

# Das asymptotische Verhalten der gewichteten Höhe des zweifach gewichteten Verzweigungsprozesses

Diplomarbeit

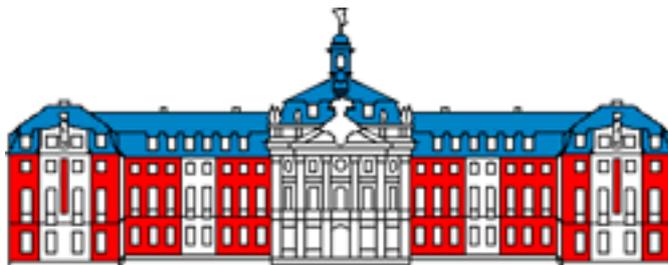
vorgelegt von

Sebastian Gebennus

Thema gestellt von

Prof. Dr. G. Alsmeyer

Münster, den 14. Mai 2006



Westfälische  
Wilhelms-Universität  
Münster

Fachbereich Mathematik und Informatik  
Institut für Mathematische Statistik



# Inhaltsverzeichnis

<b>Einleitung</b>	<b>v</b>
<b>1 Grundlagen der Theorie der Großen Abweichungen</b>	<b>1</b>
1.1 Die grundlegende Problematik . . . . .	1
1.2 Eigenschaften der kumulantenerzeugende Funktion . . . . .	3
1.3 Eigenschaften der Cramér-Funktion . . . . .	8
1.4 Der Satz von Cramér . . . . .	16
<b>2 Der zweifach gewichtete Verzweigungsprozess</b>	<b>21</b>
2.1 Modellbeschreibung . . . . .	21
<b>3 Asymptotik der gewichteten Höhe des zweifach gewichteten Verzweigungsprozesses</b>	<b>25</b>
3.1 Vorbereitungen . . . . .	25
3.2 Das asymptotische Verhalten von $H_n$ . . . . .	30
<b>4 Verallgemeinerung der Situation auf den Fall einer beliebigen Reproduktionsverteilung</b>	<b>37</b>
4.1 Verallgemeinerung des ersten Hauptresultates . . . . .	37
<b>5 Die gewichtete Höhe sukzessiv errichteter Bäume</b>	<b>43</b>
5.1 Vorbereitungen . . . . .	43
5.2 Das asymptotische Verhalten des maximalen kumulierten Astgewichtes nach $n$ Sukzessionsschritten . . . . .	49
<b>6 Anwendungen</b>	<b>53</b>
6.1 Der zufällige binäre Suchbaum . . . . .	53
6.2 Der zufällige rekursive Baum . . . . .	57
6.3 Der random lopsided tree . . . . .	63
6.4 Weitere Beispiele . . . . .	65
<b>Symbolverzeichnis</b>	<b>71</b>
<b>Literaturverzeichnis</b>	<b>73</b>



# Einleitung

In vielen Anwendungen der Theorie der Gewichteten Verzweigungsprozesse ergibt sich die folgende Fragestellung: Wie schnell wächst eine bestimmte Population von Individuen, beziehungsweise irgendein anderer reproduktiver Prozess, im Hinblick auf eine seiner/ihrer immanenten Eigenschaften (vgl. die Ausführungen in [BuD], [Dev1], [Dev2], [Dev4], [Dev5], [Dev6], [Knu], [Mah], [Pit]). Dabei ist zunächst die Frage nach einer geeigneten Vergleichsskala für die Größe zu klären. In den meisten Fällen beobachtet man die größenmäßige Entwicklung des Prozesses im Laufe der Zeit. Das bedeutet, für die Wachstumsgeschwindigkeit ist zum einen die Größe der Population zu einem Zeitpunkt  $n$ , und zum anderen das asymptotische Verhalten dieser Größe für  $n \rightarrow \infty$  interessant. Dabei kann als Vergleichsskala neben dem Alter aber auch jede andere Eigenschaft des Prozesses dienen, sofern sie sich mathematisch durch gewisse stochastische Gewichte modellieren lässt. Wir nehmen in dieser Arbeit einen allgemeineren Standpunkt an, indem wir einen *zweifach gewichteten Verzweigungsprozess* betrachten. Jedes Individuum wird dabei mit zwei stochastisch unabhängigen Gewichtsfunktionen versehen, welche zwei beliebige, voneinander unabhängige Eigenschaftsausprägungen des zugrunde liegenden Prozesses repräsentieren. Fortan interessieren wir uns für das asymptotische Verhalten der maximalen Ausprägung des einen Gewichtes relativ zur Größe des anderen. Das zweite Gewicht liefert uns also die Vergleichsgröße (in obigem speziellen Fall das Alter einer Population). Für die Untersuchung dieses Verhaltens erweisen sich einige Methoden der Theorie der Großen Abweichungen als ungemein hilfreich. Im speziellen sei hier die *Fenchel-Legendre-Transformierte* der *kumulantenerzeugenden Funktion* einer Zufallsgröße  $X$  (*Cramér-Funktion* von  $X$ ) und der Satz von Cramér genannt.

Um die Hauptresultate dieser Arbeit beweisen zu können, stellen wir zunächst das notwendige Rüstzeug zusammen. Dazu führen wir im ersten Kapitel einige Grundlagen der Theorie der Großen Abweichungen ein, sofern sie für diese Arbeit relevant sind. Neben dem Satz von Cramér sind das vor allem die Eigenschaften der Kumulanten erzeugenden Funktion sowie der Cramérfunktion. Für einen tieferen Einblick in das Fachgebiet der Großen Abweichungen verweisen wir auf die Standardwerke von A. Dembo und O. Zeitouni, R. Rockafellar und F. den Hollander ([DuZ], [Roc], [Hol]).

Im zweiten Kapitel definieren wir den oben erwähnten zweifach gewichteten Verzwei-

## Einleitung

gungsprozess und stellen seine wichtigsten Eigenschaften vor. Ausgehend von einem unendlich verzweigten  $b$ -adischen Baum versehen wir jede seiner Kanten mit zwei Gewichten, die ihrerseits gewissen Unabhängigkeitseigenschaften genügen. Außerdem befassen wir uns mit den Eigenschaften der beiden zugrunde liegenden Gewichtsfunktionen.

Das dritte Kapitel stützt sich auf die Arbeit [BuD] von N. Broutin und L. Devroye. Aufbauend auf den Definitionen und Ergebnissen der beiden ersten, vorbereitenden Kapitel liefern wir das erste Hauptresultat über das oben angesprochene asymptotische Verhalten. Wir zeigen, dass sich die gewichtete Höhe  $H_n$  eines zweifach gewichteten Verzweigungsprozesses für  $n \rightarrow \infty$  asymptotisch so verhält wie  $(nc)_{n \in \mathbb{N}}$  mit einem eindeutig bestimmten  $c > 0$ , welches wir mit Hilfe der zu den beiden Gewichtsfunktionen gehörenden Cramér-Funktionen genau spezifizieren.

Im vierten Kapitel verallgemeinern wir die Situation des vorherigen Kapitels von  $b$ -adischen Bäumen auf solche mit einer beliebigen Reproduktionsverteilung  $(p_n)_{n \in \mathbb{N}_0}$  mit Erwartungswert  $\mu > 1$ . Die Methoden aus dem dritten Kapitel lassen sich mit einigen Abwandlungen auch auf die verallgemeinerte Situation anwenden, so dass wir auch in diesem Fall eine Aussage über die Asymptotik der gewichteten Höhe erhalten.

Das fünfte Kapitel befasst sich mit der Fragestellung nach der gewichteten Höhe eines Baumes mit  $n$  Knoten. In vielen Anwendungen werden stochastische Bäume sukzessiv - das heißt Knoten für Knoten durch Anhängen an den bereits vorhandenen Baum - aufgebaut. Äußerst interessant ist deshalb in diesem Zusammenhang die Frage nach dem asymptotischen Verhalten der gewichteten Höhe eines Baumes nach  $n$  Sukzessionsschritten (nach dem Anhängen des  $n$ -ten Knotens), welche wir mit  $H_n^*$  bezeichnen. Mit Hilfe der Ergebnisse des dritten Kapitels ergibt sich, dass  $H_n^*$  logarithmisch wächst, und zwar verhält sich  $H_n^*$  für  $n \rightarrow \infty$  wie  $\frac{c}{\mu} \log n$  mit dem  $c$  aus Kapitel 3 und einem  $\mu$ , welches wir noch näher spezifizieren.

Beispiele dazu sind der *zufällige binäre Suchbaum* und der *zufällige rekursive Baum*, deren Eigenschaften in den Arbeiten von D. E. Knuth, H. Mahmoud, L. Devroye und B. Pittel, ([Knu], [Dev1], [Dev2], [Dev4], [Dev5], [Dev6], [Pit]) genau studiert worden sind. Diese Beispiele befinden sich im sechsten und letzten Kapitel dieser Arbeit.

Für die Vergabe der Diplomarbeit und die Beratung und Unterstützung während der Entstehungsphase möchte ich Herrn Prof. Dr. G. Alsmeyer meinen Dank aussprechen. Außerdem danke ich all denen, die mich auf die eine oder andere Weise bei der Erstellung dieser Arbeit unterstützt haben.

# 1 Grundlagen der Theorie der Großen Abweichungen

## 1.1 Die grundlegende Problematik

Die Theorie der Großen Abweichungen beschäftigt sich mit der Asymptotik der Wahrscheinlichkeiten von sehr seltenen Ereignissen. Die exponentielle Abfallrate dieser Wahrscheinlichkeiten wird in Termen einer Variationsformel ausgedrückt, welche meist interessante Rückschlüsse auf die betrachteten Ereignisse zulässt. Dabei ist zunächst der Begriff „selten“ näher zu spezifizieren. Dazu betrachten wir zunächst die Standardsituation, welche uns zu einer grundlegenden Problematik führt (vgl. dazu die Ausführungen in [Kön]):

Gegeben seien unabhängige und identisch verteilte reellwertige Zufallsgrößen  $X_1, X_2, \dots$  mit Erwartungswert  $\mu = 0$  und Varianz  $\sigma^2 < \infty$ . Wir betrachten die Partialsummenfolge  $(S_n)_{n \in \mathbb{N}}$  mit  $S_n = \sum_{i=1}^n X_i$ . Die Zufallsgröße  $\frac{1}{n}S_n$  wird auch manchmal *empirischer Durchschnitt* von  $X_1, \dots, X_n$  genannt. Für uns interessant ist das Verhalten dieser Folge für große  $n$ . Betrachte zunächst die folgenden drei bekannten Aussagen der Wahrscheinlichkeitstheorie:

das *schwache Gesetz der großen Zahlen*:

$$\lim_{n \rightarrow \infty} \mathbb{P}(|\frac{1}{n}S_n| \geq \varepsilon) = 0 \quad \text{für jedes } \varepsilon > 0,$$

das *starke Gesetz der großen Zahlen*:

$$\mathbb{P}(\lim_{n \rightarrow \infty} \frac{1}{n}S_n = 0) = 1,$$

der *zentrale Grenzwertsatz*:

$$\lim_{n \rightarrow \infty} \mathbb{P}(\frac{1}{\sqrt{n\sigma^2}}S_n \leq C) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^C e^{-\frac{x^2}{2}} dx \quad \text{für jedes } C \in \mathbb{R}.$$

Während diese klassischen Aussagen das ‘übliche’, das ‘normale’ Verhalten von  $S_n$  beschreiben, will die Theorie der Großen Abweichungen das ‘untypische’, das Abweichungsverhalten analysieren. Präziser ausgedrückt, einer der Hauptgegenstände dieser

## 1 Grundlagen der Theorie der Großen Abweichungen

Theorie ist die asymptotische Analyse der Wahrscheinlichkeit des Ereignisses  $\{\frac{1}{n}S_n > x\}$  bzw.  $\{\frac{1}{n}S_n < -x\}$  für  $n \rightarrow \infty$ , wobei  $x > 0$ . Im starken Gesetz der großen Zahlen werden zwar gleichzeitig diese beiden Abweichungen betrachtet, aber man ist dort zufrieden mit der wenig tiefgründigen Aussage, dass die Wahrscheinlichkeiten gegen Null konvergieren. Im zentralen Grenzwertsatz entspräche dies der Wahl  $C = -x\sqrt{n}$ , über die allerdings dort keinerlei Aussagen gemacht werden.

Mit Hilfe der Theorie der Großen Abweichungen stellt sich heraus, dass unter geeigneten Annahmen an die Integrierbarkeit der  $X_i$  diese Wahrscheinlichkeit sogar *exponentiell* abfällt, also

$$\mathbb{P}\left(\frac{1}{n}S_n > x\right) \approx e^{-n\Lambda_X^*(x)} \quad \text{und} \quad \mathbb{P}\left(\frac{1}{n}S_n < -x\right) \approx e^{-n\Lambda_X^*(-x)},$$

wobei  $\Lambda_{X_1}^*(x)$  bzw.  $\Lambda_{X_1}^*(-x)$  die *Rate* des exponentiellen Abfalls ist.  $\Lambda_{X_1}^*(x)$  steht dabei für die so genannte *Cramér-Funktion* von  $X_1$ , die wir später noch spezifizieren werden.

Als erstes werden wir uns mit der so genannten *kumulantenerzeugenden Funktion* einer Zufallsgröße  $X$  befassen, deren Eigenschaften beim Studium der *Cramér-Funktion* eine entscheidende Rolle spielen.

## 1.2 Eigenschaften der kumulantenerzeugende Funktion

Wir beschränken uns in diesem Kapitel manchmal auf den Fall einer reellwertigen Zufallsgröße  $X \geq 0$ , da es sich bei den in den folgenden Kapiteln betrachteten Gewichtsfunktionen grundsätzlich um positive Zufallsgrößen handelt.

**Definition 1.1** Sei  $X$  eine reellwertige Zufallsgröße. Dann heißt für  $\lambda \in \mathbb{R}$

$$M_X(\lambda) \stackrel{\text{def}}{=} \mathbb{E}(e^{\lambda X})$$

die *momenterzeugende Funktion* von  $X$ .

**Lemma 1.2** Sei  $X$  eine nichtnegative reellwertige Zufallsgröße. Dann ist der Definitionsbereich ihrer momenterzeugenden Funktion von der Form  $(-\infty; \lambda^*)$  bzw.  $(-\infty; \lambda^*]$ , mit  $\lambda^* \stackrel{\text{def}}{=} \sup\{\lambda \in \mathbb{R} \mid M_X(\lambda) < \infty\} \in [0; \infty]$ . Außerdem ist  $M_X(\lambda)$  monoton wachsend auf  $\{\lambda \in \mathbb{R} \mid M_X(\lambda) < \infty\}$ , und es gilt  $M_X(0) = 1$ .

BEWEIS: (vgl. [Als1], Lemma 40.2 und 40.3 Definition)

Seien  $X \geq 0$  und  $\lambda \in \mathbb{R}$  mit  $M_X(\lambda) < \infty$  und sei  $\lambda_0 \leq \lambda$ . Dann ist  $e^{\lambda_0 X} \leq e^{\lambda X}$ , da  $X$  nicht-negativ und die Exponentialfunktion monoton wachsend ist. Deshalb gilt

$$M_X(\lambda_0) = \mathbb{E}(e^{\lambda_0 X}) \leq \mathbb{E}(e^{\lambda X}) = M_X(\lambda)$$

aufgrund der Monotonie des Erwartungswertes. Daraus ergibt sich die behauptete Gestalt des Definitionsbereiches von  $M_X$ . Außerdem gilt  $\lambda^* \geq 0$  wegen  $M_X(0) = \mathbb{E}(e^{0X}) = \mathbb{E}(1) = 1 < \infty$ . Dabei ist es selbstverständlich möglich, dass  $\lambda^* = 0$  bzw.  $\lambda^* = \infty$  gilt. □

**Bemerkung 1.3** Die Monotonie von  $M_X(\lambda)$  mit  $X$  wie in Lemma 1.2 ist genau dann streng, wenn  $X$  nicht fast sicher konstant gleich 0 ist.

**Lemma 1.4 (Existenz der Momente)** Sei  $X$  eine nichtnegative reellwertige Zufallsgröße und  $M_X(\lambda) < \infty$  für ein  $\lambda > 0$ . Dann gilt

$$\mathbb{E}(X^r) < \infty \quad \forall r > 0.$$

## 1 Grundlagen der Theorie der Großen Abweichungen

BEWEIS: Sei  $\lambda > 0$  mit  $M_X(\lambda) < \infty$  und sei  $r > 0$  beliebig, dann existiert ein  $s > 0$  mit  $e^{\lambda t} \geq t^r$  für alle  $t \geq s$ . Dann liefert eine einfache Abschätzung die Behauptung.  $\square$

**Definition 1.5** Sei  $X$  eine reellwertige Zufallsgröße. Dann heißt für  $\lambda \in \mathbb{R}$

$$\Lambda_X(\lambda) \stackrel{\text{def}}{=} \log(M_X(\lambda))$$

die *kumulantenerzeugende Funktion* von  $X$ . Ihren Definitionsbereich bezeichnen wir mit  $\mathcal{D}_X$ .

Die kumulantenerzeugende Funktion weist einige Eigenschaften auf, welche uns bei der Betrachtung der *Cramér-Funktion* zu interessanten Ergebnissen führen werden. Neben der Differenzierbarkeit ist eine dieser Eigenschaften die der (strikten) Konvexität, welche wir in folgendem Lemma beweisen:

**Lemma 1.6 (Strikte Konvexität der kumulantenerzeugenden Funktion)** *Die kumulantenerzeugende Funktion  $\Lambda_X$  einer reellwertigen Zufallsgröße  $X$  ist konvex, und es gilt strikte Konvexität, falls  $X$  nicht fast sicher konstant ist.*

BEWEIS: Wie man sehr leicht einsieht, ist der Definitionsbereich von  $\Lambda_X$  gleich dem Definitionsbereich von  $M_X$ . Außerdem folgt mit Hilfe der strengen Monotonie der Logarithmusfunktion und der (strengen) Monotonie von  $M_X$  die (strenge) Monotonie von  $\Lambda_X$  innerhalb ihres Definitionsbereiches.

Die (strikte) Konvexität von  $\Lambda_X$  gewinnt man durch Anwendung der Hölderschen Ungleichung. Für  $\lambda_1$  und  $\lambda_2$  aus dem Definitionsbereich von  $\Lambda_X$  mit  $\lambda_1 < \lambda_2$  und  $t \in (0; 1)$  beliebig gilt

$$\begin{aligned} \Lambda_X(t\lambda_1 + (1-t)\lambda_2) &= \log \mathbb{E}[(e^{\lambda_1 X})^t (e^{\lambda_2 X})^{1-t}] \\ &\leq \log [(\mathbb{E}(e^{\lambda_1 X}))^t (\mathbb{E}(e^{\lambda_2 X}))^{1-t}] \\ &= t\Lambda_X(\lambda_1) + (1-t)\Lambda_X(\lambda_2). \end{aligned}$$

Dabei steht in der zweiten Zeile das  $<$ - Zeichen genau dann, wenn  $X$  nicht f.s. konstant ist.  $\square$

**Lemma 1.7 (Differenzierbarkeit der kumulantenerzeugenden Funktion)** *Die Funktion  $\Lambda_X$  ist stetig differenzierbar auf dem offenen Kern  $\mathcal{D}_X^\circ$  ihres Definitionsbereiches.*

## 1.2 Eigenschaften der kumulantenerzeugende Funktion

Insbesondere ist  $\Lambda_X$  stetig und es gilt

$$\Lambda'_X(\lambda) = \frac{1}{M_X(\lambda)} \mathbb{E}(X e^{\lambda X}).$$

BEWEIS: Die Differenzierbarkeit von  $\Lambda_X(\lambda) = \log \mathbb{E}(e^{\lambda X})$  und die Gestalt der Ableitung folgen aus dem Satz von der *Differentiation unter dem Integralzeichen* (vgl. Kapitel 4, Satz 5.7 und Zusatz in [Els]). Es gilt nämlich:

- (i) Für  $\lambda \in \mathcal{D}_X^\circ$  ist  $e^{\lambda X} \in \mathcal{L}^1$ .
- (ii) Für alle  $\omega \in \Omega$  existiert die partielle Ableitung  $\frac{\partial}{\partial \lambda} e^{\lambda X(\omega)} = X(\omega) e^{\lambda X(\omega)}$ .
- (iii) Für  $\lambda \in \mathcal{D}_X^\circ$  beliebig und  $\epsilon > 0$  so klein, dass  $[\lambda - \epsilon; \lambda + \epsilon] \subseteq \mathcal{D}_X^\circ$  gilt:  
Es existiert ein  $n \in \mathbb{N}$  mit  $x < e^{\epsilon x} \quad \forall x > n$ . Mit diesem  $n$  folgt

$$X e^{\lambda X} = \mathbf{1}_{[-n;n]}(X) X e^{\lambda X} + \mathbf{1}_{(-\infty;-n)}(X) X e^{\lambda X} + \mathbf{1}_{(n;\infty)}(X) X e^{\lambda X}.$$

Dabei ist der erste Summand beschränkt. Für den zweiten und den dritten Summand gelten die folgenden Ungleichungsketten:

$$0 \geq \mathbf{1}_{(-\infty;-n)}(X) X e^{\lambda X} \geq \mathbf{1}_{(-\infty;-n)}(X) (-e^{\epsilon|X|}) e^{\lambda X} = \mathbf{1}_{(-\infty;-n)}(X) (-e^{(\lambda-\epsilon)X}),$$

beziehungsweise

$$0 \leq \mathbf{1}_{(n;\infty)}(X) X e^{\lambda X} \leq \mathbf{1}_{(n;\infty)}(X) e^{\epsilon X} e^{\lambda X} = \mathbf{1}_{(n;\infty)}(X) e^{(\lambda+\epsilon)X}.$$

Dabei sind  $e^{(\lambda-\epsilon)X}$  und  $e^{(\lambda+\epsilon)X}$  integrierbar nach Wahl von  $\epsilon$  und damit sind auch  $X e^{\lambda X}$  und  $|X e^{\lambda X}|$  integrierbar.

Die Stetigkeit der Ableitung folgt aus dem Satz von der *Stetigen Abhängigkeit des Integrals von einem Parameter* (vgl. Kapitel 4, Satz 5.6 in [Els]), denn es gilt:

- (i) Für  $\lambda \in \mathcal{D}_X^\circ$  ist  $X e^{\lambda X} \in \mathcal{L}^1$  (siehe oben).
- (ii) Für alle  $\omega \in \Omega$  und  $\lambda \in \mathcal{D}_X^\circ$  ist  $X(\omega) e^{\lambda X(\omega)}$  stetig in  $\lambda$ .
- (iii) Nach obiger Überlegung gilt:  $\mathbb{E}(|X e^{\lambda X}|) < \infty$ .

□

**Lemma 1.8** Für die kumulantenerzeugende Funktion  $\Lambda_X$  einer reellwertigen Zufallsgröße  $X$  mit  $\mathbb{E}(X) = \mu$  gilt:

(i)  $\Lambda_X(0) = 0$ .

(ii)  $\Lambda_X(\lambda) \geq \lambda\mu$  für alle  $\lambda \in \mathcal{D}_X$ .

BEWEIS: Aussage (i) folgt aus Lemma 1.2 und Anwendung der Logarithmusfunktion. Aufgrund der Konkavität der Logarithmusfunktion folgt mit Hilfe der Jensen-Ungleichung

$$\Lambda_X(\lambda) = \log \mathbb{E}(e^{\lambda X}) \geq \mathbb{E}(\log e^{\lambda X}) = \lambda \mathbb{E}(X) = \lambda\mu.$$

Und damit ist auch (ii) gezeigt. □

Für die Untersuchung des Definitionsbereiches der Cramér-Funktion im folgenden Abschnitt ist ein genaueres Studium der Asymptotik der ersten Ableitung der kumulantenerzeugenden Funktion notwendig. Dafür benötigen wir zunächst zwei weitere Begriffe der Wahrscheinlichkeitstheorie.

**Definition 1.9** Für eine reellwertige Zufallsgröße  $X$  heißen

$$\text{esssup}(X) \stackrel{\text{def}}{=} \inf_{N: P(N)=0} \sup_{\omega \in \Omega \setminus N} X(\omega) \in (-\infty; \infty]$$

das *essentielle Supremum* von  $X$  und

$$\text{essinf}(X) \stackrel{\text{def}}{=} \sup_{N: P(N)=0} \inf_{\omega \in \Omega \setminus N} X(\omega) \in [-\infty; \infty)$$

das *essentielle Infimum* von  $X$ .

**Lemma 1.10** Die Abbildung  $\lambda \rightarrow \Lambda'_X(\lambda)$  hat die Asymptoten  $\text{esssup}(X)$  für  $\lambda \rightarrow \infty$  und  $\text{essinf}(X)$  für  $\lambda \rightarrow -\infty$ .

BEWEIS: Definiere  $\alpha \stackrel{\text{def}}{=} \text{essinf}(X)$  und  $\beta \stackrel{\text{def}}{=} \text{esssup}(X)$ . Dann gilt

$$\Lambda'_X(\lambda) = \frac{\mathbb{E}(X e^{\lambda X})}{\mathbb{E}(e^{\lambda X})} = \begin{cases} \geq \frac{\mathbb{E}(\alpha e^{\lambda X})}{\mathbb{E}(e^{\lambda X})} = \alpha \\ \leq \frac{\mathbb{E}(\beta e^{\lambda X})}{\mathbb{E}(e^{\lambda X})} = \beta. \end{cases}$$

## 1.2 Eigenschaften der kumulantenerzeugende Funktion

Daraus folgt  $\Lambda'_X(\lambda) \in [\alpha; \beta] \quad \forall \lambda \in \mathbb{R}$ . Sei nun  $\varepsilon > 0$  beliebig und  $M_\varepsilon \stackrel{\text{def}}{=} \{X > \beta - \varepsilon\}$ . Nach Definition von  $\text{esssup}(X)$  ist  $\mathbb{P}(M_\varepsilon) > 0$ . Dann gilt für  $\lambda > 0$

$$\begin{aligned} \frac{\mathbb{E}(Xe^{\lambda X})}{\mathbb{E}(e^{\lambda X})} &\geq \frac{\mathbb{E}(\mathbf{1}_{M_\varepsilon}(X)(\beta - \varepsilon)e^{\lambda X})}{\mathbb{E}(\mathbf{1}_{M_\varepsilon}(X)e^{\lambda X}) + \mathbb{E}(\mathbf{1}_{M_\varepsilon^c}(X)e^{\lambda X})} \\ &= (\beta - \varepsilon) \frac{\mathbb{E}(\mathbf{1}_{M_\varepsilon}(X)e^{\lambda X})}{\mathbb{E}(\mathbf{1}_{M_\varepsilon}(X)e^{\lambda X}) + \mathbb{E}(\mathbf{1}_{M_\varepsilon^c}(X)e^{\lambda X})} \\ &\geq (\beta - \varepsilon) \frac{\mathbb{E}(\mathbf{1}_{M_\varepsilon}(X)e^{\lambda X})}{\mathbb{E}(\mathbf{1}_{M_\varepsilon}(X)e^{\lambda X}) + \mathbb{P}(M_\varepsilon^c)e^{\lambda(\beta - \varepsilon)}} \\ &\xrightarrow{\lambda \rightarrow \infty} (\beta - \varepsilon), \end{aligned}$$

denn

$$\begin{aligned} \frac{\mathbb{E}(\mathbf{1}_{M_\varepsilon}(X)e^{\lambda X})}{e^{\lambda(\beta - \varepsilon)}} &\geq \frac{\mathbb{E}(\mathbf{1}_{M_{\frac{\varepsilon}{2}}}(X)e^{\lambda X})}{e^{\lambda(\beta - \varepsilon)}} \\ &\geq \mathbb{P}(M_{\frac{\varepsilon}{2}})e^{\lambda \frac{\varepsilon}{2}} \\ &\xrightarrow{\lambda \rightarrow \infty} \infty \quad (\text{da } \mathbb{P}(M_{\frac{\varepsilon}{2}}) > 0). \end{aligned}$$

Daraus ergibt sich

$$\frac{e^{\lambda(\beta - \varepsilon)}}{\mathbb{E}(\mathbf{1}_{M_\varepsilon}(X)e^{\lambda X})} \longrightarrow 0.$$

Es folgt obige Behauptung und damit, da  $\varepsilon > 0$  beliebig gewählt war,

$$\lim_{\lambda \rightarrow \infty} \frac{\mathbb{E}(Xe^{\lambda X})}{\mathbb{E}(e^{\lambda X})} \geq \beta.$$

Also besitzt  $\lambda \rightarrow \Lambda'_X(\lambda)$  die Aymptote  $\text{esssup}(X)$  für  $\lambda \rightarrow \infty$ . Analog zeigt man, dass  $\lambda \rightarrow \Lambda'_X(\lambda)$  für  $\lambda \rightarrow -\infty$  die Aymptote  $\text{essinf}(X)$  besitzt. Dafür definiert man sich Mengen  $N_\varepsilon \stackrel{\text{def}}{=} \{X < \alpha + \varepsilon\}$  und verfährt ähnlich wie oben.

□

## 1.3 Eigenschaften der Cramér-Funktion

In diesem Abschnitt beschäftigen wir uns mit einer der wichtigsten Funktionen aus dem Gebiet der Großen Abweichungen, der *Fenchel-Legendre-Transformierten* der kumulantenerzeugenden Funktion einer reellwertigen Zufallsgröße  $X$ . Wir werden diese zunächst definieren und daraufhin ihre wichtigsten Eigenschaften zusammentragen, welche sich für uns im Hinblick auf den *Satz von Cramér* und die späteren Untersuchungen in Kapitel 3 als sehr nützlich erweisen werden. Wir werden zeigen, dass die Fenchel-Legendre-Transformierte für die Beschreibung der eingangs erwähnten exponentiellen Abfallrate von  $\mathbb{P}(\frac{1}{n}S_n > x)$  bestens geeignet ist. Die Eigenschaften dieser im Folgenden auch als *Cramér-Funktion* bezeichneten Funktion - insbesondere ihre Differenzierbarkeit, ihre strikte Konvexität und ihre genaue Gestalt auf gewissen Intervallen - fließen maßgeblich in den Beweis des ersten Hauptresultates dieser Arbeit ein.

Selbstverständlich ließen sich die folgenden Definitionen und einige der Lemmata auch auf den Fall von Zufallsvariablen mit Werten in  $\mathbb{R}^n$  verallgemeinern. Diese Betrachtung ist für die vorliegende Arbeit jedoch irrelevant. Weitergehende Untersuchungen finden sich zum Beispiel in [Roc].

**Definition 1.11** Für die kumulantenerzeugende Funktion  $\Lambda_X$  einer reellwertigen Zufallsgröße  $X$  heißt

$$\Lambda_X^*(t) \stackrel{\text{def}}{=} \sup_{\lambda \in \mathbb{R}} (\lambda t - \Lambda_X(\lambda)), \quad t \in \mathbb{R}, \quad (1.1)$$

die *Fenchel-Legendre-Transformierte* von  $\Lambda_X$ . Im Folgenden wird  $\Lambda_X^*$  auch *Cramér-Funktion* von  $X$  genannt. Ihren Definitionsbereich bezeichnen wir mit  $\mathcal{D}_X^*$ .

**Definition 1.12** Für  $X$  und  $\Lambda_X$  wie oben heißt

$$\Lambda_{l,X}^*(t) \stackrel{\text{def}}{=} \sup_{\lambda \leq 0} (\lambda t - \Lambda_X(\lambda)), \quad t \in \mathbb{R},$$

die *linksseitige Cramér-Funktion* und

$$\Lambda_{r,X}^*(t) \stackrel{\text{def}}{=} \sup_{\lambda \geq 0} (\lambda t - \Lambda_X(\lambda)), \quad t \in \mathbb{R},$$

die *rechtsseitige Cramér-Funktion* von  $X$ .

Aus diesen Definitionen ergeben sich bereits einige grundlegende Eigenschaften, welche wir im folgenden Satz zusammenfassen. Im Hinblick auf die späteren Betrachtungen

in Kapitel 3 interessiert uns dabei vor allem die Gestalt von  $\Lambda_X^*$  auf den Intervallen  $(-\infty; \mu]$  (für den Fall, dass  $\Lambda_X(\lambda) < \infty$  für ein  $\lambda < 0$  gilt) und  $[\mu; \infty)$  (für den Fall, dass  $\Lambda_X(\lambda) < \infty$  für ein  $\lambda > 0$  gilt).

**Satz 1.13 (Grundlegende Eigenschaften)** Sei  $\Lambda_X^*$  die Cramér-Funktion einer reellwertigen Zufallsgröße  $X$  mit Erwartungswert  $\mathbb{E}(X) = \mu \in \mathbb{R}$ . Dann gilt:

- (i) Im Falle  $\mathcal{D}_X = \{0\}$  ist  $\Lambda_X^*$  konstant gleich 0.
- (ii) Es gilt  $\Lambda_X^*(\mu) = 0$  und  $\Lambda_X^*(t) \geq 0$  für alle  $t \in \mathbb{R}$ .
- (iii) Ist  $\Lambda(\lambda_0) < \infty$  für ein  $\lambda_0 > 0$ , dann gilt für alle  $t \geq \mu$

$$\Lambda_X^*(t) = \sup_{\lambda \geq 0} (\lambda t - \Lambda_X(\lambda)) = \Lambda_{r,X}^*(t),$$

und  $\Lambda_{r,X}^*(t)$  ist monoton wachsend in  $t$ . Gilt ferner  $\lambda^* = \sup \{\lambda \geq 0 \mid \Lambda_X(\lambda) < \infty\} < \infty$ , so ist

$$\Lambda_X^*(t) = \sup_{0 \leq \lambda \leq \lambda^*} (\lambda t - \Lambda_X(\lambda)).$$

- (iv) Ist  $\Lambda(\lambda_0) < \infty$  für ein  $\lambda_0 < 0$ , dann gilt für alle  $t \leq \mu$

$$\Lambda_X^*(t) = \sup_{\lambda \leq 0} (\lambda t - \Lambda_X(\lambda)) = \Lambda_{l,X}^*(t),$$

und  $\Lambda_{l,X}^*(t)$  ist monoton fallend in  $t$ .

- (v) Es gilt  $\Lambda_{r,X}^*(t) = 0$  für alle  $t < \mu$  und  $\Lambda_{l,X}^*(t) = 0$  für alle  $t > \mu$ .

BEWEIS: Zu (i): Sei  $\mathcal{D}_X = \{0\}$ , also  $\Lambda_X(\lambda) = \infty$  für alle  $\lambda \in \mathbb{R} \setminus \{0\}$  und damit  $\Lambda_X^*(t) = 0t - \Lambda_X(0) = 0$  für alle  $t \in \mathbb{R}$ .

Zu (ii): Es gilt zum einen

$$\Lambda_X^*(t) = \sup_{\lambda \in \mathbb{R}} (\lambda t - \Lambda_X(\lambda)) \geq 0 - \Lambda_X(0) = 0$$

und zum anderen

$$\Lambda_X^*(\mu) = \sup_{\lambda \in \mathbb{R}} (\lambda \mu - \Lambda_X(\lambda)) \stackrel{1.8}{\leq} \sup_{\lambda \in \mathbb{R}} (\lambda \mu - \lambda \mu) = 0.$$

Zu (iii): Sei  $\Lambda_X(\lambda_0) < \infty$  für ein  $\lambda_0 > 0$  und sei  $t \geq \mu$  beliebig. Dann gilt für jedes  $\lambda < 0$

$$\lambda t - \Lambda_X(\lambda) \leq \lambda \mu - \Lambda_X(\lambda) \leq \Lambda_X^*(\mu) \stackrel{(i)}{=} 0.$$

Mit  $\Lambda_X^*(t) \geq 0$  folgt  $\Lambda_X^*(t) = \Lambda_{r,X}^*(t)$ . Offensichtlich ist  $\Lambda_{r,X}^*(t)$  monoton wachsend in  $t$ . Gelte ferner  $\lambda^* = \sup \{\lambda \geq 0 \mid \Lambda_X(\lambda) < \infty\} < \infty$ . Für jedes  $\lambda > \lambda^*$  ist dann  $\Lambda_X(\lambda) = \infty$  und damit  $\lambda t - \Lambda_X(\lambda) = -\infty$  unabhängig von der Wahl von  $t \in \mathbb{R}$ . Daraus folgt  $\Lambda_X^*(t) = \sup_{0 \leq \lambda \leq \lambda^*} (\lambda t - \Lambda_X(\lambda))$ .

Zu (iv): Sei  $\Lambda_X(\lambda_0) < \infty$  für ein  $\lambda_0 < 0$  und sei  $t \leq \mu$  beliebig. Analog zu (iii) gilt dann für jedes  $\lambda > 0$

$$\lambda t - \Lambda_X(\lambda) \leq \lambda \mu - \Lambda_X(\lambda) \leq \Lambda_X^*(\mu) \stackrel{(i)}{=} 0.$$

Daraus folgt  $\Lambda_X^*(t) = \Lambda_{l,X}^*(t)$ , und  $\Lambda_{l,X}^*(t)$  ist offensichtlich monoton fallend in  $t$ .

Zu (v): Nach Lemma 1.9 b) gilt für  $t < \mu$  beliebig

$$\Lambda_{r,X}^*(t) = \sup_{\lambda \geq 0} (\lambda t - \Lambda_X(\lambda)) \leq \sup_{\lambda \geq 0} (\lambda t - \lambda \mu) = \sup_{\lambda \geq 0} \lambda(t - \mu) = 0$$

und für  $t > \mu$  beliebig

$$\Lambda_{l,X}^*(t) = \sup_{\lambda \leq 0} (\lambda t - \Lambda_X(\lambda)) \leq \sup_{\lambda \leq 0} (\lambda t - \lambda \mu) = \sup_{\lambda \leq 0} \lambda(t - \mu) = 0.$$

Mit  $\Lambda_{r,X}^*, \Lambda_{l,X}^* \geq 0$  folgt die Behauptung. □

Als nächstes wollen wir uns mit dem Definitionsbereich der Cramér-Funktion auseinandersetzen. Wie sich herausstellen wird, besteht diesbezüglich ein sehr enger Zusammenhang zu dem Wertebereich der ersten Ableitung der kumulantenerzeugenden Funktion.

**Lemma 1.14 (Existenz und Eindeutigkeit des Maximierers)** *Sei  $\Lambda_X^*$  die Cramér-Funktion einer nicht fast sicher konstanten reellwertigen Zufallsgröße  $X$ . Dann wird für jedes  $t \in (\text{essinf}(X); \text{esssup}(X))$  das Supremum in (1.1) in genau einem  $\lambda_t \in \mathbb{R}$  angenommen, das heißt, es gilt*

$$\Lambda_X^*(t) = \lambda_t t - \Lambda_X(\lambda_t).$$

Für dieses  $\lambda_t$  gilt ferner

$$\Lambda_X'(\lambda_t) = t.$$

BEWEIS:

*Eindeutigkeit:* Sei  $\lambda_t$  ein Maximierer, das heißt  $\Lambda_X^*(t) = \lambda_t t - \Lambda_X(\lambda_t)$ . Dann gilt  $0 = t - \Lambda_X'(\lambda_t)$  und es folgt  $\Lambda_X'(\lambda_t) = t$ .  $\Lambda_X$  ist strikt konvex, da nach Voraussetzung  $X$  nicht fast sicher konstant ist (siehe Lemma 1.6). Damit ist  $\Lambda_X'$  streng monoton wachsend und es folgt die Eindeutigkeit.

*Existenz:* Nach Lemma 1.10 existiert zu  $t \in (\text{essinf}(X); \text{esssup}(X))$  ein  $\lambda_t \in \mathbb{R}$  mit  $\Lambda'_X(\lambda_t) = t$ . Also ist  $\lambda_t$  eine Nullstelle der Ableitung von  $\lambda \mapsto \lambda t - \Lambda_X(\lambda)$  und damit wegen der strengen Monotonie von  $\Lambda'_X$  ein Maximierer.

□

**Lemma 1.15** *Für die Cramér-Funktion  $\Lambda_X^*$  einer Zufallsgröße  $X$  gilt*

$$\Lambda_X^*(t) = \begin{cases} \infty & \text{für } t \notin [\text{essinf}(X); \text{esssup}(X)] \\ < \infty & \text{für } t \in (\text{essinf}(X); \text{esssup}(X)). \end{cases}$$

BEWEIS: Sei  $t > \text{esssup}(X)$  beliebig. Dann gilt für alle  $\lambda > 0$

$$\begin{aligned} \Lambda_X^*(t) &\geq \lambda t - \Lambda_X(\lambda) \\ &= \lambda \left( t - \frac{1}{\lambda} \Lambda_X(\lambda) \right) \end{aligned}$$

und folglich

$$\begin{aligned} \Lambda_X^*(t) &\geq \limsup_{\lambda \rightarrow \infty} \lambda \left( t - \frac{1}{\lambda} \Lambda_X(\lambda) \right) \\ &= \limsup_{\lambda \rightarrow \infty} \lambda \left( t - \frac{\log \mathbb{E}(e^{\lambda X})}{\lambda} \right) \\ &= \infty, \end{aligned}$$

denn

$$\begin{aligned} t - \frac{\log \mathbb{E}(e^{\lambda X})}{\lambda} &= \frac{\log e^{\lambda t}}{\lambda} - \frac{\log \mathbb{E}(e^{\lambda X})}{\lambda} \\ &= -\frac{1}{\lambda} (\log \mathbb{E}(e^{\lambda X}) - \log e^{\lambda t}) \\ &= -\frac{1}{\lambda} \log e^{-\lambda t} \mathbb{E}(e^{\lambda X}) \\ &= -\frac{1}{\lambda} \log \mathbb{E}(e^{\lambda(X-t)}) \\ &\stackrel{\lambda > 0}{\geq} -\frac{1}{\lambda} \log \mathbb{E}(e^{\lambda(\text{esssup}(X)-t)}) \\ &= -(\text{esssup}(X) - t) \\ &= t - \text{esssup}(X) > 0. \end{aligned}$$

Ein ähnliches Argument liefert  $\Lambda_X^*(t) = \infty$  für  $t < \text{essinf}(X)$ . Die Aussage  $\Lambda_X^*(t) < \infty$  für  $t \in (\text{essinf}(X); \text{esssup}(X))$  folgt aus Lemma 1.14.

□

**Bemerkung 1.16** Nach Lemma 1.10 und 1.15 gilt

$$\text{cl}(\mathcal{D}_X^*) = [\text{essinf}(X); \text{esssup}(X)] = \text{cl}(\{\Lambda'_X(\lambda) \mid \lambda \in \mathbb{R}\}),$$

wobei mit  $\text{cl}$  der topologische Abschluss gemeint ist.

**Lemma 1.17** Für die Cramér-Funktion  $\Lambda_X^*$  einer reellwertigen Zufallsgröße  $X$  mit Erwartungswert  $\mathbb{E}(X) = \mu$  gilt

$$\Lambda_X^*(t) = 0 \quad \Leftrightarrow \quad t = \mu.$$

BEWEIS: Ist  $X$  fast sicher konstant gleich  $\mu$ , so ist  $\Lambda_X^*(t) = \infty$  für  $t \neq \mu$ , und es ist nichts weiter zu zeigen. Sei also jetzt  $X$  nicht fast sicher konstant und sei  $t \in \mathbb{R}$  mit  $\Lambda_X^*(t) = 0$ . Dann ist  $\lambda_t = 0$  der eindeutig bestimmte Maximierer im Sinne von Lemma 1.14. Mittels Lemma 1.14 folgt ferner

$$\Lambda'_X(\lambda_t) = \Lambda'_X(0) = \frac{\mathbb{E}(X e^{0X})}{\mathbb{E}(e^{0X})} = \mathbb{E}(X) = \mu.$$

□

Ebenfalls von sehr großer Bedeutung für die späteren Ergebnisse sind die strikte Konvexität und die Differenzierbarkeit der Cramér-Funktion. Mit Hilfe der (strengen) Monotonie von  $\Lambda'_X$  lässt sich das folgende Resultat beweisen:

**Satz 1.18 (Strikte Konvexität der Cramér-Funktion)** Die Cramér-Funktion  $\Lambda_X^*$  einer reellwertigen Zufallsgröße  $X$  ist strikt konvex auf  $(\text{essinf}(X); \text{esssup}(X))$ .

BEWEIS: Für den Beweis der strikten Konvexität benutzen wir die in Lemma 1.14 gewonnene geometrische Interpretation von  $\Lambda_X^*$  auf  $(\text{essinf}(X); \text{esssup}(X))$ : Für  $t_1, t_2 \in (\text{essinf}(X); \text{esssup}(X))$  mit  $t_1 < t_2$  und  $\theta \in (0, 1)$  beliebig gilt

$$\begin{aligned} \theta \Lambda_X^*(t_1) + (1 - \theta) \Lambda_X^*(t_2) &= \theta \sup_{\lambda \in \mathbb{R}} (\lambda t_1 - \Lambda_X(\lambda)) + (1 - \theta) \sup_{\lambda \in \mathbb{R}} (\lambda t_2 - \Lambda_X(\lambda)) \\ &= \sup_{\lambda \in \mathbb{R}} (\theta \lambda t_1 - \theta \Lambda_X(\lambda)) + \sup_{\lambda \in \mathbb{R}} ((1 - \theta) \lambda t_2 - (1 - \theta) \Lambda_X(\lambda)) \\ &\geq \sup_{\lambda \in \mathbb{R}} ((\theta t_1 + (1 - \theta) t_2) \lambda - \Lambda_X(\lambda)) \\ &= \Lambda_X^*(\theta t_1 + (1 - \theta) t_2). \end{aligned}$$

Dabei steht hier in der vorletzten Zeile das Gleichheitszeichen genau dann, wenn beide Suprema in der drittletzten Zeile im gleichen Punkt angenommen werden. Das bedeutet,

für die Gleichheit muss  $\lambda_{t_1} = \lambda_{t_2}$  gelten (mit  $\lambda_{t_i}$  ( $i = 1, 2$ ) wie in Lemma 1.14). Nach Lemma 1.14 gilt aber  $\Lambda'_X(\lambda_{t_1}) = t_1 < t_2 = \Lambda'_X(\lambda_{t_2})$  und damit  $\lambda_{t_1} < \lambda_{t_2}$ , weil  $\Lambda'_X$  (streng) monoton wachsend ist. Es gilt also das  $>$ -Zeichen und damit ist  $\Lambda_X^*$  strikt konvex.  $\square$

**Satz 1.19 (Differenzierbarkeit der Cramér-Funktion)** *Die Cramér-Funktion  $\Lambda_X^*$  einer reellwertigen Zufallsgröße  $X$  ist differenzierbar auf  $(\text{essinf}(X); \text{esssup}(X))$ .*

BEWEIS: Der Beweis stützt sich auf den Satz von der *Differenzierbarkeit implizit definierter Funktionen* (vgl. Satz 170.1 auf Seite 295 in [Heu]).

Für  $t \in (\text{essinf}(X); \text{esssup}(X))$  ist  $\Lambda_X^*(t) = \lambda_t t - \Lambda_X(\lambda_t)$  mit  $\lambda_t$  wie in Lemma 1.14. Dabei ist  $\lambda_t$  nach Lemma 1.14 und wegen der strengen Monotonie von  $\Lambda'_X$  genau dann ein Maximierer von (1.1), wenn  $\lambda_t$  Nullstelle der Funktion  $\lambda \mapsto t - \Lambda'_X(\lambda)$  ist. Wir definieren

$$\mathcal{F} : U \times V \rightarrow \mathbb{R}, \quad \mathcal{F}(t, \lambda) \stackrel{\text{def}}{=} t - \Lambda'_X(\lambda),$$

wobei  $U \subseteq (\text{essinf}(X); \text{esssup}(X))$  eine offene Umgebung von  $t$  und  $V$  eine offene Umgebung von  $\lambda_t$  darstellt, auf welcher die Funktion  $\Lambda'_X$  stetig differenzierbar ist. Die stetige Differenzierbarkeit von  $\Lambda'_X$  zeigt man analog zum Beweis von Lemma 1.7 (Differenzierbarkeit der kumulantenenerzeugenden Funktion) mit Hilfe der Sätze 5.6 und 5.7, Kapitel 4 in [Els]. Zum einen gilt wegen der strengen Monotonie von  $\Lambda'_X$

$$\frac{\partial}{\partial \lambda} \mathcal{F}(t, \lambda) = -\Lambda''_X(\lambda) < 0,$$

und zum anderen gilt

$$\frac{\partial}{\partial t} \mathcal{F}(t, \lambda) = 1.$$

Daher ist die Funktion  $\mathcal{F}(t, \lambda)$  stetig partiell differenzierbar. Zusammen mit  $\mathcal{F}(t, \lambda_t) = 0$  und  $\frac{\partial}{\partial \lambda} \mathcal{F}(t, \lambda) \neq 0$  folgt aus dem Satz von der Differenzierbarkeit implizit definierter Funktionen die Existenz einer an der Stelle  $t$  differenzierbaren Funktion  $f : G \rightarrow H$  mit geeigneten offenen Teilmengen  $G \subseteq U$  und  $H \subseteq V$  und  $\mathcal{F}(t, f(t)) = 0 \quad \forall t \in G$ . Nach Lemma 1.14 (insbesondere der Eindeutigkeit des Maximierers) gilt  $\Lambda_X^*(t) = f(t)t - \Lambda_X(f(t))$  und damit ist  $\Lambda_X^*$  differenzierbar als Verknüpfung von differenzierbaren Funktionen. Da  $t \in (\text{essinf}(X); \text{esssup}(X))$  beliebig gewählt war, folgt die Behauptung.  $\square$

**Folgerung 1.20 (Ableitung an der Stelle  $\mu$ )** *Für die Cramér-Funktion  $\Lambda_X^*$  einer nicht fast sicher konstanten reellwertigen Zufallsgröße  $X$  mit Erwartungswert  $\mathbb{E}(X) = \mu$  gilt*

$$\frac{d}{dt} \Lambda_X^*(t)|_{t=\mu} = 0.$$

BEWEIS: Die Behauptung folgt mit Hilfe der Ergebnisse des Lemmas 1.17 und den Sätzen 1.18 und 1.19. Die Cramér-Funktion  $\Lambda_X^*$  ist strikt konvex und differenzierbar auf  $(\text{essinf}(X); \text{esssup}(X))$ . Dabei liegt  $\mu$  in  $(\text{essinf}(X); \text{esssup}(X))$ , weil  $X$  nach Voraussetzung nicht fast sicher konstant ist. Außerdem gilt  $\Lambda_X^*(t) \geq 0$ , wobei die Gleichheit nur für  $t = \mu$  gilt. Demzufolge ist  $\mu$  ein lokales Minimum von  $\Lambda_X^*$  und daraus folgt die Behauptung.  $\square$

Im Hinblick auf die späteren Ausführungen benötigen wir noch ein weiteres Lemma, das auf die Situation in Kapitel 3 zugeschnitten ist:

**Lemma 1.21** *Sei  $X$  eine nicht fast sicher konstante, nichtnegative Zufallsgröße mit  $\mathbb{E}(X) < \infty$  und  $\mathbb{P}(X = 0) = 0$ . Dann gilt für ihre linksseitige Cramér-Funktion*

$$\lim_{t \downarrow \text{essinf}(X)} \Lambda_{t,X}^*(t) = \infty.$$

BEWEIS: Es gilt  $\Lambda_{t,X}^*(t) = \sup_{\lambda \leq 0} (\lambda t - \Lambda_X(\lambda)) \geq \lambda_0 t - \log \mathbb{E}(e^{\lambda_0 X})$  für alle  $\lambda_0 \leq 0$ . Für  $K > 0$  beliebig folgt daraus

$$\begin{aligned} \lim_{t \downarrow \text{essinf}(X)} \Lambda_{t,X}^*(t) &\geq \lim_{t \downarrow \text{essinf}(X)} \lambda_0 t - \log \mathbb{E}(e^{\lambda_0 X}) \\ &= \lambda_0 \text{essinf}(X) - \log \mathbb{E}(e^{\lambda_0 X}) > K \end{aligned}$$

für  $\lambda_0$  hinreichend klein, denn nach Voraussetzung ist  $\mathbb{P}(X = 0) = 0$ . Also existiert ein  $\varepsilon > 0$  und eine Menge  $A$  mit  $\mathbb{P}(A) > 0$  und  $X(\omega) > \text{essinf}(X) + \varepsilon$  für alle  $\omega \in A$ . Setze  $d \stackrel{\text{def}}{=} \text{essinf}(X)$ . Unter Beachtung von  $\mathbb{P}(X > 0) = 1$  gilt für  $\lambda < 0$  beliebig die Abschätzung

$$\begin{aligned} \lim_{\lambda \rightarrow -\infty} (\lambda d - \log \mathbb{E}(e^{\lambda X})) &= \lim_{\lambda \rightarrow -\infty} \left( \lambda d - \log \left( \int_A e^{\lambda X} d\mathbb{P} + \int_{A^c} e^{\lambda X} d\mathbb{P} \right) \right) \\ &= \lim_{\lambda \rightarrow -\infty} \lim_{\theta \rightarrow -\infty} \left( \lambda d - \log \left( \int_A e^{\lambda X} d\mathbb{P} + \int_{A^c} e^{\theta X} d\mathbb{P} \right) \right) \\ &= \lim_{\lambda \rightarrow -\infty} \left( \lambda d - \log \left( \int_A e^{\lambda X} d\mathbb{P} + \lim_{\theta \rightarrow -\infty} \int_{A^c} e^{\theta X} d\mathbb{P} \right) \right) \\ &= \lim_{\lambda \rightarrow -\infty} \left( \lambda d - \log \int_A e^{\lambda X} d\mathbb{P} \right) \\ &\geq \lim_{\lambda \rightarrow -\infty} \left( \lambda d - \log \int_A e^{\lambda(d+\varepsilon)} d\mathbb{P} \right) \end{aligned}$$

### 1.3 Eigenschaften der Cramér-Funktion

$$\begin{aligned}
 &= \lim_{\lambda \rightarrow -\infty} (\lambda d - \log(\mathbb{P}(A)e^{\lambda(d+\varepsilon)})) \\
 &= \lim_{\lambda \rightarrow -\infty} (\lambda d - \log \mathbb{P}(A) - \lambda(d + \varepsilon)) \\
 &= \lim_{\lambda \rightarrow -\infty} (-\log \mathbb{P}(A) - \lambda\varepsilon) \\
 &= \infty,
 \end{aligned}$$

wobei wir die Stetigkeit der Logarithmusfunktion und der Abbildung  $\lambda \mapsto \int_A e^{\lambda X} d\mathbb{P}$  benutzt haben. Daraus folgt die Behauptung des Lemmas. □

An dieser Stelle haben wir alle für die folgenden Kapitel benötigten Eigenschaften der Cramér-Funktion einer reellwertigen Zufallsgröße  $X$  zusammengetragen und somit ihre Charakterisierung abgeschlossen. Die folgende Abbildung veranschaulicht die Gestalt einer typischen Cramér-Funktion (Die zugehörige Zufallsgröße  $X$  wird hier als nicht fast sicher konstant und fast sicher beschränkt angenommen, so dass  $-\infty < \text{essinf}(X) < \text{esssup}(X) < \infty$  gilt).

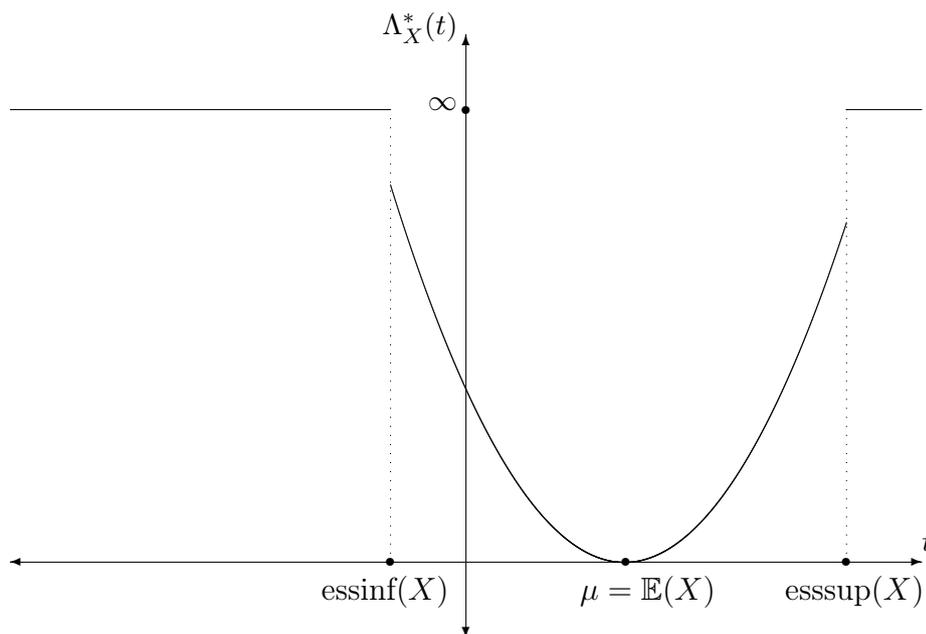


Abbildung 1: Geometrische Interpretation einer typischen Cramér-Funktion

## 1.4 Der Satz von Cramér

Dieser Abschnitt ist dem Satz von Cramér, einem fundamentalen Ergebnis der Theorie der Großen Abweichungen, gewidmet. Der Satz von Cramér löst das in Abschnitt 1.1 vorgestellte Grundproblem der exponentiellen Abfallrate. Wie in Abschnitt 1.1 sei für unabhängige und identisch verteilte reellwertige Zufallsgrößen  $X_1, X_2, \dots$  die Partialsummenfolge  $(S_n)_{n \in \mathbb{N}}$  mit  $S_n = \sum_{i=1}^n X_i$  gegeben. Mit Hilfe der Markov-Ungleichung (vgl. Satz 17.4 auf Seite 72 in [Als1]) können wir zeigen:

**Lemma 1.22** *Sei  $(X_i)_{i \in \mathbb{N}}$  eine Familie von stochastisch unabhängigen und identisch verteilten reellwertigen Zufallsgrößen mit Partialsummenfolge  $(S_n)_{n \in \mathbb{N}}$  und kumulanten-erzeugenden Funktion  $\Lambda_{X_1} = \log M_{X_1}$ . Ferner sei  $t > 0$  beliebig. Dann gilt für jedes  $\lambda \geq 0$  die Abschätzung*

$$\mathbb{P}(S_n \geq nt) \leq e^{-n(\lambda t - \Lambda_{X_1}(\lambda))}.$$

BEWEIS: Seien  $t > 0$  und  $\lambda \geq 0$  beliebig. Eine Anwendung der Markov-Ungleichung liefert

$$\mathbb{P}(S_n \geq nt) = \mathbb{P}(e^{\lambda S_n} \geq e^{\lambda nt}) \leq e^{-\lambda nt} \mathbb{E}(e^{\lambda S_n}) = e^{-\lambda nt} M_{X_1}(\lambda)^n = e^{-n(\lambda t - \Lambda_{X_1}(\lambda))}.$$

Dabei haben wir für die vorletzte Gleichheit benutzt, dass die  $(X_i)_{i \in \mathbb{N}}$  stochastisch unabhängig und identisch verteilt sind. □

Da diese Abschätzung für alle  $\lambda \geq 0$  und alle  $n \in \mathbb{N}$  gilt, folgern wir

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(S_n \geq nt) \leq -\sup_{\lambda \geq 0} (\lambda t - \Lambda_{X_1}(\lambda)) = \Lambda_{r, X_1}^*(t), \quad t > 0.$$

Unter der Zusatzvoraussetzung  $\Lambda(\lambda_0) < \infty$  für ein  $\lambda_0 > 0$  gilt somit nach Satz 1.13 für  $t \geq \mu$

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(S_n \geq nt) \leq -\Lambda_{X_1}^*(t). \quad (1.2)$$

Kommen wir nun zum angekündigten Satz von Cramér. Dieser sagt aus, dass obige Ungleichung auf der exponentiellen Skala scharf, das heißt, die rechte Seite von (1.2) auch eine Abschätzung nach unten ist.

**Satz 1.23 (Satz von Cramér)** *Es sei  $(X_i)_{i \in \mathbb{N}}$  eine Familie von stochastisch unabhängigen und identisch verteilten reellwertigen Zufallsgrößen mit Erwartungswert  $\mu \in \mathbb{R}$  und Varianz  $\sigma^2 \geq 0$ , deren kumulantenenerzeugende Funktion  $\Lambda_{X_1}(\lambda)$  an einer Stelle  $\lambda_0 > 0$  endlich ist. Setze  $S_n = \sum_{i=1}^n X_i$ . Dann gilt für jedes  $t \geq \mu$*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(S_n \geq nt) = -\Lambda_{X_1}^*(t),$$

wobei  $\Lambda_{X_1}^*$  die Cramér-Funktion von  $X_1$  ist.

BEWEIS: Es reicht,  $t = 0$  zu behandeln, denn andernfalls gehen wir über zu den Zufallsvariablen  $X_i - t$  (beachte:  $t = 0$  impliziert  $\mu \leq 0$ ). Also müssen wir zeigen, dass  $\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(S_n \geq 0) = -\Lambda_{X_1}^*(0) = \inf_{\lambda \in \mathbb{R}} \Lambda_{X_1}(\lambda) = \log \rho$  gilt, wobei wir

$$\rho \stackrel{\text{def}}{=} \inf_{\lambda \in \mathbb{R}} M_{X_1}(\lambda)$$

definieren. Für die letzte Gleichheit haben wir die Stetigkeit und die Monotonie der Logarithmusfunktion benutzt.

Wir verwenden im Folgenden die Konvention  $\log 0 \stackrel{\text{def}}{=} -\infty$ . Aus der strikten Konvexität von  $\Lambda_{X_1}$ , der strengen Monotonie der Logarithmusfunktion und der Ungleichung zwischen dem geometrischen und dem arithmetischen Mittel folgt die strikte Konvexität von  $M_{X_1}$  (vgl. Beweis von Lemma 1.6) und es gilt

$$M'_{X_1}(0) = \mathbb{E}(X_1 e^{0X_1}) = \mathbb{E}(X_1) = \mu \leq 0.$$

Ferner dürfen wir voraussetzen, dass die Verteilung der Zufallsgröße  $X_1$  Masse sowohl im Intervall  $(-\infty; 0)$  als auch im Intervall  $(0; \infty)$  besitzt. Dies sieht man folgendermaßen ein:

Wenn  $\mathbb{P}(X_1 < 0) = 1$  gilt, so ist  $M_{X_1}$  streng fallend mit  $\lim_{\lambda \rightarrow \infty} M_{X_1}(\lambda) = \rho = 0$ , und es folgt die Behauptung, denn es ist  $\mathbb{P}(S_n \geq 0) = 0$ . Falls  $\mathbb{P}(X_1 \leq 0) = 1$  und  $\mathbb{P}(X_1 = 0) > 0$ , so ist  $M_{X_1}$  fallend mit  $\lim_{\lambda \rightarrow \infty} M_{X_1}(\lambda) = \rho = \mathbb{P}(X_1 = 0) > 0$ . Aufgrund der stochastischen Unabhängigkeit der  $(X_i)_{i \in \mathbb{N}}$  gilt

$$\mathbb{P}(S_n \geq 0) = \mathbb{P}(X_1 = X_2 = \dots = X_n = 0) = \rho^n,$$

woraus wiederum die Behauptung folgt. Wegen  $\mathbb{E}(X_1) \leq 0$  kann  $X_1$  nicht in  $[0; \infty)$  konzentriert sein, sofern  $X_1$  nicht konstant gleich 0 ist. Wie man leicht einsieht, folgt die Behauptung auch in dem Trivialfall, dass  $X_1$  konstant gleich 0 ist. Also können wir jetzt davon ausgehen, dass  $\mathbb{P}(X_1 > 0) > 0$  und  $\mathbb{P}(X_1 < 0) > 0$  gelten. Insbesondere ist  $0 \in (\text{essinf}(X_1); \text{esssup}(X_1))$  und es gilt  $\lim_{\lambda \rightarrow \infty} M_{X_1}(\lambda) = \infty = \lim_{\lambda \rightarrow -\infty} M_{X_1}(\lambda)$ .

Wegen der strikten Konvexität nimmt  $M_{X_1}$  sein Minimum  $\rho$  in genau einem Punkt

## 1 Grundlagen der Theorie der Großen Abweichungen

$\lambda_0 \in \mathbb{R}$  an, und es gilt  $M'_{X_1}(\lambda_0) = 0$ . Die obere Schranke wird genau wie in Lemma 1.22 hergeleitet, was wir hier nicht wiederholen wollen. Also muss nur noch

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(S_n \geq 0) \geq \log \rho$$

gezeigt werden.

Der nun folgende Beweis der unteren Schranke benutzt einen exponentiellen Maßwechsel und ist fundamental in der Theorie der Großen Abweichungen. Wir betrachten ein Maß  $\hat{\mathbb{P}}$  mit der Eigenschaft

$$\hat{\mathbb{P}}(X_i \in dx) = \frac{1}{\rho} e^{\lambda_0 x} \mathbb{P}(X_i \in dx). \quad (1.3)$$

Die Verteilung von  $X_i$  wird also exponentiell mit Parameter  $\lambda_0$  transformiert. Die rechte Seite von (1.3) ist tatsächlich eine Wahrscheinlichkeitsverteilung, denn es gilt

$$\rho = M_{X_1}(\lambda_0) = \mathbb{E}(e^{\lambda_0 X_i}).$$

Für dieses  $\hat{\mathbb{P}}$  gilt wegen der stochastischen Unabhängigkeit der  $(X_i)_{i \in \mathbb{N}}$

$$\hat{\mathbb{P}}(X_1 \in dx_1, \dots, X_n \in dx_n) = \rho^{-n} e^{\lambda_0(x_1 + \dots + x_n)} \mathbb{P}(X_1 \in dx_1) \dots \mathbb{P}(X_n \in dx_n)$$

und mit Hilfe des Transformationsatzes

$$\hat{\mathbb{P}}(S_n \in ds) = \rho^{-n} e^{\lambda_0 s} \mathbb{P}(S_n \in ds). \quad (1.4)$$

Wir haben

$$\lim_{n \rightarrow \infty} \hat{\mathbb{E}}(e^{-\lambda_0 S_n} \mathbf{1}_{\{S_n \geq 0\}})^{\frac{1}{n}} \geq 1$$

nach dem zentralen Grenzwertsatz für  $S_n$ , denn wir können wie folgt abschätzen:

$$\hat{\mathbb{E}}(e^{-\lambda_0 S_n} \mathbf{1}_{\{S_n \geq 0\}})^{\frac{1}{n}} \geq \hat{\mathbb{E}}(e^{-\lambda_0 S_n} \mathbf{1}_{\{S_n \in (0, \sqrt{n}\sigma)\}})^{\frac{1}{n}} \geq \begin{cases} (e^{-\lambda_0 \sqrt{n}\sigma} \hat{\mathbb{P}}(\frac{S_n}{\sqrt{n}\sigma} \in [0; 1]))^{\frac{1}{n}} & \text{für } \lambda_0 > 0 \\ \hat{\mathbb{P}}(\frac{S_n}{\sqrt{n}\sigma} \in [0; 1])^{\frac{1}{n}} & \text{für } \lambda_0 \leq 0 \end{cases}$$

Dabei gilt  $\lim_{n \rightarrow \infty} (e^{-\lambda_0 \sqrt{n}\sigma})^{\frac{1}{n}} = \lim_{n \rightarrow \infty} e^{\frac{-\lambda_0 \sigma}{\sqrt{n}}} = 1$  und  $\lim_{n \rightarrow \infty} \hat{\mathbb{P}}(\frac{S_n}{\sqrt{n}\sigma} \in [0; 1])^{\frac{1}{n}} = 1$ , denn  $\hat{\mathbb{P}}(\frac{S_n}{\sqrt{n}\sigma} \in [0; 1])$  konvergiert nach dem zentralen Grenzwertsatz für  $n \rightarrow \infty$  gegen eine positive Zahl (weil  $\frac{S_n}{\sqrt{n}\sigma}$  für  $n \rightarrow \infty$  in Verteilung gegen eine  $N(0, 1)$ -Verteilung konvergiert).

Unter Verwendung von (1.4) erhalten wir schließlich

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(S_n \geq 0) \geq \log \rho.$$

Damit ist die untere Schranke bewiesen und der Beweis des Satzes von Cramér abgeschlossen. □

**Bemerkung 1.24** (i) Die Aussage des Satzes von Cramér gilt auch für  $\mathbb{P}(S_n \leq nt)$  an Stelle von  $\mathbb{P}(S_n \geq nt)$  für  $t \leq \mathbb{E}(X_1)$ , sofern ein  $\lambda_0 < 0$  existiert mit  $\Lambda_{X_1}(\lambda_0) < \infty$ . Dies folgt leicht aus einem Übergang von  $X_i$  zu  $-X_i$ .

(ii) Die Verteilung  $\hat{\mathbb{P}}^{X_i}$  im Beweis von Satz 1.24 nennt man die *Cramér-Transformierte* der Verteilung  $\mathbb{P}^{X_i}$ . Diese Transformation wandelt das „untypische“ Verhalten, das Verhalten der Großen Abweichung  $S_n \geq 0$ , in ein „typisches“ um, auf das der zentrale Grenzwertsatz angewendet werden kann.

(iii) Für eine Folge von stochastisch unabhängigen und identisch verteilten reellwertigen Zufallsgrößen  $(X_i)_{i \in \mathbb{N}}$  mit  $\Lambda_{X_1}(\lambda_0) < \infty$  sowie  $\Lambda_{X_1}(\lambda_1) < \infty$  für gewisse  $\lambda_0 < 0 < \lambda_1$  (was genau der Situation in den folgenden Kapiteln entspricht), lässt sich die Aussage des Satzes von Cramér umformulieren zu

$$\mathbb{P}(S_n \geq nt) = e^{-n\Lambda_{X_1}^*(t)+o(n)} = e^{-n\Lambda_{r,X_1}^*(t)+o(n)} \quad \text{für } t \geq \mathbb{E}(X_1),$$

und

$$\mathbb{P}(S_n \leq nt) = e^{-n\Lambda_{X_1}^*(t)+o(n)} = e^{-n\Lambda_{l,X_1}^*(t)+o(n)} \quad \text{für } t \leq \mathbb{E}(X_1).$$



## 2 Der zweifach gewichtete Verzweigungsprozess

In diesem Kapitel stellen wir den *zweifach gewichteten Verzweigungsprozess* vor, welcher der Untersuchungsgegenstand der folgenden Kapitel sein wird. Im Gegensatz zu der zugrunde liegenden Arbeit [BuD] verwenden wir hier eine leicht abgewandelte Notation, welche an die in [Als3] angelehnt ist.

### 2.1 Modellbeschreibung

Wir betrachten einen  $b$ -adischen, unendlich verzweigten stochastischen Baum  $T_\infty$ . Dabei bedeutet „ $b$ -adisch“, dass jeder Knoten (jedes Individuum) des Baumes genau  $b$  Nachkommen besitzt, wobei  $b$  eine natürliche Zahl  $\geq 2$  ist. Somit handelt es sich bei  $T_\infty$  streng genommen gar nicht um einen *stochastischen* Baum, da der Zufall hier überhaupt keine Rolle spielt. Der Baum stellt gewissermaßen das Rohmaterial für den zweifach gewichteten Verzweigungsprozess dar.

Die Wurzel des Baumes bezeichnen wir mit  $\emptyset$  und für  $n \in \mathbb{N}$  einen Knoten der  $n$ -ten Generation mit  $v = (v_1, \dots, v_n)$ . Dabei nummerieren wir die Kinder des Knotens  $v$  von links nach rechts und bezeichnen sie mit  $(v, 1), \dots, (v, b)$ , wobei  $(v, i)$  für  $i \in \{1, \dots, b\}$  eine verkürzte Schreibweise von  $(v_1, \dots, v_n, i)$  darstellt. Außerdem bezeichnen wir für einen Knoten  $v = (v_1, \dots, v_n)$  dessen Urahnen der  $i$ -ten Generation mit  $v|i \stackrel{\text{def}}{=} (v_1, \dots, v_i)$  (Insbesondere gilt  $v|0 = \emptyset$ ). Im Fall  $b = 3$  ergibt sich das folgende Bild:

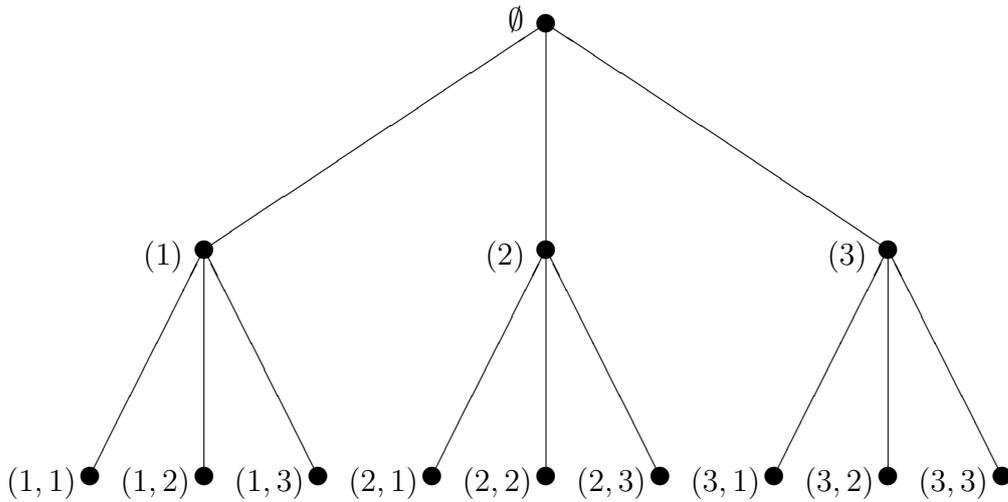


Abbildung 2: Ausschnitt des Baumes  $T_\infty$  bis zur zweiten Generation für  $b = 3$

Wir bezeichnen mit  $\mathcal{K}$  die Menge der Knoten von  $T_\infty$ , und für einen einzelnen Knoten  $v = (v_1, \dots, v_n) \in \mathcal{K}$  sei  $|v| = n$ . Das heißt,  $|v|$  bezeichnet die Generation, der  $v$  angehört. Speziell bedeutet dies  $|\emptyset| = 0$ . Für einen beliebigen Knoten  $v$  mit  $|v| \geq 1$  bezeichne ferner  $\pi(v)$  die Menge der Kanten des Astes von der Wurzel  $\emptyset$  bis zum Knoten  $v$ .

Jeder Knoten  $v$  von  $T_\infty$  werde zudem mit einem  $b$ -dimensionalen Zufallsvektor

$$(Z(v), E(v)) \stackrel{\text{def}}{=} ((Z_1(v), E_1(v)), (Z_2(v), E_2(v)), \dots, (Z_b(v), E_b(v)))$$

bestehend aus  $b$  Paaren von reellwertigen Zufallsgrößen versehen. Dabei seien die Zufallsvariablen  $(Z(v), E(v))_{v \in \mathcal{K}}$  stochastisch unabhängig, und jedes Paar  $(Z_i(v), E_i(v))$ ,  $i \in \{1, \dots, b\}$ ,  $v \in \mathcal{K}$  sei verteilt wie  $(Z, E)$  für nichtnegative und stochastisch unabhängige Zufallsgrößen  $Z$  und  $E$ . Die Familien  $(Z_i(v))_{v \in \mathcal{K}, i \in \{1, \dots, b\}}$  und  $(E_i(v))_{v \in \mathcal{K}, i \in \{1, \dots, b\}}$  seien ebenfalls stochastisch unabhängig. Man beachte aber, dass für ein festes  $v \in \mathcal{K}$  innerhalb des Zufallsvektors  $(Z(v), E(v))$  die Paare  $(Z_i(v), E_i(v))_{i=1, \dots, b}$  nicht stochastisch unabhängig zu sein brauchen.

Die Zufallsgröße  $E$  sei zudem nicht fast sicher konstant, und es gelte  $\mathbb{E}(E) < \infty$  und  $\mathbb{E}(Z) < \infty$ , sowie  $\mathbb{P}(E = 0) = 0$ . Außerdem benötigen wir noch eine weitere, im Hinblick auf die Anwendung der im ersten Kapitel gewonnenen Ergebnisse sehr wichtige Eigenschaft: Die Zufallsgröße  $Z$  erfülle die Bedingung  $\Lambda_Z(\lambda) < \infty$  für ein  $\lambda > 0$ , wobei  $\Lambda_Z$  die

kumulantenerzeugende Funktion von  $Z$  ist.

Die  $b$  Zufallsvariablen-Paare  $(Z_1(v), E_1(v)), \dots, (Z_b(v), E_b(v))$  werden den  $b$  Kanten, die den Knoten  $v$  mit seinen  $b$  Kindern  $(v, 1), \dots, (v, b)$  verbinden, zugewiesen. Im Fall  $b = 3$  ergibt sich das folgende Bild:

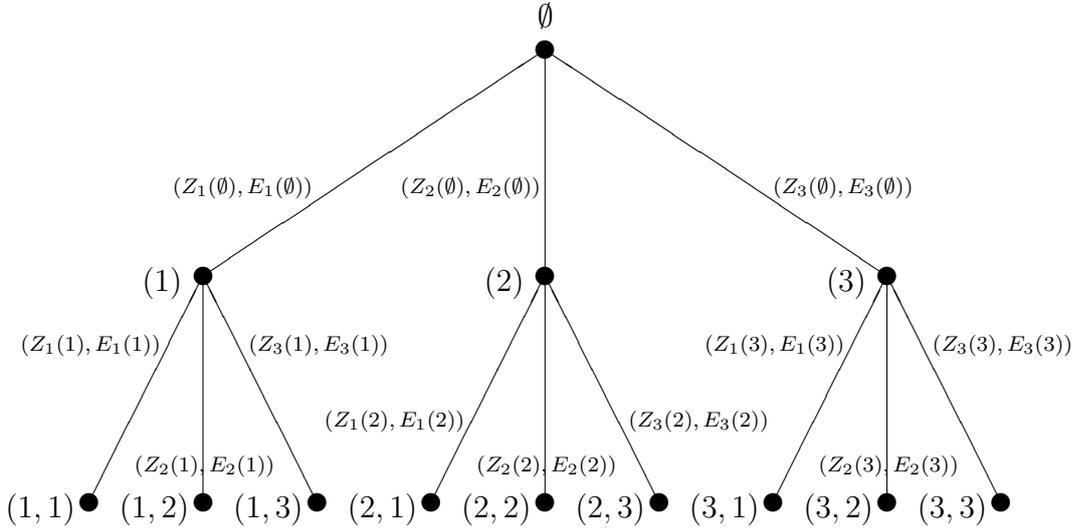


Abbildung 3: Ausschnitt des zweifach gewichteten Baumes  $\mathcal{T}_\infty$  bis zur zweiten Generation für  $b = 3$

Wir bezeichnen ferner mit  $\mathcal{E}$  die Menge der Kanten von  $T_\infty$ . Für eine beliebige Kante  $e \in \mathcal{E}$  bezeichnen wir die ihr zugewiesenen Gewichte auch mit der kürzeren Schreibweise  $(Z_e, E_e)$ . Für einen Knoten  $v = (v_1, \dots, v_n)$  seien

$$G_v \stackrel{\text{def}}{=} \sum_{e \in \pi(v)} E_e = \sum_{i=1}^n E_{v_i}(v|(i-1)) \quad (\text{Alter des Knotens } v)$$

und

$$D_v \stackrel{\text{def}}{=} \sum_{e \in \pi(v)} Z_e = \sum_{i=1}^n Z_{v_i}(v|(i-1)) \quad (\text{Gewichtete Höhe des Knotens } v)$$

die beiden *kumulierten Gewichte des Astes*  $\pi(v)$ . In den meisten Fällen interpretieren wir  $G_v$  als den Geburtszeitpunkt von  $v$  und  $D_v$  als die räumliche Länge des Astes von der Wurzel  $\emptyset$  bis zum Knoten  $v$ .

## 2 Der zweifach gewichtete Verzweigungsprozess

Nun haben wir einen  $b$ -adischen Baum, dessen Knoten  $v$  jeweils mit zwei stochastisch unabhängigen, nicht-negativen Gewichten  $G_v$  und  $D_v$  versehen sind. Einen in der oben beschriebenen Art und Weise gewichteten Baum nennen wir fortan einen ( $b$ -adischen) *zweifach gewichteten Verzweigungsprozess* und verwenden für ihn die Notation  $\mathcal{T}_\infty$ . Schließlich definieren wir für  $n \in \mathbb{N}$  die Zufallsvariable  $\mathcal{T}_n$  als den zufälligen Baum der Knoten  $v$  von  $\mathcal{T}_\infty$ , für die  $G_v \leq n$  gilt. Ferner sei  $\mathcal{K}_n$  die zufällige Menge der Knoten von  $\mathcal{T}_n$ .

Im Folgenden interessieren wir uns für die gewichtete Höhe

$$H_n \stackrel{\text{def}}{=} \max\{D_v \mid v \in \mathcal{K}_n\}$$

des zufälligen Teilbaumes  $\mathcal{T}_n$  von  $\mathcal{T}_\infty$ . Die Zufallsvariable  $H_n$  kann also als die *gewichtete Höhe des Baumes  $\mathcal{T}_\infty$  zum Zeitpunkt  $n$*  interpretiert werden.

In Kapitel 3 geben wir eine vollständige Charakterisierung des asymptotischen Verhaltens von  $H_n$  an, welche gänzlich unabhängig von den Verteilungen der Zufallsgrößen  $Z$  und  $E$  ist. Für den Beweis dieses ersten Hauptresultates bemühen wir die in Kapitel 1 erarbeiteten grundlegenden Eigenschaften der Cramér-Funktion und den Satz von Cramér.

# 3 Asymptotik der gewichteten Höhe des zweifach gewichteten Verzweigungsprozesses

## 3.1 Vorbereitungen

Damit wir unser erstes Hauptresultat beweisen können, müssen wir zunächst noch weitere technische Hilfsmittel bereitstellen. Im Folgenden sei  $\mathcal{T}_\infty$  ein  $b$ -adischer zweifach gewichteter Verzweigungsprozess mit Gewichtsfunktionen  $Z$  und  $E$  und allen Eigenschaften aus Kapitel 2. Um das asymptotische Verhalten der gewichteten Höhe  $H_n$  des Baumes  $\mathcal{T}_\infty$  zum Zeitpunkt  $n$  untersuchen zu können, definieren wir zunächst die Menge

$$\mathcal{C}_{Z,E} \stackrel{\text{def}}{=} \{(\rho, \alpha) \mid \Lambda_Z^*(\alpha) + \Lambda_E^*(\rho) = \log b, \alpha \geq \mathbb{E}(Z), \rho \leq \mathbb{E}(E)\}.$$

Dabei stehen  $\Lambda_Z^*$  und  $\Lambda_E^*$  für die Cramér-Funktionen der Gewichte  $Z$  und  $E$ . Wir untersuchen zunächst die geometrische Gestalt der Menge  $\mathcal{C}_{Z,E}$ . Mit Hilfe der Ergebnisse aus Kapitel 1 lassen sich einige Charakteristika herleiten, welche von der Wahl der Zufallsgrößen  $Z$  und  $E$  völlig unabhängig sind.

Als erstes halten wir fest, dass aufgrund der Nichtnegativität der Zufallsgröße  $E$  (siehe Kapitel 2)  $\Lambda_E(\lambda) < \infty$  für jedes  $\lambda < 0$  und damit  $\Lambda_E^*(\rho) = \Lambda_{l,E}^*(\rho)$  für jedes  $\rho \leq \mathbb{E}(E)$  gilt (vgl. Satz 1.13). Außerdem existiert nach Voraussetzung ein  $\lambda > 0$  mit  $\Lambda_Z(\lambda) < \infty$ , woraus auch  $\Lambda_Z^*(\alpha) = \Lambda_{r,Z}^*(\alpha)$  für jedes  $\alpha \geq \mathbb{E}(Z)$  folgt. Es gilt also

$$\mathcal{C}_{Z,E} = \{(\rho, \alpha) \mid \Lambda_{r,Z}^*(\alpha) + \Lambda_{l,E}^*(\rho) = \log b, \alpha \geq \mathbb{E}(Z), \rho \leq \mathbb{E}(E)\}.$$

Aus diesem Grund können wir in den folgenden Ausführungen die Notationen  $\Lambda_E^*$  und  $\Lambda_{l,E}^*$  beziehungsweise  $\Lambda_Z^*$  und  $\Lambda_{r,Z}^*$  gleichermaßen verwenden. Wie sich herausstellen wird, beschreibt die Menge  $\mathcal{C}_{Z,E}$  eine strikt konkave Kurve in dem Fall, dass  $Z$  nicht fast sicher konstant ist. Den trivialen Fall, dass  $Z$  fast sicher konstant ist, behandelt das folgende Lemma.

**Lemma 3.1** *Sei  $Z$  fast sicher konstant. Dann besteht die Menge  $\mathcal{C}_{Z,E}$  aus genau einem Punkt.*

BEWEIS: Sei  $Z$  fast sicher konstant. Nach Satz 1.13 gilt  $\Lambda_Z^*(\mathbb{E}(Z)) = 0 = \Lambda_E^*(\mathbb{E}(E))$ , und nach Satz 1.19 ist  $\Lambda_E^*$  stetig auf  $(\text{essinf}(E); \text{esssup}(E))$ . Außerdem gilt

$$\lim_{\rho \downarrow \text{essinf}(E)} \Lambda_{l,E}^*(\rho) = \infty$$

nach Lemma 1.21. Deshalb existiert ein  $\rho^* \leq \mathbb{E}(E)$  mit  $\Lambda_{l,E}^*(\rho^*) = \log b > 0$  (denn  $b \geq 2$ ) und damit  $(\rho^*, \mathbb{E}(Z)) \in \mathcal{C}_{Z,E}$ .  $\rho^*$  ist dabei eindeutig bestimmt, denn nach den Sätzen 1.13 und 1.18 ist  $\Lambda_E^*$  streng monoton fallend auf dem Intervall  $(\text{essinf}(E); \mathbb{E}(E))$ . □

**Lemma 3.2** *Ist  $Z$  nicht fast sicher konstant, so ist die Menge  $\mathcal{C}_{Z,E}$  der Graph einer stetigen und streng monoton wachsenden Kurve  $\rho \mapsto \alpha(\rho)$ .*

BEWEIS: Sei  $Z$  nicht fast sicher konstant. Dann gilt  $\text{essinf}(Z) < \mathbb{E}(Z) < \text{esssup}(Z)$ . Nach Lemma 3.1 ist  $(\rho^*, \mathbb{E}(Z)) \in \mathcal{C}_{Z,E}$  und  $\lim_{\rho \downarrow \text{essinf}(E)} \Lambda_E^*(\rho) = \infty$ . Wegen der Geometrie der Cramér-Funktionen von  $Z$  und  $E$  auf  $(\text{essinf}(Z); \text{esssup}(Z))$  beziehungsweise  $(\text{essinf}(E); \text{esssup}(E))$  besteht  $\mathcal{C}_{Z,E}$  aus mehr als nur einem Punkt (siehe Abbildung 1 im ersten Kapitel). Seien  $(\rho_1, \alpha_1), (\rho_2, \alpha_2) \in \mathcal{C}_{Z,E}$  mit  $\rho_1 < \rho_2$  beliebig. Dann gilt aufgrund der strengen Monotonie der Cramér-Funktionen  $\Lambda_E^*$  und  $\Lambda_Z^*$

$$\Lambda_E^*(\rho_1) > \Lambda_E^*(\rho_2)$$

und daher

$$\Lambda_Z^*(\alpha_1) < \Lambda_Z^*(\alpha_2).$$

Daraus ergibt sich  $\alpha_1 < \alpha_2$ , denn wegen der strengen Monotonie der Cramér-Funktionen ist für ein  $\rho_1$  mit  $(\rho_1, \alpha_1) \in \mathcal{C}_{Z,E}$  das zugehörige  $\alpha_1$  eindeutig bestimmt.

Die Stetigkeit der Kurve  $\mathcal{C}_{Z,E}$  folgt aus der Stetigkeit der Cramér-Funktionen. □

Wir sprechen im Folgenden anstatt von der Menge  $\mathcal{C}_{Z,E}$  auch von der Kurve  $\mathcal{C}_{Z,E}$  und kommen nun zu ihrer hinsichtlich unserer späteren Ergebnisse wichtigsten Eigenschaft, der strikten Konkavität:

**Lemma 3.3 (Strikte Konkavität der Kurve  $\mathcal{C}_{Z,E}$ )** *Ist  $Z$  nicht fast sicher konstant, so ist die Kurve  $\mathcal{C}_{Z,E}$  strikt konkav.*

BEWEIS: Wie in Lemma 3.1 sei  $\rho^* \geq 0$  so gewählt, dass  $(\rho^*, \mathbb{E}(Z)) \in \mathcal{C}_{Z,E}$ . Für  $(\rho, \alpha) \in \mathcal{C}_{Z,E}$  beliebig gilt dann  $\rho \in [\rho^*; \mathbb{E}(E)] \subseteq (\text{essinf}(E); \text{esssup}(E))$  und nach Satz 1.18 ist  $\Lambda_E^*$  strikt konvex auf  $(\text{essinf}(E); \text{esssup}(E))$ . Seien nun  $\alpha_1, \alpha_2 \geq \mathbb{E}(Z)$  und  $\rho_1, \rho_2 \leq \mathbb{E}(E)$  mit  $(\rho_1, \alpha_1), (\rho_2, \alpha_2) \in \mathcal{C}_{Z,E}$  und  $x \in (0; 1)$  beliebig.

Setze  $\alpha \stackrel{\text{def}}{=} x\alpha_1 + (1-x)\alpha_2$ . Nach Lemma 3.2 gibt es dann genau ein  $\rho \in [\rho^*; \mathbb{E}(E)]$  mit  $(\rho, \alpha) \in \mathcal{C}_{Z,E}$ .

Für die strikte Konkavität von  $\mathcal{C}_{Z,E}$  genügt es zu zeigen, dass  $\rho < x\rho_1 + (1-x)\rho_2$  gilt. Mit Hilfe der strikten Konkavität von  $\Lambda_E^*$  und  $\Lambda_Z^*$  erhalten wir

$$\begin{aligned} \Lambda_E^*(x\rho_1 + (1-x)\rho_2) &< x\Lambda_E^*(\rho_1) + (1-x)\Lambda_E^*(\rho_2) \\ &= \log b - (x\Lambda_Z^*(\alpha_1) + (1-x)\Lambda_Z^*(\alpha_2)) \\ &< \log b - \Lambda_Z^*(x\alpha_1 + (1-x)\alpha_2) \\ &= \Lambda_E^*(\rho). \end{aligned}$$

Da  $\Lambda_E^*$  monoton fallend ist, folgt  $\rho < x\rho_1 + (1-x)\rho_2$ . □

In den folgenden Ausführungen sei

$$c \stackrel{\text{def}}{=} \max_{(\rho, \alpha) \in \mathcal{C}_{Z,E}} \frac{\alpha}{\rho}.$$

Ist  $Z$  fast sicher konstant, so ist trivialerweise  $c = \frac{\mathbb{E}(Z)}{\rho^*}$  nach Lemma 3.1. Sei daher im Folgenden  $Z$  nicht fast sicher konstant. Falls ein  $\alpha^* \geq \mathbb{E}(Z)$  existiert mit  $\Lambda_Z^*(\alpha^*) = \log b$ , so liegt der Punkt  $(\mathbb{E}(E), \alpha^*)$  in  $\mathcal{C}_{Z,E}$  und bildet gleichzeitig den rechten Endpunkt der Kurve  $\mathcal{C}_{Z,E}$ . In diesem Fall verläuft  $\mathcal{C}_{Z,E}$  von  $(\rho^*, \mathbb{E}(Z))$  bis  $(\mathbb{E}(E), \alpha^*)$ , wobei beide Endpunkte zu  $\mathcal{C}_{Z,E}$  gehören. Die Kurve  $\mathcal{C}_{Z,E}$  ist dann kompakt. Aus der Kompaktheit folgt wiederum die Existenz eines Punktes  $(\rho_0, \alpha_0) \in \mathcal{C}_{Z,E}$  mit  $c = \frac{\alpha_0}{\rho_0}$ . Das Maximum wird aber auch in dem Fall angenommen, dass kein  $\alpha \geq \mathbb{E}(Z)$  existiert mit  $\Lambda_Z^*(\alpha) = \log b$ . Aufgrund der strengen Monotonie und der strikten Konkavität von  $\Lambda_Z^*$  auf dem Intervall  $(\mathbb{E}(Z); \text{esssup}(Z))$  ist dies allerdings nur möglich, wenn  $\text{esssup}(Z) < \infty$  gilt, denn sonst würde  $\lim_{\alpha \uparrow \text{esssup}(Z)} \Lambda_Z^*(\alpha) = \lim_{\alpha \rightarrow \infty} \Lambda_Z^*(\alpha) = \infty$  gelten. Ist  $\text{esssup}(Z) < \infty$ , so gilt  $\lim_{\alpha \uparrow \text{esssup}(Z)} \Lambda_Z^*(\alpha) = \Lambda_Z^*(\text{esssup}(Z)) < \log b$  und damit gehört auch in diesem Fall der rechte Endpunkt  $(\rho_1, \text{esssup}(Z))$  (mit einem geeigneten  $\rho_1$ ) zu  $\mathcal{C}_{Z,E}$ . Damit ist die Kurve wiederum kompakt und das Maximum wird angenommen.

Von großem Interesse ist an dieser Stelle die Frage, ob der Punkt  $(\rho_0, \alpha_0)$  mit der Eigenschaft  $\frac{\alpha_0}{\rho_0} = c$  auch eindeutig bestimmt ist, und ferner, an welcher Stelle er in  $\mathcal{C}_{Z,E}$

liegt. Die erste Frage ist mit „ja“ zu beantworten und auch auf die zweite gibt es eine Antwort, wie das folgende Lemma zeigt:

**Lemma 3.4 (Eindeutigkeit von  $(\rho_0, \alpha_0)$  und Gestalt von  $\mathcal{C}_{Z,E}$ )** Die Funktion  $f : \mathcal{C}_{Z,E} \rightarrow \mathbb{R}_{\geq 0}$ ,  $f(\rho, \alpha) \stackrel{\text{def}}{=} \frac{\alpha}{\rho}$  nimmt ihr Maximum in einem eindeutig bestimmten Punkt  $(\rho_0, \alpha_0)$  an. Existiert ferner ein  $\alpha^* > \mathbb{E}(Z)$  mit  $\Lambda_Z^*(\alpha^*) = \log b$ , so liegt  $(\rho_0, \alpha_0)$  im Inneren von  $\mathcal{C}_{Z,E}$ .

BEWEIS: Zur Existenz eines Punktes  $(\rho_0, \alpha_0)$  mit  $c = \frac{\alpha_0}{\rho_0}$  ist nach der Vorbemerkung nichts mehr zu sagen und die Eindeutigkeit folgt aus der strikten Konkavität von  $\mathcal{C}_{Z,E}$ . Gelte  $\Lambda_Z^*(\alpha^*) = \log b$  für ein  $\alpha^* > \mathbb{E}(Z)$ . Dann besitzt  $\mathcal{C}_{Z,E}$  die beiden Endpunkte  $(\rho^*, \mathbb{E}(Z))$  und  $(\mathbb{E}(E), \alpha^*)$ . Wegen  $\Lambda_E^*(\mathbb{E}(E)) = 0$  gilt  $\rho^* < \mathbb{E}(E)$  und daher  $\frac{d}{dt}\Lambda_E^*(t)|_{t=\rho^*} < 0$  wegen der strikten Konvexität von  $\Lambda_E^*$ . Daraus folgt zusammen mit  $\frac{d}{dt}\Lambda_Z^*(t)|_{t=\mathbb{E}(Z)} = 0$  (vgl. Folgerung 1.20), dass  $\mathcal{C}_{Z,E}$  im Punkt  $(\rho^*, \mathbb{E}(Z))$  die Steigung  $\infty$  besitzt.

Analog zeigt man, dass  $\mathcal{C}_{Z,E}$  im Punkt  $(\mathbb{E}(E), \alpha^*)$  die Steigung 0 besitzt. Daraus folgt die Behauptung mit Hilfe der strikten Konkavität. □

Ein wichtiges Beispiel für eine Zufallsgröße  $Z$  mit  $\Lambda_Z^*(\alpha) < \log b$  für alle  $\alpha \in [\mathbb{E}(Z); \text{esssup}(Z)]$  ist eine Bernoulli-verteilte Zufallsgröße. Diese Möglichkeit wird in der Arbeit [BuD] von den Autoren außer Acht gelassen, obwohl sich einige ihrer Ausführungen genau auf diesen Fall beziehen. In Kapitel 6 behandeln wir unter anderem den *zufälligen rekursiven Baum*, dessen Gewichtsfunktion  $Z$   $B(1, \frac{1}{2})$ -verteilt ist.

**Beispiel 3.5** Für eine  $B(1, \frac{1}{2})$ -verteilte Zufallsgröße  $Z$  existiert kein  $\alpha \in [\mathbb{E}(Z); \text{esssup}(Z)] = [\frac{1}{2}; 1]$  mit  $\Lambda_Z^*(\alpha) \geq \log 3$ .

BEWEIS: Sei  $\alpha \in [\frac{1}{2}; 1]$ . Dann gilt für jedes  $\lambda \geq 0$

$$\begin{aligned} \lambda\alpha - \Lambda_Z(\lambda) &= \lambda\alpha - \log \mathbb{E}(e^{\lambda Z}) \\ &= \log e^{\lambda\alpha} - \log \frac{1}{2}(1 + e^\lambda) \\ &= \log \frac{2e^{\lambda\alpha}}{1 + e^\lambda} \\ &\leq \log 2 \\ &< \log 3. \end{aligned}$$

□

In den meisten Fällen tritt jedoch die Situation aus Lemma 3.4 auf. Dann hat  $\mathcal{C}_{Z,E}$  die folgende Gestalt:

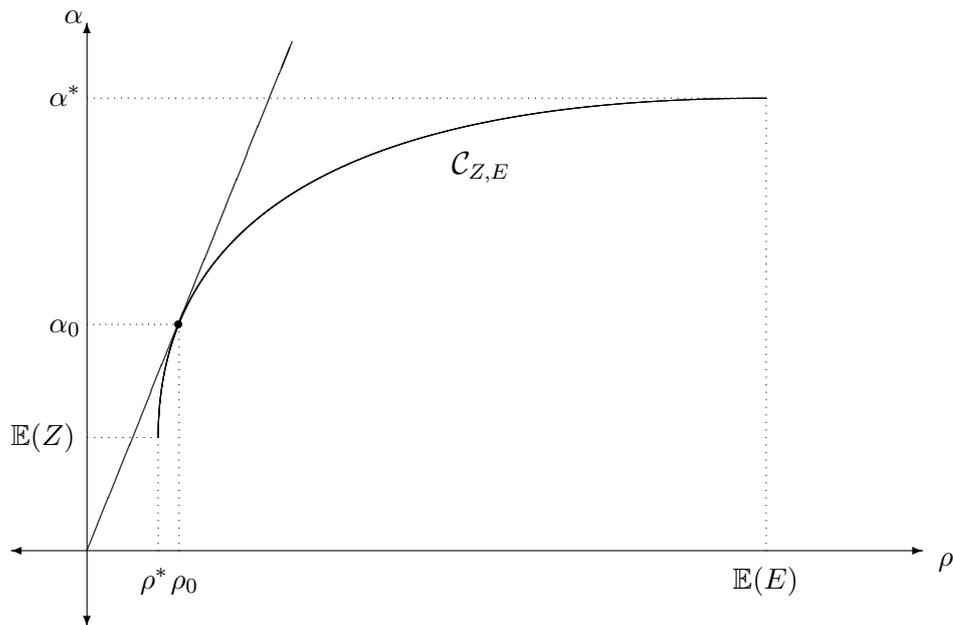


Abbildung 4: Geometrische Interpretation der Kurve  $\mathcal{C}_{Z,E}$  zusammen mit der Konstante  $c = \frac{\alpha_0}{\rho_0}$ . Beachte, dass  $\mathbb{E}(Z) < \alpha < \alpha^*$  und  $\rho^* < \rho < \mathbb{E}(E)$ . Die Steigung der Tangente an die Kurve  $\mathcal{C}_{Z,E}$  ist  $c$ .

## 3.2 Das asymptotische Verhalten von $H_n$

Die Konstante  $c$  aus dem vorherigen Abschnitt stellt sich im Folgenden als die einzig richtige Wahl für den Grenzwert nach Wahrscheinlichkeit der Folge  $(\frac{1}{n}H_n)_{n \in \mathbb{N}}$  heraus. Unabhängig von der Wahl der Gewichtsfunktionen  $Z$  und  $E$  gilt:

**Satz 3.6 (Das asymptotische Verhalten der gewichteten Höhe  $H_n$ )** Sei  $\mathcal{T}_\infty$  ein  $b$ -adischer zweifach gewichteter Verzweigungsprozess mit Gewichtsfunktionen  $Z$  und  $E$ . Dann gilt

$$\frac{1}{n}H_n \xrightarrow[n \rightarrow \infty]{} c$$

nach Wahrscheinlichkeit. Dabei ist  $c = \frac{\alpha_0}{\rho_0}$  das Maximum von  $(\rho, \alpha) \mapsto \frac{\alpha}{\rho}$  in  $\mathcal{C}_{Z,E}$  mit eindeutig bestimmten  $\rho_0$  und  $\alpha_0$ .

Wir spalten den Beweis von Satz 3.6 in zwei Lemmata auf, welche eine obere und eine untere Schranke für den Wahrscheinlichkeitslimes der Folge  $(\frac{1}{n}H_n)_{n \in \mathbb{N}}$  bereitstellen.

**Lemma 3.7 (Obere Schranke)** Mit den Notationen aus Satz 3.6 und  $\varepsilon > 0$  beliebig gilt

$$\lim_{n \rightarrow \infty} \mathbb{P}(H_n \geq c(1 + \varepsilon)n) = 0.$$

BEWEIS: Für  $n, k \in \mathbb{N}$  sei  $L_{n,k}$  die Menge der Knoten  $v$  von  $\mathcal{T}_n$  mit  $|v| = k$ . Zur Erinnerung: In Kapitel 2 haben wir mit  $\mathcal{T}_n$  den Baum der Knoten  $v$  von  $\mathcal{T}_\infty$  mit  $G_v \leq n$  bezeichnet. Dann gilt für  $\alpha > \mathbb{E}(Z)$  beliebig

$$\begin{aligned} \mathbb{P}(\{\exists v \in L_{n,k} : D_v \geq \alpha k\}) &= \mathbb{P}\left(\bigcup_{v \in L_{n,k}} \{D_v \geq \alpha k\}\right) \\ &\leq b^k \mathbb{P}\left(\sum_{i=1}^k Z_i \geq \alpha k, \sum_{i=1}^k E_i \leq n\right) \\ &= b^k \mathbb{P}\left(\sum_{i=1}^k Z_i \geq \alpha k\right) \mathbb{P}\left(\sum_{i=1}^k E_i \leq n\right), \end{aligned}$$

wobei wir bei der vorletzten Abschätzung benutzt haben, dass  $\mathbb{P}$   $\sigma$ -subadditiv und  $\mathcal{T}_\infty$   $b$ -adisch ist, sowie  $D_v = \sum_{i=1}^k Z_{v_i}(v|(i-1))$  und  $G_v = \sum_{i=1}^k E_{v_i}(v|(i-1))$  gilt. Dabei bezeichnen  $Z_i, E_i$  für  $i = 1, \dots, k$  stochastisch unabhängige Kopien von  $Z$  und  $E$ . Nach Definition des zweifach gewichteten Verzweigungsprozesses sind die Familien  $(Z_i(v))_{v \in \mathcal{K}, i \in \{1, \dots, b\}}$  und  $(E_i(v))_{v \in \mathcal{K}, i \in \{1, \dots, b\}}$  stochastisch unabhängig. Das berechtigt uns

zu der Annahme der stochastischen Unabhängigkeit von  $\sum_{i=1}^k Z_i$  und  $\sum_{i=1}^k E_i$ . Wir setzen  $\rho \stackrel{\text{def}}{=} \frac{n}{k}$  und schätzen obige beiden Wahrscheinlichkeiten weiter ab. Beachte, dass  $\Lambda_{r,Z}^*(\alpha) = 0$  für  $\alpha < \mathbb{E}(Z)$  und  $\Lambda_{l,E}^*(\rho) = 0$  für  $\rho > \mathbb{E}(E)$  gilt (vgl. Satz 1.13 (v)). Mittels

$$\mathbb{P}\left(\sum_{i=1}^k E_i \leq n\right) = \mathbb{P}\left(\lambda \sum_{i=1}^k E_i \geq \lambda \rho k\right), \quad \forall \lambda < 0,$$

und des Lemmas 1.22 für die beiden Wahrscheinlichkeiten erhalten wir

$$\begin{aligned} b^k \mathbb{P}\left(\sum_{i=1}^k Z_i \geq \alpha k\right) \mathbb{P}\left(\sum_{i=1}^k E_i \leq \rho k\right) &\leq b^k e^{-k\Lambda_{r,Z}^*(\alpha)} e^{-k\Lambda_{l,E}^*(\rho)} \\ &= e^{k(\log b - \Lambda_{r,Z}^*(\alpha) - \Lambda_{l,E}^*(\rho))} \\ &= e^{-k\gamma(\alpha, \rho)}. \end{aligned}$$

Dabei ist

$$\gamma(\alpha, \rho) \stackrel{\text{def}}{=} \Lambda_{r,Z}^*(\alpha) + \Lambda_{l,E}^*(\rho) - \log b.$$

Für  $\varepsilon > 0$  wählen wir nun  $\alpha = c(1 + \varepsilon)\rho$ . Dann gilt

$$\gamma(\alpha, \rho) = \gamma(c(1 + \varepsilon)\rho, \rho) \geq \beta \stackrel{\text{def}}{=} \inf_{\rho \geq 0} \gamma(c(1 + \varepsilon)\rho, \rho) > \inf_{\rho \geq 0} \gamma(c\rho, \rho) = 0. \quad (3.1)$$

Zur Begründung setzen wir  $f(\rho) \stackrel{\text{def}}{=} \gamma(c\rho, \rho)$  und  $f_\varepsilon(\rho) \stackrel{\text{def}}{=} \gamma(c(1 + \varepsilon)\rho, \rho)$ , sowie  $g(\rho) \stackrel{\text{def}}{=} c\rho$  und  $g_\varepsilon(\rho) \stackrel{\text{def}}{=} c(1 + \varepsilon)\rho$ . Nach den Ergebnissen des vorherigen Abschnitts behandeln wir den häufigeren Fall, dass ein  $\alpha \geq \mathbb{E}(Z)$  existiert mit  $\Lambda_Z^*(\alpha) = \log b$ , wie folgt: Wir haben  $\inf_{\rho \geq 0} f(\rho) = f(\rho_0) = 0$  mit dem  $\rho_0$  aus Abschnitt 3.1. Denn innerhalb von  $(\rho^*; \mathbb{E}(E)) \times (\mathbb{E}(Z); \alpha^*)$  gilt  $\Lambda_{r,Z}^*(\alpha) + \Lambda_{l,E}^*(\rho) = \Lambda_Z^*(\alpha) + \Lambda_E^*(\rho)$  und außerhalb von  $(\rho^*; \mathbb{E}(E)) \times (\mathbb{E}(Z); \alpha^*)$  nimmt  $\Lambda_{r,Z}^*(\alpha) + \Lambda_{l,E}^*(\rho)$  höchstens größere Werte an, da  $\Lambda_{l,E}^*$  monoton fallend auf  $(\text{essinf}(E); \rho^*)$  und  $\Lambda_{r,Z}^*(\alpha) = 0$  für  $\alpha < \mathbb{E}(Z)$  beziehungsweise  $\Lambda_{r,Z}^*$  monoton wachsend auf  $(\alpha^*; \text{esssup}(Z))$  und  $\Lambda_{l,E}^*(\rho) = 0$  für  $\rho > \mathbb{E}(E)$  gilt.

Sei jetzt  $I = [a; b]$  ein kompaktes Intervall mit  $\rho^* < a < \rho_0 < b < \mathbb{E}(E)$  so klein, dass  $\mathbb{E}(Z) < c\rho < c(1 + \varepsilon)\rho < \alpha^*$  für alle  $\rho \in I$  gilt. Für  $\rho \in I$  gilt also  $\Lambda_{l,E}^*(\rho) = \Lambda_E^*(\rho)$ ,  $\Lambda_{r,Z}^*(c(1 + \varepsilon)\rho) = \Lambda_Z^*(c(1 + \varepsilon)\rho)$  sowie  $\Lambda_{r,Z}^*(c\rho) = \Lambda_Z^*(c\rho)$ . Aufgrund der Stetigkeit der Cramér-Funktionen von  $Z$  und  $E$  können wir wiederum  $\varepsilon$  so klein wählen, dass  $f_\varepsilon$  sein Infimum ebenfalls im Inneren von  $I$  annimmt. Das bedeutet, es existiert ein  $\rho' \in I$  mit

$$\inf_{\rho \geq 0} f_\varepsilon(\rho) = \inf_{\rho \in I} f_\varepsilon(\rho) = f_\varepsilon(\rho') > f(\rho') \geq \inf_{\rho \in I} f(\rho) = \inf_{\rho \geq 0} f(\rho) = f(\rho_0) = 0,$$

wobei wir benutzt haben, dass  $\Lambda_Z^*$  streng monoton wachsend ist. Zur Veranschaulichung der Situation betrachten wir das folgende Bild:

### 3 Asymptotik der gewichteten Höhe des zweifach gewichteten Verzweigungsprozesses

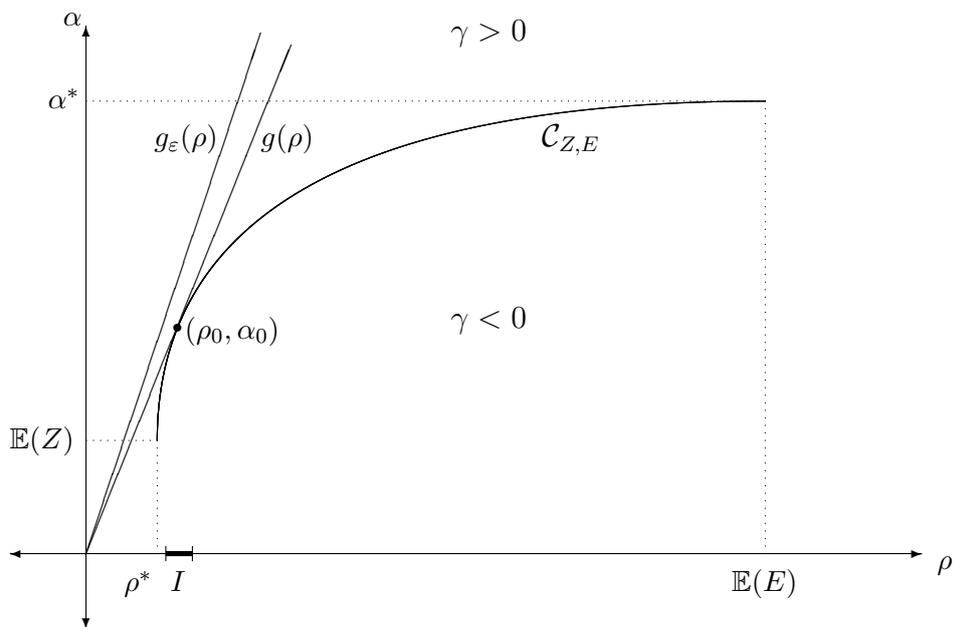


Abbildung 5: Veranschaulichung der Situation mit den Geraden  $g_\varepsilon(\rho)$  und  $g(\rho)$ , sowie dem Intervall  $I$

Der zweite Fall, dass kein  $\alpha \geq \mathbb{E}(Z)$  existiert mit  $\Lambda_Z^*(\alpha) = \log b$ , ist etwas komplizierter. Liegt der Punkt  $(\rho_0, \alpha_0)$  im Inneren von  $\mathcal{C}_{Z,E}$ , so geht der Beweis völlig analog. Ist aber  $c = \frac{\text{esssup}(Z)}{\rho_1}$  mit den Notationen aus Abschnitt 3.1, so gilt  $\inf_{\rho \geq 0} f(\rho) = f(\rho_1)$ . Anders als in obigem Fall lässt sich hier nicht einfach ein Intervall  $I$  konstruieren mit  $\rho^* < a < \rho_0 < b < \mathbb{E}(E)$  und  $\mathbb{E}(Z) < c(1 + \varepsilon)\rho < \alpha^*$  für alle  $\rho \in I$ . Wir haben die folgende Situation:

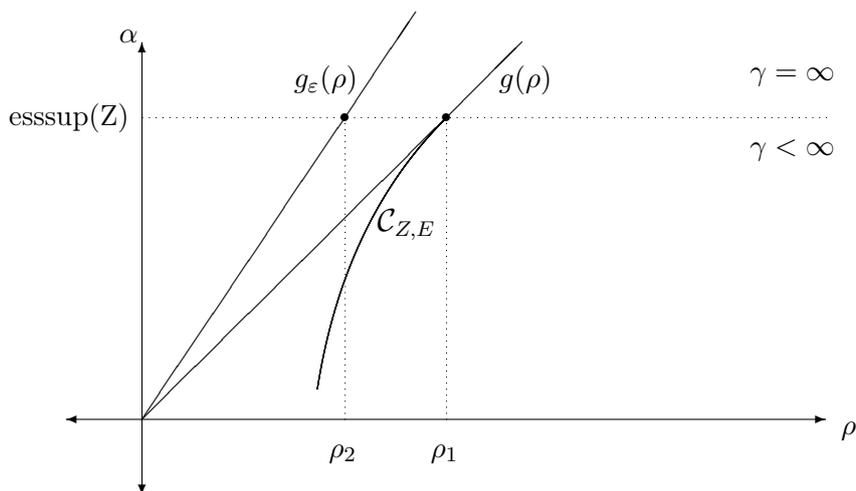


Abbildung 6: Veranschaulichung der Situation im zweiten Fall

In diesem Fall ist es angebracht, die Geraden  $g$  und  $g_\varepsilon$  als Funktionen von  $\alpha$  aufzufassen. Wir haben  $g(\alpha) = \frac{\alpha}{c}$  und  $g_\varepsilon(\alpha) = \frac{\alpha}{c(1+\varepsilon)}$ . Mit diesen Notationen gilt

$$f(\rho_1) = \inf_{\rho \geq 0} f(\rho) = \lim_{\alpha \uparrow \text{esssup}(Z)} \gamma\left(\alpha, \frac{\alpha}{c}\right).$$

Sei nun  $a < \text{esssup}(Z)$  beliebig nahe an  $\text{esssup}(Z)$ . Aufgrund der Stetigkeit der Cramér-Funktion können wir wieder  $\varepsilon$  so klein wählen, dass auch  $\alpha \mapsto \gamma\left(\alpha, \frac{\alpha}{c(1+\varepsilon)}\right)$  sein Infimum innerhalb von  $(a; \text{esssup}(Z)]$  annimmt. Für jedes  $\delta > 0$  mit  $\text{esssup}(Z) - \delta > a$  gilt dann

$$\gamma\left(\text{esssup}(Z) - \delta, \frac{\text{esssup}(Z) - \delta}{c(1+\varepsilon)}\right) > \gamma\left(\text{esssup}(Z) - \delta, \frac{\text{esssup}(Z) - \delta}{c}\right)$$

wegen der strengen Monotonie von  $\Lambda_E^*$ . Da aber  $\Lambda_E^*$  auch strikt konvex ist, gilt sogar

$$\lim_{\alpha \uparrow \text{esssup}(Z)} \gamma\left(\alpha, \frac{\alpha}{c(1+\varepsilon)}\right) > \lim_{\alpha \uparrow \text{esssup}(Z)} \gamma\left(\alpha, \frac{\alpha}{c}\right)$$

(denn der Abstand wächst für  $\delta \downarrow 0$ ), und damit folgt die Behauptung (3.1).

Für jedes  $K \in \mathbb{N}$  erhalten wir nun

$$\begin{aligned} \mathbb{P}(H_n \geq c(1+\varepsilon)n) &= \mathbb{P}(\exists k \in \mathbb{N} \exists v \in L_{n,k} : D_v \geq c(1+\varepsilon)n) \\ &\leq \sum_{k=0}^{K-1} \mathbb{P}(\exists v \in L_{n,k} : D_v \geq c(1+\varepsilon)n) \\ &\quad + \sum_{k=K}^{\infty} \mathbb{P}(\exists v \in L_{n,k} : D_v \geq c(1+\varepsilon)n). \end{aligned}$$

Für den zweiten Term folgt mittels (3.1)

$$\sum_{k=K}^{\infty} \mathbb{P}(\exists v \in L_{n,k} : D_v \geq c(1+\varepsilon)n) \leq \sum_{k=K}^{\infty} e^{-k\beta} = \mathcal{O}(e^{-K\beta}).$$

Sei  $L_{\infty,k}$  die Menge der Knoten  $v$  von  $\mathcal{T}_\infty$  mit  $|v| = k$ . Dann liefert eine grobe Abschätzung mit Hilfe der Markov-Ungleichung (vgl. [Als1], Satz 17.4(a) auf Seite 72)

$$\begin{aligned} \sum_{k=0}^{K-1} \mathbb{P}(\exists v \in L_{n,k} : D_v \geq c(1+\varepsilon)n) &\leq \sum_{k=0}^{K-1} \mathbb{P}(\exists v \in L_{\infty,k} : D_v \geq c(1+\varepsilon)n) \\ &\leq Kb^K \sup_{0 \leq k < K, v \in L_{\infty,k}} \mathbb{P}(D_v \geq c(1+\varepsilon)n) \\ &\leq Kb^K \sup_{0 \leq k < K, v \in L_{\infty,k}} \frac{\mathbb{E}(D_v)}{c(1+\varepsilon)n} \\ &\leq Kb^K \frac{K \mathbb{E}(Z)}{c(1+\varepsilon)n} \xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

Für die letzte Ungleichheit haben wir  $D_v = \sum_{i=1}^n Z_{v_i}(v|(i-1))$  benutzt. Fasst man diese beiden Abschätzungen zusammen, so erhält man

$$\mathbb{P}(H_n \geq c(1 + \varepsilon)n) \leq \mathcal{O}(e^{-K\beta}) + o(1).$$

Da  $K$  beliebig groß gewählt werden kann und  $\beta > 0$  ist, folgt die Behauptung des Lemmas. □

**Lemma 3.8 (Untere Schranke)** *Mit den Notationen aus Satz 3.6 und  $\varepsilon > 0$  beliebig gilt*

$$\lim_{n \rightarrow \infty} \mathbb{P}(H_n \geq c(1 - 2\varepsilon)n) = 1.$$

BEWEIS: Sei  $\varepsilon > 0$  beliebig. Im Folgenden heben wir zunächst einen Ast von der Wurzel  $\emptyset$  bis zu einem Knoten  $v$  hervor, für den  $D_v \geq c(1 - 2\varepsilon) \log n$  gilt. Um dies zu erreichen, konstruieren wir innerhalb von  $\mathcal{T}_\infty$  einen bestimmten Galton-Watson-Prozess mit Aussterbewahrscheinlichkeit  $q < 1$ .

Die Konstruktion startet in der Wurzel  $\emptyset$  von  $\mathcal{T}_\infty$ , welche auch die Wurzel unseres Galton-Watson-Prozesses ist. Sei  $L_{\infty, K}$  die Menge der Knoten  $v$  von  $\mathcal{T}_\infty$  mit  $|v| = K \in \mathbb{N}$ . Einen Knoten  $v \in L_{\infty, K}$  bezeichnen wir als *guten Knoten* genau dann, wenn  $D_v \geq \alpha K$  und  $G_v \leq \rho K$  für gewisse  $\alpha \geq \mathbb{E}(Z)$  und  $\rho \leq \mathbb{E}(E)$  gilt, welche wir später genau festlegen werden.

Die guten Knoten  $v \in L_{\infty, K}$  seien nun die Kinder von  $\emptyset$  in unserem Galton-Watson-Prozess. Diese Kinder reproduzieren sich wiederum unabhängig voneinander gemäß derselben Reproduktionsverteilung wie  $\emptyset$ . Das bedeutet, ein Knoten  $v$  genau  $K$  Generationen unterhalb eines Knotens  $u$  aus unserem Galton-Watson-Prozess ist genau dann ein Kind von  $u$  (also ein guter Knoten), wenn die Bedingungen  $D_v - D_u \geq \alpha K$  und  $G_v - G_u \leq \rho K$  gelten.

Nach Kapitel 2, 14.7 in [Als2] überlebt der *Galton-Watson-Prozess der guten Knoten* mit positiver Wahrscheinlichkeit, wenn die erwartete Anzahl der Kinder eines Individuums (also der Erwartungswert der Reproduktionsverteilung) echt größer als 1 ist. Wir schreiben  $N_K$  für die *zufällige Anzahl der guten Knoten in der  $K$ -ten Generation von  $\mathcal{T}_\infty$* . Also ist  $N_K$  die Anzahl der Individuen der ersten Generation unseres Galton-Watson-Prozesses der guten Knoten. Dann gilt analog zum Beweis von Lemma 3.7 mit Hilfe des Satzes von Cramér (vgl. Bemerkung 1.24 und beachte, dass hier im Gegensatz zum

Beweis von Lemma 3.7  $\alpha \geq \mathbb{E}(Z)$  und  $\rho \leq \mathbb{E}(E)$  gilt)

$$\begin{aligned} \mathbb{E}(N_K) &= \mathbb{E}\left(\sum_{v:|v|=K} \mathbf{1}_{\{D_v \geq \alpha K, G_v \leq \rho K\}}\right) \\ &= b^K \mathbb{P}\left(\sum_{i=1}^K Z_i \geq \alpha K, \sum_{i=1}^K E_i \leq \rho K\right) \\ &= e^{-\gamma(\alpha, \rho)K + o(K)}. \end{aligned}$$

Für die zweite Gleichheit haben wir benutzt, dass die Zufallsgrößen  $(D_v)_{v:|v|=K}$  beziehungsweise  $(G_v)_{v:|v|=K}$  identisch verteilt sind.

Für  $K$  hinreichend groß gilt also  $\mathbb{E}(N_K) > 1$ , sofern  $\gamma(\alpha, \rho) < 0$  gilt. Setze nun  $\alpha = \alpha_0$  und  $\rho = \frac{\rho_0}{\sqrt{1-\epsilon}}$  und beachte, dass für  $\epsilon$  hinreichend klein  $\rho \leq \mathbb{E}(E)$  gilt. Aufgrund der strengen Monotonie von  $\Lambda_E^*$  folgt  $\gamma(\alpha, \rho) < 0$  und damit  $\mathbb{E}(N_K) > 1$  für ein hinreichend großes  $K$ . Für diese  $\alpha, \rho$  und  $K$  sei nun  $q < 1$  die Aussterbewahrscheinlichkeit unseres Galton-Watson-Prozesses der guten Knoten. Der Prozess überlebt also mit Wahrscheinlichkeit  $1 - q$ .

Im nächsten Schritt vergrößern wir diese Wahrscheinlichkeit, so dass sie beliebig nahe an 1 liegt. Um dies zu erreichen, starten wir den Galton-Watson-Prozess mit  $b^{tK}$  Individuen in der  $tK$ -ten Generation von  $\mathcal{T}_\infty$  ( $t \in \mathbb{N}$ ). Damit vergrößern wir die Chance, dass wenigstens einer der dort startenden unabhängigen Prozesse der guten Knoten überlebt. Wir betrachten dafür die Verteilung des Gewichtes

$$E(v) = (E_1(v), E_2(v), \dots, E_b(v))$$

eines Knotens  $v$ . Für  $\beta > 0$  beliebig klein können wir ein  $a > 0$  finden mit

$$\mathbb{P}(E_1(v) \leq a, E_2(v) \leq a, \dots, E_b(v) \leq a) \geq 1 - \beta.$$

Setze nun

$$A \stackrel{\text{def}}{=} \{E_e \leq a \text{ für alle Kanten } e \text{ oberhalb der } tK\text{-ten Generation von } \mathcal{T}_\infty\}.$$

Dann liefert eine sehr grobe Abschätzung

$$\mathbb{P}(A^c) \leq \beta b^{tK+K} = \beta b^{(t+1)K}.$$

Damit ereignet sich  $A$ , kontrolliert durch die Variable  $\beta$ , mit Wahrscheinlichkeit beliebig nahe an 1. Außerdem gilt  $G_v \leq atK$  für alle Knoten  $v$  mit  $|v| = tK$ , falls  $A$  eintritt.

### 3 Asymptotik der gewichteten Höhe des zweifach gewichteten Verzweigungsprozesses

Sei jetzt  $B$  das Ereignis, dass einer der  $b^{tK}$  in der  $tK$ -ten Generation von  $\mathcal{T}_\infty$  startenden Galton-Watson-Prozesse der guten Knoten überlebt. Dann gilt wegen der stochastischen Unabhängigkeit der Prozesse

$$\mathbb{P}(B^c) = q^{b^{tK}}.$$

Wenn  $A \cap B$  eintritt, dann existiert für jedes  $k \in \mathbb{N}$  ein Knoten  $v$  in  $\mathcal{T}_\infty$  mit  $|v| = tK + kK$ , so dass  $G_v \leq atK + \rho kK$  und  $D_v = D_v - D_u + D_u \geq \alpha kK + 0 = \alpha kK$  (Dabei ist  $u$  der Urahne von  $v$  in der  $tK$ -ten Generation). Setze

$$k \stackrel{\text{def}}{=} \left\lceil \frac{n(1-\varepsilon)}{\rho_0 K} \right\rceil \quad ([ \ ] \text{ Gaußklammer}).$$

Mit  $\alpha = \alpha_0$  und  $\rho = \frac{\rho_0}{\sqrt{1-\varepsilon}}$  gilt dann für obigen Knoten  $v$

$$G_v \leq atK + \rho kK \leq atK + n\sqrt{1-\varepsilon} < n$$

und

$$D_v \geq \alpha kK \geq \left( \frac{n(1-\varepsilon)}{\rho_0 K} - 1 \right) \alpha_0 K = c(1-\varepsilon)n - \alpha_0 K \geq c(1-2\varepsilon)n$$

für  $n$  hinreichend groß. Es folgt

$$\mathbb{P}(H_n \geq c(1-2\varepsilon)n) \geq \mathbb{P}(A \cap B) \geq 1 - \mathbb{P}(A^c) - \mathbb{P}(B^c).$$

Diese untere Schranke können wir mit Hilfe der Konstanten  $\beta$  und  $t$  so kontrollieren, dass sie beliebig nahe an 1 liegt, indem wir zunächst  $t$  hinreichend groß und danach  $\beta$  in Abhängigkeit von  $t$  hinreichend klein wählen. Daraus folgt die Behauptung des Lemmas.  $\square$

Mit der oberen Schranke aus Lemma 3.7 und der unteren Schranke aus Lemma 3.8 ist schließlich auch Satz 3.6 bewiesen.

# 4 Verallgemeinerung der Situation auf den Fall einer beliebigen Reproduktionsverteilung

## 4.1 Verallgemeinerung des ersten Hauptresultates

In den bisherigen Abschnitten haben wir als Rohmaterial für unseren zweifach gewichteten Verzweigungsprozess einen  $b$ -adischen Baum verwendet. Das ist im Hinblick auf die Ergebnisse der dieser Arbeit zugrundeliegenden Ausarbeitung [BuD] vollkommen ausreichend. Trotzdem kann man sich die interessante Frage stellen, ob die Resultate aus Kapitel 3 auf den Fall übertragbar sind, in dem man  $b \in \mathbb{N}_{\geq 2}$  durch eine beliebige Reproduktionsverteilung  $(p_n)_{n \in \mathbb{N}_0}$  mit Erwartungswert  $\mu > 1$  ersetzt. Der Fall  $\mu \leq 1$  führt zum fast sicheren Aussterben der Population, weshalb er im Hinblick auf das asymptotische Verhalten der gewichteten Höhe uninteressant ist. Sei also  $B \sim (p_n)_{n \in \mathbb{N}_0}$  eine beliebige Reproduktionsverteilung mit  $\mu \stackrel{\text{def}}{=} \mathbb{E}(B) > 1$ .

Wir konstruieren jetzt den *verallgemeinerten zweifach gewichteten Verzweigungsprozess*  $\mathcal{T}_\infty^*$ : Die Notationen  $\mathcal{K}$ ,  $\mathcal{K}_n$ ,  $\mathcal{E}$  und  $\pi(v)$  aus Kapitel 2 bleiben erhalten. Im Gegensatz zu den Ausführungen in Kapitel 2 ist der Rohbaum jetzt aber ein Galton-Watson-Prozess mit Reproduktionsverteilung  $B$ . Diesen bezeichnen wir mit  $T_\infty^*$ . Dessen zufällige Kanten  $e \in \mathcal{E}$  versehen wir wiederum mit zwei stochastisch unabhängigen Zufallsgrößen  $Z_e$  und  $E_e$  mit derselben Verteilung wie  $Z$  beziehungsweise  $E$ . Dabei besitzen die Gewichtsfunktionen  $Z$  und  $E$  alle Eigenschaften aus Kapitel 2. Den Knoten  $v$  ordnen wir unabhängig die Gewichte

$$(Z(v), E(v)) \stackrel{\text{def}}{=} ((Z_1(v), E_1(v)), (Z_2(v), E_2(v)), \dots, (Z_{B(v)}(v), E_{B(v)}(v)))$$

zu, nur dass jetzt  $B(v)$  eine zufällige Zahl aus  $\mathbb{N}_0$  ist. Wie in Kapitel 2 seien die Familien  $(Z_i(v))_{v \in \mathcal{K}, i \geq 0}$  und  $(E_i(v))_{v \in \mathcal{K}, i \geq 0}$  stochastisch unabhängig. Für jedes  $v \in \mathcal{K}$  sei  $B(v)$  eine Kopie der Reproduktionsverteilung  $B$  (Zur Erinnerung:  $\mathcal{K}$  bezeichnet die (hier

zufällige) Menge der Knoten von  $\mathcal{T}_\infty^*$ ). Außerdem seien die Zufallsgrößen  $(B(v))_{v \in \mathcal{K}}$  stochastisch unabhängig, und die Familie  $(B(v))_{v \in \mathcal{K}}$  sei stochastisch unabhängig von den Familien  $(Z_i(v))_{v \in \mathcal{K}, i \geq 0}$  und  $(E_i(v))_{v \in \mathcal{K}, i \geq 0}$ . Die Reproduktion von Individuen vollzieht sich also *unabhängig* von den beiden Kantengewichten.

Für einen Knoten  $v$  werden die kumulierten Gewichte des Astes  $\pi(v)$  wieder mit  $D_v$  und  $G_v$  bezeichnet. Außerdem sei  $\mathcal{T}_n^*$  der zufällige Baum der Knoten  $v$  von  $\mathcal{T}_\infty^*$ , für die  $G_v \leq n$  gilt. Wir können uns also den verallgemeinerten zweifach gewichteten Verzweigungsprozess  $\mathcal{T}_\infty^*$  als einen gewöhnlichen Galton-Watson-Prozess vorstellen, dessen Kanten wir zusätzlich mit zwei unabhängigen Gewichten versehen haben.

Wie wir an späterer Stelle sehen werden, sind die Ergebnisse von Kapitel 3 nur ein Spezialfall derer aus diesem Kapitel.

Zunächst müssen wir in Bezug auf die Notationen von Kapitel 3 einige Anpassungen vornehmen. Im Folgenden sei

$$\begin{aligned} \mathcal{C}_{Z,E}^* &\stackrel{\text{def}}{=} \{(\rho, \alpha) \mid \Lambda_Z^*(\alpha) + \Lambda_E^*(\rho) = \log(\mu), \rho \leq \mathbb{E}(E), \alpha \geq \mathbb{E}(Z)\} \\ &= \{(\rho, \alpha) \mid \Lambda_{r,Z}^*(\alpha) + \Lambda_{l,E}^*(\rho) = \log(\mu), \rho \leq \mathbb{E}(E), \alpha \geq \mathbb{E}(Z)\}. \end{aligned}$$

Dann besitzt die Kurve  $\mathcal{C}_{Z,E}^*$  dieselben Eigenschaften wie  $\mathcal{C}_{Z,E}$  aus Abschnitt 3.1. Da  $\mu > 1$  gilt, übertragen sich sämtliche Beweise der Lemmata 3.1 bis 3.4 auf den verallgemeinerten Fall, indem wir einfach  $b$  durch  $\mu$  ersetzen. Deshalb können wir analog zu Satz 3.6 den folgenden Satz formulieren:

**Satz 4.1 (Das asymptotische Verhalten von  $H_n$  im verallgemeinerten Fall)**

Sei  $\mathcal{T}_\infty^*$  ein verallgemeinerter zweifach gewichteter Verzweigungsprozess mit Reproduktionsverteilung  $B \sim (p_n)_{n \in \mathbb{N}_0}$  und Gewichtsfunktionen  $Z$  und  $E$ . Es gelte  $\mu = \mathbb{E}(B) > 1$ . Dann gilt für die gewichtete Höhe  $H_n = \max\{D_v \mid v \in \mathcal{K}_n\}$  des zufälligen Teilbaumes  $\mathcal{T}_n^*$  von  $\mathcal{T}_\infty^*$

$$\frac{1}{n} H_n \xrightarrow[n \rightarrow \infty]{} c^*$$

nach Wahrscheinlichkeit, wobei  $c^* \stackrel{\text{def}}{=} \frac{\alpha_0}{\rho_0}$  das Maximum von  $f : \mathcal{C}_{Z,E}^* \rightarrow \mathbb{R}_{\geq 0}$ ,  $f(\alpha, \rho) \stackrel{\text{def}}{=} \frac{\alpha}{\rho}$  mit dem eindeutig bestimmten Punkt  $(\rho_0, \alpha_0) \in \mathcal{C}_{Z,E}^*$  ist.

Genau wie bei Satz 3.6 in Kapitel 3 spalten wir den Beweis von Satz 4.1 in zwei Lemmata auf:

**Lemma 4.2 (Obere Schranke)** *Mit den Notationen aus Satz 4.1 und  $\varepsilon > 0$  beliebig gilt*

$$\lim_{n \rightarrow \infty} \mathbb{P}(H_n \geq c^*(1 + \varepsilon)n) = 0.$$

BEWEIS: Analog zu Lemma 3.7 betrachten wir für  $\alpha > \mathbb{E}(Z)$  die Wahrscheinlichkeit, dass in der  $k$ -ten Generation von  $\mathcal{T}_n^*$  ein Knoten  $v$  existiert mit  $D_v \geq \alpha k$ . Seien also  $\alpha > \mathbb{E}(Z)$  und  $n \in \mathbb{N}$  beliebig und bezeichne  $M_k$  die zufällige Anzahl der Individuen in der  $k$ -ten Generation von  $\mathcal{T}_n^*$ . Dann gilt

$$\begin{aligned} \mathbb{P}(\{\exists v \in \mathcal{T}_n^*, |v| = k : D_v \geq \alpha k\}) &= \mathbb{P}\left(\bigcup_{v \in \mathcal{T}_n^*, |v|=k} \{D_v \geq \alpha k\}\right) \\ &= \mathbb{P}\left(\bigcup_{j \geq 0} (\{M_k = j\} \cap \bigcup_{v \in \mathcal{T}_n^*, |v|=k} \{D_v \geq \alpha k\})\right) \\ &= \sum_{j \geq 0} \mathbb{P}\left(\{M_k = j\} \cap \bigcup_{v \in \mathcal{T}_n^*, |v|=k} \{D_v \geq \alpha k\}\right) \\ &= \sum_{j \geq 0} \mathbb{P}\left(\{M_k = j\} \cap \bigcup_{v_1, \dots, v_j, |v_i|=k} \{D_{v_i} \geq \alpha k\}\right) \\ &\leq \sum_{j \geq 0} j \mathbb{P}\left(M_k = j, \sum_{i=1}^k Z_i \geq \alpha k, \sum_{i=1}^k E_i \leq n\right) \\ &= \sum_{j \geq 0} j \mathbb{P}(M_k = j) \mathbb{P}\left(\sum_{i=1}^k Z_i \geq \alpha k\right) \mathbb{P}\left(\sum_{i=1}^k E_i \leq n\right) \\ &= \mu^k \mathbb{P}\left(\sum_{i=1}^k Z_i \geq \alpha k\right) \mathbb{P}\left(\sum_{i=1}^k E_i \leq n\right), \end{aligned}$$

wobei wir die folgenden Tatsachen benutzt haben: Erstens gilt  $\{M_k = j\} \cap \{M_k = i\} = \emptyset$  für  $j \neq i$ . Zweitens sind die Zufallsgrößen  $(D_v)_{v:|v|=k}$  beziehungsweise  $(G_v)_{v:|v|=k}$  identisch verteilt, so dass wir an ihrer Stelle  $\sum_{i=1}^k Z_i$  beziehungsweise  $\sum_{i=1}^k E_i$  mit  $Z_i$  und  $E_i$  wie im Beweis von Lemma 3.7 schreiben können. Drittens sind  $M_k$ ,  $(Z_i)_{i=1, \dots, k}$  und  $(E_i)_{i=1, \dots, k}$  stochastisch unabhängig, da nach Voraussetzung die Kantengewichte keinerlei Einfluss auf die Reproduktion nehmen. Viertens beträgt für einen Galton-Watson-Prozess mit Reproduktionsverteilung  $B \sim (p_n)_{n \in \mathbb{N}_0}$  und  $\mu = \mathbb{E}(B)$  die erwartete Anzahl der Individuen der  $k$ -ten Generation  $\mu^k$  (vgl. dazu [Als2], Abschnitt 14.7 *Der Galton-Watson-Verzweigungsprozess*, Seite 116-120).

Der Rest des Beweises funktioniert analog zu dem Beweis von Lemma 3.7 mit dem Unterschied, dass wir  $b$  durch  $\mu$  ersetzen müssen und in der letzten Abschätzung (mit

der Markov-Ungleichung) zunächst analog zu obiger Überlegung

$$\begin{aligned} \sum_{k=0}^{K-1} \mathbb{P}(\{\exists v \in \mathcal{T}_n^*, |v| = k : D_v \geq c^*(1 + \varepsilon)k\}) &\leq \sum_{k=0}^{K-1} \mu^k \mathbb{P}(D_v \geq c^*(1 + \varepsilon)k) \\ &\leq K \mu^K \sup_{0 \leq k < K, |v|=k} \mathbb{P}(D_v \geq c^*(1 + \varepsilon)k) \end{aligned}$$

erhalten.

□

**Lemma 4.3 (Untere Schranke)** *Mit den Notationen aus Satz 4.1 und  $\varepsilon > 0$  beliebig gilt*

$$\lim_{n \rightarrow \infty} \mathbb{P}(H_n \geq c^*(1 - 2\varepsilon)n) = 1.$$

BEWEIS: Sei  $\varepsilon > 0$  beliebig. Analog zu Lemma 3.8 konstruieren wir den Galton-Watson-Prozess der guten Knoten. Einen Knoten  $v$  der  $K$ -ten Generation von  $\mathcal{T}_\infty^*$  nennen wir dabei wieder einen *guten Knoten* genau dann, wenn  $D_v \geq \alpha K$  und  $G_v \leq \rho K$  für gewisse  $\alpha \geq \mathbb{E}(Z)$  und  $\rho \leq \mathbb{E}(E)$  (welche wir später genau spezifizieren werden) gilt. Wie im Beweis von Lemma 3.8 nennen wir einen Knoten  $v$   $K$  Generationen unterhalb eines Knotens  $u$  aus unserem neu konstruierten Galton-Watson-Prozess einen guten Knoten genau dann, wenn  $D_v - D_u \geq \alpha K$  und  $G_v - G_u \leq \rho K$  gilt.

Sei wieder  $N_K$  die zufällige Anzahl der guten Knoten und  $M_K$  die zufällige Anzahl aller Knoten in der  $K$ -ten Generation von  $\mathcal{T}_\infty^*$  (vgl. Beweis von Lemma 4.2). Dann gilt

$$\begin{aligned} \mathbb{E}(N_K) &= \mathbb{E}\left(\sum_{j \geq 0} \mathbf{1}_{\{j\}}(M_K) N_K\right) \\ &= \sum_{j \geq 0} \mathbb{E}(\mathbf{1}_{\{j\}}(M_K) N_K) \\ &= \sum_{j \geq 0} \mathbb{E}\left(\sum_{v_1, \dots, v_j, |v_i|=K} \mathbf{1}_{\{j\}}(M_K) \mathbf{1}_{[\alpha K; \infty) \times (-\infty; \rho K]}(D_{v_i}, G_{v_i})\right) \\ &= \sum_{j \geq 0} j \mathbb{E}\left(\mathbf{1}_{\{j\}}(M_K) \mathbf{1}_{[\alpha K; \infty) \times (-\infty; \rho K]} \left(\sum_{i=1}^K Z_i, \sum_{i=1}^K E_i\right)\right) \\ &= \sum_{j \geq 0} j \mathbb{P}(M_K = j) \mathbb{P}\left(\sum_{i=1}^k Z_i \geq \alpha K, \sum_{i=1}^k E_i \leq \rho K\right) \\ &= \mu^K \mathbb{P}\left(\sum_{i=1}^k Z_i \geq \alpha K\right) \mathbb{P}\left(\sum_{i=1}^k E_i \leq \rho K\right) \\ &= e^{-\gamma^*(\alpha, \rho)K + o(K)} \end{aligned}$$

mit

$$\gamma^*(\alpha, \rho) \stackrel{\text{def}}{=} \Lambda_{r,Z}^*(\alpha) + \Lambda_{l,E}^*(\rho) - \log \mu = \Lambda_Z^*(\alpha) + \Lambda_E^*(\rho) - \log \mu$$

(für  $\alpha \geq \mathbb{E}(Z)$  und  $\rho \leq \mathbb{E}(E)$ ). Neben den in den entsprechenden Erläuterungen innerhalb der Beweise der Lemmata 3.7 und 3.8 erwähnten Voraussetzungen und Sätzen haben wir zusätzlich noch folgende Tatsachen verwendet:

Die zweite Gleichheit gilt aufgrund des Satzes von der monotonen Konvergenz. Für die vierte Gleichheit haben wir verwendet, dass die Zufallsgrößen  $(D_v)_{v:|v|=K}$  beziehungsweise  $(G_v)_{v:|v|=K}$  identisch verteilt sind, und bei der fünften und der sechsten Gleichheit haben wir die stochastische Unabhängigkeit von  $M_K$ ,  $(Z_i)_{i=1,\dots,K}$  und  $(E_i)_{i=1,\dots,K}$  ausgenutzt. Der Rest folgt analog zu dem Beweis von Lemma 3.8 mit Hilfe des Satzes von Cramér.

Setze nun  $\alpha \stackrel{\text{def}}{=} \alpha_0$  und  $\rho \stackrel{\text{def}}{=} \frac{\rho_0}{\sqrt{1-\varepsilon}}$ . Dann folgt wegen der strengen Monotonie von  $\Lambda_{l,E}^*$ , dass  $\gamma^*(\alpha, \rho) < 0$  gilt. Damit gilt auch  $\mathbb{E}(N_K) > 1$ , sofern  $K$  hinreichend groß gewählt wird. Für diese  $\alpha, \rho$  und  $K$  gilt  $q < 1$  für die Aussterbewahrscheinlichkeit  $q$  des Galton-Watson-Prozesses der guten Knoten.

Im nächsten Schritt vergrößern wir die Überlebenswahrscheinlichkeit  $1 - q$ , indem wir den Prozess analog zum Beweis von Lemma 3.8 in der  $tK$ -ten Generation von  $\mathcal{T}_\infty^*$  ( $t \in \mathbb{N}$ ) starten. Für einen Knoten  $v$  mit  $|v| = tK$  und  $\beta > 0$  beliebig klein wählen wir ein  $a > 0$ , so dass  $\mathbb{P}(G_v \leq atK) \geq 1 - \beta$  gilt. Da die Zufallsgrößen  $(G_v)_{v:|v|=tK}$  identisch verteilt sind, gilt diese Abschätzung für jeden Knoten der  $tK$ -ten Generation. Definiere

$$A \stackrel{\text{def}}{=} \{G_v \leq atK \ \forall v \in \mathcal{K}_n \text{ mit } |v| = tK\}.$$

Dann gilt analog zum Beweis von Lemma 4.2

$$\begin{aligned} \mathbb{P}(A^c) &= \mathbb{P}(\{\exists v \in \mathcal{K}_n, |v| = tK \text{ mit } G_v > atK\}) \\ &\leq \mathbb{P}\left(\bigcup_{v \in \mathcal{K}: |v|=tK} \{G_v > atK\}\right) \\ &= \sum_{n \geq 0} n \mathbb{P}(M_{tK} = n) \mathbb{P}(G_v > atK) \\ &= \mu^{tK} \mathbb{P}(G_v > atK) \\ &\leq \mu^{tK} \beta. \end{aligned}$$

Sei ferner  $B$  das Ereignis, dass einer der in der  $tK$ -ten Generation von  $\mathcal{T}_\infty^*$  startenden

Galton-Watson-Prozesse der guten Knoten überlebt. Dann gilt

$$\begin{aligned}
 \mathbb{P}(B^c) &= \mathbb{P}(\{\text{Alle in der } tK\text{-ten Generation startenden Prozesse sterben aus}\}) \\
 &= \mathbb{P}\left(\bigcup_{n \geq 0} \{M_{tK} = n\} \cap \bigcap_{i=1}^n \{\text{Prozess } \mathcal{P}_i \text{ stirbt aus}\}\right) \\
 &= \sum_{n \geq 0} \mathbb{P}(M_{tK} = n)q^n,
 \end{aligned}$$

wobei wir wie im Beweis von Lemma 3.8 die stochastische Unabhängigkeit der in der  $tK$ -ten Generation von  $\mathcal{T}_\infty^*$  startenden Prozesse der guten Knoten  $\mathcal{P}_1, \mathcal{P}_2, \dots$ , sowie die Unabhängigkeit der Zufallsgrößen  $(B(v))_{v \in \mathcal{K}}$  und die Unabhängigkeit der Familien  $(B(v))_{v \in \mathcal{K}}$ ,  $(Z_i(v))_{v \in \mathcal{K}, i \geq 0}$  und  $(E_i(v))_{v \in \mathcal{K}, i \geq 0}$  ausgenutzt haben. Für  $t$  hinreichend groß gilt dann

$$\sum_{n=0}^{\infty} \mathbb{P}(M_{tK} = n)q^n < \varepsilon.$$

Begründung: Sei  $n_0 \in \mathbb{N}$  so gewählt, dass  $\sum_{n \geq n_0} q^n < \frac{\varepsilon}{2}$  (Die Reihe konvergiert, da  $0 \leq q < 1$ ), und  $t$  so groß, dass  $\sum_{n=1}^{n_0-1} \mathbb{P}(M_{tK} = n)q^n < \frac{\varepsilon}{2}$  (möglich, da  $\mu = \mathbb{E}(B) > 1$  und  $M_{tK} \approx \prod_{i=1}^{tK} B_i$  mit  $(B_i)_{i=1, \dots, tK}$  stochastisch unabhängige Kopien von  $B$ ). Mit diesem  $t$  folgt

$$\begin{aligned}
 \sum_{n=0}^{\infty} \mathbb{P}(M_{tK} = n)q^n &= \sum_{n=1}^{n_0-1} \mathbb{P}(M_{tK} = n)q^n + \sum_{n=n_0}^{\infty} \mathbb{P}(M_{tK} = n)q^n \\
 &< \frac{\varepsilon}{2} + \sum_{n=n_0}^{\infty} q^n \\
 &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\
 &= \varepsilon.
 \end{aligned}$$

Wie im Beweis von Lemma 3.8 folgt

$$\mathbb{P}(H_n \geq c^*(1 - 2\varepsilon)n) \geq \mathbb{P}(A \cap B) \geq 1 - \mathbb{P}(A^c) - \mathbb{P}(B^c),$$

und diese untere Schranke kann durch Anpassung der Konstanten  $\beta$  und  $t$  so kontrolliert werden, dass sie beliebig nahe an 1 liegt. Daraus folgt die Behauptung des Lemmas.  $\square$

Die Lemmata 4.2 und 4.3 liefern zusammen die Aussage von Satz 4.1.

# 5 Die gewichtete Höhe sukzessiv errichteter Bäume

## 5.1 Vorbereitungen

In diesem Kapitel werden wir ein Resultat für die gewichtete Höhe eines sukzessiv (Individuum für Individuum) errichteten Baumes beweisen. In Kapitel 3 haben wir die Höhe  $H_n$  des zufälligen Teilbaumes  $\mathcal{T}_n$  eines  $b$ -adischen zweifach gewichteten Verzweigungsprozesses betrachtet. Diese ergab sich als das Maximum der kumulierten Gewichte  $D_v$  der Äste  $\pi(v)$  seiner Knoten  $v$ . Wird ein Baum nach einem bestimmten Schema sukzessiv aufgebaut, so kann man mit gleichem Interesse nach dem asymptotischen Verhalten des maximalen kumulierten Astgewichtes nach  $n$  Sukzessionsschritten (nach dem Anhängen des  $n$ -ten Knotens) fragen. Das größte Problem dabei ist die Anzahl der Knoten. Um unsere Ergebnisse aus den früheren Kapiteln anwenden zu können, müssen wir uns zunächst mit der zufälligen Anzahl  $N$  der Knoten von  $\mathcal{T}_m$  beschäftigen. Diese Anzahl ist abhängig von den Ergebnissen der Zufallsgrößen  $(G_v)_{v \in \mathcal{K}}$ . Die Vorgehensweise sieht im Folgenden so aus:

Wir konzentrieren uns für ein zuvor festgelegtes  $n \in \mathbb{N}$  zunächst darauf,  $m \in \mathbb{N}$  so zu wählen, dass die Anzahl  $N$  der Knoten von  $\mathcal{T}_m$  ungefähr  $n$  beträgt. Danach werden wir das zweite Hauptresultat über oben genanntes asymptotisches Verhalten beweisen. Dieses führt in vielen Anwendungen zu interessanten Ergebnissen. Ein Beispiel für einen sukzessiv errichteten Baum ist der so genannte *zufällige rekursive Baum*: Die Rekursion startet mit einem einzigen Individuum, der Wurzel  $\emptyset$ . Im ersten Schritt bekommt diese dann den ersten Nachkommen. Ab dem zweiten Schritt regiert der Zufall, denn nun bekommt entweder die Wurzel mit Wahrscheinlichkeit  $\frac{1}{2}$  ein weiteres Kind, oder das neue Individuum aus dem ersten Schritt bekommt ebenfalls mit Wahrscheinlichkeit  $\frac{1}{2}$  sein erstes Kind. Dieser Vorgang wird iterativ fortgesetzt. Im  $k$ -ten Schritt wird mit Hilfe einer Laplace-Verteilung auf den bereits vorhandenen  $k$  Individuen der Population entschieden, welches von ihnen das nächste Kind (das  $(k + 1)$ -te Mitglied der Population) bekommt.

In den folgenden Ausführungen benötigen wir einen bestimmten Typ des *Crump-Mode-Jagers-Prozess* (vgl. Ausführungen in [CuM]), nämlich den so genannten *Bellman-Harris-Prozess* (vgl. Ausführungen in [AuN]).

Sei  $X$  eine  $\mathbb{N}_0$ -wertige Zufallsgröße mit  $\mu = \mathbb{E}(X)$ . Wir stellen uns einen Verzweigungsprozess vor, der mit einem einzigen Individuum  $\emptyset$  startet. Dieses Individuum stirbt zu einem zufälligen Zeitpunkt  $M_1$  und gebärt im gleichen Moment  $X + 1$  voneinander unabhängige Individuen, die sich daraufhin genauso verhalten wie  $\emptyset$ . Wir nennen diese Ereignisse, wenn ein Individuum der Population stirbt und gleichzeitig Nachkommen hervorbringt, *Ersetzungen*. Als Verteilung für die Lebensdauern der Individuen unseres Prozesses wählen wir eine Exponentialverteilung, welche im Hinblick auf die weiter unten geschilderten Betrachtungen die einzig vernünftige Wahl darstellt, da sie den Prozess aufgrund der Gedächtnislosigkeit der Exponentialverteilung (vgl. [Als1], Kapitel 5, Satz 31.6 und Abschnitt 32 „Lebensdauerverteilungen“ auf Seite 146-148) markovsch macht. Wir stellen uns also vor, dass die Lebensdauer eines jeden Individuums der Population *Exp*(1)-verteilt ist, also  $M_1 \sim \text{Exp}(1)$  gilt. Sei  $M_k$  der *zufällige Zeitpunkt der k-ten Ersetzung* und sei  $N_k$  die Größe der Population unmittelbar vor  $M_k$ . Wegen der Gedächtnislosigkeit der Exponentialverteilung können wir uns vorstellen, zum Zeitpunkt  $M_k$  einen neuen Prozess mit  $N_{k+1} = N_k + X_k$  zum Zeitpunkt  $M_k$  geborenen Individuen zu starten, wobei die  $(X_k)_{k \geq 1}$  stochastisch unabhängige Kopien von  $X$  darstellen. Das bedeutet, dass die bereits verstrichene Lebenszeit der vor  $M_k$  geborenen und zum Zeitpunkt  $M_k$  noch lebenden Individuen ignoriert werden darf. Aus Symmetriegründen besitzt dann jedes der  $N_{k+1}$  Individuen dieselbe Wahrscheinlichkeit, als nächstes zu sterben.

Sei nun  $(E_i)_{i \geq 1}$  eine Familie von stochastisch unabhängigen, *Exp*(1)-verteilten Zufallsgrößen. Ein einfache Rechnung mit den jeweiligen Verteilungsfunktionen liefert

$$\min\{E_1, E_2, \dots, E_m\} \sim \frac{E_1}{m}. \quad (5.1)$$

Wir setzen  $X_0 \stackrel{\text{def}}{=} 1$  und erhalten:

**Lemma 5.1** *Für  $E_i, X_j$  und  $M_k$  ( $j \geq 0; i, k \geq 1$ ) wie oben gilt*

$$M_k \sim \sum_{i=1}^k \frac{E_i}{\sum_{j=0}^{i-1} X_j}.$$

BEWEIS: Die Aussage folgt aus der Gedächtnislosigkeit der Exponentialverteilung. Für  $k = 1$  gilt  $M_1 \sim E_1$  nach Modellannahme. Für  $k = 2$  erhalten wir

$$M_2 \sim E_1 + \frac{E_2}{1 + X_1}$$

nach (5.1), denn  $\frac{E_2}{1+X_1}$  entspricht in Verteilung der minimalen Lebensdauer eines der  $1 + X_1$  Individuen der Population unmittelbar nach  $M_1$ . Sukzessiv erhält man so die Behauptung für beliebiges  $k$ , denn nach (5.1) entspricht

$$\frac{E_k}{1 + X_1 + \cdots + X_{k-1}} \sim \min\{E_1, \dots, E_{\sum_{i=0}^{k-1} X_i}\}$$

in Verteilung der minimalen Lebensdauer eines der  $1 + X_1 + \cdots + X_{k-1}$  Individuen der Population unmittelbar nach  $M_{k-1}$ . □

Um die konkrete Anzahl der entstandenen Knoten abzuschätzen, nachdem wir den Prozess zum Zeitpunkt  $m_n$  gestoppt haben, benötigen wir eine Aussage über das asymptotische Verhalten der Ersetzungszeitpunkte  $M_k$  für  $k \rightarrow \infty$ . Dafür wiederum brauchen wir die folgenden Lemmata.

**Lemma 5.2 (Lemma von Kronecker)** Für  $n \in \mathbb{N}$  seien  $a_n$  und  $c_n$  reelle Zahlen mit  $0 < a_1 < a_2 < \cdots$  und  $\lim_{n \rightarrow \infty} a_n = \infty$ . Außerdem sei  $\sum_{n=1}^{\infty} \frac{c_n}{a_n}$  konvergent. Dann gilt

$$\lim_{n \rightarrow \infty} \frac{1}{a_n} \sum_{i=1}^n c_i = 0.$$

(vgl. Lemma 35.9 auf Seite 177 in [Als1] und Beweis in [Sch])

BEWEIS: 1. Schritt: Sei  $(b_n)_{n \in \mathbb{N}}$  eine Folge in  $\mathbb{R}$  mit  $b_n \xrightarrow[n \rightarrow \infty]{} b \in \mathbb{R}$ . Dann gilt

$$\lim_{n \rightarrow \infty} \sum_{i=1}^{n-1} \frac{a_{i+1} - a_i}{a_n} b_i = b.$$

Begründung: Es gilt

$$\sum_{i=1}^{n-1} \frac{a_{i+1} - a_i}{a_n} b_i - b = \sum_{i=1}^{n-1} \frac{a_{i+1} - a_i}{a_n} (b_i - b) - \frac{a_1}{a_n} b$$

und ferner

$$\sum_{i=1}^{n-1} \frac{a_{i+1} - a_i}{a_n} = \frac{a_n - a_1}{a_n} \xrightarrow[n \rightarrow \infty]{} 1.$$

Sei nun  $\varepsilon > 0$  beliebig und  $n_0 \in \mathbb{N}$  so gewählt, dass  $|b_n - b| < \varepsilon$  für alle  $n \geq n_0$ . Dann gilt

$$\left| \sum_{i=1}^{n-1} \frac{a_{i+1} - a_i}{a_n} b_i - b \right| \leq \sum_{i=1}^{n_0-1} \frac{a_{i+1} - a_i}{a_n} |b_i - b| + \varepsilon \frac{a_n - a_{n_0}}{a_n} + \frac{a_1}{a_n} |b| < 2\varepsilon$$

## 5 Die gewichtete Höhe sukzessiv errichteter Bäume

für  $n$  hinreichend groß, da  $\lim_{n \rightarrow \infty} a_n = \infty$  nach Voraussetzung.

2. Schritt: Für  $a_n, c_n$  wie oben setzen wir nun  $b_0 \stackrel{\text{def}}{=} 0$  und  $b_n \stackrel{\text{def}}{=} \sum_{i=1}^n \frac{c_i}{a_i}$ . Daraus folgt  $c_i = a_i(b_i - b_{i-1})$  und ferner, aufgrund der Konvergenz der Reihe  $\sum_{i=1}^{\infty} \frac{c_i}{a_i}$ ,

$$\frac{1}{a_n} \sum_{i=1}^n c_i = b_n - \sum_{i=1}^{n-1} \frac{a_{i+1} - a_i}{a_n} b_i \xrightarrow{n \rightarrow \infty} \sum_{i=1}^{\infty} \frac{c_i}{a_i} - \sum_{i=1}^{\infty} \frac{c_i}{a_i} = 0$$

nach dem ersten Schritt. □

**Lemma 5.3 (Verallgemeinertes starkes Gesetz der großen Zahlen)** Sei  $(E_i)_{i \geq 1}$  eine Familie von stochastisch unabhängigen und identisch verteilten Zufallsgrößen mit  $\mathbb{P}(E_1 \geq 0) = 1$ ,  $\mu \stackrel{\text{def}}{=} \mathbb{E}(E_1) < \infty$  und  $\sigma^2 \stackrel{\text{def}}{=} \text{Var}(E_1) \in (0; \infty)$ . Dann gilt

$$\frac{1}{\log k} \sum_{i=1}^k \frac{E_i}{i} \xrightarrow{k \rightarrow \infty} \mu$$

fast sicher. (Weiterführende Betrachtungen findet man zum Beispiel in [CuT].)

BEWEIS: Für  $n \geq 2$  setze

$$W_n \stackrel{\text{def}}{=} \sum_{i=2}^n \frac{E_i - \mu}{i \log i}.$$

Dann ist  $(W_n)_{n \geq 2}$  ein  $\mathcal{L}_2$ -Martingal bezüglich der Filtration  $\mathcal{F}_n \stackrel{\text{def}}{=} \sigma\{E_i \mid i = 1, \dots, n\}$  ( $n \geq 1$ ), denn für  $n \geq 3$  gilt

$$\mathbb{E}(W_n | \mathcal{F}_{n-1}) = \mathbb{E}\left(W_{n-1} + \frac{E_n - \mu}{n \log n} \middle| \mathcal{F}_{n-1}\right) = W_{n-1} + \mathbb{E}\left(\frac{E_n - \mu}{n \log n}\right) = W_{n-1}$$

fast sicher aufgrund der  $\mathcal{F}_{n-1}$ -Messbarkeit von  $W_{n-1}$  und der stochastischen Unabhängigkeit von  $E_n$  und  $\sigma\{E_1, \dots, E_{n-1}\}$ . Die quadratische Integrierbarkeit der  $W_n$  ( $n \geq 2$ ) folgt aus der quadratischen Integrierbarkeit und der stochastischen Unabhängigkeit der  $E_i$  ( $i \geq 1$ ).

Mit Hilfe von Lemma 23.1 in [Als2] erhalten wir daher für  $n \geq 3$

$$\begin{aligned} \mathbb{E}(W_n^2) &= \mathbb{E}(W_2^2) + \sum_{i=3}^n \mathbb{E}((W_i - W_{i-1})^2) \\ &= \mathbb{E}\left(\left(\frac{E_2 - \mu}{2 \log 2}\right)^2\right) + \sum_{i=3}^n \mathbb{E}\left(\left(\frac{E_i - \mu}{i \log i}\right)^2\right) \\ &= \frac{\sigma^2}{(2 \log 2)^2} + \sum_{i=3}^n \frac{\sigma^2}{(i \log i)^2} \\ &= \sigma^2 \sum_{i=2}^n \frac{1}{i^2 (\log i)^2}. \end{aligned}$$

Unter Beachtung von  $\sum_{n=2}^{\infty} \frac{1}{i^2 (\log i)^2} < \infty$  folgt aus dieser Abschätzung  $\sup_{n \geq 2} \mathbb{E}(W_n^2) < \infty$ . Insbesondere ist  $(W_n)_{n \geq 2}$  gleichgradig integrierbar (vgl. [Als1], Definition 50.1 und Satz 50.2(e) auf Seite 276).

Wegen der gleichgradigen Integrierbarkeit von  $(W_n)_{n \geq 2}$  folgt mit Hilfe von Satz 22.2 in [Als2]

$$W_n \xrightarrow[n \rightarrow \infty]{} W_\infty$$

fast sicher für eine  $\mathcal{F}_\infty \stackrel{\text{def}}{=} \sigma\{E_i \mid i \geq 1\}$ -messbare Zufallsgröße  $W_\infty$ . Dann folgt mittels Lemma 5.2

$$\frac{1}{\log n} \sum_{i=1}^n \frac{E_i - \mu}{i} \xrightarrow[n \rightarrow \infty]{} 0$$

fast sicher, und daraus wiederum

$$\frac{1}{\log n} \sum_{i=1}^n \frac{E_i}{i} \xrightarrow[n \rightarrow \infty]{} \mu = \lim_{k \rightarrow \infty} \frac{1}{\log k} \sum_{i=1}^k \frac{\mu}{i}$$

fast sicher. Dabei haben wir für die letzte Gleichheit  $\gamma = \lim_{k \rightarrow \infty} \left(\sum_{i=1}^k \frac{1}{i} - \log k\right)$  (Eulersche Konstante) verwendet.

□

Kommen wir nun zu dem oben angesprochenen Satz über das asymptotische Verhalten der Ersetzungszeitpunkte:

**Satz 5.4 (Asymptotisches Verhalten der Ersetzungszeitpunkte)** *Mit  $\mu = \mathbb{E}(X)$  gilt für die Ersetzungszeitpunkte  $(M_k)_{k \geq 1}$*

$$\frac{M_k}{\log k} \xrightarrow[k \rightarrow \infty]{} \frac{1}{\mu}$$

*fast sicher.*

## 5 Die gewichtete Höhe sukzessiv errichteter Bäume

BEWEIS: Nach dem starken Gesetz der großen Zahlen gilt

$$\frac{1}{i} \sum_{j=0}^{i-1} X_j = \underbrace{\frac{1}{i}}_{\rightarrow 0} + \underbrace{\frac{X_1 + \dots + X_{i-1}}{i-1}}_{\rightarrow \mu \text{ f.s.}} \underbrace{\frac{i-1}{i}}_{\rightarrow 1} \xrightarrow{i \rightarrow \infty} \mu = \mathbb{E}(X)$$

fast sicher. Die Behauptung des Satzes folgt mit Hilfe des Lemmas 5.3.

Nach Lemma 5.1 gilt

$$\frac{M_k}{\log k} \sim \frac{1}{\log k} \sum_{i=1}^k \frac{E_i}{\sum_{j=0}^{i-1} X_j} = \frac{1}{\log k} \sum_{i=1}^k \frac{E_i}{i^{\frac{1}{i}} \sum_{j=0}^{i-1} X_j}.$$

Seien  $\varepsilon > 0$  beliebig und  $\omega \in \Omega$  beliebig mit

$$\frac{1}{i} \sum_{j=0}^{i-1} X_j(\omega) \xrightarrow{i \rightarrow \infty} \mu \quad \text{und} \quad \frac{1}{\log k} \sum_{i=1}^k \frac{E_i(\omega)}{i} \xrightarrow{k \rightarrow \infty} 1.$$

Die zweite Bedingung gilt dabei nach Lemma 5.2 fast sicher, denn  $\mathbb{E}(E_i) = 1$  für alle  $i \geq 1$ . Zu  $\varepsilon$  wähle nun  $l \in \mathbb{N}$  so groß, dass  $\frac{1}{i} \sum_{j=0}^{i-1} X_j(\omega) \in (\mu - \varepsilon; \mu + \varepsilon)$  für alle  $i > l$  gilt. Für  $k > l$  erhalten wir

$$\begin{aligned} \frac{1}{\log k} \sum_{i=1}^k \frac{E_i(\omega)}{\sum_{j=0}^{i-1} X_j(\omega)} &= \frac{1}{\log k} \sum_{i=1}^l \frac{E_i(\omega)}{\sum_{j=0}^{i-1} X_j(\omega)} + \frac{1}{\log k} \sum_{i=l+1}^k \frac{E_i(\omega)}{\sum_{j=0}^{i-1} X_j(\omega)} \\ &\in \left( \frac{1}{\mu + \varepsilon} \frac{1}{\log k} \sum_{i=l+1}^k \frac{E_i(\omega)}{i} - \varepsilon; \frac{1}{\mu - \varepsilon} \frac{1}{\log k} \sum_{i=l+1}^k \frac{E_i(\omega)}{i} + \varepsilon \right) \\ &\subseteq \left( \frac{1 - \varepsilon}{\mu + \varepsilon} - \varepsilon; \frac{1 + \varepsilon}{\mu - \varepsilon} + \varepsilon \right) \end{aligned}$$

für  $k$  hinreichend groß. Dabei haben wir in der zweiten Zeile benutzt, dass  $\frac{1}{\log k} \sum_{i=1}^l \frac{E_i(\omega)}{\sum_{j=0}^{i-1} X_j(\omega)}$  für  $k$  hinreichend groß in  $(-\varepsilon; \varepsilon)$  liegt. Obige Abschätzung gilt für fast alle  $\omega \in \Omega$ . Daraus folgt

$$\frac{M_k}{\log k} \xrightarrow{k \rightarrow \infty} \frac{1}{\mu}$$

fast sicher. □

## 5.2 Das asymptotische Verhalten des maximalen kumulierten Astgewichtes nach $n$ Sukzessionsschritten

Nach den Vorbereitungen im vorherigen Abschnitt sind wir nun in der Lage, die Anzahl der Knoten des Baumes genauer zu studieren. Wir setzen

$$m_n \stackrel{\text{def}}{=} \frac{1}{\mu} \log n$$

und schreiben  $N(t)$  für die zufällige Anzahl der Knoten des Baumes zum Zeitpunkt  $t > 0$ . Es wird sich herausstellen, dass  $N(m_n)$  nahe genug bei  $n$  liegt, um Gebrauch von den Ergebnissen des dritten Kapitels machen zu können. Das asymptotische Verhalten der Anzahl der Knoten untersucht der folgende Satz:

**Satz 5.5 (Asymptotisches Verhalten der Anzahl der Knoten)** Für die Folge  $(N(m_n))_{n \geq 1}$  der Anzahlen der Knoten bis zu den Zeitpunkten  $(m_n)_{n \geq 1}$  gilt

$$\frac{\log N(m_n)}{\log n} \xrightarrow{n \rightarrow \infty} 1$$

fast sicher.

BEWEIS: Zunächst liefert das starke Gesetz der großen Zahlen

$$\frac{1 + X_1 + \cdots + X_{k-1}}{k} = \frac{N_k}{k} \xrightarrow{k \rightarrow \infty} \mu$$

fast sicher (siehe Beweis von Satz 5.4). Daraus folgt

$$\frac{\log N_k}{\log k} \xrightarrow{k \rightarrow \infty} 1$$

fast sicher, denn aus  $\frac{N_k}{k} \xrightarrow{k \rightarrow \infty} \mu$  fast sicher folgt sogar

$$\log N_k - \log k = \log \left( \frac{N_k}{k} \right) \xrightarrow{k \rightarrow \infty} \log \mu$$

fast sicher aufgrund der Stetigkeit des Logarithmus. Mit Hilfe von Satz 5.4 erhalten wir ferner

$$\frac{\log N_k}{\mu M_k} \xrightarrow{k \rightarrow \infty} 1$$

fast sicher, und daher

$$\frac{\log N(t)}{\mu t} \xrightarrow[t \rightarrow \infty]{} 1$$

fast sicher. Mit  $t = m_n$  folgt die Behauptung. □

Der nun folgende Satz ist von großem Interesse in vielen Anwendungen (siehe Kapitel 6) und stellt das zweite Hauptresultat dieser Arbeit dar. Die Beweisidee stützt sich auf die Ergebnisse aus Kapitel 4, insbesondere auf den Satz 4.1.

Einen sukzessiv aufgebauten Baum mit  $n$  Knoten generiert durch sukzessives Anhängen von  $1 + X$  Kindern an einen Knoten, der zufällig unter den bereits bestehenden Knoten ausgewählt wird, bezeichnen wir im Folgenden mit  $\mathcal{T}'_n$ . Die Lebensdauern seiner Individuen interpretieren wir im Folgenden als  $Exp(1)$ -verteilte Astgewichte.

Um das Resultat im Hinblick auf die Anwendungen in Kapitel 6 so allgemein wie möglich zu halten, versehen wir die Kanten  $e$  des sukzessiv aufgebauten Baumes mit zusätzlichen Gewichten  $Z_e$ , welche die gleichen Eigenschaften wie die Gewichte  $Z_e$  aus Kapitel 2 besitzen. Wie beim Modell für den zweifach gewichteten Verzweigungsprozess bezeichne für einen Knoten  $v$  des sukzessiv aufgebauten Baumes  $D_v = \sum_{e \in \pi(v)} Z_e = \sum_{i=1}^{|v|} Z_{v_i}(v|(i-1))$  die kumulierten Astgewichte des Astes von der Wurzel bis zum Knoten  $v$ .

Im Beweis von Satz 5.6 wird nun ausgenutzt, dass der so konstruierte (unendlich fortgesetzte) sukzessiv aufgebaute Baum  $\mathcal{T}'_\infty$  einem verallgemeinerten zweifach gewichteten Verzweigungsprozess  $\mathcal{T}^*_\infty$  mit der Reproduktionsverteilung  $(p_k)_{k \in \mathbb{N}_0}$  mit  $p_0 = 0$  und  $p_k \stackrel{\text{def}}{=} \mathbb{P}(X = k - 1)$  für  $k \in \mathbb{N}$  und den Gewichtsfunktionen  $Z$  und  $E \sim Exp(1)$  entspricht. Mit  $\mu = \mathbb{E}(X)$  gilt dann insbesondere  $\mathbb{E}((p_k)_{k \in \mathbb{N}_0}) = 1 + \mu$ .

Sei im Folgenden

$$H_n^* \stackrel{\text{def}}{=} \max\{D_v \mid v \text{ Knoten von } \mathcal{T}'_n\}$$

die zufällige Höhe eines sukzessiv aufgebauten Baumes mit  $n$  Knoten (bezüglich der Kantengewichte  $Z_e$ ,  $e$  Kante von  $\mathcal{T}'_n$ ).

**Satz 5.6 (Asymptotisches Verhalten der gewichteten Höhe  $H_n^*$ )** *Sei  $\mathcal{T}'_n$  wie oben ein sukzessiv aufgebauter Baum mit  $n$  Knoten und  $H_n^*$  seine zufällige Höhe bezüglich des Kantengewichtes  $Z$ . Ferner sei  $c^*$  die Konstante aus Satz 4.1 bezüglich eines verallgemeinerten zweifach gewichteten Verzweigungsprozesses  $\mathcal{T}^*_\infty$  mit der Reproduktionsverteilung*

## 5.2 Das asymptotische Verhalten des maximalen kumulierten Astgewichtes nach $n$ Sukzessionsschritten

$(p_k)_{k \in \mathbb{N}_0}$  mit  $p_0 = 0$  und  $p_k = \mathbb{P}(X = k - 1)$  für  $k \in \mathbb{N}$  und den Gewichtsfunktionen  $Z$  und  $E \sim \text{Exp}(1)$ . Dann gilt

$$\frac{H_n^*}{\log n} \xrightarrow{n \rightarrow \infty} \frac{c^*}{\mu}$$

nach Wahrscheinlichkeit.

BEWEIS: Nach Satz 5.5 gilt  $\frac{\log N(m_n)}{\log n} \xrightarrow{n \rightarrow \infty} 1$  fast sicher. Mit Wahrscheinlichkeit beliebig nahe an 1 (für  $n$  hinreichend groß) folgt daraus

$$n^{1-\varepsilon} \leq N(m_n) \leq n^{1+\varepsilon},$$

und daraus weiter mit Wahrscheinlichkeit beliebig nahe an 1 (für  $n$  hinreichend groß)

$$\mu(1 - \varepsilon) \frac{H_{[n^{1-\varepsilon}]}^*}{(1 - \varepsilon) \log n} = \frac{H_{[n^{1-\varepsilon}]}^*}{m_n} \leq \frac{H_{N(m_n)}^*}{m_n} \leq \frac{H_{[n^{1+\varepsilon}]}^*}{m_n} = \mu(1 + \varepsilon) \frac{H_{[n^{1+\varepsilon}]}^*}{(1 + \varepsilon) \log n}. \quad (5.2)$$

Die gewichtete Höhe  $H_{N(m_n)}^*$  des sukzessiv errichteten Baumes mit  $N(m_n)$  Knoten entspricht nach Definition von  $N(m_n)$  der gewichteten Höhe des sukzessiv errichteten Baumes zum Zeitpunkt  $m_n$ . Da die Lebensdauern der Individuen von  $\mathcal{T}'_\infty$  aber durch  $\text{Exp}(1)$ -verteilte Kantengewichte repräsentiert werden, gilt

$$H_{N(m_n)}^* = H_{m_n}$$

mit  $H_{m_n}$  der gewichteten Höhe des Teilbaumes  $\mathcal{T}_{m_n}^*$  eines verallgemeinerten zweifach gewichteten Verzweigungsprozesses  $\mathcal{T}_\infty^*$  mit der Reproduktionsverteilung  $(p_k)_{k \in \mathbb{N}_0}$  und den Gewichtsfunktionen  $Z$  und  $E \sim \text{Exp}(1)$ . Nach Satz 4.1 gilt

$$\frac{H_{\log n}}{\log n} \mu = \frac{H_{m_n}}{m_n} \xrightarrow{n \rightarrow \infty} c^*$$

nach Wahrscheinlichkeit. Zusammen mit der Abschätzung (5.2) folgt daraus die Behauptung des Satzes. □

**Bemerkung 5.7** Da die Exponentialverteilung die einzige gedächtnislose Verteilung ist, führt die Wahl einer anderen Verteilung für die Gewichtsfunktion  $E$  zu ungleichen Wahrscheinlichkeiten unter den einzelnen Individuen der Population in Bezug auf ihre verbleibende Lebensdauer. Für diesen allgemeineren Fall ist noch kein vergleichbares Resultat bekannt.



# 6 Anwendungen

In diesem Kapitel tragen wir einige Beispiele bestimmter Verzweigungsprozesse zusammen. Eine Anwendung der in den ersten fünf Kapiteln gewonnenen Ergebnisse führt dabei zu tieferen Einblicken in die Natur dieser Prozesse. Die Cramér-Funktionen  $\Lambda_Z^*$  und  $\Lambda_E^*$  sind in der Regel schwer in eine geschlossene Form zu bringen. Dadurch erweist es sich oft als problematisch, den optimalen Punkt  $(\rho_0, \alpha_0) \in \mathcal{C}_{Z,E}$  für unsere Konstante  $c$  aus den Sätzen 3.6 und 5.6 zu bestimmen.

## 6.1 Der zufällige binäre Suchbaum

Der *zufällige binäre Suchbaum* (Abkürzung RBST für *random binary search tree*) ist ein Modell, welches den Vorgang beschreibt,  $n$  numerische Elemente der Größe nach zu sortieren. Es handelt sich um eine Version des *Quicksort-Algorithmus* (vgl. Kapitel 4, Abschnitt 3 *Beispiele* in [Als3]). Dabei wird eine zufällige Permutation  $Y_1, \dots, Y_n$  der Zahlen  $\{1, 2, \dots, n\}$  Element für Element als Knoten in einen anfänglich noch leeren binären (2-adischen) Baum eingefügt. Wir definieren zunächst den *relativen Rang*  $R_i$  des Elementes  $Y_i$  als den Rang von  $Y_i$  innerhalb der Menge  $\{Y_1, \dots, Y_i\}$ . Es gilt also

$$R_i \stackrel{\text{def}}{=} |\{1 \leq j \leq i : Y_j \leq Y_i\}|.$$

Im ersten Schritt der Konstruktion setzen wir  $Y_1$  als Wurzel in unseren noch leeren binären Baum ein und schicken alle  $Y_i$  mit  $Y_i < Y_1$  der Reihe nach in den linken und die anderen  $Y_i$  mit  $Y_i > Y_1$  der Reihe nach in den rechten Teilbaum. Im zweiten Schritt wenden wir dasselbe Verfahren auf die Mengen der Elemente im linken und im rechten Teilbaum an. Diesen Prozess setzen wir sukzessiv fort, bis sich nur noch einelementrige Teilmengen der ursprünglichen Menge  $\{Y_1, \dots, Y_n\}$  in den jeweiligen Teilbäumen befinden. Diese Elemente stellen zugleich die Knoten an jenen Stellen dar. Zur Veranschaulichung dient das folgende Bild:

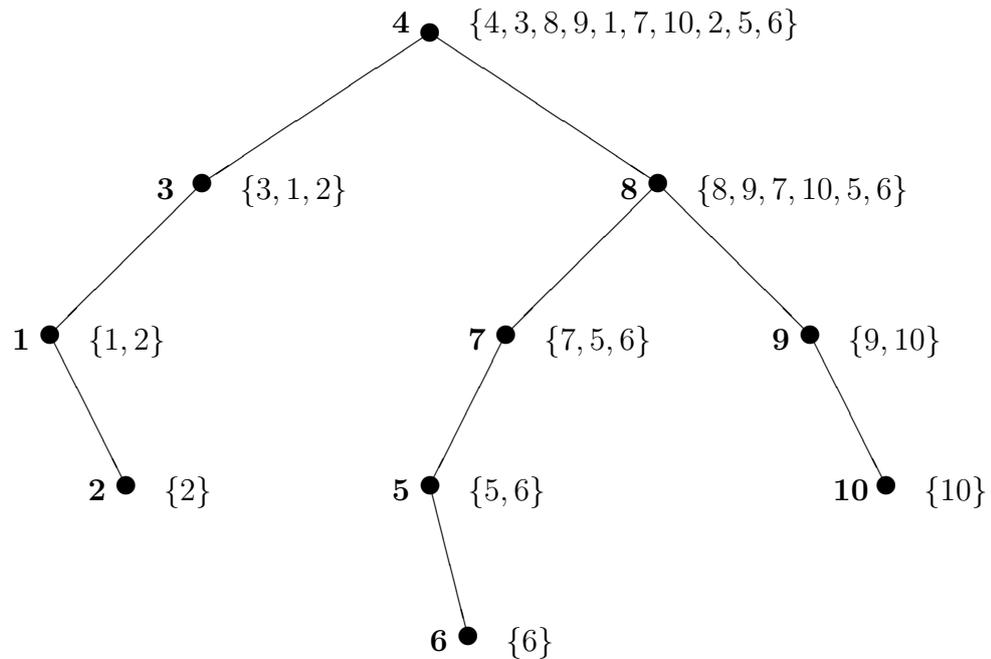


Abbildung 7: Zufälliger binärer Suchbaum der Permutation  
4, 3, 8, 9, 1, 7, 10, 2, 5, 6 für  $n = 10$

Nach  $i - 1$  Sukzessionsschritten bezeichnen wir ferner einen bis dahin leer gebliebenen Knoten als *externen Knoten*, sofern er Kind eines bereits besetzten Knotens ist, er also im  $i$ -ten Sukzessionsschritt als Zielort für das Element  $Y_i$  in Frage kommt. Außerdem bezeichnen wir nach  $i - 1$  Sukzessionsschritten einen Knoten als *internen Knoten*, wenn er bis zu diesem Zeitpunkt bereits besetzt worden ist.

Uns interessiert im Folgenden das asymptotische Verhalten der gewichteten Höhe  $H_n^*$  eines RBST mit  $n$  Knoten für  $n \rightarrow \infty$ . Eine Anwendung von Satz 5.6 liefert:

**Satz 6.1 (Gewichtete Höhe des zufälligen binären Suchbaumes)** *Für die gewichtete Höhe  $H_n^*$  eines zufälligen binären Suchbaumes mit  $n$  Knoten gilt*

$$\frac{H_n^*}{\log n} \xrightarrow[n \rightarrow \infty]{} c$$

nach Wahrscheinlichkeit, wobei  $c$  die eindeutig bestimmte Lösung größer als 1 von  $\frac{2e}{c} = e^{\frac{1}{c}}$  ist.

BEWEIS: Bezeichne  $\mathcal{T}'_\infty$  einen unendlich fortgesetzten RBST. L. Devroye zeigt in seinen Arbeiten [Dev1], [Dev2] und [Dev5], dass ein binärer zweifach gewichteter Verzweigungsprozess mit Gewichtsfunktionen  $E \sim \text{Exp}(1)$  und  $Z \equiv 1$  ein geeignetes Modell für  $\mathcal{T}'_\infty$

darstellt:

Für die in Kapitel 5 verwendeten Notationen erhalten wir speziell für den unendlich fortgesetzten RBST  $1 + X = 2$  und  $Z = 1$  mit Wahrscheinlichkeit 1, denn es handelt sich um einen binären Baum und jede Kante des Baumes steht für genau einen Sukzessions-schritt. Damit gilt für die gewichtete Höhe  $D_v$  eines Knotens  $v$ , dass  $D_v = |v|$  fast sicher. Für die Rechtfertigung der  $Exp(1)$ -Verteilung für das andere Astgewicht im unendlich fortgesetzten RBST  $\mathcal{T}'_\infty$  verweisen wir auf die oben genannten Literaturquellen.

Der umfangreiche Beweis stützt sich zum einen auf die Eigenschaften der relativen Ränge  $(R_i)_{i=1,\dots,n}$  einer zufälligen Permutation  $Y_1, \dots, Y_n$  der Zahlen  $\{1, 2, \dots, n\}$ : Nach Lemma 25.24 in [Als2] gilt, dass die Zufallsgrößen  $R_1, \dots, R_n$  stochastisch unabhängig und für  $i \in \{1, \dots, n\}$  der relative Rang  $R_i$  Laplace-verteilt ist auf der Menge  $\{1, \dots, i\}$  (siehe auch [Mah], [Knu], [Dev3] und [Gli]). Daher besitzen in jedem Sukzessionsschritt alle *externen Knoten* die gleiche Wahrscheinlichkeit, ausgewählt und durch ein  $Y_i$  ersetzt zu werden. Stellt man sich nun vor, dass sich die Individuen des unendlich fortgesetzten RBST  $\mathcal{T}'_\infty$  nach Ablauf ihrer Lebensdauer (welche durch das Kantengewicht  $E$  repräsentiert werden soll) jeweils in zwei neue Individuen aufteilen (eine unendliche Liste von verschiedenen ganzen Zahlen wird auf zwei unendliche Teillisten aufgeteilt), so besitzen also die Restlebensdauern der Individuen im unendlich fortgesetzten RBST unmittelbar nach einem Sukzessionsschritt alle dieselbe Verteilung. Damit müssen die Verteilungen der Lebensdauern der Individuen identisch und gedächtnislos sein.

Zum anderen werden für einen Knoten  $v$  eines RBST  $\mathcal{T}'_n$  mit  $n$  Knoten die Größen (das heißt die Anzahl der Knoten) des linken und des rechten Teilbaumes mit Wurzel  $(v, 1)$  beziehungsweise  $(v, 2)$  betrachtet. Tiefgründige Überlegungen führen schließlich dazu, dass im unendlich fortgesetzten RBST  $\mathcal{T}'_\infty$  jedem Kantenpaar  $e$  und  $f$ , das einen Knoten  $v$  mit seinen beiden Kindern  $(v, 1)$  und  $(v, 2)$  verbindet, Gewichte  $U_v \sim R(0, 1)$  und  $1 - U_v$  zugeordnet werden, wobei  $(U_v)_{v \in \mathcal{K}}$  (mit  $\mathcal{K}$  der Menge der Knoten von  $\mathcal{T}'_\infty$ ) eine Familie von stochastisch unabhängigen Zufallsgrößen ist. Für einen Knoten  $v$  von  $\mathcal{T}'_\infty$  ergibt sich die Höhe bezüglich dieser Gewichte (die relative Größe des Teilbaumes mit Wurzel in  $v$  relativ zur Größe des gesamten Baumes) als ein Produkt von stochastisch unabhängigen  $R(0, 1)$ -verteilten Zufallsgrößen. Eine Anwendung des Logarithmus liefert einen Übergang von diesem multiplikativen Modell zu einem additiven Modell. Schließlich führt eine Grenzwertbetrachtung und eine Identifikation der gewichteten Höhe eines Knotens  $v$  bezüglich des Kantengewichtes  $E$  mit der relativen Größe des Teilbaumes mit Wurzel  $v$  unter Beachtung von  $-\log U \sim Exp(1)$  für  $U \sim R(0, 1)$  zu der gewünschten  $Exp(1)$ -Verteilung des Kantengewichtes  $E$  (Einen detaillierten Beweis dazu findet man in [Dev1], [Dev2] und [Dev5]).

## 6 Anwendungen

Da uns im Kontext dieses Satzes nur das asymptotische Verhalten der gewichteten Höhe eines RBST mit  $n$  Knoten für  $n \rightarrow \infty$  interessiert, genügt es nach obigen Überlegungen, das Ergebnis von Satz 5.6 auf den Fall  $E \sim \text{Exp}(1)$ ,  $1 + X = 2$  und  $Z = 1$  mit Wahrscheinlichkeit 1 anzuwenden.

Für eine  $\text{Exp}(1)$ -verteilte Gewichtsfunktion  $E$  gilt für ihre momenterzeugende Funktion für  $\lambda \neq 1$

$$M_E(\lambda) = \mathbb{E}(e^{\lambda E}) = \mathbb{E}(U^{-\lambda}) = \int_0^1 \left(\frac{1}{x}\right)^\lambda dx = \frac{1}{1-\lambda}.$$

Daher ist ihre kumulantenerzeugende Funktion für  $\lambda \leq 0$

$$\Lambda_E(\lambda) = -\log(1-\lambda).$$

Die Cramér-Funktion von  $E$  ergibt sich für  $\text{essinf}(E) = 0 \leq \rho \leq 1 = \mathbb{E}(E)$  als

$$\Lambda_E^*(\rho) = \Lambda_{l,E}^*(\rho) = \sup_{\lambda \leq 0} \{\lambda \rho - \Lambda_E(\lambda)\} = \sup_{\lambda \leq 0} \{\lambda \rho + \log(1-\lambda)\}.$$

Dabei wird das Supremum nach Lemma 1.14 in  $\lambda = 1 - \frac{1}{\rho} < 0$  ( $\Leftrightarrow \rho = \Lambda_E'(\lambda) = \frac{1}{1-\lambda}$ ) angenommen und daher gilt für  $\rho > 0$

$$\Lambda_E^*(\rho) = \rho - 1 - \log \rho.$$

Da zudem

$$\Lambda_Z^*(\alpha) = \begin{cases} 0 & \text{für } \alpha = 1 \\ \infty & \text{für } \alpha \neq 1 \end{cases}$$

gilt, wird das Maximum von  $(\rho, \alpha) \mapsto \frac{\alpha}{\rho}$  in  $\mathcal{C}_{Z,E}$  im Punkt  $(\rho_0, \alpha_0)$  mit  $\alpha_0 = 1$  und  $\rho_0$  der eindeutig bestimmten Lösung kleiner als  $1 = \mathbb{E}(E)$  von  $\rho - 1 - \log \rho = \log 2$  angenommen (denn  $\Lambda_Z^*(1) = 0$ ). Deswegen ist  $c = \frac{1}{\rho_0}$ .

Das Ergebnis von Satz 5.6 und die Äquivalenz  $\rho - 1 - \log \rho = \log 2 \Leftrightarrow 2e\rho = e^\rho$  liefern die Behauptung des Satzes. Eine numerische Approximation liefert  $\rho_0 \approx 0,2319$  und damit gilt  $c \approx 4,3122$ .

□

## 6.2 Der zufällige rekursive Baum

Beim *zufälligen rekursiven Baum* (Abkürzung RRT für *random recursive tree*) handelt es sich um einen der einfachsten rekursiv aufgebauten Bäume. Wie wir bereits in der Einleitung des fünften Kapitels erwähnt haben, startet die Rekursion mit einem einzigen Individuum, der Wurzel  $v_1$ . Im ersten Schritt gebärt dieses dann das erste Kind  $v_2$ , wonach im zweiten Schritt der Zufall ins Spiel kommt: Entweder bekommt mit Wahrscheinlichkeit  $\frac{1}{2}$  die Wurzel  $v_1$  ein weiteres Kind, oder das neue Individuum  $v_2$  aus dem ersten Schritt bekommt mit Wahrscheinlichkeit  $\frac{1}{2}$  sein erstes Kind. Dieser Vorgang wird sukzessiv fortgesetzt, so dass im  $k$ -ten Schritt mit Hilfe einer Laplace-Verteilung auf den bereits vorhandenen  $k$  Individuen der Population  $v_1, \dots, v_k$  entschieden wird, welches von ihnen das  $(k+1)$ -te Mitglied der Population  $v_{k+1}$  gebären darf. Unser Hauptinteresse liegt wie beim vorangegangenen Beispiel auf dem asymptotischen Verhalten der Höhe  $H_n^*$  eines RRT mit  $n$  Knoten für  $n \rightarrow \infty$ . Das folgende Bild veranschaulicht einen RRT nach 8 Sukzessionsschritten.

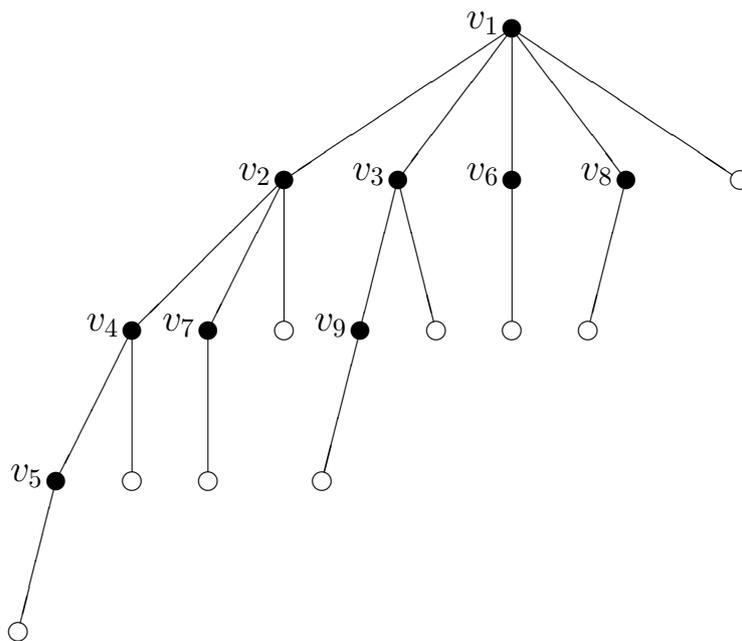


Abbildung 8: Ein zufälliger rekursiver Baum nach 8 Sukzessionsschritten. Die ausgefüllten Kreise stehen dabei für bereits besetzte (interne) Knoten, wohingegen die transparenten Kreise die externen Knoten bezeichnen.

**Satz 6.2 (Gewichtete Höhe des zufälligen rekursiven Baumes)** Für die gewichtete Höhe  $H_n^*$  eines zufälligen rekursiven Baumes mit  $n$  Knoten gilt

$$\frac{H_n^*}{\log n} \xrightarrow{n \rightarrow \infty} e$$

nach Wahrscheinlichkeit.

BEWEIS: Um Gebrauch von Satz 5.6 machen zu können, interpretieren wir den zufälligen rekursiven Baum als binären Baum (vgl. Abbildung 9). Im Beweis von Satz 6.1 haben wir gesehen, dass wir einen RBST aufbauen können, indem wir in jedem Sukzessionsschritt einen externen Knoten mittels einer Laplace-Verteilung auf der Menge der externen Knoten auswählen und durch das entsprechende Element der zufälligen Permutation ersetzen. Das funktioniert aufgrund der Eigenschaften der relativen Ränge: Für jedes  $i \in \{1, \dots, n\}$  ist  $R_i$  Laplace-verteilt auf  $\{1, \dots, i\}$  (vgl. Lemma 25.24 in [Als2]). Daher besteht von Natur aus ein sehr enger Zusammenhang zwischen dem sukzessiven Auswahlprozess des RRT und dem des RBST.

Im ersten Schritt des Beweises wollen wir die Knoten des RRT auf die externen Knoten eines korrespondierenden binären Baumes abbilden. Dabei müssen wir zunächst sicherstellen, dass die Information der Abstände der einzelnen Knoten von der Wurzel erhalten bleiben, denn schließlich wollen wir später eine Aussage über die gewichtete Höhe des Baumes machen. Sei  $\mathcal{T}'_n$  ein RRT mit  $n$  Knoten und sei  $S_n = \{d_1, d_2, \dots, d_n\}$  die (zufällige) Menge der Abstände der Knoten  $v_1, v_2, \dots, v_n$  von der Wurzel. Zur Veranschaulichung der oben angesprochenen Abbildung beschreiben wir die Konstruktion des korrespondierenden binären Baumes  $\mathcal{T}_n^b$  mit  $n$  externen Knoten zusammen mit der Menge  $S_n^b$  der gewichteten Distanzen seiner externen Knoten von der Wurzel (vgl. Abbildung 9).

$\mathcal{T}'_1$  besitzt nur einen einzigen Knoten, die Wurzel  $v_1$ , und somit ist  $S_1 = \{0\}$ . Das Anhängen eines zweiten Knotens  $v_2$  führt zu einem Baum mit den Knotenabständen  $S_2 = \{0, 1\}$ . Sei nun  $\mathcal{T}_2^b$  ein binärer Baum mit zwei externen Knoten (in der ersten Generation). Seine beiden Kanten bezeichnen wir mit  $e$  und  $f$  und gewichten diese mit  $z_e = 1$  (linke Kante) und  $z_f = 0$  (rechte Kante). Außerdem bezeichnen wir seinen linken externen Knoten mit  $v_2$  und seinen rechten externen Knoten mit  $v_1$ . Das bedeutet, dass im ersten Sukzessionsschritt im korrespondierenden binären Baum der Knoten  $v_1$  um eine Generation nach rechts unten wandert und dessen Kind  $v_2$  die Position links neben ihm einnimmt. Die tatsächlichen Abstände im RRT entsprechen also gerade den gewichteten Abständen im korrespondierenden binären Baum. Somit erhalten wir für die Menge der

gewichteten Knotenabstände  $S_2^b = \{0, 1\} = S_2$ .

Nach  $n - 1$  Konstruktionsschritten wie oben erhalten wir für einen RRT  $\mathcal{T}'_n$  mit  $n$  Knoten den zugehörigen binären Baum  $\mathcal{T}_n^b$  mit  $n$  externen Knoten. Dabei stimmen die Distanzen  $S_n$  der Knoten des RRT mit den wie oben gewichteten Distanzen  $S_n^b$  der externen Knoten des binären Baumes überein. Ein weiter Sukzessionsschritt stellt sich dann wie folgt dar: Das Anhängen eines Knotens  $v$  an einen bereits bestehenden Knoten  $u$  von  $\mathcal{T}'_n$  führt dazu, dass die nun neue Menge der Distanzen  $S_{n+1}$  gleich  $S_n \cup \{d + 1\}$  mit  $d$  dem Abstand des Knotens  $u$  von der Wurzel ist. Für den zugehörigen binären Baum bedeutet das, dass wir den externen Knoten  $u$  durch einen internen Knoten  $x$  ersetzen, wobei zwei neue externe Knoten entstehen. Die beiden von  $x$  ausgehenden Kanten bezeichnen wir wieder mit  $e$  (linke Kante) und  $f$  (rechte Kante) und gewichten diese wieder mit  $z_e = 1$  und  $z_f = 0$ . Das linke Kind von  $x$  bezeichnen wir mit  $v$  und das rechte entsprechend mit  $u$ . Damit haben wir in der Tat  $S_{n+1}^b = S_n^b \cup \{d + 1\} = S_{n+1}$  und damit stimmen für alle  $n \in \mathbb{N}$  die Distanzen  $S_n$  der Knoten des RRT mit den gewichteten Distanzen  $S_n^b$  der externen Knoten des binären Baumes überein.

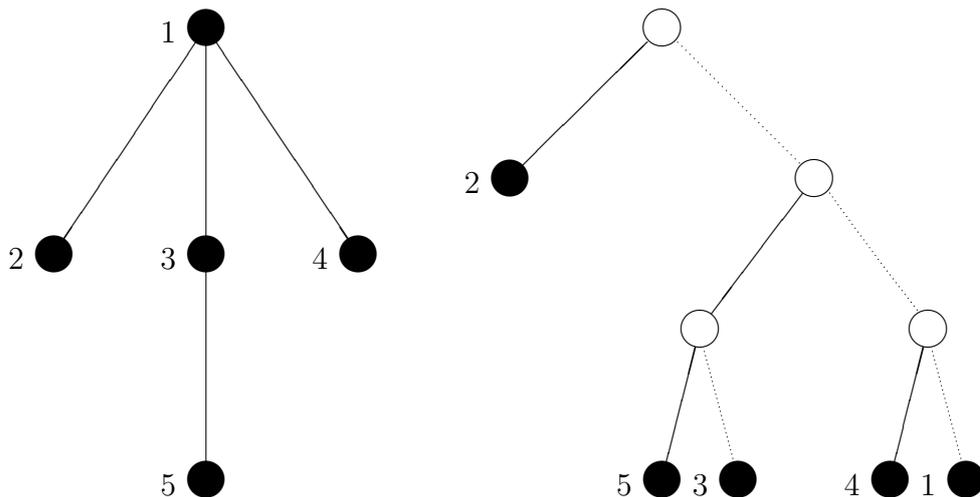


Abbildung 9: Ein zufälliger rekursiver Baum und sein korrespondierender binärer Baum. Die weißen (internen) Knoten wurden aus Konstruktionsgründen hinzugefügt. Gepunktete Linien stehen für Kanten mit  $Z = 0$  und durchgezogene Linien für Kanten mit  $Z = 1$ .

Wir haben nun unseren RRT erfolgreich mit einem binären sukzessiv aufgebauten Baum identifiziert. Zur Bestimmung der Verteilung der Lebensdauern der Individuen des binären Baumes stellen wir die folgende Überlegung an:

Zunächst interpretieren wir den binären Baum als einen Prozess, bei dem sich jedes Individuum nach Ablauf seiner Lebenszeit in zwei neue Individuen aufsplittet (Dass wir dabei das linke Kind als von der Mutter verschieden und das rechte Kind als die Mutter interpretieren, spielt für die Untersuchung der „Splittingzeiten“ keine Rolle). Da nach Voraussetzung nach  $n$  Sukzessionsschritten jedes der  $n + 1$  Individuen des RRT die gleiche Wahrscheinlichkeit  $\frac{1}{n+1}$  besitzt, als nächstes ein Kind zu gebären, gilt, dass die Restlebensdauern der Individuen des korrespondierenden binären Baumes unmittelbar nach der  $n$ -ten Splittingzeit dieselbe Verteilung besitzen. Die Verteilungen der Lebensdauern der Individuen müssen also identisch und gedächtnislos sein. Aus diesem Grund ist die Annahme einer  $Exp(\lambda)$ -Verteilung für das erste Kantengewicht  $E$  des korrespondierenden binären Baumes gerechtfertigt. Analog zu den Ausführungen im Beweis von Satz 6.1 erhält man  $\lambda = 1$  und damit  $E \sim Exp(1)$  (siehe Ausführungen in [Dev1], [Dev2] und [Dev5]).

Den zugehörigen zeitstetigen Markov-Sprungprozess  $(X_t)_{t \geq 0}$  (bei dem  $X_t$  für  $t \geq 0$  die Anzahl der Individuen zum Zeitpunkt  $t$  angibt) eines solchen binären Baumes, dessen Individuen sich nach einer  $Exp(1)$ -verteilten Lebensdauer in zwei neue Individuen aufsplitten, nennt man auch einen *Yule Prozess* (Nähere Informationen zu den Eigenschaften des Yule-Prozesses findet man in [AuN], Kapitel 3. Man beachte, dass hier im Gegensatz zum RBST die Wahl der  $Exp(1)$ -Verteilung für das erste Kantengewicht nicht nur asymptotisch für den unendlich fortgesetzten RRT, sondern auch für dessen endliche Teilbäume der ersten  $n$  Knoten gerechtfertigt ist.).

Nach den bisher angestellten Überlegungen ist ein binärer zweifach gewichteter Verzweigungsprozess mit den Gewichtsfunktionen  $E \sim Exp(1)$  und  $Z \sim B(1, \frac{1}{2})$  als Modell für einen unendlich fortgesetzten RRT  $\mathcal{T}'_\infty$  geeignet.

Begründung für  $Z \sim B(1, \frac{1}{2})$ : Zur Abschätzung der gewichteten Höhe interessiert uns (wie wir im Beweis von Satz 5.6 gesehen haben) einzig und allein die Zufallsgröße

$$H_{m_n} = \max\{D_v \mid v \in \mathcal{K}_{m_n}\}.$$

Wie man leicht einsieht, nähert sich die gewichtete Höhe  $H_{m_n}$  mit zugrunde liegender Gewichtsfunktion  $Z \sim B(1, \frac{1}{2})$  für  $m_n \rightarrow \infty$  asymptotisch der „gewichteten“ Höhe an, bei der wir statt dessen unterstellen, dass nach links gerichtete Kanten das Gewicht 1 und nach rechts gerichtete Kanten das Gewicht 0 tragen. Denn wenn wir einen beliebigen unendlichen Kantenpfad in  $\mathcal{T}'_\infty$  auswählen, so ist das Verhältnis zwischen seinen nach

links und seinen nach rechts gerichteten Kanten mit Wahrscheinlichkeit 1 ausgeglichen. Das gilt nach dem starken Gesetz der großen Zahlen, da sich die Auswahl eines zufälligen Kantenpfades im unendlich fortgesetzten binären Baum mit Hilfe einer Folge von stochastisch unabhängigen und identisch  $B(1, \frac{1}{2})$ -verteilten Zufallsgrößen beschreiben lässt. Auf jeder Ebene des Baumes wird entschieden, ob der Weg nach links oder nach rechts fortgesetzt wird, und beide Alternativen haben die Wahrscheinlichkeit  $\frac{1}{2}$ .

Nach Satz 5.6 gilt

$$\frac{H_n^*}{\log n} \xrightarrow{n \rightarrow \infty} c,$$

wobei  $c = \frac{\alpha_0}{\rho_0}$  mit  $(\rho_0, \alpha_0)$  dem eindeutig bestimmten Punkt in  $\mathcal{C}_{Z,E}$ , so dass

$\frac{\alpha_0}{\rho_0} = \max_{(\rho, \alpha) \in \mathcal{C}_{Z,E}} \frac{\alpha}{\rho}$ . Wir müssen also  $\frac{\alpha}{\rho}$  auf der Kurve

$$\mathcal{C}_{Z,E} = \{(\rho, \alpha) \mid \Lambda_Z^*(\alpha) + \Lambda_E^*(\rho) = \log 2, \rho \leq 1 = \mathbb{E}(E), \alpha \geq \frac{1}{2} = \mathbb{E}(Z)\}$$

maximieren. Im Beweis von Satz 6.1 haben wir gesehen, dass  $\Lambda_E^*(\rho) = \rho - 1 - \log \rho$  (für  $\rho > 0$ ) gilt. Wir bestimmen zunächst die Cramér-Funktion  $\Lambda_Z^*$  einer  $B(1, \frac{1}{2})$ -verteilten Zufallsgröße  $Z$ :

Für die momenterzeugende Funktion der Zufallsgröße  $Z$  ergibt sich

$$M_Z(\lambda) = \mathbb{E}(e^{\lambda Z}) = \frac{1}{2}(e^0 + e^\lambda) = \frac{1}{2}(1 + e^\lambda).$$

Daher ist ihre kumulantenerzeugende Funktion

$$\Lambda_Z(\lambda) = -\log 2 + \log(1 + e^\lambda),$$

und ihre Cramér-Funktion

$$\Lambda_Z^*(\alpha) = \sup_{\lambda \in \mathbb{R}} \{\lambda \alpha - \Lambda_Z(\lambda)\} = \sup_{\lambda \in \mathbb{R}} \{\lambda \alpha + \log 2 - \log(1 + e^\lambda)\}.$$

Dabei wird das Supremum nach Lemma 1.14 in  $\lambda = \log \alpha - \log(1 - \alpha)$  angenommen, denn es gilt

$$\Lambda_Z'(\lambda) = \frac{e^\lambda}{1 + e^\lambda} = \alpha \Leftrightarrow \lambda = \log \alpha - \log(1 - \alpha).$$

Daher haben wir

$$\Lambda_Z^*(\alpha) = \alpha \log \alpha + (1 - \alpha) \log(1 - \alpha) + \log 2$$

und ferner

$$\mathcal{C}_{Z,E} = \{(\rho, \alpha) \mid \alpha \log \alpha + (1 - \alpha) \log(1 - \alpha) + \rho - \log \rho = 1, \rho \in (0; 1), \alpha \geq \frac{1}{2}\}.$$

## 6 Anwendungen

Die Steigung der Kurve  $\mathcal{C}_{Z,E}$  ist

$$\left. \frac{df}{d\alpha} \right|_{\alpha} : \left. \frac{dg}{d\rho} \right|_{\rho} = \frac{\log \alpha - \log(1 - \alpha)}{\frac{1}{\rho} - 1},$$

wobei wir  $f(\alpha) \stackrel{\text{def}}{=} \alpha \log \alpha + (1 - \alpha) \log(1 - \alpha)$  und  $g(\rho) \stackrel{\text{def}}{=} \log \rho - \rho + 1$  definiert haben. Mit Hilfe der in Abschnitt 3.1 gewonnenen geometrischen Interpretation von  $\mathcal{C}_{Z,E}$  erhalten wir für  $(\rho_0, \alpha_0)$  (vgl. Abbildung 4)

$$\begin{aligned} \left( \left. \frac{df}{d\alpha} \right|_{\alpha_0} : \left. \frac{dg}{d\rho} \right|_{\rho_0} \right) \alpha_0 = \rho_0 & \quad ( \Leftrightarrow \alpha_0 \log \alpha_0 - \alpha_0 \log(1 - \alpha_0) = 1 - \rho_0 ) \\ & \quad ( \Leftrightarrow \rho_0 = 1 - \alpha_0 \log \alpha_0 + \alpha_0 \log(1 - \alpha_0) ). \end{aligned}$$

Einsetzen in die Gleichung für  $\mathcal{C}_{Z,E}$  liefert

$$\begin{aligned} 1 - \rho_0 + \log(1 - \alpha_0) = \log \rho_0 + 1 - \rho_0 & \Rightarrow \log(1 - \alpha_0) = \log \rho_0 \\ & \Rightarrow 1 - \alpha_0 = \rho_0. \end{aligned}$$

Dabei haben wir benutzt, dass die Logarithmusfunktion streng monoton wachsend ist. Erneutes Einsetzen in die Gleichung für  $\mathcal{C}_{Z,E}$  liefert schließlich

$$\begin{aligned} \alpha_0 = \alpha_0(\log \alpha_0 - \log(1 - \alpha_0)) & \Rightarrow 1 = \log \alpha_0 - \log(1 - \alpha_0) \\ & \Rightarrow e = \frac{\alpha_0}{1 - \alpha_0} = \frac{\alpha_0}{\rho_0} = c \end{aligned}$$

und die Behauptung des Satzes. □

## 6.3 Der random lopsided tree

Die Informationstheorie beschäftigt sich mit der Erschaffung von optimalen Codes im Hinblick auf die unterschiedlichsten Ansprüche. Vorsilbenfreie Codes verdienen dabei ein besonders großes Interesse, da man diese auf direktem Wege decodieren kann, indem man einem bestimmten Pfad durch einen stochastischen Baum folgt. Ein Knoten  $u$  eines solchen Baumes repräsentiert ein Symbol  $p$  und die Nachkommen des Knotens repräsentieren die Wörter, welche man durch Anhängen weiterer Symbole an  $p$  erhalten kann. In digitalen Anwendungen werden Symbole mit Bits und Bytes codiert und in der Regel erfordert jedes Symbol die gleiche Rechenzeit. Wir können uns die erforderlichen Rechenzeiten der einzelnen Symbole als die Längen der Kanten des Baumes vorstellen. In diesem Fall haben also alle Kanten die gleiche Länge. Es gibt jedoch auch Codes, bei denen die verschiedenen Symbole stark voneinander abweichende Rechenzeiten benötigen. Diese Codes nennt man *Varn Codes* (für eine tiefer gehende Einführung vergleiche die Ausführungen in [Var]) und führen zu Bäumen, deren Kanten unterschiedliche Längen besitzen - so genannten *random lopsided trees* (Abkürzung RLT) (Die deutsche Übersetzung „zufälliger schiefer Baum“ erscheint hier nicht adäquat. Daher verwenden wir den englischen Namen). Beachte dazu die Ausführungen in [CuG] und in [KuR].

Seien  $c_1 \leq c_2 \leq \dots \leq c_b$  feste, positive reelle Zahlen, wobei die  $(c_i)_{i=1, \dots, b}$  nicht alle gleich groß sind. Einen  $b$ -adischen Baum nennt man lopsided, wenn für jeden seiner Knoten und  $1 \leq i \leq b$  die Kante zu dessen  $i$ -tem Kind die Länge  $c_i$  besitzt. Wir definieren einen RLT wie folgt:

Wir starten mit einem  $b$ -adischen Baum, dessen Wurzel sein einziger interner Knoten ist. Schritt für Schritt wählen wir einen beliebigen externen Knoten mit Hilfe einer Laplace-Verteilung auf der Menge der externen Knoten aus und ersetzen ihn durch einen internen Knoten. Die  $b$  Kanten, welche von jedem der internen Knoten ausgehen (sofern die jeweiligen Nachkommen in der nächsten Generation schon besetzt sind) besitzen die Gewichte  $c_1, c_2, \dots, c_b$ . Für unser Modell wählen wir einen  $b$ -adischen zweifach gewichteten Verzweigungsprozess  $\mathcal{T}_\infty$  mit Gewichtsfunktionen  $E \sim Exp(1)$  (mit demselben Argument wie für den RRT, da hier derselbe Auswahlmechanismus vorliegt) und  $Z$  Laplace-verteilt auf  $\{c_1, c_2, \dots, c_b\}$  (mit demselben Argument wie für  $Z \sim B(1, \frac{1}{2})$  im Beweis von Satz 6.2). Dann gilt der folgende Satz:

**Satz 6.3 (Gewichtete Höhe des random lopsided tree)** *Für die gewichtete Höhe  $H_n^*$  eines  $b$ -adischen random lopsided tree mit  $n$  Knoten und Kantenlängen  $\{c_1, c_2, \dots, c_b\}$  gilt*

$$\frac{H_n^*}{\log n} \xrightarrow{n \rightarrow \infty} \frac{c}{b-1}$$

## 6 Anwendungen

nach Wahrscheinlichkeit. Dabei ist  $c$  das eindeutig bestimmte Maximum von  $\frac{\alpha}{\rho}$  unter der Bedingung

$$\alpha\lambda(\alpha) - \log\left(\sum_{i=1}^b c_i e^{\lambda(\alpha)c_i}\right) + \rho - 1 - \log\rho = 0,$$

wobei  $\lambda(\alpha)$  eindeutig bestimmt ist durch

$$\sum_{i=1}^b (\alpha - c_i) e^{\lambda(\alpha)c_i} = 0. \quad (6.1)$$

BEWEIS: Nach obigen Vorüberlegungen können wir wie in den Beweisen der Sätze 6.1 und 6.2 das Ergebnis von Satz 5.6 anwenden, wobei wir hier den Spezialfall  $X = b-1$  fast sicher,  $Z$  Laplace-verteilt auf  $\{c_1, c_2, \dots, c_b\}$  und  $E \sim Exp(1)$  vorliegen haben. Daher gilt

$$\Lambda_E^*(\rho) = \rho - 1 - \log\rho$$

(für  $\rho > 0$ ) und

$$\Lambda_Z(\lambda) = \log\mathbb{E}(e^{\lambda Z}) = \log\left(\sum_{i=1}^b e^{\lambda c_i}\right) - \log b.$$

Nach Lemma 1.14 wird das Supremum in  $\Lambda_Z^*(\alpha) = \sup_{\lambda \in \mathbb{R}} \{\lambda\alpha - \Lambda_Z(\lambda)\}$  angenommen für ein  $\lambda(\alpha) \in \mathbb{R}$  mit

$$\alpha = \Lambda'_Z(\lambda(\alpha)) = \frac{\sum_{i=1}^b c_i e^{\lambda(\alpha)c_i}}{\sum_{i=1}^b e^{\lambda(\alpha)c_i}} \quad (\Leftrightarrow (6.1)),$$

und dieses  $\lambda(\alpha)$  ist eindeutig bestimmt, da die  $(c_i)_{i=1, \dots, b}$  nach Voraussetzung nicht alle gleich groß sind und somit  $Z$  nicht fast sicher konstant ist (vgl. Lemma 1.6 (Strikte Konvexität der kumulantenerzeugenden Funktion)). Dann gilt

$$\mathcal{C}_{Z,E} = \{(\rho, \alpha) \mid \alpha\lambda(\alpha) - \log\left(\sum_{i=1}^b c_i e^{\lambda(\alpha)c_i}\right) + \rho - 1 - \log\rho = 0, \rho \in (0; 1), \alpha \geq \mathbb{E}(Z)\}$$

und die Behauptung folgt unmittelbar aus Satz 5.6. □

## 6.4 Weitere Beispiele

Nachfolgend stellen wir noch einige weitere, einfachere Anwendungen unserer Ergebnisse vor. Meistens genügt es dabei, die Verteilung der Gewichtsfunktion  $Z$  entsprechend den vorgestellten Spezialfällen anzupassen und daraufhin die Ergebnisse der vorangegangenen Kapitel zu nutzen.

### 1. Richtungswechsel in zufälligen binären Suchbäumen

Als erstens beschäftigen wir uns mit Richtungswechseln in zufälligen binären Suchbäumen. Sei  $\mathcal{T}'_\infty$  ein unendlich fortgesetzter RBST und für einen seiner Knoten  $v$  sei  $\pi(v)$  die Menge der Kanten des Astes von der Wurzel  $\emptyset$  bis zum Knoten  $v$  (siehe Kapitel 2). Dann setzen wir  $D(\pi(v))$  für die *Anzahl der Richtungswechsel des Pfades der Kanten  $\pi(v)$* . Codieren wir die nach links gerichteten Kanten mit 1 und entsprechend die nach rechts gerichteten Kanten mit 0, so gilt beispielsweise  $D(\pi(v)) = 5$  für einen durch 0100101 codierten Kantenpfad  $\pi(v)$ . Gezählt werden also die Stellen 01 und 10 innerhalb der Codesequenz. Wir interessieren uns insbesondere für den *Maximalwert*

$$D_{\mathcal{T}'_n} = \max\{D(\pi(v)) \mid v \text{ Knoten von } \mathcal{T}'_n\}$$

der *Richtungswechsel unter allen Pfaden des Baumes  $\mathcal{T}'_n$* . Aus Konstruktionsgründen (vgl. Abschnitt 6.1) gleichen sich die Richtungswechsel nach links und die Richtungswechsel nach rechts bei der Auswahl eines zufälligen Kantenpfades im unendlich fortgesetzten RBST  $\mathcal{T}'_\infty$  mit Wahrscheinlichkeit 1 aus (siehe auch Beweis von Satz 6.2). Ein Richtungswechsel nach links besitzt also asymptotisch die gleiche Wahrscheinlichkeit wie ein Richtungswechsel nach rechts. Das bedeutet, ein passendes Modell zur Beschreibung der Richtungswechsel innerhalb eines RBST ist ein zweifach gewichteter Verzweigungsprozess mit Gewichtsfunktionen  $Z \sim B(1, \frac{1}{2})$  und  $E \sim Exp(1)$ . Daher verhält sich die maximale Anzahl der Richtungswechsel innerhalb eines RBST asymptotisch wie die gewichtete Höhe eines RRT (vgl. Modell in Abschnitt 6.2). Mit Hilfe von Satz 6.2 erhalten wir deshalb das folgende Resultat:

**Satz 6.4 (Richtungswechsel in zufälligen binären Suchbäumen)** Für den Maximalwert  $D_{\mathcal{T}'_n}$  der Richtungswechsel unter allen Pfaden eines zufälligen binären Suchbaumes  $\mathcal{T}'_n$  mit  $n$  Knoten gilt

$$\frac{D_{\mathcal{T}'_n}}{\log n} \xrightarrow{n \rightarrow \infty} e$$

nach Wahrscheinlichkeit.

## 2. Elemente mit zwei Lebensdauern

Wir stellen uns einen binären Baum vor, dessen Kanten zwei stochastisch unabhängige,  $Exp(1)$ -verteilte Lebensdauern  $Z$  und  $E$  besitzen. Seien  $D_v$  und  $G_v$  wie in Kapitel 2 die Alter des Knotens  $v$  in Bezug auf die Gewichte  $Z$  und  $E$  und sei wieder  $\mathcal{T}_n$  der zufällige Teilbaum der Knoten  $v \in \mathcal{K}$  mit  $G_v \leq n$ . Interessant für uns ist nun der maximale Wert  $H_n$  des Alters  $D_v$  unter den Knoten  $v \in \mathcal{K}_n$ . Da die Gewichtsfunktionen  $Z$  und  $E$  den Erwartungswert 1 und zudem dieselbe Cramér-Funktion  $\Lambda_Z^*(x) = \Lambda_E^*(x) = x - 1 - \log x$  (für  $x > 0$ ) besitzen, folgt mit Hilfe von Satz 3.6:

**Satz 6.5 (Elemente mit zwei Lebensdauern)** Sei  $\mathcal{T}_\infty$  ein binärer zweifach gewichteter Verzweigungsprozess, dessen Kanten zwei voneinander unabhängige,  $Exp(1)$ -verteilte Lebensdauern  $Z$  und  $E$  besitzen. Dann gilt für das Maximum  $H_n$  der Alter der Knoten  $v \in \mathcal{K}_n$  bezüglich der Gewichtsfunktion  $Z$

$$\frac{H_n}{n} \xrightarrow{n \rightarrow \infty} c$$

nach Wahrscheinlichkeit. Dabei ist  $c \approx 5,82840157$  das Maximum von  $\frac{\alpha}{\rho}$  auf der Kurve

$$\mathcal{C}_{Z,E} = \{(\rho, \alpha) \mid \alpha - 1 - \log \alpha + \rho - 1 - \log \rho = \log 2, 0 < \rho \leq 1 \leq \alpha\}.$$

Gemessen an ihrem ersten Alter  $G_v$  existieren also Knoten  $v$ , deren zweites Alter  $D_v$  nahezu 6 mal größer ist.

## 3. Zufällige $k$ -Färbung von Kanten in einem zufälligen binären Suchbaum

Wir stellen uns einen zufälligen binären Suchbaum vor, dessen Kanten wir zufällig mit jeweils einer von  $k$  zur Verfügung stehenden Farben einfärben. Dabei habe für eine Kante  $e$  jede der  $k$  Farben die gleiche Auftrittswahrscheinlichkeit  $\frac{1}{k}$ . Unser Interesse liegt nun auf dem asymptotischen Verhalten der maximalen Anzahl gleicher Farben auf dem Pfad der

Kanten  $\pi(v)$  für einen Knoten  $v \in \mathcal{K}_n$  (also mit  $G_v \leq n$ ). Aus Symmetriegründen können wir gleich nach dem asymptotischen Verhalten der maximalen Anzahl des Auftretens einer *bestimmten* Farbe auf jenem Pfad fragen. Um diese Frage zu beantworten, versehen wir die Kanten des Baumes mit den Gewichtsfunktionen  $E \sim \text{Exp}(1)$  und  $Z \sim B(1, \frac{1}{k})$  (da das Auftreten einer bestimmten Farbe bei einer Kante die Wahrscheinlichkeit  $\frac{1}{k}$  und der ursprüngliche RBST nach Abschnitt 6.1 die Gewichte  $E \sim \text{Exp}(1)$  und  $Z = 1$  fast sicher besitzt. In dem ursprünglichen Gewicht  $Z = 1$  fast sicher verbirgt sich keine Information, die bei einer Änderung des Gewichtes verloren gehen könnte.). Nach [DuZ] gilt für die Cramér-Funktion einer  $B(1, \frac{1}{k})$ -verteilten Zufallsgröße  $Z$

$$\Lambda_Z^*(\alpha) = \alpha \log(k\alpha) + (1 - \alpha)(\log(1 - \alpha) - \log(k - 1)) + \log k, \quad (6.2)$$

und daher gilt für die relevante Kurve

$$\begin{aligned} \mathcal{C}_{Z,E} &= \{(\rho, \alpha) \mid \alpha \log(k\alpha) + (1 - \alpha)(\log(1 - \alpha) - \log(k - 1)) + \log k + \rho - 1 - \log \rho \\ &= \log 2, 0 < \rho \leq 1, k\alpha \geq 1\}. \end{aligned}$$

Mittels Satz 5.6 folgt:

**Satz 6.6 (Zufällige  $k$ -Färbung von Kanten in einem zufälligen binären Suchbaum)**

Sei  $c_k$  die Konstante aus Satz 3.6 bezüglich eines binären, zweifach gewichteten Verzweigungsprozesses  $\mathcal{T}_\infty$  mit Gewichtsfunktionen  $Z \sim B(1, \frac{1}{k})$  und  $E \sim \text{Exp}(1)$ . Dann gilt für die maximale Anzahl  $H_n^*$  des Auftretens einer bestimmten von  $k$  Kantenfarben auf den Pfaden eines zufälligen binären Suchbaumes mit  $n$  Knoten

$$\frac{H_n^*}{\log n} \xrightarrow[n \rightarrow \infty]{} c_k$$

nach Wahrscheinlichkeit.

Nachfolgende Tabelle zeigt die Werte von  $c_k$  für  $k = 1, \dots, 10$ . Für  $k = 1$  und  $k = 2$  ergeben sich die bekannten Resultate für das asymptotische Verhalten der Höhe des RBST und des RRT (vgl. Sätze 6.1 und 6.2).

k	1	2	3	4	5
$c_k$	4,3110	2,7182	2,1206	1,7955	1,5869
k	6	7	8	9	10
$c_k$	1,4397	1,3292	1,2426	1,1725	1,1148

**Bemerkung 6.7** Selbstverständlich kann man auch allgemeiner  $Z \sim B(1, p)$  mit  $p \in (0; 1)$  anstatt  $p = \frac{1}{k}$  für ein  $k \in \mathbb{N}$  annehmen. Nach [DuZ] muss man dazu in (6.2) nur das  $k$  durch  $\frac{1}{p}$  ersetzen und kann genauso verfahren wie in Satz 6.6.

Für das Studium des asymptotischen Verhaltens der maximalen Anzahl des Auftretens einer bestimmten von  $k$  Kantenfarben bei einem *zufälligen rekursiven Baum* genügt es (wie man sehr leicht einsieht) in obigem Modell die Gewichtsfunktion  $Z \sim B(1, \frac{1}{2k}) = B(1, \frac{1}{2}) * B(1, \frac{1}{k})$  zu wählen.

#### 4. Maximale Links-Minus-Rechts-Abweichung

Für die Knoten  $v$  eines binären Baumes schreiben wir

$$D_v \stackrel{\text{def}}{=} \sum_{e \in \pi(v)} (L(e) - R(e)),$$

wobei  $L : \mathcal{E} \rightarrow \{0, 1\}$  der Indikator dafür sei, ob eine Kante  $e \in \mathcal{E}$  nach links gerichtet, und entsprechend  $R : \mathcal{E} \rightarrow \{0, 1\}$  der Indikator dafür sei, ob eine Kante  $e \in \mathcal{E}$  nach rechts gerichtet ist. Wir wollen im Folgenden die maximale Links-Minus-Rechts-Abweichung  $H_n^*$  von  $D_v$  für den Fall studieren, dass der zugrunde liegende Baum ein RBST ist. Analog zu den vorangegangenen Beispielen ist ein zweifach gewichteter Verzweigungsprozess mit Gewichtsfunktionen  $Z$  mit  $\mathbb{P}(Z = -1) = \frac{1}{2} = \mathbb{P}(Z = 1)$  und  $E \sim \text{Exp}(1)$  ein geeignetes Modell zur Beschreibung der hier vorgestellten Situation. Für die Gewichtsfunktion  $Z$  erhalten wir

$$\Lambda_Z(\lambda) = \log(e^\lambda + e^{-\lambda}) - \log 2,$$

und daher gilt für die Cramérfunktion von  $Z$

$$\Lambda_Z^*(\alpha) = \begin{cases} \frac{\alpha}{2} \log \left( \frac{1+\alpha}{1-\alpha} \right) + \log 2 - \log \left( \sqrt{\frac{1+\alpha}{1-\alpha}} + \sqrt{\frac{1-\alpha}{1+\alpha}} \right) & \text{für } \mathbb{E}(Z) = 0 \leq \alpha < 1 \\ \infty & \text{für } \alpha \geq 1, \end{cases}$$

was sich durch Anwendung von Lemma 1.14 und Nachrechnen ergibt. Mit Hilfe von Satz 5.6 folgt ferner:

**Satz 6.8 (Maximale Links-Minus-Rechts-Abweichung)** Für die maximale Links-Minus-Rechts-Abweichung  $H_n^*$  eines zufälligen binären Suchbaumes mit  $n$  Knoten gilt

$$\frac{H_n^*}{\log n} \xrightarrow{n \rightarrow \infty} c$$

nach Wahrscheinlichkeit. Dabei ist  $c = \frac{\alpha_0}{\rho_0}$  das Maximum von  $(\rho, \alpha) \mapsto \frac{\alpha}{\rho}$  in  $\mathcal{C}_{Z,E}$  mit eindeutig bestimmten  $\rho_0$  und  $\alpha_0$  und  $E, Z$  wie oben. Durch Anwendung numerischer Werkzeuge lässt sich  $c \approx 2,07345$  bestimmen.



# 7 Symbolverzeichnis

Symbol	Bedeutung
$(a, b)$	Punkt $(a, b) \in \mathbb{R}^2$ mit $a, b \in \mathbb{R}$
$(a; b)$	Intervall $(a; b) \subseteq \mathbb{R}$ mit $a, b \in \mathbb{R}$ , $a < b$
$\alpha^*$	Konstante $\geq \mathbb{E}(Z)$ mit $(\mathbb{E}(E), \alpha^*) \in \mathcal{C}_{Z,E}$
$c$	Maximum von $(\alpha, \rho) \mapsto \frac{\alpha}{\rho}$ in $\mathcal{C}_{Z,E}$
$c^*$	Maximum von $(\alpha, \rho) \mapsto \frac{\alpha}{\rho}$ in $\mathcal{C}_{Z,E}^*$ im verallgemeinerten Fall
$\mathcal{C}_{Z,E}$	$\{(\rho, \alpha) \mid \Lambda_Z^*(\alpha) + \Lambda_E^*(\rho) = \log b, \alpha \geq \mathbb{E}(Z), \rho \leq \mathbb{E}(E)\}$
$\mathcal{C}_{Z,E}^*$	$\{(\rho, \alpha) \mid \Lambda_Z^*(\alpha) + \Lambda_E^*(\rho) = \log \mu, \alpha \geq \mathbb{E}(Z), \rho \leq \mathbb{E}(E)\}$ mit $\mu$ dem Erwartungswert der Reproduktionsverteilung im verallgemeinerten Fall
$D_v$	Gewichtete Höhe des Knotens $v$
$\mathcal{D}_X$	Definitionsbereich der Kumulanten erzeugenden Funktion $\Lambda_X$ der Zufallsgröße $X$
$\mathcal{D}_X^*$	Definitionsbereich der Cramérfunktion $\Lambda_X^*$ der Zufallsgröße $X$
$\text{essinf}(X)$	Essentielles Infimum der Zufallsgröße $X$
$\text{esssup}(X)$	Essentielles Supremum der Zufallsgröße $X$
$\mathbb{E}(X)$	Erwartungswert der Zufallsgröße $X$
$\mathcal{E}$	Menge der Kanten eines zweifach gewichteten Verzweigungsprozesses $\mathcal{T}_\infty$
$E_e$ und $Z_e$	Gewichte der Kante $e \in \mathcal{E}$
$E_i(v)$ und $Z_i(v)$	Gewichte der Verbindungskante des Knotens $v$ und seinem $i$ -ten Kind $(v, i)$
$G_v$	Alter des Knotens $v$
$\gamma(\alpha, \rho)$	$\Lambda_{r,Z}^*(\alpha) + \Lambda_{l,E}^*(\rho) - \log b$
$\gamma^*(\alpha, \rho)$	$\Lambda_{r,Z}^*(\alpha) + \Lambda_{l,E}^*(\rho) - \log \mu$ mit $\mu$ dem Erwartungswert der Reproduktionsverteilung im verallgemeinerten Fall
$H_n$	Gewichtete Höhe des Baumes $\mathcal{T}_\infty$ zum Zeitpunkt $n$
$\mathcal{K}$	Menge der Knoten eines zweifach gewichteten Verzweigungsprozesses $\mathcal{T}_\infty$
$\mathcal{K}_n$	Zufällige Menge der Knoten von $\mathcal{T}_n$
$L_{n,k}$	Menge der Knoten $v$ von $\mathcal{T}_n$ mit $ v  = k$

## 7 Symbolverzeichnis

$L_{\infty,k}$	Menge der Knoten $v$ von $\mathcal{T}_{\infty}$ mit $ v  = k$
$\lambda^*$	$\sup\{\lambda \in \mathbb{R} \mid M_X(\lambda) < \infty\}$
$\Lambda_X$	Kumulante erzeugende Funktion der Zufallsgröße $X$
$\Lambda_X^*$	Fenchel-Legendre-Transformierte von $\Lambda_X$ bzw. Cramér-funktion der Zufallsgröße $X$
$\Lambda_{l,X}^*$	Linksseitige Cramérfunktion der Zufallsgröße $X$
$\Lambda_{r,X}^*$	Rechtsseitige Cramérfunktion der Zufallsgröße $X$
$M_X$	Momenterzeugende Funktion der Zufallsgröße $X$
$\mathbb{N}$	Menge der natürlichen Zahlen
$o(n)$	Symbol für eine Nullfolge (für $n \rightarrow \infty$ )
$\mathcal{O}(n)$	Symbol für eine Folge $(a_n)_{n \geq 1}$ mit $\frac{a_n}{n} \xrightarrow{n \rightarrow \infty} d \in \mathbb{R}$
$(p_n)_{n \geq 0}$	Reproduktionsverteilung eines Galton-Watson-Prozesses
$\pi(v)$	Menge der Kanten des Astes von der Wurzel $\emptyset$ bis zum Knoten $v$
$\mathbb{R}$	Menge der reellen Zahlen
$\rho^*$	Konstante $\leq \mathbb{E}(E)$ mit $(\rho^*, \mathbb{E}(Z)) \in \mathcal{C}_{Z,E}$
$(\rho_0, \alpha_0)$	Eindeutig bestimmter Punkt in $\mathcal{C}_{Z,E}$ mit $\frac{\alpha_0}{\rho_0} = c$
$T_{\infty}$	$B$ -adischer Baum
$T_{\infty}^*$	Galton-Watson-Prozess mit Reproduktionsverteilung $(p_n)_{n \geq 0}$
$\mathcal{T}_{\infty}$	( $b$ -adischer) Zweifach gewichteter Verzweigungsprozess
$\mathcal{T}_{\infty}^*$	Verallgemeinerter zweifach gewichteter Verzweigungsprozess
$\mathcal{T}_n$	Zufälliger Baum der Knoten $v$ von $\mathcal{T}_{\infty}$ , für die $G_v \leq n$ gilt
$\mathcal{T}_n^*$	Zufälliger Baum der Knoten $v$ von $\mathcal{T}_{\infty}^*$ , für die $G_v \leq n$ gilt (im verallgemeinerten Fall)
$\mathcal{T}'_n$	Sukzessiv aufgebauter Baum mit $n$ Knoten
$v = (v_1, \dots, v_n)$	Knoten der $n$ -ten Generation eines zweifach gewichteten Verzweigungsprozesses
$(v, i)$	$i$ -tes Kind des Knotens $v$
$ v $	Generation, der $v$ angehört
$v i$	Urahn von $v$ in der $i$ -ten Generation
$\emptyset$	Wurzel eines Verzweigungsprozesses, leere Menge

# Literaturverzeichnis

- [Als1] G. ALSMEYER, *Wahrscheinlichkeitstheorie*. Skripten zur Mathematischen Statistik, Nr. 30, 3. Auflage, Universität Münster (2003)
- [Als2] G. ALSMEYER, *Stochastische Prozesse, Teil 1*. Skripten zur Mathematischen Statistik, Nr. 33, Universität Münster (2002)
- [Als3] G. ALSMEYER, *Verzweigungsprozesse*. Universität Münster (2001)
- [AuN] K. B. ATHREYA UND P. E. NEY, *Branching Processes*. Springer, Berlin (1972)
- [BuD] N. BROUTIN UND L. DEVROYE, *Large Deviations for the Weighted Height of an Extended Class of Trees*. *Algorithmica* (2005)
- [CuG] V. CHOI UND M. J. GOLIN, *Lopsided trees 1: a combinatorial analysis*. *Algorithmica*, Nr. 31, Seite 240-290 (2001)
- [CuM] K. S. CRUMP UND C. J. MODE, *A general age-dependent branching process*. *Journal of Mathematical Analysis and its Applications*, Nr. 24, Seite 494-508 (1968)
- [CuT] Y. S. CHOW UND H. TEICHER, *Probability Theorie: Independence, Interchangeability, Martingales*. Springer (1978)
- [DuZ] A. DEMBO UND O. ZEITOUNI, *Large Deviation Techniques and Applications*. 2. Auflage, Springer (1998)
- [Dev1] L. DEVROYE, *A note on the height of binary search trees*. *Journal of the ACM*, Nr. 33, Seite 489-498 (1986)
- [Dev2] L. DEVROYE, *Branching processes in the analysis of the height of trees*. *Acta Informatica*, Nr. 24, Seite 277-298 (1987)
- [Dev3] L. DEVROYE, *Applications of the theory of records in the study of random trees*. *Acta Informatica*, Nr. 26, Seite 123-130 (1988)

- [Dev4] L. DEVROYE, *On the expected height of fringe balanced trees*. Acta Informatica, Nr. 30, Seite 459-466 (1993)
- [Dev5] L. DEVROYE, *Branching processes and their applications in the analysis of tree structures and tree algorithms* in *Probabilistic Methods for Algorithmic Discrete Mathematics*, Nr. 16, Springer Series on Algorithms and Combinatorics, Springer, Seite 249-314, Berlin (1998)
- [Dev6] L. DEVROYE, *Universal limit laws for depth in random trees*. SIAM Journal on Computing, Nr. 28(2), Seite 409-432 (1998)
- [Els] J. ELSTRODT, *Maß- und Integrationstheorie*. 3. erweiterte Auflage, Springer, Berlin (2002)
- [Gli] N. GLICK, *Breaking records and breaking boards*. American Mathematical Monthly, Nr. 85, Seite 2-26 (1978)
- [GuS] G. R. GRIMMETT UND D. R. STIRZAKER, *Probability and Random Processes*. 2. Auflage, Oxford University Press (2001)
- [Har] T. E. HARRIS, *The Theory of Branching Processes*. Springer, Berlin (1963)
- [Heu] H. HEUSER, *Lehrbuch der Analysis, Teil 2*. 11. Auflage, B.G.Teubner, Stuttgart, Leipzig, Wiesbaden (2000)
- [Hol] F. DEN HOLLANDER, *Large Deviations*. American Mathematical Society (2000)
- [Knu] D. E. KNUTH, *The Art of Computer Programming: Sorting and Searching*. 3. Auflage, Addison-Wesley, Reading, MA (1973)
- [Kön] W. KÖNIG, *Grosse Abweichungen, Techniken und Anwendungen*. Vorlesungsskript, Universität Leipzig (2005)
- [KuR] S. KAPOOR UND E. REINGOLD, *Optimum lopsided binary trees*. Journal of the ACM, Nr. 36(3), Seite 573-590 (1989)
- [Mah] H. MAHMOUD, *Evolution of Random Search Trees*. Wiley, New York (1992)
- [Pit] B. PITTEL, *Note on the height of recursive trees and m-ary search trees*. Random Structures and Algorithms, Nr. 5, Seite 337-347 (1994)
- [Roc] R. ROCKAFELLAR, *Convex Analysis*. Princeton University Press, Princeton, NJ (1970)

- [Sch] V. SCHMIDT, *Stochastik*. Vorlesungsskript, Universität Ulm (2004)
- [Var] B. F. VARN, *Optimal variable length codes (arbitrary symbol costs and equal code word probabilities)*. Informat. Contr., Nr. 19, Seite 289-301 (1971)

Münster, den 14. Mai 2006

Ich versichere hiermit, die vorliegende Arbeit selbständig verfasst und keine anderen Hilfsmittel als die angegebenen verwendet zu haben.

Sebastian Gebennus