The Largest Face in Maps with Few Faces

by

Koen van Greevenbroek

Honours Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Bachelor of Science

in the
Department of Mathematics
Faculty of Science

SIMON FRASER UNIVERSITY
Fall 2016

Approval

Name: **Koen van Greevenbroek**

Degree: **Bachelor of Science (Mathematics)**

Title: ***The Largest Face in Maps with Few Faces***

Supervisory Committee:

Marni Mishna
Supervisor
Professor

Éric Fusy
Supervisor
Professor

Date Approved:

Abstract

The study of large random planar maps and their properties is important for the understanding of the objects that maps encode. We approximate the expected degree of the largest face in a random planar map, when restricting the number of faces. While much is known about maps on n edges, the constraint on the number of faces is new and reveals a spectrum of planar maps, from plane trees to their duals. To obtain the approximation of the expected degree of the largest face, we adapt known bijections between maps, trees, binary sequences and integer compositions to keep track of the degrees and number of faces of the maps. This allows us to use more classical tools to analyze the resulting compositions, and translate the results back to planar maps. The techniques in this thesis suggest a new perspective on map parameters.

Keywords: planar maps, bijective combinatorics, largest face degree, asymptotic analysis, generating functions.

Acknowledgements

I am very thankful to Professor Marni Mishna for her help and guidance in writing this thesis. She has always given me valuable feedback, and has helped me get the best result possible out of the thesis. Of course, I am also very grateful for her funding of the research I did this summer that lead to the thesis.

I am equally thankful to Professor Éric Fusy from LIX, École Polytechnique, for providing a great topic, and always being available to answer any technical questions. In me, you have recruited another fan of planar maps and combinatorics in general!

I would also like to thank my fellow students in the MATH 498 class, who have given me good feedback on my communication of the material in this thesis.

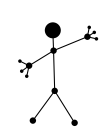


Table of Contents

Approval **ii**

Abstract **iii**

Acknowledgements **iv**

Table of Contents **v**

List of Figures **vi**

1 Introduction **1**

1.1 Planar Maps . . . 3

1.2 Combinatorial Classes & Generating Functions . . . 4

1.3 Background and Context . . . 6

1.3.1 Planar Graphs . . . 7

1.3.2 Maps Beyond Combinatorics . . . 8

1.3.3 Similar Result . . . 8

2 The Bijection **9**

2.1 Proof of Theorem 1 . . . 11

2.1.1 From Maps to Trees: The Master Bijection . . . 11

2.1.2 From Trees to Binary Sequences . . . 16

3 Analysis & Results **20**

3.1 The Boltzmann Distribution . . . 21

3.2 $\mathcal{C}^{(r)}$ Under the Boltzmann Distribution . . . 22

3.3 Approximation of $\mathbb{E}(X)$ . . . 24

3.3.1 Limiting Distribution of X_i . . . 24

3.3.2 Maximum of Exponential Random Variables . . . 26

3.4 Results . . . 27

4 Conclusion **29**

List of Figures

Figure 1.1 Spectrum of planar maps 1
Figure 1.2 Example of planar map 3
Figure 1.3 Planar map with dual 4
Figure 1.4 All 4 planar maps on 2 edges. 5

Figure 2.1 Main bijection example 10
Figure 2.2 Main bijection overview 12
Figure 2.3 Master bijection example 13
Figure 2.4 Local transformation in the master bijection. 13
Figure 2.5 Inverse of master bijection example 15
Figure 2.6 Inverse of master bijection detail 15
Figure 2.7 Tree transformation example 16
Figure 2.8 Breadth first search encoding example 17
Figure 2.9 Dyck word to binary sequence example 18
Figure 2.10 Dyck path to lattice walk example 18
Figure 2.11 Maps with corresponding compositions 19

Figure 3.1 Comparison of uniform and Boltzmann distribution 21
Figure 3.2 Comparison between numerical data and prediction 28

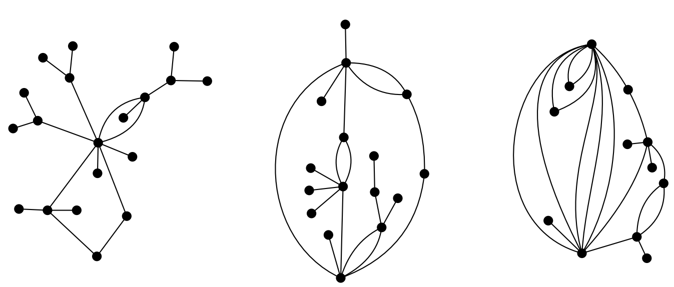
Chapter 1

Introduction

Planar maps are natural objects in combinatorics that can be defined as planar graphs embedded in a surface. We consider asymptotic behaviour of large random maps — what does a typical map look like? The usual approach is to pick some parameter of planar maps and study its distribution over maps with n edges. Our focus is the expected degree of the largest face. However, we not only fix the number of edges but also the number of faces. In particular, we consider bipartite maps with n edges and n^α faces, for a fixed α between 0 and 1.

In this way, we study a "spectrum" of maps, parameterized by α, from maps with few faces to maps with many faces. Our motivating example is the set of three bipartite planar maps show in Figure 1.1, which were generated uniformly at random on 20 edges and a given numbers of faces. As we can see, there seems to be an inverse correlation between the number of faces and the degree of the largest face. This is the expected behaviour, since the degrees of the faces must sum up to the twice the number of edges, which is 20 in this case. We would like to quantify this behaviour. The main question becomes:

> **Question:** what is the expected degree of the largest face in random map with n edges and n^α faces?



Number of faces:	3	6	10
Degree of largest face:	28	14	8

FIGURE 1.1: The spectrum of planar maps. Fewer to more face from left to right.

In the spectrum of maps, the extreme cases are trees and their duals. For $\alpha = 0$ we have trees, with only one face of degree $2n$. When α approaches 1, we get maps with $O(n)$ faces, the largest of which has expected degree $\log n$[1]. The behaviour in these cases is know, but we look at what happens in between the extremes, with $0 < \alpha < 1$.

The main result in this thesis is evidence for the following conjecture:

Conjecture 1. *The expected degree of the largest face in vertex-rooted bipartite maps with n edges and n^α faces is*

$$c\,\alpha\,n^{1-\alpha}\log n + d\,n^{1-\alpha} + O(1),$$

where c and d are constants.

Note that for $\alpha = 0$ and $\alpha = 1$ the conjecture recovers the edge cases of a maximum face of expected degree proportional to n and $\log n$, respectively.

While this is not the first study of the expected degree of the largest face in planar maps[2], this is the first time that the number of faces is also restricted. It is important to note that we still relate the number of edges and faces by considering maps with n edges and n^α faces. This makes it possible to study maps with not too few and not too many faces. Here, "too few" is understood as a constant number, while "too many" means in the order of $O(n)$.

We are interested in what "typical" large planar maps look like because planar maps and their many variations come up in a number of different contexts. First of all, maps have beautiful and simple counting formulas. This in itself makes the typical properties of maps a relevant and interesting question. However, maps are also useful outside the purely combinatorial realm. For example, they appear in some areas of physics where they essentially act as a "discretization" of a surface. Another example where maps find an interesting application is the study of so-called *Brownian surfaces*, which are in some way the surface analogues of particles with Brownian motion. In Section 1.3 we review these connections in some more detail.

This thesis is split into two different parts:

Part 1 We present a bijection between bipartite maps and a type of weighted binary sequences. The main ingredient in this bijection is the "master bijection" from [2] by Bernardi and Fusy; this is a bijection from classes of maps to decorated trees. These decorated trees are put in bijection with binary sequences, and a last simplification is to map the result to integer compositions.

[1] In [12], it is shown that the degree of the largest face in rooted maps with n edges is of the order $\log n$. Since planar maps with n edges have $n/2$ faces on average, it follows that maps with $O(n)$ edges also have a largest face of degree $\log n$.

[2] See e.g. [12]. Since maps are in bijection with trees in various ways, papers such as [17] are also relevant, in addition to techniques from [10].

Part 2 The problem is now reduced to finding the expected size of the largest part in our weighted integer compositions. We find an approximation by considering a Boltzmann distribution over the weighted integer compositions. Analysis of the compositions under the Boltzmann distribution suggests that the sizes of the parts are distributed close to exponentially, and this leads to the conjecture.

However, for the remainder of this chapter we go over definitions and background in detail.

1.1 Planar Maps

We assume familiarity with basic concepts of graph theory, see the definitions in any undergraduate textbook on graph theory.

Definition 1. A *planar map* is a connected planar graph embedded in a surface, considered up to continuous deformation.

In this thesis, we only consider maps embedded in spheres. Furthermore, we only concern ourselves with unlabelled maps, technically equivalence classes of labelled maps. Lastly, to facilitate counting, we almost always consider our maps rooted in some way, either at a vertex, an edge or a face.

A different way to look at planar maps is that it is a planar graph where the order of edges around each vertex is given. Crucially, every map has a well-defined set of faces, each identified by the edges it is incident to. Figure 1.2 show an example of several different drawings of the same planar map. Note that we may choose the outer face arbitrarily when we draw planar maps. However, if the map is rooted at a face, we usually choose to draw the root face as the outer face. Likewise, if the map is rooted at a directed edge, we usually choose to draw the face on the right (or left) of the root edge as the outer face.

Many definitions and terms from graph theory carry over to maps directly. For example, a map is *bipartite* if its underlying graph is bipartite, or *Eularian* if its underlying graph is Eularian. Two vertices or two faces in a map are *adjacent* if they are incident to the same edge, and so forth. One term that is new for maps is the concept of a *corner*; this is (informally) the angle between two edges at a vertex. Formally, a corner is a vertex v together with two endpoints of edges incident to v, that occur consecutively in the order of

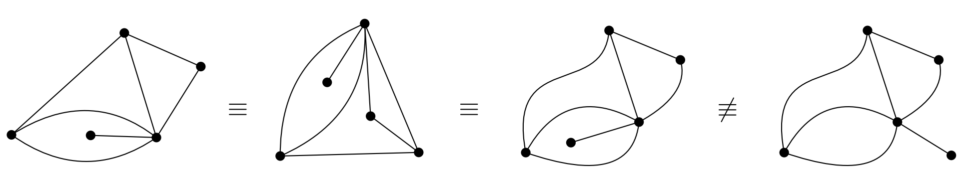


FIGURE 1.2: Three drawings of the same planar maps (left), and one different planar map (right).

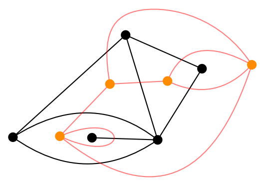


FIGURE 1.3: A planar map in black with its dual in orange.

edges around v. Note that when we refer to rooted plane trees, we usually mean plane trees that are rooted at a corner. Rooting plane trees only a vertex would still allow cycling the children of the root vertex.

Lastly, we will also use the concept of a *dangling half-edge*. This is simply a single endpoint of an edge inserted into a corner of a map. We will not make it contribute to the degree of the vertex it is incident to, and also not allow it as one of the edge endpoints in a corners.

The additional structure of maps (the embedding) allows us to do things with maps that are not possible with graphs in general. In particular, we will see that the concept of dual graphs is much more powerful for maps than for graphs.

Definition 2. The *dual* of a planar map M is another planar map M', obtained by putting a vertex in each of the faces of M, and for each edge e in M draw an edge between the vertices in the two faces that e is incident to. None of the vertices or edges of M are part of M'.

See Figure 1.3 for an example of a map with its dual. Note that Figure 1.3 shows examples of both loops and multiple edges. To be as flexible as possible when it comes to taking duals, we allow this and do not enforce our maps to be simple.

One nice thing about working with maps is that every map has a unique dual. Furthermore, taking the dual is an involution: taking the dual twice is guaranteed to give back the same map. This is useful as it allows us to translate results back and forth between maps and their duals. For example, we can compare our results about the largest face to results from [12], which are about the largest vertex. Where we have results about bipartite maps, the dual results are about the duals of bipartite maps: Eularian maps where every vertex has an even degree.

The duality between vertices and faces is also the reason for counting maps by *edges*, not by vertices or faces. This way, results about maps with n edges have dual results that are also about maps with n edges, since edges are self-dual.

1.2 Combinatorial Classes & Generating Functions

We are interested in maps as combinatorial objects. This means that we attach some notion of size to maps, and consider maps of a given size. What follows is a very brief overview of

FIGURE 1.4: All 4 planar maps on 2 edges.

the basics of combinatorial families and their generating functions, which we will need in Chapter 3. For a more complete reference, see *Analytic Combinatorics* [10].

Definition 3. A *combinatorial class* $\mathcal{A}$ is a set together with a size function $|\cdot|_{\mathcal{A}} \colon \mathcal{A} \to \mathbb{Z}_{\geq 0}$ such that there are only finitely many objects in $\mathcal{A}$ of any given size. We denote by $\mathcal{A}_n$ the (finite) set of objects in $\mathcal{A}$ of size n, and write $a_n = |\mathcal{A}_n|$. The sequence $a_0, a_1, a_2, a_3, \ldots$ is called the *counting sequence* of $\mathcal{A}$.

Example. Let $\mathcal{M}$ be the combinatorial family of planar maps, counted by edges. The counting sequence starts $1, 2, 4, 14, 52, 248, \ldots$ (A006385). For example, $m_2 = |\mathcal{M}_2| = 4$ since there are 4 maps with 2 edges, listed in Figure 1.4.

The easiest way to work with counting sequences is often using a generating function.

Definition 4. The *generating function* $A(z)$ of a combinatorial class $\mathcal{A}$ is the formal power series

$$A(z) = \sum_{n=0}^{\infty} a_n z^n.$$

We may use existing combinatorial classes to construct new ones with a few elementary constructions.

Definition 5. Let $\mathcal{A}$ and $\mathcal{B}$ be two combinatorial classes. We define

(i) $\mathcal{C} = \mathcal{A} \cup \mathcal{B}$ is the union of $\mathcal{A}$ and $\mathcal{B}$, with size function

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

(ii) $\mathcal{C} = \mathcal{A} \times \mathcal{B}$ is the cartesian product of $\mathcal{A}$ and $\mathcal{B}$, with size function $|(a, b)| = |a| + |b|$.

(iii) $\mathcal{C} = \textsc{Seq}(\mathcal{A})$ is the set of ordered sequences of objects from $\mathcal{A}$, with additive size as in the cartesian product.

When using these constructions, we often use $\mathcal{Z}$ to denote the atomic class with a single object of size 1, and ϵ to denote the class with a single element of size 0. The generating functions of these elementary classes are z and 1 respectively.

The combinatorial constructions have straightforward consequences in the context of generating functions:

(i) If $\mathcal{C} = \mathcal{A} \cup \mathcal{B}$, then $C(z) = A(z) + B(z)$.

(ii) If $\mathcal{C} = \mathcal{A} \times \mathcal{B}$, then $C(z) = A(z)B(z)$.

(iii) If $\mathcal{C} = \textsc{Seq}(\mathcal{A})$, then $C(z) = 1/(1 - A(z))$.

Example. Let $\mathcal{B}$ be the class of binary sequences counted by length, so $\mathcal{B} = \textsc{Seq}(\mathcal{Z}_0 \cup \mathcal{Z}_1)$, where $\mathcal{Z}_0$ and $\mathcal{Z}_1$ are atomic classes representing a 0 and a 1, respectively. The generating function is $B(z) = 1/(1 - 2z)$.

Example. Let $\mathcal{T}$ be the class of unlabelled rooted plane trees counted by number of nodes. The recursive specification $\mathcal{T} = \mathcal{Z} \times \textsc{Seq}(\mathcal{T})$ leads to the functional equation $T(z) = z/(1 - T(z))$ for the generating function. Solving this functional equation as a quadratic gives two solutions, but only

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

has a power series representation with positive coefficients. The plane trees are counted by the *Catalan numbers* C_n,

$$t_n = C_n = \frac{1}{n+1}\binom{2n}{n}.$$

Example. Let $\mathcal{C}$ be the class of integer composition, that is, the class of ordered sequence of positive integers, counted by the sum of the integers. Then we have $\mathcal{C} = \textsc{Seq}(\mathcal{I})$ where $\mathcal{I}$ is the combinatorial class of positive integers, so $\mathcal{I} = \textsc{Seq}_{\geq 1}(\mathcal{Z})$. The generating function is

$$C(z) = \frac{1}{1 - \frac{z}{1-z}} = \frac{1 - z}{1 - 2z}.$$

One of the most powerful ways to use generating functions is to regard them as power series representations for complex functions. Then, the singularities of the complex function tell us about the asymptotic growth of the coefficients of the associated power series. In particular, the *location* of the singularity closest to the origin, called the dominant singularity, tells us about the exponential growth. Meanwhile, the *type* of the dominant singularity dictates the sub-exponential growth. For proofs and a more in-depth reference for these techniques, see [10].

1.3 Background and Context

Tutte is generally accepted as the first to study maps seriously in the 60s. With a series of papers that culminated with “On the enumeration of planar maps” [23] in 1968, he laid the groundwork for map enumeration. Tutte found recursive decompositions for many types of maps, which led to the generating functions for these classes and hence their counting sequences. For example, Tutte found that the number of planar maps rooted at a directed edge with n edges is

$$\frac{2 \cdot 3^n}{(n+1)(n+2)}\binom{2n}{n} = \frac{2 \cdot 3^n}{n+2} C_n \sim \frac{2}{\sqrt{\pi}} 12^n n^{-5/2}. \tag{1.1}$$

This counting formula is considered nice: it is explicit, compact and involves Catalan numbers. Counting formulas for other classes of maps that Tutte found are similarly nice. Moreover, the generating functions that Tutte found are *algebraic* (i.e. the roots of polynomials) and hence relatively easy to study. The method that Tutte used to find these generating function was to find a recursive decomposition of maps, and translate the result into a functional equation for the generating function. In the case of maps, a catalytic variable keeping track of the degree of the outer face is needed, and because of this some manipulation of the resulting functional equation is required, using the *quadratic method* or the *kernel method*. See [23] and previous papers by Tutte in the same series for details.

Although Tutte obtained nice generating functions for various families of maps, his approach does not shed any light on why the counting formulas and generating functions should be so well-behaved. This is because techniques such as the quadratic method or the kernel method don't have combinatorial interpretations. The issue was resolved by Schaeffer and others in the 90s as bijections were discovered between planar maps and some types of (decorated) trees. See for example [14], [19] and [20]. Through these bijections, we have purely combinatorial proofs of the counting formulas that don't rely on some black box algebraic technique. By now, there are bijections to trees for many classes of maps, for example 2 and 3-connected maps, maps with a given girth (shortest cycle), maps with restricted face or vertex degrees, etc. Since it is known that trees always have algebraic generating functions (see for example Section I.5.4 in [10]), this explains the algebraicity of the generating functions of maps.

It is also worth noting that the counting formulas that Tutte discovered suggest a *universality class* of planar maps, which means that the asymptotic growth of many different classes of planar maps is similar. In particular, the sub-exponential term $n^{-5/2}$ that we see in Equation (1.1) is universal for most types of maps. This too is explained by the bijections between trees and maps: trees themselves form a universality class and share a sub-exponential growth term of $n^{-3/2}$. The universality class of trees is in turn explained by the type of generating functions that trees have. As mentioned in the previous section, the type of the dominant singularity of a generating function determines the sub-exponential growth of the associated counting sequence, and all generating functions associated to tree classes have the same type of dominant singularity. The reference for this material is [10].

Next, we will look at a few different ways that maps can be useful.

1.3.1 Planar Graphs

Some of the motivation for studying planar maps comes from their connections with planar graphs. One basic problem in graph theory is to find the counting sequence for the combinatorial family of labelled planar graphs. Planar graphs are harder to study than maps, since they lack the additional structure of an explicit embedding. However, Whitney's theorem [24] tells us that 3-connected planar graphs have a *unique* embedding up to reflection, hence being basically equivalent to planar maps.

We know the generating function $F(z)$ for 3-connected planar maps and hence 3-connected planar graphs; this was first studied by Tutte in [22]. Now we may decompose 2-connected planar graphs into 3-connected components to get a generating functions for 2-connected planar graphs, depending on $F(z)$. Decompose further to get a generating function for 1-connected planar graphs, and finally decompose planar graphs into 1-connected components to get a generating function expression for general labelled planar graphs. This approach was followed in [13], and the authors conclude that the number of labelled planar graphs on n vertices is asymptotically

$$g_n = g \cdot n^{-7/2} \gamma^n n!$$

where g and γ are computable constants. This result is further extended to graphs embedded in surfaces of arbitrary genus in [6].

1.3.2 Maps Beyond Combinatorics

One area where maps have found an application is in theoretical physics, more specifically, in *quantum gravity*, where maps can be regarded as a discretization of a surface. Physicists have developed their own interesting methods of map enumerations based on integrals over a space of matrices. For an introduction to these matrix integrals and the connection between maps and physics, see [25] and its references. A bijective approach to some physical models on planar graphs is presented in [3].

A different area in mathematics where maps are being applied is the study of so-called *Brownian surfaces*, which are similar to particles with Brownian motion. Random maps are used to defined a random normalized metric, which tends to some continuous limit as the maps grow bigger and bigger. See for example [16].

1.3.3 Similar Result

Some of the most directly related work is by Gao and Wormald in [12], who show that the expected degree of the largest face in a random map with n edges is in the order of $\log n$. The present thesis treats the extension of this result to maps with n edges and n^α faces.

Work in a similar spirit includes [4], [7] and [9] which all involve the study of parameters of random maps, or by extension, trees. Even a paper such as [15] deals with a problem that is somewhat similar in nature to ours.

Chapter 2

The Bijection

In this section, we present a bijection between bipartite planar maps and a type of weighted binary sequences. Afterwards, we can ignore some of the structure of the binary sequences, and end up with weighted integer compositions. For our purposes, it is critical that the bijection keeps track of the faces and their degree, since we want to study the degree of the largest face. The bijection that we present mainly consists of a specialization of the "master bijection" by Bernardi and Fusy from [2]. The master bijection is then composed with a few more standard bijections.

To state the main theorem and corollary, we need to introduce some new combinatorial families[1].

- Let $\mathcal{M}_n$ denote the class of vertex-rooted bipartite planar maps on n edges, with one marked edge and one marked (non-root) vertex or face.
- Let $\mathcal{B}_n$ denote the class of weighted binary strings with n 0s and n 1s, where every run of 0s of length d has an integer weight between 1 and $\binom{2d-1}{d}$.
- Let $\mathcal{C}_n$ denote the class of integer compositions of size n, where every part of size d has an integer weight between 1 and $\binom{2d-1}{d}$.

As a short-hand notation, we will also use superscripts such as $\mathcal{M}_n^{(r)}$ to denote all maps in $\mathcal{M}_n$ with r faces. Use the same interpretation for $\mathcal{B}_n^{(r)}$ and $\mathcal{C}_n^{(r)}$, replacing 'faces' by 'runs of 0s' and 'parts' respectively.

Theorem 1. *There is a bijection between $\mathcal{M}_n$ and $\mathcal{B}_n$ which sends a map with r faces of degrees $2d_1, 2d_2, \ldots, 2d_r$ to a weighted binary string with r runs of 0s of lengths $d_1, d_2, \ldots, d_r$.*

The proof of this theorem is constructive, and is completed in Section 2.1. We will refer to the bijection that is constructed in the proof as Φ. And example of two objects in $\mathcal{M}_{10}$ and $\mathcal{B}_{10}$ that are in bijection under Φ is given in Figure 2.1.

[1]Note that these are not quite the same as the classes we introduced in the introduction.

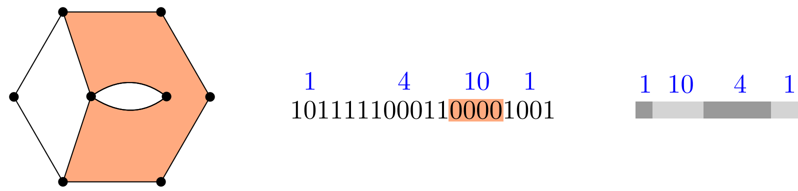


FIGURE 2.1: Left, two objects in $\mathcal{M}_{10}$ and $\mathcal{B}_{10}$ that are in bijection under Φ. The largest face in the map and the corresponding longest run of 0s is marked in orange. Right: the weighted integer composition corresponding to the objects on the left. The length of the grey blocks correspond to the size of the parts.

By Theorem 1, the expected degree of the largest face in a random map in $\mathcal{M}_n$ is twice the expected length of the longest run of 0s in a random weighted binary sequence in $\mathcal{B}_n$. Since weighted binary sequences are more explicit objects than planar maps (they are easier to represent; it is easy to see what information they contain), this has made our problem easier.

However, we can simplify even further. Note that only the 0s in binary sequences in $\mathcal{B}_n$ carry any information that relates to face degree; the 1s do not. Hence, we can strip the 1s away altogether, and regard the remaining runs of 0s as weighted parts in an integer composition of size n. This is a surjection from $\mathcal{B}_n$ to $\mathcal{C}_n$. In Figure 2.1 we see a map in $\mathcal{M}_{10}$ with the corresponding integer composition of size 10 on the right. We draw integer compositions as a sequence of bars, where a bar of length k indicates a part in the composition of size k. The blue integers above each part are the weights. We now get the following corollary:

Corollary 1. *The expected degree of the largest face of a random map in $\mathcal{M}_n^{(r)}$ is twice the expected size of the largest part of a random composition in $\mathcal{C}_n^{(r)}$.*

This works because for every element in $\mathcal{C}_n^{(r)}$, there are exactly $\binom{n+r}{n}$ ways to put n 1s between the r parts to get binary strings in $\mathcal{B}_n^{(r)}$. Hence the surjection from $\mathcal{B}_n^{(r)}$ to $\mathcal{C}_n^{(r)}$ is a $\binom{n+r}{n}$-to-1 map preserving runs of 0s, and results about the runs of 0s in $\mathcal{B}_n^{(r)}$ carry over to results about parts in $\mathcal{C}_n^{(r)}$.

In Chapter 3, we obtain results about weighted integer compositions in $\mathcal{C}_n$. We then use Corollary 1 to transfer these results back to maps in $\mathcal{M}_n$.

Finally, note that the maps in $\mathcal{M}_n$ are marked in a number of ways; this may seem to limit our results to an obscure family of maps. However, we may un-mark the marked edge and the marked face or vertex to obtain a correspondence between a single vertex-rooted bipartite map and a number of weighted binary sequences in $\mathcal{B}_n$. The number of ways to mark an edge and a face or vertex in a map with n edges and r faces is always the same, if the map is asymmetric. In [18] it is shown that almost all maps are asymmetric. It follows that any results about the expected degree of the largest face in maps in $\mathcal{M}_n^{(r)}$ must also hold asymptotically for vertex-rooted bipartite maps with n edges and r faces, or even just bipartite maps with n edges and r faces.

2.1 Proof of Theorem 1

We construct a bijection Φ to prove Theorem 1 as a functional composition of 4 individual bijections. This way, we "chain" several different classes of objects together with bijections, starting with $\mathcal{M}_n$ and ending with $\mathcal{B}_n$. The three classes between $\mathcal{M}_n$ and $\mathcal{B}_n$ are two types of trees and a class of weighted Dyck words; we define these classes shortly.

In Section 2.1.1, we cover the first and most extensive of these bijections, the one from maps to trees. This is the so-called "master bijection" introduced by Bernardi and Fusy in [1] and [2]. In Section 2.1.2 we cover the rest of the bijections. The bijections in the latter section are basically folklore, and can be regarded as merely encodings to get the trees from the master bijection in the form that is the most convenient for us. Throughout the following sections, we will use the map seen in Figure 2.1 as a running example to apply our various bijections to.

Let's define the different objects between $\mathcal{M}_n$ and $\mathcal{B}_n$ that we will show are all in bijection:

- Let $\mathcal{T}_n$ denote the class of plane rooted trees with n edges, and one marked vertex. Furthermore, every vertex of degree d at even distance from the root has an integer label between 1 and $\binom{2d-1}{d}$. We refer to vertices at even distance from the root simply as *even distance* vertices, and draw them in orange.

- Let $\mathcal{T}_n^*$ be the class of plane rooted trees with n edges, and one marked vertex. Furthermore, every internal vertex with d children has an integer weight between 1 and $\binom{2d-1}{d}$.

- Let $\mathcal{D}_n$ be the class of Dyck words with n 0s and $n+1$ 1s, and a marked 1. Furthermore, every run of 0s of length d has an integer weight between 1 and $\binom{2d-1}{d}$. Recall that a Dyck words are binary sequences where every prefix except the whole sequence contains at least as many 0s as 1s.

For a compact summary of the various objects and bijections involved in the proof of Theorem 1, see Figure 2.2.

2.1.1 From Maps to Trees: The Master Bijection

We now present the application of the master bijection from [2] to our problem. Although we do not provide a complete proof of this bijection, we will go through all the necessary steps to apply the bijection and its inverse.

Lemma 1 (Master bijection)**.** *There is a bijection between $\mathcal{M}_n$ and $\mathcal{T}_n$ which sends a map with r faces of degrees $2d_1, 2d_2, \ldots, 2d_r$ to a tree in $\mathcal{T}_n$ with r even distance vertices of degrees $d_1, d_2, \ldots, d_r$.*

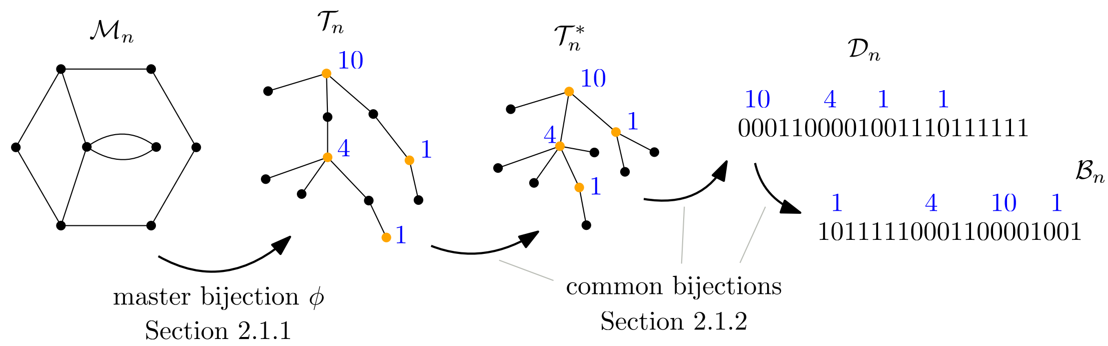


FIGURE 2.2: A visual example of the objects involved in the bijection Φ.

The master bijection is in fact a framework that can be specialized to many different kinds of maps, and produces different kinds of trees in different cases. Although the bijection is interesting in its full generality, we will concentrate on it's specialization to vertex-rooted bipartite maps. This specialization is given by Proposition 14 in [11], and gives us Lemma 1. However, note that the bijection for bipartite maps has been studied before, in almost identical form, in [5]. We will loosely follow the relevant section in [11] for the remained of this section. The only difference is that our bipartite maps, in addition to being rooted at a vertex, also have a marked edge and a marked vertex or face. These properties ensure that the resulting tree is rooted and has a marked vertex, respectively.

Let ϕ denote the specialization of the master bijection that proves Lemma 1. The bijection ϕ consists of a few different steps:

1. Put a canonical orientation on the edges of a map.
2. Apply a local transformation to each of the oriented edges of the map, resulting in a tree with dangling half-edges.
3. Remove the root vertex of the original map.
4. Replace the dangling half-edges by integer weights.

We now go through each step in more detail. For an illustrated example, see Figure 2.3.

The Orientation

We start by choosing a canonical orientation that is well-defined and unique for each map in $\mathcal{M}_n$. The orientation is obtained by directing every edge *away* from the root vertex, and is sometimes called the geodesic orientation. It is well defined since we are working with bipartite maps, so there is no edge whose endpoints are both at equal distance from the root. The orientation is also clearly unique.

The geodesic orientation has several properties that we use. Firstly, note that every face of a map M under the geodesic orientation will have an equal number of clockwise and

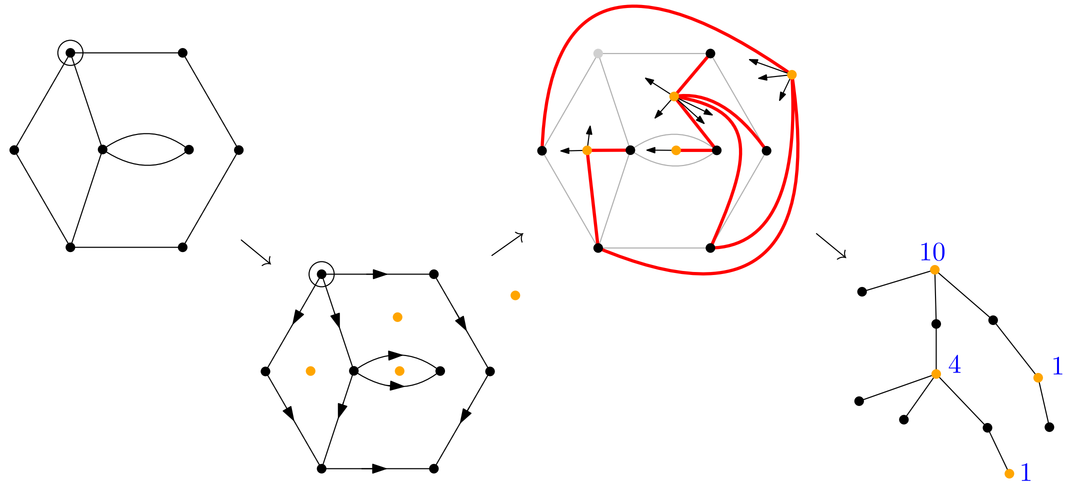


FIGURE 2.3: The master bijection ϕ applied to a map. The additional markings of the map have been left out.

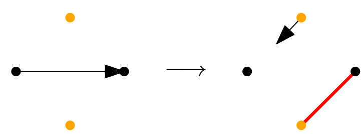


FIGURE 2.4: Local transformation in the master bijection.

counter-clockwise oriented edges around it. This also implies that the orientation is acyclic. Secondly, note that every vertex except the root vertex has at least one edge going into it.

The Local Transformation

The master bijection now proceeds by applying a local transformation to every oriented edge in the map. The local transformation is illustrated in Figure 2.4. On the left, we have a directed edge in M, and the orange vertices represent the (not necessarily distinct) faces incident to the edge on either side. On the right, the result of the transformation is show: we delete the directed edge and insert one new undirected edge, and one dangling half-edge, indicated by an arrow in the figure. By applying the local transformation to each edge of $M \in \mathcal{M}_n$, we obtain a map on black and orange vertices with n undirected edges and n dangling half-edges. We conclude by removing the root vertex of the original map, which is now isolated; the result is $\phi(M)$.

It remains to show that $\phi(M)$ is a plane tree in $\mathcal{T}_n$ which preserved the information of the face degrees in M as we require. Using Euler's formula, we easily show that the number of edges in $\phi(M)$ is one less than the number of vertices. Furthermore, one can use properties of the geodesic orientation to show that $\phi(M)$ is connected, or one can show that it is acyclic. Either way, we conclude that $\phi(M)$ is in fact a plane tree. Furthermore, the marked edge of M roots $\phi(M)$ at a corner of an orange vertex (the conventional way to root plane trees), while the marked vertex of face in M carries over to a marked vertex in $\phi(M)$.

The face degrees of M are encoded in the degrees of the orange vertices in $\phi(M)$. This is because as noted before, every face of degree $2d$ in M is incident to exactly d clockwise oriented edges. These d oriented edges lead to d undirected edges incident to the orange vertex that represents the face. Note also that by the same argument, each orange vertex of degree d has d dangling half-edges; these come from the counter-clockwise oriented edges around the corresponding face. Consider any particular face (the left-most for instance) in Figure 2.3 for an example of the above argument. Note also that since every edge is between a black and orange vertex, the orange vertices are all the vertices at even distance from the root vertex as is required in $\mathcal{T}_n$.

Finally, we replace the dangling half-edges by integer weights. Around an even distance vertex of degree d there are d dangling half-edges; canonically enumerate all possible ways to distribute the d half-edges among the d corners. There are $\binom{2d-1}{d}$ ways to distribute d non-distinguishable half-edges among d distinguishable corners (with repetition). Hence each configuration of dangling half-edges gets a unique integer weight between 1 and $\binom{2d-1}{d}$.

This shows that if $M \in \mathcal{M}_n$ then $\phi(M) \in \mathcal{T}_n$, and ϕ sends maps with face degrees $2d_1, \dots 2d_r$ to trees in $\mathcal{T}_n$ with orange vertices of degrees $d_1, \dots, d_2$. To complete the presentation of Lemma 1, we consider the inverse of ϕ to show that ϕ is a bijection.

The Inverse of the Master Bijection

The inverse of ϕ consists of a few steps:

1. Replace integer weights by dangling half-edges.
2. Complete the dangling half-edges to full edges by leading them to nearest corners of orange vertices in a counter-clockwise fashion; remove the original edges of the tree.
3. Take the dual of the resulting map; root the vertex in the outer face.

See Figure 2.5 for an example, and note that the dual map we see in black at the end is isomorphic to the map we have seen in previous examples, for example in Figure 2.3.

Let us go through step 2 in some more detail. It works by pairing up the n dangling half-edges and the n corners of the orange vertices, and drawing an edge from each dangling half-edge to its corresponding corner. To pair up the half-edges and corners, start with any half-edge and pair it with the first corner in a counter-clockwise traversal of the tree that is not yet paired up. Repeat this process for each dangling half-edge; the order doesn't matter. Figure 2.6 shows an example of one such step on the left, with the traversal and final corner in green.

We will not go through the complete proof that this is in fact the inverse of ϕ. However, for some support of this claim, consider the right of Figure 2.6. This illustration show what

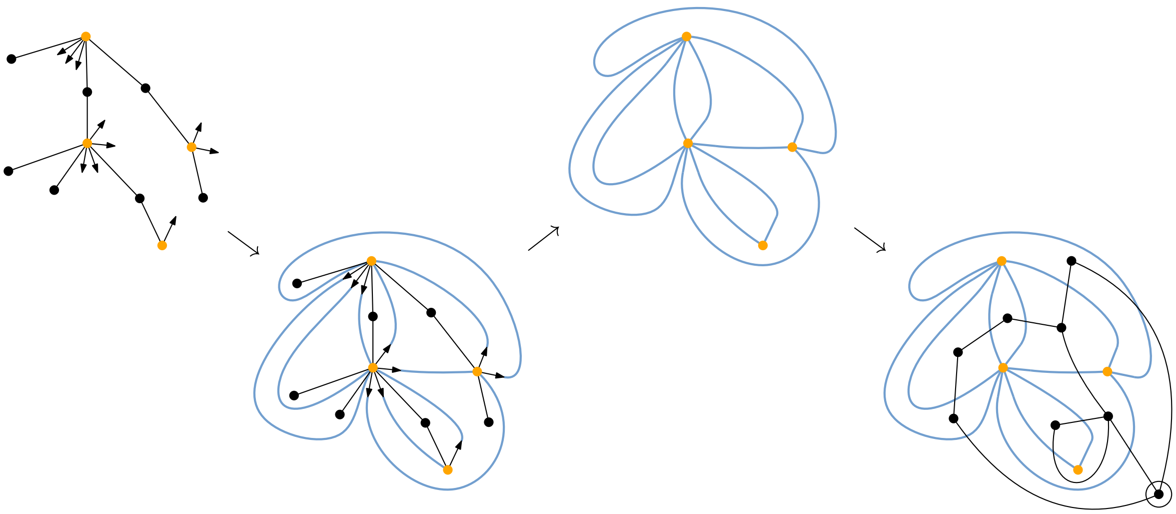


FIGURE 2.5: The inverse of the master bijection applied to a tree in $\mathcal{T}_n$. Additional markings have been left out.

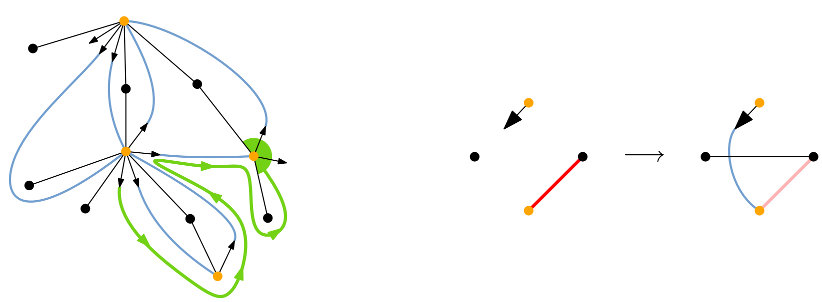


FIGURE 2.6: Left, one step in pairing dangling half-edges and corners in the process of the inverse of ϕ. Right, the "inverse local transformation".

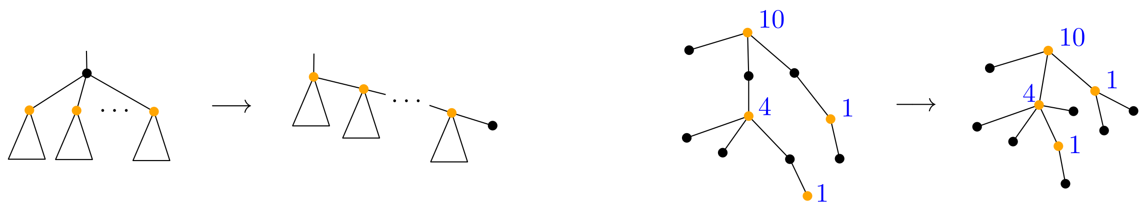


FIGURE 2.7: Left: the local transformation in α. Right: the bijection α applied to our example. Additional markings are left out.

happens to the four vertices that we applied our local transformation to in ϕ, Figure 2.4. Essentially, with the way that we have defined the pairing of dangling half-edges and corners, orange vertices that used to be adjacent as faces (incident to the same edge) will be connected, which by definition gives us the dual of the original map.

2.1.2 From Trees to Binary Sequences

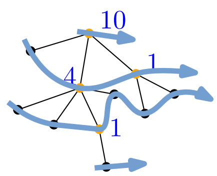
In this section, we present bijections between $\mathcal{T}_n$, $\mathcal{T}_n^*$, $\mathcal{D}_n$ and $\mathcal{B}_n$. The key step here is the bijection from trees ($\mathcal{T}_n^*$) to weighted binary sequences ($\mathcal{D}_n$), which essentially rests on a breadth-first search encoding. However, we first apply one transformation to get the trees in a suitable format for the breadth first search encoding, and afterwords we still need to go from Dyck words to general binary sequences.

Lemma 2 (Local tree transformation)**.** *There is a bijection between $\mathcal{T}_n$ and $\mathcal{T}_n^*$ which maps a tree in $\mathcal{T}_n$ with r orange vertices of degrees $d_1, d_2, \ldots, d_r$ to a tree in $\mathcal{T}_n^*$ with r internal vertices with $d_1, d_2, \ldots, d_r$ children respectively.*

Proof. We present a bijection that we will call α and that proves the lemma. The idea of the bijection α is to apply a local and reversible transformation to each black vertex, which will make all black vertices leaves and all orange vertices internal. The transformation is illustrated in Figure 2.7, together with an example of the bijection α applied to the tree from our previous examples. The local transformation to a black vertex b is to remove all edges incident to b, and create a path out of the (orange) children of b, in the order in which they appear as children. Then, the left-most vertex in the path is connected to the parent of b, while b itself is attached to the right-most vertex in the path, as shown in the figure. The weights on the orange vertices are simply preserved.

The result is that each black vertex becomes a leaf, and each orange vertex must become an internal vertex. Furthermore, it is easily seen that each orange vertex except for the root gains one child. It follows that all orange vertices (including the root) of degree d in $T \in \mathcal{T}_n$ have d children in $\mathcal{T}_n^*$.

The inverse of α is straightforward. Define a path of vertices $v_1, v_2, \ldots, v_k$ in a tree $T^* \in \mathcal{T}_n^*$ where v_k is a leaf to be a *right path* if each v_i is the rightmost child of its parent. Note that any two maximal right paths in T with different endpoints are disjoint. Now, for



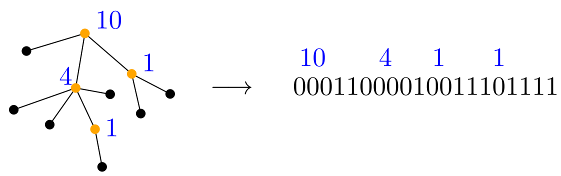


FIGURE 2.8: Left: an example of a the breadth first search order. Right: the bijection β applied to our familiar example. Additional markings are left out.

each maximal right path $v_1, v_2, \ldots, v_k$, replace the edge between v_1 and its parent p with an edge between p and v_k, and make $v_1, \ldots, v_{k-1}$ the children of v_k, in that order. $\square$

We now look at the bijection between $\mathcal{T}_n^*$ and $\mathcal{D}_n$.

Lemma 3 (Breadth first search encoding)**.** *There is a bijection between $\mathcal{T}_n^*$ and $\mathcal{D}_n$ which maps a tree in $\mathcal{T}_n^*$ with r internal vertices with $d_1, d_2, \ldots, d_r$ children respectively to a weighted Dyck word with r runs of 0s of lengths $d_1, d_2, \ldots, d_r$.*

Proof. We present a bijection β that proves the lemma. This is a relatively standard encoding of plane trees as binary sequences. A binary sequence is built up incrementally as we traverse the tree. One way to view this bijection is that we traverse the tree $T^* \in \mathcal{T}_n^*$ by height from top to bottom, and for each vertex v we add as many 0s to our binary string as v has children, followed by a 1. The traversal for our running example is shown on the left in Figure 2.8.

Another way to view this bijection, which may be more familiar to computer scientists, is that we perform a breadth first search in T^* with our queue initially containing the root vertex. Now, every time we add a vertex to the queue, we add a 0 to our binary sequence. Every time we finish an iteration, we add a 1 to our binary sequence. Recall that a breadth first search is an iteration over a queue until the queue is empty. In each iteration we pop a vertex v from the front of the queue, and then add the children of v to the back of the queue.

Either way, we end up with a binary sequence with n 0s and $n + 1$ 1s. The marked vertex in T^* is mapped to a marked 1 in $\beta(T^*)$. Each internal vertex in T^* is mapped to a unique run of 0s in $\beta(T^*)$, so we simply transfer the weights from internal vertices in T^* to runs of 0s in $\beta(T^*)$.

An illustration of β applied to our running example is given on the right in Figure 2.8.

Note that $\beta(T^*)$ is not just any binary sequence: it is Dyck word. Recall that a Dyck word is a binary sequence such that any prefix except the whole sequence contains no more 1s than 0s. The reason that $\beta(T^*)$ is a Dyck word is because each vertex in T^* is responsible for a 0 and a 1 in $\beta(T^*)$, and the 0 is always added first.

The inverse of β is straightforward. $\square$

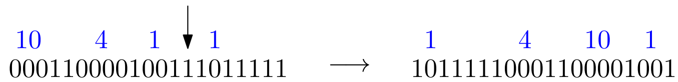


FIGURE 2.9: The bijection γ applied to our running example.

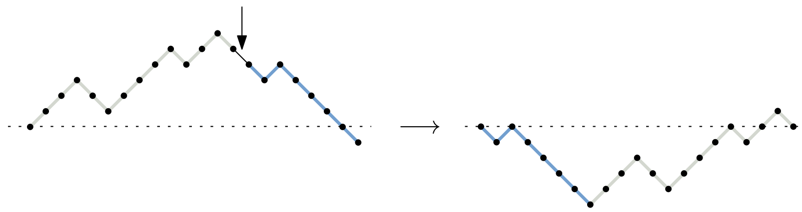


FIGURE 2.10: A more visual example of the bijection γ.

Finally, we consider the bijection between $\mathcal{D}_n$ and $\mathcal{B}_n$. Incidentally, this bijection is otherwise known as providing a way to count Dyck words bijectively, and would tell us that Dyck words are counted by the Catalan numbers. This is not quite the case for the class $\mathcal{D}_n$, since our runs of 0s carry integer weights.

Lemma 4 (Cycle lemma)**.** *There is a bijection between $\mathcal{D}_n$ and $\mathcal{B}_n$ which maps a weighted Dyck word with r runs of 0s of lengths $d_1, d_2, \ldots, d_r$ to a weighted binary sequence with r runs of 0s of length $d_1, d_2, \ldots, d_r$.*

Proof. We present a bijection γ which proves the lemma; this bijection is sometimes called the cycle lemma. The idea behind this bijection is to cut a Dyck word $D \in \mathcal{D}_n$ just after the marked 1, then switch the order of the two resulting halves, and strip the last 1. Clearly γ maps Dyck words in $\mathcal{D}_n$ to binary sequences in $\mathcal{B}_n$. An example of is shown in Figure 2.9.

This bijection creates a $(n+1)$-to-1 correspondence between Dyck words on n 0s and $n+1$ 1s and binary sequences on n 0s and n 1s. The $n+1$ Dyck words corresponding to a single binary sequence are differentiated by the $n+1$ ways to mark a 1 in the Dyck words. For normal (non-weighted) Dyck words, this is actually one of the combinatorial proofs for the formula for the nth Catalan number: $\frac{1}{n+1}\binom{2n}{n}$.

This bijection is easier to understand when we represent Dyck words and binary sequences as lattice walks from $(0,0)$ to $(2n,0)$ or $(2n,-1)$ with the steps $(1,1)$ and $(1,-1)$, such that a 0 corresponds to a $(1,1)$ step while a 1 correspondings to a $(1,-1)$ step. In this context, Dyck words are just walks that never go below the x-axis except for the very last step. The lattice walk version of the bijection γ is shown in Figure 2.10.

The inverse of γ is simple: find the first point at which a binary sequence (lattice walk) reaches its lowest height, and cut the sequence at that point, swap the resulting two halves and insert an extra marked 1 (or $(1,-1)$ step) between the two halves.

Lastly, note that γ preserves runs of 0s. We always cut just after a 1, so a run of 0s is never cut in two. Furthermore, it is impossible to merge two runs of 0s together in the procedure of γ, since a Dyck word $D \in \mathcal{D}_n$ always ends with a 1. Finally, it is also impossible

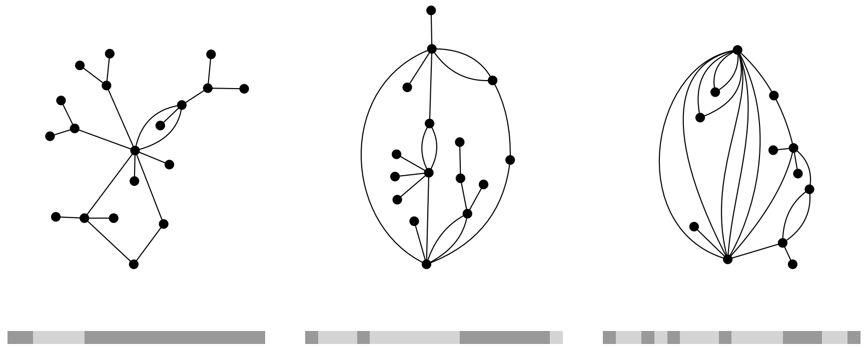


FIGURE 2.11: Three examples of maps with their corresponding compositions. Weights of the compositions have been left out.

to merge two runs of 0s together in the inverse of γ, since we insert a 1 between the two halves when we glue them back together. This way, every run of 0s of length d in $D \in \mathcal{D}_n$ gets mapped to a unique run of 0s in $\gamma(D)$ of length d with the same integer weight. $\square$

Together, Lemmas 1, 2, 3 and 4 prove the bijection between $\mathcal{M}_n$ and $\mathcal{B}_n$.

See Figure 2.11 for the compositions corresponding to the three maps we saw in the introduction. As before, we draw compositions as a sequence of bars, where each bar indicates a part, and the lengths of the bars correspond the sizes of the parts.

Chapter 3

Analysis & Results

In this part, we study the integer compositions $\mathcal{C}_n^{(r)}$, and use Corollary 1 from Chapter 2 to transfer the results back to planar maps. In particular, with Corollary 1, an equivalent statement of Conjecture 1 is that the expected size of the largest part in compositions in $\mathcal{C}_n^{(n^\alpha)}$ is

$$c\,\alpha\,n^{1-\alpha}\log n + d\,n^{1-\alpha} + O(1),$$

where c and d are constants.

Our approach for finding an approximation to the expected size of the largest part in weighted integer compositions in $\mathcal{C}_n^{(r)}$ is to attach a random variable X_i to each of the r parts in a random composition from $\mathcal{C}_n^{(r)}$. The problem is then reduced to find the expected value of the maximum of the X_i, that is, $\mathbb{E}(X)$ where $X = \max\{X_1, \ldots, X_r\}$. However, one difficulty is that the X_i are not independent, since they all have to add up to n.

To deal with this difficulty, we consider compositions $\mathcal{C}^{(r)}$ under a probability distribution called the Boltzmann distribution, introduced in [8]. This is a distribution on compositions of all sizes, not just on $\mathcal{C}_n^{(r)}$. However, the distribution can be chosen such that the expected size of an element under the distribution is n. Also, elements of the same size are assigned the same probability under the Boltzmann distribution.

Instead of find the expected size of the largest part in compositions in $\mathcal{C}_n^{(r)}$, we find the expected size of the largest part in a composition drawn at random from the Boltzmann distribution. Since the Boltzmann distribution is concentrated around compositions of size n, this gives strong evidence for the conjecture.

The advantage of considering compositions under the Boltzmann distribution is that it allows us to regard parts in the composition as *independent*. This means that the random variables $X_1, \ldots, X_r$ are distributed *independently* and *identically* as long as we consider our weighted compositions under the Boltzmann distribution. This is illustrated in Figure 3.1. On the left, we have a number of compositions (the weights have been left out) all of the same size; the sizes of their parts are dependent. On the right, we consider the sizes of the parts as independent, and we get compositions who's sizes vary.

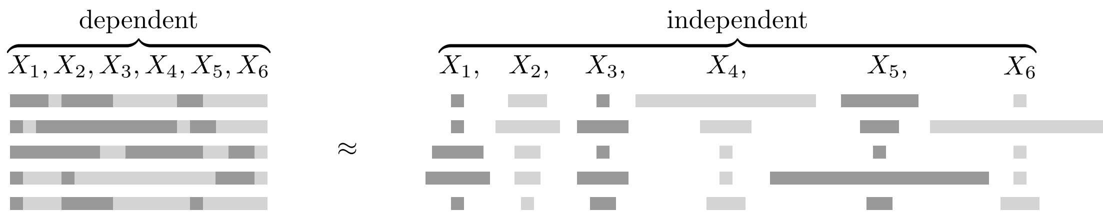


FIGURE 3.1: A comparison between compositions of fixed size, and compositions under the Boltzmann distribution.

In Section 3.1, we give an introduction to the Boltzmann distribution. In Section 3.2, we go on to finding the distribution of the X_i. Next, in Section 3.3 we go on to finding an approximate asymptotic expression for $\mathbb{E}(X)$. Finally, we compare the result with some numerical data from randomly generated compositions in Section 3.4. We are not able to prove Conjecture 1, but we do present reasonable approximations.

3.1 The Boltzmann Distribution

The Boltzmann distribution (here in the context of combinatorics, not mechanical physics) is a discreet probability distribution on a combinatorial class that uses evaluations of the generating function of the class to assign probabilities. It was originally introduced as part of a very efficient method of random generating, see [8]. This presentation of the Boltzmann distribution follows [8].

Definition 6. Let $\mathcal{A}$ be an unlabelled combinatorial family with generating function $A(x)$. The Boltzmann distribution with the positive parameter x is a distribution on the entire class $\mathcal{A}$ which assigns the probability

$$\mathbb{P}_{\mathcal{A}}(\alpha) = \frac{x^{|\alpha|}}{A(x)}$$

to the object $\alpha \in \mathcal{A}$.

Note that this is a proper probability distribution, since

$$\sum_{\alpha \in \mathcal{A}} \frac{x^{|\alpha|}}{A(x)} = \frac{A(x)}{A(x)} = 1.$$

However, since (somewhat unusually) we evaluate the generating function at x, we need x to be between 0 and the radius of convergence of $A(x)$. As a general rule, the distribution is skewed more and more to larger objects as x gets closer and closer to the radius of convergence of $A(x)$. In some cases it is possible to choose x to be *equal* to the radius of convergence.

Let M be the random variable that indicates the size of an object under the Boltzmann distribution. The Boltzmann distribution is clearly uniform on objects with the same size. However, the probability of getting an object of size n is

$$\mathbb{P}(M = n) = \sum_{\substack{\alpha \in \mathcal{A} \\ |\alpha| = n}} \frac{x^{|\alpha|}}{A(x)} = \frac{A_n x^n}{A(x)}.$$

We can then get the expected size of an object under the Boltzmann distribution:

$$\mathbb{E}(M) = \sum_{n=0}^{\infty} n \frac{A_n x^n}{A(x)} = x \frac{A'(x)}{A(x)}.$$

Not only that, we can also obtain the variance of the size of an object under the Boltzmann distribution:

$$\mathbb{V}(M) = x^2 \frac{A''(x)}{A(x)}. \tag{3.1}$$

It is easy to derive a Boltzmann distribution of some class given by a combinatorial specification. For example, consider the product specification $\mathcal{C} = \mathcal{A} \times \mathcal{B}$. Then $C(x) = A(x)B(x)$, and the probability assigned to an element $\gamma = (\alpha, \beta) \in \mathcal{C}$ is

$$\mathbb{P}_{\mathcal{C}}(\alpha, \beta) = \frac{x^{|\alpha|+|\beta|}}{C(x)} = \frac{x^{|\alpha|}}{A(x)} \cdot \frac{x^{|\beta|}}{B(x)} = \mathbb{P}_{\mathcal{A}}(\alpha) \cdot \mathbb{P}_{\mathcal{B}}(\beta). \tag{3.2}$$

The Boltzmann distribution is similarly easy to work out for other composite classes such as $\mathcal{A} \cup \mathcal{B}$ and $\textsc{Seq}(\mathcal{A})$.

Equation (3.2) shows us, by definition, that the components of elements in $\mathcal{C} = \mathcal{A} \times \mathcal{B}$ are *independent* under the Boltzmann distribution. This is the motivation for considering integer compositions under the Boltzmann distribution: the parts of the compositions are now independent.

3.2 $\mathcal{C}^{(r)}$ Under the Boltzmann Distribution

Let us apply the theory of Boltzmann distributions to our weighted compositions $\mathcal{C}^{(r)}$. Take $\mathcal{P}$ to be the combinatorial class of weighted parts, such that $\mathcal{P}$ contains $\binom{2k-1}{k}$ elements of size k. With this definition, we have $\mathcal{C}^{(r)} = \mathcal{P}^r$. By comparison with the generating function of the Catalan numbers (see example of plane trees on page 6), we find that the generating function of $\mathcal{P}$ is

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

and hence

$$C^{(r)}(z) = \left(\frac{1}{2\sqrt{1-4z}} - \frac{1}{2} \right)^r$$

Let X_i be the random variable corresponding to the size of the ith path in compositions in $\mathcal{C}^{(r)}$, under the Boltzmann distribution with parameter x. Then, due to the nature of the Boltzmann distribution, X_1 through X_r are *independent* identically distributed random variables with

$$\mathbb{P}(X_i = k) = \frac{P_k x^k}{P(x)}.$$

Now, we would like to consider integer compositions of size n with r parts. This means that we need to choose the parameter x such that the expected size of a single part is n/r. Let the expected size of a part in the composition be N. Then we have

$$N = \mathbb{E}(X_i) = x\frac{P'(x)}{P(x)} = \frac{2x}{(1-4x)\left(1-\sqrt{1-4x}\right)}. \tag{3.3}$$

Solve this equation for x to get

$$\begin{aligned} x &= \frac{1}{4}\left(1 - \frac{1}{2N} + \frac{\sqrt{8N+1}}{8N^2} - \frac{1}{8N^2}\right) \\ &\sim \frac{1}{4}\left(1 - \frac{1}{2N}\right). \end{aligned} \tag{3.4}$$

Now, we can take a Boltzmann distribution that gives us parts of any expected size we want, as long as we pick x according to the above formula. Note also that the radius of converge of $P(x)$ is $1/4$, and the larger we want our parts, the closer x gets to $1/4$. This is the expected behaviour for a Boltzmann distribution.

With this expression for x in terms of N, we can revisit the probability distribution of X_i, which we now let depend on N. Substituting $x = \frac{1}{4}\left(1 - \frac{1}{2N}\right)$ into equation (3.3) gives the approximate distribution

$$\mathbb{P}(X_i = k) \approx \binom{2k-1}{k}\frac{2}{\sqrt{2N}-1}\left(1 - \frac{1}{2N}\right)^k. \tag{3.5}$$

This expression could be made precise by using the exact instead of asymptotic form of x in equation 3.4.

We would also like to know the variance of the size of the parts. This is given by equation (3.1), and we find the asymptotic expression

$$\mathbb{V}(X_i) \sim 3N^2 + O(N^{3/2}). \tag{3.6}$$

This means that the distribution of the sizes of the parts is somewhat concentrated.

3.3 Approximation of $\mathbb{E}(X)$

Recall that our goal is to evaluate $\mathbb{E}(X) = \mathbb{E}(\max\{X_1, \dots, X_r\})$. However, as we can see in equation (3.5), the distribution of the X_i is not very nice. Our next step is to show that the X_i, when normalized suitably, converge to a continuous limiting distribution which is similar to the exponential distribution (Section 3.3.1). We then find the maximum of r exponential random variables (Section 3.3.2), and use this as an approximation to the maximum of r of our close-to-exponential random variables.

3.3.1 Limiting Distribution of X_i

A limiting distribution is defined in terms of the cumulative distribution function, who's definition we quickly recall.

Definition 7. The *cumulative distribution function* (cdf) of a random variable W is defined as $F(u) = \mathbb{P}(W \leq u)$. In general, a function F is a cdf of some distribution if F is non-decreasing, $\lim_{x\to-\infty} F(x) = 0$ and $\lim_{x\to\infty} = 1$. If $f(u)$ is the probability density function of some distribution, then the cdf of that distribution is $F(u) = \int_{-\infty}^{u} f(s)ds$.

We now present the concept of a limiting distribution, as seen in Part C of *Analytic Combinatorics* [10].

Definition 8. A sequence of probability distributions $\delta_1, \delta_2, \delta_3, \dots$ with cumulative distribution functions $F_1, F_2, F_3, \dots$ has a *limiting distribution* δ if the sequence $F_1, F_2, F_3 \dots$ converges pointwise to a limit function F, and F is a cumulative distribution function. Then the limiting distribution δ is defined as having the cumulative distribution function F.

We consider the sequence of distributions of X_i/N indexed by N, as N tends to infinity. Here, the $1/N$ factor is a normalization factor since the expected value of X_i is N. This way, $\mathbb{E}(X_i/N) = 1$.

Lemma 5. *The distribution of X_i/N tends to a continuous limiting distribution as N tends to infinity. The probability density function of this limiting distribution is*

$$f(u) = \frac{1}{\sqrt{2\pi}} \exp(-u/2) \frac{1}{\sqrt{u}}.$$

Proof. To find the limiting distribution, we need the cumulative distribution function of X_i/N, which is

$$F_N(y) = \mathbb{P}\left(\frac{X_i}{N} \leq y\right) = \mathbb{P}\left(X_i \leq yN\right).$$

We express this as the summation

$$F_N(y) = \sum_{k=0}^{yN} \mathbb{P}(X_i = k).$$

Now, with the expression for $\mathbb{P}(X_i = k)$ from equation (3.5), and using Stirling's approximation on the $\binom{2k-1}{k}$ term, we obtain

$$F_N(y) \approx \frac{1}{\sqrt{2\pi N} - \sqrt{\pi}} \sum_{k=0}^{yN} \left(1 - \frac{1}{2N}\right)^k \frac{1}{\sqrt{k}}.$$

To evaluate the limit $\lim_{N\to\infty} F_N(y)$, we regard the above summation as a Riemann sum representing an integral from 0 to y. For this Riemann summation, the interval $(0, y)$ is partitioned into yN intervals of length $\Delta u = 1/N$. With $u_k = k/N$, we rewrite

$$F_N(y) \approx \frac{1}{\sqrt{2\pi}} \sum_{k=0}^{yN} g_N(u_k)\Delta u$$

where

$$g_N(u) \sim \frac{1}{1 - \sqrt{1/(2N)}} \left(\left(1 - \frac{1}{2N}\right)^{2N}\right)^{u/2} \frac{1}{\sqrt{u}}.$$

Now, the sequence of functions g_N indexed by N converges uniformly to

$$g(u) = \exp(-u/2)\frac{1}{\sqrt{u}}$$

as N tends to infinity. Since the convergence is uniform and we are taking a Riemann sum over a bounded interval, we have

$$F(y) = \lim_{N\to\infty} F_N(y) = \frac{1}{\sqrt{2\pi}} \int_0^y \exp(-u/2)\frac{1}{\sqrt{u}}du.$$

It follows that the probability density function of the limiting distribution is

$$f(u) = \frac{1}{\sqrt{2\pi}} \exp(-u/2)\frac{1}{\sqrt{u}} \tag{3.7}$$

as desired. □

Let $Y_1, \ldots, Y_r$ be r random variables distributed with probability density function $f(u)$ as given in the lemma. With Lemma 5, we then have that $X_i/N \to Y_i$ as N tends to infinity, and hence

$$\lim_{N\to\infty} \mathbb{E}\left(\max\left\{\frac{X_1}{N}, \ldots, \frac{X_r}{N}\right\}\right) = \mathbb{E}(\max\{Y_1, \ldots, Y_r\}).$$

Now, we may factor out the $\frac{1}{N}$ term from the left hand side of the above equation to get

$$\mathbb{E}(X) = \mathbb{E}(\max\{X_1, \ldots, X_r\}) \sim N \cdot \mathbb{E}(\max\{Y_1, \ldots, Y_r\}). \tag{3.8}$$

That is, we have reduced the problem to find the expected value of the maximum of the random variables $Y_1, \ldots, Y_r$. We do not have a closed form asymptotic expression for the maximum of $Y_1, \ldots, Y_r$, but we in the next section we solve the related problem of finding the maximum of r exponential random variables.

3.3.2 Maximum of Exponential Random Variables

The probability density function $f(u)$ looks a lot like the probability density function of the exponential distribution, which we define here:

Definition 9. The exponential distribution with parameter λ is a continuous probability distribution taking positive values with probability density function $h(u) = \lambda \exp(-\lambda u)$ and cumulative distribution function $H(u) = 1 - \exp(-\lambda u)$.

Lemma 6 (Maximum of exponential random variables)**.** *Let $E_1, E_2, \ldots, E_r$ be i.i.d. exponential random variables with parameter λ. Then*

$$\mathbb{E}(\max\{E_1, E_2, \ldots, E_r\}) = \frac{1}{\lambda} H_r$$

where H_r is the rth harmonic number.

Proof. The cumulative distribution function of the random variable $E = \max\{E_1, E_2, \ldots, E_r\}$ is $(1 - \exp(-\lambda u))^r$. Then, since E only takes positive values, we may use the formula

$$\mathbb{E}(E) = \int_0^\infty 1 - (1 - \exp(-\lambda u))^r du.$$

Upon expanding $(1 - \exp(-\lambda u))^r$ using the binomial theorem and evaluating the integral, we get

$$\mathbb{E}(E) = -\sum_{k=1}^{r} \binom{r}{k} (-1)^k \frac{1}{k\lambda}.$$

By an identity for binomial coefficients, this equals H_r/λ as desired. $\square$

Also recall that H_r/λ is asymptotically

$$\frac{1}{\lambda} H_r = \frac{1}{\lambda} \log r + \frac{\gamma}{\lambda} + O\left(\frac{1}{r}\right)$$

where $\gamma \approx 0.577$ is the Euler-Mascheroni constant.

The probability density function $f(u)$ from equation (3.7) is close in behaviour to the probability density function of an exponential random variable with parameter $\lambda = \frac{1}{2}$. If $E_1, \ldots, E_r$ are r independent exponential random variables with parameter $\frac{1}{2}$, then we could conjecture that

$$\mathbb{E}(\max\{Y_1, \ldots, Y_r\}) \approx \mathbb{E}(\max\{E_1 \ldots, E_r\}) = 2H_r. \tag{3.9}$$

In the next section, we will see that this approximation is not very good, but possibly only off by a constant or asymptotically small factor.

3.4 Results

To summarize, we now know that the expected size of the largest part in a composition with r parts, each of expected size N, is asymptotically equivalent to N times the expected maximum of r independent random variables $Y_1, \ldots, Y_r$ with probability density function $f(u)$ given in equation (3.7). Note that this is only rigorous when r is fixed and N tends to infinity.

If we let $F(u)$ be the cumulative distribution function defined by

$$F(u) = \int_0^u f(s)ds,$$

Then we have

$$\mathbb{E}(\max\{Y_1, \ldots, Y_r\}) = \int_0^\infty 1 - F(u)^r du.$$

We are interested in weighted compositions in $\mathcal{C}_n$ with n^α parts, so we take $N = n^{1-\alpha}$ and $r = n^\alpha$ in equation (3.8) to get

$$\mathbb{E}(X) \sim n^{1-\alpha}\mathbb{E}(\max\{Y_1, \ldots, Y_{n^\alpha}\}) = n^{1-\alpha} \int_0^\infty 1 - F(u)^{n^\alpha} du. \tag{3.10}$$

However, since $r = n^\alpha$ is not fixed in this case, equation (3.10) only provides an approximation and is not rigorous. Still, it is a plausible approximation especially when $\alpha < 1/2$ in which case $N \gg r$. We do not have a closed form (asymptotic) expression for the quantity in equation (3.10), but it is possible to use Maple to obtain numerical values.

Another possible possible approximation comes from equation (3.9), with $N = n^{1-\alpha}$ and $r = n^\alpha$:

$$\mathbb{E}(X) \approx 2\,\alpha\, n^{1-\alpha} \log n + 2\,\gamma\, n^{1-\alpha}. \tag{3.11}$$

The two approximations in equations (3.10) and (3.11), are shown in dashed lines in Figure 3.2. Furthermore, the points connected by solid lines are the results of taking the average size of the largest part of 200 integer compositions generated randomly from the Boltzmann distribution. Each point corresponds to compositions generated around a different size. In Figure 3.2, we have taken $\alpha = 0.3$.

As we can see, the integral expression from equation (3.10) seems to be very accurate, and agrees with the data from the generated weighted compositions. However, the approximation in equation (3.11) is less accurate, which we expected. Still, it seems plausible that the approximation in (3.11) may be accurate to the first term, but off by a constant factor, or possibly a factor that is asymptotically small (e.g. off by a factor of $\log n$ or $\log\log n$). Hence Conjecture 1.

Recall that Corollary 1 states that the expected degree of the largest face in maps in $\mathcal{M}_n^{(r)}$ is twice the expected size of the largest part in weighted compositions in $\mathcal{C}_n^{(r)}$. Hence, we conclude that the expected degree of the largest face in maps in $\mathcal{M}$ under the Boltzmann

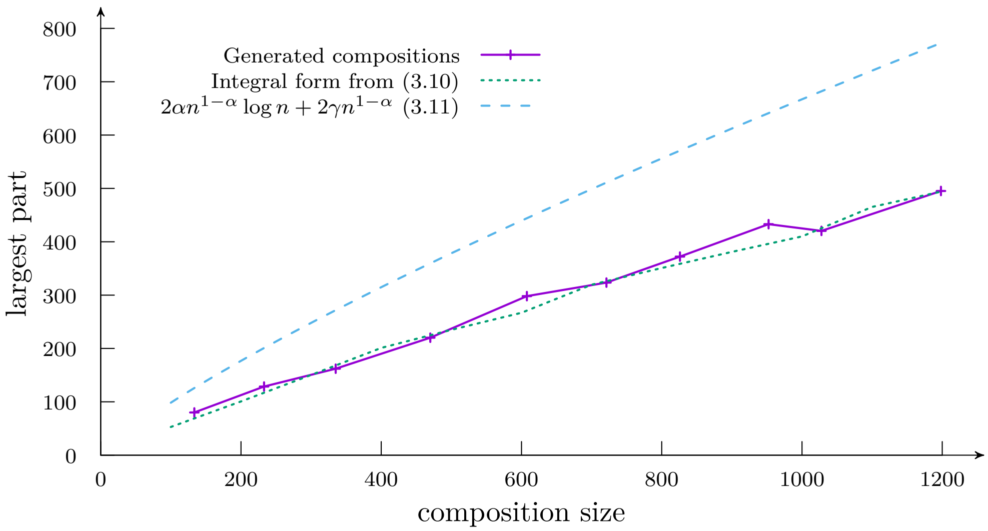


FIGURE 3.2: A comparison between data from generated compositions, and two approximations. Here $\alpha = 0.3$.

distribution with around n edges and n^α faces tends to

$$2n^{1-\alpha}\int_0^\infty 1 - F(u)^{n^\alpha} du$$

for large n. Moreover, as remarked before, since almost all maps are asymmetric [18], the above formula holds not only for maps in $\mathcal{M}$ (which are marked in a number of ways) but also for unmarked bipartite maps under the Boltzmann distribution with around n edges and n^α faces.

Finally, note again that we have been studying maps under a Boltzmann distribution centred around n in this section, not maps with exactly n edges. How valid this approximation is, depends on how concentrated the Boltzmann distribution is. Recall that we found, in equation (3.6), that the variance of the size of a part in $\mathcal{P}$ (equivalently, a single face in a map) under the Boltzmann distribution is of the order $3N^2$, where N is the expected size of the part. Hence, the variance of the size of a composition in $\mathcal{C}(r)$ is

$$\mathbb{V}(X_1 + X_2 + \cdots + X_r) \sim 3rN^2 = 3n^{2-\alpha},$$

where we have taken $N = n^{1-\alpha}$ and $r = n^\alpha$. Now, the standard deviation of the size of integer compositions under the Boltzmann distribution is in the order of $\sqrt{3}\cdot n^{1-\alpha/2}$, which is much smaller than n. Hence, the Boltzmann distribution is concentrated.

Chapter 4

Conclusion

The goal of this thesis is to study the expected degree of the largest face of bipartite planar maps. In Chapter 2, we used the bijection from Theorem 1 to transfer the problem to that of finding the expected size of the largest part in a type of weighed integer compositions. In Chapter 3, we worked with the integer compositions to find the approximate formula given in Conjecture 1.

The main difficulty with our problem is finding a closed form for the expected value of the maximum of r random variables $Y_1, \ldots, Y_r$, and in general formalizing the arguments made in Chapter 3. Many approximations are made, and though they are plausible, they are sometimes hard to justify. This is because we are working with formulas that include two different parameters (expected size of parts and number of parts) that both go to infinity. This is the price we pay for not considering, say, maps with n edges and a constant number of faces. Moreover, we did not rigorously prove that the answer to our questions is the same for maps on n edges, and maps considered under the Boltzmann distribution centred around n. However, this too seems like a plausible proposition, since the Boltzmann distribution is concentrated with a variance of the order $n^{2-\alpha}$.

In future work, the first priority would be to work out a closed form asymptotic expression for equation (3.10); as seen from Figure 3.2 this would give us a much sharper and better justified conjecture than we have now. Continuing from the work in this thesis, this may be possible with additional techniques from analysis and probability theory. Following this, more effort could be spent on making some of the arguments in Chapter 3 more rigorous.

Apart from that, there are related problems that may yield to a strategy similar to the one presented here. In particular, using the bijection presented in this thesis, it could be feasible to investigate additional properties of bipartite maps with few faces. Interesting additional properties include the expected degree of the smallest face, the expected distance from a vertex to the root, or the number and size of 2-connected components in the map.

Bibliography

[1] Olivier Bernardi and Éric Fusy. "A bijection for triangulations, quadrangulations, pentagulations, etc". *J. Combin. Theory Ser. A* 119.1 (2012).

[2] Olivier Bernardi and Éric Fusy. "Unified bijections for maps with prescribed degrees and girth". *J. Combin. Theory Ser. A* 119.6 (2012).

[3] Mireille Bousquet-Mélou and Gilles Schaeffer. "The degree distribution in bipartite planar maps: applications to the Ising model". *ArXiv Mathematics e-prints* (Nov. 2002).

[4] J. Bouttier, P. Di Francesco, and E. Guitter. "Geodesic distance in planar graphs". *Nuclear Phys. B* 663.3 (2003).

[5] J. Bouttier, P. Di Francesco, and E. Guitter. "Planar maps as labeled mobiles". *Electron. J. Combin.* 11.1 (2004).

[6] Guillaume Chapuy, Éric Fusy, Omer Giménez, Bojan Mohar, and Marc Noy. "Asymptotic enumeration and limit laws for graphs of fixed genus". *J. Combin. Theory Ser. A* 118.3 (2011).

[7] Guillaume Chapuy, Éric Fusy, Omer Giménez, and Marc Noy. "On the diameter of random planar graphs". *Combin. Probab. Comput.* 24.1 (2015).

[8] Philippe Duchon, Philippe Flajolet, Guy Louchard, and Gilles Schaeffer. "Boltzmann samplers for the random generation of combinatorial structures". *Combin. Probab. Comput.* 13.4-5 (2004).

[9] Philippe Flajolet, Zhicheng Gao, Andrew Odlyzko, and Bruce Richmond. "The distribution of heights of binary trees and other simple trees". *Combin. Probab. Comput.* 2.2 (1993).

[10] Philippe Flajolet and Robert Sedgewick. *Analytic combinatorics.* Cambridge University Press, Cambridge, 2009.

[11] Éric Fusy. "Combinatoire des cartes planaires par meta-bijection". Habilitation thesis. Laboratoire d'informatique de l'Ecole Polytechnique, 2015.

[12] Zhicheng Gao and Nicholas C. Wormald. "The distribution of the maximum vertex degree in random planar maps". *J. Combin. Theory Ser. A* 89.2 (2000).

[13] Omer Giménez and Marc Noy. "Asymptotic enumeration and limit laws of planar graphs". *J. Amer. Math. Soc.* 22.2 (2009).

[14] Benjamin Jacquard and Gilles Schaeffer. "A bijective census of nonseparable planar maps". *J. Combin. Theory Ser. A* 83.1 (1998).

[15] Svante Janson, Donald E. Knuth, Tomasz Luczak, and Boris Pittel. "The birth of the giant component". *Random Structures Algorithms* 4.3 (1993). With an introduction by the editors.

[16] Jean-François Marckert and Abdelkader Mokkadem. "Limit of normalized quadrangulations: the Brownian map". *Ann. Probab.* 34.6 (2006).

[17] A. Meir and J. W. Moon. "On the maximum out-degree in random trees". *Australas. J. Combin.* 2 (1990). Combinatorial mathematics and combinatorial computing, Vol. 2 (Brisbane, 1989).

[18] L. B. Richmond and N. C. Wormald. "Almost all maps are asymmetric". *J. Combin. Theory Ser. B* 63.1 (1995).

[19] Gilles Schaeffer. "Bijective census and random generation of Eulerian planar maps with prescribed vertex degrees". *Electron. J. Combin.* 4.1 (1997).

[20] Gilles Schaeffer. "Conjugaison d'arbres et cartes combinatoires aléatoires". PhD thesis. University of Bordeaux, 1998.

[21] Neil James Alexander Sloane. *Online Encyclopedia of Integer Sequences.* 2016. URL: `https://oeis.org/` (visited on 12/07/2016).

[22] W. T. Tutte. "A census of planar maps". *Canad. J. Math.* 15 (1963).

[23] W. T. Tutte. "On the enumeration of planar maps". *Bull. Amer. Math. Soc.* 74 (1968).

[24] Hassler Whitney. "Congruent Graphs and the Connectivity of Graphs". *Amer. J. Math.* 54.1 (1932).

[25] A. Zvonkin. "Matrix integrals and map enumeration: an accessible introduction". *Math. Comput. Modelling* 26.8-10 (1997). Combinatorics and physics (Marseilles, 1995).