

Zerlegung von Markov-Prozessen mit Hilfe regulärer Funktionen

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften

vorgelegt beim Fachbereich Mathematik
der Johann Wolfgang Goethe-Universität
in Frankfurt am Main

von
Kaya Memişoğlu
aus Hanau am Main

Frankfurt 2004
(D F 1)

vom Fachbereich Mathematik der Johann Wolfgang Goethe-Universität
als Dissertation angenommen.

Dekan : Prof. Dr. Klaus Johannson

Gutachter : Prof. Dr. Götz Kersting
Prof. Dr. Anton Wakolbinger

Datum der Disputation : 15. November 2004

Inhaltsverzeichnis

Einleitung	III
1 Zeitdiskreter Pfadraum	1
1.1 Notation	1
1.2 Hauptsatz	2
1.3 Beweis des Hauptsatzes	5
1.4 Erkennung der Zerlegungszeit	11
1.5 Verteilung des post- τ -Prozesses	16
2 Anwendungen zeitdiskreter Zerlegungen	19
2.1 Beispiel: Irrfahrten mit Drift	19
2.2 Beispiel: Driftlose Irrfahrt	23
2.3 Beispiel: Driftlose Irrfahrt mit absorbierendem Rand	25
2.4 Beispiel: Isotropische Irrfahrten auf Homogenen Bäumen	26
2.5 Beispiel: Pólyas Urnenschema	31
2.6 Beispiel: Verzweigungsprozesse	33
3 Zeitkontinuierlicher Pfadraum	43
3.1 Koterminale Zeiten	44
3.2 Hauptresultate	51
3.3 Beweis Hauptsatz 30	60
3.4 Erkennung der Zerlegungszeit	69
3.5 Der post- τ -Prozess im stetigen Fall	70
4 Anwendungen zeitkontinuierlicher Zerlegungen	73
4.1 Beispiel: Lévy-Prozesse	73
4.2 Beispiel: Driftlose Brownsche Bewegung	82
4.3 Beispiel: Brownsche Brücke	84
4.4 Beispiel: Bessel-Prozess	85
4.5 Beispiel: Verzweigungsprozesse	87

Einleitung

Zerlegung von Markov-Prozessen mit Hilfe regulärer Funktionen

Der historische Kontext

Spätestens seit den Arbeiten von Williams [Wil74] und Pitman [Pit75] sind Pfadzerlegungen ein wichtiges Arbeits- und Forschungsgebiet innerhalb der Stochastik. Die eben genannten Arbeiten entdeckten eine Zerlegung einer eindimensionalen Brownschen Bewegung mit negativem Drift in ein Anfangsstück bis zum Maximum und in ein daran anschließendes Endstück, wie folgt:

Hauptsatz (Zerlegung einer Brownschen Bewegung) Sei $(\hat{X}_t)_{t \geq 0}$ eine in 0 startende Brownsche Bewegung mit Drift 1 und sei M eine unabhängige exponentiell verteilte Zufallsvariable mit Erwartungswert $1/2$. Definiere die Stoppzeit

$$\tau := \sup\{t \geq 0 : \hat{X}_s < M \text{ für alle } s < t\}.$$

Sei weiter $(\check{X}_t)_{t \geq 0}$ ein in $\check{X}_0 = \hat{X}_\tau (= M)$ startender Prozess, dessen bedingte Verteilung gegeben $(\hat{X})_t$ und M der einer Brownschen Bewegung mit Drift -1 entspricht, die darauf bedingt ist, unterhalb der Schranke M zu bleiben. Dann ist der zusammengesetzte Prozess $(\bar{X}_t)_{t \geq 0}$ gegeben durch

$$\bar{X}_t := \begin{cases} \hat{X}_t & t < \tau \\ \check{X}_{t-\tau} & t \geq \tau \end{cases}$$

eine Brownsche Bewegung mit Drift -1 . ■

Es werden in den beiden Arbeiten noch weitere verwandte Zerlegungen dargestellt, der untersuchte Prozess ist jeweils immer eine Brownsche Bewegung. Diese Resultate inspirierten weitere Arbeiten; die ersten naheliegenden Verallgemeinerungen waren Zerlegungen einiger Klassen von Lévy-Prozessen (siehe [Ber91, Ber93, Ber92, Cha96, GP80]) – die verwendeten Beweistechniken waren jedoch sehr stark auf diese Prozesse zugeschnitten, es gab somit keinen Zugang zu anderen allgemeineren Prozessen. Pitman hat in einer unveröffentlichten Arbeit [Pit74] bereits harmonische Funktionen als wichtiges Hilfsmittel zu einer umfassenderen Zerlegungstheorie erkannt, jedoch beschränkte er sich weiterhin auf Brownsche Bewegungen. Später wurden noch weitere Zerlegungen vollkommen anderer Natur untersucht [Tan89, Tan90], und es gab auch Arbeiten hin zu allgemeinen Prinzipien und Konstruktionen für Pfadzerlegungen [Jac74, Mil77, PS73], die sich allerdings allesamt hauptsächlich mit der Charakterisierung von Zerlegungszeiten und mit der Existenz von bedingten Verteilungen beschäftigen, und diese auch weitreichend beantworten. Das Interesse an konstruktiven Resultaten ist weiterhin groß.

Gegenstand dieser Arbeit

Diese Arbeit eröffnet einen konstruktiven und sehr allgemeinen Zugang zu einer neuen großen Klasse an Zerlegungen von Markov-Prozessen. Ein Blick in das Inhaltsverzeichnis unterstreicht die Allgemeinheit und Anwendbarkeit der Zerlegung durch die vielfältigen angeführten Beispiele, die sich leicht noch weiter fortsetzen ließen.

Für die Zerlegung betrachten wir eine Markov-Kette (X_n) mit Kern P und mit Verteilung \mathbf{P} auf einem allgemeinen Raum S und eine positive harmonische Funktion h , d.h. es gilt $\mathbf{E}[h(X_1) | X_0] = h(X_0)$. Mit der Funktion h ist eine weitere Verteilung \mathbf{P}^h mit Kern P^h assoziiert, die allgemein als Doobsche h -Transformation bekannt ist. Diese Objekte reichen auch schon aus, um den folgenden Hauptsatz zu formulieren, der den Kern der Arbeit darstellt:

Hauptsatz (Zerlegung einer Markov-Kette) *Sei (\hat{X}_n) eine Markov-Kette mit Kern P^h auf einem Wahrscheinlichkeitsraum mit Maß \mathbb{P} und sei der Startzustand $\hat{X}_0 := o$, so dass $h(o) \neq 0$ ist. Weiter sei Y eine unabhängige Zufallsvariable mit Werten im Intervall $(h(o), \infty)$ und mit Verteilung*

$$\mathbb{P}(Y > y) = \frac{h(o)}{y}.$$

Definiere die zufällige Zeit τ durch

$$\tau := \sup\{n \geq 0 : h(\hat{X}_m) < h(\hat{X}_n) \leq Y \text{ für alle } m < n\}.$$

Sei außerdem $\check{X} = (\check{X}_n)$ eine weitere in $\check{X}_0 = \hat{X}_\tau$ startende Markov-Kette, deren bedingte Verteilung gegeben \hat{X} und Y übereinstimmt mit der einer Markov-Kette mit Übergangskern P , die darauf bedingt ist, innerhalb der Menge $\{x \in S : h(x) \leq h(\hat{X}_\tau)\}$ zu bleiben. Dann entspricht der Prozess $\bar{X} = (\bar{X}_n)$ definiert durch

$$\bar{X}_n := \begin{cases} \hat{X}_n & n < \tau \\ \check{X}_{n-\tau} & n \geq \tau \end{cases}$$

einer Markov-Kette mit Verteilung \mathbf{P} . ■

Es wird also eine Markov-Kette (X_n) mit Verteilung \mathbf{P} am Maximum von $(h(X_n))$ zerlegt. Die harmonische Funktion impliziert dadurch eine Ordnung auf dem Zustandsraum S . Die Harmonizität ist hierbei von zentraler Bedeutung, der Satz ist keinesfalls für allgemeine reellwertige Funktionen richtig. Dies stellt aber nur eine vergleichbar kleine Einschränkung dar, denn es werden viele Resultate für Irrfahrten abgedeckt, die eng verwandt mit der klassischen Zerlegung einer Brownschen Bewegung sind. Und auch neue Zerlegungen für Irrfahrten auf Bäumen und Verzweigungsprozesse sind einfache Anwendungen des Hauptsatzes.

Ähnliche Sätze gelten auch für zeitkontinuierliche Prozesse, allerdings sind dessen Formulierung aufgrund möglicher Sprünge ein ganzes Stück aufwendiger, weshalb an dieser Stelle nur auf das entsprechende Kapitel 3 dieser Arbeit verwiesen wird. Diese Sätze erlauben es dann, viele weitere klassische Zerlegungen in den Kontext dieser Arbeit einzubetten und zu erweitern. Darüber hinaus werden mit einfachen Mitteln neue, interessante Zerlegungen konstruiert, die Einblicke in das typische Pfadverhalten wichtiger Prozesse gestatten.

Zusammenarbeit

Ein Teil dieser Arbeit ist gemeinsam mit Prof. Götz Kersting in den „Annals of Probability“ veröffentlicht siehe [KM04] (zur Zeit im Druck).

Kapitel 1

Zeitdiskreter Pfadraum

1.1 Notation

Bevor wir weiter unten gleich den Hauptsatz für die Zerlegung eines zeitdiskreten Markov-Prozesses (im folgenden als Markov-Kette bezeichnet) präsentieren, sind noch einige Notationen einzuführen und grundlegende Annahmen und Vereinbarungen zu treffen. Wir wollen gleich den allgemeinsten Fall betrachten, dies impliziert glücklicherweise keine Verkomplizierung der Aussagen oder Beweise.

Sei $P(x, dy)$ ein Übergangskern einer Markov-Kette auf einem Zustandsraum (S, \mathcal{S}) , an den wir keinerlei weiteren topologische Forderungen stellen. Im Folgenden bezeichnet $(X_n)_{n \in \mathbb{N}_0}$ eine Markov-Kette auf S , adaptiert an eine Filtration (\mathcal{F}_n) , mit Übergangskern P und Zeitparameter $n \in \mathbb{N}_0 := \{0, 1, 2, 3, \dots\}$. Das zugehörige Wahrscheinlichkeitsmaß bezeichnen wir mit \mathbf{P}_x und die Erwartung mit \mathbf{E}_x , wobei x den Startzustand von (X_n) angibt.

Desweiteren sei

$$h : S \rightarrow \mathbb{R} \quad , \quad 0 \leq h < \infty$$

eine nicht-negative P -harmonische Funktion, d.h. es gilt für alle $x \in S$

$$Ph(x) := \int P(x, dy)h(y) = h(x).$$

Für eine solche Funktion h wird der h -TRANSFORMIERTE Kern P^h nach Doob definiert durch

$$P^h(x, dy) := \frac{1}{h(x)}P(x, dy)h(y)$$

für alle $x \in S$ mit $h(x) > 0$. Dann ist P^h wieder ein Wahrscheinlichkeitskern auf dem eingeschränkten Zustandsraum

$$S^h := \{x \in S : h(x) > 0\}.$$

Dieser Kern definiert auf einem geeigneten Wahrscheinlichkeitsraum durch Fortsetzung (Satz von Daniell Kolmogorov, vergleiche Abschnitt II.30 in [RW94], bzw. Satz von Ionescu Tulcea, vergleiche Theorem 5.17 in [Kal97]) eine Familie von Wahrscheinlichkeitsmaßen \mathbf{P}_x^h für Startpunkte $x \in S^h$ mit den Erwartungen \mathbf{E}_x^h , die gegeben sind durch

$$\mathbf{E}_x^h \phi(X_1, \dots, X_n) = h(x)^{-1} \mathbf{E}_x \phi(X_1, \dots, X_n) h(X_n) \quad (1.1)$$

für messbare nicht-negative Funktionen $\phi : S^n \rightarrow \mathbb{R}$. Damit ergibt sich die bekannte Tatsache, dass unter dem Maß \mathbf{P}_x^h der Prozess (X_n) eine Markov-Kette mit Übergangskern P^h ist.

1.2 Hauptsatz

Diese wenigen Voraussetzungen und Notationen reichen schon aus, um den Hauptsatz zu formulieren, der eine Zerlegung einer Markov-Kette (X_n) in ein Anfangsstück (\hat{X}_n) und ein Endstück (\check{X}_n) mit Hilfe einer harmonischen Funktion h beschreibt. Die Verteilung des Anfangsstückes geht aus einem Maßwechsel mittels h aus dem ursprünglichen Übergangskern der Kette hervor. Die Zerlegung findet am Maximum des Prozesses $(h(\hat{X}_n))$ unterhalb einer zufälligen Schranke statt; die Funktion h dient somit zugleich als Erzeuger einer geeigneten größer-kleiner-Relation auf dem Zustandsraum.

Hauptsatz 1 (Zerlegung einer Markov-Kette) *Sei $h \geq 0$ harmonisch und $o \in S$ so dass $h(o) > 0$ und sei (\hat{X}_n) eine Markov-Kette mit Übergangskern P^h und Startzustand o definiert auf einem Wahrscheinlichkeitsraum mit Maß \mathbb{P} . Weiter sei Y eine von (\hat{X}_n) unabhängige Zufallsvariable mit Werten im Intervall $(h(o), \infty)$ und mit Verteilung*

$$\mathbb{P}(Y > y) = \frac{h(o)}{y},$$

in anderen Worten: Y^{-1} ist uniform im Intervall $(0, h(o)^{-1})$ verteilt. Definiere die zufällige Zeit τ durch

$$\tau := \sup\{n \geq 0 : h(\hat{X}_m) < h(\hat{X}_n) \leq Y \text{ für alle } m < n\}.$$

Sei außerdem (\check{X}_n) eine weitere in $\check{X}_0 = \hat{X}_\tau$ startende Markov-Kette, deren bedingte Verteilung gegeben (\hat{X}_n) und Y übereinstimmt mit der einer Markov-Kette mit Übergangskern P , die darauf bedingt ist, innerhalb der Menge $\{x \in S : h(x) \leq h(\hat{X}_\tau)\}$ zu bleiben. Dann ist der Prozess (\bar{X}_n) definiert durch

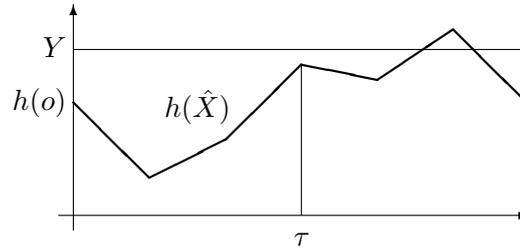
$$\bar{X}_n := \begin{cases} \hat{X}_n & n < \tau \\ \check{X}_{n-\tau} & n \geq \tau \end{cases}$$

eine Markov-Kette mit Übergangskern P , d.h. die Verteilung ist identisch zu einer Kette (X_n) mit Verteilung \mathbf{P}_o .

τ ist gerade der Zeitpunkt, bei dem der Prozess $(h(\bar{X}_n))$ sein globales Maximum zum ersten Mal annimmt. Es gibt natürlich auch Fälle, in denen τ den Wert ∞ mit positiver Wahrscheinlichkeit annehmen kann – in diesem Fall wird dann kein globales Maximum in endlicher Zeit angenommen, aber der Hauptsatz 1 umfasst auch diesen Fall, bei dem $\bar{X}_n = \hat{X}_n$ für alle n ist – dies bedeutet einfach, dass keine echte Zerlegung stattfindet und kein weiterer Prozess (\check{X}_n) an das Anfangsstück (\hat{X}_n) angeheftet wird.

Zerlegungszeiten

Eine wichtige Beobachtung ist, dass τ im Allgemeinen keine Stoppzeit sein muss, weder für den Prozess (\bar{X}_n) noch für (\hat{X}_n) : τ ist der letzte Augenblick, bei dem der reellwertige Prozess $(h(\hat{X}_n))$ alle vorigen Werte übertrifft bevor er die Schranke Y das erste mal überschreitet.



Die Zeit τ beinhaltet also offenbar Information über die Zukunft der Markov-Kette (\hat{X}_n) und kann somit im Allgemeinen keine Stoppzeit sein. Derartige zufällige Zeiten wurden bereits an unterschiedlicher Stelle untersucht; für den diskreten Fall ziehen wir die Definitionen einer ZERLEGUNGSZEIT (engl. SPLITTING TIME) von Jacobsen (siehe [Jac74]) bzw. Williams (unveröffentlicht, siehe ebenfalls [Jac74]) heran.

Definition 2 (Zerlegungszeit (Williams)) Sei (X_n) ein zeithomogener Markov-Prozess, und es bezeichne \mathcal{F}_n die von $(X_i)_{0 \leq i \leq n}$ erzeugte σ -Algebra und \mathcal{G}_n bezeichne die von $(X_i)_{n \leq i < \infty}$ erzeugte σ -Algebra. Eine zufällige Zeit τ heißt Zerlegungszeit, falls für jedes $i \in \mathbb{N}_0$ Ereignisse $F_i \in \mathcal{F}_i$ und $G_i \in \mathcal{G}_i$ existieren, so dass gilt

$$\{\tau = i\} = F_i \cap G_i.$$

Jacobsen hat später eine andere zu der obigen äquivalente Definition eingeführt, die die Idee der Zerlegungszeit noch stärker an die Pfade des Prozesses bindet.

Definition 3 (Zerlegungszeit (Jacobsen)) Eine zufällige Zeit τ heißt ZERLEGUNGSZEIT eines Prozesses (X_n) , falls sie im Sinne eines Funktionales auf dem Raum der Pfade die folgende Eigenschaft besitzt: Für jeweils zwei Pfade $\omega_1, \omega_2 : \mathbb{N}_0 \rightarrow S$ mit $\tau\omega_1 = \tau\omega_2$ und $\omega_1(\tau\omega_1) = \omega_2(\tau\omega_2)$ folgt für den zusammengesetzten Pfad $\omega : \mathbb{N}_0 \rightarrow S$ definiert durch

$$\omega(i) := \begin{cases} \omega_1(i) & \text{für } i \leq \tau\omega_1 \\ \omega_2(i) & \text{für } i \geq \tau\omega_1 \end{cases}$$

wiederum $\tau\omega = \tau\omega_1 (= \tau\omega_2)$.

Die in der Definition verwendete charakterisierende Eigenschaft wird auch als KREUZUNGSEIGENSCHAFT bezeichnet. Eine solche Zerlegungszeit τ für einen Prozess (X_n) besitzt eine verallgemeinerte starke Markov-Eigenschaft in dem folgenden Sinne:

$$\mathbf{P}(X_{\tau+1} \in A_1, \dots, X_{\tau+n} \in A_n \mid \mathcal{F}_\tau) = \mathbf{P}_{X_\tau}(X_1 \in A_1, \dots, X_n \in A_n \mid \tau = 0).$$

Der post- τ -Prozess ist also bedingt auf X_τ unabhängig vom Anfangsstück bis zum Zeitpunkt τ . Wie später gezeigt wird, ist der post- τ -Prozess selbst wieder eine Markov-Kette allerdings mit einem anderen Übergangskern (der natürlich von X_τ abhängen wird).

Wichtige Beispiele für Zerlegungszeiten sind der Zeitpunkt des ersten oder letzten Auftretens eines Maximums oder Minimums oder auch letzte Austrittszeiten aus einer Menge. Natürlich fällt auch die in Hauptsatz 1 konstruierte Zeit τ in die Kategorie einer Zerlegungszeit.

Duale Zerlegung

Aus der Harmonizität von h bezüglich des Kernes P folgt für den h -transformierten Kern P^h (der nur auf der Menge $S^h = \{x \in S : h(x) > 0\}$ definiert ist)

$$P^h 1/h(x) = \int_{S^h} P^h(x, dy) \frac{1}{h(y)} = \int_{S^h} \frac{1}{h(x)} P(x, dy) \leq 1/h(x),$$

also ist $1/h$ superharmonisch bezüglich des Kernes P^h . Aus dieser Tatsache folgt eine weitere Zerlegung, die im Sinne der h -Transformation in einer Dualitätsbeziehung zu der oben im Hauptsatz 1 dargestellten Zerlegung steht.

Hauptsatz 4 (Duale Zerlegung) *Sei $h \geq 0$ harmonisch bezüglich eines Kernes P und $o \in S$, so dass $h(o) > 0$ und sei (\tilde{X}_n^*) eine Markov-Kette mit Übergangskern P und Startzustand o definiert auf einem Wahrscheinlichkeitsraum mit Maß \mathbb{P} , und sei U eine unabhängige, uniform im Intervall $(0, h(o))$ verteilte Zufallsvariable. Definiere die Zufallsvariable τ^* durch*

$$\tau^* := \sup\{n \geq 0 : h(\tilde{X}_m^*) > h(\tilde{X}_n^*) \geq U \text{ für alle } m < n\}.$$

Sei weiter (\hat{X}_n^) eine in $\hat{X}_0^* = \tilde{X}_{\tau^*}^*$ startende Markov-Kette, deren bedingte Verteilung gegeben (\tilde{X}_n^*) und U übereinstimmt mit der einer Markov-Kette mit Übergangskern P^h , die darauf bedingt ist, innerhalb der Menge $\{x \in S : h(x) \geq h(\tilde{X}_{\tau^*}^*)\}$ zu bleiben. Dann ist der Prozess (\bar{X}_n^*) definiert durch*

$$\bar{X}_n^* := \begin{cases} \tilde{X}_n^* & n < \tau^* \\ \hat{X}_{n-\tau^*}^* & n \geq \tau^* \end{cases}$$

eine Markov-Kette mit Übergangskern P^h , d.h. die Verteilung ist identisch zu der einer Kette (X_n) mit Verteilung \mathbf{P}_o^h .

Bemerkungen

Bemerkung Die Zerlegung aus Hauptsatz 1 kann vielen Fällen dafür verwendet werden, um $\sup_n h(X_n)$ direkt in endlicher Zeit algorithmisch zu sampeln. Während ein naiver Ansatz in jedem Fall unendlich viele Simulationsschritte benötigen würde, liefert die Zerlegung einen alternativen Ansatz, sofern die Zerlegungszeit τ fast sicher mit nur endlich vielen Simulationsschritten bestimmt werden kann. In diesem Fall stimmen die Verteilungen von $h(\hat{X}_\tau)$ und $\sup_n h(X_n)$ nach Aussage des Hauptsatzes überein.

Allerdings ist nicht in jedem Fall die Zerlegungszeit τ fast sicher endlich, es werden später noch hinreichende Bedingungen angegeben werden, diese zusätzliche Forderung sicherstellen. \square

Bemerkung Mit h ist offenbar auch $g := ch$ für $c > 0$ harmonisch, und die Kerne P^h und P^g stimmen überein. Auch die Zerlegung bleibt bis auf einen Skalierungsfaktor identisch.

Die Situation gestaltet sich allerdings anders, falls in der Konstruktion h durch $g := h + c$ mit $c > 0$ ersetzt wird. Da (X_n) und (\bar{X}_n) identisch verteilt sind, bleibt die Verteilung von τ unberührt, und $\sup g(\bar{X}_n) = \sup h(\bar{X}_n) + c$. Allerdings stimmen die Kerne P^h und P^g

nicht mehr überein, und auch die algorithmische Bestimmbarkeit (im Sinne der Endlichkeit der benötigten Simulationsschritte) von τ wird im Allgemeinen verloren gehen, da g sich in einigen wesentlichen Eigenschaften von h unterscheiden kann. Insbesondere können nicht sowohl g als auch h minimal im Raum der harmonischen Funktionen sein, eine Eigenschaft die in Abschnitt 1.4 näher untersucht wird. \square

Bemerkung (Doob-Ungleichung) Als eine direkte Konsequenz der Zerlegung ergibt sich auch die Doob-Ungleichung angewandt auf das Martingal $(h(\bar{X}_n))$ bzw. $(h(X_n))$

$$\mathbb{P}\left(\sup_{1 \leq i \leq n} h(\bar{X}_i) \geq \lambda\right) \leq \mathbb{P}(Y > \lambda) = \frac{h(\bar{X}_0)}{\lambda} = \frac{\mathbb{E}h(\bar{X}_n)}{\lambda}.$$

\square

1.3 Beweis des Hauptsatzes

In diesem Abschnitt beweisen wir die beiden zentralen Hauptsätze 1 und 4 in mehreren Schritten. Wir übernehmen die bisher eingeführten Notationen, insbesondere bezeichnet (X_n) weiterhin eine Markov-Kette auf dem Raum (S, \mathcal{S}) und mit Übergangskern P , und die zugehörige Familie von Wahrscheinlichkeitsmaßen wird mit (\mathbf{P}_x) bezeichnet.

Die Verteilung des Endstückes

Als erstes untersuchen wir die Verteilung des post- τ -Prozesses, also der bedingten Kette (\check{X}_n) wie sie in Hauptsatz 1 eingeführt worden ist. Diese kann als eine Markov-Kette mit einem vom Startpunkt \check{X}_0 abhängigen Übergangskern beschrieben werden. Um den passenden Kern zu konstruieren, definieren wir zunächst die Funktionenfamilie q_s auf S durch

$$q_s(x) := \mathbf{P}_x(h(X_i) \leq s \forall i \in \mathbb{N}_0)$$

und die Stoppzeiten σ_s durch

$$\sigma_s := \inf\{i \geq 0 : h(X_i) > s\}.$$

Mit dieser Notation ist $q_s(x) = \mathbf{P}_x(\sigma_s = \infty)$, und die Funktion q_s ist zudem für alle $s \geq 0$ subharmonisch, denn es folgt aufgrund der Markov-Eigenschaft der Kette (X_n)

$$\begin{aligned} q_s(x) &= \mathbf{P}_x(h(X_i) \leq s \forall i \in \mathbb{N}_0) \\ &= \int_{\{x: h(x) \leq s\}} P(x, dy) \mathbf{P}_y(h(X_i) \leq s \forall i \in \mathbb{N}_0) \\ &\leq \int_S P(x, dy) \mathbf{P}_y(h(X_i) \leq s \forall i \in \mathbb{N}_0) \\ &= Pq_s(x). \end{aligned}$$

Auf der Menge $S_s := \{x \in S : h(x) \leq s\}$ gilt sogar Gleichheit, d.h. q_s ist harmonisch bezüglich des auf S_s eingeschränkten Kerns P . Eine weitere wichtige Eigenschaft von q_s stellt das folgende Lemma heraus

Lemma 5 Für alle $x \in S$ mit $h(x) \leq s$ gilt $q_s(x) > 0$.

Beweis Angenommen es wäre $q_s(x) = 0$, d.h. $\sigma_s < \infty$ fast sicher. Dann folgt aus dem Lemma von Fatou und der Martingal-Eigenschaft von $(h(X_n))$

$$\mathbf{E}_x h(X_{\sigma_s}) = \mathbf{E}_x \lim_{n \rightarrow \infty} h(X_{\sigma_s \wedge n}) \leq \liminf_{n \rightarrow \infty} \mathbf{E}_x h(X_{\sigma_s \wedge n}) = h(x) \leq s.$$

Da allerdings nach Definition der Stoppzeit σ_s fast sicher $h(X_{\sigma_s}) > s$ gilt, stellt dies ein Widerspruch dar. Also ist $q_s(x) = \mathbf{P}_x(\sigma_s = \infty) > 0$ und das Lemma ist bewiesen. ■

Dieses Lemma erlaubt es nun, für jedes s einen Kern Q_s auf der Menge S_s zu definieren, wobei dieser Kern durch eine Transformation mit q_s aus dem Kern P der Ausgangskette hervorgeht. Setze

$$Q_s(x, dy) := \frac{1}{q_s(x)} P(x, dy) q_s(y) = \mathbf{P}_x(X_1 \in dy \mid h(X_i) \leq s \forall i \in \mathbb{N}_0),$$

dann sind aufgrund der Harmonizität von q_s auf der Menge S_s die Kerne Q_s stochastisch, d.h. es gilt $Q_s(x, S_s) = 1$ für alle $x \in S_s$. Die Kerne Q_s beschreiben dann gerade die Verteilung der bedingten Kette (\check{X}_n) aus Hauptsatz 1, denn es gilt nach deren Konstruktion für $\check{X}_0 = x$ mit $h(x) = s$

$$\begin{aligned} \mathbb{P}(\check{X}_1 \in dx_1, \dots, \check{X}_n \in dx_n \mid \check{X}_0 = x) \\ &= \mathbf{P}_x(X_1 \in dx_1, \dots, X_n \in dx_n \mid h(X_i) \leq s \forall i \in \mathbb{N}_0) \\ &= \frac{P(x, dx_1) \cdots P(x_{n-1}, dx_n) q_s(x_n)}{q_s(x)} \quad \text{für } h(x_i) \geq s \text{ für } i = 1, \dots, n \\ &= Q_s(x, dx_1) \cdots Q_s(x_{n-1}, dx_n). \end{aligned}$$

Diese Beobachtung halten wir in dem folgenden Lemma fest.

Lemma 6 Für $s \geq 0$ ist unter dem Maß \mathbf{P}_x der Prozess (X_n) , bedingt darauf innerhalb der Menge S_s zu bleiben, eine Markov-Kette mit Übergangskern $Q_s(x, dy)$ und Zustandsraum S_s .

Analog kann für die duale Zerlegung wie in Hauptsatz 4 die Verteilung der bedingten h -transformierten Kette (\hat{X}_n^*) beschrieben werden mit den Kernen

$$Q_s^*(x, dy) := \frac{1}{q_s^*(x)} P^h(x, dy) q_s^*(y) \quad \text{mit} \quad q_s^*(x) := \mathbf{P}_x^h(h(X_i) \geq s \forall i \in \mathbb{N}_0).$$

Später wird für einige spezielle Fälle noch eine einfache Methode zur exakten Bestimmung der Kerne Q_s und Q_s^* angegeben (siehe Abschnitt 1.5).

Bemerkung Bei der Bestimmung von n -Schritt Übergangswahrscheinlichkeiten der bedingten Kette ist ein wenig Vorsicht geboten, denn im Allgemeinen gilt für $x, y \in S_s$ nicht notwendigerweise $Q_s^n(x, dy) = \frac{q_s(y)}{q_s(x)} P^n(x, dy)$. Schon für $n = 2$ gilt nur noch die Ungleichung

$$\begin{aligned} Q_s^2(x, dy) &= \int_{S_s} Q(x, dz) Q(z, dy) = \int_{S_s} \frac{q_s(z)}{q_s(x)} P(x, dz) \frac{q_s(y)}{q_s(z)} P(z, dy) \\ &\leq \int_S \frac{q_s(y)}{q_s(x)} P(x, dz) P(z, dy) = \frac{q_s(y)}{q_s(x)} P^2(x, dy). \end{aligned}$$

Falls man jedoch den Kern P auf die Menge S_s einschränkt mit der Notation $P|_{S_s}(x, A) := P(x, A \cap S_s)$ für $x \in S_s$, dann erhalten wir

$$Q_s^n(x, dy) = \frac{q_s(y)}{q_s(x)} P|_{S_s}^n(x, dy) \quad \text{für } x \in S_s.$$

Eine einfache Rechnung zeigt weiter, dass q_s harmonisch bezüglich des Kernes $P|_{S_s}$ ist, und somit lässt sich auch die von Q_s auf dem Pfadraum implizierte bedingte Verteilung durch eine Transformation von \mathbf{P} analog zu Gleichung (1.1) darstellen als

$$\begin{aligned} & \mathbf{E}_x[\phi(X_1, \dots, X_n) \mid X_i \in S_s \forall i \in \mathbb{N}] \\ &= \frac{1}{q_s(x)} \mathbf{E}_x[\phi(X_1, \dots, X_n) q_s(X_n); X_1, \dots, X_n \in S_s]. \end{aligned}$$

□

Bemerkung Betrachten wir die Ausgangskette (X_n) , und bezeichne L den Zeitpunkt des ersten globalen Supremums von $(h(X_n))$, also

$$L := \inf\{n \in \mathbb{N}_0 : h(X_n) = \sup_{i \in \mathbb{N}_0} h(X_i)\}.$$

L ist damit eine Zerlegungszeit im Sinne von Definition 2 oder Definition 3. Im vorigen Lemma wurde die Verteilung des post- L -Prozesses bestimmt; mit einer geeigneten Definition des Teilfeldes $\mathcal{F}(L)$ lässt sich die Aussage weiter zu einer schwachen Markov-Eigenschaft wie folgt verschärfen; in Abschnitt 3.1 wird dies für zeitkontinuierliche Zerlegungen in einem allgemeinen Kontext genauer diskutiert.

Satz 7 Sei (X_n) eine Markov-Kette und h eine harmonische Funktion und bezeichne L den Zeitpunkt des ersten globalen Supremums, dann gilt

$$\mathbf{E} \left[\prod_{i=1}^n f_i(X_{L+i}) \mid \mathcal{F}(L) \right] = \int \cdots \int Q_{h(X_L)}(X_L, dx_1) f_1(x_1) \cdots Q_{h(X_L)}(x_{n-1}, dx_n) f_n(x_n).$$

□

Die Verteilung des Anfangsstücks

Bezeichne im folgenden L weiterhin den Zeitpunkt des ersten globalen Supremums des ursprünglichen Prozesses $(h(X_n))$, also

$$L := \inf\{n \in \mathbb{N}_0 : h(X_n) = \sup_{i \in \mathbb{N}_0} h(X_i)\}.$$

Der post- τ - bzw. post- L -Prozess besitzen jeweils eine relativ einfache Struktur, deren Verteilung in Lemma 6 beschrieben wurde. Unter der Annahme, dass die Anfangsstücke bis zum Zeitpunkt τ bzw. L identisch verteilt sind, ist auch sofort aus der Konstruktion der Zerlegung ersichtlich, dass die Verteilungen der post- τ - und post- L -Prozesse übereinstimmen. Die wichtige und nicht direkt ersichtliche Aussage der Hauptsätze 1 und 4 betrifft vielmehr die Anfangsstücke der Prozesse. Ein zentrales Hilfsmittel zum Beweis stellt das folgende Lemma dar.

Lemma 8 Sei h harmonisch, $s > 0$ und $o, x \in S$ mit $0 < h(o) \leq h(x) \leq s$, und sei U uniform im Intervall $(0, h(o)^{-1})$ verteilt und unter dem Maß \mathbf{P}_x^h unabhängig vom Prozess (X_n) . Dann gilt

$$\begin{aligned}\mathbf{P}_x^h\left(U \leq \frac{1}{h(X_{\sigma_s})} I_{\{\sigma_s < \infty\}}\right) &= \frac{h(o)}{h(x)} \mathbf{P}_x(\sigma_s < \infty) \\ \mathbf{P}_x^h\left(\frac{1}{h(X_{\sigma_s})} I_{\{\sigma_s < \infty\}} < U \leq \frac{1}{h(x)}\right) &= \frac{h(o)}{h(x)} q_s(x).\end{aligned}$$

Beweis Da für die Stoppzeit σ_s per Definition $h(X_{\sigma_s}) > 0$ auf dem Ereignis $\{\sigma_s < \infty\}$ gilt, und der Prozess $(h(X_n))$ ein \mathbf{P}_x -Martingal ist, folgt

$$\mathbf{P}_x(\sigma_s \leq n) = \mathbf{E}_x\left(\frac{h(X_{\sigma_s})}{h(X_{\sigma_s})}; \sigma_s \leq n\right) = \mathbf{E}_x\left(\frac{h(X_n)}{h(X_{\sigma_s})}; \sigma_s \leq n\right).$$

Der in Gleichung (1.1) beschriebene Maßwechsel erlaubt es nun diese Gleichung wie folgt umzuschreiben:

$$\mathbf{P}_x(\sigma_s \leq n) = h(x) \mathbf{E}_x^h\left(\frac{1}{h(X_{\sigma_s})}; \sigma_s \leq n\right). \quad (1.2)$$

Und für $n \rightarrow \infty$ erhalten wir

$$\mathbf{P}_x(\sigma_s < \infty) = h(x) \mathbf{E}_x^h\left(\frac{1}{h(X_{\sigma_s})} I_{\{\sigma_s < \infty\}}\right), \quad (1.3)$$

und schließlich

$$q_s(x) = 1 - \mathbf{P}_x(\sigma_s < \infty) = h(x) \mathbf{E}_x^h\left(\frac{1}{h(x)} - \frac{1}{h(X_{\sigma_s})} I_{\{\sigma_s < \infty\}}\right). \quad (1.4)$$

Da nach Voraussetzung $h(X_{\sigma_s})^{-1} \leq h(x)^{-1} \leq h(o)^{-1}$ ist, und U unabhängig in $(0, h(o))$ verteilt ist, folgt

$$\begin{aligned}\mathbf{P}_x^h\left(U \leq \frac{1}{h(X_{\sigma_s})} I_{\{\sigma_s < \infty\}} \mid (X_n)\right) &= \frac{h(o)}{h(X_{\sigma_s})} I_{\{\sigma_s < \infty\}}, \\ \mathbf{P}_x^h\left(\frac{1}{h(X_{\sigma_s})} I_{\{\sigma_s < \infty\}} < U \leq \frac{1}{h(x)} \mid (X_n)\right) &= \frac{h(o)}{h(x)} - \frac{h(o)}{h(X_{\sigma_s})} I_{\{\sigma_s < \infty\}}.\end{aligned}$$

Durch Übergang zu Erwartungswerten, erhalten wir die Behauptung aus den Gleichungen (1.3) und (1.4). ■

Beweis von Hauptsatz 1 Die zufällige Zeit

$$L := \inf\{i \geq 0 : h(X_j) \leq h(X_i) \forall j \geq i\} = \inf\{i \geq 0 : h(X_i) = \sup_{n \in \mathbb{N}_0} h(X_n)\}$$

bezeichne den Zeitpunkt, zu dem der Prozess $(h(X_n))$ sein globales Maximum zum ersten Mal annimmt. Die Aussage des Hauptsatzes und die Konstruktion des Prozesses $(h(\bar{X}_n))$

implizieren, dass diese Zeit in Verteilung mit τ übereinstimmen muss. Im folgenden wird sogar die noch stärkere Aussage

$$\mathbf{P}_o((X_1, \dots, X_n) \in B, L = m) = \mathbb{P}_o((\bar{X}_1, \dots, \bar{X}_n) \in B, \tau = m) \quad (1.5)$$

für natürliche Zahlen $0 \leq m \leq n$ bewiesen. Seien hierzu die Mengen $B_{m,n}$ wie folgt definiert

$$B_{m,n} := \{(x_1, \dots, x_n) : h(x_1), \dots, h(x_{m-1}) < h(x_m) \geq h(x_{m+1}), \dots, h(x_n)\},$$

so dass die n -Tupel (x_1, \dots, x_n) unter der Funktion h an der Stelle m das erste Mal ihr Maximum annehmen. Hiermit erhalten wir eine weitere Beschreibung des Ereignisses auf der linken Seite in (1.5):

$$\begin{aligned} & \{(X_1, \dots, X_n) \in B, L = m\} \\ &= \{(X_1, \dots, X_n) \in B \cap B_{m,n}, h(X_j) \leq h(X_m) \forall j \geq n\}. \end{aligned}$$

Damit folgt

$$\begin{aligned} & \mathbf{P}_o((X_1, \dots, X_n) \in B, L = m) \\ &= \int_{B \cap B_{m,n}} P(o, dx_1) \dots P(x_{n-1}, dx_n) q_{h(x_m)}(x_n). \end{aligned}$$

Betrachten wir nun die Kette (\hat{X}_n) . Für diese gilt fast sicher $h(\hat{X}_\tau) = h(\check{X}_0) \geq h(\check{X}_1), h(\check{X}_2), \dots$ und zudem nach Konstruktion $h(\hat{X}_\tau) \leq Y$, so dass folgt

$$\begin{aligned} & \{(\bar{X}_1, \dots, \bar{X}_n) \in B, \tau = m\} \\ &= \{(\hat{X}_1, \dots, \hat{X}_m, \check{X}_1, \dots, \check{X}_{n-m}) \in B \cap B_{m,n}\} \cap \{\tau = m\} \end{aligned}$$

Das Ereignis $\{\tau = m\}$ zerfällt in zwei weitere Teilereignisse, je nachdem ob (\hat{X}_n) die Schranke Y überschreitet oder nicht:

$$\begin{aligned} &= \{(\hat{X}_1, \dots, \hat{X}_m, \check{X}_1, \dots, \check{X}_{n-m}) \in B \cap B_{m,n}\} \\ &\quad \cap (\{h(\hat{X}_m) \leq Y < h(\hat{X}_{\hat{\sigma}_m}), \hat{\sigma}_m < \infty\} \\ &\quad \cup \{h(\hat{X}_m) \leq Y, \hat{\sigma}_m = \infty\}), \end{aligned}$$

wobei die Stoppzeiten $\hat{\sigma}_m$ wie folgt definiert sind:

$$\hat{\sigma}_m := \inf\{i \geq m : h(\hat{X}_i) > h(\hat{X}_m)\}.$$

Damit folgt nun insgesamt

$$\begin{aligned} & \mathbb{P}_o((\bar{X}_1, \dots, \bar{X}_n) \in B, \tau = m) \\ &= \mathbb{P}_o((\hat{X}_1, \dots, \hat{X}_m, \check{X}_{m+1}, \dots, \check{X}_n) \in B, \tau = m) \\ &= \int_{B \cap B_{m,n}} P^h(o, dx_1) \dots P^h(x_{m-1}, dx_m) \\ &\quad \cdot Q_{h(x_m)}(x_m, dx_{m+1}) \dots Q_{h(x_m)}(x_{n-1}, dx_n) \\ &\quad \cdot \mathbb{P}\left(\frac{1}{h(\hat{X}_{\hat{\sigma}_m})} I_{\{\hat{\sigma}_m < \infty\}} < \frac{1}{Y} \leq \frac{1}{h(\hat{X}_m)} \mid \hat{X}_m = x_m\right) \end{aligned}$$

Ein Maßwechsel von P^h nach P und von $Q_{h(x_m)}$ nach P liefert

$$= \int_{B \cap B_{m,n}} P(o, dx_1) \dots P(x_{n-1}, dx_n) \frac{h(x_m) q_{h(x_m)}(x_n)}{h(o) q_{h(x_m)}(x_m)} \\ \cdot \mathbf{P}_{x_m}^h \left(\frac{1}{h(X_{\sigma_{h(x_m)}})} I_{\{\sigma_{h(x_m)} < \infty\}} < \frac{1}{Y} \leq \frac{1}{h(x_m)} \right)$$

Da Y^{-1} unabhängig und uniform im Intervall $(0, h(o))$ verteilt ist, führt Lemma 8 mit $s = h(x_m)$ schließlich zu

$$= \int_{B \cap B_{m,n}} P(o, dx_1) \dots P(x_{n-1}, dx_n) q_{h(x_m)}(x_n) \\ = \mathbf{P}_o((X_1, \dots, X_n) \in B, L = m),$$

womit die Behauptung (1.5) bewiesen ist.

Als nächstes zeigen wir

$$\mathbf{P}_o((X_1, \dots, X_n) \in B, L > n) = \mathbb{P}_o((\bar{X}_1, \dots, \bar{X}_n) \in B, \tau > n). \quad (1.6)$$

Aus der Gleichheit

$$\{(X_1, \dots, X_n) \in B, L > n\} = \{(X_1, \dots, X_n) \in B\} \\ \cap \{h(X_i) > \max(h(X_1), \dots, h(X_n)) \text{ für ein } i > n\}$$

folgt zunächst

$$\mathbf{P}_o((X_1, \dots, X_n) \in B, L > n) \\ = \int_B P(o, dx_1) \dots P(x_{n-1}, dx_n) \mathbf{P}_{x_n}(\sigma_{\max(h(x_1), \dots, h(x_n))} < \infty). \quad (1.7)$$

Andererseits gilt

$$\{(\bar{X}_1, \dots, \bar{X}_n) \in B, \tau > n\} \\ = \{(\hat{X}_1, \dots, \hat{X}_n) \in B\} \cap \{\bar{\sigma}_n < \infty, h(\hat{X}_{\bar{\sigma}_n}) \leq Y\},$$

wobei die Stoppzeiten $\bar{\sigma}_n$ definiert sind als

$$\bar{\sigma}_n := \inf\{i > n : h(\hat{X}_i) > \max(h(\hat{X}_1), \dots, h(\hat{X}_n))\}.$$

Damit folgt

$$\mathbb{P}_o((\bar{X}_1, \dots, \bar{X}_n) \in B, \tau > n) \\ = \int_B P^h(o, dx_1) \dots P^h(x_{n-1}, dx_n) \\ \mathbb{P}\left(\frac{1}{Y} \leq \frac{1}{h(\hat{X}_{\bar{\sigma}_n})} \cdot I_{\{\bar{\sigma}_n < \infty\}} \mid \hat{X}_1 = x_1, \dots, \hat{X}_n = x_n\right) \\ = \int_B P(o, dx_1) \dots P(x_{n-1}, dx_n) \frac{h(x_n)}{h(o)} \\ \mathbf{P}_{x_n}^h \left(\frac{1}{Y} \leq \frac{1}{h(X_{\sigma_{\max(h(x_1), \dots, h(x_n))}})} I_{\{\sigma_{\max(h(x_1), \dots, h(x_n))} < \infty\}} \right).$$

Da Y^{-1} unabhängig und inform im Intervall $(0, h(o))$ verteilt ist, folgt zusammen mit Gleichung (1.7) und mit Lemma 8 schließlich Gleichung (1.6).

Die Gleichungen (1.5) und (1.6) kombiniert ergeben nun die Behauptung des Hauptsatzes:

$$\mathbb{P}_o((\bar{X}_1, \dots, \bar{X}_n) \in B) = \mathbf{P}_o((X_1, \dots, X_n) \in B)$$

■

Als nächstes soll die duale Zerlegung in Hauptsatz 4 bewiesen werden; es stehen hierfür prinzipiell zwei Wege zur Verfügung. Der Satz kann analog zu den obigen Ausführungen bewiesen werden, oder auf die Zerlegung in Hauptsatz 1 zurückgeführt werden. Um die Dualität der beiden Zerlegungen zu unterstreichen, wird hier der zweite Weg gewählt.

Beweis von Hauptsatz 4 Zunächst nehmen wir an $h(x) > 0$ für alle $x \in S$. Dann gilt

$$\int_S P^h(x, dy) 1/h(y) = \int_S \frac{1}{h(x)} P(x, dy) h(y) 1/h(y) = 1/h(x),$$

also ist die Funktion $1/h$ harmonisch bezüglich dem Kern P^h , und eine Transformation mit $1/h$ liefert wieder den ursprünglichen Kern P , d.h. $(P^h)^{1/h} = P$. Setzen wir $Y := 1/U$, so ist $1/Y$ unabhängig und uniform in $(0, h(o))$ verteilt. Damit liefert Hauptsatz 1 eine Zerlegung eines Prozesses (X_n^*) mit Verteilung \mathbf{P}_o^h am Maximum von $(1/h(X_n^*))$ in ein Anfangsstück (\check{X}_n^*) mit Kern $(P^h)^{1/h} = P$ und ein Endstück (\hat{X}_n^*) mit einer aus P^h hervorgehenden bedingten Verteilung. Diese Zerlegung ist dann genau die in Hauptsatz 4 dargestellte Konstruktion.

Gibt es nun $x \in S$ mit $h(x) = 0$, so funktioniert obige Argumentationskette nicht mehr, da $1/h$ nicht mehr auf dem gesamten Raum S definiert ist. Allerdings ist für $\varepsilon > 0$ die Funktion $h(x) + \varepsilon$ wieder auf dem gesamten Raum S harmonisch bezüglich P , und es folgt für das assoziierte Maß $\mathbf{P}_o^{h+\varepsilon}$ mittels des in Gleichung (1.1) angegebenen Maßwechsels

$$\mathbf{P}_o^{h+\varepsilon} = \frac{h(o)}{h(o) + \varepsilon} \mathbf{P}_o^h + \frac{\varepsilon}{h(o) + \varepsilon} \mathbf{P}_o. \quad (1.8)$$

Diese Gleichung erlaubt es nun, die Konstruktion gemäß Hauptsatz 4 von $h + \varepsilon$ auf h zu übertragen. Denn es konvergieren einerseits die Maße $\mathbf{P}_o^{h+\varepsilon}$ in Verteilung gegen \mathbf{P}_o^h für $\varepsilon \rightarrow 0$, andererseits konvergiert auch die Verteilung der Konstruktion der dualen Zerlegung eines $\mathbf{P}_o^{h+\varepsilon}$ Prozesse (und insbesondere die darin verwendeten Zufallsvariable U , die Markov-Ketten (\check{X}_n^*) und (\hat{X}_n^*) und die Zerlegungszeit τ^*) für $\varepsilon \rightarrow 0$ gegen \mathbf{P}_o^h . ■

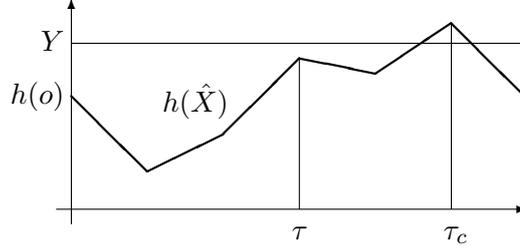
1.4 Erkennung der Zerlegungszeit

Nicht immer ist eine Zerlegung eines Prozesses möglich in dem Sinne, dass $\tau < \infty$ bzw. $\tau^* < \infty$ fast sicher gilt. Selbst die fast sicher zerlegbaren Prozesse können aus algorithmischer Sicht immer noch unerwünschte Eigenschaften besitzen. In diesem Abschnitt wollen wir die Konstruktion von einem algorithmischen Standpunkt aus untersuchen und der Frage nachgehen, unter welchen Voraussetzungen der Wert von τ fast sicher mit nur endlich vielen Schritten der Markov-Kette (\hat{X}_n) ermittelt werden kann.

Wir haben schon bemerkt, dass im Allgemeinen τ keine Stoppzeit ist und demnach nicht direkt durch Kenntnis der τ -Vergangenheit der Kette (\hat{X}_n) ermittelt werden kann; es ist vielmehr ein Blick in die Zukunft notwendig. Kennt man allerdings den Augenblick τ_c des Überschreitens der Schranke Y durch den Prozess $(h(\hat{X}_n))$, also

$$\tau_c := \inf\{n \geq 0 : h(\hat{X}_n) > Y\},$$

dann ist spätestens zu diesem Zeitpunkt der Wert von τ durch das Anfangsstück $(\hat{X}_n)_{n \leq \tau_c}$ festgelegt, denn τ_c dominiert τ in dem Sinne, dass $\tau < \tau_c$ fast sicher gilt.



Die Zerlegungszeit τ kann mit Hilfe der ÜBERKREUZUNGSZEIT τ_c nun auch wie folgt ermittelt werden

$$\tau = \inf \{ n \geq 0 : h(\hat{X}_n) = \sup_{m < \tau_c} h(\hat{X}_m) \}.$$

Damit kann der Wert von τ auf dem Ereignis $\{\sup_n h(\hat{X}_n) > Y\} = \{\tau_c < \infty\}$ mit nur endlich vielen Schritten festgestellt werden.

Die Situation sieht allerdings vollkommen anders im Falle $\{Y \geq \sup h\}$ aus. In diesem Fall kann τ durchaus einen endlichen Wert annehmen, während $\tau_c = \infty$ ist. Der Zeitpunkt τ kann in einem solchen Fall also nicht durch τ_c erkannt werden. Wir fassen dies in einen allgemeineren mathematischen Rahmen und definieren: *Eine Stoppzeit σ erkennt den Wert der Zerlegungszeit τ , falls*

$$\{\tau < \infty\} = \{\sigma < \infty\} = \{\tau \leq \sigma < \infty\} \quad \mathbb{P} - \text{f.s.}$$

Eine solche erkennende Stoppzeit σ steht in dem folgenden Zusammenhang zu τ und τ_c

$$\tau = \sup \{ n \leq \sigma : n < \tau_c, h(X_n) > \max_{i < n} h(X_i) \} \quad \mathbb{P} - \text{f.s.},$$

und τ_c selbst bestimmt den Wert τ genau dann, falls $\{\tau < \infty\} = \{\tau_c < \infty\}$ \mathbb{P} -fast sicher gilt.

Für eine große Klasse harmonischer Funktionen treten solche unerwünschten Phänomene glücklicherweise nicht auf, und für diese Klasse gelingt eine echte Pfadzerlegung (in dem Sinne $\tau < \infty$) fast sicher, zusätzlich wird die Zerlegungszeit τ von der in diesen Fällen fast sicher endlichen Stoppzeit τ_c dominiert. Um diese Funktionsklasse zu identifizieren, stellen wir zunächst fest, dass die Bedingung

$$\sup_n h(X_n) = \infty \quad \mathbf{P}_x^h - \text{f.s. } \forall x \in S^h \quad (1.9)$$

fast sicher zu einer Zerlegung führt, denn diese Bedingung impliziert in der Konstruktion $\sup_n h(\hat{X}_n) = \sup h = \infty$ und damit $\tau_c < \infty$ \mathbb{P} -fast sicher, also ist die Zerlegungszeit τ zum (fast sicher) endlichen Zeitpunkt τ_c erkennbar. Etwas allgemeiner betrachten wir die schwächere Bedingung

$$\sup_n h(X_n) = \sup h \quad \mathbf{P}_x^h - \text{f.s. } \forall x \in S^h, \quad (1.10)$$

aus dieser folgt $\sup_n h(\hat{X}_n) = \sup h$ \mathbb{P} -fast sicher. Falls nun unter dieser Voraussetzung $Y > \sup h$ ist, folgt sofort $\{\tau_c = \infty\}$ und die Zerlegungszeit τ nimmt entweder den Wert ∞ an, und

es ist dann $h(\hat{X}_n) < \sup h$ für alle n . Oder τ entspricht dem Zeitpunkt des ersten Auftretens des Supremums von h , d.h. $\tau = \inf\{n : h(\hat{X}_n) = \sup h\}$. Setzen wir

$$\tau_s := \inf\{n \geq 0 : h(\hat{X}_n) = \sup h\},$$

dann ist $\tau = \tau_s$ auf dem Ereignis $\{\tau_c = \infty\}$. Also wird die Zerlegungszeit τ von der (nicht notwendigerweise endlichen) Stoppzeit $\tau_c \wedge \tau_s$ erkannt.

Minimale harmonische Funktionen

Eine bekannte und wichtige Funktionsklasse, die die gewünschte Eigenschaft (1.10) besitzt, ist die Klasse der minimalen harmonischen Funktionen, die wie folgt definiert sind:

Definition 9 (Minimale harmonische Funktion) *Eine positive P -harmonische Funktion h heißt genau dann MINIMAL, falls für jede andere harmonische Funktion g mit $0 \leq g \leq h$ folgt $g = c \cdot h$ mit einer Konstanten $c \in \mathbb{R}_+$.*

Für einen abzählbaren Zustandsraum ist das folgende Resultat eine Folgerung aus Doobs Theorem (siehe Abschnitt III.29 in [RW94] über die Martin-Rand Theorie. Der hier angegebene Beweis gilt für allgemeine Zustandsräume.

Satz 10 *Sei h eine minimale P -harmonische Funktion, dann gilt für alle Startzustände $x \in S^h$*

$$h(X_n) \rightarrow \sup h \quad \mathbf{P}_x^h - \text{f.s.}$$

Beweis Es gilt die Ungleichung

$$\int_{S^h} P^h(x, dy) h(y)^{-1} = \int_{S^h} h(x)^{-1} P(x, dy) \leq h(x)^{-1} \int_S P(x, dy) = h(x)^{-1},$$

also ist $1/h(X_n)$ ein positives \mathbf{P}_x^h -Supermartingal. Aus Doobs Konvergenzsatz für Supermartingale folgt damit

$$1/h(X_n) \rightarrow H' \quad \mathbf{P}_x^h - \text{f.s.} \quad \text{bzw.} \quad h(X_n) \rightarrow H := 1/H' \quad \mathbf{P}_x^h - \text{f.s.}$$

für eine Zufallsvariable H mit Werten in $(0, \sup h]$. Es bleibt nun zu zeigen, dass $H = \sup h$ \mathbf{P}_x^h -fast sicher gilt.

Für $\gamma \geq 0$ definiere die Funktion $k_\gamma(x) : S^h \rightarrow \mathbb{R}_+$ durch

$$k_\gamma(x) := \mathbf{P}_x^h(H < \gamma).$$

Diese Funktion ist harmonisch bezüglich dem Kern P^h , denn es gilt

$$\begin{aligned} k_\gamma(x) &= \mathbf{P}_x^h(\lim_n h(X_n) < \gamma) \\ &= \int_{S^h} P^h(x, dy) \mathbf{P}_y^h(\lim_n h(X_n) < \gamma) = \int_{S^h} P^h(x, dy) k_\gamma(y). \end{aligned}$$

Betrachten wir das Produkt $h \cdot k_\gamma$, für das gilt

$$h(x)k_\gamma(x) = \int_{S^h} P(x, dy) h(y)k_\gamma(y) = \int_S P(x, dy) h(y)k_\gamma(y),$$

wobei $k_\gamma(x)$ für $x \notin S^h$ beliebig gewählt werden kann. Diese letzte Gleichung besagt gerade, dass die Funktion $h \cdot k_\gamma$ harmonisch bezüglich dem Kern P ist, außerdem gilt nach Definition von k_γ

$$h \cdot k_\gamma \leq h$$

Da nach Voraussetzung die harmonische Funktion h minimal ist, folgt daraus

$$h \cdot k_\gamma = \alpha h$$

für ein $\alpha \in \mathbb{R}_+$, und dies bedeutet

$$\mathbf{P}_x^h(H < \gamma) = \alpha$$

für alle $x \in S^h$ für eine (von γ abhängige) Konstante $\alpha = \alpha_\gamma \in [0, 1]$.

Als nächstes zeigen wir, dass α nur die Werte 0 oder 1 haben kann. Betrachten wir hierzu für $n \in \mathbb{N}$ die folgende Abschätzungen

$$\mathbf{P}_x^h\left(\{H < \gamma\} \cap \bigcap_{m \geq n} \{h(X_m) < \gamma\}\right) \leq \mathbf{P}_x^h(H < \gamma, h(X_n) < \gamma) \leq \mathbf{P}_x^h(H < \gamma).$$

Da $h(X_n) \rightarrow H$ \mathbf{P}_x^h -fast sicher, konvergiert die linke Seite gegen $\mathbf{P}_x^h(H < \gamma)$ für $n \rightarrow \infty$. Demnach konvergiert auch der mittlere Term gegen α , d.h.

$$\mathbf{P}_x^h(h(X_n) < \gamma, H < \gamma) \rightarrow \alpha. \quad (1.11)$$

Aus der Markov-Eigenschaft der Kette (X_n) folgt andererseits

$$\begin{aligned} \mathbf{P}_x^h(h(X_n) < \gamma, H < \gamma) &= \int_{y: h(y) < \gamma} \mathbf{P}_x^h(h(X_n) \in dy) \mathbf{P}_y^h(H < \gamma) \\ &= \int_{y: h(y) < \gamma} \mathbf{P}_x^h(h(X_n) \in dy) \alpha \\ &= \alpha \mathbf{P}_x^h(h(X_n) < \gamma). \end{aligned}$$

Falls γ ein Stetigkeitspunkt der Verteilung von H ist, so folgt für $n \rightarrow \infty$

$$\mathbf{P}_x^h(h(X_n) < \gamma, H < \gamma) \rightarrow \alpha \mathbf{P}_x^h(H \leq \gamma) = \alpha \mathbf{P}_x^h(H < \gamma) = \alpha^2. \quad (1.12)$$

Aus den Gleichungen (1.11) und (1.12) folgt also $\alpha = \alpha^2$ bzw. $\alpha = 0$ oder 1 für alle γ (bis auf höchstens abzählbar viele Unstetigkeitspunkte der Verteilung von H), dies bedeutet dass H eine degenerierte Verteilung besitzt, d.h.

$$H = \beta \quad \mathbf{P}_x^h - \text{f.s.}$$

für eine Konstante $\beta \in (0, \sup h]$. Außerdem hängt β nicht vom Startzustand x der Kette (X_n) ab. Da $1/h(X_n)$ ein nicht-negatives \mathbf{P}_x^h -Supermartingal ist, folgt mit Hilfe des Lemmas von Fatou

$$h(x)^{-1} \geq \liminf_n \mathbf{E}_x^h h(X_n)^{-1} \geq \mathbf{E}_x^h H^{-1} = \beta^{-1}.$$

bzw. $h(x) \leq \beta \leq \sup h$ für alle $x \in S^h$. Deshalb folgt $H = \sup h$ \mathbf{P}_x^h -fast sicher, was zu zeigen war. \blacksquare

Falls eine minimale harmonische Funktion zusätzlich unbeschränkt ist, dann ist in der Zerlegung $\tau < \tau_c < \infty$ fast sicher garantiert, da die Schranke Y fast sicher von $(h(\hat{X}_n))$ in endlicher Zeit überschritten wird. Der wichtigen Bedeutung für die Zerlegung wegen halten wir diese einfache Beobachtung fest.

Folgerung 11 *Sei h eine unbeschränkte minimale P -harmonische Funktion, dann gilt für alle Startzustände $x \in S^h$*

$$h(X_n) \rightarrow \infty \mathbf{P}_x^h - f.s. \quad \text{und} \quad \tau < \tau_c < \infty \mathbb{P}_x - f.s.$$

Grenzverhalten harmonischer Funktionen

Die Konvergenz gegen $\sup h$ von $(h(X_n))$ unter \mathbf{P}^h wie in Gleichung (1.10) angegeben stellt eine wichtige Voraussetzung für eine erfolgreiche Anwendung des Zerlegungssatzes dar. Mit den minimalen harmonischen Funktionen haben wir im vorigen Teilabschnitt schon eine wichtige Klasse kennengelernt, die dieser Bedingung genügen. In Beispielen und Anwendungen ist es darüber hinaus in vielen Fällen hilfreich, weitere äquivalente Bedingungen zur Verfügung stehen zu haben, die unter Umständen leichter zu verifizieren sind als Gleichung (1.10) oder gar die Minimalität einer gegebenen harmonischen Funktion. Dies leisten die beiden folgenden Sätze.

Satz 12 *Für $x \in S^h$ sind folgende Bedingungen äquivalent:*

1. $\sup_n h(X_n) = \sup h \quad \mathbf{P}_x^h - f.s.$
2. $\lim_n h(X_n) = \sup h \quad \mathbf{P}_x^h - f.s.$
3. $\lim_n h(X_n) \in \{0, \sup h\} \quad \mathbf{P}_x - f.s.$

Beweis Die Implikation 2. \Rightarrow 1. ist klar. Umgekehrt ist $1/h(X_n)$ unter \mathbf{P}_x^h ein nicht-negatives Supermartingal und konvergiert als solches. Damit existiert auch \mathbf{P}_x^h -fast sicher der Grenzwert $\lim h(X_n)$. Falls $h(X_n) < \sup h$ für alle n ist, folgt 2. aus 1., und falls $h(X_n) = \sup h$ für ein n ist, dann folgt aus dem Maximumprinzip für harmonische Funktionen $h(X_m) = \sup h$ \mathbf{P}_x^h -fast sicher für alle folgenden $m \geq n$, und damit ebenfalls 1.

Da $(h(X_n))$ ein nicht negatives \mathbf{P}_x -Martingal ist, und $(1/h(X_n))$ ein nicht negatives \mathbf{P}_x^h -Supermartingal ist, folgt aus dem Martingalkonvergenzsatz, dass $(h(X_n))$ fast sicher gegen eine Zufallsvariable H konvergiert, sowohl unter der Verteilung \mathbf{P}_x als auch unter \mathbf{P}_x^h . Dabei ist $H < \infty$ \mathbf{P}_x -f.s. und $H > 0$ \mathbf{P}_x^h -fast sicher. Die Maße \mathbf{P}_x und \mathbf{P}_x^h selbst stehen mittels H wie folgt zueinander in Beziehung

$$d\mathbf{P}_x^h(\cdot \cap \{H < \infty\}) = \frac{H}{h(x)} d\mathbf{P}_x(\cdot), \quad d\mathbf{P}_x(\cdot \cap \{H > 0\}) = \frac{h(x)}{H} d\mathbf{P}_x^h(\cdot),$$

siehe Theorem 1, Abschnitt VII.6 in [Shi95]. Damit folgt die Implikation 2. \Rightarrow 3., denn es gilt

$$\mathbf{P}_x(0 < H < \sup h) = \int_{0 < y < \sup h} \frac{h(x)}{y} \mathbf{P}_x^h(H \in dy) = 0.$$

Und umgekehrt folgt aus 3. auch 2., denn

$$\begin{aligned} \mathbf{P}_x^h(H < \sup h) &= \int_{0 \leq y < \sup h} \frac{y}{h(x)} \mathbf{P}_x(H \in dy) \\ &= \int_{0 < y < \sup h} \frac{y}{h(x)} \mathbf{P}_x(H \in dy) = 0. \end{aligned}$$

■

Das duale Gegenstück stellt der folgende Satz dar, der auf dem selben Weg bewiesen werden kann, und uns für die Wahl einer harmonischen Funktion in der dualen Zerlegung von Nutzen sein kann.

Satz 13 Für $h(x) < \sup h$ sind folgende Bedingungen äquivalent:

1. $\inf_n h(X_n) = \inf h$ \mathbf{P}_x -f.s.
2. $\lim_n h(X_n) = \inf h$ \mathbf{P}_x -f.s.
3. $\lim_n h(X_n) \in \{\inf h, \infty\}$ \mathbf{P}_x^h -f.s.

Eine Kombination der Bedingungen aus den beiden letzten Sätzen führt zu einer äquivalenten Charakterisierung von Bedingung (1.9), die sich nun auch als hinreichend für eine fast sichere duale Zerlegung erweist:

Folgerung 14 Für $x \in S_h$ gilt die folgende Äquivalenz

$$\sup_n h(X_n) = \infty \quad \mathbf{P}_x^h - \text{f.s.} \quad \Leftrightarrow \quad \inf_n h(X_n) = 0 \quad \mathbf{P}_x - \text{f.s.}$$

Beweis Aus der linken Bedingung folgt mittels Satz 12 $\lim h(X_n) \in \{0, \infty\}$ \mathbf{P}_x -fast sicher. Gleichzeitig folgt mittels Satz 13 $\lim h(X_n) = \inf h$ \mathbf{P}_x -fast sicher. Also muss $\inf h = 0$ sein, und wir erhalten insgesamt $\inf h(X_n) = 0$ \mathbf{P}_x -fast sicher. Die Umkehrung folgt analog. ■

Diese Eigenschaft kann auch mit Hilfe der Stopzeit τ_c bzw. mit ihrem dualen Analogon

$$\tau_c^* := \inf\{n \geq 0 : h(\check{X}_n^*) < U\}$$

wie folgt charakterisiert werden:

$$\tau_c < \infty \quad \mathbb{P} - \text{f.s.} \quad \Leftrightarrow \quad \tau_c^* < \infty \quad \mathbb{P} - \text{f.s.}$$

1.5 Verteilung des post- τ -Prozesses

In Lemma 6 wurde bereits die Verteilung des Prozesses (\check{X}_n) mit Hilfe der harmonischen Funktion q_s bestimmt. Die dort konstruierten Kerne Q_s besitzen in einigen Fällen übersichtliche, explizite Darstellungen, in denen die Funktion q_s die Gestalt eines Erwartungswertes besitzt.

Die Verteilung von (\check{X}_n)

Da der Prozess $(h(X_n))$ unter dem Maß \mathbf{P}_x ein nicht-negatives Martingal ist, konvergiert dieser fast sicher gegen einen Grenzwert $H := \lim h(X_n)$. Aus der Martingaleigenschaft folgt außerdem für alle $s > 0$ und alle $x \in S$ mit $h(x) \leq s$ und alle $m \in \mathbb{N}$

$$\begin{aligned} h(x) &= \mathbf{E}_x h(X_{m \wedge \sigma_s}) \\ &= \mathbf{E}_x [h(X_{\sigma_s}); \sigma_s \leq m] + \mathbf{E}_x [h(X_m); \sigma_s > m] \end{aligned}$$

mit $\sigma_s := \inf\{i \geq 0 : h(X_i) > s\}$. Da der erste Integrand $h(X_{\sigma_s}) \cdot I(\sigma_s \leq m)$ in m wächst, konvergiert dessen Erwartungswert mittels monotoner Konvergenz für $m \rightarrow \infty$. Der zweite Integrand ist nicht-negativ und durch s beschränkt, und dessen Erwartungswert konvergiert ebenfalls für $m \rightarrow \infty$ mittels dominierter Konvergenz gegen $H \cdot I(\sigma_s = \infty)$. Wir erhalten

$$h(x) = \mathbf{E}_x [h(X_{\sigma_s}); \sigma_s < \infty] + \mathbf{E}_x [H; \sigma_s = \infty]. \quad (1.13)$$

Falls zusätzlich die harmonische Funktion h die Eigenschaft (1.9) bzw. eine der beiden äquivalenten Eigenschaften aus Folgerung 14 besitzt, gilt für den Grenzwert \mathbf{P}_x -fast sicher $H = 0$, womit der zweite Summand in Gleichung (1.13) verschwindet. Durch Übergang zur bedingten Erwartung erhalten wir die Gleichung

$$h(x) = \mathbf{P}_x(\sigma_s < \infty) \mathbf{E}_x [h(X_{\sigma_s}) | \sigma_s < \infty],$$

und hiermit ergibt sich die folgende Gleichung für q_s

$$q_s(x) = 1 - \mathbf{P}_x(\sigma_s < \infty) = 1 - \frac{h(x)}{\mathbf{E}_x [h(X_{\sigma_s}) | \sigma_s < \infty]}. \quad (1.14)$$

Hiermit lassen sich die bedingten Kerne nun darstellen als

$$\begin{aligned} Q_s(x, dy) &= \frac{q_s(y)}{q_s(x)} P(x, dy) \\ &= \frac{\mathbf{E}_x [h(X_{\sigma_s}) | \sigma_s < \infty] (\mathbf{E}_y [h(X_{\sigma_s}) | \sigma_s < \infty] - h(y))}{\mathbf{E}_y [h(X_{\sigma_s}) | \sigma_s < \infty] (\mathbf{E}_x [h(X_{\sigma_s}) | \sigma_s < \infty] - h(x))} P(x, dy), \end{aligned}$$

wobei Q_s nur auf der Menge $S_s := \{z \in S : h(z) \leq s\}$ wohldefiniert ist, denn für alle $x \in S$ mit $h(x) > s$ gilt $q_s(x) = 0$.

Denken wir nun schon vorausgreifend an einen zeitkontinuierlichen Prozess mit stetigen Pfaden, dann gilt mit entsprechender Notation $h(X_{\sigma_s}) = s$ fast sicher auf dem Ereignis $\{\sigma_s < \infty\}$, es gilt also $\mathbf{E}_x [h(X_{\sigma_s}) | \sigma_s < \infty] = s$. Diese Beobachtung ermöglicht eine besonders einfache Darstellung der bedingten Übergangskerne im zeitkontinuierlichen Fall. Diese Aussage stimmt so natürlich für keinen zeitdiskreten Prozess, aber in einigen Beispielen kann der Erwartungswert dennoch einfach berechnet werden.

Die Verteilung von (\hat{X}_n^*)

Betrachten wir die Verteilung des bedingten h -transformierten Prozesses aus der dualen Zerlegung in Hauptsatz 4. Falls zusätzlich $h(x) \neq 0$ für alle $x \in S$, dann ist $1/h$ harmonisch bezüglich des Kernes P^h , also ist $(1/h(X_n))$ ein Martingal unter dem Maß \mathbf{P}^h . Für eine

Funktion h wie in Folgerung 14 erhalten wir somit analog für alle $s > 0$ und alle $x \in S$ mit $h(x) \geq s$

$$1/h(x) = \mathbf{P}_x^h(\sigma_s^* < \infty) \mathbf{E}_x^h[1/h(X_{\sigma_s^*}) | \sigma_s^* < \infty] \quad \text{mit} \quad \sigma_s^* := \inf\{i \geq 0 : h(X_i) < s\}.$$

Damit folgt

$$q_s^*(x) = 1 - \mathbf{P}_x^h(\sigma_s^* < \infty) = 1 - \frac{1}{\mathbf{E}_x^h[1/h(X_{\sigma_s^*}) | \sigma_s^* < \infty] h(x)},$$

und somit besitzt der Übergangskern der bedingten Verteilung des post- τ^* -Prozesses die Darstellung

$$\begin{aligned} Q_s^*(x, dy) &= \frac{q_s^*(y)}{q_s^*(x)} P^h(x, dy) \\ &= \frac{\mathbf{E}_x^h[1/h(X_{\sigma_s^*}) | \sigma_s^* < \infty] (\mathbf{E}_y^h[1/h(X_{\sigma_s^*}) | \sigma_s^* < \infty] h(y) - 1)}{\mathbf{E}_y^h[1/h(X_{\sigma_s^*}) | \sigma_s^* < \infty] (\mathbf{E}_x^h[1/h(X_{\sigma_s^*}) | \sigma_s^* < \infty] h(x) - 1)} P(x, dy) \end{aligned} \quad (1.15)$$

für alle $x \in S_s^* := \{z \in S : h(z) \geq s\}$.

Falls es Punkte x gibt mit $h(x) = 0$, dann ist $1/h$ nicht mehr harmonisch bezüglich P^h . Diesen Fall können wir jedoch wieder durch die harmonische Funktion $h + \varepsilon$ für $\varepsilon > 0$ abfangen. Es gilt dann vermittels des in Gleichung (1.1) angegebenen Maßwechsels, wie auch schon im Beweis der dualen Zerlegung (Hauptsatz 4) angegeben

$$\mathbf{P}_x^{h+\varepsilon} = \frac{h(x)}{h(x) + \varepsilon} \mathbf{P}_x^h + \frac{\varepsilon}{h(x) + \varepsilon} \mathbf{P}_x.$$

Damit erhalten wir analog zu Gleichung (1.13)

$$\begin{aligned} 1/(h(x) + \varepsilon) &= \mathbf{P}_x^{h+\varepsilon}(\sigma_s^* < \infty) \mathbf{E}_x^{h+\varepsilon}[1/(h(X_{\sigma_s^*}) + \varepsilon) | \sigma_s^* < \infty] \\ &\quad + \mathbf{P}_x^{h+\varepsilon}(\sigma_s^* = \infty) \mathbf{E}_x^{h+\varepsilon}[1/(H + \varepsilon) | \sigma_s^* = \infty], \end{aligned}$$

und da die Maße $\mathbf{P}^{h+\varepsilon}$ für $\varepsilon \rightarrow 0$ schwach gegen \mathbf{P}^h konvergieren, und da $h(X_n) \rightarrow H = \infty$ \mathbf{P}^h -fast sicher konvergiert, erhalten wir schließlich wieder

$$1/h(x) = \mathbf{P}_x^h(\sigma_s^* < \infty) \mathbf{E}_x^h[1/h(X_{\sigma_s^*}) | \sigma_s^* < \infty].$$

Damit können kann auch in diesem Fall der Verteilung des bedingten h -transformierten Prozesses mit Hilfe der Kerne in Gleichung (1.15) beschrieben werden.

Kapitel 2

Anwendungen zeitdiskreter Zerlegungen

In diesem Kapitel sollen einige Beispiele zu den Zerlegungssätzen dargestellt und untersucht werden. Viele Beispiele sind schon bekannt, allerdings wurden diese noch nicht in einen solch allgemeinen Rahmen eingebettet, wie es hier geschieht. Diese Beispiele finden dennoch ihren Platz in dieser Arbeit, da sie einerseits selbst schon einen ästhetischen Wert besitzen, und andererseits die vielfältige Anwendbarkeit der gefundenen Zerlegungen unterstreichen. Gleichzeitig wird durch geeignete Interpretationen der Verteilungen und Konstruktionen der jeweiligen Prozesse auch die Konkretheit der Zerlegungen in der Anwendung herausgearbeitet.

Weitere neue Beispiele lassen sich schnell finden, denn hierfür wird nicht viel benötigt: Eine Markov Kette und eine harmonische Funktion. Die zahlreichen Arbeiten über extremale harmonische Funktionen bieten daher eine reiche Quelle vielfältiger Beispiele für Zerlegungen.

Eine wichtige und gut beschreibbare Klasse von Beispielen ergeben sich aus den verschiedenen Varianten von Irrfahrten; einige der folgenden wurden durch die Übersichtsarbeit von Wolfgang Woess [Woe94] inspiriert.

2.1 Beispiel: Irrfahrten mit Drift

In diesem Abschnitt werden die Hauptresultate auf Irrfahrten (X_n) in $S = \mathbb{R}^d$ angewendet, also auf Markov-Ketten deren Kern invariant gegenüber Verschiebungen ist,

$$P(x, dy) = P(0, -x + dy) \quad \forall x, y \in \mathbb{R}^d.$$

Im folgenden gehen wir davon aus, dass der Kern nicht degeneriert ist, d.h. dass er keine deterministische Irrfahrt beschreibt.

Verteilung und h -Transformation

Zunächst betrachten wir eine Irrfahrt mit Drift, d.h. die Zuwächse haben eine endliche Erwartung ungleich Null,

$$\mu := \mathbf{E}_0 X_1 = \int x P(0, dx) \neq 0.$$

Für die Existenz wenigstens einer nicht trivialen minimalen harmonischen Funktion wird noch eine weitere Voraussetzung bezüglich der Laplace Transformierten benötigt, nämlich

$$\varphi(u) := \int P(0, dx) e^{\langle u, x \rangle} < \infty$$

für alle $u \in \mathbb{R}^d$, wobei $\langle \cdot, \cdot \rangle$ das Skalarprodukt in \mathbb{R}^d bezeichnet. Damit ist φ die Laplace-Transformierte und es gilt für den Drift $\mu = \text{grad } \varphi(0)$. Definieren wir die Funktionen

$$h_u(x) := e^{\langle u, x \rangle}, \quad u \in C := \{u : \varphi(u) = 1\}, \quad (2.1)$$

so sind diese harmonisch bezüglich P , denn es gilt aufgrund der Translationsinvarianz von P und der Definition der Menge C

$$Ph_u(x) = \int P(x, dy) e^{\langle u, y \rangle} = \int P(0, dy) e^{\langle u, y+x \rangle} = e^{\langle u, x \rangle} \varphi(u) = h_u(x),$$

siehe auch [NS66] und [DSW60]. Die h_u -transformierten Kerne sind gegeben durch

$$P^{h_u}(x, dy) = P(x, dy) e^{\langle u, y-x \rangle}.$$

Diese Kerne sind wiederum invariant gegenüber Translationen, damit ist (X_n) unter dem Maß $\mathbf{P}_x^{h_u}$ ebenfalls eine Irrfahrt. Der zugehörige Drift ist dann

$$\mu_u := \mathbf{E}_0^{h_u} X_1 = \int x P^{h_u}(0, dx) = \int x e^{\langle u, x \rangle} P(0, dx) = \text{grad } \varphi(u).$$

Eigenschaften der h -Transformation

Um für $u \in C \setminus \{0\}$ das asymptotische Verhalten von $h_u(X_n)$ unter den Maßen \mathbf{P}_0 und $\mathbf{P}_0^{h_u}$ zu untersuchen, reicht es den Exponenten der Funktion h_u – also $\langle u, X_m \rangle$ – zu betrachten. Der Erwartungswert $\mathbf{E}_0 \langle u, X_1 \rangle$ wird mit Hilfe der Erzeugendenfunktion der reellwertigen Zufallsvariablen $\langle u, X_1 \rangle$ bestimmt, die gegeben ist durch

$$\varphi_u(\lambda) := \mathbf{E}_0 e^{\lambda \langle u, X_1 \rangle} = \int e^{\lambda \langle u, x \rangle} P(0, dx) = \varphi(\lambda \cdot u).$$

Genauso berechnet sich die Erzeugendenfunktion von $\langle u, X_1 \rangle$ unter dem Maß $\mathbf{P}_0^{h_u}$ als

$$\mathbf{E}_0^{h_u} e^{\lambda \langle u, X_1 \rangle} = \int e^{\lambda \langle u, x \rangle} e^{\langle u, x \rangle} P(0, dx) = \varphi_u(\lambda + 1).$$

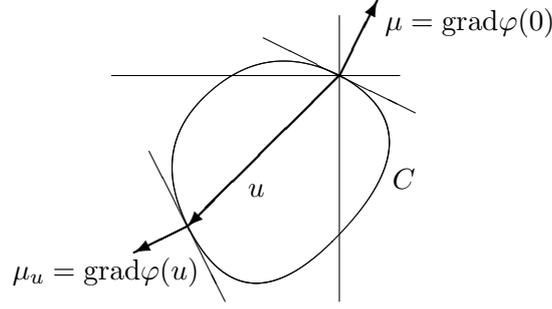
Die Erwartungswerte lassen sich nun mit Hilfe der Kettenregel berechnen,

$$\begin{aligned} \mathbf{E}_0 \langle u, X_1 \rangle &= \varphi'_u(0) = \langle u, \text{grad } \varphi(0) \rangle \\ \mathbf{E}_0^{h_u} \langle u, X_1 \rangle &= \varphi'_u(1) = \langle u, \text{grad } \varphi(u) \rangle. \end{aligned}$$

Nach Voraussetzung ist P nicht degeneriert und somit ist φ_u streng konvex. Somit besitzt die Gleichung $\varphi_u(\lambda) = 1$ genau nur die beiden Lösungen $\lambda = 0$ und $\lambda = 1$, und wegen der Konvexität ist $\varphi'_u(0) < 0$ und $\varphi'_u(1) > 0$. Aus diesen Ungleichungen folgt schließlich für alle $u \in C \setminus \{0\}$

$$\mathbf{E}_0 \langle u, X_1 \rangle = \langle u, \mu \rangle < 0 \quad \text{und} \quad \mathbf{E}_0^{h_u} \langle u, X_1 \rangle = \langle u, \mu_u \rangle > 0 \quad (2.2)$$

Es ist bekannt, dass die Menge C eine konvexe, glatte Mannigfaltigkeit in \mathbb{R}^d mit Kodimension 1 ist (Siehe [NS66] und [Hen63]), und es ergibt sich insgesamt folgendes Bild, das den Zusammenhang zwischen der Menge C , einem Punkt u , und der Drifte μ_u und μ darstellt.



Zerlegung

Aus dem starken Gesetz der großen Zahlen und der Irrfahrtsstruktur von $\langle u, X_n \rangle$ folgt aus den Ungleichungen (2.2) $\langle u, X_n \rangle \rightarrow -\infty$ \mathbf{P}_x -fast sicher und $\langle u, X_n \rangle \rightarrow +\infty$ $\mathbf{P}_x^{h_u}$ -fast sicher. Somit gilt für die harmonische Funktion h_u

$$h_u(X_n) \rightarrow 0 \quad \mathbf{P}_x - \text{f.s.}, \quad \text{und} \quad h_u(X_n) \rightarrow \infty \quad \mathbf{P}_x^{h_u} - \text{f.s.}$$

Also ist die Bedingung in Gleichung (1.9) erfüllt, und somit sind in Hinsicht auf eine Pfadzerlegung die Zerlegungszeiten τ und τ^* und die Kreuzungszeiten τ_c und τ_c^* fast sicher endlich,

$$\tau < \tau_c < \infty \quad \mathbb{P} - \text{f.s.} \quad \text{und} \quad \tau^* < \tau_c^* < \infty \quad \mathbb{P} - \text{f.s.}$$

Die einfache Struktur der Irrfahrt erlaubt es diese Pfadzerlegung weiter im Detail zu beschreiben. Wir wählen den Startpunkt $o := 0 \in \mathbb{R}^d$, dann ist $h_u(o) = 1$. Sei die Zufallsvariable Y mit Werten in $[1, \infty)$ wie in Hauptsatz 1, also unabhängig und mit der Verteilung $\mathbb{P}(Y > y) = 1/y$ für $y > 1$. Gehen wir zu der Zufallsvariable M über definiert durch

$$M := \|u\|^{-1} \log Y.$$

Dann ist $\mathbb{P}(M > m) = \mathbb{P}(Y > e^{\|u\|m}) = e^{-\|u\|m}$, also ist M exponentiell verteilt mit Erwartungswert $\|u\|^{-1}$. Außerdem gilt

$$h_u(x) > Y \quad \Leftrightarrow \quad \frac{\langle u, x \rangle}{\|u\|} > M \quad \Leftrightarrow \quad x \in H_u(M),$$

wobei $H_u(M) := \{x : \langle u, x \rangle > \|u\|M\}$ ein Halbraum in \mathbb{R}^d ist mit Abstand M vom Ursprung. Wir erhalten für jedes $u \in C, u \neq 0$ jeweils ein Verfahren zur Konstruktion der ursprünglichen Irrfahrt mit Startpunkt 0

1. Sei (\hat{X}_n) eine Irrfahrt mit Übergangskern P^{h_u} , und sei M eine unabhängige exponentiell verteilte Zufallsvariable mit Erwartungswert $\|u\|^{-1}$.

2. Setze

$$\tau := \sup\{n \geq 0 : \hat{X}_n \notin H_u(M), \langle u, \hat{X}_m \rangle < \langle u, \hat{X}_n \rangle \forall m < n\}.$$

Dies ist der erste Zeitpunkt, zu dem die Irrfahrt (\hat{X}_n) am nächsten am Halbraum $H_u(M)$ ist, aber noch nicht in diesen eingetreten ist.

3. Sei (\check{X}_n) eine Irrfahrt mit Übergangskern P und Startzustand $\check{X}_0 = \hat{X}_\tau$, die darauf bedingt ist, in dem Halbraum $\{x : \langle u, x \rangle \leq \langle u, \check{X}_0 \rangle\}$ zu bleiben. Setze

$$\bar{X}_n := \begin{cases} \hat{X}_n & n < \tau \\ \check{X}_{n-\tau} & n \geq \tau. \end{cases}$$

Dann ist (\bar{X}_n) eine Irrfahrt mit Übergangskern P .

Insbesondere falls (X_n) eine 1-dimensionale Irrfahrt mit einem negativen Drift ist, dann beinhaltet C neben 0 nur noch eine weitere Lösung u . Der h -transformierte Prozess (\hat{X}_n) ist dann wieder eine Irrfahrt mit einem positiven Drift und τ ist der letzte Zeitpunkt eines aufsteigenden Leiterpunktes bevor (\hat{X}_n) die Schranke M überschreitet. Die duale Zerlegung besitzt eine analoge Darstellung. In Dimension $d \geq 2$ existieren wesentlich mehr minimale harmonische Funktionen, die jeweils wieder zu einer Pfadzerlegung führen.

Beispiel: Irrfahrt mit normalverteilten Zuwächsen

Ein besonders einfaches Beispiel stellt eine Irrfahrt (X_n) auf \mathbb{R} mit normalverteilten Zuwächsen mit einem negativen Driftterm -1 dar, d.h.

$$\mathcal{L}(X_{n+1} - X_n) = N(-1, 1) \quad \text{bzw.} \quad P(0, dx) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x+1)^2}{2}} dx.$$

Die Laplace Transformierte φ ist in diesem Fall gegeben durch

$$\varphi(u) = \exp\left(-u + \frac{u^2}{2}\right),$$

und die Menge C der Lösungen der Gleichung $\varphi(u) = 1$ besteht aus den Punkten $C = \{0, 2\}$, also ist die Funktion $h : \mathbb{R} \rightarrow \mathbb{R}_+$ gegeben durch

$$h(x) := \exp(2x)$$

minimal und harmonisch. Die h -transformierte Verteilung ist dann wieder eine Normalverteilung, allerdings mit Erwartungswert 1, d.h.

$$P^h(0, dx) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-1)^2}{2}} dx.$$

Wir erhalten nun eine zeitdiskrete Variante der Williams' Zerlegung einer Brownschen Bewegung mit Drift.

1. Sei (\hat{X}_n) eine Irrfahrt mit $N(1, 1)$ verteilten Zuwächsen und sei M eine unabhängige exponentiell verteilte Zufallsvariable mit Erwartungswert $1/2$.

2. Setze

$$\tau := \sup\{n \geq 0 : \hat{X}_n < M, \hat{X}_m < \hat{X}_n \forall m < n\}.$$

3. Sei (\check{X}_n) eine Irrfahrt mit $N(-1, 1)$ verteilten Zuwächsen und Startzustand $\check{X}_0 = \hat{X}_\tau$, die darauf bedingt ist, die Schranke \hat{X}_τ nicht zu überschreiten. Setze

$$\bar{X}_n := \begin{cases} \hat{X}_n & n < \tau \\ \check{X}_{n-\tau} & n \geq \tau. \end{cases}$$

Dann ist (\bar{X}_n) eine Irrfahrt mit normalverteilten Zuwächsen mit Drift -1 .

2.2 Beispiel: Driftlose Irrfahrt

Driftlose Irrfahrten besitzen im Allgemeinen keine nicht trivialen positiven harmonischen Funktionen, wie sie in Gleichung (2.1) konstruiert wurden, da dort die Menge C dann nur aus dem Nullpunkt besteht. Betrachtet man allerdings Raum-Zeit-harmonischen Funktionen $h : S \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$, dann ändert sich die Situation.

Verteilung der Irrfahrt

Betrachten wir als Beispiel eine einfache Irrfahrt $(R_n) = (X_n, Y_n)$ auf $\mathbb{Z} \times \mathbb{Z}$ mit Übergangswahrscheinlichkeiten

$$\mathbf{P}(R_{n+1} = (x+1, y) \mid R_n = (x, y)) = \frac{1}{2}$$

und $\mathbf{P}(R_{n+1} = (x, y+1) \mid R_n = (x, y)) = \frac{1}{2},$

d.h. zu jedem Zeitpunkt bewegt sich die Irrfahrt auf dem Gitter entweder um einen Schritt nach rechts oder nach oben, es gilt immer $X_n + Y_n = n$. Durch eine einfache Transformation kann man dies auch als eindimensionale Irrfahrt (Z_n) auf \mathbb{Z} interpretieren mit

$$Z_n := n - 2X_n = Y_n - X_n,$$

(Z_n) ist damit eine symmetrische Irrfahrt.

Harmonische Funktionen

Für die Markov-Kette (R_n) existiert eine reichhaltige Menge minimaler harmonischer Funktionen (siehe [Saw97]) der Gestalt

$$h_\alpha(x, y) := 2^{x+y} \alpha^x (1-\alpha)^y, \quad 0 \leq \alpha \leq 1$$

mit den Konventionen $0^0 = 1$ und $0^x = 0$ für $x > 0$. Die h_α -transformierte Verteilung besitzt dann die Übergangswahrscheinlichkeiten

$$\mathbf{P}^{h_\alpha}(R_{n+1} = (x+1, y) \mid R_n = (x, y)) = \alpha$$

und $\mathbf{P}^{h_\alpha}(R_{n+1} = (x, y+1) \mid R_n = (x, y)) = 1 - \alpha.$

Aus dem starken Gesetz der großen Zahlen folgt

$$\lim_{n \rightarrow \infty} \frac{X_n}{X_n + Y_n} = \mathbf{E}^{h_\alpha} \frac{X_n}{n} = \alpha \quad \mathbf{P}^{h_\alpha}\text{-fast sicher.}$$

Die harmonische Funktion h_α und die assoziierte Verteilung \mathbf{P}^{h_α} übersetzen sich einfach auf die Irrfahrt (Z_n) zu einer Raum-Zeit-harmonischen Funktion $h'_\alpha : \mathbb{Z}_+ \times \mathbb{Z} \rightarrow \mathbb{R}_+$

$$h'_\alpha(n, z) = 2^n \alpha^{(n-z)/2} (1-\alpha)^{(n+z)/2},$$

die die symmetrische Irrfahrt (Z_n) transformiert sich in eine Irrfahrt mit Verteilung

$$\mathbf{P}^{h'_\alpha}(Z_{n+1} = z+1 \mid Z_n = z) = 1 - \alpha$$

und $\mathbf{P}^{h'_\alpha}(Z_{n+1} = z-1 \mid Z_n = z) = \alpha.$

Es gilt dann auch hier ein starkes Gesetz der großen Zahlen,

$$\lim_{n \rightarrow \infty} \frac{Z_n}{n} = 1 - 2\alpha \quad \mathbf{P}^{h'_\alpha}\text{-fast sicher.}$$

Zerlegung

Da jede der Funktionen h_α und damit auch h'_α minimal und unbeschränkt ist, eröffnen diese die Möglichkeit zur Zerlegung einer symmetrischen Irrfahrt (Z_n) auf \mathbb{Z} , deren Anfangsstück durch eine Irrfahrt (\hat{Z}_n) mit einem beliebigen Drift $1 - 2\alpha$ mit $\alpha \in [0, 1]$ beschrieben werden kann. Die für die Zerlegung mitentscheidenden Niveaulinien der auf $\mathbb{R}_+ \times \mathbb{R}$ fortgesetzten harmonischen Funktion h'_α sind implizit als Lösungen (t, z) der Gleichungen

$$2^t \alpha^{(t-z)/2} (1-\alpha)^{(t+z)/2} = c \quad c \in \mathbb{R}_+$$

gegeben; für festes c bildet die Lösungsmenge jeweils eine Gerade mit der Gleichung

$$t \mapsto t \cdot \frac{\log(4(1-\alpha)\alpha)}{\log((1-\alpha)/\alpha)} - \frac{2 \log c}{\log((1-\alpha)/\alpha)}.$$

Diese Geraden haben für $0 \leq \alpha < 1/2$ eine positive und für $1/2 < \alpha \leq 1$ eine negative Steigung, die jeweils mit c_α bezeichnet wird, also

$$c_\alpha := \frac{\log(4(1-\alpha)\alpha)}{\log((1-\alpha)/\alpha)}.$$

Dieser Vorzeichenwechsel von c_α bei $\alpha = \frac{1}{2}$ passt zu der Tatsache, dass die h'_α -transformierte Irrfahrt für $\alpha < \frac{1}{2}$ einen positiven und für $\alpha > \frac{1}{2}$ einen negativen Drift besitzt.

Wir wollen für eine Zerlegung wieder von der Zufallsvariablen Y zu der exponentiell verteilten Zufallsvariablen $\log Y$ übergehen. Damit erhalten wir für $z \in \mathbb{Z}$ und $n \in \mathbb{N}$

$$\begin{aligned} h'_\alpha(n, z) > Y &\Leftrightarrow \log h'_\alpha(n, z) > \log Y \\ &\Leftrightarrow z + n \cdot \frac{\log(4\alpha(1-\alpha))}{\log((1-\alpha)/\alpha)} > \frac{2 \log Y}{\log((1-\alpha)/\alpha)} \quad \text{für } \alpha < 1/2 \\ \text{und } z + n \cdot \frac{\log(4\alpha(1-\alpha))}{\log((1-\alpha)/\alpha)} &< \frac{2 \log Y}{\log((1-\alpha)/\alpha)} \quad \text{für } \alpha > 1/2. \end{aligned}$$

Wir erhalten damit folgende durch $\alpha \in (0, 1), \alpha \neq \frac{1}{2}$ parametrisierte Zerlegung einer freien symmetrischen Irrfahrt mit Zuwächsen ± 1 jeweils mit Wahrscheinlichkeit $1/2$.

1. Sei M exponentiell verteilt mit Erwartungswert $2/\log((1-\alpha)/\alpha)$, und sei (\hat{Z}_n) eine von M unabhängige Irrfahrt auf einem gemeinsamen Wahrscheinlichkeitsraum mit Maß \mathbb{P} , so dass gilt

$$\begin{aligned} \mathbb{P}(\hat{Z}_{n+1} = z + 1 \mid \hat{Z}_n = z) &= 1 - \alpha \\ \text{und } \mathbb{P}(\hat{Z}_{n+1} = z - 1 \mid \hat{Z}_n = z) &= \alpha. \end{aligned}$$

2. Setze für $\alpha < 1/2$

$$\tau := \sup\{n \geq 0 : \hat{Z}_n + nc_\alpha < M \text{ und } \hat{Z}_m + mc_\alpha < \hat{Z}_n + nc_\alpha \forall m < n\},$$

und setze für $\alpha > 1/2$ entsprechend

$$\tau := \sup\{n \geq 0 : \hat{Z}_n + nc_\alpha > M \text{ und } \hat{Z}_m + mc_\alpha > \hat{Z}_n + nc_\alpha \forall m < n\},$$

3. Sei (\check{Z}_n) eine symmetrische Irrfahrt auf \mathbb{Z} mit Zuwächsen ± 1 jeweils mit Wahrscheinlichkeit $1/2$ und Startpunkt $\check{Z}_0 = \check{Z}_\tau$, die auf $h'_\alpha(n + \tau, \check{Z}_n) \leq h'_\alpha(\tau, \check{Z}_0)$ für alle $n \geq 0$ bedingt ist. Setze

$$\bar{Z}_n := \begin{cases} \hat{Z}_n & n < \tau \\ \check{Z}_{n-\tau} & n \geq \tau. \end{cases}$$

Dann ist (\bar{Z}_n) eine symmetrische Irrfahrt auf \mathbb{Z} mit mit Zuwächsen ± 1 .

Bemerkung Als Mittel der oben angegebenen harmonischen Funktionen erhält man (siehe [Saw97])

$$h(x, y) = \int_0^1 h_\alpha(x, y) d\alpha = 2^{x+y} \frac{x!y!}{(x+y+1)!},$$

und unter dem Maß \mathbf{P}^h ergeben sich dann die Übergangswahrscheinlichkeiten der Markov-Kette (R_n) als

$$\mathbf{P}^h(R_{n+1} = (x+1, y) \mid R_n = (x, y)) = \frac{x+1}{x+y+2}$$

und $\mathbf{P}^h(R_{n+1} = (x, y+1) \mid R_n = (x, y)) = \frac{y+1}{x+y+2},$

damit entspricht diese Kette einem Pólya Urnen Modell, bei dem zu einem Zeitpunkt $x+1$ rote und $y+1$ blaue Kugeln in der Urne befinden, es wird eine Kugel rein zufällig gezogen und mit einer weiteren Kugel der gleichen Farbe zurückgegeben.

Diese harmonische Funktion h ist natürlich nicht mehr minimal, und erfüllt auch sonst nicht mehr die Eigenschaften, die zu einer fast sicher endlichen Zerlegungszeit τ führen. In Abschnitt 2.5 kommen wir auf ein verwandtes Beispiel zurück. \square

2.3 Beispiel: Driftlose Irrfahrt mit absorbierendem Rand

Eine Irrfahrt (Z_n) auf der reellen Achse \mathbb{R} mit erwarteten Zuwächsen

$$\mathbf{E}_0 Z_1 = \int x P(0, dx) = 0$$

besitzt zunächst keine interessanten harmonischen Funktionen. Die folgende Modifikation ändert dies: Alle negativen Zustände absorbieren die Irrfahrt, d.h. wir betrachten eine driftlose Irrfahrt mit Absorption in $(-\infty, 0)$,

$$X_n := Z_{\tau \wedge n}, \quad \text{wobei } \tau := \inf\{i \geq 0 : Z_i < 0\}.$$

Dann gilt $P(x, dy) = P(0, -x + dy)$ falls $x \geq 0$, und $P(x, dy) = \delta_x(dy)$ falls $x < 0$, wobei δ_x das Dirac-Maß in x ist. Es ist bekannt [BD94, AD01, Don98], dass

$$h(x) := \begin{cases} \sum_{k \geq 0} \mathbb{P}(H_k \leq x), & x \geq 0, \\ 0, & x < 0, \end{cases}$$

eine P -harmonische Funktion definiert, wobei H_k die Zeitpunkte der strikt absteigenden Leiterpunkte der Irrfahrt (Z_n) mit Startpunkt 0 sind. Zudem ist die Funktion h offenbar monoton wachsend, d.h. das Maximum bzw. Minimum der Irrfahrt (Z_n) und des Prozesses $(h(Z_n))$ fallen zeitlich zusammen.

Der Prozess (X_n) besitzt auch unter dem Maß \mathbf{P}_x^h eine einfache Interpretation, er ist in Verteilung identisch zu dem Prozess (Z_n) bedingt darauf, nicht in das Intervall $(-\infty, 0)$ einzutreten [BD94].

Hiermit eröffnen sich nun zwei zueinander duale Pfadzerlegungen: Eine Zerlegung am Maximum einer driftlosen Irrfahrt mit Absorbtion im Negativen und am Minimum einer driftlosen Irrfahrt bedingt im Positiven zu bleiben.

2.4 Beispiel: Isotropische Irrfahrten auf Homogenen Bäumen

Sei T_r ein homogener Baum mit $r \geq 3$ Kanten an jedem Knoten. Dieser Baum entspricht dem Cayley-Graphen einer freien Gruppe G , die von r Elementen $A := \{a_1, \dots, a_r\}$ erzeugt wird mit $a_1^2 = \dots = a_r^2 = e$, wobei e das neutrale Element ist. Jeder Knoten x entspricht genau einem reduzierten formalen Wort

$$x = a_{i_1} a_{i_2} \cdots a_{i_d}$$

mit $i_j \neq i_{j+1}$ für alle $1 \leq j \leq d-1$. Die Distanz eines Knoten x vom Zentrum (bzw. der Wurzel) e des Baumes ist definiert als

$$\begin{aligned} |x| &:= d \text{ für reduziertes } x = a_{i_1} a_{i_2} \cdots a_{i_d} \\ \text{und } |e| &:= 0, \end{aligned}$$

und für $x, y \in T_r$ sei $c(x, y)$ der letzte gemeinsame Knoten in dem Graphen. Zwei reduzierte Knoten $x = a_{i_1} a_{i_2} \cdots a_{i_k}$ und $y = a_{j_1} a_{j_2} \cdots a_{j_l}$ sind Nachbarn genau dann, wenn $|k-l|=1$ und $|y^{-1}x|=1$, wobei $y^{-1} = a_{j_l} a_{j_{l-1}} \cdots a_{j_1}$.

Bezeichnen wir die Menge der reduzierten unendlichen Folgen in $A^{\mathbb{N}}$ mit R und die Menge der reduzierten Worte der Länge d mit R_d , d.h.

$$\begin{aligned} R &:= \{a_{i_1} a_{i_2} \cdots : i_j \neq i_{j+1} \forall j \in \mathbb{N}\} \\ R_d &:= \{a_{i_1} a_{i_2} \cdots a_{i_d} : i_j \neq i_{j+1} \forall 1 \leq j < d\}. \end{aligned} \tag{2.3}$$

Mit dieser Definition entspricht R_d der Menge der Knoten des Baumes T_r und R der Menge der Endpunkte, die im Unendlichen liegen.

Als ein praktische Hilfsmittel für das weitere Vorgehen sei für einen Endpunkt $\omega \in R$ die Funktion d_ω definiert durch $d_\omega(x) := 2|c(x, \omega)| - |x|$. Jeder Knoten x besitzt jeweils genau einen Nachbar y mit $d_\omega(y) = d_\omega(x) + 1$ und genau $r-1$ Nachbarn y mit $d_\omega(y) = d_\omega(x) - 1$, die Funktion $d_\omega(x)$ ist um so größer, je näher x bei dem Randpunkt ω liegt.

Betrachten wir nun eine isotropische Irrfahrt auf dem Graphen T_r , d.h. eine Irrfahrt (X_n) mit Werten in T_r und

$$P(x, xa_i) = \frac{1}{r} \quad \forall 1 \leq i \leq r.$$

Harmonische Funktionen

Sawyer and Cartwright haben gezeigt [Saw97, CS91], dass im isotropischen Fall alle minimalen harmonischen Funktionen gegeben sind durch

$$h_\omega(x) = (r-1)^{d_\omega(x)} = (r-1)^{2|c(x,\omega)|-|x|}$$

wobei $\omega \in R$ ein Randpunkt im Unendlichen des Baumes T_r ist. Der Raum der minimalen harmonischen Funktionen ist somit isomorph zu der in Gleichung (2.3) definierten Menge R der reduzierten unendlichen Folgen in A .

Ein Maßwechsel des Kerns P nach P^h bedingt die Irrfahrt darauf, im Punkt ω zu enden, wie weiter unten noch gezeigt wird. Die konstante harmonische Funktion ist hier nicht minimal, sie ist vielmehr eine Mischung über alle möglichen Endpunkte in T_r und ist gegeben durch

$$1 = \int_R h_\omega(x) \mu(d\omega), \tag{2.4}$$

wobei μ die Gleichverteilung auf R mit Gesamtmasse 1 ist. Dies kann mit einer einfachen Rechnung nachvollzogen werden. Für ein reduziertes und endliches Wort x hängt die Länge des gemeinsamen Anfangsstückes $c(x,\omega)$ nur von den ersten $|x|$ Stellen von ω ab, damit ist auch der Wert von $h_\omega(x)$ als Funktion in ω durch die ersten $|x|$ Stellen bestimmt. Insgesamt gibt es $r(r-1)^{|x|-1}$ verschiedene reduzierte Worte der Länge $|x|$. Diese Überlegungen helfen, das Integral in Gleichung (2.4) zu vereinfachen.

$$\begin{aligned} \int_{A^{\mathbb{N}}} h_\omega(x) \mu(d\omega) &= \frac{1}{r(r-1)^{|x|-1}} \sum_{\omega \in R_{|x|}} (r-1)^{2|c(x,\omega)|-|x|} \\ &= \frac{1}{r(r-1)^{|x|-1}} \sum_{k=0}^{|x|} \sum_{\substack{\omega \in R_{|x|} \\ |c(x,\omega)|=k}} (r-1)^{2k-|x|} \end{aligned}$$

Es gilt nun die Anzahl der Summanden in der zweiten Summe zu bestimmen. Für $k=0$ gibt es $(r-1)^{|x|}$ Worte der Länge $|x|$, die sich schon im ersten Element von ω unterscheiden. Für $k \geq 1$ stimmen die ersten k Elemente überein, und die jeweils $k+1$ -ten Elemente von x und ω unterscheiden sich; es bleiben für x an dieser Stelle noch $r-2$ Elemente übrig, ohne dass die Reduktionsbedingung verletzt wird. Für die restlichen $|x|-k-1$ Stellen stehen wieder jeweils $r-1$ Elemente zur Verfügung, also gibt es hier $(r-2)(r-1)^{|x|-k-1}$ Summanden. Und für $k=|x|$ stimmen alle Stellen mit ω überein, hierfür gibt es nur eine Möglichkeit. Wir erhalten also

$$\begin{aligned} &= \frac{1}{r(r-1)^{|x|-1}} \left(\underbrace{(r-1)^{|x|}(r-1)^{-|x|}}_{k=0} + \right. \\ &\quad \left. \sum_{k=1}^{|x|-1} \underbrace{(r-2)(r-1)^{|x|-k-1}(r-1)^{2k-|x|}}_{k=1,\dots,|x|-1} + \underbrace{(r-1)^{|x|}}_{k=|x|} \right) \end{aligned}$$

Die Summe ist geometrisch, also folgt

$$\begin{aligned} &= \frac{1}{r(r-1)^{|x|-1}} \left(1 + (r-1)^{|x|-1} - 1 + (r-1)^{|x|} \right) \\ &= 1. \end{aligned}$$

Damit kann eine isotropische Irrfahrt auf T_r konstruiert werden, indem als erstes rein zufällig ein Endpunkt ω gewählt wird und dann eine h_ω -transformierte Irrfahrt gestartet wird.

h -Transformation

Der Übergangskern P^{h_ω} einer solchen minimalen harmonischen Funktion h_ω ist gegeben durch

$$P^{h_\omega}(x, y) = \begin{cases} \frac{r-1}{r} & \text{falls } d_\omega(x) < d_\omega(y) \\ \frac{1}{r(r-1)} & \text{falls } d_\omega(x) > d_\omega(y) \end{cases}$$

für benachbarte Knoten x, y . Da jeder Knoten x jeweils genau einen Nachbar y besitzt mit $d_\omega(y) = d_\omega(x) + 1$ und genau $r - 1$ Nachbarn y besitzt mit $d_\omega(y) = d_\omega(x) - 1$, verändert sich in jedem Schritt der Irrfahrt der Wert der Funktion d_ω um ± 1 , und der Prozess $(d_\omega(X_n))$ ist selbst eine Irrfahrt auf \mathbb{Z} mit positivem Drift. Daraus folgt insbesondere, dass in diesem Fall eine Transformation mittels h_ω die Irrfahrt (X_n) darauf bedingt, im Punkt ω des Randes R zu enden. Denn je größer der Wert von d_ω in einem Punkt x ist, desto näher liegt x bei ω .

Der reichhaltige harmonische Funktionsraum erlaubt es sowohl die direkte Zerlegung als auch die duale Zerlegung durchzuführen. Genauso wie im Falle der Irrfahrt gilt

$$\inf_n h_\omega(X_n) = 0 \quad \mathbf{P} - \text{f.s.} \quad \text{und} \quad \sup_n h_\omega(X_n) = \infty \quad \mathbf{P}^{h_\omega} - \text{f.s.}$$

Dies bedeutet für die zugehörigen Zerlegungszeiten τ und τ^*

$$\tau < \tau_c < \infty \quad \mathbb{P} - \text{f.s.} \quad \text{und} \quad \tau^* < \tau_c^* < \infty \quad \mathbb{P} - \text{f.s.}$$

Bedingte Verteilung Q_s

In diesem Beispiel sind wir in der glücklichen Situation auch für den bedingten post- τ -Prozess die Verteilung exakt angeben zu können. Die Bestimmung erfolgt kanonisch über die Kerne Q_s und damit mittels der Funktion

$$q_s(x) := \mathbf{P}_x(h_\omega(X_i) \leq s \quad \forall i \in \mathbb{N}_0).$$

Die Kerne Q_s der bedingten P -Kette sind dann gegeben durch

$$Q_s(x, y) = 1/q_s(x) P(x, y) q_s(y)$$

Um die Kerne explizit anzugeben, suchen wir eine geschlossene Darstellung für q_s . Aus der Definition von q_s folgt sofort

$$q_s(x) = \mathbf{P}_x((r-1)^{d_\omega(X_i)} \leq s \quad \forall i \in \mathbb{N}_0).$$

Da sowohl die Kerne P und P^{h_ω} als auch die harmonische Funktion h_ω nur vom Abstand $d_\omega(x) = 2|c(x, \omega)| - |x|$ von x zu ω abhängen, ist auch q_s im Wesentlichen eine Funktion von $d_\omega(x)$, und kann dargestellt als

$$\begin{aligned} q_s(x) &= \mathbf{P}\left(d_\omega(X_i) \leq \frac{\log s}{\log(r-1)} \mid \forall i \in \mathbb{N}_0 \mid d_\omega(X_0) = d_\omega(x)\right) \\ &= \lambda \tilde{q}\left(d_\omega(x) - \lfloor \frac{\log s}{\log(r-1)} \rfloor\right) \end{aligned} \quad (2.5)$$

mit $\tilde{q} : \mathbb{Z} \rightarrow \mathbb{R}_+$ definiert als

$$\tilde{q}(n) := \frac{1}{\lambda} \mathbf{P}(d_\omega(X_i) \leq 0 \mid \forall i \in \mathbb{N}_0 \mid d_\omega(X_0) = n)$$

mit der zusätzlichen Randbedingung $\tilde{q}(0) = 1$, die damit auch den (für die Bestimmung von Q_s unerheblichen) Skalierungsfaktor $\lambda > 0$ festlegt. Als eine direkte Konsequenz aus der Harmonizität von q_s muss \tilde{q} die folgende Rekurrenzgleichung auf \mathbb{Z} erfüllen:

$$\tilde{q}(n) = \frac{1}{r} \tilde{q}(n+1) + \frac{r-1}{r} \tilde{q}(n-1).$$

Lösen dieser Rekurrenz mit den beiden Randbedingungen $\tilde{q}(0) = 1$ und $\tilde{q}(n) = 0$ für $n > 0$ führt zu

$$\tilde{q}(n) = \begin{cases} \frac{(r-1)^n - r + 1}{2-r} & \text{falls } n \leq 0 \\ 0 & \text{sonst,} \end{cases}$$

und die gesuchte Funktion q_s ergibt sich als

$$q_s(x) = \frac{\lambda}{2-r} \left((r-1)^{d_\omega(x) - \lfloor \frac{\log s}{\log(r-1)} \rfloor} - r + 1 \right) \quad \text{für } d_\omega(x) \leq \lfloor \frac{\log s}{\log(r-1)} \rfloor$$

Damit erhalten wir schließlich die Übergangskerne der bedingten Irrfahrt als als eine q_s -Transformation des Kernes P der unbedingten freien Irrfahrt.

Dieses Beispiel gestattet auch die Bestimmung der Übergangskerne der bedingten Irrfahrt mit Hilfe der in Abschnitt 1.5 dargestellten Methode. Die für die Transformation benötigte Funktion q_s aus Gleichung (1.14) ist definiert als

$$\begin{aligned} q_s(x) &= 1 - \frac{h_\omega(x)}{\mathbf{E}_x[h_\omega(X_{\sigma_s}) \mid \sigma_s < \infty]} \\ &= 1 - \frac{(r-1)^{d_\omega(x)}}{\mathbf{E}_x[(r-1)^{d_\omega(X_{\sigma_s})} \mid \sigma_s < \infty]}, \end{aligned}$$

wobei die Stoppzeit σ_s wie üblich definiert ist als $\sigma_s := \inf\{i \geq 0 : h_\omega(X_i) > s\}$. Da $d_\omega(X_i)$ sich in jedem Schritt nur um ± 1 ändert und nur ganzzahlige Werte annehmen kann, gilt bei Start in $X_0 = x$ mit $h_\omega(x) \leq s$ (so dass $\sigma_s > 0$ ist) auf dem Ereignis $\{\sigma_s < \infty\}$

$$(r-1)^{d_\omega(X_{\sigma_s})} > s \quad \Leftrightarrow \quad d_\omega(X_{\sigma_s}) = \lfloor \frac{\log s}{\log(r-1)} \rfloor + 1 \quad \mathbf{P}\text{-fast sicher.}$$

Damit erhalten wir insgesamt für $h_\omega(x) \leq s$, d.h. $d_\omega(x) \leq \lfloor \frac{\log s}{\log(r-1)} \rfloor$,

$$q_s(x) = 1 - (r-1)^{d_\omega(x) - \lfloor \frac{\log s}{\log(r-1)} \rfloor - 1} = \frac{r-2}{r-1} \tilde{q}_s\left(d_\omega(x) - \lfloor \frac{\log s}{\log(r-1)} \rfloor\right),$$

womit nun auch der Skalierungsfaktor $\lambda = \frac{r-2}{r-1}$ aus Gleichung (2.5) bestimmt ist.

Bedingte Verteilung Q_s^*

Auch hier werden die Kerne Q_s^* für die duale Zerlegung bestimmt durch

$$Q_s^*(x, y) = 1/q_s^*(x)P^h(x, y)q_s^*(y) \quad \text{mit} \quad q_s^*(x) := \mathbf{P}_x^{h_\omega}(h_\omega(X_i) \geq s \forall i \in \mathbb{N}_0).$$

Wie im vorigen Unterabschnitt führen ähnliche Überlegungen zur Bestimmung der Kerne Q_s^* zu

$$\begin{aligned} q_s^*(x) &= \mathbf{P}_x^{h_\omega}((r-1)^{d_\omega(X_i)} \geq s \forall i \in \mathbb{N}_0) \\ &= \lambda^* \tilde{q}^* \left(d_\omega(x) - \lceil \frac{\log s}{\log(r-1)} \rceil \right) \end{aligned}$$

mit \tilde{q}^* definiert als

$$\tilde{q}^*(n) := \frac{1}{\lambda^*} \mathbf{P}^{h_\omega}(d_\omega(X_i) \geq 0 \forall i \in \mathbb{N}_0 \mid d_\omega(X_0) = n).$$

Dies führt zu der folgenden Rekurrenzgleichung für \tilde{q}^*

$$\tilde{q}^*(i) = \frac{r-1}{r} \tilde{q}^*(i+1) + \frac{1}{r} \tilde{q}^*(i-1),$$

mit den Randbedingungen $\tilde{q}^*(0) = 1$ und $\tilde{q}^*(i) = 0$ für $i < 0$. Die Lösung ist gegeben durch

$$\tilde{q}^*(i) = \begin{cases} \frac{r-1-(r-1)^{-i}}{r-2} & \text{falls } i \geq 0 \\ 0 & \text{sonst.} \end{cases}$$

Eine Anwendung der Methode aus Abschnitt 1.5 erlaubt ebenfalls eine exakte Bestimmung der Funktion q_s^* für x mit $h_\omega(x) \geq s$ als

$$\begin{aligned} q_s^*(x) &= 1 - \frac{1}{\mathbf{E}_x^{h_\omega} [1/h_\omega(X_{\sigma_s^*}) \mid \sigma_s^* < \infty] h_\omega(x)} \\ &= 1 - (r-1)^{\lceil \frac{\log s}{\log(r-1)} \rceil - d_\omega(x) - 1}, \end{aligned}$$

und damit ist $\lambda^* = \frac{r-2}{r-1}$. Der Kern Q_s^* ist wiederum eine q_s^* -Transformation des Kernes P^h .

Zerlegung

Nun können wir die erste Zerlegung einer freien isotropischen Irrfahrt auf einem homogenen Baum T_r wie folgt konstruieren (siehe die Zerlegung einer Irrfahrt in Abschnitt 2.1 für weitere Details):

1. Sei (\hat{X}_n) eine Irrfahrt auf dem Graphen T_r mit Übergangskern P^{h_ω} , und sei M eine unabhängige exponentiell verteilte Zufallsvariable mit Erwartungswert 1.
2. Setze

$$\tau := \sup\{n \geq 0 : d_\omega(X_n) \leq M, d_\omega(X_m) < d_\omega(X_n) \forall m < n\}.$$

3. Sei (\check{X}_n) eine Irrfahrt mit Übergangskern $Q_{h_\omega(\hat{X}_\tau)}$ und mit Startposition $\check{X}_0 = \hat{X}_\tau$, und definiere

$$\bar{X}_n := \begin{cases} \hat{X}_n & n < \tau \\ \check{X}_{n-\tau} & n \geq \tau. \end{cases}$$

Dann ist der Prozess (\bar{X}_n) eine isotropische Irrfahrt auf T_τ .

Die duale Zerlegung einer P^{h_ω} -Kette startet als erstes eine isotropische Irrfahrt auf T_τ und fügt dann eine bedingte P^{h_ω} -Kette mit Kern Q_s^* an.

2.5 Beispiel: Pólyas Urnenschema

Im Pólyas Urnenexperiment mit Kugeln zweier verschiedener Farben wird jede gezogene Kugel mit einer weiteren Kugel der gleichen Farbe zurückgelegt. Hierdurch ändert sich die Besetzung der Urne ständig. Zur Beschreibung dieses Prozesses betrachten wir die Markov-Kette $X_n = (R_n, T_n)$, $n \geq 0$, mit Zuständen $x = (r, t) \in \mathbb{N}^2$, $0 < r < t$, wobei r die Anzahl der roten Kugeln und t die Gesamtanzahl Kugeln in der Urne bezeichnet. Die Übergangswahrscheinlichkeiten sind gegeben durch

$$P(x, y) = \begin{cases} \frac{r}{t} & \text{für } x = (r, t), y = (r+1, t+1), \\ 1 - \frac{r}{t} & \text{für } x = (r, t), y = (r, t+1). \end{cases}$$

Es ist bekannt, dass das Verhältnis der Kugeln einer Farbe zur Gesamtanzahl Kugeln approximativ uniform verteilt ist, d.h. $R_n/T_n \rightarrow U$ in Verteilung, wobei U uniform in $[0, 1]$ verteilt ist.

Pólyas Urnenschema, Teil 1

Für gegebene Werte $0 < p < 1$ und $q = 1 - p$ sind die Funktionen

$$h_p(r, t) := (t-1) \binom{t-2}{r-1} p^{r-1} q^{t-r-1}$$

harmonisch, und es gilt offenbar $\sup h_p = \infty$. Es ist nicht schwer direkt zu sehen, dass Bedingung (1.9) erfüllt ist. Dies folgt allerdings auch aus Satz 10, da die Funktionen h_p minimal sind (siehe [BK64]). Folglich ist die Zerlegungszeit τ in der assoziierten Pfadzerlegung durch die Überkreuzungszeit τ_c erkennbar.

Es bleibt die Frage nach einer Interpretation einer solchen Pfadzerlegung. Eine einfache Rechnung zeigt, dass $P^{h_p}(x, y) = p$ bzw. q gilt, dass also die Anzahl der roten Kugeln des Anfangsstücks einer Irrfahrt (R_n) mit dem Maß \mathbf{P}^{h_p} entspricht, also

$$\mathbf{P}^{h_p}(R_{n+1} = x+1 \mid R_n = x) = p \quad \text{und} \quad \mathbf{P}^{h_p}(R_{n+1} = x \mid R_n = x) = q$$

entspricht. Damit folgt mit dem starken Gesetz der großen Zahlen

$$\lim_{n \rightarrow \infty} R_n/n = p \quad \mathbf{P}^{h_p} - \text{f.s.}$$

Eine aussagekräftige Interpretation der Zerlegungszeit selbst scheint in diesem Beispiel aufgrund der komplizierten Struktur von h_p allerdings schwierig.

Pólyas Urnenschema, Teil 2

Die Pólya-Urne ist allerdings noch reichhaltiger, denn auch die Funktion

$$h(r, t) := \frac{r}{t}$$

ist harmonisch. In diesem Falle gilt $P^h(x, y) = \frac{r+1}{t+1}$ bzw. $1 - \frac{r+1}{t+1}$, folglich können die verschobenen Zufallsvariablen $X'_n := X_n + (1, 1)$, $n \geq 0$ wiederum als eine Pólya-Urne interpretiert werden mit einer zusätzlichen roten Kugel am Anfang. Die Zerlegungszeit τ der zugehörige Pfadzerlegung entspricht dann gerade dem Zeitpunkt, beim dem der relative Anteil roter Kugeln maximal ist. Allerdings ist die Zerlegungszeit τ nicht durch eine Stoppzeit erkennbar, wie im folgenden gezeigt werden soll.

Das beschränkte Martingal $h(X_n) = R_n/T_n$ konvergiert fast sicher, und gegeben den Grenzwert W sind die Zufallsvariablen $Z_n := R_n - R_{n-1}$, $n \geq 1$ unabhängige Bernoulli Variablen mit Erfolgswahrscheinlichkeit W . Diese Aussagen gelten sowohl bezüglich des Maßes \mathbf{P}_x als auch bezüglich \mathbf{P}_x^h .

Zunächst zeigen wir, dass die Zerlegungszeit τ \mathbb{P} -fast sicher endlich ist. Betrachten wir das Anfangsstück (\hat{X}_n) der Konstruktion mit dem Übergangskern P^h . Wir führen die Notation

$$\hat{X}_n = (\hat{R}_n, \hat{T}_n) \quad \hat{Z}_n := \hat{R}_n - \hat{R}_{n-1} \quad \hat{W} := \lim_{n \rightarrow \infty} h(\hat{X}_n) \quad \mathbb{P} - \text{f.s.}$$

ein. Aus $\hat{R}_n = \hat{R}_0 + \sum_{i=1}^n \hat{Z}_i$ und aus $\hat{T}_n = n + \hat{T}_0$ folgt dann

$$\left\{ \sup_n h(\hat{X}_n) \leq \hat{W} \right\} = \left\{ \sup_n \sum_{i=1}^n (\hat{Z}_i - \hat{W}) \leq \hat{W} \hat{T}_0 - \hat{R}_0 \right\}.$$

Durch die Fluktuation der Summe der unabhängigen, identisch verteilten und zentrierten Zufallsvariablen $(\hat{Z}_i - \hat{W})$ hat das Ereignis $\{\sup_n h(\hat{X}_n) \leq \hat{W}\}$ Wahrscheinlichkeit Null, und \hat{W} ist somit \mathbb{P} -fast sicher kleiner als das globale Maximum von $h(\hat{X}_n)$. Damit ist der Zeitpunkt

$$T := \inf\{n \geq 0 : h(\hat{X}_n) = \sup_m h(\hat{X}_m)\},$$

des ersten Erreichens des globalen Maximums \mathbb{P} -fast sicher endlich, und dies impliziert, dass die Zerlegungszeit τ ebenfalls fast sicher endlich ist. Allerdings ist $\sup_n h(\hat{X}_n) \leq 1$, so dass das Ereignis $\{\tau_c = \infty\} = \{\sup_n h(\hat{X}_n) \leq Y\}$ eine strikt positive Wahrscheinlichkeit besitzt, deshalb kann τ von τ_c nicht erkannt werden.

Tatsächlich existiert keine Stoppzeit, die τ erkennen kann, wie im Folgenden gezeigt wird. Für jede \mathbb{P} -fast sicher endliche Stoppzeit σ gilt

$$\{T \leq \sigma\} = \left\{ \sup_{j \geq 0} h(\hat{X}_{\sigma+j}) \leq M \right\} \mathbb{P} - \text{f.s.} \quad \text{mit} \quad M := \max_{n \leq \sigma} h(\hat{X}_n).$$

Aus der starken Markov-Eigenschaft erhalten wir damit

$$\mathbb{P}(T \leq \sigma) = \mathbb{E} \psi(\hat{X}_\sigma, M) \quad \text{mit} \quad \psi(x, m) := \mathbf{P}_x^h(\sup_{j \geq 0} h(X_j) \leq m).$$

Da $\psi(x, m) < 1$ ist für $m < 1$, und da $M < 1$ \mathbb{P} -fast sicher ist, folgt $\mathbb{P}(T \leq \sigma) < 1$ und allgemeiner $\mathbb{P}(T \leq \sigma \mid Y) < 1$ \mathbb{P} -fast sicher (σ darf schließlich von Y abhängig sein). Da

$\tau = T$ auf dem Ereignis $\{Y \geq 1\}$ gilt, erhalten wir

$$\begin{aligned} \mathbb{P}(\tau \leq \sigma) &\leq \mathbb{P}(Y \geq 1, \tau \leq \sigma) \\ &= \mathbb{P}(Y \geq 1, T \leq \sigma) \\ &= \mathbb{E}(\mathbb{P}(T \leq \sigma | Y); Y \geq 1) \\ &< \mathbb{P}(Y \geq 1) < 1. \end{aligned}$$

Da $\tau < \infty$ \mathbb{P} -fast sicher gilt, kann die Stoppzeit σ somit die Zerlegungszeit τ nicht erkennen.

2.6 Beispiel: Verzweigungsprozesse

In diesem Abschnitt beschäftigen wir uns mit einem einfachen Modell für zeitdiskrete Verzweigungsprozesse im Raum.

Modellbeschreibung

Wir betrachten ein relativ einfaches Modell eines Verzweigungsprozesses, das aber schon interessante Ergebnisse liefert. Der Grundraum, in dem sich endlich viele Partikel bewegen, wird als diskret angenommen und mit S bezeichnet. Als eigentlichen Zustandsraum des Prozesses ist es zweckmäßig, den Raum \mathcal{M}_S der endlichen Zählmaße über S zu verwenden. Eine Zufallsvariable \mathfrak{X} mit Werten in \mathcal{M}_S kann dann als endliche Summe

$$\mathfrak{X} = \sum_{i=1}^{|\mathfrak{X}|} \dot{X}_i$$

dargestellt werden, wobei $|\mathfrak{X}| := \mathfrak{X}(S)$ die Anzahl der Partikel und $\dot{X}_i := \delta_{X_i}$ eine Punktmasse im (zufälligen) Punkt X_i bezeichnet, die genau einem Partikel der Konfiguration \mathfrak{X} entspricht. Es bezeichnet also \dot{X}_i das Partikel, und X_i den Ort des Partikels. Die Ortspunkte X_i müssen nicht notwendigerweise verschieden sein, d.h. es können sich durchaus mehrere Partikel an einem Ort befinden.

Der zeithomogene Verzweigungsprozess folgt nun in jedem Schritt dem folgenden Mechanismus:

1. Sei $\mathfrak{X}_n = \dot{X}_1 + \dots + \dot{X}_k$ der Zustand des Prozesses zum Zeitpunkt n . Als erstes erzeugt jedes Partikel \dot{X}_i unabhängig von allen anderen Partikeln eine zufällige Anzahl D_i Nachkommen $\dot{Y}_{D_1+\dots+D_{i-1}+1}, \dots, \dot{Y}_{D_1+\dots+D_i}$ gemäß einer möglicherweise vom Ort X_i abhängigen Verteilung $(p_d)_{d \geq 0}$.
2. Die Partikel \dot{X}_i selbst sterben alle, es verbleiben also nur noch jeweils die D_i Nachkommen der einzelnen Partikel am gleichen Ort X_i wie jeweils ihre Vorfahren.
3. Alle neu geborenen Partikel $\dot{Y}_1, \dots, \dot{Y}_{D_1+\dots+D_k}$ bewegen sich jeweils unabhängig im Raum S gemäß eines Bewegungskernes P .
4. Der neue Zustand des Verzweigungssystems ist gegeben durch

$$\mathfrak{X}_{n+1} := \sum_{j=1}^{D_1+\dots+D_k} \dot{Y}_j$$

Im nächsten Zeitschritt startet der Mechanismus wieder von vorne mit der neuen Konfiguration \mathfrak{X}_{n+1} .

Um die Verwandtschaftsverhältnisse der Partikel besser beschreiben zu können, wird die folgende Vaterschaftsfunktion eingeführt. Mit den Bezeichnungen von oben kann der Index des Vaters eines Partikels \dot{Y}_j ermittelt werden durch

$$\phi(j) := \min\{i : D_1 + \dots + D_i \geq j\},$$

so dass $\phi(j)$ gerade den Index i des Vaters \dot{X}_i von Partikel \dot{Y}_j ergibt. Eine weitere eng verwandte Funktion $\sigma : \mathbb{N} \rightarrow \mathbb{N}$ ist definiert als

$$\sigma(i) := D_1 + \dots + D_{i-1} + 1,$$

und gibt gerade den Index j des ersten Nachkommens von Partikel \dot{X}_i an. Damit gilt dann $\phi(\sigma(i)) = i$ für alle i .

Sobald man den Verzweigungsprozess nur von außen beobachtet, ist nicht mehr klar, wie die Verwandtschaftsverhältnisse der Partikel aussehen, denn gegeben $\mathfrak{X}_n = \dot{X}_1 + \dots + \dot{X}_k$ und $\mathfrak{X}_{n+1} = \dot{Y}_1 + \dots + \dot{Y}_l$ können wir im Allgemeinen keine Aussage mehr darüber treffen, welches Partikel \dot{X}_i der Vorfahre eines Partikels \dot{Y}_j ist. Bei der Berechnung der Übergangswahrscheinlichkeiten müssen wir vielmehr alle möglichen Verwandtschaftsverhältnisse berücksichtigen. Hierzu sei für ein (möglicherweise zufälliges) Zählmaß $\mathfrak{X} = \dot{X}_1 + \dots + \dot{X}_k$ eine induzierte Relation $\sim_{\mathfrak{X}}$ definiert auf der Menge $S_{|\mathfrak{X}|}$ der Permutationen $\pi : \{1, \dots, |\mathfrak{X}|\} \rightarrow \{1, \dots, |\mathfrak{X}|\}$ als

$$\pi \sim_{\mathfrak{X}} \pi' \quad :\Leftrightarrow \quad \dot{x}_{\pi(i)} = \dot{x}_{\pi'(i)} \quad \forall i \in \{1, \dots, |\mathfrak{X}|\}.$$

Die Menge aller daraus erzeugten Äquivalenzklassen $[\pi]_{\mathfrak{X}}$ sei mit $V[\mathfrak{X}]$ bezeichnet.

Im folgenden bezeichne \mathbf{P} das Wahrscheinlichkeitsmaß eines Verzweigungsprozesses $(\mathfrak{X}_n)_{n \in \mathbb{N}}$ mit Zustandsraum \mathcal{M}_S auf einer diskreten Menge S mit Bewegungskern P . Bedingt auf ein gegebenes Tupel (d_1, \dots, d_k) von Anzahlen an Nachkommen der einzelnen Partikel in $\mathfrak{X}_n = \dot{X}_1 + \dots + \dot{X}_k$, können wir nun einfach die Übergangswahrscheinlichkeit von \mathfrak{X}_n nach $\mathfrak{X}_{n+1} = \dot{Y}_1 + \dots + \dot{Y}_l$ bestimmen als

$$\begin{aligned} \mathbf{P}(\mathfrak{X}_{n+1} = \dot{y}_1 + \dots + \dot{y}_l \mid \mathfrak{X}_n = \dot{x}_1 + \dots + \dot{x}_k, D_1 = d_1, \dots, D_k = d_k) \\ = \sum_{[\pi] \in V[\dot{y}_1 + \dots + \dot{y}_l]} P(x_{\phi(1)}, y_{\pi(1)}) \cdots P(x_{\phi(l)}, y_{\pi(l)}). \end{aligned}$$

Damit ist nun auch klar, wie die Übergangswahrscheinlichkeiten des gesamten Verzweigungsprozesses $(\mathfrak{X}_n)_{n \in \mathbb{N}}$ mit ortsunabhängigen Geburtsraten $(p_d)_{d \in \mathbb{N}}$ und Bewegungskern P berechnet werden können:

$$\begin{aligned} \mathbf{P}(\mathfrak{X}_{n+1} = \dot{y}_1 + \dots + \dot{y}_l \mid \mathfrak{X}_n = \dot{x}_1 + \dots + \dot{x}_k) \\ = \sum_{d_1 + \dots + d_k = l} \mathbf{P}(D_1 = d_1, \dots, D_k = d_k) \sum_{[\pi] \in V[\dot{y}_1 + \dots + \dot{y}_l]} P(x_{\phi(1)}, y_{\pi(1)}) \cdots P(x_{\phi(l)}, y_{\pi(l)}) \\ = \sum_{d_1 + \dots + d_k = l} p_{d_1} \cdots p_{d_k} \sum_{[\pi] \in V} P(x_{\phi(1)}, y_{\pi(1)}) \cdots P(x_{\phi(l)}, y_{\pi(l)}). \end{aligned} \tag{2.6}$$

Harmonische Funktionen

Wir wollen nun zwei unterschiedliche Konstruktionen harmonischer Funktionen studieren, die jeweils auf einer andere Teilklasse von Verzweigungsprozessen wie oben definiert sind. Die erste Klasse harmonischer Funktionen benötigt ortsunabhängige Kinderzahlverteilungen, während die zweite nur für kritische Verzweigungsprozesse definiert ist. Beide Klassen führen auch zu unterschiedlichen Zerlegungen, insbesondere die h -transformierten Prozesse besitzen verschiedene Gestalten. Siehe auch [BK04] zu der folgenden, ersten Konstruktion einer harmonischen Funktion und zu deren assoziierten Maßwechsel.

Satz 15 *Sei (\mathfrak{X}_n) ein diskreter Verzweigungsprozess wie oben mit Verteilung \mathbf{P} und mit ortsunabhängigen Kinderverteilung (p_d) mit Sterbewahrscheinlichkeit $p_0 = 0$. Sei weiter h eine nicht-negative harmonische Funktion bezüglich des Bewegungskernes P . Dann ist die Funktion $\tilde{h} : \mathcal{M}(S) \rightarrow \mathbb{R}_+$ definiert durch*

$$\tilde{h}\left(\sum_{i=1}^k \dot{x}_i\right) := \frac{1}{k} \sum_{i=1}^k h(x_i)$$

harmonisch bezüglich der in Gleichung (2.6) beschriebenen Übergangsdynamik des Prozesses.

Beweis Als erstes sei bemerkt, dass die Voraussetzung $p_0 = 0$ wirklich notwendig ist, da die Funktion \tilde{h} nicht wohldefiniert ist auf einem Maß der Masse 0.

Bezeichne \mathcal{P} den Übergangskern des gesamten Verzweigungsprozesses (\mathfrak{X}_n) . In der folgenden Rechnung sei \mathfrak{X} die Startgeneration gegeben durch $\mathfrak{X} = \dot{X}_1 + \dots + \dot{X}_{|\mathfrak{X}|}$, und jedes Partikel \dot{X}_i habe jeweils die direkten Nachkommen $\dot{Y}_{i,1}, \dots, \dot{Y}_{i,D_i}$. Die vollständige nächste Generation werde mit $\mathfrak{Y} := \sum \dot{Y}_{i,j}$ bezeichnet. Damit gilt

$$\begin{aligned} \mathcal{P}\tilde{h}(\mathfrak{X}) &= \mathbf{E}[\tilde{h}(\mathfrak{Y}) \mid \mathfrak{X}] \\ &= \mathbf{E}[\mathbf{E}[\tilde{h}(\mathfrak{Y}) \mid \mathfrak{X}, D_1, \dots, D_{|\mathfrak{X}|}] \mid \mathfrak{X}] \end{aligned}$$

Aus der Definition von \tilde{h} folgt:

$$\begin{aligned} &= \mathbf{E}\left[\frac{1}{D_1 + \dots + D_{|\mathfrak{X}|}} \mathbf{E}\left[\sum_{i=1}^{|\mathfrak{X}|} \sum_{j=1}^{D_i} h(Y_{i,j}) \mid \mathfrak{X}, D_1, \dots, D_{|\mathfrak{X}|}\right] \mid \mathfrak{X}\right] \\ &= \mathbf{E}\left[\frac{1}{D_1 + \dots + D_{|\mathfrak{X}|}} \sum_{i,j} \mathbf{E}[h(Y_{i,j}) \mid \mathfrak{X}] \mid \mathfrak{X}\right] \end{aligned}$$

Es gilt $\mathbf{E}[h(Y_{i,j}) \mid X_i] = Ph(X_i)$ unabhängig von j , und h ist harmonisch bezüglich P , d.h. $Ph(X_i) = h(X_i)$, also erhalten wir

$$\begin{aligned} &= \mathbf{E}\left[\frac{1}{D_1 + \dots + D_{|\mathfrak{X}|}} \sum_{i=1}^{|\mathfrak{X}|} D_i \cdot h(X_i) \mid \mathfrak{X}\right] \\ &= \sum_{i=1}^{|\mathfrak{X}|} h(X_i) \mathbf{E}\left[\frac{D_i}{D_1 + \dots + D_{|\mathfrak{X}|}} \mid \mathfrak{X}\right] \end{aligned}$$

Da die Kinderzahlen D_i unabhängige Zufallsvariablen mit identischer (da ortsunabhängiger) Verteilung sind, und da diese außerdem unabhängig von X sind, erhalten wir

$$= \frac{1}{|\mathfrak{X}|} \sum_{i=1}^{|\mathfrak{X}|} h(X_i) = \tilde{h}(\mathfrak{X})$$

■

Satz 16 Sei (\mathfrak{X}_n) ein diskreter Verzweigungsprozess wie oben mit Verteilung \mathbf{P} und erwarteter Nachkommenzahl 1 pro Partikel. Sei weiter h eine positive harmonische Funktion bezüglich des Bewegungskernes P . Dann ist die Funktion $\tilde{h} : \mathcal{M}(S) \rightarrow \mathbb{R}_+$ definiert durch

$$\tilde{h}\left(\sum_{i=1}^k \dot{x}_i\right) := \sum_{i=1}^k h(x_i)$$

harmonisch bezüglich der in Gleichung (2.6) beschriebenen Übergangsdynamik des Prozesses.

Beweis Bezeichne \mathcal{P} den Übergangskern des gesamten Verzweigungsprozesses (\mathfrak{X}_n) mit Bewegungskern P und einer möglicherweise ortsabhängigen Kinderzahlverteilung mit Erwartungswert 1. In der folgenden Rechnung sei X die Startgeneration gegeben durch $\mathfrak{X} = \dot{X}_1 + \dots + \dot{X}_{|\mathfrak{X}|}$, und jedes Partikel \dot{X}_i habe jeweils die direkten Nachkommen $\dot{Y}_{i,1}, \dots, \dot{Y}_{i,D_i}$. Die vollständige nächste Generation werde mit $\mathfrak{Y} := \sum \dot{Y}_{i,j}$ bezeichnet. Damit gilt

$$\begin{aligned} \mathcal{P}\tilde{h}(\mathfrak{X}) &= \mathbf{E}[\tilde{h}(\mathfrak{Y}) \mid \mathfrak{X}] \\ &= \mathbf{E}[\mathbf{E}[\tilde{h}(\mathfrak{Y}) \mid \mathfrak{X}, D_1, \dots, D_{|\mathfrak{X}|}] \mid \mathfrak{X}] \end{aligned}$$

Die Definition von \tilde{h} liefert:

$$\begin{aligned} &= \mathbf{E}\left[\mathbf{E}\left[\sum_{i=1}^{|\mathfrak{X}|} \sum_{j=1}^{D_i} h(Y_{i,j}) \mid \mathfrak{X}, D_1, \dots, D_{|\mathfrak{X}|}\right] \mid \mathfrak{X}\right] \\ &= \mathbf{E}\left[\sum_{i,j} \mathbf{E}[h(Y_{i,j}) \mid \mathfrak{X}] \mid \mathfrak{X}\right] \end{aligned}$$

Es gilt $\mathbf{E}[h(Y_{i,j}) \mid X_i] = Ph(X_i)$ unabhängig von j , und h ist P -harmonisch, d.h. $Ph(X_i) = h(X_i)$, damit erhalten wir

$$= \mathbf{E}\left[\sum_{i=1}^{|\mathfrak{X}|} D_i \cdot h(X_i) \mid \mathfrak{X}\right]$$

Da die Kinderzahlen D_i unabhängig von \mathfrak{X} sind mit Erwartungswert $\mathbf{E}[D_i] = 1$, erhalten wir

$$\begin{aligned} &= \sum_{i=1}^{|\mathfrak{X}|} \mathbf{E}[D_i] h(X_i) \\ &= \sum_{i=1}^{|\mathfrak{X}|} h(X_i) = \tilde{h}(\mathfrak{X}) \end{aligned}$$

Man beachte, dass wir im letzten Schritt nur $\mathbf{E}[D_i] = 1$ benötigen – es wird nicht vorausgesetzt, dass die D_i identisch verteilt sind, d.h. die Kinderzahlverteilungen können ortsabhängig sein, es ist sogar eine Kopplung der D_i untereinander erlaubt. ■

h -Transformation

Beide Konstruktionen harmonischer Funktionen führen zu unterschiedlichen \tilde{h} -Transformationen. Für beide Klassen kann wieder eine alternative Konstruktion des jeweiligen Prozesses angegeben werden, die den Unterschied verdeutlicht.

Satz 17 (Alternative Konstruktion) *Bezeichne (\mathfrak{X}_n) einen Verzweigungsprozess mit Verteilung \mathbf{P} im Raum der endlichen Zählmaße über eine diskrete Menge S . Der Prozess besitze eine homogene Kinderzahlverteilung mit Sterbewahrscheinlichkeit 0, und P bezeichne den Bewegungskern. Weiter sei h eine positive P -harmonische Funktion und sei \tilde{h} definiert wie in Satz 15 als*

$$\tilde{h}\left(\sum_{i=1}^k \dot{x}_i\right) := \frac{1}{k} \sum_{i=1}^k h(x_i).$$

Dann kann jeweils ein Übergangsschritt des Prozesses (\mathfrak{X}_n) unter der \tilde{h} -transformierten Verteilung $\mathbf{P}^{\tilde{h}}$ wie folgt konstruiert werden:

1. Sei $\mathfrak{X}_n := \dot{X}_1 + \dots + \dot{X}_k$ der Prozess zum Zeitpunkt n . Wähle ein Partikel mit Index R unter den k Partikeln zufällig aus, so dass Partikel \dot{X}_r mit Wahrscheinlichkeit proportional zu $h(X_r)$ gezogen wird, d.h.

$$\mathbf{P}^{\tilde{h}}(R = r \mid \mathfrak{X}_n = \dot{x}_1 + \dots + \dot{x}_k) = \frac{h(x_r)}{\sum h(x_i)}.$$

2. Gegeben den Wert r von R ermittle die Anzahl der Nachkommen D_1, \dots, D_k gemäß der folgenden bedingten Verteilung

$$\begin{aligned} \mathbf{P}^{\tilde{h}}(D_1 = d_1, \dots, D_k = d_k \mid R = r, |\mathfrak{X}_n| = k) \\ = \frac{k \cdot d_r}{d_1 + \dots + d_k} \cdot \mathbf{P}(D_1 = d_1, \dots, D_k = d_k), \end{aligned} \quad (2.7)$$

und generiere dann die Nachkommenschaft $\dot{Y}_1, \dots, \dot{Y}_{D_1+\dots+D_k}$.

3. Bewege alle Partikel \dot{Y}_j bis auf den ersten Nachfahre $\dot{Y}_{\sigma(R)}$ von \dot{X}_R gemäß des Bewegungskernes P , und bewege den ersten Nachfahre $\dot{Y}_{\sigma(R)}$ von \dot{X}_R gemäß des h -transformierten Bewegungskernes P^h .
4. Die nächste Generation \mathfrak{X}_{n+1} ist nun gegeben durch die neuen und schon bewegten Partikel $\mathfrak{X}_{n+1} := \dot{Y}_1 + \dots + \dot{Y}_{D_1+\dots+D_k}$.

Verfahre im nächsten Zeitschritt genauso. Der daraus entstandene Prozess besitzt die gewünschte Verteilung $\mathbf{P}^{\tilde{h}}$.

Beweis

$$\begin{aligned} \mathbf{P}^{\tilde{h}}(\mathfrak{X}_{n+1} = \dot{y}_1 + \dots + \dot{y}_l \mid \mathfrak{X}_n = \dot{x}_1 + \dots + \dot{x}_k) \\ = \frac{1}{\tilde{h}(\dot{x}_1 + \dots + \dot{x}_k)} \mathbf{P}(\mathfrak{X}_{n+1} = \dot{y}_1 + \dots + \dot{y}_l \mid \mathfrak{X}_n = \dot{x}_1 + \dots + \dot{x}_k) \tilde{h}(\dot{y}_1 + \dots + \dot{y}_l) \end{aligned}$$

Die Definition der Funktion \tilde{h} und die explizite Konstruktion des Übergangskernes von \mathbf{P} aus Gleichung (2.6) liefert

$$\begin{aligned} &= \frac{k}{\sum_{i=1}^k h(x_i)} \sum_{d_1+\dots+d_k=l} \mathbf{P}(D_1 = d_1, \dots, D_k = d_k) \\ &\quad \cdot \sum_{[\pi] \in V[\tilde{y}_1+\dots+\tilde{y}_l]} P(x_{\phi(1)}, y_{\pi(1)}) \cdots P(x_{\phi(l)}, y_{\pi(l)}) \cdot \frac{1}{l} \sum_{j=1}^l h(y_j) \end{aligned}$$

Die Summe $\sum h(y_j)$ ist unter Vertauschung der Summanden invariant, also folgt

$$\begin{aligned} &= \frac{1}{\sum h(x_i)} \sum_{d_1 \dots d_k} \frac{k}{l} \mathbf{P}(D_1 = d_1, \dots, D_k = d_k) \\ &\quad \cdot \sum_{j=1}^l \sum_{[\pi] \in V} h(y_{\pi(j)}) P(x_{\phi(1)}, y_{\pi(1)}) \cdots P(x_{\phi(l)}, y_{\pi(l)}) \end{aligned}$$

Nun setzen wir $h(y_{\pi(j)}) P(x_{\phi(j)}, y_{\pi(j)}) = h(x_{\phi(j)}) P^h(x_{\phi(j)}, y_{\pi(j)})$ ein, und erhalten somit

$$\begin{aligned} &= \frac{1}{\sum h(x_i)} \sum_{d_1 \dots d_k} \frac{k}{l} \mathbf{P}(D_1 = d_1, \dots, D_k = d_k) \\ &\quad \cdot \sum_{j=1}^l h(x_{\phi(j)}) \sum_{[\pi] \in V} P(x_{\phi(1)}, y_{\pi(1)}) \cdots P^h(x_{\phi(j)}, y_{\pi(j)}) \cdots P(x_{\phi(l)}, y_{\pi(l)}) \end{aligned}$$

Anstatt über die Nachkommen zu summieren, können wir auch über deren Vorfahren summieren, wobei jedes Partikel \dot{x}_r jeweils D_r Nachfahren besitzt, und somit erhalten wir

$$\begin{aligned} &= \frac{1}{\sum h(x_i)} \sum_{d_1 \dots d_k} \frac{k}{l} \mathbf{P}(D_1 = d_1, \dots, D_k = d_k) \\ &\quad \cdot \sum_{r=1}^k d_r h(x_r) \sum_{[\pi] \in V} P(x_{\phi(1)}, y_{\pi(1)}) \cdots P^h(x_{\phi(\sigma(r))}, y_{\pi(\sigma(r))}) \cdots P(x_{\phi(l)}, y_{\pi(l)}) \end{aligned}$$

Die Gleichung $\phi(\sigma(r)) = r$ und einfache Vertauschung liefern weiter

$$\begin{aligned} &= \sum_{r=1}^k \frac{h(x_r)}{\sum h(x_i)} \sum_{d_1 \dots d_k} \frac{k \cdot d_r}{l} \mathbf{P}(D_1 = d_1, \dots, D_k = d_k) \\ &\quad \cdot \sum_{[\pi] \in V} P(x_{\phi(1)}, y_{\pi(1)}) \cdots P^h(x_r, y_{\pi(\sigma(r))}) \cdots P(x_{\phi(l)}, y_{\pi(l)}) \end{aligned}$$

Und schließlich ergibt dies mit den im Satz für R und D_i angegebenen Verteilungen unter Berücksichtigung von $d_1 + \dots + d_k = l$

$$\begin{aligned} &= \sum_{r=1}^k \mathbf{P}^{\tilde{h}}(R = r \mid \mathfrak{X}_n = \dot{x}_1 + \dots + \dot{x}_k) \\ &\quad \cdot \sum_{d_1, \dots, d_k} \mathbf{P}^{\tilde{h}}(D_1 = d_1, \dots, D_k = d_k \mid R = r, |\mathfrak{X}_n| = k) \\ &\quad \cdot \sum_{[\pi] \in V} P(x_{\phi(1)}, y_{\pi(1)}) \cdots P^h(x_r, y_{\pi(\sigma(r))}) \cdots P(x_{\phi(l)}, y_{\pi(l)}) \end{aligned}$$

Dieser Ausdruck stellt gerade die maßtheoretische Formulierung der im Satz angegebenen Konstruktion dar. Eine einfache Rechnung zeigt außerdem

$$\mathbf{E} \frac{k \cdot D_r}{D_1 + \dots + D_k} = 1,$$

also ist die in Gleichung (2.7) angegebene Verteilung tatsächlich eine Wahrscheinlichkeitsverteilung mit Gesamtmasse 1. \blacksquare

Bemerkung Die erwartete Anzahl an Nachkommen bleibt unter der Verteilung $\mathbf{P}^{\tilde{h}}$ unverändert gegenüber der Verteilung \mathbf{P} , wie man in der folgenden einfachen Rechnung sieht:

$$\begin{aligned} \mathbf{E}^{\tilde{h}}[|\mathfrak{X}_{n+1}| \mid \mathfrak{X}_n] &= \mathbf{E}^{\tilde{h}}[D_1 + \dots + D_{|\mathfrak{X}_n}| \mid \mathfrak{X}_n] \\ &= \mathbf{E}^{\tilde{h}}[\mathbf{E}^{\tilde{h}}[D_1 + \dots + D_{|\mathfrak{X}_n}| \mid \mathfrak{X}_n, R] \mid \mathfrak{X}_n] \end{aligned}$$

Nach Definition der bedingten Verteilung von $D_1 + \dots + D_{|\mathfrak{X}_n}|$ in Gleichung (2.7) folgt

$$= \mathbf{E}^{\tilde{h}}\left[\mathbf{E}\left[(D_1 + \dots + D_{|\mathfrak{X}_n|}) \frac{D_R \cdot |\mathfrak{X}_n|}{D_1 + \dots + D_{|\mathfrak{X}_n|}} \mid \mathfrak{X}_n, R\right] \mid \mathfrak{X}_n\right]$$

Da unter \mathbf{P} alle D_i unabhängig und identisch verteilt sind, folgt

$$\begin{aligned} &= \mathbf{E}^{\tilde{h}}\left[\mathbf{E}[D_1 \cdot |\mathfrak{X}_n| \mid \mathfrak{X}_n, R] \mid \mathfrak{X}_n\right] \\ &= |\mathfrak{X}_n| \cdot \mathbf{E}[D_1] \end{aligned}$$

Da unter \mathbf{P} jedes Partikel die gleiche Nachkommenverteilung besitzt, ist dies gerade die erwartete Anzahl Partikel in der $(n+1)$ -ten Generation gegeben die n -te Generation \mathfrak{X}_n , also

$$= \mathbf{E}[|\mathfrak{X}_{n+1}| \mid \mathfrak{X}_n]$$

\square

Satz 18 (Alternative Konstruktion) *Bezeichne (\mathfrak{X}_n) einen Verzweigungsprozess mit Verteilung \mathbf{P} im Raum der endlichen Zählmaße über eine diskrete Menge S . Der Prozess besitze eine möglicherweise räumlich inhomogene Kinderzahlverteilung mit Erwartungswert 1, und*

P bezeichne den Bewegungskern. Weiter sei h eine positive P -harmonische Funktion und sei \tilde{h} definiert wie in Satz 16 als

$$\tilde{h}\left(\sum_{i=1}^k \dot{x}_i\right) := \sum_{i=1}^k h(x_i).$$

Dann kann jeweils ein Übergangsschritt des Prozesses (\mathfrak{X}_n) unter der \tilde{h} -transformierten Verteilung $\mathbf{P}^{\tilde{h}}$ wie folgt konstruiert werden:

1. Sei $\mathfrak{X}_n := \dot{X}_1 + \dots + \dot{X}_k$ der Prozess zum Zeitpunkt n . Wähle ein Partikel mit Index R unter den k Partikeln zufällig aus, so dass Partikel \dot{X}_r mit Wahrscheinlichkeit proportional zu $h(x_r)$ gezogen wird, d.h.

$$\mathbf{P}^{\tilde{h}}(R = r \mid \mathfrak{X}_n = \dot{x}_1 + \dots + \dot{x}_k) = \frac{h(x_r)}{\sum h(x_i)}.$$

2. Gegeben den Wert r von R ermittle die Anzahl von Nachkommen D_1, \dots, D_k gemäß der folgenden bedingten Verteilung

$$\begin{aligned} \mathbf{P}^{\tilde{h}}(D_1 = d_1, \dots, D_k = d_k \mid R = r, |\mathfrak{X}_n| = k) \\ = d_r \cdot \mathbf{P}(D_1 = d_1, \dots, D_k = d_k), \end{aligned} \quad (2.8)$$

d.h. die Nachkommen werden immer noch unabhängig erzeugt, nur das Partikel mit Index R erzeugt eine größenverzerrte Anzahl Nachkommen. Generiere dann die Nachkommenschaft $\dot{Y}_1, \dots, \dot{Y}_{D_1+\dots+D_k}$.

3. Bewege alle Partikel \dot{Y}_j bis auf den ersten Nachfahre $\dot{Y}_{\sigma(R)}$ von \dot{X}_R gemäß des Bewegungskernes P , und bewege den ersten Nachfahre $\dot{Y}_{\sigma(R)}$ von \dot{X}_R gemäß des h -transformierten Bewegungskernes P^h .
4. Die nächste Generation \mathfrak{X}_{n+1} ist nun gegeben durch die neuen und schon bewegten Partikel $\mathfrak{X}_{n+1} := \dot{Y}_1 + \dots + \dot{Y}_{D_1+\dots+D_k}$.

Verfahre im nächsten Zeitschritt genauso. Der daraus entstandene Prozess besitzt die gewünschte Verteilung $\mathbf{P}^{\tilde{h}}$.

Beweis Eine dem Beweis von Satz 17 ähnliche Rechnung liefert

$$\begin{aligned} \mathbf{P}^{\tilde{h}}(\mathfrak{X}_{n+1} = \dot{y}_1 + \dots + \dot{y}_l \mid \mathfrak{X}_n = \dot{x}_1 + \dots + \dot{x}_k) \\ = \sum_{r=1}^k \frac{h(x_r)}{\sum h(x_i)} \sum_{d_1, \dots, d_k} d_r \mathbf{P}(D_1 = d_1, \dots, D_k = d_k) \\ \cdot \sum_{[\pi] \in V} P(x_{\phi(1)}, y_{\pi(1)}) \cdots P^h(x_r, y_{\pi(\sigma(r))}) \cdots P(x_{\phi(l)}, y_{\pi(l)}) \end{aligned}$$

Mit den im Satz angegebenen bedingten Verteilungen erhalten wir

$$\begin{aligned}
&= \sum_{r=1}^k \mathbf{P}^{\tilde{h}}(R = r \mid \mathfrak{X}_n = \dot{x}_1 + \dots + \dot{x}_k) \\
&\quad \cdot \sum_{d_1 \dots d_k} \mathbf{P}^{\tilde{h}}(D_1 = d_1, \dots, D_k = d_k \mid R = r, |\mathfrak{X}_n| = k) \\
&\quad \cdot \sum_{[\pi] \in V} P(x_{\phi(1)}, y_{\pi(1)}) \cdots P^h(x_r, y_{\pi(\sigma(r))}) \cdots P(x_{\phi(l)}, y_{\pi(l)}).
\end{aligned}$$

■

Bemerkung Der in Satz 18 angegebene Ausgangsprozess mit Verteilung \mathbf{P} ist ein kritischer Verzweigungsprozess in dem Sinne, dass die erwartete Anzahl direkter Nachkommen eines jeden Partikels gleich 1 ist. Bekanntlich stirbt ein solcher Verzweigungsprozess fast sicher in endlicher Zeit aus (siehe [Har02]). Die angegebene harmonische Funktion \tilde{h} bedingt diesen Prozess unter anderem darauf, nicht auszusterben, denn es gilt $\mathbf{P}^{\tilde{h}}(D_R = 0) = 0$.

Aus der Konstruktion des \tilde{h} -transformierten Prozesses geht hervor, dass für die erwartete Anzahl von Partikeln unter dem Maß $\mathbf{P}^{\tilde{h}}$ gilt

$$\mathbf{E}^{\tilde{h}}[|\mathfrak{X}_{n+1}| \mid \mathfrak{X}_n] = |\mathfrak{X}_n| - 1 + \mathbf{E}^{\tilde{h}}[|\mathfrak{X}_1| \mid |\mathfrak{X}_0| = 1].$$

Im nicht-degenerierten Fall, d.h. die Geburtenrate ist nicht konstant 1, stellt der Prozess $(|\mathfrak{X}_n|)$ ein striktes $\mathbf{P}^{\tilde{h}}$ -Submartingal dar. Damit wachsen die Größen der Generationen des \tilde{h} -transformierten Prozesses. Im Spezialfall einer räumlich homogenen Kinderzahlverteilung impliziert dies, dass das $\mathbf{P}^{\tilde{h}}$ -Submartingal $(|\mathfrak{X}_n|)$ fast sicher gegen Unendlich konvergiert. □

Satz 19 Sei (\mathfrak{X}_n) ein Verzweigungsprozess mit Verteilung \mathbf{P} im Raum der endlichen Zählmaße über eine diskrete Menge S . Der Prozess besitze eine möglicherweise räumlich inhomogene aber nicht degenerierte Kinderzahlverteilung mit Erwartungswert 1, und P bezeichne den Bewegungskern. Weiter sei h eine positive P -harmonische Funktion und sei \tilde{h} definiert wie in Satz 16 als

$$\tilde{h}\left(\sum_{i=1}^k \dot{x}_i\right) := \sum_{i=1}^k h(x_i).$$

Dann erfüllt \tilde{h} die äquivalenten Eigenschaften in Folgerung 14, d.h. es gilt

$$\sup_n \tilde{h}(\mathfrak{X}_n) = \infty \quad \mathbf{P}^{\tilde{h}} - f.s. \quad \text{und} \quad \inf_n \tilde{h}(\mathfrak{X}_n) = 0 \quad \mathbf{P} - f.s.$$

Beweis Es genügt wegen der Äquivalenz in Folgerung 14 eine der beiden Eigenschaften nachzuweisen, wir wollen die zweite zeigen. Zunächst gilt offenbar $\sup \tilde{h} = \infty$ und $\inf \tilde{h} = 0$. Da (\mathfrak{X}_n) unter \mathbf{P} ein kritischer und nicht degenerierter Verzweigungsprozess ist, stirbt dieser fast sicher aus (siehe [Har02]), und damit folgt sofort

$$\inf_n \tilde{h}(\mathfrak{X}_n) = 0 \quad \mathbf{P} - f.s.,$$

womit der Satz bewiesen ist. ■

Bemerkenswerterweise wird in Satz 19 an die P -harmonische Funktion h keinerlei Voraussetzung gestellt, unabhängig von der Wahl von h ist der Satz gültig.

Zerlegung

Wir erhalten nun folgende Zerlegungen bezüglich einer harmonischen Funktion \tilde{h} eines Verzweigungsprozesses $(\hat{\mathfrak{X}}_n)$ mit Übergangskern \mathcal{P} :

1. Sei $(\hat{\mathfrak{X}}_n)$ ein in $\hat{\mathfrak{X}}_0 := \circ$ startender \tilde{h} -transformierter Verzweigungsprozess mit Übergangskern $\mathcal{P}^{\tilde{h}}$ und Y eine unabhängige reellwertige Zufallsvariable mit Verteilung

$$\mathbb{P}(Y > y) = \frac{h(o)}{y}.$$

2. Setze

$$\tau := \sup\{n \geq 0 : \tilde{h}(\hat{\mathfrak{X}}_n) < Y, \tilde{h}(\hat{\mathfrak{X}}_m) < \tilde{h}(\hat{\mathfrak{X}}_n) \forall m < n\}.$$

3. Sei $(\check{\mathfrak{X}}_n)$ ein bedingter Verzweigungsprozess mit geeignetem Übergangskern $\mathcal{Q}_{\tilde{h}(\hat{\mathfrak{X}}_\tau)}$ und Startzustand $\check{\mathfrak{X}}_0 = \hat{\mathfrak{X}}_\tau$. Setze

$$\bar{\mathfrak{X}}_n := \begin{cases} \hat{\mathfrak{X}}_n & n < \tau \\ \check{\mathfrak{X}}_{n-\tau} & n \geq \tau. \end{cases}$$

Dann ist $(\bar{\mathfrak{X}}_n)$ ein Verzweigungsprozess mit Übergangskern \mathcal{P} .

Beispiel: Kritischer Galton-Watson-Baum

Das vorgestellte Verzweigungsmodell stellt eine Verallgemeinerung des bekannten Galton-Watson-Modells dar, es wurde um die Raumkomponente erweitert. Falls man jedoch den Raum S als nur einelementig wählt, dann ist der Bewegungskern P degeneriert, und es existiert nur die konstante harmonische Funktion für P . Dann gibt im kritischen Fall (d.h. die erwartete Kinderzahl eines Partikels ist gleich 1) die harmonische Funktion \tilde{h} wie in Satz 16 gerade ein Generationsgröße an. Somit können wir mittels \tilde{h} einen kritischen Galton-Watson-Prozess an seiner maximalen Generation zerlegen, so dass dessen Anfangsstück bis zum Maximum in Verteilung dem Prozess entspricht mit der zusätzlichen Bedingung, nicht auszusterben.

Umgekehrt kann das Anfangsstück eines kritischer Galton-Watson-Prozess, der bedingt ist nicht auszusterben, bis zum Zeitpunkt der ersten minimalen Generationsgröße mit der dualen Zerlegung aus einem unbedingten kritischen Galton-Watson-Prozess erhalten werden.

Kapitel 3

Zeitkontinuierlicher Pfadraum

In diesem Kapitel werden die zeitdiskreten Zerlegungen in den Hauptsätzen 1 und 4 auf zeitkontinuierliche Prozesse verallgemeinert. Es gab schon einige Versuche in diese Richtung, bekannt sind hiervon vor allem die Zerlegung der Brownschen Bewegung nach Williams (siehe [Wil74, Pit75]) oder die Zerlegung von Lévy-Prozessen (siehe [Pit75, Ber91, Ber93, Ber92, Cha96]). Es gibt darüber hinaus auch Arbeiten (siehe [Jac74, Mil78, Mil77, PS73]), die in einem allgemeinen Kontext ein passendes Konzept einer Zerlegungszeit einführen und die Verteilung des Prozesses nach der Zerlegungszeit untersuchen. All diesen Arbeiten ist jedoch gemein, dass sie keine derart konstruktiven Zerlegungen wie die hier dargestellten beinhalten.

Pitman selbst brachte die Idee der harmonischen Funktion schon 1974 ein, und entwickelte eine Zerlegung, die ein Spezialfall der hier untersuchten darstellt. Er beschränkte sich in der betreffenden unveröffentlichten Arbeit [Pit74] auf Brownsche Bewegungen als Grundprozesse. Er entdeckte die folgende Zerlegung, die in dieser Arbeit einer stetigen Version der dualen Zerlegung in Hauptsatz 33 entspricht. Zusätzlich stellte Pitman noch die weiter unten folgende Vermutung auf.

Hauptsatz 20 (Pitman 1974) *Sei \mathbb{D} ein Greensches Gebiet in \mathbb{R}^d für $d \in \mathbb{N}$ und sei $b \in \mathbb{D}$. Sei weiter h eine positive unbeschränkte harmonische Funktion auf \mathbb{D} bezüglich einer Brownschen Bewegung auf \mathbb{D} . Sei (X_t^*) eine in b startende h -transformierte Brownsche Bewegung auf \mathbb{D} und definiere die Zufallsvariablen M und L durch*

$$M := \inf\{h(X_t^*) : t \geq 0\} \quad \text{und} \quad L := \inf\{t \geq 0 : h(X_t^*) = M\}.$$

Dann ist M uniform auf $[0, h(b)]$ verteilt, und bedingt auf $M = m$ ist der pre- L -Prozess $(X_t^)_{0 \leq t \leq L}$ identisch verteilt zu einer freien Brownschen Bewegung auf \mathbb{D} mit Startpunkt b bis zum Zeitpunkt des ersten Treffens in die Menge $\{y \in \mathbb{D} : h(y) = m\}$.*

Vermutung (Pitman 1974) *Der pre- L - und der post- L -Prozess sind bedingt unabhängig gegeben X_L^* und bedingt auf das pre- L -Stück mit $X_L^* = x$ und $M = h(x) = m$ ist der post- L -Prozess $(X_{L+t}^*)_{t \geq 0}$ ein $(h - m)$ -transformierter Prozess mit Startpunkt x auf dem Gebiet $\{y : h(y) > m\}$.*

Die postulierte Unabhängigkeit wird in dieser Arbeit ebenfalls vollständig in Hauptsatz 33 bewiesen. Die vermutete einfache Darstellung der Verteilung des post- L -Prozesses wird in

Abschnitt 3.5 ausführlich diskutiert – es wird dort allerdings eine stärkere Voraussetzungen als in Pitmans Formulierung benötigt.

Dieses Kapitel – das Herzstück dieser Arbeit – wird einige Brücken zwischen diesen Arbeiten schlagen, und eine allgemeine Zerlegung von Markov-Prozessen mit Hilfe harmonischer Funktionen bereitstellen. Das Hauptaugenmerk liegt hierbei immer auf dem Anfangsstück bis zu einem Zerlegungszeitpunkt τ , die Verteilung des post- τ -Prozesses wurde bereits detailliert in [PS73] untersucht.

3.1 Koterminale Zeiten

Sei im folgenden $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$ eine Filtration, die die üblichen Bedingungen erfüllt, und sei $(X_t)_{t \in \mathbb{R}_+}$ ein adaptierter, starker Markov-Prozess mit Werten in einem kompakten metrischen Raum (S, \mathcal{S}) mit Metrik d und mit rechtsseitig stetigen Pfaden mit linksseitigen Limiten. Der Wertebereich des gesamten Prozesses (X_t) liege im mit \mathcal{D} bezeichneten Raum der rechtsseitig stetigen Pfade mit linksseitigen Limiten. Das zugrundeliegende Wahrscheinlichkeitsmaß wird mit \mathbf{P} und die Erwartung wird mit \mathbf{E} bezeichnet, und die Familie der Übergangskerne mit (P_t) . Bei Start in einem Punkt $x \in S$ verwenden wir zudem die üblichen Schreibweisen \mathbf{P}_x bzw. \mathbf{E}_x . Der Raum S besitze zusätzlich einen isolierten Punkt Δ mit $\mathbf{P}_\Delta(X_t = \Delta \forall t > 0) = 1$.

Zufällige Zeiten

Um allgemein mit einer zufälligen Zeit R (die nicht notwendigerweise eine Stoppzeit ist) eines Prozesses $(X_t)_{t \in \mathbb{R}_+}$ zu arbeiten, sind die mit R und (X_t) verbundenen σ -Algebren von Interesse. Es ist per se nicht klar, wie die σ -Algebra einer solchen allgemeinen zufälligen Zeit sinnvoll definiert werden soll, so dass R und X_R bezüglich dieser σ -Algebra messbar sind. Halten wir uns hier an das Konzept und an die Definitionen aus [PS73] und setzen

$$\begin{aligned}\mathcal{F}_{R-} &:= \bigvee_{t \geq 0} \mathcal{F}_t \cap \{R > t\} \\ \mathcal{F}_R &:= \mathcal{F}_{R-} \vee \bigvee_{S \text{ Stoppzeit}} \mathcal{F}_S \cap \{S = R\}\end{aligned}$$

mit der üblichen Notation $\mathcal{F}_t \cap A := \{A \cap B : B \in \mathcal{F}_t\}$.

Diese Definitionen sind äquivalent zu den folgenden Darstellungen (siehe Kapitel VI in [RW00])

$$\begin{aligned}\mathcal{F}_{R-} &= \sigma\{\{T < R\} : T \text{ ist eine Stoppzeit}\} \\ \mathcal{F}_R &= \sigma\{\{T \leq R\} : T \text{ ist eine Stoppzeit}\}.\end{aligned}$$

Die so definierten Teilfelder \mathcal{F}_{R-} und \mathcal{F}_R erfüllen gerade die gewünschten Messbarkeitsbedingungen, denn sowohl die zufällige Zeit R als auch der Wert X_R des Prozesses sind dann in geeigneter Art und Weise messbar, wie der folgende Satz zeigt:

Satz 21 *Sei R eine zufällige Zeit. Dann gilt*

$$\sigma(R) \subset \mathcal{F}_{R-} \subset \mathcal{F}_R \quad \text{und} \quad \sigma((X_t)_{t < R}) \subset \mathcal{F}_{R-} \quad \text{und} \quad \sigma((X_t)_{t \leq R}) \subset \mathcal{F}_R.$$

Beweis Siehe Lemma 4.1 in [PS73]. ■

Koterminaler Zeiten

Auch in diesem Kapitel der zeitkontinuierlichen Prozesse stehen nicht allgemeine zufällige Zeiten im Mittelpunkt, sondern wiederum eine spezielle Klasse, die im Falle zeitdiskreter Markov-Ketten durch die Zerlegungszeiten bereit gestellt wurde. Rufen wir uns die Definition 3 einer diskreten Zerlegungszeit nach Jacobsen und Williams kurz zurück in Erinnerung, so lautet die analoge Definition für zeitkontinuierliche Prozesse wie folgt:

Eine zufällige Zeit L heißt ZERLEGUNGSZEIT eines Prozesses (X_t) , falls sie als Funktional auf den Pfaden die folgende Kreuzungseigenschaft besitzt: Für jeweils zwei Pfade $\omega_1, \omega_2 : \mathbb{R}_+ \rightarrow S$ mit $L\omega_1 = L\omega_2$ und $\omega_1(L\omega_1) = \omega_2(L\omega_2)$ und den wie folgt zusammengesetzten Pfad $\omega : \mathbb{N}_0 \rightarrow S$

$$\omega(i) := \begin{cases} \omega_1(t) & \text{für } t \leq L\omega_1 \\ \omega_2(t) & \text{für } t > L\omega_1 \end{cases}$$

gilt $L\omega = L\omega_1 (= L\omega_2)$.

Eine solche Definition für zeitkontinuierliche Pfade ist allerdings problematisch und nur eingeschränkt sinnvoll. Denn in vielen Fällen, die durch eine geeignete Definition mit erfasst werden sollen, gilt $X_L \neq X_{L-}$, allerdings beeinflusst X_{L-} den Wert der Zerlegungszeit maßgeblich. Denken wir zum Beispiel an einen Lévy-Prozess mit nur positiven Sprüngen und einem negativen linearen Drift, und bezeichne L den Zeitpunkt des globalen Infimums, so wird dieses (falls $L < \infty$ ist) durch X_{L-} angenommen, während der Wert X_L in diesem Kontext von geringerem Interesse ist. Die obige Definition einer Zerlegungszeit kann dieses Beispiel nicht erfassen. Für stetige Prozesse treten diese Probleme natürlich nicht auf.

Eine naheliegende Möglichkeit wäre eine Erweiterung der Definition einer Zerlegungszeit um zusätzliche X_{L-} betreffende Bedingungen, wir wollen jedoch einen anderen Weg einschlagen, der von Pittenger und Shih in ihrer Arbeit „Coterminal Families and the strong Markov Property“ (siehe [PS73]) durch die Einführung sogenannter KOTERMINALER ZEITEN als Verallgemeinerung von Stoppzeiten erfolgreich besprochen wurde. Neben dem Begriff der koterminalen Zeit können wir auf Resultate über die Verteilung eines auf \mathcal{F}_L bzw. eines auf \mathcal{F}_{L-} bedingten Prozesses zurückgreifen. Diese Verteilungen werden zur Beschreibung des zweiten Teiles der Konstruktion benötigt.

In einem ersten Schritt wollen wir den Begriff der koterminalen Zeit als ein deterministisches Funktional $l : \mathcal{D} \rightarrow \mathbb{R}_+$ einführen. Hierfür werden zwei geläufige Operatoren benötigt. Der Killing-Operator $k_s : \mathcal{D} \rightarrow \mathcal{D}$ operiert wie folgt auf einem Pfad ω

$$k_s\omega(t) := \omega(t) \quad \text{für } t < s \quad \text{und} \quad k_s\omega(t) := \Delta \quad \text{für } t \geq s,$$

und der Shift-Operator $\theta_s : \mathcal{D} \rightarrow \mathcal{D}$ ist definiert durch

$$\theta_s\omega(t) := \omega(s+t) \quad \text{für } t \geq 0.$$

Definition 22 (Koterminaler Zeit) *Ein Funktional $l : \mathcal{D} \rightarrow \mathbb{R}_+$ heißt KOTERMINAL, falls die folgenden Eigenschaften besitzt:*

1. $l \circ \theta_s = \max(l - s, 0)$,
2. $l \circ k_s = l$ auf der Menge $\{\omega \in \mathcal{D} : l\omega < s\}$,
3. $l \circ k_s \leq s$ für alle $s \geq 0$,

4. $l = \lim_{s \nearrow \infty} l \circ k_s$ pfadweise.

Eine zufällige Zeit L heißt EXAKTE KOTERMINALE ZEIT, falls diese darstellbar ist als $L = l(X)$, wobei (X_t) ein an (\mathcal{F}_t) adaptierter Prozess mit Pfaden fast sicher in \mathcal{D} ist, und l ein koterminales Funktional ist.

Zu jeder exakten koterminalen Zeit $L = l(X)$ ist die zugehörige TERMINALE ZEIT T definiert durch

$$T := \inf\{t > 0 : L_t > 0\} \quad \text{mit} \quad L_t := l \circ k_t(X).$$

T ist damit eine Stoppzeit (siehe Lemma 3.2 in [PS73]).

Die Definition einer koterminalen Zeit lässt sich auf natürliche Weise auf nicht-deterministische Funktionale l erweitern. Sei hierfür l ein zufälliges \mathcal{F}_0 -messbares Funktional, so dass jede Realisierung ein koterminales Funktional auf \mathcal{D} ist. Dann kann l dargestellt werden als $l = l_Y$, wobei Y eine geeignete \mathcal{F}_0 -messbare Zufallsvariable mit Werten in einem Raum R ist, und $(l_y)_{y \in R}$ eine Familie deterministischer koterminaler Funktionale ist. Als Funktional auf dem Pfadraum $R \times \mathcal{D}$ aller Pfade (y, ω) mit Werten in $R \times S$ mit konstanter erster Komponente y ist $l(\cdot)$ wieder koterminal. Durch die einfache Erweiterung des Prozesses (X_t) zu (Y, X_t) ist deshalb $L := l_Y(X)$ ebenfalls eine exakte koterminale Zeit, die aus Sicht der Pfade von (X_t) ein zufälliges \mathcal{F}_0 -messbares Funktional ist. Wir werden von nun an auch solche zusätzlich randomisierte Zeiten $L = l_Y(X)$ ohne expliziten Hinweis als koterminal bezeichnen.

Beispiel Ein prominentes Beispiel einer koterminalen Zeit ist der letzte Austrittszeitpunkt L_A eines Prozesses (X_t) aus einer Menge A

$$L_A := \sup\{t \in \mathbb{R}_+ : X_t \in A\},$$

mit der Vereinbarung $\sup \emptyset = 0$. Diese Zeit erfüllt offenbar die Bedingungen der Definition einer koterminalen Zeit. Welche Bedeutung kommt in diesem Beispiel der terminalen Zeit T_A zu? Es gilt nach Definition

$$T_A = \inf\{t \geq 0 : l_A \circ k_t(X) > 0\} \quad \text{mit} \quad l_A \omega := \sup\{t \in \mathbb{R}_+ : \omega(t) \in A\}$$

und es ist $l_A \circ k_t(X) > 0$ genau dann, wenn der Prozess vor dem Zeitpunkt t innerhalb der Menge A gewesen ist. Es ist also

$$T_A = \inf\{t \geq 0 : X_t \in A\}.$$

Wie wir später sehen werden, ist der Zeitpunkt des ersten oder letzten Supremums eines reellwertigen Prozesses keine koterminale Zeit, da die zweite Bedingung der Definition im Allgemeinen verletzt wird; durch Übergang zu einem erweiterten Prozess mit einem nicht-kanonischen Shift-Operator kann dieser Zeitpunkt jedoch wieder in diesen Rahmen gefasst werden. \square

Bemerkung Im Vergleich zwischen einer koterminalen Zeit nach Pittenger und Shih und einer Zerlegungszeit nach Williams wird klar, dass beides unterschiedliche Objekte sind.

So ist eine koterminaler Zeit nicht notwendigerweise eine Zerlegungszeit, nicht einmal im Falle stetiger Pfade. Betrachten wir hierfür einen Prozess (X_t) mit stetigen Pfaden, und definieren wir die Zeit L durch

$$L := l(X) \quad \text{mit} \quad l\omega := \sup \left\{ t \geq 0 : \lim_{s \searrow t} \frac{\omega(s) - \omega(t)}{s - t} = \lim_{s \nearrow t} \frac{\omega(s) - \omega(t)}{s - t} \right\},$$

dann bezeichnet L also gerade den letzten Zeitpunkt, bei dem die rechtsseitige und linksseitige Ableitung von (X_t) übereinstimmen. Es ist leicht einzusehen, dass L eine koterminaler Zeit ist – allerdings erfüllt L offenbar nicht die Kreuzungseigenschaft in der Definition einer Zerlegungszeit nach Jacobsen und Williams. Denn das Kreuzen zweier Pfade ω_1, ω_2 mit $l\omega_1 = l\omega_2$ führt im Allgemeinen zu einem Pfad, dessen rechtsseitige und linksseitige Ableitungen sich an der Stelle $l\omega_1$ unterscheiden.

Es ist auch nicht jede Zerlegungszeit auch eine koterminaler Zeit, wie das folgende Beispiel illustriert. Betrachten wir einen Poisson-Prozess (X_t) mit Sprüngen der Größe 1 und einem zusätzlichen negativen linearen Drift, und setzen wir

$$T := \sup\{t \in \mathbb{R}_+ : X_t = 1\} \quad T' := \sup\{t < T : X_t = 1\},$$

dann ist T' eine Zerlegungszeit im Sinne von Williams, nicht aber eine koterminaler Zeit, da Bedingung 3 verletzt ist, sobald der Prozess zwischen den Zeitpunkten T' und T gekillt wird. \square

Bemerkung Dieser Abschnitt 3.1 ist der einzige in diesem und dem folgenden Kapitel, in dem wir auf Williams' Definition einer Zerlegungszeit zurückgreifen. In den folgenden Abschnitten wird der Begriff einer Zerlegungszeit anhand der Konstruktion einer Pfadzerlegung vollkommen anders definiert werden. \square

Bedingte Verteilungen

Nun soll der post- L -Prozess eingehend untersucht werden, insbesondere stellt sich natürlich die Frage nach dessen Verteilung. Diese Frage ist auch der zentrale Punkt in der Arbeit von Pittenger und Shih [PS73]. Es werden hier nur möglichst knapp die für uns relevanten Resultate wiedergegeben; die Beweise sind technisch aufwendig und umfangreich.

Bei der Frage nach der Verteilung des Prozesses (X_t) nach dem Zeitpunkt L gibt es zwei Möglichkeiten zu bedingen: Entweder man bedingt auf \mathcal{F}_L oder auf \mathcal{F}_{L-} . Diese beiden Ansätze führen auch zu zwei verschiedenen Eintrittsmaßen, denn während zu dem Zeitpunkt L der Wert X_L schon bekannt ist, stellt dieser Wert zu dem Zeitpunkt $L-$ noch einen relevanten Freiheitsgrad dar. Für die in dieser Arbeit vorgestellte Zerlegung sind beide Bedingungen bzw. Eintrittsmaße von Interesse und kommen in jeweils unterschiedlichen Situationen zum Tragen.

Im Zusammenhang mit einer koterminalen Zeit L und der assoziierten terminalen Zeit T teilen wir den Grundraum S in zwei unterschiedlich zu behandelnde Teilmengen ein. Es sei $S^r := \{x \in S : \mathbf{P}_x(T > 0) = 0\}$ die Menge der REGULÄREN PUNKTE und $S^i := \{x \in S : \mathbf{P}_x(T > 0) = 1\}$ die Menge der IRREGULÄREN PUNKTE.

Bedingen auf \mathcal{F}_L

Betrachten wir zunächst das Eintrittsmaß des post- L -Prozesses gegeben \mathcal{F}_L . Setze für $x \in S^i$ und $s > 0$

$$D_s(x, f) := \mathbf{E}^x[f(X_s) | T > s]$$

und für $x \in S^r$ und $s, u > 0$ und $n \in \mathbb{N}$ setze

$$D_n(x, u, s, f) := \mathbf{E}^x[f(X_s) | T \circ \theta_u > s - u, d(x, X_u) < 1/n].$$

Weiter sind die folgenden Grenzwerte zunächst nur formal definiert als

$$\begin{aligned} \overline{D}_n(x, s, f) &:= \limsup_{u \rightarrow 0} D_n(x, u, s, f) & \underline{D}_n(x, s, f) &:= \liminf_{u \rightarrow 0} D_n(x, u, s, f) \\ \overline{D}_s(x, f) &:= \limsup_{n \rightarrow \infty} \overline{D}_n(x, s, f) & \underline{D}_s(x, f) &:= \liminf_{n \rightarrow \infty} \underline{D}_n(x, s, f). \end{aligned}$$

Falls $\overline{D}_s(x, f) = \underline{D}_s(x, f)$, dann wird der gemeinsamen Wert mit $D_s(x, f)$ bezeichnet. Für eine messbare Funktion f sei $W_D(f) := \{(x, s) : D_s(x, f) \text{ existiert}\}$, d.h.

$$W_D(f) := \{(x, s) \in S \times (0, \infty) : \overline{D}_s(x, f) = \underline{D}_s(x, f)\},$$

und für $(x, s) \in W_D(f)$ sei schließlich

$$D_s^L(x, f) := \frac{D_s(x, f q_\infty)}{D_s(x, q_\infty)}$$

mit der Funktion q_∞ definiert durch

$$q_\infty(x) := \mathbf{P}^x(T = \infty).$$

Falls $f = I_A$ die Indikatorvariable einer messbaren Menge $A \in \mathcal{S}$ ist, dann verwenden wir die abkürzende Notation

$$D_s^L(x, A) := D_s^L(x, I_A).$$

Mit dieser Konstruktion stellt die Familie $(D_s^L)_{s>0}$ gerade das Eintrittsmaß des auf \mathcal{F}_L bedingten post- L -Prozesses $(X_t)_{t>L}$ dar, wie der folgende Satz zeigt.

Satz 23 (Pittenger, Shih: Theorem 6) *Auf dem Ereignis $\{L < \infty\}$ sind die Maße D_t^L fast sicher für alle reellen und messbaren Funktionen wohldefiniert, und es gilt \mathbf{P} -fast sicher*

$$\mathbf{E}[f(X_{L+t}) | \mathcal{F}_L] = D_t^L(X_L, f) \quad \text{bzw.} \quad \mathbf{P}(X_{L+t} \in A | \mathcal{F}_L) = D_t^L(X_L, A).$$

Bedingen auf \mathcal{F}_{L-}

Wie schon angedeutet, führt Bedingen auf \mathcal{F}_{L-} zu einem anderen Eintrittsmaß für den post- L -Prozess. Dieses soll nun konstruiert werden. Für $x \in S$ setze

$$Q(x, u, s, g, f) := \mathbf{E}^x[g(X_u)f(X_s) | T \leq u, T \circ \theta_u > s - u]$$

$$\overline{Q}_s(x, g, f) := \limsup_{u \rightarrow 0} Q(x, u, s, g, f) \quad \underline{Q}_s(x, g, f) := \liminf_{u \rightarrow 0} Q(x, u, s, g, f).$$

Und falls $\overline{Q}_s(x, g, f) = \underline{Q}_s(x, g, f)$, dann bezeichnen wir den gemeinsamen Wert mit $Q_s(x, g, f)$. Für messbare Funktionen h und f sei $W_Q(g, f)$ der Definitionsbereich von $Q(\cdot, g, f)$, für feste Funktionen f und g sei

$$W_Q(g, f) := \{(x, s) \in S \times (0, \infty) : \overline{Q}_s(x, g, f) = \underline{Q}_s(x, g, f)\},$$

und für $(x, s) \in W_Q(g, f)$ sei schließlich

$$Q_s^L(x, g, f) := \frac{Q_s(x, g, f q_\infty)}{Q_s(x, 1, q_\infty)}.$$

Falls $f = I_A$ und $g = I_B$ die Indikatorvariablen messbarer Menge $A, B \in \mathcal{S}$ sind, dann verwenden wir wieder die abkürzende Notation

$$Q_s^L(x, B, A) := Q_s^L(x, I_B, I_A).$$

Mit dieser Konstruktion stellt die Familie $(Q_s^L)_{s>0}$ gerade das Eintrittsmaß des auf \mathcal{F}_{L-} bedingten post- L -Prozesses $(X_t)_{t \geq L}$ dar, wie der folgende Satz zeigt.

Satz 24 (Pittenger, Shih: Theorem 6) *Auf dem Ereignis $\{L < \infty, X_{L-} \in S^r\}$ sind die Maße Q_t^L fast sicher für alle reellen und messbaren Funktionen wohldefiniert, und es gilt \mathbf{P} -fast sicher*

$$\begin{aligned} \mathbf{E}[g(X_L)f(X_{L+t})|\mathcal{F}_{L-}] &= Q_t^L(X_{L-}, g, f) \\ \text{bzw. } \mathbf{P}(X_L \in A, X_{L+t} \in B|\mathcal{F}_{L-}) &= Q_t^L(X_{L-}, A, B) \end{aligned}$$

Es besteht ein natürlicher Zusammenhang zwischen den Maßfamilien (D_s^L) und (Q_s^L) , wie der folgende Satz zeigt.

Satz 25 (Pittenger, Shih: Theorem 7) *Es existiert ein Kern $\eta : S \times \mathcal{S} \rightarrow [0, 1]$ mit*

$$Q_s^L(x, g, f) = \int \eta(x, dy)g(y)D_s^L(y, f).$$

Die Verteilung des post- L -Prozesses

Bis jetzt wurden nur die Eintrittsmaße (D_s^L) und (Q_s^L) des post- L -Prozesses bestimmt, es fehlen jedoch noch die Übergangskerne, um die Verteilung vollständig zu charakterisieren. Diese Kerne lassen sich schnell konstruieren:

$$\begin{aligned} H_s f(x) &:= \mathbf{E}_x[f(X_s) : T > s] \\ K_s f(x) &:= H_s f q_\infty(x) / q_\infty(x) = \mathbf{E}_x[f(X_s) | T = \infty] \end{aligned}$$

Damit sind dann (D_s^L) bzw. (Q_s^L) Eintrittsmaße bezüglich der Kernfamilie (K_s) in dem Sinne, dass für $s, t > 0$ gilt (siehe Theorem 7 in [PS73])

$$D_s^L(x, K_t f) = D_{s+t}^L(x, f) \quad \text{bzw.} \quad Q_s^L(x, h, K_t f) = Q_{s+t}^L(x, h, f).$$

Das zentrale Resultat der Arbeit von Pittenger und Shih ist der nun folgende Satz, der die Verteilung des post- L -Prozesses charakterisiert.

Satz 26 (Pittenger, Shih: Theorem 8) Seien $f_1, \dots, f_n : S \rightarrow \mathbb{R}$ messbare und beschränkte Funktionen und $0 < s_1 < \dots < s_n$. Dann gilt fast sicher auf $\{L < \infty\}$

$$\begin{aligned} & \mathbf{E} \left[\prod_{i=1}^n f_i(X_{L+s_i}) \middle| \mathcal{F}_L \right] \\ &= \int \cdots \int D_{s_1}^L(X_L, dz_1) f_1(z_1) K_{s_2-s_1}(z_1, dz_2) f_2(z_2) \cdots \\ & \quad \cdots \cdot K_{s_n-s_{n-1}}(z_{n-1}, dz_n) f_n(z_n) \end{aligned}$$

Außerdem gilt für stetiges $h : S \rightarrow \mathbb{R}$ fast sicher auf $\{L < \infty, X_{L-} \in S^r\}$

$$\begin{aligned} & \mathbf{E} \left[g(X_L) \prod_{i=1}^n f_i(X_{L+s_i}) \middle| \mathcal{F}_{L-} \right] \\ &= \int \cdots \int Q_{s_1}^L(X_{L-}, g, dz_1) f_1(z_1) K_{s_2-s_1}(z_1, dz_2) f_2(z_2) \cdots \\ & \quad \cdots \cdot K_{s_n-s_{n-1}}(z_{n-1}, dz_n) f_n(z_n) \\ &= \int \cdots \int \eta(X_{L-}, dy) g(y) D_{s_1}^L(y, dz_1) f_1(z_1) K_{s_2-s_1}(z_1, dz_2) f_2(z_2) \cdots \\ & \quad \cdots \cdot K_{s_n-s_{n-1}}(z_{n-1}, dz_n) f_n(z_n) \end{aligned}$$

Verzweigungspunkte

Die Theorie und Methoden von Pittenger und Shih sind selbst in einem sehr allgemeinen Rahmen anwendbar, es wird sogar die Existenz von VERZWEIGUNGSPUNKTEN in S erlaubt, die Sätze gelten bis auf einige Einschränkungen weiterhin. Diese Punkte sind $x \in S$ mit der Eigenschaft $\mathbf{P}_x(X_0 \neq x) = 1$. Die Menge aller Verzweigungspunkte wird mit S^b bezeichnet (b steht für BRANCHING POINT). Das VERZWEIGUNGSMASS ν ist für $x \in S$ und $A \in \mathcal{S}$ definiert als

$$\nu(x, A) := \mathbf{P}_x(X_0 \in A)$$

Falls $x \in S \setminus S^b$ ist, dann gilt offenbar $\nu(x, y) = \delta_{x,y}$. Für einen tieferen Einblick in die Theorie der Verzweigungspunkte und damit verbundenen Ray-Prozesse sei auf den Abschnitt „Ray Processes“ in [RW94] verwiesen. Die Verzweigungspunkte erfordern zusätzlich eine modifizierte Behandlung in der Definition der Quasi-Linksstetigkeit, die sich in der folgenden abgeschwächten starken Markov-Eigenschaft zeigt.

Definition 27 (Beinahe Quasi-Linksstetig) Ein Prozess (X_t) mit Verzweigungsmaß ν heißt genau dann BEINAHE QUASI-LINKSSTETIG, falls für jede aufsteigende und konvergente Folge von Stoppzeiten $T_n \nearrow T$ Stoppzeiten gilt:

$$\mathbf{E}^\mu \left[f(X_T) \middle| \bigvee_n \mathcal{F}_{T_n} \right] = \int \nu(X_{T-}, dy) f(y),$$

wobei $\bigvee \mathcal{F}_{T_n} := \sigma\{\mathcal{F}_{T_n}, n \in \mathbb{N}_0\}$.

Falls der Prozess keine Verzweigungspunkte besitzt (also $\nu(x, dy) = \delta_{x,y}$ ist), fällt diese Definition offenbar mit der üblichen Quasi-Linksstetigkeit zusammen.

3.2 Hauptresultate

Sei im folgenden $(X_t)_{t \in \mathbb{R}_+}$ wieder ein an eine Filtration $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$ adaptierter Markov-Prozess mit Werten in einem kompakten metrischen Raum (S, \mathcal{S}) und sei (P_t) die zugehörige Kernfamilie. Die Pfade von (X_t) seien wie üblich fast sicher rechtsseitig stetig mit linksseitigen Limiten.

An die Stelle der positiven harmonischen Funktionen im diskreten Fall treten nun im Kontinuierlichen die regulären Funktionen, die wie folgt definiert sind:

Definition 28 (Reguläre Funktion) *Eine Funktion $h : S \rightarrow \mathbb{R}_+$ heißt REGULÄR (bezüglich einer Kernfamilie (P_t)), falls sie die folgenden Bedingungen erfüllt:*

$$0 \leq h < \infty \quad \text{und} \quad P_t h = h \quad \forall t > 0 \quad \text{und} \quad h(\Delta) = 0.$$

Analog der h -Transformation im diskreten Fall ist die h -Transformation einer Kernfamilie wie folgt definiert.

Definition 29 *Für eine reguläre Funktion h ist die h -Transformation der Kernfamilie auf der Menge S^h gegeben durch*

$$P_t^h(x, dy) := \frac{h(y)}{h(x)} P_t(x, dy) \quad x \in S^h.$$

Gegeben eine Verteilung \mathbf{P} mit Erwartung \mathbf{E} eines Markov-Prozesses mit Übergangskernen (P_t) induziert die h -transformierte Kernfamilie (P_t^h) eine Verteilung \mathbf{P}^h mit Erwartung \mathbf{E}^h , die als Maßwechsel aus \mathbf{P} wie folgt implizit hervorgeht:

$$\mathbf{E}^h \phi(X_{t_1}, \dots, X_{t_n}) = \mathbf{E} \frac{1}{h(X_0)} \phi(X_{t_1}, \dots, X_{t_n}) h(X_{t_n})$$

für alle $0 < t_1 < \dots < t_n$ und alle nicht-negativen messbaren Funktionen $\phi : S^n \rightarrow \mathbb{R}$.

Beschreibung des Supremums

Für eine stetige, reguläre Funktion h sei das laufende Supremum des Prozesses $(h(X_t))$ im Folgenden mit (M_t) bezeichnet, und M_∞ sei das globale Supremum, also

$$M_t := \sup_{s \in [0, t]} h(X_s) \quad \text{und} \quad M_\infty := \lim_{t \rightarrow \infty} M_t = \sup_{s \geq 0} h(X_s).$$

Da der Prozess (X_t) selbst rechtsseitig stetige Pfade mit linksseitigen Limiten besitzt, und h als stetig vorausgesetzt wurde, besitzen sowohl der Prozess $(h(X_t))$ als auch (M_t) beide ebenfalls rechtsseitig stetige Pfade mit linksseitigen Limiten. Der Zeitpunkt des ersten Supremums bis zur Zeit t werde mit L_t bezeichnet und durch den Grenzübergang $t \rightarrow \infty$ erhalten wir den ersten Zeitpunkt L des globalen Supremums:

$$\begin{aligned} L_t &:= \inf \{s \in [0, t] : M_s = M_t\} \\ L &:= \inf \{s \in \mathbb{R}_+ : M_s = M_\infty\}. \end{aligned}$$

Die Zeitpunkte L_t erzeugen eine koterminale Familie für den erweiterten Markov-Prozess $(X_t, M_t)_{t \in \mathbb{R}_+}$ zusammen mit dem Shift-Operator θ_s , der beide Komponenten gemeinsam verschiebt, also

$$\theta_s(X, M)(t) := (X_{t+s}, M_{t+s}).$$

Setzen wir $(\tilde{X}, \tilde{M})(t) := \theta_s(X, M)(t) = (X_{t+s}, M_{t+s})$ für ein $s > 0$, dann besteht im Allgemeinen nur noch der schwächerer Zusammenhang zwischen den Komponenten

$$\tilde{M}_t \geq \sup_{s \in [0, t]} h(\tilde{X}_s) \quad \text{im Vergleich zu} \quad M_t = \sup_{s \in [0, t]} h(X_s).$$

Im geshifteten Prozess wird anfangs ein Supremum vorgegeben, das oberhalb vom Startpunkt liegen darf und nicht mit diesem zusammenfallen muss, d.h.

$$\tilde{M}_t = \tilde{M}_0 \vee \sup_{s \in [0, t]} h(\tilde{X}_s).$$

In dieser Konstruktion ist L eine koterminalen Zeit bezüglich des zusammengestzten Prozesses, nicht jedoch notwendigerweise bezüglich des einfacheren Prozesses (\tilde{X}_t) , denn die zweite definierende Eigenschaft einer koterminalen Zeit (nämlich $L \circ \theta_s = \max(L - s, 0)$) wäre verletzt.

Die mit L assoziierte terminale Zeit T ist definiert als

$$T := \inf\{t > 0 : L_t > 0\}.$$

Sowohl der Prozess (M_t) als auch (L_t) sind monoton wachsend. Aus der Definition von L_t folgt zusätzlich, dass der Prozess (L_t) gekoppelt mit dem Prozess (M_t) wächst in dem Sinne, dass für alle $t > s \geq 0$ gilt

$$M_t > M_s \Leftrightarrow L_t > L_s,$$

denn nur sobald der Prozess $(h(X_t))$ ein neues Supremum erreicht, wachsen sowohl L_t als auch M_t . Insbesondere gilt $L_t > L_0 = 0$ genau dann, wenn $M_t > M_0 = h(X_0)$ ist. Dies bedeutet für T

$$T = \inf\{t > 0 : M_t > M_0\} = \inf\{t > 0 : h(X_t) > h(X_0)\},$$

wobei die zweite Gleichheit allerdings nur für den ungeschifteten Prozess mit $M_0 = h(X_0)$ gilt.

Damit können die mit L assoziierten bedingten Eintrittsmaße D_t^L, Q_t^L und die Kernfamilie K_t^L definiert werden. Da diese sich jedoch auf den erweiterten Prozess (X_t, M_t) beziehen, erleichtern wir den Umgang mit den Objekten durch folgende abgewandelte Notation, die nur Bezug auf das erreichte Supremum M_L nimmt. Für $m \geq 0$ und $x, y \in S$ mit $h(x) = m$ und $h(y) > m$ setze

$$\begin{aligned} q_m(x) &:= \mathbf{P}_x(h(X_s) \leq m \forall s \geq 0) \\ D_t^m(x, f) &:= \mathbf{E}[f(X_t) | X_0 = x, h(X_s) \leq m \forall s \geq 0] \\ Q_t^m(x, g, f) &:= \mathbf{E}[g(X_0)f(X_t) | X_{0-} = x, h(X_s) \leq m \forall s \geq 0] \\ K_t^m(x, f) &:= \mathbf{E}[f(X_t) | X_0 = x, h(X_s) \leq m \forall s \geq 0]. \end{aligned} \tag{3.1}$$

Beschreibung der Konstruktion

Ziel ist es nun, analog zur zeitdiskreten Zerlegung eine alternative Konstruktion für den Prozess (X_t) zu finden, die die gleiche Verteilung besitzt und zusätzlich den Prozess am Supremum von $(h(X_t))$ zerlegt. Das Anfangsstück wird wieder durch einen h -transformierten Prozess (\hat{X}_t) mit Startpunkt o beschrieben bis zu einem zufälligen Zeitpunkt τ , der in Verteilung mit dem Zeitpunkt des ersten Supremums übereinstimmt.

Achtung: τ wird im folgenden als ZERLEGUNGSZEIT bezeichnet, obgleich τ keine Zerlegungszeit im Sinne von Williams ist (siehe Abschnitt 3.1). Da wir allerdings letztere Definition nicht weiter verfolgen, sollte dies zu keinen Missverständnissen führen.

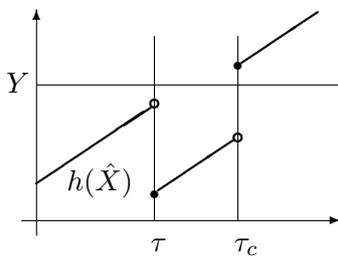
Die Zerlegungszeit (in dem Sinne der folgenden Konstruktion) τ bezeichnet den Zeitpunkt des ersten globalen Supremums des Anfangsstückes ($h(\hat{X}_t)$) vor dem ersten Überschreiten einer zufälligen Schranke Y ; die exakte Definition ist ein wenig technisch aufwendiger als die der zeitdiskreten Zerlegung, da potentielle Sprünge von (X_t) verschiedene Fälle ermöglichen. Sei als erstes die Überkreuzungszeit τ_c definiert als

$$\tau_c := \inf\{t \in \mathbb{R}_+ : h(\hat{X}_t) > Y\},$$

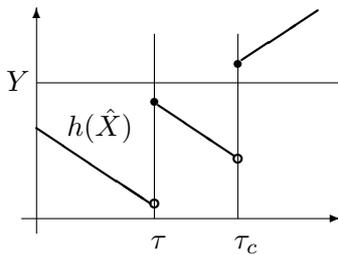
und hiermit wird die Zerlegungszeit τ definiert als

$$\tau := \inf\{t \in \mathbb{R}_+ : h(\hat{X}_t) \vee h(\hat{X}_{t-}) = \sup_{s < \tau_c} h(\hat{X}_s)\},$$

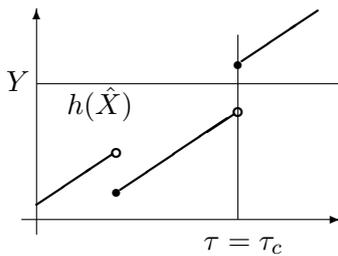
wobei sowohl τ als auch τ_c den Wert ∞ mit positiver Wahrscheinlichkeit annehmen dürfen. Diese Definition von τ zeigt deutliche Parallelen zur Definition des Zeitpunkts des ersten globalen Supremums L von $(h(X_t))$. Da für (\hat{X}_t) nur rechtsseitige Stetigkeit und die Existenz von linksseitigen Limiten vorausgesetzt ist, ist es im Allgemeinen nicht möglich $h(\hat{X}_\tau)$, $h(\hat{X}_{\tau-})$ und Y direkt in Beziehung zueinander zu setzen. Vielmehr sind die vier folgenden Fälle denkbar, bei denen auch jeweils eine etwas andere Fortsetzung in der Konstruktion benötigt wird. Es sind noch weitere Varianten mit $\tau < \tau_c = \infty$ denkbar, diese sind für die Beschreibung der Konstruktion jedoch unerheblich.



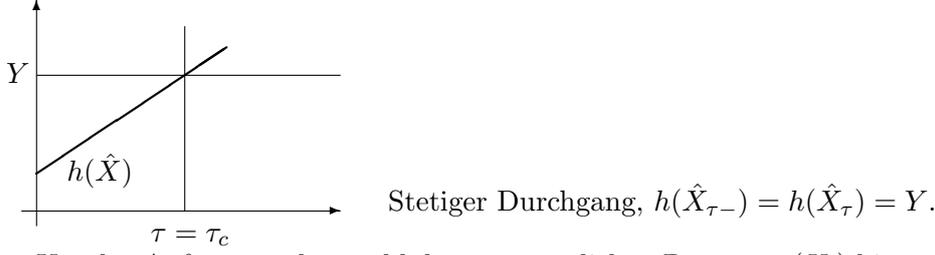
Sprung aus dem Supremum, $h(\hat{X}_\tau) < h(\hat{X}_{\tau-}) < Y$.



Sprung in das Supremum, $h(\hat{X}_{\tau-}) < h(\hat{X}_\tau) < Y$.



Sprung aus dem Supremum, $h(\hat{X}_{\tau-}) < Y < h(\hat{X}_\tau)$.



Um das Anfangsstück sowohl des ursprünglichen Prozesses (X_t) bis zum ersten Auftreten des globalen Supremum zur Zeit L , als auch das Anfangsstück des neu konstruierten Prozesses (\hat{X}_t) bis zum Zerlegungszeitpunkt τ vergleichen zu können unter Berücksichtigung dieser verschiedener Fälle, führen wir die folgende Notation ein:

$$X_{t\perp L} := \begin{cases} X_t & \text{falls } t < L \\ X_{L-} & \text{falls } t \geq L \text{ und } h(X_{L-}) \geq h(X_L) \\ X_L & \text{falls } t \geq L \text{ und } h(X_{L-}) < h(X_L) \end{cases} \quad (3.2)$$

$$\hat{X}_{t\perp\tau} := \begin{cases} \hat{X}_t & \text{falls } t < \tau \\ \hat{X}_{\tau-} & \text{falls } t \geq \tau \text{ und } h(\hat{X}_{\tau-}) \geq h(\hat{X}_\tau) \\ \hat{X}_{\tau-} & \text{falls } t \geq \tau \text{ und } h(\hat{X}_{\tau-}) \leq Y < h(\hat{X}_\tau) \\ \hat{X}_\tau & \text{falls } t \geq \tau \text{ und } h(\hat{X}_{\tau-}) < h(\hat{X}_\tau) \leq Y \end{cases} \quad (3.3)$$

Da nach Konstruktion auf dem Ereignis $\{\tau < \infty\}$ mindestens einer der Fälle $\{h(\hat{X}_\tau) \leq Y\}$ oder $\{h(\hat{X}_{\tau-}) \leq Y\}$ eintritt, ist die Fallunterscheidung in der Definition von $\hat{X}_{t\perp\tau}$ auch wirklich erschöpfend, $\{Y < h(X_{\tau-}) \leq h(X_\tau)\}$ kann nach Konstruktion nicht eintreten. Dies stellt eine für unsere Zwecke abgewandelte Form der gängigen Notation $(X_{t\wedge L})$ bzw. $(\hat{X}_{t\wedge\tau})$ dar. Den jeweils letzten Punkt dieser Prozesse bezeichnen wir mit

$$X_{\perp L} := X_{L\perp L} \in \{X_L, X_{L-}\} \quad \text{und} \quad \hat{X}_{\perp\tau} := \hat{X}_{\tau\perp\tau} \in \{\hat{X}_\tau, \hat{X}_{\tau-}\}.$$

Bemerkung Der so definierte Prozess $(X_{t\perp L})_{t \in \mathbb{R}_+}$ erzeugt eine ungewöhnliche σ -Algebra, denn es ist

$$\mathcal{F}_{\perp L} := \sigma((X_{t\perp L})) = \sigma(\mathcal{F}_L \cap \{h(X_L) > h(X_{L-})\}) \vee \sigma(\mathcal{F}_{L-} \cap \{h(X_L) \leq h(X_{L-})\}).$$

Bemerkenswerterweise liegt die σ -Algebra $\mathcal{F}_{\perp L}$ damit zwischen \mathcal{F}_{L-} und \mathcal{F}_L . \square

Hauptsatz

Die eigentliche Kernaussage der Zerlegung betrifft die Konstruktion des Anfangsstück bis zum Supremum L . Die Verteilung des post- L -Prozesses ist implizit als bedingter Prozess gegeben, und als solcher im Allgemeinen schwierig zu beschreiben; wir greifen hierfür auf die Resultate von Pittenger und Shih zurück, die ausführlich die Konstruktion dieser bedingten Verteilung untersucht haben. Damit bleibt als Hauptaufgabe die Darstellung des Anfangsstückes, mit der hier begonnen werden soll.

Um das Hauptaugenmerk auf die Konstruktion der Zerlegung und auf die Hauptresultate an sich zu richten, wird ein Großteil des Beweises der Zerlegung in den Abschnitt 3.3 verlagert.

Hauptsatz 30 Sei $(X_t)_{t \in \mathbb{R}_+}$ ein in o startender Markov-Prozess mit Kernhalbgruppe (P_t) und mit rechtsseitig stetigen Pfaden mit linken Limiten auf einem Wahrscheinlichkeitsraum mit Maß \mathbf{P} . Desweiteren sei h eine stetige, (P_t) -reguläre Funktion mit $h(o) > 0$ und L der Zeitpunkt des ersten Supremums von $(h(X_t))$, also

$$L := \inf\{t > 0 : h(X_t) \vee h(X_{t-}) = \sup_{s \geq 0} h(X_s)\}.$$

Sei $(\hat{X}_t)_{t \in \mathbb{R}_+}$ ein Markov-Prozess mit Startpunkt o mit den h -transformierten Übergangskernen (P_t^h) und mit rechtsseitig stetigen Pfaden mit linken Limiten auf einem Wahrscheinlichkeitsraum mit Maß \mathbf{P}^h , und sei zusätzlich Y eine unabhängige reelle Zufallsvariable mit der Verteilung

$$\mathbf{P}^h(Y > y) = \frac{h(o)}{y}, \quad y \geq h(o),$$

und sei die Zerlegungszeit τ definiert als

$$\tau := \inf\{t \in \mathbb{R}_+ : h(\hat{X}_t) \vee h(\hat{X}_{t-}) = \sup_{s < \tau_c} h(\hat{X}_s)\} \quad \tau_c := \inf\{t \in \mathbb{R}_+ : h(\hat{X}_t) > Y\}$$

Dann sind die Prozesse $((X_{t \perp L}), L)$ und $((\hat{X}_{t \perp \tau}), \tau)$ identisch verteilt.

Beweis Siehe Abschnitt 3.3. ■

Nachdem das Anfangsstück konstruiert worden ist, können wir einfach einen geeigneten bedingten Prozess mit der richtigen Eintrittsverteilung anhängen, und erhalten somit die vollständige Zerlegung.

Regularitätsannahme. Bevor die Zerlegung im folgenden Hauptsatz dargestellt wird, benötigen wir noch eine milde Regularitätsbedingung, die im Beweis des Satzes die Existenz der bedingten Eintrittsmaße Q_s^L sicherstellt. Sei hierfür eine weitere (im Beweis benötigte) terminale Zeit L° definiert durch

$$L^\circ := L \cdot I\{M_{L-} = M_L\},$$

wobei $I\{\cdot\}$ die Indikatorvariable eines Ereignisses bezeichnet. Die zugehörige terminale Zeit T° lässt sich darstellen als

$$T^\circ := \inf\{t > 0 : L_t \cdot I\{M_{L_t-} = M_{L_t}\} > 0\}.$$

Mit diesen Definitionen lässt sich die benötigte Regularitätsbedingung formulieren als

$$\mathbf{P}(X_{L-} \in S^\circ \mid M_{L-} = M_L) = 1 \quad \text{mit} \quad S^\circ := \{x \in S : \mathbf{P}_x(T^\circ > 0) = 0\}. \quad (3.4)$$

Die Menge S° bezeichnet hierbei die regulären Punkte bezüglich L° . Diese Bedingung besagt: Falls das Supremum des Prozesses $(h(X_t))$ durch einen linksseitigen Grenzwert angenommen wird, dann soll dieser Punkt regulär bezüglich T° sein.

Hauptsatz 31 (Zerlegung eines Markov-Prozesses) Sei h eine stetige, reguläre Funktion mit $h(o) > 0$ und sei $(\hat{X}_t)_{t \in \mathbb{R}_+}$ ein in o startender Markov-Prozess mit der h -transformierten

Kernfamilie (P_t^h) auf einem Wahrscheinlichkeitsraum mit Maß \mathbb{P} , und sei Y eine unabhängige reellwertige Zufallsvariable mit der Verteilung

$$\mathbb{P}(Y > y) = \frac{h(o)}{y}.$$

Die Zerlegungszeit τ wird definiert als

$$\tau := \inf\{t \in \mathbb{R}_+ : h(\hat{X}_t) \vee h(\hat{X}_{t-}) = \sup_{s < \tau_c} h(\hat{X}_s)\} \quad \tau_c := \inf\{t \in \mathbb{R}_+ : h(\hat{X}_t) > Y\}.$$

Für das Supremum \hat{M} des Prozesses $(h(\hat{X}_t))$ unterhalb der Schranke Y setze

$$\hat{M} := \begin{cases} h(\hat{X}_{\tau-}) & \text{falls } h(\hat{X}_\tau) \leq h(\hat{X}_{\tau-}) \leq Y \\ h(\hat{X}_{\tau-}) & \text{falls } h(\hat{X}_{\tau-}) < Y < h(\hat{X}_\tau) \\ h(\hat{X}_\tau) & \text{falls } h(\hat{X}_{\tau-}) < h(\hat{X}_\tau) \leq Y. \end{cases}$$

Damit wird der zusammengesetzte Prozess (\bar{X}_t) konstruiert durch

$$\bar{X}_t := \begin{cases} \hat{X}_t & t < \tau \\ \check{X}_{t-\tau} & t \geq \tau, \end{cases}$$

wobei die Verteilung des Prozesses (\bar{X}_t) gegeben ist durch die bedingte Kernhalbgruppe $(K_t^{\hat{M}})$ mit dem Eintrittsmaß

$$\begin{aligned} Q_s^{\hat{M}}(\hat{X}_{\tau-}, \cdot, \cdot) & \text{ falls } h(\hat{X}_\tau) \leq h(\hat{X}_{\tau-}) \leq Y \\ Q_s^{\hat{M}}(\hat{X}_{\tau-}, \cdot, \cdot) & \text{ falls } h(\hat{X}_{\tau-}) < Y < h(\hat{X}_\tau) \\ D_s^{\hat{M}}(\hat{X}_\tau, \cdot) & \text{ falls } h(\hat{X}_{\tau-}) < h(\hat{X}_\tau) \leq Y \end{aligned}$$

Unter der Annahme der Regularitätsbedingung (3.4) ist dann (\bar{X}_t) ein Markov-Prozess mit Kernfamilie (P_t) , d.h. der Prozess ist identisch verteilt mit einem in o startenden Markov-Prozess (X_t) mit Kernen (P_t) , und der Zeitpunkt τ entspricht in Verteilung dem Zeitpunkt L des ersten globalen Supremums des Prozesses $(h(X_t))$.

Beweis Die Zufallsvariable L bezeichne den ersten Zeitpunkt des globalen Supremums von $(h(X_t))$, d.h.

$$L := \inf\{t > 0 : h(X_{t-}) \vee h(X_t) = \sup_{s \geq 0} h(X_s)\},$$

damit ist L eine koterminale Zeit. Aus dem Hauptsatz 30 folgt zunächst, dass die Prozesse $((X_{t \perp L}), L)$ und $((\hat{X}_{t \perp \tau}), \tau)$ und die Suprema M_∞ und \hat{M} identisch verteilt sind. Aus der Definition dieser Prozesse folgt weiter, dass die beiden Ereignisse

$$F := \{h(X_{L-}) \geq h(X_L)\} \quad \text{und} \quad \hat{F} := \{h(\hat{X}_\tau) \leq h(\hat{X}_{\tau-}) \leq Y\} \cup \{h(\hat{X}_{\tau-}) < Y < h(\hat{X}_\tau)\}$$

ebenfalls gleichwahrscheinlich sind, denn es gilt

$$F = \{X_{\perp L} = X_{L-}\} \mathbf{P}\text{-f.s.} \quad \text{und} \quad \hat{F} = \{\hat{X}_{\perp \tau} = \hat{X}_{\tau-}\} \mathbb{P}\text{-f.s.} \quad (3.5)$$

Insgesamt erhalten wir die Verteilungsgleichheit

$$((X_{t\perp L}), M_\infty, I_F, L) \stackrel{\mathcal{L}}{=} ((\hat{X}_{t\perp\tau}), \hat{M}, I_{\hat{F}}, \tau), \quad (3.6)$$

wobei I_F und $I_{\hat{F}}$ die mit den jeweiligen Ereignissen assoziierten Indikatorvariablen bezeichnen.

Sei $(\hat{\mathcal{F}}_t)_{t \in \mathbb{R}_+}$ die dem neu konstruierten Prozess (\bar{X}_t) zugrunde liegende Filtration, und $\hat{\mathcal{F}}_{\perp\tau}$ bezeichne analog zu $\mathcal{F}_{\perp L}$ die σ -Algebra

$$\hat{\mathcal{F}}_{\perp\tau} := \sigma(\hat{\mathcal{F}}_{\tau-} \cap \hat{F}) \vee \sigma(\hat{\mathcal{F}}_{\tau} \cap \mathbb{C}\hat{F}).$$

Aus Gleichung (3.5) folgt, dass jeweils das Ereignis F in $\mathcal{F}_{\perp L}$ und \hat{F} in $\hat{\mathcal{F}}_{\perp\tau}$ liegt. Abhängig vom Eintreten des Ereignisses F müssen wir im Folgenden zwei Fälle unterscheiden, wie die Prozesse fortgesetzt werden.

Fall 1: Betrachten wir die zufällige Zeit L° definiert durch

$$L^\circ := L \cdot I\{M_{L-} = M_L\},$$

dann ist L° wieder eine koterminaler Zeit mit $L^\circ \in \{0, L\}$ fast sicher. Für alle $t \geq 0$ gilt $\{L^\circ > t\} = \{L > t\} \cap F$, und daraus folgt gemäß Definition der entsprechenden σ -Algebra die Gleichheit (wieder in etwas ungenauer Notation)

$$\mathcal{F}_{L^\circ-} = \sigma(\mathcal{F}_{L-} \cap F).$$

Da L° koterminal ist, existiert (unter der milden Regularitätsannahme (3.4), also $X_{L^\circ-} \in S^\circ$) fast sicher die auf $\mathcal{F}_{L^\circ-}$ bedingte Verteilung des post- L° -Prozesses $((X_{L+t}) | \mathcal{F}_{L^\circ-})$. Auf dem Ereignis $\{L^\circ = L < \infty\}$ und gegeben das Supremum $M = h(X_{L^\circ-})$ wird diese Verteilung durch die Kernhalbgruppe (K_t^M) mit dem Eintrittsmaß Q_s^M beschrieben, wie in Gleichung (3.1) definiert. Andererseits ist dies nach Konstruktion auch gerade die Verteilung des bedingten Prozesses $((\bar{X}_{\tau+t}) | \sigma(\hat{\mathcal{F}}_{\tau-} \cap \hat{F}))$, also gilt für den post- L - und für den post- τ -Prozess

$$((X_{L+t}) | \sigma(\mathcal{F}_{L-} \cap F)) \stackrel{\mathcal{L}}{=} ((X_{L+t}) | \mathcal{F}_{L^\circ-}) \stackrel{\mathcal{L}}{=} ((\bar{X}_{\tau+t}) | \sigma(\hat{\mathcal{F}}_{\tau-} \cap \hat{F})).$$

Da nach Gleichung (3.6) zusätzlich die Anfangsstücke $(X_{t\perp L} | F)$ und $(\bar{X}_{t\perp\tau} | \hat{F})$ zusammen mit I_F und $I_{\hat{F}}$ ebenfalls identisch verteilt sind, folgt für den gesamten Prozess

$$((X_t) | F) \stackrel{\mathcal{L}}{=} ((\bar{X}_t) | \hat{F}). \quad (3.7)$$

Fall 2: Auf den komplementären Ereignissen $\mathbb{C}F$ respektive $\mathbb{C}\hat{F}$ erhalten wir analog

$$((X_{L+t}) | \sigma(\mathcal{F}_L \cap \mathbb{C}F)) \stackrel{\mathcal{L}}{=} ((\bar{X}_{\tau+t}) | \sigma(\hat{\mathcal{F}}_{\tau} \cap \mathbb{C}\hat{F})),$$

wobei hier die bedingte Verteilung gegeben dem Supremum M durch die Kernhalbgruppe (K_t^M) mit dem Eintrittsmaß D_s^M beschrieben wird. Damit folgt wie im vorigen Fall

$$((X_t) | \mathbb{C}F) \stackrel{\mathcal{L}}{=} ((\bar{X}_t) | \mathbb{C}\hat{F}),$$

woraus zusammen mit (3.7) und $\mathbf{P}(F) = \mathbb{P}(\hat{F})$ die Behauptung folgt. ■

Falls der Prozess $h(X_t)$ fast sicher nur stetige Pfade besitzt, oder wenigstens nur negative Sprünge besitzt, dann überträgt sich diese Eigenschaft in natürlicher Weise auch auf den h -transformierten Prozess, somit wird die Schranke Y in der Konstruktion fast sicher stetig durchstoßen. Dies erlaubt die folgende wesentlich einfacherere Formulierung des Hauptsatzes 31.

Satz 32 (Halbstetige Zerlegung) *Sei h eine stetige, reguläre Funktion mit $h(o) > 0$, so dass $(h(X_t))$ \mathbf{P} -fast sicher keine positiven Sprünge besitzt. Sei $(\hat{X}_t)_{t \in \mathbb{R}_+}$ ein in o startender Markov-Prozess mit der h -transformierte Kernfamilie (P_t^h) auf einem Wahrscheinlichkeitsraum mit Maß \mathbb{P} . Desweiteren sei Y eine unabhängige reellwertige Zufallsvariable mit der Verteilung*

$$\mathbb{P}(Y > y) = \frac{h(o)}{y}.$$

Die Zerlegungszeit τ und das Supremum \hat{M} sind definiert als

$$\tau := \inf\{t \in \mathbb{R}_+ : h(\hat{X}_{t-}) = Y\} \quad \hat{M} := h(\hat{X}_{\tau-}).$$

Damit wird der zusammengesetzte Prozess (\bar{X}_t) konstruiert durch

$$\bar{X}_t := \begin{cases} \hat{X}_t & t < \tau \\ \check{X}_{t-\tau} & t \geq \tau \end{cases}$$

wobei die Verteilung des Prozesses (\bar{X}_t) gegeben ist durch die bedingte Kernhalbgruppe $(K_t^{\hat{M}})$ mit dem Eintrittsmaß $Q_s^{\hat{M}}(\hat{X}_{\tau-}, \cdot, \cdot)$.

Unter der Annahme der Regularitätsbedingung (3.4) ist dann (\bar{X}_t) ein Markov-Prozess mit Kernfamilie (P_t) , d.h. der Prozess ist identisch verteilt mit einem in o startenden Markov-Prozess (X_t) mit Kernen (P_t) , und der Zeitpunkt τ entspricht in Verteilung den Zeitpunkt L des ersten globalen Supremums des Prozesses $(h(X_t))$.

Duale Zerlegung

Analog zur zeitdiskreten Zerlegung existiert wiederum eine weitere, duale Zerlegung, die einen Prozess (X_t^*) mit h -transformierter Kernhalbgruppe (P_t^h) konstruiert. Die assoziierte Zerlegungszeit L^* beschreibt den ersten Zeitpunkt des globalen Infimums von $(h(X_t^*))$, d.h.

$$M_\infty^* := \inf_{s > 0} h(X_s^*) \quad L^* := \inf\{s \in \mathbb{R}_+ : \min(h(X_s^*), h(X_{s-}^*)) = M_\infty^*\}.$$

Die mit L^* und M_∞^* assoziierten bedingten Kerne zur Beschreibung des post- L^* -Prozesses seien wie folgt bezeichnet:

$$\begin{aligned} D_t^{*m}(x, f) &:= \mathbf{E}^h[f(X_t) \mid X_0 = x, h(X_s) \geq m \forall s \geq 0] \\ Q_t^{*m}(x, g, f) &:= \mathbf{E}^h[g(X_0)f(X_t) \mid X_{0-} = x, h(X_s) \geq m \forall s \geq 0] \\ K_t^{*m}(x, f) &:= \mathbf{E}^h[f(X_t) \mid X_0 = x, h(X_s) \geq m \forall s \geq 0]. \end{aligned}$$

Hauptsatz 33 (Duale Zerlegung) Sei $(\check{X}_t^*)_{t \in \mathbb{R}_+}$ ein in $o \in S$ startender Markov-Prozess mit Kernfamilie (P_t) und mit rechtsseitig stetigen Pfaden mit linksseitigen Limiten auf einem Wahrscheinlichkeitsraum mit Maß \mathbb{P} . Sei h eine stetige, reguläre Funktion bezüglich (P_t) mit $h(o) > 0$, und sei U eine unabhängige, uniform in $(0, h(o))$ verteilte Zufallsvariable. Definiere die Zerlegungszeit τ^* durch

$$\tau^* := \inf\{t \in \mathbb{R}_+ : \min(h(\check{X}_t^*), h(\check{X}_{t-}^*)) = \inf_{s < \tau_c^*} h(\check{X}_s^*)\} \quad \tau_c^* := \inf\{t \in \mathbb{R}_+ : h(\check{X}_t^*) < U\}.$$

Für das Infimum \check{M}^* oberhalb der Schranke U setze

$$\check{M}^* := \begin{cases} h(\check{X}_{\tau^*-}^*) & \text{falls } h(\check{X}_{\tau^*}^*) \geq h(\check{X}_{\tau^*-}^*) \geq U \\ h(\check{X}_{\tau^*-}^*) & \text{falls } h(\check{X}_{\tau^*-}^*) > U > h(\check{X}_{\tau^*}^*) \\ h(\check{X}_{\tau^*}^*) & \text{falls } h(\check{X}_{\tau^*-}^*) > h(\check{X}_{\tau^*}^*) \geq U. \end{cases}$$

Damit wird der zusammengesetzte Prozess (\bar{X}_t^*) konstruiert durch

$$\bar{X}_t^* := \begin{cases} \check{X}_t & t < \tau^* \\ \hat{X}_{t-\tau^*} & t \geq \tau^*, \end{cases}$$

wobei die Verteilung des Prozesses (\hat{X}_t) gegeben ist durch die bedingte Kernfamilie $K_s^{*\check{M}^*}$ mit dem Eintrittsmaß

$$\begin{aligned} Q_s^{*\check{M}^*}(\check{X}_{\tau^*-}, \cdot, \cdot) & \text{ falls } h(\check{X}_{\tau^*}^*) \geq h(\check{X}_{\tau^*-}^*) \geq U \\ Q_s^{*\check{M}^*}(\check{X}_{\tau^*-}, \cdot, \cdot) & \text{ falls } h(\check{X}_{\tau^*-}^*) > U > h(\check{X}_{\tau^*}^*) \\ D_s^{*\check{M}^*}(\check{X}_{\tau^*}, \cdot) & \text{ falls } h(\check{X}_{\tau^*-}^*) > h(\check{X}_{\tau^*}^*) \geq U. \end{aligned}$$

Unter der Annahme einer Regularitätsbedingung analog zu (3.4) ist dann (\bar{X}_t^*) ein Markov-Prozess mit Kernfamilie (P_t^h) , d.h. der Prozess ist identisch verteilt mit einem in o startenden Markov-Prozess (X_t^*) mit Kernen (P_t^h) , und der Zeitpunkt τ^* entspricht in Verteilung dem erste globale Infimum von $(h(X_t^*))$.

Beweis Siehe Beweis der zeitdiskreten dualen Zerlegung in Hauptsatz 4. ■

Bemerkung Die Zerlegungen funktionieren auch mit zeitinhomogenen Markov-Prozessen und entsprechenden zeitabhängigen regulären Funktionen, denn durch eine einfache Zustandsraumerweiterung $S' := S \times \mathbb{R}_+$ kann ein zeitinhomogener Prozess mit Zustandsraum S als zeithomogener Prozess mit Zustandsraum S' dargestellt werden, wobei die Zustände in S' zusätzlich um eine Zeitkomponente erweitert werden. Eine Funktion $h : \mathbb{R}_+ \times S \rightarrow \mathbb{R}_+$ heißt dann Raum-Zeit regulär, falls für alle $s, t > 0$ und $x \in S$ gilt

$$h(t, x) = \int_{\mathbb{R}^d} P_s(t, x, dy) h(s+t, y) \quad \text{bzw.} \quad \mathbf{E}[h(s+t, X_{s+t}) | X_t] = h(t, X_t)$$

wobei $P_s(t, x, dy)$ der von t abhängige Kern des zeitinhomogenen Prozesses bezeichnet. Aber auch im Falle eines zeithomogenen Prozesses führt eine Zeitabhängigkeit oft zu einer größeren Menge minimaler zeitabhängiger regulärer Funktionen.

Der einfacheren Notation wegen beschränken wir uns auf die zeithomogene Variante, als Beispiel für einen zeitinhomogenen Prozess sei auf die Konstruktion einer Brownschen Brücke verwiesen (siehe Abschnitt 4.3). □

3.3 Beweis Hauptsatz 30

Beweis (Von Hauptsatz 30) Der Beweis ist in mehrere Schritte eingeteilt, die ersten beiden diskretisieren den Prozess sowohl räumlich als auch zeitlich, die folgenden Schritte zeigen dann die Konvergenz der diskretisierten Zerlegungen.

Schritt 1: ε -Approximation. Zunächst sollen der Zeitpunkt des Supremums L und die Zerlegungszeit τ räumlich um eine ε -Ungenauigkeit verschmiert werden. Dies wird für festes $\varepsilon > 0$ durch eine induktive Konstruktion erreicht:

$$\begin{aligned} \sigma_0^\varepsilon &:= 0 & H_0^\varepsilon &:= h(X_0) \\ \sigma_i^\varepsilon &:= \inf\{t > \sigma_{i-1}^\varepsilon : h(X_t) \geq H_{i-1}^\varepsilon + \varepsilon\} & H_i^\varepsilon &:= h(X_{\sigma_i^\varepsilon}). \end{aligned}$$

Die Stoppzeiten σ_i^ε heißen ε -LEITERZEITEN und die Werte H_i^ε heißen ε -LEITERHÖHEN des Prozesses $(h(X_t))_{t \in \mathbb{R}_+}$, die Tupel $(\sigma_i^\varepsilon, H_i^\varepsilon)$ schließlich heißen ε -LEITERPUNKTE. Als Approximation des Zeitpunkts L des ersten Supremums dient

$$L^\varepsilon := \sup\{\sigma_i^\varepsilon : \sigma_i^\varepsilon < \infty\}.$$

Analog ist für den h -transformierten Prozess (\hat{X}_t) definiert:

$$\begin{aligned} \hat{\sigma}_0^\varepsilon &:= 0 & \hat{H}_0^\varepsilon &:= h(\hat{X}_0) \\ \hat{\sigma}_i^\varepsilon &:= \inf\{t > \hat{\sigma}_{i-1}^\varepsilon : h(\hat{X}_t) \geq \hat{H}_{i-1}^\varepsilon + \varepsilon\} & \hat{H}_i^\varepsilon &:= h(\hat{X}_{\hat{\sigma}_i^\varepsilon}), \end{aligned}$$

und als Approximation der Zerlegungszeit τ dient nun

$$\tau^\varepsilon := \sup\{\hat{\sigma}_i^\varepsilon : \hat{\sigma}_i^\varepsilon < \infty \text{ und } \hat{H}_i^\varepsilon \leq Y\}.$$

Für die Zeitpunkte L^ε und τ^ε folgt aus der Konstruktion

$$0 \leq \sup_{t \in \mathbb{R}_+} h(X_t) - h(X_{L^\varepsilon}) \leq \varepsilon \quad \text{falls } L^\varepsilon < \infty \quad (3.8)$$

$$\text{und } 0 \leq \sup_{0 \leq t < \tau^\varepsilon} h(\hat{X}_t) - h(\hat{X}_{\tau^\varepsilon}) \leq \varepsilon \quad \text{falls } \tau^\varepsilon < \infty. \quad (3.9)$$

Die Bedingung $\tau^\varepsilon < \infty$ in Gleichung (3.9) ist hier allerdings keine echte Einschränkung, da nach Konstruktion $\tau^\varepsilon < \infty$ fast sicher gilt. Denn entweder ist der Prozess $(h(\hat{X}_t))$ nach oben unbeschränkt, und überschreitet damit fast sicher in endlicher Zeit die Schranke Y , oder er ist nach oben beschränkt, und damit gibt es fast sicher einen letzten und größten ε -Leiterpunkt. Beide Fälle implizieren $\tau^\varepsilon < \infty$. Die Bedingung $L^\varepsilon < \infty$ ist ebenfalls fast sicher erfüllt, da $(h(X_t))$ ein nicht-negatives Martingal ist, und als solches fast sicher konvergiert. Also sind auch diese Pfade fast sicher beschränkt und somit $L^\varepsilon < \infty$.

Schritt 2: Zeitdiskretisierung. Im zweiten Schritt wird die ε -Approximation zusätzlich zeitlich diskretisiert, so dass wir dann mit den schon bekannten diskreten Methoden arbeiten können. Als Ausgangspunkt für die Zeitdiskretisierung dienen der jeweils ursprüngliche Prozess, allerdings werden nur Stützstellen im zeitlichen Abstand 2^{-n} betrachtet, d.h. aus dem Prozess $(X_t)_{t \in \mathbb{R}_+}$ wird im n -ten Diskretisierungsschritt der Prozess $(X_{2^{-n}i})_{i \in \mathbb{N}}$ mit Übergangskern $P_{2^{-n}}$. Die ε -Leiterpunkte der zeitlich diskretisierten Prozesse sind gegeben durch

$$\begin{aligned} \sigma_0^{\varepsilon,n} &:= 0 & H_0^{\varepsilon,n} &:= h(X_0) \\ \sigma_i^{\varepsilon,n} &:= \inf\{2^{-n}j > \sigma_{i-1}^{\varepsilon,n} : h(X_{2^{-n}j}) > H_{i-1}^{\varepsilon,n} + \varepsilon\} & H_i^{\varepsilon,n} &:= h(X_{\sigma_i^{\varepsilon,n}}), \end{aligned}$$

und die diskrete Variante des approximativen Zeitpunkts L^ε des ersten Supremums ist

$$L^{\varepsilon,n} := \sup\{\sigma_i^{\varepsilon,n} : \sigma_i^{\varepsilon,n} < \infty\}.$$

Analog ist für den h -transformierten Prozess (\hat{X}_t) definiert:

$$\begin{aligned} \hat{\sigma}_0^{\varepsilon,n} &:= 0 & \hat{H}_0^{\varepsilon,n} &:= h(\hat{X}_0) \\ \hat{\sigma}_i^{\varepsilon,n} &:= \inf\{2^{-n}j > \hat{\sigma}_{i-1}^{\varepsilon,n} : h(\hat{X}_{2^{-n}j}) > \hat{H}_{i-1}^{\varepsilon,n} + \varepsilon\} & \hat{H}_i^{\varepsilon,n} &:= h(\hat{X}_{\hat{\sigma}_i^{\varepsilon,n}}), \end{aligned}$$

und die diskrete Variante der approximativen Zerlegungszeit τ^ε ist nun

$$\tau^{\varepsilon,n} := \sup\{\hat{\sigma}_i^{\varepsilon,n} : \hat{\sigma}_i^{\varepsilon,n} < \infty \text{ und } \hat{H}_i^{\varepsilon,n} \leq Y\}.$$

Behauptung: Mit dieser Notation gilt die folgende Verteilungsgleichheit:

$$((X_{2^{-n}i \wedge L^{\varepsilon,n}})_{i \in \mathbb{N}}, L^{\varepsilon,n}) \stackrel{\mathcal{L}}{=} ((\hat{X}_{2^{-n}i \wedge \tau^{\varepsilon,n}})_{i \in \mathbb{N}}, \tau^{\varepsilon,n}) \quad \text{für alle } n \in \mathbb{N} \text{ und } \varepsilon > 0. \quad (3.10)$$

Für den Beweis der Behauptung betrachten wir für ein $k \in \mathbb{N}$ jeweils die zwei Ereignisse $\{L^{\varepsilon,n} = 2^{-n}k\}$ und $\{L^{\varepsilon,n} > 2^{-n}k\}$. Es ergeben sich die beiden folgenden Fälle.

Fall 2.1: $L^{\varepsilon,n} = 2^{-n}k$. Auf diesem Ereignis ist der Zeitpunkt $2^{-n}k$ die letzte endliche ε -Leiterzeit, und außerdem ist dann $h(X_{2^{-n}j}) \leq h(X_{2^{-n}k}) + \varepsilon$ für alle $j > k$. In mathematischer Schreibweise heißt das für $B \in \mathcal{S}^k$

$$\begin{aligned} &\{(X_{2^{-n}1}, \dots, X_{2^{-n}k}) \in B \text{ und } L^{\varepsilon,n} = 2^{-n}k\} \\ &= \{(X_{2^{-n}1}, \dots, X_{2^{-n}k}) \in B \cap B_k^\varepsilon \text{ und } h(X_{2^{-n}j}) \leq h(X_{2^{-n}k}) + \varepsilon \forall j > k\} \end{aligned}$$

mit

$$B_k^\varepsilon := \{(x_1, \dots, x_k) : (k, h(x_k)) \text{ ist } \varepsilon\text{-Leiterpunkt in der Folge } h(o), h(x_1), \dots, h(x_k)\}.$$

Es folgt damit

$$\begin{aligned} &\mathbf{P}_o((X_{2^{-n}1}, \dots, X_{2^{-n}k}) \in B \text{ und } L^{\varepsilon,n} = 2^{-n}k) \\ &= \int_{B \cap B_k^\varepsilon} P_{2^{-n}}(o, dx_1) \dots P_{2^{-n}}(x_{k-1}, dx_k) q_{h(x_k) + \varepsilon}^n(x_k), \end{aligned} \quad (3.11)$$

wobei die Funktion $q_s^n : S \rightarrow [0, 1]$ definiert ist als

$$q_s^n(x) := \mathbf{P}_x(h(X_{2^{-n}i}) \leq s \forall i \in \mathbb{N}_0).$$

Andererseits tritt das Ereignis $\tau^{\varepsilon,n} = 2^{-n}k$ genau dann ein, wenn $2^{-n}k$ die letzte endliche ε -Leiterzeit des diskretisierten Prozesses $(\hat{X}_{2^{-n}i})$ ist mit $h(\hat{X}_{2^{-n}k}) \leq Y$, also

$$\begin{aligned} &\{(\hat{X}_{2^{-n}1}, \dots, \hat{X}_{2^{-n}k}) \in B \text{ und } \tau^{\varepsilon,n} = 2^{-n}k\} \\ &= \{(\hat{X}_{2^{-n}1}, \dots, \hat{X}_{2^{-n}k}) \in B \cap B_k^\varepsilon\} \\ &\quad \cap \{h(\hat{X}_{2^{-n}k}) \leq Y \text{ und der folgende } \varepsilon\text{-Leiterpunkt liegt über } Y\} \\ &\quad \cup \{h(\hat{X}_{2^{-n}k}) \leq Y \text{ und die folgende } \varepsilon\text{-Leiterzeit ist unendlich}\}. \end{aligned}$$

Hiermit folgt nun

$$\begin{aligned} & \mathbf{P}_o^h((\hat{X}_{2^{-n}1}, \dots, \hat{X}_{2^{-n}k}) \in B, \tau^{\varepsilon,n} = 2^{-n}k) \\ &= \int_{B \cap B_k^\varepsilon} P_{2^{-n}}^h(o, dx_1) \dots P_{2^{-n}}^h(x_{k-1}, dx_k) \\ & \mathbf{P}_{x_k}^h\left(\frac{1}{\hat{H}_1^{\varepsilon,n}} I_{\{\hat{\sigma}_1^{\varepsilon,n} < \infty\}} < \frac{1}{Y} \leq \frac{1}{h(\hat{X}_k)}\right) \end{aligned}$$

Ein Wechsel von P^h nach P liefert

$$\begin{aligned} &= \int_{B \cap B_k^\varepsilon} P_{2^{-n}}(o, dx_1) \dots P_{2^{-n}}(x_{k-1}, dx_k) \frac{h(x_k)}{h(o)} \\ & \mathbf{P}_{x_k}^h\left(\frac{1}{\hat{H}_1^{\varepsilon,n}} I_{\{\hat{\sigma}_1^{\varepsilon,n} < \infty\}} < \frac{1}{Y} \leq \frac{1}{h(\hat{X}_k)}\right). \end{aligned}$$

An dieser Stelle kommt die zweite Aussage des in Kapitel 1 aufgetretenen Lemmas 8 zum Zuge, das hier in der folgenden Form benötigt wird:

$$\mathbf{P}_x^h\left(\frac{1}{\hat{H}_1^{\varepsilon,n}} I_{\{\hat{\sigma}_1^{\varepsilon,n} < \infty\}} < \frac{1}{Y} \leq \frac{1}{h(x)}\right) = \frac{h(o)}{h(x)} q_{h(x)+\varepsilon}^n(x).$$

Zusammen mit der letzten Gleichung und Gleichung (3.11) folgt damit

$$\begin{aligned} & \mathbf{P}_o((X_{2^{-n}1}, \dots, X_{2^{-n}k}) \in B, L^{\varepsilon,n} = 2^{-n}k) \\ &= \mathbf{P}_o^h((\hat{X}_{2^{-n}1}, \dots, \hat{X}_{2^{-n}k}) \in B, \tau^{\varepsilon,n} = 2^{-n}k). \end{aligned} \quad (3.12)$$

Fall 2.2: $L^{\varepsilon,n} > 2^{-n}k$. Um mit den ε -Leiterpunkten besser arbeiten zu können, führen wir die folgende Funktionsfamilie H^ε ein, die die größten ε -Leiterhöhen unter den $k+1$ Werten o, x_1, \dots, x_k bestimmt:

$$\begin{aligned} k=0: & \quad H^\varepsilon() := h(o) \\ k>0: & \quad H^\varepsilon(x_1, \dots, x_k) := \begin{cases} H^\varepsilon(x_1, \dots, x_{k-1}) & \text{falls } h(x_k) \leq H^\varepsilon(x_1, \dots, x_{k-1}) + \varepsilon \\ h(x_k) & \text{falls } h(x_k) > H^\varepsilon(x_1, \dots, x_{k-1}) + \varepsilon, \end{cases} \end{aligned}$$

und weiter sei $\sigma^{\varepsilon,n}(x_1, \dots, x_k)$ bzw. $\hat{\sigma}^{\varepsilon,n}(x_1, \dots, x_k)$ die nächste ε -Leiterzeit in Abhängigkeit von den ersten k Punkten, d.h.

$$\sigma^{\varepsilon,n}(x_1, \dots, x_k) := \inf\{2^{-ni} : h(X_{2^{-ni}}) > H^\varepsilon(x_1, \dots, x_k) + \varepsilon\}.$$

Mit Hilfe dieser Definitionen erhalten wir

$$\begin{aligned} & \{(X_{2^{-n}1}, \dots, X_{2^{-n}k}) \in B, L^{\varepsilon,n} > k\} \\ &= \{(X_{2^{-n}1}, \dots, X_{2^{-n}k}) \in B\} \cap \{h(X_{2^{-ni}}) > H^\varepsilon(X_{2^{-n}1}, \dots, X_{2^{-n}k}) + \varepsilon \text{ für ein } i > k\}, \end{aligned}$$

aus dieser Gleichheit folgt zunächst für den Prozess $(X_{2^{-ni}})$

$$\begin{aligned} & \mathbf{P}_o((X_{2^{-n}1}, \dots, X_{2^{-n}k}) \in B, L^{\varepsilon,n} > k) \\ &= \int_B P_{2^{-n}}(o, dx_1) \dots P_{2^{-n}}(x_{k-1}, dx_k) \mathbf{P}_{x_k}(\sigma^{\varepsilon,n}(x_1, \dots, x_k) < \infty). \end{aligned} \quad (3.13)$$

Andererseits gilt für den Prozess (\hat{X}_{2-n_i})

$$\begin{aligned} & \{(\hat{X}_{2-n_1}, \dots, \hat{X}_{2-n_k}) \in B, \tau^{\varepsilon, n} > k\} \\ &= \{(\hat{X}_{2-n_1}, \dots, \hat{X}_{2-n_k}) \in B\} \\ & \quad \cap \{\hat{\sigma}^{\varepsilon, n}(\hat{X}_{2-n_1}, \dots, \hat{X}_{2-n_k}) < \infty\} \\ & \quad \cap \{Y \geq h(\hat{X}_{\hat{\sigma}^{\varepsilon, n}(\hat{X}_{2-n_1}, \dots, \hat{X}_{2-n_k})})\} \end{aligned}$$

und damit folgt

$$\begin{aligned} & \mathbf{P}_o^h((\hat{X}_{2-n_1}, \dots, \hat{X}_{2-n_k}) \in B, \tau^{\varepsilon, n} > k) \\ &= \int_B P_{2-n}^h(o, dx_1) \dots P_{2-n}^h(x_{k-1}, dx_k) \\ & \quad \mathbf{P}_o^h\left(\frac{1}{Y} \leq \frac{1}{h(\hat{X}_{\hat{\sigma}^{\varepsilon, n}(x_1, \dots, x_k)})} \cdot I_{\{\hat{\sigma}^{\varepsilon, n}(x_1, \dots, x_k) < \infty\}} \mid \hat{X}_{2-n_1} = x_1, \dots, \hat{X}_{2-n_k} = x_k\right) \\ &= \int_B P_{2-n}(o, dx_1) \dots P_{2-n}(x_{k-1}, dx_k) \frac{h(x_k)}{h(o)} \\ & \quad \mathbf{P}_{x_k}^h\left(\frac{1}{Y} \leq \frac{1}{h(\hat{X}_{\hat{\sigma}^{\varepsilon, n}(x_1, \dots, x_k)})} \cdot I_{\{\hat{\sigma}^{\varepsilon, n}(x_1, \dots, x_k) < \infty\}}\right). \end{aligned}$$

Hier wird die erste Aussage von Lemma 8 in der folgenden Form eingesetzt:

$$\mathbf{P}_{x_k}^h\left(\frac{1}{Y} \leq \frac{1}{h(\hat{X}_{\hat{\sigma}^{\varepsilon, n}(x_1, \dots, x_k)})} \cdot I_{\{\hat{\sigma}^{\varepsilon, n}(x_1, \dots, x_k) < \infty\}}\right) = \frac{h(o)}{h(x_k)} \mathbf{P}_{x_k}(\sigma^{\varepsilon, n}(x_1, \dots, x_k) < \infty),$$

und somit erhalten zusammen mit Gleichung (3.13)

$$\begin{aligned} & \mathbf{P}_o((X_{2-n_1}, \dots, X_{2-n_k}) \in B, L^{\varepsilon, n} > 2^{-n}k) \\ &= \mathbf{P}_o^h((\hat{X}_{2-n_1}, \dots, \hat{X}_{2-n_k}) \in B, \tau^{\varepsilon, n} > 2^{-n}k). \end{aligned} \quad (3.14)$$

Aus beiden Fällen erhalten wir insgesamt die in Gleichung (3.10) behauptete Verteilungsideutität:

$$((X_{2-n_i \wedge L^{\varepsilon, n}})_{i \in \mathbb{N}}, L^{\varepsilon, n}) \stackrel{\mathcal{L}}{=} ((\hat{X}_{2-n_i \wedge \tau^{\varepsilon, n}})_{i \in \mathbb{N}}, \tau^{\varepsilon, n}) \quad \text{für alle } n \in \mathbb{N} \text{ und } \varepsilon > 0.$$

Schritt 3: Zulässige ε . Ziel soll es nun sein, dass sowohl die ε -Approximationen als auch die zeitdiskretisierten Zerlegungszeiten $L^{\varepsilon, n}$ und $\tau^{\varepsilon, n}$ gegen L bzw. gegen τ konvergieren. Dies ist jedoch nicht für jede Wahl von ε der Fall. In diesem Schritt werden einige Kriterien für zulässige ε aufgestellt, die später im Beweis die Konvergenz sicherstellen. Außerdem wird gezeigt, dass die Menge der störenden ε fast sicher nur abzählbar ist.

Wir ziehen uns in diesem Schritt auf den Prozess (X_t) zurück, exakt die gleiche Konstruktion funktioniert auch für den Prozess (\hat{X}_t) . Für einen festen Pfad (X_t) sollen nun diejenigen $\varepsilon > 0$ von den Konvergenzbetrachtungen ausgeschlossen werden, für die gilt

$$\begin{aligned} & H_i^\varepsilon + \varepsilon \in \mathcal{W} \quad \text{für ein } i \in \mathbb{N} \\ \text{mit } \mathcal{W} &:= \left\{ \sup_{t \in \mathbb{R}_+} h(X_t) \right\} \\ & \quad \cup \{X_s, X_{s-} : X_s \neq X_{s-} \text{ und } s \in \mathbb{R}_+\} \\ & \quad \cup \{X_s : s \text{ lokales Maximum oder } s \text{ lokal konstante Stelle}\}, \end{aligned}$$

d.h. es sollen ε -Leiterpunkte ausgeschlossen werden, die exakt auf Unstetigkeitsstellen, lokale Maxima oder lokal konstante Stellen fallen. Die Menge \mathcal{W} selbst ist pfadabhängig und damit zufällig, und sie ist fast sicher abzählbar.

Bezeichne $\mathcal{Z}_0 := \mathbb{R}_+$ und bezeichne \mathcal{Z}_1 die Menge aller ε , so dass $H_1^\varepsilon + \varepsilon$ nicht in \mathcal{W} liegt. Induktiv bezeichne dann \mathcal{Z}_i die Menge aller $\varepsilon > 0$, so dass $H_1^\varepsilon + \varepsilon, \dots, H_i^\varepsilon + \varepsilon$ allesamt nicht in \mathcal{W} liegen, also

$$\mathcal{Z}_i := \bigcap_{j=0 \dots i} \{\varepsilon > 0 : H_j^\varepsilon + \varepsilon \notin \mathcal{W}\} \quad \text{und setze} \quad \mathcal{Z} := \bigcap_{i \in \mathbb{N}_0} \mathcal{Z}_i = \mathcal{Z}_0 \setminus \bigcup_{i \in \mathbb{N}} (\mathcal{Z}_{i-1} \setminus \mathcal{Z}_i). \quad (3.15)$$

Diese Mengen \mathcal{Z}_i bilden eine absteigende Folge, es stellt sich aber die Frage, ob \mathcal{Z} überhaupt noch genügend viele zulässige ε enthält. Die Menge \mathcal{W} enthält zwar fast sicher nur abzählbar viele Werte, aber es ist per se nicht ausgeschlossen, dass für verschiedene ε der jeweils i -te Leiterpunkt H_i^ε einen gleichen Wert in \mathcal{W} annehmen kann. Seien $\varepsilon, \varepsilon' \in \mathcal{Z}_{i-1} \setminus \mathcal{Z}_i$, d.h.

$$H_i^\varepsilon + \varepsilon \in \mathcal{W} \quad \text{und} \quad H_i^{\varepsilon'} + \varepsilon' \in \mathcal{W}.$$

Ohne Beschränkung der Allgemeinheit sei $\varepsilon < \varepsilon'$, dann folgt aus der Konstruktion $H_i^\varepsilon \leq H_i^{\varepsilon'}$, und damit folgt

$$H_i^\varepsilon + \varepsilon \leq H_i^{\varepsilon'} + \varepsilon < H_i^{\varepsilon'} + \varepsilon'.$$

Also sind für verschiedene $\varepsilon, \varepsilon' \in \mathcal{Z}_{i-1} \setminus \mathcal{Z}_i$ auch die Werte $H_i^\varepsilon + \varepsilon$ und $H_i^{\varepsilon'} + \varepsilon'$ unterschiedlich. Diese liegen jedoch nach Voraussetzung in der abzählbaren Menge \mathcal{W} , also liegen auch nur abzählbar viele ε in $\mathcal{Z}_{i-1} \setminus \mathcal{Z}_i$. Somit werden in Gleichung (3.15) auch nur abzählbar viele Werte aus \mathcal{Z}_0 entfernt, d.h. das Komplement (in \mathbb{R}_+) der Menge \mathcal{Z} besitzt Lebesgue-Maß 0.

Eine wichtige Eigenschaft der Menge \mathcal{Z} ist, dass für alle $\varepsilon \in \mathcal{Z}$ und $\sigma_i^\varepsilon < \infty$ gilt

$$\sup_{s < t} h(X_s) < H_i^\varepsilon \quad \text{für alle } t < \sigma_i^\varepsilon, \quad (3.16)$$

d.h. der nächste ε -Leiterpunkt kann nicht schon früher durch einen linksseitigen Grenzwert erreicht werden, dies widerspräche der Definition von \mathcal{Z} .

Wie oben bemerkt, wurde hier nur die zulässige Menge \mathcal{Z} für den Pfad (X_t) konstruiert, die gleiche Konstruktion liefert eine weitere Menge $\hat{\mathcal{Z}}$ zulässiger ε für den Prozess (\hat{X}_t) , für die ebenfalls in analoger Weise gilt

$$\sup_{s < t} h(\hat{X}_s) < \hat{H}_i^\varepsilon \quad \text{für alle } t < \hat{\sigma}_i^\varepsilon < \infty.$$

Schritt 4: Konvergenz der zeitdiskreten ε -Leiterpunkte. In diesem Schritt ist zu zeigen, dass für $\varepsilon \in \mathcal{Z}$ die zeitdiskreten ε -Leiterpunkte konvergieren, d.h. dass für alle $i \in \mathbb{N}$ gilt

$$\lim_{n \rightarrow \infty} \sigma_i^{\varepsilon, n} = \sigma_i^\varepsilon \mathbf{P}_o - \text{f.s.} \quad \text{und} \quad \lim_{n \rightarrow \infty} H_i^{\varepsilon, n} = H_i^\varepsilon \mathbf{P}_o - \text{f.s.}$$

Analoge Aussagen gelten auch für den Prozess (\hat{X}_t) , wir ziehen uns hier auf (X_t) zurück, da der Beweis nur auf Pfadeigenschaften und nicht auf Verteilungsargumente baut.

Für $i = 0$ ist die Behauptung klar, betrachten wir also $\sigma_{i+1}^{\varepsilon, n}$ unter der Annahme, dass

$$\lim_{n \rightarrow \infty} \sigma_j^{\varepsilon, n} = \sigma_j^\varepsilon < \infty \quad \text{und} \quad \lim_{n \rightarrow \infty} H_j^{\varepsilon, n} = H_j^\varepsilon \quad \text{für alle } j < i + 1$$

gilt. Für jedes n ist der nächste Leiterpunkt gegeben durch

$$\sigma_{i+1}^{\varepsilon,n} = \inf\{2^{-n}j > \sigma_i^{\varepsilon,n} : h(X_{2^{-n}j}) > H_i^{\varepsilon,n} + \varepsilon\} \quad \text{und} \quad H_{i+1}^{\varepsilon,n} = h(X_{\sigma_{i+1}^{\varepsilon,n}}),$$

wobei $H_{i+1}^{\varepsilon,n}$ nur dann wohldefiniert ist, falls $\sigma_{i+1}^{\varepsilon,n} < \infty$ ist.

Fall 4.1: $\sigma_{i+1}^{\varepsilon} = \infty$. Betrachten wir zunächst den Fall $\sigma_{i+1}^{\varepsilon} = \infty$, dann folgt aus dieser Bedingung nach Definition der ε -Leiterpunkte

$$\sup_{t > \sigma_i^{\varepsilon}} h(X_t) \leq \sup_{t \in \mathbb{R}_+} h(X_t) \leq H_i^{\varepsilon} + \varepsilon,$$

wegen $\varepsilon \in \mathcal{Z}$ gilt sogar

$$\sup_{t \in \mathbb{R}_+} h(X_t) < H_i^{\varepsilon} + \varepsilon. \quad \text{Setze} \quad \alpha := \alpha(\varepsilon) := (H_i^{\varepsilon} + \varepsilon) - \sup_{t \in \mathbb{R}_+} h(X_t) > 0.$$

Da nach Induktionsvoraussetzung $H_i^{\varepsilon,n} \rightarrow H_i^{\varepsilon}$ für $n \rightarrow \infty$, ist für alle hinreichend große n

$$(H_i^{\varepsilon,n} + \varepsilon) - \sup_{t \in \mathbb{R}_+} h(X_t) > \alpha/2 \quad \text{d.h.} \quad H_i^{\varepsilon,n} + \varepsilon > \sup_{t \in \mathbb{R}_+} h(X_t),$$

und damit ist $\sigma_{i+1}^{\varepsilon,n} = \infty$ für alle hinreichend große n , also

$$\lim_{n \rightarrow \infty} \sigma_{i+1}^{\varepsilon,n} = \sigma_{i+1}^{\varepsilon} = \infty.$$

Es gilt dann natürlich auch $\sigma_j^{\varepsilon,n} = \sigma_j^{\varepsilon} = \infty$ für alle hinreichend großen n und alle $j > i$.

Fall 4.2: $\sigma_{i+1}^{\varepsilon} < \infty$. Sei nun $\sigma_{i+1}^{\varepsilon} < \infty$. Es gilt dann wegen $\varepsilon \in \mathcal{Z}$ und der Minimalität der ε -Leiterzeit $\sigma_{i+1}^{\varepsilon}$ für alle $t < \sigma_{i+1}^{\varepsilon}$

$$(H_i^{\varepsilon} + \varepsilon) > \sup_{s \leq t} h(X_s). \quad \text{Setze} \quad \alpha := \alpha(\varepsilon, t) := (H_i^{\varepsilon} + \varepsilon) - \sup_{s \leq t} h(X_s) > 0.$$

Für $t < \sigma_{i+1}^{\varepsilon}$ und alle hinreichend große n ist wegen der Induktionsvoraussetzung $|H_i^{\varepsilon,n} - H_i^{\varepsilon}| < \alpha/2$, und damit folgt

$$H_i^{\varepsilon,n} + \varepsilon \geq H_i^{\varepsilon} - \alpha/2 + \varepsilon = \sup_{s \leq t} h(X_s) + \alpha/2 > \sup_{s \leq t} h(X_s).$$

Für alle hinreichend große n ist deshalb die $(i+1)$ -te Leiterzeit $\sigma_{i+1}^{\varepsilon,n} > t$, und da $t \in [0, \sigma_{i+1}^{\varepsilon})$ beliebig gewählt war, erhalten wir insgesamt die Ungleichung

$$\liminf_{n \rightarrow \infty} \sigma_{i+1}^{\varepsilon,n} \geq \sigma_{i+1}^{\varepsilon}. \quad (3.17)$$

Durch die Wahl von ε aus \mathcal{Z} gilt im Falle, dass $\sigma_{i+1}^{\varepsilon}$ ein lokales Maximum, eine Unstetigkeitsstelle oder eine lokal konstante Stelle von $(h(X_t))$ ist, die strikte Ungleichung $H_{i+1}^{\varepsilon} > H_i^{\varepsilon} + \varepsilon$. In allen anderen Fällen gilt mindestens noch $H_{i+1}^{\varepsilon} \geq H_i^{\varepsilon} + \varepsilon$, und in jeder Umgebung um $\sigma_{i+1}^{\varepsilon}$ existieren Werte größer der Leiterhöhe H_{i+1}^{ε} , da dann nach Voraussetzung $\sigma_{i+1}^{\varepsilon}$ kein lokales Maximum ist. Es existiert somit in allen Fällen zu jedem $\delta > 0$ stets ein t_{δ} mit

$$t_{\delta} \in [\sigma_{i+1}^{\varepsilon} - \delta, \sigma_{i+1}^{\varepsilon} + \delta] \quad \text{und} \quad h(X_{t_{\delta}}) > h(X_{\sigma_{i+1}^{\varepsilon}}) + \varepsilon = H_i^{\varepsilon} + \varepsilon. \quad (3.18)$$

Im Falle einer Unstetigkeitsstelle oder einer lokal konstanten Stelle erfüllt schon einfach $t_\delta = \sigma_{i+1}^\varepsilon$ die Ungleichung (3.18). Setzen wir nun

$$\alpha := \alpha(\varepsilon, \delta) := h(X_{t_\delta}) - (H_i^\varepsilon + \varepsilon) > 0,$$

dann existiert wegen der rechtsseitigen Stetigkeit des Pfades ein $0 < \gamma < \delta$ mit

$$|h(X_u) - h(X_{t_\delta})| < \alpha/3 \quad \text{für alle } u \in [t_\delta, t_\delta + \gamma)$$

Für alle hinreichend große n gilt nach Induktionsvoraussetzung $|H_i^{\varepsilon, n} - H_i^\varepsilon| < \alpha/3$, und damit folgt für alle $u \in [t_\delta, t_\delta + \gamma)$ und alle hinreichend große n

$$\begin{aligned} h(X_u) - H_i^{\varepsilon, n} &= (h(X_u) - h(X_{t_\delta})) + (h(X_{t_\delta}) - H_i^\varepsilon) + (H_i^\varepsilon - H_i^{\varepsilon, n}) \\ &\geq -\alpha/3 + (\alpha + \varepsilon) - \alpha/3 \\ &> \varepsilon. \end{aligned}$$

Also sind alle u aus dem Intervall $[t_\delta, t_\delta + \gamma)$ gute Kandidaten für den nächsten ε -Leiterzeitpunkt, d.h. für alle hinreichend große n folgt

$$\sigma_{i+1}^{\varepsilon, n} \leq t_\delta + \gamma \leq \sigma_{i+1}^\varepsilon + \delta + \gamma \leq \sigma_{i+1}^\varepsilon + 2\delta.$$

Diese Ungleichung bleibt im Übergang zu $\limsup_{n \rightarrow \infty} \sigma_{i+1}^{\varepsilon, n}$ erhalten und da zudem $\delta > 0$ beliebig war, folgt

$$\limsup_{n \rightarrow \infty} \sigma_{i+1}^{\varepsilon, n} \leq \sigma_{i+1}^\varepsilon.$$

Zusammen mit der ersten Ungleichung (3.17) erhalten wir schließlich für alle $\varepsilon \in \mathcal{Z}$

$$\lim_{n \rightarrow \infty} \sigma_{i+1}^{\varepsilon, n} = \sigma_{i+1}^\varepsilon.$$

Konvergenz der ε -Leiterhöhen. Falls σ_{i+1}^ε keine Unstetigkeitsstelle ist, konvergieren die zeitdiskreten ε -Leiterhöhen $H_{i+1}^{\varepsilon, n}$ selbst auch,

$$\lim_{n \rightarrow \infty} H_{i+1}^{\varepsilon, n} = \lim_{n \rightarrow \infty} h(X_{\sigma_{i+1}^{\varepsilon, n}}) = H_{i+1}^\varepsilon.$$

Im Falle einer Sprungstelle σ_{i+1}^ε , d.h. $h(X_{\sigma_{i+1}^\varepsilon-}) \neq h(X_{\sigma_{i+1}^\varepsilon})$, gilt aufgrund der zeitlichen Minimalität in der Definition der ε -Leiterzeit σ_{i+1}^ε und wegen $\varepsilon \in \mathcal{Z}$

$$h(X_{\sigma_{i+1}^\varepsilon-}) \leq \sup_{t < \sigma_{i+1}^\varepsilon} h(X_t) < H_i^\varepsilon + \varepsilon \quad \text{und} \quad h(X_{\sigma_{i+1}^\varepsilon}) \neq H_i^\varepsilon + \varepsilon$$

Aus der definierenden Eigenschaft der ε -Leiterzeit folgt somit zwangsläufig $h(X_{\sigma_{i+1}^\varepsilon}) = H_{i+1}^\varepsilon > H_i^\varepsilon + \varepsilon$. Setze nun

$$\alpha := (H_i^\varepsilon + \varepsilon) - \sup_{t < \sigma_{i+1}^\varepsilon} h(X_t) > 0$$

Aufgrund der Induktionsvoraussetzung konvergiert $H_i^{\varepsilon, n} \rightarrow H_i^\varepsilon$ für $n \rightarrow \infty$ und es gilt damit für alle hinreichend große n

$$(H_i^{\varepsilon, n} + \varepsilon) - \sup_{t < \sigma_{i+1}^\varepsilon} h(X_t) > \alpha/2 > 0,$$

und dies impliziert

$$H_i^{\varepsilon,n} + \varepsilon > \sup_{t < \sigma_{i+1}^{\varepsilon}} h(X_t) \geq \sup_{2^{-n}j < \sigma_{i+1}^{\varepsilon}} h(X_{2^{-n}j}) > 0.$$

Zusätzlich zur Konvergenz der ε -Leiterzeiten folgt daraus $\sigma_{i+1}^{\varepsilon,n} \geq \sigma_{i+1}^{\varepsilon}$ für alle hinreichend große n , und wegen der rechtsseitigen Stetigkeit des Pfades ($h(X_t)$) gilt somit auch in diesem Falle $H_{i+1}^{\varepsilon,n} \rightarrow H_{i+1}^{\varepsilon}$ für $n \rightarrow \infty$.

Schritt 5: Konvergenz des diskretisierten ε -Supremums. Für die Konvergenz $L^{\varepsilon,n} \rightarrow L^{\varepsilon}$ für $n \rightarrow \infty$ mit $\varepsilon \in \mathcal{Z}$ genügt es, wenn die diskretisierten ε -Leiterpunkte $(\sigma_i^{\varepsilon,n}, H_i^{\varepsilon,n})$ konvergieren, wie an der folgenden Fallunterscheidung deutlich wird.

Fall 5.1: $L^{\varepsilon} < \infty$. Sei i der Index des letzten echten ε -Leiterpunkts, d.h.

$$\sigma_i^{\varepsilon} < \infty \quad \text{und} \quad \sigma_{i+1}^{\varepsilon} = \infty.$$

Wie im Schritt 4.1 gezeigt, gilt dann $\sigma_{i+1}^{\varepsilon,n} = \infty$ für alle hinreichend große n und damit gilt nach Definition auch $L^{\varepsilon,n} = \sigma_i^{\varepsilon,n}$ für diese n . Da außerdem $\sigma_i^{\varepsilon,n}$ gegen die i -te ε -Leiterzeit σ_i^{ε} konvergiert, folgt damit

$$\lim_{n \rightarrow \infty} L^{\varepsilon,n} = \lim_{n \rightarrow \infty} \sigma_i^{\varepsilon,n} = \sigma_i^{\varepsilon} = L^{\varepsilon} \quad \text{für alle } \varepsilon \in \mathcal{Z}.$$

Fall 5.2: $L^{\varepsilon} = \infty$. Im Falle von $L^{\varepsilon} = \infty$ sind die ε -Leiterzeiten σ_i^{ε} für alle $i \in \mathbb{N}$ endlich, und dies impliziert, dass der Pfad ($h(X_t)$) nach oben unbeschränkt ist. Damit sind dann aber auch alle zeitdiskreten ε -Leiterzeiten $\sigma_i^{\varepsilon,n}$ für jeweils hinreichend große n endlich, und somit $L^{\varepsilon,n} = \infty$ für alle n .

Schritt 6: Konvergenz der diskretisierten ε -Zerlegungszeit. Die fast sichere Konvergenz der Zerlegungszeiten $\tau^{\varepsilon,n} \rightarrow \tau^{\varepsilon}$ für $n \rightarrow \infty$ muss diesmal gesondert behandelt werden, da die Definition der Zerlegungszeit etwas von der des Supremums L^{ε} abweicht. Bemerkenswerterweise ist $\tau^{\varepsilon} < \infty$ für alle $\varepsilon \in \hat{\mathcal{Z}}$, denn nach Konstruktion gilt $\hat{H}_i^{\varepsilon} \geq i\varepsilon + h(\hat{X}_0)$ falls $\hat{\sigma}_i^{\varepsilon} < \infty$, also existiert immer ein i mit $\hat{H}_i^{\varepsilon} > Y$ oder $\hat{\sigma}_i^{\varepsilon} = \infty$, und somit muss die Zerlegungszeit τ^{ε} endlich sein.

Sei also $\varepsilon \in \hat{\mathcal{Z}}$, und sei i der letzte echte ε -Leiterpunkt kleiner oder gleich Y , d.h.

$$\hat{\sigma}_i^{\varepsilon} < \infty \text{ und } \hat{H}_i^{\varepsilon} \leq Y \text{ und } (\hat{\sigma}_{i+1}^{\varepsilon} = \infty \text{ oder } \hat{H}_{i+1}^{\varepsilon} > Y),$$

dann gilt fast sicher sogar $\hat{H}_i^{\varepsilon} < Y$. Für alle hinreichend große n ist dann auch fast sicher $\hat{H}_i^{\varepsilon,n} < Y$ und $\hat{\sigma}_i^{\varepsilon,n} < \infty$ und $\hat{\sigma}_{i+1}^{\varepsilon,n} = \infty$ oder $\hat{H}_{i+1}^{\varepsilon,n} > Y$, und damit gilt nach Definition $\tau^{\varepsilon,n} = \hat{\sigma}_i^{\varepsilon,n}$ für alle hinreichend große n . Es folgt damit aus der Konvergenz der diskreten ε -Leiterpunkte

$$\lim_{n \rightarrow \infty} \tau^{\varepsilon,n} = \lim_{n \rightarrow \infty} \hat{\sigma}_i^{\varepsilon,n} = \hat{\sigma}_i^{\varepsilon} = \tau^{\varepsilon} \quad \text{für alle } \varepsilon \in \hat{\mathcal{Z}}.$$

Schritt 7: Konvergenz der ε -Zeitpunkte. Für alle $\varepsilon \in \mathcal{Z}$ gilt nach Gleichung (3.16) $\sup_{s < t} h(X_s) < h(X_{L^{\varepsilon}})$ für alle $t < L^{\varepsilon}$, und damit ist auch $L^{\varepsilon} \leq L$, also folgt

$$\limsup_{\varepsilon \rightarrow 0, \varepsilon \in \mathcal{Z}} L^{\varepsilon} \leq L. \quad (3.19)$$

Andererseits sei $t < L$ und sei

$$0 < \alpha < \sup_{s \geq 0} h(X_s) - \sup_{s < t} h(X_s), \quad (3.20)$$

eine solche Wahl von α ist wegen der Minimalität von L stets möglich. Für alle $\varepsilon < \alpha$ mit $\varepsilon \in \mathcal{Z}$ gilt gemäß Beobachtung (3.8)

$$\sup_{s \geq 0} h(X_s) - \sup_{s \leq L^\varepsilon} h(X_s) \leq \sup_{s \geq 0} h(X_s) - h(X_{L^\varepsilon}) \leq \varepsilon < \alpha,$$

andererseits ist wegen Ungleichung (3.20) $\sup_{s \geq 0} h(X_s) - \sup_{s < t} h(X_s) > \alpha$, also muss $L^\varepsilon > t$ gelten. Wir erhalten

$$\liminf_{\varepsilon \rightarrow 0, \varepsilon \in \mathcal{Z}} L^\varepsilon \geq t \quad \text{für alle } t < L, \text{ also } \liminf_{\varepsilon \rightarrow 0, \varepsilon \in \mathcal{Z}} L^\varepsilon \geq L. \quad (3.21)$$

Aus den Gleichungen (3.19) und (3.21) folgt insgesamt die Konvergenz $L^\varepsilon \rightarrow L$ für $\varepsilon \rightarrow 0$ und $\varepsilon \in \mathcal{Z}$, und mit einer vollkommen analogen Betrachtung folgt schließlich auch die Konvergenz $\tau^\varepsilon \rightarrow \tau$ für $\varepsilon \rightarrow 0$ und $\varepsilon \in \hat{\mathcal{Z}}$.

Schritt 8: Konvergenz der Endpunkte X_{L^ε} und $\hat{X}_{\tau^\varepsilon}$. In Schritt 7 wurde gezeigt, dass die Zeitpunkte L^ε der approximierenden Suprema und die Zerlegungszeiten τ^ε konvergieren, allerdings folgt daraus nicht notwendigerweise eine Konvergenz der Form $X_{L^\varepsilon} \rightarrow X_L$ oder $\hat{X}_{\tau^\varepsilon} \rightarrow \hat{X}_\tau$. Vielmehr ist eine Fallunterscheidung anhand der Zerlegungszeit L bzw. τ notwendig. Es sind hierbei drei Fälle für τ zu unterscheiden.

Fall 1: $Y \geq h(\hat{X}_{\tau-}) \geq h(\hat{X}_\tau)$. In diesem Fall ist τ entweder eine Stetigkeitsstelle von $(h(\hat{X}_t))$ oder es ist $Y \geq h(\hat{X}_t) > h(\hat{X}_\tau)$ in einem ganzen Intervall links von τ , so dass dann nach Konstruktion τ^ε von links gegen τ konvergiert. In beiden Fällen folgt $\hat{X}_{\tau^\varepsilon} \rightarrow \hat{X}_{\tau-}$, und dies entspricht der zweiten Zeile in Gleichung 3.3.

Fall 2: $h(\hat{X}_{\tau-}) \leq Y < h(\hat{X}_\tau)$. Da in diesem Fall nach Konstruktion $\tau^\varepsilon < \tau$ für alle $\varepsilon > 0$ mit $\varepsilon \in \mathcal{Z}$ gilt, konvergiert auch hier τ^ε von links gegen τ . Somit folgt $\hat{X}_{\tau^\varepsilon} \rightarrow \hat{X}_{\tau-}$, und dies entspricht der dritten Zeile in Gleichung 3.3.

Fall 3: $h(\hat{X}_{\tau-}) < h(\hat{X}_\tau) \leq Y$. In diesem Fall ist fast sicher $Y > h(\hat{X}_t) > h(\hat{X}_{\tau-})$ in einem ganzen Intervall rechts von τ , demnach konvergiert nach Konstruktion τ^ε von rechts gegen τ , und somit $\hat{X}_{\tau^\varepsilon} \rightarrow \hat{X}_\tau$. Dies entspricht der vierten Zeile in Gleichung 3.3.

Es gilt also insgesamt $\hat{X}_{\tau^\varepsilon} \rightarrow \hat{X}_{\perp\tau}$, und eine vollkommen analoge Betrachtung führt zu $X_{L^\varepsilon} \rightarrow X_{\perp L}$ für den Prozess (X_t) . Es folgt damit schließlich auch auf dem Raum der Pfade in der Skohorod-Metrik

$$((X_{t \wedge L^\varepsilon}), L^\varepsilon) \rightarrow ((X_{t \perp L}), L) \quad \text{und} \quad ((\hat{X}_{t \wedge \tau^\varepsilon}), \tau^\varepsilon) \rightarrow ((\hat{X}_{t \perp \tau}), \tau) \quad \text{für } \varepsilon \rightarrow 0, \varepsilon \in \mathcal{Z}.$$

Schritt 9: Zusammensetzen der Beweisschritte.

$$X_t^n := X_{2^{-n} \lceil 2^n t \rceil} \quad \text{und} \quad \hat{X}_t^n := \hat{X}_{2^{-n} \lceil 2^n t \rceil}$$

Damit gilt dann in der Skohorod-Metrik d_S auf dem Raum der Pfade

$$\lim_{n \rightarrow \infty} (X_t^n) = (X_t) \text{ f.s.} \quad \text{und} \quad \lim_{n \rightarrow \infty} (\hat{X}_t^n) = (\hat{X}_t) \text{ f.s.} \quad (3.22)$$

Ersetzen wir in den ε -Approximationen den festen Wert ε jeweils durch eine Zufallsvariable E_ε , definiert durch

$$E_\varepsilon := \varepsilon + \varepsilon U_\varepsilon \text{ mit } U_\varepsilon \text{ unabhängig und uniform verteilt auf } [0, 1],$$

dann gilt jeweils $E_\varepsilon \in \mathcal{Z}$ fast sicher, da das Komplement von \mathcal{Z} abzählbar ist, und außerdem konvergiert $E_\varepsilon \rightarrow 0$ fast sicher für $\varepsilon \rightarrow 0$. Daraus folgt nun

$$\begin{array}{ccccc} ((X_{t \wedge L}^n, L^{E_\varepsilon, n}), L^{E_\varepsilon, n}) & \xrightarrow[\textcircled{5}]{n \rightarrow \infty} & ((X_{t \wedge L}^{E_\varepsilon}, L^{E_\varepsilon}), L^{E_\varepsilon}) & \xrightarrow[\textcircled{8}]{\varepsilon \rightarrow 0} & ((X_{t \perp L}, L), L) \\ \parallel \textcircled{2} & & \parallel \textcircled{1} & & \parallel \textcircled{2} \\ ((\hat{X}_{t \wedge \tau}^n, \tau^{E_\varepsilon, n}), \tau^{E_\varepsilon, n}) & \xrightarrow[\textcircled{6}]{n \rightarrow \infty} & ((\hat{X}_{t \wedge \tau}^{E_\varepsilon}, \tau^{E_\varepsilon}), \tau^{E_\varepsilon}) & \xrightarrow[\textcircled{8}]{\varepsilon \rightarrow 0} & ((\hat{X}_{t \perp \tau}, \tau), \tau) \end{array}$$

Die Gleichverteilung $\textcircled{2}$ der zeitdiskreten Prozesse folgt aus Beweisschritt 2, die zusätzliche Randomisierung von ε mittels E_ε ändert daran natürlich nichts, da die auf ε bedingten Prozesse identisch verteilt sind. Die fast sichere Konvergenzen $\textcircled{5}$ und $\textcircled{6}$ jeweils in der Skohorod-Metrik folgen aus den Beweisschritten 5 und 6 zusammen mit den Gleichungen (3.22), und der Tatsache, dass $E_\varepsilon \in \mathcal{Z}$ fast sicher für alle $\varepsilon > 0$ ist. Damit folgt dann auch die Verteilungsgleichheit $\textcircled{1}$ aus Gleichung $\textcircled{2}$. Die beiden fast sicheren Konvergenzen $\textcircled{8}$ für $\varepsilon \rightarrow 0$ folgen aus Beweisschritt 8, unter Berücksichtigung der Tatsache, dass mit $\varepsilon \rightarrow 0$ auch $E_\varepsilon \rightarrow 0$ fast sicher gilt. Hiermit folgt schließlich die Verteilungsgleichheit $\textcircled{2}$ aus der Gleichheit $\textcircled{1}$. ■

3.4 Erkennung der Zerlegungszeit

Wie im zeitdiskreten Fall stellt sich wieder die Frage, unter welchen Bedingungen eine Zerlegung fast sicher stattfindet (d.h. wann $\tau < \infty$ fast sicher gilt), und wie der Zerlegungszeitpunkt in fast sicher endlicher Zeit bestimmt werden kann. Dies ist offenbar immer dann möglich, wenn eine passende Überkreuzungszeit fast sicher endlich ist. Diese Konzepte sind identisch mit denen der diskreten Zerlegung. Setze also

$$\tau_c := \inf\{t \geq 0 : h(\hat{X}_t) > Y\}.$$

Mit dieser Definition gilt sicher $\tau \leq \tau_c \leq \infty$. Interessant ist offenbar der Fall, in dem $\tau_c < \infty$ fast sicher gilt. Die Bedingungen hierfür sind in natürlicher Weise die gleichen wie in der zeitdiskreten Zerlegung, d.h. es ist $\tau_c < \infty$ fast sicher genau dann, wenn eine der beiden folgenden äquivalenten Bedingungen erfüllt ist:

$$\sup_t h(X_t) = \infty \mathbf{P}^h - \text{f.s.} \quad \text{oder} \quad \inf_t h(X_t) = 0 \mathbf{P} - \text{f.s.} \quad (3.23)$$

Diese Bedingungen sind auch hinreichend für die duale Zerlegung, so dass dann die entsprechende Zeit τ_c^* fast sicher endlich ist mit

$$\tau_c^* := \inf\{t \geq 0 : h(\check{X}_t^*) < U\}.$$

3.5 Der post- τ -Prozess im stetigen Fall

Im von Pitman zitierten Satz 20 stellte dieser 1974 in seinem Preprint [Pit74] die Vermutung auf, dass die pre- L - und der post- L -Prozesse bedingt unabhängig sind, und dass gegeben den Wert $m := M_\infty$ die Verteilung des post- L -Prozesses als eine $(h - m)$ -Transformation auf dem Gebiet $\{x : h(x) \geq m\}$ dargestellt werden kann. Dies trifft dann natürlich auch auf den post- τ -Prozess zu, da dieser nach Konstruktion die gleiche Verteilung besitzt. Die aufgestellte Vermutung lässt sich für stetige Prozesse (bis auf eine Eintrittsverteilung) mit einer (in vielen Beispielen erfüllten) Zusatzbedingung einfach verifizieren, wie in diesem Abschnitt gezeigt wird.

Die Verteilung von (\check{X}_t)

Im zeitdiskreten Fall sind wir bereits in Abschnitt 1.5 auf eine Darstellung der bedingten Kerne (Q_s) gestoßen, die im wesentlichen durch eine harmonische Transformation aus dem ursprünglichen Kern P hervorgingen. Die dort verwendeten Ideen lassen sich fast direkt auf den zeitkontinuierlichen Fall übertragen. Betrachten wir für $m \geq 0$ die Funktion

$$q_m(x) := \mathbf{P}_x(h(X_s) \leq m \forall s \geq 0),$$

dann gilt analog zum zeitdiskreten Lemma 5 mit einer kleinen Modifikation der Aussage das folgende Lemma:

Lemma 34 Für alle $x \in S$ mit $h(x) < m$ gilt $q_m(x) > 0$.

Beweis Setz $\sigma_m := \inf\{t \geq 0 : h(X_t) > m\}$. Angenommen es wäre $q_m(x) = 0$, d.h. $\sigma_m < \infty$ fast sicher. Dann folgt aus dem Lemma von Fatou und der Martingal-Eigenschaft von $(h(X_t))$

$$\mathbf{E}_x h(X_{\sigma_m}) = \mathbf{E}_x \lim_{n \rightarrow \infty} h(X_{\sigma_m \wedge n}) \leq \liminf_{n \rightarrow \infty} \mathbf{E}_x h(X_{\sigma_m \wedge n}) = h(x) < m.$$

Da allerdings nach Definition der Stoppzeit σ_s fast sicher $h(X_{\sigma_m}) \geq s$ gilt, stellt dies ein Widerspruch dar. Also ist $q_m(x) = \mathbf{P}_x(\sigma_m = \infty) > 0$ und das Lemma ist bewiesen. ■

Dieses Lemma gestattet es nun wieder die Kerne des auf $\{\sigma_m = \infty\}$ bedingten Prozesses mit einfachen Mitteln zu bestimmen, da bei geeigneter Wahl des Startpunktes die Bedingung kein Nullereignis ist. Für $x, y \in S$ mit $h(x), h(y) < m$ gilt

$$K_t^m f(x) = \frac{1}{q_m(x)} \mathbf{E}_x[f(X_t) q_m(X_t) : h(X_s) \leq m \forall 0 \leq s \leq t]. \quad (3.24)$$

Für einen Startpunkt x mit $h(x) = m$ ist in vielen Fällen allerdings $q_m(x) = 0$, schon bei den meisten stetigen Prozessen tritt dieses Problem auf. Hierfür werden zusätzlich die Eintrittsmaße (D_t^m) bzw. (Q_t^m) benötigt wie in den Gleichungen (3.1) angegeben. Unter Ausnutzung der Martingaleigenschaft von h erhalten wir vollkommen analog zum Abschnitt 1.5 die folgende Gleichung für $x \in S$ mit $h(x) < m$:

$$h(x) = \mathbf{E}_x[h(X_{\sigma_m}); \sigma_m < \infty] + \mathbf{E}_x[H; \sigma_m = \infty] \quad (3.25)$$

mit $H := \lim h(X_t)$ fast sicher (da $(h(X_t))$ ein nicht-negatives Martingal ist, existiert dieser Grenzwert \mathbf{P}_x -fast sicher). Falls zusätzlich die reguläre Funktion h die Eigenschaft (3.23)

besitzt, gilt für den Grenzwert $H = 0$ \mathbf{P}_x -fast sicher, womit der zweite Summand in Gleichung (3.25) verschwindet, und durch Übergang zur bedingten Erwartung erhalten wir

$$h(x) = \mathbf{P}_x(\sigma_m < \infty) \mathbf{E}_x[h(X_{\sigma_m}) | \sigma_m < \infty].$$

Hiermit ergibt sich die folgende Darstellung von $q_m(x)$

$$q_m(x) = 1 - \mathbf{P}_x(\sigma_m < \infty) = 1 - \frac{h(x)}{\mathbf{E}_x[h(X_{\sigma_m}) | \sigma_m < \infty]}.$$

Zusammen mit Gleichung (3.24) und der Definition $P_t^m f(x) := \mathbf{E}_x[f(X_t) : h(X_s) \leq m \forall 0 \leq s \leq t]$ erhalten wir (weiter unter der Voraussetzung $h(X_t) \rightarrow 0$ \mathbf{P}_σ -fast sicher)

$$\begin{aligned} K_t^m(x, dy) &= \frac{q_m(y)}{q_m(x)} P_t^m(x, dy) \\ &= \frac{\mathbf{E}_x[h(X_{\sigma_m}) | \sigma_m < \infty] (\mathbf{E}_y[h(X_{\sigma_m}) | \sigma_m < \infty] - h(y))}{\mathbf{E}_y[h(X_{\sigma_m}) | \sigma_m < \infty] (\mathbf{E}_x[h(X_{\sigma_m}) | \sigma_m < \infty] - h(x))} P_t^m(x, dy) \end{aligned}$$

für alle $x, y \in \{z \in S : h(z) < m\}$. Falls der Prozess zudem fast sicher stetige Pfade besitzt, dann gilt $\mathbf{E}_x[h(X_{\sigma_m}) | \sigma_m < \infty] = m$ für alle x mit $h(x) < m$, und die Gleichung vereinfacht sich zu

$$K_t^m(x, dy) = \frac{m - h(y)}{m - h(x)} P_t^m(x, dy).$$

Die Verteilung von (\hat{X}_t^*)

Die gleiche Frage nach der Verteilung stellt sich auch für die duale Zerlegung. Unter der Bedingung $h(X_t) \rightarrow \infty$ \mathbf{P}_σ^h -fast sicher ergeben sich auch hier die Kerne der Verteilung des post- τ^* -Prozesses durch ähnlich Überlegungen als

$$\begin{aligned} K_t^{*m}(x, dy) &= \frac{q_m^*(y)}{q_m^*(x)} P_t^h(x, dy) \\ &= \frac{\mathbf{E}_x^h[1/h(X_{\sigma_m^*}) | \sigma_m^* < \infty] (\mathbf{E}_y^h[1/h(X_{\sigma_m^*}) | \sigma_m^* < \infty] h(y) - 1)}{\mathbf{E}_y^h[1/h(X_{\sigma_m^*}) | \sigma_m^* < \infty] (\mathbf{E}_x^h[1/h(X_{\sigma_m^*}) | \sigma_m^* < \infty] h(x) - 1)} P_t^m(x, dy) \end{aligned}$$

für alle $x, y \in \{z \in S : h(z) > m\}$, wobei die Stoppzeit σ_m^* analog definiert ist durch $\sigma_m^* := \inf\{t \geq 0 : h(X_t) < m\}$. Falls der zugrunde liegende Prozess zudem fast sicher stetige Pfade besitzt, vereinfacht sich die Darstellung wieder zu

$$K_t^{*m}(x, dy) = \frac{h(y) - m}{h(x) - m} P_t^m(x, dy).$$

Dies entspricht gerade der von Pitman in Hauptsatz 20 aufgestellten Vermutung, dass sich die Verteilung des post- L^* -Prozesses als eine $(h - m)$ -Transformation eines auf $\{x : h(x) \geq m\}$ eingeschränkten Prozesses darstellen lässt. Vorausgesetzt wurden allerdings stetige Pfade und $h(X_t) \rightarrow \infty$ \mathbf{P}_σ^h -fast sicher – eine stärkere Bedingung im Vergleich zu der von Pitman geforderten Unbeschränktheit von h . Zusätzlich wird selbst auch im stetigen Fall noch ein Eintrittsmaß benötigt, welches entweder durch (Q_t^{*m}) oder (D_t^{*m}) beschrieben wird.

Bemerkung Einige Beispiele für Zerlegungen behandeln Diffusionen - in diesem Fall ist es eher unüblich mit Kernen zu arbeiten, andere Darstellungen der Verteilungen (stochastische Differentialgleichungen oder Generatoren) werden oft bevorzugt. Mit den Methoden von Jacobsen [Jac74] kann eine konservative, reguläre Diffusion in Dimension 1, die bedingt ist, eine Schranke m nicht mehr zu unterschreiten, wiederum selbst als Diffusion konstruiert werden. Für eine unbedingte Diffusion mit Skalierung $S : \mathbb{R} \rightarrow \mathbb{R}$ und Geschwindigkeitsmaß μ auf \mathbb{R} ist der bedingte Prozess unter geeigneten Annahmen wieder eine Diffusion mit Skalierung S_m und Geschwindigkeit μ_m gegeben durch

$$S_m := -\frac{1}{S - S(m)} \quad \text{und} \quad \mu_m(dx) = (S(x) - S(m))^2 \mu(dx).$$

□

Kapitel 4

Anwendungen zeitkontinuierlicher Zerlegungen

4.1 Beispiel: Lévy-Prozesse

Eines der prominentesten Beispiele für eine Zerlegung gemäß dem Hauptsatz 31 sind Lévy-Prozesse. Ein Lévy-Prozess (X_t) ist definiert als zeithomogener Markov-Prozess in \mathbb{R}^d mit rechtsseitig stetigen Pfaden mit linksseitigen Limiten auf einem Wahrscheinlichkeitsraum mit Maß \mathbf{P} , so dass die Zuwächse von (X_t) eine stationäre Verteilung besitzen, d.h. es gilt

$$\mathcal{L}(X_{t+s} - X_t) = \mathcal{L}(X_s - X_0) \quad \text{für alle } s, t > 0,$$

und zudem sind für alle $t_0 < t_1 < \dots < t_n$ die Zuwächse $X_{t_1} - X_{t_0}, \dots, X_{t_n} - X_{t_{n-1}}$ unabhängig. Durch diese beiden Eigenschaften ergibt sich in natürlicher Weise eine direkte Beziehung zu den unbegrenzt teilbaren Verteilungen (siehe z.B. [Bre86]). Bekanntlich (siehe [Ber96, Bre86, RW94]) ist die charakteristische Funktion $\psi(\lambda) := \mathbf{E}_0 \exp(i\langle \lambda, X_1 \rangle)$ eines Lévy-Prozesses gegeben durch die Lévy-Khinchin-Formel

$$\log \psi(\lambda) = i\langle \lambda, b \rangle - \frac{1}{2} \lambda^T \Sigma \lambda + \int_{\mathbb{R}^d} (e^{i\langle \lambda, x \rangle} - 1 - i\langle \lambda, x \rangle 1_{\{\|x\| < 1\}}) \mu(dx) \quad , \lambda \in \mathbb{R}^d, \quad (4.1)$$

wobei $b \in \mathbb{R}^d$ der lineare Drift, $\Sigma \in \mathbb{R}^{d \times d}$ die Kovarianzmatrix der eingebetteten Brownschen Bewegung und μ das Lévy-Maß ist mit der Bedingung

$$\int_{\mathbb{R}^d} (\|x\|^2 \wedge 1) \mu(dx) < \infty. \quad (4.2)$$

Die charakteristische Funktion $\psi(\lambda)$ ist endlich für alle $\lambda \in \mathbb{R}^d$.

Interpretation der charakteristischen Funktion

Weshalb die charakteristische Funktion für Lévy-Prozesse ein derart wichtiges Hilfsmittel ist, offenbart sich in der wohlbekannten sehr direkten Interpretation der einzelnen Komponenten des charakteristischen Exponenten $\log \psi$. Jeder Lévy-Prozess (X_t) ist in drei grundlegende unabhängige Bestandteile zerlegbar

$$X_t = tb + W_t + S_t$$

wobei b der lineare Drift des Prozesses ist, (W_t) eine d -dimensionale Brownsche Bewegung mit Kovarianzmatrix Σ und (S_t) eine Sprungprozess ist, der als Grenzwert einer Folge kompensierter Compound Poisson-Prozesse zu verstehen ist. Das Lévy-Maß μ steuert hierbei die Verteilung und Intensität der Sprünge. Für Details siehe [Ber96, Bre86].

Falls das Lévy-Maß μ endlich ist, d.h. $\mu(\mathbb{R}^d) < \infty$, ist eine Sprungkompensation nicht erforderlich, und somit vereinfacht sich ψ zu (allerdings mit einem anderen Driftterm $b \in \mathbb{R}^d$)

$$\log \psi(\lambda) = i\langle \lambda, b \rangle - \frac{1}{2} \lambda^T \Sigma \lambda + \int_{\mathbb{R}^d} (e^{i\langle \lambda, x \rangle} - 1) \mu(dx). \quad (4.3)$$

Falls der Prozess schließlich gar keine Sprünge besitzt, also das Lévy-Maß Gesamtmasse 0 besitzt, vereinfacht sich die charakteristische Funktion weiter zu der einer d -dimensionalen Brownschen Bewegung mit Drift:

$$\log \psi(\lambda) = i\langle \lambda, b \rangle - \frac{1}{2} \lambda^T \Sigma \lambda.$$

Reguläre Funktionen

Das Vorgehen im Falle der Lévy-Prozesse läuft nun vollkommen analog zu der Zerlegung von Irrfahrten im Abschnitt 2.1. Betrachten wir uns hierzu wieder die Laplace Transformierte $\varphi(\lambda) := \mathbf{E}_0 \exp(\langle \lambda, X_1 \rangle)$. Mittels der Zusammenhänge

$$\varphi(\lambda) = \psi(-i\lambda) \quad \text{und} \quad \psi(\lambda) = \varphi(i\lambda) \quad \text{mit } \lambda \in \mathbb{R}^d \quad (4.4)$$

können wir zwischen der charakteristischen Funktion und der Laplace Transformierten jederzeit wechseln – allerdings ist die Laplace Transformierte φ nicht mehr notwendigerweise für alle $\lambda \in \mathbb{R}^d$ endlich. Falls jedoch $|\varphi(\lambda)| < \infty$ ist, dann besitzt die Laplace Transformierte vermöge (4.4) die folgende Gestalt

$$\log \varphi(\lambda) = \langle \lambda, b \rangle + \frac{1}{2} \lambda^T \Sigma \lambda + \int_{\mathbb{R}^d} e^{\langle \lambda, x \rangle} - 1 - \langle \lambda, x \rangle 1_{\{\|x\| < 1\}} \mu(dx).$$

Falls das erste Moment von $X_1 - X_0$ existiert, dann kann der Erwartungswert der Zuwächse auch wieder mit Hilfe der Ableitung der Laplace Transformierten bestimmt werden. Unter Verwendung der Kettenregel und $\varphi(0) = 1$ erhalten wir

$$\begin{aligned} \mathbf{E}(X_1 - X_0) &= \text{grad} \varphi(0) = \text{grad} \exp \log \varphi(0) = \text{grad} \log \varphi(0) \cdot \varphi(0) \\ &= b + \int_{\mathbb{R}^d} x - x 1_{\{\|x\| < 1\}} \mu(dx) \\ &= b + \int_{\|x\| \geq 1} x \mu(dx). \end{aligned} \quad (4.5)$$

Bezeichne C die Menge der Nullstellen des Exponenten der Laplace Transformierten, also

$$C := \{u \in \mathbb{R}^d : \log \varphi(u) = 0\} = \{u \in \mathbb{R}^d : \log \psi(-iu) = 0\},$$

Die Menge C kann unter Umständen nur aus der triviale Lösung $u = 0$ bestehen. Aufgrund der mannigfaltigen Natur von Lévy-Prozesses ist es schwierig, geeignete allgemeine hinreichende

und notwendige Bedingungen für eine nichttriviale Lösungsmenge anzugeben. Aus der Menge C lassen sich nun leicht reguläre Funktionen gewinnen; für $u \in C$ setze

$$h_u(\lambda) = e^{\langle u, x \rangle}, \quad u \in C.$$

Es gilt dann wegen der Translationsinvarianz der Zuwächse und wegen $\varphi(u) = 1$

$$P_1 h_u(x) = \mathbf{E}_x e^{\langle u, X_1 \rangle} = \mathbf{E}_0 e^{\langle u, x + X_1 \rangle} = e^{\langle u, x \rangle} \varphi(u) = h_u(x),$$

also ist h_u regulär bezüglich der Kernfamilie (P_t) .

Maßwechsel

Wir können nun eine h_u -Transformation der Verteilung durchführen, die mit h_u assoziierte Verteilung bezeichnen wir mit \mathbf{P}^{h_u} , und entsprechend die Erwartung mit \mathbf{E}^{h_u} . Um die transformierte Verteilung genau zu charakterisieren, ist es ausreichend, mit der charakteristischen Funktion bzw. mit der Laplace Transformierten anstelle direkt mit den Kernen zu arbeiten, da diese die Verteilung vollständig und eindeutig bestimmen. Bezeichne also φ^{h_u} die von der Verteilung implizierte Laplace Transformierte, dann ergibt sich die folgende einfache Beziehung zu φ :

$$\varphi^{h_u}(\lambda) := \mathbf{E}_0^{h_u} e^{\langle \lambda, X_1 \rangle} = \mathbf{E}_0 e^{\langle u, X_1 \rangle} e^{-\langle u, 0 \rangle} e^{\langle \lambda, X_1 \rangle} = \mathbf{E}_0 e^{\langle \lambda + u, X_1 \rangle} = \varphi(\lambda + u). \quad (4.6)$$

Mit Hilfe der Laplace Transformierten kann der erwartete Zuwachs auch einfach bestimmt werden, denn es gilt

$$\mathbf{E}_0^{h_u} X_1 = \mathbf{E}_0^{h_u} X_1 e^{\langle 0, X_1 \rangle} = \text{grad} \varphi^{h_u}(0) = \text{grad} \varphi(u) \quad (4.7)$$

Es kann sogar die exakte Verteilung des h_u -transformierten Prozesses hinsichtlich Drift, Brownscher Anteil und Sprüngen bestimmt werden. Da nach Konstruktion $u \in C$ und damit $\varphi(u) = 1$ ist, gilt

$$0 = \log \varphi(u) = \langle u, b \rangle + \frac{1}{2} u^T \Sigma u + \int_{\mathbb{R}^d} e^{\langle u, x \rangle} - 1 - \langle u, x \rangle 1_{\{\|x\| < 1\}} \mu(dx). \quad (4.8)$$

Aus den Gleichungen (4.6) und (4.8) erhalten wir

$$\begin{aligned} \log \varphi^{h_u}(\lambda) &= \langle \lambda + u, b \rangle + \frac{1}{2} (\lambda + u)^T \Sigma (\lambda + u) + \int_{\mathbb{R}^d} e^{\langle \lambda + u, x \rangle} - 1 - \langle \lambda + u, x \rangle 1_{\{\|x\| < 1\}} \mu(dx) \\ &= \langle \lambda, b \rangle + \frac{1}{2} \lambda^T \Sigma \lambda + u^T \Sigma \lambda + \int_{\mathbb{R}^d} e^{\langle \lambda + u, x \rangle} - e^{\langle u, x \rangle} - \langle \lambda, x \rangle 1_{\{\|x\| < 1\}} \mu(dx), \end{aligned}$$

durch einen Maßwechsel von μ nach $\mu^{h_u}(dx) := e^{\langle u, x \rangle} \mu(dx)$ erhalten wir

$$\begin{aligned} &= \langle \lambda, b \rangle + \frac{1}{2} \lambda^T \Sigma \lambda + u^T \Sigma \lambda + \int_{\mathbb{R}^d} e^{\langle \lambda, x \rangle} - 1 - \langle \lambda, x \rangle 1_{\{\|x\| < 1\}} \mu^{h_u}(dx) \\ &\quad + \int_{\|x\| < 1} (e^{\langle u, x \rangle} - 1) \langle \lambda, x \rangle \mu(dx), \end{aligned}$$

und dies lässt sich nun wie folgt in die Lévy-Khinchin-Darstellung pressen

$$\begin{aligned} &= \langle \lambda, b + u^T \Sigma + \int_{\|x\| < 1} (e^{\langle u, x \rangle} - 1) x \mu(dx) \rangle \\ &\quad + \frac{1}{2} \lambda^T \Sigma \lambda + \int_{\mathbb{R}^d} e^{\langle \lambda, x \rangle} - 1 - \langle \lambda, x \rangle 1_{\{\|x\| < 1\}} \mu^{h_u}(dx) \\ &= \langle \lambda, b^{h_u} \rangle + \frac{1}{2} \lambda^T \Sigma \lambda + \int_{\mathbb{R}^d} e^{\langle \lambda, x \rangle} - 1 - \langle \lambda, x \rangle 1_{\{\|x\| < 1\}} \mu^{h_u}(dx) \end{aligned}$$

mit dem Driftterm b^{h_u} und dem Lévy-Maß μ^{h_u} definiert durch

$$b^{h_u} := b + u^T \Sigma + \int_{\|x\| < 1} (e^{\langle u, x \rangle} - 1) x \mu(dx) \quad \text{und} \quad \mu^{h_u}(dx) := e^{\langle u, x \rangle} \mu(dx).$$

Damit scheint der h_u -transformierte Prozess wieder ein Lévy-Prozess, dessen Verteilung vollständig charakterisiert wurde. Es fehlt allerdings noch der Nachweis, dass b^{h_u} endlich ist, und dass μ^{h_u} tatsächlich ein zulässiges Lévy-Maß ist, d.h. dass Bedingung (4.2) erfüllt ist,

$$\left\| \int_{\|x\| < 1} (e^{\langle u, x \rangle} - 1) x \mu(dx) \right\| < \infty \quad \text{und} \quad \int_{\mathbb{R}^d} (\|x\|^2 \wedge 1) \mu^{h_u}(dx) < \infty. \quad (4.9)$$

Betrachten wir uns die erste der beiden Bedingungen (4.9):

$$\begin{aligned} &\left\| \int_{\|x\| < 1} (e^{\langle u, x \rangle} - 1) x \mu(dx) \right\| \leq \int_{\|x\| < 1} \|x\| \cdot |e^{\langle u, x \rangle} - 1| \mu(dx) \\ &= \int_{\substack{\|x\| < 1 \\ \|u\| \cdot \|x\| \leq 1}} \|x\| \cdot |e^{\langle u, x \rangle} - 1| \mu(dx) + \int_{\substack{\|x\| < 1 \\ \|u\| \cdot \|x\| > 1}} \|x\| \cdot |e^{\langle u, x \rangle} - 1| \mu(dx) \end{aligned}$$

Mit den Abschätzungen $e^z - 1 \leq z + z^2$ für $-1 \leq z \leq 1$ und $\langle u, x \rangle \leq \|u\| \|x\|$ folgt im ersten Summanden

$$\leq \max(\|u\|, \|u\|^2) \int_{\substack{\|x\| < 1 \\ \|u\| \cdot \|x\| \leq 1}} \|x\|^2 + \|x\|^3 \mu(dx) + \int_{\substack{\|x\| < 1 \\ \|u\| \cdot \|x\| > 1}} \|x\| \cdot |e^{\langle u, x \rangle} - 1| \mu(dx).$$

Da μ ein zulässiges Lévy-Maß ist, ist der erste Summand wegen Bedingung (4.2) endlich. Der zweiten Summanden kann einfach abgeschätzt werden durch

$$\int_{\substack{\|x\| < 1 \\ \|u\| \cdot \|x\| > 1}} \|x\| \cdot |e^{\langle u, x \rangle} - 1| \mu(dx) \leq c \int_{\|u\| \cdot \|x\| > 1} \mu(dx) \quad \text{mit} \quad c := \sup_{\substack{\|x\| < 1 \\ \|u\| \cdot \|x\| > 1}} \|x\| \cdot |e^{\langle u, x \rangle} - 1| < \infty.$$

Eine einfach Überlegung zeigt, dass das Integral $\int_{\|u\| \cdot \|x\| > 1} \mu(dx)$ ebenfalls wegen Bedingung (4.2) endlich ist. Insgesamt ist hiermit die erste Ungleichung in 4.9 gezeigt.

Bleibt noch die zweite Bedingung in (4.9) zu überprüfen. Aus der Definition des Maßes μ^{h_u} folgt

$$\int_{\mathbb{R}^d} (\|x\|^2 \wedge 1) \mu^{h_u}(dx) = \int_{\|x\| < 1} \|x\|^2 e^{\langle u, x \rangle} \mu(dx) + \int_{\|x\| \geq 1} e^{\langle u, x \rangle} \mu(dx),$$

und aus der Abschätzung $e^{\langle u, x \rangle} \leq e^{\|u\|}$ folgt weiter

$$\leq e^{\|u\|} \int_{\|x\| < 1} \|x\|^2 \mu(dx) + \int_{\|x\| \geq 1} e^{\langle u, x \rangle} \mu(dx).$$

Da μ ein zulässiges Lévy-Maß ist, ist der erste Summand wegen Ungleichung (4.2) endlich. Es bleibt nun nur noch zu zeigen, dass auch der zweite Summand endlich ist. Nach Voraussetzung ist $\log \varphi(u) = 0$, speziell folgt aus Gleichung (4.8) sofort

$$\left| \int_{\|x\| \geq 1} (e^{\langle u, x \rangle} - 1) \mu(dx) \right| < \infty,$$

und da aus der Lévy-Bedingung (4.2) für das Maß μ sofort folgt $\int_{\|x\| \geq 1} \mu(dx) < \infty$, gilt somit insgesamt für den zweiten Summanden

$$\left| \int_{\|x\| \geq 1} e^{\langle u, x \rangle} \mu(dx) \right| \leq \left| \int_{\|x\| \geq 1} (e^{\langle u, x \rangle} - 1) \mu(dx) \right| + \left| \int_{\|x\| \geq 1} \mu(dx) \right| < \infty.$$

Damit erfüllt nun auch das Maß μ^{h_u} die Bedingung (4.2) und ist somit ein zulässiges Lévy-Maß.

Mit Hilfe der Identitäten (4.4) fassen wir das Resultat wie folgt zusammen:

Satz 35 Sei \mathbf{P} die Verteilung eines Lévy-Prozesses in \mathbb{R}^d mit charakteristischer Funktion ψ , dessen Exponent die Form hat

$$\log \psi(\lambda) = i\langle \lambda, b \rangle - \frac{1}{2} \lambda^T \Sigma \lambda + \int_{\mathbb{R}^d} (e^{i\langle \lambda, x \rangle} - 1 - i\langle \lambda, x \rangle 1_{\{\|x\| < 1\}}) \mu(dx) \quad , \lambda \in \mathbb{R}^d,$$

dabei ist $b \in \mathbb{R}^d$ der Driftterm, $\Sigma \in \mathbb{R}^{d \times d}$ die Kovarianzmatrix der eingebetteten Diffusion und μ das Lévy-Maß für die Sprünge. Bezeichne $\varphi(\lambda) := \psi(-i\lambda)$ die Laplace Transformierte, und sei $u \in C := \{\lambda \in \mathbb{R}^d : \varphi(\lambda) = 1\}$. Dann ist die Funktion h_u definiert durch

$$h_u(x) := e^{\langle u, x \rangle} \quad x \in \mathbb{R}^d$$

regulär und die h_u -transformierte Verteilung \mathbf{P}^{h_u} ist wieder ein Lévy-Prozess mit der charakteristischen Funktion ψ^{h_u} der Form

$$\log \psi^{h_u}(\lambda) = i\langle \lambda, b^{h_u} \rangle - \frac{1}{2} \lambda^T \Sigma \lambda + \int_{\mathbb{R}^d} (e^{i\langle \lambda, x \rangle} - 1 - i\langle \lambda, x \rangle 1_{\{\|x\| < 1\}}) \mu^{h_u}(dx)$$

mit dem Driftterm b^{h_u} und dem Lévy-Maß μ^{h_u} definiert durch

$$b^{h_u} := b + u^T \Sigma + \int_{\|x\| < 1} (e^{\langle u, x \rangle} - 1) x \mu(dx) \quad \text{und} \quad \mu^{h_u}(dx) := e^{\langle u, x \rangle} \mu(dx).$$

Bemerkung Falls das Lévy-Maß μ endlich ist, und somit die charakteristische Funktion die einfachere Gestalt (4.3) besitzt, vereinfacht sich auch die charakteristische Funktion ψ^{h_u} der h_u -transformierten Verteilung zu

$$\log \psi^{h_u}(\lambda) = i\langle \lambda, b^{h_u} \rangle - \frac{1}{2} \lambda^T \Sigma \lambda + \int_{\mathbb{R}^d} e^{i\langle \lambda, x \rangle} - 1 \mu^{h_u}(dx),$$

mit

$$b^{h_u} := b + u^T \Sigma \quad \text{und} \quad \mu^{h_u}(dx) := e^{\langle u, x \rangle} \mu(dx). \quad (4.10)$$

□

Erkennbarkeit der Zerlegungszeit

Es stellt sich noch die Frage nach der Erkennbarkeit der Zerlegungszeit. Um diese Eigenschaft nachzuweisen betrachten wir die reellwertige Zufallsvariable $\langle u, X_1 \rangle$ unter dem Maß \mathbf{P}^{h_u} und deren Laplace Transformierte γ die sich mit Hilfe von Gleichung (4.6) wie folgt ergibt:

$$\gamma(\lambda) := \mathbf{E}^{h_u} e^{\lambda \langle u, X_1 \rangle} = \varphi^{h_u}(\lambda u) = \varphi(\lambda u + u). \quad (4.11)$$

Wegen (4.11) ist $\gamma(\lambda) = 1$ für genau die beiden Werte $\lambda = 0$ und $\lambda = -1$, und da γ konvex ist folgt daraus $\gamma'(0) > 0$. Damit gilt

$$\mathbf{E}^{h_u} \langle u, X_1 \rangle = \gamma'(0) > 0,$$

und es folgt aus dem starken Gesetz der großen Zahlen angewandt auf die Summe X_t unabhängiger und identisch verteilter Zuwächse

$$\lim_{t \rightarrow \infty} h_u(X_t) = \lim_{t \rightarrow \infty} \exp \langle u, X_t \rangle = \infty \quad \mathbf{P}^{h_u} \text{ - f.s.}$$

Damit ist wie in Abschnitt 3.4 beschrieben die Zerlegungszeit τ fast sicher endlich und zudem durch die Überkreuzungszeit τ_c erkennbar.

Zerlegung

Für eine Zerlegung gemäß Hauptsatz 31 haben wir damit die Verteilung des Anfangsstücks exakt beschrieben. Die Zerlegungszeit τ kann analog der zeitdiskreten Irrfahrt ebenfalls einfacher beschrieben werden. Sei Y die zufällige Schranke auf einem Wahrscheinlichkeitsraum mit Maß \mathbb{P} wie im Hauptsatz 31 beschrieben mit Verteilung $\mathbb{P}(Y > y) = 1/y$ für $y > 1$, und sei h_u eine reguläre Funktion mit $u \in C$. Setze

$$Z := \|u\|^{-1} \log Y,$$

dann ist $\mathbb{P}(Z > z) = \mathbb{P}(Y > e^{\|u\|z}) = e^{-\|u\|z}$, d.h. Z ist exponentiell verteilt zum Parameter $\|u\|$ und mit Erwartungswert $\|u\|^{-1}$. Außerdem gilt

$$h_u(x) > Y \quad \Leftrightarrow \quad \frac{\langle u, x \rangle}{\|u\|} > Z \quad \Leftrightarrow \quad x \in H_u(Z),$$

wobei $H_u(Z) := \{x : \langle u, x \rangle > \|u\|Z\}$ der Halbraum in \mathbb{R}^d ist, der gerade Abstand Z vom Ursprung hat. Damit erhalten wir für einen in 0 startenden Lévy-Prozesses in \mathbb{R}^d mit charakteristischer Funktion ψ die folgende Zerlegung bezüglich einer regulären Funktion h_u mit $u \in C$:

1. *Starte einen h_u -transformierten Lévy-Prozess (\hat{X}_t) mit charakteristischer Funktion ψ^{h_u} in 0.*
2. *Erzeuge eine exponentiell verteilte und unabhängige Zufallsvariable Z mit Erwartungswert $\|u\|^{-1}$, und setze*

$$\tau := \sup\{t \geq 0 : \hat{X}_t \notin H_u(Z), \langle u, \hat{X}_s \rangle < \langle u, \hat{X}_t \rangle \forall s < t\}.$$

Dies ist der erste Zeitpunkt, zu dem der Prozess (\hat{X}_t) am nächsten am Halbraum $H_u(Z)$ ist, aber noch nicht in diesen eingetreten ist.

3. *Starte einen weiteren Lévy-Prozess (\check{X}_t) mit charakteristischer Funktion ψ im Startzustand $\check{X}_0 = \hat{X}_\tau$, der darauf bedingt ist, nicht in den Halbraum $H_u(Z)$ einzutreten. Setze*

$$\bar{X}_t := \begin{cases} \hat{X}_t & t < \tau \\ \check{X}_{t-\tau} & t \geq \tau. \end{cases}$$

Dann ist (\bar{X}_t) ein Lévy-Prozess mit charakteristischer Funktion ψ .

Die duale Zerlegung bietet in diesem Fall keine wesentlichen neuen Einblicke, weshalb die Details ausgespart bleiben. Allerdings ist es durchaus sehr lohnenswerte einzelne spezielle Beispiele näher zu untersuchen, auch um alte Resultate direkt in diese Arbeit aus neuer Sicht einzubetten.

Beispiel 1: Brownsche Bewegung mit Drift

Das bekannte von Williams und Pitman untersuchte ([Wil74, Pit75]) Beispiel stellt eine Brownsche Bewegung dar. Klassisch wurde eine eindimensionale Brownsche Bewegung mit Drift am Supremum oder Infimum zerlegt, wir können dies nun sofort auf höhere Dimensionen verallgemeinern.

Sei (B_t) eine d -dimensionale Brownsche Bewegung auf einem Wahrscheinlichkeitsraum mit Maß \mathbf{P} , und sei $\Sigma \in \mathbb{R}^{d \times d}$ die Kovarianzmatrix und $b \in \mathbb{R}^d \setminus \{0\}$ der Drift. Wir setzen Σ als regulär voraus, ansonsten kann (B_t) durch affin-lineare Transformation aus einer niedriger-dimensionalen Brownschen Bewegung erhalten werden. Die Laplace Transformierte φ ist gegeben durch

$$\log \varphi(\lambda) = \langle \lambda, b \rangle + \frac{1}{2} \lambda^T \Sigma \lambda.$$

Sei $u \in C$, d.h. es gilt $\log \varphi(u) = 0$, und dies bedeutet

$$\langle u, b \rangle + \frac{1}{2} u^T \Sigma u = 0. \quad (4.12)$$

Dann ist nach Satz 35 der h_u -transformierte Prozess wieder eine Brownsche Bewegung mit Kovarianzmatrix Σ und Drift $b^{h_u} = b + u^T \Sigma$. Diese Verteilung beschreibt gerade das Anfangsstück der Zerlegung.

Betrachten wir die Lösungsmenge der Gleichung (4.12) etwas genauer. Für $b = 0$ besitzt diese Gleichung nur die triviale Lösung $u = 0$, deshalb wurde dieser Fall oben ausgeschlossen. Sei also $b \neq 0$. In der Dimension $d = 1$ erhalten wir neben der trivialen Lösung genau eine weitere Lösung $u = -2b/\sigma^2$, wobei $\sigma^2 := \text{Var}(B_1)$ die Varianz ist. Der Drift der h_u -transformierten Brownschen Bewegung ist nach Gleichung (4.10) gegeben durch

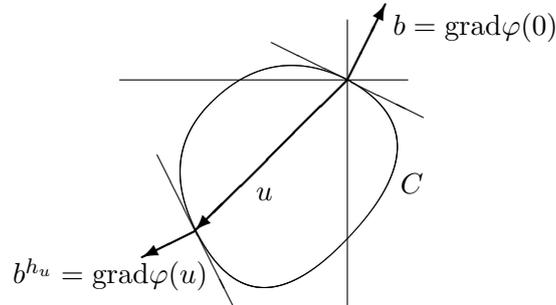
$$b^{h_u} = b + u\sigma^2 = -b.$$

Die obige Konstruktion entspricht der Zerlegung am Supremum (für negative b) bzw. am Infimum (für positive b), und das Anfangsstück hat gerade den entgegengesetzte Drift.

Für Dimensionen $d \geq 2$ und $b \neq 0$ besitzt die Gleichung (4.12) unendlich viele Lösungen, und für jede Lösung $u \in C$ gilt mit Gleichung (4.7)

$$\mathbf{E}_0^{h_u} X_1 = \text{grad} \varphi(u).$$

Da die Funktion φ nach Definition strikt konvex ist, ist die Lösungsmenge C der glatte Rand einer konvexen Menge, und in jedem Punkt u steht der Gradient $\text{grad}\varphi(u)$ senkrecht zur Tangente der Menge C im Punkt u .



Damit ist die Menge $\{\text{grad}\varphi(u) : u \in C\}$ isomorph zur $(d-1)$ -Sphäre S^{d-1} in \mathbb{R}^d , d.h. zu jeder Richtung $v \in S^{d-1}$ existiert ein $u \in C$, so dass gilt

$$\lim_{t \rightarrow \infty} \frac{B_t}{\|B_t\|} = \frac{b^{h_u}}{\|b^{h_u}\|} = v \quad \mathbf{P}^{h_u} \text{ - f.s.}$$

Damit ergibt sich die folgende Zerlegung einer d -dimensionalen Brownschen Bewegung mit Drift $b \neq 0$ und Kovarianzmatrix Σ bezüglich einer harmonischen Funktion h_u mit $u \in C$, wobei sich die Zerlegungszeit τ aufgrund der Stetigkeit wesentlich einfacher darstellen lässt.

1. *Starte eine Brownsche Bewegung (\hat{B}_t) mit Kovarianzmatrix Σ und Drift $\hat{b} = b + u^T \Sigma$.*
2. *Erzeuge eine exponentiell verteilte und unabhängig Zufallsvariable Z mit Erwartungswert $\|u\|^{-1}$, und setze*

$$\tau := \inf\{t \geq 0 : \langle u, \hat{B}_t \rangle \geq \|u\|Z\}.$$

3. *Starte eine weitere Brownsche Bewegung (\check{B}_t) mit Startpunkt $\check{B}_0 = \hat{B}_\tau$ und mit Drift b und Kovarianzmatrix Σ , die darauf bedingt ist, nicht in den Halbraum $H_u(Z)$ einzutreten. Setze*

$$\bar{B}_t := \begin{cases} \hat{B}_t & t < \tau \\ \check{B}_{t-\tau} & t \geq \tau. \end{cases}$$

Dann ist (\bar{B}_t) eine Brownsche Bewegung mit Drift b und Kovarianzmatrix Σ .

Diese Zerlegung entspricht für $d = 1$ exakt den von Williams in [Wil74] und von Pitman in [Pit75] angegebenen Zerlegungen; diese haben zusätzlich das zweite Teilstück der Zerlegung (also den post- τ -Prozess) charakterisiert; demnach ist der Prozess (\check{B}_t) eine Diffusion in $[\hat{B}_\tau, \infty)$ mit Startpunkt \hat{B}_τ und infinitesimalen Generator

$$\frac{1}{2} \frac{d^2}{dx^2} + \coth(x - \hat{B}_\tau) \frac{d}{dx}.$$

In Abschnitt 4.4 wird noch eine verwandte Zerlegung eines Bessel-Prozesses untersucht.

Für $d \geq 2$ sei die Kovarianzmatrix Σ die Einheitsmatrix und es sei $b \neq 0$ der Drift einer Brownschen Bewegung. Dann stellt $u := -2b$ eine Lösung der Gleichung 4.12 dar, und die zugehörige h_u -transformierte Brownsche Bewegung besitzt dann den Drift $b^{h_u} = -b$ in die genau entgegengesetzte Richtung. Der für die Zerlegung entscheidenden Halbraum

$H_u(Z) = H_{-2b}(Z)$ wird durch die Hyperebene $\{x \in \mathbb{R}^d : \langle b, x \rangle = -\|b\|Z\}$ definiert; diese liegt in diesem Fall auch orthogonal zum Drift b . Damit stellt eine solche spezielle Zerlegung eine einfache und direkte Verallgemeinerung in höhere Dimensionen der Zerlegung von Williams dar.

Beispiel 2: Poisson-Prozess mit Drift

Ein weiteres interessantes Beispiel stellt ein Poisson-Prozess dar. Ein solcher Prozess stellt den denkbar einfachsten Lévy-Prozess mit Sprüngen dar, diese sind immer der Größe 1 und treten in unabhängig exponentiell verteilten Zeitabständen mit der Rate $s > 0$ auf. Da ein solcher Prozess zunächst monoton wachsend ist (und somit sein Minimum zur Zeit 0 annimmt), benötigen wir für eine Zerlegung zusätzlich einen Drift b , so dass $-s < b < 0$ ist. Damit konvergiert der Prozess fast sicher gegen $+\infty$, und durch den negativen linearen Drift c nimmt der Prozess sein Minimum fast sicher zu einem endlichen Zeitpunkt $L > 0$ an.

Der Exponent der Laplace Transformierten für den oben beschriebenen Poisson-Prozess ist gegeben durch

$$\log \varphi(\lambda) = \lambda b + s(e^\lambda - 1).$$

Aufgrund der Konvexität von $\log \varphi$ und der Bedingung $-s < b < 0$ besitzt die Menge C der Nullstellen von $\log \varphi$ genau zwei Lösungen $u = 0$ und $v < 0$. Mit Hilfe von Gleichung (4.10) kann der h_v -transformierte Prozess wieder als ein Poisson-Prozess charakterisiert werden mit Drift b und Sprungrate $e^v s$, d.h. das Anfangsstück (\hat{X}_t) der Zerlegung ist ein Poisson-Prozess mit identischer linearer Drift aber mit einer kleineren Sprungrate. Wegen der strikten Konvexität gilt außerdem

$$\mathbb{E}_0 \hat{X}_1 = \varphi'(v) < 0,$$

also gilt $\hat{X}_t \rightarrow -\infty$ für $t \rightarrow \infty$. Damit haben wir die Verteilung des Anfangsstücks exakt beschrieben.

In diesem Fall kann mit Hilfe der Technik aus Abschnitt 3.5 die Verteilung des post- τ -Prozesses (\check{X}_t) als Transformation der eingeschränkten Kerne (P_t^m) dargestellt werden mit

$$P_t^m(x, dy) := \mathbf{P}_x(X_t \in dy : X_s \geq m \forall s \in [0, t]),$$

denn die zufällige Schranke Z kann vom Prozess (X_t) unter \mathbf{P} nur stetig nach unten durchbrochen werden. Gegeben das Anfangsstück (\hat{X}_t) der Zerlegung bis zum Zeitpunkt τ muss der Prozess (\check{X}_t) deshalb sofort einen Sprung nach oben machen, d.h. es ist $\check{X}_0 = \hat{X}_\tau + 1$. Bedingt auf $Z = m$ sind die Übergangskerne (K_t^m) des post- τ -Prozesses dann gegeben durch

$$K_t^m(x, dy) = \frac{m - h_v(y)}{m - h_v(x)} P_t^m(x, dy).$$

Bemerkung Salminen hat in [Sal81] für einfache Poisson-Prozesse, die keinen linearen Drift dafür aber sowohl negativen als auch positiven Sprüngen besitzen, weitere minimale Raum-Zeit-reguläre Funktionen k_z mit $z \in (-\infty, \infty)$ angegeben, die sowohl die Sprungrate als auch deren Verteilung ändern. Diese eröffnen weitere Zerlegungen ähnlicher Gestalt. \square

4.2 Beispiel: Driftlose Brownsche Bewegung

Im vorigen Abschnitt wurde mittels einer Brownschen Bewegung mit Drift eine weitere Brownsche Bewegung mit einem anderen Drift konstruiert. Will man eine driftlose Brownsche Bewegung in \mathbb{R}^d zerlegen, so benötigt man andere reguläre Funktionen, da für $b = 0$ die Gleichung (4.12) nur die triviale Lösung $u = 0$ besitzt, und damit auf diesem Weg auch nur die triviale konstante reguläre Funktion erhalten werden kann.

Den Ausweg bilden Raum-Zeit-reguläre Funktionen $h : \mathbb{R}_+ \times \mathbb{R}^d$, die die folgende definierende Eigenschaft besitzen

$$h(t, x) = \int_{\mathbb{R}^d} P_s(x, dy) h(s + t, y) \quad \text{für alle } s, t \in \mathbb{R}_+ \text{ und } x \in \mathbb{R}^d.$$

Bekanntlich ist für eine driftlose Brownsche Bewegung (B_t) in Dimension 1 der Prozess $\exp(bB_t - \frac{1}{2}b^2t)$ für alle $b \in \mathbb{R}$ ein Martingal, also ist $h_b(t, x) := \exp(bx - \frac{1}{2}b^2t)$ eine Raum-Zeit-reguläre Funktion, und allgemein ist für Dimension $d \geq 1$ und $b \in \mathbb{R}^d$ die folgende Funktion Raum-Zeit-regulär

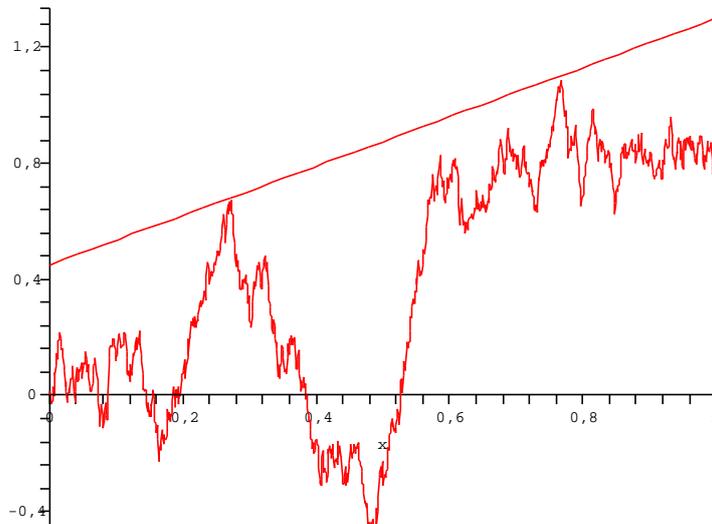
$$h_b(t, x) := \exp(\langle b, x \rangle - \frac{1}{2}\|b\|^2t), \quad t \in \mathbb{R}_+ \text{ und } x \in \mathbb{R}^d. \quad (4.13)$$

Die zugehörigen Kerne eines h_b -transformierten Prozesses müssen nun per se nicht mehr zeithomogen sein, es gilt jedoch in diesem Spezialfall für alle $s, t \in \mathbb{R}_+$ und $x, y \in \mathbb{R}^d$

$$\frac{h(t+s, y)}{h(t, x)} P_s(x, dy) = \frac{1}{(2\pi s)^{d/2}} \exp\left(-\frac{\|(y-sb) - x\|^2}{2s}\right) =: P_s^b(x, dy),$$

also ist auch der h_b -transformierte Prozess wieder zeithomogen, und entspricht gerade einer Brownschen Bewegung mit Drift b .

Für eine Zerlegung sind die Niveau-Mengen von h_b von besonderem Interesse. Aus der Definition in Gleichung (4.13) wird klar, dass zu einem festen Zeitpunkt t jede Niveaumenge von $h_b(t, \cdot)$ eine Hyperebenen im \mathbb{R}^d bildet, deren Abstand vom Ursprung linear mit der Zeit mit Geschwindigkeit $\|b\|/2$ wächst. Also sind die Niveau-Mengen von h_b in der Raum-Zeit wiederum Hyperebenen.



Brownsche Bewegung in Dimension 1 mit Drift $b = 1,7$ und eine Niveaulinie von h_b , als zufällige Schranke mit linear wachsendem Abstand vom Ursprung.

Mit dieser Erkenntnis kann der Zerlegungszeitpunkt τ einfacher beschrieben werden. Wählen wir den Startpunkt $o = 0 \in \mathbb{R}^d$, dann ist $h_b(o) = 1$, und als einen geeigneteren Ersatz für die Zufallsvariable Y mit $\mathbb{P}(Y > y) = 1/y$ setzen wir

$$Z := \|b\|^{-1}Y,$$

so dass $\mathbb{P}(Z > z) = \exp(-\|b\|z)$ gilt, also Z exponentiell zum Parameter $\|b\|$ verteilt ist. Betrachten wir nun den Abstand eines Punktes x zur Hyperebene zum Zeitpunkt t , dann gilt

$$\begin{aligned} h_b(t, x) > Y &\Leftrightarrow \frac{\langle b, x \rangle}{\|b\|} > \frac{\log Y}{\|b\|} + \frac{1}{2}\|b\|t \\ &\Leftrightarrow \frac{\langle b, x \rangle}{\|b\|} > Z + \frac{1}{2}\|b\|t \\ &\Leftrightarrow (t, x) \in H_b(Z), \end{aligned}$$

wobei $H_b(Z) := \{(t, x) \in \mathbb{R}_+ \times \mathbb{R}^d : \langle b, x \rangle - \frac{1}{2}\|b\|^2t > \|b\|Z\}$ ein Halbraum in $\mathbb{R}_+ \times \mathbb{R}^d$ mit Abstand Z vom Ursprung ist. Damit ergibt sich folgende Zerlegung einer freien Brownschen Bewegung:

1. *Starte eine Brownsche Bewegung (\hat{B}_t) in \mathbb{R}^d mit Drift b und mit Kovarianzen $\Sigma_{i,j} = \delta_{i,j}$.*
2. *Erzeuge eine exponentiell verteilte und unabhängig Zufallsvariable Z mit Erwartungswert $\|b\|^{-1}$, und setze*

$$\tau := \inf\{t \geq 0 : \langle b, \hat{B}_t \rangle - \frac{1}{2}\|b\|^2t \geq \|b\|Z\}.$$

3. *Starte eine weitere driftlose Brownsche Bewegung (\check{B}_t) mit Startpunkt $\check{B}_0 = \hat{B}_\tau$ und mit Kovarianzmatrix Σ , die darauf bedingt ist, nicht in den Halbraum $H_b(Z + \frac{1}{2}\|b\|\tau)$ in der Raum-Zeit $\mathbb{R}_+ \times \mathbb{R}^d$ einzutreten. Setze*

$$\bar{B}_t := \begin{cases} \hat{B}_t & t < \tau \\ \check{B}_{t-\tau} & t \geq \tau. \end{cases}$$

Dann ist (\bar{B}_t) eine driftlose Brownsche Bewegung mit Kovarianzmatrix $\Sigma = \delta_{i,j}$.

Mittels der dualen Zerlegung kann eine im Ursprung startende Brownsche Bewegung mit Drift $b \neq 0$ in Dimension $d \geq 1$ konstruiert werden, indem zuerst eine driftlose Brownsche Bewegung gestartet wird, bis diese in den Halbraum $H_b(1/V)$ durchstößt, wobei V unabhängig zum Parameter $\|b\|$ exponentialverteilt ist. Dann wird eine weitere Brownsche Bewegung wieder mit Drift b angehängt, die darauf bedingt ist, nicht in den Halbraum $H_b(1/V)$ einzutreten.

Bemerkung Diese Zerlegungen können auch als Spezialfall der Zerlegung einer Brownschen Bewegung mit Drift gesehen werden, wie sie als Beispiel 1 in Abschnitt 4.1 dargestellt wurde. Denn betrachten wir die Zerlegung einer Brownschen Bewegung (B'_t) mit Drift $-b/2$ in ein Anfangsstück mit entgegengesetztem Drift $b/2$ und in ein bedingtes Endstück mit Drift $-b/2$.

Dies entspricht vermittels einer Zentrierung $B_t := B'_t + t\frac{b}{2}$ gerade einer Zerlegung der nunmehr driftlosen Brownschen Bewegung (B_t) in ein Anfangsstück mit Drift b und bedingtem Endstück ohne Drift wie oben angegeben.

Eine ähnliche Interpretation gilt auch für die duale Zerlegung, diese kann aus einer Zerlegung einer Brownschen Bewegung (B'_t) mit Drift $b/2$ durch vermittels $B_t^* := B'_t + t\frac{b}{2}$ gewonnen werden. \square

4.3 Beispiel: Brownsche Brücke

Eine d -dimensionale Brownsche Brücke $(X_t)_{0 \leq t \leq 1}$ ist definiert als der folgende Prozess auf dem Zeitintervall $[0, 1]$

$$X_t := B_t - tB_1 \quad \text{wobei } (B_t) \text{ eine Brownsche Bewegung in } \mathbb{R}^d \text{ ist.}$$

Die Verteilung der Brownschen Brücke kann als h -Transformation des Wienermaßes \mathbf{P} beschrieben werden mit der Raum-Zeit-regulären Funktion $h : [0, 1] \times \mathbb{R}^d \rightarrow \mathbb{R}$ definiert durch

$$h(t, x) := \frac{1}{(1-t)^{d/2}} \exp\left(-\frac{\|x\|^2}{2(1-t)}\right)$$

(siehe Abschnitt 40 in [RW00]).

Betrachten wir das Konvergenzverhalten von $h(t, B_t)$ für $t \rightarrow 1$ für eine standard Brownschen Bewegung (B_t) in \mathbb{R}^d auf einem Wahrscheinlichkeitsraum mit Maß \mathbf{P} . Es ist fast sicher $B_1 \neq 0$, und damit existiert fast sicher ein zufälliges (da pfadabhängiges) $\varepsilon > 0$, so dass die Brownsche Bewegung im ganzen Zeitintervall $[1 - \varepsilon, 1]$ einen Mindestabstand I von 0 hält, also:

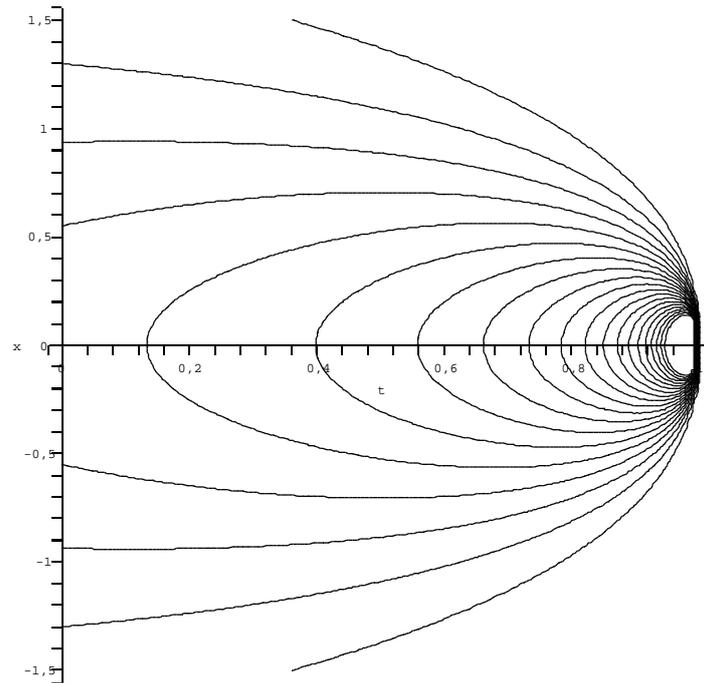
Es existiert \mathbf{P} -fast sicher ein $\varepsilon > 0$, so dass $I := \inf\{\|B_s\| : s \in [1 - \varepsilon, 1]\} > 0$.

Damit gilt dann für alle Realisierungen (B_t) außer einer \mathbf{P} -Nullmenge

$$\begin{aligned} 0 &\leq \lim_{t \rightarrow 1} h(t, B_t) \leq \limsup_{t \rightarrow 1} h(t, B_t) = \limsup_{t \rightarrow 1, t > 1 - \varepsilon} h(t, B_t) \quad \mathbf{P} - \text{f.s.} \\ &= \limsup_{t \rightarrow 1, t > 1 - \varepsilon} \frac{1}{(1-t)^{d/2}} \exp\left(-\frac{\|B_t\|^2}{2(1-t)}\right) \\ &\leq \limsup_{t \rightarrow 1, t > 1 - \varepsilon} \frac{1}{(1-t)^{d/2}} \exp\left(-\frac{I^2}{2(1-t)}\right) \quad \mathbf{P} - \text{f.s.} \\ &= 0. \end{aligned}$$

Also gilt $h(t, B_t) \rightarrow 0$ für $t \rightarrow 1$ \mathbf{P} -fast sicher, und somit erfüllt die reguläre Funktion h die Bedingung 3.23, und die Zerlegungszeit τ ist somit fast sicher endlich. Durch die folgende Fortsetzung von h auf den größeren Raum $\mathbb{R}_+ \times \mathbb{R}^d$ kann die Brownsche Brücke auch als h -Transformation einer freien Brownschen Bewegung auf der gesamten Zeitachse \mathbb{R}_+ verstanden werden, wobei der transformierte Prozess zur Zeit 1 gekillt wird.

$$h(t, x) := \begin{cases} \frac{1}{(1-t)^{d/2}} \exp\left(-\frac{\|x\|^2}{2(1-t)}\right) & \text{für } t < 1 \\ \infty & \text{für } t = 1 \text{ und } x = 0 \\ 0 & \text{für } t = 1 \text{ und } x \neq 0 \\ 0 & \text{für } t > 1. \end{cases}$$



Niveaulinien der Funktion $h(t, x)$. Kleinere Kurven liegen höher, und im Punkt $(1, 0)$ besitzt die Funktion einen Pol.

Aus der dualen Zerlegung ergibt sich damit die folgende Konstruktion einer Brownschen Brücke:

1. *Starte eine driftlose Brownsche Bewegung (\hat{B}_t) mit Varianz 1 im Punkt 0 auf einem Wahrscheinlichkeitsraum mit Maß \mathbb{P} .*
2. *Erzeuge eine unabhängige, uniform in $[0, 1]$ verteilte Zufallsvariable U , und setze*

$$\tau := \inf\{t \geq 0 : h(t, \hat{B}_t) = U\}.$$
3. *Starte eine Brownsche Brücke $(\check{B}_t)_{t \in [\tau, 1]}$ von (τ, \hat{B}_τ) nach $(1, 0)$, die darauf bedingt ist, dass $h(t, \check{B}_t) \geq U$ für alle $t \in [\tau, 1]$ ist. Setze*

$$\bar{B}_t := \begin{cases} \hat{B}_t & t < \tau \\ \check{B}_t & t \geq \tau. \end{cases}$$

Dann ist (\bar{B}_t) eine Brownsche Brücke von 0 nach 0.

Die duale Zerlegung stellt eine alternative Konstruktion einer freien Brownschen Bewegung dar, deren Anfangsstück durch eine Brownsche Brücke beschrieben wird, und deren Endstück eine Brownsche Bewegung ist darauf bedingt, eine Niveaulinie von $h(t, x)$ nicht mehr zu unterschreiten.

4.4 Beispiel: Bessel-Prozess

Ein weiteres interessantes Beispiel sind dreidimensionale Bessel-Prozesse (kurz Bessel-3-Prozesse), also eine eindimensionale Brownsche Bewegung, die darauf bedingt ist, sich nur in der positiven Halbebene zu bewegen.

Verteilung eines Bessel-Prozesses

Um die Verteilung eines Bessel-3-Prozesses zu beschreiben greifen wir auf [Wil74] zurück. Bezeichnen wir mit $\text{BES}^b(3, 0)$ einen dreidimensionalen Bessel-Prozess mit Startpunkt $b \geq 0$ und Barriere 0, dann wird die Verteilung durch die Kernfamilie (K_t) zusammen mit einem Eintrittsmaß charakterisiert, wobei die Übergangskerne als h -Transformation der Kernfamilie (P_t^-) einer in 0 gestoppten freien Brownschen Bewegung darstellbar sind. Die für diese gestoppte Brownsche Bewegung reguläre Funktion h ist gerade $h(x) = x$. Es gilt also

$$K_t(x, dy) = \frac{y}{x} P_t^-(x, dy)$$

$$P_t^-(x, dy) = P_t(x, dy) - P_t(-x, dy).$$

Die Konstruktion von (P_t^-) geht aus dem Spiegelungsprinzip hervor, so dass diese Kerne nur noch Masse auf Pfaden haben, die sich vollständig innerhalb der positiven Halbebene befinden. Falls der Bessel-Prozess im Punkt $b = 0$ startet, wird zusätzlich ein Eintrittsmaß (D_t) benötigt, dieses ist gegeben durch

$$D_t(0, dy) = 2y \mathbf{P}(T_y \in dt) / dt$$

$$\mathbf{P}(T_x \in dt) / dt = \frac{x}{(2\pi t^3)^{1/2}} \exp\left(-\frac{x^2}{2t}\right),$$

siehe Abschnitt 3 in [Wil74] und § 1.7 in [IM65]. Allgemein bezeichnen wir einen Bessel-3-Prozess (B_t) mit Startpunkt b und Barriere c mit $\text{BES}^b(3, c)$, es ist dann der verschobene Prozess $(B_t - c)$ wieder ein $\text{BES}^{b-c}(3, 0)$, und somit sind die passenden Kerne (K_t^c) gegeben durch

$$K_t^c(x, dy) = \frac{y - c}{x - c} P_t^-(x - c, dy - c).$$

Zerlegung

Aus der dualen Zerlegung bezüglich der regulären Funktion $h(x) = x$ erhalten wir die folgende Konstruktion eines Bessel-3-Prozesses, die exakt der Williams Zerlegung (Theorem 3.1 in [Wil74]) entspricht.

1. Sei $b > 0$ und (\check{B}_t^*) eine freie, in b startende Brownsche Bewegung.
2. Sei weiter U eine unabhängige, uniform in $[0, b]$ verteilte Zufallsvariable und setze

$$\tau^* := \inf\{t \geq 0 : \check{B}_t^* = U\}.$$

3. Starte in 0 einen Bessel-3-Prozess (\hat{B}_t^*) und setze

$$\bar{B}_t^* := \begin{cases} \check{B}_t^* & t < \tau^* \\ \hat{B}_{t-\tau^*}^* + \check{B}_\tau^* & t \geq \tau^*. \end{cases}$$

Dann ist (\bar{B}_t^*) eine dreidimensionaler Bessel-Prozess mit Startpunkt b .

Umgekehrt erhalten wir aus der direkten Zerlegung eine Konstruktion für eine in $b > 0$ startende und in 0 gestoppte Brownsche Bewegung. Dabei wird das Anfangsstück durch einen $\text{BES}^b(3, 0)$ -Prozess beschrieben, bis dieser zu einem Zeitpunkt τ eine (bis auf eine Verschiebung b) exponentiell verteilte Schranke Y trifft, und der post- τ -Prozess ist wiederum ein gespiegelter $\text{BES}^Y(3, Y)$ -Prozess (der also immer unterhalb der Schranke Y bleibt), der in 0 gestoppt wird.

4.5 Beispiel: Verzweigungsprozesse

In Abschnitt 2.6 wurde schon das Modell eines zeitdiskreten Verzweigungsprozesses in einem Typenraum dargestellt. Es existieren ebenfalls zeitkontinuierliche Modelle für Verzweigungsprozesse, wir wollen hier zwei Beispiele kurz anreißen. Den Beispielen liegt die Arbeit [Ove94] zugrunde, dort wurden entsprechende reguläre Funktionen bestimmt und die jeweils assoziierte transformierte Verteilung beschrieben. Alle Beispiele in diesem Abschnitt beziehen sich rein auf die Populationsgröße, es liegt kein Typenraum zu Grunde.

Kritische Verzweigung

Betrachten wir einen kritischen Verzweigungsprozess mit einem simplen binären Verzweigungsmechanismus, bei dem jedes Teilchen eine exponentiell verteilte Lebenszeit besitzt, und dann mit Wahrscheinlichkeit $1/2$ stirbt oder sich mit Wahrscheinlichkeit $1/2$ verdoppelt. Der Einfachheit halber sei die Intensität der Lebenszeit 2. Damit kann der Prozess als ein Punktprozess mit Sprüngen der Größe ± 1 interpretiert werden, denn fast sicher sterben keine zwei Teilchen zu einem gemeinsamen Zeitpunkt.

Die Kernfamilie (P_t) , die den Verzweigungsmechanismus beschreibt, ist gegeben durch

$$P_t(x, y) = \begin{cases} \sum_{i=1}^x \binom{x}{i} \binom{y-1}{i-1} \frac{t^{x+y-2i}}{(t+1)^{x+y}} & \text{für } y > 0 \\ \frac{t^x}{(t+1)^x} & \text{für } y = 0, \end{cases}$$

und es bezeichne \mathbf{P} die gesamte Verteilung des Prozesses. Dann ist die Funktion $h(x) := x$ eine minimale unbeschränkte reguläre Funktion, und als solche für eine Zerlegung geeignet. Es existieren noch weitere Raum-Zeit-reguläre Funktionen, siehe [Ove94]. Die Wahl dieser Funktion ist auf die besonders einfache Darstellung einer Zerlegung zurückzuführen.

Der h -transformierte Prozess besitzt die folgenden verschiedenen Interpretationen. 1. Der Prozess (X_t) mit der h -transformierten Verteilung \mathbf{P}^h entspricht einem interagierendem Verzweigungsprozess in dem folgenden Sinne: Falls zu einem Zeitpunkt k Teilchen am Leben sind, dann besitzt jedes Teilchen eine exponentielle Lebenszeit mit dem Parameter 2, und sobald das Teilchen stirbt, produziert es mit der Wahrscheinlichkeit $\frac{k+1}{2k}$ zwei und mit Wahrscheinlichkeit $\frac{k-1}{2k}$ keine neue Nachkommen. 2. Der Prozess mit Verteilung \mathbf{P}^h kann auch als ein Verzweigungsprozess mit unabhängiger Immigration mit Rate 1 interpretiert werden. In exponentiell verteilten Abständen ergibt sich eines der beiden folgenden Ereignisse: Mit Wahrscheinlichkeit $1/k$ immigriert ein neues Teilchen in die Population und mit Wahrscheinlichkeit $1 - 1/k$ findet wieder eine zufällige binäre Verzweigung statt. 3. Eine dritte Interpretation sieht die Verteilung \mathbf{P}^h als Verteilung des ursprünglichen Verzweigungsprozesses, unter der Bedingung, nicht auszusterben.

Richten wir unser Blick auf eine Zerlegung, dann ist die Zerlegungszeit τ eines in $\hat{X}_0 = 1$ startenden Anfangsstücks (\hat{X}_t) zunächst definiert als

$$\tau := \sup\{t \geq 0 : \hat{X}_t < Y, \hat{X}_s < \hat{X}_t \forall s < t\},$$

wobei die Zufallsvariable Y die Verteilung $\mathbb{P}(Y > y) = 1/y$ besitzt. Da der Verzweigungsprozess nur ganzzahlige Werte annehmen kann, ist die Zeit τ auch wie folgt darstellbar

$$\tau = \inf\{t \geq 0 : \hat{X}_t = Z\},$$

mit der ganzzahligen Zufallsvariable $Z := \lfloor Y \rfloor$, die die Verteilung $\mathbb{P}(Z = z) = 1/(z^2 + z)$ für $z \in \mathbb{N}$ besitzt.

Wir erhalten folgende Zerlegung eines binären Verzweigungsprozesses am Zeitpunkt der maximalen Populationsgröße

1. Sei (\hat{X}_t) ein in $X_0 := 1$ startender h -transformierter binärer Verzweigungsprozess mit Verteilung \mathbf{P}^h wie oben auf einem Wahrscheinlichkeitsraum mit Maß \mathbb{P} und sei Z eine unabhängige ganzzahlige Zufallsvariable mit Verteilung

$$\mathbb{P}(Z = z) = \frac{1}{z(z+1)} \quad z \in \mathbb{N}.$$

2. Setze

$$\tau := \inf\{t \geq 0 : \hat{X}_t = Z\}.$$

3. Sei (\check{X}_t) ein binärer Verzweigungsprozess mit Startzustand $\check{X}_0 = \hat{X}_\tau$ und mit der Verteilung \mathbf{P} bedingt auf $\{\check{X}_t \leq Z \forall t \geq 0\}$. Setze

$$\bar{X}_t := \begin{cases} \hat{X}_t & t < \tau \\ \check{X}_{t-\tau} & t \geq \tau. \end{cases}$$

Dann ist (\bar{X}_n) ein kritischer binärer Verzweigungsprozess mit Verteilung \mathbf{P} .

Feller-Prozess

Ein weiteres Beispiel eines zeitkontinuierlichen Verzweigungsprozesses ist durch den Feller-Prozess gegeben (siehe wieder [Ove94]). Diesen erhält man als Grenzwert einer Folge passende skalierte kritischer Verzweigungsprozesse wie im vorigen Teilabschnitt dargestellt. Im n -ten Approximationsschritt besitzt jedes Teilchen Masse $1/n$ und eine exponentiell verteilte Lebenszeit mit Intensität $2n$. Der daraus resultierende Grenzprozess ist schließlich ein Feller-Prozess auf \mathbb{R}_+ . Dieser kann auch als Lösung der stochastischen Differentialgleichung

$$dX_t = \sqrt{2X_t} dW_t$$

dargestellt werden, wobei (W_t) eine Brownsche Bewegung ist.

Es ist in diesem Beispiel wieder $h(x) = x$ eine minimale, unbeschränkte, reguläre Funktion, mit der eine Zerlegung des Prozesses durchgeführt werden kann. Der mit h assoziierte h -transformierte Prozess ist Lösung der stochastischen Differentialgleichung

$$dX_t = \sqrt{2X_t} dW_t + 2dt,$$

und dieser kann wieder als ein auf Überleben bedingter Feller-Prozess interpretiert werden.

Die Zerlegung eines Feller-Prozesses mit h findet analog dem vorigen Beispiel zum Zeitpunkt der maximalen Populationsgröße statt. Die Details unterscheiden sich nicht wesentlich von den anderen Beispielen, weshalb wir sie an dieser Stelle auslassen.

Literaturverzeichnis

- [AD01] ALILI, L. und R. A. DONEY: *Martin boundaries associated with a killed random walk*. Ann. Inst. H. Poincaré Probab. Statist., 37(3):313–338, 2001.
- [BD94] BERTOIN, JEAN und RONALD A. DONEY: *On conditioning a random walk to stay nonnegative*. Ann. Probab., 22(4):2152–2167, 1994.
- [Ber91] BERTOIN, JEAN: *Sur la décomposition de la trajectoire d'un processus de Lévy spectralement positif en son infimum*. Ann. Inst. Henri Poincaré, 27(4):537–547, 1991.
- [Ber92] BERTOIN, JEAN: *An Extension of Pitman's Theorem for Spectrally Positive Lévy Processes*. The Annals of Probability, 20(3):1464–1483, 1992.
- [Ber93] BERTOIN, JEAN: *Splitting at the Infimum and Excursions in Half-Lines for Random Walks and Lévy Processes*. Stochastic Processes and their Applications, 47:17–35, 1993.
- [Ber96] BERTOIN, JEAN: *Lévy Processes*. Cambridge Tracts in Mathematics, 1996.
- [BK64] BLACKWELL, DAVID und DAVID KENDALL: *The Martin Boundary for Polya's Urn Scheme, and an Application to Stochastic Population Growth*. Journal of Applied Probability, 1:284–296, 1964.
- [BK04] BIGGINS, J.D. und ANDREAS E. KYPRIANOU: *Measure change in multitype branching*. Advances in Applied Probability, 36(2), 2004.
- [Bre86] BREIMAN, LEO: *Probability*. Addison-Wesley, 1986.
- [Cha96] CHAUMONT, LOÏC: *Conditioning and Path Decompositions for Lévy Processes*. Stochastic Processes and their Applications, 64:39–54, 1996.
- [CS91] CARTWRIGHT, DONALD I. und STANLEY SAWYER: *The Martin boundary for general isotropic random walks in a tree*. J. Theoret. Probab., 4(1):111–136, 1991.
- [Don98] DONEY, R. A.: *The Martin boundary and ratio limit theorems for killed random walks*. J. London Math. Soc. (2), 58(3):761–768, 1998.
- [DSW60] DOOB, J. L., J. L. SNELL und R. E. WILLIAMSON: *Application of boundary theory to sums of independent random variables*. In: *Contributions to probability and statistics*, Seiten 182–197. Stanford Univ. Press, Stanford, Calif., 1960.
- [GP80] GREENWOOD, PRISCILLA und JIM PITMAN: *Fluctuation identities for Lévy processes and splitting at the maximum*. Adv. in Appl. Probab., 12(4):893–902, 1980.

- [Har02] HARRIS, THEODORE E.: *The theory of branching processes*. Dover Phoenix Editions. Dover Publications Inc., Mineola, NY, 2002. Corrected reprint of the 1963 original [Springer, Berlin; MR **29** #664].
- [Hen63] HENNEQUIN, PAUL-LOUIS: *Processus de Markoff en cascade*. Ann. Inst. H. Poincaré, 18:109–195 (1963), 1963.
- [IM65] ITÔ, KIYOSHI und HENRY P. MCKEAN, JR.: *Diffusion processes and their sample paths*. Die Grundlehren der Mathematischen Wissenschaften, Band 125. Academic Press Inc., Publishers, New York, 1965.
- [Jac74] JACOBSEN, MARTIN: *Splitting Times for Markov Processes and a Generalised Markov Property for Diffusions*. Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete, 30:27–43, 1974.
- [Kal97] KALLENBERG, OLAV: *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, 1997.
- [KM04] KERSTING, GÖTZ und MEMİŞOĞLU, KAYA: *Path decompositions for Markov chains*. The Annals of Probability, 32(2):1370–1390, 2004.
- [Mil77] MILLAR, P. WARWICK: *Zero-One Laws and the Minimum of a Markov Process*. Transactions of the American Mathematical Society, 226:365–391, 1977.
- [Mil78] MILLAR, P. WARWICK: *A Path Decomposition For Markov Processes*. The Annals of Probability, 6(2):345–348, 1978.
- [NS66] NEY, P. und SPITZER, F.: *The Martin Boundary for Random Walk*. Trans. Amer. Math. Soc., 121:116–132, 1966.
- [Ove94] OVERBECK, LUDGER: *Martin boundaries of some branching processes*. Ann. Inst. H. Poincaré Probab. Statist., 30(2):181–195, 1994.
- [Pit74] PITMAN, JAMES W.: *Path Decomposition for Conditional Brownian Motion*. Preprint No. 11, 1974. Institute of Mathematical Statistics, University of Copenhagen.
- [Pit75] PITMAN, JAMES W.: *One-Dimensional Brownian Motion and the Three-Dimensional Bessel Process*. Adv. Appl. Prob., 7:511–526, 1975.
- [PS73] PITTINGER, A.O. und SHIH, C.T.: *Coterminal Families and the Strong Markov Property*. Transactions of the American Mathematical Society, 182:1–42, 1973.
- [RW94] ROGERS, L. C. G. und WILLIAMS, DAVID: *Diffusions, Markov Processes and Martingales*, Band 1 der Reihe *Cambridge Mathematical Library*. Cambridge University Press, 1994.
- [RW00] ROGERS, L. C. G. und WILLIAMS, DAVID: *Diffusions, Markov Processes and Martingales*, Band 2 der Reihe *Cambridge Mathematical Library*. Cambridge University Press, 2000.
- [Sal81] SALMINEN, P.: *Martin boundaries for some space-time Markov processes*. Z. Wahrsch. Verw. Gebiete, 55(1):41–53, 1981.

- [Saw97] SAWYER, STANLEY A.: *Martin Boundary and Random Walks*. Contemporary Mathematics, 206:17–44, 1997.
- [Shi95] SHIRYAEV, A.N.: *Probability*, Band 95 der Reihe *Graduate Texts in Mathematics*. Springer-Verlag New York Inc., 1995.
- [Tan89] TANAKA, HIROSHI: *Time Reversal of Random Walks in One-Dimension*. Tokyo Journal of Mathematics, 12(1):159–174, 1989.
- [Tan90] TANAKA, HIROSHI: *Time Reversal of Random Walks in R^d* . Tokyo Journal of Mathematics, 13(2):375–389, 1990.
- [Wil74] WILLIAMS, DAVID: *Path Decomposition and Continuity of Local Time for One-Dimensional Diffusions*. Proc. London Math. Soc., 28:738–768, 1974.
- [Woe94] WOESS, WOLFGANG: *Random Walks on Infinite Graphs and Groups – A Survey on Selected Topics*. Bull. London Math. Soc., 26:1–60, 1994.

LEBENS LAUF

KAYA MEMIŞOĞLU

Geburtsdatum:

15. Oktober 1975 (Hanau am Main)

1982 bis 1988:

Johannes Gutenberg Schule
Hainburg

1988 bis 1995:

Gymnasialzweig der Einhardschule
Seligenstadt

Juni 1995:

Abitur

1995 bis 2001:

Studium Mathematik mit Nebenfach Informatik
an der Johann Wolfgang Goethe-Universität
Frankfurt am Main

Dezember 2001:

Diplom Mathematik mit Nebenfach Informatik
Diplomarbeit „ Zerlegung eines Lévy Prozesses am Infimum “
betreut von Prof. Kersting mit Zweitgutachter Prof. Wakolbinger.
Diplomprüfung bei Prof. Kersting, Prof. Schwarz
und Prof. Schnitger (Nebenfach)

Seit Januar 2002:

Wissenschaftlicher Mitarbeiter
am Institut für Stochastik und mathematische Informatik
Johann Wolfgang Goethe-Universität
Frankfurt am Main