



universität
wien

DIPLOMARBEIT

Titel der Diplomarbeit

„Regenerative und Harris-rekurrente Markovketten“

Verfasserin

Veronika Huemer-Kals

angestrebter akademischer Grad

Magistra der Naturwissenschaften (Mag.rer.nat.)

Wien, 2013

Studienkennzahl lt. Studienblatt: A 405

Studienrichtung lt. Studienblatt: Diplomstudium Mathematik

Betreuer: Privatdozent Dr. Roland Zweimüller

Danksagung

Meinen herzlichsten Dank an

Herrn Privatdozent Dr. Roland Zweimüller für die großartige Betreuung, das richtige Maß an hilfreicher Unterstützung und Freiraum und für die Bereitschaft, trotz räumlicher Distanz die Betreuung zu übernehmen

meine KollegInnen vom IBO, dafür, dass sich Studium und Arbeit zeitlich immer vereinbaren ließen und auch sonst mein Studienabschluss unterstützt wurde

Herrn Dr. Hans-Peter Ammerer, ohne den ich nicht in der Lage gewesen wäre, diese Diplomarbeit zu schreiben

Angelika Geroldinger und Marina Rehrl für die persönliche, fachliche und LateX-nische Unterstützung

Elisabeth Englmaier, Karina Fallent, Sonja Messeritsch und viele weitere Freundinnen und Freunde für persönliches Anteilnehmen am Fortschritt der Diplomarbeit

meine Eltern Karl und Elfriede und meine Geschwister Kornelia und Severin – für alles.

Inhaltsverzeichnis

1	Grundbegriffe	1
1.1	Einführung	1
1.2	Kerne	1
1.3	Wiederholung aus Maßtheorie	2
2	Markovketten – Definition und grundlegende Eigenschaften	3
2.1	Definition/Konstruktion einer Markovkette	3
2.2	Varianten der Markov-Eigenschaft	9
3	Diskrete Markovketten	11
3.1	Rekurrenz und Transienz	12
3.2	Langzeitverhalten	15
4	Markovketten auf allgemeineren Räumen	23
4.1	Irreduzibilität	23
4.2	Minorisierung	26
4.3	Periodizität	33
4.4	Rekurrenz und Harris-Rekurrenz	34
	Literaturverzeichnis	43
A	Zusammenfassung	45
B	Lebenslauf	47

Kapitel 1

Grundbegriffe

1.1 Einführung

Eine Markovkette ist ein stochastischer Prozess – wir bewegen uns also in einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P_0)$ und einem messbaren Raum $(E, \mathcal{E})$, dem – im einfachsten Fall endlichen oder abzählbaren – Zustandsraum. Zufallsvariablen $(X_n)_{n \geq 0}$ bilden zu jedem Zeitpunkt $n \geq 0$ ersteren auf letzteren ab. In welchem Zustand $x \in E$ man sich zu Beginn des Prozesses befindet, ist durch eine Startverteilung gegeben. Die Wahrscheinlichkeit, von x in eine Menge $A \in \mathcal{E}$ zu gelangen, wird durch einen Kern ausgedrückt.

Der Raum reellwertiger $\mathcal{E}$ - $\mathcal{B}_{\mathbf{R}}$ -messbarer Funktionen wird ebenso wie die σ -Algebra selber mit $\mathcal{E}$ bezeichnet. $\mathcal{M}^+(E)$ sei der Raum der nichttrivialen Maße auf E .

1.2 Kerne

Die folgenden grundlegenden Definitionen und Aussagen finden sich in dieser oder ähnlicher Form etwa in Revuz [1984], S. 8 ff. oder Nummelin [1984], S. 1 f.

Definition. Unter einem *Kern auf E* verstehen wir eine Abbildung $\kappa : E \times \mathcal{E} \rightarrow [0, \infty]$ mit den Eigenschaften

- a) $\kappa(\cdot, A)$ ist messbar $\forall A \in \mathcal{E}$ und
- b) $\kappa(x, \cdot)$ ist ein Maß $\forall x \in E$.

Ein Kern heißt *substochastisch* oder eine *Übergangswahrscheinlichkeit*, wenn

$$\kappa(x, E) \leq 1 \quad \forall x \in E$$

und *stochastisch* oder *Markovkern*, wenn

$$\kappa(x, E) = 1 \quad \forall x \in E$$

Ein Kern κ und eine nichtnegative messbare Funktion f definieren eine weitere nichtnegative messbare Funktion durch

$$\kappa f(x) = \int_E \kappa(x, dy) f(y)$$

Das heißt, κ wirkt als Operator auf $\mathcal{E}_+$, dem Raum der nichtnegativen messbaren Funktionen – und ebenso erhält man aus einem Maß μ auf $\mathcal{E}$ durch

$$\mu\kappa(A) = \int_E \kappa(x, A) \mu(dx)$$

wieder ein Maß. Entspricht f einem Kern $\kappa_2(\cdot, A)$ bzw. μ einem Kern $\kappa_1(x, \cdot)$, dann ist das Produkt zweier (stochastischer) Kerne

$$\kappa_1 \kappa_2(x, A) = \int_E \kappa_1(x, dy) \kappa_2(y, A)$$

wieder ein (stochastischer) Kern. Die Potenzen eines Kerns sind iterativ definiert durch

$$\kappa^n = \kappa \kappa^{n-1} \quad \forall n \geq 1 \quad \text{und} \quad \kappa^0(x, A) = I(x, A) = \mathbf{1}_A(x)$$

Ihre Summe ist der Potenzialkern¹ $G = \sum_{n=0}^{\infty} \kappa^n$.

Im folgenden werden oft die Potenzen einer Übergangswahrscheinlichkeit P betrachtet und dann mit P_n statt P^n bezeichnet.

1.3 Wiederholung aus Maßtheorie

Definition. Maße μ und ν auf $(E, \mathcal{E})$ heißen zueinander singulär², wenn $E = A \cup B$, $A \cap B = \emptyset$, $\mu(A) = 0$ und $\nu(B) = 0$. Schreibweise: $\mu \perp \nu$.

Satz 1.3.1 (Lebesgue-Zerlegung³). *Sind μ, ν zwei σ -endliche Maße, dann kann man ν eindeutig zerlegen in σ -endliche Maße ρ und σ mit $\rho \ll \mu$ und $\sigma \perp \mu$. Analog, wenn ν ein signiertes Maß ist.*

Die folgenden Konvergenzbegriffe für Maße werden in Hernández-Lerma und Lasserre [2003], S. 2–4, eingeführt:

Definition. Sei $\{\mu_n\}$ eine Folge in $\mathcal{M}^+(E)$. $\{\mu_n\}$ konvergiert *mengenweise* gegen μ , wenn

$$\mu_n(A) \rightarrow \mu(A) \quad \forall A \in \mathcal{E}$$

Definition. Die *Totalvariation* $\|\cdot\|_{TV}$ eines signierten Maßes μ auf $(E, \mathcal{E})$ ist gegeben durch $\|\mu\|_{TV} = (\mu^+ + \mu^-)(E)$, wobei $\mu^+(A) = \sup\{\mu(B) \mid B \subseteq A\}$ und $\mu^-(A) = -\inf\{\mu(B) \mid B \subseteq A\}$.

Bemerkung. Die Totalvariation ist eine Norm.

¹ Nummelin [1984], S. 9

² Elstrodt [2005], S. 273

³ Elstrodt [2005], S. 285

Kapitel 2

Markovketten – Definition und grundlegende Eigenschaften

2.1 Definition/Konstruktion einer Markovkette

Markovketten können auf verschiedene Arten definiert werden. Zunächst als Folge von Zufallsvariablen mit einer gewissen Eigenschaft, der (elementaren) Markoveigenschaft⁴:

Definition. Sei $(X_n)_{n \geq 0}$ eine Folge von Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P_0)$ mit Werten im messbaren Raum $(E, \mathcal{E})$. (X_n) heißt *Markovkette*, wenn

$$P_0(X_{n+1} \in A | X_0, \dots, X_n) = P_0(X_{n+1} \in A | X_n) \quad (2.1)$$

und *homogen*, wenn diese Wahrscheinlichkeit nicht von n abhängt.

Im folgenden werden wir ausschließlich homogene Markovketten betrachten.

Durch

$$P(x, A) := P_0(X_{n+1} \in A | X_n = x)$$

ist ein stochastischer Kern auf E gegeben, da die bedingte Wahrscheinlichkeit $P_0(\cdot | X_n = x)$ die entsprechenden Eigenschaften erfüllt. Die Markovkette wird also beschrieben durch ihren Übergangskern – und ihr Startmaß, das Wahrscheinlichkeitsmaß

$$\nu(A) := P_0(X_0 \in A)$$

Man kann aber auch „den Spieß umdrehen“ und von einem stochastischen Kern P auf $(E, \mathcal{E})$ ausgehend eine Markovkette konstruieren.⁵ Als Wahrscheinlichkeitsraum $(\Omega, \mathcal{F})$ wählen wir den unendlichen Produktraum $\Omega = \prod_{n=0}^{\infty} E$ mit der Produkt- σ -Algebra $\mathcal{F} = \bigotimes_{n \geq 0} \mathcal{E}$. Unsere Folge von Zufallsvariablen seien die Koordinatenprojektionen $X_n(\omega) = \omega_n$. Gewünscht ist ein Wahrscheinlichkeitsmaß auf $(\Omega, \mathcal{F})$, das diese Folge zur Markovkette macht.

⁴ Hernández-Lerma und Lasserre [2003], S. 22

⁵ vgl. Revuz [1984], S. 16 ff.

Werfen wir einen genaueren Blick auf die Mengensysteme auf Ω : Für alle n enthält $\mathcal{F} = \sigma(X_0, X_1, \dots)$ die σ -Algebra $\mathcal{F}_n := \sigma(X_0, \dots, X_n)$ der in den ersten n Projektionen enthaltenen Informationen und daher auch deren Vereinigung $\mathcal{A} = \bigcup_{n=0}^{\infty} \mathcal{F}_n$, die Informationen aus endlich vielen Projektionen.

$\mathcal{A}$ ist eine Algebra (wie man einfach nachrechnet), aber keine σ -Algebra – so ist etwa für ein beliebiges $A \in \mathcal{E}$ die Menge

$$\{X_n \in A \text{ oft}\} = \bigcap_{n \geq 1} \bigcup_{m \geq n} \underbrace{\{X_m \in A\}}_{\in \mathcal{F}_m \subseteq \mathcal{A}}$$

nicht in $\mathcal{A}$ enthalten. Die von $\mathcal{A}$ erzeugte σ -Algebra ist schon $\mathcal{F}$, denn einerseits ist $\mathcal{A} \subseteq \mathcal{F}$ und somit $\sigma(\mathcal{A}) \subseteq \sigma(\mathcal{F}) = \mathcal{F}$. Die umgekehrte Inklusion ist ebenfalls leicht zu zeigen:

$$\begin{aligned} X_k^{-1}\mathcal{E} &\subseteq \mathcal{F}_k \subseteq \mathcal{A} \quad \forall k \Rightarrow \\ \bigcup_{k \geq 0} X_k^{-1}\mathcal{E} &\subseteq \mathcal{A} \Rightarrow \\ \underbrace{\sigma\left(\bigcup_{k \geq 0} X_k^{-1}\mathcal{E}\right)}_{\mathcal{F}} &\subseteq \sigma(\mathcal{A}) \end{aligned}$$

Schließlich ist in $\mathcal{A}$ der Halbring der messbaren Rechtecke

$$\mathcal{S} = \left\{ A \in \mathcal{F} : \exists n \in \mathbb{N} : A = \prod_{j=0}^n A_j \times \prod_{j=n+1}^{\infty} E \mid A_j \in \mathcal{E} \right\}$$

enthalten.

Satz 2.1.1. *Für alle $x \in E$ existiert ein eindeutig bestimmtes Wahrscheinlichkeitsmaß⁶ P_x auf $(\Omega, \mathcal{F})$, sodass für beliebige Rechtecke $A = \prod_{j=0}^n A_j \times \prod_{j=n+1}^{\infty} E$ gilt:*

$$\begin{aligned} P_x(A) &= \mathbf{1}_{A_0}(x) \int_{A_1} P(x, dx_1) \int_{A_2} P(x_1, dx_2) \dots \\ &\quad \int_{A_{n-1}} P(x_{n-2}, dx_{n-1}) P(x_{n-1}, A_n) \end{aligned} \tag{2.2}$$

Beweis. Die Mengenfunktion P_x entspricht, auf den messbaren Rechtecken der Dimension n betrachtet, den in Klenke [2008] definierten Produkten⁷

$$\epsilon_x \otimes \bigotimes_{i=1}^n P$$

und nach Korollar 14.24 (ebd.) ist dieses Produkt ein Wahrscheinlichkeitsmaß auf dem endlichen Produktraum $(\prod_{i=0}^n E, \bigotimes_{i=0}^n \mathcal{E})$. Letzteres ist also die eindeutig bestimmte

⁶ Revuz [1984], S. 17, Theorem 2.8

⁷ Klenke [2008], Satz 14.22 sowie Korollar 14.23

Fortsetzung⁸ von $P_x^{(n)} \equiv P_x$ auf die Produkt- σ -Algebra $\mathcal{F}_n = \bigotimes_{i=0}^n \mathcal{E}$, welche ja auch von den messbaren Rechtecken erzeugt wird. Die Familie dieser Maße ist konsistent, das heißt $P_x^{(n+1)}|_{\mathcal{F}_n} = P_x^{(n)}$. Somit ist P_x auf ganz $\mathcal{A}$ eindeutig festgelegt.

Fortsetzungssatz und Eindeutigkeitssatz⁹ garantieren, dass ein σ -endliches Prämaß auf einem Halbring eindeutig auf die erzeugte σ -Algebra (hier $\sigma(\mathcal{A}) = \mathcal{F}$) fortgesetzt werden kann.

Zuerst definieren wir, analog zu P_x , die Mengenfunktion $P_{(x_0, \dots, x_m)}^m$ auf dem zu Ω isomorphen Raum $\Omega^{(m)} = \prod_{k=m}^{\infty} E$. Für eine $\mathcal{A}$ -Menge A mit einem festgelegten Anfangsstück

$$A(x_0, \dots, x_{m-1}) := \left\{ \omega^{(m)} \in \Omega^{(m)} : (x_0, \dots, x_{m-1}, \omega^{(m)}) \in A \right\}$$

ist $P_{(x_0, \dots, x_m)}^m$ definiert durch

$$P_{(x_0, \dots, x_m)}^m(A(x_0, \dots, x_{m-1})) = \delta_{(x_0, \dots, x_m)} \otimes (P \otimes \dots \otimes P)(A)$$

Falls A ein messbares Rechteck ist, lässt sich $P_x(A)$ schreiben als

$$\begin{aligned} P_x(A) &= \mathbf{1}_{A_0}(x) \underbrace{\int_{A_1} P(x, dx_1)}_{\int_E P(x, dx_1) \mathbf{1}_{A_1}(x_1)} \int_{A_2} P(x_1, dx_2) \dots \int_{A_n} P(x_{n-1}, dx_n) = \\ &= \int_E P(x, dx_1) P_{(x, x_1)}^1(A(x)) \end{aligned} \quad (2.3)$$

Für $A \in \mathcal{A}$ beliebig gilt diese Identität ebenfalls: A ist dann in einer σ -Algebra $\mathcal{F}_n$ enthalten. Die linke Seite in (2.3), $P_x(A)$, ist wie bereits erwähnt für beliebiges x ein Maß auf $\mathcal{F}_n$, ebenso die rechte Seite, was sich dank der σ -Additivität von $P_{(x, x_1)}^1$ leicht nachweisen lässt. Auf den messbaren Rechtecken $\left\{ \prod_{j=0}^n A_j \times \prod_{j=n+1}^{\infty} E : A_j \in \mathcal{E} \right\}$ stimmen diese Mengenfunktionen überein, also auch auf der erzeugten σ -Algebra $\mathcal{F}_n$.

Es bleibt die σ -Additivität von P_x auf $\mathcal{A}$ zu zeigen. Im Fall einer Algebra mit einer endlich additiven Mengenfunktion, die dem ganzen Raum einen endlichen Wert zuweist, wird die σ -Additivität der Mengenfunktion durch ihre Stetigkeit in $\emptyset$ charakterisiert.¹⁰

Wir nehmen nun eine Folge (A_j) in $\mathcal{A}$ mit $(A_j) \searrow \emptyset$ und Grenzwert $\lim_j P_x(A_j) > 0$ an. Wegen (2.3) und aufgrund des Lemmas von Fatou gilt

$$\begin{aligned} 0 < \lim_j P_x(A_j) &= \lim_j \int_E P(x, dx_1) P_{(x, x_1)}^1(A_j(x)) \\ &\leq \int_E P(x, dx_1) \lim_j \underbrace{P_{(x, x_1)}^1(A_j(x))}_{\in [0,1]} \end{aligned}$$

⁸ siehe z.B. Elstrodt [2005], S. 61

⁹ ebd.

¹⁰ Elstrodt [2005], S. 32, Satz 1.10

woraus wir schließen, dass es ein $\overline{x_1} \in E$ gibt mit

$$\lim_j P_{(x, \overline{x_1})}^1(A_j(x)) = \inf_j P_{(x, \overline{x_1})}^1(A_j(x)) > 0$$

die Mengen $A_j(x)$ also für alle j nicht leer sind. Auf dieselbe Weise können wir zu $P_{(x, \overline{x_1})}^1(A_j(x))$ ein $\overline{x_2}$ finden mit $\lim_j P_{(x, \overline{x_1}, \overline{x_2})}^2(A_j(x, \overline{x_1})) > 0$. So konstruieren wir eine Folge $(\overline{x_1}, \overline{x_2}, \overline{x_3}, \dots)$ mit

$$A_j(x, \overline{x_1}, \overline{x_2}, \overline{x_3}, \dots, \overline{x_n}) \neq \emptyset \quad \forall j \quad \forall n$$

das heißt

$$\forall j \exists \omega^{(n+1)} \in \Omega^{(n+1)} : (x, \overline{x_1}, \overline{x_2}, \overline{x_3}, \dots, \overline{x_n}, \omega^{(n+1)}) \in A_j$$

Die Mengen A_j stammen aus der Algebra $\mathcal{A}$ und hängen daher nur von endlich vielen Ereignissen ab. Das bewirkt, dass für jedes $j \geq 1$ ab einem gewissen Index k_j alle Folgen $(x, \overline{x_1}, \overline{x_2}, \overline{x_3}, \dots, \overline{x_{k_j}}, \omega^{(k_j+1)})$ in A_j liegen. Daher ist $(x, \overline{x_1}, \overline{x_2}, \overline{x_3}, \dots, \overline{x_n}, \dots)$ in allen A_j enthalten – ein Widerspruch zur Annahme, der Schnitt der Folge (A_j) wäre leer. $\square$

Satz 2.1.2. *Die Abbildung $x \mapsto P_x(A)$ ist $\mathcal{E}$ -messbar für alle $A \in \mathcal{F}$.¹¹*

Beweis. Für den Beweis nützen wir den Satz über monotone Klassen. Die Menge $\mathcal{B} := \{A \in \mathcal{F} : x \mapsto P_x(A) \text{ messbar}\}$ ist eine solche¹² und damit identisch mit der von ihr erzeugten monotonen Klasse $\mathcal{S}(\mathcal{B})$.

Messbare Rechtecke $A = \prod_{j=0}^n A_j \times \prod_{j=n+1}^\infty E$ sind in der monotonen Klasse $\mathcal{B}_n := \{A \in \mathcal{F}_n : x \mapsto P_x(A) \text{ messbar}\}$ enthalten. Offensichtlich ist $\mathcal{B}_n \subseteq \mathcal{F}_n$. Umgekehrt ist nach dem Satz über monotone Klassen¹³ – und aufgrund der Durchschnittstabilität der messbaren Rechtecke – die von ihnen erzeugte σ -Algebra $\mathcal{F}_n$ gleich der erzeugten monotonen Klasse. Deshalb gilt auch $\mathcal{F}_n \subseteq \mathcal{B}_n$ für alle n und somit $\mathcal{A} \subseteq \mathcal{B}$.

Da $\mathcal{A}$ durchschnittstabil ist, können wir nochmals den Satz über monotone Klassen anwenden und erhalten

$$\mathcal{F} = \sigma(\mathcal{A}) = \mathcal{S}(\mathcal{A}) \subseteq \mathcal{S}(\mathcal{B}) = \mathcal{B}$$

$\square$

Zusammenfassend erhalten wir folgenden

Satz 2.1.3. *Sei $(E, \mathcal{E})$ ein messbarer Raum mit einem stochastischen Kern P und einem Wahrscheinlichkeitsmaß ν . Dann existiert auf dem unendlichen Produktraum*

¹¹ Revuz [1984], S. 17, Theorem 2.8

¹² Die definierenden Eigenschaften aus Revuz [1984], S. 4, Def. 3.1, sind einfach nachzuweisen.

¹³ Revuz [1984], S. 4, Satz 3.2

$\Omega = \prod_{n=0}^{\infty} E$ mit der Produkt- σ -Algebra $\mathcal{F}$ eine homogene Markovkette, die eindeutig bestimmte kanonische Markovkette¹⁴ mit Übergangswahrscheinlichkeit P und Startmaß ν .

Beweis. Die gewünschte Markovkette erhalten wir, indem wir für ν ein – wegen Satz 2.1.2 wohldefiniertes – Maß P_ν auf Ω festlegen durch

$$P_\nu(A) = \int_E P_x(A) \nu(dx)$$

denn bezüglich P_ν erfüllt die Folge der Koordinatenprojektionen (X_n) die Markoveigenschaft

$$\underbrace{P_\nu(X_{n+1} \in A_{n+1} | X_0, \dots, X_n)}_{\mathbf{E}_\nu[\mathbf{1}_{A_{n+1}}(X_{n+1}) | \mathcal{F}_n]} = \underbrace{P_\nu(X_{n+1} \in A_{n+1} | X_n)}_{\mathbf{E}_\nu[\mathbf{1}_{A_{n+1}}(X_{n+1}) | \sigma(X_n)]}$$

Um dies zu beweisen, zeigen wir erst

$$P_\nu(X_{n+1} \in A_{n+1} | X_0, \dots, X_n) = P(x_n, A_{n+1})$$

O.B.d.A. sei $P_\nu = P_{x_0}$ für ein $x_0 \in E$. Für messbare Rechtecke gilt

$$\begin{aligned} P_{x_0}(A_0 \times \dots \times A_{n+1}) &= \left(\epsilon_{x_0} \otimes \bigotimes_{i=1}^{n+1} P \right) (A_0 \times \dots \times A_{n+1}) = \\ &= \int_{A_0 \times \dots \times A_n} \left(\epsilon_{x_0} \otimes \bigotimes_{i=1}^n P \right) (dx_0, \dots, dx_n) \underbrace{P((x_0, \dots, x_n), A_{n+1})}_{P(x_n, A_{n+1})} = \\ &= \int_{A_0 \times \dots \times A_n} P_{x_0}(dx_0, \dots, dx_n) P(x_n, A_{n+1}) = \\ &= \int_A P(x_n, A_{n+1}) P_{x_0}(dx) \end{aligned}$$

wobei $A = A_0 \times \dots \times A_n \in \mathcal{F}_n$ ein messbares Rechteck und $x = (x_0, \dots, x_n)$ ist. Anders ausgedrückt, haben wir

$$\underbrace{\int_A \mathbf{1}_{\{X_{n+1} \in A_{n+1}\}} dP_{x_0}}_{P_{x_0}(A_0 \times \dots \times A_{n+1})} = \int_A P(x_n, A_{n+1}) dP_{x_0}$$

Mittels des Fortsetzungssatzes kann man diese Eigenschaft auf beliebige $A \in \mathcal{F}_n$ übertragen. Da $P(x_n, A_{n+1})$ auch messbar bezüglich $\mathcal{F}_n$ ist, erfüllt es die Eigenschaften der bedingten Erwartung, und die Behauptung ist bewiesen.

¹⁴ Revuz [1984], S. 18, Proposition 2.10

Nun ist $P(x_n, A_{n+1})$ sogar messbar bezüglich $\sigma(X_n)$, und

$$\int_A \mathbf{1}_{\{X_{n+1} \in A_{n+1}\}} dP_{x_0} = \int_A P(x_n, A_{n+1}) dP_{x_0}$$

gilt insbesondere für $A \in \sigma(X_n) \subseteq \mathcal{F}_n$. Daraus folgt

$$P(x_n, A_{n+1}) = P_{x_0}(X_{n+1} \in A_{n+1} | X_n)$$

Die Übergangswahrscheinlichkeit ist wie gewünscht durch den Kern gegeben und die Markovkette homogen.

Schließlich gilt noch

$$P_\nu(X_0 \in A_0) = \int_E \underbrace{P_x(A_0)}_{\mathbf{1}_{A_0}(x)} \nu(dx) = \nu(A_0)$$

□

Eine alternative Definition einer Markovkette findet sich in Revuz [1984]¹⁵ und ist äquivalent zu (2.1)¹⁶:

Definition. Der stochastische Prozess $X = (X_n)_{n \geq 0}$ heißt Markovkette, wenn für alle $A \in \mathcal{F}_n = \sigma(X_1, \dots, X_n)$ und alle $B \in \sigma(X_m : m \geq n)$

$$P_0(A \cap B | X_n) = P_0(A | X_n) P_0(B | X_n)$$

fast sicher (f.s.) gilt – also die vergangene und die zukünftige Entwicklung der Kette unabhängig voneinander sind.

Schließlich können Markovketten wie in Hernández-Lerma und Lasserre [2003]¹⁷ über eine Zustandsgleichung definiert werden:

Definition. Eine Markovkette $X = (X_n)_{n \geq 0}$ ist ein dynamisches System, das eine zeithomogene Zustandsgleichung

$$X_{n+1} = F(X_n, \psi_n)$$

erfüllt. Hier ist $F : E \times Y \rightarrow E$ messbar und $\{\psi_n\}$ eine Folge unabhängiger, identisch verteilter Y -wertiger Zufallsvariablen, unabhängig vom Anfangszustand X_0 , auf einem messbaren Raum Y .

Beispiel. a) Ein einfaches Beispiel für eine Markovkette ist ein Prozess mit Werten in einem abzählbaren Raum¹⁸ E , dessen Übergangswahrscheinlichkeiten durch eine Matrix $(p(x, y))$ gegeben sind: $P(x, \{y\}) = p(x, y)$.

¹⁵ Revuz [1984], S. 14, Definition 2.1

¹⁶ Athreya und Lahiri [2006], Proposition 14.2.1

¹⁷ Hernández-Lerma und Lasserre [2003], S. 23

¹⁸ siehe Kapitel 3

- b) Sei (z_n) eine Familie von endlichen, unabhängigen, identisch verteilten Zufallsvariablen und Z_0 eine von (z_n) unabhängige, endliche Zufallsvariable. Der durch $Z_n = Z_0 + \sum_{k=1}^n z_k$ definierte Prozess (Z_n) wird eine *Irrfahrt* genannt und ist ebenfalls eine homogene Markovkette.

Bemerkung (Definition bezüglich allgemeiner Filtration). Sei $(\mathcal{G}_n)$ eine aufsteigende Folge von Teil- σ -Algebren von $\mathcal{F}$. Dann definiert $X = (X_n)$ mit den Eigenschaften

- a) X ist adaptiert an $(\mathcal{G}_n)$ und
 b) $P_0(X_{n+1} \in A | \mathcal{G}_n) = P_0(X_{n+1} \in A | X_n)$

eine Markovkette bezüglich $(\mathcal{G}_n)$.¹⁹

Bemerkung. Ist die Übergangswahrscheinlichkeit P nicht stochastisch auf E , so erweitert man E um den „Friedhofspunkt“ $\Delta \notin E$ und setzt²⁰

$$\begin{aligned} P(x, \{\Delta\}) &= 1 - P(x, E) \quad \forall x \in E \\ P(\Delta, \{\Delta\}) &= 1 \end{aligned}$$

wodurch P zu einem stochastischen Kern auf $(E_\Delta, \mathcal{E}_\Delta) = (E \cup \{\Delta\}, \sigma(E, \{\Delta\}))$ wird.

2.2 Varianten der Markov-Eigenschaft

Definition. Der *Shift-Operator*²¹ ist die Abbildung $\theta : \Omega \rightarrow \Omega$ mit $(\omega_0, \omega_1, \dots) \mapsto (\omega_1, \omega_2, \dots)$. Wir schreiben θ_n für $\underbrace{\theta \circ \dots \circ \theta}_{n \text{ mal}}$.

Satz 2.2.1 (Markov-Eigenschaft, zweite Version). *Sei Z eine nichtnegative Zufallsvariable auf $(\Omega, \mathcal{F})$. Dann gilt²²*

$$\mathbf{E}_\nu[Z \circ \theta_n | \mathcal{F}_n] = \mathbf{E}_{X_n}[Z] \quad P_\nu\text{-f.s.}$$

Bemerkung. Wie schon im Beweis von Satz 2.1.3 verwenden wir die Schreibweise $\mathbf{E}_\nu$ für den Erwartungswert bezüglich P_ν , insbesondere $\mathbf{E}_x$ für den Erwartungswert bezüglich P_x .

Beweis. Falls Z von der Form $Z = \mathbf{1}_{\{X_0 \in A_0, \dots, X_m \in A_m\}}$ ist, gilt der Satz. Von diesen Zufallsvariablen kann man das Ergebnis leicht auf Treppenfunktionen erweitern und allgemeine nichtnegative messbare Funktionen wie üblich durch letztere approximieren. $\square$

Definition. Eine *Zufallszeit* ist eine Zufallsvariable $T : \Omega \rightarrow \mathbf{N} \cup \{\infty\}$. Unter einer *Stoppzeit* für die Markovkette (X_n) versteht man eine Zufallszeit mit $\{T = n\} \in \mathcal{F}_n$ für alle n .²³

¹⁹ Nummelin [1984], Definition 1.2

²⁰ Nummelin [1984], S. 6

²¹ Revuz [1984], S. 19, Definition 2.12

²² Revuz [1984], S. 20, Proposition 2.13

²³ Nummelin [1984], S. 31

Zu T definieren wir²⁴

a) die σ -Algebra $\mathcal{F}_T := \{A \in \mathcal{F} : A \cap \{T = n\} \in \mathcal{F}_n \forall n\}$.

b) die Zufallsvariable X_T , die die Position zur Zeit T angibt:

$$X_T(\omega) = \begin{cases} X_n(\omega) & \text{falls } T(\omega) = n \\ \Delta & \text{falls } T(\omega) = \infty \end{cases}$$

c) den zufälligen Pfad θ_T : $\theta_T(\omega) = \begin{cases} \theta_n(\omega) & \text{falls } T(\omega) = n \\ \omega_\Delta = \{\Delta, \Delta, \dots\} & \text{falls } T(\omega) = \infty \end{cases}$

Beispiel. Für $A \in \mathcal{E}$ ist $T_A = \inf \{n \geq 1 : X_n \in A\}$, die *Eintrittszeit* in A , eine Stoppzeit (Konvention: $\inf \emptyset = \infty$).

Satz 2.2.2 (Starke Markoveigenschaft²⁵). *Sei T eine Stoppzeit für (X_n) , Z eine nicht-negative Zufallsvariable. Dann gilt*

$$\mathbf{E}_\nu [Z \circ \theta_T | \mathcal{F}_T] = \mathbf{E}_{X_T} [Z] \quad P_\nu\text{-f.s.}$$

auf $\{T < \infty\}$.

Beweis. Für alle $A \in \mathcal{F}_T$ soll

$$\int_A Z \circ \theta_T \cdot \mathbf{1}_{\{X_T \neq \Delta\}} dP_\nu = \int_A \mathbf{E}_{X_T} [Z] dP_\nu$$

oder äquivalent

$$\sum_{n \geq 0} \int_{A \cap \{T=n\}} Z \circ \theta_T \cdot \mathbf{1}_{\{X_n \neq \Delta\}} dP_\nu = \sum_{n \geq 0} \int_{A \cap \{T=n\}} \mathbf{E}_{X_n} [Z] dP_\nu$$

gelten. Da $A \in \mathcal{F}_T$, liegt $A \cap \{T = n\}$ in $\mathcal{F}_n$, und es genügt zu zeigen, dass für beliebige $B \in \mathcal{F}_n$

$$\int_B Z \circ \theta_n \cdot \mathbf{1}_{\{X_n \in E\}} dP_\nu = \int_B \mathbf{E}_{X_n} [Z] dP_\nu \quad \forall n \geq 0$$

was genau der Markoveigenschaft aus Satz 2.2.1 entspricht. □

²⁴ Revuz [1984], S. 23 f.

²⁵ Revuz [1984], S. 24, Theorem 3.5

Kapitel 3

Diskrete Markovketten

In diesem Abschnitt halten wir einige Resultate über Markovketten mit endlichem oder abzählbarem Zustandsraum E fest.²⁶

Die Übergangswahrscheinlichkeiten entsprechen den Einträgen einer Matrix $P = (p_{ij})$, wobei i und j Zustände in E sind.

Bemerkung. a) Die Einträge der Matrix P sind Wahrscheinlichkeiten, also nichtnegativ. Die Zeilensummen ergeben 1, denn $\sum_{j \in E} p_{ij} = P(i, \cup_{j \in E} \{j\}) = P(i, E)$. Eine Matrix mit diesen Eigenschaften wird als *stochastische Matrix* bezeichnet.

b) Es gilt $P_n(i, \{j\}) = p_{ij}^{(n)}$, dem entsprechenden Eintrag der Matrix P^n .

Beispiel. a) Sei $E = \mathbf{N}_+$ mit den Übergangswahrscheinlichkeiten

$$p_{ij} = \begin{cases} 1 & \text{für } j = i - 1 \\ 0 & \text{sonst} \end{cases}$$

für alle $i > 1$ und p_{1j} gegeben durch eine beliebige Wahrscheinlichkeitsverteilung $(p_{1j})_{j \geq 1}$. Die Markovkette springt von 1 in einen Zustand $j \in \mathbf{N}_+$ und wandert in Einserschritten schließlich wieder nach 1 zurück.

b) (Irrfahrt auf $\mathbf{Z}$). Sei $(p_{0j})_{j \in \mathbf{Z}}$ eine beliebige Wahrscheinlichkeitsverteilung. Die Übergangswahrscheinlichkeiten

$$p_{ij} = p_{0, j-i}$$

beschreiben eine Irrfahrt, dh. die Summe von unabhängigen, identisch (gemäß (p_{0j})) verteilten Zufallsvariablen.

Definition. Der Zustand j heißt vom Zustand i aus *erreichbar* (symbolisch $i \rightarrow j$), wenn $\exists n \geq 1 : p_{ij}^{(n)} > 0$. Gelten $i \rightarrow j$ und $j \rightarrow i$, so *kommunizieren* i und j . Eine *abgeschlossene Teilmenge* F ist eine nichtleere $\mathcal{E}$ -Menge mit $i \nrightarrow j$ für alle $i \in F, j \in F^c$. Wenn ein abgeschlossenes F keine echte abgeschlossene Teilmenge enthält, spricht man von einer *minimalen* Menge.

Mit der *Irreduzibilität* einer diskreten Markovkette ist gemeint, dass alle $i, j \in E$ kommunizieren, somit E die einzige abgeschlossene Menge ist. Dies entspricht Irreduzibilität

²⁶ Die hier verwendete Notation und grundlegende Definitionen finden sich in Hofbauer [2002], Athreya und Lahiri [2006] bzw. Hernández-Lerma und Lasserre [2003].

bezüglich des Zählmaßes gemäß Abschnitt 4.1.

Die *Periode* d_i eines Zustands i ist der größte gemeinsame Teiler der Menge $\{n \geq 1 : p_{ii}^{(n)} > 0\}$. Alle Zustände einer minimalen Teilmenge haben die gleiche Periode, insbesondere haben alle Zustände die gleiche Periode d , wenn P irreduzibel ist.²⁷

Die Besuchszeiten im Zustand i werden bezeichnet mit

$$T_i = T_i^{(1)} = \inf \{n > 0 : X_n = i\} \text{ bzw.}$$

$$T_i^{(k)} = \inf \{n > T_i^{(k-1)} : X_n = i\}, T_i^{(0)} := 0$$

die Anzahl der Besuche bis zum Zeitpunkt n mit

$$N_n(i) = \# \{k : 0 \leq k \leq n : X_k = i\}$$

und die Gesamtanzahl als $N(i) = \sum_{n \geq 0} \mathbf{1}_{\{i\}}(X_n)$.

Weiters definieren wir $f_{ij}^{(n)} = P_0(X_n = j, X_{n-1} \neq j, \dots, X_1 \neq j | X_0 = i)$ und deren Summe $f_{ij} = \sum_{n \geq 1} f_{ij}^{(n)}$. Dies sind die Wahrscheinlichkeiten, von i ausgehend zum Zeitpunkt n das erste Mal bzw. zu irgendeinem Zeitpunkt n in j zu landen – das heißt, dass $T_j < \infty$ ist.

3.1 Rekurrenz und Transienz

Definition. Der Zustand i heißt *rekurrent*, wenn $f_{ii} = 1$ und *transient*, wenn $f_{ii} < 1$. Äquivalent dazu kann man $\sum_{n \geq 0} p_{ii}^{(n)} = \infty$ für Rekurrenz bzw. $\sum_{n \geq 0} p_{ii}^{(n)} < \infty$ für Transienz verlangen.

Ein rekurrenter Zustand i heißt *nullrekurrent*, wenn $\mathbf{E}_i[T_i] = \infty$ und *positiv rekurrent*, wenn $\mathbf{E}_i[T_i] < \infty$.

Bemerkung. Die in der Definition erwähnte Äquivalenz wird in Satz 3.1.2 mithilfe der folgenden Proposition bewiesen. Die Summe $\sum_{n \geq 0} p_{ii}^{(n)}$ entspricht der erwarteten Anzahl an Besuchen in i bei Start in i :²⁸

$$\mathbf{E}_i[N(i)] = \sum_{n \geq 0} \mathbf{E}_i[\mathbf{1}_{\{i\}}(X_n)] = \sum_{n \geq 0} p_{ii}^{(n)}$$

Proposition 3.1.1. Für beliebige i und j in E gilt²⁹

$$P_i(T_j^{(k)} < \infty) = f_{ij} f_{jj}^{k-1} \quad \forall k \geq 1$$

Beweis. Für $k = 1$ gilt die Aussage nach Definition von f_{ij} . Unter Verwendung der

²⁷ Hofbauer [2002], S. 13, Satz 15

²⁸ vgl. Hernández-Lerma und Lasserre [2003], S. 43

²⁹ Klenke [2008], Satz 17.29

starken Markoveigenschaft und der Induktionsannahme für $k-1$ zeigen wir den Induktionsschritt nach k :

$$\begin{aligned}
 P_i \left(T_j^{(k)} < \infty \right) &= \mathbf{E}_i \left[\mathbf{1}_{\{T_j^{(k)} < \infty\}} \right] = \mathbf{E}_i \left[\mathbf{1}_{\{T_j^{(k-1)} < \infty\}} \cdot \mathbf{1}_{\{T_j^{(k)} < \infty\}} \right] = \\
 &= \int_{\{T_j^{(k-1)} < \infty\}} \mathbf{1}_{\{T_j^{(k)} < \infty\}} dP_i = \int_{\{T_j^{(k-1)} < \infty\}} \mathbf{E}_i \left[\mathbf{1}_{\{T_j^{(k)} < \infty\}} | \mathcal{F}_{T_j^{(k-1)}} \right] dP_i = \\
 &= \int_{\{T_j^{(k-1)} < \infty\}} \underbrace{\mathbf{E}_i \left[\mathbf{1}_{\{T_j^{(1)} < \infty\}} \circ \theta_{T_j^{(k-1)}} | \mathcal{F}_{T_j^{(k-1)}} \right]}_{= \mathbf{E}_{X_{T_j^{(k-1)}}} \left[\mathbf{1}_{\{T_j^{(1)} < \infty\}} \right]} dP_i = \int_{\{T_j^{(k-1)} < \infty\}} \underbrace{P_j \left(T_j^{(1)} < \infty \right)}_{= f_{jj}} dP_i = \\
 &= f_{jj} \cdot P_i \left(T_j^{(k-1)} < \infty \right) = f_{ij} f_{jj}^{k-1}
 \end{aligned}$$

□

Satz 3.1.2. *Ein $i \in E$ ist rekurrent genau dann³⁰, wenn $\sum_{n \geq 0} p_{ii}^{(n)} = \infty$, folglich transient genau dann, wenn $\sum_{n \geq 0} p_{ii}^{(n)} < \infty$.*

Beweis. Wie schon für den Fall $i = j$ erwähnt gibt $\sum_{n \geq 0} p_{ij}^{(n)}$ die erwartete Gesamtanzahl der Besuche in j bei Start in i an:

$$\begin{aligned}
 \sum_{n \geq 0} p_{ij}^{(n)} &= \mathbf{E}_i \left[\sum_{n \geq 0} \mathbf{1}_{\{j\}} (X_n) \right] = \sum_{k \geq 1} P_i \left(\sum_{n \geq 0} \mathbf{1}_{\{j\}} (X_n) \geq k \right) = \\
 &= \sum_{k \geq 1} \left(P_i \left(i = j, T_j^{(k-1)} < \infty, T_j^{(k)} = \infty \right) + P_i \left(T_j^{(k)} < \infty \right) \right) = \\
 &= P_i \left(\{i = j\} \cap \underbrace{\left(\bigcup_{k \geq 1} \left(\{T_j^{(k-1)} < \infty\} \cap \{T_j^{(k)} = \infty\} \right) \right)}_{\Omega} \right) + \sum_{k \geq 1} P_i \left(T_j^{(k)} < \infty \right) = \\
 &= \mathbf{1}_{\{i=j\}} + \sum_{k \geq 1} P_i \left(T_j^{(k)} < \infty \right) = \\
 &= \mathbf{1}_{\{i=j\}} + \sum_{k \geq 1} f_{ij} f_{jj}^{k-1}
 \end{aligned}$$

wobei im letzten Schritt die vorangegangene Proposition benützt wurde. Falls $i = j$,

³⁰ Klenke [2008], Satz 17.34

ergibt sich

$$\sum_{n \geq 0} p_{ii}^{(n)} = 1 + \sum_{k \geq 1} f_{ii}^k = \frac{1}{1 - f_{ii}}$$

und mit der Konvention $\frac{1}{0} := \infty$ folgt die Aussage. $\square$

Proposition 3.1.3 (Solidaritätseigenschaft). *Sei i rekurrent und $i \rightarrow j$. Dann ist $f_{ji} = 1$ und j ebenfalls rekurrent.*

Beweis. Siehe Athreya und Lahiri [2006], Proposition 14.1.7. $\square$

Satz 3.1.4. *a) In einer minimalen Teilmenge kommunizieren alle Zustände.³¹*

b) Kommunizierende Zustände haben die gleiche Klassifikation (transient, null- oder positiv rekurrent).

c) Minimale Teilmengen sind entweder paarweise disjunkt oder gleich.³²

d) Liegt ein Zustand in keiner minimalen Teilmenge, so ist er transient.³³

Beweis. Für den Beweis setzen wir $E(i) := \{j \in E : i \rightarrow j\}$. Diese Menge ist abgeschlossen, denn wäre $i \rightarrow j$ und $j \rightarrow k$ für ein $k \notin E(i)$, dann gilt widersprüchlicherweise auch $i \rightarrow k$. Liegt i in einer abgeschlossenen Menge F , dann sind alle $j \in E(i)$ auch in F enthalten (sonst $i \nrightarrow j$), also ist $E(i)$ die kleinste abgeschlossene Menge, die i enthält. Für Zustände i und j in einer minimalen Teilmenge F impliziert $E(i) = F = E(j)$, dass $i \rightarrow j$ und vice versa.

Wegen der Solidaritätseigenschaft (Proposition 3.1.3) sind zwei kommunizierende Zustände beide rekurrent oder beide transient. Die entsprechende Aussage über Null- bzw. positive Rekurrenz folgt aus Proposition 3.2.5 unten.

Nehmen wir zwei minimale Teilmengen F_1 und F_2 mit einem $i \in F_1 \cap F_2$ an. Dann gilt $E(i) \subseteq F_1 \cap F_2$, da der Schnitt ebenfalls eine abgeschlossene Menge ist. Wegen der Minimalität gilt aber auch $F_1 \subseteq E(i)$ und $F_2 \subseteq E(i)$, also Gleichheit.

Wenn ein Zustand i in keiner minimalen Teilmenge liegt, gibt es eine abgeschlossene echte Teilmenge $F \subset E(i)$ mit $i \notin F$. Für $j \in F$ beliebig gilt $i \rightarrow j$ wegen $F \subset E(i)$ und $j \nrightarrow i$, weil F abgeschlossen ist. Deshalb sind alle $f_{ji}^{(n)} \leq p_{ji}^{(n)} = 0$, daraus folgt $f_{ji} = 0$. Aufgrund der Solidaritätseigenschaft ist i transient. $\square$

Man kann den Zustandsraum also in disjunkte minimale Teilmengen (jeweils transient, positiv rekurrent oder nullrekurrent) sowie transiente Zustände außerhalb minimaler Teilmengen einteilen.

Im Spezialfall einer irreduziblen Markovkette mit einem positiv rekurrenten Zustand sind alle Zustände positiv rekurrent; man nennt die Markovkette dann *ergodisch*. Ist

³¹ Hofbauer [2002], S. 8, Satz 8

³² Hofbauer [2002], S. 8, Satz 9

³³ Hofbauer [2002], S. 12, Satz 14

der Zustandsraum endlich, so existiert zumindest ein positiv rekurrenter Zustand.³⁴ Eine irreduzible Markovkette auf einem endlichen Zustandsraum ist daher ergodisch.

3.2 Langzeitverhalten

Definition. Wir fassen die Projektionen zwischen den Besuchszeiten in einem Zustand i zusammen und bezeichnen die Mengen

$$\eta_k = \left\{ X_n : T_i^{(k)} \leq n \leq T_i^{(k+1)} - 1 \right\}$$

als *Zyklen* oder *Exkursionen*.

Satz 3.2.1. Sei i rekurrent. Unter dem Maß P_i sind die Zyklen $(\eta_k)_{k \geq 0}$ unabhängige und identisch verteilte Zufallsvektoren mit zufälliger Länge.³⁵ Das bedeutet für beliebige Vektoren $(x_{k1}, \dots, x_{km_k}, m_k)$ in $E^{m_k} \times \mathbf{N}$ für alle $n \in \mathbf{N}$

$$P_i \left(\eta_k = (x_{k1}, \dots, x_{km_k}), T_i^{(k+1)} - T_i^{(k)} = m_k, 0 \leq k \leq n \right) = \prod_{k=0}^n P_i \left(\eta_0 = (x_{k1}, \dots, x_{km_k}), T_i^{(1)} = m_k \right)$$

Beweis. Sei $n \in \mathbf{N}$ beliebig. Unterteilen wir die Markovkette in den Teil vor und nach der n . Besuchszeit in i und setzen

$$A := \left\{ \eta_k = (x_{kl})_{l=1}^{m_k}, T_i^{(k+1)} - T_i^{(k)} = m_k \text{ für } 0 \leq k \leq n-1 \right\} \text{ sowie}$$

$$Z := \mathbf{1}_{\{X_0=x_{n1}, \dots, X_{m_n-1}=x_{nm_n}, X_{m_n}=i\}}$$

Da $A \in \mathcal{F}_{T_i^{(n)}}$, folgt mit der starken Markoveigenschaft

$$\begin{aligned} P_i \left(\eta_k = (x_{kl})_{l=1}^{m_k}, T_i^{(k+1)} - T_i^{(k)} = m_k, 0 \leq k \leq n \right) &= \\ P_i \left(\left\{ \eta_n = (x_{nl})_{l=1}^{m_n}, T_i^{(n+1)} - T_i^{(n)} = m_n \right\} \cap A \right) &= \\ \int_A Z \circ \theta_{T_i^{(n)}} dP_i &= \int_A \mathbf{E}_i[Z] dP_i = \\ P_i \left(\eta_0 = (x_{nl})_{l=1}^{m_n}, T_i^{(1)} = m_n \right) \cdot P_i(A) & \end{aligned}$$

Genauso verfährt man mit A bzw. mit allen weiteren Besuchszeiten vor n . □

Satz 3.2.1 ermöglicht es, aus den Zyklen Schlüsse über das Verhalten der gesamten Markovkette zu ziehen: Da diese unabhängig und identisch verteilt sind, steht uns das

³⁴ vgl. Hofbauer [2002], S. 18, Satz 19

³⁵ Athreya und Lahiri [2006], Theorem 14.1.3

Starke Gesetz der Großen Zahl zur Verfügung. Ein Beispiel für eine solche Anwendung zeigt der folgende Satz aus Athreya und Lahiri [2006]³⁶:

Satz 3.2.2. *Sei i positiv rekurrent und $P_0(X_0 = i) = 1$. Der Anteil der Besuche in j bis zur Zeit n sei $L_n(j) := \frac{N_n(j)}{n+1}$. Dann gilt fast sicher*

$$L_n(j) \longrightarrow \frac{\mathbf{E}_i[N_{T_i-1}(j)]}{\mathbf{E}_i[T_i]} =: \pi_{ij}$$

Beweis. Betrachten wir eine beliebige Realisierung der Markovkette, so können wir zu jeder natürlichen Zahl n angeben, wie oft i bis zum Zeitpunkt n besucht wurde:
 $\forall n \exists k \equiv k(n) : T_i^{(k)} \leq n < T_i^{(k+1)}$. Klarerweise ist $N_{T_i^{(k)}}(j) \leq N_n(j) < N_{T_i^{(k+1)}}(j)$.
 Nun zählen wir mit ξ_r die Anzahl der Besuche in j während des r . Zyklus und formulieren die Ungleichung als

$$\sum_{r=0}^{k-1} \xi_r \leq N_n(j) \leq \sum_{r=0}^k \xi_r$$

bzw.

$$\frac{k}{n} \frac{1}{k} \sum_{r=0}^{k-1} \xi_r \leq \frac{N_n(j)}{n} \leq \frac{k+1}{n} \frac{1}{k+1} \sum_{r=0}^k \xi_r \quad (3.1)$$

Wir berechnen die Limiten von $\frac{1}{k} \sum_{r=0}^{k-1} \xi_r$ und $\frac{k(n)}{n}$:

Mengen der Form $\{\xi_r = l_r\}$ lassen sich schreiben als

$$\{\xi_r = l_r\} = \cup_{\#\{x_{rs} : x_{rs}=j, 1 \leq s \leq m_r\}=l_r} \{\eta_r = (x_{r1}, \dots, x_{rm_r})\}$$

Aus Satz 3.2.1 folgt, dass die Zufallsvariablen (ξ_r) unter P_i unabhängig und identisch verteilt sind. Da der Erwartungswert

$$\mathbf{E}_i[\xi_0] = \mathbf{E}_i[N_{T_i-1}(j)] \leq \mathbf{E}_i[T_i] < \infty$$

ist, gilt nach dem starken Gesetz der Großen Zahl

$$\frac{1}{k(n)} \sum_{r=0}^{k(n)-1} \xi_r \longrightarrow \mathbf{E}_i[\xi_0] \quad \text{f.s.}$$

Das starke Gesetz der Großen Zahl impliziert auch für die nach Satz 3.2.1 unabhängigen und identisch verteilten Zufallsvariablen $T_i^{(l+1)} - T_i^{(l)}$ (mit Erwartungswert

³⁶ Athreya und Lahiri [2006], Theorem 14.1.9

$\mathbf{E}_i [T_i^{(1)} - T_i^{(0)}] = \mathbf{E}_i [T_i] < \infty$, dass

$$\frac{1}{k(n)} \underbrace{\sum_{l=0}^{k(n)-1} (T_i^{(l+1)} - T_i^{(l)})}_{T_i^{(k(n))}} \longrightarrow \mathbf{E}_i [T_i] \quad \text{f.s.}$$

und wegen $T_i^{(k)} \leq n < T_i^{(k+1)}$ konvergiert auch $\frac{n}{k(n)}$ gegen $\mathbf{E}_i [T_i]$.

Mittels (3.1) sieht man, dass

$$L_n(j) = \frac{n}{n+1} \frac{N_n(j)}{n} \longrightarrow \frac{\mathbf{E}_i [\xi_0]}{\mathbf{E}_i [T_i]} = \pi_{ij} \quad \text{f.s.}$$

□

Satz 3.2.3. a) Die Aussage von Satz 3.2.2 gilt auch für eine beliebige Anfangsverteilung μ mit $P_\mu(T_i < \infty) = 1$.

b) Die Aussage von Satz 3.2.2 gilt auch für nullrekurrentes i und P_μ wie soeben, in diesem Fall ist $\mathbf{E}_i [N_{T_i-1}(j)] < \infty$ und $\lim L_n(j) = 0$ fast sicher.³⁷

Beweis. Der Beweis von Satz 3.2.1 funktioniert analog für ein P_μ wie angegeben. Daher gilt

$$P_\mu \left(\eta_k = (x_{k1}, \dots, x_{km_k}), T_i^{(k+1)} - T_i^{(k)} = m_k, 0 \leq k \leq n \right) = \prod_{k=0}^n P_i \left(\eta_0 = (x_{k1}, \dots, x_{km_k}), T_i^{(1)} = m_k \right)$$

Insbesondere sind die Zufallsvariablen (ξ_r) auch unter P_μ unabhängig und identisch verteilt mit $\mathbf{E}_\mu(\xi_0) = \mathbf{E}_i(\xi_0) < \infty$, und die Aussage kann wie für P_i bewiesen werden.

Falls i nullrekurrent ist, benötigen wir eine andere Form des Starken Gesetzes der Großen Zahl und behaupten:

Für eine Folge von unabhängigen, identisch verteilten Zufallsvariablen $(X_j)_{j \geq 1}$ mit $X_j \geq 0$ und $\mathbf{E}[X_j] = \infty$ gilt fast sicher $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n X_j = \infty$.

Dazu wählen wir ein $N \geq 1$ sowie ein $M > 0$ beliebig und betrachten die integrierbaren Zufallsvariablen $Y_j^{(M)} := \min(X_j, M)$. Wie (X_j) ist $(Y_j^{(M)})_{j \geq 1}$ eine Folge unabhängiger, identisch verteilter Zufallsvariablen. Lässt man M gegen unendlich gehen, dann wächst $Y_1^{(M)}$ gegen X_1 ; aus dem Satz von der monotonen Konvergenz folgt $\mathbf{E}[Y_1^{(M)}] \nearrow \mathbf{E}[X_1] = \infty$ für $M \rightarrow \infty$. Daher findet man ein M mit $\mathbf{E}[Y_1^{(M)}] \geq N$.

³⁷ Theorem 14.1.11 in Athreya und Lahiri [2006] enthält a) und b).

Aufgrund der Definition von $Y_j^{(M)}$ gilt für alle n

$$\frac{1}{n} \sum_{j=1}^n X_j \geq \frac{1}{n} \sum_{j=1}^n Y_j^{(M)}$$

und nach dem Starken Gesetz der Großen Zahl fast sicher

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n X_j \geq \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n Y_j^{(M)} = \mathbf{E} [Y_1^{(M)}] \geq N$$

Da N beliebig gewählt war, folgt die Behauptung.

Nun kann man wie im Beweis von Satz 3.2.2 schließen, dass fast sicher $\frac{T_i^{(k(n))}}{k(n)}$ gegen $\mathbf{E}_i [T_i]$ bzw. $\frac{k}{T_i^{(k)}}$ gegen 0 konvergiert. Benützen wir dies im Fall $j = i$: Zu jedem n ist $k(n)$ so gewählt, dass $T_i^{(k)} \leq n < T_i^{(k+1)}$ und deshalb $\frac{k}{n} \leq \frac{k}{T_i^{(k)}} \rightarrow 0$. Andererseits ist $N_n(i) = k$ und daher (f.s.)

$$0 \leq L_n(i) = \frac{N_n(i)}{n+1} = \frac{k}{n+1} \rightarrow 0$$

was insbesondere mit $\pi_{ii} = \frac{\mathbf{E}_i[N_{T_i-1}(i)]}{\mathbf{E}_i[T_i]} = \frac{1}{\mathbf{E}_i[T_i]}$ übereinstimmt.

Sei j beliebig. Nach Voraussetzung ist $P_\mu(T_i < \infty) = 1$. Der (endliche) Anteil der Besuche in j vor T_i wirkt sich auf den Grenzwert von $L_n(j)$ nicht aus; es genügt, die Markovkette ab dem Zeitpunkt T_i zu betrachten.

Der Zustand i ist nach Voraussetzung nullrekurrent und liegt in einer minimalen Teilmenge F . Ist j kein Element der abgeschlossenen Teilmenge F , so wird j von i ausgehend nicht mehr besucht und

$$L_n(j) \rightarrow 0 = \frac{\mathbf{E}_i[N_{T_i-1}(j)]}{\mathbf{E}_i[T_i]}$$

fast sicher. Hingegen ist ein $j \in F$ ebenfalls nullrekurrent, und wie soeben gezeigt gilt

$$L_n(j) \rightarrow \frac{\mathbf{E}_j[N_{T_j-1}(j)]}{\mathbf{E}_j[T_j]} = 0$$

Es bleibt zu zeigen, dass auch $\frac{\mathbf{E}_i[N_{T_i-1}(j)]}{\mathbf{E}_i[T_i]} = 0$ ist:

Definieren wir die Menge $A := \{i, j\}$ und deren Besuchszeiten

$$T_A^{(0)} := 0, \quad T_A^{(1)} = \inf \{n > 0 : X_n \in A\}, \quad T_A^{(k)} = \inf \{n > T_A^{(k-1)} : X_n \in A\}$$

wie für einzelne Zustände. Ebenso wie (X_n) ist $(\widetilde{X}_n)_{n \geq 0} := (X_{T_A^{(n)}})_{n \geq 0}$ eine Markovkette – aber auf einem endlichen Zustandsraum. Nun sind i und j positiv rekurrent,

wie man einfach nachrechnet. Die Anzahl $N_{T_i-1}(j)$ verändert sich in der neuen Markovkette jedoch nicht, und deren Erwartungswert ist $\mathbf{E}_i[N_{T_i-1}(j)] \leq \mathbf{E}_i[T_i] < \infty$. Mit der Nullrekurrenz von i in (X_n) gilt $\mathbf{E}_i[T_i] = \infty$ und $\frac{\mathbf{E}_i[N_{T_i-1}(j)]}{\mathbf{E}_i[T_i]} = 0$. $\square$

Korollar 3.2.4. *Sei i rekurrent und $P_\mu(T_i < \infty) = 1$. Es folgt³⁸*

$$\lim_{n \rightarrow \infty} \frac{1}{n+1} \sum_{k=0}^n P_\mu(X_k = j) = \lim_{n \rightarrow \infty} \mathbf{E}_\mu[L_n(j)] = \pi_{ij}$$

Beweis. Der Erwartungswert $\mathbf{E}_\mu[L_n(j)]$ ist gleich

$$\mathbf{E}_\mu \left[\frac{N_n(j)}{n+1} \right] = \frac{1}{n+1} \sum_{k=0}^n \mathbf{E}_\mu[\mathbf{1}_{\{j\}}(X_k)] = \frac{1}{n+1} \sum_{k=0}^n P_\mu(X_k = j)$$

$|L_n(j)| \leq 1$ ermöglicht, den Satz von der dominierten Konvergenz zu benützen:

$$\lim_{n \rightarrow \infty} \mathbf{E}_\mu[L_n(j)] = \lim_{n \rightarrow \infty} \int L_n(j) dP_\mu = \int \underbrace{\lim_{n \rightarrow \infty} L_n(j)}_{=\pi_{ij}} dP_\mu = \pi_{ij}$$

$\square$

Mit der folgenden Proposition ist Satz 3.1.4 vollständig bewiesen:

Proposition 3.2.5. *Sei i rekurrent.*

a) *i ist positiv rekurrent genau dann, wenn³⁹*

$$\lim_{n \rightarrow \infty} \mathbf{E}_i[L_n(i)] = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^n p_{ii}^{(k)} > 0.$$

b) *Wenn i positiv rekurrent und j von i aus erreichbar ist, folgt, dass auch j positiv rekurrent ist.⁴⁰*

Beweis. Nach Korollar 3.2.4 ist $\lim_{n \rightarrow \infty} \mathbf{E}_i[L_n(i)] = \frac{1}{\mathbf{E}_i[T_i]}$. Dieser Wert ist genau dann größer als 0, wenn $\mathbf{E}_i[T_i] < \infty$, also i positiv rekurrent ist.

Im zweiten Teil impliziert die Solidaritätseigenschaft erstens die Rekurrenz von j , zweitens $P_j(T_i < \infty) = 1$. Wir verwenden wieder Korollar 3.2.4 und erhalten einerseits

$$\lim_{n \rightarrow \infty} \mathbf{E}_j[L_n(j)] = \pi_{jj} = \frac{\mathbf{E}_i[\xi_0]}{\mathbf{E}_i[T_i]}$$

Andererseits gilt, da auch j rekurrent ist,

$$\lim_{n \rightarrow \infty} \mathbf{E}_j[L_n(j)] = \pi_{jj} = \frac{1}{\mathbf{E}_j[T_j]}$$

³⁸ Athreya und Lahiri [2006], Korollar 14.1.13

³⁹ Athreya und Lahiri [2006], Korollar 14.1.10(iii)

⁴⁰ Athreya und Lahiri [2006], S. 451

Dieser Grenzwert ist größer als 0: Da j von i aus erreichbar ist, gibt es ein n mit $p_{ij}^{(n)} > 0$, und so gilt

$$0 < \sum_{n \geq 0} p_{ij}^{(n)} = \mathbf{E}_i [N(j)] = \sum_{r \geq 0} \mathbf{E}_i [\xi_r] = \sum_{r \geq 0} \mathbf{E}_i [\xi_0] \Rightarrow \mathbf{E}_i [\xi_0] > 0$$

Der Nenner erfüllt $\mathbf{E}_i [T_i] < \infty$ aufgrund der positiven Rekurrenz von i . Aus $\frac{1}{\mathbf{E}_j [T_j]} > 0$ folgt, dass j positiv rekurrent ist. $\square$

Definition. Eine *stationäre Verteilung* bzw. ein *invariantes Maß* für eine Markovkette ist eine Wahrscheinlichkeitsverteilung $(\pi_j)_{j \in E}$ mit folgender Eigenschaft: Ist X_0 gemäß $(\pi_j)_{j \in E}$ verteilt, dann auch alle X_n ($n \geq 1$).

Lemma 3.2.6. Eine Wahrscheinlichkeitsverteilung $(\pi_j)_{j \in E}$ ist ein invariantes Maß genau dann, wenn $\pi_j = \sum_{i \in E} \pi_i p_{ij}$ für alle j .

Beweis. ($\Leftarrow$): Für die Startverteilung $\pi = (\pi_j)$ ist

$$\begin{aligned} P_\pi (X_n = j) &= \sum_{i \in E} \pi(i) P_i (X_n = j) = \\ &= \sum_{i \in E} \pi_i \sum_{i_1 \in E} p_{ii_1} \sum_{i_2 \in E} p_{i_1 i_2} \cdots \sum_{i_{n-1} \in E} p_{i_{n-1} j} = \\ &= \sum_{i_{n-1} \in E} \cdots \sum_{i_1 \in E} \underbrace{\sum_{i \in E} \pi_i p_{ii_1} p_{i_1 i_2} \cdots p_{i_{n-1} j}}_{\pi_{i_1}} = \cdots = \pi_j \end{aligned}$$

($\Rightarrow$): Wenn (π_j) die Verteilung von X_0 ist, dann auch von $X_1 \Rightarrow$

$$\pi_j = P_\pi (X_1 = j) = \sum_{i \in E} \pi(i) P_i (X_1 = j) = \sum_{i \in E} \pi_i p_{ij}$$

$\square$

Satz 3.2.7. Sei i positiv rekurrent und $\pi_j := \pi_{ij}$ für alle $j \in E$. Dann ist $(\pi_j)_{j \in E}$ eine stationäre Verteilung.

Beweis. Siehe Athreya und Lahiri [2006], Theorem 14.1.14. $\square$

Wir fassen zusammen:⁴¹

Satz 3.2.8. Sei i positiv rekurrent, μ eine Startverteilung mit $P_\mu (T_i < \infty) = 1$, j ein beliebiger Zustand und π_j wie bisher gegeben durch $\pi_j = \frac{\mathbf{E}_i [N_{T_i-1}(j)]}{\mathbf{E}_i [T_i]}$.

a) $(\pi_j)_{j \in E}$ ist ein invariantes Maß.

⁴¹ Athreya und Lahiri [2006], Theorem 14.1.15

$$b) L_n(j) = \frac{N_n(j)}{n+1} \rightarrow \pi_j \quad (P_\mu\text{-fast sicher})$$

$$c) \frac{1}{n+1} \sum_{k=0}^n P_\mu(X_k = j) \rightarrow \pi_j$$

Bemerkung. Falls $j = i$ und $P_\mu = P_i$, bedeutet c)

$$\frac{1}{n+1} \sum_{k=0}^n p_{ii}^{(k)} \rightarrow \pi_i = \frac{1}{\mathbf{E}_i[T_i]}$$

Der folgende Satz geht auf den Spezialfall einer ergodischen Markovkette ein.⁴²

Satz 3.2.9. *Sei (X_n) eine irreduzible Markovkette mit einem positiv rekurrenten Zustand.*

- a) *Alle Zustände sind positiv rekurrent.*
- b) $(\pi_j)_{j \in E}$ *ist das eindeutig bestimmte invariante Maß, wobei $\pi_j = \frac{1}{\mathbf{E}_j[T_j]}$.*
- c) *Für alle j und ein beliebiges Startmaß μ gelten $L_n(j) = \frac{N_n(j)}{n+1} \rightarrow \pi_j$ (P_μ -fast sicher) und*
- d) $\frac{1}{n+1} \sum_{k=0}^n P_\mu(X_k = j) \rightarrow \pi_j$

Beweis. Punkt a) gilt wegen der Irreduzibilität und Satz 3.1.4. Alle Zustände kommunizieren, daher ist für beliebige μ und i die Wahrscheinlichkeit $P_\mu(T_i < \infty) = 1$. Nach Satz 3.2.7 ist (π_j) invariant, nach Satz 3.2.8 die (eindeutige) Grenzverteilung von $\frac{1}{n+1} \sum_{k=0}^n P_\mu(X_k = j)$ sowie der Grenzwert von $L_n(j)$. $\square$

Beispiel. Auf einer reduzierten Markovkette kann es mehrere invariante Maße geben, auch wenn alle Zustände positiv rekurrent sind: Sei $E = \{A, B, a, b, c\}$ mit den Übergangswahrscheinlichkeiten $p_{AB} = p_{BA} = p_{ab} = p_{bc} = p_{ca} = 1$ gegeben. Der Zustandsraum kann in die minimalen Teilmengen $\{A, B\}$ und $\{a, b, c\}$ zerlegt werden; alle Zustände sind offensichtlich positiv rekurrent. Jedoch sind alle Wahrscheinlichkeitsmaße der Form

$$\pi = q\pi' + (1-q)\pi'' \quad 0 \leq q \leq 1$$

wobei $\pi = (\delta_A + \delta_B) \cdot \frac{1}{2}$ und $\pi'' = (\delta_a + \delta_b + \delta_c) \cdot \frac{1}{3}$, invariant: Für jedes $j \in E$ ist

$$\pi(j) = \frac{q}{2} \cdot \mathbf{1}_{\{A,B\}} + \frac{1-q}{3} \cdot \mathbf{1}_{\{a,b,c\}}$$

gleich $\sum_{i \in E} \pi_i p_{ij}$, denn in dieser Summe hat nur ein i positive Übergangswahrscheinlichkeit $p_{ij} = 1$ und nimmt unter π den gleichen Wert wie j an.

Umgekehrt zu Satz 3.2.9 zeigen wir⁴³

⁴² Athreya und Lahiri [2006], Theorem 14.1.16

⁴³ Athreya und Lahiri [2006], Theorem 14.1.18

Satz 3.2.10. Sei (X_n) eine irreduzible Markovkette und $(\pi_j)_{j \in E}$ eine stationäre Verteilung. Dann gelten die Aussagen a) bis d) aus Satz 3.2.9.

Beweis. Nach Lemma 3.2.6 gilt $\pi_j = \sum_{l \in E} \pi_l p_{lj}$ und daher

$$\pi_j = \sum_{l \in E} \sum_{i \in E} \pi_i p_{il} p_{lj} = \sum_{i \in E} \pi_i p_{ij}^{(2)}$$

bzw. induktiv $\pi_j = \sum_{i \in E} \pi_i p_{ij}^{(k)}$ für alle $k \geq 0$. Es folgt für alle n

$$\pi_j = \frac{1}{n+1} \sum_{k=0}^n \pi_j = \frac{1}{n+1} \sum_{k=0}^n \sum_{i \in E} \pi_i p_{ij}^{(k)} = \sum_{i \in E} \pi_i \left(\frac{1}{n+1} \sum_{k=0}^n p_{ij}^{(k)} \right) \quad (3.2)$$

Wenn j transient ist, gilt für eine beliebige Startverteilung μ , dass $P_\mu(N(j) < \infty) = 1$. Somit konvergiert $L_n(j) = \frac{N_n(j)}{n+1}$ fast sicher gegen 0, und mit dem Satz von der dominierten Konvergenz (wie im Beweis von Korollar 3.2.4) auch $\mathbf{E}_i[L_n(j)]$ für ein beliebiges i . Nochmalige Anwendung des Satzes von der dominierten Konvergenz liefert

$$\pi_j = \lim_{n \rightarrow \infty} \sum_{i \in E} \pi_i \underbrace{\left(\frac{1}{n+1} \sum_{k=0}^n p_{ij}^{(k)} \right)}_{=\mathbf{E}_i[L_n(j)]} = \sum_{i \in E} \pi_i \cdot 0 = 0$$

Ist $\pi_j > 0$, dann ist j also rekurrent, und zwar positiv rekurrent:

Die Rekursion $p_{ij}^{(k)} = \sum_{l=0}^k f_{ij}^{(l)} p_{jj}^{(k-l)}$ gilt klarerweise. Wegen $\sum_{k \geq 0} |f_{ij}^{(k)}| = f_{ij} < \infty$ und $\lim_{n \rightarrow \infty} \sum_{k=0}^n \left| \frac{p_{jj}^{(k)}}{n+1} \right| = \lim_{n \rightarrow \infty} \mathbf{E}_j[L_n(j)] = \frac{1}{\mathbf{E}_j[T_j]} < \infty$ ist

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n+1} \sum_{k=0}^n p_{ij}^{(k)} &= \lim_{n \rightarrow \infty} \sum_{k=0}^n \left(\sum_{l=0}^k f_{ij}^{(l)} \frac{p_{jj}^{(k-l)}}{n+1} \right) = \\ &= \lim_{n \rightarrow \infty} \sum_{k=0}^n f_{ij}^{(k)} \cdot \lim_{n \rightarrow \infty} \frac{1}{n+1} \sum_{k=0}^n p_{jj}^{(k)} = f_{ij} \cdot \frac{1}{\mathbf{E}_j[T_j]} \end{aligned}$$

Mit der Formel aus (3.2) und dem Satz von der dominierten Konvergenz folgt

$$0 < \pi_j = \lim_{n \rightarrow \infty} \sum_{i \in E} \pi_i \left(\frac{1}{n+1} \sum_{k=0}^n p_{ij}^{(k)} \right) = \underbrace{\left(\sum_{i \in E} \pi_i f_{ij} \right)}_{\geq 0} \frac{1}{\mathbf{E}_j[T_j]}$$

Somit ist $\frac{1}{\mathbf{E}_j[T_j]} > 0$ und $\mathbf{E}_j[T_j] < \infty$.

Wegen $\sum \pi_j = 1$ gibt es ein $\pi_j > 0$, und wir können alle Aussagen aus Satz 3.2.9 folgern. $\square$

Kapitel 4

Markovketten auf allgemeineren Räumen

4.1 Irreduzibilität

Sei $(E, \mathcal{E})$ ein beliebiger messbarer Raum. Zusätzlich nehmen wir einen abzählbaren Erzeuger der σ -Algebra $\mathcal{E}$ an (dh. $(E, \mathcal{E})$ ist *separabel*⁴⁴).

Definition. Eine Menge $A \in \mathcal{E}$ ist vom Zustand $x \in E$ aus *erreichbar*⁴⁵, wenn $\exists n \geq 1 : P_n(x, A) > 0$ (Schreibweise: $x \rightarrow A$, andernfalls $x \nrightarrow A$).

Gilt $x \nrightarrow F^c$ für alle Elemente x einer nichtleeren $\mathcal{E}$ -Menge F , so heißt F *abgeschlossen*⁴⁶.

Definition (φ -Irreduzibilität⁴⁷). Sei φ ein σ -endliches Maß in $\mathcal{M}^+(E)$ und $B \in \mathcal{E}$ mit $\varphi(B) > 0$. B heißt *φ -kommunizierend*, wenn von beliebigem $x \in B$ ausgehend jede Teilmenge $A \subseteq B$ mit $\varphi(A) > 0$ erreichbar ist. Falls $B = E$, wird der Kern P als *φ -irreduzibel* bezeichnet. Ist P φ -irreduzibel für ein Maß φ , so heißt P *irreduzibel* und φ ein *Irreduzibilitätsmaß* für P .

Mit φ ist auch jedes absolutstetige Maß $\rho \ll \varphi$ ein Irreduzibilitätsmaß⁴⁸, denn $\rho(A) > 0$ impliziert $\varphi(A) > 0$ und damit die gewünschte Bedingung. Mithilfe der folgenden Lemmata können wir zeigen, dass ein maximales Irreduzibilitätsmaß ψ existiert: Damit ist gemeint, dass alle Irreduzibilitätsmaße absolutstetig bezüglich ψ sind.

Lemma 4.1.1. $\varphi P \ll \varphi$ impliziert für den Potenzialkern G , dass $\varphi G \sim \varphi$.⁴⁹

Beweis. ($\varphi G \ll \varphi$): Mittels vollständiger Induktion zeigt man $\varphi P_n \ll \varphi$ für alle $n \geq 0$. Für $n = 0$ gilt dies trivial, für $n = 1$ nach Voraussetzung. Sei $\varphi P_n \ll \varphi$ erfüllt und $\varphi(A) = 0$. Dann gilt wegen

$$\varphi P_n(A) = \int \varphi(dx) P_n(x, A) = 0$$

⁴⁴ Hernández-Lerma und Lasserre [2003], S. 47

⁴⁵ Nummelin [1984], S. 8

⁴⁶ ebd.

⁴⁷ Nummelin [1984], S. 11

⁴⁸ Nummelin [1984], S. 13

⁴⁹ Nummelin [1984], Lemma 2.2 (i)

dass die Menge $\{x : P_n(x, A) > 0\}$ eine φ - und daher nach Voraussetzung auch eine φP -Nullmenge ist. Wir schließen

$$\varphi P_{n+1}(A) = \int (\varphi P)(dx) P_n(x, A) = 0$$

Somit ist für φ -Nullmengen A auch das Maß

$$\varphi G(A) = \int \varphi(dx) \sum_{n \geq 0} P_n(x, A) = \sum_{n \geq 0} \varphi P_n(A) = 0$$

($\varphi \ll \varphi G$): Aus der letzten Gleichung sieht man, dass im Fall $\varphi G(A) = 0$ auch alle $\varphi P_n(A) = 0$ sind, insbesondere für $n = 0$. $\square$

Lemma 4.1.2. *Für jedes nichttriviale, σ -endliche Maß φ auf $\mathcal{E}$ gibt es ein äquivalentes Wahrscheinlichkeitsmaß $\tilde{\varphi} \sim \varphi$.⁵⁰*

Beweis. Da φ σ -endlich ist, kann man eine Zerlegung von $E = \cup_{n \geq 1} E_n$ mit jeweils $0 < \varphi(E_n) < \infty$ wählen. Die Summe von Maßen

$$\tilde{\varphi}(A) = \sum_{n \geq 1} \frac{1}{2^n} \frac{\varphi(A \cap E_n)}{\varphi(E_n)}$$

ist wieder ein Maß, das die gleichen Nullmengen wie φ besitzt und

$$\tilde{\varphi}(E) = \sum_{n \geq 1} \frac{1}{2^n} \frac{\varphi(E_n)}{\varphi(E_n)} = \sum_{n \geq 1} \frac{1}{2^n} = 1$$

erfüllt. $\square$

Satz 4.1.3. *Sei P φ -irreduzibel. Dann existiert ein maximales Irreduzibilitätsmaß ψ , charakterisiert durch $\psi P \ll \psi$ und eindeutig bis auf Äquivalenz von Maßen.⁵¹*

Beweis. Erst zeigen wir, dass ein Irreduzibilitätsmaß ψ genau dann maximal ist, wenn $\psi P \ll \psi$.

($\Leftarrow$): Nach Voraussetzung ist φ ein beliebiges Irreduzibilitätsmaß für P . Wir zeigen $\varphi \ll \psi$:

$$\begin{aligned} \varphi(A) > 0 &\Rightarrow x \rightarrow A \quad \forall x \in E \Leftrightarrow \forall x \exists n : P_n(x, A) > 0 \Rightarrow \\ &\int \psi(dx) \sum_{n \geq 0} P_n(x, A) = \psi G(A) > 0 \Rightarrow \psi(A) > 0 \end{aligned}$$

Die letzte Implikation gilt, da nach Lemma 4.1.1 ψG absolutstetig bezüglich ψ ist.

⁵⁰ vgl. Nummelin [1984], Lemma 2.1

⁵¹ Nummelin [1984], Proposition 2.4

($\Rightarrow$): Mit ψ ist auch ψP ein Irreduzibilitätsmaß:

$$0 < \psi P(A) = \int \psi(dx) P(x, A) \Leftrightarrow \psi \left(\underbrace{\{x \in E : P(x, A) > 0\}}_{=: B} \right) > 0 \Rightarrow \\ y \rightarrow B \quad \forall y \in E, \quad x \rightarrow A \quad \forall x \in B \Rightarrow y \rightarrow A \quad \forall y \in E$$

Wegen der Maximalität gilt $\psi P \ll \psi$.

Es bleibt die Existenz von ψ zu beweisen. Das Irreduzibilitätsmaß

$$\psi = \sum_{n=0}^{\infty} 2^{-(n+1)} \tilde{\varphi} P_n$$

wobei $\tilde{\varphi} \sim \varphi$ ein Wahrscheinlichkeitsmaß ist, erfüllt $\psi P \ll \psi$ und ist somit maximal.

Dass ψ ein Irreduzibilitätsmaß ist, sieht man folgendermaßen:

$$\begin{aligned} \psi(A) > 0 &\Leftrightarrow \\ \exists n : 2^{-(n+1)} \tilde{\varphi} P_n(A) > 0 &\Leftrightarrow \\ \exists n : 2^{-(n+1)} \int_E \tilde{\varphi}(dx) P_n(x, A) > 0 &\Leftrightarrow \\ \tilde{\varphi}(\{x \in E : P_n(x, A) > 0\}) > 0 &\Leftrightarrow \\ \varphi(\{x \in E : P_n(x, A) > 0\}) > 0 &\Rightarrow \\ \forall y \in E : y \rightarrow \{x \in E : P_n(x, A) > 0\} &\Rightarrow \\ \forall y \in E : y \rightarrow A & \end{aligned}$$

Für die Absolutstetigkeit $\psi P \ll \psi$ wählen wir eine Menge A mit

$$0 = \psi(A) = \sum_{n \geq 0} 2^{-(n+1)} \int_E \tilde{\varphi}(dx) P_n(x, A) \Leftrightarrow \forall n : P_n(\cdot, A) = 0 \quad \tilde{\varphi}\text{-f.s.}$$

Dann ist auch

$$\begin{aligned} \psi P(A) &= \int_E \psi(dx) P(x, A) = \\ \sum_{n \geq 0} 2^{-(n+1)} \int_E (\tilde{\varphi} P_n)(dx) P(x, A) &= \\ \sum_{n \geq 0} 2^{-(n+1)} \tilde{\varphi} P_{n+1}(A) &= \\ \sum_{n \geq 0} 2^{-(n+1)} \int_E \tilde{\varphi}(dx) \underbrace{P_{n+1}(x, A)}_{=0 \text{ f.s.}} &= 0 \end{aligned}$$

□

4.2 Minorisierung

Sei P irreduzibel mit dem maximalen Irreduzibilitätsmaß ψ .

Definition. $\mathcal{E}^+$ bezeichnet die Teilmenge von Funktionen⁵² f in $\mathcal{E}_+$ mit $\int f d\psi > 0$.

P erfüllt die *Minorisierungsbedingung*⁵³ $M(m_0, \beta, s, \nu)$, wenn es eine natürliche Zahl $m_0 \geq 1$, eine reelle Konstante $\beta > 0$, eine Funktion $s \in \mathcal{E}^+$ sowie ein Maß $\nu \in \mathcal{M}^+(E)$ gibt, sodass

$$P_{m_0}(x, A) \geq \beta s(x) \nu(A) \quad \forall x \in E \quad \forall A \in \mathcal{E}$$

Dann heißt s eine *kleine Funktion*, ν ein *kleines Maß* und $C \in \mathcal{E}$ im Fall $s = \mathbf{1}_C$ eine *kleine Menge*. Die Menge der kleinen Funktionen wird mit $\mathcal{S}^+$ bezeichnet.

Bemerkung (Zur Minorisierungsbedingung).⁵⁴

- a) Jedes kleine Maß ν ist auch ein Irreduzibilitätsmaß, denn für beliebiges $x \in E$ und ν -positives $A \in \mathcal{E}$ gilt

$$P_{m_0}(x, A) \geq \underbrace{\beta}_{>0} s(x) \underbrace{\nu(A)}_{>0}$$

Ist $s(x) > 0$, so gilt offensichtlich $x \rightarrow A$. Ansonsten benützt man die Eigenschaft $s \in \mathcal{E}^+$ und die ψ -Irreduzibilität, um von x in die ψ -positive Menge $\{s > 0\}$ zu gelangen (und von dort nach A).

- b) Insbesondere ist ν absolutstetig bezüglich ψ .
- c) Meist kann man annehmen, dass $\int s d\nu > 0$.
- d) O.B.d.A. kann man $\beta = 1$ annehmen.

Nach Nummelin [1984], Theorem 2.1, gilt:

Satz 4.2.1. *Für einen irreduziblen Kern P ist die Menge $\mathcal{S}^+$ nichtleer.*

Der Beweis erfordert einige Vorüberlegungen und mehrere Lemmata:

Wir haben vorausgesetzt, dass $\mathcal{E}$ abzählbar erzeugt werden kann, etwa durch $\mathcal{G} = \{G_1, G_2, \dots\}$ mit $G_i \subseteq E$. Daraus konstruieren wir eine Folge von endlichen Partitionen des Zustandsraums E mittels

$$\begin{aligned} \mathcal{E}_1 &= \{G_1, G_1^C\} \\ \mathcal{E}_2 &= \{G_2 \cap G_1, G_2 \cap G_1^C, G_2^C \cap G_1, G_2^C \cap G_1^C\} \\ &\vdots \end{aligned}$$

⁵² Nummelin [1984], S. 14

⁵³ Nummelin [1984], Definition 2.3

⁵⁴ Nummelin [1984], Bemerkung 2.1

Für die Partition $\mathcal{E}_k$ werden also G_k und G_k^C mit allen $\mathcal{E}_{k-1}$ -Mengen geschnitten, dadurch werden die Partitionen immer feiner, und es gilt $\sigma(\mathcal{E}_i) \subseteq \sigma(\mathcal{E}_{i+1})$ für alle i . Außerdem ist $\sigma(\cup_{i \geq 1} \mathcal{E}_i) = \sigma(\mathcal{G}) = \mathcal{E}$. Für alle $x \in E$ und beliebiges $i \geq 1$ liegt x in einem eindeutig bestimmten Teil der Partition $\mathcal{E}_i$, den wir mit E_x^i bezeichnen.⁵⁵

Wir benötigen den folgenden Differentiationssatz für Maße. Seien ab nun λ und φ endliche Maße auf E und $\lambda = \lambda_a + \lambda_s$ die Lebesgue-Zerlegung von λ bezüglich φ (mit $\lambda_a \ll \varphi$ und $\lambda_s \perp \varphi$).

Lemma 4.2.2.

$$\lim_{i \rightarrow \infty} \mathbf{1}_{\{\varphi(E_x^i) > 0\}} \frac{\lambda(E_x^i)}{\varphi(E_x^i)} = \frac{d\lambda_a}{d\varphi}(x)$$

für φ -fast alle $x \in E$.⁵⁶

Beweis. Das Lemma wird in Doob [1990], S. 343 ff., mit Hilfe von Martingaltheorie bewiesen. $\square$

Für beliebig fixes x ist das Maß $P(x, \cdot)$ σ -endlich und besitzt somit eine Lebesgue-Zerlegung bezüglich φ :

$$P(x, A) = \mu_x(A) + \nu_x(A)$$

wobei $\nu_x \perp \varphi$ und $\mu_x \ll \varphi$. Aufgrund der Absolutstetigkeit existiert eine Radon-Nikodým-Ableitung $k(x, y) = \frac{d\mu_x}{d\varphi}$:

$$P(x, A) = \int_A k(x, y) \varphi(dy) + \nu_x(A)$$

Beispiel. Eine kleine Funktion existiert, falls auf Mengen X und Y positiven Maßes die Dichte $k(x, y)$ größer oder gleich einer Konstanten $c > 0$ ist:

$$\begin{aligned} P(x, A) &\geq \int_A k(x, y) \varphi(dy) \geq \int_{A \cap Y} k(x, y) \varphi(dy) \geq c \cdot \varphi(A \cap Y) \quad \forall x \in X \\ &\Rightarrow P(x, A) \geq c \cdot \mathbf{1}_X(x) \cdot \varphi(A \cap Y) \end{aligned}$$

Lemma 4.2.3. Die Dichte k kann nichtnegativ und $\mathcal{E} \otimes \mathcal{E}$ -messbar gewählt werden.⁵⁷

Beweis. Die Funktionen

$$k_i(x, y) := \mathbf{1}_{\{\varphi(E_y^i) > 0\}} \frac{P(x, E_y^i)}{\varphi(E_y^i)}$$

⁵⁵ Nummelin [1984], S. 16

⁵⁶ Nummelin [1984], Lemma 2.4

⁵⁷ Nummelin [1984], Lemma 2.5

sind nichtnegativ und als Produkt bzw. Quotient von $\mathcal{E} \otimes \mathcal{E}$ -messbaren Funktionen selber messbar. Beispielsweise ist das Urbild einer Menge $B \in \mathcal{B}_{\mathbf{R}}$ unter der Funktion $(x, y) \mapsto P(x, E_y^i)$

$$\{(x, y) : P(x, E_y^i) \in B\} = \bigcup_{y: \exists x: P(x, E_y^i) \in B} \left\{ (x, y) : x \in \underbrace{P(\cdot, E_y^i)^{-1}(B)}_{=: A_y \in \mathcal{E}} \right\} = \bigcup_{E_y^i: \exists x: P(x, E_y^i) \in B} \underbrace{\{A_y \times E_y^i\}}_{\in \mathcal{E} \otimes \mathcal{E}}$$

in $\mathcal{E} \otimes \mathcal{E}$ enthalten. Insgesamt ist auch der Limes für $i \rightarrow \infty$ nichtnegativ und messbar, und nach dem Differentiationssatz von soeben entspricht er der Dichte k . $\square$

Ebenso können wir für die Kerne P_n – bzw. deren absolutstetigen Teil – nichtnegative, messbare Versionen der Dichten $k^{(n)}$ finden, die zusätzlich die folgende Ungleichung erfüllen:

Lemma 4.2.4.

$$k^{(m+n)}(x, z) \geq \int P_m(x, dy) k^{(n)}(y, z) \geq \int k^{(m)}(x, y) k^{(n)}(y, z) \varphi(dy)$$

für alle $m, n \geq 1$ und $x, z \in E$.⁵⁸

Beweis. Sei für alle n und beliebig gewähltes x

$$P_n(x, A) = \mu_x^{(n)}(A) + \nu_x^{(n)}(A)$$

die Lebesgue-Zerlegung von $P_n(x, \cdot)$. Wie im vorigen Lemma gibt es für alle $\mu_x^{(n)}(\cdot)$ eine nichtnegative, $\mathcal{E} \otimes \mathcal{E}$ -messbare Dichte $k_0^{(n)}$. Die iterativ definierten Funktionen

$$k^{(1)} = k_0^{(1)}$$

$$k^{(n)} = \max \left(k_0^{(n)}(x, z), \max_{1 \leq m \leq n-1} \int P_m(x, dy) k^{(n-m)}(y, z) \right) \quad \forall n \geq 2$$

bilden ebenfalls eine Dichte:

Für $k^{(1)} = k_0^{(1)}$ ist dies klar. Nehmen wir an, dass alle $k^{(n-m)}$ Dichten für $\mu_x^{(n-m)}(\cdot)$ sind. O.B.d.A. sei $k^{(n)} = \max(k_0^{(n)}, k_*^{(n)})$, wobei $k_*^{(n)} = \int P_m(x, dy) k^{(n-m)}(y, z)$ für ein $m < n$. Wir definieren nun das Maß $Q_x(A) = \int_A \varphi(dz) k_*^{(n)}(x, z)$, welches wir

⁵⁸ Nummelin [1984], Lemma 2.6

mithilfe des Satzes von Fubini folgendermaßen abschätzen können:

$$\begin{aligned} Q_x(A) &= \int_A \varphi(dz) k_*^{(n)}(x, z) = \\ &= \int_A \varphi(dz) \int_E P_m(x, dy) k^{(n-m)}(y, z) = \\ &= \int_E P_m(x, dy) \underbrace{\int_A \varphi(dz) k^{(n-m)}(y, z)}_{\leq P_{n-m}(y, A)} \leq P_n(x, A) \end{aligned}$$

Aus der Definition von $Q_x(\cdot)$ ist ersichtlich, dass $Q_x \ll \varphi$. Hingegen ist $\mu_x^{(n)}$ in der Lebesgue-Zerlegung genau der maximale bezüglich φ absolutstetige Anteil von $P_n(x, \cdot)$, deshalb gilt $Q_x \leq \mu_x^{(n)}$. Damit ist auch das Integral von $k^{(n)}$ kleiner oder gleich dem Maß $\mu_x^{(n)}$:

$$\begin{aligned} &\int_A \varphi(dz) k^{(n)}(x, z) = \\ &= \int_{A \cap \{k_0^{(n)} \geq k_*^{(n)}\}} \varphi(dz) k^{(n)}(x, z) + \int_{A \cap \{k_0^{(n)} < k_*^{(n)}\}} \varphi(dz) k^{(n)}(x, z) = \\ &= \underbrace{\int_{A \cap \{k_0^{(n)} \geq k_*^{(n)}\}} \varphi(dz) k_0^{(n)}(x, z)}_{= \mu_x^{(n)}(A \cap \{k_0^{(n)} \geq k_*^{(n)}\})} + \underbrace{\int_{A \cap \{k_0^{(n)} < k_*^{(n)}\}} \varphi(dz) k_*^{(n)}(x, z)}_{\leq \mu_x^{(n)}(A \cap \{k_0^{(n)} < k_*^{(n)}\})} \leq \mu_x^{(n)}(A) \end{aligned}$$

Umgekehrt ist

$$\int_A \varphi(dz) k^{(n)}(x, z) \geq \int_A \varphi(dz) k_0^{(n)}(x, z) = \mu_x^{(n)}(A)$$

und insgesamt $k^{(n)}$ wie behauptet eine Dichte des absolutstetigen Teils von $P_n(x, \cdot)$.

Alle $k^{(n)}$ erfüllen $k^{(n)} \geq 0$ und die gewünschten Ungleichungen – die erste per definitionem. Die zweite ergibt sich mit der Lebesgue-Zerlegung von $P_m(x, \cdot)$ bezüglich φ :

$$\begin{aligned} P_m(x, A) &= \int_A k^{(m)}(x, y) \varphi(dy) + \nu_{m,x}(A) \Rightarrow \\ &\int P_m(x, dy) k^{(n)}(y, z) \geq \int k^{(m)}(x, y) k^{(n)}(y, z) \varphi(dy) \end{aligned}$$

Die $\mathcal{E} \otimes \mathcal{E}$ -Messbarkeit von $k^{(n)}$ ist mittels vollständiger Induktion etwas aufwändiger zu beweisen:

Die Funktion $k^{(1)} = k_0^{(1)}$ ist $\mathcal{E} \otimes \mathcal{E}$ -messbar. Nehmen wir induktiv an, dass für alle $m \leq n-1$ die Funktionen $k^{(n-m)} \in \mathcal{E} \otimes \mathcal{E}$ sind. Wir zeigen⁵⁹, dass das Integral einer

⁵⁹ analog zum Beweis von Lemma 14.20 in Klenke [2008]

nichtnegativen, $\mathcal{E} \otimes \mathcal{E}$ -messbaren Funktion $\tilde{f}$ bezüglich eines stochastischen Kerns P

$$f(x, z) = \int P(x, dy) \tilde{f}(y, z)$$

messbar ist bezüglich $\mathcal{E} \otimes \mathcal{E}$. Aufgrund der Induktionsannahme ist dann $k^{(n)}$ das Maximum messbarer Funktionen, also selber $\mathcal{E} \otimes \mathcal{E}$ -messbar.

Ist $\tilde{f} = \mathbf{1}_{A_1 \times A_2}$ die Indikatorfunktion eines messbaren Rechtecks, dann gilt die Behauptung:

$$f(x, z) = \int P(x, dy) \tilde{f}(y, z) = \int P(x, dy) \mathbf{1}_{A_1}(y) \cdot \mathbf{1}_{A_2}(z) = P(x, A_1) \cdot \mathbf{1}_{A_2}(z)$$

Messbare Rechtecke liegen also in der Menge

$$\mathcal{D} := \left\{ A \in \mathcal{E} \otimes \mathcal{E} : \int P(x, dy) \mathbf{1}_A(y, z) \in \mathcal{E} \otimes \mathcal{E} \right\}$$

$\mathcal{D}$ ist ein Dynkin-System:

$E \times E$ liegt in $\mathcal{D}$, da alle (x, z) auf $\int P(x, dy) \mathbf{1}_{E \times E}(y, z) = P(x, E) = 1$ abgebildet werden. Für $A \in \mathcal{D}$ ist $f(x, z) = P(x, A_2(z))$ messbar, wobei $A_2(z) = \{y : (y, z) \in A\}$ der z -Schnitt von A ist. Daher ist auch A^c ein Element von $\mathcal{D}$, denn

$$P(x, (A^c)_2(z)) = P(x, (A_2(z))^c) = 1 - P(x, A_2(z))$$

ist $\mathcal{E} \otimes \mathcal{E}$ -messbar. Schließlich liegt die Vereinigung disjunkter $A_n \in \mathcal{D}$ wieder in $\mathcal{D}$, was man mit $\mathbf{1}_{\cup A_n} = \sum \mathbf{1}_{A_n}$ und der Linearität des Integrals einfach nachweist.

Die messbaren Rechtecke der Dimension 2 sind ein durchschnittstabiler Erzeuger von $\mathcal{D}$. Nach dem π - λ -Theorem⁶⁰ stimmt die von ihnen erzeugte σ -Algebra $\mathcal{E} \otimes \mathcal{E}$ mit dem erzeugten Dynkin-System überein. Daher gilt $\mathcal{E} \otimes \mathcal{E} \subseteq \mathcal{D}$, und die Integrale aller $\tilde{f} = \mathbf{1}_A$ sind $\mathcal{E} \otimes \mathcal{E}$ -messbar.

Für Treppenfunktionen folgt die Messbarkeit von $f(x, z)$ mit der Linearität des Integrals. Allgemeine messbare nichtnegative Funktionen kann man durch Treppenfunktionen approximieren; $f(x, z)$ ist dann als Limes messbarer Funktionen messbar. $\square$

Definition. Seien A und B Elemente aus $\mathcal{E} \otimes \mathcal{E}$. Wir schreiben für die Schnitte⁶¹

$$A_1(x) := \{y \in E : (x, y) \in A\}, B_2(z) := \{y \in E : (y, z) \in B\}$$

und für die Zusammensetzung⁶²

$$A \circ B := (A \times E) \cap (E \times B) = \{(x, y, z) \in E \times E \times E : (x, y) \in A, (y, z) \in B\}$$

Weiters sei φ^n das n -fache Produktmaß $\varphi \times \dots \times \varphi$.

⁶⁰ Elstrodt [2005], S. 25, Satz 6.7

⁶¹ Nummelin [1984], S. 17

⁶² Nummelin [1984], S. 18

Lemma 4.2.5. *Sei $\varphi^3(A \circ B) > 0$. Dann gibt es φ -positive Mengen C und D in $\mathcal{E}$, sodass⁶³*

$$\gamma := \inf_{x \in C, z \in D} \varphi(A_1(x) \cap B_2(z)) > 0$$

Beweis. Betrachte die Maße

$$F \mapsto \varphi^2(A \cap F) \text{ und } F \mapsto \varphi^2(F)$$

auf $\mathcal{E} \times \mathcal{E}$. Es ist $\varphi^2(A \cap \cdot)$ absolutstetig bezüglich φ^2 . Die Lebesgue-Zerlegung besteht dann aus $\varphi^2(A \cap \cdot)$ selber und dem Maß, das trivial den Wert 0 annimmt. Die Radon-Nikodým-Ableitung bezüglich φ^2 ist offensichtlich $\mathbf{1}_A$.

Nach Voraussetzung gilt $\varphi^3(A \circ B) > 0$, dh. es gibt ein (x, y, z) mit $(x, y) \in A$ und $(y, z) \in B$. Setzen wir nun $E_{x,y}^i = E_x^i \times E_y^i$, dann gilt nach Lemma 4.2.2

$$\lim_{i \rightarrow \infty} \mathbf{1}_{\{\varphi^2(E_{x,y}^i) > 0\}} \frac{\varphi^2(A \cap E_{x,y}^i)}{\varphi^2(E_{x,y}^i)} = \mathbf{1}_A(x, y)$$

für φ^2 -fast alle (x, y) . Somit gibt es eine φ^2 -Nullmenge N_1 mit

$$\lim_{i \rightarrow \infty} \frac{\varphi^2(A \cap E_{x,y}^i)}{\varphi^2(E_{x,y}^i)} = 1$$

für alle (x, y) in $A \setminus N_1$.

Auf die gleiche Art zeigt man die Existenz einer φ^2 -Nullmenge N_2 , die

$$\lim_{i \rightarrow \infty} \frac{\varphi^2(B \cap E_{y,z}^i)}{\varphi^2(E_{y,z}^i)} = 1$$

für alle (y, z) in $B \setminus N_2$ erfüllt. Wir wählen (u, v, w) aus der nichtleeren Menge $(A \setminus N_1) \circ (B \setminus N_2)$. Sei j groß genug, dass

$$\frac{\varphi^2(A \cap E_{u,v}^j)}{\varphi^2(E_{u,v}^j)} \geq \frac{3}{4} \text{ und } \frac{\varphi^2(B \cap E_{v,w}^j)}{\varphi^2(E_{v,w}^j)} \geq \frac{3}{4}$$

gelten. Die linke Ungleichung ist äquivalent zu

$$\int \varphi(\{y : (x, y) \in A \cap E_{u,v}^j\}) d\varphi(x) \geq \frac{3}{4} \int \varphi(\{y : (x, y) \in E_{u,v}^j\}) d\varphi(x)$$

und dies bedeutet, dass auf einer φ -positiven Teilmenge von E_u^j

$$\varphi(A_1(x) \cap E_v^j) \geq \frac{3}{4} \varphi(E_v^j)$$

⁶³ Nummelin [1984], Lemma 2.7

gilt. Auf diese Weise erhalten wir die gesuchten Mengen

$$C = \left\{ x \in E_u^j : \varphi(A_1(x) \cap E_v^j) \geq \frac{3}{4} \varphi(E_v^j) \right\}$$

$$D = \left\{ z \in E_w^j : \varphi(B_2(z) \cap E_v^j) \geq \frac{3}{4} \varphi(E_v^j) \right\}$$

für deren Elemente $x \in C$ und $z \in D$ gilt:

$$\begin{aligned} \varphi(E_v^j) &\geq \varphi(E_v^j \cap (A_1(x) \cup B_2(z))) = \\ \varphi(A_1(x) \cap E_v^j) + \varphi(B_2(z) \cap E_v^j) - \varphi(A_1(x) \cap B_2(z) \cap E_v^j) &\geq \\ \frac{3}{4} \varphi(E_v^j) + \frac{3}{4} \varphi(E_v^j) - \varphi(A_1(x) \cap B_2(z) \cap E_v^j) &\Leftrightarrow \\ \underbrace{\frac{1}{2} \varphi(E_v^j)}_{>0} &\leq \underbrace{\varphi(A_1(x) \cap B_2(z) \cap E_v^j)}_{\leq \varphi(A_1(x) \cap B_2(z))} \Rightarrow \\ &0 < \varphi(A_1(x) \cap B_2(z)) \end{aligned}$$

□

Nun können wir Satz 4.2.1 beweisen:

Beweis. Sei ψ das maximale Irreduzibilitätsmaß für P und o.B.d.A. ein Wahrscheinlichkeitsmaß. Aufgrund der Irreduzibilität ist für beliebige $x \in E$ und für ψ -fast alle $y \in E$ die Summe

$$\sum_{m=1}^{\infty} k^{(m)}(x, y) > 0$$

Daher gibt es m_1 und m_2 mit

$$\int_E \int_E \int_E k^{(m_1)}(x, y) k^{(m_2)}(y, z) \psi(dx) \psi(dy) \psi(dz) > 0$$

Wir können nun ein $\delta > 0$ klein genug wählen, dass die Zusammensetzung der Mengen $A := \{(x, y) : k^{(m_1)}(x, y) \geq \delta\}$ und $B := \{(y, z) : k^{(m_2)}(y, z) \geq \delta\}$

$$\psi^3(A \circ B) > 0$$

erfüllt. Nach Lemma 4.2.5 existieren Mengen C und D , für deren Elemente $x \in C$ und $z \in D$ wegen der Ungleichungen aus Lemma 4.2.4 und Lemma 4.2.5

$$\begin{aligned} k^{(m_1+m_2)}(x, z) &\geq \int_E k^{(m_1)}(x, y) k^{(m_2)}(y, z) \psi(dy) \geq \\ &\geq \int_{A_1(x) \cap B_2(z)} k^{(m_1)}(x, y) k^{(m_2)}(y, z) \psi(dy) \geq \gamma \delta^2 \end{aligned}$$

gilt. P erfüllt somit für alle $x \in C$

$$\begin{aligned} P_{m_1+m_2}(x, A) &\geq \int_A k^{(m_1+m_2)}(x, z) d\psi(z) \geq \\ &\geq \int_{A \cap D} k^{(m_1+m_2)}(x, z) d\psi(z) \geq \gamma \delta^2 \psi(A \cap D) \end{aligned}$$

also die Minorisierungsbedingung $M(m_1 + m_2, \gamma \delta^2, \mathbf{1}_C, \psi(\cdot \cap D))$. $\square$

4.3 Periodizität

Sei P ein irreduzibler Kern mit maximalem Irreduzibilitätsmaß ψ , der die Minorisierungsbedingung $M(m_0, \beta, s, \nu)$ erfüllt (nach Satz 4.2.1 existiert eine kleine Funktion).

Definition. Ein *Zyklus*⁶⁴ der Länge m ist eine Folge disjunkter nichtleerer $\mathcal{E}$ -Mengen $(E_0, \dots, E_{m-1})$, sodass

$$\forall x \in E_i : P(x, E_j^c) = 0$$

wobei $j \equiv i + 1 \pmod{m}$ und $0 \leq i \leq m - 1$.

Bemerkung. a) Von x in E_i sind also nur Zustände in $E_{i+1} \pmod{m}$ erreichbar. Die Vereinigung $\cup_{i=0}^{m-1} E_i$ ist abgeschlossen, ihr Komplement muss wegen der Irreduzibilität eine ψ -Nullmenge sein.⁶⁵

b) Für die Kerne P_n und x aus E_i ergibt sich $P_n(x, E_j^c) = 0$ mit $j = i + n \pmod{m}$, insbesondere $P_m(x, E_i^c) = 0$.⁶⁶

Sei I die Menge der Indizes, für die die Minorisierungsbedingung gilt:

$$I := \{m \geq 1 : M(m, \beta_m, s, \nu) \text{ für ein } \beta_m > 0\}$$

und d ihr größter gemeinsamer Teiler. Unter den angegebenen Umständen mit der Zusatzannahme $\int s d\nu > 0$ gilt:⁶⁷

Satz 4.3.1. Zum Kern P gibt es einen d -Zyklus $(E_0, \dots, E_{d-1})$. Die Zahl d hängt nicht von der Wahl von s und ν ab. Gibt es einen weiteren Zyklus $(E'_0, \dots, E'_{d'-1})$, so teilt d' die Zahl d , und die Mengen E'_i sind Vereinigungen der E_i .

Die Mengen E_i werden *zyklische Mengen* genannt, d die *Periode* von P . Falls $d = 1$, heißt P *aperiodisch*, ansonsten *periodisch*.

Bemerkung. „Beschleunigt“ man die Kette, indem man nur Vielfache eines gewissen $m \geq 1$ betrachtet, so erhält man eine Markovkette $(X_{nm})_{n \geq 0}$ mit Übergangswahrscheinlichkeit P_m . Je nachdem, in welcher Zyklusmenge gestartet wird, ergeben sich

⁶⁴ Nummelin [1984], S. 20

⁶⁵ ebd.

⁶⁶ ebd.

⁶⁷ Nummelin [1984], Theorem 2.2

$c_m = \text{ggT}\{m, d\}$ verschiedene Zustandsräume, in jedem Fall existiert ein Zyklus der Länge $\frac{d}{c_m}$.⁶⁸

4.4 Rekurrenz und Harris-Rekurrenz

Wir erinnern uns an die Definition von T_A als Eintrittszeit in eine Menge $A \in \mathcal{E}$. Wie für diskrete Markovketten bezeichnen wir die k . Besuchszeit in A mit $T_A^{(k)}$ und die Anzahl der Besuche mit $N(A) = \sum_{n \geq 1} \mathbf{1}_A(X_n)$.

φ -Irreduzibilität bedeutet, dass für φ -positive Mengen A und für alle x ein $P_n(x, A) > 0$ ist, also auch $\mathbf{E}_x[N(A)] = \sum_{n \geq 1} P_n(x, A) > 0$.

Definition (Rekurrenz). Eine Markovkette heißt *rekurrent*⁶⁹, wenn sie φ -irreduzibel ist und

$$\varphi(A) > 0 \Rightarrow \mathbf{E}_x[N(A)] = \infty$$

für alle $x \in E$, $A \in \mathcal{E}$ erfüllt.

Definition (Harris-Rekurrenz). Eine irreduzible Markovkette (X_n) mit maximalem Irreduzibilitätsmaß ψ ist *Harris-rekurrent*⁷⁰, wenn

$$\psi(A) > 0 \Rightarrow P_x(T_A < \infty) = 1 \quad \forall x \in E \quad (4.1)$$

Bemerkung. a) Alternativ kann man die Eigenschaft (4.1) für alle Irreduzibilitätsmaße fordern:⁷¹ Gilt sie für alle, dann insbesondere für ψ . Sei umgekehrt φ ein beliebiges Irreduzibilitätsmaß, $\varphi(A) > 0$. Dann ist die gewünschte Implikation erfüllt, weil $\varphi \ll \psi$.

b) Eine weitere Möglichkeit ist, $P_x(T_A < \infty) = 1$ durch $P_x(N(A) = \infty) = 1$ zu ersetzen.⁷²

Proposition 4.4.1. $P_x(T_A < \infty) = 1 \quad \forall x \in E \Leftrightarrow P_x(N(A) = \infty) = 1 \quad \forall x \in E$.

Beweis. ($\Rightarrow$): Nach Voraussetzung gilt $P_x(T_A < \infty) = 1$ für alle $x \in E$. Nehmen wir dies auch für $T_A^{(k-1)}$ – die $(k-1)$. Besuchszeit in A – an, dann folgt mit der starken

⁶⁸ Nummelin [1984], S. 22

⁶⁹ Hernández-Lerma und Lasserre [2003], Definition 4.2.9

⁷⁰ vgl. Athreya und Lahiri [2006], S. 463

⁷¹ Tatsächlich ist das Irreduzibilitätsmaß in der Definition von Athreya und Lahiri [2006], S. 463, beliebig.

⁷² Nummelin [1984], Definition 3.5., oder Hernández-Lerma und Lasserre [2003], Definition 4.2.10

Markoveigenschaft

$$\begin{aligned}
 \mathbf{E}_x \left[\mathbf{1}_{\{T_A^{(k)} < \infty\}} | \mathcal{F}_{T_A} \right] &= \mathbf{E}_x \left[\mathbf{1}_{\{T_A^{(k-1)} < \infty\}} \circ \theta_{T_A} | \mathcal{F}_{T_A} \right] = \\
 &= \mathbf{E}_{X_{T_A}} \left[\mathbf{1}_{\{T_A^{(k-1)} < \infty\}} \right] = P_{X_{T_A}} \left(T_A^{(k-1)} < \infty \right) = 1 \\
 \Rightarrow P_x \left(T_A^{(k)} < \infty \right) &= \int_E \mathbf{1}_{\{T_A^{(k)} < \infty\}} dP_x = \int_E 1 dP_x = 1
 \end{aligned}$$

Somit wird die Menge A fast sicher unendlich oft besucht.

($\Leftarrow$): Klar. □

Proposition 4.4.2. *Ist (X_n) Harris-rekurrent, dann auch $(X_{nn_0})_{n \geq 0}$ für ein $n_0 \geq 1$.*

Beweis. Siehe Nummelin [1984], Proposition 3.14. □

Die Eigenschaft der Harris-Rekurrenz ist hilfreich, um Ergebnisse über das Langzeitverhalten von diskreten Markovketten auf allgemeine Markovketten zu übertragen. Im diskreten Fall sind die Ausflüge zwischen den Rückkehrzeiten in einen Zustand i unabhängig und identisch verteilt, wodurch man das Starke Gesetz der Großen Zahl nützen kann, um Konvergenzaussagen zu erzielen. Für die Rückkehrzeiten in eine Menge A muss dies nicht erfüllt sein. Jedoch findet man für Harris-rekurrente Markovketten „gute“ Rückkehrzeiten, sodass man ebenfalls das Starke Gesetz der Großen Zahl anwenden kann.

Zu Beginn zeigen wir die Existenz von unabhängigen, identisch verteilten Zyklen unter speziellen Voraussetzungen:⁷³

Satz 4.4.3. *Sei (X_n) eine Markovkette mit Übergangswahrscheinlichkeit P und Zustandsraum $(E, \mathcal{E})$. Angenommen, es gibt eine Menge A_0 in $\mathcal{E}$ sowie eine Konstante $0 < \alpha < 1$ und ein Wahrscheinlichkeitsmaß ν auf $\mathcal{E}$, sodass*

- a) $P(x, A) \geq \alpha \nu(A) \quad \forall x \in A_0 \quad \forall A \in \mathcal{E}$
- b) $P_x(T_{A_0} < \infty) = 1 \quad \forall x \in E$

Dann existiert zu jeder Startverteilung μ eine Folge von Zufallszeiten $(T_i)_{i \geq 1}$, sodass die Zyklen $\eta_k = \{X_r : T_k \leq r < T_{k+1}\}$ unter P_μ unabhängig und identisch verteilt und alle X_{T_i} gemäß ν verteilt sind.

Das heißt für beliebige $A_{kl} \in \mathcal{E}$, $m_k \in \mathbf{N}$ und für alle $n \in \mathbf{N}$

$$\begin{aligned}
 P_\mu(T_{k+1} - T_k = m_k, X_{T_k+l} \in A_{kl}, 0 \leq l < m_k, 1 \leq k \leq n) &= \\
 \prod_{k=1}^n P_\nu(T_2 - T_1 = m_k, X_{T_1+l} \in A_{kl}, 0 \leq l < m_k) & \quad (4.2)
 \end{aligned}$$

⁷³ Athreya und Lahiri [2006], Theorem 14.2.6

Bemerkung. Unabhängige und identisch verteilte Zyklen sind genau durch (4.2) definiert, wobei im allgemeinen auf beiden Seiten das gleiche Wahrscheinlichkeitsmaß P auftritt⁷⁴ (hier soll die zusätzliche Bedingung $X_{T_i} \sim \nu$ erfüllt sein). Ein stochastischer Prozess auf $(\Omega, \mathcal{F}, P)$, der eine solche Folge von Zufallszeiten besitzt, heißt regenerativ.

Beweis. Zuerst definiert man für beliebiges $x \in A_0$ das Maß $Q(x, \cdot)$ auf $\mathcal{E}$ durch

$$Q(x, \cdot) = \frac{P(x, \cdot) - \alpha \nu(\cdot)}{1 - \alpha}$$

Mit $P(x, \cdot)$ und ν ist auch $Q(x, \cdot)$ σ -additiv und wegen $Q(x, E) = 1$ ein Wahrscheinlichkeitsmaß.

Für beliebiges $x \in A_0$ und Startmaß μ gilt natürlich $P_\mu(X_{n+1} \in A | X_n = x) = P(x, A)$. Betrachtet man hingegen unabhängige Zufallsvariable

$$\begin{array}{ll} \eta_{n+1} & \text{mit } P_\mu(\eta_{n+1} \in A) = \nu(A) \\ \delta_{n+1} & \text{Bernoulli-verteilt mit Parameter } \alpha \\ Y_{n+1,x} & \text{mit } P_\mu(Y_{n+1,x} \in A) = Q(x, A) \end{array}$$

und setzt

$$X_{n+1} = \begin{cases} \eta_{n+1} & \text{für } \delta_{n+1} = 1 \\ Y_{n+1,x} & \text{für } \delta_{n+1} = 0 \end{cases}$$

so erhält man die gleiche Übergangswahrscheinlichkeit von x nach A :

$$\begin{aligned} P_\mu(X_{n+1} \in A | X_n = x) &= \\ P_\mu(\eta_{n+1} \in A, \delta_{n+1} = 1) + P_\mu(Y_{n+1,x} \in A, \delta_{n+1} = 0) &= \\ \nu(A) \alpha + Q(x, A) (1 - \alpha) &= P(x, A) \end{aligned}$$

Daher ist die Position der Kette nach einem Besuch in A_0 mit einer Wahrscheinlichkeit von α gemäß ν verteilt – unabhängig von der Vergangenheit.

Für alle $x \in E$ gilt nach Voraussetzung $P_x(T_{A_0} < \infty) = 1$, und nach Proposition 4.4.1 finden P_x -fast sicher unendlich viele Besuche in A_0 statt – dann auch P_μ -fast sicher. Bezeichnen wir die Besuchszeiten mit $\tau_1, \tau_2, \dots$, so besitzen die Zufallsvariablen X_{τ_i+1} mit Wahrscheinlichkeit $\alpha > 0$ die Verteilung ν . Die Folge $(\delta_{\tau_i+1})_{i \geq 1}$ entspricht einer Folge von Münzwürfen mit Erfolgswahrscheinlichkeit α .

$$\begin{aligned} P_\mu(\exists i_0 : X_{\tau_{i_0}+1} \sim \nu) &= P_\mu(\exists i_0 : \delta_{\tau_{i_0}+1} = 1) = \\ \sum_{i \geq 1} P_\mu(\delta_{\tau_j+1} = 0 \ \forall j < i, \delta_{\tau_i+1} = 1) &= \\ \sum_{i \geq 1} (1 - \alpha)^{i-1} \alpha &= \alpha \sum_{i \geq 0} (1 - \alpha)^i = 1 \end{aligned}$$

⁷⁴ siehe Gleichung (4.3) unten

(X_n) beginnt also fast sicher zum Zeitpunkt $\tau_{i_0} + 1$ mit Anfangsverteilung ν neu und unabhängig von den vorigen Zeitpunkten zu laufen. Wir setzen $T := T_1 := \tau_{i_0} + 1$ und berechnen

$$\begin{aligned}
 & P_\mu(X_j \in A_j : n+1 \leq j \leq n+k \mid X_j \in A_j : 0 \leq j \leq n, T = n+1) = \\
 & P_\mu(X_j \in A_j : n+1 \leq j \leq n+k \mid X_n \in A_n, T = n+1) = \\
 & \frac{1}{P_\mu(X_n \in A_n, T = n+1)} \cdot P_\mu(X_j \in A_j : n \leq j \leq n+k, T = n+1) = \\
 & \frac{1}{\int \mu(dx) \mathbf{1}_{A_0 \cap A_n}(x) \int_E \nu(dx_{n+1})} \cdot \int \mu(dx) \mathbf{1}_{A_0 \cap A_n}(x) \int_{A_{n+1}} \nu(dx_{n+1}) \cdot \\
 & \int_{A_{n+2}} P(x_{n+1}, dx_{n+2}) \dots \int_{A_{n+k}} P(x_{n+k-1}, dx_{n+k}) = \\
 & \frac{\mu(A_0 \cap A_n)}{\mu(A_0 \cap A_n)} \cdot \int_{A_{n+1}} \nu(dx_0) \int_{A_{n+2}} P(x_0, dx_1) \dots \int_{A_{n+k}} P(x_{k-2}, dx_{k-1}) = \\
 & P_\nu(X_0 \in A_{n+1}, \dots, X_{k-1} \in A_{n+k})
 \end{aligned}$$

Somit gilt

$$\begin{aligned}
 & P_\mu(X_j \in A_j : 0 \leq j \leq n+k, T = n+1) = \\
 & P_\mu(X_j \in A_j : 0 \leq j \leq n, T = n+1) \cdot \\
 & P_\mu(X_j \in A_j : n+1 \leq j \leq n+k \mid X_j \in A_j : 0 \leq j \leq n, T = n+1) = \\
 & P_\mu(X_j \in A_j : 0 \leq j \leq n, T = n+1) \cdot P_\nu(X_0 \in A_{n+1}, \dots, X_{k-1} \in A_{n+k})
 \end{aligned}$$

Durch Beginn zur Zeit T_1 und wiederholtes Anwenden im Fall $\mu = \nu$ folgt die Behauptung. $\square$

Korollar 4.4.4. *Wir nehmen (X_n) , A_0 , α und ν wie in Satz 4.4.3 an und ändern die Bedingung a) in*

$$\exists n_0 \geq 1 : P_{n_0}(x, A) \geq \alpha \nu(A) \text{ und } P_x(\exists n \geq 1 : X_{nn_0} \in A_0) = 1 \quad \forall x \in A_0 \quad \forall A \in \mathcal{E}$$

Für $(Y_n) = (X_{nn_0})$ und beliebiges μ bzw. P_μ gibt es dann Zufallszeiten $(T_i)_{i \geq 1}$ mit unabhängigen, identisch verteilten Zyklen und mit $Y_{T_i} \sim \nu$.⁷⁵

Beweis. Sei μ eine Startverteilung mit $\mu(A_0) = 1$. Die Voraussetzungen von 4.4.3 sind für (Y_n) mit der Übergangswahrscheinlichkeit P_{n_0} erfüllt und das Resultat bewiesen.

Ansonsten folgt aus der Voraussetzung $P_x(T_{A_0} < \infty) = 1$ für alle $x \in E$, dass auch $P_\mu(T_{A_0} < \infty) = 1$. Betrachtet man die Markovkette ab dem Zeitpunkt T_{A_0} , so gilt wieder der erste Fall, und man findet Zufallszeiten (T_i) , die (4.2) erfüllen. $\square$

Für $(Y_n) = (X_{nn_0})$ haben wir nun unabhängige, identisch verteilte Zyklen. Die im

⁷⁵ Athreya und Lahiri [2006], Theorem 14.2.7

vollständigen Zyklus $\tilde{\eta}_k = \{X_j : n_0 T_k \leq j < n_0 T_{k+1}\}$ enthaltenen Y_n ($T_k \leq n < T_{k+1}$) sind also identisch verteilt für alle Zyklen, ebenso die Länge der Zyklen. Die Stellen X_j , wobei $j \neq nn_0$, können wir gemäß der bedingten Verteilung

$$\begin{aligned} P_\mu(X_{kn_0+l} \in A_{kl}, 1 \leq l < n_0 | X_{kn_0} = x, X_{(k+1)n_0} = y) = \\ P_\mu(X_l \in A_{kl}, 1 \leq l < n_0 | X_0 = x, X_{n_0} = y) \end{aligned}$$

auffüllen.⁷⁶ Daher sind die Zyklen $(\tilde{\eta}_k)$ identisch verteilt.

Unabhängig sind sie nicht: Die Verteilung von $\tilde{\eta}_k$ ist von $X_{n_0 T_{k+1}}$ abhängig, und dieser Wert gehört zum Zyklus $\tilde{\eta}_{k+1}$. Daher sind aufeinander folgende Zyklen abhängig, alle anderen aber wegen der Markoveigenschaft unabhängig.⁷⁷

Diese Resultate erhalten wir insbesondere für Harris-rekurrente Markovketten:⁷⁸

Satz 4.4.5. *Sei (X_n) eine ψ -irreduzible, Harris-rekurrente Markovkette auf einem separablen Zustandsraum $(E, \mathcal{E})$. Dann existieren Zufallszeiten $(T_i)_{i \geq 1}$ und ein $n_0 \geq 1$, sodass die Zyklen $\tilde{\eta}_k$ identisch verteilt sind. Aufeinander folgende Zyklen sind abhängig, alle anderen Kombinationen von Zyklen unabhängig.*

Beweis. Wir überprüfen die Voraussetzungen von Korollar 4.4.4:

Nach Satz 4.2.1 existieren eine Funktion $s = \mathbf{1}_{A_0}$, ein $n_0 \geq 1$, $\alpha > 0$ und ein Maß ν , sodass

$$P_{n_0}(x, A) \geq \alpha s(x) \nu(A) \quad \forall x \in E \quad \forall A \in \mathcal{E} \quad \text{bzw.}$$

$$P_{n_0}(x, A) \geq \alpha \nu(A) \quad \forall x \in A_0 \quad \forall A \in \mathcal{E}$$

Es gilt $\psi(A_0) > 0$ und daher wegen der Harris-Rekurrenz $P_x(T_{A_0} < \infty) = 1$ für alle $x \in E$. Nach Proposition 4.4.2 ist auch $(X_{nn_0})_{n \geq 0}$ Harris-rekurrent und daher die Eintrittszeit in A_0 bezüglich der Übergangswahrscheinlichkeit P_{n_0} fast sicher endlich, das heißt

$$P_x(\exists n \geq 1 : X_{nn_0} \in A_0) = 1 \quad \forall x \in A_0$$

□

Nun untersuchen wir noch das Langzeitverhalten von regenerativen Prozessen. Wie bereits in der Bemerkung nach Satz 4.4.3 angedeutet, sind diese folgendermaßen definiert:

Definition. Ein stochastischer Prozess (X_n) auf $(\Omega, \mathcal{F}, P)$ heißt *regenerativ*⁷⁹, wenn

⁷⁶ Athreya und Lahiri [2006], S. 466

⁷⁷ Athreya und Lahiri [2006], Theorem 14.2.8

⁷⁸ Athreya und Lahiri [2006], Theorem 14.2.9

⁷⁹ Athreya und Lahiri [2006], Definition 14.2.7

es eine Folge von Zufallszeiten $(T_k)_{k \geq 1}$ (*Erneuerungszeiten*) gibt mit

$$P(T_{k+1} - T_k = m_k, X_{T_k+l} \in A_{kl}, 0 \leq l < m_k, 1 \leq k \leq n) = \prod_{k=1}^n P(T_2 - T_1 = m_k, X_{T_1+l} \in A_{kl}, 0 \leq l < m_k) \quad (4.3)$$

für beliebige $A_{kl} \in \mathcal{E}$, $m_k \in \mathbf{N}$ und für alle $n \in \mathbf{N}$.

Beispiel. a) Eine irreduzible und rekurrente Markovkette auf einem diskreten Zustandsraum.

b) (X_n) sei Harris-rekurrent und erfülle Bedingung a) in Satz 4.4.3.

Satz 4.4.6. $(X_n)_{n \geq 0}$ sei ein regenerativer Prozess mit Erneuerungszeiten $(T_i)_{i \geq 1}$. Die Maße $\tilde{\pi}$ und π auf $(E, \mathcal{E})$ seien definiert durch

$$\tilde{\pi}(A) = \mathbf{E} \left[\sum_{j=T_1}^{T_2-1} \mathbf{1}_A(X_j) \right] \quad \text{und} \quad \pi(A) = \frac{\tilde{\pi}(A)}{\tilde{\pi}(E)}$$

unter der Zusatzannahme, dass $\tilde{\pi}(E) = \mathbf{E}[T_2 - T_1] < \infty$. Dann folgt⁸⁰

$$a) \quad P\left(\frac{1}{n} \sum_{j=0}^{n-1} f(X_j) \longrightarrow \int f d\pi\right) = 1 \quad \forall f \in L^1(E, \mathcal{E}, \pi)$$

$$b) \quad \mu_n(\cdot) := \frac{1}{n} \sum_{j=0}^{n-1} P(X_j \in \cdot) \longrightarrow \pi(\cdot) \quad \text{mengenweise und bezüglich } \|\cdot\|_{TV}$$

Beweis. a) Sei o.B.d.A. $f \geq 0$ (sonst betrachte f^+, f^-). Für jedes n existiert ein $k = N_n$, sodass $T_k \leq n < T_{k+1}$ (setze $T_0 := 0$). Definiert man $Y_0 = \sum_{j=0}^{T_1-1} f(X_j)$ und $Y_i = \sum_{j=T_i}^{T_{i+1}-1} f(X_j)$, dann gilt

$$\frac{k}{n} \frac{1}{k} \left(Y_0 + \sum_{i=1}^{k-1} Y_i \right) \leq \frac{1}{n} \sum_{j=0}^{n-1} f(X_j) \leq \frac{k}{n} \frac{1}{k} \left(Y_0 + \sum_{i=1}^k Y_i \right) \quad (4.4)$$

Die (Y_n) sind unabhängig und identisch verteilt. Mit dem Starken Gesetz der Großen Zahl können wir schließen, dass $\frac{1}{k} \sum_{i=1}^{k-1} Y_i$ und $\frac{1}{k} \sum_{i=1}^k Y_i$ gegen $\mathbf{E}[Y_1]$ konvergieren. Dieser Grenzwert entspricht, wenn f eine Treppenfunktion ist,

$$\begin{aligned} \int f d\tilde{\pi} &= \int \sum_i \alpha_i \mathbf{1}_{A_i} d\tilde{\pi} = \sum_i \alpha_i \tilde{\pi}(A_i) = \sum_i \alpha_i \mathbf{E} \left[\sum_{j=T_1}^{T_2-1} \mathbf{1}_{A_i}(X_j) \right] = \\ &= \mathbf{E} \left[\sum_{j=T_1}^{T_2-1} \underbrace{\sum_i \alpha_i \mathbf{1}_{A_i}(X_j)}_{f(X_j)} \right] = \mathbf{E} \left[\sum_{j=T_1}^{T_2-1} f(X_j) \right] = \mathbf{E}[Y_1] \end{aligned}$$

⁸⁰ Athreya und Lahiri [2006], Theorem 14.2.10

was natürlich durch den Satz von der monotonen Konvergenz ebenso für allgemeine messbare Funktionen gilt. Weiters berechnen wir mithilfe des Starken Gesetzes der Großen Zahl, dass $\frac{k}{n} \rightarrow \frac{1}{\mathbf{E}[T_2 - T_1]}$: Wie im Beweis von Satz 3.2.2 argumentiert man mit $T_k \leq n < T_{k+1}$ und $\frac{1}{k}T_k = \frac{1}{k} \sum_{i=0}^{k-1} (T_{i+1} - T_i) \rightarrow \mathbf{E}[T_2 - T_1]$. Insgesamt folgt wegen (4.4)

$$\frac{1}{n} \sum_{j=0}^{n-1} f(X_j) \longrightarrow \frac{\mathbf{E}[Y_1]}{\mathbf{E}[T_2 - T_1]} = \frac{\int f d\tilde{\pi}}{\tilde{\pi}(E)} = \int f d\pi \text{ f.s.}$$

b) Wähle in Punkt a) $f = \mathbf{1}_A$ für ein beliebiges $A \in \mathcal{E} \Rightarrow$

$$\underbrace{\frac{1}{n} \sum_{j=0}^{n-1} \mathbf{1}_A(X_j)}_{=: f_n} \longrightarrow \int \mathbf{1}_A d\pi = \pi(A) \text{ f.s.}$$

Da $|f_n| \leq 1$, impliziert der Satz von der dominierten Konvergenz

$$\mu_n(A) = \frac{1}{n} \sum_{j=0}^{n-1} P(X_j \in A) = \int f_n dP \longrightarrow \int \pi(A) dP = \pi(A)$$

Der Beweis der Konvergenz bezüglich der Totalvariations-Norm verlangt starke funktionalanalytische Argumente. Hier sei auf Nummelin [1984], Kapitel 6, verwiesen.

□

Bemerkung. Der Satz gilt auch, wenn die Zyklen m -abhängig sind (wobei $1 \leq m < \infty$), dh. dass für beliebige $k \in \mathbf{N}$ die Mengen $\{\eta_1, \dots, \eta_k\}$ und $\{\eta_{k+m+1}, \dots\}$ unabhängig sind.⁸¹ Das Starke Gesetz der Großen Zahl in der allgemeinen Form des Ergodensatzes lässt sich in diesem Fall ebenso anwenden.

Definition. Ein Wahrscheinlichkeitsmaß π auf $(E, \mathcal{E})$ heißt *stationär*⁸² für die Übergangswahrscheinlichkeit P , wenn für alle $A \in \mathcal{E}$

$$\pi(A) = \int P(x, A) \pi(dx)$$

Bemerkung. Aus der Definition folgt induktiv $X_n \sim \pi$ für alle $n \geq 1$, wenn $X_0 \sim \pi$ – daher die Bezeichnung „stationär“.

Satz 4.4.7. Sei (X_n) eine Harris-rekurrente Markovkette auf einem separablen Zustandsraum $(E, \mathcal{E})$ mit Übergangswahrscheinlichkeit P . Angenommen, es gibt ein stationäres Wahrscheinlichkeitsmaß π für P . Dann ist π eindeutig bestimmt, und es gelten

⁸¹ vgl. Athreya und Lahiri [2006], Definition 16.2.3

⁸² Athreya und Lahiri [2006], Definition 14.2.8

die folgenden Konvergenzaussagen für alle $x \in E$:⁸³

$$\begin{aligned} \forall f \in L^1(E, \mathcal{E}, \pi) : \quad & \frac{1}{n} \sum_{j=0}^{n-1} f(X_j) \longrightarrow \int f d\pi && P_x\text{-fast sicher} \\ & \frac{1}{n} \sum_{j=0}^{n-1} P_x(X_j \in \cdot) =: \mu_{n,x}(\cdot) \longrightarrow \pi(\cdot) && \text{mengenweise} \\ & \mu_{n,x}(\cdot) \longrightarrow \pi(\cdot) && \text{bezüglich } \|\cdot\|_{TV} \end{aligned}$$

Beweis. Wir werden beweisen, dass ein beliebiges stationäres Maß genau von der Form des Maßes π im Satz 4.4.6 ist. Damit steht einerseits die Eindeutigkeit fest, andererseits folgen die Konvergenzaussagen.

Wie für Satz 4.4.5 gezeigt, haben wir eine Menge $A_0 \in \mathcal{E}$, ein $n_0 \geq 1$, $\alpha > 0$ und ein Maß ν mit

$$P_{n_0}(x, A) \geq \alpha \nu(A) \quad \forall x \in A_0 \quad \forall A \in \mathcal{E}$$

und $P_x(T_{A_0} < \infty) = 1$ für alle $x \in E$. O.B.d.A. sei $n_0 = 1$ (ansonsten betrachtet man die Übergangswahrscheinlichkeit P_{n_0}). Wie früher definiert man die unabhängigen Zufallsvariablen (η_n) , (δ_n) und $(Y_{n,x})$ und findet Zufallszeiten

$$\begin{aligned} T_1 &= \min \{n : n > 0, X_{n-1} \in A_0, \delta_n = 1\} \\ T_{i+1} &= \min \{n : n > T_i, X_{n-1} \in A_0, \delta_n = 1\} \end{aligned}$$

wobei die X_{T_i} – unabhängig von der Vergangenheit – gemäß ν verteilt sind. Wieder wählen wir für jedes $n \geq 1$ die Zufallsvariable $k = N_n$ mit $T_k \leq n < T_{k+1}$ ($T_0 := 0$).

Nach Proposition 4.4.1 gilt $N(A_0) = \infty$ P_x -fast sicher für beliebiges $x \in E$. Weiters hat $\delta_n = 1$ die positive Wahrscheinlichkeit α ; somit ist $P_x(N_n \rightarrow \infty) = 1$ für alle x . Mit dem Starken Gesetz der Großen Zahl folgert man wie im Beweis von Satz 3.2.2, dass $\lim \frac{N_n}{n} = \frac{1}{\mathbf{E}_\nu[T_1]}$ P_x -fast sicher. Da $|\frac{N_n}{n}| \leq 1$, liefert der Satz von der dominierten Konvergenz

$$\mathbf{E}_x \left[\frac{N_n}{n} \right] \longrightarrow \frac{1}{\mathbf{E}_\nu[T_1]}$$

Andererseits ist

$$\mathbf{E}_x \left[\frac{N_n}{n} \right] = \frac{1}{n} \mathbf{E}_x \left[\sum_{k=1}^n \mathbf{1}_{\{k \in \{T_i\}_{i \geq 1}\}} \right] = \frac{1}{n} \sum_{k=1}^n P_x(X_{k-1} \in A_0) \alpha$$

⁸³ Athreya und Lahiri [2006], Theorem 14.2.11

und daher konvergiert für alle $x \in E$, insbesondere π -fast überall,

$$f_n(x) := \frac{1}{n} \sum_{j=0}^{n-1} P_x(X_j \in A_0) = \frac{1}{\alpha} \mathbf{E}_x \left[\frac{N_n}{n} \right] \longrightarrow \frac{1}{\alpha \mathbf{E}_\nu[T_1]} \quad (4.5)$$

Ein stationäres Maß π erfüllt für alle $j \geq 0$

$$\pi(A_0) = P_\pi(X_j \in A_0) = \int \pi(dx) P_x(X_j \in A_0)$$

und somit für beliebiges n

$$n\pi(A_0) = \int \sum_{j=0}^{n-1} P_x(X_j \in A_0) \pi(dx)$$

Wegen der Irreduzibilität der Markovkette und da $\psi(A_0) > 0$, existiert für alle $x \in E$ ein $n : P_x(X_n \in A_0) > 0$. Aus der letzten Gleichung ist nun ersichtlich, dass $\pi(A_0) > 0$ ist. Dividiert man beide Seiten durch n und lässt n gegen unendlich gehen, so folgt mit $|f_n| \leq 1$, Gleichung (4.5) und dem Satz von der dominierten Konvergenz

$$0 < \pi(A_0) = \lim_{n \rightarrow \infty} \underbrace{\int \frac{1}{n} \sum_{j=0}^{n-1} P_x(X_j \in A_0) \pi(dx)}_{f_n(x)} = \frac{1}{\alpha \mathbf{E}_\nu[T_1]} \quad (4.6)$$

Somit ist $\mathbf{E}_\nu[T_1] < \infty$, und die Voraussetzungen von Satz 4.4.6 sind erfüllt (mit $\mathbf{E}[T_2 - T_1] = \mathbf{E}_\nu[T_1]$). Wie im Beweis von Punkt b) des genannten Satzes gilt

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=0}^{n-1} P_x(X_j \in A) = \frac{\mathbf{E}_\nu \left[\sum_{j=0}^{T_1-1} \mathbf{1}_A(X_j) \right]}{\mathbf{E}_\nu[T_1]} \quad \forall A \in \mathcal{E} \quad \forall x \in E$$

Nun braucht man nur noch bezüglich π zu integrieren und wie in (4.6) den Satz von der dominierten Konvergenz anzuwenden:

$$\pi(A) = \lim_{n \rightarrow \infty} \int \frac{1}{n} \sum_{j=0}^{n-1} P_x(X_j \in A) \pi(dx) = \frac{\mathbf{E}_\nu \left[\sum_{j=0}^{T_1-1} \mathbf{1}_A(X_j) \right]}{\mathbf{E}_\nu[T_1]}$$

Hiermit ist gezeigt, dass das zunächst beliebig gewählte stationäre Maß π genau das Maß aus Satz 4.4.6 ist. Die Konvergenzaussagen entsprechen nun genau Punkt a) und b) in Satz 4.4.6. $\square$

Literaturverzeichnis

- K. B. Athreya und S. N. Lahiri. *Measure Theory and Probability Theory*. Springer Science+Business Media, LLC, New York, 2006.
- K. B. Athreya und P. Ney. A new approach to the limit theory of recurrent Markov chains. *Transactions of the American Mathematical Society*, 245:493–501, 1978.
- J. L. Doob. *Stochastic Processes*. Wiley Classics Library ed. Wiley, New York [u.a], 1990.
- J. Elstrodt. *Maß- und Integrationstheorie*. Springer, Berlin [u.a.], 4. Auflage, 2005.
- O. Hernández-Lerma und J. B. Lasserre. *Markov Chains and Invariant Probabilities*. Birkhäuser Verlag AG, Basel [u.a.], 2003.
- F. Hofbauer. Stochastische Prozesse. Vorlesungsskriptum, 2002.
- A. Klenke. *Wahrscheinlichkeitstheorie*. Springer, Berlin, Heidelberg, 2. Auflage, 2008.
- E. Nummelin. A Splitting Technique for Harris Recurrent Markov Chains. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 43(4):309–318, 1978.
- E. Nummelin. *General irreducible Markov chains and non-negative operators*, Band 83 von *Cambridge tracts in mathematics*. Cambridge University Press, Cambridge [u.a], 1984.
- D. Revuz. *Markov Chains*, Band 11 von *North-Holland Mathematical Library*. North-Holland, Amsterdam, 2. Auflage, 1984.
- R. Zweimüller. Maß- und Integrationstheorie. Vorlesungsskriptum, 2009.

Anhang A

Zusammenfassung

Diese Arbeit beschäftigt sich mit homogenen Markovketten, einer bedeutenden Klasse stochastischer Prozesse. Charakterisiert sind diese dadurch, dass die Wahrscheinlichkeit, von einem Zustand in eine gewisse Menge zu gelangen, nicht von der Vergangenheit abhängt und für alle Zeitpunkte gleich ist.

Ist der Zustandsraum abzählbar, so lässt sich das Langzeitverhalten des Prozesses sehr genau vorhersagen: Die Zustände können als transient, nullrekurrent oder positiv rekurrent klassifiziert werden, und je nachdem entwickeln sich die Wahrscheinlichkeit, in einen Zustand zurückzukehren, die durchschnittliche Anzahl der Besuche etc. Die zentrale Frage ist, welche der Konvergenzaussagen (unter welchen zusätzlichen Voraussetzungen) auf Markovketten mit überabzählbarem Zustandsraum übertragen werden können.

Im ersten Kapitel werden Grundbegriffe geklärt und eine Notation festgelegt. Grundsätzlich sind Kenntnisse der Maßtheorie vorausgesetzt, wenige wesentliche Ergebnisse werden wiederholt.

Zu Beginn des Abschnitts 2 steht die Definition von homogenen Markovketten. Diese kann auf verschiedene Arten erfolgen; unter anderem wird die Existenz der kanonischen Markovkette, ausgehend von einer Startverteilung und einem stochastischen Kern, gezeigt. Weiters wird die Markoveigenschaft in verschiedenen Varianten erläutert; die wichtigste davon ist die starke Markoveigenschaft.

Kapitel 3 geht auf den Spezialfall diskreter Markovketten ein. Im ersten Teil werden Begriffe wie Erreichbarkeit von Zuständen und Irreduzibilität der Markovkette definiert. Der Zustandsraum enthält disjunkte minimale Teilmengen, Zustände können als rekurrent oder transient eingeteilt werden. Dies wird im zweiten Teil benutzt, um das Langzeitverhalten der Markovkette zu untersuchen. Da die Markovkette in einen rekurrenten Zustand fast sicher zurückkehrt, kann man sie in die Zyklen zwischen den Besuchszeiten aufspalten. Dank der starken Markoveigenschaft stellt sich heraus, dass diese Zyklen eine Folge unabhängiger, identisch verteilter Zufallsvariablen sind, und so lassen sich Schlüsse über das Langzeitverhalten der ganzen Markovkette ziehen. Unter anderem wird die Existenz eines invarianten Maßes, sowie die Eindeutigkeit im irreduziblen Fall, für eine Markovkette mit einem positiv rekurrenten Zustand gezeigt. Dieses Maß ist der Grenzwert sowohl der durchschnittlichen Anzahl der Besuche in einem gewissen Zustand als auch des über die Zeit gemittelten Wahrscheinlichkeitsmaßes, diesen Zustand aufzusuchen.

Nach dieser eingehenden Abhandlung diskreter Markovketten wird im vierten Kapitel der Zustandsraum allgemeiner als messbarer Raum angenommen. Dies benötigt eine allgemeinere Definition von Irreduzibilität, nun bezüglich verschiedener Maße, wobei die Existenz eines maximalen Irreduzibilitätsmaßes bewiesen wird. In Abschnitt 4.2 wird die Minorisierungsbedingung eingeführt und in einem aufwändigen Beweis gezeigt, dass jede irreduzible Markovkette eine solche erfüllt. Teil 4.3 geht kurz auf Periodizität ein.

Das Herzstück dieser Arbeit ist Abschnitt 4.4 mit der Untersuchung Harris-rekurrenter Markovketten. Da diese per definitionem irreduzibel sind, kann man unter anderem die Minorisierungsbedingung verwenden, um identisch verteilte, eins-abhängige Zyklen zu finden. Mit Hilfe eines Satzes über regenerative Prozesse folgen schließlich Konvergenzaussagen über Harris-rekurrente Markovketten mit einem stationären Wahrscheinlichkeitsmaß.

Anhang B

Lebenslauf

Veronika Huemer-Kals

Geboren 1982 in Steyr, Oberösterreich

Nach der Matura 2001 an der HBLA für Kultur- und Kongressmanagement Steyr Übersiedlung nach Wien. Berufliche Tätigkeit u.a. im Leopold Museum (Administration, Assistenz Marketing).

Ab Oktober 2004 Studium der Mathematik an der Universität Wien mit Studienschwerpunkt Stochastik, weiteres Interessensgebiet Algebra. Im Februar 2007 Abschluss des ersten Studienabschnitts.

Parallel zum Studium seit April 2005 Mitarbeiterin des IBO – Österreichisches Institut für Baubiologie und -ökologie bzw. der IBO GmbH mit vielfältigen administrativen und organisatorischen Tätigkeiten in den Bereichen Sekretariat, Buchhaltung, IBOmagazin, Veranstaltungen, Öffentlichkeitsarbeit u.a.

Fremdsprachenkenntnisse in Englisch, Französisch und Italienisch sowie Grundkenntnisse in Spanisch und Kroatisch.