

Asymptotische Analyse zufälliger diskreter Strukturen mit Methoden der diskreten Potentialtheorie

Von der Fakultät für Mathematik und Physik
der Gottfried Wilhelm Leibniz Universität Hannover
zur Erlangung des Grades
Doktor der Naturwissenschaften
Dr. rer. nat.
genehmigte Dissertation
von

Dipl.-Math. Igor Michailow
geboren am 08.01.1985 in Leningrad

2015

Referent: Prof. Dr. Rudolf Grübel
Korreferent: Prof. Dr. Ludwig Baringhaus

Tag der Promotion: 4. Juni 2015

Zusammenfassung

Gegenstand dieser Arbeit ist die Analyse des asymptotischen Verhaltens zufälliger diskreter Strukturen mit dem Fokus auf zufälligen Bäumen. In der Informatik gehören Bäume zu den klassischen Datenstrukturen und werden zur Behandlung von Such- und Sortierproblemen eingesetzt. Zu dem Zweck der asymptotischen Analyse werden diskrete Strukturen als Zustände einer kombinatorischen Markov-Kette aufgefasst. Mit den Methoden der diskreten Potentialtheorie wird der Zustandsraum der Markov-Kette so erweitert, dass die Zufallsgrößen einen fast sicheren Grenzwert haben. In der Literatur ist dieser Ansatz als Doob-Martin-Kompaktifizierung bekannt. Konvergenzresultate auf der Ebene der diskreten Strukturen können, wie von R. Grübel in [13] gezeigt, auf starke Grenzwertsätze für Funktionale dieser Strukturen führen.

Zu Beginn skizzieren wir die Martinrandtheorie allgemeiner irreduzibler transienter Markov-Ketten. Anschließend überprüfen wir die Übertragbarkeit der für unsere Zwecke benötigten Resultate auf die Situation kombinatorischer Markov-Ketten. Als nächstes beschreiben wir ausführlich die Anwendung der allgemeinen Theorie auf ein spezielles Urnenmodell, die Mendel-Urne. Im Anschluss daran bestimmen wir aufbauend auf der in [8] von S. N. Evans, R. Grübel und A. Wakolbinger vorgestellten „trickle-down“-Konstruktion die Doob-Martin-Kompaktifizierung zu den zufälligen quaternären Bäumen, wobei wir auf Resultate zur Mendel-Urne zurückgreifen. Quaternäre Bäume finden als Datenstruktur in der Computergraphik eine praktische Anwendung. Ein weiterer Schwerpunkt dieser Arbeit ist der „random recursive tree“-Prozess und die dazu verwandten zufälligen Ulam-Harris-Bäume. Mittels einer geeigneten algorithmischen Konstruktion erhalten wir einen alternativen Beweis der aus [8] bekannten Konvergenzresultate und darüber hinaus eine Darstellung des Limesobjektes mit Hilfe der Inputdaten. Außerdem verwenden wir den Ansatz aus [13] zum Beweis der fast sicheren Konvergenz verschiedener Funktionale zufälliger rekursiver Bäume. Zusätzlich stellen wir ein neues Baumfunktional, die horizontale Pfadlänge, vor und untersuchen deren Asymptotik. Zum Schluss setzen wir mit Hilfe der natürlichen Korrespondenz zwischen den binären Bäumen und den Ulam-Harris-Bäumen die entsprechenden Doob-Martin-Kompaktifizierungen zueinander in Beziehung.

Schlagwörter: Randwerttheorie, zufällige Bäume, Asymptotik von Baumfunktionalen.

Abstract

This dissertation is concerned with the investigation of the asymptotic behaviour of random discrete structures with a focus on random trees. Trees are classical data structures in computer science and can be used for the treatment of sorting and searching problems. For the asymptotic analysis discrete structures are considered as states of a combinatorial Markov chain. In this context methods from discrete potential theory are applied to construct an extension of the state space of the Markov chain such that the random variables converge almost surely. In the literature this approach is known as Doob-Martin compactification. Convergence results on the level of the discrete structures can lead to strong limit theorems for functionals of these structures, as has been shown by R. Grübel in [13].

We start with a brief overview on the Martin boundary theory of general transient irreducible Markov chains. Afterwards we verify whether the required results can be applied in the case of combinatorial Markov chains. Next we describe in detail the application of the general theory to a particular urn model, the Mendel urn. Based on the trickle-down construction introduced in [8] by S. N. Evans, R. Grübel and A. Wakolbinger we obtain the Doob-Martin compactification for random quad trees using the results concerning the Mendel urn model. Quad trees are applicable as a data structure in computer graphics. A further central subject of this thesis is the random recursive tree process and the related random Ulam-Harris trees. We use an appropriate algorithmic construction for an alternative proof of the convergence results presented in [8]. Additionally, we obtain a representation of the limit object by means of the input data. We apply the approach introduced in [13] to prove the almost sure convergence of various functionals of random recursive trees. Furthermore, we introduce a new tree functional, the horizontal path length, and analyse its asymptotic behaviour. Finally we make use of the natural correspondence between binary trees and Ulam-Harris trees to relate the corresponding Doob-Martin compactifications.

Keywords: boundary theory, random trees, asymptotics of tree functionals.

Einleitung

Das Gebiet der zufälligen diskreten Strukturen findet gegenwärtig ein beachtliches Interesse. Diese Thematik tangiert mathematische Teildisziplinen wie Theoretische Informatik, Graphentheorie und Kombinatorik, in der vorliegenden Arbeit liegt das Augenmerk auf der probabilistischen Sichtweise. Eines der bedeutendsten Beispiele für die zufälligen diskreten Strukturen sind Erdős-Rényi-Zufallsgraphen, eine Standardreferenz dazu ist das Buch [16] von S. Janson, T. Łuczak und A. Rucinski. Auch zufällige Bäume, die ein wesentlicher Gegenstand dieser Arbeit sind, ziehen erhebliche Aufmerksamkeit auf sich, das Buch [6] von M. Drmota beschreibt den aktuellen Stand der Forschung.

Zufällige diskrete Strukturen, wie beispielsweise zufällige binäre und rekursive Bäume, treten häufig als Output von deterministischen Algorithmen bei zufälligem Input auf. Erhöht man schrittweise den Inputumfang, so erhält man einen stochastischen Wachstumsprozess, der in der Regel zu der Klasse der kombinatorischen Markov-Ketten gehört, für Details verweisen wir auf den Artikel [12] von R. Grübel. Eine solche Markov-Kette $X = (X_n)_{n \in \mathbb{N}}$ ist zu einer kombinatorischen Familie \mathbb{F} im folgenden Sinne adaptiert: X_n nimmt Werte in der Menge $\mathbb{F}_n \subset \mathbb{F}$ der kombinatorischen Objekte der Größe n an. Es ist hervorzuheben, dass kombinatorische Markov-Ketten hochgradig transient sind: Jeder Zustand wird höchstens einmal angenommen.

Das Hauptanliegen dieser Arbeit ist die Analyse des asymptotischen Verhaltens zufälliger diskreter Strukturen. Eine Methode dazu liefert die Martinrandtheorie allgemeiner transienter Markov-Ketten, die auf J. L. Doob zurückgeht. Dabei wird der Zustandsraum \mathbb{F} einer kombinatorischen Markov-Kette $X = (X_n)_{n \in \mathbb{N}}$ so kompaktifiziert, dass X_n mit $n \rightarrow \infty$ fast sicher gegen eine Zufallsgröße X_∞ mit Werten im Martin-Rand $\partial\mathbb{F}$ konvergiert. Im Gegensatz zu der bei den transienten Markov-Ketten möglichen Einpunktkompaktifizierung liefert die Doob-Martin-Kompaktifizierung einen Ausgangspunkt für eine Reihe interessanter Resultate. Die Ideen, die dabei zum Einsatz kommen, beruhen auf dem klassischen Zusammenhang der Theorie der Markov-Prozesse mit der diskreten Potentialtheorie. Eine weitere zentrale Anwendung der Martinrandtheorie ist die Beschreibung der nichtnegativen harmonischen Funktionen. Aus der Sicht der Wahrscheinlichkeitstheorie sind harmonische Funktionen im Hinblick auf die Doob'sche h -Transformation beachtenswert. Auf diese Weise lassen sich vertraute Markov-Ketten, wie beispielsweise Pólya-Urnen und Irrfahrten, in Beziehung setzen. Neben der Originalarbeit [5] von J. L. Doob sei an dieser Stelle die folgende Auswahl der Publikationen auf diesem Gebiet genannt: der Übersichtsartikel [26] von S. A. Sawyer und die Bücher [29] von W. Woess und [17] von J. G. Kemeny, L. J. Snell und A. W. Knapp. In einer aktuellen Arbeit [8] bestimmen S. N. Evans, R. Grübel und A. Wakolbinger die Doob-Martin-Kompaktifizierungen für „trickle-down“-Prozesse, eine umfangreiche Klasse transienter Markov-Ketten, die beispielsweise den BST-, RRT- und den China-Restaurant-Prozess als Spezialfälle enthält. Solche Konvergenzresultate auf der Ebene der diskreten Strukturen sind nicht nur für sich genommen interessant, sondern sind als Basis für die Untersuchung deren Charakteristika geeignet: Ein neuer Ansatz zum Beweis starker Grenzwertsätze für Funktionale diskreter Strukturen wurde in [13] von R. Grübel vorgeschlagen und am Beispiel zufälliger binärer Bäume im Detail ausgeführt.

Die vorliegende Arbeit knüpft an die in [8] und [13] veröffentlichten Forschungsergebnisse an und hat die Weiterentwicklung des beschriebenen Themenkreises mit dem Fokus

auf zufälligen Bäumen zum Ziel. Der Inhalt dieser Arbeit gliedert sich in fünf Teile. Nach der Einleitung stellen wir im ersten Kapitel die Martinrandtheorie allgemeiner transienter Markov-Ketten dar und untersuchen Besonderheiten, die bei den kombinatorischen Markov-Ketten auftreten. Das zweite Kapitel widmet sich der Anwendung der Martinrandtheorie auf einen konkreten Spezialfall - die Mendel-Urne. Die Motivation für die Wahl dieses Urnenmodells als Beispiel ist, dass die betreffenden Ergebnisse für das Vorgehen im darauffolgenden Kapitel benötigt werden. Im Kapitel 3 bestimmen wir die Doob-Martin-Kompaktifizierung zu den in den Anwendungen der Informatik wichtigen zufälligen quaternären Bäumen. Diese werden als Datenstruktur in der Computergraphik eingesetzt und beruhen auf einer Verallgemeinerung der Idee binärer Suchbäume für den Fall zweidimensionaler Daten. Die „trickle-down“-Konstruktion macht es möglich, den Quadtree-Prozess als Verschachtelung unabhängiger Mendel-Urnen darzustellen. Im Kapitel 4 untersuchen wir den RRT-Prozess und die dazu verwandte Markov-Kette mit Werten in der kombinatorischen Familie der Ulam-Harris-Bäume. Wir verwenden eine geeignete algorithmische Konstruktion für den alternativen Beweis der aus [8] bereits bekannten Konvergenzresultate. Ein Vorteil dieses Zugangs ist eine Darstellung des Limesobjektes mit Hilfe der Inputdaten. Ferner erhalten wir mit Hilfe des Ansatzes aus [13] starke Grenzwertsätze für die Baumfunktionale interne Pfadlänge, horizontale Pfadlänge und Wiener-Index. Bei der horizontalen Pfadlänge handelt es sich um ein neues Funktional, welches mit der Laufzeit des betreffenden Algorithmus zusammenhängt. Kapitel 5 befasst sich mit der Rolle der natürlichen Korrespondenz zwischen den binären Bäumen und den Ulam-Harris-Bäumen. Die Resultate dieses Kapitels stellen einen Zusammenhang zwischen den entsprechenden Doob-Martin-Kompaktifizierungen her. Außerdem erhalten wir eine Aussage über die Verteilungen der internen und der horizontalen Pfadlänge zufälliger Ulam-Harris-Bäume.

Inhaltsverzeichnis

1	Martinrandtheorie	11
1.1	Markov-Ketten und Transienz	11
1.2	Doob-Martin-Kompaktifizierung	12
1.3	Hauptresultate	14
1.4	Besonderheiten bei kombinatorischen Markov-Ketten	16
1.4.1	Konvergenz der Markov-Kette	17
1.4.2	h -transformierte Prozesse	19
1.4.3	Integraldarstellung nichtnegativer harmonischer Funktionen	20
2	Die Mendel-Urne	23
3	Zufällige quaternäre Bäume	33
3.1	Grundbegriffe und Bezeichnungen	33
3.2	Die „trickle-down“-Konstruktion	37
3.3	Doob-Martin Kompaktifizierung für den Quadtree-Prozess	39
3.4	Der Vier-Farben-Baum als h -Transformierte des Quadtree-Prozesses	45
3.5	Ausblick: Pfadlängen in quaternären Bäumen	47
4	Zufällige rekursive Bäume	51
4.1	Grundbegriffe und Bezeichnungen	51
4.2	China-Restaurant-Prozess	55
4.3	Martin-Rand der Harris-Kette	58
4.4	Alternativer Zugang zur Konvergenz	59
4.5	Baumfunktionale	64
4.5.1	Interne Pfadlänge	65
4.5.2	Wiener-Index	71
4.5.3	Horizontale Pfadlänge und Konstruktionsaufwand	76
5	Die Rolle der natürlichen Korrespondenz	81
5.1	Natürliche Korrespondenz und die Doob-Martin-Kompaktifizierungen	84
5.2	Natürliche Korrespondenz und der algorithmische Hintergrund	88
5.3	Natürliche Korrespondenz und Baumfunktionale	89
	Anhang	91
A.1	Wörter und Alphabete	91
A.2	Ergänzungen aus der Topologie	91
A.3	Schwache Konvergenz	92
	Literatur	95

1 Martinrandtheorie

In diesem Abschnitt geben wir eine Kurzübersicht über die Martinrandtheorie transienter Markov-Ketten, die auf eine fundamentale Arbeit [5] von J. L. Doob zurückgeht. Dabei orientieren wir uns an dem Buch [29] von W. Woess und an dem Artikel [26] von S. A. Sawyer. Wir beginnen mit Grundbegriffen aus der Theorie der Markov-Ketten. Anschließend befassen wir uns mit der abstrakten Konstruktion der Doob-Martin-Kompaktifizierung für allgemeine irreduzible transiente Markov-Ketten. Dann folgt die Zusammenstellung der für unsere Zwecke benötigten Hauptresultate der Martinrandtheorie. Abschließend betrachten wir die Besonderheiten, die bei den kombinatorischen Markov-Ketten auftauchen. Insbesondere werden wir feststellen, dass sich die Irreduzibilitätsbedingung abschwächen lässt. Die Hauptquelle zu den kombinatorischen Markov-Ketten ist der Artikel [12] von R. Grübel.

1.1 Markov-Ketten und Transienz

Sei E eine endliche oder abzählbar unendliche nichtleere Menge, ν eine Verteilung auf $(E, \mathcal{P}(E))$ und $p = (p(x, y))_{x, y \in E}$ eine stochastische Matrix, das heißt, es gilt

$$p(x, y) \geq 0 \text{ für alle } x, y \in E \text{ und } \sum_{y \in E} p(x, y) = 1 \text{ für jedes } x \in E.$$

Bekanntlich bezeichnet man einen stochastischen Prozess $(X_n)_{n \in \mathbb{N}_0}$ auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ mit Werten in E als Markov-Kette mit Zustandsraum E , Startverteilung ν und Übergangsmatrix p , falls ν die Verteilung von X_0 ist und

$$\mathbb{P}(X_{n+1} = x_{n+1} | X_0 = x_0, \dots, X_n = x_n) = p(x_n, x_{n+1}) \quad (1)$$

für alle $n \geq 0$ und $x_0, \dots, x_{n+1} \in E$ mit $\mathbb{P}(x_0, \dots, x_n) > 0$ gilt. Im Falle $\nu = \delta_x$ für ein $x \in E$, das heißt bei einem sicheren Start in x , schreiben wir \mathbb{P}_x statt \mathbb{P} für das Hintergrundwahrscheinlichkeitsmaß. Markov-Ketten können durch den zugehörigen Übergangsgraphen dargestellt werden. Dies ist ein gewichteter Digraph, bei dem die Knoten den Zuständen entsprechen. Zwei Knoten $x, y \in E$ werden durch eine gerichtete Kante verbunden, falls $p(x, y) > 0$ ist. Die Kanten (x, y) werden mit den Übergangswahrscheinlichkeiten $p(x, y)$ gewichtet.

Zu jeder vorgegebenen Startverteilung ν und einer beliebigen stochastischen Matrix p existiert eine zugehörige Markov-Kette im Sinne der obigen Definition. Es existiert also ein Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ und eine Folge $(X_n)_{n \in \mathbb{N}_0}$ von E -wertigen Zufallsvariablen darauf, sodass $\mathbb{P}^{X_0} = \nu$ und (1) gilt. Den Beweis dazu liefert die folgende kanonische Konstruktion, eine ausführliche Beschreibung findet sich in [29, S.6-10]. Es sei $\Omega = E^{\mathbb{N}_0}$ die Menge aller Folgen $\omega = (x_0, x_1, x_2, \dots)$ aus E und \mathcal{F} die von dem System der Zylindermengen erzeugte σ -Algebra. Für $a_0, a_1, \dots, a_k \in E$ ist die Zylindermenge $C_{(a_0, \dots, a_k)}$ die Menge aller Folgen, die mit a_0, \dots, a_k beginnen:

$$C_{(a_0, \dots, a_k)} := \{\omega = (x_n)_{n \in \mathbb{N}_0} \in \Omega : x_i = a_i, i = 0, \dots, k\}.$$

Dabei wird die Menge Ω auch als eine Zylindermenge betrachtet. Das Wahrscheinlichkeitsmaß \mathbb{P} wird festgelegt durch $\mathbb{P}(\Omega) = 1$ und

$$\mathbb{P}(C_{(a_0, \dots, a_k)}) = \nu(\{a_0\})p(a_0, a_1)p(a_1, a_2) \dots p(a_{k-1}, a_k).$$

Schließlich bilden die Projektionen auf die n -te Koordinate

$$X_n : \Omega \rightarrow E, \quad X_n(\omega) = x_n,$$

$n \in \mathbb{N}_0$, die gesuchte Markov-Kette.

Bemerkung 1.1. Sofern nicht anders angegeben, betrachten wir im Folgenden nur solche kanonischen Markov-Ketten.

Sei p^n die n -te Matrixpotenz von p . Die Einträge $p^n(x, y)$, $x, y \in E$ sind die n -Schritt-Übergangswahrscheinlichkeiten, genauer ist

$$\mathbb{P}_x(X_n = y) = p^n(x, y) \text{ für alle } x, y \in E.$$

Eine Markov-Kette mit der Eigenschaft

$$\forall x, y \in E \quad \exists n \geq 0 \quad \text{mit } p^n(x, y) > 0$$

nennt man irreduzibel. Der Übergangsgraph einer irreduziblen Markov-Kette ist stark zusammenhängend. Das bedeutet, dass jeder Zustand von jedem anderen Zustand mit positiver Wahrscheinlichkeit in endlich vielen Schritten erreicht werden kann. Ein Zustand $x \in E$ heißt rekurrent, falls

$$\mathbb{P}_x(\{\omega \in \Omega : \exists n > 0 \text{ mit } X_n(\omega) = x\}) = 1$$

gilt, das heißt, beim Start in x kommt die Markov-Kette mit Wahrscheinlichkeit 1 zum Startzustand x zurück. Anderenfalls heißt dieser Zustand transient. Sind alle Zustände transient, so spricht man von einer transienten Markov-Kette. Bei irreduziblen Markov-Ketten sind entweder alle Zustände transient oder es sind alle Zustände rekurrent.

1.2 Doob-Martin-Kompaktifizierung

Es sei nun $(X_n)_{n \in \mathbb{N}_0}$ eine irreduzible transiente Markov-Kette mit abzählbarem Zustandsraum E und Übergangsmatrix p . Wir erinnern an die Konstruktion der Doob-Martin-Kompaktifizierung, vergleiche Abschnitt 4 von [26] und Abschnitt 7B von [29]. Der Green'sche Kern ist definiert als

$$G(x, y) = \sum_{n=0}^{\infty} p^n(x, y), \quad x, y \in E,$$

dabei lässt sich $G(x, y)$ als die erwartete Anzahl der Besuche von $(X_n)_{n \in \mathbb{N}_0}$ in y bei Start in x interpretieren. Der Martin-Kern zum Referenzzustand x_0 ist gegeben durch

$$K(x, y) := \frac{G(x, y)}{G(x_0, y)}, \quad x, y \in E. \quad (2)$$

Dieser ist wohldefiniert, denn die Irreduzibilität und die Transienz liefern

$$0 < G(x, y) < \infty \quad \text{für alle } x, y \in E.$$

Man kann zeigen, dass die Funktion $K(x, \cdot)$ für jedes $x \in E$ durch eine Konstante C_x beschränkt ist, zum Beweis siehe Lemma 4.1. in [26]. Auf der Menge E wird durch

$$d(x, y) := \sum_{z \in E} w(z) \frac{|K(z, x) - K(z, y)| + |\delta_{zx} - \delta_{zy}|}{C_z + 1} \quad (3)$$

eine Metrik d eingeführt, wobei $w(z)$ Gewichte mit der Eigenschaft $\sum_{z \in E} w(z) < \infty$ seien. Der metrische Raum (E, d) lässt sich nun durch die Betrachtung der Cauchy-Folgen vervollständigen¹. Die Folgen $(y_n)_{n \in \mathbb{N}}$ sind Cauchy-Folgen in (E, d) genau dann, wenn $K(x, y_n)$ in \mathbb{R} mit $n \rightarrow \infty$ für jedes x konvergiert. In der vorliegenden Situation bedeutet dies, dass genau eine der folgenden Bedingungen erfüllt ist:

- $y_n = y$ für ein $y \in E$ und alle bis auf endlich viele n ,
- $y_n \rightarrow \infty$ und $(K(x, y_n))_{n \in \mathbb{N}}$ ist für jedes $x \in E$ eine in \mathbb{R} konvergente Folge.

Dabei bedeutet $y_n \rightarrow \infty$, dass es für jede endliche Teilmenge $A \subset E$ nur endlich viele n mit $y_n \in A$ gibt. Auf der Menge \mathcal{C} der Cauchy-Folgen in E wird durch

$$(x_n)_{n \in \mathbb{N}} \sim (y_n)_{n \in \mathbb{N}} \Leftrightarrow \lim_{n \rightarrow \infty} d(x_n, y_n) = 0$$

eine Äquivalenzrelation eingeführt. Für $(x_n)_{n \in \mathbb{N}} \in \mathcal{C}$ bezeichne $[(x_n)_{n \in \mathbb{N}}]$ die Äquivalenzklasse von $(x_n)_{n \in \mathbb{N}}$ bezüglich \sim . Auf dem Quotientenraum $\hat{E} := \mathcal{C}/\sim$ wird dann durch

$$\hat{d}([(x_n)_{n \in \mathbb{N}}], [(y_n)_{n \in \mathbb{N}}]) = \lim_{n \rightarrow \infty} d(x_n, y_n)$$

eine Metrik definiert. Schließlich lässt sich die Menge E als dichte Teilmenge in den Raum \hat{E} der Äquivalenzklassen durch die Abbildung $\varphi : E \rightarrow \hat{E}$, $\varphi(x) = [(x)_{n \in \mathbb{N}}]_{\sim}$, die jedem $x \in E$ die Äquivalenzklasse zur konstanten Folge $(x)_{n \in \mathbb{N}}$ zuordnet, einbetten. Da die Familie $\{K(x, \cdot) : x \in E\}$ eine abzählbare Familie von beschränkten Funktionen ist, kann man mit einem Diagonalisierungsargument die Kompaktheit von \hat{E} zeigen. Zusammenfassend stellen wir fest: Mit Hilfe der Martin-Kerne wurde eine Metrik d auf E eingeführt, so dass die Vervollständigung (\hat{E}, \hat{d}) des metrischen Raumes (E, d) ein kompakter Raum ist. Der Raum (\hat{E}, \hat{d}) wird als Doob-Martin-Kompaktifizierung bezeichnet und $\partial E = \hat{E} \setminus \varphi(E)$ als Martin-Rand.

Bemerkung 1.2. Für weitere Betrachtungen sind folgende Aussagen hilfreich.

- Die Doob-Martin-Kompaktifizierung ist eindeutig bis auf Homöomorphie.
- E liegt dicht in \hat{E} und ist diskret in der Topologie von \hat{E} .
- Die Topologie von \hat{E} wird von einer Metrik erzeugt, folglich ist \hat{E} zusätzlich separabel und Hausdorff'sch.
- Ausgestattet mit der von den offenen Mengen erzeugten σ -Algebra, der Borel'schen σ -Algebra $\mathcal{B}(\hat{E})$, wird der topologische Raum \hat{E} in kanonischer Weise zum Messraum. Insbesondere verstehen wir im Folgenden unter einem Maß auf der Doob-Martin-Kompaktifizierung \hat{E} ein auf $\mathcal{B}(\hat{E})$ definiertes Maß.

¹Zur Vervollständigung metrischer Räume verweisen wir auf Kapitel 4 aus [15].

- Für jedes x lässt sich die Funktion $K(x, \cdot)$ auf \widehat{E} eindeutig stetig fortsetzen. Wir bezeichnen die Fortsetzung als den erweiterten Martin-Kern und schreiben $K(x, \alpha)$ für die Werte des erweiterten Martin-Kerns in $\alpha \in \partial E$.

Bemerkung 1.3. Die Terme δ_{zx} , δ_{zy} in (3) stellen sicher, dass E offen und ∂E abgeschlossen in der Topologie von \widehat{E} ist, siehe Abschnitt 4 von [26]: Es kann vorkommen, dass

$$\lim_{n \rightarrow \infty} K(x, y_n) = K(x, y) \text{ für alle } x \in E$$

für eine Folge $(y_n)_{n \in \mathbb{N}}$ aus E mit $y_n \rightarrow \infty$ und ein $y \in E$ gilt. Ohne dieser Terme in (3) wäre also ein Element aus E Grenzwert einer Folge, die jede endliche Teilmenge von E irgendwann verlässt. Ist jedoch die Bedingung $\#\{y \in E : p(x, y) > 0\} < \infty$ für alle $x \in E$ an die Übergangsmatrix erfüllt, so können δ_{zx} und δ_{zy} in der Definition der Metrik weggelassen werden; zu Details verweisen wir auf [29, S. 188-189].

1.3 Hauptresultate

Die Bedeutung der Martinrandtheorie liegt, soweit es die vorliegende Arbeit betrifft, in den folgenden zentralen Resultaten:

- die Integraldarstellung der nichtnegativen harmonischen Funktionen und
- die fast-sichere Konvergenz der Markov-Kette gegen eine Zufallsgröße mit Werten auf dem Martin-Rand.

Für eine Funktion $f : E \rightarrow \mathbb{R}$ sei

$$pf(x) := \sum_{y \in E} p(x, y)f(y),$$

falls die Summe konvergiert. Eine Funktion $h : E \rightarrow \mathbb{R}$ heißt harmonisch, falls $h(x) = ph(x)$ für jedes $x \in E$ erfüllt ist.

Bemerkung 1.4. Aus der Irreduzibilität folgt für jede nichtnegative harmonische Funktion h entweder $h \equiv 0$ oder $h(x) > 0$ für alle $x \in E$, siehe Lemma 6.17 in [29].

Gilt $\#\{y \in E : p(x, y) > 0\} < \infty$ für alle $x \in E$, so ist der zugehörige Übergangsgraph lokal endlich. In diesem Fall ist für jedes $\alpha \in \partial E$ die Funktion $x \mapsto K(x, \alpha)$ positiv und harmonisch mit Wert 1 im Referenzzustand x_0 , siehe Lemma 7.18 in [29]. Jede nichtnegative harmonische Funktion besitzt eine Integraldarstellung:

Satz 1.5 ([26, Satz 4.1]). *Angenommen, es gilt*

1. $G(x, y) < \infty$ für alle $x, y \in E$ und
2. $G(x_0, y) > 0$ für alle $y \in E$ und ein $x_0 \in E$.

Dann existiert für jede nichtnegative harmonische Funktion h ein Maß μ_h auf der Doob-Martin-Kompaktifizierung \widehat{E} , das auf den Martin-Rand ∂E konzentriert ist, sodass

$$h(x) = \int_{\partial E} K(x, \alpha) \mu_h(d\alpha), \text{ für alle } x \in E,$$

gilt. Hierbei bezieht sich der Martin-Kern K auf den Zustand x_0 aus der Annahme 2.

Sei $h \neq 0$ eine nichtnegative harmonische Funktion. Dann definiert

$$p^{(h)}(x, y) := \frac{1}{h(x)} p(x, y) h(y), \quad x, y \in E$$

eine neue Übergangsmatrix $p^{(h)} = (p^{(h)}(x, y))_{x, y \in E}$. Eine Markov-Kette mit dieser Übergangsmatrix bezeichnet man als eine h -transformierte Markov-Kette, die Idee geht zurück auf Doob. Man beachte, dass die Zufallsgrößen X_n unverändert bleiben; es ändert sich nur das Hintergrundwahrscheinlichkeitsmaß auf dem Trajektorienraum $\Omega = E^{\mathbb{N}_0}$. Die h -transformierte Markov-Kette hat dieselbe Doob-Martin-Kompaktifizierung wie die Ausgangskette, denn der zugehörige Martin-Kern ist

$$K^{(h)}(x, y) = \frac{h(x_0)}{h(x)} K(x, y), \quad x, y \in E.$$

Aus dieser Darstellung geht hervor, dass eine Folge $(y_n)_{n \in \mathbb{N}}$ genau dann eine Cauchy-Folge bezüglich d ist, wenn sie auch eine Cauchy-Folge bezüglich $d^{(h)}$ ist. Dabei bezeichnet $d^{(h)}$ die wie in (3) zu $K^{(h)}$ definierte Metrik.

Bemerkung 1.6. Wir bezeichnen mit \mathcal{H}_1^+ die Menge aller nichtnegativen harmonischen Funktionen h mit $h(x_0) = 1$. Dies ist offensichtlich eine konvexe Menge. Wir erinnern daran, dass ein Element einer konvexen Menge extremal heißt, wenn es nur auf triviale Weise als Konvexkombination dargestellt werden kann. Es ist bekannt, dass alle extremalen $h \in \mathcal{H}_1^+$ von der Form $K(\cdot, \alpha)$ mit einem $\alpha \in \partial E$ sind, siehe Satz 7.50 in [29]. Die Menge

$$\partial_{\min} E := \{\alpha \in \partial E : K(\cdot, \alpha) \text{ ist extremal}\}$$

bezeichnet man als minimalen Martin-Rand. Dieser lässt sich folgendermaßen charakterisieren [29, Korollar 7.51]: Für jedes $\alpha \in \partial E$ gilt $\alpha \in \partial_{\min} E$ genau dann, wenn die Austrittsverteilung der h -transformierten Markov-Kette zur harmonischen Funktion $h = K(\cdot, \alpha)$ das Einpunktmaß in α ist. Außerdem ist $\partial_{\min} E$ nach Lemma 7.52 in [29] eine Borel'sche Teilmenge von \widehat{E} . Fordert man in Satz 1.5 zusätzlich, dass μ_h auf $\partial_{\min} E$ konzentriert ist, so ist die Integraldarstellung der nichtnegativen harmonischen Funktionen eindeutig, siehe Satz 7.53 in [29].

Für die Formulierung des nächsten Satzes benötigen wir die folgenden Hilfsüberlegungen zur Messbarkeit:

Bemerkung 1.7. Als ein metrischer Raum ist \widehat{E} , wie bereits erwähnt, versehen mit der Borel'schen σ -Algebra $\mathcal{B}(\widehat{E})$. Unter einer Zufallsgröße mit Werten in \widehat{E} verstehen wir im Folgenden eine Abbildung von dem Trajektorienraum $\Omega = E^{\mathbb{N}_0}$ nach \widehat{E} , die bezüglich der von den Zylindermengen erzeugten σ -Algebra \mathcal{F} und $\mathcal{B}(\widehat{E})$ messbar ist. Ausgestattet mit der Borel'schen σ -Algebra

$$\mathcal{B}(\partial E) = \left\{ B \cap \partial E : B \in \mathcal{B}(\widehat{E}) \right\}$$

wird der Martin-Rand selbst zum Messraum. Gilt also $Y \in \partial E$ fast sicher für eine \widehat{E} -wertige Zufallsgröße Y , so existiert eine ∂E -wertige Zufallsgröße Z mit $Y = Z$ fast sicher. Andererseits kann jede ∂E -wertige Zufallsgröße trivialerweise als eine Zufallsgröße mit Werten in \widehat{E} betrachtet werden.

Das zweite Hauptresultat beschreibt das Verhalten von $(X_n)_{n \in \mathbb{N}_0}$ mit $n \rightarrow \infty$:

Satz 1.8 ([29, Satz 7.19]). *Unter den Voraussetzungen dieses Abschnittes existiert eine Zufallsgröße X_∞ mit Werten in ∂E , sodass für jedes $x \in E$*

$$\lim_{n \rightarrow \infty} X_n = X_\infty \quad \mathbb{P}_x\text{-fast-sicher}$$

in der Topologie von \widehat{E} gilt.

Zu jedem Startzustand lässt sich die Verteilung von X_∞ folgendermaßen charakterisieren:

Satz 1.9 ([26, Satz 5.1]). *Für jedes $x \in E$ gilt*

$$\mathbb{P}_x(X_\infty \in A) = \int_A K(x, \alpha) \mu_1(d\alpha),$$

für Borel-Mengen $A \subset \partial E$, wobei μ_1 das eindeutige, auf $\partial_{\min} E$ konzentrierte Maß zur harmonischen Funktion $h(x) \equiv 1$ bezeichnet:

$$\int_{\partial_{\min} E} K(x, \alpha) \mu_1(d\alpha) = 1$$

für alle $x \in E$.

Da $K(x_0, \alpha) = 1$ für alle $\alpha \in \partial E$ gilt, ist μ_1 die Austrittsverteilung $\mathcal{L}(X_\infty)$ von $(X_n)_{n \in \mathbb{N}_0}$ beim Start im Referenzzustand x_0 . Ferner bedeutet dieser Satz, dass $K(x, \cdot)$ die Radon-Nikodym-Dichte der Austrittsverteilung zum Start in einem $x \in E$ bezüglich μ_1 ist.

1.4 Besonderheiten bei kombinatorischen Markov-Ketten

Unter einer kombinatorischen Familie versteht man eine Menge \mathbb{F} mit einer Größenfunktion $\phi : \mathbb{F} \rightarrow \mathbb{N}$ und folgenden Eigenschaften: Es gelte $\#\mathbb{F}_n < \infty$ für alle $n \in \mathbb{N}$, wobei $\mathbb{F}_n := \{x \in \mathbb{F} : \phi(x) = n\}$ die Menge der kombinatorischen Objekte der Größe n sei. Außerdem bestehe \mathbb{F}_1 aus nur einem Objekt o . Ferner sei \mathbb{F} mit einer partiellen Ordnung „ \leq “ versehen, sodass $\phi(x) \leq \phi(y)$ für alle $x, y \in \mathbb{F}$ mit $x \leq y$ gelte. Zusätzlich setzt man voraus, dass zu jedem $y \in \mathbb{F}_{n+1}$ ein $x \in \mathbb{F}_n$ mit $x \leq y$ existiere. Insbesondere ist dann die Menge \mathbb{F} abzählbar. Beispiele für kombinatorische Familien sind Zeichenketten, Zahlenpartitionen, Permutationen, Graphen und Bäume.

Eine kombinatorische Markov-Kette ist eine Markov-Kette $(X_n)_{n \in \mathbb{N}}$ mit Zustandsraum \mathbb{F} , Start in o und Übergangsmatrix p , für die $\mathbb{P}_o(X_n \in \mathbb{F}_n) = 1$ für alle $n \in \mathbb{N}$ und

$$p(x, y) > 0 \Rightarrow x \leq y$$

für alle $n \in \mathbb{N}$, $x \in \mathbb{F}_n$, $y \in \mathbb{F}_{n+1}$ gilt². Eine solche Markov-Kette ist in einem starken Sinne transient: Jeder Zustand wird höchstens einmal angenommen. Zusätzlich setzen wir die folgende schwache Irreduzibilitätsannahme voraus:

$$\mathbb{P}_o(X_m = x \text{ für ein } m \in \mathbb{N}) > 0 \text{ für alle } x \in \mathbb{F},$$

²Bei kombinatorischen Markov-Ketten wählen wir \mathbb{N} statt \mathbb{N}_0 als Zeitbereich.

das heißt, jeder Zustand kann von o aus mit positiver Wahrscheinlichkeit erreicht werden. Diese Annahme stellt sicher, dass $G(o, x) > 0$ für alle $x \in \mathbb{F}$ ist. Ferner gilt nach wie vor $G(x, y) < \infty$ für alle $x, y \in \mathbb{F}$ wegen der Transienz. Somit ist die Wahl des Startzustandes o als Referenzzustand naheliegend. Anderenfalls, bei der Wahl von $x_0 \neq o$, verkleinert man den Zustandsraum \mathbb{F} zu $\mathbb{F}' := \{x \in \mathbb{F} : \exists m \geq 0 \text{ mit } p^m(x_0, x) > 0\}$. Der Martin-Kern zum Referenzzustand o , vergleiche Gleichung (2), vereinfacht sich zu:

$$K(x, y) = \begin{cases} \frac{\mathbb{P}_o(X_n=y|X_m=x)}{\mathbb{P}_o(X_n=y)} & \text{falls } x \in \mathbb{F}_m, y \in \mathbb{F}_n \text{ und } m \leq n, \\ 0 & \text{sonst.} \end{cases}$$

Dabei wurde verwendet, dass

$$\mathbb{P}_o(X_{k+l} = y | X_k = x) = p^l(x, y)$$

für $k \in \mathbb{N}$ mit $\mathbb{P}_o(X_k = x) > 0$ gilt. Aus der Definition des Martin-Kerns folgt direkt, dass zu jedem $x \in \mathbb{F}$ eine Konstante C_x existiert, sodass $K(x, y) \leq C_x$ für alle $y \in \mathbb{F}$ gilt:

$$K(x, y) \leq \frac{1}{\mathbb{P}_o(X_m = x)} =: C_x \text{ für alle } y \in \mathbb{F}, \quad (4)$$

mit $m = \phi(x)$. Für die abstrakte Konstruktion der Doob-Martin-Kompaktifizierung $\widehat{\mathbb{F}}$ reicht also die schwache Irreduzibilität aus. Da die Bedingung

$$\#\{y \in \mathbb{F} : p(x, y) > 0\} < \infty \text{ für alle } x \in \mathbb{F}$$

im Falle kombinatorischer Markov-Ketten stets erfüllt ist, kann dabei nach Bemerkung 1.3 die zur Konstruktion benötigte Metrik wie folgt modifiziert werden:

$$d(x, y) := \sum_{z \in E} w(z) \frac{|K(z, x) - K(z, y)|}{C_z}. \quad (5)$$

Wir überprüfen nun die Übertragbarkeit der Hauptresultate der Martinrandtheorie allgemeiner irreduzibler transienter Markov-Ketten auf die Situation kombinatorischer Markov-Ketten.

1.4.1 Konvergenz der Markov-Kette

Für den Satz 1.8 über die fast sichere Konvergenz der Markov-Kette gegen eine Zufallsgröße mit Werten auf dem Martin-Rand skizzieren wir einen Beweis aus einem nicht veröffentlichten Skript von R. Grübel. Man zeigt zunächst die fast sichere Konvergenz bezüglich \mathbb{P}_o . Wie bereits erwähnt, sind die Folgen $(y_n)_{n \in \mathbb{N}}$ Cauchy-Folgen in der Doob-Martin-Topologie genau dann, wenn $K(x, y_n)$ in \mathbb{R} mit $n \rightarrow \infty$ für jedes $x \in \mathbb{F}$ konvergiert. Nehmen wir nun an, dass $K(x, X_n)$ für jedes $x \in \mathbb{F}$ bezüglich \mathbb{P}_o fast sicher konvergiert. Dann existiert eine von x abhängige \mathbb{P}_o -Nullmenge N_x , sodass $K(x, X_n(\omega))$ mit $n \rightarrow \infty$ für alle $\omega \notin N_x$ konvergiert. Die Menge $N = \bigcup_{x \in \mathbb{F}} N_x$ ist als Vereinigung abzählbar vieler Nullmengen wieder eine Nullmenge. Außerhalb von N wäre also $(X_n(\omega))_{n \in \mathbb{N}}$ eine Cauchy-Folge in der Doob-Martin-Topologie. Folglich reicht es zu zeigen, dass die Zufallsvariablen $K(x, X_n)$ für jedes $x \in \mathbb{F}$ mit $n \rightarrow \infty$ \mathbb{P}_o -fast sicher konvergieren. Dazu zeigt man zunächst die folgende Hilfsaussage: Für $x \in \mathbb{F}$, $n \in \mathbb{N}$ und $m \in \{2, \dots, n\}$ gilt

$$E[K(x, X_{m-1}) | X_m, \dots, X_n] \leq K(x, X_m).$$

Dann ist für ein festes $n \in \mathbb{N}$ der Prozess $(Y_k, \mathcal{F}_k)_{k=1, \dots, n}$ mit

$$Y_k := K(x, X_{n+1-k}), \quad \mathcal{F}_k := \sigma(X_n, \dots, X_{n+1-k}),$$

ein Supermartingal. Sei $U_n[a, b]$ als Anzahl der aufsteigenden Überquerungen des Intervalls $[a, b]$ durch Y_1, \dots, Y_n . Nach der Doob'schen Aufkreuzungsungleichung, siehe [28, S.108], gilt³

$$E U_n[a, b] \leq \frac{1}{b-a} E(Y_n - a)^- = \frac{1}{b-a} E(K(x, o) - a)^-. \quad (6)$$

Nach Konstruktion ist $U_n[a, b]$ gleich der Anzahl der absteigenden Überquerungen $D_n[a, b]$ des Intervalls $[a, b]$ durch $K(x, X_1), \dots, K(x, X_n)$. Da die Oberschranke in (6) nicht von n abhängt, gilt $\mathbb{P}_o(D_\infty = \infty) = 0$ für $D_\infty = \lim_{k \rightarrow \infty} D_k$. Ferner ist der Prozess $(K(x, X_k))_{n \in \mathbb{N}}$ wegen (4) L^1 -beschränkt. Mit Argumenten wie in Beweis des Vorwärtskonvergenzsatzes von Doob, siehe [28, S.109], folgt daraus die \mathbb{P}_o -fast sichere Konvergenz von $(K(x, X_k))_{n \in \mathbb{N}}$. Somit ist dieses Hauptresultat für \mathbb{P}_o gezeigt. Die schwache Irreduzibilität liefert die fast sichere Konvergenz bezüglich \mathbb{P}_x für jedes $x \in \mathbb{F}$.

Für den Satz 1.9 über die Austrittsverteilungen verwenden wir die Beweismethode aus [29, S. 201], wobei sich bei kombinatorischen Markov-Ketten einige Vereinfachungen ergeben. Sei μ_x die Austrittsverteilung zum Start in $x \in \mathbb{F}$, also

$$\mu_x(A) = \mathbb{P}_x(X_\infty \in A)$$

für $A \in \mathcal{B}(\widehat{\mathbb{F}})$. An dieser Stelle beachte man, dass in der vorliegenden Situation sofort $\mathbb{P}_o(X_\infty \in \mathbb{F}) = 0$ folgt. Zu zeigen ist:

$$\mu_x(A) = \int_A K(x, \xi) \mu_o(d\xi)$$

für alle Borel-Mengen $A \subset \widehat{\mathbb{F}}$. Für $k \in \mathbb{N}$ mit $\mathbb{P}_o(X_k = x) > 0$ gilt

$$\mathbb{P}_x(X_n = y) = \mathbb{P}_o(X_{k+n} = y | X_k = x)$$

für jedes $y \in \mathbb{F}$. Sei $x \in \mathbb{F}$ beliebig und $k = \phi(x)$. Sei $f : \widehat{\mathbb{F}} \rightarrow \mathbb{R}$ eine stetige Funktion. Dann gilt⁴

$$\begin{aligned} E_x(f(X_n)) &= \sum_{y \in \mathbb{F}_{k+n}} f(y) \mathbb{P}_x(X_n = y) \\ &= \sum_{y \in \mathbb{F}_{k+n}} f(y) \mathbb{P}_o(X_{k+n} = y | X_k = x) \\ &= \sum_{y \in \mathbb{F}_{k+n}} f(y) \frac{\mathbb{P}_o(X_{k+n} = y | X_k = x)}{\mathbb{P}_o(X_{k+n} = y)} \mathbb{P}_o(X_{k+n} = y) \\ &= \sum_{y \in \mathbb{F}_{k+n}} f(y) K(x, y) \mathbb{P}_o(X_{k+n} = y) = E_o(f(X_n) K(x, X_n)) \end{aligned}$$

³Für eine Funktion $f : \Omega \rightarrow \mathbb{R}$ ist das Negativteil von f definiert als $f^- := -\min\{f, 0\}$.

⁴Für eine \mathbb{P}_x -integrierbare Zufallsvariable Z bezeichnet $E_x Z = \int Z d\mathbb{P}_x$ den Erwartungswert von Z bei der Startverteilung δ_x , $x \in \mathbb{F}$.

Da f und $K(x, \cdot)$ stetig auf der kompakten Menge $\widehat{\mathbb{F}}$ und damit automatisch beschränkt sind, folgt mit dem Satz von der majorisierten Konvergenz

$$E_x(f(X_\infty)) = E_o(f(X_\infty)K(x, X_\infty)).$$

Man hat also

$$\int_{\widehat{\mathbb{F}}} f(\xi) \mu_x(d\xi) = \int_{\widehat{\mathbb{F}}} f(\xi) K(x, \xi) \mu_o(d\xi).$$

Indikatorfunktionen abgeschlossener Mengen können punktweise durch stetige Funktionen approximiert werden [2, S. 8-9], also folgt für abgeschlossene Mengen A

$$\mu_x(A) = \int_A K(x, \xi) \mu_o(d\xi).$$

Da auch die abgeschlossenen Mengen die Borel'sche σ -Algebra erzeugen, schließt das den Beweis des Satzes 1.9 für die Situation kombinatorischer Markov-Ketten ab.

1.4.2 h -transformierte Prozesse

Es ist hervorzuheben, dass für schwach irreduzible Markov-Ketten die Eigenschaft nicht-negativer harmonischer Funktionen aus der Bemerkung 1.4 nicht gilt. Bei der Betrachtung der h -transformierten Markov-Ketten geht man deshalb zu einem verkleinerten Zustandsraum über:

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

Wir folgen nun dem Ansatz aus [17, S. 345-346].

Lemma 1.10. *Sei h eine nichtnegative harmonische Funktion. Für x und y mit $h(x) = 0$ und $h(y) > 0$ gilt $K(x, y) = 0$.*

Beweis. Da h eine harmonische Funktion ist, gilt

$$0 = h(x) = p^n h(x) = \sum_{z \in \mathbb{F}} p^n(x, z) h(z)$$

für jedes n . Wegen $h \geq 0$ und $h(y) > 0$ folgt $p^n(x, y) = 0$ für jedes n und somit $K(x, y) = 0$. \square

Lemma 1.11. *Wählt man bei der Definition von $d^{(h)}$ die gleichen Gewichte wie bei d in (5), so ist die Identität von $(\mathbb{F}^{(h)}, d^{(h)})$ nach (\mathbb{F}, d) eine Isometrie, vergleiche Lemma 10-25 aus [17].*

Beweis. Die Anwendung des Lemmas 1.10 bei der vierten Umformung liefert für $x, y \in \mathbb{F}^{(h)}$:

$$\begin{aligned}
d^{(h)}(x, y) &= \sum_{z \in \mathbb{F}^{(h)}} w(z) \frac{|K^{(h)}(z, x) - K^{(h)}(z, y)|}{C_z^{(h)}} \\
&= \sum_{z \in \mathbb{F}^{(h)}} w(z) \frac{\frac{h(x_0)}{h(x)} |K(z, x) - K(z, y)|}{\frac{h(x_0)}{h(x)} C_z} \\
&= \sum_{z \in \mathbb{F}^{(h)}} w(z) \frac{|K(z, x) - K(z, y)|}{C_z} \\
&= \sum_{z \in \mathbb{F}} w(z) \frac{|K(z, x) - K(z, y)|}{C_z} = d(x, y) \quad \square
\end{aligned}$$

Da Isometrien gleichmäßig stetig sind, existiert eine Fortsetzung der Identität aus dem obigen Lemma zu einer stetigen Abbildung von $\widehat{\mathbb{F}^{(h)}}$ nach $\widehat{\mathbb{F}}$, die wieder eine Isometrie ist.

Korollar 1.12. *Die Doob-Martin-Kompaktifizierung $\widehat{\mathbb{F}^{(h)}}$ der h -transformierten Markov-Kette kann als kompakte Teilmenge von $\widehat{\mathbb{F}}$ aufgefasst werden.*

Ist h strikt positiv, so hat die h -transformierte Markov-Kette wieder dieselbe Doob-Martin-Kompaktifizierung wie die Ausgangskette.

1.4.3 Integraldarstellung nichtnegativer harmonischer Funktionen

Satz 1.5 setzt die Irreduzibilität nicht voraus, sondern stellt Bedingungen an den Green'schen Kern, die, wie bereits festgestellt, auch bei schwacher Irreduzibilität erfüllt sind. Also gilt das Resultat über die Integraldarstellung nichtnegativer harmonischer Funktionen auch in der vorliegenden Situation. Alternativ lässt sich die Beweismethode des Satzes 7.45 aus [29, S. 203] beziehungsweise des Satzes 10-26 aus [17, S. 346] anwenden. Der Vorteil dieser Vorgehensweise ist, dass sich das Maß μ_h in der Integraldarstellung konstruktiv angeben lässt. Zu zeigen ist: Zu jeder nichtnegativen harmonischen Funktion h existiert ein Maß μ_h auf $\widehat{\mathbb{F}}$, das auf den Martin-Rand $\partial\mathbb{F}$ konzentriert ist, so dass

$$h(x) = \int_{\partial\mathbb{F}} K(x, \alpha) \mu_h(d\alpha) \text{ für alle } x \in \mathbb{F}$$

gilt. Sei zunächst h eine strikt positive harmonische Funktion. Der Martin-Kern der h -transformierten Markov-Kette zum Referenzzustand o ist

$$K^{(h)}(x, y) = \frac{h(o)}{h(x)} K(x, y), \quad x, y \in \mathbb{F}. \quad (7)$$

Sei $\tilde{\mu}_o$ die Austrittsverteilung der h -transformierten Kette zum Start in $o \in \mathbb{F}$. Nach dem Satz über die Charakterisierung der Austrittsverteilung gilt:

$$\int_{\partial\mathbb{F}} K^{(h)}(x, \alpha) \tilde{\mu}_o(d\alpha) = 1 \text{ für alle } x \in \mathbb{F}. \quad (8)$$

Daraus folgt durch Multiplikation von $h(x)$ auf beiden Seiten:

$$\int_{\partial\mathbb{F}} K(x, \alpha) h(o) \tilde{\mu}_o(d\alpha) = h(x) \text{ für alle } x \in \mathbb{F}. \quad (9)$$

Also erhalten wir für das darstellende Maß μ_h :

$$\mu_h(A) = h(o)\tilde{\mu}_o(A) \text{ für } A \in \mathcal{B}(\partial\mathbb{F}).$$

Ist insbesondere h im Sinne von $h(o) = 1$ normiert, so ist das darstellende Maß μ_h die Austrittsverteilung der h -transformierten Markov-Kette zum Start in $o \in \mathbb{F}$.

Ist h nichtnegativ, aber nicht strikt positiv, so geht man zu $\mathbb{F}^{(h)}$ über und ersetzt in den obigen Aussagen (7), (8) und (9) dann \mathbb{F} durch $\mathbb{F}^{(h)}$ und $\partial\mathbb{F}$ durch $\partial\mathbb{F}^{(h)}$. In dieser Situation leistet das Maß μ_h mit

$$\mu_h(A) = h(o)\tilde{\mu}_o(A \cap \partial\mathbb{F}^{(h)}) \text{ für } A \in \mathcal{B}(\partial\mathbb{F})$$

das Verlangte. Nach Lemma 1.11 und Korollar 1.12 ist μ_h wohldefiniert. Für $x \in \mathbb{F}^{(h)}$ gilt mit obigen Argumenten

$$h(x) = \int_{\partial\mathbb{F}^{(h)}} K(x, \alpha) \mu_h(d\alpha) = \int_{\partial\mathbb{F}} K(x, \alpha) \mu_h(d\alpha).$$

Da $K(x, \cdot)$ stetig ist, folgt mit Lemma 1.10, dass $K(x, \alpha) = 0$ für alle $x \in \mathbb{F} \setminus \mathbb{F}^{(h)}$ und $\alpha \in \partial\mathbb{F}^{(h)}$ ist. Somit gilt für $x \in \mathbb{F} \setminus \mathbb{F}^{(h)}$:

$$h(x) = 0 = \int_{\partial\mathbb{F}^{(h)}} K(x, \alpha) \mu_h(d\alpha) = \int_{\partial\mathbb{F}} K(x, \alpha) \mu_h(d\alpha).$$

Also folgt die Behauptung.

2 Die Mendel-Urne

Dieser Abschnitt widmet sich der Anwendung der Martinrandtheorie auf einen konkreten Spezialfall. Eine Urne enthalte Kugeln mit zwei Merkmalen: Farbe (beispielsweise Grün oder Gelb) und Oberflächenbeschaffenheit (glatt oder rau). Anfangs sei in der Urne von jeder Farbe und von jeder Oberflächenart mindestens eine Kugel vorhanden. In jedem Schritt wird folgendes zweistufiges Zufallsexperiment durchgeführt. Es handelt sich dabei um eine Kopplung unabhängiger Einzelexperimente. Zuerst zieht man eine Kugel zufällig und gleichverteilt, notiert ihre Farbe und legt die Kugel wieder zurück. Anschließend entnimmt man wieder zufällig und gleichverteilt eine Kugel, notiert ihre Oberflächenart und legt die Kugel wieder zurück. Schließlich wird eine Kugel mit der festgestellten Merkmalskombination (Farbe der ersten Kugel und die Oberflächenart der zweiten Kugel) aus einem externen Vorrat zusätzlich in die Urne gelegt.

Dieses Urnenmodell⁵, genannt Mendel-Urne, beschreiben wir formal durch eine Markov-Kette $(X_n)_{n \in \mathbb{N}}$ mit Werten in $E \subset \mathbb{N}_0^4$ und Start in $\mathbf{b} = (b_1, \dots, b_4)$. Dabei bezeichnen wir die ursprünglichen Anzahlen der vier möglichen Kugelsorten durch b_1 für die grün-glatten, b_2 für die grün-rauen, b_3 für die gelb-rauen und b_4 für die gelb-glatten Kugeln. Die Komponenten $X_1^{(n)}, \dots, X_4^{(n)}$ von X_n stehen also für die Anzahlen der Kugeln der entsprechenden Sorte zum Zeitpunkt n . Die obige Voraussetzung an den anfänglichen Urneninhalt bedeutet somit, dass eine der zwei Bedingungen erfüllt werden muss: entweder $b_1 + b_3 \geq 2$ oder $b_2 + b_4 \geq 2$. Als Zustandsraum setzen wir⁶

$$E = \{ \mathbf{x} \in \mathbb{N}_0^4 : \mathbf{b} \leq \mathbf{x} \text{ komponentenweise} \}.$$

Für ein $\mathbf{x} = (x_1, \dots, x_4) \in E$ sei in diesem Abschnitt $|\mathbf{x}| := \sum_{i=1}^4 x_i$ und es sei \mathbf{e}_i der i -te kanonische Einheitsvektor, $i = 1, \dots, 4$. Mit diesen Bezeichnungen ist der Übergangsmechanismus der Mendel-Urne gegeben durch die Übergangsmatrix $p = (p(\mathbf{x}, \mathbf{y}))_{\mathbf{x}, \mathbf{y} \in E}$ mit

$$\begin{aligned} p(\mathbf{x}, \mathbf{x} + \mathbf{e}_1) &= \frac{(x_1 + x_2)(x_1 + x_4)}{|\mathbf{x}|^2} & p(\mathbf{x}, \mathbf{x} + \mathbf{e}_2) &= \frac{(x_1 + x_2)(x_2 + x_3)}{|\mathbf{x}|^2} \\ p(\mathbf{x}, \mathbf{x} + \mathbf{e}_3) &= \frac{(x_3 + x_4)(x_2 + x_3)}{|\mathbf{x}|^2} & p(\mathbf{x}, \mathbf{x} + \mathbf{e}_4) &= \frac{(x_3 + x_4)(x_1 + x_4)}{|\mathbf{x}|^2} \end{aligned}$$

und $p(\mathbf{x}, \mathbf{y}) = 0$ für $\mathbf{y} \neq \mathbf{x} + \mathbf{e}_i$, $i = 1, \dots, 4$. Es bezeichne $(\Omega, \mathcal{F}, \mathbb{P}_{\mathbf{b}})$ den Hintergrundwahrscheinlichkeitsraum. Wir erinnern daran, dass wir kanonische Markov-Ketten betrachten: Der Grundraum $\Omega = E^{\mathbb{N}}$ ist der Trajektorienraum, \mathcal{F} ist die von den Zylindermengen erzeugte σ -Algebra und die Zufallsgrößen X_n sind die Projektionen auf die n -te Koordinate.

Beim Übergang vom Zeitpunkt n zum Zeitpunkt $n+1$ sind nur vier räumliche Übergänge möglich: Die Anzahl der Kugeln einer der vier Kugelsorten wird um 1 erhöht. Also ist die Markov-Kette in einem starken Sinne transient: Jeder Zustand wird höchstens einmal angenommen. Somit gilt $\mathbb{P}_{\mathbf{b}}(X_n \in E_n) = 1$ für $n \geq 1$, wobei

$$E_n = \{ \mathbf{x} \in \mathbb{N}_0^4 : |\mathbf{x}| = |\mathbf{b}| + n - 1, \mathbf{b} \leq \mathbf{x} \text{ komponentenweise} \}.$$

⁵vorgeschlagen von R. Grübel

⁶Hierbei ist zu beachten, dass der Zustandsraum E von der Wahl von \mathbf{b} abhängt.

Ferner ist diese Markov-Kette schwach irreduzibel: Jeder Zustand kann von \mathbf{b} aus mit positiver Wahrscheinlichkeit erreicht werden. Definieren wir eine Größenfunktion $\phi : E \rightarrow \mathbb{N}$ durch $\phi(\mathbf{x}) = |\mathbf{x}| - |\mathbf{b}| - 1$, so können wir E als eine kombinatorische Familie auffassen. Die bezüglich der Größenfunktion isotone partielle Ordnung ist dabei das komponentenweise \leq . Wir stellen fest, dass es sich bei der Mendel-Urne um eine kombinatorische Markov-Kette im Sinne des Abschnitts 1.4 handelt.

Wie sieht die Doob-Martin-Kompaktifizierung des Zustandsraumes der Mendel-Urne aus? Um den Martin-Kern zu konstruieren betrachten wir zunächst die Mehrschritt-Übergangswahrscheinlichkeiten.

Lemma 2.1. *Für die $(n - m)$ -Schritt-Übergangswahrscheinlichkeit von einem Zustand $\mathbf{x} = (x_1, \dots, x_4) \in E_m$ zu einem Zustand $\mathbf{y} = (y_1, \dots, y_4) \in E_n$, mit $m \leq n$ und $\mathbf{x} \leq \mathbf{y}$ komponentenweise, gilt*

$$\mathbb{P}_{\mathbf{b}}(X_n = \mathbf{y} | X_m = \mathbf{x}) = C_{n-m}^{\mathbf{y}-\mathbf{x}} \cdot W(\mathbf{x}, \mathbf{y})$$

mit

$$W(\mathbf{x}, \mathbf{y}) := \frac{((|\mathbf{x}| - 1)!)^2}{((|\mathbf{y}| - 1)!)^2} \prod_{(i,j) \in D} \frac{(y_i + y_j - 1)!}{(x_i + x_j - 1)!}, \quad D = \{(1, 2), (1, 4), (2, 3), (3, 4)\},$$

und dem Multinomialkoeffizienten

$$C_{n-m}^{\mathbf{y}-\mathbf{x}} := \binom{n-m}{y_1 - x_1, y_2 - x_2, y_3 - x_3, y_4 - x_4} = \frac{(n-m)!}{\prod_{i=1}^4 (y_i - x_i)!}.$$

In allen anderen Fällen ist die entsprechende Übergangswahrscheinlichkeit 0.

Beweis. Wir betrachten den zu der Markov-Kette $(X_n)_{n \in \mathbb{N}}$ zugehörigen Übergangsgraphen. Dieser Graph ist ein unendlicher gerichteter Gittergraph mit der Knotenmenge E und der Kantenmenge

$$\{(\mathbf{x}, \mathbf{x} + \mathbf{e}_i) : \mathbf{x} \in E, i = 1, \dots, 4\}.$$

Jede Kante $(\mathbf{x}, \mathbf{x} + \mathbf{e}_i)$ wird mit der Übergangswahrscheinlichkeit $p(\mathbf{x}, \mathbf{x} + \mathbf{e}_i)$, $i = 1, \dots, 4$, gewichtet. Im Folgenden betrachten wir nur Gitterwege, die stets in die Richtung der kanonischen Einheitsvektoren gehen. Unter dem Gewicht eines solchen Gitterweges verstehen wir das Produkt der Kantengewichte entlang des Weges. Die gesuchte Übergangswahrscheinlichkeit ist somit die Summe der Gewichte aller verschiedenen Gitterwege von \mathbf{x} nach \mathbf{y} . Die Anzahl der Gitterwege von \mathbf{x} nach \mathbf{y} ist gegeben durch den Multinomialkoeffizienten $C_{n-m}^{\mathbf{y}-\mathbf{x}}$.

Es bleibt zu zeigen: Alle Wege haben das gleiche Gewicht, nämlich $W(\mathbf{x}, \mathbf{y})$. Wir führen die Zuwächse

$$Z_i = X_{m+i} - X_{m+i-1} \text{ für } i = 1, \dots, n - m,$$

ein. Die Zuwächse nehmen nur Werte in der Menge $\{\mathbf{e}_1, \dots, \mathbf{e}_4\}$ der Einheitsvektoren an. Ein konkreter Gitterweg ist durch seinen Startpunkt und die Reihenfolge der Zuwächse

eindeutig festgelegt. Es gilt für $i = 1, \dots, n - m$:

$$\begin{aligned} \mathbb{P}_{\mathbf{b}}(Z_i = \mathbf{e}_1 | Z_1 = \mathbf{z}_1, \dots, Z_{i-1} = \mathbf{z}_{i-1}, X_m = \mathbf{x}) \\ = \frac{\left(\sum_{j=1}^{i-1} z_1^{(j)} + \sum_{j=1}^{i-1} z_2^{(j)} + x_1 + x_2 \right) \left(\sum_{j=1}^{i-1} z_1^{(j)} + \sum_{j=1}^{i-1} z_4^{(j)} + x_1 + x_4 \right)}{(|\mathbf{x}| + i - 1)^2}, \end{aligned}$$

wobei $z_1^{(j)}, \dots, z_4^{(j)}$ die Komponenten von \mathbf{z}_j sind. Die drei übrigen Wahrscheinlichkeiten ergeben sich analog dazu. Im folgenden seien die \mathbf{z}_i , $i = 1, \dots, n - m$, so gewählt, dass $\mathbf{x} + \sum_{j=1}^{n-m} \mathbf{z}_j = \mathbf{y}$ gilt. Wir zeigen, dass das Gewicht

$$\mathbb{P}_{\mathbf{b}}(Z_1 = \mathbf{z}_1, \dots, Z_{n-m} = \mathbf{z}_{n-m} | X_m = \mathbf{x})$$

eines konkreten Gitterweges nicht von der Reihenfolge der \mathbf{z}_i , $i = 1, \dots, n - m$, abhängt. Mit der Bezeichnung $\mathbf{s}_k = \sum_{j=1}^k \mathbf{z}_j$ gilt:

$$\begin{aligned} \mathbb{P}_{\mathbf{b}}(Z_1 = \mathbf{z}_1, \dots, Z_{n-m} = \mathbf{z}_{n-m} | X_m = \mathbf{x}) &= \left(\prod_{j=1}^{n-m} (|\mathbf{x}| + j - 1)^2 \right)^{-1} \\ &\times \prod_{\substack{1 \leq j \leq n-m \\ \mathbf{z}_j = \mathbf{e}_1}} \left(s_1^{(j-1)} + s_2^{(j-1)} + x_1 + x_2 \right) \left(s_1^{(j-1)} + s_4^{(j-1)} + x_1 + x_4 \right) \\ &\times \prod_{\substack{1 \leq j \leq n-m \\ \mathbf{z}_j = \mathbf{e}_2}} \left(s_1^{(j-1)} + s_2^{(j-1)} + x_1 + x_2 \right) \left(s_2^{(j-1)} + s_3^{(j-1)} + x_2 + x_3 \right) \\ &\times \prod_{\substack{1 \leq j \leq n-m \\ \mathbf{z}_j = \mathbf{e}_3}} \left(s_3^{(j-1)} + s_4^{(j-1)} + x_3 + x_4 \right) \left(s_2^{(j-1)} + s_3^{(j-1)} + x_2 + x_3 \right) \\ &\times \prod_{\substack{1 \leq j \leq n-m \\ \mathbf{z}_j = \mathbf{e}_4}} \left(s_3^{(j-1)} + s_4^{(j-1)} + x_3 + x_4 \right) \left(s_1^{(j-1)} + s_4^{(j-1)} + x_1 + x_4 \right). \end{aligned}$$

Das Umgruppieren der Produkte liefert das gewünschte Ergebnis, denn es gilt beispielsweise:

$$\begin{aligned} \prod_{\substack{1 \leq j \leq n-m \\ \mathbf{z}_j = \mathbf{e}_1}} \left(s_1^{(j-1)} + s_2^{(j-1)} + x_1 + x_2 \right) \cdot \prod_{\substack{1 \leq j \leq n-m \\ \mathbf{z}_j = \mathbf{e}_2}} \left(s_1^{(j-1)} + s_2^{(j-1)} + x_1 + x_2 \right) \\ = \frac{\left(s_1^{(n-m)} + s_2^{(n-m)} + x_1 + x_2 - 1 \right)!}{(x_1 + x_2 - 1)!}. \end{aligned}$$

Die entscheidende Idee dabei ist, dass bei dieser Darstellung jeder Faktor von $x_1 + x_2$ bis $s_1^{(n-m)} + s_2^{(n-m)} + x_1 + x_2 - 1$ genau 1 mal vorkommt. Insgesamt hängt

$$\begin{aligned} \mathbb{P}_{\mathbf{b}}(Z_1 = \mathbf{z}_1, \dots, Z_{n-m} = \mathbf{z}_{n-m} | X_m = \mathbf{x}) &= \frac{((|\mathbf{x}| - 1)!)^2}{((|\mathbf{y}| - 1)!)^2} \\ &\times \frac{\left(s_1^{(n-m)} + s_2^{(n-m)} + x_1 + x_2 - 1 \right)!}{(x_1 + x_2 - 1)!} \frac{\left(s_3^{(n-m)} + s_4^{(n-m)} + x_3 + x_4 - 1 \right)!}{(x_3 + x_4 - 1)!} \\ &\times \frac{\left(s_1^{(n-m)} + s_4^{(n-m)} + x_1 + x_4 - 1 \right)!}{(x_1 + x_4 - 1)!} \frac{\left(s_2^{(n-m)} + s_3^{(n-m)} + x_2 + x_3 - 1 \right)!}{(x_2 + x_3 - 1)!} \end{aligned}$$

nur von der Summe s_{n-m} der Zuwächse und nicht von ihrer Reihenfolge ab. Da $\mathbf{x} + \sum_{i=1}^{n-m} \mathbf{z}_i = \mathbf{y}$ vorausgesetzt war, gilt:

$$\mathbb{P}_{\mathbf{b}}(Z_1 = \mathbf{z}_1, \dots, Z_{n-m} = \mathbf{z}_{n-m} | X_m = \mathbf{x}) = W(\mathbf{x}, \mathbf{y}). \quad \square$$

Als Folgerung ergibt sich für den Martin-Kern zum Referenzzustand \mathbf{b} :

$$K(\mathbf{x}, \mathbf{y}) = \frac{\mathbb{P}_{\mathbf{b}}(X_n = \mathbf{y} | X_m = \mathbf{x})}{\mathbb{P}_{\mathbf{b}}(X_n = \mathbf{y})} = \frac{C_{n-m}^{\mathbf{y}-\mathbf{x}} W(\mathbf{x}, \mathbf{y})}{C_n^{\mathbf{y}-\mathbf{b}} W(\mathbf{b}, \mathbf{y})},$$

falls \mathbf{x}, \mathbf{y} wie in Lemma 2.1, und 0 sonst. Man überzeugt sich leicht davon, dass der Quotient der Pfadgewichte nicht von \mathbf{y} abhängt:

$$\frac{W(\mathbf{x}, \mathbf{y})}{W(\mathbf{b}, \mathbf{y})} = \frac{((|\mathbf{x}| - 1)!)^2}{((|\mathbf{b}| - 1)!)^2} \prod_{(i,j) \in D} \frac{(b_i + b_j - 1)!}{(x_i + x_j - 1)!} = W(\mathbf{b}, \mathbf{x})^{-1}.$$

Insgesamt hat man:

$$K(\mathbf{x}, \mathbf{y}) = \begin{cases} W(\mathbf{b}, \mathbf{x})^{-1} \frac{C_{n-m}^{\mathbf{y}-\mathbf{x}}}{C_n^{\mathbf{y}-\mathbf{b}}} & \text{falls } \mathbf{x} \in E_m, \mathbf{y} \in E_n, m < n, \\ & \mathbf{x} \leq \mathbf{y} \text{ komponentenweise,} \\ 0 & \text{sonst.} \end{cases} \quad (10)$$

Für die nachfolgenden Überlegungen ist somit nur der Quotient der Multinomialkoeffizienten von Bedeutung.

Satz 2.2. *Es sei $(\mathbf{y}_n)_{n \in \mathbb{N}}$ eine Folge in E mit $|\mathbf{y}_n| \rightarrow \infty$. Dann konvergiert $K(\mathbf{x}, \mathbf{y}_n)$ in \mathbb{R} für alle $\mathbf{x} \in E$ mit $n \rightarrow \infty$ genau dann, wenn gilt:*

$$\lim_{n \rightarrow \infty} \frac{y_i^{(n)}}{|\mathbf{y}_n|} = \alpha_i, \quad i = 1, \dots, 4,$$

für ein Quadrupel $\alpha = (\alpha_1, \dots, \alpha_4)$ aus dem dreidimensionalen Wahrscheinlichkeitssimplex

$$\Delta^3 := \{\alpha \in [0, 1]^4 : \alpha_1 + \dots + \alpha_4 = 1\}.$$

Im Falle der Konvergenz gilt

$$\lim_{n \rightarrow \infty} K(\mathbf{x}, \mathbf{y}_n) = W(\mathbf{b}, \mathbf{x})^{-1} \prod_{i=1}^4 \alpha_i^{x_i - b_i},$$

wobei hier $0^0 := 1$ sei.

Beweis. Wie bereits erwähnt, hängt der Quotient der Gewichte nicht von \mathbf{y}_n ab. Es genügt also den Quotienten der Multinomialkoeffizienten zu betrachten. Dieser lässt sich, falls $\mathbf{x} \leq \mathbf{y}_n$ komponentenweise, wie folgt darstellen:

$$\frac{C_{|\mathbf{y}_n| - |\mathbf{x}|}^{\mathbf{y}_n - \mathbf{x}}}{C_{|\mathbf{y}_n| - |\mathbf{b}|}^{\mathbf{y}_n - \mathbf{b}}} = \frac{(|\mathbf{y}_n| - |\mathbf{x}|)!}{\prod_{i=1}^4 (y_i^{(n)} - x_i)!} \times \frac{\prod_{i=1}^4 (y_i^{(n)} - b_i)!}{(|\mathbf{y}_n| - |\mathbf{b}|)!} = \frac{\prod_{i=1}^4 (y_i^{(n)} - b_i)^{x_i - b_i}}{(|\mathbf{y}_n| - |\mathbf{b}|)^{|\mathbf{x}| - |\mathbf{b}|}}, \quad (11)$$

mit fallenden Faktoriellen

$$n^k := n(n-1)\dots(n-k+1), \quad n^0 = 1.$$

Ist $b_i = x_i = y_i^{(n)}$ für ein $i = 1, \dots, 4$, so ist der i -te Faktor in der Gleichung (11) durch 1 zu ersetzen. Das führt auf die Konvention $0^0 := 1$ für diesen Beweis. Nun lässt sich folgendes ausklammern: $|\mathbf{y}_n|^{|x|-|b|}$ im Nenner und $\prod_{i=1}^4 \left(y_i^{(n)}\right)^{(x_i-b_i)}$, falls alle $y_i^{(n)} \neq 0$, im Zähler. Für den Quotienten der Multinomialkoeffizienten ergibt sich:

$$\frac{C_{|\mathbf{y}_n|-|x|}^{\mathbf{y}_n-\mathbf{x}}}{C_{|\mathbf{y}_n|-|b|}^{\mathbf{y}_n-\mathbf{b}}} = \prod_{i=1}^4 \left(\frac{y_i^{(n)}}{|\mathbf{y}_n|}\right)^{x_i-b_i} \times \frac{\prod_{i=1}^4 \prod_{m=0}^{x_i-b_i-1} \left(1 - \frac{b_i+m}{y_i^{(n)}}\right)}{\prod_{m=0}^{|x|-|b|-1} \left(1 - \frac{|b|+m}{|\mathbf{y}_n|}\right)} \quad (12)$$

Man beachte, dass alle Produkte eine von n unabhängige endliche Anzahl von Faktoren haben.

Zu „ \Leftarrow “. Wir nehmen an, die Verhältnisse konvergieren wie angegeben, also $\lim_{n \rightarrow \infty} \frac{y_i^{(n)}}{|\mathbf{y}_n|} = \alpha_i$ für $i = 1, \dots, 4$. Wir betrachten zunächst nur diejenigen Folgen $(\mathbf{y}_n)_{n \in \mathbb{N}}$, die die Eigenschaft $\mathbf{x} \leq \mathbf{y}_n$ komponentenweise für alle bis auf endlich viele n haben. Für den ersten Term in (12) gilt dann

$$\lim_{n \rightarrow \infty} \prod_{i=1}^4 \left(\frac{y_i^{(n)}}{|\mathbf{y}_n|}\right)^{x_i-b_i} = \prod_{i=1}^4 (\alpha_i)^{x_i-b_i}.$$

Da $|\mathbf{y}_n| \rightarrow \infty$ vorausgesetzt ist, gilt für den Nenner des zweiten Terms:

$$\lim_{n \rightarrow \infty} \prod_{m=0}^{|x|-|b|-1} \left(1 - \frac{|b|+m}{|\mathbf{y}_n|}\right) = 1.$$

Nun ist für jedes i offenbar $0 \leq \alpha_i \leq 1$. Wir betrachten die folgende Fallunterscheidung.

Fall 1: Es gelte $\alpha_i > 0$ für alle i . Daraus folgt dann $y_i^{(n)} \rightarrow \infty$ mit $n \rightarrow \infty$ für jedes i . Anderenfalls existiert eine beschränkte Teilfolge $(y_i^{(n_k)})_{k \in \mathbb{N}}$ von $(y_i^{(n)})_{n \in \mathbb{N}}$, sodass $\lim_{k \rightarrow \infty} \frac{y_i^{(n_k)}}{|\mathbf{y}_n|} = 0$, Widerspruch zu $\alpha_i > 0$. Somit gilt in diesem Fall für den Zähler des zweiten Terms in (12)

$$\lim_{n \rightarrow \infty} \prod_{i=1}^4 \prod_{m=0}^{x_i-b_i-1} \left(1 - \frac{b_i+m}{y_i^{(n)}}\right) = 1,$$

und damit für den Quotienten der Multinomialkoeffizienten

$$\frac{C_{|\mathbf{y}_n|-|x|}^{\mathbf{y}_n-\mathbf{x}}}{C_{|\mathbf{y}_n|-|b|}^{\mathbf{y}_n-\mathbf{b}}} \rightarrow \prod_{i=1}^4 (\alpha_i)^{x_i-b_i} \quad \text{mit } n \rightarrow \infty.$$

Fall 2: Es existiert mindestens ein i mit $\alpha_i = 0$. In diesem Fall konvergiert der erste Term in (12) gegen 0 mit $n \rightarrow \infty$. Da der zweite Term beschränkt ist, konvergiert auch der Quotient der Multinomialkoeffizienten gegen 0 für $n \rightarrow \infty$.

Nun sei $\mathbf{x} \not\leq \mathbf{y}_n$ komponentenweise für unendlich viele n . Einerseits ist dann $K(\mathbf{x}, \mathbf{y}_n) = 0$ unendlich oft. Andererseits bedeutet diese Voraussetzung zusammen mit der Forderung nach der Konvergenz der Verhältnisse, dass $\alpha_i = 0$ für ein $i = 1, \dots, 4$. Insgesamt folgt daraus die Konvergenz von $K(\mathbf{x}, \mathbf{y}_n)$ gegen 0.

Zu „ \Rightarrow “: Angenommen, eines der Verhältnisse konvergiert nicht. Ohne Einschränkung sei das das erste Verhältnis, also

$$0 \leq \liminf_{n \rightarrow \infty} \frac{y_1^{(n)}}{|\mathbf{y}_n|} < \limsup_{n \rightarrow \infty} \frac{y_1^{(n)}}{|\mathbf{y}_n|} \leq 1.$$

Wir zeigen, dass dann ein $\mathbf{x} \in E$ existiert, sodass $K(\mathbf{x}, \mathbf{y}_n)$ in \mathbb{R} mit $n \rightarrow \infty$ nicht konvergiert. Wähle $x_2 = b_2$, $x_4 = b_4$, $x_3 = b_3$ und $x_1 = b_1 + 1$. Dann reduziert sich der Quotient der Multinomialkoeffizienten zu:

$$\frac{y_1^{(n)}}{|\mathbf{y}_n|} \times \frac{1 - \frac{b_1}{y_1^{(n)}}}{1 - \frac{|b|}{|\mathbf{y}_n|}}.$$

Wir zeigen, dass dieser Ausdruck divergiert. In der vorliegenden Situation existieren nämlich Teilfolgen $(y_1^{(n_k)})_{k \in \mathbb{N}}$ und $(y_1^{(n_j)})_{j \in \mathbb{N}}$ von $(y_1^{(n)})_{n \in \mathbb{N}}$, sodass

$$0 \leq \lim_{k \rightarrow \infty} \frac{y_1^{(n_k)}}{|\mathbf{y}_{n_k}|} = \alpha_1 < \beta_1 = \lim_{j \rightarrow \infty} \frac{y_1^{(n_j)}}{|\mathbf{y}_{n_j}|} \leq 1$$

gilt. Mit Argumenten wie im ersten Beweisteil ergibt sich damit

$$\lim_{k \rightarrow \infty} \left(\frac{y_1^{(n_k)}}{|\mathbf{y}_{n_k}|} \times \frac{1 - \frac{b_1}{y_1^{(n_k)}}}{1 - \frac{|b|}{|\mathbf{y}_{n_k}|}} \right) = \alpha_1 \neq \beta_1 = \lim_{j \rightarrow \infty} \left(\frac{y_1^{(n_j)}}{|\mathbf{y}_{n_j}|} \times \frac{1 - \frac{b_1}{y_1^{(n_j)}}}{1 - \frac{|b|}{|\mathbf{y}_{n_j}|}} \right). \quad \square$$

Wir identifizieren die Doob-Martin-Kompaktifizierung \widehat{E} mit $E \cup \Delta^3$. Der Martin-Kern lässt sich wie folgt von $E \times E$ auf $E \times \widehat{E}$ stetig fortsetzen:

$$K(\mathbf{x}, \boldsymbol{\alpha}) = W(\mathbf{b}, \mathbf{x})^{-1} \prod_{i=1}^4 \alpha_i^{x_i - b_i}, \quad (13)$$

wobei $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_4) \in \Delta^3$. Mit der Martinrandtheorie folgt nun: Es existiert eine Δ^3 -wertige Zufallsgröße X_∞ , sodass X_n mit $n \rightarrow \infty$ fast sicher gegen X_∞ in der Doob-Martin-Topologie konvergiert.

Zum Beweis des nächsten Satzes benötigen wir das folgende Hilfsresultat:

Bemerkung 2.3. Betrachten wir die harmonische Funktion $h := K(\cdot, \boldsymbol{\alpha})$ mit $K(\cdot, \boldsymbol{\alpha})$ aus (13), dann ist die h -transformierte Markov-Kette zur Mendel-Urne die Irrfahrt in Richtung kanonischer Einheitsvektoren $\mathbf{e}_1, \dots, \mathbf{e}_4$ mit Parameter $\boldsymbol{\alpha} \in \Delta^3$. Eine solche Irrfahrt hat die Übergangsmatrix $p_\boldsymbol{\alpha} = (p_\boldsymbol{\alpha}(x, y))_{x, y \in E}$ mit

$$p_\boldsymbol{\alpha}(\mathbf{x}, \mathbf{x} + \mathbf{e}_i) = \alpha_i, \quad i = 1, \dots, 4.$$

und $p(\mathbf{x}, \mathbf{y}) = 0$ für $\mathbf{y} \neq \mathbf{x} + \mathbf{e}_i$, $i = 1, \dots, 4$. Diese stimmt mit der Übergangsmatrix der h -transformierten Markov-Kette überein. Man erhält beispielsweise

$$\begin{aligned} p^{(h)}(\mathbf{x}, \mathbf{x} + \mathbf{e}_1) &= \frac{p(\mathbf{x}, \mathbf{x} + \mathbf{e}_1) h(\mathbf{x} + \mathbf{e}_1)}{h(\mathbf{x})} \\ &= \frac{(x_1 + x_2)(x_1 + x_4)}{(|\mathbf{x}|)^2} \cdot \frac{W(\mathbf{b}, \mathbf{x} + \mathbf{e}_1)^{-1} \alpha_1^{x_1+1-b_1} \alpha_2^{x_2-b_2} \alpha_3^{x_3-b_3} \alpha_4^{x_4-b_4}}{W(\mathbf{b}, \mathbf{x})^{-1} \alpha_1^{x_1-b_1} \alpha_2^{x_2-b_2} \alpha_3^{x_3-b_3} \alpha_4^{x_4-b_4}} \\ &= \frac{(x_1 + x_2)(x_1 + x_4)}{(|\mathbf{x}|)^2} \cdot \alpha_1 \cdot \frac{W(\mathbf{b}, \mathbf{x})}{W(\mathbf{b}, \mathbf{x} + \mathbf{e}_1)} = \alpha_1 = p_{\alpha}(\mathbf{x}, \mathbf{x} + \mathbf{e}_1). \end{aligned}$$

Satz 2.4. Die Austrittsverteilung $\mathcal{L}(X_{\infty})$ von $(X_n)_{n \in \mathbb{N}}$ ist das Bildmaß der Produkt-Beta-Verteilung

$$\text{Beta}(b_1 + b_2, b_3 + b_4) \otimes \text{Beta}(b_1 + b_4, b_2 + b_3)$$

auf dem Einheitsquadrat unter der Transformation

$$T : [0, 1]^2 \rightarrow \Delta^3, \quad (u, v) \mapsto (uv, u(1-v), (1-u)(1-v), (1-u)v).$$

Beweis. Es sei $\mu := \text{Beta}(b_1 + b_2, b_3 + b_4) \otimes \text{Beta}(b_1 + b_4, b_2 + b_3)$ und μ^T das Bildmaß von μ unter T . Nach Bemerkung 2.3 ist für jedes $\alpha \in \Delta^3$ die h -transformierte Markov-Kette zur harmonischen Funktion $h = K(\cdot, \alpha)$ eine Irrfahrt in Richtung kanonischer Einheitsvektoren. Nach dem starken Gesetz der großen Zahl konvergiert also eine solche h -transformierte Kette in der Doob-Martin-Topologie fast sicher gegen α . Aus der Diskussion über den minimalen Martin-Rand $\partial_{\min} E$ in Bemerkung 1.6 folgt nun, dass $\partial_{\min} E = \Delta^3$ gilt. Also ist $\mathcal{L}(X_{\infty})$ nach Satz 1.9 das eindeutige repräsentierende Maß der harmonischen Funktion $h \equiv 1$. Zu zeigen ist also:

$$\int_{\Delta^3} K(\mathbf{x}, \alpha) \mu^T(d\alpha) = 1.$$

Wende den Transformationssatz über die Integration bezüglich des Bildmaßes an:

$$\int_{\Delta^3} K(\mathbf{x}, \alpha) \mu^T(d\alpha) = \int_{[0,1]^2} K(\mathbf{x}, T(u, v)) \mu(du, dv).$$

Es reicht also zu zeigen:

$$\int_{[0,1]^2} (uv)^{x_1-b_1} (u(1-v))^{x_2-b_2} ((1-u)(1-v))^{x_3-b_3} ((1-u)v)^{x_4-b_4} \mu(du, dv) = W(\mathbf{b}, \mathbf{x}).$$

Die Beta-Verteilung $\text{Beta}(p, q)$ hat die Dichte

$$f_{p,q}(t) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} t^{p-1} (1-t)^{q-1} \text{ für } t \in [0, 1].$$

Wir erhalten:

$$\begin{aligned}
& \int_{[0,1]^2} (uv)^{x_1-b_1} (u(1-v))^{x_2-b_2} ((1-u)(1-v))^{x_3-b_3} ((1-u)v)^{x_4-b_4} \mu(du, dv) \quad (14) \\
&= \int_0^1 u^{x_1-b_1+x_2-b_2} (1-u)^{x_3-b_3+x_4-b_4} f_{b_1+b_2, b_3+b_4}(u) du \\
&\quad \times \int_0^1 v^{x_1-b_1+x_4-b_4} (1-v)^{x_2-b_2+x_3-b_3} f_{b_1+b_4, b_2+b_3}(v) dv \\
&= \frac{\Gamma(b_1+b_2+b_3+b_4)}{\Gamma(b_1+b_2)\Gamma(b_3+b_4)} \int_0^1 u^{x_1+x_2-1} (1-u)^{x_3+x_4-1} du \\
&\quad \times \frac{\Gamma(b_1+b_2+b_3+b_4)}{\Gamma(b_1+b_4)\Gamma(b_2+b_3)} \int_0^1 v^{x_1+x_4-1} (1-v)^{x_2+x_3-1} dv \\
&= \frac{\Gamma(b_1+b_2+b_3+b_4)^2}{\Gamma(b_1+b_2)\Gamma(b_3+b_4)\Gamma(b_1+b_4)\Gamma(b_2+b_3)} \cdot \frac{\Gamma(x_1+x_2)\Gamma(x_3+x_4)\Gamma(x_1+x_4)\Gamma(x_2+x_3)}{\Gamma(x_1+x_2+x_3+x_4)^2} \\
&= W(\mathbf{b}, \mathbf{x}). \quad \square
\end{aligned}$$

Satz 2.5. *Es sei $X = (X_n)_{n \in \mathbb{N}}$ eine Mendel-Urne mit Start in \mathbf{b} und X_∞ der fast-sichere Grenzwert von $(X_n)_{n \in \mathbb{N}}$ in der Doob-Martin-Kompaktifizierung. Sei außerdem X^α , $\alpha \in \Delta^3$, die Irrfahrt in Richtung kanonischer Einheitsvektoren mit Start in \mathbf{b} und $\mathbb{P}_\mathbf{b}^\alpha$ das zugehörige Wahrscheinlichkeitsmaß auf dem Trajektorienraum. Dann gilt:*

$$\mathcal{L}(X|X_\infty = \alpha) = \mathbb{P}_\mathbf{b}^\alpha.$$

Beweis. Es seien das Maß μ , die Austrittsverteilung μ^T , und die Transformation T wie in Satz 2.4. Zu zeigen ist:

$$\mathbb{P}_\mathbf{b}(X \in A, X_\infty \in B) \stackrel{!}{=} \int_B \mathbb{P}_\mathbf{b}^\alpha(A) \mu^T(d\alpha), \text{ für alle } A \in \mathcal{F}, B \in \mathcal{B}(\Delta^3).$$

Es reicht dies für Zylindermengen $A \in \mathcal{F}$ zu zeigen. Wir erinnern an die Definition der Zylindermengen:

$$C_{(\mathbf{x}_1, \dots, \mathbf{x}_n)} = \{(\mathbf{y}_1, \mathbf{y}_1, \dots) \in E^\mathbb{N} : \mathbf{y}_1 = \mathbf{x}_1, \dots, \mathbf{y}_n = \mathbf{x}_n\}.$$

Ferner kann man sich auf solche Zylindermengen beschränken, für die $(\mathbf{x}_0, \dots, \mathbf{x}_n)$ ein Gitterweg von \mathbf{b} nach \mathbf{x}_n ist, mit $\mathbf{x}_0 = \mathbf{b}$ und Einheitsvektoren $\{\mathbf{e}_1, \dots, \mathbf{e}_4\}$ als Zuwächsen. Es ist also zu zeigen:

$$\mathbb{P}_\mathbf{b}(X \in C_{(\mathbf{b}, x_2, \dots, \mathbf{x}_n)}, X_\infty \in B) \stackrel{!}{=} \int_B \mathbb{P}_\mathbf{b}^\alpha(C_{(\mathbf{b}, x_2, \dots, \mathbf{x}_n)}) \mu^T(d\alpha), \text{ für alle } B \in \mathcal{B}(\Delta^3). \quad (15)$$

Nach Satz 2.4 ist $\mathcal{L}(X_\infty|X_n = \mathbf{x}_n) = \nu^T$, wobei

$$\nu = \text{Beta}\left(x_1^{(n)} + x_2^{(n)}, x_3^{(n)} + x_4^{(n)}\right) \otimes \text{Beta}\left(x_1^{(n)} + x_4^{(n)}, x_2^{(n)} + x_3^{(n)}\right).$$

Somit erhalten wir für die linke Seite von (15):

$$\begin{aligned}\mathbb{P}_{\mathbf{b}}(X \in C_{(\mathbf{b}, x_1, \dots, x_n)}, X_\infty \in B) &= \mathbb{P}_{\mathbf{b}}(X_\infty \in B \mid X_1 = \mathbf{b}, X_2 = \mathbf{x}_2, \dots, X_n = \mathbf{x}_n) \\ &\quad \times \mathbb{P}_{\mathbf{b}}(X_1 = \mathbf{b}, X_2 = \mathbf{x}_2, \dots, X_n = \mathbf{x}_n) \\ &= W(\mathbf{b}, \mathbf{x}_n) \int_B \nu^T(d\boldsymbol{\alpha})\end{aligned}$$

Mit einer Rechnung wie im Beweis zum Satz 2.4 folgt:

$$\begin{aligned}W(\mathbf{b}, \mathbf{x}_n) \int_B \nu^T(d\boldsymbol{\alpha}) &= W(\mathbf{b}, \mathbf{x}_n) \int_{T^{-1}(B)} f_{x_1^{(n)}+x_2^{(n)}, x_3^{(n)}+x_4^{(n)}}(u) f_{x_1^{(n)}+x_4^{(n)}, x_2^{(n)}+x_3^{(n)}}(v) du dv \\ &= W(\mathbf{b}, \mathbf{x}_n) \frac{\Gamma(x_1^{(n)} + x_2^{(n)} + x_3^{(n)} + x_4^{(n)})^2}{\Gamma(x_1^{(n)} + x_2^{(n)})\Gamma(x_3^{(n)} + x_4^{(n)})\Gamma(x_1^{(n)} + x_4^{(n)})\Gamma(x_2^{(n)} + x_3^{(n)})} \\ &\quad \times \int_{T^{-1}(B)} u^{x_1^{(n)}+x_2^{(n)}-1} (1-u)^{x_3^{(n)}+x_4^{(n)}-1} du \\ &\quad \times \int_{T^{-1}(B)} v^{x_1^{(n)}+x_4^{(n)}-1} (1-v)^{x_2^{(n)}+x_3^{(n)}-1} dv \\ &= \frac{\Gamma(b_1 + b_2 + b_3 + b_4)^2}{\Gamma(b_1 + b_2)\Gamma(b_3 + b_4)\Gamma(b_1 + b_4)\Gamma(b_2 + b_3)} \\ &\quad \times \int_{T^{-1}(B)} u^{x_1^{(n)}+x_2^{(n)}-1} (1-u)^{x_3^{(n)}+x_4^{(n)}-1} du \\ &\quad \times \int_{T^{-1}(B)} v^{x_1^{(n)}+x_4^{(n)}-1} (1-v)^{x_2^{(n)}+x_3^{(n)}-1} dv.\end{aligned}$$

Für die rechte Seite von (15) gilt:

$$\begin{aligned}\int_B \mathbb{P}_{\mathbf{b}}^\alpha(C_{(\mathbf{b}, x_2, \dots, x_n)}) \mu^T(d\boldsymbol{\alpha}) &= \int_B \alpha_1^{x_1^{(n)}-b_1} \alpha_2^{x_2^{(n)}-b_2} \alpha_3^{x_3^{(n)}-b_3} \alpha_4^{x_4^{(n)}-b_4} \mu^T(d\boldsymbol{\alpha}) \\ &= \int_{T^{-1}(B)} \left((uv)^{x_1^{(n)}-b_1} (u(1-v))^{x_2^{(n)}-b_2} \right. \\ &\quad \left. \times ((1-u)(1-v))^{x_3^{(n)}-b_3} ((1-u)v)^{x_4^{(n)}-b_4} \right) \mu(du, dv) \\ &\stackrel{*}{=} \frac{\Gamma(b_1 + b_2 + b_3 + b_4)^2}{\Gamma(b_1 + b_2)\Gamma(b_3 + b_4)\Gamma(b_1 + b_4)\Gamma(b_2 + b_3)} \\ &\quad \times \int_{T^{-1}(B)} u^{x_1^{(n)}+x_2^{(n)}-1} (1-u)^{x_3^{(n)}+x_4^{(n)}-1} du \\ &\quad \times \int_{T^{-1}(B)} v^{x_1^{(n)}+x_4^{(n)}-1} (1-v)^{x_2^{(n)}+x_3^{(n)}-1} dv,\end{aligned}$$

zu (*) siehe Gleichung (14) aus Satz 2.4. Das vollendet den Beweis, denn die linke und die rechte Seite von (15) stimmen überein. \square

Bemerkung 2.6. Es seien $b_1, \dots, b_4 > 0$. Wir betrachten das Pólya'sche Urnenmodell mit 4 Farben. Bekanntlich ist die Grenzverteilung der Anteile der Kugelfarben die Dirichlet-Verteilung $\text{Dir}_{b_1, \dots, b_4}$, siehe [18, S.537]. Sie ist definiert durch

$$\text{Dir}_{b_1, \dots, b_4}(A) = \int \mathbf{1}_A(\alpha_1, \dots, \alpha_4) f_{b_1, \dots, b_4}(\alpha_1, \dots, \alpha_4) d\alpha_1 d\alpha_2 d\alpha_3$$

für ein messbares $A \subset \Delta^3$ mit

$$f_{b_1, \dots, b_4}(\alpha_1, \dots, \alpha_4) = \frac{\Gamma(b_1 + \dots + b_4)}{\Gamma(b_1) \dots \Gamma(b_4)} \prod_{i=1}^4 \alpha_i^{b_i-1}.$$

Die ganzzahligen Parameter b_1, \dots, b_4 beschreiben die Anfangsbelegung der Urne: b_i ist die Anzahl der Kugeln der i -ten Farbe zum Startzeitpunkt. Das Pólya'sche Urnenmodell mit 4 Farben ist eine h -transformierte Markov-Kette zur Mendel-Urne, wobei die harmonische Funktion h gegeben ist durch

$$\begin{aligned} h(\mathbf{x}) &= \int K(\mathbf{x}, \boldsymbol{\alpha}) \text{Dir}_{b_1, \dots, b_4}(d\boldsymbol{\alpha}) \\ &= \int \mathbf{1}_{\Delta^3}(\alpha_1, \dots, \alpha_4) W(\mathbf{b}, \mathbf{x})^{-1} \prod_{i=1}^4 \alpha_i^{x_i-b_i} \frac{\Gamma(b_1 + \dots + b_4)}{\Gamma(b_1) \dots \Gamma(b_4)} \prod_{i=1}^4 \alpha_i^{b_i-1} d\alpha_1 d\alpha_2 d\alpha_3 \\ &= W(\mathbf{b}, \mathbf{x})^{-1} \frac{\Gamma(b_1 + \dots + b_4)}{\Gamma(b_1) \dots \Gamma(b_4)} \frac{\Gamma(x_1) \dots \Gamma(x_4)}{\Gamma(x_1 + \dots + x_4)} \\ &\quad \times \int \mathbf{1}_{\Delta^3}(\alpha_1, \dots, \alpha_4) \frac{\Gamma(x_1 + \dots + x_4)}{\Gamma(x_1) \dots \Gamma(x_4)} \prod_{i=1}^4 \alpha_i^{x_i-1} d\alpha_1 d\alpha_2 d\alpha_3 \\ &= W(\mathbf{b}, \mathbf{x})^{-1} \frac{\Gamma(b_1 + \dots + b_4)}{\Gamma(b_1) \dots \Gamma(b_4)} \frac{\Gamma(x_1) \dots \Gamma(x_4)}{\Gamma(x_1 + \dots + x_4)}. \end{aligned}$$

Man kann sich davon überzeugen, indem man die zugehörige Übergangsmatrix bestimmt:

$$\begin{aligned} p^{(h)}(\mathbf{x}, \mathbf{x} + \mathbf{e}_1) &= \frac{p(\mathbf{x}, \mathbf{x} + \mathbf{e}_1) h(\mathbf{x} + \mathbf{e}_1)}{h(\mathbf{x})} \\ &= p(\mathbf{x}, \mathbf{x} + \mathbf{e}_1) \frac{W(\mathbf{b}, \mathbf{x})}{W(\mathbf{b}, \mathbf{x} + \mathbf{e}_1)} \frac{\Gamma(x_1 + 1) \Gamma(x_2) \dots \Gamma(x_4)}{\Gamma(x_1 + \dots + x_4 + 1)} \frac{\Gamma(x_1 + \dots + x_4)}{\Gamma(x_1) \dots \Gamma(x_4)} \\ &= \frac{x_1}{x_1 + \dots + x_4}. \end{aligned}$$

Analog erhält man

$$p^{(h)}(\mathbf{x}, \mathbf{x} + \mathbf{e}_i) = \frac{x_i}{|\mathbf{x}|}.$$

Das ist tatsächlich der Übergangsmechanismus des Pólya'schen Urnenmodells mit 4 Farben. Dabei bedeutet der Zustand \mathbf{x} zum Zeitpunkt n , dass die Anzahl aller Kugeln i -ter Farbe in der Urne gleich x_i ist⁷.

⁷Man beachte, dass in der Literatur oft nur die hinzugekommenen Kugeln betrachtet werden.

3 Zufällige quaternäre Bäume

In der Informatik gehören Bäume zu den klassischen Datenstrukturen. Quaternäre Bäume wurden von R. A. Finkel und J. L. Bentley in [9] als eine Verallgemeinerung der Idee binärer Suchbäume für den Fall zweidimensionaler Daten vorgeschlagen. Ziel dieses Abschnittes ist es, die Doob-Martin-Kompaktifizierung zu den zufälligen quaternären Bäumen zu bestimmen. Dabei verwenden wir die gängige Annahme, dass die Daten Realisierungen von unabhängigen, auf dem Einheitsquadrat gleichverteilten Zufallsgrößen sind⁸. Wir werden erkennen, dass die zufälligen quaternären Bäume zu den „trickle-down“-Prozessen, einer in [8] von S. N. Evans, R. Grübel und A. Wakolbinger vorgestellten Klasse transienter Markov-Ketten, gehören. Nach der Zusammenstellung der benötigten Grundbegriffe werden wir die „trickle-down“-Konstruktion am Beispiel zufälliger quaternärer Bäume erläutern und anschließend die Methoden aus [8] in der vorliegenden Situation anwenden. Außerdem werden wir feststellen, dass die Doob-Martin-Kompaktifizierung als Ausgangspunkt für die Erforschung der Baumfunktionale zufälliger quaternärer Bäume, wie beispielsweise der internen Pfadlänge, dienen kann. Die Ideen aus diesem Abschnitt können auch auf den Fall d -dimensionaler Daten mit $d > 2$ verallgemeinert werden.

3.1 Grundbegriffe und Bezeichnungen

Wir definieren zunächst quaternäre Bäume unter Anwendung des Formalismus aus der Theoretischen Informatik, siehe dazu Anhang A.1: Wörter und Alphabete.

Definition 3.1. *Ein quaternärer Baum ist eine präfixstabile endliche Teilmenge \mathbf{t} von $\{1, 2, 3, 4\}^*$.*

Die Menge der endlichen quaternären Bäume bezeichnen wir mit \mathbb{T}^{quad} . Quaternäre Bäume bilden eine kombinatorische Familie mit der Größenfunktion $\phi : \mathbb{T}^{\text{quad}} \rightarrow \mathbb{N}$, $\phi(\mathbf{t}) = \#\mathbf{t}$, und der Enthaltenseinsrelation als partieller Ordnung.

Der Grund für die Interpretation präfixstabiler Wortmengen als Bäume wird deutlich, wenn man zur graphentheoretischen Sichtweise wechselt. Auf einer Knotenmenge \mathbf{t} , die den Bedingungen aus der Definition 3.1 genügt, lässt sich mit Hilfe der folgenden Festlegung der Kantenmenge ein gerichteter Graph definieren: $(u, v) \in \mathbf{t} \times (\mathbf{t} \setminus \{\Lambda\})$ ist eine gerichtete Kante von u nach v genau dann, wenn $u = \bar{v}$ gilt. In der Sprache der Graphentheorie heißen die Elemente von \mathbf{t} interne Knoten, das leere Wort entspricht dem Wurzelknoten. Ein $u \in \{1, 2, 3, 4\}^*$ heißt externer Knoten eines nichtleeren Baumes \mathbf{t} , falls $u \notin \mathbf{t}$ und $\bar{u} \in \mathbf{t}$ gilt. Dabei nennen wir den Rumpf \bar{u} den direkten Vorgänger von u . Ein externer Knoten ist also selbst kein Element von \mathbf{t} , hat aber einen direkten Vorgänger in \mathbf{t} . Die Menge der externen Knoten von \mathbf{t} bezeichnet man mit $\partial\mathbf{t}$. Die direkten Nachfolger von $u = u_1 \dots u_n$ sind: $u \circ 1 = u_1 \dots u_n 1, \dots, u \circ 4 = u_1 \dots u_n 4$. Die Tiefe eines Knotens u ist die Länge des eindeutigen Pfades von der Wurzel Λ nach u , also gleich der Länge $|u|$. Den Teilbaum eines Knotens $u \in \mathbf{t}$ definiert man als

$$\mathbf{t}(u) = \{v \in \{1, 2, 3, 4\}^* : u \circ v \in \mathbf{t}\},$$

⁸Zufällige quaternäre Bäume zu den auf dem Einheitsquadrat gleichverteilten Daten wurden beispielsweise in [4], [10] und [23] untersucht.

dieser ist wieder ein quaternärer Baum im Sinne der Definition 3.1. Die Abbildung 1 illustriert den Baum $\{\Lambda, 1, 2, 4, 22, 43\}$ in der üblichen Weise: Die internen Knoten werden als Punkte, die externen Knoten werden als Quadrate und die Kanten werden als Linien dargestellt.

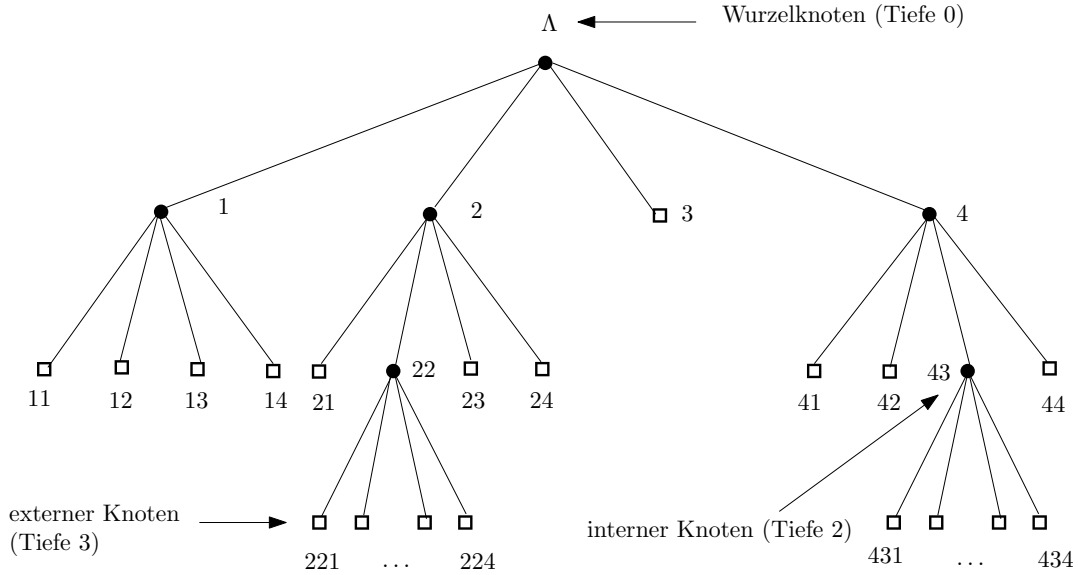


Abbildung 1: Der Baum $\{\Lambda, 1, 2, 4, 22, 43\}$ mit 6 internen und 19 externen Knoten.

Wie bereits erwähnt, werden quaternäre Bäume als Datenstruktur verwendet: Die Daten werden sequentiell in den Knoten eines quaternären Baumes mit Hilfe einer Zuordnung gespeichert. Formal benötigt man dazu die Definition des bewerteten Baumes⁹. Ein bewerteter quaternärer Baum ist ein Paar (\mathbf{t}, φ) mit einem quaternären Baum \mathbf{t} und einer Abbildung $\varphi : \mathbf{t} \rightarrow \mathbb{R}^2$. Sei nun $(x_i)_{i \in \mathbb{N}}$ eine Folge von paarweise verschiedenen Punkten $x_i \in \mathbb{R}^2$. Der folgende Algorithmus konstruiert dazu eine Folge von bewerteten Bäumen $(\mathbf{t}_n, \varphi_n)$, $n \in \mathbb{N}$ (siehe Abbildung 2):

- Für $n = 1$ gelte $\mathbf{t}_1 = \{\Lambda\}$ und $\varphi_1(\Lambda) = x_1$, das heißt, der erste Baum besteht nur aus dem Wurzelknoten bewertet mit x_1 .
- Die Bewertung des Wurzelknotens teilt den Grundraum \mathbb{R}^2 in vier Quadranten: NO, NW, SW und SO (siehe den oberen Teil von Abbildung 2). Diese bezeichnen wir nach üblicher Konvention jeweils mit $1, \dots, 4$. Im zweiten Schritt vergleicht man x_2 mit dem Inhalt x_1 des Wurzelknotens und stellt fest, in welchem Quadranten sich der Punkt x_2 befindet. Der entsprechende externe Knoten wird in einen internen umgewandelt und mit x_2 bewertet, das heißt

$$\mathbf{t}_2 = \begin{cases} \{\Lambda, (1)\}, & \text{falls } x_{2,1} > x_{1,1}, x_{2,2} > x_{1,2}, \\ \{\Lambda, (2)\}, & \text{falls } x_{2,1} < x_{1,1}, x_{2,2} > x_{1,2}, \\ \{\Lambda, (3)\}, & \text{falls } x_{2,1} < x_{1,1}, x_{2,2} < x_{1,2}, \\ \{\Lambda, (4)\}, & \text{falls } x_{2,1} > x_{1,1}, x_{2,2} < x_{1,2}. \end{cases}$$

⁹Im allgemeinen Fall sind beliebige Knotenbewertungen vorstellbar. In diesem und in den nachfolgenden Abschnitten ist die Wahl des Wertebereichs der Bewertungsabbildung auf den betreffenden Algorithmus abgestimmt.

- Sei $(\mathbf{t}_n, \varphi_n)$ schon konstruiert. Jeder Knoten teilt seinen Subquadranten in vier weitere Quadranten. Man vergleicht x_{n+1} nacheinander mit dem Inhalt $\varphi_n(u)$ eines Knotens u , beginnend mit dem Wurzelknoten. Ist x_{n+1} im i -ten Subquadrant von u , so geht man zum Knoten $u \circ i$ über und wiederholt den Test, bis man einen externen Knoten v gefunden hat. Dieser wird in einen internen Knoten umgewandelt und mit x_{n+1} bewertet. Man erhält $\mathbf{t}_{n+1} = \mathbf{t}_n \cup \{v\}$, $\varphi_{n+1}(v) = x_{n+1}$ und $\varphi_{n+1}(u) = \varphi_n(u)$ für alle $u \in \mathbf{t}_n$.

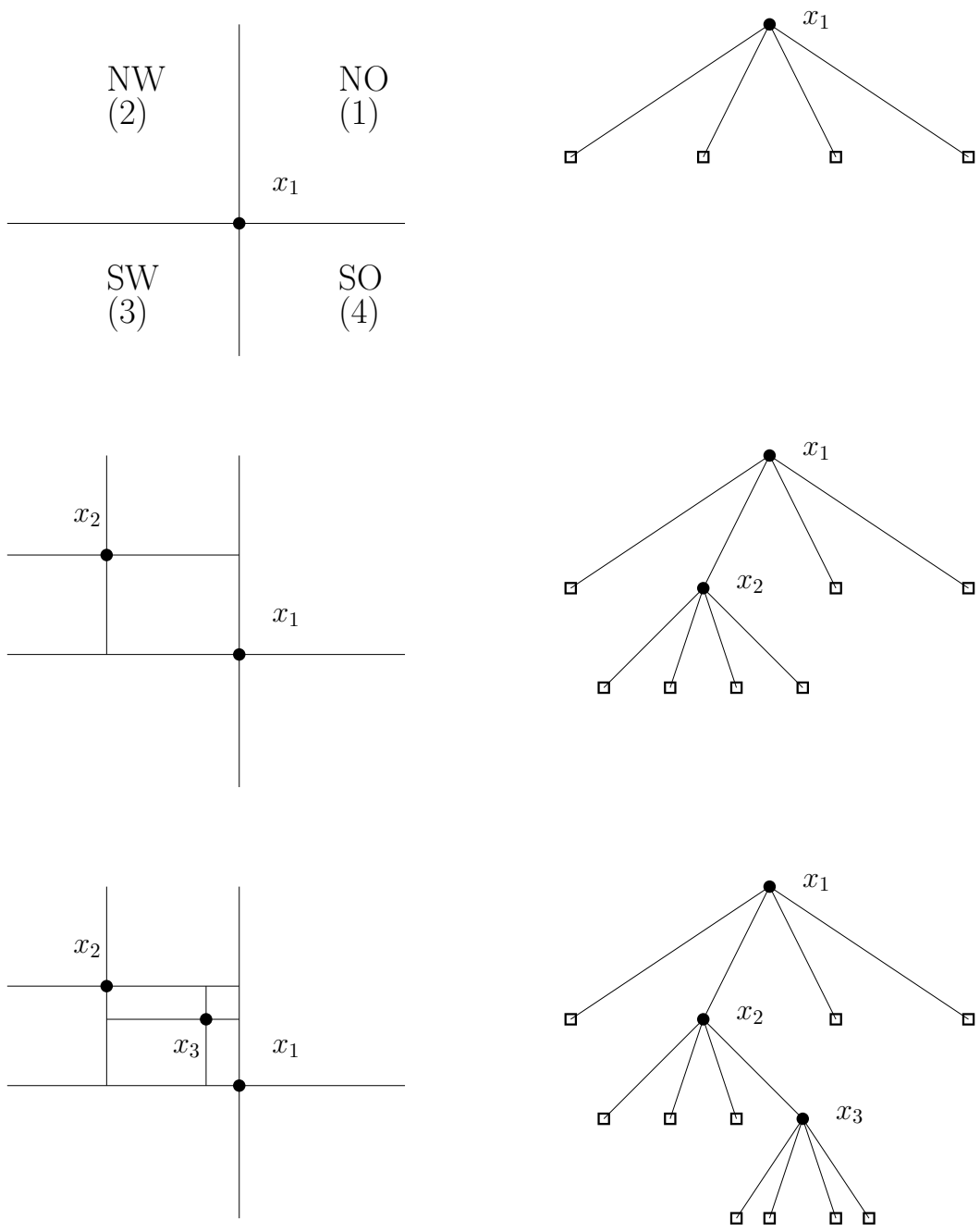


Abbildung 2: Ein Beispiel zum Quadtree-Algorithmus. Die Knoten sind mit Bewertungen beschriftet. Die Beschriftung mit den Wörtern aus $\{1, 2, 3, 4\}^*$ wurde einfachheitshalber weggelassen.

Die probabilistische Analyse untersucht das Verhalten solcher Bäume bei der Verwendung von zufälligen Daten. Es sei dazu $(\eta_i)_{i \in \mathbb{N}}$ eine Folge unabhängiger, auf dem Einheitsquadrat $[0, 1]^2$ gleichverteilter Zufallsgrößen. Die Anwendung des obigen Algorithmus auf den Input $(\eta_i)_{i \in \mathbb{N}}$ liefert eine Folge von zufälligen bewerteten quaternären Bäumen. Entfernt man die Bewertungen, so erhält man eine kombinatorische Markov-Kette $(X_n)_{n \in \mathbb{N}}$, die wir als „Quadtree-Prozess“ oder auch „Quadtree-Kette“ bezeichnen.

Für das weitere Vorgehen benötigen wir das folgende Resultat über die Entwicklung der Teilbaumgrößen an dem Wurzelknoten:

Bemerkung 3.2. Wir nennen den Teilbaum $\mathbf{t}(1)$ den Nordost-Teilbaum eines Baumes $\mathbf{t} \in \mathbb{T}^{\text{quad}}$, entsprechend werden die anderen drei Teilbäume bezeichnet. Es sei X_n der Baum zu η_1, \dots, η_n . Mit $Z_{n,i}$, $i = 1, \dots, 4$ bezeichnen wir die Anzahl der Knoten im i -ten Teilbaum von X_n . Beim Einfügen des $n + 1$ ten Wertes bestimmt der Rang der Komponenten von η_{n+1} den Teilbaum, in dem ein externer Knoten in einen internen umgewandelt wird. Es ist bekannt, dass der Rang von $\eta_{n+1,1}$ unter $\eta_{1,1}, \dots, \eta_{n,1}$ gleichverteilt auf $\{1, \dots, n + 1\}$ ist. Diese Zahl nennen wir „horizontaler Rang“ von η_{n+1} . Analog ist der Rang von $\eta_{n+1,2}$ unter $\eta_{1,2}, \dots, \eta_{n,2}$ gleichverteilt auf $\{1, \dots, n + 1\}$. Diesen bezeichnen wir als „vertikalen Rang“. Da die Komponenten von η_n unabhängig sind, sind auch die Ränge unabhängig. Angenommen, es gilt $Z_n = z_n$. Das bedeutet, dass der Baum $n = z_{n,1} + z_{n,2} + z_{n,3} + z_{n,4} + 1$ Knoten besitzt. Wir berechnen die Wahrscheinlichkeit für das Ereignis $\{Z_{n+1} = z_n + \mathbf{e}_1\}$, dass der $n + 1$ -te Knoten zum Nordost-Teilbaum hinzukommt. Das ist der Fall, wenn η_{n+1} den vertikalen Rang $\geq z_{n,3} + z_{n,4} + 1$ und den horizontalen Rang $\geq z_{n,2} + z_{n,3} + 1$ hat. Daraus folgt:

$$\mathbb{P}(Z_{n+1} = z_n + \mathbf{e}_1 | Z_n = z_n) = \frac{(z_{n,1} + z_{n,2} + 1)(z_{n,1} + z_{n,4} + 1)}{(z_{n,1} + z_{n,2} + z_{n,3} + z_{n,4} + 2)^2}.$$

Analog erhält man

$$\begin{aligned} \mathbb{P}(Z_{n+1} = z_n + \mathbf{e}_2 | Z_n = z_n) &= \frac{(z_{n,1} + z_{n,2} + 1)(z_{n,2} + z_{n,3} + 1)}{(z_{n,1} + z_{n,2} + z_{n,3} + z_{n,4} + 2)^2}, \\ \mathbb{P}(Z_{n+1} = z_n + \mathbf{e}_3 | Z_n = z_n) &= \frac{(z_{n,3} + z_{n,4} + 1)(z_{n,2} + z_{n,3} + 1)}{(z_{n,1} + z_{n,2} + z_{n,3} + z_{n,4} + 2)^2}, \\ \mathbb{P}(Z_{n+1} = z_n + \mathbf{e}_4 | Z_n = z_n) &= \frac{(z_{n,3} + z_{n,4} + 1)(z_{n,1} + z_{n,4} + 1)}{(z_{n,1} + z_{n,2} + z_{n,3} + z_{n,4} + 2)^2}. \end{aligned}$$

Zusammenfassend stellen wir fest: Der Prozess $(Z_n + (1, 0, 1, 0)')_{n \in \mathbb{N}}$ ist eine Markov-Kette mit dem gleichen Übergangsmechanismus wie die Mendel-Urne mit Start in $(1, 0, 1, 0)'$. Man beachte, dass es auch möglich ist, $(Z_n + (0, 1, 0, 1)')_{n \in \mathbb{N}}$ als eine Mendel-Urne mit Start in $(0, 1, 0, 1)'$ zu betrachten.

Der Ansatz aus [8] macht es möglich, den Quadtree-Prozess $(X_n)_{n \in \mathbb{N}}$ durch Verschachtelung unabhängiger Mendel-Urnen zu konstruieren. Diese Vorgehensweise liefert unter anderem eine Möglichkeit zur Berechnung des Martin-Kerns zur Quadtree-Kette, indem man die Resultate zur Mendel-Urne verwendet.

3.2 Die „trickle-down“-Konstruktion

Zur leichteren Lesbarkeit verzichten wir auf die Beschreibung der allgemeinen Version der „trickle-down“-Konstruktion und skizzieren stattdessen die Konstruktionsidee am Beispiel zufälliger quaternärer Bäume. Für die formale Definition verweisen auf die Originalpublikation [8].

Ausgangspunkt ist $\{1, \dots, 4\}^*$, betrachtet als vollständiger quaternärer Baum aus der Sicht der Graphentheorie. Am Anfang sind alle Knoten unbesetzt. In jedem Schritt werden Partikel in den Wurzelknoten eingespeist. Das erste Partikel besetzt die Wurzel. Die anderen Partikel sickern entlang der gerichteten Kanten durch, bis sie einen unbesetzten Knoten gefunden haben. Das „Durchsickern“ geschieht gemäß einem stochastischen Mechanismus. Jeder Knoten $u \in \{1, \dots, 4\}^*$ wird mit einem stochastischen Prozess $Y^u = (Y_n^u)_{n \in \mathbb{N}_0}$ versehen, sodass $(Y_n^u + (1, 0, 1, 0)')_{n \in \mathbb{N}_0}$ eine Mendel-Urne mit Start in $(1, 0, 1, 0)'$ ist, wobei die Unabhängigkeit der Prozesse vorausgesetzt wird. Die Realisierungen der $(Y_n^u)_{n \in \mathbb{N}_0}$ dienen als Weiterleitungsanweisungen. Der stochastische Mechanismus funktioniert folgendermaßen:

- $n = 1$: Das erste Partikel bleibt in der Wurzel stehen.
- $n = 2$: Da die Wurzel besetzt ist, wird das Partikel zum i -ten direkten Nachfolger weitergeleitet, wenn $Y_1^\Lambda = \mathbf{e}_i$ ist. Da alle direkten Nachfolger unbesetzt sind, kommt das zweite Partikel in einem davon zum Stillstand.
- $n > 2$: Das Partikel wird abhängig von dem Wert von $Y_{n-1}^\Lambda - Y_{n-2}^\Lambda$ zum einem direkten Nachfolger der Wurzel weitergeleitet und sickert durch, bis es einen unbesetzten Knoten gefunden hat. Ist es an einem besetzten Knoten u , so geht es zu $u \circ i$ über, falls der i -te Einheitsvektor \mathbf{e}_i der Wert von $Y_{A_n^u}^u - Y_{A_{n-1}^u}^u$ ist. Dabei ist A_n^u die „lokale Zeit“: die Anzahl der zur (globalen) Zeit n in u weitergeleiteten Partikel.

Das Resultat dieser Konstruktion hat dieselbe Verteilung wie der Quadtree-Prozess $(X_n)_{n \in \mathbb{N}}$. Abbildung 3 zeigt den Zustand zu den Zeitpunkten $n = 8$, $n = 9$ und $n = 10$ der „trickle-down“-Konstruktion zu den Realisierungen

$$\begin{aligned} y_7^\Lambda &= (3, 3, 0, 1)', & y_8^\Lambda &= (3, 4, 0, 1)', & y_9^\Lambda &= (3, 5, 0, 1)', \\ y_2^1 &= (0, 1, 0, 1)', \\ y_2^2 &= (1, 1, 0, 0)', & y_3^2 &= (1, 2, 0, 0)', & y_4^2 &= (1, 2, 1, 0)', \\ y_1^{22} &= (0, 0, 0, 1). \end{aligned}$$

Die Anfangsstücke der Weiterleitungsanweisungen können aus dem Zustand zum Zeitpunkt $n = 8$ nicht rekonstruiert werden. Die zugehörigen Realisierungen der relevanten „lokalen Zeitpunkte“ sind

$$\begin{array}{lll} a_8^\Lambda = 7, & a_9^\Lambda = 8, & a_{10}^\Lambda = 9, \\ a_8^1 = 2, & a_9^1 = 2, & a_{10}^1 = 2, \\ a_8^2 = 2, & a_9^2 = 3, & a_{10}^2 = 4, \\ a_8^{22} = 0, & a_9^{22} = 1, & a_{10}^{22} = 1. \end{array}$$

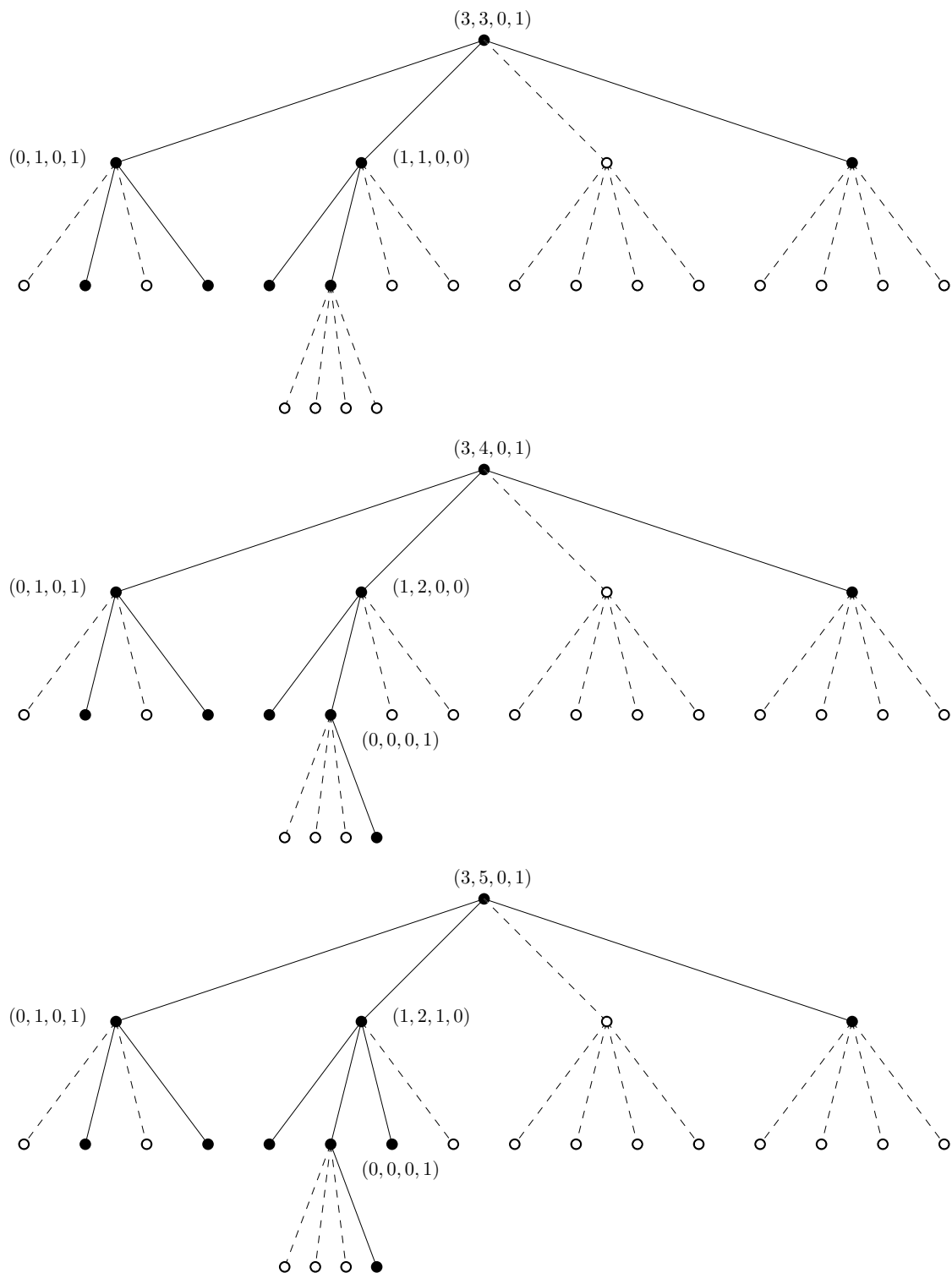


Abbildung 3: Ein Beispiel für die „trickle-down“-Konstruktion zu den Realisierungen $y_7^\Lambda = (3, 3, 0, 1)'$, $y_8^\Lambda = (3, 4, 0, 1)'$, $y_9^\Lambda = (3, 5, 0, 1)'$, $y_2^1 = (0, 1, 0, 1)'$, $y_2^2 = (1, 1, 0, 0)'$, $y_3^2 = (1, 2, 0, 0)'$, $y_4^2 = (1, 2, 1, 0)'$, $y_1^{22} = (0, 0, 0, 1)$. Die Knoten sind mit den zugehörigen Weiterleitungsanweisungen beschriftet. Die Beschriftung mit den Wörtern aus $\{1, 2, 3, 4\}^*$ wurde wieder weggelassen.

3.3 Doob-Martin Kompaktifizierung für den Quadtree-Prozess

Da die Quadtree-Kette $(X_n)_{n \in \mathbb{N}}$ ein spezieller „trickle-down“-Prozess ist, kann nun der Martin-Kern mit Hilfe der Resultate aus [8] bestimmt werden.

Lemma 3.3. *Es seien $\mathbf{s}, \mathbf{t} \in \mathbb{T}^{\text{quad}}$ mit $\mathbf{s} \subset \mathbf{t}$ und $D := \{(1, 2), (1, 4), (2, 3), (3, 4)\}$. Der Martin-Kern der Quadtree-Kette $(X_n)_{n \in \mathbb{N}}$ zum Referenzzustand $\mathbf{t}_0 = \{\Lambda\}$ ist gegeben durch*

$$K(\mathbf{s}, \mathbf{t}) = \frac{\alpha(\mathbf{s})}{\#\mathbf{s}!} \left(\frac{\#\mathbf{t}}{\#\mathbf{s}} \right)^{-1} \prod_{u \in \mathbf{s}} \#\mathbf{t}(u),$$

wobei der Faktor

$$\alpha(\mathbf{s}) := \prod_{u \in \mathbf{s}} \frac{(\#\mathbf{s}(u)!)^2}{\prod_{(i,j) \in D} (\#\mathbf{s}(u \circ i) + \#\mathbf{s}(u \circ j))!}$$

nicht von \mathbf{t} abhängt.

Beweis. Wegen der Verwandtschaft der quaternären und binären Bäume können Beweistechniken aus [8, S.17-19], Beispiel 4.3, übertragen werden. Sei $x = (x^u)_{u \in \{1, \dots, 4\}^*}$ mit

$$x^u = (x_1^u, \dots, x_4^u) := (\#\mathbf{s}(u \circ 1), \dots, \#\mathbf{s}(u \circ 4))$$

und $y = (y^u)_{u \in \{1, \dots, 4\}^*}$ mit

$$y^u = (y_1^u, \dots, y_4^u) := (\#\mathbf{t}(u \circ 1), \dots, \#\mathbf{t}(u \circ 4)).$$

Man beachte, dass $x^u = y^u = (0, 0, 0, 0)$ für alle bis auf endlich viele $u \in \{1, \dots, 4\}^*$ gilt. Nach Korollar 4.2 aus [8] ist der Martin-Kern der Markov-Kette $(X_n)_{n \in \mathbb{N}}$ gegeben durch

$$K(\mathbf{s}, \mathbf{t}) = \prod_{u \in \{1, \dots, 4\}^*} K^u(x^u, y^u),$$

wobei K^u der Martin-Kern der Markov-Kette Y^u zum Referenzzustand $(0, 0, 0, 0)$ ist. Hierbei ist hervorzuheben, dass in diesem Produkt alle bis auf endlich viele Faktoren gleich 1 sind. Da $(Y_n^u + (1, 0, 1, 0)')_{n \in \mathbb{N}_0}$ eine Mendel-Urne mit Start in $(1, 0, 1, 0)'$ ist, erhalten wir mit der Gleichung (10) aus Abschnitt 2 für den Martin-Kern von Y^u :

$$K^u(x^u, y^u) = \frac{\binom{y_1^u + \dots + y_4^u - x_1^u - \dots - x_4^u}{y_1^u - x_1^u, \dots, y_4^u - x_4^u}}{\binom{y_1^u + \dots + y_4^u}{y_1^u, \dots, y_4^u}} \frac{(x_1^u + \dots + x_4^u + 1)!^2}{(x_1^u + x_2^u)! (x_1^u + x_4^u)! (x_2^u + x_3^u)! (x_3^u + x_4^u)!}.$$

Somit erhält man für den Martin-Kern von $(X_n)_{n \in \mathbb{N}}$:

$$\begin{aligned} K(\mathbf{s}, \mathbf{t}) &= \prod_{u \in \{1, \dots, 4\}^*} K^u(x^u, y^u) \\ &= \prod_{u \in \{1, \dots, 4\}^*} \left(\frac{y_1^u! \dots y_4^u!}{(y_1^u + \dots + y_4^u + 1)!} (y_1^u + \dots + y_4^u + 1) \right. \\ &\quad \times \frac{(y_1^u + \dots + y_4^u - x_1^u - \dots - x_4^u)!}{(y_1^u - x_1^u)! \dots (y_4^u - x_4^u)!} \\ &\quad \left. \times \frac{(x_1^u + \dots + x_4^u + 1)!^2}{(x_1^u + x_2^u)! (x_1^u + x_4^u)! (x_2^u + x_3^u)! (x_3^u + x_4^u)!} \right). \end{aligned}$$

Außerdem hat man

$$\begin{aligned} x_1^u + \dots + x_4^u + 1 &= \#\mathbf{s}(u) \text{ für jedes } u \in \mathbf{s}, \\ y_1^u + \dots + y_4^u + 1 &= \#\mathbf{t}(u) \text{ für jedes } u \in \mathbf{t}. \end{aligned}$$

Für den ersten Term in $K(\mathbf{s}, \mathbf{t})$ gilt folglich:

$$\begin{aligned} \prod_{u \in \{1, \dots, 4\}^*} \frac{y_1^u! \dots y_4^u!}{(y_1^u + \dots + y_4^u + 1)!} &= \prod_{u \in \{1, \dots, 4\}^*} \frac{\#\mathbf{t}(u \circ 1)! \dots \#\mathbf{t}(u \circ 4)!}{\#\mathbf{t}(u)!} \\ &= \frac{1}{\#\mathbf{t}(\Lambda)!} = \frac{1}{\#\mathbf{t}!} \text{ (Teleskopprodukt)} \end{aligned}$$

und

$$\prod_{u \in \{1, \dots, 4\}^*} (y_1^u + \dots + y_4^u + 1) = \prod_{u \in \mathbf{t}} \#\mathbf{t}(u).$$

Für den zweiten Term erhält man wieder mit einem Teleskopargument:

$$\begin{aligned} \prod_{u \in \{1, \dots, 4\}^*} \frac{(y_1^u + \dots + y_4^u - x_1^u - \dots - x_4^u)!}{(y_1^u - x_1^u)! \dots (y_4^u - x_4^u)!} \\ &= \prod_{u \in \mathbf{s}} \frac{(\#\mathbf{t}(u) - \#\mathbf{s}(u))!}{(\#\mathbf{t}(u \circ 1) - \#\mathbf{s}(u \circ 1))! \dots (\#\mathbf{t}(u \circ 4) - \#\mathbf{s}(u \circ 4))!} \\ &\quad \times \prod_{u \in \mathbf{t} \setminus \mathbf{s}} \frac{(\#\mathbf{t}(u) - 1)!}{\#\mathbf{t}(u \circ 1)! \#\mathbf{t}(u \circ 2)! \#\mathbf{t}(u \circ 3)! \#\mathbf{t}(u \circ 4)!} \\ &= \frac{(\#\mathbf{t} - \#\mathbf{s})!}{\prod_{u \in \mathbf{t} \setminus \mathbf{s}} \#\mathbf{t}(u)}. \end{aligned}$$

Schließlich haben wir noch den dritten Term, welcher $\alpha(\mathbf{s})$ entspricht. Insgesamt ergibt sich:

$$\begin{aligned} K(\mathbf{s}, \mathbf{t}) &= \frac{1}{\#\mathbf{t}!} \prod_{u \in \mathbf{t}} \#\mathbf{t}(u) \frac{(\#\mathbf{t} - \#\mathbf{s})!}{\prod_{u \in \mathbf{t} \setminus \mathbf{s}} \#\mathbf{t}(u)} \alpha(\mathbf{s}) \\ &= \frac{\alpha(\mathbf{s})}{\#\mathbf{s}!} \frac{\#\mathbf{s}! (\#\mathbf{t} - \#\mathbf{s})!}{\#\mathbf{t}!} \prod_{u \in \mathbf{s}} \#\mathbf{t}(u). \quad \square \end{aligned}$$

In [8] findet man ein allgemeines Resultat zur Bestimmung des Martin-Randes eines „trickle-down“-Prozesses mit Hilfe der Martin-Ränder der Komponentenprozesse Y^u . In einem Spezialfall kann man diesen auch direkt erhalten.

Satz 3.4. *Eine Folge $(\mathbf{t}_n)_{n \in \mathbb{N}}$ aus \mathbb{T}^{quad} , welche $\#\mathbf{t}_n \rightarrow \infty$ erfüllt, konvergiert genau dann in der Doob-Martin Kompaktifizierung, wenn $\frac{\#\mathbf{t}_n(u)}{\#\mathbf{t}_n}$ für alle $u \in \{1, \dots, 4\}^*$ konvergiert.*

Beweis. Die Folge $(\mathbf{t}_n)_{n \in \mathbb{N}}$ mit $\#\mathbf{t}_n \rightarrow \infty$ konvergiert genau dann in der Doob-Martin Kompaktifizierung, wenn die folgende Bedingung erfüllt ist: $(K(\mathbf{s}, \mathbf{t}_n))_{n \in \mathbb{N}}$ ist für jedes

$\mathbf{s} \in \mathbb{T}^{\text{quad}}$ eine in \mathbb{R} konvergente Folge. Der Martin-Kern lässt sich wie folgt umschreiben:

$$\begin{aligned}
K(\mathbf{s}, \mathbf{t}_n) &= \frac{\alpha(\mathbf{s})}{\#\mathbf{s}!} \binom{\#\mathbf{t}_n}{\#\mathbf{s}}^{-1} \prod_{u \in \mathbf{s}} \#\mathbf{t}_n(u) \\
&= \alpha(\mathbf{s}) \frac{(\#\mathbf{t}_n - \#\mathbf{s})!}{\#\mathbf{t}_n!} \#\mathbf{t}_n^{\#\mathbf{s}} \prod_{u \in \mathbf{s}} \frac{\#\mathbf{t}_n(u)}{\#\mathbf{t}_n} \\
&= \alpha(\mathbf{s}) \frac{\#\mathbf{t}_n}{\#\mathbf{t}_n} \frac{\#\mathbf{t}_n}{\#\mathbf{t}_n - 1} \cdots \frac{\#\mathbf{t}_n}{\#\mathbf{t}_n - \#\mathbf{s} + 1} \prod_{u \in \mathbf{s}} \frac{\#\mathbf{t}_n(u)}{\#\mathbf{t}_n}. \tag{16}
\end{aligned}$$

„ \Leftarrow “ Angenommen, es gilt $\frac{\#\mathbf{t}_n(u)}{\#\mathbf{t}_n} \rightarrow r(u)$ mit $r(u) \in [0, 1]$ und $n \rightarrow \infty$ für alle $u \in \{1, \dots, 4\}^*$. Dann folgt:

$$\lim_{n \rightarrow \infty} K(\mathbf{s}, \mathbf{t}_n) = \alpha(\mathbf{s}) \prod_{u \in \mathbf{s}} r(u) \quad \text{für alle } \mathbf{s} \in \mathbb{T}^{\text{quad}}.$$

„ \Rightarrow “ Angenommen, es existiert eine Menge $A \subset \{1, \dots, 4\}^*$, sodass $\frac{\#\mathbf{t}_n(u)}{\#\mathbf{t}_n}$ für alle $u \in A$ divergiert. Zu zeigen ist: Es existiert ein quaternärer Baum \mathbf{s} , sodass $(K(\mathbf{s}, \mathbf{t}_n))_{n \in \mathbb{N}}$ keine konvergente Folge ist. Sei dazu

$$k := \min \left\{ |u| \in A : \frac{\#\mathbf{t}_n(u)}{\#\mathbf{t}_n} \text{ divergiert} \right\}.$$

Wähle ein $\tilde{u} \in A$ mit $|\tilde{u}| = k$. Schließlich sei \mathbf{s} der quaternäre Baum, der aus dem eindeutigen Pfad von der Wurzel Λ zum Knoten \tilde{u} besteht. Diese Wahl von \tilde{u} sichert, dass für alle $u \in \mathbf{s} \setminus \tilde{u}$ das Verhältnis $\frac{\#\mathbf{t}_n(u)}{\#\mathbf{t}_n}$ konvergiert. Folglich konvergiert $(K(\mathbf{s}, \mathbf{t}_n))_{n \in \mathbb{N}}$ für diesen speziellen Baum \mathbf{s} nicht in \mathbb{R} , denn für den letzten Term in (16) gilt:

$$\prod_{u \in \mathbf{s}} \frac{\#\mathbf{t}_n(u)}{\#\mathbf{t}_n} = \frac{\#\mathbf{t}_n(\tilde{u})}{\#\mathbf{t}_n} \prod_{u \in \mathbf{s} \setminus \tilde{u}} \frac{\#\mathbf{t}_n(u)}{\#\mathbf{t}_n}.$$

Man beachte dabei, dass aus $\lim_{n \rightarrow \infty} \prod_{u \in \mathbf{s} \setminus \tilde{u}} \frac{\#\mathbf{t}_n(u)}{\#\mathbf{t}_n} = 0$ auch $\lim_{n \rightarrow \infty} \frac{\#\mathbf{t}_n(\tilde{u})}{\#\mathbf{t}_n} = 0$ folgen würde. \square

Es sei $\{1, \dots, 4\}^\infty$ die Menge aller unendlichen Folgen aus $\{1, \dots, 4\}$. Der Martin-Rand $\partial\mathbb{T}^{\text{quad}}$ kann mit dem Raum der Wahrscheinlichkeitsmaße auf $\{1, \dots, 4\}^\infty$ versehen mit der von den Mengen

$$A_u := \{v \in \{1, \dots, 4\}^\infty : u < v\},$$

$u \in \{1, \dots, 4\}^*$, erzeugten σ -Algebra identifiziert werden. Wir schreiben $u < v$ für ein $u = u_1 \dots u_k \in \{1, \dots, 4\}^*$ und ein $v = v_1 v_2 \dots \in \{1, \dots, 4\}^\infty$, falls $u_1 \dots u_k = v_1 \dots v_k$ gilt. Eine Folge $(\mathbf{t}_n)_{n \in \mathbb{N}}$ quaternärer Bäume mit $\#\mathbf{t}_n \rightarrow \infty$ konvergiert genau dann gegen ein $\mu \in \partial\mathbb{T}^{\text{quad}}$, wenn gilt:

$$\lim_{n \rightarrow \infty} \frac{\#\mathbf{t}_n(u)}{\#\mathbf{t}_n} = \mu(A_u).$$

Der erweiterte Martin-Kern ist gegeben durch

$$K(\mathbf{t}, \mu) = \alpha(\mathbf{t}) \prod_{u \in \mathbf{t}} \mu(A_u).$$

Nach den Resultaten der Martinrandtheorie existiert eine $\partial\mathbb{T}^{\text{quad}}$ -wertige Zufallsgröße X_∞ , sodass X_n mit $n \rightarrow \infty$ fast sicher gegen X_∞ in der Doob-Martin-Topologie konvergiert. Zur Beschreibung der Austrittsverteilung zeigen wir zunächst das folgende Hilfsresultat:

Lemma 3.5. *Der Martin-Rand $\partial\mathbb{T}^{\text{quad}}$ der Quadtree-Kette stimmt mit dem minimalen Martin-Rand $\partial_{\min}\mathbb{T}^{\text{quad}}$ überein.*

Beweis. Wir betrachten die h -transformierte Markov-Kette zur harmonischen Funktion $h = K(\cdot, \mu)$ mit einem beliebigen $\mu \in \partial\mathbb{T}^{\text{quad}}$. Diese hat den Zustandsraum

$$\{\mathbf{t} \in \mathbb{T}^{\text{quad}} : \mu(A_u) > 0 \text{ für alle } u \in \mathbf{t}\}. \quad (17)$$

Es sei p die Übergangsmatrix der Quadtree-Kette. Wir bestimmen nun die Übergangsmatrix $p^{(h)}$ der h -transformierten Kette. Für \mathbf{s}, \mathbf{t} mit $\mathbf{t} \neq \mathbf{s} \cup \{u\}$ für ein $u \in \partial\mathbf{s}$ gilt $p(\mathbf{s}, \mathbf{t}) = 0$. Folglich reicht es, den Übergang von \mathbf{t} zu $\mathbf{t} \cup \{u\}$ aus dem Zustandsraum in (17) mit $u \in \partial\mathbf{t}$ zu betrachten:

$$\begin{aligned} p^{(h)}(\mathbf{t}, \mathbf{t} \cup \{u\}) &= p(\mathbf{t}, \mathbf{t} \cup \{u\}) \frac{K(\mathbf{t} \cup \{u\}, \mu)}{K(\mathbf{t}, \mu)} = p(\mathbf{t}, \mathbf{t} \cup \{u\}) \frac{\alpha(\mathbf{t} \cup \{u\}) \prod_{v \in \mathbf{t} \cup \{u\}} \mu(A_v)}{\alpha(\mathbf{t}) \prod_{v \in \mathbf{t}} \mu(A_v)} \\ &= p(\mathbf{t}, \mathbf{t} \cup \{u\}) \frac{\alpha(\mathbf{t} \cup \{u\})}{\alpha(\mathbf{t})} \mu(A_u). \end{aligned}$$

Beim Übergang von \mathbf{t} zu $\mathbf{t} \cup \{u\}$ erhöhen sich die Teilbaumgrößen entlang des Pfades von der Wurzel Λ zum Knoten u um 1. Teilbaumgrößen an den Knoten v mit $v \not\leq u$ bleiben unverändert. Folglich ist

$$\begin{aligned} \frac{\alpha(\mathbf{t} \cup \{u\})}{\alpha(\mathbf{t})} &= \prod_{\substack{\Lambda < v \leq u \\ \varepsilon(v)=1}} \frac{(\#\mathbf{t}(\bar{v}) + 1)^2}{(\#\mathbf{t}(\bar{v} \circ 1) + \#\mathbf{t}(\bar{v} \circ 2) + 1)(\#\mathbf{t}(\bar{v} \circ 1) + \#\mathbf{t}(\bar{v} \circ 4) + 1)} \\ &\times \prod_{\substack{\Lambda < v \leq u \\ \varepsilon(v)=2}} \frac{(\#\mathbf{t}(\bar{v}) + 1)^2}{(\#\mathbf{t}(\bar{v} \circ 1) + \#\mathbf{t}(\bar{v} \circ 2) + 1)(\#\mathbf{t}(\bar{v} \circ 2) + \#\mathbf{t}(\bar{v} \circ 3) + 1)} \\ &\times \prod_{\substack{\Lambda < v \leq u \\ \varepsilon(v)=3}} \frac{(\#\mathbf{t}(\bar{v}) + 1)^2}{(\#\mathbf{t}(\bar{v} \circ 3) + \#\mathbf{t}(\bar{v} \circ 4) + 1)(\#\mathbf{t}(\bar{v} \circ 2) + \#\mathbf{t}(\bar{v} \circ 3) + 1)} \\ &\times \prod_{\substack{\Lambda < v \leq u \\ \varepsilon(v)=4}} \frac{(\#\mathbf{t}(\bar{v}) + 1)^2}{(\#\mathbf{t}(\bar{v} \circ 3) + \#\mathbf{t}(\bar{v} \circ 4) + 1)(\#\mathbf{t}(\bar{v} \circ 1) + \#\mathbf{t}(\bar{v} \circ 4) + 1)}. \end{aligned}$$

Nach der „trickle-down“-Konstruktion ist dies der Kehrwert der Übergangswahrscheinlichkeit $p(\mathbf{t}, \mathbf{t} \cup \{u\})$ der Quadtree-Kette. Insgesamt erhalten wir:

$$p^{(h)}(\mathbf{t}, \mathbf{t} \cup \{u\}) = \mu(A_u),$$

mit \mathbf{t} und $\mathbf{t} \cup \{u\}$ aus dem Zustandsraum in (17), $u \in \partial\mathbf{t}$. Eine ähnliche Situation liegt im Fall der zufälligen binären Bäume vor, siehe [8, §5]. Man kann zeigen, dass die h -Transformierte wieder eine Markov-Kette von dem „trickle-down“-Typ ist. Mit dem starken Gesetz der großen Zahl folgt: für jedes $\mu \in \partial\mathbb{T}^{\text{quad}}$ konvergiert die h -transformierte Markov-Kette zu $h = K(\cdot, \mu)$ in der Doob-Martin-Topologie fast sicher gegen μ . Mit der Bemerkung 1.6 über den minimalen Martin-Rand erhalten wir schließlich, dass $\partial\mathbb{T}^{\text{quad}} = \partial_{\min}\mathbb{T}^{\text{quad}}$ gilt. \square

Wir werden nun zeigen, dass X_∞ ein zufälliges Wahrscheinlichkeitsmaß mit folgenden Eigenschaften ist:

Satz 3.6. Für jedes $u \in \{1, \dots, 4\}^*$ gilt

$$X_\infty(A_u) = \prod_{\Lambda < v \leq u} Z_{\bar{v}, \varepsilon(v)}$$

in Verteilung. Dabei sind die Zufallsgrößen

$$Z_u := (Z_{u,1}, Z_{u,2}, Z_{u,3}, Z_{u,4}), \quad u \in \{1, \dots, 4\}^*,$$

unabhängig. Die Verteilung¹⁰ von Z_u ist das Bildmaß von $\text{unif}(0, 1)^2$ unter

$$T : [0, 1]^2 \rightarrow \Delta^3, \quad (q, r) \mapsto (qr, q(1-r), (1-q)(1-r), (1-q)r).$$

Beweis. Es sei Z ein zufälliges Wahrscheinlichkeitsmaß, das den Mengen A_u die Masse $\prod_{\Lambda < v \leq u} Z_{\bar{v}, \varepsilon(v)}$ zuordnet. Da nach Lemma 3.5 der Martin-Rand der Quadtree-Kette mit dem minimalen Martin-Rand übereinstimmt, reicht es nach dem Satz 1.9 über die Charakterisierung der Austrittsverteilung zu zeigen:

$$EK(\mathbf{t}, Z) = 1$$

für alle $\mathbf{t} \in \mathbb{T}^{\text{quad}}$. Wir stellen fest:

$$EK(\mathbf{t}, Z) = \alpha(\mathbf{t}) E \left(\prod_{u \in \mathbf{t}} Z(A_u) \right) = \alpha(\mathbf{t}) E \left(\prod_{u \in \mathbf{t}} \prod_{\Lambda < v \leq u} Z_{\bar{v}, \varepsilon(v)} \right).$$

Also gilt

$$E \left(\prod_{u \in \mathbf{t}} Z(A_u) \right) = E \left(\prod_{u \in \mathbf{t}} Z_{u,1}^{\#\mathbf{t}(u \circ 1)} Z_{u,2}^{\#\mathbf{t}(u \circ 2)} Z_{u,3}^{\#\mathbf{t}(u \circ 3)} Z_{u,4}^{\#\mathbf{t}(u \circ 4)} \right).$$

Damit erhalten wir für $E \left(\prod_{u \in \mathbf{t}} Z(A_u) \right)$ Folgendes:

$$\begin{aligned} & \prod_{u \in \mathbf{t}} \int_0^1 \int_0^1 (qr)^{\#\mathbf{t}(u \circ 1)} (q(1-r))^{\#\mathbf{t}(u \circ 2)} ((1-q)(1-r))^{\#\mathbf{t}(u \circ 3)} ((1-q)r)^{\#\mathbf{t}(u \circ 4)} dq dr \\ &= \prod_{u \in \mathbf{t}} \left(\int_0^1 q^{\#\mathbf{t}(u \circ 1) + \#\mathbf{t}(u \circ 2)} (1-q)^{\#\mathbf{t}(u \circ 3) + \#\mathbf{t}(u \circ 4)} dq \right. \\ & \quad \left. \times \int_0^1 r^{\#\mathbf{t}(u \circ 1) + \#\mathbf{t}(u \circ 4)} (1-r)^{\#\mathbf{t}(u \circ 2) + \#\mathbf{t}(u \circ 3)} dr \right) \\ &= \prod_{u \in \mathbf{t}} \left(\frac{(\#\mathbf{t}(u \circ 1) + \#\mathbf{t}(u \circ 2))! (\#\mathbf{t}(u \circ 3) + \#\mathbf{t}(u \circ 4))!}{(\#\mathbf{t}(u \circ 1) + \dots + \#\mathbf{t}(u \circ 4) + 1)!} \right. \\ & \quad \left. \times \frac{(\#\mathbf{t}(u \circ 1) + \#\mathbf{t}(u \circ 4))! (\#\mathbf{t}(u \circ 2) + \#\mathbf{t}(u \circ 3))!}{(\#\mathbf{t}(u \circ 1) + \dots + \#\mathbf{t}(u \circ 4) + 1)!} \right) = \frac{1}{\alpha(\mathbf{t})}. \end{aligned}$$

Somit ist $EK(\mathbf{t}, Z) = 1$ für alle $\mathbf{t} \in \mathbb{T}^{\text{quad}}$ und der Beweis abgeschlossen. \square

¹⁰vergleiche Satz 2.4 über die Austrittsverteilung der Mendel-Urne.

Bemerkung 3.7. Der algorithmische Hintergrund liefert einen alternativen Ansatz zum Beweis der fast sicheren Konvergenz der Quadtree-Kette in der Doob-Martin-Topologie, vergleiche die Vorgehensweise für binäre Bäume in [13]. Sei $(X_n)_{n \in \mathbb{N}}$ die Folge zufälliger quaternärer Bäume zu einer Folge $(\eta_i)_{i \in \mathbb{N}}$ unabhängiger, auf dem Einheitsquadrat gleichverteilter Zufallsvariablen. Für ein $u = u_1 \dots u_k \in \{1, \dots, 4\}^*$ definieren wir das Reverse von u als $\overleftarrow{u} := u_k \dots u_1$ und

$$j_a(u) := \inf \{1 \leq i \leq |u| : u_i = a\} \text{ für ein } a \in \{1, \dots, 4\},$$

mit $\inf \emptyset = \infty$. Ferner benutzen wir in dieser Bemerkung die abkürzende Schreibweise $l \wedge m := \min \{l, m\}$. Der Zeitpunkt der Hinzunahme des Knotens u zum zufälligen Baum ist

$$\tau(u) = \inf \{n \in \mathbb{N} : X_n \ni u\},$$

dies ist eine fast sicher endliche Stoppzeit. Nach Konstruktion enthält der i -te Teilbaum eines Knotens Datenpunkte aus dem i -ten Subquadrant dieses Knotens. Ist i der letzte Buchstabe von u , so befindet sich u im i -ten Teilbaum seines direkten Vorgängers \bar{u} und seine Bewertung $\eta_{\tau(u)}$ in dessen i -ten Subquadranten. Der Post- $\tau(u)$ -Prozesses $(\eta_{\tau(u)+i})_{i \in \mathbb{N}}$ hat die starke Markov-Eigenschaft und ist wieder eine Folge unabhängiger, $\text{unif}(0, 1)^2$ -verteilter Zufallsvariablen. Wenden wir eine Verallgemeinerung des Satzes von Glivenko-Cantelli an, so erhalten wir die folgende Darstellung des fast-sicheren Grenzwertes X_∞ in Abhängigkeit der Inputfolge: Für alle $u = u_1 \dots u_k \in \{1, \dots, 4\}^*$ gilt $\lim_{n \rightarrow \infty} \frac{\#X_n(u)}{n} = X_\infty(A_u)$ fast sicher, wobei

- im Falle $u_k = 1$

$$X_\infty(A_u) = \left| \eta_{\tau(\bar{u}),1} - \begin{cases} 1 & \text{falls } j_2(\overleftarrow{u}) \wedge j_3(\overleftarrow{u}) = \infty \\ \eta_{\tau(u_1 \dots u_{k-j_2(\overleftarrow{u}) \wedge j_3(\overleftarrow{u})}),1} & \text{sonst} \end{cases} \right| \\ \times \left| \eta_{\tau(\bar{u}),2} - \begin{cases} 1 & \text{falls } j_3(\overleftarrow{u}) \wedge j_4(\overleftarrow{u}) = \infty \\ \eta_{\tau(u_1 \dots u_{k-j_3(\overleftarrow{u}) \wedge j_4(\overleftarrow{u})}),2} & \text{sonst} \end{cases} \right|,$$

- im Falle $u_k = 2$

$$X_\infty(A_u) = \left| \eta_{\tau(\bar{u}),1} - \begin{cases} 0 & \text{falls } j_1(\overleftarrow{u}) \wedge j_4(\overleftarrow{u}) = \infty \\ \eta_{\tau(u_1 \dots u_{k-j_1(\overleftarrow{u}) \wedge j_4(\overleftarrow{u})}),1} & \text{sonst} \end{cases} \right| \\ \times \left| \eta_{\tau(\bar{u}),2} - \begin{cases} 1 & \text{falls } j_3(\overleftarrow{u}) \wedge j_4(\overleftarrow{u}) = \infty \\ \eta_{\tau(u_1 \dots u_{k-j_3(\overleftarrow{u}) \wedge j_4(\overleftarrow{u})}),2} & \text{sonst} \end{cases} \right|,$$

- im Falle $u_k = 3$

$$X_\infty(A_u) = \left| \eta_{\tau(\bar{u}),1} - \begin{cases} 0 & \text{falls } j_1(\overleftarrow{u}) \wedge j_4(\overleftarrow{u}) = \infty \\ \eta_{\tau(u_1 \dots u_{k-j_1(\overleftarrow{u}) \wedge j_4(\overleftarrow{u})}),1} & \text{sonst} \end{cases} \right| \\ \times \left| \eta_{\tau(\bar{u}),2} - \begin{cases} 0 & \text{falls } j_1(\overleftarrow{u}) \wedge j_2(\overleftarrow{u}) = \infty \\ \eta_{\tau(u_1 \dots u_{k-j_1(\overleftarrow{u}) \wedge j_2(\overleftarrow{u})}),2} & \text{sonst} \end{cases} \right|,$$

- im Falle $u_k = 4$

$$X_\infty(A_u) = \left| \eta_{\tau(\bar{u}),1} - \begin{cases} 1 & \text{falls } j_2(\overleftarrow{u}) \wedge j_3(\overleftarrow{u}) = \infty \\ \eta_{\tau(u_1 \dots u_{k-j_2(\overleftarrow{u}) \wedge j_3(\overleftarrow{u})}),1} & \text{sonst} \end{cases} \right| \\ \times \left| \eta_{\tau(\bar{u}),2} - \begin{cases} 0 & \text{falls } j_1(\overleftarrow{u}) \wedge j_2(\overleftarrow{u}) = \infty \\ \eta_{\tau(u_1 \dots u_{k-j_1(\overleftarrow{u}) \wedge j_2(\overleftarrow{u})}),2} & \text{sonst} \end{cases} \right|.$$

Betrachten wir als Beispiel $u = 3141114$. Es ist $k = |u| = 7$, $\overleftarrow{u} = 4111413$,

$$j_1(\overleftarrow{u}) = 2, \quad j_2(\overleftarrow{u}) = \infty, \quad j_3(\overleftarrow{u}) = 7, \quad j_4(\overleftarrow{u}) = 1.$$

Da $u_k = 4$ ist, gilt wegen $\tau(\Lambda) = 1$

$$X_\infty(A_{3141114}) = |\eta_{\tau(314111),1} - \eta_{1,1}| \cdot |\eta_{\tau(314111),2} - \eta_{\tau(31411),2}|.$$

Dieser Ansatz wird im Abschnitt 4.4 für zufällige rekursive Bäume ausführlich betrachtet.

3.4 Der Vier-Farben-Baum als h -Transformierte des Quadtree-Prozesses

Wir definieren ein neues zufälliges Wahrscheinlichkeitsmaß auf $\{1, \dots, 4\}^\infty$ versehen mit der von den Mengen A_u , $u \in \{1, \dots, 4\}^*$, erzeugten σ -Algebra durch

$$S(A_u) = \prod_{\Lambda < v \leq u} Z_{\bar{v}, \varepsilon(v)},$$

wobei die Zufallsgrößen $Z_u := (Z_{u,1}, Z_{u,2}, Z_{u,3}, Z_{u,4})$ unabhängig und $\text{Dir}_{\frac{1}{3}, \dots, \frac{1}{3}}$ -verteilt seien.

Lemma 3.8. *Es sei $(Z_1, \dots, Z_4) \sim \text{Dir}_{b_1, \dots, b_4}$ mit $b_1, \dots, b_4 \in \mathbb{R}$, $b_1, \dots, b_4 > 0$. Dann gilt für $a_1, \dots, a_4 \in \mathbb{N}$:*

$$E \left(\prod_{i=1}^4 Z_i^{a_i} \right) = \frac{\Gamma \left(\sum_{i=1}^4 b_i \right)}{\prod_{i=1}^4 \Gamma(b_i)} \times \frac{\prod_{i=1}^4 \Gamma(a_i + b_i)}{\Gamma \left(\sum_{i=1}^4 (a_i + b_i) \right)}.$$

Beweis. Mit Rechnungen wie in Bemerkung 2.6 ergibt sich:

$$\begin{aligned} E \left(\prod_{i=1}^4 Z_i^{a_i} \right) &= \int \left(\prod_{i=1}^4 \alpha_i^{a_i} \right) \mathbf{1}_{\Delta^3}(\alpha_1, \dots, \alpha_4) \frac{\Gamma \left(\sum_{i=1}^4 b_i \right)}{\prod_{i=1}^4 \Gamma(b_i)} \prod_{i=1}^4 \alpha_i^{b_i-1} d\alpha_1 d\alpha_2 d\alpha_3 \\ &= \frac{\Gamma \left(\sum_{i=1}^4 b_i \right)}{\prod_{i=1}^4 \Gamma(b_i)} \times \frac{\prod_{i=1}^4 \Gamma(a_i + b_i)}{\Gamma \left(\sum_{i=1}^4 (a_i + b_i) \right)} \\ &\quad \times \int \mathbf{1}_{\Delta^3}(\alpha_1, \dots, \alpha_4) \frac{\Gamma \left(\sum_{i=1}^4 (a_i + b_i) \right)}{\prod_{i=1}^4 \Gamma(a_i + b_i)} \prod_{i=1}^4 \alpha_i^{a_i+b_i-1} d\alpha_1 d\alpha_2 d\alpha_3. \end{aligned}$$

In der letzten Zeile erkennen wir das Integral über die Dichte der Dirichlet-Verteilung mit Parametern $a_i + b_i$, $i = 1, \dots, 4$, also folgt die Behauptung. \square

Durch die Festlegung

$$h(\mathbf{t}) := EK(\mathbf{t}, S) = \alpha(\mathbf{t})E \left(\prod_{u \in \mathbf{t}} \prod_{\Lambda < v \leq u} Z_{\bar{v}, \varepsilon(v)} \right) = \alpha(\mathbf{t}) \prod_{u \in \mathbf{t}} E \left(\prod_{i=1}^4 Z_{u,i}^{\#\mathbf{t}(u \circ i)} \right)$$

für alle $\mathbf{t} \in \mathbb{T}^{\text{quad}}$ erhalten wir eine harmonische Funktion. Mit Lemma 3.8 folgt:

$$h(\mathbf{t}) = \alpha(\mathbf{t}) \prod_{u \in \mathbf{t}} \left(\frac{\Gamma(\frac{4}{3}) \prod_{i=1}^4 \Gamma(\#\mathbf{t}(u \circ i) + \frac{1}{3})}{\Gamma(\frac{1}{3})^4 \Gamma(\sum_{i=1}^4 \#\mathbf{t}(u \circ i) + \frac{4}{3})} \right).$$

Sei p wieder die Übergangsmatrix der Quadtree-Kette. Wir bestimmen nun die Übergangswahrscheinlichkeiten

$$p^{(h)}(\mathbf{s}, \mathbf{t}) = p(\mathbf{s}, \mathbf{t}) \frac{h(\mathbf{t})}{h(\mathbf{s})}, \quad \mathbf{s}, \mathbf{t} \in \mathbb{T}^{\text{quad}},$$

der h -transformierten Kette. Ähnlich wie im Beweis von Lemma 3.5 reicht es, den Übergang von \mathbf{t} zu $\mathbf{t} \cup \{u\}$ mit $\mathbf{t} \in \mathbb{T}^{\text{quad}}$ und $u \in \partial \mathbf{t}$ zu betrachten. Wir definieren $g : \mathbb{N}_0^4 \rightarrow \mathbb{R}$,

$$g((a_1, \dots, a_4)) := \left(\frac{\Gamma(\frac{4}{3}) \prod_{i=1}^4 \Gamma(a_i + \frac{1}{3})}{\Gamma(\frac{1}{3})^4 \Gamma(\sum_{i=1}^4 a_i + \frac{4}{3})} \right),$$

und stellen fest:

$$\frac{g((a_1, \dots, a_4) + \mathbf{e}_j)}{g((a_1, \dots, a_4))} = \frac{\Gamma(a_j + \frac{1}{3} + 1) \Gamma(\sum_{i=1}^4 a_i + \frac{4}{3})}{\Gamma(\sum_{i=1}^4 a_i + \frac{4}{3} + 1)} = \frac{a_j + \frac{1}{3}}{\sum_{i=1}^4 a_i + \frac{4}{3}} = \frac{3a_j + 1}{3(\sum_{i=1}^4 a_i + 1) + 1},$$

$j = 1, \dots, 4$. Es ist $\sum_{i=1}^4 \#\mathbf{t}(v \circ i) + 1 = \#\mathbf{t}(v)$ für alle $v \in \mathbf{t}$. Außerdem gilt

$$\frac{\alpha(\mathbf{t} \cup \{u\})}{\alpha(\mathbf{t})} = \frac{1}{p(\mathbf{t}, \mathbf{t} \cup \{u\})},$$

siehe Beweis von Lemma 3.5. Wir fassen zusammen:

$$\begin{aligned} p^{(h)}(\mathbf{t}, \mathbf{t} \cup \{u\}) &= p(\mathbf{t}, \mathbf{t} \cup \{u\}) \frac{h(\mathbf{t} \cup \{u\})}{h(\mathbf{t})} \\ &= p(\mathbf{t}, \mathbf{t} \cup \{u\}) \frac{\alpha(\mathbf{t} \cup \{u\})}{\alpha(\mathbf{t})} \prod_{\Lambda < v \leq u} \frac{g((\#\mathbf{t}(\bar{v} \circ 1), \dots, \#\mathbf{t}(\bar{v} \circ 4)) + \mathbf{e}_{\varepsilon(v)})}{g((\#\mathbf{t}(\bar{v} \circ 1), \dots, \#\mathbf{t}(\bar{v} \circ 4)))} \\ &= \prod_{\Lambda < v \leq u} \frac{3\#\mathbf{t}(v) + 1}{3\#\mathbf{t}(\bar{v}) + 1}. \end{aligned} \tag{18}$$

Mit einem Teleskopargument folgt für die Übergangsmatrix $p^{(h)}$:

$$p^{(h)}(\mathbf{t}, \mathbf{t} \cup \{u\}) = \frac{1}{3\#\mathbf{t} + 1} = \frac{1}{\#\partial \mathbf{t}}, \quad \text{falls } \mathbf{t} \in \mathbb{T}^{\text{quad}}, u \in \partial \mathbf{t},$$

und 0 sonst. Der h -Prozess ist also eine kombinatorische Markov-Kette mit Zustandsraum \mathbb{T}^{quad} und folgendem Übergangsmechanismus: Beim Übergang vom Zeitpunkt n zu $n + 1$

wird ein externer Knoten zufällig und gleichverteilt ausgewählt und in einen internen Knoten umgewandelt. Man beachte die Analogie zu den zufälligen binären Bäumen und den zufälligen Harris-Bäumen aus dem Abschnitt 4.1. Aus der Produktdarstellung der Übergangswahrscheinlichkeiten in (18) folgt außerdem: Der h -Prozess kann als ein „trickle-down“-Prozess aufgefasst werden. Dabei wird die Rolle der Weiterleitungsketten durch die Markov-Ketten $\tilde{Y}^u = (\tilde{Y}_n^u)_{n \in \mathbb{N}_0}$ mit Zustandsraum $E = \mathbb{N}_0^4$, Start in $(0, \dots, 0)$ und der Übergangsmatrix $q = (q(\mathbf{x}, \mathbf{y}))_{\mathbf{x}, \mathbf{y} \in \mathbb{N}_0^4}$ mit

$$q(\mathbf{x}, \mathbf{x} + \mathbf{e}_i) = \frac{3x_i + 1}{3(x_1 + \dots + x_4) + 4} \quad \text{für } i = 1, \dots, 4$$

und $q(\mathbf{x}, \mathbf{y}) = 0$ sonst, übernommen. Diese Markov-Kette beschreibt das folgende Pólya'sche Urnenmodell mit vier Farben: Am Anfang enthalte die Urne jeweils eine Kugel jeder Farbe. In jedem Schritt wird eine Kugel zufällig und gleichverteilt gezogen und zusammen mit drei weiteren Kugeln derselben Farbe zurückgelegt. Der Zustand $x \in \mathbb{N}_0^4$ bedeutet also, dass die Anzahl aller Kugeln der i -ten Farbe in der Urne $3x_i + 1$ ist.

Bemerkung 3.9. Wir erinnern daran, dass für die Komponentenprozesse Y^u in der „trickle-down“-Darstellung der Quadtree-Kette gilt: $(Y_n^u + (1, 0, 1, 0))'_{n \in \mathbb{N}_0}$ ist eine Mendel-Urne. Mit Lemma 3.8 und der Vorgehensweise wie in Bemerkung 2.6 kann man zeigen: \tilde{Y}^u ist eine h -Transformierte von Y^u und zwar zu der strikt positiven harmonischen Funktion

$$\mathbb{N}_0^4 \rightarrow \mathbb{R}, \quad (x_1, \dots, x_4) \mapsto \frac{(\sum_{i=1}^4 x_i)!^2}{\prod_{(i,j) \in D} (x_i + x_j)!} g((x_1, \dots, x_4)),$$

mit D wie im Lemma 3.3. Beide Markov-Ketten haben folglich dieselbe Doob-Martin-Kompaktifizierung. Außerdem ist $\text{Dir}_{\frac{1}{3}, \dots, \frac{1}{3}}$ nach der Diskussion im Abschnitt 1.4.3 die Austrittsverteilung von \tilde{Y}^u . Ändert man die Parameter der Dirichlet-Verteilung in der Definition des zufälligen Wahrscheinlichkeitsmaßes S , so erhält man zufällige quaternäre „trickle-down“-Bäume zu Vier-Farben-Pólya-Urnen mit anderen Anfangsbelegungen und Übergangsmechanismen.

3.5 Ausblick: Pfadlängen in quaternären Bäumen

Die interne Pfadlänge eines quaternären Baumes \mathbf{t} ist definiert als die Summe der Tiefen aller internen Knoten:

$$\text{IPL}(\mathbf{t}) = \sum_{u \in \mathbf{t}} |u|.$$

Analog ist die externe Pfadlänge die Summe der Tiefen der externen Knoten:

$$\text{EPL}(\mathbf{t}) = \sum_{u \in \partial \mathbf{t}} |u|.$$

Wie im Fall binärer Bäume (siehe [13, S.12]) kann man für $\mathbf{t} \in \mathbb{T}^{\text{quad}}$ zeigen:

$$\text{IPL}(\mathbf{t}) = \sum_{u \in \mathbf{t}} \#\mathbf{t}(u) - \#\mathbf{t}.$$

Es besteht folgender Zusammenhang zwischen der internen und externen Pfadlänge:

Lemma 3.10. Für jeden quaternären Baum \mathbf{t} mit n Knoten gilt:

$$\text{EPL}(\mathbf{t}) = 3 \text{IPL}(\mathbf{t}) + 4n$$

Beweis. Wie im Fall binärer Bäume verwenden wir das Induktionsprinzip, vergleiche [20, S.83]. Die Behauptung ist offensichtlich gültig für $n = 1$: Der einzige quaternäre Baum mit einem Knoten ist der Baum bestehend nur aus dem Wurzelknoten. Dieser hat die interne Pfadlänge 0 und externe Pfadlänge 4.

Angenommen, die Behauptung gelte für alle quaternären Bäume mit n Knoten. Zu zeigen ist: Für jedes $\mathbf{t} \in \mathbb{T}^{\text{quad}}$ mit $|\mathbf{t}| = n + 1$ gilt: $\text{EPL}(\mathbf{t}) = 3 \text{IPL}(\mathbf{t}) + 4(n + 1)$. Zu jedem solchen \mathbf{t} existiert ein $\mathbf{s} \in \mathbb{T}^{\text{quad}}$ mit $|\mathbf{s}| = n$ und ein $u \in \partial\mathbf{s}$, sodass gilt: $\mathbf{t} = \mathbf{s} \cup \{u\}$. Einen Baum \mathbf{s} mit diesen Eigenschaften erreicht man durch das Löschen eines beliebigen internen Knotens von \mathbf{t} , welcher als direkte Nachfolger nur externe Knoten hat. Es gilt:

$$\text{IPL}(\mathbf{t}) = \text{IPL}(\mathbf{s}) + |u| \tag{19}$$

$$\text{EPL}(\mathbf{t}) = \text{EPL}(\mathbf{s}) - |u| + 4(|u| + 1) \tag{20}$$

Damit gilt unter Verwendung der Induktionsvoraussetzung (IV):

$$\begin{aligned} \text{EPL}(\mathbf{t}) &= \text{EPL}(\mathbf{s}) + 3|u| + 4 && \text{(nach (20))} \\ &= 3 \text{IPL}(\mathbf{s}) + 4n + 3|u| + 4 && \text{(nach IV)} \\ &= 3(\text{IPL}(\mathbf{s}) + |u|) + 4(n + 1) \\ &= 3 \text{IPL}(\mathbf{t}) + 4(n + 1) && \text{(nach (19))} \quad \square \end{aligned}$$

R. Neininger und L. Rüschemdorf haben 1999 in [23] die Konvergenz in Verteilung der skalierten internen Pfadlänge

$$Y_n = \frac{\text{IPL}(X_n) - E(\text{IPL}(X_n))}{n}$$

zufälliger quaternärer Bäume gezeigt. Die explizite Formel für den Erwartungswert findet sich in [10, Satz 1]¹¹:

$$E(\text{IPL}(X_n)) = nH(n) + \frac{1}{3}H(n) - \frac{n+1}{6} - 1.$$

Wir werden nun zeigen, dass die Doob-Martin-Kompaktifizierung der Quadtree-Kette ein möglicher Ausgangspunkt für die Erforschung der internen Pfadlänge ist. Wir definieren $\psi(0) := 0$ und

$$\psi(n) := nH(n) + \frac{1}{3}H(n) - \frac{n+1}{6} - 1.$$

Für $u \in X_n$ setzen wir

$$C_n(u) := 1 + \frac{\sum_{i=1}^4 \psi(\#X_n(u \circ i)) - \psi(\#X_n(u))}{\#X_n(u)}.$$

¹¹Hierbei ist zu beachten, dass die Definition der internen Pfadlänge in [10] sich von unserer um n unterscheidet.

Mit diesen Bezeichnungen können wir nun Y_n wie folgt darstellen:

$$Y_n = -1 + \sum_{u \in X_n} \frac{\#X_n(u)}{n} C_n(u),$$

denn es gilt

$$\begin{aligned} \sum_{u \in X_n} \frac{\#X_n(u)}{n} C_n(u) &= \sum_{u \in X_n} \frac{\#X_n(u)}{n} + \frac{1}{n} \sum_{u \in X_n} \left(\sum_{i=1}^4 \psi(\#X_n(u \circ i)) - \psi(\#X_n(u)) \right) \\ &= \sum_{u \in X_n} \frac{\#X_n(u)}{n} - \frac{(\psi(\#X_n(\Lambda)))}{n} \\ &= \frac{\text{IPL}(X_n) - \psi(n)}{n} + 1. \end{aligned}$$

Es seien für $u \in \{1, \dots, 4\}^*$

$$\xi_u = (\xi_{u,1}, \dots, \xi_{u,4}) := \left(\frac{X_\infty(A_{u1})}{X_\infty(A_u)}, \dots, \frac{X_\infty(A_{u4})}{X_\infty(A_u)} \right).$$

Dann ist

$$X_\infty(A_u) = \prod_{\Lambda < v \leq u} \xi_{\bar{v}, \varepsilon(v)}.$$

Aus dem Abschnitt über die Doob-Martin-Kompaktifizierung der Quadtree-Kette ist bekannt:

$$\lim_{n \rightarrow \infty} \frac{\#X_n(u)}{n} = X_\infty(A_u)$$

fast sicher. Außerdem kann man mit der Vorgehensweise wie im Lemma 3.1 in [23] zeigen:

$$\lim_{n \rightarrow \infty} C_n(u) = C(\xi_u)$$

fast sicher, wobei

$$C : \Delta^3 \rightarrow \mathbb{R} \quad C(\eta) := 1 + \sum_{i=1}^4 \eta_i \ln \eta_i$$

mit $0 \ln 0 := 0$ sei. Insgesamt führen diese Beobachtungen auf die *Vermutung*, dass fast sichere Konvergenz von Y_n gegen $Y_\infty - 1$ mit

$$Y_\infty = \sum_{u \in \{1, \dots, 4\}^*} X_\infty(A_u) C(\xi_u)$$

vorliegt.

Eine Möglichkeit, diese Vermutung zu überprüfen, bietet die Methode der Martingalprojektion aus [13]. Man zeigt als Erstes, dass der Grenzwert Y_∞ der Partialsummen

$$\sum_{\substack{u \in \{1, \dots, 4\}^* \\ |u| \leq k}} X_\infty(A_u) C(\xi_u)$$

fast sicher und in L^2 existiert. Die Zufallsgröße Y_∞ wird auf die natürliche Filtration $(\mathcal{F}_n)_{n \in \mathbb{N}}$, $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$ projiziert. Dann ist $(Z_n, \mathcal{F}_n)_{n \in \mathbb{N}}$ mit $Z_n := E[Y_\infty | \mathcal{F}_n]$ ein L^2 -beschränktes Martingal, was die Konvergenz von Z_n gegen Y_∞ fast sicher und in L^2 mit $n \rightarrow \infty$ bedeutet. Man sieht leicht, dass $X_\infty(A_u)$ und $C(\xi_u)$ bedingt unabhängig gegeben \mathcal{F}_n sind. Kann man außerdem zeigen, dass für jedes feste n die Summe über $u \in \{1, \dots, 4\}^*$ und $E[\cdot | \mathcal{F}_n]$ vertauscht werden kann, so erhält man

$$Z_n = E[Y_\infty | \mathcal{F}_n] = \sum_{u \in \{1, \dots, 4\}^*} E[X_\infty(A_u) | \mathcal{F}_n] E[C(\xi_u) | \mathcal{F}_n].$$

In der vorliegenden Situation erhalten wir nach einigen Rechnungen:

$$\begin{aligned} E[X_\infty(A_u) | \mathcal{F}_n] &= \prod_{\Lambda < v \leq u} E[\xi_{\bar{v}, \varepsilon(v)} | \mathcal{F}_n] \\ &= \prod_{\substack{\Lambda < v \leq u \\ \varepsilon(v)=1}} \frac{(\#X_n(\bar{v} \circ 1) + \#X_n(\bar{v} \circ 2) + 1)(\#X_n(\bar{v} \circ 1) + \#X_n(\bar{v} \circ 4) + 1)}{(\#X_n(\bar{v}) + 1)^2} \\ &\quad \times \prod_{\substack{\Lambda < v \leq u \\ \varepsilon(v)=2}} \frac{(\#X_n(\bar{v} \circ 1) + \#X_n(\bar{v} \circ 2) + 1)(\#X_n(\bar{v} \circ 2) + \#X_n(\bar{v} \circ 3) + 1)}{(\#X_n(\bar{v}) + 1)^2} \\ &\quad \times \prod_{\substack{\Lambda < v \leq u \\ \varepsilon(v)=3}} \frac{(\#X_n(\bar{v} \circ 3) + \#X_n(\bar{v} \circ 4) + 1)(\#X_n(\bar{v} \circ 2) + \#X_n(\bar{v} \circ 3) + 1)}{(\#X_n(\bar{v}) + 1)^2} \\ &\quad \times \prod_{\substack{\Lambda < v \leq u \\ \varepsilon(v)=4}} \frac{(\#X_n(\bar{v} \circ 3) + \#X_n(\bar{v} \circ 4) + 1)(\#X_n(\bar{v} \circ 1) + \#X_n(\bar{v} \circ 4) + 1)}{(\#X_n(\bar{v}) + 1)^2} \end{aligned}$$

und

$$\begin{aligned} E[C(\xi_u) | \mathcal{F}_n] &= 1 - 2H(\#X_n(\bar{u}) + 1) \\ &\quad + \frac{1}{\#X_n(\bar{u}) + 1} \sum_{(i,j) \in D} \left((\#X_n(u \circ i) + \#X_n(u \circ j) + 1) \right. \\ &\quad \left. \times H(\#X_n(u \circ i) + \#X_n(u \circ j) + 1) \right), \end{aligned}$$

mit $D = \{(1, 2), (1, 4), (2, 3), (3, 4)\}$ wie im Lemma 3.3 und den harmonischen Zahlen $H(0) := 0$, $H(n) := \sum_{i=1}^n \frac{1}{i}$, $n \in \mathbb{N}$.

Es bleibt die *offene Frage*, ob $Y_n - Z_n$ asymptotisch vernachlässigbar ist.

4 Zufällige rekursive Bäume

Dieser Abschnitt befasst sich mit dem „random recursive tree“- Prozess und der dazu verwandten Markov-Kette mit Werten in der kombinatorischen Familie der Harris-Bäume. Zufällige Harris-Bäume entstehen als Ergebnis eines stochastischen Wachstumsprozesses oder als Output eines bestimmten deterministischen Algorithmus bei zufälligem Input. Nach der Einführung der grundlegenden Begriffe werden Resultate zur Doob-Martin Kompaktifizierung zu der Harris-Kette aus der Publikation „trickle-down processes and their boundaries“ [8] von S. N. Evans, R. Grübel und A. Wakolbinger vorgestellt. Anschließend verwenden wir den algorithmischen Hintergrund für den alternativen Beweis der Konvergenzresultate aus [8]. Zusätzlich erhalten wir eine Darstellung des Limesobjektes mit Hilfe der Inputfolge. Schließlich wird die Vorgehensweise zum Beweis der fast sicheren Konvergenz verschiedener Baumfunktionale aus dem Artikel „search trees: metric aspects and strong limit theorems“ [13] von R. Grübel in der Situation zufälliger Harris-Bäume für die interne Pfadlänge, horizontale Pfadlänge und den Wiener-Index angewendet.

4.1 Grundbegriffe und Bezeichnungen

Aufbauend auf dem Übersichtsartikel von Smythe und Mahmoud [27] werden im Folgenden Grundlagen zu den zufälligen rekursiven Bäumen bereitgestellt. Bei der Beschreibung der Ulam-Harris-Bäume halten wir uns an [8]. In der Graphentheorie wird ein Baum als ein Graph ohne Kreise definiert.

Definition 4.1. *Ein Baum mit der Knotenmenge $\{1, \dots, n\}$ wird als rekursiver Baum bezeichnet, falls der Knoten 1 als Wurzel ausgezeichnet ist und für jeden Knoten k , $2 \leq k \leq n$, der eindeutige Pfad von 1 nach k eine monotone Folge bildet.*

Bemerkung 4.2. Bei rekursiven Bäumen spielt die Anordnung der direkten Nachfolger eines Knotens keine Rolle. Jede Anordnung repräsentiert denselben Baum, siehe Abbildung 4. Es hat sich jedoch die folgende bildliche Darstellungsweise durchgesetzt: die Wurzel wird wie üblich oben gezeichnet, die direkten Nachfolger eines Knotens werden von links nach rechts in aufsteigender Reihenfolge sortiert (Abbildung 4, links). Dadurch wird eine natürliche Einbettung in die Ebene definiert.

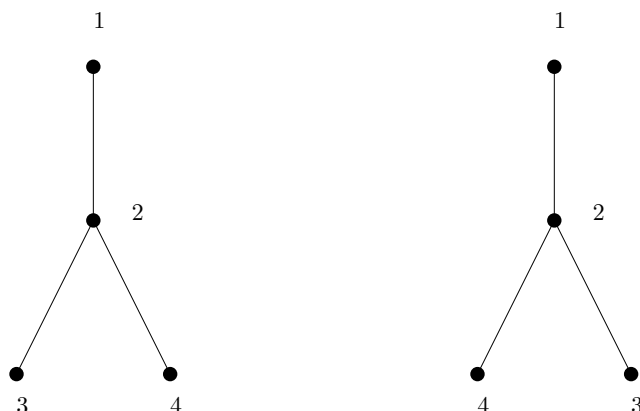


Abbildung 4: Zwei mögliche Darstellungen des rekursiven Baumes mit 4 Knoten und der Kantenmenge $\{(1, 2), (2, 3), (2, 4)\}$.

Unter einem „random recursive tree“ (RRT) - Prozess verstehen wir eine Markov-Kette $(Y_n)_{n \in \mathbb{N}}$ mit Werten in der kombinatorischen Familie der rekursiven Bäume und dem folgenden Übergangsmechanismus: Der Baum Y_1 besteht nur aus dem Wurzelknoten 1. Beim Übergang von n zu $n + 1$ wird einer der n in Y_n vorhandenen Knoten zufällig und gleichverteilt als direkter Vorgänger des neuen Knotens $n + 1$ ausgewählt. Man sieht leicht, dass Y_n auf der Menge der $(n - 1)!$ rekursiver Bäume mit n Knoten gleichverteilt ist.

Wie im Falle der quaternären Bäume, ist bei der Definition der Ulam-Harris-Bäume die Sprache der Theoretischen Informatik hilfreich.

Definition 4.3. *Ein Ulam-Harris-Baum ist eine präfixstabile endliche Teilmenge \mathbf{t} von \mathbb{N}^* mit folgender Eigenschaft: ist $u = u_1 \dots u_n \in \mathbf{t}$, so auch alle $u_1 \dots u_{n-1} \circ i \in \mathbf{t}$ mit $1 \leq i \leq u_n - 1$.*

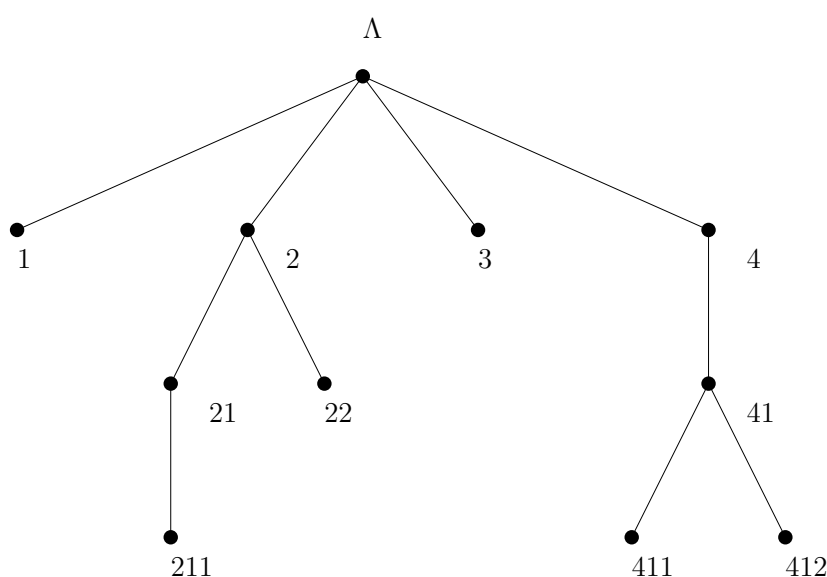


Abbildung 5: Der Ulam-Harris-Baum $\{\Lambda, 1, 2, 3, 4, 21, 22, 41, 211, 411, 412\}$.

Eine Wortmenge \mathbf{t} wie in der Definition 4.3 lässt sich in kanonischer Weise als Baum aus der Sicht der Graphentheorie interpretieren, indem man die Elemente von \mathbf{t} als Knoten auffasst und die Kantenmenge folgendermaßen definiert: Die Knoten u und v , $v \neq \Lambda$, werden durch eine gerichtete Kante (u, v) verbunden, falls $u = \bar{v}$ gilt. Viele graphentheoretische Begriffe, die für die quaternären Bäume festgelegt wurden, lassen sich leicht übernehmen. Die Definition der externen Knoten bedarf hier einer leichten Modifikation. Wir definieren einen Knoten $u \notin \mathbf{t}$ als externen Knoten eines Ulam-Harris-Baumes \mathbf{t} , falls $\bar{u} \in \mathbf{t}$ und $\mathbf{t} \cup \{u\}$ wieder ein Ulam-Harris-Baum ist, und schreiben $\partial \mathbf{t}$ für die Menge der externen Knoten von \mathbf{t} . Mit \mathbb{H}_n bezeichnen wir die Menge der Ulam-Harris-Bäume mit n Knoten und setzen $\mathbb{H} := \bigcup_{n=1}^{\infty} \mathbb{H}_n$.

Zufällige Ulam-Harris-Bäume können einerseits wie zufällige rekursive Bäume als Ergebnis eines stochastischen Wachstumsprozesses betrachtet werden. Andererseits entstehen sie als Output eines deterministischen Algorithmus bei zufälligem Input. Bei der algorithmischen Konstruktion benötigt man wieder den Begriff der bewerteten Bäume. Ein bewerteter Ulam-Harris-Baum ist ein Paar (\mathbf{t}, φ) mit $\mathbf{t} \in \mathbb{H}$ und einer Abbildung $\varphi : \mathbf{t} \rightarrow [0, 1]$. Zu einer Folge $(x_i)_{i \in \mathbb{N}}$ paarweise verschiedener $x_i \in [0, 1]$ kann eine Folge bewerteter Ulam-Harris-Bäume wie folgt konstruiert werden:

- Der erste Baum besteht nur aus dem Wurzelknoten bewertet mit 0.
- Im zweiten Schritt wird der Baum um einen direkten Nachfolger des Wurzelknotens mit der Bewertung x_1 ergänzt.
- Beim Übergang vom Zeitpunkt $n \geq 2$ zu $n + 1$ wird ein neuer Knoten hinzugefügt und mit x_n bewertet. Ist $x_n < x_1, \dots, x_{n-1}$, so wird der neue Knoten ein direkter Nachfolger des Wurzelknotens, sonst des Knotens mit der Bewertung $\max \{x_i : 1 \leq i \leq n - 1, x_i < x_n\}$.

Im n -ten Schritt erhält man also zu $n - 1$ Eingangsdaten einen bewerteten Ulam-Harris-Baum mit n Knoten. Der Wurzelknoten hat hier eine Sonderrolle: Die Bewertung wurde willkürlich gewählt und ist für die probabilistische Analyse nicht weiter interessant. Die Abbildung 6 zeigt ein Beispiel zu diesem Algorithmus mit 10 Eingangsdaten. Man beachte, dass durch diese Prozedur Bäume (\mathbf{t}, φ) mit pfadmonotonen Bewertungen erzeugt werden: für Knoten $u, v \in \mathbf{t}$ mit $u \leq v$ gilt $\varphi(u) \leq \varphi(v)$. Wie bereits erwähnt, ist der Algorithmus deterministisch. Wendet man diesen Algorithmus auf eine Folge $(\eta_n)_{n \in \mathbb{N}}$ unabhängiger und jeweils auf dem Einheitsintervall $[0, 1]$ gleichverteilter Zufallsvariablen an und entfernt die Bewertungen, so erhält man eine kombinatorische Markov-Kette $(X_n)_{n \in \mathbb{N}}$ mit \mathbb{H} als Zustandsraum und Start in $X_1 = \{\Lambda\}$. Beim Übergang von n zu $n + 1$ bestimmt der Rang von x_n unter x_1, \dots, x_{n-1} die Wahl des externen Knotens. Aus der bekannten Tatsache, dass der Rang von η_n unter $\eta_1, \dots, \eta_{n-1}$ gleichverteilt auf $\{1, \dots, n\}$ ist, ergibt sich der folgende Übergangsmechanismus: X_{n+1} wird aus X_n erhalten, indem einer der externen Knoten zufällig und gleichverteilt ausgewählt und in einen internen Knoten umgewandelt wird, in Zeichen:

$$\mathbb{P}(X_{n+1} = \mathbf{t} | X_n = \mathbf{s}) = \begin{cases} \frac{1}{n} & \text{falls } \mathbf{s} \in \mathbb{H}_n, \mathbf{t} \in \mathbb{H}_{n+1} \text{ und } \mathbf{s} \subset \mathbf{t}, \\ 0 & \text{sonst.} \end{cases}$$

Eine solche Markov-Kette bezeichnen wir im Folgenden als Harris-Kette.

Die natürliche Einbettung aus der Bemerkung 4.2 führt auf einen Zusammenhang zwischen rekursiven Bäumen und Ulam-Harris-Bäumen. Genauer wird dabei eine Abbildung ψ von der Menge der rekursiven Bäume in die Menge der Ulam-Harris-Bäume definiert (siehe Beispiel in Abbildung 7), indem die Knoten eines rekursiven Baumes \mathbf{y} auf Wörter aus \mathbb{N}^* abgebildet werden. Dem Wurzelknoten 1 wird das leere Wort Λ zugeordnet. Sei nun $j \in \mathbf{y}$ mit $j \neq 1$ und i direkter Vorgänger von j , das heißt es existiere eine Kante (i, j) in der Kantenmenge von \mathbf{y} . Sei u das zu i zugehörige Wort und r der Rang von j unter den direkten Nachfolgern von i . Dann wird der Knoten j auf das Wort $u \circ r$ abgebildet. Wir stellen fest, dass $(\psi(Y_n))_{n \in \mathbb{N}}$ wieder eine Harris-Kette ist, in dem Sinne, dass die Verteilungen von $(X_n)_{n \in \mathbb{N}}$ und $(\psi(Y_n))_{n \in \mathbb{N}}$ übereinstimmen. Betrachten wir dazu die Übergangsmechanismen. Beide Ketten starten mit dem allein aus dem Wurzelknoten bestehenden Baum. Außerdem haben wir

$$\mathbb{P}(X_{n+1} = \mathbf{t} | X_n = \mathbf{s}) = \frac{1}{n} = \mathbb{P}(\psi(Y_{n+1}) = \mathbf{t} | \psi(Y_n) = \mathbf{s}),$$

falls $\mathbf{s} \subset \mathbf{t}$, $\mathbf{s} \in \mathbb{H}_n$, $\mathbf{t} \in \mathbb{H}_{n+1}$, denn

$$\begin{aligned} \mathbb{P}(\psi(Y_{n+1}) = \mathbf{t} | \psi(Y_n) = \mathbf{s}) &= \frac{(n-1)!}{\#\{\psi^{-1}(\mathbf{s})\}} \mathbb{P}(\psi(Y_{n+1}) = \mathbf{t}, \psi(Y_n) = \mathbf{s}) \\ &= \frac{(n-1)!}{\#\{\psi^{-1}(\mathbf{s})\}} \left(\#\{\psi^{-1}(\mathbf{s})\} \frac{1}{(n-1)!} \frac{1}{n} \right) = \frac{1}{n}. \end{aligned}$$

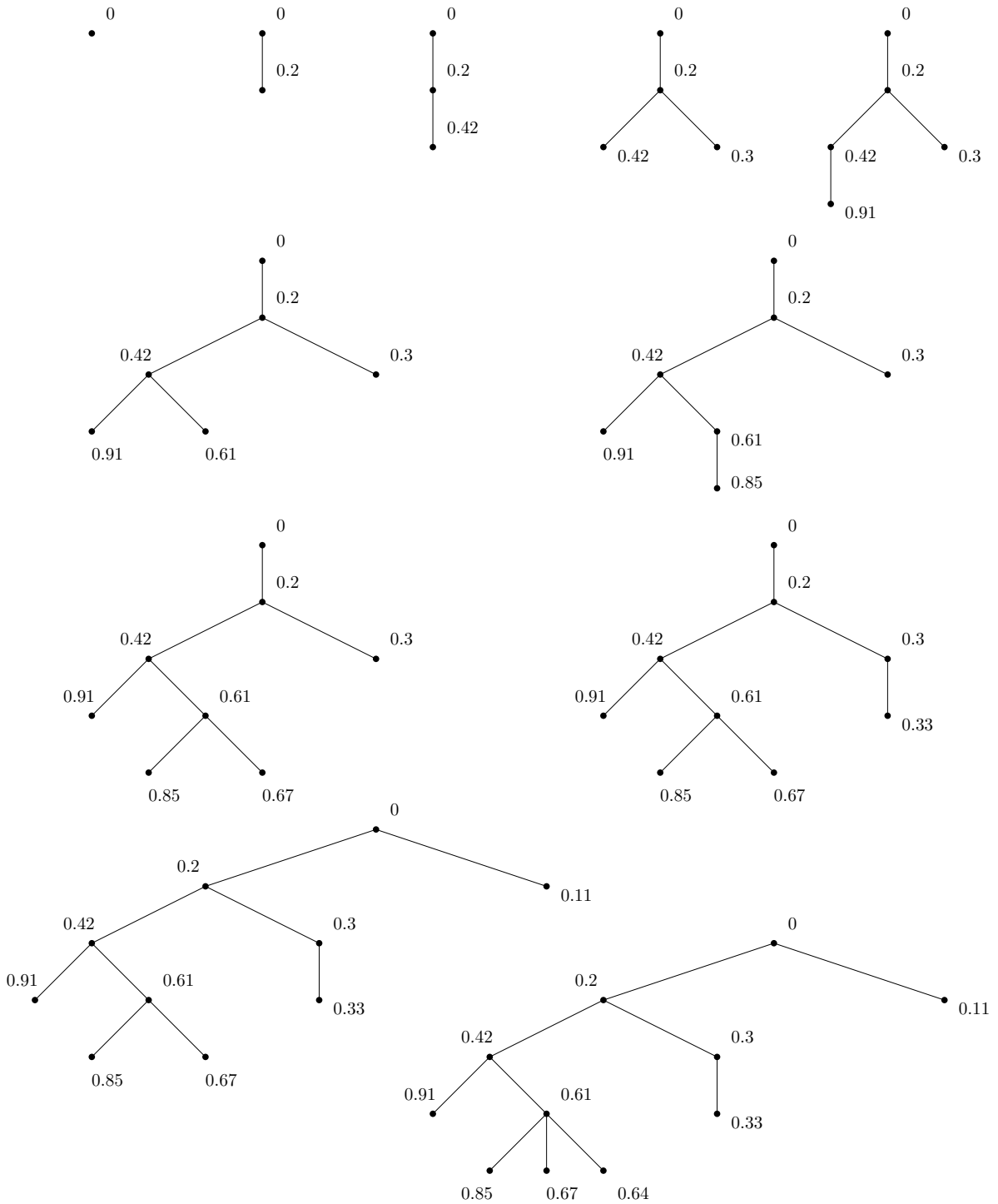


Abbildung 6: Bewertete Ulam-Harris-Bäume zu den Daten 0.2, 0.42, 0.3, 0.91, 0.61, 0.85, 0.67, 0.33, 0.11, 0.64. Die Beschriftung der Knoten mit den Wörtern aus \mathbb{N}^* wurde einfachheitshalber weggelassen.

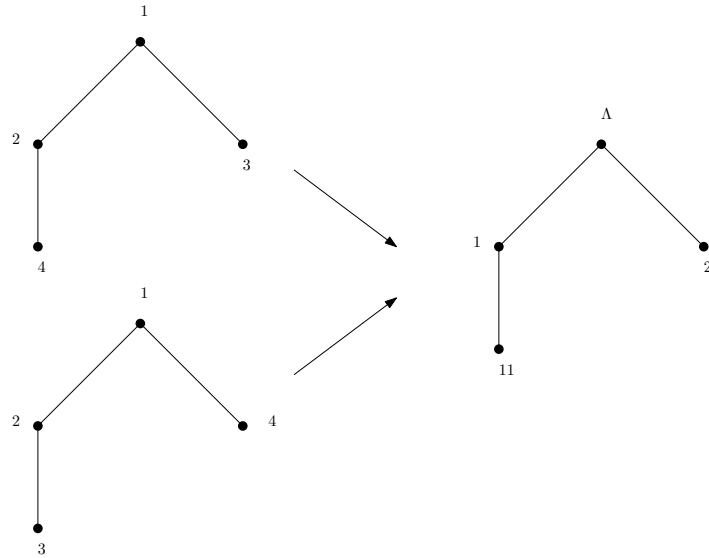


Abbildung 7: Zwei verschiedene rekursive Bäume werden auf denselben Ulam-Harris-Baum abgebildet.

Dabei haben wir verwendet, dass ein rekursiver Baum \mathbf{y} mit n Knoten und $\mathbf{s} = \psi(\mathbf{y})$ sich nur auf eine Weise zu einem rekursiven Baum \mathbf{z} mit $n + 1$ Knoten und $\mathbf{t} = \psi(\mathbf{z})$ ergänzen lässt, sodass $\mathbf{y} \subset \mathbf{z}$ gilt.

Bemerkung 4.4. Der obige Zusammenhang zwischen dem RRT-Prozess und der Harris-Kette ist insofern von Bedeutung, dass einige Baumfunktionale wie die interne Pfadlänge und der Wiener-Index von dem rekursiven Baum nur über den zugehörigen Ulam-Harris-Baum abhängen.

In [8] wurde eine Methode zur Untersuchung des asymptotischen Verhaltens der „trickle-down“-Prozesse mit Hilfe der Martinrandtheorie vorgestellt. Die „trickle-down“-Prozesse sind eine Klasse transienter Markov-Ketten, die die Harris-Kette als Spezialfall enthält. Betrachtet man die Entwicklung der Anzahl der Teilbäume und der Teilbaumgrößen an dem Wurzelknoten, so stellt man fest, dass diese mit dem China-Restaurant-Prozess zusammenhängt (siehe dazu Bemerkung 4.5 aus dem nächsten Abschnitt). Daraus ergibt sich, dass der China-Restaurant-Prozess bei der „trickle-down“-Konstruktion als Markov-Kette der Weiterleitungsanweisungen dient. Der Zusammenhang zwischen dem China-Restaurant-Prozess und der Harris-Kette entspricht dem Zusammenhang zwischen der Mendel-Urne und der Quadtree-Kette, vergleiche Abschnitt 3.2.

4.2 China-Restaurant-Prozess

Wir betrachten eine Möglichkeit zur Konstruktion von zufälligen Partitionen von $\{1, \dots, n\}$, die unter dem Namen China-Restaurant-Prozess¹² bekannt ist. Diese Konstruktion funktioniert wie folgt¹³. Das Restaurant sei anfangs leer und enthalte unendlich viele nummerierte Tische, wobei an jedem Tisch unbegrenzt viel Platz zur Verfügung stehe. Die Gäste treffen nacheinander ein. Der erste Gast setzt sich an den ersten Tisch. Seien nun $n \geq 1$ Gäste bereits eingetroffen und k Tische mit jeweils n_1, \dots, n_k Personen,

¹²Bei der Eindeutschung des Begriffes „chinese restaurant process“ folgen wir [18, S.542]

¹³Beschreibung in Anlehnung an [8, §6.2], [18, S.542] und [24]

$n = \sum_{i=1}^k n_i$, besetzt. Der $n + 1$ -te Gast setzt sich an den nächsten freien Tisch (dieser hat die Nummer $k + 1$) mit Wahrscheinlichkeit $\frac{1}{n+1}$, oder an den j -ten besetzten Tisch, $j = 1, \dots, k$, mit Wahrscheinlichkeit $\frac{n_j}{n+1}$. Die Tische werden als Blöcke einer zufälligen Partition interpretiert. Formal handelt es sich dabei um eine Markov-Kette mit dem Zustandsraum $E \subset \mathbb{N}_0^{\mathbb{N}}$,

$$E = \{(0, 0, \dots)\} \cup \{a = (a_1, \dots, a_k, 0, 0, \dots) : k \geq 1 \text{ und } a_i > 0 \text{ für } 1 \leq i \leq k\}, \quad (21)$$

Start in $(0, 0, \dots)$ und Übergangsmatrix $p = (p(a, b))_{a, b \in E}$:

$$p(a, b) = \begin{cases} \frac{\frac{1}{m}}{1 + \sum_{i=1}^m a_i} & \text{falls } b = (a_1, \dots, a_k, 1, 0, 0, \dots) \\ \frac{\frac{a_j}{m}}{1 + \sum_{i=1}^m a_i} & \text{falls } b = (a_1, \dots, a_j + 1, \dots, a_k, 0, 0, \dots), \end{cases}$$

falls a wie in (21).

Bemerkung 4.5. Die folgende Beobachtung aus [8, §6.2] stellt die Verbindung zwischen zufälligen Ulam-Harris-Bäumen und dem China-Restaurant-Prozess her und erlaubt die Anwendung der „trickle-down“-Konstruktion. Seien n_1, \dots, n_k , $n_1 + \dots + n_k = n - 1$, die Teilbaumgrößen an dem Wurzelknoten von X_n . Wie bereits erwähnt, wird beim Übergang von X_n zu X_{n+1} ein externer Knoten zufällig und gleichverteilt ausgewählt, und in einen internen Knoten umgewandelt. Der $n + 1$ -te Knoten wird also mit Wahrscheinlichkeit $\frac{n_j}{n}$ zum j -ten Teilbaum hinzugefügt, $1 \leq j \leq k$, oder bildet mit Wahrscheinlichkeit $\frac{1}{n}$ einen neuen Teilbaum. Die Anzahl der Teilbäume und die Teilbaumgrößen an dem Wurzelknoten entwickeln sich also gemäß dem Übergangsmechanismus des China-Restaurant-Prozesses.

Ausgehend von [8, §6.2] kann der Martin-Rand der China-Restaurant-Kette als der unendlichdimensionale Wahrscheinlichkeitssimplex

$$\Delta^\infty := \left\{ (\rho_n)_{n \in \mathbb{N}} : \rho_n \in [0, 1] \forall n \in \mathbb{N}, \sum_{n=1}^{\infty} \rho_n \leq 1 \right\}$$

identifiziert werden. Eine Folge $(b_i)_{i \in \mathbb{N}} = (b_{i,1}, b_{i,2}, \dots)_{i \in \mathbb{N}}$ aus E mit $\lim_{i \rightarrow \infty} \sum_{j=1}^{\infty} b_{i,j} = \infty$ konvergiert genau dann in der Doob-Martin Kompaktifizierung gegen ein $\rho \in \Delta^\infty$, wenn für jedes feste $k \in \mathbb{N}$ gilt:

$$\lim_{i \rightarrow \infty} \frac{b_{i,k}}{\sum_{j=1}^{\infty} b_{i,j}} = \rho_k.$$

Die Austrittsverteilung stimmt überein mit der Verteilung des „stick breaking“-Prozesses, betrachtet als Δ^∞ -wertige Zufallsgröße

$$(U_1, (1 - U_1)U_2, (1 - U_1)(1 - U_2)U_3, \dots),$$

wobei U_i unabhängig und jeweils $\text{unif}(0, 1)$ verteilt sind. Diese Verteilung wird als Griffiths-Engen-McCloskey (GEM)-Verteilung bezeichnet, siehe dazu auch die Überlegungen zur Verteilungskonvergenz der relativen Blockgrößen beim China-Restaurant-Prozess im Abschnitt 3.1. von [24].

Für spätere Berechnungen zu zufälligen rekursiven Bäumen bestimmen wir nun die Austrittsverteilung beim Start in $a = (a_1, \dots, a_k, 0, 0, \dots)$.

Lemma 4.6. Die Austrittsverteilung beim Start in $a = (a_1, \dots, a_k, 0, 0, \dots)$ stimmt überein mit der Verteilung des modifizierten „stick breaking“-Prozesses:

$$(U_1, (1 - U_1)U_2, (1 - U_1)(1 - U_2)U_3, \dots),$$

mit unabhängigen U_i und

$$\begin{aligned} U_i &\sim \text{Beta} \left(a_i, 1 + \sum_{j=i+1}^k a_j \right), \text{ für } i < k, \\ U_k &\sim \text{Beta} (a_k, 1), \\ U_i &\sim \text{Beta} (1, 1) = \text{unif}(0, 1), \text{ für } i > k. \end{aligned} \tag{22}$$

Beweis. Sei $(Y_n)_{n \in \mathbb{N}}$ eine China-Restaurant-Kette mit dem Einpunktmaß in a als Startverteilung und \mathbb{P}_a das zugehörige Hintergrundwahrscheinlichkeitsmaß. Sei Y_∞ der fast sichere Limes. Nach dem Satz über die Charakterisierung der Austrittsverteilung gilt für Borel-Mengen A :

$$\mathbb{P}_a \left(\lim_{n \rightarrow \infty} Y_n = Y_\infty \in A \right) = \int_A K(a, \rho) \mu_1(d\rho),$$

wobei K der erweiterte Martin-Kern ist und μ_1 das auf $\partial_{\min} E$ konzentrierte darstellende Maß zur harmonischen Funktion $h \equiv 1$. Aus [8, S.37] ist bekannt:

$$\begin{aligned} K(a, \rho) &= \frac{\left(\sum_{j=1}^k a_j \right)!}{\prod_{j=1}^k (a_j - 1)!} \rho_1^{a_1-1} \dots \rho_k^{a_k-1} \\ &\quad \times (1 - \rho_1)(1 - \rho_1 - \rho_2) \dots (1 - \rho_1 - \rho_2 - \dots - \rho_{k-1}). \end{aligned}$$

Das Maß $\mu_1 = \nu^T$ ist das Bildmaß von $\nu = \text{unif}(0, 1)^{\otimes \mathbb{N}}$ unter der Transformation

$$T : [0, 1]^{\otimes \mathbb{N}} \rightarrow \Delta^\infty, \quad (u_1, u_2, \dots) \mapsto (u_1, (1 - u_1)u_2, (1 - u_1)(1 - u_2)u_3, \dots).$$

Also folgt:

$$\begin{aligned} \mathbb{P}_a (Y_\infty \in A) &= \int_A K(a, \rho) \nu^T(d\rho) \\ &= \int_{T^{-1}(A)} K(a, T(u)) \nu(du) \\ &= \int_{T^{-1}(A)} \frac{\left(\sum_{j=1}^k a_j \right)!}{\prod_{j=1}^k (a_j - 1)!} \\ &\quad \times u_1^{a_1-1} ((1 - u_1)u_2)^{a_2-1} \dots ((1 - u_1) \dots (1 - u_{k-1})u_k)^{a_k-1} \\ &\quad \times (1 - u_1)(1 - u_1 - (1 - u_1)u_2) \times \dots \\ &\quad \times (1 - u_1 - (1 - u_1)u_2 - \dots - (1 - u_1) \dots (1 - u_{k-2})u_{k-1}) \nu(du). \end{aligned}$$

Der erste Term lässt sich mit einem Teleskopargument darstellen als

$$\frac{\left(\sum_{j=1}^k a_j\right)!}{\prod_{j=1}^k (a_j - 1)!} = \prod_{i=1}^k \frac{\left(\sum_{j=i}^k a_j\right)!}{(a_i - 1)! \left(\sum_{j=i+1}^k a_j\right)!},$$

und der zweite Term als

$$u_1^{a_1-1} \left((1-u_1)u_2\right)^{a_2-1} \cdots \left((1-u_1) \cdots (1-u_{k-1})u_k\right)^{a_k-1} = \prod_{i=1}^k u_i^{a_i-1} (1-u_i)^{\sum_{j=i+1}^k (a_j-1)}.$$

Der dritte Term vereinfacht sich zu:

$$(1-u_1)^{k-1} (1-u_2)^{k-2} \cdots (1-u_{k-1}).$$

Insgesamt hat man:

$$\mathbb{P}_a(Y_\infty \in A) = \int_{T^{-1}(A)} \prod_{i=1}^k \left(\frac{\left(\sum_{j=i}^k a_j\right)!}{(a_i - 1)! \left(\sum_{j=i+1}^k a_j\right)!} u_i^{a_i-1} (1-u_i)^{\sum_{j=i+1}^k a_j} \right) \nu(du).$$

In Klammern erkennt man die Dichten der Betaverteilungen mit Parametern a_i und $1 + \sum_{j=i+1}^k a_j$, für $i \leq k$. Des Weiteren hängt der Integrand nicht von den Komponenten u_i , $i > k$, ab. Also ist die Austrittsverteilung das Bildmaß der Verteilung von $(U_i)_{i \in \mathbb{N}}$ unter T , mit U_i wie in (22) und T wie oben. \square

4.3 Martin-Rand der Harris-Kette

Wie bereits erwähnt, lässt sich die Harris-Kette $(X_n)_{n \in \mathbb{N}}$ als ein „trickle-down“-Prozess mit dem China-Restaurant-Prozess als Markov-Kette der Weiterleitungsanweisungen betrachten. Die in [8] von S. N. Evans, R. Grübel und A. Wakolbinger vorgestellte „trickle-down“-Konstruktion erlaubt es, die Doob-Martin Kompaktifizierung des „trickle-down“-Prozesses mit Hilfe der Kompaktifizierungen der Weiterleitungsketten zu beschreiben. Die Autoren identifizieren den Martin-Rand der Harris-Kette als die Menge \mathbb{M} der Wahrscheinlichkeitsmaße auf

$$\left(\mathbb{N}^\infty \sqcup \bigsqcup_{k=0}^{\infty} (\mathbb{N}^k \times \{\diamond\}^\infty), \mathcal{A} \right),$$

mit der von den Mengen

$$A_u := \left\{ v \in \mathbb{N}^\infty \sqcup \bigsqcup_{k=0}^{\infty} (\mathbb{N}^k \times \{\diamond\}^\infty) : u < v \right\}, \quad u \in \mathbb{N}^*,$$

erzeugten σ -Algebra \mathcal{A} . Dabei schreibt man $u < v$ für ein $u = u_1 \dots u_k \in \mathbb{N}^*$ und ein $v = v_1 v_2 \dots \in \mathbb{N}^\infty \sqcup \bigsqcup_{k=0}^{\infty} (\mathbb{N}^k \times \{\diamond\}^\infty)$, falls $u_1 \dots u_k = v_1 \dots v_k$ gilt. Eine Folge $(\mathbf{t}_n)_{n \in \mathbb{N}}$ von Ulam-Harris-Bäumen konvergiert gegen ein $\mu \in \mathbb{M}$ genau dann wenn $\#\mathbf{t}_n \rightarrow \infty$ und

$$\lim_{n \rightarrow \infty} \frac{\#\mathbf{t}_n(u)}{\#\mathbf{t}_n} = \mu(A_u)$$

für alle $u \in \mathbb{N}^*$ gilt. Die Martinrandtheorie liefert die Existenz einer \mathbb{M} -wertigen Zufallsgröße X_∞ mit $X_n \rightarrow X_\infty$ fast sicher in der Doob-Martin-Kompaktifizierung, das heißt, X_∞ ist ein zufälliges Wahrscheinlichkeitsmaß auf $(\mathbb{N}^\infty \sqcup \bigsqcup_{k=0}^{\infty} (\mathbb{N}^k \times \{\diamond\}^\infty), \mathcal{A})$. Ausgehend von der Bemerkung 6.4 in [8, S.39] hat X_∞ dieselbe Verteilung wie ein Maß, welches den Mengen A_u , $u \in \mathbb{N}^*$, die Masse $\prod_{\Lambda < v \leq u} Z_{\bar{v}, \varepsilon(v)}$ zuordnet, wobei die $Z_u = (Z_{u,1}, Z_{u,2}, Z_{u,3}, \dots)$ unabhängig und jeweils GEM-verteilt sind.

4.4 Alternativer Zugang zur Konvergenz

Wir betrachten eine Folge $(\eta_n)_{n \in \mathbb{N}}$ unabhängiger, auf dem Einheitsintervall gleichverteilter Zufallsvariablen und die daraus erzeugte Folge $(X_n)_{n \in \mathbb{N}}$ zufälliger Ulam-Harris-Bäume. Der Ausgangspunkt bei der Konstruktion der Doob-Martin-Kompaktifizierung ist der der Harris-Kette zugrundeliegende Übergangsmechanismus. Der algorithmische Hintergrund ermöglicht einen alternativen Zugang zu den mit der Martinrandtheorie in [8] erhaltenen Konvergenzresultaten und eine Darstellung des Grenzwertes X_∞ mit Hilfe der Inputfolge $(\eta_n)_{n \in \mathbb{N}}$. Folgende Aussagen sind dabei hilfreich:

Bemerkung 4.7.

1. Aus dem Satz von Glivenko-Cantelli folgt

$$\sup_{0 \leq a < b \leq 1} \left| \frac{1}{n} \sum_{i=1}^n 1_{(a,b]}(\eta_i) - (b - a) \right| \rightarrow 0$$

mit $n \rightarrow \infty$ fast sicher.

2. Sei $(\mathcal{G}_n)_{n \in \mathbb{N}}$ mit $\mathcal{G}_n = \sigma(\eta_1, \dots, \eta_n) \subset \mathcal{G}$ für jedes $n \in \mathbb{N}$ die natürliche Filtration von $(\eta_n)_{n \in \mathbb{N}}$ und τ eine fast sicher endliche Stoppzeit. Der Markov-Prozess $(\eta_n)_{n \in \mathbb{N}}$ hat die starke Markov-Eigenschaft: Der Post- τ -Prozess $(\eta_{\tau+n})_{n \in \mathbb{N}}$ ist unabhängig von der σ -Algebra $\mathcal{G}_\tau = \{A \in \mathcal{G} : A \cap \{\tau \leq n\} \in \mathcal{G}_n \text{ für alle } n \in \mathbb{N}\}$ der τ -Vergangenheit.
3. Durch Zerlegung nach den Werten von τ lässt sich zeigen, dass der Post- τ -Prozess $(\eta_{\tau+n})_{n \in \mathbb{N}}$ eine Familie unabhängiger und jeweils $\text{unif}(0, 1)$ -verteilter Zufallsvariablen ist.

Um das asymptotische Verhalten der Teilbaumgrößen an der Wurzel zu untersuchen, betrachten wir noch einmal den Algorithmus. Beim Übergang von n zu $n+1$ wird der neue Knoten bewertet mit η_n ein direkter Nachfolger des Wurzelknotens falls $\eta_n < \eta_1, \dots, \eta_{n-1}$ gilt. Wir setzen $\tau(1) := 1$ und definieren für jedes $u \in \mathbb{N}$, $u \geq 2$

$$\tau(u) := \min \{j : \eta_j < \eta_{\tau(u-1)}\}.$$

Die $\tau(u) + 1$ sind die Zeitpunkte der Hinzunahme der Nachfolgerknoten der Wurzel zum zufälligen Baum, da der erste Baum aus dem Wurzelknoten bewertet mit 0 besteht. Wir stellen fest, dass die Folge $(\eta_{\tau(u)})_{u \in \mathbb{N}}$ der Bewertungen der direkten Nachfolgerknoten die Folge der absteigenden Rekorde ist. Wird beim Schritt $n + 1$ der neue Knoten mit der Bewertung η_n kein direkter Nachfolger der Wurzel, so wird der bereits existierende Knoten mit der Bewertung

$$\max \{\eta_i : 1 \leq i \leq n - 1, \eta_i < \eta_n\}$$

als direkter Vorgänger ausgewählt.

Betrachten wir zunächst den einfachen Fall $u = 1$. Zum Teilbaum von $u = 1$ kommen nur die Knoten mit der Bewertung größer als der Trivialrekord η_1 hinzu. Da η_2, η_3, \dots

unabhängige $\text{unif}(0, 1)$ -verteilte Zufallsvariablen sind, folgt mit dem Satz von Glivenko-Cantelli (siehe erste Aussage aus der Bemerkung 4.7):

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{\#X_n(1)}{n} &= \lim_{n \rightarrow \infty} \frac{1}{n} \# \{2 < i \leq n : \eta_i \in (\eta_1, 1)\} \\ &= \lim_{n \rightarrow \infty} \frac{n-1}{n} \frac{1}{n-1} \sum_{i=1}^{n-1} 1_{\{\eta_1 < \eta_{1+i} < 1\}} \\ &= 1 - \eta_1 \end{aligned}$$

mit $n \rightarrow \infty$ fast sicher.

Betrachten wir nun den Teilbaum vom Knoten $u \in \mathbb{N}$, $u \geq 2$. Nach Konstruktion kommt ein Knoten genau dann zu dem Teilbaum von $u \geq 2$ hinzu, falls seine Bewertung zwischen den benachbarten Rekorden $\eta_{\tau(u)}$ und $\eta_{\tau(u-1)}$ liegt. In diesem Fall ist also der Post- $\tau(u)$ -Prozess $(\eta_{\tau(u)+n})_{n \in \mathbb{N}}$ zu untersuchen. Dieser ist wieder eine Folge unabhängiger und $\text{unif}(0, 1)$ -verteilter Zufallsvariablen (siehe Bemerkung 4.7). Also lässt sich hier wieder der Satz von Glivenko-Cantelli anwenden:

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{\#X_n(u)}{n} &= \lim_{n \rightarrow \infty} \frac{1}{n} \# \{ \tau(u) < i \leq n : \eta_i \in (\eta_{\tau(u)}, \eta_{\tau(u-1)}) \} \\ &= \lim_{n \rightarrow \infty} \frac{n - \tau(u)}{n} \frac{1}{n - \tau(u)} \sum_{i=1}^{n - \tau(u)} 1_{\{\eta_{\tau(u)} < \eta_{\tau(u)+i} < \eta_{\tau(u-1)}\}} \\ &= \eta_{\tau(u-1)} - \eta_{\tau(u)} \end{aligned}$$

mit $n \rightarrow \infty$ fast sicher.

Insgesamt konvergieren die relativen Teilbaumgrößen $(\#X_n(u)/n)_{u \in \mathbb{N}}$ an dem Wurzelknoten mit $n \rightarrow \infty$ fast sicher gegen die Folge der Rekorddifferenzen

$$(1 - \eta_1, \eta_1 - \eta_{\tau(2)}, \eta_{\tau(2)} - \eta_{\tau(3)}, \dots).$$

Die Folge der absteigenden Rekorde hat eine einfache Darstellung (siehe [1, S.21]): Für $u = 1, 2, \dots$ gilt

$$\eta_{\tau(u)} = \prod_{i=1}^u V_i \quad \text{in Verteilung,}$$

wobei $(V_i)_{i \in \mathbb{N}}$ eine Folge unabhängiger $\text{unif}(0, 1)$ -verteilter Zufallsvariablen ist. Mit V_i ist auch $1 - V_i$ auf dem Einheitsintervall gleichverteilt. Somit hat die Folge der Rekorddifferenzen dieselbe Verteilung wie der „stick breaking“-Prozesses, also die GEM-Verteilung. Die Theorie der Rekorde liefert also einen alternativen Zugang zum bereits bekannten (siehe [8] und Abschnitt 4.2) asymptotischen Verhalten der relativen Teilbaumgrößen an der Wurzel.

Satz 4.8. *Es sei $(X_n)_{n \in \mathbb{N}}$ eine Folge zufälliger Ulam-Harris-Bäume konstruiert zu einer Folge $(\eta_n)_{n \in \mathbb{N}}$ unabhängiger und jeweils auf dem Einheitsintervall $[0, 1]$ gleichverteilter Zufallsvariablen. Für jedes $u \in \mathbb{N}^*$ sei $\tau(u) := \inf \{n \in \mathbb{N} : X_n \ni u\} - 1$. Dann existiert eine \mathbb{M} -wertige Zufallsgröße X_∞ , sodass*

$$\lim_{n \rightarrow \infty} \frac{\#X_n(u)}{n} = X_\infty(A_u)$$

für alle $u \in \mathbb{N}^*$ fast sicher gilt. Es seien außerdem

$$0 =: \eta_{(\tau(u):0)} < \eta_{(\tau(u):1)} < \dots < \eta_{(\tau(u):\tau(u))} < \eta_{(\tau(u):\tau(u)+1)} := 1$$

erweiterte Ordnungsstatistiken zu $\eta_1, \dots, \eta_{\tau(u)}$. Dann gilt

$$X_\infty(A_u) = \eta_{(\tau(u):\kappa(u)+1)} - \eta_{(\tau(u):\kappa(u))},$$

wobei $\kappa(u) := 1 + \sum_{k=1}^{\tau(u)} \mathbf{1}_{\{\eta_k < \eta_{\tau(u)}\}}$ der Rang von $\eta_{\tau(u)}$ unter $\eta_1, \dots, \eta_{\tau(u)}$ bei Anordnung nach wachsender Größe sei.

Beweis. Wir betrachten einen beliebigen Knoten $u \in \mathbb{N}^*$. Wegen der Sonderrolle des Wurzelknotens ist $\inf\{n \in \mathbb{N} : X_n \ni u\}$ der Zeitpunkt der Hinzunahme des Knotens u zum zufälligen Baum und $\eta_{\tau(u)}$ seine Bewertung. Ein weiterer Knoten v kommt zum Teilbaum von u hinzu, wenn $\tau(v) > \tau(u)$ und $\eta_{\tau(v)} \in (\eta_{\tau(u)}, \eta_{(\tau(u):\kappa(u)+1)})$ gilt, da der Algorithmus Bäume mit pfadmonotonen Bewertungen erzeugt. Wäre die obige Bedingung also nicht erfüllt, so wäre v Nachfolger eines Knotens mit der Bewertung $\geq \eta_{(\tau(u):\kappa(u)+1)}$. Man beachte, dass $\eta_{\tau(u)} = \eta_{(\tau(u):\kappa(u))}$ gilt und dass $\tau(u)$ eine fast sicher endliche Stoppzeit ist. Mit dem Satz von Glivenko-Cantelli erhalten wir wieder:

$$X_\infty(A_u) = \lim_{n \rightarrow \infty} \frac{\#X_n(u)}{n} = \eta_{(\tau(u):\kappa(u)+1)} - \eta_{(\tau(u):\kappa(u))}$$

mit $n \rightarrow \infty$ fast sicher. Für $u \in \mathbb{N}^*$ ist die Zufallsgröße

$$\xi_u = (\xi_{u,1}, \xi_{u,2}, \dots) := \left(\frac{X_\infty(A_{u1})}{X_\infty(A_u)}, \frac{X_\infty(A_{u2})}{X_\infty(A_u)}, \dots \right)$$

die Folge der Differenzen der standardisierten absteigenden Rekorde im Intervall $(\eta_{(\tau(u):\kappa(u))}, \eta_{(\tau(u):\kappa(u)+1)})$ der Ordnungsstatistiken und nimmt Werte in dem unendlich-dimensionalen Wahrscheinlichkeitssubsimplex

$$\Delta^\infty = \left\{ (\rho_n)_{n \in \mathbb{N}} : \rho_n \in [0, 1] \forall n \in \mathbb{N}, \sum_{n=1}^{\infty} \rho_n \leq 1 \right\}$$

an. Ein $\rho \in \Delta^\infty$ kann als ein Wahrscheinlichkeitsmaß auf $\mathbb{N} \cup \{\diamond\}$ aufgefasst werden, welcher den Punkten $n \in \mathbb{N}$ die Masse ρ_n und dem Punkt \diamond die Masse $1 - \sum_{n=1}^{\infty} \rho_n$ zuordnet. Es gilt

$$X_\infty(A_u) = \prod_{\Lambda < v \leq u} \xi_{\bar{v}, \varepsilon(v)}.$$

Also ist X_∞ als ein zufälliges Wahrscheinlichkeitsmaß auf $(\mathbb{N}^\infty \sqcup \bigsqcup_{k=0}^{\infty} (\mathbb{N}^k \times \{\diamond\}^\infty), \mathcal{A})$ interpretierbar. \square

Für den alternativen Beweis der Konvergenzresultate für die zufälligen Ulam-Harris-Bäume brauchen wir zwei weitere Hilfsmittel.

Lemma 4.9. *Die mit $(b - a)$, $0 \leq a < b \leq 1$, standardisierte Folge der Differenzen der absteigenden Rekorde im Intervall $(a, b] \subset [0, 1]$ hat dieselbe Verteilung wie die Folge der Differenzen der absteigenden Rekorde im Intervall $[0, 1]$.*

Beweis. Es seien

$$\begin{aligned}\sigma(1) &:= \min \{j : \eta_j \in (a, b]\}, \\ \sigma(n) &:= \min \{j : \eta_j \in (a, b], j > \sigma(n-1)\} \text{ für } n \geq 2 \text{ und} \\ Y_n &:= \frac{\eta_{\sigma(n)} - a}{b - a} \text{ für } n \in \mathbb{N}.\end{aligned}$$

Es reicht zu zeigen, dass $(Y_n)_{n \in \mathbb{N}}$ eine Folge unabhängiger $\text{unif}(0, 1)$ -verteilter Zufallsvariablen ist. Es ist

$$\mathcal{L} \left(\frac{\eta_n - a}{b - a} \mid \eta_n \in (a, b] \right) = \text{unif}(0, 1).$$

Wir zerlegen nach den Werten der $\sigma(n)$:

$$\begin{aligned}\mathbb{P}(Y_1 \leq y_1, \dots, Y_n \leq y_n) \\ = \sum_{i_1=1}^{\infty} \sum_{i_2=i_1+1}^{\infty} \dots \sum_{i_n=i_{n-1}+1}^{\infty} \mathbb{P}(Y_1 \leq y_1, \dots, Y_n \leq y_n, \sigma(1) = i_1, \dots, \sigma(n) = i_n),\end{aligned}$$

wobei $y_i \in [0, 1]$ für $1 \leq i \leq n$ seien. Man hat

$$\begin{aligned}\mathbb{P}(Y_1 \leq y_1, \dots, Y_n \leq y_n, \sigma(1) = i_1, \dots, \sigma(n) = i_n) \\ = \mathbb{P} \left(\frac{\eta_{\sigma(1)} - a}{b - a} \leq y_1, \dots, \frac{\eta_{\sigma(n)} - a}{b - a} \leq y_n, \sigma(1) = i_1, \dots, \sigma(n) = i_n \right) \\ = \mathbb{P} \left(\frac{\eta_{i_1} - a}{b - a} \leq y_1, \dots, \frac{\eta_{i_n} - a}{b - a} \leq y_n, \sigma(1) = i_1, \dots, \sigma(n) = i_n \right)\end{aligned}$$

und

$$\begin{aligned}\{\sigma(1) = i_1, \dots, \sigma(n) = i_n\} = \{ \eta_{i_j} \in (a, b] \text{ für alle } 1 \leq j \leq n \\ \text{und } \eta_i \notin (a, b] \text{ für alle } i \in \{1, \dots, i_n\} \setminus \{i_1, \dots, i_n\} \}.\end{aligned}$$

Also folgt mit der Unabhängigkeit der η_i :

$$\begin{aligned}\mathbb{P}(Y_1 \leq y_1, \dots, Y_n \leq y_n, \sigma(1) = i_1, \dots, \sigma(n) = i_n) \\ = \mathbb{P} \left(\frac{\eta_{i_1} - a}{b - a} \leq y_1, \dots, \frac{\eta_{i_n} - a}{b - a} \leq y_n, \mid \eta_{i_j} \in (a, b] \text{ für alle } 1 \leq j \leq n \right) \\ \times \mathbb{P}(\eta_{i_j} \in (a, b] \text{ für alle } 1 \leq j \leq n) \\ \times \mathbb{P}(\eta_i \notin (a, b] \text{ für alle } i \in \{1, \dots, i_n\} \setminus \{i_1, \dots, i_n\}) \\ = \left(\prod_{i=1}^n y_i \right) (b - a)^n (1 - (b - a))^{i_n - n}\end{aligned}$$

Schließlich:

$$\begin{aligned}\mathbb{P}(Y_1 \leq y_1, \dots, Y_n \leq y_n) &= \left(\prod_{i=1}^n y_i \right) (b - a)^n \sum_{i_1=1}^{\infty} \sum_{i_2=i_1+1}^{\infty} \dots \sum_{i_n=i_{n-1}+1}^{\infty} (1 - (b - a))^{i_n - n} \\ &= \left(\prod_{i=1}^n y_i \right) (b - a)^n \sum_{j=n}^{\infty} \binom{j-1}{n-1} (1 - (b - a))^{j-n}\end{aligned}$$

$$\begin{aligned}
&= \left(\prod_{i=1}^n y_i \right) (b-a)^n \left(\frac{1}{(1 - (1 - (b-a)))^n} \right) \\
&= \prod_{i=1}^n y_i.
\end{aligned}$$

Daraus folgt, dass die Y_i unabhängig und jeweils $\text{unif}(0, 1)$ -verteilt sind und damit auch die Behauptung. \square

Das zweite Hilfsmittel ist die folgende von R. Grübel vorgeschlagene Zerlegung.

Bemerkung 4.10. Für ein $\mu \in \mathbb{M}$ mit $\mu(A_1) \in (0, 1)$ definiert man $\mu^b, \mu^\sharp \in \mathbb{M}$ durch

$$\begin{aligned}
\mu^b(A_u) &= \frac{\mu(A_{1 \circ u_1 \dots u_k})}{\mu(A_1)}, \\
\mu^\sharp(A_u) &= \frac{\mu(A_{(1+u_1) \circ u_2 \dots u_k})}{1 - \mu(A_1)}
\end{aligned}$$

für $u = u_1 \dots u_k \in \mathbb{N}^* \setminus \{\Lambda\}$ und $\mu^b(A_\Lambda) = \mu^\sharp(A_\Lambda) = 1$. Man kann zeigen, dass $X_\infty(A_1)$, X_∞^b und X_∞^\sharp unabhängig sind und dass X_∞^b und X_∞^\sharp dieselbe Verteilung wie X_∞ haben.

Satz 4.11. Die Zufallsgrößen

$$\xi_u = \left(\frac{X_\infty(A_{u_1})}{X_\infty(A_u)}, \frac{X_\infty(A_{u_2})}{X_\infty(A_u)}, \dots \right), \quad u \in \mathbb{N}^*,$$

sind unabhängig und jeweils GEM-verteilt.

Beweis. Wir haben bereits festgestellt, dass ξ_u für $u \in \mathbb{N}^*$ die Folge der Differenzen der standardisierten absteigenden Rekorde im Intervall $(\eta_{(\tau(u):\kappa(u))}, \eta_{(\tau(u):\kappa(u)+1)})$ der entsprechenden Ordnungsstatistiken ist. Also ist ξ_u nach Bemerkung 4.7 und Lemma 4.9 GEM-verteilt. Es bleibt die Unabhängigkeit der $(\xi_u)_{u \in \mathbb{N}^*}$ zu zeigen. Wir betrachten dazu die Ulam-Harris-Bäume $\mathbf{t}_k = \{u : |u|_1 \leq k\}$, $k \in \mathbb{N}$, als Indexmengen. Dabei ist $|u|_1 := u_1 + u_2 + \dots + u_{|u|}$. Sind für alle $k \in \mathbb{N}$ die ξ_u mit $u \in \mathbf{t}_k$ unabhängig, so besteht jede endliche Teilfamilie von $(\xi_u)_{u \in \mathbb{N}^*}$ aus unabhängigen Zufallsgrößen, was die zweite Aussage beweist.

Sei k fest. Man beachte, dass in \mathbf{t}_k genau k Knoten die Tiefe 1 haben. Wir definieren $\mathbf{t}_{k,u} = \{v \in \mathbf{t}_k : u \leq v\}$. Die k -fache Wiederholung der Zerlegung aus der Bemerkung 4.10 ergibt die gemeinsame Unabhängigkeit von

$$\left(X_\infty(A_1), X_\infty^\sharp(A_1), \dots, X_\infty^{\sharp^{k-1}}(A_1), X_\infty^b, X_\infty^{\sharp b}, \dots, X_\infty^{\sharp^{k-1}b}, X_\infty^{\sharp^k} \right),$$

wobei $\sharp^j = \underbrace{\sharp \dots \sharp}_j$ sei. Wir stellen fest:

- Die ξ_u mit $1 \leq u$ sind Funktionen von X_∞^b . Für $u = 1u_2 \dots u_j$ gilt nämlich

$$\xi_u = \left(\frac{X_\infty^b(A_{u_2 \dots u_j 1})}{X_\infty^b(A_{u_2 \dots u_j})}, \frac{X_\infty^b(A_{u_2 \dots u_j 2})}{X_\infty^b(A_{u_2 \dots u_j})}, \dots \right).$$

- Die ξ_u mit $l \leq u$, $1 < l \leq k$ sind Funktionen von $X_\infty^{\#^{l-1}b}$, denn für $u = lu_2 \dots u_j$ gilt

$$\xi_u = \left(\frac{X_\infty^{\#^{l-1}b}(A_{u_2 \dots u_j 1})}{X_\infty^{\#^{l-1}b}(A_{u_2 \dots u_j})}, \frac{X_\infty^{\#^{l-1}b}(A_{u_2 \dots u_j 2})}{X_\infty^{\#^{l-1}b}(A_{u_2 \dots u_j})}, \dots \right).$$

- ξ_Λ ist eine Funktion von $X_\infty(A_1), X_\infty^\#(A_1), \dots, X_\infty^{\#^{k-1}}(A_1)$ und $X_\infty^{\#^k}$:

$$\begin{aligned} \xi_\Lambda = & \left(X_\infty(A_1), (1 - X_\infty(A_1))X_\infty^\#(A_1), \dots, \right. \\ & (1 - X_\infty(A_1)) \dots (1 - X_\infty^{\#^{k-1}}(A_1))X_\infty^{\#^k}(A_1), \\ & \left. (1 - X_\infty(A_1)) \dots (1 - X_\infty^{\#^{k-1}}(A_1))X_\infty^{\#^k}(A_2), \dots \right). \end{aligned}$$

Daraus ergibt sich die Unabhängigkeit von $\xi_\Lambda, (\xi_u)_{u \in t_{k,1}}, \dots, (\xi_u)_{u \in t_{k,k}}$. Die so entstandenen unabhängigen Teilfamilien von $(\xi_u)_{u \in t_k}$ sowie die entsprechenden Indexmengen lassen sich durch die Anwendung der gesamten obigen Prozedur auf die dazugehörigen $X_\infty^{\dots b}$ -Objekte weiter rekursiv aufspalten, siehe das Beispiel für $k = 3$ in der Abbildung 8.

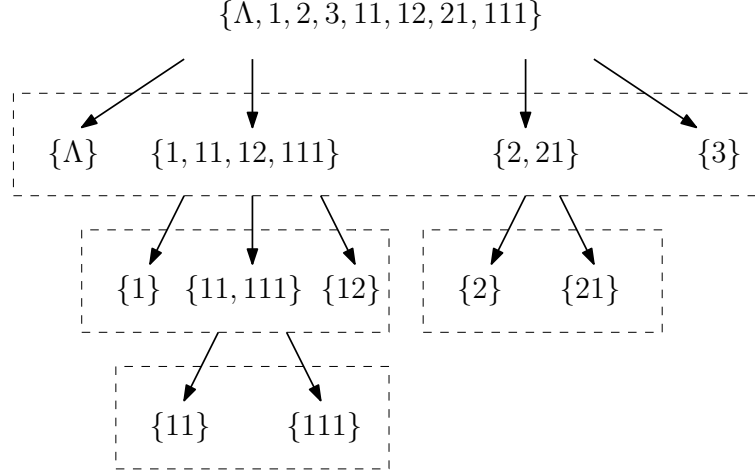


Abbildung 8: Zerlegung der Indexmenge t_3 . Die Prozedur aus dem Beweis wurde vier mal angewendet.

Diese Zerlegung beweist für alle $k \in \mathbb{N}$ die Unabhängigkeit der ξ_u mit $u \in t_k$. □

4.5 Baumfunktionale

Es sei $(X_n)_{n \in \mathbb{N}}$ eine Harris-Kette, X_∞ der fast sichere Grenzwert dieser Markov-Kette, und $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$. Wir stellen zunächst einige für die Analyse der Baumfunktionale relevante Aussagen zusammen. Mit obigen Überlegungen zur Harris-Kette und dem China-Restaurant-Prozess folgt:

- Die Zufallsgrößen $\xi_u = \left(\frac{X_\infty(A_{u1})}{X_\infty(A_u)}, \frac{X_\infty(A_{u2})}{X_\infty(A_u)}, \dots \right)$ sind unabhängig jeweils GEM-verteilt.
- Es gilt

$$X_\infty(A_u) = \prod_{\Lambda < v \leq u} \xi_{\bar{v}, \varepsilon(v)}.$$

- $\mathcal{L}(\xi_u | \#X_n(u \circ i) = a_i \text{ für } i \leq k, \#X_n(u \circ i) = 0 \text{ für } i > k), a_1, \dots, a_k > 0$ stimmt überein mit der Verteilung von

$$(U_1, (1 - U_1)U_2, (1 - U_1)(1 - U_2)U_3, \dots)$$

mit U_i unabhängig und

$$\begin{aligned} U_i &\sim \text{Beta}\left(a_i, 1 + \sum_{j=i+1}^k a_j\right), \text{ für } i < k, \\ U_k &\sim \text{Beta}(a_k, 1), \\ U_i &\sim \text{unif}(0, 1), \text{ für } i > k. \end{aligned}$$

- Für $i \leq k$ gilt:

$$\begin{aligned} &\mathcal{L}(\xi_{u,i} | \#X_n(u \circ 1) = a_1, \dots, \#X_n(u \circ k) = a_k, \#X_n(u \circ (k+1)) = 0, \dots) \\ &= \text{Beta}\left(a_i, 1 + \sum_{\substack{j=1 \\ j \neq i}}^k a_j\right). \end{aligned}$$

- Für $i > k$ gilt:

$$\mathcal{L}(\xi_{u,i} | \#X_n(u \circ 1) = a_1, \dots, \#X_n(u \circ k) = a_k, \#X_n(u \circ (k+1)) = 0, \dots)$$

stimmt überein mit der Verteilung von $Y \prod_{j=1}^{i-k} V_j$, wobei Y, V_1, \dots, V_{i-k} unabhängig sind, $Y \sim \text{Beta}(1, \sum_{j=1}^k a_j)$ und $V_j \sim \text{unif}(0, 1)$ für $1 \leq j \leq i - k$.

4.5.1 Interne Pfadlänge

Die interne Pfadlänge eines Ulam-Harris-Baumes \mathbf{t}

$$\text{IPL}(\mathbf{t}) := \sum_{u \in \mathbf{t}} |u|$$

lässt sich mit Hilfe der Teilbaumgrößen wie im Fall binärer Bäume (siehe [13, S.12]) auch schreiben als

$$\text{IPL}(\mathbf{t}) = \sum_{u \in \mathbf{t}} \#\mathbf{t}(u) - \#\mathbf{t}.$$

Dies erlaubt es, die Aussagen über die fast sichere Konvergenz für zufällige Ulam-Harris-Bäume einzusetzen, um die fast sichere Konvergenz für die interne Pfadlänge zu beweisen. Dazu werden im Folgenden einige Hilfsgrößen eingeführt und Hilfsaussagen gezeigt.

Lemma 4.12. *Es gilt für $u \in X_n$: $E[X_\infty(A_u) | \mathcal{F}_n] = \frac{\#X_n(u)}{n}$.*

Beweis. Mit einem Teleskopproduktargument erhält man unter Verwendung von $1 + \sum_{u \circ i \in X_n} \#X_n(u \circ i) = \#X_n(u)$ für $u \in X_n$:

$$\begin{aligned} E[X_\infty(A_u) | \mathcal{F}_n] &= \prod_{\Lambda < v \leq u} E[\xi_{\bar{v}, \varepsilon(v)} | \mathcal{F}_n] \\ &= \prod_{\Lambda < v \leq u} \frac{\#X_n(v)}{1 + \sum_{\bar{v} \circ i \in X_n} \#X_n(\bar{v} \circ i)} \\ &= \prod_{\Lambda < v \leq u} \frac{\#X_n(v)}{\#X_n(\bar{v})} = \frac{\#X_n(u)}{n}. \quad \square \end{aligned}$$

Definition 4.13. Wir definieren

$$C : \Delta^\infty \rightarrow [-\infty, 1], \quad C(s) := 1 + \sum_{i=1}^{\infty} s_i \ln s_i$$

mit der Vereinbarung $0 \ln 0 := 0$.

Eine Folge $s = (s_i)_{i \in \mathbb{N}} \in \Delta^\infty$ mit $\sum_{i=1}^{\infty} s_i = 1$ kann als ein Wahrscheinlichkeitsmaß auf \mathbb{N} interpretiert werden. Der Ausdruck $-\sum_{i=1}^{\infty} s_i \ln s_i$ wird dann als Entropie von s bezeichnet. In der vorliegenden Situation muss die Entropie nicht zwingend endlich sein [18, S.120].

Bemerkung 4.14. Das folgende Hilfsmittel aus [13, S.14] erweist sich beim Beweis des nächsten Lemmas als nützlich: Es sei Z eine Beta(a, b)-verteilte Zufallsvariable. Dann gilt:

$$EZ \log Z = \frac{a}{a+b} (H(a) - H(a+b)),$$

wobei $H(0) := 0$, $H(n) := \sum_{i=1}^n \frac{1}{i}$, $n \in \mathbb{N}$, die harmonischen Zahlen sind.

Lemma 4.15. Es gilt für $u \notin X_n$

$$E[C(\xi_u) | \mathcal{F}_n] = 0$$

und für $u \in X_n$

$$E[C(\xi_u) | \mathcal{F}_n] = 1 + \frac{\sum_{u \circ i \in X_n} \#X_n(u \circ i) H(\#X_n(u \circ i))}{\#X_n(u)} - H(\#X_n(u)).$$

Außerdem hat die Zufallsvariable $C(\xi_u)$ für jedes $u \in \mathbb{N}^*$ ein endliches zweites Moment.

Beweis. Zur ersten Behauptung: Für $u \notin X_n$ und ein festes i hat die Zufallsvariable $\xi_{u,i} \log \xi_{u,i}$ dieselbe Verteilung wie die Zufallsgröße $\left(\prod_{j=1}^i Z_j\right) \left(\log \prod_{j=1}^i Z_j\right)$, wobei Z_j unabhängig und jeweils $\text{unif}(0, 1)$ -verteilt sind. Es gilt also unter Verwendung der Bemerkung 4.14:

$$\begin{aligned} E(\xi_{u,i} \log \xi_{u,i}) &= E\left(\left(\prod_{j=1}^i Z_j\right) \left(\log \prod_{j=1}^i Z_j\right)\right) = i (EZ_1)^{i-1} E(Z_1 \log Z_1) \\ &= -\frac{1}{4} i \left(\frac{1}{2}\right)^{i-1}. \end{aligned}$$

Aufsummieren liefert (mit dem Satz von der monotonen Konvergenz angewendet auf Funktionenreihen, die Zufallsvariablen $-\xi_{u,i} \ln \xi_{u,i}$ sind nichtnegativ):

$$1 + \sum_{n=1}^{\infty} n (EZ_1)^{n-1} E(Z_1 \log Z_1) = 1 - \frac{1}{4} \sum_{n=1}^{\infty} n \left(\frac{1}{2}\right)^{n-1} = 0.$$

Zur dritten Behauptung: Es sei $C_k(\xi_u) := 1 + \sum_{i=1}^k \xi_{u,i} \log \xi_{u,i}$. Es ist $\|C_{k+1}(\xi_u) - C_k(\xi_u)\|_2 = \|\xi_{u,k+1} \log \xi_{u,k+1}\|_2$. Mit obigen Bezeichnungen gilt

$$\begin{aligned} \sum_{i=1}^n \|\xi_{u,i} \log \xi_{u,i}\|_2 &\leq \sum_{i=1}^n \sum_{j=1}^i \left\| \left(\prod_{l=1}^i Z_l \right) \log Z_j \right\|_2 \\ &= \sum_{i=1}^n \sum_{j=1}^i \left(\frac{2}{27} 3^{-(i-1)} \right)^{1/2} \\ &= \sum_{i=1}^n i \left(2 \left(\frac{1}{3} \right)^{i+2} \right)^{1/2}. \end{aligned}$$

Dabei wurde verwendet, dass

$$\left\| \left(\prod_{l=1}^i Z_l \right) \log Z_j \right\|_2^2 = E \left(\prod_{\substack{l=1 \\ l \neq j}}^i Z_l^2 \right) EZ_j^2 (\log Z_j)^2 = \left(\frac{1}{3} \right)^{i-1} EZ_j^2 (\log Z_j)^2$$

und

$$\begin{aligned} EZ_j^2 (\log Z_j)^2 &= \int_0^1 z^2 (\log z)^2 dz = \frac{z^3}{3} \log z \Big|_0^1 - \frac{2}{3} \int_0^1 z^2 \log z dz \\ &= -\frac{2}{3} \left(\frac{z^3}{3} \log z \Big|_0^1 - \frac{1}{3} \int_0^1 z^2 dz \right) = \frac{2}{27} \end{aligned}$$

gilt. Mit

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n i \left(2 \left(\frac{1}{3} \right)^{i+2} \right)^{1/2} = \lim_{n \rightarrow \infty} \frac{\sqrt{2}}{3} \sum_{i=1}^n i \left(\frac{1}{\sqrt{3}} \right)^i = \frac{\sqrt{2}}{3} \frac{\sqrt{3}}{(\sqrt{3}-1)^2} < \infty$$

folgt schließlich, dass die Partialsummen $C_k(\xi_u) = 1 + \sum_{i=1}^k \xi_{u,i} \log \xi_{u,i}$ eine Cauchy-Folge in L^2 bilden, also: $EC(\xi_u)^2 = \|C(\xi_u)\|_2^2 < \infty$.

Zur zweiten Behauptung: Sei wieder $\#X_n(u \circ 1) = a_1, \dots, \#X_n(u \circ k) = a_k, \#X_n(u \circ j) = 0$ für $j > k$, wobei $a_1, \dots, a_k > 0$. Dann gilt für $i \leq k$: $\mathcal{L}(\xi_{u,i} | \mathcal{F}_n) = \text{Beta} \left(a_i, 1 + \sum_{\substack{j=1 \\ j \neq i}}^k a_j \right)$, man erhält also mit der Bemerkung 4.14:

$$E\xi_{u,i} \log \xi_{u,i} = \frac{a_i}{1 + \sum_{j=1}^k a_j} \left(H(a_i) - H\left(1 + \sum_{j=1}^k a_j\right) \right).$$

Im Fall $i > k$ hat $\xi_{u,i}$ unter \mathcal{F}_n dieselbe Verteilung wie die Zufallsvariable YZ , Y und Z unabhängig, wobei $Y \sim \text{Beta}(1, \sum_{j=1}^k a_j)$ und Z ein Produkt von $i - k$ unabhängigen $\text{unif}(0, 1)$ -verteilten Zufallsgrößen ist. Es gilt:

$$\begin{aligned} EYZ \log YZ &= E(Y \log Y)Z + EY(Z \log Z), \\ E(Y \log Y)Z &= \frac{1}{1 + \sum_{j=1}^k a_j} \left(H(1) - H(1 + \sum_{j=1}^k a_j) \right) \left(\frac{1}{2} \right)^{i-k}, \\ EY(Z \log Z) &= \frac{1}{1 + \sum_{j=1}^k a_j} \left((i - k) \left(\frac{1}{2} \right)^{i-k-1} \left(-\frac{1}{4} \right) \right). \end{aligned}$$

Aufsummieren liefert:

$$\begin{aligned} \sum_{i=k+1}^{\infty} \frac{1}{1 + \sum_{j=1}^k a_j} \left(H(1) - H(1 + \sum_{j=1}^k a_j) \right) \left(\frac{1}{2} \right)^{i-k} \\ &= \frac{1}{1 + \sum_{j=1}^k a_j} \left(H(1) - H(1 + \sum_{j=1}^k a_j) \right) \sum_{i=0}^{\infty} \left(\frac{1}{2} \right)^{i+1} \\ &= \frac{1}{1 + \sum_{j=1}^k a_j} \left(H(1) - H(1 + \sum_{j=1}^k a_j) \right) \end{aligned}$$

und

$$\begin{aligned} \sum_{i=k+1}^{\infty} \frac{1}{1 + \sum_{j=1}^k a_j} \left((i - k) \left(\frac{1}{2} \right)^{i-k-1} \left(-\frac{1}{4} \right) \right) &= \frac{1}{1 + \sum_{j=1}^k a_j} \left(-\frac{1}{4} \right) \sum_{i=0}^{\infty} i \left(\frac{1}{2} \right)^{i-1} \\ &= -\frac{1}{1 + \sum_{j=1}^k a_j}. \end{aligned}$$

Somit:

$$\sum_{i=k+1}^{\infty} E[\xi_{u,i} \log \xi_{u,i} | \mathcal{F}_n] = -\frac{H(1 + \sum_{j=1}^k a_j)}{1 + \sum_{j=1}^k a_j}.$$

Insgesamt:

$$\begin{aligned} E[C(\xi_u) | \mathcal{F}_n] &= 1 + \sum_{i=1}^k \frac{a_i}{1 + \sum_{j=1}^k a_j} \left(H(a_i) - H(1 + \sum_{j=1}^k a_j) \right) \\ &\quad - \frac{H(1 + \sum_{j=1}^k a_j)}{1 + \sum_{j=1}^k a_j} \\ &= 1 + \frac{1}{1 + \sum_{j=1}^k a_j} \left(\sum_{i=1}^k a_i H(a_i) \right) - H(1 + \sum_{j=1}^k a_j). \quad \square \end{aligned}$$

Lemma 4.16. *Der Grenzwert*

$$Y_{\infty} := \sum_{u \in \mathbb{N}^*} X_{\infty}(A_u) C(\xi_u)$$

existiert fast sicher und in L^2 .

Beweis. Es wird die Beweismethode zum Satz 6 aus [13] auf die vorliegende Situation übertragen. Dabei ist zu beachten, dass im Gegensatz zu binären Bäumen bei Ulam-Harris-Bäumen ein Knoten unendlich viele potentielle direkte Nachfolger hat. Daher definieren wir hier \mathcal{G}_k als $\sigma(\xi_u : |u|_1 \leq k)$ wobei $|u|_1 := u_1 + u_2 + \dots + u_{|u|}$ und setzen

$$Y_k := \sum_{\substack{u \in \mathbb{N}^* \\ |u|_1 \leq k}} X_\infty(A_u)C(\xi_u).$$

Nach Konstruktion ist die Zufallsvariable $X_\infty(A_u)$ messbar bezüglich $\mathcal{G}_{|u|_1-1}$, die Zufallsvariable $C(\xi_u)$ ist unabhängig von $\mathcal{G}_{|u|_1}$ und es gilt $E(C(\xi_u)) = 0$. Das alles liefert wie in der Situation binärer Bäume:

$$\begin{aligned} E[Y_{k+1} | \mathcal{G}_k] &= E \left[Y_k + \sum_{|u|_1=k+1} X_\infty(A_u)C(\xi_u) \mid \mathcal{G}_k \right] \\ &= Y_k + \sum_{|u|_1=k+1} X_\infty(A_u)EC(\xi_u) \\ &= Y_k. \end{aligned}$$

Für $u, v \in \mathbb{N}^*$, $u \neq v$ mit $|u|_1 = |v|_1 = k+1$ gilt mit obigen Argumenten:

$$E[X_\infty(A_u)C(\xi_u)X_\infty(A_v)C(\xi_v) | \mathcal{G}_k] = 0.$$

Daraus erhalten wir:

$$\begin{aligned} E[(Y_{k+1} - Y_k)^2 | \mathcal{G}_k] &= E \left[\left(\sum_{|u|_1=k+1} X_\infty(A_u)C(\xi_u) \right)^2 \mid \mathcal{G}_k \right] \\ &= \sum_{|u|_1=k+1} X_\infty(A_u)^2 EC(\xi_u)^2 \end{aligned}$$

und

$$E(Y_{k+1} - Y_k)^2 = \sum_{|u|_1=k+1} EX_\infty(A_u)^2 EC(\xi_u)^2.$$

Es ist $EX_\infty(A_u)^2 = (\frac{1}{3})^{|u|_1}$. Nach Lemma 4.15 gilt für jedes $u \in \mathbb{N}^*$: $\kappa := EC(\xi_u)^2 < \infty$. Außerdem gibt es $\binom{n-1}{j-1}$ geordnete j -Partitionen der Zahl n . Also hat die obige Summe

$$\sum_{j=1}^{k+1} \binom{k}{j-1} = 2^k$$

Summanden. Somit gilt

$$E(Y_{k+1} - Y_k)^2 = 2^k 3^{-(k+1)} \kappa$$

und

$$\sum_{k=1}^{\infty} E(Y_{k+1} - Y_k)^2 < \infty.$$

Daraus folgt, dass (Y_k, \mathcal{G}_k) ein L^2 -beschränktes Martingal ist, und dass $Y_k \rightarrow Y_\infty$ mit $k \rightarrow \infty$ fast sicher und in L^2 gilt, siehe beispielsweise Abschnitt 12.1. in [28]. \square

Die fast sichere Konvergenz der geeignet skalierten internen Pfadlänge zufälliger rekursiver Bäume wurde 1991 von H. M. Mahmoud in [21] gezeigt. Wie bereits erwähnt, ist dieses Baumfunktional invariant unter dem Übergang von rekursiven zu Harris-Bäumen. Der folgende Satz befasst sich mit der fast sicheren Konvergenz der internen Pfadlänge der zufälligen Ulam-Harris-Bäume und liefert zusätzlich die Darstellung der Grenzzufallsvariable als Funktion von X_∞ .

Satz 4.17. *Mit $n \rightarrow \infty$ gilt fast sicher und in L^2 :*

$$\frac{1}{n} \text{IPL}(X_n) - \log n \rightarrow \gamma - 1 + Y_\infty$$

wobei Y_∞ wie in Lemma 4.16 und γ die Eulersche Konstante ist.

Beweis. Wir setzen die Argumentation wie in der Situation binärer Bäume fort. Lemma 4.16 zeigt, dass Y_∞ wohldefiniert ist und ein endliches zweites Moment hat. Folglich ist (Z_n, \mathcal{F}_n) mit $Z_n := E[Y_\infty | \mathcal{F}_n]$, $n \in \mathbb{N}$, ein L^2 -beschränktes Martingal. Das bedeutet, dass Z_n gegen Y_∞ fast sicher und in L^2 mit $n \rightarrow \infty$ konvergiert.

Wir bestimmen nun $E[Y_\infty | \mathcal{F}_n]$. Sei Y_k wie in Lemma 4.16. Wir zeigen zunächst, dass für jedes feste $n \in \mathbb{N}$: $E[Y_k | \mathcal{F}_n] \rightarrow E[Y_\infty | \mathcal{F}_n]$ fast sicher und in L^2 mit $k \rightarrow \infty$ gilt. Die L^2 -Konvergenz ergibt sich aus der Jensen'schen Ungleichung, siehe Korollar 8.20 aus [18]. Für die fast sichere Konvergenz reicht es zu zeigen, dass

$$\sum_{k=1}^{\infty} \|E[Y_k | \mathcal{F}_n] - E[Y_\infty | \mathcal{F}_n]\|_2 < \infty$$

gilt, siehe beispielsweise Satz 6.12. aus [18]. Da die Abbildung $E[\cdot | \mathcal{F}_n]$ eine Kontraktion ist, reicht es, $\|Y_k - Y_\infty\|_2$ zu untersuchen. Als quadratintegrierbares Martingal hat $(Y_k)_{k \in \mathbb{N}}$ unkorrelierte Zuwächse. Also gilt

$$E(Y_{k+l} - Y_k)^2 = \sum_{m=k}^{k+l-1} E(Y_{m+1} - Y_m)^2.$$

Mit dem Lemma von Fatou und der Aussage $E(Y_{k+1} - Y_k)^2 = 2^k 3^{-(k+1)} \kappa$ aus dem Beweis von Lemma 4.16 folgt:

$$E(Y_\infty - Y_k)^2 \leq \sum_{m \geq k} E(Y_{m+1} - Y_m)^2 = \sum_{m \geq k} 2^m 3^{-(m+1)} \kappa = \kappa \left(\frac{2}{3}\right)^k,$$

und somit

$$\sum_{k=1}^{\infty} \|E[Y_k | \mathcal{F}_n] - E[Y_\infty | \mathcal{F}_n]\|_2 \leq \sum_{k=1}^{\infty} \|Y_k - Y_\infty\|_2 < \infty.$$

Wir haben gezeigt, dass für jedes feste $n \in \mathbb{N}$

$$Z_n = E[Y_\infty | \mathcal{F}_n] = \lim_{k \rightarrow \infty} E[Y_k | \mathcal{F}_n] = \lim_{k \rightarrow \infty} \sum_{\substack{u \in \mathbb{N}^* \\ |u|_1 \leq k}} E[X_\infty(A_u)C(\xi_u) | \mathcal{F}_n]$$

gilt. Außerdem sind $X_\infty(A_u)$ und $C(\xi_u)$ bedingt unabhängig gegeben \mathcal{F}_n und es ist $E[C(\xi_u)|\mathcal{F}_n] = 0$ für $u \notin X_n$ nach Lemma 4.15. Insgesamt erhält man:

$$\begin{aligned}
Z_n &= \sum_{u \in \mathbb{N}^*} E[X_\infty(A_u)|\mathcal{F}_n] E[C(\xi_u)|\mathcal{F}_n] \\
&= \sum_{u \in X_n} \frac{\#X_n(u)}{n} \left(1 + \frac{\sum_{uoi \in X_n} \#X_n(u \circ i) H(\#X_n(u \circ i))}{\#X_n(u)} - H(\#X_n(u)) \right) \\
&= \frac{1}{n} (\text{IPL}(X_n) + n) - \frac{1}{n} \sum_{u \in X_n} \left(\#X_n(u) H(\#X_n(u)) - \sum_{uoi \in X_n} \#X_n(u \circ i) H(\#X_n(u \circ i)) \right) \\
&= \frac{1}{n} (\text{IPL}(X_n) + n) - \frac{1}{n} (nH(n)) \\
&= \frac{1}{n} (\text{IPL}(X_n) + n) - H(n).
\end{aligned}$$

Bei der dritten Gleichheit wurde ein Analogon zum Lemma 5 aus [13] verwendet: Für jeden Ulam-Harris-Baum \mathbf{t} und jede Funktion $\varphi : \mathbb{N}^* \rightarrow \mathbb{R}$ mit $\varphi(u) = 0$ für alle $u \in \partial \mathbf{t}$ gilt

$$\sum_{u \in \mathbf{t}} \left(\varphi(u) - \sum_{uoi \in \mathbf{t}} \varphi(u \circ i) \right) = \varphi(\Lambda).$$

Dieses wurde in Verbindung mit $\#X_n(\Lambda)H(\#X_n(\Lambda)) = nH(n)$ und $\#X_n(u)H(\#X_n(u)) = 0$ für alle $u \in \partial X_n$ genutzt. Der Rest folgt nun aus der bekannten Tatsache

$$\lim_{n \rightarrow \infty} (H(n) - \log n) = \gamma. \quad \square$$

Bemerkung 4.18. Die obigen Überlegungen liefern eine weitere Möglichkeit zur Bestimmung von $E(\text{IPL}(X_n))$: Da $E(Y_\infty) = 0$ ist, gilt $EZ_n = \frac{1}{n}(E(\text{IPL}(X_n)) + n) - H(n) = E[E[Y_\infty|\mathcal{F}_n]] = 0$, also

$$E(\text{IPL}(X_n)) = n(H(n) - 1)$$

in Übereinstimmung mit dem bereits bekannten Resultat, siehe beispielsweise [27, S.6]¹⁴.

4.5.2 Wiener-Index

Unter dem Wiener-Index eines Ulam-Harris-Baumes \mathbf{t} versteht man

$$\text{WI}(\mathbf{t}) := \frac{1}{2} \sum_{(u,v) \in \mathbf{t} \times \mathbf{t}} d(u,v),$$

wobei $d(u,v)$ den kanonischen Graph-Abstand zweier Knoten u und v bezeichnet. Dieser ist definiert als die Länge des kürzesten Weges von u nach v und kann wie folgt berechnet werden: $d(u,v) = |u| + |v| - 2|u \wedge v|$. Dabei ist $u \wedge v$ der längste gemeinsame Präfix beziehungsweise letzter gemeinsamer Vorfahr von u und v .

Lemma 4.19. *Für den Wiener-Index eines Ulam-Harris-Baumes \mathbf{t} mit $\#\mathbf{t} = n$ gilt:*

$$\text{WI}(\mathbf{t}) = n\text{IPL}(\mathbf{t}) + n^2 - \sum_{u \in \mathbf{t}} \#\mathbf{t}_n(u)^2.$$

¹⁴Die Definition der IPL entspricht der Definition der „total path length“ (TPL) in [27].

Beweis. Für einen Ulam-Harris-Baum \mathbf{t} mit $\#\mathbf{t} = n$ gilt:

$$\begin{aligned} \text{WI}(\mathbf{t}) &= \frac{1}{2} \sum_{(u,v) \in \mathbf{t} \times \mathbf{t}} d(u,v) \\ &= \frac{1}{2} \left(\sum_{u,v \in \mathbf{t}} |u| + \sum_{u,v \in \mathbf{t}} |v| - 2 \sum_{u,v \in \mathbf{t}} |u \wedge v| \right) \\ &= n\text{IPL}(\mathbf{t}) - \sum_{u,v \in \mathbf{t}} |u \wedge v|. \end{aligned}$$

Es bleibt also zu zeigen:

$$\sum_{u,v \in \mathbf{t}} |u \wedge v| = \sum_{u \in \mathbf{t}} \#\mathbf{t}(u)^2 - n^2.$$

Das kann wie im Fall binärer Bäume mit Hilfe des Induktionsprinzips gezeigt werden (vergleiche [3, S.70]). Die Aussage gilt offenbar für $\#\mathbf{t} = 1$. Wir nehmen nun an, die Aussage wäre für alle \mathbf{t} mit $\#\mathbf{t} < n$ bereits gezeigt. Beim Induktionsschritt wird eine für Ulam-Harris-Bäume geeignete Zerlegung benötigt. Durch das Entfernen der Kante zwischen der Wurzel Λ und dem Knoten 1 zerfällt \mathbf{t} in folgende zwei Teilbäume, siehe das Beispiel in Abbildung 9:

$$\mathbf{t}^b = \{u \in \mathbb{N}^* : 1 \circ u \in \mathbf{t}\}$$

und

$$\mathbf{t}^\sharp = \{\Lambda\} \cup \{u \in \mathbb{N}^* \setminus \{\Lambda\} : (1 + u_1) \circ u_2 \dots u_{|u|} \in \mathbf{t}\}.$$

Die Mengen \mathbf{t}^b und \mathbf{t}^\sharp sind wieder Ulam-Harris-Bäume im Sinne der Definition 4.3. Informell ausgedrückt ist dies die endliche Version der Zerlegung aus der Bemerkung 4.10.

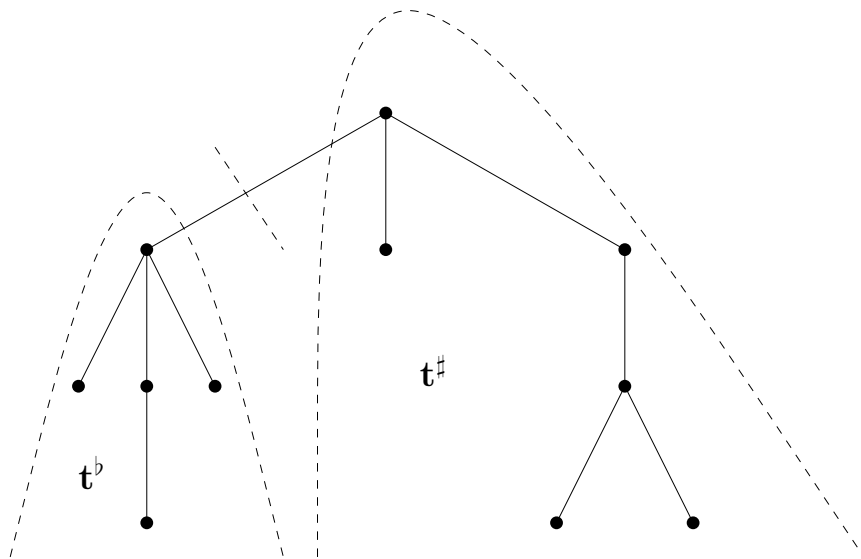


Abbildung 9: Durch das Entfernen der durchgestrichelten Kante erhalten wir $\mathbf{t}^b = \{\Lambda, 1, 2, 21, 3\}$ und $\mathbf{t}^\sharp = \{\Lambda, 1, 2, 21, 211, 212\}$.

Wir erhalten:

$$\begin{aligned}
\sum_{u,v \in \mathbf{t}} |u \wedge v| &= \sum_{u,v \in \mathbf{t}^b} (|u \wedge v| + 1) + \sum_{u,v \in \mathbf{t}^\sharp} |u \wedge v| \\
&= (\#\mathbf{t}^b)^2 + \sum_{u,v \in \mathbf{t}^b} |u \wedge v| + \sum_{u,v \in \mathbf{t}^\sharp} |u \wedge v| \\
&= \sum_{u \in \mathbf{t}^b} (\#\mathbf{t}^b(u))^2 + \sum_{u \in \mathbf{t}^\sharp} (\#\mathbf{t}^\sharp(u))^2 - (\#\mathbf{t}^\sharp)^2 \\
&= \left(\sum_{u \in \mathbf{t}^b} (\#\mathbf{t}^b(u))^2 + \sum_{u \in \mathbf{t}^\sharp} (\#\mathbf{t}^\sharp(u))^2 + 2\#\mathbf{t}^b\#\mathbf{t}^\sharp + (\#\mathbf{t}^b)^2 \right) \\
&\quad - 2\#\mathbf{t}^b\#\mathbf{t}^\sharp - (\#\mathbf{t}^b)^2 - (\#\mathbf{t}^\sharp)^2 \\
&= \sum_{u \in \mathbf{t}} \#\mathbf{t}(u)^2 - n^2,
\end{aligned}$$

wobei die Induktionsvoraussetzung bei der dritten Gleichheit verwendet wurde. Bei den letzten beiden Umformungen ist zu beachten:

$$n = \#\mathbf{t} = \#\mathbf{t}(\Lambda) = \#\mathbf{t}^b + \#\mathbf{t}^\sharp. \quad \square$$

Satz 4.20. *Mit $n \rightarrow \infty$ gilt fast sicher und in L^2 :*

$$\frac{1}{n^2} \text{WI}(X_n) - \log n \rightarrow \gamma - 1 + Y_\infty - Z_\infty$$

wobei Y_∞ wie im Satz 4.17 ist und

$$Z_\infty = \sum_{u \in \mathbb{N}^*} X_\infty(A_u)^2.$$

Beweis. Wir zeigen zunächst $EZ_\infty^2 < \infty$, sodass $Z_n := E[Z_\infty | \mathcal{F}_n] \rightarrow Z_\infty$ fast sicher und in L^2 gilt. Sei dazu

$$\tilde{Z}_k = \sum_{\substack{u \in \mathbb{N}^* \\ |u|_1 \leq k}} X_\infty(A_u)^2.$$

Mit der Abschätzung

$$\begin{aligned}
\left\| \tilde{Z}_{k+1} - \tilde{Z}_k \right\|_2 &\leq \sum_{\substack{u \in \mathbb{N}^* \\ |u|_1 = k+1}} \|X_\infty(A_u)^2\|_2 = \sum_{\substack{u \in \mathbb{N}^* \\ |u|_1 = k+1}} \sqrt{E X_\infty(A_u)^4} \\
&= \sum_{\substack{u \in \mathbb{N}^* \\ |u|_1 = k+1}} 5^{-\frac{k+1}{2}} = 2^k 5^{-\frac{k+1}{2}}
\end{aligned}$$

erhalt man, dass $(\tilde{Z}_k)_{k \in \mathbb{N}}$ eine Cauchy-Folge in L^2 ist und insbesondere $EZ_\infty^2 < \infty$.

Nun zum Beweis der Hauptaussage: Die Anwendung der bedingten Version des Satzes von der monotonen Konvergenz ergibt

$$\begin{aligned}
Z_n &= E \left[\sum_{u \in \mathbb{N}^*} X_\infty(A_u)^2 | \mathcal{F}_n \right] = \sum_{u \in \mathbb{N}^*} E [X_\infty(A_u)^2 | \mathcal{F}_n] \\
&= \sum_{u \in X_n} E [X_\infty(A_u)^2 | \mathcal{F}_n] + \sum_{u \notin X_n} E [X_\infty(A_u)^2 | \mathcal{F}_n]
\end{aligned}$$

für jedes feste $n \in \mathbb{N}$. Es ist $X_\infty(A_u)^2 = \prod_{\Lambda < v \leq u} \xi_{\bar{v}, \varepsilon(v)}^2$. Für $u \in X_n$ erhält man mit der Formel für das zweite Moment der Betaverteilung:

$$\begin{aligned} E [X_\infty(A_u)^2 | \mathcal{F}_n] &= \prod_{\Lambda < v \leq u} E [\xi_{\bar{v}, \varepsilon(v)}^2 | \mathcal{F}_n] \\ &= \prod_{\Lambda < v \leq u} \frac{\#X_n(v)(\#X_n(v) + 1)}{(1 + \sum_{\bar{v}i \in X_n} \#X_n(\bar{v}i))(2 + \sum_{\bar{v}i \in X_n} \#X_n(\bar{v}i))} \\ &= \prod_{\Lambda < v \leq u} \frac{\#X_n(v)(\#X_n(v) + 1)}{\#X_n(\bar{v})(\#X_n(\bar{v}) + 1)}. \end{aligned}$$

Also gilt für $u \in X_n$:

$$E [X_\infty(A_u)^2 | \mathcal{F}_n] = \frac{\#X_n(u)(\#X_n(u) + 1)}{n(n + 1)}.$$

Somit hat man

$$\begin{aligned} \sum_{u \in X_n} E [X_\infty(A_u)^2 | \mathcal{F}_n] &= \frac{1}{n(n + 1)} \sum_{u \in X_n} \#X_n(u)(\#X_n(u) + 1) \\ &= \frac{1}{n(n + 1)} \left(\sum_{u \in X_n} \#X_n(u)^2 + \text{IPL}(X_n) + n \right). \end{aligned}$$

Für einen zu X_n externen Knoten u gilt:

$$\begin{aligned} E [X_\infty(A_u)^2 | \mathcal{F}_n] &= E [X_\infty(A_{\bar{u}})^2 | \mathcal{F}_n] E [\xi_{\bar{u}, \varepsilon(u)}^2 | \mathcal{F}_n] \\ &= \frac{\#X_n(\bar{u})(\#X_n(\bar{u}) + 1)}{n(n + 1)} E [\xi_{\bar{u}, \varepsilon(u)}^2 | \mathcal{F}_n], \end{aligned}$$

da der direkte Vorgänger \bar{u} von u zu X_n gehört. Die Verteilung der $\xi_{\bar{u}, \varepsilon(u)}$ unter \mathcal{F}_n ist dank den Betrachtungen zum China-Restaurant-Prozess im Zusammenhang mit den Teilbaumgrößen bekannt (siehe Zusammenfassung am Anfang des Abschnittes 4.5), also ist

$$E [\xi_{\bar{u}, \varepsilon(u)}^2 | \mathcal{F}_n] = \frac{2}{\#X_n(\bar{u})(\#X_n(\bar{u}) + 1)} \left(\frac{1}{3} \right).$$

Wir erhalten für $u \in \partial X_n$:

$$E [X_\infty(A_u)^2 | \mathcal{F}_n] = \frac{2}{n(n + 1)} \left(\frac{1}{3} \right).$$

Zu jedem $v \notin X_n$ existiert genau ein externer Knoten $u \in \partial X_n$, sodass $v = \bar{u} \circ (\varepsilon(u) + j) \circ w$ mit einem $j \geq 0$ und $w \in \mathbb{N}^*$ gilt. Für v von dieser Form gilt also

$$E [X_\infty(A_v)^2 | \mathcal{F}_n] = \frac{2}{n(n + 1)} \left(\frac{1}{3} \right)^{1+j+|w|_1}.$$

Für das leere Wort gilt $|\Lambda|_1 = 0$ und es gibt 2^{k-1} Wörter mit $|w|_1 = k$, $k \geq 1$, siehe Überlegungen zu den Zahl-Partitionen im Beweis des Satzes 4.17. Durch das Aufspalten

nach $u \in \partial X_n$ (siehe Abbildung 10) erhält man

$$\begin{aligned}
\sum_{u \notin X_n} E[X_\infty(A_u)^2 | \mathcal{F}_n] &= \sum_{u \in \partial X_n} \sum_{j=0}^{\infty} \sum_{w \in \mathbb{N}^*} E[X_\infty(A_{\bar{u} \circ (\varepsilon(u)+j) \circ w})^2 | \mathcal{F}_n] \\
&= \frac{2}{n(n+1)} \sum_{u \in \partial X_n} \sum_{j=0}^{\infty} \left(\frac{1}{3}\right)^{j+1} \left(1 + \sum_{k=1}^{\infty} 2^{k-1} \left(\frac{1}{3}\right)^k\right) \\
&= \frac{4}{n(n+1)} \sum_{u \in \partial X_n} \sum_{j=0}^{\infty} \left(\frac{1}{3}\right)^{j+1} \\
&= \frac{4}{n(n+1)} \sum_{u \in \partial X_n} \frac{1}{2} = \frac{2}{n+1}.
\end{aligned}$$

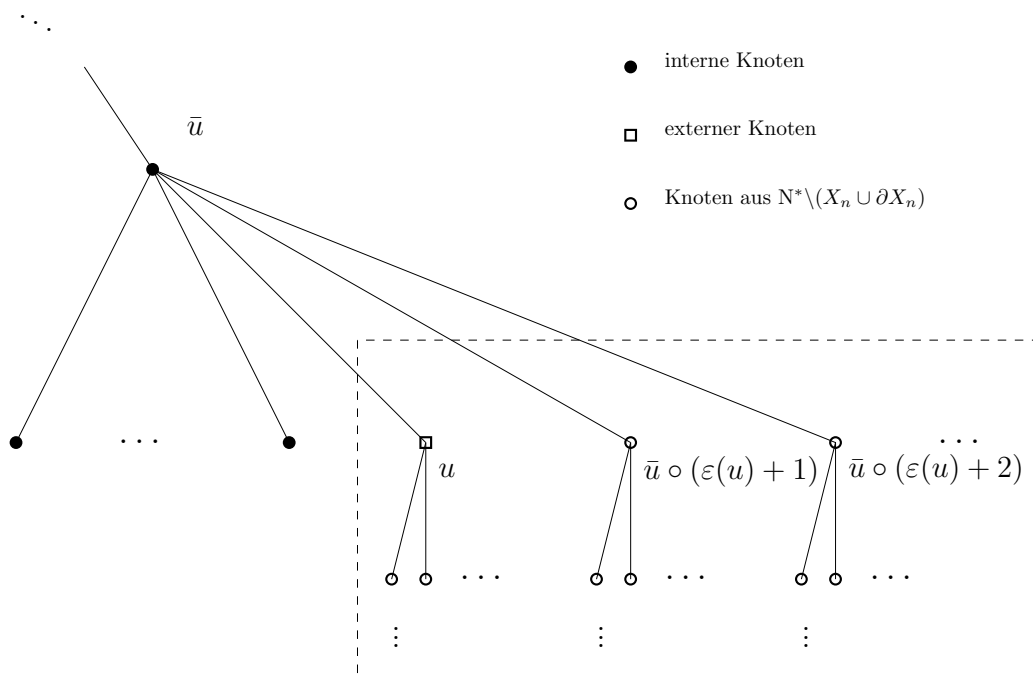


Abbildung 10: Die Summe $\sum_{u \notin X_n} E[X_\infty(A_u)^2 | \mathcal{F}_n]$ wird nach $u \in \partial X_n$ aufgespalten. Im Kasten befinden sich die Elemente, die zu dem Knoten $u \in \partial X_n$ gehörenden Summanden zusammengefasst werden.

Insgesamt:

$$Z_n = E[Z_\infty | \mathcal{F}_n] = \frac{1}{n(n+1)} \left(\sum_{u \in X_n} \#X_n(u)^2 + \text{IPL}(X_n) + n \right) + \frac{2}{n+1}.$$

Schließlich hat man, siehe Lemma 4.19,

$$\text{WI}(X_n) = n \text{IPL}(X_n) + n^2 - n(n+1)Z_n + \text{IPL}(X_n) + 3n.$$

Mit dem Satz 4.17 über die interne Pfadlänge folgt die Behauptung. □

Die fast sichere Konvergenz impliziert bekanntlich die Konvergenz in Verteilung. Die Konvergenz in Verteilung des geeignet skalierten Wiener-Index zufälliger rekursiver Bäume wurde 2002 von R. Neininger in [22] erhalten.

Bemerkung 4.21. Auch hier ist die Darstellung des Wiener-Index aus dem Beweis dazu geeignet, um $E(\text{WI}(X_n))$ zu bestimmen. Mit obigen Überlegungen folgt, dass $E(Z_\infty) = 1 + \sum_{k=1}^{\infty} 2^{k-1} (\frac{1}{3})^k = 2$ gilt, und somit ist auch $EZ_n = E(E[Z_\infty | \mathcal{F}_n]) = 2$. Wir erhalten

$$\begin{aligned} E(\text{WI}(X_n)) &= nE(\text{IPL}(X_n)) + n^2 - n(n+1)EZ_n + E(\text{IPL}(X_n)) + 3n \\ &= n(nH(n) - n) + n^2 - 2n(n+1) + (nH(n) - n) + 3n \\ &= n^2H(n) + nH(n) - 2n^2 \end{aligned}$$

in Übereinstimmung mit dem bereits bekannten Resultat für $E(\text{WI}(X_n))$, vergleiche Satz 1.2 aus [22].

4.5.3 Horizontale Pfadlänge und Konstruktionsaufwand

Wir definieren die horizontale Pfadlänge eines Ulam-Harris-Baumes \mathbf{t} als

$$\text{HPL}(\mathbf{t}) := \sum_{u \in \mathbf{t}} (|u|_1 - |u|) = \sum_{u \in \mathbf{t}} |u|_1 - \text{IPL}(\mathbf{t})$$

mit $|u|_1 = u_1 + \dots + u_n$ für $u = u_1 \dots u_n \in \mathbb{N}^*$. Hierbei interpretieren wir $|u|_1 - |u|$ als die horizontale Position eines Knotens u , siehe Abbildung 11.

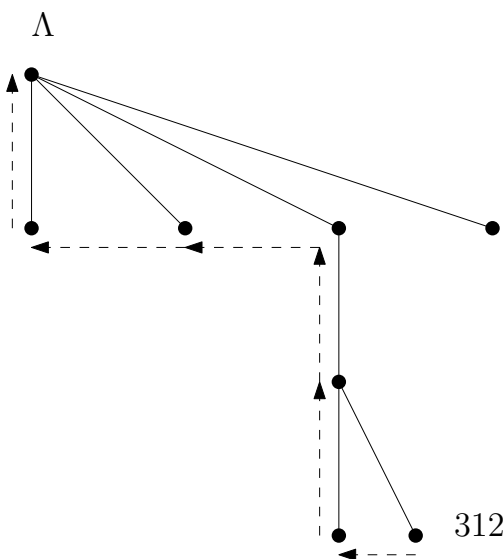


Abbildung 11: Die horizontale Position des Knotens 312 ist $|312|_1 - |312| = 3$. Sie entspricht der Anzahl der Schritte nach links im eingezeichneten Pfad von 312 zur Wurzel Λ .

Die horizontale Pfadlänge lässt sich mit Hilfe der folgenden Aussage aus den Teilbaumgrößen berechnen.

Lemma 4.22. Sei \mathbf{t} ein Ulam-Harris-Baum. Dann gilt:

$$\sum_{u \in \mathbf{t}} |u|_1 = \sum_{u \in \mathbf{t}} \sum_{uoi \in \mathbf{t}} i \# \mathbf{t}(u \circ i).$$

Beweis. Wir verwenden wieder das Induktionsprinzip. Die Aussage gilt für $\#\mathbf{t} = 1$. Für den Induktionsschritt erinnern wir an die Zerlegung von \mathbf{t} in Teilbäume

$$\mathbf{t}^b = \{u \in \mathbb{N}^* : 1 \circ u \in \mathbf{t}\}$$

und

$$\mathbf{t}^\# = \{\Lambda\} \cup \{u \in \mathbb{N}^* \setminus \{\Lambda\} : (1 + u_1, u_2, \dots, u_{|u|}) \in \mathbf{t}\}.$$

Sei die Aussage für kleinere Bäume bereits gezeigt. Die Anwendung der Induktionsvoraussetzung bei der zweiten Umformung ergibt

$$\begin{aligned} \sum_{u \in \mathbf{t}} |u|_1 &= \sum_{u \in \mathbf{t}^b} |u|_1 + \#\mathbf{t}^b + \sum_{u \in \mathbf{t}^\#} |u|_1 + (\#\mathbf{t}^\# - 1) \\ &= \sum_{u \in \mathbf{t}^b} \sum_{u \circ i \in \mathbf{t}^b} i \#\mathbf{t}^b(u \circ i) + \sum_{u \in \mathbf{t}^\#} \sum_{u \circ i \in \mathbf{t}^\#} i \#\mathbf{t}^\#(u \circ i) + (\#\mathbf{t} - 1) \\ &= \sum_{u \in \mathbf{t}} \sum_{u \circ i \in \mathbf{t}} i \#\mathbf{t}(u \circ i), \end{aligned}$$

denn aus der Definition der Zerlegung folgt

$$\sum_{u \in \mathbf{t}} \sum_{u \circ i \in \mathbf{t}} i \#\mathbf{t}(u \circ i) = \sum_{u \in \mathbf{t}^b} \sum_{u \circ i \in \mathbf{t}^b} i \#\mathbf{t}(u \circ i) + \sum_{u \in \mathbf{t}^\#} \sum_{u \circ i \in \mathbf{t}^\#} i \#\mathbf{t}(u \circ i) + \sum_{j \in \mathbb{N} \cap \mathbf{t}} \#\mathbf{t}(j). \quad \square$$

Warum ist die horizontale Pfadlänge relevant? Der Konstruktionsaufwand („construction time“) $\text{CT}(X_n)$ für einen zufälligen Ulam-Harris-Baum X_n zu $n - 1$ Daten, gemessen als Anzahl der benötigten Vergleiche, ist

$$\text{CT}(X_n) = \sum_{u \in X_n} |u|_1 - n.$$

Der Konstruktionsaufwand kann also bis auf einen Korrekturterm in die Beiträge der horizontalen und der internen Pfadlänge zerlegt werden:

$$\text{CT}(X_n) = \text{HPL}(X_n) + \text{IPL}(X_n) - n.$$

Lemma 4.23. *Es sei*

$$\tilde{C} : \Delta^\infty \rightarrow [-2, +\infty) \quad \tilde{C}(s) := -2 + \sum_{i=1}^{\infty} i s_i.$$

Dann hat die Zufallsvariable $\tilde{C}(\xi_u)$ für jedes $u \in \mathbb{N}^$ ein endliches zweites Moment. Außerdem gilt für $u \notin X_n$*

$$E \left[\tilde{C}(\xi_u) | \mathcal{F}_n \right] = 0$$

und für $u \in X_n$

$$E \left[\tilde{C}(\xi_u) | \mathcal{F}_n \right] = -2 + \frac{\sum_{u \circ i \in X_n} i \#X_n(u \circ i)}{\#X_n(u)} + \frac{\#\{k \in \mathbb{N} : u \circ k \in X_n\} + 2}{\#X_n(u)}.$$

Beweis. Zur ersten Aussage: Wir betrachten die Partialsummen $\tilde{C}_k(\xi_u) := -2 + \sum_{i=1}^k i \xi_{u,i}$. Da ξ_u für jedes $u \in \mathbb{N}^*$ GEM-verteilt ist, gilt für $k \in \mathbb{N}$

$$\left\| \tilde{C}_k(\xi_u) - \tilde{C}_{k-1}(\xi_u) \right\|_2^2 = \|k \xi_{u,k}\|_2^2 = k^2 \left(\frac{1}{3} \right)^k.$$

Wie im Lemma 4.15 folgt daraus, dass die Partialsummen $\tilde{C}_k(\xi_u)$ eine Cauchy-Folge in L^2 bilden, insbesondere ist $E\tilde{C}_k(\xi_u)^2 < \infty$.

Für $u \notin X_n$ ist ξ_u unter \mathcal{F}_n GEM - verteilt. Also:

$$E [i \xi_{u,i} | \mathcal{F}_n] = i \left(\frac{1}{2} \right)^i$$

und

$$-2 + \sum_{n=1}^{\infty} n \left(\frac{1}{2} \right)^n = 0.$$

Also folgt die zweite Aussage.

Für die dritte Aussage sei wieder $\#X_n(u \circ 1) = a_1, \dots, \#X_n(u \circ k) = a_k, \#X_n(u \circ j) = 0$ für $j > k$, wobei $a_1, \dots, a_k > 0$. Bekanntlich gilt für $i \leq k$:

$$\mathcal{L}(\xi_{u,i} | \mathcal{F}_n) = \text{Beta} \left(a_i, 1 + \sum_{\substack{j=1 \\ j \neq i}}^k a_j \right).$$

Mit den Rechenregeln für Betaverteilungen folgt:

$$E \left[\sum_{i=1}^k i \xi_{u,i} | \mathcal{F}_n \right] = \sum_{i=1}^k \frac{i a_i}{1 + \sum_{j=1}^k a_j}.$$

Wie schon dargelegt, hat $\xi_{u,i}$ im Fall $i > k$ unter \mathcal{F}_n dieselbe Verteilung wie die Zufallsvariable YZ , Y und Z unabhängig, wobei $Y \sim \text{Beta}(1, \sum_{j=1}^k a_j)$ und Z ein Produkt von $i - k$ unabhängigen $\text{unif}(0, 1)$ -verteilten Zufallsgrößen ist. Es gilt:

$$E i Y Z = \frac{1}{1 + \sum_{j=1}^k a_j} i \left(\frac{1}{2} \right)^{i-k}$$

Aufsummieren liefert:

$$\frac{1}{1 + \sum_{j=1}^k a_j} \sum_{i=k+1}^{\infty} i \left(\frac{1}{2} \right)^{i-k} = \frac{k+2}{1 + \sum_{j=1}^k a_j}.$$

Insgesamt hat man:

$$\sum_{i=1}^k \frac{i a_i}{1 + \sum_{j=1}^k a_j} + \frac{k+2}{1 + \sum_{j=1}^k a_j}.$$

Damit ist die dritte Aussage gezeigt. □

Satz 4.24. *Mit $n \rightarrow \infty$ gilt fast sicher und in L^2 :*

$$\frac{1}{n} \text{CT}(X_n) - 2 \log n \rightarrow H_\infty + 2Y_\infty + 2\gamma - 4$$

wobei Y_∞ wie im Satz 4.17 und

$$H_\infty = \sum_{u \in \mathbb{N}^*} X_\infty(A_u) \tilde{C}(\xi_u).$$

Beweis. Wir stellen fest, dass H_∞ wie Y_∞ aufgebaut ist, wobei die Hilfsfunktion C durch die Hilfsfunktion \tilde{C} mit gleichen Eigenschaften ersetzt wurde. Um zu zeigen, dass der Grenzwert H_∞ fast sicher und in L^2 existiert, lässt sich also der Beweis vom Lemma 4.16 direkt übertragen. Der weitere Beweis läuft analog zum Beweis des Satzes 4.17 über die interne Pfadlänge: Mit $n \rightarrow \infty$ konvergiert $V_n := E[H_\infty | \mathcal{F}_n]$ fast sicher und in L^2 gegen H_∞ . Außerdem kann man mit den Argumenten aus dem Satz 4.17 zeigen:

$$E \left[\sum_{u \in \mathbb{N}^*} X_\infty(A_u) \tilde{C}(\xi_u) | \mathcal{F}_n \right] = \sum_{u \in \mathbb{N}^*} E[X_\infty(A_u) | \mathcal{F}_n] E[\tilde{C}(\xi_u) | \mathcal{F}_n]$$

für jedes feste $n \in \mathbb{N}$. Wir erhalten:

$$\begin{aligned} E[H_\infty | \mathcal{F}_n] &= \sum_{u \in \mathbb{N}^*} E[X_\infty(A_u) | \mathcal{F}_n] E[\tilde{C}(\xi_u) | \mathcal{F}_n] \\ &= \sum_{u \in X_n} \frac{\#X_n(u)}{n} \left(-2 + \frac{\sum_{u \circ i \in X_n} i \#X_n(u \circ i)}{\#X_n(u)} + \frac{\#\{k \in \mathbb{N} : u \circ k \in X_n\} + 2}{\#X_n(u)} \right) \\ &= -\frac{2}{n} \sum_{u \in X_n} \#X_n(u) + \frac{1}{n} \sum_{u \in X_n} \sum_{u \circ i \in X_n} i \#X_n(u \circ i) \\ &\quad + \frac{1}{n} \sum_{u \in X_n} (\#\{k \in \mathbb{N} : u \circ k \in X_n\}) + \frac{1}{n} \sum_{u \in X_n} 2 \\ &= -\frac{2}{n} (\text{IPL}(X_n) + n) + \frac{1}{n} (\text{CT}(X_n) + n) + \frac{(n-1)}{n} + 2 \\ &= \frac{1}{n} \text{CT}(X_n) - \frac{2}{n} \text{IPL}(X_n) + \frac{(n-1)}{n} + 1. \end{aligned}$$

Aus dem Satz 4.17 ist bekannt:

$$\frac{1}{n} \text{IPL}(X_n) - \log n \rightarrow \gamma - 1 + Y_\infty$$

mit $n \rightarrow \infty$ gilt fast sicher und in L^2 . Insgesamt:

$$\frac{1}{n} \text{CT}(X_n) - 2 \log n \rightarrow H_\infty + 2Y_\infty + 2\gamma - 4. \quad \square$$

Korollar 4.25. *Mit $n \rightarrow \infty$ gilt fast sicher und in L^2 :*

$$\frac{1}{n} \text{HPL}(X_n) - \log n \rightarrow H_\infty + Y_\infty + \gamma - 2$$

Beweis. Es ist

$$\frac{1}{n}\text{HPL}(X_n) - \log n = \frac{1}{n}\text{CT}(X_n) - 2 \log n - \left(\frac{1}{n}\text{IPL}(X_n) - \log n \right) + 1.$$

Wende nun den obigen Satz an. □

Bemerkung 4.26. Wir bestimmen nun $E(\text{CT}(X_n))$ und $E(\text{HPL}(X_n))$. Es ist $EH_\infty = 0$, also auch $EV_n = E(E[H_\infty|\mathcal{F}_n]) = 0$. Somit gilt für den erwarteten Konstruktionsaufwand:

$$\begin{aligned} E(\text{CT}(X_n)) &= 2E(\text{IPL}(X_n)) - (n-1) - n \\ &= 2n(H(n) - 1) - 2n + 1 \\ &= 2nH(n) - 4n + 1. \end{aligned}$$

Für die horizontale Pfadlänge ergibt sich:

$$\begin{aligned} E(\text{HPL}(X_n)) &= E(\text{CT}(X_n)) - E(\text{IPL}(X_n)) + n \\ &= nH(n) - 2n + 1. \end{aligned}$$

5 Die Rolle der natürlichen Korrespondenz

In diesem Kapitel befassen wir uns mit der Bedeutung der natürlichen Korrespondenz zwischen den binären Bäumen und den Ulam-Harris-Bäumen. Dazu erinnern wir zunächst an einige Grundbegriffe.

Definition 5.1. *Ein binärer Baum ist eine präfixstabile endliche Teilmenge von $\{0, 1\}^*$.*

Es sei \mathbb{B}_n die Menge der binären Bäume mit n Knoten und $\mathbb{B} := \bigcup_{n=1}^{\infty} \mathbb{B}_n$. Binäre Bäume werden als Datenstruktur zur Behandlung von Suchproblemen eingesetzt. Die Daten werden dabei Schritt für Schritt in die Knoten des Baumes mit Hilfe einer Zuordnung gespeichert, was wieder auf das Konzept bewerteter Bäume führt. Gegeben sei eine Folge $(x_i)_{i \in \mathbb{N}}$ von paarweise verschiedenen Punkten $x_i \in \mathbb{R}$. Der „binary search tree“ (BST) - Algorithmus¹⁵ konstruiert dazu folgendermaßen eine Folge von bewerteten Bäumen (z_n, φ_n) , $n \in \mathbb{N}$:

- Für $n = 1$ gelte $z_1 = \{\Lambda\}$ und $\varphi_1(\Lambda) = x_1$, das heißt, der erste Baum besteht nur aus dem Wurzelknoten bewertet mit x_1 .
- Sei (z_n, φ_n) schon konstruiert. Den Baum z_{n+1} erhält man durch Umwandlung eines externen Knoten von z_n in einen internen. Man vergleicht x_{n+1} nacheinander mit dem Inhalt $\varphi_n(u)$ eines Knoten u , beginnend mit dem Wurzelknoten. Bei $x_{n+1} < \varphi_n(u)$ geht man zu $u \circ 0$ über und zu $u \circ 1$ sonst. Die Prozedur wird wiederholt, bis man einen externen Knoten v gefunden hat. Dieser wird in einen internen Knoten umgewandelt und mit x_{n+1} bewertet. Man erhält $z_{n+1} = z_n \cup \{v\}$, $\varphi_{n+1}(v) = x_{n+1}$ und $\varphi_{n+1}(u) = \varphi_n(u)$ für alle $u \in z_n$.

Abbildung 12 veranschaulicht die Funktionsweise des BST-Algorithmus unter Verwendung der Daten aus dem Abschnitt 4.1 über Ulam-Harris-Bäume, siehe Abbildung 6. Die Anwendung des BST-Algorithmus auf eine Inputfolge $(\eta_i)_{i \in \mathbb{N}}$ unabhängiger $\text{unif}(0, 1)$ -verteilter Zufallsvariablen liefert eine Folge zufälliger bewerteter binärer Bäume. Ohne Berücksichtigung der Bewertungen handelt es sich dabei um eine kombinatorische Markov-Kette mit Zustandsraum \mathbb{B} und folgendem Übergangsmechanismus: Beim Übergang vom Schritt n zu $n + 1$ wählt man einen externen Knoten zufällig und gleichverteilt aus und wandelt diesen in einen internen Knoten um. Dieser stochastische Prozess wird als BST-Prozess bezeichnet.

Wir betrachten eine bijektive Abbildung zwischen \mathbb{H}_{n+1} und \mathbb{B}_n , die in der Literatur als „natural rotation“ oder als „rotation correspondence“ [19, §2.3.2] [11, S. 73] bekannt ist. Zu einem Ulam-Harris Baum $t \neq \{\Lambda\}$ erhält man wie folgt einen binären Baum $R(t)$:

- Die Wurzel und alle dazu inzidenten Kanten werden gelöscht,
- die aufeinanderfolgenden Geschwister einer Familie werden durch Kanten verbunden,
- die Kanten des Ulam-Harris-Baumes werden entfernt mit Ausnahme der Kanten zwischen den Knoten und ihren ersten direkten Nachfolgern,

¹⁵Die Beschreibung des BST-Algorithmus ist aus [13, §2.2] entnommen.

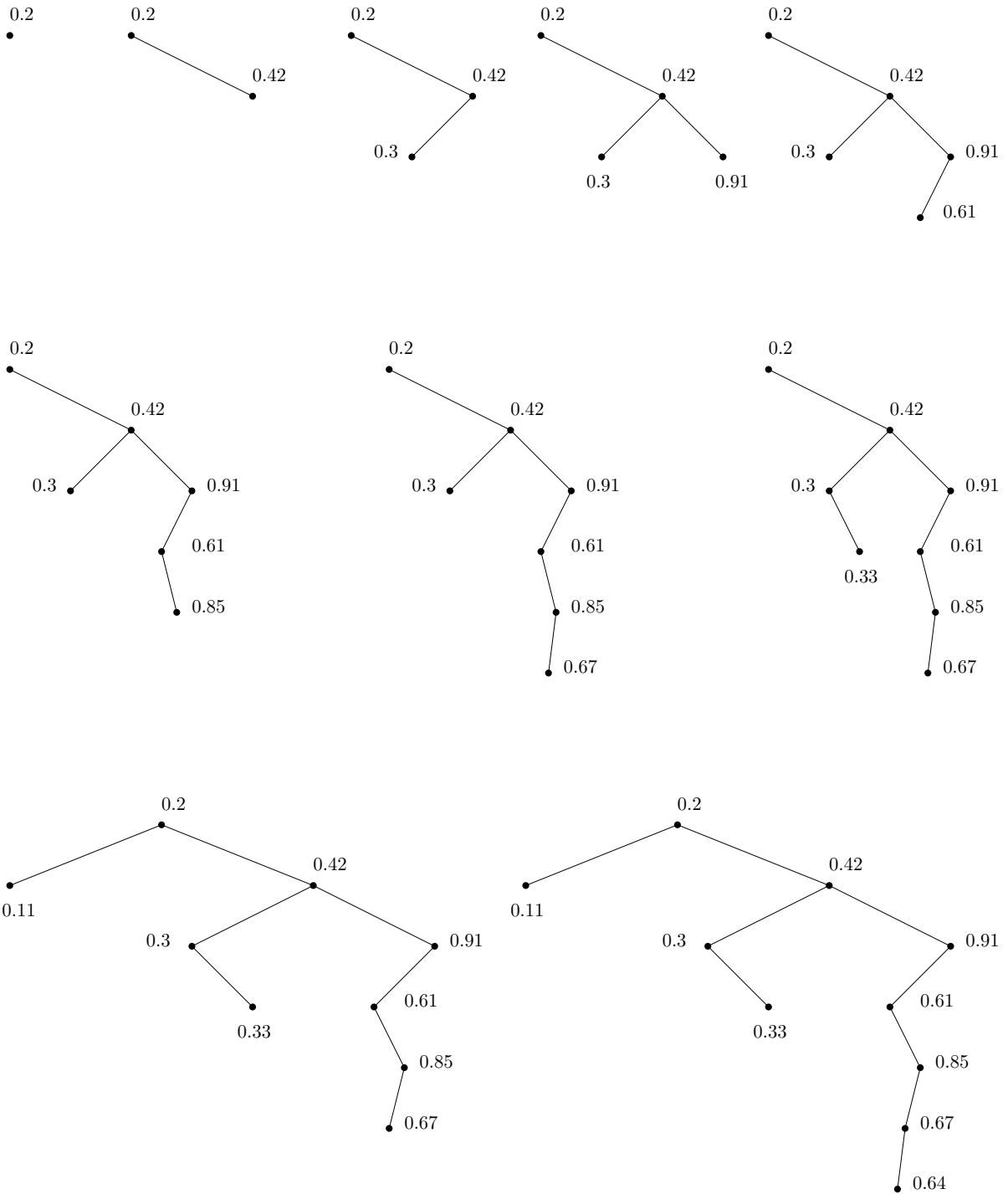


Abbildung 12: Bewertete binäre Bäume, die als Output des BST-Algorithmus zu den Daten 0.2, 0.42, 0.3, 0.91, 0.61, 0.85, 0.67, 0.33, 0.11, 0.64 entstehen. Die Beschriftung der Knoten mit den Wörtern aus $\{0, 1\}^*$ wurde weggelassen.

- das Bild wird schließlich um 45° gedreht.

Abbildung 13 veranschaulicht diese Prozedur, die im Folgenden als natürliche Korrespondenz bezeichnet wird, graphisch. Wir geben nun eine präzise Beschreibung dazu an. Wir definieren zunächst die Abbildung $R : \mathbb{N}^* \setminus \{\Lambda\} \rightarrow \{0, 1\}^*$ durch $R(1) := \Lambda$ und falls $u \neq 1$:

$$R(u) := \begin{cases} R(u_1 \dots u_k) \circ 0, & \text{für } u = u_1 \dots u_k \circ 1, \\ R(u_1 \dots u_k \circ j) \circ 1, & \text{für } u = u_1 \dots u_k \circ (j + 1), j > 0. \end{cases}$$

Durch

$$R(\mathbf{t}) := \{R(u) : u \in \mathbf{t}, u \neq \Lambda\}$$

für Bäume $\mathbf{t} \in \mathbb{H}$, $\mathbf{t} \neq \{\Lambda\}$, wird R zur Bijektion zwischen \mathbb{H}_{n+1} und \mathbb{B}_n .

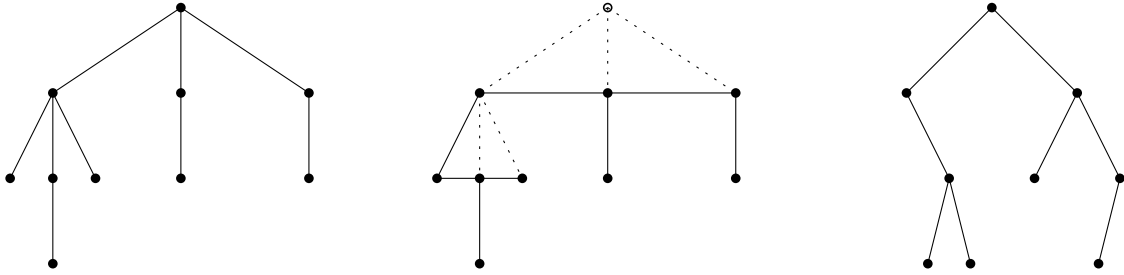


Abbildung 13: Ein Beispiel für die Durchführung der „natural rotation“.

Bemerkung 5.2. Folgende Aussagen erweisen sich als nützlich:

1. Es ist $R(u_1 \dots u_k) = 1^{u_1-1} 0 1^{u_2-1} 0 \dots 1^{u_k-1}$ wobei $1^0 := \Lambda$ sei.
2. Die natürliche Korrespondenz ist nach Konstruktion verträglich mit der Enthaltensein-Relation „ \subset “: Ist $\mathbf{s} \subset \mathbf{t}$ mit $\mathbf{s}, \mathbf{t} \in \mathbb{H}$, so ist $R(\mathbf{s}) \subset R(\mathbf{t})$. Das folgt aus der Tatsache $R(\mathbf{s} \cup \{u\}) = R(\mathbf{s}) \cup \{R(u)\}$ mit einem zu \mathbf{s} externen Knoten u , denn \mathbf{s} kann in $\#\mathbf{t} - \#\mathbf{s}$ Schritten durch Hinzunahme externer Knoten zu \mathbf{t} ergänzt werden. Umgekehrt hat man $R^{-1}(\mathbf{y}) \subset R^{-1}(\mathbf{z})$ für binäre Bäume \mathbf{y} und \mathbf{z} mit $\mathbf{y} \subset \mathbf{z}$. Dabei ist für $\mathbf{z} \in \mathbb{B}$

$$R^{-1}(\mathbf{z}) := \{R^{-1}(u) : u \in \mathbf{z}\} \cup \{\Lambda\}.$$

3. Es besteht folgende Beziehung zwischen den Teilbaumgrößen: Für $\mathbf{t} \in \mathbb{H}$, $\mathbf{t} \neq \{\Lambda\}$, $\mathbf{z} = R(\mathbf{t})$ und $u \in \mathbf{t}$ gilt $\#\mathbf{t}(u) = 1 + \#\mathbf{z}(R(u \circ 1))$, denn die Nachfolger von u in \mathbf{t} werden auf den Teilbaum vom Knoten $R(u \circ 1)$ in $R(\mathbf{t})$ abgebildet.

Bemerkung 5.3. Sei nun $(Z_n)_{n \in \mathbb{N}}$ ein BST-Prozess. Dann ist $(X_n)_{n \in \mathbb{N}}$ mit $X_1 \equiv \{\Lambda\}$ und $X_n = R^{-1}(Z_{n-1})$ für $n \geq 2$ eine Harris-Kette. Dazu betrachten wir die Übergangswahrscheinlichkeiten: für $\mathbf{s} \in \mathbb{H}_n$ und $\mathbf{t} \in \mathbb{H}_{n+1}$, $n > 1$ gilt

$$\begin{aligned} \mathbb{P}(X_{n+1} = \mathbf{t} | X_n = \mathbf{s}) &= \mathbb{P}(R^{-1}(Z_n) = \mathbf{t} | R^{-1}(Z_{n-1}) = \mathbf{s}) \\ &= \mathbb{P}(Z_n = R(\mathbf{t}) | Z_{n-1} = R(\mathbf{s})) = \frac{1}{n} \end{aligned}$$

wegen der oben erwähnten Verträglichkeit von R mit „ \subset “. Analog lässt sich zeigen: Ist $(X_n)_{n \in \mathbb{N}}$ eine Harris-Kette, so ist $(R(X_{n+1}))_{n \in \mathbb{N}}$ ein BST-Prozess.

5.1 Natürliche Korrespondenz und die Doob-Martin-Kompaktifizierungen

Die Beobachtung aus der Bemerkung 5.3 führt zu der Frage, ob es einen Zusammenhang zwischen den Doob-Martin-Kompaktifizierungen der BST- und der Harris-Kette gibt. Der Martin-Rand $\partial\mathbb{B}$ des BST-Prozesses ist laut [8] die Menge der Wahrscheinlichkeitsmaße auf $\{0, 1\}^\infty$, versehen mit der von den Mengen

$$B_u := \{v \in \{0, 1\}^\infty : u < v\},$$

$u \in \{0, 1\}^*$, erzeugten σ -Algebra. Die Konvergenz einer Folge $(z_n)_{n \in \mathbb{N}}$ binärer Bäume mit der Eigenschaft $\#z_n \rightarrow \infty$ gegen ein Randelement $\nu \in \partial\mathbb{B}$ bedeutet

$$\lim_{n \rightarrow \infty} \frac{\#z_n(u)}{\#z_n} = \nu(B_u)$$

für alle $u \in \{0, 1\}^*$. In [13] wird gezeigt, dass die Konvergenz in der Doob-Martin-Kompaktifizierung der BST-Kette sich als schwache Konvergenz von Maßen auf einem geeigneten metrischen Raum interpretieren lässt. Zum besseren Verständnis der Resultate aus [13] führen wir bei der folgenden Darstellung einige Zwischenschritte aus. Auf $\{0, 1\}^* \cup \{0, 1\}^\infty$ wird durch

$$d(u, v) := \begin{cases} 2^{-|u \wedge v|} & \text{für } u \neq v, \\ 0 & \text{für } u = v \end{cases}$$

eine Metrik definiert¹⁶. Diese Metrik ist sogar eine Ultrametrik, das heißt

$$d(u, v) \leq \max\{d(u, w), d(w, v)\}$$

für alle $u, v, w \in \{0, 1\}^* \cup \{0, 1\}^\infty$. Auf diese Weise erhält man die Endenkompaktifizierung von $\{0, 1\}^*$, betrachtet als unendlicher Baum aus der Sicht der Graphentheorie, siehe dazu [29, S.232-237]. Der Raum $\{0, 1\}^* \cup \{0, 1\}^\infty$ ist ein kompakter separabler metrischer Raum. Die Mengen B_u werden zu den Mengen

$$\bar{B}_u := \{v \in \{0, 1\}^* \cup \{0, 1\}^\infty : u \leq v\},$$

$u \in \{0, 1\}^*$, erweitert. Wegen

$$\begin{aligned} \bar{B}_u &= \{v \in \{0, 1\}^* \cup \{0, 1\}^\infty : d(u, v) \leq 2^{-|u|}\} \\ &= \{v \in \{0, 1\}^* \cup \{0, 1\}^\infty : d(u, v) < 2^{-|u|+1}\} \end{aligned}$$

sind diese Mengen offen und abgeschlossen. Das Mengensystem $\{\emptyset\} \cup \{\bar{B}_u : u \in \{0, 1\}^*\}$ ist ein durchschnittsstabiles Erzeugendensystem der Borel'schen σ -Algebra $\mathcal{B}(\{0, 1\}^* \cup \{0, 1\}^\infty)$. Außerdem ist es eine konvergenzbestimmende Klasse: Es seien ν_n, ν Wahrscheinlichkeitsmaße auf $(\{0, 1\}^* \cup \{0, 1\}^\infty, \mathcal{B}(\{0, 1\}^* \cup \{0, 1\}^\infty))$, dann folgt aus $\nu_n(B) \rightarrow \nu(B)$ für alle $B \in \{\bar{B}_u : u \in \{0, 1\}^*\}$ die schwache Konvergenz $\nu_n \rightarrow \nu$. Diese Behauptung kann wie folgt bewiesen werden. Angenommen, es gilt $\nu_n(\bar{B}_u) \rightarrow \nu(\bar{B}_u)$ für alle $u \in \{0, 1\}^*$. Dann gilt auch $\nu_n(\{u\}) \rightarrow \nu(\{u\})$ für alle $u \in \{0, 1\}^*$, denn man hat $\{u\} = \bar{B}_u \setminus (\bar{B}_{u_0} \cup \bar{B}_{u_1})$. Wir betrachten das System

$$\mathcal{U} := \{\emptyset\} \cup \{\{u\} : u \in \{0, 1\}^*\} \cup \{\bar{B}_u : u \in \{0, 1\}^*\}.$$

¹⁶Wir benutzen hier eine von [13] abweichende Definition der Metrik.

Wir zeigen nun, dass \mathcal{U} die Bedingungen des Korollars 1 zum Satz 2.2 aus [2], siehe Anhang A.3, erfüllt. Offensichtlich ist \mathcal{U} abgeschlossen bezüglich endlicher Durchschnittsbildung. Es bleibt zu zeigen: Für jedes $v \in \{0, 1\}^* \cup \{0, 1\}^\infty$ und jedes positive ε existiert ein $B \in \mathcal{U}$ mit $v \in B^\circ \subset B \subset S(v, \varepsilon)$. Dabei bezeichnet $S(v, \varepsilon)$ die offene ε -Kugel um v .

- Für $u \in \{0, 1\}^*$ ist die Aussage klar. Man beachte dabei, dass $S(u, \varepsilon) = \{u\}$ für $\varepsilon \leq 2^{-|u|}$ gilt.
- Sei nun $u \in \{0, 1\}^\infty$. Wähle ein $k \in \mathbb{N}$ mit $2^{-k} < \varepsilon$ und setze $w = u_1 \dots u_k$. Ist $v \in \bar{B}_w$, so ist $d(w, v) \leq 2^{-k}$. Außerdem ist $d(u, w) = 2^{-k}$. Nach der Ultrametrik-Ungleichung gilt $d(u, v) \leq \max\{d(u, w), d(w, v)\} < \varepsilon$. Also ist $\bar{B}_w \subset S(u, \varepsilon)$.

Damit ist der Beweis abgeschlossen. Wir stellen fest, dass \mathcal{U} auch die Bedingungen des Satzes 2.2 aus [2] erfüllt. Das bedeutet, dass jede offene Menge eine endliche oder abzählbare Vereinigung der Mengen aus \mathcal{U} ist. Das Mengensystem \mathcal{U} ist eine abzählbare Basis der Topologie des metrischen Raumes $(\{0, 1\}^* \cup \{0, 1\}^\infty, d)$.

Schließlich lassen sich Bäume $\mathbf{z} \in \mathbb{B}$ als Gleichverteilungen auf der Knotenmenge \mathbf{z} interpretieren, wobei $(\{0, 1\}^* \cup \{0, 1\}^\infty, \mathcal{B}(\{0, 1\}^* \cup \{0, 1\}^\infty))$ der zugrundeliegende Messraum ist. Formal setzt man dabei $\text{unif}(\mathbf{z})(\bar{B}_u) := \frac{\#\mathbf{z}(u)}{\#\mathbf{z}}$ für alle $u \in \{0, 1\}^*$. Einem Maß ν aus dem Martin-Rand kann durch die Festlegung

$$\tilde{\nu}(\bar{B}_u) := \nu(\{0, 1\}^\infty \cap \bar{B}_u), \quad u \in \{0, 1\}^*,$$

ein Maß $\tilde{\nu}$ auf $(\{0, 1\}^* \cup \{0, 1\}^\infty, \mathcal{B}(\{0, 1\}^* \cup \{0, 1\}^\infty))$ eindeutig zugeordnet werden, da $\{0, 1\}^\infty \cap \bar{B}_u = B_u$ für alle $u \in \{0, 1\}^*$ gilt. Ein solches Maß erfüllt die Bedingung $\tilde{\nu}(\{0, 1\}^*) = 0$. Insgesamt folgt aus der Konvergenz in der Doob-Martin-Kompaktifizierung die schwache Konvergenz der zugehörigen Wahrscheinlichkeitsmaße. Es ist aber zu beachten, dass es Maße gibt, die kein Urbild in der Doob-Martin-Kompaktifizierung haben, wie beispielsweise Gleichverteilungen auf endlichen nicht präfixstabilen Mengen. Es wurde also eine stetige und injektive Abbildung

$$\iota : \mathbb{B} \cup \partial\mathbb{B} \rightarrow \mathbb{M}(\{0, 1\}^* \cup \{0, 1\}^\infty)$$

konstruiert, wobei $\mathbb{M}(\{0, 1\}^* \cup \{0, 1\}^\infty)$ die Menge der Wahrscheinlichkeitsmaße auf $(\{0, 1\}^* \cup \{0, 1\}^\infty, \mathcal{B}(\{0, 1\}^* \cup \{0, 1\}^\infty))$ bezeichne. Da $\mathbb{B} \cup \partial\mathbb{B}$ kompakt und $\mathbb{M}(\{0, 1\}^* \cup \{0, 1\}^\infty)$ versehen mit der schwachen Topologie Hausdorff'sch ist, folgt, dass die Umkehrabbildung $\iota^{-1} : \mathbb{M}(\{0, 1\}^* \cup \{0, 1\}^\infty) \rightarrow \mathbb{B} \cup \partial\mathbb{B}$ stetig ist, siehe Satz A.2 im Anhang. Somit ist die Doob-Martin-Kompaktifizierung homöomorph zu $\mathbb{M}_r(\{0, 1\}^* \cup \{0, 1\}^\infty) := \iota(\mathbb{B} \cup \partial\mathbb{B})$ mit der Unterraumtopologie.

Wir erinnern an den Abschnitt 4.3: Der Martin-Rand $\partial\mathbb{H}$ der Harris-Kette ist die Menge der Wahrscheinlichkeitsmaße auf $(\mathbb{N}^\infty \sqcup \bigsqcup_{k=0}^\infty (\mathbb{N}^k \times \{\diamond\}^\infty), \mathcal{A})$, mit der von den Mengen

$$A_u := \left\{ v \in \mathbb{N}^\infty \sqcup \bigsqcup_{k=0}^\infty (\mathbb{N}^k \times \{\diamond\}^\infty) : u < v \right\},$$

$u \in \mathbb{N}^*$, erzeugten σ -Algebra \mathcal{A} . Ferner konvergiert eine Folge $(\mathbf{t}_n)_{n \in \mathbb{N}}$ aus \mathbb{H} genau dann gegen ein $\mu \in \partial\mathbb{H}$, wenn $\#\mathbf{t}_n \rightarrow \infty$ und $\lim_{n \rightarrow \infty} \frac{\#\mathbf{t}_n(u)}{\#\mathbf{t}_n} = \mu(A_u)$ für alle $u \in \mathbb{N}^*$ gilt. Die Doob-Martin-Kompaktifizierung ist bekanntlich bis auf Homöomorphie festgelegt.

Die obige Vorgehensweise motiviert die Suche nach einem geeigneten Repräsentanten. Es ist dabei zu beachten, dass \mathbb{N}^* , betrachtet als unendlicher Baum, nicht lokal endlich ist: Jeder Knoten besitzt unendlich viele Nachbarn. Als erstes setzen wir die Abbildung $R : \mathbb{N}^* \setminus \{\Lambda\} \rightarrow \{0, 1\}^*$ auf $\widehat{\mathbb{N}^*} \setminus \{\Lambda\}$ fort, wobei wir die abkürzende Schreibweise $\widehat{\mathbb{N}^*} := \mathbb{N}^* \sqcup \mathbb{N}^\infty \sqcup \bigsqcup_{k=0}^\infty (\mathbb{N}^k \times \{\diamond\}^\infty)$ verwenden. Für Punkte von der Form $u \diamond^\infty$, $u \in \mathbb{N}^*$, setzen wir $R(u \diamond^\infty) = R(u) \circ 111 \dots$, für $u \in \mathbb{N}^\infty$ definieren wir $R(u_1 u_2 \dots) = 1^{u_1-1} 0 1^{u_2-1} 0 \dots$, vergleiche Bemerkung 5.2. Dann definieren wir durch

$$d'(u, v) := d(R(u), R(v))$$

für $u, v \in \widehat{\mathbb{N}^*} \setminus \{\Lambda\}$ mit d wie oben eine Metrik. Sie hat folgende Eigenschaften. Ist $u \leq v$, $u \in \mathbb{N}^* \setminus \{\Lambda\}$, $v \in (\mathbb{N}^* \setminus \{\Lambda\}) \sqcup \mathbb{N}^\infty$, dann ist $d'(u, v) = 2^{-(|u|_1-1)}$. Sind die Knoten von der Form $u \circ i \circ v$, $u \circ j \circ w$ mit $u \in \mathbb{N}^*$, $i, j \in \mathbb{N}$, $i \neq j$ und $v, w \in \widehat{\mathbb{N}^*}$, dann ist

$$d'(u \circ i \circ v, u \circ j \circ w) = 2^{-(|u|_1-1)-\min\{i,j\}}.$$

Dadurch wird die Abbildung R zu einer Isometrie. Die betrachteten Räume $(\widehat{\mathbb{N}^*} \setminus \{\Lambda\}, d')$ und $(\{0, 1\}^* \cup \{0, 1\}^\infty, d)$ haben also die gleichen topologischen Eigenschaften, insbesondere sind beide kompakt und separabel. Wir erweitern die Mengen A_u zu

$$\bar{A}_u := \left\{ v \in \widehat{\mathbb{N}^*} : u \leq v \right\},$$

$u \in \mathbb{N}^*$, und betrachten das Mengensystem

$$\mathcal{D} := \{D_u : u \in \mathbb{N}^* \setminus \{\Lambda\}\} \quad \text{mit} \quad D_u := \bar{A}_u \setminus \left(\{\bar{u}\} \cup \bigcup_{k=1}^{\varepsilon(u)-1} \bar{A}_{u \circ k} \right).$$

Die offenen ε -Kugeln sind entweder Einpunktmengen oder Mengen aus \mathcal{D} , sie bilden eine abzählbare Basis der Topologie des metrischen Raumes $(\widehat{\mathbb{N}^*} \setminus \{\Lambda\}, d')$. Die Hinzunahme des Punktes Λ , sodass $\{\Lambda\}$ offen ist, ändert die interessierenden topologischen Eigenschaften nicht. Die Borel'sche σ -Algebra $\mathcal{B}(\widehat{\mathbb{N}^*})$ wird also von $\mathcal{U}_1 := \{\emptyset\} \cup \{\{u\} : u \in \mathbb{N}^*\} \cup \mathcal{D}$ erzeugt. Ein anderes durchschnittsstabiles Erzeugendensystem ist

$$\mathcal{U}_2 := \{\emptyset\} \cup \{\{u\} : u \in \mathbb{N}^*\} \cup \{\bar{A}_u : u \in \mathbb{N}^*\}.$$

Es bleibt zu zeigen, dass \mathcal{U}_2 eine konvergenzbestimmende Klasse ist.

Lemma 5.4. *Es seien μ_n, μ Wahrscheinlichkeitsmaße auf $(\widehat{\mathbb{N}^*}, \mathcal{B}(\widehat{\mathbb{N}^*}))$. Gilt $\mu_n(A) \rightarrow \mu(A)$ für alle $A \in \mathcal{U}_2$, so konvergiert μ_n gegen μ schwach.*

Beweis. Gilt $\mu_n(A) \rightarrow \mu(A)$ für alle $A \in \mathcal{U}_2$, so gilt wegen der Definition von \mathcal{D} auch $\mu_n(A) \rightarrow \mu(A)$ für alle $A \in \mathcal{U}_1$. Das Mengensystem \mathcal{U}_1 erfüllt die Bedingungen des Satzes 2.2 aus [2], siehe Satz A.4. \square

Wie in der Situation binärer Bäume werden Ulam-Harris-Bäume als Gleichverteilungen auf ihrer Knotenmenge interpretiert. Einem Randelement μ wird ein Maß $\tilde{\mu}$ auf $(\widehat{\mathbb{N}^*}, \mathcal{B}(\widehat{\mathbb{N}^*}))$ wie folgt zugeordnet: man setzt $\tilde{\mu}(\{u\}) := 0$ für $u \in \mathbb{N}^*$ und

$$\tilde{\mu}(A) := \mu \left(A \cap \left(\mathbb{N}^\infty \sqcup \bigsqcup_{k=0}^\infty (\mathbb{N}^k \times \{\diamond\}^\infty) \right) \right)$$

für $A \in \{\bar{A}_u : u \in \mathbb{N}^*\}$. Aus der Konvergenz in Doob-Martin-Kompaktifizierung der Harris-Kette folgt also auch die schwache Konvergenz der zugehörigen Wahrscheinlichkeitsmaße. Erneut ist es zu beachten, dass es Maße gibt, die keinen Urbild in der Doob-Martin-Kompaktifizierung haben. Wie oben erhält man, dass die Doob-Martin-Kompaktifizierung homöomorph ist zu der Menge $\mathbb{M}_r(\widehat{\mathbb{N}^*})$ der zulässigen Maße auf $(\widehat{\mathbb{N}^*}, \mathcal{B}(\widehat{\mathbb{N}^*}))$ mit der Unterraumtopologie.

Da nun geeignete Repräsentanten der Doob-Martin-Kompaktifizierungen gefunden sind, können wir die natürliche Korrespondenz auf den Martin-Rand fortsetzen. Wir definieren eine Abbildung $\Psi : \mathbb{M}_r(\{0, 1\}^* \cup \{0, 1\}^\infty) \rightarrow \mathbb{M}_r(\widehat{\mathbb{N}^*})$, $\nu \mapsto \Psi(\nu)$ folgendermaßen: Ist $\nu = \text{unif}(\mathbf{z})$ für ein $\mathbf{z} \in \mathbb{B}$, so setzen wir $\Psi(\nu) := \text{unif}(R^{-1}(\mathbf{z}))$. Für Elemente ν aus dem Martin-Rand der BST-Kette sei

$$\Psi(\nu)(\bar{A}_u) := \nu(\bar{B}_{R(u \circ 1)}), \quad u \in \mathbb{N}^*$$

Satz 5.5. *Die Abbildung Ψ ist ein Homöomorphismus zwischen $\mathbb{M}_r(\{0, 1\}^* \cup \{0, 1\}^\infty)$ und $\mathbb{M}_r(\widehat{\mathbb{N}^*}) \setminus \{\text{unif}(\{\Lambda\})\}$.*

Beweis. Angenommen es gilt $\nu_n \rightarrow \nu$ schwach, wobei $\nu_n = \text{unif}(\mathbf{z}_n)$ und $(\mathbf{z}_n)_{n \in \mathbb{N}}$ eine Folge binärer Bäume ist. Ist $\nu = \text{unif}(\mathbf{z})$ für ein $\mathbf{z} \in \mathbb{B}$, so gilt $\mathbf{z}_n = \mathbf{z}$ für alle bis auf endlich viele n . Daraus folgt $\Psi(\nu_n) \rightarrow \Psi(\nu)$ schwach.

Sei nun ν ein Element aus dem Martin-Rand der BST-Kette, das heißt $\nu(\{u\}) = 0$ und $\nu_n(\bar{B}_u) \rightarrow \nu(\bar{B}_u)$ für alle $u \in \{0, 1\}^*$. Außerdem gilt in diesem Fall $\#\mathbf{z}_n \rightarrow \infty$ mit $n \rightarrow \infty$. Mit Aussage 3 aus der Bemerkung 5.2 folgt:

$$\begin{aligned} \Psi(\nu)(\bar{A}_u) &= \nu(\bar{B}_{R(u \circ 1)}) = \lim_{n \rightarrow \infty} \nu_n(\bar{B}_{R(u \circ 1)}) \\ &= \lim_{n \rightarrow \infty} \frac{\#\mathbf{z}_n(R(u \circ 1))}{\#\mathbf{z}_n} = \lim_{n \rightarrow \infty} \frac{1 + \#\mathbf{z}_n(R(u \circ 1))}{1 + \#\mathbf{z}_n} = \lim_{n \rightarrow \infty} \frac{\#R^{-1}(\mathbf{z}_n)(u)}{\#R^{-1}(\mathbf{z}_n)} \\ &= \lim_{n \rightarrow \infty} \text{unif}(R^{-1}(\mathbf{z}_n))(\bar{A}_u) \end{aligned}$$

für alle $u \in \mathbb{N}^*$. Das bedeutet aber die Konvergenz von $\Psi(\nu_n) = \text{unif}(R^{-1}(\mathbf{z}_n))$ gegen $\Psi(\nu)$ in der Doob-Martin-Kompaktifizierung der Harris-Kette. \square

Sei wieder $(Z_n)_{n \in \mathbb{N}}$ ein BST-Prozess und Z_∞ der fast sichere Grenzwert. Dann gilt $\Psi(Z_n) \rightarrow \Psi(Z_\infty)$ mit $n \rightarrow \infty$ fast sicher in der Doob-Martin-Kompaktifizierung der Harris-Kette. Auf diese Weise kann die Austrittsverteilung der Harris-Kette also auch mit Hilfe der Austrittsverteilung der BST-Kette bestimmt werden.

Bemerkung 5.6. Wir sind von einer aus [8] bekannten Beschreibung der Doob-Martin-Kompaktifizierung $\mathbb{H} \cup \partial\mathbb{H}$ der Harris-Kette ausgegangen. Wäre diese nicht bekannt, so könnte man einen Repräsentanten von $\partial\mathbb{H}$ mit dem folgenden Ansatz bestimmen: Die natürliche Korrespondenz R ist als Funktion von $(\mathbb{N}^* \setminus \{\Lambda\}, d')$ nach $(\{0, 1\}^*, d)$ eine Isometrie. Nach Satz A.3 lässt sie sich zu einer Isometrie zwischen den jeweiligen Vervollständigungen fortsetzen. Man betrachte nun für zulässige Maße $\mu \in \mathbb{M}_r(\{0, 1\}^* \cup \{0, 1\}^\infty)$ die Bildmaße bezüglich der Umkehrabbildung der Fortsetzung von R und nutze das Continuous Mapping Theorem, siehe beispielsweise [18, Satz 13.25]. Die Sonderrolle von Λ stört dabei nicht.

5.2 Natürliche Korrespondenz und der algorithmische Hintergrund

Sowohl zufällige binäre Bäume als auch zufällige Harris-Bäume können algorithmisch aus einer Folge $(\eta_n)_{n \in \mathbb{N}}$ unabhängiger, auf dem Einheitsintervall gleichverteilter Zufallsvariablen erzeugt werden. Was passiert, wenn die entsprechenden Algorithmen auf dieselbe Inputfolge angewendet werden? Eine graphische Darstellung des Outputs zu 10 Beobachtungsdaten findet man in den Abbildungen 6 und 12. Um diese Frage zu beantworten, führen wir zu einem binären Baum \mathbf{z} wie folgt die „gespiegelte Version“ $S(\mathbf{z})$ ein. Zunächst setzen wir $S(\Lambda) = \Lambda$ und $S(u) = (1 - u_1) \circ (1 - u_2) \circ \dots \circ (1 - u_n)$ für ein $u = u_1 u_2 \dots u_n \in \{0, 1\}^*$, $u \neq \Lambda$ (bitweise Invertierung). Dann definieren wir $S(\mathbf{z}) := \{S(u) : u \in \mathbf{z}\}$, siehe Abbildung 14.

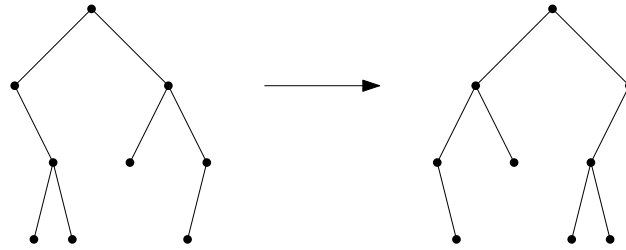


Abbildung 14: Ein binärer Baum und seine „gespiegelte Version“.

Satz 5.7. Sei $(x_i)_{i \in \mathbb{N}}$ eine Folge paarweise verschiedener Punkte $x_i \in [0, 1]$ und $(\mathbf{t}_i, \varphi_i)_{i \in \mathbb{N}}$ die zugehörige Folge bewerteter Ulam-Harris-Bäume. Sei $(\mathbf{z}_i, \gamma_i)_{i \in \mathbb{N}}$ die aus $(x_i)_{i \in \mathbb{N}}$ durch den BST-Algorithmus konstruierte Folge bewerteter binärer Bäume. Dann gilt für $n \geq 2$: $S(R(\mathbf{t}_n)) = \mathbf{z}_{n-1}$.

Beweis. Wir zeigen die Aussage per Induktion. Es ist $\mathbf{t}_2 = \{\Lambda, 1\}$, $\varphi_2(\Lambda) = 0$, $\varphi_2(1) = x_1$ und $\mathbf{z}_1 = \{\Lambda\}$, $\gamma_1(\Lambda) = x_1$. Ist $x_2 < x_1$ so gilt $\mathbf{t}_3 = \{\Lambda, 1, 2\}$, $\varphi_3(2) = x_2$ und $\mathbf{z}_1 = \{\Lambda, 0\}$, $\gamma_2(0) = x_2$. Anderenfalls hat man $\mathbf{t}_3 = \{\Lambda, 1, 11\}$, $\varphi_3(11) = x_2$ und $\mathbf{z}_1 = \{\Lambda, 1\}$, $\gamma_2(1) = x_2$. Die Behauptung gilt also für $n = 2$ und $n = 3$.

Beim Induktionsschritt n zu $n+1$ hat man $\mathbf{t}_{n+1} = \mathbf{t}_n \cup \{u\}$ mit einem zu \mathbf{t}_n externen Knoten $u \in \mathbb{N}^*$ und $\mathbf{z}_n = \mathbf{z}_{n-1} \cup \{v\}$ mit einem zu \mathbf{z}_{n-1} externen Knoten $v \in \{0, 1\}^*$. Die Knoten u und v erhalten die Bewertung x_n . Nach Induktionsvoraussetzung gilt $S(R(\mathbf{t}_n)) = \mathbf{z}_{n-1}$. Es bleibt zu zeigen: $S(R(u)) = v$. Es existiert ein k mit $1 \leq k \leq n-1$, sodass x_k die Bewertung von \bar{u} ist. Es seien $0 =: x_{(0)} < x_{(1)} < \dots < x_{(k-1)} < x_{(k)} := 1$ erweiterte Ordnungsstatistiken zu x_1, \dots, x_{k-1} und sei j so, dass $x_{(j)} < x_k < x_{(j+1)}$ gilt. Nach Konstruktion gehören zum Teilbaum von \bar{u} in \mathbf{t}_{n+1} Knoten mit Bewertungen im Intervall $(x_k, x_{(j+1)})$, siehe Beweis zum Satz 4.8. Nach Induktionsvoraussetzung existiert ein mit x_k bewerteter Knoten w in \mathbf{z}_{n-1} mit $S(R(\bar{u})) = w$. Zum Teilbaum von w in \mathbf{z}_n gehören Knoten mit Bewertungen im Intervall $(x_{(j)}, x_{(j+1)})$. Insbesondere enthält der rechte Teilbaum von w Knoten mit Bewertungen in $(x_k, x_{(j+1)})$. Ist $\varepsilon(u) = 1$, so ist x_n der erste absteigende Rekord der Folge $(x_i)_{i \in \mathbb{N}}$ im Intervall $(x_k, x_{(j+1)})$. Also ist $v = w \circ 1 = S(R(u))$. Ist $\varepsilon(u) = l$ mit $l > 1$, so ist x_n der l -te absteigende Rekord. In diesem Fall ist $v = w \circ 1 \circ 0^{l-1} = S(R(u))$. \square

Sei nun $(Z_n)_{n \in \mathbb{N}}$ eine mit dem BST-Algorithmus aus $(\eta_n)_{n \in \mathbb{N}}$ erzeugte Folge binärer Bäume und sei Z_∞ die aus Satz 1 in [13] bekannte Darstellung des fast sicheren Grenzwertes.

Sei $(X_n)_{n \in \mathbb{N}}$ die aus $(\eta_n)_{n \in \mathbb{N}}$ erzeugte Folge der Harris-Bäume und X_∞ wie im Satz 4.8. Unter diesen Voraussetzungen gilt: $S(R(X_n)) = Z_{n-1}$ für $n \geq 2$. Folglich ist $X_\infty(A_u) = Z_\infty(B_{S(R(u \circ 1))})$, $u \in \mathbb{N}^*$.

5.3 Natürliche Korrespondenz und Baumfunktionale

Die natürliche Korrespondenz ist auch bei der Betrachtung der internen und der horizontalen Pfadlänge hilfreich. Wir definieren eine neue Abbildung $T : \mathbb{H} \rightarrow \mathbb{H}$ durch $T(\{\Lambda\}) := \{\Lambda\}$ und $T(\mathbf{t}) := R^{-1}(S(R(\mathbf{t})))$ für $\mathbf{t} \neq \{\Lambda\}$, siehe Beispiel in Abbildung 15.

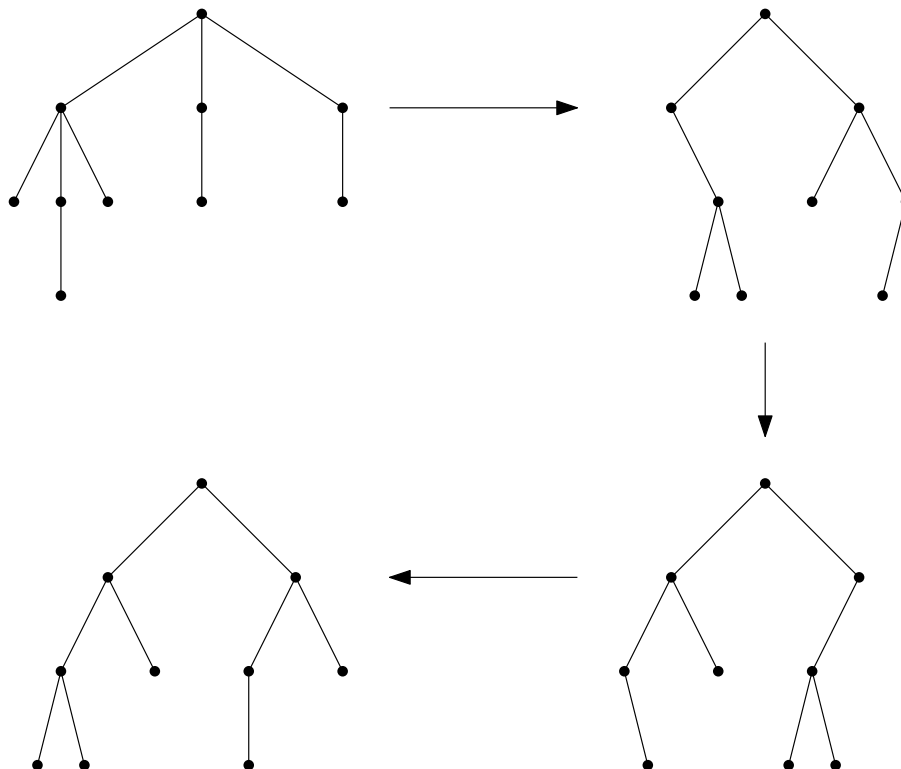


Abbildung 15: Die Abbildung T angewendet auf einen Ulam-Harris-Baum \mathbf{t} mit 10 Knoten. Es gilt $\text{IPL}(\mathbf{t}) = 16$, $\text{IPL}(T(\mathbf{t})) = 19$ und $\text{HPL}(T(\mathbf{t})) = 7 = \text{IPL}(\mathbf{t}) - (\#\mathbf{t} - 1)$.

Bemerkung 5.8. Folgende Aussagen zum Verhalten der interessierenden Baumfunktionale unter den oben eingeführten Abbildungen erweisen sich als nützlich:

1. Die Spiegelung S ändert die interne Pfadlänge bei binären Bäumen \mathbf{z} nicht:

$$\text{IPL}(\mathbf{z}) = \text{IPL}(S(\mathbf{z})).$$

2. Für ein $u \in \{0, 1\}^*$ gilt $|u| = |u|_1 + |S(u)|_1$, mit anderen Worten setzt sich die Tiefe eines Knotens im binären Baum aus der Anzahl der Schritte nach links und nach rechts zusammen. Da die interne Pfadlänge die Summe der Tiefen interner Knoten ist, folgt daraus für binäre Bäume \mathbf{z}

$$\text{IPL}(\mathbf{z}) = \sum_{u \in \mathbf{z}} |u|_1 + \sum_{u \in S(\mathbf{z})} |u|_1.$$

3. Für $\mathbf{t} \in \mathbb{H}$ gilt

$$\text{IPL}(\mathbf{t}) = \text{IPL}(R(\mathbf{t})) + (\#\mathbf{t} - 1) - \sum_{u \in R(\mathbf{t})} |u|_1,$$

denn ein $u \in \mathbb{N}^* \setminus \{\Lambda\}$ wird nach Konstruktion auf ein $v \in \{0, 1\}^*$ mit $v = R(u)$ abgebildet, sodass $|u| = 1 + |v| - |v|_1$ gilt.

4. Außerdem gilt für $\mathbf{t} \in \mathbb{H}$

$$\sum_{u \in \mathbf{t}} |u|_1 = \text{IPL}(R(\mathbf{t})) + (\#\mathbf{t} - 1),$$

denn für $u \in \mathbb{N}^*$ hat man $|u|_1 = |R(u)|_1 + 1$, wieder nach Konstruktion.

Satz 5.9. Sei $(X_n)_{n \in \mathbb{N}}$ eine Harris-Kette. Dann gilt

$$\mathcal{L}(\text{HPL}(X_n)) = \mathcal{L}(\text{IPL}(X_n) - (n - 1)).$$

Beweis. Zunächst stellen wir fest, dass mit $(X_n)_{n \in \mathbb{N}}$ auch $(T(X_n))_{n \in \mathbb{N}}$ eine Harris-Kette ist: Wie bereits in der Bemerkung 5.3 erwähnt, ist in dieser Situation $(R(X_{n+1}))_{n \in \mathbb{N}}$ ein BST-Prozess. Offensichtlich ist die Verteilung eines BST-Prozesses invariant unter der Spiegelung S , also ist auch $(S(R(X_{n+1})))_{n \in \mathbb{N}}$ ein BST-Prozess. Schließlich folgt diese erste Hilfsaussage wieder mit der Bemerkung 5.3 aus $T(X_1) \equiv \{\Lambda\}$ und $T(X_n) = R^{-1}(S(R(X_n)))$ für $n \geq 2$. Es bleibt zu zeigen:

$$\text{HPL}(T(\mathbf{t})) = \text{IPL}(\mathbf{t}) - (\#\mathbf{t} - 1)$$

für $\mathbf{t} \in \mathbb{H}$. Nach der Definition der horizontalen Pfadlänge eines Baumes $\mathbf{t} \in \mathbb{H}$ folgt mit den Hilfsaussagen 3 und 4 aus der Bemerkung 5.8:

$$\begin{aligned} \text{HPL}(\mathbf{t}) &= \sum_{u \in \mathbf{t}} |u|_1 - \text{IPL}(\mathbf{t}) \\ &= \text{IPL}(R(\mathbf{t})) + (\#\mathbf{t} - 1) - \text{IPL}(R(\mathbf{t})) - (\#\mathbf{t} - 1) - \sum_{u \in R(\mathbf{t})} |u|_1 \\ &= \sum_{u \in R(\mathbf{t})} |u|_1. \end{aligned}$$

Mit den Hilfsaussagen 2 und 3 aus der Bemerkung 5.8 folgt daraus die zweite Behauptung:

$$\begin{aligned} \text{HPL}(T(\mathbf{t})) &= \text{HPL}(R^{-1}(S(R(\mathbf{t})))) = \sum_{u \in S(R(\mathbf{t}))} |u|_1 \\ &= \sum_{u \in S(R(\mathbf{t}))} |u|_1 + \sum_{u \in R(\mathbf{t})} |u|_1 + (\#\mathbf{t} - 1) - \sum_{u \in R(\mathbf{t})} |u|_1 - (\#\mathbf{t} - 1) \\ &= \text{IPL}(R(\mathbf{t})) + (\#\mathbf{t} - 1) - \sum_{u \in R(\mathbf{t})} |u|_1 - (\#\mathbf{t} - 1) \\ &= \text{IPL}(\mathbf{t}) - (\#\mathbf{t} - 1). \end{aligned} \quad \square$$

Anhang

A.1 Wörter und Alphabete

Der folgende Formalismus ist aus den Lehrbüchern zur Theoretischen Informatik übernommen, unter anderem [14] und [7]. Unter einem Alphabet Σ versteht man üblicherweise eine endliche, nicht-leere Menge. Wir lassen zusätzlich abzählbar unendliche Alphabete zu. Die Elemente eines Alphabets nennt man Symbole oder Buchstaben. Ein Wort ist eine endliche, möglicherweise leere, Folge von Buchstaben aus Σ . Für ein Wort $w = (w_1, w_2, \dots, w_n)$ verwenden wir die Kurzschreibweise $w = w_1 w_2 \dots w_n$. Σ^* bezeichnet die Menge aller Wörter über dem Alphabet Σ , einschließlich des leeren Wortes Λ . Die Länge $|w|$ eines Wortes w ist die Anzahl der Buchstaben in w . Das leere Wort hat die Länge 0. Für ein Wort $w = w_1 \dots w_n$ ist der Rumpf von w definiert als

$$\bar{w} := \begin{cases} w_1 \dots w_{n-1}, & \text{falls } n > 0, \\ \Lambda, & \text{falls } n = 0, \end{cases}$$

und der letzte Buchstabe von w ist gegeben durch

$$\varepsilon(w) := \begin{cases} w_n, & \text{falls } n > 0, \\ \Lambda, & \text{falls } w = \Lambda. \end{cases}$$

Die Verkettungsoperation für Wörter (Konkatenation) bezeichnen wir wie üblich mit „ \circ “:

$$v \circ w = v_1 \dots v_k \circ w_1 \dots w_l := v_1 \dots v_k w_1 \dots w_l.$$

Für $v \circ w$ schreibt man auch vw . Wir verwenden die Notation

$$w^n = \underbrace{w w \dots w}_{n\text{-mal}}, \quad w \in \Sigma^*, n \in \mathbb{N}$$

und $w^0 = \Lambda$. Ein Wort v nennt man ein Präfix von w und schreibt $v \leq w$ genau dann, wenn $v \circ u = w$ für ein $u \in \Sigma^*$ gilt. Ist zusätzlich $v \neq w$, so ist v ein echtes Präfix von w , in Zeichen $v < w$. Eine Wortmenge $L \subset \Sigma^*$ heißt präfixstabil, falls mit $w \in L$ auch jedes v mit $v \leq w$ ein Element von L ist.

A.2 Ergänzungen aus der Topologie

Folgende Definitionen, Aussagen und die zugehörigen Beweise finden sich in [15] und [25]. Sei (X, \mathcal{O}) ein topologischer Raum und $A \subset X$. Die Menge aller inneren Punkte von A wird mit A° bezeichnet, \bar{A} ist die Bezeichnung für den Abschluss von A . Man sagt A liegt dicht in X , falls $\bar{A} = X$ gilt. Im Folgenden heben wir einige besondere Topologien hervor:

- Für eine beliebige Menge X definiert die Potenzmenge von X eine Topologie auf X , genannt diskrete Topologie.
- Ist $X_0 \subset X$, so heißt $\mathcal{O}_{X_0} := \{U \cap X_0 : U \in \mathcal{O}\}$ Unterraumtopologie oder Spurtopologie, (X_0, \mathcal{O}_{X_0}) heißt Unterraum von X .
- Sei (X, d) ein metrischer Raum. Die Menge $S(x, \varepsilon) := \{y \in X : d(x, y) < \varepsilon\}$ heißt offene ε -Kugel um x . Die Topologie eines metrischen Raumes besteht aus Vereinigungsmengen von offenen Kugeln. Ein topologischer Raum X heißt metrisierbar, wenn es auf X eine Metrik gibt, die die Topologie von X induziert.

Seien X und Y topologische Räume.

- Eine bijektive Abbildung $f : X \rightarrow Y$ heißt Homöomorphismus, wenn f und f^{-1} stetig sind. Die Räume X und Y heißen dann homöomorph.
- X heißt ein Hausdorff-Raum, wenn je zwei verschiedene Punkte aus X disjunkte Umgebungen besitzen. Jeder metrisierbare Raum ist Hausdorff'sch.
- X heißt separabel, falls er eine abzählbare dichte Teilmenge enthält. Die Existenz einer abzählbaren Basis der Topologie impliziert die Separabilität. Dabei ist eine Teilmenge $\mathcal{U} \subset \mathcal{O}$ Basis der Topologie, wenn jede offene Menge von (X, \mathcal{O}) Vereinigung von Mengen aus \mathcal{U} ist.
- X heißt kompakt, falls jede offene Überdeckung eine endliche Teilüberdeckung besitzt:

$$\bigcup_{i \in I} U_i = X, U_i \text{ offen in } X \quad \Rightarrow \quad \exists \quad I' \subset I : |I'| < \infty, \bigcup_{i \in I'} U_i = X.$$

Eine Teilmenge $A \subset X$ heißt kompakt, wenn der Unterraum A kompakt ist.

Ein Hausdorff-Raum X heißt folgenkompakt, wenn jede Folge von Punkten aus X eine konvergente Teilfolge besitzt. In einem metrischen Raum stimmen die Begriffe kompakt und folgenkompakt überein. Stetige Bilder kompakter Räume sind kompakt:

Satz A.1 ([25, Satz 8.10]). *Ist X kompakt und $f : X \rightarrow Y$ stetig, dann ist $f(X)$ kompakt.*

Folgende Aussage folgt direkt aus [25, Satz 8.11]:

Satz A.2. *Sei Y Hausdorff'sch, $D \subset X$ kompakt, $f : D \rightarrow Y$ stetig und injektiv. Dann ist die Umkehrfunktion $f^{-1} : f(D) \rightarrow X$ stetig.*

Ein metrischer Raum heißt vollständig, wenn jede Cauchy-Folge konvergiert. Jeder metrische Raum (X, d) lässt sich dicht in einen vollständigen metrischen Raum $(\widehat{X}, \widehat{d})$ einbetten. Für weitere Details zur Vervollständigung metrischer Räume siehe [15, Kapitel 4]. Der nachfolgende Satz ist aus [15, S.68] entnommen:

Satz A.3. *Seien (X, d) und (Y, d') metrische Räume und $f : X \rightarrow Y$ eine gleichmäßig stetige Abbildung. Sind dann $(\widehat{X}, \widehat{d})$ und $(\widehat{Y}, \widehat{d}')$ Vervollständigungen von (X, d) und (Y, d') , so gibt es genau eine Fortsetzung von f zu einer stetigen Abbildung $f : \widehat{X} \rightarrow \widehat{Y}$.*

A.3 Schwache Konvergenz

Sei X ein metrischer Raum und \mathcal{B} die von den offenen Mengen erzeugte σ -Algebra. Wir bezeichnen mit $C_b(X)$ den Raum der stetigen und beschränkten Funktionen $f : X \rightarrow \mathbb{R}$. Seien μ_n, μ Wahrscheinlichkeitsmaße auf (X, \mathcal{B}) . Man sagt, dass μ_n schwach gegen μ konvergiert, falls

$$\int f d\mu_n \rightarrow \int f d\mu \text{ für jedes } f \in C_b(X)$$

gilt. Ist X separabel, so kann man zeigen, dass die Topologie der schwachen Konvergenz von einer Metrik induziert ist [18, S.252]. Folgende Aussagen und die zugehörigen Beweise findet man in [2].

Satz A.4 ([2, Satz 2.2]). Sei $\mathcal{U} \subset \mathcal{B}$, sodass

1. \mathcal{U} abgeschlossen bezüglich endlicher Durchschnittsbildung ist und
2. jede offene Menge aus X eine endliche oder abzählbare Vereinigung der Mengen aus \mathcal{U} ist.

Gilt $\mu_n(A) \rightarrow \mu(A)$ für jedes $A \in \mathcal{U}$, dann folgt $\mu_n \rightarrow \mu$ schwach.

Korollar A.5. Sei $\mathcal{U} \subset \mathcal{B}$, sodass

1. \mathcal{U} abgeschlossen bezüglich endlicher Durchschnittsbildung ist und
2. für jedes $x \in X$ und jedes $\varepsilon > 0$ ein $A \in \mathcal{U}$ mit $x \in A^\circ \subset A \subset S(x, \varepsilon)$ existiert.

Ist X separabel und gilt $\mu_n(A) \rightarrow \mu(A)$ für jedes $A \in \mathcal{U}$, dann folgt $\mu_n \rightarrow \mu$ schwach.

Literatur

- [1] Arnold, Barry C., N. Balakrishnan, and H. N. Nagaraja: *Records*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, 1998, ISBN 0-471-08108-6. A Wiley-Interscience Publication.
- [2] Billingsley, Patrick: *Convergence of probability measures*. John Wiley & Sons, Inc., New York-London-Sydney, 1968.
- [3] Dennert, Florian: *Zufällige binäre Bäume: Algorithmen, Asymptotik und Statistik*. PhD thesis, Leibniz Universität Hannover, 2009.
- [4] Devroye, Luc and Louise Laforest: *An analysis of random d -dimensional quad trees*. SIAM J. Comput., 19(5):821–832, 1990, ISSN 0097-5397.
- [5] Doob, J. L.: *Discrete potential theory and boundaries*. J. Math. Mech., 8:433–458; erratum 993, 1959.
- [6] Drmota, Michael: *Random trees*. SpringerWienNewYork, Vienna, 2009, ISBN 978-3-211-75355-2. An interplay between combinatorics and probability.
- [7] Erk, Katrin and Lutz Priese: *Theoretische Informatik: Eine umfassende Einführung*. Springer, 2008, ISBN 9783540763192.
- [8] Evans, S. N., R. Grübel, and A. Wakolbinger: *Trickle-down processes and their boundaries*. Electronic Journal of Probability, 17, 2012.
- [9] Finkel, R.A. and J.L. Bentley: *Quad trees: a data structure for retrieval on composite keys*. Acta Informatica, 4(1):1–9, 1974, ISSN 0001-5903.
- [10] Flajolet, Philippe, Gaston Gonnet, Claude Puech, and J. M. Robson: *Analytic variations on quadtrees*. Algorithmica, 10(6):473–500, 1993, ISSN 0178-4617.
- [11] Flajolet, Philippe and Robert Sedgewick: *Analytic combinatorics*. Cambridge University Press, Cambridge, 2009, ISBN 978-0-521-89806-5.
- [12] Grübel, Rudolf: *Kombinatorische Markov-Ketten*. Math. Semesterber., 60(2):185–215, 2013, ISSN 0720-728X.
- [13] Grübel, Rudolf: *Search trees: metric aspects and strong limit theorems*. Ann. Appl. Probab., 24(3):1269–1297, 2014, ISSN 1050-5164.
- [14] Hromkovič, J.: *Theoretische Informatik: Formale Sprachen, Berechenbarkeit, Komplexitätstheorie, Algorithmik, Kommunikation und Kryptographie*. Springer Vieweg, 2014, ISBN 9783658064327.
- [15] Jänich, Klaus: *Topologie*. Springer-Verlag, Berlin, 8th edition, 2005, ISBN 978-3-540-21393-2.
- [16] Janson, Svante, Tomasz Łuczak, and Andrzej Ruciński: *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000, ISBN 0-471-17541-2.

- [17] Kemeny, John G., J. Laurie Snell, and Anthony W. Knapp: *Denumerable Markov chains*. Springer-Verlag, New York-Heidelberg-Berlin, 2nd edition, 1976. With a chapter on Markov random fields, by David Griffeath, Graduate Texts in Mathematics, No. 40.
- [18] Klenke, A.: *Wahrscheinlichkeitstheorie*. Springer-Verlag Berlin Heidelberg, 2nd edition, 2008, ISBN 978-3540763178.
- [19] Knuth, Donald E.: *The art of computer programming. Vol. 1*. Addison-Wesley, Reading, MA, 3rd edition, 1997, ISBN 0-201-89683-4.
- [20] Mahmoud, H. M.: *Evolution of random search trees*. John Wiley & Sons Inc, 1992.
- [21] Mahmoud, Hosam M.: *Limiting distributions for path lengths in recursive trees*. Probab. Engrg. Inform. Sci., 5(1):53–59, 1991, ISSN 0269-9648.
- [22] Neininger, Ralph: *The Wiener index of random trees*. Combinatorics, Probability & Computing, 11(6):587–597, 2002.
- [23] Neininger, Ralph and Ludger Rüschemdorf: *On the internal path length of d -dimensional quad trees*. Random Structures Algorithms, 15(1):25–41, 1999, ISSN 1042-9832.
- [24] Pitman, J.: *Combinatorial stochastic processes*. Combinatorial Stochastic Processes: École D'Été de Probabilités de Saint-Flour XXXII - 2002. Springer, 2006, ISBN 9783540309901.
- [25] Querenburg, Boto von: *Mengentheoretische Topologie*. Springer-Verlag Berlin Heidelberg, 2001, ISBN 9783540677901.
- [26] Sawyer, S. A.: *Martin boundaries and random walks*. Contemporary Mathematics, 206, 1997.
- [27] Smythe, R. T. and H. M. Mahmoud: *A survey of recursive trees*. Theor. Probability and Math. Statist., 51, 1995.
- [28] Williams, David: *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991, ISBN 0-521-40455-X; 0-521-40605-6.
- [29] Woess, Wolfgang: *Denumerable Markov chains. Generating functions, boundary theory, random walks on trees*. EMS Textbooks in Mathematics. European Mathematical Society (EMS), Zürich, 2009, ISBN 978-3-03719-071-5.

Danksagung

Mein besonderer Dank gilt meinem Doktorvater Herrn Prof. Dr. Rudolf Grübel für die umfassende Betreuung bei der Entstehung dieser Arbeit.

Herrn Prof. Dr. Ludwig Baringhaus danke ich für die Übernahme des Korreferats.