Stochastische Prozesse

Sommersemester 2011

Prof. Dr. Karl-Theodor Sturm

Letzte Änderung: 22. September 2011

Inhaltsverzeichnis

Kapitel 1

Bedingte Erwartungen

1.1. Definition der bedingten Erwartung

Gegeben: Ein Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ und σ-Algebra $\mathcal{A}_0 \subset \mathcal{A}$ sowie eine Zufallsvariable $X \in \mathcal{L}^0(\Omega, \mathcal{A})$.

Ziel: Definition von $\mathbb{E}[X|A_0]$ der besten Schätzung/Approximation von X unter allen $\mathcal{A}_0$-messbaren Zufallsvariablen.

Anwendung: Folge $(Y_n)_{n\in\mathbb{N}}$ von Zufallsvariablen auf $(\Omega, \mathcal{A}, \mathbb{P})$ für $k \in \mathbb{N}$ fix:

$$\mathcal{A}_0 = \sigma(Y_1, \dots, Y_n), \quad A_1 = \sigma(Y_n), \quad X = Y_{k+1},$$

Dann ist $(Y_n)_{n\in\mathbb{N}}$ eine Markov-Kette wenn gilt:

$$\mathbb{E}[X|\mathcal{A}_0] = \mathbb{E}[X|\mathcal{A}_1]$$

Beispiel 1.1.1. *1.*

$$\mathcal{A}_0 = \mathcal{A} \qquad \mathbb{E}[X|\mathcal{A}] = X$$

2.

$$\mathcal{A}_0 = \{\emptyset, \Omega\} \qquad \mathbb{E}[X|\mathcal{A}_0] = \mathbb{E}[X]$$

3.

$$\mathcal{A}_0 = \{\emptyset, B, B^c, \Omega\} \qquad B \in \mathcal{A}$$

$\mathcal{A}_0$-messbare Zufallsvariablen sind von der Form:

$$Y = \alpha 1_B + \beta 1_{B^c}$$

Falls $\omega \in B$:

$$\mathbb{E}[X|\mathcal{A}_0](\omega) = \frac{\int_B X d\mathbb{P}}{\mathbb{P}(B)}$$

Falls $\omega \notin B$:

$$\mathbb{E}[X|\mathcal{A}_0](\omega) = \frac{\int_{B^c} X d\mathbb{P}}{\mathbb{P}(B^c)}$$

$\Rightarrow$

$$X_0(\cdot) = \mathbb{E}[X|A_0](\cdot) = \left(\frac{1}{\mathbb{P}(B)} \int_B X d\mathbb{P}\right) 1_B(\cdot) + \left(\frac{1}{\mathbb{P}(B^c)} \int_{B^c} X d\mathbb{P}\right) 1_{B^c}(\cdot)$$

X_0 ist eine $\mathcal{A}_0$-messbare Zufallsvariable.

4. elementare bedingte Erwartung von X unter $\mathcal{A}_0$:

$$\mathcal{A}_0 = \sigma(B_i : i \in \mathbb{N}) \qquad \textit{mit } \Omega = \dot{\bigcup_{i\in\mathbb{N}}} B_i \quad \textit{und} \quad \mathbb{P}(B_i) > 0$$

$$X_0(\omega) = \sum_{i=1}^{\infty} \mathbb{E}[X|B_i] 1_{B_i}(\omega) \qquad \mathbb{E}[X|B_i] = \frac{1}{\mathbb{P}(B_i)} \int_{B_i} X d\mathbb{P} \quad \in \mathbb{R}$$

$X_0 = \mathbb{E}[X|\mathcal{A}_0]$ ist eine Zufallsvariable, falls $X \in \mathcal{L}^0(\Omega, \mathcal{A})$ entweder nichtnegativ oder integrierbar ist.

Definition 1.1.2. *$X_0 \in \mathcal{L}_0(\Omega, \mathcal{A}_0)$ heißt bedingte Erwartung von X unter $\mathcal{A}_0$ falls für alle $B \in \mathcal{A}_0$ gilt:*

$$\int_B X d\mathbb{P} = \int_B X_0 d\mathbb{P} \tag{1.1}$$

Dabei ist vorausgesetzt, dass $X_0 := \mathbb{E}[X|\mathcal{A}_0]$ entweder nichtnegativ oder integrierbar ist.

Bemerkung 1.1.3.

- *Bedingte Erwartungen sind Zufallsvariablen, keine Zahlen!*
- *Die bedingte Erwartung $\mathbb{E}[X|\mathcal{A}_0]$ ist nur bis auf $\mathbb{P}$-Nullstellen eindeutig bestimmt.*
  Genauer: $\mathbb{E}[X|\mathcal{A}_0]$ bezeichnet die Äquivalenzklasse der $\mathcal{A}_0$-messbaren Zufallsvariablen die 1.1 erfüllen.
- *Bedingte Erwartungen verhalten sich in vielerlei Hinsicht wie Erwartungen.*

Proposition 1.1.4. *Gegeben seien Zufallsvariablen $X, Y \in \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P})$ (oder $\mathcal{L}^0_+(\Omega, \mathcal{A}, \mathbb{P})$). Dann gilt:*

(i) Für $\alpha, \beta \in \mathbb{R}$ (bzw. $\alpha, \beta \in \mathbb{R}_+$) gilt:

$$\mathbb{E}[\alpha X + \beta Y|\mathcal{A}_0] = \alpha\mathbb{E}[X|\mathcal{A}_0] + \beta\mathbb{E}[X|\mathcal{A}_0] \qquad \mathbb{P}\text{-fast sicher}$$

(ii) Aus $X \leq Y$ $\mathbb{P}$-fast sicher folgt:

$$\mathbb{E}[X|\mathcal{A}_0] \leq \mathbb{E}[Y|\mathcal{A}_0] \qquad \mathbb{P}\text{-fast sicher} \qquad \text{(Monotonie)}$$

(iii) Für eine isotone Folge $(X_n)_{n\in N}$ in $\mathcal{L}^0_+(\Omega, \mathcal{A}, \mathbb{P})$ gilt:

$$\sup_n \mathbb{E}[X_n|\mathcal{A}_0] = \mathbb{E}\left[\sup_n X_n|\mathcal{A}_0\right] \qquad \mathbb{P}\text{-fast sicher} \qquad \text{(monotone Konvergenz)}$$

Aus (ii) folgt insbesondere:

- *X=Y $\mathbb{P}$-fast sicher $\Rightarrow \mathbb{E}[X|A_0] = \mathbb{E}[Y|A_0]$ $\mathbb{P}$-fast sicher*
- *$X \geq 0$ $\mathbb{P}$-fast sicher $\Rightarrow \mathbb{E}[X|A_0] \geq 0$ $\mathbb{P}$-fast sicher*
- *$|\mathbb{E}[X|\mathcal{A}_0]| \leq \mathbb{E}[|X||\mathcal{A}_0]$ $\mathbb{P}$-fast sicher*

Beweis. *(i) Gegeben seien $\mathbb{E}[X|A_0]$ und $\mathbb{E}[Y|A_0]$, $B \in \mathcal{A}_0$ beliebig:*

$$\begin{aligned}\int_B [\alpha X + \beta Y]\, d\mathbb{P} &\overset{Lin.}{=} \alpha \int_B X d\mathbb{P} + \beta \int_B Y d\mathbb{P} \overset{1.1}{=} \alpha \int_B \mathbb{E}[X|\mathcal{A}_0]\, d\mathbb{P} + \beta \int_B \mathbb{E}[Y|\mathcal{A}_0]\, d\mathbb{P} \\ &= \int_B \underbrace{(\alpha\mathbb{E}[X|\mathcal{A}_0] + \beta\mathbb{E}[X|\mathcal{A}_0])}_{\mathcal{A}_0\text{-messbare Zufallsvariable}} d\mathbb{P}\end{aligned}$$

$\Rightarrow (i)$

(ii) Monotonie in 1.1

(iii) Sei $X_\infty = \sup_n X_n$. Dann gilt mit monotoner Konvergenz für alle $B \in \mathcal{A}$:

$$\int_B X_\infty d\mathbb{P} = \int_B \sup_n X_n d\mathbb{P} = \int_B \sup_n (X_n 1_B) d\mathbb{P}$$

$$= \sup_n \int_B X_n d\mathbb{P} = \sup_n \int_B X_n^0 d\mathbb{P} = \int_B \sup_n X_n^0 d\mathbb{P} = \int_B X_\infty^0 d\mathbb{P}$$

mit $X_n^0 = \mathbb{E}\left[X_n|\mathcal{A}_0\right], \qquad X_\infty^0 = \sup_n X_n^0 \ \in \mathcal{L}^0(\mathcal{A}_0).$

□

Korollar 1.1.5. *(Jensen-Ungleichung für bedingte Erwartungen)*

Sei $\phi : \mathbb{R} \longrightarrow \mathbb{R}$ *konvex und* X, $\phi(X) \in \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P})$.
Dann gilt:

$$\phi\left(\mathbb{E}\left[X|\mathcal{A}_0\right]\right) \le \mathbb{E}\left[\phi(X)|\mathcal{A}_0\right] \qquad \mathbb{P}\text{-fast sicher}$$

Insbesondere für alle $p \ge 1$

$$\mathbb{E}\left[X|\mathcal{A}_0\right] \le \mathbb{E}\left[|X|^p|\mathcal{A}_0\right]^{\frac{1}{p}} \qquad \mathbb{P}\text{-fast sicher}$$

Beweis. Jedes konvexe ϕ lässt sich darstellen als Supremum abzählbar vieler affin linearer Funktionen:

$$\phi(X) = \sup_{y \in Y} l_y(X) \qquad \text{mit} \quad x \mapsto l_y(x) = ay_+ b_y x \qquad Y \text{ abzählbar}$$

$$\Rightarrow \mathbb{E}\left[\phi(X)|\mathcal{A}_0\right] = \mathbb{E}[\sup l_y(X)|\mathcal{A}_0] \overset{Mon.}{\ge} \sup_y \mathbb{E}\left[l_y(X)|\mathcal{A}_0\right] = \sup_y l_y(\mathbb{E}\left[X|\mathcal{A}_0\right]) = \phi(\mathbb{E}\left[X|\mathcal{A}_0\right])$$

□

Satz 1.1.6. *(i) Für* $X \in L^2(\Omega, \mathcal{A}, \mathbb{P})$ *und* $\mathcal{A}_0 \subset \mathcal{A}$ *existiert ein eindeutiges* $Y \in L^2(\Omega, \mathcal{A}_0, \mathbb{P})$ *mit minimalem Abstand* $\|X - Y\|_{L_2}$.
Die Funktion Y *ist die „Projektion von* X *auf* $L^2(\Omega, \mathcal{A}_0, \mathbb{P})$*"*

(ii) Diese Projektion stimmt mit der bedingten Erwartung von X *unter* $\mathcal{A}_0$ *überein.*

Bemerkung: *Hierbei ist* $L^2(\Omega, \mathcal{A}, \mathbb{P})$ *der Raum der Äquivalenzklassen in* $\mathcal{L}^2(\Omega, \mathcal{A}, \mathbb{P})$.
Es ist ein Hilbert-Raum (d.h. vollständig und mit innerem Produkt) mit Norm $\left(\|Y\|_2 = \mathbb{E}\left[Y^2\right]\right)^{\frac{1}{2}}$

Beweis. *(i) Sei* $\alpha = \inf\left\{\|X - Y\|_{L^2}^2 : Y \in L^2(\Omega, \mathcal{A}_0, \mathbb{P})\right\}$ *sowie eine minimierende Folge* $(Y_n)_n \in L^2(\Omega, \mathcal{A}_0, \mathbb{P})$ *mit* $\|X - Y_n\|_{L_2} = \alpha + \epsilon_n$, $\epsilon_n \searrow 0$.

Behauptung: $(Y_n)_n$ *ist eine Cauchy-Folge.*
Setze $Y_{n,k} = \frac{1}{2}(Y_n + Y_k) \quad \in L^2(\Omega, \mathcal{A}_0, \mathbb{P})$.

$$\Rightarrow \alpha \le \|X - Y_{n,k}\|_{L_2}^2 \le \frac{1}{2}\|X - Y_n\|_{L_2}^2 + \frac{1}{2}\|X - Y_k\|_{L_2}^2 - \frac{1}{4}\|Y_k - Y_n\|_{L_2}^2$$

$$\le \frac{1}{2}(\alpha + \epsilon_n) + \frac{1}{2}(\alpha + \epsilon_k) - \frac{1}{4}\|Y_n - Y_k\|_{L_2}^2$$

$$\Rightarrow \|Y_n - Y_k\|^2 \le 2(\epsilon_n + \epsilon_k) \longrightarrow 0 \qquad \text{für } n, k \to \infty$$

$\Rightarrow (Y_n)_n$ *ist eine Cauchy Folge in* $L^2(\Omega, \mathcal{A}_0, \mathbb{P})$. *Dies ist ein Hilbert-Raum.*
$\Rightarrow$ *es existiert ein eindeutiges* $Y \in L^2(\Omega, \mathcal{A}_0, \mathbb{P})$ *mit* $Y_n \longrightarrow Y$ *in* $L^2(\Omega, \mathcal{A}_0, \mathbb{P})$. *Man kann zeigen:* X-Y *ist orthogonal zu* $L^2(\Omega, \mathcal{A}_0, \mathbb{P})$, *das heißt:*
Für alle $Z \in L^2(\Omega, \mathcal{A}_0, \mathbb{P})$ *gilt:*

$$\mathbb{E}\left[(X - Y)Z\right] = 0 \tag{1.2}$$

(Dadurch ist Y *eindeutig bestimmt.)*

(ii) Aus der Jensen-Ungleichung mit $\phi(t) = |t|^2$ folgt:

$$\mathbb{E}\left[X|\mathcal{A}_0\right]^2 \leq \mathbb{E}\left[X^2|\mathcal{A}_0\right]$$

und damit:

$$\mathbb{E}[\underbrace{\mathbb{E}\left[X|\mathcal{A}_0\right]}_{X_0}{}^2] \leq \mathbb{E}\left[\mathbb{E}\left[X^2|\mathcal{A}_0\right]\right]$$

$\Rightarrow X_0 \in L^2(\Omega, \mathcal{A}_0, \mathbb{P})$ *falls* $X \in L^2(\Omega, \mathcal{A}, \mathbb{P})$ *(gezeigt: quadratintegrabel)*
Behauptung: $X - X_0$ *ist orthogonal zu* $L^2(\Omega, \mathcal{A}_0, \mathbb{P})$.
Sei $B \in \mathcal{A}_0$, *dann gilt:*

$$\mathbb{E}\left[(X - X_0)1_B\right] = \int_B X d\mathbb{P} - \int_B X_0 d\mathbb{P} = 0 \Rightarrow 1.2$$

□

Korollar 1.1.7. *Für alle X in* $\mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P})$ *existiert eine eindeutige Version der bedingten Erwartung* X_0 *in* $L^1(\Omega, \mathcal{A}_0, \mathbb{P})$.

Beweis. *o.B.d.A* $X \geq 0$
Wähle $X_n = \min\{X, n\} \in L^2$

$$X_n \nearrow X \qquad \mathbb{E}\left[X_n|\mathcal{A}_0\right] \in L^2\left(\Omega, \mathcal{A}_0, \mathbb{P}\right)$$

$$\Rightarrow \mathbb{E}\left[X|\mathcal{A}_0\right] = \sup_n \mathbb{E}\left[X_n|\mathcal{A}_0\right] \qquad \textit{existiert und hat die gewünschten Eigenschaften.}$$

Alternativer Beweis: (für die Existenz der bedingten Erwartungen)
Gegeben sei $X \in \mathcal{L}^1_+(\Omega, \mathcal{A}, \mathbb{P})$, $\mathcal{A}_0 \subset \mathcal{A}$. *Betrachte zwei Maße* $\mathbb{P}_0$, $\mathbb{Q}_0$ *auf* $(\Omega, \mathcal{A})$:

$$\mathbb{P}_0 = \mathbb{P}|_{\mathcal{A}_0} \qquad \mathbb{Q}_0(B) = \int\limits_B X d\mathbb{P} \qquad \textit{für } B \in \mathcal{A}_0$$

Es gilt: $\mathbb{Q}_0$ *ist absolut stetig bezüglich* $\mathbb{P}_0$ *(*$\mathbb{Q}_0 << \mathbb{P}_0$*). im folgenden Sinne:*

$$\mathbb{P}_0(B) = 0 \Rightarrow \mathbb{Q}_0(B) = 0 \qquad (\forall B \in \mathcal{A}_0) \tag{1.3}$$

Verwende Satz von Radon Nikodym:
aus 1.3 folgt:

$$\exists X_0 \in \mathcal{L}^1_+(\Omega, \mathcal{A}_0, \mathbb{P}_0) \quad \textit{sodass } \mathbb{Q}_0 = X_0\mathbb{P}_0$$

im Sinne: $\mathbb{Q}_0(B) = \int\limits_B X_0 d\mathbb{P}_0 \qquad (\forall B \in \mathcal{A}_0)$
Zusammen erhält man:

$$\int\limits_B X_0 d\mathbb{P} = \int\limits_B X_0 d\mathbb{P}_0 = \mathbb{Q}_0(B) = \int\limits_B X d\mathbb{P} \qquad (\forall B \in \mathcal{A}_0)$$

□

Proposition 1.1.8. *Bedingte Erwartungen haben folgende Eigenschaften:*

(i) Für alle σ-Algebren $\mathcal{A}_1 \subset \mathcal{A}_0 \subset \mathcal{A}$ *und Zufallsvariablen* $X \in \mathcal{L}^1(X, \mathcal{A}, \mathbb{P})$ *gilt:*

$$\mathbb{E}\left[\mathbb{E}\left[X|\mathcal{A}_0\right]|\mathcal{A}_1\right] = \mathbb{E}\left[\mathbb{E}\left[X|\mathcal{A}_1\right]|A_0\right] = \mathbb{E}\left[X|\mathcal{A}_1\right] \qquad \mathbb{P}\textit{-fast sicher}$$

Insbesondere:

$$\mathbb{E}\left[\mathbb{E}\left[X|\mathcal{A}_0\right]\right] = \mathbb{E}\left[X\right]$$

(ii) a) Falls X unabhängig von $\mathcal{A}_0$ *ist, gilt:* $\mathbb{E}\left[X|\mathcal{A}_0\right] = \mathbb{E}\left[X\right] \qquad \mathbb{P}$*-fast sicher*

b) Falls X messbar bezüglich $\mathcal{A}_0$ ist, gilt: $\mathbb{E}[X|\mathcal{A}_0] = X \qquad \mathbb{P}$*-fast sicher*

(iii) Falls X $\mathcal{A}_0$-messbar ist, Y beliebige Zufallsvariable (d.h. $\mathcal{A}$-messbar) und X und Y nichtnegativ oder in $\mathcal{L}^2(x, \mathcal{A}, \mathbb{P})$:

$$\mathbb{E}[XY|\mathcal{A}_0] = X\mathbb{E}[Y|\mathcal{A}_0]$$

Beweis. *(i) Da* $\mathbb{E}[X|A_1]$ *bereits $\mathcal{A}_0$-messbar ist folgt*

$$\mathbb{E}[\mathbb{E}[X|\mathcal{A}_0]|\mathcal{A}_1] = \mathbb{E}[\mathbb{E}[X|\mathcal{A}_1]|\mathcal{A}_0] = \mathbb{E}[X|\mathcal{A}_1]$$

direkt aus (ii)b

(ii) a) Da X unabhängig von $\mathcal{A}_0$ und 1_B messbar bezüglich $\mathcal{A}_0$ ist, folgt, dass X und 1_B unabhängig sind. Damit ergibt sich:

$$\int_B X d\mathbb{P} = \mathbb{E}[X \cdot 1_B] = \mathbb{E}[X] \qquad \mathbb{P}(B) = \int_B \mathbb{E}[X]\, d\mathbb{P}$$

b) folgt direkt aus der Definition.

(iii) Sei zunächst $X = 1_A$ mit $A \in \mathcal{A}_0$. Dann gilt für alle B in $\mathcal{A}_0$

$$\int_B XY d\mathbb{P} = \int_{A\cap B} Y d\mathbb{P} = \int_{A\cap B} \mathbb{E}[Y|\mathcal{A}_0]\, d\mathbb{P} = \int_B \underbrace{X\mathbb{E}[Y|\mathcal{A}_0]}_{\mathcal{A}_0-messbar} d\mathbb{P}$$

Daraus folgt die Behauptung für $X = 1_A$. Durch Erweiterung auf Linearkombinationen und lineare Limiten lässt sich (iii) für alle X in $L^0_+(\Omega, \mathcal{A}_0, \mathbb{P})$ beweisen.

□

Beispiel 1.1.9. *(i) $\mathcal{A}_0 = \sigma(B_i, i \in \mathbb{N})$ mit disjunkter Zerlegung $\Omega = \bigcup_{i=1}^{\infty} B_i \qquad$ mit $\mathbb{P}(B_i) > 0$.*

Dann gilt:

$$\mathbb{E}[X|\mathcal{A}_0] = \sum_{i=1}^{\infty} \underbrace{\frac{1}{\mathbb{P}(B_i)}\mathbb{E}[X \cdot 1_{B_i}]}_{ren.\ Erw. von\ X\ auf\ B_i} 1_{B_i}$$

Zu zeigen:

$$\forall B \in \mathcal{A}_0 : \quad \int_B X d\mathbb{P} = \int_B X_0 d\mathbb{P}$$

Dafür genügt es zu zeigen:

$$\int_{B_i} X d\mathbb{P} = \int_{B_i} X_0 d\mathbb{P} \qquad (\forall i \in \mathbb{N}) \qquad mit\ X_0 = \sum_{i\in S} \frac{1}{\mathbb{P}(B_i)}\mathbb{E}[X \cdot 1_{B_i}]\, 1_{B_i}$$

es gilt:

$$\int_{B_i} X d\mathbb{P} = \mathbb{E}[1_{B_i} X] = \int_{B_i} \left[\frac{1}{\mathbb{P}(B_i)}\mathbb{E}[1_{B_i} X]\, 1_{B_i}\right] d\mathbb{P}$$

(ii) <u>Ziegenproblem:</u>

$\Omega = \Omega^{(1)} \times \Omega^{(2)}$
$\Omega^{(1)} = \{1, 2, 3\}$ *mit Wahrscheinlichkeiten* p_1, p_2, p_3
Entscheidung des Showmasters/Technikers: Wo ist der Porsche?
$\Omega^{(2)} = \{1, 2, 3\}$ *mit Wahrscheinlichkeiten* q_1, q_2, q_3
Entscheidung des Kandidaten (er hat keine Informationen über p_1, p_2, p_3*)*

Strategie ohne Wechseln:
Erfolg auf $\Omega_{ohne} = \{(1,1),(2,2),(3,3)\}$
Wahrscheinlichkeit für Erfolg: $p_1q_1 + p_2q_2 + p_3q_3$
Für $q = \frac{1}{3}$*:* $\frac{1}{3}(p_1 + p_2 + p_3) = \frac{1}{3}$
Strategie mit Wechsel:
Erfolg auf $\Omega_{mit} = \{(1,2),(1,3),(2,1),(2,3),(3,1),(3,2)\}$
Erfolg bei „Wechselstrategie“: $p_1(q_2 + q_3) + p_2(q_1 + q_3) + p_3(q_1 + q_2)$
Für $q = \frac{1}{3}$*:* $\frac{2}{3}(p_1 + p_2 + p_3) = \frac{2}{3}$

1.2. Bedingte Wahrscheinlichkeiten

Definition 1.2.1. *Analog zu bedingten Erwartungen lassen sich bedingte Wahrscheinlichkeiten und bedingte Verteilungen definieren.*

$$\forall C \in \mathcal{A}: \quad \mathbb{P}(C|\mathcal{A}_0) = \mathbb{E}[1_C|\mathcal{A}_0] \qquad \text{bedingte Wahrscheinlichkeit von } C \text{ unter } \mathcal{A}_0$$

Bemerkung: Dies ist eine Zufallsvariable, keine Zahl!
Bedingte Verteilung von $X \in \mathcal{L}^0(\Omega, \mathcal{A}, \mathbb{P})$*:*

$$\mathbb{P}(X \in B|\mathcal{A}_0) = \mathbb{E}[1_B(X)|\mathcal{A}_0] \qquad \text{für } B \in \mathcal{B}(\mathbb{R})$$

Beispiel 1.2.2. $\mathcal{A}_0 = \sigma(B_i, i \in \mathbb{N})$ *mit* $\Omega = \bigcup_{i=1}^{\infty} B_i$ *und* $\mathbb{P}(B_i) > 0$
Dann gilt:

$$\boxed{\mathbb{P}(C|\mathcal{A}_0) = \sum_{i=1}^{\infty} \mathbb{P}(C|B_i) 1_{B_i} \qquad \text{fast sicher}}$$

mit der „klassischen“ bedingten Wahrscheinlichkeit:

$$\mathbb{P}(C|B_i) = \frac{\mathbb{P}(C \cap B_i)}{\mathbb{P}(B_i)}$$

Anwendungen:

Satz 1.2.3. *Satz von der totalen Wahrscheinlichkeit (Bayes)*
$\mathbb{P}(C) = \mathbb{E}[1_C] = \sum_{i=1}^{\infty} \mathbb{P}(C|B_i)\mathbb{P}(B_i)$

Beweis.

$$\mathbb{P}(C) = \mathbb{E}[\mathbb{E}[1_C|A_0]] = \mathbb{E}\left[\sum_{i=1}^{\infty} \mathbb{P}(C|B_i) 1_{B_i}\right]$$

□

Exkurs: *Eine Universität stellt fest:*

$$63\% = \mathbb{P}(\text{Frauen bestehen Examen}) < \mathbb{P}(\text{Männer bestehen Examen}) = 74\%$$

aber in jedem Fach gilt:

$$\mathbb{P}(\text{Frauen bestehen Examen}) > \mathbb{P}(\text{Männer bestehen Examen})$$

konkret (GW: Geisteswissenschaften, NW: Naturwissenschaften):

$$60\% = \mathbb{P}(\text{Frauen bestehen GW}) > \mathbb{P}(\text{Männer bestehen GW}) = 50\%$$

$$90\% = \mathbb{P}(\text{Frauen bestehen NW}) > \mathbb{P}(\text{Männer bestehen NW}) = 80\%$$

Wie ist das zu erklären?

$$90\% = \mathbb{P}(\textit{Frauen studieren GW}) > \mathbb{P}(\textit{Männer studieren GW}) = 20\%$$

Wir betrachten zwei Wahrscheinlichkeitsmaße $\mathbb{P}_F$ und $\mathbb{P}_M$ und die Ereignisse:
G: studiert Geisteswissenschaften,
N: studiert Naturwissenschaften $\qquad G \dot{\bigcup} N = \Omega$
E: Erfolg im Examen

$$\mathbb{P}_F(E) = \mathbb{P}_F(E|N)\mathbb{P}_F(N) + \mathbb{P}_F(E|G)\mathbb{P}_F(G) = 0,9 \cdot 0,1 + 0,6 \cdot 0,9 = 0,63$$

$$\mathbb{P}_M(E) = \mathbb{P}_M(E|N)\mathbb{P}_M(N) + \mathbb{P}_M(E|G)\mathbb{P}_M(G) = 0,8 \cdot 0,8 + 0,5 \cdot 0,2 = 0,74$$

Bemerkung 1.2.4. *Für die bedingte Wahrscheinlichkeit $\mathbb{P}(\cdot|\mathcal{A}_0)$ gelten folgende Eigenschaften:*

(i) $\forall C \in \mathcal{A}_0$:
- $0 \leq \mathbb{P}(C|\mathcal{A}_0) \leq 1 \qquad \mathbb{P}-$ *fast sicher*
- $\mathbb{P}(\emptyset|\mathcal{A}_0) = 0 \qquad \mathbb{P}-$ *fast sicher*
- $\mathbb{P}(\Omega|\mathcal{A}_0) = 1 \qquad \mathbb{P}-$ *fast sicher*

(ii) $\forall C_1 \subset C_2 \qquad \mathbb{P}(C_1|\mathcal{A}_0) \leq \mathbb{P}(C_2|\mathcal{A}_0) \qquad \mathbb{P}-$*fast sicher*

(iii) $\forall$ *disjunkte* $(C_n)_{n\in\mathbb{N}} \qquad \mathbb{P}(\bigcup_{n=1}^{\infty} C_n|\mathcal{A}_0) = \sum_{n=1}^{\infty} \mathbb{P}(C_n|\mathcal{A}_0) \qquad \mathbb{P}$*-fast sicher*

Alles folgt aus entsprechenden Eigenschaften der bedingten Erwartungen. Würden (i)-(iii) für alle Mengen ohne Ausnahmemengen auf ganz Ω gelten, so wäre für jedes $\omega \in \Omega$ die Abbildung $C \mapsto \mathbb{P}(C|\mathcal{A}_0)(\omega)$ ein Wahrscheinlichkeitsmaß.
Für gegebene $C, C_1, C_2, (C_n)_{n\in\mathbb{N}}$ lassen sich Versionen finden, für die (i)-(iii) gelten. Für sogenannte Standard-Borel-Räume (z.B. $\mathbb{R}$) existiert eine Version von

$$\Omega \times \mathcal{A} \to [0,1] \qquad (\omega, C) \mapsto \mathbb{P}(C|\mathcal{A}_0)(\omega),$$

sodass für $\mathbb{P}$-fast alle $\omega \in \Omega$ die Eigenschaften (i)-(iii) für jede Wahl von $C, C_1, C_2, (C_n)_{n\in\mathbb{N}}$ gelten. Im Allgemeinen existiert eine solche Version nicht.
In vielen Fällen betrachtet man bedingte Erwartungen, bedingte Wahrscheinlichkeiten und bedingte Verteilungen unter der Bedingung $\mathcal{A}_0 = \sigma(Y)$.
Für eine Zufallsvariable $Y : \Omega \to \mathbb{R}$ schreibe: $\mathbb{E}[X|Y]$, $\mathbb{P}(C|Y)$ etc.
Aus dem Faktorisierungslemma folgt:

$$X_0 \in \mathcal{L}^0(\Omega, \sigma(Y)) \iff X_0 = g(Y) \qquad \textit{für ein } g \in \mathcal{L}^0(\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

Damit folgt insbesondere, dass für alle $X \in \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P})$ ein g existiert mit:

$$\mathbb{E}[X|Y] = g(Y) \qquad \mathbb{P}\textit{-fast sicher}$$

Dabei ist g eindeutig bis auf $\mathbb{P}_Y$-Nullmengen auf $\mathbb{R}$. Schreibweise:

$$g(Y) = \mathbb{E}[X|Y = y] \quad \textit{für } y \in \mathbb{R}$$

$$\boxed{\textit{Faktorisierte bedingte Erwartung von } X \textit{ unter der Bedingung } Y = y}$$

Es gilt:

$$\mathbb{E}[X|Y](\omega) = g(Y(\omega)) = \mathbb{E}[X|Y = y]\,|_{y=Y(\omega)} \qquad \textit{für } \mathbb{P}\textit{-fast alle } \omega$$

beziehungsweise:

$$\mathbb{E}[X|Y](\cdot) = \mathbb{E}[X|Y = Y(\cdot)] \qquad \mathbb{P}\textit{-fast überall}$$

Beispiel 1.2.5. *Seien $X, Y : \Omega \to \mathbb{R}$ Zufallsvariablen mit gemeinsamer Verteilung $\mathbb{P}_{(X,Y)}$ auf $\mathbb{R}^2$ gegeben durch $f \in \mathcal{L}^2$ mit Dichte-Funktion $f : \mathbb{R}^2 \to [0, \infty]$ nichtnegativ und integrierbar. Das heißt:*

$$\mathbb{P}\left(((X,Y) \in C)\right) = \iint\limits_C f(x,y)dxdy \qquad \text{für alle Borel Mengen } C \subset \mathbb{R}^2$$

Sei ferner X integrierbar. Dann gilt für $\mathbb{P}_Y$-fast alle $y \in \mathbb{R}$:

$$0 < F(y) = \int\limits_{\mathbb{R}} f(x,y)dx < \infty$$

und

$$\mathbb{E}\left[X|Y = y\right] = \frac{1}{F(y)} \int\limits_{\mathbb{R}} xf(x,y)dx =: g(y)$$

Insbesondere also: $\mathbb{E}\left[X|Y\right] = g(y)$ $\mathbb{P}$-fast überall

Beweis. Zunächst sei $0 < F < \infty$ $\mathbb{P}_Y$-fast überall auf $\mathbb{R}$, denn für alle $C \in \mathcal{B}(\mathbb{R})$ gilt:

$$\mathbb{P}_Y(C_2) = P(y \in C_2) = \int\limits_{\mathbb{R}} \left[\int\limits_{C_2} f(x,y)dy\right] dx = \int\limits_{C_2} F(y)dy$$

Mit $C_2 := \{F = 0\}$ folgt die erste Behauptung, analog folgt: $F < \infty$ $\mathbb{P}_y$-fast sicher.
Sei nun $B \in \sigma(Y)$, also z.B. $B = Y^{-1}(A)$ mit $A \in \mathcal{B}(\mathbb{R})$:

$$\int\limits_B g(y)d\mathbb{P} = \int\limits_\Omega 1_A(y)g(y)d\mathbb{P} = \int\limits_{\mathbb{R}} 1_A(y)g(y)d\mathbb{P}_Y(y) = \int\limits_{\mathbb{R}}\int\limits_{\mathbb{R}} 1_A(y)g(y)f(x,y)dxdy$$

$$= \int\limits_A g(y)F(y)dy = \int\limits_A\int\limits_{\mathbb{R}} x\underbrace{f(x,y)dxdy}_{d\mathbb{P}_{(x,y)}} = \int\limits_\Omega x1_A(y)d\mathbb{P} = \int\limits_B Xd\mathbb{P}$$

<u>Also:</u> $g(Y) = \mathbb{E}\left[X|\sigma(Y)\right]$ fast sicher ☐

Definition 1.2.6. *(i) Zwei Messräume $(\Omega, \mathcal{A})$, $(\Omega', \mathcal{A}')$ heißen <u>messbar isometrisch</u> falls es eine Bijektion $\psi : \Omega \to \Omega'$ gibt mit ψ und ψ^{-1} messbar.*

(ii) Ein Messraum $(\Omega, \mathcal{A})$ heißt Standard-Borel-Raum, falls er messbar isometrisch zu einer Borel-Teilmenge von $[0,1]$ (oder $\mathbb{R}$) ist (mit induzierter Borel-σ-Algebra)

Definition 1.2.7. *Ein topologischer Raum heißt <u>polnischer Raum</u>, falls er separabel und vollständig metrisierbar ist.*

Proposition 1.2.8. *(i) Jeder polnische Raum ist homöomorph zu einer G_δ-Menge eines kompakten metrischen Raumes, welcher o.B.d.A als $[0,1]^{\mathbb{N}}$ angenommen werden kann.*

(ii) Jeder polnische Raum ist ein Standard-Borel-Raum.

Lemma 1.2.9. *$\mathbb{R}^d$ ist Standard-Borel.*

Beweis. zeige: $[0,1)^d$ ist Standard-Borel.
Wähle für alle $x \in [0,1)$ eine Darstellung $x = 0.x_1x_2x_3...$ als Binärzahl mit unendlich vielen Nullen
Für $(x^1, ..., x^d) \in [0,1)^d$ analog: $x^i = 0.x_1^ix_2^ix_3^i...$
Gegeben ein solches $(x^1, ..., x^d) \in [0,1)^d$ bilde:

$$x = x_1^1x_1^2x_1^3, ..., x_1^d, x_2^1, x_2^2, x_2^3, ... \quad \in [0,1)$$

Die so definierte Abbildung $(0,1]^d \to (0,1]$ ist bijektiv, messbar und ϕ^{-1} ist messbar. ☐

Lemma 1.2.10. $[0,1)^{\mathbb{N}}$ *ist Standart-Borel.*

Beweis. Gegeben $x \in [0,1)$, $x = x_1x_2x_3\ldots$
$x^1 = 0, x_1x_2x_4x_7\ldots,\qquad x^2 = 0, x_3x_5x_8\ldots,\qquad x^3 = 0, x_6x_9\ldots$
$\Rightarrow (X^i)_{i\in\mathbb{N}} \in [0,1)^{\mathbb{N}}$

$$\phi : [0,1) \longrightarrow [0,1)^{\mathbb{N}}, \quad x \longmapsto (x^i)_{i\in\mathbb{N}} \qquad \textit{bijektiv, messbar, } \phi^{-1} \textit{ messbar}$$

□

Bedingte Wahrscheinlichkeiten
$X_0 = \mathbb{P}(C|\mathcal{A}_0)(\omega) = \mathbb{E}\left[1_c|\mathcal{A}_0\right](\omega)$ definiert durch:

$$\int_A 1_C d\mathbb{P} = \int_A X_0 d\mathbb{P} \qquad (\forall A \in \mathcal{A}_0)$$

Ziel: Wahl von Versionen $\mathbb{P}(C|\mathcal{A}_0)(\omega)$, sodass für $\mathbb{P}$-fast alle $\omega \in \Omega$ gilt:

$$C \longmapsto \mathbb{P}(C|\mathcal{A}_0)(\omega) \qquad \text{ist ein Wahrscheinlichkeitsmaß auf } (\Omega, \mathcal{A}, \mathbb{P})$$

Definition 1.2.11. *Gegeben seien zwei Messräume $(S_1, \mathcal{S}_1)$,$(S_2, \mathcal{S}_2)$. Ein Markov-Kern von $(S_1, \mathcal{S}_1)$ nach $(S_2, \mathcal{S}_2)$ ist eine Abbildung*

$$k : S_1 \times \mathcal{S}_2 \to [0,1]$$

mit den folgenden Eigenschaften:

(i) $\forall C \in \mathcal{S}_2 \qquad x \mapsto k(x, C) \qquad$ *ist ein $\mathcal{S}_1$-messbar.*

(ii) $\forall x \in S_1 \qquad C \mapsto k(x, C) \qquad$ *ist ein Wahrscheinlichkeitsmaß auf $(S_2, \mathcal{S}_2)$*

Definition 1.2.12. *Gegeben sei ein Wahrscheinlichtsraum $(\Omega, \mathcal{A}, \mathbb{P})$ und eine σ-Algebra $\mathcal{A}_0 \subset \mathcal{A}$. Dann heißt ein Markov-Kern k von $(\Omega, \mathcal{A}_0)$ nach $(\Omega, \mathcal{A})$ reguläre bedingte Wahrscheinlichkeit unter $\mathcal{A}_0$ falls gilt:*

$$\forall C \in \mathcal{A}, \qquad \mathbb{P}(C|\mathcal{A}_0)(\omega) = k(\omega, C) \quad \textit{für } \mathbb{P}\textit{-fast alle } \omega \in \Omega$$

Je zwei solche Markov-Kerne k, k' stimmen für $\mathbb{P}$-fast alle ω überein falls gilt:

$$k(\omega, C) = k'(\omega, C) \qquad (\forall C \in \mathcal{A})$$

Zu Markov-Kernen
Ein Markov-Kern k ist ein parameterabhängiges Wahrscheinlichkeitsmaß auf $(S_2, \mathcal{S}_2)$.
Wir schreiben $k(x, dy)$ und definieren für messbare Funktionen $f : S_2 \to \mathbb{R}$:

$$kf(x) = \int_S f(y)k(x, dy)$$

Für einen Markov-Kern k von $(S_1, \mathcal{S}_1)$ nach $(S_2, \mathcal{S}_2)$ und k' von $(S_2, \mathcal{S}_2)$ nach $(S_3, \mathcal{S}_3)$ ist

$$\bar{k}(x, C) = \int_S k'(y, C)k(x, dy)$$

ein Markov-Kern von $(S_1, \mathcal{S}_1)$ nach $(S_3, \mathcal{S}_3)$.
Für S_1, S_2 diskret (z.B. endlich) ist jeder Markov-Kern durch eine stochastische Matrix $(k_{ij})_{i\in S_1, j\in S_2}$ gegeben mit

$$k(i, C) = \sum_{j\in C} k_{ij} \qquad \text{mit} \sum_{j\in S_2} k_{ij} = 1$$

Für Wahrscheinlichkeitsmaße μ auf $(S_1, \mathcal{S}_1)$ definiert man ein Wahrscheinlichkeitsmaß μk auf $(S_2, \mathcal{S}_2)$ durch:

$$(\mu k)(C) = \int_{S_1} k(x, C)\mu(dx)$$

1.3. Bedingte Verteilungen

Gegeben sei ein Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$, ein Standard-Borel-Raum $(S, \mathcal{S})$ und eine Zufallsvariable $X : \Omega \to S$
Ziel: Konstruktion einer regulären Version der bedingten Verteilung $\mathbb{P}_x(\cdot|\mathcal{A}_0)$

Satz 1.3.1. *Es existiert ein Markov-Kern k, sodass für alle $C \in \mathcal{S}$ gilt:*

$$\mathbb{P}_X(C|\mathcal{A}_0)(\omega) = k(\omega, C) \qquad \textit{für } \mathbb{P}\textit{-fast alle } \omega \in \Omega$$

Beweis. 1) Reduktion auf $S = \mathbb{R}$:
Da $(S, \mathcal{S})$ ein Standard-Borel-Raum ist, existiert eine bijektive und messbare Abbildung $\phi : S \to B \in \mathcal{B}(\mathbb{R})$ sodass auch ϕ^{-1} messbar ist. Ersetze nun X durch $\phi(X)$ und S durch $\mathbb{R}$:

$$\Omega \xrightarrow{X} S \xrightarrow{\phi} \mathbb{R}$$

und ersetze $C \in \mathcal{S}$ durch $B \in \mathcal{B}(\mathbb{R})$. Jetzt können wir o.B.d.A annehmen: $S = \mathbb{R}$

2) Konstruktion einer Verteilungsfunktion
Es existiert ein $\Omega_1 \in \mathcal{A}$ mit $\mathbb{P}(\Omega_1) = 1$ und es existiert eine Funktion $G : \Omega \times \mathbb{Q} \to \mathbb{R}$ mit:

(i) $\forall \omega \in \Omega_1 \qquad q \longmapsto G(\omega, q) \qquad$ *isoton*

(ii) $\forall q \in \mathbb{Q} \qquad \omega \longmapsto G(\omega, q) \qquad$ *ist eine Version von* $\mathbb{P}\left(X \leq q|\mathcal{A}_0\right) = \mathbb{P}(\left(-\infty, q\right]|\mathcal{A}_0))$

Für alle $q_1 < q_2 \in \mathbb{Q}$ gilt: $\mathbb{P}(X \leq q_1|A_0) \leq \mathbb{P}(X \leq q_2|A_0) \qquad \mathbb{P} -$ *fast überall*

3) Fortsetzung auf $\Omega_1 \times \mathbb{R}$:
Betrachte die Funktion $F : \Omega_1 \times \mathbb{R} \longrightarrow \mathbb{R}$ definiert durch:

$$F(\omega, x) := \inf_{q \in \mathbb{Q}, q > x} G(\omega, q)$$

Daraus folgt, dass die Abbildung $x \mapsto F(q, x)$ für alle $\omega \in \Omega$ isoton und rechtsstetig ist mit:

$$\lim_{x \to -\infty} F(\omega, x) = 0 \qquad\qquad \lim_{x \to \infty} F(\Omega, q)$$

Das bedeutet, dass die Abbildung $x \mapsto F(\omega, x)$ eine Verteilungsfunktion eines eindeutig festgelegten Wahrscheinlichkeitsmaßes $k(\omega, \cdot)$ auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ ist, mit:

$$k(\omega, (-\infty, x]) = F(\omega, x) \qquad (\forall \omega \in \Omega, \forall x \in \mathbb{R})$$

Für jedes $\omega \in \Omega \setminus \Omega_1$ wähle beliebiges Wahrscheinlichkeitsmaß ν und setze $k(\omega, \cdot) = \nu(\cdot)$.

4) Identifizieren als bedingte Wahrscheinlichkeit
Für alle $x \in \mathbb{R}$ und für $\mathbb{P}$-fast alle $\omega \in \Omega$ gilt wegen der monotonen Konvergenz von bedingten Erwartungen:

$$k(\omega, (.\infty, x]) = F(\omega, x) = \inf_{q \in \mathbb{Q}, q > x} G(\omega, q) = \inf_{q \in \mathbb{Q}, q > x} P(X \leq q|\mathcal{A}_0)(\omega) = \mathbb{P}(X \leq x|\mathcal{A}_0)(\omega)$$

Betrachte nun $\mathcal{D} = \{C \in \mathcal{B}(\mathbb{R}) : k(\omega, C|_0) = \mathbb{P}(X \in C|\mathcal{A}_0)(\omega)$ für $\mathbb{P} -$ fast $alle \omega in \Omega\}$
$\Rightarrow \qquad (-\infty, x] \in \mathcal{D}$, *$\mathcal{D}$ ist stabil unter $\cap$ und ist stabil unter disjunkten Vereinigungen:*

$$C = \dot{\bigcup_n} C_n \ \textit{mit } C_n \in \mathcal{D} : k(\omega, C) = \sum_n k(\omega, C_n) = \sum_n \mathbb{P}(X \in C_n|\mathcal{A}_0)(\omega) = stackrel \mathbb{P} - f.s. = \mathbb{P}(X \in C|\mathcal{A}_0)(\omega)$$

$\Rightarrow \qquad \mathcal{D}$ *enthält ein Dynkin-System, das von den Intervallen $(-\infty, x]$ erzeugt wird.*
$\Rightarrow \qquad \mathcal{D} = \mathcal{B}(\mathbb{R})$

$\square$

Korollar 1.3.2. *Sei* $(\Omega, \mathcal{A})$ *ein Standard-Borel-Raum. So existiert für alle* $\mathcal{A}_0 \subset \mathcal{A}$ *und für alle Wahrscheinlichkeitsmaße* $\mathbb{P}$ *auf* $(\Omega, \mathcal{A})$ *eine reguläre bedingte Wahrscheinlichkeit* $\mathbb{P}(\cdot|\mathcal{A}_0)$*. Das heißt, es existiert ein Markov-Kern* $k : (\Omega, \mathcal{A}) \to [0,1]$ *mit:*

$$\mathbb{P}(C|\mathcal{A}_0) = k(\cdot, C) \qquad (\forall C \in \mathcal{A})$$

Beweis. *Satz mit* $X = id$ *und* $\Omega = \mathcal{S}$ □

Desintegration

Gegeben sei ein Wahrscheinlichkeitsmaß μ auf einem Produkt-Raum $S = S_1 \times S_2$ mit der Produkt σ-Algebra. Die Bildmaße μ_1 bzw. μ_2 auf S_1 bzw. S_2 unter den Projektionen:

$$\pi_1 : S \longrightarrow S_1 \qquad (x_1, x_2) \longmapsto x_1$$

$$\pi_2 : S \longrightarrow S_2 \qquad (x_1, x_2) \longmapsto x_2$$

heißen Marginalien oder Randverteilungen von μ, also:

$$\mu_1 = (\pi_1)_* \mu, \qquad \mu_1(B) = \mu(B \times S_2), \qquad \mu_2(C) = \mu(S_1 \times C)$$

Ziel: Konstruktion eines Markov-Kerns $k : S_1 \times \mathcal{S}_2 \to [0,1]$ mit

$$d\mu(x_1, x_2) = k(x_1, dx_2) d\mu_1(x) \qquad \text{(„Desintegration von } \mu \text{ nach } \mu_1\text{“)}$$

Dann gilt für alle messbaren und nichtnegativen Funktionen f auf S:

$$\int_S f d\mu = \int_{S_1} \left[\int_{S_2} f(x_1) k(x_1, dx_2) \right] d\mu(x_1)$$

Korollar 1.3.3. *Seien* $(S_1, \mathcal{S}_1)$ *und* $(S_2, \mathcal{S}_2)$ *Standard-Borel Räume auf* $S = S_1 \times S_2$*. Dann existiert zu jedem Wahrscheinlichkeitsmaß* μ *auf* $(S, \mathcal{S})$ *ein Markov-Kern* $k : S_1 \times \mathcal{S}_2 \to [0,1]$ *mit:*

$$d\mu(x_1, x_2) = k(x_1, dx_2) d\mu_1(x_1)$$

Wobei μ_1 *eine Marginalie ist und* $k(x, \cdot)$ *eindeutig bestimmt ist für* μ_1*-fast alle* x*.*

Beweis. *mit Satz über reguläre bedingte Verteilungen für*

$$\Omega = S = S_1 \times S_2 \; (X = \pi_1, S_1 \textit{ anstelle von } S \textit{ im vorherigen Satz }, \mathbb{P} = \nu, \; \mathcal{A}_0 = \{A \times S_2, A \in \mathcal{S}_1\}$$

□

Kapitel 2

Markov-Ketten

Gegeben sei ein Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ und ein Standard-Borel-Raum $(S, \mathcal{S})$.

Definition 2.0.4. *Eine Folge von Zufallsvariablen $(X_n)_{n\in N_0}$ mit $X_i : \Omega \to S$ und Werten in S heißt Markov-Kette, falls:*

$$\mathbb{P}(X_{n+1} \in B|\mathcal{A}_n) = \mathbb{P}(X_{n+1} \in B|X_n) \qquad (\forall n \in \mathbb{N}_0, \forall B \in \mathcal{S}) \qquad \textit{mit } A_n = \sigma(X_1, ..., X_n) \tag{2.1}$$

Definition 2.0.5. *Das Wahrscheinlichkeitsmaß $\mu_0 = (X_0)_*\mathbb{P}$ heißt „Startverteilung" der Markov-Kette. Definiere analog $\mu_n = (X_n)_*\mathbb{P}$.*
Die Markov-Kerne $p_n(x, B) := \mathbb{P}(x_n \in B|X_{n-1} = x), (n \in \mathbb{N}, x \in S, B \in \mathcal{S})$ heißen „Übergangswahrscheinlichkeiten im n-ten Schritt". Es sind faktorisierte reguläre bedingte Wahrscheinlichkeiten.
Es gilt:

$$\mu_n = \mu_{n-1}p_n \qquad \textit{(Verteilung nach dem n-ten Schritt)} \tag{2.2}$$

denn:

$$\mu_n(B) = \mathbb{P}(X_n \in B) = \int_S \mathbb{P}(X_n \in B|X_{n-1} = x)d\mathbb{P} \quad X_{n-1}(x) = \int_S p_{n-1}(x, B)d\mu_{n-1}(x) = (\mu_{n-1}p_n)(B)$$

Iteration von 2.2 ergibt:

$$\mu_n = \mu_{n-1}p_n = (\mu_{n-2}p_{n-1})p_n = \underbrace{\mu_0}_{Startv.} \underbrace{p_1p_2 \cdots p_n}_{:=p_{0,n}} = \mu_0 p_{0,n}$$

Setzt man für alle $m < n : p_{m,n} = p_{m+1}p_{m+2} \cdots p_{n-1}p_n$, so gilt: $\mu_n = \mu_m p_{m,n}$.
Mit anderen Worten gilt:

$$\mathbb{P}(X_n \in B) = \int_S \underbrace{\mathbb{P}(X_n \in B|X_m = x)}_{p_{m,n}(x,B)} d\mathbb{P}_{x_n}(x)$$

Wiederholung (Verknüpfung von Markov-Kernen)

Gegeben seien Markov-Kerne $p, q \in (S, \mathcal{S})$, dann existiert ein Markov-Kern pq auf $(S, \mathcal{S})$ definiert als:

$$(pq)(x, B) = \int_S p(x, dy)q(y, B)$$

Es gilt:

$$(pq)f = \int_S (pq)(x, dy)f(y) = \int_S \underbrace{\int_S f(y)q(z, dy)}_{(qf)(z)} p(x, dz) = p(qf)(x) \qquad \textit{(assoziativ)}$$

Jeder Markov-Kern p auf $(S, \mathcal{S})$ definiert einen linearen Operator auf $\mathcal{L}^\infty(S, \mathcal{S})$ mit:

$$f \geq 0 \Rightarrow pf \geq 0, \; p1 = 1$$

Neutrales Element: $p(x,dy) = \delta_x(dy) \qquad \Rightarrow pq = q \quad \forall q.$
Für abzählbare S *ist jeder Markov-Kern* $p(x,dy)$ *gegeben durch:*

$$p(i,B) = \sum_{j \in B} p_{ij} \qquad (\forall i \in S, \forall B \subset S) \qquad bzw.\ p_{ij} = p(i,\{j\})$$

Zum Markov-Kern $(pq)(x,dy)$ *gehört die Matrix* $(pq)_{ij} = \left(\sum_{k \in S} p_{ik} q_{kj} \right)_{ij}$

Definition 2.0.6. *(i) Eine Markov-Kette* $(X_n)_{n \in N_0}$ *heißt* <u>*homogen*</u> *(=zeithomogen), falls die Übergangswahrscheinlichkeiten* p_n *iid sind.*
Mit anderen Worten: Es existiert <u>*ein*</u> *Markov-Kern auf* $(S, \mathcal{S})$ *mit der Eigenschaft:*

$$\mathbb{P}(X_n \in B | X_0, \ldots, X_{n-1}) = p(X_{n-1}, B) \qquad \forall n \in \mathbb{N}, \forall B \in \mathcal{S}$$

(ii) Die Markov-Kette $(X_n)_{n \in N_0}$ *heißt* <u>*translationsinvariant*</u> *(=räumlich homogen), falls* S *eine additive Gruppe ist, und es gilt:*

$$p_n(x,B) = p_n(0, B-x) \qquad \forall n \in \mathbb{N}, \forall x \in S,\ \forall B \subset S \qquad (B - x := \{b - x | b \in B\})$$

Beispiel 2.0.7. *(i) Summe von iid Zufallsvariablen*
Gegeben sei eine Folge $(\xi_i)_{\cdot \in N}$ *unabhängiger, identisch verteilter Zufallsvariablen* ξ_i *auf* $(\Omega, \mathcal{A}, \mathbb{P})$ *und eine beliebige Zufallsvariable* X_0. *Definiere:*

$$X_n := \sum_{i=1}^{n} \xi_i + X_0 \qquad \nu = (\xi_i)_* \mathbb{P}, \mu = (X_0)_* \mathbb{P}$$

Dann ist $(X_n)_{n \in \mathbb{N}_0}$ *Markov-Kette, denn:*

$$\mathbb{P}(X_n \in B | X_0, \ldots, X_{n-1}) = \mathbb{P}(X_n \in B | X_{n-1})$$

Da ξ_n *unabhängig von* X_0 *und von* $\xi_i (i < n)$ *ist, ist sie homogen und translationsinvariant:*

$$\mathbb{P}(X_n \in B | X_{n-1} = x) = \mathbb{P}(x + \xi_n \in B | X_0 + \sum_{i_1}^{n-1} \xi_i = x) = \mathbb{P}(x + \xi_n \in B) = \nu(B - x)$$

Es gilt auch die Umkehrung:
Sei $(X_n)_{n \in N_0}$ *eine homogene, translationsinvariante Markov-Kette, so betrachte für alle* i*:* $\xi_i := X_i - X_{i-1}$ *Dann sind* ξ_i *unabhängig und identisch verteilte Zufallsvariablen.*

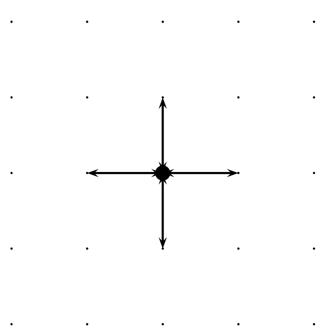
(ii) Irrfahrten („random walks“) auf Graphen
Gegeben sei ein Graph (V,E), wobei V die Menge der Knoten („vertices“) und E die Menge der Kanten (=Paare auf V) („edges“) ist.

$$x \sim y \iff (x,y) \in E,\ deg(x) = \sharp \{y \in S | y \sim x\}$$

Eine Irrfahrt auf S ist eine homogene Markov-Kette auf S mit:

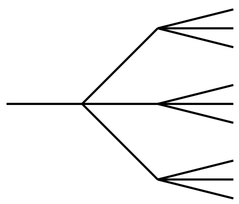
$$p_{ij} = \begin{cases} \frac{1}{deg(i)} & \text{für } j \sim i \\ 0 & \text{sonst} \end{cases}$$

a) Symmetrische Irrfahrt auf $\mathbb{Z}^d$, $S = \mathbb{Z}^d$



$x \sim y \iff |x-y| = 1, \qquad \Rightarrow deg(G) = \frac{1}{2d}$

b) Baum T_k, deg(G)=$\frac{1}{k}$



c) Perkolationscluster, Zufallsgraphen

d) Galton-Watson Verzweigungsprozess
Sei q ein Wahrscheinlichkeitsmaß auf $\mathbb{N}_0$ (=Verteilung der Anzahl der Nachkommen eines Individuums) (z.B. $q = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$)
Ferner seien $(Y_{n,i})_{n,i\in\mathbb{N}_0}$ iid Zufallsvariablen mit:

$$\mathbb{P}(Y_{n,i} = k) = q_{ik} \qquad \forall k \in \mathbb{N}_0, i \in \mathbb{N}$$

$Y_{n,i}$:=Anzahl der Nachkommen der (n+1)-ten Generation des i-ten Individuums aus der n-ten Generation
Setzte nun:

$$X_0 = 1, \quad \dots \quad X_{n+1} = \sum_{i=1}^{X_n} Y_{n,i} = \text{ Anzahl der Individuen in der (n+1)-ten Generation}$$

Der Prozess $(X_n)_{n\in N_0}$ heißt Galton-Watson-Prozess mit Nachkommensverteilung q.

Proposition 2.0.8. *Der Galton-Watson-Prozess $(X_n)_{n\in N_0}$ ist eine homogene Markov-Kette mit Übergangswahrscheinlichkeit*

$$p(k,\cdot) = q^{*k} \qquad \text{k-faches Faltungsprodukt von } q \text{ auf } \mathbb{N}_0$$

Das heißt:

$$p(k,j) = q^{*k}(\{j\}) = \sum_{i=0}^{j} q^{*(k-1)}(j-i)q(i) \qquad (\text{mit } q^{*0} = \delta_0)$$

Beweis. *Wir wählen $S = \mathbb{N}_0$ und betrachten:*

$$\mathbb{P}(X_1 = x_1 | X_0 = x_0, \dots, X_{n-1} = x_{n-1}) = \mathbb{P}(\underbrace{Y_{n+1} + \cdots Y_{n-1,X_{n-1}} = x_1}_{\text{unabhämgige ZV}}) = q^{*(x_{n-1})}(x_n) = p(x_{n-1}, \dots, x_n)$$

□

Proposition 2.0.9. *Eine Folge von Zufallsvariablen $(X_n)_{n\in N_0}$ ist genau dann eine Markov-Kette, wenn für alle $n \in \mathbb{N}$ ein Markov-Kern p_n auf $(S, \mathcal{S})$ existiert mit:*

$$\mathbb{P}\underbrace{(X_{n+1} \in B | X_0, \dots, X_n)}_{:=Z} = p_{n+1}(X_n, B) \qquad \forall B \in \mathcal{S} \tag{2.3}$$

Beweis. „$\Leftarrow$“ *Aus 2.3 folgt: Z ist $\sigma(X_n)$-messbar, und damit:*

$$Z = \mathbb{E}[Z|X_n] = \mathbb{E}[\mathbb{E}[1_B(X_{n+1})|X_0, \dots, X_n]|X_n] = \mathbb{E}[1_B(X_{n+1})|X_n] = \mathbb{P}(X_{n+1} \in B|X_n)$$

„⇒“ *Seien* $X, Y : \Omega \to S$ *messbar und* $\mu = (X,Y)_*\mathbb{P}$ *deren gemeinsame Verteilung auf* S^2. *Ferner sei* μ_1 *die erste Marginalie:*

$$\mu_1 = (\pi)_*\mu = (\pi \circ (X,y))_*\mathbb{P} = X_*P)$$

Aus Desintegration von Wahrscheinlichkeitsmaßen auf Produkträumen von Standard-Borel Räumen folgt die Existenz eines Markov-Kerns p *auf* $(S, \mathcal{S})$ *mit* $d\mu(x,y) = p(x,dy)d\mu_1(x)$. *Genauer:*

$$\int_{S^2} f d\mu = \int_S \left[\int_f (x,y)p(x,dy) \right] d\mu_1(x) \qquad (\forall f \in \mathcal{L}_+0(S^2, \mathcal{S}^2))$$

Wende dies nun auf $X = X_n$ *und* $Y = X_{n-1}$ *an:*

$$\mathbb{P}(X_n \in A, X_{n-1} \in B) = \mu(A \times B) = \int_A p(x,B)d\mu_1(x) \qquad (\forall A, B \in \mathcal{S})$$

Speziell gilt für $C \in \sigma(X_n)$, *also* $C = X_n^{-1}(A)$ *mit* $A \in \mathcal{S}$:

$$\int_C 1_B(X_{n+1})d\mathbb{P} = \mathbb{P}(C \cap \{X_{n+1} \in B\}) = \int_{X_n^{-1}=C} Cp(X_n, B)d\mathbb{P}$$

Daraus folgt:

$$\mathbb{P}(X_{n+1} \in B | X_n) = \mathbb{E}\left[1_B(X_{n+1}) | X_n\right] = p(X_n, B) \qquad \mathbb{P}\text{-fast überall auf } \Omega$$

Und damit:

$$\mathbb{P}(X_{n+1} \in B | X_n = x) = p(x,B) \qquad \mathbb{P}_{X_n}\text{-fast überall auf } S$$

□

Definition 2.0.10. *Zwei Markov-Kerne* $(X_n)_{n \in N_0}$ *definiert auf* $(\Omega, \mathcal{A}, \mathbb{P})$, *bzw.* $(Y_n)_{n\in\mathbb{N}_0}$ *auf* $(\Lambda, \mathcal{B}, \mathbb{Q})$ *auf dem selben Zustandsraum* $(S, \mathcal{S})$*heißen* *äquivalent* *(oder Versionen voneinander), falls ihre Verteilungen gleich sind:*

$$(X.)_*\mathbb{P} = (Y.)_*\mathbb{Q} \qquad \text{auf } (S^{\mathbb{N}_0}, \mathcal{S}^{\otimes \mathbb{N}})$$

Die Markov-Kette $(\pi_n)_{n\in\mathbb{N}_0}$ *definiert auf* $(S^{\mathbb{N}_0}, \mathcal{S}^{\otimes\mathbb{N}}, \mu)$ *heißt kanonische Version. Hierbei ist:*

$$\pi_n : S^{\mathbb{N}_0} \longrightarrow S, x = (x_n)_n \longmapsto x_n$$

Äquivalente Versionen besitzen die selbe Startverteilung μ_0, *sowie die selben Übergangswahrscheinlichkeiten* p_n, $n \in \mathbb{N}$.

Ziel: Konstruktion einer Markov-Kette aus gegebener Startverteilung und Übergangswahrscheinlichkeiten $p_n, n \in \mathbb{N}$.

Definition 2.0.11. *Ein Wahrscheinlichkeitsmaß* μ_n *auf* $\left(S^{1+n}, \mathcal{S}^{1+n}\right)$ *ist gegeben durch:*

$$\mu_n(B) = \int 1_B(x_0, x_1, \ldots, x_n)\mu_0(dx_0)p_1(x_0, dx_1)\cdots p_n(x_{n-1}, dx_n) = \int_{S^{1+n}} 1_B(y,z)\mu_{n-1}(d\mu)p_n(y_n, dz)$$

$$= \int_{S^n} \left[\int_{S^1} 1_B(y,z)p_n(y_{n-1}, dz) \right] \mu_{n-1}(dy) \qquad \Rightarrow \text{rekursive Definition}$$

Also für $f \in \mathcal{L}^0(S^{1+n}, \mathcal{S}^{1+n})$:

$$\int\limits_{S^{1+n}} f d\mu = \int\limits_{S^n} \left[\int\limits_S f(y,z) p_n(y_{n-1}, dz) \right] d\mu_{n-1}(y)$$

Für $B = A \times S$ mit $A \in \mathcal{S}^{1+n}$ gilt: $\mu_{n+1}(B) = \mu_n(A)$, und damit:

$$\mu_{n+k}(A \times S^k) = \mu_n(A) \qquad \forall n,k \in \mathbb{N}, \forall A \in \mathcal{L}^{1+n}$$

Wir definieren ein Wahrscheinlichkeitsmaß $\tilde{\mu_n}$ auf $\tilde{\mathcal{S}^{1+n}}$ durch:

$$\tilde{\mu}_n(A \times S \times S \times \cdots) = \mu_n(A) \qquad (\forall A \in \mathcal{S}^{1+n})$$

Die Maße $\tilde{\mu_n}$ für $n \in \mathbb{N}$ sind <u>konsistent</u>. Für alle $n,k \in \mathbb{N}$ und für alle $B \in \tilde{\mathcal{S}^{1+n}}$ gilt:$\tilde{\mu_n(B)} = \tilde{\mu}_{n+l}$
Daher existiert eine wohldefinierte Fortsetzung μ von $\tilde{\mu_n}$ auf $\tilde{\mathcal{S}}^\infty$ mit

$$\tilde{\mu}(B) := \tilde{\mu_n}(B) \quad \textit{falls } B \in \tilde{\mathcal{S}}^{1+n}$$

Das Maß $\tilde{\mu}$ ist ein <u>Inhalt</u> auf der <u>Algebra</u> $\tilde{\mathcal{S}}^\infty$.

Satz 2.0.12. *(Ionescu Tulcea)*
Zu gegebenem Wahrscheinlichkeitsmaß μ_n und Markov-Kern p_n, $n \in \mathbb{N}$ auf $(S, \mathcal{S})$ existiert genau ein Wahrscheinlichkeitsmaß μ auf $(S^{\mathbb{N}_0}, \mathcal{S}^{\otimes \mathbb{N}_0})$ mit:

$$\mu(B) = \tilde{\mu}(B) \qquad \forall B \in \tilde{\mathcal{S}}^\infty$$

Beweis. **(i)<u>Eindeutigkeit</u>** $\tilde{\mathcal{S}}^\infty$ *ist $\cap$-stabiler Erzeuger von $\mathcal{S}^{\otimes \mathbb{N}_=}$ und μ ist eindeutig festgelegt auf $\tilde{\mathcal{S}}^\infty$*

(ii)<u>Existenz</u> *Es genügt zu zeigen, dass der Inhalt $\tilde{\mu}$ $\emptyset$-stetig auf $\tilde{\mathcal{S}}^\infty$ ist, denn daraus folgt bereits, dass $\tilde{\mu}$ ein Prä-Maß auf $\tilde{\mathcal{S}}^\infty$ ist, für das eine Fortsetzung auf $\sigma(\tilde{\mathcal{S}}^\infty)$ existiert.*

<u>Wiederholung:</u>

(i) $\tilde{\mu}$ ist Prä-Maß auf $\tilde{\mathcal{S}}^\infty$ $\iff \forall A_n \in \tilde{\mathcal{S}}^\infty, \bigcup A_n = A_\infty \in \tilde{\mathcal{S}}^\infty \Rightarrow \tilde{\mu}(\bigcup A_n) = \sum \tilde{\mu}(A_n)$

(ii) $\tilde{\mu}$ ist $\emptyset$-stetig $\iff \forall A_n \in \tilde{\mathcal{S}}^\infty, A_n \supset A_{n-1} \bigcap\limits_n A_n = \emptyset, A_n \searrow \emptyset \Rightarrow \tilde{\mu}(A_n) \searrow 0$

Wir zeigen:

$$\forall B_0 \supset B_1 \supset \ldots \textit{ in } \tilde{\mathcal{S}}^\infty, \textit{ mit } \inf_n \tilde{\mu}(B_n) = \alpha > 0 \Rightarrow \bigcap_n B_n \neq \emptyset$$

Wir können o.B.d.A annehmen, dass für alle $B_n \in \tilde{\mathcal{S}}^{1+n}$ gilt:

$$B_n = A_n \times S \times S \times \ldots \qquad \textit{mit } A_n \in \mathcal{S}^{1+n}$$

Gegeben eine solche Folge von Mengen $(A_n)_{n\in\mathbb{N}}$ definieren wir für alle $n \geq m \in \mathbb{N}$ Funktionen $h_{m,n}$ und h_m auf S^{m+1} durch:

$$h_{m,n}(x_0, \ldots, x_m) := \int\limits_{S^{n-m}} 1_{A_n}(x_0, \ldots, x_m, x_{m+1}, \ldots, x_n) p_{m+1}(x_{m+1}, d_{m+2}) \cdots p_{n-1}(x_{n-1}, dx_n)$$

Sowie : $h_n = \inf\limits_{n \geq m} h_{m,n}$
Aus $A_n \times S \supset A_{n+1}$ folgt:

$$h_{m,n+1}(x_0, \ldots, x_n) = \int\limits_{S^{n+1-m}} 1_{A_{n+1}}(x_0, \ldots, x_{n+1}) p_{m+1}(x_m, dx_{m+1}) \cdots p_n(x_n, dx_{n+1})$$

$$\leq \int 1_{A_m \times S}(x_0, \ldots, x_{n+1}) p_{m+1}(x_m, dx_{m+1}) \cdots p_n(x_n, dx_{n+1})$$

$$= \int\limits_{S^{n-m}} 1_{A_n}(x_0, \dots, x_n p_{m+1}(x_m, dx_{m+1}) \cdots p_n(x_n, dx_{n+1}) = h_{m,n}(x_0, \dots, x_n)$$

Daraus folgt: $h_{m,n} \searrow h_m$ *für* $n \to \infty$. *Daher erhalten wir mit monotoner Konvergenz (und* $d\mu_n = d_{\mu_0} p_1(x_0, dx_1) \cdots p_n(x_{n+1}, dx_n)$*):*

$$\int\limits_{S^{1+n}} h_n d\mu_n = \inf_{n \geq m} \int\limits_{S^{1+m}} h_{m,n} d\mu_n = \inf_{n \geq m} \int\limits_{S^{1+n}} 1_{A_n}(x_0, \dots, x_n) d\mu_n(x_0, \dots, x_n)$$

$$= \inf_{n \geq m} \mu_n(A_n) = \inf_{n \geq m} \tilde{\mu}(B_n) = \alpha > 0 \quad \textit{(nach Voraussetzung)}$$

□

Satz 2.0.13. *(Ionescu Tulcea 2. Version)*
Zu jeder Folge von Markov-Kernen $(p_n)_n$ *auf* $(S, \mathcal{S})$ *und jedem Wahrscheinlichkeitsmaß* μ_0 *auf* $(S, \mathcal{S})$ *existiert genau ein Wahrscheinlichkeitsmaß* μ *auf* $(S^{\mathbb{N}_0}, \mathcal{S}^{\otimes \mathbb{N}_0})$ *sodass für alle* $n \in \mathbb{N}$ *gilt:*

$$\mu(B) = \tilde{\mu}_n(B) \qquad (\forall B \in \tilde{\mathcal{S}}^{1+n}) \tag{2.4}$$

beziehungsweise:

$$\mu(A \times S \times S \times \cdots) = \mu_n(A) = \int\limits_A \mu_0(dx_0) p_1(x_0, dx_1) \cdots p_n(x_{n-1}, x_n) \qquad (\forall A \in \tilde{\mathcal{S}}^{1+n})$$

Insbesondere gilt für $m = 0$*:*

$$\int h_0 d\mu_0 = \alpha \quad \Rightarrow \quad \exists x_0 \in S : h_0(x_0) \geq \alpha$$

Behauptung: Es existiert eine Folge $x_0, x_1, \dots$ *in* S*, sodass für alle* $m \in \mathbb{N}_0$ *gilt:* $h_m(x_0, \dots, x_m) \geq \alpha$

Beweis. Beweis mittels Induktion über m*.*
$m = 0$*: klar*
Betrachte also:

$$\int\limits_S h_{m+1}(x_0, \dots, x_m, y_{m+1}) p_{m+1}(x_m, dy_{m+1}) \overset{\textit{(Fubini)}}{=} \inf_{n \geq m+1} \int\limits_S h_{m+1,n}(x_0, \dots, x_m, y_{m+1}) p_{m+1}(x_m, dy_{m+1})$$

$$= \inf_{n \geq m+1} h_{m,n}(x_0, \dots, x_m) \overset{\textit{mon. fallend}}{=} h_m(x_0, \dots, x_m) \overset{\textit{IA}}{\geq} \alpha$$

Daraus folgt die Existenz eines x_{m+1} *mit* $h(x_0, \dots, x_m, x_{m+1}) \geq \alpha$.
Betrachte also $(x_0, x_1, \dots) =: x \in S^{\mathbb{N}_0}$.
Nach Konstruktion gilt für alle m*:*

$$0 < \alpha \leq h_{m,m}(x_0, \dots, x_m) = 1_{A_m}(x_0, \dots, x_m)$$

Daraus folgt $(x_0, \dots, x_m) \in A_m$ *und daraus:* $x \in B_m$.
Also gilt wie gewünscht:

$$x \in \bigcap_m B_m \Rightarrow \bigcap_m B_m \neq \emptyset$$

□

Korollar 2.0.14. *Für alle* μ_0 *und alle* $(p_n)_n$ *wie im vorherigen Satz existiert eine Markov-Kette mit* μ_0 *als Startverteilung und* p_n *als Übergangswahrscheinlichkeiten.*
Zum Beispiel:

$$\Omega = S^{\mathbb{N}_0}, \quad \mathcal{A} = \mathcal{S}^{\otimes \mathbb{N}_0}, \quad \mathbb{P} = \mu \quad \textit{und } X_n(\omega) = \omega(n) \quad \textit{Projektionsabbildung}$$

Beweis. Zunächst 2.4 für $n = 0$:

$$(X_0)_*\mathbb{P}(A) = \mathbb{P}(X_0 \in A) = \mu_0(A)$$

Für $n \geq 1$ hängt die Übergangswahrscheinlichkeit $(X_0, \dots, X_n)$ nach X_{n+1} nur von X_n ab:

$$(X_0, \dots, X_n, X_{n+1})_*\mathbb{P}(A) = \mathbb{P}((X_0, \dots, X_{n+1}) \in A == \mu_{n+1}(A) = \int\limits_A d\mu_1(x_0, \dots, x_n)p_{n+1}(x_n, dx_n + 1)$$

□

2.1. Markov-Eigenschaft

Betrachte im Folgenden nur noch homogene Markov-Ketten. Diese sind (bis auf Äquivalenz) eindeutig charakterisiert durch die Startverteilung ν und Übergangswahrscheinlichkeit p.
Sei p vorgegeben, dann existiert für alle Wahrscheinlichkeitsmaße ν auf $(S, \mathcal{S})$ ein eindeutiges Wahrscheinlichkeitsmaß $\mathbb{P}_\nu$ auf $(\Omega = S^{\mathbb{N}_0}, \mathcal{A} = \mathcal{S}^{\otimes \mathbb{N}_0})$, sodass die Projektionsabbildungen $(X_n)_{n \in N_0}$ eine homogene Markov-Kette mit Startverteilung ν und Übergangswahrscheinlichkeit p bilden.

Sei $\mathbb{P}_x := \mathbb{P}_{\delta_x}$.

Lemma 2.1.1. *(i) $\mathbb{P}_x(d\omega)$ ist Markov-Kern von $(S, \mathcal{S})$ nach $(\Omega, \mathcal{A})$*

(ii) $\forall A \in \mathcal{A}$: $\quad \mathbb{P}_\nu(A) = \int\limits_S \mathbb{P}_x(A)\nu(dx)$

Beweis. *(i) $\forall n \in \mathbb{N}$, $\forall A \in \tilde{\mathcal{S}}^{1+n}$ (also $A = B \times S \times S \times \cdots$ mit $B \in \mathcal{S}^{1+n}$) ist*

$$x \mapsto \mathbb{P}_x(A) = \mu_n(B) = \int\limits_B p(x, dx_1)p(x_1, dx_2) \cdots p(x_{n-1}, dx_n)$$

messbar. Also betrachten wir nun das System von Mengen

$$\mathcal{D} = \{A \in \mathcal{A} : x \mapsto \mathbb{P}_x(A) messbar\}$$

Daraus folgt:
1) $\bigcup \tilde{\mathcal{S}}^{1+n} \subset \mathcal{D}$
2) $\mathcal{D}$ ist Dynkin-System
Aus 1) & 2) folgt: $\mathcal{D}$ enthält $\sigma(\bigcup\limits_n \tilde{\mathcal{S}}^{n+1}) = \tilde{\mathcal{S}}^\infty$

(ii) klar

□

Lemma 2.1.2. *Für alle $n \in \mathbb{N}$, $f \in \mathcal{L}_+^\infty(S)$ und alle Wahrscheinlichkeitsmaße ν auf $(S, \mathcal{S})$ gilt:*

$$\mathbb{E}_\nu \left[f(X_{k+1}) | \mathcal{A}_k\right] = \mathbb{E}_{X_k}\left[f(X_1)\right] \qquad \mathbb{P}_\nu\text{-fast sicher}$$

Dabei ist:

$$\mathbb{E}_\nu\left[\cdot\right] = \int \cdot \, d\mathbb{P}_\nu \quad \text{und} \quad \mathbb{E}_{X_k}\left[f(X_1)\right] = \nu \cdot X_k \text{ mit } \nu(x) := \mathbb{E}_x\left[f(X_1)\right]$$

Beweis. Betrachte die linke Seite:

$$\mathbb{E}_\nu\left[f(X_{k+1})|A_k\right](\omega) = \mathbb{E}_\nu\left[f(X_{k+1})|X_k\right](\omega) = \int\limits_S f(y)p(X_k(\omega), dy) = (pf)(X_k(\omega))$$

mit $pf(x) := \int\limits_S p(x,dy)f(y)$. *Dafür gilt:*

$$\nu(x) = \int\limits_\Omega f(x_o)d\mathbb{P}_x = \int\limits_{S\times S} f(x_o,x_1)d\mu_1(x_0,x_1) = \int f(x_0,x_1)\delta_x(dx_0)p(x_0,dx_1) = pf(x)$$

Damit gilt für $\mathbb{P}_\nu$*-fast alle* $\omega \in \Omega$*:*

$$\mathbb{E}_\nu\left[f(X_{k+1})|\mathcal{A}_k\right](\omega) = (pf)(X_k(\omega)) = \nu(X_k(\omega)) = \mathbb{E}_{X_k}\left[f(X_1)\right] = (\textit{rechte Seite})$$

□

Bemerkung 2.1.3. *Die linke Seite ist die bedingte Erwartung, also eine Äquivalenzklasse von Zufallsvariablen. Die rechte Seite ist der Erwartungswert bezüglich eines zufällig gewählten Wahrscheinlichkeitsmaßes.*

Satz 2.1.4. *In obigem Setting gilt die schwache Markov-Eigenschaft, das heißt:*

$$\mathbb{E}_\nu\left[f(X_{n+k})_{n\in\mathbb{N}_0}|\mathcal{A}_n\right] = \mathbb{E}_{X_k}\left[f(X_n)_{n\in\mathbb{N}}\right] \qquad \forall k\in\mathbb{N}, \forall f\in\mathcal{L}^0_+(\Omega,\mathcal{A}) \tag{2.5}$$

Mit anderen Worten: Die bedingte Verteilung der Folge $(X_{k+n})_{n\in\mathbb{N}_0}$ *bedingt auf* $\mathcal{A}_k$*, stimmt mit der Verteilung von* $(X_n)_{n\in N_0}$ *gestartet in* X_k *überein.*

Beweis. *Sei o.B.d.A:*

$$f = 1_A \quad \textit{mit} \quad A\in\mathcal{A} \quad \textit{und} \quad A\in\bigcup_m \tilde{\mathcal{S}}^{1+m}.$$

und: $A = B_0\times\cdots\times B_m\times S\times S\times\cdots$ *mit* $B_i\in\mathcal{S}$
Wir benutzen Induktion nach m:
m=1: *vorheriges Lemma*
$m\to m+1$
Betrachte $\mathbb{E}_\nu\left[f(X_{\cdot+k})|\mathcal{A}_k\right] = \mathbb{P}(X_k\in B_0, X_{k+1}\in B_1,\ldots,X_{k+m}\in B_m|\mathcal{A}_k)$

$$= \mathbb{E}_\nu\left[\mathbb{E}\left[1_{B_0}(X_k)\cdots 1_{B_m}(X_{k+m})|\mathcal{A}_{k+m-1}\right]|\mathcal{A}_k\right] \overset{1_{B_i}(X_{k+i})mb.}{=} \mathbb{E}_\nu\left[1_{B_0}(X_k)\cdots 1_{B_{m-1}}(X_{k+m-1})\mathbb{E}\left[1_{B_m}(X_{m+k}|\mathcal{A}_{m+k}\right]|A_k\right]$$

$$\overset{2.1.2}{=} \mathbb{E}\left[1_{B_0}(X_k)\cdots 1_{B_{m-1}}(X_{k+m-1}]\right)\mathbb{E}_{X_{k+m-1}}\left[1_{B_m}(X_1)|\mathcal{A}_k\right] \overset{(IA)}{=} \mathbb{E}_{X_k}\left[1_{B_0}(X_0)\cdots 1_{B_{m-1}}(X_{m-1})\mathbb{E}_{X_{m-1}}\left[1_{B_m}(X_1)\right]\right]$$

$$\overset{2.1.2}{=} \mathbb{E}_{X_k}\left[1_{B_0}(X_0)\cdots 1_{B_{m-1}}(X_{m-1})\mathbb{E}_{X_k}\left[1_{B_m}(X_m)|\mathcal{A}_{m-1}\right]\right] = \mathbb{E}_{X_k}\left[\mathbb{E}_{X_k}\left[1_{B_0}(X_0)\cdots 1_{B_{m-1}}(X_{m-1})1_{B_m}(X_m)|\mathcal{A}_{m-1}\right]\right]$$

$$= \mathbb{E}_{X_k}\left[1_{B_0}(X_0)\cdots 1_{B_m}(X_m)\right] = \mathbb{E}_{X_k}\left[f(X_1)\right]$$

□

Satz 2.1.5. *Für alle Markov-Kerne* $p: S\times\mathcal{S}\to\mathbb{R}$ *existiert ein Markov-Kern* $\mathbb{P}: S\times\mathcal{A}\to\mathbb{R}$*,* $(x,A)\mapsto\mathbb{P}_x(A)$ *(mit* $\Omega = S^{\mathbb{N}_0}, \mathcal{A} = \mathcal{S}^{\otimes\mathbb{N}_0}$*).*

(i) Für alle Wahrscheinlichkeitsmaße ν *auf* $(S,\mathcal{S})$ *bilden die Projektionen* $X_n(\omega) = \omega(n)$ *eine homogene Markov-Kette auf* $(\Omega,\mathcal{A},\mathbb{P})$ *mit Startverteilung* ν *und Übergangswahrscheinlichkeit* $P_\nu(A)$ *gegeben durch:*

$$\mathbb{P}_\nu(A) := \int\limits_S \mathbb{P}_x(A)\nu(dx) \qquad (\forall A\in\mathcal{A})$$

(ii) Für alle $k\in\mathbb{N}$*, alle Verteilungen* ν *und alle* $f\in\mathcal{L}^0_+(\Omega,\mathbb{R})$ *gilt die Markov-Eigenschaft:*

$$\mathbb{E}_\nu\left[f(X_{n+k})_{n\in\mathbb{N}_0}|\mathcal{A}_k\right](\omega) = \mathbb{E}_{X_k(\omega)}\left[f(X_n)_{n\in\mathbb{N}_0}\right] \quad \textit{für } \mathbb{P}_\nu\textit{-fast alle } \omega\in\Omega$$

Definiert man Shift-Abbildungen

$$\theta_k : S^{\mathbb{N}_0} \longrightarrow S^{\mathbb{N}_0} \qquad x = (x_n)_{n\in\mathbb{N}_0} \longmapsto \theta_k(x) = (x_{k+n})_{n\in\mathbb{N}_0}$$

so lässt sich die Markov-Eigenschaft auch äquivalent wie folgt formulieren:

$$\boxed{\mathbb{E}_\nu\left[f\circ\theta_k|\mathcal{A}_k\right] = \mathbb{E}_{X_k}\left[f\right] \quad \mathbb{P}_\nu\textit{-fast überall}} \qquad \forall k\in\mathbb{N}_0, \forall\nu, \forall f\in\mathcal{L}^0_+(\Omega) \tag{2.6}$$

Beachte: *Die Abbildung* $(X_n)_{n\in\mathbb{N}_0} : \Omega \longrightarrow S^{\mathbb{N}_0}$*,* $\omega\longmapsto (X_n(\omega))_{n\in\mathbb{N}_0} = (\omega(n))_{n\in\mathbb{N}_0} = \omega$ *ist im kanonischen Modell die Identität.*

Definition 2.1.6. *Eine <u>Markovsche Familie</u> besteht aus:*

- *einem Messraum* $(\Omega, \mathcal{A})$ *und einem Standard-Borel-Raum* $(S, \mathcal{S})$
- *Wahrscheinlichkeitsmaßen* $\mathbb{P}_x$ *auf* $(\Omega, \mathcal{A})$ *für jedes* $x \in S$ *mit*

$$(x, A) \longmapsto \mathbb{P}_x(A) \qquad \text{ist ein Markov-Kern}$$

- *Die Shift Abbildungen* $\theta_k : \Omega \longmapsto \Omega$ *sind messbar mit* $\theta_{k+l} = \theta_k \circ \theta_l \quad \forall n, k \in \mathbb{N}_0$ *(Halbgruppe)*
- *Die Auswertungsabbildung* $X_n : \Omega \longrightarrow S$ *ist messbar mit:*

$$X_{n+k} = X_n \circ \theta_k \qquad \text{für alle } n, k \in \mathbb{N}_0 \text{ für die 2.6 gilt}$$

Hierbei ist $\mathcal{A}_k = \sigma(X_0, \ldots, X_n)$.
Man kann die Bedingung auch allgemeiner fassen. Dann gehört zu der Markovschen Familie auch die <u>Filtrierte Familie</u> von $(\mathcal{A}_n)_{n \in \mathbb{N}_0}$ *von* σ*-Algebren* $\mathcal{A}_n \subset \mathcal{A}$ *auf* Ω *mit:* $\mathcal{A}_{k+1} \supset \mathcal{A}_k \supset \sigma(X_0, \ldots, X_n)$ *(äquivalent zu:* X_n *ist* $\mathcal{A}_k$*-messbar* $\forall n \leq k$*).*

<u>Ziel: Auflösung der Asymmetrie zwischen Vergangenheit und Zukunft</u>

Proposition 2.1.7. *Für jede Markovsche Familie* $((X_n)_{n \in \mathbb{N}_0}, (\theta_n)_{n \in \mathbb{N}_0}, (\mathbb{P}_x)_x)$ *und jedes Wahrscheinlichkeitsmaß auf dem Zustandsraum S gilt:*
$(X_n)_{n \in \mathbb{N}_0}$ *bildet eine homogene Markov-Kette auf* $(\Omega, \mathcal{A}, \mathbb{P}_\nu)$ *mit Startverteilung* ν.

Beweis. *(i)* $(X_n)_{n \in \mathbb{N}_0}$ *ist eine Markov-Kette, da* $\mathbb{E}[\cdot|\mathcal{A}_n]$ *zu etwas führt, was* X_n*-messbar ist. Daher gilt:*

$$\mathbb{E}_\nu \left[1_B(X_{n+1}|\mathcal{A}_n\right] = \mathbb{E}_\nu \left[\underbrace{1_B(X_{n+1})}_{= f \circ \theta_n, f = 1_B X_1} |X_n \right]$$

(ii) $(X_n)_{n \in \mathbb{N}_0}$ *ist homogen. Wähle dazu* $f = 1_B(X_1)$ *mit* $B \in \mathcal{S}$*:*

$$\mathbb{P}_\nu(X_{n+1} \in \mathbb{N}|X_n) = \mathbb{P}_\nu(X_{n+1} \in B|X_0, \ldots X_n) = \mathbb{E}_\nu \left[1_B(\underbrace{X_1 \circ \theta_n}_{= X_{n+1}} |X_0, \ldots, X_n) \right] \stackrel{ME}{=} \mathbb{E}_{X_n}\left[1_B(X_1)\right] = p_1(X_n, B)$$

□

Definition 2.1.8. *Gegeben sei ein Wahrscheinlichkeitsraum* $(\Omega, \mathcal{A}, \mathbb{P})$, *eine Folge von messbarer Abbildungen* $(X_n)_{n \in N_0}$ *und ein Standard-Borel-Raum* $(S, \mathcal{S})$. *Wir definieren:*

1. $\mathcal{A}_k = \sigma(X_0, \ldots, X_k)$ *k-Vergangenheit*
2. $\mathcal{B}_k = \sigma(X_n, n \geq k)$ *k-Zukunft*
3. $\sigma(X_k)$ *k-Gegenwart*

Satz 2.1.9. *Gegeben sei eine Folge messbarer Abbildungen* $(X_n)_{n \in N_0}$ *mit Werten in einem Standard-Borel-Raum* $(S, \mathcal{S})$ *definiert auf einem Wahrscheinlichkeitsraum* $(\Omega, \mathcal{A}, \mathbb{P})$.
Dann sind äquivalent:

(i) $(X_n)_{n \in N_0}$ *ist eine Markov-Kette, das heißt:*

$$\mathbb{P}(X_{n+1} \in C|\mathcal{A}_n) = \mathbb{P}(X_{n+1} \in C|X_n) \qquad \forall n \in \mathbb{N}_0, \forall C \in \mathcal{S}$$

(ii)

$$\mathbb{P}(B|\mathcal{A}_n) = \mathbb{P}(B|X_n) \qquad \forall n \in \mathbb{N}_0, \forall B \in \mathcal{B}_n$$

(iii) Für alle $k \in \mathbb{N}_0$ sind die σ-Algebren $\mathcal{A}_n$ und $\mathcal{B}_n$ bedingt unabhängig bezüglich $\sigma(X_n)$:

$$\mathbb{P}(A \cap B|X_n) = \mathbb{P}(A|X_n)\mathbb{P}(B|X_n) \quad \mathbb{P}-\textit{fast sicher} \qquad \forall A \in \mathcal{A}_n, \forall B \in \mathcal{B}_n$$

Mit anderen Worten: „Vergangenheit und Zukunft hängen voneinander nur mittels der Gegenwart ab“

Beweis. *(i) ⇒ (ii) Sei $B \in \mathcal{B}_k$. Wir können o.B.d.A schreiben: $B = \{X_k \in C_k\} \cap \{X_{k+1} \in C_{k+1}\} \cap \cdots \cap \{X_{k+l} \in C_{k+l}\}$ mit $C_i \in \mathcal{S}$*
Zur Vereinfachung nehmen wir $l = 2$ an (ansonsten Induktion nach l), und erhalten:

$$\mathbb{P}(X_k \in C_k, X_{k+1} \in C_{k+1}, X_{k+2} \in C_{k+2}|\mathcal{A}_k) = 1_{C_k}(X_k)\mathbb{E}\left[1_{C_k+1}(X_{k+1})1_{C_{k+2}}(X_{k+2})|\mathcal{A}_k\right]$$

$$= 1_{C_k}(X_k)\mathbb{E}\left[\mathbb{E}\left[1_{C_k+1}(X_{k+1})1_{C_{k+2}}(X_{k+2})|\mathcal{A}_{k+1}\right]|\mathcal{A}_k\right] = 1_{C_k}(X_k)\mathbb{E}\left[1_{C_k+1}(X_{k+1})\mathbb{E}\left[1_{C_{k+2}}(X_{k+2})|\mathcal{A}_{k+1}\right]|\mathcal{A}_k\right]$$

$$\overset{(i)\text{ für } k+1}{=} 1_{C_k}(X_k)\mathbb{E}\left[\underbrace{1_{C_k+1}(X_{k+1})\mathbb{E}\left[1_{C_{k+2}}(X_{k+2})|X_{k+1}\right]}_{=f(X_{k+1})}|\mathcal{A}_k\right]$$

$$\overset{(i)\text{ für } k}{=} 1_{C_k}(X_k)\mathbb{E}\left[1_{C_k+1}(X_{k+1})\mathbb{E}\left[1_{C_{k+2}}(X_{k+2})|X_{k+1}\right]|X_k\right]$$

$$\overset{(i)\text{ für } k+1}{=} 1_{C_k}(X_k)\mathbb{E}\left[\mathbb{E}\left[1_{C_k+1}(X_{k+1})1_{C_{k+2}}(X_{k+2})|\mathcal{A}_{k+1}\right]|X_k\right]$$

$$= 1_{C_k}(X_k)\mathbb{E}\left[1_{C_{k+1}}(X_{k+1})1_{C_{k+2}}(X_{k+2})|X_k\right] = \mathbb{P}(X_k \in C_k, X_{k+1} \in C_{k+1}, X_{k+2} \in C_{k+2}|X_k)$$

(ii) ⇒ (i) trivial (wähle $B = X_{k+1}^{-1}(C)$)

(ii) ⇒ (iii) Seien A,B wie gefordert. Dann gilt:

$$\mathbb{P}(A \cap B|X_k) = \mathbb{E}\left[\mathbb{E}\left[1_A 1_B|\mathcal{A}_k\right]|X_k\right] = \mathbb{E}\left[1_A \underbrace{\mathbb{E}\left[1_B|\mathcal{A}_k\right]}_{\overset{(ii)}{=}\mathbb{E}[1_B|X_k]}|X_k\right]$$

$$\overset{(ii)}{=} \mathbb{E}\left[1_A\mathbb{E}\left[1_B|X_k\right]|X_k\right] = \mathbb{E}\left[1_B|X_k\right]\mathbb{E}\left[1_A|X_k\right]$$

(iii) ⇒ (ii) Sei $B \in \mathcal{B}_k$. Zu zeigen:

$$\int_A 1_B d\mathbb{P} = \int_A \mathbb{P}(B|X_k)d\mathbb{P} \qquad \forall A \in \mathcal{A}_n$$

Für die rechte Seite gilt:

$$\mathbb{E}\left[\mathbb{E}\left[1_A\mathbb{E}\left[1_B|X_k\right]|X_k\right]\right] = \mathbb{E}\left[\mathbb{E}\left[1_A|X_k\right]\mathbb{E}\left[1_B|X_k\right]\right] = \mathbb{E}\left[\mathbb{E}\left[1_A 1_B|X_k\right]\right] = \mathbb{E}\left[1_A 1_B\right] = \textit{ linke Seite}$$

□

Beispiel 2.1.10. *Gegeben sei ein Graph (V, E) und ein Potential $H : V \longrightarrow \mathbb{R}$. Ziel ist es, auf V eine Irrfahrt zu definieren, die das Minimum von H findet (bei einem deterministischen Gradientenverfahren ergeben sich Probleme durch die Nebenminima)*
Wir definieren die freie Irrfahrt durch:

$$p(x, y) = \begin{cases} \frac{1}{deg(x)} & y \sim x \\ 0 & sonst \end{cases} \qquad mit\ deg(x) := \sharp\{y : y \sim x\} \quad x \sim y \iff (y, x) \in E$$

Wir definieren die Gewichtete Irrfahrt auf V mit Inverser Temperatur $\beta > 0$ durch:

$$p^{\beta}(x,y) = \begin{cases} \frac{1}{deg(x)} exp\left(-\beta(H(y)-H(x))_+\right) & y \sim x \\ 1 - \sum\limits_{y, y\sim x} p^{\beta}(x,y) & y = x \\ 0 & sonst \end{cases}$$

Dies führt zu einer homogenen Markov-Kette(für jedes festgelegt β)(Metropolis Algorithmus) Wähle $\beta = \beta_n$ abhängig von n mit $\beta_n \nearrow \infty$. Das Ergebnis ist eine inhomogene Markov-Kette (simulated annealing)

Beispiel 2.1.11. *Sei $S = \{-1,1\}^V$ ein Zustandsraum mit endlicher Menge V, z.B. $V = (\mathbb{Z}/_{n\mathbb{Z}})^d$ (diskreter d-dimensionaler Torus).*
Definiere eine Nachbarschaftsrelation auf S:

$$x = (x_i)_{i\in V} \sim y = (y_i)_{i \in V} \qquad \Longleftrightarrow \qquad \exists j \in V : \; x_i = \begin{cases} y_i & i \neq j \\ -y_j & i = j \end{cases}$$

Definiere die Übergangswahrscheinlichkeit p auf $S = \{-1,1\}^V$ durch:

$$p(x,y) = \begin{cases} \frac{1}{|v|} & x \sim y \\ 0 & sonst \end{cases}$$

Mit anderen Worten:

$$p(x, x^{(j)}) = \frac{1}{|v|} \qquad \text{für alle } j \in V \text{ mit } x^{(j)} \in S \text{ und } x_i^j = \begin{cases} x_i & i \neq j \\ -x_i & i = j \end{cases}$$

Dies ist eine freie Irrfahrt.
Beachte: für $V = \{0,1,\ldots,N-1\}^d$ mit $d = 2$, $N = 1000$ ist $|S| = 2^{1000\times 1000} \approx 1,27 \cdot 10^{3000000}$.

Beispiel 2.1.12. *(Ising-Modell)*
Sei (V,E) ein Graph.

$$H(x) = \sum_{\substack{k,l\in V \\ k\sim l}} |x_k - x_l|^2 = |V| - \sum_{\substack{k,l\in V \\ k\sim l}} x_k x_l$$

Für $y = x^{(i)}$ gilt:

$$H(x^{(i)}) - H(x) = \sum_{\substack{k,l\in V \\ k\sim l}} x_k x_j$$

$$p^{\beta}(x,y) = \begin{cases} \frac{1}{|v|} \exp(-4\beta \left[\sum_{k\sim l}^{k,l\in V} x_k x_j\right]_*) & \text{falls } y = x^j \text{ für ein } j \in V \\ 1 - \sum_{j\in V} p^{\beta}(x, y^{(j)}) & y = x \\ 0 & sonst \end{cases}$$

Dies ist eine homogene Markov-Kette auf S, dem Raum aller Konfigurationen.

2.2. Stoppzeiten und starke Markov-Eigenschaft

Erinnerung: schwache Markov-Eigenschaft:

$$\mathbb{E}_{\nu}\left[Z \circ \theta_k | \mathcal{A}_k\right] = \mathbb{E}_{X_k}\left[Z\right] \qquad \forall k \in \mathbb{N}, \forall Z \text{ ZV auf } \Omega$$

Ziel: Markov-Eigenschaft mit zufälligen $k = k(\omega)$. Dies ist im Allgemeinen falsch, zum Beispiel für $k(\omega) =$ {letzter Zeitpunkt an dem ... > 0 war}

Definition 2.2.1. *Sei $(\Omega, \mathcal{A})$ ein Messraum, $(\mathcal{A}_n)_{n\in\mathbb{N}_0}$ eine Filtrierung (d.h. $\forall n$: $\mathcal{A}_n$ ist σ-Algebra auf Ω, $\mathcal{A}_n \subset \mathcal{A}_{n+1} \subset \cdots \subset \mathcal{A}$). Mit dieser Filtrierung messen wir den Zuwachs an Informationen. Typischerweise: $\mathcal{A}_n = \sigma(Y_n : k \leq n)$ mit messbaren $Y_k : \Omega \longrightarrow S$
Eine Abbildung $\tau : \Omega \longrightarrow \mathbb{N}_0 \cup \{\infty\}$ heißt Stoppzeit , falls gilt.:*

$$\{\tau = n\} \in \mathcal{A}_n \qquad \forall n \in \mathbb{N}_0 \tag{2.7}$$

*Typische Stoppzeiten: erstes Eintreten von ... , keine Stoppzeiten: letztes Eintreten von ...
Mit 2.7 gilt auch:*

$$\{\tau \leq n\} \in \mathcal{A}_n \qquad \forall n \in \mathbb{N}_0 \tag{2.8}$$

2.7 gilt genau dann, wenn auch 2.8 gilt, denn:

$$\{\tau \leq n\} = \bigcup_k \underbrace{\{\tau = k\}}_{\in\mathcal{A}_n\subset\mathcal{A}_m} \quad \in \mathcal{A}_n$$

$$\{\tau = n\} = \underbrace{\{\tau \leq n\}}_{\mathcal{A}_n} \setminus \underbrace{\{\tau \leq n-1\}}_{\in\mathcal{A}_{n-1}} \quad \in \mathcal{A}_n$$

Beispiel 2.2.2. $\mathcal{A}_n = \sigma\left(Y_k : k \leq n\right), B \in \mathcal{S}$

- *„erste Trefferzeit von B“: $\tau := \inf\{n \in \mathbb{N}_0 : Y_n \in B\}$ mit $(\inf\emptyset = +\infty)$ ist eine Stoppzeit. Denn:*

$$\{\tau = k\} = \underbrace{\{X_k \in B\}}_{\in\mathcal{A}_k} \cap \bigcap_{n\leq k-1} \underbrace{\{X_n \notin B\}}_{\mathcal{A}_n\subset\mathcal{A}_k} \quad \in \mathcal{A}_k$$

- *Mit σ und τ ist auch $\sigma \wedge \tau := \min\{\sigma, \tau\}$ eine Stoppzeit*
- *Jede konstante Zufallsvariable τ ist eine Stoppzeit.(„Fixe Zeiten sind Stoppzeiten“)*

Definition 2.2.3. *Für eine Stoppzeit τ auf $(\Omega, \mathcal{A}, (\mathcal{A}_n)_{n\in\mathbb{N}_0})$ definieren wir die σ-Algebra der τ-Vergangenheit:*

$$\mathcal{A}_\tau = \{A \in \mathcal{A} : \forall n \in \mathbb{N}_0 : A \cap \{\tau = n\} \in \mathcal{A}_n\} = \bigcap_{n\in\mathbb{N}_0} \{A : A \cap \{\tau = n\}\} \qquad \text{(„Spuren von } \sigma\text{-Algebren“)}$$

Beachte: Für $\tau = k$ gilt: $\mathcal{A}_\tau = \mathcal{A}_k$

Sei $(X_n)_{n\in N_0}$ eine Folge von Abbildungen $X_n : \Omega \longrightarrow S$ jeweils $\mathcal{A}_k$-messbar. Auf $\{\tau < \infty\} \subset \Omega$ definieren wir

$$X_\tau : \{\tau < \infty\} \longrightarrow S, \qquad \omega \mapsto X_{\tau(\omega)}(\omega) =: X_\tau(\omega)$$

als Komposition der Abbildungen $(\omega, k) \longmapsto X_k(\omega), \omega \longmapsto (\tau(\omega), \omega)$ Daraus folgt:

$$X_\tau : \{\tau < \infty\} \longrightarrow S \text{ ist messbar bezüglich } \mathcal{A}_\tau (\text{insbesondere bezüglich } \mathcal{A})$$

Denn für alle $B \in \mathcal{S}$ gilt:

$$\{X_\tau \in B\} \cap \{\tau < \infty\} = \bigcup_{n\in\mathbb{N}_0} \{X_n \in B\} \cap \{\tau = n\} \qquad \in \mathcal{A}_n$$

$$\{X_\tau \in B\} \cap \{\tau = n\} = \underbrace{\{X_n \in B\}}_{\in\mathcal{A}_n} \cap \underbrace{\{\tau = n\}}_{\in\mathcal{A}_n} \qquad \in \mathcal{A}_n$$

$\Rightarrow \{X_\tau \in B\} \in \mathcal{A}_\tau$

Satz 2.2.4. *Gegeben sei eine Markovsche Familie $(\mathbb{P}_x)_{x\in S}$, $(X_n)_{n\in N_0}$ im kanonischen Modell $\Omega = S^{\mathbb{N}_0}$,$\mathcal{A} = \mathcal{S}^{\otimes n}$. Dann gilt für alle Wahrscheinlichkeitsmaße ν auf S, alle Stoppzeiten τ und alle Zufallsvariablen $Z : \Omega \longmapsto \mathbb{R}_+$ die starke Markov-Eigenschaft:*

$$\mathbb{E}_\nu\left[Z \circ \theta_\tau | \mathcal{A}_\tau\right] = \mathbb{E}_{X_\tau}\left[Z\right] \quad \mathbb{P}_\nu\text{-fast sicher auf } \{\tau < \infty\} \tag{2.9}$$

Beweis. *Zu zeigen:* $\int\limits_{A\cap\{\tau<\infty\}} Z\circ\theta_\tau d\mathbb{P} = \int\limits_{A\cap\{\tau<\infty\}} \mathbb{E}_x[Z]\,d\mathbb{P} \qquad \forall A\in\mathcal{A}_\tau$

$$LS = \sum_{k=0}^{\infty} \int\limits_{A\cap\{\tau=k\}} Z\circ\theta_\tau d\mathbb{P} = \sum_{k=0}^{\infty} \int\limits_{A\cap\{\tau=k\}\,\in\mathcal{A}_k} Z\circ\theta_k d\mathbb{P}$$

$$\overset{sw.\ ME}{=} \sum_{k=0}^{\infty} \int\limits_{A\cap\{\tau=k\}} \mathbb{E}_{X_k}[Z]\,d\mathbb{P} = \int\limits_{A\cap\{\tau<\infty\}} \mathbb{E}_{X_\tau}[Z]\,d\mathbb{P} = RS$$

□

Definition 2.2.5. *Gegeben sei ein Markov-Kern p auf dem Zustandsraum $(S,\mathcal{S})$ und eine endliche Stoppzeit T (im kanonischen Modell). Definiere den Markov-Kern p^T auf $(S,\mathcal{S})$ durch:*

$$p^T(x,A) = \mathbb{P}_x(X_T\in\mathcal{A})$$

Falls T nicht notwendig endlich ist, dann definiere einen Sub-Markov-Kern ($p^T(x,S)\le 1$):

$$p^T(x,A) = \mathbb{P}_x(X_T\in\mathcal{A}, T<\infty)$$

Mit anderen Worten: $p^T(x,\cdot)$ ist die Verteilung von X_T unter dem Maß $\mathbb{P}_x$.

Proposition 2.2.6. *Für alle endlichen Stoppzeiten S,T gilt:*

$$p^S\circ p^T = p^{S+T\circ\theta_S}$$

Beweis. *Übungen.* □

Beispiel 2.2.7.

- *Ist $T=k$ eine konstante Stoppzeit, dann ist p^T das k-fache Produkt von p:*

$$p^k(x,A) = \mathbb{P}_x(X_k\in A)$$

- *Speziell: $B\subset S$. Teilmenge, $T=T_B=\inf\limits_n\{n\in\mathbb{N}: X_n\in B\}$ „erste Trefferzeit von B“*

$$p_b(x,A) = p^{T_B}(x,B) = \mathbb{P}_x(X_{T_B}\in A, T_B<\infty) \qquad \textit{(Sub-Markov-Kern)}$$

2.3. Das Reflektionsprinzip

Ziel: Vergleich zwischen dem Maximalprozess $\sup\limits_{m\le n} X_n$ und X_n für $X_n = \sum\limits_{i=1}^{n}\xi_i$mit iid ξ_i.

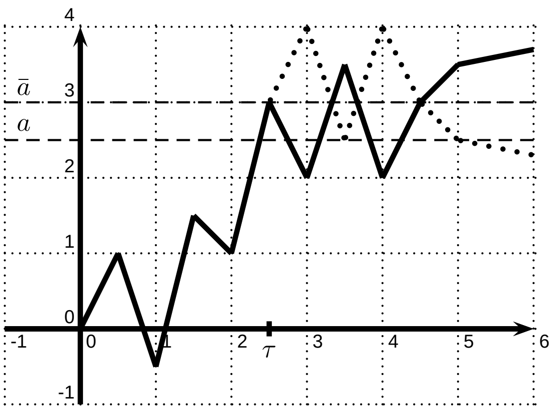
Satz 2.3.1. *Gegeben seien Zufallsvariablen $(\xi_i)_{i\in\mathbb{N}}$ mit symmetrischer Verteilung auf $\mathbb{R}$ und $X_n := \sum\limits_{i\le n}\xi_i$. Dann gilt:*

$$\mathbb{P}(\sup_{m\le n} X_n > a) \le 2\mathbb{P}(X_n > a)$$

Beweis. 1. *Intuition mittels Bild*

τ = 1. Zeitpunkt mit $X_n > a$ (Stoppzeit)

$\bar{a} = x_\tau$



2. *formal*
Sei τ eine Stoppzeit mit $\tau = \inf\{m \in \{0,1,\dots,n\} : X_n > a\}$ und $\inf \emptyset = \infty$. Dann gilt $\sup\limits_{m \leq n} X_n > a$ genau dann, wenn $\tau \leq n$. Verwende im Folgenden die Tatsache, dass $(X_n)_{n \leq 0}$ eine Markov-Kette ist. Definiere die Zufallsvariable Y_m für $m \in \mathbb{N}_0$ durch $Y_m(n) = 1_{\{m \leq n\}} 1_{\{X_{n-m} > a\}}$. Dann gilt für alle $y > a$ und alle $m \leq n$:

$$\frac{1}{2} \leq \mathbb{P}_y(X_{n-m} \geq y) \leq \mathbb{P}_y(X_{n-m} > a) = \mathbb{E}_y\left[Y_m\right]$$

Verwende nun:

$$Y_\tau \circ \theta_\tau = \begin{cases} 1 & \text{falls } X_m > a \text{ (und damit } \tau \leq n) \\ 0 & \text{sonst} \end{cases}$$

Da $X_m > 0$ auf $\{\tau = m\}$folgt mit starker Markov-Eigenschaft:

$$\mathbb{E}_0\left[Y_t \circ \theta_\tau | \mathcal{A}_\tau\right] = \mathbb{E}_{X_\tau}\left[Y_\tau\right] = \sum_{m=0}^{n} 1_{\{\tau = m\}} \mathbb{E}_{X_m}\left[Y_m\right] \geq \frac{1}{2} \quad \text{auf } \{\tau < \infty\} = \{\tau \leq n\}$$

Integration über die Menge $\{\tau \leq n\}$ liefert:

$$\frac{1}{2}\mathbb{P}(\tau \leq n) \leq \mathbb{E}_0\left[\mathbb{E}_0\left[Y_\tau \circ \theta_\tau | \mathcal{A}_\tau\right] \underbrace{1_{\{\tau \leq n\}}}_{\text{mb. bzgl. } \mathcal{A}_\tau}\right] = \mathbb{E}_0\left[\mathbb{E}_0\left[Y_\tau \circ \theta_\tau 1_{\{\tau \leq n\}} | \mathcal{A}_\tau\right]\right]$$

$$= \mathbb{E}_0\left[Y_\tau \circ \theta_\tau 1_{\{\tau \leq n\}}\right] = \mathbb{E}\left[1_{\{X_n > a\}} 1_{\{\tau \leq n\}}\right] = \mathbb{P}(X_n > a) \quad \text{denn } (X_n > a \Rightarrow \tau \leq n)$$

Daraus folgt bereits die Behauptung:

$$\frac{1}{2}\mathbb{P}\left((\max_{m \leq n} X_m > a)\right) = \frac{1}{2}\mathbb{P}(\tau \leq n) \leq \mathbb{P}(X_n > a)$$

□

2.4. Rekurrenz und Transienz

Frage: Kehrt ein Markov-Prozess mit positiver Wahrscheinlichkeit (oder Wahrscheinlichkeit 1) im Laufe der Zeit wieder zum Startpunkt zurück?
Hierfür gehen wir im Folgenden von einem abzählbaren Zustandsraum S, mit $\mathcal{S} = 2^S$ aus. Dann entsprechen die Markov-Kerne $p(x, dy)$ genau den stochastischen Matrizen $(p(x,y))_{x,y \in S}$ mit abzählbar vielen Einträgen.In diesem Setting gilt für alle $n \in \mathbb{N}$:

$$p^n(x,y) = \sum_{s \in S} p^{k-1}(x,z)p(z,y), \qquad T_x^0 = 0$$

Definition 2.4.1. *Definiere für $x \in S$ und $k \in \mathbb{N}$ die k-te Eintrittszeit T_x^k von x durch:*

$$T_x^k = \inf\left\{n > T_x^{k-1} X_n = x\right\}$$

Insbesondere die erste Eintrittszeit:

$$T_x := T_x^1 = \inf\{n \geq 1 X_n = x\}$$

Es gilt im kanonischen Modell: $T_x^k = T_x^1 \circ \theta_{T_x^{k-1}} + T_x^{k-1}$. Ferner definieren wir:

$$F(x,y) = \mathbb{P}_x(\exists n \in \mathbb{N} : X_n = y) = \mathbb{P}_x(T_y^1 < \infty)$$

Im Fall $x = y$ heißt dies Rückkehrwahrscheinlichkeit nach x

Proposition 2.4.2. *Für alle Zustände $x, y \in S$ und alle $k \in \mathbb{N}$ gilt:*

$$\mathbb{P}_x(T_y^k < \infty) = F(x,y)F(y,y)^{k-1}$$

<u>*intuitiv:*</u> *Wenn man, von x ausgehend, y k-mal treffen will, muss man zunächst y treffen und dann (k-1)-mal, von y ausgehend, y treffen.*

Beweis. *Induktion nach k:*
Die Aussage ist für $k = 1$ klar.
Sei also $k \leq 2$ und die Behauptung für $k = 1$ gegeben. Mit starker Markov-Eigenschaft folgt:

$$\mathbb{E}_x\left[1_{\{T_y^k<\infty\}}\right] = \mathbb{E}_x\left[\mathbb{E}_x\left[1_{\{T_y^k<\infty\}}|\mathcal{A}_{T_{k-1}}\right]1_{\{T_y^{k-1}<\infty\}}\right]$$

$$= \mathbb{E}_x\left[\mathbb{E}_x\left[1_{\left\{T_y^1\circ\theta_{T_y^{k-1}}<\infty\right\}}|\mathcal{A}_{T_y^{k-1}}\right]1_{T_y^{k-1}<\infty}\right] = \mathbb{E}_x\left[\mathbb{E}_{X_{T_{k-1}}=y}\left[1_{\{T_y^1<\infty\}}\right]1_{\{T_y^{k-1}<\infty\}}\right]$$

$$= \mathbb{E}_x\left[\underbrace{\mathbb{E}_y\left[1_{\{T_y<\infty\}}\right]}_{=F(y,y)}1_{\{T_y^{k-1}<\infty\}}\right] = F(y,y)\mathbb{P}_x(T_y^{k-1}<\infty) \overset{IA}{=} F(y,y)F(x,y)F(y,y)^{k-2} = F(x,y)F(y,y)^{k-1}$$

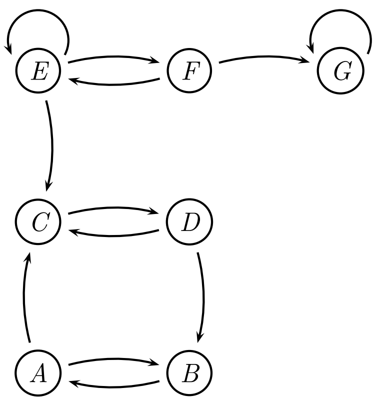
□

Definition 2.4.3. *Ein Zustand $x \in S$ heißt:*

- *absorbierend, falls $p(x,x) = 1$*
- *rekurrent, falls $F(x,x) = 1$*
- *transient, falls $F(x,x) < 1$*

Es gilt stets: x absorbierend $\Rightarrow$ x rekurrent

Beispiel 2.4.4. *(i)*



absorbierdend: G *transient: E, F* *rekurrent: A, B, C, D*

(ii) <u>*Galton-Watson-Verzweigungsprozess*</u> *mit Nachwuchsverteilung q auf $\mathbb{N}_0$, $g(\{0\}) > 0$ und $p(x,y) = (q^{*x})(y)$. Dann gilt für alle $k \in \mathbb{N}$:*

$$p(0,k) = 0, \quad p(k,0) > 0$$

Also ist 0 ein absorbierender Zustand und alle anderen Zustände $k \in \mathbb{N}$ sind transient.

Definition 2.4.5. *Definiere die Gesamtzahl der Besuche in einem Zustand y*

$$N(Y) = \sum_{k=0}^{\infty} 1_{\{X_k = y\}} \ : \Omega \longrightarrow \mathbb{N}_0 \cup \{\infty\}$$

Sowie die Greenfunktion:

$$G(x,y) = \mathbb{E}_x\left[N(y)\right] = \sum_{k=0}^{\infty} p^k(x,y)$$

Lemma 2.4.6. *Aufgrund der Definition gilt für die Greenfunktion:*

$$G(y,y) = \frac{1}{1-F(y,y)} \qquad G(x,y) = 1_{\{x=y\}} + F(x,y)G(y,y)$$

Beweis.

$$G(x,y) = \mathbb{E}_x\left[N(y)\right] = \sum_{k=1}^{\infty} \mathbb{P}_x(N(y) \geq k) = 1_{\{x=y\}} + \sum_{k=1}^{\infty} \mathbb{P}_x(T_y^k < \infty)$$

$$\overset{2.4.2}{=} 1_{\{x=y\}} + \sum_{k=1}^{\infty} F(x,y)F(y,y)^{k-1} = 1\{x=y\} + \frac{F(x,y)}{1-F(y,y)}$$

□

Satz 2.4.7.

a) Für jedes $x \in S$ sind äquivalent:

- **a$_1$)** *x ist rekurrent, d.h.* $\mathbb{P}_x(\exists n \geq 1 : X_n = x) = 1$
- **a$_2$)** $\mathbb{P}_x($*für unendlich viele* $n \in \mathbb{N} : X_n = x) = 1$
- **a$_3$)** $\mathbb{E}_x\left[N(x)\right] = \infty$

b) Für jedes $x \in S$ sind äquivalent:

- **b$_1$)** *x ist transient, d.h.* $\mathbb{P}_x(\exists n \geq 1 : X_n = x) < 1$ *bzw.* $F(x,x) < 1$
- **b$_2$)** $\mathbb{P}_x($*für unendlich viele* $n \in \mathbb{N} : X_n = x) = 0$
- **b$_3$)** $\mathbb{E}_x\left[N(x)\right] < \infty$*, d.h.* $G(x,x) < \infty$

Beweis. a) $a_2 \Rightarrow a_1$ *trivial*
$a_2 \iff N_x = \infty$ $\mathbb{P}_x$*-fast sicher* $\Rightarrow a_3$
Mit Lemma 2.4.6 folgt:

$$F(x,x) = 1 \iff G(x,x) = \infty \Rightarrow a_3$$

Mit Proposition 2.4.2 folgt aus $F(x,x) = 1$ für alle $k \in \mathbb{N}$:

$$\mathbb{P}_x(T_x^k < \infty) = F(x,x)^k = 1 \Rightarrow \mathbb{P}(x \text{ wird unendlich oft angenommen}) = 1$$

b) $b_1 \iff \neg a_1 \iff \neg a_3 \iff b_2$
Ferner: $b_3 \Rightarrow N_x < \infty$ $\mathbb{P}_x$*-fast sicher* $\iff b_2 \iff \neg a_2 \iff \neg a_3 \iff b_3$

□

Satz 2.4.8. *Sei x ein rekurrenter Zustand und es gelte $F(x,y) > 0$ für einen Zustand y. Dann ist auch y rekurrent mit $F(x,y) = F(y,x) = 1$.*

Beweis. Sei $F(x,y) > 0$ und $y \neq x$. Dann existiert ein $k \in \mathbb{N}$ mit $p^k(x,y) > 0$ und folglich gibt es $x_1, \dots x_k \in S$ mit $x_k = y, x_i \neq x$ und $\mathbb{P}_x(X_i = x_i) > 0$ für $i = 1, \dots, k$. Dann gilt:

$$1 - F(x,x) = \mathbb{P}_x(T_x^1 = \infty) \geq \mathbb{P}_x(X_1 = x_1, \dots, X_n = x_n, T_x^1 = \infty)$$

$$\overset{ME}{=} \mathbb{P}_x(X_1 = x_1, \dots, X_n = x_n)\mathbb{P}_y(T_x^1 = \infty) = \mathbb{P}_x(X_1 = x_1, \dots, X_n = x_n)\,(1 - F(y,x))$$

Wegen $\mathbb{P}_x(X_1 = x_1, \dots, X_n = x_n) > 0$ *gilt:* $F(x,x) = 1 \Rightarrow F(y,x) = 1$ *Da* $F(y,x) > 0$ *ist, existiert ein* $l \in \mathbb{N}$ *mit* $p^l(y,x) > 0$ *und daher gilt:*

$$p^{l+n+k}(y,y) \geq p^l(y,x)p^n(x,x)p^k(x,y)$$

und somit:

$$G(y,y) = \sum_{m=0}^{\infty} p^m(y,y) \geq \sum_{m=0}^{\infty} p^{l+n+k}(y,y) \geq \underbrace{p^l(y,x)}_{>0} \quad \underbrace{G(x,x)}_{=\infty,\ da\ F(x,x)=1} \quad \underbrace{p^k(x,y)}_{>0} = \infty$$

Daraus folgt, dass y *rekurrent ist. Vertauschen der Rollen von* x *und* y *liefert:* $F(x,y) = 1$ □

Lemma 2.4.9. *Für alle Zustände* $x, y, z \in S$ *gilt:*

$$F(x,y) \geq F(x,z)F(z,y)$$

Beweis. starke Markov-Eigenschaft. □

Definition 2.4.10. *Die Markov-Kette* $(X_n)_{n \in N_0}$ *auf dem Zustandsraum* S *heißt* <u>*irreduzibel*</u> *, falls für alle Zustände* $x, y \in S$ *gilt:*

$$F(x,y) > 0 \qquad (\text{oder äquivalent: } G(x,y) > 0 \quad \forall x, y \in S)$$

Satz 2.4.11. *Sei* $(X_n)_{n \in N_0}$ *eine irreduzible Markov-Kette. Dann gilt:*

(i) Es sind entweder alle Zustände $x \in S$ *rekurrent oder alle Zustände* $x \in S$ *transient.*

(ii) Falls $|S| > 1$ *ist, gibt es keine absorbierenden Zustände.*

(iii) Falls S *endlich ist, sind alle Zustände rekurrent.*

Beweis. *(i) Falls ein rekurrenter Zustand* $x \in S$ *existiert, so ist wegen* $F(x,y) > 0$ *und dem Satz 2.4.8 auch jeder andere Zustand* y *rekurrent.*

(ii) Falls x *absorbierend ist, dann gilt* $F(x,y) = 0 \forall y \neq x$

(iii) Für alle Zustände $x \in S$ *gilt:*

$$\sum_{y \in S} G(x,y) = \sum_{y \in S} \sum_{k \in \mathbb{N}_0} p^k(x,y) \overset{Fubini}{=} \sum_{k \in \mathbb{N}_0} \underbrace{\sum_{y \in S} p^k(x,y)}_{=1} = \infty$$

Da S *endlich ist existiert ein* $y \in S$ *mit* $G(x,y) = \infty$*. Nach Voraussetzung ist* $F(y,x) > 0$ *und wir erhalten:*

$$\exists k \in \mathbb{N} : p^k(y,x) > 0 \Rightarrow p^{n+k}(x,x) \geq p^n(x,y)p^k(y,x)$$

$$\Rightarrow G(x,x) \geq \sum_{n=0}^{\infty} p^n(x,y)p^k(y,x) = \underbrace{G(x,y)}_{=\infty} \underbrace{p^k(y,x)}_{>0} = \infty$$

$\Rightarrow$ *x rekurrent.*

□

Definition 2.4.12. *Zwei Zustände* $x, y \in S$ *heißen* <u>*kommunizierend*</u> *(kurz* $x \leftrightarrow y$*) falls gilt:*

$$F(x,y)F(y,x) > 0$$

Dies definiert eine Äquivalenzrelation auf S*. Für alle Zustände* $x \in S$ *definiert die Menge* C_x *aller mit* x *kommunizierender Zustände*

$$C_x = \{y \in S : x \leftrightarrow y\}$$

eine <u>kommunizierende Klasse</u>. Definiert man $S_0 = S/_{\leftrightarrow}$ so lässt sich S als disjunkte Vereinigung kommunizierender Klassen schreiben:

$$S = \dot{\bigcup_{x \in S_0}} C_x$$

Bemerkung 2.4.13. *(i) Für jede kommunizierende Klasse $C \subset S$ gilt: Entweder sind alle $x \in C$ rekurrent, oder alle $x \in C$ sind transient.*

(ii) Eine Teilmenge $C \subset S$ heißt abgeschlossen, falls für alle $x \in C$ gilt:

$$F(x,y) > 0 \Rightarrow y \in C$$

Mit anderen Worten:

$$\mathbb{P}_x(\exists n \geq 1 : X_n = y) = 0 \qquad \forall x \in C, \forall y \notin C,$$

Lemma 2.4.14. *Falls C eine kommunizierende Klasse rekurrenter Zustände ist, dann ist C abgeschlossen.*

Beweis. Angenommen, C sei nicht abgeschlossen. Dann existieren Zustände $i \in C$, $j \in S \setminus C$ und eine positive Zahl m sodass gilt: $\mathbb{P}_i(X_m = j) > 0$.
Es gilt aber $\mathbb{P}_i(\{X_m = j\} \cap \{X_n = i \text{ für unendlich viele } n\}) = 0$ und daraus folgt:

$$\mathbb{P}_i(X_n = i \text{ für unendlich viele } n) < 1$$

Dies bedeutet, dass i nicht rekurrent sein kann, was einen Widerspruch zu den Voraussetzungen darstellt. □

Proposition 2.4.15. *Sei eine Teilmenge R von S gegeben durch:*

$$R = \{x \in S | x \text{ rekurrent}\} = \{x \in S | F(x,x) = 1\}.$$

Dann existiert eine disjunkte Zerlegung $R = \dot{\bigcup}_{i \in I} R_i$ mit irreduzibel und abgeschlossenen R_i.

Beweis. Für alle $x \in \Omega$ definiere die Menge $C_x = \{y \in S : F(x,y) > 0\}$. Nach dem letzten Satz gilt $C_x \subset R$ (d.h. y rekurrent $\forall y \in C_x$) und $F(y,x) > 0$.
<u>Behauptung:</u> Es gilt entweder $C_x \cap C_y = \emptyset$ oder $C_x = C_y$.
Angenommen $C_x \cap C_y \neq \emptyset$. Dann existiert ein $z \in R$ mit $F(x,z) > 0$ und $F(y,z) > 0$. Es gilt aber:

$$F(x,y) \geq \underbrace{F(x,z)}_{>0} \underbrace{F(z,y)}_{>0} > 0$$

Sei also $\omega \in C_y$. Dann ist:

$$F(x,\omega) \geq \underbrace{F(x,y)}_{\geq 0} F(y,\omega) > 0 \qquad \Rightarrow \omega \in C_x$$

Daraus folgt $C_y \subset C_x$ (und analog $C_x \subset C_y$) und damit $C_x = C_y$. Dadurch bekommen wir eine disjunkte Zerlegung $R = \bigcup_{x \in R_0} C_x$ für eine geeignete Teilmenge $R_0 \subset R$. □

2.5. Transienz und Rekurrenz für Irrfahrten auf $\mathbb{Z}^d$

Sei also $S = \mathbb{Z}^d$ mit Übergangswahrscheinlichkeit p gegeben durch

$$p(x,y) = \begin{cases} \frac{1}{2d} & |x-y| = 1 \\ 0 & \text{sonst} \end{cases}.$$

Wir verwenden: $(X_n)_{n\in N_0}$ rekurrent $\iff \sum\limits_{n=0}^{\infty} p^n(0,0) = +\infty$.
Zunächst ein semi-rigoroses Argument für (rekurrent $\iff d \leq 2$):
Die Behauptung folgt aus $p^n(0,0) \sim c_d n^{-\frac{d}{2}}$ für $n \to \infty$.
Letzteres folgt aus dem Zentralen Grenzwertsatz, kann aber nur gelten für gerades n, für ungerades n gilt nämlich $p^n(0,0) = 0$.
Betrachte $(\xi_k)_{k\in\mathbb{N}}$ iid Zufallsvariablen auf $\mathbb{R}^d$ mit Verteilung $p^2(0,x)$. Daraus folgt:

$$S_n = \sum_{k=1}^{n} \xi_k = X_{2n} \qquad \text{Zufallsvariable auf Irrfahrt}$$

$$\Rightarrow G(0,0) = \sum_{n=0}^{\infty} p^{2n}(0,0) = \sum_{n=0}^{\infty} \mathbb{P}(S_n = 0) = \sum_{n=0}^{\infty} \mathbb{P}\left(S_n \in \left[-\frac{1}{2}, \frac{1}{2}\right)^d\right)$$

Sei also n fest. Unser <u>Ziel</u> ist es zu zeigen, dass gilt:

$$\mathbb{P}\left(S_n \in \left[-\frac{1}{2}, \frac{1}{2}\right)^d\right) \sim c_d\, n^{\frac{1}{2}}$$

Der Multivariante Zentrale Grenzwertsatz liefert:

$$\mathbb{P}\left(\frac{1}{\sqrt{n}} S_n \in \left[-\frac{1}{2}, \frac{1}{2}\right)^d\right) \longrightarrow N_{0,A}\left(\left[-\frac{1}{2}, \frac{1}{2}\right)^d\right) \qquad \text{für } n \to \infty$$

mit Kovarianzmatrix A von $\xi_1 = (\xi_1^1, \ldots, c_1^d)$ (da iid). Daraus folgt:

$$A_{ij} = \operatorname{Cov}(\xi_1^i, \xi_1^j) = \mathbb{E}\left[\xi_1^i \xi_1^j\right] = \begin{cases} 0 & i \neq j \\ b_d \text{ (Konstante)} & i = j \end{cases}$$

$\Rightarrow N_{0,A}\left(\left[-\frac{1}{2}, \frac{1}{2}\right)^d\right) = c_d$ ist konstant.
Daraus folgt:

$$LS = \mathbb{P}\left(S_n \in \left[-\frac{\sqrt{n}}{2}, \frac{\sqrt{n}}{2}\right)^d\right) \overset{\text{(Zerlegung)}}{=} \sum_{\substack{z\in\mathbb{Z}^d \\ |z|\leq \frac{\sqrt{n}}{2}}} \mathbb{P}\left(S_n \in \left[z_i - \frac{1}{2}, z_i + \frac{1}{2}\right)^d\right) \tag{2.10}$$

Verwende: $\exists C :\ \forall z, z' \in \mathbb{Z}^d,\ |z| < \dfrac{\sqrt{n}}{2},\ z - z' \ \in \mathbb{Z}^d$:

$$\frac{1}{c} \leq \mathbb{P}\left(S_n \in \left[z_i - \frac{1}{2}, z_i + \frac{1}{2}\right)^d\right) \leq C$$

$$\Rightarrow \mathbb{P}(X_{2n} = 0) = \mathbb{P}_0(Y_{2n}^1 = 0)\mathbb{P}_0(Y_{2n}^1 = 0) \sim \frac{1}{\sqrt{\pi n}} \frac{1}{\sqrt{\pi n}} = \frac{1}{\pi n} \qquad \text{da } p^{2n}(0,0) \sim \frac{1}{\pi n}$$

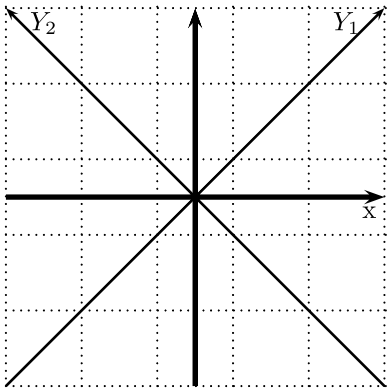
$\Rightarrow \sum\limits_{n\in\mathbb{N}} p^n(0,0) = +\infty$, denn $\sum\limits_{n\in\mathbb{N}} \dfrac{1}{n} = +\infty$ $\qquad \Rightarrow$ *rekurrent.*

<u>$d = 1$</u>

$$p^{2n+1}(x,x) = 0, \qquad p^{2n}(x,x) \sim \frac{1}{\sqrt{\pi}} \frac{1}{\sqrt{n}}$$

$\Rightarrow$ rekurrent.

<u>$d=2$</u> Wir betrachten die Koordinaten Y_1, Y_2 mit (Y_n^1), (Y_n^2) unabhängigen symmetrischen Irrfahrten auf $\frac{1}{2}\mathbb{Z}$.



Dann ist $X_n := Y_n^1 + Y_n^2$ symmetrische Irrfahrt auf $\mathbb{Z}^2$. Es gilt:

$$p^{2n+1}(0,0) = 0, \qquad p^{2n}(0,0) \sim \frac{1}{\sqrt{\pi n}}$$

$\Rightarrow$ rekurrent.

Satz 2.5.1. *Die symmetrische Irrfahrt auf $\mathbb{Z}^d$ ist für $d \geq 3$ transient.*

Beweis. für d=3

$$p(x,y) = \begin{cases} \frac{1}{6} & |x-y| = 1 \\ 0 & \textit{sonst} \end{cases}$$

Wir suchen nach Bedingungen für Rückkehr zu 0 in N Schritten:

- *N gerade, $N = 2n$*
- *i Schritte nach rechts $\Rightarrow$ i Schritte nach links*
- *j Schritte nach oben $\Rightarrow$ j Schritte nach unten*
- *k Schritte nach vorne $\Rightarrow$ k Schritte nach hinten*

$\Rightarrow i + j + k = n$

$$p^{2n}(0,0) = \sum_{\substack{i,j,k \in \mathbb{N}_0 \\ i+j+k=n}} \frac{(2n)!}{(i!j!k!)^2} \left(\frac{1}{6}\right)^{2n} = \binom{2n}{n} \left(\frac{1}{2}\right)^{2n} \sum_{\substack{i,j,k \in \mathbb{N}_0 \\ i+j+k=n}} \binom{n}{i\ j\ k}^2 \left(\frac{1}{3}\right)^{2n} \qquad \textit{mit } \binom{n}{i\ j\ k} = \frac{n!}{i!j!k!}$$

Es gilt:

$$\sum_{\substack{i,j,k \in \mathbb{N}_0 \\ i+j+k=n}} \binom{n}{i\ j\ k} a^i b^j c^k = (a+b+c)^n \qquad \textit{insbesondere:} \sum_{\substack{i,j,k \in \mathbb{N}_0 \\ i+j+k=n}} \binom{n}{i\ j\ k} \left(\frac{1}{3}\right)^n = 1$$

Ferner gilt für $n = 3m$ und alle i, j, k mit $(i + j + k = n)$:

$$\binom{n}{i\ j\ k} \leq \binom{n}{\frac{n}{3}\ \frac{n}{3}\ \frac{n}{3}}$$

Falls $n = 3m$:

$$p^{6n}(0,0) = p^{2n}(0,0) \leq \binom{2n}{n} \left(\frac{1}{2}\right)^{2n} \binom{n}{\frac{n}{3}\ \frac{n}{3}\ \frac{n}{3}} \left(\frac{1}{3}\right)^n \overset{Stirling}{\sim} \frac{1}{2\sqrt{2\pi}^3} \left(\frac{6}{n}\right)^{\frac{3}{2}} \qquad \textit{für } n \to \infty$$

Ferner:

$$p^{6n}(0,0) \geq \frac{1}{6} p^{6n-2}(0,0) \geq \frac{1}{6^2} p^{6n-4}(0,0)$$

$\Rightarrow \sum p^{6n} < \infty$, $\sum p^{6n-2} < \infty$, $\sum p^{6n-4} < \infty$
$\Rightarrow \sum p^n < \infty \Rightarrow$ *transient.* □

Proposition 2.5.2. *Sei S ein Zustandsraum und $C \subset S$ eine kommunizierende Klasse. Dann gilt:*

(i) Entweder sind alle Zustände $y \in C$ rekurrent oder alle $y \in C$ sind transient.

(ii) Sei $x \in C$ ein rekurrenter Zustand, dann ist C abgeschlossen.

Beweis. Angenommen C sei eine kommunizierende, nicht abgeschlossene Klasse. Dann existieren Zustände $y \in C, z \notin C$ mit $F(z,y) = 0$ (da $z \notin C$) und $F(y,z) > 0$. Also kann y nicht rekurrent sein und es existiert ein $n \in \mathbb{N}$ mit $p^n(y,z) = \delta > 0$ und daraus folgt: $F(y,y) \leq 1 - \delta$. □

2.6. Potentialtheorie

Sei S ein beliebiger Zustandsraum und p ein Markov-Kern. Wir interpretieren den Operator $L : \mathcal{L}^\infty(S) \longrightarrow \mathcal{L}^\infty(S)$:

$$-Lu(x) = (I-p)u(x) = u(x) - \int_S u(y)p(x,dy) = \int_S [u(x) - u(y)]\, p(x,dy)$$

als "Laplace-Operator".
Für S diskret gilt:

$$-Lu(x) = \sum_{y \in S} [u(x) - u(y)]\, p(x,y)$$

Mit anderen Worten: u (als Spaltenvektor) ist Eigenvektor zu p zum Eigenwert 1.

Definition 2.6.1. *Für $A \in \mathcal{S}$ heißt u harmonisch außerhalb A, falls*

$$Lu = 0 \qquad \text{in } A^C := S \setminus A$$

Das heißt, für alle $x \in A^C$ gilt:

$$u(x) = pu(x) = \int_S u(y)p(x,dy)$$

Für $A = \emptyset$ ist u harmonisch. („Eigenfunktion von p zum Eigenwert 1").

Definition 2.6.2. *Dirichlet-Problem*
Sei $A \in \mathcal{S}$ und $f \in \mathcal{L}^\infty(S)$. Dann heißt $u \in \mathcal{L}^\infty(S)$ Lösung des Dirichlet-Problems zu L, f, A falls gilt:

$$Lu = 0 \quad \text{auf } A^C \tag{2.11}$$
$$u = f \quad \text{auf } A \tag{2.12}$$

Betrachte den Sub-Markov-Kern der Trefferverteilung:

$$p_A(x,B) = \mathbb{P}_x(X_{T_A} \in B, T_A < \infty), \qquad p_A(x,S) = \mathbb{P}_X(T_A < \infty) = h_A(x) \quad \text{„Gleichgewichtsverteilung"}$$

Satz 2.6.3. *(i) Für alle $f \in \mathcal{L}^\infty(S)$ ist $u := p_A f$ eine Lösung des obigen Dirichlet-Problems. Es gilt:*

$$u(x) = \mathbb{E}_x\left[f(X_{T_A}) 1_{\{T_A < \infty\}}\right] \qquad \text{mit } T_A = \inf\{n \geq 0 : X_n \in A\}$$

(ii) Falls $f \geq 0$ ist, so ist $h_A = p_A f$ die minimale nicht-negative Lösung des Dirichlet-Problems.

(iii) Falls für alle Zustände $\mathbb{P}$-fast sicher $T_A < \infty$ ist,so ist die Lösung des Dirichlet-Problems eindeutig.

Beweis.

(i) $\underline{x \in A}$

$$T_A = 0 \Rightarrow u(x) = \mathbb{E}_x\left[f(X_0)\right] = f(x) \qquad \textit{mit } X_0 = x \textit{ unter } \mathbb{P}_x$$

$\underline{x \notin A}$

$$T_A \circ \theta_1 + 1 = T_A, \qquad X_{T_A} \circ \theta_1 = X_{T_A}$$

Daher folgt mit der schwachen Markov-Eigenschaft:

$$\begin{aligned} pu(x) &= \mathbb{E}_x\left[u(X_1)\right] = \mathbb{E}_x\left[\mathbb{E}_{X_1}\left[f(X_{T_A})1_{\{T_A<\infty\}}\right]\right] \overset{ME}{=} \mathbb{E}_x\left[\mathbb{E}_x\left[f(X_{T_A}\circ\theta_1)1_{\{T_A\circ\theta_1<\infty\}}|\mathcal{A}_1\right]\right] \\ &= \mathbb{E}_x\left[f(X_{T_A})1_{\{T_A<\infty\}}\right] = u(x) \end{aligned}$$

$$u(x) = pu(x) = \sum_{y\in S} p(x,y)u(y) = \mathbb{E}_x\left[u(X_1)\right]$$

(ii) *Für alle $x \notin A$ und alle $n \in \mathbb{N}$ gilt:*

$$u(x) = pu(x) = \int_A u(y)p(x,dy) + \int_{A^C} u(y)p(x,dy)$$

$$= \int_A f(y)p(x,dy) + \int_{A^C}\left[\int_A \underbrace{u(z)}_{=f(z)} p(y,dz) + \int_{A^C} u(z)p(y,dz)\right]p(x,dy) = \dots$$

$$= \mathbb{E}_x\left[f(X_1)1_{\{X_1\in A\}} + f_2(X_2)1_{\{X_1\notin A, X_2\in A\}} + \dots + f(X_n)1_{\{X_1\notin A,\dots,X_{n-1}\notin A, X_n\in A\}} + u(X_n)1_{\{X_1\notin A,\dots,X_n\notin A\}}\right]$$

Damit gilt:

$$u(x) = \mathbb{E}_x\left[f(X_{T_A})1_{\{T_A\le n\}}\right] + \mathbb{E}_x\left[u(X_n)1_{\{T_A>n\}}\right] \tag{2.13}$$

Falls nun $u \geq 0$ ist gilt:

$$u(x) \geq \mathbb{E}_x\left[f(X_{T_A})1_{\{T_A\le n\}}\right] \overset{n\to\infty}{\longrightarrow} \mathbb{E}_x\left[f(X_{T_A})1_{\{T_A<\infty\}}\right] = p_A f(x)$$

Also gilt: $u(x) \geq p_A f(x)$.

(iii) *Falls für alle Zustände x $\mathbb{P}_x$-fast sicher $(T_A < \infty)$ ist und u, f beschränkt sind, folgt:*

$$\mathbb{E}_x\left[f(X_{T_A})1_{\{T_A\le n\}}\right] \overset{n\to\infty}{\longrightarrow} \mathbb{E}_x\left[f(X_{T_A})\right] \quad \textit{und} \quad \mathbb{E}_x\left[u(X_n)1_{\{T_A>n\}}\right] \overset{n\to\infty}{\longrightarrow} 0$$

Damit erhalten wir: $u(x) = \mathbb{E}_x\left[f(X_{T_A})\right]$.

□

Bemerkung 2.6.4.

(i) Falls $\mathbb{P}_x(T_A = \infty) > 0$ für gegebenen Zustand $x \in S$ ist, dann ist für alle $\alpha \in \mathbb{R}$ eine weitere Lösung des Dirichlet-Problems (2.11) gegeben durch:

$$u(x) = p_A f(x) + \alpha \underbrace{\mathbb{P}_x(T_A = \infty)}_{=1-h_A}$$

(ii) Für die symmetrische Irrfahrt auf $\mathbb{Z}^d$ und jede Teilmenge A mit beschränktem Komplement A^C ist $T_A < \infty$ $\mathbb{P}_x$-fast sicher für alle $x \in \mathbb{Z}^d$.

Beispiel 2.6.5. *„Geburts- und Sterbe-Ketten“*

$$S = \mathbb{N}_0, \quad A = \{0\}, \quad p(0,0) = 1, \quad p(i,i+1) = p, \quad p(i,i-1) = (1-p) = q$$

Gesucht ist $u \geq 0$ mit:

$$\begin{aligned} u(0) &= \quad 1 \\ u(i) &= \quad pu(i+1) + q(i-1) \quad \forall i \geq 1 \end{aligned}$$

Für $p \neq q$ gilt: $\qquad u(i) = A + B\left(\frac{q}{p}^i\right).$

1. $p < \frac{1}{2} \Rightarrow B = 0 \Rightarrow A = 1 \Rightarrow u(i) = 1 \ \forall i \geq 0$
   $\Rightarrow \mathbb{P}_i(T_0 < \infty) = 1 \ \forall i \Rightarrow$ *sicheres Aussterben der Population*
2. $p > \frac{1}{2}, A = 0, B = 1 \ \Rightarrow \mathbb{P}_i(T_0 < \infty) = u(i) = \left(\frac{q}{p}\right)^i$
3. $p = \frac{1}{2}, u(i) = A + B_i.$*wegen* $0 \leq u \leq 1$ *folgt:* $u(i) = 1$, *d.h.* $\mathbb{P}_i(T_i < \infty) = 1$. *(„Gambler's Ruin“).*

Beispiel 2.6.6.

$$S = \mathbb{N}_0, \quad A = \{0\}, \quad p(i,i+1) = p_i, \quad p(i,i-1) = q_i = 1 - p_i \ (\forall i \geq 1) \quad p(0,0) = 1$$

Gesucht ist eine Funktion u mit:

$$u(i) = \begin{cases} 1 & i = 0 \\ u(i+1)p_i + u(i-1)q_i & i \neq 0 \end{cases}$$

Definiere eine neue Funktion $v(i) = u(i-1) - u(i)$. Dann gilt:

$$p_i v(i+1) = q_i v(i) \quad \Rightarrow \quad v(i+1) = \frac{q_i}{p_i} v(i) = \underbrace{\frac{q_i \cdots q_1}{p_i \cdots p_1}}_{=:\gamma_i} v(1) = \gamma_i v(1) \qquad \textit{mit } \gamma_0 := 1$$

Zurück zu $u(i)$:

$$u(i) = u(i-1) - v(i) = u(0) - \sum_{1 \leq k \leq i} v(k) = 1 - \sum_{0 \leq k < i} \gamma_k v(1)$$

mit $v(1)$ als freie Variable.

Poisson-Gleichung

Aufgrund der Linearität des Laplace-Operators L gilt das Superpositionsprinzip.
Die Lösung zu (f,h) ist gegeben durch die Summe der Lösung zu $(f,0)$ und zu $(0,h)$:

$$u_{f,h} = u_f + u_h$$

Deswegen gelte o.B.d.A: $f = 0$.
Gegeben h, suchen wir ein u, das die Poisson-Gleichung löst:

$$-Lu = \quad h \quad \text{auf } A^C \tag{2.14}$$
$$u = \quad 0 \quad \text{auf } A \tag{2.15}$$

Sei v eine Lösung zu

$$-Lv = \quad h \quad \text{auf } S$$

und w eine Lösung zu

$$\begin{aligned} -Lw = \quad & h \quad \text{auf } A^C \\ w = \quad & v \quad \text{auf } A, \end{aligned}$$

dann ist auch $u = v - w$ eine Lösung von 2.14. Daher lösen wir die Poisson-Gleichung zunächst für $A = \emptyset$. Betrachte die Greenfunktion und den Green-Operator:

$$G(x, dy) = \sum_{n=0}^{\infty} p^n(x, dy) \qquad \text{also } Gh(x) = \sum_{n=0}^{\infty} \underbrace{p^n h(x)}_{\geq 0} \quad (\forall h \in \mathcal{L}_+(S))$$

<u>Diskret</u>:

$$Gh(x) = \sum_{y \in S} G(x, y) h(y) \qquad G(x, y) = \sum_{n=0}^{\infty} p^n(x, y)$$

In vielen Fällen ist $G(x, y) = \infty$ bzw. $Gh(x) = \infty$.

Satz 2.6.7. *Für alle $h \in \mathcal{L}_+(S)$, für die $Gh(x) < \infty$ für alle $x \in S$ gilt, ist $u(x) = Gh(x)$ Lösung der Poisson-Gleichung $(-Lu = h)$ auf S.*

Beweis. Nach Voraussetzung gilt: $Gh(x) < \infty$ daraus folgt:

$$\sum_{n=0}^{\infty} p^n h(x) \qquad \textit{ist absolut summierbar,}$$

denn mit der Jensen-Ungleichung folgt:

$$\sum_{n=0}^{\infty} |p^n h(x)| \leq \sum_{n=0}^{\infty} p^n |h|(x) = G|h|(x)$$

$\Rightarrow \quad u(x) = Gh(x) = h(x) + p(Gh)(x) = h(x) + pu(x)$ □

Definition 2.6.8. *Für alle $A, B \in \mathcal{S}$ definiere die Anzahl der Besuche in B vor dem Treffer in A durch*

$$N_A(B) = \sum_{n=0}^{T_A - 1} 1_B(X_n)$$

und dadurch den <u>Green-Operator zu A</u>:

$$G_A(x, B) := \mathbb{E}_x \left[N_A \right](B) = \mathbb{E}_x \left[\sum_{n=0}^{T_A - 1} 1_B(X_n) \right]$$

Für $h \geq 0$ gilt:

$$G_A h(x) = \mathbb{E}_x \left[\sum_{n=0}^{T_A - 1} h(X_n) \right]$$

Speziell für $A = \emptyset$:

$$G_A(x, B) = \mathbb{E}_x \left[\sum_{n=0}^{\infty} 1_B(X_n) \right] = \sum_{n=0}^{\infty} p^n(x, B) = G(x, B)$$

Proposition 2.6.9. *Es gilt $G = G_A + p_A G$ $(v = u + w)$ im Sinne von:*

$$G(x, B) = G_A(x, B) + p_A(G(\cdot\,, B))(x)$$

Beweis. *Zu Zeigen*

$$Gh(X) = G_a H(X) + p_A(Gh)(x) \qquad (\forall h \geq 0)$$

Wir verwenden die starke Markov-Eigenschaft:

$$p_A(Gh)(x) = \mathbb{E}_x\left[Gh(X_{T_A})1_{\{T_A<\infty\}}\right] = \mathbb{E}_x\left[\sum_{n=0}^{\infty} p^n h(X_{T_A})1_{\{T_A<\infty\}}\right] = \mathbb{E}_x\left[\mathbb{E}_{X_{T_A}}\left[\sum_{n=0}^{\infty} h(X_n)\right]1_{\{T_A<\infty\}}\right]$$

$$\overset{SME}{=} \mathbb{E}_x\left[\sum_{n=0}^{\infty} p^n h(X_{T_A+n})1_{\{T_A<\infty\}}\right] = \mathbb{E}_x\left[\sum_{n=T_A}^{\infty} h(X_n)\right]$$

Daraus folgt:

$$G_A h(x) + p_A G(x) = \mathbb{E}_x\left[\sum_{n=0}^{T_A-1} h(X_n)\right] + \mathbb{E}_x\left[\sum_{n=T_A}^{\infty} h(X_n)\right] = \mathbb{E}_x\left[\sum_{n=0}^{\infty} h(X_n)\right] = Gh(x)$$

□

Satz 2.6.10. *Für alle $h \in \mathcal{L}(S)$ mit ($G^A|h| < \infty$ auf S) ist $u := G_A h$ Lösung der Poisson-Gleichung:*

$$-Lu = h \quad \text{auf } A^C \tag{2.16}$$

$$u = 0 \quad \text{auf } A \tag{2.17}$$

Beweis. *Für alle $x \notin A$ gilt:*

$$p(G^A h)(x) = \mathbb{E}_x\left[\mathbb{E}_{x_1}\left[\sum_{n=0}^{T_A-1} h(X_n)\right]\right] = \mathbb{E}_x\left[\sum_{n=1}^{T_A-1} h(X_n)\right] = G_A h(x) - \underbrace{h(x)}_{=\mathbb{E}_x[h(X_0)]}$$

□

Bemerkung 2.6.11. *Gegeben sei eine symmetrische oder asymmetrische Irrfahrt auf $\mathbb{Z}^d$ mit beschränkten A^C und $\mathbb{E}_X[T_A] < \infty$. Dann gilt:*

$$G_A|h|(x) \leq ||h||_\infty \mathbb{E}_x[T_A] < \infty \qquad (\forall h \text{ beschränkt})$$

Zusammenfassung:
Falls $\mathbb{E}_x[T_A] < \infty$ ist, so gibt es für alle $f \in \mathcal{L}^\infty(S)$ ein $u \in \mathcal{L}^\infty(S)$ mit:

$$-Lu = h \quad \text{auf } A^C \tag{2.18}$$

$$u = f \quad \text{auf } A \tag{2.19}$$

Nämlich:

$$u(x) = \mathbb{E}_x[f(X_{T_A})] + \mathbb{E}_x\left[\sum_{n=0}^{T_A-1} h(X_n)\right]$$

2.7. Invariante Maße, invariante Verteilungen

Gegeben sei ein Markov-Kern p auf dem Zustandsraum S.

Definition 2.7.1. *Ein Maß μ auf S heißt invariantes Maß, falls $\mu p = \mu$ gilt. Das bedeutet:*

$$\int_S p(x,B)\mu(dx) = \mu(B) \qquad (\forall B \in \mathcal{S})$$

Das Maß μ heißt invariante Verteilung , falls es zusätzlich ein Wahrscheinlichkeitsmaß ist.
Für endlichen/diskreten Zustandsraum S:
μ ist als „Zeilenvektor“ Eigenvektor zum Eigenwert 1 bezüglich der Multiplikation von links.

$$\mu(y) = \sum_{x \in S} \mu(x) p(x, y) \qquad (\forall y \in S)$$

Beispiel 2.7.2. *Die Gleichverteilung auf $\mathbb{Z}^d$ ist ein invariantes Maß für die symmetrische Irrfahrt.*

Lemma 2.7.3. *Eine Gleichverteilung auf einem diskreten Zustandsraum S ist genau dann invariantes Maß, wenn $(p(x, y))_{x,y \in S}$ doppel-stochastisch ist. Das heißt:*

$$\sum_{y \in S} p(\hat{x}, y) = 1, \quad \sum_{x \in S} p(x, \hat{y}) = 1 \qquad (\forall \hat{x}, \hat{y} \in S)$$

Beispiel 2.7.4. *Für die asymmetrische Irrfahrt auf $\mathbb{Z}$ mit. $p(i, i + 1) = p$, $p(i, i - 1) = q$ ist*

$$\mu(x) = \left(\frac{p}{q}\right)^x$$

ein invariantes Maß.

Proposition 2.7.5. *Für jedes invariante Maß μ definiert*

$$\hat{p}(x, y) := \frac{\mu(y)}{\mu(x)} p(y, x)$$

einen Markov-Kern (“Dualer Kern„) wiederum mit μ als invariantem Maß.

Beweis. *(i)* $\hat{p} \geq 0$

(ii) $\sum_{y \in S} \hat{p}(x, y) = \frac{1}{\mu(x)} \underbrace{\sum_{y \in S} \mu(y) p(y, x)}_{\mu(x)} = 1$

(iii) $\sum_{y \in S} \mu(x) \hat{p}(x, y) = \sum_{y \in S} \mu(y) p(y, x) = \mu(x)$

□

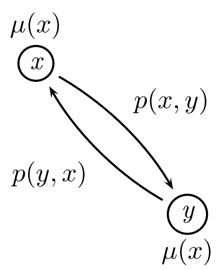
Reversible Maße

Wir wollen nun untersuchen, unter welchen Bedingungen invariante Maße existieren und “eindeutig sind„. Eine einfache Bedingung für die Existenz invarianter Maße ist die Existenz von reversiblen Maßen.

Definition 2.7.6. *Ein Maß μ auf S heißt reversibles Maß zur Markov-Kette mit Übergangswahrscheinlichkeit p, falls:*

$$\mu(x) p(x, y) = \mu(y) p(y, x) \qquad \forall x, y \in S \tag{2.20}$$

Die Gleichung 2.20 heißt „detailed balance equation“ .



Nun gilt:

$$\sum_{x\in S}\mu(x)p(x,y)=\sum_{y\in S}\mu(y)p(y,x)=\mu(y)$$

Bemerkung 2.7.7. *Sei nun μ ein invariantes Maß der Markov-Kette $(X_j)_{0\leq j\leq n}$, X_0 sei „verteilt“ nach μ und $Y_m = X_{n-m}$ für $0 \leq m \leq n$. Dann ist $(Y_m)_{0\leq m\leq n}$ wieder eine Markov-Kette mit der dualen Übergangswahrscheinlichkeit:*

$$q(x,y)=\frac{\mu(y)p(y,x)}{\mu(x)}$$

Wenn μ ein reversibles Maß der Kette X_n ist, dann ist $q(x,y)=p(x,y)$. Die Kette stimmt mit der Kette rückwärts in der Zeit betrachtet überein.

Beispiel 2.7.8. *Gegeben sei ein Graph $G=(S,E)$ mit Ecken S und Kanten E. Für $i,j\in S$ sei $a_{ij}=a_{ji}=1 \iff (i,j)\in E$ ($a_{ij}=0$ sonst). Ferner sei für jeden Zustand $i\in S$: $\mu(i)=\sum_j a_{ij}<\infty$. Betrachte nun die symmetrische Irrfahrt auf G gegeben durch: $p_{ij}=\dfrac{a_{ij}}{\mu(i)}$. Dann gilt*

$$\mu(i)p(i,j)=a_{ij}=a_{ji}=\mu(j)p(j,i),$$

was μ zu einem reversiblen Maß macht.

Für den Rest dieses Abschnitts sei stets X_n eine Markov-Kette mit Übergangswahrscheinlichkeit p auf dem Zustandsraum S.

Satz 2.7.9. *Sei $x\in S$ ein rekurrenter Zustand und $\tau=\inf\{n\geq 1 : X_n=x\}$ die Trefferzeit von x. Dann definiert*

$$\mu(y)=\mathbb{E}_x\left[\sum_{n=0}^{\tau-1}1_{\{X_n=y\}}\right]=\sum_{n=0}^{\infty}\mathbb{P}_x(X_n=y,\tau>n)$$

ein invariantes Maß.

Beweis. *Schreibe $\bar{p}_n(x,y)=\mathbb{P}_x(X_n=y,\tau>n)$. Mit Fubini folgt:*

$$\sum_{y\in S}\mu(y)p(y,z)=\sum_{n=0}^{\infty}\sum_{y\in S}\bar{p_n}(x,y)p(y,z)$$

1.Fall: $z\neq x$

$$\sum_{y\in S}\bar{p}_n(x,y)p(y,z)=\sum_{y\in S}\mathbb{P}_x(X_n=y,\tau>n,X_{n+1}=z)=\mathbb{P}_x(\tau>n+1,X_{n+1}=z)=\bar{p}_{n+1}(x,z)$$

$$\sum_{n=0}^{\infty}\bar{p}_{n+1}(x,z)=\sum_{n=0}^{\infty}\bar{p}_n(x,z)=\mu(z) \qquad da\ p_0(x,z)=0\ wegen\ x\neq z$$

2.Fall: $z=x$

$$\sum_{y\in S}\bar{p}_n(x,y)p(y,x)=\sum_{y\in S}\mathbb{P}_x(X_n=y,\tau>n,X_{n+1}=x)=\mathbb{P}_x(\tau=n+1)$$

$$\sum_{n=0}^{\infty}\mathbb{P}_x(\tau=n+1)=\sum_{n=0}^{\infty}\mathbb{P}_x(\tau=n)\overset{rek.}{=}1=\mu(x)$$

$\Rightarrow \mu p=\mu$
Zu zeigen: $\mu(y)<\infty$
Aus $\mu p=\mu$ folgt $\mu p^k=\mu\ \forall k$. Zusätzlich gilt: $\mu(x)=1$.
Falls $p^n(y,x)>0$, dann gilt wegen $\sum_{y\in S}\mu(y)p^n(y,x)=\mu(x)$ auch $\mu(y)<\infty$.
Somit folgt aus $\delta_{xy}=\mathbb{P}_x(\tau_y<\infty)>0$, dass auch δ_{yx} gilt. Ist aber $\delta_{xy}=0$, dann ist $\mu(y)=0$.

□

Bemerkung 2.7.10. *Dies ist der sogenannte cycle „trick“.*
$\mu(y)$ ist die erwartete Anzahl von Besuchen in y zwischen 0 und $\tau - 1$ und $\mu p(y)$ ist die erwartete Anzahl von Besuchen in y zwischen 1 und τ. Daraus folgt: $\mu = \mu p$

Satz 2.7.11. *Falls p irreduzibel und rekurrent ist, dann ist das invariante Maß bis auf einen Faktor eindeutig.*

Beweis. *Aus Satz 2.7.9 wissen wir, dass ein invariantes Maß ν existiert. Sei nun $a \in S$ ein fixer Zustand. Dann gilt für alle Zustände $z \in S$:*

$$\nu(z) = \sum_{y \in S} \nu(y)p(y,z) = \nu(a)p(a,z) + \sum_{y \in S \setminus \{a\}} \nu(a)p(a,y)p(y,z) + \sum_{y \in S \setminus \{a,b\}} \nu(x)p(x,y)p(y,z)$$

$$= \nu(a)\mathbb{P}_a(X_1 = z) + \nu(a)\mathbb{P}_a(X_1 \neq a, X_2 = z) + \mathbb{P}_\nu(X_0 \neq a, X_1 \neq a, X_2 = z)$$

Setzt man dies induktiv fort erhält man:

$$\nu(z) = \nu(a) \sum_{m=1}^{n} \mathbb{P}_a(X_k \neq a, 1 \leq k \leq n, X_n = z) + \underbrace{\mathbb{P}_\nu(X_j \neq a, 1 \leq j < n, X_m = z)}_{\geq 0}$$

$\Rightarrow \nu(z) \geq \nu(a)\mu(z)$ *wobei:*

$$\mu(z) = \sum_{m=1}^{\infty} \mathbb{P}_a(X_m \neq a, 1 \leq k \leq m, X_m = z) = \begin{cases} \sum\limits_{n=1}^{\infty} \mathbb{P}_a(\tau > m-1, X_m = a) = \sum\limits_{n=1}^{\infty} \mathbb{P}(\tau = n) & z = a \\ \sum\limits_{n=1}^{\infty} \mathbb{P}_s(\tau > m-1, X_m = z) = \sum\limits_{n=1}^{\infty} \mathbb{P}(\tau > n, X_n = 2) & z \neq a \end{cases}$$

Aus Satz 2.7.9 folgt nun, dass μ ein invariantes Maß mit $\mu(a) = 1$ ist.

$$\nu(x) = \nu(a)\mu(x) \qquad \forall x \in S$$

□

Invariante Maße können für transiente Ketten existieren (z.B. Irrfahrt auf $\mathbb{Z}^d$). Es existieren aber keine invarianten Verteilungen für transiente Ketten.

Satz 2.7.12. *Falls eine invariante Verteilung π existiert, dann sind alle Zustände mit $\pi(x) > 0$ rekurrent.*

Beweis.

$$G(x,y) = \frac{\delta_{xy}}{1 - \delta_{yy}} \qquad \delta_{xy} = \mathbb{P}_x(T_y < \infty)$$

Für einen Zustand $y \in S$ mit $\pi(y) > 0$ gilt mit Fubini und $\pi^n = \pi$:

$$\infty = \sum_{n=1}^{\infty} \pi(y) = \sum_{x \in S} \pi(x) \sum_{n=1}^{\infty} p^n(x,y) = \sum_{x \in S} \pi(x) \frac{\delta_{xy}}{1 - \delta_{yy}}$$

$$\leq \sum_{x \in S} \pi(x) \frac{1}{1 - \delta_{yy}} = \frac{1}{1 - \delta_{yy}}$$

$\Rightarrow \delta_{yy} = 1$ *und y rekurrent* □

Satz 2.7.13. *Falls p irreduzibel ist und eine invariante Verteilung π hat, gilt:*

$$\pi(x) = \frac{1}{\mathbb{E}_x[\tau_x]} \qquad \textit{mit } \tau_x = \inf\{n \geq 1 : X_n = x\}$$

Beweis. *Wegen $\pi(x) = \sum_{y\in S} \pi(y)p^n(y,x)$ und Irreduzibilität folgt für alle $x \in S$:$\pi(x) > 0$.*
Mit 2.7.12 folgt: Alle Zustände $x \in S$ sind rekurrent. Dann können wir durch:

$$\mu(y) = \sum_{n=0}^{\infty} \mathbb{P}_x(X_n = y, \tau > n)$$

ein invariantes Maß mit $\mu(x) = 1$ definieren. Mit Fubini gilt:

$$\sum_{y\in S} \mu(y) = \sum_{n=0}^{\infty} \mathbb{P}_x(\tau_x \geq n) = \mathbb{E}_x[\tau_x]$$

Wir wissen 2.7.11: $\pi = c\mu$, also $c = \dfrac{1}{\mathbb{E}_x[\tau_x]}$. □

Definition 2.7.14. *Ein Zustand x heißt null-rekurrent, falls $\mathbb{E}_x[\tau_x] = \infty$ und positiv-rekurrent, falls $\mathbb{E}_x[\tau_x] < \infty$.*

Satz 2.7.15. *Sei p irreduzibel. Dann sind äquivalent:*

1) Es existiert ein Zustand $x \in S$ sodass x positiv rekurrent ist.

2) Es existiert eine invariante Verteilung π.

3) alle Zustände in S sind positiv rekurrent.

Beweis. **1)⇒2)** *Falls x positiv-rekurrent ist, definiert:*

$$\pi(y) = \frac{\sum_{n=0}^{\infty} \mathbb{P}_x(X_n = y, \tau_x > x)}{\mathbb{E}_x[\tau_x]}$$

eine invariante Verteilung.

2)⇒3) *Aus 2.7.13 folgt: $\pi(y) = \dfrac{1}{\mathbb{E}_y[\tau_y]}$. und wegen Irreduzibilität gilt für alle Zustände $y \in S$ mit $\pi(y) > 0$. Daraus folgt bereits: $\mathbb{E}_y[\tau_y] < \infty$.*

3)⇒1) *klar*

□

Beispiel 2.7.16. *Die symmetrische Irrfahrt auf einem Graphen ist genau dann irreduzibel, wenn der Graph zusammenhängend ist. Falls G zusammenhängend ist, gilt für alle Zustände $i \in S$:$\mu(i) \geq 1$. Die Kette ist genau dann positiv rekurrent, wenn $|C| < \infty$ ist.*

2.8. Konvergenz gegen Gleichgewicht

Ziel: Aussagen, die das invariante Maß μ konstruktiv liefern.
Ansatz: Wähle bedingte Startverteilung ν und betrachte Folge $\nu, \nu p, \nu p^2, , \nu p^3, \ldots$
Falls diese Folge konvergent ist, dann konvergiert sie gegen eine invariante Verteilung . Finde geeignete Voraussetzungen unter denen für Wahrscheinlichkeitsmaße ν gilt :

$$\lim_{n\to\infty} \nu p^n(y) = \mu(y)$$

beziehungsweise für $x \in S$:

$$\lim_{n\to\infty} p^n(x,y) = \mu(y)$$

Alternativ: Konvergenz der Cesaro-Mittel:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} p^k(x,y) = \mu(y)$$

beziehungsweise für beliebige Wahrscheinlichkeitsmaße ν:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \nu p^k(y) = \mu(y)$$

Lemma 2.8.1. *(i) Falls für ein gegebenes Wahrscheinlichkeitsmaß ν der Limes*

$$\mu(y) = \lim_{n\to\infty} \nu p^n(y)$$

für alle $y \in S$ existiert, dann definiert μ eine invariante Verteilung.

(ii) Analog, falls $\mu(y) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \nu p^k(y)$ existiert.

Beweis. *(i)*

$$\begin{aligned}\mu p(z) &= \sum_{y\in S} \mu(y)p(y,z) = \sum_{y\in S} \lim_{n\to\infty} \left[\sum_{x\in S} \nu(x)p^n(x,y)\right] p(y,z)\\ &= \lim_{n\to\infty} \sum_{x\in S} \nu(x) \underbrace{\left[\sum_{y\in S} p^n(x,y)p(y,z)\right]}_{p^{n+1}(x,z)}\\ &= \lim_{n\to\infty} \sum_{x\in S} \nu(x)p^{n+1}(x,y) = \mu(z)\end{aligned}$$

$\Rightarrow \mu$ *ist invariant.*

(ii) Analog:

$$\mu p(z) = \sum_{y\in S} \left[\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} p^k(y)\right] p(y,z) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \nu p^{k+1}(z) = \mu(z)$$

□

Sein nun $N_n(y) = \sum_{k=1}^{n} 1_{\{y\}}(X_k)$ die Anzahl der Besuche in y bis zur Zeit n.

Satz 2.8.2. *Falls ein Zustand $y \in S$ rekurrent ist, dann gilt für alle Zustände $x \in S$ $\mathbb{P}$-fast sicher*

(i)

$$\frac{1}{n} N_n(y) \longrightarrow \frac{1}{\mathbb{E}_y[T_y]} 1_{\{T_y<\infty\}} \qquad \textit{für } n \to \infty$$

(ii) Für alle Zustände $x, y \in S$ gilt:

$$\frac{1}{n} \sum_{k=1}^{n} p^k(x,y) \longmapsto \frac{1}{\mathbb{E}_y[T_y]} F(x,y)$$

Wiederholung: $F(x,y) = \mathbb{P}_x(T_y < \infty)$, falls irreduzibel und rekurrent: $F \equiv 1$, falls y transient: $\mathbb{E}_y[T_y] = \infty$
Insbesondere für irreduzibel und rekurrent:

$$\frac{1}{n} \sum_{k=1}^{n} p^k(x,y) \longrightarrow \mu(y) \quad (\forall x \in \Omega) \qquad \textit{invariantes Wahrscheinlichkeitsmaß}$$

Allgemein: Für alle Wahrscheinlichkeitsmaße ν gilt:

$$\frac{1}{n}\sum_{k=1}^{n}\nu p^k(y) \longrightarrow \mu(y)$$

Beweis. *(i) Sei y rekurrent und zunächst als Startpunkt $x = y$ gewählt. Wiederhole:*

$$T_y^k = \inf\left\{n > T_y^{k-1} : X_n = y\right\} \quad T_y^0 = 0, T_y = T_y^1 \qquad \textit{Zeitpunkt des k-ten Besuchs in } y$$

Aus der starken Markov-Eigenschaft (mit $t_y^k = T_y^k - T_y^{k-1} = T_y^1 \circ \theta - T_y^{k-1}$) folgt:
Die Zufallsvariable $t_y^1, t_y^2, t_y^3, \dots$ sind unabhängig und identisch verteilt (unter $\mathbb{P}_y$). Mit dem starken Gesetz der Großen Zahlen folgt:

$$\frac{1}{n}T_y^k = \frac{1}{n}\sum_{k=1}^{n}T_y^k \longrightarrow \mathbb{E}_y\left[T_y\right] \quad \mathbb{P}_y\textit{-fast sicher} \qquad \textit{für } k \to \infty$$

Nun ist:

$$T_y^n(\omega) = \sum_{k=1}^{n}T_y^k = \inf\left\{k \geq 1 : N_n(y) = n\right\} \qquad \textit{Zeitpunkt des n-ten Besuchs in } y$$

In gewisser Weise sind also $n \longmapsto T^n$ und $n \longmapsto N_n$ invers zueinander (für feste y, n).
<u>Genauer:</u>

$$T^{N_n} \leq n < T^{N_n+1} \qquad \frac{T^{N_n}}{N_n} \leq \frac{n}{N_n} \leq \frac{T^{N_{n+1}}}{N_n+1}\frac{N_n+1}{N_n}$$

Also gilt für $n \to \infty$ (wobei $N_n \to \infty$ wegen Rekurrenz von y):

$$\frac{n}{N_n} \longrightarrow \mathbb{E}_y\left[T_y\right] \qquad \mathbb{P}_y\textit{-fast sicher}$$

Sei nun $x \in S$ ein beliebiger Startpunkt.
Auf der Menge $\{T_y = \infty\}$ ist für alle $n \in \mathbb{N}$ $(N_n(y) = 0)$ und damit $\frac{1}{n}N_n(y) = 0$.
Die starke Markov-Eigenschaft impliziert weiterhin, dass die $t_y^2, t_y^3, t_y^4, \dots$ iid sind (unter $\mathbb{P}_x$) mit $\mathbb{P}_x(t_y^k = n) = \mathbb{P}_y(T_y = m)\ \forall k, n \in \mathbb{N}_{\geq 2}$. Damit:

$$\frac{1}{k}T_y^k = \frac{1}{k}T_y^1 + \frac{1}{k}(t_y^2 + \cdots t_y^k) \longrightarrow \begin{cases} 0 + \mathbb{E}_y\left[T_y\right] & \mathbb{P}_x\textit{-fast sicher auf } \{T_y < \infty\} \\ \infty & \textit{auf } \{T_y = \infty\} \end{cases}$$

(ii) Sei y transient. Dann gilt $G(y,y) < \infty$ sowie für alle $x \in S$: $G(x,y) = \sum\limits_{k=1}^{\infty}p^k(x,y) < \infty$.
Daraus folgt:

$$p^n(x,y) \longrightarrow 0 \quad \textit{für } n \to \infty \quad \textit{und} \quad \frac{1}{n}\sum_{k=1}^{n}p^k(x,y) \longrightarrow 0 \quad \textit{für } n \to \infty$$

Also LS=0, ferner RS=0, da $\mathbb{E}_y\left[T_y\right] = \infty$. Sei nun y rekurrent. Wegen $0 \leq \frac{1}{n}N_n(y) \leq 1$ folgt mit majorisierter Konvergenz aus (i):

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}p^k(x,y) = \lim_{n\to\infty}\mathbb{E}_x\left[\frac{1}{n}N_n(y)\right] \overset{\textit{Maj. Krit.}}{=} \mathbb{E}_x\left[\frac{1}{\mathbb{E}_y\left[T_y\right]}1_{\{T_y<\infty\}}\right]$$

$$= \frac{1}{\mathbb{E}_y\left[T_y\right]}\underbrace{\mathbb{P}_x(T_y < \infty)}_{=F(x,y)}$$

<u>Zusatz:</u>

- *In (i) gilt analog für alle Wahrscheinlichkeitsmaße ν auf S:* $\frac{1}{n}N_n \longrightarrow \frac{1}{\mathbb{E}_y\left[T_y\right]}1_{\{T_y<\infty\}}$ $\mathbb{P}_\nu$*-fast sicher*
- *In (ii) gilt für alle Wahrscheinlichkeitsmaße ν auf S:* $\frac{1}{n}\sum\limits_{k=1}^{n}\nu p^k(y) \longrightarrow \mu(y)$ $\mathbb{P}_\nu$*-fast sicher*

$\square$

Korollar 2.8.3. *Version des Ergodensatzes*
Sei nun $(X_n)_{n\in N_0}$ eine irreduzible, positiv rekurrente Markov-Kette mit beliebiger Startverteilung ν und Übergangswahrscheinlichkeit p. Dann gilt für alle beschränkten Funktionen $f : S \longrightarrow \mathbb{R}$

$$\frac{1}{n}\sum_{k=1}^{n} f(X_n) \overset{n\to\infty}{\longrightarrow} \bar{f} \qquad \mathbb{P}_v\text{-fast sicher}$$

wobei $\bar{f} = \sum\limits_{x\in S}\mu(x)f(x)$ mit eindeutiger invarianter Verteilung μ ist. („Mittel über Zeit"=„Mittel über Raum")

Beweis. *Sei o,B.d.A $|f| \leq 1$. Sei μ die eindeutige invariante Verteilung. Zu gegebenen $\epsilon > 0$ existiert ein* <u>endliches</u> *$S_0 \subset S$ mit $\mu(S\setminus S_0) \leq \frac{\epsilon}{4}$. Wir wissen aus dem Satz 2.8.2:*

$$\frac{1}{n}N_n(y) \overset{n\to\infty}{\longrightarrow} \frac{1}{\mathbb{E}_y\left[T_y\right]}1_{\{T_y<\infty\}} = \mu(y) \qquad \mathbb{P}_\nu\text{-fast sicher}$$

Also existiert $\bar{n} = \bar{n}(n,\epsilon) \in \mathbb{N}$ sodass für alle $n \geq \bar{n}$:

$$\sum_{y\in S_0} |\frac{1}{n}N(y)\mu(y)| \leq \frac{\epsilon}{4}$$

Damit gilt $\mathbb{P}_\nu$-fast sicher:

$$|\frac{1}{n}\sum_{k=1}^{n} f(X_n) - \bar{f}| = |\sum_{y\in S}(\frac{1}{n}N_n(y) - \mu(y)f(y)| \leq |\sum_{y\in S_0}|\frac{1}{n}N_n(y) - \mu(y)| + \sum_{y\notin S_0}\underbrace{|\frac{1}{n}N_n(y) - \mu(y)|}_{\leq\frac{1}{n}N_n(y)+\mu(y)}$$

$$\leq \sum_{y\in S_0}|\frac{1}{n}N_n(y) - \mu(y)| + \underbrace{\sum_{y\notin S_0}(\frac{1}{n}N_n(y) - \mu(y))}_{=-\sum\limits_{y\in S_0}(\frac{1}{n}N_n(y)-\mu(y))} + 2\sum_{y\notin S_0}\mu(y)$$

$$\leq 2\sum_{y\in S_0}|\frac{1}{n}N(y) - \mu(y)| + 2\sum_{y\notin S_0}\mu(y) \leq 2\frac{\epsilon}{4} + 2\frac{\epsilon}{4} \leq \epsilon$$

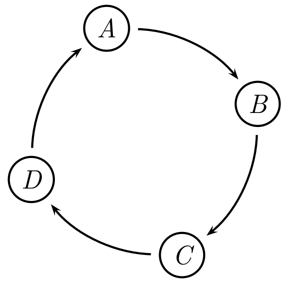
$\square$

Beispiel 2.8.4. $S = \{a, b\}$,

$$p = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \Rightarrow \quad p^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = E$$

$$\Rightarrow p^{2n} = E, \quad p^{2n+1} = p \qquad \Rightarrow \lim_{n\to\infty} p^n(x,y) \text{ existiert nicht!}$$

<u>*Analog:*</u>



Definition 2.8.5. *(i) Sei $x \in S$ ein rekurrenter Zustand. Die Periode von x d_x ist der ggT der Menge $I_x = \{n \in \mathbb{N} : p^n(x,x) > 0\}$.*

(ii) Die Markov-Kette heißt aperiodisch , falls für alle $x \in S$ gilt: $d_x = 1$.

Lemma 2.8.6. *Sei $x \in S$ ein rekurrenter Zustand und $y \in S$ ein Zustand mit $F(x,y) > 0$. Dann gilt:*

$$d_y = d_x$$

Die Periode ist also eine Klasseneigenschaft.

Beweis. Wegen der Rekurrenz von x gilt dann auch $F(y,x) > 0$. Es existieren daher $k, l \in \mathbb{N}$ mit $p^k(x,y) > 0$ und $p^l(y,x) > 0$. Daraus folgt:

$$p^{l+k}(y,y) \geq p^l(y,x)p^k(x,y) > 0$$

Also ist d_y Teiler von $k+l$. Sei nun $n \in I_x$ (also $p^n(x,x) = 1$) daraus folgt:

$$p^{l+n+k}(y,y) \geq p^l(y,x)p^n(x,x)p^k(x,y)$$

Dann ist d_y ein Teiler von $l+n+k$. Da d_y auch $k+l$ teilt, ist es ein Teiler von n. Da d_y der ggT aller Elemente in I_x ist, ist d_y ein Teiler von d_x. Da der Zustand y auch rekurrent ist, lassen sich die Rollen von x und y vertauschen um zu zeigen, dass auch d_x ein Teiler von d_y ist. Also ist $d_x = d_y$. $\square$

Lemma 2.8.7. *Sei $x \in S$ ein rekurrenter Zustand mit $d_x = 1$. Dann existiert ein $m_0 \in \mathbb{N}$ sodass gilt:*

$$\{n \in \mathbb{N}, n \geq m_0\} \subset I_x$$

Satz 2.8.8. *Sei p irreduzibel und aperiodisch mit invarianter Verteilung μ. Dann gilt für alle Startverteilungen ν und alle Zustände $y \in S$:*

$$\nu p^n(y) \longrightarrow \mu(y) \qquad \text{für } n \to \infty \tag{2.21}$$

Es gilt sogar in totaler Variationsnorm ($||\nu - \mu||_{tv} := \sum_{y \in S} |\nu(y) - \mu(y)|$):

$$||\nu p^n - \mu||_{tv} \longrightarrow 0 \qquad \text{für } n \to \infty$$

Beweis. Genereller Plan: Betrachte zwei Markov-Ketten mit Übergangswahrscheinlichkeit p:

$(X_n)_{n \in N_0}$ mit Startverteilung ν, $(Y_n)_{n \in \mathbb{N}_0}$ mit Startverteilung μ auf dem selben Wahrscheinlichkeitsraum

Zeige: $(X_n)_{n \in N_0}$ und $(Y_n)_{n \in \mathbb{N}_0}$ treffen sich fast sicher.

$$T = \inf\{n \geq 0, X_n = Y_n\}$$

Ersetze X nach T durch Y und zeige, dass das neue X in Verteilung gegen μ konvergiert.
Im Detail: Setze $\bar{S} = S \times S$ und den Markov-Kern $\bar{p}$ mit:

$$\bar{p}((x_0,y_0),(x_1,y_1)) = \begin{cases} p(x_0,x_1)p(y_0,y_1) & x_0 \neq y_0 \\ p(x_0,x_1)\delta_{x_1}(y_1) & x_0 = y_0 \end{cases}$$

Also:

- *Bis zu ihrem Zusammentreffen sind* $(X_n)_{n\in N_0}$ *und* $(Y_n)_{n\in\mathbb{N}_0}$ *unabhängiger Koordinatenprozesse.*
- *Danach werden sie „verklebt“.*
- *Jeder Koordinatenprozess hat Übergangswahrscheinlichkeit p.*

Sei $(Z_n)_{n\in\mathbb{N}_0} = (X_n, Y_n)_{n\in\mathbb{N}_0}$ *eine Markov-Kette auf* $\bar{S}$ *zu* $\bar{p}$ *mit Startverteilung* $\bar{\nu} = \nu \otimes \mu$. *Betrachte die Stoppzeit* T *mit:*

$$T = \inf\{n \geq 0 : X_n = Y_n\} = \inf\left\{n \geq 0 : Z_n \in diag(\bar{S})\right\}$$

Aufgrund der Konstruktion gilt $X_n = Y_n$ *für alle* $n \geq T$.

$$|\mathbb{P}(X_n = y) - \mathbb{P}(Y_n = y)| = |\mathbb{P}(X_y = y, n \geq T| + |\mathbb{P}(X_n = y, n < T) - \mathbb{P}(Y_n = y, n \geq T) - \mathbb{P}(Y_n = y, n < T)|$$

$$\leq \mathbb{P}(X_n = y, n \geq T) + \mathbb{P}(Y_n = y, n < T)$$

Daraus folgt:

$$\sum_{y\in S} |\mathbb{P}(X_n = y) - \mathbb{P}(Y_n = y)| \leq 2\mathbb{P}(n < T)$$

Falls $\underline{\mathbb{P}(T < \infty) = 1}$ *dann gilt:*

$$\sum_{y\in S} |\nu p^n(y) - \mu(y)| = \sum_{y\in S} |\mathbb{P}(X_n = y) - \mathbb{P}(Y_n = y)| \overset{n\to\infty}{\longrightarrow} 0$$

Es bleibt zu zeigen, dass fast sicher $T < \infty$ *gilt (d.h. „die Kopplung war erfolgreich“). Um dies zu zeigen, brauchen wir* X_n, Y_n *nur bis zum ersten Treffen zu kennen. Hierzu betrachten wir eine leicht modifizierte Kette:*

$$\tilde{p}((X_0, y_0), (x_1, y_1)) = p(x_0, x_1)p(y_0, y_1), \quad \tilde{\nu} = \nu \otimes \mu, \quad \tilde{Z}_n = (\tilde{X}_n, Y_n), \quad \tilde{T} = \inf\left\{n \geq 0 : \tilde{X}_n = \tilde{Y}_n\right\}$$

Behauptung: *Es gilt fast sicher* $\tilde{T} < \infty$.
Dafür genügt es zu zeigen, dass $\tilde{T}_{(x,x)} < \infty$ *für einen Zustand* $x \in S$ *ist. Wir wissen, dass* $\tilde{p}$ *eine invariante Verteilung* $\tilde{\mu} = \mu \otimes \mu$ *besitzt. Ferner ist* $\tilde{p}$ *irreduzibel, denn es gilt:*

$$\forall x_0, x_1, y_0, y_1 \in S \;\exists\, k, l \in \mathbb{N} : \quad p^k(x_1, x_1) > 0,\ p^l(y_0, y_1) > 0$$

Mit Lemma 2.8.7 folgt aus der Aperiodizität:

$$\exists m \in \mathbb{N} : \ p^{m+l}(x_0, x_1) > 0, \qquad p^{m+k}(y_1, y_1) > 0$$

Somit ist auch $\tilde{p}^{k+l+n}((x_0, x_1), (y_0, y_1)) > 0$, *also ist auch* $\tilde{p}$ *irreduzibel. Da* $\tilde{p}$ *irreduzibel ist und eine invariante Verteilung* $\tilde{\mu}$ *existiert, ist* $\tilde{p}$ *rekurrent.*

$$\Rightarrow \forall x \in S : \ \tilde{T}_{(x,x)} < \infty \text{ fast sicher} \quad \Rightarrow \tilde{T} < \infty \text{ fast sicher} \quad \Rightarrow T < \infty \text{ fast sicher}$$

□

Bemerkung 2.8.9. Ausblick:

1. *Für periodische Markov-Ketten mit der Periode d gilt:*

$$S = \bigcup_{r=0}^{d-1} S_r$$

Für gegebenes $x \in S_0$ *gilt:*

$$\lim_{n\to\infty} p^{nd+r}(x, y) = d\mu(y) \qquad (\textit{Übungen})$$

2. *Konvergenzgeschwindigkeit in 2.21*
Es gilt $||\nu p^n - \mu|| \leq Cr^n$ *für eine Konstante* $C \in \mathbb{R}_+$ *und* $r < 1$.

Reversible Markov-Ketten

Gegeben sei ein beliebiger Zustandsraum S, ein Markov-Kern $p(x, dy)$ und ein Maß μ.

Definition 2.8.10. *(i) Ein Markov-Kern $\hat{p}$ aus S heißt dual zu p (bzgl. μ) gilt:*

$$p(x, dy)\mu(dx) = \hat{p}(y, dx)\mu(dy) \qquad \text{als Maße auf } S \times S \tag{2.22}$$

Das heißt für alle $A, B \in \mathcal{S}$ gilt:

$$\int_A p(x, B)\mu(dx) = \int_{A\times B} p(x, dy)\mu(dx) = \int_{A\times B} \hat{p}(y, dx)\mu(dy) = \int_B \hat{p}(y, A)\mu(dy)$$

(ii) Der Markov-Kern p heißt reversibel (bzgl. μ) oder μ heißt reversibles Maß für p, falls p dual zu sich selbst ist.

Bemerkung 2.8.11. *Offenbar:*

a) Aus der Existenz eines dualen $\hat{p}$ folgt, dass μ invariant ist. Mit $A = S$ folgt für alle $B \in S$:

$$\int_S p(x, B)\mu(dx) = \mu(B)$$

b) Für diskreten Zustandsraum S heißt 2.22:

$$\mu(x)p(x, y) = \mu(y)\hat{p}(y, x) \qquad (\forall x, y \in S)$$

p ist reversibel, falls gilt:

$$\boxed{\mu(x)p(x, y) = \mu(y)p(y, x) \quad (\forall x, y \in S)} \qquad \text{„detaillierte Balance“}$$

Hierbei gilt o.B.d.A $\mu > 0$. Ansonsten betrachte Einschränkung auf $S_0 = \{x : \mu(x) > 0\}$, invariante Teilmenge für p (d.h. $\forall x \in S_0, y \notin S_0 :\ p(x, y) = 0$).

$$0 = \mu(y) = \sum_{x\in S_0} \underbrace{\mu(x)}_{>0} p(x, y)$$

Unter der Voraussetzung $\mu > 0$ gilt: Zu p existiert genau ein duales $\hat{p}$:

$$\hat{p}(y, x) = \frac{\mu(x)}{\mu(y)} p(x, y)$$

Falls μ konstant auf S ist (und S endlich), ist $\hat{p}$ die adjungierte Matrix zu p. Dann ist p genau dann reversibel, wenn es eine symmetrische Matrix ist.

Beispiel 2.8.12. *(i) Gegeben sei eine abzählbare Menge V („vertices“,„Knoten“) und eine symmetrische Funktion $c : V \times V \longrightarrow \mathbb{R}_+$ („Kantengewichte“) gegeben durch:*

$$c(x) = \sum_{y\in V} c(x, y) \qquad S = \{x \in V : c(x) > 0\}$$

Wähle $\mu(x) := c(x)$ und $p(x, y) = \dfrac{c(x, y)}{c(x)}$. Dann ist p ein reversibler Markov-Kern bezüglich μ. Denn:

$$\sum_{y\in S} p(x, dy) = \frac{1}{c(x)} \sum_{y\in S} c(x, y) = 1, \qquad \mu(x)p(x, y) = c(x, y) \quad \text{symmetrisch in } (x, y)$$

Setze $E = \{(x, y) \in V \times V.\ c(x, y) > 0\}$ („Knoten“). Dann ist (V, E, C) ein gewichteter Graph.

(ii) $S = \mathbb{R}^d$, $p(x,dy) = f(x-y)dy$, $\mu(dy) = dy$. *Für festes* $t > 0$ *sei* f *gegeben durch:*

$$f(x) = \frac{1}{(2\pi t)^{\frac{d}{2}}} \exp\left(-\frac{|x|^2}{2t}\right)$$

Dann ist p *ein Markov-Kern*

$$\int\limits_{\mathbb{R}^d} p(x,dy) = \int\limits_{\mathbb{R}^d} f(y)dy = 1$$

und reversibel, da $(p(x,dy) = f(x-y)dydx)$ *symmetrisch in* (x,y) *ist.*

Lemma 2.8.13. *Sei* p *ein Markov-Kern, der reversibel bezüglich* μ *ist. Dann definiert*

$$Pu(x) = \int\limits_S u(y)p(x,dy)$$

für $u \in \mathcal{L}^1(S,\mu) \cap L^\infty(S)$ *einen beschränkten, linearen und symmetrischen Operator auf* $L^2(S,\mu)$ *mit Norm* ≤ 1.

Beweis. **Linearität** *klar.*

Beschränktheit *folgt aus Norm-Abschätzung.*

Symmetrie *Für alle* $u, v \in \mathcal{L}^1(S,\mu) \cap L^\infty(S)$ *gilt:*

$$< v, Pu >_{L^2} = \int\limits_S v(x)Pu(x)\mu(dx) = \int\limits_S\int\limits_S v(x)u(y)\underbrace{p(x,dy)\mu(dx)}_{sym.\ in\ (x,y)} = \int\limits_S\int\limits_S v(y)u(x)p(x,dy)\mu(dx) = < u, Pv >_{L^2}$$

Normabschätzung *Für alle* $u, v \in \mathcal{L}^1(S,\mu) \cap L^\infty(S)$ *gilt:*

$$\begin{aligned} |< v, Pu >_{L_2}|^2 &= \left|\int\limits_S\int\limits_S v(x)u(y)p(x,dy)\mu(dx)\right|^2 \overset{CSU}{\leq} \left|\int\limits_S\int\limits_S v^2(x)p(x,dy)\mu(dx)\right| \cdot \left|\int\limits_S\int\limits_S u^2(y)p(x,dy)\mu(dx)\right| \\ &= ||v||_{L^2}^2 \cdot ||u||_{L^2}^2 \end{aligned}$$

Das heißt:

$$| < v, Pu > | \leq ||v|| \cdot ||u|| \qquad mit\ v = Pu: \ ||Pu|| \leq ||u||$$

$\square$

Bemerkung 2.8.14. *Sei* S *ein endlicher Zustandsraum, also o.B.d.A.* $S = \{1, \ldots, N\}$ *und* p *reversibel bezüglich* μ *mit* $\mu > 0$. *Dann gilt:*

(i) $q(x,y) = \sqrt{\frac{\mu(x)}{\mu(y)}}p(x,y)$ *ist symmetrische Matrix.*
Dann existieren reelle Eigenwerte $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$, *denn:*

$$q(x,y) = \sqrt{\frac{\mu(x)}{\mu(y)}}p(x,y) = \sqrt{\frac{\mu(y)}{\mu(x)}}p(y,x) = q(y,x) \qquad \Longleftrightarrow \qquad \mu(x)p(x,y) = \mu(y)p(y,x)$$

(ii) Sei u *eine* q*-Eigenfunktion zum Eigenwert* λ. *Dann ist* $\hat{u}(y) = \frac{1}{\sqrt{\mu(y)}}u(y)$ p*-Eigenfunktion (bzgl. Rechtsmultiplikation) zum Eigenwert* λ *und* $\check{u}(x) = \sqrt{\mu(x)}u(x)$ *eine* p*-Eigenfunktion (bzgl. Linksmultiplikation) zum Eigenwert* λ.
In der Tat:

$$p\hat{u}(x) = \sum_{y\in S} p(x,y)\frac{1}{\sqrt{\mu(y)}}u(y) = \frac{1}{\sqrt{\mu(x)}}\lambda u(x) = \lambda\hat{u}(x)$$

Analog: $\check{u}p(y) = \cdots = \lambda\check{u}(y)$

(iii) P hat reelle Eigenwerte $\lambda_1 \geq \cdots \geq \lambda_N$ mit $|\lambda_k| \leq 1$ $(\forall k \in \{1, \ldots, N\})$, da $||Pu||_{L_2} \leq ||u||_{L_2}$ gilt. Wegen $\mu p = \mu$ ist $\lambda_1 = 1$ Eigenwert zum Eigenvektor μ.

Proposition 2.8.15. *Sei S ein beliebiger Zustandsraum und p ein bezüglich μ reversibler Markov-Kern. Dann ist $\{\psi_n\}_{1 \leq n \leq N}$ eine vollständige Orthonormalbasis von $L^2(S, \mu)$ bestehend aus Eigenfunktionen zu p $\underbrace{(p\psi_k = \lambda_k \psi_k)}$ mit $(\lambda_1 = 1)$ und $(\psi_1(x) = 1)$, denn:*

$$p\psi_1(x) = \int_S \underbrace{\psi_1(y)}_{=1} p(x, dy) = 1$$

<u>Annahme:</u> $\lambda_* = \sup\limits_{k \neq 1} |\lambda_l|$. Dann gilt für alle $u \in L^2(S, \mu)$:

$$||p^n u - \bar{u}||_{L^2} \leq \lambda_*^n ||u - \bar{u}||_{L^2} \longrightarrow 0 \quad \text{für } n \to \infty$$

mit $\bar{u} =< u, \psi_1 >= \int\limits_S u(y)\mu(dy)$.
<u>Bemerkung:</u> Vollständige Orthonormalbasis heißt:

$$\forall u \in L^2: \quad \sum_{k=1}^{\infty} < u, \psi_k >_{L^2} \psi_k \quad \text{und} \quad ||u||^2 = \sum_{k=1}^{\infty} < u, \psi_k >^2$$

Beweis. *Setzte $u_0 = u - \bar{u} = \sum\limits_{k=2}^{\infty} < u, \psi_k > \psi_k$, dann gilt:*

$$p^n u_0 = \sum_{k=2}^{\infty} < u, \psi_k > p^n \psi_k = \sum_{k=2}^{\infty} < u, \psi_k > \lambda_k^n \psi_k$$

Daraus folgt:

$$||p^n u - \bar{u}||^2 = \sum_{k=2}^{\infty} < u, \psi_k >^2 \lambda_k^{2k} \leq \lambda_*^{2k} \sum_{k=2}^{\infty} < u, \psi_k >^2 = \lambda_*^{2k} ||u - \bar{u}||^2$$

□

Beispiel 2.8.16. *$S = [0, 2\pi R]$, $\mu(dx) = \frac{1}{2\pi R} dx$ (normiertes Lebesque-Maß), $f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$, $F(x) = \sum\limits_{k \in \mathbb{Z}} f(x + Rk2\pi)$. Dann ist p mit*

$$p(x, dy) = F(x - y)dy$$

ein Markov-Kern auf S, da

$$\int_S p(x, dy) = \int_0^{2\pi} \sum_{k \in \mathbb{Z}} f(x - y - k2\pi R)dy = \int_{\mathbb{R}} f(x - y)dy = 1$$

und reversibel bezüglich μ.
Sei $\psi_k(x) = \exp\left(\frac{ikx}{R}\right)$ für $k \in \mathbb{Z}$ und $x \in S$. Dies ist eine vollständige Orthonormalbasis von $L^2_{\mathbb{C}}(S, \mu)$. Die Funktionen ψ_k, $k \in \mathbb{Z}$ sind Eigenfunktionen zu p mit Eigenwert $\lambda_k = \exp\left(-\frac{k^2}{2}\right)$, denn

$$p\psi_k(x) = \int_S \psi_k(y) F(x - y)dy = \int_{\mathbb{R}} \psi_k(y) f(x - y)dy = \int_{\mathbb{R}} \psi_k(x + y) f(y)dy$$

$$= \exp\left(\frac{ikx}{R}\right) \int_{\mathbb{R}} \underbrace{\exp\left(\frac{iky}{R}\right) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2R^2}\right) dy}_{\text{Fourier Transformation von } f} = \exp(ikx) \exp\left(-\frac{k^2}{2R^2}\right) = \lambda_k \psi_k(x)$$

Hierbei ist:

$$\lambda_0 = 1 \quad \psi_1 = 1 \quad \lambda_* = \sup_{k \neq 0} |\lambda_k| = e^{-\frac{1}{2R^2}} \overset{\mathbb{R}\to\infty}{\longrightarrow} 1$$

Daher gilt für alle $u \in L^2(S, \mu)$*:*

$$||p^n u - \bar{u}||_{L^2} \leq e^{-\frac{n}{2R^2}} ||u - \bar{u}||_{L_2}$$

In dualer Darstellung: Betrachte für alle Anfangsverteilungen ν *mit* $\delta = \dfrac{d\mu}{d\nu}$ *auf* S *die Markov-Kette* $(X_n)_{n \in N_0}$ *mit Wahrscheinlichkeitsmaß* $\mathbb{P}_\nu$*. Die Verteilung zur Zeit* n *ist gegeben durch:*

$$\nu_n = (X_n)_* \mathbb{P}_\nu = \mathbb{P}_\nu(X_n \in \cdot)$$

. Dann gilt:

$$\frac{d\nu_n}{dn} = \delta_n, \quad ||\delta_n - 1||_{L^2} \leq e^{-\frac{n}{2}} \cdot C$$

Beispiel:
$S = R$, $p(x, dy) = f(x-y)dy$, $\mu(dy) = dy \quad \Rightarrow$ *keine exponentielle Konvergenz gegen das Gleichgewicht.*

Für andere Kriterien, die exponentielles Wachstum verifizieren, betrachten wir die Energie.

Definition 2.8.17. *Für alle* $u, v \in L^2(S, \mu)$ *ist die Energie* $\epsilon(u, v)$ *gegeben durch:*

$$\epsilon(u, v) = -< Au, v >_{L^2} = \int_S u(x)v(x)\mu(dx) - \int_S u(y)v(x)p(x, dy)\mu(dx)$$

Mit $A = P - I$ *gilt:*

$$\epsilon(u, v) = \frac{1}{2} \int_S \int_S [u(x) - u(y)]^2 p(x, dy)\mu(dx) \qquad \epsilon(u, u) = \frac{1}{2} \int_S \int_S [u(x) - u(y)]^2 p(x, dy)\mu(dx)$$

Definition 2.8.18. *Sei* μ *ein reversibles Maß. Definiere* δ *durch:*

$$\delta := \inf_{\substack{u \in L^2, u \neq 0 \\ \int u d\mu = 0}} \frac{\epsilon(u, u)}{||u||^2_{L^2}}$$

Für diskreten Zustandsraum S *definiere:*

$$\delta := \inf_{\substack{u \\ \sum u(i)\mu(i) = 0}} \frac{\frac{1}{2} \sum\limits_{i,j \in S} |u(i) - u(j)|^2 p(i, j)\mu(i)\mu(j)}{\sum\limits_{i \in S} |u(i)|^2 \mu(i)}$$

Satz 2.8.19. *Sei* μ *ein reversibles Wahrscheinlichkeitsmaß und* $\delta > 0$*. Dann gilt für alle* $u \in L^2$*:*

$$||p^n u - \bar{u}||_{L^2} \leq (1 - \delta)^n ||u - \bar{u}||_{L^2} \longrightarrow 0 \quad \text{für } n \to \infty$$

Beweis. *Für alle* $u \in L^2$ *mit* $< u, 1 >= 0$ *gilt:*

$$< u, u > - < u, Pu >= \epsilon(u, u) \geq \delta ||u||^2$$

Also gilt für geeignetes $P^{\frac{1}{2}}$ *mit* $P^{\frac{1}{2}} \circ P^{\frac{1}{2}} = P$*:*

$$< u, Pu >=< P^{\frac{1}{2}} u, P^{\frac{1}{2}} u > \leq (1 - \delta) ||u||^2$$

und:

$$||P^{\frac{1}{2}} u|| \leq (1 - \delta)^{\frac{1}{2}} ||u|| \quad \Rightarrow \quad ||Pu|| \leq (1 - \delta)||u||, \ ||P^n u|| \leq (1 - \delta)^n ||u||$$

Für alle $u \in L^2$ *gilt:*

$$u = u_0 - \bar{u}, \quad \bar{u} =< u, 1 >, \quad < u_0, 1 >= 0$$

□

Kapitel 3

Brown'sche Bewegung

3.1. Definition und Eigenschaften

Definition 3.1.1. *Ein stochastischer Prozess ist eine Familie $(X_t)_{t\geq 0}$ von Zufallsvariablen $X_i : \Omega \longrightarrow \mathbb{R}$, definiert auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$.*
Allgemein kann man einen stochastischen Prozess $(X_t)_{t\in I}$ für eine beliebige Indexmenge I betrachten mit Zufallsvariablen $X_i : \Omega \longrightarrow S$ und $(S, \mathcal{S})$ einem beliebigen Messraum. Typischerweise wird $t \in I$ als Zeit und $x \in S = \mathbb{R}$ (oder $S = \mathbb{R}^d$) als Ort interpretiert.

Im Folgenden sei $I \subset \mathbb{R}_+$ ($T = \mathbb{R}_+$ oder $I = [0,1]$) und der Zustandsraum $S = \mathbb{R}$ (oder $S = \mathbb{R}^d$). Das X_t eine Zufallsvariable ist bedeutet, dass die Abbildung $\omega \longmapsto X_t(\omega)$ messbar ist. Im Allgemeinen gibt es keine Restriktion für die Trajektorie $t \longmapsto X_t(\omega)$.
Jeder stochastische Prozess lässt sich alternativ beschreiben als Abbildung

$$X. : I \times \Omega \longrightarrow S \qquad (t, \omega) \longrightarrow X_t(\omega)$$

oder als Abbildung:

$$X_i : \Omega \longrightarrow S^I \qquad \omega \longmapsto (t \mapsto X_t(\omega))$$

Definition 3.1.2. *Ein stochastischer Prozess $(X_t)_{t\geq 0}$ mit Werten in $\mathbb{R}$ definiert auf einem beliebigen Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ heißt Brown'sche Bewegung, falls gilt:*

(i) $(X_t)_{t\geq 0}$ hat unabhängige Zuwächse, das heißt für alle $0 \leq t_0 < t_1 < \cdots < t_n$ gilt:

$$X_{t_0}, X_{t_1} - X_{t_0}, \ldots, X_{t_n} - X_{t_{n-1}} \qquad \textit{sind unabhängig.}$$

(ii) Die Zuwächse sind Gauß-verteilt:

$$(X_t - X_s)_* \mathbb{P} = N(0, t - s)$$

(iii) Für $\mathbb{P}$-fast alle ω ist $t \longmapsto X_t(\omega)$ stetig

$(X_t)_{t\geq 0}$ heißt standardisierte Brown'sche Bewegung, falls $X_0 = 0$ ist. Für jede Brown'sche Bewegung $(X_t)_{t\geq 0}$ ist $((\tilde{X}_t)_{t\geq 0} = X_t - X_0)$ eine standardisierte Brown'sche Bewegung.

Bemerkung 3.1.3. *Die Brown'sche Bewegung wird über ihre Eigenschaften und nicht über eine Konstruktion definiert. Wir werden sehen, dass die Brown'sche Bewegung existiert und dass sie sich auf mindestens drei verschiedene Arten konstruieren lässt.*

Lemma 3.1.4. *Sei $(X_t)_{t\geq 0}$ eine standardisierte Brown'sche Bewegung . Dann existiert für jedes $\alpha > 0$ eine Konstante $C(\alpha)$, sodass gilt:*

$$\mathbb{E}\left[|X_t|^{2\alpha}\right] = C(\alpha) t^{\alpha}$$

Allgemein gilt für jede Brown'sche Bewegung:

$$\mathbb{E}\left[|X_t - X_s|^{2\alpha}\right] = C(\alpha) |t - s|^{\alpha}$$

Beweis. Jede Zufallsvariable X_t ist $N(0,t)$-verteilt. Daraus folgt:

$$\mathbb{E}\left[|X_t|^{2\alpha}\right] = t^\alpha \int\limits_{\mathbb{R}} \frac{|x|^{2\alpha}}{t^\alpha} \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{x^2}{2t}\right) dx \overset{\frac{x}{\sqrt{t}}=z}{=} t^\alpha \underbrace{\int\limits_{\mathbb{R}} |z|^{2\alpha} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz}_{=:c(\alpha)} = t^\alpha C(\alpha)$$

$\square$

Lemma 3.1.5. *Für jede $Brown'scheBewegung$ gilt:*

(i) $\mathbb{E}[X_t] = 0$, $\mathbb{E}\left[X_t^2\right] = t$

(ii) $\mathbb{E}[X_s, X_t] = s \wedge t := \min\{s,t\}$

Beweis. *(i)* $X_t - X_0$ *ist* $N(0,t)$*-verteilt.*

(ii) Für $s \leq t$ gilt:

$$\mathbb{E}[X_s X_t] = \mathbb{E}[X_s X_s] + \mathbb{E}[X_s(X_t - X_s)] \overset{(i)}{=} s + \mathbb{E}[X_s] \underbrace{\mathbb{E}[X_t - X_s]}_{=0} = s$$

$\square$

Proposition 3.1.6. *Sei $(X_t)_{t\geq 0}$ eine Brown'sche Bewegung auf $(\Omega, \mathcal{A}, \mathbb{P})$. Dann sind auch die folgenden Prozesse Brown'sche Bewegungen.*

(i) $(X_{s+t})_{t\geq 0}$ *für ein $s > o$ („Shift").*

(ii) $(X_{s+t} - X_s)_{t\geq 0}$ *für ein $s > 0$. (sogar standardisierte Brown'sche Bewegung)*

(iii) $(-X_t)_{t\geq 0}$

(iv) $(cX_{\frac{t}{c^2}})_{t\geq 0}$ *für ein $c > 0$ („Scherung")*

(v) $(\tilde{X}_t)_{t\geq 0}$ *mit $X_0 = 0$ und $\tilde{X}_t = tX_{\frac{1}{t}}$ für $t > 0$ („Inversion")*

Beweis. **(i)-(iii)** *Stetigkeit der Trajektorie, Unabhängigkeit der Zuwächse und Verteilung der Zuwächse klar*

(iv) *Stetigkeit, Unabhängigkeit klar*
Verteilung der Zuwächse $\bar{X}_t = cX_{\frac{t}{c^2}}$:

$$(\bar{X}_t - \bar{X}_s)_*\mathbb{P}(B) = \mathbb{P}(\bar{X}_t - \bar{X}_s \in B) = \mathbb{P}\left(X_{\frac{t}{c^2}} - X_{\frac{s}{c^2}} \in \frac{1}{c}B\right) = N\left(0, \frac{t+s}{c^2}\right)\left(\frac{1}{c}B\right)$$

$$= \frac{1}{\sqrt{2\pi\frac{t-s}{c}}} \int\limits_B \exp\left(-\frac{-x^2c^2}{2(t-s)}\right) dxc = \frac{1}{\sqrt{2\pi(t-s)}} \int\limits_B \exp\left(-\frac{x^2}{2(t-s)}\right) dx = N(0, t-s)(B)$$

(v) *Zeige zunächst nur:*

- *Stetigkeit von $t \longmapsto \tilde{X}_t(\omega)$ für $t > 0$*
- *Unabhängigkeit der Zuwächse.*

Statt der Verteilung der Zuwächse bestimmen wir $\mathbb{E}\left[\tilde{X}_s\tilde{X}_t\right]$ für $0 < s \leq t$:

$$\mathbb{E}\left[\tilde{X}_s\tilde{X}_t\right] = st\mathbb{E}\left[X_{\frac{1}{s}}X_{\frac{1}{t}}\right] = st\frac{1}{t} = s = \min\{s,t\}.$$

Mit Satz 3.3.4 folgt, dass $(\tilde{X}_t)_{t\geq 0}$ eine Brown'sche Bewegung ist.

$\square$

Exkurs:

Definition 3.1.7. *(i) Zwei stochastische Prozesse* $(X_t)_{t\in I}$ *auf* $(\Omega,\mathcal{A},\mathbb{P})$ *bzw.* $(\tilde{X}_t)_{t\in I}$ *auf* $(\tilde{\Omega},\tilde{\mathcal{A}},\tilde{\mathbb{P}})$ *heißen äquivalent, falls für alle endlichen* $J\subset I$ *gilt:*

$$p_J = \tilde{p}_J \qquad \text{auf } S^J$$

Hierbei ist $\mathbb{P}_J = [(X_t)_{t\in J}]_* \mathbb{P}$*, das heißt für* $J=(t_1,\dots t_n)$ *und* $B = B_1\times\dots\times B_n \subset S^n = S^J$ *gilt:*

$$\mathbb{P}_J(B) = \mathbb{P}(X_{t_1}\in B_1,\dots,X_{t_n}\in B_n) \qquad \underline{\text{endlich-dimensionale Verteilungen}}$$

(ii) Die Prozesse $(X_t)_{t\geq 0}$ *und* $(\tilde{X}_t)_{t\geq 0}$ *heißen Modifikationen voneinander, falls:*

$$(\Omega,\mathcal{A},\mathbb{P}) = (\tilde{\Omega},\tilde{\mathcal{A}},\tilde{\mathbb{P}}) \quad \text{und} \quad \mathbb{P}(X_t=\tilde{X}_t)=1\ (\forall t\in I)$$

(iii) Die Prozesse $(X_t)_{t\geq 0}$ *und* $(\tilde{X}_t)_{t\geq 0}$ *heißen ununterscheidbar (indistinguishable), falls:*

$$(\Omega,\mathcal{A},\mathbb{P}) = (\tilde{\Omega},\tilde{\mathcal{A}},\tilde{\mathbb{P}}) \quad \text{und} \quad \mathbb{P}(\forall t\in I: X_t=\tilde{X}_t)=1$$

Bemerkung 3.1.8. *Es gilt: (iii)* $\Rightarrow$ *(ii)* $\Rightarrow$ *(i), und falls I abzählbar ist auch (ii)* $\Rightarrow$ *(iii). Im Allgemeinen ist (ii)* $\not\Rightarrow$ *(iii).*

Beispiel 3.1.9. $\Omega=\mathbb{R}$, $\mathbb{P}=N(0,1)$, $I=\mathbb{R}_+$, $S=\mathbb{R}$.

$$X_t(\omega) = \begin{cases} 0 & t\neq|\omega| \\ 1 & t=|\omega| \end{cases}$$

Dann ist $\mathbb{P}$*-fast sicher jede Trajektorie unstetig.*

Proposition 3.1.10. *Seien* $(X_t)_{t\in\mathbb{R}_+}$ *und* $(\tilde{X}_t)_{t\in\mathbb{R}_+}$ *stochastische Prozesse mit Werten in* $\mathbb{R}$ *und fast sicher stetigen Pfaden. Dann sind* X *und* $\tilde{X}$ *genau dann Modifikationen voneinander, wenn sie ununterscheidbar sind.*

Beweis. „$\Leftarrow$" *stets*

„$\Rightarrow$" *Da* X *und* $\tilde{X}$ *Modifikationen voneinander sind gilt:*

$$\mathbb{P}(\forall t\in\mathbb{Q}_+:\ X_t=\tilde{X}_t)=1, \quad \text{und} \quad \mathbb{P}(t\to X_t,\ t\to\tilde{X}_t \text{ stetig})=1$$

$$\Rightarrow \mathbb{P}(\forall f\in\mathbb{R}_+:\ X_t=\tilde{X}_t)=1$$

□

Proposition 3.1.11. *Sei* $(X_t)_{t\in I}$ *ein stochastischer Prozess definiert auf dem Wahrscheinlichkeitsraum* $(\Omega,\mathcal{A}_0,\mathbb{P})$ *mit Werten in* $(S,\mathcal{S})$*. Dann definieren die Koordinatenabbildungen*

$$\pi_i: \quad S^I \longrightarrow S$$
$$\omega \longmapsto \omega(t)$$

einen äquivalenten Prozess $(\pi_t)_{t\in I}$ *auf dem Wahrscheinlichkeitsraum* $(S^I,\mathcal{S}^I,\mathbb{P}_x)$*. Dieser Prozess wird auch kanonische Version genannt. Dabei ist:*

$$S^I = \prod_{t\in I} S = \{\text{Abbildungen } \omega:\ I\longrightarrow S\} \quad \text{und} \quad \mathcal{S}^I = \bigotimes_{t\in I}\mathcal{S}^t$$

Die σ*-Algebra* $\mathcal{S}^I$ *ist die kleinste* σ*-Algebra auf* S^I *bezüglich der die Koordinatenabbildungen für alle* $t\in I$ *messbar sind. Das ist die kleinste* σ*-Algebra auf* S^I*, die alle Zylindermengen enthält. Eine Zylindermenge* A *ist gegeben durch:*

$$A=\prod_{t\in I} A_t \quad \text{mit} \quad A_t\in\mathcal{S}\ (\forall t\in I) \quad \text{und} \quad A_t = S \text{ für alle bis auf endlich viele } t$$

Das Maß P_x *ist das Bildmaß von* $\mathbb{P}$ *unter der Abbildung:*

$$X: \quad \Omega\longrightarrow S^I$$
$$\omega\longmapsto X(\omega)$$

Beweis. *Für alle $t_1, \dots, t_n \in I$ und alle $B_1, \dots, B_n \in \mathcal{S}$ gilt:*

$$\mathbb{P}_x(v \in S^I : \pi_{t_1}(v) \in B_1, \dots, \pi_{t_n}(v) \in B_n) = \mathbb{P}(\omega \in \Omega : \pi_{t_1}(X(\omega)) \in B_1, \dots, \pi_{t_n}(X(\omega)) \in B_n)$$
$$= \mathbb{P}(\omega \in \Omega : X_{t_1}(\omega) \in B_1, \dots, X_{t_n}(\omega) \in B_n)$$

□

einige weitere Prozesse
Sei $(X_t)_{t\in\mathbb{R}_+}$ ein stochastischer Prozess definiert auf dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ mit Werten in $(S, \mathcal{S})$. Definiere analog für jede endliche Teilmenge $J \subset I$

$$S^J, \quad \mathcal{S}^J, \quad X_J, \quad \mathbb{P}_J = (X_j)_*\mathbb{P}$$

und für endliche Teilmengen J, H von I, $J \subset H \subset I$, die Projektionen $\pi_J^H : S^H \longrightarrow S^J$ als Restriktion der Abbildungen $(v_t)_{t\in H}$ zu Abbildungen $(v_t)_{t\in J}$. Dann gilt:

$$\pi_J^H \circ X_k = X_j \quad \text{und daher: } \boxed{\mathbb{P}_J = (\pi_J^H)_*\mathbb{P}_H} \tag{3.1}$$

Definition 3.1.12. *Eine Familie $(p_J)_{J \text{ endl.}\subset I}$ von Wahrscheinlichkeitsmaßen p_J auf $(S^J, \mathcal{S}^J)$ heißt projektive Familie, falls gilt:*

$$\boxed{p_J = \left(\pi_J^H\right)_* \mathbb{P}_H \qquad \forall J \subset H \subset I \quad J, H \text{ endlich}}$$

Für jeden stochastischen Prozess $(X_t)_{t\in I}$ definieren die endlich-dimensionalen Verteilungen $(p_J)_{J \text{ endl.}\subset I}$ eine projektive Familie.

Frage: Gilt auch die Umkehrung? Gibt es zu jeder projektiven Familie $(p_J)_J$ einen stochastischen Prozess mit $\mathbb{P}_J = p_J$?

Satz 3.1.13. *Kolmogorov*
Sei S ein Standard-Borel-Raum und I eine beliebige Menge. Zu jeder projektiven Familie $(p_J)_{J \text{ endl.}\subset I}$ von Wahrscheinlichkeitsmaßen auf $\left(S^J, \mathcal{S}^J\right)$ existiert genau ein Wahrscheinlichkeitsmaß p_I auf $\left(S^I, \mathcal{S}^I\right)$ mit:

$$p_J = \left(\pi_J^I\right)_* p_I \qquad (\forall J \text{endl. } J \subset I)$$

Das Maß p_I heißt projektiver Limes der Wahrscheinlichkeitsmaße p_J. Schreibweise:

$$P_I = \varprojlim_{J \text{ endl. } \subset I} P_J$$

Beweis. *(i) Für $\{t_1, \dots, t_n\} = J$ endl. $\subset I$ ist $\mathcal{S}^J = \mathcal{S}^{\otimes n}$ eine σ-Algebra in S^J und $\mathcal{S}_J^I = \left(\pi_J^I\right)^{-1} \mathcal{S}^J$ ist eine σ-Algebra in S^I.*

Also ist $\tilde{\mathcal{S}}^I$, gegeben durch

$$\tilde{\mathcal{S}}^I = \bigcup_{J \text{ endl. } \subset I} \mathcal{S}_J^I,$$

eine $\cap$-stabile Algebra auf S^I, die Algebra der Zylinder-Mengen. Dann ist ferner: $\mathcal{S}^I = \sigma(\tilde{\mathcal{S}}^I)$.

(ii) Eindeutigkeit
Sei p_I ein Wahrscheinlichkeitsmaß auf $\left(S^I, \mathcal{S}^I\right)$ mit 3.1. Dann gibt es für alle $A \in \tilde{\mathcal{S}}^I$ eine endliche Teilmenge $J \subset I$ und Mengen $B \in \mathcal{S}^J$, $A \in \mathcal{S}_J^I$ mit $A = \left(\pi_J^I\right)^{-1}(B)$. Also gilt:

$$p_I(A) = p_I\left(\left(\pi_J^I\right)^{-1}(B)\right) = \left(\left(\pi_J^I\right)_* p_I\right)(B) \overset{3.1}{=} p_J(B) \tag{3.2}$$

$p_J(B)$ ist eindeutig festgelegt durch die projektive Familie $(P_J)_{J \text{ endl. } \subset I}$. Damit ist p_I eindeutig festgelegt auf einem $\cap$-stabilen Erzeuger der $\sigma-$Algebra $\mathcal{S}^I$ und damit eindeutig auf $\mathcal{S}^I$.

(iii) *Existenz*
Definiere p_I zunächst auf $\tilde{\mathcal{S}}^I$ durch 3.2. Dies ist konsistent wegen der Projektionen: $p_H = \left(\pi_J^H\right)_* p_J$. Dies definiert einen normierten Inhalt auf $\left(S^I, \tilde{\mathcal{S}}^I\right)$ (also additiv, aber nicht notwendigerweise σ-additiv. Man kann zeigen, dass p_I $\emptyset$-stetig ist:

$$\forall A_n \in \tilde{\mathcal{S}}^I,\ A_n \searrow \emptyset \text{ gilt: } p_I(A_n) \searrow 0$$

Da S ein Standard-Borel-Raum ist, lässt sich der Inhalt p_I kanonisch zu einem Wahrscheinlichkeitsmaß auf $\left(S^I, \mathcal{S}^I\right)$ fortsetzen.

□

Korollar 3.1.14. *Im obigen Setting gilt:*
Es existiert ein stochastischer Prozess $(X_t)_{t\in I}$ auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ mit vorgegebener endlich-dimensionaler Verteilung:

$$p_J = (X_J)_* \mathbb{P} \qquad \forall J \subset I \text{ endlich}$$

Beweis. *Wähle dazu:*

$$\Omega = S^I, \quad X_t = \pi_t = \mathbb{P} = p_I$$

□

3.2. Identifizieren der endlich-dimensionalen Verteilungen

Satz 3.2.1. *(i) Ist $(X_t)_{t\geq 0}$ eine Brown'sche Bewegung mit Startverteilung ν, so sind die endlich-dimensionalen Verteilungen für $J = \{t_0, t_1, \dots, t_n\}$ mit $0 = t_0 < t_1 \dots < t_n$ und $B \subset \mathcal{B}(\mathbb{R}^{1+n})$ gegeben durch:*

$$\mathbb{P}_X(B) = \int_B p_{t_n - t_{n-1}}(x_{n-1}, dx_n) \cdots p_{t_1 - t_0}(x_0, dx_1)\nu(dx_0) \quad \text{mit} \quad p_t(x, dy) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{|x-y|^2}{2t}\right) dy$$

(ii) Definiert man $p_J(B)$ für alle endlichen J wie oben, so ist $(p_J)_{J \text{ endl.} \subset I}$ eine projektive Familie.

Beweis. *(i) Gegeben sei ein endliches $J = \{t_0, t_1, \dots, t_n\} \subset I$ (o.B.d.A: $0 = t_0 < t_1 < \dots < t_n$) sowie eine Brown'sche Bewegung $(X_t)_{t\geq 0}$ definiert auf $(\Omega, \mathcal{A}, \mathbb{P})$. Dann ist $(Y_k)_{k=0,\dots,n}$ mit $Y_k = X_{t+k}$ eine Markov-Kette mit Übergangswahrscheinlichkeiten $q_k(x, dy) = p_{t_k - t_{k-1}}(x, dy)$ und Startverteilung ν. Hierbei ist:*

$$p_t(x, dy) = \underbrace{\frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(x-y)^2}{2t}\right)}_{=:\gamma_t(x,y)} dy$$

Daraus folgt für alle $B \in \mathcal{B}(\mathbb{R}^{1+n})$:

$$\mathbb{P}\left((X_{t_0}, \dots, X_{t_n}) \in B\right) = \mathbb{P}\left((Y_0, \dots, Y_n) \in B\right) = \int_B q_{n-1}(x_{n-1}, dx_n) \cdots q_1(x_0, dx_1)\nu(dx_0)$$

$$= \int_B p_{t_n - t_{n-1}}(x_{n-1}, dx_n) \cdots p_{t_1 - t_0}(x_0, dx_1)\nu(dx_0)$$

(ii) Gegeben seien endliche Teilmengen $J \subset H \subset I$, also z.B: $H = \{t_0, \dots, t_n\}$, $J = \{t_0, \dots t_i, t_{i+1}, \dots, t_n\}$.
zu zeigen:

$$\int_{B_0} \cdots \int_{B_i} \cdots \int_{B_n} p_{t_n - t_{n-1}}(x_{n-1}, dx_n) \cdots p_{t_1 - t_0}(x_0, dx_1)\nu(dx_0)$$

$$= \int\limits_{B_0} \cdots \int\limits_{B_{i-1}} \int\limits_{\mathbb{R}} \int\limits_{B_{i+1}} \cdots \int\limits_{B_n} p_{t_n - t_{n-1}}(x_{n-1}, dx_n) \cdots p_{t_1 - t_0}(x_0, dx_1)\nu(dx_0)$$

$$\int\limits_{\mathbb{R}} p_{t_{i+1}-t_i}(x_i, dx_{i+1}) p_{t_i - t_{i-1}}(x_{i-1}, dx_i) = \underbrace{\left[\int\limits_{\mathbb{R}} \gamma_{t_{i+1}-t_i}(x_{i+1}, x_i)\gamma_{t_i - t_{i-1}}(x_i, x_{i-1}) dx_i \right]}_{=\gamma_{t_{i+1}-t_{i-1}}(x_{i+1}, x_{i-1})} dx_{i+1}$$

$$= p_{t_{i+1}-t_{i-1}}(x_{i-1}, dx_{i+1})$$

□

Bemerkung 3.2.2. *Satz 3.2.1 gilt auch für die Rohversion der Brown'schen Bewegung (wir verwenden keine Stetigkeit).*

Korollar 3.2.3. *Zu beliebiger Startverteilung ν existiert eine Rohversion der Brown'schen Bewegung.*

Beweis. *Nach Satz 3.2.1(ii) ist $(p_J)_J$ eine projektive Familie. Der projektive Limes p_I existiert nach dem Satz von Kolmogorov (3.1.13). Er ist ein Wahrscheinlichkeitsmaß auf $\Omega = \mathbb{R}^{\mathbb{R}^+}$ mit σ-Algebra $\mathcal{B}(\mathbb{R})^{\mathbb{R}^+}$. Die Projektionen $(\pi_t)_{t\geq 0}$ definieren eine Rohversion der Brown'schen Bewegung.* □

Es bleibt zu zeigen, dass $(X_t)_{t\geq 0}$ fast sicher stetige Trajektorien besitzt, <u>oder</u> dass eine Modifikation $(\tilde{X}_t)_t$ von $(X_t)_t$ mit fast sicher stetigen Trajektorien existiert.

Satz 3.2.4. *(Kolmogorov-Chentsov)*
Sei $(X_t)_{t\geq 0}$ ein stochastischer Prozess mit Werten in $\mathbb{R}$ und für alle $T > 0$ existieren α, β, C mit:

$$\mathbb{E}\left[|X_t - X_s|^\alpha\right] \leq C|t-s|^{1+\beta} \qquad (\forall s, t \in [0, T])$$

Dann existiert eine Modifikation $(\tilde{X}_t)_{t\geq 0}$ von $(X_t)_{t\geq 0}$ deren Trajektorien lokal Hölder-stetig von jeder Ordnung $\gamma < \dfrac{\beta}{\alpha}$ sind, das heißt:

$$\left|\tilde{X}_t(\omega) - \tilde{X}_s(\omega)\right| \leq C|t-s|^\gamma \qquad \textit{für } s, t \in [0, T] \quad \textit{und} \quad |t-s| < \Delta(w)$$

Beweis. *(i) Falls für alle $T \in \mathbb{N}$ eine Modifikation X^T mit lokal Hölder-stetigen Trajektorien existiert, dann sind X^S und X^t ununterscheidbar auf dem Intervall $[0, S \wedge T]$. Es existiert daher eine Modifikation $\tilde{X}$ mit lokal stetigen Trajektorien auf $[0, \infty]$.*

(ii) Im Folgenden gelte o.B.d.A: $T = 1$. Zeige, dass eine Modifikation $\hat{X}$ mit lokal stetigen Trajektorien auf $[0, 1]$ existiert. Mit Tschebyschow folgt aus 3.3:

$$\mathbb{P}(|X_t - X_s| > \epsilon) \leq C\epsilon^{-\alpha} |t-s|^{1+\beta} \qquad (\forall s, t \in [0, 1])$$

Insbesondere konvergiert X_s gegen X_t $\mathbb{P}$-stochastisch für $s \to t$.

(iii) Mit Borel-Cantelli folgt:

$$\exists \Omega_1 \subset \Omega,\ \mathbb{P}(\Omega_1) = 1,\ \textit{sodass}\ \forall \omega \in \Omega_1,\ \forall \gamma \leq \frac{\beta}{\alpha} :\ \exists k, \Delta = \Delta(\omega, \gamma):$$

$$|X_t(\omega) - X_s(\omega)| \leq k\,|t-s|^\gamma \qquad \forall s, t \in \mathbb{D} \textit{ mit } |s-t| < \Delta$$

Hierbei ist $\mathbb{D} = \left\{\frac{k}{2^n},\ k = 0, 1, \ldots, 2^n\ \ n \in \mathbb{N}\right\}$ die Menge der dyadischen Zahlen.
<u>Hinweis:</u>
Sei $I \subset \mathbb{R}$ ein reelles Intervall und $f:\ I \longrightarrow \mathbb{R}$ eine Funktion. Dann ist f genau dann Hölder-stetig, falls $k, \gamma \in \mathbb{R}$ existieren sodass gilt:

$$|f(t) - f(s)| \leq k\,|t-s|^\gamma \qquad (\forall s, t \in I) \tag{3.3}$$

Dann ist insbesondere $t \longmapsto X_t(\omega)$ gleichmäßig stetig und lokal Hölder-stetig auf $\mathbb{D}$.

(iv) Wir definieren nun $(\tilde{X}_t)_{t\in[0,1]}$ wie folgt:

$$\tilde{X}_t(\omega) = \begin{cases} 0 \ \ \forall t \in [0,1] & \textit{falls } \omega \in \Omega \setminus \Omega_1 \\ \lim_{\substack{s\to t \\ s\in\mathcal{D}}} X_s(\omega) & \omega \in \Omega_1 \end{cases}$$

Der Limes $\lim_{\substack{s\to t \\ s\in\mathcal{D}}} X_s(\omega))$ existiert, da $s \longmapsto X_s(\omega)$ gleichmäßig stetig ist.
Dann gilt:

- *$t \longmapsto \tilde{X}_t(\omega)$ ist stetig auf $[0,1]$ für jedes $\omega \in \Omega$.*
- *$\omega \longmapsto \tilde{X}_t(\omega)$ ist messbar, da es Limes von messbaren Abbildungen X_{t_n} mit $t_n \in \mathcal{D}$, $t_n \to t$ ist.*
- *Für jedes $t \in [0,1]$ gilt $\tilde{X}_t = X_t$, denn für $t \in \mathcal{D}$ gilt dies per Definition und für allgemeines $f \in [0,1]$ sei $s_n \in \mathcal{D}$, $s_n \to t$. Dann gilt:*

$$X_{s_n} \longrightarrow X_t \quad \mathbb{P}-\textit{stochastisch}, \quad X_{s_n} \longrightarrow \tilde{X}_t \quad \mathbb{P}-\textit{fast sicher} \quad \Rightarrow X_t = \tilde{X}_t \quad \mathbb{P}-\textit{fast sicher}$$

□

Korollar 3.2.5. *(i) Zu jeder Startverteilung existiert eine Brown'sche Bewegung .*

(ii) Jede Brown'sche Bewegung hat fast sicher Trajektorien, die lokal Hölder-stetig von jeder Ordnung $\gamma < \frac{1}{2}$ sind.

Beweis. *(i) Für jedes Wahrscheinlichkeitsmaß ν auf $\mathbb{R}$ existiert eine Rohversion $(X_t)_{t\geq 0}$ der Brown'schen Bewegung. Dann gilt 3.3 für alle $\alpha > 2$ (dann ist: $1+\beta = \frac{\alpha}{2}$ und $\gamma = \frac{1}{2} - \frac{1}{\alpha}$). Daher existiert eine Modifikation von $(X_t)_{t\geq 0}$ mit Hölder-stetigen Trajektorien.*

(ii) Gegeben sei eine Brown'sche Bewegung $(X_t)_{t\geq 0}$. Dann existiert eine Modifikation $(\tilde{X}_t)_{t\geq 0}$ mit Hölder-stetigen Trajektorien.

□

Bemerkung 3.2.6. *Wir wissen:*

- *Die endlich-dimensionalen Verteilungen von X und $\tilde{X}$ stimmen überein.*
  ($\mathbb{P}_x = \mathbb{P}_{\tilde{X}}$ auf $(\mathbb{R}^{\mathbb{R}^*}, \mathcal{B}(\mathbb{R})^{R^+})$)
- *Es gilt:*

$$X, \tilde{X}:\ \Omega \longrightarrow \mathbb{R}^{\mathbb{R}_+} \quad \textit{und} \quad \tilde{X}:\ \Omega \longrightarrow \mathcal{C} := \mathcal{C}(\mathbb{R}_+, \mathbb{R}) \subset \mathbb{R}^{\mathbb{R}_+}$$

Daher ist es naheliegend zu behaupten:

$$\mathbb{P}_{\tilde{X}}(\mathcal{C}) = 1$$

Die Menge $\mathcal{C}$ ist aber nicht messbar! Es gilt aber trotzdem folgender Satz:

Satz 3.2.7. *Sei $(X_t)_{t\geq 0}$ wie im Satz 3.2.4. Dann existiert ein Wahrscheinlichkeitsmaß $\hat{\mathbb{P}}$ auf $(\mathcal{C}, \mathcal{B}(\mathcal{C}))$ sodass der Projektionsprozess $(\pi_t 4)_{t\geq 0}$ auf $(\mathcal{C}, \mathcal{B}(\mathcal{C}), \hat{\mathbb{P}})$ äquivalent zum Ausgangsprozess ist.*
Hinweis: Im Fall der Brown'sche Bewegung heißt $\hat{\mathbb{P}}$ Wiener-Maß.

3.3. Konkrete Konstruktion der Brown'schen Bewegung

1) Kurzer Exkurs: der Hilbertraum $L^2[0,1]$:

- Es sei $L^2[0,1]$ der Hilbertraum der quadratisch integrierbaren Funktionen $f:\ [0,1] \longrightarrow \mathbb{R}$.

$$L^2[0,1] = \left\{ f:\ [0,1] \longrightarrow \mathbb{R} \mid f \text{ messbar}, \int_{[0,1]} f^2(x)dx < \infty \right\} / \sim$$

$$\text{wobei } f \sim g \iff f(x) = g(x) \ \text{ fast überall auf } [0,1]$$

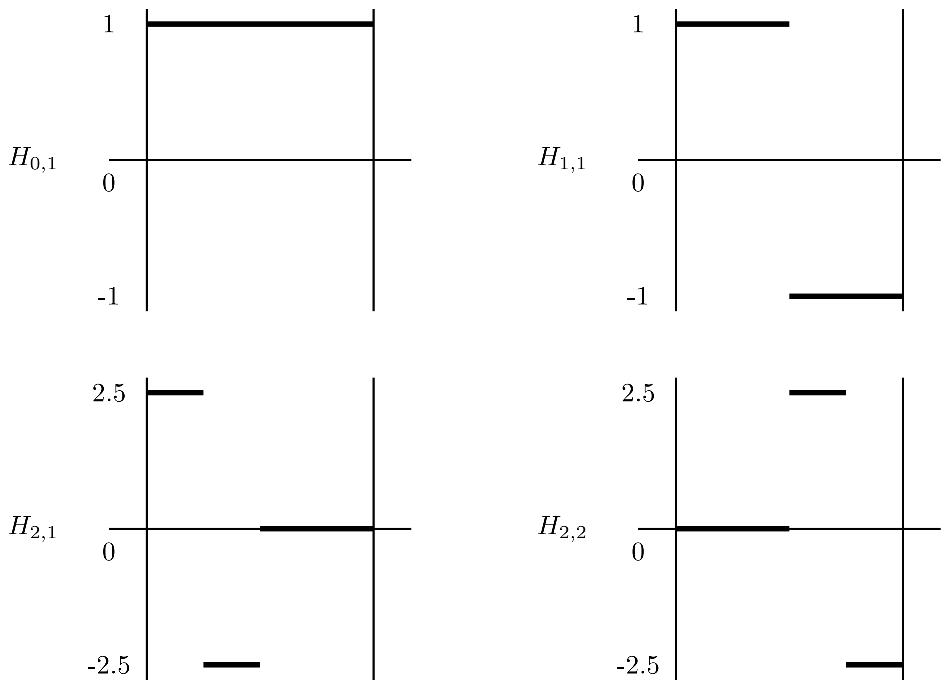
- mit Skalarprodukt:
$$< f, g >= \int_0^1 f(x)g(x)dx$$
und Norm:
$$||f||_2 = \sqrt{< f, f >}$$
- Eine Folge $(\phi_n) \in L^2[0,1]$ heißt Orthonormalsystem , falls gilt: $< \phi_n, \phi_m >= \delta_{nm}$.
- Ein Orthonormalsystem heißt vollständig (oder Orthonormalbasis (ONB)), falls
$$\left\{ \sum_{n=1}^{N} a_n \phi_n \mid N \in \mathbb{N},\ a_1, \ldots, a_N \in \mathbb{R} \right\} \text{ dicht in } L[0,1] \text{ ist.}$$

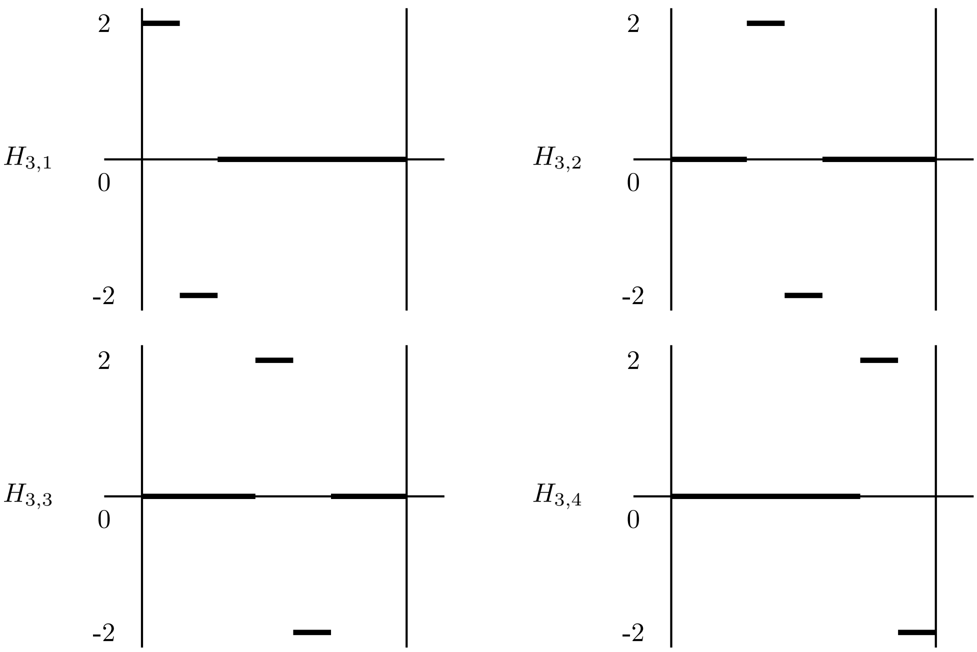
Zwei Tatsachen über Orthonormalbasen:
Sei (ϕ_n) eine Orthonormalbasis. Dann gilt:

(1) Für alle Funktionen $f \in L^2[0,1]$ gilt:
$$f = \sum_{n=1}^{\infty} < f, \phi_n > \phi_n \qquad \text{i.e.} \left\| f - \sum_{n=1}^{N} < f, \phi_n > \phi_n \right\|_2 \overset{N \to \infty}{\longrightarrow} 0$$

(2) Persevalsche Gleichung
Für alle Funktionen $f, g \in L^2[0,1]$ gilt:
$$< f, g >= \sum_{n=1}^{\infty} < f, \phi_n >< g, \phi_n > \qquad \text{Insbesondere: } ||f||_2^2 = \sum_{n=1}^{\infty} < f, \phi_n >^2$$

2) Die Haarbasis von $L^2[0,1]$:
Setze:
$$H_{0,1} := 1_{[0,1]}, \qquad H_{1,1}(t) := \begin{cases} 1 & \text{falls } 0 \le t \le \frac{1}{2} \\ -1 & \text{falls } \frac{1}{2} \le t < 1 \\ 0 & \text{sonst} \end{cases}$$
$$H_{n,k}(t) := 2^{\frac{(n-1)}{2}} H_{1,1}\left(2^{n-1}t - k + 1\right) \qquad \text{für } n \ge 1,\ 1 \le k \le 2^{n-1}$$





Proposition 3.3.1. $\{H_{0,1}\} \cup \{H_{n,k}\}_{n \geq 1, 1 \leq k \leq 2^{n-1}}$ *ist eine Orthonormalbasis für* $L^2[0,1]$.

Beweis. (Idee)

- *Orthogonalität: Übung!*
- $\int_0^1 H_{n,k}(x)^2 dx = \left(2^{\frac{n-1}{2}}\right)^2 \frac{1}{2^n} + \left(-2^{\frac{n-1}{2}}\right)^2 \frac{1}{2^n} = 1$
- *Vollständigkeit:*

  1. *Jede dyadische Indikatorfunktion*

  $$1_{\left[\frac{k}{2^n}, \frac{k+1}{2^n}\right]} \qquad (n \in \mathbb{N},\ k = 0, \ldots, 2^n - 1)$$

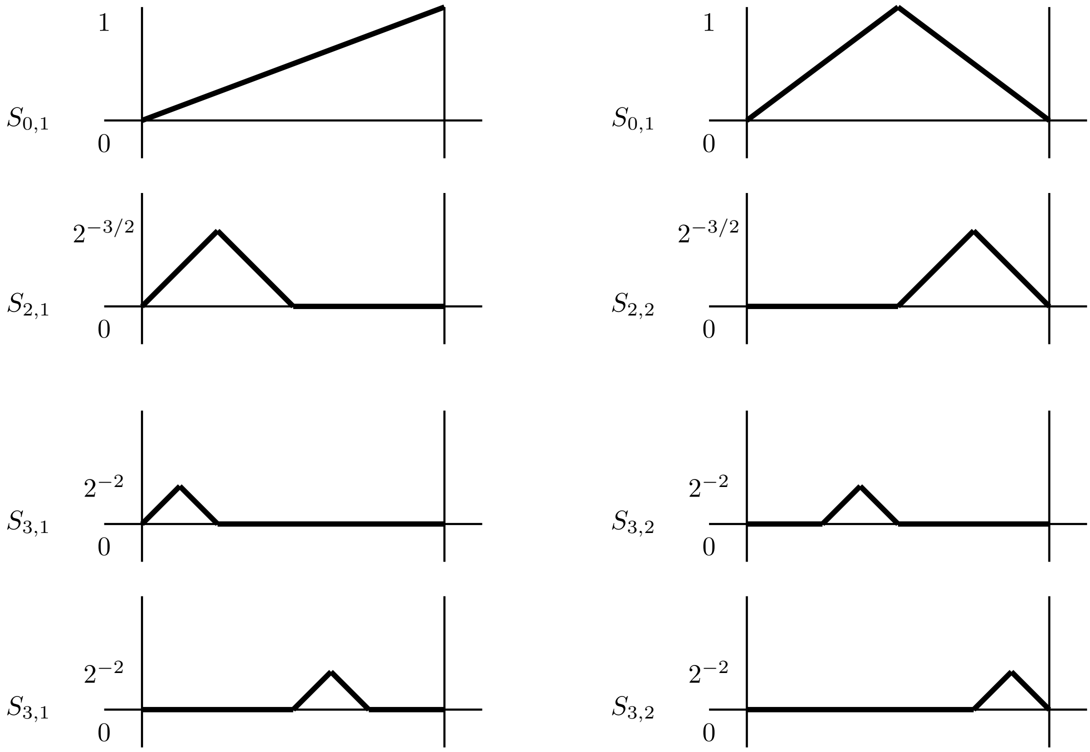
     ist eine endliche Linearkombination von Haarfunktionen.
  2. *Jede Indikatorfunktion* 1_A*, wobei* A *eine Borel-Menge ist, ist in* $L^2[0,1]$ *approximierbar durch dyadische Indikator-Funktionen.*
  3. *Jede* $L^2[0,1]$*-Funktion ist approximierbar durch Linearkombinationen von Indikatorfunktionen.*

$\square$

Bemerkung 3.3.2. *Eine Basis, die aus Translationen und Dilatationen von einer festen Funktion besteht, heißt Waveletbasis.*

3) Schauderfunktionen
Die Schauderfunktionen $S_{n,k}$ sind definiert durch:

$$S_{n,k}(t) = \int_0^\infty H_{n,k}(s) ds$$



Satz 3.3.3. *Seien* $\{\xi_{0,1}\} \cup \{\xi_{n,k}\}_{n\geq 1, 1\leq k\leq 2^{n-1}}$ *unabhängige und* $N(0,1)$*-verteilte Zufallsvariablen. Definiere:*

$$X^N(t) := \xi_{0,1}S_{0,1} + \sum_{n=1}^{N}\sum_{k=1}^{2^{N-1}} \xi_{n,k}S_{n,k}$$

Dann gilt:

1. *Es existiert fast sicher der Grenzwert* $X := L^\infty - \lim_{N\to\infty} X^N$.
2. X *ist eine Brown'sche Bewegung.*

Um Satz 3.3.3 zu beweisen, benutzen wir:

Satz 3.3.4. *Für einen stochastischen Prozess* $X = (X_t)_{t\in[0,\infty)}$ *sind äquivalent:*

(i) X *ist eine Brown'sche Bewegung.*

(ii) X *ist ein stetiger, zentrierter, Gauß'scher Prozess mit:*

$$Cov[X_s, X_t] = s \wedge t \qquad (\forall s, t \geq 0)$$

Beweis. (i) $\Rightarrow$ (ii) *Sei* X *eine Brown'sche Bewegung. Für* $0 \leq s \leq t$ *gilt:*

$$Cov[X_s, X_t] = \mathbb{E}[X_sX_t] = \mathbb{E}\left[X_s^2\right] + \mathbb{E}[X_s(X_t - X_s)] = s + \underbrace{\mathbb{E}[X_s]}_{=0}\mathbb{E}[X_t - X_s] = s \wedge t$$

(ii) $\Rightarrow$ (i) *Weil* X *ein Gauß-Prozess ist und* $Cov[X_s, X_t] = s \wedge t$ *ist, gilt:*

$$(X_{t_1}, \ldots, X_{t_n}) \sim N(0, Q) \qquad \text{wobei} \quad Q_{ij} = t_i \wedge t_j$$

Also sind die endlich-dimensionalen Verteilungen von X *eindeutig bestimmt. Da* X *außerdem stetig ist, ist* X *eine Brown'sche Bewegung.*

□

Beweis. von Satz 3.3.3

1. Sei (b_n) eine Orthonormalbasis für $L^2[0,1]$. Setze:

$$B_n(t) = \int\limits_0^t \underbrace{b_n(s)}_{=<b_n, 1_{[0,t]}>} ds \quad \text{und} \quad X^N(t) = \sum_{n=1}^{N} \xi B_n(t) \quad \text{mit } (\xi_n) \text{ iid Gauß}$$

<u>Behauptung:</u> Für festes $t \in [0,1]$ gilt:

$$(X^n(t))_{n\in\mathbb{N}} \text{ ist eine Cauchy-Folge in } L^2(\mathbb{R})$$

Beweis. für $m \leq n$
Beachte:

$$\mathbb{E}[\xi_k \xi_l] = \delta_{kl} \tag{3.4}$$

$$\mathbb{E}\left[|X^m(t) - X^n(t)|^2\right] = \mathbb{E}\left[\left|\sum_{k=m+1}^{n} \xi_k B_k(t)\right|^2\right] = \mathbb{E}\left[\sum_{k,l=m+1} \xi_k \xi_l B_k(t) B_l(t)\right] \overset{3.4}{=} \sum_{k=m+1}^{n} B_n(t)^2$$

$$= \sum_{k=m+1}^{n} < b_n, 1_{[0,t]} >^2 \leq \sum_{k=m+1}^{\infty} < b_n, 1_{[0,t]} >^2 \overset{m\to\infty}{\longrightarrow} 0,$$

weil mit Perseval gilt:

$$\sum_{k=1}^{\infty} < b_n, 1_{[0,t]} >^2 = ||1_{[0,t]}||_2^2 = t$$

Also ist die Folge $(X^n(t))_n$ eine Cauchy-Folge in $L^2(\mathbb{R})$. □

2. Für $t_1, \ldots, t_N \geq 0$ gilt:

$$\underline{X}_N^n = \left(X_{t_1}^n, \ldots, X_{t_N}^n\right) \quad \text{ist Gauß-verteilt mit Kovarianz} \quad \mathbb{E}\left[X_{t_i}^n, X_{t_j}^n\right] \overset{n\to\infty}{\longrightarrow} t_i \wedge t_j$$

Beweis.

$$\mathrm{Cov}(X_s^n, X_t^n) = \mathbb{E}\left[\left(\sum_{k=1}^{n} \xi_k B_n(s)\right)\left(\sum_{l=1}^{n} \xi_l B_l(s)\right)\right] \overset{3.4}{=} \sum_{k=1}^{n} B_k(s) B_k(t)$$

$$= \sum_{k=1}^{n} < b_k, 1_{[0,s]} >< b_k, 1_{[0,t]} > \overset{\text{Perseval}}{\longrightarrow} < 1_{[0,s]}, 1_{[0,t]} > = \int\limits_0^1 1_{[0,s]}(r) 1_{[0,t]}(r) dr = s \wedge t$$

□

□

Satz 3.3.5. *Seien $\{\xi_{n,k}\}$ unabhängige und $N(0,1)$-verteilte Zufallsvariablen. Definiere:*

$$X^N(t) = \xi_{0,1} S_{0,1}(t) + \sum_{n=1}^{N} \sum_{k=1}^{2^{n-1}} \xi_{n,k} S_{n,k}(t)$$

Dann gilt:

- $X = L^\infty = \lim\limits_{N\to\infty} X^N$ *existiert fast sicher.*
- X *ist eine Brown'sche Bewegung.*

Beweis. <u>**step 1**</u> *Für festes $t \in [0,1]$ ist $\left(X^N(t)\right)_N$ eine Cauchy-Folge in $L^2(\mathbb{R})$.*

step 2 *Für $t_1, \ldots, t_k \in [0,1]$ ist $\underline{X}_k^N = \left(X_{t_1}^N, \ldots, X_{t_k}^N\right)$ zentriert Gauß-verteilt mit Kovarianz*

$$Q_{ij}^N := \mathbb{E}\left[X_{t_i}^N, X_{t_j}^N\right] \overset{N\to\infty}{\longrightarrow} t_i \wedge t_j.$$

step 3 *Für $t_1, \ldots, t_k \in [0,1]$ ist $\underline{X}_k(X_{t_1}, \ldots, X_{t_j})$ zentriert Gauß-verteilt mit Kovarianz*

$$Q_{ij}^N := \mathbb{E}\left[X_{t_i}, X_{t_j}\right] = t_i \wedge t_j.$$

Beweis. *Es gibt eine Teilfolge N_j, sodass X_k^N fast sicher gegen X_k konvergiert. Es gilt also mit step 2 für $\xi \in \mathbb{R}^K$.*

$$\mathbb{E}\left[e^{i\xi \underline{X}_k}\right] = \lim_{j\to\infty} \mathbb{E}\left[e^{i\xi \underline{X}_k^{N_j}}\right] = \lim_{j\to\infty} e^{<Q^{N_j}\xi,\xi>} = e^{<Q\xi,\xi>}$$

$\Rightarrow \underline{X}_k$ ist zentriert Gauß-verteilt mit Kovarianz Q, also: $\mathbb{E}\left[X_{t_i}, X_{t_j}\right] = t_i \wedge t_j$. ☐

step 4 *Wir zeigen:*

$$\lim_{N\to\infty} ||X^N - X||_\infty = 0 \qquad \mathbb{P}\text{-fast sicher}$$

Beweis. *Da $\mathcal{C}([0,1], ||\cdot||_\infty)$ vollständig ist, reicht es zu zeigen, dass $\left\{X_k^N\right\}$ fast sicher eine Cauchy-Folge auf $\mathcal{C}([0,1], ||\cdot||_\infty)$ ist.*
Beachte:

- $||S_{n,k}||_\infty = 2^{-\frac{n+1}{2}} \leq 2^{-\frac{n}{2}}$
- $S_{n,k} \cdot S_{n,l} = 0$ *falls* $k \neq l$.

Also gilt:

$$||X^n - X^{m-1}||_\infty = \left|\left|\sum_{k=1}^{2^{n-1}} \xi_{n,k} S_{n,k}\right|\right|_\infty \leq 2^{-\frac{n}{2}} \max_{1\leq k\leq 2^{n-1}} \{|\xi_{n,k}|\}$$

Daraus folgt für alle $c > 0$:

$$\mathbb{P}(||X^n - X^{n-1}||_\infty > c) \leq \mathbb{P}\left(2^{-\frac{n}{2}} \max_{1\leq k\leq 2^{n-1}} \{|\xi_{n,k}|\} > c\right) \leq \sum_k^{2^{n-1}} \mathbb{P}\left(|\xi_{n,k}| > c\cdot 2^{\frac{n}{2}}\right)$$

$$= 2^{n-1}\frac{2}{\sqrt{2\pi}} \int_{c\cdot 2^{\frac{n}{2}}}^{\infty} e^{-\frac{x^2}{2}} dx \overset{trick}{\leq} 2^{n-1}\frac{2}{\sqrt{2\pi}} e^{-\frac{c^2 2^n}{2}} \leq 2^n e^{-c^2\cdot 2^{n-1}}$$

Verwende dazu den folgenden Trick: Für alle $y > 1$ gilt:

$$\int_y^\infty e^{-\frac{x^2}{2}} dx \leq \int_y^\infty x e^{-\frac{x^2}{2}} dx = \left[e^{-\frac{x^2}{2}}\right]_j^\infty = e^{-\frac{y^2}{2}}$$

Setze $c = 2^{-\frac{n}{4}}$. Dann gilt:

$$\mathbb{P}(||X^n - X^{x-1}||_\infty > 2^{-\frac{n}{4}}) \leq 2^k \exp(-2^{\frac{n}{2}-1})$$

Insbesondere gilt:

$$\sum_{k=1}^\infty \mathbb{P}\left(||X^n - X^{n-1}|| > 2^{-\frac{n}{4}}\right) \leq \sum_{k=1}^\infty 2^k \exp(-2^{\frac{n}{2}-1}) \overset{\exp(-x)\leq Cx^{-4}}{\leq} C\sum_{n=1}^\infty 2^n \cdot 2^{-2n+4} = \tilde{C}\sum_{n=1}^\infty 2^{-n} < \infty$$

Also, nach dem Borel-Cantelli Lemma:

$$\mathbb{P}(||X^n - x^{n-1}||_\infty > 2^{\frac{n}{4}} \textit{ höchstens endlich oft}) = 1$$

Es existiert also fast sicher ein $N \in \mathbb{N}$, sodass für alle $N \leq m \leq n$ gilt:

$$||X^n - X^m||_\infty \leq \sum_{k=m+1}^{n} ||X^k - X^{k-1}||_\infty \leq \sum_{k=m+1}^{n} 2^{-\frac{k}{n}} \overset{m,n\to\infty}{\longrightarrow} 0$$

Die Folge (X_k^n) ist also eine Cauchy-Folge in $\mathcal{C}\,[0,1]$ bezüglich $||\cdot||_\infty$. □

□

Satz 3.3.6. *Sei $(B_t)_{t\geq 0}$ eine Brown'sche Bewegung. Definiere:*

$$X_t = \frac{1}{a} B_{a^2 t}, \qquad \textit{für } a > 0$$

$$Y_t = \begin{cases} tB_{\frac{1}{t}} & \textit{falls } t > 0 \\ 0 & \textit{falls } t = 0 \end{cases}$$

Dann sind X und Y wieder Brown'sche Bewegungen.

Beweis. Die Prozesse X und Y sind zentrierte Gauß-Prozesse.

- *X ist stetig auf $[0,\infty)$ und es gilt:*

$$\mathbb{E}\left[X_s X_t\right] = \frac{1}{a^2}\mathbb{E}\left[B_{a^2 s} B_{a^2 t}\right] = \frac{1}{a^2}(a^2 s \wedge a^2 t) = s \wedge t$$

Wegen Satz 3.3.4 ist X also eine Brown'sche Bewegung.

- *Y ist stetig auf $(0,\infty)$ und für $s,t > 0$ gilt:*

$$\mathbb{E}\left[Y_s Y_t\right] = st\mathbb{E}\left[B_{\frac{1}{s}} B_{\frac{1}{t}}\right] = st\left(\frac{1}{s} + \frac{1}{t}\right) = st\frac{1}{s \vee t} = s \wedge t$$

Es bleibt zu zeigen, dass Y stetig in 0 ist. Betrachte dazu:

1) $Y|_{\mathbb{Q}}$ hat die gleiche Verteilung wie $B|_{\mathbb{Q}}$. Es gilt also:

$$\lim_{t\downarrow 0, tt\in\mathbb{Q}} Y_t = 0 \qquad \textit{fast sicher}$$

2) Y ist fast sicher stetig auf $(0,\infty)$, also:

$$0 = \lim_{t\downarrow 0, t\in\mathbb{Q}} Y_t = \lim_{t\searrow 0} Y_t$$

□

3.4. Invarianzprinzip

Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und $(\xi_k)_{k\in\mathbb{N}}$ iid Zufallsvariablen $\in L^2$ mit $\mathbb{E}\left[\xi_k\right] = 0$, $\mathrm{Var}(\xi_k) = \sigma^2 > 0$.

- Partialsummen:

$$X_n = \sum_{k=1}^{n} \xi_k$$

- stetige Interpolation:
$$X_t := X_{\lfloor t\rfloor} + (t - \lfloor t\rfloor)\xi_{\lfloor t\rfloor+1}$$
- Reskalierung (von Raum und Zeit:)
$$X_t^{(n)} = \frac{1}{\sqrt{n}\sigma}X_{nt} \qquad X^{(n)} \text{ ist Zufallsvariable auf } \mathcal{C}_0 = \{v \in \mathcal{C}(\mathbb{R},\mathbb{R}),\ v(0) = 0\}$$

Dann ist $X^{(n)} := (X_t^{(n)})_{t\geq 0}$ ein stetiger stochastischer Prozess mit Werten in $\mathbb{R}$, definiert auf $(\Omega, \mathcal{A}, \mathbb{P})$ (z.B. $\Omega = \{0,1\}^{\mathbb{N}}$, $\mathbb{P}$ Bernoulli-Maß $\xi_k = \pm 1$.)

Satz 3.4.1. *Invarianzprinzip von Donsker*
Die Zufallsvariable $X^{(n)}$ konvergiert für $n \to \infty$ in Verteilung gegen eine Brown'sche Bewegung.
Mit anderen Worten: Die Wahrscheinlichkeitsmaße $X_*^{(n)}\mathbb{P}$ konvergieren im Sinne der schwachen Konvergenz von Wahrscheinlichkeitsmaßen auf $(\mathcal{C}_0, \mathcal{B}(\mathcal{C}_0))$ gegen das Wiener Maß $\hat{\mathbb{P}}$.
Explizit bedeutet das:
$$\lim_{n\to\infty} \int_\Omega f(X^{(n)})d\mathbb{P} = \int_{\mathcal{C}_0} f(\cdot)d\hat{\mathbb{P}} \qquad (\forall f \in \mathcal{C}_b(\mathcal{C}_0))$$

Erste Vorüberlegungen und Bemerkungen:

1. *Für $s = \frac{k}{n}$ und $t = \frac{k+l}{n}$ mit $k, l \in \mathbb{N}$ sind die Zuwächse $X_t^{(n)} - X_s^{(n)} = \frac{1}{\sqrt{n}\sigma}(\xi_{k+1} + \xi_{k+2} + \cdots + \xi_{k+l})$ unabhängig von $\xi_1, \ldots, \xi_n$ und damit insbesondere damit auch unabhängig von $(X_r^{(n)})_{0\leq r\leq s}$ (und damit von deren Zuwächsen) mit Erwartungswert 0 und Varianz $\frac{l}{n} = t - s$. Es ist im Limes $n, l \to \infty$ Gauß-verteilt.*
2. *Aus dem Invarianzprinzip folgt der Zetrale Grenzwertsatz. Denn für alle $t \in (0, \infty)$ ist die Abbildung $\pi_t : \mathcal{C}_0 \longrightarrow \mathbb{R},\ v \longmapsto v(t)$ stetig. Es gilt also insbesondere für $t = 1$:*
$$X_*^n\mathbb{P} \longrightarrow \hat{\mathbb{P}} \quad \text{auf } \mathcal{C}_0 \quad \Rightarrow \quad (X_1^n)_*\mathbb{P} \longrightarrow N(0,1) \quad \text{auf } \mathbb{R}$$
Das heißt:
$$\left(\frac{1}{\sqrt{n}\sigma}\sum_{k=1}^n \xi_k\right)_* \mathbb{P} \longrightarrow N(0,1) \qquad \textit{Zentraler Grenzwertsatz}$$

Der Beweis des Theorems beruht auf folgendem allgemeinen Satz über die Konvergenz von Wahrscheinlichkeitsmaßen auf $\mathcal{C}_0$.

Satz 3.4.2. *Sei $(\mu_n)_n$ Folge von Wahrscheinlichkeitsmaßen auf $\mathcal{C}_0$. Dann sind äquivalent:*

(i) *$(\mu_n)_n$ ist schwach konvergent.*

(ii) *$\{\mu_n\}_n$ ist relativ kompakt (bzgl. der schwachen Topologie in $\mathcal{P}(\mathcal{C}_0)$) und für jedes endliche $J = \{t_1, \ldots, t_k\}$ gilt: Die endlich-dimensionale Verteilung $((\pi_J)_*\mu_n)_{n\in\mathbb{N}}$ ist schwach konvergent als Wahrscheinlichkeitsmaß auf $\mathbb{R}^k$.*

Beweis. *(i) $\Rightarrow$ (ii) Für alle endlichen Mengen $J = \{t_1, \ldots, t_k\}$ und alle Funktionen $f \in \mathcal{C}_b(\mathbb{R}^k)$ mit $\hat{f} := f \circ \pi_J \in \mathcal{C}_b(\mathcal{C}_0)$ gilt nach Voraussetzung:*
$$\int_{\mathbb{R}^k} fd(\pi_J)_*\mu_n = \int_{\mathcal{C}_0} f(\pi_J(v))d\mu_n(v) = \int_{\mathcal{C}_0} \hat{f}(v)d\mu_n(v) \overset{n\to\infty}{\longrightarrow} \int_{\mathcal{C}_0} \hat{f}(v)d\mu(v) = \int_{\mathbb{R}^k} fd(\pi_J)_*\mu$$
Das heißt, $(\pi_J)_*\mu_n$ konvergiert als Wahrscheinlichkeitsmaß auf $\mathbb{R}^k$ schwach gegen $(\pi_J)_*\mu$.

(ii) ⇒ (i) *Sei $\{\mu_n\}_n$ schwach relativ kompakt, das heißt für jede Teilfolge $(\mu'_n)_n$ existiert eine Teilfolge $(\mu''_n)_n$, die schwach gegen μ'' konvergiert.*
Behauptung: μ'' ist eindeutig.
Sei μ^* ein weiteres Maß, gegen das eine Teilfolge $(\mu^*_n)_n$ konvergiert. Dann gilt für alle endlichen J:

$$(\pi_J)_* \mu'' = \lim_{n\to\infty} (\pi_J)_* \mu''_n = \underbrace{\lim_{n\to\infty} (\pi_J)_* \mu_n}_{ex.\ n.V.} = \lim_{n\to\infty} (\pi_J)_* \mu^*_n = (\pi_J)_* \mu^* = (\pi_J)_* \mu^*$$

Das bedeutet, dass μ'' und μ^* die selben endlich-dimensionalen Verteilungen haben und sie damit übereinstimmen. Es bleibt zu zeigen, dass die komplette Folge $(\mu_n)_n$ (und nicht nur $(\mu''_n)_n$) gegen μ'' konvergiert.
Annahme: Die Folge $(\mu_n)_n$ konvergiere nicht gegen μ''.
Dann existieren eine Funktion $f \in \mathcal{C}_b(\mathcal{C}_0)$ sodass $\int f d\mu_n$ nicht gegen $\int f d\mu^*$ konvergiert, wobei entweder der Grenzwert von $\int f d\mu_n$ nicht existiert oder ungleich $\int f d\mu^*$ ist. Dann existiert eine Teilfolge $(\hat{\mu}_n)_n$ sodass $\lim_{n\to\infty} \int df \hat{\mu}_n$ existiert und ungleich $\int f d\mu''$ ist.
Dann existiert aber keine Teilfolge $(\hat{\hat{\mu}}_n)_n$ von $(\hat{\mu}_n)_n$ mit $\lim_{n\to\infty} \int f d\hat{\hat{\mu}}_n = \int f d\mu''$.

□

Beispiel 3.4.3. *für die Notwendigkeit der Bedingung „relativ kompakt“*
Sei $X^n(\omega)$ unabhängig von ω gegeben durch:

$$X^n(\omega) = \begin{cases} nt & \text{falls } t \in \left[0, \frac{1}{2n}\right) \\ \frac{1}{2} - n(t - \frac{1}{2n}) & \text{falls } t \in \left[\frac{1}{2n}, \frac{1}{n}\right) \\ 0 & \text{falls } t \geq \frac{1}{n} \end{cases}$$

Dann gilt für alle $t \geq 0$:

$$X^n_t(\omega) \overset{n\to\infty}{\longrightarrow} 0 \qquad X^{(n)}_\cdot(\omega) \overset{n\to\infty}{\not\longrightarrow} 0$$

und:

$$(X^n_t)_* \mathbb{P} \overset{n\to\infty}{\longrightarrow} \delta_0 \quad \text{Dirac-Maß in } 0 \in \mathbb{R}$$

aber:

$$\left(X^{(n)}_\cdot\right)_* \mathbb{P} \overset{n\to\infty}{\not\longrightarrow} \delta_0 \quad \text{Dirac-Maß in } \vec{0} \in \mathcal{C}_0$$

Beweis. (des Invarianzprinzips)

I. Konvergenz der endlich-dimensionalen Verteilungen
Behauptung: Für alle endlichen $J = \{t_1, \dots, t_k\}$ konvergiert $(\pi_J)_*\mu_n$ gegen $(\pi_J)_* \mu$. Betrachte dazu zunächst $k = 1$ und zeige:

$$(\pi_t)_*\mu_n \overset{n\to\infty}{\longrightarrow} (\pi_t)_* \mu \qquad \text{schwach als Wahrscheinlichkeitsmaß auf } \mathbb{R}$$

RS: $(\pi_t)_*\mu = N(0,t) = (X_t)_*\tilde{\mathbb{P}}$ ist normalverteilt mit Erwartungswert 0 und Varianz t.

LS: $(\pi_t)_* \mu_n = \left(X^{(n)}_t\right)_* \mathbb{P}$

Es gilt:

- $$\left| X^{(n)}_t - \frac{1}{\sigma\sqrt{\lfloor nt \rfloor}} X_{\lfloor nt \rfloor} \right| \leq \underbrace{\frac{1}{\sigma\sqrt{n}} \left|\xi_{\lfloor nt \rfloor + 1}\right|}_{||\cdot||_{L^2} = \frac{1}{\sqrt{n}}} \overset{n\to\infty}{\longrightarrow} 0 \qquad \text{in } L^2 \text{ und damit } \mathbb{P}\text{-stochastisch}$$
- $\frac{\lfloor nt \rfloor}{n} \overset{n\to\infty}{\longrightarrow} t.$

Daraus folgt: $\left(X_t^{(n)}\right)_* \mathbb{P} \overset{n\to\infty}{\longrightarrow} N(0,t)$. Verwende nun den Zentralen Grenzwertsatz:

$$\frac{1}{\sigma\sqrt{\lfloor nt\rfloor}} X_{\lfloor nt\rfloor} \longrightarrow N(0,1) \quad \Rightarrow \quad \frac{1}{\sigma\sqrt{n}} X_{\lfloor nt\rfloor} \longrightarrow N(0,t) \quad \Rightarrow \quad X_t^{(n)} \longrightarrow N(0,t)$$

Dies zeigt die Behauptung für $k=1$. Sei nun $J=\{s,t\}$ zweielementig mit $s<t$. Behauptung:

$$\left(X_s^{(n)}, X_t^{(n)}\right)_* \mathbb{P} \overset{n\to\infty}{\longrightarrow} \left(X_s, X_t)_* \tilde{P}\right) \tag{3.5}$$

Betrachte die lineare Transformation $\phi : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, $(s,t) \longmapsto (s, t-s)$. Dann ist 3.5 äquivalent zu:

$$\left(X_s^{(n)}, X_t^{(n)} - X_s^{(n)}\right)_* \mathbb{P} \overset{n\to\infty}{\longrightarrow} \left(X_s, X_t - X_s)_* \tilde{P}\right) \tag{3.6}$$

RS: Nach Definition der Brown'schen Bewegung sind X_s und $X_t - X_s$ unabhängig, also gilt:

$$\left(X_s, X_t - X_s\right)_* \tilde{P} = N(0,s) \otimes N(0,t-s)$$

LS:

$$\left(X_s^{(n)} - \frac{1}{\sigma\sqrt{n}} X_{\lfloor ns\rfloor}\right) \overset{n\to\infty}{\longrightarrow} 0 \quad \mathbb{P}\text{-stochastisch}$$
$$\left(X_t^{(n)} - \frac{1}{\sigma\sqrt{n}} X_{\lfloor nt\rfloor}\right) \overset{n\to\infty}{\longrightarrow} 0 \quad \mathbb{P}\text{-stochastisch}$$

$$\left(\frac{1}{\sigma\sqrt{n}} X_{\lfloor ns\rfloor}, \frac{1}{\sigma\sqrt{n}}\left(X_{\lfloor nt\rfloor} - X_{\lfloor ns\rfloor}\right)\right)_* \mathbb{P} = \left(\frac{1}{\sigma\sqrt{n}} \sum_{k=1}^{\lfloor ns\rfloor} \xi_k, \frac{1}{\sigma\sqrt{n}} \sum_{j=\lfloor ns\rfloor+1}^{\lfloor nt\rfloor} \xi_j\right)_* \mathbb{P}$$
$$\overset{\text{unabh.}}{=} \left(\frac{1}{\sigma\sqrt{n}} \sum_{k=1}^{\lfloor ns\rfloor} \xi_k\right)_* \mathbb{P} \otimes \left(\frac{1}{\sigma\sqrt{n}} \sum_{j=\lfloor ns\rfloor+1}^{\lfloor nt\rfloor} \xi_j\right)_* \mathbb{P} \overset{n\to\infty}{\longrightarrow} N(0,s) \otimes N(0,t-s)$$

Analog für allgemeines $J = \{t_1, \ldots, t_n\}$.

II. Relative Kompaktheit von $\{\mu_n\}_n$
Dies beruht auf folgenden fundamentalen Resultaten:

Lemma 3.4.4. *(Prohorov)*
Eine Familie $\{\mu_i\}_{i\in I}$ von Wahrscheinlichkeitsmaßen auf einem polnischen Raum E ist genau dann relativ kompakt (bzgl. der schwachen Topologie), wenn für alle $\epsilon > 0$ eine kompakte Menge $M \subset E$ existiert sodass für alle $i \in I$ gilt:

$$\mu_i(E \setminus K) \le \epsilon$$

Anwendung mit $E = \mathcal{C}_0 = \{v : \mathbb{R}_+ \longrightarrow \mathbb{R} \text{ stetig}, v(0) = 0\}$.
Ziel: Finde geeignetes kompaktes $K \subset \mathcal{C}_0$.

Lemma 3.4.5. *(Arzela-Ascoli)*
Eine Menge $A \subset \mathcal{C}_0$ ist relativ-kompakt genau dann, wenn für alle $T > 0$ gilt:

$$\lim_{t\to\infty} \sup_{\omega\in A} m_T(\omega,\delta) = 0 \tag{3.7}$$

mit Stetigkeitsmodul

$$m_T(\omega,\delta) = \sup_{\substack{s,t\in[0,T]\\ |s-t|<\delta}} |\omega(s) - \omega(t)|\,.$$

Lemma 3.4.6. *Eine Folge $(\mu_n)_n$ von Wahrscheinlichkeitsmaßen auf $\mathcal{C}_0$ ist genau dann relativ kompakt, wenn für alle $T > 0$ und alle $\epsilon >$ o gilt:*

$$\lim_{t\to 0} \sup_n \mu_n \left(\omega : \ m_T(\omega, \delta) > \epsilon\right) = 0$$

Beweis. von "$\Leftarrow$"
Für alle $T > 0$, alle $\eta > 0$ und alle $k \in \mathbb{N}$ wähle $\delta_k > 0$ mit:

$$\sup_n \mu_n \left(\omega : \ m_T(\omega, \delta_w) \leq \frac{1}{k} \ \ \forall k \in \mathbb{N}\right) \leq \eta \cdot 2^{-(T+k)}$$

Definiere $A_T = \left\{\omega : \ m_T(\omega, \delta_k) > \frac{1}{k} \ \ \forall k \in \mathbb{N}\right\}$ und $A = \bigcap_{T\in\mathbb{N}} A_T$. Da $A_T \subset A$ abgeschlossen ist, ist A nach Arzela-Ascoli kompakt und es gilt:

$$\mu_n(A_T) \geq 1 - \eta \sum 2^{-(T+k)} = 1 - \eta 2^{-T} \quad \textit{und} \quad \mu_n(A) \geq 1 - \eta, \quad \textit{uniform in } \eta$$

$\Rightarrow$ *$\{\mu_n\}_n$ ist relativ kompakt in $\mathcal{P}(\mathcal{C}_0)$.* □

Bleibt zu zeigen: die $\mu_n = \left(X^{(n)}\right)_* \mathbb{P}$ erfüllen 3.7.
(Siehe dazu: Karatzas/Shreve: Brownian motion and stochastic calculus)

□

Anwendung des Invarianzprinzips

Korollar 3.4.7. *Sei $\phi : \ \mathcal{C}_0 \longrightarrow \mathbb{R}^d$ stetig, z.B.:*

$$\phi(v) = \left(v(1), \max_{1\leq t\leq 2} v(t), \min_{2\leq s,t\leq 3} |v(s) - v(t)|\right)$$

Dann gilt mit $X^{(n)}$ wie im Invarianzprinzip und einer Brown'schen Bewegung X: Die Verteilung von $\phi(X^{(n)})$ konvergiert gegen die Verteilung von $\phi(x)$.

Beweis. Für alle Funktionen $f \in \mathcal{C}_b(\mathbb{R}^d)$ gilt:

$$\int_{\mathbb{R}^d} f d\phi(X^{(n)})_*\mathbb{P} = \int_{\Omega} f(\phi(X^{(n)}))d\mathbb{P} = \int_{\mathcal{C}_0} \underbrace{f \circ \phi}_{\in\mathcal{C}_b} d\left(X_*^{(n)}\mathbb{P}\right)$$

$$\overset{\textit{Inv Pr.}}{\longrightarrow} \int_{\mathcal{C}_0} f \circ \phi dX_*\hat{\mathbb{P}} = \int_{\mathbb{R}^d} f d\left(\phi(X)_*\hat{\mathbb{P}}\right)$$

□

Korollar 3.4.8. *Das Reflektionsprinzip*

(i) Sei $(X_t)_{t\geq 0}$ eine standardisierte Brown'sche Bewegung und $X_t^* = \sup_{0\leq s\leq t} X_s$. Dann gilt für alle $\alpha > 0$:

$$\mathbb{P}(X_t^* \geq \alpha) = 2\mathbb{P}(X_t \geq \alpha) = 2\int_{\alpha}^{\infty} \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{x^2}{2t}\right) dx$$

(ii) Für jedes $X_n = \sum_{k=1}^n \xi_k$ mit iid ξ_k, $\mathbb{E}[\xi_k] = 0, var(\xi_k) = 1$, gilt mit $X_n^* = \max_{0\leq k\leq n} X_k$:

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{1}{\sqrt{n}} X_n^* \geq \alpha\right) = 2 \lim_{n\to\infty} \mathbb{P}\left(\frac{1}{\sqrt{n}} X_n \geq \alpha\right) = 2 \int_{\alpha}^{\infty} \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{x^2}{2t}\right) dx$$

Beweis. (ii) ist für die symmetrische Irrfahrt bereits in 2.3.1 bewiesen worden und (i) folgt aus (ii) mit dem Invarianzprinzip. □

Index