Westfälische Wilhelms-Universität Münster

Institut für Mathematische Statistik

Cut-Off bei Geburts- und Todesprozessen

Diplomarbeit

von

Christoph Diehl

Datum

29. Januar 2009

Westfälische Wilhelms-Universität Münster

Institut für Mathematische Statistik

Cut-Off bei Geburts- und Todesprozessen

Diplomarbeit

von

Christoph Diehl

Datum

29. Januar 2009

Betreut durch Professor Dr. G. Alsmeyer

Danksagung

Ich danke Herrn Prof. G.Alsmeyer für die Vergabe des interessanten Themas und die Betreuung und Unterstützung während der Erstellung dieser Arbeit. Des Weiteren danke ich meinen Eltern für die langjährige Unterstützung während des Studiums.

Inhaltsverzeichnis

1 Einleitung

Einige ergodische Markov-Prozesse zeigen einen scharfen Übergang bei der Konvergenz gegen ihre stationäre Verteilung. Sie bleiben bis zu einem festen Zeitpunkt in weitem Abstand von ihrer stationären Verteilung und nähern sich ihr im Anschluss exponentiell schnell an. Anfang der achtziger Jahre wurde bei zufälligen Transpositionen auf der symmetrischen Gruppe in [DS81] erstmalig eine solche abrupte Konvergenz gegen die stationäre Verteilung festgestellt. Diesem Phänomen wurde von Aldous und Diaconis in [AD85] der Name Cut-Off-Effekt gegeben. Einer der interessantesten und präzisesten Cut-Offs wurde in [DB92] von Diaconis und Beyer bewiesen. In dieser Arbeit wurde die Frage beantwortet, wie oft ein anfänglich geordnetes Kartenspiel mit der Mischmethode Riffle-Shuffle gemischt werden muss, damit es als gemischt betrachtet werden kann. Wir zitieren hier das Hauptergebnis dieser Arbeit:

Theorem 1.1. *Bezeichne P_n^l die Verteilung eines anfänglich geordneten, aus n Karten bestehenden Kartenspiels nach l Mischungen mit der Mischmethode Riffle-Shuffle. Sei u_n die Gleichverteilung über allen Permutationen und setze $l = (3/2)\log_2 n + c$. Dann gilt für großes n*

$$\left\|P_n^l - u_n\right\|_{TV} = 1 - 2\Phi\left(-\frac{2^{-c}}{4\sqrt{3}}\right) + O\left(\frac{1}{n^{1/4}}\right), \quad \textit{wobei} \quad \Phi(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t} e^{-s^2/2}ds.$$

Eine Reihe weiterer Beispiele für den Auftritt des Cut-Off-Effekts liefert Diaconis in der Übersichtsarbeit [Dia96]. Gemeinsamkeit dieser Beispiele ist, dass für den Beweis des Vorliegens des Cut-Off-Effekts jeweils gründliche Kenntnisse der zugrundeliegenden Familie von Markov-Prozessen vorliegen mussten. Insbesondere die Zeit, zu der ein Cut-Off auftritt, musste zum Beweis der Existenz eines Cut-Offs bekannt sein.

Diaconis fragte deshalb in dem Übersichtsartikel [Dia96] nach einem allgemeinen Kriterium für das Auftreten eines Cut-Offs in einer Familie endlicher, ergodischer Markov-Prozesse. Ein erstes solches Kriterium bei der Untersuchung des Cut-Off-Phänomens gab Peres im Jahr 2004. Die Konvergenzrate wird bekanntermaßen vom betraglich zweitgrößten Eigenwert des Markov-Prozesses bestimmt. Naheliegend ist demnach die Vermutung, dass die Spektrallücke, welche im reversiblen Fall die Differenz zwischen Spektralradius und dem betraglich zweitgrößten Eigenwert ist, in einem derartigen Kriterium auftaucht. Als zweite Größe integriert sich die ϵ-Mischzeit, also die erste Zeit, zu der der Abstand zur stationären Verteilung einen fest vorgegebenen Wert $\epsilon > 0$ unterschreitet. Konkret formulierte Peres: Ein notwendiges — in vielen Fällen auch hinreichendes — Kriterium für Cut-Off in einer Familie endlicher reversibler Markov-Prozesse ist das Folgende: Das Produkt aus Spektrallücke und Mischzeit geht gegen unendlich. Bei diesem Kriterium muss also die Zeit, zu der der Cut-Off auftritt, nicht notwendig bekannt sein, um Aussagen über die Existenz eines Cut-Offs machen zu können. Eine Reihe von Arbeiten zur Gültigkeit des Peres-Kriteriums in verschiedenen Klassen von Markov-Prozessen und bei Zugrundelegung verschiedener Abstandsbegriffe sind in jüngster Zeit erschienen. Es zeigt sich, dass bei Abstandsmessung bezüglich der L^p-Norm, $1 < p \leq \infty$, das Kriterium von Peres bei reversiblen Prozessen in der Tat notwendig und hinreichend ist. Dies wurde kürzlich in [CSC08] und [Che06] bewiesen. Misst man den Abstand zur stationären Verteilung in totaler Variation, ist es im Allgemeinen nicht hinreichend. Ein auf Aldous zurückgehendes Beispiel für eine reversible Markov-Kette, welche das Kriterium erfüllt, aber keinen Cut-Off zeigt, werden wir in Kapitel 7 vorstellen. Bei der ob ihrer mannigfaltigen Anwendungsgebiete sehr bedeutenden Klasse der Geburts- und Todesprozesse ist das Kriterium hingegen auch in Totalvariation notwendig und hinreichend, wie Peres, Lubetzky und Ding in [DLP08] zeigten.

Wir betrachten in unserer Arbeit ebenfalls die Klasse der Geburts- und Todesprozesse und messen den Abstand zur stationären Verteilung in Separation. Hauptergebnis unserer Arbeit wird der Beweis sein, dass bei in 0 gestarteten, endlichen Geburts- und Todesprozessen in stetiger Zeit und Abstandsmessung in Separation der Vorschlag von Peres ein notwendiges und hinreichendes Kriterium für die Existenz eines Cut-Offs ist.

Wir formulieren jetzt dieses Hauptergebnis für Geburts- und Todesprozesse in stetiger Zeit. Für jedes n bezeichne γ_n^t die Verteilung eines ergodischen, in 0 gestarteten, Geburts- und Todesprozesses auf $\Omega_n = \{0, 1, ..., n\}$ zur Zeit t. ξ_n^* sei die zugehörige stationäre Verteilung. Die

Separation zwischen γ_n^t und der Zielverteilung ξ_n^* ist definiert durch

$$\mathrm{s}(\gamma_n^t, \xi_n^*) = \sup_{\omega \in \Omega_n} \{1 - \gamma_n^t(\omega)/\xi_n^*(\omega)\}.$$

Seien $\lambda_{n,i} \in [0,2]$, $i = 0, 1, 2, ..., n$, die Eigenwerte, in nichtfallender Reihenfolge, von $-Q_n$ (wobei Q_n die zugehörige Q-Matrix ist, also die infinitesimale erzeugende Matrix des zugehörigen Markov-Prozesses darstellt). $\lambda_{n,1} = \lambda_n$ heißt Spektrallücke. Setze

$$t_n = \sum_{i=1}^{n} (\lambda_{n,i})^{-1}.$$

Der Hauptsatz lautet folgendermaßen:

Theorem 1.2. *In obiger Situation hat die Familie* $(\Omega_n, \xi_n^*, (\gamma_n^t)_{t>0})_{n=1,2,...}$ *genau dann einen Separations-Cut-Off , wenn* $N_n = \lambda_n t_n \xrightarrow{n} \infty$.

t_n kann man äquivalent durch die ϵ-Separations-Mischzeit

$$\tau_n^{\mathrm{s}} = \tau_n^{\mathrm{s}}(\epsilon) = \inf\{t \geq 0 : \mathrm{s}(\gamma_n^t, \xi_n^*) \leq \epsilon\}$$

ersetzen, also die erste Zeit, zu der der Abstand in Separation zwischen der Verteilung des Markov-Prozesses und der stationären Verteilung geringer als ϵ ist. Dieses Theorem werden wir in Kapitel 5 beweisen.

Unsere Arbeit ist folgendermaßen aufgebaut: Zunächst werden wir in Kapitel 2 notwendige Grundbegriffe der Theorie der Markov-Prozesse bereitstellen, verschiedene Abstandsbegriffe zwischen Wahrscheinlichkeitsverteilungen vorstellen und diese miteinander vergleichen, bevor wir den Cut-Off-Effekt definieren. Dabei werden wir auch den Begriff der Fenstergröße eines Cut-Offs und mit dem Pre-Cut-Off eine schwächere Form des Cut-Off-Effekts präsentieren.

In Kapitel 3 werden wir eine Verbindung zwischen stark stationären Zeiten und der Abstandsmessung durch Separation herstellen. Stark stationäre Zeiten eines Markov-Prozesses entsprechen der Zeit bis zur Absorption eines sogenannten stark stationären dualen Prozesses zu diesem Markov-Prozess. Die Entwicklung dieser Dualitätstheorie wird einen großen Teil dieser Arbeit einnehmen und geht auf Diaconis und Fill [DF90b] sowie Fill [Fil92] zurück. Wir werden den dualen Prozess erst im diskreten Fall untersuchen und danach den stetigen Fall behandeln. Wir erhalten mit dem Algebraischen Dualitätstheorem eine notwendige und hinreichende Bedingung für die Existenz eines solchen dualen Prozesses und geben eine Pfad-Konstruktion des

dualen Prozesses an. Bei Geburts- und Todesprozessen vereinfacht sich die Dualitätstheorie signifikant und der von uns konstruierte duale Prozess hat eine besonders einfache Gestalt: Der duale Prozess ist selbst ein Geburts- und Todesprozess mit den gleichen Eigenwerten. Unser durch die Pfad-Konstruktion entstandene stark stationäre duale Prozess ist scharf, das heißt die Zeit bis zur Absorption des dualen Prozesses liefert eine minimale stark stationäre Zeit für den zugrundeliegenden Markov-Prozess. Der Separationsabstand zur Zeit t ist damit gleich der Wahrscheinlichkeit, dass die stark stationäre Zeit größer als t ist.
Durch Untersuchung der Verteilung der Ersteintrittszeit des absorbierenden Geburts- und Todesprozesses in Kapitel 4 erhalten wir eine Darstellung des Separationsabstands als Summe unabhängiger, exponentialverteilter Zufallsgrößen. Die Eigenwerte des zugrundeliegenden Markov-Prozesses parametrisieren dabei diese Zufallsgrößen. Dies ist ein klassisches Resultat, welches auf Keilson [Kei79] bzw. Karlin und Mcgregor [KM59] zurückgeht. Mittlerweile gibt es einen neuen, stochastischen Beweis von Fill [Fil07], wobei wir aber den klassischen Beweis geben. Mit Hilfe dieser Darstellung des Separationsabstands und einigen elementaren Folgerungen aus der Chebychev-Ungleichung werden wir in Kapitel 5 unseren Hauptsatz über den Separations-Cut-Off bei Geburts- und Todesprozessen beweisen. Des Weiteren erhalten wir Schranken für den Separationsabstand und die Äquivalenz von Pre-Cut-Off und Cut-Off bei Abstandsmessung mit Separation.
Aus der Darstellung des Separationsabstand als Summe unabhängiger, exponentialverteilter Zufallsgrößen erhalten wir zusätzlich Aussagen über die Form des Cut-Offs, also über das Verhalten der Abstandsfunktion im Zeitfenster des abrupten Übergangs zur stationären Verteilung. Einen Satz, der die Form des Cut-Offs bei Geburts- und Todesprozessen und Separationsabstandsmessung beschreibt, werden wir in Kapitel 6 mit Hilfe der Theorie unendlich teilbarer Verteilungen beweisen. In Kapitel 5 und 6 orientieren wir uns dabei wesentlich an den Resultaten von Diaconis und Saloff-Coste in [DSC06].
In Kapitel 7 präsentieren wir das erwähnte Beispiel von Aldous als Beispiel dafür, dass in Totalvariation die Bedingung $\lambda_n \tau_n \to \infty$ nicht notwendigerweise einen Cut-Off impliziert. Des Weiteren werden wir neueste Ergebnisse im Zusammenhang mit dem Peres-Kriterium und dem Cut-Off-Effekt vorstellen. Insbesondere ziehen wir Verbindungen zwischen dem Totalvariations-Cut-Off und dem Separations-Cut-Off bei Geburts- und Todesprozessen [DLP08]. Ein Überblick über noch offene, verwandte Fragen rundet unsere Arbeit ab.

2 Einführung in die Markov-Theorie in diskreter und stetiger Zeit

2.1 Präliminarien

Eine stochastische Folge $M = (M_n)_{n \geq 0}$ auf einem W-Raum $(\Omega, \mathfrak{A}, P)$ mit Werten in dem *Zustandsraum* S heißt *Markov-Kette* (in diskreter Zeit), falls

$$P(M_{n+1} = x_{n+1} | M_0 = x_0, ..., M_n = x_n) = P(M_{n+1} = x_{n+1} | M_n = x_n)$$

für alle $x_i \in S$ mit $0 \leq i \leq n$ und $n \geq 0$. Wir sprechen von einer *endlichen Markov-Kette* (EMK) bzw. einer *diskreten Markov-Kette* (DMK), falls S endlich oder diskret ist. Wir betrachten ausschließlich die Klasse zeitlich homogener Markov-Ketten mit höchstens abzählbar unendlichem Zustandsraum. *Zeitliche Homogenität* bedeutet in diesem Zusammenhang, dass $P(M_{n+1} = x_{n+1} | M_n = x_n)$ nicht von dem Zeitparameter n abhängt.
In diesem Fall ist die Verteilung einer Markov-Kette $M = (M_n)_{n \geq 0}$ wegen des Satzes von Ionescu-Tulcea durch ihre *Startverteilung* λ (die Verteilung von M_0) und die *Elementarwahrscheinlichkeiten*

$$p_{xy} = P(M_{n+1} = y | M_n = x)$$

für alle $x, y \in S$ vollständig determiniert. Die Elementarwahrscheinlichkeiten werden in der sogenannten *Übergangsmatrix* $\boldsymbol{P} = (p_{xy})_{x,y \in S}$ zusammengefasst. Für eine EMK $(M_n)_{n \geq 0}$ mit Übergangsmatrix $\boldsymbol{P}$ und Startverteilung λ, also $P(M_0 = x) = \lambda(x)$ für alle $x \in S$, ist die Verteilung von M_n gegeben durch

$$P(M_n = x) = \lambda \boldsymbol{P}^n(x) = \sum_{y \in S} \lambda(y) \boldsymbol{P}^n(y, x) \qquad \text{für alle } x \in S,$$

wobei $\boldsymbol{P}^n$ die *n-Schritt-Übergangsmatrix* ist, welche folgendermaßen iterativ definiert wird:

$$\boldsymbol{P}^n(x,y) = \sum_{z\in S} \boldsymbol{P}^{n-1}(x,z)\boldsymbol{P}(z,y) \qquad \text{für alle } x,y\in S.$$

In naheliegender Verallgemeinerung betrachten wir nun Markov-Ketten in stetiger Zeit, welche wir allgemein als *Markov-Prozesse* (MP) und im Fall eines höchstens abzählbaren Zustandsraums S als *Markov-Sprungprozesse* (MSP) bezeichnen. Die Gefahr der Explosion eines MSPs, also des Auftretens unendlich vieler Übergänge in endlicher Zeit, schließen wir dadurch aus, dass wir zunächst nur endliche Zustandsräume betrachten. Es gelte also $|S| < \infty$. Markov-Sprungprozesse werden mittels eines *infinitesimalen Erzeugers* definiert, einer sogenannten *Q-Matrix*. Wir betrachten in weiten Teilen dieser Arbeit MSP, deren Q-Matrix folgende spezielle Gestalt hat: Gegeben eine Übergangsmatrix $\boldsymbol{P}$, sei $M = (M_t)_{t\geq 0}$ ein MSP mit infinitesimalem Erzeuger $Q = \boldsymbol{P} - I$. Es genügt —wegen zeitlicher Transformationsmöglichkeiten endlicher MSP— bei der Untersuchung des Cut-Off-Phänomens diese speziellen MSP zu untersuchen, wie wir zu Beginn von Kapitel 5 erläutern. Bei einem MSP verweilt man eine exponentialverteilte Zeit in einem Zustand $x \in S$ und wechselt dann gemäß der Q-Matrix den Zustand. Die Verteilung von M_t ist somit im endlichen Fall eindeutig bestimmt durch die Startverteilung λ und die *zeitstetige Halbgruppe* $H_t = e^{-t(I-\boldsymbol{P})}$ durch die Formel

$$P(M_t = x) = \sum_{y\in S} \lambda(y)H_t(y,x) \qquad \text{für alle } x\in S, t\geq 0.$$

Dabei ist $H_t(x,y) = e^{-t}\sum_{n=0}^{\infty} \frac{t^n\boldsymbol{P}^n(x,y)}{n!}$ für $x,y\in S$, $t\geq 0$ und $\boldsymbol{P}^0 = I$. $\mathfrak{M}(S)$ bezeichne die Menge der endlichen Maße auf S, $\mathfrak{W}(S)$ die Menge der Verteilungen auf S, also der W-Maße auf S. Gegeben eine Übergangsmatrix $\boldsymbol{P}$, nennen wir ein Maß $\xi \in \mathfrak{M}(S)$ *stationär* oder *invariant* bezüglich $\boldsymbol{P}$, falls $\xi \not\equiv 0$ und $\xi\boldsymbol{P} = \xi$ oder äquivalent $\xi \not\equiv 0$ und

$$\sum_{y\in S} \xi(y)\boldsymbol{P}(y,x) = \xi(x) \qquad \text{für alle } x\in S.$$

Ein Maß $\xi \in \mathfrak{M}(S)$ heißt *reversibel*, falls $\xi \not\equiv 0$ und ξ die *detaillierten Gleichgewichtsgleichungen*

$$\xi(x)\boldsymbol{P}(x,y) = \xi(y)\boldsymbol{P}(y,x) \qquad \text{für alle } x,y\in S$$

erfüllt. Ein reversibles Maß ist notwendigerweise stationär. Falls ξ stationär (reversibel) bezüglich $\boldsymbol{P}$ ist, gilt für alle $t\geq 0$, $\xi H_t = \xi$, oder äquivalent $\sum_{y\in S}\xi(y)H_t(y,x) = \xi(x)$ für alle $x\in S$ (im reversiblen Fall: $\xi(x)H_t(x,y) = \xi(y)H_t(y,x)$ für alle $x,y\in S$). Im Fall eines endlichen

Zustandsraums S existiert immer ein stationäres Maß $\xi \in \mathfrak{M}(S)$, die Normierung $\xi^* = \xi/\xi(S)$ ist somit eine stationäre Verteilung, also $\xi^* \in \mathfrak{W}(S)$.

Eine Übergangsmatrix $\boldsymbol{P}$ heißt *irreduzibel*, falls für alle $x, y \in S$ ein $n = n(x, y)$ existiert, so dass $\boldsymbol{P}^n(x, y) > 0$. Ein Zustand $x \in S$ heißt *aperiodisch*, falls $\boldsymbol{P}^n(x, x) > 0$ für genügend großes n, und $\boldsymbol{P}$ heißt aperiodisch, falls alle Zustände aperiodisch sind. Aperiodizität ist eine *Solidaritätseigenschaft*, das heißt aus der Aperiodizität eines Zustands folgt die Aperiodizität aller mit diesem Zustand kommunizierenden Zustände. Bei Irreduzibilität von $\boldsymbol{P}$ existiert eine eindeutige stationäre Verteilung ξ^*, welche überall positiv ist (Satz 10.4 in [Als05a]). Der folgende Satz liefert den Zusammenhang zwischen stationären Verteilungen und dem asymptotischem Verhalten von Markov-Ketten und Markov-Sprungprozessen (siehe Satz 11.1 in [Als05a] und Satz 10.1 in [Als05b]).

Theorem 2.1. ***(Ergodensatz)*** *Sei $\boldsymbol{P}$ eine irreduzible Übergangsmatrix auf einer endlichen Menge S mit stationärer Verteilung ξ^*. Dann gilt*

$$\lim_{t \to \infty} H_t(x, y) = \xi^*(y) \qquad \textit{für alle } x, y \in S.$$

Falls $\boldsymbol{P}$ irreduzibel und aperiodisch ist, dann gilt

$$\lim_{n \to \infty} \boldsymbol{P}^n(x, y) = \xi^*(y) \qquad \textit{für alle } x, y \in S.$$

Bei abzählbar unendlichen Zustandsräumen gelten die Aussagen des Satzes jeweils, falls zusätzlich die Solidaritätseigenschaft *positive Rekurrenz* vorliegt. Ein Zustand $x \in S$ heißt positiv rekurrent, falls die erwartete Rückkehrzeit in diesen Zustand endlich ist. Eine Übergangsmatrix nennen wir *ergodisch*, falls die zugehörige Markov-Kette konvergiert. Bei endlichen Zustandsräumen und unserer Definition einer stationären Verteilung als positive Verteilung ist Ergodizität demnach im zeitstetigen Fall äquivalent zu Irreduzibilität und im zeitdiskreten Fall zu Irreduzibilität und Aperiodizität.

In dieser Arbeit betrachten wir ergodische, endliche MSP und ergodische EMK und untersuchen die Abstandsfunktion der Verteilung des MSPs bzw. der EMK zu ihrer stationären Verteilung. Qualitativ ist die Konvergenz nach Voraussetzung der Ergodizität klar, wir interessieren uns vielmehr für eine quantitative Untersuchung der Konvergenzgeschwindigkeit und insbesondere der Form des Übergangs zur Stationarität. Dazu stellen wir nun geeignete Abstandsbegriffe zwischen Verteilungen vor und definieren im Anschluss den Cut-Off, also die abrupte Konvergenz gegen die stationäre Verteilung.

2.2 Abstände

Seien $\mu, \nu \in \mathfrak{M}(\Omega)$, wobei Ω endlich ist. Wir betrachten ν als Referenzmaß, später wird ν als stationäre Verteilung ξ^* eines MSPs bzw. einer EMK gewählt und der Abstand der Verteilung des MSPs bzw. der EMK zu dieser Verteilung untersucht. Wir möchten einen Abstandsbegriff zwischen den beiden Maßen μ und ν haben, der im Fall $\mu, \nu \in \mathfrak{W}(\Omega)$ jeder endlichen Menge Ω und jedem Paar μ und ν eine reelle Zahl $D(\mu, \nu) \in [0, 1]$ zuordnet, so dass

$$\sup_{\Omega} \sup_{\mu,\nu} D(\mu, \nu) = 1 \tag{2.1}$$

und $D(\mu, \nu) = 0$ genau dann, wenn $\mu = \nu$. Die Variationsnorm ist ein Abstandsbegriff, der diese Eigenschaften besitzt und folgendermaßen definiert ist.

Definition 2.1. Seien $\mu, \nu \in \mathfrak{M}(\Omega)$, wobei Ω endlich ist. Der *Abstand in totaler Variation* (oder kurz die *Variationsnorm*) zwischen μ und ν ist folgendermaßen definiert:

$$D(\mu, \nu) = \|\mu - \nu\|_{TV} \stackrel{def}{=} \max_{A \subset \Omega} |\mu(A) - \nu(A)| \,. \tag{2.2}$$

Ein sensiblerer Abstandsbegriff zwischen W-Maßen, der die Eigenschaften besitzt, ist Separation. Diese ist folgendermaßen definiert:

Definition 2.2. Seien $\mu, \nu \in \mathfrak{W}(\Omega)$, wobei Ω endlich ist. Die *Separation* von μ und ν ist definiert durch

$$s(\mu, \nu) \stackrel{def}{=} \sup_{\omega \in \Omega} \left\{ 1 - \frac{\mu(\omega)}{\nu(\omega)} \right\}. \tag{2.3}$$

Wir nutzen auch die Notation

$$s(t) = s(\mu_t, \nu),$$

falls μ_t die Verteilung einer Markov-Kette bzw. eines Markov-Prozesses zur Zeit t ist. Es gilt $0 \leq s(\mu, \nu) \leq 1$, im Allgemeinen $s(\mu, \nu) \neq s(\nu, \mu)$ und Separation majorisiert Totalvariation:

Bemerkung 2.1. Die Separation ist stets eine obere Schranke für den Abstand in totaler Variation:

$$\|\mu - \nu\|_{TV} \leq \mathrm{s}(\mu, \nu). \tag{2.4}$$

Beweis. Setze $A = \{\omega \in \Omega : \nu(\omega) > \mu(\omega)\}$. Wir erhalten

$$\|\mu - \nu\|_{TV} = \sum_{\omega \in A} (\nu(\omega) - \mu(\omega)) = \sum_{\omega \in A} \nu(\omega) \left(1 - \frac{\mu(\omega)}{\nu(\omega)}\right) \leq \mathrm{s}(\mu, \nu)$$

und damit die Behauptung. □

Verzichtet man auf die Beschränktheit des Abstands für W-Maße durch (2.1), kann man den Abstand in L^p bzw. L^∞ definieren.

Definition 2.3. Seien $\mu, \nu \in \mathfrak{M}(\Omega)$, wobei Ω endlich und ν als positiv vorausgesetzt ist. Der *Abstand in L^p* für $1 \leq p \leq \infty$ zwischen μ und ν ist definiert durch

$$\|\mu - \nu\|_p = \begin{cases} \left(\sum_{\omega \in \Omega} \left|\frac{\mu(\omega)}{\nu(\omega)} - 1\right|^p \nu(\omega)^p\right)^{\frac{1}{p}}, & \text{falls } 1 \leq p < \infty, \\ \max_{\omega \in \Omega} \left|\frac{\mu(\omega)}{\nu(\omega)} - 1\right|, & \text{falls } p = \infty. \end{cases} \tag{2.5}$$

Bemerkung 2.2. Gilt $\mu(\Omega) = \nu(\Omega)$, so ist der Abstand in totaler Variation die Hälfte des Abstands in L^1, weil :

$$\|\mu - \nu\|_1 = \sum_{\omega \in \Omega} \left|\frac{\mu(\omega)}{\nu(\omega)} - 1\right| \nu(\omega) = \sum_{\omega \in \Omega} |\mu(\omega) - \nu(\omega)| = 2 \|\mu - \nu\|_{TV}.$$

Sowohl der Abstand in totaler Variation als auch der Abstand in L^p werden im Laufe dieser Arbeit immer wieder eine Rolle spielen. In dieser Arbeit messen wir den Abstand zwischen W-Maßen in Separation, falls wir es nicht explizit anders angeben. Wir werden in Kapitel 3 eine Verbindung zwischen Separation und stark stationären Zeiten vorstellen und die darauf aufbauende Dualitätstheorie nach [DF90b] und [Fil92] entwickeln. Damit werden wir dann in Kapitel 5 unseren Hauptsatz über Separations-Cut-Off bei Geburts- und Todesprozessen beweisen. Hier stellen wir zunächst die formale Definition des Cut-Offs, also des abrupten Übergangs zur stationären Verteilung, vor.

2.3 Cut-Off

Es sei (Ω_n, ν_n), n=1,2,..., eine Folge (endlicher) W-Räume, wobei jeder einzelne W-Raum mit einer Folge von W-Maßen (μ_n^k), $k = 0, 1, 2, ...$, ausgestattet ist, so dass $\lim_{k \to \infty} D(\mu_n^k, \nu_n) = 0$.

D erfülle die in Abschnitt 2.2 geforderten Eigenschaften. Später werden wir Familien ergodischer Markov-Ketten bzw. Familien ergodischer Markov-Sprungprozesse auf einem Zustandsraum der Form $\Omega_n = \{0, ..., m_n\}$ betrachten, welche zur Zeit k die Verteilung μ_n^k bzw. zur Zeit t die Verteilung γ_n^t besitzen und gegen die stationäre Verteilung $\nu_n = \xi_n^*$ konvergieren.

Definition 2.4. Eine Familie $\big(\Omega_n, \nu_n, (\mu_n^k)_{k=0,1,2,...}\big)_{n=1,2,...}$ zeigt einen *Cut-Off* (genauer: einen *D-Cut-Off*), falls eine Folge $(t_n)_{n\geq 1}$ positiver reeller Zahlen existiert, so dass für alle $\epsilon \in (0,1)$ folgendes gilt:

$$(a) \lim_{n\to\infty} D(\mu_n^{k_n}, \nu_n) = 0, \qquad \text{falls} \quad k_n > (1+\epsilon)t_n \quad \text{für alle genügend großen } n.$$

$$(b) \lim_{n\to\infty} D(\mu_n^{k_n}, \nu_n) = 1, \qquad \text{falls} \quad k_n < (1-\epsilon)t_n \quad \text{für alle genügend großen } n.$$

In der nächsten Definition wird mit der Folge $(b_n)_{n\geq 1}$ das Zeitfenster beschrieben, in dem der Cut-Off stattfindet.

Definition 2.5. Gegeben Folgen $(t_n)_{n\geq 1}$ und $(b_n)_{n\geq 1}$ positiver reeller Zahlen, sagen wir, dass die Familie

$$\big(\Omega_n, \nu_n, (\mu_n^k)_{k=0,1,2,...}\big)_{n=1,2,...}$$

einen (t_n, b_n)-*Cut-Off* (genauer: einen (t_n, b_n)-*D-Cut-Off*) zeigt, falls $\frac{b_n}{t_n} \overset{n\to\infty}{\longrightarrow} 0$ gilt und

$$(a) \quad f_+(c) = \limsup_{n\to\infty} D(\mu_n^{\lceil t_n + cb_n \rceil}, \nu_n) \quad \text{erfüllt} \quad \lim_{c\to\infty} f_+(c) = 0, \tag{2.6}$$

$$(b) \quad f_-(c) = \liminf_{n\to\infty} D(\mu_n^{\lfloor t_n - cb_n \rfloor}, \nu_n) \quad \text{erfüllt} \quad \lim_{c\to\infty} f_-(c) = 1. \tag{2.7}$$

Falls statt der diskreten Familie $(\mu_n^k)_{k=0,1,2,...}$ von W-Maßen eine stetige Familie $(\gamma_n^t)_{t\geq 0}$ vorliegt, modifizieren wir die Definition in naheliegender Weise und definieren f_+ und f_- ohne Gaussklammer.

Die Existenz eines (t_n, b_n)-Cut-Offs impliziert die Existenz eines Cut-Offs. Eine Abschwächung des Cut-Off Begriffs ist der Pre-Cut-Off.

Definition 2.6. Die Familie $(\Omega_n, \nu_n, (\mu_n^k)_{k=0,1,2,...})_{n=1,2,...}$ zeigt einen *Pre-Cut-Off* (genauer: einen *D-Pre-Cut-Off*), falls eine Folge $(t_n)_{n\geq 1}$ positiver reeller Zahlen und Konstanten $0 < c \leq 1 \leq C < \infty$ existieren, so dass

$$(a) \quad \lim_{n\to\infty} D(\mu_n^{k_n}, \nu_n) \longrightarrow 0, \quad \text{falls} \quad k_n \geq Ct_n,$$

$$(b) \quad \lim_{n\to\infty} D(\mu_n^{k_n}, \nu_n) \longrightarrow 1, \quad \text{falls} \quad k_n \leq ct_n.$$

Bemerkung 2.3. (a) In den Definitionen 2.4 und 2.5 muss nicht $\lim_{n\to\infty} t_n = \infty$ gelten. Dies wird bei der zeitlichen Transformation am Anfang von Kapitel 5 von Bedeutung sein.
(b) Weist eine Familie (μ_n^t) mit stetigem Zeitparameter sowohl einen (s_n)-Cut-Off als auch einen (t_n)-Cut-Off auf, so folgt $s_n \sim t_n$ ($\lim_{n\to\infty} \frac{s_n}{t_n} = 1$). Im diskreten Fall gilt dieses Resultat nur, falls $\lim_{n\to\infty} s_n = \infty$. Vergleiche Lemma 5.1.

3 Stark stationäre Zeiten und Dualitätstheorie

Im Hauptsatz dieser Arbeit werden wir die Existenz eines Cut-Offs bei endlichen in 0 gestarteten Geburts- und Todesprozessen in stetiger Zeit nachweisen, wenn wir den Abstand zur stationären Verteilung in Separation messen. Wir werden im ersten Abschnitt dieses Kapitels zunächst den Separationsabstand mit stark stationären Zeiten in Verbindung bringen. In den folgenden beiden Abschnitten werden wir mit der Dualitätstheorie von Diaconis und Fill ein allgemeines Konstruktionsverfahren für stark stationäre Zeiten vorstellen, welches das Problem der Separationsabstandsbestimmung in ein Absorptionszeit-Problem, also ein Ersteintrittszeit-Problem, verwandelt.

3.1 Stark stationäre Zeiten und Separation

Sei $M = (M_t)_{t \geq 0}$ eine diskrete Markov-Kette oder ein Markov-Sprungprozess mit Startpunkt i_0 und stationärer Verteilung ξ^*. Zur Erklärung des Begriffs der stark stationären Zeit stellen wir zunächst den Begriff der randomisierten Stoppzeit vor. Eine randomisierte Stoppzeit für M ist eine Stoppregel für M, so dass die Entscheidung zur Zeit t zu stoppen nur auf der Basis der Entwicklung von M bis zur Zeit t und (möglicher) unabhängiger Randomisierung getroffen wird. Wir formalisieren den Begriff der randomisierten Stoppzeit.

Sei $(\Omega, \mathfrak{A}, P)$ ein gegebener W-Raum. Der Zustandsraum S sei abzählbar. $\mathfrak{F} = (\mathfrak{F}_t)_{0 \leq t < \infty}$ sei eine Filtration von $(\Omega, \mathfrak{A})$ und $\mathfrak{F}_\infty := \sigma \langle \mathfrak{F}_t : 0 \leq t < \infty \rangle$ die kleinste σ-Algebra, welche jedes $\mathfrak{F}_t$ enthält, also die aymptotische Gesamtinformation des Beobachters.

Definition 3.1. Sei $T : \Omega \to [0, \infty]$. Falls in obiger Situation eine Unter-σ-Algebra $\mathfrak{G}$ von $\mathfrak{A}$ existiert, welche unabhängig von $\mathfrak{F}_\infty$ ist, so dass

$$\{T \le t\} \in \sigma \langle \mathfrak{F}_t, \mathfrak{G} \rangle \qquad \text{für alle } 0 \le t < \infty, \tag{3.1}$$

so nennen wir T eine *randomisierte Stoppzeit* bezüglich $(\mathfrak{F}_t)_{t \ge 0}$.

T ist demnach eine Stoppzeit bezüglich der Filtration $(\mathfrak{F}_t, \mathfrak{G})_{t \ge 0}$ von $(\Omega, \mathfrak{A})$. Nun sind wir in der Lage den Begriff der stark stationären Zeit zu definieren.

Definition 3.2. Sei $\mathfrak{F} = (\mathfrak{F}_t)_{t \ge 0}$ eine Filtration des W-Raums $(\Omega, \mathfrak{A}, P)$, bezüglich derer $M = (M_t)_{t \ge 0}$ die Markov-Eigenschaft besitze. M habe die eindeutige stationäre Verteilung ξ^* und nehme Werte in $(S, \mathfrak{S})$ an. Sei T eine randomisierte Stoppzeit bezüglich $\mathfrak{F}$. Wir nennen T eine *stark stationäre Zeit* (SST) für M, falls, bedingt unter $\{T < \infty\}$, der gestoppte Prozess M_T die stationäre Verteilung ξ^* hat und unabhängig von T ist, also

$$P(T \le t, M_T = y) = P(T \le t)\xi^*(y) \tag{3.2}$$

für jedes $0 \le t < \infty$ und $y \in S$.

Satz 3.1. *Die folgenden drei Aussagen sind für eine randomisierte Stoppzeit äquivalent:*
(a) T *ist eine stark stationäre Zeit.*
(b) $P(T \le t, M_t = y) = P(T \le t)\xi^*(y)$ *für alle* $0 \le t < \infty$ *und* $y \in S$.
(c) $P(T \le t, M_u = y) = P(T \le t)\xi^*(y)$ *für alle* $0 \le t < u < \infty$ *und* $y \in S$.

Beweis. Zum Beweis siehe Proposition 2.4 in [Fil91]. □

Bemerkung 3.1. Aus technischen Gründen müssen wir bei der Definition der stark stationären Zeit im zeitstetigen Fall die Vollständigkeit und Rechtsstetigkeit der Filtration $\mathfrak{F} = (\mathfrak{F}_t)_{t \ge 0}$ fordern. Dabei bedeutet Rechtsstetigkeit in diesem Zusammenhang: $\mathfrak{F}$ ist rechtsstetig, falls $\mathfrak{F}_t = \bigcap_{u > t} \mathfrak{F}_u$ für alle $0 \le t < \infty$, und Vollständigkeit, dass $\mathfrak{F}_0$ und damit $F_t, 0 \le t < \infty$, das System aller $\mathfrak{A}$-Nullmengen enthält. Vergleiche dazu [Fil91].

Nun kommen wir zu dem bereits angekündigten Satz, der Separation und stark stationäre Zeiten miteinander verbindet. Dabei betrachten wir zunächst diskrete Markov-Ketten (DMK).

Satz 3.2. *Sei $M = (M_n)_{n\geq 0}$ eine positiv rekurrente DMK bezüglich einer Filtration $(\mathfrak{F}_n)_{n\geq 0}$ mit Zustandsraum S, stationärer Verteilung ξ^* und beliebiger Anfangsverteilung.*
(a) Ist T eine stark stationäre Zeit für M, so gilt

$$s(n) \leq P(T > n)$$

für alle $n \geq 0$.
(b) Es existiert bei endlichem Zustandsraum eine stark stationäre Zeit T für M, so dass

$$s(n) = P(T > n)$$

für alle $n \geq 0$. In diesem Fall heißt T ***minimale stark stationäre Zeit.***

Beweis. Der Beweis ist im Anhang nachzulesen. □

Der folgende Satz ist das stetige Pendant zu Satz 3.2.

Satz 3.3. *Sei $M = (M_t)_{t\geq 0}$ ein ergodischer, nichtexplodierender Markov-Sprungprozess bezüglich einer Filtration $(\mathfrak{F}_t)_{t\geq 0}$ mit Zustandsraum S, stationärer Verteilung ξ^* und beliebiger Anfangsverteilung.*
(a) Ist T eine stark stationäre Zeit für M, so gilt

$$s(t) \leq P(T > t)$$

für alle $0 \leq t < \infty$.
(b) Es existiert bei endlichem Zustandsraum eine stark stationäre Zeit T für M, so dass

$$s(t) = P(T > t)$$

für alle $0 \leq t < \infty$ gilt.

Beweis. Der Beweis ist das Hauptergebnis von [Fil91]. □

Bemerkung 3.2. Verzichtet man bei Satz 3.2(b) auf die Voraussetzung der Endlichkeit des Zustandsraums, kann man die Existenz einer minimalen stark stationären Zeit für gewisse Klassen von Startverteilungen zeigen, wie zum Beispiel in [Als05a] ausgeführt wird. Eine analoge Verallgemeinerungsmöglichkeit von Satz 3.3(b) auf abzählbar unendliche Zustandsräume wird vermutet, der Nachweis ist aber noch offen.

3.2 Dualitätstheorie in diskreter Zeit

3.2.1 Existenz und Eigenschaften der stark stationären dualen Kette

Die Ergebnisse des vorangegangenen Abschnitts erlauben uns eine Bestimmung des Separationsabstands durch minimale stark stationäre Zeiten, also solcher reellwertiger nichtnegativer Zufallsvariablen T, für die $s(t) = P(T > t)$ gilt. Folglich befassen wir uns nun mit der Konstruktion von SST, insbesondere minimaler SST. Dazu stellen wir eine vereinigende Theorie für die Konstruktion von SST vor, welche auf Arbeiten von Diaconis und Fill zurückgeht (siehe [Fil92] und [DF90b]). Diese Dualitätstheorie liefert einen probabilistischen Ansatz für die Beschränkung der Konvergenzgeschwindigkeit gegen die stationäre Verteilung, welche zu Ergebnissen führt, die zum Teil nicht mit anderen Techniken zur Beschränkung der Konvergenzrate wie Kopplung, Fourieranalysis oder Eigenwertuntersuchungen erzielt werden können. Wir konstruieren dazu einen sogenannten stark stationären dualen Prozess (SSD) für unsere zugrundeliegende Markov-Kette (resp. unseren MSP), dessen Zeit zur Absorption eine SST für die ursprüngliche Markov-Kette liefert. Wir transformieren somit das Problem der Bestimmung einer stark stationären Zeit in ein Ersteintrittszeitproblem. Darüber hinaus geben wir Kriterien für die Minimalität der so erhaltenen SST an und betrachten speziell die uns in erster Linie interessierende Klasse der Geburts- und Todesprozesse, für die der SSD eine besonders einfache Gestalt hat. In dieser Klasse liefert uns die Dualitätstheorie den genauen Separationsabstand. Aus beweistechnischen Gründen und wegen der historischen Entwicklung betrachten wir zunächst diskrete Markov-Ketten, bevor wir Markov-Sprungprozesse behandeln.

Stark stationäre Zeiten und Dualität

$M \sim (\lambda, \boldsymbol{P})$ stehe abkürzend dafür, dass $M = (M_n)_{n\geq 0}$ eine zeithomogene Markov-Kette mit Startverteilung λ und Übergangsfunktion $\boldsymbol{P}$ ist. Der zugehörige Zustandsraum S sei abzählbar und der Zeitparameter n diskret. Sei nun $M \sim (\lambda, \boldsymbol{P})$ eine ergodische Markov-Kette auf einem W-Raum $(\Omega, \mathfrak{A}, P)$ mit stationärer Verteilung ξ^*.

Definition 3.3. Sei $X^* = (X_n^*)_{n\geq 0}$ ein stochastischer Prozess auf $(\Omega, \mathfrak{A}, P)$ mit Werten in einem diskreten Zustandsraum S^*. Die folgenden drei Bedingungen seien erfüllt:

(a) Für jedes $n \geq 0$ gilt

$$X_n^* \text{ und die Kette } M \text{ sind bedingt unabhängig unter } M_0, M_1, ..., M_n. \tag{3.3}$$

(b) Es existiert ein Zustand $\infty \in S^*$, so dass

$$P(M_n \in A | X_0^* = x_0^*, X_1^* = x_1^*, ..., X_{n-1}^* = x_{n-1}^*, X_n^* = \infty) = \xi^*(A) \tag{3.4}$$

für alle $n \geq 0$, $A \subset S$ und jeden möglichen Wert von $(X_0^*, ..., X_n^*)$ der Form
$(x_0^*, ..., x_{n-1}^*, x_n^* = \infty)$.
(c) Für den Zustand ∞ gilt:

$$\infty \text{ ist ein absorbierender Zustand für } X^*, \tag{3.5}$$

also $X_m^* = \infty \quad \Rightarrow X_k^* = \infty$ für alle $k \geq m$.
Dann heißt X^* *stark stationärer dualer Prozess* für M.

Der nächste Satz zeigt uns, wie wir einen stark stationären dualen Prozess nutzen können, um eine stark stationäre Zeit zu erhalten. Er zeigt ebenfalls umgekehrt, dass prinzipiell jede stark stationäre Zeit aus einer solchen Konstruktion resultiert.

Satz 3.4. *(a) Sei X^* ein stark stationärer dualer Prozess zu M. Sei $T = T_\infty^*$ die Zeit bis zur Absorption in ∞ für X^*. Dann ist T eine stark stationäre Zeit für M.*
(b) Sei T umgekehrt eine stark stationäre Zeit für M. Sei $S^* = \{0, 1, ...\} \cup \{\infty\}$ und

$$X_n^* = \begin{cases} n, & \textit{falls } T > n, \\ \infty, & \textit{falls } T \leq n. \end{cases} \tag{3.6}$$

Dann ist X^* ein stark stationärer dualer Prozess für M und es gilt $T = T_\infty^*$.

Beweis. (a) $T = T_\infty^*$ ist nach (3.3) eine randomisierte Stoppzeit für M. Für jedes $n \geq 0$ und $A \subset S$ gilt nach (3.4)

$$P(M_n \in A | T = n) = P(M_n \in A | X_0^* \neq \infty, ..., X_{n-1}^* \neq \infty, X_n^* = \infty) = \xi^*(A).$$

(b) Die bedingte Unabhängigkeitsbedingung folgt, weil T eine SST ist. ∞ ist ein absorbierender Zustand nach Definition von X^*. Die möglichen Werte von $(X_0^*, ..., X_n^*)$ sind von der Form (0,..., l-1,∞,...,∞) mit $0 \leq l \leq n + 1$. Falls $0 \leq l \leq n$, folgt

$$P(M_n \in A | X_0^* = 0, ..., X_{l-1}^* = l - 1, X_l^* = \infty, ..., X_n^* = \infty) = P(M_n \in A | T = l) = \xi^*(A)$$

für alle $A \subset S$. Damit ist (3.4) nachgewiesen.

Insgesamt haben wir also gezeigt, dass X^* ein SSD für M ist. $T = T^*_\infty$ folgt direkt aus (3.6). $\square$

In diesem Satz wird also die Einführung des stark stationären dualen Prozesses gerechtfertigt, denn zum einen wird gezeigt, wie man aus dem dualen Prozess eine stark stationäre Zeit für die zugrundeliegende Kette erhält und zum anderen wird bewiesen, dass jede stark stationäre Zeit aus der Konstruktion eines solchen dualen Prozesses resultiert.

Im Beweis der Rückrichtung wird ein dualer Prozess X^* aus einer stark stationären Zeit T konstruiert ohne Berücksichtigung der Verteilungen $P(M_n = \cdot | X^*_0 = x^*_0, ..., X^*_n = x^*_n)$, außer im Fall $x^*_n = \infty$. Diese Konstruktion ist nicht eindeutig, es gibt im Allgemeinen viele stark stationäre duale Prozesse zu einer stark stationären Zeit. Die Benutzung eines dualen Prozesses bietet den Vorteil, dass Kenntnisse über die Verteilung des dualen Prozesses vor der Absorption herangezogen werden können, um bessere Schranken für den Abstand in Totalvariation als den Separationsabstand zu finden. Einzelheiten dazu finden sich in Abschnitt 2.5 in [DF90b].

Als erstes Beispiel für einen stark stationären dualen Prozess erinnern wir an den im Anhang nachlesbaren Beweis von Satz 3.2 und die dortige Konstruktion einer minimalen SST. Wir beschreiben diese Konstruktion in der Terminologie des dualen Prozesses. Sei $M = (M_n)_{n\geq 0}$ eine positiv rekurrente DMK mit stationärer Verteilung ξ^*. $s(n)$ bezeichne den Separationsabstand von $M = (M_n)_{n\geq 0}$ zur stationären Verteilung ξ^* zur Zeit n. Gegeben $M_0 = x_0$, setze

$$M^*_0 = \infty \quad \text{mit Wahrscheinlichkeit} \quad (1 - s(0))\xi^*(x_0)/\lambda(x_0), \quad \text{und} \tag{3.7}$$

$$M^*_0 = 0 \quad \text{mit Wahrscheinlichkeit} \quad 1 - P(M^*_0 = \infty), \tag{3.8}$$

mit von M unabhängiger Randomisierung. Induktiv definieren wir M^* folgendermaßen: Angenommen $M_0 = x_0, ..., M_{n-1} = x_{n-1}$ sind gegeben und $M^*_0 = x^*_0, ..., M^*_{n-1} = x^*_{n-1}$ wurden festgelegt. Gegeben $M_n = x_n$, setze

$$M^*_n = \infty \qquad \text{falls} \quad M^*_{n-1} = \infty, \quad \text{und} \tag{3.9}$$

$$M^*_n = \infty \quad \text{mit Wahrscheinlichkeit} \quad \frac{s(n-1) - s(n)}{s(n-1) - s(n, x_n)}, \text{ falls } M^*_{n-1} \neq \infty. \tag{3.10}$$

Dann ist M^* nach dem Beweis von Satz 3.2 ein stark stationärer dualer Prozess für M und die erhaltene stark stationäre Zeit ist minimal für M.

Diese Konstruktion eines SSDs von Aldous und Diaconis benötigt die Separationsfunktion als Input und damit die Kenntnis der Verteilung der Markov-Kette zu jeder Zeit n. In aller Regel ist der Grund für die Konstruktion stark stationärer Zeiten aber gerade eine a priori nicht bekannte Separationsfunktion s mittels Satz 3.2 zu beschränken. Also ist diese Konstruktion nur von theoretischem Interesse. In Abschnitt 3.2.2 geben wir eine praktische Konstruktion eines stark stationären dualen Prozesses für eine Vielzahl von Klassen von Markov-Ketten, welche auch die Klasse der Geburts- und Todesprozesse umfasst.

Bemerkung 3.3. Im Allgemeinen ist ein stark stationärer dualer Prozess nicht notwendigerweise markovsch. Der duale Prozess aus dem vorangegangenen Beispiel besitzt jedoch die Eigenschaft, dass sowohl der bivariate Prozess $(M^*, M) = (M_n^*, M_n)_{n\geq 0}$ als auch der duale Prozess M^* Markov-Ketten sind. Startverteilung λ^* und Übergangsmatrix $\boldsymbol{P}^*$ für M^* sind gegeben durch

$$\lambda^*(0) = s(0) = 1 - \lambda^*(\infty),$$

$$\boldsymbol{P}^*(n-1, n) = s(n)/s(n-1) = 1 - \boldsymbol{P}^*(n-1, \infty), \quad n = 1, 2, ...,$$

$$\boldsymbol{P}^*(\infty, \infty) = 1.$$

Ein Übergangskern —im Folgenden auch *Link* genannt— zwischen M und M^* kann folgendermaßen definiert werden:

$$\Lambda(x^*, x) := P(M_n = x | M_0^* = x_0^*, M_1^* = x_1^*, ..., M_{n-1}^* = x_{n-1}^*, M_n^* = x^*).$$

Dieser hängt im Beispiel nicht von $x_0^*, x_1^*, ..., x_{n-1}^*$ ab und ist definiert durch

$$\begin{aligned} \Lambda(n, \cdot) &= [P(M_n = \cdot) - (1 - s(n))\xi^*(\cdot)]/s(n), \quad n = 0, 1, ..., \\ \Lambda(\infty, \cdot) &= \xi^*. \end{aligned}$$

Nun werden wir allgemein versuchen, zu gegebener Markov-Kette M, einen stark stationären dualen Prozess M^* zu konstruieren, welcher selber eine Markov-Kette ist und (M^*,M) zu einer bivariaten Markov-Kette macht. Zunächst werden wir dazu die parallele Konstruktion von (M^*,M) besprechen. Sobald wir geklärt haben, unter welchen Umständen eine bivariate Markov-Kette mit den gewünschten Eigenschaften existiert, werden wir zeigen, wie man M^* aus einer Realisierung von M konstruiert. Die Bedingungen für die Existenz einer bivariaten

Markov-Kette mit den Koordinatenprozessen M und M^* liefert das Algebraische Dualitätstheorem.

Seien $\boldsymbol{P}$ und $\boldsymbol{P^*}$ stochastische Matrizen auf diskreten Mengen S bzw. S^* und $\lambda \in \mathfrak{W}(S)$ bzw. $\lambda^* \in \mathfrak{W}(S^*)$. Λ sei ein Link zwischen S und S^*. Wir suchen eine bivariate Markov-Kette $(M^*,M)=(M_n^*, M_n)_{n\geq 0}$, so dass M eine zeithomogene, ergodische Markov-Kette mit Startverteilung λ und Übergangsmatrix $\boldsymbol{P}$ ist, M^* eine zeithomogene, ergodische Markov-Kette mit Startverteilung λ^* und Übergangsmatrix $\boldsymbol{P}^*$ ist, und M mit M^* durch Λ folgendermaßen verknüpft ist:

$$P(M_n = \cdot | M_0^* = x_0^*, ..., M_n^* = x_n^*) = \Lambda(x_n^*, \cdot). \tag{3.11}$$

Dies impliziert

$$P(M_n = \cdot | M_n^* = x_n^*) = \Lambda(x_n^*, \cdot). \tag{3.12}$$

Die Problemstellung ist motiviert durch das Ziel einen stark stationären dualen Prozess im Sinne von Definition 3.3 zu konstruieren, M^* wird später der duale Prozess von M sein.
Falls (3.3) erfüllt sein soll, muss notwendigerweise gelten

$$M_{n-1}^* \text{ und } M_n \text{ sind bedingt unabhängig unter } M_{n-1}. \tag{3.13}$$

Nach (3.11) muss außerdem gelten

$$M_{n-1}^* \text{ und } M_n \text{ sind bedingt unabhängig unter } M_n^*. \tag{3.14}$$

Folgendes kommutatives Diagramm dient dem Verständnis der beiden bedingten Unabhängigkeitsbedingungen in (3.13) und (3.14).

$$\begin{tikzcd}
M_{n-1}^* \arrow[r, "P^*"] \arrow[d, "\Lambda"'] & M_n^* \arrow[d, "\Lambda"] \\
M_{n-1} \arrow[r, "P"'] & M_n
\end{tikzcd}$$

Das Algebraische Dualitätstheorem gibt für die Existenz einer bivariaten Markov-Kette, welche obige Bedingungen erfüllen soll, gewisse Anforderungen an die Beziehungen zwischen λ, λ^*, $\boldsymbol{P}$,$\boldsymbol{P}^*$ und Λ. Sind diese erfüllt, konstruieren wir im Beweis eine solche bivariate Markov-Kette. Um Probleme mit Nullereignissen zu vermeiden, nehmen wir im Folgenden an, dass jedes $x^* \in S^*$ erreichbar ist, d.h. $P(M_n^* = x^*) > 0$ für ein $n \geq 0$.

Satz 3.5. ***(Algebraisches Dualitätstheorem)*** *Seien* $(\lambda, \boldsymbol{P})$ *auf* S, $(\lambda^*, \boldsymbol{P}^*)$ *auf* S^* *und eine Übergangsmatrix* Λ *von* S^* *nach* S *gegeben. Es existiert genau dann eine bivariate Markov-Kette* (M^*, M) *mit* $M^* \sim (\lambda^*, \boldsymbol{P}^*)$ *und* $M \sim (\lambda, \boldsymbol{P})$*, welche die bedingten Verteilungen* (3.12) *besitzt und die Unabhängigkeitsbedingungen* (3.13) *und* (3.14) *erfüllt, wenn zwischen* $(\lambda^*, \boldsymbol{P}^*)$ *und* $(\lambda, \boldsymbol{P})$ *bezüglich des Links* Λ *die beiden folgenden algebraischen Dualitätsbedingungen erfüllt sind:*

$$\lambda = \lambda^* \Lambda, \tag{3.15}$$

$$\Lambda \boldsymbol{P} = \boldsymbol{P}^* \Lambda. \tag{3.16}$$

Beweis. Nehmen wir zunächst an, dass (M^*, M) mit den gewünschten Eigenschaften existiert. Wir zeigen die Gültigkeit von (3.15) und (3.16). Zunächst zu (3.15), also $\lambda = \lambda^* \Lambda$:

$$\lambda(x) = P(M_0 = x) = \sum_{x^* \in S^*} P(M_0^* = x^*) P(M_0 = x | M_0^* = x^*) = \sum_{x^* \in S^*} \lambda^*(x^*) \Lambda(x^*, x).$$

Für die letzte Gleichheit haben wir (3.12) genutzt. Damit gilt also $\lambda = \lambda^* \Lambda$.

Nun zeigen wir (3.16), also $\Lambda \boldsymbol{P} = \boldsymbol{P}^* \Lambda$. Dazu bedingen wir $P(M_n = y | M_{n-1}^*)$ einmal unter M_{n-1} und einmal unter M_n^*. Damit ergibt sich

$$\begin{aligned} P(M_n = y | M_{n-1}^* = x^*) &= \sum_{x \in S} P(M_{n-1} = x | M_{n-1}^* = x^*) P(M_n = y | M_{n-1}^* = x^*, M_{n-1} = x) \\ &= \sum_{y^* \in S^*} P(M_n^* = y^* | M_{n-1}^* = x^*) P(M_n = y | M_{n-1}^* = x^*, M_n^* = y^*). \end{aligned}$$

Wir sehen mit (3.12) und (3.13), dass in der ersten Summe der Eintrag (x^*, y) von $\Lambda \boldsymbol{P}$ steht. Mit (3.12) und (3.14) ergibt sich in der zweiten Summe der Eintrag (x^*, y) von $\boldsymbol{P}^* \Lambda$. Damit ist $\Lambda \boldsymbol{P} = \boldsymbol{P}^* \Lambda$ gezeigt.

Nun zur Rückrichtung. Wir werden, gegeben die Dualitätsgleichungen (3.15) und (3.16), eine bivariate Markov-Kette auf $\boldsymbol{S} := ((x^*, x) : \Lambda(x^*, x) > 0)$ angeben, welche die Bedingungen erfüllt. Als Startverteilung wählen wir

$$\boldsymbol{\lambda}(x^*, x) = \lambda^*(x^*) \Lambda(x^*, x) \tag{3.17}$$

und als Übergangsfunktion

$$\mathbf{P}((x^*, x), (y^*, y)) = \begin{cases} \boldsymbol{P}(x, y) \boldsymbol{P}^*(x^*, y^*) \Lambda(y^*, y) / \Delta(x^*, y), & \text{falls } \Delta(x^*, y) > 0, \\ 0, & \text{sonst.} \end{cases} \tag{3.18}$$

Hierbei ist Δ definiert als

$$\Delta(x^*, y) := \sum_{y^* \in S^*} \boldsymbol{P}^*(x^*, y^*)\Lambda(y^*, y) = \sum_{x \in S} \Lambda(x^*, x)\boldsymbol{P}(x, y), \tag{3.19}$$

also $\Delta = \boldsymbol{P}^*\Lambda = \Lambda\boldsymbol{P}$. Nachrechnen zeigt, dass $(M^*, M) \sim (\boldsymbol{\lambda}, \mathbf{P})$ die Bedingungen erfüllt. $\square$

Bemerkung 3.4. Die Kette $(M^*, M) \sim (\boldsymbol{\lambda}, \mathbf{P})$ erfüllt (3.11) und es gilt

$$P(M_0^* = x_0^* | M_0 = x_0) = \frac{P(M_0^* = x_0^*, M_0 = x_0)}{P(M_0 = x_0)} = \frac{\lambda^*(x_0^*)\Lambda(x_0^*, x_0)}{\lambda(x_0)}. \tag{3.20}$$

Unter der Voraussetzung $\Delta(x_{n-1}^*, x_n) > 0$, gilt

$$\begin{aligned} P(M_n^* = x_n^* | M_0^* = x_0^*, M_0 = x_0; \ldots; M_{n-1}^* = x_{n-1}^*, M_{n-1} = x_{n-1}; M_n = x_n) \\ = \frac{\boldsymbol{P}^*(x_{n-1}^*, x_n^*)\Lambda(x_n^*, x_n)}{\Delta(x_{n-1}^*, x_n)}, \qquad n \geq 1. \end{aligned} \tag{3.21}$$

Außerdem gilt wegen (3.13)

$$P(M_n = x_n | M_0^* = x_0^*, M_0 = x_0; \ldots; M_{n-1}^* = x_{n-1}^*, M_{n-1} = x_{n-1}) = \boldsymbol{P}(x_{n-1}, x_n), \tag{3.22}$$

d.h., dass M bezüglich $(\sigma(M_k, M_k^*)_{k \leq n})_{n \geq 0}$ die Markov-Eigenschaft besitzt.

Konstruktion des dualen Prozesses

Seien $(\lambda, \boldsymbol{P})$ auf S, $(\lambda^*, \boldsymbol{P^*})$ auf S^* und ein Link Λ von S^* nach S gegeben wie im Algebraischen Dualitätstheorem. Angenommen, die Dualitätsbedingungen $\lambda = \lambda^*\Lambda$ und $\Lambda\boldsymbol{P} = \boldsymbol{P}^*\Lambda$ sind erfüllt. Im Beweis des Algebraischen Dualitätstheorems haben wir gezeigt, wie man eine bivariate Kette (M^*, M) mit $M \sim (\lambda, \boldsymbol{P})$ und $M^* \sim (\lambda^*, \boldsymbol{P}^*)$ konstruiert, so dass (3.11) erfüllt ist. Nun klären wir die Frage, wie man aus einer Realisierung von M den dualen Prozess M^* konstruiert, welcher uns die gesuchte stark stationäre Zeit für M mittels Satz 3.4 liefert.

Nehmen wir also an, dass eine Realisierung der Kette $M \sim (\lambda, \boldsymbol{P})$ vorliegt. Wir möchten aus jedem Pfad von M einen Pfad von M^* konstruieren. Dabei konstruieren wir M^* zeitgleich zur Entwicklung von M und es soll (3.3) gelten. Unter der Voraussetzung, dass M eine ergodische Markov-Kette mit stationärer Verteilung ξ^* ist und M^* den absorbierenden Zustand ∞ mit $\Lambda(\infty, \cdot) = \xi^*$ besitzt, sind wir in der Lage die konstruierte Kette M^* als dualen Prozess zu M zu nutzen, dessen Absorptionszeit eine stark stationäre Zeit für M ist. Sollte der absorbierende Zustand ∞ nicht bereits existieren, kann man ihn zu S^* adjungieren und $\lambda^*(\infty) = 0$,

$P^*(\infty, \infty) = 1$ sowie $P^*(\infty, x^*) = P^*(x^*, \infty) = 0$ für $x^* \neq \infty$ setzen. Wir nennen einen stark stationären dualen Prozess, welcher (3.11) erfüllt, einen Λ-*verlinkten dualen Prozess*. Explizit konstruiert man M^* mit Blick auf (3.20) und (3.21) in Bemerkung 3.4 mittels einer algorithmischen Form des Satzes von Bayes. Es gilt

$$\lambda(\cdot) = \sum_{x_0^* \in S^*} \lambda^*(x_0^*)\Lambda(x_0^*, \cdot).$$

λ ist also eine Mischung der Verteilungen $\Lambda(x_0^*, \cdot)$. Wir möchten nun, nachdem wir mit der Verteilung λ ein M_0 gewählt haben, ein M_0^* erzeugen, so dass

$$P(M_0 = \cdot | M_0^* = x_0^*) = \Lambda(x_0^*, \cdot).$$

Der Satz von Bayes liefert die folgende explizite Konstruktion:
Wenn $M_0 = x_0$ beobachtet wird, dann setze

$$M_0^* = x_0^* \quad \text{mit Wahrscheinlichkeit} \quad \lambda^*(x_0^*)\Lambda(x_0^*, x_0)/\lambda(x_0). \tag{3.23}$$

Induktiv setzen wir diese Konstruktion fort. Angenommen $M_0 = x_0$,..., $M_{n-1} = x_{n-1}$ wurden beobachtet und entsprechend wurden $M_0^* = x_0^*$,..., $M_{n-1}^* = x_{n-1}^*$ gesetzt. Wenn nun $M_n = x_n$ beobachtet wird, setze

$$M_n^* = x_n^* \quad \text{mit Wahrscheinlichkeit} \quad P^*(x_{n-1}^*, x_n^*)\Lambda(x_n^*, x_n)/\Delta(x_{n-1}^*, x_n). \tag{3.24}$$

Dabei ist wie vorher $\Delta = P^*\Lambda$. Wir erhalten mit der Konstruktion (3.17) und (3.18), dass $(M^*, M) \sim (\boldsymbol{\lambda}, \mathbf{P})$. Wir haben also, gegeben die algebraischen Dualitätsbedingungen, einen stark stationären dualen Prozess $M^* \sim (\lambda^*, \boldsymbol{P}^*)$ konstruiert, welcher mit M über Λ verlinkt ist. Mit Satz 3.4 erhalten wir eine stark stationäre Zeit für M. Besonders interessieren uns minimale stark stationäre Zeiten, also solche Zeiten T, bei denen der Separationsabstand zur Zeit n genau gleich der Wahrscheinlichkeit ist, dass der Prozess erst nach n mit T gestoppt wird, also dass

$$s(n) = P(T > n)$$

gilt. Dies motiviert die folgende Definition.

Definition 3.4. Sei X^* ein stark stationärer dualer Prozess zu M. Sei $T = T_\infty^*$ die zugehörige stark stationäre Zeit. Falls T eine minimale stark stationäre Zeit ist, nennen wir den dualen Prozess X^* *scharf*.

Bemerkung 3.5. Wann ist also der von uns in (3.23) und (3.24) konstruierte duale Prozess M^* scharf? Es gilt $P(M_n = \cdot) = P(M_n^* = \cdot)\Lambda$. Hiermit und mit der Definition der Separation sieht man, dass $s(n) = P(T_\infty^* > n)$ gilt, wenn (bei endlichem S genau dann, wenn) ein $x \in S$ mit der Eigenschaft existiert, dass für jedes $S^* \ni x^* \neq \infty$ entweder $P(M_n^* = x^*) = 0$ oder $\Lambda(x^*, x) = 0$ gilt. Ist dies gegeben, kann wegen der Dualitätsbedingung $P(M_n = \cdot) = P(M_n^* = \cdot)\Lambda$ der Separationsabstand von x nur für $x^* = \infty$ kleiner 1 sein. Dann folgt aber

$$1 - \frac{P(M_n = x)}{\xi^*(x)} = 1 - \frac{P(M_n^* = \infty)\Lambda(\infty, x)}{\xi^*(x)} = 1 - P(M_n^* = \infty) = 1 - P(T \leq n) = P(T > n).$$

Damit ist T eine minimale stark stationäre Zeit und der duale Prozess M^* scharf.

3.2.2 Dualität bei Ketten mit monotonem Likelihoodquotienten

Bei der bisherigen Konstruktion stark stationärer dualer Prozesse haben wir keine Einschränkungen bezüglich der Wahl von $S^*, \boldsymbol{P}^*, \lambda^*$ und Λ getroffen. Nur die Beziehungen $\Lambda\boldsymbol{P} = \boldsymbol{P}^*\Lambda$ sowie $\lambda = \lambda^*\Lambda$ mussten nach dem Algebraischen Dualitätstheorem erfüllt sein.

Die nun folgenden Restriktionen bei der Wahl von S^* und Λ erlauben uns weitergehende Aussagen über den dualen Prozess zu treffen. Der Zustandsraum S^* des dualen Prozesses bestehe aus Teilmengen des Zustandsraums S der zugrundeliegenden Markov-Kette M und $\Lambda(x^*, \cdot)$ sei die bei x^* abgeschnittene stationäre Verteilung ξ^* von M. Des Weiteren werden wir voraussetzen, dass die Kette M eine noch zu präzisierende Monotonieeigenschaft besitzt.

Sei S eine endliche, linear geordnete Menge. Es kann also bei 2 Elementen aus S eindeutig bestimmt werden, welches das größere und welches das kleinere ist. Sei $M \sim (\lambda, \boldsymbol{P})$ eine ergodische Markov-Kette mit Zustandsraum $S = \{0, 1, ..., d\}$ und stationärer Verteilung ξ^*. Sei $\frac{\lambda(x)}{\xi^*(x)}$ monoton in x, bei $\lambda = \delta_0$ ist dies offensichtlich erfüllt. Dann ist die Startverteilung λ^* der von uns in diesem Abschnitt konstruierten, mengenwertigen dualen Kette (also der dualen auf dem aus Teilmengen von S bestehenden Zustandsraum S^*) auf Intervalle der Form $\{0, ..., z^*\}$ mit $z^* \in S$ konzentriert. Ein analoges Resultat ergibt sich für die dualen Übergänge $\boldsymbol{P}^*$ bei einer geeigneten Monotoniebedingung an $\boldsymbol{P}$. Durch Identifikation der Mengen der Form $\{0, ..., z^*\}$ mit ihren rechten Endpunkten lässt sich der duale Prozess so interpretieren, dass er Werte auf dem Zustandsraum S annimmt, also auf dem gleichen Zustandsraum wie M lebt. Bei Geburtsund Todesprozessen ergeben sich weitere Vereinfachungen.

Wir werden nun zunächst eine notwendige und hinreichende Bedingung dafür herleiten, dass $(\lambda, \boldsymbol{P})$ bezüglich des Links Λ ein algebraisch duales Paar $(\lambda^*, \boldsymbol{P^*})$ besitzt mit Zustandsraum S^*, der aus Teilmengen von S der Form $\{0, ..., z^*\}$ besteht. Der Übergangskern Λ sei hierbei die Familie der abgeschnittenen stationären Verteilungen

$$\Lambda(x^*, x) = I_{\{0,\dots,x^*\}}(x)\xi^*(x)/H(x^*), \qquad x, x^* \in S. \tag{3.25}$$

Es bezeichnet

$$H(x^*) := \sum_{x \in S: x \leq x^*} \xi^*(x), \quad x^* \in \mathbb{R}, \tag{3.26}$$

die kumulierte Verteilungsfunktion für die stationäre Verteilung ξ^*.

Sobald wir algebraische Dualität vorliegen haben, können wir vermöge der allgemeinen Konstruktion (3.23) und (3.24) einen Λ-verlinkten dualen Prozess M^* zu M konstruieren. Wir werden zeigen, dass dieser duale Prozess scharf ist im Sinne von Definition 3.4.

Nach dem Algebraischen Dualitätstheorem müssen für die algebraische Dualität die Dualitätsbedingungen $\lambda = \lambda^*\Lambda$ und $\Lambda\boldsymbol{P} = \boldsymbol{P}^*\Lambda$ erfüllt sein. Ausgeschrieben bedeutet das für die erste der beiden Bedingungen

$$\lambda(x) = \sum_{x^* \geq x} \lambda^*(x^*)\xi^*(x)/H(x^*), \quad x \in S,$$

umgeformt bedeutet das:

$$\lambda(x^*) = H(x^*) \left[\frac{\lambda(x^*)}{\xi^*(x^*)} - \frac{\lambda(x^*+1)}{\xi^*(x^*+1)}\right], \quad x^* \in S, \tag{3.27}$$

mit der Konvention $\frac{\lambda(d+1)}{\xi^*(d+1)} = 0$. Die Lösung ist nichtnegativ genau dann, wenn der Klammerausdruck in (3.27) nichtnegativ ist, also genau dann, wenn der Likelihoodquotient nicht wachsend in x ist.

Die Beziehung $\Lambda\boldsymbol{P} = \boldsymbol{P}^*\Lambda$ liefert

$$\sum_{x \leq x^*} \xi^*(x)\boldsymbol{P}(x,y)/H(x^*) = \sum_{y^* \geq y} \boldsymbol{P}^*(x^*, y^*)\xi^*(y)/H(y^*),$$

das heißt

$$\Leftrightarrow \left(\sum_{x \leq x^*} \xi^*(x)\boldsymbol{P}(x,y)/\xi^*(y)\right)/H(x^*) = \sum_{y^* \geq y} \boldsymbol{P}^*(x^*, y^*)/H(y^*).$$

Zur Interpretation sei

$$\widetilde{\boldsymbol{P}} = \xi^*(x)\boldsymbol{P}(x,y)/\xi^*(y), \quad x, y \in S, \tag{3.28}$$

der zeitlich invertierte Prozess zu $\boldsymbol{P}$. Sei $\widetilde{M} = (\widetilde{M}_n)_{n\geq 0}$ eine Markov-Kette mit Startverteilung $\widetilde{\lambda}$ und Übergangsfunktion $\widetilde{\boldsymbol{P}}$. Falls $\widetilde{M}$ deterministisch in y gestartet wird, also $\widetilde{\lambda}(y) = 1$ gilt, schreiben wir P_y für das W-Maß P. Die Beziehung $\Lambda\boldsymbol{P} = \boldsymbol{P}^*\Lambda$ sagt uns dann

$$P_y(\widetilde{M}_1 \leq x^*)/H(x^*) = \sum_{y^* \geq y} \boldsymbol{P}^*(x^*, y^*)/H(y^*), \qquad x^*, y \in S. \tag{3.29}$$

Elementare Umformungen liefern den zu (3.29) äquivalenten Ausdruck

$$\boldsymbol{P}^*(x^*, y^*) = \frac{H(y^*)}{H(x^*)} \left[P_{y^*}(\widetilde{M}_1 \leq x^*) - P_{y^*+1}(\widetilde{M}_1 \leq x^*) \right] \qquad x^*, y^* \in S. \tag{3.30}$$

Dabei haben wir $P_{d+1}(\widetilde{M}_1 \leq x^*) = 0$ für alle $x^* \in S$ gesetzt. Ergebnis ist, dass $\Lambda\boldsymbol{P} = \boldsymbol{P}^*\Lambda$ genau dann eine nichtnegative Lösung besitzt, wenn $P_{y^*}(\widetilde{M}_1 \leq x^*)$ in y^* fällt für jedes feste x^*. In diesem Fall ist $\boldsymbol{P}^*$ eindeutig gegeben durch (3.30). Die Eindeutigkeit folgt aus der Dualitätsbedingung $\Lambda\boldsymbol{P} = \boldsymbol{P}^*\Lambda$ und der Tatsache, dass Λ als untere Dreiecksmatrix mit strikt positiven Einträgen auf der Diagonalen invertierbar ist, also gilt $\boldsymbol{P}^* = \Lambda\boldsymbol{P}\Lambda^{-1}$.

Bemerkung 3.6. Die Bezeichung (stochastische) Monotonie für $P_y(\widetilde{M}_1 \leq x)$, falls $P_y(\widetilde{M}_1 \leq x)$ in y fällt für jedes x, geht auf Daley zurück.

Wir erhalten folgendes Anologon zum Algebraischen Dualitätstheorem.

Satz 3.6. *Sei M eine irreduzible und aperiodische Markov-Kette auf $S = \{0, ..., d\}$ mit Startverteilung λ und Übergangsfunktion $\boldsymbol{P}$. ξ^* bezeichne die stationäre Verteilung und H die kumulierte Verteilungsfunktion von ξ^*. Der zeitlich invertierte Prozess zu $\boldsymbol{P}$ sei $\widetilde{\boldsymbol{P}}(x,y) = \frac{\xi^*(y)\boldsymbol{P}(y,x)}{\xi^*(x)}$. Dann hat $(\lambda, \boldsymbol{P})$ genau dann ein algebraisch duales Paar $(\lambda^*, \boldsymbol{P}^*)$ auf $S^* = S$ bezüglich des Links $\Lambda(x^*, x) = I_{\{0,\dots,x^*\}}(x)\xi^*(x)/H(x^*)$, wenn*

$$\lambda(x)/\xi^*(x) \text{ monoton fallend in } x \tag{3.31}$$

und

$$\widetilde{\boldsymbol{P}} \text{ stochastisch monoton ist.} \tag{3.32}$$

In diesem Fall ist das algebraisch duale Paar $(\lambda^*, \boldsymbol{P}^*)$ eindeutig bestimmt durch

$$\lambda(x^*) = H(x^*) \left[\frac{\lambda(x^*)}{\xi^*(x^*)} - \frac{\lambda(x^*+1)}{\xi^*(x^*+1)} \right], \quad x^* \in S, \tag{3.33}$$

und

$$\boldsymbol{P}^*(x^*, y^*) = \frac{H(y^*)}{H(x^*)}\left[P_{y^*}(\widetilde{M}_1 \leq x^*) - P_{y^*+1}(\widetilde{M}_1 \leq x^*)\right], \quad x^*, y^* \in S. \tag{3.34}$$

Bemerkung 3.7. Der Satz gibt uns also bei den vorgenommenen Restriktionen bezüglich der Wahl von Λ und S^* als notwendiges und hinreichendes Kriterium für die Existenz eines algebraisch dualen Paars die Monotoniebedingungen (3.31) und (3.32). Sind diese erfüllt, liefert uns der Satz mit (3.33) und (3.34) die Gestalt der algebraisch dualen Paars.

Analog zur Vorgehensweise im vorangegangenen Abschnitt 3.2.1 konstruieren wir nun pfadweise einen Λ-verlinkten stark stationären dualen Prozess M^* zu einer Kette mit monotonem Likelihoodquotienten $M \sim (\lambda, \boldsymbol{P})$. Zunächst betrachten wir den Induktionsschritt. Mit (3.34) und (3.25) erhalten wir für den Zähler in (3.24)

$$\boldsymbol{P}^*(x^*_{n-1}, x^*_n)\Lambda(x^*_n, x_n)$$

$$= \frac{H(x^*_n)}{H(x^*_{n-1})H(x^*_n)}\left[P_{x^*_n}(\widetilde{M}_1 \leq x^*_{n-1}) - P_{x^*_n+1}(\widetilde{M}_1 \leq x^*_{n-1})\right] I_{\{0,\ldots,x^*_n\}}(x_n)\xi^*(x_n)$$

$$= \frac{\xi^*(x_n)}{H(x^*_{n-1})}\left[P_{x^*_n}(\widetilde{M}_1 \leq x^*_{n-1}) - P_{x^*_n+1}(\widetilde{M}_1 \leq x^*_{n-1})\right] I_{\{0,\ldots,x^*_n\}}(x_n). \tag{3.35}$$

Für den Nenner in (3.24) ergibt sich

$$\Delta(x^*_{n-1}, x_n) = \sum_{y^* \in S} \boldsymbol{P}^*(x^*_{n-1}, y^*)\Lambda(y^*, x_n)$$

$$= \sum_{y^* \in S} \frac{H(y^*)}{H(x^*_{n-1})H(y^*)}\left[P_{y^*}(\widetilde{M}_1 \leq x^*_{n-1}) - P_{y^*+1}(\widetilde{M}_1 \leq x^*_{n-1})\right] I_{\{0,\ldots,y^*\}}(x_n)\xi^*(x_n)$$

$$= \frac{\xi^*(x_n)}{H(x^*_{n-1})} \sum_{y^* \in S}\left[P_{y^*}(\widetilde{M}_1 \leq x^*_{n-1}) - P_{y^*+1}(\widetilde{M}_1 \leq x^*_{n-1})\right] I_{\{0,\ldots,y^*\}}(x_n)$$

$$= \frac{\xi^*(x_n)}{H(x^*_{n-1})} P_{x_n}(\widetilde{M}_1 \leq x^*_{n-1}). \tag{3.36}$$

Also erhalten wir in (3.34) durch Einsetzen von (3.36) und (3.35) und anschliessendes Kürzen

$$P(M^*_n = x^*_n | M^*_{n-1} = x^*_{n-1}, M_n = x_n)$$

$$= \frac{\left[P_{x^*_n}(\widetilde{M}_1 \leq x^*_{n-1}) - P_{x^*_n+1}(\widetilde{M}_1 \leq x^*_{n-1})\right] I_{\{0,\ldots,x^*_n\}}(x_n)}{P_{x_n}(\widetilde{M}_1 \leq x^*_{n-1})}. \tag{3.37}$$

Nun zur Startverteilung von M^*, also zu (3.23). Mit (3.25) und (3.33) erhalten wir

$$P(M_0^* = x_0^* | M_0 = x_0)$$

$$= \frac{H(x_0^*)\xi^*(x_0)}{H(x_0^*)\lambda(x_0)} \left[\frac{\lambda(x_0^*)}{\xi^*(x_0^*)} - \frac{\lambda(x_0^*+1)}{\xi^*(x_0^*+1)} \right] I_{\{0,\ldots,x_0^*\}}(x_0)$$

$$= \frac{\left[\frac{\lambda(x_0^*)}{\xi^*(x_0^*)} - \frac{\lambda(x_0^*+1)}{\xi^*(x_0^*+1)} \right] I_{\{0,\ldots,x_0^*\}}(x_0)}{\frac{\lambda(x_0)}{\xi^*(x_0)}}. \tag{3.38}$$

Bemerkung 3.8. Der in (3.37) und (3.38) konstruierte stark stationäre duale Prozess M^* ist scharf im Sinne von Definition 3.4. Dies lässt sich mit der Bemerkung 3.5 einsehen. Der einzige Zustand, der so groß ist wie $y = d \in S$, ist $y^* = d$. Da aber $\Lambda(x^*, \cdot) = \xi^*$ genau dann gilt, wenn $x^* = d$, entspricht $y^* = d$ der Rolle von ∞ aus der Definition eines stark stationären dualen Prozesses. Es gilt also $\Lambda(y^*, d) = 0$ für alle $y^* \neq \infty$ und mit der gleichen Argumentation wie in Bemerkung 3.5 folgt die Schärfe.

Nun werden wir die Klasse der Markov-Ketten vorstellen, für die wir den Separations-Cut-Off beweisen.

3.2.3 Geburts- und Todesprozesse

Ein *Geburts- und Todesprozess* in diskreter Zeit ist eine Markov-Kette M auf einem W-Raum $(\Omega, \mathfrak{A}, P)$ mit Werten in dem Zustandsraum $S = \{0, 1, ..., d\}$ (endlicher Fall) oder $S = \{0, 1, ...\}$ (abzählbar unendlicher Fall), welche nur von einem Zustand in einen benachbarten Zustand springen oder im gegenwärtigen Zustand verharren kann. Wir behandeln nur den endlichen Fall und setzen die Irreduzibilität voraus. Die Startverteilung von M bezeichnen wir mit λ. Die zu M gehörige Übergangsmatrix $\boldsymbol{P}$ hat *Tridiagonalgestalt*, das heißt nur auf der Diagonalen und auf den beiden Nebendiagonalen können positive Einträge stehen und alle anderen Einträge sind gleich 0. Wir schreiben $\boldsymbol{P}(x, x+1) = p_x$, $\boldsymbol{P}(x, x-1) = q_x$ und $\boldsymbol{P}(x, x) = r_x$ für die jeweiligen Übergangswahrscheinlichkeiten für den Zustand x. Wir setzen $p_d = q_0 = 0$. Die

Übergangsmatrix $\boldsymbol{P}$ hat somit die folgende Gestalt:

$$\boldsymbol{P} = \begin{pmatrix} r_0 & p_0 & 0 & 0 & 0 & \dots\dots\dots\dots & 0 \\ q_1 & r_1 & p_1 & 0 & 0 & \dots\dots\dots\dots & 0 \\ 0 & q_2 & r_2 & p_2 & 0 & \dots\dots\dots\dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \dots\dots\dots\dots & \vdots \\ 0 & .. & \dots\dots\dots & 0 & q_{d-1} & r_{d-1} & p_{d-1} \\ 0 & \dots\dots & \dots\dots\dots & & 0 & q_d & r_d \end{pmatrix}$$

Irreduzibilität lässt sich bei Geburts- und Todesprozessen äquivalent durch die Forderung $p_x > 0$ für alle $0 \leq x < d$ und $q_x > 0$ für alle $0 < x \leq d$ ausdrücken. Irreduzible Geburts- und Todesprozesse sind reversibel, also gilt $\boldsymbol{P} = \widetilde{\boldsymbol{P}}$, und die stationäre Verteilung ist gegeben durch

$$\xi^*(x) = c \prod_{y=1}^{x} \frac{p_{y-1}}{q_y},$$

wobei $c = \xi^*(0)$ eine normalisierende Konstante ist. Ist $\boldsymbol{P}$ irreduzibel und aperiodisch (=ergodisch), folgt mit Satz 3.6, dass $(\lambda, \boldsymbol{P})$ genau dann eine S-wertige, algebraisch Duale $(\lambda^*, \boldsymbol{P}^*)$ besitzt, wenn die beiden Monotoniebedingungen

$$\frac{\lambda(x)}{\xi^*(x)} \text{ ist fallend in } x, \text{ d.h. } q_{x+1}\lambda(x+1) \leq \ p_x\lambda(x) \text{ für } x < d, \tag{3.39}$$

und

$$\boldsymbol{P} \text{ ist monoton, d.h. } p_x + q_{x+1} \leq 1, \qquad x < d, \tag{3.40}$$

erfüllt sind.

Bemerkung 3.9. Auf die Forderung der Aperiodizität kann verzichtet werden, da (3.40) diese zusammen mit der Irreduzibilität wegen $r_0 = 1 - p_0 > 1 - (p_0 + q_1) \geq 0$ impliziert.

Wie sieht nun der stark stationäre duale Prozess M^* zu einem Geburts- und Todesprozess M auf $S = \{0, ..., d\}$ aus? Satz 3.6 und die Konstruktion (3.37) und (3.38) liefern uns einen dualen Prozess auf $S = \{0, ..., d\}$, welcher ebenfalls ein Geburts- und Todesprozess ist.

Die Startverteilung von M^* ist

$$\lambda^*(x^*) = \begin{cases} \frac{H(x^*)}{\xi^*(x^*)p_{x^*}}[p_{x^*}\lambda(x^*) - q_{x^*+1}\lambda(x^*+1)], & x^* < d, \\ \frac{\lambda(d)}{\xi^*(d)}, & x^* = d. \end{cases} \tag{3.41}$$

Die Übergangsparameter sind gegeben durch

$$\begin{aligned} \boldsymbol{P}^*(x^*, x^*-1) &= q^*_{x^*} = \frac{H(x^*-1)}{H(x^*)} p_{x^*}, \\ \boldsymbol{P}^*(x^*, x^*+1) &= p^*_{x^*} = \frac{H(x^*+1)}{H(x^*)} q_{x^*+1}, \\ \boldsymbol{P}^*(x^*, x^*) &= r^*_{x^*} = 1 - (p_{x^*} + q_{x^*+1}), \\ \boldsymbol{P}^*(x^*, x^*+i) &= \boldsymbol{P}(x^*, x^*-i) = 0 \qquad \text{sonst.} \end{aligned}$$

Der Klammerausdruck in (3.34) liefert, dass M^* nur in benachbarte Zustände springen kann. Es gelte die übliche Konvention $p^*_d = q^*_0 = 0$. H bezeichne die kumulierte Verteilungsfunktion der stationären Verteilung ξ^*. Mit Blick auf (3.37) erhalten wir bei reversiblen Prozessen bei der Konstruktion des stark stationären dualen Prozesses

$$\begin{aligned} &P(M^*_n = x^*_n | M_n = x_n, M^*_{n-1} = x^*_{n-1}) \\ &= \frac{[P_{x^*_n}(M_1 \le x^*_{n-1}) - P_{x^*_n+1}(M_1 \le x^*_{n-1})] I_{\{0,\dots,x^*_n\}}(x_n)}{P_{x_n}(M_1 \le x^*_{n-1})}, \end{aligned} \tag{3.42}$$

wobei der Nenner eine normalisierende Konstante ist.

Bemerkung 3.10. Bei Geburts- und Todesprozessen vereinfacht sich (3.42). Wir unterscheiden drei Fälle:

(a) Falls $x^*_{n-1} \ge x_n + 1$ gilt, so ist (3.42)

$p_{x^*_{n-1}}, 1 - (p^*_{x_{n-1}} + q_{x^*_{n-1}+1}), q_{x^*_{n-1}+1}$ oder 0 , je nachdem, ob $x^*_{n-1} = x^*_n - 1$,

$x^*_{n-1} = x^*_n$, $\quad x^*_{n-1} = x^*_n + 1$ oder $|x^*_{n-1} - x^*_n| > 1$.

(b) Falls $x^*_{n-1} = x_n$ gilt, so ist (3.42) $[1 - (p_{x^*_{n-1}} + q_{x^*_{n-1}+1})]/[1 - p_{x^*_{n-1}}]$, $\quad q_{x^*_{n-1}+1}/[1 - p_{x^*_{n-1}}]$ oder 0, je nachdem, ob $x^*_n = x^*_{n-1}$, $x^*_n = x^*_{n-1} + 1$ oder weder noch ist.

(c) Ist $x^*_{n-1} = x_n - 1$, so ist (3.42) 1, falls $x^*_n = x_n$ und sonst 0.

3.3 Dualitätstheorie in stetiger Zeit

3.3.1 Existenz und Eigenschaften des stark stationären dualen Prozesses

Die Dualitätstheorie im stetigen Fall kommt in weiten Teilen zu ähnlichen Ergebnissen wie die Dualitätstheorie in diskreter Zeit. Insbesondere erhalten wir im Algebraischen Dualitätstheorem die gleichen Dualitätsbedingungen $\lambda = \lambda^*\Lambda$ und $\Lambda\boldsymbol{P}(\cdot) = \boldsymbol{P}^*(\cdot)\Lambda$. Neu ist dabei die äquivalente Darstellung der zweiten Gleichung durch die Erzeuger G und G^*: $\Lambda G = G^*\Lambda$. Für die Klasse der irreduziblen Geburts- und Todesprozesse ergibt sich ein dualer Prozess, welcher ein absorbierender Geburts- und Todesprozess auf dem gleichen Zustandsraum ist. Dieser duale Prozess ist scharf, liefert also eine minimale stark stationäre Zeit für den zugrundeliegenden Geburts- und Todesprozess, und es lässt sich mit der Dualitätsbedingung $\Lambda G = G^*\Lambda$ eine Verbindung zwischen den Eigenwerten der beiden Prozesse herstellen.

Beginnen wir nun mit unserer Definition des stark stationären dualen Prozesses in stetiger Zeit. Wir schreiben $M \sim (\lambda, G, \boldsymbol{P}(\cdot), S)$ als Abkürzung dafür, dass $M = (M_t)_{0\leq t<\infty}$ ein nicht-explodierender Markov-Sprungprozess mit Zustandsraum S, Startverteilung λ, Erzeuger G und Übergangsfunktion $(\boldsymbol{P}(t))_{0\leq t<\infty}$ ist. Wir bezeichnen in diesem Kapitel den Erzeuger mit G und dessen Einträge mit $g_{x,y}$, damit keine Verwechslungsgefahr mit den Todesraten bzw. Todeswahrscheinlichkeiten q_x bei Geburts- und Todesprozessen besteht. Wir sprechen von einer Explosion eines MSPs, falls unendlich viele Übergänge des MSPs in endlicher Zeit auftreten. Hinreichend für die Nichtexplosivität ist die Endlichkeit des Zustandsraums S. Wir setzen S als diskret voraus, also als endlich oder abzählbar unendlich. Es sei $M \sim (\lambda, G, \boldsymbol{P}(\cdot), S)$ ein ergodischer Markov-Sprungprozess mit stationärer Verteilung ξ^* auf dem W-Raum $(\Omega, \mathfrak{A}, P)$. Sei $(\mathfrak{F}_t)_{t\geq 0}$ eine Filtration von $(\Omega, \mathfrak{A})$, bezüglich derer M die Markov-Eigenschaft besitzt, und sei $\mathfrak{F}_\infty := \sigma\left\langle \mathfrak{F}_t : 0 \leq t < \infty \right\rangle$ die asymptotische Gesamtinformation des Beobachters. Aus technischen Gründen und ohne Beschränkung der Allgemeinheit nehmen wir an, dass $(\mathfrak{F}_t)_{t\geq 0}$ eine vollständige, rechtsstetige Filtration ist. In der folgenden Definition ist die σ-Algebra, welche zu dem Zustandsraum S^* gehört, beliebig.

Definition 3.5. Sei $X^* = (X_t^*)_{0\leq t<\infty}$ ein stochastischer Prozess auf $(\Omega, \mathfrak{A}, P)$, der Werte in einem messbaren Zustandsraum S^* annimmt. Die folgenden drei Bedingungen seien erfüllt:
(a) Es existiert eine von $\mathfrak{F}_\infty$ unabhängige Sub-σ-Algebra $\mathfrak{G}$ von $\mathfrak{F}$, so dass

$$X_t^* \text{ messbar ist bezüglich } \sigma\left\langle \mathfrak{F}_t, \mathfrak{G}\right\rangle \text{ für alle } 0 \leq t < \infty. \tag{3.43}$$

(b) Definiere $\mathfrak{F}_t^* := \sigma\left\langle X_s^* : s \leq t\right\rangle, 0 \leq t < \infty$. Es existiert ein Zustand $\infty \in S^*$, so dass

$$P(M_t = x|\mathfrak{F}_t^*) = \xi^*(x) \text{ fast sicher auf } \{X_t^* = \infty\} \tag{3.44}$$

für alle $0 \leq t < \infty$ und $x \in S$.

(c) Für den Zustand ∞ gilt:

$$\infty \text{ ist ein absorbierender Zustand für } X^*, \tag{3.45}$$

also $X_s^* = \infty \quad \Rightarrow X_t^* = \infty$ für alle $t \geq s$.
Dann heißt X^* *stark stationärer dualer Prozess* zu M.

Nun stellen wir analog zum Vorgehen in diskreter Zeit eine Verbindung zwischen stark stationären Zeiten des zugrundeliegenden Markov-Sprungprozesses mit der Zeit bis zur Absorption des stark stationären dualen Prozesses her.

Satz 3.7. *(a) Sei X^* ein stark stationärer dualer Prozess von M. Sei $T = T_\infty^*$ die Zeit bis zur Absorption in ∞ für X^*. Dann ist T eine stark stationäre Zeit für M.*
(b) Sei T umgekehrt eine stark stationäre Zeit für M. Sei $S^* = [0, \infty]$ (mit der Borelschen σ-Algebra) und definiere

$$X_t^* = \begin{cases} t, & \textit{falls } T > t, \\ \infty, & \textit{falls } T \leq t. \end{cases} \tag{3.46}$$

Dann ist X^* ein stark stationärer dualer Prozess von M, und es gilt $T = T_\infty^*$.

Beweis. (a) Für alle $0 \leq t < \infty$ gilt $\{T \leq t\} = \{X_t^* = \infty\} \in \sigma \langle \mathfrak{F}_t, \mathfrak{G} \rangle$ wegen (3.43) in der Definition des stark stationären dualen Prozesses. Also erfüllt T die technische Definition einer randomisierten Stoppzeit aus Abschnitt 3.1. Weiter gilt für $0 \leq t < \infty$

$$P(T \leq t, M_t = x) = P(X_t^* = \infty, M_t = x) = E(P(X_t^* = \infty, M_t = x | \mathfrak{F}_t^*))$$

$$= E(P(M_t = x | \mathfrak{F}_t^*); X_t^* = \infty) = E(\xi^*(x); X_t^* = \infty)$$

$$= P(X_t^* = \infty)\xi^*(x) = P(T \leq t)\xi^*(x).$$

Nach Satz 3.1 ($b \Rightarrow a$) aus Abschnitt 3.1 ist T damit eine stark stationäre Zeit.
(b) Als randomisierte Stoppzeit erfüllt T die Messbarkeitsbedingung (3.43). Der Zustand ∞ ist wegen (3.46) ein absorbierender Zustand für X^*. Es bleibt also zum Nachweis, dass X^* ein stark stationärer dualer Prozess zu M ist, die Bedingung (3.44) zu zeigen. Es gilt $\mathfrak{F}_t^* = \sigma \langle \{T \leq s\} : s \leq t \rangle = \sigma \left\langle T I_{\{T \leq t\}} \right\rangle$. Wenn wir nun beide Seiten von (3.44) mit $I_{\{X_t^* = \infty\}}$ multiplizieren, so sind beide Seiten von (3.44) nichtnegative Zufallsvariablen, welche bezüglich der

σ-Algebra $\mathfrak{F}^*_t$ messbar sind und den Erwartungswert $P(M_t = x, X^*_t = \infty) = P(T \leq t, M_t = x) = P(T \leq t)\xi^*(x) = \xi^*(x)P(X^*_t = \infty)$ haben. Weiter bilden die Ereignisse $\{T \leq s\}$ ein π-System, welches $\mathfrak{F}^*_t$ erzeugt und mit einer analogen Rechnung wie zuvor erhalten wir, dass sich beide Seiten von (3.44) zu $\xi^*(x)P(T \leq s)$ integrieren. Damit ist (3.44) gezeigt und der Beweis, dass X^* ein stark stationärer dualer Prozess zu M ist, erbracht.
Die Gültigkeit von $T = T^*_\infty$ folgt direkt aus (3.46). □

Bemerkung 3.11. Ein *π-System* auf einer Menge Ω ist eine Menge $\mathfrak{P}$ bestehend aus Teilmengen von Ω, so dass
(a)$\mathfrak{P}$ nicht leer ist,
(b)$A \cap B \in \mathfrak{P}$, wenn A und B in $\mathfrak{P}$ sind.
Ein endliches Maß, also insbesondere jedes W-Maß, ist eindeutig bestimmt durch die Werte, die es auf einem, die zugehörige σ-Algebra erzeugenden, π-System annimmt.

Im folgenden Abschnitt stellen wir das Algebraische Dualitätstheorem für Markov-Prozesse vor.

Algebraische Dualität

Seien $\boldsymbol{P}(\cdot)$ und $\boldsymbol{P}^*(\cdot)$ nichtexplodierende Markov-Übergangsfunktionen auf den diskreten Mengen S und S^* und $\lambda \in \mathfrak{W}(S)$ bzw. $\lambda^* \in \mathfrak{W}(S^*)$. Die zugehörigen Erzeuger bezeichnen wir mit G und G^*. Λ sei wie in der diskreten Dualitätstheorie ein Link, also ein Übergangskern, von S^* nach S. Wir möchten einen bivariaten Markov-Prozess $(M^*, M) = (M^*_t, M_t)_{0 \leq t < \infty}$ mit den Randverteilungen

$$M^* \sim (\lambda^*, G^*, \boldsymbol{P}^*(\cdot), S^*) \qquad M \sim (\lambda, G, \boldsymbol{P}(\cdot), S) \tag{3.47}$$

konstruieren, so dass M und M^* derart durch Λ verlinkt sind, dass für alle $0 \leq t < \infty, x^* \in S^*$ und $x \in S$

$$P(M_t = x|\mathfrak{F}^*_t) = \Lambda(x^*, x) \text{ fast sicher auf } \{M^*_t = x^*\} \tag{3.48}$$

gilt. Dabei sei $\mathfrak{F}^*_t := \sigma\langle M^*_s : s \leq t\rangle, 0 \leq t < \infty$. Falls wir $\mathfrak{F}^*_{=t} = \sigma\langle M^*_t\rangle$ definieren, dann erhalten wir aus (3.48)

$$P(M_t = x|\mathfrak{F}^*_{=t}) = \Lambda(x^*, x) \text{ fast sicher auf } \{M^*_t = x^*\}. \tag{3.49}$$

Um Probleme mit Nullereignissen zu vermeiden, nehmen wir die Erreichbarkeit von $x^* \in S^*$

an, d.h. $P(M_t^* = x^*) > 0$ für ein $t \geq 0$. Dann können wir (3.49) schreiben als

$$P(M_t = x | M_t^* = x^*) = \Lambda(x^*, x). \tag{3.50}$$

Wir interessieren uns für den bivariaten Markov-Sprungprozess, da wir zu einem gegebenen Markov-Sprungprozess einen stark stationären dualen Prozess konstruieren möchten, der ebenfalls markovsch ist. Satz 3.7 erlaubt durch Untersuchung der Absorptionszeit des stark stationären dualen Markov-Sprungprozesses Aussagen über stark stationäre Zeiten des zugrundeliegenden Markov-Sprungprozesses zu treffen. Besonders interessieren uns dabei scharfe duale Prozesse, welche minimale stark stationäre Zeiten liefern.

Damit M^* ein stark stationärer dualer Markov-Sprungprozess im Sinne von Definition 3.5 ist, müssen für $0 \leq s \leq t < \infty$ nach (3.43) und (3.48) die beiden folgenden bedingten Unabhängigkeitsbedingungen gelten:

$$M_s^* \text{ und } M_t \text{ sind bedingt unabhängig gegeben } M_s, \tag{3.51}$$

und

$$M_s^* \text{ und } M_t \text{ sind bedingt unabhängig gegeben } M_t^*. \tag{3.52}$$

Wie im diskreten Fall liefert das Algebraische Dualitätstheorem ein notwendiges und hinreichendes Kriterium für die Existenz eines bivariaten Markov-Sprungprozesses, welcher (3.47), (3.50), (3.51) und (3.52) erfüllt. Neu ist dabei die Charakterisierung mittels Erzeugern.

Wir beschränken uns im Folgenden auf endliche Zustandsräume, also $|S| < \infty$ und $|S^*| < \infty$.

Satz 3.8. ***(Algebraisches Dualitätstheorem)*** *Seien $(\lambda, G, \boldsymbol{P}(\cdot))$ auf einem endlichen Zustandsraum S, $(\lambda^*, G^*, \boldsymbol{P}^*(\cdot))$ auf einem endlichen Zustandsraum S^* und eine Übergangsmatrix Λ von S^* nach S gegeben. Genau dann existiert ein bivariater Markov-Sprungprozess (M^*, M) mit $M^* \sim (\lambda^*, G^*, \boldsymbol{P}^*(\cdot))$ und $M \sim (\lambda, G, \boldsymbol{P}(\cdot))$, welcher die bedingten Verteilungen* (3.50) *besitzt und die Unabhängigkeitsbedingungen* (3.51) *und* (3.52) *erfüllt, wenn zwischen $(\lambda^*, \boldsymbol{P}^*(\cdot))$ und $(\lambda, \boldsymbol{P}(\cdot))$ bezüglich des Links Λ die beiden folgenden Dualitätsbedingungen erfüllt sind:*

$$\lambda = \lambda^* \Lambda, \tag{3.53}$$

$$\Lambda \boldsymbol{P}(t) = \boldsymbol{P}^*(t)\Lambda, \qquad \textit{für alle } 0 \leq t < \infty. \tag{3.54}$$

Des Weiteren ist (3.54) *äquivalent mittels Erzeugern ausdrückbar:*

$$\Lambda G = G^* \Lambda. \tag{3.55}$$

Beweis. Die Notwendigkeit der Bedingungen (3.53) und (3.54) für die Existenz des bivariaten Markov-Sprungprozesses (M^*, M) mit den vorgegebenen Eigenschaften wird wie im diskreten Fall gezeigt, siehe Abschnitt 3.2. Die Äquivalenz von (3.54) und (3.55) liefert das Lemma 3.3. Wir zeigen, dass die Bedingungen (3.53) und (3.54) hinreichend sind. Nehmen wir also an, dass (3.53) und (3.54) gelten. Wir möchten einen bivariaten Markov-Sprungprozess konstruieren, welcher (3.50), (3.51) und (3.52) erfüllt und die Randverteilungen (3.47) besitzt. Wie im diskreten Fall wählen wir als bivariaten Zustandsraum $\mathbf{S} := \{(x^*, x) : \Lambda(x^*, x) > 0\}$ und als Startverteilung

$$\boldsymbol{\lambda}(x^*, x) := \lambda^*(x^*)\Lambda(x^*, x). \tag{3.56}$$

Natürliche Wahl für die bivariate Übergangsfunktion ist

$$\mathbf{P}(t) := \lim_{h \downarrow 0} (\mathbf{P}^{(h)})^{\lfloor t/h \rfloor} \quad , \quad 0 \leq t < \infty. \tag{3.57}$$

$\mathbf{P}^{(h)}$ bezeichnet die bivariate Ein-Schritt-Übergangsmatrix für das h-Skelett und ist folgendermaßen definiert

$$\mathbf{P}^{(h)}_{(x^*,x),(y^*,y)} = \begin{cases} \boldsymbol{P}_{x,y}(h)\boldsymbol{P}^*_{x^*,y^*}(h)\Lambda(y^*, y)/\Delta_{x^*,y}(h), & \text{falls } \Delta_{x^*,y}(h) > 0, \\ 0, & \text{für } x = 0. \end{cases} \tag{3.58}$$

Dabei ist $\Delta(h) = \boldsymbol{P}^*(h)\Lambda = \Lambda \boldsymbol{P}(h)$. Vergleiche dazu (3.18) im Beweis des Algebraischen Dualitätstheorems im diskreten Fall. Wir werden zeigen, dass der Grenzwert in (3.57) existiert und dass der bivariate Markov-Sprungprozess mit Übergangsfunktion $\mathbf{P}(\cdot)$ die gewünschten Eigenschaften aufweist. Wir müssen dazu mit (3.57) arbeiten, da die Matrizen $\mathbf{P}^{(t)}$ nicht die Kolmogorov-Chapman Gleichungen erfüllen. Wir werden in Lemma 3.2 zeigen, dass

$$\mathbf{P}^{(h)} = \mathbf{I} + h\mathbf{G} + o(h) \qquad \text{für } h \downarrow 0 \tag{3.59}$$

für die Einheitsmatrix I und eine Matrix $\mathbf{G}$ gilt, die wir in Lemma 3.2 explizit angeben werden. Es folgt aus (3.59), dass der Grenzwert in (3.57) existiert und dass

$$\mathbf{P}(t) = \exp(t\mathbf{G}) \tag{3.60}$$

gilt. Aus (3.59) und der Tatsache, dass $\mathbf{P}^{(h)}$ eine stochastische Matrix ist, folgt, dass $\mathbf{G}$ ein Erzeuger ist. Demnach ist (3.60) eine Übergangsfunktion eines Markov-Sprungprozesses mit Erzeuger $\mathbf{G}$. Sei nun $\mathbf{M} = (M^*, M)$ ein Markov-Sprungprozess mit Startverteilung $\boldsymbol{\lambda}$ und Übergangsfunktion $\mathbf{P}(\cdot)$. Wir behaupten, dass (3.47) erfüllt ist, dass also $M^* \sim (\lambda^*, G^*, \boldsymbol{P}^*(\cdot))$ und $M \sim (\lambda, G, \boldsymbol{P}(\cdot))$ gilt. Wir setzen $\mathbf{x}_k = (x_k^*, x_k)$. Sei $0 = t_0 \leq t_1 \leq ... \leq t_n$. Wir erhalten

$$
\begin{aligned}
&P(M_{t_0} = x_0, ..., M_{t_n} = x_n) \\
&= \sum_{x_0^*,...,x_n^*} \boldsymbol{\lambda}(\mathbf{x}_0) \prod_{k=1}^{n} \mathbf{P}_{\mathbf{x}_{k-1},\mathbf{x}_k}(t_k - t_{k-1}) \\
&= \lim_{h \downarrow 0} \sum_{x_0^*,...,x_n^*} \boldsymbol{\lambda}(\mathbf{x}_0) \prod_{k=1}^{n} ((\mathbf{P}^{(h)})^{\lfloor h^{-1}(t_k - t_{k-1}) \rfloor})_{\mathbf{x}_{k-1},\mathbf{x}_k} \\
&= \lim_{h \downarrow 0} \lambda(x_0) \prod_{k=1}^{n} ((\boldsymbol{P}(h))^{\lfloor h^{-1}(t_{k-1} - t_k) \rfloor})_{x_{k-1},x_k} \\
&= \lim_{h \downarrow 0} \lambda(x_0) \prod_{k=1}^{n} \boldsymbol{P}_{x_{k-1},x_k}(h \lfloor h^{-1}(t_k - t_{k-1}) \rfloor) \\
&= \lambda(x_0) \prod_{k=1}^{n} \boldsymbol{P}_{x_{k-1},x_k}(t_k - t_{k-1})
\end{aligned}
$$

Dabei haben wir für die zweite Gleichung das Algebraische Dualitätstheorem im diskreten Fall verwendet. Für M^* erhalten wir das gewünschte Resultat durch analoge Vorgehensweise:

$$
\begin{aligned}
&P(M_{t_0}^* = x_0^*, ..., M_{t_n}^* = x_n^*, M_{t_n} = x_n) \\
&= \left(\sum_{x_0,...,x_n} \boldsymbol{\lambda}(\mathbf{x}_0) \prod_{k=1}^{n} \mathbf{P}_{\mathbf{x}_{k-1},\mathbf{x}_k}(t_k - t_{k-1}) \right) \Lambda(x_n^*, x_n) \\
&= \left(\lim_{h \downarrow 0} \sum_{x_0,...,x_n} \boldsymbol{\lambda}_0(\mathbf{x}_0) \prod_{k=1}^{n} (\mathbf{P}^{(h)\lfloor h^{-1}(t_k - t_{k-1}) \rfloor})_{\mathbf{x}_{k-1},\mathbf{x}_k} \right) \Lambda(x_n^*, x_n) \\
&= \left(\lim_{h \downarrow 0} \lambda^*(x_0^*) \prod_{k=1}^{n} ((\boldsymbol{P}(h))^{\lfloor \frac{t_k - t_{k-1}}{h} \rfloor})_{x_{k-1},x_k} \right) \Lambda(x_n^*, x_n) \\
&= \left(\lim_{h \downarrow 0} \lambda^*(x_0^*) \prod_{k=1}^{n} \boldsymbol{P}_{x_{k-1},x_k}(h \lfloor h^{-1}(t_k - t_{k-1}) \rfloor) \right) \Lambda(x_n^*, x_n) \\
&= \lambda^*(x_0^*) \prod_{k=1}^{n} \boldsymbol{P}^*_{x_{k-1}^*,x_k^*}(t_k - t_{k-1}) \times \Lambda(x_n^*, x_n)
\end{aligned}
\tag{3.61}
$$

Diese Resultate liefern die gewünschten Ergebnisse $M^* \sim (\lambda^*, \boldsymbol{P}^*(\cdot))$, $M \sim (\lambda, \boldsymbol{P}(\cdot))$ und die Verlinkung (3.48) ist erfüllt. Aus (3.48) folgt wiederum (3.50) und (3.52). Es bleibt also allein noch nachzuweisen, dass (3.51) gilt. Diese Aussage folgt aus dem nachfolgenden Lemma. □

Lemma 3.1. *Sei* $\mathfrak{F}_t := \sigma \langle M_s : 0 \leq s \leq t \rangle, 0 \leq t < \infty$. *Analog definieren wir* $\mathfrak{F}_t^*$. *Dann sind* $\mathfrak{F}_t^*$ *und* M *bedingt unabhängig gegeben* $\mathfrak{F}_t$.

Beweis. Es genügt wegen der Markov-Eigenschaft von M die folgende Gleichheit

$$P(B^* \cap B \cap C) = P(B^* \cap B)p \tag{3.62}$$

für Ereignisse der Form

$$B^* = \{M^*(t_1) = x_1^*, ..., M^*(t_n) = x_n^*\}$$
$$B = \{M(t_1) = x_1, ..., M(t_n) = x_n\}$$
$$C = \{M(u_1) = y_1, ..., M(u_m) = y_m\}$$

zu zeigen, wobei $0 \leq t_1 \leq ... \leq t_n = u_0 \leq u_1 \leq ... \leq u_m, x_n = y_0$, und

$$p = \prod_{k=1}^{n} \boldsymbol{P}_{y_{k-1},y_k}(u_k - u_{k-1})$$

Das diskrete Analogon dieses Lemmas ist (3.22) in Bemerkung 3.4. Anwenden von (3.22) auf das Skelett und anschließende Grenzwertbildung liefert den Beweis. □

Dieses Lemma zeigt also insbesondere die Gültigkeit von (3.51) in dem Beweis des Algebraischen Dualitätstheorems. Das folgende Lemma wird (3.59) zeigen, also

$$\mathbf{P}^{(h)} = \mathbf{I} + h\mathbf{G} + o(h) \qquad \text{für } h \downarrow 0.$$

Sei $\mathbf{S} := \{(x^*, x) : \Lambda(x^*, x) > 0\}$ und $\mathbf{x} = (x^*, x) \in \mathbf{S}$. Wir definieren den Erzeuger $\mathbf{G}(\mathbf{x}, \mathbf{y})$. Für festes $\mathbf{x}$ machen wir dies nun in fünf Fällen für $\mathbf{y} = (y^*, y) \in \mathbf{S}$, welche alle möglichen Fälle abdecken. Wir schreiben Γ für die Matrix $\Lambda G = G^* \Lambda$.

1. $y = x, y^* = x^*$. Dann setze

$$\mathbf{G}(\mathbf{x},\mathbf{y}) := -\left[g_{x^*}^* + g_x + \frac{\Gamma(x^*, x)}{\Lambda(x^*, x)}\right]. \tag{3.63}$$

2. $y \neq x, y^* = x^*$. Dann setze

$$\mathbf{G}(\mathbf{x},\mathbf{y}) := g_{x,y}. \tag{3.64}$$

3. $y = x, y^* \neq x^*$. Dann setze

$$\mathbf{G}(\mathbf{x},\mathbf{y}) := \frac{g_{x^*,y^*}^* \Lambda(y^*, x)}{\Lambda(x^*, x)}. \tag{3.65}$$

4. $y \neq x$, $\Lambda(x^*,y) = 0$, $g^*_{x^*,y^*} + g_{x,y} > 0$. Der Schnitt mit 2. ist leer, denn $\Lambda(x^*,y) = 0$ impliziert $y^* \neq x^*$ wegen der Voraussetzung $\Lambda(y^*,y) > 0$. Des Weiteren gilt in diesem Fall, dass $\Gamma(x^*,y) > 0$ gilt. Wenn zum Beispiel $g^*_{x^*,y^*} > 0$ gilt, dann folgt

$$\Gamma(x^*,y) = \sum_{z^* \in S} g^*_{x^*,z^*} \Lambda(z^*,y)$$

$$= \sum_{z^* \neq x^*} g^*_{x^*,z^*} \Lambda(z^*,y) \geq g^*_{x^*,y^*} \Lambda(y^*,y) > 0.$$

Wir setzen

$$\mathbf{G}(\mathbf{x{,}y}) := \frac{g_{x,y} g^*_{x^*,y^*} \Lambda(y^*,y)}{\Gamma(x^*,y)}. \tag{3.66}$$

5. Für alle anderen Zustände y setzen wir

$$\mathbf{G}(\mathbf{x{,}y}) := 0. \tag{3.67}$$

Lemma 3.2. *Seien $\boldsymbol{P}^{(h)}$ und $\boldsymbol{G}$ auf dem Zustandsraum $\boldsymbol{S} = \{\boldsymbol{x} = (x^*,x) : \Lambda(x^*,x) > 0\}$ definiert durch* (3.58) *beziehungsweise* (3.63) *bis* (3.67)*. Dann gilt* (3.59)*:*

$$\mathbf{P}^{(h)} = \mathbf{I} + h\mathbf{G} + o(h) \qquad \text{für } h \downarrow 0.$$

Beweis. Fixiere $\mathbf{x}$ und bezeichne mit C_i die Klasse der bivariaten Zustände $\mathbf{y}$, welche unter Fall $i, i = 1, ..., 5$, fallen. Wende die asymptotischen Beziehungen $\boldsymbol{P}(h) = I + hG + o(h)$, $\boldsymbol{P}^*(h) = I^* + G^* + o(h)$ und $\Delta(h) = \boldsymbol{P}^*(h)\Lambda = \Lambda + hG^*\Lambda + o(h)$ auf (3.58) an, um die $\mathbf{x{,}y}$ Einträge in (3.59) für $\mathbf{x}$ fest und $\mathbf{y}$ in den Fällen 1 bis 4 zu erhalten. Im fünften Fall bemerken wir zunächst, dass

$$h^{-1} \sum_{\mathbf{y} \in C_5} \mathbf{P}^{(h)}_{\mathbf{xy}} = h^{-1} \left[1 - \sum_{\mathbf{y} \notin C_5} \mathbf{P}^{(h)}_{\mathbf{xy}} \right] \qquad \rightarrow \quad -\sum_{i=1}^{4} \sum_{\mathbf{y} \in C_i} \mathbf{G}(\mathbf{x{,}y})$$

Einsetzen der Ausdrücke (3.63) - (3.66) liefert nach einiger Rechnung folgendes Ergebnis:

$$h^{-1} \sum_{\mathbf{y} \in C_5} \mathbf{P}^{(h)}_{\mathbf{xy}} \rightarrow 0$$

für $h \downarrow 0$. Dies zeigt (3.59) für $\mathbf{y}$ in Fall 5 und das Lemma ist damit bewiesen. $\square$

Bleibt zuletzt noch die Äquivalenz der Formulierung mit Erzeugern und Übergangsmatrizen zu beweisen. Dann ist das Algebraische Dualitätstheorem vollständig bewiesen.

Lemma 3.3. *In der Situation von Lemma 3.2 gilt*

$$\Lambda \boldsymbol{P}(t) = \boldsymbol{P}^*(t)\Lambda \qquad \textit{für alle } 0 \leq t < \infty \tag{3.68}$$

genau dann, wenn

$$\Lambda G = G^* \Lambda. \tag{3.69}$$

Beweis. Falls (3.68) gilt, dann bilde die (rechtsseitigen) Ableitungen in t=0 um (3.69) zu erhalten. Falls umgekehrt (3.69) erfüllt ist, gilt $\Lambda G^n = (G^*)^n \Lambda$ für $n \geq 0$. Multipliziere nun beide Seiten mit $t^n/n!$ und summiere über n, um (3.68) zu erhalten. □

Wir lassen nun die Voraussetzung der Endlichkeit der Zustandsräume S und S^* fallen. Es stellt sich die Frage, welche Ergebnisse der Dualitätstheorie im endlichen Fall auf den abzählbar unendlichen Fall übertragen werden können. Die Notwendigkeit der beiden algebraischen Dualitätsbedingungen hängt nicht von der Endlichkeit von S und S^* ab. In der Rückrichtung des Beweises zeigt sich jedoch das Problem, dass die für den Beweis entscheidende Darstellung der Übergangsfunktionen in Form der Matrixexponentialfunktionen in (3.60) nur für endliche Zustandsräume gilt. Ein möglicher Beweis der Rückrichtung erfordert demnach einen anderen Ansatz. Es zeigt sich, dass die für das weitere Vorgehen notwendigen Resultate weiterhin Bestand haben. Da wir in den nächsten Kapiteln unserer Arbeit nur den endlichen Fall betrachten, verweisen wir für eine ausführlichere Diskussion auf [Fil91], Abschnitt 2.2b, und notieren das für die weitere Entwicklung entscheidende Ergebnis:

Satz 3.9. *Die Dualitätsbedingungen $\lambda = \lambda^* \Lambda$ und $\Lambda G = G^* \Lambda$ seien erfüllt. Falls $\boldsymbol{M} = (M^*, M)$ ein bivariater Markov-Sprungprozess mit Startverteilung $\boldsymbol{\lambda}$ aus* (3.56) *und Erzeuger $\boldsymbol{G}$ aus* (3.63)-(3.67) *ist, dann gilt $M^* \sim (\lambda^*, G^*)$ und* (3.48), (3.50) *sowie* (3.52) *sind erfüllt.*

Für die Herleitung des Satzes wurde folgende Situation vorausgesetzt, welche auch in unseren nachfolgenden Überlegungen Bestand haben soll. S und S^* seien diskret, M und M^* nichtexplodierend. Wir nehmen an, dass genau ein Zustand $\infty \in S^*$ existiert, so dass $\Lambda(\infty, \cdot)$ eine stationäre Verteilung ξ^* für M ist und ∞ für M^* absorbierend ist. Abgesehen von der möglichen Ausnahme $S(\infty)$ seien die Mengen $S(x^*)$ alle endlich. Dabei gelte für jedes $x^* \in S^*$ die Notation $S(x^*) = \{x \in S : \Lambda(x^*, x) > 0\}$.

Wir konstruieren nun mit Hilfe des Algebraischen Dualitätstheorems bzw. Satz 3.9 zu einem Markov-Sprungprozess $M \sim (\lambda, G)$ einen stark stationären dualen Markov-Sprungprozess M^*.

Konstruktion des dualen Prozesses

Seien $(\lambda, G, \boldsymbol{P}(\cdot))$ auf S, $(\lambda^*, G^*, \boldsymbol{P}^*(\cdot))$ auf S^* und ein Link von S^* nach S gegeben. $\boldsymbol{P}(\cdot)$ sei ergodisch mit stationärer Verteilung ξ^* und die Dualitätsbedingungen $\lambda = \lambda^*\Lambda$ und $\Lambda G = G^*\Lambda$ seien erfüllt. Sei $M \sim (\lambda, G, \boldsymbol{P}(\cdot))$ gegeben auf einem W-Raum $(\Omega, \mathfrak{A}, P)$. Wir konstruieren nun einen stochastischen Prozess M^* auf dem selben W-Raum $(\Omega, \mathfrak{A}, P)$, so dass $\mathbf{M} = (M^*, M)$ ein bivariater Markov-Sprungprozess ist mit Startverteilung $\boldsymbol{\lambda}$ in (3.56) und Erzeuger $\mathbf{G}$ aus (3.63) - (3.67). Nach Satz 3.9 gilt $M^* \sim (\lambda^*, G^*)$ und die Verlinkungsbedingung (3.48) ist erfüllt. Aus der Konstruktion wird ersichtlich sein, dass (3.43) erfüllt ist; der Markov-Sprungprozess M^* ist also ein die Verlinkungsbedingung (3.48) erfüllender, stark stationärer dualer Markov-Sprungprozess zu M.

Beginnen wir nun mit der Konstruktion des dualen Markov-Sprungprozesses M^*. Sei $(Y_k)_{k\geq 0}$ die in M eingebettete, diskrete Markov-Kette, also der Pfad der von M besuchten Zustände. Ohne Beschränkung der Allgemeinheit nehmen wir an, dass $\lambda(M_0) > 0$ und $g_{Y_{k-1},Y_k} > 0$, $k \geq 1$, für jede Realisierung von M gilt. Bei Beobachtung von $M_0 = x_0$, nutze unabhängige Randomisierung und setze

$$M_0^* = x_0^* \qquad \text{mit Wahrscheinlichkeit } \lambda^*(x_0^*)\Lambda(x_0^*, x_0)/\lambda(x_0). \tag{3.70}$$

Wegen der Dualitätsbedingung $\lambda = \lambda^*\Lambda$ summieren sich die Wahrscheinlichkeiten in (3.70) zu eins. Des Weiteren sieht man mit (3.70) direkt, dass fast sicher $x_0 \in S(x_0^*)$ gilt und dass (3.56) die gemeinsame Verteilung von (M_0^*, M_0) ist.
Iterativ setzen wir die Definition des dualen Prozesses M^* fort. Sei $n \geq 1$ und angenommen, dass M^* bis zur Zeit τ_{n-1} des $(n-1)$-ten Übergangs des bivariaten Prozesses (M^*, M) konstruiert wurde und $\tau_0 := 0$ gesetzt wurde. Wir beschreiben nun die Definition von τ_n und von $M^*(\tau_n)$ mit einer exponentialverteilten Zufallsvariable V_n^*. Wir schreiben $\mathbf{x} = (x^*, x)$ für den Wert von

(M^*, M) zur Zeit τ_{n-1}, induktiv erhalten wir $x \in S(x^*)$ fast sicher. Des Weiteren schreiben wir

$$\mathbf{G}(\mathbf{x}) := |\mathbf{G}(\mathbf{x},\mathbf{x})| = -\mathbf{G}(\mathbf{x},\mathbf{x}) \tag{3.71}$$

genauso wie $g_x = |g_{xx}| = -g_{xx}$.

Sei $V^* = V_n^* \sim \operatorname{Exp}(\mathbf{G}(\mathbf{x}) - g_x)$ unabhängig von $V_1^*, ..., V_{n-1}^*$ und dem Prozess M. Wir unterscheiden zwischen dem Fall $(\tau_{n-1} + V^*) < \sigma_n$ oder $(\tau_{n-1} + V^*) > \sigma_n$, wobei σ_n die erste Wechselzeit für M nach τ_{n-1} ist. Gleichheit tritt mit Wahrscheinlichkeit 0 auf wegen der Stetigkeit der Exponentialverteilung. Die Idee für die iterative Definition von M^* ist nun die folgende: Wir lassen eine exponentielle Uhr unabhängig von der exponentiellen Uhr für M laufen. Die Uhr, die zuerst abläuft, zeigt einen Wechsel für den nächsten Übergang von $\mathbf{M}$ an. Falls die Uhr für M zuerst abläuft, findet bei M ein Zustandswechsel statt, aber nicht bei M^*, außer dieser wird notwendig, um die Beziehung $x \in S(x^*)$ zu erhalten. Falls die andere Uhr zuerst abläuft, so verändern wir nur den Wert von M^* und nicht von M. Formal bedeutet das folgendes:

(a) Falls $\tau_{n-1} + V^* > \sigma_n$, setze $\tau_n = \sigma_n$. Falls zur Zeit τ_n M nach $y \neq x$ springt (so dass $g_{xy} > 0$), setze

$$M_{\tau_n}^* = y^* \qquad \text{mit Wahrscheinlichkeit } \mathbf{G}(\mathbf{x},\mathbf{y})/g_{xy} \tag{3.72}$$

mit $\mathbf{y} = (y^*, y)$. Mit (3.64) folgt — falls $y \in S(x^*)$ — , dass $M_{\tau_n}^*$ fast sicher auf x^* gesetzt wird und nur M den Zustand zur Zeit τ_n wechselt. Falls $y \notin S(x^*)$, dann wird der Wert y^*, der $y \in S(y^*)$ erfüllt für $M_{\tau_n}^*$ gemäß (3.66) mit folgender Wahrscheinlichkeit gewählt:

$$g_{x^*,y^*}^* \Lambda(y^*, y)/\Gamma(x^*, y). \tag{3.73}$$

Da $\Gamma = G^*\Lambda$ ist, summiert sich (3.73) zu eins. In diesem Fall verändern sowohl M als auch (zwangsweise) M^* den Zustand zur Zeit τ_n.

(b) Falls $\tau_{n-1} + V^* < \sigma_n$, dann ist der exponentielle Parameter
$\mathbf{G}(\mathbf{x}) - g_x = \sum_{y^* \neq x^* : x \in S(y^*)} \mathbf{G}((x^*, x), (y^*, y))$ ohne Beschränkung der Allgemeinheit strikt positiv. Setze $\tau_n = \tau_{n-1} + V^*$. Dann setze für $y^* \neq x^*$ mit $x \in S(y^*)$

$$M_{\tau_n}^* = y^* \qquad \text{mit Wahrscheinlichkeit } \frac{\mathbf{G}((x^*, x), (y^*, y))}{\mathbf{G}(\mathbf{x}) - g_x}. \tag{3.74}$$

In diesem Fall verändert nur M^* den Zustand zur Zeit τ_n und M bleibt im gleichen Zustand.

Wir erhalten wie gewünscht $\mathbf{M} \sim (\boldsymbol{\lambda}, \mathbf{G})$. Besonders interessieren wir uns für stark stationäre duale Prozesse, welche scharf sind im Sinne von Definition 3.4, also deren Zeit zur Absorption eine minimale stark stationäre Zeit für den zugrundeliegenden Markov-Sprungprozess liefert.

Bemerkung 3.12. Der duale Prozess M^*, den wir in diesem Teilabschnitt in (3.72), (3.73) und (3.74) definiert haben, ist für endliches S genau dann scharf, wenn ein $y \in S$ existiert, welches zu keinem $S(y^*)$ gehört, außer zu $S(\infty)$ (wobei die Erreichbarkeitsbedingung jedes Zustands für M^* unterstellt wird). Dies bedeutet, dass ein $y \in S$ existiert, für das $\Lambda(y^*, y) = 0$ für alle $S^* \ni y^* \neq \infty$. Mit der Begründung aus Bemerkung 3.5 folgt die Schärfe.

Wie im diskreten Fall werden wir nun nach der allgemeinen Konstruktion eines stark stationären dualen Markov-Sprungprozesses die bisherigen Resultate sukzessive auf Geburts- und Todesprozesse zuschneiden. Zunächst stellen wir in Analogie zu Abschnitt 3.2.2 den mengenwertigen dualen Prozess vor. Dazu sei $M \sim (\lambda, G, \boldsymbol{P}(\cdot), S)$ ein ergodischer Markov-Sprungprozess mit stationärer Verteilung ξ^*. Sei S^* eine Kollektion nichtleerer Teilmengen von S. Wir nehmen an, dass $S \in S^*$ und dass alle anderen Elemente von S^* endliche Teilmengen von S sind. Für jedes $x^* \in S^*$ sei $\Lambda(x^*, x)$ die Renormierung von ξ^* bezüglich der Menge x^*:

$$\Lambda(x^*, x) = \frac{I_{x^*}(x)\xi^*(x)}{\sum_{y \in x^*} \xi^*(y)}, \qquad x \in S. \tag{3.75}$$

In (3.75) ist I_{x^*} der Indikator der Menge x^*.

3.3.2 Dualität bei Prozessen mit monotonem Likelihoodquotienten

Wir nehmen nun an, dass der diskrete Zustandsraum S linear geordnet ist. Wir können also bei 2 beliebigen Elementen aus S eindeutig entscheiden, welches größer und welches kleiner ist. Wir wählen $S^* = \{0, 1, 2, ...\} \cup \{\infty\}$ und definieren

$$\Lambda(x^*, x) = I_{\{x \leq x^*\}} \xi^*(x)/H(x^*), \qquad x^* \in S^*, x \in S, \tag{3.76}$$

wobei H die Verteilungsfunktion für ξ^* ist. Wir lassen also in (3.75) S^* aus den Elementen S (identifiziert mit ∞) und Teilmengen von S der Form $\{0, 1, ..., x^*\}$ (identifiziert mit dem größten Element x^*) bestehen.

Wir haben im Algebraischen Dualitätstheorem gesehen, dass ein durch Λ verlinkter, nichtexplodierender, stark stationärer dualer Markov-Prozess $M^* \sim (\lambda^*, G^*, \boldsymbol{P}^*(\cdot), S^*)$ für M genau dann konstruiert werden kann, wenn die Dualitätsbedingungen $\lambda = \lambda^* \Lambda$ und $\Lambda G = G^* \Lambda$ erfüllt sind. Für gegebenes (λ, G) und unsere Wahl von Λ in (3.76) kann man diese beiden Gleichungen eindeutig für (λ^*, G^*) lösen. Diese Lösung ist allerdings nicht immer zulässig, d.h. ein Paar einer Startverteilung und eines Erzeugers. Notwendige und hinreichende Bedingungen für die Zulässigkeit der Lösung können wir mittels Übergangsfunktionen ausdrücken wie im diskreten Theorem 2.2 von Diaconis und Fill in [DF90a] (Vergleiche dazu auch Satz 4.6 in [DF90b]). Wir formulieren unseren zusammenfassenden Satz 3.10 mit Erzeugern und bezeichnen den zeitlich invertierten Prozess zu G mit $\widetilde{G}$:

$$\widetilde{g}_{x,y} = \xi^*(y) g_{yx} / \xi^*(x). \tag{3.77}$$

Der Begriff MLR wird in der Bemerkung nach dem Satz erklärt.

Satz 3.10. ***(Dualitätstheorem für MLR-Prozesse.)*** *Der Markov-Sprungprozess* $(\lambda, G, \boldsymbol{P}(\cdot), S)$ *besitzt genau dann einen algebraischen, nichtexplodierenden, dualen Markov-Sprungprozess* $(\lambda^*, G^*, \boldsymbol{P}^*(\cdot), S^*)$ *bezüglich des Links* Λ *aus* (3.76)*, wenn die folgenden Monotonie-Bedingungen*

$$\lambda(x)/\xi^*(x) \quad \textit{fällt in } x, \tag{3.78}$$

$$\sum_{x \le x^*} \widetilde{g}_{y^*,x} \begin{cases} \textit{fällt} \\ \textit{wächst} \\ \textit{fällt} \end{cases} \textit{mit } y^* \textit{ auf} \begin{cases} \{0, ..., x^*\} \\ \{x^*, x^*+1\} \\ \{x^*+1, x^*+2, ...\} \end{cases} \tag{3.79}$$

für jedes feste $x^* = 0, 1, 2, ...$ *erfüllt sind und der Erzeuger* G^* *nichtexplodierend ist. Die Einträge der erzeugenden Matrix des dualen Prozesses sind dabei folgendermaßen definiert:*

$$g^*_{x^*,y^*} := \frac{H(y^*)}{H(x^*)} \sum_{x \le x^*} (\widetilde{g}_{y^*,x} - \widetilde{g}_{y^*+1,x}), \qquad x^*, y^* = 0, 1, ..., \tag{3.80}$$

$$g^*_{x^*,\infty} = \frac{1}{H(x^*)} \lim_{y \uparrow \infty} \downarrow \sum_{x \le x^*} \widetilde{g}_{yx}, \qquad x^* = 0, 1, 2, ..., \tag{3.81}$$

$$g^*_{\infty,y^*} = 0 \qquad , y^* \in S^*. \tag{3.82}$$

In diesem Fall ist die duale Startverteilung eindeutig bestimmt durch

$$\lambda^*(x^*) = H(x^*) \left[\frac{\lambda(x^*)}{\xi^*(x^*)} - \frac{\lambda(x^*+1)}{\xi^*(x^*+1)} \right], \qquad x^* = 0, 1, 2, ..., \tag{3.83}$$

$$\lambda^*(\infty) = \lim_{x\uparrow\infty} \downarrow \frac{\lambda(x)}{\xi^*(x)}, \tag{3.84}$$

und der duale Erzeuger ist eindeutig bestimmt durch (3.80) - (3.82).

Bemerkung 3.13. Der Ausdruck MLR-Kette (Kette mit monotonem Likelihoodquotienten) kann wie in diskreter Zeit erklärt werden (vgl. Bemerkung 4.11 bei Diaconis und Fill in [DF90b]): die Übergangsfunktion $\boldsymbol{P}(\cdot)$ erhält genau dann fallenden Likelihoodquotienten bezüglich ξ^*, wenn eine Übergangsfunktion $\boldsymbol{P}^*(\cdot)$ auf S^* existiert, so dass die algebraische Dualitätsbedingung $\Lambda\boldsymbol{P}(t) = \boldsymbol{P}^*(t)\Lambda$ für alle $t \geq 0$ erfüllt ist.

Wenn die algebraischen Dualitätsbedingungen erfüllt sind, kann man mittels der Konstruktion aus Abschnitt 3.3.1 einen stark stationären dualen Markov-Sprungprozess M^* konstruieren. Wie in diskreter Zeit, muss man dafür die kumulierte Verteilungsfunktion H nicht berechnen. Außerdem ist der duale Markov-Sprungprozess bei endlichem Zustandsraum scharf im Sinne von Definition 3.4, da Λ untere Dreiecksgestalt hat. Es existiert damit ein Zustand, der zu keinem $S(x^*)$ außer zu $S(\infty)$ gehört.

3.3.3 Geburts- und Todesprozesse

In diesem Abschnitt betrachten wir die Klasse der Geburts- und Todesprozesse. Ein Geburts- und Todesprozess in stetiger Zeit ist ein Markov-Sprungprozess M auf einem W-Raum $(\Omega, \mathfrak{A}, P)$ mit Zustandsraum $S = \{0, 1, ..., d\}$ (endlicher Fall) oder $S = \{0, 1, ...\}$ im abzählbar unendlichen Fall, welcher eine exponentialverteilte Zeit in einem Zustand $x \in S$ verweilt und im Anschluss in einen benachbarten Zustand springt. Wir betrachten irreduzible Geburts- und Todesprozesse. Die Startverteilung von M bezeichnen wir mit λ. Die zu M gehörige Q-Matrix hat Tridiagonalgestalt, das bedeutet nur auf der Hauptdiagonalen und auf den beiden Nebendiagonalen können Einträge stehen, welche ungleich 0 sind, und alle anderen Einträge sind gleich 0. Wir schreiben $G(x, x+1) = p_x$ für die Geburtsrate im Zustand x und $G(x, x-1) = q_x$ für die Todesrate im Zustand x, wobei $q_0 = 0$. Die Beträge der Diagonaleinträge $|G(x, x)| = g_x = r_x$ geben die Parameter der exponentialverteilten Verweilzeit an, die Übergangswahrscheinlichkeiten von x nach $x+1$ bzw. von x nach $x-1$ werden durch p_x/r_x bzw. q_x/r_x angegeben. Die erzeugende Matrix G hat also die Gestalt

$$G = \begin{pmatrix} r_0 & p_0 & 0 & 0 & 0 & \dots & 0 \\ q_1 & r_1 & p_1 & 0 & 0 & \dots & 0 \\ 0 & q_2 & r_2 & p_2 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \end{pmatrix}$$

Irreduzibilität lässt sich bei Geburts- und Todesprozessen äquivalent durch die Forderung $p_x > 0$ für $x \geq 0$ und $q_x > 0$ für $x \geq 1$ ausdrücken. Im Fall eines abzählbar unendlichen Zustandsraums besteht Explosionsgefahr des Markov-Sprungprozesses M, im endlichen Fall kann M nicht explodieren. Sei $M \sim (\lambda, G, \boldsymbol{P}(\cdot), S)$ nun ein irreduzibler, positiv rekurrenter Geburts- und Todesprozess. Irreduzible Geburts- und Todesprozesse sind zeitlich reversibel, also gilt $\widetilde{G} = G$. Die stationäre Verteilung ist proportional zu dem Vektor der potenziellen Koeffizienten

$$\xi^*(x) = c \prod_{y=1}^{x} \frac{p_{y-1}}{q_y},$$

wobei $c = \xi^*(0)$ eine normalisierende Konstante ist.

Wir erhalten in Satz 3.10 für $0 \leq x^*, y^* < \infty$,

$$\sum_{x \leq x^*} g_{y^*,x} = \begin{cases} -p_{x^*}, & \text{falls } y^* = x^*, \\ q_{x^*+1}, & \text{falls } y^* = x^* + 1, \\ 0 & \text{sonst.} \end{cases} \tag{3.85}$$

$M = \widetilde{M}$ erfüllt also die stochastische Monotoniebedingung (3.79). Die Matrix G^* aus (3.80) - (3.82) ist der Erzeuger eines Geburts- und Todesprozesses auf $\{0, 1, ...\}$ mit Sterberaten

$$q^*_{x^*} = \frac{H(x^* - 1)}{H(x^*)} p_{x^*}, \qquad x^* = 1, 2, 3, ..., \tag{3.86}$$

und Geburtsraten

$$p^*_{x^*} = \frac{H(x^* + 1)}{H(x^*)} q_{x^*+1}, \qquad x^* = 0, 1, 2, ... \tag{3.87}$$

Mit Reuters [Reu57] (Theorem 11) notwendiger und hinreichender Bedingung für die Nichtexplosivität eines Geburts- und Todesprozesses und einigen einfachen Berechnungen [Bré99] (S.352f.) sieht man, dass der in (3.86) und (3.87) definierte Geburts- und Todesprozess genau dann nichtexplodierend ist, wenn

$$\sum_{x=1}^{\infty} \left(\frac{1}{q_{x+1}} + \frac{p_x}{q_{x+1} q_x} + ... + \frac{p_x \cdots p_1}{q_{x+1} \cdots q_1} \right) = \infty. \tag{3.88}$$

Erfüllt der duale Prozess diese Bedingung (3.88), so ist der Zustand ∞ in endlicher Zeit nicht von einem endlichen Zustand erreichbar und kann deshalb vernachlässigt werden.

Wir starten nun den Geburts- und Todesprozess M deterministisch in 0, für die Startverteilung λ gelte also $\lambda = \delta_0$. Damit ist die in Satz 3.10 formulierte Monotoniebedingung (3.78) an die Startverteilung erfüllt und wegen (3.83) gilt für die Startverteilung λ^* des dualen Prozesses M^* ebenfalls $\lambda^* = \delta_0$. Also existiert nach (3.88) ein nichtexplodierender dualer Markov-Sprungprozess. Um die Anwendung zu vereinfachen, beschreiben wir nun die Konstruktion von τ_n und $M^*_{\tau_n}$ aus Abschnitt 3.3.1 in Abhängigkeit davon, ob $x^* \geq x+1$ oder $x^* = x$. Dabei gelte $(M^*, M) = (x^*, x)$ zur vorherigen Übergangsszeit τ_{n-1} des bivariaten Markov-Sprungprozesses. Wie in Abschnitt 3.3.1 bezeichnen wir mit σ_n die erste M-Übergangszeit nach τ_{n-1}. Die Zufallsvariable $V^* = V^*_n$ sei
(1) exponentialverteilt mit Parameter α und
(2) unabhängig von $V^*_1, ..., V^*_{n-1}$ und M.

1. Angenommen, dass $x^* \geq x + 1$. Sei $\alpha = p_{x^*} + q_{x^*+1}$.
(a) Falls $\tau_{n-1} + V^* > \sigma_n,$ setze $\tau_n = \sigma_n$ und $M^*(\tau_n) = x^*$.
(b) Falls $\tau_{n-1} + V^* < \sigma_n,$ setze $\tau_n = \tau_{n-1} + V^*$ und

$$M^*_{\tau_n} = \begin{cases} x^* - 1 & \text{mit Wahrscheinlichkeit } \dfrac{p_{x^*}}{p_{x^*} + q_{x^*+1}}, \\ x^* + 1 & \text{mit Wahrscheinlichkeit } \dfrac{q_{x^*+1}}{p_{x^*} + q_{x^*+1}}. \end{cases}$$

2. Angenommen, dass $x^* = x$. Sei $\alpha = q_{x+1}$.
(a) Falls $\tau_{n-1} + V^* > \sigma_n$, setze $\tau_n = \sigma_n$. Falls $M_{\tau_n} = x - 1$, dann lasse den Wert von M^* unverändert: $M^*_{\tau_n} = x$. Falls $M_{\tau_n} = x + 1$, setze $M^*_{\tau_n} = x + 1$.
(b) Falls $\tau_{n-1} + V^* < \sigma_n,$ setze $\tau_n = \tau_{n-1} + V^*$ und $M^*_{\tau_n} = x + 1$.

Zusammenfassend liefert die Dualitätstheorie für irreduzible Geburts- und Todesprozesse die Existenz eines dualen Prozesses, welcher ein absorbierender Geburts- und Todesprozess auf dem gleichen Zustandsraum S mit den gleichen Eigenwerten ist und die Schärfebedingung in Definition 3.4 erfüllt. In diskreter Zeit ist dieses Ergebnis auf monotone Ketten beschränkt. Die Dualitätstheorie gestattet uns demnach bei Geburts- und Todesprozessen eine Bestimmung des Separationsabstands mittels der Absorptionszeiten des dualen Prozesses, welche minimale stark stationäre Zeiten für den zugrundeliegenden Prozess liefern und somit Gleichheit in Satz 3.2 bzw. Satz 3.3 erfüllen.

Im folgenden Kapitel untersuchen wir Ersteintrittszeiten in 0 gestarteter Geburts- und Todesprozesse. Bei absorbierenden Geburts- und Todesprozessen mit genau einem absorbierenden Zustand ist die Ersteintrittszeit in diesen absorbierenden Zustand die Zeit bis zur Absorption. Wir erhalten eine Summendarstellung der Absorptionszeit und damit der minimalen stark stationären Zeit mit exponentialverteilten Zufallsgrößen als Summanden, welche von den Eigenwerten der Geburts- und Todesprozesse parametrisiert werden.

4 Ersteintrittszeiten bei Geburts- und Todesprozessen

Sei M ein Geburts- und Todesprozess in stetiger Zeit mit Zustandsraum $S = \{0, ..., d\}$, Geburtsrate $p_x > 0$ für $0 \leq x \leq d-1$ und Todesrate $q_x > 0$ für $1 \leq x \leq d$, wobei wie gehabt $q_0 = p_d = 0$ gesetzt werde. In diesem Kapitel untersuchen wir die Verteilung der folgendermaßen definierten Ersteintrittszeit.

Definition 4.1. Die Zufallsvariable

$$T_{mn} = \inf_{t \geq 0} \{M_t = n | M_0 = m\}$$

heißt *Ersteintrittszeit in den Zustand n bei Start in m.*

Wir interessieren uns nun für die Verteilung von T_{0n}, wählen also als Anfangsverteilung $\lambda = \delta_0$. Die Dichte $s_{0n}(\tau)$ der Ersteintrittszeit T_{0n} ist:

$$s_{0n}(\tau) = -\frac{d}{d\tau} P(M_t < n, 0 < t < \tau | M_0 = 0). \tag{4.1}$$

Analog ist

$$s_n^+(\tau) = s_{nn+1}(\tau) = -\frac{d}{d\tau} P(M_t < n+1, 0 < t < \tau | M_0 = n) \tag{4.2}$$

die Dichte der Zufallsvariable T_{nn+1}. Wir drücken nun die Dichte $s_{0n}(\tau)$ von T_{0n} als Faltung der Dichten $s_i^+(\tau)$ der $T_{ii+1}, 0 \leq i \leq n-1$, aus:

$$s_{0n}(\tau) = s_0^+(\tau) * s_1^+(\tau) * ... * s_{n-1}^+(\tau), \tag{4.3}$$

in Zufallsvariablen bedeutet dies $T_{0n} = T_{01} + ... + T_{n-1n}$. Diese Zufallsvariablen sind wegen der Markov-Eigenschaft von M unabhängig. Mittels eines rekursiven probabilistischen Arguments erhalten wir $s_{0n}(\tau)$ folgendermaßen: Sei $v_n = p_n + q_n$ die Verweilzeit im Zustand n. Die Dichte der Verweilzeit hat die Form $v_n e^{-v_n t}$ und mit Wahrscheinlichkeit q_n/v_n findet ein nachfolgender Wechsel nach $n-1$ statt, mit Wahrscheinlichkeit p_n/v_n ein Wechsel nach $n+1$. Somit erhalten wir:

$$s_n^+(\tau) = \frac{p_n}{v_n} v_n e^{-v_n \tau} + \frac{q_n}{v_n} v_n e^{-v_n \tau} * s_{n-1}^+(\tau) * s_n^+(\tau), \quad n \geq 1. \tag{4.4}$$

Nun betrachten wir die Laplace-Transformierte (L.T.), welche mit dem jeweiligen griechischen Kleinbuchstaben bezeichnet wird. Bei der Berechnung der L.T. von s_n^+ erhalten wir

$$\begin{aligned} \sigma_n^+(s) &= \frac{p_n}{s + p_n + q_n} + \frac{q_n}{s + p_n + q_n} \sigma_{n-1}^+(s) \sigma_n^+(s) \\ \Rightarrow 1 &= \frac{p_n}{(s + p_n + q_n)\sigma_n^+(s)} + \frac{q_n}{s + p_n + q_n} \sigma_{n-1}^+(s) \\ \Rightarrow 1 &= \frac{1}{s + p_n + q_n} \left(\frac{p_n}{\sigma_n^+(s)} + q_n \sigma_{n-1}^+(s) \right) \\ \Rightarrow s + p_n + q_n &= \frac{p_n}{\sigma_n^+(s)} + q_n \sigma_{n-1}^+(s) \\ \Rightarrow \sigma_n^+(s) &= \frac{p_n}{s + p_n + q_n - q_n \sigma_{n-1}^+(s)}, \end{aligned} \tag{4.5}$$

wobei wir für die erste Gleichung den Multiplikationssatz für L.T. und deren Linearität ausgenutzt haben (siehe [Als05c], Kapitel 40-42). Wegen $q_0 = 0$ gilt

$$\sigma_0^+(s) = \frac{p_0}{p_0 + s}, \tag{4.6}$$

und wir erhalten σ_n^+ iterativ aus (4.5) und (4.6). Mit dem Multiplikationssatz für L.T. erhalten wir aus (4.3):

$$\sigma_{0n}(s) = \sigma_0^+(s) ... \sigma_{n-1}^+(s) \tag{4.7}$$

und

$$\sigma_{0n+1}(s) = \sigma_{0n}(s) \sigma_n^+(s). \tag{4.8}$$

Der folgende Satz liefert uns nun Einblick in die Verteilung der Ersteintrittszeit T_{0N}.

Satz 4.1. *Sei M ein irreduzibler Geburts- und Todesprozess mit Zustandsraum $S = \{0, ..., d\}$ und $N \in S$. Dann gilt für die L.T. der Ersteintrittszeit T_{0N}*

$$\sigma_{0N}(s) = \frac{\theta_{N1}}{\theta_{N1} + s} \frac{\theta_{N2}}{\theta_{N2} + s} ... \frac{\theta_{NN}}{\theta_{NN} + s}; \qquad \theta_{Nj} > 0, \tag{4.9}$$

wobei $\theta_{Nj} \neq \theta_{Ni}$ für alle $i \neq j$.

Bemerkung 4.1. Eine $\Gamma(\alpha, \beta)$ verteilte Zufallsgröße X hat die L.T. $\sigma_{\alpha,\beta}(t) = (\frac{\beta}{\beta+t})^\alpha$. Ferner bestimmt die L.T. eines endlichen Maßes, also insbesondere eines W-Maßes, dieses eindeutig, sofern die L.T. für mindestens ein t existiert und endlich ist. Das Theorem zeigt somit wegen der Struktur der L.T. von T_{0N} unter Benutzung des Multiplikationssatzes für L.T., dass die Ersteintrittszeit T_{0N} eine Summe N unabhängiger exponentialverteilter Zufallsvariablen ist. Für einen reinen Geburtsprozess ist dies klar, aber da ein irreduzibler Geburts- und Todesprozess sich nach Erreichen eines Zustands n mit $0 < n < d$ auch wieder nach unten bewegen kann, ist dieses Ergebnis durchaus überraschend und es existiert keine naheliegende stochastische Interpretation. Wir werden dieses Resultat zunächst mit analytischen Mitteln beweisen und im Anschluss eine Verbindung der Parameter mit den Eigenwerten der Q-Matrix M herstellen. Des Weiteren geben wir mittels der Dualitätstheorie eine erste, allerdings wenig naheliegende, stochastische Interpretation des Theorems.

Wir folgen dem Beweis von Keilson [Kei79]. Zunächst beweisen wir zwei Lemmata.

Lemma 4.1. *Außer in Singularitäten ist $\sigma_n^+(s)$ monoton fallend für reelles s.*

Beweis. Wir beweisen das Lemma per Induktion. $\sigma_0^+(s) = \frac{p_0}{p_0+s}$ ist monoton fallend für reelles s. Nehmen wir also an, dass $\frac{d}{ds}\sigma_{n-1}^+(s) < 0$, $s \in \mathbb{R}$, gilt. Wir erhalten außerhalb von Singularitäten mit der Quotientenregel

$$\frac{d}{ds}\sigma_n^+(s) = \frac{-p_n(1 - q_n \frac{d}{ds}\sigma_{n-1}^+(s))}{(s + p_n + q_n - q_n\sigma_{n-1}^+(s))^2} < 0,$$

da der Klammerausdruck im Zähler nach Induktionsvoraussetzung > 0 ist. □

Lemma 4.2. *$\sigma_n^+(s)$ ist eine rationale Funktion und hat einen einfachen Pol zwischen jedem Paar von benachbarten Polen von $\sigma_{n-1}^+(s)$. Alle $n+1$ Pole von $\sigma_n^+(s)$ liegen auf der negativen Halbachse.*

Beweis. Wieder gehen wir zunächst induktiv vor. $\sigma_0^+(s)$ ist eine rationale Funktion. Aus der Rationalität von $\sigma_{n-1}^+(s)$ folgt mit $\sigma_n^+(s) = \frac{p_n}{s+p_n+q_n-q_n\sigma_{n-1}^+(s)}$ induktiv die Rationalität von $\sigma_n^+(s)$. $\sigma_n^+(s)$ hat Pole in den Nullstellen von $s + p_n + q_n - q_n\sigma_{n-1}(s)$. Das vorherige Lemma 4.1 liefert, dass zwischen jedem Paar von Polen von $\sigma_{n-1}^+(s)$ die Ableitung von $s+p_n+q_n-q_n\sigma_{n-1}^+(s)$ positiv ist, also

$$\frac{d}{ds}(s + p_n + q_n - q_n\sigma_{n-1}^+(s)) = 1 - q_n\frac{d}{ds}\sigma_{n-1}^+(s) > 0.$$

Wir erhalten mit dem Zwischenwertsatz für rationale Funktionen und der Monotonie, dass genau eine Nullstelle von $\sigma_{n-1}^{+}(s)$ in diesem Intervall liegt. Diese Nullstelle korrespondiert mit einem einfachen Pol von $\sigma_n^{+}(s)$, da der Nenner von $\sigma_n^{+}(s)$ gerade $s+p_n+q_n-q_n\sigma_{n-1}^{+}(s)$ ist. Des Weiteren liegen alle $n+1$ Pole von $\sigma_n^{+}(s)$ auf der negativen Halbachse. Dies folgt allgemein, da die L.T. nichtnegativer Zufallsgrößen auf $[0,\infty)$ positiv und monoton fallend ist. $\square$

Beweis. [Satz 4.1] Wieder gehen wir induktiv vor. Der Induktionsanfang $\sigma_{01}(s) = \frac{\theta_{11}}{\theta_{11}+s} = \frac{p_0}{p_0+s}$ ist klar. Angenommen, der Satz gilt für N. Dann sehen wir mit (4.8) und $n = N-1$, dass die Nullstellen von $\sigma_{N-1}^{+}(s)$ die Polstellen von $\sigma_{0N-1}(s)$ sind und die Polstellen von $\sigma_{N-1}^{+}(s)$ die Polstellen von $\sigma_{0N}(s)$ sind. Aus $\sigma_N^{+}(s) = \frac{p_N}{s+p_N+q_N-q_N\sigma_{N-1}^{+}(s)}$ folgt, dass die Polstellen von $\sigma_{N-1}^{+}(s)$ die Nullstellen von $\sigma_N^{+}(s)$ sind. Damit fallen die Nullstellen von $\sigma_N^{+}(s)$ mit den Polstellen von $\sigma_{0N}(s)$ zusammen. Wir erhalten also in (4.8) mit $N = n$, dass $\sigma_{0N+1}(s)$ die gewünschte Form hat. $\square$

Die Ersteintrittszeit T_{0N} besitzt also eine Darstellung als Summe unabhängiger exponentialverteilter Zufallsvariablen, wobei alle Parameter ungleich sind.

Nun werden wir Brown und Shao [BS87] folgend einen Zusammenhang zwischen den Parametern und den Eigenwerten des betrachteten Geburts- und Todesprozess M herstellen. Zunächst erinnern wir daran, dass ein Geburts- und Todesprozess reversibel ist; die Übergangsratenmatrix ist demnach ähnlich zu einer symmetrischen Matrix und wegen des Spektralsatzes für reell symmetrische Matrizen sind somit alle Eigenwerte von M reell. Wir erhalten allgemein eine Spektraldarstellung der Übergangsratenmatrix, so dass wir die Ersteintrittszeit T_{jn} von einem Zustand $0 \le j < n$ in den Zustand n folgendermaßen darstellen können:

$$P(T_{jn} > t) = \sum_{i=1}^{n} \gamma_i e^{-\lambda_i t}, \tag{4.10}$$

wobei $\lambda_1, ..., \lambda_n$ die Eigenwerte der Matrix $Q_{(n)}$ sind und $Q_{(n)}$ die $n \times n$-Matrix ist, welche aus den ersten n Spalten und Zeilen der Übergangsratenmatrix Q von M entsteht. Die γ_i in (4.10) hängen von den zugehörigen rechten und linken Eigenvektoren ab.

Ein Geburts- und Todesprozess ist ein Markov-Sprungprozess, bei dem Übergänge vom gegenwärtigen Zustand nur in benachbarte Zustände möglich sind. Man bezeichnet Geburts- und Todesprozesse deshalb auch als *sprungfrei*. Im Folgenden betrachten wir Markov-Sprungprozesse, welche die Eigenschaft der *Sprungfreiheit nach oben* besitzen und lassen die Voraussetzung der *Sprungfreiheit nach unten* fallen. Wir betrachten also Markov-Sprungprozesse auf dem Zustandsraum $S = \{0, 1, ...\}$ mit Übergangsraten

$$\begin{aligned} (i) \quad & Q_{i,i+1} = b_i > 0 \\ (ii) \quad & Q_{i,j} = 0 \quad \text{für } j \geq i+2 \\ (iii) \quad & -Q_{ii} < \infty \quad \text{für alle } i. \end{aligned} \tag{4.11}$$

Die zugehörige erzeugende Matrix Q hat demnach die folgende Gestalt

$$Q = \begin{pmatrix} * & b_0 & 0 & 0 & 0 & \dots & 0 \\ * & * & b_1 & 0 & 0 & \dots & 0 \\ * & * & * & b_2 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \end{pmatrix}.$$

Ein Geburts- und Todesprozess ist insbesondere ein nach oben sprungfreier Markov-Sprungprozess. Für die Klasse der nach oben sprungfreien Markov-Sprungprozesse kann man zeigen, dass die Matrix $Q_{(n)}$ genau dann zu einer Diagonalmatrix ähnlich ist, wenn die Eigenwerte $\lambda_1, ..., \lambda_n$ alle verschieden sind. Wir werden dieses Resultat zu einem späteren Zeitpunkt für Geburts- und Todesprozesse mit positiven Geburts- und Todesraten zeigen. Wir haben also eine Spektraldarstellung der Form (4.10) genau dann zur Verfügung, wenn die Eigenwerte alle verschieden sind. Der folgende Satz liefert uns für die Klasse der nach oben sprungfreien Markov-Sprungprozesse eine Darstellung der γ_i in (4.10). Für diese Darstellung werden nur die Eigenwerte benötigt. Eine Berechnung der rechten und linken Eigenvektoren entfällt.

Satz 4.2. *Die Übergangsratenmatrix $Q_{(n)}$ zwischen den Zuständen $\{0, ..., n-1\}$ eines Markov-Sprungprozesses aus der Klasse der nach oben sprungfreien Markov-Sprungprozesse sei diagonalisierbar mit paarweise verschiedenen Eigenwerten λ_i. Dann hat die Ersteintrittszeit T_{0n} die folgende Verteilung:*

$$P(T_{0n} > t) = \sum_{i=1}^{n} \prod_{j \neq i} \left(\frac{\lambda_j}{\lambda_j - \lambda_i} \right) e^{-\lambda_i t}. \tag{4.12}$$

Aus (4.12) folgt weiter, dass für reelle $\lambda_1, ..., \lambda_n$ T_{0n} die Faltung exponentialverteilter Zufallsvariablen mit Parametern $\lambda_1, ..., \lambda_n$ ist.

Beweis. Damit der Zustand n erreicht wird, muss wenigstens ein Übergang von i nach $i+1$

für $i = 0, ..., n-1$ stattfinden. Deshalb ist T_{0n} stochastisch größer als die Faltung n exponentialverteilter Zufallsvariablen mit Parametern $b_0, ..., b_{n-1}$, welche wiederum stochastisch größer als M ist, wobei M das Maximum n unabhängiger exponentialverteilter Zufallsvariablen mit Parametern $b_0, ..., b_{n-1}$ ist. Definiere $F_{0,n}^{(t)} = P(T_{0n} \leq t)$. Es gilt:

$$F_{0,n}^{(t)} \leq P(M \leq t) = t^n \left(\prod_{i=0}^{n-1} b_i \right) + o(t^n) \qquad \text{für } t \to 0. \tag{4.13}$$

Seien nun die Eigenwerte $\lambda_1, ..., \lambda_n$ alle verschieden und (4.10) erfüllt. Dann ist $F_{0,n}^{(t)} = 1 - P(T_{0n} > t)$ nach (4.10) analytisch und $o(t^{n-1})$ für $t \to 0$ nach (4.13). Die Betrachtung der Taylorentwicklung von $F_{0,n}^{(t)}$ in 0 liefert

$$\frac{d^k}{dt^k} F_{0,n}^{(t)}|_{t=0} = 0 \text{ für } k = 1, ..., n-1, \tag{4.14}$$

da $F_{0,n}^{(t)} = o(t^{n-1})$ ist. Aus den Gleichungen (4.10) und (4.14) folgt

$$\sum_{i=1}^{n} \gamma_i \lambda_i^k = \begin{cases} 1 & \text{für } k = 0, \\ 0 & \text{für } k = 1, ..., n-1. \end{cases} \tag{4.15}$$

Im Fall $k = 0$ folgt dies aus (4.10) und $\sum_{i=1}^{n} \gamma_i = 1$ und im Fall $k = 1, ..., n-1$ mit der Taylorentwicklung (4.14). Wir definieren $\gamma = (\gamma_1, ..., \gamma_n)$. W sei eine $n \times n$-Matrix mit den Einträgen $W_{ik} = \lambda_i^{k-1}, i, k = 1, 2, ..., n$, und δ sei ein Spaltenvektor der Länge n mit $\delta_1 = 1$ und $\delta_i = 0$ für $i = 2, ..., n$. Wir schreiben (4.15) in der Form

$$\gamma W = \delta^T. \tag{4.16}$$

Nun ist W die wohlbekannte Vandermonde-Matrix (Seite 72 in [Cul67]), welche die Determinante $\prod_{i>j} (\lambda_i - \lambda_j)$ besitzt und damit genau dann invertierbar ist, wenn $\lambda_i \neq \lambda_j$ für alle $i \neq j$ mit $1 \leq i, j \leq n$. (4.16) liefert, dass γ^T mit der ersten Zeile von W^{-1} übereinstimmt, also

$$\gamma_i = (W^{-1})_{1i} = \prod_{j \neq i} \left(\frac{\lambda_j}{\lambda_j - \lambda_i} \right). \tag{4.17}$$

Einsetzen von γ^T in (4.10) liefert (4.12).

Falls die Eigenwerte $\lambda_1, ..., \lambda_n$ alle verschieden sind, dann folgt aus (4.12) bei Betrachten der L.T. von T_{0n}

$$\sigma_{0n}(s) = \sum_{i=1}^{n} \prod_{j \neq i} \left(\frac{\lambda_j}{\lambda_j - \lambda_i} \right) \frac{\lambda_i}{\lambda_i + s} = p(s) \prod_{i=1}^{n} \frac{\lambda_i}{\lambda_i + s}, \tag{4.18}$$

wobei $p(s)$ definiert ist durch

$$p(s) = \sum_{i=1}^{n} \left[\prod_{j \neq i} \left(\frac{\lambda_j + s}{\lambda_j - \lambda_i} \right) \right]. \tag{4.19}$$

Nun ist $p(s) - 1$ nach (4.19) ein Polynom von Grad $n-1$ mit n verschiedenen Nullstellen $-\lambda_1, ..., -\lambda_n$. Das einzige Polynom in einem nullteilerfreien Ring, welches dies erfüllt, ist das Nullpolynom. Also $p(s) - 1 \equiv 0$ und es folgt aus (4.18), dass $\sigma_{0n}(s)$ die folgende Gestalt hat:

$$\sigma_{0n}(s) = \prod_{i=1}^{n} \frac{\lambda_i}{\lambda_i + s}. \tag{4.20}$$

Falls nun wie im Fall eines Geburts- und Todesprozesses die Eigenwerte $\lambda_1, ..., \lambda_n$ alle reell sind, sind die Eigenwerte wegen $P(T_{0n} > t) \to 0$ notwendigerweise positiv. In diesem Fall liefert die Gestalt der L.T. in (4.20), dass T_{0n} eine Summe exponentialverteilter Zufallsvariablen mit Parametern $\lambda_1, ..., \lambda_n$ ist. □

Betrachten wir nun den Fall verschiedener reeller Eigenwerte $\lambda_1, ..., \lambda_r$ mit Vielfachheiten $m_1, ..., m_r$. Wir können eine Folge von Vektoren $\lambda^{(k)}$ konstruieren, welche gegen den Vektor λ konvergiert, wobei jeder einzelne Vektor $\lambda^{(k)}$ n verschiedene Einträge hat. Wir wählen dazu eine Folge von Matrizen $Q_{(n)}^{(k)}$ mit Eigenwerten $\lambda^{(k)}$, welche gegen $Q_{(n)}$ mit Eigenwerten λ konvergiert. Für den Prozess mit Übergangsratenmatrix $Q_{(n)}^{(k)}$ ist die Verteilung der zugehörigen Ersteintrittszeit $T_{0n}^{(k)}$ nach den obigen Überlegungen die Summe n exponentialverteilter Zufallsvariablen und besitzt die L.T.

$$\sigma_k(s) = \prod_{i=1}^{n} \left[\frac{\lambda_i^{(k)}}{\lambda_i^{(k)} + s} \right]. \tag{4.21}$$

Grenzübergang $k \to \infty$ liefert die Konvergenz in Verteilung von $T_{0n}^{(k)}$ gegen T_{0n} und (4.21) konvergiert gegen

$$\sigma(s) = \lim_{k \to \infty} \sigma_k(s) = \prod_{i=1}^{r} \left[\frac{\lambda_i}{\lambda_i + s} \right]^{m_i}. \tag{4.22}$$

Bei der Untersuchung der Ersteintrittszeit T_{0n} bei Geburts- und Todesprozessen hatten wir gezeigt, dass T_{0n} eine Summe n unabhängiger exponentialverteilter Zufallsvariablen S_i ist, also $T = \sum_{i=1}^{n} S_i$, wobei für die Parameter θ_i der S_i gilt, dass $\theta_i \neq \theta_j$ für alle $i \neq j$. Da T_{0n} nach (4.22) nur dann eine Summe n unabhängiger, exponentialverteilter Zufallsvariablen ist, wenn die Eigenwerte alle verschieden sind, liefert uns Satz 4.1.1 insbesondere, dass bei Geburts- und Todesprozessen mit positiven Geburts- und Sterberaten die Eigenwerte $\lambda_1, ..., \lambda_n$

von $Q_{(n)}$ alle verschieden sind und für die Parameter in der Summendarstellung $T_{0n} = \sum_{i=1}^{n} S_i$ gilt $S_i \sim \text{Exp}(\lambda_i)$. Damit ist die angekündigte Verbindung der Parameter zum Spektrum des Geburts- und Todesprozesses hergestellt.

Bemerkung 4.2. Einen ersten stochastischen Beweis für Satz 4.1 liefert Fill [Fil07]. Wir skizzieren seine Überlegungen. Sei M^* ein Geburts- und Todesprozess in stetiger Zeit auf $\{0, ..., d\}$, wobei d absorbierend und alle anderen Geburts- und Todesraten positiv sind. T_{0d}^* sei die Ersteintrittszeit von 0 nach d, also die Zeit bis zur Absorption in d. Zu M^* existiert ein sogenannter "Anti-Dual" M, ein Geburts- und Todesprozess mit stark stationärem dualem Prozess M^*. T_{0d}^* ist eine minimale stark stationäre Zeit für M. Zu M kann man nun mit M^* einen stark stationären dualen Prozess $\widehat{M}$ konstruieren, welcher ein reiner Geburtsprozess ist, wobei die Geburtsraten die nichtnegativen Eigenwerte der Matrix $-Q^*$ sind und Q^* die erzeugende Matrix von M^* ist. Die S_i können somit stochastisch als Wartezeiten für aufeinanderfolgende Geburten in $\widehat{M}$ interpretiert werden. Für Details und den diskreten Fall verweisen wir auf [Fil07], insbesondere Theorem 3.1 und Theorem 4.2.

Zusammenfassend erhalten wir also mit den Ergebnissen aus diesem Kapitel und den Resultaten aus der Dualitätstheorie den folgenden Satz:

Satz 4.3. *(a) Sei M ein irreduzibler, in 0 gestarteter, Geburts- und Todesprozess mit Werten in $S = \{0, ..., d\}$, stetigem Zeitparameter t und stationärer Verteilung ξ^*. Dann gilt für die Separationsfunktion s:*

$$s(t) = s(\gamma^t, \xi^*) = \max_{0 \le x \le d} \left(1 - \frac{\gamma^t(x)}{\xi^*(x)}\right) = \sum_{i=1}^{d} \prod_{j \ne i} \frac{\lambda_j}{\lambda_j - \lambda_i} e^{-t\lambda_i} = P(T > t),$$

wobei $T = \sum_{i=1}^{d} S_i$ und jedes S_i eine exponentialverteilte Zufallsvariable mit Parameter λ_i ist und die S_i´s unabhängig sind. Dabei sind die λ_i´s die positiven Eigenwerte von $-Q$, und Q ist die zugehörige erzeugende Matrix der Form $Q = \boldsymbol{P} - I$. Insbesondere gilt

$$E(T) = \sum_{i=1}^{d} \lambda_i^{-1}, \qquad Var(T) = \sum_{i=1}^{d} \lambda_i^{-2}.$$

(b) Sei M ein irreduzibler, in 0 gestarteter, Geburts- und Todesprozess in diskreter Zeit mit Werten in $S = \{0, ..., d\}$ und stationärer Verteilung ξ^*. Die Monotoniebedingung $p_x + q_{x+1} \le 1$,

$0 \leq x < d$, sei erfüllt. Dann gilt:

$$s(k) = s(\mu^k, \xi^*) = \max_{0 \leq x \leq d} \left(1 - \frac{\mu^k(x)}{\xi^*(x)}\right) = \sum_{i=1}^{d} \prod_{j \neq i} \frac{\lambda_j}{\lambda_j - \lambda_i} (1 - \lambda_i)^k = P(T > k),$$

wobei T eine Zufallsvariable ist mit

$$E(T) = \sum_{i=1}^{d} (1 - \lambda_i)^{-1}, \qquad Var(T) = \sum_{i=1}^{d} \lambda_i (1 - \lambda_i)^{-2}.$$

Die λ_i´s sind dabei die positiven Eigenwerte von $-Q = -(\boldsymbol{P} - I)$, wobei $\boldsymbol{P}$ die Ein-Schritt-Übergangsmatrix von M ist. Die Zufallsvariable T kann als Summe $T = \sum_{i=1}^{d} S_i$ geschrieben werden mit unabhängigen Zufallsvariablen $S_i, 1 \leq i \leq m$. S_i ist geometrisch verteilt mit Erfolgswahrscheinlichkeit λ_i, falls $\lambda_i \in [0, 1]$, und S_i ist eine Bernoulli-Variable mit Parameter (λ_i^{-1}), falls $\lambda_i > 1$.

Beweis. (a) ist im Wesentlichen eine Kombination der Sätze 4.1 und 4.2 mit der Aussage aus der Dualitätstheorie, dass der duale Prozess eines endlichen Geburts- und Todesprozesses ein absorbierender Geburts- und Todesprozess auf dem selben Zustandsraum mit den selben Eigenwerten ist. Es bleibt zu zeigen, dass $Q_{(d)}$ und Q die gleichen positiven Eigenwerte $\lambda_1, ..., \lambda_d$ besitzen. Wir zeigen dies in (b) für Q-Matrizen der Form $Q = \boldsymbol{P} - I$, die von uns bei der Cut-Off Untersuchung ausschließlich betrachtet werden.

(b) Sei $\boldsymbol{P}^*_{(d)}$ die $d \times d$-Matrix, welche man durch Entfernen der letzten Zeile und Spalte der Übergangsmatrix $\boldsymbol{P}^*$ des dualen Geburts- und Todesprozesses M^* von M erhält. Da die Eigenwerte $\theta_1, ..., \theta_d$ von $\boldsymbol{P}^*_{(d)}$ durch die Eigenwerte $\lambda_1, ..., \lambda_d$ von $-Q_{(d)}$ ($Q_{(d)} = \boldsymbol{P}^*_{(d)} - I$ ist die erzeugende Matrix) eindeutig bestimmt sind ($\lambda_i = 1 - \theta_i$), folgt aus der paarweisen Verschiedenheit der λ_i, dass $\theta_i \neq \theta_j$ für $i \neq j$. Ferner gilt für die Ersteintrittszeit T^*_{0d}, dass die wahrscheinlichkeitserzeugende Funktion von $T^*_{0,d}$ die folgende Gestalt hat

$$u \to \prod_{j=1}^{d} \left[\frac{(1 - \theta_j)u}{1 - \theta_j u}\right] = \prod_{j=1}^{d} \left[\frac{\lambda_j u}{1 + \lambda_j u - u}\right]. \tag{4.23}$$

Da d ein absorbierender Zustand für M^* ist, können wir die Eigenwerte von $\boldsymbol{P}^*$ und $\boldsymbol{P}^*_{(d)}$ leicht miteinander in Beziehung setzen. Entwicklung nach der letzten Zeile, welche nur auf der Diagonalen einen Eintrag hat, liefert, dass die Eigenwerte von $\boldsymbol{P}^*$ gerade $1, \theta_1, ..., \theta_d$ sind. Nun gilt wegen des Algebraischen Dualitätstheorems $\Lambda \boldsymbol{P} = \boldsymbol{P}^* \Lambda$. Die Invertierbarkeit von Λ als untere Dreiecksmatrix mit strikt positiven Diagonaleinträgen liefert somit die Ähnlichkeit von $\boldsymbol{P}$ und $\boldsymbol{P}^*$. Folglich haben $\boldsymbol{P}$ und $\boldsymbol{P}^*$ dieselben Eigenwerte. Betrachtet man nun wieder $\boldsymbol{P} - I$, so hat $\boldsymbol{P} - I$ die Eigenwerte $0, \lambda_1, ..., \lambda_d$, welche nach Satz 4.1 zusammen mit Satz 4.2 alle verschieden sind.

Argumentation wie im Beweis von Satz 4.1 liefert für

$$P(T^*_{0,d} > k) = \sum_{i=1}^{d} \gamma_i \theta_i^k$$

die explizite Gestalt der γ_i:

$$P(T^*_{0,d} > k) = \sum_{i=1}^{d} \left(\prod_{j \neq i} \frac{1 - \theta_j}{\theta_i - \theta_j} \right) \theta_i^k = \sum_{i=1}^{d} \left(\prod_{j \neq i} \frac{\lambda_j}{\lambda_i - \lambda_j} \right) (1 - \lambda_i)^k.$$

Damit ist der Beweis vollständig. □

Bemerkung 4.3. Lassen wir den Geburts- und Todesprozess am anderen Ende des Zustandsraums d starten, erhalten wir, dass der Separationsabstand des in 0 bzw. in d gestarteten Geburts- und Todesprozesses von der stationären Verteilung zu jeder Zeit gleich ist, da die Eigenwerte sich nicht verändern. Bei Abstandsmessung in totaler Variation ist dies nicht gültig.

5 Separations-Cut-Off bei Geburts- und Todesprozessen

In diesem Kapitel werden wir das Hauptresultat — eine notwendige und hinreichende Bedingung für das Auftreten des Cut-Off-Effekts bei Geburts- und Todesprozessen — vorstellen und beweisen. Da wir in diesem und den folgenden Kapiteln Familien von Markov-Prozessen betrachten, passen wir unsere bisherige Notation an diese neue Situation an. Mit γ_n^t bezeichnen wir die Verteilung eines Markov-Prozesses M_n zur Zeit t, wobei t ein stetiger Parameter ist. Im diskreten Fall schreiben wir μ_n^k für die Verteilung einer Markov-Kette M_n zur Zeit k.

Vor dem eigentlichen Beweis geben wir nun die Begründung, warum wir uns auf Q-Matrizen der Form $\boldsymbol{P} - I$ für eine stochastische Matrix $\boldsymbol{P}$ und die Identität I beschränken können. Gegeben eine Übergangsmatrix $\boldsymbol{P}$ auf einem abzählbaren Raum, können wir den von $\boldsymbol{P} - I$ erzeugten Markov-Sprungprozess betrachten. Dieser hat bei Start in x zur Zeit t die Verteilung

$$H_t(x, \cdot) = \gamma^t(x, \cdot) = e^{-t} \sum_{k=0}^{\infty} \frac{t^k \boldsymbol{P}^n(x, \cdot)}{k!}.$$

Allgemein definiert man einen Markov-Sprungprozess mit Hilfe einer Matrix Q, für die $\sum_y Q(x, y) = 0$ und $Q(x, y) \geq 0$ für $x \neq y$ gilt. In dieser Allgemeinheit muss $\sum_{x \neq y} Q(x, y)$ nicht gleichmäßig beschränkt sein und es besteht die Gefahr der Explosion des Markov-Sprungprozesses, also des Auftretens unendlich vieler Übergänge in endlicher Zeit. Auf einem endlichen Zustandsraum besteht diese Gefahr jedoch nicht. Wir setzen $q = \max_x\{-Q(x, x)\}$ und betrachten $\boldsymbol{P}(x, y) = I(x, y) + q^{-1}(Q(x, y))$. Für die Verteilung $\gamma_Q^t(x, \cdot)$ des in x gestarteten Markov-Sprungprozesses mit Erzeuger Q gilt

$$\gamma_Q^t(x, \cdot) = \gamma^{qt}(x, \cdot) = H_{qt}(x, \cdot) = e^{-qt} \sum_{k=0}^{\infty} \frac{(qt)^k \boldsymbol{P}^k(x, \cdot)}{k!}. \tag{5.1}$$

Es folgt die Unabhängigkeit des Auftretens eines D-Cut-Offs in einer Familie ergodischer endlicher Markov-Sprungprozesse $(\Omega_n, \nu_n, \gamma^t_{Q_n}(x_n, \cdot))$ von der gewählten Zeit-Skala: (5.1) liefert

$$\gamma^t_{Q_n}(x_n, \cdot) = \gamma_n^{q_n t}(x_n, \cdot) = H_{q_n t}(x_n, \cdot) = e^{-q_n t} \sum_{k=0}^{\infty} \frac{(q_n t)^k \boldsymbol{P}^k(x_n, \cdot)}{k!}.$$

Demnach zeigt $(\Omega_n, \nu_n, \gamma^t_n(x_n, \cdot))$ genau dann einen D-Cut-Off zur Zeit t_n (resp. einen (t_n, b_n)-D-Cut-Off), wenn $(\Omega_n, \nu_n, \gamma^t_{Q_n}(x_n, \cdot))$ einen D-Cut-Off zur Zeit t_n/q_n zeigt (resp. einen $(t_n/q_n, b_n/q_n)$-D-Cut-Off). Dies würde nicht gelten, falls $\lim_{n\to\infty} t_n = \infty$ in der Cut-Off Definition vorausgesetzt worden wäre. Die zeitliche Transformationsmöglichkeit gestattet uns ausschließlich Markov-Sprungprozesse mit Q-Matrizen der Form $Q = \boldsymbol{P} - I$ zu betrachten, da der Nachweis des Cut-Offs bei dem transformierten MSP äquivalent zur Existenz eines Cut-Offs bei dem ursprünglichen MSP ist. Cut-Off Zeitpunkt t_n und Fenstergröße b_n verändern sich entsprechend der Veränderung der Zeitskala durch q_n.

Für den Beweis unseres Hauptresultates benötigen wir noch ein Lemma mit elementaren Folgerungen aus der Chebychev-Ungleichung.

Folgerungen aus der Chebychev-Ungleichung

Sei $\boldsymbol{P}_n(x, y)$ ein Markov-Kern auf einer endlichen Menge Ω_n, welcher die eindeutige stationäre Verteilung ξ^*_n besitzt. Die Ergebnisse aus Kapitel 3 liefern bei Abstandsmessung in Separation die Existenz einer Folge minimaler stark stationärer Zeiten T^D_n, also eine Folge für die

$$D(\mu^k_n, \xi^*_n) = P(T^D_n > k) \text{ bzw. } D(\gamma^t_n, \xi^*_n) = P(T^D_n > t) \tag{5.2}$$

gilt. Bezeichnet D den Totalvariationsabstand, so existiert ebenfalls eine Folge reeller nichtnegativer Zufallsvariablen T^D_n, die stochastisch als minimale Kopplungszeiten interpretiert werden können, welche (5.2) erfüllen. Den Erwartungswert der T^D_n's werden wir mit t_n bezeichnen, die Varianz mit σ^2_n. Die Chebychev-Ungleichung in der Form $P(X > t) \leq \frac{\sigma^2}{\sigma^2+t^2}$ liefert uns für alle $a > 0$

$$P(T^D_n > t_n + a\sigma_n) = P\left(\frac{T^D_n - t_n}{\sigma_n} > a\right) \leq \frac{1}{1+a^2}, \tag{5.3}$$

$$P(T^D_n < t_n - a\sigma_n) \leq \frac{1}{1+a^2}. \tag{5.4}$$

Wir erhalten folgendes Lemma:

Lemma 5.1. *(a) Für alle $\epsilon \in (0,1)$ gilt für die Mischzeit $\tau_n^D(\epsilon)$*

$$t_n - (\epsilon^{-1} - 1)^{-1/2}\sigma_n \leq \tau_n^D(\epsilon) \leq t_n + (\epsilon^{-1} - 1)^{1/2}\sigma_n.$$

(b) Falls eine Konstante $c > 0$ existiert, so dass $ct_n > \sigma_n$, und ein Cut-Off zur Zeit s_n vorliegt mit $\lim_{n\to\infty} s_n = \infty$, dann gilt $s_n \sim t_n$. Im zeitstetigen Fall gilt $s_n \sim t_n$ auch ohne die Voraussetzung $\lim_{n\to\infty} s_n = \infty$.

(c) Falls $\sigma_n^{-1} t_n \longrightarrow \infty$, dann liegt ein (t_n, σ_n)-D-Cut-Off vor.

Beweis. (a) Wir setzen $\epsilon = (1 + a^2)^{-1}$. Daraus folgt $a = (\epsilon^{-1} - 1)^{\frac{1}{2}}$ und damit

$$P(T_n^D > t_n + a\sigma_n) = D(\mu_n^{t_n + a\sigma_n}, \xi_n^*) \leq \epsilon \quad \Rightarrow \tau_n^D(\epsilon) \leq t_n + a\sigma_n.$$

Analog folgt die untere Schranke aus der zweiten Ungleichung in (5.4) durch Lösen von $1 - \epsilon = (1 + a^2)^{-1}$.

(b) Wir behandeln nur den zeitstetigen Fall. Angenommen, es liegt ein Cut-Off zur Zeit s_n vor und seien $\epsilon, \eta \in (0,1)$. Dann gilt nach Definition von Cut-Off und Mischzeit für genügend großes n

$$(1 - \eta)s_n \leq \tau_n^D(\epsilon) \leq (1 + \eta)s_n.$$

Setze $\epsilon = (1 + \eta^2)^{-1}$. Daraus folgt $\eta = \sqrt{\epsilon^{-1} - 1}$. Mit der zweiten Ungleichung in Lemma 5.1(a) erhalten wir

$$\tau_n^D(\epsilon) \leq t_n + \eta\sigma_n$$

und damit

$$(1 - \eta)s_n \leq t_n + \eta\sigma_n \leq (1 + c\eta)t_n,$$

wobei für die zweite Ungleichung die Voraussetzung $ct_n \geq \sigma_n$ gebraucht wird. Nun setze $\epsilon = (1 + \eta^{-2})^{-1}$. Daraus folgt $\eta = (\epsilon^{-1} - 1)^{-\frac{1}{2}}$. Mit der ersten Ungleichung in Lemma 5.1(a) erhalten wir

$$\tau_n^D(\epsilon) \geq t_n - \eta\sigma_n$$

und damit

$$(1 - c\eta)t_n \leq t_n - \eta\sigma_n \leq (1 + \eta)s_n.$$

Für die erste Ungleichung wurde wieder die Voraussetzung $ct_n \geq \sigma_n$ gebraucht. Damit ist $\lim_{n\to\infty} \frac{s_n}{t_n} = 1$ gezeigt und (b) bewiesen.

(c) Nachzuprüfen ist, ob Definition 2.5 erfüllt ist. Aus der Voraussetzung $\sigma_n^{-1}t_n \to \infty$ folgt sofort $\frac{\sigma_n}{t_n} \to 0$ und für f_+, f_- gilt wegen Ungleichung (5.3)

$$\lim_{a\to\infty} f_+(a) = \lim_{a\to\infty} \limsup_{n\to\infty} D(\mu_n^{t_n+a\sigma_n}, \xi_n^*)$$

$$= \lim_{a\to\infty} \limsup_{n\to\infty} P(T_n^D > t_n + a\sigma_n) \leq \lim_{a\to\infty} \frac{1}{1+a^2} = 0.$$

Analog folgt $\lim_{a\to\infty} f_-(a) = 1$ und damit die Existenz eines (t_n, σ_n)-D-Cut-Offs. □

Kommen wir nun zur Formulierung des Hauptergebnisses. Für $n = 1, 2, ...$ sei $\Omega_n = \{0, 1, ..., m_n\}$ ausgestattet mit einem irreduziblen Geburts- und Todesprozess mit stationärer Verteilung ξ_n^*. Seien $q_{n,x}$, $r_{n,x}$ und $p_{n,x}$ die zugehörigen Übergangswahrscheinlichkeiten. Sei μ_n^k die Verteilung des zugehörigen, in 0 gestarteten, Geburts- und Todesprozesses, mit diskretem Zeitparameter k. Sei γ_n^t die Verteilung des zugehörigen, in 0 gestarteten, Geburts- und Todesprozesses mit stetigem Zeitparameter t. Seien $\lambda_{n,i} \in [0,2], 0 \leq i \leq m_n$, die zugehörigen Eigenwerte. Setze

$$\lambda_n = \lambda_{n,1}, \qquad t_n = \sum_{i=1}^{m_n} \lambda_{n,i}^{-1}.$$

Den kleinsten positiven Eigenwert $\lambda_n = \lambda_{n,1}$ bezeichnet man als *Spektrallücke*. Für jedes $\epsilon \in (0,1)$ ist die *ϵ-Separations-Mischzeit* folgendermaßen definiert:

$$\tau_n(\epsilon) = \inf\{t \geq 0 : \mathrm{s}(\gamma_n^t, \xi_n^*) \leq \epsilon\}.$$

Theorem 5.1. ***(Separations-Cut-Off bei Geburts- und Todesprozessen)***

In obiger Situation hat die Familie

$$(\Omega_n, \xi_n^*, (\gamma_n^t)_{t>0})_{n=1,2,...}$$

einen Separations-Cut-Off genau dann, wenn $N_n = \lambda_n t_n \overset{n}{\to} \infty$. Für jedes c>0 gelten die Separationsschranken

$$s(\gamma_n^{(1+c)t_n}, \xi_n^*) \leq \frac{1}{1+c^2 N_n}, \tag{5.5}$$

$$s(\gamma_n^{(1-c)t_n}, \xi_n^*) \geq 1 - \frac{1}{1+c^2 N_n}. \tag{5.6}$$

Für jedes feste $\epsilon \in (0,1)$ ist die Bedingung $\lambda_n t_n \to \infty$ äquivalent zu $\lambda_n \tau_n(\epsilon) \to \infty$.

Beweis. Mit Satz 4.3 erhalten wir

$$\mathrm{s}(\gamma_n^t, \xi_n^*) = P(T_n > t),$$

wobei

$$T_n = \sum_{i=1}^{m_n} S_{n,i} \quad \text{und} \quad S_{n,i} \sim \mathrm{Exp}(\lambda_{n,i}).$$

Insbesondere gilt $E(T_n) = t_n$.

Für die Varianz $\mathrm{Var}(T_n) = \sigma_n^2$ gilt

$$\begin{aligned}\sigma_n^2 &= \sum_{i=1}^{m_n} \lambda_{n,i}^{-2} = \sum_{i=1}^{m_n} \lambda_{n,i}^{-2} \left(\frac{\lambda_n}{\lambda_n}\right)^2 = \lambda_n^{-2} \sum_{i=1}^{m_n} \left(\frac{\lambda_n}{\lambda_{n,i}}\right)^2 \\ &\le \lambda_n^{-2} \sum_{i=1}^{m_n} \frac{\lambda_n}{\lambda_{n,i}} = \lambda_n^{-2} \lambda_n \sum_{i=1}^{m_n} \frac{1}{\lambda_{n,i}} = \lambda_n^{-1} \sum_{i=1}^{m_n} \frac{1}{\lambda_{n,i}} = \lambda_n^{-1} t_n.\end{aligned}$$

Für die Ungleichung haben wir genutzt, dass $\lambda_{n,i} \ge \lambda_n$ für alle $i \ge 1$. Wegen

$$\sigma_n = \sqrt{\sum_{i=1}^{m_n} \lambda_{n,i}^{-2}} \ge \lambda_n^{-1}$$

gilt also

$$\sigma_n \le t_n \quad \text{und} \quad \sigma_n \le \sqrt{\lambda_n^{-1} t_n} = N_n^{-\frac{1}{2}} t_n. \tag{5.7}$$

Die Separationsschranken (5.5) und (5.6) folgen direkt aus (5.7), Lemma 5.1 und Satz 4.3

$$\mathrm{s}(\gamma_n^{(1+c)t_n}, \xi_n^*) = P(T_n > t_n + ct_n) \le P(T_n > t_n + c\sigma_n N_n^{\frac{1}{2}}) \le \frac{1}{1 + c^2 N_n}.$$

Entsprechend:

$$\begin{aligned}\mathrm{s}(\gamma_n^{(1-c)t_n}, \xi_n^*) &= P(T_n > t_n - ct_n) = 1 - P(T_n < t_n - ct_n) \\ &\ge 1 - P(T_n < t_n - c\sigma_n N_n^{\frac{1}{2}}) \ge 1 - \frac{1}{1 + c^2 N_n}.\end{aligned}$$

Angenommen, dass $N_n = \lambda_n t_n \xrightarrow{n} \infty$. Aus der zweiten Ungleichung in (5.7) folgt, dass $\frac{t_n}{\sigma_n} \xrightarrow{n} \infty$, da $N_n^{\frac{1}{2}} \le \frac{t_n}{\sigma_n}$.

Nach Teil(c) von Lemma 5.1 folgt also, dass ein Separations-Cut-Off zur Zeit s_n stattfindet und sogar, dass ein (t_n, σ_n)-Cut-Off stattfindet, wir also mit der Folge σ_n sogar die Fenstergröße kennen.

Falls umgekehrt zur Zeit s_n ein Cut-Off stattfindet, dann ist nach (5.7) ($\sigma_n \leq t_n$) und Lemma 5.1(b) ($\sigma_n \leq 1t_n$ und Cut-Off zur Zeit s_n) $s_n \sim t_n$ (d.h.$\frac{s_n}{t_n} = 1$), es muss also ein Cut-Off zur Zeit t_n vorliegen. Da die Separationsschranken (5.5) und (5.6) für jedes $c > 0$ gelten, folgt $N_n \xrightarrow{n} \infty$.

Sei nun $\epsilon \in (0, 1)$ fest. Mit der oberen Grenze in Lemma 5.1(a) und der Beziehung $\sigma_n \leq t_n$ ergibt sich

$$\tau_n(\epsilon) \leq t_n + (\epsilon^{-1} - 1)^{\frac{1}{2}}\sigma_n \leq (1 + (\epsilon^{-1} - 1)^{\frac{1}{2}}t_n.$$

Also gilt

$$(\lambda_n \tau_n(\epsilon) \xrightarrow{n} \infty) \Rightarrow (\lambda_n t_n \xrightarrow{n} \infty).$$

Umgekehrt, falls $t_n\lambda_n \xrightarrow{n} \infty$, liegt ein Cut-Off zur Zeit t_n vor (Begründung siehe weiter oben in diesem Beweis) und nach (5.7) und Lemma 5.1(a) gilt

$$(1 - (\epsilon^{-1} - 1)^{-\frac{1}{2}}t_n) \leq \tau_n(\epsilon) \leq (1 + (\epsilon^{-1} - 1)^{\frac{1}{2}}t_n).$$

Es folgt wegen $\epsilon \in (0, 1)$ beliebig: $\lambda_n \tau_n(\epsilon) \xrightarrow{n} \infty$. Damit ist Theorem 5.1 gezeigt. $\square$

Bemerkung 5.1. (a) Theorem 5.1 zeigt, dass bei in 0 gestarteten Geburts- und Todesprozessen mit stetigem Zeitparameter ein Separations-Cut-Off nur dann auftreten kann, wenn m_n gegen unendlich geht, da sonst die Summe $N_n = \lambda_n t_n$ nicht gegen unendlich gehen kann.

(b) Die Separationsschranken in Theorem 5.1 liefern, dass kein Separations-Pre-Cut-Off im Sinne von Definition 2.6 vorliegen kann, falls $\lambda_n t_n$ beschränkt ist. Also folgt aus der Existenz eines Separations-Pre-Cut-Offs die Unbeschränktheit von $\lambda_n t_n$ und damit wegen Theorem 5.1 die Existenz eines Separations-Cut-Offs. Da aus der Existenz eines Separations-Cut-Offs immer die Existenz eines Separations-Pre-Cut-Offs folgt (setze $c = 1 - \epsilon$ bzw. $C = 1 + \epsilon$), ist bei zeitstetigen Geburts- und Todesprozessen die Existenz eines Separations-Pre-Cut-Offs äquivalent zur Existenz eines Separations-Cut-Offs.

(c) Folgendermaßen können wir in Theorem 5.1 die Separationsschranken umschreiben, so dass wir sie in Abhängigkeit von Spektrallücke und Mischzeit erhalten:

$$\widetilde{N}_n = \lambda_n \tau_n(\epsilon) \leq \lambda_n(t_n + (\epsilon^{-1} - 1)^{1/2}\sigma_n)$$

$$= \lambda_n t_n + (\epsilon^{-1} - 1)^{1/2} \lambda_n \sigma_n \leq \lambda_n t_n + (\epsilon^{-1} - 1)^{1/2} \lambda_n t_n = (1 + (\epsilon^{-1} - 1)^{1/2}) N_n.$$

Daher gilt für alle $c > 0$ und $\epsilon \in (0,1)$

$$\mathrm{s}(\gamma^{(1+c)t_n}, \xi_n^*) \leq \frac{1}{1 + c^2(1 + (\epsilon^{-1} - 1)^{1/2})^{-1} \widetilde{N}_n}$$

und

$$\mathrm{s}(\gamma^{(1-c)t_n}, \xi_n^*) \geq 1 - \frac{1}{1 + c^2(1 + (\epsilon^{-1} - 1)^{1/2})^{-1} \widetilde{N}_n}.$$

Kommen wir nun zur diskreten Version des Hauptsatzes. Wir benötigen zusätzlich die Monotoniebedingung $p_{n,x} + q_{n,x+1} \leq 1$.

Theorem 5.2. *Bezüglich der oben eingeführten Situation und Notation und unter der Annahme, dass für alle $x \in \{0, ..., m_n - 1\}$ die Monotoniebedingung*

$$p_{n,x} + q_{n,x+1} \leq 1$$

erfüllt ist, hat die Familie

$$(\Omega_n, \xi_n^*, (\mu_n^k)_{k=0,1,\ldots})_{n=1,2,\ldots}$$

genau dann einen Separations-Cut-Off, wenn $N_n = \lambda_n t_n \xrightarrow{n} \infty$.

Die Bedingung $p_{n,x} + q_{n,x+1} \leq 1$ impliziert die Aperiodizität der Markov-Kette mit analoger Argumentation wie in Kapitel 3:

$$r_{0,0} = 0 \quad \Rightarrow \quad p_{0,1} = 1 \quad \Rightarrow \quad q_{1,0} = 0 \quad \Rightarrow \quad \lightning,$$

da der Geburts- und Todesprozess als irreduzibel vorausgesetzt ist. Also gilt $r_{0,0} > 0$ und damit ist M_n aperiodisch. Insbesondere konvergiert M_n nach dem Ergodensatz gegen eine eindeutige stationäre Verteilung ξ_n^* . Wie im stetigen Fall kann man das Theorem ebenfalls mit $\tau_n(\epsilon)$ anstelle von t_n formulieren.

Nachdem wir das angekündigte Resultat, eine notwendige und hinreichende Bedingung für die Existenz eines Separations-Cut-Offs bei Geburts- und Todesprozessen bewiesen haben, kommen wir nun dazu die Form des Cut-Offs zu beschreiben.

6 Die Form des Cut-Offs

Ist ein Cut-Off nachgewiesen — sagen wir zur Zeit t_n — stellt sich natürlicherweise als Nächstes die Frage nach der Schärfe des Cut-Offs, also nach der Abruptheit des Übergangs zur stationären Verteilung. Wenn ein (s_n, b_n)-Cut-Off vorliegt (nach möglicher Anpassung der Folge (s_n)), untersuchen wir dazu das Cut-Off Fenster, also die Folge (b_n). Die Form des Cut-Offs werden wir in Satz 6.5 näher beschreiben, in dem wir die Funktionen f_+ und f_- aus der Cut-Off Definition bestimmen. Zunächst stellen wir einige Aussagen aus der Theorie unendlich teilbarer Verteilungen vor, die wir beim Beweis von Satz 6.5 benötigen werden. Wir werden den Begriff der unendlichen Teilbarkeit erklären. Die Gamma-Verteilung und damit insbesondere die Exponentialverteilung wird als unendlich teilbare Verteilung identifiziert. Danach wird mit der Lévy-Khintschin Formel eine erschöpfende Beschreibung der unendlich teilbaren Verteilungen auf $\mathbb{R}$ gegeben. Unsere Darstellung des Exkurses basiert auf Kapitel 16 aus [Kle08] sowie auf Kapitel 48 aus [Als05c].

6.1 Exkurs: Unendliche Teilbarkeit

Definition 6.1. Ein W-Maß $\mu \in \mathfrak{W}(\mathbb{R})$ heißt *unendlich teilbar*, falls es für jedes $n \in \mathbb{N}$ ein $\mu_n \in \mathfrak{W}(\mathbb{R})$ mit der Eigenschaft $\mu_n^{*n} = \mu$ gibt.
Analog nennen wir eine charakteristische Funktion ϕ eines W-Maßes auf $\mathbb{R}$ (kurz: CFW) *unendlich teilbar*, falls es zu jedem $n \in \mathbb{N}$ eine CFW ϕ_n gibt mit $\phi_n^n = \phi$.
Eine reelle Zufallsvariable X heißt *unendlich teilbar*, falls es zu jedem $n \in \mathbb{N}$ unabhängig, identisch verteilte Zufallsvariablen $X_{n,1}, ..., X_{n,n}$ gibt mit $X \stackrel{d}{=} X_{n,1} + ... + X_{n,n}$.

Offenbar sind alle drei Begriffe der unendlichen Teilbarkeit äquivalent. Wir verzichten im Folgenden auf eine ausführliche Vorstellung der Theorie unendlich teilbarer Verteilungen und konzentrieren uns auf die Aussagen, die für den Beweis von Satz 6.5 benötigt werden.

Lemma 6.1. *Die Exponentialverteilung mit Parameter λ ist unendlich teilbar, die dazugehörigen Teiler sind gammaverteilt mit Parametern λ und $1/n$.*

Satz 6.1. *Ist $(\mu_n)_{n\in\mathbb{N}}$ eine (schwach) konvergente Folge unendlich teilbarer Wahrscheinlichkeitsmaße auf $\mathbb{R}$, so ist $\mu = \lim_{n\to\infty}\mu_n$ unendlich teilbar.*

Definition 6.2. Ein σ-endliches Maß ν auf $\mathbb{R}$ mit $\nu(\{0\}) = 0$ und $\int (x^2 \wedge 1)\nu(dx) < \infty$ heißt *kanonisches Maß*. Sind $\sigma^2 \geq 0$ und $b \in \mathbb{R}$, so heißt (σ^2, b, ν) *kanonisches Tripel*.

Satz 6.2. ***(Lévy-Khintschin Formel).*** *Sei $\mu \in \mathfrak{W}(\mathbb{R})$ und $\psi(t) = \log \int e^{itx}\mu(dx)$. μ ist genau dann unendlich teilbar, wenn es ein kanonisches Tripel (σ^2, b, ν) gibt, so dass*

$$\psi(t) = \frac{\sigma^2}{2}t^2 + ibt + \int (e^{itx} - 1 - itxI_{\{|x|<1\}})\nu(dx) \tag{6.1}$$

gilt. Dieses Tripel ist durch (6.1) eindeutig festgelegt. ν bezeichnet man als Lévy-Maß, σ^2 als Gaußschen Koeffizienten und b als Zentrierungskonstante.

Bemerkung 6.1. Die Lévy-Khintschin Formel sagt also insbesondere aus, dass eine unendlich teilbare Verteilung einen normalverteilten Anteil hat. Wenn der normalverteilte Anteil ungleich 0 ist, ist also die Dichtefunktion der unendlich teilbaren Verteilung überall positiv, weil die Normalverteilung diese Eigenschaft besitzt. Dieses Argument werden wir im Beweis von Satz 6.5(b) benötigen.

6.2 Form des Separations-Cut-Offs bei Geburts- und Todesprozessen

Im Beweis von Satz 6.5 werden wir auf den Satz von Prohorov verweisen, der von uns jetzt in der Version [Als05c] vorgestellt wird.

Definition 6.3. Eine Familie $(Q_i)_{i\in I}$ endlicher Maße auf $(\mathbb{R}^d, \mathfrak{B}^d)$ heißt straff, wenn für jedes $\epsilon > 0$ ein Kompaktum $K \subset \mathbb{R}^d$ existiert mit

$$\sup_{i\in I} Q_i(K^c) < \epsilon.$$

Satz 6.3. ***(Satz von Prohorov)*** *Eine Familie* $\mathfrak{M}$ *endlicher Maße auf einem polnischen Raum* $(S, \mathfrak{S})$ *ist genau dann schwach relativ folgenkompakt (d.h. jede Folge in* $\mathfrak{M}$ *enthält eine schwach konvergente Teilfolge), wenn sie gleichmäßig beschränkt ist* ($\sup_{Q\in\mathfrak{M}} Q(S) < \infty$) *und straff ist.*

Im Beweis genügt uns der folgende Spezialfall:

Satz 6.4. *Eine Familie* $\mathfrak{W} \subset \mathfrak{W}(\mathbb{R})$ *ist genau dann schwach relativ folgenkompakt, wenn sie straff ist.*

Nun zu dem angekündigten Satz, der uns die Gestalt der Folge (b_n) im Falle von Geburts- und Todesprozessen und Abstandsmessung in Separation näher beschreibt.

Satz 6.5. *Gegeben eine Familie von Geburts- und Todesprozessen wie im vorherigen Kapitel gelte* $N_n = \lambda_n t_n \to \infty$*, es liege also ein Separations-Cut-Off vor. Setze*

$$\sigma_n^2 = \sum_{i=1}^{m_n} \lambda_{n,i}^{-2}.$$

(a) Gelte $\lambda_n \sigma_n \to \infty$*. Dann gilt für reelles* c

$$\lim_{n\to\infty} s(\gamma_n^{t_n+c\sigma_n}, \xi_n^*) = 1 - \Phi(c), \qquad \textit{wobei } \Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-x^2/2} dx.$$

Insbesondere liegt ein (t_n, σ_n)*-Cut-Off vor, aber kein* (t_n, λ_n^{-1})*-Cut-Off.*
(b) Angenommen, dass $\lambda_n \sigma_n$ *beschränkt ist. Dann liegt ein* (t_n, σ_n)*-Cut-Off vor (äquivalent ein* (t_n, λ_n^{-1}) *-Cut-Off) und für jedes reelle* $c > 0$ *gilt*

$$\liminf_{n\to\infty} s(\gamma_n^{t_n+c\sigma_n}, \xi_n^*) > 0,$$

während für jedes reelle $c < 0$ *gilt*

$$\limsup_{n\to\infty} s(\gamma_n^{t_n+c\sigma_n}, \xi_n^*) < 1.$$

Beweis. Nach Satz 4.3 gilt $s(\gamma_n^t, \xi_n^*) = P(T_{0n} > t)$. Die momenterzeugende Funktion der Zufallsvariablen $\frac{T_{0n}-t_n}{\sigma_n}$ hat folgende Gestalt

$$M_n(t) = E(e^{t(T_{0n}-t_n)/\sigma_n}).$$

Da T_{0n} die Summe m_n unabhängiger, exponentialverteilter Zufallsvariablen mit Parametern $\lambda_{n,i}$ ist, erhalten wir

$$M_n(t) = E\left(e^{\frac{tT_{0n}}{\sigma_n} - \frac{tt_n}{\sigma_n}}\right) = e^{-\frac{tt_n}{\sigma_n}} \prod_{i=1}^{m_n} \frac{\lambda_{n,i}}{\lambda_{n,i} - t/\sigma_n} = e^{F_n(t)},$$

wobei

$$F_n(t) = \log\left(e^{-tt_n/\sigma_n} \prod_{i=1}^{m_n} \frac{\lambda_{n,i}}{\lambda_{n,i} - t/\sigma_n}\right) = -tt_n/\sigma_n + \sum_{i=1}^{m_n} \log\left(\frac{1}{1 - t/(\sigma_n \lambda_{n,i})}\right)$$

$$= -tt_n/\sigma_n - \sum_{i=1}^{m_n} \log(1 - t\sigma_n^{-1}\lambda_{n,i}^{-1}) = \sum_{k=2}^{\infty} \frac{\theta_k(n)}{k\theta_2(n)^{k/2}} t^k$$

mit

$$\theta_k(n) = \sum_{i=1}^{m_n} (\lambda_{n,1}/\lambda_{n,i})^k.$$

Da die Eigenwerte $\lambda_{n,i}$ monoton nicht fallend geordnet sind und im Intervall [0,2] liegen, gilt $1 \leq \theta_k(n) \leq \theta_2(n), k \geq 2$. Damit konvergiert die obige Reihe, zumindest für $t \in (0,1)$, und da $F_n(t) = t^2/2 + \sum_{k=3}^{\infty} \frac{\theta_k(n)}{k\theta_2(n)^{k/2}} t^k$ gilt, folgt

$$0 \leq F_n(t) - \frac{t^2}{2} = \sum_{k=3}^{\infty} \frac{t^k}{k\theta_2(n)^{(k-2)/2}} \frac{\theta_k(n)}{\theta_2(n)} \leq \sum_{k=3}^{\infty} \frac{t^k}{k\theta_2(n)^{(k-2)/2}}.$$

Es gilt ferner $\theta_2(n) = \sum_{i=1}^{m_n} \left(\frac{\lambda_{n,1}}{\lambda_{n,i}}\right)^2 = \lambda_n^2 \sum_{i=1}^{m_n} \left(\frac{1}{\lambda_{n,i}}\right)^2 \Rightarrow \theta_2^{1/2} = \lambda_n \sigma_n$.

Falls also $\lambda_n \sigma_n \to \infty$ gilt, dann folgt $M_n(t) \to e^{t^2/2}$ für jedes reelle t und damit ist $\sigma_n^{-1}(T_{0n} - t_n)$ asymptotisch verteilt wie eine standardnormalverteilte Zufallsvariable. Wegen $\mathrm{s}(\gamma_n^t, \xi_n^*) = 1 - P(T_n \leq t)$ ist (a) bewiesen.

Nun zum Beweis von (b). Angenommen $\lambda_n \sigma_n$ ist beschränkt, also $\lambda_n \sigma_n \leq A$. Dann gilt

$$\lambda_n^{-1} \leq \sigma_n \leq A\lambda_n^{-1}$$

und, für jedes k=2,3,...,

$$1 \leq \theta_k(n) \leq A < \infty.$$

Die Chebychev-Ungleichung liefert $P(|\frac{T_{0n} - t_n}{\sigma_n}| \geq \epsilon) \leq \epsilon^{-2}$ und damit die Straffheit der Verteilungen von $\sigma_n^{-1}(T_{0n} - t_n), n = 1, 2, ...$ Nach dem Satz von Prohorov können wir aus einer beliebigen, gegebenen Teilfolge (n_j) eine Teilfolge (n_{j_l}) wählen, so dass entlang dieser Teilfolge

$P(T_{0n} > t_n + c\sigma_n)$ gegen $P(T > c)$, $c \in \mathbb{R}$, konvergiert, für eine Zufallsvariable T. Nun folgt aus der vorherigen Berechnung der momenterzeugenden Funktion, dass entlang der Teilfolge (n_{j_l}) von (n_j) der Grenzwert

$$\lim_{l\to\infty} \theta_k(n_{j_l}) = \theta_k \in [1, A]$$

für jedes $k \geq 2$ existiert und T die momenterzeugende Funktion

$$\exp\left(\frac{t^2}{2} + \sum_{k\geq 3} \frac{\theta_k t^k}{k\theta_2^{k/2}}\right) \tag{6.2}$$

hat für jedes $t \in (-\theta_2^{1/2}, \theta_2^{1/2})$. Die Zufallsvariablen $\sigma_n^{-1}(T_{0n} - t_n)$ sind unendlich teilbar, da T_{0n} als Summe exponentialverteilter Zufallsvariablen unendlich teilbar ist. Nach Satz 6.1 ist also ebenfalls T unendlich teilbar. Mit Blick auf (6.2) sieht man, dass der normalverteilte Anteil von T nichttrivial ist und T also mit Bemerkung 6.1 überall positive Dichte besitzt. Damit ist Teil (b) des Satzes bewiesen. Da die Summe in (6.2) wegen der Beschränktheit von θ_2 nicht verschwindet, kann kein Grenzwert der Folge $(T_{0n} - t_n)/\sigma_n$ normalverteilt sein. Damit ist Satz 6.5 bewiesen. □

Bemerkung 6.2. Wegen des Auftretens der Normalverteilung sprechen wir in (a) auch von einem Gaussschen-Cut-Off. In (b) ist die Form des Cut-Offs nach der Schlussbemerkung im Beweis nicht gausssch. Wir haben also genau dann einen Gaussschen-Cut-Off, wenn die Größenordnung der Fenstergröße strikt größer ist als die der Inversen der Spektrallücke. Die Inverse der Spektrallücke bezeichnet man auch als *relaxation-time*.

Bemerkung 6.3. Der erste Teil von (b), also der Fakt, dass für jedes reelle c>0

$$\liminf_{n\to\infty} \mathrm{s}(\gamma_n^{t_n+c\sigma_n}, \xi_n^*) > 0$$

gilt, kann auch elementar hergeleitet werden. Um $P(T_{0n} > t_n + c\sigma_n)$ von unten zu beschränken, schreibe

$$\begin{aligned} &P(T_{0n} > t_n + c\sigma_n) \\ &\geq P\left(S_{n,1} > \lambda_n^{-1} + (c+1)\sigma_n; \sum_{i=2}^{m_n} S_{n,i} > \sum_{i=2}^{m_n} \lambda_{n,i}^{-1} - \sigma_n\right) \\ &\geq P(S_{n,1} > \lambda_n^{-1} + (c+1)\sigma_n) P\left(\sum_{i=2}^{m_n} S_{n,i} > \sum_{i=2}^{m_n} \lambda_{n,i}^{-1} - \sigma_n\right). \end{aligned}$$

Mit der zweiten Ungleichung in (5.4) in der Form $P(T_{0n} > t_n - \sigma_n) \geq 1 - \frac{1}{1+1}$ und $\sigma_n^2 \geq Var(\sum_{i=2}^{m_n} S_{n,i})$, erhalten wir

$$P\left(\sum_{i=2}^{m_n} S_{n,i} > \sum_{i=2}^{m_n} \lambda_{n,i}^{-1} - \sigma_n\right) \geq \frac{1}{2}.$$

Es folgt

$$P(T_{0n} > t_n + c\sigma_n) \geq \frac{1}{2}e^{-(\lambda_n^{-1}+(c+1)\sigma_n)}\lambda_n = \frac{1}{2}e^{-(1+(c+1)\lambda_n\sigma_n)} \geq \frac{1}{2}e^{-(1+(c+1)A)}.$$

Für den zweiten Teil von Teil (b) ist kein elementarer Beweis bekannt.

Bemerkung 6.4. Angenommen, dass in Satz 6.5(b) zusätzlich für jedes k

$$\theta_k = \lim_{n\to\infty} \left(\frac{\lambda_n}{\lambda_{n,i}}\right)^k < \infty$$

existiert. Dann gilt für jedes $c \in \mathbb{R}$

$$\lim_{n\to\infty} \mathrm{s}(\gamma_n^{t_n+c\sigma_n}, \xi_n^*) = 1 - F(c),$$

wobei $F(t)$ die Verteilungsfunktion einer unendlich teilbaren Zufallsvariable ist, deren momenterzeugende Funktion für $t \in (-\theta_2^{-1/2}, \theta_2^{-1/2})$ gegeben ist durch (6.2). Insbesondere gilt nach Satz 6.5(b), dass $0 < F(c) < 1$ für alle $c \in \mathbb{R}$.

7 Ausblick

Nachdem wir in dieser Arbeit gezeigt haben, dass mit dem Kriterium von Peres eine notwendige und hinreichende Bedingung für das Auftreten des Cut-Off-Effekts bei Geburts- und Todesprozessen und Abstandsmessung in Separation formuliert werden kann, stellt sich als Nächstes natürlicherweise die Frage, ob dieses Kriterium auch für andere Klassen reversibler Markov-Prozesse bei Abstandsmessung in Separation oder bei Geburts- und Todesprozessen und Abstandsmessung in Totalvariation notwendig und hinreichend ist. Zunächst zur zweiten Frage.

7.1 Totalvariations-Cut-Off bei Geburts- und Todesprozessen

Ding, Lubetzky und Peres bewiesen 2008 die Gültigkeit des Peres-Kriteriums in der Klasse der Geburts- und Todesprozesse bei Abstandsmessung in "worst-case"-Totalvariation [DLP08], wobei "worst-case" in diesem Zusammenhang bedeutet, dass man den Startpunkt des Geburts- und Todesprozesses so wählt, dass der Abstand der Verteilung des Geburts- und Todesprozesses zur stationären Verteilung maximal ist. Eine Folge von Geburts- und Todesprozessen weist genau dann einen Cut-Off in "worst-case"-Totalvariation auf, wenn das Produkt aus Spektrallücke und Mischzeit gegen unendlich geht (Korollar 3 in [DLP08]). Im Fall diskreter Zeit betrachten die Autoren die Klasse der "lazy chains", also die Klasse der Geburts- und Todesprozesse, für die die Wahrscheinlichkeit, in einem Zustand zu verweilen, größer als $1/2$ ist (Korollar 2 in [DLP08]). Diese sind insbesondere monoton. Des Weiteren wird in diesem Paper gezeigt, dass für Geburts- und Todesprozesse in stetiger Zeit und Geburts- und Todesprozesse in diskreter

Zeit, welche “lazy” sind, der Separationsabstand maximal ist, wenn der Geburts- und Todesprozess in einem der beiden Endpunkte des Zustandsraums gestartet wird (Theorem 4 in [DLP08]). Da die Separationsmischzeit und die Totalvariationsmischzeit dieselbe Ordnung haben, folgt die Äquivalenz des Cut-Offs bei Geburts- und Todesprozessen und Separation bzw. Totalvariation (Korollar 5 in [DLP08]):

Für eine Familie endlicher irreduzibler Geburts- und Todesprozesse in stetiger Zeit ist Cut-Off in “worst-case”-Totalvariation äquivalent zu Cut-Off in “worst-case”-Separation.

Bei Separation startet man also immer in 0 oder in dem anderen Endpunkt des Zustandsraums. Die Äquivalenz ist dabei in dem Sinne zu verstehen, dass die Existenz des einen Cut-Offs die Existenz des anderen Cut-Offs impliziert; die Cut-Off Zeitpunkte können sich jedoch unterscheiden, was sie beispielsweise beim Bernoulli-Laplace-Modell (siehe Abschnitt 7 in [DSC06]) auch tun.

Die Vermutung liegt nahe, dass das Peres-Kriterium auch für Verwandte von Geburts- und Todesprozessen gültig bleibt. Diese Vermutung ist aber im Allgemeinen nicht richtig. Wir stellen dazu nun ein auf David Aldous zurückgehendes Beispiel in der Version aus [CSC08] vor. In diesem betrachten wir eine Familie reversibler MSP, welche eng mit den Geburts- und Todesprozessen verwandt ist, bei der aber das Peres-Kriterium nicht hinreichend ist.

7.2 Beispiel von Aldous

Die von uns betrachtete Kette besteht aus drei Teilen, einem Tail und zwei Armen. Die beiden Arme sind an dem Tail befestigt und an ihrem anderen Ende miteinander verbunden. Der Tail hat die Länge n, ist also ein Abschnitt der Form $\{x_1, x_2, ..., x_n\}$. Der linke Arm hat ebenfalls die Länge n, $\{y_1, ..., y_n\}$. Der rechte Arm hat die doppelte Länge $2n$, $\{z_1, ..., z_{2n}\}$ mit $z_{2n} = y_n$. Übergänge sind wie bei einem Geburts- und Todesprozess nur in einen benachbarten Zustand möglich, falls eine Zustandsveränderung stattfindet, und wir haben in allen drei Teilen der Markov-Kette eine Drift nach oben vorliegen, also innerhalb des linken Armes in Richtung y_n, entlang des rechten Armes in Richtung z_{2n} und entlang des Tails in Richtung der Arme.
Präziser: $p_{i,n} > 1/2, q_{i,n} < 1/2, p_{i,n} + q_{i,n} = 1, i = t, l, r$, wobei t für Tail, l für links und r für

rechts steht. Entlang des Tails gehen wir demnach mit Wahrscheinlichkeit $p_{t,n}$ nach oben und mit $q_{t,n}$ nach unten. Am oberen Ende des Tails gehen wir mit Wahrscheinlichkeit $(q_{l,n}+q_{r,n})/2$ nach unten, mit $p_{l,n}/2$ nach links und mit $p_{r,n}/2$ nach rechts. Entlang der Arme bewegen wir uns hoch bzw. runter mit Wahrscheinlichkeiten $p_{i,n}, q_{i,n}, i = l$ oder r. Im Punkt $y_n = z_{2n}$, gehen wir mit Wahrscheinlichkeit $q_{l,n}$ nach y_{n-1}, mit $q_{r,n}$ nach z_{2n-1} oder bleiben mit $1-(q_{r,n}+q_{l,n})$. Wir machen diese Kette reversibel durch geeignete Wahl der Wahrscheinlichkeiten entlang des linken und rechten Armes, in diesem Fall durch $p_{l,n}/q_{l,n} = (p_{r,n}/q_{r,n})^2$.
Die Berechnung der stationären Verteilung liefert, dass diese im Wesentlichen im Übergangspunkt zwischen den beiden Armen $y_n = z_{2n}$ konzentriert ist, da wir entlang beider Arme eine Drift in Richtung dieses Punktes haben und entlang des Tails einen Drift in Richtung der Arme. Unter der Annahme, dass $p_{i,n} > 2/3, i = t, l, r$, kann man nun mit der Cheeger-Ungleichung zeigen, dass die Spektrallücke dieser Kette größer einer echt positiven Konstanten ist. Siehe dazu Prop.4.4 in [Che06].
Wir behaupten nun, dass $1 < a < b < \infty$ und $\epsilon \in (0,1)$ existieren, so dass für die in x_1 gestartete Kette der Abstand in Totalvariation geringer als $1-\epsilon$ zur Zeit an ist, aber größer als ϵ zur Zeit bn, für alle großen n. Außerdem gilt, wie man sich leicht überlegt, dass der Startpunkt x_1 derjenige Startpunkt ist, für den die Mischung der Kette am längsten dauert, also die Mischzeit am größten ist. Das Produkt aus Spektrallücke und maximaler Totalvariations-Mischzeit ist also von der Größenordnung n und geht somit gegen unendlich, da man die Spektrallücke nach unten gegen eine echt positive Konstante abschätzen kann. In diesem Fall ist dies aber nicht hinreichend für einen Cut-Off. Der Totalvariationsabstand ist zur Zeit an kleiner als $1-\epsilon$, weil man eine gute Chance hat, y_n in dieser Zeit zu erreichen. Der Grund, warum der Abstand in Totalvariation zur Zeit bn größer als ϵ ist, liegt darin, dass es länger dauert, den Punkt y_n durch den doppelt so langen rechten Arm zu erreichen.
Auch wenn die Behauptung, dass kein Cut-Off vorliegt, intuitiv ist, sind die Berechnungen, insbesondere für die Abschätzung der Spektrallücke, nicht ganz einfach. Falls man die Drift von n abhängig macht durch $p_{i,n} \geq 1 - 1/n, i = t, l, r$, vereinfachen sie sich, weil eine sehr starke Konzentration der stationären Verteilung im Verbindungspunkt der beiden Arme vorliegt. Für genauere Erläuterungen verweisen wir wiederum auf [Che06], Abschnitt 4.2.

7.3 Ausblick und offene Fragen

Das vorangegangene Beispiel zeigt, dass es bei Geburts- und Todesprozessen genügt einen Übergang von n nach 0 zu erlauben, damit das Kriterium von Peres bei Abstandsmessung in Totalvariation seine Gültigkeit verliert. Dennoch wird in einer Reihe von Klassen von Markov-Prozessen die Gültigkeit des Kriteriums von Peres vermutet; zu nennen sind insbesondere Random Walks auf endlichen Gruppen.

Eine weitere offene Frage ist, ob eine Äquivalenz von Totalvariations-Cut-Off und Separations-Cut-Off auch im Allgemeinen bei Familien irreduzibler, reversibler Markov-Prozesse analog zur Äquivalenz bei Geburts- und Todesprozessen vorliegt. In diesem Fall könnte man das Beispiel von Aldous auch als Gegenbeispiel für den Separationsfall verwenden.

Des Weiteren führt das Beispiel von Aldous zurück zur Frage von Diaconis nach einem allgemeinen Kriterium für das Auftreten eines Cut-Offs in einer Familie endlicher, ergodischer Markov-Prozesse.

A Anhang

Wir geben nun den Beweis von Satz 3.2.

Satz A.1. *Sei $M = (M_n)_{n\geq 0}$ eine positiv rekurrente DMK bezüglich einer Filtration $(\mathfrak{F}_n)_{n\geq 0}$ mit Zustandsraum S, stationärer Verteilung ξ^* und beliebiger Anfangsverteilung.*
(a) Ist T eine stark stationäre Zeit für M, so gilt

$$s(n) \leq P(T > n)$$

für alle $n \geq 0$.
(b) Es existiert bei endlichem Zustandsraum eine stark stationäre Zeit T für M, so dass

$$s(n) = P(T > n) \tag{A.1}$$

für alle $n \geq 0$. In diesem Fall heißt T ***minimale stark stationäre Zeit.***

Beweis. Es gilt mit Satz 3.1(b)

$$P(M_n = i) \geq P(M_n = i, T \leq n) = P(T \leq n)\xi^*(i) = (1 - P(T > n)\xi^*(i)$$

und folglich

$$P(T > n) \geq 1 - \frac{P(M_n = i)}{\xi^*(i)}$$

für alle $i \in S$ und $n \geq 0$

Nun konstruieren wir für den Beweis von (b) eine minimale stark stationäre Zeit. Wir setzen $\mu_n(i) = P(M_n = i)$,

$$a_n = \min_{i\in S} \frac{\mu_n(i)}{\xi^*(i)} = 1 - s(\mu_n, \xi^*).$$

und $k = \inf\{n \geq 0 : a_n > 0\}$. Sei nun T so definiert, dass $P(T < k) = 0$,

$$P(T = k | M_k = i) = \frac{a_k \xi^*(i)}{\mu_k(i)}, \quad i \in S,$$

und induktiv

$$P(T = n | M_n = i, T > n-1) = \frac{a_n - a_{n-1}}{\frac{\mu_n(i)}{\xi^*(i)} - a_{n-1}}, \quad i \in S, n > k.$$

Wenn wir zeigen können, dass

$$P(M_m = i, T = m) = \xi^*(i)(a_m - a_{m-1}) \tag{A.2}$$

für alle $m \geq k$ und $i \in S$ gilt, so ist T eine stark stationäre Zeit, denn per Summation über $i \in S$ liefert (A.2) $P(T = n) = a_n - a_{n-1}$ und damit

$$P(M_n = i, T = n) = \xi^*(i) P(T = n)$$

für alle $n \geq 0$. Außerdem folgt (A.1) aus der folgenden Rechnung

$$P(T > n) = 1 - P(T \leq n) = a_k + \sum_{l=k+1}^{n} (a_l - a_{l-1}) = 1 - a_n = s(\mu_n, \xi^*).$$

Wir beweisen nun (A.2) per Induktion über $n \geq k$. Für $n = k$ gilt dies nach Definition von $P(T = k | M_k = i)$. Nehmen wir also für den Induktionsschritt an, die Aussage sei wahr für alle $i \in S$ und $m \leq n-1$. Dies impliziert

$$P(M_n = i, T \leq n-1) = \xi^*(i) a_{n-1},$$

denn für $m \leq n-1$ gilt

$$\begin{aligned}
P(M_n = i, T = m) &= \sum_{j \in S} P(M_n = i, M_m = j, T = m) \\
&= \sum_{j \in S} P(M_n = i | M_m = j, T = m) P(M_m = j, T = m) \\
&= \sum_{j \in S} p_{ji} \xi^*(j)(a_m - a_{m-1}) \\
&= \xi^*(i)(a_m - a_{m-1}).
\end{aligned}$$

Folgendermaßen erhalten wir (A.2) für $m = n$

$$\begin{aligned}
P(M_n = i, T = n) &= P(T = n | M_n = i, T > n-1) P(M_n = i, T > n-1) \\
&= \frac{a_n - a_{n-1}}{\frac{\mu_n(i)}{\xi^*(i)} - a_{n-1}} \left(P(M_n = i) - P(M_n = i, T \leq n-1) \right) \\
&= \frac{a_n - a_{n-1}}{\frac{\mu_n(i)}{\xi^*(i)} - a_{n-1}} \left(\mu_n(i) - \xi^*(i) a_{n-1} \right) \\
&= \xi^*(i)(a_n - a_{n-1}).
\end{aligned}$$

Damit ist der Satz bewiesen. □

Literaturverzeichnis

[AD85] D. Aldous and P. Diaconis. Shuffling cards and stopping times. *Amer. Math. Monthly*, 93:333–348, 1985.

[Als05a] G. Alsmeyer. *Stochastische Prozesse*. Universität Münster, 2005.

[Als05b] G. Alsmeyer. *Stochastische Prozesse, Teil 2*. Universität Münster, 2005.

[Als05c] G. Alsmeyer. *Wahrscheinlichkeitstheorie*. Universität Münster, 2005.

[Bré99] P. Brémaud. *Markov chains. Gibbs Fields, Monte Carlo Simulation, and Queues*. Springer, New York, 1999.

[BS87] M. Brown and Y.S. Shao. Identifying coefficients in the spectral representation for first passage time distributions. *Prob. Eng. Info. Sci.*, 1:69–74, 1987.

[Che06] G.Y. Chen. *The cut off phenomenon for finite Markov chains*. Ph.D. dissertation, Cornell University, 2006.

[CSC08] G.Y. Chen and L. Saloff-Coste. The cutoff phenomenon for ergodic Markov processes. *Electronic Journal of Probability*, 13:26–78, 2008.

[Cul67] C. G. Cullen. *Matrices and Linear Transformations*. Addison-Wesley, Reading, Massachusetts, 1967.

[DB92] P. Diaconis and D. Beyer. Trailing the dovetail shuffle to its lair. *Ann. Appl. Prob*, 2:294–313, 1992.

[DF90a] P. Diaconis and J. A. Fill. Examples for the theory of strong stationary duality with countable state spaces. *Prob. Eng. Info. Sci.*, 4(1):157–180, 1990.

[DF90b] P. Diaconis and J. A. Fill. Strong stationary times via a new form of duality. *Ann. Probab.*, 18(4):1483–1522, 1990.

[Dia96] P. Diaconis. The cutoff phenomenon in finite Markov chains. *Proc. Nat. Acad. Sci.*, 93(4):1659–1664, 1996.

[DLP08] J. Ding, E. Lubetzky, and Y. Peres. Total variation cutoff in birth-death chains. 2008.

[DS81] P. Diaconis and M. Shahshahani. Generating a random permutation with random transpositions. *Z. Wahrsch. Verw. Gebiete*, 57(1):159–179, 1981.

[DSC06] P. Diaconis and L. Saloff-Coste. Separation cut-offs for birth and death chains. *Ann. Appl. Prob.*, 16(4):2098–2122, 2006.

[Fil91] J.A. Fill. Time to stationarity for a continous-time Markov chain. *Prob. Eng. Info. Sci.*, 5(1):61–76, 1991.

[Fil92] J. A. Fill. Strong stationary duality for continous-time Markov chains. part 1: theory. *J. Theoret. Probab.*, 5(1):45–70, 1992.

[Fil07] J.A. Fill. The passage time distribution for a birth-and-death distribution: Strong stationary duality gives a first stochastic proof. 2007.

[Kei79] J. Keilson. *Markov chain models — rarity and exponentiality.* Springer, New York, 1979.

[Kle08] A. Klenke. *Wahrscheinlichkeitstheorie.* Springer, 2008.

[KM59] S. Karlin and J. McGregor. Coincidence properties of birth and death processes. *Pacific J. Math*, 9:1109–1140, 1959.

[Reu57] G. E. H. Reuter. Denumerable Markov processes and the associated contraction semigroups on l. *Acta Math.*, 97(1):1–46, 1957.

Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe, dass alle Stellen der Arbeit, die wörtlich oder sinngemäß aus anderen Quellen übernommen wurden, als solche kenntlich gemacht sind und dass die Arbeit in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegt wurde.

Christoph Diehl

Münster, den 29. Januar 2009