



TECHNISCHE  
UNIVERSITÄT  
WIEN

DIPLOMARBEIT

# Vergleich von einfach erzeugten Bäumen und Pólya-Bäumen

ausgeführt am

Institut für  
Diskrete Mathematik und Geometrie  
TU Wien

unter der Anleitung von

**Ao.Univ.Prof. Dipl.-Ing. Dr.techn. Bernhard  
Gittenberger**

durch

**Nicolas Meyer**

Matrikelnummer: 01425287

Schwarzspanierstraße 11

1090, Wien

# Zusammenfassung

Diese Arbeit besteht darin, zwei Familien von Wurzelbäumen zu definieren und ihre entsprechenden Eigenschaften zu vergleichen. Die erste betrachtete Klasse ist die von den einfach erzeugten Bäumen (engl. simply generated trees), welche eine Funktionalgleichung der Form

$$S(z) = z\Phi(S(z)) \quad (1)$$

erfüllen. Diese wurde von Meir und Moon erstmals in [25] definiert. Die zweite Struktur ist die von den Pólya-Bäumen, die von Pólya in [30] präsentiert wurde. Es ist eigentlich leicht zu zeigen, dass Pólya-Bäume nicht einfach erzeugt sind (siehe [11]). Doch besitzen Pólya-Bäume und einfach erzeugte Bäume viele Ähnlichkeiten. Dabei gibt es folgende Verfahren, um die Charakteristika der studierten Familien zu vergleichen:

- das Lösen von Funktionalgleichungen und von Abzählproblemen, wobei man sogenannte erzeugende Funktionen betrachtet;
- die Singularitätsanalyse, um Koeffizienten von erzeugenden Funktionen asymptotisch zu berechnen, sowie
- die Anwendung der Stochastik, um Parameter zu analysieren, Wahrscheinlichkeitsmodelle aufzubauen und Grenzverteilungen zu ermitteln.

Diese Methoden werden im ersten Kapitel präsentiert. Im zweiten Teil werden diese beiden Klassen von Wurzelbäumen definiert, die Asymptotik ihrer Koeffizienten verglichen und ein Zusammenhang durch eine Funktionalgleichung hergestellt. Das dritte Kapitel ist der Analyse von verschiedenen Parametern von einfach erzeugten Bäumen und von Pólya-Bäumen gewidmet. Zum einen additive Parameter, die eine bestimmte Gleichung erfüllen und unter Nebenbedingungen zur asymptotischen Normalität führen, zum anderen Parameter, die die Form von Bäumen beschreiben, und zuletzt weitere Parameter, die andere Eigenschaften besitzen.

# Inhaltsverzeichnis

<b>1</b>	<b>Diskrete Methoden und analytische Kombinatorik</b>	<b>1</b>
1.1	Kombinatorische Strukturen und erzeugende Funktionen . . . . .	1
1.1.1	Unmarkierte Strukturen . . . . .	1
1.1.2	Markierte Strukturen . . . . .	5
1.2	Singularitätsanalyse und Asymptotik . . . . .	7
1.2.1	Analytische Funktionen und Singularitäten . . . . .	7
1.2.2	Algebraische Singularitäten . . . . .	10
1.2.3	Singularitätsanalyse . . . . .	13
1.3	Parameter, multivariate erzeugende Funktionen und Grenzverteilungen .	17
1.3.1	Parameter . . . . .	18
1.3.2	Grenzverteilungen . . . . .	20
<b>2</b>	<b>Einfach erzeugte Bäume und Pólya-Bäume</b>	<b>29</b>
2.1	Einfach erzeugte Bäume . . . . .	30
2.1.1	Grundlagen . . . . .	30
2.1.2	Singularitätsanalyse . . . . .	34
2.2	Pólya-Bäume . . . . .	38
2.2.1	Grundlagen . . . . .	38
2.2.2	Singularitätsanalyse . . . . .	42
2.3	Zusammenhang zwischen einfach erzeugten Bäumen und Pólya-Bäumen .	43
<b>3</b>	<b>Vergleiche von Parametern in einfach erzeugten Bäumen und Pólya-Bäumen</b>	<b>47</b>
3.1	Additive Parameter . . . . .	47
3.1.1	Definition und allgemeine Ergebnisse . . . . .	47
3.1.2	Anzahl der Blätter und Knoten vom Grad $k$ . . . . .	53
3.1.3	Die interne Pfadlänge . . . . .	65
3.2	Die Form von einfach erzeugten Bäumen und Pólya-Bäumen . . . . .	70
3.2.1	Das Profil . . . . .	72
3.2.2	Die Höhe . . . . .	90
3.3	Weitere Parameter . . . . .	95
3.3.1	Transversalen . . . . .	95
3.3.2	Protection number . . . . .	103
3.3.3	Der Wiener-Index . . . . .	113



# Kapitel 1

## Diskrete Methoden und analytische Kombinatorik

Wir können die *Kombinatorik* als Studie von endlichen und diskreten Strukturen (die sogenannten *kombinatorischen Strukturen oder Klassen*, siehe Definition 1.1.1) definieren (vgl. mit Definition im Vorwort von [15]). Der erste Schritt besteht darin, die exakte Anzahl der Objekte in einer gegebenen Klasse durch die *symbolische Methode* zu bestimmen. Diese wird in Abschnitt 1.1 präsentiert. Daher erhalten wir eine Folge, die *Zählfolge* (siehe Definition 1.1.1) genannt wird. Doch sind manche Abzählprobleme viel schwieriger zu lösen und die Darstellung von der entsprechenden Zählfolge ist manchmal kompliziert, oder existiert sogar gar nicht. Es ist aber möglich, diese Folge anhand von approximativen Verfahren zu erhalten. Diese Methode heißt *Singularitätsanalyse*, und wird in Abschnitt 1.2 erklärt. Es ist auch sinnvoll, kombinatorische Klassen als zufällige Strukturen zu betrachten, um Wahrscheinlichkeitsmodelle aufzubauen und Grenzverteilungen zu berechnen (siehe Unterkapitel 3.1).

### 1.1 Kombinatorische Strukturen und erzeugende Funktionen

Im Folgenden werden Resultate präsentiert, die nicht unbekannt sind. Sämtliche zitierte Begriffe kann man in [15] finden.

#### 1.1.1 Unmarkierte Strukturen

Die symbolische Methode besteht aus drei Schritten:

- eine kombinatorische Struktur mit einer Funktion, die ihrer Größe entspricht definieren (Definition 1.1.1);
- diese Klasse mithilfe von einfachen Operationen (Definitionen 1.1.7, 1.1.8, 1.1.10, 1.1.15 und 1.1.18) spezifizieren;
- eine Funktionalgleichung erhalten (Definitionen 1.1.7 und 1.1.18, Sätze 1.1.9, 1.1.12 und 1.1.16).

**Definition 1.1.1** Eine kombinatorische Struktur (oder Klasse) ist ein Paar  $\mathcal{A} = (A, w)$  bestehend aus:

- einer höchstens abzählbaren Menge  $A$ ;
- einer Funktion  $w : A \rightarrow \mathbb{N}$  (Größe), die so definiert sei, dass alle Mengen

$$w^{-1}(\{n\}) = \{\alpha \in A : w(\alpha) = n\}, \quad \forall n \in \mathbb{N}_0$$

endlich sind.

Wir erhalten daher eine gewisse Folge

$$a_n := |w^{-1}(\{n\})| \tag{1.1}$$

Diese wird Zählfolge der Struktur  $\mathcal{A}$  genannt.

**Definition 1.1.2** Sei  $\mathcal{A} = (A, w)$  eine kombinatorische Struktur und bezeichne  $(a_n)_{n \geq 0}$  ihre zugehörige Zählfolge. Dann heißt folgende Potenzreihe

$$A(z) = \sum_{\alpha \in A} z^{w(\alpha)} = \sum_{n \geq 0} a_n z^n \tag{1.2}$$

die (gewöhnliche) erzeugende Funktion von  $\mathcal{A}$ .

**Beispiel 1.1.3** Sei  $\mathcal{A} = (A, |\cdot|)$  eine kombinatorische Klasse und  $\epsilon$  bezeichne ein kombinatorisches Objekt der Größe 0. Die zugehörige erzeugende Funktion ist 1. Für ein Objekt  $\alpha \in A$  ist die erzeugende Funktion durch  $z^{|\alpha|}$  gegeben.

**Beispiel 1.1.4** Ein Objekt der Größe 1 nennt man Atom und wird mit  $\circ$  bezeichnet.

**Definition 1.1.5** Zwei kombinatorische Strukturen  $\mathcal{A}$  und  $\mathcal{B}$  mit den Zählfolgen  $(a_n)_{n \in \mathbb{N}}$  bzw.  $(b_n)_{n \in \mathbb{N}}$  heißen isomorph (i.Z.  $\mathcal{A} \cong \mathcal{B}$ ), wenn für jedes  $n \in \mathbb{N}$  gilt  $a_n = b_n$ .

**Bemerkung 1.1.6** Seien  $\mathcal{A}$  und  $\mathcal{B}$  zwei kombinatorische Klassen mit erzeugenden Funktionen  $A(z)$  bzw.  $B(z)$ .  $\mathcal{A} \cong \mathcal{B}$  impliziert klarerweise  $A(z) = B(z)$  für alle  $z$ .

**Definition 1.1.7** Seien  $\mathcal{C}, \mathcal{A}^{(1)}, \mathcal{A}^{(2)}, \dots, \mathcal{A}^{(m)}$  kombinatorische Strukturen mit zugehörigen Zählfolgen  $(c_n)_{n \geq 0}, (a_n^{(1)})_{n \geq 0}, (a_n^{(2)})_{n \geq 0}, \dots, (a_n^{(m)})_{n \geq 0}$ , und gelte

$$\mathcal{C} = \phi(\mathcal{A}^{(1)}, \mathcal{A}^{(2)}, \dots, \mathcal{A}^{(m)}).$$

Dann ist  $\phi$  eine zulässige kombinatorische Konstruktion genau dann, wenn  $(c_n)_{n \geq 0}$  nur von  $(a_n^{(1)})_{n \geq 0}, (a_n^{(2)})_{n \geq 0}, \dots, (a_n^{(m)})_{n \geq 0}$  abhängt.

Setze  $C(z) := \sum_{n \geq 0} c_n z^n, A_1(z) := \sum_{n \geq 0} a_n^{(1)} z^n, A_2(z) := \sum_{n \geq 0} a_n^{(2)} z^n, \dots, A_m(z) := \sum_{n \geq 0} a_n^{(m)} z^n$  als EF von  $\mathcal{C}, \mathcal{A}^{(1)}, \mathcal{A}^{(2)}, \dots, \mathcal{A}^{(m)}$ .

Das ergibt einen Operator  $\psi$ , sodass

$$C(z) = \psi(A_1(z), A_2(z), \dots, A_m(z)). \tag{1.3}$$

Nun betrachten wir folgende kombinatorische Konstruktionen:

**Definition 1.1.8** Seien  $\mathcal{A} = (A, |\cdot|_A)$ ,  $\mathcal{B} = (B, |\cdot|_B)$  und  $\mathcal{C} = (C, |\cdot|_C)$  drei kombinatorische Konstruktionen, wobei  $\mathcal{C} = \phi(\mathcal{A}, \mathcal{B})$ :

- Gilt  $A \cap B = \emptyset$ , dann bezeichnet  $C = A \cup B = A + B$  die kombinatorische Summe von  $A$  und  $B$ . Für ein  $x \in C$  gilt dann  $|x|_C = |x|_A$  für  $x \in A$  oder  $|x|_C = |x|_B$  für  $x \in B$ .
- Das Produkt von  $A$  und  $B$  ist definiert durch  $C = A \times B = \{(\alpha, \beta) : \alpha \in A, \beta \in B\}$ . Dabei gilt  $|(\alpha, \beta)|_C = |\alpha|_A + |\beta|_B$  ( $\alpha \in A, \beta \in B$ ).

**Satz 1.1.9** Seien  $\mathcal{A} = (A, |\cdot|_A)$ ,  $\mathcal{B} = (B, |\cdot|_B)$  und  $\mathcal{C} = (C, |\cdot|_C)$  drei kombinatorische Konstruktionen mit erzeugenden Funktionen  $A(z) = \sum_{n \geq 0} a_n z^n$ ,  $B(z) = \sum_{n \geq 0} b_n z^n$  bzw.  $C(z) = \sum_{n \geq 0} c_n z^n$ . Es gelte dazu  $\mathcal{C} = \phi(\mathcal{A}, \mathcal{B})$ .

- Bezeichnet  $\mathcal{C}$  eine kombinatorische Summe von  $\mathcal{A}$  und  $\mathcal{B}$ , so gilt  $c_n = a_n + b_n$  ( $n \in \mathbb{N}$ ) und daher  $C(z) = A(z) + B(z)$ ;
- Bezeichnet  $\mathcal{C}$  ein Produkt von  $\mathcal{A}$  und  $\mathcal{B}$ , so gilt  $c_n = \sum_{k=0}^n a_k b_{n-k}$  ( $n \in \mathbb{N}$ ) und daher  $C(z) = A(z)B(z)$ .

Nun betrachten wir Operatoren auf einer kombinatorischen Struktur:

**Definition 1.1.10** Seien  $\mathcal{A} = (A, |\cdot|)$  und  $\mathcal{C} = (C, |\cdot|)$  zwei kombinatorische Strukturen mit  $a_0 = 0$  und  $\mathcal{C} = \phi(\mathcal{A})$ . Daher bezeichnet  $\mathcal{C}$ :

- die Folgen von Objekten aus  $\mathcal{A}$ , falls  $C = F(A) = \{\epsilon\} + A + A \times A + \dots = \{(\alpha_1, \alpha_2, \dots, \alpha_k), \alpha_i \in A, k \in \mathbb{N}\}$ . Für die Größe gilt daher  $|(\alpha_1, \alpha_2, \dots, \alpha_k)| = \sum_{i=1}^k |\alpha_i|$ ;
- die Zyklen von Objekten aus  $\mathcal{A}$ , falls  $C = Z(A) = F(A)/\approx$ , wobei für je zwei Tupel  $(\alpha_1, \alpha_2, \dots, \alpha_k), (\beta_1, \beta_2, \dots, \beta_k) \in A^*$   $(\alpha_1, \alpha_2, \dots, \alpha_k) \approx (\beta_1, \beta_2, \dots, \beta_k)$  gilt genau dann, wenn  $(\alpha_1, \alpha_2, \dots, \alpha_k)$  und  $(\beta_1, \beta_2, \dots, \beta_k)$  sich nur durch zyklische Vertauschung der Einträge unterscheiden. Für die Größe gilt ebenfalls  $|(\alpha_1, \alpha_2, \dots, \alpha_k)| = \sum_{i=1}^k |\alpha_i|$ ;
- die Mengen von Objekten aus  $\mathcal{A}$ , falls  $\mathcal{C} = S(\mathcal{A})$ . Für eine Teilmenge  $\{\alpha_1, \alpha_2, \dots, \alpha_k\} \subseteq A$  gilt  $|\{\alpha_1, \alpha_2, \dots, \alpha_k\}| = \sum_{i=1}^k |\alpha_i|$ ;
- die Multimengen von Objekten aus  $\mathcal{A}$ , falls  $\mathcal{C} = M(\mathcal{A})$ . Für eine Teilmenge  $\{\alpha_1^{j_1}, \alpha_2^{j_2}, \dots, \alpha_k^{j_k}\} \subseteq A$  gilt  $|\{\alpha_1^{j_1}, \alpha_2^{j_2}, \dots, \alpha_k^{j_k}\}| = \sum_{i=1}^k j_i |\alpha_i|$ .  $j_1, j_2, \dots, j_k$  bezeichnen die Vielfachheiten der Elementen  $\alpha_1, \alpha_2, \dots, \alpha_k$ .

Man kann die Mengen bzw. Multimengen vermittels Isomorphie alternativ darstellen:

**Bemerkung 1.1.11** Seien  $\mathcal{A} = (A, |\cdot|)$  und  $\mathcal{C} = (C, |\cdot|)$  kombinatorische Strukturen mit  $a_0 = 0$  und  $\mathcal{C} = \phi(\mathcal{A})$ . Dann gilt:

- $\mathcal{C} = S(\mathcal{A}) \cong \prod_{\alpha \in A} (\{\epsilon\} \cup \{\alpha\})$
- $\mathcal{C} = M(\mathcal{A}) \cong \prod_{\alpha \in A} F(\{\alpha\})$

**Satz 1.1.12** Seien  $\mathcal{A} = (A, |\cdot|)$  und  $\mathcal{C} = (C, |\cdot|)$  kombinatorische Konstruktionen mit erzeugenden Funktionen  $A(z) = \sum_{n \geq 0} a_n z^n$  bzw.  $C(z) = \sum_{n \geq 0} c_n z^n$ . Setze  $\mathcal{C} = \phi(\mathcal{A})$ . Dann gilt für

- Folgen:  $C(z) = 1 + A(z) + A(z)^2 + \dots = \frac{1}{1 - A(z)}$ ;
- Zyklen:  $C(z) = \sum_{k \geq 1} \frac{\varphi(k)}{k} \log \left( \frac{1}{1 - A(z^k)} \right)$  mit  $\varphi(k) = |\{j : 1 \leq j < k : \text{ggT}(j, k) = 1\}|$  (Eulersche Phi-Funktion);
- Mengen:  $C(z) = \prod_{\alpha \in A} (1 + z^{|\alpha|}) = \prod_{n \geq 1} (1 + z^n)^{a_n}$ ;
- Multimengen:  $C(z) = \prod_{\alpha \in A} \frac{1}{1 - z^{|\alpha|}} = \prod_{n \geq 1} \left( \frac{1}{1 - z^n} \right)^{a_n}$ ;

**Bemerkung 1.1.13** Bezeichne mit  $\mathcal{C}$  die Mengen von Objekten aus  $\mathcal{A}$ , so gilt folgende alternative Darstellung für die erzeugende Funktion von  $\mathcal{C}$

$$C(z) = \exp \left( \sum_{k \geq 1} \frac{(-1)^{k-1}}{k} A(z^k) \right) \quad (1.4)$$

Dies zeigt man wie folgt:

$$\begin{aligned} C(z) &= \prod_{n \geq 1} (1 + z^n)^{a_n} \\ &= \exp \left( \log \prod_{n \geq 1} (1 + z^n)^{a_n} \right) \\ &= \exp \left( \sum_{n \geq 1} a_n \log(1 + z^n) \right) \\ &= \exp \left( \sum_{n \geq 1} a_n \sum_{k \geq 1} \frac{(-1)^{k-1}}{k} z^{nk} \right) \\ &= \exp \left( \sum_{k \geq 1} \frac{(-1)^{k-1}}{k} \sum_{n \geq 1} a_n (z^k)^n \right) \\ &= \exp \left( \sum_{k \geq 1} \frac{(-1)^{k-1}}{k} A(z^k) \right) \end{aligned}$$



**Bemerkung 1.1.14** Seien  $\mathcal{C}$  die Multimengen von Objekten aus  $\mathcal{A}$  und bezeichne mit  $C(z)$  die erzeugende Funktion von  $\mathcal{C}$ , so haben wir

$$C(z) = \exp \left( \sum_{k \geq 1} \frac{A(z^k)}{k} \right) \quad (1.5)$$

Der Beweis geht ganz analog wie in der vorigen Bemerkung.

Wir können zu den obigen Konstruktionen *Restriktionen* einsetzen:

**Definition 1.1.15** Sei  $K$  eine zulässige Konstruktion, die Folgen, Zyklen, Mengen oder Multimengen von Objekten einer kombinatorischen Struktur entspricht. Dann ist die Restriktion  $K_k$  ( $K_{\leq k}$  bzw.  $K_{\geq k}$ ) eine Konstruktion mit genau/höchstens bzw. mindestens  $k$  Elementen.

**Satz 1.1.16** Seien  $\mathcal{A}, \mathcal{C}$  zwei kombinatorische Klassen mit  $a_0 = 0$  und erzeugenden Funktion  $A(z)$  bzw.  $C(z)$  und  $k \in \mathbb{N}$ . Dann gilt:

- Ist  $\mathcal{C} = F_k(\mathcal{A})$ , so gilt  $C(z) = A(z)^k$ ;
- Ist  $\mathcal{C} = Z_k(\mathcal{A})$ , so gilt  $C(z) = [u^k] \sum_{l \geq 1} \frac{\varphi(l)}{l} \log \left( \frac{1}{1 - u^l A(z^l)} \right)$ ;
- Ist  $\mathcal{C} = S_k(\mathcal{A})$ , so gilt  $C(z) = [u^k] \exp \left( \sum_{l \geq 1} (-1)^{l-1} \frac{u^l A(z^l)}{l} \right)$ ;
- Ist  $\mathcal{C} = M_k(\mathcal{A})$ , so gilt  $C(z) = [u^k] \exp \left( \sum_{l \geq 1} \frac{u^l A(z^l)}{l} \right)$ .

*Beweis:* Siehe [15, Theorem I.3]. □

**Definition 1.1.17** Seien  $\mathcal{A} = (A, |\cdot|)$  und  $\mathcal{B} = (B, |\cdot|)$  zwei kombinatorische Klassen mit  $b_0 = 0$ . Setze  $A_n = \{\alpha \in A : |\alpha| = n\}$ . Dann ist die Komposition von  $A$  und  $B$  definiert durch  $C = A(B) = A_0 + A_1 \times B + A_2 \times B \times B + \dots$ . Sind  $A(z)$ ,  $B(z)$  und  $C(z)$  die erzeugende Funktionen von  $\mathcal{A}$ ,  $\mathcal{B}$  bzw.  $\mathcal{C}$ , so gilt  $C(z) = A(B(z))$ .

## 1.1.2 Markierte Strukturen

Wir können spezielle Klassen, sogenannte *markierte Strukturen* definieren. In diesem Fall sind die Objekte aus Atomen aufgebaut. Jedes Atom bekommt eine *Markierung* (Definition 1.1.19). Prominente Beispiele sind (zyklische) Permutationen und Urnen.

**Definition 1.1.18** Sei  $\mathcal{A} = (A, |\cdot|)$  eine markierte Klasse und  $\alpha \in A$  ein Objekt mit  $|\alpha| = n$ . Daher bekommt dieses Objekt Markierungen, das heißt die natürlichen Zahlen  $1, 2, \dots, n$ .

**Definition 1.1.19** Sei  $\mathcal{A} = (A, |\cdot|)$  eine markierte Struktur mit Zählfolge  $(a_n)_{n \in \mathbb{N}}$ . Dann ist die zugehörige exponentielle erzeugende Funktion durch

$$\hat{A}(z) = \sum_{\alpha \in A} \frac{z^{|\alpha|}}{|\alpha|!} = \sum_{n \geq 0} a_n \frac{z^n}{n!}$$

definiert.

Alternativ zu unmarkierten Klassen definieren wir folgende Konstruktionen. Die Summe funktioniert genauso wie im unmarkierten Fall. Das Produkt ist ein wenig komplizierter.

**Definition 1.1.20** Sei  $\alpha$  ein schwach markiertes Objekt, d.h. ein Objekt der Größe  $n$  mit Markierungen  $M \subset \mathbb{Z}$  wobei  $|M| = n$ . Weiters bezeichne  $\rho$  eine Ummarkierung, sodass  $\rho(\alpha)$  kanonisch markiert ist (zum Beispiel für die Folge  $(3, 7, 1, 4)$  gilt  $\rho((3, 7, 1, 4)) = (2, 4, 1, 3)$ ).

Seien nun  $\alpha, \beta$  zwei Objekte aus verschiedenen markierten Strukturen. Dann ist  $\alpha * \beta$  die Menge aller korrekt markierten Paaren  $(\alpha', \beta')$ , sodass  $\rho(\alpha') = \alpha$  und  $\rho(\beta') = \beta$ .

**Satz 1.1.21** Seien  $\mathcal{A} = (A, |\cdot|)$ ,  $\mathcal{B} = (B, |\cdot|)$  und  $\mathcal{C} = (C, |\cdot|)$  drei markierte Strukturen. Dann beschreibt

$$C = A * B = \bigcup_{\alpha \in A, \beta \in B} \alpha * \beta$$

ein Produkt von  $A$  und  $B$ . Sind  $A(z) = \sum_{n \geq 0} a_n \frac{z^n}{n!}$ ,  $B(z) = \sum_{n \geq 0} b_n \frac{z^n}{n!}$  und  $C(z) = \sum_{n \geq 0} c_n \frac{z^n}{n!}$  die exponentiell erzeugenden Funktionen von  $\mathcal{A}, \mathcal{B}$  bzw.  $\mathcal{C}$ , so folgt  $C(z) = A(z)B(z)$  und  $c_n = \sum_{k=0}^n \binom{n}{k} a_k b_{n-k}$ .  $\binom{n}{k} = \frac{n!}{k!(n-k)!}$  ist der Binomialkoeffizient.

**Definition 1.1.22** Seien  $\mathcal{A} = (A, |\cdot|)$  und  $\mathcal{C} = (C, |\cdot|)$  zwei markierte Strukturen mit  $a_0 = 0$  und  $C = \phi(\mathcal{A})$ . Dann bezeichnet  $\mathcal{C}$ :

- die Folgen von Objekten aus  $\mathcal{A}$ , falls  $C = F(\mathcal{A}) = \{\epsilon\} + A + A * A + \dots = \{(\alpha_1, \alpha_2, \dots, \alpha_k), \alpha_i \in A, k \in \mathbb{N}\}$ . Für die Größe gilt daher  $|(\alpha_1, \alpha_2, \dots, \alpha_k)| = \sum_{i=1}^k |\alpha_i|$ ;
- die Zyklen von Objekten aus  $\mathcal{A}$ , falls  $C = Z(\mathcal{A}) = A + \frac{1}{2}A * A + \frac{1}{3}A * A * A + \dots$ . Für die Größe gilt ebenfalls  $|(\alpha_1, \alpha_2, \dots, \alpha_k)| = \sum_{i=1}^k |\alpha_i|$ ;
- die Mengen von Objekten aus  $\mathcal{A}$ , falls  $C = S(\mathcal{A}) = \{\epsilon\} + A + \frac{1}{2!}A * A + \frac{1}{3!}A * A * A + \dots$ . Für eine Teilmenge  $\{\alpha_1, \alpha_2, \dots, \alpha_k\} \subseteq A$  gilt  $|\{\alpha_1, \alpha_2, \dots, \alpha_k\}| = \sum_{i=1}^k |\alpha_i|$ .

**Satz 1.1.23** Seien  $\mathcal{A} = (A, |\cdot|)$  und  $\mathcal{C} = (C, |\cdot|)$  zwei markierte Klassen mit exponentiell erzeugenden Funktionen  $A(z) = \sum_{n \geq 0} a_n \frac{z^n}{n!}$  bzw.  $C(z) = \sum_{n \geq 0} c_n \frac{z^n}{n!}$ . Setze  $C = \phi(\mathcal{A})$ . Dann gilt:

- *Folgen:*  $C(z) = 1 + A(z) + A(z)^2 + \dots = \frac{1}{1 - A(z)}$ ;
- *Zyklen:*  $C(z) = A(z) + \frac{A(z)^2}{2} + \frac{A(z)^3}{3} + \dots = \log\left(\frac{1}{1 - A(z)}\right)$
- *Mengen:*  $C(z) = C(z) = 1 + A(z) + \frac{A(z)^2}{2!} + \frac{A(z)^3}{3!} + \dots = e^{A(z)}$ .

**Beispiel 1.1.24** Als Permutationen bezeichnen wir die Folgen von Atomen. Daher ist die zugehörige exponentiell erzeugende Funktion durch  $P(z) = 1 + z + z^2 + \dots = \frac{1}{1 - z}$  gegeben, und wir haben  $p_n = n!$ .

Unter Urnen versteht man die Mengen von Atomen. Die exponentiell erzeugende Funktion lautet dann  $U(z) = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \dots = e^z$  und daher  $u_n = 1$ .

Die Kreise oder zyklischen Permutationen sind die Zyklen von Atomen. Wir haben dann  $K(z) = z + \frac{z^2}{2} + \frac{z^3}{3} + \dots = \log\left(\frac{1}{1 - z}\right)$  und  $k_n = (n - 1)!$ .

**Definition 1.1.25** Seien  $\mathcal{A} = (A, |\cdot|)$  und  $\mathcal{B} = (B, |\cdot|)$  zwei markierte Strukturen mit  $b_0 = 0$ . Sei  $A_n = \{\alpha \in A : |\alpha| = n\}$ . Dann ist die Komposition von  $A$  und  $B$  definiert durch  $C = A(B) = A_0 + A_1 \times B + A_2 \times (B * B) + A_3 \times (B * B * B) + \dots$ . Sind  $A(z)$ ,  $B(z)$  und  $C(z)$  die exponentiell erzeugenden Funktionen von  $\mathcal{A}$ ,  $\mathcal{B}$  bzw.  $\mathcal{C}$ , so gilt  $C(z) = A(B(z))$ .

## 1.2 Singularitätsanalyse und Asymptotik

Im vorigen Abschnitt haben wir Verfahren gesehen, um explizit die Zählfolge einer un- bzw. markierten kombinatorischen Struktur zu bestimmen. Doch haben diese Methoden einen Nachteil: manche explizite Darstellungen dieser Zählfolgen sind kompliziert oder existieren gar nicht. Es ist aber möglich einen asymptotischen Ausdruck für diese Folgen zu erhalten, indem man die zugehörige (exponentiell) erzeugende Funktion als *analytisches Objekt* (Definition 1.2.1, siehe auch [15, Kapitel IV]) betrachtet.

### 1.2.1 Analytische Funktionen und Singularitäten

Zuerst brauchen wir einige Begriffe der komplexen Analysis. Eine Funktion auf einer offenen Teilmenge  $G$  von  $\mathbb{C}$  heißt holomorph, wenn sie bei jedem  $z \in G$  komplex differenzierbar ist. Ein äquivalenter Begriff ist die Analytizität.

**Definition 1.2.1** Sei  $G \subseteq \mathbb{C}$  offen. Eine Funktion  $f : G \rightarrow \mathbb{R}(\mathbb{C})$  heißt analytisch bei  $z_0 \in G$ , falls in einer Umgebung von  $z_0$  die Potenzreihe

$$\sum_{n \geq 0} a_n (z - z_0)^n \tag{1.6}$$

mit Koeffizienten

$$a_n = \frac{f^{(n)}(z_0)}{n!}, \quad n \in \mathbb{N}_0 \tag{1.7}$$

gegen  $f$  konvergiert.

Ist  $f$  in jedem Punkt von  $G$  analytisch, so ist  $f$  analytisch.

Die Koeffizienten von  $f$  besitzen laut folgendem Satz eine alternative Darstellung:

**Satz 1.2.2** Sei  $f : G \rightarrow \mathbb{R}(\mathbb{C})$  ( $G \subseteq \mathbb{C}$  offen) analytisch und  $\Gamma$  eine geschlossene Kurve, die für jedes  $z_0 \in G$  im negativen Uhrzeigersinn einmal umläuft. Dann gilt:

$$f(z_0) = \frac{1}{2\pi i} \int_{\Gamma} \frac{f(z)}{z - z_0} dz \quad (1.8)$$

und (für Ableitungen)

$$f^{(n)}(z_0) = \frac{n!}{2\pi i} \int_{\Gamma} \frac{f(z)}{(z - z_0)^{n+1}} dz \quad (1.9)$$

**Bemerkung 1.2.3** Analytizität, Holomorphie und Formeln (1.8) – (1.9) sind äquivalent (siehe [37, Abschnitt 2.1]).

Aus (1.9) und (1.7) erhalten wir klarerweise eine alternative Darstellung für die Koeffizienten der Reihe aus der Formel (1.6):

$$a_n = \frac{1}{2\pi i} \int_{\Gamma} \frac{f(z)}{(z - z_0)^{n+1}} dz \quad (1.10)$$

Nun betrachten wir eine Funktion  $f$ , die nicht ganz in  $G$  analytisch ist:

**Definition 1.2.4** Sei  $G \subseteq \mathbb{C}$  offen und  $z_0 \in G$ . Ist  $f$  auf  $G \setminus \{z_0\}$  analytisch, dann ist  $z_0$  eine isolierte Singularität von  $f$ . Sei dazu  $N \in \mathbb{N}_0 \cup \{+\infty\}$  durch

$$N := \inf\{n \in \mathbb{N}_0 : \lim_{z \rightarrow z_0} (z - z_0)^{n+1} f(z) = 0\}$$

definiert. Dann ist  $z_0$  eine:

- hebbare Singularität, falls  $N = 0$ ;
- Polstelle (kurz: ein Pol) der Vielfachheit  $N$ , falls  $0 < N < +\infty$ ;
- wesentliche Singularität, falls  $N = +\infty$ ;

**Definition 1.2.5** Sei  $G \subseteq \mathbb{C}$  offen. Eine Funktion  $f$ , die auf  $G$  außer isolierten Singularitäten, die alle Pole sind, analytisch ist, heißt meromorph.

Funktionen, die bei  $z_0$  eine isolierte Singularität besitzen, können lokal in eine Reihe entwickelt werden. Diese ist nicht mehr eine Potenzreihe, sondern eine *Laurentreihe*:

**Definition 1.2.6** Sei  $G \subseteq \mathbb{C}$  offen und  $z_0 \in G$ . Ist  $f$  auf  $G \setminus \{z_0\}$  analytisch, dann gilt:

$$f(z) = \sum_{n \in \mathbb{Z}} a_n (z - z_0)^n \quad (1.11)$$

für  $z \neq z_0$ . Die Darstellung (1.11) heißt die Laurententwicklung von  $f$ .

**Definition 1.2.7** Ist  $z_0 \in G$  eine Polstelle der Ordnung  $N$  einer meromorphen Funktion  $f$ , so ist der Hauptteil von  $f$  durch

$$g(z) = \sum_{n=1}^N \frac{a_{-n}}{(z - z_0)^n} \quad (1.12)$$

definiert. Die Laurententwicklung von  $f$  um  $z_0$  lautet dann

$$f(z) = g(z) + \sum_{n \geq 0} a_n (z - z_0)^n. \quad (1.13)$$

**Bemerkung 1.2.8** Klarerweise ist  $f(z) - g(z)$  bei  $z_0$  analytisch (Nebenteil von  $f$ ).

**Definition 1.2.9** Sei  $f$  eine Abbildung, die eine isolierte Singularität bei  $z_0$  besitzt. Ist  $\sum_{n \in \mathbb{Z}} a_n (z - z_0)^n$  die entsprechende Laurentreihe von  $f$  um  $z_0$ , so heißt der Koeffizient  $a_{-1}$  das Residuum von  $f$  an der Stelle  $z_0$  ( $\text{Res}(f, z_0)$ ).

**Satz 1.2.10** Sei  $f$  analytisch auf  $G \setminus \{z_1, \dots, z_k\}$  ( $G \subseteq \mathbb{C}$  offen) und  $\gamma$  eine geschlossene Kurve, die die Punkte  $z_1, \dots, z_k$  im negativen Uhrzeigersinn einmal umläuft. Dann gilt:

$$\frac{1}{2\pi i} \int_{\gamma} f(z) dz = \sum_{i=1}^k \text{Res}(f, z_i). \quad (1.14)$$

*Beweis:* Siehe [37, Satz 3.5.3]. □

Nun verwenden wir unsere Definitionen, um Koeffizienten von (exponentiell) erzeugenden Funktionen zu bestimmen (vgl. mit [15, Abschnitt IV.1]). Wir beginnen zuerst mit einer exakten Methode.

**Satz 1.2.11 (Inversionsformel von Lagrange)** Seien  $\Phi(t)$  und  $T(z)$  analytisch um 0 mit  $z = \frac{T(z)}{\Phi(T(z))}$  und  $\Phi(0) \neq 0$ . Dann gilt

$$[z^n]T(z) = \frac{1}{n} [u^{n-1}] \Phi(u)^n. \quad (1.15)$$

Allgemeiner haben wir

$$[z^n]f(T(z)) = \frac{1}{n} [u^{n-1}] \Phi(u)^n f'(u) \quad (1.16)$$

für eine analytische Funktion  $f(t)$ .

*Beweis:* Setze  $t_n := [z^n]T(z)$ . Mithilfe der Formel (1.10) und der Substitution  $T(z) = T$ , was  $dT = T'(z)dz$  impliziert, haben wir

$$nt_n = \frac{1}{2\pi i} \int_{\Gamma} \frac{T'(z)}{z^n} dz = \frac{1}{2\pi i} \int_{\tilde{\Gamma}} \frac{dT}{\left(\frac{T}{\Phi(T)}\right)^n} = \frac{1}{2\pi i} \int_{\tilde{\Gamma}} \Phi(T)^n \frac{dT}{T^n} = [u^{n-1}] \Phi(u)^n.$$

□

Weiters haben wir:

- Gegeben:  $f(z) = \sum_{n \geq 0} a_n \frac{z^n}{(n!)}$  eine bei  $z = 0$  analytische Funktion mit Konvergenzradius  $R$ , die einer (exponentiell) erzeugenden Funktion entspricht.
- Gesucht: Folge der Koeffizienten  $(a_n)_{n \in \mathbb{N}}$  von  $f(z)$  oder eine weitere Folge  $(\widetilde{a}_n)_{n \in \mathbb{N}}$  für die gilt  $a_n \sim \widetilde{a}_n$ , also  $\lim_{n \rightarrow \infty} \frac{a_n}{\widetilde{a}_n} = 1$ .
- Idee: Finde die Singularität  $z_0$  mit  $|z_0|$  minimal (*dominante Singularität*).

Klarerweise gilt  $R = |z_0|$ , also

$$\frac{1}{|z_0|} = \limsup_{n \rightarrow \infty} |a_n|^{1/n} \quad (1.17)$$

(vgl. mit [15, Theorem IV.7]). Aus (1.17) folgt

$$\log \frac{1}{R^n} \sim \log |a_n| \quad \Leftrightarrow \quad a_n \asymp \frac{1}{R^n}. \quad (1.18)$$

Wir können aber einen genaueren Ausdruck für die  $a_n$  bestimmen:

**Definition 1.2.12** Die  $a_n$  besitzen folgende asymptotische Darstellung:

$$a_n = A^n \theta(n), \quad (1.19)$$

wobei:

- $A^n$  (exponentielles Wachstum): hängt von der Lokalisierung der dominanten Singularität ab;
- $\theta(n)$  (subexponentieller Faktor): hängt von dem Typ der Singularität ab.

**Bemerkung 1.2.13** Man sieht leicht, dass  $A = \frac{1}{R}$  und  $\limsup_{n \rightarrow \infty} |\theta(n)|^{1/n} = 1$ .

## 1.2.2 Algebraische Singularitäten

Nun betrachten wir eine Abbildung  $f$ , die für  $|z| < R$  analytisch ist. Seien  $z_0$  die dominante Singularität von  $f$ , also  $|z_0| = R$ , und  $\alpha \in \mathbb{C} \setminus \mathbb{Z}$ , sodass

$$f(z) = (z - z_0)^\alpha g(z), \quad (1.20)$$

wobei  $g$  bei  $z_0$  analytisch ist.

**Definition 1.2.14** Gilt  $g(z) = \sum_{n \geq 0} g_n (z - z_0)^n$ , so gilt für  $f$ :

$$f(z) = \sum_{n \geq 0} g_n (z - z_0)^{\alpha+n}. \quad (1.21)$$

Die Darstellung (1.21) heißt *Singuläre Entwicklung* oder *Puiseux-Reihe* von  $f$ .

Sei  $\Gamma$  die *Gammafunktion*. Für  $\operatorname{Re}(z) > 0$  gilt:

$$\Gamma(z) = \int_0^{+\infty} x^{z-1} e^{-x} dx$$

Die Gammafunktion besitzt einige interessante Eigenschaften:

**Lemma 1.2.15** *Für die Gammafunktion  $\Gamma$  gilt:*

- (i)  $\Gamma(z+1) = z\Gamma(z)$  für jedes  $z \in \mathbb{C}$  mit  $\operatorname{Re}(z) > 0$ ;
- (ii)  $\Gamma(n) = (n-1)!$  für jedes  $n \in \mathbb{N}$ ;
- (iii)  $\Gamma(1/2) = \sqrt{\pi}$ ;
- (iv)  $\Gamma(z+1) \sim \left(\frac{z}{e}\right)^z \sqrt{2\pi z}$  für  $\operatorname{Re}(z) \rightarrow \infty$  (Stirling-Formel für die Gammafunktion).

Die Gammafunktion spielt eine zentrale Rolle bei der Bestimmung der Asymptotik von Koeffizienten binomischer Reihen. Genauer:

**Satz 1.2.16 (Transfersatz)** *Ist  $\alpha \notin \mathbb{N}_0$ , so folgt*

$$[z^n](1-z)^\alpha \sim \frac{n^{-\alpha-1}}{\Gamma(-\alpha)}. \quad (1.22)$$

*Wir werden eine genauere Darstellung von (1.22) im nächsten Abschnitt sehen.*

*Allgemeiner gilt*

$$[z^n] \left(1 - \frac{z}{z_0}\right)^\alpha \sim \frac{n^{-\alpha-1} z_0^{-n}}{\Gamma(-\alpha)}. \quad (1.23)$$

**Lemma 1.2.17** *Seien  $(a_n)_{n \geq 0}$  und  $(b_n)_{n \geq 0}$  zwei Folgen mit  $a_n = \mathcal{O}(n^{-\gamma})$ ,  $\gamma > 0$  und  $b_n = \mathcal{O}(c^n)$ ,  $0 < c < 1$ . Dann gilt*

$$\sum_{k=0}^n a_k b_{n-k} = \mathcal{O}(n^\gamma). \quad (1.24)$$

*Beweis:* Wir haben

$$\left| \sum_{k=0}^{n/2} a_k b_{n-k} \right| \leq \max_{0 \leq k \leq n/2} |a_k| \sum_{k=0}^{n/2} K c_{n-k} \leq K_1 K c^{n/2} = \mathcal{O}(n^{-\gamma}),$$

und

$$\left| \sum_{k=n/2}^n a_k b_{n-k} \right| \leq \max_{n/2 \leq k \leq n} |a_k| \sum_{k=n/2}^n K c_{n-k} = \mathcal{O}(n^{-\gamma}).$$

□

**Lemma 1.2.18** Sei  $f(z)$  analytisch in  $|z| < 1 + \eta$  mit  $\eta > 0$ . Setze weiters  $g(z) = (1 - z)^\gamma f(z)$  mit  $\gamma \in \mathbb{C} \setminus \mathbb{Z}$ . Dann haben wir

$$[z^n]g(z) = \mathcal{O}(n^{\gamma-1}). \quad (1.25)$$

*Beweis:* Setze  $a_n = [z^n](1 - z)^\gamma$  und  $b_n = [z^n]f(z)$ . Aus Satz 1.2.16 folgt, dass  $a_n = \mathcal{O}(n^{-\gamma-1})$ . Da  $f(z)$  analytisch in  $|z| < 1 + \eta$  ist, gilt  $b_n = \mathcal{O}(c^n)$  für ein  $c \in (0, 1)$ . Lemma 1.2.17 liefert uns (1.25).  $\square$

**Satz 1.2.19 (Darboux)** Sei  $f(z)$  analytisch in  $|z| < 1 + \eta$  mit  $\eta > 0$ . Lokal um  $z = 1$  gelte  $f(z) = \sum_{j \geq 0} a_j (1 - z)^j$ . Ist  $\alpha \notin \mathbb{N}_0$ , so folgt

$$[z^n](1 - z)^\alpha f(z) = \sum_{j=0}^m a_j \binom{n - \alpha - j - 1}{n} + \mathcal{O}(n^{-m-\alpha-2}). \quad (1.26)$$

*Beweis:* Wir können  $(1 - z)^\alpha f(z)$  wie folgt umschreiben:

$$(1 - z)^\alpha f(z) = \sum_{j=0}^m a_j (1 - z)^{\alpha+j} + \sum_{j>m} a_j (1 - z)^{\alpha+j}.$$

Daraus folgt, dass

$$[z^n](1 - z)^\alpha f(z) = \sum_{j=0}^m a_j \binom{n - \alpha - j - 1}{n} + [z^n](1 - z)^{\alpha+m+1} h(z).$$

Die Funktion  $h(z)$  ist ebenfalls analytisch in  $|z| < 1 + \eta$ . Aus Lemma 1.2.18 folgt damit  $[z^n](1 - z)^{\alpha+m+1} h(z) = \mathcal{O}(n^{-m-\alpha-2})$ .  $\square$

Nun haben wir einen weiteren Satz von Darboux.

**Satz 1.2.20 (Methode von Darboux)** Sei  $f(z)$  analytisch in  $|z| < 1$  und  $k$ -mal stetig differenzierbar auf  $|z| = 1$ . Dann gilt

$$[z^n]f(z) = o(n^{-k}). \quad (1.27)$$

*Beweis:* siehe [15, Theorem VI.14].  $\square$

Ein weiteres Lemma wurde von Meir und Moon in [25] ausformuliert.

**Lemma 1.2.21 (Meir und Moon, [25, Lemma 3.1])** Seien  $A(z) = \sum_{n \geq 0} a_n z^n$  und  $B(z) = \sum_{n \geq 0} b_n z^n$  zwei Potenzreihen und setze  $C(z) = \sum_{n \geq 0} c_n z^n := A(z)B(z)$ . Seien weiters  $a, b, k_1, k_2$  und  $\rho$  Konstanten, wobei  $a, b, \rho > 0$ ,

$$a_n \sim a \rho^{-n} n^{-k_1}, \quad n \rightarrow \infty$$

und

$$b_n \sim b \rho^{-n} n^{-k_2}, \quad n \rightarrow \infty$$

Dann gilt:

- (i) Wenn  $k_1 = 0$  und  $k_2 = 1/2$ , so folgt  $c_n \sim 2ab\rho^{-n} n^{1/2}$
- (ii) Wenn  $k_1 = k_2 = 1/2$ , so folgt  $c_n \sim \pi ab\rho^{-n}$
- (iii) Wenn  $k_1 = k_2 = 3/2$ , so folgt  $c_n \sim A(\rho)b_n + B(\rho)a_n$
- (iv) Wenn  $k_1 = 3/2$ ,  $k_2 = 1/2$  und  $A(\rho) \neq 0$ , so folgt  $c_n \sim A(\rho)b_n$



### 1.2.3 Singularitätsanalyse

Seien  $f_n = [z^n]f(z)$ . Gesucht ist eine analytische Funktion  $g(z)$ , deren Koeffizienten  $g_n = [z^n]g(z)$  eine leichtere Darstellung besitzen. Wir haben (siehe Formeln  $T_1, T_2, T_3$  und  $T_0$  in [12]):

- $f(z) = \mathcal{O}(g(z)) \Rightarrow f_n = \mathcal{O}(g_n)$ ;
- $f(z) = o(g(z)) \Rightarrow f_n = o(g_n)$ ;
- $f(z) \sim g(z) \Rightarrow f_n \sim g_n$ .

Gibt es zusätzlich weitere analytische Funktionen  $h_0(z), h_1(z), \dots, h_k(z)$  mit

$$f(z) = \sum_{i=0}^k h_i(z) + \mathcal{O}(g(z)) \quad \text{und} \quad h_{i+1}(z) = o(h_i(z)), i = 0, \dots, k-1$$

so gilt

$$f(z) = \sum_{i=0}^k h_{i,n} + \mathcal{O}(g_n) \quad \text{und} \quad h_{i+1,n} = o(h_{i,n}), i = 0, \dots, k-1$$

Nun wiederholen wir den Transfersatz:

**Satz 1.2.22 (Flajolet, Odlyzko)** *Ist  $\alpha \notin \mathbb{N}_0$ , so gilt*

$$[z^n](1-z)^\alpha \sim \frac{n^{-\alpha-1}}{\Gamma(-\alpha)} \left( 1 + \sum_{k \geq 1} \frac{e_k(\alpha)}{n^k} \right) \quad (1.28)$$

wobei  $\Gamma$  die Gammafunktion bezeichnet und  $e_k(\alpha) = \sum_{i=k}^{2k} (-1)^i c_{k,i}(\alpha+1)(\alpha+2) \dots (\alpha+i)$

mit  $\sum_{k,l \geq 0} c_{k,l} x^l y^k = \frac{e^x}{(1+xy)^{-1-1/y}}$ .

*Beweis:* Die Idee ist, einen Kreis  $C$  um Null vom Radius  $R > 1$  zu betrachten, diesen in eine *Hankel-Kurve*  $\mathcal{H}_n$  vom Radius  $R$  um Eins (die entsprechende Singularität von  $(1-z)^\alpha$  und von Schlitz-Breite  $\frac{1}{n}$  zu verzerren (vgl. mit [13, Proposition 1] oder [15, Theorem VI.1]). Dies ist ein geschlossener Weg bestehend aus den vier folgenden Teilen:

- $\mathcal{H}_n^- = \left\{ z = 1 - \frac{it}{n}, |z| \leq R \right\}$
- $\mathcal{H}_n^{(0)} = \left\{ z = 1 - \frac{e^{i\phi}}{n}, |\phi| \leq \frac{\pi}{2} \right\}$
- $\mathcal{H}_n^+ = \left\{ z = 1 + \frac{it}{n}, |z| \leq R \right\}$
- $\tilde{\mathcal{H}}_n$  der Kreis vom Radius  $R$  ohne die geschlitze Kurve  $\mathcal{H}_n^- \cup \mathcal{H}_n^{(0)} \cup \mathcal{H}_n^+$  um die Achse  $\mathbb{R}_{\geq 1}$

Wir integrieren dann  $(1-z)^\alpha$  entlang dieser neuen Kurve. Entlang  $\tilde{\mathcal{H}}_n$  haben wir (wegen  $|z|=R$ )  $|z|^{-n-1} = R^{-n-1}$ ,  $|1-z|^\alpha = \mathcal{O}(R^\alpha) = \mathcal{O}(1)$  und daher

$$\int_{\tilde{\mathcal{H}}_n} \frac{(1-z)^\alpha}{z^{n+1}} dz = \mathcal{O}(R^{-n})$$

Klarerweise konvergiert dann das Integral gegen 0 wenn  $R$  gegen  $\infty$  strebt.

Setze andererseits  $z = 1 + \frac{t}{n}$ . Wir erhalten  $(1-z)^\alpha = \frac{(-t)^\alpha}{n^\alpha}$ ,

$$z^{-n-1} = \left(1 + \frac{t}{n}\right)^{-n-1} = e^{-t} e^t \left(1 + \frac{t}{n}\right)^{-n-1} = e^{-t} \sum_{k,l \geq 0} c_{k,l} t^l n^{-k}.$$

Lasse  $R \rightarrow \infty$ . Wir haben dann eine neue Kurve  $\overline{\mathcal{H}}_n$ , und entlang dieser gilt

$$\begin{aligned} \frac{1}{2\pi i} \int_{\overline{\mathcal{H}}_n} \frac{(1-z)^\alpha}{z^{n+1}} dz &= \frac{1}{2\pi i} \int_{\overline{\mathcal{H}}_n} \frac{(-t)^\alpha}{n^\alpha} e^{-t} \sum_{k,l \geq 0} c_{k,l} t^l n^{-k} \frac{dt}{n} \\ &= \frac{n^{-\alpha-1}}{2\pi i} \int_{\overline{\mathcal{H}}_n} (-t)^\alpha e^{-t} \sum_{k,l \geq 0} c_{k,l} t^l n^{-k} dt \\ &= n^{-\alpha-1} \sum_{k,l \geq 0} c_{k,l} (-1)^l n^{-k} \frac{1}{2\pi i} \int_{\overline{\mathcal{H}}_n} (-t)^{\alpha+l} e^{-t} dt \\ &= n^{-\alpha-1} \sum_{k,l \geq 0} c_{k,l} (-1)^l n^{-k} \frac{1}{\Gamma(-\alpha-l)}. \end{aligned}$$

Wegen  $\Gamma(z+1) = z\Gamma(z)$  gilt

$$\frac{1}{\Gamma(-\alpha-l)} = \frac{-\alpha-l}{\Gamma(-\alpha-l+1)} = \frac{(-1)^2(\alpha+l-1)(\alpha+l)}{\Gamma(-\alpha-l+2)} = \dots = \frac{(-1)^l(\alpha+1)(\alpha+2)\dots(\alpha+l)}{\Gamma(-\alpha)}$$

Mit  $c_{0,0} = 1$  und  $c_{k,l} = 0$  für  $l < k$  und  $l > 2k$  erhalten wir (1.28).  $\square$

Eine Variante von diesem Satz existiert mit dem Logarithmus:

**Satz 1.2.23 (Flajolet, Odlyzko)** *Ist  $\alpha \notin \mathbb{N}_0$  und  $\beta$  eine komplexe Zahl, so gilt*

$$[z^n](1-z)^\alpha \left(\frac{1}{z} \log\left(\frac{1}{1-z}\right)\right)^\beta \sim \frac{n^{-\alpha-1}}{\Gamma(-\alpha)} (\log(n))^\beta \left(1 + \sum_{k \geq 1} \frac{c_k(\alpha, \beta)}{(\log(n))^k}\right) \quad (1.29)$$

$$\text{mit } c_k(\alpha, \beta) = \binom{\beta}{k} \Gamma(-\alpha) \left. \frac{d^k}{ds^k} \frac{1}{\Gamma(s)} \right|_{s=-\alpha}.$$

*Beweis:* Analog wie im vorigen Satz. Siehe [13, Theorem 3A] oder [15, Theorem VI.2].  $\square$

Sei  $f(z)$  eine Funktion analytisch bei  $z=0$  und singularär bei  $z=z_0$ . Die Idee von Flajolet (vgl. mit [13]) ist, einen bestimmten Analytizitätsbereich in der Nähe von  $z_0$  aufzubauen, um die Informationen der Singularität genauer zu machen, und daher die Asymptotik der Koeffizienten von  $f(z)$  zu bestimmen:

**Definition 1.2.24** Eine Funktion  $f$  heißt  $\Delta$ -analytisch, falls sie auf  $\Delta = \Delta(\phi, \eta) = \{z \in \mathbb{C} : |z| \leq 1 + \eta, |\arg(z - 1)| \geq \phi\}$  mit  $\eta > 0$  und  $0 < \phi < \frac{\pi}{2}$  analytisch ist.

Nun formulieren wir weitere Transfersätze für  $\Delta$ -analytischen Funktionen.

**Satz 1.2.25 (Flajolet, Odlyzko)** Sei  $f$   $\Delta$ -analytisch und singulär bei  $z = 1$ . Gilt zusätzlich

$$f(z) = \mathcal{O}(|1 - z|^\alpha) \quad (1.30)$$

mit  $\alpha \in \mathbb{R}$  in  $\Delta \setminus \{1\}$ , so ist

$$a_n = [z^n]f(z) = \mathcal{O}(n^{-\alpha-1}). \quad (1.31)$$

*Beweis:* (siehe auch Flajolet und Odlyzko, [13, Theorem 1] oder Flajolet und Sedgewick [15, Theorem VI.3]) Sei  $K \subseteq \Delta \setminus \{1\}$  kompakt. Dann gibt es eine Konstante  $C > 0$ , sodass  $|1 - z|^\alpha \geq C$  für alle  $z \in K$ . Wegen dieser Tatsache und der Analytizität von  $f(z)$  in solch einer Menge ist (1.30) äquivalent zu

$$|f(z)| \leq C_1 |1 - z|^\alpha,$$

wobei  $C_1 > 0$  eine Konstante ist. Mithilfe von (1.10) haben wir (mit  $z_0 = 0$ )

$$a_n = \frac{1}{2\pi i} \int_{\Gamma} \frac{f(z)}{z^{n+1}} dz,$$

wobei  $\Gamma$  eine positiv orientierte geschlossene Kurve in  $\Delta$  ist, die den Ursprung umschließt. Wir setzen  $\Gamma = \gamma_1 \cup \gamma_2 \cup \gamma_3 \cup \gamma_4$  mit

$$\gamma_1 = \{z : |z - 1| = \frac{1}{n}, |\arg(z - 1)| \geq \phi\}$$

$$\gamma_2 = \{z : \frac{1}{n} \leq |z - 1|, |z| \leq 1 + \eta, |\arg(z - 1)| = \phi\}$$

$$\gamma_3 = \{z : |z - 1| = 1 + \eta, |\arg(z - 1)| \geq \phi\}$$

$$\gamma_4 = \{z : \frac{1}{n} \leq |z - 1|, |z| \leq 1 + \eta, |\arg(z - 1)| = -\phi\}$$

und daher

$$a_n^{(j)} = \frac{1}{2\pi} \int_{\gamma_j} \frac{|f(z)|}{|z|^{n+1}} dz.$$

Klarerweise haben wir  $|a_n| \leq a_n^{(1)} + a_n^{(2)} + a_n^{(3)} + a_n^{(4)}$ .

Entlang  $\gamma_1$  haben wir  $|z|^{n+1} \geq \left(1 - \frac{1}{n}\right)^{n+1}$  und  $|f(z)| \leq C_1 |1 - z|^\alpha = C_1 n^{-\alpha}$ . Daraus folgt

$$a_n^{(1)} \leq \frac{1}{2\pi} C_1 n^{-\alpha} \left(1 - \frac{1}{n}\right)^{-n-1} \frac{2\pi}{n} = \mathcal{O}(n^{-\alpha-1}).$$

Für  $\gamma_2$  und  $\gamma_4$  sei  $w = e^{i\phi}$  und substituieren  $z$  durch  $1 + \frac{wt}{n}$  mit  $1 \leq t \leq E \cdot n$ , wobei  $E > 0$  so definiert ist, dass  $|1 + wE| = 1 + \eta$ . Dann gilt (für  $a_n^{(4)}$  analog)

$$\begin{aligned} a_n^{(2)} &\leq \frac{1}{2\pi} \int_1^{E \cdot n} C_1 \left(\frac{t}{n}\right)^\alpha \left|1 + \frac{wt}{n}\right|^{-n-1} \frac{dt}{n} \\ &\leq C_1 n^{-\alpha-1} \frac{1}{2\pi} \int_1^{E \cdot n} t^\alpha \left|1 + \frac{wt}{n}\right|^{-n-1} dt = \mathcal{O}(n^{-\alpha-1}), \end{aligned}$$

wobei das Integral im obigen Teil der Ungleichung in  $n$  gleichmäßig beschränkt ist.

Bei  $\gamma_3$  haben wir  $|z| = 1 + \eta$  und  $|z - 1| \leq 3$ . Wir bekommen

$$a_n^{(3)} \leq \frac{1}{2\pi} C_1 3^\alpha (1 + \eta)^{-n-1} 2\pi (1 + \eta) = \mathcal{O}((1 + \eta)^{-n}) = \mathcal{O}(n^{-\alpha-1})$$

Aufsummieren liefert uns (1.31). □

Analoge Sätze gelten auch für  $o$ - und  $\sim$ -Transfer (siehe [13, Corollary 1], [13, Corollary 2], [15, Theorem VI.3] und [15, Corollary VI.1]).

**Bemerkung 1.2.26** *Hat  $f$  ihre dominante Singularität bei  $z_0$ , und gilt  $f(z) = \mathcal{O}\left(\left|1 - \frac{z}{z_0}\right|^\alpha\right)$ ,  $\alpha \in \mathbb{R}$  für  $z \in \Delta_{z_0} \setminus \{z_0\}$ , wobei*

$$\Delta_{z_0} := \{z \in \mathbb{C} : |z| \leq z_0 + \eta, |\arg(z - z_0)| \geq \phi\}$$

so folgt  $[z^n]f(z) = \mathcal{O}(z_0^{-n} n^{-\alpha-1})$ . Analoge Aussagen gelten für  $o$ - und  $\sim$ -Transfer. (siehe [15, Theorem VI.4]).

Jetzt betrachten wir Funktionen  $y(z)$ , die eine Funktionalgleichung der Form

$$y(z) = \Psi(y(z)) \Leftrightarrow F(z, y(z)) := \Psi(y(z)) - y(z) \equiv 0 \quad (1.32)$$

erfüllen. Dabei bezeichnet  $\Psi$  einen Operator auf  $y(z)$ . Eine Methode, um die Koeffizienten von  $y(z)$  zu bestimmen, ist der *Hauptsatz über implizite Funktionen*:

**Satz 1.2.27 (Hauptsatz über implizite Funktionen)** *Sei  $F(z, y)$  bei  $(z_0, y_0)$  analytisch. Gesucht ist eine Funktion  $y(z)$ , die bei  $z_0$  stetig ist, und*

$$F(z, y(z)) \equiv 0 \quad (1.33)$$

erfüllt. Gelten zusätzlich

$$\begin{aligned} F(z_0, y_0) &= 0; & F_y(z_0, y_0) &= 0 \\ F_z(z_0, y_0) &\neq 0; & F_{yy}(z_0, y_0) &\neq 0 \end{aligned}$$

so gibt es eine Umgebung  $U$  von  $z_0$ , sodass für jedes  $z \in U \setminus \{z_0\}$  die Gleichung  $F(z, y) = 0$  genau zwei analytische Lösungen besitzt, wobei  $(z \rightarrow z_0)$

$$y_{1,2}(z) = y_0 \pm \sqrt{\frac{2z_0 F_z(z_0, y_0)}{F_{yy}(z_0, y_0)} \sqrt{1 - \frac{z}{z_0}}} + \mathcal{O}\left(1 - \frac{z}{z_0}\right). \quad (1.34)$$

*Beweis:* Sei  $y(z)$ , die lokal um  $z_0$  analytisch ist mit  $y(z_0) = y_0$ , und (1.33) löst. Dabei gilt:

$$\begin{aligned} 0 &= F(z, y(z)) \\ &= F_z(z_0, y_0)(z - z_0) + F_y(z_0, y_0)(y - y_0) \\ &\quad + \frac{1}{2} F_{zz}(z_0, y_0)(z - z_0)^2 + F_{zy}(z_0, y_0)(z - z_0)(y - y_0) + \frac{1}{2} F_{yy}(z_0, y_0)(y - y_0)^2 \\ &= F_z(z_0, y_0)(z - z_0) + \frac{1}{2} F_{zz}(z_0, y_0)(z - z_0)^2 + F_{zy}(z_0, y_0)(z - z_0)(y - y_0) + \frac{1}{2} F_{yy}(z_0, y_0)(y - y_0)^2 \\ &= F_z(z_0, y_0)(z - z_0) + \mathcal{O}(|z - z_0|) + \frac{1}{2} F_{yy}(z_0, y_0)(y - y_0)^2 \end{aligned}$$

Daraus folgt (mit  $y \rightarrow y_0$  und  $z \rightarrow z_0$ )

$$\begin{aligned}(y - y_0)^2 &= -\frac{F_z(z_0, y_0)}{\frac{1}{2}F_{yy}(z_0, y_0)}(z - z_0) + o(|z - z_0|) \\ &= \frac{2z_0 F_z(z_0, y_0)}{F_{yy}(z_0, y_0)} \left(1 - \frac{z}{z_0}\right) + o(|z - z_0|)\end{aligned}$$

und daher

$$y - y_0 = \pm \sqrt{\frac{2z_0 F_z(z_0, y_0)}{F_{yy}(z_0, y_0)}} \sqrt{1 - \frac{z}{z_0}} + o(|z - z_0|)$$

□

Nun kümmern wir uns um *inverse Funktionen*:

- Gegeben: Funktion  $\Psi(y)$  analytisch bei  $y = y_0$  mit  $\Psi(y_0) = z_0$ ;
- Gesucht: Lösung(en) der Gleichung  $\Psi(y) = z$  mit  $z$  bzw.  $y$  hinreichend nahe  $z_0$  bzw.  $y_0$

Dazu formulieren wir folgende Lemmata aus:

**Lemma 1.2.28 (Analytische Inversion)** Sei  $\Psi(y)$  analytisch bei  $y = y_0$  mit  $\Psi(y_0) = z_0$ . Gilt zusätzlich  $\Psi'(y_0) \neq 0$ , dann gibt es eine Umgebung  $U$  von  $z_0$  und eine analytische Funktion  $y(z)$ , die die Gleichung  $\Psi(y) = z$  auf  $U$  löst, und für die gilt  $y(z_0) = y_0$ .

*Beweis:* siehe [15, Lemma IV.2].

**Lemma 1.2.29 (Singuläre Inversion)** Sei  $\Psi(y)$  analytisch bei  $y = y_0$  mit  $\Psi(y_0) = z_0$ . Gelten zusätzlich  $\Psi'(y_0) = 0$  und  $\Psi''(y_0) \neq 0$ , so existiert eine Umgebung  $U$  von  $z_0$ , für die gilt: für jede feste Richtung  $\theta$  gibt es zwei Funktionen  $y_1(z)$  und  $y_2(z)$  auf  $U_{\neq \theta} := \{z \in U \mid \arg(z - z_0) \not\equiv \theta \pmod{2\pi}, z \neq z_0\}$ , die die Gleichung  $\Psi(y) = z$  auf  $U_{\neq \theta}$  lösen. Diese sind auf dieser Menge analytisch und erfüllen  $\lim_{z \rightarrow z_0} y(z) = y_0$ .

*Beweis:* siehe [15, Lemma IV.3].

Wir wiederholen das Prinzip der Singularitätsanalyse: Gegeben ist  $f(z)$  für  $|z| < R$  analytisch

- 1. Schritt: Lokalisier die dominante Singularität  $z_0$  von  $f(z)$ , das heißt  $|z_0| = R$ , und lege fest, dass  $f(z)$  im Gebiet  $\Delta_{z_0}$  analytisch ist;
- 2. Schritt: Bestimme die singuläre Entwicklung von  $f(z)$  als  $z \rightarrow z_0$
- 3. Schritt: Wende einen Transfersatz an, um die Asymptotik der Koeffizienten zu bestimmen.

## 1.3 Parameter, multivariate erzeugende Funktionen und Grenzverteilungen

Nun betrachten wir kombinatorische Klassen als zufällige Strukturen.

### 1.3.1 Parameter

Um kombinatorische Klassen als zufällige Strukturen zu definieren brauchen wir weitere Begriffe, die man auch in [15] finden kann.

**Definition 1.3.1** Sei  $\mathcal{A} = (A, |\cdot|)$  eine kombinatorische Struktur. Unter einem (eindimensionalen) Parameter von  $\mathcal{A}$  versteht man eine Funktion  $\chi : A \rightarrow \mathbb{N}_0$ . Wir erhalten dann eine doppelt indizierte Folge nämlich

$$a_{n,k} = |\{\alpha \in \mathcal{A} : |\alpha| = n, \chi(\alpha) = k\}| \quad (1.35)$$

und daher folgende bivariate erzeugende Funktion:

$$A(z, u) = \sum_{\alpha \in \mathcal{A}} \frac{z^{|\alpha|}}{\omega_{|\alpha|}} u^{\chi(\alpha)} = \sum_{n \geq 0} \sum_{k \geq 0} a_{n,k} \frac{z^n}{\omega_n} u^k, \quad (1.36)$$

wobei

$$\omega_n = \begin{cases} 1, & \text{für unmarkierte Strukturen} \\ n!, & \text{für markierte Strukturen.} \end{cases}$$

Dabei markiert  $z$  die Größe und  $u$  den Parameter  $\chi$ .

Parameter können auch mehrdimensional sein.

**Definition 1.3.2** Sei  $\mathcal{A} = (A, |\cdot|)$  eine kombinatorische Struktur. Ein  $d$ -dimensionaler Parameter ist ein  $d$ -dimensionaler Vektor  $\boldsymbol{\chi} = (\chi_1, \chi_2, \dots, \chi_d)$  mit  $\chi_i : A \rightarrow \mathbb{N}$  ( $i = 1, \dots, d$ ). Dabei bekommen wir folgende Zählfolge (mit  $\mathbf{k} = (k_1, k_2, \dots, k_d)$ ):

$$a_{n,\mathbf{k}} = |\{\alpha \in \mathcal{A} : |\alpha| = n, \chi_i(\alpha) = k_i \text{ für } i = 1 \dots d\}| \quad (1.37)$$

und daher folgende multivariate erzeugende Funktion (mit  $\mathbf{u}^{\mathbf{k}} = \prod_{i=1}^d u_i^{k_i}$ ):

$$A(z, \mathbf{u}) = \sum_{\alpha \in \mathcal{A}} \frac{z^{|\alpha|}}{\omega_{|\alpha|}} \mathbf{u}^{\chi(\alpha)} = \sum_{n \geq 0} \sum_{\mathbf{k} \in \mathbb{N}^d} a_{n,\mathbf{k}} \frac{z^n}{\omega_n} \mathbf{u}^{\mathbf{k}}. \quad (1.38)$$

Wir definieren folgenden Operator:

**Definition 1.3.3** Seien  $\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_d$  kombinatorische Strukturen und  $\mathcal{A} = \phi(\mathcal{B}_1, \dots, \mathcal{B}_d)$  eine zulässige Konstruktion. Dann ist das Kennzeichnen der Strukturen  $\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_d$  eine Substitution der Form  $\mathcal{B}_i \rightarrow \mu_i \mathcal{B}_i := \{\mu_i\} \times \mathcal{B}_i, i = 1 \dots, d$  auf die Objekte in  $\mathcal{A}$ , wobei die  $\mu_i$  Objekte der Größe 0 sind. Aus der Gleichung  $A(z) = \psi(B_1(z), \dots, B_d(z))$  folgt dann

$$A(z, u_1, \dots, u_d) = \psi(u_1 B_1(z), \dots, u_d B_d(z))$$

Wir betrachten wieder eine kombinatorische Klasse  $\mathcal{A} = (A, |\cdot|)$ . Nun setzen wir  $\mathcal{A}_n := \{\alpha \in \mathcal{A} : |\alpha| = n\}$ . Dann gilt  $a_n = |\mathcal{A}_n| = \omega_n [z^n] A(z, 1)$ . Stattet man  $\mathcal{A}$  mit allen Teilmengen  $\mathcal{A}_n$  mit der (diskreten) Gleichverteilung aus, und bezeichne  $X_n$  die zugehörige Zufallsvariable, so gilt  $\mathbb{P}_n(X_n = \alpha) = \frac{1}{a_n}$  für ein  $\alpha \in \mathcal{A}_n$ .

**Definition 1.3.4** Sei  $\chi$  ein Parameter von  $\mathcal{A}$ . Dieser bildet eine (diskrete) Zufallsvariable auf den Wahrscheinlichkeitsraum  $\mathcal{A}_n$  mit:

$$\mathbb{P}_n(\chi = k) = \frac{a_{n,k}}{a_n} = \frac{a_{n,k}}{\sum_{k \geq 0} a_{n,k}}. \quad (1.39)$$

Daher ist die zugehörige Wahrscheinlichkeitserzeugende Funktion wie folgt definiert:

$$p(u) = \sum_{k \geq 0} \mathbb{P}_n(\chi = k) u^k \quad (1.40)$$

**Satz 1.3.5** Sei  $\mathcal{A} = (A, |\cdot|)$  eine kombinatorische Struktur, und  $A(z, u)$  die zugehörige bivariate erzeugende Funktion mit einem Parameter  $\chi$ . Dann gilt:

$$p(u) = \mathbb{E}(u^X) = \frac{[z^n]A(z, u)}{[z^n]A(z, 1)}. \quad (1.41)$$

*Beweis:* Wir wissen schon, dass

$$\omega_n[z^n]A(z, 1) = a_n.$$

Andererseits gilt aus (1.36)

$$\omega_n[z^n]A(z, u) = \sum_{k \geq 0} a_{n,k} u^k,$$

also insgesamt

$$\frac{[z^n]A(z, u)}{[z^n]A(z, 1)} = \sum_{k \geq 0} \frac{a_{n,k}}{a_n} u^k = \sum_{k \geq 0} \mathbb{P}_n(\chi = k) u^k = p(u).$$

□

**Definition 1.3.6** Sei  $X$  eine diskrete Zufallsvariable mit Wertebereich  $\mathbb{N}_0$  und  $P(u) = \sum_{k \geq 0} \mathbb{P}(X = k) u^k$  ihre zugehörige Wahrscheinlichkeitserzeugende Funktion. Dann sind der Erwartungswert von  $X$  unter  $\mathbb{P}$  durch:

$$\mathbb{E}(X) = p'(1) = \sum_{k \geq 0} k \mathbb{P}(X = k). \quad (1.42)$$

definiert. Allgemeiner gilt für eine Funktion  $f$

$$\mathbb{E}(f(X)) = \sum_{k \geq 0} f(k) \mathbb{P}(X = k) \quad (1.43)$$

Ist  $f(t) = t^r$  mit  $r \in \mathbb{N}$ , so heißt der folgende Erwartungswert

$$\mathbb{E}(X^r) = \left( u \frac{\partial}{\partial u} \right)^r p(u) \Big|_{u=1} = \sum_{k \geq 0} k^r \mathbb{P}(X = k) \quad (1.44)$$

das  $r$ -te Moment von  $X$  unter  $\mathbb{P}$ .

Die Varianz von  $X$  unter  $\mathbb{P}$  lässt sich daher durch:

$$\mathbb{V}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 \quad (1.45)$$

definieren.

**Satz 1.3.7** Seien  $\mathcal{A}$ ,  $A(z, u)$  und  $\chi$  wie im Satz 1.3.5. Dann gilt:

$$\mathbb{E}_n(\chi) = \frac{[z^n]A_u(z, 1)}{[z^n]A(z, 1)} \quad (1.46)$$

und

$$\mathbb{V}_n(\chi) = \frac{[z^n]A_{uu}(z, 1)}{[z^n]A(z, 1)} + \mathbb{E}_n(\chi) - (\mathbb{E}_n(\chi))^2. \quad (1.47)$$

*Beweis:* Die Darstellung des Erwartungswertes kommt direkt aus (1.41) und (1.42). Für die Varianz gilt (mithilfe von (1.45), (1.44), (1.41) und (1.42)):

$$\begin{aligned} \mathbb{V}_n(\chi) &= \mathbb{E}(\chi^2) - (\mathbb{E}(\chi))^2 \\ &= \left( u \frac{\partial}{\partial u} \right)^2 p(u) \Big|_{u=1} - (\mathbb{E}(\chi))^2 \\ &= (u^2 p''(u) + u p'(u))|_{u=1} - (\mathbb{E}(\chi))^2 \\ &= p''(1) + p'(1) - (\mathbb{E}(\chi))^2 \\ &= \frac{[z^n]A_{uu}(z, 1)}{[z^n]A(z, 1)} + \mathbb{E}_n(\chi) - (\mathbb{E}_n(\chi))^2 \end{aligned}$$

## 1.3.2 Grenzverteilungen

Wir betrachten nun eine diskrete Zufallsvariable  $X_n$  mit entsprechenden Wahrscheinlichkeiten  $p_{n,k} := \mathbb{P}(X_n = k)$  und lassen  $n$  gegen  $\infty$  streben. Interessant ist, wenn die  $p_{n,k}$  konvergieren. Wir bekommen eine *Grenzverteilung*, die diskret, aber auch stetig sein kann.

### Diskrete Grenzverteilungen

**Definition 1.3.8** Sei  $X_n$  eine diskrete Zufallsvariable mit Verteilung  $p_{n,k} = \mathbb{P}(X_n = k)$ . Dann besitzt  $X_n$  eine diskrete Grenzverteilung genau dann, wenn für jedes  $k \in \mathbb{N}_0$ , der folgende Grenzwert  $p_k := \lim_{n \rightarrow \infty} p_{n,k}$  mit  $\sum_{k \geq 0} p_k = 1$  existiert.

**Satz 1.3.9** Sei  $X_n$  eine diskrete Zufallsvariable mit Verteilung  $(p_{n,k})_{k \geq 0}$  und diskreter Grenzverteilung  $(p_k)_{k \geq 0}$ . Dann ist für jedes  $k \in \mathbb{N}_0$  die Konvergenz von  $p_{n,k}$  gegen  $p_k$  gleichmäßig.

*Beweis:* Wegen  $\sum_{k \geq 0} p_k = 1 < \infty$  gibt es für jedes  $\epsilon > 0$  ein  $k_0 \in \mathbb{N}$ , sodass:

$$\sum_{k \geq k_0} p_k < \epsilon \quad \text{und} \quad \sum_{k < k_0} p_k > 1 - \epsilon$$



Da  $p_{n,k}$  gegen  $p_k$  für jedes  $k$  konvergiert, kann man ein  $n_0$  finden, sodass für alle  $n > n_0$  und  $k < k_0$   $|p_{n,k} - p_k| < \frac{\epsilon}{k_0}$  gilt. Dies impliziert

$$\sum_{k < k_0} p_{n,k} > 1 - 2\epsilon \quad \text{und} \quad \sum_{k \geq k_0} p_k < 2\epsilon$$

und daher

$$\sum_{k \geq 0} |p_{n,k} - p_k| = \sum_{k < k_0} |p_{n,k} - p_k| + \sum_{k \geq k_0} |p_{n,k} - p_k| < 4\epsilon.$$

□

**Satz 1.3.10 (Stetigkeitssatz)** Sei  $\Omega$  eine Teilmenge des Einheitskreises mit einem Häufungspunkt im Inneren des Kreises. Setze  $p_n(u) = \sum_{k \geq 0} p_{n,k} u^k$  und  $p(u) = \sum_{k \geq 0} p_k u^k$ .

Gilt für jedes  $u \in \Omega$   $p(u) = \lim_{n \rightarrow \infty} p_n(u)$ , dann gelten für alle  $k \in \mathbb{N}_0$

- $\lim_{n \rightarrow \infty} p_{n,k} = p_k$
- $\lim_{n \rightarrow \infty} \sum_{i < k} p_{n,i} = \sum_{i < k} p_i$ .

*Beweis:* Siehe [15, Theorem IX.1].

□

## Stetige Grenzverteilungen

Im vorigen Abschnitt haben wir gesehen, dass eine diskrete Zufallsvariable  $X_n$  eine diskrete Grenzverteilung besitzen könnte. Wir betrachten nun den Fall, wenn  $\lim_{n \rightarrow \infty} \mathbb{E}(X_n) = \infty$  und  $\lim_{n \rightarrow \infty} \mathbb{V}(X_n) = \infty$ . Zuerst wiederholen wir ein paar Grundlagen der Wahrscheinlichkeitstheorie:

**Definition 1.3.11** Die Verteilungsfunktion einer reellen Zufallsvariable  $X$  ist gegeben durch  $F_X(x) = \mathbb{P}(X \leq x)$ . Ist  $F_X(x)$  stetig, so hat  $X$  eine stetige Verteilung und wird stetige Zufallsvariable genannt.

Wir brauchen folgenden Konvergenzbegriff:

**Definition 1.3.12** Sei  $X$  eine stetige Zufallsvariable mit Verteilungsfunktion  $F_X$ , und  $X_n, n \in \mathbb{N}$  eine Folge von Zufallsvariablen mit Verteilungsfunktion  $F_{X_n}$ . Dann konvergiert  $X_n$  gegen  $X$  in Verteilung oder schwach (im Zeichen  $X_n \xrightarrow{d} X$ ), falls

$$\lim_{n \rightarrow \infty} F_{X_n}(x) = F_X(x), \quad \forall x \in \mathbb{R}. \quad (1.48)$$

**Definition 1.3.13** Sei  $X$  eine Zufallsvariable mit Mittelwert  $\mu = \mathbb{E}(X)$  und Varianz  $\sigma^2 = \mathbb{V}(X)$ . Dann ist die zentrierte Zufallsvariable  $X^*$  durch

$$X^* = \frac{X - \mu}{\sigma} \quad (1.49)$$

gegeben. Die entsprechende Verteilungsfunktion lautet dann  $F_{X^*}(x) = F_X(\mu + \sigma x)$  für jedes  $x \in \mathbb{R}$ .

**Bemerkung 1.3.14** Man sieht unmittelbar, dass  $F_X(x) = F_{X^*}\left(\frac{x-\mu}{\sigma}\right)$ .

**Definition 1.3.15** Seien  $X_n$  eine diskrete Zufallsvariable und  $Y$  eine stetige Zufallsvariable. Gilt  $X_n^* \xrightarrow{d} Y$ , so erfüllt  $X_n$  einen stetigen Grenzwertsatz.

**Bemerkung 1.3.16** Wegen Definitionen 1.3.12 und 1.3.13 mit  $\mu_n = \mathbb{E}(X_n)$  und  $\sigma_n^2 = \mathbb{V}(X_n)$  ist  $X_n^* \xrightarrow{d} Y$  äquivalent zu

$$\lim_{n \rightarrow \infty} F_{X_n^*}(x) = \lim_{n \rightarrow \infty} F_{X_n}(\mu_n + \sigma_n^2 x) = F_Y(x), \quad x \in \mathbb{R}.$$

Wir haben im vorigen Abschnitt gesehen, dass eine Zufallsvariable eine diskrete Grenzverteilung besitzt, wenn der Stetigkeitssatz 1.3.10 erfüllt ist. Es gibt auch eine Version dieses Satzes für stetige Grenzverteilungen. Doch brauchen wir zuerst weitere Definitionen, die hilfreich sein können:

**Definition 1.3.17** Sei  $X$  eine Zufallsvariable. Dazu werden folgende Integraltransformationen:

- $\lambda_X(s) = \mathbb{E}(e^{sX}) = \int_{-\infty}^{\infty} e^{sx} dF_X(x)$ , (Laplace-Transformation) momenterzeugende Funktion;
- $\phi_X(t) = \mathbb{E}(e^{itX}) = \int_{-\infty}^{\infty} e^{itx} dF_X(x)$ , (Fourier-Transformation) charakteristische Funktion

genannt.

**Bemerkung 1.3.18** Man kann alternativ das  $r$ -te Moment definieren:

$$\mathbb{E}(X^r) = \left. \frac{d^r}{ds^r} \lambda_X(s) \right|_{s=0} = (-i)^r \left. \frac{d^r}{dt^r} \phi_X(t) \right|_{t=0}. \quad (1.50)$$

Man sieht leicht, dass  $\lambda_{X^*}(s) = e^{-\frac{\mu s}{\sigma}} \lambda_X\left(\frac{s}{\sigma}\right)$  und  $\phi_{X^*}(t) = e^{-\frac{i\mu t}{\sigma}} \phi_X\left(\frac{t}{\sigma}\right)$ .

Nun können wir eine Version des Stetigkeitssatzes für die stetige Grenzverteilung darstellen:

**Satz 1.3.19 (Stetigkeitssatz)** Seien  $X_n, n \in \mathbb{N}$  und  $X$  Zufallsvariablen mit charakteristischen Funktionen  $\phi_n$  bzw.  $\phi$ . Weiters habe  $X$  eine stetige Verteilungsfunktion. Dann sind äquivalent:

- $X_n \xrightarrow{d} X$ ;
- $\lim_{n \rightarrow \infty} \phi_n(t) = \phi(t)$  für alle  $t \in \mathbb{R}$ ;
- $\lim_{n \rightarrow \infty} \phi_n(t) = \phi(t)$  für alle  $t \in (-t_0, t_0)$  und  $t_0$  reell.

*Beweis:* Siehe [5, Theorem 26.3]. □

**Beispiel 1.3.20** Das prominenteste Beispiel einer stetigen Verteilung ist die Normalverteilung (oder Gauß-Verteilung)  $(\mu, \sigma^2)$  mit  $\mu \in \mathbb{R}$  und  $\sigma^2 > 0$ . Die ist gegeben durch

$$F_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt = \Phi\left(\frac{x-\mu}{\sigma}\right) \quad (1.51)$$

Dabei bezeichnet  $\Phi$  die Verteilungsfunktion der sogenannten Standardnormalverteilung, also die zentrierte Normalverteilung mit Mittelwert 0 und Varianz 1.

Die charakteristische Funktion der (Standard-) Normalverteilung ist gegeben durch  $\phi_X(t) = \exp\left(i\mu t - \frac{\sigma^2 t^2}{2}\right)$  ( $\phi_{X^*}(t) = \exp\left(-\frac{t^2}{2}\right)$ ).

Die Standardnormalverteilung spielt eine wichtige Rolle als stetige Grenzverteilung. Es gibt nämlich eine Reihe von Sätzen, in der eine zentrierte (diskrete) Zufallsvariable gegen eine standardnormalverteilte Zufallsvariable schwach konvergiert. Man spricht von *asymptotischer Normalität*.

**Satz 1.3.21 (Zentraler Grenzwertsatz)** Seien  $X_n, n \in \mathbb{N}$  unabhängige und identisch verteilte Zufallsvariablen mit Mittelwert  $\mu = \mathbb{E}(X_n)$  und Varianz  $\sigma^2 = \mathbb{V}(X_n)$ . Setze  $S_n = \sum_{i=1}^n X_i$ . Dann konvergiert die zentrierte Zufallsvariable  $S_n^*$  schwach gegen eine standard normalverteilte Zufallsvariable  $N$ , also

$$S_n^* = \frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{d} N \sim \mathcal{N}(0, 1) \quad (1.52)$$

*Beweis:* Siehe [5, Theorem 27.1] oder [15, Theorem IX.6]. □

**Satz 1.3.22 (Quasi-Power Theorem, Hwang)** Sei  $X_n$  eine nichtnegative diskrete Zufallsvariable mit wahrscheinlichkeitserzeugender Funktion  $p_n(u)$ . Es gelte

$$p_n(u) = A(u)B(u)^{\beta_n} \left(1 + \mathcal{O}\left(\frac{1}{\kappa_n}\right)\right) \quad (1.53)$$

gleichmäßig in fester Umgebung von  $u = 1$ , wobei  $\beta_n, \kappa_n \rightarrow \infty$ ,  $A(u), B(u)$  um  $u = 1$  analytisch sind mit  $A(1) = B(1) = 1$ . Weiters erfülle  $B(u)$  die sogenannte Variabilitätsbedingung, das heißt

$$\text{Var}(B) = B''(1) + B'(1) - B'(1)^2 \neq 0.$$

Dann ist  $X_n$  asymptotisch normalverteilt mit Mittelwert

$$\mu_n := \mathbb{E}(X_n) \sim \beta_n B'(1) + A'(1) + \mathcal{O}\left(\frac{1}{\kappa_n}\right)$$

und Varianz

$$\sigma_n^2 := \mathbb{V}(X_n) \sim \beta_n \text{Var}(B) + \text{Var}(A) + \mathcal{O}\left(\frac{1}{\kappa_n}\right).$$

Dazu ist die Verteilung der zentrierten Zufallsvariable (asymptotisch) gegeben durch

$$\mathbb{P}\left(\frac{X_n - \mu_n}{\sigma_n} \leq x\right) = \phi(x) + \mathcal{O}\left(\frac{1}{\kappa_n} + \frac{1}{\sqrt{\beta_n}}\right), \quad (1.54)$$

wobei  $\phi$  die Verteilungsfunktion der Standardnormalverteilung bezeichnet.

*Beweis:* Siehe [8, Theorem 2.22].

Wir werden sehen, dass eine Variante des Hauptsatzes für implizite Funktionen zu asymptotischer Normalität führt (vgl. mit [8, Theorem 2.21])

**Proposition 1.3.23** Sei  $F(z, y, u) = \sum_{n \geq 0} \sum_{k \geq 0} f_{n,k}(u) z^n y^k$  eine in einer Umgebung von  $(0, 0, 0)$  analytische Funktion, für die gilt:

$$\begin{aligned} F(0, y, u) &\equiv 0 \\ F(z, 0, u) &\neq 0 \\ f_{n,k}(1) &\in \mathbb{R}_{\geq 0}, \quad \forall n, k \in \mathbb{N}_0 \end{aligned}$$

Dann ist die eindeutig bestimmte Lösung  $y = y(z, u) = \sum_{n \geq 0} y_n(u) z^n$  der Funktionalgleichung

$$y = F(z, y, u)$$

mit  $y(0, u) = 0$  um 0 analytisch, und wir haben  $y_n(1) \in \mathbb{R}_{\geq 0}, n \in \mathbb{N}_0$ . Weiters existieren im Konvergenzradius von  $F(z, y, u)$  nichtnegative Lösungen  $z = z_0$  und  $y = y_0$  sodass

$$\begin{aligned} F(z_0, y_0, 1) &= y_0 \\ F_y(z_0, y_0, 1) &= 1 \\ F_z(z_0, y_0, 1) &\neq 0 \\ F_{yy}(z_0, y_0, 1) &\neq 0 \end{aligned}$$

Daher gibt es um  $(z_0, 1)$  analytische Funktionen  $a(z, u) = a(u)$ ,  $b(z, u)$  und  $c(z, u)$  sodass  $y(z, u)$  für  $|z| < z_0$  und  $|u| \leq 1$  analytisch ist, und lokal um  $(z_0, 1)$  folgenden Ausdruck besitzt

$$y(z, u) = b(z, u) - c(z, u) \sqrt{1 - \frac{z}{a(u)}}. \quad (1.55)$$

Außerdem haben wir  $a(1) = z_0$ ,  $b(a(u), u) = y(a(u), u)$  und

$$c(a(u), u) = \sqrt{\frac{2a(u)F_z(a(u), b(a(u), u), u)}{F_{yy}(a(u), b(a(u), u), u)}}.$$

*Beweis:* Analog wie im Satz 1.2.27. □

**Satz 1.3.24** Sei  $F(z, y, u) = \sum_{n \geq 0} \sum_{k \geq 0} f_{n,k}(u) z^n y^k$  mit allen Bedingungen von Proposition 1.3.22 erfüllt. Ist  $(X_n)_{n \geq 1}$  die Folge der Zufallsvariablen definiert durch die Wahrscheinlichkeitserzeugende Funktion

$$\mathbb{E}(u^{X_n}) = \frac{[z^n]y(z, u)}{[z^n]y(z, 1)},$$

so besitzen der Mittelwert und die Varianz folgende asymptotische Darstellungen:

$$\mathbb{E}(X_n) = \mu n + \mathcal{O}(1) \quad \mathbb{V}(X_n) = \sigma^2 n + \mathcal{O}(1), \quad (1.56)$$

wobei

$$\mu = \frac{F_u}{z_0 F_z}(z_0, y_0, 1)$$

und

$$\sigma^2 = \left[ \left( \frac{F_u}{z_0 F_z} \right)^2 + \frac{1}{z_0} \left( \frac{F_u}{F_z} - \frac{1}{F_{yy} F_z} \left( \frac{F_u F_{zy}}{F_z} - F_{yu} \right) + \frac{1}{F_z} \left( \frac{F_u^2 F_{zz}}{F_z^2} - \frac{2F_u F_{zu}}{F_z} + F_{uu} \right) \right) \right] (z_0, y_0, 1).$$

Gilt zusätzlich  $\sigma^2 > 0$ , so konvergiert  $X_n$  in Verteilung gegen eine standardnormalverteilte Zufallsvariable, also

$$\frac{X_n - \mathbb{E}(X_n)}{\sqrt{\mathbb{V}(X_n)}} \xrightarrow{d} N \sim \mathcal{N}(0, 1). \quad (1.57)$$

*Beweis:* Siehe [8, Theorem 2.23].

## Brownsche Exkursionen und Konvergenz von stochastischen Prozessen

Nun kümmern wir uns nicht mehr um die Konvergenz einer Zufallsvariable, sondern von einem sogenannten *stochastischen Prozess*. Folgende Begriffe kann man in [4], [6], [21], [22] oder [33] finden.

**Definition 1.3.25** Sei  $T \subseteq \mathbb{R}$ . Unter einem stochastischen Prozess versteht man eine Familie von Zufallsvariablen  $(X(t))_{t \in T}$  mit Werten in  $Y \subseteq \mathbb{R}^k$ . Der Definitionsbereich  $T$  bzw. Wertebereich  $Y$  wird Parameterraum bzw. Zustandsraum genannt. Ist  $T$  diskret bzw. stetig, so redet man von einem Prozess in diskreter bzw. stetiger Zeit.

Prominente Beispiele von stochastischen Prozessen sind der *Markov-Prozess* und der *Wiener-Prozess*.

**Definition 1.3.26** Unter einem Markov-Prozess versteht man einen stochastischen Prozess  $X(t)$  mit folgender Eigenschaft: für  $s_1 < \dots < s_n < s < t$  und jede Borelmenge gilt

$$\mathbb{P}(X(t) \in B | X(s_1), \dots, X(s_n), X(s)) = \mathbb{P}(X(t) \in B | X(s))$$

**Definition 1.3.27** Ein Prozess  $(B(t))_{t \geq 0}$  mit Parameterraum  $T = [0, \infty)$  und Zustandsraum  $X = \mathbb{R}$  heißt *Wiener-Prozess* oder *Brownsche Bewegung*, falls:

- $B(0) = 0$ ;
- für  $t_0 < t_1 < \dots < t_d$  sind die Inkremente  $B(t_i) - B(t_{i-1})$ ;  $i = 1, \dots, d$  unabhängig
- für  $t_0 < t_1$  gilt  $B(t_1) - B(t_0) \sim \mathcal{N}(0, t_1 - t_0)$ .

Mithilfe des Wiener-Prozesses kann man die *Brownsche Exkursion* definieren (vgl. mit Definition von Cohen und Hooghiemstra in [6]):

**Definition 1.3.28** Sei  $(B(t))_{t \geq 0}$  ein eindimensionaler Wiener-Prozess. Setze  $\tau^- := \sup\{s \leq 1 : B(s) = 0\}$  und  $\tau^+ := \inf\{s \geq 1 : B(s) = 0\}$ . Dann heißt der Prozess  $(e(t))_{0 \leq t \leq 1}$  gegeben durch

$$e(t) := \frac{|B(\tau^- + t(\tau^+ - \tau^-))|}{(\tau^+ - \tau^-)^{1/2}}, \quad 0 \leq t \leq 1 \quad (1.58)$$

Brownsche Exkursion.

**Bemerkung 1.3.29** Eine Brownsche Exkursion kann als Wiener Prozess mit  $\mathbb{P}(e(0) = 0) = \mathbb{P}(e(1) = 0) = 1$  alternativ definiert werden.

**Definition 1.3.30** Die lokale Zeit einer Brownschen Exkursion  $(e(t))_{0 \leq t \leq 1}$  der Dauer 1 ist definiert durch

$$l(t) = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \int_0^1 \mathbb{1}_{[t, t+\epsilon]} e(s) ds. \quad (1.59)$$

Daher erhalten wir einen weiteren stochastischen Prozess  $(l(t))_{t \geq 0}$ .

**Satz 1.3.31** Die charakteristische Funktion für die lokale Zeit einer Brownschen Exkursion der Dauer 1 ist gegeben durch

$$\mathbb{E}e^{it\kappa} = 1 + \frac{\sqrt{2}}{\sqrt{\pi}} \int_{c-i\infty}^{c+i\infty} \frac{t\sqrt{-x} \exp(-\kappa\sqrt{-2x})}{\sqrt{-x} \exp(-\kappa\sqrt{-2x}) - it\sqrt{2} \sinh(\kappa\sqrt{-2x})} e^{-x} dx, \quad (1.60)$$

wobei  $c < 0$  eine Konstante ist. Für den Vektor  $(l(\kappa_1), \dots, l(\kappa_d))$  haben wir:

$$\mathbb{E}e^{it_1\kappa_1 + \dots + it_d\kappa_d} = 1 + \frac{\sqrt{2}}{i\sqrt{\pi}} \int_{c-i\infty}^{c+i\infty} f_{\kappa_1, \dots, \kappa_d}(x, t_1, \dots, t_d) e^{-x} dx, \quad (1.61)$$

wobei

$$f_{\kappa_1, \dots, \kappa_d}(x, t_1, \dots, t_d) = \Psi_{\kappa_1}(z, t_1 + \Psi_{\kappa_2 - \kappa_1}(\dots \Psi_{\kappa_{d-1} - \kappa_{d-2}}(x, t_{d-1} + \Psi_{\kappa_d - \kappa_{d-1}}(x, t_d)) \dots)) \quad (1.62)$$

mit

$$\Psi_{\kappa}(x, t) = \frac{it\sqrt{-x} \exp(-\kappa\sqrt{-2x})}{\sqrt{-x} \exp(-\kappa\sqrt{-2x}) - it\sqrt{2} \sinh(\kappa\sqrt{-2x})}. \quad (1.63)$$

*Beweis:* Siehe [22]. □

Nun kümmern wir uns um die schwache Konvergenz eines stochastischen Prozesses (vgl. mit [4, Theorem 7.5] und [8, Theorem 4.14]).

**Satz 1.3.32** Sei  $C[0, 1]$  der Raum der stetigen Funktionen auf  $[0, 1]$ . Weiters seien  $(X_n(t))_{0 \leq t \leq 1}$  eine Folge von stochastischen Prozessen und  $(X(t))_{0 \leq t \leq 1}$  ein stochastischer Prozess auf  $C[0, 1]$ , für die gilt

- (i)  $(X_n(\kappa_1), \dots, X_n(\kappa_d)) \xrightarrow{d} (X(\kappa_1), \dots, X(\kappa_d)), n \rightarrow \infty$  für jedes  $d \in \mathbb{N}$  und  $0 \leq \kappa_1 < \kappa_2 < \dots < \kappa_d$ ;
- (ii) (Straffheit) es gibt eine Konstante  $C > 0$  und Exponente  $\alpha > 1$  und  $\beta > 0$ , sodass  $\mathbb{E}(|X_n(0)|^\beta) \leq C$  und

$$\mathbb{E}(|X_n(s) - X_n(t)|^\beta) \leq C|s - t|^\alpha, \quad \forall s, t \in [0, 1].$$

Dann folgt

$$(X_n(t))_{0 \leq t \leq 1} \xrightarrow{d} (X(t))_{0 \leq t \leq 1}.$$

*Beweis:* Siehe [4, Theorem 7.5]. □

Wir präsentieren einen weiteren Satz, der bei die Konvergenz von Zufallsvariable oder von stochastischen Prozessen eine wesentliche Rolle spielt:

**Satz 1.3.33 (Continuous mapping theorem)** *Seien  $X_n$  eine Folge diskreter Zufallsvariablen und  $X$  eine Zufallsvariable, sodass  $X_n \xrightarrow{d} X$ . Weiters sei  $F : \mathbb{R}^k \mapsto \mathbb{R}^m$  ein fast überall stetiges Funktional (d.h. für die Menge der Unstetigkeitsstellen  $N_F$  von  $F$  gilt  $\mathbb{P}(X \in N_F) = 0$ ). Dann gilt*

$$F(X_n) \xrightarrow{d} F(X). \quad (1.64)$$

*Beweis:* Siehe [33, Theorem 2.3]. □





# Kapitel 2

## Einfach erzeugte Bäume und Pólya-Bäume

Wir kommen nun zum Kern der Arbeit. Zuerst wiederholen wir ein paar wesentliche Definitionen der Graphentheorie.

**Definition 2.0.1** Ein Graph  $G = (V, E)$  ist eine Menge von Knoten  $V$  (engl. vertices) und von Kanten  $E$  (engl. edges), wobei  $E \subseteq V \times V$ . Ist  $e \in E$ , so schreibt man  $e = (u, v)$  bei einem gerichteten Graphen und  $e = uv$  bei einem ungerichteten Graphen.

**Definition 2.0.2** Seien  $u$  und  $v$  zwei Knoten in einem Graph  $G = (V, E)$ . Dann heißt  $v$  Nachbar von  $u$ , wenn  $uv \in E$ . Ist der Graph gerichtet, so wird  $v$  Nachfolger bzw. Vorgänger von  $u$  genannt, falls  $(u, v) \in E$  bzw.  $(v, u) \in E$ .

**Definition 2.0.3** Sei  $G = (V, E)$  ein Graph und  $u \in V$ . Unter dem Grad von  $u$  versteht man die Anzahl aller Nachbarn von  $u$ . Bei einem gerichteten Graph bezeichnet man mit  $d^+(u)$  ( $d^-(u)$ ) den Hingrad (Weggrad) von  $u$ , d.h. die Anzahl aller Nachfolger (Vorgänger) von  $u$ .

Ist  $G = (V, E)$  ein Graph, so wird ein Teilgraph  $K \subseteq G$  der Gestalt  $K := (\{v_1, \dots, v_k\} \in V, \{v_i v_{i+1} \mid i = 1, \dots, k-1\} \cup v_k v_1)$  für  $k \geq 3$  Kreis genannt.

**Definition 2.0.4** Ist ein Graph kreislos und zusammenhängend (d.h. zwischen je zwei Knoten gibt es einen Weg, also eine Folge von Kanten, wobei kein Knoten mehrfach verwendet wird), so spricht man von einem Baum. Besitzt ein Baum als gerichteter Graph einen ausgezeichneten Knoten  $r$  (die sogenannte Wurzel) vom Hingrad  $d^-(r) = 0$ , so spricht man von einem Wurzelbaum.

Ein Knoten vom Hingrad  $d^+(v) = 0$  heißt *Blatt*.

Wir können für Wurzelbäume mit bestimmten Eigenschaften Klassen definieren. Das Ziel dieses Kapitels ist, folgende zwei Klassen von Wurzelbäumen zu vergleichen: die einfach erzeugten Bäume (Teil 2.1) und die Pólya-Bäume (Teil 2.2).

## 2.1 Einfach erzeugte Bäume

Die einfach erzeugten Bäume wurden zum ersten Mal von Meir und Moon in [23] definiert. Es geht um eine wichtige Klasse von Bäume, die unter anderem die ebenen Wurzelbäume oder Catalan-Bäume (Beispiel 2.1.7), die markierten ebenen Wurzelbäume oder Cayley-Bäume (Beispiel 2.1.8), die Binärbäume (Beispiele 2.1.9 und 2.1.10) und die Motzkin-Bäume (Beispiel 2.1.11) umfasst.

### 2.1.1 Grundlagen

In diesem Abschnitt kümmern wir uns um eine kombinatorische Struktur bestehend aus einer Menge  $\mathcal{T}$  und einer Größenfunktion  $|\cdot| : \mathcal{T} \rightarrow \mathbb{N}_0$ . Dazu definieren wir:

**Definition 2.1.1** Sei  $\mathcal{T}$  eine Klasse von Bäumen und  $(\phi_k)_{k \geq 0}$  eine Folge nichtnegativer Ganzzahlen mit  $\phi_0 > 0$  und  $\phi_j > 0$  für ein  $j \geq 2$ . Das Gewicht eines Baumes  $T \in \mathcal{T}$  (siehe u. a. [2, Abschnitt 2,1], [8, Teil 1.2.7], [23, Unterkapitel 2.3] oder [25, Teil 3]) ist dann gegeben durch

$$\omega(T) = \prod_{k \geq 0} \phi_k^{N_k(T)} = \prod_{v \in T} \phi_{d^+(v)}. \quad (2.1)$$

Dabei beschreibt  $N_k(T)$  die Anzahl der Knoten mit  $k$  Kindern und  $d^+(v)$  den Weggrad (d.h die Anzahl der Nachfolger von  $v$ ) des Knotens  $v \in T$ . Wir setzen

$$\Phi(t) = \sum_{k \geq 0} \phi_k t^k \quad (2.2)$$

als die gewöhnliche erzeugende Funktion der Folge  $(\phi_k)_{k \geq 0}$ .

Dazu definieren wir:

**Definition 2.1.2** Eine Familie  $\mathcal{S}$  von Wurzelbäumen ist eine Familie von einfach erzeugten Bäumen (engl.: *simply generated trees*), falls ihre zugehörige erzeugende Funktion

$$S(z) := \sum_{S \in \mathcal{S}} \omega(S) z^{|S|} = \sum_{n \geq 1} s_n z^n \quad (2.3)$$

folgende Funktionalgleichung erfüllt (siehe u.a. [8, Abschnitt 1.2.7], [23, Unterkapitel 2.3] oder [25, Teil 3]):

$$S(z) = z\Phi(S(z)), \quad (2.4)$$

wobei  $\Phi$  die erzeugende Funktion der Folge  $(\phi_k)_{k \geq 0}$  ist.

Die Koeffizienten  $s_n$  beschreiben nicht mehr die Anzahl der Bäume der Größe  $n$ , sondern die Summe der Gewichte von Bäumen der Größe  $n$ , also gilt

$$s_n = \sum_{S \in \mathcal{S}_n} \omega(S) \quad (2.5)$$

wobei  $\mathcal{S}_n = \{S \in \mathcal{S} : |S| = n\}$ . Wegen der Funktionalgleichung und mithilfe der Lagrange-Inversionsformel erhalten wir (vgl. mit [8, Formel (3.2)])

$$s_n = \frac{1}{n} [u^{n-1}] \Phi(u)^n. \quad (2.6)$$

**Definition 2.1.3** Sei  $\mathcal{S}_n$  die Menge aller einfach erzeugten Bäume mit  $n$  Knoten und sei  $S_n \in \mathcal{S}_n$  ein zufälliger Baum. Dann definieren wir mit

$$\mathbb{P}(S_n = T) = \frac{\omega(T)}{\sum_{S \in \mathcal{S}_n} \omega(S)} = \frac{\omega(T)}{s_n} \quad (2.7)$$

eine Wahrscheinlichkeitsverteilung auf  $\mathcal{S}_n$  (siehe [8, Abschnitt 1.2.7] und [23, Formel (2.4)]).

Es sei bemerkt, dass die Folge der Gewichte  $(\phi_k)_{k \geq 0}$  auch als Wahrscheinlichkeitsverteilung interpretiert werden kann (vgl. mit [8, Abschnitt 1.2.7] und [23, Theorem 7.1]).

**Definition 2.1.4** Unter einem Galton-Watson Verzweigungsprozess versteht man eine Zufallsvariable  $\xi$  zusammen mit einer Wahrscheinlichkeitsverteilung  $(\pi_k)_{k \geq 0}$  auf  $\mathbb{N}_0$  (siehe [23, Abschnitt 2.2]). Dabei wird ein Baum wie folgt rekursiv aufgebaut: man startet mit der Wurzel, dann bekommt jeder Knoten in jedem Rekursionsschritt eine gewisse Anzahl von Nachfolgern, die einer unabhängigen Kopie von  $\xi$  entsprechen. Wir erhalten dann folgende Wahrscheinlichkeitserzeugende Funktion:

$$\mathbb{E}(e^{\xi t}) = \sum_{k \geq 0} \pi_k t^k. \quad (2.8)$$

Der entsprechende Baum wird Galton-Watson Baum genannt.  $\xi$  heißt die Nachkommenverteilung (engl. offspring distribution) des Galton-Watson Verzweigungsprozesses.

**Definition 2.1.5** Die erwartete Anzahl der Kinder in einem zufälligen Galton-Watson Baum ist gegeben durch

$$\mathbb{E}(\xi) = \left. \frac{d}{dt} \mathbb{E}(e^{\xi t}) \right|_{t=1} = \sum_{k \geq 0} k \pi_k. \quad (2.9)$$

Dabei heißt der Galton-Watson Verzweigungsprozess

- subkritisch, falls  $\mathbb{E}(\xi) < 1$ ;
- kritisch, falls  $\mathbb{E}(\xi) = 1$ ;
- superkritisch, falls  $\mathbb{E}(\xi) > 1$ ;

Die Beziehung zwischen Galton-Watson Prozessen und einfach erzeugten Bäumen ist die folgende: bezeichne  $\xi$  die von dem Prozess induzierte Zufallsvariable und  $\nu(T)$  die Wahrscheinlichkeit, dass ein Baum  $T$  auftritt. Dann erfüllt die erzeugende Funktion der Folge  $(s_n)_{n \geq 0}$  mit

$$s_n = \mathbb{P}(|T| = n) = \sum_{T: |T|=n} \nu(T)$$

die Funktionalgleichung (2.4) (vgl. mit [8, Abschnitt 1.2.7]) wobei  $\Phi(t) = \mathbb{E}(t^\xi) = \sum_{k \geq 0} \phi_k t^k$ . Dabei ist  $\phi_k = \mathbb{P}(\xi = k)$  und daher  $\nu(T) = \omega(T)$ . Wenn der Galton-Watson Baum  $T$  die Größe  $n$  hat, dann gilt die Wahrscheinlichkeitsverteilung (2.7) auf  $\mathcal{S}_n$  (siehe [8, Abschnitt 1.2.7]). Hier haben wir  $\Phi(1) = 1$ . Wenn allgemeiner  $\Phi(t) < \infty$  gilt, so können wir  $\phi_k$  durch  $\frac{t^k \phi_k}{\Phi(t)}$  ersetzen. Wir bekommen eine neue Wahrscheinlichkeitsverteilung gegeben durch  $\mathbb{P}(\xi = k) = \frac{t^k \phi_k}{\Phi(t)}$  (siehe [23, Lemma 4.1]). Diese erfüllt folgende Eigenschaften.

**Satz 2.1.6** Sei  $(\phi_k)_{k \geq 0}$  eine Folge von Gewichten mit  $\phi_0 > 0$  und  $\phi_j > 0$  für mindestens ein  $j \geq 2$ . Bezeichne mit  $\rho$  den (positiven) Konvergenzradius der erzeugende Funktion  $\Phi(t)$  der Folge. Setze  $\tilde{\Phi}(t) = \frac{t\Phi'(t)}{\Phi(t)}$  und  $\nu = \lim_{t \rightarrow \rho} \tilde{\Phi}(t)$ . Daher gilt:

- ist  $\nu \geq 1$ , dann gibt es eine eindeutig bestimmte Zahl  $\tau \in [0, \rho]$  mit  $\tilde{\Phi}(\tau) = 1$ ;
- ist  $\nu < 1$ , dann setze  $\tau = \rho$ .

In beiden Fällen ist  $\tau \in \mathbb{R}_{\geq 0}$  und  $0 < \Phi(\tau) < \infty$ . Dann wird durch

$$\pi_k := \frac{\tau^k \phi_k}{\Phi(\tau)}, \quad k \geq 0 \quad (2.10)$$

eine Wahrscheinlichkeitsverteilung eines Galton-Watson Verzweigungsprozesses definiert (siehe [8, Abschnitt 1.2.7] und [23, Theorem 7.1]). Der Mittelwert und die Varianz dieses Prozesses sind gegeben durch  $\mu = \min(\nu, 1)$  und

$$\sigma^2 = \tau \tilde{\Phi}'(\tau) = \frac{\tau^2 \Phi''(\tau)}{\Phi(\tau)}. \quad (2.11)$$

*Beweis:* Siehe [23, Unterkapitel 16] □

**Beispiel 2.1.7** Ein ebener Wurzelbaum oder Catalan-Baum kann als ein einfach erzeugter Baum mit  $\phi_k = 1$  interpretiert werden (siehe [23, Example 10.1]). Dabei haben wir

$$\Phi(t) = \sum_{k \geq 0} t^k = \frac{1}{1-t}$$

und daher

$$S(z) = \frac{z}{1-S(z)} = \frac{1 - \sqrt{1-4z}}{2}.$$

Die Koeffizienten sind dann durch

$$s_n = c_{n-1} = \frac{1}{n} \binom{2n-2}{n-1}$$

gegeben. Dabei bezeichnet  $c_n$  die  $n$ -te Catalan-Zahl.

**Beispiel 2.1.8** Ein Cayley-Baum oder markierter ebener Wurzelbaum ist ein markierter einfach erzeugter Baum mit  $\phi_k = 1$  (siehe [23, Example 10.2]). Daher erhalten wir

$$\hat{\Phi}(t) = \sum_{k \geq 0} \frac{t^k}{k!} = e^t$$

sowie

$$\hat{S}(z) = ze^{\hat{S}(z)}.$$

Wir bekommen dann

$$s_n = n^{n-1}.$$

**Beispiel 2.1.9** Ein Binärbaum 1. Art (oder strikter Binärbaum) (siehe [23, Example 10.3]) ist ein Wurzelbaum, dessen Knoten den Weggrad 2 (interne Knoten) oder 0 (externe Knoten oder Blätter) haben können. Es gilt dann  $\phi_0 = \phi_2 = 1$  und  $\phi_j = 0$  sonst. Daher haben wir

$$\Phi(t) = 1 + t^2$$

und

$$S(z) = z(1 + S(z)^2) = \frac{1 - \sqrt{1 - 4z^2}}{2z}.$$

Für die Koeffizienten bekommen wir

$$s_n = \begin{cases} 0, & n = 1 \text{ oder } n \text{ gerade,} \\ c_{(n-1)/2} = \frac{2}{n+1} \binom{n-1}{(n-1)/2}, & n > 1 \text{ ungerade.} \end{cases}$$

**Beispiel 2.1.10** Ein Binärbaum 2. Art (oder unvollständiger Binärbaum) (siehe [8, Example 1.3] oder [23, Example 10.4]) ist ein Wurzelbaum, der aus Knoten, die höchstens zwei Kinder (Nachfolger) besitzen, besteht. Die Folge der Gewichte lautet dann  $\phi_0 = 1$ ,  $\phi_1 = 2$ ,  $\phi_2 = 1$  und  $\phi_j = 0$  für  $j > 2$  also  $\phi_k = \binom{2}{k}$ ,  $k \geq 0$ . Wir bekommen dann

$$\Phi(t) = \sum_{k \geq 0} \binom{2}{k} t^k = (1 + t)^2$$

sowie

$$S(z) = z(1 + S(z))^2 = \frac{1 - \sqrt{1 - 4z}}{2z} - 1$$

und daher

$$s_n = c_n = \frac{1}{n+1} \binom{2n}{n}.$$

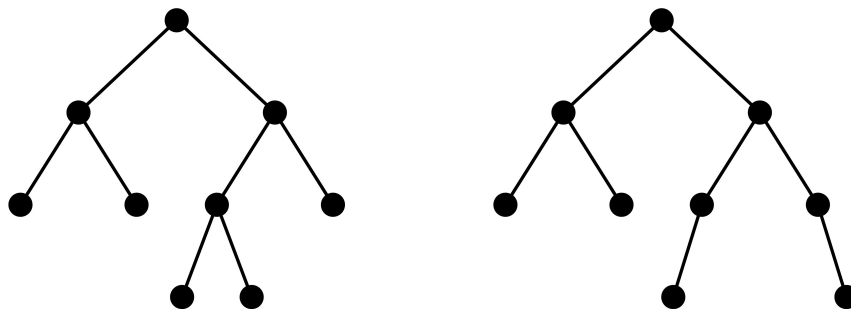


Abbildung 2.1: Ein Binärbaum erster Art (links) und ein Binärbaum zweiter Art (rechts) jeweils mit neun Knoten.

**Beispiel 2.1.11** *Unter einem Motzkin-Baum (siehe [23, Example 10.5]) versteht man einen planaren Wurzelbaum, wobei für jeden Knoten  $v$   $d^+(v) \leq 2$  gilt. Für die Folge  $(\phi_k)_{k \geq 0}$  gilt dann  $\phi_0 = \phi_1 = \phi_2 = 1$  und  $\phi_j = 0 \forall j > 2$ . Daher erhalten wir:*

$$\Phi(t) = 1 + t + t^2$$

und

$$S(z) = z(1 + S(z) + S(z)^2).$$

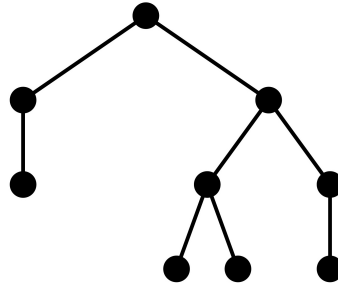


Abbildung 2.2: Ein Motzkin-Baum mit neun Knoten.

## 2.1.2 Singularitätsanalyse

Als Erinnerung erfüllt die erzeugende Funktion  $S(z)$  der einfach erzeugten Bäume folgende Funktionalgleichung:

$$S(z) = z\Phi(S(z)).$$

Die Gewichtsfunktion  $\Phi(t)$  ist bei  $t = 0$  klarerweise analytisch und es gilt  $\Phi(0) \neq 0$ . Daher können wir die Funktionalgleichung wie folgt umschreiben:

$$z = \frac{S(z)}{\Phi(S(z))} =: \Psi(S(z)). \quad (2.12)$$

Andererseits sind alle Gewichte nichtnegativ. Daher erfüllt  $\Phi(t)$  die Bedingungen folgender Proposition (vgl. mit [15, Proposition IV.5] und [25, Theorem 3.1]):

**Proposition 2.1.12** *Sei  $\Phi(t)$  eine bei  $t = 0$  analytische Funktion mit Konvergenzradius  $R \in (0, +\infty]$ , die folgende Eigenschaften besitzt:*

- $\Phi(0) \neq 0$ ;
- Die Koeffizienten von  $\Phi$  sind nichtnegativ;
- Unter der Nebenbedingung  $\lim_{x \rightarrow R^-} \frac{x\Phi'(x)}{\Phi(x)} > 1$  besitzt die Gleichung  $\frac{t\Phi'(t)}{\Phi(t)} = 1$  eine eindeutige Lösung  $\tau \in (0, R)$ ;

Dann ist  $S(z)$  bei  $z = 0$  analytisch, und es gilt

$$[z^n]S(z) \asymp \left(\frac{1}{\rho}\right)^n, \quad \text{wobei} \quad \rho = \frac{\tau}{\Phi(\tau)} = \frac{1}{\Phi'(\tau)}. \quad (2.13)$$

*Beweis:* Siehe [15, Proposition IV.5].

Es folgt, dass  $\rho$  der Konvergenzradius und daher die dominante Singularität von  $S(z)$  ist.

**Satz 2.1.13** *Seien  $\rho$  und  $\tau$  wie in der vorigen Proposition. Dann gilt  $S(\rho) = \tau$ , oder genauer  $\lim_{z \rightarrow \rho} S(z) = \tau$ .*

*Beweis:* Wir betrachten die Funktion  $\Psi(S) = \frac{S}{\Phi(S)}$ . Aus (2.12) zusammen mit (2.13) folgt  $\Psi(\tau) = \rho$ . Für die 1. Ableitung haben wir (mithilfe von (2.13))

$$\Psi'(\tau) = \frac{\Phi(\tau) - \tau\Phi'(\tau)}{\Phi(\tau)^2} = \frac{\Phi(\tau) - \Phi(\tau)}{\Phi(\tau)^2} = 0.$$

Für die zweite Ableitung erhalten wir

$$\Psi''(\tau) = \frac{(\Phi'(\tau) - \Phi'(\tau) - \tau\Phi''(\tau))\Phi(\tau)^2 - 2(\Phi(\tau) - \tau\Phi'(\tau))\Phi(\tau)\Phi'(\tau)}{\Phi(\tau)^4} = -\frac{\tau\Phi''(\tau)}{\Phi(\tau)^2} \neq 0.$$

Daher sind die Bedingungen von Lemma 1.2.29 erfüllt, und der gewünschte Grenzwert gilt für  $S(z)$  als Lösung von  $\Psi(S) = z$ .  $\square$

Interessant ist das Verhalten von  $S(z)$  in der Nähe von  $z = \rho$  (vgl. mit [8, Theorem 3.6] [15, Theorem 6.6] oder [25, Theorem 3.1]):

**Satz 2.1.14** *Die erzeugende Funktion der einfach erzeugten Bäume besitzt folgende Puiseux-Reihe:*

$$S(z) = \tau - d_1 \left(1 - \frac{z}{\rho}\right)^{1/2} + \mathcal{O}\left(1 - \frac{z}{\rho}\right), \quad z \rightarrow \rho \quad (2.14)$$

mit  $d_1 = \sqrt{\frac{2\Phi(\tau)}{\Phi''(\tau)}}$ .

*Beweis:* Wir wenden den Satz 1.2.27 an. Setze dabei

$$F(z, y) = z\Phi(y) - y \quad (2.15)$$

und  $(z_0, y_0) = (\rho, \tau)$ . Es gilt

$$\begin{aligned} F(\rho, \tau) &= \rho\Phi(\tau) - \tau = 0 \\ F_y(\rho, \tau) &= \rho\Phi'(\tau) - 1 = 0 \\ F_z(\rho, \tau) &= \Phi(\tau) \\ F_{yy}(\rho, \tau) &= \rho\Phi''(\tau). \end{aligned}$$

Dann folgt

$$\begin{aligned} y(z) = S(z) &= \tau - \sqrt{\frac{2\rho F_z(\rho, \tau)}{F_{yy}(\rho, \tau)}} \sqrt{1 - \frac{z}{\rho}} + \mathcal{O}\left(1 - \frac{z}{\rho}\right) \\ &= \tau - \sqrt{\frac{2\Phi(\tau)}{\Phi''(\tau)}} \sqrt{1 - \frac{z}{\rho}} + \mathcal{O}\left(1 - \frac{z}{\rho}\right). \end{aligned}$$

□

**Bemerkung 2.1.15** Wir erinnern uns, dass die Varianz der kritischen Nachkommenverteilung eines Galton-Watson Verzweigungsprozesses durch  $\sigma^2 = \frac{\tau^2 \Phi''(\tau)}{\Phi(\tau)}$  gegeben ist.

Daher können wir (2.14) wie folgt umschreiben:

$$S(z) = \tau - \frac{\tau\sqrt{2}}{\sigma} \left(1 - \frac{z}{\rho}\right)^{1/2} + \mathcal{O}\left(1 - \frac{z}{\rho}\right), \quad z \rightarrow \rho. \quad (2.16)$$

Nun können wir einen asymptotischen Ausdruck für die Koeffizienten  $s_n$  berechnen (vgl. mit [8, Theorem 3.6] [15, Theorem 6.6] oder [25, Theorem 3.1])

**Satz 2.1.16** Sei  $d = \text{ggT}\{j > 0 : \phi_j > 0\}$  die Periode. Dann haben die Koeffizienten  $s_n = [z^n]S(z)$  folgendes asymptotische Verhalten:

$$s_n = d \frac{d_1 \rho^{-n}}{2\sqrt{\pi} n^{3/2}} \left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right), \quad n \rightarrow \infty \quad (2.17)$$

falls  $n \equiv 1(d)$ , und  $s_n = 0$  sonst.

*Beweis:* Wir zeigen die Aussage für  $d = 1$  (aperiodischer Fall). Laut Satz 1.2.16 gilt

$$[z^n] \left(1 - \frac{z}{\rho}\right)^{1/2} \sim \frac{n^{-1/2-1}}{\Gamma(-1/2)} \rho^{-n} = \frac{n^{-3/2} \rho^{-n}}{-2\sqrt{\pi}}$$

und laut Satz 1.2.25

$$\mathcal{O}\left([z^n] \left(1 - \frac{z}{\rho}\right)^{3/2}\right) = \mathcal{O}\left(n^{-5/2} \rho^{-n}\right).$$

Aus Satz 2.1.14 haben wir

$$S(z) = \tau - d_1 \sqrt{1 - \frac{z}{\rho}} + d_2 \left(1 - \frac{z}{\rho}\right) + \mathcal{O}\left(\left(1 - \frac{z}{\rho}\right)^{3/2}\right),$$

wobei  $d_1 = \sqrt{\frac{2\Phi(\tau)}{\Phi''(\tau)}}$  ist, und  $\tau, d_2 \left(1 - \frac{z}{\rho}\right)$  keine Rolle mehr für  $n > 1$  spielen. Insgesamt bekommen wir dann

$$\begin{aligned} s_n &= \sqrt{\frac{2\Phi(\tau)}{\Phi''(\tau)}} \frac{\rho^{-n}}{2\sqrt{\pi} n^{3/2}} + \mathcal{O}\left(\frac{\rho^{-n}}{n^{5/2}}\right) \\ &= \sqrt{\frac{2\Phi(\tau)}{\Phi''(\tau)}} \frac{\rho^{-n}}{2\sqrt{\pi} n^{3/2}} \left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right). \end{aligned}$$

□



Als Folgerung betrachten wir folgendes Korollar:

**Korollar 2.1.17** Die Koeffizienten des Ausdrucks  $\frac{zS'(z)}{S(z)}$  erfüllen

$$[z^n] \frac{zS'(z)}{S(z)} \sim \frac{ns_n}{\tau}, \quad n \rightarrow \infty. \quad (2.18)$$

*Beweis:* Aus (2.14) folgt

$$zS'(z) \sim \frac{d_1}{2} \left(1 - \frac{z}{\rho}\right)^{-1/2}, \quad z \rightarrow \rho,$$

und daher

$$\frac{zS'(z)}{S(z)} \sim \frac{d_1}{2\tau} \left(1 - \frac{z}{\rho}\right)^{-1/2}, \quad z \rightarrow \rho.$$

Transfersatz 1.2.16 und (2.17) liefert uns

$$[z^n] \frac{d_1}{2\tau} \left(1 - \frac{z}{\rho}\right)^{-1/2} = \frac{d_1 n^{-1/2} \rho^{-n}}{2\sqrt{\pi}\tau} \sim \frac{ns_n}{\tau}.$$

□

**Bemerkung 2.1.18** Sei  $\sigma^2$  die Varianz der induzierten Verteilung eines Galton-Watson Verzweigungsprozesses. Dann kann man aus (2.17)  $s_n$  wie folgt alternativ darstellen (mit  $d = 1$ ):

$$\frac{\tau \rho^{-n}}{\sqrt{2\pi\sigma n^{3/2}}} \left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right), \quad n \rightarrow \infty. \quad (2.19)$$

**Beispiel 2.1.19** Als erstes Beispiel betrachten wir die ebenen Wurzelbäume (also gilt  $\Phi(t) = \frac{1}{1-t}$  und daher  $S(z) = \frac{1 - \sqrt{1-4z}}{2}$ ). Es ist leicht zu merken, dass  $S(z)$  eine einzige Singularität bei  $z = \frac{1}{4}$  besitzt. Aus (2.13) folgt  $\tau = \frac{1}{2}$ . Es gelten dazu  $\Phi'(\tau) = \frac{1}{(1-\frac{1}{2})^2} = 4$ ,  $\Phi''(\tau) = \frac{2}{(1-\frac{1}{2})^3} = 16$  und  $d_1 = \sqrt{\frac{4}{16}} = \frac{1}{2}$ . Da

$$S(z) = \frac{1}{2} - \frac{1}{2}\sqrt{1-4z} = \tau - d_1\sqrt{1-\rho},$$

erhalten wir

$$s_n = \frac{1}{2} \frac{4^n}{2\sqrt{\pi n^{3/2}}} \left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right) = \frac{4^{n-1}}{\sqrt{\pi n^{3/2}}} \left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right).$$

**Beispiel 2.1.20** Jetzt betrachten wir die Cayley-Bäume (also gilt  $\Phi(t) = e^t$  und  $\hat{S}(z) = ze^{\hat{S}(z)}$ ).  $S(z)$  besitzt ihre einzige Singularität bei  $z = e^{-1}$ . Es folgt  $\Phi'(\tau) = e$  und daher  $\tau = 1$ . Wir erhalten dann

$$S(z) = 1 - \sqrt{\frac{2 \cdot 1 \cdot e}{e}} \sqrt{1-ze} + \mathcal{O}(1-ze) = 1 - \sqrt{2} \sqrt{1-ze} + \mathcal{O}(1-ze)$$

und

$$s_n = \sqrt{2} \frac{e^n}{2\sqrt{\pi n^{3/2}}} \left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right) = \frac{e^n}{\sqrt{2\pi n^{3/2}}} \left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right).$$

**Beispiel 2.1.21** Die erzeugende Funktion  $(S(z) = \frac{1 - \sqrt{1 - 4z^2}}{2z})$  der Binärbäume erster Art (mit  $\Phi(t) = 1 + t^2$ ) besitzt zwei Singularitäten bei  $z_1 = -\frac{1}{2}$  und  $z_2 = \frac{1}{2}$ . Daher ist  $z_2$  die dominante Singularität von  $S(z)$  und wir bekommen (mithilfe von (2.13))  $2 = \Phi'(\tau) = 2\tau$  also  $\tau = 1$ . Daher haben wir

$$S(z) \sim 1 - \sqrt{\frac{2 \cdot 1 \cdot 2}{2}} \sqrt{1 - 2z} = 1 - \sqrt{2} \sqrt{1 - 2z}$$

und

$$s_n = \sqrt{2} \frac{2^n}{2\sqrt{\pi n^{3/2}}} \left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right) = \frac{2^n}{\sqrt{2\pi n^{3/2}}} \left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right).$$

**Beispiel 2.1.22** Die erzeugende Funktion  $(S(z) = \frac{1 - \sqrt{1 - 4z}}{2z} - 1)$  der Binärbäume erster Art (mit  $\Phi(t) = 1 + 2t + t^2$ ) hat ihre dominante Singularität bei  $z = \frac{1}{4}$  und es gilt  $4 = \Phi'(\tau) = 2 + 2\tau$  also  $\tau = 1$ . Wir haben dann

$$S(z) \sim 1 - \sqrt{\frac{2 \cdot 1 \cdot 4}{2}} \sqrt{1 - 4z} = 1 - 2\sqrt{1 - 4z}$$

und

$$s_n = 2 \frac{4^n}{2\sqrt{\pi n^{3/2}}} \left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right) = \frac{4^n}{\sqrt{\pi n^{3/2}}} \left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right).$$

**Beispiel 2.1.23** Für die Motzkin-Bäume gilt  $\Phi(t) = 1 + t + t^2$  und  $S(z) = \frac{1 - z - \sqrt{(1+z)(1-3z)}}{2z}$ . Daher besitzt  $S(z)$  zwei Singularitäten bei  $z_1 = -1$  und  $z_2 = \frac{1}{3}$ . Klarerweise ist  $z_2$  die dominante Singularität, es gilt  $3 = \Phi'(\tau) = 1 + 2\tau$  und daher  $\tau = 1$ . Wir erhalten dann

$$S(z) \sim 1 - \sqrt{\frac{2 \cdot 1 \cdot 3}{2}} \sqrt{1 - 3z} = 1 - \sqrt{3} \sqrt{1 - 3z}$$

und

$$s_n = \sqrt{3} \frac{3^n}{2\sqrt{\pi n^{3/2}}} \left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right).$$

## 2.2 Pólya-Bäume

Pólya-Bäume wurden zum ersten Mal in 1937 von Pólya in [30] präsentiert und analysiert. Wie bei einfach erzeugten Bäumen besitzt die erzeugende Funktion  $P(z)$  der Pólya-Bäume eine implizite Darstellung der Form  $P(z) = z\Psi(P(z))$  mit einem geeigneten Operator  $\Psi$ , der im Satz 2.2.2 definiert wird. Doch hat dieser Operator ganz andere Eigenschaften als  $\Phi$ , und die Bestimmung der Koeffizienten ist viel komplizierter.

### 2.2.1 Grundlagen

Ein Wurzelbaum heißt *eben*, wenn die Kinder der Wurzel eine links-nach-rechts Reihenfolge besitzen.

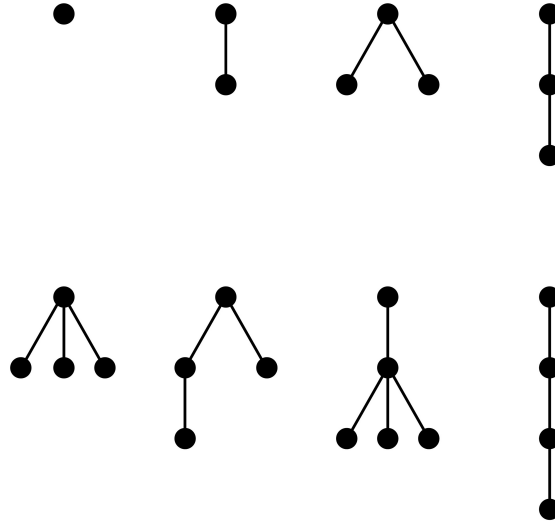


Abbildung 2.3: Sämtliche Pólya-Bäume der Größe  $n = 1, 2, 3, 4$ .

**Definition 2.2.1** *Unter einem Pólya-Baum (oder ungeordnetem Baum) versteht man einen unmarkierten nichtebenen Wurzelbaum.*

**Satz 2.2.2** *Sei  $\mathcal{P}$  die Klasse der Pólya-Bäume. Die zugehörige erzeugende Funktion  $P(z)$  erfüllt dann folgende Funktionalgleichung (siehe u.a [15, Formel (73)] oder [20, Formel (1.1)]):*

$$P(z) = z \exp \left( \sum_{k \geq 1} \frac{P(z^k)}{k} \right). \tag{2.20}$$

*Beweis:* Da Pólya-Bäume eine Wurzel besitzen und nichteben sind, kann man die Klasse  $\mathcal{P}$  wie folgt definieren:

$$\mathcal{P} = \{\circ\} \times M(\mathcal{P}), \tag{2.21}$$

also als eine Multimenge von Bäumen solcher Art, die an der Wurzel hängen (vgl. mit [15, Formel (73)]). Wir erinnern uns, dass

$$M(\mathcal{P}) \simeq \prod_{\alpha \in \mathcal{P}} \{\alpha\}^*.$$

Aus (1.5) erhalten wir (mit  $M(z)$  als erzeugende Funktion von  $M(\mathcal{P})$ )

$$M(z) = \exp \left( \sum_{k \geq 1} \frac{P(z^k)}{k} \right)$$

und insgesamt

$$P(z) = zM(z) = z \exp \left( \sum_{k \geq 1} \frac{P(z^k)}{k} \right).$$

□

**Bemerkung 2.2.3** Wir können die Funktion  $M(z)$  alternativ wie folgt bestimmen: Betrachte  $\Gamma$  als Untergruppe der symmetrischen Gruppe  $\text{Sym}_n$ ,  $n \in \mathbb{N}$ . Dann ist der Zykluszeiger von  $\Gamma$  definiert durch

$$Z(\Gamma, x_1, \dots, x_n) = \frac{1}{|\Gamma|} \sum_{\pi \in \Gamma} x_1^{l_1(\pi)} x_2^{l_2(\pi)} \dots x_n^{l_n(\pi)}, \quad (2.22)$$

wobei  $l_i(\pi)$  ( $i = 1, \dots, n$ ) die Anzahl der Zyklen der Länge  $i$  in der Permutation  $\pi \in \Gamma$  bezeichnet (vgl. mit [28, Teil 3.3.2] oder [29, Unterkapitel 2.2]).

Setze nun  $\Gamma = \text{Sym}_n$ . Wir definieren  $a_{l_1, l_2, \dots, l_n}$  als die Anzahl der Permutationen in  $\text{Sym}_n$  mit  $l_i$  Zyklen der Länge  $i$ . Es gilt  $n = \sum_{i=1}^n i l_i$ ,

$$a_{l_1, l_2, \dots, l_n} = \frac{n!}{l_1! l_2! \dots l_n! 1^{l_1} 2^{l_2} \dots n^{l_n}}$$

und daher (mit  $|\text{Sym}_n| = n!$  und  $\mathbf{l} = (l_1, l_2, \dots, l_n)$ )

$$\begin{aligned} Z(\text{Sym}_n, x_1, \dots, x_n) &= \frac{1}{n!} \sum_{\mathbf{l} \in \mathbb{N}_0^n, \sum_{i=1}^n i l_i = n} a_{l_1, l_2, \dots, l_n} x_1^{l_1} x_2^{l_2} \dots x_n^{l_n} \\ &= \sum_{\mathbf{l} \in \mathbb{N}_0^n, \sum_{i=1}^n i l_i = n} \prod_{i=1}^n \frac{x_i^{l_i}}{l_i! i^{l_i}}. \end{aligned}$$

Aufsummieren und  $P(z), P(z^2), \dots, P(z^n)$  einsetzen liefert uns

$$\begin{aligned} M(z) &= \sum_{n \geq 1} Z(\text{Sym}_n, P(z), P(z^2), \dots, P(z^n)) \\ &= \sum_{n \geq 1} \left( \sum_{\mathbf{l} \in \mathbb{N}_0^n, \sum_{i=1}^n i l_i = n} \prod_{i=1}^n \frac{P(z^i)^{l_i}}{l_i! i^{l_i}} \right) \\ &= \prod_{i \geq 1} \left( \sum_{l_i \geq 0} \frac{\left( \frac{P(z^i)}{i} \right)^{l_i}}{l_i!} \right) \\ &= \prod_{i \geq 1} \exp \left( \frac{P(z^i)}{i} \right) \\ &= \exp \left( \sum_{i \geq 1} \frac{P(z^i)}{i} \right). \end{aligned}$$

□

Die Koeffizienten  $p_n$  besitzen keine explizite Darstellung. Doch können wir mithilfe der Funktionalgleichung (2.20) einen rekursiven Ausdruck berechnen (siehe [16, Fact 4] und [20, Abschnitt 1.1]):

**Satz 2.2.4** Die Anzahl  $p_n$  der Pólya-Bäume mit  $n$  Knoten ist durch

$$p_n = \begin{cases} n, & n \in \{0, 1\} \\ \frac{1}{n-1} \sum_{i=1}^{n-1} p_{n-i} \sum_{m|i} mp_m, & n > 1 \end{cases} \quad (2.23)$$

gegeben.

*Beweis:* Mithilfe von (2.20) differenzieren wir  $P(z)$  nach  $z$ . Wir erhalten

$$P'(z) = \left( 1 + \sum_{i \geq 1} z^i P'(z^i) \right) \exp \left( \sum_{i \geq 1} \frac{P(z^i)}{i} \right).$$

Wir betrachten dann die logarithmische Ableitung von  $P(z)$  und wir bekommen

$$\begin{aligned} z \frac{P'(z)}{P(z)} &= 1 + \sum_{i \geq 1} z^i P'(z^i) \\ &= 1 + \sum_{i \geq 1} \sum_{m \geq 1} mp_m z^{mi} \\ &= 1 + \sum_{m \geq 1} mp_m \frac{z^m}{1 - z^m}. \end{aligned}$$

Daher gilt

$$zP'(z) = P(z) \left( 1 + \sum_{m \geq 1} mp_m \frac{z^m}{1 - z^m} \right)$$

und beim Koeffizientenvergleich

$$\begin{aligned} np_n &= p_n + \sum_{i=1}^{n-1} \left( [z^i] \sum_{m \geq 1} mp_m \frac{z^m}{1 - z^m} \right) p_{n-i} \\ \Leftrightarrow (n-1)p_n &= \sum_{i=1}^{n-1} \left( [z^i] \sum_{m \geq 1} mp_m \frac{z^m}{1 - z^m} \right) p_{n-i} \\ &= \sum_{i=1}^{n-1} \left( \sum_{m|i} mp_m \right) p_{n-i}. \end{aligned}$$

Dividieren beider Seiten durch  $(n-1)$  liefert uns das gewünschte Ergebnis.  $\square$

## 2.2.2 Singularitätsanalyse

Als Erinnerung erfüllt die erzeugende Funktion der Pólya-Bäume folgende Funktionalgleichung:

$$P(z) = z \exp \left( \sum_{k \geq 1} \frac{P(z^k)}{k} \right).$$

Wir können die Gleichung wie folgt aufschreiben (siehe [15, Proposition VII.5]):

$$P(z) = \zeta(z)e^{P(z)} \quad \text{mit} \quad \zeta(z) = z \exp \left( \sum_{k \geq 2} \frac{P(z^k)}{k} \right). \quad (2.24)$$

Mit dieser Überlegung bekommen wir dann eine alternative Darstellung für die EF der Pólya-Bäume (vgl. u.a. mit [8, Theorem 3.8], [15, Proposition VII.5], [16, Theorem 2] oder [20, Formel (1.3)]):

**Satz 2.2.5** *Die erzeugende Funktion  $P(z)$  der Pólya-Bäume hat den Konvergenzradius  $\rho \approx 0.3383219$  und besitzt daher folgende Puiseux-Reihe:*

$$P(z) = 1 - d_1 \left( 1 - \frac{z}{\rho} \right)^{1/2} + \mathcal{O} \left( \left( 1 - \frac{z}{\rho} \right) \right), \quad z \rightarrow \rho, \quad (2.25)$$

wobei  $d_1 = \sqrt{2\rho e\zeta'(\rho)} \approx 2.6811266$ .

*Beweis:* Betrachte zuerst die exponentiell erzeugende Funktion  $C(z)$  der Cayley-Bäume. Wir wissen schon, dass sie die Lösung folgender Funktionalgleichung

$$C(z) = ze^{C(z)}$$

ist. Aus Beispiel 2.1.20 ist sie analytisch für  $|z| < e^{-1} \approx 0.36787944117$ . Andererseits ist die Anzahl der Pólya-Bäume durch die Anzahl der ebenen Wurzelbäume nach unten beschränkt. Aus Beispiel 2.1.19 ist der Konvergenzradius der zugehörigen erzeugenden Funktion gleich  $\frac{1}{4}$ . Insgesamt erhalten wir  $\frac{1}{4} \leq \rho \leq e^{-1}$ . Wegen (2.24) gilt

$$P(z) = C(\zeta(z)). \quad (2.26)$$

$\zeta(z)$  ist analytisch für  $|z| < \rho^{1/2}$ . Da  $\rho < 1$  gilt, ist der Konvergenzkreis von  $P(z)$  im Konvergenzkreis von  $\zeta(z)$  enthalten, und  $\zeta(z)$  ist daher bei  $z = \rho$  analytisch. Weiters hat Otter in [27] gezeigt, dass  $P(\rho) = 1$ . Nun setzen wir  $F(z, y) = e^y \zeta(z) - y$ . Es gelten (mit  $z_0 = \rho$  und  $y_0 = P(\rho) = 1$ ):

$$\begin{aligned} F(\rho, P(\rho)) &= 0 \\ F_y(\rho, P(\rho)) &= e^{P(\rho)} \zeta(\rho) - 1 = P(\rho) - 1 = 1 - 1 = 0 \\ F_z(\rho, P(\rho)) &= e^{P(\rho)} \zeta'(\rho) = e\zeta'(\rho) \neq 0 \\ F_{yy}(\rho, P(\rho)) &= e^{P(\rho)} \zeta(\rho) = P(\rho) = 1. \end{aligned}$$

Damit sind die Bedingungen von Satz 1.2.27 erfüllt, und wir bekommen

$$\begin{aligned} y(z) = P(z) &= P(\rho) - \sqrt{\frac{2\rho F_z(\rho, P(\rho))}{F_{yy}(\rho, P(\rho))}} \left(1 - \frac{z}{\rho}\right)^{1/2} + \mathcal{O}\left(1 - \frac{z}{\rho}\right) \\ &= 1 - \sqrt{2\rho e\zeta'(\rho)} \left(1 - \frac{z}{\rho}\right)^{1/2} + \mathcal{O}\left(1 - \frac{z}{\rho}\right). \end{aligned}$$

□

**Satz 2.2.6** Für die Anzahl  $p_n = [z^n]P(z)$  der Pólya-Bäume gilt

$$p_n = \frac{d_1 \rho^{-n}}{2\sqrt{\pi n^{3/2}}} \left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right). \quad (2.27)$$

*Beweis:* Eine ähnliche Argumentation wie im Satz 2.1.16 liefert uns das gewünschte Ergebnis. □

Für die ganze Asymptotik von  $P(z)$  bzw.  $p_n$  siehe [16, Theorem 2] bzw. [16, Theorem 3]. Wir fassen unsere Eigenschaften zusammen. Insgesamt haben wir

Bäume	Dom. Singularität	Koeffizienten
Einfach erzeugt ( $d = 1$ )	$\rho = \frac{\tau}{\Phi(\tau)}$ mit $\tau\Phi'(\tau) = \Phi(\tau)$	$s_n \sim \frac{d_1 \rho^{-n}}{2\sqrt{\pi n^{3/2}}}$ mit $d_1 = \sqrt{\frac{2\Phi(\tau)}{\Phi''(\tau)}}$
Catalan	$\rho = \frac{1}{4}$ ( $\tau = \frac{1}{2}$ )	$s_n \sim \frac{4^{n-1}}{\sqrt{\pi n^{3/2}}}$
Cayley	$\rho = e^{-1}$ ( $\tau = 1$ )	$s_n \sim \frac{\sqrt{2}e^{-n}}{\sqrt{\pi n^{3/2}}}$
Binär 1	$\rho = \frac{1}{2}$ ( $\tau = 1$ )	$s_n \sim \frac{2^n}{\sqrt{2\pi n^{3/2}}}$
Binär 2	$\rho = \frac{1}{4}$ ( $\tau = 1$ )	$s_n \sim \frac{4^n}{\sqrt{\pi n^{3/2}}}$
Motzkin	$\rho = \frac{1}{3}$ ( $\tau = 1$ )	$s_n \sim \frac{\sqrt{3} \cdot 3^n}{2\sqrt{\pi n^{3/2}}}$
Pólya	$\rho \approx 0.3383219$	$p_n \sim \frac{d_1 \rho^{-n}}{2\sqrt{\pi n^{3/2}}}$ mit $d_1 \approx 2.6811266$

## 2.3 Zusammenhang zwischen einfach erzeugten Bäumen und Pólya-Bäumen

Es gibt einen Zusammenhang zwischen einfach erzeugten Bäumen und Pólya-Bäumen. Dies wurde von Genitrini in [16] und Gittenberger, Yu Jin und Wallner in [20] gezeigt. Doch ist es leicht zu zeigen, dass Pólya-Bäume nicht einfach erzeugt sind (vgl. [11, Theorem 1]):

**Satz 2.3.1** Pólya-Bäume sind nicht einfach erzeugt.

*Beweis:* Angenommen die Pólya-Bäume wären einfach erzeugt, dann gäbe es eine Potenzreihe  $\Phi(t)$  bestehend aus nichtnegativen Koeffizienten, sodass

$$P(z) = z\Phi(P(z)). \quad (2.28)$$

Da  $p_0 = 0$  und  $p_1 \neq 0$  ist  $P(z)$  invertierbar, und wegen  $p_1 = 1$  gilt sogar  $P(z) \sim z$  und  $P^{-1}(z) \sim z$  für  $z \rightarrow 0$ . Es folgt  $\Phi(0) = 1$ . Mithilfe von (2.28) haben wir (mit  $P(z) = y$  und daher  $z = P^{-1}(P(z)) = P^{-1}(y)$ )

$$y = P^{-1}(y)\Phi(y) \Leftrightarrow \Phi(y) = \frac{y}{P^{-1}(y)}.$$

Diese Reihe besitzt aber negative Koeffizienten, was die Nichtnegativität der Gewichte verletzt.  $\square$

Bevor wir einen Zusammenhang zwischen einfach erzeugten Bäumen erhalten, brauchen wir weitere Definitionen (vgl. mit [20, Teil 2]):

**Definition 2.3.2** Sei  $\mathcal{C}_1$  die Klasse der Cayley-Bäume, also der markierten ebenen Wurzelbäume. Aus Beispiel 2.1.8 ist die zugehörige exponentiell erzeugende  $C(z)$  durch

$$C(z) = \sum_{n \geq 0} n^{n-1} \frac{z^n}{n!} \quad (2.29)$$

gegeben. Dazu definieren wir eine weitere Klasse von einfach erzeugten Bäume  $\mathcal{C}_2$ , insbesondere von unmarkierten ebenen Wurzelbäumen mit Gewichten  $\phi_k = \frac{1}{k!}$ . Die zugehörige erzeugende Funktion ist dann  $C(z)$ .

**Definition 2.3.3** Sei  $\mathcal{C}_3$  die Klasse der C-Bäume, also von unmarkierten nichtebenen Wurzelbäumen mit Gewichtsfunktion  $\omega(T) = e(T) \prod_{k \geq 0} \left(\frac{1}{k!}\right)^{N_k(T)}$  ( $T \in \mathcal{C}_3$ ), wobei  $e(T)$  die Anzahl der Einbettungen von  $T$  auf eine Ebene bezeichnet. Die zugehörige erzeugende Funktion ist dann auch  $C(z)$  (Beweis in [20, Lemma 1]).

Um eine alternative Darstellung für die Funktion  $P(z)$  zu finden, brauchen wir eine zusätzliche kombinatorische Struktur (siehe [20, Definition 2]):

**Definition 2.3.4** Als kombinatorische Struktur der D-Wälder betrachten wir die Klasse  $M_{\geq 2}(\mathcal{P})$  aller Multimengen von Pólya-Bäumen, für welche jede Komponente zumindest verdoppelt ist. Daher kann man eine zugehörige Gewichtsfunktion wie folgt definieren:

$$\omega(F) = \frac{|\{\sigma \in \text{Aut}(F) : \sigma_1 = 0\}|}{|\text{Aut}(F)|}, \quad F \in M_{\geq 2}(\mathcal{P}), \quad (2.30)$$

wobei  $\text{Aut}(F)$  die Automorphismengruppe von  $F$  und  $\sigma_i$  die Anzahl der Zyklen der Länge  $i$  im Automorphismus  $\sigma \in \text{Aut}(F)$  bezeichnen.

Damit erhalten wir folgende erzeugende Funktion (vgl. mit [20, Theorem 2]):

**Satz 2.3.5** Für die erzeugende Funktion  $D(z) = \sum_{n \geq 0} d_n z^n = \sum_{F \in M_{\geq 2}(\mathcal{P})} \omega(F) z^{|F|}$  gilt

$$D(z) = \exp \left( \sum_{i \geq 2} \frac{P(z^i)}{i} \right). \quad (2.31)$$



*Beweis:* Wir können eine ähnliche Überlegung wie in Bemerkung 2.2.3 anwenden. Bezeichne wieder  $Z(\text{Sym}_n, x_1, x_2, \dots, x_n)$ ,  $n \in \mathbb{N}$  den Zyklenzeiger der symmetrischen Gruppe  $\text{Sym}_n$ . Da wir aus der Definition von  $M_{\geq 2}(\mathcal{P})$  keinen Zyklus der Länge eins haben, gilt

$$D(z) = \sum_{n \geq 0} Z(\text{Sym}_n, 0, P(z^2), \dots, P(z^n)) = \prod_{i \geq 2} \exp\left(\frac{P(z^i)}{i}\right) = \exp\left(\sum_{i \geq 2} \frac{P(z^i)}{i}\right).$$

□

**Satz 2.3.6** *Die erzeugende Funktion  $P(z)$  der Pólya-Bäume besitzt folgende alternative Darstellung:*

$$P(z) = C(zD(z)), \tag{2.32}$$

wobei  $C(z)$  bzw.  $D(z)$  die erzeugende Funktion der  $C$ -Bäume bzw. der  $D$ -Wälder bezeichnet.

*Beweis:* Vergleiche mit (2.26), wobei  $D(z) = z\zeta(z)$ .

□



# Kapitel 3

## Vergleiche von Parametern in einfach erzeugten Bäumen und Pólya-Bäumen

Im Unterkapitel 1.3 haben wir gesehen, dass kombinatorische Klassen als zufällige Strukturen betrachtet werden können, indem man sogenannte Parameter (siehe Definitionen 1.3.1 und 1.3.2) definiert. Damit werden multivariate erzeugende Funktionen aufgebaut, um eine Wahrscheinlichkeitsverteilung zu erhalten und Momente zu berechnen. Zweck dieses Kapitel ist, Parameter für einfach erzeugte Bäume und Pólya-Bäume zu betrachten, und die entsprechenden Ergebnisse zu vergleichen. Im ersten Abschnitt werden wir mit Parametern, die eine bestimmte Funktionalgleichung erfüllen (*additive Parameter*), arbeiten. Dann werden wir im Teil 3.2 bestimmte Parameter der Struktur der Bäume analysieren. Im Abschnitt 3.3 werden wir Ergebnisse für weitere Parameter berechnen.

### 3.1 Additive Parameter

*Additive Parameter* sind Parameter, die eine bestimmte Funktionalgleichung erfüllen (Definition 3.1.1). Im Abschnitt 3.1.1 betrachten wir allgemeine Ergebnisse, die von Wagner in [34] gezeigt wurden. In den Teilen 3.2 und 3.3 berechnen wir Ergebnisse für die folgenden zwei additive Parameter: die Anzahl der Knoten vom Weggrad  $k$  ( $k \in \mathbb{N}_0$ ) und die interne Pfadlänge.

#### 3.1.1 Definition und allgemeine Ergebnisse

**Definition 3.1.1** Sei  $\mathcal{T}$  eine Struktur von Wurzelbäumen, und  $T \in \mathcal{T}$  ein Wurzelbaum mit von der Wurzel hängenden Teilbäumen  $T_1, T_2, \dots, T_k$ . Unter einem additiven Baum-Parameter versteht man einen Parameter  $F : \mathcal{T} \rightarrow \mathbb{N}_0$ , der folgende Gleichung erfüllt:

$$F(T) = \sum_{i=1}^k F(T_i) + f(T) \tag{3.1}$$

Dabei wird  $f$  Maut-Funktion genannt.

**Bemerkung 3.1.2** Bei  $T = \{\circ\}$  gilt offensichtlich  $F(T) = f(T)$ .

Prominente Beispiele für Parameter sind:

- die *Anzahl der Blätter* und allgemeiner *der Knoten vom Weggrad  $k$*  (wird im Abschnitt 3.1.2 präsentiert, siehe auch [8, Teil 3.2], [10] oder [17]);
- die *interne Pfadlänge* (wird im Abschnitt 3.1.3 dargestellt, siehe auch Example III.15, III.20 und Example VII.9 in [15] oder [35]);
- die *Anzahl der vollen Teilbäume von  $T$*  (d.h. Teilbäume bestehend aus einem Knoten und seinen Nachfolgern in  $T$ ; siehe [34]). Dabei ist die Maut-Funktion gegeben durch:

$$f(T) = \begin{cases} 1, & \text{falls } |T| = k \\ 0, & \text{sonst} \end{cases}$$

- das *Log-Produkt der Größen der Teilbäume von  $T$*  (siehe [26] oder [34]) mit Maut-Funktion  $f(T) = \log |T|$ .

### Einfach erzeugte Bäume

Wir rufen in Erinnerung, dass

$$S(z) = \sum_{S \in \mathcal{S}} \omega(S) z^{|S|} = z\Phi(S(z)),$$

und definieren die entsprechende bivariate erzeugende Funktion durch

$$S(z, u) = \sum_{S \in \mathcal{S}} \omega(S) z^{|S|} u^{F(S)}. \quad (3.2)$$

Daher gilt

$$\begin{aligned} z\Phi(S(z, u)) &= z \sum_{j \geq 0} \phi_j S(z, u)^j \\ &= z \sum_{j \geq 0} \phi_j \sum_{S_1, S_2, \dots, S_j} \prod_{i=1}^j \omega(S_i) z^{|S_i|} u^{F(S_i)} \\ &= \sum_{S \in \mathcal{S}} \omega(S) z^{|S|} u^{F(S) - f(S)}, \end{aligned}$$

und

$$S(z, 1) = z\Phi(S(z, 1)). \quad (3.3)$$

Wir betrachten additive Parameter mit hinreichend kleiner Maut-Funktion (siehe [34, Theorem 2.1]):

**Satz 3.1.3** *Sei  $F$  ein additiver Parameter mit Maut-Funktion  $f$ , und  $\mathcal{S}$  eine Familie von einfach erzeugten Bäumen, für die gilt*

$$\frac{\sum_{S \in \mathcal{S}_n} \omega(S) |f(S)|}{\sum_{S \in \mathcal{S}_n} \omega(S)} = \mathcal{O}(c^n), \quad c \in (0, 1). \quad (3.4)$$

Ist  $S_n$  ein zufälliger einfach erzeugter Baum aus  $\mathcal{S}_n$ , so haben wir

$$\mathbb{E}(F(S_n)) = \mu n + \mathcal{O}(1), \quad (3.5)$$

mit  $\mu = \frac{1}{\tau} \sum_{S \in \mathcal{S}} \omega(S) f(S) \rho^{|S|}$ , und

$$\mathbb{V}(F(S_n)) = \sigma^2 n + \mathcal{O}(1), \quad (3.6)$$

mit  $\sigma^2 = \mu^2 \left(1 - \frac{\Phi(\tau)}{\tau^2 \Phi''(\tau)}\right) + \frac{1}{\tau} \sum_{S \in \mathcal{S}} \omega(S) f(S) (2F(S) - f(S)) \rho^{|S|} - \frac{2\mu}{\tau} \sum_{S \in \mathcal{S}} \omega(S) f(S) |S| \rho^{|S|}$   
wobei  $\tau$  und  $\rho$  wie in Proposition 2.1.12 sind.

Gilt zusätzlich  $\sigma^2 \neq 0$ , so folgt

$$\frac{F(S_n) - \mathbb{E}(F(S_n))}{\sqrt{\mathbb{V}(F(S_n))}} \xrightarrow{d} N \sim \mathcal{N}(0, 1). \quad (3.7)$$

*Beweis:* Wir leiten  $S(z, u)$  und  $z\Phi(S(z, u))$  nach  $u$  ab und setzen  $u = 1$ . Einerseits gilt:

$$S_u(z, 1) = \sum_{S \in \mathcal{S}} \omega(S) F(S) z^{|S|}$$

Andererseits haben wir:

$$\begin{aligned} \frac{\partial}{\partial u} (z\Phi(S(z, u))) \Big|_{u=1} &= z\Phi'(S(z, 1)) S_u(z, 1) \\ &= \frac{\partial}{\partial u} \left( \sum_{S \in \mathcal{S}} \omega(S) z^{|S|} u^{F(S)-f(S)} \right) \Big|_{u=1} \\ &= \sum_{S \in \mathcal{S}} \omega(S) (F(S) - f(S)) z^{|S|} \\ &= \sum_{S \in \mathcal{S}} \omega(S) F(S) z^{|S|} - \sum_{S \in \mathcal{S}} \omega(S) f(S) z^{|S|} \\ &= S_u(z, 1) - \sum_{S \in \mathcal{S}} \omega(S) f(S) z^{|S|} \end{aligned}$$

Daher bekommen wir (mit  $H(z) := \sum_{S \in \mathcal{S}} \omega(S) f(S) z^{|S|}$ )

$$S_u(z, 1) = \frac{H(z)}{1 - z\Phi'(S(z, 1))}.$$

Jetzt leiten wir  $z\Phi(S(z, 1))$  nach  $z$  ab:

$$\frac{\partial}{\partial z} (z\Phi(S(z, 1))) = \Phi(S(z, 1)) + z\Phi'(S(z, 1)) S_z(z, 1) = \frac{S(z, 1)}{z} + z\Phi'(S(z, 1)) S_z(z, 1)$$

Wegen (2.4) und  $S(z, 1) = S(z)$  gilt auch

$$\frac{\partial}{\partial z} (z\Phi(S(z, 1))) = \frac{d}{dz} (z\Phi(S(z))) = S'(z).$$

Nun haben wir

$$S'(z) = \frac{S(z)}{z} + z\Phi'(S(z))S'(z) \quad \Leftrightarrow \quad z\Phi'(S(z)) = 1 - \frac{S(z)}{zS'(z)}.$$

Insgesamt gilt

$$\begin{aligned} S_u(z, 1) &= \frac{H(z)}{1 - z\Phi'(S(z, 1))} \\ &= \frac{H(z)}{1 - \left(1 - \frac{S(z)}{zS'(z)}\right)} \\ &= \frac{zS'(z)}{S(z)} H(z). \end{aligned}$$

Wir erinnern uns (vgl. Korollar 2.17), dass

$$[z^n] \frac{zS'(z)}{S(z)} \sim \frac{ns_n}{\tau}, \quad n \rightarrow \infty.$$

Wegen (3.4) hat die Funktion  $H(z)$  einen größeren Konvergenzkreis als  $S(z)$ . Dann ist  $H(z)$  bei  $z = \rho$  analytisch, und es folgt

$$[z^n] S_u(z, 1) \sim \frac{ns_n}{\tau} H(\rho), \quad n \rightarrow \infty$$

und insgesamt gilt

$$\begin{aligned} \mathbb{E}(F(S_n)) &= \frac{[z^n] S_u(z, 1)}{[z^n] S(z, 1)} \\ &\sim \frac{\frac{ns_n}{\tau} H(\rho)}{s_n} \\ &= \frac{1}{\tau} H(\rho)n. \end{aligned}$$

Für die Varianz muss man die zweite partielle Ableitung nach  $u$   $S_{uu}(z, u)$  bestimmen, und dann  $u = 1$  setzen. Die Grenzverteilung folgt direkt aus Satz 1.3.24.  $\square$

## Pólya-Bäume

Wir wissen schon, dass

$$P(z) = \sum_{P \in \mathcal{P}} z^{|P|} = z \exp \left( \sum_{k \geq 1} \frac{P(z^k)}{k} \right).$$

Die entsprechende bivariate erzeugende Funktion lautet dann

$$P(z, u) = \sum_{P \in \mathcal{P}} z^{|P|} u^{F(P)}. \quad (3.8)$$

Weiters gilt (vgl. mit [34, Teil 3.1]):

$$z \exp \left( \sum_{k \geq 1} \frac{P(z^k, u^k)}{k} \right) = \sum_{P \in \mathcal{P}} z^{|P|} u^{F(P) - f(P)}. \quad (3.9)$$

Ähnlich wie im vorigen Abschnitt betrachten wir einen additiven Parameter mit hinreichend kleiner Maut-Funktion (vgl. mit [34, Theorem 3.1]).

**Satz 3.1.4** Sei  $F$  ein additiver Parameter mit Maut-Funktion  $f$ , für die gilt

$$\frac{\sum_{P \in \mathcal{P}_n} |f(P)|}{p_n} = \mathcal{O}(c^n), \quad c \in (0, 1) \quad (3.10)$$

wobei  $\mathcal{P}_n$  die Klasse der Pólya-Bäume mit  $n$  Knoten und  $p_n$  ihre Kardinalität bezeichnen. Weiters sei  $P_n \in \mathcal{P}_n$  ein zufälliger Pólya-Baum. Dann gilt:

$$\mathbb{E}(F(P_n)) = \mu n + \mathcal{O}(1) \quad (3.11)$$

mit  $\mu = \frac{\sum_{P \in \mathcal{P}} \left( f(P) \rho^{|P|} + F(P) \frac{\rho^{2|P|}}{1 - \rho^{|P|}} \right)}{1 + \sum_{n \geq 1} n p_n \frac{\rho^{2n}}{1 - \rho^n}}$  und

$$\mathbb{V}(F(P_n)) = \sigma^2 n + \mathcal{O}(1) \quad (3.12)$$

wobei  $\sigma^2 \geq 0$  (eine explizite aber komplizierte Formel kann berechnet werden, siehe [34, Remark 4]).

Ist zusätzlich  $\sigma^2 \neq 0$ , so folgt

$$\frac{F(P_n) - \mathbb{E}(F(P_n))}{\sqrt{\mathbb{V}(F(P_n))}} \xrightarrow{d} N \sim \mathcal{N}(0, 1) \quad (3.13)$$

*Beweis:* Analog wie bei Satz 3.1.3 leiten wir  $P(z, u)$  und  $z \exp \left( \sum_{k \geq 1} \frac{P(z^k, u^k)}{k} \right)$  nach  $u$  ab, und setzen  $u = 1$ . Einerseits gilt

$$P_u(z, 1) = \sum_{P \in \mathcal{P}} F(P) z^{|P|}$$

und andererseits

$$\begin{aligned} \frac{\partial}{\partial u} \left( z \exp \left( \sum_{k \geq 1} \frac{P(z^k, u^k)}{k} \right) \right) \Big|_{u=1} &= P(z, 1) \left( P_u(z, 1) + \sum_{k \geq 2} P_u(z^k, 1) \right) \\ &= \frac{\partial}{\partial u} \left( \sum_{P \in \mathcal{P}} z^{|P|} u^{F(P) - f(P)} \right) \Big|_{u=1} \\ &= \sum_{P \in \mathcal{P}} (F(P) - f(P)) z^{|P|} \\ &= \sum_{P \in \mathcal{P}} F(P) z^{|P|} - \sum_{P \in \mathcal{P}} f(P) z^{|P|} \\ &= P_u(z, 1) - \sum_{P \in \mathcal{P}} f(P) z^{|P|} \end{aligned}$$

also (mit  $H(z) := \sum_{P \in \mathcal{P}} f(P)z^{|P|}$ )

$$P_u(z, 1) = \frac{H(z) + P(z, 1) \sum_{k \geq 2} P_u(z^k, 1)}{1 - P(z, 1)}.$$

Außerdem gilt

$$\begin{aligned} zP_z(z, 1) &= P(z, 1) \left( 1 + \sum_{k \geq 1} z^k P_z(z^k, 1) \right) \\ &= P(z, 1) \left( 1 + zP_z(z, 1) + \sum_{k \geq 2} z^k P_z(z^k, 1) \right) \\ &= \frac{P(z, 1) \left( 1 + \sum_{k \geq 2} z^k P_z(z^k, 1) \right)}{1 - P(z, 1)} \\ \Leftrightarrow \frac{1}{1 - P(z, 1)} &= \frac{zP_z(z, 1)}{P(z, 1) \left( 1 + \sum_{k \geq 2} z^k P_z(z^k, 1) \right)} \end{aligned}$$

und daher

$$\begin{aligned} P_u(z, 1) &= zP_z(z, 1) \frac{H(z) + P(z, 1) \sum_{k \geq 2} P_u(z^k, 1)}{P(z, 1) \left( 1 + \sum_{k \geq 2} z^k P_z(z^k, 1) \right)} \\ &= zP'(z) \frac{H(z) + P(z) \sum_{k \geq 2} P_u(z^k, 1)}{P(z) \left( 1 + \sum_{k \geq 2} z^k P'(z^k) \right)} \\ &= zP'(z) \frac{H(z) + P(z) \sum_{k \geq 2} \sum_P F(P) z^{k|P|}}{P(z) \left( 1 + \sum_{k \geq 2} \sum_{n \geq 1} np_n (z^n)^k \right)} \\ &= \frac{zP'(z)}{P(z)} \frac{\sum_P \left( f(P) z^{|P|} + P(z) F(P) \sum_{k \geq 2} z^{k|P|} \right)}{1 + \sum_{n \geq 1} np_n \sum_{k \geq 2} (z^k)^n} \\ &= \frac{zP'(z)}{P(z)} \frac{\sum_P \left( f(P) z^{|P|} + P(z) F(P) \frac{z^{2|P|}}{1 - z^{|P|}} \right)}{1 + \sum_{n \geq 1} np_n \frac{z^{2n}}{1 - z^n}}. \end{aligned}$$



Wir interessieren uns für  $P_u(z, 1)$  wenn  $z \rightarrow \rho$ . Aus Satz 2.2.5 (Beweis) und [27] wissen wir, dass  $P(\rho) = 1$ , und  $P(z) \sim 1 - d_1 \left(1 - \frac{z}{\rho}\right)^{1/2}$ , was  $zP'(z) \sim \frac{d_1}{2} \left(1 - \frac{z}{\rho}\right)^{-1/2}$  impliziert.

Nun gilt

$$\frac{zP'(z)}{P(z)} \sim \frac{d_1}{2} \left(1 - \frac{z}{\rho}\right)^{-1/2}, \quad z \rightarrow \rho.$$

Wegen (3.10) liegt  $H(z)$  in einem größeren Konvergenzkreis als  $P(z)$ , und ist daher bei  $z = \rho$  analytisch. Daraus folgt

$$P_u(z, 1) \sim \frac{d_1}{2} \left(1 - \frac{z}{\rho}\right)^{-1/2} \frac{\sum_P \left( f(P)\rho^{|P|} + F(P)\frac{\rho^{2|P|}}{1 - \rho^{|P|}} \right)}{1 + \sum_{n \geq 1} np_n \frac{\rho^{2n}}{1 - \rho^n}}, \quad z \rightarrow \rho$$

Insgesamt haben wir dann (mit Transfersatz 1.2.16 und (2.27)):

$$\begin{aligned} \mathbb{E}(F(P_n)) &= \frac{[z^n]P_u(z, 1)}{[z^n]P(z, 1)} \\ &\sim \frac{d_1 \rho^{-n}}{2\sqrt{\pi}n^{1/2}} \frac{\sum_P \left( f(P)\rho^{|P|} + F(P)\frac{\rho^{2|P|}}{1 - \rho^{|P|}} \right)}{1 + \sum_{n \geq 1} np_n \frac{\rho^{2n}}{1 - \rho^n}} \left( \frac{d_1 \rho^{-n}}{2\sqrt{\pi}n^{3/2}} \right)^{-1} \\ &= \frac{\sum_P \left( f(P)\rho^{|P|} + F(P)\frac{\rho^{2|P|}}{1 - \rho^{|P|}} \right)}{1 + \sum_{n \geq 1} np_n \frac{\rho^{2n}}{1 - \rho^n}} n \\ &= \mu n. \end{aligned}$$

Die Formel für die Varianz wird ähnlich gezeigt, indem man die zweite Ableitung  $P_{uu}(z, 1)$  berechnet.

Die asymptotische Normalität von  $\frac{F(P_n) - \mathbb{E}(F(P_n))}{\sqrt{\mathbb{V}(F(P_n))}}$  folgt direkt aus Satz 1.3.24.  $\square$

### 3.1.2 Anzahl der Blätter und Knoten vom Grad $k$

In einem zufälligen Wurzelbaum  $T$  ist die Anzahl der Knoten vom Weggrad  $k \in \mathbb{N}_0$  ein additiver Parameter mit Maut-Funktion

$$f(T) = \begin{cases} 1, & \text{falls der Wurzelgrad von } T \text{ gleich } k \text{ ist} \\ 0, & \text{sonst} \end{cases}$$

(siehe [34]). Insbesondere hat die Anzahl der Blätter (Knoten vom Weggrad 0) folgende Maut-Funktion:

$$f(T) = \begin{cases} 1, & T = \{\circ\} \\ 0, & \text{sonst.} \end{cases}$$

In diesem Abschnitt bezeichnen wir die Anzahl der Blätter mit  $X_n$  und die Anzahl der Knoten vom Weggrad  $k \in \mathbb{N}$  mit  $X_n^{(k)}$ .

### Einfach erzeugte Bäume

**Definition 3.1.5** Die bivariate erzeugende Funktion für die Anzahl der Knoten vom Weggrad  $k$  in einfach erzeugten Bäume lautet

$$S^{(k)}(z, u) = \sum_{n \geq 0} \sum_{m \geq 0} s_{n,m}^{(k)} z^n u^m \quad (3.14)$$

mit  $s_{n,m}^{(k)} = \sum_{S \in \mathcal{S}_n, X_n^{(k)}(S)=m} \omega(S)$ , wobei  $X_n^{(k)}(S)$  die Anzahl der Knoten mit  $k$  Kinder in  $S$  bezeichnet. Wie üblich markiert  $z$  die Größe, und  $u$  die Anzahl der Knoten vom Weggrad 0.

Wir haben (vgl. mit [8, Abschnitt 3.2.1] oder [15, III.18])

**Satz 3.1.6** Die bivariate erzeugende Funktion für die Anzahl der Blätter in einfach erzeugten Bäumen erfüllt folgende Funktionalgleichung

$$S^{(0)}(z, u) = z\Phi(S^{(0)}(z, u)) + \phi_0 z(u-1). \quad (3.15)$$

Für die bivariate erzeugende Funktion der Anzahl von Knoten vom höherem Grad gilt

$$S^{(k)}(z, u) = z\Phi(S^{(k)}(z, u)) + \phi_k z(u-1)S^{(k)}(z, u)^k. \quad (3.16)$$

Die  $\phi_k$  sind die Gewichte aus Definition 2.1.1.

*Beweis:* Symbolische Methode, indem man die Blätter bzw. die Knoten mit  $k$  Kinder kennzeichnet und  $\mathcal{S}$  durch die Struktur von  $S^{(0)}(z, u)$  bzw.  $S^{(k)}(z, u)$  ersetzt (siehe [15, III.18] für die Catalan-Bäume).  $\square$

Mithilfe dieser Funktionalgleichung sind für manche Baumklassen der Erwartungswert und die Varianz relativ leicht zu berechnen:

**Beispiel 3.1.7** Für die ebenen Wurzelbäume gilt

$$S^{(0)}(z, u) = \frac{z}{1 - S^{(0)}(z, u)} + z(u-1) = \frac{1}{2} \left( 1 + (u-1)z - \sqrt{1 - 2(u+1)z + (u-1)^2 z^2} \right).$$

Ableiten nach  $u$  und  $u = 1$  setzen liefert

$$S_u^{(0)}(z, 1) = \frac{1}{2}z + \frac{1}{2} \frac{z}{\sqrt{1-4z}}.$$

Daher gilt

$$\mathbb{E}(X_n) = \frac{[x^n] S_u^{(0)}(z, 1)}{[x^n] S^{(0)}(z, 1)} = \frac{n}{2}.$$

Wir leiten nochmals nach  $u$  ab, setzen  $u = 1$ , und bekommen:

$$S_{uu}^{(0)}(z, 1) = \frac{2z^3}{(1-4z)^{3/2}} = 2 \sum_{n \geq 0} (2n-5) \binom{2n-6}{n-3} z^n$$

und daher

$$\begin{aligned} \mathbb{V}(X_n) &= \frac{[x^n] S_{uu}^{(0)}(z, 1)}{[x^n] S^{(0)}(z, 1)} + \mathbb{E}(X_n) - (\mathbb{E}(X_n))^2 \\ &= \frac{2(2n-5) \binom{2n-6}{n-3}}{\frac{1}{n} \binom{2n-2}{n-1}} + \frac{n}{2} - \frac{n^2}{4} \\ &= \frac{n^3 - 3n^2 + 2n}{4n-6} + \frac{n}{2} - \frac{n^2}{4} \\ &= \frac{n^2 - 2n}{8n-12} \\ &\sim \frac{n}{8}. \end{aligned}$$

**Beispiel 3.1.8** Für die Binärbäume erster Art haben wir:

$$S^{(0)}(z, u) = z(1 + S^{(0)}(z, u)^2) + z(u-1) = \frac{1 - \sqrt{1-4z^2u}}{2z}.$$

Daraus folgt

$$S_u^{(0)}(z, 1) = \frac{z}{(1-4z^2)^{1/2}} = \sum_{n \geq 0} \binom{2n}{n} z^{2n+1}$$

und

$$S_{uu}^{(0)}(z, 1) = \frac{2z^3}{(1-4z^2)^{3/2}} = 2 \sum_{n \geq 0} (2n-1) \binom{2n-2}{n-1} z^{2n+1}.$$

Wir bekommen (nur für ungerades  $n$ ):

$$\mathbb{E}(X_n) = \frac{\binom{n-1}{(n-1)/2}}{\frac{2}{n+1} \binom{n-1}{(n-1)/2}} = \frac{n+1}{2} = \frac{n}{2} + \mathcal{O}(1)$$

und

$$\begin{aligned} \mathbb{V}(X_n) &= \frac{2(n-2) \binom{n-3}{(n-3)/2}}{\frac{2}{n+1} \binom{n-1}{(n-1)/2}} + \mathbb{E}(X_n) - (\mathbb{E}(X_n))^2 \\ &= \frac{n^2-1}{4} + \frac{n+1}{2} - \frac{n^2+2n+1}{4} = 0. \end{aligned}$$

Mithilfe von Satz 1.3.24 können wir sogar allgemeine asymptotische Ausdrücke für  $\mathbb{E}(X_n)$  und  $\mathbb{V}(X_n)$  bestimmen (vgl. mit [8, Theorem 3.13]):

**Satz 3.1.9** Sei  $X_n$  die Anzahl der Blätter in einem zufälligen einfach erzeugten Baum mit  $n$  Knoten. Für den Mittelwert und die Varianz gilt dann

$$\mathbb{E}(X_n) = \frac{\phi_0}{\Phi(\tau)}n + \mathcal{O}(1) \quad (3.17)$$

und

$$\mathbb{V}(X_n) = \left( \frac{\phi_0}{\Phi(\tau)} - \frac{\phi_0^2}{\Phi(\tau)^2} - \frac{\phi_0^2}{\tau^2\Phi(\tau)\Phi''(\tau)} \right) n + \mathcal{O}(1). \quad (3.18)$$

Allgemeiner gilt

$$\mathbb{E}(X_n^{(k)}) = \frac{\phi_k\tau^k}{\Phi(\tau)}n + \mathcal{O}(1) \quad (3.19)$$

und

$$\mathbb{V}(X_n^{(k)}) = \left( \frac{\phi_k\tau^k}{\Phi(\tau)} - \frac{\phi_k^2\tau^{2k}}{\Phi(\tau)^2} - \frac{(k-1)^2\phi_k^2\tau^{2k-2}}{\Phi(\tau)\Phi''(\tau)} \right) n + \mathcal{O}(1) \quad (3.20)$$

wobei  $X_n^{(k)}$  die Anzahl der Knoten vom Grad  $k \in \mathbb{N}$  in einem zufälligen einfach erzeugten Baum mit  $n$  Knoten bezeichnet. Die  $\phi_k$  sind die Gewichte aus Definition 2.1.1.  $\tau$  ist wie in Proposition 2.1.12.

*Beweis:* Wir wenden Proposition 1.3.23 und Satz 1.3.24 an. Setze  $F(z, y, u) = z\Phi(y) + \phi_0z(u-1)$ . Wir haben

$$\begin{aligned} F(0, y, u) &= 0 \\ F(z, 0, u) &= z\Phi(0) + \phi_0z(u-1) \neq 0 \\ [z^n y^k]F(z, y, 1) &= [z^n y^k]z\Phi(y) \geq 0. \end{aligned}$$

Weiters ist  $y = S^{(0)}(z, u)$  die eindeutige Lösung der Funktionalgleichung  $F(z, y, u) = y$  und es gelten (wobei  $(z_0, y_0) = (\rho, \tau)$  und wegen Proposition 2.1.12)

$$\begin{aligned} S^{(0)}(0, u) &= F(0, y, u) = 0 \\ F(\rho, \tau, 1) &= \rho\Phi(\tau) = \tau \\ F_y(\rho, \tau, 1)\rho\Phi'(\tau) &= 1 \\ F_z(\rho, \tau, 1) &= \Phi(\tau) \neq 0 \\ F_{yy}(\rho, \tau, 1) &= \rho\Phi''(\tau) \neq 0. \end{aligned}$$

Wir erhalten dann (mit  $F_u(\rho, \tau, 1) = \phi_0\rho$ ,  $F_{zy}(\rho, \tau, 1) = \Phi'(\tau)$ ,  $F_{zu}(\rho, \tau, 1) = \phi_0$  und  $F_{zz}(\rho, \tau, 1) = F_{yu}(\rho, \tau, 1) = F_{uu}(\rho, \tau, 1) = 0$ )

$$\mathbb{E}(X_n) = \frac{\phi_0\rho}{\rho\Phi(\tau)}n + \mathcal{O}(1) = \frac{\phi_0}{\Phi(\tau)}n + \mathcal{O}(1)$$

und

$$\begin{aligned} \mathbb{V}(X_n) &\sim \left( \frac{\phi_0^2}{\Phi(\tau)^2} + \frac{1}{\rho} \left( \frac{\phi_0\rho}{\Phi(\tau)} - \frac{1}{\rho\Phi''(\tau)\Phi(\tau)} \left( \frac{\phi_0\rho\Phi(\tau)}{\Phi(\tau)} \right)^2 - \frac{1}{\Phi(\tau)} \frac{2\phi_0^2\rho}{\Phi(\tau)} \right) \right) n \\ &= \left( \frac{\phi_0^2}{\Phi(\tau)^2} + \frac{\phi_0}{\Phi(\tau)} - \frac{\phi_0^2}{\tau^2\Phi(\tau)\Phi''(\tau)} - \frac{2\phi_0^2}{\Phi(\tau)^2} \right) n \\ &= \left( \frac{\phi_0}{\Phi(\tau)} - \frac{\phi_0^2}{\tau^2\Phi(\tau)\Phi''(\tau)} - \frac{\phi_0^2}{\Phi(\tau)^2} \right) n. \end{aligned}$$

(3.19) und (3.20) werden analog berechnet mit  $F(z, y, u) = z\Phi(y) + \phi_k z(u-1)y^k$ .  $\square$

**Bemerkung 3.1.10** Sei  $\pi_{n,k} = \frac{\mathbb{E}(X_n^{(k)})}{n}$  die Wahrscheinlichkeit, dass ein zufälliger Knoten in einem zufälligen einfach erzeugten Baum der Größe  $n$  den Weggrad  $k$  hat. Dann sieht man leicht, dass

$$\pi_{n,k} \rightarrow \pi_k = \frac{\phi_k \tau^k}{\Phi(\tau)}. \quad (3.21)$$

(siehe [8, Teil 3.2.1]). Die  $\pi_k, k \geq 0$  bilden genau die Wahrscheinlichkeitsverteilung des Galton-Watson Verzweigungsprozesses (vgl. mit Satz 2.1.6, insbesondere (2.10)).

**Beispiel 3.1.11** Für die ebenen Wurzelbäume haben wir  $\phi_0 = 1, \tau = \frac{1}{2}, \Phi(\tau) = \frac{1}{1 - \frac{1}{2}} = 2$  und  $\Phi''(\tau) = \frac{2}{(1 - \frac{1}{2})^3} = 16$ . (3.17) liefert uns also auch den exakten Mittelwert. Laut (3.18) haben wir

$$\mathbb{V}(X_n) \sim \left( \frac{1}{2} - \frac{1}{\frac{1}{4} \cdot 16} - \frac{1}{4} \right) n = \left( \frac{1}{2} - \frac{1}{8} - \frac{1}{4} \right) n = \frac{n}{8}.$$

Allgemeiner haben wir für die Anzahl der Knoten vom Weggrad  $k$  (mit jeweils  $\phi_k = 1$ )

$$\mathbb{E}(X_n^{(k)}) = \left( \frac{1}{2} \right)^k \frac{n}{2}$$

und

$$\mathbb{V}(X_n^{(k)}) \sim \left( \frac{1}{2} \left( \frac{1}{2} \right)^k - \frac{1}{4} \left( \frac{1}{2} \right)^{2k} - \frac{(k-1)^2}{2 \cdot 16} \left( \frac{1}{2} \right)^{2k-2} \right) n.$$

**Beispiel 3.1.12** Wegen Satz 3.1.6 hat die bivariate erzeugende Funktion für die Anzahl der Blätter in Cayley-Bäumen folgende Darstellung:

$$S^{(0)}(z, u) = ze^{S^{(0)}(z, u)} + z(u-1).$$

Für den entsprechenden Mittelwert und die Varianz gelten daher (mit  $\phi_0 = 1, \tau = 1$  und  $\Phi(\tau) = \Phi''(\tau) = e^1$ )

$$\mathbb{E}(X_n) \sim e^{-1}n$$

und

$$\mathbb{V}(X_n) \sim \left( \frac{1}{e} - \frac{1}{e^2} - \frac{1}{e^2} \right) n = \frac{e-2}{e^2}n.$$

Allgemeiner gilt (mit  $\phi_k = \frac{1}{k!}$ )

$$\mathbb{E}(X_n^{(k)}) \sim \frac{e^{-1}}{k!}n$$

und

$$\mathbb{V}(X_n^{(k)}) \sim \left( \frac{1}{ek!} - \frac{1}{(ek!)^2} - \frac{(k-1)^2}{(ek!)^2} \right) n.$$

**Beispiel 3.1.13** Wir betrachten Binärbäume erster Art. Da  $\phi_0 = 1$ ,  $\tau = 1$  und  $\Phi(\tau) = \Phi''(\tau) = 2$ , haben wir

$$\mathbb{E}(X_n) \sim \frac{1}{2}n$$

und

$$\mathbb{V}(X_n) \sim \left(\frac{1}{2} - \frac{1}{4} - \frac{1}{2 \cdot 2}\right)n = 0.$$

Der hier ermittelte Wert für die Varianz ist sogar exakt (vgl. Beispiel 3.1.8).

Für die Anzahl der Knoten vom Weggrad 2 (interne Knoten) haben wir (mit  $\phi_2 = 1$ )

$$\mathbb{E}(X_n^{(2)}) \sim \frac{1}{2}n$$

und

$$\mathbb{V}(X_n^{(2)}) \sim \left(\frac{1}{2} - \frac{1}{4} - \frac{1}{4}\right)n = 0.$$

Wegen  $\phi_k = 0$  für jedes  $k \in \mathbb{N}_0 \setminus \{0, 2\}$  gilt  $\mathbb{E}(X_n^{(k)}) = \mathbb{V}(X_n^{(k)}) = 0$  für diese  $k$ .

Ergebnisse für Binärbäume zweiter Art und Motzkin-Bäume werden ähnlich berechnet (siehe Tabelle am Ende des Abschnitts). Nun können wir unsere Grenzverteilung bestimmen (vgl. mit [8, Theorem 3.13])

**Proposition 3.1.14** Es gelten

$$\frac{X_n - \mathbb{E}(X_n)}{\sqrt{\mathbb{V}(X_n)}} \xrightarrow{d} N \sim \mathcal{N}(0, 1)$$

und

$$\frac{X_n^{(k)} - \mathbb{E}(X_n^{(k)})}{\sqrt{\mathbb{V}(X_n^{(k)})}} \xrightarrow{d} N \sim \mathcal{N}(0, 1)$$

für alle oben betrachteten Klassen von einfach erzeugte Bäume mit Knoten vom geeigneten Weggrad, bis auf Binärbäume erster Art.

*Beweis:* Das Ergebnis folgt direkt aus Satz 3.1.3. Für Binärbäume gilt aber  $\mathbb{V}(X_n) = \mathbb{V}(X_n^{(2)}) = 0$  (andere Fälle laut Beispiel 3.1.13 irrelevant), was die Bedingung  $\sigma^2 \neq 0$  verletzt.  $\square$

## Pólya-Bäume

Für Pólya-Bäume sind solche Resultate viel komplizierter zu bestimmen. In diesem Abschnitt wiederholen wir die Ergebnisse von Robinson und Schwenk in [31] und von Drmota und Gittenberger in [10] und Gittenberger in [17]. Bezeichne  $p_{n,m}^{(k)}$  die Anzahl Pólya-Bäume mit  $n$  Knoten, wobei  $m$  davon vom Weggrad  $k$  sind. Die zugehörige bivariate erzeugende Funktion lautet dann

$$P^{(k)}(z, u) = \sum_{n \geq 0} \sum_{m \geq 0} p_{n,m}^{(k)} z^n u^m. \quad (3.22)$$

Um den Mittelwert und die Varianz für die Anzahl der Knoten vom Grad  $k$  in einem Pólya-Baum von  $n$  Knoten zu bestimmen betrachtet man weitere Klassen von Bäumen (siehe [10, Teil 2.1], [17, Abschnitt 2.2] oder [31, Teil 3]):

**Definition 3.1.15** Sei  $\tilde{\mathcal{P}}$  die kombinatorische Struktur der gepflanzten Bäume (d.h. Wurzelbäume, deren Wurzeln ausnahmensweise einen Vorgänger besitzt, der aber nicht zur Größe beiträgt). Wir definieren die entsprechende bivariate erzeugende Funktion wie folgt:

$$p^{(k)}(z, u) = \sum_{n \geq 0} \sum_{m \geq 0} \tilde{p}_{n,m}^{(k)} z^n u^m. \quad (3.23)$$

**Satz 3.1.16** Ist  $Z(\text{Sym}_k, x_1, x_2, \dots, x_k)$  der Zyklenzeiger der symmetrischen Gruppe  $\text{Sym}_k$  (vgl. mit Bemerkung 2.2.3), so besitzen  $P_k(z, u)$  und  $p_k(z, u)$  folgende alternative Darstellungen:

$$p^{(k)}(z, u) = z \exp \left( \sum_{i \geq 1} \frac{p^{(k)}(z^i, u^i)}{i} \right) + z(u-1)Z(\text{Sym}_{k-1}, p^{(k)}(z, u), p^{(k)}(z^2, u^2), \dots, p^{(k)}(z^{k-1}, u^{k-1})),$$

$$P^{(k)}(z, u) = z \exp \left( \sum_{i \geq 1} \frac{p^{(k)}(z^i, u^i)}{i} \right) + z(u-1)Z(\text{Sym}_k, p^{(k)}(z, u), p^{(k)}(z^2, u^2), \dots, p^{(k)}(z^k, u^k)).$$

Es ist leicht zu erkennen, dass (mit  $P^{(k)}(z, 1) = P(z)$  und  $p(z) := p^{(k)}(z, 1)$ ) gilt  $p(z) = P(z)$  (vgl. mit [31, Formel (4)]).

*Beweis:* Siehe [31, Teil 3]. □

Nun können wir den Mittelwert von  $X_n^{(k)}$  bestimmen.

**Satz 3.1.17** Sei  $X_n^{(k)}$  die Anzahl der Knoten vom Weggrad  $k \in \mathbb{N}$  in einem zufälligen Pólya-Baum. Dann ist der Erwartungswert von  $X_n^{(k)}$  durch

$$\mathbb{E}(X_n^{(k)}) \sim \mu_k n \quad (3.24)$$

gegeben, wobei

$$\mu_k = \frac{2}{d_1^2} \left( \sum_{i \geq 2} p_u^{(k)}(\rho^i, 1) + \rho Z(\text{Sym}_{k-1}; P(\rho), P(\rho^2), \dots, P(\rho^{k-1})) \right) \quad (3.25)$$

und  $\rho$  die dominante Singularität von  $P(z)$  ist.  $d_1$  ist wie im Satz 2.2.5.

*Beweis:* Wir leiten  $p^{(k)}(z, u)$  nach  $u$  ab. Mithilfe von Satz 3.1.16 bekommen wir

$$\begin{aligned} p_u^{(k)}(z, u) &= z \exp \left( \sum_{i \geq 1} \frac{p^{(k)}(z^i, u^i)}{i} \right) \sum_{i \geq 1} p_u^{(k)}(z^i, u^i) u^{i-1} \\ &\quad + z Z(\text{Sym}_{k-1}, p^{(k)}(z, u), p^{(k)}(z^2, u^2), \dots, p^{(k)}(z^{k-1}, u^{k-1})) \\ &\quad + z(u-1)(Z(\text{Sym}_{k-1}, p^{(k)}(z, u), p^{(k)}(z^2, u^2), \dots, p^{(k)}(z^{k-1}, u^{k-1})))_u. \end{aligned}$$

Setze nun  $u = 1$ . Wegen  $p^{(k)}(z, 1) = P(z)$  bekommen wir

$$\begin{aligned} p_u^{(k)}(z, 1) &= z \exp \left( \sum_{i \geq 1} \frac{P(z^i)}{i} \right) \sum_{i \geq 1} p_u^{(k)}(z^i, 1) + z Z(\text{Sym}_{k-1}, P(z), P(z^2), \dots, P(z^{k-1})) \\ &= P(z) \sum_{i \geq 1} p_u^{(k)}(z^i, 1) + z Z(\text{Sym}_{k-1}, P(z), P(z^2), \dots, P(z^{k-1})). \end{aligned}$$

Analog haben wir

$$P_u^{(k)}(z, 1) = P(z) \sum_{i \geq 1} p_u^{(k)}(z^i, 1) + zZ(\text{Sym}_k, P(z), P(z^2), \dots, P(z^k)).$$

Sei  $\rho$  die dominante Singularität von  $P(z)(= p^{(k)}(z, 1))$ . In der Nähe von  $\rho$  gilt

$$Z(\text{Sym}_k; P(z), \dots, P(z^k)) \sim P(z)Z(\text{Sym}_{k-1}; P(z), \dots, P(z^{k-1}))$$

(vgl. [31, Formel (30)]). Wegen  $P(\rho) = 1$  (vgl. Beweis von Satz 2.2.5 und [27]) gilt sogar  $Z(\text{Sym}_k; P(\rho), \dots, P(\rho^k)) \sim Z(\text{Sym}_{k-1}; P(\rho), \dots, P(\rho^{k-1}))$  und daher

$$P_u^{(k)}(\rho, 1) \sim p_u^{(k)}(\rho, 1).$$

Es genügt also, einen asymptotischen Ausdruck von  $p_u^{(k)}(z, 1)$  für  $z \rightarrow \rho$  zu berechnen. Wir können  $p_u^{(k)}(z, 1)$  wie folgt umschreiben:

$$p_u^{(k)}(z, 1) = \frac{1}{1 - P(z)} \left( P(z) \sum_{i \geq 2} p_u^{(k)}(z^i, 1) + zZ(\text{Sym}_{k-1}, P(z), P(z^2), \dots, P(z^{k-1})) \right).$$

Aus Satz 2.2.5 und wegen  $P(\rho) = 1$  haben wir

$$p_u^{(k)}(z, 1) \sim \frac{1}{d_1 \left(1 - \frac{z}{\rho}\right)^{1/2}} \left( \sum_{i \geq 2} p_u^{(k)}(\rho^i, 1) + \rho Z(\text{Sym}_{k-1}, P(\rho), P(\rho^2), \dots, P(\rho^{k-1})) \right), \quad z \rightarrow \rho$$

und insgesamt gilt (mithilfe der Sätze 1.2.16 und 2.2.6)

$$\begin{aligned} \mathbb{E}(X_n^{(k)}) &= \frac{[z^n] P_u^{(k)}(z, 1)}{p_n} \\ &\sim \frac{[z^n] p_u^{(k)}(z, 1)}{p_n} \\ &\sim \frac{\rho^{-n}}{\sqrt{\pi d_1 n^{1/2}}} \left( \sum_{i \geq 2} T_u^{(k)}(\rho^i, 1) + \rho Z(\text{Sym}_{k-1}, P(\rho), P(\rho^2), \dots, P(\rho^{k-1})) \right) \\ &\quad \frac{d_1 \rho^{-n}}{2\sqrt{\pi n^{3/2}}} \\ &= \frac{2}{d_1^2} \left( \sum_{i \geq 2} T_u^{(k)}(\rho^i, 1) + \rho Z(\text{Sym}_{k-1}, P(\rho), P(\rho^2), \dots, P(\rho^{k-1})) \right) n \\ &= \mu_k n. \end{aligned}$$

□

Mithilfe eines Ergebnisses von Schwenk (siehe [32, Corollary 4.1]) können wir den Mittelwert  $\mathbb{E}(X_n^{(k)})$  für große  $k$  wie folgt alternativ darstellen.

**Korollar 3.1.18** *Der Erwartungswert  $\mathbb{E}(X_n^{(k)})$  besitzt für große  $k$  folgenden asymptotischen Ausdruck:*

$$\mathbb{E}(X_n^{(k)}) \sim K \rho^k n \tag{3.26}$$



wobei  $K = \frac{2C}{d_1^2} \approx 6.380045$  und insbesondere

$$C = \exp\left(\sum_{i \geq 1} \left(\frac{P(\rho^i)}{\rho^i} - 1\right) \frac{1}{i}\right) \approx 7.7581604.$$

$\rho$  ist die dominante Singularität von  $P(z)$  und  $d_1$  ist wie im Satz 2.2.5.

*Beweis:* Folgt aus Theorem 3.1 in [32] und aus der Tatsache, dass  $\sum_{i \geq 2} p^{(k)}(\rho^i, 1)$  mithilfe grober Abschätzung schneller als  $\rho^k$  fällt. Für einen vollständigen Beweis siehe [32, Corollary 4.1]  $\square$

Nun formulieren wir unser Ergebnis für die Grenzverteilung (vgl. mit [17, Theorem 1] und [10, Theorem 2.1] für vektorwertiges  $k$ ):

**Satz 3.1.19** Sei  $X_n^{(k)}$  die Anzahl der Knoten vom Weggrad  $k$  in einem zufälligen Pólya-Baum. Definiere  $F(z, y, u)$  durch

$$F(z, y, u) = ze^y \exp\left(\sum_{i \geq 2} \frac{T^{(k)}(z^i, u^i)}{i}\right) + z(u-1)Z(\text{Sym}_{k-1}, y, T^{(k)}(z^2, u^2), \dots, T^{(k)}(z^{k-1}, u^{k-1})).$$

Setze weiters

$$f_1 = -\frac{F_u}{F_z}(\rho, 1, 1),$$

$$f_2 = \left(\frac{1}{F_{yy}F_z} \left(\frac{F_u F_{zy}}{F_z} - F_{zy}\right)^2 - \frac{1}{F_z} \left(\frac{F_u^2 F_{zz}}{F_z^2} - \frac{2F_u F_{uz}}{F_z} + F_{uu}\right)\right)(\rho, 1, 1).$$

Dann folgt (mit  $n \rightarrow \infty$  und  $k$  fest)

$$\mathbb{E}(X_n^{(k)}) = \mu_k n + \mathcal{O}(1) \tag{3.27}$$

und

$$\mathbb{V}(X_n^{(k)}) = \sigma_k^2 n + \mathcal{O}(1) \tag{3.28}$$

mit

$$\mu_k = -\frac{f_1}{\rho} \quad \text{und} \quad \sigma_k^2 = \left(\frac{f_1}{\rho}\right)^2 - \frac{f_2}{\rho}. \tag{3.29}$$

Insbesondere gilt

$$\mu_k \sim \frac{2C}{d_1^2} \rho^k \quad \text{und} \quad \sigma_k^2 \sim \frac{2C}{d_1^2} \rho^k \tag{3.30}$$

für hinreichend großes  $k$  (vgl. mit [17, Theorem 1]).  $\rho$  ist die dominante Singularität von  $P(z)$ ,  $d_1$  ist wie im Satz 2.2.5 und

$$C = \exp\left(\sum_{i \geq 1} \frac{1}{i} \left(\frac{P(\rho^i)}{\rho^i} - 1\right)\right) \approx 7.7581604.$$

Wir wiederholen die Idee von Drmota und Gittenberger in [10] (beachte nur, dass  $k \in \mathbb{N}$ ). Dazu brauchen wir zusätzliche Eigenschaften. Die erste Beobachtung lautet (vgl. mit [10, Lemma 3.2]):

**Lemma 3.1.20** Sei  $k \in \mathbb{N}$ . Dann gibt es eine Konstante  $\eta > 0$  und Funktionen  $b_i(z, u), c_i(z, u), i = 1, 2$  und  $f(u)$  für die gilt:

- $b_i(z, u), c_i(z, u), i = 1, 2$  und  $f(u)$  sind analytisch für  $|u - 1| < \eta$  und  $|z - f(u)| < \eta$ ;
- $b(\rho, 1) = 1$  und  $c(\rho, 1) = d_1$ , wobei  $d_1$  wie im Satz 2.2.5 ist;
- für  $(z, u) \in G, |u - 1| < \eta$  und  $|z - f(u)| < \eta$  besitzen  $p^{(k)}(z, u)$  und  $P^{(k)}(z, u)$  folgende Darstellungen

$$p^{(k)}(z, u) = b_1(z, u) - c_1(z, u) \left(1 - \frac{z}{f(u)}\right)^{1/2} \quad (3.31)$$

und

$$P^{(k)}(z, u) = b_2(z, u) - c_2(z, u) \left(1 - \frac{z}{f(u)}\right)^{1/2} \quad (3.32)$$

wobei  $G := \left\{ (z, u) \in \mathbb{C}^2 : |u| \leq 1 + \frac{\eta}{2}, |z| \leq \rho + \frac{\eta}{2}, \arg(z - f(u)) \neq 0 \right\}$ .

Beweis: Siehe [10, Lemma 3.2] □

Für Funktionen, die nichtnegative Koeffizienten haben und eine Darstellung der Form (3.31) oder (3.32) erfüllen, gilt folgende Proposition (vgl. mit [10, Proposition 4.1]):

**Proposition 3.1.21** Sei  $\eta > 0$ . Sei weiters  $y(z, u) = \sum_{n,m \geq 0} y_{m,n} z^n u^m$  eine analytische Funktion mit nichtnegativen Koeffizienten  $y_{m,n}$ , die für  $(z, u) \in G, |u - 1| < \eta$  und  $|z - \rho| < \eta$  die Darstellung

$$y(z, u) = b(z, u) - c(z, u) \left(1 - \frac{z}{f(u)}\right)^{1/2}$$

hat, wobei  $b(z, u), c(z, u)$  und  $f(u)$  analytisch für  $|u - 1| < \eta$  und  $|z - \rho| < \eta$  sind.  $\rho > 0$  bezeichne der Konvergenzradius von  $y(z, 1)$  und  $G$  sei wie im Lemma 3.1.20. Dann besitzt  $y_n(u) := [z^n]y(z, u) = \sum_{m \geq 0} y_{m,n} u^m$  folgende asymptotische Darstellung:

$$y_n(u) = \frac{c(f(u), u)}{2\sqrt{\pi n^{3/2}}} f(u)^{-n+1} + \mathcal{O}\left(\frac{f(u)^{-n}}{n^{5/2}}\right)$$

gleichmäßig für  $|u - 1| < \eta$ .

Wir betrachten nichtnegative Zahlen  $y_{m,n}$ , sodass

$$\sum_{m \geq 0} y_{m,n} < \infty, \quad \forall n > 0.$$

Wir interessieren uns für die Grenzverteilung einer Zufallsvariable  $X_n$  mit der Wahrscheinlichkeit

$$\mathbb{P}(X_n = m) = \frac{y_{m,n}}{y_n}.$$

Dafür formulieren wir eine Proposition, die von Bender und Richmond in [3] gezeigt wurde (Theorem 1).

**Proposition 3.1.22** Seien  $y_{m,n} \geq 0$  und Funktionen  $H(u), f(u)$  definiert für  $u = e^{it}$ ,  $|t| < \epsilon$  für ein  $\epsilon > 0$  und  $t \in \mathbb{R}$ , sodass  $H(1) \neq 0$ ,  $H(u)$  gleichmäßig stetig,  $f(1) = \rho > 0$  und  $f(e^{it})$  dreimal stetig differenzierbar, mit

$$y_n(u) = \sum_{m \geq 0} y_{m,n} u^m \sim a_n H(u) f(u)^{-n}$$

gleichmäßig für  $|t| < \epsilon$ . Die  $a_n$  bilden eine Folge positiver Zahlen. Definiere weiters  $\mu$  und  $\sigma$  durch

$$\mu = i \frac{\partial}{\partial t} \log f(e^{it}) \Big|_{t=0} \quad \text{und} \quad \sigma^2 = - \frac{\partial^2}{\partial t^2} \log f(e^{it}) \Big|_{t=0} \neq 0.$$

Dann gilt

$$\frac{X_n - n\mu}{\sqrt{n}} \xrightarrow{d} N \sim \mathcal{N}(0, \sigma^2). \quad (3.33)$$

*Beweis:* Siehe [3, Theorem 1]. □

Nun können wir den Satz 3.1.19 beweisen.

*Beweis (Satz 3.1.19):* Ersetze  $u_0 = 1$  in Proposition 1.3.23 durch ein  $u$  aus einer hinreichend kleinen Umgebung von 1. Dann sind  $y = y(f(u), u)$ ,  $z = f(u)$  die Lösungen der Gleichungen

$$\begin{aligned} y &= F(z, y, u), \\ 1 &= F_y(z, y, u). \end{aligned}$$

Wir können  $\mu$  und  $\sigma^2$  aus Proposition 3.1.22 wie folgt umschreiben:

$$\mu = i \frac{\partial}{\partial t} \log f(e^{it}) \Big|_{t=0} = i \frac{f_u(e^{it})}{f(e^{it})} i e^{it} \Big|_{t=0} = - \frac{f_u(1)}{f(1)}$$

und

$$\sigma^2 = - \frac{\partial^2}{\partial t^2} \log f(e^{it}) \Big|_{t=0} = - \frac{\partial}{\partial t} \frac{f_u(e^{it})}{f(e^{it})} i e^{it} \Big|_{t=0} = \frac{f_u(1)^2 - f_{uu}(1) f(1)}{f(1)^2} - \frac{f_u(1)}{f(1)}.$$

Daraus folgt

$$y_u = F_z f_u + F_u + F_y y_u.$$

Da  $F_y = 1$  ist folgt  $F_z f_u + F_u \equiv 0$ , was  $f_u = -\frac{F_u}{F_z}$  impliziert und daher gilt

$$\mu = \frac{F_u(z_0, y(f(1), 1), 1)}{f(u_0) F_z(z_0, y(f(1), 1), 1)} = \frac{F_u(z_0, y_0, 1)}{z_0 F_z(z_0, y_0, 1)}.$$

Nochmaliges implizites Ableiten liefert uns

$$f_{uu} = \frac{1}{F_{yy} F_z} \left( \frac{F_u F_{yz}}{F_z} - F_{yu} \right)^2 - \frac{1}{F_z} \left( \frac{F_u^2 F_{zz}}{F_z^2} - \frac{2 F_u F_{zu}}{F_z} + F_{uu} \right).$$

Nun können wir die Varianz bestimmen. Sei bemerkt, dass  $\mu$  und  $\sigma^2$  an der Stelle  $(\rho, 1, 1)$  genau dieselbe Form wie  $\mu_k$  bzw.  $\sigma_k^2$  in (3.29) besitzen, wobei  $\rho$  die dominante Singularität

von  $P(z)$  ist.

Wir betrachten wieder die Funktionen  $P^{(k)}(z, u)$  und  $p^{(k)}(z, u)$ . Wegen Lemma 3.1.20 und der Tatsache dass die Koeffizienten  $p_{n,m}^{(k)}$  und  $\tilde{p}_{n,m}^{(k)}$  nichtnegativ sind kann man das Lemma 3.1.21 anwenden. Daher sind die Voraussetzungen der Proposition 3.1.22 erfüllt, und wegen remark 3 in [10] hat  $X_n^{(k)}$  genau den Mittelwert  $\mathbb{E}(X_n^{(k)}) = \mu n + \mathcal{O}(1)$  und die Varianz  $\mathbb{V}(X_n^{(k)}) = \sigma^2 n + \mathcal{O}(1)$ . Aus Proposition 3.1.22 erhalten wir daher die asymptotische Normalität. Für großes  $k$  wurde das  $\mu_k$  in (3.30) in Korollar schon bewiesen. Für die Varianz sollen wir die partiellen Ableitungen  $F_z, F_y, F_u, F_{zz}, F_{zy}, F_{zu}, F_{yy}$  und  $F_{uu}$  an der Stelle  $(\rho, 1, 1)$  noch berechnen, und diese mithilfe von Rechenregeln für den Zykluszeiger (siehe [32]) vereinfachen. Einsetzen in (3.29) liefert uns (3.30). Für einen vollständigen Beweis siehe [10, Teil 4]  $\square$

Unter den Voraussetzungen des Satzes 3.1.19 ist  $k$  fest. Nun betrachten wir den Fall, dass  $k$  von  $n$  abhängt (siehe [17, Theorem 2]).

**Bemerkung 3.1.23** Sei  $X_n^{(k)}$  die Anzahl der Knoten vom Weggrad  $k$  in einem zufälligen Pólya-Baum der Größe  $n$ . Betrachte nun den Fall  $k = k(n)$ , d.h.  $k$  ist eine Folge, die von  $n$  abhängt. Dann haben wir:

- wenn  $\lim_{n \rightarrow \infty} \mathbb{E}(X_n^{(k)}) = \infty$ , dann ist  $X_n^{(k)}$  immer noch asymptotisch normalverteilt mit dem gleichen Mittelwert und der Varianz wie im Satz 3.1.19.
- wenn es eine Teilfolge  $n_i$  von  $n$  gibt, sodass  $\lim_{i \rightarrow \infty} \mathbb{E}(X_{n_i}^{(k(n_i))}) = \lambda > 0$ , dann hat  $X_{n_i}^{(k(n_i))}$  eine Poisson-Grenzverteilung, d.h. für jedes  $x \geq 0$  gilt

$$\lim_{i \rightarrow \infty} \mathbb{P}(X_{n_i}^{(k(n_i))} = x) = \frac{\lambda^x}{x!} e^{-\lambda}.$$

- wenn  $\lim_{n \rightarrow \infty} \mathbb{E}(X_n^{(k)}) = 0$ , dann ist  $X_n^{(k)}$  degeneriert, d. h.  $\lim_{n \rightarrow \infty} \mathbb{P}(X_n^{(k)} = 0) = 1$ .

Diese Aussagen wurden in [17, Teil 3] bewiesen.

Wir vergleichen unsere Ergebnisse. Insgesamt haben wir:

Bäume	$\mathbb{E}(X_n^{(k)})$	$\mathbb{V}(X_n^{(k)})$	Grenzverteilung
Einfach erzeugt	$\frac{\phi_k \tau^k}{\Phi(\tau)} n$	$\left( \frac{\phi_k \tau^k}{\Phi(\tau)} - \frac{\phi_k^2 \tau^{2k}}{\Phi(\tau)^2} - \frac{(k-1)^2 \phi_k^2 \tau^{2k-2}}{\Phi(\tau) \Phi''(\tau)} \right) n$	$\mathcal{N}(0, 1)$
Catalan	$\left(\frac{1}{2}\right)^k \frac{n}{2}$	$\left( \left(\frac{1}{2}\right)^{k+1} - \left(\frac{1}{2}\right)^{2k+2} - \frac{(k-1)^2}{32} \left(\frac{1}{2}\right)^{2k-2} \right) n$	$\mathcal{N}(0, 1)$
Cayley	$\frac{e^{-1}}{k!} n$	$\left( \frac{1}{ek!} - \frac{1}{(ek!)^2} - \left(\frac{k-1}{ek!}\right)^2 \right) n$	$\mathcal{N}(0, 1)$
Binär 1	$\frac{1}{2} n$ für $k = 0, 2$	0 (exaktes Wert)	
Binär 2	$\binom{2}{k} \frac{1}{4} n$	$\left( \binom{2}{k} \frac{1}{4} - \binom{2}{k}^2 \frac{1}{16} - \binom{2}{k}^2 \frac{(k-1)^2}{8} \right) n$	$\mathcal{N}(0, 1)$
Motzkin	$\frac{1}{3} n$ für $k = 0, 1, 2$	$\frac{1}{18} n$ für $k = 0, 2$ ; $\frac{2}{9} n$ für $k = 1$	$\mathcal{N}(0, 1)$
Pólya	$K \rho^k n$ (für großes $k$ )	$K \rho^k n$ (für großes $k$ )	$\mathcal{N}(0, 1)$

mit  $\rho \approx 0.3383219$  und  $K \approx 6.380045$ .

Als ein weiteres Beispiel für einen additiven Parameter betrachten wir nun die interne Pfadlänge:

### 3.1.3 Die interne Pfadlänge

**Definition 3.1.24** Sei  $T$  ein Wurzelbaum der Grösse  $n$  und beschreiben  $T_1, T_2, \dots, T_k$  die von der Wurzel von  $T$  hängenden Teilbäume. Dann ist die interne Pfadlänge des Baumes wie folgt definiert (siehe [35, Formel (3)]):

$$F(T) = \sum_{i=1}^k (F(T_i) + |T_i|) = \sum_{i=1}^k F(T_i) + |T| - 1. \quad (3.34)$$

$F$  ist dann ein additiver Parameter mit Maut-Funktion  $f(T) = |T| - 1$ .

Aus [15, Seite 184] können wir eine allgemeine Darstellung für die entsprechende bivariate erzeugende Funktion bestimmen:

**Satz 3.1.25** Sei  $\mathcal{T}$  eine Klasse von Wurzelbäumen mit  $\mathcal{T} = \phi(\mathcal{T})$ . Weiters sei  $\psi$  ein Operator, sodass  $T(z) = \psi(T(z))$ , wobei  $T(z)$  die erzeugende Funktion von  $\mathcal{T}$  bezeichnet. Ist  $T(z, u)$  die entsprechende bivariate erzeugende Funktion von  $\mathcal{T}$  für die interne Pfadlänge, so gilt

$$T(z, u) = \psi(T(zu, u)). \quad (3.35)$$

*Beweis:* Mithilfe von (3.34) können wir den Parameter  $F(T)$  wie folgt umschreiben:

$$F(T) = \tilde{F}(T) + |T|,$$

wobei  $\tilde{F}(T) = \sum_{i=1}^k F(T_i) - 1$ . Ist  $T_F(z, u)$  die entsprechende bivariate erzeugende Funktion, so gilt

$$T_F(z, u) = \sum_{T \in \mathcal{T}} z^{|T|} u^{F(T)} = \sum_{T \in \mathcal{T}} (zu)^{|T|} u^{\tilde{F}(T)} = T_{\tilde{F}}(zu, u).$$

Die induktive Definition (3.34) von  $F(T)$  und einsetzen in  $\mathcal{T} = \phi(\mathcal{T})$  liefern uns (3.35).  $\square$

#### Einfach erzeugte Bäume

Zuerst definieren wir folgende bivariate erzeugende Funktion (vgl. mit [15, III.20]):

**Satz 3.1.26** Die bivariate erzeugende Funktion  $S(z, u)$  für die interne Pfadlänge in einfach erzeugten Bäumen erfüllt folgende Funktionalgleichung

$$S(z, u) = z\Phi(S(zu, u)). \quad (3.36)$$

*Beweis:* (3.35) mit  $\psi(S(z, u)) = z\Phi(S(zu, u))$ .  $\square$

**Beispiel 3.1.27** Für ebene Wurzelbäume haben wir

$$S(z, u) = \frac{z}{1 - S(zu, u)}.$$

Ableiten nach  $u$  und  $u = 1$  setzen liefert:

$$\begin{aligned} S_u(z, 1) &= \frac{z^2 S'(z) + z S_u(z, 1)}{(1 - S(z))^2} \\ \Leftrightarrow S_u(z, 1) &= \frac{z^2 S'(z)}{(1 - S(z))^2 - z} \\ &= \frac{z}{2(1 - 4z)} - \frac{z}{2\sqrt{1 - 4z}} \\ &= \sum_{n \geq 0} 2^{2n-3} z^n - \frac{1}{2} \sum_{n \geq 0} \binom{2n-2}{n-1} z^n. \end{aligned}$$

Der exakte Erwartungswert der interne Pfadlänge ist dann durch

$$\begin{aligned} \mathbb{E}(F(S_n)) &= \frac{[z^n] S_u(z, 1)}{[z^n] S(z, 1)} \\ &= \frac{2^{2n-3} - \binom{2n-2}{n-1}}{\frac{1}{n} \binom{2n-2}{n-1}} \\ &= n \left( \frac{(n-1)! 2^{2n-3}}{(2n-2)!} - 1 \right) \end{aligned}$$

gegeben. Wir können das Ergebnis mithilfe der Stirlingschen Formel ( $n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$ ) vereinfachen, und erhalten

$$\begin{aligned} \mathbb{E}(F(S_n)) &\sim \frac{\sqrt{\pi(n-1)} n 2^{2n-3}}{2^{2n-2}} - n \\ &\sim \frac{\sqrt{\pi n n} 2^{2n-3}}{2^{2n-2}} - n \\ &= \frac{\sqrt{\pi} n^{3/2}}{2} - n. \end{aligned}$$

Manche Ergebnisse sind aber viel schwieriger zu erhalten. Dafür benötigen wir folgenden Satz (vgl. mit [15, Proposition VII.3]):

**Satz 3.1.28** Der Mittelwert und die Varianz der internen Pfadlänge in einem zufälligen einfach erzeugten Baum  $S_n$  mit  $n$  Knoten besitzen folgende asymptotische Darstellung

$$\mathbb{E}(F(S_n)) = \frac{d_1 \sqrt{\pi} n^{3/2}}{2\tau} (1 + \mathcal{O}(n^{-1/2})) \quad (3.37)$$

und

$$\mathbb{V}(F(S_n)) = \frac{(10 - 3\pi) d_1^2 n^3}{12\tau^2} (1 + \mathcal{O}(n^{-1/2})) \quad (3.38)$$

wobei  $\tau$  bzw.  $d_1$  wie in Proposition 2.1.12 bzw. im Satz 2.1.14 sind.

*Beweis:* Wir leiten  $S(z, u)$  nach  $u$  ab, und setzen  $u = 1$ . Wir erhalten:

$$\begin{aligned} S_u(z, 1) &= z\Phi'(S(z, 1))(zS_z(z, 1) + S_u(z, 1)) \\ \Leftrightarrow S_u(z, 1) &= \frac{z\Phi'(S(z, 1))zS_z(z, 1)}{1 - z\Phi'(S(z, 1))} \\ &= \frac{z\Phi'(S(z))zS'(z)}{1 - z\Phi'(S(z))}. \end{aligned}$$

Wegen  $S'(z) = \Phi(S(z)) + z\Phi'(S(z))S'(z) = \frac{S(z)}{z} + z\Phi'(S(z))S'(z)$  gilt  $z\Phi'(S(z)) = 1 - \frac{S(z)}{zS'(z)}$  und daher

$$\begin{aligned} S_u(z, 1) &= (zS'(z) - S(z)) \frac{zS'(z)}{S(z)} \\ &= \frac{z^2S'(z)^2}{S(z)} - zS'(z). \end{aligned}$$

Wir wissen schon (vgl. mit Satz 2.1.14), dass  $S(z) = \tau - d_1 \left(1 - \frac{z}{\rho}\right)^{1/2} + \mathcal{O}\left(1 - \frac{z}{\rho}\right)$ . Daher haben wir  $zS'(z) \sim -\frac{d_1}{2} \left(1 - \frac{z}{\rho}\right)^{-1/2}$  und  $(zS'(z))^2 \sim \frac{d_1^2}{4} \left(1 - \frac{z}{\rho}\right)^{-1}$ . Es folgt:

$$S_u(z, 1) = \frac{d_1^2}{4\tau} \left(1 - \frac{z}{\rho}\right)^{-1} + \mathcal{O}\left(\left(1 - \frac{z}{\rho}\right)^{-1/2}\right).$$

Wir bekommen dann mithilfe der Sätze 1.2.16 und 1.2.25 (mit  $\Gamma(1) = 1$ )

$$[z^n]S_u(z, 1) = \frac{d_1^2 n^0 \rho^{-n}}{4\tau\Gamma(1)} \left(1 + \mathcal{O}\left(n^{1/2-1}\right)\right) = \frac{d_1^2 \rho^{-n}}{4\tau} \left(1 + \mathcal{O}\left(n^{-1/2}\right)\right).$$

Nun können wir den Mittelwert bestimmen:

$$\begin{aligned} \mathbb{E}(F(S_n)) &= \frac{[z^n]S_u(z, 1)}{[z^n]S(z, 1)} \\ &= \frac{\frac{d_1^2 \rho^{-n}}{4\tau} \left(1 + \mathcal{O}\left(n^{-1/2}\right)\right)}{\frac{d_1 \rho^{-n}}{2\sqrt{\pi n^{3/2}}} \left(1 + \mathcal{O}\left(n^{-1}\right)\right)} \\ &= \frac{d_1 \sqrt{\pi n^{3/2}}}{2\tau} \left(1 + \mathcal{O}\left(n^{-1/2}\right)\right). \end{aligned}$$

Die Varianz wird analog berechnet, indem man die zweite Ableitung  $S_{uu}(z, 1)$  berechnet und dann (1.45) anwendet.  $\square$

Wegen der Konstruktion von  $F(S_n)$  ist die Grenzverteilung offensichtlich.

**Proposition 3.1.29** Für die interne Pfadlänge  $F(S_n)$  eines einfach erzeugten Baumes mit  $n$  Knoten gilt folgende schwache Konvergenz

$$\frac{F(S_n) - \mathbb{E}(F(S_n))}{\sqrt{\mathbb{V}(F(S_n))}} \xrightarrow{d} N \sim \mathcal{N}(0, 1).$$

*Beweis:* Die Maut-Funktion  $f(S_n) = |S_n| - 1$  erfüllt klarerweise die Bedingung (3.4). Weiters ist wegen  $10 - 3\pi > 0$  und  $d_1, \tau \neq 0$  die Varianz immer positiv. (3.7) liefert uns das gewünschte Ergebnis.  $\square$

**Beispiel 3.1.30** *Wir bestimmen asymptotischen Mittelwert und Varianz der internen Pfadlänge in einem zufälligen ebenen Wurzelbaum und vergleichen das Ergebnis von  $\mathbb{E}(F(S_n))$  mit der exakten Formel. Wir haben  $\tau = d_1 = \frac{1}{2}$  (vgl. Beispiel 2.1.19). Mit Hilfe von (3.37) erhalten wir*

$$\mathbb{E}(F(S_n)) = \frac{\sqrt{\pi}n^{3/2}}{2} \left(1 + \mathcal{O}(n^{-1/2})\right).$$

Für die Varianz gilt

$$\mathbb{V}(F(S_n)) = \frac{(10 - 3\pi)n^3}{12} \left(1 + \mathcal{O}(n^{-1/2})\right).$$

Die Ergebnisse für die weiteren spezifischen einfach erzeugten Bäume werden genauso berechnet (siehe die Tabelle am Ende des Abschnitts).

## Pólya-Bäume

Das Verfahren, um Ergebnisse für Pólya-Bäume zu ermitteln, ist dasselbe wie bei einfach erzeugten Bäumen. Der folgende Abschnitt ist von [35] motiviert. Wie üblich definieren wir eine bivariate erzeugende Funktion (vgl. mit [35, Teil 2]).

**Satz 3.1.31** *Die bivariate erzeugende Funktion  $S(z, u)$  für die interne Pfadlänge in Pólya-Bäume erfüllt folgende Funktionalgleichung*

$$P(z, u) = z \exp \left( \sum_{i \geq 1} \frac{P(z^i u^i, u^i)}{i} \right). \quad (3.39)$$

*Beweis:* (3.35) mit  $\psi(P(z, u)) = z \exp \left( \sum_{i \geq 1} \frac{P(z^i, u^i)}{i} \right)$ .  $\square$

Dann haben wir (vgl. mit [35, Abschnitt 3] mit  $r = 1$  und  $s = 0$ ):

**Satz 3.1.32** *Der Erwartungswert und die Varianz der interne Pfadlänge bei einem zufälligen Pólya-Baum  $P_n$  mit  $n$  Knoten erfüllen*

$$\mathbb{E}(F(P_n)) = \frac{\sqrt{\pi}n^{3/2}}{d_1} (1 + \mathcal{O}(n^{-1/2})), \quad (3.40)$$

und

$$\mathbb{V}(F(P_n)) = \frac{(10 - 3\pi)n^3}{3d_1^2} (1 + \mathcal{O}(n^{-1/2})), \quad (3.41)$$

wobei  $d_1$  wie im Satz 2.2.5 ist.



*Beweis:* Wir zeigen nur (3.40). Wir leiten  $P(z, u)$  nach  $u$  ab, und setzen  $u = 1$ . Wir bekommen

$$\begin{aligned} P_u(z, 1) &= z \exp \left( \sum_{i \geq 1} \frac{1}{i} P(z^i, 1) \right) \sum_{i \geq 1} \left( z^i P_z(z^i, 1) + P_u(z^i, 1) \right) \\ &= P(z, 1) \sum_{i \geq 1} \left( z^i P_z(z^i, 1) + P_u(z^i, 1) \right) \\ &= P'(z) \sum_{i \geq 1} \left( z^i P'(z^i) + P_u(z^i, 1) \right). \end{aligned}$$

Daher haben wir

$$\begin{aligned} (1 - P(z))P_u(z, 1) &= P(z) \left( zP'(z) + \sum_{i \geq 2} \left( z^i P'(z^i) + P_u(z^i, 1) \right) \right) \\ \Leftrightarrow P_u(z, 1) &= \frac{P(z)}{1 - P(z)} \left( zP'(z) + \sum_{i \geq 2} \left( z^i P'(z^i) + P_u(z^i, 1) \right) \right). \end{aligned}$$

Die Summe ist klarerweise in einem Konvergenzkreis mit Radius  $\tilde{\rho} > \rho$  analytisch. Aus Satz 2.2.5 haben wir  $P(z) \sim 1 - d_1 \left(1 - \frac{z}{\rho}\right)^{1/2}$ ,  $zP'(z) \sim \frac{1}{2}d_1 \left(1 - \frac{z}{\rho}\right)^{-1/2}$ , also

$$\frac{P(z)}{1 - P(z)} \sim \frac{1 - d_1 \left(1 - \frac{z}{\rho}\right)^{1/2}}{d_1 \left(1 - \frac{z}{\rho}\right)^{1/2}} = \frac{1}{d_1} \left(1 - \frac{z}{\rho}\right)^{-1/2} + \mathcal{O}(1),$$

und daher

$$\begin{aligned} \frac{P(z)}{1 - P(z)} zP'(z) &\sim \left( \frac{1}{d_1} \left(1 - \frac{z}{\rho}\right)^{-1/2} + \mathcal{O}(1) \right) \frac{1}{2}d_1 \left(1 - \frac{z}{\rho}\right)^{-1/2} \\ &= \frac{1}{2} \left(1 - \frac{z}{\rho}\right)^{-1} + \mathcal{O} \left( \left(1 - \frac{z}{\rho}\right)^{-1/2} \right). \end{aligned}$$

Wir erhalten

$$P_u(z, 1) = \frac{1}{2} \left(1 - \frac{z}{\rho}\right)^{-1} + \mathcal{O} \left( \left(1 - \frac{z}{\rho}\right)^{-1/2} \right).$$

Für die Koeffizienten von  $P_u(z, 1)$  bekommen wir dann (analog wie im Satz 3.1.28):

$$[z^n]P_u(z, 1) = \frac{1}{2}n^0\rho^{-n}(1 + \mathcal{O}(n^{-1/2})) = \frac{1}{2}\rho^{-n}(1 + \mathcal{O}(n^{-1/2}))$$

Daraus folgt

$$\begin{aligned} \mathbb{E}(F(P_n)) &= \frac{[z^n]P_u(z, 1)}{[z^n]P(z, 1)} \\ &= \frac{\frac{1}{2}\rho^{-n}(1 + \mathcal{O}(n^{-1/2}))}{\frac{d_1\rho^{-n}}{2\sqrt{\pi}n^{3/2}}(1 + \mathcal{O}(n^{-1}))} \\ &= \frac{\sqrt{\pi}n^{3/2}}{d_1}(1 + \mathcal{O}(n^{-1/2})). \end{aligned}$$

□

**Proposition 3.1.33** Die interne Pfadlänge  $F(P_n)$  eines zufälligen Pólya-Baumes der Größe  $n$  erfüllt folgenden Grenzwertsatz:

$$\frac{F(P_n) - \mathbb{E}(F(P_n))}{\sqrt{\mathbb{V}(F(P_n))}} \xrightarrow{d} N \sim \mathcal{N}(0, 1)$$

*Beweis:* Analog wie in Proposition 3.1.29. □

Wir vergleichen unsere Ergebnisse:

Bäume	$\mathbb{E}(F(T_n))$	$\mathbb{V}(F(T_n))$
Einfach erzeugt	$\sim \frac{d_1 \sqrt{\pi} n^{3/2}}{2\tau}$ mit $d_1 = \sqrt{\frac{2\Phi(\tau)}{\Phi''(\tau)}}$	$\sim \frac{(10-3\pi)d_1^2 n^3}{12\tau^2}$
Catalan	$\sim \frac{\sqrt{\pi} n^{3/2}}{2}$	$\sim \frac{(10-3\pi)n^3}{12}$
Cayley	$\sim \sqrt{\frac{\pi}{2}} n^{3/2}$	$\sim \frac{(10-3\pi)n^3}{6}$
Binär 1	$\sim \sqrt{\frac{\pi}{2}} n^{3/2}$	$\sim \frac{(10-3\pi)n^3}{6}$
Binär 2	$\sim \sqrt{\pi} n^{3/2}$	$\sim \frac{(10-3\pi)n^3}{3}$
Motzkin	$\sim \frac{\sqrt{3\pi} n^{3/2}}{2}$	$\sim \frac{(10-3\pi)n^3}{4}$
Pólya	$\sim \frac{\sqrt{\pi} n^{3/2}}{d_1}$ mit $d_1 = \sqrt{2e\rho\zeta'(\rho)} \approx 2.6811266$	$\sim \frac{(10-3\pi)n^3}{3d_1^2}$

Wir werden im Abschnitt 3.3.3 die interne Pfadlänge wieder anwenden, um Resultate für den Wiener-Index zu berechnen.

## 3.2 Die Form von einfach erzeugten Bäumen und Pólya-Bäumen

Nun kümmern wir uns um Parameter, die der Form eines zufälligen ebenen Wurzelbaumes entsprechen. Die folgenden Begriffe wurden zum ersten Mal von Aldous in [1] und [2] definiert und danach von Flajolet und Odlyzko in [14], Drmota und Gittenberger in [9] und [11] und Drmota in [8, Kapitel 4] analysiert. Zuerst definieren wir:

**Definition 3.2.1** Die Höhe  $H(v)$  eines Knotens  $v$  in einem ebenen Wurzelbaum ist der Abstand zwischen der Wurzel und  $v$ . Allgemeiner ist die Höhe  $H(T)$  eines zufälligen Wurzelbaumes  $T$  der maximale Abstand von der Wurzel zu einem Blatt. Besitzt der Baum  $T$  genau  $n$  Knoten, so bezeichnen wir die entsprechende Höhe durch  $H_n$ . Die Höhe ist eine Zufallsvariable.

Wir können die Höhe alternativ beschreiben. Dafür brauchen wir zusätzliche Begriffe:

**Definition 3.2.2** Sei  $T$  ein ebener Wurzelbaum mit Wurzel  $r$ . Die Tiefensuche (engl. *depth-first-search*) kann als ein Pfad  $(v(i), 1 \leq i \leq 2n - 1)$  über die Knoten von  $T$  folgender Bauart definiert werden:

- Beginne mit  $v(1) = r$ .
- Sei  $v(i)$  der aktuelle Knoten. Untersuche die am weitesten links liegende Kante  $e$ , die zu  $v(i)$  inzident ist und von  $r$  wegführt, welche noch nicht untersuchte ausgehende Knoten hat, und setze daher  $(v(i), v(i + 1)) := e$ .
- Wenn dies nicht möglich ist, setze  $v(i + 1)$  als Elternknoten von  $v(i)$ .
- Wiederhole die zwei vorigen Schritte, bis alle Knoten des Baumes besucht wurden.

Die Tiefensuche induziert einen Parameter  $X_n(t) := H(v(t))$ . Wir können diesen mit  $\sqrt{n}$  skalieren. Wir erhalten folgenden weiteren Parameter (vgl. mit [2, Formel (11)]):

$$C_n(t) := \frac{1}{\sqrt{n}} X_n(2nt), \quad 0 \leq t \leq 1. \quad (3.42)$$

Die zugehörige Folge  $(C_n(t))_{0 \leq t \leq 1}$  bildet klarerweise einen stochastischen Prozess und wird Contour-Prozess genannt.

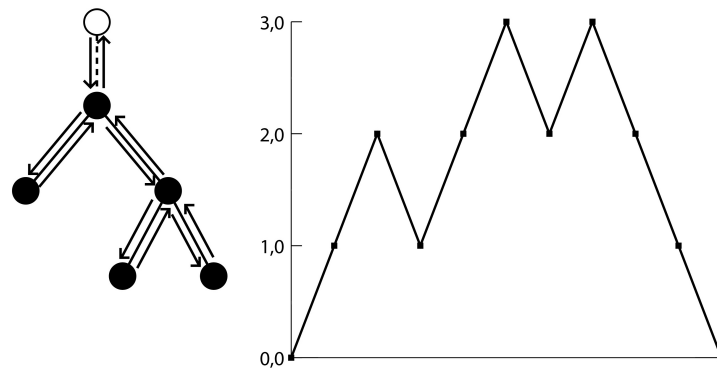


Abbildung 3.1: Die Tiefensuche und der entsprechende Contour-Prozess eines Wurzelbaumes; Die  $y$ -Achse bezeichnet die Höhe der Knoten. Auf der Nullten Ebene wird ein zusätzlichen Knoten gezeichnet, der nicht gezählt wird.

**Bemerkung 3.2.3** Für  $t = 0$  oder  $t = 2n$  setzen wir  $X_n(t) = 0$ . Man sieht ohne Schwierigkeiten, dass  $C_n(t) = 0$  für  $t = 0$  oder  $t = 1$ .

Man sieht leicht, dass die Höhe eines zufälligen Wurzelbaumes erfüllt

$$H_n = \max_{1 \leq t \leq 2n-1} X_n(t) \Leftrightarrow \frac{1}{\sqrt{n}} H_n = \max_{0 \leq t \leq 1} C_n(t). \quad (3.43)$$

Diese Darstellung wird eine wesentliche Rolle spielen im Beweis von bestimmten Konvergenzsätzen im Abschnitt 3.2.2.

**Definition 3.2.4** Unter dem Profil eines zufälligen ebenen Wurzelbaumes  $T$  versteht man eine Folge  $(L_T(t))_{t \geq 0}$  wobei  $L_T(t)$  die Anzahl der Knoten der Höhe  $t$  bezeichnet. Gilt  $|T| = n$ , so schreiben wir  $(L_n(t))_{t \geq 0}$ . Das Profil ist per Definition ein stochastischer Prozess in diskreter Zeit mit der Höhe als Parameterraum. Damit können wir die Breite (engl. width) wie folgt definieren:

$$W_n := \sup_{t \geq 0} L_n(t) = \max_{t \geq 0} L_n(t). \quad (3.44)$$

Interessant ist für  $L_n(t)$  der Fall  $t \notin \mathbb{N}_0$  (vgl. mit [11, Formel (8)]):

**Bemerkung 3.2.5** Falls  $t$  keine nichtnegative Ganzzahl ist, dann wird  $L_n(t)$  iterativ durch

$$L_n(t) = (\lceil t \rceil - t)L_n(\lfloor t \rfloor) + (t - \lfloor t \rfloor)L_n(\lceil t \rceil)$$

berechnet.

Wir können  $L_n(t)$  mit  $\sqrt{n}$  skalieren (siehe [8, Theorem 4.10] oder [11, Theorem 3]). Daher erhalten wir mit

$$l_n(t) = \frac{1}{\sqrt{n}}L_n(t\sqrt{n}), \quad t \geq 0 \quad (3.45)$$

einen weiteren stochastischen Prozess  $(l_n(t))_{t \geq 0}$ , der für einfach erzeugte und Polya Bäume interessante Ergebnisse liefert.

### 3.2.1 Das Profil

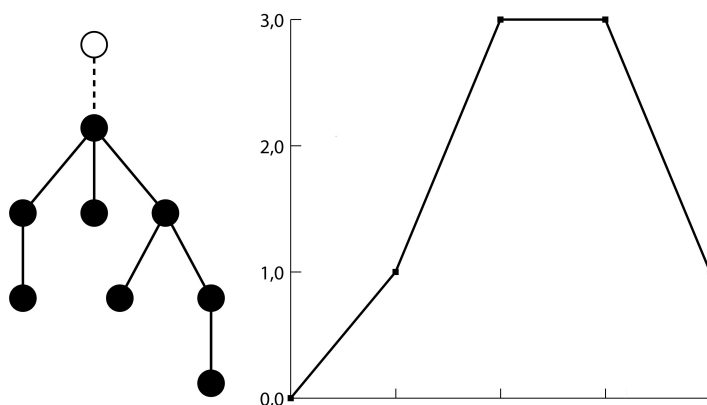


Abbildung 3.2: Das Profil eines Wurzelbaumes  $T$ ; Die  $x$ -Achse bezeichnet die Höhe  $k$  der Knoten, und die  $y$ -Achse bezeichnet  $L_T(k)$ . Auf der Nullten Ebene wird ein Knoten gezeichnet, der aber nicht gezählt wird.

#### Einfach erzeugte Bäume

In [25] haben Meir und Moon Ausdrücke für den Mittelwert und das zweite faktorielle Moment (und daher für die Varianz) von  $L_n(k)$  berechnet. Später hat Aldous (siehe [2]) einen Grenzwertsatz für den Prozess  $(l_n(t))_{t \geq 0}$  ausformuliert, der später von Drmota und Gittenberger in [9] bewiesen wurde.

Wir erinnern uns, dass die Nachkommenverteilung  $\xi$  eines Galton-Watson Prozesses durch

$$\mathbb{P}(\xi = k) = \frac{\tau^k \phi_k}{\Phi(\tau)}$$

gegeben ist, und diese ist kritisch, falls  $\mathbb{E}(\xi) = 1$ .

Unser Ziel ist, folgenden Satz zu zeigen (vgl. mit [8, Theorem 4.10] oder [9, Theorem 1.1]):

**Satz 3.2.6** Ist die Nachkommenverteilung  $\xi$  eines Galton-Watson Prozesses kritisch, aperiodisch und von endlicher Varianz  $\sigma^2 = \mathbb{V}(\xi)$ , so gilt

$$(l_n(t))_{t \geq 0} \xrightarrow{d} \left( \frac{\sigma}{2} l \left( \frac{\sigma}{2} t \right) \right)_{t \geq 0}, \quad n \rightarrow \infty \quad (3.46)$$

in  $C([0, \infty))$ , wobei  $l(t)$  die lokale Zeit einer Brownsche Exkursion der Dauer 1 ist.

Der erste Schritt ist, folgende Aussage zu beweisen:

**Satz 3.2.7** Seien  $\sigma$   $l_n(t)$  und  $l(t)$  wie im vorigen Satz. Weiters seien  $t_1, \dots, t_d$  beliebige aber feste nichtnegative Zahlen. Dann gilt:

$$(l_n(t_1), \dots, l_n(t_d)) \xrightarrow{d} \frac{\sigma}{2} \left( l \left( \frac{\sigma}{2} t_1 \right), \dots, l \left( \frac{\sigma}{2} t_d \right) \right), \quad n \rightarrow \infty \quad (3.47)$$

Die Beweisart von Drmota und Gittenberger (siehe [8, Abschnitt 4.2] und [9]) wird später durchgeführt. Wir berechnen zuerst die Darstellungen von  $\mathbb{E}(L_n(k))$  und  $\mathbb{V}(L_n(k))$ . Wir benötigen weitere erzeugende Funktionen (vgl. mit [8, Unterkapitel 4.2.3] [9, Teil 2] und [25, Abschnitt 2]):

**Definition 3.2.8** Sei  $\mathcal{S}_n$  eine Klasse von einfach erzeugten Bäumen mit  $n$  Knoten. Setze

$$s_{n,m}^{(k)} = \sum_{S \in \mathcal{S}_n, L_S(k)=m} \omega(S) \quad (3.48)$$

als die Anzahl der Bäume aus  $\mathcal{S}_n$ , die auf der Höhe  $k$  genau  $m$  Knoten besitzen. Dann ist die Wahrscheinlichkeit, dass ein zufälliger einfach erzeugter Baum der Größe  $n$  genau  $m$  Knoten der Höhe  $k$  besitzt, gegeben durch

$$\mathbb{P}(L_n(k) = m) = \frac{s_{n,m}^{(k)}}{s_n}. \quad (3.49)$$

Wir definieren dazu

$$s_k(z, u) = \sum_{n,m \geq 0} s_{n,m}^{(k)} z^n u^m = \sum_{n \geq 0} s_n \mathbb{E}(u^{L_n k}) z^n. \quad (3.50)$$

Ist  $\psi$  der Operator für die Funktionalgleichung der einfach erzeugten Bäume (also  $\mathcal{S} = \psi(\mathcal{S}) \leftrightarrow S(z) = z\Phi(S(z))$ ), und  $\tilde{\mathcal{S}}^{(k)}$  die Struktur von  $s_k(z, u)$ , so gilt

$$\tilde{\mathcal{S}}^{(k)} = \psi^k(\{\mu \circ\} \times \mathcal{S}), \quad (3.51)$$

wobei  $\mu$  das Kennzeichen bezeichnet. Daher besitzt die Reihe  $s_k(z, u)$  folgende rekursive Darstellung:

$$s_0(z, u) = uS(z) \quad (3.52)$$

$$s_{k+1}(z, u) = z\Phi(s_k(z, u)), \quad k \geq 0 \quad (3.53)$$

Wir können  $s_{n,m}^{(k)}$  und  $s_k(z, u)$  wie folgt verallgemeinern (siehe [8, Abschnitt 4.2.3]):

**Definition 3.2.9** Sei  $d \in \mathbb{N}$ . Seien weiters  $k_i$  nichtnegative Ganzzahlen mit  $0 \leq k_1 < k_2 < \dots < k_d$  und  $m_i \in \mathbb{N}_0$  ( $i = 1, \dots, d$ ). Die Anzahl der Bäume aus  $\mathcal{S}_n$ , die auf der Höhe  $k_i$  genau  $m_i$  ( $i = 1, \dots, d$ ) Knoten haben, ist gegeben durch:

$$s_{n;m_1, \dots, m_d}^{(k_1, \dots, k_d)} := \sum_{\substack{S \in \mathcal{S}_n, L_S(k_i) = m_i \\ i=1, \dots, d}} \omega(S) \quad (3.54)$$

Analog zur vorigen Definition haben wir

$$\mathbb{P}(L_n(k_1) = m_1, L_n(k_2) = m_2, \dots, L_n(k_d) = m_d) = \frac{s_{n;m_1, \dots, m_d}^{(k_1, \dots, k_d)}}{s_n} \quad (3.55)$$

und

$$s_{k_1, \dots, k_d}(z, u_1, \dots, u_d) = \sum_{n, m_1, \dots, m_d \geq 0} s_{n;m_1, \dots, m_d}^{(k_1, \dots, k_d)} z^n \mathbf{u}^{\mathbf{m}} = \sum_{n \geq 0} s_n \mathbb{E}(u_1^{L_n(k_1)} \dots u_d^{L_n(k_d)}) z^n. \quad (3.56)$$

Diese Funktion erfüllt ebenfalls eine Rekursion, nämlich:

$$s_{0, k_2, \dots, k_d}(z, u_1, \dots, u_d) = u_1 s_{k_2, \dots, k_d}(z, u_2, \dots, u_d), \quad (3.57)$$

$$s_{k_1+1, k_2+1, \dots, k_d+1}(z, u_1, \dots, u_d) = z \Phi(s_{k_1, \dots, k_d}(z, u_1, \dots, u_d)). \quad (3.58)$$

**Bemerkung 3.2.10** Wir können die Funktionen  $s_{k_1, \dots, k_d}(z, u_1, \dots, u_d)$  vereinfachen (siehe [8, Remark 4.19]). Seien nämlich die Funktionen  $S_k(z, u)$  definiert durch

$$S_0(z, u) = u \quad (3.59)$$

$$S_{k+1}(z, u) = z \Phi(S_k(z, u)), \quad k \geq 0 \quad (3.60)$$

Der Zusammenhang zwischen diesen und  $s_{k_1, \dots, k_d}(z, u_1, \dots, u_d)$  ist gegeben durch

$$s_{k_1, \dots, k_d}(z, u_1, \dots, u_d) = S_{k_1} \left( z, u_1 S_{k_2 - k_1}(z, u_2 S_{k_3 - k_2}(z, \dots, u_{d-1} S_{k_d - k_{d-1}}(z, u_d S(z)))) \right). \quad (3.61)$$

Wir betrachten zuerst den Fall  $d = 1$ . Mithilfe dieser Funktion können wir aus [25, Theorem 2.1] weitere definieren, nämlich:

**Definition 3.2.11** Sei  $L_n(k)$  die Anzahl der Knoten der Höhe  $k$  in einem zufälligen einfach erzeugten Baum der Größe  $n$ . Für den entsprechenden Erwartungswert bzw. das zweite faktorielle Moment definieren wir:

$$E_k(z) = \sum_{n \geq 0} \mathbb{E}(L_n(k)) s_n z^n = \left. \frac{\partial}{\partial z} s_k(z, u) \right|_{u=1} \quad (3.62)$$

und

$$E_k^{(2)}(z) = \sum_{n \geq 0} \mathbb{E}(L_n(k)(L_n(k) - 1)) s_n z^n = \left. \frac{\partial^2}{\partial z^2} s_k(z, u) \right|_{u=1}. \quad (3.63)$$

Diese Funktionen besitzen Ausdrücke, die nur von  $S(z)$ ,  $\Phi(t)$  und  $k$  abhängen (siehe [25, Theorem 4.1]).

**Satz 3.2.12** Sei  $k \in \mathbb{N}_0$ . Dann gelten

$$E_k(z) = S(z)(z\Phi'(S(z)))^k \quad (3.64)$$

und

$$E_k^{(2)}(z) = S(z)^2(z\Phi''(S(z))) \sum_{j=k-1}^{2k-2} ((z\Phi'(S(z)))^j) \quad (3.65)$$

*Beweis:* Wegen (3.62) und (3.53) haben wir:

$$\begin{aligned} E_{k+1}(z) &= \frac{\partial}{\partial u} s_{k+1}(z, 1) \\ &= \frac{\partial}{\partial u} (z\Phi(s_k(z, 1))) \\ &= z\Phi'(s_k(z, 1)) \frac{\partial}{\partial u} s_k(z, 1) \\ &= z\Phi'(S(z)) E_k(z) \\ &= (z\Phi'(S(z)))^2 E_{k-1}(z) \\ &= (z\Phi'(S(z)))^{k+1} E_0(z) \end{aligned}$$

Da  $s_0(z, u) = uS(z)$  (3.52) haben wir klarerweise  $E_0(z) = S(z)$  und daher (3.64). Mit ähnlicher Überlegung und (3.63) erhalten wir (3.65).  $\square$

Meir und Moon haben in [25, Theorem 4.2] gezeigt, dass die mittlere Anzahl der Knoten der Höhe  $k$  in einem zufälligen einfach erzeugten Baum folgende (erste) asymptotische Darstellung besitzt:

$$\mathbb{E}(L_n(k)) \sim Ak + 1, \quad n \rightarrow \infty \quad (3.66)$$

wobei  $A = \tau\rho\Phi''(\tau)$  ( $\tau$  und  $\rho$  sind wie in Proposition 2.1.12). Insbesondere gilt

$$\mathbb{V}(L_n(k)) \sim \frac{1}{2}A^2k(k-3) + k(A + \tau^2\rho\Phi'''(\tau)), \quad n \rightarrow \infty \quad (3.67)$$

für die Varianz. Wir haben also einen Ausdruck für  $\mathbb{E}(L_n(k))$  gefunden. Doch können wir eine genauere Darstellung berechnen, nämlich:

**Satz 3.2.13** Setze  $k = \lfloor \kappa\sqrt{n} \rfloor$  für  $\kappa > 0$ . Dann erfüllt die Mittlere Anzahl der Knoten der Höhe  $k$  in einem zufälligen einfach erzeugten Baum folgende Asymptotik:

$$\mathbb{E}(L_n(k)) \sim \sigma^2 k e^{-\sigma^2 k^2 / 2n} \quad (3.68)$$

wobei  $\sigma^2$  die Varianz der entsprechenden kritischen Nachkommenverteilung  $\xi$  eines Galton-Watson Prozesses bezeichnet.

Die Aussage wurde von Meir und Moon bewiesen (siehe [25, Theorem 4.3]). Doch gibt es eine effizientere Weise, diesen Ausdruck zu zeigen (vgl. [8, Abschnitt 4.2.5]). Der Beweis wird später durchgeführt.

Wir kümmern uns nun wieder um die Funktion  $s_k(z, u)$ , um eine Grenzverteilung für das Profil zu bestimmen. Wir sollen danach die Bedingungen des Satzes 1.3.32 überprüfen, also die schwache Konvergenz von  $l_n(\kappa)$  gegen  $\frac{\sigma}{2}l\left(\frac{\sigma}{2}\kappa\right)$  für ein nichtnegatives  $\kappa$

beweisen, und zeigen, dass der Prozess  $(l_n(t))_{t \geq 0}$  in  $C[0, \infty)$  straff ist. Aus der Darstellung (2.16) von  $S(z)$  erhalten wir einen asymptotischen Ausdruck für  $s_k(z, u)$  und allgemeiner für  $S_k(z, u)$ . Wir erinnern uns, dass

$$S(z) = \tau - \frac{\tau\sqrt{2}}{\sigma} \left(1 - \frac{z}{\rho}\right)^{1/2} + \mathcal{O}\left(\left|1 - \frac{z}{\rho}\right|\right), \quad z \rightarrow \rho.$$

Nun setzen wir  $\alpha := \alpha(z) = z\Phi'(S(z))$  und  $\beta = \beta(z) = z\Phi''(S(z))/2$ . Aus der vorigen Formel und dem Beweis von Korollar 2.1.17 wobei  $\alpha(z) = 1 - \frac{S(z)}{zS'(z)}$  und  $\sigma = \frac{\tau\sqrt{2}}{d_1}$  bekommen wir

$$\alpha(z) = 1 - \sigma\sqrt{2} \left(1 - \frac{z}{\rho}\right)^{1/2} + \mathcal{O}\left(\left|1 - \frac{z}{\rho}\right|\right), \quad z \rightarrow \rho. \quad (3.69)$$

Die 2. Ableitung von  $S(z)$  führt uns zu

$$\beta(z) = \frac{\sigma^2}{2\tau} + \mathcal{O}\left(\left(1 - \frac{z}{\rho}\right)^{1/2}\right), \quad z \rightarrow \rho. \quad (3.70)$$

Nun haben wir hinreichend Informationen, um einen asymptotischen Ausdruck für  $s_k(z, u)$  zu bestimmen (vgl. mit Lemma 4.20 und Proposition 4.21 in [8] oder mit Lemma 2.1 und Theorem 2.1 in [9]):

**Lemma 3.2.14** *Setze  $z = \rho\left(1 + \frac{x}{n}\right)$ ,  $u = e^{it\sqrt{n}}$  und  $k = \lfloor \kappa\sqrt{n} \rfloor$ , wobei  $\rho$  die dominante Singularität von  $S(z)$  ist. Weiters sei angenommen, dass  $|u - S(z)| = \mathcal{O}\left(\frac{1}{\sqrt{n}}\right)$ ,  $\frac{x}{n} \rightarrow 0$ ,  $\arg(-x) < \pi$  und  $\left|1 - \sqrt{\frac{-x}{n}}\right| \leq 1 + \frac{C}{\sqrt{n}}$  wobei  $C > 0$  eine Konstante ist. Dann erfüllt die Folge  $w_k(z, u) = s_k(z, u) - S(z)$*

$$w_k(z, u) = \frac{1}{\sqrt{n}} \frac{it\sqrt{-2x}\tau \exp\left(-\frac{1}{2}\sigma\kappa\sqrt{-2x}\right) \left(1 + \mathcal{O}\left(\frac{x}{\sqrt{n}}\right)\right)}{\sqrt{-2x} \exp\left(\frac{1}{2}\sigma\kappa\sqrt{-2x}\right) \left(1 + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right)\right) - it\sigma \sinh\left(\frac{1}{2}\sigma\kappa\sqrt{-2x}\right)} \quad (3.71)$$

gleichmäßig für  $k = \mathcal{O}(\sqrt{n})$ . Dabei ist  $\sigma^2 = \frac{\tau^2\Phi''(\tau)}{\Phi(\tau)}$  die Varianz von  $\xi$ .  $\tau$  ist wie in Proposition 2.1.12.

*Beweis:* Sei  $W_k(z, u) := S_k(z, u) - S(z)$ . Es gilt  $W_0 = u - S(z)$ . Per Induktion kann man zeigen, dass es eine Konstante  $c > 0$  gibt, für die gilt

$$\frac{|W_0||\alpha|^k}{1 + ck|W_0|} \leq |W_k| \leq \frac{|W_0||\alpha|^k}{1 - ck|W_0|},$$

wobei  $k|W_0| \leq 1/(2C)$  (siehe [8, Formel (4.36)] oder [9, Lemma 3.1]).

Daher haben wir

$$\begin{aligned} W_{k+1} &= \alpha W_k + \beta W_k^2 + \mathcal{O}\left(|W_k|^3\right) \\ \Leftrightarrow \frac{1}{W_{k+1}} &= \frac{1}{\alpha W_k} \frac{1}{1 + \beta W_k/\alpha + \mathcal{O}\left(|W_k|^2\right)} \\ &= \frac{1}{\alpha w_k} - \frac{\beta}{\alpha^2} + \mathcal{O}\left(\frac{|w_k|}{|\alpha|}\right). \end{aligned}$$



Multiplikation mit  $\alpha^{k+1}$  liefert uns  $\frac{\alpha^{k+1}}{W_{k+1}} = \frac{\alpha^k}{W_k} - \beta\alpha^{k-1} + \mathcal{O}(|W_k||\alpha|^k)$  und wegen  $W_k = \mathcal{O}(W_0\alpha^k)$  folgt

$$\frac{\alpha^k}{W_k} = \frac{1}{W_0} - \frac{\beta}{\alpha} \frac{1 - \alpha^k}{1 - \alpha} + \mathcal{O}\left(|W_0| \frac{\beta}{\alpha} \left| \frac{1 - \alpha^{2k}}{1 - \alpha^2} \right|\right).$$

Wir erhalten

$$S_k(z, u) = S(z) + \frac{(u - S(z))\alpha^k}{1 - \frac{(u-S(z))\beta}{\alpha(1-\alpha)} + \frac{(u-S(z))\beta}{\alpha(1-\alpha)}\alpha^k + \mathcal{O}\left(|u - S(z)|^2 \frac{1 - |\alpha|^{2k}}{1 - |\alpha|^2}\right)} \quad (3.72)$$

für  $k|u - S(z)| \leq \epsilon$  und  $z \in \Delta_\epsilon$  wobei

$$\Delta_\epsilon = \{z \in \mathbb{C} : |z - \rho| < \epsilon, |\arg(z - \rho)| > \theta\}, \quad (3.73)$$

wobei  $\epsilon > 0$  und  $0 < \theta < \frac{\pi}{2}$ . (vgl. mit [8, Lemma 4.20] und [11, Teil 4.1]). Nun betrachten wir den Fall  $d = 1$ . Wegen (3.52)-(3.53) und (3.59)-(3.60) haben wir  $s_k(z, u) = S_k(z, uS(z))$ .

Wir können mit  $z = \rho\left(1 + \frac{x}{n}\right)$  unsere Darstellungen (2.16), (3.69) und (3.70) umschreiben. Es gelten

$$S(z) = \tau - \frac{\tau\sqrt{2}}{\sigma} \sqrt{\frac{-x}{n}} + \mathcal{O}\left(\frac{|x|}{n}\right) \quad (3.74)$$

$$\alpha = 1 - \sigma\sqrt{2} \sqrt{\frac{-x}{n}} + \mathcal{O}\left(\frac{|x|}{n}\right) \quad (3.75)$$

$$\beta = \frac{\sigma^2}{2\tau} + \mathcal{O}\left(\sqrt{\frac{-x}{n}}\right) \quad (3.76)$$

und

$$\alpha^k = \exp(-\sigma\kappa\sqrt{-2x})(1 + \mathcal{O}(\kappa x/\sqrt{n})). \quad (3.77)$$

Dabei gilt

$$\frac{\beta(u-1)S(z)}{\alpha(1-\alpha)} \sim \frac{(\exp(it\sqrt{n}) - 1)\sigma\sqrt{n}}{2\sqrt{-2x}}$$

und (mithilfe von (3.72))

$$\begin{aligned}
w_k(z, u) &= \frac{(u-1)S(z)\alpha^k}{1 - \frac{(u-1)S(z)\beta}{\alpha(1-\alpha)} + \frac{(u-1)S(z)\beta}{\alpha(1-\alpha)}\alpha^k + \mathcal{O}\left(\frac{|(u-1)|^2 1-|\alpha|^{2k}}{1-|\alpha|^2}\right)} \\
&= \frac{(\exp^{it\sqrt{n}} - 1)\tau \exp(-\sigma\kappa\sqrt{-2x}) \left(1 + \mathcal{O}\left(\frac{\kappa x}{\sqrt{n}}\right)\right)}{1 - \frac{(\exp(it\sqrt{n})-1)\sigma\sqrt{n}}{2\sqrt{-2x}} + \frac{(\exp(it\sqrt{n})-1)\sigma\sqrt{n}}{2\sqrt{-2x}} \exp(-\sigma\kappa\sqrt{-2x}) + \mathcal{O}\left(\frac{|x|^{3/2}}{\sqrt{n}}\right)} \\
&= \frac{1}{\sqrt{n}} \frac{it\tau \exp(-\sigma\kappa\sqrt{-2x}) \left(1 + \mathcal{O}\left(\frac{\kappa x}{\sqrt{n}}\right)\right)}{1 - \frac{it\sigma}{2\sqrt{-2x}} + \frac{it\sigma}{2\sqrt{-2x}} \exp(-\sigma\kappa\sqrt{-2x}) + \mathcal{O}\left(\frac{|x|^{3/2}}{\sqrt{n}}\right)} \\
&= \frac{1}{\sqrt{n}} \frac{it\sqrt{-2x}\tau \exp(-\sigma\kappa\sqrt{-2x}) \left(1 + \mathcal{O}\left(\frac{\kappa x}{\sqrt{n}}\right)\right)}{\sqrt{-2x} - \frac{it\sigma}{2} + \frac{it\sigma}{2} \exp(-\sigma\kappa\sqrt{-2x}) + \mathcal{O}\left(\sqrt{\frac{|x|}{n}}\right)} \\
&= \frac{1}{\sqrt{n}} \frac{it\sqrt{-2x}\tau \exp\left(-\frac{1}{2}\sigma\kappa\sqrt{-2x}\right) \left(1 + \mathcal{O}\left(\frac{x}{\sqrt{n}}\right)\right)}{\sqrt{-2x} \exp\left(\frac{1}{2}\sigma\kappa\sqrt{-2x}\right) \left(1 + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right)\right) - it\sigma \sinh\left(\frac{1}{2}\sigma\kappa\sqrt{-2x}\right)}.
\end{aligned}$$

□

Nun können wir die Grenzverteilung ermitteln (siehe [8, Proposition 4.21] oder [9, Proposition 4.1]):

**Satz 3.2.15** *Ist  $\kappa$  eine nichtnegative Zahl, so gilt*

$$\lim_{n \rightarrow \infty} \mathbb{E} \left( e^{it \frac{1}{\sqrt{n}} L_n(\kappa\sqrt{n})} \right) = \mathbb{E} \left( e^{it \frac{\sigma}{2} l\left(\frac{\sigma}{2}\kappa\right)} \right), \quad (3.78)$$

wobei  $|t| \leq \frac{c}{\tau\kappa}$ . Wegen des Stetigkeitssatzes für stetige Grenzverteilungen (Satz 1.3.19) ist die Aussage äquivalent zu

$$l_n(\kappa) \xrightarrow{d} \frac{\sigma}{2} l\left(\frac{\sigma}{2}\kappa\right). \quad (3.79)$$

*Beweis:* Zuerst sei bemerkt, dass

$$\mathbb{E} \left( e^{it \frac{1}{\sqrt{n}} L_n(k)} \right) = \frac{[z^n] s_k(z, e^{it/\sqrt{n}})}{s_n}.$$

Die Idee ist, diesen Wert entlang einer geeigneten Contour  $\Gamma$  zu bestimmen (vgl. [8, Proposition 4.21]). Es gilt dann:

$$\mathbb{E} \left( e^{it \frac{1}{\sqrt{n}} L_n(k)} \right) = \frac{1}{2\pi i s_n} \int_{\Gamma} s_k(z, e^{it/\sqrt{n}}) \frac{dz}{z^{n+1}} = 1 + \frac{1}{2\pi i s_n} \int_{\Gamma} w_k(z, e^{it/\sqrt{n}}) \frac{dz}{z^{n+1}}$$

Wir wählen den Weg  $\Gamma = \gamma \cup \Gamma'$  bestehend aus:

- einer Linie  $\gamma = \{z = \rho \left(1 + \frac{x}{n}\right) : \operatorname{Re}(x) = c, |\operatorname{Im}(x)| \leq C \log^2(n)\}$ , wobei  $C > 0$  und  $c < 0$  Konstanten sind;
- einem Kreisbogen  $\Gamma'$  um 0, der  $\Gamma$  zu einer geschlossenen Kurve macht.

Für  $z \in \Gamma'$  ist  $\frac{|z^{-n-1}|}{s_n} = \mathcal{O}(n^{3/2}e^{-\log^2(n)})$  und  $|s_k(z, e^{it/\sqrt{n}})| = \mathcal{O}(1)$ , also der Beitrag ist exponentiell klein. Der Ausdruck konvergiert dann gegen Null wenn  $n \rightarrow \infty$ .

Für  $z \in \gamma$  ist die Darstellung (3.72) gültig. Wir können dann  $s_k(z, u)$  entlang der Linie integrieren. Mithilfe der Substitution  $z = \rho \left(1 + \frac{x}{n}\right)$ ,  $dz = \frac{\rho}{n} dx$  und  $z^{-n} = \rho^{-n} e^{-x} \left(1 + \mathcal{O}\left(\frac{|x|^2}{n}\right)\right)$  bekommen wir:

$$\begin{aligned} \mathbb{E}\left(e^{it\frac{1}{\sqrt{n}}L_n(k)}\right) &\sim 1 + \frac{\sigma\sqrt{2\pi}\rho^n n^{3/2}}{2\pi\tau i} \int_{c-i\log^2(n)}^{c+i\log^2(n)} \frac{1}{n^{3/2}} \frac{it\sqrt{-2x}\tau \exp\left(-\frac{1}{2}\sigma\kappa\sqrt{-2x}\right) \rho^{-n} e^{-x}}{\sqrt{-2x} \exp\left(\frac{1}{2}\sigma\kappa\sqrt{-2x}\right) - it\sigma \sinh\left(\frac{1}{2}\sigma\kappa\sqrt{-2x}\right)} dx \\ &= 1 + \frac{\sigma}{\sqrt{2\pi}} \int_{c-i\log^2(n)}^{c+i\log^2(n)} \frac{t\sqrt{-2x} \exp\left(-\frac{1}{2}\sigma\kappa\sqrt{-2x}\right) e^{-x}}{\sqrt{-2x} \exp\left(\frac{1}{2}\sigma\kappa\sqrt{-2x}\right) - it\sigma \sinh\left(\frac{1}{2}\sigma\kappa\sqrt{-2x}\right)} dx \\ &= 1 + \frac{\sigma\sqrt{2}}{2\sqrt{\pi}} \int_{c-i\log^2(n)}^{c+i\log^2(n)} \frac{t\sqrt{-2x} \exp\left(-\frac{1}{2}\sigma\kappa\sqrt{-2x}\right) e^{-x}}{\sqrt{-2x} \exp\left(\frac{1}{2}\sigma\kappa\sqrt{-2x}\right) - it\sigma \sinh\left(\frac{1}{2}\sigma\kappa\sqrt{-2x}\right)} dx \end{aligned}$$

lässt man  $n$  gegen  $\infty$  streben, so erhalten wir die charakteristische Funktion von  $\frac{\sigma}{2}l\left(\frac{\sigma}{2}\right)$  wobei  $l(t)$  die lokale Zeit einer Brownsche Exkursion der Dauer 1 ist (vgl. mit (1.60)).  $\square$

Mithilfe dieses Satzes kann man den Satz 3.2.13 beweisen:

*Beweis* (Satz 3.2.13): Aus (3.64) gilt  $E_k(z) = S(z)\alpha(z)^k \sim \tau \exp(-\sigma\kappa\sqrt{-2x})$  wobei  $z = \rho \left(1 + \frac{x}{n}\right)$ . Mithilfe der Cauchyschen Integralformel, indem man die gleiche Kurve  $\Gamma$  wie im Beweis des vorigen Satzes betrachtet, folgt:

$$\begin{aligned} \mathbb{E}(L_n(k)) &= \frac{[z^n]E_k(z)}{s_n} \\ &= \frac{1}{2\pi i s_n} \int_{\Gamma} \frac{E_k(z)}{z^{n+1}} dz \\ &\sim \frac{\sigma\sqrt{2\pi}\rho^n n^{3/2}}{2\pi\tau i} \int_{c-i\infty}^{c+i\infty} \frac{1}{n} \tau \exp(-\sigma\kappa\sqrt{-2x}) \rho^{-n} e^{-x} dx \\ &= \frac{\sigma\sqrt{2\pi}n^{1/2}}{2\pi i} \int_{c-i\infty}^{c+i\infty} \exp(-\sigma\kappa\sqrt{-2x}) e^{-x} dx. \end{aligned}$$

Aus Lemma 4.22 in [8] gilt

$$\frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \exp(-r\sqrt{-2x}) e^{-x} dx = \frac{r}{\sqrt{2\pi}} e^{-r^2/2}$$

für ein positives  $r$ . Einsetzen von  $r = \sigma\kappa > 0$  liefert

$$\mathbb{E}(L_n(k)) \sim \sigma^2 \kappa n^{1/2} e^{-\sigma^2 \kappa^2 / 2} = \sigma^2 k e^{-\sigma^2 k^2 / 2n}$$

mit  $k = \lfloor \kappa\sqrt{n} \rfloor$ , was zu zeigen war.  $\square$

Wir wollen noch die Aussage für  $0 \leq \kappa_1 < \kappa_2 < \dots < \kappa_d$  zeigen. Wir betrachten

nun den Fall  $d = 2$ . Zur Erinnerung ist die charakteristische Funktion der lokalen Zeit  $(l(\kappa_1), l(\kappa_2))$ ,  $\kappa_1 < \kappa_2$  einer 2-dimensionalen Brownschen Exkursion durch

$$\mathbb{E}e^{it_1 l(\kappa_1) + it_2 l(\kappa_2)} = 1 + \frac{\sqrt{2}}{i\sqrt{\pi}} \int_{c-i\infty}^{c+i\infty} \Psi_{\kappa_1}(x, it_1 + \Psi_{\kappa_2 - \kappa_1}(x, it_2)) e^{-x} dx$$

gegeben, wobei  $c < 0$  und

$$\begin{aligned} & \Psi_{\kappa_1}(x, it_1 + \Psi_{\kappa_2 - \kappa_1}(x, it_2)) \\ &= \frac{\sqrt{-2x} e^{-\kappa_1 \sqrt{-2x}} \left( it_1 + \frac{it_2 \sqrt{-2x} e^{-(\kappa_2 - \kappa_1) \sqrt{-2x}}}{\sqrt{-2x} e^{(\kappa_2 - \kappa_1) \sqrt{-2x}} - 2it_2 \sinh((\kappa_2 - \kappa_1) \sqrt{-2x})} \right)}{\sqrt{-2x} e^{-\kappa_1 \sqrt{-2x}} - \left( it_1 + \frac{it_2 \sqrt{-2x} e^{-(\kappa_2 - \kappa_1) \sqrt{-2x}}}{\sqrt{-2x} e^{(\kappa_2 - \kappa_1) \sqrt{-2x}} - 2it_2 \sinh((\kappa_2 - \kappa_1) \sqrt{-2x})} \right) \sinh(\kappa_1 \sqrt{-2x})}. \end{aligned}$$

Wir wollen diese Formel benützen, um die schwache Konvergenz zu zeigen:

**Satz 3.2.16** *Seien  $0 \leq \kappa_1 < \kappa_2$ . Dann gilt (mit gleicher Notation wie im vorigen Satz)*

$$(l_n(\kappa_1), l_n(\kappa_2)) \xrightarrow{d} \left( \frac{\sigma}{2} l \left( \frac{\sigma}{2} \kappa_1 \right), \frac{\sigma}{2} l \left( \frac{\sigma}{2} \kappa_2 \right) \right) \quad (3.80)$$

oder äquivalent

$$\lim_{n \rightarrow \infty} \mathbb{E}e^{it_1 l_n(\kappa_1) + it_2 l_n(\kappa_2)} = \mathbb{E}e^{it_1 \frac{\sigma}{2} l \left( \frac{\sigma}{2} \kappa_1 \right) + it_2 \frac{\sigma}{2} l \left( \frac{\sigma}{2} \kappa_2 \right)}. \quad (3.81)$$

*Beweis:* Analog zum eindimensionalen Fall, indem man die Funktion  $s_{\kappa_1, \kappa_2}(z, u_1, u_2) = S_{\kappa_1}(z, u_1 S_{\kappa_2 - \kappa_1}(z, u_2 S(z)))$  entlang der gleichen Kurve  $\Gamma = \Gamma' \cup \gamma$  wie im Beweis des Satzes 3.2.15 integriert, und  $n$  gegen  $\infty$  streben lässt. Für einen vollständigen Beweis siehe [8, Proposition 4.23].  $\square$

Für ein höheres  $d$  können wir das Resultat mithilfe dieses Beweises iterativ erhalten. Damit ist der Satz 3.2.7 bewiesen.

Wir wollen noch zeigen, dass der Prozess  $(l_n(t))_{t \geq 0}$  straff in  $C[0, \infty)$  ist. Das ist genau dann der Fall, wenn  $(L_n(t))_{t \geq 0}$  für  $0 \leq t \leq A\sqrt{n}$  ( $A > 0$  Konstante) straff ist (siehe [9, Teil 6]). Wir zeigen:

**Satz 3.2.17** *Es existiert eine positive Konstante  $C$ , sodass:*

$$\mathbb{E}(L_n(k) - L_n(k+h))^4 \leq Ch^2 n, \quad \forall n, k, h \in \mathbb{N}_0. \quad (3.82)$$

Die Idee ist, eine alternative Darstellung für  $\mathbb{E}(L_n(k) - L_n(k+h))^4$  zu bestimmen, und daher die obige asymptotische obere Schranke zu erhalten (siehe [8, Unterkapitel 4.2.6] und [9, Teil 6]). Zuerst ist die Anzahl der Bäume aus  $\mathcal{S}_n$  mit  $m$  Knoten der Höhe  $k$  und  $l$  Knoten der Höhe  $k+h$  gegeben durch

$$s_{nml, kh} = [z^n u^m v^l] S_k(z, u S_h(z, v S(z))).$$

Die zugehörige Wahrscheinlichkeit ist dann

$$\mathbb{P}(L_n(k) = m, L_n(k+h) = l) = \frac{s_{nml, kh}}{s_n}.$$

Daher haben wir

$$\mathbb{P}(L_n(k) - L_n(k+h) = r) = \frac{[z^n u^r] S_k \left( z, u S_h \left( z, \frac{S(z)}{u} \right) \right)}{s_n}$$

und

$$\mathbb{E}(L_n(k) - L_n(k+h))^4 = \frac{[z^n] H_{kh}(z)}{s_n}$$

wobei

$$H_{kh}(z) = \left( \frac{\partial}{\partial u} + 7 \frac{\partial^2}{\partial u^2} + 6 \frac{\partial^3}{\partial u^3} + \frac{\partial^4}{\partial u^4} \right) S_k \left( z, u S_h \left( z, \frac{S(z)}{u} \right) \right) \Big|_{u=1}, \quad k, h \geq 0.$$

Einen vollständigen Beweis von Satz 3.2.17 kann man in [8, Unterkapitel 4.2.6] oder [9, Teil 6] finden. In diesem werden eine asymptotische Darstellung für  $H_{k,h}(z)$  auf einem bestimmten  $\Delta$ -Gebiet wie in Bemerkung 1.2.26 und eine asymptotische obere Schranke für die Koeffizienten  $[z^n] H_{k,h}(z)$  und daher für  $\mathbb{E}(L_n(k) - L_n(k+h))^4$  berechnet.

$(l_n(t))_{t \geq 0}$  ist straff in  $C[0, \infty)$ . Dann ist der Prozess für jedes  $T > 0$  straff in  $C[0, T]$ . Dann können wir den Satz 1.3.32 anwenden (siehe [8, Seite 134]), womit der Satz 3.2.6 gezeigt ist. Als Folgerung haben wir (vgl. mit [8, Theorem 4.11]):

**Satz 3.2.18** *Die Breite  $W_n$  eines zufälligen einfach erzeugten Baumes erfüllt folgende schwache Konvergenz:*

$$\frac{1}{\sqrt{n}} W_n \xrightarrow{d} \frac{\sigma}{2} \sup_{0 \leq t \leq 1} l(t) \quad (3.83)$$

wobei  $l(t)$  und  $\sigma$  wie im Satz 3.2.6 sind.

*Beweis:* Aus (3.44) und (3.45) haben wir

$$\frac{1}{\sqrt{n}} W_n = \sup_{t \geq 0} l_n(t).$$

Das Supremum von stochastischen Prozessen bildet ein stetiges Funktional. Damit können wir den Satz 1.3.33 auf (3.46) anwenden, und wir bekommen (3.83).  $\square$

## Pólya-Bäume

Für Pólya-Bäume erfüllt unser Prozess  $(l_n(t))_{t \geq 0}$  folgenden Grenzwertsatz (siehe [11, Theorem 3]):

**Satz 3.2.19** *Ist  $(l_n(t))_{t \geq 0}$  das durch  $\sqrt{n}$  skalierte Profil eines zufälligen Pólya-Baum der Größe  $n$ , so gilt*

$$((l_n(t))_{t \geq 0}) \xrightarrow{d} \left( \frac{d_1}{2\sqrt{2}} l \left( \frac{d_1}{2\sqrt{2}} t \right) \right)_{t \geq 0}, \quad n \rightarrow \infty \quad (3.84)$$

wobei  $d_1$  die Konstante der Gleichung (2.25) sind.  $l(t)$  bezeichne wieder die lokale Zeit einer Brownsche Exkursion der Dauer 1.

Der erste Schritt, um den Satz zu beweisen, ist folgende Aussage zu zeigen (vgl. von [11, Theorem 4]):

**Satz 3.2.20** *Seien  $d_1, l_n(t)$  und  $l(t)$  wie im vorigen Satz. Dann gilt*

$$((l_n(t_1), \dots, l_n(t_d)) \xrightarrow{d} \frac{d_1}{2\sqrt{2}} \left( l \left( \frac{d_1}{2\sqrt{2}} t_1 \right), \dots, l \left( \frac{d_1}{2\sqrt{2}} t_d \right) \right), \quad n \rightarrow \infty \quad (3.85)$$

für beliebige aber feste nichtnegative  $t_1, \dots, t_d$ .

Zuerst definieren wir eine bivariate erzeugende Funktion  $p_k(z, u)$ , die eine ähnliche Bedeutung wie  $s_k(z, u)$  hat (vgl. mit [8, Unterkapitel 4.3.1], [11, Teil 3] oder [25, Abschnitt 2]).

**Definition 3.2.21** *Bezeichne mit  $p_{n,m}^{(k)}$  die Anzahl der Pólya-Bäume der Größe  $n$ , die auf der Höhe  $k$  genau  $m$  Knoten besitzen. Die entsprechende Wahrscheinlichkeit ist gegeben durch:*

$$\mathbb{P}(L_n(k) = m) = \frac{p_{n,m}^{(k)}}{p_n}. \quad (3.86)$$

Daher haben wir,

$$p_k(z, u) = \sum_{n,m \geq 0} p_{n,m}^{(k)} z^n u^m = \sum_{n \geq 0} p_n \mathbb{E}(u^{L_n(k)}) z^n. \quad (3.87)$$

Per Konstruktion haben wir

$$p_0(z, u) = uP(z), \quad (3.88)$$

$$p_{k+1}(z, u) = z \exp \left( \sum_{i \geq 1} \frac{p_k(z^i, u^i)}{i} \right), \quad k \geq 0. \quad (3.89)$$

Wie im vorigen Abschnitt kann man diese Rekursion allgemeiner machen (vgl. mit [8, Lemma 4.43]):

**Definition 3.2.22** *Seien  $d \in \mathbb{N}$ ,  $k_i$  nichtnegative Ganzzahlen mit  $0 \leq k_1 < k_2 < \dots < k_d$  und  $m_i \in \mathbb{N}_0$ ,  $i = 1, \dots, d$ . Wir definieren  $p_{n;m_1, \dots, m_d}^{(k_1, \dots, k_d)}$  als die Anzahl der Pólya-Bäume der Größe  $n$  mit  $m_i$  Knoten der Höhe  $k_i$  für jedes  $1 \leq i \leq d$ . Die zugehörige Wahrscheinlichkeit lautet dann*

$$\mathbb{P}(L_n(k_1) = m_1, L_n(k_2) = m_2, \dots, L_n(k_d) = m_d) = \frac{p_{n;m_1, \dots, m_d}^{(k_1, \dots, k_d)}}{p_n}. \quad (3.90)$$

Wir definieren

$$p_{k_1, \dots, k_d}(z, u_1, \dots, u_d) = \sum_{n, m_1, \dots, m_d \geq 0} p_{n;m_1, \dots, m_d}^{(k_1, \dots, k_d)} \mathbf{u}^m = \sum_{n \geq 0} p_n \mathbb{E}(u_1^{L_n(k_1)} \dots u_d^{L_n(k_d)}) z^n, \quad (3.91)$$

und diese Funktion erfüllt

$$p_{0, k_2, \dots, k_d}(z, u_1, \dots, u_d) = u_1 p_{k_2, \dots, k_d}(z, u_2, \dots, u_d), \quad (3.92)$$

$$p_{k_1+1, \dots, k_d+1}(z, u_1, \dots, u_d) = z \exp \left( \sum_{i \geq 1} \frac{p_{k_1, \dots, k_d}(z^i, u_1^i, \dots, u_d^i)}{i} \right). \quad (3.93)$$

Wir betrachten den Fall  $d = 1$ . Sei  $E_k(z, u)$  und  $E_k^{(2)}(z, u)$  die 1. bzw. 2. partielle Ableitung von  $p_k(z, u)$  nach  $u$  und daher  $E_k(z) = E_k(z, 1)$  und  $E_k^{(2)}(z) = E_k^{(2)}(z, 1)$  wie im vorigen Abschnitt. Mithilfe des rekursiven Ausdrucks von  $p_k(z, u)$  (3.88)-(3.89) bekommen wir (siehe [25, Theorem 6.1]):

**Satz 3.2.23** Die Reihe  $E_k(z, u)$  besitzt folgende rekursive Darstellung:

$$E_0(z, u) = P(z), \quad (3.94)$$

$$E_{k+1}(z, u) = p_{k+1}(z, u) \sum_{i \geq 1} E_k(z^i, u^i) u^{i-1}, \quad k \geq 0. \quad (3.95)$$

*Beweis:* Die Aussage für  $k = 0$  ist trivial. Betrachte  $k \geq 0$ . Wegen (3.89) gilt:

$$\begin{aligned} E_{k+1}(z) &= \frac{\partial}{\partial u} p_{k+1}(z, u) \\ &= \frac{\partial}{\partial u} \left( z \exp \left( \sum_{i \geq 1} \frac{p_k(z^i, u^i)}{i} \right) \right) \\ &= z \exp \left( \sum_{i \geq 1} \frac{p_k(z^i, u^i)}{i} \right) \sum_{i \geq 1} \frac{\partial}{\partial u} p_k(z^i, u^i) \\ &= p_{k+1}(z, u) \sum_{i \geq 1} E_k(z^i, u^i) u^{i-1}. \end{aligned}$$

□

**Korollar 3.2.24** Die Funktion  $E_k(z) = \sum_{n \geq 1} n \mathbb{E}(L_n(k)) p_n z^n$  erfüllt folgende Rekursion:

$$E_0(z) = P(z), \quad (3.96)$$

$$E_{k+1}(z) = P(z) \sum_{i \geq 1} E_k(z^i), \quad k \geq 0. \quad (3.97)$$

*Beweis:* Voriger Satz mit  $u = 1$ . □

Für  $E_k^{(2)}(z)$  haben wir:

**Korollar 3.2.25** Die Funktion  $E_k^{(2)}(z) = \sum_{n \geq 1} n \mathbb{E}(X_{n,k}(X_{n,k} - 1)) p_n z^n$  hat folgende rekursive Darstellung:

$$E_0^{(2)}(z) = 0, \quad (3.98)$$

$$E_{k+1}^{(2)}(z) = P(z) \left( \left( \sum_{i \geq 1} E_k(z^i) \right)^2 + \sum_{i \geq 1} i E_k(z^i) - \sum_{i \geq 1} E_k(z^i) + \sum_{i \geq 1} E_k^{(2)}(z^i) \right). \quad (3.99)$$

*Beweis:* Ähnlich wie im vorigen Satz. □

Nun können wir einen Ausdruck für  $\mathbb{E}(L_n(k))$  bestimmen.

**Satz 3.2.26** Für die mittlere Anzahl der Knoten der Höhe  $k$ ,  $\mathbb{E}(L_n(k))$  in einem zufälligen Pólya-Baum der Größe  $n$  gilt:

$$\lim_{n \rightarrow \infty} \mathbb{E}(L_n(1)) = 2 + \sum_{i \geq 2} P(\rho^i), \quad (3.100)$$

$$\lim_{n \rightarrow \infty} \mathbb{E}(L_n(k+1)) = \mathbb{E}(L_n(k)) + \sum_{i \geq 1} E_k(\rho^i), \quad k \geq 1. \quad (3.101)$$

Insbesondere gilt für die Varianz von  $L_n(1)$

$$\lim_{n \rightarrow \infty} \mathbb{V}(L_n(1)) = \sum_{i \geq 1} iP(\rho^i). \quad (3.102)$$

*Beweis:* Siehe [25, Corollary 6.1]. □

Wir werden später eine explizite Darstellung für  $\mathbb{E}(L_n(k))$  bestimmen.

Nun kümmern wir uns um die Grenzverteilung von  $L_n(k)$ . Wie im vorigen Abschnitt ist es sinnvoll, einen asymptotischen Ausdruck für  $w_k(z, u) := p_k(z, u) - P(z)$  zu bekommen. Wir interessieren uns für das asymptotische Verhalten von  $E_k(z)$  und  $E_k^{(2)}(z)$ . Erst brauchen wir weitere Analytizitätsbereiche (vgl. mit [11, Teil 4.1]):

$$\Theta = \{z \in \mathbb{C} : |z - \rho| < \rho + \eta, |\arg(z - \rho)| \neq 0\}, \quad (3.103)$$

$$\Xi_k = \{u \in \mathbb{C} : |u| \leq 1, k|u - 1| \leq \tilde{\eta}\}. \quad (3.104)$$

mit  $\eta, \tilde{\eta} > 0$  und  $0 < \theta < \frac{\pi}{2}$ . Außerdem sei erinnert, dass  $P(z) \sim 1 - d_1 \left(1 - \frac{z}{\rho}\right)^{1/2}$ . Damit haben wir:

**Lemma 3.2.27** Sei  $z \in \Theta$ . Dann erfüllt  $E_k(z)$ :

$$E_k(z) = C_k(z)P(z)^{k+1} \quad (3.105)$$

wobei  $C_k(z)$  eine konvergente Folge analytischer Funktionen bezeichnet. Ist  $C(z)$  die entsprechende Grenzfunktion, so gilt  $C(\rho) = \frac{1}{2}d_1^2$  mit  $d_1$  und  $\rho$  wie im Satz 2.2.5, und es gibt ein  $L \in (0, 1)$  mit

$$C_k(z) = C(z) + \mathcal{O}(L^k). \quad (3.106)$$

Außerdem kann man Konstanten  $\epsilon, \theta$  und  $\tilde{\eta}$  wie in (3.73) und (3.104) finden, sodass

$$|E_k(z, u)| = \mathcal{O}(|P(z)|^{k+1}) \quad (3.107)$$

gleichmäßig für  $z \in \Delta_\epsilon$  und  $u \in \Xi_k$ .

*Beweis:* Siehe [8, Lemma 4.46] oder [11, Lemma 3].

Wir erhalten einen Ausdruck von  $w_k(z, u)$ :

$$w_k(z, u) \sim (u - 1)E_k(z) \sim (u - 1)C(z)P(z)^k, \quad u \rightarrow 1.$$

Mithilfe folgendem Lemma können wir diese Darstellung genauer machen:



**Lemma 3.2.28** Seien  $|z| \leq \rho - \eta$  für  $\eta > 0$  und  $|u| \leq 1$ . Dann gilt gleichmäßig

$$E_k^{(2)}(z, u) = \mathcal{O}(P(|z|)^{k+1}) \quad (3.108)$$

Weiters kann man Konstanten  $\epsilon, \theta, \tilde{\eta}$  wie in (3.73) und (3.104) finden, sodass

$$E_k^{(2)}(z, u) = \mathcal{O}(kP(|z|)^{k+1}) \quad (3.109)$$

gleichmäßig für  $u \in \Xi_k$  und  $z \in \Delta_\epsilon$ .

*Beweis:* Siehe [8, Lemma 4.47] oder [11, Lemma 4].

Wir setzen  $w_k(z, u) = p_k(z, u) - P(z)$  und

$$\Sigma_k(z, u) = \sum_{i \geq 2} \frac{w_k(z^i, u^i)}{i}. \quad (3.110)$$

Wir haben dann:

**Lemma 3.2.29** Seien  $\epsilon, \theta, \tilde{\eta}$  und  $C_k = \frac{E_k(z)}{P(z)^{k+1}}$  wie im Lemma 3.2.28. Dann folgt

$$w_k(z, u) = C_k(z)(u-1)P(z)^{k+1}(1 + \mathcal{O}(k|u-1|)) \quad (3.111)$$

gleichmäßig für  $u \in \Xi_k$  und  $z \in \Delta_\epsilon$ .

Weiters gilt

$$\Sigma_k(z, u) = \tilde{C}(z)(u-1)P(z^2)^{k+1} + \mathcal{O}(u-1)^2 P(|z|^2)^k \quad (3.112)$$

wobei  $|z| \leq \rho + \eta$ ,  $|u| \leq 1$  und  $\tilde{C}_k(z) = \sum_{i \geq 2} C_k(z^i) \left( \frac{P(z^i)}{P(z^2)} \right)^{k+1}$  eine Folge analytischer Funktionen, die gegen eine weitere Funktion  $\tilde{C}(z)$  gleichmäßig konvergiert mit

$$\tilde{C}_k(z) = \tilde{C}(z) + \mathcal{O}(L^k), \quad 0 < L < 1. \quad (3.113)$$

*Beweis:* Siehe [8, Lemma 4.48] oder [11, Lemma 5]. □

Nun können wir eine asymptotische Darstellung für  $w_k(z, u)$  bestimmen (vgl. mit [8, Lemma 4.49] und [8, Proposition 4.50] oder [11, Lemme 6]):

**Satz 3.2.30** Seien  $z = \rho \left(1 + \frac{x}{n}\right)$  mit  $|\arg(x)| > 0$  und  $x = \mathcal{O}(\log^2(n))$  für  $n \rightarrow \infty$ ,  $u = e^{it/\sqrt{n}}$  und  $k = \lfloor \kappa\sqrt{n} \rfloor$  für fixe  $\kappa$  und  $t$ . Dann haben wir

$$w_k(z, u) \sim \frac{d_1^2}{2\sqrt{n}} \frac{it\sqrt{-x} \exp\left(-\frac{1}{2}\kappa d_1 \sqrt{-x}\right)}{\sqrt{-x} \exp\left(\frac{1}{2}\kappa d_1 \sqrt{-x}\right) - \frac{itd_1}{2} \sinh\left(\frac{1}{2}\kappa d_1 \sqrt{-x}\right)}. \quad (3.114)$$

gleichmäßig für  $k = \mathcal{O}(\sqrt{n})$ . Dabei sind  $d_1$  und  $\rho$  wie im Satz 2.2.5.

*Beweis:* Wegen dem vorigen Lemma gilt  $\Sigma_k(z, u) = \mathcal{O}(w_k(z, u))$ . Aus der rekursiven Darstellung (3.89) von  $p_k(z, u)$  und  $e^x = 1 + x + \frac{x^2}{2} + \mathcal{O}(x^3)$  folgt

$$\begin{aligned} w_{k+1}(z, u) &= P(z)(\exp(w_k(z, u) + \Sigma_k(z, u)) - 1) \\ &= P(z)\left(w_k(z, u) + \frac{w_k(z, u)^2}{2} + \Sigma_k(z, u) + \mathcal{O}(w_k(z, u)^3) + \mathcal{O}(\Sigma_k(z, u)^2)\right) \\ &= P(z)w_k(z, u) \left(1 + \frac{w_k(z, u)}{2} + \frac{\Sigma_k(z, u)}{w_k(z, u)} + \mathcal{O}(w_k(z, u)^2) + \mathcal{O}\left(\frac{\Sigma_k(z, u)^2}{w_k(z, u)}\right)\right) \\ &= P(z)w_k(z, u) \left(1 + \frac{w_k(z, u)}{2} + \mathcal{O}(w_k(z, u)^2) + \mathcal{O}(\Sigma_k(z, u))\right) \left(1 + \frac{\Sigma_k(z, u)}{w_k(z, u)}\right). \end{aligned}$$

Dies ist äquivalent zu (vgl. z.B. [11, Lemma 6])

$$\begin{aligned} \frac{P(z)}{w_{k+1}(z, u)} + \frac{P(z)\Sigma_k(z, u)}{w_k(z, u)w_{k+1}(z, u)} &= \frac{1}{w_k(z, u)} \left(1 - \frac{w_k(z, u)}{2} + \mathcal{O}(w_k(z, u)^2) + \mathcal{O}(\Sigma_k(z, u))\right) \\ &= \frac{1}{w_k(z, u)} - \frac{1}{2} + \mathcal{O}(w_k(z, u)) + \mathcal{O}\left(\frac{\Sigma_k(z, u)}{w_k(z, u)}\right). \end{aligned}$$

und multiplizieren mit  $P(z)^k$  liefert

$$\begin{aligned} \frac{P(z)^{k+1}}{w_{k+1}(z, u)} &= \frac{P(z)^k}{w_k(z, u)} - \frac{P(z)^{k+1}\Sigma_k(z, u)}{w_k(z, u)w_{k+1}(z, u)} - \frac{P(z)^k}{2} + \mathcal{O}(P(z)^k w_k(z, u)) + \mathcal{O}\left(\frac{P(z)^k \Sigma_k(z, u)}{w_k(z, u)}\right) \\ &= \frac{P(z)^0}{w_0(z, u)} - \sum_{i=0}^k \frac{P(z)^{i+1}\Sigma_i(z, u)}{w_i(z, u)w_{i+1}(z, u)} - \frac{1}{2} \sum_{i=0}^k P(z)^i \\ &\quad + \mathcal{O}\left(\sum_{i=0}^k P(z)^i w_i(z, u)\right) + \mathcal{O}\left(\sum_{i=0}^k \frac{P(z)^i \Sigma_i(z, u)}{w_i(z, u)}\right) \\ &= \frac{1}{w_0(z, u)} - \sum_{i=0}^k \frac{P(z)^{i+1}\Sigma_i(z, u)}{w_i(z, u)w_{i+1}(z, u)} - \frac{1}{2} \frac{1 - P(z)^{k+1}}{1 - P(z)} \\ &\quad + \mathcal{O}\left(\frac{1 - L^{k+1}}{1 - L}\right) + \mathcal{O}\left(w_0(z, u) \frac{1 - P(z)^{2(k+1)}}{1 - P(z)^2}\right). \end{aligned}$$

Nun können wir die 1. Summe wie folgt umschreiben. Wegen  $w_k(z, u) = C_k(z)(u - 1)P(z)^{k+1}(1 + \mathcal{O}(k|u - 1|))$  für  $u \in \Xi_k$ ,  $z \in \Delta_\epsilon$  und  $\Sigma_k(z, u) = \tilde{C}_k(z)(u - 1)P(z^2)^{k+1} + \mathcal{O}(|u - 1|^2 P(|z|^2)^k)$  für  $|z| < \rho + \eta$ ,  $|u| \leq 1$  (siehe Lemma 3.2.27) folgt

$$\begin{aligned} \sum_{i=0}^{k-1} \frac{P(z)^{i+1}\Sigma_i(z, u)}{w_i(z, u)w_{i+1}(z, u)} &= \sum_{i=0}^{k-1} \frac{\tilde{C}_i(z)P(z^2)^{i+1} + \mathcal{O}(|u - 1|P(|z|^2)^i)}{C_i(z)C_{i+1}(z)P(z)^{i+1}(1 + \mathcal{O}(i|u - 1|))} \\ &= \sum_{i=0}^{k-1} \frac{\tilde{C}_i(z)}{C_i(z)C_{i+1}(z)} \left(\frac{P(z^2)}{P(z)}\right)^{k+1} + \mathcal{O}(|u - 1|). \end{aligned}$$

Für  $\tilde{C}_k(z)$  gilt sogar (mithilfe der Formel (27) im Beweis von Lemma 3 in [11])

$$\begin{aligned} \tilde{C}_k(z) &= \sum_{j \geq 2} C_k(z^j) \left(\frac{P(z^j)}{P(z^2)}\right)^{k+1} \\ &= (C_{k+1}(z) - C_k(z)) \left(\frac{P(z)}{P(z^2)}\right)^{k+1}. \end{aligned}$$

Daher haben wir

$$\begin{aligned} \sum_{i=0}^{k-1} \frac{P(z)^{i+1} \Sigma_i(z, u)}{w_i(z, u) w_{i+1}(z, u)} &= \sum_{i=0}^{k-1} \frac{C_{i+1}(z) - C_i(z)}{C_i(z) C_{i+1}(z)} + \mathcal{O}(|u-1|) \\ &= \frac{1}{C_0(z)} - \frac{1}{C_k(z)} + \mathcal{O}(|u-1|) \\ &= 1 - \frac{1}{C_k(z)} + \mathcal{O}(|u-1|). \end{aligned}$$

Insgesamt gilt

$$w_k(z, u) = \frac{C_k(z) w_0(z, u) P(z)^k}{1 - \frac{1}{2} C_k(z) w_0(z, u) \frac{1-P(z)^k}{1-P(z)} + \mathcal{O}(|u-1|)} \quad (3.115)$$

für  $z \in \Delta_\epsilon$  und  $u \in \Xi_k$  (vgl. mit [11, Lemma 6] oder [8, Lemma 4.49]).

Aus (3.106) wissen wir, dass die Folge  $C_k(z)$  gegen  $C(z)$  gleichmäßig konvergiert. Daher ist sie auch gleichmäßig stetig mit  $C_k(z) \sim C(\rho) = \frac{1}{2} d_1^2$ . Wegen  $P(z) = 1 - d_1(1 - \frac{z}{\rho})^{1/2} + \mathcal{O}(|1 - \frac{z}{\rho}|)$  folgt

$$P(z)^k = \exp(-\kappa d_1 \sqrt{-x}) \left( 1 + \mathcal{O}\left(\frac{\kappa}{\sqrt{n}}\right) \right), \quad (3.116)$$

mit  $\kappa = \frac{k}{\sqrt{n}}$  und  $z = \rho \left(1 + \frac{x}{n}\right)$ . Mit  $u = e^{it/\sqrt{n}}$  erhalten wir

$$\begin{aligned} w_k(z, u) &= \frac{C_k(z)(u-1)(1-P(z))P(z)^{k+1}}{1 - P(z) - \frac{1}{2} C_k(z)(u-1)(1-P(z)^k)P(z) + \mathcal{O}(|u-1| \cdot |1-P(z)|)} \\ &\sim \frac{C(\rho)(\exp(it/\sqrt{n}) - 1) d_1 \sqrt{-x} \exp(-\kappa d_1 \sqrt{-x})}{d_1 \sqrt{-x} - \frac{1}{2} C(\rho)(\exp(it/\sqrt{n}) - 1)(1 - \exp(-\kappa d_1 \sqrt{-x}))} \\ &\sim \frac{d_1^2}{2\sqrt{n}} \frac{it d_1 \sqrt{-x} \exp(-\frac{1}{2} \kappa d_1 \sqrt{-x})}{d_1 \sqrt{-x} \exp(\frac{1}{2} \kappa d_1 \sqrt{-x}) - \frac{it d_1^2}{4\sqrt{n}} (\exp(\frac{1}{2} \kappa d_1 \sqrt{-x}) - \exp(-\frac{1}{2} \kappa d_1 \sqrt{-x}))} \\ &= \frac{d_1^2}{2\sqrt{n}} \frac{it \sqrt{-x} \exp(-\frac{1}{2} \kappa d_1 \sqrt{-x})}{\sqrt{-x} \exp(\frac{1}{2} \kappa d_1 \sqrt{-x}) - \frac{it d_1}{2} \sinh(\frac{1}{2} \kappa d_1 \sqrt{-x})}. \end{aligned}$$

□

Nun können wir die Grenzverteilung ermitteln (vgl. mit [8, Proposition 4.50]):

**Satz 3.2.31** Seien  $\kappa > 0$  und  $t$  mit  $|\kappa t| \leq c$ . Dann gilt

$$\lim_{n \rightarrow \infty} \mathbb{E} e^{it l_n(\kappa)} = \mathbb{E} e^{it \frac{d_1}{2\sqrt{2}} l\left(\frac{d_1}{2\sqrt{2}} \kappa\right)} \quad (3.117)$$

oder äquivalent

$$l_n(\kappa) \xrightarrow{d} \frac{d_1}{2\sqrt{2}} l\left(\frac{d_1}{2\sqrt{2}} \kappa\right). \quad (3.118)$$

$l(t)$  ist die lokale Zeit einer Brownsche Exkursion der Dauer 1.  $d_1$  ist wie im Satz 2.2.5.

*Beweis* : Wie im vorigen Abschnitt (Satz 3.2.15) wenden wir die Cauchysche Integralformel entlang des Weges  $\Gamma = \Gamma' \cup \gamma$  an, um die Koeffizienten von  $p_k(z, e^{it/\sqrt{n}})$  zu berechnen,

und daher einen Ausdruck für  $\mathbb{E}e^{itl_n(\kappa)}$  zu erhalten. Der Beitrag von  $\Gamma'$  ist exponentiell klein und konvergiert gegen 0 für  $n \rightarrow \infty$  (vgl. Beweis von Satz 3.2.15). Für  $z \in \gamma$  haben wir mit der Substitution  $z = \rho \left(1 + \frac{x}{n}\right)$ ,  $dz = \frac{\rho}{n} dx$  und  $z^{-n} \sim \rho^{-n} e^{-x}$

$$\begin{aligned} & \frac{1}{2\pi i p_n} \int_{\gamma} p_k(z, e^{it/\sqrt{n}}) \frac{dz}{z^{n+1}} \\ &= 1 + \frac{1}{2\pi i p_n} \int_{\gamma} w_k(z, e^{it/\sqrt{n}}) \frac{dz}{z^{n+1}} \\ &= 1 + \frac{2\sqrt{\pi} \rho^n n^{3/2}}{2\pi d_1 i} \int_{c-i \log^2(n)}^{c+i \log^2(n)} \frac{d_1^2}{2n^{3/2}} \frac{it\sqrt{-x} \exp\left(-\frac{1}{2}\kappa d_1 \sqrt{-x}\right) \rho^{-n} e^{-x}}{\sqrt{-x} \exp\left(\frac{1}{2}\kappa d_1 \sqrt{-x}\right) - \frac{itd_1}{2} \sinh\left(\frac{1}{2}\kappa d_1 \sqrt{-x}\right)} dx \\ &= 1 + \frac{d_1}{2\sqrt{\pi}} \int_{c-i \log^2(n)}^{c+i \log^2(n)} \frac{t\sqrt{-x} \exp\left(-\frac{1}{2}\kappa d_1 \sqrt{-x}\right) e^{-x}}{\sqrt{-x} \exp\left(\frac{1}{2}\kappa d_1 \sqrt{-x}\right) - \frac{itd_1}{2} \sinh\left(\frac{1}{2}\kappa d_1 \sqrt{-x}\right)} dx \\ &= 1 + \frac{d_1}{2\sqrt{2}} \frac{\sqrt{2}}{\sqrt{\pi}} \int_{c-i \log^2(n)}^{c+i \log^2(n)} \frac{t\sqrt{-x} \exp\left(-\frac{1}{2}\kappa d_1 \sqrt{-x}\right) e^{-x}}{\sqrt{-2x} \exp\left(\frac{1}{2}\kappa d_1 \sqrt{-x}\right) - \frac{itd_1}{\sqrt{2}} \sinh\left(\frac{1}{2}\kappa d_1 \sqrt{-x}\right)} dx. \end{aligned}$$

Lässt man  $n \rightarrow \infty$ , so haben wir  $\mathbb{E}e^{it \frac{d_1}{2\sqrt{2}} l\left(\frac{d_1}{2\sqrt{2}} \kappa\right)}$ .  $\square$

Nun berechnen wir einen alternativen Ausdruck für den Mittelwert  $\mathbb{E}(L_n(k))$ .

**Satz 3.2.32** Für  $k = \lfloor \kappa \sqrt{n} \rfloor$ ,  $\kappa > 0$  erfüllt die mittlere Anzahl der Knoten der Höhe  $k$  in einem zufälligen Pólya-Baum der Größe  $n$  folgende asymptotische Darstellung:

$$\mathbb{E}(L_n(k)) \sim \frac{d_1^2 k}{2} e^{-d_1^2 k^2 / 4n}. \quad (3.119)$$

*Beweis:* Im Beweis des vorigen Satzes haben wir gesehen, dass  $C_k(z) \sim C(\rho) = \frac{d_1^2}{2}$ . Zusammen mit  $P(\rho) = 1$ , (3.105) und (3.116) erhalten wir

$$E_k(z) = C_k(z)P(z)^{k+1} \sim \frac{d_1^2}{2} \exp(-\kappa d_1 \sqrt{-x}).$$

wobei  $z = \rho \left(1 + \frac{x}{n}\right)$ . Die Bestimmung der Koeffizienten ist genauso wie im Beweis von Satz 3.2.13, indem man  $E_k(z)$  entlang der gleichen Kurve  $\Gamma$  wie im Beweis von Satz 3.2.15 integriert und Lemma 4.22 aus [8] anwendet.  $\square$

Wie bei einfach erzeugten Bäumen können wir eine analoge Aussage für höhere Dimensionen beweisen. Wir zeigen den Fall  $d = 2$ , indem wir zuerst das asymptotische Verhalten von  $w_{k_1, k_2}(z, u_1, u_2) = p_{k_1, k_2}(z, u_1, u_2) - P(z)$  lokal analysieren (vgl. mit [11, Proposition 1]):

**Proposition 3.2.33** Seien  $\kappa_1, \kappa_2$  mit  $0 \leq \kappa_1 \leq \kappa_2$  und  $t_1, t_2$  mit  $|\kappa_j t_j| \leq C$ ,  $j = 1, 2$  ( $C > 0$  Konstante). Weiters setze  $z = \rho \left(1 + \frac{x}{n}\right)$  mit  $\operatorname{Re}(x) = c < 0$ ,  $u_j = e^{it_j/\sqrt{n}}$  und  $k_j = \lfloor \kappa_j \sqrt{n} \rfloor$  für  $j = 1, 2$ . Dann haben wir

$$w_{k_1, k_2}(z, u_1, u_2) \sim \frac{d_1 \sqrt{2}}{\sqrt{n}} \Psi_{\frac{\kappa_1 d_1}{2\sqrt{2}}} \left(x, \frac{it_1 d_1}{2\sqrt{2}}\right) + \Psi_{\frac{(\kappa_2 - \kappa_1) d_1}{2\sqrt{2}}} \left(z, \frac{it_2 d_1}{2\sqrt{2}}\right), \quad n \rightarrow \infty \quad (3.120)$$

gleichmäßig für  $k_i|u_i - 1| \leq C$ ,  $i = 1, 2$  und  $x = \mathcal{O}(\log^2(n))$ .  $\Psi_k(x, t)$  ist die durch (1.63) gegebene Funktion.

*Beweis:* Siehe [8, Proposition 4.51] oder [11, Proposition 1]. □

**Satz 3.2.34** Seien  $\kappa_1 > \kappa_2 > 0$  und  $t_1, t_2$  mit  $|\kappa_i t_i| \leq C$ . Dann gilt

$$(l_n(\kappa_1), l_n(\kappa_2)) \xrightarrow{d} \left( \frac{d_1}{2\sqrt{2}} l \left( \frac{d_1}{2\sqrt{2}} \kappa_1 \right), \frac{d_1}{2\sqrt{2}} l \left( \frac{d_1}{2\sqrt{2}} \kappa_2 \right) \right). \quad (3.121)$$

*Beweis:* Analog wie in Satz 3.2.31, indem man die Funktion  $\frac{p_{k_1, k_2}(z, u_1, u_2)}{p_n}$  entlang  $\Gamma$  integriert und die Darstellung (3.120) von  $w_{k_1, k_2}(z, u_1, u_2) = p_{k_1, k_2}(z, u_1, u_2) - P(z)$  anwendet. Wenn  $n$  gegen  $\infty$  strebt, dann bekommen wir das gewünschte Ergebnis. □

Wir wollen noch zeigen, dass der Prozess  $(l_n(t))_{t \geq 0}$  in  $C[0, \infty)$  straff ist. Unserer Ziel ist, folgende Aussage zu beweisen (vgl. mit [8, Theorem 4.52] und [11, Theorem 5]):

**Satz 3.2.35** Es gilt

$$\mathbb{E}(L_n(k) - L_n(k+h)^4) \leq Ch^2 n \quad (3.122)$$

für jedes  $n, k, h \in \mathbb{N}_0$ .

*Beweis:* Sei zuerst bemerkt, dass das Moment aus (3.122) folgende Darstellung hat:

$$\mathbb{E}(L_n(k) - L_n(k+h)^4) = \frac{1}{p_n} [z^n] \left[ \left( \frac{\partial}{\partial u} + 7 \frac{\partial^2}{\partial u^2} + 6 \frac{\partial^3}{\partial u^3} + \frac{\partial^4}{\partial u^4} \right) p_{k, k+h} \left( z, u, \frac{1}{u} \right) \right]_{u=1}$$

Wegen  $p_n \sim \frac{d_1 \rho^{-n}}{2\sqrt{\pi n^{3/2}}}$  genügt es zu zeigen, dass

$$[z^n] \left[ \left( \frac{\partial}{\partial u} + 7 \frac{\partial^2}{\partial u^2} + 6 \frac{\partial^3}{\partial u^3} + \frac{\partial^4}{\partial u^4} \right) p_{k, k+h} \left( z, u, \frac{1}{u} \right) \right]_{u=1} = \mathcal{O}(h^2 \rho^{-n} n^{-1/2})$$

Dafür betrachten wir folgende partielle Ableitungen:

$$E_{k, h}^{(j)}(z) := \frac{\partial^j}{\partial u^j} p_{k, k+h} \left( z, u, \frac{1}{u} \right) \Big|_{u=1}$$

Die Idee ist, eine asymptotische obere Schranke für diese Funktionen in bestimmten Analytizitätsbereich zu finden (siehe [8, Abschnitt 4.3.4] oder [11, Teil 6] für einen vollständigen Beweis). □

Daher ist unser Prozess  $(l_n(t))_{t \geq 0}$  straff in  $C[0, \infty)$ . Es sei erinnert, dass die Straffheit in  $C[0, \infty)$  äquivalent zu der Straffheit in  $C[0, T]$  für jedes  $T > 0$  ist (siehe [8, Seite 134]). Damit ist laut dem Satz 1.3.32 der Satz 3.2.19 endlich bewiesen. Wir erhalten folgende Konsequenz:

**Satz 3.2.36** Seien  $l(t)$  und  $d_1$  wie im Satz 3.2.19, und bezeichne  $W_n$  die Breite eines zufälligen Pólya-Baum der Größe  $n$ . Dann gilt

$$\frac{1}{\sqrt{n}} W_n \xrightarrow{d} \frac{d_1}{2\sqrt{2}} \sup_{0 \leq t \leq 1} l(t). \quad (3.123)$$

*Beweis:* Genauso wie der Beweis von Satz 3.2.18. □

Wir fassen unsere Ergebnisse zusammen:

	Einfach erzeugte Bäume	Pólya-Bäume
$\mathbb{E}(L_n(k))$	$\sim \sigma^2 k e^{-\sigma^2 k^2 / 2n}$	$\sim \frac{d_1^2 k}{2} e^{-d_1^2 k^2 / 4n}$
Grenzverteilung von $(l_n(t))_{t \geq 0}$	$\left(\frac{\sigma}{2} l\left(\frac{\sigma}{2} t\right)\right)_{t \geq 0}$	$\left(\frac{d_1}{2\sqrt{2}} l\left(\frac{d_1}{2\sqrt{2}} t\right)\right)_{t \geq 0}$
Grenzverteilung von $\frac{1}{\sqrt{n}} W_n$	$\frac{\sigma}{2} \sup_{0 \leq t \leq 1} l(t)$	$\frac{d_1}{2\sqrt{2}} \sup_{0 \leq t \leq 1} l(t)$

Dabei sind  $\sigma$  und  $d_1$  wie in (2.11) bzw. (2.25) und  $l(t)$  ist die lokale Zeit einer Brownschen Exkursion der Dauer 1.

## 3.2.2 Die Höhe

### Einfach erzeugte Bäume

Resultate für die Höhe  $H_n$  eines zufälligen einfach erzeugten Baumes der Größe  $n$  wurden schon von Flajolet in [14] gezeigt. Laut Drmota [8, Unterkapitel 4.2.7] können wir eine ähnliche Überlegung wie im vorigen Abschnitt anwenden, um die Momente  $\mathbb{E}(H_n^r)$  mit  $r \in \mathbb{N}$  zu bestimmen. Wir definieren folgende erzeugende Funktion:

**Definition 3.2.37** Die erzeugende Funktion  $s_k(z)$  für einfach erzeugte Bäume der Höhe  $< k$  ist definiert durch

$$s_k(z) = \sum_{n \geq 1} \left( \sum_{S \in \mathcal{S}_n, H(S) < k} w(S) \right) z^n. \quad (3.124)$$

Da  $L_S(k) = 0 \Leftrightarrow H(S) < k$  sieht man leicht, dass  $s_k(z) = s_k(z, 0)$  und daher (wegen (3.52)-(3.53))

$$s_0(z) = 0, \quad (3.125)$$

$$s_{k+1}(z) = z\Phi(s_k(z)). \quad (3.126)$$

Wir erhalten dann folgendes Ergebnis:

**Proposition 3.2.38** Wir setzen  $e_k(z) = s(z) - s_k(z)$ ,  $k \geq 0$ . In einem  $\Delta$ -Gebiet  $\Delta$  wie in Bemerkung 1.2.24 gilt sogar:

$$e_k(z) = \frac{\alpha(z)^k}{\frac{\beta(z)}{\alpha(z)} \frac{1-\alpha(z)^k}{1-\alpha(z)} + \mathcal{O}\left(\min\left\{\log(k), \log\left(\frac{1}{1-|\alpha(z)|}\right)\right\}\right)} \quad (3.127)$$

wobei  $\alpha(z) = z\Phi'(S(z))$  und  $\beta(z) = \frac{z}{2}\Phi''(S(z))$ .

Den Beweis von dieser Proposition kann man in [8, Proposition 4.28] finden.

Setze wieder  $z = \rho(1 + \frac{x}{n})$  mit  $\operatorname{Re}(x) = c < 0$  und  $\rho$  wie in Proposition 2.1.12. Wegen (3.75), (3.77) und  $\beta(\rho) = \frac{\sigma}{\tau}$  (vgl [8, Seite 147]) haben wir:

$$\begin{aligned} e_k(z) &\sim \frac{1}{\sqrt{n}} \frac{\exp(-\sigma\kappa\sqrt{-2x})}{\frac{\sigma}{\tau} \frac{1 - \exp(-\sigma\kappa\sqrt{-2x})}{\sigma\sqrt{-2x}}} \\ &= \frac{1}{\sqrt{n}} \frac{\tau\sqrt{-2x} \exp(-\sigma\kappa\sqrt{-2x})}{1 - \exp(-\sigma\kappa\sqrt{-2x})} \\ &= \frac{1}{\sqrt{n}} \frac{\tau\sqrt{-2x}}{\exp(\sigma\kappa\sqrt{-2x}) - 1}. \end{aligned}$$

Wir bekommen dann einen Ausdruck für die Verteilung von  $H_n$  (vgl. mit [8, Theorem 4.29]):

**Satz 3.2.39** Sei  $\kappa > 0$  und setze  $k = \lfloor \kappa\sqrt{n} \rfloor$ . Dann gilt

$$\mathbb{P}(H_n \geq k) = \frac{\sigma}{i\sqrt{2\pi}} \int_{c-i\infty}^{c+i\infty} \frac{\sqrt{-2x}e^{-x}}{\exp(\sigma\kappa\sqrt{-2x}) - 1} dx + o(1) \quad (3.128)$$

wobei  $c < 0$  eine Konstante ist.  $\sigma$  ist wie in (2.11).

*Beweis:* Man sieht leicht, dass

$$\begin{aligned} \mathbb{P}(H_n \geq k) &= 1 - \mathbb{P}(H_n < k) \\ &= 1 - \frac{[z^n]s_k(z)}{[z^n]S(z)} \\ &= \frac{[z^n](S(z) - s_k(z))}{[z^n]S(z)} \\ &= \frac{[z^n]e_k(z)}{s_n} \\ &= \frac{1}{2\pi i s_n} \int_{\Gamma} e_k(z) \frac{dz}{z^{n+1}}. \end{aligned}$$

Wir betrachten wieder die Contour  $\Gamma = \Gamma' \cup \gamma$  (siehe Beweis von Satz 3.2.15). Der Beitrag von  $\Gamma'$  ist wieder exponentiell klein und hat  $\mathcal{O}(\rho^{-n}e^{-C\log^2(n)})$ . Für  $z \in \gamma$  haben wir (mit  $s_n \sim \frac{\tau\rho^{-n}}{\sigma\sqrt{2\pi n^{3/2}}}$ ,  $z = \rho(1 + \frac{x}{n})$ ,  $z^{-n} \sim \rho^{-n}e^{-x}$  und  $dz = \frac{\rho}{n}dx$ )

$$\begin{aligned} \frac{1}{2\pi i s_n} \int_{\gamma} e_k(z) \frac{dz}{z^{n+1}} &\sim \frac{\sigma\sqrt{2\pi}\rho^n n^{3/2}}{2\tau\pi i} \int_{c-i\log^2(n)}^{c+i\log^2(n)} \frac{1}{n^{3/2}} \frac{\tau\sqrt{-2x}\rho^{-n}e^{-x}}{\exp(\sigma\kappa\sqrt{-2x}) - 1} dx \\ &= \frac{\sigma}{i\sqrt{2\pi}} \int_{c-i\log^2(n)}^{c+i\log^2(n)} \frac{\sqrt{-2x}e^{-x}}{\exp(\sigma\kappa\sqrt{-2x}) - 1} dx \\ &= \frac{\sigma}{i\sqrt{2\pi}} \int_{c-i\infty}^{c+i\infty} \frac{\sqrt{-2x}e^{-x}}{\exp(\sigma\kappa\sqrt{-2x}) - 1} dx + o(1). \end{aligned}$$

□

Nun können wir asymptotische Darstellungen für die Momente  $\mathbb{E}(H_n^r)$ ,  $r \geq 1$  bestimmen (vgl. mit [8, Theorem 4.29]):

**Satz 3.2.40** Die mittlere Höhe eines zufälligen einfach erzeugten Baumes der Größe  $n$  erfüllt

$$\mathbb{E}(H_n) \sim \frac{\sqrt{2\pi}}{\sigma} n^{1/2}, \quad (3.129)$$

Wobei  $\sigma$  wie in (2.11) ist. Für  $r > 1$  gilt

$$\mathbb{E}(H_n^r) \sim \left(\frac{2}{\sigma}\right)^r 2^{-r/2} r(r-1)\Gamma(r/2)\zeta(r)n^{r/2}. \quad (3.130)$$

*Beweis:* Wir definieren

$$H_r(z) = \sum_{n \geq 1} \mathbb{E}(H_n^r) s_n z^n. \quad (3.131)$$

Sei bemerkt, dass  $\mathbb{P}(H_n = k) = \frac{[z^n](s_{k+1}(z) - s_k(z))}{s_n}$ . Daraus folgt (mithilfe von (1.42)), dass

$$\begin{aligned} H_1(z) &= \sum_{n \geq 1} \sum_{k \geq 0} k \mathbb{P}(H_n = k) s_n z^n \\ &= \sum_{k \geq 1} k (s_{k+1}(z) - s_k(z)) \\ &= \sum_{k \geq 1} (S(z) - s_k(z)) \\ &= \sum_{k \geq 1} e_k(z) \end{aligned}$$

und allgemeiner

$$H_r(z) = \sum_{k \geq 1} (k^r - (k-1)^r) e_k(z).$$

Wegen (3.127) und  $k^r - (k-1)^r \sim rk^{r-1}$  haben wir

$$H_r(z) \sim r \frac{\alpha}{\beta} (1-\alpha) \sum_{k \geq 1} k^{r-1} \frac{\alpha^k}{1-\alpha^k}, \quad z \rightarrow \rho.$$

Die Summe erfüllt folgende Asymptotik (vgl. [8, Lemma 4.37])

$$\sum_{k \geq 1} k^{r-1} \frac{\alpha^k}{1-\alpha^k} \sim \begin{cases} \frac{1}{1-\alpha} \log\left(\frac{1}{1-\alpha}\right), & r-1 = 0 \Leftrightarrow r = 1, \\ \frac{(r-1)! \zeta(r)}{(1-\alpha)^r}, & r-1 > 0 \Leftrightarrow r > 1. \end{cases}$$

- 1. Fall:  $r = 1$

Für  $r = 1$  haben wir (mithilfe von (3.69))

$$H_1(z) \sim \frac{\alpha}{\beta} \log\left(\frac{1}{1-\alpha}\right) \sim \frac{1}{\beta} \log\left(\frac{1}{(1-z/\rho)^{1/2}}\right) = \frac{1}{2\beta} \log\left(\frac{1}{1-z/\rho}\right)$$



und daher (mit  $\beta \sim \frac{\sigma^2}{2\tau}$  (vgl. mit (3.70)) und  $s_n \sim \frac{\tau\rho^{-n}}{\sigma\sqrt{2\pi n^{3/2}}}$ )

$$\begin{aligned}\mathbb{E}(H_n) &= \frac{[z^n]H_1(z)}{s_n} \\ &\sim \frac{\tau}{\sigma^2} \frac{\sigma\sqrt{2\pi}\rho^n n^{3/2}}{\tau} [z^n] \log\left(\frac{1}{1-z/\rho}\right) \\ &= \frac{\sqrt{2\pi}\rho^n n^{3/2}}{\sigma} \rho^{-n} n^{-1} \\ &= \frac{\sqrt{2\pi}}{\sigma} n^{1/2}.\end{aligned}$$

• 2. Fall:  $r > 1$ :

Für  $r > 1$  gilt (wieder mithilfe von (3.69))

$$H_r(z) \sim \frac{r!\zeta(r)\alpha}{\beta(1-\alpha)^{r-1}} \sim \frac{r!\zeta(r)}{\beta(\sigma\sqrt{2})^{r-1}} \left(1 - \frac{z}{\rho}\right)^{-\frac{r-1}{2}}.$$

Wir bekommen (mit Transferlemma 1.2.16):

$$\begin{aligned}\mathbb{E}(H_n^r) &= \frac{[z^n]H_r(z)}{s_n} \\ &\sim \frac{2\tau}{\sigma^2} \frac{r!\zeta(r)}{(\sigma\sqrt{2})^{r-1}} \frac{\sigma\sqrt{2\pi}\rho^n n^{3/2}}{\tau} \frac{\rho^{-n} n^{(r-3)/2}}{\Gamma((r-1)/2)} \\ &= \frac{2r!\zeta(r)\sqrt{\pi}}{(\sigma\sqrt{2})^r \Gamma((r-1)/2)} n^{r/2}.\end{aligned}$$

$\Gamma(z+1) = z\Gamma(z)$  liefert uns (3.130). □

Die Höhe erfüllt folgenden Grenzwertsatz (vgl. mit [8, Theorem 4.8]):

**Satz 3.2.41** *Die durch  $\sqrt{n}$  skalierte Höhe eines zufälligen einfach erzeugten Baumes der Größe  $n$  konvergiert in Verteilung gegen den Maximum einer Brownschen Exkursion der Dauer 1, genauer:*

$$\frac{1}{\sqrt{n}} H_n \xrightarrow{d} \frac{2}{\sigma} \max_{0 \leq t \leq 1} e(t) \quad (3.132)$$

wobei  $\sigma$  die Varianz des entsprechenden Nachkommenverteilung eines Galton-Watson Verzweigungsprozesses ist.

*Beweis:* Aldous hat in [2, Theorem 2] gezeigt, dass für den Contour-Prozess  $(C_n(t))_{0 \leq t \leq 1}$  folgende schwache Konvergenz vorliegt:

$$(C_n(t))_{0 \leq t \leq 1} \xrightarrow{d} \left(\frac{2e(t)}{\sigma}\right)_{0 \leq t \leq 1}, \quad n \rightarrow \infty. \quad (3.133)$$

Weiters wissen wir schon, dass  $\frac{1}{\sqrt{n}} H_n = \max_{0 \leq t \leq 1} C_n(t)$ . Das Maximum von stochastischen Prozessen ist ein stetiges Funktional. Satz 1.3.33 liefert uns das gesuchte Ergebnis. □

## Pólya-Bäume

Die Idee, um die Momente  $\mathbb{E}(H_n^r)$  zu bestimmen ist ähnlich wie bei einfach erzeugten Bäumen (siehe [8, Unterkapitel 4.3.5] oder [11, Teil 7]). Wir definieren zuerst eine zu  $s_k(z)$  analoge erzeugende Funktion:

**Definition 3.2.42** Sei  $p_n^{(k)}$  die Anzahl der Pólya-Bäumen mit  $n$  Knoten der Höhe  $< k$ . Die entsprechende erzeugende Funktion ist dann

$$p_k(z) = p_k(z, 0) = \sum_{n \geq 1} p_n^{(k)} z^n. \quad (3.134)$$

(vgl. mit [8, Abschnitt 4.3.5] oder [11, Teil 7]). Wegen (3.89)-(3.90) sieht man leicht, dass

$$p_0(z) = 0, \quad (3.135)$$

$$p_{k+1}(z) = z \exp\left(\sum_{i \geq 1} \frac{p_k(z^i)}{i}\right), \quad k \geq 0. \quad (3.136)$$

Wir setzen  $e_k(z) = P(z) - p_k(z)$ . Der wichtigste Unterschied zwischen einfach erzeugenden Bäumen und Pólya-Bäumen ist das Analytizitätsgebiet von  $e_k(z)$  (siehe [8, Formel (4.118)] oder [11, Proposition 3]):

**Proposition 3.2.43** Für  $z \in \Delta_\epsilon$  (vgl. (3.73)) gilt

$$e_k(z) = \frac{P(z)^k}{\frac{1}{2} \frac{1-P(z)^k}{1-P(z)} + \mathcal{O}\left(\min\left\{\log(k), \log\left(\frac{1}{1-|P(z)|}\right)\right\}\right)}. \quad (3.137)$$

*Beweis:* Siehe [8, Proposition 3] □

Setze  $z = \rho \left(1 + \frac{x}{n}\right)$  und  $k = \lfloor \kappa \sqrt{n} \rfloor$ . Aus  $P(z) \sim 1 - d_1 \left(1 - \frac{z}{\rho}\right)^{1/2} = 1 - d_1 \left(-\frac{x}{n}\right)^{1/2}$  und (3.116) folgt

$$e_k(z) \sim \frac{1}{\sqrt{n}} \frac{\exp(-\kappa d_1 \sqrt{-x})}{\frac{1}{2} \frac{1 - \exp(-\kappa d_1 \sqrt{-x})}{d_1 \sqrt{-x}}} = \frac{2}{\sqrt{n}} \frac{d_1 \sqrt{-x}}{\exp(\kappa d_1 \sqrt{-x}) - 1}.$$

Nun gilt

**Satz 3.2.44** Die Wahrscheinlichkeit, dass ein zufälliger Pólya-Baum der Größe  $n$  mindestens die Höhe  $k$  hat, ist gegeben durch

$$\mathbb{P}(H_n \geq k) = \frac{2}{\sqrt{\pi} i} \int_{c-i\infty}^{c+i\infty} \frac{\sqrt{-x} e^{-x}}{\exp(\kappa d_1 \sqrt{-x}) - 1} dx \quad (3.138)$$

wobei  $c$  eine negative Konstante ist.

*Beweis:* Genauso wie der Beweis von Satz 3.2.39. □

Nun können wir die Momente  $\mathbb{E}(H_n^r)$  endlich bestimmen (vgl. mit [11, Theorem 2]):

**Satz 3.2.45** Die mittlere Höhe eines Knotens in einem zufälligen Pólya-Baum der Größe  $n$  hat folgenden asymptotischen Ausdruck:

$$\mathbb{E}(H_n) \sim \frac{2\sqrt{\pi}}{d_1} n^{1/2}. \quad (3.139)$$

Allgemeiner gilt

$$\mathbb{E}(H_n^r) \sim \left(\frac{2}{d_1}\right)^r r(r-1)\Gamma(r/2)\zeta(r)n^{r/2}, \quad (3.140)$$

wobei  $\zeta(z)$  die Riemannsche Zetafunktion ist.

*Beweis:* Genauso wie der Beweis von Satz 3.2.40. □

Nun formulieren wir einen Grenzwertsatz aus (vgl. mit [8, Theorem 4.59]):

**Satz 3.2.46** Die mit  $\sqrt{n}$  skalierte Höhe eines zufälligen Pólya-Baum der Größe erfüllt folgende Konvergenz in Verteilung:

$$\frac{1}{\sqrt{n}}H_n \xrightarrow{d} \frac{2\sqrt{2}}{d_1} \max_{0 \leq t \leq 1} e(t), \quad (3.141)$$

wobei  $e(t)$  eine Brownsche Exkursion der Dauer 1 ist.  $d_1$  ist wie in (2.25).

Wie üblich vergleichen wir unsere Resultate:

	Einfach erzeugte Bäume	Pólya-Bäume
$\mathbb{E}(H_n)$	$\sim \frac{\sqrt{2\pi}}{\sigma} n^{1/2}$	$\sim \frac{2\sqrt{\pi}}{d_1} n^{1/2}$
$\mathbb{E}(H_n^r)$	$\sim \left(\frac{2}{\sigma}\right)^r 2^{-r/2} r(r-1)\Gamma(r/2)\zeta(r)n^{r/2}$	$\sim \left(\frac{2}{d_1}\right)^r r(r-1)\Gamma(r/2)\zeta(r)n^{r/2}$
Grenzverteilung von $\frac{1}{\sqrt{n}}H_n$	$\frac{2}{\sigma} \sup_{0 \leq t \leq 1} e(t)$	$\frac{2\sqrt{2}}{d_1} \sup_{0 \leq t \leq 1} e(t)$

Dabei sind  $\sigma$  und  $d_1$  wie in (2.11) bzw. (2.25) und  $e(t)$  ist die lokale Zeit einer Brownschen Exkursion der Dauer 1. Wir bezeichnen mit  $\Gamma(t)$  die Gammafunktion und mit  $\zeta(t)$  die Riemannsche Zetafunktion.

### 3.3 Weitere Parameter

Wir betrachten nun Parameter, die weder additiv sind noch mit der Form eines Wurzelbaumes zu tun haben.

#### 3.3.1 Transversalen

Interessante Resultate für *Transversalen* wurden von Gittenberger und Kraus in [19] gezeigt.

**Definition 3.3.1** Sei  $T$  ein Wurzelbaum. Eine *Transversale*  $V$  ist eine Teilmenge der Knotenmenge von  $T$  wobei für jeden Pfad von der Wurzel in einem Blatt  $V$  mindestens ein Element des Pfades besitzt.

Verschiedene Werte können als diskrete Zufallsvariablen betrachtet werden: die Anzahl  $X_n$  der Transversalen in einem zufälligen Baum mit  $n$  Knoten, die Anzahl  $X_{n,k}$  der Transversalen der Größe  $k$  in einem zufälligen Baum mit  $n$  Knoten und die Größe  $Y_n$  einer Transversale.

**Definition 3.3.2** Sei  $T$  ein Wurzelbaum. Unter einer zufälligen  $p$ -Menge von  $T$  versteht man eine Teilmenge der Knotenmenge von  $T$ , deren Elemente unabhängig mit Wahrscheinlichkeit  $p$  (mit  $0 < p < 1$ ) gewählt werden.

Die Frage ist, ob eine zufällige  $p$ -Menge in einem Wurzelbaum  $T$  der Größe  $n$  eine Transversale ist (Ereignis  $M_n(T)$ ). Die entsprechende Wahrscheinlichkeit ist gegeben durch

$$\mathbb{P}(M_n(T)) = \sum_{k=0}^n X_{n,k}(T) p^k (1-p)^{n-k} \quad (3.142)$$

(siehe [19, Formel (1)]), wobei  $X_{n,k}(T)$  die Anzahl der Transversalen von  $T$  mit  $k$  Elementen bezeichnet.

### Einfach erzeugte Bäume

Sei  $\mathcal{S}$  eine Familie von einfach erzeugten Bäumen, und  $\mathcal{S}_n := \{S \in \mathcal{S} : |S| = n\}$ . Wir konstruieren für die Transversalen folgende bivariate erzeugende Funktion (siehe [19, Abschnitt 3]):

**Definition 3.3.3** Sei  $X_{n,k}(S)$  die Anzahl der Transversalen mit  $k$  Elementen eines zufälligen Baumes  $S \in \mathcal{S}_n$ . Die bivariate erzeugende Funktion für die Transversalen ist dann gegeben durch

$$S(z, u) = \sum_{n \geq 0} \sum_{k=0}^n s_{n,k} z^n u^k \quad \text{mit} \quad s_{n,k} = \sum_{S \in \mathcal{S}_n} X_{n,k}(S). \quad (3.143)$$

Klarerweise besitzt der Mittelwert für die Anzahl  $X_{n,k}$  der Transversalen mit  $k$  Elementen in einem zufälligen einfach erzeugten Baum der Größe  $n$  folgenden Ausdruck:

$$\mathbb{E}(X_{n,k}) = \frac{s_{n,k}}{s_n}. \quad (3.144)$$

Für die mittlere Größe  $Y_n$  einer Transversalen in einem einfach erzeugten Baum mit  $n$  Knoten haben wir laut Satz 1.3.7 und (3.143)

$$\mathbb{E}(Y_n) = \frac{[z^n] S_u(z, 1)}{[z^n] S(z, 1)} = \frac{\sum_{k=0}^n k s_{n,k}}{\sum_{k=0}^n s_{n,k}},$$

und

$$\begin{aligned} \mathbb{V}(Y_n) &= \frac{[z^n] S_{uu}(z, 1)}{[z^n] S(z, 1)} + \mathbb{E}(Y_n) - (\mathbb{E}(Y_n))^2 \\ &= \frac{\sum_{k=0}^n k(k-1) s_{n,k}}{\sum_{k=0}^n s_{n,k}} + \frac{\sum_{k=0}^n k s_{n,k}}{\sum_{k=0}^n s_{n,k}} - (\mathbb{E}(Y_n))^2 \\ &= \frac{\sum_{k=0}^n k^2 s_{n,k}}{\sum_{k=0}^n s_{n,k}} - (\mathbb{E}(Y_n))^2. \end{aligned}$$

Wir wollen diese Werte (asymptotisch) bestimmen. Dazu brauchen wir zusätzliche Eigenschaften. Zuerst eine kleine Bemerkung über Schranken von  $X_{n,k}(S)$  (siehe auch [19, Formel (8)]).

**Bemerkung 3.3.4** Sei  $S$  ein einfach erzeugter Baum der Größe  $n$ . Jede Teilmenge der Knotenmenge von  $S$ , die die Wurzel enthält ist klarerweise eine Transversale. Andererseits ist jede Transversale eine Teilmenge der Knotenmenge von  $S$ . Daher gilt

$$\binom{n-1}{k-1} \leq X_{n,k}(S) \leq \binom{n}{k}. \quad (3.145)$$

Wir unterscheiden Transversalen, die die Wurzel eines Baumes enthalten und jene, die sie nicht enthalten. Daher zerlegen wir die Klasse der einfach erzeugten Bäume in zwei Teilstrukturen (vgl. mit [19, Abschnitt 3])

$$\mathcal{S} = \mathcal{S}^{(1)} \cup \mathcal{S}^{(2)}, \quad (3.146)$$

wobei:

- $\mathcal{S}^{(1)}$ : Familie von einfach erzeugten Bäume zusammen mit einer Transversalen, die die Wurzel enthält
- $\mathcal{S}^{(2)}$ : Familie von einfach erzeugten Bäume zusammen mit einer Transversalen, die die Wurzel nicht enthält

Für die zweite Klasse gilt sogar

$$\mathcal{S}^{(2)} = \{\circ\} \times \Psi_{\geq 1}(\mathcal{S}), \quad (3.147)$$

wobei  $\Psi$  den Operator für die Gewichtsfunktion  $\Phi(t)$  bezeichnet.

Daher können wir die bivariate erzeugende Funktion (3.143) wie folgt alternativ darstellen:

**Satz 3.3.5** Die bivariate erzeugende Funktion  $S(z, u)$  für die Transversalen in einfach erzeugten Bäume erfüllt:

$$S(z, u) = uz\Phi(S(z(1+u))) + z\Phi(S(z, u)) - z\phi_0. \quad (3.148)$$

*Beweis:* Setze  $s_{n,k}^{(i)} := |\mathcal{S}^{(i)}|$ ,  $i = 1, 2$  und die entsprechenden bivariaten erzeugenden Funktionen  $S^{(i)}(z, u)$  ( $i = 1, 2$ ). Per Konstruktion gilt  $s_{n,k}^{(1)} = s_n \binom{n-1}{k-1}$  und daher

$$\begin{aligned} S^{(1)}(z, u) &= \sum_{n \geq 0} \sum_{k=0}^n s_n \binom{n-1}{k-1} z^n u^k \\ &= \sum_{n \geq 0} s_n u (1+u)^{n-1} z^n \\ &= \frac{u}{1+u} S(z(1+u)). \end{aligned}$$

Für den zweiten Teil erhalten wir aus (3.147)

$$S^{(2)}(z, u) = z\Phi(S(z, u)) - z\phi_0.$$

(3.146) liefert uns das gewünschte Ergebnis.  $\square$

Nun haben wir (siehe [7, Theorem 1] und [19, Theorem 2]):

**Satz 3.3.6** *Für die mittlere Wahrscheinlichkeit, dass eine zufällige  $p$ -Menge eine Transversale ist, gilt*

$$\lim_{n \rightarrow \infty} \mathbb{E}(\mathbb{P}(M_n)) = \frac{p}{1 - (1-p)\rho\Phi'(S_0)} \quad (3.149)$$

wobei  $S_0$  die Lösung der Gleichung

$$S_0 = p\tau + (1-p)\tau \left( \frac{\Phi(S_0)}{\Phi(\tau)} - \frac{\phi_0}{\Phi(\tau)} \right) \quad (3.150)$$

bezeichnet, und  $\rho$  die dominante Singularität von  $S(z)$  ist.

*Beweis:* Der gesuchte Erwartungswert  $\mathbb{E}(\mathbb{P}(P_n))$  erfüllt wegen (3.142) und (3.143)

$$\begin{aligned} \mathbb{E}(\mathbb{P}(M_n)) &= \frac{\sum_{S \in \mathcal{S}_n} \mathbb{P}(M_n(S))}{s_n} \\ &= \frac{\sum_{S \in \mathcal{S}_n} \sum_{k=0}^n X_{n,k}(S) p^k (1-p)^{n-k}}{s_n} \\ &= \frac{[z^n] S \left( (1-p)z, \frac{p}{1-p} \right)}{s_n}. \end{aligned} \quad (3.151)$$

Die Idee wäre, das Gleichungssystem

$$S(z, u) = uz\Phi(S(z(1+u))) + z\Phi(S(z, u)) - z\phi_0 \quad (3.152)$$

$$1 = z\Phi'(S(z, u)) \quad (3.153)$$

bezüglich  $z$  und  $S$  für fixes  $u > 0$  mithilfe des Hauptsatz über implizite Funktionen zu lösen. Aus [19, Teil 3] wissen wir, dass  $\rho_0(u) = \frac{\rho}{1+u}$  die Singularität von  $S(z, u)$  ist. Da aber die Funktion

$$F(z, u, s) = s - uz\Phi(S(z(1+u))) + z\Phi(s) - z\phi_0$$

bei  $z = \rho_0(u)$  singulär ist, ist der Hauptsatz über implizite Funktionen nicht direkt anwendbar.

Setze  $u_0 = \frac{p}{1-p}$ ,  $\rho_0 := \rho_0(u_0) = (1-p)\rho$  und  $S_0 := S(\rho_0, u_0)$ . Einsetzen in (3.152) liefert uns (3.150). Aus (2.14) folgt

$$S(z(1+u_0)) = \tau - d_1 \left( 1 - \frac{z(1+u_0)}{\rho} \right)^{1/2} + \mathcal{O} \left( \left| 1 - \frac{z(1+u_0)}{\rho} \right| \right), \quad z \rightarrow \rho_0 \quad (3.154)$$

und daher (vgl. mit [19, Formel (15)])

$$S(z, u_0) = S_0 - d_1 \frac{p}{1 - \rho_0 \Phi'(S_0)} \left(1 - \frac{z}{\rho_0}\right)^{1/2} + \mathcal{O}(|z - \rho_0|) \quad (3.155)$$

lokal um  $z = \rho_0$ . Der Transfersatz 1.2.16 und Dividieren durch  $s_n \sim \frac{d_1 \rho^{-n}}{2\sqrt{\pi n^{3/2}}}$  liefert uns (3.149). Für einen vollständigen Beweis siehe [7, Theorem 1] oder [19, Theorem 2].  $\square$

**Korollar 3.3.7** Sei  $X_n = \sum_{k=0}^n X_{n,k}$  die Anzahl der Transversalen in einem zufälligen einfach erzeugten Baum mit  $n$  Knoten. Dann gilt

$$\mathbb{E}(X_n) = \frac{\sum_{k=0}^n s_{n,k}}{s_n} \sim \frac{2^n}{2 - \rho \Phi'(S_0)}. \quad (3.156)$$

*Beweis:* Sei  $S \in \mathcal{S}_n$ . Setze  $p = 1/2$  in (3.142). Damit bekommen wir

$$\mathbb{P}(M_n(S)) = 2^{-n} \sum_{k=0}^n X_{n,k}(S).$$

Mithilfe von (3.143), (3.151) und (3.149) erhalten wir

$$\mathbb{E}(X_n) = \frac{\sum_{k=0}^n s_{n,k}}{s_n} = \frac{\sum_{S \in \mathcal{S}_n} \sum_{k=0}^n X_{n,k}(S)}{s_n} = 2^n \mathbb{E}(\mathbb{P}(M_n)) \sim \frac{2^n}{2 \left(1 - \frac{1}{2} \rho \Phi'(S_0)\right)}.$$

$\square$

**Satz 3.3.8** Ist  $k = \lambda n$  mit  $(\lambda \in [\epsilon, 1 - \epsilon])$ , dann ist die erwartete Anzahl der Transversalen der Größe  $k$  in einem zufälligen einfach erzeugten Baum mit  $n$  Knoten gegeben durch

$$\mathbb{E}(X_{n,k}) \sim \frac{\rho(1 - \lambda)}{1 - \rho(1 - \lambda)\Phi'(S_0)} \sqrt{\frac{1 - \lambda}{2\pi\lambda n}} \frac{1}{\lambda^{\lambda n} (1 - \lambda)^{(1 - \lambda)n}}. \quad (3.157)$$

*Beweis:* siehe [19, Theorem 3].  $\square$

**Satz 3.3.9** Die mittlere Größe  $Y_n$  einer Transversale in einem einfach erzeugten Baum mit  $n$  Knoten ist asymptotisch normalverteilt mit Mittelwert

$$\mathbb{E}(Y_n) = \frac{\sum_{k=0}^n k s_{n,k}}{\sum_{k=0}^n s_{n,k}} \sim \frac{1}{2} n, \quad (3.158)$$

und Varianz

$$\mathbb{V}(Y_n) = \frac{\sum_{k=0}^n k^2 s_{n,k}}{\sum_{k=0}^n s_{n,k}} - (\mathbb{E}(Y_n))^2 \sim \frac{1}{4} n. \quad (3.159)$$

*Beweis:* Im Beweis von Satz 3.3.6 haben wir vorausgesetzt, dass  $u$  ein reeller Parameter ist. Hier sollen wir aber beachten, dass  $u$  in einer komplexen Umgebung von 1 liegt, um den Fehler abzuschätzen und damit die asymptotische Normalität zu zeigen. Die Idee ist die folgende:

Zuerst stellen wir die Funktion  $S(z(1 + u))$  lokal dar. Diese Darstellung impliziert, dass

der Fehlerterm in (3.155) für  $|u - 1| < \epsilon$  ( $\epsilon > 0$ ) und  $z \rightarrow \rho_0(u) = \frac{\rho}{1+u}$  in einem  $\Delta$ -Gebiet der Form  $\frac{1}{1+u}\Delta_\rho$  ( $\Delta_\rho$  ist wie in Bemerkung 1.2.26) gleichmäßig ist (siehe Beweis von Theorem 1 in [19]). Daher können wir das System (3.152)-(3.153) umschreiben und damit den Hauptsatz über implizite Funktionen und Theorem 2.18 in [8] anwenden. Wir bekommen dann

$$[z^n]S(z, u) = \frac{d_1}{(1+u)(1-\rho_0(u)\Phi'(S(\rho_0(u), u)))} \frac{\rho_0(u)^{-n}}{2\sqrt{\pi n^{3/2}}} (1 + \mathcal{O}(n^{-1})) \quad (3.160)$$

für  $|u - 1| < \epsilon$  (vgl. mit [19, Abschnitt 3]). Aus (1.41) erhalten wir

$$\begin{aligned} \mathbb{E}(u^{Y_n}) &= \frac{[z^n]S(z, u)}{[z^n]S(z, 1)} \\ &= \frac{2(1-\rho_0(1)\Phi'(S(\rho_0(1), 1)))}{(1+u)(1-\rho_0(u)\Phi'(S(\rho_0(u), u)))} \left(\frac{\rho_0(1)}{\rho_0(u)}\right)^n (1 + \mathcal{O}(n^{-1})). \end{aligned}$$

Nun können wir das Quasi-Power Theorem von Hwang (Satz 1.3.22) anwenden, und wir erhalten den Mittelwert

$$\begin{aligned} \mathbb{E}(Y_n) &= -\frac{\rho'_0(1)}{\rho_0(1)}n + \mathcal{O}(1) \\ &= -\frac{\frac{\rho}{2^2}}{\frac{\rho}{2}}n + \mathcal{O}(1) \\ &= \frac{n}{2} + \mathcal{O}(1) \end{aligned}$$

und die Varianz

$$\begin{aligned} \mathbb{V}(Y_n) &= \left(-\frac{\rho''_0(1)}{\rho_0(1)} - \frac{\rho'_0(1)}{\rho_0(1)} + \left(\frac{\rho'_0(1)}{\rho_0(1)}\right)^2\right)n + \mathcal{O}(1) \\ &= \left(-\frac{\frac{2\rho}{2^3}}{\frac{\rho}{2}} + \frac{1}{2} + \frac{1}{4}\right)n + \mathcal{O}(1) \\ &= \frac{n}{4} + \mathcal{O}(1) \end{aligned}$$

sowie die asymptotische Normalität von  $Y_n$ . Für einen vollständigen Beweis siehe [19, Theorem 1].  $\square$

## Pólya-Bäume

Die Ausarbeitung für Pólya-Bäume ist ähnlich wie bei einfach erzeugten Bäumen. Zuerst definieren wir eine bivariate erzeugende Funktion, die eine ähnliche Darstellung wie  $S(z, u)$  besitzt (vgl. mit [19, Abschnitt 4]).

**Definition 3.3.10** *Genauso wie bei einfach erzeugten Bäumen ist die bivariate erzeugende Funktion  $P(z, u)$  für die Transversalen in Pólya-Bäume gegeben durch*

$$P(z, u) = \sum_{n \geq 0} \sum_{k=0}^n p_{n,k} z^n u^k \quad \text{mit} \quad p_{n,k} = \sum_{P \in \mathcal{P}_n} X_{n,k}(P) \quad (3.161)$$

wobei  $X_{n,k}(P)$  die Anzahl der Transversalen in einem Pólya-Baum der Größe  $n$  bezeichnet.



Mit analoger Argumentation wie in vorigen Abschnitt kann man die bivariate kombinatorische Klasse  $\mathcal{P}$  der Pólya-Bäume wie folgt zerlegen (siehe [19, Formel (17)]):

$$\mathcal{P} = \mathcal{P}^{(1)} \cup \mathcal{P}^{(2)} \quad (3.162)$$

wobei:

- $\mathcal{P}^{(1)}$ : Familie von Pólya-Bäume zusammen mit einer Transversalen, die die Wurzel enthält
- $\mathcal{P}^{(2)}$ : Familie von Pólya-Bäume zusammen mit einer Transversalen, die die Wurzel nicht enthält

Dabei gilt

$$\mathcal{P}^{(2)} = \{\circ\} \times M_{\geq 1}(\mathcal{P}), \quad (3.163)$$

wobei  $M_{\geq 1}(\mathcal{P})$  Multimengen von Objekten aus  $\mathcal{P}$  bezeichnet, in welche jede Komponente von  $\mathcal{P}$  mindestens einmal vorkommt. Wegen der Nichtplanarität in  $\mathcal{P}$  und sogar in  $\mathcal{P}^{(i)}$  ( $i = 1, 2$ ) ist es schwieriger eine so gute untere Schranke wie in (3.145) zu bestimmen (die beste wäre 1, vgl. mit [19, Abschnitt 4]). Daher definieren wir eine zusätzliche Klasse von Pólya-Bäumen  $\tilde{\mathcal{P}}$ , in der die Anzahl der ausgezeichneten Knoten beliebig ist. Wir erhalten (siehe Formeln (19) und (20) in [19])

$$\tilde{\mathcal{P}} = (\{\circ\} + \mu\{\circ\}) \times M(\tilde{\mathcal{P}}) \quad (3.164)$$

$$\mathcal{P}^{(1)} = \mu\{\circ\} \times M(\tilde{\mathcal{P}}) \quad (3.165)$$

und damit (vgl. mit [19, Formel (22)]):

**Satz 3.3.11** *Die bivariate erzeugende Funktion  $P(z, u)$  für die Transversalen in Pólya-Bäumen besitzt folgende alternative Darstellung:*

$$P(z, u) = \frac{u}{1+u} \tilde{P}(z, u) + z \left( \exp \left( \sum_{i \geq 1} \frac{P(z^i, u^i)}{i} \right) - 1 \right), \quad (3.166)$$

wobei  $\tilde{P}(z, u)$  die bivariate erzeugende Funktion von  $\tilde{\mathcal{P}}$  ist.

*Beweis:* Seien  $P^{(i)}(z, u)$  die bivariaten erzeugenden Funktionen von  $\mathcal{P}^{(i)}$  ( $i = 1, 2$ ). Wegen (3.163) gilt

$$P^{(2)}(z, u) = z \left( \exp \left( \sum_{i \geq 1} \frac{P(z^i, u^i)}{i} \right) - 1 \right).$$

Aus (3.164) bekommen wir

$$\tilde{P}(z, u) = z(1+u) \exp \left( \sum_{i \geq 1} \frac{\tilde{P}(z^i, u^i)}{i} \right)$$

und aus (3.165)

$$P^{(1)}(z, u) = zu \exp \left( \sum_{i \geq 1} \frac{\tilde{P}(z^i, u^i)}{i} \right) = \frac{u}{1+u} \tilde{P}(z, u).$$

(3.162) liefert uns das gewünschte Ergebnis. □

**Satz 3.3.12** Die mittlere Wahrscheinlichkeit, dass eine zufällige  $p$ -Menge in einem Pólya-Baum mit  $n$  Knoten eine Transversale ist, erfüllt:

$$\lim_{n \rightarrow \infty} \mathbb{E}(\mathbb{P}(M_n)) = p \left( 1 - \rho_0 \exp \left( \sum_{i \geq 1} \frac{P(\rho_0^i, (p/(1-p))^i)}{i} \right) \right)^{-1}. \quad (3.167)$$

Dabei ist  $\rho_0$  die einzige positive Singularität von  $\tilde{P}(z, p/(1-p))$  wobei  $\tilde{P}(z, u)$  die bivariate erzeugende Funktion der Struktur  $\tilde{\mathcal{P}}$  ist.

*Beweis:* Setze  $u > 0$ . In [19, Lemma 2] wird gezeigt, dass die Abbildungen  $z \mapsto P(z, u)$  und  $z \mapsto \tilde{P}(z, u)$  dieselbe dominante Singularität  $\rho_0(u)$  besitzen. Damit ist der Satz 1.2.27 (Hauptsatz über implizite Funktionen) für (3.166) nicht direkt anwendbar. Mit analoger Argumentation wie im Satz 3.3.6 können wir (3.167) beweisen. Für einen vollständigen Beweis siehe [19, Theorem 4].  $\square$

Mithilfe dieses Satzes können wir folgende Aussage mit  $p = 1/2$  beweisen.

**Korollar 3.3.13** Die mittlere Anzahl  $X_n$  der Transversalen in einem Pólya-Baum der Größe  $n$  hat folgenden Ausdruck

$$\mathbb{E}(X_n) \sim \frac{2^n}{2 - 2\tilde{\rho} \exp \left( \sum_{i \geq 1} \frac{P(\tilde{\rho}^i, 1)}{i} \right)} \quad (3.168)$$

wobei  $\tilde{\rho} \approx 0.180343$  die dominante Singularität von  $\tilde{P}(z) := \tilde{P}(z, 1)$  ist. Aus [16, Abschnitt 4] wissen wir, dass

$$\frac{1}{2 - 2\tilde{\rho} \exp \left( \sum_{i \geq 1} \frac{P(\tilde{\rho}^i, 1)}{i} \right)} \approx 0.6126998.$$

*Beweis:* Genauso wie in Korollar 3.3.7.  $\square$

Nun formulieren wir unser Ergebnis aus (vgl. mit [19, Theorem 5]).

**Satz 3.3.14** Die mittlere Größe  $Y_n$  einer Transversalen in Pólya-Bäumen mit  $n$  Knoten ist asymptotisch normalverteilt mit Mittelwert

$$\mathbb{E}(Y_n) \sim \mu n, n \rightarrow \infty \quad (3.169)$$

und Varianz

$$\mathbb{V}(Y_n) \sim \sigma^2 n, n \rightarrow \infty. \quad (3.170)$$

Dabei sind  $\mu \approx 0.505903$  und  $\sigma^2$  numerisch berechenbare Konstanten.

*Beweis:* Mit ähnlicher Überlegung wie im Satz 3.3.9 können wir den Hauptsatz über implizite Funktionen anwenden. Wir bekommen für  $Y_n$  eine wahrscheinlichkeitserzeugende Funktion der Form (1.53). Aus dem Satz 1.3.22 (Quasi-Power Theorem) folgt die asymptotische Normalität. Für den Mittelwert berechnen wir die Ableitung  $P_u(z, 1)$  mithilfe von (3.166), um damit den Ausdruck

$$\mathbb{E}(Y_n) = \frac{[z^n] P_u(z, 1)}{p_n}$$

zu approximieren.

Für die Varianz berechnen wir  $\frac{\partial}{\partial u} \left( u \frac{\partial}{\partial u} P(z, u) \right) \Big|_{u=1}$  und bekommen damit wegen (1.45)

$$\mathbb{V}(Y_n) = \frac{[z^n] \frac{\partial}{\partial u} \left( u \frac{\partial}{\partial u} P(z, u) \right) \Big|_{u=1}}{p_n} - \mathbb{E}(Y_n)^2.$$

Für eine genaue Ausarbeitung siehe [19, Theorem 5]. □

Wir fassen unsere Resultate zusammen:

	Einfach erzeugte Bäume	Pólya-Bäume
$\mathbb{E}(X_n)$	$\sim \frac{2^n}{2 - \rho \Phi'(S_0)}$	$\sim \frac{2^n}{2 - 2\tilde{\rho} \exp(\sum_{i>1} P(\tilde{\rho}^i, 1)/i)}$
$\mathbb{E}(Y_n)$	$\sim \frac{1}{2}n$	$\sim \mu n$
$\mathbb{V}(Y_n)$	$\sim \frac{1}{4}n$	$\sim \sigma^2 n$
Grenzverteilung von $\frac{Y_n - \mathbb{E}(Y_n)}{\sqrt{\mathbb{V}(Y_n)}}$	$\mathcal{N}(0, 1)$	$\mathcal{N}(0, 1)$

Dabei gilt  $\tilde{\rho} \approx 0.180343$ ,  $\mu \approx 0.505903$  und  $\rho$  bzw.  $S_0$  sind wie in (2.13) bzw. (3.150).

### 3.3.2 Protection number

Der Parameter, der in diesem Abschnitt präsentiert wird, wurde für einfach erzeugte Bäume und Pólya-Bäume in [18] analysiert.

**Definition 3.3.15** *In einem Wurzelbaum bezeichnet die protection number eines Knoten die Länge des kürzesten Pfades dieses Knotens zu einem Blatt. Ist der betroffene Knoten die Wurzel, so spricht man von der protection number des Baumes.*

#### Einfach erzeugte Bäume

Um eine Verteilung der protection number von einfach erzeugten Bäumen zu definieren brauchen wir ein zusätzliche erzeugende Funktion (siehe [18, Formel (1)]).

**Definition 3.3.16** *Wir nehmen an, dass die Gewichtsfunktion  $\Phi(t)$  aperiodisch ist. Dann ist die erzeugende Funktion für einfach erzeugte Bäume, die mindestens die protection number  $k$  haben, definiert durch*

$$S_{\geq k}(z) = z(\Phi(S_{\geq k-1}(z)) - \phi_0). \tag{3.171}$$

**Bemerkung 3.3.17** *Offensichtlich gelten  $S_{\geq 0}(z) = S(z)$  und*

$$S_{\geq k}(z) = \Omega^k(S(z)) \tag{3.172}$$

wobei  $\Omega(t) = z(\Phi(t) - \phi_0)$  und  $\Omega^k(\cdot)$  die  $k$ -fache Hintereinanderausführung davon ist.

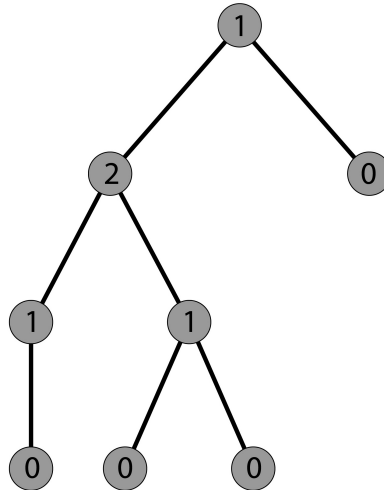


Abbildung 3.3: Die protection numbers der Knoten eines Motzkin-Baumes der Größe 8.

Interessant ist das asymptotische Verhalten von  $S_{\geq k}(z)$  (siehe Beweis von Theorem 1 in [18])

**Lemma 3.3.18** *Für jedes  $k \in \mathbb{N}_0$  hat  $S_{\geq k}(z)$  die gleiche dominante Singularität wie  $S(z)$  und eine entsprechende Puiseux-Reihe der Gestalt*

$$S_{\geq k}(z) = d_{0,k} + d_{1,k} \sqrt{1 - \frac{z}{\rho}} + \mathcal{O}\left(1 - \frac{z}{\rho}\right) \quad (3.173)$$

wobei  $d_{1,k} = \Omega'(d_{0,k-1})d_{1,k-1} = (\Omega')^k(\tau)d_1$ . Dabei ist  $d_1$  ist wie im Satz 2.1.14.

*Beweis:* Wir zeigen die Aussage für  $k = 1$  (der Rest folgt wegen (3.172) durch Induktion). Die Funktion  $\Omega(t)$  ist in  $S(\rho) = \tau$  analytisch. Wir erinnern uns, dass

$$S(z) = \tau - d_1 \left(1 - \frac{1}{\rho}\right)^{1/2} + \mathcal{O}\left(1 - \frac{z}{\rho}\right), \quad z \rightarrow \rho$$

mit  $d_1 = \sqrt{\frac{2\Phi(\tau)}{\Phi''(\tau)}}$ .

Daher hat die Komposition  $S_{\geq 1}(z) = \Omega(S(z))$  bei  $z = \rho$  folgende singuläre Entwicklung:

$$S_{\geq 1}(z) = \Omega(\tau) - \Omega'(\tau)d_1 \left(1 - \frac{1}{\rho}\right)^{1/2} + \mathcal{O}\left(1 - \frac{z}{\rho}\right), \quad z \rightarrow \rho. \quad (3.174)$$

Wir wissen schon, dass die Gewichtsfunktion  $\Phi(t)$  nur aus nichtnegativen Koeffizienten besteht. Da  $\tau, \rho > 0$  gilt, sind  $\Omega(\tau)$  und  $\Omega'(\tau)$  ebenfalls nichtnegativ. Damit ist die Aussage gezeigt.  $\square$

Die Folgerung ist offensichtlich.

**Lemma 3.3.19** Die Koeffizienten von  $S_{\geq k}(z)$  besitzen folgende asymptotische Darstellung:

$$s_{n,k} = [z^n]S_{\geq k}(z) \sim \frac{d_{1,k}\rho^{-n}}{2\sqrt{\pi n^{3/2}}} \left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right), \quad n \rightarrow \infty \quad (3.175)$$

mit  $d_{1,k} = \Omega'(d_{0,k-1})d_{1,k-1} = (\Omega')^k(\tau)d_1$ .

*Beweis:* Folgt aus den Transfersätzen 1.2.16 und 1.2.25. □

**Bemerkung 3.3.20** Es gilt

$$d_{0,k} = \Omega(d_{0,k-1}) = \Omega^k(\tau) = \Omega^k(S(\rho)) = S_{\geq k}(\rho) \quad (3.176)$$

für alle  $k \in \mathbb{N}_0$ .

Die protection number eines einfach erzeugten Baumes induziert eine Zufallsvariable  $X_n$ . Die entsprechende Verteilung ist gegeben durch:

$$\mathbb{P}(X_n \geq k) = \frac{[z^n]S_{\geq k}(z)}{[z^n]S(z)} \quad (3.177)$$

Unserer Ziel ist den Erwartungswert

$$\mathbb{E}(X_n) = \sum_{k \geq 1} \mathbb{P}(X_n \geq k) \quad (3.178)$$

asymptotisch zu bestimmen. Der folgende Satz liefert uns das gesuchte Ergebnis (siehe auch [18, Theorem 1])

**Satz 3.3.21** Sei  $X_n$  die protection number eines zufälligen einfach erzeugten Baumes mit  $n$  Knoten. Für den Mittelwert und die Varianz von  $X_n$  gelten dann

$$\lim_{n \rightarrow \infty} \mathbb{E}(X_n) = \sum_{k \geq 1} \rho^{k-1} \prod_{i=1}^{k-1} \Phi'(S_{\geq i}(\rho)) \quad (3.179)$$

und

$$\lim_{n \rightarrow \infty} \mathbb{V}(X_n) = \sum_{k \geq 1} (2k-1)\rho^{k-1} \prod_{i=1}^{k-1} \Phi'(S_{\geq i}(\rho)) - \left(\lim_{n \rightarrow \infty} \mathbb{E}(X_n)\right)^2. \quad (3.180)$$

*Beweis:* Wegen (3.176) kann man  $d_{1,k}$  wie folgt umschreiben:

$$\begin{aligned} d_{1,k} &= \Omega'(S_{\geq k-1}(\rho))d_{1,k-1} \\ &= \rho\Phi'(S_{\geq k-1}(\rho))d_{1,k-1} \\ &= \rho^2\Phi'(S_{\geq k-1}(\rho))\Phi'(S_{\geq k-2}(\rho))d_{1,k-2} \\ &= \rho^{k-1} \prod_{i=1}^{k-1} \Phi'(S_{\geq i}(\rho))d_{1,0}. \end{aligned}$$

Da  $d_{1,0} = d_1$  gilt, erhalten wir

$$\begin{aligned}\mathbb{E}(X_n) &= \sum_{k \geq 1} \frac{[z^n] S_{\geq k}(z)}{[z^n] S(z)} \\ &\sim \sum_{k \geq 1} \frac{\frac{d_{1,k} \rho^{-n}}{2\sqrt{\pi n^{3/2}}}}{\frac{d_1 \rho^{-n}}{2\sqrt{\pi n^{3/2}}}} \\ &= \sum_{k \geq 1} \frac{d_{1,k}}{d_1} \\ &= \sum_{k \geq 1} \rho^{k-1} \prod_{i=1}^{k-1} \Phi'(S_{\geq i}(\rho)).\end{aligned}$$

Der Ausdruck für die Varianz folgt aus (1.45) mit  $\mathbb{E}(X_n^2) = \sum_{k \geq 1} (2k-1) \mathbb{P}(X_n \geq k)$ .  $\square$

Nun interessieren wir uns für die protection number eines zufälligen Knoten in einem einfach erzeugten Baum der Größe  $n$ . Die Idee, um Ergebnisse zu berechnen, ist, ein Blatt zu markieren, und dieses durch einen Baum der protection Number  $k$  zu ersetzen. Dazu brauchen wir die bivariate erzeugende Funktion  $S^{(0)}(z, u)$  aus Satz 3.1.6. Wir erinnern uns, dass  $z$  bzw.  $u$  die Größe bzw. die Anzahl der Blätter markiert. Mit dieser kann man einen Ausdruck für eine weitere erzeugende Funktion bestimmen (siehe auch [18, Formel (5)]):

**Definition 3.3.22** Sei  $\tilde{S}_k(z)$  die erzeugende Funktion der  $k$ -protected Knoten. Diese erfüllt per Definition folgende Gleichung:

$$\tilde{S}_k(z) = \sum_{n \geq 0} \tilde{s}_{n,k} z^n = \frac{S_{\geq k}(z)}{z} \frac{\partial}{\partial u} S^{(0)}(z, 1). \quad (3.181)$$

Sei  $\tilde{X}_n$  die mittlere protection Number eines zufälligen Knoten in einem einfach erzeugten Baum der Größe  $n$ . Das Verfahren, um den Mittelwert  $\mathbb{E}(\tilde{X}_n)$  zu bestimmen, ist das gleiche wie bei  $X_n$ . Die Wahrscheinlichkeit  $\mathbb{P}(\tilde{X}_n \geq k)$  wird bestimmt, und über alle  $k \in \mathbb{N}$  aufsummiert. Diese Wahrscheinlichkeit  $\mathbb{P}(\tilde{X}_n \geq k)$  ist gegeben durch:

$$\mathbb{P}(\tilde{X}_n \geq k) = \frac{[z^n] \tilde{S}_k(z)}{n [z^n] S(z)}. \quad (3.182)$$

(vgl. mit [18, Formel (6)]). Dann haben wir (siehe auch [18, Theorem 2]):

**Satz 3.3.23** Für die mittlere protection Number  $\tilde{X}_n$  eines zufälligen Knoten in einem einfach erzeugten Baum der Größe  $n$  gilt

$$\lim_{n \rightarrow \infty} \mathbb{E}(\tilde{X}_n) = \frac{\phi_0}{S(\rho)} \sum_{k \geq 1} S_{\geq k}(\rho) \quad (3.183)$$

und

$$\lim_{n \rightarrow \infty} \mathbb{V}(\tilde{X}_n) = \frac{\phi_0}{S(\rho)} \sum_{k \geq 1} (2k-1) S_{\geq k}(\rho) - \left( \lim_{n \rightarrow \infty} \mathbb{E}(\tilde{X}_n) \right)^2. \quad (3.184)$$

*Beweis:* Wir nutzen die Darstellung (3.181) von  $\tilde{S}_k(z)$ , um die Koeffizienten dieser Reihe zu bestimmen. Aus (3.15) folgt:

$$\begin{aligned} S_u^{(0)}(z, 1) &= z\Phi'(S^{(0)}(z, 1))S_u^{(0)}(z, 1) + \phi_0 z \\ &= z\Phi'(S(z))S_u^{(0)}(z, 1) + \phi_0 z \\ &= \frac{\phi_0 z}{1 - z\Phi'(S(z))}. \end{aligned}$$

Wir können dies ein wenig umschreiben. Aus  $S'(z) = \Phi(S(z)) + z\Phi'(S(z))S'(z)$  und  $\Phi(S(z)) = \frac{S(z)}{z}$  folgt

$$z\Phi'(S(z)) = \frac{S'(z) - \Phi(S(z))}{S'(z)} = 1 - \frac{S(z)}{zS'(z)}$$

und

$$S_u^{(0)}(z, 1) = \frac{\phi_0 z^2 S'(z)}{S(z)}.$$

Dabei gilt

$$\tilde{S}_k(z) = \frac{\phi_0 z S'(z) S_{\geq k}(z)}{S(z)}.$$

Wir wissen, dass  $[z^n]S_{\geq k}(z) \sim \frac{d_{1,k}\rho^{-n}}{2\sqrt{\pi n^{3/2}}}$  und  $[z^n]\frac{zS'(z)}{S(z)} \sim \frac{ns_n}{\tau} \sim \frac{d_1\rho^{-n}}{2\sqrt{\pi\tau n^{1/2}}}$ . Nun können wir Lemma 1.2.21 (iv) anwenden, und wir bekommen

$$[z^n]\tilde{S}_k(z) \sim \frac{\phi_0 S_{\geq k}(\rho) ns_n}{\tau}$$

und daher

$$\begin{aligned} \mathbb{E}(\tilde{X}_n) &= \sum_{k \geq 1} \mathbb{P}(\tilde{X}_n \geq k) \\ &= \sum_{k \geq 1} \frac{[z^n]\tilde{S}_k(z)}{n[z^n]S(z)} \\ &\sim \sum_{k \geq 1} \frac{\phi_0 S_{\geq k}(\rho) ns_n}{\tau ns_n} \\ &= \frac{\phi_0}{\tau} \sum_{k \geq 1} S_{\geq k}(\rho). \end{aligned}$$

Die Varianz erhält man wieder mit  $\mathbb{V}(\tilde{X}_n) = \sum_{k \geq 1} (2k-1)\mathbb{P}(\tilde{X}_n \geq k) - \left(\mathbb{E}(\tilde{X}_n)\right)^2$ .  $\square$

**Beispiel 3.3.24** Als erstes Beispiel betrachten wir die Catalan-Bäume. Wir haben  $\rho = \frac{1}{4}$ ,  $\tau = \frac{1}{2}$ ,  $\Phi(t) = \frac{1}{1-t}$  und damit  $\Phi'(t) = \frac{1}{(1-t)^2}$ . Wir bekommen folgende Rekursion:

$$S_{\geq k}(\rho) = \frac{1}{4} \left( \frac{1}{1 - S_{\geq k-1}(\rho)} - 1 \right), S_{\geq 1}(\rho) = \frac{1}{4}.$$

Explizit haben wir

$$S_{\geq k}(\rho) = \frac{3}{2(4^k + 2)}$$

und daher

$$\lim_{n \rightarrow \infty} \mathbb{E}(X_n) = \sum_{k \geq 1} \left(\frac{1}{4}\right)^{k-1} \prod_{i=1}^{k-1} \frac{1}{\left(1 - \frac{3}{2(4^i+2)}\right)^2} \approx 1.622971384715353.$$

**Beispiel 3.3.25** Nun betrachten wir die Cayley-Bäume. Zur Erinnerung, hier sind  $\tau = 1$ ,  $\rho = e^{-1}$ ,  $\phi_0 = 1$  und  $\Phi(t) = \Phi'(t) = e^t$ . Daraus folgt:

$$S_{\geq k}(\rho) = e^{-1} \left( e^{S_{\geq k-1}(\rho)} - 1 \right), S_{\geq 1}(\rho) = e^{-1}(e^1 - 1) = 1 - e^{-1}.$$

Diese Rekursion kann nicht explizit gelöst werden (siehe [ , Example(Cayley trees)]). Doch erhalten wir numerische Ergebnisse:

$$\lim_{n \rightarrow \infty} \mathbb{E}(X_n) \approx 2.286198316708012$$

$$\lim_{n \rightarrow \infty} \mathbb{V}(X_n) \approx 1.598472890455086.$$

**Beispiel 3.3.26** Für strikte Binärbäume sind  $\rho = \frac{1}{2}$ ,  $\tau = 1$ ,  $\phi_0 = 1$ ,  $\Phi(t) = 1 + t^2$  und  $\Phi'(t) = 2t$ . Dann gilt

$$S_{\geq k}(\rho) = \frac{1}{2} S_{\geq k-1}(\rho)^2 = \frac{S_0(\rho)^{2^k}}{2^{\sum_{j=0}^{k-1} 2^j}} = 2^{1-2^k}$$

und daher

$$\lim_{n \rightarrow \infty} \mathbb{E}(X_n) = \sum_{k \geq 1} 2^{1-k} \prod_{i=1}^{k-1} 2^{2-2^i} \approx 3.536472483525321.$$

Für Binärbäume zweiter Art und die Motzkin Bäume gibt es keine explizite Darstellung für  $S_{\geq k}(\rho)$ . Doch können die Grenzwerte von  $\mathbb{E}(X_n)$  und  $\mathbb{V}(X_n)$  numerisch berechnet werden (siehe Tabelle am Ende des Abschnittes).

## Pólya-Bäume

Analog wie bei einfach erzeugten Bäumen brauchen wir eine zusätzliche erzeugende Funktion (vgl. mit [18, Formel (8)]):

**Definition 3.3.27** Die erzeugende Funktion  $P_{\geq k}(z)$  von Pólya Bäume, die  $k \geq 1$  als minimale protected number hat, besitzt folgende Darstellung:

$$P_{\geq k}(z) = z \exp \left( \sum_{i \geq 1} \frac{P_{\geq k-1}(z)}{i} \right) - z. \quad (3.185)$$

**Bemerkung 3.3.28** Klarerweise gilt  $P_{\geq 0}(z) = P(z)$ .

Wie vorher berechnen wir die singuläre Entwicklung von  $P_{\geq k}(z)$  (siehe Beweis von Lemma 3 in [18]):



**Lemma 3.3.29** Für jedes  $k \in \mathbb{N}_0$  besitzt  $P_{\geq k}(z)$  die gleiche dominante Singularität  $\rho$  wie  $P(z)$  und eine Puiseux-Reihe der Form

$$P_{\geq k}(z) = P_{\geq k}(\rho) - d_{1,k} \left(1 - \frac{z}{\rho}\right)^{1/2} + \mathcal{O}\left(1 - \frac{z}{\rho}\right), \quad z \rightarrow \rho \quad (3.186)$$

mit Konstante  $d_{1,k} > 0$ .

*Beweis:* Mit Induktion wobei  $P_{\geq 0}(z) = P(z)$  und wegen der rekursiven Darstellung von  $P_{\geq k}(z)$  (3.185).  $\square$

Wie im vorigen Abschnitt induziert die protection Number  $X_n$  von Pólya-Bäume eine diskrete Zufallsvariable. Dazu definieren wir folgende bedingte Wahrscheinlichkeit:

$$\mathbb{P}(X_n \geq k | X_n \geq k-1) = \frac{[z^n]P_{\geq k}(z)}{[z^n]P_{\geq k-1}(z)}. \quad (3.187)$$

Wir können diese Koeffizienten asymptotisch bestimmen (vgl. mit [18, Lemma 3]):

**Lemma 3.3.30** Für die Koeffizienten von  $P_{\geq k}(z)$  und  $P_{\geq k-1}(z)$  gilt

$$p_{n,k-1} = [z^n]P_{\geq k-1}(z) = \frac{d_{1,k-1}\rho^{-n}}{2\sqrt{\pi n^{3/2}}} \left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right), \quad n \rightarrow \infty \quad (3.188)$$

und damit

$$p_{n,k} = [z^n]P_{\geq k}(z) = \frac{(P_{\geq k}(\rho) + \rho)d_{1,k-1}\rho^{-n}}{2\sqrt{\pi n^{3/2}}} \left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right), \quad n \rightarrow \infty. \quad (3.189)$$

*Beweis:* Die erste Gleichung folgt aus der asymptotischen Darstellung von  $P_{\geq k-1}(z)$  (3.186) und den Transfersätzen 1.2.16 und 1.2.25.

Für die zweite Gleichung nutzen wir die rekursive Formel (3.185) von  $P_{\geq k}(z)$ , die Puiseux-Reihe (3.186) von  $P_{\geq k-1}(z)$ ,  $e^{-x} \sim 1 - x$ , und wir bekommen:

$$\begin{aligned} P_{\geq k}(z) &\sim \rho \exp\left(\sum_{i \geq 2} \frac{P_{\geq k-1}(\rho^i)}{i}\right) e^{P_{\geq k-1}(\rho) - d_{1,k-1}\left(1 - \frac{z}{\rho}\right)^{1/2}} \\ &= \rho \exp\left(\sum_{i \geq 1} \frac{P_{\geq k-1}(\rho^i)}{i}\right) e^{-d_{1,k-1}\left(1 - \frac{z}{\rho}\right)^{1/2}} \\ &= (P_{\geq k}(\rho) + \rho) e^{-d_{1,k-1}\left(1 - \frac{z}{\rho}\right)^{1/2}} \\ &\sim (P_{\geq k}(\rho) + \rho) \left(1 - d_{1,k-1} \left(1 - \frac{z}{\rho}\right)^{1/2}\right). \end{aligned}$$

Transfersatz liefert uns (3.189).  $\square$

Nun können wir einen asymptotischen Ausdruck für den Mittelwert von  $X_n$  bestimmen. Zuerst eine Bemerkung:

**Bemerkung 3.3.31** Sei  $X_n$  die protection number eines zufälligen Pólya-Baumes mit  $n$  Knoten. Dann erfüllt der Mittelwert

$$\begin{aligned}\mathbb{E}(X_n) &= \sum_{k \geq 1} \mathbb{P}(X_n \geq k) \\ &= \sum_{k \geq 1} \frac{[z^n]P_{\geq k}(z)}{[z^n]P(z)} \\ &= \sum_{k \geq 1} \prod_{i=1}^k \frac{[z^n]P_{\geq i}(z)}{[z^n]P_{\geq i-1}(z)} \\ &= \sum_{k \geq 1} \prod_{i=1}^k \mathbb{P}(X_n \geq i | X_n \geq i-1).\end{aligned}$$

Nun formulieren wir unser gewünschtes Ergebnis (vgl. mit [18, Theorem 3])

**Satz 3.3.32** Sei  $X_n$  die protection number eines zufälligen Pólya-Baumes mit  $n$  Knoten. Dann besitzen der Erwartungswert und die Varianz von  $X_n$  folgende asymptotische Darstellungen:

$$\lim_{n \rightarrow \infty} \mathbb{E}(X_n) = \sum_{k \geq 1} \prod_{i=1}^k (P_{\geq i}(\rho) + \rho) \approx 2.154889671973873 \quad (3.190)$$

und

$$\lim_{n \rightarrow \infty} \mathbb{V}(X_n) = \sum_{k \geq 1} (2k-1) \prod_{i=1}^k (P_{\geq i}(\rho) + \rho) - (\mathbb{E}(X_n))^2 \approx 1.369993017502652. \quad (3.191)$$

*Beweis:* Mithilfe von Satz 3.2.30 und Bemerkung 3.2.31 haben wir

$$\begin{aligned}\mathbb{E}(X_n) &= \sum_{k \geq 1} \prod_{i=1}^k \frac{[z^n]P_{\geq i}(z)}{[z^n]P_{\geq i-1}(z)} \\ &= \sum_{k \geq 1} \prod_{i=1}^k \frac{(P_{\geq i}(\rho) + \rho) b_{1,i-1} \rho^{-n}}{2\sqrt{\pi n^{3/2}}} \\ &= \sum_{k \geq 1} \prod_{i=1}^k (P_{\geq i}(\rho) + \rho).\end{aligned}$$

Für die Varianz benötigen wir (1.45) und

$$\mathbb{E}(X_n^2) = \sum_{k \geq 1} (2k-1) \mathbb{P}(X_n \geq k) = \sum_{k \geq 1} (2k-1) \prod_{i=1}^k \mathbb{P}(X_n \geq i | X_n \geq i-1).$$

□

Bezeichne  $\tilde{P}_k(z)$  die erzeugende Funktion für  $k$ -protected Knoten in Polya-Bäume. Im Gegensatz zu einfach erzeugte Bäume können wir wegen der Nichtplanarität der Pólya-Bäume nicht die gleiche Methode wie im vorigen Abschnitt anwenden (siehe [18, Abschnitt 3.2]). Unsere Methode geht wie folgt. Sei  $F(P)$  die Anzahl der  $k$ -protected Knoten in einem zufälligen Pólya-Baum  $P$ . Laut [18, Abschnitt 3.2] ist  $F(P)$  ein additiver Parameter mit Maut-Funktion

$$f(T) = \begin{cases} 1, & \text{wenn } P \text{ mindestens die protection number } k \text{ hat} \\ 0, & \text{sonst} \end{cases}$$

Bezeichne  $P_k(z, u) = \sum_{P \in \mathcal{P}} z^{|P|} u^{F(P)}$  die entsprechende bivariate erzeugende Funktion. Dann ist  $\tilde{P}_k(z)$  durch

$$\tilde{P}_k(z) = \frac{\partial}{\partial u} P_k(z, 1) \quad (3.192)$$

gegeben.

Sei wieder  $\tilde{X}_n$  die mittlere protection number eines zufälligen Knoten in einem Pólya-Baum der Größe  $n$ . Diese definiert wieder eine Zufallsvariable. Unser erster Ziel ist, die Wahrscheinlichkeit  $\mathbb{P}(\tilde{X}_n \geq k)$  zu bestimmen (siehe [18, Formel (16)]).

**Satz 3.3.33** *Die Wahrscheinlichkeit  $\mathbb{P}(\tilde{X}_n \geq k)$  erfüllt folgende Asymptotik*

$$\mathbb{P}(\tilde{X}_n \geq k) \sim \frac{2}{d_1^2} \left( \sum_{i \geq 2} \tilde{P}_k(\rho^i) + P_{\geq k}(\rho) \right) \quad (3.193)$$

wobei  $d_1$  und  $\rho$  wie im Satz 2.2.5 sind.

*Beweis:* Da  $F(P)$  ein additiver Parameter ist, können wir (3.9) anwenden, diese Gleichung nach  $u$  differenzieren und  $u = 1$  setzen. Wegen (3.192), der Definition von  $f(P)$  und einer ähnlichen Argumentation wie im Beweis von Satz 3.1.4 (mit  $H(z) = P_{\geq k}(z)$ ) bekommen wir

$$\tilde{P}_k(z) = \frac{P(z) \sum_{i \geq 2} \tilde{P}_k(z^i) + P_{\geq k}(z)}{1 - P(z)}.$$

Weiters sind, wegen dem Beweis des Satzes 3.1.4, die Funktionen  $P_{\geq k}(z)$  und  $\tilde{P}_k(z^i)$ ,  $i \geq 2$  bei  $z = \rho$  analytisch. Da  $P(z) \sim 1 - d_1 \left(1 - \frac{z}{\rho}\right)^{1/2}$  haben wir

$$\tilde{P}_k(z) \sim \frac{\sum_{i \geq 2} \tilde{P}_k(\rho^i) + P_{\geq k}(\rho)}{d_1 \left(1 - \frac{z}{\rho}\right)^{1/2}}$$

und damit (wegen  $p_n \sim \frac{d_1 \rho^{-n}}{2\sqrt{\pi n^{3/2}}}$ )

$$[z^n] \tilde{P}_k(z) \sim \frac{\left(\sum_{i \geq 2} \tilde{P}_k(\rho^i) + P_{\geq k}(\rho)\right) \rho^{-n}}{d_1 \sqrt{\pi n^{1/2}}}.$$

Daher gilt

$$\mathbb{P}(\tilde{X}_n \geq k) = \frac{[z^n] \tilde{P}_k(z)}{n[z^n] P(z)} \sim \frac{\left(\sum_{i \geq 2} \tilde{P}_k(\rho^i) + P_{\geq k}(\rho)\right) \rho^{-n}}{d_1 \sqrt{\pi n^{1/2}} n \frac{d_1 \rho^{-n}}{2\sqrt{\pi n^{3/2}}}} = \frac{2}{d_1^2} \left(\sum_{i \geq 2} \tilde{P}_k(\rho^i) + P_{\geq k}(\rho)\right).$$

□

Nun können wir eine Formel für den Erwartungswert  $\mathbb{E}(\tilde{X}_n)$  und die Varianz bestimmen (vgl. mit [18, Theorem 4]).

**Satz 3.3.34** Sei  $\tilde{X}_n$  die mittlere protection number eines zufälligen Knoten in einem Polya-Baum der Größe  $n$ . Dann gilt

$$\lim_{n \rightarrow \infty} \mathbb{E}(\tilde{X}_n) = \sum_{k \geq 1} \frac{2}{d_1^2} \left(\sum_{i \geq 2} \tilde{P}_k(\rho^i) + P_{\geq k}(\rho)\right) \approx 0.9953254987 \quad (3.194)$$

und

$$\lim_{n \rightarrow \infty} \mathbb{V}(\tilde{X}_n) \approx 1.3818769746. \quad (3.195)$$

*Beweis:* Wegen des vorigen Satzes haben wir

$$\mathbb{E}(\tilde{X}_n) = \sum_{k \geq 1} \mathbb{P}(\tilde{X}_n \geq k) \sim \sum_{k \geq 1} \frac{2}{d_1^2} \left(\sum_{i \geq 2} \tilde{P}_k(\rho^i) + P_{\geq k}(\rho)\right).$$

Für die Varianz betrachten wir wie üblich (1.45) mit

$$\mathbb{E}(\tilde{X}_n^2) = \sum_{k \geq 1} (2k - 1) \mathbb{P}(\tilde{X}_n \geq k).$$

□

Insgesamt haben wir:

Bäume	$\lim_{n \rightarrow \infty} \mathbb{E}(X_n)$	$\lim_{n \rightarrow \infty} \mathbb{V}(X_n)$
Einfach erzeugt	$\sum_{k \geq 1} \rho^{k-1} \prod_{i=1}^{k-1} \Phi'(S_{\geq i}(\rho))$	$\sum_{k \geq 1} (2k - 1) \rho^{k-1} \prod_{i=1}^{k-1} \Phi'(S_{\geq i}(\rho)) - \left(\lim_{n \rightarrow \infty} \mathbb{E}(X_n)\right)^2$
Catalan	1.622971384715353	0.7156950717833327
Cayley	2.286198316708012	1.598472890455086
Binär 1	3.536472483525321	3.763883442795153
Binär 2	1.562988296151161	0.372985688954940
Motzkin	2.546378248338912	1.679348871220563
Pólya	2.154889671973873	1.369993017502652

### 3.3.3 Der Wiener-Index

Der *Wiener-Index* wurde zum ersten Mal vom Chemiker H. Wiener in 1947 in [36] definiert, um Strukturen von Molekülen zu analysieren. Unser Ziel ist, diesen Parameter mathematisch zu betrachten (Definition 3.3.32, siehe auch [35, Teil 1]) und für einfach erzeugte Bäume und Pólya-Bäume zu analysieren.

**Definition 3.3.35** Sei  $T$  ein Baum mit Knotenmenge  $V(T)$ . Unter dem Wiener-Index  $W(T)$  von  $T$  versteht man die Summe aller Abstände  $d(u, v)$  zwischen zwei Knoten  $u, v \in V(T)$ , also

$$W(T) = \sum_{u, v \in V(T)} d(u, v) = \frac{1}{2} \sum_{u \in V(T)} \sum_{v \in V(T)} d(u, v). \quad (3.196)$$

Ist  $F(T)$  die interne Pfadlänge (vgl. Abschnitt 3.1.3), so kann man den Wiener Index wie folgt rekursiv darstellen (vgl. mit [35, Teil 2]):

$$W(T) = F(T) + \sum_{j=1}^k W(T_j) + \sum_{1 \leq j_1, j_2 \leq k} |T_{j_1}| (F(T_{j_2}) + |T_{j_2}|). \quad (3.197)$$

Wir können Ergebnisse für den Wiener Index bestimmen. Dazu brauchen wir den zusätzlichen Parameter

$$Q(T) = |T|F(T) - W(T). \quad (3.198)$$

(vgl. mit [24, Formel 3.2] oder [35, Formel 4]). Dieser Parameter besitzt folgende alternative Darstellung (siehe [24, Abschnitt 3]):

**Satz 3.3.36** Seien  $T$  ein Wurzelbaum mit Wurzel  $r$  und  $u, v \in V(T)$ . Wir bezeichnen mit  $u \wedge v$  den letzten gemeinsamen Vorgänger von  $u$  und  $v$ . Dann folgt:

$$Q(T) = \sum_{u \in V(T)} \sum_{v \in V(T)} d(r, u \wedge v). \quad (3.199)$$

*Beweis:* Wir haben

$$d(u, v) = d(u, u \wedge v) + d(u \wedge v, v) = d(u, r) - d(u \wedge v, r) + d(v, r) - d(u \wedge v, r).$$

Aufsummieren, (3.196) und (3.198) liefern uns dann (3.199).  $\square$

#### Einfach erzeugte Bäume

Sei  $S_n$  ein einfach erzeugter Baum mit  $n$  Knoten. Das erste Moment des Wiener-Indexes  $\mathbb{E}(W(S_n))$  wurde schon in [12] berechnet. Viel komplizierter ist es, höhere Momente zu erhalten, indem man auch die Parameter  $F(S_n)$  und  $Q(S_n)$  betrachtet. Das allgemeine Ergebnis wurde von Janson in [24, Theorem 3.4] präsentiert. Ziel ist es eine asymptotische Darstellung für  $\mathbb{E}(F(S_n)^k Q(S_n)^l W(S_n)^m)$  zu bestimmen, und damit eine Grenzverteilung für den Wiener-Index zu erhalten. Dafür wenden wir eine Beweisart an, die ähnlich wie in [35] ist, um die Ergebnisse in [24] zu erhalten. Zuerst definieren wir eine multivariate erzeugende Funktion, die von der interne Pfadlänge  $F(S_n)$  und von dem Parameter  $Q(S_n)$  abhängt. Sei erinnert, dass die bivariate erzeugende Funktion  $S(z, u) = \sum_{S \in \mathcal{S}} \omega(S) z^{|S|} u^{F(S)}$  für die interne Pfadlänge die Funktionalgleichung

$$S(z, u) = z\Phi(S(zu, u))$$

erfüllt, wobei  $\Phi(t)$  wie in (2.4) ist (vgl. mit Abschnitt 3.1.3). Allgemeiner haben wir:

**Definition 3.3.37** Sei  $\mathcal{S}$  eine Klasse von einfach erzeugten Bäumen. Bezeichne  $S \in \mathcal{S}$  die Größe bzw. die interne Pfadlänge durch  $|S|$  bzw.  $F(S)$ . Weiters sei  $Q(S)$  der Parameter wie in (3.198). Dann definieren wir folgende multivariate erzeugende Funktion:

$$S(z, u, x, v) := \sum_{S \in \mathcal{S}} \omega(S) z^{|S|} x^{|S|^2} u^{F(S)} v^{Q(S)}. \quad (3.200)$$

Für  $x = 1$  erfüllt diese Funktion

$$S(z, u, 1, v) = z\Phi(S(zu, u, v, v)) \quad (3.201)$$

(vgl. mit [35, Formel (5)]).

Diese Funktion spielt eine wesentliche Rolle, um unsere Resultate zu erhalten. Die Idee ist, folgenden Differentialoperator anzuwenden (vgl. mit [35, Teil 3]):

**Definition 3.3.38** Sei  $F(z, u, x, v)$  eine multivariate erzeugende Funktion. Bezeichne mit  $\Theta_z$  der Operator definiert durch

$$(\Theta_z F)(z, u, x, v) = zF_z(z, u, x, v) \quad (3.202)$$

und analog für  $u, x$  und  $v$ .

Mithilfe dieser Definition angewendet auf  $S(z, 1, 1, 1)$  ist leicht zu erkennen, dass

$$S_{k,l}(z) := (\Theta_u^k \Theta_v^l S)(z, 1, 1, 1) = \sum_{S \in \mathcal{S}} F(S)^k Q(S)^l z^{|S|}. \quad (3.203)$$

Sei noch bemerkt, dass aus (3.200)

$$\Theta_x S(z, u, 1, v) = \Theta_z^2 S(z, u, 1, v) \quad (3.204)$$

folgt. Mithilfe dieses Ausdrucks und der Funktionalgleichung (3.201) gilt:

**Satz 3.3.39** Die Funktionen  $S_{1,0}(z)$  und  $S_{0,1}(z)$  erfüllen

$$S_{1,0}(z) = \frac{(zS'(z))^2}{S(z)} - zS'(z) \quad (3.205)$$

und

$$S_{0,1}(z) = \left( \frac{zS'(z)}{S(z)} - 1 \right) (zS'(z) + z^2S''(z)). \quad (3.206)$$

*Beweis:* Wir zeigen nur (3.205) ((3.206) geht ganz analog). Direkt aus der Gleichung (3.201) zusammen mit (3.202) für  $u$  haben wir

$$\begin{aligned} (\Theta_u S)(z, u, 1, v) &= uS_u(z, u, 1, v) \\ &= uz\Phi'(S(zu, u, v, v))(zS_z(zu, u, v, v) + S_u(zu, u, v, v)) \end{aligned}$$

Setze nun  $u = v = 1$ . Wir bekommen (mit  $S_z(z, 1, 1, 1) = S'(z)$  und  $S_u(z, 1, 1, 1) = S_{1,0}(z)$ )

$$S_{1,0}(z) = z\Phi'(S(z))(zS'(z) + S_{1,0}(z))$$

und daher

$$S_{1,0}(z) = \frac{z\Phi'(S(z))}{1 - z\Phi'(S(z))} zS'(z). \quad (3.207)$$

Aus (2.4) haben wir  $z\Phi'(S(z)) = 1 - \frac{S(z)}{zS'(z)}$ . Einsetzen in (3.207) liefert uns (3.205).  $\square$

Mit ähnlichen Überlegungen wie in [35, Lemma 3] können die Operatoren  $\Theta_u$  und  $\Theta_v$  auf  $S(z, u, 1, v)$  mehrmals anwenden werden, und wir bekommen

$$\begin{aligned} (\Theta_u^k \Theta_v^l S)(z, u, 1, v) &= z \sum_{i=0}^{k-1} \sum_{j=0}^l \binom{k-1}{i} \binom{l}{j} (\Theta_u^i \Theta_v^j (\Phi' \circ S))(zu, u, v, v) \\ &\quad \times ((\Theta_z + \Theta_u)^{k-i} (\Theta_x + \Theta_v)^{l-j} S)(zu, u, v, v) \end{aligned} \quad (3.208)$$

für  $k \geq 1, l \geq 0$ , sowie

$$\begin{aligned} (\Theta_u^k \Theta_v^l S)(z, u, 1, v) &= z \sum_{i=0}^k \sum_{j=0}^{l-1} \binom{k}{i} \binom{l-1}{j} (\Theta_u^i \Theta_v^j (\Phi' \circ S))(zu, u, v, v) \\ &\quad \times ((\Theta_z + \Theta_u)^{k-i} (\Theta_x + \Theta_v)^{l-j} S)(zu, u, v, v) \end{aligned} \quad (3.209)$$

für  $k \geq 0, l \geq 1$ . Aus dieser Darstellungen erhalten wir einen asymptotischen Ausdruck für  $S_{k,l}(z)$ .

**Satz 3.3.40** Die Funktion  $S_{k,l}(z)$  besitzt die gleiche dominante Singularität  $\rho$  wie  $S(z)$ , und es gilt

$$S_{k,l}(z) = d_{k,l} \left(1 - \frac{z}{\rho}\right)^{-(3k+5l-1)/2} + \mathcal{O}\left(\left(1 - \frac{z}{\rho}\right)^{-(3k+5l-2)/2}\right), \quad z \rightarrow \rho, \quad (3.210)$$

wobei die Koeffizienten  $d_{k,l}$  durch folgende Rekursion

$$\begin{aligned} d_{k,l} &= \frac{1}{2\sqrt{2}\sigma} \left( (3k+5l-4)k d_{k-1,l} + \frac{1}{2}(3k+5l-4)(3k+5l-6)l d_{k,l-1} \right. \\ &\quad \left. + \frac{1}{\tau} \sum_{i=0}^k \sum_{j=0}^l \binom{k}{i} \binom{l}{j} d_{i,j} d_{k-i,l-j} \right), \quad k, l \geq 1 \end{aligned} \quad (3.211)$$

mit Anfangswerten  $d_{1,0} = \frac{\tau}{2\sigma^2}$  und  $d_{0,1} = \frac{\tau}{4\sigma^2}$  definiert sind. Dabei sind  $\tau$  und  $\sigma$  sind wie in (2.13) bzw. (2.11).

*Beweis:* Wir zeigen die Aussage für  $k=0, l=1$  und  $k=1, l=0$  (Für höheren  $k, l$  bekommen wir (3.210) und (3.211) aus (3.208)-(3.209) und Induktion. Eine analoge Ausarbeitung für Pólya-Bäume finden wir in [35, Proposition 5]). Im Beweis vom Korollar 2.1.17 haben wir gezeigt, dass  $zS'(z) \sim \frac{d_1}{2} \left(1 - \frac{z}{\rho}\right)^{-1/2} = \frac{\tau}{\sigma\sqrt{2}} \left(1 - \frac{z}{\rho}\right)^{-1/2}$  und  $\frac{zS'(z)}{S(z)} \sim \frac{d_1}{2\tau} \left(1 - \frac{z}{\rho}\right)^{-1/2} = \frac{1}{\sigma\sqrt{2}} \left(1 - \frac{z}{\rho}\right)^{-1/2}$ . Aus (3.205) folgt daher

$$S_{1,0}(z) \sim \frac{\tau}{2\sigma^2} \left(1 - \frac{z}{\rho}\right)^{-1}.$$

Weiters gilt  $z^2 S''(z) \sim \frac{d_1}{4} \left(1 - \frac{z}{\rho}\right)^{-3/2} = \frac{\tau}{2\sqrt{2}\sigma} \left(1 - \frac{z}{\rho}\right)^{-3/2}$  und daher wegen (3.206)

$$S_{0,1}(z) \sim \frac{\tau}{4\sigma^2} \left(1 - \frac{z}{\rho}\right)^{-2}.$$

□

Nun ist es möglich, das asymptotische Moment  $\mathbb{E}(F(S_n)^k Q(S_n)^l)$  zu berechnen.

**Satz 3.3.41** Für  $k, l \geq 1$  haben die Momente  $\mathbb{E}(F(S_n)^k Q(S_n)^l)$  folgende asymptotische Darstellung:

$$\mathbb{E}(F(S_n)^k Q(S_n)^l) = \frac{\sqrt{2\pi}\sigma d_{k,l}}{\Gamma((3k+5l-1)/2)\tau} n^{(3k+5l)/2} \left(1 + \mathcal{O}(n^{-1/2})\right). \quad (3.212)$$

$\tau$  und  $\sigma$  sind wie in (2.13) bzw. (2.11).

*Beweis:* Wegen  $s_n \sim \frac{\tau\rho^{-n}}{\sqrt{2\pi\sigma n^{3/2}}}$  und Transfersatz 1.2.16 haben wir

$$\mathbb{E}(F(S_n)^k Q(S_n)^l) = \frac{[z^n]S_{k,l}}{s_n} \sim \frac{\frac{d_{k,l}n^{(3k+5l-3)/2}\rho^{-n}}{\Gamma((3k+5l-1)/2)}}{\frac{\tau\rho^{-n}}{\sqrt{2\pi\sigma n^{3/2}}}} = \frac{\sqrt{2\pi}\sigma d_{k,l}}{\Gamma((3k+5l-1)/2)\tau} n^{(3k+5l)/2}.$$

□

Wir schreiben die Darstellung (3.212) um. Betrachte die Zahlen  $\omega_{k,l}$  definiert durch  $\omega_{1,0} = \omega_{0,1} = 1$ ,  $\omega_{k,l} = 0$  für  $k < 0$  oder  $l < 0$  und

$$\omega_{k,l} = \frac{2^{(5k+7l-3)/2}\sigma^{k+l+1}}{k!!l!\tau} d_{k,l}, \quad k, l \geq 1.$$

Aus (3.211) folgt

$$\omega_{k,l} = 2(3k+5l-4)\omega_{k-1,l} + 2(3k+5l-4)(3k+5l-6)\omega_{k,l-1} + \sum_{\substack{i=0 \\ 0 < i+j < k+l}}^k \sum_{j=0}^l \omega_{i,j}\omega_{k-i,l-j}. \quad (3.213)$$

Wir erhalten

$$\mathbb{E}(F(S_n)^k Q(S_n)^l) = \frac{k!!l!\sqrt{\pi}\omega_{k,l}}{2^{(5k+7l-4)/2}\Gamma((3k+5l-1)/2)} \left(\frac{1}{\sigma}\right)^{k+l} n^{(3k+5l)/2} \left(1 + \mathcal{O}(n^{-1/2})\right).$$

In [24, Theorem 3.3] hat Janson gezeigt, dass

$$\mathbb{E}(\zeta_1^k \zeta_2^l) = \frac{k!!l!\sqrt{\pi}\omega_{k,l}}{2^{(5k+7l-4)/2}\Gamma((3k+5l-1)/2)} \quad (3.214)$$

wobei  $\zeta_1, \zeta_2$  zwei Zufallsvariablen bezeichnen, die wie folgt definiert sind (siehe [24, Theorem 3.1] und [35, Theorem 1]):

**Definition 3.3.42** Wir definieren  $(\zeta_1, \zeta_2, \zeta)$  als ein Tripel von Zufallsvariablen, sodass

$$\zeta_1 = 2 \int_0^1 e(t)dt, \quad (3.215)$$

$$\zeta_2 = 4 \int \int_{0 < s < t < 1} \min_{s \leq u \leq t} e(u)dsdt, \quad (3.216)$$

$$\zeta = \zeta_1 - \zeta_2 = 2 \int \int_{0 < s < t < 1} (e(s) + e(t) - 2 \min_{s \leq u \leq t} e(u))dsdt \quad (3.217)$$

wobei  $(e(t))_{0 \leq t \leq 1}$  eine Brownsche Exkursion der Dauer 1 bezeichnet.



Wir haben also

$$\mathbb{E}(F(S_n)^k Q(S_n)^l) = \mathbb{E}(\zeta_1^k \zeta_2^l) \left(\frac{1}{\sigma}\right)^{k+l} n^{(3k+5l)/2} (1 + \mathcal{O}(n^{-1/2})). \quad (3.218)$$

Die Momente von  $\zeta$  können mithilfe von  $\zeta = \zeta_1 - \zeta_2$  berechnet werden. Die zwei ersten Werte sind (vgl. mit [24, Theorem 3.3])

$$\begin{aligned} \mathbb{E}(\zeta_1) &= \sqrt{\pi/2} & \mathbb{E}(\zeta_2) &= \sqrt{\pi/8} & \mathbb{E}(\zeta) &= \sqrt{\pi/8} \\ \mathbb{E}(\zeta_1^2) &= 5/3 & \mathbb{E}(\zeta_2^2) &= 7/15 & \mathbb{E}(\zeta^2) &= 2/5. \end{aligned}$$

Damit bekommen wir:

**Satz 3.3.43** *Der mittlere Wiener-Index eines zufälligen einfach erzeugten Baumes  $S_n$  mit  $n$  Knoten erfüllt:*

$$\mathbb{E}(W(S_n)) = \sqrt{\frac{\pi}{8\sigma^2}} n^{5/2} (1 + \mathcal{O}(n^{-1/2})). \quad (3.219)$$

*Beweis:* Aus (3.198) folgt  $W(S_n) = nF(S_n) - Q(S_n)$  und daher wegen (3.218)

$$\begin{aligned} \mathbb{E}(W(S_n)) &= n\mathbb{E}(F(S_n)) - \mathbb{E}(Q(S_n)) \\ &\sim n\mathbb{E}(\zeta_1) \frac{n^{3/2}}{\sigma} - \mathbb{E}(\zeta_2) \frac{n^{5/2}}{\sigma} \\ &= (\mathbb{E}(\zeta_1) - \mathbb{E}(\zeta_2)) \frac{n^{5/2}}{\sigma} \\ &= \mathbb{E}(\zeta) \frac{n^{5/2}}{\sigma} \\ &= \sqrt{\frac{\pi}{8\sigma^2}} n^{5/2}. \end{aligned}$$

□

Allgemeiner haben wir (vgl. mit [24, Theorem 3.4])

$$\mathbb{E}(F(S_n)^k Q(S_n)^l W(S_n)^m) = \mathbb{E}(\zeta_1^k \zeta_2^l \zeta^m) \left(\frac{1}{\sigma}\right)^{k+l+m} n^{(3k+5l+5m)/2} (1 + \mathcal{O}(n^{-1/2})). \quad (3.220)$$

Daher gilt für die Varianz von  $W(S_n)$ :

**Satz 3.3.44** *Die Varianz des Wiener-Indexes eines zufällig einfach erzeugten Baumes erfüllt folgende Asymptotik:*

$$\mathbb{V}(W(S_n)) = \frac{16 - 5\pi}{40\sigma^2} n^5 (1 + \mathcal{O}(n^{-1/2})). \quad (3.221)$$

*Beweis:* Aus (3.220) und (3.219) haben wir zusammen mit den zwei ersten Momenten von  $\zeta$

$$\begin{aligned} \mathbb{V}(W(S_n)) &= \mathbb{E}(W(S_n)^2) - (\mathbb{E}(W(S_n)))^2 \\ &\sim \frac{\mathbb{E}(\zeta^2)}{\sigma^2} n^5 - \frac{\pi}{8\sigma^2} n^5 \\ &= \left(\frac{2}{5\sigma^2} - \frac{\pi}{8\sigma^2}\right) n^5 \\ &= \frac{16 - 5\pi}{40\sigma^2} n^5. \end{aligned}$$

□

Zum Schluss bestimmen wir eine Grenzverteilung für den Wiener-Index. Genauer

**Satz 3.3.45** Seien  $S_n$  und  $\sigma$  wie im vorigen Satz. Der normalisierte Wiener-Index  $n^{-5/2}W(S_n)$  erfüllt dann folgenden schwachen Grenzwertsatz:

$$n^{-5/2}W(S_n) \xrightarrow{d} \sigma^{-1}\zeta. \quad (3.222)$$

*Beweis:* Die Idee ist, die Tiefensuche  $X_n(t), 0 \leq t \leq 2n$  und die Contour  $C_n(t) = \frac{1}{\sqrt{n}}X_n(2nt), 0 \leq 1$  von  $S_n$  (siehe Definition 3.2.2) anzuwenden. Betrachte  $F(S_n)$ . In [24, Teil 4] wird gezeigt, dass

$$\frac{1}{n^{3/2}}F(S_n) = \frac{1}{2n^{3/2}} \int_0^{2n} X_n(x)dx + \mathcal{O}(n^{-1/2}) = \frac{1}{n^{1/2}} \int_0^1 X_n(2nt)dt + \mathcal{O}(1) = \int_0^1 C_n(t)dt + \mathcal{O}(1).$$

Das Integral bildet ein stetiges lineares Funktional. Daher können wir Satz 1.3.33 auf (3.133) anwenden und erhalten mit (3.215)

$$\frac{1}{n^{3/2}}F(S_n) \xrightarrow{d} \int_0^1 \frac{2e(t)}{\sigma} dt = \frac{\zeta_1}{\sigma}.$$

Betrachte nun  $Q(S_n)$ . Aus [24, Teil 4] haben wir

$$\begin{aligned} \frac{1}{n^{5/2}}Q(S_n) &= \frac{1}{4n^{5/2}} \int_0^{2n} \int_0^{2n} \min_{x_1 \leq x \leq x_2} X_n(x) dx_1 dx_2 + \mathcal{O}(n^{-1/2}) \\ &= \int_0^1 \int_0^1 \min_{s \leq u \leq t} C_n(u) ds dt + \mathcal{O}(1) \\ &= 2 \int_0^1 \int_{0 < s < t < 1} \min_{s \leq u \leq t} C_n(u) ds dt + \mathcal{O}(1). \end{aligned}$$

Das Minimum ist wieder ein stetiges Funktional. Daraus folgt (mit Satz 1.3.33 angewendet auf (3.133) und (3.216))

$$\frac{1}{n^{5/2}}Q(S_n) \xrightarrow{d} 2 \int_0^1 \int_{0 < s < t < 1} \min_{s \leq u \leq t} \frac{2e(u)}{\sigma} ds dt = \frac{\zeta_2}{\sigma}.$$

(3.198) und (3.217) liefern uns das gewünschte Ergebnis. . □

## Pólya-Bäume

Sei  $P_n$  ein Pólya-Baum mit  $n$  Knoten. Die entsprechenden Resultate wurden in [35] gezeigt. Wir führen eine ähnliche Beweisart wie im vorigen Teil, um einen Ausdruck für die Momente  $\mathbb{E}(F(P_n)^k Q(P_n)^l W(P_n)^m)$  zu berechnen, und eine Grenzverteilung zu erhalten. Analog wie im vorigen Abschnitt definieren wir eine multivariate erzeugende Funktion, die von der Größe  $|P_n|$ , der internen Pfadlänge  $F(P_n)$  und  $Q(P_n)$  abhängt (vgl. mit [35, Teil 2]).

**Definition 3.3.46** Die multivariate erzeugende Funktion  $P(z, u, x, v) = \sum_{P \in \mathcal{P}} z^{|P|} x^{|P|^2} u^{F(P)} v^{Q(P)}$  erfüllt folgende Funktionalgleichung:

$$P(z, u, 1, v) = z \exp \left( \sum_{j \geq 1} \frac{1}{j} P((zu)^j, u^j, v^j, v^j) \right). \quad (3.223)$$

Unser erster Ziel ist es, eine asymptotische Darstellung für das Moment  $\mathbb{E}(F(P_n)^k Q(P_n)^l)$  zu bestimmen. Man sieht leicht, dass

$$P_{k,l}(z) := (\Theta_u^k \Theta_v^l P)(z, 1, 1, 1) = \sum_{P \in \mathcal{P}} F(P)^k Q(P)^l z^{|P|}. \quad (3.224)$$

Wobei der  $\Theta_u$  und  $\Theta_v$  die im vorigen Abschnitt definierten Operatoren sind.

**Bemerkung 3.3.47** Wegen  $P(z, u, x, v) = \sum_{P \in \mathcal{P}} z^{|P|} x^{2|P|} u^{F(P)} v^{Q(P)}$  gilt  $\Theta_x P(z, u, 1, v) = \Theta_z^2 P(z, u, 1, v)$ .

**Satz 3.3.48** Es gilt

$$P_{1,0}(z) = P(z) \sum_{j \geq 1} (z^j P'(z^j) + P_{1,0}(z^j)) \quad (3.225)$$

und

$$P_{0,1}(z) = P(z) \sum_{j \geq 1} (z^j P'(z^j) + z^{2j} P''(z^j) + P_{0,1}(z^j)). \quad (3.226)$$

*Beweis:* Wir wenden den Operator  $\Theta_u$  an. Mithilfe von (3.223) haben wir

$$\begin{aligned} & \Theta_u P(z, u, 1, v) \\ &= u P_u(z, u, 1, v) \\ &= uz \exp\left(\sum_{j \geq 1} \frac{1}{j} P(z^j u^j, u^j, v^j, v^j)\right) \sum_{j \geq 1} (u^{j-1} z^j P_z(z^j u^j, u^j, v^j, v^j) + u^{j-1} P_u(z^j u^j, u^j, v^j, v^j)) \\ &= z P(z, u, 1, v) \sum_{j \geq 1} (u^j z^j P_z(z^j u^j, u^j, v^j, v^j) + u^j P_u(z^j u^j, u^j, v^j, v^j)) \\ &= z P(z, u, 1, v) \sum_{j \geq 1} ((\Theta_z P)(z^j u^j, u^j, v^j, v^j) + (\Theta_u P)(z^j u^j, u^j, v^j, v^j)). \end{aligned}$$

Setzt man  $u = v = 1$  so erhalten wir (3.225). Mit Anwendung des Operators  $\Theta_v$  auf  $P(z, u, 1, v)$ , Bemerkung 3.3.47 und  $u = v = 1$  bekommen wir (3.226).  $\square$

Allgemeiner haben wir (siehe Lemma 3 in [35])

$$\begin{aligned} (\Theta_u^k \Theta_v^l P)(z, u, 1, v) &= \sum_{i=0}^{k-1} \sum_{h=0}^l \binom{k-1}{i} \binom{l}{h} (\Theta_u^i \Theta_v^h P)(z, u, 1, v) \sum_{j \geq 1} j^{k+l+i+h-1} \\ &\quad \times ((\Theta_z + \Theta_u)^{k-i} (\Theta_x + \Theta_v)^{l-h} P)(z^j) \end{aligned} \quad (3.227)$$

für  $k \geq 1, l \geq 0$  und

$$\begin{aligned} (\Theta_u^k \Theta_v^l P)(z, u, 1, v) &= \sum_{i=0}^k \sum_{h=0}^{l-1} \binom{k}{i} \binom{l-1}{h} (\Theta_u^i \Theta_v^h P)(z, u, 1, v) \sum_{j \geq 1} j^{k+l+i+h-1} \\ &\quad \times ((\Theta_z + \Theta_u)^{k-i} (\Theta_x + \Theta_v)^{l-h} P)(z^j) \end{aligned} \quad (3.228)$$

für  $k \geq 0, l \geq 1$ . Nun können wir eine asymptotische Darstellung für  $P_{k,l}(z)$  ermitteln (siehe auch [35, Proposition 5]):

**Satz 3.3.49** Für  $k, l \geq 0$  und  $k + l \geq 1$  besitzt die Funktion  $P_{k,l}(z)$  folgende Puiseux-Reihe:

$$P_{k,l}(z) = d_{k,l} \left(1 - \frac{z}{\rho}\right)^{-(3k+5l-1)/2} + \mathcal{O}\left(\left(1 - \frac{z}{\rho}\right)^{-(3k+5l-2)/2}\right) \quad (3.229)$$

mit Koeffizienten  $d_{1,0} = \frac{1}{2}$ ,  $d_{0,1} = \frac{1}{4}$ , und

$$d_{k,l} = \frac{1}{d_1} \left( \frac{3k+5l-4}{2} k d_{k-1,l} + \frac{(3k+5l-4)(3k+5l-6)}{4} l d_{k,l-1} + \frac{1}{2} \sum_{\substack{i=0 \\ 0 < i+h < k+l}}^k \sum_{h=0}^l \binom{k}{i} \binom{l}{h} d_{i,h} d_{k-i,l-h} \right). \quad (3.230)$$

$d_1$  ist wie im Satz 2.2.5.

*Beweis:* Analog wie im Satz 3.3.40, per Induktion. Siehe [35, Proposition 5].

Daher bekommen wir:

**Satz 3.3.50** Die Momente  $\mathbb{E}(F(P_n)^k Q(P_n)^l)$  haben folgende asymptotische Darstellung:

$$\mathbb{E}(F(P_n)^k Q(P_n)^l) = \mathbb{E}(\zeta_1^k \zeta_2^l) \left(\frac{\sqrt{2}}{d_1}\right)^{k+l} n^{(3k+5l)/2} (1 + \mathcal{O}(n^{-1/2})), \quad (3.231)$$

wobei  $\zeta_1, \zeta_2$  wie in Definition 3.3.42 sind.  $d_1$  ist wie im Satz 2.2.5.

*Beweis:* Aus (3.231) und Transfersätzen 1.2.16 und 1.2.25 gilt

$$\begin{aligned} [z^n]P_{k,l}(z) &= \frac{d_{k,l} \rho^{-n} n^{(3k+5l-1)/2-1}}{\Gamma((3k+5l-1)/2)} + \mathcal{O}\left(\frac{\rho^{-n} n^{(3k+5l-2)/2-1}}{\Gamma((3k+5l-2)/2)}\right) \\ &= \frac{d_{k,l} \rho^{-n} n^{(3k+5l-3)/2}}{\Gamma((3k+5l-1)/2)} (1 + \mathcal{O}(n^{-1/2})). \end{aligned}$$

Insgesamt haben wir (mit  $p_n \sim \frac{d_1 \rho^{-n}}{2\sqrt{\pi} n^{3/2}}$ )

$$\begin{aligned} \mathbb{E}(F(P_n)^k Q(P_n)^l) &= \frac{[z^n]P_{k,l}(z)}{p_n} \\ &= \frac{\frac{d_{k,l} \rho^{-n} n^{(3k+5l-3)/2}}{\Gamma((3k+5l-1)/2)}}{\frac{d_1 \rho^{-n}}{2\sqrt{\pi} n^{3/2}}} (1 + \mathcal{O}(n^{-1/2})) \\ &= \frac{2\sqrt{\pi} d_{k,l}}{d_1 \Gamma((3k+5l-1)/2)} n^{(3k+5l)/2} (1 + \mathcal{O}(n^{-1/2})). \end{aligned}$$

Setze nun (vgl. [35, Abschnitt 3])

$$\omega_{k,l} = \frac{d_1^{k+l-1} 2^{2k+3l-1}}{k!l!} d_{k,l}. \quad (3.232)$$

Aus (3.230) folgt, dass die Zahlen  $\omega_{k,l}$  dieselbe Darstellung wie in (3.213) besitzen (insbesondere  $\omega_{1,0} = \omega_{0,1} = 1$  und  $\omega_{k,l} = 0$  für  $k < 0$  oder  $l < 0$ ; vgl. mit [35, Formel 15]). Wegen (3.214) (siehe auch [24, Theorem 3.3]) erhalten wir

$$\begin{aligned}\mathbb{E}(F(P_n)^k Q(P_n)^l) &= \frac{2\sqrt{\pi}k!l!\omega_{k,l}}{d_1^{k+l}2^{2k+3l-1}\Gamma((3k+5l-1)/2)}n^{(3k+5l)/2}(1 + \mathcal{O}(n^{-1/2})) \\ &= \frac{\sqrt{\pi}k!l!\omega_{k,l}}{2^{(5k+7l-4)/2}\Gamma((3k+5l-1)/2)}\left(\frac{\sqrt{2}}{d_1}\right)^{k+l}n^{(3k+5l)/2}(1 + \mathcal{O}(n^{-1/2})) \\ &= \mathbb{E}(\zeta_1^k \zeta_2^l)\left(\frac{\sqrt{2}}{d_1}\right)^{k+l}n^{(3k+5l)/2}(1 + \mathcal{O}(n^{-1/2})).\end{aligned}$$

□

Nun haben wir folgenden Satz:

**Satz 3.3.51** *Der Erwartungswert für den Wiener-Index eines zufälligen Pólya-Baumes der Größe  $n$  ist gegeben durch*

$$\mathbb{E}(W(P_n)) = \frac{\sqrt{\pi}}{2d_1}n^{5/2}(1 + \mathcal{O}(n^{-1/2})) \quad (3.233)$$

wobei  $\zeta = \zeta_1 - \zeta_2$ .

*Beweis:* Genauso wie im Satz 3.3.43 mithilfe von (3.231). □

Allgemeiner gilt

$$\mathbb{E}(F(P_n)^k Q(P_n)^l W(S_n)^m) = \mathbb{E}(\zeta_1^k \zeta_2^l \zeta^m)\left(\frac{\sqrt{2}}{d_1}\right)^{k+l+m}n^{(3k+5l+5m)/2}(1 + \mathcal{O}(n^{-1/2})) \quad (3.234)$$

(vgl. mit [35, Teil 3]). Daher können wir die Varianz von  $W(P_n)$  bestimmen:

**Satz 3.3.52** *Die Varianz des Wiener-Indexes eines zufälligen Pólya-Baumes lautet*

$$\mathbb{V}(W(P_n)) = \frac{16 - 5\pi}{20d_1^2}n^5(1 + \mathcal{O}(n^{-1/2})). \quad (3.235)$$

*Beweis:* genauso wie im Satz 3.3.44 mithilfe der Formeln (3.234) und (3.233). □

Für die Grenzverteilung gilt folgende Aussage:

**Satz 3.3.53** *Der normalisierte Wiener-Index eines zufälligen Pólya-Baumes mit  $n$  Knoten erfüllt folgenden schwachen Grenzwertsatz:*

$$n^{-5/2}W(P_n) \xrightarrow{d} \frac{\sqrt{2}}{d_1}\zeta. \quad (3.236)$$

Die Beweisidee ist dieselbe wie in [24, Theorem 3.1] für einfach erzeugte Bäume. Insgesamt haben wir für den Wiener-Index:

Bäume	Mittelwert	Varianz	Konvergenz
Einfach erzeugt	$\sim \sqrt{\frac{\pi}{8\sigma^2}} n^{5/2}$	$\sim \frac{16-5\pi}{40\sigma^2} n^5$	$n^{-5/2}W(S_n) \xrightarrow{d} \frac{1}{\sqrt{\sigma}}\zeta$
Catalan	$\sim \frac{\sqrt{\pi}}{4} n^{5/2}$	$\sim \frac{16-5\pi}{80} n^5$	$n^{-5/2}W(S_n) \xrightarrow{d} \frac{1}{\sqrt{2}}\zeta$
Cayley	$\sim \sqrt{\frac{\pi}{8}} n^{5/2}$	$\sim \frac{16-5\pi}{40} n^5$	$n^{-5/2}W(S_n) \xrightarrow{d} \zeta$
Binär 1	$\sim \sqrt{\frac{\pi}{8}} n^{5/2}$	$\sim \frac{16-5\pi}{40} n^5$	$n^{-5/2}W(S_n) \xrightarrow{d} \zeta$
Binär 2	$\sim \frac{\sqrt{\pi}}{2} n^{5/2}$	$\sim \frac{16-5\pi}{20} n^5$	$n^{-5/2}W(S_n) \xrightarrow{d} \sqrt{2}\zeta$
Motzkin	$\sim \frac{\sqrt{3\pi}}{4} n^{5/2}$	$\sim \frac{48-15\pi}{80} n^5$	$n^{-5/2}W(S_n) \xrightarrow{d} \sqrt{\frac{3}{2}}\zeta$
Pólya	$\sim \frac{\sqrt{\pi}}{2d_1} n^{5/2}$	$\sim \frac{16-5\pi}{20d_1^2} n^5$	$n^{-5/2}W(P_n) \xrightarrow{d} \frac{\sqrt{2}}{d_1}\zeta$

# Literaturverzeichnis

- [1] D. Aldous: *The continuum random tree. I*, Ann. Probab., (1991), 1-28, **19**(1).
- [2] D. Aldous: *The continuum random tree. II*, An overview. Stochastic analysis (Durham, 1990), 23-70, London Math. Soc. Lecture Note Ser. **167**, Cambridge Univ. Press, Cambridge, 1991.
- [3] E. A. Bender und L. B. Richmond: *Central and local limit theorems applied to asymptotic enumeration II: multivariate generating functions*, J. Combin. Theory Ser. A **34** (1983), 255-265.
- [4] P. Billingsley: *Convergence of Probability Measures. Second edition*, Wiley Series in Probability and Statistics: Probability and Statistics. A Wiley-Interscience Publication. John Wiley Sons Inc., New York, 1999. x+277 Seiten. ISBN: 0-471-19745-9
- [5] P. Billingsley: *Probability and Measure*, Wiley Series in Probability and Statistics. John Wiley Sons, Inc., Hoboken, NJ, 2012. xviii+624 Seiten. ISBN: 978-1-118-12237-2
- [6] J. W. Cohen, G. Hooghiemstra: *Brownian excursion, the M/M/1 queue and their occupation times*, Math. Oper. Res. **6** (1981), no. 4, 608-629.
- [7] L. Devroye: *A note on the probability of cutting a Galton-Watson tree*, Electron. J. Probab. **16**: 2001-2019, 2011.
- [8] M. Drmota: *Random Trees*, Springer, Wien New York, Vienna, 2009. xviii+458 Seiten. ISBN: 978-3-211-75355-2.
- [9] M. Drmota, B. Gittenberger: *On the profile of random trees*, Random Structures Algorithms **10** (1997), no. 4, 421-451.
- [10] M. Drmota, B. Gittenberger: *The distribution of nodes of given degree in random trees*, J. Graph Theory **31** (1999), no. 3, 227-253.
- [11] M. Drmota, B. Gittenberger: *The shape of unlabeled rooted random trees*, European J. Combin. **31** (2010), no. 8, 2028-2063.

- [12] R. C. Entringer, A. Meir, J. W. Moon, L. A. Székely: *The Wiener index of trees from certain families*, Australas. J. Combin. **10** (1994), 211-224.
- [13] P. Flajolet, A. Odlyzko: *Singularity analysis of generating functions*, SIAM J. Discrete Math. **3** (1990), no. 2, 216-240.
- [14] P. Flajolet, A. Odlyzko: *The average height of binary trees and other simple trees*, J. Comput. System Sci. **25** (1982), no. 2, 171-213.
- [15] P. Flajolet, R. Sedgewick: *Analytic Combinatorics*, Cambridge University Press, Cambridge, 2009. xiv+810 Seiten. ISBN: 978-0-521-89806-5.
- [16] A. Genitrini: *Full asymptotic expansion for Pólya structures*, Proceedings of the 27th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms?AofA'16, 12 Seiten, Jagiellonian Univ., Dep. Theor. Comput. Sci., Krak016.
- [17] B. Gittenberger: *Nodes of large degree in random trees and forests*, Random Structures Algorithms **28** (2006), no. 3, 374-385.
- [18] B. Gittenberger, Z. Golebiewski, I. Larcher, M. Sulkowska: *Protection number in simply generated trees and Pólya trees*, Vorabdruck.
- [19] B. Gittenberger, V. Kraus: *On the number of transversals in random trees*, 23rd Intern. Meeting on Probabilistic, Combinatorial, and Asymptotic Methods for the Analysis of Algorithms (AofA'12), 141-154, Discrete Math. Theor. Comput. Sci. Proc., AQ, Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2012.
- [20] B. Gittenberger, E. Yu Jin, M. Wallner: *On the shape of random Pólya structures*, Discrete Math. **341** (2018), no. 4, 896-911.
- [21] K. Grill: *Theorie stochastischer Prozesse*, Skript, Institut für Statistik und Wahrscheinlichkeitstheorie, TU Wien.
- [22] G. Hooghiemstra: *On the explicit form of the density of Brownian excursion local time*, Proc. Amer. Math. Soc. **84** (1982), no. 1, 127-130.
- [23] S. Janson: *Simply generated trees, conditioned Galton-Watson trees, random allocations and condensation*, Probab. Surv. **9** (2012), 103-252.
- [24] S. Janson: *The Wiener index of simply generated random trees*, Random Structures Algorithms **22** (2003), no. 4, 337-358.



- [25] A. Meir, J. W. Moon: *On the altitude of nodes in random trees*, Canadian J. Math. **30** (1978), no. 5, 997-1015.
- [26] A. Meir, J. W. Moon: *On the log-product of the subtree-sizes of random trees*, Random Structures Algorithms **12** (1998), no. 2, 197-212.
- [27] R. Otter: *The number of trees*, Ann. of Math. (2) **49** (1948), 583-599.
- [28] A. Panagiotou, B. Stuffer: *Scaling limits of random Pólya trees*, Probab. Theory Related Fields **170** (2018), no. 3-4, 801-820.
- [29] A. Panholzer: *The climbing depth of random trees*, Random Structures Algorithms **26** (2005), no. 1-2, 84-109.
- [30] G. Pólya: *Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und chemische Verbindungen*, Acta Math. **68** (1937), no. 1, 145-254.
- [31] R. W. Robinson, A. Schwenk: *The distribution of degrees in a large random tree*, Discr. Math. **12** (1975), 359-372
- [32] A. Schwenk: *An asymptotic evaluation of the cycle index of a symmetric group*, Discrete Math. **18** (1977), no. 1, 71-78.
- [33] A. W. van der Vaart: *Asymptotic Statistics*, New York: Cambridge University Press (1998). ISBN: 0-521-49603-9.
- [34] S. Wagner: *Central limit theorems for additive tree parameters with small toll functions*, Combin. Probab. Comput. **24** (2015), no. 1, 329-353.
- [35] S. Wagner: *On the Wiener index of random trees*, Discrete Math. **312** (2012), no. 9, 1502-1511.
- [36] H. Wiener: *Structural determination of paraffin boiling points*, J. Amer. Chem. Soc. **69** (1947), 17-20.
- [37] H. Woracek: *Komplexe Analysis*, SS 2015, Skript, Institut für Analysis und Scientific Computing, TU Wien.