Stochastische Prozesse

Die Mitarbeiter von http://mitschriebwiki.nomeata.de/

28. September 2017

Inhaltsverzeichnis

Vorwort

Über dieses Skriptum

Dies ist ein Mitschrieb der Vorlesung „Stochastische Prozesse“ von Prof. Dr. Bäuerle im Sommersemester 08 an der Universität Karlsruhe (TH). Die Mitschriebe der Vorlesung werden mit ausdrücklicher Genehmigung von Prof Dr. Bäuerle hier veröffentlicht, Prof. Dr. Bäuerle ist für den Inhalt nicht verantwortlich.

Wer

Gestartet wurde das Projekt von Joachim Breitner. Weiter haben Felix Wellen und Michael Walter beim Mitschreiben geholfen.

Wo

Alle Kapitel inklusive LaTeX-Quellen können unter http://mitschriebwiki.nomeata.de abgerufen werden. Dort ist ein von Joachim Breitner programmiertes *Wiki*, basierend auf http://latexki.nomeata.de installiert. Das heißt, jeder kann Fehler nachbessern und sich an der Entwicklung beteiligen. Auf Wunsch ist auch ein Zugang über *Subversion* möglich.

I. Markov-Ketten mit diskretem Zeitparameter

1. Elementare Eigenschaften von Markov-Ketten

Gegeben sei eine Folge von Zufallsvariablen (X_n) auf dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ mit $X_n : \Omega \to S$ wobei S nicht leer, und endlich oder abzählbar unendlich ist.

Definition
Eine $S \times S$-Matrix $P = (p_{ij})$ heißt *stochastische Matrix*, falls $p_{ij} \geq 0$ ist und für alle $i \in S$ die Zeilensumme $\sum_{j \in S} p_{ij} = 1$ ist.

Definition
Sei P eine stochastische Matrix. Eine (endliche oder unendliche) Folge $X_0, X_1, X_2, \ldots$ von S-wertigen Zufallsvariablen heißt (homogene[1]) *Markov-Kette* mit Übergangsmatrix P, falls für alle $n \in \mathbb{N}$[2] und für alle Zustände $i_k \in S$ mit

$$P(X_0 = i_0, \ldots, X_n = i_n) > 0$$

gilt

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

Die p_{ij} heißen Übergangswahrscheinlichkeiten und die *Startverteilung* ν der Kette ist definiert durch $\nu(i) \coloneqq P(X_0 = i)$ für $i \in S$.

Bemerkung: Jede Folge von unabhängigen Zufallsvariablen ist eine Markov-Kette.

Satz 1.1 (Eigenschaften von Markov-Ketten)
(X_n) ist genau dann eine Markov-Kette mit Übergangsmatrix P, falls gilt:

$$P(X_k = i_k,\, 0 \leq k \leq n) = P(X_0 = i_0) \prod_{k=0}^{n-1} p_{i_k i_{k+1}} \quad \forall n \in \mathbb{N}_0\ \forall i_k \in S$$

genau dann wenn gilt:

$$P(X_k = i_k,\, 1 \leq k \leq n \mid X_0 = i_0) = \prod_{k=0}^{n-1} p_{i_k i_{k+1}} \quad \forall n \in \mathbb{N}_0\ \forall i_k \in S \text{ mit } P(X_0 = i_0) > 0$$

[1] kurz für zeit-homogen. Die Übergangswahrscheinlichkeiten hängen nicht vom aktuellen Zeitpunkt ab.
[2] Hier ist $\mathbb{N} = 1, 2, \ldots$

genau dann wenn gilt:

$$P(X_k = i_k,\, m \le k \le m+n) = P(X_m = i_m) \prod_{k=m}^{m+n-1} p_{i_k i_{k+1}} \quad \forall m, n \in \mathbb{N}_0 \ \forall i_k \in S$$

Beweis
Zur ersten Äquivalenz. Sei $A_k \coloneqq [X_k = i_k]$, $k \in \mathbb{N}_0$.

„$\Longrightarrow$" Induktion über n: $n = 0$ ✓, $n \curvearrowright n+1$:

$$\begin{aligned} P(A_0 A_1 \dots A_n A_{n+1}) &= P(A_0 \dots A_n) \cdot P(A_{n+1} \mid A_0 \dots A_n) \\ &= P(A_0 \dots A_n) \cdot p_{i_n i_{n+1}} && \text{(Markov-Eigenschaft)} \\ &= P(X_0 = i_0) \prod_{k=0}^{n} p_{i_k i_{k+1}} && \text{(I.V.)} \end{aligned}$$

„$\Longleftarrow$"

$$\begin{aligned} P(A_{n+1} \mid A_0 \dots A_n) &= \frac{P(A_0 \dots A_n A_{n+1})}{P(A_0 \dots A_n)} \\ &= p_{i_n i_{n+1}} && \text{(Vor.)} \end{aligned}$$

Die weiteren Äquivalenzen sind ähnlich zu beweisen. ∎

Konstruktion einer Markov-Kette. Seien (Y_n) Zufallsvariablen, unabhängig und identisch verteilt (u.i.v.), in Z. Weiter ist $g : S \times Z \to S$ eine messbare Abbildung. Definiere die Folge (X_n) mit

$$X_0 = c \in S, \quad X_n = g(X_{n-1}, Y_n).$$

Die so konstruierte Folge (X_n) ist eine Markov-Kette mit Werten in S und Übergangsmatrix $P = (p_{ij})$ mit $p_{ij} = P(g(i, Y_n) = j)$.

Beweis
Die Variablen $X_0, \dots, X_n$ hängen nur von $X_0, Y_1, \dots, Y_n$ ab, sind also unabhängig von Y_{n+1}.

$$\begin{aligned} P(X_{n+1} = i_{n+1} \mid X_k = i_k, 0 \le k \le n) &= \frac{P(X_k = i_k,\, 0 \le k \le n+1)}{P(X_k = i_k,\, 0 \le k \le n)} \\ &= \frac{P(X_k = i_k,\, 0 \le k \le n,\, g(i_n, Y_{n+1}) = i_{n+1})}{P(X_k = i_k,\, 0 \le k \le n)} \\ &= \frac{P(X_k = i_k,\, 0 \le k \le n) \cdot P(g(i_n, Y_{n+1}) = i_{n+1})}{P(X_k = i_k,\, 0 \le k \le n)} \\ &= P(g(i_n, Y_{n+1}) = i_{n+1}) \\ &= \frac{P(g(i_n, Y_{n+1}) = i_{n+1}) \cdot P(X_n = i_n)}{P(X_n = i_n)} \\ &= \frac{P(g(i_n, Y_{n+1}) = i_{n+1}, X_n = i_n)}{P(X_n = i_n)} \\ &= P(g(i_n, Y_{n+1}) = i_{n+1} \mid X_n = i_n) \\ &= P(X_{n+1} = i_{n+1} \mid X_n = i_n) \end{aligned}$$

∎

Bemerkung: Umgekehrt kann zu jeder stochastischen Matrix P eine Markov-Kette (X_n) konstruiert werden mit $X_n = g(X_{n-1}, Y_n)$, wobei (Y_n) u.i.v. und o.B.d.A. $Y_n \sim U[0,1]$.

Beispiel 1.1 (Lagerhaltung)
Sei Y_n die Nachfrage nach einem gelagerten Produkt im Zeitintervall $(n-1, n]$. (Y_n) sei u.i.v. und $Y_n \in \mathbb{N}_0$. Die Auffüll-Politik sei eine (z, Z)-Politik mit $z \leq Z$, $z, Z \in \mathbb{N}$, die wie folgt funktioniert: Falls der Lagerbestand zur Zeit $n \leq z$ ist, dann fülle auf Z auf, sonst tue nichts.

Sei X_n der Lagerbestand zum Zeitpunkt n, $S = \mathbb{N}_0$. Es gilt

$$X_n = \begin{cases} (Z - Y_n)^+, & X_{n-1} \leq z \\ (X_{n-1} - Y_n)^+, & X_{n-1} > z \end{cases}$$

Also ist (X_n) eine Markov-Kette mit Übergangsmatrix $P = (p_{ij})$ und

$$p_{ij} = \begin{cases} P((Z - Y_n)^+ = j), & i \leq z \\ P((i - Y_n)^+ = j), & i > z \end{cases}$$

Beispiel 1.2 (Ruinspiel)
Zwei Spieler mit Startkapital $B \in \mathbb{N}$ Euro spielen in Runden um jeweils einen Euro, etwa mit einem Münzwurf. Spieler I gewinnt dabei mit Wahrscheinlichkeit p. Sei $Y_n = 1$, falls Spieler I die n-te Runde gewinnt, und $Y_n = -1$, falls er die n-Runde verliert. Wir nehmen an, dass Y_n u.i.v. ist.

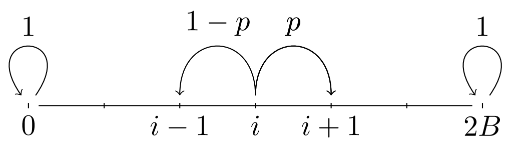
Wir interessieren uns für das Kapital X_n von Spieler I nach der n-ten Runde. Damit ist der Zustandsraum $S = \{0, 1, \ldots, 2B\}$.

Es gilt $X_0 = B$ und

$$X_n = \begin{cases} 2B, & X_{n-1} = 2B \\ X_{n-1} + Y_n, & 0 < X_{n-1} < 2B \\ 0, & X_{n-1} = 0. \end{cases}$$

Es folgt aus der Konstruktion direkt dass (X_n) eine Markov-Kette ist mit Übergangsmatrix $P = (p_{ij})$, und $p_{00} = p_{2B,2B} = 1$ sowie

$$p_{ij} = \begin{cases} p, & j = i+1 \\ 1-p, & j = i-1 \end{cases} \quad \text{für } 0 < i < 2B.$$



Beispiel 1.3 (Wartesystem)
Zu jedem Zeitpunkt $n = 0, 1, \ldots$ können maximal m Kunden bedient werden. Y_n sei die Anzahl der zufällig im Zeitintervall $(n-1, n]$ eintreffenden Kunden und sei u.i.v.

Sei X_n die Anzahl der zur Zeit n wartenden Kunden, $S = \mathbb{N}_0$. Es gilt $X_0 = c$ und $X_n = (X_{n-1} - m)^+ + Y_n$. Also ist (X_n) eine Markov-Kette mit Übergangsmatrix $P = (p_{ij})$ und $p_{ij} = P(Y_n = j - (i-m)^+)$, $i, j \in \mathbb{N}_0$.

Definition
Sei P eine stochastische $S \times S$-Matrix. Dann heißen die Elemente $p_{ij}^{(n)}$ von P^n die n-Schritt-Übergangswahrscheinlichkeiten zu P. Wir definieren $P^0 = E$, also $p_{ij}^{(0)} = \delta_{ij}$.

Satz 1.2
Sei (X_n) eine Markov-Kette mit Übergangsmatrix P. Dann gilt:

a) $P(X_{n+m} = j \mid X_m = i) = p_{ij}^{(n)}$ für alle $i, j \in S$, $m, n \in \mathbb{N}_0$ mit $P(X_m = i) > 0$.

b) $P(X_n = j) = \sum_{i \in S} P(X_0 = i) p_{ij}^{(n)}$, $j \in S$, $n \in \mathbb{N}$.

Beweis
a)

$$\begin{aligned} P(X_{n+m} = i_{n+m}, X_m = i_m) &= \sum_{i_{m+1}, \ldots, i_{n+m-1} \in S} P(X_m = i_m) \prod_{k=m}^{m+n-1} p_{i_k i_{k+1}} \\ &= P(X_m = i_m) p_{i_m i_{m+n}}^{(n)} \end{aligned}$$

b)

$$\begin{aligned} P(X_n = j) &= \sum_{i \in S} P(X_n = j, X_0 = i) \\ &= \sum_{i \in S} P(X_n = j \mid X_0 = i) \cdot P(X_0 = i) \\ &= \sum_{i \in S} P(X_0 = i) p_{ij}^{(n)} \end{aligned}$$

■

Bemerkung: i) Wegen $P^{n+m} = P^n \cdot P^m$ gilt:

$$p_{ij}^{(n+m)} = \sum_{k \in S} p_{ik}^{(n)} p_{kj}^{(m)} \text{ für } i, j \in S$$

Dies ist die „Chapman-Kolmogorov-Gleichung".

ii) Ist $X_0 \sim \nu$, so gilt $X_n \sim \nu \cdot P^n$.

Satz 1.3 (Existenzsatz für Markov-Ketten)
Sei ν ein Wahrscheinlichkeitsmaß auf S und P eine stochastische $S \times S$-Matrix. Sei X_n die n-te Projektion auf $\Omega := S^{\mathbb{N}_0}$, also $X_n : \Omega \to S$, $n \in \mathbb{N}_0$ mit $X_n(\omega) = X_n((i_0, i_1, \ldots)) = i_n$.

Dann existiert ein Wahrscheinlichkeitsmaß P auf $\mathcal{F} = \oplus_{n=0}^{\infty} \mathcal{P}(S)$, sodass (X_n) eine Markov-Kette mit Übergangsmatrix P und Startverteilung ν ist, d.h:

- $P(X_0 = i_0) = \nu(i_0)$, $i_0 \in S$
- $P(X_{n+1} = j \mid X_n = i) = p_{ij}$, $i, j \in S$, $P(X_n = i) > 0$.

Beweis
Satz von Ionescu-Tulcea über die Fortsetzung von Maßen und die Existenz zufälliger Folgen. ■

2. Klassifikation von Zuständen, Rekurrenz und Transienz

In diesem Paragraphen widmen wir uns Fragestellungen wie diesen: Welche Zustände in S werden von der Markov-Kette mit Sicherheit besucht und welche nicht? Wenn sie besucht werden, wie oft?

Definition
Sei (X_n) eine Markov-Kette mit Übergangsmatrix $P = (p_{ij})$.

a) $i \in S$ *führt nach* $j \in S$ (kurz $i \rightsquigarrow j$), falls es ein $n \in \mathbb{N}$ gibt mit $p_{ij}^{(n)} > 0$.

b) $i \in S$ *kommuniziert mit* $j \in S$ (kurz $i \leftrightarrow j$) falls sowohl $i \rightsquigarrow j$ als auch $j \rightsquigarrow i$ gilt.

Bemerkung: Für $i, j \in S$ sei $i \sim j$ definiert als $(i \leftrightarrow j) \vee (i = j)$. Diese Relation ist eine Äquivalenzrelation auf S, da sie reflexiv, symmetrisch und transitiv ist.

Dies liefert uns eine Partition von S mit den Äquivalenzklassen $K(i) := \{j \in S \mid i \sim j\}$. Die Äquivalenzklasse $K(i)$ enthält i selbst und die mit i kommunizierenden Zustände.

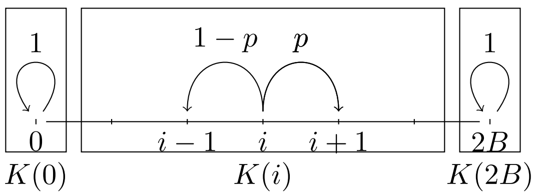
Definition
Sei (X_n) eine Markov-Kette mit Übergangsmatrix $P = (p_{ij})$.

a) $J \subset S$ heißt *abgeschlossen*, wenn es keine zwei Zustände $j \in J$ und $i \in S \setminus J$ gibt mit $j \rightsquigarrow i$.

b) Die Markov-Kette (X_n) beziehungsweise die Übergangsmatrix P heißen *irreduzibel*, falls S nur aus einer Klasse besteht, also für alle $i, j \in S$, $i \neq j$, gilt $i \leftrightarrow j$.

Beispiel 2.1
Skizze, hier ausgelassen

Beispiel 2.2 (Ruinspiel)



Lemma 2.1
$J \subset S$ ist genau dann abgeschlossen, wenn $(p_{ij},\ i, j \in J)$ stochastisch ist.

Beweis
„$\Longrightarrow$“: Klar. „$\Longleftarrow$“: Es gilt: $(p_{ij},\ i, j \in J)$ stochastisch $\iff$ $(p_{ij}^{(n)},\ i, j \in J)$ stochastisch für alle $n \in \mathbb{N}$. ■

Sei (X_n) eine Markov-Kette mit Übergangsmatrix $P = (p_{ij})$. Es sei

$$T_i := \inf\{n \in \mathbb{N} \mid X_n = i\}$$

die (zufällige) Ersteintrittszeit der Markov-Kette in den Zustand i.

Wir setzen dabei $\inf \emptyset := \infty$. Weiter sei für $i, j \in S$, $n \in \mathbb{N}$:

$$\begin{aligned} f_{ij}^{(n)} &:= P(T_j = n \mid X_0 = i) = P_i(T_j = n) \\ &= P(X_n = j,\, X_\nu \neq j \text{ für } 1 \leq \nu < n \mid X_0 = i) \\ f_{ij}^{(0)} &:= 0 \end{aligned}$$

Offenbar ist $f_{ij}^{(1)} = p_{ij}$. Weiter definieren wir

$$f_{ij}^* := \sum_{n=0}^{\infty} f_{ij}^{(n)} = \sum_{n=0}^{\infty} P_i(T_j = n) = P_i(T_j < \infty) = P_i(\exists n \in \mathbb{N} : X_n = j) \in [0,1]$$

Definition
Ein Zustand $i \in S$ heißt *rekurrent*, falls $f_{ii}^* = 1$ und *transient* sonst.

Lemma 2.2
Für alle $n \in \mathbb{N}$, $i, j \in S$ gilt:

$$p_{ij}^{(n)} = \sum_{k=1}^{n} f_{ij}^{(k)} p_{jj}^{(n-k)}$$

Beweis (Methode des ersten Besuches)
Unter Verwendung der Formel $P(AB \mid C) = P(B \mid C) \cdot P(A \mid BC)$ für Ereignisse A, B, C zeigen wir:

$$\begin{aligned} p_{ij}^{(n)} &= P_i(X_n = j) = \sum_{k=1}^{n} P(X_n = j,\, X_\mu \neq j,\, 1 \leq \mu < k,\, X_k = j \mid X_0 = i) \\ &= \sum_{k=1}^{n} P_i(X_\mu \neq j,\, 1 \leq \mu < k,\, X_k = j) \cdot \\ &\qquad \underbrace{P(X_n = j \mid X_0 = i,\, X_\mu \neq j,\, 1 \leq \mu < k,\, X_k = j)}_{=P(X_n = j \mid X_k = j)} \\ &= \sum_{k=1}^{n} f_{ij}^{(k)} p_{jj}^{(n-k)} \end{aligned}$$

■

Satz 2.3
$i \in S$ ist rekurrent genau dann, wenn gilt:

$$\sum_{n=0}^{\infty} p_{ii}^{(n)} = \infty$$

Beweis
Für $s \in (0,1)$ erhalten wir aus Lemma 2.2:

$$\begin{aligned}\sum_{n=0}^{\infty} p_{ii}^{(n)} s^n &= 1 + \sum_{n=1}^{\infty} p_{ii}^{(n)} s^n = \\ &= 1 + \sum_{n=1}^{\infty} s^n \sum_{k=1}^{n} f_{ii}^{(k)} p_{ii}^{(n-k)} \\ &= 1 + \sum_{k=1}^{\infty} f_{ii}^{(k)} s^k \sum_{n=k}^{\infty} p_{ii}^{(n-k)} s^{n-k} \\ &= 1 + \sum_{k=1}^{\infty} f_{ii}^{(k)} s^k \sum_{n=0}^{\infty} p_{ii}^{(n)} s^n\end{aligned}$$

Abkürzend schreiben wir $F(s) := \sum_{k=1}^{\infty} f_{ii}^{(k)} s^k$ und $P(s) = \sum_{n=0}^{\infty} p_{ii}^{(n)} s^n$, also gilt

$$P(s) = 1 + F(s) \cdot P(s).$$

Nun sei $s \to 1$ (monotone Konvergenz!), und wir erhalten

$$P(1) = 1 + f_{ii}^* \cdot P(1).$$

Es folgt: Ist $f_{ii}^* = 1$, so gilt $P(1) = 1 + P(1)$, also ist $P(1) = \sum_{n=0}^{\infty} p_{ii}^{(n)} = \infty$. Ist ansonsten $f_{ii}^* < 1$, so gilt $P(1) = \frac{1}{1-f_{ii}^*} < \infty$. ∎

Bemerkung: Die im Satz 2.3 auftretende Reihe kann wie folgt interpretiert werden:

$$\sum_{n=0}^{\infty} p_{ii}^{(n)} = \sum_{n=0}^{\infty} E_i[1_{[X_n=i]}] = E_i(\sum_{n=0}^{\infty} 1_{[X_n=i]})$$

Sie bezeichnet also die erwartete Anzahl der Besuche des Zustandes $i \in S$.

Satz 2.4 (Solidaritätsprinzip)
Ist ein Zustand $i \in S$ rekurrent (bzw. transient), so ist jeder Zustand in $K(i)$ rekurrent (bzw. transient).

Beweis
Sei i rekurrent und $j \in K(i)$, $j \neq i$, das heißt es gibt $n, m \in \mathbb{N}$, sodass $p_{ij}^{(m)} \cdot p_{ji}^{(n)} > 0$. Mit der Abschätzung

$$\sum_{k=0}^{\infty} p_{jj}^{(k)} \geq \sum_{k=0}^{\infty} p_{jj}^{(m+n+k)} \geq \sum_{k=0}^{\infty} p_{ji}^{(n)} p_{ii}^{(k)} p_{ij}^{(m)} = p_{ij}^{(m)} p_{ji}^{(n)} \sum_{k=0}^{\infty} p_{ii}^{(k)}$$

und Satz 2.3 ist $\sum_{k=0}^{\infty} p_{jj}^{(k)} = \infty$ und j rekurrent. ∎

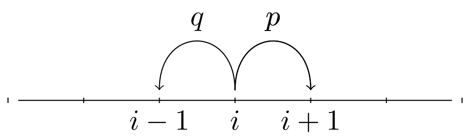
Bemerkung: Ist $i \in S$ rekurrent (bzw. transient), so sagen wir $K(i)$ ist rekurrent (bzw. transient).

Ist (X_n) irreduzibel und ein $i \in S$ ist rekurrent (bzw. transient), so sagen wir (X_n) ist rekurrent (bzw. transient).

Beispiel 2.3 (Irrfahrt auf den ganzen Zahlen, „Random Walk“)
Es sei $X_n = \sum_{k=1}^{n} Y_k = X_{n-1} + Y_n$ und $X_0 = 0$, wobei (Y_n) u.i.v. mit

$$P(Y_n = 1) = p = 1 - P(Y_n = -1) = 1 - q, \quad p \in (0,1)$$

ist ($S = \mathbb{Z}$).



(X_n) ist nach Konstruktion eine irreduzible Markov-Kette. Ist (X_n) rekurrent oder transient?

Wir wenden Satz 2.3 an und untersuchen o.B.d.A. $i = 0$. Es gilt für alle $n \in \mathbb{N}_0$:

$$\begin{aligned} p_{00}^{(2n+1)} &= 0 \\ p_{00}^{(2n)} &= p^n q^n \binom{2n}{n} = p^n q^n \frac{(2n)!}{(n!)^2} \end{aligned}$$

Mit der Stirling-Formel ($n! \simeq (\frac{n}{e})^n \cdot \sqrt{2\pi n}$) erhält man dann

$$p_{00}^{(2n)} \approx (pq)^n \cdot \frac{(\frac{2n}{e})^{2n}\sqrt{4\pi n}}{(\frac{n}{e})^{2n} 2\pi n} = \frac{(4pq)^n}{\sqrt{\pi n}}.$$

Fall 1 Ist $p = q = \frac{1}{2}$, so ist $p_{00}^{(2n)} \approx \frac{1}{\sqrt{\pi n}}$, also ist $\sum_{n=0}^{\infty} p_{00}^{(2n)} = \infty$ und die Markov-Kette ist rekurrent.

Fall 2 Ist dagegen $p \neq q$, also $pq < \frac{1}{4}$, so ist $\sum_{n=0}^{\infty} p_{00}^{(2n)} \leq c \cdot \sum_{n=0}^{\infty} (4pq)^n < \infty$, also ist die Markov-Kette transient.

Bemerkung: Betrachtet man die symmetrische Irrfahrt auf $\mathbb{Z}^d$, also $p_{ij} = \frac{1}{2d}$ für $\|i - j\| = 1$, mit $\|\cdot\|$ der l^1-Norm und $i, j \in \mathbb{Z}^d$, so ist die Irrfahrt rekurrent für $d = 1, 2$ und transient sonst.

Lemma 2.5
Liegen i und j in der selben rekurrenten Klasse, so gilt $f_{ij}^* = f_{ji}^* = 1$.

Lemma 2.6
Für alle $i, j \in S$ gilt: Wenn j transient ist, dann gilt

$$\sum_{n=0}^{\infty} p_{ij}^{(n)} < \infty$$

Insbesondere ist

$$\lim_{n \to \infty} p_{ij}^{(n)} = 0.$$

Beweis
Summiere die Gleichung in Lemma 2.2 über alle n:

$$\begin{aligned}\sum_{n=0}^{\infty} p_{ij}^{(n)} &= \delta_{ij} + \sum_{n=1}^{\infty}\sum_{k=1}^{n} f_{ij}^{(k)} p_{jj}^{(n-k)} \\ &= \delta_{ij} + \sum_{k=1}^{\infty} f_{ij}^{(k)} \sum_{n=k}^{\infty} p_{jj}^{(n-k)} \\ &= \delta_{ij} + f_{ij}^{*} \underbrace{\sum_{n=0}^{\infty} p_{jj}^{(n)}}_{<\infty \text{ da } j \text{ transient}} < \infty\end{aligned}$$

■

Satz 2.7
Ist eine Klasse $K \subseteq S$ rekurrent, so ist K abgeschlossen bzw. $(p_{ij},\ i,j \in K)$ ist stochastisch.

Beweis
Wir zeigen: Ist $i \in K$ rekurrent und $i \rightsquigarrow j$, $i \neq j$, dann gilt $j \rightsquigarrow i$ und damit $j \in K$.

Angenommen, $j \rightsquigarrow i$ gelte nicht, also $p_{ji}^{(n)} = 0$ für alle $n \in \mathbb{N}$, und sei $N \in \mathbb{N}$ die kleinste Zahl mit $p_{ij}^{(N)} > 0$. Es gilt nun für alle $n \in \mathbb{N}$

$$P_i(X_N = j, X_n = i) = 0.$$

Denn für $n > N$ gilt: $P_i(X_N = j, X_n = i) = p_{ij}^{(N)} p_{ji}^{(n-N)} = 0$ und für $n < N$ gilt: $P_i(X_N = j, X_n = i) = p_{ii}^{(n)} p_{ij}^{(N-n)} = 0$, da $N - n < N$.

Also ist $P_i(T_i \leq m, X_N = j) = \sum_{n=1}^{m} P_i(T_i = n, X_N = j) \leq \sum_{n=1}^{m} P_i(X_n = i, X_N = j) = 0$ und damit

$$\begin{aligned}\sum_{n=1}^{m} f_{ii}^{(n)} &= P_i(T_i \leq m) \\ &= P_i(T_i \leq m, X_N \neq j) \\ &\leq P_i(X_N \neq j) \\ &= 1 - P_i(X_N = j) = 1 - p_{ij}^{(N)}.\end{aligned}$$

Für $m \to \infty$ folgt dann

$$1 = f_{ii}^{*} = \sum_{n=1}^{\infty} f_{ii}^{(n)} \leq 1 - p_{ij}^{(N)} < 1,$$

was ein Widerspruch ist. ■

Satz 2.8
Ist die Klasse K endlich und abgeschlossen, so ist K rekurrent.

Beweis
Da $(p_{ij},\, i,j \in K)$ stochastisch ist, folgt induktiv, dass $(p_{ij}^{(n)},\, i,j \in K)$ stochastisch für alle $n \in \mathbb{N}$ ist. Angenommen, K wäre transient. Sei dann $j \in K$, dann ist nach Lemma 2.6 $p_{ij}^{(n)} \to 0$ für $n \to \infty$ und alle $i \in S$. Für $i \in K$ folgt also: $1 = \sum_{j \in K} p_{ij}^{(n)} \to 0$ für $n \to \infty$. Widerspruch. ■

Bemerkung: Insbesondere gilt: Ist S endlich und P irreduzibel, so ist die Markov-Kette rekurrent.

Beispiel 2.4 (Irrfahrt mit reflektierenden Grenzen)
Die Irrfahrt ist irreduzibel und rekurrent nach Satz 2.8.

Absorbtionswahrscheinlichkeiten

Sei $(X_n)_{n\in\mathbb{N}}$ eine Markov-Kette mit Zustandsraum S und Übergangsmatrix $P = (p_{ij})$. Aufgrund der bisherigen Ergebnisse können wir S zerlegen in rekurrente Klassen $K_1, K_2, \ldots$ und eine Menge von transienten Zuständen T, also $S = T \cup K_1 \cup K_2 \cup \ldots$.

Es sei $\tau \coloneqq \inf\{n \geq 0 \mid X_n \notin T\}$ die Austrittszeit aus der Menge der transienten Zustände. Für $i \in T$, $k \in T^c$ interessiert uns $u_{ik} = P_i(X_\tau = k)$, vorausgesetzt $P_i(\tau < \infty) = 1$.

Wir unterteilen $P = (p_{ij})$ in

$$P = \begin{pmatrix} Q & R \\ 0 & \tilde{P} \end{pmatrix}$$

wobei Q die Einschränkung von P auf die transienten Zustände ist, also $Q = (q_{ij}) = (p_{ij}, i, j \in T)$.

Satz 2.9
Für $i \in T$, $j \in T^c$ gilt:

$$u_{ij} = \sum_{k \in T} q_{ik} u_{kj} + p_{ij}.$$

Beweis
Sei $i \in T$, $j \in T^c$.

$$\begin{aligned}
u_{ij} &= P_i(X_\tau = j) \\
&= \sum_{k\in S} P_i(X_\tau = j, X_1 = k) \\
&= \underbrace{\sum_{k\in T} P_i(X_\tau = j, X_1 = k)}_{=:A} + \underbrace{\sum_{k\in T^c} P_i(X_\tau = j, X_1 = k)}_{=p_{ij}} \\
A &= \sum_{k\in T}\sum_{n\geq 2} P_i(\tau = n, X_n = j, X_1 = k) \\
&= \sum_{k\in T}\sum_{n\geq 2} P_i(X_2 \in T, \dots, X_{n-1} \in T, X_n = j, X_1 = k) \\
&= \sum_{k\in T}\sum_{n\geq 2} P_i(X_2 \in T, \dots, X_{n-1} \in T, X_n = j \mid X_1 = k) \cdot P_i(X_1 = k) \\
&= \sum_{k\in T}\sum_{n\geq 2} p_{ik} P_k(X_1 \in T, \dots, X_{n-2} \in T, X_{n-1} = j) \\
&= \sum_{k\in T}\sum_{n\geq 2} p_{ik} P_k(\tau = n-1, X_\tau = j) \\
&= \sum_{k\in T} p_{ik} P_k(X_\tau = j) \\
&= \sum_{k\in T} p_{ik} u_{kj}
\end{aligned}$$

Da für $i, k \in T$ gilt: $p_{ik} = q_{ik}$, folgt die Behauptung. ■

Bemerkung: Es sei $U = (u_{ij})_{i\in T, j\in T^c}$. Dann lässt sich Satz 2.9 schrieben als $U = QU + R$ bzw. $U - QU = R$, also $(I - Q)U = R$. Falls $I - Q$ invertierbar ist, ist $U = (I - Q)^{-1}R$

3. Stationäre Verteilungen

Sei (X_n) eine Markov-Kette mit Übergangsmatrix P und Startverteilung ν.

Dann ist $X_n \sim \nu \cdot P^n$. Im Allgemeinen hängt diese Verteilung von n ab. Es gibt aber spezielle Verteilungen ν, sodass die mit dieser Verteilung gestartete Kette zu jedem Zeitpunkt n die selbe Verteilung ν besitzt. Man sagt dann, die Kette ist *im Gleichgewicht* bzw. *stationär*.

Definition
Eine Abbildung $\nu : S \to \mathbb{R}_{\geq 0}$ heißt *Maß*.

NB: Ein Maß ν definiert ein Maß $\mathcal{P}(S) \to \mathbb{R}_{\geq 0}, A \mapsto \sum_{a\in A} \nu(a)$ im gewöhnlichen Sinne. Falls $\sum_{a\in S} \nu(a) = 1$, so definiert es sogar eine Verteilung.

Definition
Sei (X_n) eine Markov-Kette mit Übergangsmatrix $P = (p_{ij})$. Ein Maß ν heißt *invariant* für P, falls $\nu \cdot P = \nu$, d.h. falls gilt:

$$\sum_{i\in S} \nu(i) \cdot p_{ij} = \nu(j).$$

Ist ν eine Verteilung und invariant, so nennt man ν auch *stationäre Verteilung* oder *Gleichgewichtsverteilung.*

Bemerkung: a) Ist S endlich, so kann jedes (nichtdegenerierte) invariante Maß zu einer stationären Verteilung normiert werden.

b) Ist ν invariant, so gilt $\nu \cdot P^n = \nu$ $(n \in \mathbb{N}_0)$.

Ist ν eine stationäre Verteilung, so gilt

$$P_\nu(X_n = j) = \sum_{i \in S} \nu(i) \cdot p_{ij}^{(n)} = \nu(j),$$

d.h. die mit ν gestartete Kette hat zu jedem Zeitpunkt die Verteilung ν.

c) Ist P irreduzibel und $\nu \neq 0$ ein invariantes Maß, so ist $\nu(j) > 0$ für jedes $j \in S$.

Denn: $\nu \neq 0$, also existiert $i_0 \in S$ mit $\nu(i_0) > 0$. Wegen der Irreduzibilität gibt es ferner für jedes $j \in S$ ein $n \in \mathbb{N}_0$ mit $p_{i_0 j}^{(n)} > 0$. Zusammen:

$$\nu(j) = \sum_{i \in S} \nu(i) \cdot p_{ij}^{(n)} \geq \nu(i_0) p_{i_0 j}^{(n)} > 0$$

Gibt es immer ein invariantes Maß bzw. eine stationäre Verteilung? Ist es eindeutig?

Im Folgenden sei P irreduzibel. Wir definieren für ein beliebiges $k \in S$ das Maß γ_k wie folgt:

$$\gamma_k(i) := E_k[\sum_{n=1}^{T_k} 1_{[X_n = i]}]$$

Satz 3.1
Sei (X_n) irreduzibel und rekurrent, $k \in S$. Dann gilt:

a) γ_k ist ein invariantes Maß

b) $0 < \gamma_k < \infty$

c) γ_k ist das einzige invariante Maß mit $\gamma_k(k) = 1$ (d.h. γ_k ist eindeutig bis auf Vielfache).

Beweis
a) Zunächst gilt:

$$\begin{aligned}\gamma_k(i) &= E_k[\sum_{n=1}^{T_k} 1_{[X_n = i]}] = E_k[\sum_{n=1}^{\infty} 1_{[X_n = i, n \leq T_k]}] = \sum_{n=1}^{\infty} P_k(X_n = i, n \leq T_k) \\ &= \sum_{n=1}^{\infty} \sum_{j \in S, j \neq k} P_k(X_n = i, X_{n-1} = j, n \leq T_k)\end{aligned}$$

Für $j \neq k$ erhält man

$$\begin{aligned}
&P_k(X_n = i, X_{n-1} = j, n \leq T_k) \\
=&P(X_n = i, X_{n-1} = j, X_{n-2} \neq k, \dots, X_1 \neq k, X_0 = k)/P(X_0 = k) \\
=&P(X_n = i \mid X_{n-1} = j, X_{n-2} \neq k, \dots, X_1 \neq k, X_0 = k) \cdot P_k(X_{n-1} = j, X_{n-2} \neq k, \dots, X_1 \neq k) \\
=&p_{ji} \cdot P_k(X_{n-1} = j, n \leq T_k)
\end{aligned}$$

Die so erhaltene Identität gilt auch für $j = k$ und es folgt

$$\begin{aligned}
\gamma_k(i) &= \sum_{n=1}^{\infty} \sum_{j \in S} P_k(X_n = i, X_{n-1} = j, n \leq T_k) \\
&= \sum_{n=1}^{\infty} \sum_{j \in S} p_{ji} \cdot P_k(X_{n-1} = j, n \leq T_k) \\
&= \sum_{j \in S} p_{ji} \sum_{n=0}^{\infty} P_k(X_n = j, n+1 \leq T_k) \\
&= \sum_{j \in S} p_{ji} E_k[\sum_{n=0}^{T_k - 1} 1_{[X_n = j]}] \\
&= \sum_{j \in S} p_{ji} E_k[\sum_{n=1}^{T_k} 1_{[X_n = j]}] = \sum_{j \in S} \gamma_k(j) p_{ji}
\end{aligned}$$

b) Nach Bemerkung c) oben gilt $\gamma_k > 0$, denn $\gamma_k(k) = 1$. Wegen $\gamma_k = \gamma_k \cdot P^n$ folgt

$$1 = \gamma_k(k) \geq \gamma_k(j) \cdot p_{jk}^{(n)}$$

für jedes $j \in S$. Es gibt allerdings mindestens ein $p_{jk}^{(n)} > 0$, denn die Markov-Kette ist irreduzibel; daran erkennt man $\gamma_k < \infty$.

c) Sei λ ein weiteres, invariantes Maß mit $\lambda(k) = 1$. Es gilt also:

$$\begin{aligned}
\lambda(j) &= \sum_{i \in S \setminus \{k\}} \lambda(i) \cdot p_{ij} + 1 \cdot p_{kj} \\
&= \sum_{i \in S \setminus \{k\}} \left(\sum_{l \in S \setminus \{k\}} \lambda(l) \cdot p_{li} + p_{ki} \right) p_{ij} + p_{kj} \\
&= \sum_{i,l \in S \setminus \{k\}} \lambda(l) \cdot p_{li} \cdot p_{ij} + \sum_{i \in S \setminus \{k\}} p_{ki} \cdot p_{ij} + p_{kj} \\
&= \sum_{i,l \in S \setminus \{k\}} \lambda(l) \cdot p_{li} \cdot p_{ij} + P_k(X_2 = j, T_k \geq 2) + P_k(X_1 = j, T_k \geq 1)
\end{aligned}$$

Iterativ erhalten wir für $n \in \mathbb{N}$:

$$\begin{aligned}
\lambda(j) &= \sum_{i_0, i_1, \dots, i_n \in S \setminus \{k\}} \lambda(i_n) \prod_{r=1}^{n} p_{i_r i_{r-1}} p_{i_0 j} + \sum_{r=1}^{n+1} P_k(X_r = j, T_k \geq r) \\
&\geq \sum_{r=1}^{n+1} P_k(X_r = j, T_k \geq r) = E_k[\sum_{r=1}^{\min(n+1, T_k)} 1_{[X_r = j]}] \xrightarrow{n \to \infty} \gamma_k(j).
\end{aligned}$$

Es ist also $\lambda - \gamma_k$ ebenfalls ein invariantes Maß mit $(\lambda - \gamma_k)(k) = 0$; nach Bemerkung c) folgt $\lambda - \gamma_k = 0$, d.h. $\lambda = \gamma_k$. ■

Bemerkung: a) Ist S endlich und P irreduzibel, so folgt aus Satz 3.1, dass eine stationäre Verteilung existiert.

b) Ist (X_n) irreduzibel und transient, so kann keine stationäre Verteilung existieren, denn:

Angenommen es existiert eine stationäre Verteilung π. Dann ist

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

Für $n \to \infty$ wird daraus (mit majorisierter Konvergenz und der bekannten Eigenschaft transienter Übergangswahrscheinlichkeiten):

$$\begin{aligned}\pi(j) &= \sum_{i \in S} \pi(i) \lim_{n \to \infty} p_{ij}^{(n)} \\ &= \sum_{i \in S} \pi(i) \cdot 0 \\ &= 0\end{aligned}$$

Definition
Für $i \in S$ sei

$$\begin{aligned}m_i := E_i[T_i] &= \sum_{n=1}^{\infty} n \cdot P_i(T_i = n) + \infty \cdot (1 - f_{ii}^*) \\ &= \sum_{n=1}^{\infty} n \cdot f_{ii}^{(n)} + \infty \cdot (1 - f_{ii}^*)\end{aligned}$$

die *mittlere Rückkehrzeit* des Zustands i.

Bemerkung: Ist j transient, so folgt $m_j = \infty$.

Definition
Ein Zustand $i \in S$ heißt *positiv rekurrent*, falls $m_i < \infty$ und *nullrekurrent*, falls i rekurrent und $m_i = \infty$ ist.

Bemerkung: Jeder positiv rekurrente Zustand ist auch rekurrent.

Satz 3.2
Sei (X_n) eine irreduzible Markov-Kette. Die folgenden Aussagen sind äquivalent:

i) Es existiert eine stationäre Verteilung.

ii) Es gibt einen positiv rekurrenten Zustand $i \in S$.

iii) Alle Zustände in S sind positiv rekurrent.

Sind diese Bedingungen erfüllt, so ist die stationäre Verteilung eindeutig und durch

$$\pi(i) = \frac{1}{m_i}$$

gegeben.

Beweis

(iii) $\implies$ (ii) ✓

(ii) $\implies$ (i) Sei $k \in S$ mit $m_k < \infty$, dann ist (X_n) rekurrent und mit Satz 3.1 ist γ_k ein invariantes Maß. Dann gilt:

$$\begin{aligned}\sum_{j\in S}\gamma_k(j) &= \sum_{j\in S} E_k[\sum_{n=1}^{T_k} 1_{[X_n=j]}] \\ &= E_k[\sum_{n=1}^{T_k}\sum_{j\in S} 1_{[X_n=j]}] \\ &= E_k[\sum_{n=1}^{T_k} 1] \\ &= E_k[T_k] = m_k < \infty\end{aligned}$$

also ist γ_k normierbar.

(i) $\implies$ (iii) Sei π eine stationäre Verteilung und $k \in S$. Insbesondere sind $\pi(j) > 0$ für alle $j \in S$. Dann ist $\gamma := \frac{\pi}{\pi(k)}$ ein invariantes Maß mit $\gamma(k) = 1$. Nach Satz 3.1 c) folgt $\gamma = \gamma_k$. Beachte dass im Beweis von 3.1 c) die Voraussetzung (X_n) rekurrent nicht verwendet wurde. Wie oben ist

$$m_k = \sum_{j\in S}\gamma_k(j) = \frac{1}{\pi(k)}\sum_{j\in S}\pi(j) = \frac{1}{\pi(k)} < \infty.$$

Da $k \in S$ beliebig ist, folgt die Behauptung (iii).

Außerdem ist gezeigt, dass $\pi(k) = \frac{1}{m_k}$. ■

Bemerkung: i) Es gilt also die folgende Trichotomie für irreduzible Markov-Ketten, das heißt eine irreduzible Markov-Kette gehört immer zu genau einem der folgenden Fälle:

- Die Markov-Kette ist transient, es gibt keine stationäre Verteilung.
- Die Markov-Kette ist nullrekurrent, insbesondere gilt für alle $i, j \in S$:

  $$P_i(T_j < \infty) = 1 \text{ und } E_i[T_j] = \infty$$

  und es gibt ein (bis auf Vielfache) eindeutiges invariantes Maß, aber keine stationäre Verteilung.
- Die Markov-Kette ist positiv rekurrent, für alle $i, j \in S$ ist $E_i[T_j] < \infty$ und es gibt eine stationäre Verteilung.

ii) Ist S endlich und die Markov-Kette irreduzibel, so ist sie automatisch positiv rekurrent.

iii) Ist π eine stationäre Verteilung, so gilt:

$$\pi(i) = \frac{\gamma_k(i)}{\sum_{j\in S}\gamma_k(j)} = \frac{E_k[\sum_{n=1}^{T_k} 1_{[X_n=i]}]}{E_k[T_k]}.$$

$\pi(i)$ ist also der durchschnittliche Bruchteil der Zeit, den die Markov-Kette im Zustand i verbringt, während sie einen Zyklus durchläuft.

Beispiel 3.1
Sei $S = \{1,2\}$. Die Übergangsmatrix P sei gegeben durch

$$P = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}, \quad \alpha, \beta \in (0,1].$$

Also ist die Markov-Kette irreduzibel und positiv rekurrent. Die stationäre Verteilung ist die Lösung des linearen Gleichungssystems

$$\pi P = \pi$$

unter Berücksichtigung von $\pi \geq 0$ und $\pi(1) + \pi(2) = 1$, also

$$\pi(1) = \frac{\beta}{\alpha+\beta}, \quad \pi(2) = \frac{\alpha}{\alpha+\beta}.$$

Beispiel 3.2 (Irrfahrt)
Siehe Beispiel 2.3: Für $p \neq q$ ist die Markov-Kette transient. Existiert ein invariantes Maß? Ansatz:

$$\gamma(j) = \sum_{i\in\mathbb{Z}} \gamma(i) p_{ij} = \gamma(j+1)\cdot q + \gamma(j-1)\cdot p$$

$$\implies \gamma(j+1) - \gamma(j) = \frac{p}{q}(\gamma(j) - \gamma(j-1))$$

Also: $\gamma_1(j) = 1$ und $\gamma_2(j) = (\frac{p}{q})^j$, $j \in \mathbb{Z}$, sind verschiedene invariante Maße.

Ist $p = q = \frac{1}{2}$, so ist die Markov-Kette rekurrent. Ist sie nullrekurrent oder positiv rekurrent? Es gibt keine stationäre Verteilung (siehe oben), die Markov-Kette ist also nullrekurrent.

Beispiel 3.3 (Geburts- und Todesprozess in diskreter Zeit)
Es sei (X_n) eine Markov-Kette mit $S = \mathbb{N}_0$ und Übergangsmatrix

$$P = \begin{pmatrix} p_{00} & p_{01} & 0 & \cdots & \\ p_{10} & p_{11} & p_{12} & 0 & \cdots \\ 0 & p_{21} & p_{22} & p_{23} & \ddots \\ \vdots & 0 & p_{32} & p_{33} & \ddots \\ & \vdots & \ddots & \ddots & \ddots \end{pmatrix}$$

mit $p_{01} > 0$, $p_{i,i+1} > 0$, $p_{i,i-1} > 0$ für alle $i \geq 1$, also ist (X_n) irreduzibel. Wann ist (X_n) positiv rekurrent?

Der Ansatz $\pi P = \pi$ liefert:

$$\begin{aligned}
\pi(0) &= p_{00}\cdot\pi(0) + p_{10}\cdot\pi(1) \\
\pi(i) &= p_{i-1,i}\cdot\pi(i-1) + p_{ii}\cdot\pi(i) + p_{i+1,i}\cdot\pi(i+1) \\
\iff p_{i,i-1}\cdot\pi(i) + p_{i,i+1}\pi(i) &= p_{i-1,i}\cdot\pi(i-1) + p_{i+1,i}\cdot\pi(i+1) \\
\iff p_{i+1,i}\cdot\pi(i+1) - p_{i,i+1}\cdot\pi(i) &= p_{i,i-1}\cdot\pi(i) - p_{i-1,i}\cdot\pi(i-1) \\
&= \ldots = p_{10}\cdot\pi(1) - p_{01}\cdot\pi(0)
\end{aligned}$$

Aus der ersten Gleichung folgt $p_{01} \cdot \pi(0) = p_{10} \cdot \pi(1)$ und damit:

$$\begin{aligned} p_{i+1,i} \cdot \pi(i+1) - p_{i,i+1} \cdot \pi(i) = 0 \\ \implies \pi(i+1) &= \frac{p_{i,i+1}}{p_{i+1,i}} \pi(i) \\ &= \ldots = \pi(0) \cdot \prod_{k=0}^{i} \frac{p_{k,k+1}}{p_{k+1,k}}. \end{aligned}$$

Für $\pi(0) > 0$ erhält man ein invariantes Maß. (X_n) ist positiv rekurrent, falls

$$\sum_{i=1}^{\infty} \prod_{k=0}^{i} \frac{p_{k,k+1}}{p_{k+1,k}} < \infty.$$

Im Spezialfall $p_{k,k+1} = p$, $p_{k,k-1} = q = 1 - p$, $k \geq 1$ und $p_{01} = p$, $p_{00} = 1 - p$ gilt

$$(X_n) \text{ ist positiv rekurrent } \iff \sum_{i=1}^{\infty} \left(\frac{p}{q}\right)^{i+1} < \infty \iff p < q.$$

4. Konvergenz gegen die stationäre Verteilung

Sei (X_n) eine Markov-Kette mit Übergangsmatrix P. Wir nehmen an, dass (X_n) bzw. P aperiodisch ist, das heißt: Für alle Zustände $i \in S$ gilt:

$$d_i := \operatorname{ggT}\{n \in \mathbb{N} \mid p_{ii}^{(n)} > 0\} = 1$$

Lemma 4.1
P ist genau dann irreduzibel und aperiodisch, wenn für alle $i, j \in S$ eine Zahl $n_0 \in \mathbb{N}$ existiert, sodass für alle $n \geq n_0$ gilt: $p_{ij}^{(n)} > 0$.

(ohne Beweis)

Satz 4.2 (Konvergenzsatz)
Es sei (X_n) irreduzibel, aperiodisch und positiv rekurrent mit stationärer Verteilung π. Dann gilt für alle $i, j \in S$:

$$\lim_{n\to\infty} p_{ij}^{(n)} = \pi(j) = \frac{1}{m_j}$$

Beweis
Wir benutzen ein sogenanntes “Kopplungsargument”.

(1) Sei (Y_n) eine weitere Markov-Kette, unabhängig von (X_n) mit gleicher Übergangsmatrix und Startverteilung π, also $Y_n \sim \pi$ für alle $n \in \mathbb{N}$. Es sei

$$T := \inf\{n \in \mathbb{N} \mid X_n = Y_n\}$$

die Treffzeit der Markov-Ketten. Wir zeigen zunächst: $P(T < \infty) = 1$.

Offenbar ist $(X_n, Y_n)_{n\in\mathbb{N}_0}$ eine Markov-Kette auf S^2 mit Übergangsmatrix $\hat{P} = (\hat{p}_{(ij)(kl)})$, wobei $\hat{p}_{(ij)(kl)} = p_{ik} \cdot p_{jl}$. Weiter ist $\hat{p}^{(n)}_{(ij)(kl)} = p^{(n)}_{ik} \cdot p^{(n)}_{jl}$ und mit Lemma 4.1: Die Kette (X_n, Y_n) ist irreduzibel und aperiodisch. Man kann nachrechnen:

$$\hat{\pi}(i, j) := \pi(i) \cdot \pi(j)$$

ist eine stationäre Verteilung für (X_n, Y_n), also ist sie nach Satz 3.2 positiv rekurrent.

Sei $X_0 = i$ und die Startverteilung $\hat{\nu}$ von (X_n, Y_n) gegeben durch $\hat{\nu}(k, l) = \delta_i(k) \cdot \pi(l)$. Für $b \in S$ sei

$$T_{(b,b)} := \inf\{n \in \mathbb{N} \mid (X_n, Y_n) = (b, b)\}.$$

Offenbar ist $T \leq T_{(b,b)}$ und $P_{\hat{\nu}}(T_{(b,b)} < \infty) = 1$. Daraus folgt, dass $P_{\hat{\nu}}(T < \infty) = 1$.

(2) Betrachte $(Z_n)_{n\in\mathbb{N}_0}$ definiert durch

$$Z_n = \begin{cases} X_n, & \text{für } n \leq T \\ Y_n, & \text{für } n > T. \end{cases}$$

Es ist (Z_n) eine Markov-Kette mit Übergangsmatrix P und $Z_0 = i$, denn:

Nach Satz 1.1 genügt es zu zeigen, dass

$$\forall n \in \mathbb{N}, i_k \in S : P_{\hat{\nu}}(Z_k = i_k,\, 0 \leq k \leq n) = \delta_i(i_0) \prod_{k=0}^{n-1} p_{i_k, i_{k+1}}$$

Es gilt

$$\begin{aligned}
&P_{\hat{\nu}}(Z_k = i_k,\, 0 \leq k \leq n) \\
&= \sum_{r=0}^{n} P_{\hat{\nu}}(Z_k = i_k,\, 0 \leq k \leq n,\, T = r) \\
&\quad + P_{\hat{\nu}}(Z_k = i_k,\, 0 \leq k \leq n,\, T > n) \\
&= \sum_{r=0}^{n} \underbrace{P_{\hat{\nu}}(X_k = i_k,\, 0 \leq k \leq r,\, Y_k = i_k,\, r+1 \leq k \leq n,\, Y_0 \neq i_0, \ldots, Y_{r-1} \neq i_{r-1},\, Y_r = i_r)}_{=:\mathrm{I}} \\
&\quad + \underbrace{P_{\hat{\nu}}(X_k = i_k,\, 0 \leq k \leq n,\, Y_0 \neq i_0, \ldots, Y_n \neq i_n)}_{=:\mathrm{II}}
\end{aligned}$$

mit

$$\begin{aligned}
\mathrm{I} = &P_{\hat\nu}(X_k = i_k, 0 \le k \le r) \cdot P_{\hat\nu}(Y_k = i_k, r+1 \le k \le n | Y_0 \ne i_0, \dots, Y_{r-1} \ne i_{r-1}, Y_r = i_r) \cdot \\
&P_{\hat\nu}(Y_0 \ne i_0, \dots, Y_{r-1} \ne i_{r-1}, Y_r = i_r) \\
= &\delta_i(i_0) \cdot \prod_{k=0}^{r-1} p_{i_k,i_{k+1}} \cdot \prod_{k=r}^{n-1} p_{i_k,i_{k+1}} \cdot P_{\hat\nu}(Y_0 \ne i_0, \dots, Y_{r-1} \ne i_{r-1}, Y_r = i_r) \\
= &\delta_i(i_0) \cdot \prod_{k=0}^{n-1} p_{i_k,i_{k+1}} \cdot P_{\hat\nu}(Y_0 \ne i_0, \dots, Y_{r-1} \ne i_{r-1}, Y_r = i_r), \\
\mathrm{II} = &\ldots = \delta_i(i_0) \cdot \prod_{k=0}^{n-1} p_{i_k,i_{k+1}} \cdot P_{\hat\nu}(Y_0 \ne i_0, \dots, Y_{n-1} \ne i_{n-1}, Y_n \ne i_n)
\end{aligned}$$

Tatsächlich gilt also

$$P_{\hat\nu}(Z_k = i_k, 0 \le k \le n) = \delta_i(i_0) \cdot \prod_{k=0}^{n-1} p_{i_k,i_{k+1}}$$

(3) Es gilt nun

$$\begin{aligned}
p_{i,j}^{(n)} &= P_{\hat\nu}(Z_n = j) = P_{\hat\nu}(Z_n = j, T \le n) + P_{\hat\nu}(Z_n = j, T > n) \\
\pi(j) &= P_{\hat\nu}(Y_n = j) = \underbrace{P_{\hat\nu}(Y_n = j, T \le n)}_{=P_{\hat\nu}(Z_n = j, T \le n)} + P_{\hat\nu}(Y_n = j, T > n) \\
\Rightarrow \quad |p_{i,j}^{(n)} - \pi(j)| &\le 2 \cdot P_{\hat\nu}(\underbrace{T > n}_{\downarrow\{T=\infty\}}) \longrightarrow 0
\end{aligned}$$

■

Satz 4.3
Seien i, $j \in S$ Zustände sowie d_j die Periode von j. Dann gilt:

$$\lim_{n\to\infty} p_{i,j}^{(n\cdot d_j + r)} = \frac{d_j}{m_j} \sum_{k=0}^{\infty} f_{i,j}^{(k\cdot d_j + r)} \qquad r = 1, \dots, d_j$$

Speziell:

a) Ist j transient oder nullrekurrent (d.h. $m_j = \infty$), so gilt $p_{i,j}^{(n)} \longrightarrow 0$

b) Ist j aperiodisch (d.h. $d_j = 1$), so gilt $p_{i,j}^{(n)} \longrightarrow \frac{1}{m_j} f_{i,j}^*$

5. Markov-Ketten und Martingale

Erinnerung: (X_n) heißt *Martingal*, falls

a) $E|X_n| < \infty$

b) $E[X_{n+1}|X_1,\dots,X_n] = X_n$, bzw.

b') $E[X_{n+1}|\mathcal{F}_n] = X_n$, wobei $\mathcal{F}_n := \sigma(X_1,\dots,X_n)$ die *natürliche Filtration* bezeichnet.

Erinnerung: Sei $(\Omega,\mathcal{F},P)$ W-Raum, X Zufallsvariable mit $E|X| < \infty$, $\mathcal{G} \subseteq \mathcal{F}$ Unter-σ-Algebra. $Z \coloneqq E[X|\mathcal{G}]$ heißt bedingter Erwartungswert von X bzgl. $\mathcal{G}$, falls

a) Z ist $\mathcal{G}$-messbar

b) $\int_A Z \cdot dP = \int_A X \cdot dP \quad \forall A \in \mathcal{G}$

Sei (X_n) eine Markov-Kette auf dem Zustandsraum S mit Übergangsmatrix P sowie $\mathcal{F}_n := \sigma(X_1,\dots,X_n)$ die natürliche Filtration. Weiter sei $h : S \to \mathbb{R}$ und $P : (S \to \mathbb{R}) \to (S \to \mathbb{R})$ definiert durch

$$(Ph)(i) := \sum_{j\in S} p_{i,j} \cdot h(j) \qquad \forall i \in S$$

NB: "Ph" macht Sinn im Sinne einer "Matrix·Vektor"-Multiplikation.

Lemma 5.1
Sei $h : S \to \mathbb{R}$ eine Funktion mit $\|P\,|h|\|_\infty < \infty$. Dann gilt:

$$(Ph)(X_n) = E[h(X_{n+1})|\mathcal{F}_n]$$

Beweis
Wir prüfen nach, dass $(Ph)(X_n)$ ein bedingter Erwartungswert von $h(X_{n+1})$ bzgl. $\mathcal{F}_n$ ist.

a) $(Ph)(X_n)$ ist $\mathcal{F}_n$-messbar, denn X_n ist $\mathcal{F}_n$-messbar.

b) Sei $A := \{X_0 = i_0,\dots,X_n = i_n\} \in \mathcal{F}_n$. Dann:

$$\begin{aligned}
&\int_A (Ph)(X_n)\cdot dP = \int 1_A \cdot (Ph)(X_n)\cdot dP = \int 1_{\{X_0=i_0,\dots,X_n=i_n\}}(Ph)(i_n)\cdot dP \\
=&(Ph)(i_n)\cdot P(A) = \sum_{j\in S} p_{i_n,j}\cdot h(j)\cdot P(X_0=i_0)\cdot \prod_{k=0}^{n-1} p_{i_k,i_{k+1}} \\
=&\sum_{j\in S} h(j)\cdot P(X_0=i_0,\dots,X_n=i_n,X_{n+1}=j) = \sum_{j\in S}\int_{A\cap\{X_{n+1}=j\}} h(X_{n+1})\cdot dP \\
=&\sum_{j\in S}\int_A 1_{\{X_{n+1}=j\}} h(X_{n+1})\cdot dP \overset{LDC}{=} \int_A h(X_{n+1})\cdot dP
\end{aligned}$$

Da jedes $A \in \mathcal{F}_n$ abzählbare Vereinigung solcher "Elementarereignisse" ist, folgt die Behauptung. ■

Satz und Definition 5.2
Seien P eine Übergangsmatrix auf S und $h : S \to \mathbb{R}$ eine Funktion mit $\|P\,|h|\|_\infty < \infty$. Gilt dann $Ph = h$, so nennen wir h *harmonisch*. Im Fall $Ph \geq h$ bzw. $Ph \leq h$ heißt h *sub*- bzw. *superharmonisch*.

Ist (X_n) eine Markov-Kette mit Übergangsmatrix P und h [sub-/super-]harmonisch, so ist $(h(X_n))$ ein $(\mathcal{F}_n)$-[Sub-/Super-]Martingal.

Beweis
Nach Lemma 5.1 gilt

$$E[h(X_{n+1})|\mathcal{F}_n] = Ph(X_n) = h(X_n) + (Ph - h)(X_n)$$

$\rightsquigarrow$ Behauptung. ■

Bemerkung: Ist $h : S \to \mathbb{R}$ eine beschränkte Funktion, so wird durch

$$\begin{aligned} Z_n^h &:= h(X_n) - \sum_{k=1}^{n} (E[h(X_k) \mid \mathcal{F}_{k-1}] - h(X_{k-1})) \\ &= h(X_n) - \sum_{k=1}^{n} (Ph - h)(X_{k-1}) \end{aligned}$$

ein Martingal definiert, genannt Levi-Martingal zu (X_n).

Die Markoveigenschaft lässt sich über Levi-Martingale charakterisieren:

Satz 5.3
Es sei $(X_n)_{n\in\mathbb{N}_0}$ ein stochastischer Prozess mit Werten in S und natürlicher Filtration $(\mathcal{F}_n)$, d.h. $\mathcal{F}_n = \sigma(X_0, \ldots, X_n)$, und P sei eine stochastische Matrix auf S. Ist dann für alle beschränkten $h : S \to \mathbb{R}$ der Prozess

$$Z_n^h := h(X_n) - \sum_{k=1}^{n} (Ph - h)(X_{k-1})$$

ein $(\mathcal{F}_n)$-Martingal, so ist X_n eine homogene Markov-Kette mit Übergangsmatrix P.

Beweis
Aus $E[Z_{n+1}^h \mid \mathcal{F}_n] = Z_n^h$ erhält man

$$E[h(X_{n+1}) \mid \mathcal{F}_n] = (Ph)(X_n)$$

bzw.

$$\int_A h(X_{n+1}) dP = \int_A (Ph)(X_n) dP \text{ für alle } A \in \mathcal{F}_n$$

Es sei $i_0, i_1, \ldots, i_{n+1} \in S$ und $A \coloneqq \{X_0 = i_0, \ldots, X_n = i_n\}$ und wir setzen $h \coloneqq 1_{\{i_{n+1}\}}$. Die linke Seite der letzten Gleichung ergibt dann

$$\int_A h(X_{n+1})dP = P(X_0 = i_0, \ldots, X_n = i_n, X_{n+1} = i_{n+1})$$

und die rechte Seite ergibt

$$\begin{aligned}\int_A (Ph)(X_n)dP &= \int_{\{X_0=i_0,\ldots,X_n=i_n\}} (Ph)(i_n)dP \\ &= \int_{\{X_0=i_0,\ldots,X_n=i_n\}} p_{i_n i_{n+1}} dP \\ &= p_{i_n i_{n+1}} \cdot P(X_0 = i_0, \ldots, X_n = i_n).\end{aligned}$$

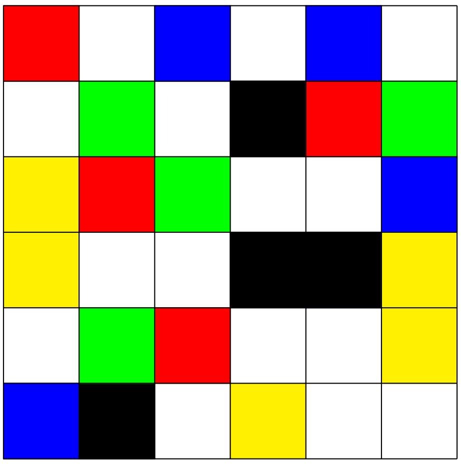
Durch Teilen der rechten Wahrscheinlichkeit erhält man

$$P(X_{n+1} = i_{n+1} \mid X_0 = i_0, \ldots, X_n = i_n) = p_{i_n i_{n+1}}. \qquad \blacksquare$$

Bemerkung: Es gilt: Ist $(X_n)_{n\in\mathbb{N}_0}$ ein nicht-negatives Supermartingal, so gibt es eine Zufallsvariable X_∞ mit

$$\lim_{n\to\infty} X_n = X_\infty \quad P-\text{f.s.}$$

Beispiel 5.1
Jedem Feld eines schachbrettartigen $N \times N$-Gitters wird eine von L möglichen Farben zugewiesen. Diese Färbung wird mittels Zufallsexperimenten modifiziert: Eine Zelle wird gleichverteilt gewählt, daraufhin wird einer der vier Nachbarn (modulo N) zufällig gewählt und dessen Farbe dem zuerst gewählten Feld zugewiesen. Offensichtlich ist dies eine Markov-Kette und die monochromen Zustände sind die absorbierenden.



Formal seien $L, N \in \mathbb{N}$, $L, N \geq 2$. $I \coloneqq \{1, \ldots, N\}^2$, $S \coloneqq \{1, \ldots, L\}^I = \{f : I \to \{1, \ldots, L\}\}$. Sei (X_n) die Markov-Kette in S, die die Zustandsfolge angibt.

Wie verhält sich die Folge für $n \to \infty$? Sei $l \in \{1, \ldots, L\}$ fest und Y_n sei die Anzahl der Felder mit Farbe l (zum Beispiel: blau) im Zustand X_n. Sei (A, B) ein Nachbarpaar im Gitter. Wäre dies die Wahl in einem Zustandsübergang, so gelte: Ist $X_n(A) = X_n(B)$ oder $X_n(A) \neq l, X_n(B) \neq l$, so gilt auch $Y_{n+1} = Y_n$. Ist dagegen $X_n(A) = l$ und $X_n(B) \neq l$, so ist $Y_{n+1} = Y_n - 1$. Ist letztlich $X_n(A) \neq l$ und $X_n(B) = l$, so ist $Y_{n+1} = Y_n + 1$.

Die Wahrscheinlichkeit, erst A, dann B zu wählen ist gleich der Wahrscheinlichkeit, erst B und dann A zu wählen. Damit ist

$$E[Y_{n+1} \mid X_1, \dots, X_n] = Y_n.$$

Sei $\mathcal{F}_n \coloneqq \sigma(X_0, \dots, X_n)$. Dann ist $(Y_n)_{n\in\mathbb{N}_0}$ ein $(\mathcal{F}_n)$-Martingal. Nach der Bemerkung folgt: $Y_n \to Y_\infty$ für ein $n \to \infty$ P-fast-sicher. Da (Y_n) ganzzahlig ist, ist $Y_n(\omega)$ konstant ab einem $n \geq n_0(\omega)$. Als Konstanten kommen nur 0 und N^2 in Frage, denn für $k \in \{1, \dots, N^2 - 1\}$ gilt

$$\begin{aligned} P(Y_{n+j} = k \mid Y_n = \dots = Y_{n+j-1} = k) &\leq 1 - \frac{1}{N^2 4} \\ \implies P(Y_n = \dots = Y_{n+j} = k) &\leq (1 - \frac{1}{N^2 4})^j \\ \implies P(\underbrace{Y_m = k, \forall m \geq n}_{=:A_n}) &= 0 \end{aligned}$$

Es gilt $\{\omega \in \Omega \mid \exists n\ \forall m \geq n : Y_m(\omega) = k\} = \bigcup_{n=0}^\infty A_n$. Damit ist

$$P(\lim_{n\to\infty} Y_n = k) = P(\exists n\ \forall m \geq n : Y_n = k) \leq \sum_{n=0}^\infty P(A_n) = 0$$

und wir folgern, dass $P(Y_\infty \in \{0, N^2\}) = 1$.

Außerdem gilt noch, da Y_n beschränkt ist:

$$EY_\infty = \lim_{n\to\infty} EY_n = EY_0.$$

Damit ist die Wahrscheinlichkeit, dass das Feld irgendwann komplett blau ist, gleich

$$P(Y_\infty = N^2) = \frac{1}{N^2} EY_\infty = \frac{1}{N^2} EY_0 = \frac{1}{N^2} \#\{A \in I \mid X_0(A) = l\}.$$

Anwendungen dieses Modells findet man in der Physik (Vielteilchensysteme), in der Biologie (Ausbreitung von Infektionen) oder in der Finanzmathematik (Kreditrisiken).

6. Die starke Markov-Eigenschaft

Gegeben sei eine Markov-Kette $(X_n)_{n\in\mathbb{N}_0}$ mit Zustandsraum S auf dem Wahrscheinlichkeitsraum $(\Omega = S_0^{\mathbb{N}_0}, \mathcal{F} \coloneqq \otimes_{n=0}^\infty \mathcal{P}(S), P)$. Beachte, dass die Mengen

$$Z(i_0, i_1, \dots, i_n) \coloneqq \{i_0\} \times \{i_1\} \times \dots \times \{i_n\} \times S \times S \times \dots$$

für $n \in \mathbb{N}_0$, $i_0, \dots i_n \in S$ ein durchschnittsstabiles Erzeugendensystem für $\mathcal{F}$ bilden. Weiter sei $(\mathcal{F}_n)_{n\in\mathbb{N}_0}$ mit $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$ die natürliche Filtration von (X_n).

$\tau : \Omega \to \mathbb{N}_0$ sei eine $(\mathcal{F}_n)$-Stoppzeit, das heißt $\{\tau \leq n\} \in \mathcal{F}_n$ und $P(\tau < \infty) = 1$. Die gestoppte Markov-Kette $X^\tau = (X_n^\tau)_{n\in\mathbb{N}_0}$ ist

$$X_n^\tau \coloneqq X_{\min(\tau,n)}$$

für $n \in \mathbb{N}_0$ und $Y = (Y_n)_{n\in\mathbb{N}_0}$ definiert durch

$$Y_n := X_{\tau+n}$$

heißt der Post-τ-Prozess.

Satz 6.1 (Starke Markov-Eigenschaft)
Mit den oben eingeführten Bezeichnungen gilt:

a) Y ist eine Markov-Kette mit Übergangsmatrix P und Startverteilung ν, wobei $X_\tau \backsim \nu$.

b) X^τ und Y sind unter X_τ bedingt unabhängig.

Beweis
a.) Es gilt:

$$\begin{aligned}
&P(Y_0 = i_0, \dots, Y_n = i_n, Y_{n+1} = i_{n+1})\\
=&\sum_{k=0}^{\infty} P(\tau = k, X_k = i_0, \dots, X_{k+n+1} = i_{n+1})\\
=&\sum_{k=0}^{\infty} P(X_{k+n+1} = i_{n+1} \mid X_k = i_0, \dots, X_{k+n} = i_n, \tau = k) \cdot P(X_k = i_0, \dots, X_{k+n} = i_n, \tau = k)\\
=&p_{i_n i_{n+1}} \sum_{k=0}^{\infty} P(X_k = i_0, \dots, X_{k+n} = i_n, \tau = k)\\
=&p_{i_n i_{n+1}} P(Y_0 = i_0, \dots, Y_n = i_n) \implies \text{ Behauptung.}
\end{aligned}$$

b.) Seien

$$\begin{aligned}
A &:= Z(i_0, \dots, i_m) = \{i_0\} \times \dots \times \{i_m\} \times S \times S \times \dots\\
B &:= Z(j_0, \dots, j_n),
\end{aligned}$$

dann gilt:

$$\begin{aligned}
&P(X^\tau \in A, Y \in B, X_\tau = j)\\
=&\sum_{k=0}^{\infty} P(\tau = k, X_{0\wedge k} = i_0, \dots, X_{m\wedge k} = i_m, X_k = j_0, \dots, X_{k+n} = j_n, X_k = j)\\
=&\sum_{k=0}^{\infty} P(X_k = j_0, \dots, X_{k+n} = j_n \mid X_{0\wedge k} = i_0, \dots, X_{m\wedge k} = i_m, X_k = j, \tau = k)\\
&\quad \cdot P(X_{0\wedge k} = i_0, \dots, X_{m\wedge k} = i_m, X_k = j, \tau = k)\\
=&P(X_0 = j_0, \dots, X_n = j_n \mid X_0 = j) \cdot P(X^\tau \in A, X_\tau = j)\\
=&P(Y \in B \mid X_\tau = j) \cdot P(X^\tau \in A \mid X_\tau = j) \cdot P(X_\tau = j)
\end{aligned}$$

Teilen durch $P(X_\tau = j) \implies$ Behauptung. ■

II. Markov-Ketten in stetiger Zeit

7. Ein wichtiger Spezialfall: der Poisson-Prozess

Gegeben sei ein Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$. Wir betrachten jetzt einen stochastischen Prozess $N = (N_t)_{t\geq 0}$ mit Zustandsraum $\mathbb{N}_0$, d.h. $(N_t)_{t\geq 0}$ ist eine Familie von $(\mathcal{F}, \mathcal{P}(\mathbb{N}_0))$-messbaren ZV, der bestimmte Ereignisse zählen soll (z.B. Emission von Partikeln beim radioaktiven Zerfall, Ankünfte von Kunden in einem Bediensystem, Schäden bei einer Versicherung).
Wir nehmen an, dass mindestens ein Ereignis eintritt und die Anzahl der Ereignisse in einem kompakten Intervall soll endlich sein.
Wir stellen folgende Forderungen an N:

(A1) Alle Pfade $t \mapsto N(t, \omega)$ liegen in

$$D_0 := \{f : [0, \infty) \longrightarrow \mathbb{N}_0 \mid f(0) = 0, f \uparrow, f \text{ stetig von rechts } \}.$$

(A2) $(N_t)_{t\geq 0}$ hat unabhängige Zuwächse, d.h. für alle $0 \leq t_0 \leq t_1 \leq \cdots \leq t_n, n \in \mathbb{N}$ sind die Zufallsvariablen

$$N_{t_0}, N_{t_1} - N_{t_0}, \ldots, N_{t_n} - N_{t_{n-1}}$$

stochastisch unabhängig.

(A3) $(N_t)_{t\geq 0}$ hat stationäre Zuwächse, d.h. $\forall t > 0$ hängt die Verteilung von $N_{s+t} - N_s$ nicht von s ab.

(A4) Ereignisse treten einzeln auf, d.h. $P(N_h \geq 2) = o(h)$ mit $h \downarrow 0$.

Bemerkung: Der stochastische Prozess kann als Zufallsgröße mit Werten in einem Funktionenraum aufgefasst werden, d.h.

$$N : \Omega \ni \omega \mapsto N(\cdot, \omega) \in D_0$$

Als σ-Algebra auf D_0 wählen wir

$$\mathfrak{B}(D_0) := \sigma(\{\pi_t : t \geq 0\}),$$

wobei $\pi_t : D_0 \longrightarrow \mathbb{N}_0, f \mapsto f(t)$ die t-te Projektion ist.
Es gilt: $N : \Omega \longrightarrow D_0$ ist $(\mathcal{F}, \mathfrak{B}(D_0))$-messbar $\iff N_t = \pi_t \circ N : \Omega \longrightarrow \mathbb{N}_0$ sind $(\mathcal{F}, \mathcal{P}(\mathbb{N}_0))$-messbar $\forall t \geq 0$. (Stochastik II Übungsaufgabe 21)
Die Mengen der Form

$$A(t_1, \ldots, t_n, i_1, \ldots, i_n) := \{f \in D_0 \mid f(t_j) - f(t_{j-1}) = i_j \text{ für } j = 1, \ldots, n\}$$

mit $0 =: t_0 \leq t_1 \leq \cdots \leq t_n < \infty, i_1, \ldots, i_n \in \mathbb{N}_0$ bilden ein durchschnittsstabiles Erzeugendensystem von $\mathfrak{B}(D_0)$.

Satz 7.1
Es sei $N = (N_t)_{t\geq 0}$ ein stochastischer Prozess, der den Bedingungen (A1)-(A4) genügt. Dann hat N mit Wahrscheinlichkeit 1 nur Sprünge der Höhe 1 und es existiert ein $\lambda > 0$, so dass:

a) $\forall s, t \geq 0$ ist $N_{s+t} - N_s$ Poisson-verteilt mit Parameter λt.

b) Die Zeiten zwischen aufeinanderfolgenden Sprüngen des Prozesses sind unabhängig und exponentialverteilt mit Parameter λ.

Beweis
Sprünge der Höhe 1: Es ist also zu zeigen:

$$P(N_s - N_{s-} \geq 2 \text{ für ein } s > 0) = 0$$

wobei N_{s-} der linksseitige Grenzwert ist.
Für festes $t > 0$ gilt:

$$\begin{aligned} &P(N_s - N_{s-} \geq 2 \text{ für ein } s \in (0, t]) \\ \leq\, &P(N_{\frac{kt}{n}} - N_{\frac{(k-1)t}{n}} \geq 2 \text{ für ein } k \in \{1, \ldots, n\}) \\ \leq\, &n \cdot P(N_{\frac{t}{n}} \geq 2) \end{aligned}$$

(das letzte $\leq$ gilt wegen (A3)).
Aus (A4) folgt:

$$\lim_{n\to\infty} \frac{P(N_{\frac{t}{n}} \geq 2)}{\frac{t}{n}} \cdot t = 0.$$

Weiter gilt mit Stetigkeit des Wahrscheinlichkeitsmaßes:

$$P(N_s - N_{s-} \geq 2 \text{ für ein } s > 0) = \lim_{t\to\infty} P(N_s - N_{s-} \geq 2 \text{ für ein } s \in (0, t]) = 0.$$

a) Betrachte dazu

$$\Phi : [0, \infty) \to [0, 1] \text{ mit } \Phi(t) := P(N_t = 0).$$

Für alle $s, t > 0$ folgt:

$$\begin{aligned} \Phi(s + t) &= P(N_s = 0, N_{s+t} - N_s = 0) \\ &= P(N_s = 0) \cdot P(N_t = 0) \qquad \text{(nach (A3) und (A2))} \\ &= \Phi(s) \cdot \Phi(t). \end{aligned}$$

Daraus folgt, dass $\Phi(\frac{k}{n}) = (\Phi(\frac{1}{n}))^k$ und $\Phi(1) = \Phi(\frac{1}{n} + \cdots + \frac{1}{n}) = (\Phi(\frac{1}{n}))^n$ für $k, n \in \mathbb{N}$ gilt. Das heißt, dass $\Phi(\frac{k}{n}) = (\Phi(1))^{\frac{k}{n}}$.

Da Φ fallend ist, folgt mit einem Einschachtelungsargument:

$$\Phi(t) = (\Phi(1))^t \text{ für alle } t > 0.$$

Weiter ist $\Phi(1) \in (0,1)$, da $\Phi(1) \in [0,1]$ und falls $\Phi(1) = 1$, dann wäre

$$P(N_t = 0 \text{ für alle } t \geq 0) = \lim_{t\to\infty} \Phi(t) = 1$$

im Widerspruch zur Forderung, dass mindestens ein Ereignis eintritt, und wäre $\Phi(1) = 0$, dann gälte für alle $n \in \mathbb{N}$

$$\begin{aligned} P(N_1 \geq n) &\geq P(N_{\frac{k}{n}} - N_{\frac{k-1}{n}} \geq 1 \text{ für } k = 1,\dots,n) \\ &= \Big(1 - \Phi(\frac{1}{n})\Big)^n && \text{((A2), (A3))} \\ &= 1 \end{aligned}$$

im Widerspruch zur Forderung, dass in kompakten Intervallen nur endlich viele Ereignisse eintreten.

Sei jetzt

$$\lambda := -\log \Phi(1).$$

Es ist $0 < \lambda < \infty$ und $P(N_t = 0) = (\Phi(1))^t = e^{-\lambda t}$. Weiter sei $t > 0$ und für $n \in \mathbb{N}$ definieren wir

$$\begin{aligned} X_{nk} &:= 1_{\mathbb{N}}\big(N_{\frac{kt}{n}} - N_{\frac{(k-1)t}{n}}\big) \\ &= \begin{cases} 1, & \text{falls mindestens ein Ereignis in } \big[\frac{(k-1)t}{n}, \frac{kt}{n}\big) \text{ eintritt} \\ 0, & \text{sonst} \end{cases} \end{aligned}$$

Wegen (A2) sind $X_{n1},\dots,X_{nn}$ unabhängig und wegen (A3) identisch verteilt mit $B(1, 1-e^{-\lambda\frac{t}{n}})$. Dann gilt

$$\sum_{k=1}^n X_{nk} \sim B(n, 1-e^{-\lambda\frac{t}{n}}) \xrightarrow{d} \mathrm{Po}(\lambda t)$$

da $n(1-e^{-\lambda\frac{t}{n}}) \xrightarrow{n\to\infty} \lambda t$.

Aus (A4) folgt:

$$\begin{aligned} P(\sum_{k=1}^n X_{nk} \neq N_t) &= P(\bigcup_{k=1}^n \{N_{\frac{kt}{n}} - N_{\frac{(k-1)t}{n}} \geq 2\}) \\ &\leq n \cdot P(N_{\frac{t}{n}} \geq 2) \xrightarrow{n\to\infty} 0 \end{aligned}$$

und damit haben wir gezeigt, dass

$$N_t \sim \mathrm{Po}(\lambda t).$$

b) Sei

$$T_1 := \inf\{t > 0 \mid N_t \neq 0\}$$

der erste Sprungzeitpunkt. Wir erhalten

$$P(T_1 > t) = P(N_t = 0) = e^{-\lambda t} \text{ mit } t \geq 0$$

und damit

$$T_1 \sim \mathrm{Exp}(\lambda).$$

Die weiteren Aussagen werden später gezeigt. ∎

Bemerkung: Der Prozess $N = (N_t)_{t\geq 0}$ in Satz 7.1 heißt *Poisson-Prozess* mit Parameter λ.

Die Bedingungen (A2) und (A3) können etwas „abstrakter" gefasst werden: Für $u \geq 0$ definiere

$$\begin{aligned} S_u &: D_0 \to D_0, && f \mapsto f(\cdot \wedge u) \\ Z_u &: D_0 \to D_0, && f \mapsto f(u + \cdot) - f(u) \end{aligned}$$

wobei $(\cdot \wedge u)$ die Abbildung $v \mapsto \min(v, u)$ darstellt.

Beide Abbildungen sind $(\mathfrak{B}(D_0), \mathfrak{B}(D_0))$-messbar.

Wir können also auch stochastische Prozesse $S_u(N) = (N_{t\wedge u})_{t\geq 0}$ und $Z_u(N) = (N_{u+t} - N_u)_{t\geq 0}$ als Zufallsgrößen mit Werten in $(D_0, \mathfrak{B}(D_0))$ auffassen.

Weiter gilt mit

$$A(t_1, \dots, t_n; i_1, \dots, i_n) := \{f \in D_0 \mid f(t_j) - f(t_{j-1}) = i_j \text{ für } j = 1, \dots, n\},$$

wobei $0 =: t_0 \leq t_1 \leq \dots \leq t_n < \infty$, $i_1, \dots, i_n \in \mathbb{N}_0$, dass

$$\{S_u(N) \in A(t_1, \dots, t_n; i_1, \dots, i_n)\} = \{N_{t_1\wedge u} - N_{t_0\wedge u} = i_1, \dots, N_{t_n\wedge u} - N_{t_{n-1}\wedge u} = i_n\}$$

und

$$\{Z_u(N) \in A(s_1, \dots, s_l; j_1, \dots, j_l)\} = \{N_{u+s_1} - N_{u+s_0} = j_1, \dots, N_{u+s_l} - N_{u+s_{l-1}} = j_l\}$$

Wegen (A2) folgt:

$$\begin{aligned} &P\big(S_u(N) \in A(t_1, \dots, t_n; i_1, \dots, i_n),\, Z_u(N) \in A(s_1, \dots, s_l; j_1, \dots, j_l)\big) \\ &\qquad = P\big(S_u(N) \in A(t_1, \dots, t_n; i_1, \dots, i_n)\big) \cdot P\big(Z_u(N) \in A(s_1, \dots, s_l; j_1, \dots, j_l)\big) \end{aligned}$$

und daraus die Unabhängigkeit der Prozesse $S_u(N)$ und $Z_u(N)$.

Analog folgt mit (A3):

$$P\big(Z_u(N) \in A(s_1, \dots, s_l; j_1, \dots, j_l)\big) = P\big(N \in A(s_1, \dots, s_l; j_1, \dots, j_l)\big)$$

Also haben $Z_u(N)$ und N dieselbe Verteilung.

Diese Aussagen können jetzt auf Stoppzeiten verallgemeinert werden. Es sei dazu $\mathcal{F}_t := \sigma(\{N_s, s \leq t\})$ für $t \geq 0$ und $(\mathcal{F}_t)_{t\geq 0}$ die natürliche Filtration.

Definition
Sei τ eine $(\mathcal{F}_t)_{t\geq 0}$-Stoppzeit, das heißt $\{\tau \leq t\} \in \mathcal{F}_t$. Der Prozess

$$S_\tau(N) = (N_{\tau\wedge t})_{t\geq 0}$$

heißt *Prä-τ-Prozess* und der Prozess

$$Z_\tau(N) = (N_{\tau+t} - N_\tau)_{t\geq 0}$$

heißt *Post-τ-Prozess*.

Lemma 7.2
Ist τ eine endliche Stoppzeit, so sind die Prozesse $(N_{\tau\wedge t})_{t\geq 0}$ und $(N_{\tau+t}-N_\tau)_{t\geq 0}$ stochastisch unabhängig. Außerdem hat $(N_{\tau+t}-N_\tau)_{t\geq 0}$ dieselbe Verteilung wie $(N_t)_{t\geq 0}$.

Beweis
Zunächst nehmen wir an, dass τ Werte in $\mathbb{Q}_+$ annimmt. Für $0 =: s_0 \leq s_1 \leq \dots \leq s_k < \infty$, $0 =: t_0 \leq t_1 \leq \dots \leq t_l < \infty$ und beliebige $i_1, \dots, i_k, j_1, \dots, j_l \in \mathbb{N}_0$ gilt

$$\begin{aligned} C &:= P\big(S_\tau(N) \in \underbrace{A(s_1, \dots, s_k; i_1, \dots, i_k)}_{=:A},\, Z_\tau(N) \in \underbrace{A(t_1, \dots, t_l; j_1, \dots, j_l)}_{=:B}\big) \\ &= \sum_{u\in\mathbb{Q}_+} P(\tau = u,\, S_u(N) \in A,\, Z_u(N) \in B) \end{aligned}$$

Da $\{\tau = u\} \in \mathcal{F}_u$ kann dieses Ereignis durch $S_u(N)$ ausgedrückt werden. Da $S_u(N)$ und $Z_u(N)$ unabhängig sind, folgt:

$$\begin{aligned} \dots &= \sum_{u\in\mathbb{Q}_+} P(\tau = u, S_u(N) \in A) \cdot \underbrace{P(Z_u(N) \in B)}_{=P(N\in B)} \\ &= P(N \in B) \cdot P(S_\tau(N) \in A) \end{aligned}$$

Im Fall $k = 1, s_1 = 0, i_1 = 0$ folgt:

$$P(Z_\tau(N) \in B) = P(N \in B)$$

also gilt $Z_\tau(N) \overset{\mathrm{d}}{=} N$ und damit

$$P(S_\tau(N) \in A,\, Z_\tau(N) \in B) = P(Z_\tau(N) \in B) \cdot P(S_\tau(N) \in A)$$

also sind $Z_\tau(N)$ und $S_\tau(N)$ unabhängig.

Weiter sei τ beliebig. Betrachte die Folge

$$\tau_n := \frac{\lceil 2^n\tau\rceil}{2^n},\ n \in \mathbb{N},$$

für die $\tau_n \in \mathbb{Q}_+$ P-fast-sicher gilt und $\tau_n \to \tau$ für $n \to \infty$. Mit (A1) folgt für alle $t \geq 0$

$$N_{\tau_n\wedge t} \xrightarrow{n\to\infty} N_{\tau\wedge t}, \quad N_{\tau_n+t} \xrightarrow{n\to\infty} N_{\tau+t}$$

P-fast-sicher und

$$\begin{aligned} P(S_\tau(N) \in A,\, Z_\tau(N) \in B) &= \lim_{n\to\infty} P(S_{\tau_n}(N) \in A,\, Z_{\tau_n}(N) \in B) \\ &= \lim_{n\to\infty} P(S_{\tau_n}(n) \in A) \cdot P(Z_{\tau_n}(N) \in B) \\ &= P(S_\tau(N) \in A) \cdot P(Z_\tau(N) \in B) \end{aligned}$$

Analog kann man die Aussage $Z_\tau(N) \overset{\mathrm{d}}{=} N$ auf beliebige τ erweitern. ∎

Damit können wir den Beweis von Satz 7.1 abschließen:

Es seien

$$\begin{aligned} T_1 &:= \inf\{t > 0 \mid N_t \neq 0\} \\ S_1 &:= T_1 \\ S_k &:= \inf\{t > S_{k-1} \mid N_t \neq N_{t-}\},\ k = 2, 3, \dots \\ T_k &:= S_k - S_{k-1},\ k = 2, 3, \dots \end{aligned}$$

Wir wissen bereits, dass $T_1 \sim \exp(\lambda)$. Induktiv nehmen wir an, dass $T_1, \dots, T_k$ unabhängig und identisch exponential-verteilt seien. $\tau := T_1 + \dots + T_k = S_k$ ist eine Stoppzeit, da $\{\tau \leq t\} = \{N_t \geq k\} \in \mathcal{F}_t$. Da $P(\tau > t) = P(N_t < k) \to 0$ für $t \to \infty$ gilt, ist τ P-fast-sicher endlich.

T_{k+1} ist die Zeit bis zum ersten Sprung im Post-τ-Prozess. Nach Lemma 7.2 ist $T_{k+1} \stackrel{\mathrm{d}}{=} N_1 \stackrel{\mathrm{d}}{=} N_1 - N_0 \stackrel{\mathrm{d}}{=} T_1 \sim \exp(\lambda)$ und T_{k+1} ist unabhängig von $S_\tau(N)$, also auch von $T_1, \dots, T_k$.

Beispiel 7.1 (Bedingte Gleichverteilungseigenschaft)
Es seien $K, X_1, X_2, \dots$ unabhängige Zufallsvariablen mit $K \sim \mathrm{Po}(\lambda T)$, $T > 0$, und $X_i \sim U(0, T)$.

Wir definieren

$$N_t := \#\{1 \leq i \leq K \mid X_i \leq t\},\ t \geq 0.$$

Es sei $0 =: t_0 \leq t_1 \leq \dots \leq t_n = T$ eine Zerlegung des Intervalls $[0, T]$. Betrachte für $i_1, \dots, i_n \in \mathbb{N}_0$ das Ereignis

$$A := \{N_{t_1} - N_{t_0} = i_1, \dots, N_{t_n} - N_{t_{n-1}} = i_n\}.$$

Für $\omega \in A$ folgt $k := K(\omega) = i_1 + \dots + i_n$. Man erhält (siehe Definition der Multinomialverteilung, Henze, Stochastik I, S. 121)

$$\begin{aligned} P(N_{t_1} - N_{t_0} = i_1, \dots, N_{t_n} - N_{t_{n-1}} = i_n) &= e^{-\lambda T} \frac{(\lambda T)^k}{k!} \cdot \frac{k!}{i_1! \cdots i_n!} \cdot \left(\frac{t_1 - t_0}{T}\right)^{i_1} \cdots \left(\frac{t_n - t_{n-1}}{T}\right)^{i_n} \\ &= \prod_{l=1}^{n} \underbrace{e^{-\lambda(t_l - t_{l-1})} \frac{(\lambda(t_l - t_{l-1}))^{i_l}}{i_l!}}_{=P(M_l = i_l)} \end{aligned}$$

wobei $M_l \sim \mathrm{Po}(\lambda(t_l - t_{l-1}))$. Das zeigt: Die Zuwächse des Prozesses $\tilde{N} = (N_t)_{0 \leq t \leq T}$ sind unabhängig und stationär und Poisson-verteilt mit Parameter λ·Intervalllänge, und (A1), (A4) ist erfüllt, also ist $\tilde{N}$ ein bei T gestoppter Poisson-Prozess.

Beispiel 7.2 (Das Inspektions-Paradoxon)
Es seien $X_1, X_2, \dots$ unabhängig und identisch verteilte Zufallsvariablen mit $X_k \sim \exp(\lambda)$, welche die Lebensdauer von Glühbirnen modellieren.

Es sei $S_n := \sum_{k=1}^{n} X_k$ der Zeitpunkt, an dem die n-te Birne kaputt geht und eine neue Birne eingesetzt wird, und N_t die Anzahl der Erneuerungen bis zum Zeitpunkt t und damit ein Poisson-Prozess mit Parameter λ.

In der Erneuerungstheorie interessiert man sich für

$$\begin{aligned} V_t &:= S_{N_{t+1}} - t = \text{Restlebensdauer} \\ W_t &:= t - S_{N_t} = \text{Alter} \\ L_t &:= W_t + V_t = \text{Gesamtlebensdauer} \end{aligned}$$

der zum Zeitpunkt t in Gebrauch befindlichen Glühbirne.

Es ist $V_t \sim \exp(\lambda)$, da die Exponentialverteilung gedächtnislos ist. Die Variable W_t kann höchstens t sein, das heißt $P(W_t > s) = 0$ für $s > t$. Für $0 \leq s \leq t$ gilt $P(W_t \geq s) = P(N_t - N_{t-s} = 0) = e^{-\lambda s}$. Damit ist

$$F_{W_t}(s) = \begin{cases} 1 - e^{-\lambda s}, & 0 \leq s \leq t \\ 1, & s \geq t. \end{cases}$$

Außerdem sind nach der starken Markoveigenschaft die Zufallsvariablen W_t und V_t unabhängig. L_t ergibt sich als Faltung dieser Variablen. Die Dichte f_{L_t} für $s \geq t$:

$$\begin{aligned} f_{L_t}(s) &= \int_0^s f_{V_t}(s-u) F_{W_t}(du) \\ &= \int_0^t \lambda e^{-\lambda(s-u)} \lambda e^{-\lambda u} du + \lambda e^{-\lambda(s-t)} \cdot e^{-\lambda t} \\ &= \lambda(1+\lambda t) e^{-\lambda s} \end{aligned}$$

und für $s < t$:

$$\begin{aligned} f_{L_t}(s) &= \int_0^s \lambda e^{-\lambda(s-u)} \lambda e^{-\lambda u} du \\ &= \lambda^2 s e^{-\lambda s} \end{aligned}$$

Also ist L_t *nicht* $\exp(\lambda)$-verteilt! Für großes t gilt

$$EL_t \approx 2EX_i = \frac{2}{\lambda}.$$

Dieses Ergebnis lässt sich erklären: Inspiziert man zu einer zufälligen Zeit die aktuell leuchtende Glühbirne, so ist die Wahrscheinlichkeit, eine länger lebende zu erwischen, größer, als eine kurzlebige Birne anzutreffen.

8. Der allgemeine Fall im Schnelldurchgang

Definition
Ein stochastischer Prozess $(X_t)_{t\geq 0}$ mit abzählbarem Zustandsraum S heißt (homogene) Markov-Kette, falls gilt: Für alle $n \in \mathbb{N}$ und $0 \leq t_0 < t_1 < \cdots < t_n$, $t, h > 0$ und $i_k \in S$ mit $P(X_{t_k} = i_k,\, 0 \leq k \leq n) > 0$:

$$\begin{aligned} P(X_{t_n+h} = i_{n+1} \mid X_{t_k} = i_k,\, 0 \leq k \leq n) &= P(X_{t_n+h} = i_{n+1} \mid X_{t_n} = i_n) \\ &= P(X_{t+h} = i_{n+1} \mid X_t = i_n) \end{aligned}$$

Bemerkung: Definieren wir für alle $i, j \in S$

$$p_{ij}(t) := P(X_t = j \mid X_0 = i)$$

und die Matrix

$$P(t) = \big(p_{ij}(t)\big)_{i,j \in S},$$

so gelten analog zum diskreten Fall die Chapman-Kolmogorov-Gleichungen für alle $i, j \in S$, $s, t > 0$:

$$p_{ij}(t+s) = \sum_{k \in S} p_{ik}(t) p_{kj}(s)$$

Mit $P(0) = E$ wird $\{P(t) : t \geq 0\}$ zu einer Halbgruppe von stochastischen Matrizen. Diese nennen wir *Übergangsmatrizenfunktion*.

Falls zusätzlich für $i, j \in S$

$$\lim_{t \downarrow 0} p_{ij}(t) = \delta_{ij}$$

gilt, also $P(t)$ rechtsseitig stetig in 0 ist, dann heißt sie *Standardübergangsmatrizenfunktion*.

Satz und Definition 8.1
Sei $\{P(t) : t \geq 0\}$ eine Standardübergangsmatrizenfunktion. Dann ist jedes $p_{ij}(t)$ in 0 rechtseitig differenzierbar, das heißt es existiert für alle $i, j \in S$

$$q_{ij} := \lim_{t \downarrow 0} \frac{1}{t}(p_{ij}(t) - \delta_{ij}).$$

Die Matrix $Q := (q_{ij})$ heißt *Intensitätsmatrix* oder *infinitesimaler Erzeuger* (Generator) zu $\{P(t) : t \geq 0\}$.

Beispiel 8.1
Es sei $N = (N_t)_{t \geq 0}$ ein Poisson-Prozess mit Intensität λ. Dann gilt für $i_{n+1} \geq i_n$:

$$\begin{aligned} P(N_{t_{n+1}} = i_{n+1} \mid N_{t_0} = i_0, \ldots, N_{t_n} = i_n) &= P(N_{t_{n+1}} - N_{t_n} = i_{n+1} - i_n) \\ &= e^{-\lambda(t_{n+1} - t_n)} \frac{(\lambda(t_{n+1} - t_n))^{i_{n+1} - i_n}}{(i_{n+1} - i_n)!} \end{aligned}$$

Also ist N eine homogene Markov-Kette mit Übergangsmatrix

$$p_{ij}(t) = \begin{cases} e^{-\lambda t} \frac{(\lambda t)^{j-i}}{(j-i)!}, & \text{falls } j \geq i \\ 0, & \text{sonst} \end{cases}$$

weiter gilt

$$\lim_{t \downarrow 0} \frac{1}{t}(p_{ij}(t) - \delta_{ij}) = \begin{cases} \lambda, & \text{falls } j = i+1 \\ -\lambda, & \text{falls } j = i \\ 0, & \text{sonst.} \end{cases}$$

Lemma 8.2
Sei $\{P(t) : t \geq 0\}$ eine Standardübergangsmatrizenfunktion. Dann gilt für $q_{ij} := p'_{ij}(0)$:

a) $0 \leq q_{ij} < \infty$, $i \neq j$, $-\infty \leq q_{ii} \leq 0$.

b) $\sum_{j \neq i} q_{ij} \leq -q_{ii} =: q_i$. Falls S endlich ist, gilt für alle $i \in S$: $q_i = \sum_{j \neq i} q_{ij}$. In diesem Fall heißt die Standardübergangsmatrizenfunktion *konservativ*.

Beweis

a) $0 \le q_{ij}$ für $i \ne j$ und $q_{ii} \le 0$ klar, da $0 \le p_{ij} \le 1$. Schwieriger ist zu zeigen, dass $q_{ij} < \infty$, hier wird auf die Literatur verwiesen.

b) Es gilt für $t \ge 0$:

$$\sum_{j\in S, j\ne i} \frac{p_{ij}(t)}{t} = \frac{1 - p_{ii}(t)}{t}$$

und damit mit Fatou

$$-q_{ii} = \lim_{t\downarrow 0} \frac{1 - p_{ii}(t)}{t} \ge \sum_{j\in S, j\ne i} \limsup_{t\downarrow 0} \frac{p_{ij}(t)}{t} = \sum_{j\in S, j\ne i} q_{ij}$$

und Gleichheit für endliche S. ■

Satz 8.3
Sei $\{P(t) : t \ge 0\}$ eine Standardübergangsmatritzenfunktion und S sei endlich. Dann gilt das sogenannte *Kolmogorovsche Rückwärtsdifferentialgleichungssystem*: Für $t \ge 0$ ist

$$P'(t) = QP(t)$$

das heißt für alle $i, j \in S$ ist

$$p'_{ij}(t) = -q_i p_{ij}(t) + \sum_{k\in S, k\ne i} q_{ik} p_{kj}(t).$$

Beweis
Wegen Chapman-Kolmogorov gilt für $i, j \in S$, $t, h > 0$:

$$\begin{aligned} p_{ij}(t+h) &= \sum_{k\in S} p_{ik}(h) p_{kj}(t) \\ \Longrightarrow \quad \frac{p_{ij}(t+h) - p_{ij}(t)}{h} &= \frac{p_{ii}(h) - 1}{h} p_{ij}(t) + \sum_{k\in S, k\ne i} \frac{p_{ik}(h)}{h} p_{kj}(t) \end{aligned}$$

Für $h \downarrow 0$ geht die rechte Seite gegen

$$-q_i p_{ij}(t) + \sum_{k\in S, k\ne i} q_{ik} p_{kj}(t)$$

woraus die Behauptung folgt. ■

Bemerkung: a) Unter milden Bedingungen ($\{P(t) : t \ge 0\}$ eine konservative Standardübergangsmatrizenfunktion und q_i für alle $i \in S$ endlich) gilt Satz 8.3 auch für abzählbares S. Weitere Bedingungen sind nötig für das Kolmogorovsche Vorwärtsdifferentialgleichungssystem

$$P'(t) = P(t)Q.$$

b) Falls S endlich ist, ist die Lösung von $P'(t) = QP(t)$, $P(0) = E$ gegeben durch

$$P(t) = e^{tQ} = \sum_{n=0}^{\infty} \frac{(tQ)^n}{n!}.$$

Beispiel 8.2

$$S = \{0,1\}, 0 < q_0, q_1 < \infty$$
$$Q = \begin{pmatrix} -q_0 & q_0 \\ q_1 & -q_1 \end{pmatrix}$$

Rückwärtsdifferentialgleichung $P'(t) = QP(t)$

$$P(t) = \begin{pmatrix} p_{00}(t) & p_{01}(t) \\ p_{10}(t) & p_{11}(t) \end{pmatrix}$$

$$\begin{aligned}
(1) \quad & p'_{00}(t) = -q_0 p_{00}(t) + q_0 p_{10}(t) \\
(2) \quad & p'_{01}(t) = -q_0 p_{01}(t) + q_0 p_{11}(t) \\
(3) \quad & p'_{10}(t) = -q_1 p_{10}(t) + q_1 p_{00}(t) \\
(4) \quad & p'_{11}(t) = -q_1 p_{11}(t) + q_1 p_{01}(t)
\end{aligned}$$

Es gilt: $p_{01}(t) = 1 - p_{00}(t), p_{11}(t) = 1 - p_{10}(t)$
Eingesetzt in (2): $-p'_{00}(t) = -q_0 + q_0 p_{00}(t) + q_0 - q_0 p_{10}(t) = (1)$
also ist (2) überflüssig. Analog: (4) ist überflüssig. Sei

$$y(t) = \begin{pmatrix} p_{00}(t) \\ p_{10}(t) \end{pmatrix}$$

Zu lösen: $y'(t) = Qy(t), y(0) = (1,0)^T$. Die Eigenwerte von Q sind: $\lambda_1 = 0, \lambda_2 = -(q_0 + q_1)$
Eigenvektoren: $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} -q_0 \\ q_1 \end{pmatrix}$

$$\Longrightarrow \; y(t) = \alpha \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \beta \begin{pmatrix} -q_0 \\ q_1 \end{pmatrix} e^{-(q_0+q_1)t}$$

$y(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ liefert $\alpha = \frac{q_1}{q_0+q_1}, \beta = -\frac{1}{q_0+q_1}$
Insgesamt also:

$$\begin{aligned}
p_{00}(t) &= \frac{q_1}{q_0+q_1} + \frac{q_0}{q_0+q_1} e^{-(q_0+q_1)t} \\
p_{10}(t) &= \frac{q_1}{q_0+q_1} - \frac{q_1}{q_0+q_1} e^{-(q_0+q_1)t}
\end{aligned}$$

Hier fehlt ein kleines Bild, das zeigt, wie p_{00} und p_{10} so aussehen.

Wir wollen weiter annehmen, dass die Pfade der Markov-Kette $(X_t)_{t\geq 0}$ in

$$D(S) := \{f : [0,\infty) \longrightarrow S \mid f(t+) = f(t) \; \forall t \geq 0, f(t-) \text{ existiert } \forall t > 0\}$$

liegen (diskrete Topologie auf S). Die Eigenschaft der f in $D(S)$ wird auch mit RCLL oder càdlàg bezeichnet.

Gegeben sei jetzt eine Intensitätsmatrix $Q = (q_{ij}), q_{ij} \in \mathbb{R}$ mit

(Q1) $q_{ij} \geq 0 \; \forall i, j \in S, i \neq j, q_{ii} \leq 0 \; \forall i \in S$

(Q2) $\sum_{j \in S} q_{ij} = 0 \; \forall i \in S$

(Q3) $0 < \sup_{i \in S} |q_{ii}| =: \lambda < \infty$

Satz 8.4
Für eine Matrix Q gelte (Q1)-(Q3). Es sei $N = (N_t)_{t \geq 0}$ ein Poisson-Prozess mit Intensität λ und $Y = (Y_n)_{n \in \mathbb{N}_0}$ eine von N unabhängige Markov-Kette mit Start in $i_0 \in S$ und Übergansmatrix $\tilde{P} = (\tilde{p}_{ij})_{i,j \in S}, \tilde{P} := E + \frac{1}{\lambda} Q$, also

$$\tilde{p}_{ij} = \delta_{ij} + \frac{q_{ij}}{\lambda} \quad \forall i, j \in S.$$

Dann ist $X = (X_t)_{t \geq 0}$ mit $X_t := Y_{N_t} \quad \forall t \geq 0$ eine Markov-Kette mit Start in i_0, Pfaden in $D(S)$ und Intensitätsmatrix Q.

Beweis
Offenbar ist $\tilde{P}$ eine stochastische Matrix, $X_0 = i_0$ und die Pfade in $D(S)$. Für $0 \leq t_0 < t_1 < \cdots < t_{n+1}, \; i_0, \ldots, i_{n+1} \in S$ gilt:

$$\begin{aligned}
&P(X_{t_m} = i_m, 0 \leq m \leq n+1) \\
=& \sum_{k_0, \ldots, k_{n+1} \in \mathbb{N}_0} P(N_{t_m} = k_m, Y_{k_m} = i_m, 0 \leq m \leq n+1) \\
=& \sum_{k_0, \ldots, k_{n+1} \in \mathbb{N}_0} P(N_{t_m} = k_m, 0 \leq m \leq n+1) \cdot P(Y_{k_m} = i_m, 0 \leq m \leq n+1)
\end{aligned}$$

Betrachte die Faktoren:

$$\begin{aligned}
P(N_{t_m} = k_m, 0 \leq m \leq n+1) &= P(N_{t_{n+1} - t_n} = k_{n+1} - k_n) \cdot P(N_{t_m} = k_m, 0 \leq m \leq n) \\
P(Y_{k_m} = i_m, 0 \leq m \leq n+1) &= P(Y_{k_m} = i_m, 0 \leq m \leq n) \cdot \tilde{p}^{(k_{n+1} - k_n)}_{i_n i_{n+1}}
\end{aligned}$$

Sei jetzt $l := k_{n+1} - k_n$, dann

$$\begin{aligned}
&P(X_{t_m} = i_m, 0 \leq m \leq n+1) \\
=&P(X_{t_m} = i_m, 0 \leq m \leq n) \cdot \sum_{l=0}^{\infty} \tilde{p}^{(l)}_{i_n i_{n+1}} \cdot P(N_{t_{n+1} - t_n} = l)
\end{aligned}$$

$\implies$ X ist eine Markov-Kette in stetiger Zeit (und homogen) mit Übergangsmatrizenfunktion:

$$p_{ij}(t) = \sum_{l=0}^{\infty} \tilde{p}^{(l)}_{ij} e^{-\lambda t} \frac{(\lambda t)^l}{l!}$$

Bestimme Ableitungen $p'_{ij}(0) \implies Q$ ist wie gewünscht. ■

Eine "Umkehrung" des Satzes gilt:
Dazu sei $X = (X_t)_{t\geq 0}$ eine Markov-Kette mit Zustandsraum S und Sprungzeiten:

$$\begin{aligned} S_0 &:= 0 \\ S_1 &:= \inf\{t \geq 0 \mid X_t \neq X_0\} \\ S_n &:= \inf\{t > S_{n-1} \mid X_t \neq X_{S_{n-1}}\}, n \geq 2. \end{aligned}$$

Verweildauern: $n \geq 1$:

$$T_n := S_n - S_{n-1}$$

Weiter sei die eingebettete Markov-Kette $Y = (Y_n)$ gegeben durch: $Y_n := X_{S_n}, n \in \mathbb{N}_0$.

Satz 8.5
Es sei $X = (X_t)_{t\geq 0}$ eine Markov-Kette mit Zustandsraum S und Intensitätsmatrix Q, wobei (Q1) bis (Q3) erfüllt seien.

Dann gilt:

a) $Y = (Y_n)$ ist eine (zeitdiskrete) Markov-Kette mit Übergangsmatrix $P = (p_{ij})$, wobei

$$p_{ij} = \begin{cases} \frac{q_{ij}}{q_i} & \text{, falls } i \neq j \\ 0 & \text{, falls } i = j \end{cases}, \text{ falls } q_i > 0$$
$$p_{ij} = \delta_{ij}, \text{ falls } q_i = 0$$

b) $T_1, T_2, \ldots$ sind bedingt unter (Y_n) stochastisch unabhängig mit

$$T_n \sim \text{Exp}(q_{Y_{n-1}})$$

Hier fehlt ein Bild.

Beispiel 8.3
Sei $S = \mathbb{N}_0$ und $Q = (q_{ij})_{i,j\in S}$ mit

$$q_{ij} := \begin{cases} -(i+1)^2, & \text{falls } j = i \\ (i+1)^2, & \text{falls } j = i+1 \\ 0, & \text{sonst.} \end{cases}$$

Es gilt: $Y_n = n$ für $n \in \mathbb{N}$ bei Start in 0 und $\sup_{i\in\mathbb{N}_0} q_i = \sup_{i\in\mathbb{N}_0}(i+1)^2 = \infty$, das heißt dass die Bedingung (Q3) nicht erfüllt ist.

Weiter ist $T_n \sim \exp(n^2)$. Für die Gesamtlebensdauer $T = \sum_{n=1}^{\infty} T_n$ gilt:

$$ET = \sum_{n=1}^{\infty} ET_n = \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} < \infty$$

Also liegen die Pfade nicht in $D(S)$, da sie keinen linksseitigen Grenzwert in T haben.

Definition
Sei $(X_t)_{t\geq 0}$ eine Markov-Kette mit Standardübergangsmatrizenfunktion $\{P(t),\, t \geq 0\}$. Ein Maß $\mu : S \to \mathbb{R}_+$ heißt *invariantes Maß*, falls für alle $t \geq 0$ gilt:

$$\mu = \mu P(t)$$

also für jedes $j \in S$ gilt:

$$\mu(j) = \sum_{i \in S} \mu(i) p_{ij}(t)$$

Gilt $\sum_{j \in S} \mu(j) = 1$, so heißt μ *stationäre Verteilung*.

Satz 8.6
Sei $X = (X_t)_{t\geq 0}$ eine Markov-Kette mit Intensitätsmatrix Q, wobei (Q1) bis (Q3) erfüllt seien. Dann ist μ genau dann ein invariantes Maß, wenn

$$\mu Q = 0$$

gilt.

Beweis
Unter den Voraussetzungen gelten die Vorwärts- und Rückwärtsdifferentialgleichungssysteme:

$$P'(t) = P(t)Q \qquad\qquad P'(t) = QP(t)$$

Damit folgt

$$\begin{aligned} P(t) &= E + \int_0^t P'(s) ds \\ &= E + \int_0^t P(s) ds\, Q \quad (*) \\ &= E + Q \int_0^t P(s) ds \quad (\triangle) \end{aligned}$$

Sei μ ein invariantes Maß für X, also $\forall t \geq 0 : \mu = \mu P(t)$. Aus $(*)$ folgt für $t \geq 0$:

$$\mu = \mu P(t) = \mu + \int_0^t \mu P(s) ds Q = \mu + t \cdot \mu Q$$

Damit gilt $0 = t\mu Q$ und, wegen $t \geq 0$, auch $\mu Q = 0$.

Falls $\mu Q = 0$ gilt, so folgt aus $(\triangle)$: $\forall t \geq 0$: $\mu P(t) = \mu$. ∎

Definition
Sei $X = (X_t)_{t\geq 0}$ eine Markov-Kette mit Zustandsraum S und $i \in S$ mit $q_i < \infty$.

a) i heißt *rekurrent*, falls i rekurrent für die eingebettete Kette (Y_n) ist.

b) i heißt *positiv rekurrent*, falls i rekurrent ist und

$$m_i := E_i[\tilde{T}_i] < \infty$$

wobei $\tilde{T}_i$ der erste Rückkehrzeitpunkt in den Zustand i ist.

c) X heißt *irreduzibel*, falls die eingebettete Kette (Y_n) irreduzibel ist.

Satz 8.7
Sei X eine Markov-Kette mit Intensitätsmatrix Q, wobei (Q1) bis (Q3) erfüllt seien. Sind alle Zustände rekurrent und ist die Markov-Kette irreduzibel, so gilt:

a) $\lim_{t\to\infty} p_{ij}(t) = \frac{1}{m_j q_j}$ für alle $i, j \in S$, unabhängig von $i \in S$.

b) X besitzt genau dann eine stationäre Verteilung $\pi = (\pi_i)_{i\in S}$, wenn es einen positiv rekurrenten Zustand j gibt. In diesem Fall ist $\pi_j = \frac{1}{m_j q_j} > 0$ für jedes $j \in S$ und alle Zustände sind positiv rekurrent.

Beweis
Nicht hier. Wird teilweise in der Übung untersucht. ■

Beispiel 8.4 (Geburts- und Todesprozesse in stetiger Zeit)
Sei X eine Markov-Kette mit $S = \mathbb{N}_0$ und Intensitätsmatrix

$$Q = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & \cdots & & \\ \mu_1 & -(\mu_1+\lambda_1) & \lambda_1 & 0 & \cdots & \\ 0 & \mu_2 & -(\mu_2+\lambda_2) & \lambda_2 & 0 & \cdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}$$

mit $\lambda_0, \lambda_i, \mu_i > 0$. Weiter sei $\sup_{i\in\mathbb{N}}(\lambda_i + \mu_i) < \infty$.

Die Übergangswahrscheinlichkeiten der eingebetteten Kette (Y_n) sind

$$p_{i,i+1} = \frac{\lambda_i}{\mu_i + \lambda_i}, \qquad p_{i,i-1} = \frac{\mu_i}{\mu_i + \lambda_i}, \qquad p_{01} = 1.$$

Also ist X irreduzibel.

Wie sieht eine stationäre Verteilung aus, falls sie existiert? Der Ansatz ist die Gleichung $\pi Q = 0$:

$$\begin{aligned} 0 &= -\lambda_0\pi_0 + \mu_1\pi_1 \\ 0 &= \lambda_0\pi_0 - (\mu_1+\lambda_1)\pi_1 + \mu_2\pi_2 \\ &\vdots \\ 0 &= \lambda_i\pi_i - (\mu_{i+1}\lambda_{i+1})\pi_{i+1} + \mu_{i+2}\pi_{i+2}, \quad i \geq 1 \\ \implies \quad \pi_1 &= \frac{\lambda_0}{\mu_1}\pi_0 \\ \pi_2 &= \frac{1}{\mu_2}\left(\frac{(\mu_1+\lambda_1)\lambda_0}{\mu_1}\pi_0 - \frac{\lambda_0\mu_1}{\mu_1}\pi_0\right) = \frac{\lambda_0\lambda_1}{\mu_1\mu_2}\pi_0 \end{aligned}$$

Dies legt die Vermutung nahe, dass für $i \geq 0$

$$\pi_{i+1} = \frac{\lambda_0 \cdots \lambda_i}{\mu_1 \cdots \mu_{i+1}} \pi_0$$

gilt, welche sich leicht durch Einsetzen bestätigen lässt.

$\pi = (\pi_i)$ kann zu einer stationären Verteilung normiert werden, falls

$$\sum_{i=1}^{\infty} \frac{\lambda_0 \cdots \lambda_i}{\mu_1 \cdots \mu_{i+1}} < \infty. \quad (*)$$

Nach Beispiel 3.3 ist (Y_n) positiv rekurrent, falls

$$\sum_{n=1}^{\infty} \prod_{k=0}^{n} \frac{p_{k,k+1}}{p_{k+1,k}} < \infty \iff \sum_{i=1}^{\infty} \frac{\lambda_0 \cdots \lambda_i}{\mu_1 \cdots \mu_{i+1}} < \infty.$$

Also ist nach Satz 8.7 X positiv rekurrent, falls $(*)$ gilt, und π ist die Grenzverteilung.

Fals $\lambda_i = \lambda$ und $\mu_i = \mu$ ist, dann ist X eine M/M/1-Warteschlange mit stationärer Verteilung (falls $\rho := \frac{\lambda}{\mu} < 1$): $\pi_i = (1-\rho)\rho^i$ für $i \in \mathbb{N}_0$.

III. Die Brownsche Bewegung

9. Definition und erste Eigenschaften

Es sei $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum und $T \subseteq \mathbb{R}$, $T \neq \emptyset$ eine Zeitindexmenge. Wir betrachten jetzt einen stochastischen Prozess $X = (X_t)_{t \in T}$ mit Zustandsraum $\mathbb{R}$, das heißt dass $(X_t)_{t \in T}$ eine Familie von $(\mathcal{F}, \mathfrak{B})$-messbaren Zufallsvariablen ist.

X können wir betrachten als Abbildung

$$X : \Omega \to \mathbb{R}^T = \{f : T \to \mathbb{R}\}.$$

Wir schreiben dabei statt $\big(X(\omega)\big)(t)$ auch $X_t(\omega)$ oder $X(t, \omega)$.

Weiter sei $\pi_t : \mathbb{R}^T \to \mathbb{R}$, $f \mapsto f(t)$ die t-te Projektion. Eine σ-Algebra $\mathfrak{B}^T$ auf $\mathbb{R}^T$ ist gegeben durch

$$\begin{aligned}
\mathfrak{B}^T &:= \sigma(\{\pi_t \mid t \in T\}) = \bigoplus_{t \in T} \mathfrak{B} \\
&= \sigma(\{f : T \to \mathbb{R} \mid f(t_1) \in B_1, \dots, f(t_n) \in B_n,\, t_i \in T,\, B_i \in \mathfrak{B},\, i = 1, \dots, n\}) \\
&= \sigma(\{f : T \to \mathbb{R} \mid (f(t_1), \dots, f(t_n)) \in B,\, t_i \in T,\, i = 1, \dots, n,\, B \in \mathfrak{B}(\mathbb{R}^n)\})
\end{aligned}$$

X ist genau dann $(\mathcal{F}, \mathfrak{B}^T)$-messbar, wenn $\pi_t \circ X$ für alle $t \in T$ $(\mathcal{F}, \mathfrak{B})$-messbar ist (Siehe Stochastik II, Übungsaufgabe 21), also ist X $(\mathcal{F}, \mathfrak{B}^T)$-messbar.

Jede Menge $A \in \mathfrak{B}^T$ besitzt folgende Struktur: Es gibt abzählbar viele Zeitpunkte $t_1, t_2, \ldots \in T$ und ein $B \in \mathfrak{B}(\mathbb{R}^\infty)$, so dass

$$A = \{f : T \to \mathbb{R} \mid \big(f(t_1), f(t_2), \dots\big) \in B\}.$$

X induziert ein Wahrscheinlichkeitsmaß P^X auf $(\mathbb{R}^T, \mathfrak{B}^T)$.

Im Folgenden sei $T_0 := \{t_1, \dots, t_n\}$ mit $t_1 < t_2 < \cdots < t_n$, $t_i \in T$, $n \in \mathbb{N}$. Die Abbildung

$$\pi_{T_0} : \mathbb{R}^T \to \mathbb{R}^{T_0}, \quad f \mapsto (f(t_1), \dots, f(t_n))$$

sei die Projektion auf die T_0-Koordinaten.

Die Verteilung von $(X_{t_1}, \dots, X_{t_n})$ ist dann ein Wahrscheinlichkeitsmaß auf $(\mathbb{R}^{T_0}, \mathfrak{B}^{T_0})$ und ergibt sich als Bild von P^X unter π_{T_0}.

Die Gesamtheit dieser Verteilungen nennt man Familie der endlich-dimensionalen Verteilungen des Prozesses (engl.: finite dimensional distribution, Fidis).

Definition
Ein Prozess, bei dem alle endlich-dimensionale Verteilungen Normalverteilungen sind, heißt *Gauss-Prozess*.

Ist $(X_t)_{t\in T}$ ein Gauss-Prozess, so nennen wir die Abbildung

$$t \mapsto EX_t$$

die *Erwartungswertfunktion* und

$$(s,t) \mapsto \operatorname{Cov}(X_s, X_t)$$

die *Kovarianzfunktion* zu X.

Bemerkung: Beim Gauss-Prozess definiert die Erwartungswertfunktion und die Kovarianzfunktion bereits alle endlich-dimensionalen Verteilungen. Da für $T_0 = \{t_1, \ldots, t_n\}$ und $B_1, \ldots, B_n \in \mathfrak{B}$ die Mengen

$$\pi_{T_0}^{-1}(B_1 \times \cdots \times B_n) = \{f \in B^T \mid f(t_1) \in B_1, \ldots, f(t_n) \in B_n\}$$

ein durchschnittsstabiles Erzeugendensystem von $\mathfrak{B}^T$ bilden, ist P^X auf $(\mathbb{R}^T, \mathfrak{B}^T)$ durch die endlich-dimensionalen Verteilungen bzw. durch die Erwartungswertfunktion und Kovarianzfunktion festgelegt.

Im Folgenden wollen wir weiter voraussetzen, dass die Pfade von X stetig sind, also

$$X : \Omega \to C(T) \coloneqq \{f : T \to \mathbb{R} \mid f \text{ stetig}\}$$

wobei $T = [0,1]$ oder $T = [0, \infty)$ ist.

Jetzt ist $\pi_t : C(T) \to \mathbb{R}$ und $\mathfrak{B}(C(T)) \coloneqq \sigma(\{\pi_t \mid t \in T\})$. Auch hier ist die Verteilung eines Prozesses durch seine endlich-dimensionalen Verteilungen festgelegt.

Definition
Es sei $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum, $T = \mathbb{R}_+$ oder $T = [0, R]$ mit $R > 0$, $(\mathcal{F}_t)_{t\in T}$ eine Filtration und $(B_t)_{t\in T}$ ein $(\mathcal{F}_t)_{t\in T}$ -adaptiver Prozess, d.h. B_t ist $(\mathcal{F}_t, \mathfrak{B})$-messbar. Gilt

a) $P(B_0 = 0) = 1$

b) P-fast-alle Pfade von $(B_t)_{t\in T}$ sind stetig.

c) Für alle $s, t \in T$ mit $s < t$ ist $B_t - B_s$ unabhängig von $\mathcal{F}_s$ (d.h. $B_t - B_s$ unabhängig von 1_A, $A \in \mathcal{F}_s$) und $\mathcal{N}(0, t-s)$-verteilt,

so heißt $(B_t, \mathcal{F}_t)_{t\in T}$ eine (eindimensionale) *Brownsche Bewegung* oder auch *Wiener-Prozess*.

Bemerkung: a) Falls $\mathcal{F}_t = \sigma(\{B_s,\, s \leq t\})$, so spricht man von *der* Brownschen Bewegung.

b) Wir können B auf einer Nullmenge abändern zu $\tilde{B}$ und $\tilde{B}$ ist wieder eine Brownsche Bewegung. Dies hat aber Auswirkungen auf die Filtration.

c) Eine Brownsche Bewegung hat unabhängige Zuwächse (folgt aus Teil c) der Definition).

d) 1827 beobachtet Brown die Bewegung von Pollen in einer Flüssigkeit. 1900 verwendet Bachelier die Brownsche Bewegung zur Modellierung von Aktienkursen. 1905 beschrieb Einstein die Molekülbewegung mit der Brownschen Bewegung. 192? baute Wieder das mathematische Fundament der Brownschen Bewegung.

Lemma 9.1
Die Brownsche Bewegung $(B_t, \mathcal{F}_t)_{t \in T}$ ist ein $(\mathcal{F}_t)$-Martingal.

Beweis
Offenbar ist $E|B_t| < \infty$. Für $s \leq t$ gilt:

$$\begin{aligned} E[B_t \mid \mathcal{F}_s] &= E[B_s + (B_t - B_s) \mid \mathcal{F}_s] \\ &= B_s + E[B_t - B_s \mid \mathcal{F}_s] \\ &= B_s + 0 \end{aligned}$$

■

Satz 9.2
Die Brownsche Bewegung $(B_t)_{t \in T}$ ist ein Gauss-Prozess mit Erwartungswertfunktion $EB_t = 0$ und Kovarianzfunktion $\operatorname{Cov}(B_s, B_t) = s \wedge t = \min\{s, t\}$.

Ist umgekehrt $(X_t)_{t \in T}$ ein Gauss-Prozess mit $EX_t = 0$ und $\operatorname{Cov}(X_s, X_t) = s \wedge t$ und fast-sicher stetigen Pfaden, so ist $(X_t)_{t \in T}$ eine Brownsche Bewegung.

Beweis
Ist $(B_t)_{t \in T}$ die Brownsche Bewegung, so gilt $B_0 = 0$ und $B_t - B_0 = B_t \sim \mathcal{N}(0, t)$, also ist $EB_t = 0$ für alle $t \in T$. Da Zuwächse unabhängig sind folgt für $s \leq t$:

$$\begin{aligned} \operatorname{Cov}(B_s, B_t) &= EB_sB_t - EB_sEB_t \\ &= EB_s(B_s + (B_t - B_s)) - 0 \\ &= EB_s^2 + EB_s(B_t - Bs) \\ &= EB_s^2 + 0 \\ &= \operatorname{Var}(B_s) \\ &= s \\ &= s \wedge t \end{aligned}$$

Außerdem ist für $0 \leq t_1 \leq \cdots \leq t_n$ der Vektor $(B_{t_1}, B_{t_2} - B_{t_1}, \ldots, B_{t_n} - B_{t_{n-1}}) \sim \mathcal{N}(0, \Sigma)$ multivariat normalverteilt. Aus diesem Vektor erhalten wir durch Multiplikation mit einer geeigneten Matrix den Vektor $(B_{t_1}, \ldots, B_{t_n})$, der demnach auch normalverteilt ist.

Sei jetzt $(X_t)_{t \in T}$ ein Gauss-Prozess mit stetigen Pfaden. Da $X_0 \sim \mathcal{N}(0, 0)$ ist $X_0 = 0$. Also bleibt c) zu zeigen mit der natürlichen Filtration. Sei $0 \leq s_1 \leq s_2 \leq \cdots \leq s_n \leq s \leq t$. Da X ein Gauss-Prozess ist, gilt

$$(X_{s_1}, X_{s_2}, \ldots, X_{s_n}, X_s, X_t) \sim \mathcal{N}(0, \cdot).$$

Es gibt eine Matrix, die diesen Vektor in den Vektor

$$(X_{s_1}, X_{s_2} - X_{s_1}, \ldots, X_{s_n} - X_{s_{n-1}}, X_t - X_s) \sim \mathcal{N}(0, \Sigma) \quad (*)$$

transformiert. Man kann nachrechnen, dass Σ Diagonalform hat, das heißt, dass $X_t - X_s$ von $(X_{s_1}, X_{s_2} - Xs_1, \ldots, X_{s_n} - X_{s_{n-1}})$ und damit auch von $(X_{s_1}, \ldots, X_{s_n})$ unabhängig ist. Da

$\{X_{s_1} \in B_1, \ldots, X_{s_n} \in B_n\}$ für $n \in \mathbb{N}$, $0 \leq s_1 \leq \cdots \leq s_n \leq s$ und $B_1, \ldots, B_n \in \mathfrak{B}$ ein durchschnittsstabiles Erzeugendensystem von $\mathcal{F}_s$ bilden (siehe Henze, Stochastik II, Seite 103), ist $X_t - X_s$ unabhängig von $\mathcal{F}_s$. In $(*)$ kann man ablesen, dass $X_t - X_s \sim \mathcal{N}(0, t-s)$. ■

Satz 9.3
Ist $(B_t)_{t\geq 0}$ eine Brownsche Bewegung, so auch

a) $(-B_t)_{t\geq 0}$

b) $(B_{a+t} - B_a)_{t\geq 0}$ mit festem $a \geq 0$.

c) $(cB_{\frac{t}{c^2}})_{t\geq 0}$ mit festem $c \neq 0$. In diesem Sinne ist die Brownsche Bewegung *selbstähnlich*.

d) $(\tilde{B}_t)_{t\geq 0}$ mit $\tilde{B}_t \coloneqq t \cdot B_{\frac{1}{t}}$ für $t > 0$ und $\tilde{B}_0 \coloneqq 0$.

Beweis
Wir beweisen nur Teil d):

Wir verwenden den Satz 9.2. Es seien $t_1, \ldots, t_n > 0$. Dann ist

$$(B_{\frac{1}{t_1}}, \ldots, B_{\frac{1}{t_n}})$$

normalverteilt, und damit ist auch

$$(t_1 \cdot B_{\frac{1}{t_1}}, \ldots, t_n \cdot B_{\frac{1}{t_n}})$$

normalverteilt, da dies eine lineare Transformation ist. Also ist $\tilde{B}$ ein Gauss-Prozess.

Weiter gilt für $t > 0$:

$$E\tilde{B}_t = tEB_{\frac{1}{t}} = t \cdot 0 = 0$$

und $E\tilde{B}_0 = E0 = 0$. Sei $s, t > 0$. Es ist

$$\begin{aligned} \operatorname{Cov}(\tilde{B}_s, \tilde{B}_t) &= \operatorname{Cov}(sB_{\frac{1}{s}}, tB_{\frac{1}{t}}) \\ &= s \cdot t \cdot (\frac{1}{s} \wedge \frac{1}{t}) \\ &= s \wedge t \end{aligned}$$

und dies gilt auch für $s = 0$ und für $t = 0$.

Es bleibt zu zeigen, dass $\tilde{B}$ P-fast-sicher stetige Pfade hat. Nach Voraussetzung gibt es eine Menge $A \in \mathcal{F}$ mit $P(A) = 1$ und $t \mapsto B_t(\omega)$ für alle $\omega \in A$ stetig auf $[0, \infty)$. Für diese $\omega \in A$ ist dann auch $t \mapsto \tilde{B}_t(\omega)$ stetig auf $(0, \infty)$. Insbesondere gilt für $\omega \in A$ und $m \in \mathbb{N}$:

$$\sup_{0<t\leq \frac{1}{m}} |\tilde{B}_t(\omega)| = \sup_{q\in\mathbb{Q}\cap(0,\frac{1}{m}]} |\tilde{B}_q(\omega)|$$

Damit ist $t \mapsto \tilde{B}_t(\omega)$ in $t = 0$ stetig für alle $\omega \in \tilde{F}$ mit

$$\tilde{F} \coloneqq \bigcap_{n=1}^{\infty} \bigcup_{m=1}^{\infty} \bigcap_{q\in\mathbb{Q}\cap(0,\frac{1}{m}]} \{|\tilde{B}_q| \leq \frac{1}{n}\}.$$

Weiter gilt

$$\begin{aligned}
P(\tilde{F}) &= \lim_{n\to\infty} P\Big(\bigcup_{m=1}^{\infty} \bigcap_{q\in\mathbb{Q}\cap(0,\frac{1}{m}]} \{|\tilde{B}_q| \leq \frac{1}{n}\}\Big) \\
&= \lim_{n\to\infty} \lim_{m\to\infty} P\Big(\bigcap_{q\in\mathbb{Q}\cap(0,\frac{1}{m}]} \{|\tilde{B}_q| \leq \frac{1}{n}\}\Big) \\
&= \lim_{n\to\infty} \lim_{m\to\infty} \lim_{K\to\infty} P\big(\bigcap_{j=1}^{K} \{|\tilde{B}_{q_j}| \leq \frac{1}{n}\}\big) \\
&= \lim_{n\to\infty} \lim_{m\to\infty} \lim_{K\to\infty} P\big(\bigcap_{j=1}^{K} \{|B_{q_j}| \leq \frac{1}{n}\}\big) \\
&= P(F)
\end{aligned}$$

mit

$$F := \bigcap_{n=1}^{\infty} \bigcup_{m=1}^{\infty} \bigcap_{q\in\mathbb{Q}\cap(0,\frac{1}{m}]} \{|B_q| \leq \frac{1}{n}\}.$$

Wegen $A \subseteq F$ gilt $P(\tilde{F}) = 1$, damit $P(\tilde{F} \cap A) = 1$. Also sind P-fast-sicher alle Pfade von $\tilde{B}$ auch stetig in $t = 0$. ■

10. Existenz

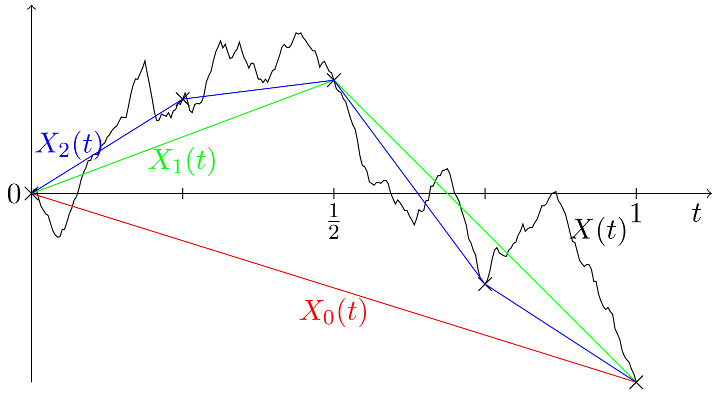
Die Existenz der Brownschen Bewegung wurde erst 1923 von Norbert Wiener bewiesen. Heute gibt es drei Standardverfahren, diese zu beweisen. Wir wählen hier einen funktional-analytischen Zugang.

Satz 10.1
Es gibt einen Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$, auf dem ein stochastischer Prozess $(B_t)_{0\leq t\leq 1}$ definiert werden kann, mit den Eigenschaften

a) $B_0(\omega) = 0$ für alle $\omega \in \Omega$

b) $t \mapsto B_t(\omega)$ ist stetig für alle $\omega \in \Omega$

c) Für alle $0 \leq s \leq t \leq 1$ ist $B_t - B_s$ unabhängig von $\sigma(\{B_u,\, u \leq s\})$ und $\mathcal{N}(0, t-s)$-verteilt.



Beweis
Es sei $L^2 = L^2([0,1], \mathfrak{B}_{[0,1]}, \lambda_{[0,1]}) := \{f : [0,1] \to \mathbb{R} \mid f \text{ messbar}, \int_0^1 f^2(x)dx < \infty\}$ beziehungsweise deren Äquivalenzklassen. Versehen mit

$$\langle f, g\rangle := \int_0^1 f(x) \cdot g(x)dx$$

ist L^2 ein Hilbertraum. Weiter sei für alle $n \in \mathbb{N}_0$

$$S_n := \{(n,k) \mid 1 \leq k \leq 2^n, k \text{ ungerade}\}$$

und $S := \bigcup_{n=0}^{\infty} S_n$. Definiere für $(n,k) \in S$

$$g_{01}(t) := 1$$
$$g_{nk}(t) := 2^{\frac{(n-1)}{2}} \left(1_{[(k-1)2^{-n},\, k2^{-n})}(t) - 1_{[k2^{-n},\, (k+1)2^{-n})}(t)\right) \quad \text{für } n \neq 0,\ 0 \leq t \leq 1.$$

Diese Funktionen bilden eine Orthonormalbasis (die *Haar-Basis*) von L^2, das heißt $\|g_{nk}\| = 1$, $\langle g_{nk}, g_{ml}\rangle = 0$ für $(n,k) \neq (m,l)$ und es gilt die sogenannte Parsevalsche Gleichung:

$$\langle f, g\rangle = \sum_{n=0}^{\infty} \sum_{(n,k)\in S_n} \langle f, g_{nk}\rangle \langle g, g_{nk}\rangle$$

Es sei nun $\{Z_{nk} \mid (n,k) \in S\}$ eine Familie von unabhängigen $\mathcal{N}(0,1)$-verteilten Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$.

Wir definieren $(X_n(t))_{0\leq t\leq 1}$, $n \in \mathbb{N}_0$, durch

$$X_n(t) := \sum_{m=0}^{n} \sum_{(m,k)\in S_m} f_{mk}(t) \cdot Z_{mk}$$

wobei für alle $t \in [0,1], (m,k) \in S$

$$f_{mk}(t) := \int_0^t g_{mk}(s)ds$$

die sogenannte *Schauder-Funktionen* sind. Diese $(X_n(t))_{0\leq t\leq 1}$ sind stetige stochastische Prozesse auf $(\Omega, \mathcal{F}, P)$.

Es gilt $0 \leq f_{nk} \leq 2^{-\frac{(n+1)}{2}}$. Wir behaupten, dass für P-fast-alle $\omega \in \Omega$ die $(X_n(t,\omega))_{0\leq t\leq 1}$ für $n \to \infty$ gleichmäßig konvergiert. Dann wäre die Grenzfunktion X automatisch stetig.

Für $Z \sim \mathcal{N}(0,1)$ verwenden wir ohne Beweis die Abschätzung

$$P(|Z| \geq a) \leq \sqrt{\frac{2}{\pi}}\frac{1}{a}e^{-\frac{a^2}{2}} \text{ für alle } a > 0.$$

Sei $\|f\|_\infty := \sup_{0\leq t\leq 1} |f(t)|$ für $f \in C[0,1]$. Dann gilt für alle $n \in \mathbb{N}$:

$$\begin{aligned}
P(\|X_n - X_{n-1}\|_\infty > a_n) &= P(\|\sum_{(n,k)\in S_n} f_{nk}(t)Z_{nk}\|_\infty > a_n)\\
&= P(\sup_{(n,k)\in S_n} |Z_{nk}| > 2^{\frac{(n+1)}{2}}a_n)\\
&\leq \sum_{(n,k)\in S_n} P(|Z_{nk}| > 2^{\frac{(n+1)}{2}}a_n)\\
&= |S_n| \cdot P(|Z_{01}| > 2^{\frac{(n+1)}{2}}a_n)\\
&\leq 2^n \cdot P(|Z_{01}| > 2^{\frac{(n+1)}{2}}a_n)\\
&\leq 2^n \cdot \sqrt{\frac{2}{\pi}}\frac{1}{a_n}2^{-\frac{(n+1)}{2}}e^{-\frac{1}{2}2^{(n+1)}a_n^2}
\end{aligned}$$

mit $a_n := \sqrt{n2^{-n}}$ wird daraus

$$\begin{aligned}
&= \sqrt{\frac{2}{\pi}}2^n 2^{-\frac{n+1}{2}}n^{-\frac{1}{2}}2^{\frac{n}{2}}e^{-\frac{1}{2}2^{n+1}n2^{-n}}\\
&= \frac{1}{\sqrt{\pi}}2^n n^{-\frac{1}{2}}e^{-n}
\end{aligned}$$

das heißt dass mit $A_n := \{\|X_n - X_{n-1}\|_\infty > (n2^{-n})^{\frac{1}{2}}\}$ gilt

$$\sum_{n=1}^\infty P(A_n) < \infty.$$

Nach dem Lemma von Borel-Cantelli gilt dann

$$P(\underbrace{\limsup_{n\to\infty} A_n}_{=:N}) = 0.$$

Das heißt, für jedes $\omega \notin N$ gibt es einen Index $n_0 = n_0(\omega) \in \mathbb{N}$, sodass für alle $m, n \geq n_0$ gilt:

$$\begin{aligned}
\|X_m(\omega) - X_n(\omega)\|_\infty &\leq \sum_{k=(m\wedge n)+1}^{m\vee n} \|X_k(\omega) - X_{k-1}(\omega)\|_\infty\\
&\leq \sum_{k=(m\wedge n)+1}^{\infty} \sqrt{k2^{-k}} \to 0 \text{ für } m,n\to\infty
\end{aligned}$$

Also ist $(X_n(\omega))_{n\in\mathbb{N}}$ eine Cauchy-Folge in $(C[0,1], \|\cdot\|_\infty)$ für $\omega \notin N$. Dieser Raum ist vollständig, also existiert eine Grenzfunktion $X(\omega)$, wobei wir $X(\omega) := 0$ für $\omega \in N$ setzen.

Der Prozess $(X_t)_{0\leq t\leq 1}$ hat P-fast-sicher stetige Pfade. Wir zeigen nun, dass dieser Prozess ein Gauss-Prozess wie in Satz 9.2 ist:

Es gilt jetzt für alle $n \in \mathbb{N}$, $0 \le s, t \le 1$

$$\begin{aligned} EX_n(t) &= \sum_{m=0}^{n} \sum_{(m,k)\in S_m} f_{nk}(t) \cdot EZ_{mk} \\ &= \sum_{m=0}^{n} \sum_{(m,k)\in S_m} f_{nk}(t) \cdot 0 \\ &= 0 \end{aligned}$$

und, wobei die gemischten Terme verschwinden, da die Z_{nk} unabhängig sind,

$$\begin{aligned} EX_n(t)X_n(s) &= \sum_{m=0}^{n} \sum_{(m,k)\in S_m} f_{nk}(t) f_{nk}(s) E{Z_{mk}}^2 \\ &= \sum_{m=0}^{n} \sum_{(m,k)\in S_m} f_{nk}(t) f_{nk}(s) \end{aligned}$$

Offenbar ist $f_{nk}(s) = \langle 1_{[0,s]}, g_{nk} \rangle$ und mit der Parsevalschen Gleichung folgt:

$$(*) \qquad \begin{aligned} \lim_{n\to\infty} EX_n(t)X_n(s) &= \sum_{m=0}^{\infty} \sum_{(m,k)\in S_m} \langle 1_{[0,s]}, g_{nk} \rangle \langle 1_{[0,t]}, g_{nk} \rangle \\ &= \langle 1_{[0,s]}, 1_{[0,t]} \rangle \\ &= s \wedge t \end{aligned}$$

Sei $t_1, \dots, t_k \in [0,1]$. Es gilt

$$(X_n(t_1), \dots, X_n(t_k)) \to (X(t_1), \dots, X(t_k)) \quad P\text{-f.s.}$$

woraus die Konvergenz in Verteilung folgt und, dazu äquivalent, die Konvergenz der charakteristischen Funktionen $\varphi_n : \mathbb{R}^k \to \mathbb{C}$, $\varphi_n(\theta) = E \exp(i \sum_{\nu=1}^{k} \theta_\nu X_n(t_\nu))$, also

$$\varphi_n(\theta) \xrightarrow{n\to\infty} \varphi(\theta) = \exp(i \sum_{\nu=1}^{k} \theta_\nu X(t_\nu)).$$

Es gilt $\varphi_n(\theta) = \exp(-\frac{1}{2}\theta^\top \Sigma_n \theta)$ mit $(\Sigma_n)_{i,j} = EX_n(t_i)X_n(t_j)$. Mit $(*)$ folgt dann

$$\lim_{n\to\infty} (\Sigma_n)_{i,j} = t_i \wedge t_j =: (\Sigma)_{i,j}$$

also

$$\varphi_n(\theta) \to \varphi(\theta) = \exp(-\frac{1}{2}\theta^\top \Sigma \theta)$$

Daraus folgt, dass $(X_{t_1}, \dots, X_{t_k})$ normalverteilt ist mit $EX_t = 0$ für alle $t \in [0,1]$ und $\mathrm{Cov}(X_s, X_t) = s \wedge t$ für alle $s, t \in [0,1]$. Mit Satz 9.2 folgt die Behauptung. ■

Bemerkung: Auf $(C[0,1], \mathfrak{B}(C[0,1]))$ definiert die Verteilung P^B von $(B_t)_{0\le t\le 1}$ ein Wahrscheinlichkeitsmaß, das sogenannte Wiener-Maß. Mit $(\Omega, \mathcal{F}, P) = (C[0,1], \mathfrak{B}(C[0,1]), P^B)$ und $B_t = \pi_t$ hat man ein explizites Modell für $(B_t)_{0\le t\le 1}$, die sogenannte Kanonische Konstruktion.

11. Pfadeigenschaften

Satz 11.1
Ist $(B_t)_{t\geq 0}$ eine Brownsche Bewegung auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$, so gilt für P-fast-alle $\omega \in \Omega$, dass

$$\sup_{t\geq 0} B_t(\omega) = \infty, \qquad \inf_{t\geq 0} B_t(\omega) = -\infty.$$

Beweis
Sei $Z := \sup_{t\geq 0} B_t$. Nach Satz 9.3 c) gilt für alle $c > 0$: $Z \overset{\mathrm{d}}{=} c \cdot Z$. Da $Z \geq 0$ folgt $P(Z \in \{0, \infty\}) = 1$.

Nach Satz 9.3 b) ist $(B_{1+t} - B_1)_{t\geq 0}$ wieder eine Brownsche Bewegung, das Supremum davon ist also wieder P-f.s. 0 oder ∞.

$$\begin{aligned} P(Z = 0) &\leq P(B_1 \leq 0 \text{ und } B_t \leq 0 \text{ für alle } t \geq 1) \\ &= P(B_1 \leq 0 \text{ und } \sup_{t\geq 0}(B_{1+t} - B_1) \leq |B_1|) \\ &= P(B_1 \leq 0 \text{ und } \sup_{t\geq 0}(B_{1+t} - B_1) = 0) \end{aligned}$$

$(B_{1+t} - B_1)$ ist unabhängig von B_1, also gilt weiter

$$\begin{aligned} &= P(B_1 \leq 0) \cdot P(\sup_{t\geq 0}(B_{1+t} - B_1) = 0) \\ &= \frac{1}{2} \cdot P(Z = 0) \end{aligned}$$

Das heißt, dass $P(Z = 0) \leq \frac{1}{2} P(Z = 0)$, also muss $P(Z = 0) = 0$ und $P(\sup_{t\geq 0} B_t = \infty) = 1$ gelten.

Die Aussage für das Infimum folgt dann aus Satz 9.3 a). ■

Bemerkung: Für alle $a \in \mathbb{R}$ ist $\{t \geq 0 \mid B_t = a\}$ P-fast-sicher unbeschränkt, da stetige Funktionen auf kompakten Intervallen beschränkt sind. Insbesondere kehrt $(B_t)_{t\geq 0}$ mit Wahrscheinlichkeit 1 unendlich oft nach 0 zurück, und es gibt für P-fast-alle $\omega \in \Omega$ eine Folge $(t_n)_{n\in\mathbb{N}} = (t_n(\omega))_{n\in\mathbb{N}}$ mit $t_n \to \infty$ und $B_{t_n}(\omega) = 0$ für alle $n \in \mathbb{N}$. Mit der Zeitumkehr $\tilde{B}_t = tB_{\frac{1}{t}}$, $\tilde{B}_0 = 0$ folgt, dass für P-fast alle $\omega \in \Omega$ die Nullstellen des Pfades $t \to B_t(\omega)$ einen Häufungspunkt in 0 haben.

Definition
Eine Funktion $f : \mathbb{R}_+ \to \mathbb{R}$ heißt Lipschitz-stetig in $x \in \mathbb{R}_+$ wenn es ein $\delta > 0$ und ein $K < \infty$ gibt mit

$$|f(x) - f(y)| \leq K|x - y|$$

für alle $y \in [x - \delta, x + \delta] \cap \mathbb{R}_+$.

Satz 11.2
P-fast-alle Pfade einer Brownschen Bewegung sind in keinem Punkt Lipschitz-stetig.

Beweis
Es seien $0 < a < b < \infty$ fest und $K > 0$. Für $n \in \mathbb{N}$ mit $a - \frac{2}{n} > 0$ sei

$$A_n := \{\omega \in \Omega \mid \exists s \in (a,b)\ \forall t \in [s - \frac{2}{n}, s + \frac{2}{n}] : |B_t(\omega) - B_s(\omega)| \leq K|t-s|\}$$

Für jedes $I = [s - \frac{2}{n}, s + \frac{2}{n}]$ gibt es ein $k \in \mathbb{N}$ mit $\frac{k-2}{n}, \frac{k-1}{n}, \frac{k}{n}, \frac{k+1}{n} \in I$. Auf A_n gilt für $j = k-1, k, k+1$

$$\begin{aligned}
\underbrace{|B(\frac{j}{n}) - B(\frac{j-1}{n})|}_{=:\Delta_{n,j}} &\leq |B(\frac{j}{n}) - B(s)| + |B(s) - B(\frac{j-1}{n})| \\
&\leq K \cdot \frac{2}{n} + K \cdot \frac{2}{n} \\
&= \frac{4K}{n}
\end{aligned}$$

Also ist

$$A_n \subseteq \bigcup_{k=\lfloor na \rfloor}^{\lceil nb+1 \rceil} \{\omega \in \Omega \mid \Delta_{n,j}(\omega) \leq \frac{4K}{n} \text{ für } j = k-1, k, k+1\}$$

Die Zuwächse der Brownschen Bewegung sind unabhängig, also folgt:

$$\begin{aligned}
P(A_n) &\leq \sum_{k=\lfloor na \rfloor}^{\lceil nb+1 \rceil} P\Big(\Delta_{n,k-1} \leq \frac{4K}{n}\Big) \cdot P\Big(\Delta_{n,k} \leq \frac{4K}{n}\Big) \cdot P\Big(\Delta_{n,k+1} \leq \frac{4K}{n}\Big) \\
&\leq \lceil nb + 1 \rceil \cdot P\Big(\Delta_{n,1} \leq \frac{4K}{n}\Big)^3
\end{aligned}$$

Da $\sqrt{n}(B(\frac{1}{n}) - B(\frac{0}{n})) \sim \mathcal{N}(0,1)$ folgt

$$\begin{aligned}
P\Big(\Delta_{n,1} \leq \frac{4k}{n}\Big) &= P\Big(\Delta_{n,1}\sqrt{n} \leq \frac{4k}{\sqrt{n}}\Big) \\
&= \int_{-\frac{4k}{\sqrt{n}}}^{\frac{4k}{\sqrt{n}}} \frac{1}{\sqrt{2\pi}} e^{\frac{x^2}{2}} dx \leq \frac{8k}{\sqrt{2\pi n}}
\end{aligned}$$

und insgesamt $P(A_n) \longrightarrow 0$ für $n \longrightarrow \infty$.
Sei $n_0(a) := \inf \{n \in \mathbb{N} \mid a - \frac{2}{n} > 0\}$. Wegen $A_n \subset A_{n+1}$ folgt für $A := \bigcup_{n=n_0}^{\infty} A_n$, dass $P(A) = 0$.
$A = A(k)$ hängt von k ab. $N := \bigcup_{k \in \mathbb{N}} A(k)$ ist wieder eine Nullmenge, so dass $\forall \omega \notin N$ gilt: zu keinem $s \in (a,b)$ und keinem $\delta > 0$, $\exists k < \infty$ mit

$$|B_t(\omega) - B_s(\omega)| \leq k|t-s| \text{ mit } t \in [s-\delta, s+\delta]$$

$\Longrightarrow$ Die Pfade, die in einem $s \in (a,b)$ lipschitzstetig sind, sind in einer P-Nullmenge.

$$N = \mathcal{N}(a,b) \implies \tilde{N} = \bigcup_{\substack{k \in \mathbb{N} \\ a = \frac{1}{k}, b = k}} \mathcal{N}(a,b) \text{ ist wieder Nullmenge}$$

■

$\Longrightarrow$ P-f.a. Pfade sind in keinem $s > 0$ lipschitzstetig. Bleibt $s = 0$ (Übung)

Bemerkung: Ist eine Funktion $f : \mathbb{R}_+ \longrightarrow \mathbb{R}$ in $s \in \mathbb{R}$ differenzierbar, so existiert

$$\lim_{t \to s} \frac{f(t) - f(s)}{t - s}$$

und f wäre in s lipschitzstetig. Satz 11.2 impliziert also, dass P-f.a. Pfade der Brownschen Bewegung in keinem Punkt differenzierbar sind.

Definition
Sei $f : \mathbb{R}_+ \longrightarrow \mathbb{R}$. Die Totalvariation $V_a^b f$ von f auf $[a,b] \subset \mathbb{R}_+$ ist

$$V_a^b f = \sup_{\mathcal{Z}} \sum_{j=1}^{n} |f(t_j) - f(t_{j-1})|,$$

wobei $\mathcal{Z} = \{a = t_0 < t_1 < \cdots < t_n = b\}$ eine Zerlegung von $[a,b]$ ist.

Bemerkung: Funktionen mit endlicher Totalvariation sind fast überall differenzierbar. Also haben die Pfade der Brownschen Bewegung P-f.s. unbeschränkte Totalvariation, d.h. auch kleine Stücke sind "unendlich lang".

Definition
Sind die ZV $X, X_1, X_2, \dots \in L^2(\Omega, \mathcal{F}, P)$, so heißt die Folge $(X_n)_{n\in\mathbb{N}}$ im *quadratischen Mittel konvergent* gegen X ($X_n \xrightarrow{L^2} X$), falls gilt:

$$\lim_{n \to \infty} E(X_n - X)^2 = 0$$

Bemerkung: $X_n \xrightarrow{L^2} X \implies X_n \xrightarrow{P} X$.

Satz 11.3
Es sei $(\mathcal{Z}_n), \mathcal{Z}_n = \{0 = t_{n_0} \leq t_{n_1} \leq \cdots \leq t_{n_{k_n}} = t\}$ eine Folge von Zerlegungen von $[0,t]$ mit Feinheitsgraden

$$\delta(\mathcal{Z}_n) = \max_{j=1,\dots,k_n} |t_{n_j} - t_{n_{j-1}}|$$

so dass $\lim_{n\to\infty} \delta(\mathcal{Z}_n) = 0$. Dann gilt mit $n \to \infty$:

$$\underbrace{\sum_{j=1}^{k_n} (B(t_{n_j}) - B(t_{n_{j-1}}))^2}_{=:V_{\mathcal{Z}_n}} \xrightarrow{L^2} t$$

Beweis
Zu zeigen ist: $E(V_{\mathcal{Z}_n} - t)^2 \to 0$ für $n \to \infty$. Es gilt:

(1) $EV_{\mathcal{Z}_n} = \sum_{j=1}^{k_n}(t_{n_j} - t_{n_{j-1}}) = t$

(2) $E(B(t_{n_j}) - B(t_{n_{j-1}}))^4 = 3(t_{n_j} - t_{n_{j-1}})^2$

Insgesamt:

$$\begin{aligned}
EV_{\mathcal{Z}_n}^2 &= \sum_{j=1}^{k_n} E(B(t_{n_j}) - B(t_{n_{j-1}}))^4 + 2\sum_{j<l} E(B(t_{n_j}) - B(t_{n_{j-1}}))^2 \cdot E(B(t_{n_l}) - B(t_{n_{l-1}}))^2 \\
&= 3\sum_{j=1}^{k_n}(t_{n_j} - t_{n_{j-1}})^2 + 2\sum_{j<l}(t_{n_j} - t_{n_{j-1}})(t_{n_l} - t_{n_{l-1}}) \\
&= \left(\sum_{j=1}^{k_n}(t_{n_j} - t_{n_{j-1}})\right)^2 + 2\sum_{j=1}^{k_n}(t_{n_j} - t_{n_{j-1}})^2 \\
&= t^2 + 2\sum_{j=1}^{k_n}(t_{n_j} - t_{n_{j-1}})^2
\end{aligned}$$

$$\begin{aligned}
\implies E(V_{\mathcal{Z}_n} - t)^2 &= EV_{\mathcal{Z}_n}^2 - 2t^2 + t^2 \\
&= 2\sum_{j=1}^{k_n}(t_{n_j} - t_{n_{j-1}})^2 \\
&\le 2\delta(\mathcal{Z}_n) \cdot t \to 0 \text{ für } n \to \infty
\end{aligned}$$

■

Bemerkung: Beachte, dass die Zerlegungsfolge $\mathcal{Z}_n$ nicht von ω abhängt. Gilt zusätzlich $\mathcal{Z}_n \subset \mathcal{Z}_{n+1}$, so gilt sogar $V_{\mathcal{Z}_n} \xrightarrow{f.s.} t$. Dieser Grenzwert heißt quadratische Variation des Prozesses.

Bemerkung: Eine stetige Funktion f auf einem kompakten Intervall $[a, b]$ mit $V_a^b f < \infty$, hat stets quadratische Variation 0, da

$$\sum_{k=1}^{n}(f(t_k) - f(t_{k-1}))^2 \le \max_{1\le k\le n} |f(t_k) - f(t_{k-1})| \cdot \sum_{i=1}^{n} |f(t_k) - f(t_{k-1})|$$

Also $V_a^b B(\cdot, \omega) = \infty \quad \forall 0 \le a < b < \infty$.

12. Die Brownsche Bewegung als Markov-Prozess

Definition
Ein *stochastischer Kern* (Übergangswahrscheinlichkeit) $P(x, A)$ von $(\mathbb{R}, \mathfrak{B})$ nach $(\mathbb{R}, \mathfrak{B})$ ist eine Funktion, so dass

a) für alle $x \in \mathbb{R}$ ist $A \mapsto P(x, A)$ ein Wahrscheinlichkeitsmaß und

b) für alle $A \in \mathfrak{B}$ ist $x \mapsto P(x, A)$ messbar.

Definition
Eine Familie $(P_t)_{t\geq 0}$ von stochastischen Kernen von $(\mathbb{R}, \mathfrak{B})$ nach $(\mathbb{R}, \mathfrak{B})$ heißt *Markov-Halbgruppe* oder *Übergangs-Halbgruppe*, wenn für alle $s, t \geq 0$ und für alle $A \in \mathfrak{B}$ gilt:

$$P_{s+t}(x, A) = \int P_t(y, A) P_s(x, dy)$$

Bemerkung: Wir schreiben kurz: $P_{s+t} = P_s \otimes P_t$.

Beispiel 12.1
Für $t \geq 0$ sei $P_t(x, \cdot) = \mathcal{N}(x, t)$ die Normalverteilung mit Erwartungswert x und Varianz t, wobei $\mathcal{N}(x, 0) := \delta_x$.

$$P_t(x, A) = \int_A \frac{1}{\sqrt{2\pi t}} e^{-\frac{1}{2t}(y-x)^2} dy$$

ist messbar (Fubini) und es gilt

$$\begin{aligned}
\int P_t(y, A) P_s(x, dy) &= \int_A \Big(\int \underbrace{\frac{1}{\sqrt{2\pi t}} e^{-\frac{1}{2t}(z-y)^2}}_{=: f_2(z-y)} \underbrace{\frac{1}{\sqrt{2\pi s}} e^{-\frac{1}{2s}(y-x)^2}}_{=: f_1(y)} dy \Big) dz \\
&= \int_A \Big(\int f_2(z-y) \cdot f_1(y) dy \Big) dz \\
&= P_{s+t}(x, A).
\end{aligned}$$

Definition
Es seien $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum mit Filtration $(\mathcal{F}_t)_{t\geq 0}$, $(X_t)_{t\geq 0}$ ein dazu adaptierter reellwertiger Prozess, ν ein Wahrscheinlichkeitsmaß auf $(\mathbb{R}, \mathfrak{B})$ und $(P_t)_{t\geq 0}$ eine Markov-Halbgruppe. Gilt dann $X_0 \sim \nu$ und für alle $0 \leq s \leq t$, $A \in \mathfrak{B}$

$$P(X_t \in A \mid \mathcal{F}_s) = P_{t-s}(X_s, A) \qquad (*)$$

so heißt $(X_t, \mathcal{F}_t)_{t\geq 0}$ ein (homogener) *Markov-Prozess* mit Übergangshalbgruppe $(P_t)_{t\geq 0}$ und Startverteilung ν.

Bemerkung: a) In der Regel ist $(\mathcal{F}_t)_{t\geq 0}$ die natürliche Filtration zu $(X_t)_{t\geq 0}$.

b) Zum Nachweis von $(*)$ genügt zu zeigen, dass für alle $F \in \mathcal{F}_s$, $A \in \mathfrak{B}$

$$\underbrace{\int_F 1_{[X_t \in A]} dP}_{P(F \cap \{X_t \in A\})} = \int_F (P_{t-s})(X_s, A) dP$$

da $P_{t-s}(X_s, A)$ als messbare Funktion von X_s $\mathcal{F}_s$-messbar ist.

Besitzt $\mathcal{F}_s$ einen durchschnittsstabilen Erzeuger $\mathcal{E}$, so genügt es, die Gleichung für alle $F \in \mathcal{E}$ nachzuprüfen.

Satz 12.1
Es sei $(B_t)_{t\geq 0}$ die Brownsche Bewegung mit zugehöriger natürlicher Filtration $(\mathcal{F}_t)_{t\geq 0}$. Dann ist $(B_t, \mathcal{F}_t)_{t\geq 0}$ ein Markov-Prozess mit Startverteilung δ_0 und stochastischen Kernen $P_t(x, \cdot) = \mathcal{N}(x, t)$.

Beweis
Wir zeigen zunächst eine Hilfsaussage: Seien $0 = t_0 \leq t_1 \leq \cdots \leq t_n$ und $g : \mathbb{R}^n \to \mathbb{R}$ eine beschränkte $(\mathfrak{B}^n, \mathfrak{B})$-messbare Funktion. Setze $Y_i = B_{t_i} - B_{t_{i-1}}$ für $i = 1, \ldots, n$. $Y_1, \ldots, Y_n$ sind unabhängig und $Y_i \sim \mathcal{N}(0, t_i - t_{i-1})$. Also gilt:

$$\begin{aligned} Eg(B_{t_1}, \ldots, B_{t_n}) &= Eg(Y_1, Y_1 + Y_2, \ldots, Y_1 + \cdots + Y_n) \qquad (\triangle) \\ &= \int \cdots \int g(y_1, y_1 + y_2, \ldots, y_1 + \cdots + y_n) \\ &\quad \mathcal{N}(0, t_n - t_{n-1})(dy_n)\mathcal{N}(0, t_{n-1} - t_{n-2})(dy_{n-1}) \cdots \mathcal{N}(0, t_1)(dy_1) \\ &= \int \cdots \int g(x_1, \ldots, x_n)\mathcal{N}(x_{n-1}, t_n - t_{n-1})(dx_n) \cdots \mathcal{N}(0, t_1)(dx_1) \end{aligned}$$

Es sei $A \in \mathfrak{B}$, $0 \leq s \leq t$. Zu zeigen ist, dass $P(B_t \in A \mid \mathcal{F}_s) = \mathcal{N}(B_s, t-s)(A)$, das heißt

$$\int_F 1_{[B_t \in A]} dP = \int_F \mathcal{N}(B_s, t-s)(A) dP \qquad (\square)$$

für alle $F \in \mathcal{F}_s$, beziehungsweise auf einem durchschnittsstabilen Erzeugendensystem von $\mathcal{F}_s$. Ein solcher ist

$$\bigcap_{k=1}^n \{B_{t_k} \in A_k\} \quad 0 \leq t_1 \leq \cdots \leq t_n = s,\ A_1 \ldots, A_n \in \mathfrak{B}.$$

Setze $g(x_1, \ldots, x_n, x) = 1_A(x) \cdot \prod_{k=1}^n (1_{A_k}(x_k))$ in $(\triangle)$:

$$\begin{aligned} Eg(B_{t_1}, \ldots, B_{t_n}) &= Eg(Y_1, Y_1 + Y_2, \ldots, Y_1 + Y_n) \\ &= \int_{A_1} \cdots \int_{A_n} \int_A \mathcal{N}(x_n, t-s)(dx)\mathcal{N}(x_{n-1}, t_n - t_{n-1})(dx_n) \cdots N(0, t_1)(dx_1) \\ &= \text{linke Seite von } (\square) \text{ mit obigen } F \end{aligned}$$

Setze jetzt $g(x_1, \ldots, x_n) = \mathcal{N}(x_n, t-s)(A) \prod_{k=1}^n 1_{A_k}(x_k)$ in $(\triangle)$:

$$\begin{aligned} Eg(B_{t_1}, \ldots, B_{t_n}) &= \int \cdots \int_{\{B(t_1) \in A_1, \ldots, B(t_n) \in A_n\}} \mathcal{N}(x_n, t-s)(A) dP \\ &= \int_{A_1} \cdots \int_{A_n} \mathcal{N}(x_n, t-s)(A) N(x_{n-1}, t_n - t_{n-1})(dx_n) \cdots \mathcal{N}(0, t_1)(dx_1) \\ &= \text{rechte Seite von } (\square) \text{ mit obigen } F \end{aligned}$$

Da $\int_A \mathcal{N}(x_n, t-s)(dx) = \mathcal{N}(x_n, t-s)(A)$ ist $(\square)$ gezeigt. ∎

Bemerkung: Die Markov-Eigenschaft lässt sich schreiben als

$$P(X_{s+t} \in A \mid \mathcal{F}_s) = P_t(X_s, A) \qquad (*)$$

für alle $A \in \mathfrak{B}$, $s, t \geq 0$.

Definiere für $f : \mathbb{R} \to \mathbb{R}$ (oder $\to \mathbb{C}$), falls existent, für alle $x \in \mathbb{R}$:

$$\big(P_t f\big)(x) := \int f(y) P_t(x, dy)$$

Also kann man schreiben:

$$P_t(X_s, A) = \int 1_A(y) P_t(X_s, dy) = \big(P_t 1_A\big)(X_s)$$

für $f = 1_A$ kann man $(*)$ schreiben als

$$E[f(X_{s+t}) \mid \mathcal{F}_s] = \big(P_t f\big)(X_s) \tag{$\triangle$}$$

Mittels Algebraischer Induktion und der monotonen Konvergenz in bedingter Version kann man daraus folgern, dass die Gleichung $(\triangle)$ für alle messbaren, beschränkten $f : \mathbb{R} \to \mathbb{R}$ gilt.

Umgekehrt stellt sich die Frage: Für welche Funktionenklasse muss die Gleichung $(\triangle)$ gelten, um $(*)$ zu folgern?

Für $s = 0$ und $\mathcal{F}_s = \{\emptyset, \Omega\}$ folgt $(*)$, wenn $(\triangle)$ für alle $f : \mathbb{R} \to \mathbb{C}$ der Form $f(x) = e^{i\theta x}$, $\theta \in \mathbb{R}$ gilt. (Eindeutigkeitssatz für charakteristische Funktionen)

Dies gilt auch für bedingte charakteristische Funktionen, das heißt $(*)$ ist erfüllt, wenn für alle $\theta \in \mathbb{R}$ gilt:

$$E[e^{i\theta X_{s+t}} \mid \mathcal{F}_s] = \int e^{i\theta y} P_t(X_s, dy).$$

Damit kann man Satz 12.1 direkt zeigen:

$$\begin{aligned}
E[e^{i\theta B_{s+t}} \mid \mathcal{F}_s] &= E[e^{i\theta(B_s + B_{s+t} - B_s)} \mid \mathcal{F}] \\
&= e^{i\theta B_s} \cdot E[e^{i\theta B_{s+t} - B_s}] \\
&= e^{i\theta B_s} \cdot e^{-\theta^2 \frac{t}{2}} \\
&= \int e^{i\theta y} P_t(B_s, dy)
\end{aligned}$$

In Kapitel II haben wir gesehen, dass die Übergangsmatrizenfunktion $(P(t))_{t\geq 0}$ durch die Intensitätsmatrix Q in bestimmten Fällen charakterisiert werden kann.
Insbesondere ist:

$$q_{ij} := \lim_{t \downarrow 0} \frac{p_{ij}(t) - \delta_{ij}}{t}$$

Unter bestimmten Voraussetzungen kann auch $(P_t)_{t\geq 0}$ bei allgemeinen Markov-Prozessen aus dem sogenannten Generator

$$\mathcal{G} := \lim_{t \downarrow 0} \frac{1}{t}(P_t - P_0)$$

bestimmt werden. Wir betrachten hier die Brownsche Bewegung.

$$C_b^2(\mathbb{R}) := \{f : \mathbb{R} \to \mathbb{R} \mid f \text{ ist zweimal stetig differenzierbar, } f \text{ beschränkt}\}$$

Satz 12.2
Es sei $(P_t)_{t\geq 0}$ die Übergangshalbgruppe zur Brownschen Bewegung. Dann gilt:

$$\mathcal{G}f := \lim_{t \downarrow 0} \frac{1}{t}(P_t f - f) = \frac{1}{2} f'' \text{ für alle } f \in C_b^2(\mathbb{R})$$

Beweis
Sei $f \in C_b^2(\mathbb{R})$.

$$\begin{aligned}\frac{1}{t}(P_t f - f)(x) &= \int \frac{1}{t}(f(y) - f(x))\mathcal{N}(x,t)(dy) \\ &= \frac{1}{\sqrt{2\pi}} \int t^{-\frac{3}{2}}(f(y) - f(x))e^{-\frac{(y-x)^2}{2t}}dy = (*)\end{aligned}$$

Substitution mit $z = \frac{y-x}{\sqrt{t}}$ liefert:

$$(*) = \frac{1}{\sqrt{2\pi}} \int \frac{1}{t}(f(x + z\sqrt{t}) - f(x))e^{-\frac{1}{2}z^2}dz$$

Taylor-Entwicklung:

$$f(x + z\sqrt{t}) = f(x) + z\sqrt{t}f'(x) + \frac{1}{2}z^2 t f''(\xi) \text{ mit } \xi \in (x, x + z\sqrt{t})$$

Daraus folgt:

$$\frac{1}{t}(P_t f - f)(x) = \frac{1}{\sqrt{2\pi}} \frac{f'(x)}{\sqrt{t}} \int z e^{-\frac{1}{2}z^2}dz + \frac{1}{2}\int \frac{1}{\sqrt{2\pi}} z^2 f''(\xi) e^{-\frac{1}{2}z^2}dz$$

Es gilt: $f''(\xi) \to f''(x)$ für $t \to 0$. Da f'' beschränkt ist, folgt mit majorisierter Konvergenz:

$$\lim_{t \downarrow 0} \frac{1}{t}(P_t f - f)(x) = \frac{1}{2}f''(x) \underbrace{\int \frac{1}{\sqrt{2\pi}} z^2 e^{-\frac{1}{2}z^2}dz}_{=1}$$

■

13. Die starke Markov-Eigenschaft mit Anwendungen

Wir benötigen zunächst folgende Hilfsüberlegungen:
Ist $\tau : \Omega \longrightarrow [0, \infty]$ eine Stoppzeit, so ist

$$\mathcal{F}_\tau := \{A \in \mathcal{F} \mid A \cap \{\tau \leq t\} \in \mathcal{F}_t \text{ für alle } t \geq 0\}$$

die σ-Algebra der τ-Vergangenheit (vgl. Stochastik II).

Lemma 13.1

a) Ist τ eine Stoppzeit, so wird durch

$$\tau_n(\omega) := \begin{cases} \frac{\lceil 2^n \tau(\omega) \rceil}{2^n} & \text{, falls } \tau(\omega) < \infty \\ \infty & \text{, falls } \tau(\omega) = \infty \end{cases}$$

eine Folge $(\tau_n)_{n \in \mathbb{N}}$ von Stoppzeiten definiert, die für $n \to \infty$ fast sicher gegen τ konvergiert.
Ist τ beschränkt, so nimmt jedes τ_n nur endlich viele Werte an.

b) Ist $X = (X_t)_{t\geq 0}$ ein $(\mathcal{F}_t)_{t\geq 0}$ adaptierter Prozess mit rechtsseitig stetigen Pfaden, so ist für alle $t \geq 0$ die Abbildung

$$[0,t] \times \Omega \ni (s,\omega) \mapsto X_s(\omega) \in \mathbb{R}$$

$\mathfrak{B}[0,t] \otimes \mathcal{F}_t$-messbar.

c) Ist τ eine endliche Stoppzeit und $(X_t)_{t\geq 0}$ ein adaptierter Prozess mit rechtsseitig stetigen Pfaden, so ist

$$X_\tau : \Omega \longrightarrow \mathbb{R}, \omega \mapsto X_{\tau(\omega)}(\omega)$$

$\mathcal{F}_\tau$-messbar.

Bemerkung: Einen stochastischen Prozess X mit der Messbarkeitseigenschaft aus Teil b.) nennt man *progressiv messbar.*

Beweis

a) Es ist zu zeigen, dass τ_n eine Stoppzeit ist:

$$\{\tau_n \leq t\} = \left\{ \tau \leq \underbrace{\frac{\lfloor 2^n t \rfloor}{2^n}}_{\leq t} \right\} \in \mathcal{F}_t \text{ für alle } t \geq 0.$$

Die anderen Aussagen sind “klar”.

b) Wir zeigen zunächst:

$$(s,\omega) \mapsto 1_{[u,v)}(s) \cdot X_v(\omega)$$

ist $\mathfrak{B}[0,t] \otimes \mathcal{F}_t$-messbar für $0 \leq u < v \leq t$. Denn: Sei $A \in \mathfrak{B}$ mit $0 \notin A$, dann ist

$$(1_{[u,v)} X_v)^{-1}(A) = \underbrace{[u,v)}_{\in \mathfrak{B}[0,t]} \times \underbrace{X_v^{-1}(A)}_{\mathcal{F}_v \subset \mathcal{F}_t}$$

Analog: $(s,\omega) \mapsto 1_t(s) X_t(\omega)$ ist messbar. Sei jetzt:

$$X^{(n)}(s,\omega) := \sum_{k=0}^{n-1} 1_{[\frac{kt}{n}, \frac{(k+1)t}{n})}(s) \cdot X(\frac{(k+1)t}{n}, \omega) + 1_t(s) \cdot X_t(\omega)$$

$X^{(n)}$ ist eine Linearkombination von Abbildungen vom obigen Typ, also $\mathfrak{B}[0,t] \otimes \mathcal{F}_t$-messbar.

c) Es ist zu zeigen: Für alle $A \in \mathfrak{B}$ und $t \geq 0$ ist

$$X_\tau^{-1}(A) \cap \{\tau \leq t\} \in \mathcal{F}_t$$

Wir betrachten die Abbildungen:

$$g : \Omega \longrightarrow [0,t] \times \Omega, g(\omega) = (\tau(\omega) \wedge t, \omega)$$
$$h : [0,t] \times \Omega \longrightarrow \mathbb{R}, h(s,\omega) = X_s(\omega)$$

mit der σ-Algebra $\mathfrak{B}[0,t] \otimes \mathcal{F}_t$ auf $[0,t] \times \Omega$. g ist $(\mathcal{F}_t, \mathfrak{B}[0,t] \otimes \mathcal{F}_t)$-messbar, da für alle $A \in \mathcal{F}_t, s \leq t$ gilt:

$$\{\omega \in \Omega \mid (\tau(\omega) \wedge t, \omega) \in [0,s] \times A\} = A \cap \underbrace{\{\tau \leq s\}}_{\in \mathcal{F}_s} \in \mathcal{F}_t$$

Die Messbarkeit von h folgt aus b.). Es gilt also:

$$\begin{aligned} & h(g(\omega)) = h(\tau(\omega) \wedge t, \omega) = X_{\tau(\omega)\wedge t}(\omega) \text{ ist } (\mathcal{F}_t, \mathfrak{B})\text{-messbar} \\ \Longrightarrow\ & X_{\tau\wedge t}^{-1}(A) \in \mathcal{F}_t \text{ für alle } A \in \mathfrak{B} \\ \Longrightarrow\ & X_\tau^{-1}(A) \cap \{\tau \leq t\} = \underbrace{X_{\tau\wedge t}^{-1}(A)}_{\in \mathcal{F}_t} \cap \underbrace{\{\tau \leq t\}}_{\in \mathcal{F}_t} \in \mathcal{F}_t \end{aligned}$$

■

Satz 13.2
Es sei $(B_t, \mathcal{F}_t)_{t\geq 0}$ eine Brownsche Bewegung, τ eine endliche Stoppzeit. Dann sind die σ-Algebra $\mathcal{F}_\tau$ und der Prozess $(B_{\tau+t} - B_\tau)_{t\geq 0}$ stochastisch unabhängig und $(B_{\tau+t} - B_\tau)_{t\geq 0}$ hat dieselbe Verteilung wie $(B_t)_{t\geq 0}$.

Beweis
Wir zeigen zunächst: Für alle $t \geq 0$ ist $B_{\tau+t} - B_\tau$ unabhängig von $\mathcal{F}_\tau$ und hat Verteilung $\mathcal{N}(0,t)$. Angenommen, τ nimmt nur abzählbar viele Werte an. Dann gilt für alle $A \in \mathcal{F}_\tau$, $C \in \mathfrak{B}$, dass

$$\begin{aligned} P(A \cap \{B_{\tau+t} - B_\tau \in C\}) &= \sum_{s\geq 0} P(\underbrace{A \cap \{\tau = s\}}_{\in \mathcal{F}_s} \cap \{B_{\tau+t} - B_\tau \in C\}) \\ &= \sum_{s\geq 0} P(A \cap \{\tau = s\}) \cdot \underbrace{P(B_{s+t} - B_s \in C)}_{\mathcal{N}(0,t)(C)} \\ &= P(A) \cdot \mathcal{N}(0,t)(C). \end{aligned}$$

Mit $A = \Omega$ ergibt das:

$$P(B_{\tau+t} - B_\tau \in C) = \mathcal{N}(0,t)(C)$$

und für alle $A \in \mathcal{F}_\tau$ und $C \in \mathfrak{B}$ gilt

$$P(A \cap \{B_{\tau+t} - B_\tau \in C\}) = P(A) \cdot P(B_{\tau+t} - B_\tau \in C).$$

Sei jetzt τ eine beliebige (endliche) Stoppzeit und

$$\tau_n := \frac{\lceil 2^n \tau \rceil}{2^n}.$$

Wegen $\mathcal{F}_\tau \subset \mathcal{F}_{\tau_n}$ folgt für alle $A \in \mathcal{F}_\tau$:

$$P(A \cap \{B_{\tau_n+t} - B_{\tau_n} \in C\}) = P(A) \cdot \mathcal{N}(0,t)(C)$$

Für die Unabhängigkeit genügt es, $C = (-\infty, c)$, $c \in \mathbb{R}$, zu betrachten, da diese ein durchschnittsstabiles Erzeugendensystem von $\mathfrak{B}$ bilden.

Es bleibt zu zeigen:

$$\lim_{n\to\infty}\int_A 1_{(-\infty,c)}(B_{\tau_n+t}-B_{\tau_n})dP=\int_A 1_{(-\infty,c)}(B_{\tau+t}-B_\tau)dP \tag{$*$}$$

Stetigkeit der Pfade impliziert $B_{\tau_n}\to B_\tau$, $B_{\tau_n+t}\to B_{\tau+t}$ mit $n\to\infty$ und mit majorisierter Konvergenz ($1_{(-\infty,c)}\le 1$) folgt

$$\lim_{n\to\infty}\int_A 1_{(-\infty,c)}(B_{\tau_n+t}-B_{\tau_n})dP=\int_A \lim_{n\to\infty} 1_{(-\infty,c)}(B_{\tau_n+t}-B_{\tau_n})dP$$

Das Problem ist nun, dass die Indikatorfunktion nicht stetig ist. Jedoch ist $B_{\tau_n+t}-B_{\tau_n}\sim\mathcal{N}(0,t)$, also ist $B_{\tau+t}-B_\tau\sim\mathcal{N}(0,t)$ und $P(B_{\tau+t}-B_\tau=c)=0$, und damit gilt ($*$). Damit ist die Anfangsbehauptung gezeigt.

Sei jetzt $A\in\mathcal{F}_\tau$, $0=:t_0<t_1<\cdots<t_n$, $C_1,\dots,C_n\in\mathfrak{B}$. Wende die bisher bewiesene Aussage auf die Stoppzeiten $\tau+t_{n-1},\tau+t_{n-2},\dots,\tau+t_0$ an, so folgt:

$$P(A\cap\bigcap_{k=1}^n\{B(\tau+t_k)-B(\tau+t_{k-1})\in C_k\})=P(A)\cdot\prod_{k=1}^n P(B(\tau+t_k)-B(\tau+t_{k-1})\in C_k)$$

Da die Mengen

$$\bigcap_{k=1}^n\{B(\tau+t_k)-B(\tau+t_{k-1})\in C_k\}$$

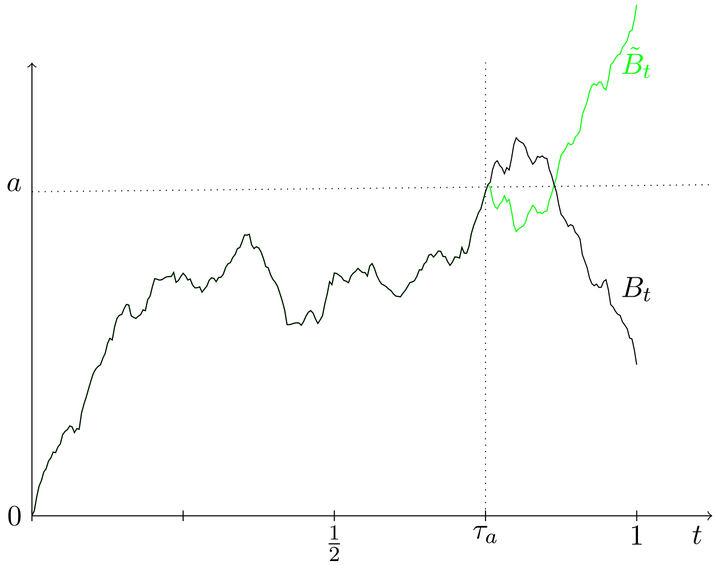
ein durchschnittsstabiles Erzeugendensystem der von $(B_{\tau+t}-B_\tau)_{t\ge0}$ erzeugten σ-Algebra bilden, folgt die Unabhängigkeit. Analog folgt die Verteilungsaussage. ∎

Bemerkung: $(B_{\tau\wedge t})_{t\ge0}$ heißt Prä-τ-Prozess, $(B_{\tau+t}-B_\tau)_{t\ge0}$ heißt Post-τ-Prozess.

Satz 13.3 (Das Spiegelungsprinzip)
Es sei $(B_t)_{t\ge0}$ die Brownsche Bewegung, τ eine endliche Stoppzeit. Dann ist auch der ab τ gespiegelte Prozess $(\tilde{B}_t)_{t\ge0}$:

$$\tilde{B}_t(\omega):=\begin{cases}2B_{\tau(\omega)}(\omega)-B_t(\omega), & t\ge\tau(\omega)\\ B_t(\omega), & t<\tau(\omega)\end{cases}$$

eine Brownsche Bewegung.



Beweis
Es sei $X_t := B_{\tau\wedge t}$, $Y_t := B_{\tau+t} - B_\tau$, $t \geq 0$. Sei jetzt

$$\begin{aligned}\psi : [0,\infty) \times C[0,\infty) \times C[0,\infty) &\to C[0,\infty)\\ (a,f,g) &\mapsto h\end{aligned}$$

wobei

$$h(t) := \begin{cases} f(t), & \text{falls } t \leq a \\ f(a) + g(t-a), & \text{sonst}\end{cases}$$

Nun ist $B = \psi(\tau, X, Y)$. Nach Satz 13.2 ist (τ, X) unabhängig von Y und Y ist eine Brownsche Bewegung, also ist (τ, X) unabhängig von $-Y$ und $-Y$ ist eine Brownsche Bewegung, das heißt

$$(\tau, X, Y) \stackrel{\mathrm{d}}{=} (\tau, X, -Y)$$

also ist $B = \psi(\tau, X, Y) \stackrel{\mathrm{d}}{=} \psi(\tau, X, -Y) = \tilde{B}$. ■

Eine klassische Anwendung des Spiegelungsprinzips ist die Bestimmung der Verteilung von

$$M_t := \sup_{0\leq s\leq t} B_s.$$

Dazu betrachten wir die Stoppzeit

$$\tau_a := \inf\{t \geq 0 \mid B_t = a\}$$

für ein $a \in \mathbb{R}$ (wobei wir $\inf \emptyset = \infty$ setzen).

Satz 13.4
Es sei $t \geq 0$. Dann haben M_t, $M_t - B_t$ und $|B_t|$ dieselbe Verteilung.

Beweis
Es sei $x > 0$ und $y < x$. Sei $f : [0,t] \to \mathbb{R}$ so, dass $f(0) = 0$, $f(t) \leq y$, $\max_{0\leq s\leq t} f(s) \geq x$. Es sei

$$g(s) = \begin{cases} f(s), & s \leq \tau_x \\ 2x - f(s), & s > \tau_x \end{cases}$$

also ist $g(0) = 0$ und $g(t) \geq 2x - y$. Dann ist

$$\begin{aligned} P(M_t \geq x, B_t \leq y) &= P(\tilde{B}_t \geq 2x - y) \\ &\overset{13.3}{=} P(B_t \geq 2x - y) \\ &= P\left(\frac{B_t}{\sqrt{t}} \geq \frac{2x-y}{\sqrt{t}}\right) \\ &= 1 - \Phi\left(\frac{2x-y}{\sqrt{t}}\right) \end{aligned}$$

für alle $x > 0$, $y < x$. Die gemeinsame Dichte von (M_t, B_t) ist

$$\begin{aligned} &-\frac{\partial^2}{\partial x \partial y} P(M_t \geq x, B_t \leq y) \\ =&\frac{\partial}{\partial}\left(-\phi\left(\frac{2x-y}{\sqrt{t}}\right)\frac{1}{\sqrt{t}}\right) \\ =&-\phi'\left(\frac{2x-y}{\sqrt{t}}\right)\frac{2}{t} \qquad \text{, für } x \geq 0, y \leq x, \text{ sonst } 0 \end{aligned}$$

wobei $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$. Dichtetransformation liefert:

$$\begin{aligned} & f_{M_t, M_t - B_t}(x,y) \\ &= -\frac{2}{t}\phi'\left(\frac{x+y}{\sqrt{t}}\right), \quad x, y \geq 0 \\ \Longrightarrow \quad f_{M_t}(x) &= \int f_{M_t, M_t-B_t}(x,y) dy \\ &= -\int \frac{2}{t}\phi'\left(\frac{x+y}{\sqrt{t}}\right) dy \\ &= -\frac{2}{\sqrt{t}}\phi\left(\frac{x+y}{\sqrt{t}}\right)\Big|_0^\infty \\ &= \frac{2}{\sqrt{t}}\phi\left(\frac{x}{\sqrt{t}}\right), \quad x \geq 0. \end{aligned}$$

Analog:$f_{M_t - B_t}(y) = \frac{2}{\sqrt{t}}\phi\left(\frac{y}{\sqrt{t}}\right), \quad y \geq 0$ und daraus folgt $M_t \overset{d}{=} M_t - B_t$. Die Dichte von $|B_t|$ kann man direkt bestimmen und man erhält wieder denselben Ausdruck. ∎

Bemerkung: Satz 13.3 kann verwendet werden, um die Verteilung von bestimmten Eintrittszeiten der Brownschen Bewegung zu bestimmen:

$$\begin{aligned} &\{\tau_a \leq t\} = \{M_t \geq a\}, \quad a > 0 \\ \Longrightarrow\ & P(\tau_a \leq t) = P(M_t \geq a) = P(|B_t| \geq a) = P(|B_1| \geq \frac{a}{\sqrt{t}}), \qquad \text{für alle } t > 0. \end{aligned}$$

Also gilt für die Dichte f_{τ_a} von τ_a:

$$\begin{aligned} f_{\tau_a}(t) =& \frac{d}{dt}\left(\int_{\frac{a}{\sqrt{t}}}^{\infty} 2\phi(y)dy\right) \\ =& \frac{a}{\sqrt{2\pi t^3}} e^{-\frac{a^2}{2t}}, \quad t > 0. \end{aligned}$$

Es gilt: $E\tau_a = \infty$.
Wir betrachten jetzt $Z(\omega) := \{t \geq 0 \mid B_t(\omega) = 0\}$ die Nullstellen der Pfade $t \mapsto B_t(\omega)$ einer Brownschen Bewegung.

Definition
Eine Menge $A \subset \mathbb{R}$ heißt *perfekt*, wenn sie mit der Menge ihrer Häufungspunkte übereinstimmt.

Satz 13.5
Für P-fast alle $\omega \in \Omega$ ist $Z(\omega)$ perfekt.

Beweis
Nach Voraussetzung existiert ein $C \in \mathcal{F}$ mit $P(C) = 1$ und $t \mapsto B_t(\omega)$ ist stetig für alle $\omega \in C$. Da $Z(\omega) = B^{-1}(\omega)(\{0\})$ ist $Z(\omega)$ abgeschlossen für alle $\omega \in C$, da $\{0\}$ abgeschlossen und $t \mapsto B_t(\omega)$ stetig ist.
Es bleibt noch zu zeigen, dass für P-fast alle $\omega \in \Omega$ jedes $t \in Z(\omega)$ Häufungspunkt von $Z(\omega)$ ist. Wir wissen bereits (Vergleich: Bemerkung nach Satz 11.1), dass ein $A_0 \in \mathcal{F}$ existiert, mit $P(A_0) = 1$ und für alle $\omega \in A_0$ ist $t_0 = 0$ Häufungspunkt von $Z(\omega)$. Für $q > 0, q \in \mathbb{Q}$ seien

$$\begin{aligned} \tau_q(\omega) :=& \inf\{t \geq q \mid B_t(\omega) = 0\} \\ A_q :=& \{\omega \in \Omega \mid \tau_q(\omega) \text{ ist Häufungspunkt von } Z(\omega) \cap (\tau_q(\omega), \infty)\} \end{aligned}$$

Da τ_q eine Stoppzeit ist, ist $\tilde{B}_t := B_{\tau_q+t} - \underbrace{B_{\tau_q}}_{=0, P-f.s.} = B_{\tau_q+t}$ nach Satz 13.2 wieder eine Brownsche Bewegung. Daraus folgt: $P(A_q) = 1$.
Sei

$$A := C \cap \bigcap_{\substack{q \in \mathbb{Q}, \\ q \geq 0}} A_q$$

Da $P(A) = 1$ genügt es zu zeigen, dass $Z(\omega)$ für alle $\omega \in A$ perfekt ist.
Annahme: Es existiert ein $t_0 \in Z(\omega)$, das für ein $\omega \in A$ kein Häufungspunkt von $Z(\omega)$ ist. Dann ist $t_0 > 0$ und es existiert $\varepsilon > 0$ mit $(t_0 - \varepsilon, t_0) \cap Z(\omega) = \emptyset$. Sei $q \in (t_0 - \varepsilon, t_0) \cap \mathbb{Q}$. Da $t_0 \in Z(\omega)$ ist $\tau_q(\omega) = t_0$. Daraus folgt, dass $\tau_q(\omega) = t_0$ kein Häufungspunkt von $Z(\omega) \cap (t_0, \infty)$ ist. Widerspruch!
Also gilt die Behauptung. ∎

Bemerkung: Man kann zeigen, dass $Z(\omega)$ für P-fast alle $\omega \in \Omega$ eine Lebesgue-Nullmenge ist (also insbesondere kein Intervall $[a, b]$ enthält). Weiter gilt, dass nichtleere perfekte Mengen überabzählbar sind.

14. Das Invarianzprinzip von Donsker

Eine Brownsche Bewegung kann durch eine geeignet skalierte Irrfahrt approximiert werden:
Sei $\xi_1, \xi_2, \ldots$ eine Folge von unabhängig identisch verteilten Zufallsvariablen, mit $E\xi_i = 0, \mathrm{Var}(\xi_i) = \sigma^2, 0 < \sigma^2 < \infty$ (z.B. $P(\xi_i = 1) = P(\xi_i = -1) = \frac{1}{2}$). Weiter sei

$$S_k := \sum_{j=1}^{k} \xi_j, k \in \mathbb{N},\ S_0 = 0$$

und

$$Y_t := S_{\lfloor t \rfloor} + (t - \lfloor t \rfloor)\xi_{\lfloor t \rfloor + 1} \text{ für } t \geq 0.$$

Hier fehlt ein Bild, auf dem zu sehen ist, dass Y_t das lineare Interpolationspolynom der Punkte S_k ist.

Wir skalieren jetzt (Y_t) in der Zeit und im Zustandsraum

$$X_t^{(n)} := \frac{1}{\sigma\sqrt{n}} Y_{n \cdot t}, \quad t \geq 0$$

Bemerkung: Offenbar ist

$$D_k^{(n)} = X_{\frac{(k+1)}{n}}^{(n)} - X_{\frac{k}{n}}^{(n)} = \frac{1}{\sigma\sqrt{n}} \xi_{k+1}$$

unabhängig von $\sigma(\xi_1, \ldots, \xi_k)$ und es gilt $ED_k^{(n)} = 0$ und $\mathrm{Var}(D_k^{(n)}) = \frac{1}{n}$, also ist $(X^{(n)})_{t \geq 0}$ der Brownschen Bewegung nicht unähnlich.

Lemma 14.1
Seien (X_n), X und (Y_n) Zufallsvariablen auf einem gemeinsamen Wahrscheinlichkeitsraum mit Werten in einem metrischen Raum (S, ρ). Falls $X_n \xrightarrow{\mathrm{d}} X$ und $\rho(X_n, Y_n) \xrightarrow{\mathrm{P}} 0$, dann gilt $Y_n \xrightarrow{\mathrm{d}} X$.

Satz 14.2
Mit $(X_t^{(n)})_{t \geq 0}$ definiert wie oben und Zeitpunkten $0 \leq t_1 < \cdots < t_l < \infty$ gilt:

$$(X_{t_1}^{(n)}, \ldots, X_{t_l}^{(n)}) \xrightarrow{\mathrm{d}} (B_{t_1}, \ldots, B_{t_l}),\ n \to \infty,$$

wobei $(B_t)_{t \geq 0}$ die Brownsche Bewegung ist.

Beweis
Wir betrachten hier den Fall $l = 2$ (Rest analog) und zeigen $(X_s^{(n)}, X_t^{(n)}) \xrightarrow{\mathrm{d}} (B_s, B_t)$ für $0 \leq s < t$.

Es gilt:

$$\begin{aligned} P(|X_t^{(n)} - \frac{1}{\sigma\sqrt{n}} S_{\lfloor tn \rfloor}| > \varepsilon) &\leq P(|\frac{1}{\sigma\sqrt{n}} \cdot 1 \cdot \xi_{\lfloor tn \rfloor + 1}| > \varepsilon) \\ &\leq \frac{1}{\varepsilon^2 n} \to 0 \text{ für } n \to \infty \end{aligned}$$

Also folgt

$$\begin{aligned} &P(\|(X_s^{(n)}, X_t^{(n)}) - \frac{1}{\sigma\sqrt{n}}(S_{\lfloor sn \rfloor}, S_{\lfloor tn \rfloor})\|^2 > \varepsilon) \\ \leq\ &P(|X_t^{(n)} - \frac{1}{\sigma\sqrt{n}} S_{\lfloor tn \rfloor}|^2 > \frac{\varepsilon}{2}) + P(|X_s^{(n)} - \frac{1}{\sigma\sqrt{n}} S_{\lfloor sn \rfloor}|^2 > \frac{\varepsilon}{2}) \end{aligned}$$

und damit genügt es nach 14.1 zu zeigen, dass

$$\frac{1}{\sigma\sqrt{n}}(S_{\lfloor sn \rfloor}, S_{\lfloor tn \rfloor}) \xrightarrow{\text{d}} (B_s, B_t)$$

Nach dem „Continuous Mapping Theorem“ genügt es zu zeigen:

$$\frac{1}{\sigma\sqrt{n}}(\sum_{j=1}^{\lfloor sn \rfloor} \xi_j, \sum_{\lfloor sn \rfloor + 1}^{\lfloor tn \rfloor} \xi_j) \xrightarrow{\text{d}} (B_s, B_t - B_s)$$

Wir betrachten die charakteristische Funktion:

$$\begin{aligned} &\lim_{n\to\infty} E(\exp(\frac{i\theta}{\sigma\sqrt{n}} \sum_{j=1}^{\lfloor sn \rfloor} \xi_j + \frac{i\gamma}{\sigma\sqrt{n}} \sum_{\lfloor sn \rfloor + 1}^{\lfloor tn \rfloor} \xi_j)) \\ &\overset{\xi_i \text{ unabh.}}{=} \lim_{n\to\infty} E(\exp(\frac{i\theta}{\sigma\sqrt{n}} \sum_{j=1}^{\lfloor sn \rfloor} \xi_j)) \cdot E(\exp(\frac{i\gamma}{\sigma\sqrt{n}} \sum_{\lfloor sn \rfloor + 1}^{\lfloor tn \rfloor} \xi_j)) \end{aligned}$$

Nach dem Zentralem Grenzwertsatz können wir verwenden, dass

$$\frac{\sqrt{s}}{\sigma\sqrt{\lfloor sn \rfloor}} \sum_{j=1}^{\lfloor sn \rfloor} \xi_j \xrightarrow{\text{d}} Z\sqrt{s}, \quad Z \sim \mathcal{N}(0,1).$$

Da

$$|(\frac{1}{\sigma\sqrt{n}} - \frac{\sqrt{s}}{\sigma\sqrt{\lfloor sn \rfloor}}) \sum_{j=1}^{\lfloor sn \rfloor} \xi_j| \xrightarrow{\text{P}} 0 \text{ für } n \to \infty,$$

folgt mit 14.1, dass

$$\lim_{n\to\infty} E(\exp(\frac{i\theta}{\sigma\sqrt{n}} \sum_{j=1}^{\lfloor sn \rfloor} \xi_j)) = e^{-\theta^2 \frac{s}{2}}$$

und analog für den zweiten Grenzwert. Mit dem Stetigkeitssatz für charakteristische Funktionen folgt die Behauptung. ■

Also konvergieren alle endlich-dimensionalen Verteilungen gegen die entsprechenden endlich-dimensionalen Verteilungen der Brownschen Bewegung.

Schwache Konvergenz kann man aber noch allgemeiner auffassen:

Definition
(vergleiche Stochastik 2) Sei (S, ρ) ein metrischer Raum mit Borel-σ-Algebra $\mathfrak{B}(S)$. Seien $\{P_n\}, P$ Wahrscheinlichkeitsmaße auf $(S, \mathfrak{B}(S))$. $\{P_n\}$ konvergiert schwach gegen P ($P_n \xrightarrow{w} P$) genau dann, wenn

$$\lim_{n\to\infty} \int_S f(s) dP_n(s) = \int_S f(s) dP(s)$$

für alle beschränkten, stetigen Funktionen f auf S gilt.

Auf dem gegeben Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ fassen wir die Brownsche Bewegung $B = (B_t)_{t\geq 0}$ als Abbildung $B : \Omega \to C[0,\infty) = \{f : [0,\infty) \to \mathbb{R} \mid f \text{ stetig}\}$ auf. Auf $C[0,\infty)$ betrachten wir die Metrik

$$\rho(f,g) := \sum_{n=1}^{\infty} \frac{1}{2^n} \max_{0\leq t\leq n} (|f(t) - g(t)| \wedge 1).$$

Offenbar gilt: $X^{(n)} \xrightarrow{w} X$ impliziert $(X^{(n)}_{t_1}, \ldots, X^{(n)}_{t_l}) \xrightarrow{d} (X_{t_1}, \ldots, X_{t_l})$, da für $f : \mathbb{R}^l \to \mathbb{R}$, f stetig und beschränkt, auch $f \circ \pi_{T_0} : C[0,\infty) \to \mathbb{R}$, $T_0 = \{t_1, \ldots, t_l\}$, wieder stetig und beschränkt ist und also gilt

$$\begin{aligned}
\lim_{n\to\infty} \int f(X^{(n)}_t, \ldots, X^{(n)}_{t_l}) dP_n &= \lim_{n\to\infty} \int (f \circ \pi_{T_0})(X^{(n)}) dP_n \\
&= \int (f \circ \pi_{T_0})(X) dP \\
&= \int f(X_{t_1}, \ldots, X_{t_n}) dP.
\end{aligned}$$

Reicht jetzt die Konvergenz der Fidis aus, damit $X^{(n)} \xrightarrow{w} X = B$?

Definition
(vergleiche Stochastik 2) Sei (S, ρ) ein metrischer Raum mit Borel-σ-Algebra $\mathfrak{B}(S)$, $\{P_n, n \in \mathbb{N}\}$ eine Familie von Wahrscheinlichkeitsmaßen auf $(S, \mathfrak{B}(S))$. Diese Familie heißt straff, falls für alle $\varepsilon > 0$ eine kompakte Menge $K \subseteq S$ existiert, so dass für alle $P \in \{P_n, n \in \mathbb{N}\}$ gilt dass $P(K) > 1 - \varepsilon$.

Beispiel 14.1
Sei $(X^{(n)}_t)_{t\geq 0}$ definiert durch

$$X^{(n)}_t = \begin{cases} nt, & 0 \leq t \leq \frac{1}{2n} \\ 1 - nt, & \frac{1}{2n} \leq t \leq \frac{1}{n} \\ 0, & t > \frac{1}{n} \end{cases}$$

Hier fehlt ein Bild von $X^{(n)}_t$
eine Folge von (deterministischen) Prozessen in $C[0,\infty)$ und $X \equiv 0$. Offenbar gilt für alle

$0 \leq t_1 < \cdots < t_l$, dass $(X_{t_1}^{(n)}, \ldots, X_{t_l}^{(n)}) \xrightarrow{n\to\infty} (X_{t_1}, \ldots, X_{t_l}) = (0, \ldots, 0)$.
Allerdings ist für $H : C[0,\infty) \longrightarrow \mathbb{R}$ gegeben durch

$$H(f) := \left(\max_{0\leq t\leq 1} f(t)\right) \wedge 1 \quad \text{(stetig und beschänkt)}$$

$$\int H(f)dP_n = \frac{1}{2} \text{ konvergiert für } n \to \infty \text{ nicht gegen } \int H(f)dP(f) = 0$$

Also gilt $P_n \overset{w}{\to} P$ nicht.

Satz 14.3
Sei $\{X_{(n)}, n \in \mathbb{N}\}$ eine straffe Familie von stetigen Prozessen, so dass für alle $0 \leq t_1 < \cdots < t_l < \infty$ die Folge $((X_{t_1}^{(n)}, \ldots, X_{t_l}^{(n)}))_n$ in Verteilung konvergiert. Sei P_n das von $X^{(n)}$ induzierte Wahrscheinlichkeitsmaß auf $(C[0,\infty), \mathfrak{B}(C[0,\infty)))$. Dann existiert ein Wahrscheinlichkeitsmaß P mit $P_n \overset{w}{\to} P$ unter dem für die Projektionen $X_t : C[0,\infty) \longrightarrow \mathbb{R}, X_t(\omega) = \omega_t$ gilt: $(X_{t_1}^{(n)}, \ldots, X_{t_l}^{(n)}) \overset{d}{\to} (X_{t_1}, \ldots, X_{t_l})$

Beweis
Offenbar ist jede Teilfolge $\{\hat{X}^{(n)}\}$ von $\{X^{(n)}\}$ straff. Nach dem Satz von Prohorov (Stochastik II, Satz 5.8 oder Karatzas/Shrewe S. 67) hat jede straffe Folge eine schwach konvergente Teilfolge. Seien $\{\hat{X}^{(n)}\}$ und $\{\check{X}^{(n)}\}$ verschiedene Teilfolgen, die Maße auf $C[0,\infty)$ induzieren, die gegen ein Maß P beziehungsweise Q konvergieren (d.h. $\hat{P}^n \overset{w}{\to} P, \check{P}^n \overset{w}{\to} Q)$).
Nach Voraussetzung gilt:

$$\begin{aligned} &P(f \in C[0,\infty)|(f(t_1), \ldots, f(t_l)) \in A) \\ =&Q(f \in C[0,\infty)|(f(t_1), \ldots, f(t_l)) \in A) \end{aligned}$$

für alle $A = (-\infty, x_1] \times \cdots \times (-\infty, x_l], l \geq 1, x_i \in \mathbb{R}, 0 \leq t_1 < \cdots < t_l < \infty$. Wegen dem Maßeindeutigkeitssatz folgt $P = Q$.
Annahme: Es existiert eine Folge $\{P^n\}$ von $\{X^{(n)}\}$ induzierten Maßen, die nicht schwach gegen P konvergieren. Dann existiert eine beschränkte stetige Funktion $h : C[0,\infty) \longrightarrow \mathbb{R}$ mit $\int h(f)P_n(df) \nrightarrow \int h(f)P(df), n \to \infty$.
Sei jetzt $\{\tilde{P}^n\}$ eine Teilfolge, so dass $\lim_{n\to\infty} \int h(f)\tilde{P}^n(df)$ existiert, aber ungleich $\int h(f)P(df)$ ist. Für jede Teilfolge $\{\hat{P}^n\}$ von $\{\tilde{P}^n\}$ gilt $\lim_{n\to\infty} \int h(f)\hat{P}^n(df) \neq \int h(f)P(df)$ und der Limes existiert. Das ist ein Widerspruch zu obiger Überlegung. ∎

Wie kann man jetzt Straffheit für $S = C[0,\infty)$ charakterisieren? Dazu sei für $f \in C[0,\infty)$ $T > 0, \delta > 0$ das "Stetigkeitsmodul" definiert als

$$m^T(f,\delta) := \max_{\substack{|s-t| \leq \delta \\ 0 \leq s,t \leq T}} |f(s) - f(t)|$$

Für $m^T(f,\delta)$ gilt jetzt:

Lemma 14.4
Das Stetigkeitsmodul $m^T(f,\delta)$ besitzt folgende Eigenschaften:

a) $f \mapsto m^T(f,\delta)$ ist stetig bezüglich der Metrik ρ

b) $\delta \mapsto m^T(f,\delta) \uparrow$

c) $\lim_{\delta \to 0} m^T(f,\delta) = 0$ für alle $f \in C[0,\infty)$.

ohne Beweis.
Wir benötigen folgenden Satz:

Satz 14.5
Sei $A \subset C[0,\infty)$. Der Abschluss $\bar{A}$ ist kompakt, genau dann wenn

$$\sup_{f \in A} |f(0)| < \infty$$

und

$$\lim_{\delta \downarrow 0} \sup_{f \in A} m^T(f,\delta) = 0$$

für alle $T > 0$.

ohne Beweis (Karatzas/Shrewe S. 62)
Die Charakterisierung der Straffheit auf $C[0,\infty)$ ist jetzt wie folgt:

Satz 14.6
Eine Folge $\{P_n\}_{n\in\mathbb{N}}$ von Wahrscheinlichkeitsmaßen auf $(C[0,\infty), \mathfrak{B}(C[0,\infty)))$ ist straff, genau dann wenn

a) $\lim_{a \to 0} \sup_{n\in\mathbb{N}} P_n(\{f : |f(0)| > a\}) = 0$

b) $\lim_{\delta \downarrow 0} \sup_{n\in\mathbb{N}} P_n(\{f : m^T(f,\delta) > \varepsilon\}) = 0$ für alle $T > 0, \varepsilon > 0$.

Beweis
Sei zunächst $\{P_n\}_{n\in\mathbb{N}}$ straff. Nach Definition existiert zu jedem $\eta > 0$ ein kompaktes K mit $P_n(K) \geq 1-\eta$ für alle $n \in \mathbb{N}$.

$$\overset{14.3}{\Longrightarrow} \forall f \in K \text{ ist } |f(0)| < a \text{ für großes } a \in \mathbb{R}$$
$$\Longrightarrow \text{ a.) ist erfüllt .}$$

Weiter gilt nach Satz 14.5 für T und $\varepsilon > 0$:
Es existiert $\delta_0 > 0$ mit

$$\forall f \in K, 0 \leq \delta < \delta_0 : m^T(f, \delta) \leq \varepsilon$$

$\implies$ b.) ist erfüllt.
Umgekehrt gilt jetzt a.) und b.). Für $T \in \mathbb{N}$ und $\eta > 0$ wählen wir $a > 0$ so, dass

$$\sup_{n \in \mathbb{N}} P_n(\{f : |f(0)| > a\}) \leq \frac{\eta}{2^{T+1}}$$

Wähle $\delta_k > 0, k = 1, 2, \ldots$ so, dass

$$\sup_{n \in \mathbb{N}} P_n(\{f : m^T(f, \delta_k) > \frac{1}{k}\}) \leq \frac{\eta}{2^{T+k+1}}$$

Definiere $A_T := \{f : |f(0)| \leq a, m^T(f, \delta_k) \leq \frac{1}{k}, k = 1, 2, \ldots\}$,

$$A := \bigcap_{T=1}^{\infty} A_T$$

Die Mengen A_T, A sind abgeschlossen (vergleiche Lemma 14.4 a.)). Es gilt

$$P_n(A_T) = 1 - P_n(A_T^c) \geq 1 - \sum_{k=0}^{\infty} \frac{\eta}{2^{T+k+1}} = 1 - \frac{\eta}{2^T}$$

für alle $n \in \mathbb{N}$. Daraus folgt $P_n(A_T^c) \leq \frac{\eta}{2^T}$ für alle $n \in \mathbb{N}$.

„$\Longleftarrow$“

$$A := \bigcap_{T=1}^{\infty} A_T$$

ist abgeschlossen und für alle $n \in \mathbb{N}$ gilt

$$P_n(A_T^c) \leq \frac{\eta}{2^T}.$$

Daraus folgt, dass für alle $n \in \mathbb{N}$ gilt:

$$\begin{aligned} P_n(A) &= P_n(\bigcap_{T=1}^{\infty} A_T) \\ &= 1 - P_n(\bigcup_{T=1}^{\infty} A_T^c) \\ &\geq 1 - \sum_{T=1}^{\infty} P_n(A_T^c) \\ &\geq 1 - \sum_{T=1}^{\infty} \frac{\eta}{2^T} \\ &= 1 - \eta \end{aligned}$$

Nach Satz 14.5 ist A kompakt, womit die Behauptung gezeigt ist. ∎

Satz 14.7 (Invarianzprinzip von Donsker (1951))
Sei $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum und $\{\xi_k\}_{k\in\mathbb{N}}$ eine Folge von unabhängig und identisch verteilten Zufallsvariablen mit $E\xi_k = 0$ und $0 < \mathrm{Var}(\xi_k) = \sigma^2 < \infty$.

Definiere $X^{(n)} := (X_t^{(n)})_{t\geq 0}$ wie zu Beginn des Abschnitts, also

$$X_t^{(n)} := \frac{1}{\sigma\sqrt{n}} Y_{nt}, \quad t \geq 0$$

wobei

$$Y_t := S_{\lfloor t \rfloor} + (t - \lfloor t \rfloor)\xi_{\lfloor t \rfloor + 1}$$

für $t \geq 0$ und $S_k := \sum_{j=1}^k \xi_j$ und P_n sei das auf $(C[0,\infty), \mathfrak{B}(C[0,\infty))$ induzierte Wahrscheinlichkeitsmaß.

Dann gilt $P_n \xrightarrow{\mathrm{w}} P$, wobei P ein Wahrscheinlichkeitsmaß ist, so dass $W_t := \pi_t : C[0,\infty) \to \mathbb{R}$, $f \mapsto f_t$, die Brownsche Bewegung ist.

Beweis
Nach Satz 14.2, Satz 14.3 reicht es zu zeigen, dass $\{P_n\}_{n\in\mathbb{N}}$ straff ist. Dazu verwenden wir Satz 14.6.

Da $X_0^{(n)} = 0$ fast sicher für alle $n \in \mathbb{N}$ gilt, ist (1) klar. Es bleibt zu zeigen, dass für alle $\varepsilon > 0$ und $T > 0$ gilt:

$$\lim_{\delta\to 0} \sup_{n\in\mathbb{N}} P(\max\{|X_s^{(n)} - X_t^{(n)}| \mid |s-t| < \delta,\, 0 \leq s,t \leq T\} > \varepsilon) = 0$$

beziehungsweise

$$\lim_{\delta\to 0} \limsup_{n\to\infty} P(\max\{|X_s^{(n)} - X_t^{(n)}| \mid |s-t| < \delta,\, 0 \leq s,t \leq T\} > \varepsilon) = 0$$

da für festes $n \in \mathbb{N}$ gilt, dass die Wahrscheinlichkeit für $\delta \to \infty$ gegen null geht.

Es gilt jetzt:

$$\begin{aligned}
&P(\max\{|X_s^{(n)} - X_t^{(n)}| \mid |s-t| < \delta,\, 0 \leq s,t \leq T\} > \varepsilon) \\
&\qquad\qquad = P(\max\{|Y_s - Y_t| \mid |s-t| \leq n\delta,\, 0 \leq s,t \leq nT\} > \varepsilon\sigma\sqrt{n})
\end{aligned}$$

und

$$\begin{aligned}
&\max\{|Y_s - Y_t| \mid |s-t| \leq n\delta,\, 0 \leq s,t \leq nT\} \\
&\leq \max\{|Y_s - Y_t| \mid |s-t| \leq \lfloor n\delta \rfloor + 1,\, 0 \leq s,t \leq \lfloor nT \rfloor + 1\} \\
&\leq \max\{|S_{j+k} - S_k| \mid 1 \leq j \leq \lfloor n\delta \rfloor + 1,\, 0 \leq k \leq \lfloor nT \rfloor + 1\}
\end{aligned}$$

Wir zeigen jetzt

$$\lim_{\delta\to 0} \limsup_{n\to\infty} P(\max\{|S_{j+k} - S_k| \mid 1 \leq j \leq \lfloor n\delta \rfloor + 1,\, 0 \leq k \leq \lfloor nT \rfloor + 1\} > \varepsilon\sigma\sqrt{n}) = 0.$$

Etwa eine Seite an in der Vorlesung ausgelassener Rechnungen zeigt, dass:

$$P(\max\{|S_{j+k} - S_k| \mid 1 \leq j \leq \lfloor n\delta \rfloor + 1,\, 0 \leq k \leq \lfloor nT \rfloor + 1\} > \varepsilon\sigma\sqrt{n}) \leq 2\frac{T}{\delta} \underbrace{P(\max_{1 \leq j \leq \lfloor n\delta \rfloor + 1} |S_j| > \frac{1}{3}\varepsilon\sigma\sqrt{n})}_{=:A}$$

Es bleibt zu zeigen:

$$\lim_{\delta \to 0} \limsup_{n \to \infty} \frac{1}{\delta} P(A) = 0.$$

Nach einer weiteren halben Seite Rechnung weiß man, dass

$$P(\max_{1 \leq j \leq \lfloor n\delta \rfloor + 1} |S_j| > \varepsilon\sigma\sqrt{n}) \leq 2P(|S_{\lfloor n\delta \rfloor + 1}| \geq \sigma\sqrt{n}(\varepsilon - \sqrt{2\delta}))$$

Nach dem zentralen Grenzwertsatz gilt

$$\frac{1}{\sigma\sqrt{\lfloor n\delta \rfloor + 1}} S_{\lfloor n\delta \rfloor + 1} \xrightarrow{\mathrm{d}} Z \sim \mathcal{N}(0,1)$$

also mit Lemma 14.1:

$$\frac{1}{\sigma\sqrt{n\delta}} S_{n\delta} \xrightarrow{\mathrm{d}} Z$$

Sei $\lambda > 0$ und $\{\varphi_k\}_{k \in \mathbb{N}}$, $\varphi_k : \mathbb{R} \to \mathbb{R}$ stetig und beschränkt mit $\varphi_k \to \varphi$ und punktweise fallend und

$$\varphi(x) = 1_{(-\infty, -\lambda] \cup [\lambda, \infty)}(x)$$

Für alle $k \in \mathbb{N}$ gilt nun

$$\limsup_{n \to \infty} P(|S_{\lfloor n\delta \rfloor + 1}| \geq \lambda\sigma\sqrt{n\delta}) \leq \lim_{n \to \infty} E\varphi_k(\frac{1}{\sigma\sqrt{n\delta}} S_{\lfloor n\delta \rfloor + 1}) = E\varphi_k(Z).$$

Mit $k \to \infty$ gilt dann

$$\begin{aligned} \limsup_{n \to \infty} P(|S_{\lfloor n\delta \rfloor + 1}| \geq \lambda\sigma\sqrt{n\delta}) &\leq E\varphi(Z) = P(|Z| \geq \lambda) \\ &\leq \frac{E|Z|^3}{\lambda^3} \end{aligned}$$

Setzen wir $\lambda := \frac{\varepsilon - \sqrt{2\delta}}{\sqrt{\delta}}$, so folgt:

$$\limsup_{n \to \infty} \frac{1}{\delta} P(|S_{\lfloor n\delta \rfloor + 1}| \geq \sigma\sqrt{n}(\varepsilon - \sqrt{2\delta})) \leq \frac{\sqrt{\delta}}{(\varepsilon - \sqrt{2\delta})^3} E|Z|^3$$

Mit $\delta \to 0$ folgt dann die Behauptung. ∎

Satz um Satz (hüpft der Has)