

DIPLOMARBEIT

Kombinatorische Abzählprobleme bei markierten k-Bäumen

zur Erlangung des akademischen Grades

Diplom-Ingenieur

im Rahmen des Studiums

Technische Mathematik

eingereicht von

Patrick Reiner

Matrikelnummer 01225756

ausgeführt am Institut für Diskrete Mathematik und Geometrie
der Fakultät für Mathematik und Geoinformation der Technischen Universität Wien

Betreuung

Betreuer/in: Ao.Univ.Prof. Dipl.-Ing. Dr.techn. Alois Panholzer

Wien, 22.03.2018

(Unterschrift Verfasser/in)

(Unterschrift Betreuer/in)

Abstract

Die k -Bäume sind in der Graphentheorie eine Verallgemeinerung der Bäume. Durch ihren rekursiven Aufbau sind die k -Bäume interessant als Modelle für komplexe Netzwerke, welche in verschiedenen Formen vielerorts in der Wissenschaft auftreten. In der Arbeit wird für verschiedene Arten von k -Bäumen das kombinatorische Abzählungsproblem gelöst. Mit rekursiven Konstruktionen, welche nach verschiedenen probabilistischen Regeln ablaufen, werden Wahrscheinlichkeitsverteilungen auf den k -Bäumen erzeugt. Diese Verteilungen können zum Teil kombinatorisch interpretiert werden und in Zusammenhang mit bestimmten Typen von k -Bäumen gebracht werden. Ausgestattet mit einem Wachstum dieses Typs werden für die k -Bäume die Verteilungen für den Knotengrad, die Nachfahren und die Vorfahren bestimmt. Diese Größen werden für einen beliebigen Knoten auch auf ihre Grenzverteilung untersucht.

In graph theory k -trees are a generalisation of ordinary trees. As recursive defined structures they can be seen as complex network models. Complex networks play an important role in different parts of science. In this thesis we will count various types of k -trees. These graphs can be generated with probabilistic growth rules. Distinct forms of k -trees can be identified with some growth rules. Formulas will be derived for the degree, the descendants and the ancestors. In particular we will study for each model the exact distribution of a node and calculate the limiting distribution of a random node.

Vorwort

Die vorliegende Diplomarbeit habe ich im Zuge meines Studiums Technische Mathematik an der Technischen Universität Wien verfasst. Für die Mathematik habe ich schon seit der Schule ein großes Interesse. Das Studium hat mir diese Naturwissenschaft in neuem Licht erscheinen lassen und neue Aspekte und Wissensgebiete beleuchtet.

Im Zuge dieses Studiums durfte ich in der Vorlesung über Diskrete Methoden die erzeugenden Funktionen kennen lernen. Ich finde es faszinierend, dass man mit diesem technischen Instrument kombinatorische Probleme durch ihre „Strukturgleichungen“ einfach beschreiben kann. Die entsprechenden Gleichungen für die Potenzreihen können dann mit den verschiedensten Methoden gelöst werden und auf diese Weise wird auch Wissen über Differentialgleichungen oder Komplexe Analysis verwendet, um reine kombinatorische Probleme zu lösen. Mit Interesse an dieser Thematik hat mich mein Betreuer Alois Panholzer zu den k -Bäumen geführt. Diese sind aufgrund ihrer baumartigen Struktur prädestiniert für eine solche Analyse.

An dieser Stelle möchte ich mich auch noch bei meinem Betreuer Alois Panholzer bedanken. Aufkommende Fragen konnten immer schnell und unkompliziert geklärt werden. Das Arbeitsklima war entgegenkommend und die Abwicklung von diversen organisatorischen Dingen war immer sehr schnell und einfach erledigt. Danke für die Unterstützung und nützlichen Anmerkungen bei dieser Arbeit!

Durch das Studium haben mich auch viele Studienkollegen und Freunde begleitet, deren Aufzählung hier viel zu lange wäre. Mit ihrer Hilfe konnten auch die anstrengenden Phasen des Studiums immer mit angenehmen Auszeiten bewältigt werden. Die mathematischen Diskussionen und Gespräche, welche je nach Motivation auch vom Thema abgedriftet sind, gehörten bei der Vorbereitung für die nächsten Übungen zum Alltag und haben das Studium bereichert. Diese haben uns manche Nachmittage in den Lernräumen im Freihaus und auch mit entsprechendem Equipment anderenorts beschäftigt.

Zu guter Letzt möchte ich noch meiner Familie Danke sagen. Sie hat mich immer in meinem Weg vorbehaltlos unterstützt und ihre Hilfe hat mir die Möglichkeit gegeben, dieses Studium zu absolvieren.

Patrick Reiner
Wien, März 2018

Inhaltsverzeichnis

1 Einleitung und Grundlagen	7
1.1 Graphentheorie	7
1.2 k -Bäume	9
1.2.1 k -Bäume als Graphen	11
1.3 Kombinatorische Strukturen und erzeugende Funktionen	12
1.3.1 Nützliche Potenzreihen und formale Topologie	16
1.3.2 Lagrange'sche Inversionsformel und weitere Sätze	17
1.4 Asymptotisches Verhalten und Hilfsresultate	18
2 Abzählprobleme bei k-Bäumen	26
2.1 Markierte k -Bäume	26
2.1.1 Weitere Beweise	29
2.2 Markierte k -Wurzelbäume	29
2.3 Ansteigende k -Bäume	30
2.4 Geordnete Wurzelbäume	32
2.5 Geordnete ansteigende k -Bäume	34
2.6 Markierte d -äre k -Bäume	36
2.7 Ansteigende d -äre k -Bäume	38
2.8 (k, m) -Bäume	40
2.9 Markierte k -gonale 2-Bäume	44
2.9.1 Fall: k ungerade	48
2.9.2 Fall: k gerade	50
3 Weitere Betrachtung von k-Bäumen	52
3.1 Randomisierte Modelle	52
3.2 Knotengrad	54
3.2.1 Geordnete, ansteigende k -Bäume	55
3.2.2 Ansteigende k -Bäume	63
3.2.3 Ansteigende d -äre k -Bäume	69
3.2.4 CU-Modell	75
3.3 Lokaler Clusterkoeffizient	79
3.4 Nachfahren	81
3.4.1 Geordnete ansteigende k -Bäume	81
3.4.2 Ansteigende k -Bäume	84
3.4.3 Ansteigende d -äre k -Bäume	86
3.4.4 CU-Modell	88
3.5 Vorfahren	90
3.5.1 Geordnete ansteigende k -Bäume	91
3.5.2 Ansteigende k -Bäume	94
3.5.3 Ansteigende d -äre k -Bäume	95
3.5.4 CU-Modell	97
3.6 Vergleich der Modelle	100
Literaturverzeichnis	101

A Resultate	103
A.1 Abzählprobleme	103
A.2 Knotengrad	104
A.3 Lokaler Clusterkoeffizient	105
A.4 Nachfahren	105
A.5 Vorfahren	106

Abbildungsverzeichnis

1 Beispiele für Bäume	8
2 Beispiele für k -Bäume	9
3 Beispiel für einen markierten k -Baum	10
4 Beispiel für einen chordalen Graphen	11
5 Beispiel für einen ansteigenden k -Baum	31
6 Beispiel für einen geordneten k -Wurzelbaum	33
7 Beispiel für einen geordneten ansteigenden k -Baum	34
8 Beispiel für einen d -ären k -Baum	37
9 Beispiel für einen ansteigenden d -ären k -Baum	39
10 Beispiel für einen (k, m) -Baum	41
11 Weiteres Beispiel für einen (k, m) -Baum	41
12 Beispiele für k -gonale 2-Bäume	45
13 Veranschaulichung des Beweises von Lemma 2.7	46
14 Anschauliche Darstellung der Abstandsfunktion d für k -gonale 2-Bäume	49
15 Darstellung von Lemma 2.11(v)	50
16 k -gonalen 2-Bäume als Wurzelbäume	51
17 Vergleich randomisierte Modelle	54
18 Knotengrad	55
19 Nachfahren	81
20 Vorfahren	91

1 Einleitung und Grundlagen

Komplexe Netzwerke treten in vielen verschiedenen Disziplinen der Wissenschaft in unterschiedlicher Form auf. So gibt es technische, physische Netzwerke wie beispielsweise die Stromnetze und das Internet. Die Webseiten bilden hingegen mit den Links als Verbindung ein virtuelles Netzwerk. Die sozialen Medien und Facebook im speziellen bilden Beispiele für Netzwerke in der Soziologie. In der Biologie sind sie auch zu finden, unter anderem bei der Untersuchung der Ausbreitung einer Krankheit. Die Analyse von Netzwerken ist daher von großer Bedeutung (vgl. [25], S. 1). Wie sich herausgestellt hat, haben viele reale Netzwerke typische Eigenschaften. Diese umfassen einen hohen lokalen Clusterkoeffizienten und die Eigenschaft, dass zwei Knoten über einen relativ kurzen Pfad verbunden sind. Des Weiteren erfüllt der Grad eines Knotens ein Potenzgesetz (vgl. [25], S. 1 & [6] Abschnitt 1). Die Modellierung solcher Netzwerke ist unter anderem möglich, indem man Graphen rekursiv aufbaut. Das Hinzufügen von Knoten geschieht dabei nach einer wahrscheinlichkeitsbasierten Regel. Die k -Bäume, welche in dieser Arbeit betrachtet werden, können wie auch die bekannten Bäume der Graphentheorie definitionsgemäß rekursiv aufgebaut werden. In Kapitel 3 werden verschiedene probabilistische Regeln für das Aufbauen eines k -Baumes betrachtet und unter anderem ein Potenzgesetz für den Knotengrad hergeleitet. Der Exponent dieses Gesetzes liegt jeweils zwischen zwei und drei, das sind für reale Netzwerke typische Werte (vgl. [2] & [19], Abschnitt 1). Davor wird in Kapitel 2 für verschiedene Modelle von k -Bäumen gezählt, wie viele es abhängig von der Knotenanzahl gibt. Die wichtigsten Ergebnisse dieser zwei Kapitel sind am Ende der Arbeit übersichtlich aufgelistet.

In diesem Kapitel werden grundlegende Definitionen der Graphentheorie angegeben. Dabei liegt das Hauptaugenmerk auf der Einführung der Notation und der Sprechweise. Die Auflistung verzichtet auf die säuberliche Definition aller verwendeter Begriffe. Wie zum Beispiel „Teilgraph“ sind viele Begriffe selbsterklärend, beziehungsweise werden auch nur elementare Bezeichnungen, die in jedem Werk zur Graphentheorie (z. B. [7]) zu finden sind, verwendet. Abschnitt 1.3 widmet sich den kombinatorischen Strukturen und erzeugenden Funktionen, welche in der Arbeit verwendet werden, um Rekursionen zu lösen. Im letzten Abschnitt dieses Kapitels wird auf die Asymptotik des Binomialkoeffizienten eingegangen und es werden ein paar Hilfsresultate gezeigt, die in Kapitel 3 verwendet werden.

1.1 Graphentheorie

Definition 1.1 (vgl. [7], S. 2f & S. 30f). Ein (un-)gerichteter Graph $G = (V, E)$ besteht aus einer Knotenmenge $V = V(G)$ und einer Kantenmenge $E = E(G)$. Eine (un-)gerichtete Kante e ist ein (un-)geordnetes Paar $e = (v_1, v_2)$ von zwei Knoten $v_1, v_2 \in V$.¹

Meistens werden wir es mit ungerichteten Graphen zu tun haben, daher beziehen sich die weiteren Definitionen nur auf diesen Fall. Allerdings können diese meistens mit Berücksichtigung der Orientierung analog für gerichtete Graphen definiert werden. Wir befassen uns auch hauptsächlich mit schlichten Graphen, das heißt, es gibt keine Schlingen, das sind Kanten der Art (v, v) , und auch keine Mehrfachkanten zwischen zwei Knoten. So können

¹Im weiteren Verlauf bezeichnet die Schreibweise (a, b) wie üblich ein geordnetes Paar.

wir die Kanten als eine Relation auf den Knoten auffassen, welche im ungerichteten Fall symmetrisch ist. Auf diese Weise wird mit den geordneten Paaren (v, w) und (w, v) dieselbe Kante beschrieben.

Definition 1.2 (vgl. [7], S. 3 & S. 5). Für $v \in V(G)$ wird die Menge $\Gamma(v) = \{w \in V : (v, w) \in E(G)\}$ als Nachbarn und $d(v) = |\Gamma(v)|$ als der Grad des Knoten v bezeichnet.

Falls $e = (v, w)$ für $v, w \in V(G)$ gilt, heißen die Knoten v und w adjazent und sie inzidieren mit der Kante e . Aufgrund der vorigen Definition werden v und w auch als benachbart bezeichnet und wir sagen auch, dass die Kante e die Knoten v und w verbindet.

Definition 1.3 (vgl. [7], S. 3). Ein Graph heißt vollständiger Graph, falls je zwei Knoten des Graphens durch eine Kante verbunden sind. Für $n \in \mathbb{N}$ bezeichne K_n einen vollständigen Graphen mit $|V(K_n)| = n$.

Eine Folge von Kanten $(e_1, e_2, \dots, e_k) \in E(G)^k$ heißt Kantenfolge, falls die eine Kante bei der darauffolgenden endet. Das heißt, es gibt Knoten $v, v_1, v_2, \dots, v_{k-1}, w$ mit:

$$e_1 = (v, v_1), e_2 = (v_1, v_2), \dots, e_k = (v_{k-1}, w). \quad (1.1)$$

Man sagt auch, dass die Kantenfolge die Knoten v und w verbindet. Stimmt v mit w überein, wird die Kantenfolge als geschlossen bezeichnet. Sind zudem noch die Knoten in (1.1) paarweise verschieden, spricht man von einem Kreis (vgl. [7], S. 8 & S. 11).

Gibt es für je zwei Knoten $v, w \in V(G)$ eine Kantenfolge, welche diese beiden verbindet, heißt der Graph G zusammenhängend. Sind bei einer Kantenfolge die Knoten paarweise verschieden, wird von einem Weg beziehungsweise Pfad gesprochen (vgl. [7], S. 11 & S. 7). Mit diesen Begriffen ist es möglich, einen Spezialfall der k -Bäume zu beschreiben.

Definition 1.4 (vgl. [7], S. 14). Ein Graph, der keine Kreise enthält, wird als Wald bezeichnet. Ein Wald, der zusammenhängend ist, heißt Baum.

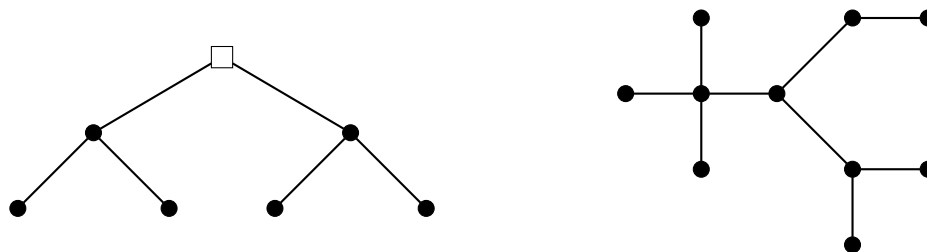


Abbildung 1: Zwei Beispiele für Bäume, wobei beim linken Baum ein Knoten als Wurzel ausgezeichnet ist.

In der Abbildung 1 ist bei einem Baum ein Knoten ausgezeichnet. Dieser wird Wurzel des Baumes genannt. Pfade können nun eine Richtung haben, sie können sich von der Wurzel wegbewegen oder zur Wurzel hinführen. Allerdings haben nicht alle Pfade eine Richtung, sie können auch abschnittsweise zur Wurzel führen und wieder weg. Ist $P = (v_1, v_2, \dots, v_l)$ ein Pfad zur Wurzel und v_l die Wurzel, dann nennen wir v_i ein Kind von v_{i+1} , für $1 \leq i < l$. Für $i > j \geq 1$ nennen wir v_i Vorfahre von v_j beziehungsweise v_j Nachfahre von v_i .

1.2 k -Bäume

Bevor wir hier mit der Definition eines k -Baumes starten, soll noch auf die Bezeichnung eingegangen werden. Im Englischen werden sie als „ k -dimensional trees“ und „ k -trees“ bezeichnet (siehe [3]). Wir werden diese nur als k -Bäume bezeichnen, da wir eine Verwechslungsgefahr mit den k -dimensionalen Suchbäumen zur Speicherung von Punkten aus \mathbb{R}^k unterbinden möchten.

Ein k -Baum wird rekursiv definiert und startet bei K_k . Der Graph K_k tritt auch oft aufgrund der Konstruktion als Teilgraph in einem k -Baum auf. Daher werden die vollständigen Graphen mit k Knoten im Zusammenhang mit k -Bäumen meistens anders bezeichnet. In der Arbeit von Beineke und Pippert ([3]) werden sie mit r -Zellen (engl. „ r -cells“) bezeichnet, wobei eine r -Zelle mit K_{r+1} übereinstimmt. Wir werden hingegen den mittlerweile gebräuchlichen Ausdruck einer k -Clique (engl. „ k -clique“) verwenden. Eine k -Clique ist isomorph zum vollständigen Graphen K_k .

Definition 1.5 (vgl. [3]). Eine k -Clique ist ein k -Baum und ein k -Baum mit $(n + 1)$ Knoten wird erzeugt, wenn ein $(n + 1)$ -ter Knoten mit jedem Knoten einer k -Clique eines k -Baumes mit n Knoten verbunden wird.

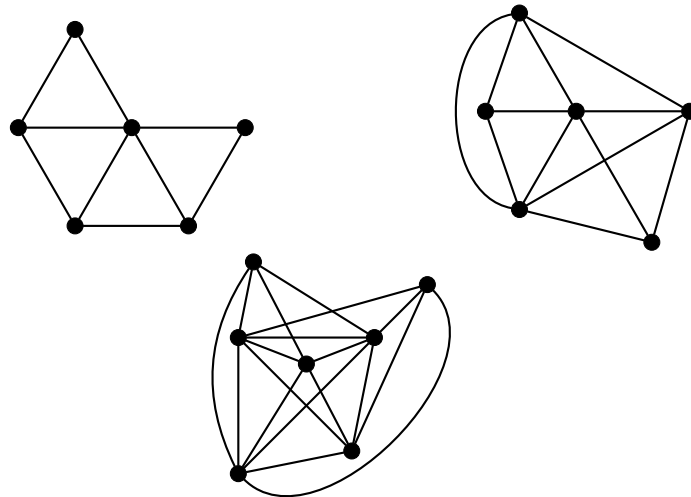


Abbildung 2: Beispiele für k -Bäume mit $k = 2, 3, 4$.

Wenn wir den Spezialfall $k = 1$ betrachten, ist ein 1-Baum über den rekursiven Aufbau sicher zusammenhängend. Ein Knoten wird nur bei einem anderen angehängt, daher ist es auch nicht möglich, dass Kreise entstehen. Ein 1-Baum ist also ein Baum aufgrund der Definition 1.4. Dementsprechend verallgemeinern wir auch die Sprechweisen, wie beispielsweise „Kind von“ auf k -Bäume und deren Cliques.

Ähnlich wie bei einem Baum, bei dem man die Kantenzahl einfach aus den Knoten bestimmen kann, ist es bei k -Bäumen möglich, die Anzahl der r -Cliques zu bestimmen.

Lemma 1.1 (vgl. [3]). *Es gibt in einem k -Baum mit n Knoten genau*

$$\binom{k}{r} + (n - k) \binom{k}{r - 1} = \binom{k}{r - 1} \frac{n + (r - 1)(n - k - 1)}{r} \quad (1.2)$$

r -Cliques, mit $1 \leq r \leq \min(k + 1, n)$. Insbesondere gibt es $\frac{1}{2}k(2n - k - 1)$ Kanten und $(k(n - k) + 1)$ k -Cliques.

Beweis: Wir unterscheiden zwei Fälle: $r = k + 1$ und $r \leq k$. Betrachten wir zuerst $r = k + 1$. Für $n = k + 1$ gibt es genau eine r -Clique, welche alle Knoten umfasst. Dies stimmt mit (1.2) überein. Für jeden weiteren Knoten v , der hinzugefügt wird, gibt es genau eine neue r -Clique. Diese besteht aus v und der k -Clique, bei der v angehängt wird. Es gibt somit $(n - k) (k + 1)$ -Cliques in einem k -Baum. Der zweite Fall $r \leq k$ wird auch über Induktion bewiesen. Für $n = k$ gibt es genau $\binom{k}{r}$ r -Cliques, denn es sind alle k Knoten paarweise miteinander verbunden, sodass wir beliebige r Stück für eine r -Clique auswählen können. Wenn ein weiterer Knoten v dem k -Baum hinzugefügt wird, muss jede neu entstehende r -Clique den Knoten v umfassen. Dieser ist genau mit k anderen Knoten verbunden, aus denen noch weitere $r - 1$ Knoten ausgewählt werden müssen, um eine r -Clique zu erzeugen. Dies bedeutet, dass sich mit jedem Knoten die Anzahl der r -Cliques um $\binom{k}{r-1}$ erhöht. Insgesamt ergibt das:

$$\begin{aligned} \binom{k}{r} + (n - k) \binom{k}{r-1} &= \binom{k}{r-1} \frac{k - r + 1}{r} + (n - k) \binom{k}{r-1} \\ &= \binom{k}{r-1} \frac{k + 1 + r(n - k - 1) \pm (n - k - 1)}{r} \\ &= \binom{k}{r-1} \frac{n + (r - 1)(n - k - 1)}{r}. \end{aligned}$$

Die Behauptung für die Kanten und die k -Cliques erfolgt durch Einsetzen. □

Wir werden in den folgenden Abschnitten nicht k -Bäume im Sinne der Definition 1.5 untersuchen, sondern jedem Knoten eines k -Baumes eine natürliche Zahl zuordnen.

Definition 1.6 (vgl. [3]). Ein markierter k -Baum mit n Knoten ist ein k -Baum mit n Knoten, wobei den Knoten bijektiv die Zahlen $\{1, \dots, n\}$ zugeordnet sind.

Die Zuordnung der Zahlen $1, \dots, n$ ist vollkommen willkürlich, man könnte dies auch mit einer anderen Beschriftung umsetzen, solange die Bezeichnung für jeden Knoten anders ist.

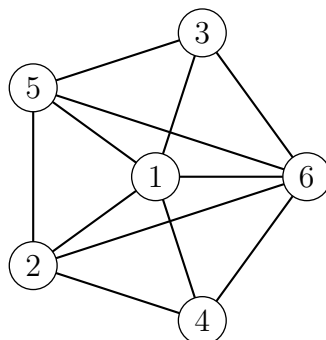


Abbildung 3: Ein markierter 3-Baum mit 6 Knoten.

1.2.1 k -Bäume als Graphen

Bei der Definition eines Baumes wird nicht berücksichtigt, dass er rekursiv aufgebaut werden kann. Dies wird durch die geforderten Eigenschaften erreicht. Bei unserer Definition von k -Bäumen ist es genau umgekehrt, jedoch kann auch hier eine Charakterisierung angegeben werden, welche den rekursiven Aufbau nicht explizit erwähnt.

Satz 1.2 ([22]). *Ein Graph G ist genau dann ein k -Baum, wenn:*

1. G ist zusammenhängend,
2. G hat eine k -Clique, aber keine $(k + 2)$ -Clique,
3. jeder minimale x - y -Trenner von G ist eine k -Clique.

Dabei ist ein x - y -Trenner für $(x, y) \notin E$ eine Teilmenge X der Knotenmenge V des Graphen $G = (V, E)$, sodass jeder Kantenzug von x nach y mindestens ein Element aus dem induzierten Graphen G' von X enthält. Der von X induzierte Graph ist folgendermaßen definiert durch: $G' = (X, \{e \in E : e = (v_1, v_2) \text{ mit } v_1, v_2 \in X\})$.

Die Bedingung 1 kam auch bei der Definition von Bäumen vor. Hier gibt es keine Kreisfreiheit mehr, da es schon in einer k -Clique viele Kreise gibt. Allerdings kann man die Bedingungen 2 und 3 informell so deuten, dass es bezüglich k -Cliquen keine Kreise gibt.

Definition 1.7 (vgl. [23], Lemma 1 und Theorem 1). Ein Graph G mit n Knoten heißt chordal oder trianguliert, wenn eine der folgenden äquivalenten Bedingungen gilt:

1. Jeder Kreis mit vier oder mehr Kanten besitzt eine Sehne (engl. „chord“), das ist eine Kante, welche nicht zum Kreis gehört und 2 Kanten des Kreises verbindet.
2. Jeder minimale x - y -Trenner ist eine Clique.
3. Der Graph besitzt eine perfekte Eliminationsordnung (engl. „perfect elimination ordering“). Dies ist eine Reihenfolge der Knoten (v_1, v_2, \dots, v_n) , sodass für $i = 1, \dots, n - 1$ die Nachbarn von v_i in $\{v_{i+1}, \dots, v_n\}$ eine Clique bilden.

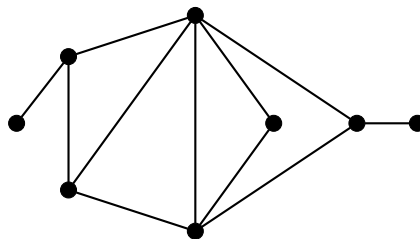


Abbildung 4: Beispiel für einen chordalen Graphen.

Es gibt noch weitere äquivalente Formulierungen für chordale Graphen. Hier sind nur jene aufgelistet, die in natürlicher Weise mit den k -Bäumen zusammenfallen.

Mit der Charakterisierung durch Satz 1.2 wissen wir, dass ein k -Baum chordal ist. Die perfekte Eliminationsordnung ist auch schnell gefunden, wenn man weiß, wie der k -Baum rekursiv aufgebaut ist. Wenn man die Reihenfolge des Aufbaus umdreht, erhält man die perfekte Eliminationsordnung.

Definition 1.8 (vgl. [1]). Ein partieller k -Baum (engl. „partial k -tree“) ist ein Teilgraph eines k -Baumes.

Partielle k -Bäume und chordale Graphen sind interessant, da es ein paar NP-schwere Probleme gibt, welche für diese Klasse von Graphen in polynomieller Laufzeit lösbar sind, wenn k fixiert wird. Die zulässige Färbung des Graphens ist ein Beispiel für ein solches Problem (vgl. [1] und [12], S. 65f).

1.3 Kombinatorische Strukturen und erzeugende Funktionen

Um die k -Bäume weiter zu untersuchen, benötigen wir unter anderem kombinatorische Strukturen und erzeugende Funktionen als mathematisches Werkzeug. Diese werden hier kurz eingeführt, wobei für eine umfangreichere Betrachtung auf [8] verwiesen sei.

Definition 1.9 (vgl. [8], Definition I.1 und I.2, S16). Eine kombinatorische Struktur \mathcal{C} ist ein Paar (C, w) , bestehend aus einer höchstens abzählbaren Menge C und einer Größenfunktion $w : C \rightarrow \mathbb{N}$ mit der Eigenschaft, dass die Mengen

$$w^{-1}(n) = \{c \in C : w(c) = n\}$$

für beliebiges $n \in \mathbb{N}$ endlich sind. Es bezeichne c_n die Anzahl der Elemente der Größe n : $c_n := |w^{-1}(n)|$.

Später wird die Bewertungsfunktion meistens nicht mehr explizit erwähnt und es findet in der Notation keine Unterscheidung mehr zwischen der Menge C und der dazugehörigen kombinatorischen Struktur statt. Die Bewertungsfunktion wird manchmal auch mit Betragstrichen notiert.

Definition 1.10 (vgl. [8], Definition I.3, S. 19). Zwei kombinatorische Strukturen \mathcal{A}, \mathcal{B} sind isomorph (i.Z.: $\mathcal{A} \cong \mathcal{B}$), falls $a_n = b_n \forall n \in \mathbb{N}$, gilt.

Definition 1.11 (vgl. [8] Definition I.4 S. 19 und II.2 S. 97). Die (gewöhnlich) erzeugende Funktion einer kombinatorischen Struktur \mathcal{C} ist die formale Potenzreihe

$$C_g(z) = \sum_{c \in \mathcal{C}} z^{|c|} = \sum_{n=0}^{\infty} c_n z^n.$$

Die exponentiell erzeugende Funktion ist in analoger Weise definiert:

$$C(z) = \sum_{c \in \mathcal{C}} \frac{z^{|c|}}{|c|!} = \sum_{n=0}^{\infty} c_n \frac{z^n}{n!}.$$

Wenn man nun eine erzeugende Funktion für eine kombinatorische Struktur kennt, möchte man auch die Koeffizienten ablesen. Den Operator bezeichnen wir wie üblich mit $[z^n]$:

$$[z^n]C_g(x) = [z^n] \left(\sum_{n=0}^{\infty} c_n z^n \right) = c_n.$$

Ein Vorteil von erzeugenden Funktionen ist deren Verhalten unter kombinatorischen Konstruktionen. So kann zum Beispiel sehr leicht die erzeugende Funktion von der disjunkten Vereinigung zweier kombinatorischen Strukturen berechnet werden. Dies wird später noch betrachtet werden, davor muss noch eine Unterteilung der kombinatorischen Strukturen vorgenommen werden. Wir nennen eine kombinatorische Struktur \mathcal{C} markiert, wenn mit jedem Element c der Größe $n \in \mathbb{N}$ eine Permutation der Zahlen $1, \dots, n$ verknüpft ist, indem Teilobjekte von c mit diesen Zahlen markiert sind. Wir können so ein Element c mit Permutationen der Bezeichnungen in $n! - 1$ andere Objekte von \mathcal{C} überführen. Ein Beispiel für solch eine kombinatorische Struktur wären die markierten k -Bäume mit der Knotenanzahl als Bewertungsfunktion. Wenn dies nicht der Fall ist, nennen wir die Struktur unmarkiert. Dies wären beispielsweise k -Bäume oder die Potenzmenge einer endlichen Menge mit der Mächtigkeit als Bewertungsfunktion. Bei unmarkierten Strukturen werden die gewöhnlichen erzeugten Funktionen zur Beschreibung herangezogen und bei markierten die exponentiell Erzeugenden.

Wenn wir zwei kombinatorische Strukturen \mathcal{A} und \mathcal{B} haben, wollen wir die disjunkte Vereinigung dieser beider Strukturen bilden. Wenn diese \mathcal{C} ist, so gilt $\mathcal{C} = \mathcal{A} \cup \mathcal{B}$ und für die Bewertungsfunktion:

$$|c|_{\mathcal{C}} = \begin{cases} |c|_{\mathcal{A}}, & \text{falls } c \in \mathcal{A}, \\ |c|_{\mathcal{B}}, & \text{falls } c \in \mathcal{B}, \end{cases} \quad \forall c \in \mathcal{C}.$$

Es ist üblich, die Vereinigung als Addition zu schreiben ($\mathcal{C} = \mathcal{A} + \mathcal{B}$). Dies ist auf das Verhalten der erzeugenden Funktionen zurückzuführen. Die erzeugende Funktion von \mathcal{C} ist nämlich die Summe der erzeugenden Funktionen von \mathcal{A} und \mathcal{B} .

Ähnlich geht dies auch beim Produkt zweier Strukturen. So ist bei $\mathcal{D} = \mathcal{A} \times \mathcal{B}$ die Bewertungsfunktion als $|(a, b)|_{\mathcal{D}} = |a|_{\mathcal{A}} + |b|_{\mathcal{B}}$ definiert. Die erzeugende Funktion von \mathcal{D} lässt sich auch als Produkt schreiben.

Mit diesen beiden einfachen Konstruktionen können schon kompliziertere Objekte, wie zum Beispiel Mengen von Elementen einer kombinatorischen Struktur, dargestellt werden. In den Tabellen 1 und 2 sind bekannte Grundkonstruktionen für die unmarkierten und markierten Strukturen aufgelistet. Die exakte Definition kann in [8] Kapitel I und II nachgelesen werden. Wir möchten nur noch darauf hinweisen, dass in markierten Strukturen das Produkt ein wenig anders definiert werden muss. Das Produkt zweier markierter Elemente muss auch eine gültige Markierung besitzen. Dies geschieht in natürlicher Weise, indem die Reihenfolge der Markierung, eingeschränkt auf die einzelnen Elemente, beibehalten wird, aber die Markierungen aus der Summe der Größen ausgewählt werden. Das bedeutet für zwei Objekte der Größe m und n : Man wählt eine m -elementige Teilmenge M von $\{1, 2, \dots, n + m\}$ aus und das Element vom ersten Objekt, welches mit 1 markiert war, wird nun mit der kleinsten Zahl von M markiert. Der Zweier wird durch die zweitkleinste Zahl ersetzt usw. Das zweite Objekt wird analog mit $\{1, 2, \dots, n + m\} \setminus M$ markiert. So erhalten wir eine gültige Markierung. Da jede beliebige m -elementige Auswahl eine gültige Beschriftung liefert, ergibt dies $\binom{n+m}{m}$ Objekte im Produkt. Wie man nachrechnen kann, passt dies auch für die exponentiell erzeugende Funktion in dem Sinne, dass es genau dem gewöhnlichen Cauchy-Produkt entspricht.

Wir möchten noch darauf hinweisen, dass für manche Konstruktionen gewisse Voraussetzungen gelten. Zum Beispiel wird beim Box-Produkt $\mathcal{C} = \mathcal{A}^{\square} \star \mathcal{B}$ gefordert, dass $a_0 = 0$ ist. Beim Box-Produkt wird ein Element aus \mathcal{A} mit dem Einser markiert (vgl. [8], S. 139).

Konstruktion		EF
Vereinigung	$\mathcal{C} = \mathcal{A} + \mathcal{B}$	$C_g(z) = A_g(z) + B_g(z)$
Produkt	$\mathcal{C} = \mathcal{A} \times \mathcal{B}$	$C_g(z) = A_g(z) \cdot B_g(z)$
Folgen	$\mathcal{C} = \text{SEQ}(\mathcal{A})$	$C_g(z) = \frac{1}{1-A_g(z)}$
Mengen	$\mathcal{C} = \text{SET}(\mathcal{A})$	$C_g(z) = \exp(A_g(z) - \frac{1}{2}A_g(z^2) + \frac{1}{3}A_g(z^3) - \dots)$
Auswählen	$\mathcal{C} = \Theta\mathcal{A}$	$C_g(z) = z \frac{d}{dz} A_g(z)$
Verknüpfung	$\mathcal{C} = \mathcal{A}(\mathcal{B})$	$C_g(z) = A_g(B_g(z))$

Tabelle 1: Zulässige Konstruktionen für unmarkierte Strukturen (vgl. [8], S. 93).

Konstruktion		EF
Vereinigung	$\mathcal{C} = \mathcal{A} + \mathcal{B}$	$C(z) = A(z) + B(z)$
Produkt	$\mathcal{C} = \mathcal{A} \star \mathcal{B}$	$C(z) = A(z) \cdot B(z)$
Folgen	$\mathcal{C} = \text{SEQ}(\mathcal{A})$	$C(z) = \frac{1}{1-A(z)}$
Mengen	$\mathcal{C} = \text{SET}(\mathcal{A})$	$C(z) = \exp(A(z))$
Zyklus	$\mathcal{C} = \text{CYC}(\mathcal{A})$	$C(z) = \log \frac{1}{1-A(z)}$
Auswählen	$\mathcal{C} = \Theta\mathcal{A}$	$C(z) = z \frac{d}{dz} A(z)$
Verknüpfung	$\mathcal{C} = \mathcal{A}(\mathcal{B})$	$C(z) = A(B(z))$
Box-Product	$\mathcal{C} = \mathcal{A}^\square \star \mathcal{B}$	$C(z) = \int_0^z \frac{d}{dt} A(t) \cdot B(t) dt$

Tabelle 2: Zulässige Konstruktionen für markierte Strukturen (vgl. [8], S. 148).

Für die genauen formalen Konstruktionen sei auf die Fachliteratur verwiesen.

Beispiele: Für die Verdeutlichung der kombinatorischen Konstruktionen und dem Unterschied zwischen markierten und unmarkierten Objekten betrachten wir noch zwei Beispiele.

Beginnen wir mit den Binärbäumen \mathcal{B} als unmarkierte Struktur und die Bewertungsfunktion gibt die Anzahl der internen Knoten an. Ein Binärbaum kann rekursiv definiert werden: Entweder ist er ein externer Knoten (\square) oder er besteht aus einem internen

Knoten (\circ), der sich in zwei Binärbäume verzweigt. Es gilt somit:

$$\mathcal{B} = \{\square\} + \{\circ\} \times \mathcal{B} \times \mathcal{B}.$$

Nun benützen wir die erzeugenden Funktionen, um die Koeffizienten zu berechnen.

$$\begin{aligned} B_g(z) &= 1 + zB_g(z)^2, \\ 0 &= zB_g(z)^2 - B_g(z) + 1, \\ B_g(z) &= \frac{1 \pm \sqrt{1 - 4z}}{2z}. \end{aligned}$$

Da $B_g(z)$ eine Potenzreihe ist, fällt eine der Lösungen der quadratischen Gleichungen weg, was zu

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

führt. Die Koeffizienten, die uns interessieren, können mittels der binomischen Reihe bestimmt werden (vgl. [8], S. 67):

$$\begin{aligned} b_n &= [z^n]B_g(z) = [z^n] \frac{1 - \sqrt{1 - 4z}}{2z} \\ &= \frac{1}{2}[z^{n+1}] \left((-1) \sum_{k=0}^{\infty} \binom{1/2}{k} (-4z)^k \right) = (-1) \frac{(-4)^{n+1}}{2} \binom{1/2}{n+1} \\ &= \frac{(-1)^n 4^{n+1} \frac{1}{2} (\frac{1}{2} - 1) \cdots (\frac{1}{2} - n)}{2 (n+1)!} \\ &= 2^n \frac{(2n-1)(2n-3) \cdots 3 \cdot 1}{(n+1)!} \underbrace{\frac{2n(2n-2) \cdots 2}{2^n n!}}_{=1} = \frac{1}{n+1} \binom{2n}{n}. \end{aligned}$$

Als ein Beispiel für markierte kombinatorische Strukturen betrachten wir Permutationen. Je nach ihrer Darstellung können wir sie unterschiedlich auffassen. Die Permutationen sind aufgrund der Tupelschreibweise eine Folge von einfach markierten Elementen (\circ) der Größe 1. Mit der Zyklendarstellung können wir sie auch als Menge von Zyklen einfach markierter Elemente auffassen. In der Notation der kombinatorischen Strukturen erhalten wir für die Permutationen \mathcal{P} die Gleichung

$$\mathcal{P} = \text{SEQ}(\{\circ\}) = \text{SET}(\text{CYC}(\{\circ\})).$$

Dies ergibt für die erzeugenden Funktionen

$$\begin{aligned} P(z) &= \frac{1}{1-z} = \exp(\log(\frac{1}{1-z})) \\ &= \sum_{i=0}^{\infty} i! \frac{z^i}{i!}, \end{aligned}$$

und für die Koeffizienten (vgl. [8], S. 104)

$$p_n = n![z^n]P(z) = n!.$$

1.3.1 Nützliche Potenzreihen und formale Topologie

Bei den vorigen Beispielen wurde schon zweimal das Wissen über die Reihenentwicklung von Funktionen eingesetzt. Hier listen wir noch einmal zur Erinnerung ein paar Funktionen auf, die im Folgenden öfters eingesetzt werden. Eine umfassendere Liste ist in dem Werk von Wilf [27], aus welchem der kommende Ausschnitt von 5 Gleichungen stammt, ab Seite 52 angegeben.

$$\begin{aligned}\frac{1}{1-z} &= \sum_{n \geq 0} z^n, \\ e^z &= \sum_{n \geq 0} \frac{z^n}{n!}, \\ \log \frac{1}{1-x} &= \sum_{n \geq 0} \frac{z^n}{n}, \\ (1+z)^\alpha &= \sum_{n \geq 0} \binom{\alpha}{n} z^n, \\ \frac{1}{(1-z)^m} &= \sum_{n \geq 0} \binom{m+n-1}{m-1} z^n.\end{aligned}$$

Eine weitere Darstellung einer Funktion durch eine Potenzreihe geht auf die vorzeichenlosen Stirling-Zahlen erster Art zurück. Die Stirling-Zahlen $\begin{bmatrix} n \\ k \end{bmatrix}$ geben an, wie viele Permutationen es mit k Zyklen gibt. Für sie gilt die Rekursion $\begin{bmatrix} n+1 \\ k \end{bmatrix} = \begin{bmatrix} n \\ k-1 \end{bmatrix} + n \begin{bmatrix} n \\ k \end{bmatrix}$ mit den Startwerten $\begin{bmatrix} n \\ n \end{bmatrix} = 1$ und $\begin{bmatrix} n \\ k \end{bmatrix} = 0$ für $k = 0 < n$ oder $n < k$ (vgl. [11], S.245ff). Die Stirling-Zahlen erster Art erfüllen folgende Gleichung für bivariate erzeugende Funktionen (vgl. [8], S. 736):

$$(1-z)^{-v} = \sum_{n \geq 0} \sum_{k \geq 0} \begin{bmatrix} n \\ k \end{bmatrix} \frac{z^n}{n!} v^k. \quad (1.3)$$

Definition 1.12 (vgl. [8], S. 157). Für eine beliebige diskrete Zufallsvariable X bezeichnen wir $P(v) = \sum_{k \geq 0} \mathbb{P}\{X = k\} v^k$ als die wahrscheinlichkeitserzeugende Funktion.

Wenn man eine kombinatorische Struktur \mathcal{A} mit einem Parameter χ untersucht, verwendet man die bivariate erzeugende Funktion $A(z, v) = \sum_{n \geq 0} \sum_{k \geq 0} a_{n,k} \frac{z^n}{n!} v^k$. Diese markiert mit v diesen Parameter, sodass $a_{n,k}$ die Anzahl der Elemente von \mathcal{A} der Größe n mit $\chi = k$ angibt. Unter der Voraussetzung, dass die Objekte von $A_n = \{a \in \mathcal{A} : w(a) = n\}$ uniform verteilt sind, können wir die dadurch induzierte Wahrscheinlichkeitsverteilung von χ angeben mit:

$$\mathbb{P}_{A_n}\{\chi = k\} = \frac{a_{n,k}}{a_n}.$$

Für den Parameter χ beziehungsweise seiner entsprechenden Zufallsvariablen können wir die wahrscheinlichkeitserzeugenden Funktionen aus der bivariaten erzeugenden Funktion herleiten (vgl. [8], S. 157):

$$P_{A_n}(v) = \frac{n! [z^n] A(z, v)}{a_n} = \frac{[z^n] A(z, v)}{[z^n] A(z, 1)}.$$

Formale Topologie

Auf den formalen Potenzreihen P wird über eine Metrik eine Topologie, die sogenannte formale Topologie, definiert. Dabei wird der Abstand von zwei Elementen über den ersten Koeffizienten, der unterschiedlich ist, definiert. Dazu wird die Funktion $\nu : P \rightarrow \mathbb{N} \cup \{\infty\}$ mit $\nu(0) = \infty$ und $\nu(A(z)) = \min\{n \in \mathbb{N} : a_n \neq 0\}$, falls $A(z) \neq 0$, benützt. Mit der Konvention $2^{-\infty} = 0$ definieren wird die Metrik $d : P^2 \rightarrow \mathbb{R}$ über:

$$d(A(z), B(z)) := 2^{-\nu(A(z)-B(z))}.$$

Dies bedeutet für zwei Potenzreihen $A(z), B(z)$ mit $d(A(z), B(z)) < 2^{-k}$, dass die ersten k Koeffizienten übereinstimmen. Dieses Argument können wir auf eine Menge von Potenzreihen erweitern, falls der paarweise Abstand gering genug ist. Es kann für gegebenes $m \in \mathbb{N}$ bei einer Cauchyfolge $(A_n(z))_{n \in \mathbb{N}}$ ein $\epsilon_m > 0$ mit $-\log_2 \epsilon_m > m$ gewählt werden. So gibt es ein $n_m \in \mathbb{N}$ mit $d(A_n, A_{n'}) < \epsilon_m$ für $n, n' \geq n_m$, respektive $[z^k]A_{n'} = [z^k]A_n$ für $k \leq m$. Die Potenzreihe $B(z)$ mit $b_k := [z^k]A_{n_k}(z)$ ist der Grenzwert der Cauchyfolge. Somit bilden die Potenzreihen mit der formalen Metrik einen vollständigen metrischen Raum (vgl. [8], S. 730f).

1.3.2 Lagrange'sche Inversionsformel und weitere Sätze

In diesem Abschnitt werden noch ein paar Sätze angegeben, welche im weiteren Teil der Arbeit verwendet werden. Auf die Beweise wird verzichtet und dafür auf die Literatur verwiesen. Der folgende Satz, die Lagrange'sche Inversionsformel, ist sehr hilfreich beim Lösen von Gleichungen über kombinatorischen Strukturen. Die Gleichungen können meistens über einen rekursiven Aufbau schnell hergeleitet werden.

Satz 1.3 (Lagrange'sche Inversionsformel, vgl. [8], S. 732). *Es sei φ eine Potenzreihe mit $\varphi_0 \neq 0$ und $y(z)$ eine Potenzreihe mit $y = z\varphi(y)$. Dann gilt für eine beliebige Potenzreihe h :*

$$[z^n]h(y(z)) = \frac{1}{n}[u^{n-1}]\varphi(u)^n h'(u). \quad (\text{LIF})$$

Ein weiteres Hilfsmittel, welches benötigt wird, ist das Lemma von Burnside. Zunächst sei es hier in der allgemeinen Formulierung für Gruppenaktionen erwähnt.

Satz 1.4 (vgl. [5] §119 und [4]). *Wenn eine endliche Gruppe G auf einer Menge Ω agiert, dann kann die Anzahl k der Orbits mit Hilfe der Fixpunkte unter einem Element von G bestimmt werden:*

$$k = \frac{1}{|G|} \sum_{g \in G} |\{\omega \in \Omega : \omega^g = \omega\}|. \quad (1.4)$$

Wenn wir als endliche Gruppe eine Untergruppe \mathfrak{S} der Permutationsgruppe \mathfrak{S}_M einer endlichen Menge M wählen, kann eine Relation \sim mit $m_1 \sim m_2 \Leftrightarrow \exists \sigma \in \mathfrak{S} : \sigma(m_1) = m_2$ definiert werden. Die Äquivalenzklassen dieser Relation sind genau die Orbits und die Formel (1.4) ergibt:

$$|M/\sim| = \frac{1}{|\mathfrak{S}|} \sum_{\sigma \in \mathfrak{S}} |\{m \in M : \sigma(m) = m\}| = \frac{1}{|\mathfrak{S}|} \sum_{\sigma \in \mathfrak{S}} |\text{Fix}_M(\sigma)|,$$

wobei mit $\text{Fix}_M(\sigma)$ die Menge der Fixpunkte der Funktion $\sigma : M \rightarrow M$ bezeichnet wird. Bei der asymptotischen Betrachtung von Vorfahren in k -Bäumen wird das „Quasi-Power Theorem“ von Hwang verwendet.² Für eine um 1 analytische Funktion $f(z)$ mit $f(1) = 1$ schreiben wir

$$\mathbb{E}(f(z)) = f'(1), \quad \mathbb{V}(f(z)) = f''(1) + f'(1) - f'(1)^2.$$

Wenn $f(z)$ eine wahrscheinlichkeitserzeugende Funktion von X wäre, dann würden diese beiden Ausdrücke mit dem Erwartungswert und der Varianz von X übereinstimmen. Dies legt obige Schreibweise nahe.

Satz 1.5 („Quasi-Power Theorem“, vgl. [8], Theorem IX.8). *Sei $(X_n)_{n \in \mathbb{N}}$ eine Folge von diskreten nichtnegativen Zufallsvariablen mit den wahrscheinlichkeitserzeugenden Funktionen $(p_n)_{n \in \mathbb{N}}$ und $(\beta_n)_{n \in \mathbb{N}}, (\kappa_n)_{n \in \mathbb{N}}$ Folgen, die gegen $+\infty$ divergieren. Ist*

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

mit $A(v), B(v)$ um $v = 1$ analytische Funktionen mit $A(1) = B(1) = 1$ und ist die Bedingung

$$\mathbb{V}(B(v)) = B''(1) + B'(1) - B'(1)^2 \neq 0$$

erfüllt, so gilt:

$$\begin{aligned} \mathbb{E}(X_n) &= \beta_n \mathbb{E}(B(v)) + \mathbb{E}(A(v)) + \mathcal{O}(\kappa_n^{-1}), \\ \mathbb{V}(X_n) &= \beta_n \mathbb{V}(B(v)) + \mathbb{V}(A(v)) + \mathcal{O}(\kappa_n^{-1}). \end{aligned}$$

Darüber hinaus ist die Verteilung von X_n nach Standardisierung asymptotisch normalverteilt:

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

mit $\phi(x)$ der Verteilungsfunktion der Normalverteilung:

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{y^2}{2}} dy.$$

1.4 Asymptotisches Verhalten und Hilfsresultate

Viele Ergebnisse werden eine binomische Reihe beinhalten, sodass wir hier noch einmal die allgemeine Definition des Binomialkoeffizienten in Erinnerung rufen möchten. Bei $\binom{n}{k}$ kann n recht einfach auf einen Wert aus \mathbb{C} verallgemeinert werden. Dazu wird die Schreibweise der fallenden Faktoriellen benötigt:

$$n^{\underline{k}} := n(n-1) \cdots (n-k+1),$$

²Die Definition der \mathcal{O} -Notation wird in 1.4 gegeben.

womit der Binomialkoeffizient kurz als

$$\binom{\alpha}{k} = \frac{\alpha^k}{k!}$$

geschrieben werden kann für $\alpha \in \mathbb{C}$. Für $n \in \mathbb{N}$ stimmen die fallenden Faktoriellen n^k auch mit $\frac{n!}{(n-k)!}$ überein (vgl. [11], S. 47f).

Für Vereinfachungen ist es auch sinnvoll, die Darstellung über die Gamma-Funktion zu wählen. So definieren wir für $\alpha, \beta \in \mathbb{C}$ mit $-\alpha, -\beta, \beta - \alpha \notin \mathbb{N}$ (vgl. [18], Abschnitt 4.3):

$$\binom{\alpha}{\beta} = \frac{\Gamma(\alpha + 1)}{\Gamma(\beta + 1)\Gamma(\alpha - \beta + 1)}.$$

Wegen $\Gamma(x + 1) = x\Gamma(x)$ stimmt diese Definition auf den natürlichen Zahlen mit der herkömmlichen Definition der Faktoriellen überein. Es sind Resultate, die auf der Eigenschaft $n! = n(n - 1)!$ für Faktorielle beruhen, auch für den allgemeinen Binomialkoeffizienten gültig. Dies betrifft unter anderem $\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1}$ und $\sum_{k=0}^m \binom{n+k}{k} = \binom{n+m+1}{n+1}$. Um die Asymptotik von Termen anzugeben, verwenden wir die \mathcal{O} -Notation.

Definition 1.13 (vgl. [11], S. 428ff). Es seien $(a_n)_{n \in I}, (b_n)_{n \in I}$ zwei Folgen (bzw. Netze über \mathbb{R}) in \mathbb{C} . Wir sagen, dass a_n nicht schneller wächst als b_n ($a_n = \mathcal{O}(b_n)$ bzw. $a_n \in \mathcal{O}(b_n)$), wenn:

$$\frac{|a_n|}{|b_n|} \leq C, \quad \text{für } n \geq n_0 \in I.$$

Wir nennen die beiden Folgen asymptotisch äquivalent ($a_n \sim b_n$), falls:

$$\lim_{n \rightarrow \infty} \frac{|a_n|}{|b_n|} = 1.$$

Nach der Stirling-Formel gilt somit für die Gamma-Funktion die Darstellung (vgl. [8], S. 747):

$$\Gamma(x) = \sqrt{\frac{2\pi}{x}} \left(\frac{x}{e}\right)^x \left(1 + \mathcal{O}\left(\frac{1}{x}\right)\right), \quad \text{für } x \in \mathbb{R},$$

und insbesondere

$$\Gamma(x) \sim \sqrt{\frac{2\pi}{x}} \left(\frac{x}{e}\right)^x, \quad \text{für } x \in \mathbb{R}.$$

Mit der Stirling-Formel ist es auch möglich, den Binomialkoeffizienten asymptotisch zu behandeln. Im Laufe der Arbeit werden wir mehrere Male Binomialkoeffizienten der Form $\binom{n+\alpha}{n}$ bezüglich der Asymptotik von $n \rightarrow \infty$ behandeln. Als Vorbereitung betrachten wir drei Zwischenresultate. Zuerst schätzen wir den Fehler von $(1 + \frac{x}{n})^n$ ab. Diese Folge konvergiert bekanntermaßen gegen e^x und zwar monoton steigend. Mit Exponent $n + 1$ ist die Folge fallend und konvergiert gegen denselben Grenzwert. So haben wir:

$$\left(1 + \frac{x}{n}\right)^n \leq e^x \leq \left(1 + \frac{x}{n}\right)^{n+1},$$

was uns für Fehlerabschätzungen hilft:

$$\left| \frac{(1+xn^{-1})^n}{e^{\frac{x}{n}}} - 1 \right| = n \left(1 - \frac{(1+\frac{x}{n})^n}{e^x} \right) \leq n - \frac{n}{1+\frac{x}{n}} \leq C, \quad \text{für } n \geq n_0$$

$$\Rightarrow \left(1 + \frac{x}{n}\right)^n = e^x (1 + \mathcal{O}(n^{-1})).$$

Ein weiterer Term, der bei der Rechnung für die Asymptotik des Binomialkoeffizienten auftritt, ist $(1 + \frac{x}{n})^\alpha$. Es sei α und x beschränkt. Dann gilt $(1 + \frac{x}{n})^\alpha = 1 + \mathcal{O}(n^{-1})$. Um die binomische Reihe verwenden zu können, betrachten wir nur $n \in \mathbb{N}$ mit $|x| \leq C_x < n$:

$$\left(1 + \frac{x}{n}\right)^\alpha = \sum_{l \geq 0} \binom{\alpha}{l} \frac{x^l}{n^l} = 1 + \mathcal{O}(n^{-1}). \quad (1.5)$$

Mit folgendem Lemma soll kurz festgehalten werden, wie sich der Fehler in der \mathcal{O} -Notation bei einer Division verhält.

Lemma 1.6. Sei $a_n = \tilde{a}_n(1 + \mathcal{O}(n^{-\alpha}))$ und $b_n = \tilde{b}_n(1 + \mathcal{O}(n^{-\beta}))$ mit $\alpha, \beta > 0$. Dann gilt:

$$\frac{a_n}{b_n} = \frac{\tilde{a}_n}{\tilde{b}_n} (1 + \mathcal{O}(n^{-\min(\alpha, \beta)})).$$

Beweis: Sei ohne Beschränkung $\alpha \leq \beta$ und C_1, C_2 die Konstanten mit $|\frac{a_n}{\tilde{a}_n} - 1|n^\alpha \leq C_1$ und $|\frac{b_n}{\tilde{b}_n} - 1|n^\beta \leq C_2$. Für n groß genug gilt:

$$\begin{aligned} \left| \frac{\frac{a_n \tilde{b}_n}{\tilde{a}_n b_n} - 1 \right| &= \frac{\left| \left(\frac{a_n}{\tilde{a}_n} - 1 \right) \left(\frac{\tilde{b}_n}{b_n} - 1 \right) + \frac{\tilde{b}_n}{b_n} + \frac{a_n}{\tilde{a}_n} - 2 \right|}{n^{-\alpha}} \\ &\leq \frac{\left| \left(\frac{a_n}{\tilde{a}_n} - 1 \right) \left(\frac{\tilde{b}_n}{b_n} - 1 \right) \right| + \left| \frac{\tilde{b}_n}{b_n} - 1 \right| + \left| \frac{a_n}{\tilde{a}_n} - 1 \right|}{n^{-\alpha}} \\ &\leq C_1 + C_2 + C_1. \end{aligned}$$

□

Mit diesen Vorbereitungen können wir den Binomialkoeffizienten asymptotisch beschreiben.

Lemma 1.7. Für $n \in \mathbb{N}$ und $\alpha \in \mathbb{R}$ gilt:

$$\binom{n + \alpha}{n} = \frac{1}{\Gamma(\alpha + 1)} n^\alpha (1 + \mathcal{O}(n^{-1})) \sim \frac{1}{\Gamma(\alpha + 1)} n^\alpha.$$

Beweis:

$$\begin{aligned}
\binom{n+\alpha}{n} &= \frac{\Gamma(n+\alpha+1)}{\Gamma(n+1)\Gamma(\alpha+1)} = \frac{1}{\Gamma(\alpha+1)} \frac{\sqrt{\frac{2\pi}{n+\alpha+1}} \left(\frac{n+\alpha+1}{e}\right)^{n+\alpha+1}}{\sqrt{\frac{2\pi}{n+1}} \left(\frac{n+1}{e}\right)^{n+1}} (1 + \mathcal{O}(n^{-1})) \\
&= \frac{1}{\Gamma(\alpha+1)} \frac{n^{n+\alpha+1} \left(1 + \frac{\alpha+1}{n}\right)^{n+\alpha+1}}{n^{n+1} \left(1 + \frac{1}{n}\right)^{n+1} e^\alpha} (1 + \mathcal{O}(n^{-1})) \\
&\stackrel{(1.5)}{=} \frac{1}{\Gamma(\alpha+1)} \frac{n^\alpha e^{\alpha+1}}{e^1 e^\alpha} (1 + \mathcal{O}(n^{-1})) \\
&= \frac{1}{\Gamma(\alpha+1)} n^\alpha (1 + \mathcal{O}(n^{-1})) \sim \frac{1}{\Gamma(\alpha+1)} n^\alpha.
\end{aligned}$$

□

Bei der Grenzwertberechnung werden wir auch das Wissen über das asymptotische Verhalten einsetzen. Dies soll mit folgendem Lemma noch formal beschrieben werden.

Lemma 1.8. *Es seien $(a_n)_{n \in \mathbb{N}}, (b_n)_{n \in \mathbb{N}}, (\tilde{a}_n)_{n \in \mathbb{N}}, (\tilde{b}_n)_{n \in \mathbb{N}}$ Folgen, wobei $a_n \sim \tilde{a}_n, b_n \sim \tilde{b}_n$ und $b_n \neq 0 \neq \tilde{b}_n, \forall n \in \mathbb{N}$. Existiert einer der beiden folgenden Limiten, dann gilt:*

$$\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = \lim_{n \rightarrow \infty} \frac{\tilde{a}_n}{\tilde{b}_n}.$$

Beweis: Falls $\lim_{n \rightarrow \infty} \frac{\tilde{a}_n}{\tilde{b}_n}$ existiert, gilt:

$$\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = \lim_{n \rightarrow \infty} \frac{a_n \tilde{a}_n \tilde{b}_n}{\tilde{a}_n \tilde{b}_n b_n} = \lim_{n \rightarrow \infty} \frac{\tilde{a}_n}{\tilde{b}_n}.$$

□

Nun folgen noch drei Gleichungen für Summen von Binomialkoeffizienten, die wir in Abschnitt 3 mehrere Male benötigen, um die Ergebnisse umzuformen. Folgendes Resultat wird bei der Berechnung des Knotengrades hilfreich sein.

Lemma 1.9 (vgl. [26]). *Für $-z \notin \mathbb{N}$ und $m \in \mathbb{N}$ gilt:*

$$\sum_{l=0}^m \frac{(-1)^l \binom{m}{l}}{l+z+1} = \frac{1}{(m+z+1) \binom{m+z}{m}}.$$

Beweis: Um diese Gleichung zu zeigen, wird eine Induktion nach m durchgeführt. Für

$m = 0$ stimmt es. Der Induktionsschritt wird nachgewiesen mit:

$$\begin{aligned}
\sum_{l=0}^{m+1} \frac{(-1)^l \binom{m+1}{l}}{l+z+1} &= \sum_{l=1}^m \frac{(-1)^l \left(\binom{m}{l-1} + \binom{m}{l} \right)}{l+z+1} + \frac{1}{z+1} + \frac{(-1)^{m+1}}{m+1+z+1} \\
&= \sum_{l=0}^m \frac{(-1)^l \binom{m}{l}}{l+z+1} - \sum_{l=0}^m \frac{(-1)^l \binom{m}{l}}{l+(z+1)+1} \\
&\stackrel{\text{(IV)}}{=} \frac{1}{(m+z+1) \binom{m+z}{m}} - \frac{1}{(m+z+2) \binom{m+z+1}{m}} \\
&= \frac{m!}{(m+1+z)^{m+1}} - \frac{m!(z+1)}{(m+z+2)(m+1+z)^{m+1}} \\
&= \frac{\frac{m+1+z+1}{m+1}}{(m+z+2) \binom{m+1+z}{m+1}} - \frac{\frac{z+1}{m+1}}{(m+z+2) \binom{m+1+z}{m+1}} \\
&= \frac{1}{(m+1+z+1) \binom{m+1+z}{m+1}}.
\end{aligned}$$

□

Mit Hilfe des Differenzenoperators können wir das nächste Zwischenresultat zeigen, welches im Beweis von Lemma 1.11 benötigt wird.

Lemma 1.10 (vgl. [18], Abschnitt 4.3). *Für $n, m, j \in \mathbb{N}$ gilt:*

$$(n-j)^m = \sum_{l=0}^m \binom{m}{l} (-1)^l (n+\alpha-l)^{m-l} (j+\alpha)^l.$$

Beweis: Wir definieren $a_{n,m} := (n-j)^m$ und $b_{n,m} := \sum_{l=0}^m \binom{m}{l} (-1)^l (n+\alpha-l)^m (j+\alpha)^l$. Für den Differenzenoperator Δ für Folgen (nach n) ergibt sich:

$$\begin{aligned}
\Delta(a_{n,m}) &= a_{n+1,m} - a_{n,m} = (n+1-j)^m - (n-j)^m \\
&= (n-j)^{m-1} (n+1-j - (n-j-m+1)) = m(n-j)^{m-1} = ma_{n,m-1},
\end{aligned}$$

beziehungsweise für die rechte Seite:

$$\begin{aligned}
\Delta(b_{n,m}) &= \sum_{l=0}^m \binom{m}{l} (-1)^l (j+\alpha)^l \left((n+1+\alpha-l)^{m-l} - (n+\alpha-l)^{m-l} \right) \\
&= \sum_{l=0}^{m-1} \binom{m}{l} (-1)^l (j+\alpha)^l (n+\alpha-l)^{m-1-l} (-l+m) \\
&= \sum_{l=0}^{m-1} \left(m \binom{m}{l} - m \binom{m-1}{l-1} \right) (-1)^l (j+\alpha)^l (n+\alpha-l)^{m-1-l} \\
&= m \sum_{l=0}^{m-1} \binom{m-1}{l} (-1)^l (j+\alpha)^l (n+\alpha-l)^{m-1-l} = mb_{n,m-1}.
\end{aligned}$$

Es erfüllen beide Größen die Rekursion $c_{n+1,m} = c_{n,m} + mc_{n,m-1}$ mit den Bedingungen $c_{j,m} = 0, \forall m \in \mathbb{N}$ und $c_{n,0} = 1, \forall n \in \mathbb{N}$. Daher stimmen sie überein. □

Aus diesem Lemma können wir eine Darstellung für eine Summe mit Brüchen gewisser Binomialkoeffizienten herleiten. Diese wird später beim Bestimmen der Größe der Nachfahren eines k -Baumes benötigt. Ein Zwischenresultat ist das folgende Lemma.

Lemma 1.11 (vgl. [18], Abschnitt 4.3). *Es sei $-\alpha, -\beta \in \mathbb{R} \setminus \mathbb{N}$ mit $\alpha - \beta \notin \mathbb{N}$. Für $n, j, m \in \mathbb{N}$ stimmt die Gleichung:*

$$\sum_{j=1}^n \frac{\binom{n-m+\alpha}{n-m-j}}{\binom{n+\beta}{n-j}} = \frac{n+1+\beta}{(m+1)\binom{m+\beta-\alpha+1}{m+1}} - \frac{1}{\binom{n+\beta}{n}} \sum_{l=0}^m \binom{m}{l} (-1)^l \frac{\beta+1}{l+\beta-\alpha+1} \binom{n+\alpha-l}{n}.$$

Beweis: Zuerst zeigen wir das Zwischenresultat $\frac{\binom{n-m+\alpha}{n-m-j}}{\binom{n+\beta}{n-j}} = \sum_{l=0}^m \binom{m}{l} (-1)^l \frac{\binom{j+\beta}{\beta-\alpha+l}}{\binom{n+\beta}{\beta-\alpha+l}}$:

$$\begin{aligned} \frac{\binom{n+\alpha-m}{n-m-j}}{\binom{n+\beta}{n-j}} &= \frac{\Gamma(n+\alpha-m+1)}{\Gamma(n+\beta+1)} (n-j)^m \frac{\Gamma(j+\beta+1)}{\Gamma(j+\alpha+1)} \\ &= \sum_{l=0}^m \binom{m}{l} (-1)^l \frac{\Gamma(n+\alpha-m+1)}{\Gamma(n+\beta+1)} \frac{\Gamma(n+\alpha-l+1)}{\Gamma(n+\alpha-m+1)} \\ &\quad \cdot \frac{\Gamma(j+\alpha+1)}{\Gamma(j+\alpha-l+1)} \frac{\Gamma(j+\beta+1)}{\Gamma(j+\alpha+1)} \\ &= \sum_{l=0}^m \binom{m}{l} (-1)^l \frac{\Gamma(n+\alpha-l+1)}{\Gamma(n+\beta+1)} \frac{\Gamma(j+\beta+1)}{\Gamma(j+\alpha-l+1)} \frac{\Gamma(\beta-\alpha+l+1)}{\Gamma(\beta-\alpha+l+1)} \\ &= \sum_{l=0}^m \binom{m}{l} (-1)^l \frac{\binom{j+\beta}{\beta-\alpha+l}}{\binom{n+\beta}{\beta-\alpha+l}}. \end{aligned}$$

Mit der Beziehung $\binom{a+b}{k} = \binom{a+b+1}{k+1} - \binom{a+b}{k+1}$ erhalten wir für die Summe:

$$\begin{aligned}
\sum_{j=1}^n \frac{\binom{n+\alpha-m}{n-m-j}}{\binom{n+\beta}{n-j}} &= \sum_{l=0}^m \binom{m}{l} \frac{1}{\binom{n+\beta}{\beta-\alpha+l}} \sum_{j=1}^n \binom{j+\beta}{\beta-\alpha+l} \\
&= \sum_{l=0}^m \binom{m}{l} \frac{1}{\binom{n+\beta}{\beta-\alpha+l}} \left(\binom{n+1+\beta}{\beta-\alpha+l+1} - \binom{\beta+1}{\beta-\alpha+l+1} \right) \\
&= \sum_{l=0}^m \binom{m}{l} \frac{\Gamma(n+\alpha-l+1)}{\Gamma(n+\beta+1)(l+\beta-\alpha+1)} \left(\frac{\Gamma(n+\beta+2)}{\Gamma(n+\alpha-l+1)} \right. \\
&\quad \left. - \frac{\Gamma(\beta+2)}{\Gamma(\alpha-l+1)} \right) \\
&= \frac{1}{\binom{n+\beta}{n}} \sum_{l=0}^m \binom{m}{l} \frac{\beta+1}{l+\beta-\alpha+1} \left(\frac{\Gamma(n+\beta+2)}{\Gamma(\beta+2)\Gamma(n+1)} \right. \\
&\quad \left. - \frac{\Gamma(n+\alpha-l+1)}{\Gamma(\alpha-l+1)\Gamma(n+1)} \right) \\
&= \frac{1}{\binom{n+\beta}{n}} \sum_{l=0}^m \binom{m}{l} \frac{\beta+1}{l+\beta-\alpha+1} \left(\binom{n+1+\beta}{n} - \binom{n+\alpha-l}{n} \right).
\end{aligned}$$

Die erste Summe können wir noch mit Lemma 1.9 umformen:

$$\begin{aligned}
\frac{\binom{n+1+\beta}{n}}{\binom{n+\beta}{n}} \sum_{l=0}^m \binom{m}{l} \frac{\beta+1}{l+\beta-\alpha+1} &= \frac{n+\beta+1}{\beta+1} \frac{\beta+1}{(m+\beta-\alpha+1)\binom{m+\beta-\alpha}{m}} \\
&= \frac{n+\beta+1}{(m+1)\binom{m+\beta-\alpha+1}{m+1}},
\end{aligned}$$

womit wir insgesamt das Gewünschte erhalten. \square

Bei der wahrscheinlichkeitserzeugenden Funktion für die Vorfahren tritt folgende Summe mehrere Male auf.

Lemma 1.12 (vgl. [18], Abschnitt 5.3). *Es sei $\alpha, \beta \in \mathbb{R}$ mit $-\alpha, -\beta \notin \mathbb{N}$ und $\beta - \alpha \notin \mathbb{N}^\times$. Dann gilt:*

$$\sum_{j=1}^n \frac{\binom{j-1+\alpha}{j-1}}{\binom{j-1+\beta}{j-1}} = \frac{\beta}{\alpha - \beta + 1} \left(\frac{\binom{n+\alpha}{n}}{\binom{n+\beta-1}{n}} - 1 \right).$$

Beweis: Um diese Gleichung zu zeigen, wird die Darstellung des Binomialkoeffizienten

über die Gamma-Funktion verwendet, sowie die Eigenschaft $\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1}$:

$$\begin{aligned}
\sum_{j=1}^n \frac{\binom{j-1+\alpha}{j-1}}{\binom{j-1+\beta}{j-1}} &= \sum_{j=1}^n \frac{\Gamma(j+\alpha)\Gamma(j)\Gamma(\beta+1)}{\Gamma(j+\beta)\Gamma(j)\Gamma(\alpha+1)} \\
&= \frac{\Gamma(\beta+1)\Gamma(\alpha-\beta+1)}{\Gamma(\alpha+1)} \sum_{j=1}^n \frac{\Gamma(j+\alpha)}{\Gamma(j+\beta)\Gamma(\alpha-\beta+1)} \\
&= \frac{1}{\binom{\alpha}{\alpha-\beta}} \sum_{j=0}^{n-1} \binom{j+\alpha}{\alpha-\beta} = \frac{1}{\binom{\alpha}{\alpha-\beta}} \left(\binom{n+\alpha}{\alpha-\beta+1} - \binom{\alpha}{\alpha-\beta+1} \right) \\
&= \frac{\Gamma(\alpha-\beta+1)\Gamma(\beta+1)}{\Gamma(\alpha+1)} \left(\frac{\Gamma(n+\alpha+1)}{\Gamma(\alpha-\beta+2)\Gamma(n+\beta)} - \frac{\Gamma(\alpha+1)}{\Gamma(\alpha-\beta+2)\Gamma(\beta)} \right) \\
&= \frac{\beta}{\alpha-\beta+1} \left(\frac{\Gamma(n+\alpha+1)\Gamma(\beta)}{\Gamma(\alpha+1)\Gamma(n+\beta)} - 1 \right) \\
&= \frac{\beta}{\alpha-\beta+1} \left(\frac{\binom{n+\alpha}{n}}{\binom{n+\beta-1}{n}} - 1 \right).
\end{aligned}$$

□

2 Abzählprobleme bei k -Bäumen

Wir werden zu Beginn eine Formel für die Anzahl markierter k -Bäume herausarbeiten. Der Beweis dafür basiert auf einer Arbeit von Moon [17]. Im Anschluss wird ein Beweis von Foata [9] skizziert, welcher eine Beziehung zu gewissen azyklischen Funktionen ausnützt. Des Weiteren werden verschiedene Variationen der markierten k -Bäume betrachtet, wobei unter anderem gewisse Bedingungen an die Markierung gestellt werden. In 2.8 und 2.9 werden noch zwei zu den k -Bäumen eng verwandte kombinatorische Strukturen betrachtet. Davon kann eine als Verallgemeinerung der k -Bäume aufgefasst werden.

2.1 Markierte k -Bäume

Der Beweis basiert, wie auch der Ursprüngliche von Beineke und Pippert [3], auf der Auszeichnung einer k -Clique als Wurzel des k -Baumes. Zuerst werden k -Bäume mit der Einschränkung, dass eine bestimmte Anzahl an Knoten mit der Wurzelclique verbunden ist, betrachtet.

Lemma 2.1 (vgl. [17], Theorem 2). *Es bezeichne $W_k(n, d)$ die Anzahl der k -Bäume mit n Knoten, bei denen genau d verschiedene Knoten mit einer ausgezeichneten k -Clique benachbart sind. Dann gilt für $1 \leq d \leq n - k$:*

$$W_k(n, d) = \binom{n-k-1}{d-1} (k(n-k))^{n-d-k}. \quad (2.1)$$

Beweis: Zuerst wird folgende Gleichung gezeigt:

$$W_k(n, d) = \binom{n-k}{d} \sum_{i=1}^{n-d-k} W_k(n-d, i) (kd)^i, \quad \text{falls } 1 \leq d \leq n-k. \quad (\text{I})$$

Wir zeigen dies, indem wir schrittweise ausgehend von der ausgezeichneten k -Clique W einen k -Baum betrachten. Mit dieser Clique sind d Knoten verbunden und es gibt $\binom{n-k}{d}$ Möglichkeiten, diese auszuwählen. Der Teilgraph C soll der erzeugte Graph von diesen $k+d$ Knoten sein. Wenn wir nun kurz C als eine einzige k -Clique K auffassen, ergibt sich mit den übrigen Knoten noch ein k -Baum der Größe $n-d$, wobei i Knoten mit $1 \leq i \leq n-d-k$ mit der Clique K adjazent sind. Für diesen übrigen k -Baum gibt es $W_k(n-d, i)$ Möglichkeiten, allerdings sind ursprünglich die i Knoten mit einer k -Clique aus C benachbart. Mit jedem Knoten, der W angehängt wird, erhöht sich die Anzahl der k -Cliquen in C um k . Insgesamt ergibt dies $(kd)^i$ Möglichkeiten, die i Knoten mit jeweils einer k -Clique von C zu verbinden und so folgt (I).

Nun führen wir eine Induktion nach n durch, um (2.1) zu zeigen. Für den Induktionsanfang $n = k+1$ stimmt die Formel. Nehmen wir nun an, sie gilt für alle n mit $n < m$ und $1 \leq d \leq n-1$. Es gilt für $1 \leq d \leq m-1$:

$$\begin{aligned}
W_k(m, d) &= \binom{m-k}{d} \sum_{i=1}^{m-d-k} W_k(m-d, i) (kd)^i \\
&\stackrel{\text{(IV)}}{=} \binom{m-k}{d} \sum_{i=1}^{m-d-k} \binom{m-d-k-1}{i-1} \underbrace{(k(m-d-k))}_{=:l}^{m-d-i-k} (kd)^i \\
&= \binom{m-k}{d} \sum_{i=1}^l \binom{l-1}{i-1} (kl)^{l-i} (kd)^i \\
&= \binom{m-k}{d} \sum_{i=0}^{l-1} \binom{l-1}{i} (kl)^{l-1-i} (kd)^{i+1} \\
&= \binom{m-k}{d} kd (kl + kd)^{l-1} = \binom{m-k}{d} kd (k(m-k))^{m-d-k-1} \\
&= \frac{kd(m-k)!}{k(m-k)d!(m-k-d)!} (k(m-k))^{m-d-k} = \binom{m-k-1}{d-1} (k(m-k))^{m-d-k}.
\end{aligned}$$

□

Mit dieser Vorarbeit kann das Abzählproblem für markierte k -Bäume gelöst werden.

Satz 2.2 (vgl. [17], Theorem 3). *Für $1 \leq k \leq n$ gilt:*

$$T_k(n) = \binom{n}{k} (k(n-k) + 1)^{n-k-2}. \quad (2.2)$$

Beweis: Wir bezeichnen mit $R_k(n)$ die Anzahl an markierten k -Bäumen, bei denen vorgegebene k Knoten als Clique R miteinander verbunden sind. Nach Lemma 2.1 gilt für diese:

$$\begin{aligned}
R_k(n) &= \sum_{d=1}^{n-k} W_k(n, d) = \sum_{d=1}^{n-k} \binom{n-k-1}{d-1} (k(n-k))^{n-d-k} \\
&= (k(n-k) + 1)^{n-k-1}.
\end{aligned} \quad (2.3)$$

Um die k Knoten der Clique R auszuwählen, gibt es $\binom{n}{k}$ Möglichkeiten. Wenn wir hingegen einen markierten k -Baum betrachten, enthält dieser nach Lemma 1.1 $k(n-k) + 1$ k -Cliquen und wir erhalten die Gleichung

$$\binom{n}{k} R_k(n) = (k(n-k) + 1) T_k(n),$$

welche sich zu

$$T_k(n) = \binom{n}{k} (k(n-k) + 1)^{-1} R_k(n) = \binom{n}{k} (k(n-k) + 1)^{n-k-2}$$

umformen lässt.

□

Für den Fall $k = 1$ stimmt Formel (2.2) natürlich mit Caley's Formel überein.

Korollar 2.3. *Es gibt n^{n-2} markierte Bäume.*

Die exponentiell erzeugende Funktion der markierten k -Bäume werden wir später noch für Abschätzungen brauchen und daher werden wir sie an dieser Stelle ein wenig genauer studieren. Insbesondere möchten wir zeigen, dass die Potenzreihe einen Konvergenzradius größer 0 hat und somit auch holomorph in einem Bereich um 0 ist. Wir bezeichnen die exponentiell erzeugende Funktion mit $T(z)$, wobei diese natürlich auch von dem Parameter k abhängig ist.

$$T(z) = \sum_{n=k}^{\infty} \binom{n}{k} (k(n-k) + 1)^{n-k-2} \frac{z^n}{n!} = \sum_{n=k}^{\infty} \underbrace{\frac{(k(n-k) + 1)^{n-k-2}}{k!(n-k)!}}_{:= a_n} z^n.$$

Der Konvergenzradius stimmt mit $(\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|})^{-1}$ überein. Der Radius wird nach unten abgeschätzt. Dafür möchten wir

$$\lim_{n \rightarrow \infty} \sqrt[n]{n} = \lim_{n \rightarrow \infty} \exp\left(\frac{1}{n} \log n\right) = \exp\left(\lim_{n \rightarrow \infty} \frac{1}{n} \log n\right) \stackrel{\text{de l'Hospital}}{=} \exp\left(\lim_{n \rightarrow \infty} \frac{1}{n}\right) = 1$$

und die Stirling-Formel

$$n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n, \quad n \rightarrow \infty$$

in Erinnerung rufen. Um den Radius nach unten abzuschätzen, wird nur der Limes Superior betrachtet, welcher dementsprechend nach oben abgeschätzt werden muss:

$$\begin{aligned} \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} &= \limsup_{n \rightarrow \infty} \sqrt[n]{\frac{(k(n-k) + 1)^{n-k-2}}{k!(n-k)!}} \\ &= \limsup_{n \rightarrow \infty} \sqrt[n]{\frac{(k(n-k) + 1)^{n-k-2} e^{n-k}}{k!(n-k)^{n-k} \sqrt{2\pi(n-k)}}}. \end{aligned}$$

Für n groß genug ist der Bruch innerhalb der Wurzel sicherlich größer als 1. Indem wir nur die $(n-k)$ -te Wurzel ziehen, schätzen wir nach oben ab. Sei $m = n - k$, dann ist

$$\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} \leq \limsup_{m \rightarrow \infty} \sqrt[m]{\frac{(km+1)^{m-2} e^m}{k! m^m \sqrt{2\pi m}}} = \limsup_{m \rightarrow \infty} \frac{e(km+1)}{m} = ek \quad (2.4)$$

und der Konvergenzradius von $T(z)$ größer gleich $\frac{1}{ek}$. So repräsentiert die exponentiell erzeugende Funktion nicht nur eine formale Potenzreihe, sondern auch eine in der komplexen Ebene um 0 holomorphe Funktion.

2.1.1 Weitere Beweise

Markierte Bäume sind natürlich ein Spezialfall der markierten k -Bäume und können dementsprechend ähnlich rekursiv definiert werden. Um die Formel von Caley zu beweisen, kann auch ein Zusammenhang mit Funktionen auf endlichen Mengen ausgenützt werden. Interessanterweise gibt es auch so einen Beweisansatz für k -Bäume. Wir wollen diesen Ansatz, welcher auf Foata [9] zurückgeht, skizzenhaft ausführen, um die Ähnlichkeit der Strukturen von Wurzelbäumen und k -Wurzelbäumen zu verdeutlichen. Auf technische Details wird verzichtet. Für einen vollständigen Beweis sei auf die Arbeit von Foata ([9]), dessen Notation wir im nächsten Abschnitt übernehmen, verwiesen. Das Kernstück bilden sogenannte (X, Y) -azyklische Funktionen.

Definition 2.1 (vgl. [9], §2). Es seien A, B, C endliche disjunkte Mengen und $X = A \cup C$, $Y = B \cup C$ und $|X| = n \in \mathbb{N}^\times$. Die Funktion $\gamma : Y \rightarrow X$ sei surjektiv und erfülle $\gamma(c) = c$, $\forall c \in C$ und $\gamma(B) = A$.

Eine Funktion $f : X \rightarrow X$ nennen wir azyklisch, falls $f^n = f^{n-1}$ gilt. Die Funktion $g : X \rightarrow Y$ wird als (X, Y) -azyklisch bezeichnet, falls $\gamma \circ g$ azyklisch ist und die Menge der Fixpunkte mit C übereinstimmt. Die Menge dieser Funktionen wird mit $F(A, B, C)$ bezeichnet.

Mit einer Bijektion kann $F(A, B, C)$ in eine andere Funktionenmenge abgebildet werden, welche sehr leicht abgezählt werden kann. So erhält man $F(A, B, C) = |C|(|B| + |C|)^{|A|-1}$ (vgl. [9], Theorem 1).

Die markierten k -Wurzelbäume können nun mit (X, Y) -azyklischen Funktionen identifiziert werden. Dazu bezeichnen wir die Knoten von der Wurzel weg, das heißt $c = \{x_1, x_2, \dots, x_k\}$ sei die Wurzel und für $j \in \{k+1, k+2, \dots, n\}$ ist x_j mit genau einer k -Clique K_j aus $\{x_1, x_2, \dots, x_{j-1}\}$ verbunden. Sei $A = \{x_{k+1}, x_{k+2}, \dots, x_n\}$, $C = \{c\}$ und B die Menge aller k -Cliquen des k -Baumes ohne die Wurzel. Die Funktion g ordnet jedem Knoten x_j aus A die k -Clique K_j zu. γ bildet jede k -Clique auf den beim rekursiven Aufbau als letztes hinzugefügten Knoten ab, dies ist in unserer Bezeichnung der Knoten der Clique mit dem höchsten Index. Der Definitionsbereich beider Funktionen umfasst natürlich auch noch die Wurzel, welche ein Fixpunkt beider Funktionen ist. Obwohl bei unserer Beschreibung der Funktion die Aufzählung der Knoten verwendet wird, sind g und γ nur von der Wurzelclique abhängig und nicht von unserer Nummerierung. Für das Bild eines Knotens $x_j \in A$ gilt entweder $g(x_j) = c$ oder $g(x_j) \in \gamma^{-1}(x_i)$ mit $k < i < j$. Damit ist $(\gamma \circ g)^{n-k}(A) = \{c\}$ und insbesondere ist $\gamma \circ g$ eine azyklische Funktion mit Fixpunkt c . Wurzelbäume entsprechen also (X, Y) -azyklischen Funktionen. Wie im Beweis von Satz 2.2 ergibt dies (vgl. [9], §3):

$$R_k(n) = |C|(|B| + |C|)^{|A|-1} = (k(n-k) + 1)^{n-k-1}.$$

2.2 Markierte k -Wurzelbäume

Im Beweis von Satz 2.2 sind k -Bäume aufgetreten, bei denen eine Clique ausgezeichnet war. Diese werden wir nun analog zu den Bäumen als Wurzelclique bezeichnen. Insbesondere wird beim rekursiven Aufbau eines k -Baumes mit der Wurzelclique gestartet, bei

der dann weitere Knoten angehängt werden. Die k -Bäume mit einer Wurzelclique werden im Folgenden als k -Wurzelbäume bezeichnet. Bei den k -Wurzelbäumen werden nur die Knoten, die nicht zur Wurzelclique gehören, gezählt. Dies vereinfacht die Darstellung durch kombinatorische Strukturen. Die Knoten der Wurzelclique werden wir mit $0_1, \dots, 0_k$ markieren.

Es sei \mathcal{R} die kombinatorische Struktur der k -Wurzelbäume und \mathcal{S} umfasse die Elemente von \mathcal{R} , bei welchen die Wurzelclique genau ein Kind besitzt. Wir können nun jedes Element von \mathcal{R} , welches l Kinder hat, als Menge von l Elementen von \mathcal{S} identifizieren, indem wir die Wurzelcliquen gleichsetzen und unter Berücksichtigung der Ordnung die Knoten ummarkieren. Wenn wir hingegen ein Element von \mathcal{S} betrachten, besteht dieses aus einer Wurzelclique, welche mit genau einem Knoten verbunden ist. Dieser Graph aus $k + 1$ Knoten besitzt k k -Cliques, welche nicht mit der Wurzelclique übereinstimmen. Diese können nun als Wurzel eines k -Wurzelbaumes interpretiert werden, um \mathcal{S} durch \mathcal{R} zu beschreiben. Bei dem Produkt der k Elemente von \mathcal{R} kommt es natürlich auf die Reihenfolge an, da sie je nach Abfolge mit unterschiedlichen Cliques verbunden werden (vgl. [19], Abschnitt 2). Diese Überlegungen ergeben die Gleichungen

$$\begin{aligned}\mathcal{R} &= \text{SET}(\mathcal{S}), \\ \mathcal{S} &= \mathcal{Z} \star \mathcal{R}^k,\end{aligned}$$

wobei \mathcal{Z} nur aus einem Element der Größe 1 besteht. Für die exponentiell erzeugenden Funktionen bedeutet dies:

$$\begin{aligned}R(z) &= \exp(S(z)), \\ S(z) &= zR^k(z) = z \exp(kS(z)).\end{aligned}$$

Mit der Beziehung $\frac{S}{\exp(kS)} = z$ kann die Lagrange'sche Inversionsformel benützt werden, um die Koeffizienten von \mathcal{R} zu bestimmen:

$$\begin{aligned}r_m &= m![z^m]R(z) = m![z^m] \exp(S(z)) \stackrel{(\text{LIF})}{=} m! \frac{1}{m} [u^{m-1}] \exp(kum) \exp(u) \\ &= (m-1)! [u^{m-1}] \exp(u(km+1)) = (m-1)! \frac{(km+1)^{m-1}}{(m-1)!} \\ &= (km+1)^{m-1}.\end{aligned} \tag{2.5}$$

Dies liefert mit $m = n - k$ in der in Abschnitt 2.1 verwendeten Bezeichnung dasselbe Ergebnis wie in Formel (2.3). Für die Hilfsgröße \mathcal{S} können wir dies analog durchführen:

$$\begin{aligned}s_m &= m![z^m]S(z) \stackrel{(\text{LIF})}{=} m! \frac{1}{m} [u^{m-1}] \exp(kum) \\ &= (m-1)! \frac{(km)^{m-1}}{(m-1)!} = (km)^{m-1}.\end{aligned}$$

2.3 Ansteigende k -Bäume

In diesem Abschnitt werden k -Bäume betrachtet, bei denen die Markierungen der Kinder einer Clique größer als alle Markierungen der Elternclique sind. Dies führt dazu, dass beim

rekursiven Aufbau der erste Knoten, welcher nicht zur Wurzelclique gehört, mit einer 1 markiert wird. Insbesondere erhält man eine aufsteigende Folge von Zahlen, wenn man einen Pfad von der Wurzel weg betrachtet, bei dem der darauffolgende Knoten immer ein Kind des Vorherigen ist. Wir werden diese k -Bäume als (markierte) ansteigende k -Bäume bezeichnen. Dies wird motiviert durch die englische Bezeichnung „increasing Caley tree“ für den Fall $k = 1$ (vgl. [8], S. 145) beziehungsweise „increasing k -tree“ für den allgemeinen Fall (vgl. [18], Abschnitt 2).

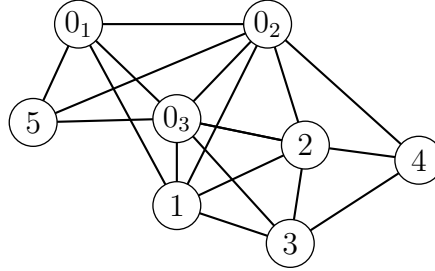


Abbildung 5: Ein ansteigender 3-Baum der Größe 5.

Bei der Beschreibung durch kombinatorische Strukturen ändert sich im Vergleich zu den Wurzelbäumen nur die rekursive Beschreibung von \mathcal{S} . Der erste Knoten muss mit 1 markiert werden, was durch das Box-Produkt beschrieben wird:

$$\begin{aligned}\mathcal{T} &= \text{SET}(\mathcal{S}), \\ \mathcal{S} &= \mathcal{Z}^\square \star \mathcal{T}^k.\end{aligned}$$

Übersetzt für die exponentiell erzeugenden Funktionen heißt das

$$T(z) = \exp(S(z)), \quad (2.6)$$

$$S(z) = \int_0^z 1 \cdot T^k(t) dt \quad \Leftrightarrow \quad S'(z) = T^k(z), \quad S(0) = 0. \quad (2.7)$$

Wir möchten darauf hinweisen, dass die formalen Potenzreihen $T(z)$ und $S(z)$ einen Konvergenzradius größer als 0 haben, da ihre Koeffizienten kleiner als die entsprechenden Werte in Satz 2.2 sind. Bei den markierten k -Bäumen ergibt sich nach (2.4) ein Konvergenzradius größer gleich $\frac{1}{ek}$. Dies ist auch eine untere Schranke für die Konvergenzradien von S und T . So haben wir analytische Funktionen und können gewöhnlich differenzieren. Die Gleichung (2.6) in (2.7) eingesetzt, führt zu

$$S' = \exp(kS) \quad \Leftrightarrow \quad S' \exp(-kS) = 1,$$

was integriert

$$\begin{aligned}\int S' \exp(-kS) &= \int (\exp(-kS)(-k)^{-1})' dz = z + \tilde{c}, \\ (-k)^{-1} \exp(-kS) &= z + \tilde{c}, \\ -kS &= \log(k(c - z)), \\ S(z) &= \left(-\frac{1}{k}\right) \log(k(c - z))\end{aligned}$$

ergibt. Die Integrationskonstante c kann mit Hilfe von $S(0) = 0$ bestimmt werden:

$$\begin{aligned} 0 &= \log(kc), \\ c &= k^{-1}. \end{aligned}$$

Die Funktionen $S(z)$ und $T(z)$ lassen sich nun darstellen als

$$\begin{aligned} S(z) &= \left(-\frac{1}{k}\right) \log(1 - kz) = \log\left((1 - kz)^{-\frac{1}{k}}\right), \\ T(z) &= \exp(S(z)) = (1 - kz)^{-\frac{1}{k}}. \end{aligned}$$

Für das kombinatorische Abzählproblem von \mathcal{T} bedeutet dies:

$$\begin{aligned} t_n &= n![z^n]T(z) = n! \binom{-\frac{1}{k}}{n} (-1)^n k^n \\ &= n! k^n (-1)^n \frac{(-\frac{1}{k})(-\frac{1}{k}-1)\cdots(-\frac{1}{k}-n+1)}{n!} \\ &= n! k^n \frac{(n-1+\frac{1}{k})\cdots\frac{1}{k}}{n!} \\ &= n! k^n \binom{n-\frac{k-1}{k}}{n} = k^n \left(n-1+\frac{1}{k}\right)^n, \end{aligned}$$

sowie für die Hilfsgröße \mathcal{S} , für $n \geq 1$:

$$\begin{aligned} s_n &= n![z^n]S(z) = n![z^n] \frac{1}{k} \sum_{l=0}^{\infty} \frac{(kz)^n}{n} \\ &= n! \frac{1}{k} \frac{k^n}{n} = (n-1)! k^{n-1}. \end{aligned}$$

Wir haben mit dieser Forderung an die Markierung die Familie der k -Bäume stark eingeschränkt. So können wir beim rekursiven Aufbau davon ausgehen, dass wir zuerst den Knoten mit der Markierung 1 anhängen usw. Dies bedeutet, wir können für t_n eine Rekursion der Ordnung 1 aufstellen. Wir müssen dazu nur wissen, wie viele Cliques ein ansteigender k -Baum hat. Die Markierung hat keinen Einfluss auf die Anzahl der Cliques und somit stimmt diese nach Lemma 1.1 mit $(1 + k(n-1))$ für einen ansteigenden k -Baum der Größe $n-1$ ($= n-1+k$ Knoten) überein. Der mit n markierte Knoten kann nun ein Kind von jeder dieser Cliques sein und es ergeben sich immer unterschiedliche Bäume, womit sich die Rekursion $t_n = (1 + k(n-1))t_{n-1}$ ergibt. Mit $t_1 = 1$ bekommen wir die Produktdarstellung (vgl. [18], Proposition 1)

$$t_n = \prod_{l=0}^{n-1} (1 + kl).$$

2.4 Geordnete Wurzelbäume

Bei den geordneten k -Wurzelbäumen werden im Gegensatz zu den k -Wurzelbäumen die Kinder einer Clique linear geordnet. Das heißt, wir können von einem ersten, zweiten

Kind etc. sprechen. Die Markierung der Knoten hat keinen Einfluss auf die Ordnung der Kinder. Für zwei Kinder einer Clique kann aufgrund der Markierung noch nicht auf deren Reihenfolge in der Ordnung geschlossen werden. Die Totalordnung wird graphisch so umgesetzt, dass das erste Kind ganz links ist und jedes Element größer ist als alle Elemente, die sich links davon befinden.

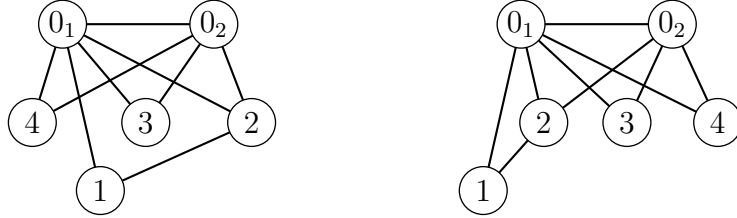


Abbildung 6: Unterschiedliche geordnete 2-Wurzelbäume der Größe 4. Wenn man auf die Ordnung der Kinder verzichtet, stellen sie den gleichen 2-Wurzelbaum dar.

Für die kombinatorische Beschreibung der Struktur bedeutet dies, dass im Vergleich zu den k -Wurzelbäumen die Struktur \mathcal{T} der geordneten k -Wurzelbäume aus einer Folge von geordneten k -Wurzelbäumen \mathcal{S} , mit genau einem Kind besteht:

$$\begin{aligned}\mathcal{T} &= \text{SEQ}(\mathcal{S}), \\ \mathcal{S} &= \mathcal{Z} \star \mathcal{T}^k,\end{aligned}$$

wobei \mathcal{Z} nur aus einem Element der Größe 1 besteht. Dies lässt sich für die exponentiell erzeugenden Funktionen in

$$\begin{aligned}T(z) &= \frac{1}{1 - S(z)}, \\ S(z) &= zT^k(z) = \frac{z}{(1 - S(z))^k}\end{aligned}$$

übersetzen. Wie bei den k -Wurzelbäumen können wir auch hier mit der Lagrange'schen Inversionformel über die Beziehung $z = \frac{S}{(1-S)^k}$ die Koeffizienten der exponentiell erzeugenden Funktion bestimmen. Für $n \geq 1$ gilt:

$$\begin{aligned}t_n &= n![z^n]T(z) = n![z^n]\frac{1}{1 - S(z)} \stackrel{\text{(LIF)}}{=} n!\frac{1}{n}[u^{n-1}](1 - u)^{-kn} \frac{1}{(1 - u)^2} \\ &= (n - 1)![u^{n-1}](1 - u)^{-(kn+2)} = (n - 1)![u^{n-1}] \sum_{m=0}^{\infty} \binom{kn + 2 + m - 1}{kn + 2 - 1} u^m \\ &= (n - 1)! \binom{(k + 1)n}{kn + 1} = (n - 1)! \frac{((k + 1)n)!}{(kn + 1)!(n - 1)!} \\ &= \frac{((k + 1)n)!}{(kn + 1)!}, \\ s_n &= n![z^n]S(z) \stackrel{\text{(LIF)}}{=} n!\frac{1}{n}[u^{n-1}](1 - u)^{-kn} = (n - 1)! \binom{kn + n - 1 - 1}{kn - 1} \\ &= (n - 1)! \binom{(k + 1)n - 2}{kn - 1} = \frac{((k + 1)n - 2)!}{(kn - 1)!}.\end{aligned}$$

2.5 Geordnete ansteigende k -Bäume

Bei den geordneten ansteigenden k -Bäumen (engl. „ordered increasing k -trees“) werden zusätzlich zu den ansteigenden k -Bäumen auch noch die Kinder linear geordnet. Das heißt, wir können bei einer Clique beispielsweise von einem ersten und zweiten Kind sprechen (vgl. [18], Abschnitt 2). Für die Markierungen der Knoten gelten die gleichen Bedingungen wie bei den ansteigenden k -Bäumen. In der folgenden Abbildung wird die Reihenfolge der Kinder durch die graphische Reihenfolge von links nach rechts umgesetzt.

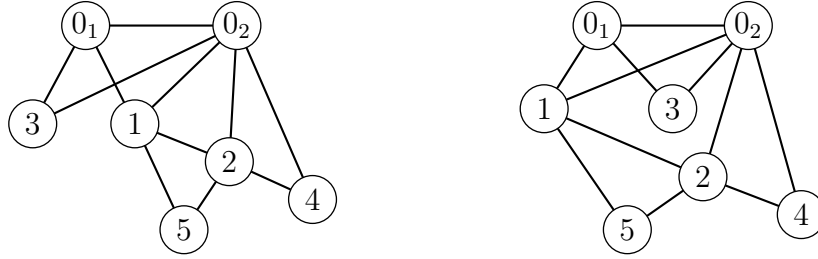


Abbildung 7: Unterschiedliche geordnete ansteigende 2-Bäume der Größe 5, welche den gleichen ansteigenden 2-Baum darstellen.

Für die Beschreibung der kombinatorischen Strukturen hat dies zur Folge, dass die geordneten ansteigenden k -Bäume \mathcal{T} nicht mehr aus einer Menge von einfacheren Objekten, sondern aus einer Folge von Objekten aus \mathcal{S} besteht. Die Struktur \mathcal{S} umfasst wieder die Teilmenge von \mathcal{T} , bei denen die Wurzelclique nur ein Kind besitzt (vgl. [19], Abschnitt 2). Es ergeben sich die Gleichungen

$$\begin{aligned}\mathcal{T} &= \text{SEQ}(\mathcal{S}), \\ \mathcal{S} &= \mathcal{Z}^\square \star \mathcal{T}^k.\end{aligned}$$

Dies lässt sich auch in Gleichungen für die exponentiell erzeugenden Funktionen überführen. Man beachte wieder, dass die Koeffizienten von $T(z)$ und $S(z)$ sicher kleiner als die von den gewöhnlichen k -Bäumen sind. So haben wir jeweils Potenzreihen mit Konvergenzradius größer als 0, welche somit auch differenzierbar sind.

$$T(z) = \frac{1}{1 - S(z)}, \quad (2.8)$$

$$S(z) = \int_0^z 1 \cdot T^k(t) dt \quad \Leftrightarrow \quad S'(z) = T^k(z), \quad S(0) = 0. \quad (2.9)$$

Die Gleichung (2.8) führt zu

$$\begin{aligned}(1 - S(z))T(z) &= 1, & \left| \frac{d}{dz} \right. \\ -S'T + (1 - S)T' &= 0.\end{aligned}$$

Mit (2.9) kann dies vereinfacht werden:

$$\Rightarrow \quad -T^{k+1} + \frac{T'}{T} = 0 \quad \Leftrightarrow \quad T' = T^{k+2} \quad \Leftrightarrow \quad \frac{T'}{T^{k+2}} = 1.$$

Da die Ableitung von $T(z)$ an der Stelle 0 verschieden von 0 ist ($t_1 = 1$), kann $T(z)$ lokal invertiert werden. Mit der Substitution $u = T(x)$ kann die Stammfunktion bestimmt werden:

$$\begin{aligned} \int 1 dz &= \int \frac{T'}{T^{k+2}} dz, \\ z + C &= \int \frac{1}{u^{k+2}} du = -\frac{u^{-(k+1)}}{k+1}, \\ \Rightarrow T^{k+1}(z) &= -\frac{1}{(z+C)(k+1)}. \end{aligned}$$

Es gibt nur einen geordneten ansteigenden k -Baum der Größe 0 und dies bedeutet für die Integrationskonstante:

$$\begin{aligned} 1 &= -\frac{1}{C(k+1)} \quad \Leftrightarrow \quad C(k+1) = -1, \\ T^{k+1}(z) &= \frac{1}{1 - (k+1)z}, \\ T(z) &= (1 - (k+1)z)^{-\frac{1}{k+1}}. \end{aligned}$$

Eingesetzt in (2.9) ergibt dies für die erzeugende Funktion der geordneten ansteigenden k -Bäume mit einer Wurzel, die genau ein Kind hat:

$$S(z) = 1 - \frac{1}{T(z)} = 1 - (1 - (k+1)z)^{\frac{1}{k+1}}.$$

Die Koeffizienten können auf die übliche Weise mit Hilfe der binomischen Reihe berechnet werden:

$$\begin{aligned} T(z) &= (1 - (k+1)z)^{-\frac{1}{k+1}} = \sum_{n=0}^{\infty} \binom{-\frac{1}{k+1}}{n} (-1)^n ((k+1)z)^n, \\ t_n &= n![z^n]T(z) = n!(k+1)^n (-1)^n \binom{-\frac{1}{k+1}}{n} \\ &= n!(k+1)^n (-1)^n \frac{(-\frac{1}{k+1})(\frac{1}{k+1}-1)\cdots(\frac{1}{k+1}-n+1)}{n!} \\ &= n!(k+1)^n \frac{(n-1+\frac{1}{k+1})\cdots\frac{1}{k+1}}{n!} \\ &= n!(k+1)^n \binom{n-\frac{k}{k+1}}{n} = \left(n - \frac{k}{k+1}\right)^n (k+1)^n. \end{aligned}$$

Analog geht dies für s_n mit $n \geq 1$. Elemente der Größe 0 gibt es in \mathcal{S} keine, da immer

ein Knoten mit 1 markiert werden muss.

$$\begin{aligned}
S(z) &= 1 - \sum_{n=0}^{\infty} \binom{\frac{1}{k+1}}{n} (-1)^n ((k+1)z)^n, \\
s_n &= n! [z^n] S(z) = n! (-1)^n \binom{\frac{1}{k+1}}{n} (-1)^n (k+1)^n \\
&= n! (k+1)^n (-1)^{n-1} \frac{\frac{1}{k+1} (\frac{1}{k+1} - 1) \cdots (\frac{1}{k+1} - n + 1)}{n!} \\
&= n! (k+1)^{n-1} \frac{(n-1 - \frac{1}{k+1}) \cdots (1 - \frac{1}{k+1})}{n!} \\
&= (n-1)! (k+1)^{n-1} \binom{n-1 - \frac{1}{k+1}}{n-1} \\
&= \left(n-1 - \frac{1}{k+1} \right)^{n-1} (k+1)^{n-1}.
\end{aligned}$$

Wie bei den ansteigenden k -Bäumen ist es möglich, eine Rekursion der Ordnung 1 für die geordneten ansteigenden k -Bäume aufzustellen. Dazu schauen wir uns an, wie sich die Anzahl der möglichen Plätze, bei denen ein Knoten hinzugefügt werden kann, sich beim Hinzufügen eines Knotens verändert. Zum einen entstehen k neue k -Cliques, was auch k neue mögliche Plätze zur Folge hat. Zum anderen gibt es auch einen neuen Platz bei der k -Clique, an welcher der Knoten angefügt wird. Die Anzahl der möglichen Plätze erhöht sich mit jedem Knoten um $k+1$. Der n -te Knoten kann mit der Wurzel verbunden werden oder nimmt einen der Plätze, welche die vorigen Knoten erzeugt haben, ein. Für die Rekursion bedeutet dies: $t_n = (1 + (k+1)(n-1))t_{n-1}$, mit dem Ergebnis (vgl. [18], Proposition 1):

$$t_n = \prod_{l=0}^{n-1} (1 + (k+1)l).$$

2.6 Markierte d -äre k -Bäume

Wir können wie bei den Bäumen nur solche k -Bäume betrachten, welche eine bestimmte maximale Anzahl an Kindern pro Clique besitzen. Wir stellen uns das so vor, dass jede Clique eine bestimmte Anzahl $d \in \mathbb{N}$ an geordneten freien Plätzen für die Kinder hat. Diese k -Bäume nennen wir markierte d -äre k -Bäume.

Ähnlich zu vorigen Abschnitten bezeichnet \mathcal{S} die Struktur der d -ären k -Bäume mit einem oder keinem Kind der Wurzelclique. So können die d -ären k -Bäume \mathcal{T} als Produkt von d Elementen aus \mathcal{S} aufgefasst werden. Für die kombinatorischen Strukturen ergibt dies die Gleichungen

$$\begin{aligned}
\mathcal{T} &= \mathcal{S}^d, \\
\mathcal{S} &= \{\emptyset\} + \mathcal{Z} \star \mathcal{T}^k.
\end{aligned}$$

Für die exponentiell erzeugenden Funktionen stimmt dies mit

$$\begin{aligned}
T(z) &= S^d(z), \\
S(z) &= 1 + zT^k(z)
\end{aligned}$$

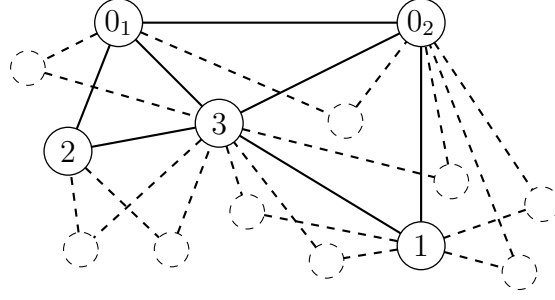


Abbildung 8: Ein binärer 2-Baum der Größe 3, bei dem die freien Plätze eingezeichnet sind. So sehen wir, dass beispielsweise der mit 3 markierte Knoten das erste Kind der Wurzelclique und der Knoten mit der 2 das zweite Kind beziehungsweise das Kind an der zweiten Position der Clique $\{0_1, 3\}$ ist.

überein. Für $S(z)$ und $T(z)$ ergeben sich durch gegenseitiges Einsetzen zwei Fixpunktgleichungen:

$$\begin{aligned} T(z) &= (1 + zT^k(z))^d, \\ S(z) &= 1 + zS^{dk}(z). \end{aligned} \tag{2.10}$$

Üblicherweise löst man im Zusammenhang mit kombinatorischen Strukturen solche Gleichungen mit der Lagrange'schen Inversionsformel. Dies ist im vorliegenden Fall nicht direkt möglich, da der konstante Term der Potenzreihen S und T ungleich 0 ist. So können S und T auch nicht als Produkt $z\varphi(\cdot)$ mit einer Potenzreihe φ geschrieben werden. Allerdings lässt sich dies durch eine Substitution $\tilde{S} = S - 1$ erreichen. Für die Potenzreihe \tilde{S} gilt:

$$\tilde{S} = z(\tilde{S} + 1)^{dk} \quad \Leftrightarrow \quad z = \frac{\tilde{S}}{(\tilde{S} + 1)^{dk}}.$$

Für $n \geq 1$ stimmen die Koeffizienten von S und \tilde{S} überein und diese können mit der Lagrange'schen Inversionsformel berechnet werden. Sei $n \geq 1$:

$$\begin{aligned} s_n &= n![z^n]S(z) = n![z^n]\tilde{S}(z) \stackrel{\text{(LIF)}}{=} \frac{n!}{n}[u^{n-1}](1+u)^{ndk} \\ &= (n-1)! \binom{ndk}{n-1}. \end{aligned}$$

Ein markierter d -ärer k -Baum, der nur aus der Wurzel besteht, ist der einzige d -äre k -Baum der Größe 0 und daher $t_0 = 1$. Mit Hilfe der Funktion \tilde{S} können auch die übrigen Koeffizienten von T berechnet werden. Für $n \geq 1$ gilt:

$$\begin{aligned} t_n &= n![z^n]T(z) = n![z^n]S^d(z) = n![z^n](1 + \tilde{S})^d \stackrel{\text{(LIF)}}{=} \frac{n!}{n}[u^{n-1}]d(1+u)^{d-1}(1+u)^{ndk} \\ &= (n-1)!d[u^{n-1}](1+u)^{ndk+d-1} = (n-1)!d \binom{ndk+d-1}{n-1}. \end{aligned}$$

Die Koeffizienten von S und T sind abgesehen von einer Multiplikation mit $n!$ die Fuss-Catalan-Zahlen. Diese lösen, wie Graham, Knuth und Patashnik in [11] angeführt haben,

die Frage „Wie schauen die Koeffizienten von $G(z)$ mit $G(z) = zG^m(z) + 1$ aus?“ (vgl. [11], S. 347). Die Antwort würde mit entsprechendem Parameter auch Formel (2.10) erfüllen. In [11] werden allerdings die Fuss-Catalan-Zahlen mit einem Parameter weniger, als wir es im Folgenden tun, betrachtet ($r = 1$).³

Definition 2.2 (vgl. [16], S. 940f). Die Fuss-Catalan-Zahlen $A_m(p, r)$ mit den Parametern $p, r \in \mathbb{R}$ sind für $m \in \mathbb{N}$ definiert durch: $A_0(p, r) := 1$ und

$$A_m(p, r) := \frac{r}{m!} \prod_{i=1}^{m-1} (mp + r - i).$$

Falls $mp + r \neq 0$ ist, können die Fuss-Catalan-Zahlen auch als $\frac{r}{mp+r} \binom{mp+r}{m}$ geschrieben werden.

Die Fuss-Catalan-Zahlen erfüllen selbst eine Rekursion und es gelten weitere Beziehungen zwischen den verschiedenen Parametern. Des Weiteren erfüllt die erzeugende Funktion eine Gleichung, die von der gleichen Bauart wie (2.10) ist. Auf das Nachrechnen des folgenden Lemmas wird verzichtet und für einen Beweis sei auf die Literatur verwiesen (siehe [16], S. 941ff).

Lemma 2.4. *Es sei $B_{p,r} = \sum_{m=0}^{\infty} A_m(p, r) z^m$ die erzeugende Funktion der Fuss-Catalan-Zahlen. Dann gilt:*

1. $B_{p,1}(z)^r = B_{p,r}(z)$,
2. $B_{p,1}(z) = 1 + zB_{p,1}(z)^p = 1 + zB_{p,p}(z)$,
3. $B_{p,r}(z) = \left(1 + zB_{p,r}(z)^{\frac{p}{r}}\right)^r$.

In der Schreibweise der Fuss-Catalan-Zahlen bedeutet dies für die exponentiell erzeugenden Funktionen $S(z)$ und $T(z)$:

$$\begin{aligned} T(z) &= B_{kd,d}(z), \\ S(z) &= B_{kd,1}(z), \end{aligned}$$

beziehungsweise für ihre Koeffizienten:

$$\begin{aligned} t_n &= n!A_n(kd, d) = \frac{n!d}{nkd + d} \binom{nkd + d}{n}, \\ s_n &= n!A_n(kd, 1) = \frac{n!}{nkd + 1} \binom{nkd + 1}{n}. \end{aligned}$$

2.7 Ansteigende d -äre k -Bäume

Zusätzlich zu den d -ären k -Bäumen werden wir hier eine aufsteigende Reihenfolge der Markierungen voraussetzen. Die k -Bäume, welche wir so erhalten, nennen wir ansteigende d -äre k -Bäume (engl. „ d -ary increasing k -trees“, vgl. [18], Abschnitt 2).

³In [11] erscheint unsere Definition der Fuss-Catalan-Zahlen als Ausdruck im Zusammenhang mit den Fuss-Catalan-Zahlen, allerdings werden sie nicht als solche bezeichnet (siehe S. 349).

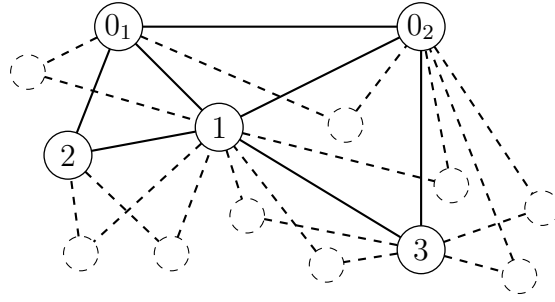


Abbildung 9: Ein ansteigender binärer 2-Baum der Größe 3, bei dem auch die freien Plätze eingezeichnet sind.

So lautet die Beziehung für die kombinatorischen Strukturen

$$\begin{aligned}\mathcal{T} &= \mathcal{S}^d, \\ \mathcal{S} &= \{\emptyset\} + \mathcal{Z}^\square \star \mathcal{T}^k,\end{aligned}$$

welche sich für die exponentiell erzeugenden Funktionen in

$$\begin{aligned}T(z) &= S^d(z), \\ S(z) &= 1 + \int_0^z 1 \cdot T^k(t) dt \quad \Leftrightarrow \quad S'(z) = T^k(z), \quad S(0) = 1\end{aligned}$$

übersetzen lässt. Diese Differentialgleichung lässt sich wieder mit der üblichen Methode lösen:

$$S' = S^{kd} \quad \Leftrightarrow \quad \frac{S'}{S^{kd}} = 1.$$

Integration auf beiden Seiten der Gleichung liefert mit der Substitution $u = S(z)$:

$$\begin{aligned}\int 1 dz &= \int \frac{S'}{S^{kd}} dz = \int u^{-kd} du, \\ z + c &= \frac{1}{1 - kd} u^{1 - kd} = \frac{1}{1 - kd} S^{1 - kd}, \\ \Rightarrow S(z) &= ((1 - kd)(z + c))^{\frac{1}{1 - kd}},\end{aligned}$$

wobei c aus der Anfangsbedingung berechnet werden kann:

$$\begin{aligned}1 &= S(0) = ((1 - kd)c)^{\frac{1}{1 - kd}}, \\ 1 &= (1 - kd)c, \\ \Rightarrow S(z) &= (1 + (1 - kd)z)^{\frac{1}{1 - kd}}, \\ T(z) &= (1 + (1 - kd)z)^{\frac{d}{1 - kd}}.\end{aligned}$$

Für die Koeffizienten bedeutet dies:

$$\begin{aligned}
t_n &= n![z^n]T(z) = n!(1 - kd)^n \binom{\frac{d}{1-kd}}{n} \\
&= n!(1 - kd)^n (-1)^n \binom{n - 1 + \frac{d}{kd-1}}{n} \\
&= n!(kd - 1)^n \binom{n - 1 + \frac{d}{kd-1}}{n} \\
&= (kd - 1)^n \left(n - 1 + \frac{d}{kd - 1} \right)^n.
\end{aligned}$$

Analog geht die Rechnung für die Hilfsgröße \mathcal{S} :

$$\begin{aligned}
s_n &= n![z^n]S(z) = \dots = n!(kd - 1)^n \binom{n - 1 + \frac{1}{kd-1}}{n} \\
&= (kd - 1)^n \left(n - 1 + \frac{1}{kd - 1} \right)^n.
\end{aligned}$$

Wie bei den vorigen ansteigenden Strukturen hätte man auch hier eine einfache Rekursion zum Lösen des kombinatorischen Abzählproblems aufstellen können. Wir schauen uns wieder an, wie sich die Anzahl der freien Plätze beim Erweitern des k -Baums verhält. Durch neu erzeugte Cliques kommen kd Plätze dazu, aber der Knoten, welcher zum k -Baum hinzugefügt wird, nimmt auch einen Platz ein. Bei der Wurzel gibt es nach Definition d Plätze und daher gilt $t_n = (d + (kd - 1)(n - 1))t_{n-1}$ und (vgl. [18], Proposition 1)

$$t_n = \prod_{l=0}^{n-1} (d + (kd - 1)l).$$

2.8 (k, m) -Bäume

In den nächsten zwei Abschnitten werden wir Verallgemeinerungen von markierten k -Bäumen betrachten. Wenn wir uns die Konstruktion der k -Bäume in Erinnerung rufen, sehen wir, dass abgesehen von der Startclique immer $k + 1$ Knoten paarweise miteinander verbunden werden und nie weniger. Insbesondere bilden die k Knoten, mit denen der neue Knoten adjazent ist, schon eine k -Clique, das heißt, sie sind schon über Kanten verbunden. Dies führt zur Überlegung, dass wir Kanten definieren, die nicht nur 2 Knoten verbinden sondern $k + 1$ viele.⁴ Dies führt zur allgemeineren Definition von markierten (k, m) -Bäumen.

Definition 2.3 (vgl. [21]). Es seien $k, m \in \mathbb{N}$ und k die Größe der Kanten. Ein nichtleerer Graph wird als (k, m) -Baum bezeichnet, wenn die Kanten zu E_1, E_2, \dots, E_e angeordnet werden können, sodass für jedes i mit $2 \leq i \leq e$ es ein i' mit $1 \leq i' \leq i$ gibt, welches $|E_i \cap E_{i'}| = m$ und $(E_i \setminus E_{i'}) \cap (\cup_{j=1}^{i-1} E_j) = \emptyset$ erfüllt. Dabei wird m als Überlappungsgröße

⁴Diese „verallgemeinerten Kanten“ stimmen nicht mit dem graphentheoretischen Begriff überein, jedoch wird im folgenden Abschnitt die Bezeichnung „Kante“ auch für diese Verallgemeinerung verwendet.

(engl. „overlap size“) bezeichnet und eine Teilmenge einer Kante der Größe m nennt man Überlappung (engl. „lap“).

Um einen k -Baum zu erhalten, müssen die Kanten eine Überlappungsgröße von k besitzen. Daher kann ein gewöhnlicher k -Baum mit einem markierten $(k + 1, k)$ -Baum identifiziert werden. Ein (k, m) -Baum kann nach Definition rekursiv aufgebaut werden, indem man bei einer Überlappung oder Kante startet und immer weitere Kanten dazugibt, wobei diese eine Überlappung einer bestehenden Kante überdecken.

Wenn man eine Kante zu einem (k, m) -Baum dazugibt, erhöht sich die Anzahl der Überlappungen um $\binom{k}{m} - 1$. So hat ein (k, m) -Baum mit e Kanten insgesamt $e \left(\binom{k}{m} - 1 \right) + 1$ Überlappungen. Die Knotenanzahl verhält sich ähnlich. Mit jeder neuen Kante gibt es $k - m$ neue Knoten und daher hat ein (k, m) -Baum mit e Kanten $e(k - m) + m$ Knoten. Dementsprechend werden wir eine Formel für die Anzahl der (k, m) -Bäume in Abhängigkeit der Kantenanzahl angeben.

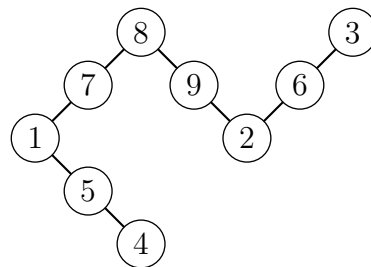


Abbildung 10: Ein $(3, 1)$ -Baum mit 4 Kanten. Es soll beachtet werden, dass die Kanten, nicht so wie es die graphische Darstellung möglicherweise suggeriert, keinen „Anfang“ und kein „Ende“ besitzen beziehungsweise die Knoten nicht auf ihnen geordnet sind.

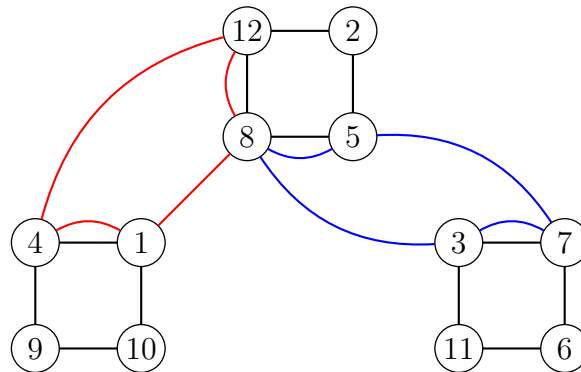


Abbildung 11: Ein $(4, 2)$ -Baum mit 12 Knoten, wobei der Knoten 8 auf drei Kanten liegt.

Für den Beweis wird noch ein technisches Lemma benötigt, welches Beineke und Pippert in [3] verwendet haben, um das Abzählungsproblem für k -Bäume zu lösen. Sie haben dies mit einer Folgerung aus der Abelschen Verallgemeinerung des Binomischen Lehrsatzes⁵ nachgerechnet. Wir haben schon die Anzahl der k -Wurzelbäume bestimmt und so können wir mit Hilfe dieser folgendes Resultat zeigen.

⁵ $(x + y)^n = x \sum_{j=0}^n (x + jz)^{j-1} (x - jz)^{n-j-1}$ (siehe [24], S. 102).

Lemma 2.5 ([3], Lemma 2). *Es gilt*

$$\sum_{(n-k-1)} \binom{n-k-1}{i_1, i_2, \dots, i_{k+1}} \prod_{j=1}^{k+1} (ki_j + 1)^{i_j-1} = (k+1)(kn - k^2 + 1)^{n-k-2},$$

wobei über alle $(k+1)$ -elementigen Kompositionen $i_1 + i_2 + \dots + i_{k+1} = n - k - 1$ summiert wird.

Beweis: Im Beweis von Satz 2.2 haben wir für die Anzahl der k -Wurzelbäume $R_k(n)$ und der k -Bäume die Beziehung

$$\binom{n}{k} R_k(n) = (k(n-k) + 1) T_k(n) \quad (2.11)$$

erhalten. Wir können k -Wurzelbäume mit n Knoten konstruieren, indem wir mit einer $(k+1)$ -Clique C starten und zu jeder k -elementigen Teilmenge einen k -Wurzelbaum dazugeben. Diese $k+1$ k -Wurzelbäume entsprechen einer $(k+1)$ -elementige Partition der restlichen Knoten, wobei ein Element der Partition mit einem k -Wurzelbaum übereinstimmt. Betrachten wir die Möglichkeiten solch einer Konstruktion: Um die $(k+1)$ -Clique auszuwählen, gibt es $\binom{n}{k+1}$ Möglichkeiten. Die Reihenfolge der k -elementigen Teilmengen dieser Clique können wir auf $(k+1)!$ verschiedene Arten auswählen. Die Anzahl der Partitionen stimmt mit dem Multinomialkoeffizient überein. Mit der k -Teilmenge verbundene k -Wurzelbäume, die einer Partitionsklasse der Größe i_j entsprechen, gibt es $R_k(k+i_j)$ viele. Insgesamt ergibt dies

$$\binom{n}{k+1} (k+1)! \sum_{(n-k-1)} \binom{n-k-1}{i_1, i_2, \dots, i_{k+1}} \prod_{j=1}^{k+1} R_k(k+i_j), \quad (2.12)$$

wobei über alle möglichen $(k+1)$ -elementigen Kompositionen von $n-k-1$ summiert wird. Das heißt $i_1 + i_2 + \dots + i_{k+1} = n - k - 1$. Die so konstruierten k -Wurzelbäume sind nicht paarweise verschieden. Wenn die Partitionsklassen permutiert werden und die Auswahl der k -elementigen Teilmenge der Wurzelclique entsprechend angepasst wird, erhalten wir den gleichen k -Wurzelbaum. Die Clique der Größe $k+1$, mit welcher wir bei der Konstruktion gestartet sind, kann nach Lemma 1.1 aus $n-k$ verschiedenen Möglichkeiten ausgewählt werden. So stimmt (2.12) mit $(n-k)(k+1)! T_k(n)$ überein und dies führt mit der Beziehung (2.11) zu:

$$\begin{aligned} \binom{n}{k} R_k(n) &= (k(n-k) + 1) \frac{1}{n-k} \binom{n}{k+1} \sum_{(n-k-1)} \binom{n-k-1}{i_1, i_2, \dots, i_{k+1}} \prod_{j=1}^{k+1} R_k(k+i_j), \\ (k+1) R_k(n) &= (k(n-k) + 1) \sum_{(n-k-1)} \binom{n-k-1}{i_1, i_2, \dots, i_{k+1}} \prod_{j=1}^{k+1} R_k(k+i_j). \end{aligned}$$

Nun können wir unser Wissen über k -Wurzelbäume in Form von Formel (2.3) bemühen und erhalten $R_k(k+i_j) = (ki_j + 1)^{i_j-1}$ beziehungsweise $R_k(n) = (k(n-k) + 1)^{n-k-1}$. Eingesetzt in die obige Gleichung erhalten wir das Gewünschte:

$$(k+1)(k(n-k) + 1)^{n-k-2} = \sum_{(n-k-1)} \binom{n-k-1}{i_1, i_2, \dots, i_{k+1}} \prod_{j=1}^{k+1} (ki_j + 1)^{i_j-1}.$$

□

Der folgende Beweis für das Abzählungsproblem stammt von Pikhurko aus [21], aber die Idee ist, wie auch dort erwähnt wird, die gleiche, die Beineke und Pippert in [3] verwendet haben, um die markierten k -Bäume abzuzählen. Mit $T_{k,m}(e)$ wird die Anzahl der (k, m) -Bäume mit e Kanten bezeichnet. Bei $R_{k,m}(e)$ zählen wir die Menge, bei der wir zusätzlich eine Überlappung als Wurzel ausgezeichnet haben.

Satz 2.6 (vgl. [21], Theorem 1). *Seien $k, m, e \in \mathbb{N}$ mit $0 \leq m \leq k - 1$ und $e \geq 1$. Mit den Bezeichnungen $n = e(k - m) + m, l = \binom{k}{m}$ und $f = e(l - 1) + 1$ gilt für die Anzahl der (k, m) -Bäume mit n Knoten*

$$T_{k,m}(e) = \frac{n! f^{e-2}}{e! m! ((k - m)!)^e} = \frac{(e(k - m) + m)! (e \binom{k}{m} - e + 1)^{e-2}}{e! m! ((k - m)!)^e}. \quad (2.13)$$

Beweis: Wie auch schon bei den k -Bäumen können wir (k, m) -Wurzelbäume abzählen, indem wir eine m -elementige Menge als Wurzel definieren oder aus einem (k, m) -Baum eine Überlappung als Wurzel auswählen. Da es genau f Überlappungen in einem (k, m) -Baum mit e Kanten gibt, gilt

$$\binom{n}{m} R_{k,m}(e) = f T_{k,m}(e). \quad (2.14)$$

Wir können einen (k, m) -Wurzelbaum auch rekursiv aufbauen, indem wir bei einer Kante $E = \{v_1, v_2, \dots, v_k\}$ mit $v_i \in \{1, 2, \dots, n\}$ für $i = 1, \dots, k$ starten. Die Überlappungen L_1, L_2, \dots, L_l von E können nun jeweils als Wurzel für den entsprechenden (k, m) -Teilbaum, welcher über den Schnitt mit E verbunden ist, interpretiert werden. Die $e - 1$ übrigen Kanten müssen auf die l Teilbäume aufgeteilt werden. So erhalten wir eine Komposition der natürlichen Zahl $e - 1 = e_1 + e_2 + \dots + e_l$. Gleiches gilt auch für die Knoten und es ergibt sich eine Partition für die Knoten, welche nicht zur Kante E gehören, in die Klassen X_1, X_2, \dots, X_l mit den Größen $e_1(k - m), e_2(k - m), \dots, e_l(k - m)$. Das heißt, jeder der l Teilbäume hat als Knotenmenge $L_i \cup X_i$ für ein bestimmtes $0 \leq i \leq l$. Natürlich wird hier jeder (k, m) -Baum mehrmals konstruiert, sogar e -mal, da wir jede Kante als Wurzel auswählen können. Anders betrachtet kann die Wurzelkante als k -elementige Teilmenge aus $\{1, 2, \dots, n\}$ ausgewählt werden und die Anzahl der Partitionen der übrigen Menge entspricht genau dem Multinomialkoeffizient. Bei jeder Überlappung L_i gibt es für den $(k - m)$ -Wurzelbaum auch noch $R_{k,m}(e_i)$ Möglichkeiten. Insgesamt erhalten wir

$$\begin{aligned} e T_{k,m}(e) &= \binom{n}{k} \sum_{(e-1)} \binom{n - k}{e_1(k - m), e_2(k - m), \dots, e_l(k - m)} \prod_{i=1}^l R_{k,m}(e_i) \\ &= \frac{n!}{k!} \sum_{(e-1)} \prod_{i=1}^l \frac{R_{k,m}(e_i)}{(e_1(k - m))! \cdots (e_l(k - m))!} \\ &\stackrel{(2.14)}{=} \frac{n!}{k!} \sum_{(e-1)} \prod_{i=1}^l \frac{(e_i(l - 1) + 1) T_{k,m}(e_i)}{(e_1(k - m))! \cdots (e_l(k - m))!} \frac{m!(m + e_i(k - m) - m)!}{(e_i(k - m) + m)!} \\ &= \frac{n!}{k!} \sum_{(e-1)} \prod_{i=1}^l \frac{m!(e_i(l - 1) + 1) T_{k,m}(e_i)}{(e_i(k - m) + m)!}, \end{aligned}$$

wobei über alle Kompositionen der Größe l von $e - 1$ summiert wird. Für $e = 0$ stimmt die Formel (2.13). Wenn wir sie in obige Gleichung einsetzen, ergibt sich

$$\begin{aligned} \frac{n!ef^{e-2}}{e!m!((k-m)!)^e} &= \frac{n!}{k!} \sum_{(e-1)} \prod_{i=1}^l \frac{m!(e_i(l-1)+1)}{(e_i(k-m)+m)!} \frac{(e_i(k-m)+m)!(e_i(l-1)+1)^{e_i-2}}{e_i!m!((k-m)!)^{e_i}}, \\ \frac{f^{e-2}}{(e-1)!m!(k-m)!} &= \frac{1}{k!} \sum_{(e-1)} \prod_{i=1}^l \frac{(e_i(l-1)+1)^{e_i-1}}{e_i!}, \\ l(e(l-1)+1)^{e-2} &= \sum_{(e-1)} \frac{(e-1)!}{e_1!e_2!\cdots e_l!} \prod_{i=1}^l (e_i(l-1)+1)^{e_i-1}. \end{aligned}$$

Dies ist genau die Aussage von Lemma 2.5 unter Berücksichtigung der unterschiedlichen Notation. \square

Bei einem k -Baum wird mit jedem Knoten, der hinzugefügt wird, eine neue $(k+1)$ -Clique erzeugt, was bei einem $(k+1, k)$ -Baum einer Kante entspricht. Ein k -Baum mit n Knoten besitzt $n-k$ Kanten und so stimmt $T_{k+1,k}(n-k)$ mit der hergeleiteten Formel (2.2) überein:

$$\begin{aligned} T_{k+1,k}(n-k) &= \frac{(n-k+k)!((n-k)(k+1) - (n-k) + 1)^{n-k-2}}{(n-k)!k!} \\ &= \binom{n}{k} (k(n-k) + 1)^{n-k-2}. \end{aligned}$$

2.9 Markierte k -gonale 2-Bäume

Im letzten Abschnitt dieses Kapitels wird noch eine Verallgemeinerung von 2-Bäumen vorgestellt. Als Motivation betrachten wir den graphischen Aufbau eines 2-Baumes. Nach der rekursiven Definition wird mit zwei Knoten, die über eine Kante verbunden sind, gestartet. Diese bilden mit dem nächsten Knoten, der hinzugefügt wird, ein Dreieck. Jeder weitere Knoten, welcher noch zum 2-Baum ergänzt wird, bildet ein weiteres Dreieck. Das heißt, wir könnten zur Erstellung eines 2-Baumes auch folgendermaßen vorgehen: Wir starten mit einem Dreieck und fügen rekursiv weitere Dreiecke hinzu, die jeweils nur eine gemeinsame Kante mit dem vorigen 2-Baum haben. So können wir einen 2-Baum als aneinandergeliebte Dreiecke auffassen. Anstatt von Dreiecken könnte man auch andere beliebige Polygone mit k Seiten (k -gone) nehmen. Dies führt uns zu der Definition der k -gonalen 2-Bäume.⁶

Definition 2.4 (vgl. [14], Abschnitt 1). Sei $3 \leq k \in \mathbb{N}$. Ein k -gon ist ein Kreis (als Teilgraph) mit k Knoten respektive Kanten. Ein k -gonaler 2-Baum ist folgendermaßen definiert:

- Ein k -gon ist ein k -gonaler 2-Baum.

⁶Wenn hier von aneinandergeliebten Dreiecken und Polygonen gesprochen wird, setzt dies nicht voraus, dass es eine entsprechende graphische Darstellung in der Ebene gibt, in welcher sich Knoten und Kanten als planare Dreiecke und Polygone ergeben. Labelle, Lamathe und Leroux haben im Zusammenhang mit Molekülverbindungen planare 2-Bäume betrachtet (siehe [13]).

- Ein k -gonaler 2-Baum T , bei dem $k - 2$ Knoten mit $k - 1$ Kanten hinzugefügt werden, sodass diese mit zwei in T verbundenen Knoten ein k -gon bilden, ist auch ein k -gonaler 2-Baum.

Ein markierter k -gonaler 2-Baum ist ein k -gonaler 2-Baum, bei dem die k -gone im kombinatorischen Sinne markiert sind.

Die 3-gonalen 2-Bäume stimmen damit mit den (unmarkierten) 2-Bäumen überein. Beim markierten Fall gibt es aufgrund der unterschiedlichen Beschriftung keine Überschneidung der Definitionen. Wir werden im Folgenden wie bisher nur den markierten Fall betrachten. Dementsprechend sei im Weiteren, wenn wir von einem k -gonalen 2-Baum sprechen, dieser immer markiert.

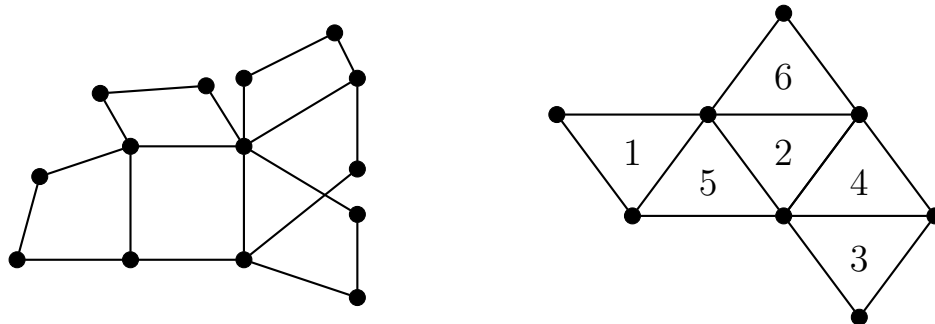


Abbildung 12: Links ist ein unmarkierter 4-gonaler 2-Baum abgebildet. Rechts befindet sich ein 3-gonaler 2-Baum der Größe 6.

Mit der unterschiedlichen Markierung „ändert“ sich natürlich auch die Größe, nach welcher die Objekte abgezählt werden. Bisher haben wir immer die Knoten gezählt, bei den k -gonalen 2-Bäumen werden es die k -gone sein. Um das Abzählungsproblem zu lösen, betrachten wir auch noch eine zusätzliche Struktur auf den Bäumen. Wir werden die Kanten orientieren, sodass jedes k -gon einen orientierten Zyklus formt. Dies bedeutet, dass durch die Orientierung einer Kante e schon die Orientierungen aller Kanten, die zu einem k -gon gehören, welches e enthält, festgelegt sind. Dies lässt sich transitiv fortsetzen und damit bestimmt eine einzige Kante, wie der Baum orientiert ist⁷. Wie wir es bisher auch schon öfters getan haben, werden wir auch hier zuerst eine kombinatorische Struktur betrachten, bei der eine „Wurzel“ ausgezeichnet ist. Die Wurzel wird in unserem Fall eine Kante sein (vgl. [14], Abschnitt 1). Es ergibt sich folgende Liste von Strukturen:

- \mathcal{A} , die markierten k -gonalen 2-Bäume,
- \mathcal{A}_o , die orientierten (und markierten) k -gonalen 2-Bäume,
- \mathcal{A}_o^- , die orientierten (markierten) k -gonalen 2-Wurzelbäume, das heißt bei denen eine Kante als Wurzel ausgezeichnet ist.

⁷Obwohl graphentheoretisch ein orientierter k -gonaler 2-Baum einem orientierten Graphen entspricht, wird die Verbindung von zwei Knoten weiterhin als Kante und nicht als Bogen bezeichnet. Die Orientierung wird als zusätzliche Eigenschaft von Kanten betrachtet und schränkt uns nicht in der „Bewegung“ ein. So werden auch Pfade entgegen der Orientierung betrachtet.

Wir werden nun die Abzählprobleme für die Liste „von unten nach oben“ lösen. Dabei werden wir zuerst mit Hilfe einer rekursiven Beschreibung der orientierten k -gonalen Wurzelbäume eine Beziehung zu den markierten Wurzelbäumen herstellen. Mit dieser ist es möglich, die Anzahl der orientierten Bäume zu bestimmen. Dann müssen wir nur noch die Orientierung der Kanten loswerden, dabei wird interessanterweise die Parität von k eine Rolle spielen.

Lemma 2.7 (vgl. [14], Theorem 2). *Es sei $\mathcal{B} = \mathcal{A}_o^-$, dann gelten die Gleichungen*

$$\begin{aligned} \mathcal{B} &= \text{SET}(\mathcal{Z}\mathcal{B}^{k-1}), \\ B(z) &= \exp(zB^{k-1}(z)), \end{aligned} \tag{2.15}$$

wobei \mathcal{Z} die kombinatorische Struktur eines einzelnen k -gons ist.

Beweis: Wir teilen den orientierten Baum an der Wurzel auf, indem wir jeweils maximale Teilgraphen T_i auswählen, die nur ein k -gon K_i , mit der Wurzel gemeinsam haben. Aufgrund der rekursiven, baumartigen Aufbauweise von k -gonalen 2-Bäumen sind diese Teilgraphen disjunkt. Bei den K_i 's bildet jede Kante verschieden zur Wurzel mit den übrigen verbundenen k -gonen ein Objekt von \mathcal{B} , indem die Kante als Wurzel definiert wird. Für jede Kante entstehen unterschiedliche orientierte Wurzelbäume. Somit lässt sich ein orientierter k -gonaler Wurzelbaum als Menge von kombinatorischen Produkten aus einem k -gon und einem $(k-1)$ -Tupel von orientierten k -gonalen Wurzelbäumen auffassen.⁸ \square

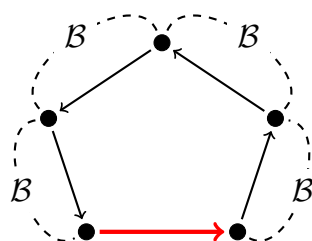


Abbildung 13: Ein K_i aus dem Beweis von Lemma 2.7, dabei ist die rote Kante die Wurzel (vgl. [14], Abb. 2).

Bisher haben wir die markierten Wurzelbäume noch nicht betrachtet. Da wir bei jedem Baum einen beliebigen Knoten als Wurzel auswählen können, gibt es nach Caley's Formel (Korollar 2.3) n^{n-1} markierte Wurzelbäume. Die kombinatorische Struktur \mathcal{C} dieser Bäume erfüllt die Gleichung $\mathcal{C} = \mathcal{Z} \star \text{SET}(\mathcal{C})$, denn wir können jedes Kind der Wurzel als Wurzel des Teilbaumes, welcher über das Kind mit der ursprünglichen Wurzel verbunden ist, ansehen. So besteht ein Wurzelbaum aus der Wurzel, einem einfach markierten Objekt, und einer Menge von markierten Wurzelbäumen. Aus dieser Gleichung lassen sich

⁸In [14] wird jeweils nur die Wurzelkante orientiert und nicht der ganze Baum. Dementsprechend können die Kanten in den K_i 's beliebig orientiert werden, solange dies nach einer Vorschrift konsistent durchgeführt wird. Dort werden die Kanten „von der Wurzel weg“ orientiert. Dies hat den Vorteil, dass die Wurzelkante nicht orientiert sein muss, aber für k gerade bleibt die Orientierung der gegenüberliegenden Kante unklar (vgl. [14], Abbildung 3).

mit der Lagrange'schen Inversionsformel auch die Koeffizienten berechnen. Die Gleichung, welche \mathcal{C} erfüllt, ist ähnlich aufgebaut wie die der orientierten k -gonalen Wurzelbäume und es gibt tatsächlich eine Beziehung zwischen diesen beiden Größen.

Proposition 2.8 (vgl. [14], Proposition 1). *Für die exponentiell erzeugende Funktion der orientierten markierten Wurzelbäume \mathcal{B} und markierten Wurzelbäume \mathcal{C} gilt:*

$$B(z) = {}^{k-1}\sqrt{\frac{C((k-1)z)}{(k-1)z}}.$$

Beweis: Durch Manipulation der Formel (2.15) erhalten wir mit $\varphi(f(z)) := (k-1)ze^{f(z)}$:

$$\begin{aligned} (k-1)zB^{k-1}(z) &= (k-1)z \left(\exp(zB^{k-1}(z)) \right)^{k-1} = (k-1)ze^{(k-1)zB^{k-1}(z)}, \\ \Leftrightarrow \frac{(k-1)zB^{k-1}(z)}{(k-1)ze^{(k-1)zB^{k-1}(z)}} &= \frac{(k-1)zB^{k-1}(z)}{\varphi((k-1)zB^{k-1}(z))} = z. \end{aligned}$$

Für die markierten Wurzelbäume ergibt sich mit dem Argument $(k-1)z$ dieselbe Formel:

$$\begin{aligned} C((k-1)z) &= (k-1)ze^{C((k-1)z)}, \\ \Leftrightarrow \frac{C((k-1)z)}{(k-1)e^{C((k-1)z)}} &= \frac{C((k-1)z)}{\varphi(C((k-1)z))} = z. \end{aligned}$$

Wenn wir nun die Lagrange'sche Inversionsformel zu Hilfe nehmen, sehen wir, dass die Potenzreihen $(k-1)zB^{k-1}(z)$ und $C((k-1)z)$ dieselben Koeffizienten besitzen und somit übereinstimmen. Es ergibt sich

$$B^{k-1}(z) = \frac{C((k-1)z)}{(k-1)z},$$

wobei die rechte Seite eine Potenzreihe mit konstantem Term 1 ist ($c_0 = 0, c_1 = 1$). Dies trifft auch für die Reihe $B(z)$ zu. Mit dieser Bedingung (${}^{k-1}\sqrt{1} = 1$) ist eine $(k-1)$ -te Wurzelfunktion in einer Umgebung um 1 eindeutig definiert. Dementsprechend können wir die Wurzel ziehen und dies beendet den Beweis. \square

Mit dieser Vorarbeit können wir nun die Anzahl orientierter k -gonaler Wurzelbäume bestimmen.

Satz 2.9 (vgl. [14], Corollary 1). *Wenn wir mit $m = (k-1)n+1$ die Anzahl der Kanten bezeichnen, gilt für die Koeffizienten der orientierten k -gonalen Wurzelbäume:*

$$a_{o,n}^- = ((k-1)n+1)^{n-1} = m^{n-1}.$$

Beweis: Zuerst zeigen wir ein Hilfsresultat, welches auf die Arbeit [10] von Fowler, Gessel, Labelle und Leroux zurückgeht:

$$\left(\frac{C(z)}{z} \right)^\mu = \sum_{n \geq 0} \mu(\mu+n)^{n-1} \frac{z^n}{n!}, \quad \mu > 0. \quad (2.16)$$

Dies können wir mit Hilfe der Lagrange'schen Inversionsformel nachrechnen. Zur Erinnerung, es gilt $C(z) = ze^{C(z)}$.

$$\begin{aligned} [z^n] \left(\frac{C(z)}{z} \right)^\mu &= [z^n] \exp(\mu C(z)) \stackrel{\text{(LIF)}}{=} \frac{1}{n} [u^{n-1}] e^{u\mu} \mu e^{\mu u} \\ &= \frac{\mu (n + \mu)^{n-1}}{n (n-1)!}, \end{aligned}$$

womit (2.16) gezeigt ist. Nun setzen wir $\mu = (k-1)^{-1}$ und erhalten auf der linken Seite, wenn wir $(k-1)z$ als Argument nehmen, die gesuchten Koeffizienten für $n \geq 1$:

$$\begin{aligned} a_{o,n}^- &= n! [z^n] \sqrt[k-1]{\frac{C((k-1)z)}{(k-1)z}} = (k-1)^{-1} ((k-1)^{-1} + n)^{n-1} (k-1)^n \\ &= (1 + (k-1)n)^{n-1} = m^{n-1}. \end{aligned}$$

□

Korollar 2.10 (vgl. [14], Corollary 2). *Es gilt*

$$a_{o,n} = ((k-1)n + 1)^{n-2} = m^{n-2}, \quad \text{für } n \geq 2.$$

Beweis: Bei einem orientierten k -gonalen Baum können wir jede Kante als Wurzel auswählen, daher $a_{o,n}^- = ma_{o,n}$. □

Um die gewöhnlichen k -gonalen 2-Bäume zu erhalten, müssen wir die Orientierten nur noch nach der Orientierung faktorisieren. Dies werden wir mit dem Lemma von Burnside 1.4 erledigen. Wie wir sehen werden, unterscheidet sich die Anzahl der Fixpunkte in Abhängigkeit von der Parität k .

2.9.1 Fall: k ungerade

Die Permutationsgruppe \mathfrak{S} auf den orientierten k -gonalen Bäumen besteht aus zwei Elementen, aus der Identität id und der Permutation τ , bei der jede Orientierung umgedreht wird. Wenn wir die Orientierungen nur bei einer Teilmenge T der Kanten mit $\emptyset \neq T \neq E$ umdrehen, sind nicht mehr alle k -gone wie ein Zyklus orientiert, das heißt, es liegt nicht mehr ein orientierter k -gonaler 2-Baum vor (vgl. [14], Abschnitt 4). Somit wissen wir nach dem Lemma von Burnside für $n \geq 2$:

$$a_n = \frac{1}{|\mathfrak{S}|} (|\text{Fix}_{\mathcal{A}_{o,n}}(\text{id})| + |\text{Fix}_{\mathcal{A}_{o,n}}(\tau)|) = \frac{1}{2} (m^{n-2} + |\text{Fix}_{\mathcal{A}_{o,n}}(\tau)|). \quad (2.17)$$

Um genauer zu untersuchen, ob ein Element ein Fixpunkt von τ ist, definieren wir eine Funktion d , welche den Abstand von zwei k -gonen untersucht, die einen gemeinsamen Nachbarn haben. Um diese zu definieren, sind noch zwei Hilfsgrößen notwendig. Mit $F \subseteq V^k$ bezeichnen wir die Menge der k -gone eines orientierten k -gonalen 2-Baumes und $N \subseteq F^2$ sei die Teilmenge, welche aus Paaren (K_1, K_2) von k -gonen besteht, die über ein k -gon verbunden sind. Das heißt, es existiert ein $K \in F$ mit $|K \cap K_1| = |K \cap K_2| = 2$. Für zwei k -gone $(K_1, K_2) \in N$ gibt es bei K_1 genau eine orientierte Kante, die zu einem

Knoten von K zeigt und unterschiedlich ist zur Kante, welche beide k -gone gemeinsam besitzen. Die Funktion $h : N \rightarrow V$ soll auf diesen Knoten abbilden. So können wir die Funktion $d : N \rightarrow \mathbb{N}$ folgendermaßen definieren, wobei $|K \cap K_1| = |K \cap K_2| = 2$ gilt:

$$d(K_1, K_2) := \min(\{|P| : P \text{ ist ein Pfad von } h(K_1, K) \text{ nach } K_2\}).$$

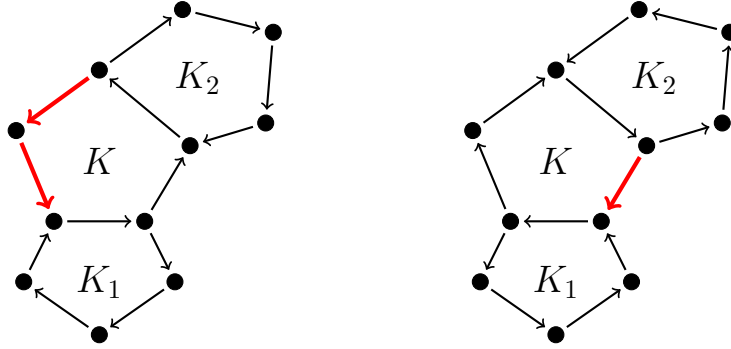


Abbildung 14: Es wird hier die Funktion d graphisch an einem Beispiel und deren Bild unter der Permutation τ dargestellt. Der rote Pfad ist jeweils der Kürzeste, wobei es links zwei gleich lange Pfade gibt.

Das Minimum wird im Prinzip nur von zwei sinnvollen Pfaden gebildet, denn man kann links- bzw. rechtsherum auf dem k -gon K entlanggehen, bis man auf einen Knoten von K_2 trifft. Nun betrachten wir, wie sich $d(K_1, K_2)$ unter τ verhält. Dabei weist, wenn es entscheidend ist, der Index der Funktion d_\times hin, in welchem orientierten k -gonalen 2-Baum der Abstand gemessen wird. Obwohl wir im weiteren Verlauf auch vom Abstand im Zusammenhang mit dem Wert von d sprechen, ist dies keine Metrik. Wie man sich leicht überlegen kann, ist d sogar in den meisten Fällen nicht symmetrisch.

Lemma 2.11. Sei $T \in \mathcal{A}_o$ und $(K_1, K_2) \in N_T$. Dann gelten die Aussagen:

(i) $d(K_1, K_2) \leq \lfloor \frac{k-1}{2} \rfloor$.

(ii) $d_T(K_1, K_2) = 0 \Rightarrow d_{\tau(T)}(K_1, K_2) = \begin{cases} 0, & \text{falls } |K_1 \cap K_2| = 2, \\ 1, & \text{falls } |K_1 \cap K_2| = 1. \end{cases}$

(iii) $0 \neq d(K_1, K_2) < \lfloor \frac{k}{2} \rfloor - 1 \Rightarrow d_{\tau(T)}(K_1, K_2) \in \{d(K_1, K_2) - 1, d(K_1, K_2) + 1\}$.

(iv) Falls k ungerade ist: $d(K_1, K_2) = \frac{k-1}{2} \Rightarrow d_{\tau(T)}(K_1, K_2) = \frac{k-1}{2} - 1$.

(v) Falls k gerade ist: $d(K_1, K_2) = \frac{k-2}{2} \Rightarrow d_{\tau(T)}(K_1, K_2) = \frac{k-2}{2}$.

Beweis: Bezeichne $K \in F$ das k -gon mit $|K \cap K_1| = |K \cap K_2| = 2$. Wenn wir links und rechtsrum jeweils $\lfloor \frac{k-1}{2} \rfloor$ Schritte von $h(K_1, K)$ gehen, erreichen wir inklusive dem Startknoten $2 \lfloor \frac{k-1}{2} \rfloor + 1 \geq k - 1$ Knoten und damit sicher auch einen von K_2 .

Damit $d_T(K_1, K_2) = 0$ gilt, müssen K_1 und K_2 einen Knoten gemeinsam haben. Falls sie eine Kante gemeinsam haben, bleibt der Abstand unter der Permutation τ gleich, ansonsten erhöht er sich auf 1.

Für (iii) müssen wir uns überlegen, dass der Knoten, welcher weiter weg ist, auch nach

der Permutation weiter weg ist. Wenn wir auf K in die gleiche Richtung wie der minimale Pfad in T von $h(K_1, K)$ nach K_2 gehen, stimmt das. In der anderen Richtung hat er den Abstand $k - 1 - d_T(K_1, K_2)$, welcher sich durch die Permutation um 1 erhöhen oder verringern kann. Das bedeutet, dass der Abstand in die andere Richtung größer ist als

$$k - d_T(K_1, K_2) - 2 > k - \left\lfloor \frac{k}{2} \right\rfloor - 1 = \left\lceil \frac{k}{2} \right\rceil - 1 = \left\lfloor \frac{k-1}{2} \right\rfloor.$$

Daher ist der Abstand zum anderen Knoten immer noch kürzer und es gilt (iii). Falls k ungerade ist und der Knoten $h(K_1, K)$ genau gegenüber der Kante, welche K und K_2 gemeinsam haben, liegt, verringert sich der Abstand durch einen Orientierungswechsel um 1 (für den Fall $k = 5$ ist dies in der Abbildung 14 ersichtlich). Dies entspricht (iv). Wenn k gerade ist, besitzt ein k -gon gegenüberliegende Seiten. Falls K_1, K_2 an gegenüberliegenden Kanten in K angrenzen, ändert sich deren Abstand nicht, was mit der Aussage (v) übereinstimmt. \square

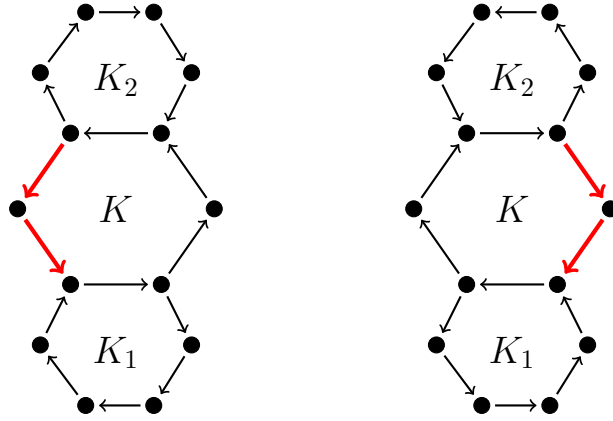


Abbildung 15: Für den Fall $k = 6$ ist hier die Aussage von Lemma 2.11(v) dargestellt.

Mit diesem Hilfsmittel kann die Anzahl der Fixpunkte von τ bestimmt werden. Für einen Fixpunkt im Falle eines ungeraden k 's folgt aus Lemma 2.11, dass alle k -gone eine gemeinsame Kante besitzen. Da wir nur die k -gone markieren, ist durch diese Bedingung ein k -gonaler 2-Baum eindeutig festgelegt. Mit diesem einzigen Fixpunkt kann mit Hilfe der Formel (2.17) die Anzahl der k -gonalen 2-Bäume angegeben werden.

Satz 2.12 (vgl. [14], Proposition 4). *Falls k ungerade ist, gilt für die markierten k -gonalen 2-Bäume \mathcal{A} :*

$$a_n = \frac{1}{2} (m^{n-2} + 1), \quad n \geq 2,$$

wobei $m = (k - 1)n + 1$ die Anzahl der Kanten ist.

2.9.2 Fall: k gerade

Betrachten wir im Fall, dass k gerade ist, ein Element aus der Fixpunktmenge $\text{Fix}_{\mathcal{A}_{o,n}}(\tau)$. Aufgrund von Lemma 2.11 sind die Nachbarn eines k -gons auf nur zwei gegenüberliegenden

Kanten aufgeteilt. An den anderen Kanten sind keine k -gone angeklebt. Wenn nun die Kanten, die zu mindestens zwei verschiedenen k -gonen gehören, als Knoten und die k -gone als Kanten interpretiert werden, erhält man einen gewöhnlichen Baum mit markierten Kanten. Damit am Ende eines Astes nicht eine Kante ins nirgendwo führt, muss die gegenüberliegende Kante als Knoten dazugegeben werden. Die Fixpunkte können also bijektiv auf die Bäume, welche markierte Kanten haben, abgebildet werden (vgl. [14], Abschnitt 4.2).

Wenn wir bei einem Baum mit n markierten Kanten einen Knoten als Wurzel auswählen und die Markierung der Kanten den Knoten in Richtung „von der Wurzel weg“ zuordnen, erhalten wir einen Wurzelbaum von der Größe $n + 1$, bei dem abgesehen von der Wurzel alle Knoten markiert sind. Die Wurzel markieren wir nun mit einer Zahl aus der Menge $\{1, 2, \dots, n + 1\}$, wobei die restlichen Markierungen konsistent angepasst werden (Zum Beispiel erhöhe um 1 alle Markierungen, die größer gleich als die ist, welche für die Wurzel ausgewählt wird.). Mit Caley's Formel (2.3) beziehungsweise mit unserer Formel (2.5) für k -Wurzelbäume⁹ können wir die Anzahl der Wurzelbäume mit $n + 1$ Knoten bestimmen. Dies ist $(n + 1)^n$. Für die Abbildung zwischen den Bäumen mit markierten Kanten und markierten Wurzelbäumen wurde zweimal beliebig aus einer $(n + 1)$ -elementigen Menge ausgewählt. Mit der entsprechenden Division ergibt dies für die Fixpunkte $\text{Fix}_{\mathcal{A},n}(\tau) = (n + 1)^{n-2}$, womit die a_n von \mathcal{A} für gerades k bestimmt werden können.

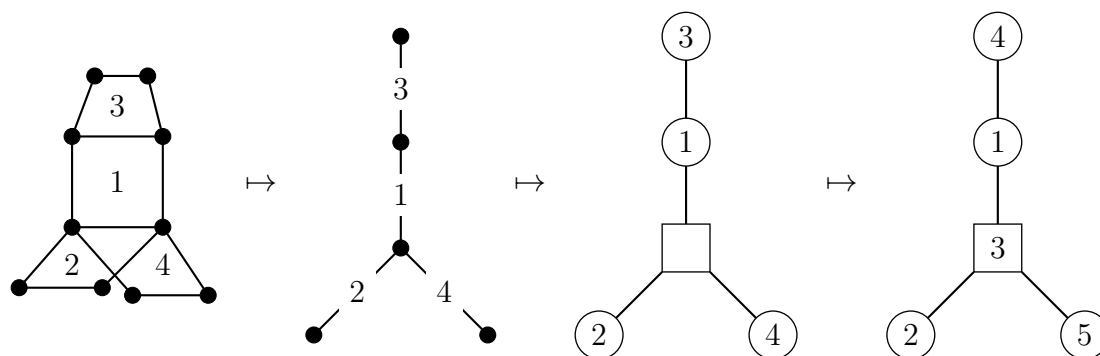


Abbildung 16: Exemplarische Darstellung der Identifikation von den k -gonalen 2-Bäumen mit den markierten Wurzelbäumen.

Satz 2.13 (vgl. [14], Proposition 9). *Falls k gerade ist, gilt für die markierten k -gonalen 2-Bäume \mathcal{A} :*

$$a_n = \frac{1}{2} (m^{n-2} + (n + 1)^{n-2}), \quad n \geq 2,$$

wobei $m = (k - 1)n + 1$ die Anzahl der Kanten ist.

⁹Man beachte, dass bei den k -Wurzelbäumen nur die externen Knoten gezählt wurden, aber im vorliegenden Fall alle gezählt werden.

3 Weitere Betrachtung von k -Bäumen

Im weiteren Teil werden spezifische Parameter von k -Bäumen untersucht. Zuerst wird der Knotengrad analysiert, aus welchem der lokale Clusterkoeffizient berechnet werden kann. Dieser gibt Auskunft darüber, wie stark die Vernetzung um einen Knoten herum ist. Der Knotengrad ist abhängig von der Anzahl der Kinder einer Clique, respektive eines Knotens. Es wird auch die transitive Erweiterung der „Kinder-Relation“, welche uns zu der Definition der Nachfahren führt, betrachtet. Abgeschlossen wird die Betrachtung der k -Bäume durch die Analyse der Vorfahren eines Knotens.

Der Berechnung dieser Größen liegt immer der rekursive Aufbau von k -Bäumen zu Grunde, insbesondere werden wir von dem n -ten hinzugefügten Knoten sprechen. Deshalb schränken wir uns im Folgenden auf die ansteigenden Modelle von k -Bäumen mit einer Wurzel ein. So stimmt die kombinatorische Markierung mit dem Zeitpunkt des Hinzufügens beim rekursiven Aufbau überein. Wir werden auch die Knoten an Hand ihrer Markierung benennen. Das heißt, wir werden von einem Knoten $n \in \mathbb{N}$ sprechen. Die Rekursionen für den Knotengrad und die Nachfahren sind „ähnlich“ aufgebaut und können aus der Beobachtung was passiert, wenn ein Knoten hinzugefügt wird, hergeleitet werden. Sie lassen sich mit der gleichen Methodik lösen. Hingegen geschieht bei der Berechnung der Vorfahren der Schritt in die andere Richtung. Man untersucht wie sich die Größe beim Entfernen eines Knotens ändert.

3.1 Randomisierte Modelle

Um beispielsweise den durchschnittlichen Knotengrad zu studieren, ist eine Wahrscheinlichkeitsverteilung der k -Bäume notwendig. Man könnte diese einfach unter der Gleichverteilung betrachten. Bei realen Netzwerken werden diese meistens von klein auf aufgebaut. Bei den k -Bäumen ist dies über die rekursive Struktur auch möglich. Dies werden wir im Weiteren verwenden und mit möglicherweise unterschiedlichen Wahrscheinlichkeiten den Cliques einen weiteren Knoten zuordnen. Je nach Präferenzen beim Hinzufügen von Knoten können die bisher betrachteten ansteigenden Modelle konstruiert werden, sodass diese gleichverteilt sind.

Satz 3.1 (vgl. [18], Proposition 2). *Wenn man das Hinzufügen von Knoten einer Wahrscheinlichkeitsverteilung unterwirft, ist es möglich, die ansteigenden, die geordneten ansteigenden und die ansteigenden d -ären k -Bäume zu erzeugen, sodass diese gleich verteilt sind. Man startet mit einer Wurzelclique $0_1, 0_2, \dots, 0_k$ und gibt Knoten mit einer gewissen Wahrscheinlichkeit bei Cliques hinzu. Im Speziellen sind die Wahrscheinlichkeiten für das Hinzufügen des n -ten Knotens bei den verschiedenen Modellen:*

(i) *ansteigende k -Bäume:* $p(C) = \frac{1}{1+k(n-1)},$

(ii) *geordnete ansteigende k -Bäume:* $p(C) = \frac{d^+(C)+1}{1+(k+1)(n-1)},$

(iii) *ansteigende d -äre k -Bäume:* $p(C) = \frac{d-d^+(C)}{d+(kd-1)(n-1)},$

wobei mit Wahrscheinlichkeit $p(C)$ ein Knoten mit der Clique C verbunden wird und $d^+(C)$ die Anzahl der Kinder von C bezeichnet.

Beweis: Das Abzählungsproblem der drei Modelle konnten wir in Abschnitt 2 mit einer Rekursion lösen. Dabei haben wir betrachtet, wie viele Plätze es gibt, um weitere Knoten hinzuzufügen. Um Satz 3.1 zu zeigen, müssen wir uns nur noch überlegen, dass die Anzahl dieser Plätze mit dem Zähler im jeweiligen Bruch übereinstimmt. Der Nenner stimmt für jedes Modell, wie wir uns in Abs. 2 überlegt haben, mit der Zahl für mögliche Plätze für den n -ten Knoten überein. Wenn der Zähler jeweils mit der Anzahl an Plätzen pro Clique übereinstimmt, heißt dies, dass der n -te Knoten jeden Platz mit gleicher Wahrscheinlichkeit wählt. Der Platz des n -ten Knotens wird also mit einer uniformen Verteilung unter allen möglichen Plätzen ausgewählt.

Ad (i): Bei einem ansteigenden k -Baum kommt es nicht auf die Reihenfolge der Kinder an und so hat jede k -Clique einen Platz bei dem die Knoten eingefügt werden können.

Ad (ii): Wenn eine Clique bei einem geordneten ansteigenden k -Baum kein Kind hat, gibt es noch keine Reihenfolge der Kinder und die Clique hat damit einen möglichen Platz für ein Kind. Wenn die Clique m Kinder hat, kann ein weiterer Knoten vor dem ersten Kind, zwischen dem ersten und zweiten Kind und so weiter eingefügt werden. Es gibt somit $m + 1$ Plätze, um ein weiteres Kind bei der Clique hinzuzufügen.

Ad (iii): Die Anzahl der Kinder einer Clique ist bei den d -ären k -Bäumen mit d beschränkt und jedes Kind nimmt genau einen möglichen Platz ein. \square

Bei den Parametern, die wir berechnen werden, hat die Reihenfolge der Kinder keine Bedeutung. Daher könnte man beispielsweise die berechneten Werte für die geordneten ansteigenden k -Bäume mit den Werten für ansteigende k -Bäume, welche beim Hinzufügen eines Knotens das Wahrscheinlichkeitsmodell (ii) aus Satz 3.1 verwenden, identifizieren. Die ansteigende Markierung benötigen wir, um die Knoten nach ihrer Reihenfolge, in der sie auftreten, zu unterscheiden. So sind gemittelte Werte, wie zum Beispiel der Knotengrad eines beliebig hinzugefügten Knotens, auch für (unmarkierte) k -Bäume, welche mit der entsprechenden Verteilung rekursiv aufgebaut werden, gültig.

Ein weiteres Modell zum Erzeugen von k -Bäumen, welches wir im Folgenden betrachten, stammt aus einer Arbeit von Cooper und Uehara ([6]). Deshalb werden wir es im weiteren Teil der Arbeit als CU-Modell in Anlehnung an die Autoren bezeichnen. Um den Algorithmus von Cooper und Uehara auszuführen, wird als Startpunkt eine $(k + 1)$ -Clique gewählt. Um die gleiche kombinatorische Größenfunktion zu haben, wie bei den übrigen Modellen, nehmen wir an, dass diese $(k + 1)$ -Clique aus den Knoten $0_1, 0_2, \dots, 0_k, 1$ besteht und die Größe 1 aufweist. Das Hinzufügen eines weiteren Knotens erfolgt mit den folgenden zwei Schritten:

- Unter der Gleichverteilung der $(k + 1)$ -Cliquen wird eine $(k + 1)$ -Clique ausgewählt.
- Diese $(k + 1)$ -Clique enthält $k + 1$ k -Cliquen und aus diesen wird wieder unter uniformer Verteilung eine k -Clique ausgewählt, bei welcher ein Knoten hinzugefügt wird.

Für dieses randomisierte Erzeugen von k -Bäumen haben wir kein entsprechendes Modell, welches gleichverteilt erzeugt wird. Wir werden diese k -Bäume an Hand der getroffenen Auswahlen in den obigen zwei Schritten unterscheiden. Genaueres dazu folgt in Abschnitt 3.2.4.

Im weiteren Teil der Arbeit werden wir diese mögliche randomisierte Erzeugung der verschiedenen k -Baummodelle nicht mehr explizit erwähnen, sondern unterscheiden die

berechneten Parameter an dem Modell aus Kapitel 2, welches uniform erzeugt wird. Wie wir gezeigt haben, stimmen diese überein. Für die Praxis relevanter ist natürlich die Beschreibung durch die probabilistische Erzeugung von k -Bäumen als eine angenommene Gleichverteilung der verschiedenen Modelle.

Um die vier verschiedenen randomisierten Modelle zur Erzeugung von k -Bäumen ein wenig genauer unterscheiden zu können, wollen wir uns überlegen, welche Cliques bevorzugt werden. Das heißt, bei welchen Cliques wird mit höherer Wahrscheinlichkeit ein weiterer Knoten hinzugefügt werden. Bei den ansteigenden k -Bäumen sind alle Cliques gleichberechtigt, das heißt, bei jeder Clique wird mit gleicher Wahrscheinlichkeit ein Knoten hinzugefügt. Bei den geordneten ansteigenden k -Bäumen werden Cliques mit vielen Kindern präferiert, wohingegen bei den d -ären Bäumen genau das Gegenteil passiert. Dort werden Cliques mit wenigen Kindern mit größerer Wahrscheinlichkeit ausgewählt. Dementsprechend liegt die Vermutung nahe, dass zum Beispiel der Knotengrad bei den geordneten ansteigenden k -Bäumen mehr Ausreißer mit hohem Knotengrad hat als die d -ären Bäume. Beim CU-Modell werden auch Cliques mit vielen Kindern beim Hinzufügen priorisiert. Denn für jedes Kind ist eine Clique in einer weiteren $(k + 1)$ -Clique und wird damit wahrscheinlicher ausgewählt. Dementsprechend erwarten wir ähnliche Werte für diese zwei Modelle.¹⁰ Abbildung 17 zeigt, dass die geordneten ansteigenden k -Bäume und das CU-Modell unterscheiden und nicht übereinstimmen.

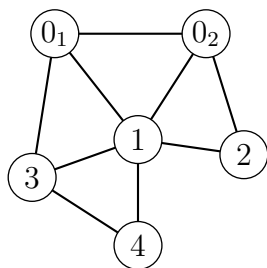


Abbildung 17: Beim Modell der geordneten ansteigenden k -Bäume wird bei der Clique $\{0_1, 1\}$ mit einer Wahrscheinlichkeit von $\frac{2}{1+3+4} = \frac{2}{13}$ und bei der Clique $\{0_1, 0_2\}$ mit $\frac{2}{13}$ ein Knoten hinzugefügt. Beim CU-Modell lauten die Wahrscheinlichkeiten $\frac{2}{4} \cdot \frac{1}{3} = \frac{1}{6}$ und $\frac{1}{12}$ für die zwei Cliques.

3.2 Knotengrad

In einem k -Baum hat jeder Knoten verschieden zu einem Wurzelknoten mindestens k Nachbarn. Für den Knoten x sind dies die k Knoten, die eine k -Clique C bilden, bei der x hinzugefügt wurde. Deshalb werden wir im Folgenden von dem ausgehenden Knotengrad beziehungsweise der Anzahl der Kinder von einem Knoten sprechen. Dabei werden nur Knoten, welche nicht zu C gehören, gezählt. Dies sind natürlich auch die Knoten die eine größere Markierung als x aufweisen. Für einen Wurzelknoten sieht es anders aus,

¹⁰Mit Hinblick auf die Wahrscheinlichkeiten von Abbildung 17 könnte man die Überlegungen von Abs. 3.2.4 auch für Cliques durchführen, um zu sehen, dass für das CU-Modell als Verteilung $p(C) = \frac{d^+(C)+1}{(n-1)(k+1)}$ für Cliques ungleich der Wurzel herauskommt. Dementsprechend unterscheiden sich die beiden Verteilungen für den gleichen k -Baum abgesehen von der Wurzel nur im „ungekürzten Nenner“ um 1.

dieser ist mindestens mit $k - 1$ Knoten verbunden. Dies hat auch Auswirkungen auf die Rekursion, sodass wir im Folgenden die verschiedenen Fälle für Wurzelknoten und hinzugefügte Knoten betrachten. Im Weiteren wird zuerst für Hinzugefügte und danach für Wurzelknoten der Knotengrad berechnet. Am Ende jedes Modells wird für einen beliebig hinzugefügten Knoten der Grad berechnet und auf ein Potenzgesetz geschlossen.

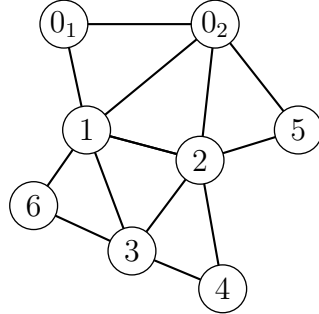


Abbildung 18: Der Knoten 2 hat in diesem ansteigenden k -Baum der Größe 6 einen Grad von 5. Die Kinder von 2 sind die Knoten 3, 4, 5 und somit hat der Knoten 2 einen ausgehenden Knotengrad von 3.

3.2.1 Geordnete, ansteigende k -Bäume

Hinzugefügter Knoten: Nun betrachten wir den ausgehenden Knotengrad für geordnete ansteigende k -Bäume. Zuerst werden wir den Knotengrad für Knoten außerhalb der Wurzel berechnen. Dieser unterscheidet sich zu dem Grad der Wurzelknoten, da diese abhängig vom ausgehenden Knotengrad unterschiedlich vielen k -Cliques angehören. Früher wurde schon festgestellt, dass für das Hinzufügen des n -ten Knotens $1 + (k + 1)(n - 1)$ Möglichkeiten vorhanden sind. Es sei $t_{n,j,m}$ die Anzahl der geordneten ansteigenden k -Bäume mit n Knoten, bei dem der Knoten j genau m Kinder besitzt. Um eine Rekursion für $t_{n,j,m}$ aufzustellen, müssen wir wissen, wie viele Möglichkeiten es gibt, einen Knoten bei einer Clique, die den Knoten j umfasst, anzuhängen. Der n -te Knoten kann entweder bei Knoten j angehängt werden oder nicht. Da alle Möglichkeiten mit gleicher Wahrscheinlichkeit ausgewählt werden, führt dies zu einer Beschreibung der Rekursion mit

$$t_{n,j,m} = |\{\text{Knoten } n \text{ ist kein Kind von } j\}| \cdot t_{n-1,j,m} + |\{n \text{ ist Kind von } j\}| \cdot t_{n-1,j,m-1}.$$

Wenn ein Knoten x mit j verbunden wird, entstehen auf zwei Arten Möglichkeiten für neue Plätze. Zum Einen ergibt sich ein neuer Platz in der k -Clique, bei welcher x hinzugefügt wird. Zum Anderen entstehen $k - 1$ neue k -Cliques, welche die Knoten x und j umfassen. In diesen gibt es genau einen möglichen Platz. Das heißt, mit jedem Kind erhöht sich die Anzahl der Plätze um k . Wenn der Knoten j noch kein Kind hat, gibt es, da er ein Teil von k Cliques ist, k Möglichkeiten einen Knoten bei j anzuhängen. So gibt es bei einem Knoten mit m Kindern $(m + 1)k$ mögliche Plätze für neue Kinder.

Für die Rekursion bedeutet dies, dass der n -te Knoten km Möglichkeiten hat, ein Kind von Knoten j zu werden, respektive $1 + (k + 1)(n - 1) - (m + 1)k = (k + 1)n - km - 2k$

Möglichkeiten, kein Kind von j zu sein. Dies führt zur Rekursion (vgl. [19], Abschnitt 4.1):

$$t_{n,j,m} = ((k+1)n - km - 2k)t_{n-1,j,m} + kmt_{n-1,j,m-1}, \quad \text{für } n > j \geq 1 \text{ und } m \geq 0.$$

Dabei gilt $t_{j,j,0} = t_j$ und $t_{j,j,m} = 0$ für $j \geq 1, m > 0$, wobei t_j die Anzahl der geordneten ansteigenden k -Bäume der Größe j ist. Diese Rekursion führt zu einer Gleichung für die bivariate erzeugende Funktion $T^{[z]}(z, v) = \sum_{n \geq j} \sum_{m \geq 0} t_{n,j,m} \frac{z^{n-j}}{(n-j)!} v^m$. Dazu multiplizieren wir obige Gleichung mit $\frac{z^{n-1-j}}{(n-1-j)!} v^m$ und summieren auf. Um es übersichtlich zu halten, werden die Terme zuerst einzeln betrachtet und erst am Ende in die Gleichung eingesetzt. Mit der üblichen Schreibweise, bei der die Variable im Index die Ableitung in die genannte Richtung bezeichnet, ergibt sich für die linke Seite:

$$\sum_{\substack{n \geq j+1 \\ m \geq 0}} \frac{t_{n,j,m}}{(n-1-j)!} z^{n-1-j} v^m = \sum_{\substack{n \geq j+1 \\ m \geq 0}} \frac{(n-j)t_{n,j,m}}{(n-j)!} z^{n-1-j} v^m = T_z^{[j]}(z, v)(z, v),$$

beziehungsweise für die rechte Seite:

$$\begin{aligned} \sum_{\substack{n \geq j+1 \\ m \geq 0}} \frac{(k+1)nt_{n-1,j,m}}{(n-1-j)!} z^{n-1-j} v^m &= (k+1) \sum_{\substack{n \geq j \\ m \geq 0}} \left(\frac{t_{n,j,m}}{(n-1-j)!} + \frac{(j+1)t_{n,j,m}}{(n-j)!} \right) z^{n-j} v^m \\ &= (k+1)(zT_z^{[j]}(z, v) + (j+1)T^{[j]}(z, v)), \\ \sum_{\substack{n \geq j+1 \\ m \geq 0}} \frac{(-km)t_{n-1,j,m}}{(n-1-j)!} z^{n-1-j} v^m &= (-kv) \sum_{\substack{n \geq j+1 \\ m \geq 1}} \frac{mt_{n-1,j,m}}{(n-1-j)!} z^{n-1-j} v^{m-1} \\ &= (-kv)T_v^{[j]}(z, v) \\ \sum_{\substack{n \geq j+1 \\ m \geq 0}} \frac{(-2k)t_{n-1,j,m}}{(n-1-j)!} z^{n-1-j} v^m &= (-2k)T^{[j]}(z, v), \\ \sum_{\substack{n \geq j+1 \\ m \geq 0}} \frac{kmt_{n-1,j,m-1}}{(n-1-j)!} z^{n-1-j} v^m &= kv \sum_{\substack{n \geq j \\ m \geq 1}} \frac{m}{(n-j)!} t_{n-1,j,m-1} z^{n-j-1} v^m \\ &= kv \left(v \sum_{\substack{n \geq j \\ m \geq 1}} \frac{(m-1)}{(n-j)!} t_{n,j,m-1} z^{n-j-1} v^{m-1} \right. \\ &\quad \left. + \sum_{\substack{n \geq j \\ m \geq 1}} \frac{1}{(n-j)!} t_{n,j,m-1} z^{n-j-1} v^m \right) \\ &= kv^2 T_v^{[j]}(z, v) + kv T^{[j]}(z, v). \end{aligned}$$

Dies führt zu einer partiellen Differenzialgleichung:

$$T_z^{[j]}(z, v) = (k+1)zT_z^{[j]}(z, v) + kv(v-1)T_v^{[j]}(z, v) + ((k+1)(j+1) - 2k + kv)T^{[j]}(z, v),$$

beziehungsweise mit der Randbedingung:

$$(1 - (k + 1)z)T^{[j]} + kv(1 - v)T^{[j]} - (k(j - 1 + v) + j + 1)T^{[j]} = 0, \quad T^{[j]}(0, v) = t_j. \quad (3.1)$$

Diese partielle Differentialgleichung wird mit der Methode der Charakteristiken auf eine gewöhnliche Differentialgleichung zurückgeführt und gelöst (vgl. [19], Abschnitt 4.1). Dazu sei $z = z(r, s)$, $v = v(r, s)$ und $\tilde{T}^{[j]}(r, s) = T^{[j]}(z(r, s), v(r, s))$. Wenn die beiden Gleichungen

$$\begin{aligned} \frac{\partial z}{\partial r} &= 1 - (k + 1)z, \\ \frac{\partial v}{\partial r} &= kv(1 - v) \end{aligned} \quad (3.2)$$

erfüllt sind, vereinfacht sich Formel (3.1) zu

$$\frac{\partial}{\partial r} \tilde{T}^{[j]}(r, s) = \frac{\partial T^{[j]}}{\partial z} \frac{\partial z}{\partial r} + \frac{\partial T^{[j]}}{\partial v} \frac{\partial v}{\partial r} = (k(j - 1 + v) + j + 1)T^{[j]}(z(s, r), v(s, r)),$$

wobei auf der rechten Seite auch noch die Koordinatentransformation vollzogen werden muss. Lösen wir zuerst die gewöhnlichen Differentialgleichungen (3.2). Bei z ist eine homogene Lösung leicht ersichtlich:

$$\begin{aligned} z_r^{[h]} &= -(k + 1)z, \\ \Rightarrow z^{[h]} &= ce^{-(k+1)r}. \end{aligned}$$

Mit der Methode der Variation der Konstanten ($c = c(r)$) ergibt sich:

$$\begin{aligned} c_r e^{-(k+1)r} - c(k + 1)e^{-(k+1)r} &= 1 - (k + 1)ce^{-(k+1)r}, \\ c_r &= e^{(k+1)r}, \\ \Rightarrow c(r) &= \frac{1}{k + 1} e^{(k+1)r}. \end{aligned}$$

Zusammengefasst ergibt sich für die Lösung:

$$z(r, s) = e^{-(k+1)r} \left(\frac{e^{k+1r}}{k + 1} + c_1(s) \right) = \frac{1}{k + 1} + c_1(s)e^{-(k+1)r}.$$

Analoges führen wir für $v(r, s)$ durch:

$$v_r = kv(1 - v) = (-k)v^2 + kv,$$

wobei dies eine Bernoullische Differentialgleichung ist. Mit der Transformation $y := v^{-1}$ ergibt sich

$$\begin{aligned} \frac{-1}{y^2} y_r &= (-k) \frac{1}{y^2} + \frac{k}{y}, \\ y_r &= k - ky. \end{aligned}$$

Die homogene Lösung ist wieder mit $y^{[h]} = ce^{-kr}$ ersichtlich und für eine Partikuläre gilt:

$$\begin{aligned} c_r e^{-kr} &= k, \\ \Rightarrow c(r) &= e^{kr}. \end{aligned}$$

Für die Hilfsfunktion y erhalten wir:

$$y = e^{-kr}(e^{kr} + c_2(s)),$$

beziehungsweise für die ursprüngliche Koordinate:

$$v(r, s) = \frac{e^{kr}}{c_2(s) + e^{kr}}.$$

Die Konstanten wählen wir so, dass für die Transformation der Randbedingung $T^{[j]}(0, v) = \tilde{T}^{[j]}(0, s)$ gilt:

$$\begin{aligned} 0 = z(0, s) &= \frac{1}{k+1} + c_1(s) \quad \Rightarrow \quad c_1(s) = -\frac{1}{k+1}, \\ s = v(0, s) &= \frac{1}{c_2(s) + 1} \quad \Rightarrow \quad c_2(s) = \frac{1-s}{s}. \end{aligned}$$

Bevor wir uns der Differentialgleichung zuwenden, betrachten wir noch die Rücktransformationen.

$$\begin{aligned} z(r, s) &= \frac{1 - e^{-(k+1)r}}{k+1}, \\ 1 - (k+1)z &= e^{-(k+1)r}, \\ r &= -\frac{1}{k+1} \log(1 - (k+1)z) = \log\left(\left(1 - (k+1)z\right)^{-\frac{1}{k+1}}\right). \end{aligned}$$

Der Term $1 - (k+1)z$ kommt im Weiteren öfters vor, daher definieren wir $q(z) := 1 - (k+1)z$ und damit stellen wir auch s durch die ursprünglichen Koordinaten dar:

$$\begin{aligned} v &= \frac{e^{kr}}{\frac{1-s}{s} + e^{kr}} = \frac{e^{kr}}{s^{-1} - 1 + e^{kr}}, \\ \frac{e^{kr}}{v} &= s^{-1} - 1 + e^{kr}, \\ s^{-1} &= \frac{e^{kr} + v(1 - e^{kr})}{v}, \\ s &= \frac{v}{e^{kr} + v(1 - e^{kr})} = \frac{v}{q^{-\frac{k}{k+1}} + v(1 - q^{-\frac{k}{k+1}})} \\ &= \frac{v}{q^{-\frac{k}{k+1}}(1 - v(1 - q^{\frac{k}{k+1}}))}. \end{aligned}$$

Mit diesen Vorbereitungen kann die separable Differentialgleichung gelöst werden:

$$\begin{aligned} \frac{\partial}{\partial r} \tilde{T}^{[j]} &= \left(k \left(j - 1 + \frac{e^{kr}}{c_2 + e^{kr}} \right) + j + 1 \right) \tilde{T}^{[j]}, \\ \frac{\tilde{T}_r^{[j]}}{\tilde{T}^{[j]}} &= \left(k \left(j - 1 + \frac{e^{kr}}{c_2 + e^{kr}} \right) + j + 1 \right). \end{aligned}$$

Die Integration der rechten Seite liefert:

$$\int k(j-1) + j + 1 + k \frac{e^{kr}}{c_2(s) + e^{kr}} dr = r(k(j-1) + j + 1) + \log(c_2 + e^{kr}) + \widehat{c}(s),$$

wobei dies für die Gleichung folgendes bedeutet:

$$\widetilde{T}^{[j]} = c(s) \exp(r(k(j-1) + j + 1) + \log(c_2 + e^{kr})).$$

Mit der Randbedingung kann die Integrationskonstante bestimmt werden:

$$\begin{aligned} t_j &= T^{[j]}(0, v(0, s)) = \widetilde{T}^{[j]}(0, s) = c(s)(c_2(s) + 1) \\ &= c(s)(s^{-1} - 1 + 1) = \frac{c(s)}{s}, \\ &\Rightarrow c(s) = t_j s. \end{aligned}$$

Nun werden die Koordinaten wieder zurücktransformiert:

$$\begin{aligned} \widetilde{T}^{[j]} &= t_j s \exp(r(k(j-1) + j + 1) + \log(c_2 + e^{kr})) \\ &= t_j s \exp(r(k(j-1) + j + 1) + \log(v) - \log(e^{kr})), \\ T^{[j]} &= \frac{vt_j}{q^{-\frac{k}{k+1}} \left(1 - v \left(1 - q^{\frac{k}{k+1}}\right)\right)} q^{-\frac{kj-k+j+1}{k+1}} v^{-1} q^{-\frac{k}{k+1}} \\ &= \frac{t_j}{\left(1 - v \left(1 - q^{\frac{k}{k+1}}\right)\right) q^{\frac{kj-k+j+1}{k+1}}} \\ &= \frac{t_j}{\left(1 - v \left(1 - (1 - (k+1)z)^{\frac{k}{k+1}}\right)\right) (1 - (k+1)z)^{\frac{kj-k+j+1}{k+1}}}. \end{aligned}$$

Bezeichne $Y_{n,j}$ die Zufallsvariable des Knotengrades vom Knoten $j \geq 1$ in einem geordneten, ansteigenden k -Baum der Größe n , wenn diese gleichverteilt sind. Aufgrund der Gleichverteilung können wir aus der bivariaten erzeugenden Funktion die Verteilung von $Y_{n,j}$ bestimmen:

$$\mathbb{P}\{Y_{n,j} = m\} = \frac{t_{n,j,m}}{t_n} = \frac{(n-j)!}{t_n} [z^{n-j} v^m] T^{[j]}(z, v).$$

Die Berechnung wird der Übersicht halber aufgeteilt. Zuerst betrachten wir den von z und v unabhängigen Faktor:

$$\frac{(n-j)! t_j}{t_n} = \frac{(n-j)! j! (k+1)^j \binom{j - \frac{k}{k+1}}{j}}{n! (k+1)^n \binom{n - \frac{k}{k+1}}{n}} = \frac{\binom{j - \frac{k}{k+1}}{j}}{\binom{n}{j} \binom{n - \frac{k}{k+1}}{n} (k+1)^{n-j}}.$$

Die Variable v , welche den ausgehenden Grad markiert, kommt nur in einem Faktor des Bruchs von $T^{[j]}$ vor. Daher betrachten wir nur diesen Term und berechnen den m -ten Koeffizienten von v :

$$[v^m] \frac{1}{1 - v \left(1 - q^{\frac{k}{k+1}}\right)} = \left(1 - q^{\frac{k}{k+1}}\right)^m = \sum_{l=0}^m \binom{m}{l} (-1)^l (1 - (k+1)z)^{\frac{lk}{k+1}}.$$

Für das Ablesen der Koeffizienten bedeutet dies:

$$\begin{aligned}
[z^{n-j}v^m] \frac{T^{[j]}(z, v)}{t_j} &= [z^{n-j}] \sum_{l=0}^m \binom{m}{l} (-1)^l (1 - (k+1)z)^{\frac{kl+kj-k+j+1}{k+1}} \\
&= \sum_{l=0}^m \binom{m}{l} (-1)^l (-1)^{n-j} (k+1)^{n-j} \binom{\frac{k(l+1)-1}{k+1} - j}{n-j} \\
&= \sum_{l=0}^m \binom{m}{l} (-1)^l (-1)^{n-j} (k+1)^{n-j} \\
&\quad \cdot \frac{\left(\frac{k(l+1)-1}{k+1} - j\right) \cdots \left(\frac{k(l+1)-1}{k+1} - j - n + j + 1\right)}{(n-j)!} \\
&= \sum_{l=0}^m \binom{m}{l} (-1)^l (-1)^{n-j} (k+1)^{n-j} \frac{\left(\frac{k(l+1)-1}{k+1} - j\right) \cdots \left(\frac{k(l+2)}{k+1} - n\right)}{(n-j)!} \\
&= \sum_{l=0}^m \binom{m}{l} (-1)^l (k+1)^{n-j} \binom{n - \frac{k(l+2)}{k+1}}{n-j},
\end{aligned}$$

und dies ergibt insgesamt

$$\begin{aligned}
\mathbb{P}\{Y_{n,j} = m\} &= \frac{\binom{j - \frac{k}{k+1}}{j}}{\binom{n}{j} \binom{n - \frac{k}{k+1}}{n-j}} (k+1)^{n-j} \sum_{l=0}^m \binom{m}{l} (-1)^l (k+1)^{n-j} \binom{n - \frac{k(l+2)}{k+1}}{n-j} \\
&= \frac{\binom{j - \frac{k}{k+1}}{j}}{\binom{n}{j} \binom{n - \frac{k}{k+1}}{n-j}} \sum_{l=0}^m \binom{m}{l} (-1)^l \binom{n - \frac{k(l+2)}{k+1}}{n-j}.
\end{aligned}$$

Wurzelknoten: Es bleibt noch übrig, den ausgehenden Knotengrad für die Wurzelknoten zu berechnen. Aufgrund der Symmetrie ist die Wahrscheinlichkeitsverteilung gleich für jeden Knoten der Wurzel. Wir berechnen den Grad für den Knoten 0_1 und bezeichnen mit $t_{n,0,m}$ die Mächtigkeit der geordneten ansteigenden k -Wurzelbäume, bei denen der Knoten 0_1 genau m Kinder hat. Das einzig Essentielle, welches sich ändert, sind die freien Plätze für Kinder, wenn der Knoten 0_1 noch keine Kinder hat. Dies ist genau ein Platz, denn, wenn der Knoten 0_1 keine Kinder hat, besteht der k -Baum nur aus der Wurzel. Mit jedem Kind von 0_1 erhöhen sich wie bei den anderen Knoten auch die Plätze für Kinder um km . Dies ergibt die Rekursion:

$$t_{n,0,m} = (1 + (k-1)(n-1) - (1+km))t_{n-1,0,m} + (1+k(m-1))t_{n-1,0,m-1},$$

für $n \geq 1, m \geq 0$ und mit den Startwerten $t_{0,0,0} = 1$ und $t_{0,0,m} = 0$ für $m \geq 1$. Wir verwenden nun keine „verschobene“ bivariate erzeugende Funktion, sondern $T^{[0]}(z, v) =$

$\sum_{n \geq 0} \sum_{m \geq 0} t_{n,0,m} \frac{z^n}{n!} v^m$. Aufsummiert ergeben die einzelnen Terme:

$$\begin{aligned} \sum_{\substack{n \geq 1 \\ m \geq 0}} \frac{t_{n,0,m}}{(n-1)!} z^{n-1} v^m &= T_z^{[0]}, \\ \sum_{\substack{n \geq 1 \\ m \geq 0}} \frac{(k+1)(n-1)t_{n-1,0,m}}{(n-1)!} z^{n-1} v^m &= (k+1)zT_z^{[0]}, \\ \sum_{\substack{n \geq 1 \\ m \geq 0}} \frac{km t_{n-1,0,m}}{(n-1)!} z^{n-1} v^m &= kvT_v^{[0]}, \\ \sum_{\substack{n \geq 1 \\ m \geq 0}} \frac{t_{n-1,0,m-1}}{(n-1)!} z^{n-1} v^m &= vT_v^{[0]}, \\ \frac{k(m-1)t_{n-1,0,m-1}}{(n-1)!} z^{n-1} v^m &= kv^2T_v^{[0]}. \end{aligned}$$

Dies liefert die partielle Differentialgleichung

$$(1 - (k+1)z)T_z^{[0]} + kv(1-v)T_v^{[0]} - vT^{[0]} = 0, \quad T^{[0]}(0, v) = 1.$$

Die Faktoren bei den Ableitungen sind die gleichen wie in der Gleichung (3.1) für einen beliebigen Knoten $j \geq 1$. Daher können wir auch die gleiche Methode mit derselben Transformation zum Lösen verwenden. In den Koordinaten (r, s) lautet die Differentialgleichung:

$$\tilde{T}_r^{[0]}(r, s) = \frac{e^{kr}}{c_2(s) + e^{kr}} \tilde{T}^{[0]}(r, s).$$

Die Lösung dieser separablen Differentialgleichung ist:

$$\begin{aligned} \tilde{T}^{[0]} &= c(s) \exp\left(\frac{1}{k} \log(c_2 + e^{kr})\right) \\ &= c(s) \exp\left(\frac{1}{k} \log\left(\frac{c_2 + e^{kr}}{e^{kr}} + \log(e^{kr})\right)\right). \end{aligned}$$

Mit der Randbedingung können wir die Integrationskonstante bestimmen:

$$1 = \tilde{T}^{[0]}(0, s) = c(s)(s^{-1} - 1 + 1)^{\frac{1}{k}} \Rightarrow c(s) = s^{\frac{1}{k}}.$$

Wenn wir den Koordinatenwechsel zurück durchführen, ergibt sich für die erzeugende Funktion folgende Lösung:

$$\begin{aligned} T^{[0]}(z, v) &= \frac{v^{\frac{1}{k}}}{q^{-\frac{1}{k+1}} \left(1 - v \left(1 - q^{\frac{k}{k+1}}\right)\right)^{\frac{1}{k}}} v^{-\frac{1}{k}} q^{-\frac{1}{k+1}} \\ &= \frac{1}{\left(1 - v \left(1 - q^{\frac{k}{k+1}}\right)\right)^{\frac{1}{k}}}. \end{aligned}$$

Daraus können wir die Koeffizienten ablesen und die Zufallsvariable $Y_{n,0}$ des ausgehenden Knotengrades vom Knoten 0_1 in einem geordneten ansteigenden k -Baum bestimmen:

$$\mathbb{P}\{Y_{n,0} = m\} = \frac{t_{n,j,m}}{t_n} = \frac{n!}{t_n} [z^n v^m] T^{[0]}(z, v),$$

wobei für die Koeffizienten gilt:

$$\begin{aligned} [z^n v^m] T^{[0]}(z, v) &= [z^n v^m] \frac{1}{\left(1 - v \left(1 - q^{\frac{k}{k+1}}\right)\right)^{\frac{1}{k}}} \\ &= [z^n] \binom{-\frac{1}{k}}{m} (-1)^m \left(1 - q^{\frac{k}{k+1}}\right)^m \\ &= [z^n] \binom{m-1 + \frac{1}{k}}{m} \sum_{l=0}^m \binom{m}{l} (-1)^l q^{\frac{lk}{k+1}} \\ &= \binom{m - \frac{k-1}{k}}{m} \sum_{l=0}^m \binom{m}{l} (-1)^l \binom{\frac{lk}{k+1}}{n} (-1)^n (k+1)^n \\ &= \binom{m - \frac{k-1}{k}}{m} \sum_{l=0}^m \binom{m}{l} (-1)^l (k+1)^n \binom{n-1 - \frac{lk}{k+1}}{n}. \end{aligned}$$

Mit $t_n = n!(k+1)^n \binom{n-\frac{k-1}{k}}{n}$ ergibt sich für die Verteilung (vgl. [19], Theorem 2):

$$\mathbb{P}\{Y_{n,0} = m\} = \frac{\binom{m - \frac{k-1}{k}}{m}}{\binom{n - \frac{k-1}{k}}{n}} \sum_{l=0}^m \binom{m}{l} (-1)^l \binom{n-1 - \frac{lk}{k+1}}{n}.$$

Beliebig hinzugefügter Knoten: Bisher haben wir immer einen bestimmten Knoten j oder 0_1 betrachtet und für diesen den ausgehenden Grad berechnet. Nun werden wir einen beliebigen eingefügten Knoten im k -Baum betrachten und dafür eine Verteilungsfunktion angeben. Die Grenzverteilung für $n \rightarrow \infty$ werden wir auch noch bestimmen. Es sei \bar{Y}_n die Zufallsvariable des ausgehenden Grades eines beliebigen eingefügten Knotens. Dann gilt:¹¹

$$\begin{aligned} \mathbb{P}\{\bar{Y}_n = m\} &= \frac{1}{n} \sum_{j=1}^n \mathbb{P}\{Y_{n,j} = m\} \\ &= \frac{1}{n} \sum_{j=1}^n \frac{\binom{j - \frac{k}{k+1}}{j}}{\binom{n}{j} \binom{n - \frac{k}{k+1}}{n}} \sum_{l=0}^m \binom{m}{l} (-1)^l \binom{n - \frac{k(l+2)}{k+1}}{n-j} \\ &= \frac{1}{n \binom{n - \frac{k}{k+1}}{n}} \sum_{l=0}^m \binom{m}{l} (-1)^l \sum_{j=1}^n \frac{\binom{j - \frac{k}{k+1}}{j} \binom{n - \frac{k(l+2)}{k+1}}{n-j}}{\binom{n}{j}} \\ &= \frac{1}{n \binom{n - \frac{k}{k+1}}{n}} \sum_{l=0}^m \frac{(-1)^l \binom{m}{l}}{k(l+2) + 1} \left(\binom{n + \frac{1}{k+1}}{n} - \binom{n - \frac{k(l+2)}{k+1}}{n} \right). \end{aligned}$$

¹¹Der letzte Umformungsschritt verwendet eine hypergeometrische Identität und wurde mit Maple 18.01 durchgeführt (siehe <https://www.maplesoft.com/products/Maple/>). Für weitere Informationen über Hypergeometrische Funktionen und Algorithmen in diesem Zusammenhang siehe [20].

Um die Konvergenz in der Verteilung von \bar{Y}_n zu bestimmen, betrachten wir zuerst die Limiten der einzelnen Brüche der Binomialkoeffizienten:

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{\binom{n + \frac{1}{k+1}}{n}}{\binom{n - \frac{k}{k+1}}{n}} &= \lim_{n \rightarrow \infty} \frac{\left(n + \frac{1}{k}\right) \left(n - \frac{k}{k+1}\right)^{n-1}}{n \left(n - \frac{k}{k+1}\right)^{n-1} \left(1 - \frac{k}{k+1}\right)} = k + 1, \\ \lim_{n \rightarrow \infty} \frac{\binom{n - \frac{k(l+2)}{k+1}}{n}}{\binom{n - \frac{k}{k+1}}{n}} &\stackrel{\text{L. 1.8}}{=} \lim_{n \rightarrow \infty} \frac{c\left(\frac{k(l+2)}{k+1}, \frac{k}{k+1}\right)}{n} \frac{n^{-\frac{k(l+2)}{k+1}}}{n^{-\frac{k}{k+1}}} = 0. \end{aligned}$$

Für den Limes \bar{Y} bezüglich der Konvergenz in Verteilung von \bar{Y}_n ($\bar{Y}_n \xrightarrow{(d)} \bar{Y}$) ergibt dies:

$$\begin{aligned} \mathbb{P}\{\bar{Y} = m\} &= \lim_{n \rightarrow \infty} \mathbb{P}\{\bar{Y}_n = m\} = (k + 1) \sum_{l=0}^m \frac{(-1)^l \binom{m}{l}}{k(l+2) + 1} \\ &= \frac{k + 1}{k} \sum_{l=0}^m \frac{(-1)^l \binom{m}{l}}{l + \left(\frac{2k+1}{k} - 1\right) + 1}, \end{aligned}$$

was sich mit Satz 1.9 umformen lässt zu

$$\begin{aligned} \mathbb{P}\{\bar{Y} = m\} &= \frac{k + 1}{k} \frac{1}{\left(\frac{2k+1}{k} + m\right) \binom{\frac{2k+1}{k} - 1 + m}{m}} \\ &= \frac{k + 1}{k(m + 1) \binom{m+2+\frac{1}{k}}{m+1}}. \end{aligned}$$

Bei $\mathbb{P}\{\bar{Y} = m\}$ interessieren wir uns noch für das asymptotische Verhalten bezüglich m . Dazu verwenden wir die Darstellung des Binomialkoeffizientens über die Gammafunktion und deren Annäherung mit der Stirling-Formel:

$$\begin{aligned} \mathbb{P}\{\bar{Y} = m\} &= \frac{k + 1}{k(m + 1) \binom{m+2+\frac{1}{k}}{m+1}} \stackrel{\text{L. 1.7}}{\sim} \frac{k + 1}{k(m + 1)} \Gamma\left(2 + \frac{1}{k}\right) m^{-1-\frac{1}{k}} \\ &= \frac{k + 1}{k} \Gamma\left(2 + \frac{1}{k}\right) m^{-2-\frac{1}{k}}. \end{aligned}$$

Somit erfüllt $\mathbb{P}\{\bar{Y} = m\}$ asymptotisch ein Potenzgesetz mit Exponent $2 + \frac{1}{k}$ (vgl. [19], Theorem 3).

3.2.2 Ansteigende k -Bäume

Hinzugefügter Knoten: Bei den ansteigenden k -Bäumen kann analog zu den geordneten ansteigenden k -Bäumen vorgegangen werden. Die Möglichkeiten, einen weiteren Knoten einem ansteigenden k -Baum hinzuzufügen, entsprechen genau den k -Cliques. Aus Lemma 1.1 wissen wir, dass ein ansteigender k -Baum mit $n - 1$ Knoten genau $1 + k(n - 1)$ Cliques besitzt. Ein Knoten j ist durch das Verbinden mit einer k -Clique beim Hinzufügen von j Teilmenge von mindestens k Cliques. Durch jedes Kind v von j entstehen $k - 1$ neue Cliques, die die Knoten j und v umfassen. Daher ist ein Knoten

mit einem ausgehenden Grad m eine Teilmenge von $k + (k - 1)m$ Cliques. Dies führt zur Rekursion

$$t_{n,j,m} = (1 + k(n - 2) - (k - 1)m)t_{n-1,j,m} + (k + (k - 1)(m - 1))t_{n-1,j,m-1},$$

für $n > j \geq 1$ und $m \geq 0$, wobei $t_{n,j,m}$ die Anzahl der ansteigenden k -Bäume ist, bei denen der Knoten $j \geq 1$ genau m Kinder hat. Die Anfangsbedingung für die Rekursion sind $t_{j,j,0} = t_j$ und $t_{j,j,m} = 0$ für $j \geq 1, m > 0$. Dabei sei t_j die Anzahl der ansteigenden k -Bäume. Dies kann in eine Gleichung für die erzeugende Funktion $T^{[j]}(z, v) = \sum_{n \geq j} \sum_{m \geq 0} t_{n,j,m} \frac{z^{n-j}}{(n-j)!} v^m$ übergeführt werden. Für die einzelnen Terme gilt:

$$\begin{aligned} \sum_{\substack{n \geq j+1 \\ m \geq 0}} \frac{k(n-2)t_{n-1,j,m}}{(n-j-1)!} z^{n-j-1} v^m &= k \sum_{\substack{n \geq j+1 \\ m \geq 0}} \frac{(n-j-1+j-1)t_{n-1,j,m}}{(n-j-1)!} z^{n-j-1} v^m \\ &= kzT_z^{[j]} + k(j-1)T^{[j]}, \\ \sum_{\substack{n \geq j+1 \\ m \geq 0}} \frac{(k-1)mt_{n-1,j,m}}{(n-j-1)!} z^{n-j-1} v^m &= (k-1)vT_v^{[j]}, \\ \sum_{\substack{n \geq j+1 \\ m \geq 0}} \frac{(k-1)(m-1)t_{n-1,j,m-1}}{(n-j-1)!} z^{n-j-1} v^m &= (k-1)v^2T_v^{[j]}, \\ \sum_{\substack{n \geq j+1 \\ m \geq 0}} \frac{kt_{n-1,j,m-1}}{(n-j-1)!} z^{n-j-1} v^m &= kvT^{[j]}, \\ \sum_{\substack{n \geq j+1 \\ m \geq 0}} \frac{t_{n-1,j,m}}{(n-j-1)!} z^{n-j-1} v^m &= T^{[j]}, \\ \sum_{\substack{n \geq j+1 \\ m \geq 0}} \frac{t_{n,j,m}}{(n-j-1)!} z^{n-j-1} v^m &= T_z^{[j]}. \end{aligned}$$

Unter der Verwendung der Rekursion ergibt dies:

$$T_z^{[j]} = T^{[j]} + kzT_z^{[j]} + k(j-1)T^{[j]} - (k-1)vT_v^{[j]} + kvT^{[j]} + (k-1)v^2T_v^{[j]}$$

beziehungsweise

$$(1 - kz)T_z^{[j]} + (k-1)v(1-v)T_v^{[j]} - (1 + k(j-1) + kv)T^{[j]} = 0, \quad \text{mit } T^{[j]}(0, v) = t_j. \quad (3.3)$$

Diese Gleichung ist ähnlich aufgebaut wie die Gleichung (3.1) für die geordneten ansteigenden k -Bäume. In den ersten beiden Termen müsste man nur k durch $k + 1$ ersetzen. Dieser Unterschied um 1 ist zurückzuführen auf den einen zusätzlichen Platz, der bei den geordneten ansteigenden k -Bäumen entsteht, wenn man weitere Knoten zu einer Clique mit mindestens einem Kind hinzufügt. Da sich die Gleichungen kaum unterscheiden, wird im Folgenden nicht mehr alles noch einmal vollständig durchgerechnet, insbesondere für die Koordinatentransformation bei der Methode der Charakteristiken. Es können die

Rechnungen von den geordneten ansteigenden k -Bäumen einfach adaptiert werden. Für die Koordinatentransformation ($z = z(r, s)$, $v = v(r, s)$) ergibt sich:

$$\begin{aligned}\frac{\partial z}{\partial r} = 1 - kz &\Leftrightarrow z = \frac{1}{k} + c_1(s)e^{-kr}, \\ \frac{\partial v}{\partial r} = (k-1)v(1-v) &\Leftrightarrow v = \frac{e^{(k-1)r}}{c_2(s) + e^{(k-1)r}},\end{aligned}$$

mit den Konstanten:

$$\begin{aligned}0 = z(0, s) = \frac{1}{k} + c_1(s) &\Rightarrow c_1(s) = -\frac{1}{k}, \\ s = v(0, s) = \frac{1}{c_2(s) + 1} &\Rightarrow c_2(s) = \frac{1-s}{s}.\end{aligned}$$

Mit der Bezeichnung $q(z) := 1 - kz$ ergibt sich für die Rücktransformation:

$$\begin{aligned}r &= \log\left((1 - kz)^{-\frac{1}{k}}\right), \\ s &= \frac{v}{q^{-\frac{k-1}{k}}\left(1 - v\left(1 - q^{\frac{k-1}{k}}\right)\right)}.\end{aligned}$$

Die partielle Differentialgleichung transformiert sich zu der gewöhnlichen Differentialgleichung:

$$\tilde{T}_r^{[j]}(r, s) = \left(1 + k(j-1) + k\frac{e^{(k-1)r}}{c_2 + e^{(k-1)r}}\right)\tilde{T}^{[j]},$$

dessen Lösung

$$\begin{aligned}\tilde{T}^{[j]} &= c(s) \exp\left(r(1 + k(j-1)) + \frac{k}{k-1} \log(c_2 + e^{(k-1)r})\right) \\ &= c(s) \exp\left(r(1 + k(j-1)) + \frac{k}{k-1} \log\left(\frac{c_2 + e^{(k-1)r}}{e^{(k-1)r}}\right) + \frac{k}{k-1} \log(e^{(k-1)r})\right)\end{aligned}$$

ist. Für die Integrationskonstante gilt:

$$\begin{aligned}t_j = \tilde{T}^{[j]}(0, s) &= c(s)(c_2(s) + 1)^{\frac{k}{k-1}} = c(s)(s^{-1} - 1 + 1)^{-\frac{k}{k-1}} \\ &= c(s)s^{\frac{k}{k-1}}, \\ \Rightarrow c(s) &= t_j s^{\frac{k}{k-1}}.\end{aligned}$$

Wenn dies nun zurück transformiert wird, erhält man:

$$\begin{aligned}T^{[j]}(z, v) &= t_j \frac{v^{\frac{k}{k-1}}}{q^{-1}\left(1 - v\left(1 - q^{\frac{k-1}{k}}\right)\right)^{\frac{k}{k-1}}} q^{-\frac{1+k(j-1)}{k}} v^{-\frac{k}{k-1}} q^{-1} \\ &= \frac{t_j}{\left(1 - v\left(1 - q^{\frac{k-1}{k}}\right)\right)^{\frac{k}{k-1}} q^{\frac{1+k(j-1)}{k}}}.\end{aligned}$$

Es sei $Y_{n,j}$ die Zufallsvariable des Knotengrades vom Knoten $j \geq 1$ in einem ansteigenden k -Baum der Größe n , wenn diese gleichverteilt sind. So gilt:

$$\mathbb{P}\{Y_{n,j} = m\} = \frac{t_{n,j,m}}{t_n} = \frac{(n-j)!}{t_n} [z^{n-j} v^m] T^{[j]}(z, v).$$

Betrachten wir zuerst den Faktor $\frac{(n-j)!t_j}{t_n}$, dieser ergibt:

$$\frac{(n-j)!t_j}{t_n} = (n-j)! \frac{j!k^j \binom{j-\frac{k-1}{k}}{j-\frac{k-1}{k}}}{n!k^n \binom{n-\frac{k-1}{k}}{n-\frac{k-1}{k}}} = \frac{\binom{j-\frac{k-1}{k}}{j-\frac{k-1}{k}}}{k^{n-j} \binom{n}{j} \binom{n-\frac{k-1}{k}}{n-\frac{k-1}{k}}}.$$

Beim restlichen Faktor extrahieren wir zuerst nur einmal den m -ten Koeffizienten von v :

$$\begin{aligned} [v^m] \frac{1}{\left(1 - v(1 - q^{\frac{k-1}{k}})\right)^{\frac{k}{k-1}}} &= \binom{-\frac{k}{k-1}}{m} (-1)^m \left(1 - q^{\frac{k-1}{k}}\right)^m \\ &= \binom{m-1 + \frac{k}{k-1}}{m} \sum_{l=0}^m \binom{m}{l} (-1)^l q^{l\frac{k-1}{k}}. \end{aligned}$$

Für das Extrahieren der Koeffizienten bedeutet dies:

$$\begin{aligned} [z^{n-j} v^m] \frac{T^{[j]}(z, v)}{t_j} &= \binom{m + \frac{1}{k-1}}{m} \sum_{l=0}^m \binom{m}{l} (-1)^l [z^{n-j}] q^{\frac{l(k-1)-1-k(j-1)}{k}} \\ &= \binom{m + \frac{1}{k-1}}{m} \sum_{l=0}^m \binom{m}{l} (-1)^l \binom{\frac{l(k-1)-1-k(j-1)}{k}}{n-j} (-1)^{n-j} k^{n-j} \\ &= \binom{m + \frac{1}{k-1}}{m} \sum_{l=0}^m \binom{m}{l} (-1)^l \binom{n-j-1 - \frac{l(k-1)-1-k(j-1)}{k}}{n-j} k^{n-j} \\ &= \binom{m + \frac{1}{k-1}}{m} \sum_{l=0}^m \binom{m}{l} (-1)^l \binom{n-2 - \frac{l(k-1)-1}{k}}{n-j} k^{n-j}. \end{aligned}$$

Das ergibt insgesamt:

$$\mathbb{P}\{Y_{n,j} = m\} = \frac{\binom{j-\frac{k-1}{k}}{j-\frac{k-1}{k}} \binom{m+\frac{1}{k-1}}{m}}{\binom{n}{j} \binom{n-\frac{k-1}{k}}{n-\frac{k-1}{k}}} \sum_{l=0}^m \binom{m}{l} (-1)^l \binom{n-2 - \frac{l(k-1)-1}{k}}{n-j} k^{n-j}.$$

Wurzelknoten: Wie es ebenso bei den geordneten ansteigenden k -Bäumen der Fall war, ist auch bei den ansteigenden k -Bäumen die Berechnung des ausgehenden Grades für einen Wurzelknoten sehr analog zu den übrigen Knoten. Es ändert sich in der Rekursion wieder nur der Term, welcher den möglichen Plätzen für Kinder bei Grad 0 entspricht. Ein ausgehender Grad von 0 für einen Wurzelknoten bedeutet, dass der ansteigende k -Baum nur aus der Wurzel besteht. Dort gibt es nur eine Möglichkeit, um ein Kind hinzuzufügen. Mit $t_{n,0,m}$ bezeichnen wir die Anzahl der ansteigenden k -Bäume der Größe n , bei denen der Knoten mit der Markierung 0_1 ausgehenden Grad m hat. Die Rekursion für diese Größe lautet:

$$t_{n,0,m} = (1 + k(n-1) - (1 + (k-1)m))t_{n-1,0,m} + (1 + (k-1)(m-1))t_{n-1,0,m-1},$$

mit den Bedingungen $t_{0,0,0} = 1$ und $t_{0,0,m} = 0$ für $m \geq 1$. Für die bivariate erzeugende Funktion $T^{[0]}(z, v) = \sum_{n \geq 0} \sum_{m \geq 0} t_{n,0,m} \frac{z^n}{n!} v^m$ kann die Rekursion mit Hilfe von

$$\begin{aligned} \sum_{\substack{n \geq 1 \\ m \geq 0}} \frac{t_{n,0,m}}{(n-1)!} z^{n-1} v^m &= T_z^{[0]}, \\ \sum_{\substack{n \geq 1 \\ m \geq 0}} \frac{k(n-1)t_{n-1,0,m}}{(n-1)!} z^{n-1} v^m &= kzT_z^{[0]}, \\ \sum_{\substack{n \geq 1 \\ m \geq 0}} \frac{(k-1)t_{n-1,0,m}}{(n-1)!} z^{n-1} v^m &= (k-1)vT_v^{[0]}, \\ \sum_{\substack{n \geq 1 \\ m \geq 0}} \frac{t_{n-1,0,m-1}}{(n-1)!} z^{n-1} v^m &= vT^{[0]}, \\ \sum_{\substack{n \geq 1 \\ m \geq 0}} \frac{(k-1)(m-1)t_{n-1,0,m}}{(n-1)!} z^{n-1} v^m &= (k-1)v^2T_v^{[0]}, \end{aligned}$$

auf eine partielle Differentialgleichung zurückgeführt werden:

$$(1 - kz)T_z^{[0]} + (k-1)v(1-v)T_v^{[0]} - vT^{[0]} = 0, \quad T^{[0]}(0, v) = 1.$$

Diese Gleichung unterscheidet sich nur beim Faktor bei der Funktion $T^{[0]}$ und der Randbedingung von der Gleichung (3.3). So können wir dieselbe Transformation zum Lösen verwenden. Es ergibt sich dann die gewöhnliche Differentialgleichung:

$$\tilde{T}^{[0]}(r, s)_r = \frac{e^{(k-1)r}}{c_2(s) + e^{(k-1)r}} \tilde{T}^{[0]}(r, s),$$

deren Lösung

$$\begin{aligned} \tilde{T}^{[0]} &= c(s) \exp\left(\frac{1}{k-1} \log(c_2 + e^{(k-1)r})\right) \\ &= c(s) \left(\frac{c_2 + e^{(k-1)r}}{e^{(k-1)r}}\right)^{\frac{1}{k-1}} e^r \end{aligned}$$

ist. Die Randbedingung liefert die Integrationskonstante:

$$1 = \tilde{T}^{[0]}(0, s) = c(s) (s^{-1} + 1 - 1)^{\frac{1}{k-1}} \Rightarrow c(s) = s^{\frac{1}{k-1}}.$$

In den ursprünglichen Koordinaten ergibt dies:

$$\begin{aligned} T^{[0]}(z, v) &= \frac{v^{\frac{1}{k-1}}}{q^{-\frac{1}{k}} \left(1 - v \left(1 - q^{\frac{k-1}{k}}\right)\right)^{\frac{1}{k-1}}} v^{-\frac{1}{k-1}} q^{-\frac{1}{k}} \\ &= \frac{1}{\left(1 - v \left(1 - q^{\frac{k-1}{k}}\right)\right)^{\frac{1}{k-1}}}. \end{aligned}$$

Die Zufallsvariable $Y_{n,0}$ des ausgehenden Knotengrades vom Knoten 0_1 in einem ansteigenden k -Baum kann damit bestimmt werden:

$$\mathbb{P}\{Y_{n,0} = m\} = \frac{t_{n,j,m}}{t_n} = \frac{n!}{t_n} [z^n v^m] T^{[0]}(z, v),$$

wobei für das Ablesen der Koeffizienten gilt:

$$\begin{aligned} [z^n v^m] T^{[0]} &= [z^n] \binom{-\frac{1}{k-1}}{m} (-1)^m \sum_{l=0}^m \binom{m}{l} (-1)^l q^{l \frac{k-1}{k}} \\ &= \binom{m-1+\frac{1}{k-1}}{m} \sum_{l=0}^m \binom{m}{l} (-1)^l (-1)^n k^n \binom{l(k-1)}{n} \\ &= \binom{m-\frac{k-2}{k-1}}{m} \sum_{l=0}^m \binom{m}{l} (-1)^l k^n \binom{n-1-\frac{l(k-1)}{k}}{n}. \end{aligned}$$

So ergibt sich für die Verteilung folgendes:

$$\begin{aligned} \mathbb{P}\{Y_{n,0} = m\} &= \frac{n! \binom{m-\frac{k-2}{k-1}}{m}}{n! k^n \binom{n-\frac{k-1}{k}}{n}} \sum_{l=0}^m \binom{m}{l} (-1)^l k^n \binom{n-\frac{l(k-1)+k}{k}}{n} \\ &= \frac{\binom{m-\frac{k-2}{k-1}}{m}}{\binom{n-\frac{k-1}{k}}{n}} \sum_{l=0}^m \binom{m}{l} (-1)^l \binom{n-\frac{l(k-1)+k}{k}}{n}. \end{aligned}$$

Beliebig hinzugefügter Knoten: Nachdem wir die Wurzel betrachtet haben, wenden wir uns wieder den hinzugefügten Knoten zu. Mit \bar{Y}_n soll die Zufallsvariable des ausgehenden Grades eines beliebigen eingefügten Knotens bezeichnet werden. Diese Größe können wir über die Zufallsvariable $Y_{n,j}$ darstellen:¹²

$$\begin{aligned} \mathbb{P}\{\bar{Y}_n = m\} &= \frac{1}{n} \sum_{j=1}^n \mathbb{P}\{Y_{n,j} = m\} \\ &= \frac{1}{n} \sum_{j=1}^n \frac{\binom{j-\frac{k-1}{k}}{j} \binom{m+\frac{1}{k-1}}{m}}{\binom{n}{j} \binom{n-\frac{k-1}{k}}{n}} \sum_{l=0}^m \binom{m}{l} (-1)^l \binom{n-2-\frac{l(k-1)-1}{k}}{n-j} \\ &= \frac{\binom{m+\frac{1}{k-1}}{m} \sum_{l=0}^m (-1)^l \binom{m}{l} \sum_{j=1}^n \frac{\binom{j-\frac{k-1}{k}}{j} \binom{n-2-\frac{l(k-1)-1}{k}}{n-j}}{\binom{n}{j}}}{n \binom{n-\frac{k-1}{k}}{n}} \\ &= \frac{\binom{m+\frac{1}{k-1}}{m} \sum_{l=0}^m \frac{(-1)^l \binom{m}{l}}{l(k-1)+2k} \left(\binom{n+\frac{1}{k}}{n} + \binom{n-2-\frac{l(k-1)-1}{k}}{n} \right)}{n \binom{n-\frac{k-1}{k}}{n}} \end{aligned}$$

Um die Grenzverteilung \bar{Y} zu bestimmen, betrachten wir die beiden von n abhängigen

¹²siehe Fußnote 11 für den letzten Umformungsschritt.

Summanden separat:

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{\binom{n+\frac{1}{k}}{n}}{n \binom{n-1+\frac{1}{k}}{n}} &= \lim_{n \rightarrow \infty} \frac{n + \frac{1}{k}}{n^{\frac{1}{k}}} = k, \\ \lim_{n \rightarrow \infty} \frac{\binom{n-2-\frac{l(k-1)-1}{k}}{n}}{n \binom{n-1+\frac{1}{k}}{n}} &\stackrel{\text{L. 1.8}}{=} c_l \lim_{n \rightarrow \infty} \frac{n^{-2-\frac{l(k-1)-1}{k}}}{n \cdot n^{-1+\frac{1}{k}}} = c_l \lim_{n \rightarrow \infty} n^{-2-\frac{k-1}{k}} = 0. \end{aligned}$$

Damit erhalten wir die Verteilungsfunktion von \bar{Y} in Form von:

$$\begin{aligned} \mathbb{P}\{\bar{Y} = m\} &= k \binom{m + \frac{1}{k-1}}{m} \sum_{l=0}^m \frac{(-1)^l \binom{m}{l}}{l(k-1) + 2k} = \frac{k}{k-1} \binom{m + \frac{1}{k-1}}{m} \sum_{l=0}^m \frac{(-1)^l \binom{m}{l}}{l + \frac{2k}{k-1}} \\ &\stackrel{\text{L. 1.9}}{=} \frac{k}{k-1} \frac{\binom{m+\frac{1}{k-1}}{m}}{\left(m + \frac{2k}{k-1}\right) \binom{m-1+\frac{2k}{k-1}}{m}} = \frac{k}{k-1} \frac{\binom{m+\frac{1}{k-1}}{m}}{(m+1) \binom{m+\frac{2k}{k-1}}{m+1}}. \end{aligned}$$

Die Grenzverteilung untersuchen wir wie auch schon bei den geordneten ansteigenden k -Bäumen auf ihre Asymptotik bezüglich m :

$$\mathbb{P}\{\bar{Y} = m\} \sim \frac{k\Gamma(\frac{2k}{k-1})}{(k-1)\Gamma(\frac{k}{k-1})} \frac{m^{\frac{1}{k-1}}}{m^{2+\frac{2}{k-1}}} = \frac{k\Gamma(\frac{2k}{k-1})}{(k-1)\Gamma(\frac{k}{k-1})} m^{-2-\frac{1}{k-1}}.$$

Der ausgehende Knotengrad der ansteigenden k -Bäume erfüllt daher asymptotisch ein Potenzgesetz mit dem Koeffizienten $2 + \frac{1}{k-1}$.

3.2.3 Ansteigende d -äre k -Bäume

Hinzugefügter Knoten: Nun betrachten wir den ausgehenden Knotengrad für ansteigende d -äre k -Bäume. Zuerst berechnen wir dies für einen Knoten mit der Markierung $j \geq 1$. Dazu sei t_n die Anzahl der ansteigenden d -ären k -Bäume der Größe n , wie in Kapitel 2 berechnet und mit $t_{n,j,m}$ bezeichnen wir die Mächtigkeit der Teilmenge, bei denen der j -te Knoten m Kinder besitzt. Um eine Rekursion für $t_{n,j,m}$ aufzustellen, wird benötigt, wie viele mögliche Plätze für Kinder von j abhängig vom Grad m es gibt. Wenn der Knoten j dem Baum hinzugefügt wird, entstehen k neue Cliques, die den Knoten j umfassen und sie haben jeweils d freie Plätze für mögliche Kinder von j . Beim Hinzufügen eines Kindes v von j ergeben sich $k-1$ neue Cliques, die diese beiden Knoten umfassen und jeweils d freie Plätze aufweisen. Der Knoten v nimmt auch einen Platz ein und somit vergrößern sich mit jedem Kind die Plätze für Kinder um $(k-1)d-1$. Damit gibt es, wenn der Knoten j den Grad m hat, $kd + ((k-1)d-1)m$ Plätze für weitere Kinder. Beim Aufstellen der Rekursion in Abschnitt 2.7 haben wir uns schon überlegt, dass bei einem d -ären k -Baum es $d + (kd-1)(n-1)$ Plätze für den n -ten Knoten gibt. Dies führt zur Rekursion:

$$\begin{aligned} t_{n,j,m} &= ((1-k)d + (kd-1)(n-1) - m((k-1)d-1))t_{n-1,j,m} \\ &\quad + (kd + (m-1)((k-1)d-1))t_{n-1,j,m-1}, \end{aligned}$$

für $n > j \geq 1$ und $m \geq 0$. Die Startwerte sind $t_{j,j,0} = t_j$ für $j \geq 1$ und $t_{j,j,m} = 0$ für $m > 0$. Für die bivariate erzeugende Funktion $T^{[j]}(z, v) = \sum_{n \geq j} \sum_{m \geq 0} t_{n,j,m} \frac{z^{n-j}}{(n-j)!} v^m$ ergibt sich mittels der Nebenrechnung

$$\begin{aligned}
& \sum_{\substack{n \geq j+1 \\ m \geq 0}} \frac{(1-k)dt_{n-1,j,m}}{(n-j-1)!} z^{n-j-1} v^m = (1-k)dT^{[j]}, \\
& \sum_{\substack{n \geq j+1 \\ m \geq 0}} \frac{(kd-1)(n-j+j-1)t_{n-1,j,m}}{(n-j-1)!} z^{n-j-1} v^m = (kd-1)(zT_z^{[j]} + jT^{[j]}), \\
& \sum_{\substack{n \geq j+1 \\ m \geq 0}} \frac{m((k-1)d-1)t_{n-1,j,m}}{(n-j-1)!} z^{n-j-1} v^m = ((k-1)d-1)vT_v^{[j]}, \\
& \sum_{\substack{n \geq j+1 \\ m \geq 0}} \frac{kdt_{n-1,j,m-1}}{(n-j-1)!} z^{n-j-1} v^m = vkdT^{[j]}, \\
& \sum_{\substack{n \geq j+1 \\ m \geq 0}} \frac{(m-1)((k-1)d-1)t_{n-1,j,m-1}}{(n-j-1)!} z^{n-j-1} v^m = v^2((k-1)d-1)T_v^{[j]}, \\
& \sum_{\substack{n \geq j+1 \\ m \geq 0}} \frac{t_{n,j,m}}{(n-j-1)!} z^{n-j-1} v^m = T_z^{[j]},
\end{aligned}$$

folgende partielle Differentialgleichung:

$$(1 - (kd-1)z)T_z^{[j]} + ((k-1)d-1)v(1-v)T_v^{[j]} - ((1-k)d + (kd-1)j + vkd)T^{[j]}, \quad (3.4)$$

mit der Randbedingung $T^{[j]}(0, v) = t_j$. Diese Gleichung hat ähnliche Form wie die Gleichungen (3.1) und (3.3), sodass wir auch hier die gleiche Methode zum Lösen verwenden. Folglich werden nicht alle Schritte vollständig ausgeführt. Für die Koordinatentransformationen $z = z(r, s)$ und $v = v(r, s)$ bedeutet dies:

$$\begin{aligned}
\frac{\partial z}{\partial r} = 1 - (kd-1)z & \Leftrightarrow z = \frac{1}{kd-1} + c_1(s)e^{-(kd-1)r}, \\
\frac{\partial v}{\partial r} = ((k-1)d-1)v(1-v) & \Leftrightarrow v = \frac{e^{((k-1)d-1)r}}{c_2(s) + e^{((k-1)d-1)r}},
\end{aligned}$$

wobei sich die Konstanten ergeben durch:

$$\begin{aligned}
0 = z(0, s) = \frac{1}{kd-1} + c_1(s) & \Rightarrow c_1(s) = -\frac{1}{kd-1}, \\
s = v(0, s) = \frac{1}{c_2(s) + 1} & \Rightarrow c_2(s) = \frac{1-s}{s}.
\end{aligned}$$

Für die Rücktransformation kürzen wir mit $q(z) := 1 - (kd-1)z$ ab:

$$\begin{aligned}
r &= \log \left((1 - (kd-1)z)^{-\frac{1}{kd-1}} \right), \\
s &= \frac{v}{q^{-\frac{(k-1)d-1}{kd-1}} \left(1 - v \left(1 - q^{\frac{(k-1)d-1}{kd-1}} \right) \right)}.
\end{aligned}$$

Die Differentialgleichung (3.4) transformiert sich unter den Koordinaten (r, s) zu:

$$\tilde{T}_r^{[j]}(r, s) = \left((1-k)d + (kd-1)j + kd \frac{e^{((k-1)d-1)r}}{c_2(s) + e^{((k-1)d-1)r}} \right) \tilde{T}^{[j]}.$$

Die Lösung davon ist

$$\begin{aligned} \tilde{T}^{[j]} &= c(s) \exp \left(r((1-k)d + (kd-1)j) + \frac{kd}{(k-1)d-1} \log(c_2 + e^{((k-1)d-1)r}) \right) \\ &= c(s) \exp \left(r((1-k)d + (kd-1)j) + \frac{kd}{(k-1)d-1} \right. \\ &\quad \left. \left(\log \left(\frac{c_2 + e^{((k-1)d-1)r}}{e^{((k-1)d-1)r}} \right) - \log(e^{((k-1)d-1)r}) \right) \right), \end{aligned}$$

wobei die Integrationskonstante über $t_j = \tilde{T}^{[j]}(0, s)$ bestimmt werden kann:

$$\begin{aligned} t_j &= \tilde{T}^{[j]}(0, s) = c(s)(c_2(s) + 1)^{\frac{kd}{(k-1)d-1}} = c(s)(s^{-1} - 1 + 1)^{\frac{kd}{(k-1)d-1}} \\ &= c(s)s^{-\frac{kd}{(k-1)d-1}}, \\ &\Rightarrow c(s) = t_j s^{\frac{kd}{(k-1)d-1}}. \end{aligned}$$

Nun setzen wir die ursprünglichen Koordinaten ein und erhalten:

$$\begin{aligned} T^{[j]}(z, v) &= \frac{t_j v^{\frac{kd}{(k-1)d-1}}}{q^{-\frac{kd}{k-1}} \left(1 - v \left(1 - q^{\frac{(k-1)d-1}{k-1}} \right) \right)^{\frac{kd}{(k-1)d-1}}} q^{-\frac{(1-k)d+(kd-1)j}{k-1}} v^{-\frac{kd}{(k-1)d-1}} q^{-\frac{kd}{(k-1)d-1}} \\ &= \frac{t_j}{\left(1 - v \left(1 - q^{\frac{(k-1)d-1}{k-1}} \right) \right)^{\frac{kd}{(k-1)d-1}} q^{\frac{(1-k)d+(kd-1)j}{k-1}}}. \end{aligned}$$

Wir bezeichnen mit $Y_{n,j}$ die Zufallsvariable des Knotengrades vom Knoten $j \geq 1$ in einem ansteigenden d -ären k -Baum der Größe n , wenn diese gleichverteilt sind. Das Ablesen der Koeffizienten für die Berechnung von

$$\mathbb{P}\{Y_{n,j} = m\} = \frac{t_{n,j,m}}{t_n} = \frac{(n-j)!}{t_n} [z^{n-j} v^m] T^{[j]}(z, v)$$

teilen wir wieder in mehrere Schritte auf. Für den von z und v unabhängigen Faktor gilt:

$$\frac{(n-j)! t_t}{t_n} = \frac{(n-j)! j! (kd-1)^j \binom{j-1+\frac{d}{k-1}}{j}}{n! (kd-1)^n \binom{n-1+\frac{d}{k-1}}{n}} = \frac{\binom{j-1+\frac{d}{k-1}}{j}}{\binom{n}{j} (kd-1)^{n-j} \binom{n-1+\frac{d}{k-1}}{n}}.$$

Als Nächstes werden vorerst nur die Koeffizienten von v betrachtet:

$$\begin{aligned} [v^m] &= \frac{1}{\left(1 - v \left(1 - q^{\frac{(k-1)d-1}{k-1}} \right) \right)^{\frac{kd}{(k-1)d-1}}} = \binom{-\frac{kd}{(k-1)d-1}}{m} (-1)^m \left(1 - q^{\frac{(k-1)d-1}{k-1}} \right)^m \\ &= \binom{m-1+\frac{kd}{(k-1)d-1}}{m} \sum_{l=0}^m \binom{m}{l} (-1)^l q^{l \frac{(k-1)d-1}{k-1}}. \end{aligned}$$

Das ergibt für das Berechnen des $z^{n-j}v^m$ -Koeffizienten:

$$\begin{aligned}
[z^{n-j}v^m] \frac{T^{[j]}(z, v)}{t_j} &= \binom{m-1 + \frac{kd}{(k-1)d-1}}{m} \sum_{l=0}^m \binom{m}{l} (-1)^l [z^{n-j}] q^{\frac{l((k-1)d-1) - (1-k)d - (kd-1)j}{kd-1}} \\
&= \binom{m-1 + \frac{kd}{(k-1)d-1}}{m} \sum_{l=0}^m \binom{m}{l} (-1)^l (kd-1)^{n-j} (-1)^{n-j} \\
&\quad \binom{\frac{l((k-1)d-1) - (1-k)d - (kd-1)j}{kd-1}}{n-j} \\
&= \binom{m-1 + \frac{kd}{(k-1)d-1}}{m} \sum_{l=0}^m \binom{m}{l} (-1)^l (kd-1)^{n-j} \\
&\quad \binom{n-j-1 - \frac{l((k-1)d-1) + (k-1)d - (kd-1)j}{kd-1}}{n-j},
\end{aligned}$$

was sich mit der Nebenrechnung $l((k-1)d-1) + (k-1)d + kd - 1 = (l+1)((k-1)d-1) + kd$ zu

$$\begin{aligned}
[z^{n-j}v^m] \frac{T^{[j]}(z, v)}{t_j} &= \binom{m-1 + \frac{kd}{(k-1)d-1}}{m} \sum_{l=0}^m \binom{m}{l} (-1)^l (kd-1)^{n-j} \\
&\quad \binom{n - \frac{(l+1)((k-1)d-1) + kd}{kd-1}}{n-j}
\end{aligned}$$

vereinfacht. Für die Verteilung ergibt sich:

$$\mathbb{P}\{Y_{n,j} = m\} = \frac{\binom{j-1 + \frac{d}{kd-1}}{j} \binom{m-1 + \frac{kd}{(k-1)d-1}}{m}}{\binom{n}{j} \binom{n-1 + \frac{d}{kd-1}}{n}} \sum_{l=0}^m \binom{m}{l} (-1)^l \binom{n - \frac{(l+1)((k-1)d-1) + kd}{kd-1}}{n-j}.$$

Wurzelknoten: Für den Fall eines Wurzelknotens ändern sich wieder nur die möglichen Plätze am Anfang, wenn der Knoten noch keine Kinder hat. Dies sind bei den d -ären k -Bäumen genau d Stück. Es bezeichne $t_{n,0,m}$ die Anzahl der ansteigenden d -ären k -Bäume der Größe n , bei denen der Knoten, welcher mit 0_1 markiert ist, m Kinder hat. Ansonsten mit den gleichen Überlegungen wie für einen Knoten $j \geq 1$ erhalten wir die Rekursion:

$$\begin{aligned}
t_{n,0,m} &= ((kd-1)(n-) - ((k-1)d-1)m)t_{n-1,0,m} \\
&\quad + (d + ((k-1)d-1)(m-1))t_{n-1,0,m-1}, \quad n \geq 1, m \geq 0.
\end{aligned}$$

Für den Fall $n = 0$ gilt: $t_{0,0,0} = 1$ und $t_{0,0,m} = 0$ für $m \geq 1$. In der Sprache der bivariaten

erzeugenden Funktionen ergibt die Rekursion mittels der Gleichungen

$$\begin{aligned}
\sum_{\substack{n \geq 1 \\ m \geq 0}} \frac{t_{n,0,m}}{(n-1)!} z^{n-1} v^m &= T_z^{[0]}, \\
\sum_{\substack{n \geq 1 \\ m \geq 0}} \frac{(kd-1)(n-1)t_{n-1,0,m}}{(n-1)!} z^{n-1} v^m &= (kd-1)zT_z^{[0]}, \\
\sum_{\substack{n \geq 1 \\ m \geq 0}} \frac{((k-1)d-1)mt_{n-1,0,m}}{(n-1)!} z^{n-1} v^m &= ((k-1)d-1)vT_v^{[0]}, \\
\sum_{\substack{n \geq 1 \\ m \geq 0}} \frac{dt_{n-1,0,m-1}}{(n-1)!} z^{n-1} v^m &= dvT^{[0]}, \\
\sum_{\substack{n \geq 1 \\ m \geq 0}} \frac{((k-1)d-1)(m-1)t_{n-1,0,m-1}}{(n-1)!} z^{n-1} v^m &= ((k-1)d-1)v^2T_v^{[0]},
\end{aligned}$$

eine partielle Differentialgleichung für $T^{[0]}(z, v) = \sum_{n \geq 0} \sum_{m \geq 0} t_{n,0,m} \frac{z^n}{n!} v^m$:

$$(1 - (kd-1)z)T_z^{[0]} + ((k-1)d-1)v(1-v)T_v^{[0]} - dvT^{[0]}, \quad T^{[0]}(0, v) = 1.$$

Um diese partielle Differentialgleichung zu lösen, können wir die gleiche Transformation wie bei der Gleichung (3.4) verwenden. Für die transformierte Funktion $\tilde{T}^{[0]}(r, s)$ erhalten wir die gewöhnliche Differentialgleichung

$$\tilde{T}_r^{[0]} = \frac{e^{((k-1)d-1)r}}{c_2 + e^{((k-1)d-1)r}} \tilde{T}^{[0]}.$$

Die Lösung dieser separablen Differentialgleichung ist:

$$\begin{aligned}
\tilde{T}^{[0]}(r, s) &= c(s) \exp\left(\frac{1}{(k-1)d-1} \log(c_2 + e^{((k-1)d-1)r})\right) \\
&= c(s) \left(\frac{c_2 + e^{((k-1)d-1)r}}{e^{((k-1)d-1)r}}\right)^{\frac{1}{(k-1)d-1}} e^r.
\end{aligned}$$

Die Randbedingung führt für $c(s)$ zur Gleichung

$$1 = \tilde{T}^{[0]}(0, s) = c(s)(s^{-1} - 1 + 1)^{\frac{1}{(k-1)d-1}} \Rightarrow c(s) = s^{\frac{1}{(k-1)d-1}}.$$

Wenn wir dies einsetzen und die Koordinaten zurück transformieren, erhalten wir:

$$\begin{aligned}
T^{[0]}(z, v) &= \frac{v^{\frac{1}{(k-1)d-1}}}{q^{-\frac{1}{kd-1}} \left(1 - v \left(1 - q^{\frac{(k-1)d-1}{kd-1}}\right)\right)^{\frac{1}{(k-1)d-1}}} v^{-\frac{1}{(k-1)d-1}} q^{-\frac{1}{kd-1}} \\
&= \frac{1}{\left(1 - v \left(1 - q^{\frac{(k-1)d-1}{kd-1}}\right)\right)^{\frac{1}{(k-1)d-1}}}.
\end{aligned}$$

Mit dieser Vorarbeit kann nun die Zufallsvariable $Y_{n,0}$ des ausgehenden Knotengrades vom Knoten 0_1 in einem ansteigenden d -ären k -Baum bestimmt werden:

$$\mathbb{P}\{Y_{n,0} = m\} = \frac{t_{n,j,m}}{t_n} = \frac{n!}{t_n} [z^n v^m] T^{[0]}(z, v).$$

Die Koeffizienten können mit Hilfe der Binomischen Reihe ermittelt werden:

$$\begin{aligned} [z^n v^m] &= T^{[0]}(z, v) [z^n v^m] \frac{1}{\left(1 - v \left(1 - q^{\frac{(k-1)d-1}{kd-1}}\right)\right)^{\frac{1}{(k-1)d-1}}} \\ &= [z^n] \binom{-\frac{1}{(k-1)d-1}}{m} (-1)^m \sum_{l=0}^m \binom{m}{l} (-1)^l q^{l \frac{(k-1)d-1}{kd-1}} \\ &= \binom{m-1 + \frac{1}{(k-1)d-1}}{m} \sum_{l=0}^m \binom{m}{l} (-1)^l (kd-1)^n (-1)^n \binom{l \frac{(k-1)d-1}{kd-1}}{n} \\ &= \binom{m - \frac{(k-1)d-2}{(k-1)d-1}}{m} \sum_{l=0}^m \binom{m}{l} (-1)^l (kd-1)^n \binom{n-1 - \frac{l((k-1)d-1)}{kd-1}}{n}. \end{aligned}$$

Für die Verteilung ergibt sich damit:

$$\begin{aligned} \mathbb{P}\{Y_{n,0} = m\} &= \frac{n! \binom{m - \frac{(k-1)d-2}{(k-1)d-1}}{m}}{n! (kd-1)^n \binom{n - \frac{(k-1)d-1}{kd-1}}{n}} \sum_{l=0}^m \binom{m}{l} (-1)^l (kd-1)^n \binom{n - \frac{(l+1)(k-1)d-1}{kd-1}}{n} \\ &= \frac{\binom{m - \frac{(k-1)d-2}{(k-1)d-1}}{m}}{\binom{n - \frac{(k-1)d-1}{kd-1}}{n}} \sum_{l=0}^m \binom{m}{l} (-1)^l \binom{n - \frac{(l+1)(k-1)d-1}{kd-1}}{n}. \end{aligned}$$

Beliebig hinzugefügter Knoten: Wir nehmen nun an, dass wir einen Knoten ungleich der Wurzelknoten zufällig auswählen. Für die Verteilungsfunktion der Zufallsvariable \bar{Y}_n des ausgehenden Grades eines beliebigen eingefügten Knotens gilt:¹³

$$\begin{aligned} \mathbb{P}\{\bar{Y}_n = m\} &= \frac{1}{n} \sum_{j=1}^n \mathbb{P}\{Y_{n,j} = m\} \\ &= \sum_{j=1}^n \frac{\binom{j-1 + \frac{d}{kd-1}}{j} \binom{m-1 + \frac{kd}{(k-1)d-1}}{m}}{n \binom{n}{j} \binom{n-1 + \frac{d}{kd-1}}{n}} \sum_{l=0}^m \binom{m}{l} (-1)^l \binom{n - \frac{(l+1)((k-1)d-1) + kd}{kd-1}}{n-j} \\ &= \frac{\binom{m-1 + \frac{kd}{(k-1)d-1}}{m}}{n \binom{n-1 + \frac{d}{kd-1}}{n}} \sum_{l=0}^m \binom{m}{l} (-1)^l \sum_{j=1}^n \frac{\binom{j-1 + \frac{d}{kd-1}}{j} \binom{n - \frac{(l+1)((k-1)d-1) + kd}{kd-1}}{n-j}}{\binom{n}{j}} \\ &= \frac{\binom{m-1 + \frac{kd}{(k-1)d-1}}{m}}{n \binom{n-1 + \frac{d}{kd-1}}{n}} \sum_{l=0}^m \frac{\binom{m}{l} (-1)^l d}{l((k-1)d-1) + 2kd + 1} \left(\binom{n + \frac{d}{kd-1}}{n} \right. \\ &\quad \left. + \binom{n-2 - \frac{l((k-1)d-1) - d + 1}{kd-1}}{n} \right). \end{aligned}$$

¹³siehe Fußnote 11 für den letzten Umformungsschritt.

Um die Grenzverteilung zu bestimmen, betrachten wir vorerst nur einmal folgende Limiten:

$$\lim_{n \rightarrow \infty} \frac{\binom{n + \frac{d}{kd-1}}{n}}{n \binom{n-1 + \frac{d}{kd-1}}{n}} = \lim_{n \rightarrow \infty} \frac{\left(n + \frac{d}{kd-1}\right) \left(n-1 + \frac{d}{kd-1}\right)^{n-1}}{n \left(n-1 + \frac{d}{kd-1}\right)^{n-1} \frac{d}{kd-1}} = \frac{kd-1}{d},$$

für den Zweiten kürzen wir mit $\alpha_l := \frac{l((k-1)d-1)-d-1}{kd-1}$ ab und erhalten mit Lemma 1.7:

$$\lim_{n \rightarrow \infty} \frac{\binom{n-2-\alpha_l}{n}}{n \binom{n-1 + \frac{d}{kd-1}}{n}} = \lim_{n \rightarrow \infty} \frac{c\left(-1 - \alpha_l, \frac{d}{kd-1}\right)}{n} n^{-2-\alpha_l+1-\frac{d}{kd-1}} = 0.$$

Für die Grenzverteilung \bar{Y} ($\bar{Y}_n \xrightarrow{(d)} \bar{Y}$) folgt daraus:

$$\begin{aligned} \mathbb{P}\{\bar{Y} = m\} &= \binom{m-1 + \frac{kd}{(k-1)d-1}}{m} (kd-1) \sum_{l=0}^m \frac{\binom{m}{l} (-1)^l}{l((k-1)d-1) + 2kd+1} \\ &= \binom{m-1 + \frac{kd}{(k-1)d-1}}{m} (kd-1) \sum_{l=0}^m \frac{\binom{m}{l} (-1)^l}{l + \frac{2kd+1}{(k-1)d-1}} \\ &\stackrel{\text{L. 1.9}}{=} \frac{\binom{m-1 + \frac{kd}{(k-1)d-1}}{m} (kd-1)}{\left(m + \frac{2kd+1}{(k-1)d-1}\right) \binom{m-1 + \frac{2kd+1}{(k-1)d-1}}{m}}. \end{aligned}$$

Die Grenzverteilung \bar{Y} verhält sich asymptotisch folgendermaßen:

$$\begin{aligned} \mathbb{P}\{\bar{Y} = m\} &\stackrel{\text{L. 1.7}}{\sim} \frac{(kd-1)\Gamma\left(\frac{2kd+1}{(k-1)d-1}\right) m^{-1 + \frac{kd}{(k-1)d-1}}}{m\Gamma\left(\frac{kd}{(k-1)d-1}\right) m^{-1 + \frac{2kd+1}{(k-1)d-1}}} = \frac{(kd-1)\Gamma\left(\frac{2kd+1}{(k-1)d-1}\right)}{\Gamma\left(\frac{kd}{(k-1)d-1}\right)} m^{-1 - \frac{kd+1}{(k-1)d-1}} \\ &= \frac{(kd-1)\Gamma\left(\frac{2kd+1}{(k-1)d-1}\right)}{\Gamma\left(\frac{kd}{(k-1)d-1}\right)} m^{-2 - \frac{d+2}{(k-1)d-1}}. \end{aligned}$$

Die Verteilungsfunktion erfüllt ein Potenzgesetz mit Exponent $2 + \frac{d+2}{(k-1)d-1}$. Für reguläre Werte von k und d ($(k-1)d-1 > 0$) stimmt die Ungleichung $\frac{d+1}{(k-1)d-1} > \frac{1}{k}$. Dies bedeutet, dass der Exponent der Grenzverteilung der ansteigenden d -ären k -Bäume größer ist als jener der geordneten ansteigenden k -Bäume. Damit treten Knoten mit hohem ausgehenden Knotengrad in einem geordneten ansteigenden k -Baum mit höherer Wahrscheinlichkeit auf als in einem ansteigenden d -ären k -Baum. Dies deckt sich mit den randomisierten Modellen zur Konstruktion von diesen zwei Arten von k -Bäumen. Bei den geordneten ansteigenden k -Bäumen erhöht sich mit jedem Kind einer Clique die Wahrscheinlichkeit den Knoten der Clique hinzuzufügen, während dies bei den d -ären k -Bäumen genau umgekehrt ist. So gibt es bei den geordneten ansteigenden k -Bäumen mehr Ausreißer mit hohem Knotengrad.

3.2.4 CU-Modell

Hinzugefügter Knoten: Für die vorigen drei randomisierten Modelle zum Aufbau eines k -Baumes gab es immer ein entsprechendes k -Baum-Modell, von welchem jeder Baum

mit der gleichen Wahrscheinlichkeit erzeugt wurde. Das heißt, die entsprechenden Modelle waren immer uniform verteilt und wir konnten uns bei der Rekursion auf die Mächtigkeit der Menge der k -Bäume mit der gewünschten Eigenschaft einschränken. Aus diesen konnten wir einfach die Wahrscheinlichkeitsverteilung berechnen. Für das CU-Modell zur Erzeugung eines k -Baumes wählen wir den gleichen Weg, obwohl wir im Hintergrund kein explizites Modell für die k -Bäume haben. Wir werden im Folgenden von den verschiedenen Möglichkeiten sprechen. Damit möchten wir zählen, wie viele verschiedene Möglichkeiten es bei der Auswahl der Cliques gibt. Beim Einfügen eines Knotens gibt es $|\{C : C \text{ ist eine } (k+1)\text{-Clique}\}|$ Möglichkeiten, um eine $(k+1)$ -Clique auszuwählen und weitere $(k+1)$ Möglichkeiten, um die k -Clique zu bestimmen, bei welcher der Knoten hinzugefügt wird. Nach Lemma 1.1 gibt es in einem k -Baum mit n Knoten, die jeweils verschieden zu den Wurzelknoten sind, $n(k+1)$ -Cliques. Das heißt, für den n -ten Knoten ($n \geq 2$) gibt es im CU-Modell $(n-1)(k+1)$ Möglichkeiten, diesen einzufügen, respektive gibt es $(n-1)!(k+1)^{n-1}$ Möglichkeiten, einen k -Baum der Größe n zu erstellen. Um nun eine Rekursion für den ausgehenden Knotengrad aufzustellen, müssen wir nur noch wissen, wie viele davon den Knotengrad erhöhen. Dazu betrachten wir den Knoten j . Wenn dieser hinzugefügt wird beziehungsweise der ausgehende Knotengrad gleich 0 ist, ist der Knoten j Teil einer einzigen $(k+1)$ -Clique. Damit nun ein Kind bei j angefügt wird, muss genau diese Clique ausgewählt werden und im zweiten Schritt eine der k -Cliques, welche j umfasst. Es gibt somit k Möglichkeiten für das Anfügen eines neuen Knotens, damit sich der Grad von j erhöht. Jedes Kind von j erzeugt eine weitere $(k+1)$ -Clique, welche j als Teilobjekt hat. Dies bedeutet, dass es $(m+1)k$ Möglichkeiten gibt, bei denen sich die Kinderanzahl erhöht, gibt, wenn der Knoten bereits m Kinder hat. Damit können wir folgende Rekursion aufstellen:

$$t_{n,j,m} = ((n-1)(k+1) - (m+1)k)t_{n-1,j,m} + mkt_{n-1,j,m-1}, \quad n \geq j+1, m \geq 0,$$

wobei $t_{n,j,m}$ die Anzahl der Möglichkeiten bezeichnet mit dem CU-Modell einen k -Baum, bei dem der Knoten m Kinder hat, zu erzeugen. Die Startwerte für die Rekursion sind $t_{j,j,0} = (j-1)!(k+1)^{j-1} =: t_j$ und $t_{j,j,m} = 0$ für $m \geq 1$. Für die bivariate erzeugende Funktion $T^{[j]}(z, v) = \sum_{n \geq j} \sum_{m \geq 0} t_{n,j,m} \frac{z^{n-j}}{(n-j)!} v^m$ können wir diese auf eine partielle Differentialgleichung führen:

$$\begin{aligned} \sum_{\substack{n \geq j+1 \\ m \geq 0}} \frac{t_{n,j,m}}{(n-j-1)!} z^{n-j-1} v^m &= T_z^{[j]}, \\ \sum_{\substack{n \geq j+1 \\ m \geq 0}} \frac{(k+1)(n-1)t_{n-1,j,m}}{(n-j-1)!} z^{n-j-1} v^m &= (k+1)(zT_z^{[j]} + jT^{[j]}), \\ \sum_{\substack{n \geq j+1 \\ m \geq 0}} \frac{k(m+1)t_{n-1,j,m}}{(n-j-1)!} z^{n-j-1} v^m &= (kvT_v^{[j]} + kT^{[j]}), \\ \sum_{\substack{n \geq j+1 \\ m \geq 0}} \frac{mkt_{n-1,j,m-1}}{(n-j-1)!} z^{n-j-1} v^m &= kv^2T_v^{[j]} + kvT^{[j]}, \end{aligned}$$

womit sich die Differentialgleichung

$$(1 - (k+1)z)T_z^{[j]} + kv(1-v)T_v^{[j]} - ((k+1)j - k + kv)T^{[j]} = 0$$

mit der Randbedingung $T^{[j]}(0, v) = t_j$ ergibt. Die Faktoren bei den Ableitungen sind genau die gleichen wie bei den geordneten ansteigenden k -Bäumen. Deshalb kann die Gleichung auch mit derselben Koordinatentransformation gelöst werden. Zur Erinnerung:

$$z = \frac{1 - e^{-(k+1)z}}{k+1}, \quad r = -\frac{1}{k+1} \log(q),$$

$$v = \frac{e^{kr}}{s^{-1} - 1 + e^{kr}}, \quad s = \frac{v}{q^{-\frac{k}{k+1}} \left(1 - v \left(1 - q^{\frac{k}{k+1}}\right)\right)},$$

mit $q(z) := 1 - (k+1)z$. So erhalten wir in den Koordinaten (r, s) die gewöhnliche Differentialgleichung:

$$\tilde{T}_r^{[j]}(r, s) = \left((k+1)j - k + k \frac{e^{kr}}{s^{-1} - 1 + e^{kr}} \right) \tilde{T}^{[j]}(r, s),$$

mit der Lösung

$$\tilde{T}^{[j]}(r, s) = c(s) \exp \left(r((k+1)j - k) + \log \left(\frac{s^{-1} - 1 + e^{kr}}{e^{kr}} \right) + kr \right).$$

Mit der Randbedingung erhalten wir für die Integrationskonstante:

$$t_j = \tilde{T}^{[j]}(0, s) = c(s)s^{-1} \quad \Rightarrow \quad c(s) = \frac{t_j}{s}.$$

In den ursprünglichen Koordinaten ergibt sich die Lösung

$$T^{[j]}(z, v) = \frac{t_j v}{q^{-\frac{k}{k+1}} \left(1 - v \left(1 - q^{\frac{k}{k+1}}\right)\right)} q^{-j + \frac{k}{k+1}} v^{-1} q^{-\frac{k}{k+1}}$$

$$= \frac{t_j}{\left(1 - v \left(1 - q^{\frac{k}{k+1}}\right)\right)} q^{j - \frac{k}{k+1}}.$$

Um nun die Verteilung $Y_{n,j}$ der Kinder des Knotens j zu ermitteln, werden die Koeffizienten von $T^{[j]}(z, v)$ benötigt:

$$[z^{n-j} v^m] \frac{T^{[j]}(z, v)}{t_j} = [z^{n-j}] \left(1 - q^{\frac{k}{k+1}}\right)^m q^{-j + \frac{k}{k+1}} = \sum_{l=0}^m \binom{m}{l} (-1)^l [z^{n-j}] q^{-j + \frac{(l+1)k}{k+1}}$$

$$= \sum_{l=0}^m \binom{m}{l} (-1)^l (k+1)^{n-j} (-1)^{n-j} \binom{\frac{(l+1)k}{k+1} - j}{n-j}$$

$$= \sum_{l=0}^m \binom{m}{l} (-1)^l (k+1)^{n-j} \binom{n - \frac{(l+2)k+1}{k+1}}{n-j}.$$

Die Wahrscheinlichkeit einer bestimmten Anzahl an Kindern lässt sich dadurch berechnen

mit:

$$\begin{aligned}
\mathbb{P}\{Y_{n,j} = m\} &= \frac{t_{n,j,m}}{t_n} = \frac{(n-j)!}{t_n} [z^{n-j} v^m] T^{[j]}(z, v) \\
&= \frac{(j-1)!(k+1)^{j-1}(n-j)!}{(n-1)!(k+1)^{n-j}} \sum_{l=0}^m \binom{m}{l} (-1)^l (k+1)^{n-j} \binom{n - \frac{(l+2)k+1}{k+1}}{n-j} \\
&= \frac{1}{\binom{n-1}{j-1}} \sum_{l=0}^m \binom{m}{l} (-1)^l \binom{n - \frac{(l+2)k+1}{k+1}}{n-j}.
\end{aligned}$$

Wurzelknoten: Durch die Besonderheit, dass beim CU-Modell ausgehend von einer $(k+1)$ -Clique weitere Knoten hinzugefügt werden, ist der Fall eines Wurzelknotens schon berechnet. Denn der Knoten mit der Markierung 1 ist auch bei der „Startclique“ der Größe $k+1$ dabei. Die weitere Auswahl nimmt keine Rücksicht auf die Markierung, sodass alle diese $k+1$ Knoten gleichberechtigt sind beziehungsweise gilt:

$$\mathbb{P}\{Y_{n,0} = m\} = \mathbb{P}\{Y_{n,1} = m\} = \sum_{l=0}^m \binom{m}{l} (-1)^l \binom{n - \frac{(l+2)k+1}{k+1}}{n-1}.$$

Beliebig hinzugefügter Knoten: Beim CU-Modell wird der Knoten mit der Markierung 1 auch als hinzugefügter Knoten bezeichnet, obwohl er im Algorithmus ein Element der Startclique der Größe $k+1$ ist. So hat ein mit dem CU-Modell erzeugter k -Baum der Größe n genau n hinzugefügte Knoten. Dies gilt auch für die anderen Modelle. Für die Verteilung \bar{Y}_n eines beliebigen hinzugefügten Knotens gilt:¹⁴

$$\begin{aligned}
\mathbb{P}\{\bar{Y}_n = m\} &= \frac{1}{n} \sum_{j=1}^n \mathbb{P}\{Y_{n,j} = m\} = \frac{1}{n} \sum_{j=1}^n \frac{1}{\binom{n-1}{j-1}} \sum_{l=0}^m \binom{m}{l} (-1)^l \binom{n - \frac{(l+2)k+1}{k+1}}{n-j} \\
&= \frac{1}{n} \sum_{l=0}^m \binom{m}{l} (-1)^l \sum_{j=1}^n \frac{\binom{n - \frac{(l+2)k+1}{k+1}}{n-j}}{\binom{n-1}{j-1}} \\
&= \frac{1}{n} \sum_{l=0}^m \binom{m}{l} \frac{(-1)^l}{(l+2)k+1} \left(n(k+1) + (k(l+2) - k) \binom{n - \frac{(l+2)k+1}{k+1}}{n-1} \right).
\end{aligned}$$

Um die Grenzverteilung \bar{Y} zu bestimmen, betrachten wir die Konvergenz der beiden von n abhängigen Summanden einzeln:

$$\begin{aligned}
\lim_{n \rightarrow \infty} \frac{n(k+1)}{n} &= k+1, \\
\lim_{n \rightarrow \infty} \frac{1}{n} \binom{n - \frac{(l+2)k+1}{k+1}}{n-1} &\stackrel{\text{L. 1.8}}{=} c \left(1 - \frac{(l+2)k+1}{k+1} \right) \lim_{n \rightarrow \infty} \frac{n^{1 - \frac{(l+2)k+1}{k+1}}}{n} = 0.
\end{aligned}$$

Die Verteilungen \bar{Y}_n konvergieren demnach in Verteilung gegen eine Zufallsvariable \bar{Y} mit:

$$\mathbb{P}\{\bar{Y} = m\} = (k+1) \sum_{l=0}^m \binom{m}{l} \frac{(-1)^l}{k(l+1)} = \frac{k+1}{k} \sum_{l=0}^m \binom{m}{l} \frac{(-1)^l}{l+1}.$$

¹⁴siehe Fußnote 11 für den letzten Umformungsschritt.

Diese Summe ist auch schon bei den geordneten ansteigenden k -Bäumen aufgetreten und so wissen wir:

$$\mathbb{P}\{\bar{Y} = m\} = \frac{k+1}{k(m+1)\binom{m+2+\frac{1}{k}}{m+1}} \sim \frac{k+1}{k} \Gamma\left(2 + \frac{1}{k}\right) m^{-2-\frac{1}{k}}.$$

Die Grenzverteilungsfunktionen der Kinder stimmen für das CU-Modell und das der geordneten ansteigenden k -Bäume überein. Bei beiden Modellen werden Cliques mit einer hohen Kinderanzahl bevorzugt, in dem Sinne, dass diese mit einer höheren Wahrscheinlichkeit ein weiteres Kind bekommen. Die genauen Wahrscheinlichkeiten für weitere Kinder unterscheidet sich. Jedoch schlägt sich die Tatsache der gleichen Bevorzugung in der Grenzverteilung nieder und liefert, dass beide Modelle das Potenzgesetz mit dem gleichen Koeffizienten $2 + \frac{1}{k}$ erfüllen.¹⁵

3.3 Lokaler Clusterkoeffizient

Für die im vorigen Abschnitt untersuchten ansteigenden Modelle von k -Bäumen wird nun der lokale Clusterkoeffizient eines Knotens berechnet. Dieser gilt als wichtiger Parameter bei realen und skalenfreien Netzwerken. Er berechnet sich aus der Umgebung eines Knotens und ist direkt proportional zu den Verbindungen zwischen zwei Nachbarn (vgl. [19], Abschnitt 3.1 und [6], Abschnitt 1).

Definition 3.1 (vgl. [19], Abschnitt 3.1 und [6], Abschnitt 1). Der lokale Clusterkoeffizient $C(v)$ eines Knoten $v \in G$ ist definiert durch:

$$C(v) = \begin{cases} \frac{|\{e \in E(G) : e = (x, y) \text{ mit } x, y \in \Gamma(v)\}|}{\binom{d(v)}{2}}, & \text{falls } d(v) \geq 2, \\ 0, & \text{falls } d(v) = 0 \text{ oder } d(v) = 1. \end{cases}$$

Der lokale Clusterkoeffizient misst, wie stark sich die Verbindungen in der Nähe eines Knotens häufen. Er liegt definitionsgemäß zwischen 0 und 1.

Beim rekursiven Aufbau eines k -Baumes wird beim Hinzufügen eines Nachbarn eines Knotens, indem er mit einer k -Clique verbunden wird, auch festgelegt, wie dieser Nachbar mit den übrigen Nachbarn verbunden ist. Deshalb kann der lokale Clusterkoeffizient für k -Bäume in Abhängigkeit vom Knotengrad angegeben werden.

Lemma 3.2 (vgl. [19], Lemma 1). *Der lokale Clusterkoeffizient ist für einen Knoten v eines k -Baumes nur vom Grad abhängig. Es gilt für $d(v) \geq k \geq 2$:*

$$C(v) = \frac{2(k-1)}{d(v)} - \frac{(k-1)(k-2)}{d(v)(d(v)-1)}.$$

Beweis: Die Bedingung $d(v) \geq k$ unterbindet, dass v ein Wurzelknoten ist. Denn jeder Knoten, welcher einem k -Baum hinzugefügt wird, wird gleich mit einer k -Clique respektive k Knoten verbunden. Es soll $M(v)$ die Kanten zwischen Nachbarn von v zählen, das heißt $M(v) = |\{e \in E(G) : e = (x, y) \text{ mit } x, y \in \Gamma(v)\}|$. Wenn $d(v) = k$, hat der Knoten v keine Kinder und ist Teil einer $(k+1)$ -Clique und so ist $M(v) = \binom{k}{2}$. Jedes Kind von

¹⁵Siehe [6] für eine weitere Berechnungsweise des Koeffizienten für das Potenzgesetz.

v wird wie jeder Knoten in einem k -Baum mit einer k -Clique, wobei v ein Teil davon ist, verbunden. Dies bedeutet, dass sich $M(v)$ für jedes Kind um $k - 1$ erhöht und wir erhalten $M(v) = \binom{k}{2} + (k - 1)(d(v) - k)$. Für den lokalen Clusterkoeffizienten ergibt das:

$$\begin{aligned} C(v) &= \frac{\binom{k}{2}}{\binom{d(v)}{2}} + \frac{(k - 1)(d(v) - k)}{\binom{d(v)}{2}} = \frac{k(k - 1)}{d(v)(d(v) - 1)} + \frac{2(k - 1)(d(v) - 1 + 1 - k)}{d(v)(d(v) - 1)} \\ &= \frac{2(k - 1)(d(v) - 1)}{d(v)(d(v) - 1)} + \frac{(k - 1)(k + 2 - 2k)}{d(v)(d(v) - 1)} = \frac{2(k - 1)}{d(v)} - \frac{(k - 1)(k - 2)}{d(v)(d(v) - 1)} \end{aligned}$$

□

Daher ist der lokale Clusterkoeffizient in einem k -Baum nur vom Knotengrad abhängig, welchen wir in Form des ausgehenden Knotengrades schon ausgiebig studiert haben. Insbesondere ist auch die Verteilung des lokalen Clusterkoeffizienten aus den vorigen Ergebnissen berechenbar. Dazu sei \tilde{Y}_n die Zufallsvariable des Knotengrades eines beliebig ausgewählten Knotens. Aufgrund der Gleichverteilung beim Auswählen gilt unter der Berücksichtigung der Notation vom letzten Abschnitt:

$$\tilde{Y}_{n+k} = \frac{k}{n+k} (Y_{n,0} + k - 1) + \frac{n}{n+k} \left(\sum_{j=1}^n Y_{n,j} + k \right),$$

sowie für den Clusterkoeffizienten eines zufällig ausgewählten Knotens:

$$\tilde{C}_n = \frac{2(k - 1)}{\tilde{Y}_n} - \frac{(k - 1)(k - 2)}{\tilde{Y}_n(\tilde{Y}_n - 1)}.$$

Für die vier betrachteten Modelle konvergiert für jedes $m \geq 0$ die Wahrscheinlichkeit $\mathbb{P}\{Y_{n,0} = m\}$ gegen 0 und somit $\tilde{Y}_n \xrightarrow{(d)} \bar{Y} + k$. Da $f(m) = \frac{2(k-1)}{m} - \frac{(k-1)(k-2)}{m(m-1)}$ beschränkt ist für $m \geq k$, gilt (vgl. [19], Abschnitt 4.2):

$$\lim_{n \rightarrow \infty} \mathbb{E}(\tilde{C}_n) = \mathbb{E}(f(\bar{Y} + k)) = \sum_{m \geq k} \mathbb{P}\{\bar{Y} + k = m\} \left(\frac{2(k - 1)}{m} - \frac{(k - 1)(k - 2)}{m(m - 1)} \right).$$

Explizit ergibt das für die einzelnen Modelle:

- Geordnete ansteigende k -Bäume und CU-Modell:

$$\lim_{n \rightarrow \infty} \mathbb{E}(\tilde{C}_n) = \sum_{m \geq 0} \frac{(k + 1)(k - 1)}{k(m + 1)(m + k) \binom{m+2+\frac{1}{k}}{m+1}} \left(2 - \frac{k - 2}{m + k + 1} \right).$$

- Ansteigende k -Bäume:

$$\lim_{n \rightarrow \infty} \mathbb{E}(\tilde{C}_n) = \sum_{m \geq 0} \frac{k \binom{m+\frac{1}{k-1}}{m}}{(m + 1)(m + k) \binom{m+\frac{2k}{k-1}}{m+1}} \left(2 - \frac{k - 2}{m + k + 1} \right).$$

- Ansteigende d -äre k -Bäume:

$$\lim_{n \rightarrow \infty} \mathbb{E}(\tilde{C}_n) = \sum_{m \geq 0} \frac{(k - 1)(kd - 1) \binom{m-1+\frac{kd}{(k-1)d-1}}{m}}{((m + k) \binom{m+\frac{2kd+1}{(k-1)d-1}}{m+1}) \binom{m+1+\frac{2kd-1}{(k-1)d-1}}{m}} \left(2 - \frac{k - 2}{m + k + 1} \right).$$

3.4 Nachfahren

Wie bei „normalen“ Bäumen bilden die Nachfahren die transitive Erweiterung der Kind-Beziehung. Wie sich herausstellen wird, ist es vorteilhaft, einen Knoten als Nachfahren von sich selbst zu bezeichnen. Dies führt zur Definition eines Nachfahren: Ein Nachfahre eines Knotens ist der Knoten selbst und jedes Kind eines Nachfahrens (vgl. [18], Abschnitt 3). Mit dieser rekursiven Definition können wir wie beim Bestimmen der Kinder vorgehen. Wir werden eine Rekursion aufstellen, welche zu einer Differentialgleichung sehr ähnlicher Bauart wie bei Abschnitt 3.2 führt. Diese kann auf dieselbe Weise gelöst werden, weshalb die Rechnungen nicht in aller Ausführlichkeit angegeben sind.

Bei den Kindern war eine Fallunterscheidung bezüglich der Wurzelknoten notwendig. Jeder hinzugefügte Knoten j ist natürlich ein Nachfahre eines Wurzelknoten. Er ist sogar Nachfahre jedes Wurzelknoten. Denn das Kind eines Wurzelknoten, welches auf einem Weg liegt, der durch die Nachfahren-Relation von der Wurzel zum Knoten j besucht wird, hat als Elternclique entweder die Wurzelclique oder eine Menge von Knoten, die einen Knoten umfasst, der ein Kind der Wurzelclique ist.

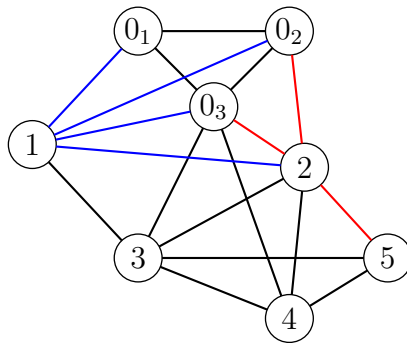


Abbildung 19: Ein 3-Baum, bei dem man mit dem „kürzesten Weg“ vom Knoten 5 zur Wurzel noch nicht einsieht, dass er Nachfahre von jedem Wurzelknoten ist.

3.4.1 Geordnete ansteigende k -Bäume

Hinzugefügter Knoten: Wie wir schon wissen, gibt es für das Hinzufügen des n -ten Knotens $1 + (k + 1)(n - 1)$ Möglichkeiten. Für die Rekursion muss bestimmt werden, wie viele Plätze davon einem weiteren Nachfolger für den Knoten $j \geq 1$ entsprechen. Falls der Knoten j nur einen Nachfolger, das heißt sich selbst, hat, ist er ein Teil von k Cliques, wobei jede davon noch kein Kind hat. Dies liefert k Möglichkeiten für Nachfolger. Jeder weitere Nachfolger erzeugt einen Platz bei der Clique, bei welcher er hinzugefügt wird. Zudem entstehen k neue k -Cliques, wobei jedes Kind dieser Cliques ein Nachfolger von j wäre. Es entstehen insgesamt pro Nachfolger $k + 1$ neue Plätze für weitere Nachfolger. Für die Anzahl $t_{n,j,m}$ der geordneten ansteigenden k -Bäume der Größe n , wobei der Knoten j genau m Nachfolger besitzt, erhalten wir die Rekursion (vgl. [18], Abschnitt 4.1):

$$t_{n,j,m} = (1 - k + (k - 1)(n - m))t_{n-1,j,m} + (k + (k + 1)(m - 2))t_{n-1,j,m-1},$$

für $n \geq j + 1, m \geq 1$ und den Bedingungen $t_{j,j,1} = 1$ und $t_{j,j,m} = 0$ für $j \geq 1$ und $1 \neq m \in \mathbb{N}$. Wie schon bei den Kindern leiten wir daraus eine Differentialgleichung

für die bivariate erzeugende Funktion $T^{[j]}(z, v) = \sum_{n \geq j} \sum_{m \geq 1} t_{n,j,m} \frac{z^{n-j}}{(n-j)!} v^m$ her. Diese Transformation der Gleichung geschieht auf die gleiche Weise und da bezüglich n und m die Terme abgesehen von den Vorfaktoren gleich sind, verweisen wir auf die Berechnung des Knotengrades und geben hier nur die partielle Differentialgleichung an:

$$(1 - (k+1)z)T_z^{[j]} + (k+1)v(1-v)T_v^{[j]} - (2 + (k+1)j - v)T^{[j]} = 0, \quad T^{[j]}(0, v) = vt_j,$$

wobei t_j die Gesamtanzahl der geordneten ansteigenden k -Bäume der Größe j bezeichnet. Diese Gleichung lösen wir auf die übliche Weise. Dadurch, dass dieses Mal bei der Ableitung nach z und v jeweils der Faktor $(k+1)$ steht, vereinfacht dies das Ergebnis. Die Koordinatentransformationen sind analog zu den Berechnungen beim Knotengrad durchzuführen, womit wir erhalten:

$$\begin{aligned} z &= \frac{1 - e^{-(k+1)z}}{k+1}, & r &= -\frac{1}{k+1} \log(q), \\ v &= \frac{e^{(k+1)r}}{s^{-1} - 1 + e^{(k+1)r}}, & s &= \frac{v}{q^{-1}(1 - v(1 - q))}, \end{aligned}$$

wobei $q := 1 - (k+1)z$. Die transformierte Lösung $\tilde{T}^{[j]}(r, s) = T^{[j]}(z(r, s), v(r, s))$ ergibt sich auch mittels analoger Rechnung:

$$\tilde{T}^{[j]} = c(s) \exp\left(r(2 + (k+1)j) - \frac{1}{k+1} \log(v^{-1}(r, s)) - r\right),$$

wobei die Randbedingung andere Gestalt hat:

$$\begin{aligned} v(0, s)t_j &= T^{[j]}(0, v(0, s)) = \tilde{T}^{[j]}(0, s) = c(s) (s^{-1} - 1 + 1)^{-\frac{1}{k+1}}, \\ \Rightarrow c(s) &= t_j s^{\frac{k}{k+1}}. \end{aligned}$$

Wenn wir die Lösung zurück transformieren, ergibt sich:

$$\begin{aligned} T^{[j]}(z, v) &= t_j \left(\frac{v}{q^{-1}(1 - v(1 - q))} \right)^{\frac{k}{k+1}} q^{-\frac{2+(k+1)j}{k+1}} v^{\frac{1}{k+1}} q^{\frac{1}{k+1}} \\ &= \frac{t_j v}{(1 - (k+1)vz)^{\frac{k}{k+1}} (1 - (k+1)z)^{j-1+\frac{2}{k+1}}}. \end{aligned}$$

Für die Zufallsvariable $D_{n,j}$ der Nachfolger des Knotens j für k -Bäume der Größe n möchten wir die Verteilungsfunktion angeben. Diese berechnet sich aus den Koeffizienten von $T^{[j]}$:

$$\mathbb{P}\{D_{n,j} = m\} = \frac{t_{n,j,m}}{t_j} = \frac{(n-j)! t_j}{t_n} [z^{n-j} v^m] \frac{v}{(1 - (k+1)vz)^{\frac{k}{k+1}} (1 - (k+1)z)^{j-1+\frac{2}{k+1}}}.$$

Beim Ablesen der Koeffizienten ergibt sich:

$$\begin{aligned} [z^{n-j} v^{m-1}] \frac{T^{[j]}}{t_j v} &= \binom{-\frac{k}{k+1}}{m-1} (-1)^{m-1} (k+1)^{m-1} [z^{n-j-m+1}] \frac{1}{q^{j-1+\frac{2}{k+1}}} \\ &= \binom{m-2+\frac{k}{k+1}}{m-1} (k+1)^{m-1} \binom{-(j-1+\frac{2}{k+1})}{n-j-m+1} ((-1)(k+1))^{n-j-m+1} \\ &= (k+1)^{n-j} \binom{m-1-\frac{1}{k-1}}{m-1} \binom{n-m-1+\frac{2}{k+1}}{n-j-m+1}. \end{aligned}$$

Der restliche Teil kann vereinfacht werden zu:

$$\begin{aligned} \frac{(n-j)!t_j}{t_n} &= \frac{(n-j)!j!(k+1)^j \binom{j-\frac{k}{k+1}}{j}}{n!(k+1)^n \binom{n-\frac{k}{k+1}}{n-j}} = \frac{n-j! \left(j - \frac{k}{k+1}\right)^j}{(k+1)^{n-j} \left(n - \frac{k}{k+1}\right)^n} \\ &= \frac{1}{(k+1)^{n-j} \binom{n-\frac{k}{k+1}}{n-j}}. \end{aligned}$$

Insgesamt erhalten wir für die diskrete Verteilungsfunktion (vgl. [18], Abschnitt 4.1):

$$\mathbb{P}\{D_{n,j} = m\} = \frac{\binom{m-1-\frac{1}{k+1}}{m-1} \binom{n-m-1+\frac{2}{k+1}}{n-m-j+1}}{\binom{n-1+\frac{1}{k+1}}{n-j}}.$$

Beliebig hinzugefügter Knoten: Die Zufallsvariable \bar{D}_n eines zufälligen eingefügten Knotens kann nun über das vorherige Ergebnis berechnet werden:

$$\mathbb{P}\{\bar{D}_n = m\} = \frac{1}{n} \sum_{j=1}^n \mathbb{P}\{D_{n,j} = m\} = \frac{\binom{m-1-\frac{1}{k+1}}{m-1}}{n} \sum_{j=1}^n \frac{\binom{n-m-1+\frac{2}{k+1}}{n-m-j+1}}{\binom{n-1+\frac{1}{k+1}}{n-j}},$$

wobei wir die Summe mit Lemma 1.11 vereinfachen zu:

$$\frac{\binom{m-1-\frac{1}{k+1}}{m-1}}{n} \left(\frac{n + \frac{1}{k+1}}{m \binom{m+1-\frac{1}{k+1}}{m}} - \frac{1}{\binom{n-1+\frac{1}{k+1}}{n}} \sum_{l=0}^{m-1} \binom{m-1}{l} \frac{(-1)^l \frac{1}{k+1}}{l+2-\frac{1}{k+1}} \binom{n-2-l+\frac{2}{k+1}}{n} \right).$$

Der Bruch des ersten Summanden kann noch gekürzt werden:

$$\frac{\binom{m-1-\frac{1}{k+1}}{m-1}}{n} \frac{n + \frac{1}{k+1}}{m \binom{m+1-\frac{1}{k+1}}{m}} = \frac{\left(m-1-\frac{1}{k+1}\right)^{m-1} \left(n + \frac{1}{k+1}\right)}{n \left(m+1-\frac{1}{k+1}\right)^m} = \frac{\left(1-\frac{1}{k+1}\right) \left(n + \frac{1}{k+1}\right)}{n \left(m + \frac{k}{k+1}\right) \left(m - \frac{1}{k+1}\right)},$$

womit wir insgesamt für die Verteilung folgende Darstellung finden:

$$\begin{aligned} \mathbb{P}\{\bar{D}_n = m\} &= \frac{k \left(n + \frac{1}{k+1}\right)}{(k+1)n \left(m + \frac{k}{k+1}\right) \left(m - \frac{1}{k+1}\right)} \\ &\quad - \frac{\binom{m-1-\frac{1}{k+1}}{m-1}}{\binom{n-1+\frac{1}{k+1}}{n}} \sum_{l=0}^{m-1} \frac{\binom{m-1}{l} (-1)^l}{(k+1)(l+1) + k} \binom{n-2-l+\frac{2}{k+1}}{n}. \end{aligned}$$

Um die Grenzverteilung \bar{D} für $n \rightarrow \infty$ von \bar{D}_n zu bestimmen, betrachten wir die von n abhängigen Faktoren der beiden Summanden einzeln:

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{k \left(n + \frac{1}{k+1}\right)}{(k+1)n \left(m + \frac{k}{k+1}\right) \left(m - \frac{1}{k+1}\right)} &= \frac{k}{(k+1) \left(m + \frac{k}{k+1}\right) \left(m - \frac{1}{k+1}\right)}, \\ \lim_{n \rightarrow \infty} \frac{\binom{n-2-l+\frac{2}{k+1}}{n}}{\binom{n-1+\frac{1}{k+1}}{n}} &\stackrel{\text{L. 1.8}}{=} c\left(-l - \frac{k}{k+1}, -\frac{k}{k+1}\right) \lim_{n \rightarrow \infty} \frac{n^{-2-l+\frac{2}{k+1}}}{n^{-1+\frac{1}{k+1}}} = 0. \end{aligned}$$

Somit gilt für die Grenzverteilung (vgl. [18], Abschnitt 4.3):

$$\mathbb{P}\{\bar{D} = m\} = \frac{k}{(k+1) \left(m + \frac{k}{k+1}\right) \left(m - \frac{1}{k+1}\right)}.$$

3.4.2 Ansteigende k -Bäume

Hinzugefügter Knoten: Bei den ansteigenden k -Bäumen bezeichne $t_{n,j,m}$ die Bäume mit n Knoten, bei denen der Knoten j genau m Nachfahren besitzt. Jeder zusätzliche Nachfahre eines Knotens erzeugt k neue Cliques, welchen auch k neuen Plätzen für weitere Nachfahren entsprechen. So ergibt sich die Rekursion (vgl. [18], Abschnitt 4.1):

$$\begin{aligned} t_{n,j,m} &= (1 + k(n-1) - k - k(m-1))t_{n-1,j,m} + (k + (m-2)k)t_{n-1,j,m-1} \\ &= (k(n-1-m) + 1)t_{n-1,j,m} + (m-1)kt_{n-1,j,m-1}, \quad n > j \geq 1, m \geq 1. \end{aligned}$$

Die Anfangsbedingungen für die Rekursion sind $t_{j,j,1} = 1$ und $t_{j,j,m} = 0$ für $j \geq 1$ und $1 \neq m \in \mathbb{N}$. Für die erzeugende Funktion $T^{[j]}(z, v) = \sum_{n \geq j} \sum_{m \geq 1} t_{n,j,m} \frac{z^{n-j}}{(n-j)!} v^m$ ergibt sich eine partielle Differenzialgleichung. Die ist im Vergleich zu den geordneten ansteigenden k -Bäume einfacher aufgebaut, da nur konstante Faktoren beim Term von $T^{[j]}$ vorkommen:

$$(1 - kz)T_z^{[j]} + kv(1 - v)T_v^{[j]} - (kj + 1)T^{[j]} = 0, \quad T^{[j]}(0, v).$$

Mit den üblichen Transformationen ($q := 1 - kz$)

$$\begin{aligned} z &= \frac{1 - e^{-(kd-1)z}}{kd - 1}, & r &= -\frac{1}{kd - 1} \log(q), \\ v &= \frac{e^{(kd-1)r}}{s^{-1} - 1 + e^{(kd-1)r}}, & s &= \frac{v}{q^{-1}(1 - v(1 - q))}, \end{aligned}$$

erhält man die gewöhnliche Differenzialgleichung für $\tilde{T}^{[j]}$:

$$\frac{\tilde{T}_r^{[j]}}{\tilde{T}^{[j]}} = kj + 1,$$

mit der Lösung

$$\tilde{T}^{[j]}(r, s) = c(s) \exp((kj + 1)r).$$

Die Randbedingung transformiert sich zu $c(s) = st_j$, wobei t_j die Mächtigkeit der ansteigenden k -Bäumen mit j Knoten ist. Wir erhalten folgende Lösung in den ursprünglichen Variablen:

$$\begin{aligned} T^{[j]}(z, v) &= \frac{t_j v}{q^{-1}(1 - v(1 - q))} q^{-(j + \frac{1}{k})} \\ &= \frac{t_j v}{(1 - kvz)(1 - kz)^{j - 1 + \frac{1}{k}}}. \end{aligned}$$

Für die Verteilung der Zufallsvariable $D_{n,j}$ der Nachfolger ergibt sich damit (vgl. [18],

Abschnitt 4.1):

$$\begin{aligned}
\mathbb{P}\{D_{n,j} = m\} &= \frac{t_{n,j,m}}{t_n} = \frac{(n-j)!}{t_n} [z^{n-j} v^m] T^{[j]}(z, v) \\
&= \frac{(n-j)! t_j}{t_n} k^{m-1} [z^{n-j-m+1}] \frac{1}{(1-kz)^{j-1+\frac{1}{k}}} \\
&= \frac{(n-j)! t_j}{t_n} k^{n-j} (-1)^{n-j-m+1} \binom{-(j-1+\frac{1}{k})}{n-j-m+1} \\
&= \frac{(n-j)! j! \binom{j-1+\frac{1}{k}}{j}}{n! \binom{n-1+\frac{1}{k}}{n}} \binom{n-m-1+\frac{1}{k}}{n-j-m+1} = \frac{\binom{n-m-1+\frac{1}{k}}{n-m-j+1}}{\binom{n-1+\frac{1}{k}}{n-j}}.
\end{aligned}$$

Beliebig hinzugefügter Knoten: Um die Verteilung der Nachfahren eines beliebigen Knotens zu berechnen, verwenden wir eine andere Darstellung:

$$\begin{aligned}
\frac{\binom{n-m-1+\frac{1}{k}}{n-m-j+1}}{\binom{n-1+\frac{1}{k}}{n-j}} &= \frac{(n-j)! (n-m-1+\frac{1}{k})^{n-m-j+1}}{(n-j-m+1)! (n-1+\frac{1}{k})^{n-j}} \\
&= \frac{(n-j)^{m-1} (j-1+\frac{1}{k})}{(n-1+\frac{1}{k})^m} \\
&= \frac{(n-j)^{m-1} (n-m+\frac{1}{k})}{(n-1+\frac{1}{k})^m} - \frac{(n-j)^{m-1} (n-j-m+1)}{(n-1+\frac{1}{k})^m} \\
&= \frac{\binom{n-j}{m-1}}{\binom{n-1+\frac{1}{k}}{m-1}} - \frac{\binom{n-j}{m}}{\binom{n-1+\frac{1}{k}}{m}}.
\end{aligned}$$

Mit der Identität $\sum_{j=0}^n \binom{j}{m} = \binom{n+1}{m+1}$ erhalten wir für die Verteilung der Zufallsvariable \bar{D}_n eines beliebigen eingefügten Knotens eine geschlossene Formel:

$$\begin{aligned}
\mathbb{P}\{\bar{D}_n = m\} &= \frac{1}{n} \sum_{j=1}^n \mathbb{P}\{\bar{D}_{n,j} = m\} = \sum_{j=1}^n \frac{\binom{j-1}{m-1}}{\binom{n-1+\frac{1}{k}}{m-1}} - \frac{\binom{j-1}{m}}{\binom{n-1+\frac{1}{k}}{m}} \\
&= \frac{\binom{n}{m}}{n \binom{n-1+\frac{1}{k}}{m-1}} - \frac{\binom{n}{m+1}}{n \binom{n-1+\frac{1}{k}}{m}}.
\end{aligned}$$

Diese kann noch auf einen Bruch zusammengeführt werden:

$$\begin{aligned}
\frac{\binom{n}{m}}{n \binom{n-1+\frac{1}{k}}{m-1}} - \frac{\binom{n}{m+1}}{n \binom{n-1+\frac{1}{k}}{m}} &= \frac{\Gamma(n+1) \Gamma(n-m+\frac{1}{k}+1)}{nm \Gamma(n+\frac{1}{k}) \Gamma(n-m+1)} - \frac{\Gamma(n+1) \Gamma(n-m+\frac{1}{k})}{n(m+1) \Gamma(n+\frac{1}{k}) \Gamma(n-m)} \\
&= \frac{\Gamma(n-m+\frac{1}{k}) \Gamma(n) ((m+1)(n-m+\frac{1}{k}) - (n-m)m)}{m(m+1) \Gamma(n+\frac{1}{k}) \Gamma(n-m)} \\
&= \frac{\binom{n-m-1+\frac{1}{k}}{n-m} \Gamma(n) \Gamma(\frac{1}{k}) (\frac{m+1}{k} + n - m) \frac{1}{k}}{m(m+1) \Gamma(n+\frac{1}{k}) \frac{1}{k}} \\
&= \frac{(k(n-m)+m+1) \binom{n-m-1+\frac{1}{k}}{n-m}}{m(m+1) \binom{n-1+\frac{1}{k}}{n-1}}.
\end{aligned}$$

Um die Grenzverteilung zu bestimmen, betrachten wir zuerst nur einmal den Bruch der Binomialkoeffizienten:

$$\frac{\binom{n-m-1+\frac{1}{k}}{n-m}}{\binom{n-1+\frac{1}{k}}{n-1}} \stackrel{\text{L. 1.7}}{\sim} \frac{\Gamma\left(1+\frac{1}{k}\right)(n-m)^{-1+\frac{1}{k}}}{\Gamma\left(\frac{1}{k}\right)(n-1)^{\frac{1}{k}}} = \frac{1}{k(n-m)} \left(\frac{n-m}{n-1}\right)^{\frac{1}{k}} \sim \frac{1}{kn}.$$

Die Zufallsvariablen \bar{D}_n konvergieren daher in der Verteilung gegen eine Grenzverteilung \bar{D} mit folgender Gestalt (vgl. [18], Abschnitt 4.3):

$$\mathbb{P}\{\bar{D} = m\} = \frac{1}{m(m+1)}.$$

3.4.3 Ansteigende d -äre k -Bäume

Hinzugefügter Knoten: Bei den ansteigenden d -ären k -Bäumen kann für die Berechnung der Verteilungen $D_{n,j}$ der Nachfahren des Knotens j analog zu den geordneten ansteigenden k -Bäumen vorgegangen werden. Jeder neue Nachfahre erzeugt kd neue Plätze für Nachfahren und nimmt einen Platz ein. Für die Anzahl $t_{n,j,m}$ der ansteigenden d -ären k -Bäume, bei denen der Knoten j genau m Nachfahren hat, bedeutet dies (vgl. [18], Abschnitt 4.1):

$$\begin{aligned} t_{n,j,m} &= (d + (kd - 1)(n - 1) - kd - (kd - 1)(m - 1))t_{n-1,j,m} \\ &\quad + (kd + (kd - 1)(m - 2))t_{n-1,j,m-1} \\ &= ((kd - 1)(n - 1 - m) + d - 1)t_{n-1,j,m} + (kd + (kd - 1)(m - 2))t_{n-1,j,m-1}, \end{aligned}$$

für $m \geq 1, n \geq j + 1$. Für die erzeugende Funktion $T^{[j]}(z, v) = \sum_{n \geq j} \sum_{m \geq 1} t_{n,j,m} \frac{z^{n-j}}{(n-j)!} v^m$ erhalten wir damit die Differentialgleichung

$$(1 - (kd - 1)z)T_z^{[j]} + (kd - 1)v(1 - v)T_v^{[j]} - ((kd - 1)j + d - 1 + v)T^{[j]} = 0,$$

mit $T^{[j]}(0, v) = vt_j$, bei der t_j die Anzahl der ansteigenden d -ären k -Bäume ist. Mit den Transformationen

$$\begin{aligned} z &= \frac{1 - e^{-(kd-1)z}}{kd - 1}, & r &= -\frac{1}{kd - 1} \log(q), \\ v &= \frac{e^{(kd-1)r}}{s^{-1} - 1 + e^{(kd-1)r}}, & s &= \frac{v}{q^{-1}(1 - v(1 - q))}, \end{aligned}$$

mit $q := 1 - (kd - 1)z$ ergibt sich die gewöhnliche Differentialgleichung:

$$\tilde{T}_r^{[j]} = \left((kd - 1)j + d - 1 + \frac{e^{(kd-1)r}}{s^{-1} - 1 + e^{(kd-1)r}} \right) \tilde{T}^{[j]},$$

deren Lösung ist:

$$\tilde{T}^{[j]}(r, s) = c(s) \exp \left(r((kd - 1)j + d - 1) + \frac{1}{kd - 1} \log(v^{-1}(r, s)) + r \right).$$

Das Erfüllen der Randbedingung lässt sich berechnen durch:

$$\begin{aligned} v(0, s)t_j &= T^{[j]}(0, v(0, s)) = c(s)s^{-\frac{1}{kd-1}}, \\ \Rightarrow c(s) &= t_j s^{\frac{kd}{kd-1}}. \end{aligned}$$

So erhalten wir insgesamt die Lösung der partiellen Differentialgleichung:

$$\begin{aligned} T^{[j]}(z, v) &= t_j \left(\frac{v}{q^{-1}(1-v(1-q))} \right)^{\frac{kd}{kd-1}} q^{-j-\frac{d-1}{kd-1}} v^{-\frac{1}{kd-1}} q^{-\frac{1}{kd-1}} \\ &= \frac{vt_j}{(1-v(1-q))^{\frac{kd}{kd-1}} q^{j-1+\frac{d-1}{kd-1}}} \\ &= \frac{vt_j}{(1-(kd-1)vz)^{\frac{kd}{kd-1}} (1-(kd-1)z)^{j-1+\frac{d-1}{kd-1}}}. \end{aligned}$$

Die Koeffizienten dieser Funktion sind:

$$\begin{aligned} [z^{n-j}v^m]T^{[j]}(z, v) &= t_j \binom{-\frac{kd}{kd-1}}{m-1} (-1)^{m-1} (kd-1)^{m-1} [z^{n-j-m+1}]q^{-(j-1+\frac{d-1}{kd-1})} \\ &= t_j \binom{m-2+\frac{kd}{kd-1}}{m-1} (kd-1)^{n-j} (-1)^{n-j-m+1} \binom{-(j-1+\frac{d-1}{kd-1})}{n-j-m+1} \\ &= t_j \binom{m-1+\frac{1}{kd-1}}{m-1} (kd-1)^{n-j} \binom{n-m-1+\frac{d-1}{kd-1}}{n-j-m+1}. \end{aligned}$$

Es ergibt sich damit für die Zufallsvariable $D_{n,j}$ der Nachfahren des j -ten Knotens folgende Formel (vgl. [18], Abschnitt 4.1):

$$\begin{aligned} \mathbb{P}\{D_{n,j} = m\} &= \frac{t_{n,j,m}}{t_n} = \frac{(n-j)!j! \binom{j-1+\frac{d}{kd-1}}{j}}{n! \binom{n-1+\frac{d}{kd-1}}{n}} \binom{m-1+\frac{1}{kd-1}}{m-1} \binom{n-m-1+\frac{d-1}{kd-1}}{n-j-m+1} \\ &= \frac{\binom{m-1+\frac{1}{kd-1}}{m-1} \binom{n-m-1+\frac{d-1}{kd-1}}{n-j-m+1}}{\binom{n-1+\frac{d}{kd-1}}{n-j}}. \end{aligned}$$

Beliebig hinzugefügter Knoten: Wie man am Ergebnis von $\mathbb{P}\{D_{n,j} = m\}$ ersieht, werden auch die darauf „aufbauenden“ Zufallsvariablen, abgesehen von den Konstanten die gleiche Form wie bei den geordneten ansteigenden k -Bäumen haben. So wird auch beim Berechnen von der Zufallsvariablen \bar{D}_n eines beliebigen eingefügten Knotens Lemma

1.11 verwendet:

$$\begin{aligned}
\mathbb{P}\{\bar{D}_n = m\} &= \frac{1}{n} \sum_{j=1}^n \mathbb{P}\{D_{n,j} = m\} = \frac{1}{n} \binom{m-1 + \frac{1}{kd-1}}{m-1} \sum_{j=1}^n \frac{\binom{n-m-1 + \frac{d-1}{kd-1}}{n-j-m+1}}{\binom{n-1 + \frac{d}{kd-1}}{n-j}} \\
&= \frac{\binom{m-1 + \frac{1}{kd-1}}{m-1} \left(n + \frac{d}{kd-1}\right)}{nm \binom{m+1 + \frac{1}{kd-1}}{m-1}} \\
&\quad - \frac{\binom{m-1 + \frac{1}{kd-1}}{m-1}}{n \binom{n-1 + \frac{d}{kd-1}}{n}} \sum_{l=0}^{m-1} \frac{\binom{m-1}{l} (-1)^l \frac{d}{kd-1}}{l+2 + \frac{1}{kd-1}} \binom{n-l-2 + \frac{d-1}{kd-1}}{n} \\
&= \frac{kd \left(n + \frac{d}{kd-1}\right)}{n(kd-1) \left(m + \frac{kd}{kd-1}\right) \left(m + \frac{1}{kd-1}\right)} \\
&\quad - \frac{\binom{m-1 + \frac{1}{kd-1}}{m-1}}{n \binom{n-1 + \frac{d}{kd-1}}{n}} \sum_{l=0}^{m-1} \frac{\binom{m-1}{l} (-1)^l d}{(kd-1)(l+1) + kd} \binom{n-l-2 + \frac{d-1}{kd-1}}{n}.
\end{aligned}$$

Für die Grenzverteilung \bar{D} bleibt wieder nur der erste Summand übrig:

$$\begin{aligned}
\lim_{n \rightarrow \infty} \frac{kd \left(n + \frac{d}{kd-1}\right)}{n(kd-1) \left(m + \frac{kd}{kd-1}\right) \left(m + \frac{1}{kd-1}\right)} &= \frac{kd}{(kd-1) \left(m + \frac{kd}{kd-1}\right) \left(m + \frac{1}{kd-1}\right)}, \\
\lim_{n \rightarrow \infty} \frac{\binom{n-l-2 + \frac{d-1}{kd-1}}{n}}{n \binom{n-1 + \frac{d}{kd-1}}{n}} &\stackrel{\text{L. 1.8}}{=} \frac{\Gamma\left(\frac{d}{kd-1}\right)}{\Gamma\left(-l-1 + \frac{d-1}{kd-1}\right)} \lim_{n \rightarrow \infty} \frac{n^{-l-2 + \frac{d-1}{kd-1}}}{nn^{-1 + \frac{d}{kd-1}}} = 0,
\end{aligned}$$

womit sich für die Verteilung von \bar{D} folgendes ergibt (vgl. [18], Abschnitt 4.3):

$$\mathbb{P}\{\bar{D} = m\} = \frac{kd}{(kd-1) \left(m + \frac{kd}{kd-1}\right) \left(m + \frac{1}{kd-1}\right)}.$$

3.4.4 CU-Modell

Beim CU-Modell betrachten wir wie schon bei den Kindern die Möglichkeiten, einen k -Baum zu konstruieren und welche davon die Anzahl der Nachfahren erhöhen. Wenn ein Knoten nur sich selbst als Nachfahre hat, ist er Teil einer $(k+1)$ -Clique und damit gibt es k Möglichkeiten, einen Nachfahren hinzuzufügen. Jeder weitere Nachfahre erzeugt eine neue $(k+1)$ -Clique, wobei jeder Knoten, der zu einer k -Teilclique dieser Clique hinzugefügt wird, ein Nachfahre wäre. Das heißt, jeder Nachfahre erhöht die Möglichkeiten eines weiteren Nachfahrens um $k+1$. Für die Möglichkeiten $t_{n,j,m}$ einen k -Baum mit n Knoten zu konstruieren, bei dem der Knoten j genau m Nachfahren hat, bedeutet dies:

$$t_{n,j,m} = ((n-1)(k+1) - k - (k+1)(m-1))t_{n-1,j,m} + (k + (k+1)(m-2))t_{n-1,j,m-1},$$

mit $t_{j,j,1} = 1$ und $t_{j,j,m} = 0$ für $j \geq 1$ und $1 \neq m \in \mathbb{N}$. Für die bivariate erzeugende Funktion $T^{[j]}(z, v) = \sum_{n \geq j} \sum_{m \geq 1} t_{n,j,m} \frac{z^{n-j}}{(n-j)!} v^m$ erhält man daraus die partielle Differentialgleichung:

$$(1 - (k+1)z)T_z^{[j]} + (k+1)v(1-v)T_v^{[j]} - (j(k+1) + 1 - v)T^{[j]} = 0,$$

mit der Randbedingung $T^{[j]}(0, v) = vt_j$. Wir können die gleiche Koordinatentransformation wie bei den geordneten ansteigenden k -Bäumen durchführen, um die Gleichung zu lösen.

$$\begin{aligned} z &= \frac{1 - e^{-(k+1)z}}{k+1}, & r &= -\frac{1}{k+1} \log(q), \\ v &= \frac{e^{(k+1)r}}{s^{-1} - 1 + e^{(k+1)r}}, & s &= \frac{v}{q^{-1}(1 - v(1 - q))}, \end{aligned}$$

wobei $q := 1 - (k+1)z$, was für die Differentialgleichung zur Folge hat:

$$\tilde{T}_r^{[j]} = j(k+1) + 1 + \frac{e^{(k+1)r}}{s^{-1} - 1 + e^{(k+1)r}} \tilde{T}^{[j]}$$

für die transformierte Funktion $\tilde{T}^{[j]}(r, s) = T^{[j]}(z(r, s), v(r, s))$. Die Lösung davon ist:

$$\tilde{T}^{[j]}(r, s) = c(s) \exp\left(r(j(k+1) + 1) - \frac{1}{k+1} \log(v^{-1}(r, s)) - r\right),$$

wobei die Randbedingung noch zur Berechnung der Integrationskonstante benützt werden kann:

$$t_j s = \tilde{T}^{[j]}(0, s) = c(s) s^{\frac{1}{k+1}} \quad \Rightarrow \quad c(s) = t_j s^{\frac{k}{k+1}}.$$

So erhalten wir:

$$\begin{aligned} T^{[j]}(z, v) &= t_j \left(\frac{v}{q^{-1}(1 - v(1 - q))} \right)^{\frac{k}{k+1}} q^{-j} v^{\frac{1}{k+1}} \\ &= \frac{t_j v}{(1 - (k+1)vz)^{\frac{k}{k+1}} (1 - (k+1)z)^{j-1 + \frac{1}{k+1}}}. \end{aligned}$$

Die Koeffizienten von $T^{[j]}(z, v)$ lauten:

$$\begin{aligned} [z^{n-j} v^m] \frac{T^{[j]}(z, v)}{t_j} &= (k+1)^{m-1} (-1)^{m-1} \binom{-\frac{k}{k+1}}{m-1} [z^{n-j-m+1}] q^{-j+1 - \frac{1}{k+1}} \\ &= (k+1)^{n-j} \binom{m-2 + \frac{k}{k+1}}{m-1} (-1)^{n-m-j+1} \binom{-j+1 - \frac{1}{k+1}}{n-m-j+1} \\ &= (k+1)^{n-j} \binom{m-1 - \frac{1}{k+1}}{m-1} \binom{n-m-1 + \frac{1}{k+1}}{n-m-j+1}. \end{aligned}$$

Die Verteilungsfunktion der Zufallsvariable $D_{n,j}$ der Nachfolger des j -ten Knotens können wir durch die Koeffizienten ausdrücken:

$$\begin{aligned} \mathbb{P}\{D_{n,j} = m\} &= \frac{t_{n,j,m}}{t_n} = \frac{(n-j)!}{t_n} [z^{n-j} v^m] T^{[j]}(z, v) \\ &= \frac{(j-1)!(k+1)^{j-1} (n-j)!}{(n-1)!(k+1)^{n-1}} (k+1)^{n-j} \binom{m-1 - \frac{1}{k+1}}{m-1} \binom{n-m-1 + \frac{1}{k+1}}{n-m-j+1} \\ &= \frac{1}{\binom{n-1}{j-1}} \binom{m-1 - \frac{1}{k+1}}{m-1} \binom{n-m-1 + \frac{1}{k+1}}{n-m-j+1}. \end{aligned}$$

Für die Zufallsvariable \bar{D}_n der Nachfahren eines beliebigen eingefügten Knotens ergibt dies die Summendarstellung:

$$\mathbb{P}\{\bar{D}_n = m\} = \frac{1}{n} \sum_{j=1}^n \mathbb{P}\{D_{n,j} = m\} = \frac{1}{n} \binom{m-1-\frac{1}{k+1}}{m-1} \sum_{j=1}^n \frac{\binom{n-m-1+\frac{1}{k+1}}{n-m-j+1}}{\binom{n-1}{j-1}}.$$

Wegen $\binom{n-1}{n-j} = \binom{n+\beta}{n-j}$, also $-\beta \in \mathbb{N}$, kann nicht direkt Lemma 1.11 verwendet werden, aber bis auf die letzten Umformungsschritte kann analog zum Beweis dieses Lemmas vorgegangen werden. Dies liefert:

$$\mathbb{P}\{\bar{D}_n = m\} = \frac{\binom{m-1-\frac{1}{k+1}}{m-1}}{n} \sum_{l=0}^{m-1} \frac{\binom{m-1}{l} (-1)^l}{\binom{n-1}{l+1-\frac{1}{k+1}}} \left(\binom{n}{l+2-\frac{1}{k+1}} - \binom{0}{l+2-\frac{1}{k+1}} \right).$$

Mit

$$\frac{\binom{n}{l+2-\frac{1}{k+1}}}{n \binom{n-1}{l+1-\frac{1}{k+1}}} = \frac{\Gamma(n+1)\Gamma(l+2-\frac{1}{k+1})\Gamma(n-1-l+\frac{1}{k+1})}{n\Gamma(n)\Gamma(l+3-\frac{1}{k+1})\Gamma(n-1-l+\frac{1}{k+1})} = \frac{1}{l+2-\frac{1}{k+1}}$$

erhalten wir für die Verteilung \bar{D} ($\bar{D}_n \xrightarrow{(d)} \bar{D}$):

$$\begin{aligned} \mathbb{P}\{\bar{D} = m\} &= \binom{m-1-\frac{1}{k+1}}{m-1} \sum_{l=0}^{m-1} \frac{\binom{m-1}{l} (-1)^l}{l+2-\frac{1}{k+1}} \stackrel{\text{L. 1.9}}{=} \frac{\binom{m-1-\frac{1}{k+1}}{m-1}}{(m+1-\frac{1}{k+1})\binom{m-\frac{1}{k+1}}{m-1}} \\ &= \frac{k+1}{k(m+1-\frac{1}{k+1})(m-\frac{1}{k+1})}. \end{aligned}$$

Die Verteilung für $n \rightarrow \infty$ stimmt für das CU-Modell wie auch schon beim Grad eines Knotens mit jener für die geordneten ansteigenden k -Bäume überein.

3.5 Vorfahren

In diesem Abschnitt werden noch die Vorfahren von Knoten untersucht. Ein Knoten x ist Vorfahre von y , wenn y Nachfahre von x ist. Somit ist insbesondere jeder Knoten Vorfahre von sich selbst und jeder Wurzelknoten ist Vorfahre eines jeden Knotens (abgesehen von den anderen Knoten in der Wurzel) in einem k -Baum. Damit ist die Anzahl der Vorfahren für einen eingefügten Knoten nach unten beschränkt mit $k+1$ und für Wurzelknoten stimmt sie mit 1 überein.

Für die Bestimmung der Vorfahren wird eine Rekursion gelöst. Bisher haben wir immer einen Knoten festgehalten und haben in der Rekursion untersucht, ob ein weiterer Knoten zu einem gewissen Teil hinzugefügt wird. Bei den Vorfahren müssen wir umgekehrt vorgehen, da sich die Anzahl der Vorfahren eines Knotens nach dessen Einfügung in den k -Baum nicht mehr ändert. In der Darstellung als Graphen gehen wir im Folgenden in der Rekursion eine Kante in Richtung Wurzel und damit verringern sich die Vorfahren sicherlich um mindestens 1. Wie man in der Abbildung 20 sieht, können Vorfahren „übersprungen“ werden, wenn nicht der Knoten mit der höchstmöglichen Markierung gewählt wird. Wenn ein Knoten x Kind einer Clique C , bei welcher der Knoten y die

höchste Markierung innerhalb der Clique aufweist, ist, gewährleistet der rekursive Aufbau von k -Bäumen, dass der Knoten y Kind einer Clique ist, welche $C \setminus \{y\}$ umfasst. Somit sind auch die Vorfahren von y genau um eines geringer als von x .

Deshalb unterscheiden wir, ob der n -te Knoten ein Kind vom Knoten mit der Markierung $n - 1$ ist. Wenn er es ist, weist der Knoten $n - 1$ klarerweise die höchste Markierung in der Elternclique von n auf und die Anzahl der Vorfahren weist einen Unterschied um 1 auf. Ist der Knoten n kein Nachfahre von $n - 1$, können wir die Markierung der beiden Knoten vertauschen und durch Löschen erhalten wir einen k -Baum der Größe $n - 1$, bei dem wir die Vorfahren von $n - 1$ untersuchen müssen, um die Anzahl der Vorfahren von Knoten n im ursprünglichen Baum zu bestimmen. Das heißt, für das Aufstellen der Rekursion wird nur noch abhängig vom betrachteten Modell die Zahl der Plätze beziehungsweise Möglichkeiten, sodass Knoten n Kind von $n - 1$ ist, benötigt (vgl. [18], Abschnitt 5.1). Im Folgenden werden die Verteilungsfunktionen der Vorgänger für den n -ten Knoten bestimmt. Zudem wird auch die wahrscheinlichkeitserzeugende Funktion eines beliebig eingefügten Knotens in einem Baum der Größe n berechnet, womit das asymptotische Verhalten für einen zufällig ausgewählten Knoten bestimmt werden kann.

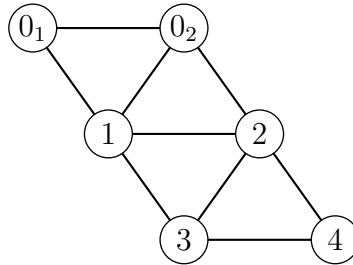


Abbildung 20: Ein ansteigender 2-Baum, bei dem der Knoten 4 sechs Vorfahren hat, während es beim Knoten 2 vier sind.

3.5.1 Geordnete ansteigende k -Bäume

Hinzugefügter Knoten: Bei den geordneten ansteigenden k -Bäumen entstehen beim Hinzufügen eines Knotens k neue Plätze, jeweils einer bei einer neu entstandenen Clique und ein Platz bei der Clique, bei welcher der Knoten hinzugefügt wird. Wenn wir den Knoten $n - 1$ hinzufügen, sind davon nur die k Plätze bei den neuen Cliquen mögliche Kandidaten für Knoten n , sodass dieser ein Kind von $n - 1$ wird. Für die Rekursion bedeutet dies, dass es k Plätze gibt, bei denen sich die Anzahl der Vorfahren um 1 verringert und bei den restlichen die Vorfahren gleich bleiben. In einer Formel ausgedrückt heißt dies für die Anzahl $t_{n,m}$ der geordneten, ansteigenden k -Bäume, bei denen der Knoten n genau m Vorfahren hat (vgl. [18], Abschnitt 5.1):

$$\begin{aligned}
 t_{n,m} &= kt_{n-1,m-1} + (1 + (k + 1)(n - 1) - k)t_{n-1,m} & n \geq 2, m \geq 1, \\
 &= kt_{n-1,m-1} + ((k + 1)(n - 2) + 2)t_{n-1,m}
 \end{aligned}$$

wobei $t_{1,k+1} = 1$ und $t_{1,m} = 0$ für $k + 1 \neq m \in \mathbb{N}$ gilt. Diese Rekursion wird so wie die bisherigen mit der bivariaten erzeugenden Funktion $T(z, v) = \sum_{n \geq 1} \sum_{m \geq 1} t_{n,m} \frac{z^{n-1}}{(n-1)!} v^m$ gelöst. Wie beim Berechnen des Knotengrades kann die Rekursion aufsummiert werden,

mit dem Ergebnis:

$$(1 - (k + 1)z)T_z - (kv + 2)T = 0, \quad T(0, v) = v^{k+1}.$$

Im Gegensatz zum Knotengrad und zu den Nachfahren tritt bei der Berechnung des Vorfahrens nur eine gewöhnliche Differentialgleichung auf und keine Partielle. Diese separable Gleichung können wir direkt lösen, ohne dass weitere Vorarbeiten nötig sind:

$$\begin{aligned} \frac{T_z}{T} &= \frac{kv + 2}{1 - (k + 1)z}, \\ \log(T(z, v)) &= -\frac{kv + 2}{k + 1} \log(1 - (k + 1)z) + \tilde{c}(v), \\ T(z, v) &= c(v) \exp\left(-\frac{kv + 2}{k + 1} \log(1 - (k + 1)z)\right), \\ &= \frac{c(v)}{(1 - (k + 1)z)^{\frac{kv+2}{k+1}}}. \end{aligned}$$

Mit Hilfe der Randbedingung $T(0, v) = v^{k+1}$ kann die Integrationskonstante bestimmt werden ($c(v) = v^{k+1}$). Damit erhalten wir:

$$T(z, v) = \frac{v^{k+1}}{(1 - (k + 1)z)^{\frac{kv+2}{k+1}}}.$$

Die Verteilungsfunktion der Zufallsvariable A_n der Vorfahren des n -ten Knotens in einem geordneten ansteigenden k -Baum kann über die Koeffizienten berechnet werden:

$$\mathbb{P}\{A_n = m + k\} = \frac{t_{n, m+k}}{t_n} = \frac{(n-1)!}{t_n} [z^{n-1} v^{m+k}] T(z, v).$$

Für die Berechnung der Koeffizienten wird die Gleichung (1.3) der Stirling-Zahlen erster Art verwendet:

$$\begin{aligned} [z^{n-1} v^{m+k}] T(z, v) &= [z^{n-1} v^{m-1}] \frac{1}{(1 - (k + 1)z)^{\frac{kv+2}{k+1}}} \\ &= \left(\frac{k}{k+1}\right) \sum_{i \geq m-1} \frac{\begin{bmatrix} i \\ m-1 \end{bmatrix}}{i!} (k+1)^i [z^{n-i-1}] \frac{1}{(1 - (k + 1)z)^{\frac{2}{k+1}}} \\ &= \left(\frac{k}{k+1}\right) (k+1)^{n-1} \sum_{i=m-1}^{n-1} \frac{\begin{bmatrix} i \\ m-1 \end{bmatrix}}{i!} (-1)^{n-i-1} \binom{-\frac{2}{k+1}}{n-i-1} \\ &= \left(\frac{k}{k+1}\right) (k+1)^{n-1} \sum_{i=m}^n \frac{\begin{bmatrix} i-1 \\ m-1 \end{bmatrix}}{(i-1)!} \binom{n-i-2 + \frac{2}{k+1}}{n-i}. \end{aligned}$$

Die Verteilungsfunktion von A_n lautet damit (vgl. [18], Abschnitt 5.1):

$$\begin{aligned}\mathbb{P}\{A_n = m + k\} &= \frac{(n-1)!(k+1)^{n-1}}{n!(k+1)^n \binom{n-1+\frac{1}{k+1}}{n}} \binom{k}{k+1} \sum_{i=m}^n \frac{\lfloor \frac{i-1}{m-1} \rfloor}{(i-1)!} \binom{n-i-2+\frac{2}{k+1}}{n-i} \\ &= \frac{k^{m-1}}{(k+1)^{m-1} n \binom{n-1+\frac{1}{k+1}}{n}} \sum_{i=m}^n \frac{\lfloor \frac{i-1}{m-1} \rfloor}{(i-1)!} \binom{n-i-2+\frac{2}{k+1}}{n-i} \\ &= \frac{k^{m-1}}{(k+1)^{m-1} \binom{n-1+\frac{1}{k+1}}{n-1}} \sum_{i=m}^n \frac{\lfloor \frac{i-1}{m-1} \rfloor}{(i-1)!} \binom{n-i-2+\frac{2}{k+1}}{n-i}.\end{aligned}$$

Beliebig hinzugefügter Knoten: Für die Zufallsvariable \bar{A}_n der Vorfahren eines zufälligen hinzugefügten Knotens werden wir die wahrscheinlichkeitserzeugende Funktion $\bar{P}_n(v) = \sum_{m \geq 0} \mathbb{P}\{\bar{A}_n = m\} v^m$ berechnen, um mit dem „Quasi-Power Theorem“ von Hwang das asymptotische Verhalten der Zufallsvariable zu betrachten. Da sich \bar{P}_n aus der wahrscheinlichkeitserzeugenden Funktion $P_n(v) = \sum_{m \geq 0} \mathbb{P}\{A_n = m\} v^m$ von A_n berechnen lässt, untersuchen wir diese zuvor:

$$\begin{aligned}P_n(v) &= \frac{(n-1)!}{t_n} [z^{n-1}] T(z, v) = \frac{(n-1)! v^{k+1}}{t_n} (k+1)^{n-1} \binom{-\frac{kv+2}{k+1}}{n-1} \\ &= \frac{v^{k+1}}{n(k+1) \binom{n-1+\frac{1}{k+1}}{n}} \binom{n-2+\frac{kv+2}{k+1}}{n-1} = v^{k+1} \frac{\binom{n-2+\frac{kv+2}{k+1}}{n-1}}{\binom{n-1+\frac{1}{k+1}}{n-1}}.\end{aligned}$$

Für die wahrscheinlichkeitserzeugende Funktion von \bar{A}_n ergibt das:

$$\bar{P}_n(v) = \frac{1}{n} \sum_{j=1}^n P_j(v) = \frac{v^{k+1}}{n} \sum_{j=1}^n \frac{\binom{j-2+\frac{kv+2}{k+1}}{j-1}}{\binom{n-1+\frac{1}{k+1}}{j-1}}.$$

Für v in einer Umgebung um 1 ist $1 - \frac{kv+1}{k+1} \notin \mathbb{N}^\times$ und mit Lemma 1.12 erhalten wir:

$$\bar{P}_n(v) = \frac{1}{n} \sum_{j=1}^n P_j(v) = \frac{v^{k+1}}{(kv+1)n} \left(\frac{\binom{n-1+\frac{kv+2}{k+1}}{n}}{\binom{n-1+\frac{1}{k+1}}{n}} - 1 \right),$$

und die bezüglich v gleichmäßige asymptotische Näherung:

$$\bar{P}_n(v) = \frac{v^{k+1} \Gamma\left(\frac{1}{k+1}\right)}{(kv+1) \Gamma\left(\frac{kv+2}{k+1}\right)} n^{\frac{kv+1}{k+1}-1} (1 + \mathcal{O}(n^{-1})).$$

Mit $\beta_n = \log n$, $B(v) = \exp\left(\frac{kv+1}{k+1} - 1\right)$, wobei $B''(1) + B'(1) - B'(1)^2 = \frac{k}{k+1} \neq 0$ ist, können wir das „Quasi-Power Theorem“ anwenden. Damit ist die standardisierte Zufallsvariable asymptotisch normalverteilt:

$$\mathbb{P}\left\{ \frac{\bar{A}_n - \mathbb{E}(\bar{A}_n)}{\sqrt{\mathbb{V}(\bar{A}_n)}} \leq x \right\} = \phi(x) + \mathcal{O}\left(\frac{1}{\sqrt{\log n}}\right),$$

wobei $\phi(x)$ die Verteilungsfunktion der Normalverteilung ist und für den Erwartungswert gilt (vgl. [18], Abschnitt 5.2):

$$\mathbb{E}(\bar{A}_n) = \log(n) B'(1) + \mathcal{O}(1) = \frac{k}{k+1} \log(n) + \mathcal{O}(1).$$

3.5.2 Ansteigende k -Bäume

Hinzugefügter Knoten: Bei den ansteigenden k -Bäumen besitzt jede Clique einen Platz für weitere Knoten. Da ein Knoten beim Hinzufügen Teil von k Cliquen ist, gibt es k Plätze, sodass der n -te Knoten beim $(n-1)$ -ten als Kind hinzugefügt wird. Für die Anzahl $t_{n,m}$ der ansteigenden k -Bäume der Größe n , bei denen der n -te Knoten m Vorfahren hat, ergibt sich eine Rekursion mit der Gestalt (vgl. [18], Abschnitt 5.1):

$$t_{n,m} = kt_{n-1,m-1} + (1 + k(n-1) - k)t_{n-1,m}, \quad n \geq 2, m \geq 1.$$

Die Startwerte der Rekursion sind $t_{1,k+1} = 1$ und $t_{1,m} = 0$ für $k+1 \neq m \in \mathbb{N}$. Aus dieser Gleichung kann durch Aufsummieren und Einsetzen der erzeugenden Funktion $T(z, v) = \sum_{n \geq 1} \sum_{m \geq 1} t_{n,m} \frac{z^{n-1}}{(n-1)!} v^m$ die gewöhnliche Differentialgleichung

$$(1 - kz)T_z - (kv + 1)T = 0, \quad T(0, v) = v^{k+1}$$

hergeleitet werden. Deren Lösung kann analog zu den geordneten ansteigenden Bäumen hergeleitet werden und stimmt mit

$$T(z, v) = \frac{v^{k+1}}{(1 - kz)^{v + \frac{1}{k}}}$$

überein. Die Koeffizienten können mit Hilfe der Stirling-Zahlen erster Art ausgedrückt werden:

$$\begin{aligned} [z^{n-1}v^{m+k}]T(z, v) &= [z^{n-1}v^{m-1}] \frac{v^{k+1}}{(1 - kz)^{v + \frac{1}{k}}} = [z^{n-1}] \left(\frac{1}{(1 - kz)^{\frac{1}{k}}} \sum_{i \geq m-1} \begin{bmatrix} i \\ m-1 \end{bmatrix} k^i \frac{z^i}{i!} \right) \\ &= \sum_{i \geq m-1} \begin{bmatrix} i \\ m-1 \end{bmatrix} \frac{k^i}{i!} [z^{n-i-1}] \frac{1}{(1 - kz)^{\frac{1}{k}}} \\ &= k^{n-1} \sum_{i=m-1}^{n-1} \frac{\begin{bmatrix} i \\ m-1 \end{bmatrix}}{i!} (-1)^{n-i-1} \binom{-\frac{1}{k}}{n-i-1} \\ &= k^{n-1} \sum_{i=m-1}^{n-1} \frac{\begin{bmatrix} i \\ m-1 \end{bmatrix}}{i!} \binom{n-i-2 + \frac{1}{k}}{n-i-1}. \end{aligned}$$

Für die Zufallsvariable A_n der Vorfahren des Knotens n mit $n \geq 1$ ergibt dies für die Verteilungsfunktion die Gestalt (vgl. [18], Abschnitt 5.1):

$$\begin{aligned} \mathbb{P}\{A_n = m + k\} &= \frac{t_{n,m+k}}{t_n} = \frac{(n-1)!k^{n-1}}{n!k^n \binom{n-1+\frac{1}{k}}{n}} \sum_{i=m-1}^{n-1} \frac{\begin{bmatrix} i \\ m-1 \end{bmatrix}}{i!} \binom{n-i-2 + \frac{1}{k}}{n-i-1} \\ &= \frac{1}{\binom{n-1+\frac{1}{k}}{n-1}} \sum_{i=m}^n \frac{\begin{bmatrix} i-1 \\ m-1 \end{bmatrix}}{(i-1)!} \binom{n-i-1 + \frac{1}{k}}{n-i}. \end{aligned}$$

Beliebig hinzugefügter Knoten: Die wahrscheinlichkeitserzeugende Funktion $P_n(v) =$

$\sum_{m \geq 0} \mathbb{P}\{\bar{A}_n = m\} v^m$ kann auch direkt aus $T(z, v)$ berechnet werden. Es gilt:

$$\begin{aligned} P_n(v) &= \frac{(n-1)!}{t_n} [z^{n-1}] T(z, v) = \frac{(n-1)! v^{k+1} k^{n-1}}{n! k^n \binom{n-1+\frac{1}{k}}{n}} (-1)^{n-1} \binom{-(v+\frac{1}{k})}{n-1} \\ &= \frac{v^{k+1}}{nk \binom{n-1+\frac{1}{k}}{n}} \binom{n-2+v+\frac{1}{k}}{n-1} = v^{k+1} \frac{\binom{n-2+v+\frac{1}{k}}{n-1}}{\binom{n-1+\frac{1}{k}}{n-1}}. \end{aligned}$$

Wie bei den geordneten ansteigenden k -Bäumen können wir auch hier Lemma 1.12 für v in einer Umgebung von 1 verwenden, um die wahrscheinlichkeitserzeugende Funktion $\bar{P}_n(v) = \sum_{m \geq 0} \mathbb{P}\{\bar{A}_n = m\} v^m$ der Zufallsvariable \bar{A}_n eines beliebigen eingefügten Knotens in einem ansteigenden k -Baum der Größe n zu bestimmen.

$$\begin{aligned} \bar{P}_n(v) &= \frac{1}{n} \sum_{j=1}^n P_j(v) = \frac{v^{k+1}}{n} \sum_{j=1}^n \frac{\binom{j-2+v+\frac{1}{k}}{j-1}}{\binom{j-1+\frac{1}{k}}{j-1}} \\ &= \frac{v^k}{kn} \left(\frac{\binom{n-1+v+\frac{1}{k}}{n}}{\binom{n-1+\frac{1}{k}}{n}} - 1 \right) \end{aligned}$$

Wenn $v \in \mathbb{C}$ mit $|v-1| \leq \epsilon$ erhalten wir die Darstellung

$$\bar{P}_n(v) = \frac{v^k \Gamma\left(\frac{1}{k}\right)}{k \Gamma\left(v + \frac{1}{k}\right)} n^{v-1} \left(1 + \mathcal{O}(n^{-1})\right),$$

bei welcher die Asymptotik gleichmäßig für v gilt. Es werden alle Voraussetzungen für das „Quasi-Power Theorem“ erfüllt, sodass \bar{A}_n asymptotisch normalverteilt ist mit dem Erwartungswert (vgl. [18], Abschnitt 5.2):

$$\mathbb{E}(\bar{A}_n) = \log(n) \frac{d}{dv} \exp(v-1) \Big|_{v=1} + \mathcal{O}(1) = \log(n) + \mathcal{O}(1).$$

3.5.3 Ansteigende d -äre k -Bäume

Hinzugefügter Knoten: Es sei $t_{n,m}$ die Anzahl der ansteigenden d -ären k -Bäume, bei denen der Knoten n genau m Vorfahren hat. Wenn ein Knoten x bei einer Clique hinzugefügt wird, entstehen k neue Cliques mit d Plätzen für Kinder von x . Das heißt, es gibt kd Plätze, bei denen der Knoten n ein Kind von $n-1$ ist. So erhalten wir die Rekursion (vgl. [18], Abschnitt 5.1):

$$\begin{aligned} t_{n,m} &= kdt_{n-1,m-1} + (d + (kd-1)(n-1) - kd)t_{n-1,m} \\ &= kdt_{n-1,m-1} + (d-1 + (kd-1)(n-2))t_{n-1,m}, \end{aligned} \quad n \geq 2, m \geq 1.$$

Da die Wurzelclique d verschiedene Plätze für Kinder hat, sind die Startwerte der Rekursion gegeben mit $t_{1,k+1} = d$ und $t_{1,m} = 0$ für $k+1 \neq m \in \mathbb{N}$. Die erzeugende Funktion $T(z, v) = \sum_{n \geq 1} \sum_{m \geq 1} t_{n,m} \frac{z^{n-1}}{(n-1)!} v^m$ erfüllt wegen der Rekursion die Differentialgleichung:

$$(1 - (kd-1)z)T_z - (d-1 + vkd)T = 0, \quad T(0, v) = dv^{k+1}.$$

Die Lösung dieser separablen Differentialgleichung lässt sich mit analogen Rechnungen wie bei den vorigen Resultaten erhalten und ist gegeben durch (vgl. [18], Abschnitt 5.1):

$$T(z, v) = \frac{dv^{k-1}}{(1 - (kd - 1)z)^{\frac{kdv+d-1}{kd-1}}}.$$

Um die Verteilungsfunktion der Vorfahren A_n in einem d -ären k -Baum zu bestimmen, werden die Koeffizienten der erzeugenden Funktion $T(z, v)$ benötigt:

$$\begin{aligned} [z^{n-1}v^{m+k}]T(z, v) &= [z^{n-1}v^{m-1}] \frac{d}{(1 - (kd - 1)z)^{\frac{kdv+d-1}{kd-1}}} \\ &= d \left(\frac{kd}{kd-1} \right)^{m-1} \sum_{i \geq m-1} \frac{\begin{bmatrix} i \\ m-1 \end{bmatrix}}{i!} (kd-1)^i [z^{n-i-1}] \frac{1}{(1 - (kd-1)z)^{\frac{d-1}{kd-1}}} \\ &= d \left(\frac{kd}{kd-1} \right)^{m-1} (kd-1)^{n-1} \sum_{i=m-1}^{n-1} \frac{\begin{bmatrix} i \\ m-1 \end{bmatrix}}{i!} (-1)^{n-i-1} \binom{-\frac{d-1}{kd-1}}{n-i-1} \\ &= d \left(\frac{kd}{kd-1} \right)^{m-1} (kd-1)^{n-1} \sum_{i=m-1}^{n-1} \frac{\begin{bmatrix} i \\ m-1 \end{bmatrix}}{i!} \binom{n-i-2+\frac{d-1}{kd-1}}{n-i-1}. \end{aligned}$$

Die Verteilung von A_n lautet nun:

$$\begin{aligned} \mathbb{P}\{A_n = m+k\} &= \frac{t_{n,m+k}}{t_n} \\ &= \frac{(n-1)!d(kd)^{m-1}(kd-1)^{n-1}}{n!(kd-1)^n \binom{n-1+\frac{d}{kd-1}}{n-1}} \sum_{i=m}^n \frac{\begin{bmatrix} i-1 \\ m-1 \end{bmatrix}}{(i-1)!} \binom{n-i-1+\frac{d-1}{kd-1}}{n-i} \\ &= \frac{(kd)^{m-1}}{(kd-1)^{m-1} \binom{n-1+\frac{d}{kd-1}}{n-1}} \sum_{i=m}^n \frac{\begin{bmatrix} i-1 \\ m-1 \end{bmatrix}}{(i-1)!} \binom{n-i-1+\frac{d-1}{kd-1}}{n-i}. \end{aligned}$$

Beliebig hinzugefügter Knoten: Das asymptotische Verhalten der Zufallsvariable der Vorfahren eines beliebigen eingefügten Knotens wird als nächstes studiert. Dafür wird zuerst die wahrscheinlichkeitserzeugende Funktion $P_n(v) = \sum_{m \geq 0} \mathbb{P}\{A_n = m\}v^m$ von A_n berechnet:

$$\begin{aligned} P_n(v) &= \frac{(n-1)!}{t_n} [z^{n-1}]T(z, v) = \frac{(n-1)!v^{k+1}d}{n!(kd-1)^n \binom{n-1+\frac{d}{kd-1}}{n-1}} (kd-1)^{n-1} (-1)^{n-1} \binom{-\frac{kdv+d-1}{kd-1}}{n-1} \\ &= \frac{dv^{k+1}}{d \binom{n-1+\frac{d}{kd-1}}{n-1}} \binom{n-2+\frac{kdv+d-1}{kd-1}}{n-1}. \end{aligned}$$

Wie auch schon bei den vorigen zwei Modellen können wir mit Hilfe von Lemma 1.12 die wahrscheinlichkeitserzeugende Funktion $\bar{P}_n(v) = \sum_{m \geq 0} \mathbb{P}\{\bar{A}_n = m\}v^m$ von \bar{A}_n , der

Zufallsvariable eines beliebigen eingefügten Knotens, berechnen und kompakt darstellen.

$$\begin{aligned}
\bar{P}_n(v) &= \frac{1}{n} \sum_{j=1}^n P_j(v) = \frac{v^{k+1}}{n} \sum_{j=1}^n \frac{\binom{j-2+\frac{kdv+d-1}{kd-1}}{j-1}}{\binom{j-1+\frac{d}{kd-1}}{j-1}} \\
&= \frac{v^{k+1}}{n} \frac{\frac{d}{kd-1}}{-1 + \frac{kdv-1}{kd-1} + 1} \left(\frac{\binom{n-1+\frac{kdv+d-1}{kd-1}}{n}}{\binom{n-1+\frac{d}{kd-1}}{n}} - 1 \right) \\
&= \frac{v^{k+1}d}{n(kdv-1)} \left(\frac{\binom{n-1+\frac{kdv+d-1}{kd-1}}{n}}{\binom{n-1+\frac{d}{kd-1}}{n}} - 1 \right).
\end{aligned}$$

Für $v \in \mathbb{C}$ mit $|v-1| \leq \epsilon$ erhalten wir die bezüglich n asymptotische Darstellung:

$$\bar{P}_n(v) = \frac{dv^{k+1}\Gamma\left(\frac{d}{kd-1}\right)}{(kdv-1)\Gamma\left(\frac{kdv+d-1}{kd-1}\right)} n^{\frac{kdv-1}{kd-1}-1} (1 + \mathcal{O}(n^{-1})).$$

Es sind alle Voraussetzungen für das „Quasi-Power Theorem“ von Hwang erfüllt. Daher ist \bar{A}_n asymptotisch normalverteilt mit dem Erwartungswert (vgl. [18], Abschnitt 5.2):

$$\mathbb{E}(\bar{A}_n) = \log(n) \frac{d}{dv} \exp\left(\frac{kdv-1}{kd-1} - 1\right) \Big|_{v=1} + \mathcal{O}(1) = \frac{kd}{kd-1} \log(n) + \mathcal{O}(1).$$

3.5.4 CU-Modell

Hinzugefügter Knoten: Bei CU-Modell müssen die k -Bäume wieder an Hand der Möglichkeiten des Aufbaus unterschieden werden. Wenn der Knoten $n-1$ beim k -Baum angehängt wird, ist er Teil von einer $(k+1)$ -Clique und damit von k k -Cliquen. Es ergeben sich deshalb k Möglichkeiten, den n -ten Knoten bei einer Clique, welche $n-1$ umfasst, beim k -Baum hinzuzufügen. Dies führt für die Anzahl $t_{n,m}$ der k -Bäume des CU-Modells, bei dem der n -te Knoten m Vorfahren hat, zu einer Rekursion:

$$\begin{aligned}
t_{n,m} &= kt_{n-1,m-1} + ((n-1)(k+1) - k)t_{n-1,m} & n \geq 2, m \geq 1. \\
&= kt_{n-1,m-1} + ((n-2)(k+1) + 1)t_{n-1,m},
\end{aligned}$$

Die Startwerte sind $t_{1,k+1} = 1$ und $t_{1,m} = 0$ für $k+1 \neq m \in \mathbb{N}$. Aufsummieren der Rekursion ergibt die Differentialgleichung

$$(1 - (k+1)z)T_z - (kv+1)T = 0, \quad T(0, v) = v^{k+1},$$

wobei $T(z, v) = \sum_{n \geq 1} \sum_{m \geq 1} t_{n,m} \frac{z^{n-1}}{(n-1)!} v^m$ die erzeugende Funktion ist. Auf die gleiche Weise wie bei den geordneten ansteigenden k -Bäumen kann diese Gleichung mit folgendem Ergebnis gelöst werden:

$$T(z, v) = \frac{v^{k+1}}{(1 - (k+1)z)^{\frac{kv+1}{k+1}}}.$$

Die Zufallsvariable A_n der Anzahl der Vorfahren in einem k -Baum unter der Verteilung des CU-Modells nimmt nur größere Werte als $k+1$ an und für die Verteilung von A_n gilt:

$$\mathbb{P}\{A_n = m+k\} = \frac{t_{n,m+k}}{t_n} = \frac{(n-1)!}{(n-1)!(k+1)^{n-1}} [z^{n-1}v^{m-1}] \frac{1}{(1 - (k+1)z)^{\frac{kv+1}{k+1}}}.$$

Die Koeffizienten der bivariaten Funktion sind:

$$\begin{aligned}
[z^{n-1}v^{m+k}]T(z, v) &= [z^{n-1}v^{m-1}] \frac{1}{(1 - (k+1)z)^{\frac{kv+1}{k+1}}} \\
&= \left(\frac{k}{k+1}\right)^{m-1} \sum_{i \geq m-1} \frac{\lfloor m-1 \rfloor}{i!} (k+1)^i [z^{n-i-1}] \frac{1}{(1 - (k+1)z)^{\frac{1}{k+1}}} \\
&= \left(\frac{k}{k+1}\right)^{m-1} (k+1)^{n-1} \sum_{i=m-1}^{n-1} \frac{\lfloor m-1 \rfloor}{i!} (-1)^{n-i-1} \binom{-\frac{1}{k+1}}{n-i-1} \\
&= \left(\frac{k}{k+1}\right)^{m-1} (k+1)^{n-1} \sum_{i=m}^n \frac{\lfloor m-1 \rfloor}{(i-1)!} \binom{n-i-2 + \frac{1}{k+1}}{n-i},
\end{aligned}$$

womit die Verteilungsfunktion mit

$$\begin{aligned}
\mathbb{P}\{A_n = m+k\} &= \frac{1}{(k+1)^{n-1}} \left(\frac{k}{k+1}\right)^{m-1} (k+1)^{n-1} \sum_{i=m}^n \frac{\lfloor m-1 \rfloor}{(i-1)!} \binom{n-i-2 + \frac{1}{k+1}}{n-i} \\
&= \left(\frac{k}{k+1}\right)^{m-1} \sum_{i=m}^n \frac{\lfloor m-1 \rfloor}{(i-1)!} \binom{n-i-2 + \frac{1}{k+1}}{n-i}
\end{aligned}$$

übereinstimmt.

Beliebig hinzugefügter Knoten: Für die wahrscheinlichkeitserzeugende Funktion $P_n(v) = \sum_{m \geq 0} \mathbb{P}\{A_n = m\}v^m$ von A_n erhalten wir:

$$\begin{aligned}
P_n(v) &= \frac{(n-1)!}{t_n} [z^{n-1}]T(z, v) = \frac{(n-1)!}{t_n} v^{k+1} (k+1)^{n-1} (-1)^{n-1} \binom{-\frac{kv+1}{k+1}}{n-1} \\
&= v^{k+1} \binom{n-2 + \frac{kv+1}{k+1}}{n-1}.
\end{aligned}$$

Sei \bar{A}_n die Zufallsvariable der Vorfahren eines beliebigen eingefügten Knotens, wenn die k -Bäume nach dem CU-Modell verteilt sind. So ist $\mathbb{P}\{\bar{A}_n = m\} = \frac{1}{n} \sum_{j=1}^n \mathbb{P}\{A_n = m\}$ beziehungsweise für die wahrscheinlichkeitserzeugende Funktion $\bar{P}_n(v) = \sum_{m \geq 0} \mathbb{P}\{\bar{A}_n = m\}v^m$:

$$\begin{aligned}
\bar{P}_n(v) &= \frac{1}{n} \sum_{j=1}^n P_j(v) = \frac{v^{k+1}}{n} \sum_{j=1}^n \binom{j-2 + \frac{kv+1}{k+1}}{j-1} \\
&= \frac{v^{k+1}}{n} \sum_{j=0}^{n-1} \binom{j-1 + \frac{kv+1}{k+1}}{j},
\end{aligned}$$

was sich wegen $\binom{j+\alpha+1}{j+1} = \binom{j+\alpha}{j} + \binom{j+\alpha}{j+1}$ zu

$$\bar{P}_n(v) = \frac{v^{k+1}}{n} \binom{n-1 + \frac{kv+1}{k+1}}{n-1}$$

umformen lässt. Um Hwang's Satz anzuwenden, müssen wir noch das asymptotische Verhalten bezüglich n betrachten. Für $v \in \mathbb{C}$ mit $|v - 1| \leq \epsilon$ erhalten wir die bezüglich v gleichmäßige Abschätzung:

$$\bar{P}_n(v) = \frac{v^{k+1}}{n\Gamma\left(\frac{kv+1}{k+1}\right)} n^{\frac{kv+1}{k+1}} (1 + \mathcal{O}(n^{-1})),$$

womit wir Hwang's Theorem 1.5 anwenden können. Damit erhalten wir für den Erwartungswert:

$$\mathbb{E}(\bar{A}_n) = \log(n) \frac{d}{dv} \exp\left(\frac{kv+1}{k+1} - 1\right) \Big|_{v=1} + \mathcal{O}(1) = \frac{k}{k+1} \log(n) + \mathcal{O}(1).$$

3.6 Vergleich der Modelle

Wie schon in der Einleitung erwähnt, erfüllen alle vier Modelle ein Potenzgesetz für den Knotengrad. Dieser ist von k abhängig und befindet sich, abgesehen für die d -ären k -Bäume, immer zwischen 2 und 3. Im Detail sind diese in aufsteigender Reihenfolge:

- geordnete ansteigende k -Bäume und CU-Modell: $2 + \frac{1}{k}$,
- ansteigende k -Bäume: $2 + \frac{1}{k-1}$,
- ansteigende d -äre k -Bäume: $2 + \frac{d+2}{(k-1)d-1}$.

Diese Ergebnisse sind auch im Einklang mit den in Abschnitt 3.1 angestellten Überlegungen. Bei den geordneten ansteigenden k -Bäumen und dem CU-Modell werden die Knoten mit hohem Grad präferiert, sodass es mehr Ausreißer mit großem Grad gibt. Die anderen zwei Modelle machen dies nicht. Bei den d -ären k -Bäumen werden sogar die Cliques mit wenigen Kindern bevorzugt. Dies drückt sich auch im Exponenten aus.

Für $d \rightarrow \infty$ nähern sich die ansteigenden d -ären k -Bäume den ansteigenden k -Bäumen an, was sich unter anderem auch beim Koeffizienten des Potenzgesetzes oder in der Verteilung für die Nachfahren manifestiert. Dies ist interessant, da bei den d -ären k -Bäumen im Vergleich zu den ansteigenden k -Bäumen die Plätze geordnet sind. So würde man vermuten, dass sie eher Ähnlichkeit mit den geordneten ansteigenden k -Bäumen aufweisen. Wenn man sich jedoch die Wahrscheinlichkeitsverteilung der randomisierten Modelle ansieht, erkennt man, dass für $d \rightarrow \infty$ quasi keine Beschränkung der Kinder vorliegt beziehungsweise überall gleich viele freie Plätze pro Clique sind. Damit ergibt sich eine uniforme Verteilung wie bei den ansteigenden k -Bäumen.

In Abbildung 17 haben wir schon gesehen, dass das CU-Modell und die geordneten, ansteigenden k -Bäume eine sehr ähnliche Verteilung beim Hinzufügen des nächsten Knotens liefern. Dies spiegelt sich auch in den Ergebnissen wider. So hat beispielsweise der Erwartungswert der Vorfahren asymptotisch die gleiche Form, sowie auch die Grenzverteilung des Knotengrades.

Literatur

- [1] ARNBORG, STEFAN UND PROSKUROWSKI, ANDRZEJ: *Linear Time Algorithms for NP-hard Problems restricted to partial k-Trees*, Discrete Applied Mathematics Vol. 23 Is. 1 (1989), S. 11-24.
- [2] BARABASI, A.-L. UND ALBERT, R.: *Emergence of Scaling in Random Networks*, Science Vol. 286 (1999), S. 509-512.
- [3] BEINEKE, LOWELL W. UND PIPPERT, RAYMOND E.: *The Number of Labeled k-Dimensional Trees*, Journal of Combinatorial Theory 6 (1969), S. 200-205.
- [4] *Burnside Lemma*, Encyclopedia of Mathematics http://www.encyclopediaofmath.org/index.php?title=Burnside_Lemma&oldid=35303 (16.1.2018).
- [5] BURNSIDE, WILLIAM: *Theory of groups of finite order*, Cambridge University Press (1897), Online verfügbar unter <https://archive.org/download/117768378/117768378.pdf> (16.1.2018).
- [6] COOPER, COLIN UND UEHARA, RYUHEI: *Scale free properties of random k-trees*, Mathematics in Computer Science Vol 3. Is. 4 (2010), S. 489-496.
- [7] DIESTEL, REINHARD: *Graphentheorie*, Springer-Verlag, 3. Auflage, Berlin Heidelberg (2006).
- [8] FLAJOLET, PHILIPPE UND SEDGEWICK, ROBERT: *Analytic Combinatorics*, Cambridge University Press (2009).
- [9] FOATA, DOMINIQUE: *Enumerating k-Trees*, Discrete Mathematics Vol. 1 Is. 2 (1971), S. 181-186.
- [10] FOWLER, T., GESSEL, I., LABELLE, G. UND LEROUX, P.: *The Specification of 2-trees*, Advances in Applied Mathematics Vol. 28 Is. 2 (2002), S. 145-168.
- [11] GRAHAM, R.L., KNUTH, D.E. UND PATASHNIK, O.: *Concrete Mathematics*, Addison-Wesley Publishing Company, Reading [u.a.] (1989).
- [12] KRUMKE, SVEN O. UND NOLTEMEIER, HARTMUT: *Graphentheoretische Konzepte und Algorithmen* (1. Aufl.), Teubner Verlag, Wiesbaden (2005).
- [13] LABELLE, G., LAMATHE, C. UND LEROUX, P.: *A classification of plane and planar 2-trees*, Theoretical Computer Science Vol. 307 Is. 2 (2003), S. 337-363.
- [14] LABELLE, G., LAMATHE, C. UND LEROUX, P.: *Labelled and unlabelled enumeration of k-gonal 2-trees*, Journal of Combinatorial Theory, Series A Vol. 106 Is. 2 (2004), S. 193-219.
- [15] MARKENZON, L., JUSTEL, C.M. UND PACIORNIK, N.: *Subclasses of k-trees: Characterization and recognition*, Discrete Applied Mathematics Vol. 154 Is. 5 (2006), S. 818-825.

- [16] MLOTKOWSKI, WOJCIECH: *Fuss-Catalan numbers in noncommutative probability*, Documenta Mathematica Vol. 15 (2010), S. 939-955.
- [17] MOON, JOHN W.: *The Number of Labeled k -Trees*, Journal of Combinatorial Theory 6 (1969), S. 196-199.
- [18] PANHOLZER, ALOIS UND SEITZ, GEORG: *Ancestors and Descendants in evolving k -Tree Models*, Random Structures & Algorithms Vol. 44 Is. 4 (2014), S. 465-489.
- [19] PANHOLZER, ALOIS UND SEITZ, GEORG: *Ordered increasing k -trees: Introduction and analysis of a preferential attachment network model*, Discrete Mathematics & Theoretical Computer Science, DMTCS Proceedings Vol. AM (2010).
- [20] PETKOVSEK, M., WILF, H.S., ZEILBERGER, D.: *A=B*, A K Peters Ltd., Wellesley (1996). Online verfügbar unter <https://www.math.upenn.edu/~wilf/Downld.html>.
- [21] PIKHURKO, OLEG: *Enumeration of Labelled (k, m) -Trees*, Discrete Applied Mathematics Vol. 23 Is. 1 (1989), S. 11-24.
- [22] ROSE, DONALD J.: *On simple Characterizations of k -Trees*, Discrete Mathematics Vol. 7 Is. 3-4 (1974), S. 317-322.
- [23] ROSE, DONALD J.: *Triangulated Graphs and the Elimination Process*, Journal of Mathematical Analysis and Applications Vol. 32 Is. 3 (1970), S. 597-609.
- [24] SYLOW, LUDVIG UND LIE SOPHUS: *Oeuvres complètes de Niels Henrik Abel*, Christiania Imprimerie De Grondahl & Son (1881), Online verfügbar unter: <http://www.abelprize.no/c54178/artikkel/vis.html?tid=54181> (17.12.2017).
- [25] VAN DER HOFSTAD, REMCO: *Random Graphs and Complex Networks: Volume 1*, Cambridge University Press (2016), Onlineversion verfügbar unter: <http://www.win.tue.nl/~rhofstad/NotesRGCN.pdf> (3.3.2018).
- [26] WIKIPEDIA: *Binomialkoeffizient*, <https://de.wikipedia.org/wiki/Binomialkoeffizient> (11.2.2018).
- [27] WILF, HERBERT S.: *generatingfunctionology*, Academic Press, Inc., 2. Auflage (1994), Onlineversion verfügbar unter: <https://www.math.upenn.edu/~wilf/gfology2.pdf> (15.1.2018).

A Resultate

A.1 Abzählprobleme

Bei den markierten k -Bäumen werden alle Knoten gezählt und bei den übrigen k -Bäumen mit einer Wurzel werden nur die Knoten außerhalb dieser beim Zählen berücksichtigt. Für $n \in \mathbb{N}$ gibt es t_n viele k -Bäume der Größe n :

- markierte k -Bäume:

$$t_n = \binom{n}{k} (k(n-k) + 1)^{n-k-2},$$

- markierte k -Wurzelbäume

$$t_n = (kn + 1)^{n-1},$$

- ansteigende k -Bäume:

$$t_n = n! k^n \binom{n - \frac{k-1}{k}}{n},$$

- geordnete k -Wurzelbäume:

$$t_n = \frac{((k+1)n)!}{(kn+1)!}, \quad n \geq 1,$$

- geordnete ansteigende k -Bäume:

$$t_n = n!(k+1)^n \binom{n - \frac{k}{k+1}}{n},$$

- markierte d -äre k -Bäume:

$$t_n = (n-1)! d \binom{ndk + d - 1}{n-1}, \quad n \geq 1,$$

- ansteigende d -äre k -Bäume:

$$t_n = n!(kd-1)^n \binom{n - 1 + \frac{d}{kd-1}}{n}.$$

A.2 Knotengrad

Die Verteilung der diskreten Zufallsvariable $Y_{n,j}$ des ausgehenden Knotengrades eines hinzugefügten Knotens $j \geq 1$ lautet für $m \geq 0$:

- geordnete ansteigende k -Bäume:

$$\mathbb{P}\{Y_{n,j} = m\} = \frac{\binom{j - \frac{k}{k+1}}{j}}{\binom{n}{j} \binom{n - \frac{k}{k+1}}{n}} \sum_{l=0}^m \binom{m}{l} (-1)^l \binom{n - \frac{k(l+2)}{k+1}}{n - j},$$

- ansteigende k -Bäume:

$$\mathbb{P}\{Y_{n,j} = m\} = \frac{\binom{j - \frac{k-1}{k}}{j} \binom{m + \frac{1}{k-1}}{m}}{\binom{n}{j} \binom{n - \frac{k-1}{k}}{n}} \sum_{l=0}^m \binom{m}{l} (-1)^l \binom{n - 2 - \frac{l(k-1)-1}{k}}{n - j},$$

- ansteigende d -äre k -Bäume:

$$\mathbb{P}\{Y_{n,j} = m\} = \frac{\binom{j - 1 + \frac{d}{kd-1}}{j} \binom{m - 1 + \frac{kd}{(k-1)d-1}}{m}}{\binom{n}{j} \binom{n - 1 + \frac{d}{kd-1}}{n}} \sum_{l=0}^m \binom{m}{l} (-1)^l \binom{n - \frac{(l+1)((k-1)d-1) + kd}{kd-1}}{n - j},$$

- CU-Modell:

$$\mathbb{P}\{Y_{n,j} = m\} = \frac{1}{\binom{n-1}{j-1}} \sum_{l=0}^m \binom{m}{l} (-1)^l \binom{n - \frac{(l+2)k+1}{k+1}}{n - j}.$$

Für die Grenzverteilung eines unter der Gleichverteilung hinzugefügten Knotens ergibt sich:

- geordnete ansteigende k -Bäume und CU-Modell:

$$\mathbb{P}\{\bar{Y} = m\} = \frac{k+1}{k(m+1) \binom{m+2+\frac{1}{k}}{m+1}},$$

- ansteigende k -Bäume:

$$\mathbb{P}\{\bar{Y} = m\} = \frac{k}{k-1} \frac{\binom{m + \frac{1}{k-1}}{m}}{(m+1) \binom{m + \frac{2k}{k-1}}{m+1}},$$

- ansteigende d -äre k -Bäume:

$$\mathbb{P}\{\bar{Y} = m\} = \frac{\binom{m-1 + \frac{kd}{(k-1)d-1}}{m} (kd-1)}{(m + \frac{2kd+1}{(k-1)d-1}) \binom{m-1 + \frac{2kd+1}{(k-1)d-1}}{m}}.$$

Die Grenzverteilungen erfüllen ein Potenzgesetz mit den Exponenten

- geordnete ansteigende k -Bäume und CU-Modell: $2 + \frac{1}{k}$,
- ansteigende k -Bäume: $2 + \frac{1}{k-1}$,
- ansteigende d -äre k -Bäume: $2 + \frac{d+2}{(k-1)d-1}$.

A.3 Lokaler Clusterkoeffizient

Der Erwartungswert des lokalen Clusterkoeffizienten eines beliebigen ausgewählten Knotens konvergiert:

- geordnete ansteigende k -Bäume und CU-Modell:

$$\lim_{n \rightarrow \infty} \mathbb{E}(\tilde{C}_n) = \sum_{m \geq 0} \frac{(k+1)(k-1)}{k(m+1)(m+k) \binom{m+2+\frac{1}{k}}{m+1}} \left(2 - \frac{k-2}{m+k+1} \right),$$

- ansteigenden k -Bäume:

$$\lim_{n \rightarrow \infty} \mathbb{E}(\tilde{C}_n) = \sum_{m \geq 0} \frac{k \binom{m+\frac{1}{k-1}}{m}}{(m+1)(m+k) \binom{m+\frac{2k}{k-1}}{m+1}} \left(2 - \frac{k-2}{m+k+1} \right),$$

- ansteigende d -äre k -Bäume:

$$\lim_{n \rightarrow \infty} \mathbb{E}(\tilde{C}_n) = \sum_{m \geq 0} \frac{(k-1)(kd-1) \binom{m-1+\frac{kd}{(k-1)d-1}}{m}}{\left((m+k) \binom{m+\frac{2kd+1}{(k-1)d-1}}{m} \right) \binom{m+1+\frac{2kd-1}{(k-1)d-1}}{m}} \left(2 - \frac{k-2}{m+k+1} \right).$$

A.4 Nachfahren

Die Verteilung der diskreten Zufallsvariable $D_{n,j}$ der Anzahl der Nachfahren eines hinzugefügten Knotens $j \geq 1$ stellt sich für $m \geq 1$ dar durch:

- geordnete ansteigende k -Bäume:

$$\mathbb{P}\{D_{n,j} = m\} = \frac{\binom{m-1-\frac{1}{k-1}}{m-1} \binom{n-m-1+\frac{2}{k+1}}{n-m-j+1}}{\binom{n-1+\frac{1}{k+1}}{n-j}},$$

- ansteigende k -Bäume:

$$\mathbb{P}\{D_{n,j} = m\} = \frac{\binom{n-m-1+\frac{1}{k}}{n-m-j+1}}{\binom{n-1+\frac{1}{k}}{n-j}},$$

- ansteigende d -äre k -Bäume:

$$\mathbb{P}\{D_{n,j} = m\} = \frac{\binom{m-1+\frac{1}{kd-1}}{m-1} \binom{n-m-1+\frac{d-1}{kd-1}}{n-j-m+1}}{\binom{n-1+\frac{d}{kd-1}}{n-j}},$$

- CU-Modell:

$$\mathbb{P}\{D_{n,j} = m\} = \frac{\binom{m-1-\frac{1}{k+1}}{m-1} \binom{n-m-1-\frac{k}{k+1}}{n-m-j+1}}{\binom{n-1}{j-1}}.$$

Die Verteilungsfunktion für einen beliebigen hinzugefügten Knoten konvergiert gegen:

- geordnete ansteigende k -Bäume und CU-Modell:

$$\mathbb{P}\{\bar{D} = m\} = \frac{k}{(k+1) \left(m + \frac{k}{k+1}\right) \left(m - \frac{1}{k+1}\right)},$$

- ansteigende k -Bäume:

$$\mathbb{P}\{\bar{D} = m\} = \frac{1}{m(m+1)},$$

- ansteigende d -äre k -Bäume:

$$\mathbb{P}\{\bar{D} = m\} = \frac{kd}{(kd-1) \left(m + \frac{kd}{kd-1}\right) \left(m + \frac{1}{kd-1}\right)}.$$

A.5 Vorfahren

Für die Anzahl der Vorfahren eines hinzugefügten Knotens $n \geq 1$ ergibt sich für $m \geq 0$:

- geordnete ansteigende k -Bäume:

$$\mathbb{P}\{A_n = m+k\} = \frac{k^{m-1}}{(k+1)^{m-1} \binom{n-1+\frac{1}{k+1}}{n-1}} \sum_{i=m}^n \frac{\left[\begin{smallmatrix} i-1 \\ m-1 \end{smallmatrix} \right]}{(i-1)!} \binom{n-i-2+\frac{2}{k+1}}{n-i},$$

- ansteigende k -Bäume:

$$\mathbb{P}\{A_n = m+k\} = \frac{1}{\binom{n-1+\frac{1}{k}}{n-1}} \sum_{i=m}^n \frac{\left[\begin{smallmatrix} i-1 \\ m-1 \end{smallmatrix} \right]}{(i-1)!} \binom{n-i-1+\frac{1}{k}}{n-i},$$

- ansteigende d -äre k -Bäume:

$$\mathbb{P}\{A_n = m + k\} = \frac{(kd)^{m-1}}{(kd-1)^{m-1} \binom{n-1+\frac{d}{kd-1}}{n-1}} \sum_{i=m}^n \frac{\lfloor \frac{i-1}{m-1} \rfloor}{(i-1)!} \binom{n-i-1+\frac{d-1}{kd-1}}{n-i},$$

- CU-Modell:

$$\mathbb{P}\{A_n = m + k\} = \left(\frac{k}{k+1}\right)^{m-1} \sum_{i=m}^n \frac{\lfloor \frac{i-1}{m-1} \rfloor}{(i-1)!} \binom{n-i-2+\frac{1}{k+1}}{n-i}.$$

Die Verteilungsfunktion eines zufällig ausgewählten Knotens ist für alle Modelle asymptotisch normalverteilt mit:

$$\mathbb{P}\left\{\frac{\bar{A}_n - \mathbb{E}(\bar{A}_n)}{\mathbb{V}(\bar{A}_n)} \leq x\right\} = \phi(x) + \mathcal{O}\left(\frac{1}{\sqrt{\log n}}\right),$$

wobei für die Erwartungswerte gilt:

- geordnete ansteigende k -Bäume und CU-Modell: $\mathbb{E}(\bar{A}_n) = \frac{k}{k+1} \log(n) + \mathcal{O}(1)$,
- ansteigende k -Bäume: $\mathbb{E}(\bar{A}_n) = \log(n) + \mathcal{O}(1)$,
- ansteigende d -äre k -Bäume: $\mathbb{E}(\bar{A}_n) = \frac{kd}{kd-1} \log(n) + \mathcal{O}(1)$.