten in Definition 5.2.1, dann bilden die $B(x, \delta_x)$ für $x \in I$ eine offene Überdeckung von I. Dann existieren $x_1, \dots, x_n \in I$ mit $I \subset \bigcup_{i=1}^n B(x_i, \delta_{x_i})$ und ohne Einschränkung sei $x_1 < x_2 < \dots < x_n$. Beachte nun, dass die Ungleichung für die lokale Hölder-Stetigkeit auf $B(x_i, \delta_{x_i})$ global gilt mit den Konstanten c_{x_i} und δ_{x_i}. Wir nehmen weiter ohne Einschränkung an, dass $x_1, \dots, x_n$

minimal ist in dem Sinne, dass keine dieser x_i weggenommen werden kann, ohne die Überdeckungseigenschaft zu verletzen. Wir setzen nun $\delta := \frac{1}{2}\min_i x_i + \delta_{x_i} - (x_{i+1} - \delta_{x_{i+1}})$, für alle $x \in I$ existiert dann ein $i \in \{1, \dots, n\}$ mit $B(x, \delta) \subset B(x_i, \delta_{x_i})$. Für $y, z \in I$ mit $|y - z| < \delta$ gilt dann $y, z \in B(x, \delta)$ mit $x := \frac{y+z}{2}$. Dann gilt die definierende Ungleichung der Hölder-Stetigkeit mit der Konstanten c_{x_i}. Setzen wir $c' := \max_i c_{x_i}$, so erhalten wir dies global. Umgekehrt gilt für $y, z \in I$ mit $|y - z| \geq \delta$

$$|f(y) - f(z)| \leq 2\,\|f\|_\infty \leq 2\,\|f\|_\infty \left(\frac{|y-z|}{\delta}\right)^\gamma,$$

insgesamt setzen wir also $c' := \max\{c_{x_1}, \dots, c_{x_n}, 2\,\|f\|_\infty \delta^{-\gamma}\}$ und erhalten Hölder-Stetigkeit.

Für iii) sei $n := \lceil \frac{T}{\delta} \rceil$, was der Anzahl der Intervalle der Länge δ in I angibt. Für $y', z' \in I$ gilt

$$\frac{|y' - z'|}{n} \leq \delta.$$

Damit gilt für $y, z \in I$ und ohne Einschränkung $y < z$ durch Bilden einer Teleskopsumme und der Dreiecksungleichung

$$\begin{aligned}
|f(y) - f(z)| &\leq \sum_{k=1}^{n} \left| f\left(y + (y-z)\frac{k}{n}\right) - f\left(y + (y-z)\frac{k-1}{n}\right)\right| \\
&\leq \sum_{k=1}^{n} c \left|(y-z)\frac{1}{n}\right|^\gamma \\
&= c n^{1-\gamma} |y-z|^\gamma.
\end{aligned}$$

□

Satz 5.2.3 **Kolmogorov-Chentsov**
Sei $(X_t)_{t\geq 0}$ ein $\mathbf{R}$-wertiger stochastischer Prozess, für den es Konstanten $\alpha, \beta, c > 0$ gibt mit

$$P(|X_t - X_s| \geq \varepsilon) \leq c\frac{|t-s|^{1+\beta}}{\varepsilon^\alpha} \tag{*}$$

für alle $s, t \geq 0$ und $\varepsilon > 0$. Dann gelten die folgenden Aussagen:

i) Es existiert eine Version $(\tilde{X}_t)_{t\geq 0}$ von $(X_t)_{t\geq 0}$, die P-fast sicher lokal Hölder-stetige Pfade mit Ordnung γ für $0 < \gamma < \frac{\beta}{\alpha}$ hat.

ii) Für jedes $0 < \gamma < \frac{\beta}{\alpha}$ und alle $\varepsilon > 0$ und $T < \infty$ existiert ein $k > 0$ mit

$$P\left(|\tilde{X}_t - \tilde{X}_s| \leq k|t-s|^\gamma \ \forall_{s,t \in [0,T]}\right) \geq 1 - \varepsilon.$$

Die Ungleichung in (*) bedeutet im Wesentlichen, dass X_t im α-ten Mittel Hölder-stetig ist. Die erste Aussage des Satzes liefert die Existenz einer Hölder-Konstante für jeden Pfad, der zweite Teil besagt, dass für viele Pfade die selbe Konstante gewählt werden kann.

Es gibt Verallgemeinerungen des Satzes, in denen die **R**-Wertigkeit durch $\mathscr{X}$-Wertigkeit ersetzt werden kann, falls $\mathscr{X}$ ein vollständiger metrischer Raum ist. Die Zeit t kann durch mehrdimensionale Indexbereiche ersetzt werden, wobei $1+\beta$ dann zu $d+\beta$ wird.

Beweis: Zunächst werden wir zwei Vereinfachungen vornehmen. Es genügt, die Existenz von lokal Hölder-stetigen Versionen auf $[0,T]$ für alle $T>0$ zu zeigen. Dazu sei $(X_t^T)_{t\geq 0}$ eine lokal Hölder-stetige Version auf $[0,T]$. Für $S,T>0$ sind dann nach Satz 1.2.6 (X_t^T) und (X_t^S) auf $[0,\min\{S,T\}]$ nicht unterscheidbar. Daraus folgt, dass für $\Omega_{S,T} := \{\omega : \exists_{t\in[0,S\wedge T]} \text{ mit } X_t^T(\omega) \neq X_t^S(\omega)$, $X^T(\omega)$ und $X^S(\omega)$ lokal Hölder-stetig$\}$ gilt, dass $P(\Omega_{S,T})=0$ ist. Damit ist auch $\Omega_\infty := \bigcup_{S,T\in\mathbf{N}}\Omega_{S,T}$ eine Nullmenge. Wir setzen nun $\tilde{X}_t(\omega) := X_t^t(\omega)$ für $\omega\in\Omega\setminus\Omega_\infty$. Für $s,t\in[0,\infty)$ mit $s\leq t$ gibt es dann T mit $s\leq t\leq T$ und es folgt

$$\left|\tilde{X}_t(\omega)-\tilde{X}_s(\omega)\right| = \left|X_t^t(\omega)-X_s^s(\omega)\right| = \left|X_t^T(\omega)-X_s^T(\omega)\right|,$$

woraus man unmittelbar die Abschätzung der Hölder-Stetigkeit gewinnen kann. Als weitere Vereinfachung wollen wir begründen, dass es sogar genügt, lediglich $T=1$ zu betrachten. In den anderen Fällen kann die Zeit auf $[0,1]$ umskaliert werden, da dies nicht die lokale Hölder-Stetigkeit ändert.

Wir kommen nun zum eigentlichen Beweis. Wir wollen zunächst zeigen, dass (X_t) auf den Zahlen in $[0,1]\cap\mathbf{Q}$ mit endlicher Binärdarstellung lokal Hölder-stetig ist. Da diese Zahlen dicht in $[0,1]$ liegen, gibt es dann eine lokal Hölder-stetige Fortsetzung $(\tilde{X}_t)$, die sich als gesuchte Version herausstellen wird. Dazu sei $n\in\mathbf{N}$ und $k\in\{1,\dots,2^n\}$, dann gilt

$$P(|X_{k2^{-n}}-X_{(k-1)2^{-n}}|\geq 2^{-\gamma n}) \overset{(*)}{\leq} c\frac{2^{-n(1+\beta)}}{2^{-\alpha\gamma n}} = c\cdot 2^{-n(1+\beta-\alpha\gamma)}. \tag{**}$$

Wir setzen nun

$$A_n := \left\{\max_{k\in\{1,\dots,2^n\}} |X_{k2^{-n}}-X_{(k-1)2^{-n}}|\geq 2^{-\gamma n}\right\},$$
$$B_n := \bigcup_{m=n}^{\infty} A_m \qquad \text{und} \qquad N := \limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} B_n.$$

Wir wollen nun zeigen, dass $P(N)=0$ gilt. Mit (**) folgt hierzu

$$\begin{aligned} P(A_n) &\leq \sum_{k=1}^{2^n} P(|X_{k2^{-n}}-X_{(k-1)2^{-n}}|\geq 2^{-\gamma n}) \\ &\leq c\cdot 2^{n-n(1+\beta-\alpha\gamma)} \\ &= c\cdot 2^{-n(\beta-\alpha\gamma)}. \end{aligned}$$

Wegen $\gamma<\frac{\beta}{\gamma}$ folgt weiter

$$\sum_{n=1}^{\infty} P(A_n)\leq c\sum_{n=1}^{\infty} 2^{-n(\beta-\alpha\gamma)} < \infty.$$

Mit dem Satz von Borel-Cantelli folgt dann $P(N) = P(\limsup A_n) = 0$. Nun betrachten wir $\Omega \setminus N$ und fixieren daher $\omega \in \Omega \setminus N$. Dann ist $\omega \in \Omega \setminus N = \Omega \setminus \left(\bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m\right) = \bigcup_{n=1}^{\infty} \left(\Omega \setminus \bigcap_{m=n}^{\infty} A_m\right)$, also existiert n_0 mit $\omega \notin \bigcup_{m=n_0}^{\infty} A_m = B_{n_0}$. Nach Konstruktion folgt dann

$$|X_{k2^{-n}}(\omega) - X_{(k-1)2^{-n}}(\omega)| < 2^{-\gamma n} \tag{***}$$

für alle $n \geq n_0$ und $k = 1, \ldots, 2^n$.

Wir betrachten nun die dyadischen Zahlen mit endlicher Darstellung, d. h. es sei $D_m := \{k2^{-m} : k = 0, \ldots, 2^m\}$ und dann $D := \bigcup_{m=1}^{\infty} D_m$. Unser Ziel ist es, (***) auf D zu verallgemeinern. Dazu beachten wir, dass jedes $t \in D_m$ eine eindeutige Binärdarstellung

$$t = \sum_{i=0}^{m} b_i(t) 2^{-i}$$

mit geeigneten $b_i(t) \in \{0,1\}$ besitzt. Für $t \in D_m$ mit $t < 2^{-n}$ für $m \geq n$ gilt ferner $b_i(t) = 0$ für alle $i = 0, \ldots, n$. Dies folgt, da $2^{-n} > t = \sum_{i=0}^{m} b_i(t) 2^{-i} \geq b_{i^*}(t) 2^{-i^*}$ gilt, also muss $b_{i^*}(t) = 0$ für $i^* \leq n$ sein.
Sei nun $m \geq n \geq n_0$ und $s, t \in D_m$ mit $s \leq t \leq s + 2^{-n}$, d. h. es gilt $0 \leq t - s \leq 2^{-n}$. Da D_n eine Partition von $[0,1]$ der Feinheit 2^{-n} definiert, gibt es ein $u \in D_n$ mit $u \leq s < u + 2^{-n}$ und $u \leq t < s + 2^{-n} < u + 2^{1-n}$. Damit folgt $0 \leq s - u < 2^{-n}$ und $0 \leq t - u < 2^{-(n-1)}$. Für die Binärdarstellung von $t - u \in D_m$ und $s - u \in D_m$ folgt nun

$$b_i(s-u) = b_i(t-u) = 0$$

für alle $i < n$. Wir setzen nun $t_l := u + \sum_{i=n}^{l} b_i(t-u) 2^{-i}$ für $l = n-1, \ldots, m$. Dann gilt $t_{n-1} = u$, $t_m = t$, $t_l \in D_l$ und $|t_l - t_{l-1}| \leq 2^{-l}$. Wir erhalten nun schließlich

$$\begin{aligned}
|X_t(\omega) - X_u(\omega)| &\leq \sum_{l=n}^{m} |X_{t_l}(\omega) - X_{t_{l-1}}(\omega)| \leq \sum_{l=n}^{m} 2^{-\gamma l} \\
&\leq \frac{2^{-\gamma n}}{1 - 2^{-\gamma}}.
\end{aligned}$$

Analog erhalten wir auch

$$|X_s(\omega) - X_u(\omega)| \leq \frac{2^{-\gamma n}}{1 - 2^{-\gamma}}.$$

Für $m \geq n \geq n_0$ und $s, t \in D_m$ mit $|s - t| \leq 2^{-n}$ folgt also

$$|X_s(\omega) - X_t(\omega)| \leq \underbrace{\frac{2}{1 - 2^{-\gamma}}}_{=:\tilde{c}} 2^{-\gamma n}. \tag{****}$$

Der in (****) auftauchende Faktor $(2^{-n})^{\gamma}$ muss nun noch durch $|s - t|^{\gamma}$ ersetzt werden, um Hölder-Stetigkeit zu erhalten. Dazu seien $s, t \in D$ mit $|s - t| \leq 2^{-n_0}$ für ein n_0, wobei dieses n_0 maximal gewählt sein soll, es gilt also $|s - t| > 2^{-(n_0+1)} = 2^{-n_0-1}$. Für $n \geq n_0$ gilt dann

$|s-t| > 2^{-n-1}$. Ferner gibt es ein $m \geq 1$ mit $s,t \in D_m$ und ohne Einschränkung sei $m \geq n$. Damit gilt

$$|X_s(\omega) - X_t(\omega)| \overset{(****)}{\leq} \tilde{c}2^{-\gamma n_0} = \tilde{c}2^{\gamma}2^{-(n_0+1)\gamma} \leq 2^{\gamma}\tilde{c}|s-t|^{\gamma}.$$

Mit anderen Worten haben wir auf Intervallen der Länge 2^{-n_0} also γ-Hölder-Stetigkeit. Mit einer analogen Argumentation wie in Lemma 5.2.2 iii) existiert eine Konstante $k = 2^{\gamma}\tilde{c}2^{(n_0+1)(1-\gamma)}$ mit

$$|X_s(\omega) - X_t(\omega)| \leq k|s-t|^{\gamma}$$

für alle $s,t \in D$. Der betrachtete Pfad ist also global Hölder-stetig auf D. Da $D \subset [0,1]$ dicht ist, existiert eine Hölder-stetige Fortsetzung $\tilde{X}(\omega)$ von $X(\omega)$ auf $[0,1]$.

Es bleibt nun noch zu zeigen, dass der so gefundene Prozess $\tilde{X}$ eine Version von X ist. Für $t \in D$ gilt nach Konstruktion $P(X_t \neq \tilde{X}_t) = P(N) = 0$. Für $t \notin D$ mit $(t_n) \subset D$ und $t_n \to t$ folgt

$$P(X_t \neq \tilde{X}_t) = P(X_t \neq \lim_{n\to\infty} X_{t_n}) = 0,$$

da wir $X_{t_n} \to X_t$ stochastisch aus (*) erhalten.

Nun kommen wir zur zweiten Aussage. Sei $\varepsilon > 0$ und $n \geq 1$ so groß, dass

$$P(B_n) \leq \sum_{m=n}^{\infty} P(A_m) \leq c \sum_{m=n}^{\infty} 2^{-m(\beta-\alpha\gamma)} \leq \varepsilon$$

erfüllt ist. Für $\omega \notin B_n$ folgt dann mit der obigen Rechnung

$$|X_t(\omega) - X_s(\omega)| \leq k|s-t|^{\gamma}$$

für alle $s,t \in [0,1]$. Durch eine Transformation kann man dies auf allgemeine T erweitern. □

Korollar 5.2.4 **Existenz des Wiener-Prozesses**
Es existiert ein Wiener-Prozess. Ferner haben je zwei Wiener-Prozesse die gleichen endlich-dimensionalen Randverteilungen. Schließlich sind fast alle Pfade eines Wiener-Prozesses lokal Hölder-stetig zur Ordnung γ für alle $\gamma \in \left(0, \frac{1}{2}\right)$.

Beweis: Für die Existenz zeigen wir zunächst, dass ein zentrierter Gauß-Prozess mit Kovarianzfunktion $\Gamma(s,t) := \min\{s,t\}$ für alle $s,t \in \mathbf{R}$ existiert. Im Hinblick auf Lemma 5.1.5 zeigen wir zunächst, dass Γ strikt positiv definit ist. Dazu seien $t_1,\dots,t_n \in [0,\infty)$ paarweise verschieden und $\alpha \in \mathbf{R}^n \setminus \{0\}$, dann gilt

$$\begin{aligned}\sum_{i,j=1}^n \alpha_i\alpha_j\Gamma(t_i,t_j) &= \sum_{i,j=1}^n \alpha_i\alpha_j \int_0^\infty \mathbf{1}_{[0,t_i]}(x)\mathbf{1}_{[0,t_j]}(x)\,\mathrm{d}x \\ &= \int_0^\infty \sum_{i,j=1}^n \alpha_i\alpha_j\mathbf{1}_{[0,t_i]}(x)\mathbf{1}_{[0,t_j]}(x)\,\mathrm{d}x \\ &= \int_0^\infty \left(\sum_{i=1}^n \alpha_i\mathbf{1}_{[0,t_i]}(x)\right)^2 \mathrm{d}x \\ &\geq 0.\end{aligned}$$

Da für paarweise verschiedene $t_1,\dots,t_n$ die Familie $(\mathbf{1}_{[0,t_i]})_{i=1,\dots,n}$ linear unabhängig ist, folgt $\sum \alpha_i \mathbf{1}_{[0,t_i]} = 0$ λ-fast sicher genau dann, wenn $\alpha_1 = \ldots = \alpha_n = 0$ gilt. Da wir genau solche α jedoch ausgeschlossen haben, ist obiges Integral sogar strikt positiv. Nach Lemma 5.1.5 existiert also ein zentrierter Gauß-Prozess mit Kovarianzfunktion Γ. Wir wollen nun noch eine stetige Version von (X_t) finden, die dann nach Satz 5.1.6 einen Wiener-Prozess liefert. Dazu beobachten wir, dass der Beweis von Satz 5.1.6 nicht die Stetigkeit verwendet hatte, also sind die Zuwächse von (X_t) unabhängig und $X_t - X_s$ ist $\mathcal{N}(0,t-s)$-verteilt für $t \geq s$, insbesondere ist also $X_1 = X_1 - X_0 \sim \mathcal{N}(0,1)$. Für $n \geq 1$ erhalten wir damit

$$\mathbf{E}(X_t - X_s)^{2n} = \mathbf{E}(\sqrt{t-s}X_1)^{2n} = (t-s)^n\mathbf{E}X_1^{2n} =: (t-s)^n c_n.$$

Mit der Markovungleichung erhalten wir dann

$$P(|X_t - X_s| \geq \varepsilon) \leq \frac{\mathbf{E}|X_t - X_s|^{2n}}{\varepsilon^{2n}} = c_n\frac{|t-s|^n}{\varepsilon^{2n}}.$$

Setzen wir nun $1+\beta = n$ und $\alpha = 2n$, so besitzt (X_t) nach Satz 5.2.3 eine γ-Hölderstetige Version für $\gamma < \frac{\beta}{\alpha} = \frac{n-1}{2n} \to \frac{1}{2}$. □

Definition 5.2.5 Hölder-Stetigkeit in einem Punkt
Eine Funktion $f: I \to \mathbf{R}$ für ein Intervall I heißt Hölder-stetig der Ordnung $\gamma > 0$ im Punkt $x_0 \in I$ genau dann, wenn es $c,\delta > 0$ gibt, so dass für alle $x \in I$ mit $|x - x_0| < \delta$ gilt:

$$|f(x) - f(x_0)| \leq c|x - x_0|^\gamma.$$

Lokale Hölder-Stetigkeit impliziert insbesondere Hölder-Stetigkeit in allen Punkten. Ist f in x_0 differenzierbar, so ist f auch 1-Hölder-stetig in x_0.

Satz 5.2.6 **Irregularität der Pfade (Paley-Zygmund)**
Für $\gamma > \frac{1}{2}$ sind die Pfade eines Wiener-Prozesses fast sicher in keinem $t_0 \geq 0$ Hölder-stetig zur Ordnung γ. Insbesondere ist fast sicher jeder Pfad eines Wiener-Prozesses nirgends differenzierbar.

In der Analysis werden Funktionen, die stetig, aber nirgends differenzierbar sind, als Skurrilität angesehen und nur exemplarisch behandelt, bei den Wiener-Prozessen sind sie hingegen Normalität.

Beweis: Es genügt, $T = [0,1]$ zu betrachten. Für $t \in T$ und $\gamma > \frac{1}{2}$ definieren wir

$$H_{\gamma,t} := \{f : T \to \mathbf{R} \mid f \text{ stetig und } \gamma\text{-Hölder-stetig in } t\}$$

und

$$H_\gamma := \bigcup_{t \in T} H_{\gamma,t}.$$

Wir wollen zeigen, dass $W(\omega) \notin H_\gamma$ für fast alle ω gilt. Dazu sei $t \in T$ und $f \in H_{\gamma,t}$. Ferner sei $\delta > 0$ und $c > 0$ mit

$$|s-t| < \delta \qquad \Rightarrow \qquad |f(t) - f(s)| \leq c|t-s|^\gamma.$$

Wir wählen nun $k > \frac{2}{2\gamma - 1}$ und setzen $n_0 := \lceil \frac{k+2}{\delta} \rceil$, für $i = \lfloor t_n \rfloor + 1$ und $l \in \{0, \ldots, k-1\}$ sowie $n \geq n_0$ gilt dann

$$\left| f\left(\frac{i+l+1}{n}\right) - f\left(\frac{i+l}{n}\right) \right| \leq \left| f\left(\frac{i+l+1}{n}\right) - f(t) \right| + \left| f(t) - f\left(\frac{i+l}{n}\right) \right|.$$

Es gilt $\lfloor t_n \rfloor + 1 + (k-1) + 1 = \lfloor t_n \rfloor + k + 1$ und daher folgt

$$\begin{aligned}
\left| \frac{i+l+1}{n} - t \right| &\leq \left| \frac{\lfloor t_n \rfloor}{n} + \frac{k+1}{n} - t \right| \\
&\leq \left| \frac{\lfloor t_n \rfloor}{n} - t \right| + \frac{k+1}{n} \\
&= t - \frac{\lfloor t_n \rfloor}{n} + \frac{k+1}{n} \\
&\leq t - \frac{t_n - 1}{n} + \frac{k+1}{n} = \frac{1}{n} + \frac{k+1}{n} = \frac{k+2}{n} \\
&\leq \delta.
\end{aligned}$$

Insgesamt erhalten wir damit

$$\left| f\left(\frac{i+l+1}{n}\right) - f\left(\frac{i+l}{n}\right) \right| \leq 2c \left(\frac{k+2}{n}\right)^\gamma.$$

Für $N > 2c(k+2)^\gamma$ gilt also

$$f \in A_{N,n,i} := \bigcap_{l=0}^{k-1} \left\{ g : \left| g\left(\frac{i+l+1}{n}\right) - g\left(\frac{i+l}{n}\right) \right| \leq Nn^{-\gamma} \right\}.$$

Wir setzen nun $A_{N,n} := \bigcup_{i=1}^{n} A_{N,n,i}$, $A_N := \bigcap_{n \geq n_0} A_{N,n}$ und schließlich $A := \bigcup_{N=1}^{\infty} A_N$. Für $f \in A_{N,n}$ für alle $n \geq n_0$ haben wir gezeigt, dass $f \in A_N$ für unser gewähltes N gilt und dass dann $f \in A$ gilt. Insgesamt erhalten wir damit $H_\gamma \subset A$. Jetzt wollen wir zeigen, dass die Pfade fast sicher nicht in A liegen. Dazu betrachten wir

$$\begin{aligned} P\left(\{\omega : W(\omega) \in A_{N,n,i}\}\right) &= P\left(\omega : \left| W_{\frac{i+l+1}{n}}(\omega) - W_{\frac{i+l}{n}}(\omega) \right| \leq Nn^{-\gamma} \ \forall_{l=0,\dots,k-1}\right) \\ &= \left(P(|W_{\frac{1}{n}}| \leq Nn^{-\gamma})\right)^k \\ &= \left(P(|W_1| \leq Nn^{-\gamma+\frac{1}{2}})\right)^k \\ &\leq 2\left(Nn^{-\gamma+\frac{1}{2}}\right)^k. \end{aligned}$$

Damit erhalten wir nun

$$\begin{aligned} P(\omega : W(\omega) \in A_N) &= P\left(\omega : W(\omega) \in \bigcap_{m \geq n_0} A_{N,m}\right) \\ &\leq \lim_{n\to\infty} P\left(\omega : W(\omega) \in \bigcap_{m \geq n} A_{N,m}\right) \\ &\leq \limsup_{n\to\infty} P(\omega : W(\omega) \in A_{N,m}) \\ &\leq \limsup_{n\to\infty} \sum_{i=1}^{n} P(\omega : W(\omega) \in A_{N,n,i}) \\ &\leq \limsup_{n\to\infty} 2n\left(Nn^{-\gamma+\frac{1}{2}}\right)^k = 0 \end{aligned}$$

für alle $N \geq 1$ und wegen $k > \frac{2}{2\gamma-1}$, so dass schließlich $P(W(\omega) \in A) = 0$ folgt. □

5.3. Die starke Markoveigenschaft

Bis jetzt haben wir gesehen, dass $(W_{s+t} - W_s)_{t\geq 0}$ ein Wiener-Prozess ist, der von $\mathscr{F} := \sigma(W_r : r \leq s)$ unabhängig ist. Wir wollen dies nun auf Stoppzeiten verallgemeinern.

Definition 5.3.1 **Stoppzeit**
Sei $\mathscr{F} = (\mathscr{F}_t)_{t\geq 0}$ eine Filtration und $\tau : \Omega \to [0,\infty]$. Dann heißt τ $\mathscr{F}$-Stoppzeit genau dann, wenn $\{\tau \leq t\} \in \mathscr{F}_t$ für alle $t \geq 0$ gilt.

Beispiel 5.3.2 *Passierzeit von Wiener-Prozessen*
Für $a \in \mathbf{R}$ ist $\tau_a := \inf\{t \geq 0 : W_t = a\}$ eine Stoppzeit. Den Beweis, der die Stetigkeit der Pfade verwendet, überlassen wir zur Übung. //

Analog zu Definition 5.3.1 wollen wir auch andere bereits aus früheren Kapiteln bekannte Begriffe definieren:

- Der Prozess (X_t) heißt an die Filtration $\mathscr{F} = (\mathscr{F}_t)_{t\geq 0}$ adaptiert, falls $\{X_t \in B\} \in \mathscr{F}_t$ für alle $t \geq 0$ und $B \in \mathscr{B}$ gilt.
- Sei τ eine endliche $\mathscr{F}$-Stoppzeit und (X_t) ein an $\mathscr{F}$ adaptierter Prozess. Dann setzen wir $X_\tau : \Omega \to \mathbf{R}$ mit $X_\tau(\omega) := X_{\tau(\omega)}(\omega)$ für $\omega \in \Omega$.
- Die σ-Algebra der τ-Vergangenheit ist definiert durch

$$\mathscr{F}_\tau := \{A \in \mathscr{A} : A \cap \{\tau \leq t\} \in \mathscr{F}_t \quad \forall_{t\geq 0}\}.$$

Im zeitdiskreten Fall ist X_τ auch $\mathscr{F}_\tau$-messbar. Im zeitkontinuierlichen Fall gilt dies im Allgemeinen jedoch nicht. Dies führt zu folgender Definition:

Definition 5.3.3 **Progressiv-Messbarkeit**
Ein stochastischer Prozess $X = (X_t)_{t\geq 0}$ heißt $\mathscr{F}$-progressiv-messbar genau dann, wenn die Abbildung

$$\begin{aligned} \Omega \times [0,t] &\to \mathbf{R}, \\ (\omega, s) &\mapsto X_s(\omega) \end{aligned}$$

$\mathscr{F}_t \otimes \mathbf{B}([0,t])$-messbar für alle $t \geq 0$ ist.

Mit anderen Worten fordert man also gemeinsame Messbarkeit in ω und s. Ist (X_t) $\mathscr{F}$-progressiv-messbar, so ist (X_t) insbesondere $\mathscr{F}$-adaptiert.

Lemma 5.3.4
Ist (X_t) $\mathscr{F}$-adaptiert und links- oder rechtsstetig, so ist (X_t) auch progressiv-messbar. Insbesondere sind Wiener-Prozesse, homogene Poissonprozesse und inhomogene Poissonprozesse progressiv-messbar.

Beweis: Wir behandeln hier nur den rechtsstetigen Fall und nehmen ohne Einschränkung $t = 1$ an. Definiere nun

$$Y_n(\omega, s) := \begin{cases} X_{(j+1)2^{-n}}(\omega) & \text{für } s \in [j2^{-n}, (j+1)2^{-n}),\ j = 0, \dots, 2^n - 1 \\ X_s(\omega) & \text{für } s \geq 1 \end{cases}.$$

Da (X_t) rechtsstetig ist, folgt $\lim_{n\to\infty} Y_n(\omega, s) = X_s(\omega)$ für $(\omega, s) \in \Omega \times [0, \infty)$. Wir zeigen noch, dass $Y_n\big|_{\Omega\times[0,t]}$ messbar bezüglich $\mathscr{F}_1 \otimes \mathscr{B}([0,1])$ ist. Dazu betrachten wir

$$\begin{aligned}\{Y_n \in B\} = &\bigcup_{j=0}^{2^n-1} \{X_{(j+1)2^{-n}} \in B\} \times [j2^{-n}, (j+1)2^{-n}) \cup \{X_1 \in B\} \times \{1\} \\ &\in \mathscr{F}_1 \otimes \mathscr{B}([0,1]).\end{aligned}$$

□

Satz 5.3.5 **Messbarkeit von X_τ**
Sei $\mathscr{F} = (\mathscr{F}_t)$ eine Filtration, $X = (X_t)$ sei ein $\mathscr{F}$-progressiv-messbarer stochastischer Prozess und τ eine $\mathscr{F}$-Stoppzeit. Dann ist X_τ $\mathscr{F}_\tau$-messbar.

Beweis: Sei $B \subset \mathscr{B}$ und $t \geq 0$. Wir wollen zeigen, dass $\{X_\tau \in B\} \cap \{\tau \leq t\} \in \mathscr{F}_t$ gilt, diese Menge ist aber darstellbar als $\{X_{\tau\wedge t} \in B\} \cap \{\tau \leq t\}$ und es gilt $\{\tau \leq t\} \in \mathscr{F}_t$, es genügt also $\{X_{\tau\wedge t} \in B\} \in \mathscr{F}_t$ zu zeigen. Da $\tau \wedge t \leq t$ eine Stopzeit ist, können wir ohne Einschränkung annehmen, dass $\tau \leq t$ gilt, zeigen also $\{X_\tau \in B\} \in \mathscr{F}_t$, also die $\mathscr{F}_t$-Messbarkeit von X_τ. Dazu betrachten wir $\psi : \Omega \to \Omega \times [0, t]$ mit $\omega \mapsto (\omega, \tau(\omega))$. Für $A \in \mathscr{F}_t$ und $u \leq t$ gilt nun

$$\{\psi \in A \times [0, u]\} = \{\omega \in \Omega : \omega \in A \text{ und } \tau(\omega) \leq u\} = A \cap \underbrace{\{\tau \leq u\}}_{\in \mathscr{F}_u \subset \mathscr{F}_t} \in \mathscr{F}_t,$$

also ist ψ messbar bezüglich $(\mathscr{F}_t, \mathscr{F}_t \otimes \mathscr{B}([0,t]))$. Ferner gilt $X_\tau = X \circ \psi$ und da X bezüglich $(\mathscr{F}_t \otimes \mathscr{B}([0,t]), \mathscr{B})$ messbar ist folgt daher, dass X_τ als Komposition $(\mathscr{F}_t, \mathscr{B})$-messbar ist. □

Satz 5.3.6 **Starke Markoveigenschaft des Wiener-Prozesses**
Es sei $(W_t)_{t\geq 0}$ ein Wiener-Prozess mit natürlicher Filtration $\mathscr{F} = (\mathscr{F}_t)$ und $\tau : \Omega \to [0, \infty]$ eine fast sicher endliche $\mathscr{F}$-Stoppzeit. Dann ist $(X_t)_{t\geq 0}$ definiert durch $X_t := W_{t+\tau} - W_\tau$ für $t \geq 0$ ein Wiener-Prozess, der von $\mathscr{F}_\tau$ unabhängig ist.

Ist $\tau = t_0$ konstant, so entspricht dies der einfachen Markoveigenschaft aus Korollar 5.1.9.

Beweis: Wir wollen folgende Identitäten für alle $t_1 < \ldots < t_n$, $n \geq 1$, $A \in \mathscr{B}^n$ und $B \in \mathscr{F}_\tau$ zeigen:

$$\begin{array}{ccc} P((X_{t_1},\ldots,X_{t_n}) \in A \text{ und } B) & \overset{\text{i)}}{=} & P((X_{t_1},\ldots,X_{t_n}) \in A) \cdot P(B) \\ \overset{\text{ii)}}{=} & & \overset{\text{iii)}}{=} \\ & P((W_{t_1},\ldots,W_{t_n}) \in A) \cdot P(B) & \end{array}$$

Ist dies gezeigt, so folgt die Unabhängigkeit aus i) und da (W_t) und (X_t) nach iii) die selben Randverteilungen besitzen, ist (X_t) ein zentrierter Gauß-Prozess mit Kovarianzfunktion $\Gamma(s,t) = \min\{s,t\}$. Da (X_t) offensichtlich auch stetig ist, folgt nach Satz 5.1.6 also, dass (X_t) ein Wiener-Prozess ist. Der Beweis von i) – iii) erfolgt nun in zwei Schritten:

Im ersten Schritt nehmen wir an, dass τ nur abzählbar viele Werte annimmt. Sei $\tau(\Omega)$ also abzählbar und $(s_n)_{n\geq 1}$ eine entsprechende Abzählung. Für $n \geq 1$ und $B \in \mathscr{F}_\tau$ gilt dann

$$B \cap \{\tau = s_n\} = \underbrace{B \cap \{\tau \leq s_n\}}_{\in \mathscr{F}_{s_n}} \cap \{\tau = s_n\},$$

der letzte Teil lässt sich als $\{\tau = s_n\} = \{\tau \leq s_n\} \setminus \bigcup_{s_m < s_n} \{\tau \leq s_m\}$ schreiben lässt. Nun ist $\{\tau \leq s_n\} \in \mathscr{F}_{s_n}$ und $\{\tau \leq s_m\} \in \mathscr{F}_{s_m} \subset \mathscr{F}_{s_n}$, insgesamt also $B \cap \{\tau = s_n\} \in \mathscr{F}_{s_n}$. Auf $\{\tau = s_n\}$ gilt nun $X_t = W_{t+\tau} - W_\tau = W_{t+s_n} - W_{s_n} =: Y_t^{(n)}$. Beachte, dass $(Y^{(n)})_{t\geq 0}$ ein Wiener-Prozess ist, der nach Korollar 5.1.9 von $\mathscr{F}_{s_n}$ unabhängig ist. Nun folgt

$$\begin{aligned} P((X_{t_1},\ldots,X_{t_n}) \in A \text{ und } B) &= \sum_{n=1}^{\infty} P((X_{t_1},\ldots,X_{t_n}) \in A \text{ und } B \cap \{\tau = s_n\}) \\ &= \sum_{n=1}^{\infty} P((Y_{t_1}^{(n)},\ldots,Y_{t_n}^{(n)}) \in A \text{ und } B \cap \{\tau = s_n\}) \\ &= \sum_{n=1}^{\infty} P((Y_{t_1}^{(n)},\ldots,Y_{t_n}^{(n)}) \in A) P(B \cap \{\tau = s_n\}) \\ &= \sum_{n=1}^{\infty} P((W_{t_1},\ldots,W_{t_n}) \in A) P(B \cap \{\tau = s_n\}) \\ &= P((W_{t_1},\ldots,W_{t_n}) \in A) P(B). \end{aligned}$$

Damit ist iii) gezeigt. Setzt man $B := \Omega$, so folgt auch ii).

Im zweiten Schritt sei τ allgemein. Für $n \in \mathbf{N}_0$ sei $\tau_n := \sum_{k=0}^{\infty} \frac{k+1}{2^n} \mathbf{1}_{\{k2^{-n} \leq \tau < (k+1)2^{-n}\}}$ und $\tau_n(\omega) := \infty$, falls $\tau(\omega) = \infty$ ist. Offensichtlich ist $\tau_n(\Omega)$ abzählbar. Zu zeigen ist also, dass τ_n eine Stoppzeit ist. Für $0 \leq t < 2^{-n}$ gilt $\{\tau_n \leq t\} = \emptyset \in \mathscr{F}_t$. Für $2^{-n} \leq t$ gibt es ein $k_0 \geq 0$ mit $(k_0+1)2^{-n} \leq t < (k_0+2)2^{-n}$. Dann folgt

$$\begin{aligned} \{\tau_n \leq t\} &= \bigcup_{k=0}^{k_0} \left\{\tau_n = \frac{k+1}{2^n}\right\} = \bigcup_{k=0}^{k_0} \underbrace{\{k2^{-n} \leq \tau < (k+1)2^{-n}\}}_{\in \mathscr{F}_{(k+1)2^{-n}}} \\ &\in \mathscr{F}_{(k_0+1)2^{-n}} \subset \mathscr{F}_t. \end{aligned}$$

Ferner ist $\tau_n \geq \tau$ und $\tau_n \searrow \tau$. Für $n \in \mathbf{N}_0$ setzen wir nun $Y_t^{(n)} := W_{t+\tau_n} - W_{\tau_n}$, dann ist $Y_t^{(n)}$ gemäß dem ersten Schritt ein Wiener-Prozess. Da Wiener-Prozesse stetig sind, folgt $Y_t^{(n)} \overset{n\to\infty}{\longrightarrow} X_t$ P-fast sicher. Ferner sei $B \in \mathscr{F}_\tau \subset \mathscr{F}_{\tau_n}$, mit dem ersten Schritt folgt dann $P((Y_{t_1}^{(n)}, \ldots, Y_{t_m}^{(n)}) \in A \text{ und } B) = P((W_{t_1}, \ldots, W_{t_m}) \in A)P(B)$. Da P-fast sichere Konvergenz die Konvergenz in Verteilung impliziert, folgt insgesamt

$$P((X_{t_1}, \ldots, X_{t_m}) \in A)P(B) = P((W_{t_1}, \ldots, W_{t_m}) \in A)P(B).$$

Damit ist wieder iii) gezeigt und ii) folgt wiederum mit $B = \Omega$. $\square$

5.4. Das Reflektionsprinzip

Wir betrachten für $a > 0$ die erste Passierzeit $\tau_a := \inf\{t \geq 0 : W_t = a\}$ und wollen untersuchen, wie die Verteilung oder Verteilungsfunktion von τ_a aussieht. Wegen $P(W_t = a) = 0$ gilt zunächst $P(\tau_a \leq t) = P(\tau_a \leq t, W_t \geq a) + P(\tau_a \leq t, W_t \leq a)$. Ferner gilt $P(\tau_a \leq t, W_t \geq a) = P(W_t \geq a)$, da mit $W_0 = 0$ P-fast sicher und der Stetigkeit der Pfade folgt, dass $\{W_t \geq a\} \subset \{\tau_a \leq t\}$ gilt. Für den zweiten Summanden wollen wir nun eine heuristische Betrachtung durchführen:

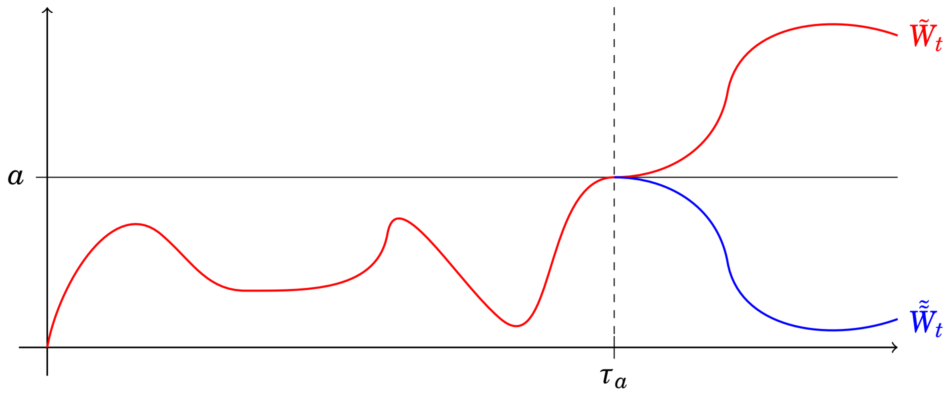

Abbildung 5.1.: Der Prozess wird in τ_a neugestartet und ist dann um a gewissermaßen symmetrisch.

Wir starten den Wiener-Prozess in τ_a neu, nach Satz 5.3.6 ist die Zukunft dann von der Vergangenheit bis τ_a unabhängig. Der neugestartete Wiener-Prozess ist zudem gewissermaßen um die a symmetrisch, d. h. W und $-W$ sind Wiener-Prozesse. Damit ist

$$P(\tau_a \leq t, W_t \leq a) = P(\tau_a \leq t, \tilde{W}_t \leq a) \stackrel{!}{=} P(\tau_a \leq t, \tilde{\tilde{W}}_t \geq a) = P(\tilde{\tilde{W}}_t \geq a) = P(W_t \geq a).$$

Insgesamt gilt also $P(\tau_a \leq t) = 2P(W_t \geq a)$. Die markierte Gleichheit bei $\stackrel{!}{=}$ ist so jedoch nicht begründbar, da Symmetrie eigentlich eine andere Eigenschaft ist. Um dies zu reparieren, benötigen wir den folgenden Satz.

Satz 5.4.1
Mit den obigen Bezeichnungen ist $(\tilde{W}_t)_{t\geq 0}$ definiert durch

$$\tilde{W}_t := \begin{cases} W_t & \text{falls } t \leq \tau_a \\ 2a - W_t & \text{sonst} \end{cases}$$

ebenfalls ein Wiener-Prozess.

Beweis: Wegen Satz 5.3.6 definiert $X_t := W_{t+\tau_a} - W_{\tau_a}$ einen von $\mathscr{F}_{\tau_a}$ unabhängigen Wiener-Prozess. Damit haben $X := (X_t)$, $(-X_t)$ und $W := (W_t)$ die selben Verteilungen und es gilt $X \stackrel{[d]}{=} -X \stackrel{[d]}{=} W$. Ferner ist X von der $\mathscr{F}_{\tau_a}$-messbaren Zufallsvariablen τ_a unabhängig. Wegen

$\tau_a \wedge s \le \tau_a$ für $s \ge 0$ gilt $\mathscr{F}_{\tau_a \wedge s} \subset \mathscr{F}_{\tau_a}$ und damit ist $Y := (W_{\tau_a \wedge s})_{s \ge 0}$ nach Satz 5.3.5 und Lemma 5.3.4 ein $(\mathscr{F}_{\tau_a \wedge s})_{s\ge 0}$-adaptierter Prozess. Wegen $\mathscr{F}_{\tau_a \wedge s} \subset \mathscr{F}_{\tau_a}$ ist dann X von Y unabhängig und da (Y, τ_a) von X bzw. $-X$ unabhängig ist, erhalten wir

$$(Y, \tau_a, X) \overset{[d]}{=} (Y, \tau_a, -X). \tag{*}$$

Wir betrachten nun $C_0 := \{g \in C([0,\infty)),\ g(0) = 0\}$ und $\psi\colon C([0,\infty)) \times [0,\infty) \times C_0 \to C([0,\infty))$ definiert durch

$$\psi(f, t_0, g) := \begin{cases} f(t) & \text{falls } t \le t_0 \\ f(t_0) + g(t - t_0) & \text{falls } t \ge t_0 \end{cases}$$

für alle $t \ge 0$. Mit (*) erhalten wir dann $\psi(Y, \tau_a, X) \overset{[d]}{=} \psi(Y, \tau_a, -X)$. Ferner ist

$$\begin{aligned}\psi(Y, \tau_a, X) &= \begin{cases} W_t & \text{falls } t \le \tau_a \\ W_{\tau_a} + X_{t-\tau_a} & \text{falls } t \ge \tau_a \end{cases} \\ &= (W_t)_{t \ge 0}\end{aligned}$$

und analog

$$\begin{aligned}\psi(Y, \tau_a, -X) &= \begin{cases} W_t & \text{falls } t \le \tau_a \\ W_{\tau_a} - X_{t-\tau_a} & \text{falls } t \ge \tau_a \end{cases} \\ &= W_{\tau_a} - W_{\tau_a + t - \tau_a} + W_{\tau_a} = 2a - W_t \\ &= (\tilde{W}_t)_{t \ge 0}.\end{aligned}$$

Damit besitzen (W_t) und $(\tilde{W}_t)$ die gleichen Verteilungen und daher ist $(\tilde{W}_t)$ ein zentrierter Gauß-Prozess mit der Kovarianzfunktion $\Gamma(s,t) = \min\{s,t\}$, der nach Konstruktion stetige Pfade besitzt, also ist er ein Wiener-Prozess. □

Satz 5.4.2 **Reflektionsprinzip**
Sei (W_t) ein Wiener-Prozess und τ_a die erste Passierzeit für ein $a > 0$. Dann gilt

$$\begin{aligned} P\left(\max_{s\in[0,t]} W_s \ge a\right) &= P(\tau_a \le t) = P(|W_t| \ge a) = 2P(W_t \ge a) \\ &= 1 - \mathscr{N}(0,t)[-a,a]. \end{aligned}$$

Man kann ferner zeigen, dass P-fast sicher $\tau_a < \infty$ gilt, was unmittelbar aus den Aussagen des Satzes folgt. Die Dichte von τ_a lässt sich auch bestimmen, hierzu verweisen wir auf [Meintrup04, Satz 13.9]. Schließlich lässt sich auch $\mathbf{E}\tau_a = \infty$ zeigen.

Satz 5.4.1 wird häufig ebenfalls oder alternativ Reflektionsprinzip genannt.

Beweis: Man kann leicht zeigen, dass $\{\tau_a \le t\} = \left\{\max_{s\in[0,t]} W_s \ge a\right\}$ gilt. Damit reicht es, zu zeigen, dass

$$P\left(\max_{s\in[0,t]} W_s \ge a\right) = 2P(W_t \ge a)$$

gilt. Dazu sei $(\tilde{W}_t)$ der Wiener-Prozess aus Satz 5.4.1. Dann gilt $\{W_t \ge a\} \subset \{\tilde{W}_t \le a\}$ und analog für $>$ und $<$, denn für $W_t \ge a$ gilt $\tau_a \le t$ und damit $\tilde{W}_t = 2a - W_t \le 2a - a = a$. Ferner erhalten wir eine disjunkte Vereinigung

$$\left\{\max_{s\in[0,t]} W_s \ge a\right\} = \{W_t \ge a\} \sqcup \{\tilde{W}_t > a\},$$

dazu sei für „$\subset$" zunächst $s \in [0,t]$ mit $W_s \ge a$, dann gilt entweder $W_t \ge a$ und wir sind fertig oder es gilt $W_t < a$. In diesem Fall folgt mit dem Zwischenwertsatz $\tau_a \le s \le t$ und damit $\tilde{W}_t = 2a - W_t > 2a - a = a$. Für „$\supset$" beobachten wir zunächst, dass $\{W_t \ge a\} \subset \left\{\max_{s\in[0,t]} W_s \ge a\right\}$ offensichtlich ist. Sei nun $\tilde{W}_t > a$, dann folgt mit unseren Vorüberlegungen $W_t < a$. Damit gilt $\tilde{W}_t \ne W_t$ und aus der Definition von $\tilde{W}_t$ folgt daher $t > \tau_a$. Dann existiert ein $s \in [0,t]$ mit $W_s = a$ und daher $\max_{s\in[0,t]} W_s \ge a$.

Damit gilt nun

$$\begin{aligned} P\left(\max_{s\in[0,t]} W_s \ge a\right) &= P(W_t \ge a) + P(\tilde{W}_t > a) = P(W_t \ge a) + P(\tilde{W}_t \ge a) \\ &= 2P(W_t \ge a). \end{aligned}$$

□

5.5. Der Wiener-Prozess als Martingal

In diesem letzten Abschnitt wollen wir eine kurze Einführung in die Theorie der Martingale in stetiger Zeit geben und die Anwendung auf Wiener-Prozesse diskutieren.

Definition 5.5.1 **Martingal**
Ein stochastischer Prozess $(X_t)_{t\geq 0}$, der an $\mathscr{F} = (\mathscr{F}_t)_{t\geq 0}$ adaptiert und integrierbar ist, heißt Martingal genau dann, wenn

$$\mathbf{E}(X_t \mid \mathscr{F}_s) = X_s$$

für alle $0 \leq s \leq t$ gilt. Sub- und Super-Martingale werden ebenfalls analog definiert.

Satz 5.5.2 **Unabhängige Zuwächse**
Sei (X_t) ein integrierbarer, stochastischer Prozess mit unabhängigen Zuwächsen. Dann ist $(X_t - \mathbf{E}X_t)_{t\geq 0}$ ein Martingal bezüglich der natürlichen Filtration.

Insbesondere ist der Wiener-Prozess ein Martingal und $(N_t - \lambda t)_{t\geq 0}$ ist ebenfalls ein Martingal, falls N_t ein homogener Poisson-Prozess mit Rate λ ist.

Beweis: Die Aussage folgt aus Lemma 5.1.8 und einigen elementaren Rechnungen. □

Satz 5.5.3 **Gleichgradig integrierbare Martingale**
Sei (X_t) ein Martingal, dann sind folgende Aussagen äquivalent:

i) Der Prozess (X_t) ist gleichgradig integrierbar.

ii) Es existiert $X_\infty \in \mathscr{L}_1$ mit $X_t = \mathbf{E}(X_\infty \mid \mathscr{F}_t)$ für alle $t \geq 0$.

Beweis: Die Richtung von i) ⇒ ii) ist eine Folgerung aus Satz 2.7.2, die andere Richtung folgt aus Satz 2.6.4. □

Satz 5.5.4 **Optional Sampling Theorem**
Sei $(X_t)_{t\geq 0}$ ein rechtsstetiges Martingal und $\sigma \leq \tau$ Stoppzeiten. Ist τ beschränkt oder (X_t) gleichgradig integrierbar, so sind $X_\sigma, X_\tau \in \mathscr{L}_1$ und es gilt

$$\mathbf{E}(X_\tau \mid \mathscr{F}_\sigma) = X_\sigma.$$

Dabei ist X_σ nach Lemma 5.3.4 und Satz 5.3.5 $\mathscr{F}_\sigma$-messbar.

Beweis: Wir wollen den Beweis an dieser Stelle lediglich skizzieren. Wie im Beweis von Satz 2.4.3 zeigen wir zunächst

$$\mathbf{E}(X_N \mid \mathscr{F}_\tau) = X_\tau$$

für $N \in [0, \infty]$ mit $\tau \leq N$. Dies beweist man in zwei Schritten: Zunächst nimmt man an, dass τ abzählbar viele, aufsteigende Werte besitzt. Dann kann man die Aussage aus Satz 2.4.3 und Satz 2.8.1 folgern. Im zweiten Schritt betrachtet man allgemeine τ und approximiert diese durch $\tau_n \searrow \tau$, wobei die τ_n wie im ersten Schritt gegeben sind. Hierfür ist die Rechtsstetigkeit nötig. Für die Details verweisen wir auf [Meintrup04]. Dort wird auch gezeigt, dass jedes Martingal eine rechtsstetige Version besitzt, wenn $\mathscr{A}$ vervollständigt wird. □

Korollar 5.5.5 **Optional Stopping Theorem (Doob)**
Sei (X_t) ein rechtsstetiges Martingal und τ eine Stoppzeit. Dann ist $(X_{\tau\wedge t})_{t\geq 0}$ ein Martingal. Ist τ beschränkt oder (X_t) gleichgradig integrierbar, so gilt $X_\tau \in \mathscr{L}_1$ und $\mathbf{E}X_\tau = \mathbf{E}X_0$.

Beweis: Die Martingaleigenschaft folgt aus Satz 5.5.4 für die beschränkten Stoppzeiten $\tau \wedge s$ und $\tau \wedge t$. Die zweite Aussage folgt aus Satz 5.5.4 für $\sigma = 0$, denn dann ist $\mathbf{E}(X_\tau \mid \mathscr{F}_0) = X_0$ und damit $\mathbf{E}X_\tau = \mathbf{E}(\mathbf{E}(X_\tau \mid \mathscr{F}_0)) = \mathbf{E}X_0$. □

Satz 5.5.6 **Wiener-Prozesse und Martingale**
Sei (W_t) ein Wiener-Prozess. Dann sind (W_t) und $(W_t^2 - t)$ stetige Martingale.

Es lassen sich auch andere Transformationen betrachten, die (stetige) Martingale ergeben. Hierzu verweisen wir auf [Meintrup04, 13.10 und 13.11].

Beweis: Die Aussagen folgen aus Satz 5.5.2 und einfachem Nachrechnen. □

Korollar 5.5.7 **Ruinwahrscheinlichkeiten des Wiener-Prozesses**
Sei (W_t) ein Wiener-Prozess und $a < 0 < b$. Wir setzen $\tau := \inf\{t \geq 0 : W_t = a \text{ oder } W_t = b\}$. Dann gilt

i) $P(W_\tau = a) = \dfrac{b}{b-a}$.

ii) $\mathbf{E}\tau = -ab$.

Beweis: Der Beweis verläuft weitgehend analog zur symmetrischen Irrfahrt aus Beispiel 2.4.2. Für die Details verweisen wir auf [Meintrup04, S. 381 ff.]. □

Der Wiener-Prozess lässt sich durch „Irrfahrten“ approximieren: Sind (Y_i) i. i. d. Zufallsvariablen mit $\mathbf{E}Y_i = 0$ und $\operatorname{Var} Y_i =: \sigma \in (0,\infty)$, so setzen wir für $t \geq 0$

$$s_t^{(n)} := \sum_{i=1}^{\lfloor nt \rfloor} Y_i \qquad \text{und} \qquad \tilde{s}_t^{(n)} := \frac{1}{\sqrt{\sigma^2 n}} s_t^{(n)}.$$

Das Donskersche Invarianzprinzip besagt dann, dass $\tilde{s}^{(n)}$ gegen einen Wiener-Prozess konvergiert. Wie diese Konvergenz und der Wiener-Prozess aussehen wollen wir hier nicht erläutern und verweisen auf [Klenke06].

Der Wiener-Prozess lässt sich auch als Fourierreihe mit zufälligen Koeffizienten darstellen. Diese Koeffizienten sind unabhängig und normalverteilt. Eine mögliche Orthonormalbasis lässt sich explizit darstellen.

Der Satz vom iterierten Logarithmus besagt

$$\limsup_{t\to\infty} \frac{|W_t|}{\sqrt{2t \log\log t}} = 1 \quad P\text{-fast sicher.}$$

Ferner gibt es ein Gesetz der großen Zahlen, welches

$$\lim_{t\to\infty} \frac{W_t}{t} = 0 = \mathbf{E}W_0 \quad P\text{-fast sicher}$$

besagt. Insbesondere ist $(tW_{\frac{1}{t}})_{t\geq 0}$ wieder ein Wiener-Prozess.

Literaturverzeichnis

[Kallenberg01] O. Kallenberg, *Foundations of Modern Probability*, Springer, 2001

[Klenke06] A. Klenke, *Wahrscheinlichkeitstheorie*, Springer, 2006

[Krylov02] N. V. Krylov, *Introduction to the Theory of Random Processes*, AMS: Graduate Studies in Mathematics, 2002

[Meintrup04] D. Meintrup und S. Schäffler, *Stochastik – Theorie und Anwendungen*, Springer, 2004

[WTSkript11] I. Steinwart, *Wahrscheinlichkeitstheorie*, Mitschrieb der Vorlesung „Wahrscheinlichkeitstheorie“ von Ingo Bürk, Wintersemester 2010/2011

[Kestelman60] Kestelman, *Modern Theory of Integration*, 1960

[Graves56] Graves, *The Theory of Functions of Real Variables*, 1956

Abbildungsverzeichnis

A
Anhang

An dieser Stelle wollen wir einige wichtige Sätze aus anderen Vorlesungen festhalten, wie mehrmals benötigt und referenziert werden. Da diese nur der Übersicht dienen, werden wir auf Beweise dieser Sätze an dieser Stelle verzichten.

Satz A.1 **Totale Wahrscheinlichkeit**
Sei $(\Omega, \mathscr{A}, P)$ ein Wahrscheinlichkeitsraum und $(B_i)_{i\in I}$ eine höchstens abzählbare Zerlegung von Ω. Ferner gelte $P(B_i) > 0$ für alle $i \in I$. Für alle $A \in \mathscr{A}$ gilt dann

$$P(A) = \sum_{i\in I} P(B_i)P(A \mid B_i).$$

Beweis: Der Beweis findet sich in [WTSkript11, Satz I.7.3]. □

Satz A.2 **Beppo Levi I / Monotone Konvergenz**
Es sei $(\Omega, \mathscr{A}, \mu)$ ein Maßraum und $f_n : \Omega \to [0,\infty]$ für $n \geq 1$ eine Folge messbarer Funktionen mit $f_n \nearrow f$. Dann folgt, dass f messbar und nicht-negativ ist. Außerdem gilt

$$\int f \,\mathrm{d}\mu = \lim_{n\to\infty} \int f_n \,\mathrm{d}\mu = \sup_{n\geq 1} \int f_n \,\mathrm{d}\mu.$$

Beweis: Der Beweis findet sich in [WTSkript11, Satz I.11.13]. □

Lemma A.3 **Fatou**
Es sei $(\Omega, \mathscr{A}, \mu)$ ein Maßraum und $f_n : \Omega \to [0,\infty]$ für $n \geq 1$ eine messbare Funktionenfolge. Dann gilt

$$\int \liminf_{n\to\infty} f_n \,\mathrm{d}\mu \leq \liminf_{n\to\infty} \int f_n \,\mathrm{d}\mu.$$

Beweis: Der Beweis findet sich in [WTSkript11, Lemma I.11.14]. □

Satz A.4 **Lebesgue / Dominierte Konvergenz**
Es sei $(\Omega, \mathscr{A}, \mu)$ ein Maßraum, $f_n : \Omega \to \mathbf{R} \cup \{\pm\infty\}$ für $n \geq 1$ eine Folge messbarer Funktionen und $f, g : \Omega \to \mathbf{R} \cup \{\pm\infty\}$ messbare Abbildungen mit $f_n \to f$ und $|f_n| \leq g$ für alle $n \geq 1$. Ist g bezüglich μ integrierbar, so gilt dies auch für f und es folgt

$$\int \lim_{n\to\infty} f_n \,\mathrm{d}\mu = \int f \,\mathrm{d}\mu = \lim_{n\to\infty} \int f_n \,\mathrm{d}\mu.$$

Beweis: Der Beweis findet sich in [WTSkript11, Lemma I.11.16]. □

Lemma A.5 **Borel-Cantelli II**
Es sei $(\Omega, \mathscr{A}, P)$ ein Wahrscheinlichkeitsraum und $(A_n)_{n\geq 1} \subset \mathscr{A}$ unabhängig. Dann gilt

$$\sum_{n=1}^{\infty} P(A_n) = \infty \Longrightarrow P\left(\limsup_{n\to\infty} A_n\right) = 1.$$

Für unabhängige Folgen $(A_n)_{n\geq 1} \subset \mathscr{A}$ gilt damit insbesondere $P(\limsup A_n) \in \{0,1\}$ und $P(\liminf A_n) \in \{0,1\}$.

Beweis: Der Beweis findet sich in [WTSkript11, Lemma II.6.1]. □

Stichwortverzeichnis