Combinatorics of oriented trees and tree-like structures

by

Isaac Owino Okoth

Dissertation approved for the degree of Doctor of Philosophy in Mathematics at Stellenbosch University

Department of Mathematical Sciences,
University of Stellenbosch,
Private Bag X1, Matieland 7602, South Africa.

Promoter: Prof. Stephan Wagner

March 2015

Declaration

By submitting this dissertation electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Signature: ....
Isaac Owino Okoth

Date: February 17, 2015

Abstract

Combinatorics of oriented trees and tree-like structures

Isaac Owino Okoth

Department of Mathematical Sciences,
University of Stellenbosch,
Private Bag X1, Matieland 7602, South Africa.

Dissertation: PhD (Mathematics)

March 2015

In this thesis, a number of combinatorial objects are enumerated. Du and Yin as well as Shin and Zeng (by a different approach) proved an elegant formula for the number of labelled trees with respect to a given indegree sequence, where each edge is oriented from a vertex of lower label towards a vertex of higher label. We refine their result to also take the number of sources (vertices of indegree 0) or sinks (vertices of outdegree 0) into account. We find formulas for the mean and variance of the number of sinks or sources in these trees. We also obtain a differential equation and a functional equation satisfied by the generating function for these trees. Analogous results for labelled trees with two marked vertices, related to functional digraphs, are also established.

We extend the work to count reachable vertices, sinks and leaf sinks in these trees. Among other results, we obtain a counting formula for the number of labelled trees on n vertices in which exactly k vertices are reachable from a given vertex v and also the average number of vertices that are reachable from a specified vertex in labelled trees of order n.

In this dissertation, we also enumerate certain families of set partitions and related tree-like structures. We provide a proof for a formula that counts connected cycle-free families of k set partitions of $\{1, \ldots, n\}$ satisfying a certain coherence condition and then establish a bijection between these families and the set of labelled free k-ary cacti with a given vertex-degree distribution. We then show that the formula also counts colou-red Husimi graphs in which there are no blocks of the same colour that are incident to one another. We extend the work to count coloured oriented cacti and coloured cacti.

Noncrossing trees and related tree-like structures are also considered in this thesis. Specifically, we establish formulas for locally oriented noncrossing trees with a given number of sources and sinks, and also with given indegree and outdegree sequences. The work is extended to obtain the average number of reachable vertices in these trees. We then generalise the concept of noncrossing trees to find formulas for the number of noncrossing Husimi graphs, cacti and oriented cacti. The study is further extended to find formulas for the number of bicoloured noncrossing Husimi graphs and the number of noncrossing connected cycle-free pairs of set partitions.

Opsomming

Kombinatorika van georiënteerde bome en boom-vormige strukture

("Combinatorics of oriented trees and tree-like structures")

Isaac Owino Okoth

Departement Wiskundige Wetenskappe,
Universiteit van Stellenbosch,
Privaatsak X1, Matieland 7602, Suid Afrika.

Proefskrif: PhD (Wiskunde)

Maart 2015

In hierdie tesis word 'n aantal kombinatoriese objekte geenumereer. Du en Yin asook Shin en Zeng (deur middel van 'n ander benadering) het 'n elegante formule vir die aantal geëtiketteerde bome met betrekking tot 'n gegewe ingangsgraadry, waar elke lyn van die nodus met die kleiner etiket na die nodus met die groter etiket toe georiënteer word. Ons verfyn hul resultaat deur ook die aantal bronne (nodusse met ingangsgraad 0) en putte (nodusse met uitgangsgraad 0) in ag te neem. Ons vind formules vir die gemiddelde en variansie van die aantal putte of bronne in hierdie bome. Ons bepaal verder 'n differensiaalvergelyking en 'n funksionaalvergelyking wat deur die voortbringende funksie van hierdie bome bevredig word. Analoë resultate vir geëtiketteerde bome met twee gemerkte nodusse (wat verwant is aan funksionele digrafieke), is ook gevind.

Ons gaan verder voort deur ook bereikbare nodusse, bronne en putte in hierdie bome at te tel. Onder andere verkry ons 'n formule vir die aantal geëtiketteerde bome met n nodusse waarin presies k nodusse vanaf 'n gegewe nodus v bereikbaar is asook die gemiddelde aantal nodusse wat bereikbaar is vanaf 'n gegewe nodus.

Ons enumereer in hierdie tesis verder sekere families van versamelingsverdelings en soortgelyke boom-vormige strukture. Ons gee 'n bewys vir 'n formule wat die aantal van samehangende siklus-vrye families van k versamelingsverdelings op $\{1, \ldots, n\}$ wat 'n sekere koherensie-vereiste bevredig, en ons beskryf 'n bijeksie tussen hierdie familie en die versameling van geëtiketteerde vrye k-êre kaktusse met 'n gegewe nodus-graad-verdeling. Ons toon ook dat hierdie formule ook gekleurde Husimi-grafieke tel waar blokke van dieselfde kleur nie insident met mekaar mag wees nie. Ons tel verder ook gekleurde georiënteerde kaktusse en gekleurde kaktusse.

Nie-kruisende bome en soortgelyke boom-vormige strukture word in hierdie tesis ook beskou. On bepaal spesifiek formules vir lokaal georiënteerde nie-kruisende bome wat 'n gegewe aantal bronne en putte het asook nie-kruisende bome met gegewe ingangs- en uitgangsgraadrye. Ons gaan voort deur die gemiddelde aantal bereikbare nodusse in hierdie bome te bepaal. Ons veralgemeen dan die konsep van nie-kruisende bome en vind formules vir die aantal nie-kruisende Husimi-grafieke, kaktusse en georiënteerde kaktusse. Laastens vind ons 'n formule vir die aantaal tweegekleurde nie-kruisende Husimi-grafieke en die aantal nie-kruisende samehangende siklus-vrye pare van versamelingsverdelings.

Acknowledgements

I would like to express my special gratitude to:

- ◇ My advisor Prof. Stephan Wagner for his generosity in sharing ideas and insights. It was an honour to work with him. It was also kind of him to accept to translate the abstract of this dissertation into Afrikaans.
- ◇ Professors Alois Panholzer (Vienna University of Technology), Augustine Munagi (University of Witwatersrand) and Helmut Prodinger (University of Stellenbosch) for accepting to serve on my jury and for their valuable comments on the previous version of this thesis.
- ◇ German Academic Exchange Service (DAAD) and African Institute for Mathematical Sciences (AIMS) for financing this project.
- ◇ Mathematics division of Stellenbosch University for hosting my PhD studies.
- ◇ Maseno University for granting me a study leave with pay. I heartily thank Prof. John Ogonji Agure who tapped me, after my first degree, and encouraged me to pursue postgraduate studies in mathematics.
- ◇ My parents Domnick Okoth, Caren Okoth and Lewnida Okoth for their unrivalled support. *Erokamano ahinya.*
- ◇ My wife Everlyne Owino and son Neil Owino for their moral support and boundless love. My siblings: Francis, Paul, Selline, Millicent and Beatrice, as well as cousins: Kennedy, Mike, Molly, Roseline, Felix and Auma, for their never ending support and love. *Asanteni sana.*
- ◇ My extended families and friends especially Valisoa, Ronalda and Titus for the company.

My heartfelt appreciation goes to the almighty God for the gift of life and privilege of education. Thank you Lord for seeing me this far.

Dedications

To my family

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Enumerative combinatorics is a branch of combinatorics which concentrates on establishing the number of certain combinatorial objects with a given set of parameters. It is considered by many as the most classical area of combinatorics. In this dissertation, quite a number of combinatorial objects are studied. These include labelled trees, noncrossing trees, cacti, Husimi graphs and set partitions. We utilise the concepts of generating functions, recurrence equations and functional equations to derive explicit formulas for the number of these objects, and in many instances we provide bijective proofs.

The class of labelled trees in which the edges are oriented from a vertex of lower label towards a vertex of higher label (called a *local orientation* [17]) is considered in Chapter 3. We prove a formula for the number of labelled trees on n vertices with a given local indegree sequence or number of sources such that a given vertex r is a sink of degree d. We find a formula for the average number of sinks in these trees as a corollary of our main theorem and after deriving a differential equation, recurrence equation and functional equation satisfied by the generating function for these trees in Section 3.3, we rediscover the formula for the mean number of sinks as well as the variance. Analogous results for labelled trees with two marked vertices are also given in Section 3.4.

Next, we introduce the idea of *reachability* of vertices in these trees. A vertex v is said to be *reachable* from a vertex u if there is an oriented path from vertex u to vertex v. In Chapter 4, among many other interesting counting formulas, we obtain the average number of vertices that can be reached from a given vertex u and the number of trees on n vertices such that exactly

k vertices are reachable from vertex u.

The main objects of study in Chapter 5 are cycle-free set partition families, a concept that is motivated by spanning tree enumeration as well as cycle factorisations of permutations. Here, we provide a new proof for a formula that counts the number of connected cycle-free families of k set partitions of $[n] := \{1, 2, \ldots, n\}$ satisfying a certain coherence condition and then establish a bijection between these families and the set of labelled free k-ary cacti with a given vertex-degree distribution. We also show that the formula counts coloured Husimi graphs in which there are no blocks (complete graphs) of the same colour that are incident to one another. We extend the work to coloured oriented cacti and coloured cacti.

Noncrossing trees, noncrossing graphs with a given number of vertices and edges, noncrossing forests, noncrossing unicyclic graphs etc. have been extensively studied in literature, see [15, 18, 22, 41, 42, 43] among many other papers. Husimi graphs were introduced by Japanese physicist Kodi Husimi in 1950. These are connected graphs whose blocks are complete graphs. Surprisingly, the number of noncrossing Husimi graphs had not been considered. In Chapter 6, we find formulas for the number of noncrossing Husimi graphs, cacti and oriented cacti. We obtain a formula for the number of bicoloured noncrossing Husimi graphs and also the number of noncrossing connected cycle-free pairs of set partitions.

Flajolet and Noy [22] enumerated noncrossing trees with a given outdegree sequence where all edges are oriented away from the root, i.e. the edges are *globally oriented*. In Chapter 7, we find a formula for the number of locally oriented noncrossing trees with given number of sources and sinks. As a corollary, we obtain the mean and variance of the number of sinks in these trees. We then extend the study to find a formula for the number of these trees, where both indegree and outdegree sequences are simultaneously given.

In Chapter 8, we extend the concept of reachability of vertices in labelled trees, introduced and studied extensively in Chapter 4, to locally oriented noncrossing trees. A formula for the number of locally oriented noncrossing trees of order n in which exactly k vertices are reachable from vertex 1 deserves to be mentioned in particular. This formula generalises a result for the number of noncrossing increasing trees on n vertices, counted by the well-known Catalan numbers.

Chapter 2

Terminologies and basic notions

In this chapter, we introduce terminologies and preliminary results that we require in the next chapters. We do this for the sake of completeness, even though the definitions and concepts presented here are standard in texts on graph theory and enumerative combinatorics.

2.1 Graph theoretic concepts

A *graph* G is a pair $(V(G), E(G))$ where $V(G)$ is a set of vertices and $E(G)$ is a set of edges of G. The cardinalities $|V(G)|$ and $|E(G)|$ are respectively, the *order* and *size* of the graph. All graphs considered in this thesis are *simple*, i.e. there are no loops and multiple edges between any two given vertices. A graph H is said to be a *subgraph* of G if $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$.

We say that two vertices are *adjacent* if they are connected by an edge and two edges are *incident* if they share a vertex. The number of vertices in G that are adjacent to a vertex v is the *degree* of v, denoted by $\deg_G(v)$. A graph, K_n, of order n in which each vertex has degree $(n-1)$ is called a *complete graph*.

If we assign labels to the vertices of the graph, the resulting graph is called a *labelled graph*. A graph in which every vertex is reachable from every other vertex is said to be *connected*. A *tree* is a connected graph without cycles. In a tree, a vertex which has degree 1 is called a *leaf*. A labelled graph which is also a tree is called a *labelled tree*. Figure 2.1 shows a labelled tree of order 6, where vertices $1, 4$ and 6 are all leaves. A *rooted tree* is a tree in which a fixed vertex has been chosen. This fixed vertex is often referred to as the *root*.

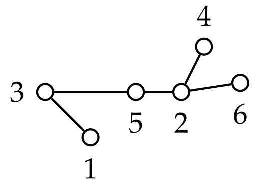

Figure 2.1: Labelled tree

A *forest* is a collection of graphs in which each component is a tree whereas a *rooted forest* is a graph whose components are rooted trees. An example of a forest is given in Figure 2.2.

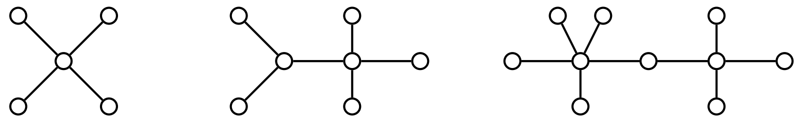

Figure 2.2: Forest

Labelled trees have been extensively studied. In 1889, Cayley [6] showed that the number of distinct spanning trees of a complete graph with vertex set $\{1, \ldots, n\}$ is given by n^{n-2} (see [39] for more details). This elegant formula is credited to Arthur Cayley though it was originally obtained by Borchardt [5] in 1860 and Sylvester [59] in 1857. Several proofs of this beautiful formula exist in the literature. Remarkable ones include Kirchhoff's matrix tree theorem [1], Joyal's bijective proof [32], Pitman's double counting argument [44] and Prüfer sequences [49] that yield a bijective proof.

Our work is motivated by the study of labelled trees of complete graphs in which edges have certain orientations. A *directed graph* (*digraph* for short) is a graph in which the edges have orientations. The number of edges that are oriented towards a vertex is the *indegree* of such a vertex. The number of edges that are oriented away from the vertex is its *outdegree*. The *indegree* (resp. *outdegree*) *sequence* of a tree is the ordered sequence of indegrees (resp. outdegrees) of the vertices of the tree. We write the indegree sequence (or outdegree sequence) of a graph in abbreviated form as $\langle a_1^{b_1} a_2^{b_2} \cdots \rangle$, where b_i is the number of vertices of indegree (outdegree) a_i and vertices of indegree (outdegree) 0 are ignored. Figure 2.3 shows a labelled tree with an indegree sequence $\langle 1^3 2^1 \rangle$, since there are three vertices (1, 3 and 5) with indegree 1 and a single vertex (6) with indegree 2. It has outdegree sequence $\langle 1^2 3^1 \rangle$, since there are two vertices (4 and 6) with outdegree 1 and one vertex (2) with outdegree 3.

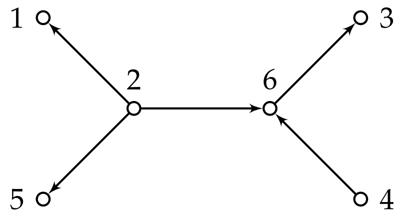

Figure 2.3: Labelled tree with oriented edges

For other definitions and notations that are not given here, we refer the reader to [16]. Further relevant definitions are also presented throughout the thesis.

2.2 Number partitions and set partitions

A *partition* of $n \in \mathbb{N}$ is a way of representing n as an unordered sum of positive integers. So $1 + 2$ and $2 + 1$ count as the same representation. We shall always sort the summands in increasing order. The partitions of 5 are

$$1+1+1+1+1, \;\; 1+1+1+2, \;\; 1+2+2, \;\; 1+1+3, \;\; 2+3, \;\; 1+4, \;\; 5.$$

Let $\mathcal{P}(n)$ be the set of number partitions of $n \in \mathbb{N}$, and let $m_k(p)$ be the number of occurrences of $k \in \mathbb{N}$ in a partition $p \in \mathcal{P}(n)$, so that

$$n = \sum_{k \in \mathbb{N}} k \cdot m_k(p)$$

and $m_k(p) = 0$ for $k > n$. Further, we define

$$|p| = \sum_{k \in \mathbb{N}} m_k(p).$$

Now, set $\mathcal{P}_i(n) = \{p \in \mathcal{P}(n) : |p| = i\}$ for $i \in \mathbb{N}$. If a number partition p of n has $k_1, k_2, \ldots, k_i$ as its distinct parts, then we write

$$p = k_1^{m_{k_1}(p)} k_2^{m_{k_2}(p)} \cdots k_i^{m_{k_i}(p)}.$$

So the number partition $1 + 1 + 1 + 2$ can also be written as $1^3 2^1$ in shorthand.

Let S be a finite set. We say that P is a *set partition* of S if P is a family of non-empty and disjoint subsets of S such that the union of these subsets gives S. The elements of S are called *blocks* or *parts* or *cells*. The block sizes of P define a number partition p of $|S|$, and the set partition P is said to be of *type* p.

As an example, consider the partition

$$P = \{\{1,7,8\}, \{2,6\}, \{3\}, \{4,5,9\}, \{10\}\}$$

of $\{1, 2, \ldots, 10\}$. The number partition associated with this set partition is given by $1 + 1 + 2 + 3 + 3$ since in P there are two subsets of cardinality 1, one subset of size 2 and two subsets of cardinality 3. This implies that the partition is of type $1^2 2^1 3^2$.

2.3 Generating functions

Generating functions are one of the most powerful and versatile tools used in enumerative combinatorics. In this thesis we make use of two kinds of generating functions.

Definition 2.3.1. Let $(g_0, g_1, \ldots)$ be a sequence of integers. The *ordinary generating function* of this sequence is

$$\sum_{i \geq 0} g_i x^i$$

and its *exponential generating function* is

$$\sum_{i \geq 0} g_i \frac{x^i}{i!}.$$

Ordinary generating functions are usually used to count unlabelled structures. On the other hand, exponential generating functions are used when one deals with labelled structures. Consider for example, the sequence $(1, 1, \ldots)$. Its ordinary generating function is

$$\sum_{i \geq 0} x^i = \frac{1}{1 - x},$$

while the corresponding exponential generating function is

$$\sum_{i\geq 0} \frac{x^i}{i!} = \exp(x).$$

It is common combinatorial practice to denote the coefficient of x^n in a generating function $g(x)$ by $[x^n]g(x)$.

Let $f(x)$ be a generating function that satisfies the functional equation $f(x) = x\phi(f(x))$. We will often use the following theorem to extract the coefficient of x^n in the generating function.

Theorem 2.3.2 (Lagrange Inversion Formula [58, Theorem 5.4.2])**.** *Let $f(x)$ be a generating function that satisfies the functional equation $f(x) = x\phi(f(x))$, where $\phi(0) \neq 0$. We have*

$$[x^n]f(x)^k = \frac{k}{n}[t^{n-k}]\phi(t)^n.$$

For more background on generating functions, the reader is referred to Wilf's *Generatingfunctionology* [63].

Chapter 3

Trees with a local orientation

3.1 Introduction

In any digraph, a vertex with indegree (outdegree) zero is called a *source* (*sink*) respectively. Labelled trees whose edges have orientations have been studied in particular in two versions; in a *global orientation* [58], all edges in a rooted labelled tree are directed towards the root as opposed to a *local orientation* [17], where each edge is oriented towards the vertex with higher label.

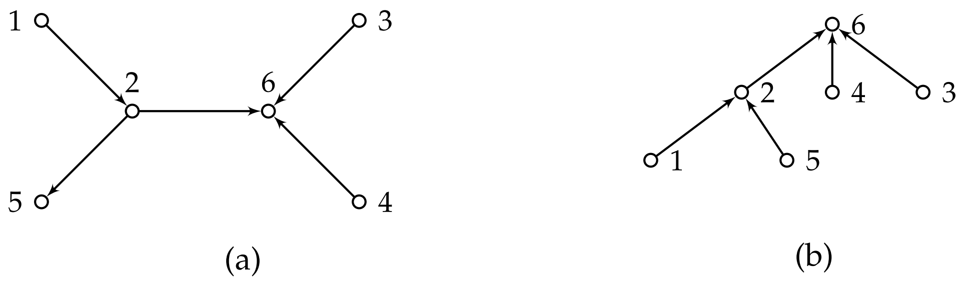

Figure 3.1: (a) Labelled tree with local orientation, (b) Labelled tree with global orientation

Trees treated in this chapter have locally oriented edges, unless stated otherwise. Recently, Du and Yin [17], Shin and Zeng [53], and Wagner [62] considered the enumeration of labelled trees, with a local orientation, with respect to their indegree sequences. Consider a labelled tree with a vertex set $[n]$ and indegree sequence $\lambda = \langle 1^{e_1} 2^{e_2} \cdots \rangle$, where e_j is the number of vertices with indegree j satisfying the conditions

$$\sum_{i\geq 0} e_i = n \text{ and } \sum_{i\geq 1} ie_i = n-1. \tag{3.1.1}$$

The said authors showed that the number of such trees is given by

$$T_\lambda = \frac{(n-1)!^2}{(n-k)!e_1!(1!)^{e_1}e_2!(2!)^{e_2}\cdots}, \tag{3.1.2}$$

where k is the number of parts of the partition λ. This result was initially conjectured by Cotterill [12] in his study of secant planes of curves in projective spaces. This expression also counts the number of labelled trees with a fixed root and global indegree sequence λ, which is a classical refinement of Cayley's formula for the number of labelled trees (see [58, Corollary 5.3.5]). We note here that Equation (3.1.2) (for the global orientation) can also be obtained by considering functional digraphs, (see [58, Section 5.3]).

In this section, we will exploit a result of Wagner [62] in which he gave a direct proof of a curious algebraic identity (pointed out by Shin and Zeng in [53]) relating a generalised form of Equation (3.1.2) and the generating function for the number of labelled trees with a given indegree sequence. To restate this result, we first define the type of a monomial, which is a partition induced by its exponents.

Example 3.1.1. Let us consider the polynomial

$$6y_2y_3y_4 + 2y_2y_4^2 + 4y_3y_4^2 + 3y_3^2y_4 + y_4^3.$$

The monomial $y_2y_3y_4$ is of type $\langle 1^3 \rangle$ since all the variables have exponent 1. Monomials $y_2y_4^2$, $y_3y_4^2$ and $y_3^2y_4$ are all of type $\langle 1^1 2^1 \rangle$ while y_4^3 is of type $\langle 3^1 \rangle$.

Theorem 3.1.2 ([62, Theorem 1]). *Let* $\lambda = \langle 1^{e_1} 2^{e_2} \cdots \rangle$ *be a partition of* $n-m+1$ *and* $e_0 = n - e_1 - e_2 - \cdots$. *The following identity holds:*

$$\sum_{\text{type } \mathbf{y}^{\mathbf{b}}=\lambda} [\mathbf{y}^{\mathbf{b}}] \prod_{i=m}^{n} (iy_i + y_{i+1} + \cdots + y_n) = \frac{n!(n-m+1)!}{e_0!(0!)^{e_0}e_1!(1!)^{e_1}e_2!(2!)^{e_2}\cdots} \tag{3.1.3}$$

where $[\mathbf{y}^{\mathbf{b}}] = [y_1^{b_1}y_2^{b_2}\cdots y_n^{b_n}]$ *is the coefficient of the monomial* $y_1^{b_1}y_2^{b_2}\cdots y_n^{b_n}$ *in the polynomial* $\prod_{i=m}^{n}(iy_i + y_{i+1} + \cdots + y_n)$.

3.2 Enumeration by sources, sinks and indegree sequences

In this section, we prove a theorem which gives the number of trees with a given indegree sequence such that a specific vertex r is a sink of degree d. This refines Equation (3.1.2). Among other corollaries, we obtain a formula for the total number of sinks in labelled trees with a given indegree sequence. We also prove an analogous result given the number of sources.

Let $\mathrm{Tree}(K_n)$ be the set of all labelled trees with labels $1, 2, \ldots, n$. We denote the indegree / outdegree of vertex i in a labelled tree $T \in \mathrm{Tree}(K_n)$ by $\mathrm{indeg}_T(i)$ / $\mathrm{outdeg}_T(i)$ respectively.

Theorem 3.2.1. *Let $N^d(n, \lambda; r)$ be the number of trees of order n with indegree sequence $\lambda = \langle 1^{e_1} 2^{e_2} \cdots \rangle$ such that vertex r is a sink of degree d. Then we have*

$$N^d(n, \lambda; r) = \frac{(r-1)!(n-2)!(n-d-1)!de_d}{(r-d-1)!e_0!(0!)^{e_0}e_1!(1!)^{e_1}e_2!(2!)^{e_2}\cdots} \tag{3.2.1}$$

where $e_0 = n - e_1 - e_2 - \cdots$.

Proof. The generating function for the number of trees with a given indegree and outdegree sequence is given by

$$\sum_{T \in \mathrm{Tree}(K_n)} \prod_{i=1}^{n} y_i^{\mathrm{indeg}_T(i)} x_i^{\mathrm{outdeg}_T(i)} = x_1 y_n \prod_{i=2}^{n-1} \left((x_1 + \cdots + x_i) y_i + x_i (y_{i+1} + \cdots + y_n) \right). \tag{3.2.2}$$

(See [36, Theorem 4] and also [51, Equation (8)] for details.) Setting $x_r = 0$ and $x_i = 1$ for all $i \neq r$, we obtain

$$\sum_{T:\, r \text{ is a sink}} \prod_{i=1}^{n} y_i^{\mathrm{indeg}_T(i)} = (r-1) y_r y_n \prod_{i=2}^{r-1} (i y_i + \cdots + y_n) \cdot \prod_{i=r+1}^{n-1} ((i-1) y_i + \cdots + y_n),$$

which is the generating function for the number of trees on n vertices such that vertex r is a sink. Since y_r only occurs as a term in the first product, we can rewrite the generating function as:

$$\sum_{T:\, r \text{ is a sink}} \prod_{i=1}^{n} y_i^{\text{indeg}_T(i)} = \frac{r-1}{n-1} \sum_{A \subseteq \{2,3,\dots,r-1\}} y_r^{|A|+1} \times \prod_{\substack{i=2 \\ i \notin A}}^{r-1} (iy_i + y_{i+1} + \dots + y_{r-1} + y_{r+1} + \dots + y_n) \times \prod_{i=r+1}^{n} ((i-1)y_i + y_{i+1} + \dots + y_n). \tag{3.2.3}$$

Setting $y_i = z_i$ if $i < r$ and $y_{i+1} = z_i$ if $i \geq r$, Equation (3.2.3) becomes

$$\begin{aligned} \sum_{T:\, r \text{ is a sink}} \prod_{i=1}^{n} y_i^{\text{indeg}_T(i)} &= \frac{r-1}{n-1} \sum_{A \subseteq \{2,3,\dots,r-1\}} y_r^{|A|+1} \cdot \prod_{\substack{i=2 \\ i \notin A}}^{r-1} (iz_i + z_{i+1} + \dots + z_{n-1}) \\ &\times \prod_{i=r}^{n-1} (iz_i + z_{i+1} + \dots + z_{n-1}) \\ &= \frac{r-1}{n-1} \sum_{A \subseteq \{2,3,\dots,r-1\}} y_r^{|A|+1} \cdot \prod_{\substack{i=2 \\ i \notin A}}^{n-1} (iz_i + z_{i+1} + \dots + z_{n-1}). \end{aligned} \tag{3.2.4}$$

Now, we are going to mimic the construction of Wagner [62], in which he proved the theorem for the case $r = n$.

We interpret the product on the right hand side of Equation (3.2.4) in terms of functions from the set $[2..n-1] = \{2, \dots, n-1\}$ to itself, with the additional requirement that $f(i) \geq i$ for all i and $i \notin A \subseteq \{2, \dots, r-1\}$. As noted in [62], the product can be written out as

$$\prod_{\substack{i=2 \\ i \notin A}}^{n-1} (iz_i + z_{i+1} + \dots + z_{n-1}) = \sum_{g:[2..n-1] \setminus A \longrightarrow [n-1]} \prod_{i=2}^{n-1} z_i^{|(\phi g)^{-1}(i)|}$$

where $\phi g(i) = \max(g(i), i)$. Let $\mathcal{S} = (S_1, S_2, \dots, S_m)$ be any set partition of $[2..n-1]$ whose type is λ. We therefore have $m = e_1 + e_2 + \dots = n - 1 - e_0$. So $e_0 = n - m - 1$. Let the largest element in S_i be denoted by s_i. We assume that $n - 1 = s_1 > s_2 > \dots > s_m$. Suppose that ϕg induces the fixed set partition $\mathcal{S}$ (i.e., the different nonempty preimages of ϕg are precisely $S_1, S_2, \dots, S_m$). Now, the only possible value for $\phi g(S_1)$ is $n - 1$. Hence $g(j) = n - 1$ for all the elements $j \in S_1 \setminus \{n-1\}$, but there are

$n-1$ different choices for $g(n-1) = g(s_1)$. There are $n - s_2 - 1$ choices for $\phi g(S_2)$, but also if $\phi g(S_2) = s_2$ then there are s_2 choices for $g(s_2)$. In total there are $n-2$ possibilities. Similarly, there are $n - s_3 - 2$ possible values for $\phi g(S_3)$, but if $\phi g(S_3) = s_3$, then there are s_3 choices for $g(s_3)$ so that we have $n-3$ choices altogether, etc.

Therefore, the number of possible functions g so that ϕg induces a fixed partition $\mathcal{S}$ is

$$(n-1)(n-2)\cdots(n-m) = \frac{(n-1)!}{(n-m-1)!} = \frac{(n-1)!}{e_0!}.$$

There are $\binom{r-2}{d-1}$ choices for A and

$$\binom{n-d-1}{e_1, e_2, \ldots, e_d - 1, e_{d+1}, \ldots} \cdot \frac{1}{d!^{(e_d-1)} \cdot \prod\limits_{\substack{j\geq 1 \\ j\neq d}} j!^{e_j}} = \frac{(n-d-1)!d!e_d}{e_1!(1!)^{e_1}e_2!(2!)^{e_2}\cdots}$$

choices for $\mathcal{S}$.

Since there are $\dfrac{(n-1)!}{e_0!}$ choices for g then the required equation is

$$N^d(n,\lambda;r) = \frac{r-1}{n-1} \cdot \frac{(n-1)!}{e_0!} \binom{r-2}{d-1} \frac{(n-d-1)!d!e_d}{e_1!(1!)^{e_1}e_2!(2!)^{e_2}\cdots}.$$

This completes the proof. □

Corollary 3.2.2. *Let $N(n,\lambda)$ be the number of trees of order n with indegree sequence $\lambda = \langle 1^{e_1} 2^{e_2} \cdots \rangle$. We have*

$$N(n,\lambda) = \frac{(n-1)!^2}{e_0!(0!)^{e_0}e_1!(1!)^{e_1}e_2!(2!)^{e_2}\cdots}.$$

Proof. By substituting $r = n$ in Equation (3.2.1), we obtain

$$N^d(n,\lambda;n) = \frac{(n-1)!(n-2)!de_d}{e_0!(0!)^{e_0}e_1!(1!)^{e_1}e_2!(2!)^{e_2}\cdots}.$$

Since vertex n is always a sink, all possible trees with indegree sequence λ are covered by this.

We now sum over d to obtain

$$N(n,\lambda) = \frac{(n-1)!(n-2)!}{e_0!(0!)^{e_0}e_1!(1!)^{e_1}e_2!(2!)^{e_2}\cdots} \sum_{d\geq 1} de_d.$$

By Condition (3.1.1) we obtain the desired result. This result is also given in Equation (3.1.2). □

Corollary 3.2.3. *Let $N(n,\lambda)$ be the number of trees of order n with an indegree sequence $\lambda = \langle 1^{e_1}2^{e_2}\cdots\rangle$. The total number of sinks in these trees is given by*

$$\frac{(n-2)!n!}{e_0!(0!)^{e_0}e_1!(1!)^{e_1}e_2!(2!)^{e_2}\cdots}\cdot\left(n-\sum_{d\geq 0}\frac{e_d}{d+1}\right).$$

Proof. We sum over r in Equation (3.2.1) to obtain the total number of sinks of degree d as:

$$\begin{aligned}
&\frac{(n-2)!(n-d-1)!de_d}{e_0!(0!)^{e_0}e_1!(1!)^{e_1}e_2!(2!)^{e_2}\cdots}\sum_{r=d+1}^{n}\frac{(r-1)!}{(r-d-1)!}\\
&=\frac{(n-2)!(n-d-1)!de_d}{e_0!(0!)^{e_0}e_1!(1!)^{e_1}e_2!(2!)^{e_2}\cdots}\sum_{r=d+1}^{n}\binom{r-1}{d}d!\\
&=\frac{d!(n-2)!(n-d-1)!de_d}{e_0!(0!)^{e_0}e_1!(1!)^{e_1}e_2!(2!)^{e_2}\cdots}\sum_{i=d}^{n-1}\binom{i}{d}\\
&=\frac{(n-2)!n!}{e_0!(0!)^{e_0}e_1!(1!)^{e_1}e_2!(2!)^{e_2}\cdots}\cdot\frac{de_d}{d+1}. \qquad (3.2.5)
\end{aligned}$$

The last equality follows by hockey stick identity, i.e.

$$\sum_{i=j}^{n-1}\binom{i}{j}=\binom{n}{j+1}. \qquad (3.2.6)$$

Summing over d in Equation (3.2.5), we obtain the number of sinks as

$$\frac{(n-2)!n!}{e_0!(0!)^{e_0}e_1!(1!)^{e_1}e_2!(2!)^{e_2}\cdots}\cdot\left(\sum_{d\geq 1}(d+1)\frac{e_d}{d+1}-\sum_{d\geq 1}\frac{e_d}{d+1}\right).$$

By Condition (3.1.1) we arrive at

$$\frac{(n-2)!n!}{e_0!(0!)^{e_0}e_1!(1!)^{e_1}e_2!(2!)^{e_2}\cdots}\cdot\left(n-e_0-\sum_{d\geq 1}\frac{e_d}{d+1}\right),$$

which is the required formula. □

From Corollaries 3.2.2 and 3.2.3, we obtain a formula for the mean number of sinks in labelled trees of order $n \geq 2$ with an indegree sequence $\langle 1^{e_1}2^{e_2}\cdots\rangle$ as

$$\mu=\frac{n}{n-1}\left(n-\sum_{d\geq 0}\frac{e_d}{d+1}\right),$$

where $e_0 = n - e_1 - e_2 - \cdots$.

In Theorem 3.2.4 below, we give a formula for the number of trees with a given number of sources such that a specific vertex is a sink of degree d. Its consequence, Corollary 3.2.5, generalizes a result proved by William Chen [9]: the total number of trees on n vertices having a given number of sources (see Corollary 3.2.7 below).

Theorem 3.2.4. *Let $T^d(n,k;r)$ be the number of trees of order n with k sources such that vertex r is a sink of degree d. Then,*

$$T^d(n,k;r) = \frac{(r-1)!(n-2)!}{(r-d-1)!(d-1)!k!} \left\{ {n-d-1 \atop n-k-1} \right\} \tag{3.2.7}$$

where $\left\{ {n \atop k} \right\}$ denotes Stirling numbers of the second kind.

Proof. We obtain the formula by summing over all possibilities of $e_1, e_2, \ldots$ in Theorem 3.2.1. □

If we sum over all k in Equation (3.2.7), and use the identity

$$p^\alpha = \sum_{i=1}^{p} \binom{p}{i} (p-i)! \left\{ {\alpha \atop p-i} \right\}, \tag{3.2.8}$$

we obtain a formula for the total number of trees in which vertex r is a sink of degree d, i.e.

$$\sum_k T^d(n,k;r) = (r-1) \cdot \binom{r-2}{d-1} (n-1)^{n-d-2}. \tag{3.2.9}$$

Corollary 3.2.5. *Let $T(n,k;r)$ be the number of trees of order n with k sources such that vertex r is a sink. Then,*

$$T(n,k;r) = (r-1) \cdot \frac{(n-2)!}{k!} \sum_{i=0}^{r-2} \binom{r-2}{i} \left\{ {n-i-2 \atop n-k-1} \right\}. \tag{3.2.10}$$

Proof. We obtain the result by summing over all d in Equation (3.2.7). □

In the following corollary, we obtain a formula for the total number of spanning trees of a complete digraph on n vertices such that vertex r is a sink.

Corollary 3.2.6.

$$\sum_k T(n,k;r) = (r-1)\cdot n^{r-2}(n-1)^{n-r-1}. \tag{3.2.11}$$

Proof. This result can be proved by summing over all d in Equation (3.2.9):

$$\begin{aligned}\sum_k T(n,k;r) &= (r-1)\sum_{d=1}^{r-1}\binom{r-2}{d-1}(n-1)^{n-d-2}\\ &= (r-1)\sum_{i=0}^{r-2}\binom{r-2}{i}(n-1)^{n-i-3}\\ &= (r-1)(n-1)^{n-r-1}\sum_{i=0}^{r-2}\binom{r-2}{i}(n-1)^{r-i-2}.\end{aligned}$$

We proceed by binomial theorem to obtain the required formula. □

This formula can also be obtained directly from the generating function (3.2.2) by setting $x_r = 0$ and $x_i = 1$ for $i \neq r$ as well as $y_i = 1$ for all i. Since vertex n is always a sink in these trees then we recover Cayley's formula for the total number of labelled trees on n vertices by substituting for $r = n$ in Equation (3.2.11), i.e.

$$\sum_k T(n,k;n) = n^{n-2}. \tag{3.2.12}$$

Corollary 3.2.7. *Let $T(n,k)$ be the number of trees of order n with k sources. Then, we have*

$$T(n,k) = \frac{(n-1)!}{k!}\left\{{n-1 \atop n-k}\right\}. \tag{3.2.13}$$

Proof. Putting $r = n$ in Equation (3.2.10) and using the vertical recurrence relation:

$$\left\{{m+1 \atop p+1}\right\} = \sum_i \binom{m}{i}\left\{{i \atop p}\right\}, \tag{3.2.14}$$

where m and p are non-negative integers (see [28, Equation (6.15)]), we obtain the desired formula. □

In the following corollary, we obtain a formula for the total number of sinks in labelled trees on n vertices with k sources.

Corollary 3.2.8. *Let $T(n,k,\ell)$ be the number of trees of order n with k sources and ℓ sinks. Then the total number of sinks in these trees is given by*

$$\sum_{\ell} \ell \cdot T(n,k,\ell) = \frac{(n-2)!}{(k-1)!} \left\{ {n \atop n-k} \right\}. \tag{3.2.15}$$

Proof. We sum over all r in Equation (3.2.10). It suffices to prove that

$$\frac{(n-2)!}{k!} \sum_{r=1}^{n} \sum_{i=0}^{r-2} (r-1) \binom{r-2}{i} \left\{ {n-i-2 \atop n-k-1} \right\} = \frac{(n-2)!}{(k-1)!} \left\{ {n \atop n-k} \right\}$$

or equivalently

$$\sum_{r=1}^{n} \sum_{i=0}^{r-2} (r-1) \binom{r-2}{i} \left\{ {n-i-2 \atop n-k-1} \right\} = k \left\{ {n \atop n-k} \right\}.$$

We have

$$\begin{aligned}
\sum_{r=1}^{n} \sum_{i=0}^{r-2} (r-1) \binom{r-2}{i} \left\{ {n-i-2 \atop n-k-1} \right\} &= \sum_{i=0}^{n-2} \sum_{r=i+2}^{n} (r-1) \binom{r-2}{i} \left\{ {n-i-2 \atop n-k-1} \right\} \\
&= \sum_{i=0}^{n-2} (i+1) \sum_{r=i+2}^{n} \binom{r-1}{i+1} \left\{ {n-i-2 \atop n-k-1} \right\} \\
&= \sum_{i=0}^{n-2} (i+1) \sum_{j=i+1}^{n-1} \binom{j}{i+1} \left\{ {n-i-2 \atop n-k-1} \right\}.
\end{aligned}$$

By hockey stick identity (3.2.6), we obtain

$$\begin{aligned}
&\sum_{r=1}^{n} \sum_{i=0}^{r-2} (r-1) \binom{r-2}{i} \left\{ {n-i-2 \atop n-k-1} \right\} = \sum_{i=0}^{n-2} (i+1) \binom{n}{i+2} \left\{ {n-i-2 \atop n-k-1} \right\} \\
&= \sum_{i=0}^{n-2} n \binom{n-1}{i+1} \left\{ {n-i-2 \atop n-k-1} \right\} - \sum_{i=0}^{n-2} \binom{n}{i+2} \left\{ {n-i-2 \atop n-k-1} \right\} \\
&= \sum_{j=1}^{n-1} n \binom{n-1}{j} \left\{ {n-j-1 \atop n-k-1} \right\} - \sum_{j=2}^{n} \binom{n}{j} \left\{ {n-j \atop n-k-1} \right\} \\
&= \sum_{j=0}^{n-1} n \binom{n-1}{j} \left\{ {n-j-1 \atop n-k-1} \right\} - n \left\{ {n-1 \atop n-k-1} \right\} - \sum_{j=0}^{n} \binom{n}{j} \left\{ {n-j \atop n-k-1} \right\} \\
&+ n \left\{ {n-1 \atop n-k-1} \right\} + \left\{ {n \atop n-k-1} \right\}.
\end{aligned}$$

By Equation (3.2.14), we have

$$\begin{aligned}\sum_{r=1}^{n}\sum_{i=0}^{r-2}(r-1)\binom{r-2}{i}\left\{{n-i-2 \atop n-k-1}\right\} &= n\left\{{n \atop n-k}\right\} - \left\{{n+1 \atop n-k}\right\} + \left\{{n \atop n-k-1}\right\} \\ &= n\left\{{n \atop n-k}\right\} - (n-k)\left\{{n \atop n-k}\right\} - \left\{{n \atop n-k-1}\right\} + \left\{{n \atop n-k-1}\right\} \\ &= k\left\{{n \atop n-k}\right\}.\end{aligned}$$

This completes the proof. □

From Equation (3.2.13), it follows that the mean number of sinks in trees of order n with k sources is given by

$$\mu = \frac{k\left\{{n \atop n-k}\right\}}{(n-1)\left\{{n-1 \atop n-k}\right\}}.$$

We will obtain the same formula from a generating function approach in Corollary 3.3.2.

In Corollary 3.2.9 below, we obtain a formula for the total number of sinks in all trees on n vertices.

Corollary 3.2.9.

$$\sum_{k,\ell} \ell \cdot T(n,k,\ell) = (n-1)^{n-1}. \tag{3.2.16}$$

Proof. By summing over all k in Equation (3.2.15) we obtain, making use of (3.2.8),

$$\begin{aligned}\sum_{k,\ell} \ell \cdot T(n,k,\ell) &= \sum_{k=1}^{n-1} \frac{(n-2)!}{(k-1)!}\left\{{n \atop n-k}\right\} \\ &= \frac{1}{n-1}\sum_{k=1}^{n-1}\binom{n-1}{n-k}(n-k)!\left\{{n \atop n-k}\right\} \\ &= \frac{1}{n-1}\cdot(n-1)^n \\ &= (n-1)^{n-1}.\end{aligned}$$

This result can also be obtained from Equation (3.2.11) by summing over all r.

$$\begin{aligned}\sum_{k,\ell} \ell \cdot T(n,k,\ell) &= \sum_{r=1}^{n} (r-1) n^{r-2} (n-1)^{n-r-1} \\ &= \frac{1}{n^2} \cdot (n-1)^{n-1} \sum_{r=1}^{n} (r-1) \left(\frac{n}{n-1}\right)^{r} \\ &= \frac{1}{n^2} \cdot (n-1)^{n-1} \cdot \frac{n}{n-1} \sum_{r=1}^{n} (r-1) \left(\frac{n}{n-1}\right)^{r-1}.\end{aligned}$$

Let $r-1=i$, then

$$\sum_{k,\ell} \ell \cdot T(n,k,\ell) = \frac{1}{n} \cdot (n-1)^{n-2} \sum_{i=0}^{n-1} i \left(\frac{n}{n-1}\right)^{i}. \tag{3.2.17}$$

By the identity

$$\sum_{i=0}^{n-1} i x^{i} = \frac{x - n x^{n} + (n-1) x^{n+1}}{(1-x)^2}$$

and some algebra we obtain

$$\sum_{i=0}^{n} i \left(\frac{n}{n-1}\right)^{i} = n(n-1). \tag{3.2.18}$$

From Equations (3.2.17) and (3.2.18) we obtain the desired result. □

Corollary 3.2.10. *The probability that a randomly chosen vertex of a random tree on n vertices is a sink is given by*

$$\left(1 - \frac{1}{n}\right)^{n-1}. \tag{3.2.19}$$

Proof. Since each tree has n vertices then there are a total of n^{n-1} vertices in the n^{n-2} spanning trees. As seen in the preceding corollary, the total number of sinks in these trees is $(n-1)^{n-1}$. The result is thus immediate. □

Letting $n \to \infty$ in Equation (3.2.19), we obtain that there are approximately $\frac{n}{e}$ sinks in a tree.

Corollary 3.2.11. *The number of spanning trees of a complete digraph with a local orientation on n vertices having k sources and vertex r as a leaf sink is given by*

$$(r-1) \cdot \frac{(n-2)!}{k!} \left\{ \begin{matrix} n-2 \\ n-k-1 \end{matrix} \right\}. \tag{3.2.20}$$

Proof. Consider a complete digraph on $n-1$ vertices with labels from the set $[n] \setminus \{r\}$. By Corollary 3.2.7, there are $\frac{(n-2)!}{k!} \left\{ {n-2 \atop n-k-1} \right\}$ spanning trees for this digraph with k sources. Vertex r can only be attached at a vertex with label t where $1 \leq t < r$ so that it becomes a leaf sink. So for each tree on $n-1$ vertices, we obtain $r-1$ trees on n vertices. Thus we obtain the required formula. □

Corollary 3.2.11 can also be obtained by setting $d = 1$ in Equation (3.2.7). By summing over all k in Equation (3.2.20), we obtain the following result.

Corollary 3.2.12. *The number of trees on n vertices having vertex r as a leaf sink is given by:* $(r-1)(n-1)^{n-3}$.

The following corollary gives the total number of leaf sinks in trees on n vertices.

Corollary 3.2.13. *The number of leaf sinks in trees on n vertices is given by the formula:*

$$\frac{n(n-1)^{n-2}}{2}. \tag{3.2.21}$$

Proof. The proof follows from Corollary 3.2.12 by summing over all r. □

Corollary 3.2.14. *The probability that a randomly chosen vertex of a random tree on n vertices is a leaf sink is given by*

$$\frac{1}{2} \cdot \left(1 - \frac{1}{n}\right)^{n-2}. \tag{3.2.22}$$

Proof. Since each tree has n vertices, there are a total of n^{n-1} vertices in the n^{n-2} trees. By Equation (3.2.21), the total number of leaf sinks in these trees is $\frac{1}{2}n(n-1)^{n-2}$. The result thus follows. □

Letting $n \to \infty$ in Equation (3.2.22), we obtain that there are approximately $\frac{n}{2e}$ leaf sinks in a tree.

As in the proof of Lemma 3.2.11, there are

$$(n-r) \cdot \frac{(n-2)!}{k!} \left\{ {n-2 \atop n-k-1} \right\} \tag{3.2.23}$$

trees of order n with k sources such that vertex r is a leaf source. By summing over all r and k in Equation (3.2.23), we obtain Equation (3.2.21) as the total number of leaf sources in these trees. In total the number of leaves in all trees of order n is given by

$$n(n-1)^{n-2}. \tag{3.2.24}$$

Equation (3.2.24) appears in Neil Sloane's OEIS [55] as A055541: total number of leaves in all labelled trees with n nodes.

In [39, Theorem 3.5], we find that the total number of labelled trees of order n with s leaves is given by

$$\frac{n!}{s!}\left\{ {n-2 \atop n-s} \right\}. \tag{3.2.25}$$

Of course, by summing over all s in Equation (3.2.25) and using the identity (3.2.8), we recover Cayley's formula, n^{n-2}, for the number of labelled trees on n vertices.

Proposition 3.2.15. *The total number of leaf sinks in labelled trees on n vertices with s leaves is given by*

$$\frac{n!}{2(s-1)!}\left\{ {n-2 \atop n-s} \right\}. \tag{3.2.26}$$

Proof. Since there are s leaves in each of the trees enumerated by Equation (3.2.25), there is a total of

$$s \cdot \frac{n!}{s!}\left\{ {n-2 \atop n-s} \right\}$$

leaves in these trees. Now, by symmetry we obtain Equation (3.2.26) as the number of leaf sinks in these trees. □

Summing over all s in Equation (3.2.26), we recover Equation (3.2.21) for the total number of leaf sinks in trees of order n.

In the following proposition, we obtain a formula for the number of trees on n vertices such that the number of sources and sinks add up to n. This formula was obtained by Alexander Postnikov [47] in his study of intransitive trees. These trees also appear in a paper by Chauve, Dulucq and Rochnitzer [8] as alternating trees. An *intransitive* or *alternating* tree is a tree in which each vertex is either a source or a sink.

Proposition 3.2.16. *Let $T(n,k,\ell)$ be the number of trees of order n with k sources and ℓ sinks. We have*

$$\sum_{k+\ell=n} T(n,k,\ell) = \frac{1}{n \cdot 2^{n-1}} \sum_{i=1}^{n} \binom{n}{i} i^{n-1}.$$

3.3 Generating function approach

In this section, we obtain a differential equation satisfied by the generating function for the number of trees with a given number of sources and sinks. This leads us to a recurrence equation for these trees. We also obtain a functional equation for the generating function. The mean number of sinks and its variance follow as corollaries.

Let $T(t)$ be the generating function for the number of unrooted labelled trees. Then it follows (by Cayley's formula) that

$$T = T(t) = \sum_{N} \frac{t^{|N|}}{|N|!} = \sum_{n \geq 1} n^{n-2} \frac{t^n}{n!}.$$

The generating function for the number of rooted trees is

$$\sum_{N} |N| \frac{t^{|N|}}{|N|!} = t \frac{\partial}{\partial t} T(t) = t T_t.$$

By Pólya's *component principle* [45, 46], T_t satisfies the functional equation $T_t = \exp(tT_t)$. See also [25] for details.

Let x, y and z mark sources, sinks, and *other vertices* which are neither sources nor sinks respectively. We define the generating function

$$T(x,y,z,t) = \sum_{N} x^{\sharp\,\text{sources}(N)} y^{\sharp\,\text{sinks}(N)} z^{\sharp\,\text{other vertices}(N)} \frac{t^{|N|}}{|N|!}.$$

The generating functions for the collection of rooted trees in which the root is a source, sink and *other vertex* are given by xT_x, yT_y and zT_z respectively (subscripts indicating partial derivatives).

Note that if we draw an oriented edge from a source to vertex 1, then the source remains a source. However, a sink becomes *other vertex* while *other vertex* remains *other vertex*. Since vertex 1 is always a sink we obtain the following equation for the generating function:

$$T_t = y \exp(zT_y + zT_z + xT_x).$$

Here, vertex 1 counts as a sink if it is the only vertex but not as a source. By symmetry we have

$$x\exp(zT_x + zT_z + yT_y) = T_t + x - y.$$

Taking logarithms on both sides and differentiating with respect to t, we obtain

$$zT_{xt} + zT_{zt} + yT_{yt} = \frac{T_{tt}}{T_t + x - y}.$$

By setting $U(x, y, z, t) = T_t + x - y$, the last equation becomes

$$U(zU_z + zU_x + yU_y + y - z) = U_t. \tag{3.3.1}$$

On the other hand, we also have

$$tT_t = xT_x + yT_y + zT_z.$$

Replacing T_t with $U - x + y$ and differentiating with respect to t we get

$$U - x + y + tU_t = xT_{xt} + yT_{yt} + zT_{zt}.$$

Therefore,

$$U + tU_t = xU_x + yU_y + zU_z. \tag{3.3.2}$$

By combining Equations (3.3.1) and (3.3.2) we obtain a quasi-linear partial differential equation, namely

$$U((z - x)U_x + tU_t + U + y - z) = U_t.$$

By setting $z = 1$, we obtain the differential equation satisfied by the generating function for the number of trees on n vertices with a given number of sources and sinks, i.e.

$$U((1 - x)U_x + tU_t + U + y - 1) = U_t. \tag{3.3.3}$$

Since T_t is the generating function for the number of rooted labelled trees, we have

$$U = T_t + x - y = \sum_{n,k,\ell} T(n, k, \ell)x^\ell y^k \frac{t^{n-1}}{n-1} + x - y.$$

Now, by Equation (3.3.3), we have

$$\begin{aligned}
&\sum_{n,k,\ell} T(n+1,k,\ell)x^{\ell}y^{k}\\
&\quad= \sum_{n,k,\ell}(n-\ell)\sum_{i=0}^{n}\sum_{j=0}^{k}\sum_{h=0}^{\ell}\binom{n}{i} T(i,j,h)\cdot T(n-i,k-j,\ell-h)x^{\ell}y^{k}\\
&\quad+\sum_{n,k,\ell}(\ell+1)\sum_{i=0}^{n}\sum_{j=0}^{k}\sum_{h=0}^{\ell}\binom{n}{i} T(i,j,h)\cdot T(n-i,k-j,\ell+1-h)x^{\ell}y^{k}\\
&\quad+\sum_{n,k,\ell} T(n,k,\ell-1)x^{\ell}y^{k}+\sum_{n,k,\ell}\ell\cdot T(n,k,\ell)x^{\ell}y^{k}-\sum_{n,k,\ell}(\ell+1)T(n,k,\ell+1)x^{\ell}y^{k}.
\end{aligned}$$

This implies that the coefficients $T(n,k,\ell)$ satisfy the recurrence equation:

$$\begin{aligned}
T(n,k,\ell) &= (n-\ell-1)\sum_{i=0}^{n-1}\sum_{j=0}^{k}\sum_{h=0}^{\ell}\binom{n-1}{i} T(i,j,h)\cdot T(n-i-1,k-j,\ell-h)\\
&+(\ell+1)\sum_{i=0}^{n-1}\sum_{j=0}^{k}\sum_{h=0}^{\ell}\binom{n-1}{i} T(i,j,h)\cdot T(n-i-1,k-j,\ell+1-h)\\
&+T(n-1,k,\ell-1)+\ell\cdot T(n-1,k,\ell)-(\ell+1)T(n-1,k,\ell+1).
\end{aligned} \tag{3.3.4}$$

Table 3.1 below gives the number of trees of order $2 \le n \le 6$ with k sources and ℓ sinks.

We remain to find a functional equation satisfied by the generating function. We use the method of characteristics (see [21, Theorem 3.1]) to solve the PDE (3.3.3). We have the characteristic equations:

$$\frac{\mathrm{d}x}{(x-1)U} = \frac{\mathrm{d}t}{1-Ut} = \frac{\mathrm{d}y}{0} = \frac{\mathrm{d}U}{U(U+y-1)}.$$

This implies that $\mathrm{d}y = 0$. Hence $y = C_1$, where C_1 is a constant. We solve the differential equation

$$\frac{\mathrm{d}x}{(x-1)U} = \frac{\mathrm{d}U}{U(U+y-1)}$$

by the method of separation of variables to obtain

$$\ln(x-1)+c_2 = \ln(U+y-1).$$

Taking the exponential on both sides we get

$$C_2(x-1) = U+y-1,$$

n=2

$k \backslash \ell$	1
1	1

n=3

$k \backslash \ell$	1	2
1	1	1
2	1	

n=4

$k \backslash \ell$	1	2	3
1	1	4	1
2	4	5	
3	1		

n=5

$k \backslash \ell$	1	2	3	4
1	1	11	11	1
2	11	44	17	
3	11	17		
4	1			

n=6

$k \backslash \ell$	1	2	3	4	5
1	1	26	66	26	1
2	26	237	288	49	
3	66	288	146		
4	26	49			
5	1				

Table 3.1: Table showing the number of trees on n vertices with k sources and ℓ sinks.

where $C_2 = e^{c_2}$ is a constant. We rewrite the differential equation

$$\frac{\mathrm{d}t}{1-Ut} = \frac{\mathrm{d}U}{U(U+y-1)}$$

as

$$\frac{\mathrm{d}t}{\mathrm{d}U} + \frac{1}{U+y-1}t = \frac{1}{U(U+y-1)}.$$

This is a linear ODE which we solve using an integrating factor,

$$\exp\left(\int \frac{1}{U+y-1}\mathrm{d}U\right) = \exp\left(\ln(U+y-1)\right) = U+y-1.$$

Therefore,

$$t(U+y-1) = \int \frac{U+y-1}{U(U+y-1)}\mathrm{d}U = \ln U + C_3,$$

where C_3 is a constant.

A solution of the PDE (3.3.3) is of the form

$$F(C_1, C_2, C_3) = F\left(y, \frac{U+y-1}{x-1}, t(U+y-1) - \ln U\right) = 0, \tag{3.3.5}$$

where F is an arbitrary function. Since $U = x$ when $t = 0$, we must have

$$F\left(y, 1 + \frac{y}{x-1}, -\ln x\right) = 0,$$

for all x and y. Thus we take

$$F(z_1, z_2, z_3) = z_2 - 1 - \frac{z_1}{e^{-z_3} - 1}. \tag{3.3.6}$$

From Equations (3.3.5) and (3.3.6), we obtain

$$\frac{U + y - 1}{x - 1} - 1 - \frac{y}{\exp\left(\ln U - t(U + y - 1)\right) - 1} = 0.$$

We set $V = U - x$ to get

$$\frac{V + x + y - 1}{x - 1} - 1 = \frac{y}{(V + x)e^{-t(V+x+y-1)} - 1}.$$

Thus, we obtain the functional equation

$$(V + x)(V + y) = (V + xy)e^{t(V+x+y-1)}. \tag{3.3.7}$$

By replacing V by $\frac{R}{t}$, we obtain,

$$(R + tx)(R + ty) = (R + txy) \cdot te^{R+tx+ty-t}.$$

This functional equation is symmetric in x and y. It would be interesting to obtain further combinatorial interpretations of this functional equation.

We summarise the discussion above with the following theorem:

Theorem 3.3.1. *Let $T(n, k, \ell)$ be the number of trees on n vertices with k sinks and ℓ sources. The generating function*

$$R = \sum_{\substack{n \geq 2 \\ k, \ell \geq 1}} T(n, k, \ell) x^\ell y^k \frac{t^n}{(n-1)!}$$

satisfies the functional equation

$$(R + tx)(R + ty) = (R + txy) \cdot te^{R+tx+ty-t}.$$

Setting $x = 1$ in Equation (3.3.7) and differentiating implicitly with respect to y we obtain

$$V_y\big|_{x=1} = \frac{t(V + 1)}{1 - t(V + 1)}. \tag{3.3.8}$$

Differentiating Equation (3.3.8) with respect to y, we obtain

$$V_{yy}\big|_{x=1} = \frac{t\, V_y\big|_{x=1}}{(1 - t(V + 1))^2}.$$

By Equation (3.3.8), we have

$$V_{yy}\big|_{x=1} = \frac{t^2(V+1)}{(1-t(V+1))^3}. \tag{3.3.9}$$

By setting $x = 1$ in Equation (3.3.7) and differentiating implicitly with respect to t we obtain

$$V_t\big|_{x=1} = \frac{(V+1)(V+y)}{1-t(V+1)}. \tag{3.3.10}$$

We set $x = 1$ in Equation (3.3.7) and differentiate with respect to x to get

$$V_x\big|_{x=1} = \frac{y(V+1)}{(V+y)(1-t(V+1))} - 1.$$

By differentiating with respect to t and plugging in Equation (3.3.10) we get

$$V_{xt}\big|_{x=1} = \frac{y(V+1)}{(1-t(V+1))^3}. \tag{3.3.11}$$

Now, by Equations (3.3.9) and (3.3.11), we have

$$t^2 \frac{\partial^2}{\partial x \partial t} V\bigg|_{x=1} = y \frac{\partial^2}{\partial y^2} V\bigg|_{x=1}. \tag{3.3.12}$$

We proceed in the same manner to obtain

$$t^4 \frac{\partial^4}{\partial x^2 \partial t^2} V\bigg|_{x=1} = y^2 \frac{\partial^4}{\partial y^4} V - t(y^2 - 2y) \frac{\partial^3}{\partial y^3} V\bigg|_{x=1}. \tag{3.3.13}$$

Corollary 3.3.2. *The mean number of sinks in trees of order n having k sources is*

$$\mu = \frac{k \left\{ {n \atop n-k} \right\}}{(n-1) \left\{ {n-1 \atop n-k} \right\}}. \tag{3.3.14}$$

Proof. We derive the formula from Equation (3.3.12). We have an exponential generating function

$$V = \sum_{n,k,\ell} T(n,k,\ell) x^\ell y^k \frac{t^{n-1}}{(n-1)!},$$

for the number of labelled trees on n vertices having ℓ sources and k sinks. Thus, x and y are marking sources and sinks respectively. We have,

$$t^2 \frac{\partial^2}{\partial x \partial t} V \bigg|_{x=1} = (n-1) \sum_{n,k,\ell} \ell \cdot T(n,k,\ell) \frac{y^k t^n}{(n-1)!} \tag{3.3.15}$$

and

$$y \frac{\partial^2}{\partial y^2} V \bigg|_{x=1} = \sum_{n,k,\ell} k(k-1) T(n,k,\ell) \frac{y^{k-1} t^{n-1}}{(n-1)!}. \tag{3.3.16}$$

Since Equations (3.3.15) and (3.3.16) are equal (by Equation (3.3.12)), we can compare the coefficients of $y^k t^n$ and use Equation (3.2.13) to arrive at

$$\sum_{\ell} \ell \cdot T(n,k,\ell) = \frac{(n-2)!}{(k-1)!} \left\{ {n \atop n-k} \right\}.$$

Since there are

$$\frac{(n-1)!}{k!} \left\{ {n-1 \atop n-k} \right\}$$

trees on n vertices with k sources by Equation (3.2.13), it follows that the mean number of sinks in these trees is given by

$$\mu = \left(\frac{(n-1)!}{k!} \left\{ {n-1 \atop n-k} \right\} \right)^{-1} \cdot \frac{(n-2)!}{(k-1)!} \left\{ {n \atop n-k} \right\}$$
$$= \frac{k \left\{ {n \atop n-k} \right\}}{(n-1) \left\{ {n-1 \atop n-k} \right\}}.$$

□

Corollary 3.3.3. *The variance of the number of sinks in trees of order n having k sources is given by*

$$\sigma^2 = \frac{1}{A} \left[(k(n-k)-1) \frac{(n-3)!}{(k-1)!} \left\{ {n \atop n-k} \right\} + (k+1) \frac{(n-3)!}{(k-1)!} \left\{ {n \atop n-k-1} \right\} \right]$$
$$- \frac{1}{A^2} \left(\frac{(n-2)!}{(k-1)!} \left\{ {n \atop n-k} \right\} \right)^2,$$

where

$$A = \frac{(n-1)!}{k!} \left\{ {n-1 \atop n-k} \right\}.$$

Proof. Consider the generating function

$$V = \sum_{n,k,\ell} T(n,k,\ell) x^\ell y^k \frac{t^{n-1}}{(n-1)!}$$

where x is marking sources and y is marking sinks in labelled trees of order n. We have

$$t^4 \frac{\partial^4}{\partial x^2 \partial t^2} V \bigg|_{x=1} = (n-1)(n-2) \sum_{n,k,\ell} \ell(\ell-1) \cdot T(n,k,\ell) \frac{y^k t^{n+1}}{(n-1)!}, \quad (3.3.17)$$

and

$$\begin{aligned} & y^2 \frac{\partial^4}{\partial y^4} V \bigg|_{x=1} - t(y^2-2y) \frac{\partial^3}{\partial y^3} V \bigg|_{x=1} \\ &\quad = \sum_{n,k,\ell} \frac{k!}{(k-4)!} T(n,k,\ell) \frac{y^{k-2} t^{n-1}}{(n-1)!} - \sum_{n,k,\ell} \frac{k!}{(k-3)!} T(n,k,\ell) \frac{y^{k-1} t^n}{(n-1)!} \\ &\quad + 2 \sum_{n,k,\ell} \frac{k!}{(k-3)!} T(n,k,\ell) \frac{y^{k-2} t^n}{(n-1)!}. \end{aligned} \quad (3.3.18)$$

By Equation (3.3.13), we know that Equations (3.3.17) and (3.3.18) are equal. We can thus compare the coefficients of $y^k t^{n+1}$ and use Equation (3.2.13) to obtain

$$\begin{aligned} & \sum_\ell \ell(\ell-1) T(n,k,\ell) \\ &\quad = \frac{(n-3)!}{(k-2)!} \left\{ {n+1 \atop n-k} \right\} - \frac{(n-3)!}{(k-2)!} \left\{ {n \atop n-k} \right\} + 2 \frac{(n-3)!}{(k-1)!} \left\{ {n \atop n-k-1} \right\} \\ &\quad = (n-k-1) \frac{(n-3)!}{(k-2)!} \left\{ {n \atop n-k} \right\} + (k+1) \frac{(n-3)!}{(k-1)!} \left\{ {n \atop n-k-1} \right\}. \end{aligned}$$

We get the second moment:

$$\begin{aligned} \sum_\ell \ell^2 T(n,k,\ell) &= \sum_\ell \ell(\ell-1) T(n,k,\ell) + \sum_\ell \ell \cdot T(n,k,\ell) \\ &= (n-k-1) \frac{(n-3)!}{(k-2)!} \left\{ {n \atop n-k} \right\} + (k+1) \frac{(n-3)!}{(k-1)!} \left\{ {n \atop n-k-1} \right\} \\ &+ \frac{(n-2)!}{(k-1)!} \left\{ {n \atop n-k} \right\} \\ &= (k(n-k)-1) \frac{(n-3)!}{(k-1)!} \left\{ {n \atop n-k} \right\} + (k+1) \frac{(n-3)!}{(k-1)!} \left\{ {n \atop n-k-1} \right\}. \end{aligned} \quad (3.3.19)$$

Now, since the variance equals

$$\sigma^2 = \frac{\sum_\ell \ell^2 T(n,k,\ell)}{\sum_\ell T(n,k,\ell)} - \left(\frac{\sum_\ell \ell \cdot T(n,k,\ell)}{\sum_\ell T(n,k,\ell)} \right)^2,$$

by Equations (3.2.13), (3.3.14) and (3.3.19), we obtain the required formula. □

3.4 Trees with two marked vertices

In this section, we extend the results in Section 3.2 to trees with two marked vertices. Let us first state Joyal's bijection as motivation.

Joyal's bijection (see [1, 32] for details): There is a bijection between the set of labelled trees on n vertices with two marked vertices and the set of functions from the set $[n]$ to itself. Since there are n^n such functions, it follows that there are n^n labelled trees on n vertices such that two (not necessarily distinct) vertices are marked. This marking is done such that there is an arrow leaving from a vertex u to nowhere, and there is also another arrow coming into a vertex v from nowhere. Vertices u and v are not necessarily distinct. The indegree and outdegree of vertices in the labelled trees are preserved after this marking. The bijection is between trees and functional digraphs. So all results in this section imply analogous statements for functional digraphs.

Since there are n choices for vertex u and also n choices for vertex v, a generating function for such trees, in terms of indegrees and outdegrees of their vertices is given by

$$\sum_{\substack{T \in \mathrm{Tree}(K_n) \\ T \text{ has 2 marked vertices}}} \prod_{i=1}^{n} y_i^{\mathrm{indeg}_T(i)} x_i^{\mathrm{outdeg}_T(i)}$$
$$= (x_1 + \cdots + x_n)(y_1 + \cdots + y_n) \sum_{T \in \mathrm{Tree}(K_n)} \prod_{i=1}^{n} y_i^{\mathrm{indeg}_T(i)} x_i^{\mathrm{outdeg}_T(i)}. \quad (3.4.1)$$

The terms $(x_1 + \cdots + x_n)$ and $(y_1 + \cdots + y_n)$ represent one arrow to nowhere and from nowhere respectively. The term

$$\sum_{T \in \mathrm{Tree}(K_n)} \prod_{i=1}^{n} y_i^{\mathrm{indeg}_T(i)} x_i^{\mathrm{outdeg}_T(i)}$$

is the generating function for the number of labelled trees with given indegree and outdegree sequence, given by Equation (3.2.2).

Theorem 3.4.1. *Let $N_2^d(n,\lambda;r)$ be the number of trees of order n with an indegree sequence $\lambda = \langle 1^{e_1}2^{e_2}\cdots\rangle$ such that vertex r is a sink of degree d and there are two marked vertices. The number of such trees is*

$$N_2^d(n,\lambda;r) = \frac{(r-1)(r-1)!(n-1)!(n-d)!de_d}{(r-d)!e_0!(0!)^{e_0}e_1!(1!)^{e_1}e_2!(2!)^{e_2}\cdots}. \tag{3.4.2}$$

Proof. In Equation (3.4.1), if we set $x_r = 0$ and $x_i = 1$ for all $i \neq r$, then we obtain

$$\sum_{\substack{T_2:\, r \text{ is a sink,}\\ 2 \text{ vertices marked}}} \prod_{i=1}^{n} y_i^{\mathrm{indeg}_T(i)} = (n-1)(r-1)y_r y_n \cdot \prod_{i=1}^{r-1}(iy_i + \cdots + y_n) \times \prod_{i=r+1}^{n-1}((i-1)y_i + \cdots + y_n) \tag{3.4.3}$$

which is the generating function for the number of trees on n vertices having two marked vertices such that vertex r is a sink. In the generating function (3.4.3), y_r does not occur as a term in the last product and thus we can rewrite it as:

$$\sum_{\substack{T_2:\, r \text{ is a sink,}\\ 2 \text{ vertices marked}}} \prod_{i=1}^{n} y_i^{\mathrm{indeg}_T(i)} = (n-1)\cdot\frac{r-1}{n-1}\sum_{B\subseteq\{1,2,\ldots,r-1\}} y_r^{|B|+1} \times \prod_{\substack{i=1\\ i\notin B}}^{r-1}(iy_i + y_{i+1} + \cdots + y_{r-1} + y_{r+1} + \cdots + y_n) \times \prod_{i=r+1}^{n}((i-1)y_i + y_{i+1} + \cdots + y_n). \tag{3.4.4}$$

Setting $y_i = z_i$ if $i < r$ and $y_{i+1} = z_i$ if $i \geq r$, Equation (3.4.4) becomes

$$\begin{aligned}\sum_{\substack{T_2:\, r \text{ is a sink,}\\ 2 \text{ vertices marked}}} \prod_{i=1}^{n} y_i^{\mathrm{indeg}_T(i)} &= (r-1)\sum_{B\subseteq\{1,2,\ldots,r-1\}} y_r^{|B|+1} \cdot \prod_{\substack{i=1\\ i\notin B}}^{r-1}(iz_i + z_{i+1} + \cdots + z_{n-1}) \\ &\times \prod_{i=r}^{n-1}(iz_i + z_{i+1} + \cdots + z_{n-1}) \\ &= (r-1)\sum_{B\subseteq\{1,2,\ldots,r-1\}} y_r^{|B|+1} \cdot \prod_{\substack{i=1\\ i\notin B}}^{n-1}(iz_i + z_{i+1} + \cdots + z_{n-1}).\end{aligned}$$

With a similar construction as in the proof of Theorem 3.2.1, we have

$$N_2^d(n,\lambda;r) = \frac{(r-1)(r-1)!(n-1)!(n-d)!de_d}{(r-d)!e_0!(0!)^{e_0}e_1!(1!)^{e_1}e_2!(2!)^{e_2}\cdots}.$$

□

A number of corollaries follow from this theorem. The following corollary follows by substituting for $r = n$ in Equation (3.4.2), and summing over all d making use of the fact that $\sum_{d\geq 1} de_d = n$ in these trees.

Corollary 3.4.2. *Let $N_2(n,\lambda;n)$ be the number of trees of order n with an indegree sequence $\lambda = \langle 1^{e_1}2^{e_2}\cdots\rangle$ such that there are two marked vertices and vertex n is a sink. There are*

$$N_2(n,\lambda;n) = \frac{(n-1)(n-1)!n!}{e_0!(0!)^{e_0}e_1!(1!)^{e_1}e_2!(2!)^{e_2}\cdots}$$

such trees.

Corollary 3.4.3. *The total number of sinks of degree d in trees of order n with an indegree sequence $\lambda = \langle 1^{e_1}2^{e_2}\cdots\rangle$ such that two vertices are marked is given by*

$$\frac{(n-1)!n!e_d}{e_0!(0!)^{e_0}e_1!(1!)^{e_1}e_2!(2!)^{e_2}\cdots}\cdot\frac{nd-1}{d+1}. \tag{3.4.5}$$

Proof. We sum over r in Equation (3.4.2) to obtain the total number of sinks as

$$\begin{aligned}
&\frac{(n-1)!(n-d)!de_d}{e_0!(0!)^{e_0}e_1!(1!)^{e_1}e_2!(2!)^{e_2}\cdots}\sum_{r=1}^{n}(r-1)\frac{(r-1)!}{(r-d)!}\\
&=\frac{(n-1)!(n-d)!de_d}{e_0!(0!)^{e_0}e_1!(1!)^{e_1}e_2!(2!)^{e_2}\cdots}\sum_{r=1}^{n}(r-1)\binom{r-1}{d-1}(d-1)!\\
&=\frac{(n-1)!(n-d)!d!e_d}{e_0!(0!)^{e_0}e_1!(1!)^{e_1}e_2!(2!)^{e_2}\cdots}\sum_{i=0}^{n-1}i\binom{i}{d-1}\\
&=\frac{(n-1)!(n-d)!d!e_d}{e_0!(0!)^{e_0}e_1!(1!)^{e_1}e_2!(2!)^{e_2}\cdots}\left[\sum_{i=0}^{n-1}(i+1)\binom{i}{d-1}-\sum_{i=0}^{n-1}\binom{i}{d-1}\right]\\
&=\frac{(n-1)!(n-d)!d!e_d}{e_0!(0!)^{e_0}e_1!(1!)^{e_1}e_2!(2!)^{e_2}\cdots}\left[d\sum_{i=0}^{n-1}\binom{i+1}{d}-\sum_{i=0}^{n-1}\binom{i}{d-1}\right]\\
&=\frac{(n-1)!(n-d)!d!e_d}{e_0!(0!)^{e_0}e_1!(1!)^{e_1}e_2!(2!)^{e_2}\cdots}\left[d\binom{n+1}{d+1}-\binom{n}{d}\right]\\
&=\frac{(n-1)!n!e_d}{e_0!(0!)^{e_0}e_1!(1!)^{e_1}e_2!(2!)^{e_2}\cdots}\left[\frac{(n+1)d}{d+1}-1\right].
\end{aligned}$$

This completes the proof. □

By summing over all d in Equation (3.4.5), we obtain a formula for the total number of sinks in trees where two vertices are marked.

Corollary 3.4.4. *Let $N_2(n, \ell, \lambda)$ be the number of trees of order n with ℓ sinks and indegree sequence $\lambda = \langle 1^{e_1} 2^{e_2} \cdots \rangle$ such that two vertices are marked. The total number of sinks in these trees is given by:*

$$\sum_{\ell} \ell \cdot N_2(n, \ell, \lambda) = \frac{n!^2(n - e_0 + 1) - (n-1)!(n+1)! \sum_{d \geq 1} \frac{e_d}{d+1}}{e_0!(0!)^{e_0} e_1!(1!)^{e_1} e_2!(2!)^{e_2} \cdots}.$$

Proof. We have

$$\sum_{\ell} \ell \cdot N_2(n, \ell, \lambda) = \frac{(n-1)!n!}{e_0!(0!)^{e_0} e_1!(1!)^{e_1} e_2!(2!)^{e_2} \cdots} \sum_{d \geq 1} \frac{(nd-1)e_d}{d+1}$$

$$= \frac{(n-1)!n!}{e_0!(0!)^{e_0} e_1!(1!)^{e_1} e_2!(2!)^{e_2} \cdots} \left[n(n - e_0 + 1) - (n+1) \sum_{d \geq 1} \frac{e_d}{d+1} \right].$$

□

Lemma 3.4.5. *Let $N_2'(n, \lambda)$ be the number of trees of order n with an indegree sequence $\lambda = \langle 1^{e_1} 2^{e_2} \cdots \rangle$ such that vertex n is not a sink and there are two marked vertices. The number of such trees is given by*

$$N_2'(n, \lambda) = \frac{(n-1)!n!}{e_0!(0!)^{e_0} e_1!(1!)^{e_1} e_2!(2!)^{e_2} \cdots}.$$

Proof. The generating function of such trees is obtained by setting $x_n = 1$ and $x_i = 0$ for all $i \neq n$ in the term $(x_1 + \cdots + x_n)$ and also setting $x_i = 1$ for all i in the term

$$\sum_{T \in \mathrm{Tree}(K_n)} \prod_{i=1}^{n} y_i^{\mathrm{indeg}_T(i)} x_i^{\mathrm{outdeg}_T(i)}$$

$$= x_1 y_n \prod_{i=2}^{n-1} \left((x_1 + \cdots + x_i) y_i + x_i (y_{i+1} + \cdots + y_n) \right)$$

of Equation (3.4.1). The generating function is therefore

$$\sum_{\substack{T_2:\, n \text{ is not a sink,} \\ 2 \text{ vertices marked}}} \prod_{i=1}^{n} y_i^{\mathrm{indeg}_T(i)} = y_n \sum_{i=1}^{n} y_i \cdot \prod_{i=2}^{n-1} (i y_i + \cdots + y_n)$$

$$= \frac{1}{n} \prod_{i=1}^{n} (i y_i + \cdots + y_n).$$

Now using the same arguments as in the proof of Theorem 3.4.1 with $r = n$, we obtain

$$\frac{(n-1)!^2 d e_d}{e_0!(0!)^{e_0} e_1!(1!)^{e_1} e_2!(2!)^{e_2}\cdots}$$

trees on n vertices such that vertex n is a non-sink vertex of degree d. Summing over d we obtain the required equation. □

Corollary 3.4.6. *Let $N_2(n, \lambda)$ be the number of trees of order n with an indegree sequence $\lambda = \langle 1^{e_1} 2^{e_2} \cdots \rangle$ such that there are two marked vertices. The number of such trees is given as*

$$N_2(n,\lambda) = \frac{n!^2}{e_0!(0!)^{e_0} e_1!(1!)^{e_1} e_2!(2!)^{e_2}\cdots}.$$

Proof. This number is the sum of the number of trees in which the vertex labelled n is a sink and those in which vertex n is not a sink, i.e.

$$N_2(n,\lambda) = N_2(n,\lambda;n) + N_2'(n,\lambda),$$

so the formula follows from Corollary 3.4.2 and Lemma 3.4.5. This result was also proved by Wagner [62], (setting $m = 1$ in Equation (3.1.3)). □

We now consider the enumeration of trees with two marked vertices by number of sources and sinks.

Theorem 3.4.7. *Let $T_2^d(n,k;r)$ be the number of trees of order n with two marked vertices and k sources such that vertex r is a sink of degree d. Then*

$$T_2^d(n,k;r) = \frac{(r-1)(r-1)!(n-1)!}{(r-d)!(d-1)!k!} \left\{ {n-d \atop n-k-1} \right\}. \tag{3.4.6}$$

Proof. The proof follows a similar construction as in the proof of Theorem 3.4.1 by considering the number of sources rather than indegree sequences. □

By summing over k in Equation (3.4.6), we obtain the formula

$$\sum_k T_2^d(n,k;r) = (r-1)\binom{r-1}{d-1}(n-1)^{n-d} \tag{3.4.7}$$

for the total number of trees on n vertices having vertex r as a sink of degree d such that two vertices are marked.

Corollary 3.4.8. *Let $T_2(n,k;r)$ be the number of trees of order n with two marked vertices and k sources such that vertex r is a sink. Then,*

$$T_2(n,k;r) = (r-1) \cdot \frac{(n-1)!}{k!} \sum_{i=0}^{r-1} \binom{r-1}{i} \left\{ {n-i-1 \atop n-k-1} \right\}. \qquad (3.4.8)$$

Proof. We obtain this formula by summing over all d in Equation (3.4.6). □

Summing over all k in Equation (3.4.8) (or over all d in Equation (3.4.7)), we obtain a formula for the number of labelled trees on n vertices having two marked vertices such that vertex r is a sink:

$$\sum_k T_2(n,k;r) = (r-1) \cdot n^{r-1}(n-1)^{n-r}. \qquad (3.4.9)$$

From Equations (3.2.11) and (3.4.9) we see that, for any r

$$\sum_k T_2(n,k;r) = n(n-1) \sum_k T(n,k;r).$$

By substituting $r = n$ in Equation (3.4.8) and using the vertical recurrence relation of the Stirling numbers of the second kind we obtain a formula, namely

$$T_2(n,k;n) = (n-1)\frac{(n-1)!}{k!} \left\{ {n \atop n-k} \right\}, \qquad (3.4.10)$$

for the number of trees on n vertices having k sinks such that there are two marked vertices and vertex n is a sink.

The following corollary gives the total number of sinks in all trees on n vertices with k sources such that there are two marked vertices.

Corollary 3.4.9. *Let $T_2(n,k,\ell)$ be the number of trees of order n with k sources and ℓ sinks such that there are two marked vertices. The total number of sinks in trees with k sources is given by*

$$\sum_\ell \ell \cdot T_2(n,k,\ell) = \frac{(n-1)!}{k!} \left\{ {n \atop n-k-1} \right\} + \frac{(n-1)!}{(k-1)!} \left\{ {n+1 \atop n-k} \right\}. \qquad (3.4.11)$$

Proof. We use Corollary 3.4.8, and we shall show that

$$\frac{(n-1)!}{k!} \sum_{r=1}^{n} \sum_{i=0}^{r-1} (r-1) \binom{r-1}{i} \left\{ {n-i-1 \atop n-k-1} \right\} = \frac{(n-1)!}{k!} \left\{ {n \atop n-k-1} \right\} + \frac{(n-1)!}{(k-1)!} \left\{ {n+1 \atop n-k} \right\}.$$

It suffices to show that

$$\sum_{r=1}^{n}\sum_{i=0}^{r-1}(r-1)\binom{r-1}{i}\left\{ {n-i-1 \atop n-k-1} \right\} = \left\{ {n \atop n-k-1} \right\} + k\left\{ {n+1 \atop n-k} \right\}.$$

We have

$$\begin{aligned}\sum_{r=1}^{n}\sum_{i=0}^{r-1}(r-1)\binom{r-1}{i}\left\{ {n-i-1 \atop n-k-1} \right\} &= \sum_{i=0}^{n-1}\sum_{r=i+1}^{n}(r-1)\binom{r-1}{i}\left\{ {n-i-1 \atop n-k-1} \right\} \\ &= \sum_{i=0}^{n-1}\sum_{r=i+1}^{n} r\binom{r-1}{i}\left\{ {n-i-1 \atop n-k-1} \right\} - \sum_{i=0}^{n-1}\sum_{r=i+1}^{n}\binom{r-1}{i}\left\{ {n-i-1 \atop n-k-1} \right\} \\ &= \sum_{i=0}^{n-1}(i+1)\sum_{r=i+1}^{n}\binom{r}{i+1}\left\{ {n-i-1 \atop n-k-1} \right\} - \sum_{i=0}^{n-1}\sum_{j=i}^{n-1}\binom{j}{i}\left\{ {n-i-1 \atop n-k-1} \right\}.\end{aligned}$$

By Equation (3.2.6), we get

$$\begin{aligned}&\sum_{r=1}^{n}\sum_{i=0}^{r-1}(r-1)\binom{r-1}{i}\left\{ {n-i-1 \atop n-k-1} \right\} \\ &\quad= \sum_{i=0}^{n-1}(i+1)\binom{n+1}{i+2}\left\{ {n-i-1 \atop n-k-1} \right\} - \sum_{i=0}^{n-1}\binom{n}{i+1}\left\{ {n-i-1 \atop n-k-1} \right\} \\ &\quad= \sum_{i=0}^{n-1}(n+1)\binom{n}{i+1}\left\{ {n-i-1 \atop n-k-1} \right\} - \sum_{i=0}^{n-1}\binom{n+1}{i+2}\left\{ {n-i-1 \atop n-k-1} \right\} \\ &\quad\quad - \sum_{i=0}^{n-1}\binom{n}{i+1}\left\{ {n-i-1 \atop n-k-1} \right\} \\ &\quad= \sum_{j=1}^{n}(n+1)\binom{n}{j}\left\{ {n-j \atop n-k-1} \right\} - \sum_{j=2}^{n+1}\binom{n+1}{j}\left\{ {n-j+1 \atop n-k-1} \right\} \\ &\quad\quad - \sum_{j=1}^{n}\binom{n}{j}\left\{ {n-j \atop n-k-1} \right\} \\ &\quad= \sum_{j=0}^{n}(n+1)\binom{n}{j}\left\{ {n-j \atop n-k-1} \right\} - (n+1)\left\{ {n \atop n-k-1} \right\} \\ &\quad\quad - \sum_{j=0}^{n+1}\binom{n+1}{j}\left\{ {n-j+1 \atop n-k-1} \right\} - \left\{ {n+1 \atop n-k-1} \right\} - (n+1)\left\{ {n \atop n-k-1} \right\} \\ &\quad\quad - \sum_{j=0}^{n}\binom{n}{j}\left\{ {n-j \atop n-k-1} \right\} - \left\{ {n \atop n-k-1} \right\}.\end{aligned}$$

Now, by identity (3.2.14), we obtain

$$\sum_{r=1}^{n}\sum_{i=0}^{r-1}(r-1)\binom{r-1}{i}\left\{{n-i-1 \atop n-k-1}\right\}$$
$$= (n+1)\left\{{n+1 \atop n-k}\right\} - (n+1)\left\{{n \atop n-k-1}\right\} - \left\{{n+2 \atop n-k}\right\}$$
$$+ (n+1)\left\{{n \atop n-k-1}\right\} + \left\{{n+1 \atop n-k-1}\right\} - \left\{{n+1 \atop n-k}\right\} + \left\{{n \atop n-k-1}\right\}$$
$$= n\left\{{n+1 \atop n-k}\right\} - (n-k)\left\{{n+1 \atop n-k}\right\} - \left\{{n+1 \atop n-k-1}\right\} + \left\{{n+1 \atop n-k-1}\right\}$$
$$+ \left\{{n \atop n-k-1}\right\}$$
$$= k\left\{{n+1 \atop n-k}\right\} + \left\{{n \atop n-k-1}\right\}.$$

This completes the proof. □

By Equations (3.4.10) and (3.4.11), we see that the average number of sinks in labelled trees of order n with two marked vertices and k sources is

$$\mu = \frac{1}{n-1}\left(\left\{{n \atop n-k}\right\}\right)^{-1}\left[\left\{{n \atop n-k-1}\right\} + k\left\{{n+1 \atop n-k}\right\}\right].$$

From Equation (3.4.11), we have

$$\sum_{\ell}\ell \cdot T_2(n,k,\ell) = (n-k)\cdot\frac{(n-1)!}{(k-1)!}\left\{{n \atop n-k}\right\} + (k+1)\frac{(n-1)!}{k!}\left\{{n \atop n-k-1}\right\}.$$

Now, from Equation (3.2.15), we obtain the following equation that relates ordinary labelled trees and those with two marked vertices:

$$\sum_{\ell}\ell \cdot T_2(n,k,\ell) = (n-1)\left[(n-k)\sum_{\ell}\ell\cdot T(n,k,\ell) + (k+1)\sum_{\ell}\ell\cdot T(n,k+1,\ell)\right].$$

If we sum over k in Equation (3.4.11) we obtain a formula for the total number of sinks in these trees, i.e.

$$\sum_{k,\ell}\ell \cdot T_2(n,k,\ell) = n\cdot(n-1)^n. \qquad (3.4.12)$$

This formula is also obtained by summing over all r in Equation (3.4.9). By Equations (3.2.16) and (3.4.12), we obtain a relation between the total number of sinks in labelled trees and the total number of sinks in labelled trees but with two marked vertices:

$$\sum_{k,\ell} \ell \cdot T_2(n,k,\ell) = n(n-1) \sum_{k,\ell} \ell \cdot T(n,k,\ell).$$

We obtain labelled trees on n vertices with k sources and ℓ sinks such that two vertices are marked from ordinary labelled trees by performing any of the four procedures described below:

1. Allow an arrow to leave a non-sink vertex v to nowhere and another arrow to enter a non-source vertex u from nowhere in a labelled tree of order n with k sources and ℓ sinks. By this procedure, we obtain a total of

   $$(n-k)(n-\ell)T(n,k,\ell)$$

   marked trees on n vertices with k sources and ℓ sinks.

2. Create an arrow from a non-sink vertex v to nowhere and another arrow to a source u from nowhere in a labelled tree on n vertices with $k+1$ sources and ℓ sinks. We obtain

   $$(n-\ell)(k+1)T(n,k+1,\ell)$$

   marked trees of order n with k sources and ℓ sinks.

3. Create an arrow from a sink v to nowhere and another arrow to a non-source vertex u from nowhere in a labelled tree of order n with k sources and $\ell+1$ sinks. Here, we obtain a total of

   $$(n-k)(\ell+1)T(n,k,\ell+1)$$

   possible marked trees on n vertices with k sources and ℓ sinks.

4. Lastly, create an arrow from a sink v to nowhere and allow another arrow to enter source u from nowhere in a labelled tree of order n with $k+1$ sources and $\ell+1$ sinks. In this case, there are

   $$(k+1)(\ell+1)T(n,k+1,\ell+1)$$

   possible marked trees on n vertices with k sources and ℓ sinks.

Therefore many of the results also follow from the relation

$$\begin{aligned} T_2(n,k,\ell) &= (n-k)(n-\ell)T(n,k,\ell) + (k+1)(n-\ell)T(n,k+1,\ell) \\ &\quad + (n-k)(\ell+1)T(n,k,\ell+1) + (k+1)(\ell+1)T(n,k+1,\ell+1). \end{aligned}$$

Lemma 3.4.10. *Let $T_2'(n,k)$ be the number of trees of order n with two marked vertices and k sources such that vertex n is not a sink. Then,*

$$T_2'(n,k) = \frac{(n-1)!}{k!} \left\{ {n \atop n-k} \right\}. \tag{3.4.13}$$

Proof. The proof is similar to that of Lemma 3.4.5 if we consider enumeration by number of sources and not by indegree sequences. □

Now we have,

$$T_2(n,k) = T_2(n,k;n) + T_2'(n,k).$$

By Equations (3.4.10) and (3.4.13) we obtain

$$T_2(n,k) = \frac{n!}{k!} \left\{ {n \atop n-k} \right\}.$$

Summing over all k we recover the formula for the total number of trees on n vertices with two marked vertices:

$$\sum_k T_2(n,k) = n^n,$$

and thus by Equation (3.2.12) we have,

$$\sum_k T_2(n,k) = n^2 \sum_k T(n,k).$$

Chapter 4

Reachability in trees with a local orientation

4.1 Introduction

In this chapter, we investigate the average number of vertices that can be reached from a given vertex i following oriented edges in trees with a local orientation [17]. We also obtain a formula for the number of trees on n vertices such that exactly k vertices are reachable from vertex i.

Definition 4.1.1. A vertex j is said to be *reachable* from a vertex i if there is an oriented path from vertex i to vertex j, and we say that a path p has length ℓ if there are ℓ edges on the path.

4.2 Enumeration of trees by path lengths

The main aim of this section is to find the average number of vertices, sinks and leaf sinks in labelled trees of order n that are reachable from a given vertex i.

Proposition 4.2.1. *Let $R(n,i,j,\ell)$ be the number of labelled trees on $[n]$ such that vertex j is reachable from vertex i by a path of length ℓ. We have*

$$R(n,i,j,\ell) = \binom{j-i-1}{\ell-1}(\ell+1)n^{n-\ell-2}. \qquad (4.2.1)$$

Proof. Let us consider a path of length ℓ starting at vertex i and ending at vertex j. There are $\binom{j-i-1}{\ell-1}$ possible paths. Now, the number of forests on n

labelled vertices rooted at the $\ell + 1$ vertices on the path is given by

$$(\ell+1)n^{n-\ell-2},$$

(see [39, Theorem 3.3, p. 17]). Therefore, there are

$$\binom{j-i-1}{\ell-1}(\ell+1)n^{n-\ell-2}$$

trees in which vertex j can be reached from vertex i in ℓ steps. □

Corollary 4.2.2. *The total number of labelled trees on n vertices such that vertex j is reachable from vertex $i \neq j$ is given by*

$$(2n+j-i+1)n^{n-j+i-2}(n+1)^{j-i-2}. \tag{4.2.2}$$

Proof. We sum over all ℓ in Equation (4.2.1).

$$\begin{aligned}\sum_{\ell} R(n,i,j,\ell) &= \sum_{\ell=1}^{j-i}\binom{j-i-1}{\ell-1}(\ell+1)n^{n-\ell-2}\\ &= \sum_{\ell=0}^{j-i-1}\binom{j-i-1}{\ell}(\ell+2)n^{n-\ell-3}.\end{aligned} \tag{4.2.3}$$

By the binomial theorem we have

$$\sum_{\ell=0}^{j-i-1}\binom{j-i-1}{\ell}x^{\ell} = (1+x)^{j-i-1}.$$

Multiplying through by x^2 and differentiating with respect to x we get

$$\sum_{\ell=0}^{j-i-1}\binom{j-i-1}{\ell}(\ell+2)x^{\ell+1} = x^2(j-i-1)(1+x)^{j-i-2} + 2x(1+x)^{j-i-1}.$$

We therefore have

$$\begin{aligned}&n^{n-2}\sum_{\ell=0}^{j-i-1}\binom{j-i-1}{\ell}(\ell+2)x^{\ell+1}\\ &\qquad = n^{n-2}(x^2(j-i-1)(1+x)^{j-i-2} + 2x(1+x)^{j-i-1}).\end{aligned}$$

Substituting $x = n^{-1}$, we obtain

$$\begin{aligned}&\sum_{\ell=0}^{j-i-1}\binom{j-i-1}{\ell}(\ell+2)n^{n-\ell-3}\\ &\qquad = n^{n-2}\left[\frac{1}{n^2}(j-i-1)\left(1+\frac{1}{n}\right)^{j-i-2} + \frac{2}{n}\left(1+\frac{1}{n}\right)^{j-i-1}\right].\end{aligned} \tag{4.2.4}$$

By Equations (4.2.3) and (4.2.4), it follows that

$$\sum_{\ell} R(n,i,j,\ell) = (2n+j-i+1)n^{n-j+i-2}(n+1)^{j-i-2}.$$

This completes the proof. □

Corollary 4.2.3. *Let $R(n,i,\ell)$ be the total number of vertices that can be reached from vertex i in ℓ steps in labelled trees on n vertices. Then we have*

$$R(n,i,\ell) = \binom{n-i}{\ell}(\ell+1)n^{n-\ell-2}. \tag{4.2.5}$$

Proof. Let us consider trees on n vertices such that there is a path of length ℓ starting at vertex i and ending at vertex j. By Proposition 4.2.1, there are

$$\binom{j-i-1}{\ell-1}(\ell+1)n^{n-\ell-2}$$

such trees. Since the range of j is $i+\ell \leq j \leq n$, the number of vertices reachable in ℓ steps from vertex i is thus

$$\sum_{j=i+\ell}^{n}\binom{j-i-1}{\ell-1}(\ell+1)n^{n-\ell-2} = \sum_{k=\ell-1}^{n-i-1}\binom{k}{\ell-1}(\ell+1)n^{n-\ell-2}.$$

By the identity

$$\sum_{j=p}^{n}\binom{j}{p} = \binom{n+1}{p+1}, \tag{4.2.6}$$

we obtain

$$\sum_{k=\ell-1}^{n-i-1}\binom{k}{\ell-1}(\ell+1)n^{n-\ell-2} = \binom{n-i}{\ell}(\ell+1)n^{n-\ell-2},$$

which completes the proof. □

By substituting for $\ell = 1$ and $i = 1$ in (4.2.5), we obtain:

Corollary 4.2.4. *There are a total of $2(n-1)n^{n-3}$ children of vertex 1 in all labelled trees of order n.*

The sequence $1, 4, 24, 200, \ldots$ in Corollary 4.2.4 appears in the OEIS [55] as A089946.

Corollary 4.2.5. *Let $R(n,i)$ be the total number of vertices that can be reached from vertex i in labelled trees on n vertices. Then we have*

$$R(n,i) = n^{i-2}(n+1)^{n-i-1}(2n-i+1). \tag{4.2.7}$$

Proof. We sum over all ℓ in Equation (4.2.5) to obtain

$$R(n,i) = \sum_{\ell=0}^{n-i} \binom{n-i}{\ell} (\ell+1) n^{n-\ell-2}. \tag{4.2.8}$$

By the binomial theorem we have

$$\sum_{\ell=0}^{n-i} \binom{n-i}{\ell} x^{\ell} = (1+x)^{n-i}.$$

Multiplying through by x and differentiating with respect to x we get

$$\sum_{\ell=0}^{n-i} \binom{n-i}{\ell} (\ell+1) x^{\ell} = x(n-i)(1+x)^{n-i-1} + (1+x)^{n-i}.$$

We therefore have

$$n^{n-2} \sum_{\ell=0}^{n-i} \binom{n-i}{\ell} (\ell+1) x^{\ell} = n^{n-2}(x(n-i)(1+x)^{n-i-1} + (1+x)^{n-i}).$$

Substituting $x = n^{-1}$, we obtain

$$\sum_{\ell=0}^{n-i} \binom{n-i}{\ell} (\ell+1) n^{n-\ell-2} = n^{n-2} \left[\frac{1}{n}(n-i) \left(1+\frac{1}{n}\right)^{n-i-1} + \left(1+\frac{1}{n}\right)^{n-i} \right]. \tag{4.2.9}$$

From Equations (4.2.8) and (4.2.9) we obtain the required equation. □

Corollary 4.2.6. *For any fixed i, the average number of vertices that can be reached from vertex i in labelled trees of order n tends to $2e$ as n tends to infinity.*

Proof. By dividing Equation (4.2.7) by n^{n-2}, the number of Cayley trees on n vertices, we obtain

$$n^{i-n}(n+1)^{n-i-1}(2n-i+1) \tag{4.2.10}$$

as the average number of vertices that can be reached from vertex i in labelled trees of order n.

Now,

$$\lim_{n\to\infty}\left(n^{i-n}(n+1)^{n-i-1}(2n-i+1)\right)=\lim_{n\to\infty}\left[\left(1+\frac{1}{n}\right)^{n-i-1}\left(2-\frac{i}{n}+\frac{1}{n}\right)\right]$$
$$=\lim_{n\to\infty}\left[\left(1+\frac{1}{n}\right)^{n}\left(1+\frac{1}{n}\right)^{-i-1}\left(2-\frac{i}{n}+\frac{1}{n}\right)\right]$$
$$=2e.$$

□

Corollary 4.2.7. *Let $R(n,i,\ell)$ be the total number of vertices that can be reached from vertex i in ℓ steps in labelled trees on n vertices. Then we have*

$$\sum_i R(n,i,\ell)=\binom{n-1}{\ell}n^{n-\ell-1}. \tag{4.2.11}$$

Proof. We sum over all i in Equation (4.2.5),

$$\sum_{i=1}^{n}\binom{n-i}{\ell}(\ell+1)n^{n-\ell-2}=(\ell+1)n^{n-\ell-2}\sum_{i=1}^{n}\binom{n-i}{\ell}.$$

By identity (4.2.6) we have

$$\sum_{i=1}^{n}\binom{n-i}{\ell}(\ell+1)n^{n-\ell-2}=(\ell+1)n^{n-\ell-2}\binom{n}{\ell+1}=n^{n-\ell-1}\binom{n-1}{\ell}.$$

□

Since Equation (4.2.11) is the number of labelled trees on $n+1$ nodes with a fixed node of degree $\ell+1$ where $0<\ell<n$, (see [11, Theorem F] for details), Corollary 4.2.7 will also be proved by constructing a bijection between the set of labelled trees on $n+1$ vertices where a given vertex, say v, has degree $\ell+1$ and the set of oriented paths of length ℓ in labelled trees on n vertices. We now give this bijection:

Proposition 4.2.8. *There is a bijection between the set of labelled trees on $n+1$ nodes with a fixed node of degree $\ell+1$ where $0<\ell<n$, and the set of oriented paths of length ℓ in labelled trees on n vertices.*

Proof. Consider a labelled tree on n vertices with a path of length ℓ. Create edges between all vertices on the path and a new vertex $n+1$. We then delete all the edges that were on the path. The degree of vertex $n+1$ is therefore $\ell + 1$. The reverse procedure is as follows: Delete vertex $n+1$ and all the edges incident to it. Now, create a path from the vertex with the smallest label that was initially attached to vertex $n+1$ to the other vertex with highest label such that the labels decrease along the path. Since vertex $n+1$ had degree $\ell+1$, the path created has length ℓ. □

Corollary 4.2.9. *There are a total of $(n+1)^{n-1}$ oriented paths in trees on $[n]$.*

Proof. The result follows by summing over all ℓ in Equation (4.2.11). □

Proposition 4.2.10. *The number of trees of order n in which vertex j is a sink reachable from label i in ℓ steps is*

$$\binom{j-i-1}{\ell-1}(n-1)^{n-j-1}n^{j-\ell-2}(\ell n + j - \ell - 1). \tag{4.2.12}$$

Proof. Let us consider trees on n vertices such that there is a path of length ℓ starting at vertex i and ending at vertex j. Let the number of neighbours of j (different from i), all of label less than j be r. There are $\binom{j-i-1}{\ell-1}$ possible paths and $\binom{j-\ell-1}{r}$ choices for the r neighbours of j. The number of forests on $n-1$ vertices rooted at the ℓ vertices of the path (with the exception of j) and the r neighbours of j is

$$(\ell+r)(n-1)^{n-\ell-r-2}.$$

Therefore, there are

$$A = \sum_{r\geq 0}\binom{j-i-1}{\ell-1}\binom{j-\ell-1}{r}(\ell+r)(n-1)^{n-\ell-r-2}$$

trees in which sink j can be reached from vertex i in ℓ steps.

Now,

$$\begin{aligned} A &= (n-1)^{n-\ell-2}\binom{j-i-1}{\ell-1}\sum_{r=0}^{j-\ell-1}\binom{j-\ell-1}{r}(\ell+r)(n-1)^{-r} \\ &= \ell(n-1)^{n-\ell-2}\binom{j-i-1}{\ell-1}\sum_{r=0}^{j-\ell-1}\binom{j-\ell-1}{r}(n-1)^{-r} \\ &+ (n-1)^{n-\ell-2}\binom{j-i-1}{\ell-1}\sum_{r=0}^{j-\ell-1}\binom{j-\ell-1}{r}r(n-1)^{-r}, \end{aligned}$$

which is the same as

$$\begin{aligned}
A &= \ell(n-1)^{n-\ell-2}\binom{j-i-1}{\ell-1}\sum_{r=0}^{j-\ell-1}\binom{j-\ell-1}{r}(n-1)^{-r} \\
&+ (j-\ell-1)(n-1)^{n-\ell-2}\binom{j-i-1}{\ell-1}\sum_{r=1}^{j-\ell-1}\binom{j-\ell-2}{r-1}(n-1)^{-r} \\
&= \ell(n-1)^{n-\ell-2}\binom{j-i-1}{\ell-1}\sum_{r=0}^{j-\ell-1}\binom{j-\ell-1}{r}(n-1)^{-r} \\
&+ (j-\ell-1)(n-1)^{n-\ell-3}\binom{j-i-1}{\ell-1}\sum_{r=0}^{j-\ell-2}\binom{j-\ell-2}{r}(n-1)^{-r}.
\end{aligned}$$

By the binomial theorem we have

$$\begin{aligned}
A &= \ell(n-1)^{n-j-1}\binom{j-i-1}{\ell-1}n^{j-\ell-1} \\
&+ (j-\ell-1)(n-1)^{n-j-1}\binom{j-i-1}{\ell-1}n^{j-\ell-2} \\
&= \binom{j-i-1}{\ell-1}(n-1)^{n-j-1}n^{j-\ell-2}(\ell n+j-\ell-1).
\end{aligned}$$

□

Corollary 4.2.11. *The total number of labelled trees on n vertices such that a sink j is reachable from vertex $i \neq j$ is given by*

$$(n^2+2jn-in-2n+i-1)n^{i-2}(n-1)^{n-j-1}(n+1)^{j-i-2}.$$

Proof. We sum over all ℓ in Equation (4.2.12).

$$\begin{aligned}
B &= (n-1)^{n-j}n^{j-3}\sum_{\ell=1}^{j-i}\binom{j-i-1}{\ell-1}\ell n^{-(\ell-1)} \\
&+ (j-1)(n-1)^{n-j-1}n^{j-3}\sum_{\ell=1}^{j-i}\binom{j-i-1}{\ell-1}n^{-(\ell-1)} \\
&= (n-1)^{n-j}n^{j-3}\sum_{\ell=1}^{j-i}\binom{j-i-1}{\ell-1}(\ell-1)n^{-(\ell-1)} \\
&+ (n-1)^{n-j}n^{j-3}\sum_{\ell=1}^{j-i}\binom{j-i-1}{\ell-1}n^{-(\ell-1)} \\
&+ (j-1)(n-1)^{n-j-1}n^{j-3}\sum_{\ell=1}^{j-i}\binom{j-i-1}{\ell-1}n^{-(\ell-1)}.
\end{aligned}$$

Thus

$$\begin{aligned}B &= (j-i-1)(n-1)^{n-j}n^{j-3}\sum_{\ell=0}^{j-i-2}\binom{j-i-2}{\ell}n^{-\ell}\\&+(n-1)^{n-j}n^{j-3}\sum_{\ell=0}^{j-i-1}\binom{j-i-1}{\ell}n^{-\ell}\\&+(j-1)(n-1)^{n-j-1}n^{j-3}\sum_{\ell=0}^{j-i-1}\binom{j-i-1}{\ell}n^{-\ell}.\end{aligned}$$

By the binomial theorem we get

$$\begin{aligned}B &= (j-i-1)(n-1)^{n-j}n^{i-1}(n+1)^{j-i-2}+(n-1)^{n-j}n^{i-2}(n+1)^{j-i-1}\\&+(j-1)(n-1)^{n-j-1}n^{i-2}(n+1)^{j-i-1}\\&= n^{i-2}(n-1)^{n-j-1}(n+1)^{j-i-2}\left[n^2+2jn-in-2n+i-1\right].\end{aligned}$$

This completes the proof. □

The following result follows by summing over all j in Equation (4.2.12).

Corollary 4.2.12. *The total number of sinks that are reachable from vertex i in ℓ steps in labelled trees on n vertices is given by*

$$\sum_{j=i+\ell}^{n}\binom{j-i-1}{\ell-1}(n-1)^{n-j-1}n^{j-\ell-2}(\ell n+j-\ell-1).$$

Corollary 4.2.13. *The total number of sinks in labelled trees having vertex i as a sink is given by*

$$(i-1)n^{i-2}(n-1)^{n-i-2}. \tag{4.2.13}$$

Proof. The formula follows by setting $\ell=0$ and $j=i$ in Equation (4.2.12). □

Corollary 4.2.14. *There are a total of $(n-1)^{n-1}$ sinks in trees of order n.*

Proof. The formula follows by summing over all i in Equation (4.2.13). □

Lemma 4.2.15. *The total number of trees of order n in which vertex j is a leaf sink reachable from vertex i in ℓ steps is*

$$\ell\binom{j-i-1}{\ell-1}(n-1)^{n-\ell-2}. \tag{4.2.14}$$

Proof. The choices for the paths of length ℓ starting at i and ending in j is $\binom{j-i-1}{\ell-1}$. Ignoring the edges on the path, a forest on $n-1$ labelled vertices rooted at the ℓ vertices of the path (with the exception of vertex j) remains. The number of such forests is $\ell(n-1)^{n-\ell-2}$. Therefore, the total number of oriented paths starting from vertex i such that the end vertex is a leaf sink is:

$$\ell\binom{j-i-1}{\ell-1}(n-1)^{n-\ell-2}.$$

□

The following result follows immediately by summing over all j in Equation (4.2.14).

Corollary 4.2.16. *There are a total of*

$$\ell\binom{n-i}{\ell}(n-1)^{n-\ell-2} \tag{4.2.15}$$

leaf sinks reachable from vertex i in ℓ steps in labelled trees of order n.

Corollary 4.2.17. *Let $L(n,i)$ be the total number of leaf sinks that can be reached from vertex i in labelled trees on n vertices. We have*

$$L(n,i) = (n-i)n^{n-i-1}(n-1)^{i-2}. \tag{4.2.16}$$

Proof. We sum over all i in Equation (4.2.15) to obtain

$$L(n,i) = \sum_{\ell=0}^{n-i} \ell\binom{n-i}{\ell}(n-1)^{n-\ell-2}.$$

By the binomial theorem we have

$$\sum_{\ell=0}^{n-i}\binom{n-i}{\ell}x^{\ell} = (1+x)^{n-i}.$$

Differentiating with respect to x we get

$$\sum_{\ell=0}^{n-i}\ell\binom{n-i}{\ell}x^{\ell-1} = (n-i)(1+x)^{n-i-1}.$$

We therefore have

$$(n-1)^{n-3}\sum_{\ell=0}^{n-i}\ell\binom{n-i}{\ell}x^{\ell-1} = (n-1)^{n-3}(n-i)(1+x)^{n-i-1}.$$

Substituting $x = (n-1)^{-1}$, we obtain

$$\sum_{\ell=0}^{n-i} \ell \binom{n-i}{\ell} (n-1)^{n-\ell-2} = (n-1)^{n-3}(n-i)\left(1+\frac{1}{n-1}\right)^{n-i-1}. \quad (4.2.17)$$

We simplify Equation (4.2.17) to obtain the required equation. □

Corollary 4.2.18. *For any fixed i, the average number of leaf sinks that can be reached from vertex i in labelled trees of order n tends to 1 as n tends to infinity.*

Proof. By dividing Equation (4.2.16) by n^{n-2}, the number of Cayley trees on n vertices, we obtain

$$(n-i)n^{1-i}(n-1)^{i-2} \quad (4.2.18)$$

as the average number of leaf sinks that can be reached from vertex i in labelled trees of order n. Now,

$$\lim_{n\to\infty}\left((n-i)n^{1-i}(n-1)^{i-2}\right) = \lim_{n\to\infty}\left[\left(1-\frac{i}{n}\right)\left(1-\frac{1}{n}\right)^{i-2}\right] = 1.$$

□

4.3 Trees with a given number of reachable vertices

Theorem 4.3.1. *The total number of rooted labelled trees of order n such that exactly k vertices are reachable from the root is given by*

$$\sum_{j=0}^{k-1}(-1)^j\binom{k-1}{j}(n-j-1)^{n-1}. \quad (4.3.1)$$

A subtree of an r-rooted tree T is said to be *decreasing* (*increasing*) if the labels in the subtree are decreasing (increasing) as one moves away from the root. A *maximal decreasing* (*increasing*) subtree [7], is a decreasing (increasing) subtree rooted at vertex r with the highest number of vertices. Seo and Shin [52] showed that Equation (4.3.1) gives the number of rooted Cayley trees on n vertices whose maximal decreasing subtree has exactly k vertices. By symmetry, it is clear that Equation (4.3.1) also gives the number of rooted Cayley trees on n vertices whose maximal increasing subtree has k vertices.

If we orient the edges from vertices of lower label towards vertices of higher label in the rooted Cayley trees, a maximal increasing subtree of order k has exactly k reachable vertices. This proves Theorem 4.3.1.

Milan Janjić [31, Theorem 2.1] showed that Equation (4.3.1) is the number of functions from an $(n-1)$-set to another $(n-1)$-set whose images contain $k-1$ fixed elements. In [33, Theorem 1.3], J. S. Kim constructed a bijection between the set of these functions and the set of rooted labelled trees of order n having maximal decreasing subtrees of order k. He thus also proved Equation (4.3.1).

Corollary 4.3.2. *The number of labelled trees of order n having exactly k vertices reachable from vertex* 1 *is given by*

$$\sum_{j=0}^{k-2}(-1)^j\binom{k-2}{j}(k-1)(n-j-2)^{n-2}. \tag{4.3.2}$$

Proof. By Theorem 4.3.1, there are

$$\sum_{j=0}^{k-2}(-1)^j\binom{k-2}{j}(n-j-2)^{n-2}$$

rooted labelled trees on $[2..n]$ such that exactly $k-1$ vertices are reachable from the root. We follow the following steps to obtain trees in which k vertices are reachable from vertex 1:

Step 1: Consider a v_1-rooted tree T with a maximal increasing subtree T_0 whose vertex-set is $\{v_1, v_2, \ldots, v_{k-1}\}$, where $v_i < v_{i+1}$ for all i. In T, delete all the edges belonging to T_0 so as to obtain non-single-vertex subtrees: $T_1, T_2, \ldots, T_r$.

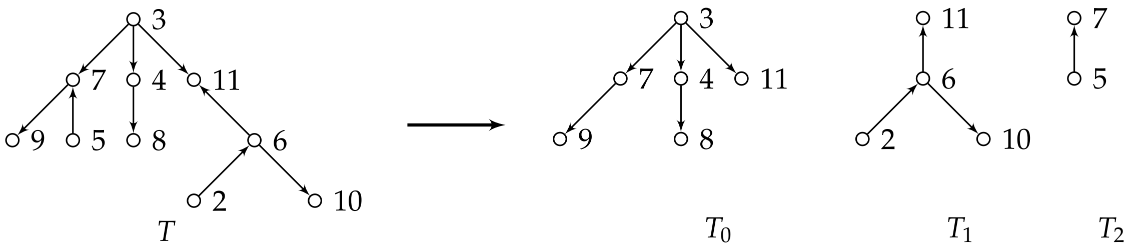

Figure 4.1: Diagram showing Step 1 in the proof of Corollary 4.3.2

Step 2: Relabel the root in T_0 as 1, vertex v_2 as v_1, vertex v_3 as v_2, and so on. Now, attach vertex v_{k-1} to any of the $k-1$ vertices in the new maximal increasing subtree rooted at vertex 1. For each maximal increasing subtree initially rooted at vertex v_1 we obtain $k-1$ new such subtrees rooted at vertex 1 with k reachable vertices.

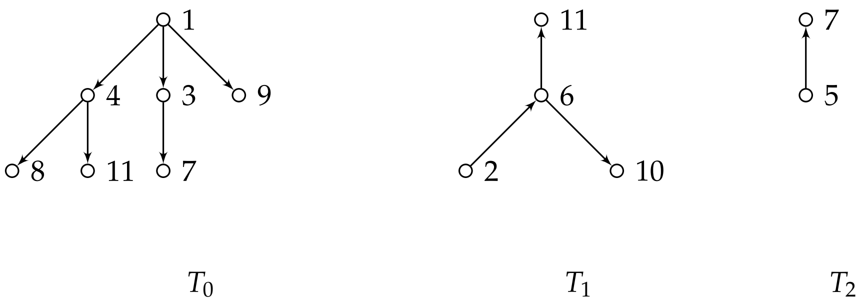

Figure 4.2: Diagram showing Step 2 in the proof of Corollary 4.3.2

Step 3: Identify vertex v_i in the subtrees $T_1, T_2, \ldots, T_r$ with vertex v_i in the new maximal increasing subtree, for all $i \in \{1, \ldots, k-1\}$ for which v_i occurs in one of the T_j.

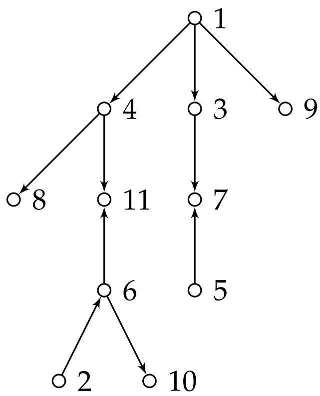

Figure 4.3: Diagram showing Step 3 in the proof of Corollary 4.3.2

Therefore, the total number of trees with exactly k vertices reachable from vertex 1 is

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

□

Proposition 4.3.3. *Let $E(n,i)$ be the total number of labelled trees on n vertices with exactly $n-i+1$ vertices reachable from vertex i. We have*

$$E(n,i) = (n-i+1)!n^{i-2}.$$

Proof. Since there are $n-(n-i+1) = (i-1)$ vertices which are not reachable from vertex i, these vertices must have labels $1,\ldots,i-1$. The reachable vertices therefore have labels $i,\ldots,n$ forming $(n-i)!$ recursive trees [38, Section 7]. The total number of forests on n vertices rooted at the $n-i+1$ vertices is $(n-i+1)n^{n-(n-i+1)-1}$. Thus,

$$E(n,i) = (n-i+1)n^{i-2} \cdot (n-i)!.$$

□

Theorem 4.3.4. *The number of labelled trees of order n having exactly k vertices reachable from vertex j_1 is given by*

$$\frac{(n-k)^n}{k\binom{n}{k}} \sum_{j_2=j_1+1}^{n-k+2} \sum_{j_3=j_2+1}^{n-k+3} \cdots \sum_{j_k=j_{k-1}+1}^{n} \left[\prod_{p=1}^{k} \left(\frac{n-p+1}{n-p} \right)^{j_p} \cdot \sum_{m=1}^{k} \frac{j_m - m}{(n-m)(n-m+1)} \right].$$

Proof. There are $(k-1)!$ recursive trees with vertex set $V = \{j_1 < j_2 < \cdots < j_k\}$. Let i_m be the number of neighbours of vertex j_m whose labels are less than j_m and do not belong to the set V.

We have $\binom{j_1-1}{i_1}$ choices for the i_1 neighbours of j_1. Excluding vertex j_1 and its neighbours, we have $\binom{j_2-i_1-2}{i_2}$ choices for the neighbours of j_2. In general, there are

$$\binom{j_m - i_1 - \cdots - i_{m-1} - m}{i_m}$$

choices for neighbours of j_m. There are thus

$$\binom{j_1-1}{i_1}\binom{j_2-i_1-2}{i_2}\cdots\binom{j_k-i_1-\cdots-i_{k-1}-k}{i_k}$$

possibilities for the choice of neighbours of elements of A. The number of forests on $n-k$ vertices (excluding the k elements in A), rooted at the neighbours of elements in A is given by

$$(i_1+\cdots+i_k)(n-k)^{n-k-1-i_1-\cdots-i_k}.$$

Hence the number of trees with exactly k reachable vertices from vertex j_1 is given by

$$(k-1)! \sum_{j_2=j_1+1}^{n-k+2} \sum_{j_3=j_2+1}^{n-k+3} \cdots \sum_{j_k=j_{k-1}+1}^{n} \sum_{i_1=0}^{j_1-1} \sum_{i_2=0}^{j_2-i_1-2} \cdots \sum_{i_k=0}^{j_k-i_1-\cdots-i_{k-1}-k} A_i \tag{4.3.3}$$

where

$$A_i = \binom{j_1-1}{i_1}\binom{j_2-i_1-2}{i_2}\cdots\binom{j_k-i_1-\cdots-i_{k-1}-k}{i_k}(i_1+i_2+\cdots+i_k) \times (n-k)^{n-k-(i_1+\cdots+i_k)-1}.$$

Let

$$B_t = \binom{j_1-1}{i_1}\binom{j_2-i_1-2}{i_2}\cdots\binom{j_t-i_1-\cdots-i_{t-1}-t}{i_t},$$

where $1 \le t \le k$.

Consider the following calculation:

$$\begin{aligned}\sum_{i_1,\ldots,i_k} B_k x^{i_1+\cdots+i_k} &= \sum_{i_1,\ldots,i_{k-1}} B_{k-1} x^{i_1+\cdots+i_{k-1}}(1+x)^{j_k-i_1-\cdots-i_{k-1}-k}\\ &= (1+x)^{j_k-k}\sum_{i_1,\ldots,i_{k-1}} B_{k-1}\left(\frac{x}{1+x}\right)^{i_1+\cdots+i_{k-1}}\\ &= (1+x)^{j_k-k}\left(1+\frac{x}{1+x}\right)^{j_{k-1}-(k-1)}\sum_{i_1,\ldots,i_{k-2}} B_{k-2}\left(\frac{x}{1+2x}\right)^{i_1+\cdots+i_{k-2}}.\end{aligned}$$

We continue with the iteration to obtain

$$\sum_{i_1,\ldots,i_k} B_k x^{i_1+\cdots+i_k} = \prod_{p=0}^{k-1}\left(1+\frac{x}{1+px}\right)^{j_{k-p}-(k-p)}. \tag{4.3.4}$$

Differentiating Equation (4.3.4) with respect to x we arrive at

$$\begin{aligned}&\sum_{i_1,\ldots,i_k} (i_1+\cdots+i_k)B_k x^{i_1+\cdots+i_k-1}\\ &\qquad= \prod_{p=0}^{k-1}\left(1+\frac{x}{1+px}\right)^{j_{k-p}-(k-p)}\sum_{m=0}^{k-1}\frac{j_{k-m}-(k-m)}{(1+mx)(1+(m+1)x)}.\end{aligned} \tag{4.3.5}$$

Multiplying Equation (4.3.5) by x and setting $x=(n-k)^{-1}$, we obtain

$$\begin{aligned}&\sum_{i_1,\ldots,i_k} (i_1+\cdots+i_k)B_k(n-k)^{-i_1-\cdots-i_k}\\ &= (n-k)\prod_{p=0}^{k-1}\left(\frac{n-k+p+1}{n-k+p}\right)^{j_{k-p}-(k-p)}\sum_{m=0}^{k-1}\frac{j_{k-m}-(k-m)}{(n-k+m)(n-k+m+1)}.\end{aligned}$$

Multiplying through by $(n-k)^{n-k-1}$ we get

$$\sum_{i_1,\dots,i_k} (i_1+\cdots+i_k)B_k(n-k)^{n-k-i_1-\cdots-i_k-1}$$
$$= (n-k)^{n-k}\prod_{p=0}^{k-1}\left(\frac{n-k+p+1}{n-k+p}\right)^{j_{k-p}-(k-p)}\sum_{m=0}^{k-1}\frac{j_{k-m}-(k-m)}{(n-k+m)(n-k+m+1)}. \tag{4.3.6}$$

Since the left hand side of Equation (4.3.6) is precisely

$$\sum_{i_1=0}^{j_1-1}\sum_{i_2=0}^{j_2-i_1-2}\cdots\sum_{i_k=0}^{j_k-i_1-\cdots-i_{k-1}-k} A_i$$

in Equation (4.3.3), we find that the number of trees in which exactly k vertices are reachable from vertex j_1 is

$$(n-k)^{n-k}(k-1)!\sum_{j_2=j_1+1}^{n-k+2}\sum_{j_3=j_2+1}^{n-k+3}\cdots\sum_{j_k=j_{k-1}+1}^{n} A$$

where

$$A=\prod_{p=0}^{k-1}\left(\frac{n-k+p+1}{n-k+p}\right)^{j_{k-p}-(k-p)}\cdot\sum_{m=0}^{k-1}\frac{j_{k-m}-(k-m)}{(n-k+m)(n-k+m+1)}.$$

This completes the proof. □

Chapter 5

Set partitions and tree-like structures

The relationship between labelled trees on n vertices and the symmetric group S_n of order n was discovered by Dénes [13] and subsequent work on the same has since been done by Moszkowski [40], Eden and Schützenberger [20], and Goulden and Pepper [27] among many other authors. In this chapter, we prove a formula for the number of cycle-free connected families of k set partitions of $[n]$. We show that the formula also counts free cacti with a given vertex-degree distribution as well as the number of coloured Husimi graphs with a given block-colour distribution.

5.1 Connected cycle-free families of set partitions

The following definition of cycle-free and connected families of set partitions is adapted from the previous work of Teufl and Wagner [60]. Let I be an index set. For $i \in I$, let S_i be a non-empty subset of a finite set S. The union of all S_i gives S. Now, let P_i be a set partition of S_i and let $\mathcal{P} = \{P_i : i \in I\}$ denote the family of such partitions. We define a multigraph $G_{\mathcal{P}}$ as having vertex set

$$\{(P, p) : P \in \mathcal{P}, p \in P\}$$

and two distinct vertices (P_1, p_1) and (P_2, p_2) are joined by $|p_1 \cap p_2|$ edges in the edge set. This ensures that $G_{\mathcal{P}}$ has no loops. The family $\mathcal{P}$ of set partitions is said to be *connected* if $G_{\mathcal{P}}$ is connected and *cycle-free* if the multigraph is cycle-free.

Theorem 5.1.1. *The number of connected, cycle-free families of k set partitions of $[n]$ is given by*

$$\frac{n!n^{k-2}\prod_{i=1}^{k}(\ell_i-1)!}{\prod_{i=1}^{k}\prod_{j\geq 1}(j-1)!^{a_{ij}}a_{ij}!} \tag{5.1.1}$$

where ℓ_i is the number of blocks in the i-th partition and a_{ij} is the number of blocks of size j in the i-th partition such that $\sum_{i=1}^{k}\ell_i=(k-1)n+1$.

Proof. The case $k=2$ was recently proved by Teufl and Wagner in [60]. Let us first consider the case in which all the k partitions are of the shape $m_i^1 1^{n-m_i}$ for $i=1,\ldots,k$ respectively. Since the partitions are cycle-free and connected, i.e. two blocks from distinct partitions have at most one point in common, the k set partitions form a 'tree' if we ignore the singletons. Such a tree will be called a *generalised tree*. Let y_i mark the number of blocks of size i in the generalised tree. Let $T(x)$ be the exponential generating function for the number of rooted labelled generalised trees. Now, the vertices in the block other than the root are given the structure of a 'forest' of rooted generalised trees. By the composition principle, the exponential generating function enumerating the forests is equal to

$$\exp\left(\sum_{i=1}^{\infty}y_{i+1}\frac{T(x)^i}{i!}\right).$$

By the product principle, we find that the exponential generating function for the number of rooted generalised trees satisfies

$$T(x)=x\exp\left(\sum_{i=1}^{\infty}y_{i+1}\frac{T(x)^i}{i!}\right).$$

By the Lagrange Inversion Formula [58, Theorem 5.4.2], we have

$$\begin{aligned}
[x^n]T(x) &= \frac{1}{n}[t^{n-1}]\exp\left(n\sum_{i=1}^{\infty}y_{i+1}\frac{t^i}{i!}\right)\\
&= \frac{1}{n}[t^{n-1}]\sum_{j\geq 0}\frac{n^j}{j!}\left(\sum_{i=1}^{\infty}y_{i+1}\frac{t^i}{i!}\right)^j\\
&= \frac{1}{n}[t^{n-1}]\sum_{j\geq 0}\frac{n^j}{j!}\left(y_2\frac{t}{1!}+y_3\frac{t^2}{2!}+y_4\frac{t^3}{3!}+\cdots\right)^j.
\end{aligned}$$

It follows that

$$\begin{aligned}
[x^n]T(x) &= \frac{1}{n}[t^{n-1}]\sum_{j\geq 0}\frac{n^j}{j!}\sum_{j_1+j_2+\cdots=j}\frac{j!}{j_1!j_2!\cdots}\left(y_2\frac{t}{1!}\right)^{j_1}\left(y_3\frac{t^2}{2!}\right)^{j_2}\cdots \\
&= \frac{1}{n}[t^{n-1}]\sum_{j\geq 0}\sum_{j_1+j_2+\cdots=j}\frac{n^j}{j_1!j_2!\cdots}\frac{y_2^{j_1}y_3^{j_2}\cdots}{(1!)^{j_1}(2!)^{j_2}\cdots}t^{j_1+2j_2+\cdots} \\
&= \sum_{j\geq 0}\sum_{\substack{j_1+j_2+\cdots=j\\ j_1+2j_2+\cdots=n-1}}\frac{n^{j-1}}{j_1!(1!)^{j_1}j_2!(2!)^{j_2}\cdots}y_2^{j_1}y_3^{j_2}\cdots.
\end{aligned}$$

The generating function for unrooted generalised trees on n vertices is therefore given by,

$$n!\prod_{i=1}^{\infty}\frac{y_{i+1}^{j_i}}{j_i!(i!)^{j_i}}n^{\sum j_i-2}.$$

We multiply by $n!$ since we are dealing with an exponential generating function. Since $\sum j_i = k$, we obtain that the number of families of k partitions of the shape $m_i^1 1^{n-m_i}$ for $i = 1,\ldots,k$ is given by

$$\frac{n!n^{k-2}}{\prod_{i=1}^{k}(m_i-1)!}.$$

From the $k = 2$ case (see [60] for details), we know that, given partitions $P_1, P_2, \ldots, P_{k-1}$ and the type of P_k, the number of possibilities for P_k to complete a connected cycle-free partition is of the form

$$\alpha(P_1, P_2, \ldots, P_{k-1})\cdot\beta(\text{type}(P_k)),$$

where β is known and only depends on the type of P_k (not $P_1, P_2, \ldots, P_{k-1}$). Thus changing the type λ_k to a type λ_k' only produces a factor $\dfrac{\beta(\lambda_k')}{\beta(\lambda_k)}$.

The number of families of two connected cycle-free set partitions of types $\lambda_0 = m_0^1 1^{n-m_0}$ and $\lambda_1 = 1^{a_{11}}2^{a_{12}}\cdots$ such that $\sum_j a_{1j} = \ell_1$ is given by

$$\gamma(\lambda_0,\lambda_1) = \frac{n!(\ell_1-1)!}{(m_0-1)!\prod_{j\geq 1}((j-1)!)^{a_{1j}}a_{1j}!}.$$

(See [60, Theorem 5.1].)

Similarly, if $\lambda_2 = m_1^1 1^{n-m_1}$ then,

$$\gamma(\lambda_0,\lambda_2) = \frac{n!}{(m_0-1)!(m_1-1)!}.$$

We therefore have

$$\frac{\gamma(\lambda_0,\lambda_1)}{\gamma(\lambda_0,\lambda_2)} = \frac{\beta(\lambda_1)}{\beta(\lambda_2)} = \frac{(m_1-1)!(\ell_1-1)!}{\prod_{j\geq 1}((j-1)!)^{a_{1j}} a_{1j}!}.$$

The number of connected cycle free families of k set partitions such that the first partition is $\lambda_1 = 1^{a_{11}} 2^{a_{12}} \cdots$ and the subsequent $k-1$ set partitions are of the shape $m_i^1 1^{n-m_i}$ for $i = 2, \ldots, k$ is therefore given by

$$\frac{n! n^{k-2}}{\prod_{i=1}^{k}(m_i - 1)!} \cdot \frac{(m_1-1)!(\ell_1-1)!}{\prod_{j\geq 1}((j-1)!)^{a_{1j}} a_{1j}!},$$

where ℓ_1 is the number of parts in λ_1, and a_{1j} is the number of parts of size j in λ_1. Iterating this idea we obtain the required equation. □

5.2 Enumeration of labelled free k-ary cacti

A *k-cactus* is a connected (simple) graph in which every edge lies on exactly one k-cycle. Such a cycle will be called a k-gon. A k-cactus is said to be *planar* if it is embedded in the plane such that every edge is part of the unbounded region. A planar k-cactus is *edge-rooted* if one of the edges is distinguished as a root and it is of *size n* if there are n k-gons. We colour the vertices of a k-cactus with k colours such that in a k-gon, colour i is used at most once. We say that $(\ell_1, \ldots, \ell_k)$ is a *vertex-colour distribution* of the k-cactus if there are ℓ_i vertices of colour i for all $i = 1, \ldots, k$. The number of k-gons that are incident to a vertex is the *degree* of such a vertex. We also say that $(a_{ij})_{1\leq i\leq k, j\geq 0}$ is a *vertex-degree distribution* of the k-cactus if $(\ell_1, \ldots, \ell_k)$, where $\ell_i = \sum_{j\geq 0} a_{ij}$ for all i, is its vertex-colour distribution and a_{ij} is the number of vertices of colour i and degree j. The following two lemmas are proved in [4].

Lemma 5.2.1. *There exits a k-ary cactus on N vertices and n k-gons if and only if $N = (k-1)n + 1$.*

Lemma 5.2.2. *Let $A = (a_{ij})_{1\leq i\leq k, j\geq 0}$ be a $k \times \infty$ matrix of non-negative integers, and set $N = \sum_{ij} a_{ij}$. There exists a k-ary cactus having N vertices and n k-gons and whose vertex-degree distribution is given by the matrix A if and only if*

1. *$n = (N-1)/(k-1)$ is an integer,*

2. *$\sum_j j a_{ij} = n$, for all i,*

3. $n \geq 1 \Rightarrow a_{i0} = 0$, *for all* i.

Let $\lambda = 1^{k_1}2^{k_2}\cdots$ be a partition of $n \geq 1$. We denote the length of the partition by $l(\lambda) = k_1 + k_2 + \cdots$. The following theorem is due to Goulden and Jackson [26]. The case $k = 2$ was proved by Bédard and Goupil [3] using an induction argument.

Theorem 5.2.3 ([26, Theorem 3.1]). *Let $\lambda_1, \ldots, \lambda_k$ be partitions of n such that $l(\lambda_1) + \cdots + l(\lambda_k) = (k-1)n + 1$. Then there is a bijection between planar edge-rooted k-cacti on n k-gons with colour i vertex-degree distribution λ_i, $i = 1, \ldots, k$, and k-tuples $(\alpha_1, \ldots, \alpha_k)$ of permutations in the symmetric group S_n with cycle distributions $\lambda_1, \ldots, \lambda_k$ respectively, such that $\alpha_1 \cdots \alpha_k = (1, 2, \ldots, n)$.*

Using Theorem 5.2.3, the aforementioned authors enumerated the number of connection coefficients for the symmetric group of order n by proving that there are

$$\frac{n^{k-1}\prod_{i=1}^{k}(\ell_i - 1)!}{\prod_{i=1}^{k}\prod_{j\geq 1} a_{ij}!}$$

plane edge-rooted k-cacti with n unlabelled vertices having a vertex-degree distribution $(a_{ij})_{1\leq i\leq k, j\geq 0}$, where $\sum_{i=1}^{k}\ell_i = \sum_{i=1}^{k}\sum_{j\geq 0} a_{ij} = (k-1)n + 1$. For $k = 2$, they obtained an equivalent result for the number of plane edge-rooted trees.

Bóna, Bousquet, Labelle and Leroux [4] defined a *free k-ary* cactus informally as a k-ary cactus without the plane embedding. They obtained a formula [4, Proposition 27] for the number of labelled free k-ary cacti on n k-gons having vertex-color distribution $(\ell_1, \ldots, \ell_k)$ as

$$n^{k-2}\prod_{i=1}^{k}\frac{(\ell_i - 1)!\,\ell_i^{n-\lambda_i}}{(n-\ell_i)!}$$

if $\sum_{i=1}^{k}\ell_i = (k-1)n + 1$.

We now give a different proof of these results based on Theorem 5.1.1:

Proposition 5.2.4. *The number of labelled free k-ary cacti on n k-gons having vertex-degree distribution $(a_{ij})_{1\leq i\leq k, j\geq 0}$ is given by*

$$\frac{n!\,n^{k-2}\prod_{i=1}^{k}(\ell_i - 1)!}{\prod_{i=1}^{k}\prod_{j\geq 1}(j-1)!^{a_{ij}}a_{ij}!} \tag{5.2.1}$$

where $\sum_{i=1}^{k}\sum_{j\geq 0} a_{ij} = \sum_{i=1}^{k}\ell_i = (k-1)n + 1$.

Proof. To prove this proposition, we need to show that there is a bijection between the set of labelled free k-cacti on n k-gons having a vertex-degree distribution $(a_{ij})_{1\leq i\leq k, j\geq 0}$ such that $\sum_{i=1}^{k}\sum_{j\geq 0} a_{ij} = (k-1)n+1$ and the set of families of k connected cycle-free set partitions of $[n]$.

Consider a labelled free k-cactus on n k-gons with labels from the set $[n]$ and having vertex-degree distribution (a_{ij}) for $i = 1,\ldots,k$ and $j \geq 0$. The degree distribution of vertices of colour i can therefore be written as $\lambda_i = 1^{a_{i1}}2^{a_{i2}}\cdots$. Let s_i be a set of labels of k-gons that are incident to a vertex s of colour i. Let

$$S_i = \{s_i : s \text{ is a vertex of colour } i\}.$$

We shall show that S_i is a set partition of $[n]$. By Lemma 5.2.2, $\sum_j ja_{ij} = n$ and $a_{i0} = 0$. This ensures that there is no empty set in S_i. The sets s_i are disjoint since the labels in $[n]$ are used only once in the labelled free cactus. Therefore λ_i is a partition of the set $[n]$. By Lemma 5.2.1, the cacti satisfy the condition $\sum_{i,j} a_{ij} = (k-1)n+1$. It follows that the partitions also satisfy this coherence condition. Since the n k-gons are connected and cycle-free then the partitions are also cycle-free and connected.

Next, we describe the reverse procedure of getting free k-cacti on n k-gons from a family of k set partitions of $[n]$ that are connected and cycle-free. Consider set partitions $\lambda_1,\ldots,\lambda_k$ of $[n]$ that are connected and cycle-free such that $\sum_{i=1}^{k}\ell(\lambda_i) = (k-1)n+1$. There is a bijection between these set partitions and cacti-like tree such that the k set partitions correspond to k vertices per block in the cacti-like tree, blocks correspond to vertices and block sizes to degrees in the cacti-like tree. Now, each block in the cacti-like tree corresponds to a k-gon in a free cactus. □

The bijection in Proposition 5.2.4 is shown by Figure 5.1, where the three partitions are shown by normal curves, dotted curves and dashed curves. From the cacti-like tree in Figure 5.1, we obtain a labelled free ternary cactus in Figure 5.2.

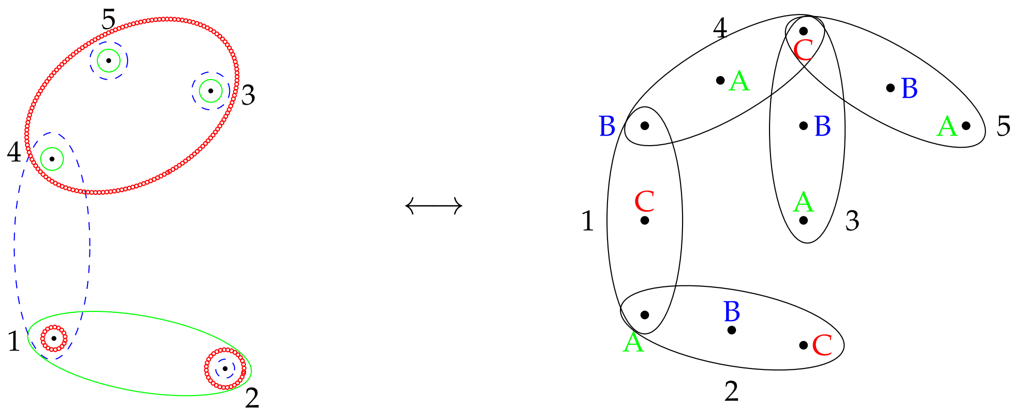

Figure 5.1: Bijection: Connected cycle-free family of 3 set partitions of [5] and cacti-like tree with 5 blocks, each having 3 vertices.

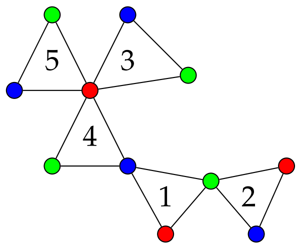

Figure 5.2: Labelled free 3-ary cactus

5.3 Coloured Husimi graphs, cacti and oriented cacti

Definition 5.3.1 ([35])**.** A *cutpoint* of a connected graph G is a vertex whose removal will disconnect G. A graph which has no cutpoint is said to be 2-connected. A *block* in a simple graph is a maximal 2-connected subgraph.

Husimi graphs were introduced by Japanese physicist Kodi Husimi in [30]. A *Husimi graph* is a connected graph whose blocks are complete graphs. If the blocks of a connected graph are polygons then the graph is called a *cactus*. Cacti were introduced by Harary and Uhlenbeck in [29] where they appeared as 'Husimi trees'. In 1996, Collin Springer [56] introduced and studied *oriented cacti*. These are connected graphs whose blocks are oriented cycles.

Lemma 5.3.2 ([35, Proposition 3.1])**.** *The number of Husimi graphs on* $[n]$ *where*

$n \geq 1$ is given by

$$\sum_{k\geq 0} \left\{ {n-1 \atop k} \right\} n^{k-1}, \tag{5.3.1}$$

where k is the number of blocks.

Lemma 5.3.3 (Husimi [30], Mayer [37]). *Let $(n_2, n_3, \ldots)$ be a sequence of non-negative integers satisfying the condition that $n = \sum_{j\geq 2}(j-1)n_j + 1$. The number $\mathrm{HG}_n(n_2, n_3, \cdots)$ of Husimi graphs on $[n]$ having n_j blocks of size j is given by*

$$\mathrm{HG}_n(n_2, n_3, \cdots) = \frac{(n-1)!n^{k-1}}{\prod_{j\geq 2}((j-1)!)^{n_j} n_j!}, \tag{5.3.2}$$

where $k = \sum_{j\geq 2} n_j$.

Formula (5.3.2) was proved by Husimi in [30] by first establishing a recurrence equation satisfied by the graphs. Later on, Mayer [37] provided a direct proof and Leroux [35] proved the formula from a generating function approach.

Lemma 5.3.4 (Ford and Uhlenbeck [23]). *Let $(n_2, n_3, \ldots)$ be a sequence of non-negative integers satisfying the condition that $n = \sum_{j\geq 2}(j-1)n_j + 1$. The number of cacti on $[n]$ having n_j polygons of size j is given by*

$$\frac{1}{2^{\sum_{i\geq 3} n_i}} \frac{(n-1)!}{\prod_{i\geq 2} n_i!} n^{k-1}, \tag{5.3.3}$$

where $k = \sum_{j\geq 2} n_j$.

Equation (5.3.2) was given in [29] as the formula for counting the number of cacti in the statement of Lemma 5.3.4. This incorrect solution was pointed out in [23] and corrected therein.

Lemma 5.3.5 ([35, Proposition 3.6]). *The number of labelled cacti on $[n]$, where $n \geq 2$, is given by*

$$\sum_{k\geq 0} \sum_{\substack{n_2+n_3+\cdots=k \\ n_2+2n_3+\cdots=n-1}} \frac{(n-1)!n^{k-1}}{2^{\sum_{i\geq 3} n_i} \prod_{i\geq 2} n_i!}. \tag{5.3.4}$$

Lemma 5.3.6 (Springer [56]). *Let $(n_2, n_3, \ldots)$ be a sequence of non-negative integers and $n = \sum_{j\geq 2}(j-1)n_j + 1$. The number $\mathrm{OC}_n(n_2, n_3, \cdots)$, of oriented cacti on $[n]$ having n_j oriented cycles of size j is given by*

$$\mathrm{OC}_n(n_2, n_3, \cdots) = \frac{(n-1)!}{\prod_{i\geq 2} n_i!} n^{k-1}, \tag{5.3.5}$$

where $k = \sum_{j\geq 2} n_j$.

Comparing Equations (5.3.2) and (5.3.5), we remark that

$$\mathrm{OC}_n(n_2, n_3, \cdots) = \prod_{i\geq 2}((i-1)!)^{n_i} \cdot \mathrm{HG}_n(n_2, n_3, \cdots). \tag{5.3.6}$$

Lemma 5.3.7 ([35, Proposition 3.3]). *The number of oriented cacti on $[n]$, where $n \geq 2$, is given by*

$$\sum_{k\geq 1} \frac{(n-1)!}{k!} \binom{n-2}{k-1} n^{k-1},$$

where k is the number of cycles.

Definition 5.3.8. A *coloured Husimi graph* is a Husimi graph whose blocks are coloured such that no blocks of the same colour are incident to one another. In the same manner, a *coloured cactus* (*coloured oriented cactus*) is a cactus (oriented cactus) whose polygons (oriented cycles) of the same colour are not incident to one another.

For the remainder of this section, we find formulas for the number of these coloured structures. Let a_{ij} denote the number of blocks of size j and colour i.

Proposition 5.3.9. *Let $(n_2, n_3, \ldots)$ be a sequence of non-negative integers satisfying the condition: $n = \sum_{j\geq 2}(j-1)n_j + 1$. The number of coloured Husimi graphs on $[n]$ having n_j blocks of size j and block-colour distribution $(a_{ij})_{1\leq i\leq k, j\geq 2}$ is given by*

$$\frac{n!n^{k-2}\prod_{i=1}^{k}(\ell_i - 1)!}{\prod_{i=1}^{k}\prod_{j\geq 1}(j-1)!^{a_{ij}} a_{ij}!} \tag{5.3.7}$$

where $a_{i1} = n - \sum_{j\geq 2} a_{ij}$ and $\ell_i = \sum_{j\geq 1} a_{ij}$.

Proof. We shall show that there is a bijection between the set of these graphs and the set of families of connected cycle-free k set partitions of $[n]$ such that there are a_{ij} blocks of size j in the i-th partition satisfying the coherence condition $\sum_{i=1}^{k} \sum_{j \geq 1} a_{ij} = (k-1)n + 1$.

Given the coloured Husimi graph, the number of blocks of colour i is given by $\sum_{j \geq 2} a_{ij}$. Since $a_{i1} = n - \sum_{j \geq 2} a_{ij}$, we have that $\sum_{j \geq 1} a_{ij} = n$. Let B_{ij} be the set of vertices in the j-th block of colour i. Since the blocks of the same colour are not incident to one another, the B_{ij}'s are disjoint. The vertices which are not part of any block are taken as singletons. Thus the family of sets comprising the singletons and the blocks do not comprise of empty sets. Hence $(a_{ij})_{j \geq 1}$ is a partition of $[n]$. We have k such partitions since $1 \leq i \leq k$. The set partitions satisfy the condition $\sum_{i=1}^{k} \sum_{j \geq 1} a_{ij} = (k-1)n + 1$.

Let us state the reverse procedure. Consider families of k cycle-free connected set partitions of $[n]$ in which there are ℓ_i blocks in the i-th partition and a_{ij} blocks of size j in the i-th partition such that $\sum_{i=1}^{k} \ell_i = (k-1)n + 1$. Allow edges between two distinct vertices in the block and remove all singletons. If v is a block in the i-th partition then colour the resultant complete graph with colour i. We thus obtain a coloured Husimi graph satisfying the condition $n = \sum_{j \geq 2} (j-1) n_j + 1$ where $n_j = \sum_{i=1}^{k} a_{ij}$. □

Using (5.3.6), we obtain that

Corollary 5.3.10. *Let $(n_2, n_3, \ldots)$ be a sequence of non-negative integers satisfying the condition: $n = \sum_{j \geq 2} (j-1) n_j + 1$. The number of coloured oriented cacti on $[n]$ having n_j cycles of size j and block-colour distribution $(a_{ij})_{1 \leq i \leq k, j \geq 2}$ is given by*

$$\frac{n! n^{k-2} \prod_{i=1}^{k} (\ell_i - 1)!}{\prod_{i=1}^{k} \prod_{j \geq 1} a_{ij}!}.$$

And for the non-oriented case,

Corollary 5.3.11. *Let $(n_2, n_3, \ldots)$ be a sequence of non-negative integers satisfying the condition: $n = \sum_{j \geq 2} (j-1) n_j + 1$. The number of coloured cacti on $[n]$ having n_j polygons of size j and block-colour distribution $(a_{ij})_{1 \leq i \leq k, j \geq 2}$ is given by*

$$\frac{n! n^{k-2} \prod_{i=1}^{k} (\ell_i - 1)!}{2^{\sum_{j \geq 3} n_j} \prod_{i=1}^{k} \prod_{j \geq 1} a_{ij}!}.$$

5.4 Families with given degree sequence

Let the degree of vertex i in the Husimi graph be the number of blocks incident to it. Let $(n_2, n_3, \ldots)$ be a sequence of non negative integers and also let $n = \sum_{j\geq 2}(j-1)n_j + 1$. We are now going to adapt the Prüfer correspondence obtained by Springer [56] for labelled oriented cacti to Husimi graphs.

A *leaf-block* l of a Husimi graph h is a block of h containing exactly one cutpoint, denoted by c_l. Set

$$L = \{x : x \text{ is a vertex in the leaf-block } l \text{ in the Husimi graph } h\}.$$

Let $l(h)$ be the leaf-block of h for which the set $L(h) \setminus \{c_{l(h)}\}$ contains the smallest element among all sets of the forms $L \setminus \{c_l\}$.

Now to each Husimi graph h on $[n]$ having k blocks we assign a pair (s, λ), where s is a sequence $(c_1, \ldots, c_{k-1})$ of elements of $[n]$ of length $k-1$, and λ is a partition of the set $[n] \setminus \{c_{k-1}\}$ into k parts. The Prüfer correspondence proceeds recursively as follows:

I. Add the cutpoint $c_{l(h)}$ to the sequence s.

II. Add the set $L(h) \setminus \{c_{l(h)}\}$ to the partition λ.

III. Remove the block $l(h)$, but not the cutpoint $c_{l(h)}$, from h.

IV. Repeat steps I, II and III with the new Husimi graph.

The procedure stops after the $(k-1)^{\text{th}}$ iteration. At the $(k-1)^{\text{th}}$ iteration, the vertices in the leaf block minus the cutpoint are added to the partition λ. The process is shown in Figure 5.3. We note that if h has block-size distribution $2^{n_2}3^{n_3}\cdots$ then λ has type $1^{n_2}2^{n_3}\cdots$. The procedure can be reversed and the bijection proves Equations (5.3.1) and (5.3.2).

We observe that vertices of the leaf-blocks other than the cutpoints do not appear in the sequence s. The resultant sequence has $k-1$ entries with a cutpoint of degree d_i appearing $d_i - 1$ times. Therefore, the number of Husimi graphs on $[n]$ having n_j blocks of size j and degree sequence $(d_1, \ldots, d_n)$ is given by

$$\frac{(n-1)!}{\prod_{j\geq 2}(j-1)!^{n_j} n_j!} \cdot \frac{(k-1)!}{\prod_{i=1}^{n}(d_i-1)!},$$

where $k = \sum_{j\geq 2} n_j$.

In Figure 5.3, we have that $k - 1 = 4$ and $\sum_{i=1}^{10}(d_i - 1) = 4$.

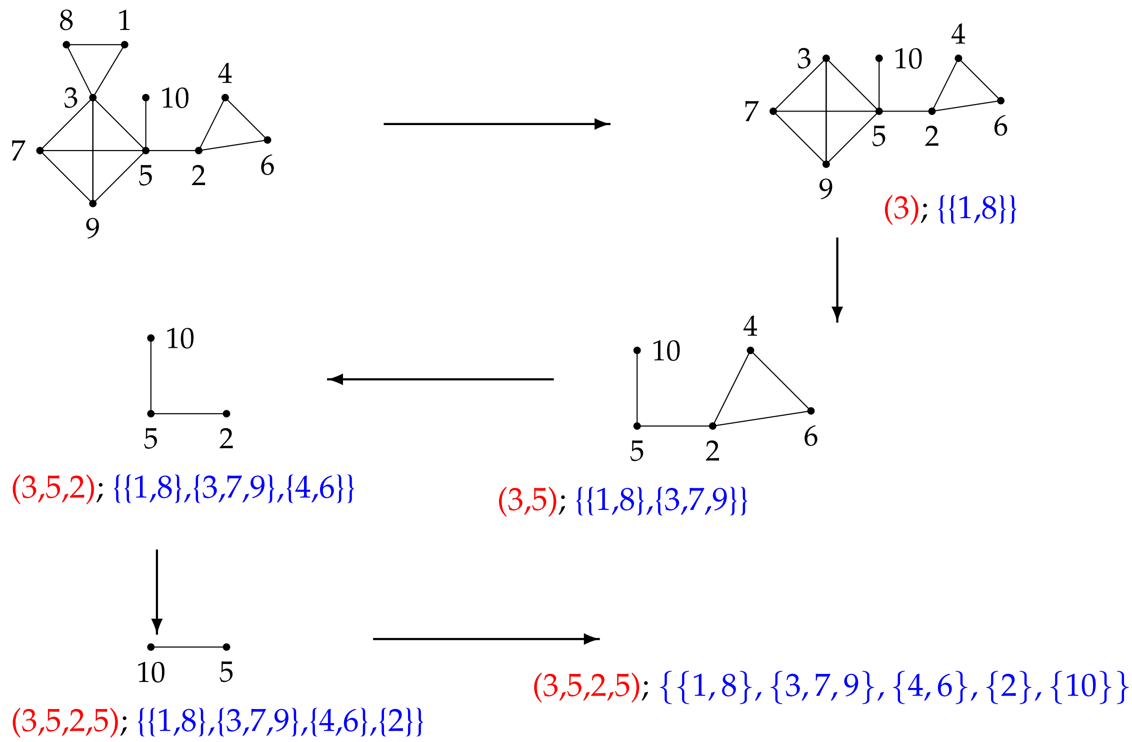

Figure 5.3: Prüfer correspondence for Husimi graphs

We pose the following two problems.

Problem 5.4.1. *What is the formula for the number of coloured Husimi graphs with a given degree sequence?*

Problem 5.4.2. *What is the total number of connected, cycle-free families of k set partitions of $[n]$ with a given degree sequence?*

A solution to Problem 5.4.1 will settle Problem 5.4.2 if we define the degree of vertex v in a connected cycle-free family of k set partitions of $[n]$ as the degree of vertex v in the corresponding coloured Husimi graph.

Chapter 6

Noncrossing tree-like structures

6.1 Introduction

A *noncrossing set partition* of $[n]$ is a partition such that if $a < b < c < d$ and both a and c are in one block then both b and d are not in the same block. The systematic study of these partitions started with the paper [34] by Germain Kreweras. He obtained that the number of noncrossing set partitions on $[n]$ of type $1^{m_1}2^{m_2}\cdots n^{m_n}$ is given by

$$\frac{n(n-1)\cdots(n-\ell+2)}{m_1!m_2!\cdots m_n!}$$

where $\ell = m_1 + m_2 + \cdots + m_n \geq 2$.

Noncrossing set partitions is one of the more than 200 structures counted by the famous Catalan numbers [57]. There are

$$\frac{1}{n+1}\binom{2n}{n}$$

noncrossing set partitions of $[n]$. Rodica Simion's survey paper [54] gives a general overview of results involving noncrossing partitions. H. Prodinger [48] obtained a bijection between the set of ordered trees on n vertices with k leaves and the set of noncrossing set partitions of $[n]$ having k parts. Dershowitz and Zaks [14] had shown that the number of these ordered trees is given by

$$\frac{1}{n}\binom{n}{k}\binom{n}{k-1},$$

the Narayana or Runyon numbers, a formula also obtained by Edelman [19] to count noncrossing partitions of $[n]$ with k parts.

6.2 Noncrossing Husimi graphs

Marc Noy [41] showed that the number of noncrossing trees on n labelled vertices is given by

$$\frac{1}{2n-1}\binom{3n-3}{n-1}.$$

This result was later generalised to connected graphs by Flajolet and Noy [22]. In this section, we extend their results to noncrossing Husimi graphs. But first off, let us review the notion of *butterfly decomposition* of noncrossing trees introduced in [22].

A *butterfly* is an ordered pair of trees that share a root. If a vertex v in a tree has degree d, then the tree can be decomposed into d butterflies hanging from v.

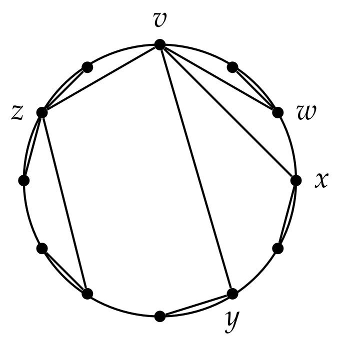

Figure 6.1: Noncrossing tree

In Figure 6.1, there are 4 butterflies rooted at w, x, y and z. The said authors showed that if $T(x)$ is the generating function for trees and $B(x)$ is the generating function for butterflies then we have the following equations:

$$T(x) = \frac{x}{1-B} \text{ and } B(x) = \frac{T^2}{x}.$$

Theorem 6.2.1. *Let $(n_2, n_3, \dots)$ be a sequence of non-negative integers satisfying the condition that $n = \sum_{j\geq 2}(j-1)n_j + 1$. The number $NHG_n(n_2, n_3, \cdots)$ of noncrossing Husimi graphs on $[n]$ having n_j blocks of size j is given by*

$$NHG_n(n_2, n_3, \dots) = \frac{(2n+k-2)!}{(2n-1)!\prod_{j\geq 2} n_j!} \tag{6.2.1}$$

where $k = \sum_{j\geq 2} n_j$.

Proof. Let $F(x)$ be the generating function for noncrossing Husimi graphs. Let y_i mark the number of vertices in each block. Adopting the butterfly decomposition of noncrossing trees to noncrossing Husimi graphs, we have that

$$F(x) = \frac{x}{1 - \sum_{i\geq 1} y_{i+1} B^i}$$

and

$$B(x) = \frac{F^2}{x}$$

where $B(x)$ is the generating function for butterflies.

Therefore the generating function $F(x)$ satisfies

$$F(x) = \frac{x}{1 - \sum\limits_{i\geq 1} y_{i+1} \left(\frac{F^2}{x}\right)^i}.$$

Thus for $G = \frac{F}{\sqrt{x}}$ we have

$$G(x) = \frac{\sqrt{x}}{1 - \sum\limits_{i\geq 1} y_{i+1} G^{2i}}.$$

By the Lagrange Inversion Formula, we obtain

$$\begin{aligned}
[x^n]\,F(x) = [x^{n-\frac{1}{2}}]G(x) &= \frac{1}{2n-1}[t^{2n-2}]\left(1 - \sum_{i\geq 1} y_{i+1} t^{2i}\right)^{-(2n-1)} \\
&= \frac{1}{2n-1}[t^{2n-2}]\sum_{k\geq 0}\binom{-(2n-1)}{k}\left(-\sum_{i\geq 1} y_{i+1} t^{2i}\right)^k \\
&= \frac{1}{2n-1}[t^{2n-2}]\sum_{k\geq 0}\binom{2n+k-2}{k}\left(\sum_{i\geq 1} y_{i+1} t^{2i}\right)^k \\
&= \frac{1}{2n-1}\sum_{k\geq 0}\binom{2n+k-2}{k}\sum_{\substack{n_2+n_3+\cdots=k \\ n_2+2n_3+\cdots=n-1}}\frac{k!\,y_2^{n_2} y_3^{n_3}\cdots}{n_2!\,n_3!\cdots}.
\end{aligned} \tag{6.2.2}$$

Therefore,

$$NHG_n(n_2, n_3, \ldots) = \frac{1}{2n-1}\binom{2n+k-2}{k}\frac{k!}{\prod_{j\geq 2} n_j!}.$$

□

Corollary 6.2.2. *The number of noncrossing Husimi graphs on $n \geq 2$ vertices is given by*

$$\frac{1}{n-1}\sum_{k=1}^{n-1}\binom{2n+k-2}{k-1}\binom{n-1}{k}.$$

Proof. We need to show that the number of noncrossing Husimi graphs on n vertices with k blocks is given by the generalised Narayana number,

$$\frac{1}{n-1}\binom{2n+k-2}{k-1}\binom{n-1}{k}. \tag{6.2.3}$$

Let $[[n,k]]$ denote the set of all types of partitions of $[n]$ of length k. Since

$$\sum_{P\in[[n-1,k]]}\frac{k!}{n_2!n_3!\dots} = \binom{n-2}{k-1},$$

the result follows from Equation (6.2.2). □

The formula (6.2.3) appears in [50] and [61] as the number of dissections of a convex polygon on $2n$ vertices with $k-1$ noncrossing diagonals such that the number of edges enclosing each interior region is even. We now construct a bijection between the set of these dissections and the noncrossing Husimi graphs.

Lemma 6.2.3. *There is a bijection between the set of dissections of a convex polygon on $2n$ vertices with $k-1$ noncrossing diagonals such that the number of edges enclosing each interior region is divisible by two and the set of noncrossing Husimi graphs on n vertices with k blocks.*

Proof. Consider a convex polygon on $2n$ vertices such that the vertices are labelled in clockwise direction as $1, 1', 2, 2', \dots, n, n'$. Let the number of noncrossing diagonal edges be $k-1$ and the number of edges of each interior region be divisible by 2. There are k such regions. Create an edge between any two vertices of label $1, \dots, n$ that are in the same region. A vertex which is incident to more than one region is considered to belong to all the incident regions. The resultant graph is a noncrossing Husimi graph on n vertices with k blocks. See Figure 6.2 for an example. The process can easily be reversed. □

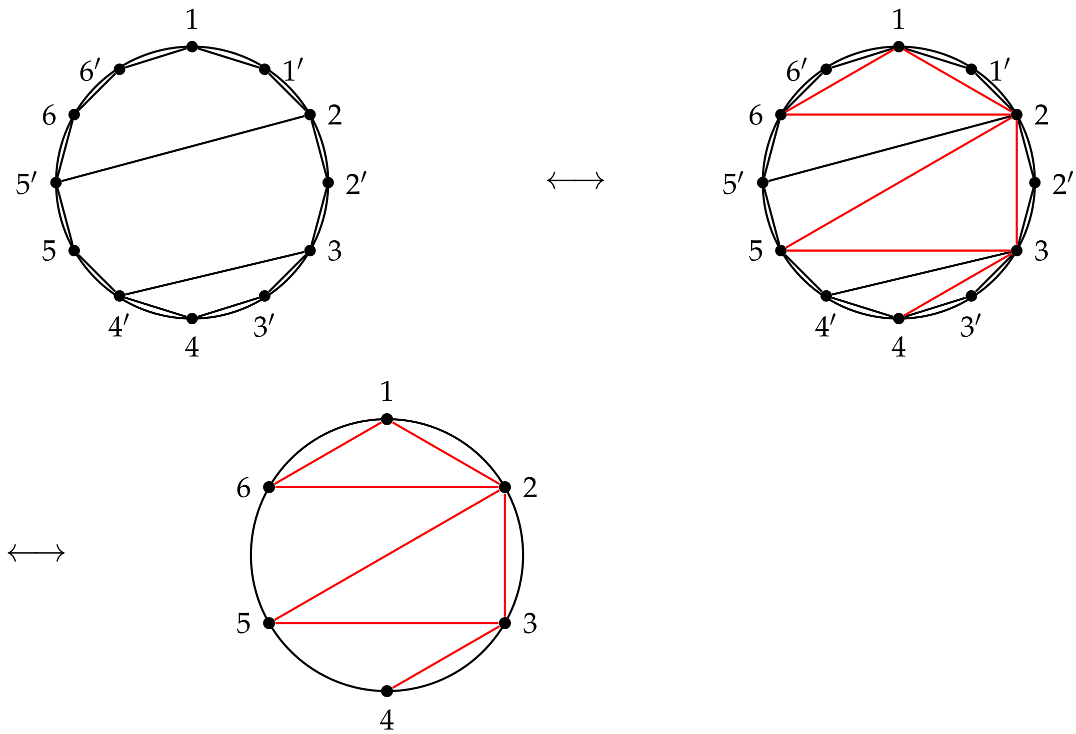

Figure 6.2: Diagram showing the bijection in the proof of Lemma 6.2.3.

We obtain further corollaries of Theorem 6.2.1.

Corollary 6.2.4. *Let* $(n_2, n_3, \ldots)$ *be a sequence of non-negative integers satisfying the condition that* $n = \sum_{j \geq 2}(j-1)n_j + 1$. *The number* $\mathrm{NC}_n(n_2, n_3, \ldots)$ *of noncrossing cacti on* $[n]$ *having* n_j *blocks of size* j *is given by*

$$\mathrm{NC}_n(n_2, n_3, \ldots) = \frac{(2n+k-2)!}{(2n-1)! \prod_{j \geq 2} n_j!}, \tag{6.2.4}$$

where $k = \sum_{j \geq 2} n_j$.

Proof. In the noncrossing setting, there is only one way to turn a complete graph into a cycle thus the required equation follows from Equation (6.2.1) i.e.,

$$NC_n(n_2, n_3, \ldots) = NHG_n(n_2, n_3, \ldots).$$

□

Corollary 6.2.5. *The number of noncrossing cacti on* $[n]$, *where* $n \geq 2$, *is*

$$\frac{1}{n-1} \sum_{k=1}^{n-1} \binom{2n+k-2}{k-1} \binom{n-1}{k}.$$

Proof. We obtain the formula by summing over all possibilities of $n_2, n_3, \ldots$ and k as in the proof of Corollary 6.2.2. □

Corollary 6.2.6. *Let $(n_2, n_3, \ldots)$ be a sequence of non-negative integers satisfying the condition that $n = \sum_{j \geq 2}(j-1)n_j + 1$. The number $\mathrm{NOC}_n(n_2, n_3, \cdots)$ of noncrossing oriented cacti on $[n]$ having n_j blocks of size j is given by*

$$\mathrm{NOC}_n(n_2, n_3, \cdots) = \frac{(2n+k-2)!2^{k-n_2}}{(2n-1)!\prod_{j\geq 2} n_j!},$$

where $k = \sum_{j\geq 2} n_j$.

Proof. Since any polygon of size ≥ 3 has 2 orientations, we have

$$NOC_n(n_2, n_3, \ldots) = 2^{k-n_2} \cdot NC_n(n_2, n_3, \ldots).$$

The formula thus follows from Equation (6.2.4). □

Corollary 6.2.7. *The number of noncrossing oriented cacti on $[n]$, where $n \geq 2$, is*

$$\sum_{k\geq 0} \sum_{\substack{n_2+n_3+\cdots=k \\ n_2+2n_3+\cdots=n-1}} \frac{(2n+k-2)!2^{k-n_2}}{(2n-1)!\prod_{j\geq 2} n_j!}.$$

6.3 Bicoloured noncrossing Husimi graphs

In the next proposition, we obtain a formula for the number of noncrossing Husimi graphs on n labelled vertices such that the degrees of the vertices are less than or equal to 2. This will make 2-colouring possible. Recall that the *degree* of a vertex v in a Husimi graph is the number of blocks that are incident to it.

Proposition 6.3.1. *Let $NHG_{n,2}(n_2, n_3, \ldots)$ be the number of noncrossing Husimi graphs on $[n]$ having n_i blocks of size i such that $\sum_{i\geq 2}(i-1)n_i + 1 = n$ and all the vertices have degree less than or equal to 2. Then*

$$NHG_{n,2}(n_2, n_3, \ldots) = \frac{n!2^{k-1}}{(n-k+1)!\prod_{j\geq 2} n_j!} \tag{6.3.1}$$

where $k = \sum_{j\geq 2} n_j$.

Proof. Let $F(x)$ be the generating function for 2-colourable noncrossing Husimi graphs with root degree 1 (or 0). Let y_i mark blocks of size i. Since each vertex in the block is to have degree less than or equal to two, the generating function satisfies

$$F(x) = x(1 + \sum_{i\geq 1} y_{i+1}(2F - x)^i). \tag{6.3.2}$$

The butterflies of these graphs must be rooted at vertices of degree 1 (or consists of a single vertex). We subtract x to cater for cases in which a butterfly consists of a single vertex.

Setting $G = 2F - x$ in Equation (6.3.2) we obtain

$$G = x(1 + 2\sum_{i\geq 1} y_{i+1}G^i).$$

G is the generating function for 2-coloured Husimi graphs with root degree 1 (in the case of a single vertex, there are no blocks, thus nothing to be coloured; otherwise there are precisely two colourings). When $y_2 = y_3 = \cdots = 1$, then we obtain the generating function for the large Schröder numbers.

Now, for arbitrary root degree, root degree 2 Husimi graphs are obtained by merging two root degree 1 Husimi graphs. We subtract F for double counting root degree 1 Husimi graphs. The generating function is thus

$$H(x) = \frac{F^2}{x} - F = \frac{G^2}{4x} - \frac{x}{4}.$$

This implies that

$$[x^n]H = \frac{1}{4}[x^{n+1}]G^2.$$

By the Lagrange Inversion Formula, we have

$$\begin{aligned}
\frac{1}{4}[x^{n+1}]G^2 &= \frac{1}{2(n+1)}[t^{n-1}]\left(1 + 2\sum_{i\geq 1} y_{i+1}t^i\right)^{n+1} \\
&= \frac{1}{2(n+1)}[t^{n-1}]\sum_{k\geq 0}\binom{n+1}{k}\left(2\sum_{i\geq 1} y_{i+1}t^i\right)^k \\
&= \frac{1}{2(n+1)}\sum_{k\geq 0} 2^k \binom{n+1}{k} \sum_{\substack{n_2+n_3+\cdots=k \\ n_2+2n_3+\cdots=n-1}} \frac{k!y_2^{n_2}y_3^{n_3}\cdots}{n_2!n_3!\cdots}
\end{aligned}$$

Therefore,

$$NHG_{n,2}(n_2, n_3, \ldots) = \frac{2^{k-1}}{n+1}\binom{n+1}{k} \cdot \frac{k!}{n_2! n_3! \cdots}. \qquad (6.3.3)$$

□

Corollary 6.3.2. *There are $n \cdot 2^{n-3}$ noncrossing trees on $n \geq 2$ vertices such that all the vertices have degree less than or equal to 2.*

Proof. The result follows from Equation (6.3.3) by taking $(n_2, n_3, \ldots) = (n-1, 0, \ldots)$ so that $k = n-1$.

Observe that these trees are also noncrossing paths. The corollary thus follows by a simple counting argument as well: first choose a root (in n ways), then 2 choices for each step. □

Corollary 6.3.3. *Let $NHG_{n,2}$ be the number of noncrossing Husimi graphs on $[n]$ in which all the vertices have degree at most 2. Then*

$$NHG_{n,2} = \frac{1}{n-1}\sum_{k=1}^{n-1} 2^{k-1}\binom{n}{k-1}\binom{n-1}{k}. \qquad (6.3.4)$$

Proof. To prove Formula (6.3.4), we need to show that the number of noncrossing Husimi graphs on n vertices with k blocks in which each vertex has degree ≤ 2 is given by

$$\frac{2^{k-1}}{n-1}\binom{n}{k-1}\binom{n-1}{k}.$$

Since

$$\sum_{P \in [[n-1,k]]} \frac{k!}{n_2! n_3! \ldots} = \binom{n-2}{k-1},$$

the result follows from Equation (6.3.3). □

Lemma 6.3.4. *The number of bicoloured noncrossing Husimi graphs on $[n]$ having n_i blocks of size i such that $\sum_{i\geq 2}(i-1)n_i + 1 = n$ is equal to*

$$\frac{n! 2^k}{(n-k+1)! \prod_{j\geq 2} n_j!}$$

where $k = \sum_{j\geq 2} n_j$.

Proof. Consider a noncrossing Husimi graph on $[n]$ having n_i blocks of size i such that $\sum_{i\geq 2}(i-1)n_i + 1 = n$ and with vertices having degree less than or equal to 2. Let b be a block in the graph. There are two choices for colouring block b and one choice for the remaining blocks. The result thus follows from Equation (6.3.1). □

Corollary 6.3.5. *The number of bicoloured noncrossing Husimi graphs on n vertices is given by*

$$\frac{1}{n-1}\sum_{k=1}^{n-1} 2^k \binom{n}{k-1}\binom{n-1}{k}. \tag{6.3.5}$$

We obtain the following special case by setting $k = n-1$ in Equation (6.3.5).

Corollary 6.3.6. *There are $n \cdot 2^{n-2}$ bicoloured noncrossing trees on $n \geq 2$ labelled vertices.*

Corollary 6.3.7. *The number of bicoloured noncrossing cacti on $[n]$ having n_i cycles of size i such that $\sum_{i\geq 2}(i-1)n_i + 1 = n$ is equal to*

$$\frac{n!2^k}{(n-k+1)!\prod_{j\geq 2} n_j!},$$

where $k = \sum_{j\geq 2} n_j$.

Corollary 6.3.8. *The number of bicoloured noncrossing cacti on $[n]$, where $n \geq 2$, is*

$$\frac{1}{n-1}\sum_{k=1}^{n-1} 2^k \binom{n}{k-1}\binom{n-1}{k}.$$

Corollary 6.3.9. *The number of bicoloured noncrossing oriented cacti on $[n]$ having n_i cycles of size i such that $\sum_{i\geq 2}(i-1)n_i + 1 = n$ is equal to*

$$\frac{n!2^{2k-n_2}}{(n-k+1)!\prod_{j\geq 2} n_j!},$$

where $k = \sum_{j\geq 2} n_j$.

Corollary 6.3.10. *The number of bicoloured noncrossing oriented cacti on $[n]$, for $n \geq 2$, is*

$$\sum_{k\geq 0} \sum_{\substack{n_2+n_3+\cdots=k \\ n_2+2n_3+\cdots=n-1}} \frac{n!2^{2k-n_2}}{(n-k+1)!\prod_{j\geq 2} n_j!}.$$

6.4 Noncrossing connected cycle-free set partitions

Proposition 6.4.1. *There is a bijection between the set of noncrossing coloured Husimi graphs on* $\{1, 2, \ldots, n\}$ *having* n_j *blocks of size* j *and block-colour distribution* $(a_{ij})_{1 \le i \le k, j \ge 2}$ *such that* $n = \sum_{i=1}^{k} \sum_{j \ge 2} (j-1) a_{ij} + 1$ *and* $a_{i1} = n - \sum_{j \ge 2} a_{ij}$*, and the set of noncrossing connected cycle-free families of* k *set partitions of* $[n]$ *such that there are* a_{ij} *blocks of size* j *in the* i*-th partition satisfying the coherence condition* $\sum_{i=1}^{k} \sum_{j \ge 1} a_{ij} = (k-1)n + 1$.

Proof. Given a noncrossing coloured Husimi graph, the number of blocks of colour i is given by $\sum_{j \ge 2} a_{ij}$. Since $a_{i1} = n - \sum_{j \ge 2} a_{ij}$, we have that $\sum_{j \ge 1} a_{ij} = n$. Let B_{ij} be the set of vertices in the j-th block of colour i. Since the blocks of the same colour are not incident to one another, the B_{ij}'s are disjoint. The vertices which are not part of any block are taken as singletons. Thus the family of sets comprising the singletons and the blocks do not have empty sets. Hence $(a_{ij})_{j \ge 1}$ is a partition of $[n]$. Since the blocks are noncrossing, we obtain a noncrossing set partition. We have k such partitions since $1 \le i \le k$. The set partitions satisfy the condition $\sum_{i=1}^{k} \sum_{j \ge 1} a_{ij} = (k-1)n + 1$.

We obtain the reverse. Consider families of cycle-free connected noncrossing k set partitions of $[n]$ in which there are a_{ij} blocks of size j in the i-th partition satisfying the condition: $\sum_{i=1}^{k} \sum_{j \ge 1} a_{ij} = (k-1)n + 1$. Allow edges between two distinct vertices in the block and remove all singletons. If v is a block in the i-th partition then colour the resultant complete graph with colour i. We thus obtain a noncrossing coloured Husimi graph satisfying the condition $n = \sum_{j \ge 2} (j-1) n_j + 1$ where $n_j = \sum_{i=1}^{k} a_{ij}$. □

Theorem 6.4.2. *Let* $\lambda_a = 1^{a_1} 2^{a_2} \cdots$ *and* $\lambda_b = 1^{b_1} 2^{b_2} \cdots$ *be partitions of* $n \ge 1$*, with* $a = a_1 + a_2 + \cdots$ *and* $b = b_1 + b_2 + \cdots$*. If* $a + b = n + 1$ *then the number of noncrossing connected, cycle-free pairs of partitions of types* λ_a *and* λ_b *is given by*

$$n 2^{n - a_1 - b_1} \cdot \frac{(a-1)!(b-1)!}{a_1! a_2! \cdots b_1! b_2! \cdots}.$$

Proof. Let $A(z)$ be the generating function for 2-colourable noncrossing Husimi graphs with root degree 1 (or 0) such that the root block is of colour 1. Let $B(z)$ be the corresponding generating function in which the root block is of colour 2. Let x_j and y_i mark blocks of sizes j and i of colours 1 and 2 respectively. Since each vertex has degree less than or equal to two, the generating functions satisfy (in a similar way as in the previous section)

$$A(z) = z(1 + \sum_{j\geq 1} x_{j+1}(2B - z)^j)$$

and

$$B(z) = z(1 + \sum_{i\geq 1} y_{i+1}(2A - z)^i).$$

Set $2B - z = V$ and $2A - z = U$ so that

$$U = z(1 + 2\sum_{j\geq 1} x_{j+1}V^j)$$

and

$$V = z(1 + 2\sum_{i\geq 1} y_{i+1}U^i).$$

We obtain a 2-colourable noncrossing Husimi graphs by identifying roots of block colour 1 and 2, or the noncrossing Husimi graph has root degree 1 of colour 1 or 2. We thus need

$$\frac{2(A - z)(B - z)}{z} + A + B - z = \frac{UV}{2z} + \frac{z}{2}.$$

We apply the multivariate Lagrange Inversion Formula [25, Theorem 1.29] to the system:

$$U = z_1(1 + 2\sum_{j\geq 1} x_{j+1}V^j),$$
$$V = z_2(1 + 2\sum_{i\geq 1} y_{i+1}U^i).$$

We have

$$UV = \sum_{s,t\geq 0} z_1^{s+1} z_2^{t+1} [\alpha^s \beta^t] \left(1 + 2\sum_{j\geq 1} x_{j+1}\alpha^j\right)^{t+1} \cdot \left(1 + 2\sum_{i\geq 1} y_{i+1}\beta^i\right)^{s+1} \Delta,$$

where Δ is the determinant given by

$$\Delta = \frac{(1 + 2\sum_{j\geq 1} x_{j+1}\alpha^j) \cdot (1 + 2\sum_{i\geq 1} y_{i+1}\beta^i) - 4\sum_{j\geq 1} jx_{j+1}\alpha^j \cdot \sum_{i\geq 1} iy_{i+1}\beta^i}{\left(1 + 2\sum_{j\geq 1} x_{j+1}\alpha^j\right)\left(1 + 2\sum_{i\geq 1} y_{i+1}\beta^i\right)}.$$

We thus have

$$UV = \sum_{s,t\geq 0} z_1^{s+1} z_2^{t+1} [\alpha^s \beta^t] \left(1 + 2\sum_{j\geq 1} x_{j+1}\alpha^j\right)^{t+1} \left(1 + 2\sum_{i\geq 1} y_{i+1}\beta^i\right)^{s+1}$$

$$- 4 \sum_{s,t\geq 0} \left[z_1^{s+1} z_2^{t+1} [\alpha^s \beta^t] \left(1 + 2\sum_{j\geq 1} x_{j+1}\alpha^j\right)^{t} \left(1 + 2\sum_{i\geq 1} y_{i+1}\beta^i\right)^{s} \sum_{j\geq 1} j x_{j+1}\alpha^j \times \sum_{i\geq 1} i y_{i+1}\beta^i \right].$$

Let us extract the coefficient of α^s.

$$[\alpha^s] \left(1 + 2\sum_{j\geq 1} x_{j+1}\alpha^j\right)^{t+1} = [\alpha^s] \sum_{k\geq 0} \binom{t+1}{k} \left(2\sum_{i\geq 1} x_{j+1}\alpha^i\right)^k$$

$$= \sum_{k\geq 0} 2^k \binom{t+1}{k} \sum_{\substack{a_2+a_3+\cdots=k \\ a_2+2a_3+\cdots=s}} \frac{k! x_2^{a_2} x_3^{a_3} \cdots}{a_2! a_3! \cdots}.$$

Similarly,

$$[\beta^t] \left(1 + 2\sum_{i\geq 1} y_{i+1}\beta^i\right)^{s+1} = [\beta^t] \sum_{\ell\geq 0} \binom{s+1}{\ell} \left(2\sum_{i\geq 1} y_{i+1}\beta^i\right)^\ell$$

$$= \sum_{\ell\geq 0} 2^\ell \binom{s+1}{\ell} \sum_{\substack{b_2+b_3+\cdots=\ell \\ b_2+2b_3+\cdots=t}} \frac{\ell! y_2^{b_2} y_3^{b_3} \cdots}{b_2! b_3! \cdots}.$$

Since $k = a - a_1$, $\ell = b - b_1$, $s = n - a$ and $t = n - b$ in our notation, we have

$$[\alpha^{n-a}] \left(1 + 2\sum_{j\geq 1} x_{j+1}\alpha^j\right)^{n-b+1}$$

$$= \sum_{a_1} 2^{a-a_1} \frac{(n-b+1)!}{(n-b-a+a_1+1)!} \sum_{\substack{a_1+a_2+\cdots=a \\ a_2+2a_3+\cdots=n-a}} \frac{x_2^{a_2} x_3^{a_3} \cdots}{a_2! a_3! \cdots}$$

and

$$[\beta^{n-b}] \left(1 + 2\sum_{i\geq 1} y_{i+1}\beta^i\right)^{n-a+1}$$

$$= \sum_{b_1} 2^{b-b_1} \frac{(n-a+1)!}{(n-a-b+b_1+1)!} \sum_{\substack{b_1+b_2+\cdots=b \\ b_2+2b_3+\cdots=n-b}} \frac{y_2^{b_2} y_3^{b_3} \cdots}{b_2! b_3! \cdots}.$$

By the coherence condition, $a+b=n+1$, we obtain

$$[\alpha^{n-a}]\left(1+2\sum_{j\geq 1}x_{j+1}\alpha^j\right)^{n-b+1}=\sum_{a_1}2^{a-a_1}a!\sum_{\substack{a_1+a_2+\cdots=a\\a_2+2a_3+\cdots=n-a}}\frac{x_2^{a_2}x_3^{a_3}\cdots}{a_1!a_2!\cdots}$$

and

$$[\beta^{n-b}]\left(1+2\sum_{i\geq 1}y_{i+1}\beta^i\right)^{n-a+1}=\sum_{b_1}2^{b-b_1}b!\sum_{\substack{b_1+b_2+\cdots=b\\b_2+2b_3+\cdots=n-b}}\frac{y_2^{b_2}y_3^{b_3}\cdots}{b_1!b_2!\cdots}.$$

We also have

$$\begin{aligned}
&[\alpha^s]\left(1+2\sum_{j\geq 1}x_{j+1}\alpha^j\right)^t\sum_{j\geq 1}jx_{j+1}\alpha^j\\
&=\sum_{w=0}^{s}[\alpha^w]\left(1+2\sum_{j\geq 1}x_{j+1}\alpha^j\right)^t[\alpha^{s-w}]\sum_{j\geq 1}jx_{j+1}\alpha^j\\
&=\sum_{w=0}^{s}[\alpha^w]\sum_{k\geq 0}\binom{t}{k}\left(2\sum_{i\geq 1}x_{j+1}\alpha^i\right)^k(s-w)x_{s-w+1}\\
&=\sum_{w=0}^{s}\sum_{k\geq 0}2^k\binom{t}{k}\sum_{\substack{a_2+a_3+\cdots=k\\a_2+2a_3+\cdots=w}}\frac{k!x_2^{a_2}x_3^{a_3}\cdots}{a_2!a_3!\cdots}(s-w)x_{s-w+1}\\
&=\sum_{w=0}^{s}\sum_{k\geq 0}2^k\binom{t}{k}\sum_{\substack{a_2+a_3+\cdots=k+1\\a_2+2a_3+\cdots=s}}\frac{k!x_2^{a_2}x_3^{a_3}\cdots}{a_2!a_3!\cdots}(s-w)a_{s-w+1}\\
&=\sum_{k\geq 0}2^k\binom{t}{k}\sum_{\substack{a_2+a_3+\cdots=k+1\\a_2+2a_3+\cdots=s}}\frac{k!x_2^{a_2}x_3^{a_3}\cdots}{a_2!a_3!\cdots}\cdot\sum_{w=0}^{s}(s-w)a_{s-w+1}\\
&=\sum_{k\geq 0}2^k\binom{t}{k}\sum_{\substack{a_2+a_3+\cdots=k+1\\a_2+2a_3+\cdots=s}}\frac{k!x_2^{a_2}x_3^{a_3}\cdots}{a_2!a_3!\cdots}\cdot s.
\end{aligned}$$

Likewise,

$$[\beta^t]\left(1+2\sum_{i\geq 1}y_{i+1}\beta^i\right)^s\sum_{i\geq 1}iy_{i+1}\beta^i=\sum_{\ell\geq 0}2^\ell\binom{s}{\ell}\sum_{\substack{b_2+b_3+\cdots=\ell+1\\a_2+2a_3+\cdots=s}}\frac{\ell!y_2^{b_2}y_3^{b_3}\cdots}{b_2!b_3!\cdots}\cdot t.$$

Again using coherence conditions $k = a - a_1 - 1$, $\ell = b - b_1 - 1$, $s = n - a$, $t = n - b$ and $a + b = n + 1$, we obtain

$$[\alpha^{n-a}] \left(1 + 2\sum_{j\geq 1} x_{j+1}\alpha^j\right)^{n-b} \sum_{j\geq 1} j x_{j+1}\alpha^j$$
$$= \sum_{a_1} 2^{a-a_1-1}(a-1)! \sum_{\substack{a_1+a_2+\cdots=a \\ a_2+2a_3+\cdots=n-a}} \frac{x_2^{a_2}x_3^{a_3}\cdots}{a_1!a_2!\cdots} \cdot (b-1)$$

and

$$[\beta^{n-b}] \left(1 + 2\sum_{i\geq 1} y_{i+1}\beta^i\right)^{n-a} \sum_{i\geq 1} i y_{i+1}\beta^i$$
$$= \sum_{b_1} 2^{b-b_1-1}(b-1)! \sum_{\substack{b_1+b_2+\cdots=b \\ b_2+2b_3+\cdots=n-b}} \frac{y_2^{b_2}y_3^{b_3}\cdots}{b_1!b_2!\cdots} \cdot (a-1).$$

Hence

$$\begin{aligned}
&[x_2^{a_2}x_3^{a_3}\cdots y_2^{b_2}y_3^{b_3}\cdots]UV \\
&= z_1^{n-a+1}z_2^{n-b+1}2^{a+b-a_1-b_1}\frac{a!b!}{a_1!a_2!\cdots b_1!b_2!\cdots} \\
&- z_1^{n-a+1}z_2^{n-b+1}4\cdot 2^{a+b-a_1-b_1-2}(a-1)(b-1)\frac{(a-1)!(b-1)!}{a_1!a_2!\cdots b_1!b_2!\cdots} \\
&= z_1^{n-a+1}z_2^{n-b+1}(a+b-1)\cdot 2^{n+1-a_1-b_1}\frac{(a-1)!(b-1)!}{a_1!a_2!\cdots b_1!b_2!\cdots} \\
&= z_1^{n-a+1}z_2^{n-b+1}n\cdot 2^{n+1-a_1-b_1}\frac{(a-1)!(b-1)!}{a_1!a_2!\cdots b_1!b_2!\cdots}.
\end{aligned}$$

Now, by setting $z_1 = z_2 = z$, we have

$$[x_2^{a_2}x_3^{a_3}\cdots y_2^{b_2}y_3^{b_3}\cdots]\frac{UV}{2z} = z^n n\cdot 2^{n-a_1-b_1}\frac{(a-1)!(b-1)!}{a_1!a_2!\cdots b_1!b_2!\cdots}.$$

Therefore

$$[z^n x_2^{a_2}x_3^{a_3}\cdots y_2^{b_2}y_3^{b_3}\cdots]\frac{UV}{2z} = n\cdot 2^{n-a_1-b_1}\frac{(a-1)!(b-1)!}{a_1!a_2!\cdots b_1!b_2!\cdots}.$$

By Proposition 6.4.1, we obtain that there are

$$n\cdot 2^{n-a_1-b_1}\frac{(a-1)!(b-1)!}{a_1!a_2!\cdots b_1!b_2!\cdots}$$

noncrossing connected, cycle-free pairs of partitions of $[n]$ of types $1^{a_1}2^{a_2}\cdots$ and $1^{b_1}2^{b_2}\cdots$. □

Chapter 7

Locally oriented noncrossing trees

7.1 Introduction

Recall that a *noncrossing tree* [15] is a tree drawn in the plane with its vertices on the boundary of a circle such that the edges are straight line segments that do not cross. The number of these trees on n labelled vertices is known to be given by

$$\frac{1}{n-1}\binom{3n-3}{n-2}. \tag{7.1.1}$$

See [15, 41] for details. In this chapter, we study labelled noncrossing trees having a local orientation as in Section 3.2, i.e., all edges are oriented towards the larger label. We shall call these trees *locally oriented noncrossing trees* (*lnc-trees* for short).

7.2 Enumeration by sources and sinks

Theorem 7.2.1. *The number of lnc-trees of order n with k sources and ℓ sinks is equal to*

$$\frac{1}{n-1}\binom{n-1}{k-1}\binom{n-1}{\ell-1}\binom{n-1}{k+\ell-1}. \tag{7.2.1}$$

Proof. Let $A(x,y,z)$ be the generating function for the number of noncrossing trees rooted at vertex 1, where x, y and z mark sinks, sources and total number of vertices respectively. Let $B(x,y,z)$ be the corresponding generating function for noncrossing trees rooted at vertex n. The root does not count as a source or sink in these generating functions.

The trees rooted at vertex 1 can be represented as shown in Figure 7.1,

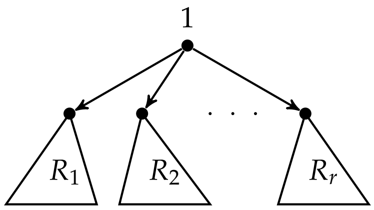

Figure 7.1: Noncrossing tree rooted at vertex 1.

where $R_1, R_2, \ldots, R_r$ are noncrossing subtrees. These subtrees are either rooted at a sink (in which case the root has the highest label in the subtree) or at a vertex which is neither a source nor a sink (in which case the subtree consists of two parts consisting of the vertices with labels higher than the root's label and lower than the root's label respectively). Thus we have,

$$A = \frac{z}{1 - (\frac{AB}{z} + (x-1)B)}. \tag{7.2.2}$$

Similarly,

$$B = \frac{z}{1 - (\frac{AB}{z} + (y-1)A)}. \tag{7.2.3}$$

Solving Equations (7.2.2) and (7.2.3) simultaneously we obtain

$$(A - z)(zx - (y-1)A^2 - z(y-1)(x-1)A) = A(A + z(x-1))^2. \tag{7.2.4}$$

Setting $A = z + xzT$ in Equation (7.2.4) and solving for z we obtain

$$z = \frac{T}{(1+T)(1+xT)(1+yT)}.$$

Thus

$$T = z(1+T)(1+xT)(1+yT).$$

By the Lagrange Inversion Formula [58, Theorem 5.4.2], one has

$$[z^n]A = x[z^{n-1}]T = \frac{x}{n-1}[t^{n-2}]((1+t)(1+xt)(1+yt))^{n-1}.$$

Now, we extract the coefficient of x^ℓ to obtain

$$\begin{aligned}[z^n x^\ell]A &= \frac{1}{n-1}[t^{n-2}x^{\ell-1}](1+t)^{n-1}(1+xt)^{n-1}(1+yt)^{n-1} \\ &= \frac{1}{n-1}\binom{n-1}{\ell-1}[t^{n-2}](1+t)^{n-1}(1+yt)^{n-1}t^{\ell-1}.\end{aligned}$$

Similarly, we extract the coefficient of y^{k-1} (since the root, whose label is 1, is not counted as a source) to get

$$\begin{aligned}[z^n x^\ell y^{k-1}]A &= \frac{1}{n-1}\binom{n-1}{\ell-1}[t^{n-2}y^{k-1}](1+t)^{n-1}(1+yt)^{n-1}t^{\ell-1} \\ &= \frac{1}{n-1}\binom{n-1}{\ell-1}\binom{n-1}{k-1}[t^{n-2}](1+t)^{n-1}t^{k+\ell-2}.\end{aligned}$$

Therefore,

$$[z^n x^\ell y^{k-1}]A = \frac{1}{n-1}\binom{n-1}{\ell-1}\binom{n-1}{k-1}\binom{n-1}{n-k-\ell}.$$

This completes the proof. □

Let T be a ternary tree which consists of a root e and 3 disjoint ternary subtrees, T_3, T_2, T_1 in this order. See Figure 7.2.

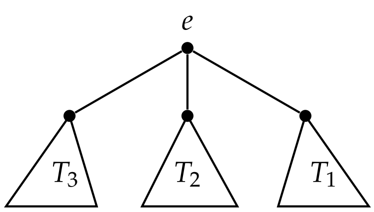

Figure 7.2: Rooted ternary tree showing left-, middle- and right-edges.

We shall call an edge from the root to T_1, T_2 and T_3 right-edge, middle-edge and left-edge respectively. J. Cigler [10, Theorem 1] showed that the number of ternary trees on $n-1$ vertices having precisely $(\ell-1)$, $(n-k-\ell)$, and $(k-1)$ right-edges, middle-edges and left-edges respectively is given by Equation (7.2.1).

Lemma 7.2.2. *There is a bijection between the set of ternary trees on $[n-1]$ such that there are $k-1$ and $\ell-1$ left-edges and right-edges respectively, and the set of lnc-trees on $[n]$ with k sources and ℓ sinks.*

Proof. Our bijection is a modification of the bijection between ternary trees on $n-1$ vertices and noncrossing trees on n vertices obtained independently by R. Simion and A. Postnikov as mentioned by R. Stanley in [58, Solution 5.46]. Other bijective proofs can be found in the literature, for example, Dulucq and Penaud [18], and Panholzer and Prodinger [43].

Given an lnc-tree T on n vertices having k sources and ℓ sinks, we obtain the corresponding ternary tree $\beta(T)$ on $n-1$ vertices ($n-1$ edges of T as vertices of $\beta(T)$) with $k-1$ left-edges and $\ell-1$ right-edges by the following steps:

I. Let s be the vertex with the smallest label in T (1 for this case), and let j be the vertex with the largest label attached to s. We label the edge sj as e. Three subtrees of T are defined as follows.

II. T_1 is the connected component containing 1 in the graph $T-e$.

III. T_2 is the graph obtained from T by removing the vertices $1, 2, \ldots, j-1$.

IV. T_3 is the connected component containing vertex j in the graph obtained from T by removing the edge e and vertices $j+1, j+2, \ldots, n$.

V. Define e as the root of $\beta(T)$ and recursively define $\beta(T_i)$ to be the i-th subtree of the root, with the subtrees read from right to left. Label left-edges, middle-edges and right-edges as l, m and r respectively in the ternary tree rooted at e. Each label r corresponds to the vertex with the largest label in subtree T_1 (i.e., this vertex was initially a sink in the lnc-tree T). Label m corresponds to the vertex with the smallest label in T_2 (this vertex was neither a source nor a sink in T), and label l corresponds to the vertex with the smallest label in T_3 (i.e., this vertex was initially a source in T). See Figure 7.3 for an example. The process is reversible.

The resultant ternary tree has $k-1$ left-edges (vertex 1, which is a source, does not correspond to any left-edge in the ternary tree since it always occurs in the subtree T_1) and $\ell-1$ right-edges (vertex n, which is a sink, does not correspond to any right-edge in the ternary tree since it always occurs in the subtree T_2). □

We obtain several corollaries of Theorem 7.2.1.

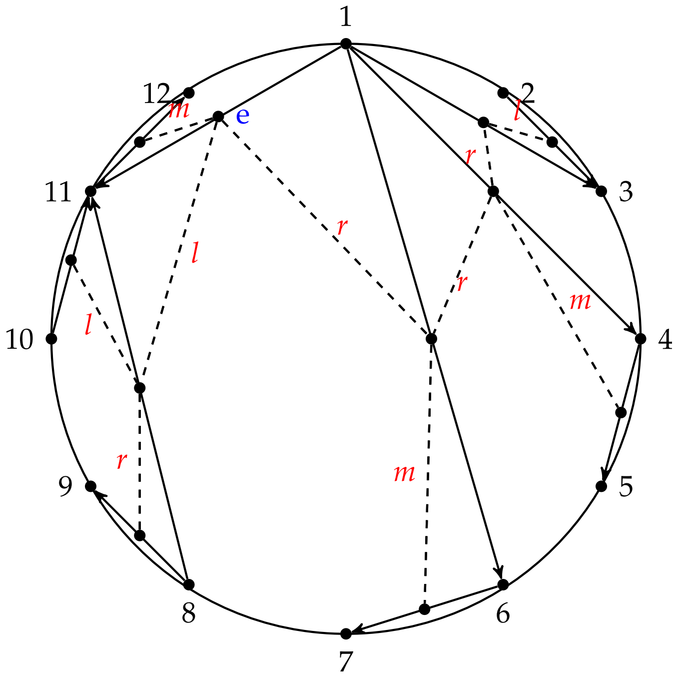

Figure 7.3: Diagram showing the bijection in the proof of Lemma 7.2.2.

Corollary 7.2.3. *There are*

$$\frac{1}{n-1}\binom{n-1}{k-1}\binom{2n-2}{n-k-1} \tag{7.2.5}$$

lnc-trees of order n with k sources.

Proof. The formula follows by summing over all ℓ in Equation (7.2.1) and using Vandermonde's Convolution Formula

$$\sum_i \binom{r}{m+i}\binom{s}{p-i} = \binom{r+s}{m+p},$$

where m and p are integers (See [28, Equation (5.22)]). □

Corollary 7.2.4. *The number of lnc-trees of order n with k sources and ℓ sinks such that $k+\ell = r$ is given by*

$$\frac{1}{n-1}\binom{n-1}{r-1}\binom{2n-2}{r-2}. \tag{7.2.6}$$

Proof. We set $k+\ell = r$ and $\ell = r-k$ in Equation (7.2.1) and sum over all k. □

In the following corollary, we obtain a formula for the number of noncrossing trees on n vertices such that each vertex is either a source or a sink. These trees are called *noncrossing alternating trees*. They were first studied by Gelfand, Graev and Postnikov in [24].

Corollary 7.2.5. *The number of noncrossing alternating trees of order n is given by the $(n-1)^{th}$ Catalan number,*

$$\frac{1}{n-1}\binom{2n-2}{n-2}.$$

Proof. Set $r = n$ in Equation (7.2.6). □

Corollary 7.2.6. *The total number of sinks in lnc-trees on n labelled vertices with k sources is*

$$\frac{n-k+1}{2n-2}\binom{n-1}{k-1}\binom{2n-2}{n-k-1}. \tag{7.2.7}$$

Proof. From Equation (7.2.1) we have that the total number of sinks is

$$\begin{aligned}
&\frac{1}{n-1}\binom{n-1}{k-1}\sum_{\ell}\ell\binom{n-1}{\ell-1}\binom{n-1}{k+\ell-1}\\
&=\frac{1}{n-1}\binom{n-1}{k-1}\left[\sum_{\ell}(\ell-1)\binom{n-1}{\ell-1}\binom{n-1}{k+\ell-1}+\sum_{\ell}\binom{n-1}{\ell-1}\binom{n-1}{k+\ell-1}\right]\\
&=\frac{1}{n-1}\binom{n-1}{k-1}\left[(n-1)\binom{2n-3}{n-k-2}+\binom{2n-2}{n-k-1}\right] \qquad (7.2.8)\\
&=\frac{n-k+1}{2n-2}\binom{n-1}{k-1}\binom{2n-2}{n-k-1}.
\end{aligned}$$

□

Corollary 7.2.7. *The mean number of sinks in lnc-trees on n labelled vertices with k sources is*

$$\mu = \frac{n-k+1}{2}. \tag{7.2.9}$$

Proof. The result follows from Equations (7.2.5) and (7.2.7). □

Corollary 7.2.8. *The variance of the number of sinks in lnc-trees on n vertices with k sources is given by*

$$\sigma^2 = \frac{n^2-k^2-2n+1}{8n-12}.$$

Proof. We have

$$
\begin{aligned}
&\sum_{\ell} \ell^2 \binom{n-1}{\ell-1}\binom{n-1}{k+\ell-1} \\
&\quad= \sum_{\ell}(\ell-2)(\ell-1)\binom{n-1}{\ell-1}\binom{n-1}{k+\ell-1} + 3\sum_{\ell}(\ell-1)\binom{n-1}{\ell-1}\binom{n-1}{k+\ell-1} \\
&\qquad+ \sum_{\ell}\binom{n-1}{\ell-1}\binom{n-1}{k+\ell-1} \\
&\quad=(n-1)(n-2)\sum_{\ell}\binom{n-3}{\ell-3}\binom{n-1}{k+\ell-1} + 3(n-1)\sum_{\ell}\binom{n-2}{\ell-2}\binom{n-1}{k+\ell-1} \\
&\qquad+ \sum_{\ell}\binom{n-1}{\ell-1}\binom{n-1}{k+\ell-1}.
\end{aligned}
$$

Now by Vandermonde's Convolution Formula,

$$
\begin{aligned}
&\sum_{\ell} \ell^2 \binom{n-1}{\ell-1}\binom{n-1}{k+\ell-1} \\
&\qquad= (n-1)(n-2)\binom{2n-4}{n-k-3} + 3(n-1)\binom{2n-3}{n-k-2} + \binom{2n-2}{n-k-1} \\
&\qquad= (n-1)\left[\frac{(n-2)(n-k)^2 + 3(n-1)(n-k) - 1}{(2n-3)(n-k-1)}\binom{2n-3}{n-k-2}\right] \\
&\qquad= \frac{(n-2)(n-k)^2 + 3(n-1)(n-k) - 1}{2(2n-3)}\binom{2n-2}{n-k-1}.
\end{aligned}
$$

Therefore

$$
\begin{aligned}
M_2 &= \frac{1}{n-1}\binom{n-1}{k-1}\sum_{\ell}\ell^2\binom{n-1}{\ell-1}\binom{n-1}{k+\ell-1} \\
&= \frac{1}{n-1}\binom{n-1}{k-1}\frac{(n-2)(n-k)^2 + 3(n-1)(n-k) - 1}{2(2n-3)}\binom{2n-2}{n-k-1}.
\end{aligned}
\tag{7.2.10}
$$

Since the variance is given by

$$
\sigma^2 = \frac{M_2}{T} - \mu^2,
$$

where T is the number of lnc-trees on n vertices with k sources, the result follows from Equations (7.2.5), (7.2.9) and (7.2.10) after some straightforward calculations. □

Corollary 7.2.9. *There are a total of*

$$
\frac{n+1}{3n-3}\binom{3n-3}{n-2} \tag{7.2.11}
$$

sinks in all lnc-trees of order n.

Proof. We sum over all k in Equation (7.2.7).

$$\begin{aligned}
&\sum_{k=1}^{n} \frac{n-k+1}{2n-2} \binom{n-1}{k-1} \binom{2n-2}{n-k-1} \\
&= \sum_{k=1}^{n} \left[\binom{n-1}{k-1} \binom{2n-3}{n-k-2} + \frac{1}{n-1} \binom{n-1}{k-1} \binom{2n-2}{n-k-1} \right] \\
&= \binom{3n-4}{n-3} + \frac{1}{n-1} \binom{3n-3}{n-2} \\
&= \frac{n+1}{3n-3} \binom{3n-3}{n-2}.
\end{aligned}$$

□

Corollary 7.2.10. *In an average noncrossing tree on n vertices, approximately one third of the vertices are sinks for large n.*

Proof. The total number of vertices in noncrossing trees of order n is equal to

$$\frac{n}{n-1} \binom{3n-3}{n-2},$$

and from Equation (7.2.11), we know that there are

$$\frac{n+1}{3n-3} \binom{3n-3}{n-2}$$

sinks. Therefore the probability that a given vertex is a sink is

$$\frac{1}{3} \left(1 + \frac{1}{n} \right).$$

The result follows by letting n tend to infinity. □

From Equations (7.1.1) and (7.2.11), it follows that

Corollary 7.2.11. *The average number of sinks in lnc-trees on n vertices is equal to*

$$\mu_t = \frac{n+1}{3}. \tag{7.2.12}$$

Corollary 7.2.12. *The variance of the number of sinks in lnc-trees on n vertices is given by*

$$\sigma_t^2 = \frac{4n^2 - 10n + 4}{27n - 36}.$$

Proof. By summing over all k in Equation (7.2.10) we obtain the second moment:

$$M = \frac{n^3 + 2n^2 - 5n}{(3n-4)(3n-3)} \binom{3n-3}{n-2}. \tag{7.2.13}$$

Since the variance equals

$$\sigma_t^2 = \frac{M}{T_t} - \mu_t^2,$$

where T_t is the number of lnc-trees on n vertices, the result follows from Equations (7.1.1), (7.2.12) and (7.2.13). □

Corollary 7.2.13. *The covariance of the number of sources and sinks in lnc-trees of order n is equal to*

$$\sigma_t(k,\ell) = \frac{-2n^2 + 5n - 2}{27n - 36}.$$

Proof. By Equation (7.2.8), we have

$$\begin{aligned}
&\frac{1}{n-1}\binom{n-1}{k-1}\sum_{\ell} k\ell \binom{n-1}{\ell-1}\binom{n-1}{k+\ell-1} \\
&= \frac{k}{n-1}\binom{n-1}{k-1}\left[(n-1)\binom{2n-3}{n-k-2} + \binom{2n-2}{n-k-1}\right] \\
&= (n-1)\binom{n-2}{k-2}\binom{2n-3}{n-k-2} + \binom{n-1}{k-1}\binom{2n-3}{n-k-2} \\
&\quad + \binom{n-2}{k-2}\binom{2n-2}{n-k-1} + \frac{1}{n-1}\binom{n-1}{k-1}\binom{2n-2}{n-k-1}.
\end{aligned} \tag{7.2.14}$$

By summing over all k in Equation (7.2.14), we obtain

$$\frac{n^3 - 2}{(3n-4)(3n-3)} \binom{3n-3}{n-2} := C. \tag{7.2.15}$$

By symmetry, the mean number of sinks and sources is given by Equation (7.2.12). Since the covariance equals

$$\sigma_t(k,\ell) = \frac{C}{T_t} - \mu_t(k)\mu_t(\ell),$$

where T_t is the number of lnc-trees on n vertices, $\mu_t(k)$ is the mean number of sources and $\mu_t(\ell)$ is the mean number of sinks, the result follows from Equations (7.1.1), (7.2.12) and (7.2.15). □

7.3 Noncrossing trees with given in- and out-degree sequences

In [22], Flajolet and Noy obtained a formula for the number of noncrossing trees on $[n]$ with a given outdegree sequence if all the edges in the tree are *globally oriented*, i.e., all edges are oriented away from the root. They showed that the number of noncrossing trees on $[n]$ with an outdegree sequence $\langle 1^{e_1}2^{e_2}\cdots\rangle$ satisfying the conditions that

$$\sum_{i\geq 0} e_i = n \text{ and } \sum_{i\geq 1} ie_i = n-1 \tag{7.3.1}$$

is given by

$$\frac{1}{n(n-1)}\binom{n}{e_0,e_1,\ldots,e_m}1^{e_0}2^{e_1}\cdots(m+1)^{e_m}\sum_{i=1}^{m}\frac{i}{i+1}e_i,$$

where $e_0 = n - e_1 - e_2 - \cdots$.

In this section, we seek a formula for the number of these trees if the edges are locally oriented as in Section 3.2.

Theorem 7.3.1. *The number of lnc-trees of order n with indegree sequence* $\langle 1^{e_1}2^{e_2}\cdots\rangle$ *and outdegree sequence* $\langle 1^{m_1}2^{m_2}\cdots\rangle$ *(both satisfying* (7.3.1)*) is equal to*

$$(n-1)\cdot\frac{(n-e_0-1)!(n-m_0-1)!}{e_1!e_2!\cdots m_1!m_2!\cdots}\binom{n-1}{e_0+m_0-1}, \tag{7.3.2}$$

where $e_0 = n - e_1 - e_2 - \cdots$ *and* $m_0 = n - m_1 - m_2 - \cdots$.

Proof. Let A be the generating function for noncrossing trees rooted at vertex 1 such that the root has degree 1. Let B be the corresponding generating function for these trees rooted at vertex n of degree 1. In each case, the root does not count towards the generating function. Let T be the generating function for all noncrossing trees. Let x_i, y_i and z mark vertices with indegree i, vertices with outdegree i and number of vertices (excluding the root in the case of A and B) respectively, for $i = 0, \ldots, n-1$.

If we consider vertex 1 as a root then we have the decomposition shown in Figure 7.4, where $R_1, R_2, \ldots, R_r$ are noncrossing subtrees.
Therefore,

$$T = zy_0\sum_{k\geq 1} x_kA^k. \tag{7.3.3}$$

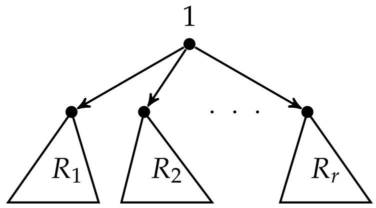

Figure 7.4: Decomposition of noncrossing tree rooted at vertex 1.

Similarly, if we consider vertex n as a root, then we have

$$T = zx_0 \sum_{k \geq 1} y_k B^k. \tag{7.3.4}$$

Moreover,

$$A = z \sum_{k \geq 0} x_k A^k \sum_{\ell \geq 0} y_{\ell+1} B^{\ell}. \tag{7.3.5}$$

This follows from the following decomposition of noncrossing trees of type A (trees of type B can be decomposed analogously).

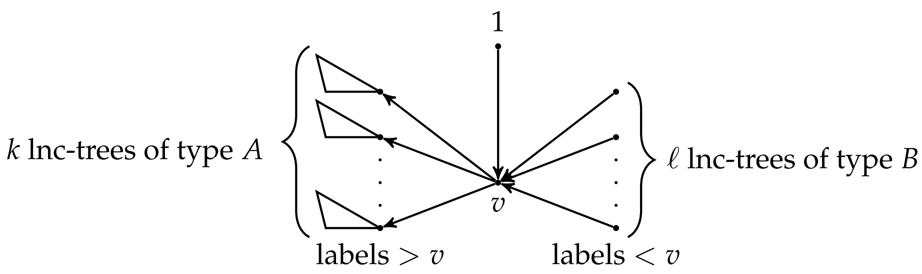

Figure 7.5: Noncrossing tree rooted at vertex 1 with degree 1.

From Equations (7.3.3), (7.3.4) and (7.3.5), we get

$$A = z \left(x_0 + \frac{T}{zy_0} \right) \left(\frac{T}{zx_0 B} \right).$$

Thus,

$$x_0 y_0 z A B = T \left(T + z x_0 y_0 \right). \tag{7.3.6}$$

Setting $T = x_0 y_0 z W$ in Equation (7.3.6), we obtain

$$AB = z x_0 y_0 W \left(1 + W \right). \tag{7.3.7}$$

From Equations (7.3.3) and (7.3.4), we have

$$x_0 W = \sum_{k\geq 1} x_k A^k \tag{7.3.8}$$

and

$$y_0 W = \sum_{k\geq 1} y_k B^k. \tag{7.3.9}$$

Equation (7.3.8) implies that $A = \phi(x_0 W)$, where ϕ is implicitly given by

$$t = \sum_{k\geq 1} x_k \phi(t)^k. \tag{7.3.10}$$

Likewise, Equation (7.3.9) implies that $B = \psi(y_0 W)$, where ψ is implicitly given by

$$u = \sum_{k\geq 1} y_k \psi(u)^k.$$

So, from Equation (7.3.7) we have

$$\phi(x_0 W)\psi(x_0 W) = z x_0 y_0 W\,(1+W)\,.$$

Hence,

$$W = z\cdot(1+W)\cdot\frac{x_0 W}{\phi(x_0 W)}\cdot\frac{y_0 W}{\psi(y_0 W)}.$$

By the Lagrange Inversion Formula [58, Theorem 5.4.2], one has

$$\begin{aligned}
[z^n x_0^{e_0} y_0^{m_0}]T &= [z^{n-1} x_0^{e_0-1} y_0^{m_0-1}]W \\
&= \frac{1}{n-1}[w^{n-2} x_0^{e_0-1} y_0^{m_0-1}]\left((1+w)\cdot\frac{x_0 w}{\phi(x_0 w)}\cdot\frac{y_0 w}{\psi(y_0 w)}\right)^{n-1} \\
&= \frac{1}{n-1}\binom{n-1}{n-e_0-m_0}[t^{e_0-1}]\left(\frac{t}{\phi(t)}\right)^{n-1}[u^{m_0-1}]\left(\frac{u}{\psi(u)}\right)^{n-1}.
\end{aligned} \tag{7.3.11}$$

Applying the Lagrange Inversion Formula again, we obtain

$$\begin{aligned}
[t^{e_0-1}]\left(\frac{t}{\phi(t)}\right)^{n-1} &= [t^{-(n-e_0)}]\phi(t)^{-(n-1)} \\
&= \frac{n-1}{n-e_0}[t^{n-1}]\left(\phi^{\langle -1\rangle}(t)\right)^{n-e_0}.
\end{aligned}$$

Now by Equation (7.3.10), we get

$$\begin{aligned}
[t^{e_0-1}]\left(\frac{t}{\phi(t)}\right)^{n-1} &= \frac{n-1}{n-e_0}[t^{n-1}]\left(\sum_{k\geq 1} x_k t^k\right)^{n-e_0} \\
&= \frac{n-1}{n-e_0}[t^{n-1}]\sum_{n\geq e_0}\sum_{e_1+e_2+\cdots=n-e_0}\frac{(n-e_0)!}{e_1!e_2!\cdots}x_1^{e_1}x_2^{e_2}\cdots t^{e_1+2e_2+\cdots} \\
&= \frac{n-1}{n-e_0}\sum_{\substack{e_1+e_2+\cdots=n-e_0 \\ e_1+2e_2+\cdots=n-1}}\frac{(n-e_0)!}{e_1!e_2!\cdots}x_1^{e_1}x_2^{e_2}\cdots . \qquad (7.3.12)
\end{aligned}$$

Similarly,

$$[u^{m_0-1}]\left(\frac{u}{\psi(u)}\right)^{n-1} = \frac{n-1}{n-m_0}\sum_{\substack{m_1+m_2+\cdots=n-m_0 \\ m_1+2m_2+\cdots=n-1}}\frac{(n-m_0)!}{m_1!m_2!\cdots}y_1^{m_1}y_2^{m_2}\cdots . \qquad (7.3.13)$$

Plugging Equations (7.3.12) and (7.3.13) into (7.3.11), we obtain

$$\begin{aligned}
&[z^n x_0^{e_0}x_1^{e_1}x_2^{e_2}\cdots y_0^{m_0}y_1^{m_1}y_2^{m_2}\cdots]T \\
&\qquad = (n-1)\binom{n-1}{n-e_0-m_0}\frac{(n-e_0-1)!}{e_1!e_2!\cdots}\frac{(n-m_0-1)!}{m_1!m_2!\cdots}.
\end{aligned}$$

This completes the proof. □

Corollary 7.3.2. *The number of lnc-trees of order n with an indegree sequence $\lambda = \langle 1^{e_1}2^{e_2}\cdots\rangle$ is equal to*

$$\frac{(2n-2)!}{(n+e_0-1)!e_1!e_2!\cdots}, \qquad (7.3.14)$$

where $e_0 = n - e_1 - e_2 - \cdots$.

Proof. Let $[[n,k]]$ denote the set of all types of partitions of $[n]$ of length k. Since

$$\sum_{P\in[[n-1,n-m_0]]}\frac{(n-m_0)!}{m_1!m_2!\cdots} = \binom{n-2}{n-m_0-1},$$

it follows from Equation (7.3.2) that there are

$$\begin{aligned}
&\frac{n-1}{n-m_0}\cdot\frac{(n-e_0-1)!}{e_1!e_2!\cdots}\binom{n-2}{n-m_0-1}\binom{n-1}{e_0+m_0-1} \\
&\quad = \frac{(n-e_0-1)!}{e_1!e_2!\cdots}\binom{n-1}{n-m_0}\binom{n-1}{e_0+m_0-1} \qquad (7.3.15)
\end{aligned}$$

noncrossing trees on n vertices with m_0 sinks and having an indegree sequence $\langle 1^{e_1}2^{e_2}\cdots\rangle$. Now, summing over all m_0 in Equation (7.3.15) we obtain the required equation. □

Corollary 7.3.3. *The mean number of sinks in lnc-trees on n labelled vertices and with an indegree sequence $\langle 1^{e_1}2^{e_2}\cdots\rangle$ is given by*

$$\mu = \frac{n-e_0+1}{2}, \tag{7.3.16}$$

where $e_0 = n - e_1 - e_2 - \cdots$.

Proof. From Equation (7.3.15) we find that the total number of sinks in trees with an indegree sequence $\langle 1^{e_1}2^{e_2}\cdots\rangle$ is equal to

$$\begin{aligned}
T =& \frac{(n-e_0-1)!}{e_1!e_2!\cdots}\sum_{m_0} m_0\binom{n-1}{m_0-1}\binom{n-1}{e_0+m_0-1}\\
=& \frac{(n-e_0-1)!}{e_1!e_2!\cdots}\left[\sum_{m_0}(m_0-1)\binom{n-1}{m_0-1}\binom{n-1}{e_0+m_0-1}\right.\\
&\left.+\sum_{m_0}\binom{n-1}{m_0-1}\binom{n-1}{e_0+m_0-1}\right]\\
=& \frac{(n-e_0-1)!}{e_1!e_2!\cdots}\left[\sum_{m_0}(n-1)\binom{n-2}{m_0-2}\binom{n-1}{e_0+m_0-1}\right.\\
&\left.+\sum_{m_0}\binom{n-1}{m_0-1}\binom{n-1}{e_0+m_0-1}\right]\\
=& \frac{(n-e_0-1)!}{e_1!e_2!\cdots}\left[(n-1)\binom{2n-3}{n-e_0-2}+\binom{2n-2}{n-e_0-1}\right]\\
=& \frac{(n-e_0+1)(2n-2)!}{2(n+e_0-1)!e_1!e_2!\cdots}.
\end{aligned}$$

Thus the mean number of sinks is

$$\mu = \frac{(n-e_0+1)(2n-2)!}{2(n+e_0-1)!e_1!e_2!\cdots}\times\frac{(n+e_0-1)!e_1!e_2!\cdots}{(2n-2)!} = \frac{n-e_0+1}{2}.$$

□

Corollary 7.3.4. *The variance of the number of sinks in lnc-trees on n vertices and with an indegree sequence $\langle 1^{e_1}2^{e_2}\cdots\rangle$ is given by*

$$\sigma^2 = \frac{n^2-e_0^2-2n+1}{8n-12},$$

where $e_0 = n - e_1 - e_2 - \cdots$.

Proof. We compute the second moment:

$$\begin{aligned}
S =& \frac{(n-e_0-1)!}{e_1!e_2!\cdots} \sum_{m_0} m_0^2 \binom{n-1}{m_0-1}\binom{n-1}{e_0+m_0-1} \\
=& \frac{(n-e_0-1)!}{e_1!e_2!\cdots} \left[\sum_{m_0} (m_0-1)(m_0-2) \binom{n-1}{m_0-1}\binom{n-1}{e_0+m_0-1} \right. \\
&\left. + 3\sum_{m_0}(m_0-1)\binom{n-1}{m_0-1}\binom{n-1}{e_0+m_0-1} + \sum_{m_0}\binom{n-1}{m_0-1}\binom{n-1}{e_0+m_0-1} \right] \\
=& \frac{(n-e_0-1)!}{e_1!e_2!\cdots} \left[(n-1)(n-2)\sum_{m_0}\binom{n-3}{m_0-3}\binom{n-1}{e_0+m_0-1} \right. \\
&\left. + 3(n-1)\sum_{m_0}\binom{n-2}{m_0-2}\binom{n-1}{e_0+m_0-1} + \sum_{m_0}\binom{n-1}{m_0-1}\binom{n-1}{e_0+m_0-1} \right] \\
=& \frac{(n-e_0-1)!}{e_1!e_2!\cdots} \left[(n-1)(n-2)\binom{2n-4}{n-e_0-3} + 3(n-1)\binom{2n-3}{n-e_0-2} \right. \\
&\left. + \binom{2n-2}{n-e_0-1} \right] \\
=& \frac{(n-e_0-1)!}{e_1!e_2!\cdots} \cdot \frac{(n-2)(n-e_0)^2+3(n-1)(n-e_0)-1}{2(2n-3)} \binom{2n-2}{n-e_0-1}.
\end{aligned} \tag{7.3.17}$$

Since the variance equals

$$\sigma^2 = \frac{S}{T} - \mu^2,$$

where T is the total number of lnc-trees on n vertices with an indegree sequence $\langle 1^{e_1}2^{e_2}\cdots\rangle$, by Equations (7.3.14) and (7.3.17) and a little algebra we obtain the required equation. □

Chapter 8

Reachability in noncrossing trees

Recall from Chapter 4 that a vertex v is *reachable* from vertex u if there is an oriented path from vertex u to vertex v and a path p has length ℓ if there are ℓ edges on the path. In this chapter, we investigate the average number of reachable vertices in locally oriented noncrossing trees (lnc-trees).

8.1 Enumeration by path lengths

Proposition 8.1.1. *Let $N(n, \ell)$ be the total number of vertices that can be reached from vertex 1 in ℓ steps in lnc-trees of order n. Then we have*

$$N(n, \ell) = \frac{3\ell + 1}{3n - 2}\binom{3n - 2}{n - \ell - 1}. \tag{8.1.1}$$

Proof. The generating function for the number of lnc-trees satisfies

$$N(x) = \frac{1}{1 - xN(x)^2},$$

where x is marking the number of non-root vertices in these trees. Consider an lnc-tree in which there is a directed path of length ℓ starting at vertex 1. Let the final vertex of this path be i. The lnc-tree can be decomposed into a path from 1 to i and several lnc-trees attached to it, see Figure 8.1. We obtain the generating function

$$(xN(x))^2(xN(x)^3)^{\ell-1} = x^{\ell+1}N(x)^{3\ell+1}.$$

Set $F(x) = \sqrt{x} \cdot N(x)$ so that

$$F(x) = \frac{\sqrt{x}}{1 - F(x)^2}.$$

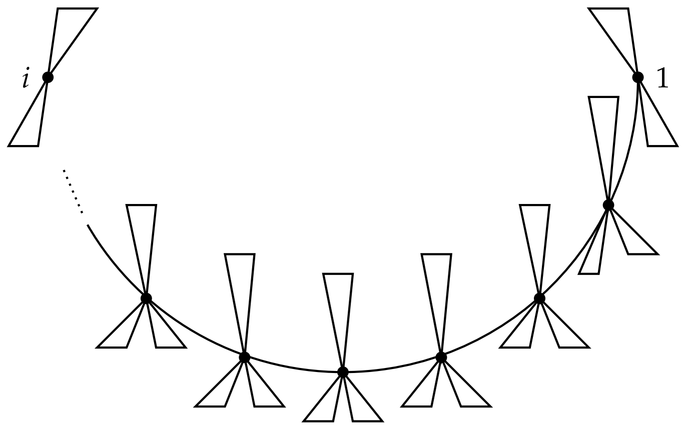

Figure 8.1: Decomposition of lnc-tree with a path of length ℓ.

By the Lagrange Inversion Formula, we have

$$\begin{aligned}
[x^n]x^{\ell+1}N(x)^{3\ell+1} &= [x^{n-\ell-1}]N(x)^{3\ell+1} = [x^{\frac{2n+\ell-1}{2}}]F(x)^{3\ell+1} \\
&= \frac{3\ell+1}{2n+\ell-1}[t^{2n-2\ell-2}](1-t^2)^{-(2n+\ell-1)} \\
&= \frac{3\ell+1}{2n+\ell-1}[t^{2n-2\ell-2}]\sum_{k\geq 0}\binom{-(2n+\ell-1)}{k}(-t^2)^k \\
&= \frac{3\ell+1}{2n+\ell-1}[t^{2n-2\ell-2}]\sum_{k\geq 0}\binom{2n+\ell-2+k}{k}t^{2k} \\
&= \frac{3\ell+1}{2n+\ell-1}\binom{3n-3}{n-\ell-1}.
\end{aligned}$$

This completes the proof. □

Bijective proof of Proposition 8.1.1. In [15, Lemma 1], we find that the number of lattice paths from (i,j) to $(n,2n)$ that do not cross the line $y=2x$ is given by

$$\frac{2i-j+1}{3n-i-j+1}\binom{3n-i-j+1}{n-i}.$$

These paths are called *good* paths. Since Equation (8.1.1) counts the number of good paths from $(0,-3\ell)$ to $(n-\ell-1, 2n-2\ell-2)$, Proposition 8.1.1 will be proved by showing that there is a bijection between these paths and the set of vertices that are reachable in ℓ steps from vertex 1 in lnc-trees on n vertices.

We obtain a complete ternary tree on $3n-2$ vertices from the ternary tree obtained using the procedure described in the proof of Lemma 7.2.2 by attaching leaves so that the initial $n-1$ vertices become internal vertices, each of outdegree three.

We now describe the well-known bijection between complete ternary trees with $n-1$ internal nodes and good paths from $(0,0)$ to $(n-1, 2n-2)$ where each step is of the form $(1,0)$ or $(0,1)$. Traversing a complete ternary tree with $n-1$ internal vertices in preorder (i.e visit vertex, right-child, middle-child, left-child in this order) and drawing a $(1,0)$ step for each internal vertex and a $(0,1)$ step for each leaf (except the last one), we obtain a good path ending at $(n-1, 2n-2)$. Thus given an lnc-tree, one obtains a good path via the associated ternary tree.

Let $k_1, k_2, \ldots, k_\ell$ be a sequence of edges on the path from vertex 1 to v so that v is reachable from 1 in ℓ steps. Mark each of these ℓ edges by an asterisk $(*)$. There are $n-\ell-1$ edges that are not marked. When obtaining a ternary tree on $n-1$ vertices mark all the vertices that are associated with marked edges in the lnc-tree. We traverse the resultant complete ternary tree in preorder to obtain a lattice path from $(0, -3\ell)$ to $(n-\ell-1, 2n-2\ell-2)$, drawing a $(1,0)$ step for each unmarked internal vertex and a $(0,1)$ step for each marked internal vertex and each leaf (except the last one).

Consider a complete ternary tree in which the first marked vertex is at distance $k \geq 0$ from the root and all the marked vertices are on a path consisting only of middle-edges such that all the children of these marked vertices are leaves. All the children of the unmarked internal vertices from the root to the first marked vertex are leaves as well, except the left-child of the root which may be a leaf or not. Such a tree is the worst possible case in that the vertical steps of the path occur at the earliest possible time. The lattice path associated with this ternary tree has a vertical path of length 3ℓ starting at $(k, -3\ell + 2k)$. The remaining part of the tree to be traversed has $n-\ell-k-1$ internal vertices. The paths associated with this part are the good paths from $(k, 2k)$ to $(n-\ell-1, 2n-2\ell-2)$. Thus the paths do not cross the line $y = 2x$.

We remark the following:

Remark 8.1.2. In the ternary tree:

1. Either the root or a vertex u, with the property that the edge connecting u to its parent is a right edge and the path from the root to u consists only of right-edges, is marked.

2. All other markings are on the subtree rooted at u_m where (u, u_m) is the middle-edge coming out of u or out of the root, and the path from a marked vertex s to u_m consists only of middle-edges and right-edges.

3. If v is marked and (v, u) is a right edge, where u is either a child or parent of v, then u is not marked.

4. If there is a path consisting only of middle-edges from a marked vertex u to another marked vertex w, where u is the root or a vertex connected to its parent by a right-edge, then all the other vertices on this path are marked as well.

We now describe the reverse procedure: Consider a lattice path from $(0, -3\ell)$ to $(n - \ell - 1, 2n - 2\ell - 2)$ such that the path does not cross the line $y = 2x$. We obtain a complete ternary tree on $3n - 2$ vertices such that out of the $n - 1$ internal vertices, ℓ are marked. The steps are outlined below:

1. We start with a vertex, which will be the root. If there is a vertical step coming out of $(0, -3\ell)$, then mark the root with an asterisk and all other $(\ell - 1)$ marked vertices will be on the subtree rooted at the middle-child of the root. Otherwise, create a right-edge to a new vertex. Continue until you find a vertical step. Mark the vertex that corresponds to this vertical step.

2. In preorder, draw an internal vertex for any horizontal step and a leaf for any vertical step until you get to the marked vertex.

3. If the next step is vertical, then draw a middle-edge coming out of the marked vertex to a new vertex. Mark this new vertex as well. Otherwise, do as in step 2. There is no marked vertex on the path consisting only of middle-edges and starting at the marked vertex. The next marked vertex is connected to its parent by a right-edge.

4. Repeat step 3 until you obtain a complete ternary tree with ℓ marked internal vertices.

□

As an example, consider the lnc-tree on [6] shown in Figure 8.2. To obtain the good path that corresponds to vertex 4, reachable from vertex 1 in two steps, we mark each edge on the path from 1 to 4 with an asterisk ($*$). The associated ternary tree has 5 vertices and its corresponding complete ternary tree has 16 vertices as shown in Figure 8.3.

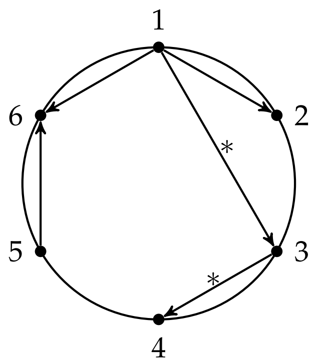

Figure 8.2: Lnc-tree on 6 vertices.

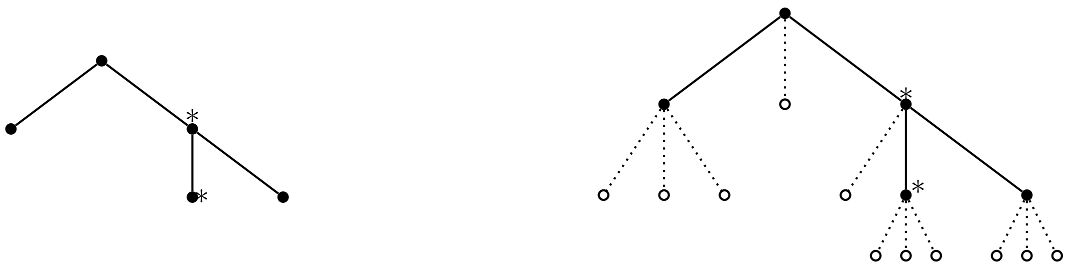

Figure 8.3: Ternary tree on 5 vertices with two marked vertices and its associated complete ternary tree with two marked vertices.

By the process described above, we obtain a lattice path that starts at $(0, -6)$ and ends at $(3, 6)$ and does not cross the line $y = 2x$. This path is shown in Figure 8.4.

By setting $\ell = 1$ in Equation (8.1.1) we obtain:

Corollary 8.1.3. *There are a total of*

$$\frac{2}{n}\binom{3n-3}{n-2} \tag{8.1.2}$$

children of vertex 1 *in all lnc-trees on* $[n]$.

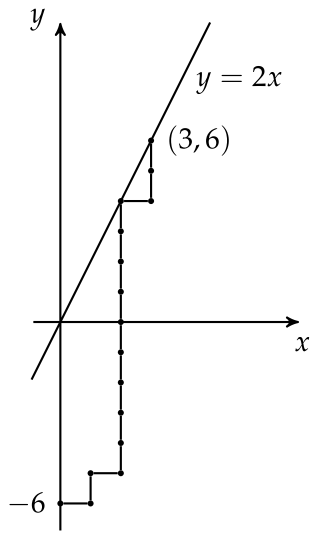

Figure 8.4: Good path

Since there are

$$\frac{1}{n-1}\binom{3n-3}{n-2}$$

lnc-trees of order n, by Equation (8.1.2) the average number of children of vertex 1 is given by

$$2\left(1-\frac{1}{n}\right).$$

By taking $n \to \infty$, it follows that in an average lnc-tree there are approximately two children of vertex 1.

Corollary 8.1.4. *The number of vertices that are reachable from vertex* 1 *in lnc-trees of order n is equal to*

$$2\binom{3n-3}{n-1} - \sum_{i=0}^{n-1}\binom{3n-3}{i}. \tag{8.1.3}$$

Proof. The result follows by summing over all ℓ in Equation (8.1.1). □

Corollary 8.1.5. *The average number of vertices that are reachable from vertex* 1 *in lnc-trees of order n is*

$$(2n-1)\left[2-\binom{3n-3}{n-1}^{-1}\sum_{i=0}^{n-1}\binom{3n-3}{i}\right].$$

Proof. We obtain the desired formula by dividing Equation (8.1.3) by the total number of lnc-trees of order n, which is

$$\frac{1}{2n-1}\binom{3n-3}{n-1}.$$

□

Proposition 8.1.6. *The number of lnc-trees on n vertices such that vertex j is reachable in $j-i$ steps from vertex i is given by*

$$\frac{j-i+1}{2n-j+i-1}\binom{3n-2j+2i-3}{n-j+i-1}. \tag{8.1.4}$$

Proof. The generating function for the number of noncrossing trees satisfies

$$N(x) = \frac{1}{1-xN(x)^2},$$

where x is marking the number of non-root vertices in the tree.

Consider an lnc-tree in which there is an oriented path of length ℓ starting at i and ending at j such that $\ell = j-i$. We have a decomposition of the noncrossing tree into noncrossing trees rooted at the $\ell+1$ vertices of the path. See Figure 8.5 below.

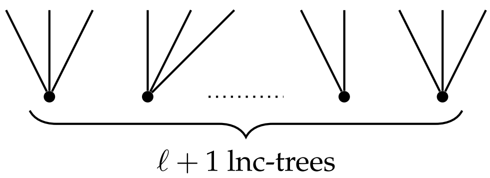

Figure 8.5: Decomposition of lnc-tree.

Therefore the generating function for the number of these trees is

$$(xN(x))^{\ell+1} = \frac{x^{\ell+1}}{(1-xN(x)^2)^{\ell+1}}.$$

Set $F(x) = \sqrt{x}\cdot N(x)$ so that

$$F(x) = \frac{\sqrt{x}}{1-F(x)^2}.$$

Now, by the Lagrange Inversion Formula [58, Theorem 5.4.2], the number of trees on n vertices such that vertex j is reachable from vertex i in $\ell = j - i$ steps is therefore given as

$$\begin{aligned}
[x^n]x^{\ell+1}N(x)^{\ell+1} = [x^{n-\ell-1}]N(x)^{\ell+1} &= [x^{n-\frac{\ell}{2}-\frac{1}{2}}]F(x)^{\ell+1} \\
&= \frac{\ell+1}{2n-\ell-1}[t^{2n-2\ell-2}](1-t^2)^{-(2n-\ell-1)} \\
&= \frac{\ell+1}{2n-\ell-1}[t^{2n-2\ell-2}]\sum_{k\geq 0}\binom{-(2n-\ell-1)}{k}(-t^2)^k \\
&= \frac{\ell+1}{2n-\ell-1}[t^{2n-2\ell-2}]\sum_{k\geq 0}\binom{2n-\ell-2+k}{k}t^{2k} \\
&= \frac{\ell+1}{2n-\ell-1}\binom{3n-2\ell-3}{n-\ell-1}.
\end{aligned}$$

We obtain the required formula by setting $\ell = j - i$. □

Corollary 8.1.7. *The number of lnc-trees of order n in which there is an oriented path of length $n-i$ from i to n is given by*

$$\frac{n-i+1}{n+i-1}\binom{n+2i-3}{i-1}.$$

Proof. The result follows from Equation (8.1.4) by setting $j = n$. □

The sequence in Corollary 8.1.7 appears as A069269 in the OEIS [55].

8.2 Trees with exact number of reachable vertices

The number of lnc-trees on n vertices in which all the vertices are reachable from vertex 1 is given by the $(n-1)^{\text{th}}$ Catalan number,

$$\frac{1}{n}\binom{2n-2}{n-1}.$$

These trees are called *noncrossing increasing trees* [2].

Corollary 8.2.1. *The number of noncrossing increasing trees on n vertices with an outdegree sequence $\langle 1^{m_1}2^{m_2}\cdots\rangle$ is given by*

$$\frac{(n-1)!}{m_0!m_1!m_2!\cdots}$$

where $m_0 = n - m_1 - m_2 - \cdots$.

Proof. Since the indegree sequence in all noncrossing increasing trees on n vertices is $\langle 1^{n-1}\rangle$ (and all trees with this indegree sequence are increasing), the result follows immediately from Equation (7.3.2). $\square$

In the next theorem, we prove a result that generalizes the formula for the number of noncrossing increasing trees.

Theorem 8.2.2. *The number of lnc-trees on n vertices such that exactly $k \geq 2$ vertices are reachable from vertex 1 is given by*

$$\frac{1}{2n-k-1}\binom{3n-2k-2}{n-k}\binom{2k-2}{k}.$$

Proof. We delete vertex 1 and all the edges that are on a path from vertex 1 to a vertex which is reachable from vertex 1. We obtain a forest of lnc-trees rooted at the $k-1$ vertices (other than vertex 1) that were reachable from vertex 1. We relabel the vertices such that the vertex with the smallest label takes label 1, second smallest label takes label 2, and so on. By setting $i = 1$ in Proposition 8.1.6, we obtain that the number of these forests (on $n-1$ vertices) is given by

$$\frac{k-1}{2n-k-1}\binom{3n-2k-2}{n-k}.$$

Since the number of noncrossing increasing trees on k vertices is

$$\frac{1}{k}\binom{2k-2}{k-1} = \frac{1}{k-1}\binom{2k-2}{k},$$

the total number of lnc-trees on n vertices such that exactly k vertices are reachable from vertex 1 is

$$\frac{k-1}{2n-k-1}\binom{3n-2k-2}{n-k}\cdot\frac{1}{k-1}\binom{2k-2}{k}$$

where $k \geq 2$. $\square$

List of references

[1] M. Aigner and G. M. Ziegler. *Proofs from THE BOOK*, volume **4**. Springer-Verlag / Berlin / Heidelberg, 2010.

[2] A. Asinowski and T. Mansour. Dyck paths with coloured ascents. *European J. Combin.*, **29**(5):1262 – 1279, 2008.

[3] F. Bédard and A. Goupil. The poset of conjugacy classes and decomposition of products in the symmetric group. *Canad. Math. Bull.*, **35**(5):152–160, 1992.

[4] M. Bóna, M. Bousquet, G. Labelle, and P. Leroux. Enumeration of m-ary cacti. *Adv. Appl. Math.*, **24**(1):22–56, 2000.

[5] C. W. Borchardt. Ueber eine der Interpolation entsprechende Darstellung der Eliminations-Resultante. *J. reine angew. Math.*, **57**:111–121, 1860.

[6] A. Cayley. A theorem on trees. *Quart. J. Math.*, **23**:376–378, 1889.

[7] C. Chauve, S. Dulucq, and O. Guibert. Enumeration of some labelled trees. In *Formal Power Series and Algebraic Combinatorics*, pages 146–157. Springer, 2000.

[8] C. Chauve, S. Dulucq, and A. Rochnitzer. Enumerating alternating trees. *J. Combin. Theory, Ser. A*, **94**:142–151, 2001.

[9] W. Y. Chen. A general bijective algorithm for trees. *Proc. Nati. Acad. Sci. USA*, **87**:9635–9639, 1990.

[10] J. Cigler. Some Remarks on Catalan Families. *European J. Combin.*, **8**(3):261 – 267, 1987.

[11] L. Comtet. *Advanced Combinatorics: The Art of Finite and Infinite Expansions*. D. Reidel Publishing Company, 1974.

[12] E. Cotterill. Geometry of curves with exceptional secant planes: linear series along the general curve. *Math. Z.*, **267**(3-4):549–582, 2011.

[13] J. Dénes. The representation of a permutation as the product of a minimal number of transpositions and its connection with the theory of graphs. *A Magyar Tudományos Akadémia Matematikai Kutató Intézetének Közleményei*, **4**:63 – 70, 1959.

[14] N. Dershowitz and S. Zaks. Enumerations of ordered trees. *Discrete Math.*, **31**(1):9–28, 1980.

[15] E. Deutsch and M. Noy. Statistics on non-crossing trees. *Discrete Math.*, **254**(1–3):75 – 87, 2002.

[16] R. Diestel. *Graph Theory,* vol. 173 of *Graduate Texts in Mathematics, 3rd ed.* Springer-Verlag, Berlin, 2005.

[17] R. R. X. Du and J. Yin. Counting labelled trees with a given indegree sequence. *J. Combin. Theory, Ser. A*, **117**(3):345 – 353, 2010.

[18] S. Dulucq and J. Penaud. Cordes, arbres et permutations. *Discrete Math.*, **117**(1-3):89–105, 1993.

[19] P. H. Edelman. Chain enumeration and non-crossing partitions. *Discrete Math.*, **31**(2):171–180, 1980.

[20] M. Eden and M. P. Schützenberger. Remark on a theorem of Dénes. *A Magyar Tudományos Akadémia Matematikai Kutató Intézetének Közleményei*, **7**:353–355, 1962.

[21] G. Evans, J. Blackledge, and P. Yardley. *Analytic Methods for Partial Differential Equations*. Springer, 2000.

[22] P. Flajolet and M. Noy. Analytic combinatorics of non-crossing configurations. *Discrete Math.*, **204**(1-3):203 – 229, 1999.

[23] G. W. Ford and G. E. Uhlenbeck. Combinatorial Problems in the Theory of Graphs, I. *Proc. Nat. Acad. Sciences*, **42**:122–128, 1956.

[24] I. M. Gelfand, M. I. Graev, and A. Postnikov. Combinatorics of hypergeometric functions associated with positive roots. In *The Arnold-Gelfand mathematical seminars*, pages 205–221. Springer, 1997.

[25] I. P. Goulden and D. M. Jackson. *Combinatorial Enumeration*. John-Wiley & Sons, 1983.

[26] I. P. Goulden and D. M. Jackson. The Combinatorial Relationship Between Trees, Cacti and Certain Connection Coefficients for the Symmetric Group. *European J. Combin.*, **13**(5):357 – 365, 1992.

[27] I. P. Goulden and S. Pepper. Labelled trees and factorizations of a cycle into transpositions. *Discrete Math.*, **113**:253 – 268, 1993.

[28] R. L. Graham, D. E. Knuth, and O. Patashnik. *Concrete Mathematics*. Addison-Wesley Publishing Company, 1990.

[29] F. Harary and G. E. Uhlenbeck. On the number of Husimi trees. *Proc. Nat. Aca. Sci.*, **39**:315–322, 1953.

[30] K. Husimi. Note on Mayers' Theory of Cluster Integrals. *J. Chem. Phys.*, **18**:682–684, 1950.

[31] M. Janjić. Enumerative formulae for some functions on finite sets. Available online at http://www.pmfbl.org/janjic/enumfun.pdf.

[32] A. Joyal. Une théorie combinatoire des séries formelles. *Adv. Math.*, **42**:1–82, 1981.

[33] J. S. Kim. Bijections on rooted trees with fixed size of maximal decreasing subtrees. *Ann. Combin.*, pages 339–352, 2013.

[34] G. Kreweras. Sur les partitions non croisees d'un cycle. *Discrete Math.*, **1**(4):333–350, 1972.

[35] P. Leroux. Enumerative problems inspired by Mayer's theory of cluster integrals. *Electron. J. Combin.*, **11**(1), 2004.

[36] J. L. Martin and V. Reiner. Factorization of some weighted spanning tree enumerators. *J. Combin. Theory, Ser. A*, **104**(2):287–300, 2003.

[37] J. E. Mayer. *Equilibrium Statistical Mechanics: The international encyclopedia of physical chemistry and chemical physics*. Pergamon Press, Oxford, 1968.

[38] A. Meir and J. W. Moon. On the altitude of nodes in random trees. *Canad. J. Math.*, **30**:997 – 1015, 1978.

[39] J. W. Moon. Counting labelled trees. *Canadian Mathematical Congress*, **1**, 1970.

[40] P. Moszkowski. A Solution to a Problem of Dénes: a Bijection Between Trees and Factorizations of Cyclic Permutations. *European J. Combin.*, **10**:13 – 16, 1989.

[41] M. Noy. Enumeration of noncrossing trees on a circle. *Discrete Math.*, **180**(1-3):301–313, 1998.

[42] S. X. M. Pang and L. Lv. A Refinement of Leaves on Noncrossing Trees. *Graphs Combin.*, **29**(1):131–143, 2013.

[43] A. Panholzer and H. Prodinger. Bijection for ternary trees and noncrossing trees. *Discrete Math.*, **250**(1-3):181–195, 2002.

[44] J. Pitman. Coalescent random forests. *J. Combin. Theory, Ser. A*, **85**:165–193, 1999.

[45] G. Pólya. Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und chemische Verbindungen. *Acta Math.*, **68**:145– 254, 1937.

[46] G. Pólya and R. C. Read. *Combinatorial Enumeration of Groups, Graphs and Chemical Compounds*. Springer-Verlag, New York, Berlin, 1987.

[47] A. Postnikov. Intransitive trees. *J. Combin. Theory, Ser. A*, **79**:360–366, 1997.

[48] H. Prodinger. A correspondence between ordered trees and noncrossing partitions. *Discrete Math.*, **46**(2):205–206, 1983.

[49] H. Prüfer. Neuer Beweis eines Satzes über Permutationen. *Archiv der Math. u. Physik*, **3**(47):142–144, 1918.

[50] J. H. Przytycki and A. S. Sikora. Polygon Dissections and Euler, Fuss, Kirkman, and Cayley Numbers. *J. Combin. Theory, Ser. A*, **92**(1):68 – 76, 2000.

[51] J. B. Remmel and S. G. Williamson. Spanning trees and function classes. *Electron. J. Combin.*, **9**(1):229–233, 2002.

[52] S. Seo and H. Shin. On the enumeration of rooted trees with fixed size of maximal decreasing trees. *Discrete Math.*, **312**(2):419–426, 2012.

[53] H. Shin and J. Zeng. A bijective enumeration of labelled trees with given indegree sequence. *J. Combin. Theory, Ser. A*, **118**(1):115–128, 2011.

[54] R. Simion. Noncrossing partitions. *Discrete Math.*, **217**(1–3):367– 409, 2000.

[55] N. J. A. Sloane. The On-Line Encyclopedia of Integer Sequences (OEIS). Available online at http://oeis.org.

[56] C. Springer. Factorizations, trees, and cacti. *Proceedings of the Eighth International Conference on Formal Power Series and Algebraic Combinatorics (FPSAC)*, University of Minnesota, **23**:427–438, 1996.

[57] R. P. Stanley. Exercises on Catalan numbers and related numbers. Available at http://www-math.mit.edu/∼rstan/ec/.

[58] R. P. Stanley. *Enumerative Combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999.

[59] J. J. Sylvester. On the change of systems of independent variables. *Quart. J. Math.*, **1**:126–134, 1857.

[60] E. Teufl and S. Wagner. The Number of Spanning Trees in Self-Similar Graphs. *Ann. Combin.*, **15**:355–380, 2011.

[61] E. Tzanaki. Polygon dissections and some generalizations of cluster complexes. *J. Combin. Theory, Ser. A*, **113**(6):1189 – 1198, 2006.

[62] S. Wagner. Labeled trees, functions, and an algebraic identity. *Electron. J. Combin.*, **18**, 2011.

[63] H. S. Wilf. *Generatingfunctionology*. A. K. Peters, Ltd., Natick, MA, USA, 2006.