TECHNISCHE UNIVERSITÄT WIEN
Vienna University of Technology

D I P L O M A R B E I T

Enumeration of Lambda-terms and Directed Acyclic Graphs

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

unter der Anleitung von

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

durch

Isabella Larcher
Friedhofgasse 5
2230 Gänserndorf

Datum

Unterschrift

D I P L O M A T H E S I S

Enumeration of Lambda-terms and Directed Acyclic Graphs

Author:

Isabella Larcher
Friedhofgasse 5
2230 Gänserndorf

Supervisor:

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

Abstract

This thesis is divided into two parts. The first part is devoted to the investigation of lambda-terms. Our main goal will be to determine the number of certain lambda-terms of a given size, as well as their asymptotic number as their size tends to infinity. We will derive asymptotic results for various subclasses of lambda-terms, while the problem of counting unrestricted lambda-terms is still unsolved. In order to give an explanation for that, we will explicate the difficulties that have to be overcome in order to be able to establish the asymptotics of lambda-terms. Finally, we will focus on some structural properties of lambda-terms, thereby showing that the structure of lambda-terms highly differs from that of trees, although lambda-terms are closely related to Motzkin-trees.

In the second part of this thesis we will investigate both labeled and unlabeled directed acyclic graphs (DAGs). The asympotics for labeled DAGs can be determined easily while the asymptotic behavior of unlabeled DAGs is still unknown. Furthermore we will show some very interesting results on the structure of DAGs, e.g. asymptotically almost all DAGs are weakly connected. At last, we will investigate two subclasses of DAGs that are of special interest, namely (labeled and unlabeled) extensional and (labeled) essential DAGs.

Zusammenfassung

Enstprechend dem Titel, besteht diese Diplomarbeit aus zwei Teilen. Der erste Teil behandelt Lambda-Terme. Insbesondere sind wir an der Anzahl bestimmter Lambda-Terme einer gewissen Größe interessiert, sowie an der asymptotischen Anzahl, wenn ihre Größe gegen unendlich geht. Wir werden die asymptotische Anzahl an Termen für spezielle Unterklassen von Lambda-Termen herleiten, jedoch nicht für die Klasse aller Lambda-Terme. Dieses Problem ist bisher noch offen und wir werden sehen, dass noch einige Schwierigkeiten bewältigt werden müssen um eine Lösung dafür zu finden. Am Ende des ersten Teils zeigen wir noch einige interessante Eigenschaften über die Struktur von Lambda-Termen. Dabei wird insbesondere auffallen, dass sich diese Struktur grundlegend von der von Bäumen unterscheidet, obwohl Lambda-Terme in enger Beziehung zu Motzkin-Bäumen stehen.
Im zweiten Teil dieser Diplomarbeit untersuchen wir sowohl markierte als auch unmarkierte gerichtete azyklische Graphen (directed acyclic graphs, DAGs). Im Fall von markierten DAGs lässt sich das asymptotische Verhalten einfach bestimmen, während das für unmarkierte DAGs noch unbekannt ist. Weiters werden wir einige interessante Resultate über die Struktur von DAGs präsentieren, z.B. dass asymptotisch fast alle DAGs schwach zusammenhängend sind. Zum Schluss werden wir noch zwei Unterklassen von DAGs untersuchen, die von besonderem Interesse sind, nämlich sogennante (markierte und unmarkierte) extensionale DAGs und (markierte) essentielle DAGs.

Acknowledgments

First of all I want to thank my supervisor, Dr. Bernhard Gittenberger, for introducing me to the fascinating field of combinatorics and for his support and encouragement during the whole writing process.

Furthermore I want to thank my family, who made all this possible, for believing in me and supporting me throughout my whole life.
I also want to thank my boyfriend for his constant encouragement and for always being there for me.

Finally, I want to thank my dear colleagues, who were never too tired to help me with any problems and who became close friends.

Contents

Introduction

The problems of enumerating lambda-terms and directed acyclic graphs (DAGs) are popular topics in combinatorics, which have been investigated thoroughly over the past years. At the first sight both lambda-terms and DAGs appear to be very simple structures, in the sense that their construction can easily be described. The fascination for these structures therefore arises from the fact that there are still a lot of unanswered questions about their number and structural properties. The lambda-calculus was invented in the 1930s by Church and Kleene for the investigation of decision problems and after more than 80 years there are still many open problems concerning the number of lambda-terms. So far no one has yet accomplished to derive the asymptotics of the number of lambda-terms due to the fact that their number increases very fast (superexponentially) with increasing size, which makes the common methods of analytic combinatorics inapplicable.

Lambda-terms can to some extent (by identifying them with a special kind of tree-like structure) be regarded as DAGs, while reversely most DAGs do not correspond to lambda-terms. Consequently, the increase of the number of DAGs is even greater than that of lambda-terms. Therefore it seems likely to assume that the problem of counting DAGs is unsolved as well. But we will see that by introducing a special generating function the asymptotics for the number of labeled DAGs can actually be determined. Unfortunately, in case of unlabeled DAGs the asymptotic behavior is still unknown.

The aim of this thesis is to give an thorough introduction to the combinatorics of lambda-terms and DAGs. We will discuss both their number and structure, and investigate some interesting subclasses.
According to the title this thesis is divided into two parts. In the first part we will investigate the number and structure of lambda-terms. Their formal definition will be given in Chapter 1, where further important definitions and notations will be introduced.
Our fundamental tools for gaining information on the number of a certain combinatorial class are generating functions together with the symbolic method, singularity analysis and the saddle-point method, which will be introduced in Chapter 2.
As mentioned before, unfortunately these concepts cannot be applied to the class of lambda-terms. In Chapter 3 we will explain in detail why this is the case and give lower and upper bounds for the number of lambda-terms of a certain size. The exact number of lambda-terms of a given size, as well as the asymptotic number, still remain unsettled and it seems as there are a lot of obstacles which need to be conquered in order to solve them. However, there are results on the asymptotic behavior of certain subclasses of lambda-terms, which allow the use of the methods of analytic combinatorics introduced in Chapter 2. An exhaustive treatment of the various subclasses that have been investigated so far will be presented in Chapter 4.
In Chapter 5 we will briefly give some results on the structure of lambda-terms. It is very interesting that although lambda-terms are somewhat related to Motzkin trees, their structure is totally different compared to that of trees, while it rather resembles that of DAGs.

The second part of this thesis is devoted to the investigation of both labeled and unlabeled DAGs. In Chapter 6 we will introduce some basic definitions and explain the relation between lambda-terms and DAGs. In Chapter 7 recurrence relations for the number of DAGs of a certain size will be set up.

In the labeled case the recurrence will be very descriptive, while the recursion for unlabeled DAGs will turn out to be much more complex. Then a special generating function for the number of labeled DAGs will be introduced that adopts an intriguing representation. Moreover, in Chapter 7 we will present a bijection between the number of labeled DAGs and the number of (0,1)-matrices with positive eigenvalues.

Chapter 8 will be concerned with the investigation of the asymptotic behavior of DAGs. Due to the explicit representation of the special generating function of labeled DAGs, their asymptotics can easily be determined. However, as mentioned before, the asymptotic behavior of unlabeled DAGs appears to remain unsolved. Nevertheless, by fixing the number of edges we establish asymptotics for both labeled and unlabeled DAGs, which will also be presented in Chapter 8.

In Chapter 9 we will investigate some properties of the structure of labeled DAGs. First we will see that almost all DAGs are weakly connected. Then we will derive an interesting result on the average number of out-points of a random DAG, namely that for sufficiently large sizes the number of out-points is less than 4. As a last result on the structure of labeled DAGs we will show that the height of DAGs is asymptotically normally distributed.

Finally, in Chapter 10, we will consider two subclasses of DAGs that are of special interest, namely extensional and essential DAGs.

Part I

Lambda-terms

Chapter 1

Basic definitions

In this chapter we introduce the basic definitions and notations that will be used throughout the first part of this thesis. This contains the definition of lambda-terms and lambda-trees, and since we are interested in the enumeration of those objects we also need to define their size. In doing so we will stick to the definitions given in [2] and [18].

First of all we have to define what a lambda-term is.

Definition 1.1 (lambda-terms, [18, Def. 3])
Let $\mathcal{V}$ be a countable set of variables. The set Λ of lambda-terms is defined by the following grammar:

1. *every variable in $\mathcal{V}$ is a lambda-term,*
2. *if T and S are lambda-terms then TS is a lambda-term,*
3. *if T is a lambda-term and x is a variable then $\lambda x.T$ is a lambda-term.*

We call TS an application and $\lambda x.T$ an abstraction.

Remark 1.2
The name application arises, since lambda-terms of the form TS can be regarded as functions $T(S)$, where the function T is applied to S, which in turn can be a function itself.

Both application and repeated abstraction are not commutative, i.e. in general the lambda-terms TS and ST, as well as $\lambda x.\lambda y.M$ and $\lambda y.\lambda x.M$, are different (with the exceptions of $T = S$ and none of the variables x or y occurring in M).

Definition 1.3 (bound, free variables, [2, Def. 2.1.6])
A variable x occurs free in a lambda-term if it is not in the scope of a $\lambda x.$. Otherwise we call it a bound variable.

Each λ binds exactly one variable (which may occur several times in the terms), and each variable can be bound by at most one λ.

Definition 1.4 (open, closed lambda-term)
A lambda-term is closed if it contains no free variables; otherwise it is called open.

Now we introduce some notational conventions that will be used throughout this thesis:

(i) $x, y, z, \ldots$ denote arbitrary variables.

(ii) $M, N, L, \ldots$ denote arbitrary lambda-terms.

(iii) The lambda-term $\lambda x_1. \ldots \lambda x_n.M$ is read as $\lambda x_1.(\lambda x_2.(\ldots(\lambda x_n.M))\ldots)$, whereas $MN_1 \ldots N_n$ is an abbreviation of $(\ldots((MN_1)N_2)\ldots N_n)$.

(iv) The symbol $\equiv$ denotes syntactic equality.

Definition 1.5 (subterm, [2, Def. 2.1.8])
The collection $S(N)$ of subterms of a lambda-term N is defined inductively as follows:

(i) $S(x) = \{x\}$,

(ii) $S(N_1N_2) = S(N_1) \cup S(N_2) \cup \{N_1N_2\}$,

(iii) $S(\lambda x.N_1) = S(N_1) \cup \{\lambda x.N_1\}$.

Lambda-terms can be seen as rooted unary-binary trees containing additional directed edges. Note that the resulting structures are in general no trees in the sense of graph theory, but due to their close relation (see Def.1.6) they are called lambda-trees.

Definition 1.6 (lambda-tree, [18, Def. 5])
With every lambda-term T, the corresponding lambda-tree $G(T)$ can be constructed in the following way:

- *If x is a variable then $G(x)$ is a single node labeled with x. Note that x is unbound.*
- *$G(PQ)$ is a lambda-tree with a binary node as root, having the two lambda-trees $G(P)$ (to the left) and $G(Q)$ (to the right) as subtrees.*
- *The tree $G(\lambda x.P)$ is obtained from $G(P)$ in four steps:*
  1. *Add a unary node as new root.*
  2. *Connect the new root by an undirected edge with the root of $G(P)$.*
  3. *Connect all leaves of $G(P)$ labelled with x by directed edges with the new root.*

4. Remove all labels x from $G(P)$. Note that now x is bound.

Obviously applications correspond to binary nodes and abstractions correspond to unary nodes.

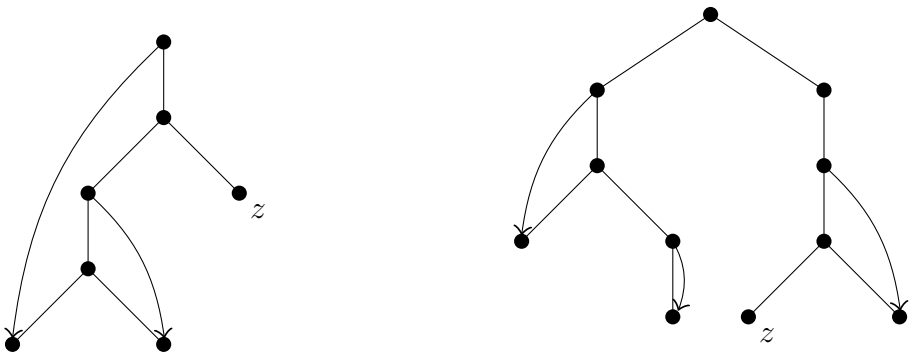


Figure 1.1: The lambda-trees representing the terms $\lambda x.((\lambda y.(xy))z)$ and $(\lambda x.(x(\lambda y.y))(\lambda x.(\lambda y.zy))$.

Removing all directed edges and labels yields the underlying unary-binary tree of the lambda-tree.

Due to this correspondence of lambda-terms and lambda-trees, we sometimes call the variables leaves, and the path connecting the root with a leaf a branch.

There are different approaches as to how one can define the size of a lambda-term, and a lambda-tree respectively. We will introduce the two most common ones and also give an additional and rather unconventional definition of the size, which leads to very interesting results.

Most commonly the size $|T|$ of a lambda-term T is defined in the following way:

Definition 1.7 (size1 of a lambda-term)

$$|x| = 1$$
$$|\lambda x.M| = 1 + |M|$$
$$|MN| = 1 + |M| + |N|$$

As we can see, $|T|$ is the total number of nodes of the lambda-tree $G(T)$.

The other popular definition for the size $\|T\|$ of a lambda-term T, that will also be used in parts of this thesis, is the following:

Definition 1.8 (size2 of a lambda-term)

$$\|x\| = 0$$
$$\|\lambda x.M\| = 1 + \|M\|$$
$$\|MN\| = 1 + \|M\| + \|N\|$$

As we can see, $\|T\|$ is the number of all internal nodes of the lambda-tree G(T).

As an example, for the terms $T = \lambda x.((\lambda y.(xy))z)$ and $S = (\lambda x.(x(\lambda y.y))(\lambda x.(\lambda y.zy))$, pictured in Figure 1.1, it holds that $|T| = 7$, $\|T\| = 4$, $|S| = 11$ and $\|S\| = 7$.

Since we are interested in the enumeration of lambda-terms and the set $\mathcal{V}$ of variables is infinite, we have to introduce further conventions in order to prevent having infinitely many lambda-terms of a certain size.

Definition 1.9 (α-equivalence)
Two lambda-terms M and N are α-equivalent, denoted by $M \equiv_\alpha N$, if they only differ by the names of their bound variables.

In the following we will consider α-equivalent terms to be equal and therefore we write for example $\lambda x.x \equiv \lambda y.y$.

Additionally, we may assume the free variables of a lambda-term to be consecutively numbered with x_1 being the left-most variable occurring in the lambda-term and x_2 being the next free variable that occurs that is not already present in the term and so on. In short, the names of the different free variables are not important in the sense that they can be renamed, i.e. $\lambda x.yx \equiv \lambda x.zx$. Since we will mainly focus on closed lambda-terms, these considerations will not concern us for too long, but for the sake of completeness they have to be mentioned.

At last, we now want to present another interesting approach of defining the size of a lambda-term, introduced in [17], which originated from the field of computer science. Yet before we can give the formal definition, we have to introduce another well-established representation of lambda-terms, called De Bruijn representation (see [14]). The purpose of this representation was to eliminate the occurrences of variable names, which simplifies the process of identifying two equivalent terms.

Definition 1.10 (De Bruijn representation)
In the De Bruijn representation of a lambda-term variables are replaced by positive integers (called De Bruijn indices), that represent the number of λ's on the path connecting a variable with its binding λ in the corresponding lambda-tree. Thus, the abstraction which binds a variable n is the n-th λ before the variable. A variable n is free if the prefix which has this variable as its last symbol contains less than n λ's.

For example, the De Bruijn representation of the lambda-term $\lambda x.\lambda y.xy$ is $\lambda\lambda 21$ and that of $\lambda x.\lambda y.\lambda z.((xz)(yz))$ is $\lambda\lambda\lambda(31)(21)$. This representation

makes the use of variables unnecessary and allows us to say that two lambda-terms are equivalent if they have the same De Bruijn representation. So, what we are actually interested in is the number of different De Bruijn representations of a given size.

Now, following the notation of [17], we will replace the variable n by a sequence of n symbols, namely a string of $n-1$ so-called 'successors' S and a so-called 'zero' 0 at the end. The examples from above can therefore be rewritten in the following form:

$\lambda\lambda 21$, corresponding to the lambda-term $\lambda x.\lambda y.xy$, can be rewritten as $\lambda\lambda(S0)0$ and the term $\lambda\lambda\lambda(31)(21)$, which corresponds to $\lambda x.\lambda y.\lambda z.(xz)(yz)$ has the representation $\lambda\lambda\lambda((SS0)0)((S0)0)$.

Now we are able to define the size of a lambda-term, as it has been done in [17]:

Definition 1.11 (size3 of a lambda-term, [17, p.2])
Considering a lambda-term T to be in the representation introduced above, the size $|T|$ can be computed as follows:
$|0| = a, \qquad |Sn| = |n| + b, \qquad |\lambda M| = |M| + c, \qquad |MN| = |M| + |N| + d.$

Using this definition of the size we have for the examples given above $|\lambda\lambda(S0)0| = 2a + b + 2c + d$ and $|\lambda\lambda\lambda((SS0)0)((S0)0)| = 4a + 3b + 3c + 3d$.
In [17] the authors made the following assumptions about the constants a, b, c, d:

1. a, b, c, d are non-negative integers
2. $a + d \geq 1$

3. $b, c \geq 1$

4. $\gcd(b, c, a + d)$=1

The second assumption arises from the fact that if the zeros and the applications both had size 0 (i.e. $a + d = 0$), this would result in an infinite number of terms of a given size, because any number of applications and zeros could be inserted into a term without increasing its size. The third assumption can be explained similarly, since successors or abstractions having size 0 (i.e. b or c equals to 0), would yield an infinite number of terms of a given size, because one could insert any length of strings of successors or abstractions into a term without increasing its size. The last assumption is a rather technical one, which provides pleasant properties for further estimations (see [17] for more details).

This definition of the size has a very interesting property: Since variables that have many successors have a greater size than those having fewer successors, it follows that the size of a given lambda-term is increased compared to the size we defined in Def. 1.7 and Def. 1.8. This corresponds to a shift of the lambda-terms to higher sizes, with the result that there are fewer terms of small sizes. Thus, the number of lambda-terms increases more slowly with increasing size, which is the reason why, using this definition of the size, they can be counted (cf. [17]). In Chapter 3 we will see why this is not possible for our combinatorial definition of the size (Def. 1.7, Def. 1.8).

Now one could try to set the variables in such a way that this new size corresponds to size1 (Def. 1.7). This would yield $a = 1$, $b = 0$, $c = 1$ and $d = 1$. Unfortunately, by the comments stated above, we are not allowed to set $b = 0$. This would only be possible if we just took closed lambda-terms into account. However, if we did so, we would lose the advantage that a variable with many successors has a bigger size, which therefore would not

lead to the desired shift of the lambda-terms and still no counting would be possible.

Remark 1.12
For $a = b = c = d = 1$ we get the so-called natural size, which has been discussed in [6].

Chapter 2

Methods

Now we want to introduce some tools and methods that will be used within this thesis. Our goal is to count combinatorial objects of a certain size n and estimate their asymptotic behavior as n tends to infinity. In order to do so the concepts of generating functions and singularity analysis are essential. Furthermore, we will introduce the saddle-point method, which is very useful to obtain some information on the asymptotics of rapidly growing sequences. The results presented in this chapter are strongly based on the book 'Analytic Combinatorics' of Flajolet and Sedgewick ([15]), to which we also refer the interested reader for gaining further information.

2.1 Symbolic Method and Generating Functions

The symbolic method has proved to be a very simple and efficient approach for various combinatorial enumeration problems. Yet before we can give an accurate explanation, we need to introduce some basic definitions, as they can be found in [15]:

Definition 2.1 (combinatorial class, [15, Def. I.1])
A combinatorial class $\mathcal{A}$ is a finite or denumerable set, on which a size function is defined, satisfying the following conditions:
(i) The size of an element is a non-negative integer.
(ii) The number of elements of any given size is finite.

We can see that any set of lambda-terms together with one of the size functions defined in Chapter 1 represents a combinatorial class.

We denote by A_n the number of objects in class $\mathcal{A}$ having size n. Our goal is to gain some information about the numbers A_n, either explicitly (if possible) or asymptotically for $n \to \infty$.

Two combinatorial classes $\mathcal{A}$ and $\mathcal{B}$ are called (combinatorially) isomorphic, if the sequences $(A_n)_{n\geq 0}$ and $(B_n)_{n\geq 0}$ are identical, i.e. $A_n = B_n \ \ \forall n \in \mathbb{N}_0$.

Definition 2.2 (ordinary generating function, OGF, [15, Def. I.4])
The ordinary generating function (OGF) of a combinatorial class $\mathcal{A}$ is the formal power series $A(z) = \sum_{n=0}^{\infty} A_n z^n$.

Equivalently, the OGF of the class $\mathcal{A}$ admits the representation $A(z) = \sum_{\alpha\in\mathcal{A}} z^{|\alpha|}$, where $|\alpha|$ denotes the size of α. We say that the variable z marks the size in the generating function $A(z)$.
Generating functions can also involve more than one variable, each of them marking different parameters (e.g. the number of free leaves, the number of unary-nodes, etc.). For our purposes bivariate generating functions, i.e. functions depending on two variables, will be sufficient:

$$A(z,u) = \sum_{n\geq 0}\sum_{m\geq 0} A_{n,m} z^n u^m.$$

By $[z^n]$ we denote the operation of coefficient extraction, i.e.

$$[z^n]A(z) = [z^n]\left(\sum_{n\geq 0} A_n z^n\right) = A_n,$$

which returns the number of objects of the corresponding combinatorial class $\mathcal{A}$ of size n.

Since we will work with generating functions throughout this thesis, we have to define the sum and the product of formal power series:

The sum is defined by

$$A(z) + B(z) = \sum_{n\geq 0} A_n z^n + \sum_{n\geq 0} B_n z^n = \sum_{n\geq 0} (A_n + B_n) z^n,$$

and the product by

$$A(z) \cdot B(z) = \sum_{n\geq 0} A_n z^n \cdot \sum_{n\geq 0} B_n z^n = \sum_{n\geq 0} \sum_{k=0}^{n} A_n B_{n-k} z^n.$$

Now we have all the knowledge that is necessary in order to explain the symbolic method, which will be one of our most important tools throughout this thesis.
Within the symbolic method combinatorial classes are built directly in terms of simpler classes by means of a collection of combinatorial constructions, which can easily be translated into generating functions. We will now introduce some of these elementary combinatorial constructions, that will be particularly important, as well as their translation into generating functions.

Definition 2.3 (union of disjoint sets, [15, p.23])
Let $\mathcal{A}$ and $\mathcal{B}$ be disjoint combinatorial classes with size functions $|\cdot|_{\mathcal{A}}$ and $|\cdot|_{\mathcal{B}}$ and let $\mathcal{C}$ denote the combinatorial class which satisfies $\mathcal{C} = \mathcal{A} \cup \mathcal{B}$. The size function $|\cdot|_{\mathcal{C}}$ of $\mathcal{C}$ is given as follows: For $c \in \mathcal{C}$

$$|c|_{\mathcal{C}} = \begin{cases} |c|_{\mathcal{A}} & \text{for } c \in \mathcal{A} \\ |c|_{\mathcal{B}} & \text{for } c \in \mathcal{B} \end{cases}$$

Therefore we get $C_n = A_n + B_n$, which yields $C(z) = A(z) + B(z)$ for the genrating function of the combinatorial class $\mathcal{C}$.

Thus, the union of disjoint sets translates to the sum of the generating functions of the combinatorial classes. Note that the construction of the combinatorial sum $\mathcal{A} + \mathcal{B}$ is essentially the same as the disjoint union, with the difference that the condition of disjointness can be omitted, since the elements in $\mathcal{B}$ which are already in $\mathcal{A}$ can be regarded as copies of themselves, hence making $\mathcal{A}$ and $\mathcal{B}$ disjoint (see [15, p.25]).
Another important construction is the cartesian product:

Definition 2.4 (cartesian product, [15, p.23])
Let $\mathcal{A}$ and $\mathcal{B}$ be combinatorial classes with size functions $|\cdot|_{\mathcal{A}}$ and $|\cdot|_{\mathcal{B}}$ and let $\mathcal{C}$ denote the the combinatorial class which satisfies $\mathcal{C} = \mathcal{A} \times \mathcal{B} := \{c = (a, b) : a \in \mathcal{A}, b \in \mathcal{B}\}$. The size function $|\cdot|_{\mathcal{C}}$ of $\mathcal{C}$ is given as follows: For $c = (a, b) \in \mathcal{C}$

$$|c|_{\mathcal{C}} = |a|_{\mathcal{A}} + |b|_{\mathcal{B}}.$$

Considering all possibilities we immediately get $C_n = \sum_{k=0}^{n} A_k B_{n-k}$, which by the definition of the product of two power series yields $C(z) = A(z) \cdot B(z)$.

We see that the construction of the cartesian product translates to the product of the generating functions of the combinatorial classes.

Two combinatorial structures that are needed frequently are the neutral set and the atomic set. The neutral set $\mathcal{E}$ consists of one element of size 0, while the atomic set $\mathcal{Z}$ contains just one element of size 1. The respecting generating functions are $E(z) = 1$ and $Z(z) = z$.

A summary of the most important constructions, that will be needed at later points in this thesis, and their translations into generating functions is given in the following table:

Combinatorial Construction		Generating Function
Neutral Set	$\mathcal{E} = \{\epsilon\}$	$E(z) = 1$
Atomic Set	$\mathcal{Z} = \{a\}$	$Z(z) = z$
Disjoint Union/ Sum	$\mathcal{C} = \mathcal{A} \cup \mathcal{B} = \mathcal{A} + \mathcal{B}$	$C(z) = A(z) + B(z)$
Cartesian Product	$\mathcal{C} = \mathcal{A} \times \mathcal{B}$	$C(z) = A(z) \cdot B(z)$

Now we will demonstrate this concept by means of counting Motzkin trees, which are ordinary unary-binary trees, whose size is defined in Def. 1.7, i.e. the size equals to the total number of nodes.

Example 2.5 (Motzkin trees)
Let $M(z) = \sum_{n=0}^{\infty} M_n z^n$ be the generating function of the class $\mathcal{M}$ of Motzkin trees, i.e. M_n denotes the number of Motzkin trees of size n.
The class of Motzkin trees can be formally specified by

$$\mathcal{M} = \mathcal{Z} + (\mathcal{Z} \times \mathcal{M}) + (\mathcal{Z} \times \mathcal{M} \times \mathcal{M}).$$

This specification results from the structure of unary-binary trees: They can either consist of a single node (which is represented by the first summand), or they are made up of a node that has either one or two children (cf. the second and the third summand).

Translating into generating functions yields

$$M(z) = z + zM(z) + zM^2(z). \tag{2.1}$$

Thus, we have a quadratic equation in $M(z)$, which has the solutions

$$M(z)_{1,2} = \frac{1 - z \pm \sqrt{1 - 2z - 3z^2}}{2z}.$$

By (2.1) it follows that $M(0) = M_0 = 0$, which implies

$$M(z) = \frac{1 - z - \sqrt{1 - 2z - 3z^2}}{2z}.$$

■

In the next section we will turn to the analysis of the coefficients of generating functions in order to gain asymptotic information on the number of structures of a certain size.

2.2 Singularity Analysis

Singularity analysis relies on the simple principle that some special points of a generating function, called singularities, are reflected in the function's coefficients. Therefore we gain information on the number of certain objects of a given size by determining the singularities of the OGF of the combinatorial class. These interesting results can be obtained by no longer considering generating functions as formal power series, but as functions in the complex plane that are analytic around 0. We will see in the next chapter by means of the example of lambda-terms that unfortunately some sequences grow too fast, with the result that they are not analytic around 0 and therefore the concept of singularity analysis cannot be applied. We refer the reader who is not familiar with basic concepts of complex analysis to [15], since we will use some of these concepts in the sequel.

Definition 2.6 (singularity, [15, Def. IV.4])
Given a function f defined in the region interior to the simple closed curve γ, a point z_0 on the boundary (γ) of the region is a singularity, if f is not analytically continuable to z_0.

In short, singularities are points where a function f is not analytic.
The singularities which are closest to the origin, are called the dominant singularities of f, and will turn out to be particularly important. Their distance to the origin is called the radius of convergence of $f(z)$.

The general form of the coefficients of a generating function looks like $[z^n]F(z) = A^n\theta(n)$, where A denotes the exponential growth factor and $\theta(n)$ a subexponential factor, i.e. $\limsup |\theta(n)|^{\frac{1}{n}} = 1$.

In [15] Flajolet and Sedgewick introduced the following two principles:

First Principle of Coefficient Asymptotics:
The location of a function's singularities dictates the exponential growth (A^n) of its coefficients.
Second Principle of Coefficient Asymptotics:
The nature of a function's singularities determines the associate subexponential factor ($\theta(n)$).

The first principle is specified by the following theorem:

Theorem 2.7 (Exponential growth formula, [15, Thm. IV.7])
If f(z) is analytic at 0 and R is the radius of convergence, i.e. $R := sup\{r \geq 0 |$ f is analytic in $|z| < r\}$, then the coefficient $f_n = [z^n]f(z)$ satisfies

$$f_n \sim \left(\frac{1}{R}\right)^n.$$

For functions with non-negative coefficients, including all combinatorial generating functions, one can also adopt

$$R := sup\{r \geq 0 |\ f \text{ is analytic at all points of } \ 0 \leq z < r\}$$

Thus, the exponential factor can easily be determined by determining the radius of convergence. In order to derive the subexponential factor, we have to distinguish between certain kinds of functions:
For rational and, more generally, for meromorphic functions, which have only polar singularities, the subexponential factor $\theta(n)$ is of polynomial growth.

Theorem 2.8 (Expansion of meromorphic functions, [15, Thm. IV.10])
Let $f(z)$ be a function meromorphic at all points of the closed disc $|z| \leq R$, with poles at points $\alpha_1, \alpha_2, \ldots, \alpha_m$. Assume that $f(z)$ is analytic at all points of $|z| = R$ and at $z = 0$. Then there exist m polynomials $\{\prod_j(x)\}_{j=1}^m$ such that

$$f_n \equiv [z^n]f(z) = \sum_{j=1}^{m} \prod_j(n)\alpha_j^{-n} + O(R^{-n}).$$

Furthermore the degree of $\prod_j$ is equal to the order of the pole of f at α_j minus one.

Now we consider functions whose singularities are of richer nature than poles. For this purpose we need the following considerations:
If $f(z)$ is singular at $z = \zeta$, then $g(z) \equiv f(z\zeta)$ satisfies

$$[z^n]f(z) = \zeta^{-n}[z^n]f(z\zeta) = \zeta^{-n}[z^n]g(z),$$

where $g(z)$ now has a singularity at $z = 1$. Consequently, in the following we will w.l.o.g. only discuss functions that are singular at $z = 1$ in order to determine the subexponential factor.

Our goal is to translate an expansion of a function near its singularity into an asymptotic approximation of its coefficients. Suppose we have a function $f(z)$ that is singular at $z = 1$ and has an asymptotic expansion of the form $f(z) = \sigma(z) + O(\tau(z))$ near $z = 1$, where $\tau(z) = o(\sigma(z))$ as $z \to 1$, with σ and τ being of the form $(1 - z)^{-\alpha}$ for $\alpha \in \mathbb{C} \setminus \mathbb{Z}_{\leq 0}$, which is called standard scale of functions. Then we get

$$f_n \equiv [z^n]f(z) = [z^n]\sigma(z) + [z^n]O(\tau(z)). \tag{2.2}$$

For functions of the standard function scale, we have the following result:

Theorem 2.9 ([15, Thm. VI.1])
Let $\alpha \in \mathbb{C} \setminus \mathbb{Z}_{\leq 0}$. Then

$$[z^n](1 - z)^{-\alpha} \sim \frac{n^{\alpha-1}}{\Gamma(\alpha)} \left(1 + \sum_{k=1}^{\infty} \frac{e_k}{n^k}\right),$$

where e_k is a polynomial in α of degree $2k$.

Thus, the term $[z^n]\sigma(z)$ in Equation (2.2) can easily be described asymptotically. What is left to do is to extract the coefficients of the error terms in the expansion of the function near the singularity, which can be achieved by the so-called transfer theorems, which guarantee that $[z^n]O(\tau(z)) = O([z^n]\tau(z))$. The basic condition for this transfer is analyticity in a certain region, called Δ-domain.

Definition 2.10 (Δ-domain, [15, Def. VI.1])
Given two numbers ϕ, R with $R > 1$ and $0 < \phi < \frac{\pi}{2}$, the open domain $\Delta(\phi, R)$, defined as

$$\Delta(\phi, R) = \{z \mid |z| < R, z \neq 1, |arg(z - 1)| > \phi\},$$

is called a Δ-domain.

Now we can formulate the transfer theorems:

Theorem 2.11 (Transfer-theorems, [15, Thm. VI.3, Cor. VI.1])
Let α be an arbitrary complex number in $\mathbb{C}\backslash\mathbb{Z}_{\leq 0}$ and suppose $f(z)$ is analytic in a Δ-domain Δ with an isolated singularity at $z = 1$. Then we get for $z \in \Delta$, $z \to 1$
Big-Oh transfer:

$$O((1-z)^{-\alpha}) \to f_n = O(n^{\alpha-1})$$

Little-Oh transfer:

$$o((1-z)^{-\alpha}) \to f_n = o(n^{\alpha-1})$$

Sim-transfer:

$$f(z) \sim (1-z)^{-\alpha} \to f_n \sim \frac{n^{\alpha-1}}{\Gamma(\alpha)}$$

Remark 2.12
We will call the number $-\alpha$ the type of the singularity.

Now we have all the tools to determine an asymptotic behavior of a function's coefficients by performing a term-by-term transfer from the asymptotic expansion of the function near its singularity.
We will explain these concepts by the example of Motzkin trees.

Example 2.13 (Motzkin trees, continuation)
From Example 2.5 we know that the generating function $M(z)$ of Motzkin trees has the representation

$$M(z) = \frac{1 - z - \sqrt{(1+z)(1-3z)}}{2z}.$$

Therefore it is singular at $z = -1$ and $z = 1/3$, with the dominant singularity being $z = 1/3$.

The important condition that $M(z)$ is analytic in a Δ-domain is fulfilled. Around the point $1/3$, a singular expansion is obtained by multiplying $\sqrt{1-3z}$ and the analytic expansion of the factor $(1+z)^{1/2}/(2z)$, which yields

$$M(z) = 1 - 3^{1/2}\sqrt{1-3z} + O((1-3z)^{3/2}).$$

Applying the term-by-term transfer, and using $\Gamma(-\frac{1}{2}) = -2\sqrt{\pi}$, we get

$$M_n = \sqrt{\frac{3}{4\pi n^3}}3^n + O(3^n n^{-2}).$$

■

2.3 Saddle-point method

The saddle-point method, also known as method of steepest descent, is a very powerful tool for obtaining asymptotic information about rapidly growing functions. It is typically applied to entire functions (that have no singularities) or functions with singularities that exhibit an exponential growth around their singularities. There are many books that include exhaustive presentations of this approach. We will stick to the descriptions given in [15] and [26].

Our goal is to derive the asymptotic behavior of the coefficients of a generating function $G(z)$, which is assumed to be analytic in the origin and has non-negative coefficients, given in integral form by

$$[z^n]G(z) = \frac{1}{2i\pi}\int_{\mathcal{C}} G(z)\frac{dz}{z^{n+1}},$$

where $\mathcal{C}$ encircles the origin, lies within the domain where G is analytic and is positively oriented (cf. Cauchy's coefficient formula, [15, Thm. IV.4]).

The saddle-point method for Cauchy coefficient integrals is a special case of the general saddle-point method with the integrand being $F(z) = G(z)/z^{n+1}$, which we will now briefly introduce.

First of all, consider that there are three different kinds of points on the surface representing the modulus of an analytic function $F(z)$, namely ordinary points, zeros and saddle-points (cf. [15, Fig. VIII.1]). Thus, there are no local maxima (cf. maximum modulus principle).

Definition 2.14 (saddle-point, [15, p. 543])
A saddle-point is a point z_0, such that $F(z_0) \neq 0$ and $F'(z_0) = 0$.

It is easy to see that for our purposes the saddle-point $z_0 > 0$ will always be unique, since in case of enumeration problems we only have non-negative coefficients, which implies $F''(z) > 0$ on the positive real line.

In order to proceed we set $F(z) = e^{f(z)}$, where both $F(z)$ and $f(z)$ depend on a large parameter n.

The saddle-point method is based on the fact that we are allowed to shift contours of integration when estimating integrals of analytic functions.
The basic idea is to choose the contour in such a way that it passes through the single saddle-point. The saddle-point then presents a local maximum of the integrand along the path, while the rest of the contour (further away from the saddle-point) is asymptotically negligible. We get essentially:

$$\text{Saddle-point method} \;=\; \text{Choice of contour} \;+\; \text{Laplace's method}$$

The first step in applying the saddle-point method is to find the saddle-point z_0, using the saddle-point equation $F'(z_0) = 0$ or equivalently $f'(z_0) = 0$.

Then we have to perform the splitting of the integration contour $\mathcal{C} = \mathcal{C}^{(0)} \cup \mathcal{C}^{(1)}$, where $\mathcal{C}^{(0)}$, called the "central part", contains z_0 (or passes sufficiently close to it), and $\mathcal{C}^{(1)}$ is formed of the remaining "tails". This splitting has to be carried out independently for each application of the saddle-point method taking into account the growth of the integrand and thereby has to fulfil the following conditions:

- The contribution of the tails has to be asymptotically negligible (cf. [15, p.552], condition $\mathbf{SP_1}$):
$$\int_{\mathcal{C}^{(1)}} F(z)dz = o\left(\int_{\mathcal{C}} F(z)dz\right).$$
- In the central region the quantity of $f(z)$ in the integrand has to be asymptotically well approximated by a quadratic function. (cf. [15, p.552], condition $\mathbf{SP_2}$)
- The tails can be completed back, introducing only negligible error terms, meaning that the incomplete Gaussian integral is asymptotically equivalent to a complete one. (cf. [15, p.553], condition $\mathbf{SP_3}$)

If all these assumptions are fulfilled, we get the following saddle-point approximation:

Theorem 2.15 (Saddle-point approximation, [15, Thm. VIII.3])
If the contour integral of a function $F(z) = e^{f(z)}$ along the path $\mathcal{C}$ fulfils the conditions listed above, we get
$$\frac{1}{2i\pi}\int_{\mathcal{C}} e^{f(z)}dz \sim \pm\frac{e^{f(z_0)}}{\sqrt{2\pi f''(z_0)}},$$
where z_0 is the saddle-point of $F(z)$, and the $\pm$ reflects the orientation of the contour.

In case of Cauchy coefficient integrals, we often make use of polar coordinates and set $z = re^{i\theta}$, where the radius r of the circle is chosen equal to (or close to) the positive saddle-point value.

Thus, we need to estimate

$$[z^n]G(z) = \frac{r^{-n}}{2\pi} \int_{-\pi}^{\pi} G(re^{i\theta})e^{-ni\theta} d\theta.$$

The saddle-point equation can now be written as

$$a(r) := r\frac{G'(r)}{G(r)} = n. \tag{2.3}$$

In order for the saddle-point method to be applicable, we would need to check all the conditions introduced above. Fortunately, for Cauchy coefficient integrals it is possible to specify a certain set of functions that fulfil the necessary conditions, namely Hayman-admissible functions.

Definition 2.16 (Hayman-admissibilty, [15, Def. VIII.1])
Let $G(z)$ have radius of convergence ρ with $0 < \rho \leq +\infty$ and be always positive on some subinterval (R_0,ρ) of $(0,\rho)$. The function $G(z)$ is Hayman-admissible, if it satisfies the following conditions, with $a(r)$ as defined in (2.3) and $b(r) := r\frac{G'(r)}{G(r)} + r^2\frac{G''(r)}{G(r)} - r^2\left(\frac{G'(r)}{G(r)}\right)^2$:

$\mathbf{H_1}$. *Capture condition:* $\lim_{r\to\rho} a(r) = +\infty$ *and* $\lim_{r\to\rho} b(r) = +\infty$

$\mathbf{H_2}$. *Locality condition: For some function $\theta_0(r)$ defined over (R_0, ρ) and satisfying $0 < \theta_0 < \pi$, one has*

$$G(re^{i\theta}) \sim G(r)e^{i\theta a(r) - \theta^2 b(r)/2} \qquad \text{as } r \to \rho,$$

uniformly in $|\theta| \leq \theta_0(r)$.

$\mathbf{H_3}$. *Decay condition: Uniformly in* $\theta_0(r) \leq |\theta| < \pi$

$$G(re^{i\theta}) = o\left(\frac{G(r)}{\sqrt{b(r)}}\right).$$

The big advantage of Hayman-admissible functions is that they fulfil a variety of closure properties. So, if we can express our function of interest in terms of other functions that we know are Hayman-admissible, we do not have to check the necessary conditions, which spares a lot of work.
For a Hayman-admissible function $G(z)$ we get the saddle-point approximation

$$[z^n]G(z) \sim \frac{G(r)}{r^n\sqrt{2\pi b(r)}} \qquad \text{as} \quad n \to \infty,$$

where r is the saddle-point and $b(r) := r\frac{G'(r)}{G(r)} + r^2\frac{G''(r)}{G(r)} - r^2\left(\frac{G'(r)}{G(r)}\right)^2 = r\left(r\frac{G'(r)}{G(r)}\right)'$.

2.4 Working process

All the concepts that have been introduced in this chapter allow us to handle a great variety of combinatorial problems. We will now give a short summary of the working process underlining the different application fields of the established approaches.

- First we set up formal specifications for the combinatorial classes of interest.
- The symbolic method transforms these specifications into equations that define generating functions.

- By treating generating functions as analytic objects in the complex plane, we are able to give estimations of their coefficients. For any of the following methods to work, we therefore need analyticity around 0.
  - For meromorphic functions, whose singularities are just poles, we can use Thm. 2.8.
  - For functions with essential singularities (not poles) that exhibit moderate (polynomial) growth around the singularity, we can apply singularity analysis.
  - Functions without any singularities or with singularities at which the function exhibits exponential growth can often be treated by the saddle-point method.

Chapter 3

Counting lambda-terms

In this chapter the problem of counting lambda-terms is introduced and it is explained why the common methods of analytic combinatorics, introduced in Chapter 2, fail for this purpose. Moreover we will present the concept of contexts (Section 3.2) and discuss the asymptotic behavior of the number of closed lambda-terms (Section 3.3). In Section 3.4 we present the polynomials $P_n(m)$ associated with the number of lambda-terms of size n with at most m distinct free variables and in the last section we give lower and upper bounds for the number of closed lambda-terms of a certain size. As mentioned in Chapter 1, there are different ways how the size of a lambda-term can be defined. Since the results presented in this thesis are based on different sources, we will sometimes have to switch between the two common sizes introduced in the first chapter (Def. 1.7, Def. 1.8). Mostly we will use one definition consistently throughout a whole chapter or section, which will be stated at the respective beginning. Throughout Chapter 3 the second definition for the size of a lambda term (Def. 1.8) is used; i.e. the size corresponds to the number of inner nodes.
Most of the results given in this chapter are based on [19], where further information can be found.

3.1 Counting lambda-terms with at most m distinct free variables

Our goal in this section is to set up a recursion for the number of lambda-terms with a bounded number of free variables.
In Chapter 3.1 - 3.4 we only consider lambda-terms modulo α-equivalence, i.e. now the terms $\lambda x.yx$ and $\lambda x.zx$ are regarded to be different. Let $M = \{x_1, \ldots, x_m\}$. Then $\mathcal{T}_{n,m}$ denotes the set of lambda-terms of size n with the property that its set of free variables is a subset of M. Let $T_{n,m}$ denote the cardinality of the set $\mathcal{T}_{n,m}$.
We get the following equations:

$$T_{0,m} = m, \tag{3.1}$$

$$T_{n+1,m} = T_{n,m+1} + \sum_{i=0}^{n} T_{i,m} T_{n-i,m} \tag{3.2}$$

The first summand of the recursion corresponds to the number of lambda-terms of size $n+1$ starting with an abstraction. Such a lambda-term results from a lambda-term of size n by adding a lambda that becomes the new root of the corresponding lambda-tree. The term of size n, out of which the new lambda-term results, can have at most one free variable more, in case this variable gets bound by the new lambda. The second summand derives from the lambda-terms of size $n+1$ starting with an application. Obviously such a term arises from two smaller lambda-terms with at most m distinct free variables whose sizes add to n, because the application node, which connects those two terms, has itself size one.

We get an induction, that relies on two variables (n and m), with the distinct feature that one is decreasing (the size n), and the other one is increasing (the maximum number of free variables m). It is this growth of m, why treatments by generating functions and classical analytic combinatorics fail.

Special interest lies in the number of closed lambda-terms of size n, which obviously corresponds to $T_{n,0}$.

The first values of $T_{n,0}$ are:
0, 1, 3, 14, 82, 579, 4741, 43977, 454283, 5159441, 63782411, ...

Remark 3.1
In case of variable size 1, the sequence $T_{n,0}$ looks like:
0, 1, 2, 4, 13, 42, 139, 506, 1915, 7558, 31092, 132170, 580466, 2624545, ...

Although the problem of determining the asymptotic number of closed lambda-terms is still unsolved, we know that it is superexponential in n (i.e. it grows faster than a^n for any a), but it is asymptotically smaller than n^n.

Before we can give an explanation for this behavior, we have to introduce the concept of so-called i-contexts (see [19]).

3.2 Counting contexts

Contexts are lambda-terms where the free variables are replaced by so-called holes, which are considered to be distinct and are basically the same as free variables. The important difference is that a variable can occur multiple times in a term, while each hole occurs only once and can be filled (substituted) independently of the other holes in the term. An i-context is a lambda-term with exactly i holes. Obviously 0-contexts correspond to closed lambda terms.

Let $c_{n,i}$ denote the number of i-contexts of size n.

Then we get

$$c_{0,1} = 1,$$

$$c_{0,i} = 0, \qquad \text{for } i \neq 1.$$

Considering how an i-context of size $n+1$ can be constructed from smaller ones yields the following recursion ([19, p.9]):

$$c_{n+1,i} = \sum_{j=i}^{n+1} \binom{j}{i} c_{n,j} + \sum_{j=0}^{i} \sum_{k=0}^{n} c_{k,j} c_{n-k,i-j}$$

In analogy to the recursion for the numbers $T_{n,m}$ (Equ. (3.2)) we distinguish between the two cases that the term starts either with an abstraction or an application.
The first summand corresponds to the i-contexts of size $n+1$ starting with an abstraction: In order to derive such a term we add a new unary root to a j-context ($j \in \{i, \ldots, n+1\}$) of size n and choose a set of $j-i$ holes among the j holes which we substitute by occurrences of the variable that is bound by the new lambda. There are $\binom{j}{i} c_{n,j}$ such i-contexts for every fixed j. Summing up yields the number of all i-contexts of size $n+1$ starting with an abstraction.
The second summand gives the number of i-contexts of size $n+1$ starting with an application. Such a term results by combining a j-context of size k and a $(i-j)$-context of size $n-k$ ($j \in \{0, \ldots, i\}$, $k \in \{0, \ldots, n\}$) with a newly added binary node. Again summing up yields the desired number.

Due to the fact that 0-contexts correspond to closed lambda-terms, we are now ready to give some results on the asymptotic behavior of closed lambda-terms.

3.3 The asymptotic behavior of closed lambda-terms

First we introduce the bivariate generating function $L(z,u) = \sum_{n,i\geq 0} c_{n,i} z^n u^i$, where z marks the size and u the number of holes. Remember that the size of a lambda-term still corresponds to the number of its inner nodes (Def. 1.8).

Lemma 3.2 ([19, Prop. 3])
Let $c_{n,i}$ denote the number of i-contexts of size n. The bivariate generating function $L(z,u) = \sum_{n,i\geq 0} c_{n,i} z^n u^i$ satisfies the following equation:

$$L(z,u) = u + zL(z,u+1) + zL(z,u)^2.$$

Proof. To prove this functional equation we use the symbolic method, as it has been done in [7]. Therefore we first have to introduce the following atomic classes: the class of application nodes $\mathcal{A}$, the class of abstraction nodes $\mathcal{U}$, the class of bound leaves $\mathcal{D}$ and the class of free leaves $\mathcal{F}$ [1].
Then the class $\mathcal{L}$ of equivalence classes of lambda-terms can be specified by

$$\mathcal{L} = \mathcal{F} + (\mathcal{A} \times \mathcal{L}^2) + (\mathcal{U} \times subs(\mathcal{F} \to \mathcal{F} + \mathcal{D}, \mathcal{L})),$$

where the substitution operator $subs(\mathcal{F} \to \mathcal{F}+\mathcal{D}, \mathcal{L})$ corresponds to replacing arbitrarily many free leaves in $\mathcal{L}$ by bounded ones.
This specification gives rise to the above functional equation for $L(z,u) = \sum_{n,i\geq 0} c_{n,i} z^n u^i$, where z counts the size and u the number of free leaves, or the number of holes respectively.

□

[1] Actually the class $\mathcal{F}$ is the class of holes, because we consider the free leaves to be anonymous.

The equation

$$L(z,u) = u + zL(z,u+1) + zL(z,u)^2$$

has the following solution

$$L(z,u) = \frac{1 - \sqrt{1 - 4z(u + zL(z,u+1))}}{2z}.$$

Proceeding equivalently as in [19] we state $M(z,u) = 2zL(z,u)$.

Then $M(z,u) = 1 - \sqrt{1 - 4zu - 2zM(z,u+1)}$ and hence $M(z,0) =$

$$1-\sqrt{1 - 2z(1 - \sqrt{1 - 4z - 2z(1 - \sqrt{1 - 8z - 2z(1 - \sqrt{1 - 12z - 2z(1 - \sqrt{1 - 16z - ..})})})}}.$$

As stated in [7], [19] and [10], the singularities of $M(z,u)$ are the singularities of $M(z,u+1)$ and $z_u > 0$ with $1 - 4z_u u - 2z_u M(z_u, u+1) = 0$, which yields $z_u < \frac{1}{4u}$ (since $z_u M(z_u, u+1) > 0$). Therefore $L(z,0)$ has a sequence of singularities $(z_u)_{u\in\mathbb{N}}$ which tends to 0, which implies that the radius of convergence is 0. Remembering Theorem 2.7 this implies that the coefficients of the generating function $L(z,0)$ grow faster than a^n for any a $\in \mathbb{R}$. As stated before, such a behavior is called superexponential growth.

This rapid growth is caused by the various possibilities of connecting the unary nodes with certain leaves. If we cancelled all those pointers we would get ordinary unary-binary trees, which are counted by the large Schröder numbers. These are asymptotically equivalent to $\left(\frac{1}{3-2\sqrt{2}}\right)^n \frac{1}{\sqrt{\pi} n^{3/2}}$ ([13]). This gives of course a first lower bound for the $T_{n,0}$, but we will derive better bounds in Section 3.5.

Remark 3.3

In case of defining the size of a lambda-tree as the number of its internal and external nodes (Def. 1.7), i.e. the size of variables is 1, the well-known

Motzkin-numbers equal the number of unary-binary trees. In Chapter 2 (Example 2.13) we derived that the number of Motzkin-trees is asymptotically equivalent to $\sqrt{\frac{3}{4\pi n^3}}3^n$.

Note that the generating function of open lambda-terms can be derived from the generating function of closed lambda-terms by $L(z,1) - L(z,0) = \frac{(1-z)L(z,0)-zL(z,0)^2}{z}$. Thus, the problems of counting open or closed lambda-terms are essentially the same and by solving one of them we immediately get the solution for the other one. Therefore in the following we will only focus on closed lambda-terms.

3.4 The polynomials P_n

In the last section it has been shown why it is very difficult to handle unrestricted lambda-terms. Now we will present an interesting approach, introduced in [19], which associates the numbers $T_{n,m}$ (cf. Section 3.1) with the following polynomials $P_n(m)$ in m:

$$P_0(m) = m,$$

$$P_{n+1}(m) = P_n(m+1) + \sum_{i=0}^{n} P_i(m)P_{n-i}(m).$$

Obviously the sequence $(P_n(0))_{n\geq 0}$ corresponds to the sequence $(T_{n,0})_{n\geq 0}$. Therefore the constant coefficient of a polynomial $P_n(m)$ is exactly equal to the number of closed lambda terms of size n.

Lemma 3.4 ([19, Lem. 1])
For every n, the degree of the polynomial P_n is equal to $n+1$.

Proof. This can easily be seen by induction on n. □

Thus, every polynomial P_n has the form

$$P_n(m) = p_n^{[1]} m^{n+1} + p_n^{[2]} m^n + \ldots + p_n^{[i]} m^{n+2-i} + \ldots + p_n^{[n+1]} m + p_n^{[n+2]}.$$

Lemma 3.5 ([19, Lem. 2])
For every $n \geq 0$ and $i \geq 0$,

$$p_0^{[1]} = 1, \quad p_0^{[i]} = 0 \quad \text{for } i > 1, \tag{3.3}$$

$$p_{n+1}^{[i]} = \sum_{j=0}^{i-2} \binom{n+1-j}{i-2-j} p_n^{[j+1]} + \sum_{k=1}^{i} \sum_{j=0}^{n} p_j^{[k]} p_{n-j}^{[i+1-k]}. \tag{3.4}$$

Proof. The Equations (3.3) are valid, because it holds that $P_0(m) = m$. To derive Equation (3.4) we have to consider that the i^{th} leading coefficient in the polynomial $P_{n+1}(m+1)$ can be derived by summing the coefficients of m^{n+3-i} in the polynomials $P_n(m+1)$ and $\sum_{j=0}^{n} P_j(m) P_{n-j}(m)$.

- Since $P_n(m+1) = p_n^{[1]}(m+1)^{n+1} + \ldots + p_n^{[i-1]}(m+1)^{n+3-i} + \ldots + p_n^{[n+2]}$ the coefficient of m^{n+3-i} is

  $$\binom{n+1}{i-2} p_n^{[1]} + \binom{n}{i-3} p_n^{[2]} + \ldots + \binom{n+3-i}{0} p_n^{[i-1]} = \sum_{j=0}^{i-2} \binom{n+1-j}{i-2-j} p_n^{[j+1]}.$$

- The j-th summand of $\sum_{j=0}^{n} P_j(m) P_{n-j}(m)$, is equal to

  $$\left(p_j^{[1]} m^{j+1} + \ldots + p_j^{[k]} m^{j+2-k} + \ldots + p_j^{[j+2]} \right)$$
  $$\cdot \left(p_{n-j}^{[1]} m^{n-j+1} + \ldots + p_{n-j}^{[i+1-k]} m^{n-j+1+k-i} + \ldots + p_{n-j}^{[n-j+2]} \right).$$

  Thus, the coefficient of m^{n+3-i} in $\sum_{j=0}^{n} P_j(m) P_{n-j}(m)$ is given by

  $$\sum_{k=1}^{i} \sum_{j=0}^{n} p_j^{[k]} p_{n-j}^{[i+1-k]}.$$

□

n	P_n
0	m
1	$m^2 + m + 1$
2	$2m^3 + 3m^2 + 5m + 3$
3	$5m^4 + 10m^3 + 22m^2 + 25m + 14$
4	$14m^5 + 35m^4 + 94m^3 + 154m^2 + 163m + 82$
5	$42m^6 + 126m^5 + 396m^4 + 838m^3 + 1277m^2 + 1235m + 579$
6	$132m^7 + 462m^6 + 1654m^5 + 4260m^4 + 8384m^3 + 11791m^2 + 10707m + 4741$
7	$429m^8 + 1716m^7 + 6868m^6 + 20742m^5 + 49720m^4 + 90896m^3 + 120628m^2 + 104055m + 43977$
8	$1430m^9 + 6435m^8 + 28396m^7 + 98028m^6 + 275886m^5 + 617096m^4 + 1068328m^3 + 1352268m^2 + 1117955m + 454283$

Table 3.1: The first nine polynomials P_n ([19, Fig. 2]).

A combinatorial interpretation of the coefficients $p_n^{[i]}$ can be given, considering how a lambda-term can be constructed from contexts. Assume we have an i-context of size n, which per definition has i holes and no free variables. By filling the i holes with free variables taken among m ones, we build a term of size n with exactly i occurences of free variables. There are $c_{n,i}m^i$ such terms. Thus, for the number of lambda terms of size n with at most m distinct free variables, we get

$$T_{n,m} = c_{n,n+1}m^{n+1} + \ldots + c_{n,i}m^i + \ldots + c_{n,0},$$

which is the polynomial $P_n(m)$. It follows that $c_{n,n+2-i} = p_n^{[i]}$, which implies that the coefficients of the polynomials P_n count the i-contexts of size n.

Remark 3.6

Using variable size 1 and therefore a slightly different definition of the polynomials P_n Lescanne provides explicit formulas for the leading coefficients, the second leading coefficients and in case of the odd polynomials also for the third leading coefficients. The interested reader is referred to [21] for more details.

3.5 Lower and upper bounds

In this section we derive lower and upper bounds for the number of closed lambda-terms of size n. One lower bound has already been mentioned, namely the 'Large Schröder numbers', which count ordinary unary-binary trees (for leaf size 0). Now we will give a more precise one, as it can be found in [13].
Let λ_n denote the number of closed lambda-terms of size n (where the size still is defined as in Def. 1.8, i.e. variables have size 0).

Theorem 3.7 ([13, Thm. 5.1])
For any $\epsilon \in (0, 4)$ we have

$$\lambda_n \gtrsim \left(\frac{(4-\epsilon)n}{\ln(n)}\right)^{n-\frac{n}{\ln(n)}}$$

Proof. A lower bound for λ_n is given by the number $\mathrm{LB(n,k)}$ of lambda-terms of size n that start with k head-lambdas and have no other lambda below. Considering that the lower part of the term is a binary tree with $n-k$ inner nodes and therefore $n-k+1$ leaves, which can be bound by k lambdas, we have $\mathrm{LB(n,k)} = \mathrm{C(n-k)k^{n-k+1}}$, where $C(i)$ denotes the i-th Catalan number. For $k = \left\lceil \frac{n}{\ln(n)} \right\rceil$ this yields

$$\lambda_n \geq C\left(n - \left\lceil \frac{n}{\ln(n)} \right\rceil\right)\left(\left\lceil \frac{n}{\ln(n)} \right\rceil\right)^{n-\left\lceil \frac{n}{\ln(n)} \right\rceil+1}.$$

With the well-known asymptotics of the Catalan numbers $C(n) \sim \frac{4^n}{n^{3/2}\sqrt{\pi}}$ (see for example [15]) it follows:

$$
\begin{aligned}
\lambda_n &\gtrsim \frac{4^{n-\left\lceil \frac{n}{\ln(n)} \right\rceil}}{\left(n - \left\lceil \frac{n}{\ln(n)} \right\rceil\right)^{3/2} \sqrt{\pi}} \left(\left\lceil \frac{n}{\ln(n)} \right\rceil \right)^{n-\left\lceil \frac{n}{\ln(n)} \right\rceil + 1} \\
&\gtrsim \left(\frac{4n}{\ln(n)} \right)^{n-\frac{n}{\ln(n)}} \frac{1}{p(n)} \qquad \text{for some positive polynomial p} \\
&\gtrsim \left(\frac{(4-\epsilon)n}{\ln(n)} \right)^{n-\frac{n}{\ln(n)}} \qquad \text{since } \left(\frac{4}{4-\epsilon} \right)^{n-\frac{n}{\ln(n)}} \gtrsim p(n).
\end{aligned}
$$

□

With similar arguments David et al. also exhibited an upper bound for λ_n.

Theorem 3.8 ([13, Thm. 5.6])
For any $\epsilon > 0$ we have

$$
\lambda_n \lesssim \left(\frac{(12+\epsilon)n}{\ln(n)} \right)^{n-\frac{n}{3\ln(n)}}
$$

Sketch of proof. The idea of the proof is to show that the number λ_n of closed lambda-terms of size n is asymptotically equal to the number T_n of closed lambda-terms of size n with less than $\frac{3n}{\ln(n)}$ and more than $\frac{n}{3\ln(n)}$ unary nodes (cf. [13, Thm. 5.4]). Consequently, such a lambda-term can have at most $n - \frac{n}{3\ln(n)}$ binary nodes and therefore at most $n - \frac{n}{3\ln(n)} + 1$ leaves. The upper bound for T_n can be obtained in the following way:

- First we consider binary trees consisting of at most $n - \left\lfloor \frac{n}{3\ln(n)} \right\rfloor$ binary nodes. With the simple estimation for the Catalan numbers, $C(n+1) \geq \sum_{i=0}^{n} C(i)$, we get that the number of these binary trees does not exceed $C\left(n - \left\lfloor \frac{n}{3\ln(n)} \right\rfloor + 1\right)$.

- Then we insert at most $\frac{3n}{\ln(n)}$ unary nodes in such a tree, which can be done in less then $\binom{3n}{\left\lceil \frac{3n}{\ln(n)} \right\rceil}$ ways.

- Finally, each of the (at most) $n+1-\frac{n}{3\ln(n)}$ leaves can be bound by at most $\frac{3n}{\ln(n)}$ lambdas, resulting in a maximum of $\left(\frac{3n}{\ln(n)}\right)^{n+1-\frac{n}{3\ln(n)}}$ possible bindings.

Therefore, we get

$$\lambda_n \sim T_n \lesssim C\left(n-\left\lceil\frac{n}{3\ln(n)}\right\rceil+1\right)\left(\begin{array}{c}3n\\ \left\lfloor\frac{3n}{\ln(n)}\right\rfloor\end{array}\right)\left(\frac{3n}{\ln(n)}\right)^{n+1-\frac{n}{3\ln(n)}}.$$

Using again the asymptotic expansion of the Catalan numbers and some simplifications similar to the proof for the lower bound (Theorem 3.7) we get the desired result. □

Remark 3.9

Note that the ratio between the upper and the lower bounds that we obtained for λ_n is exponential, while λ_n itself is superexponential ([13, p.16]).

Chapter 4

Counting restricted closed lambda-terms

This chapter gives an overview over some restrictions on the structure of lambda-terms that have been imposed in order to get a simplification of the counting problem. All of the restrictions that will be introduced involve a limitation of the possible variable bindings in a lambda-term, since we know that these various possibilities cause the fast growth of lambda-terms with the consequence that the radius of convergence is 0. Thus, unlike unrestricted lambda-terms, the ones presented in this chapter can be counted by methods introduced in Chapter 2. In Section 4.1 we will prescribe the total number of unary nodes in a lambda-tree, while in Section 4.2 we restrict the nesting levels of unary nodes. At last, in Section 4.3 we consider lambda-terms where each unary node can only bind a prescribed number of leaves, namely BCI- and BCK-lambda-terms.
For the entire Chapter 4 the first definition for the size of a lambda-term (Def. 1.7) is used, i.e. variables have size 1.

4.1 Lambda-terms with prescribed number of unary nodes

As mentioned above, the first restriction on lambda-terms that we will focus on, concerns the total number of unary nodes in a term. In Subsection 4.1.1 we will prescribe the exact number of unary nodes, while in Subsection 4.1.2 we will investigate lambda-terms that have an upper bound for the total number of unary nodes. The results given in this section are based on [11].

4.1.1 Closed lambda-terms with exactly q unary nodes

Let $\mathcal{S}_q$ denote the set of lambda-terms that have exactly q unary nodes.
Our goal is derive an equation for the bivariate generating function $S_q(z, f)$ associated with $\mathcal{S}_q$, where z marks the total size and f the number of free leaves.
Obviously $\mathcal{S}_0$ contains just ordinary binary trees, which are enumerated by the well known Catalan numbers. Thus, we have

$$S_0(z, f) = \frac{1 - \sqrt{1 - 4fz^2}}{2z}.$$

Using the symbolic method with the notations from Section 3.3 (cf. proof of Lemma 3.2) we derive the following equation for $q = 1$:

$$\mathcal{S}_1 = (\mathcal{U} \times subs(\mathcal{F} \to \mathcal{F} + \mathcal{D}, \mathcal{S}_0)) + (\mathcal{A}, \mathcal{S}_0, \mathcal{S}_1) + (\mathcal{A}, \mathcal{S}_1, \mathcal{S}_0).$$

The first summand corresponds to those terms that have the unique unary node as their root, whereby each leaf can either be bound or not. The other two summands correspond to all other terms, i.e. terms where the root is a binary node and the unique unary node appears in one of the subtrees.

This yields

$$S_1(z,f) = zS_0(z,f+1) + 2zS_0(z,f)S_1(z,f),$$

which has the following solution:

$$S_1(z,f) = \frac{zS_0(z,f+1)}{1-2zS_0(z,f)} = \frac{1-\sqrt{1-4(f+1)z^2}}{2\sqrt{1-4fz^2}}.$$

For q ≥ 1 we proceed analogously and therefore we get

$$\mathcal{S}_q = (\mathcal{U} \times subs(\mathcal{F} \to \mathcal{F} + \mathcal{D}, \mathcal{S}_{q-1})) + \sum_{l=0}^{q} (\mathcal{A}, \mathcal{S}_l, \mathcal{S}_{q-l}),$$

which translates into

$$S_q(z,f) = zS_{q-1}(z,f+1) + z\sum_{l=0}^{q} S_l(z,f)S_{q-l}(z,f).$$

Solving yields the following term for $S_q(z,f)$:

$$S_q(z,f) = \frac{z}{1-2zS_0(z,f)} \left(S_{q-1}(z,f+1) + \sum_{l=1}^{q-1} S_l(z,f)S_{q-l}(z,f) \right).$$

Remember that our special interest lies in the number of closed lambda-terms. Thus, we set $f = 0$, resulting in

$$S_q(z,0) = z \left(S_{q-1}(z,1) + \sum_{l=1}^{q-1} S_l(z,0)S_{q-l}(z,0) \right).$$

In order to solve this recursion we need the following lemma.

Lemma 4.1 ([11, Lem. 5])
Let $\sigma_q(f) = \sqrt{1-4(f+q)z^2}$ for $q \geq 0$ (σ_q is actually a function in the two variables z and f, but since z plays no role for our purpose and for the proof of the lemma, we omit it). Then, for all $q \geq 0$, there exists a rational function R_q in $q+1$ variables such that

$$S_q(z,f) = -\frac{z^{q-1}\sigma_q(f)}{2\prod_{l=0}^{q-1}\sigma_l(f)} + R_q(z,\sigma_0(f),\ldots,\sigma_{q-1}(f)).$$

Moreover, the denominator of $R_q(z,\sigma_0(f),\ldots,\sigma_{q-1}(f))$ is of the form $\prod_{0\leq l<q}\sigma_l(f)^{\alpha_{l,q}}$ where the exponents $\alpha_{0,q},\ldots,\alpha_{q-1,q}$ are positive integers.

Proof. The lemma follows by induction on q. □

For the generating function of closed lambda-terms with exactly q unary nodes $S_q(z,0)$ it follows directly by setting $f=0$ ([11, Lem. 6])

$$S_q(z,0) = -\frac{z^{q-1}\sqrt{1-4qz^2}}{2\prod_{l=0}^{q-1}\sqrt{1-4lz^2}} + R_q(z,1,\sqrt{1-4z^2},\ldots,\sqrt{1-4(q-1)z^2}). \quad (4.1)$$

Its dominant singularities are $z = \pm\frac{1}{2\sqrt{q}}$.

In order to exhibit an asymptotic behavior for the coefficients of $S_q(z,0)$ we have to consider that a lambda-term with exactly q unary nodes and i leaves has $i-1$ binary nodes and size $n = q+2i-1$. Thus, we get the following result.

Theorem 4.2 ([11, Prop. 4])
The number of closed lambda-terms with exactly q unary nodes of size n is 0 for $n = q \bmod 2$. Otherwise its asymptotic value is

$$[z^n]S_q(z,0) \sim \frac{\sqrt{2}}{2^q\sqrt{(q-1)!}\sqrt{\pi n^3}}(\sqrt{q})^n, \quad \text{as } n \to \infty.$$

Proof. The term R_q from Equ. (4.1) has singularities at $z = \pm\frac{1}{2\sqrt{l}}$ for $1 \leq l < q$, while the first term has singularities of smaller type at $z = \frac{1}{2\sqrt{q}}$. Hence the first term gives the dominant contribution to the asymptotics of $[z^n]S_q(z,0)$, which yields

$$[z^n]S_q(z,0) \sim [z^{q+2i-1}]\frac{-z^{q-1}\sqrt{1-4qz^2}}{2\prod_{l=1}^{q-1}\sqrt{1-4lz^2}} \sim [z^i]\frac{-\sqrt{1-4qz}}{2\prod_{l=1}^{q-1}\sqrt{1-4lz}}, \quad \text{as } n \to \infty.$$

Since the denominator $\prod_{l=1}^{q-1}\sqrt{1-4lz}$ contributes a multiplicative factor

$$\prod_{l=1}^{q-1}\sqrt{1-\frac{l}{q}} = q^{(1-q)/2}\sqrt{(q-1)!}$$

we obtain the desired result by using singularity analysis and the fact that $\Gamma(-\frac{1}{2}) = -2\sqrt{\pi}$. □

4.1.2 Closed lambda-terms with at most q unary nodes

Let $S_{\leq q}(z,f)$ be the generating function for lambda-terms with at most q unary nodes, where again z marks the size, i.e. the total number of nodes, and f marks the number of free leaves. Obviously, it holds that $S_{\leq q}(z,f) = \sum_{l=0}^{q} S_l(z)$ and therefore the results obtained for a fixed number of unary nodes can be applied (see [11, Section 4.4]).

The dominant singularity of $S_{\leq q}(z, f)$ arises from $S_q(z, f)$, whereas the contribution of the other summands $S_l(z, f)$ with $l < q$ to the asymptotics of the coefficients of $S_{\leq q}(z, f)$ are negligible. Therefore the asymptotic behavior for the number of closed lambda-terms of a fixed or a bounded number of unary nodes is the same.

4.2 Lambda-terms with bounds on the nesting levels of unary nodes

In this section we focus on lambda-terms that are restricted by an upper bound for the number of the nesting levels of their unary nodes. This restriction can either be local, by bounding the unary length of each binding, or global, by bounding the unary height of the lambda-term.

Consider a lambda-term T and its associated enriched tree $G(T)$.

Definition 4.3 (unary length, [11, Def. 1])
The unary length of the binding of a leaf e by some abstraction λe in T is defined as the number of unary nodes on the path connecting λe and e in the underlying Motzkin tree and it is denoted by $l_u(e)$.

Definition 4.4 (unary height, [11, Def. 2])
The unary height of a vertex v of T is defined as the number of unary nodes on the path from the root to v in the underlying Motzkin tree and is denoted by $h_u(v)$.
The unary height of the lambda-term T, denoted by $h_u(T)$, is defined by $\max_{v \text{ vertex of } T} h_u(v)$.

4.2.1 Closed lambda-terms of bounded unary length of bindings

Let $\mathcal{G}_{\leq k}$ denote the class of closed lambda-terms where all bindings have unary length less than or equal to k. In order to derive an equation for the generating function $G_{\leq k}(z)$ associated with $\mathcal{G}_{\leq k}$ we have to introduce further classes of unary-binary trees, as it has been done in [11]:
We denote by $\hat{\mathcal{P}}^{(i,k)}$ the class of unary-binary trees such that every leaf can be labeled in $\min\{h_u + i, k\}$ ways. With $\mathcal{Z}$ being the class of atoms, and $\mathcal{A}$ and $\mathcal{U}$ denoting again the class of application nodes, and unary nodes respectively, we get the following specification for the classes $\hat{\mathcal{P}}^{(i,k)}$ by using the symbolic method:

$$\hat{\mathcal{P}}^{(k,k)} = k\mathcal{Z} + (\mathcal{A} \times \hat{\mathcal{P}}^{(k,k)} \times \hat{\mathcal{P}}^{(k,k)}) + (\mathcal{U} \times \hat{\mathcal{P}}^{(k,k)})$$

and

$$\hat{\mathcal{P}}^{(i,k)} = i\mathcal{Z} + (\mathcal{A} \times \hat{\mathcal{P}}^{(i,k)} \times \hat{\mathcal{P}}^{(i,k)}) + (\mathcal{U} \times \hat{\mathcal{P}}^{(i+1,k)}), \qquad \text{for } i < k.$$

Translating into generating functions and solving for $\hat{P}^{(k,k)}(z)$ and $\hat{P}^{(i,k)}(z)$, we obtain

$$\hat{P}^{(k,k)}(z) = \frac{1 - z - \sqrt{(1-z)^2 - 4kz^2}}{2z}, \tag{4.2}$$

and

$$\hat{P}^{(i,k)}(z) = \frac{1 - \sqrt{1 - 4iz^2 - 4z^2\hat{P}^{(i+1,k)}(z)}}{2z} \qquad \text{for } i < k. \tag{4.3}$$

Clearly, the class $\hat{\mathcal{P}}^{(0,k)}$ is isomorphic to the class $\mathcal{G}_{\leq k}$ and therefore it holds that $G_{\leq k}(z) = \hat{P}^{(0,k)}(z)$.
The Equations (4.2) and (4.3) can be rewritten in the form

$$\hat{P}^{(i,k)}(z) = \frac{1}{2z}\left(1 - \mathbf{1}_{[i=k]}z - \sqrt{\hat{R}_{k-i+1,k}(z)}\right),$$

where

$$\hat{R}_{1,k}(z) = (1-z)^2 - 4kz^2,$$

$$\hat{R}_{2,k}(z) = 1 - 4(k-1)z^2 - 2z + 2z^2 + 2z\sqrt{\hat{R}_{1,k}},$$

and

$$\hat{R}_{i,k}(z) = 1 - 4(k-i+1)z^2 - 2z + 2z\sqrt{\hat{R}_{i-1,k}(z)}, \tag{4.4}$$

for $3 \leq i \leq k+1$.

Hence, we get $G_{\leq k}(z) = \frac{1-\sqrt{\hat{R}_{k+1,k}(z)}}{2z}$, which involves $k+1$ nested radicals.

As described in Chapter 2, we now have to determine the location and type of the dominant singularity of $G_{\leq k}(z)$ in order to determine the asymptotics of its counting sequence. Since the generating function is built of nested radicals the singularities are the values where at least one of the radicands vanishes.

Lemma 4.5 ([11, Lem. 9])
Let $\hat{\rho}_k$ be the dominant singularity of the function $G_{\leq k}(z)$. Then $\hat{\rho}_k = \frac{1}{1+2\sqrt{k}}$ comes from the innermost radicand and is of type $\frac{1}{2}$.

Sketch of proof. First of all, note that the dominant singularities of $G_{\leq k}(z)$ and $\sqrt{\hat{R}_{k+1,k}(z)}$ are the same.
By induction on j it can easily be seen that the function $\hat{R}_{j,k}(z)$ is strictly decreasing on the positive real line (cf. [11], Lem. 7).
The next step is to prove that the dominant singularity of the radical $\sqrt{\hat{R}_{j,k}(z)}$ is unique (cf. [11], Lem. 8): Let $z_0 > 0$ denote dominant singularity of $\hat{R}_{j,k}(z)$. Assume that z_0 is a root of $\hat{R}_{j,k}(z)$. Then $2z_0\hat{P}^{(k-j+1,k)}(z_0) + \mathbf{1}_{j=k}z_0 = 1$. If there was another (complex) root $x = z_0e^{i\theta}$ of $\hat{R}_{j,k}(z)$, having the same

modulus as z_0, we would have

$$1 = 2z_0\hat{P}^{(k-j+1,k)}(z_0) + \mathbf{1}_{j=k}z_0 = \left|2z_0e^{i\theta}\hat{P}^{(k-j+1,k)}(z_0e^{i\theta}) + \mathbf{1}_{j=k}z_0e^{i\theta}\right|.$$

Regarding $\hat{P}^{(k-j+1)}(z)$ as the generating function of some appropriate class of lambda-terms implies that its coefficients are positive for all sufficiently large n. But this yields

$$\left|2z_0e^{i\theta}\hat{P}^{(k-j+1,k)}(z_0e^{i\theta}) + \mathbf{1}_{j=k}z_0e^{i\theta}\right| < 1,$$

whenever $\theta \neq 0$, which leads to a contradiction. Hence z_0 cannot be a root of $\hat{R}_{j,k}(z)$, and therefore has to be a zero of some $\hat{R}_{j-l,k}(z)$ with $0 < l < j$. But this again results in a contradiction by using the same arguments as above. Thus, we get that the dominant singularity is unique.
Now we determine the roots of the innermost radicand $\hat{R}_{1,k}(z)$, which is a quadratic equation and has two roots: $\frac{1}{1+2\sqrt{k}}$ and $\frac{1}{1-2\sqrt{k}}$. Since k is a positive integer, $\hat{\rho}_k = \frac{1}{1+2\sqrt{k}}$ is the dominant singularity of the generating function $\hat{P}^{(1,k)}(z)$ and it is of type $\frac{1}{2}$. What is now left to do, is to show that none of the radicands $\hat{R}_{j,k}$, $2 \leq j \leq k+1$, has a positive root that is smaller or equal to $\hat{\rho}_k$. This can be seen by induction on j and the fact that $\hat{R}_{j,k}(z)$ is decreasing on $\mathbb{R}_+$.

□

By means of the following proposition we are finally able to establish the asymptotic behavior of the number of lambda-terms having only bindings of bounded unary length, as it was done in [11].

Proposition 4.6 ([11, Prop. 5])
Let $\hat{\rho}_k$ be the root of the innermost radicand $\hat{R}_{1,k}(z)$. Then

$$\hat{R}_{1,k}(\hat{\rho}_k(1-\epsilon)) = 2(1-\hat{\rho}_k)\epsilon + \mathcal{O}(\epsilon^2) \tag{4.5}$$

and

$$\hat{R}_{j,k}(\hat{\rho}_k(1-\epsilon)) = c_j\hat{\rho}_k^2 + \frac{4\hat{\rho}_k^{\frac{3}{2}}k^{\frac{1}{4}}}{\sqrt{\prod_{l=2}^{j-1} c_l}}\sqrt{\epsilon} + \mathcal{O}(\epsilon), \tag{4.6}$$

for $2 \leq j \leq k+1$, *where* $c_1 = 1$ *and* $c_j = 4j - 5 + 2\sqrt{c_{j-1}}$ *for* $2 \leq j \leq k+1$.

Sketch of proof. Using the Taylor expansion of $\hat{R}_{1,k}(z)$ around $\hat{\rho}_k$ and setting $z = \hat{\rho}_k(1-\epsilon)$ we obtain Equation (4.5). The next step is to compute an expansion of $\hat{R}_{j,k}(z)$ around $\hat{\rho}_k$, where $2 \leq j \leq k+1$. Using Equation (4.5) and the recursive relation (4.4) for $\hat{R}_{j,k}(z)$, we can show Equation (4.6) for $j = 2$. The desired result is then obtained by induction on j. □

Now we are in the position to give the final result of this subsection:

Theorem 4.7 ([11, Thm. 1])
For the generating function $G_{\leq k}(z)$ *of lambda terms where all bindings have unary length not larger than* k *it holds that*

$$[z^n]G_{\leq k}(z) \sim \sqrt{\frac{2k+\sqrt{k}}{4\pi\prod_{j=2}^{k+1} c_j}}\, n^{-\frac{3}{2}}(1+2\sqrt{k})^n, \quad \text{as } n \to \infty, \tag{4.7}$$

where

$$c_1 = 1 \quad \text{and} \quad c_j = 4j - 5 + 2\sqrt{c_{j-1}}, \quad \text{for} \quad 2 \leq j \leq k+1.$$

Proof. Remember that the generating function of $G_{\leq k}(z)$ has the representation

$$G_{\leq k}(z) = \frac{1}{2z}\left(1 - \sqrt{\hat{R}_{k+1,k}(z)}\right).$$

From Lemma 4.5 we know that its dominant singularity $\hat{\rho}_k = \frac{1}{1+2\sqrt{k}}$ is algebraic and of type $\frac{1}{2}$. Thus, using singularity analysis, we get the factor $n^{-\frac{3}{2}}(1+2\sqrt{k})^n$ in Equation (4.7).
Now we consider the constant (w.r.t. n) term of Equation (4.7). From Proposition 4.6 we get

$$\hat{R}_{k+1,k}(\hat{\rho}_k(1-\epsilon)) = c_{k+1}\hat{\rho}_k^2 + \underbrace{\frac{4\hat{\rho}_k^{\frac{3}{2}}k^{\frac{1}{4}}}{\sqrt{\prod_{l=2}^{k} c_l}}}_{:=d_{k+1}}\sqrt{\epsilon} + O(\epsilon).$$

Therefore it follows that

$$G_{\leq k}(\hat{\rho}_k(1-\epsilon)) = \frac{1-\hat{\rho}_k\sqrt{c_{k+1}}}{2\hat{\rho}_k} - \frac{d_{k+1}}{4\hat{\rho}_k^2\sqrt{c_{k+1}}}\sqrt{\epsilon} + O(\epsilon),$$

which gives

$$[z^n]G_{\leq k}(z) \sim -\frac{d_{k+1}}{4\hat{\rho}_k^2\sqrt{c_{k+1}}}[z^n]\sqrt{1-\frac{z}{\hat{\rho}_k}}, \qquad \text{as } n \to \infty.$$

Setting in the formulas for d_{k+1} and $\hat{\rho}_k$ and using some basic simplifications yields the desired result. □

Remark 4.8

In [11] the authors showed that the multiplicative constant in Equation (4.7) decreases very fast for $k \to \infty$. More precisly they established the following equation (see [11, Prop. 6]):

$$\sqrt{\frac{2k+\sqrt{k}}{4\pi\prod_{j=2}^{k+1} c_j}} = \frac{1}{D2^{k+1}e^{\sqrt{k+1}}}\sqrt{\frac{(k+1)^{\frac{1}{4}}(2k+\sqrt{k})}{k!}}\left(1+O\left(\frac{1}{\sqrt{k}}\right)\right), \quad \text{as } k \to \infty,$$

where $D = \sqrt{\pi\omega e^{\frac{1}{4}-\frac{5}{4}\gamma+\zeta(\frac{1}{2})}}$, ω is a computable constant with numerical value $\omega \approx 0,118\ldots$, and γ is the Euler-Mascheroni constant .

4.2.2 Closed lambda-terms of bounded unary height

In this subsection we are going to investigate the asymptotic behavior of closed lambda-terms with bounded unary height, i.e. each branch in the corresponding lambda-tree contains a limited number of unary-nodes.
Let $\mathcal{H}_{\leq k}$ denote the class of closed lambda-terms with unary height less than or equal to k.
In order to set up an equation for $\mathcal{H}_{\leq k}$, we again have to introduce further classes of unary-binary trees, as presented in [11].
Let $\mathcal{P}^{(i,k)}$ be the class of unary-binary trees such that $i + h_u(e) \leq k$ for every leaf e (i.e. the unary height of each leaf e is at most $k - i$) and every leaf e is colored with one out of $i + h_u(e)$ colors.
Obviously, $\mathcal{P}^{(0,k)}$ is isomorphic to the class $\mathcal{H}_{\leq k}$. For general i, $\mathcal{P}^{(i,k)}$ is isomorphic to the class of closed lambda-terms, where the underlying unary-binary tree consists of a path of i unary nodes with an appended Motzkin tree beneath that has unary height less than or equal to $k - i$. Each leaf e can then be bound in $i + h_u(e)$ ways in order for the lambda-term to become closed.

Analogously to the previous subsection, the classes $\mathcal{P}^{(i,k)}$ can be recursively specified by using the symbolic method:

$$\mathcal{P}^{(k,k)} = k\mathcal{Z} + (\mathcal{A} \times \mathcal{P}^{(k,k)} \times \mathcal{P}^{(k,k)}),$$

and

$$\mathcal{P}^{(i,k)} = i\mathcal{Z} + (\mathcal{A} \times \mathcal{P}^{(i,k)} \times \mathcal{P}^{(i,k)}) + (\mathcal{U} \times \mathcal{P}^{(i+1,k)}), \qquad \text{for } i < k.$$

Translating into generations functions and solving for $P^{(k,k)}$ and $P^{(i,k)}$ we obtain

$$P^{(k,k)}(z) = \frac{1 - \sqrt{1 - 4kz^2}}{2z},$$

and

$$P^{(i,k)}(z) = \frac{1 - \sqrt{1 - 4iz^2 - 4z^2 P^{(i+1,k)}(z)}}{2z} \qquad \text{for } i < k. \tag{4.8}$$

Due to the isomorphism mentioned above, we directly get an expression for the generating function $H_{\leq k}(z) = P^{(0,k)}(z)$, which involves $k+1$ nested radicals (just as the generating function $G_{\leq k}$ of the number of closed lambda-terms of bounded unary length did):

$$H_{\leq k}(z) = \frac{1 - \sqrt{1 - 2z + 2z\sqrt{\ldots\sqrt{1 - 4(k-i+1)z^2 - 2z + 2z\sqrt{\ldots + 2z\sqrt{1-4kz^2}}}}}}{2z}$$

We will now proceed analogously to the case of bounded unary length of bindings, by determining the dominant singularity of this generating function together with its type.

First we have to define two auxiliary sequences which will be important for our further considerations:

Definition 4.9 ([11, Def. 6])
Let $(u_i)_{i\geq 0}$ be the integer sequence defined by

$$u_0 = 0 \text{ and } u_{i+1} = u_i^2 + i + 1, \qquad \text{for } i \geq 0$$

and $(N_i)_{i\geq 0}$ by

$$N_i = u_i^2 - u_i + 1 \qquad \text{for all } i \geq 0.$$

Remark 4.10 ([11, Rem. 6])
Obviously the two sequences $(u_i)_{i\geq 0}$ and $(N_i)_{i\geq 0}$ are strictly increasing and have super-exponential growth.

Now we will show that the bounding of the unary height leads to a very unusual behaviour, namely the location and number of the dominant radicands changes with the bound k.

Remark 4.11
In [11] the authors showed that for Motzkin trees of bounded unary height the generating function also involves nested radicals, and the dominant singularity comes from the outermost one. Additionally, we just saw that the dominant singularity for lambda-trees of bounded unary length (cf. Lemma 4.5) comes from the innermost radicand.

Denoting by $R_{i,k}(z)$ the $i-th$ radicand $(1 \leq i \leq k+1)$ of $H_{\leq k}(z)$, according to the numbering from the innermost radicand outwards, we have by Equation (4.8)

$$P^{(i,k)}(z) = \frac{1 - \sqrt{R_{k-i+1,k}(z)}}{2z},$$

Writing the radicands recursively yields

$$R_{1,k}(z) := 1 - 4kz^2$$

and for $2 \leq i \leq k+1$

$$R_{i,k}(z) := 1 - 4(k-i+1)z^2 - 2z + 2z\sqrt{R_{i-1,k}}, \tag{4.9}$$

which gives

$$R_{i,k}(z) = 1-4(k-i+1)z^2-2z+2z\sqrt{1-\ldots\sqrt{1-4(k-1)z^2-2z+2z\sqrt{1-4kz^2}}}.$$

We get $H_{\leq k}(z) = P^{(0,k)}(z) = \frac{1-\sqrt{R_{k+1,k}(z)}}{2z}$, and therefore the dominant singularities of $H_{\leq k}(z)$ and $\sqrt{R_{k+1,k}(z)}$ are the same.

Theorem 4.12 ([11, Thm. 2])
Let $(N_i)_{i\geq 0}$ be the sequence defined in Def. 4.9 and k be an integer. Define j as the integer such that $k \in [N_j, N_{j+1})$.

- *If $k \neq N_j$, then the dominant radicand of $H_{\leq k}(z)$ is the j-th radicand (counted from the innermost one outwards), and the dominant singularity ρ_k is of type $\frac{1}{2}$.*
- *Otherwise, the $j-th$ and the $(j+1)-st$ radicand vanish simultaneously at the dominant singularity of $H_{\leq k}(z)$, which is equal to $\frac{1}{2u_j}$ and of type $\frac{1}{4}$.*

Sketch of proof. Analogously to the proof of Lemma 4.5 we start by showing that for every $k > 0$ and $1 \leq i \leq k+1$, the real function $R_{i,k}(z)$ is strictly decreasing on the positive real line (cf. [11, Lem. 11]) and that the dominant singularity of the radical $\sqrt{R_{i,k}(z)}$ is unique (cf. [11, Lem. 12]). Since the proof for that is very similar to that for $\hat{R}_{i,k}(z)$, introduced in the previous subsection, we will skip the details and refer the reader to the proof of Lemma 4.5, or Lemma 11 and 12 of [11] respectively.
For the integers $j \geq 1$ and $k \in [N_j, N_{j+1})$ let $\sigma_{i,k}$ $(i = 1, \ldots, j+1)$ denote the smallest positive root of the radicand $R_{i,k}(z)$. Then we get that $\sigma_{i+1,k} \leq \sigma_{i,k}$, because if x_0 is a singularity of some radical, it is also a singularity of all the radicals lying more outwards.
Now we assume that the singularity $\sigma_{j,k}$, which per definition is a root of $R_{j,k}$, is also a root of the radicand $R_{j+1,k}$. We will show that then it holds that $\sigma_{j,k} = \frac{1}{1+\sqrt{1+4(k-j)}}$ and $R_{j-p,k}(\sigma_{j,k}) = 4\alpha_p \sigma_{j,k}^2$, for all $p < j$, with the sequence α_p being defined by

$$\begin{cases} \alpha_0 = 0, \\ \alpha_p = (\alpha_{p-1} + p)^2 & \text{for } p \geq 1. \end{cases}$$

(cf. [11, Lem. 15]).
The term for $\sigma_{j,k}$ can be easily derived by shifting Equation (4.9) from j to

$j+1$ and using that for $\sigma_{i,k}$ both radicands $R_{j,k}$ and $R_{j+1,k}$ vanish. The equation for $R_{j-p,k}(\sigma_{j,k})$ can be obtained by again considering Equ. (4.9) for $z = \sigma_{j,k}$, which results in

$$0 = R_{j,k} = -4z^2 + 2z\sqrt{R_{j-1,k}(z)},$$

leading to $R_{j-1,k}(z) = 4z^2$. Iterating yields the desired form for $R_{j-p,k}$.
The next step is to show that if there is a z that cancels both radicands $R_{j,k}$ and $R_{j+1,k}$, then we must have $k = N_j$, with N_j being defined by $N_0 := 0$ and $N_i := \alpha_i - \alpha_{i-1}$ (cf. [11, Lem. 16]). Note that this sequence is precisely the sequence $(N_i)_{i\geq 0}$ introduced in Def. 4.9. We already know that that the simultaneous vanishing of both radicands implies that $z = \sigma_{j,k} = \frac{1}{1+\sqrt{1+4(k-j)}}$. Using the Equation for $R_{j-p,k}(\sigma_{j,k})$ obtained above, yields $1 - 4k\sigma_{j,k}^2 = 4\sigma_{j,k}^2\alpha_{j-1}$, which results in $\sigma_{j,k}^2 = \frac{1}{4(k+\alpha_{j-1})}$. Thus we get

$$\left(\frac{1}{1+\sqrt{1+4(k-j)}}\right)^2 = \frac{1}{4(k+\alpha_{j-1})}.$$

Setting $l = k - j$ and solving leads to $k = (j+\alpha_{j-1})^2 - \alpha_{j-1}$, which is equal to $k = \alpha_j - \alpha_{j-1}$ by the recursion for α_p.
What is left to do is now to show that no more than two radicands can vanish at the same time and if so, then they are consecutive ones (cf. [11, Lem. 17]). To do so, we assume that there were two non-consecutive radicands $R_{i,k}$ and $R_{j,k}$ vanishing simultaneously. Since we know that the zeros of the radicands decrease this would imply that all the radicands $R_{l,k}$ for $i \leq l \leq j$ vanish at the same time. But this could only happen for $z = \sigma_{i,k}$, where the polynomial part $1-4(k-j+1)z^2-2z$ simplifies to $4(j-i-1)z^2$, which is strictly positive as soon as $j > i+1$. Therefore we see that it is not possible that more than two successive radicands vanish simultaneously.
The types of the singularities are obvious, since they arise from square roots, and for the case when two radicands vanish simultaneously the radicands are successive ones. □

To finally get the asymptotics for the number of lambda-terms of bounded unary height, we need the following auxiliary lemma, which gives some information on the behavior of the radicands in a neighborhood of the dominant singularity.

Lemma 4.13 ([11, Prop. 9])
Let $\rho = \sigma_{j,N_j}$ be the dominant singularity of $H_{\leq N_j}(z)$. Then for any $\epsilon > 0$

(i)

$$R_{j,N_j}(\rho - \epsilon) = \gamma_j \epsilon + \mathcal{O}(\epsilon^2) \text{ with } \gamma_j = -\frac{d}{dz} R_{j,N_j}(\rho).$$

(ii)

$$R_{j+1,N_j}(\rho - \epsilon) = 2\rho\sqrt{\gamma_j}\epsilon^{\frac{1}{2}} + \mathcal{O}(\epsilon),$$

(iii) for $p \geq 2$,

$$R_{j+p,N_j}(\rho - \epsilon) = 4\rho^2\lambda_{p-1} + \frac{(2\rho)^{\frac{3}{2}}\gamma_j^{\frac{1}{4}}}{2^{p-2}\sqrt{\prod_{i=1}^{p-2}\lambda_i}}\epsilon^{\frac{1}{4}} + \mathcal{O}(\epsilon^{\frac{1}{2}})$$

with the sequence $(\lambda_i)_{i\geq 1}$ defined by $\lambda_0 = 0$ and $\lambda_{i+1} = i + 1 + \sqrt{\lambda_i}$ for $i \geq 0$.

Sketch of proof. The proof is similar to that of Proposition 4.6. The Taylor expansion of $R_{j,N_j}(z)$ around ρ yields (i). Using the recurrence relation (4.9) for $R_{j,k}(z)$ we immediately get (ii). By computing the expansion of R_{j+2,N_j} around ρ and proceeding by induction the last result (iii) of the lemma can be obtained. □

Now we are in the position to give the asymptotic behaviour of the number of lambda-terms with bounded unary height.

Theorem 4.14 ([11, Thm. 3])
Let $(N_i)_{i\geq 0}$ and $(u_i)_{i\geq 0}$ be as in Def 4.9.

(i) If there exists $j \geq 0$ such that $N_j < k < N_{j+1}$, then there exists a constant h_k such that

$$[z^n]H_{\leq k}(z) \sim h_k n^{-3/2}(\sigma_{j,k})^{-n}, \text{ as } n \to \infty.$$

(ii) If there exists j such that $k{=}N_j$, then the following asymptotic relation holds:

$$[z^n]H_{\leq N_j}(z) \sim h_k n^{-5/4}(\sigma_{j,k})^{-n} = h_{N_j} n^{-5/4}(2u_j)^n, \text{ as } n \to \infty,$$

where

$$h_{N_j} = \frac{\gamma_j^{1/4}(2u_j)^{1/4}}{2^{N_j-j+2}\sqrt{2}\Gamma(3/4)\sqrt{\prod_{i=1}^{N_j-j}\lambda_i}},$$

with γ_j and the sequence $(\lambda_i)_{i\geq 0}$ as defined in the previous lemma.

Sketch of proof. The expressions for $[z^n]H_{\leq k}(z)$ follow immediately from the fact that the dominant singularity for the cases $k \neq N_j$ and $k = N_j$ is of type $\frac{1}{2}$ and $\frac{1}{4}$, and then applying singularity analysis as introduced in Chapter 2. The last equation for h_{N_j} can be obtained analogously as the term in Lemma 4.7 (Equation (4.7)) by means of Lemma 4.13 and $H_{\leq N_j} = \frac{1}{2z}(1-\sqrt{R_{N_j+1,N_j}}(z))$, using also $\Gamma(-\frac{1}{4}) = -4\Gamma(\frac{3}{4})$ for further simplifications. □

Remark 4.15
The constant h_{N_j} decreases exponentially fast as $j \to \infty$. See [11, Prop. 13] for more detailed information.

4.3 BCI- and BCK-lambda-terms

Further simplifications of lambda-terms are represented by BCI(p)- and BCK(p)-lambda-terms. The restriction within BCI(p)-terms lies in the prescribed number of exactly p leaves that are bound by each lambda, while in a BCK(p)-term all lambdas bind at most p leaves. The first three subsections are devoted to the estimation of the asymptotic behavior of BCI(1)- and BCK(1)-lambda-terms, as well as to computing the ratio between BCI(1)-terms and BCK(1)-terms. The results given in those three subsections (4.3.1, 4.3.2, 4.3.3) are mainly based on [18]. In Subsection 4.3.4 we will give asymptotics for the number of BCI(p)-terms, estimated in [10]. The last subsection (4.3.5) deals with the asymptotic number of BCK(2)-terms, which has been derived in [9]. Remember that the size of a lambda-term is still defined as the total number of its nodes (Def. 1.7).

4.3.1 BCI(1)-lambda-terms

The simplest subclass of BCI(p)-lambda-terms are the BCI(1)-lambda-terms, which are often just referred to as BCI-terms.

Definition 4.16 (BCI-lambda-terms)
The class of BCI-lambda-terms is built of all closed lambda-terms, where each lambda binds exactly one leaf.

In a BCI-lambda-term the number of leaves is equal to the number of unary nodes and greater by one than the number of binary nodes. Thus, the size of BCI-terms is equal to $3k + 2$ (k binary nodes, $k + 1$ unary nodes and $k+1$ leaves). Denoting the number of BCI-terms of size n by a_n, we get that $a_n = 0$ for $n \not\equiv 2 \bmod 3$. Thus, for the number a_n^* of BCI-lamda-terms with n binary nodes, it holds that $a_n^* = a_{3n+2}$.

Lemma 4.17 ([18, Lem.37])
The sequence (a_n^*) *satisfies the recurrence:*

$$a_0^* = 1, \qquad a_1^* = 5, \tag{4.10}$$

$$a_n^* = 6na_{n-1}^* + \sum_{i=1}^{n-2} a_i^* a_{n-i-1}^*, \qquad \textit{for } n \geq 2. \tag{4.11}$$

Proof. The initial conditions (4.10) are derived by considering that the only BCI-term of size 2 (no binary nodes) is $\lambda x.x.$ and that there are five BCI-terms of size 5 (one binary node): $\lambda xy.xy$, $\lambda xy.yx$, $(\lambda x.x)(\lambda x.x)$, $\lambda x.(\lambda y.y)x$, and $\lambda x.x(\lambda y.y)$. In order to prove Equation (4.11) let P be a BCI-term with $n \geq 2$ binary nodes. As usual we distinguish between the terms starting with an application and those starting with an abstraction. Both cases are depicted in Figure 4.1.

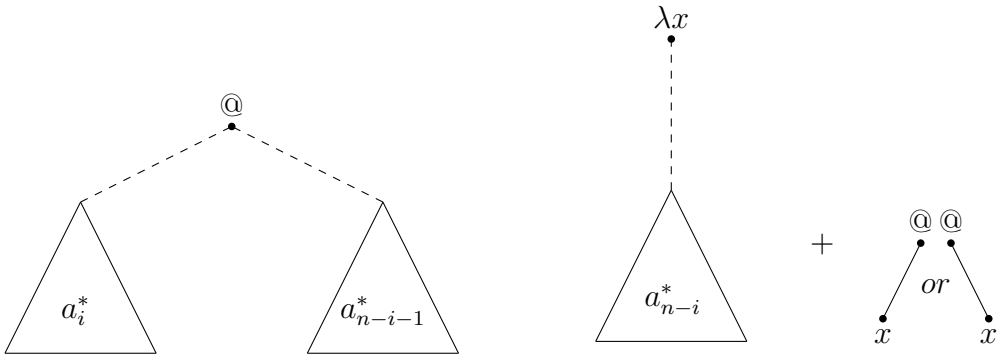


Figure 4.1: Two ways of obtaining a BCI term with $n \geq 2$ binary nodes (cf. [18], Fig.4; application nodes are labeled with @ for a better differentiation).

- 1.case: P starts with an application, i.e. $P \equiv MN$.
  Such a term results by combining two smaller BCI-terms whose sizes add to $n-1$. This gives $\sum_{i=0}^{n-1} a_i^* a_{n-i-1}^*$ possibilities.

- 2.case: P starts with an abstraction, i.e. $P \equiv \lambda x.M$, where x occurs free in M exactly once. Obviously the parent of the leaf labeled with x has to be a binary node. Thus, M can be seen as a lambda-tree for some BCI-term Q with $n-1$ binary nodes with an additional leaf labeled with x (and therefore also with an additional edge). The insertion of the edge that contains the leaf x can either appear to the right or to the left (cf. Figure 4.1). Since the number of all edges in the tree for the BCI-term Q is $3n-1$, we all together have $2(3n-1)a^*_{n-1}$ possible insertions.

Summing up we get Equation (4.11). □

The first values of (a_n) are the following ([18, p.687]):

$$0, 0, 1, 0, 0, 5, 0, 0, 60, 0, 0, 1150, 0, 0, 27120, 0, 0, 828250, 0, 0, 30220800, \ldots$$

With $A(x)$ denoting the generating function for the sequence (a^*_n), we get

$$6x^2\frac{\partial A(x)}{\partial x} + xA^2(x) + (4x-1)A(x) + 1 = 0, \quad A(0) = 1.$$

This is a non-linear Riccati differential equation, which therefore has a non-elementary function as solution. We could study this equation and thereby estimate the asymptotic behavior of the sequence (a^*_n), but instead we briefly want to introduce another approach, that has been performed in [8] and that is also applicable for BCK(1)-terms.

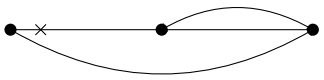
Bodini et al. showed that there is a one-to-one correspondence between BCI-terms and special classes of rooted combinatorial maps. These maps can be regarded as connected graphs (with possible loops and multi-edges) embedded in an oriented surface up to homotopy ([8, p.3]). But we will deal with another definition of maps, which will be useful for further considerations.

In [8, Def. 1.1], a map is introduced as a triplet (E, α, β), where E is a finite set of half-edges, α is a permutation on E and β is an involution on E without any fixed points. Additionally α and β have to generate a transitive subgroup of the permutation group. The permutations α and β contain information on the order of the half-edges around the vertices and of the correspondence between two half-edges. Using this definition a rooted map is a map with a distinguished half-edge.
The size of a map is defined as the number of its half-edges.

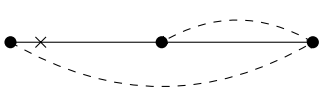
Now we are able to give the following result:

Theorem 4.18 ([8, Thm. 3.4])
The class of BCI-terms is in bijection with the class of unlabeled rooted maps with root degree of 2 and all other vertex degrees equal to 3. A BCI-term of size $3k + 2$ *corresponds to a map of size* $6k + 2$.

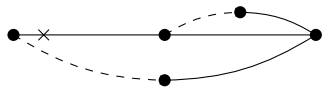
We will explain this bijection by means of an example.
We start with a map of size 8 with root degree 2 (the cross marks the half-edge corresponding to the root of the map) and all other vertex degrees equal to 3:



By performing a so-called right-depth-first traversal (see [8, p.6]), we get a unique spanning tree of the map. All the edges that are not contained in that tree are now presented as dashed lines.

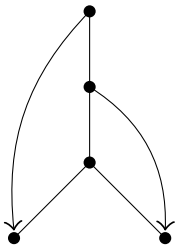


Then we insert a node in each of the dashed edges, thereby creating a new edge for each insertion. The new edges (that would be visited next in the right-depth-first traversal in order to receive a spanning tree) are now replaced by continuous lines.



In [8] this process is introduced as so-called L-morphing.

The last step is to replace the dashed edges with arcs pointing away from the root, resulting in the following BCI-term of size 5:



It can easily be seen that this process is invertible by performing each of the steps backwards (cf. [8, Lem. 2.5]).

Thus, by establishing the asymptotic behavior of the specific class of unlabeled rooted maps introduced in Theorem 4.18, the authors in [8] derived the asymptotics for the number of BCI-terms. The advantage of this approach is, that an explicit formula for the generating function $C(z)$ of rooted maps with root degree 2 and all other vertex degrees 3, as well as the asymptotics of its coefficients, is known. Thus, we calculate

$$[x^{3k+2}]A(x) = [x^{6k+2}]C(x),$$

and obtain the following result:

Theorem 4.19 ([8, Thm. 3.8])
For $n \equiv 2$ mod 3 the number a_n of BCI-terms is asymptotically

$$a_n \sim \frac{\sqrt[3]{2}\sqrt{6}}{2n^{1/6}\sqrt{\pi}} \left(\frac{2n}{e}\right)^{n/3} (1 - \frac{1}{2n} + O(n^{-2})) \quad \textit{for } n \to \infty.$$

Otherwise it holds that $a_n = 0$.

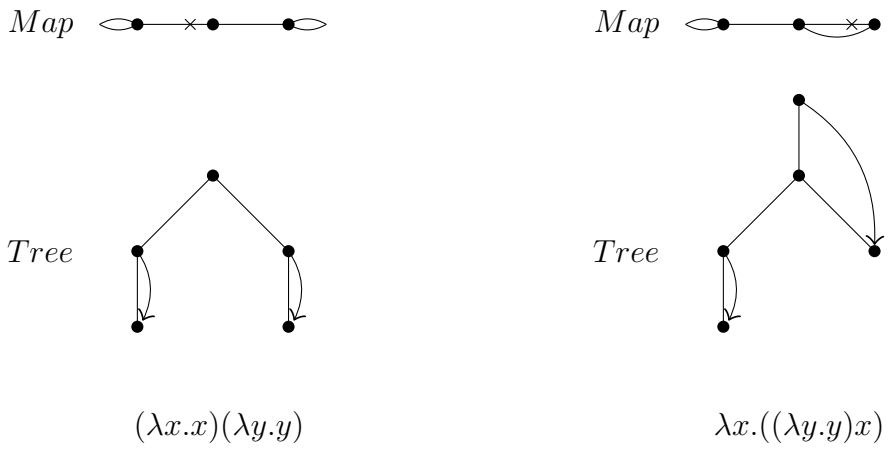


Figure 4.2: Two rooted unlabeled maps and their corresponding BCI-lambda-terms, and lambda trees respectively (cf. [8, Fig. 5]).

4.3.2 BCK(1)-lambda-terms

In contrast to the strictly prescribed number of one binding of each lambda, as it is the case for BCI-terms, BCK(1)-terms just use this number as an upper bound. Equivalently to BCI-terms we will denote the BCK(1)-terms simply as BCK-terms, as it is often done in the literature.

Definition 4.20 (BCK-lambda-term)
The class of BCK-lambda-terms contains all closed lambda-terms, where each lambda binds at most one leaf.

Let b_n denote the number of BCK-terms of size n.

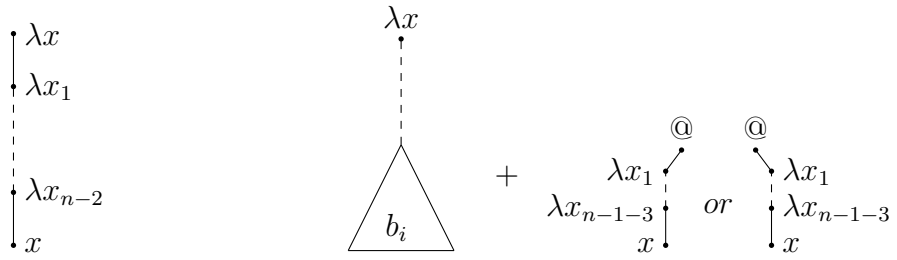
Lemma 4.21 ([18], Lem. 39)
The sequence (b_n) satisfies the following recursive equation:

$$b_0 = b_1 = 0, \quad b_2 = 1, \quad b_3 = 2, \quad b_4 = 3, \tag{4.12}$$

$$b_n = b_{n-1} + 2\sum_{i=0}^{n-3} i b_i + \sum_{i=0}^{n-1} b_i b_{n-i-1} + 1 \qquad \textit{for } n \geq 5. \tag{4.13}$$

Proof. The initial conditions (4.12) follow directly from the fact that there are are no BCK-terms of size 0 and 1, there is one term of size 2, namely $\lambda x.x$, two terms of size 3, namely $\lambda xy.x$, and $\lambda xy.y$, and three terms of size 4: $\lambda xyz.x$, $\lambda xyz.y$ and $\lambda xyz.z$.
In order to determine Equation (4.13) let P be a BCK-term of size $n \geq 5$. We perform again a distinction of cases:

- 1. case: P starts with an application, i.e. $P \equiv MN$. This case is analogously to the BCI-case (cf. 1. case in the proof of Lemma 4.17), and therefore leads to $\sum_{=0}^{n-1} b_i b_{n-i-1}$ possibilities.
- 2. case: P is in the form of an abstraction, $P \equiv \lambda x.M$ and x occurs free in M exactly once. We have to consider two subcases here, which are both depicted in Figure 4.3.
  - Case 2a: $M \equiv \lambda x_1 \ldots x_{n-2}.x$.



Case 2*a*. Case 2*b*.

Figure 4.3: Both subcases of the second case of the construction of a BCK-term of size $n \geq 5$ (cf. [18, Fig. 5]).

– Case 2b: M is a term built of a BCK-term Q of size $i = 2, \ldots, n-3$ with an additional term $\lambda x_1 \ldots x_{n-i-3}.x$ inserted on one of its edges or on the edge joining λx with Q. Since Q has $i-1$ edges and the insertion can either be carried out on the left or on the right, this case gives us $2\sum_{i=2}^{n-3} ib_i$ possibilities.

Thus, from case 2 we derive altogether $1+2\sum_{i=2}^{n-3} ib_i$ possibilities.

- 3. case: P is in the form of an abstraction, $P \equiv \lambda x.M$ and x does not occur free in M. This case gives of course b_{n-1} possibilities, since M can be any BCK-term of size $n-1$.

Summing up yields Equation (4.13). □

The first values of (b_n) are the following (cf. [18, p.688]):

$0, 0, 1, 2, 3, 9, 30, 81, 225, 702, 2187, 6561, 19602, 59049, 177633, 532170, 1594323, \ldots$

With $B(x)$ denoting the generating function for the sequence (b_n), we get

$$2x^4\frac{\partial B(x)}{\partial x} + (x - x^2)B^2(x) - (1 - x)^2B(x) + x^2 = 0, \quad B(0) = 0.$$

Again we obtain a non-linear Riccati differential equation, but solving the equation has turned out to be much more complicated than in the BCI-case. Therefore we use the approach introduced in the previous subsection, that is taking advantage of an isomorphism between a special class of rooted unlabeled maps and BCK-lambda-terms.

Theorem 4.22 ([8, Thm.3.5])
The class of BCK-terms is in bijection with the class of unlabeled rooted maps with root degree of 1 or 2 and all other vertex degrees equal to 2 or 3. In this correspondence, a BCK-term of size $k + 1$ is associated with a map of size $2k$.

The isomorphism follows the same algorithm as for the BCI case by performing the L-morphing process. Investigating asymptotics for the suitable class of maps the authors in [8] proved the following result.

Theorem 4.23 ([8, Thm. 3.10])
The number b_n of BCK-lambda-terms is asymptotically

$$b_n \sim \frac{e^{1/18}\sqrt[3]{2}}{\sqrt{6\pi}n^{1/6}}\left(\frac{2n}{e}\right)^{n/3} e^{1/2(2n)^{2/3}-1/6(2n)^{1/3}}.$$

4.3.3 Ratio of BCI to BCK

We already know that there are no BCI-terms of size $n \not\equiv 2 \bmod 3$ and that the set of BCI-terms is a proper subset of the set of BCK-terms. In this subsection we will show that the ratio of BCI-terms of size n to BCK-terms of size n, tends to 0.
Each BCK-term can be obtained from a BCI-term with some additional (possibly none) lambdas. Therefore we obtain the following formula for b_n depending on the sequence a_n:

Lemma 4.24 ([18, Lem. 41])
For $k \in \mathbb{N}$ the following formula holds

$$b_{3k+2} = \sum_{i=0}^{k} \binom{3k}{3i} a_{3i+2}$$

Proof. A BCI-term of size $3i+2$ can be transformed into a BCK-term of size $3k+2$ by inserting $3k+2-3i-2 = 3k-3i$ additional lambdas. Using that the number of possibilities to choose k elements out of n with repetition is $\binom{n+k-1}{n-1}$ gives $\binom{3k}{3k-3i} = \binom{3k}{3i}$ possible insertions of the $3k-3i$ lambdas in the $3i+1$ edges of the BCI-term. □

Theorem 4.25 ([18, Thm. 42])
The asymptotic ratio of BCI-terms to BCK-terms is equal to 0.

Proof. By Theorem 4.19 and Lemma 4.24, we get

$$\frac{a_{3k+2}}{b_{3k+2}} \leq \frac{a_{3k+2}}{a_{3k+2} + \binom{3k}{3} a_{3k-1}} \quad \to \ 0, \qquad \text{as } k \to \infty.$$

□

Remark 4.26
Using the asymptotics for BCI- and BCK-lambda-terms (Theorem 4.19 and 4.23) we are able to derive the asymptotic ratio between the number of BCI- and BCK-terms of size n for $n = 2$ mod 3 (cf. [8, Cor. 3.11]):

$$\frac{a_n}{b_n} \sim 3e^{-2^{-1/3}n^{2/3}+1/6\sqrt[3]{2}n^{1/3}-1/18}$$

4.3.4 BCI(p)-lambda-terms

The aim of this subsection is to derive asymptotics for the number of BCI(p)-terms, as it is done in [10].

Definition 4.27 (BCI(p)-lamda-term, [10, Def. 1])
BCI(p) is the (non-empty) set of closed lambda-terms where each unary node has exactly p pointers, i.e. binds exactly p occurences of its variable.

The smallest BCI(p)-terms consist of a unary node at the beginning, followed by p leaves that are bound by the lambda. Obviously, by removing the unary root and all its pointers, we get an ordinary binary tree. Therefore the number of smallest BCI(p)-terms is equal to the number of binary trees with $p-1$ binary nodes, which is exactly the Catanlan number $C(p-1) = \binom{2p-2}{p-1}/p$. The size of such a term is clearly $\underbrace{p-1}_{\textit{binary nodes}} + \underbrace{p}_{\textit{leaves}} + \underbrace{1}_{\textit{root}} = 2p$.

In Subsection 4.3.1 we derived that there are only BCI(1)-terms of size $n = 2$ mod 3. Analogously we get for the size n of a BCI(p)-term with j unary nodes that $n = (2p+1)j - 1$ (because it has pj leaves and therefore $pj - 1$ binary nodes).

Let us denote by $g_n = g_n^{(p)}$ the number of BCI(p)-terms of size n and by $G_p(z)$ the generating function of this sequence.

With the considerations above we have

$$G_p(z) = \sum_{j \geq 1} g_{j(2p+1)-1} z^{j(2p+1)-1}.$$

Proposition 4.28 ([10, Prop. 4])
The generating function of BCI(p)-terms satisfies the differential equation

$$G_p(z) = C(p-1)z^{2p} + zG_p(z)^2 + \Delta_p G_p(z) \tag{4.14}$$

where

$$\Delta_p = \sum_{l=1}^{p} \frac{\alpha_{l,p}}{l!} z^{l+2p+1} D^l$$

with constants $\alpha_{l,p}$ defined by

$$\alpha_{l,p} = \sum_{\sum_i s_i = l; \sum_i i s_i = p} \binom{l}{s_1, \ldots s_p} \prod_{m=1}^{p} \binom{2m}{m}^{s_m}. \tag{4.15}$$

Proof. In order to derive a formal equation for BCI(p)-terms, note that a BCI(p)-term falls into exactly one of the following three categories:

- It is either a smallest term (which are explained above).
- Or its root is a binary node and the two sub-terms attached to the root are themselves BCI(p)-terms.
- Or its root is a unary node and the sub-term attached to the root is an open BCI(p)-term with exactly p free leaves.

Therefore we get the following formal specification for the set $\mathcal{T}$ of BCI(p)-terms:

$$\mathcal{T} = \mathcal{S} \cup (\{\circ\} \times \mathcal{T} \times \mathcal{T}) \cup (\{\circ\} \times \tilde{\mathcal{T}}),$$

where $\mathcal{S}$ is the set of all smallest BCI(p)-terms, and $\tilde{\mathcal{T}}$ a certain set of open BCI(p)-terms.

The first summand of Equation (4.14) obviously corresponds to the first case, where we have a smallest term. Remember that we already got that there are $C(p-1)$ such terms and that they have the size $2p$. Clearly, the second summand is represented by the case that the root is a binary node, i.e. case 2. To derive the third summand we will give a step-by-step explanation how to create such a term $\tilde{T} \in \{\circ\} \times \tilde{\mathcal{T}}$:
We start with a BCI(p)-term T and mark p of its nodes, where multiple choices of the same node are allowed. With every choice of a leaf we also "hit" the edge leading to it. In case of the root the corresponding edge is the edge connecting it with the new root. Now assume that l edges are hit and s_i of them exactly i times. We then replace an edge that is hit i times by a path where at each node of the path a binary tree is attached, either to left or to the right, such that the number of leaves of all these binary trees altogether is i. Since each of the replacements creates i new leaves and i new internal nodes, the whole replacement process yields $\sum_{i=1}^{p} is_i$ new leaves and p new internal nodes in T. $\tilde{T}$ is now obviously an open BCI(p)-term of size $|T| + 2p + 1$. For a bijection we also have to consider the other direction: We start with a BCI(p)-term T with a unary node as root. By removing the root and all its pointers we get a term with p free leaves. Obviously these leaves must be children of binary nodes and therefore they form binary trees consisting of free leaves only.
What is now left to do is to count in how many ways such a construction can be done. Note that the generating function of binary trees is $T(u) = \sum_{n\geq 1} C(n-1)u^n = \frac{1-\sqrt{1-4u}}{2}$. Thus, the number of sequences of left- or right-binary trees with exactly i leaves is

$$[u^i]\frac{1}{1-2T(u)} = [u^i]\frac{1}{\sqrt{1-4u}} = \binom{2i}{i}.$$

Each of the s_i edges that is hit i times can therefore be raplaced by one of these $\binom{2i}{i}$ sequences of binary trees. This gives the factor $\prod_{i=1}^{p}\binom{2i}{i}^{s_i}$ in Equation (4.15). The factor $\binom{l}{s_1,\ldots,s_p}$ in Equation (4.15) accounts for the different possibilities how to partition the l edges according to the times that they are hit. Finally we have to consider that marking l distinct edges corresponds to applying the operator $z^l D^l/l!$ on the level of generating functions and it also arises a factor z^{2p+1}, since we created $2p+1$ new nodes during the whole replacement process. □

Regarding Equation (4.14) it is now easy to set up a recurrence relation for the coefficients of $G_p(z)$.

Proposition 4.29 ([10, Prop. 10])
The coefficients $g_{n(2p+1)-1}$ satisfy the recursion

$$g_{n(2p+1)-1} = \sum_{l=1}^{n-1} g_{j(2p+1)-1} g_{(n-1-l)(2p+1)-1} + Q_p(n-1) g_{(n-1)(2p+1)-1}, \qquad \text{for } n \geq 2,$$

with initial condition $g_{(2p+1)-1} = C(p-1)$ and where

$$Q_p(n) = \sum_{m=1}^{p} \alpha_{m,p} \binom{n(2p+1)-1)}{m},$$

with $\alpha_{m,p}$ defined as in (4.15).

Proof. The proposition follows directly by translating Equation (4.14): The first term of the equation only affects the case $n = 1$, the second term is a Cauchy product and the differential operator yields a shift of the coefficients of the power series and a multiplication by $Q_p(n-1)$. □

We will now establish another representation for the polynomials $Q_p(n)$, which will be useful for further estimations.

Lemma 4.30 ([10, Lem. 11])
The polynomials $Q_p(n)$ can be represented more explicitly as

$$Q_p(n) = 4^p \binom{(p+\frac{1}{2})n + p - \frac{3}{2}}{p}.$$

Proof. By setting $f(u) = \frac{1}{\sqrt{1-4u}}$ we can easily see that $\alpha_{m,p} = [u^p](f(u)-1)^m$. Since the occurring binomial coefficient is zero for $m > p$, we get

$$\begin{aligned} Q_p(n) &= \sum_{m=1}^{p} \alpha_{m,p} \binom{n(2p+1)-1)}{m} \\ &= [u^p] \sum_{m \geq 1} \binom{(2p+1)n-1)}{m} (f(u)-1)^m \\ &= [u^p] f(u)^{(2p+1)n-1} \\ &= 4^p \binom{(p+\frac{1}{2})n + p - \frac{3}{2}}{p}. \end{aligned}$$

□

Our goal is now to derive the asymptotic behavior of the number of BCI(p)-terms of size n. For that purpose we are going to linearize the differential equation, which is possible because of the fast growth of the coefficients of $G_p(z)$. As we have already mentioned in the last chapter the significant increase in the number of lambda-terms of a given size when compared to Motzkin trees comes from the large number of ways of binding a leaf to unary nodes. Thus, the contribution of the term G_p^2 corresponding to the binary tree-like structure, will turn out to be asymptotically negligible when compared to that of the differential term which involves the various possible bindings of leaves.

Therefore we will work with the linearized equation

$$L_p(z) = C_{p-1} z^{2p} + \Delta_p L_p(z).$$

Theorem 4.31 ([10, Thm. 18])
Set $l_{p,n} = [z^n]L_p(z)$. Then, for fixed p and $n \to \infty$

$$l_{p,n} \sim B_p \beta_p^{n-1} n^{\gamma_p} (n-1)!^p$$

where

$$B_p = C(p-1) \prod_{k=1}^{p} \frac{1}{\Gamma\left(1 + \frac{2(p-k)-1}{2p+1}\right)}$$

$$= C(p-1) exp\left(-\frac{2p+1}{2} \int_1^2 \log(\Gamma(x)) dx\right) \left(1 + O\left(\frac{1}{p}\right)\right), \quad as\ p \to \infty,$$

$$\approx C(p-1)(1.0844375142\ldots)^{(2p+1)/2} \left(1 + O\left(\frac{1}{p}\right)\right)$$

and

$$\beta_p = \frac{(4p+2)^p}{p!}, \quad \gamma_p = \frac{p(p-2)}{2p+1}.$$

Proof. The linearized equation $L_p(z) = C(p-1)z^{2p} + \Delta_p L_p(z)$ implies $l_{p,2p} = C(p-1)$ and $l_{p,n} = Q_p(n-1) l_{p,n-2p-1}$ for $n > 2p$. Thus

$$l_{p,(2p+1)n-1} = C(p-1) \prod_{j=1}^{n-1} Q_p(j)$$

$$= C(p-1) \prod_{j=1}^{n-1} 4^j \frac{\left[(j+\frac{1}{2})n + j - \frac{3}{2}\right] \cdot \left[(j+\frac{1}{2})n + j - \frac{5}{2}\right] \ldots \left[(j+\frac{1}{2})n + j - \frac{2j+1}{2}\right]}{j!},$$

which can be simplified to

$$C(p-1)\beta_p^{n-1}(n-1)!^p \prod_{j=1}^{n-1} \prod_{k=1}^{p} \left(1 + \frac{2(p-k)-1}{2p+1} \cdot \frac{1}{j}\right).$$

Now observe that for $n \to \infty$ it holds that

$$C(p-1) \prod_{j=1}^{n-1} \prod_{k=1}^{p} \left(1 + \frac{2(p-k)-1}{2p+1} \cdot \frac{1}{j}\right) \sim B_p n^{\gamma_p}.$$

The asymptotics for B_p can be derived from Euler McLaurin's formula. □

The authors in [10] also derived the asymptotic behavior for the non-linearized equation (Equation (4.14)). We will only give the result as it can be seen that it rarely differs from the solution obtained in Theorem 4.31:

Remark 4.32 ([10, Thm. 19])
For $p \geq 2$, the number of BCI(p)-terms of size $(2p+1)n-1$ is asymptotically

$$A_p \beta_p^{n-1} n^{\gamma_p} (n-1)!^p$$

where β_p and γ_p are as defined in Theorem 4.31 and $A_p = a_p B_p$ with B_p as in Theorem 4.31 and $a_p = 1 + O(1/(pe^p))$, as $p \to \infty$.

4.3.5 BCK(2)-terms

Compared to BCI(p)-terms it is a lot more difficult to enumerate BCK(p)-terms, which are a proper superset of BCI(p)-terms. In this subsection we will just give the basic idea of how to derive asymptotics for the number of BCK(2)-terms by presenting an approach introduced in [9]. For more detailed information see [9], where also a brief outlook for the problem of determining the asymptotic number of BCK(p)-terms with $p > 2$ can be found.

Definition 4.33 (BCK(p)-lambda-term, [10, Def. 1])
BCK(p) is the set of closed lambda-terms where each unary node binds at most p leaves.

Let $F_p(z)$ be the generating function associated to BCK(p)-terms.
In [10] the authors showed the following result:

Proposition 4.34 ([10, Prop.8])
Let $M(z,u)$ denote the generating function of Motzkin trees where z marks the size (i.e. the total number of nodes) and u marks the number of leaves. This function is given by the unique power series solution of $M(z,u) = uz + zM(z,u) + zM(z,u)^2$, that is

$$M(z,u) = \frac{1 - z - \sqrt{(1-z)^2 - 4uz^2}}{2z}.$$

Then $F_p(z)$ is given as the solution of

$$F_p(z) = z[u^p]\frac{M(z,u)}{1-u} + zF_p(z)^2 + z[u^p]\frac{1}{1-u}F_p\left(\frac{z}{1-2zM[z,u)}\right).$$

Sketch of proof. Similar to the BCI-case (cf. Proposition 4.28): As minimal structures we have Motzkin trees with an additional root having pointers to all leaves. The application-case is equivalent. The terms starting with a unary root node (and not being smallest terms) can be generated in the following way:
We fix the number l of pointers we want to have at the root and then do an edge hitting process as in the BCI case with the difference that now we substitute the hit edges by sequences of left or right Motzkin trees with an additional unary root node (corresponding to the nodes in the paths which substitute the hit edges) such that these trees have altogether l leaves. Recalling on the level of generating functions edge hitting corresponds to applying a differential operator, we get the differential equation for $F_p(z)$. □

Since we are interested in the number of BCK(2)-terms, we set $p = 2$ and with $F(z) := F_2(z)$ we get:

$$F(z) = z[u^2]\frac{M(z,u)}{1-u} + zF(z)^2 + z[u^2]\frac{1}{1-u}F\left(\frac{z}{1-2zM(z,u)}\right).$$

Analogously to the BCI(p)-case we proceed by omitting the quadratic term,

which yields

$$G(z) = z[u^2]\frac{M(z,u)}{1-u} + z[u^2]\frac{1}{1-u}G\left(\frac{z}{1-2zM(z,u)}\right).$$

Moreover, in [9] it has been shown, that this modification leads only to a difference by a multiplicative constant in the asymptotic result ([9, Prop.5]). Unfortunately $G(z)$ is not an analytic function, due to the fast growth of its coefficients g_n. Thus we will work with the exponential generating function (see for example [15])

$$H(z) = \sum_{n\geq 0} \frac{g_n}{n!} z^n,$$

which is an entire function, that is also Hayman-admissible ([9, Lemma 6]).

Applying the saddle point method to $H(z)$ the authors in [9] obtained an asymptotic behavior of $h_n = [z^n]H(z)$ up to a multiplicative constant.

Using that the asymptotics of the coefficients of the generating functions $G(z)$ and $F(z)$ only differ by a multiplicative constant, i.e.

$$\frac{f_n}{g_n} = \frac{f_n}{n!h_n} \to \text{const},$$

we get the asymptotics for f_n:

Theorem 4.35 ([9, Thm. 1])
The asymptotic number of BCK(2)-terms of size n satisfies

$$f_n \sim An^{2n/5}2^{n/5}e^{-2n/5}exp\left(2^{-8/5}n^{4/5} + \frac{7\cdot 2^{4/5}}{15}n^{3/5} - \frac{17\cdot 2^{1/5}}{75}n^{2/5} - \frac{41\cdot 2^{3/5}}{500}n^{1/5}\right)n^{-3/5}$$

where A is some positive constant.

Chapter 5

Structure of lambda-terms

The structure of lambda-terms strongly differs from that of trees, while it rather resembles the structure of directed acyclic graphs, which will be introduced in the second part of this thesis. This chapter is split into two sections: The first one gives results on the structure of lambda-terms obtained by experiments where lambda-terms have been randomly generated, performed in [36]. In the second section we will present some results established in [13] which have been formally proved by means of codings.

5.1 Experimental results

In [36] Jue Wang derived an algorithm for the random generation of lambda-terms, which we will not discuss in detail. The aim of this section is to simply point out a few interesting properties of lambda-terms that have been detected by experiments based on random generation. In this section the size of a lambda-term corresponds to the total number of nodes in the enriched lambda-tree (Def. 1.7), i.e. variables have size 1.

5.1.1 Top-Level Application Nodes

We start by considering the percentage of lambda-terms that start with an application among all lambda-terms (see [36, Section 5.1]). For that purpose the number of lambda-terms in the appropriate classes have been recursively computed. Fig. 5.1 shows that the percentage of terms with top-level application nodes decreases rapidly with increasing size. This will not appear to be a big surprise - as mentioned at some earlier points in this thesis, it is the number of possible variable bindings that leads to the great increase in the number of terms. Remember that for BCI(p)-terms and BCK(2)-terms we worked with the linearized equation and omitted the quadratic term (corresponding to lambda-terms starting with a binary node), which altered the asymptotic number only by a multiplicative constant. As an example for size 10 there are only 550 terms with an application node at the top-level out of a total of 7558 terms.

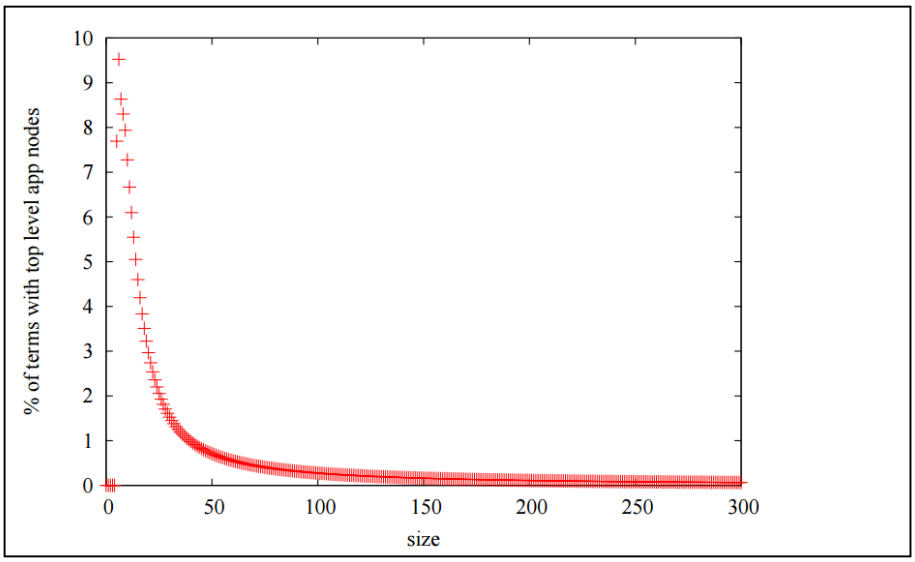


Figure 5.1: Percentage of lambda-terms with top-level application nodes from size 0 to 300 ([36, Fig.11]).

5.1.2 Ratio of application nodes to lambda nodes

Now we present some results on the ratio of the total number of application nodes to lambda nodes in an arbitrary lambda-term (see [36, Section 5.2]). Wang randomly generated 10,000 terms and then calculated the ratio of binary to unary nodes. In contrast to the result presented in the previous subsection, which has been derived explicitly by solving a recursion, the result given in this subsection is probabilistic. Consequently, the accuracy of the results presented in Fig. 5.2 depends on how well the 10,000 generated terms represent the total set of lambda-terms. In Figure 5.2 it can be seen that the ratio of application nodes to lambda nodes is steadily increasing as the size of the lambda-term increases and it exceeds the value 1 already for a small size, i.e. if the size of the lambda-term is big enough the average number of application nodes is greater than that of lambda nodes.

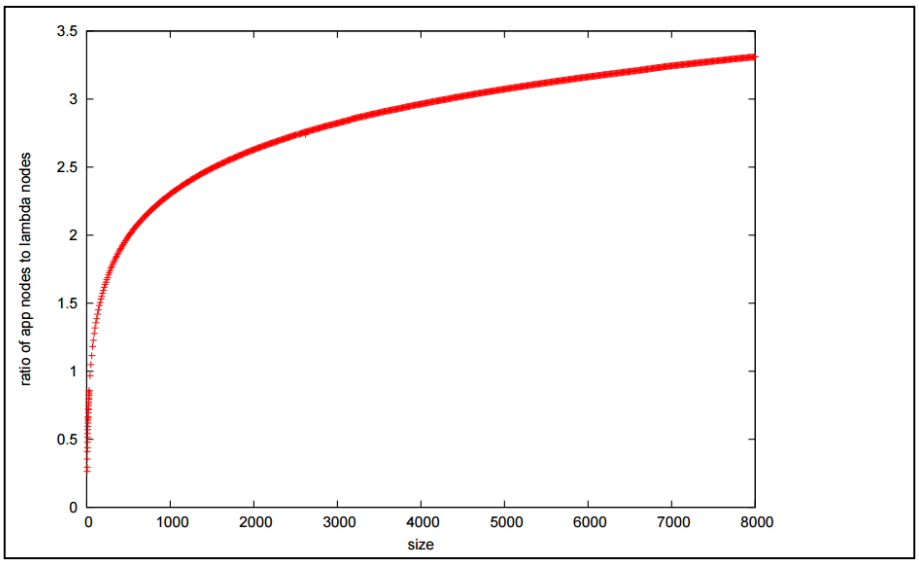


Figure 5.2: Ratio of application nodes to lambda nodes in terms from size 0 to 8000 ([36, Fig.12]).

5.1.3 Percentage of dummy variable bindings

In this subsection we consider the percentage of so-called dummy variable bindings among all variable bindings (see [36, Section 5.3]). A dummy variable binding is characterized by bound variables that do not occur in the body of the term. As an example, in the term $\lambda x(\lambda y.x)$, λy is a dummy binding while λx is not a dummy binding. Considering for example lambda-terms of size 4. There are altogether 6 variable bindings out of 10 (i.e. 60%) that are dummies (cf. [36, Fig.1]).

Fig. 5.3 shows the percentage of dummy bindings in lambda-trees for size 10 to 4000. Analogously to the result presented in the previous subsection, the data that has been used to derive Figure 5.3 is based on 10,000 randomly generated terms. It can easily be seen that with increasing size there are more lambda nodes that are actually binding variables and less dummy bindings.

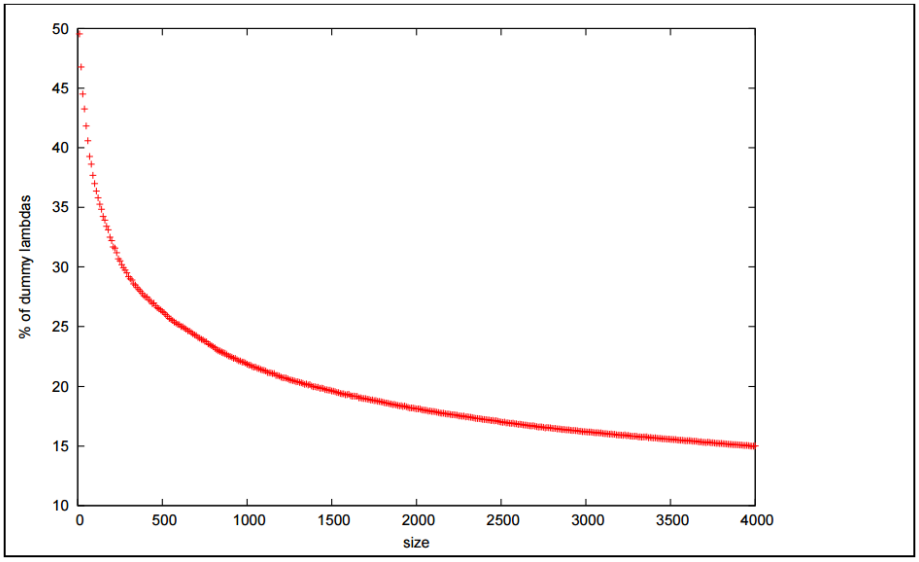


Figure 5.3: Percentage of dummy variable bindings for terms from size 0 to 4000 ([36, Fig.13]).

5.2 Results proved by codings

In this section we will give the basic idea of some results on the structure of lambda-terms, that can be formally proved by means of codings. This approach has been introduced in [13], where we refer the interested reader to, for gaining more detailed information. Throughout this section, the size of a lambda-term corresponds to the number of its inner nodes, i.e. variables have the size 0.

Within [13] the following properties of lambda-terms (amongst others, which will not be discussed any further in this thesis) have been shown:

- A random lambda-term starts with a long chain of lambdas ([13, Thm. 6.3]).

- Head lambdas bind "many" occurences, meaning that in a random lambda-term head lambdas actually bind some variables ([13, Section 6.2]).

- A random term avoids any fixed closed lambda-term, i.e. that given any fixed closed lambda-term, almost no term has this term as a subterm ([13, Section 6.3]).

- The maximum number of incomparable binding lambda nodes of typical lambda-terms, meaning that there is no branch in the lambda-tree containing more than one of those lambda-nodes (called lambda-width of a term), is very small. More precisely almost all closed lambda-terms are of width at most 2 ([13, Section 6.4]).

In order to give the exact formulations of the theorems that contain the results summarized above, we would need to introduce a lot of additional notations and definitions, which would exceed the limits of this thesis. Consequently, this also makes it very difficult to give a sketch of the proofs for the mentioned properties. Therefore we will just briefly introduce the scheme, with which the authors in [13] proved those results. The interested reader is referred to [13] for the complete proofs.

First a set $\Lambda_n(\mathcal{P})$ of closed lambda-terms of size n satisfying some property $\mathcal{P}$ is introduced. The next step is to define an injective and size-preserving function $\phi_n^{\mathcal{P}} : \Lambda_n(\mathcal{P}) \to \Lambda_n$ (called a coding) in such a way that the ratio between the number of elements in the image of $\phi_n^{\mathcal{P}}$ and all the elements in $\Lambda_n(\mathcal{P})$ is asymptotically 0. Then it follows that this property is not satisfied by a random lambda-term.
In order to show the results above, one has to consider successive sets of lambda-terms $X_1, \ldots, X_k$ with $X_{i+1} \subseteq X_i$ and prove:

(1) The ratio of the number of lambda-terms in X_1 compared to the number of all lambda-terms is asymptotically 1.

(2) The ratio between the number of lambda-terms in $X_i \setminus X_{i+1}$ and the number of all lambda-terms is asymptotically 0. This implies that the ratio of the number of lambda-terms in X_{i+1} among all closed lambda-terms is asymptotically 1.

Part II

DAGs

Chapter 6

Definitions

The second part of this thesis is devoted to the problem of counting directed acyclic graphs (DAGs), which are closely related to lambda-trees, especially each lambda-tree can be considered as a DAG. In this chapter we will introduce some basic definitions that will be used in the subsequent chapters.

Definition 6.1 (directed acyclic graph, DAG)
A directed acyclic graph (DAG), is a directed graph without any directed cycles.

It is easy to see, that every lambda-tree can be regarded as a DAG by removing all vertex labels and directing all undirected edges downwards (i.e. away from the root). Since then no arc ends at a node that is closer to the root than the start node of the arc, it follows that there cannot be any cycles (cf. Figure 6.1). Obviously the reverse is not true, i.e. there are DAGs, that do not correspond to lambda-trees (cf. for example Figure 6.2).

Definition 6.2 (in-degree, out-degree, out-point, terminal vertex)
The in-degree of a vertex is the number of its incoming arcs, i.e. the number

of directed edges pointing to the vertex. Equivalently, the out-degree of a vertex is the number of its outgoing arcs, i.e. the number of directed edges that start in this vertex. An out-point is a vertex, that has no incoming arcs, i.e. in-degree 0. Vertices that have out-degree 0 are called terminal vertices.

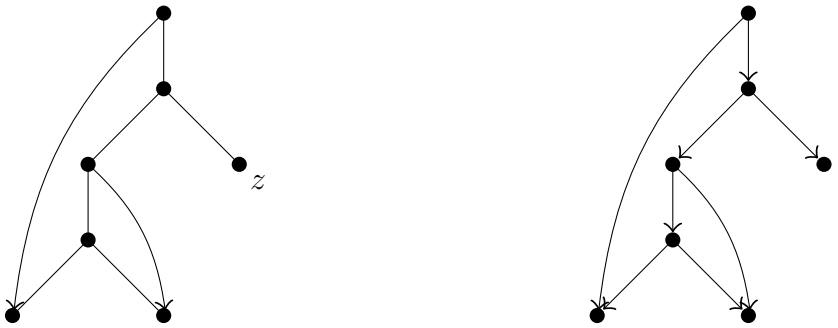


Figure 6.1: The lambda-tree representing the term $\lambda x.(\lambda y.(xy))z$ and the corresponding unlabeled DAG.

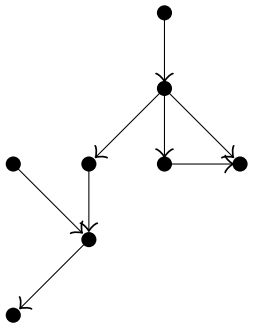


Figure 6.2: An unlabeled DAG, that can not be regarded as the enriched tree of some lambda-term.

Definition 6.3 (in-, out-neighbourhood)
The in-neighbourhood of a vertex v, is the set of all vertices that have outgoing arcs leading to the vertex v. Analogously the set of all vertices that are reached by arcs starting in v is called the out-neighbourhood of v.

We can give necessary and sufficient conditions, which all have to be fulfilled for a DAG to correspond to a lambda-tree:

- The graph is weakly connected.
  $\rightarrow$ Obvious.

- There is only one out-point.
  $\rightarrow$ The out-point of the DAG obviously corresponds to the unique root in the lambda-tree.

- Every vertex with out-degree greater than 0 (except for the unique out-point) has in-degree 1.
  $\rightarrow$ This condition prevents bindings of inner nodes.

- Every terminal vertex has in-degree at most 2.
  $\rightarrow$ This condition prevents that a variable gets bound by more than one unary node.

- For a vertex v with in-degree 2, let p_1 and p_2 be the predecessors of v. Then there has to be a path leading from p_1 to p_2 or from p_2 to p_1.
  $\rightarrow$ This condition ensures that a lambda only binds variables that are in its scope.

- For a vertex v with out-degree $k \geq 3$, $k-1$ of the out-going edges must point to terminal vertices and for all of these vertices there has to be another path connecting them with the vertex v.
  $\rightarrow$ This condition ensures that if we remove all the pointers, the underlying tree is indeed a unary-binary tree and there is no vertex that has more than two children.

So, we see, that a DAG has to fulfill a lot of constrictions in order to correspond to a lambda-tree, and therefore (as we will see in the following chapters) the number of DAGs exceeds the number of lambda-terms by far.

In contrast to the varying size of lambda-terms the size of a DAG is consistently defined as follows:

Definition 6.4 (size of a DAG)
The size of a DAG G is defined as the total number of its vertices, and is denoted by $|G|$.

In the following, when we talk about "labeled DAGs", we mean DAGs with labeled nodes and unlabeled edges, i.e. for a labeled DAG of size n we take $\{1, 2, \ldots, n\}$ to be the vertex set.

Definition 6.5 (isomorphic)
Two DAGs G and H of size n are isomorphic, if there is a permutation $\pi : \{1, 2, \ldots, n\} \rightarrow \{1, 2, \ldots, n\}$ acting on the point sets of the graphs which preserves the arcs, i.e. every edge $\langle u, v \rangle$ is in G if and only if the edge $\langle \pi(u), \pi(v) \rangle$ is in H.

We will consider two isomorphic DAGs to be equal, just like we did for lambda-terms.

Chapter 7

Counting DAGs

Now we are going to investigate the number of DAGs of a given size. We will discuss both the labeled and the unlabeled case, starting with the easier one, namely the problem of counting labeled DAGs.

7.1 Labeled DAGs

In the first part of this section we will set up a recursion for the number of labeled DAGs of size n and establish an interesting representation for the generating function of this number. The second part deals with another approach, namely we are going to show that there is a bijection between the number of labeled DAGs of size n and the number of $n \times n$ (0,1)-matrices with positive eigenvalues.

7.1.1 Recurrence and Generating Function

From now on let a_n be the number of labeled DAGs of size n.

Our goal is to set up a recursion for a_n. This has independently been done by Robinson in [28] and Stanley in [32], but we will closely stick to Rodionov's

work [30], which is based on the aforementioned ones.

First we introduce the generating functions $A_n(x) = \sum_{r=0}^{\infty} a_{n,r} x^r$, which are polynomials in x with the coefficients $a_{n,r}$ equal to the number of labeled DAGs of size n having r arcs.

Obviously, it holds that $a_n = A_n(1)$.

Theorem 7.1 ([30, Thm. 1])
The polynomials $A_n(x)$ when $n \geq 0$ satisfy the recursion

$$\sum_{m=0}^{n} (-1)^m \binom{n}{m} (1+x)^{m(n-m)} A_m(x) = \delta_{n0},$$

where δ_{n0} is Kronecker's symbol.

Proof. For $n = 0$, the formula is true by definition, since $A_0(x) = 1$. For $n \geq 1$ set $N = \{1, \ldots, n\}$ and for any non-empty subset $\alpha \subseteq N$ let $A_{n,r}^{\alpha}$ (with $r \geq 0$) denote the set of all labeled DAGs of size n with r arcs and all points v_i, with $i \in \alpha$, being out-points. Clearly, the inclusion $\emptyset \neq \alpha \subseteq \beta \subseteq N$ implies $A_{n,r}^{\beta} \subseteq A_{n,r}^{\alpha}$. Thus, using the inclusion-exclusion principle, we get

$$a_{n,r} = \sum_{\emptyset \neq \alpha \subseteq N} (-1)^{|\alpha|-1} |A_{n,r}^{\alpha}|, \tag{7.1}$$

with $|\alpha|$ denoting the number of elements in α.

Next we have to consider that the arcs of any graph $G \in A_{n,r}^{\alpha}$ can be divided into two classes:

- The first class consists of the directed edges $\langle v_i, v_j \rangle$, where $i, j \in N \setminus \alpha$.
- The second class is built of the arcs going from v_i, $i \in \alpha$, to v_j, $j \in N \setminus \alpha$.

Note that, if there are k arcs in the first class, then there are $r-k$ arcs in the second one. Therefore the following formula holds for α with $|\alpha| = m \geq 1$:

$$|A^{\alpha}_{n,r}| = \sum_{k=0}^{r} \binom{m(n-m)}{r-k} a_{n-m,k}. \tag{7.2}$$

Using (7.1) and (7.2) we get the following equality:

$$a_{n,r} = \sum_{m=1}^{n} (-1)^{m-1} \binom{n}{m} \sum_{k=0}^{r} \binom{m(n-m)}{r-k} a_{n-m,k}$$

Thus, we get

$$\begin{aligned} A_n(x) &= \sum_{m=1}^{n} (-1)^{m-1} \binom{n}{m} \sum_{r=0}^{\infty} \sum_{k=0}^{r} \binom{m(n-m)}{r-k} a_{n-m,k} x^r \\ &= \sum_{m=1}^{n} (-1)^{m-1} \binom{n}{m} \sum_{k=0}^{\infty} \left[\sum_{r=k}^{\infty} \binom{m(n-m)}{r-k} x^{r-k} \right] a_{n-m,k} x^k \\ &= \sum_{m=1}^{n} (-1)^{m-1} \binom{n}{m} (1+x)^{m(n-m)} A_{n-m}(x). \end{aligned}$$

□

By using $a_n = A_n(1)$, we get the following result:

Theorem 7.2 ([28, Equ. (13)])
The numbers a_n of labeled DAGs of size n satisfy the recursion

$$a_n = \sum_{m=1}^{n} (-1)^{m-1} \binom{n}{m} 2^{m(n-m)} a_{n-m},$$

with $a_1 = 1$.

In order to get a better combinatorial understanding of this equation (which will be particularly important in Chapter 10), we will give a brief interpretation of the occurring factors (see [28], or [16, p. 1415]):

The basic idea for setting up the recursion is to consider a DAG as an extension of another DAG of lower size by adding some additional out-points. Obviously the resulting graph is again a DAG. Now for example, consider we add one additional out-point to a DAG of size $n-1$, then the number a_{n-1} has to be multiplied by $2^{(n-1)(1)}$, which is the number of possibilities to connect the $n-1$ pre-existing vertices to the single n-th vertex, and by $\binom{n}{1}$, namely the number of possible ways to label the n-th vertex. However, thereby those new DAGs having two out-points (where one out-point already existed) have been counted twice, so we have to subtract these double-counts. But this in turn removes the number of all DAGs with three out-points, which therefore must be added back in, and so on. Using the inclusion-exclusion principle the recursive formula for a_n follows.

The first values of a_n are ([28, Table 1]):

1, 1, 3, 25, 543, 29281, 3781503, 1138779265, 783702329343, ...

Now we introduce the so-called "special generating function", as it has been done by Robinson in [28]:

$$A(t) = \sum_{n=0}^{\infty} \frac{a_n}{n!} 2^{-\binom{n}{2}} t^n$$

The big advantage of this function is that it adopts a very nice representation, due to which the asymptotic behavior of the coefficients a_n can easily be determined, as we will see in the next chapter.

Theorem 7.3 ([30, Thm. 2])
The special generating function $A(t)$ is given by the formula

$$A(t) = \left(\sum_{m=0}^{\infty} \frac{(-1)^m}{m!} 2^{-\binom{m}{2}} t^m \right)^{-1}. \tag{7.3}$$

Proof. We follow the proof of [28] and [30], while Stanley in [32] discovered this equation independently by means of the chromatic polynomial. (He used the fact that if χ_G is the chromatic polynomial of a labeled symmetric graph G on p points, then $(-1)^p\chi_G(-1)$ is the number of acyclic orientations of G.) Define

$$B(t) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m!} 2^{-\binom{m}{2}} t^m.$$

By formally multiplying the series $A(t)$ and $B(t)$ we get

$$A(t)B(t) = \sum_{n=0}^{\infty} \left(\sum_{k+m=n} \frac{a_k(-1)^m}{k!m!} 2^{-\binom{k}{2}-\binom{m}{2}} \right) t^n$$

$$= \sum_{n=0}^{\infty} \underbrace{\left[\sum_{k=0}^{n} (-1)^k \binom{n}{k} 2^{k(n-k)} a_k \right]}_{=\delta_{n0} \text{ by Theorem 7.1}} \frac{(-1)^n}{n!} 2^{-\binom{n}{2}} t^n \equiv 1$$

It follows that Equation 7.3 holds in any neighbourhood of zero in which the function $B(t)$ has no roots. □

7.1.2 A Bijection

As stated in the beginning of Section 7.1, the aim of this subsection is to show that the number of DAGs of size n is equal to the number of $n \times n$ $(0,1)$-matrices whose eigenvalues are positive real numbers.

Eric W. Weisstein ([25]) came up with the idea that there could be a relation between those two numbers, when he computed the numbers of real $n \times n$ $(0,1)$-matrices that have positive eigenvalues and observed that the resulting sequence of values coincided with the sequence A003024 in [31], which counts labeled DAGs of size n. McKay et al. ([25]) accomplished to prove the fact that there actually is a one-to-one correspondence between those numbers, which we will now present.

Theorem 7.4 ([25], Thm. 1.1)
For each $n = 1, 2, 3 \ldots$, the number of labeled DAGs of size n is equal to the number of real $n \times n$ matrices of zeros and ones whose eigenvalues are positive real numbers.

Remark 7.5
The eigenvalues of these matrices are actually all equal to one, as the proof of Thm. 7.4 will show.

Proof. Consider a labeled DAG G and let $A = A(G)$ be its vertex adjacency matrix. Obviously A has only zeros on the diagonal, because otherwise there would be cycles of length 1. Now we define $B = I + A$, where I denotes the identity matrix. Note that the resulting matrix B is again a (0,1)-matrix.

- First we show that B has only positive eigenvalues:
  Note that the eigenvalues will not change if the vertices of the graph G are renumbered in accordance with the partial order that it generates. Since then no arc would end at a vertex that has a smaller index than than its start node, $A = A(G)$ would be strictly upper triangular, and B would be upper triangular with 1's on the diagonal. Thus, all eigenvalues of B are equal to 1.

- Now we will show the reverse direction. Let B be a $n \times n$ (0,1)-matrix whose eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ are all positive real numbers. Then we have

$$
\begin{aligned}
1 &\geq \frac{1}{n}\mathrm{trace(B)} \qquad \text{(since all } B_{i,i} \leq 1) \\
&= \frac{1}{n}(\lambda_1 + \lambda_2 + \ldots + \lambda_n) \\
&\geq (\lambda_1 \lambda_2 \ldots \lambda_n)^{\frac{1}{n}} \qquad \text{(by the arithmetic-geometric mean inequality)} \\
&= (\det(B))^{\frac{1}{n}} \\
&\geq 1 \qquad\qquad\qquad \text{(since } \det(B) \text{ is a positive integer).}
\end{aligned}
$$

It follows that the arithmetic and geometric means of the eigenvalues are equal, which implies that the eigenvalues are all equal, namely all $\lambda_i(B) = 1$. Thus, B has only ones on the diagonal.
Now we can regard B as the adjacency matrix of a directed graph H, having a loop at each vertex. Since

$$\mathrm{trace(B^k)} = \sum_{\mathrm{i=1}}^{\mathrm{n}} \lambda_{\mathrm{i}}^{\mathrm{k}} = \sum_{\mathrm{i=1}}^{\mathrm{n}} 1 = \mathrm{n},$$

for all k, the number of closed walks of length k in H is n.
Since $\mathrm{trace(B^k)} = \mathrm{n}$, there are only ones on the diagonal of B^k and therefore it follows that in H there is exactly one closed walk of length k from every vertex to itself. Thus, there are no other closed walks than the loops at each vertex.
By putting $A = B - I$, we get that in A there are no closed walks of any length, which implies that A is a (0,1)-matrix representing the adjacency matrix of a labeled DAG of size n.

□

7.2 Unlabeled DAGs

In order to derive a recursion for the number of unlabeled DAGs of a given size, we follow the work of Robinson ([29]).
We denote by A_n the number of unlabeled DAGs of size n and by S_n the symmetric group of all $n!$ permutations of the point set $\{1, 2, \ldots, n\}$.

We consider a permutation $g \in S_n$ to act on the vertex set of any labeled DAG of size n by relabeling its vertices according to g. For example, if $g = (1)(2\ 3)$ then $g(\alpha) = \alpha$, $g(\beta) = \beta$, $g(\gamma) = \delta$ and $g(\delta) = \gamma$, with α, β, γ and δ as depicted in Figure 7.1.

Since two DAGs are isomorphic if and only if they can be mapped onto each other by a member of S_n, the number of orbits of all labeled DAGs of size n under S_n counts precisely the number A_n of unlabeled DAGs of size n.

Therefore we use the following lemma:

Lemma 7.6 (Burnside's Lemma, [12, p.249, Thm. B])
If G is a finite group represented as permutations on a finite object set, then

$$|\textit{orbits of } G| = \frac{1}{|G|} \sum_{g \in G} |\textit{fixed points of } g|.$$

Applying the Burnside's Lemma to S_n, acting on labeled DAGs of size n, we get

$$A_n = \frac{1}{n!} \sum_{g \in S_n} \underbrace{|\alpha \text{ such that } g(\alpha) = \alpha|}_{=:N(g)}.$$

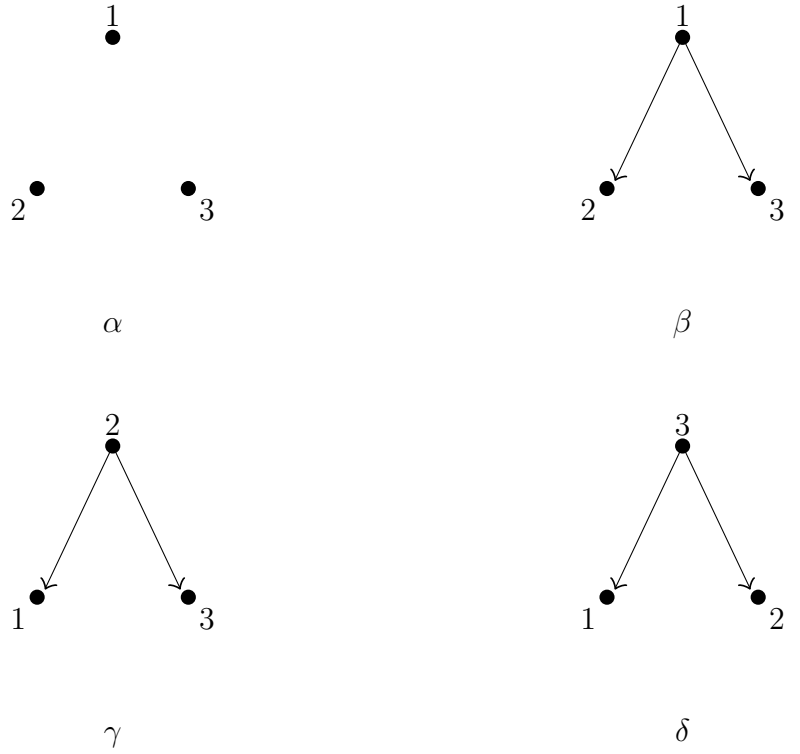


Figure 7.1: Four labeled DAGs of size 3 ([29, Fig. 1]).

Thus, for any $g \in S_n$ we have to find the number $N(g)$ of labeled DAGs α of size n such that $g(\alpha) = \alpha$.

We start by setting up a recursion for $N(g)$ using the inclusion-exclusion principle:

Let $g \in S_n$ be the product of the disjoint cycles $\gamma_1, \gamma_2, \ldots, \gamma_m$ of length $l(1)$, $l(2), \ldots, l(m)$. The condition $g(\alpha) = \alpha$ implies that the vertices in any cycle γ_i $(1 \leq i \leq m)$, must either all be out-points of α or all be non-out-points of α.

Now we define Q_i $(1 \leq i \leq m)$ to be the set of labeled DAGs of size n, such that $g(\alpha) = \alpha$ (for the fixed permutation g) with all points of γ_i being out-points of α.
Since every DAG of size $n > 0$ has at least one out-point, there must be a cycle that consists of out-points of α only. Therefore there is no labeled DAG fixed by g, that is not in any of the sets $Q_1, \ldots, Q_m$.

The inclusion-exclusion principle yields

$$0 = N(g) - \sum_{1 \leq i \leq m} |Q_i| + \sum_{1 \leq i < j \leq m} |Q_i \cap Q_j| \mp \ldots,$$

which can be written as follows:

$$N(g) = - \sum_{\emptyset \neq I \subseteq \{1,\ldots,m\}} (-1)^{|I|} |\bigcap_{k \in I} Q_k|$$

Thus, for any non-empty subset I of $\{1, \ldots, m\}$ we have to determine the number $|\bigcap_{k \in I} Q_k|$, which equals the number of labeled DAGs fixed by g, where all the cycles γ_k with $k \in I$ contain only out-points.

First take a look at the cycles γ_i and γ_j with $i \in I$ and $j \notin I$. If there is an arc $\langle x, y \rangle$ from a vertex $x \in \gamma_i$ to $y \in \gamma_j$ in a DAG α fixed by g, then α obviously also contains the arcs $\langle g(x), g(y) \rangle$, $\langle g^2(x), g^2(y) \rangle, \ldots$, which are just $\langle \gamma_i(x), \gamma_j(y) \rangle$, $\langle \gamma_i^2(x), \gamma_j^2(y) \rangle$, $\ldots$. Thus, we first arrive at $\langle \gamma_i^p(x), \gamma_j^p(y) \rangle = \langle x, y \rangle$ for $p = lcm(l(i), l(j))$, which denotes the least common multiple of $l(i)$ and $l(j)$. Since in total there are $l(i)l(j)$ arcs from γ_i to γ_j, they fall into exactly $\gcd(l(i), l(j))$ different cycles, which stands for the greatest common divisor of the point cycle lengths.
Those cycles of arcs must either all be contained in α or none of them is in α, which results in $2^{\gcd(l(i),l(j))}$ possibilities for arcs between γ_i and γ_j. Multiplying together all the independent possibilities for the choices of i and j, gives

$$2^{\sum_{i \in I} \sum_{j \notin I} \gcd(l(i),l(j))}$$

different configurations of arcs leading from out-points to the remaining vertices in a DAG fixed by g.

Now we independently investigate the sub-graph induced by the points of the cycles γ_j for $j \notin I$, which could be any DAG fixed by $\prod_{j\notin I}\gamma_j$.

Therefore we get $|\bigcap_{k\in I} Q_k| = 2^{\sum_{i\in I}\sum_{j\notin I}\gcd(l(i),l(j))} N(\prod_{j\notin I}\gamma_j)$, which yields the following recursion for $N(g)$:

$$N(g) = -\sum_{\emptyset\neq I\subseteq\{1,\ldots,m\}} (-1)^{|I|} 2^{\sum_{i\in I}\sum_{j\notin I}\gcd(l(i),l(j))} N(\prod_{j\notin I}\gamma_j). \qquad (7.4)$$

We can see by induction on the number n of vertices, that $N(g)$ only depends on the cycle length in the representation of g as disjoint cycles. Therefore we can put Equation (7.4) in terms of cycle types $Z(g) = \prod_{1\leq i\leq n} a_i^{\sigma_i}$, which are monomials in the commuting variables $a_1, a_2, \ldots$, where σ_i denotes the number of cycles of g of length i. Grouping the subsets I according to the number τ_i of cycles of length i contained in I, and by the fact that for fixed values of τ_i, there are $\prod_i \binom{\sigma_i}{\tau_i}$ such sets I with $|I| = \sum_i \tau_i$ for each subset, Equation (7.4) becomes

$$N(\prod_j a_j^{\sigma_j}) = -\sum_{0\leq\tau_i\leq\sigma_i} (-1)^{\sum_i \tau_i} N(\prod_j a_j^{\sigma_j-\tau_j}) 2^{\sum_{i,j}\tau_i(\sigma_j-\tau_i)\gcd(i,j)} \prod_i \binom{\sigma_i}{\tau_i},$$

where the term for $\tau_1 = \tau_2 = \ldots = \tau_n = 0$ is omitted in the summation. We have the initial value $N(1) = 1$ representing the case in which all the points are out-points, which includes just the single DAG with $n = \sum_i i\sigma_i$ points and no arcs.

Since for the number of unlabeled DAGs of size n we have $A_n = \frac{1}{n!}\sum_{g\in S_n} N(g)$, this can be rewritten in terms of cycle types as

$$A_n = \sum N(\prod_i a_i^{\sigma_i}) / \prod_i \sigma_i! i^{\sigma_i},$$

where we sum over all sequences $\sigma_1, \sigma_2, \ldots, \sigma_n$ of non-negative integers such that $n = \sum_i i\sigma_i$. In order to prove this, Robinson ([29]) showed that the number of elements g of S_n with $Z(g) = \prod_i a_i^{\sigma_i}$ is exactly $n!/\prod_i \sigma_i! i^{\sigma_i}$, for any such sequence:
There are $n!$ assignments of the numbers $\{1, 2, \ldots, n\}$ to a fixed sequence of cycles with σ_i of length i for all $1 \leq i \leq n$, but any two of them that represent the same permutation are considered to be equal. This explains the denominator $\prod_i \sigma_i! i^{\sigma_i}$, since equivalent assignments are generated by swapping the entire cycles of length i among themselves, which can be done in $\sigma_i!$ different ways for each i, and by rotating the assignment for each cycle of length i in any of the i possible ways keeping the cyclic order fixed, which gives i^{σ_i} different possibilities.

By determining $N(\prod_i a_i^{\sigma_i})$ with the recursion obtained above, we can estimate the numbers A_n recursively.

The first values of A_n are the following ([29, Table 2]):

1, 1, 2, 6, 31, 302, 5984, 243668, 20286025, 3424938010, 1165948612902, ...

Remark 7.7
In [29] Robinson established a more efficient way to determine the numbers A_n, by using the generating function $Z(A)$ in $a_1, a_2, a_3, \ldots$, which is given by

$$Z(A) = \sum_{\sigma_i \geq 0} N(\prod_i a_i^{\sigma_i}) \prod_i a_i^{\sigma_i} / \prod_i \sigma_i! i^{\sigma_i}.$$

A_n can then be found by simply summing the coefficients of all the monomials $\prod_i a_i^{\sigma_i}$ with $n = \sum_i i\sigma_i$.
The big advantage lies in introducing the operation $*$ for monomials defined by

$$\left(\prod_i a_i^{\tau_i}\right) * \left(\prod_j a_j^{\nu_j}\right) = 2^{\sum_{i,j} \tau_i \nu_j \gcd(i,j)} \prod_i a_i^{\tau_i} \prod_j a_j^{\nu_j},$$

with which we get (see [29, p.35])

$$e^{-\sum_{i\geq 1} a_i/i} * Z(A) = 1.$$

Since the polynomial $Z_{\leq n}(A)$ consisting of the terms of $Z(A)$ with $\sum_i i\sigma_i \leq n$ is all that contributes to the terms $Z_{n+1}(A)$ with $\sum_i i\sigma_i = n+1$, we can start with $Z_0(A) = 1$ and find $Z_n(A)$ for successive values n by taking advantage of the distributive property of $*$ over $+$ (see [29, pp.35] for more details).

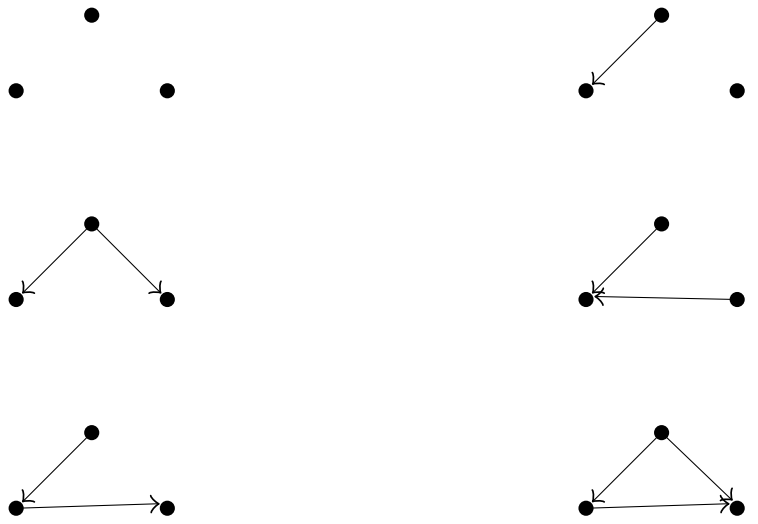


Figure 7.2: The six unlabeled DAGs of size 3 ([29, Fig. 3]).

Chapter 8

Asymptotics

As mentioned before, it is easy to establish the asymptotic behavior of the number of labeled DAGs due to the explicit representation for their generating function (cf. Theorem 7.3). This result will be presented in Section 8.1. Unfortunately we do not have such a representation for the generating function of unlabeled DAGs. In order to get some results on their asymptotics we fix the number of edges in the DAG. In Section 8.2 we investigate the asymptotic behavior of labeled DAGs with a fixed number of edges. Finally, in the last section of this chapter, Section 8.3, we derive results on the asymptotics of unlabeled DAGs with a prescribed number of edges.

8.1 Asymptotic behavior of the number of labeled DAGs

In order to obtain the asymptotic behavior of the sequence a_n enumerating labeled DAGs of size n, we analyse the special generating function $A(t)$ and therefore use its specific representation

$$A(t) = \frac{1}{B(t)},$$

where

$$B(t) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m!} 2^{-\binom{m}{2}} t^m, \tag{8.1}$$

which we derived in Chapter 7 (cf. Theorem 7.3). Note that $B(t)$ is an entire function. From [28] it is known that $B(t)$ has the zeros $\omega_0 \approx 1.488079$ and $\omega_1 \approx 4.881141$ ($\omega_3 \approx 13.56$, $\omega_4 \approx 85$ - furthermore the authors of [28] conjectured that $\omega_i = (i+1)2^i + o(2^i)$).
Therefore $A(t)$ is meromorphic inside the circle of radius ω_1, with a single pole at ω_0. Thus we can apply singularity analysis - more precisely we can apply Theorem 2.8, introduced in Chapter 2, which gives the following result (see [32],[35],[22]):

Theorem 8.1
The asymptotic behavior of the numbers a_n of labeled DAGs of size n is

$$a_n \sim \lambda n! 2^{\binom{n}{2}} \omega^{-n} \quad as \quad n \to \infty, \tag{8.2}$$

where $\omega = \omega_0 = 1.4880785455997102947\ldots$ is the least root of $B(t)$, given in Equation 8.1 and $\lambda = -\frac{1}{\omega B'(\omega)} = 1.7410611252932298403\ldots$.

Remark 8.2
Robinson ([28]) got the same result by developing and analyzing a convergent series for a_n, namely

$$a_n = n! 2^{\binom{n}{2}} \sum_{j=0}^{\infty} \frac{1}{\omega_j^{n+1} B(\omega_j/2)},$$

where $\omega_0, \omega_1, \omega_2, \ldots$ denote the zeros of B with increasing modulus.

8.2 Asymptotic behavior of the number of labeled DAGs with a fixed number of edges

In [5] the authors extended Equation (8.2) to $a_{n,q}$, which denotes the number of labeled DAGs of size n with q unlabeled edges. By denoting $N = \binom{n}{2}$ and $B(x,y) = \sum_{n=0}^{\infty} \frac{(-1)^n x^n}{n!(1+y)^N}$, they obtained the following result:

Theorem 8.3 ([5, Thm. 1])
Let $\epsilon > 0$ be given and suppose $q = q(n)$ satisfies $\epsilon N \leq q \leq (1-\epsilon)N$ for all large n. Then

$$a_{n,q} \sim n! \binom{N}{q} \frac{e^{-x^2 r}}{B(\nu\omega, r)\omega^{n+1}} \tag{8.3}$$

where

$$r = \frac{q}{N-q}, \quad \nu = \frac{N-q}{N}, \quad x = \frac{\nu\omega B(\nu^2\omega, r)}{2B(\nu\omega, r)}$$

and $\omega = \omega(r) > 0$ is the smallest solution of the equation $B(\omega, r) = 0$.

In order to prove this theorem we need the following two auxiliary lemmas (Lemma 8.4, Lemma 8.6):

Lemma 8.4 (Central limit theorem, [5, Lem. 2])
Let $X_n(r)$ be a random variable with $Pr\{X_n(r) = m\}$ proportional to $a_{n,m} r^m$ and suppose that $q = q(n)$ with $0 < \epsilon < q/N < 1 - \epsilon$, $r = \frac{q}{N-q}$ and $\omega(r)$ is as in Theorem 8.3. Then there is a function $K = K(n) = o(n)$ such that as $n \to \infty$

$$\frac{1}{K} \sum_{m=q}^{q+K-1} Pr\{X_n(r) = m\} \sim \frac{e^{-v^2/2}}{\sqrt{2\pi\sigma_n(r)^2}},$$

where

$$v = \frac{nr\omega'(r)}{\sigma_n(r)\omega_(r)}, \quad \sigma_n(r)^2 = \frac{Nr}{(1+r)^2}$$

and the rate of convergence depends on ϵ.

Idea of proof. Let

$$\mu_n(r) = \frac{Nr}{1+r} - \frac{nr\omega'(r)}{\omega(r)}.$$

The idea of the proof is to show that the characteristic function of

$$Y_n(r) = \frac{X_n(r) - \mu_n(r)}{\sigma_n(r)}$$

converges pointwise to $e^{-t^2/2}$ and then using the continuity theorem ([20, p.52, Thm. 2]) to get a central limit theorem for $Y_n(r)$.
Let us define $A_n(y) = \sum_{q=0}^{N} a_{n,q} y^q$.
The characteristic function of X_n is given by

$$\mathbb{E}e^{itX_n} = \sum_m Pr\{X_n(r) = m\} \cdot e^{itm} = c \cdot A_n(re^{it}).$$

Therefore by setting $g_n(t) = \log A_n(re^{it})$, the logarithm of the characteristic function of X_n is

$$g_n(t) - g_n(0) = g_n'(0)t + g_n''(0)\frac{t^2}{2} + O(Mt^3),$$

with M being the maximum modulus of $g_n^{(3)}(u)$ for $0 \le u \le t$.
The derivatives $g_n'(t), g_n''(t)$ and $g_n^{(3)}(t)$ can be calculated by means of Lemma 1 of [5], and by evaluating them at $t = 0$ we get

$$g_n(t) - g_n(0) = i\mu_n(r)t - \sigma_n^2(r)\frac{t^2}{2} + o(1),$$

for $t = o(n^{12/3})$.
Now we derive for the characteristic function of Y_n:

$$\mathbb{E}e^{itY_n} = \mathbb{E}e^{\frac{X_n(r)-\mu_n(r)}{\sigma_n(r)}} = e^{g_n\left(\frac{t}{\sigma_n(r)}\right) - \frac{i\mu_n(r)t}{\sigma_n(r)} - g(0)} \sim e^{-t^2/2},$$

for $t = o(n^{-2/3\sigma_n(r)})$. □

This central limit theorem enables us to formulate a corollary, which gives a very interesting result on the structure on DAGs:

Corollary 8.5
The average number of edges in a labeled DAG of size n is $\sim \frac{n^2}{4}$.

This follows immediately by calculating

$$\lim_{r\to 1} \mu_n(r) = \lim_{r\to 1} \left(\frac{Nr}{1+r} - \frac{nr\omega'(r)}{\omega(r)} \right) \sim \frac{\binom{n}{2}}{2} \sim \frac{n^2}{4}.$$

In Chapter 9 we will focus on more results on the structure of DAGs. Now we return to the proof of Theorem 8.3, for which we need the second auxiliary lemma given below, which is a more combinatorial one:

Lemma 8.6 ([5, Lemma 3])
Let $K(n) = o(n)$ and $\epsilon > 0$ be given. If $n, q \to \infty$ in such a way that $\epsilon < q/N < 1 - \epsilon$, then

$$a_{n,q+k} \sim \binom{N-q}{q}^k a_{n,q}$$

uniformly for $|k| < K(n)$ at a rate depending on $K(n)$ and ϵ.

Proof. Let P_n be the set of all ordered partitions of $\{1, 2, \ldots, n\}$. For $\pi \in P_n$ let $|\pi| = s$ denote the number of blocks of π and let $B_i = B_i(\pi)$, for $i = 1, \ldots, s$, be the blocks of π, with $b_i = |B_i|$. Additionally we will use the following notations:

$$g(\pi) = \sum_{i=1}^{s} \binom{b_i}{2} \quad \text{and} \quad h(\pi) = g(\pi) + \sum_{i=1}^{s-1} b_i b_{i+1}.$$

Now we have to introduce some further definitions that will be necessary for the proof:
Let D be a labeled DAG with n vertices. First we define an ordered partition of vertices of D inductively as follows:
$D_1 := D$, B_i is the set of out-points of D_i and $D_{i+1} := D_i - B_i$. Then the so-called tower $T(D)$ of D contains all the vertices of D with the arcs going from $B_i(\pi(D))$ to $B_{i+1}(\pi(D))$ for $i \geq 1$. In the following we denote by $q(T)$ the number of arcs in $T = T(D)$.

The proof consists of two parts.

- In the first part we show that a linear bound on $g(T)$ leads to linear bounds on $h(T)$ and $q(T)$. More precisely if T is the tower of a labeled DAG of size n with $g(T) \leq C_1 n$, it follows that $q(T) < (2C_1 + 1)n$ and $h(T) < (3C_1 + 1)n$:
  This can be shown by using the arithmetic-geometric mean inequality for b_i^2 and b_{i+1}^2, which yields

$$b_i b_{i+1} \leq \binom{b_i}{2} + \frac{1}{2} b_i + \binom{b_{i+1}}{2} b_{i+1}.$$

  Thus, we get

$$q(T) \leq h(T) - g(T) = \sum_{i=1}^{s-1} b_i b_{i+1} < 2 \underbrace{\sum_{i=1}^{s} \binom{b_i}{2}}_{\leq C_1 n} + n \leq (2C_1 + 1)n,$$

  and

$$h(T) = \underbrace{g(T)}_{\leq C_1 n} + \underbrace{\sum_{i=1}^{s-1} b_i b_{i+1}}_{\leq (2C_1+1)n} \leq (3C_1 + 1)n.$$

- In the second part we prove that most DAGs have small $g(D)$: Let us suppose that $0 < \epsilon < \frac{m}{N} < 1 - \epsilon$ for all $n > n_0$. Then there is a

$C_2 = C_2(\epsilon, n_0)$ such that the fraction of labeled DAGs of size n with m arcs and $g(D) > C_2 n$ is less than $(1-\epsilon)^n$: The number of DAGs with $g(D) = j$ is at most

$$|P_n| \binom{N-j}{m} < n^n \binom{N-j}{m}.$$

For all $b_i = 1$ we get

$$n! \binom{N-n+1}{m-n+1} > \left(\frac{n}{e}\right)^n \left(\frac{m-n}{N}\right)^n \binom{N}{m} > \left(\frac{nm}{3N}\right)^n \binom{N}{m}.$$

Thus for $n > n_0$ the fraction is less than

$$\left(\frac{3N}{m}\right)^n \sum_{j > C_2 n} \left(\frac{N-j}{N}\right)^m \leq C_3 \left(\frac{3}{e}\right)^n e^{-\epsilon C_2 n}, \tag{8.4}$$

where $C_3(\epsilon) > 0$. For sufficiently large C_2 this is less than $(1-\epsilon)^n$.

Therefore we get the following result:

$$a_{n,m} = \sum_T \binom{N-h(T)}{m-q(T)} \sim \sum_{g(T) \leq C_2 n} \binom{N-h(T)}{m-q(T)}.$$

Finally we have to show that

$$\binom{N-h(T)}{q+k-q(T)} \sim \left(\frac{N-q}{q}\right)^k \binom{N-h(T)}{q-q(T)}$$

for $g(T) \leq C_2 n$. This can be proved by means of Stirling's formula, the bounds for $h(T)$ and $q(T)$, which have been derived in the first part of the proof, and $log(1+u) = u + O(u^2)$ (see [5, p.21] for the details). Putting all together completes the proof. □

Now we are able to prove Theorem 8.3 by combining our analytical and combinatorial results.

Proof of Theorem 8.3. Define $A_n(y) = \sum_{q=0}^{N} a_{n,q} y^q$ as in the proof of Theorem 8.4. By the Lemmas 8.4 and 8.6 we get

$$\frac{e^{-v^2/2}}{\sqrt{2\pi}\sigma_n(r)} A_n(r) \sim \frac{1}{K} \sum_{m=q}^{q+K-1} a_{n,m} r^m \sim \frac{1}{K} \sum_{m=q}^{q+K-1} a_{n,q} r^q = a_{n,q} r^q.$$

Using $A_n(r) \sim n!(1+r)^N \omega^{-n} C(r)$ with $C(r) = \frac{1}{\omega B(\nu\omega, r)}$ (see [5], p.16f) and

$$\frac{(1+r)^N r^{-q} C(r)}{\sqrt{2\pi}\sigma_n(r)} = \frac{N^N}{q^q (N-q)^{N-q} \omega B(\nu\omega, r)} \sqrt{\frac{(1+r)^2}{2\pi N r}} \sim \binom{N}{q} \frac{1}{\omega B(\nu\omega, r)},$$

we get

$$a_{n,q} \sim \frac{e^{-v^2/2}}{\sqrt{2\pi}\sigma_n(r)} \frac{n!(1+r)^N \omega^{-n} C(r)}{r^q} \sim e^{-v^2/2} n! \omega^{-n} \binom{N}{q} \frac{1}{\omega B(\nu\omega, r)}.$$

Noting that $v = -nx\sqrt{\frac{r}{N}} \sim -x\sqrt{2r}$ completes the proof. □

8.3 Comparison of the asymptotics of labeled and unlabeled DAGs with a fixed number of edges

In the previous section, we derived the asymptotic behavior of the numbers $a_{n,q}$ of labeled DAGs. The aim of this section is to use the results obtained for $a_{n,q}$ to get some information on the asymptotics of the numbers $A_{n,q}$, which count unlabeled DAGs of size n with q arcs. The results given in this section are based on [4].

Theorem 8.7 ([4, Thm. 2])
Let $\epsilon > 0$ be given and suppose $q = q(n)$ satisfies $\epsilon N < q < (1-\epsilon)N$ for all large n. Then

$$a_{n,q} \sim n! A_{n,q}.$$

The proof of this theorem is very technical, therefore we will only give the idea and the basic steps.

Sketch of proof. We will use some definitions and notations introduced in the proof of Lemma 8.6.
Let π be a permutation of $\{1, \ldots, n\}$ that acts on the vertices of a labeled DAG D by permuting the vertex labels and let $F(\pi, T)$ be the number of labeled DAGs of size n with q arcs with tower T (cf. Lemma 8.6 for the definition of the tower of a DAG) that are fixed by π. By Burnside's lemma (Lemma 7.6) we get

$$n! A_{n,q} = \sum_{\pi,T} F(\pi, T) = a_{n,q} + \sum_{T} \sum_{\pi \neq 1} F(\pi, T) = a_{n,q} + \sum_{1} + \sum_{2},$$

where $\sum_1$ is the sum of $F(\pi, T)$ over T and $\pi \neq 1$ with $g(T) < n^{1.2}$ and $\sum_2$ is the same sum with $g(T) \geq n^{1.2}$ (cf. Lemma 8.6 for the definition of $g(T)$). Our goal is now to show that both sums, $\sum_1$ and $\sum_2$, are equal to $o(a_{n,q})$.

- In order to get $\sum_2 = o(a_{n,q})$ we use the same arguments as in the second part of the proof of Lemma 8.6, which yields

  $$\sum_{g(T) \geq n^{1.2}} F(1, T) \leq a_{n,q} \left(\frac{3N}{q} \right)^n C e^{-n^{1.2} \frac{q}{N}}.$$

  (cf. Equation (8.4).)
  Since $\frac{q}{N} > \epsilon$ and $F(\pi, T) \leq F(1, T)$, we get

  $$\sum_{2} \leq n! \sum_{g(T) \geq n^{1.2}} F(1, T) = o(a_{n,q}).$$

- Now we are going to show that $\sum_1 = o(a_{n,q})$, which is more extensive than the first case - therefore we will skip some of the technical details. Bender and Robinson ([4]) used the ideas of Wright ([37]) to prove the desired result. Wright showed that $F(\pi, T)$ is the coefficient of x^Q in

$\prod(1+x^i)^{P_i}$, where $P_i = P_i(\pi, T)$ denotes the orbits of size i.

Thus for any positive real x it can be shown ([4, p. 365]) that,

$$F(\pi, T) \leq x^{-Q}(1+x)^E \left(\frac{1+x^2}{(1+x)^2}\right)^{(E-P_1)/2},$$

where $E = N - h(T)$ (cf. Lemma 8.6 for the definition of $h(T)$) are the potential edges from which $Q = q - q(T)$ must be chosen once T has been specified.
By setting $x = Q/(E-Q)$ and $\beta = (1+x^2)/(1+x)^2$ we get

$$F(\pi, T) \leq \frac{E^E \beta^{(E-P_1)/2}}{Q^Q (E-Q)^{E-Q}} \leq C\sqrt{N}\binom{E}{Q}\beta^{(E-P_1)/2}$$

$$\leq CnF(1, T)\beta^{(E-P_1)/2}. \tag{8.5}$$

The fact that $g(T) \leq n^{1.2}$ implies (by the first part of the proof of Lemma 8.6)

$$b_i \leq Cn^{0.6}, \quad q(T) \leq Cn^{1.2}, \quad h(T) \leq Cn^{1.2}.$$

Therefore, and since $\epsilon < \frac{q}{N} < 1 - \epsilon$, it follows that β is bounded above by $\beta(\epsilon) < 1$ for all T with $g(T) < n^{1.2}$. Now let π, which is acting on the set $\{1, \ldots, n\}$, have $n - a$ fixed points. Then it is easy to show that

$$E - P_1 \geq \frac{an}{4} \tag{8.6}$$

(see [4, p.366]).
Thus by (8.5) and (8.6) we get the following result:

$$\sum_1 \leq \sum_{\pi \neq 1} \sum_T CnF(1, T)\beta^{an/4}$$

$$
\begin{aligned}
&\leq Cn \sum_{T} F(1,T) \sum_{a \geq 1} n^a \beta^{an/4} \\
&= Cn a_{n,q} \sum_{a \geq 1} (n\beta^{n/4})^a \\
&\leq \tilde{C} n^2 a_{n,q} \beta^{n/4} = o(a_{n,q}).
\end{aligned}
$$

□

Chapter 9

Structure of DAGs

In this chapter we investigate some properties of the structure of DAGs. We already derived that the average number of edges in a labeled DAG of size n is $\sim \frac{n^2}{4}$ (cf. Corollary 8.5). In the first section of this chapter we are going to show that asymptotically almost all DAGs (unlabeled and labeled) are weakly connected ([4]). In the next section we will show that the number of out-points in labeled DAGs decreases with increasing size ([23]), e.g. from some n onwards, more than $99,7\%$ of all labeled DAGs of size n have less than 4 out-points. At last we present an interesting result, proved by McKay in [24], namely that the height of labeled DAGs of size n is asymptotically normally distributed with mean $C_1 n$ and variance $C_2 n$.

9.1 Weakly connected DAGs

As mentioned above, in this section we will show that asymptotically almost all DAGs are weakly connected. Therefore we follow the work of Bender and Robinson ([4]).
Let $c_{n,q}$ be the number of weakly connected labeled DAGs of size n with q arcs and $C_{n,q}$ shall stand for the number of unlabeled ones.

Theorem 9.1 ([4, Thm. 2])
Let $\epsilon > 0$ be given and suppose $q = q(n)$ satisfies $\epsilon N < q < (1-\epsilon)N$ for all large n, with $N = \binom{n}{2}$. Then

$$a_{n,q} \sim c_{n,q} \quad \textit{and} \quad A_{n,q} \sim C_{n,q}.$$

We will prove $c_{n,q} \sim a_{n,q}$ and then specify the small changes that are needed to prove $C_{n,q} \sim A_{n,q}$.

Sketch of proof. To construct an unlabeled DAG of size n with q arcs, we can take n linearly ordered points and connect them with q edges in such a way that each edge is oriented towards the lower point. Since this also creates some redundant constructions, we get

$$a_{n,q} \leq n! A_{n,q} \leq n! \binom{N}{q}. \tag{9.1}$$

It is easy to see that

$$a_{n,q} - c_{n,q} \leq \sum_{\substack{i,j \\ 1 \leq i \leq n/2}} T_{ij}, \tag{9.2}$$

where

$$T_{ij} = \binom{n}{i} a_{i,j} a_{n-i,q-j}.$$

The idea of the proof is to show that the right hand side of Equation (9.2) is $o(a_{n,q})$.

By Equation (9.1) and some further estimations concerning binomial coefficients (see [4, p. 367] for the detailed estimation) we get

$$\sum_j T_{ij} \leq n! \binom{N}{q} e^{-i(n-i)\epsilon}, \qquad \text{for } i \in I = [1; \frac{n}{2}] \text{ and } j \leq \binom{i}{2}.$$

Now we perform a distinction of cases:

- First case: $I = [C\ \log n, \frac{n}{2}]$

  We get

  $$\sum_{i\in I}\sum_{j} T_{ij} \leq n!\binom{N}{q}\sum_{i\in I} e^{-i(n-i)\epsilon} \leq n!\binom{N}{q}O(e^{-Cn\ \log n}). \qquad (9.3)$$

  Since $\epsilon < \nu < 1-\epsilon$ in Theorem 8.3 and r is bounded away from 0 and ∞, it follows that x, ω, and $B(\nu\omega, r)$ are bounded away from 0 and ∞. By (8.3) and (9.3), we get

  $$\sum_{i\in I}\sum_{j} T_{ij} = o(a_{n,q}).$$

- Second case: $I = [1, C\ \log n]$

  For $i \in I$ we have (by the definition of T_{ij})

  $$\frac{1}{a_{n,q}}\sum_{j} T_{ij} = \sum_{j}\binom{n}{i} a_{i,j}\frac{a_{n-i,q-j}}{a_{n,q}}.$$

  Using the following estimation (see [4, pp. 367] for the technical details)

  $$\binom{n}{i} a_{i,j}\frac{a_{n-i,q-j}}{a_{n,q}} \leq \left(1 - \frac{i(n-i)}{N}\right)^{q} e^{O(i^2)}, \qquad (9.4)$$

  we get, for sufficiently large n,

  $$\sum_{j} T_{ij} = \sum_{j\leq i^2} T_{ij} \qquad (\text{since } j \leq \binom{i}{2} \leq i^2)$$

$$\leq \left(1 - \frac{i(n-i)}{N}\right)^q e^{O(i^2)} a_{n,q} \qquad \text{(by 9.4)}$$
$$\leq exp\left(-\frac{i(n-i)q}{N} + O(i^2)\right) a_{n,q}$$
$$\leq e^{-Cin} a_{n,q}.$$

Thus,

$$\sum_{1 \leq i \leq C \ log \ n} \sum_j T_{ij} \leq a_{n,q} \sum_{i \geq 1} (e^{-Cn})^i = o(a_{n,q}).$$

Combining the results obtained for both cases yields $a_{n,q} - c_{n,q} = o(a_{n,q})$ and therefore completes the proof of $a_{n,q} \sim c_{n,q}$.

Let us now briefly consider the unlabeled case: By replacing all occurrences of $a_{*,*}$ by $A_{*,*}$ and omitting all references to $\binom{n}{i}$ and $n!$, all equations following (9.1) remain valid. The references to Theorem 8.3 still hold, since $A_{n,q} \sim \frac{a_{n,q}}{n!}$, as it has been proved in the previous chapter (cf. Theorem 8.7). □

9.2 Number of out-points

Let $a_n^{(k)}$ be the number of labeled DAGs of size n with exactly k out-points. Obviously $a_n^{(0)} = 0$ for $n > 0$, since there is no DAG without any out-points, and $\sum_k a_n^{(k)} = a_n$.

Let $A^{(k)}(z) = \sum_{n=k}^{\infty} a_n^{(k)} \frac{z^n}{n! 2^{\binom{n}{2}}}$ be the special generating function of the sequence $a_n^{(k)}$.

Now we are going to show that this generating function has also an explicit representation similar to $A(z)$ (cf. Thm. 7.3), due to which it is easy to determine the asymptotic behavior of its coefficients.

Theorem 9.2 ([23, Cor. 1])
The generating function $A^{(k)}(z)$ is given by the formula

$$A^{(k)}(z) = \frac{z^k}{k!2^{\binom{k}{2}}} \frac{B(2^{-k})}{B(z)},$$

where $B(z)$ is defined as in Thm. 7.3.

Proof. From [23] Prop. 1 we know that the number $a_n^{[k]}$ of labeled DAGs of size n, whose set of out-points is a fixed k-element subset of $\{1, \ldots, n\}$, satisfies the recursion

$$\sum_{t=k}^{n} \binom{n-k}{t-k} a_{n-t} 2^{(t-k)(n-t)} a_t^{[k]} = 2^{k(n-k)} a_{n-k}. \tag{9.5}$$

For every fixed k the right hand side of Equation (9.5) counts the number of labeled DAGs of size n with $\{1, \ldots, k\}$ being a subset of the set of all out-points. Now, considering how such DAGs can be obtained from DAGs of lower sizes, where the set of out-points is a fixed k-element subset of $\{1, \ldots, n\}$, yields the left hand side of the equation:
Suppose the subgraph S consisting of all vertices which are accessible from any vertex of $\{1, \ldots, k\}$ consists of t vertices. Then a_{n-t} enumerates all possible subgraphs on the remaining $n - t$ vertices and $2^{(t-k)(n-t)}$ enumerates all possible sets of arcs from these vertices to the vertices of the subgraph S (taking into account the absence of arcs which leads to the vertices of $\{1, \ldots, k\}$). Since there are $\binom{n-k}{t-k}$ possibilities for the choice of the t vertices of S the recursion follows.

Additionally, it obviously holds that

$$a_n^{(k)} = \binom{n}{k} a_n^{[k]}.$$

Therefore, and by setting $t = s + k$ and $n = m + k$, Equation (9.5) can be

rewritten as follows:

$$\sum_{s=0}^{m} \frac{m!}{s!(m-s)!} a_{m-s} 2^{s(m-s)} \frac{a_{k+s}^{(k)} k! s!}{(k+s)!} = 2^{km} a_m.$$

By using the identity

$$2^{\binom{i+j}{2}} = 2^{\binom{i}{2}} 2^{\binom{j}{2}} 2^{ij}$$

we get

$$2^{s(m-s)} = \frac{2^{\binom{m}{2}} 2^{\binom{k}{2}}}{2^{\binom{k+s}{2}} 2^{\binom{m-s}{2}}} 2^{ks}.$$

Therefore we have

$$\sum_{s=0}^{m} \frac{a_{m-s}}{(m-s)! 2^{\binom{m-s}{2}}} \frac{a_{k+s}^{(k)} 2^{ks}}{(k+s)! 2^{\binom{k+s}{2}}} = \frac{a_m 2^{km}}{k! 2^{\binom{k}{2}} m! 2^{\binom{m}{2}}},$$

which is equivalent to our desired result by using $A(z) = \frac{1}{B(z)}$ and comparing coefficients. □

Now we want to determine the asymptotic behavior of the numbers $a_n^{(k)}$. Remember that for the asymptotics of labeled DAGs of size n we got the following result (Theorem 8.1):

$$a_n \sim \lambda n! 2^{\binom{n}{2}} \omega^{-n} \quad \text{as} \quad n \to \infty,$$

where $\omega \approx 1.4880785$ denotes the least root of $B(t) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m!} 2^{-\binom{m}{2}} t^m$, and $\lambda = -\frac{1}{\omega B'(\omega)} \approx 1.74106$.

Theorem 9.3 ([23, Prop. 2])
As $n \to \infty$,

$$a_n^{(k)} \sim \frac{\omega^k B(2^{-k}\omega)}{k! 2^{\binom{k}{2}}} a_n,$$

where $k \geq 1$ is fixed.

Proof. Just as $A(z)$, $A^{(k)}(z)$ has radius of convergence ω and one singular point on the circle $|z| = \omega$, namely a simple pole at ω with residue $\frac{\omega^k}{k!2^{\binom{k}{2}}}\frac{B(2^{-k}\omega)}{B'(\omega)}$ (cf. Thm 9.2). By applying singularity analysis and using the asymptotics of a_n (Thm. 8.1) the theorem follows. □

Now we introduce a uniform probability distribution, and let ξ_n be the random variable corresponding to the number of out-points in a random labeled DAG of size n. The next Theorem follows as an immediate consequence of Theorem 9.3.

Theorem 9.4 ([23, Thm. 1])
As $n \to \infty$, for any fixed natural number k,

$$\boldsymbol{P}\{\xi_n = k\} \to \phi_k = \frac{\omega^k B(2^{-k}\omega)}{k!2^{\binom{k}{2}}}.$$

The limiting probabilities ϕ_k decrease very fast with increasing k. For the first values we have $\phi_1 = 0.5743\ldots, \phi_2 = 0.3662\ldots$ and $\phi_3 = 0.0564\ldots$. Thus it follows that for sufficiently large n, more than 99.7% of all labeled DAGs of size n have less than 4 out-points.

Let $\Phi(z) = \sum_{k=0}^{\infty} \phi_k z^k$ be the generating function of the limiting distribution.

Corollary 9.5 ([23, Cor. 3])
$\Phi(z) = B(\omega(1-z))$.

Proof.

$$\begin{aligned}\sum_{k=0}^{\infty}\frac{B(2^{-k}\omega)\omega^k}{k!2^{\binom{k}{2}}}z^k &= \sum_{k=0}^{\infty}\frac{z^k}{k!2^{\binom{k}{2}}}\sum_{n=0}^{\infty}\frac{(-1)^n\omega^{n+k}}{n!2^{\binom{n}{2}}2^{kn}}\\ &= \sum_{k=0}^{\infty}\sum_{n=0}^{\infty}\frac{(-1)^n z^k\omega^{n+k}}{k!n!2^{\binom{k+n}{2}}}\\ &= \sum_{m=0}^{\infty}\frac{(-1)^m\omega^m}{m!2^{\binom{m}{2}}}\sum_{k=0}^{m}\binom{m}{k}(-1)^k z^k\\ &= \sum_{m=0}^{\infty}\frac{(-1)^m\omega^m}{m!2^{\binom{m}{2}}}(1-z)^m. \qquad \square\end{aligned}$$

Remark 9.6
Liskovets obtained the following estimate for the values of the distribution function for arbitrary n (see [23, Prop. 3]):
For all $n \geq r \geq 2$,

$$\boldsymbol{P}\{\xi_n \leq r\} \geq 1 - \frac{n-r}{n+1}\frac{2}{(r+1)!2^{\binom{r}{2}}}.$$

9.3 Height of DAGs

In this section we will investigate the asymptotic behavior of the height of a random labeled DAG of size n. The results are based on McKay's work ([24]).

Definition 9.7 (height)
The height $h(D)$ of a labeled DAG D is the length of the longest directed path in D.

We are going to prove that the values of $h(D)$ over all labeled DAGs D of size n are asymptotically normally distributed with mean $C_1 n$ and variance $C_2 n$.

Let us denote by $\mathcal{A}_n$ the set of all labeled DAGs of size n and let $\mathcal{A} = \bigcup_{n=1}^{\infty} \mathcal{A}_n$.

Definition 9.8 (asymptotically normal, [24, p.459])
A function $f : \mathcal{A} \to \mathbb{R}$ is asymptotically normal over $\mathcal{A}$ with mean $\mu = \mu(n)$ and variance $\sigma^2 = \sigma^2(n)$ if

$$\lim_{n\to\infty} \sup_x \left| P(n, \mu + \sigma x, f) - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt \right| = 0,$$

where

$$P(n, z, f) = \frac{|\{D \in \mathcal{A}_n | f(D) \leq z\}|}{|\mathcal{A}_n|} \ .$$

In order to prove the normality of the height we will use the following theorem of Bender.

Theorem 9.9 ([3, Thm. 1])
Let $f(z, w)$ have a power series expansion $f(z, w) = \sum_{n,k\geq 0} a_n(k) z^n w^k$ with non-negative coefficients. Suppose there exists

(i) an $A(s)$ continuous and non-zero near 0,

(ii) an $r(s)$ with bounded third derivative near 0,

(iii) a non-negative integer m, and

(iv) $\epsilon, \delta > 0$,

such that

$$\left(1-\frac{z}{r(s)}\right)^m f(z,e^\delta)-\frac{A(s)}{1-z/r(s)}$$

is analytic and bounded for $|s|<\epsilon$, $|z|<|r(0)|+\delta$.

Let $\mu=-\frac{r'(0)}{r(0)}$, $\sigma^2=\mu^2-\frac{r''(0)}{r(0)}$.

If $\sigma\neq 0$ *then then* $a_n(k)$ *is asymptotically normal with* $\mu_n=n\mu$ *and* $\sigma_n^2=n\sigma^2$.

McKay mentioned that it should be possible to prove stronger local limit theorems for the same quantities, using Theorems 3 or 4 of [3], but it seems very difficult to verify the additional requirements.

Definition 9.10 (layers, [24])
Let $V_0=V_0(D)$ *be the set of out-points of a DAG* D.
For $k=1,2,\ldots$, *we define* $V_k=V_k(D)$ *inductively to be the set of out-points of the graph* $V(D)\setminus(V_0\cup V_1\cup\ldots\cup V_{k-1})$. *We denote by* V_h *the last* V_i *that is not-empty. The sets* $V_0,V_1,\ldots,V_h$ *are called the layers of* D.

Clearly $h=h(D)$ is the height of D. Remember that $(V_0,V_1,\ldots,V_h)$ is called the tower of D.

Now we introduce the special generating function

$$B(x,(y_1,y_2,\ldots))=\sum_{n\geq 1}\sum_{D\in\mathcal{A}}(n!2^{\binom{n}{2}})^{-1}x^n y_1^{n_1(D)}y_2^{n_2(D)}\ldots,$$

where $n_i=n_i(D)$ denotes the number of layers of D which have size i.

Theorem 9.11 ([24, Thm. 2.1])

$$B(x,(y_1,y_2,\ldots)) = \sum_{v_1,\ldots,v_h} \prod_{i=0}^{h-1} (1-2^{-v_i})^{v_{i+1}} \prod_{i=0}^{h} \frac{x^{v_i} y_{v_i}}{v_i! 2^{\binom{v_i}{2}}},$$

where the sum is over all tuples $(v_0, v_1, \ldots, v_h)$ *such that* $h \geq 0$ *and* $v_i \geq 1$ *for* $0 \leq i \leq h$.

Proof. We want to count the number of labeled DAGs of size n with height h and $|V_i(D)| = v_i$ for a fixed $(v_0, \ldots, v_h)$ with $0 \leq i \leq h$. Obviously $v_0 + \ldots + v_n = n$. There are $\binom{n}{v_0, v_1, \ldots, v_h}$ possibilities of assigning vertex labels to the layers $V_0, V_1, \ldots, V_h$. The in-neighbourhood of each vertex of layer V_i can be a subset of $V_0 \cup \ldots \cup V_{i-1}$, which contains at least one element of V_{i-1}. Clearly this gives $2^{v_0+\ldots+v_{i-1}} - 2^{v_0+\ldots+v_{i-2}} = 2^{v_0,v_1,\ldots,v_{i-2}}(2^{v_{i-1}} - 1)$ possibilities. Thus the number of DAGs corresponding to $(v_0, v_1, \ldots, v_h)$ is

$$\binom{n}{v_0, v_1, \ldots, v_h} \prod_{i=1}^{h} (2^{v_0,v_1,\ldots,v_{i-1}})^{v_i} (1-2^{-v_{i-1}})^{v_i},$$

which completes the proof. □

Now we will introduce some auxiliary notations, as it has been done in [24]:

Let Λ be the infinite matrix

$$\Lambda = \operatorname{diag}\left(\mathrm{xy}_1, \frac{\mathrm{x}^2\mathrm{y}_2}{4}, \ldots, \frac{\mathrm{x}^{\mathrm{i}}\mathrm{y}_{\mathrm{i}}}{\mathrm{i}! 2^{\binom{\mathrm{i}}{2}}}, \ldots\right),$$

and

$$M = (m_{ij}), \qquad \text{where } m_{ij} = (1-2^{-i})^j \text{ for} 1 \leq i, j < \infty.$$

Additionally, we define for any matrix M $\det(M)$ to be the determinant of M, adj(M) to be the adjoint matrix of M and $\mathcal{S}(M)$ to be the sum of all the entries of M. With I being the infinite identity matrix, the following equality holds:

Theorem 9.12 ([24, Thm. 2.2])

$$B(x,(y_1,y_2,\ldots)) = \mathcal{S}((I-\Lambda M)^{-1}\Lambda)$$

Proof. Clearly, the sum of the terms in Theorem 9.11 corresponding to any specific value of h is $\mathcal{S}((\Lambda M)^h\Lambda)$. Thus,

$$B(x,(y_1,y_2,\ldots)) = \mathcal{S}(\Lambda + \Lambda M\Lambda + (\Lambda M)^2\Lambda + \ldots) = \mathcal{S}((I-\Lambda M)^{-1}\Lambda).$$

□

Using $(I-\Lambda M)^{-1} = \mathrm{adj(I-\Lambda M)}/\det(\mathrm{I-\Lambda M})$, it can be shown (by the fact that the entries of Λ decrease very rapidly down the diagonal, see [24, p.461]) that each entry of $\mathrm{adj(I-\Lambda M)}$, their sum, and $\det(I-\Lambda M)$ are entire functions in x for the y_i $(i \geq 1)$ being fixed and uniformly bounded values.

Now we are ready to give the main result of this section:

Theorem 9.13 ([24, Thm. 3.1])
There are constants $C_1 \approx 0,7643344264$ *and* $C_2 \approx 0,1452097407$ *such that the height function* h *is asymptotically normal over* $\mathcal{A}$ *with mean* $C_1 n$ *and variance* $C_2 n$.

Proof. Define

$$H(x,y) = \sum_{n\geq 1}\sum_{D\in\mathcal{A}_n} \frac{x^n y^{h(D)}}{n!2^{\binom{n}{2}}}.$$

By Theorem 9.12 and the comments stated above we get

$$H(x,y) = \frac{c(x,y)}{yd(x,y)},$$

where $c(x,y) = \mathcal{S}(\mathrm{adj(I-\Lambda M)\Lambda})$ and $d(x,y) = \det(I-\Lambda M)$ with $\Lambda = \Lambda(x,(y,y,\ldots))$. As mentioned before $c(x,y)$ and $d(x,y)$ are entire functions

in x for fixed y. Therefore the singularities of $H(x,y)$ are given by the zeros of $d(x,y)$ which are not cancelled by zeros of $c(x,y)$. For example at $x=0$ we have a zero of both $c(x,y)$ and $d(x,y)$.
Since $H(x,1)=\frac{1}{B(x)}$, and the smallest zero of the entire function $B(x)$ is a simple pole at $\omega \approx 0,488078$ (cf. Thm 8.1) we can apply Theorem 9.9 with $m=0$ and a function $r(s)$ with $r(0)=\omega$.
The computation of the constants was done in [24] by use of Maple. □

In [24] McKay presented another interesting result, namely that the number of layers of a labeled DAG of a given size is also asymptotically normally distributed:

Theorem 9.14 ([24, Thm. 3.3])
Let $k \geq 1$. For $D \in \mathcal{A}$, define $n_k(D)$ to be the number of layers of D with size k. Then there are positive constants $C_{k,1}$ and $C_{k,2}$ such that the function n_k is asymptotically normal over $\mathcal{A}$ with mean $C_{k,1}n$ and variance $C_{k,2}n$.

Sketch of proof. Define

$$N_k(x,y) = \sum_{n\geq 1} \sum_{D\in\mathcal{A}_n} \frac{x^n y^{n_k(D)}}{n! 2^{\binom{n}{2}}}.$$

Then

$$N_k(x,y) = \frac{e(x,y)}{f(x,y)},$$

where $e(x,y) = \mathcal{S}(adj(I-\Lambda M)\Lambda)$, and $f(x,y) = det(I-\Lambda M)$, with $\Lambda = \Lambda(x,(1,1,\ldots,y,1,1,\ldots))$, and y being the k-th entry of the second argument.
The proof now proceeds analogously to that of Theorem 9.13. □

The following table (Figure 9.1) shows the number of DAGs of size n with height h:

h \ n	1	2	3	4	5	6	7	8	9	10
0	1	1	1	1	1	1	1	1	1	1
1		1	6	43	420	5821	113946	3142903	121796520	6631278361
2			1	22	610	25355	1634801	165904172	26641660900	6765756939985
3				1	70	6050	805700	175178976	63165264960	37991918437600
4					1	204	50288	19044480	11971739136	12852596994048
5						1	560	371840	376406016	633769500672
6							1	1472	2528256	6563880960
7								1	3744	16135680
8									1	9280
9										1

Figure 9.1: Number of DAGs of size n and height h from sizes up to 10 ([24, Table 2]).

Chapter 10

Special classes

This chapter is devoted to the investigation of two special subclasses of DAGs, namely extensional and essential DAGs. We will derive recursive formulas for the number of those special labeled DAGs of a given size and estimate their asymptotic behavior as the size tends to infinity. Additionally we will briefly discuss the problem of deriving the number of unlabeled extensional DAGs. The results given in this chapter are based on [35], [27] and [34].

10.1 Extensional DAGs

We start with the definition of an extensional DAG:

Definition 10.1 (extensional DAG,)
A DAG G is called extensional if for all pairs of vertices u, v of G, $u \neq v$ implies that u and v do not have the same out-neighbourhood.

10.1.1 Labeled extensional DAGs

Let e_n denote the number of labeled extensional DAGs of size n.
Then we get the following result:

Theorem 10.2 ([35, Equ. (2)])
The numbers e_n of labeled extensional DAGs of size n satisfy the recursion

$$e_n = \sum_{k=1}^{n} (-1)^{k+1} \binom{n}{k} (2^{n-k} - n + k)^{\underline{k}} e_{n-k}, \tag{10.1}$$

with $e_0 = 1$, where $x^{\underline{n}} = x(x-1)\dots(x-n+1)$ denotes a falling factorial.

This equation has been derived by Wagner in [35], and by Policriti and Tomescu in [27]. It is quite similar to the recursion for the number a_n of labeled DAGs of size n (therefore the proof of the recursion works analogously), but with the difference that instead of the factor $2^{k(n-k)}$ we have a factor $(2^{n-k} - n + k)^{\underline{k}}$. This modification can be explained by the following considerations:
First of all there are 2^{n-k} possibilities to connect the k newly added out-points with the $n-k$ pre-existing vertices, but $n-k$ choices are forbidden by the definition of extensional DAGs, since they are already out-neighbourhoods of the pre-existing vertices. Additionally we have to use a falling factorial rather than a k-th power, because the new out-neighbourhoods have also to be pairwise distinct.

Remark 10.3
In Equation (10.1) we could just sum up until $n - \lceil log_2\ n \rceil$ instead of n, since the non-null terms of this sum are those with $2^{n-k} \geq n$ (cf. [27], Thm. 1).

The number of labeled extensional DAGs of size 1 to 7 is ([35]):
1, 2, 12, 216, 10560, 1297440, 381013920

Our next goal is to estimate the asymptotic behavior of the sequence e_n.

For this purpose we follow the work of Wagner ([35]) and rewrite Equ. (10.1)

as

$$\sum_{k=0}^{n} \frac{(-1)^{n-k}}{(n-k)!} 2^{-\binom{k}{2}-\binom{n-k}{2}} \frac{e_k}{k!} \prod_{j=k}^{n-1}(1-2^{-k}j) = \delta_{n0}.$$

Since the product $\prod_{j=k}^{n-1}(1-2^{-k}j)$ is approximately 1, we rewrite the equation again resulting in

$$\sum_{k=0}^{n} \frac{(-1)^{n-k}}{(n-k)!} 2^{-\binom{k}{2}-\binom{n-k}{2}} \frac{e_k}{k!} = \sum_{k=0}^{n-1} \frac{(-1)^{n-k}}{(n-k)!} 2^{-\binom{k}{2}-\binom{n-k}{2}} \frac{e_k}{k!} \left(1 - \prod_{j=k}^{n-1}(1-2^{-k}j)\right) + \delta_{n0}.$$

Now let

$$E(x) = \sum_{n\geq 0} \frac{e_n x^n}{2^{\binom{n}{2}} n!}$$

be the special generating function of the sequence e_n and $B(x)$ is defined as in the previous chapters by

$$B(x) = \sum_{n\geq 0} \frac{(-1)^n x^n}{2^{\binom{n}{2}} n!}.$$

Using these notations the equation above transforms to

$$E(x)B(x) = 1 + \underbrace{\sum_{n\geq 0}\sum_{k=0}^{n-1} \frac{(-1)^{n-k} e_k}{k!(n-k)!} 2^{-\binom{k}{2}-\binom{n-k}{2}} \left(1 - \prod_{j=k}^{n-1}(1-2^{-k}j)\right) x^n}_{=:\psi(x)}.$$

Now we are going to show that $\psi(x)$, which defines the sum on the right hand side of the equation, converges for $|x| < 2\omega$.

For that purpose note that the inequality $e_k \leq a_k \leq C_1 k! 2^{\binom{k}{2}} \omega^{-k}$ holds for some constant C_1, so that

$$\left|\sum_{k=0}^{n-1}\frac{(-1)^{n-k}e_k}{k!(n-k)!}2^{-\binom{k}{2}-\binom{n-k}{2}}\left(1-\prod_{j=k}^{n-1}(1-2^{-k}j)\right)x^n\right|$$

$$\leq C_1\cdot\sum_{k=0}^{n-1}2^{-\binom{n-k}{2}}\omega^{-k}\left(1-\prod_{j=k}^{n-1}(1-2^{-k}j)\right)|x|^n.$$

Using

$$\prod_{j=k}^{n-1}(1-2^{-k}j)\geq 1-\sum_{j=k}^{n-1}2^{-k}j\geq 1-n^2 2^{-k},$$

we get

$$\sum_{k=0}^{n-1}2^{-\binom{n-k}{2}}\omega^{-k}\left(1-\prod_{j=k}^{n-1}(1-2^{-k}j)\right)|x|^n\leq n^2|x|^n\sum_{k=0}^{n-1}2^{-\binom{n-k}{2}}(2\omega)^{-k}.$$

Since the product $2^{-\binom{n-k}{2}}(2\omega)^{-k}$ is maximal for $k=n-\frac{3}{2}-log_2(\omega)$, reaching a value of $C_2(2\omega)^{-n}$ for some constant C_2, we finally get

$$\left|\sum_{k=0}^{n-1}\frac{(-1)^{n-k}e_k}{k!(n-k)!}2^{-\binom{k}{2}-\binom{n-k}{2}}\left(1-\prod_{j=k}^{n-1}(1-2^{-k}k)\right)x^n\right|$$

$$\leq C_1C_2n^3\left(\frac{|x|}{2\omega}\right)^n.$$

This proves that the function $\psi(x)$ is holomorphic for $|x|<2\omega$. Therefore $E(x)=\sum_{n\geq 0}\frac{e_n x^n}{2^{\binom{n}{2}}n!}$ is meromorphic for $|x|<2\omega$, except for a simple pole at ω.

By applying singularity analysis, we obtain

$$e_n\sim\frac{n!2^{\binom{n}{2}}(1+\psi(\omega))}{-\phi'(\omega)\omega^{n+1}}.$$

Thus, we can formulate the following theorem:

Theorem 10.4 ([35, p.5])
The asymptotic behavior of the sequence e_n of extensional labeled DAGs of size n is

$$e_n \sim \alpha\omega^{-n} n! 2^{\binom{n}{2}},$$

with $\alpha = -(1 + \psi(\omega))/(\omega\phi'(\omega)) \approx 0,567952$.

Now we are able to determine the ratio between labeled extensional DAGs and all labeled DAGs:

Corollary 10.5 ([35, Thm. 1])
The ratio of labeled extensional DAGs among all labeled DAGs is

$$\lim_{n\to\infty} \frac{e_n}{a_n} = 1 + \psi(\omega) \approx 3,065509^{-1} \approx 0,326210$$

The error term for the ratio can be described by an additional factor $1 + O(\gamma^{-n})$ for any fixed $\gamma > 2$, since the next singularity has absolute value at least 2ω.

10.1.2 Unlabeled extensional DAGs

Now we turn to the problem of counting unlabeled extensional DAGs, which will turn out to be reducible to the labeled case, that has just been discussed in the previous subsection.

In [27] it has been shown that labeled extensional DAGs have the nice property that their automorphism group is always trivial:

Lemma 10.6 ([27, Lem. 2])
Given a labeled extensional DAG G, the automorphism group $\mathcal{A}ut(G)$ of G is $\{id_G : V(G) \to V(G)\}$, where $V(G)$ is the set of all vertices of G and $id_G(v) = v$, for all $v \in V(G)$.

Proof. Let $f \in \mathcal{A}ut(G)$ and let $c = (x_1, \ldots, x_r)$ be a cycle of length $r \geq 2$ in the disjoint cycle decomposition of the permutation f on $V(G)$. We will now show that this leads to a contradiction, which implies that there can only be cycles of length 1, i.e. the automorphism group is trivial. The cycle $(x_1, \ldots, x_r)$ implies that $f(x_1) = x_2$, and by the extensionality of G, we have that the out-neighbourhoods of x_1 and x_2 are not equal. As f is an automorphism, we know that the out-neighbourhoods of x_1 and x_2 must have the same cardinality, hence their out-neighbourhoods cannot be empty and therefore there is a y_1, which is in the out-neighborhood of x_1, but not in that of x_2. Now suppose that $f(y_1) = y_1$. Then $\langle f(x_1), f(y_1)\rangle = \langle x_2, y_1\rangle \in E(G)$, which is a contradiction to y_1 not being in the out-neighbourhood of x_2. Therefore $f(y_1) \neq y_1$ and hence y_1 belongs to a cycle of the permutation f on $V(G)$ of length greater than or equal to 2. We can repeat the above procedure arbitrarily many times, and, as the number of vertices of G is finite, we will reach a vertex already visited. We thus contradict the acyclicity of G. □

This property of extensional DAGs to have trivial automorphism groups, makes it very easy to set up a recursion for the number of unlabeled extensional DAGs, because their number is simply $e_n/n!$.

Let E_n be the number of unlabeled extensional DAGs of size n.

With the arguments above we derive again a recursive equation for E_n that is similar to that of the number a_n of labeled DAGs of size n:

Theorem 10.7 ([27, Thm. 2])
The numbers E_n of unlabeled extensional DAGs of size n satisfy the recursion

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

with $E_0 = 1$.

Obviously in the unlabeled case, we have to use a binomial coefficient instead of a falling factorial (cf. Equation (10.1)), since the order of the vertices is irrelevant.

The number of unlabeled extensional DAGs of size 1 to 7 is ([35], Sloane's A001192): 1, 1, 2, 9, 88, 1802, 75598

Clearly, the asymptotic behavior of the number E_n follows immediately from that of e_n:

Theorem 10.8
The asymptotic behavior of the sequence E_n of unlabeled extensional DAGs of size n is

$$E_n \sim \alpha \omega^{-n} 2^{\binom{n}{2}},$$

with $\alpha = -(1 + \psi(\omega))/(\omega \phi'(\omega)) \approx 0,567952$.

Remember that in Section 8.3 we have shown that almost all DAGs have a trivial automorphism group (cf. Theorem 8.7). Therefore Corollary 10.5 holds for unlabeled extensional DAGs as well:

Corollary 10.9
The ratio of unlabeled extensional DAGs among all unlabeled DAGs is

$$\lim_{n\to\infty} \frac{E_n}{A_n} \approx 3,065509^{-1} \approx 0,326210$$

10.2 Essential DAGs

Now we focus on another interesting subclass of labeled DAGs, namely labeled essential DAGs. They originate from the study of Bayesian networks (see for example [1]), and their enumeration has been studied by Steinsky in [34] and [33].

Definition 10.10 (essential DAG, [35, p.14])
A DAG is called essential if there is no pair of two vertices u and v such that the in-neighbourhood of u is the union of the in-neighbourhood of v and the set $\{v\}$, i.e. there is no arc from a vertex v to a vertex u such that u has the same in-neighbours as v (except for v itself).

Let $\hat{e}_n$ denote the number of labeled essential DAGs of size n.

The recursion for $\hat{e}_n$ can be set up in the same way as the recursion for a_n by using the inclusion-exclsuion principle with the difference that now we add k terminal vertices (vertices with out-degree 0) to an essential DAG of size $n-k$, instead of k out-points. This has the advantage that removing such a terminal vertex from an essential DAG produces another essential DAG, because it does not influence the parents of any remaining vertex. Obviously any DAG of size $n > 0$ has at least one terminal vertex.

Theorem 10.11 ([34, p.270])
The numbers $\hat{e}_n$ of labeled essential DAGs of size n satisfy the recursion

$$\hat{e}_n = \sum_{k=1}^{n} \binom{n}{k} (2^{n-k} - n + k)^k \hat{e}_{n-k}, \tag{10.2}$$

with $\hat{e}_0 = 1$.

Again we have a factor $2^{n-k} - n + k$, since this is the number of possibilities to connect the pre-existing $n - k$ vertices with the newly added k terminal vertices, where $n - k$ choices (namely the in-neighbourhoods of the pre-existing vertices) are forbidden by the defining property of essential DAGs. In contrast to labeled extensional DAGs (cf. Equation (10.1)) the in-neighbourhoods of the k terminal vertices do not necessarily have to be distinct, therefore we have a k-th power instead of the falling factorial.

The number $\hat{e}_n$ of labeled essential DAGs from size 1 to 7 is ([34], Table 1):
1, 1, 4, 59, 2616, 306117, 87716644

In [33] the asymptotic behavior of labeled essential DAGs has been derived.

Theorem 10.12 ([33, Thm. 1])
The asymptotic behavior of the sequence $\hat{e}_n$ of labeled essential DAGs of size n is

$$\hat{e}_n \sim \alpha n! \omega^{-n} 2^{\binom{n}{2}},$$

with

$$\alpha \approx 13,6517978^{-1} \approx 0,1275334679.$$

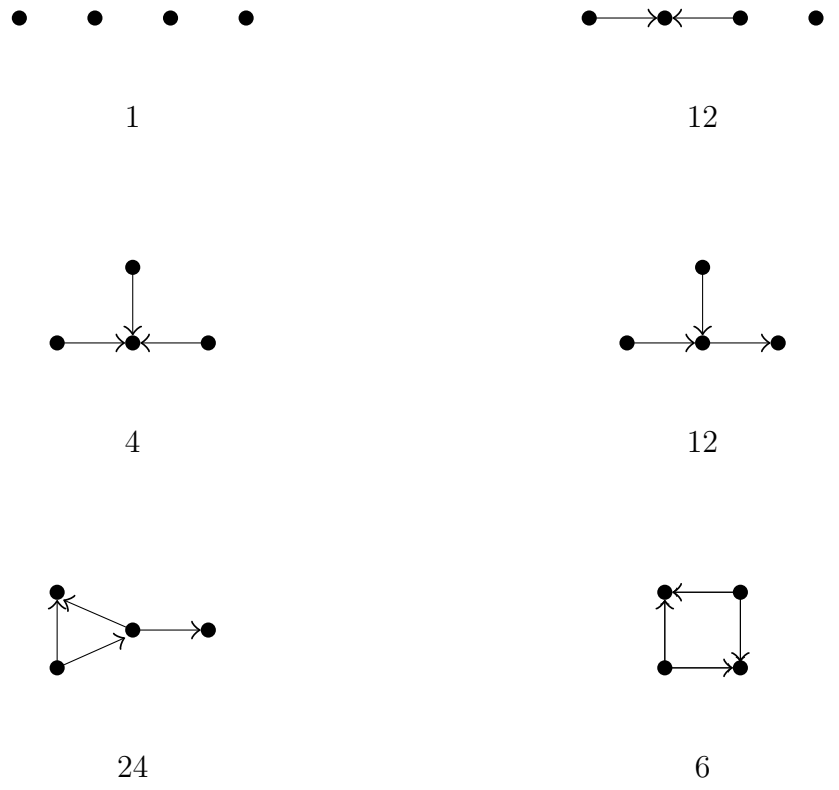


Figure 10.1: All unlabeled essential DAGs of size 4, together with the number of their possible labelings ([34, Figure 2]).

Proof. The proof is analogous to that of Theorem 10.4: The recursion (10.2) is rewritten in the form

$$B(x)\hat{E}(x) = 1 + h(x),$$

for some function $h(x)$, where $\hat{E}(x) = \sum_{n=0}^{\infty} \frac{\hat{e}_n z^n}{n! 2^{\binom{n}{2}}}$ denotes the special generating function of the sequence $\hat{e}_n$. Then it is shown that h is holomorphic

for $|x| < 2\omega$ and by applying singularity analysis, we get that

$$\hat{e}_n \sim \alpha n! \omega^{-n} 2^{\binom{n}{2}},$$

with

$$\alpha = -(1 - h(\omega))/(\omega\phi'(\omega)) \approx 0,1275334679\ldots\ .$$

□

Corollary 10.13 ([33, Cor. 1])
The proportion of labeled essential DAGs among all labeled DAGs is

$$\lim_{n\to\infty} \frac{\hat{e}_n}{a_n} \approx 0,073250.$$

In [35] it has been shown, that the distribution of the number of arcs, the number of out-points and the height are the same for extensional and essential DAGs, as well as for the family of all DAGs (The results for extensional and essential DAGs can be shown analogously to those for all DAGs with some simple modifications). This implies that all of them have essentially the same shape. This might be very different for other subclasses of DAGs, such as lambda-trees, as we have seen in the first part of this thesis.

Bibliography

[1] Steen A Andersson, David Madigan, Michael D Perlman, et al. "A characterization of Markov equivalence classes for acyclic digraphs." In: *The Annals of Statistics* 25.2 (1997), pp. 505–541.

[2] Hendrik Pieter Barendregt. *The Lambda Calculus*. Vol. 3. North-Holland Amsterdam, 1984.

[3] Edward A Bender. "Central and local limit theorems applied to asymptotic enumeration." In: *Journal of Combinatorial Theory, Series A* 15.1 (1973), pp. 91–111.

[4] Edward A Bender and Robert W Robinson. "The asymptotic number of acyclic digraphs, II." In: *Journal of Combinatorial Theory, Series B* 44.3 (1988), pp. 363–369.

[5] Edward A. Bender et al. "The asymptotic number of acyclic digraphs. I." In: *Combinatorica* 6.1 (1986), pp. 15–22.

[6] Maciej Bendkowski et al. "A natural counting of lambda terms." In: *International Conference on Current Trends in Theory and Practice of Informatics*. Springer. 2016, pp. 183–194.

[7] Olivier Bodini, Danièle Gardy, and Bernhard Gittenberger. "Lambda terms of bounded unary height." In: *ANALCO11—Workshop on Analytic Algorithmics and Combinatorics*. SIAM, Philadelphia, PA, 2011, pp. 23–32.

[8] Olivier Bodini, Daniele Gardy, and Alice Jacquot. "Asymptotics and random sampling for BCI and BCK lambda terms." In: *Theoretical Computer Science* 502 (2013), pp. 227–238.

[9] Olivier Bodini and Bernhard Gittenberger. "On the asymptotic number of BCK (2)-terms." In: *Proceedings of the Meeting on Analytic Algorithmics and Combinatorics*. Society for Industrial and Applied Mathematics. 2014, pp. 25–39.

[10] Olivier Bodini et al. "Enumeration of generalized BCI lambda-terms." In: *Electron. J. Combin.* 20.4 (2013), Paper 30, 23.

[11] Olivier Bodini et al. "On the number of unary-binary tree-like structures with restrictions on the unary height." In: *arXiv preprint arXiv:1510.01167* (2015).

[12] Louis Comtet. *Advanced Combinatorics: The Art of Finite and Infinite Expansions*. Springer Science & Business Media, 2012.

[13] René David et al. "Asymptotically almost all λ-terms are strongly normalizing." In: *Log. Methods Comput. Sci.* 9.1 (2013), 1:02, 30.

[14] Nicolaas Govert De Bruijn. "Lambda calculus notation with nameless dummies, a tool for automatic formula manipulation, with application to the Church-Rosser theorem." In: *Indagationes Mathematicae (Proceedings)*. Vol. 75. 5. Elsevier. 1972, pp. 381–392.

[15] Philippe Flajolet and Robert Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 2009.

[16] Steven B Gillispie. "Formulas for counting acyclic digraph Markov equivalence classes." In: *Journal of Statistical Planning and Inference* 136.4 (2006), pp. 1410–1432.

[17] Bernhard Gittenberger and Zbigniew Gołębiewski. "On the number of lambda terms with prescribed size of their De Bruijn representation." In: Nicolas Ollinger and Heribert Vollmer, editors, *33rd Symposium on Theoretical Aspects of Computer Science (STACS 2016)*, volume 47 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pp. 1–13, Schloss Dagstuhl–Leibniz-Zentrum für Informatik (2016).

[18] Katarzyna Grygiel, Paweł M. Idziak, and Marek Zaionc. "How big is BCI fragment of BCK logic." In: *J. Logic Comput.* 23.3 (2013), pp. 673–691.

[19] Katarzyna Grygiel and Pierre Lescanne. "Counting and generating lambda terms." In: *Journal of Functional Programming* 23.05 (2013), pp. 594–628.

[20] Andrey Nikolaevich Kolmogorov and BV Gnedenko. "Limit distributions for sums of independent random variables." In: *Addison-Wesley, Cambridge, Mass* (1954).

[21] Pierre Lescanne. "On counting untyped lambda terms." In: *Theoretical Computer Science* 474 (2013), pp. 80–97.

[22] Valery A Liskovets. "More on counting acyclic digraphs." In: *arXiv preprint arXiv:0804.2496* (2008).

[23] Valery A Liskovets. "On the number of maximal vertices of a random acyclic digraph." In: *Theory of Probability & Its Applications* 20.2 (1976), pp. 401–409.

[24] Brendan D McKay. "On the shape of a random acyclic digraph." In: *Mathematical Proceedings of the Cambridge Philosophical Society*. Vol. 106. 03. Cambridge Univ Press. 1989, pp. 459–465.

[25] Brendan D McKay et al. "Acyclic digraphs and eigenvalues of (0, 1)-matrices." In: *Journal of Integer Sequences* 7.2 (2004), p. 3.

[26] Andrew M Odlyzko. "Asymptotic enumeration methods." In: *Handbook of Combinatorics* 2 (1995), pp. 1063–1229.

[27] Alberto Policriti and Alexandru I Tomescu. "Counting extensional acyclic digraphs." In: *Information Processing Letters* 111.16 (2011), pp. 787–791.

[28] Robert W Robinson. "Counting labeled acyclic digraphs." In: *New Directions in Graph Theory (Academic Press)* (1973), pp. 239–273.

[29] Robert W Robinson. "Counting unlabeled acyclic digraphs." In: *Combinatorial Mathematics V*. Springer, 1977, pp. 28–43.

[30] VI Rodionov. "On the number of labeled acyclic digraphs." In: *Discrete Mathematics* 105.1 (1992), pp. 319–321.

[31] Neil JA Sloane et al. *The On-line Encyclopedia of Integer Sequences*. 2003.

[32] Richard P Stanley. "Acyclic orientations of graphs." In: *Discrete Mathematics* 5.2 (1973), pp. 171 –178.

[33] Bertran Steinsky. "Asymptotic behaviour of the number of labelled essential acyclic digraphs and labelled chain graphs." In: *Graphs and Combinatorics* 20.3 (2004), pp. 399–411.

[34] Bertran Steinsky. "Enumeration of labelled chain graphs and labelled essential directed acyclic graphs." In: *Discrete Mathematics* 270.1 (2003), pp. 267–278.

[35] Stephan Wagner. "Asymptotic enumeration of extensional acyclic digraphs." In: *Algorithmica* 66.4 (2013), pp. 829–847.

[36] Jue Wang. "Generating random lambda calculus terms." In: *Unpublished manuscript* (2005).

[37] Edward M Wright. "Graphs on unlabelled nodes with a given number of edges." In: *Acta Mathematica* 126.1 (1971), pp. 1–9.