Graph enumeration and random graphs

Author: Lander Ramos [1]
Advisor: Marc Noy [2]

Departament de Matemàtiques
Universitat Politècnica de Catalunya [3]

January 14, 2017

[1]E-mail: landertxu@gmail.com
[2]E-mail: marc.noy@upc.es
[3]Jordi Girona 1–3, 08034 Barcelona, Spain.

Acknowledgements

My first words of gratitude indubitably go to my advisor, Marc Noy. Thank you for trusting me to start this project, for working with me in many interesting problems, and for helping me entering the world of mathematics, by introducing many people with similar research interests to me. Also, I would like to thank Juanjo Rué for sharing so many other interesting problems with me, and for inviting me to Berlin, where we started working in a a project that eventually would become an important part of my thesis. Thanks for your patience and clear explanation of all my doubts.

Special mention to the professors that received me during my short stays in Paris an Vienna. They helped me a lot both from the academic point of view and in the process of adapting to new cities for me. Thanks to Vlady Ravelomanana, for so many mornings discussing very interesting problems. And thanks a lot to Michael Drmota, who very patiently helped me to understand so many technical details that I was unable to do by my own. An important part of my thesis is thank to you.

I am convinced that this thesis would have not been possible without Guillem Godoy, who introduced me in the world of research when I was just an undergraduate student. And, of course, the work of Josep Grané looking for potential talents was also indispensable. Thanks!

Como parte de la experiencia del doctorado, me quedo con mis comañeros de despacho, con los almuerzos con el grupo de LSI, con Carles y Adrià, con el grupo de Teoria de Nombres, con Elisa, con las innumerables charlas apasionantes durante la pausa del café con Víctor, y con mis compañeros de piso durante estos años, Narcís, Jordi, Sergi, Miquel y Rosa. Sin duda habéis ayudado a que la experiencia sea inmejorable.

Acabo agradeciendo a mis padres que siempre estén ahí. No solo cuando las cosas parecen ir bien, sino también cuando todo parece que se derrumba y saben encontrar las palabras de ánimo perfectas. I gràcies al Pau, que també hi és sempre, per la seva paciència i comprensió quan li havia de dedicar hores i hores a la tesi, o quan el doctorat em feia allunyar-me de Barcelona durant llargs períodes. Moltes gràcies!

Introduction

In this thesis we use analytic combinatorics to deal with two related problem: graph enumeration and random graphs from constrained graph families, which are particular cases of combinatorial classes, defined as (infinite) sets, together with a positive integer size function such that the number of elements of each size is finite. We are interested in drawing a general picture of some graph families by determining, first, how many elements are there of a given possible size (graph enumeration), and secondly, what is the typical behaviour of an element of fixed size chosen uniformly at random, when the size tends to infinity (random graphs).

More precisely, the classes we are interested in are graphs of a given kind, like planar graphs with vertex degree constraints. So first we give an asymptotic estimate of the total number of graphs of such classes of a given size n, when n tends to infinity. And moreover, we evaluate some parameters, like the number of edges, of a graph of size n chosen uniformly at random, when n tends to infinity as well. This means that for those classes we have an estimate of the number of graphs, as well as its typical behaviour. The reason we compute an estimate, and not the exact number of graphs, or the exact value of a parameter, is that computations are typically very difficult, or even unfeasible, but we have powerful tools that give precise asymptotic estimates when the size of the objects is large enough.

Most of the tools used in this thesis can be found in [5]. Actually, many of the papers that inspired this thesis were also based in those results, and consist on related results, concerning classes of graphs like planar graphs or series-parallel graphs, and parameters like the number of edges or the degree distribution. In order to use these results, we need to work with the *generating function* of the corresponding combinatorial class, which is a complex function associated to the class.

In Chapter 1 we introduce the notation used in the rest of the thesis. In addition some definitions regarding the *Symbolic Method* are included, in particular all the details about generating functions. The chapter is complemented with the main results that we use in this thesis, namely analytic theorems and combinatorial techniques.

In Chapter 2 we study planar graphs with minimum degree 2 and 3. This is one of the natural sequels of the work of O. Giménez and M. Noy [8], where the authors gave an enumeration of planar graphs, as well as many limit laws of random planar graphs. In particular, they describe the generating function of planar graphs implicitly as the solution of a system of explicit equations, since an explicit expression is not likely to exist. The knowledge of a generating function leads to many interesting properties regarding the combinatorial class that it counts. One of this properties is the asymptotic enumeration of objects in the class, in this case planar graphs. It is obtained from the location of the main singularity of the generating function, and the behaviour of the generating function near this singularity. After applying the corresponding analytic theorems and techniques, the authors of [8] obtained an asymptotic estimate of the number of planar graphs, given by

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

where g and γ are well defined analytic constants, and $\gamma \approx 27.2269$. Moreover, if we enrich the generating function by adding a variable that encodes edges, we obtain properties regarding the behaviour of the corresponding parameter, in this case edges, in random planar graphs. In particular they proved that the number of edges in a random planar graph with n vertices is asymptotically normal, and the mean μ_n and variance σ_n^2 satisfy

$$\mu_n \sim \kappa n, \qquad \sigma_n^2 \sim \lambda n,$$

where $\kappa \approx 2.2133$ and $\lambda \approx 0.4303$ are computable analytic constants. There are similar results for the degree sequence or the number of connected components, obtained by enriching the generating function with the corresponding parameter.

Many of our results for planar graphs with minimum degree 2 and 3 are similar to the results obtained for planar graphs, and in fact we need to use the corresponding generating function, after some modifications to remove vertices of degree 1 and 2. For this removal we need the concept of the *core* and *kernel* of a graph, defined as follows: the core of a graph is obtained by repeatedly removing vertices of degree one, and the kernel of a graph without vertices of degree one is obtained by replacing maximal paths of vertices of degree 2 with edges. We use this concept to obtain different parameters from planar graphs, like the expected size of the core, or the size of the trees attached to it, which roughly corresponds to the part of the graph not belonging to the core. We also use extremal techniques to compute the expected size of the largest of those trees.

As a warm-up we obtain the same results for embedded maps with minimum degree 2 and 3, using the fact that we have explicit expressions for their generating functions. A map, as opposed to planar graphs, is associated with an embedding in the sphere, and this gives an explicit method to obtain its generating function, and therefore to count them. These results are interesting by themselves, but we take advantage of the existence of explicit expressions to show how the combinatorics works, paying less attention to the analysis, which in the case of graphs contains many technical difficulties due to the non-existence of an explicit expression for its generating function, and we have to solve them using involved ad-hoc techniques.

In Chapter 3 we venture into a vast field: counting subgraphs in random graphs. The most general version of the problem would be: given a class of graphs $\mathcal{G}$, and a graph H, give as much information as possible about the number of occurrences of H as a subgraph of a graph in $\mathcal{G}$, taken uniformly at random. This general version is out of the reach with our current tools, so instead we tackle some specifical subcases of the problem, namely a particular and easy enough class of graphs $\mathcal{G}$, a particular simple small enough (but non trivial) subgraph H, or partial results about the behaviour of the number of occurrences in families of classes of graphs. Note that in Chapter 2 we show how to deal with the problem of counting edges in planar graphs, so the subgraph H should be more interesting than a single edge.

One of the results that inspired our work is the paper by C. McDiarmid [16], who counted the number of pendant copies of graphs in random graphs from classes of graphs satisfying a general enough property that includes a wide variety of graph constrained classes. Here, a pendant copy of a graph H in a graph G is a copy of H as a subgraph of G joined to the rest of the graph by a bridge edge, i.e., an edge such that after removing it the number of connected components increases. McDiarmid proved that if the class of graphs satisfy a number of natural properties, then a large enough graph in the class, taken uniformly at random among all the graphs of the same size, contains a linear number of pendant copies of any connected graph in the class. Since all the pendant copies are at the same time copies as subgraphs, this gives a lower bound for the number of occurrences of graphs for classes of graphs satisfying those conditions. However, the number of occurrences of a graph of k vertices as a subgraph of a graph of n vertices could be up to $O(n^k)$.

In the first part of our work we deal with subcritical graphs classes, as defined in [12]. There is a precise definition of a subcritical graph class, but its main characterization is more interesting: a subcritical class of graphs is such that the size of the largest 2-connected component of a graph in the class is typically of logarithmic size with respect to the number of vertices, as opposed to critical classes of graphs, which contain a 2-connected component of linear size. Note that a 2-connected component is a maximal 2-connected subgraph, and a 2-connected graph is a connected graph that does not contain a cut vertex. We study the number of 2-connected graphs in subcritical classes of graphs, taking advantage of the fact that the 2-connected components are small, and that a 2-connected subgraph cannot traverse cut vertices. We prove that the number of occurrences of such subgraphs follows a Gaussian limit law with linear expectation.

In the second part of the chapter we compute the exact value of the constants for given subgraphs, triangles and quadrangles, in a given subcritical class of graphs, the class of series-parallel graphs. Our previous result already proves that the number of triangles and quadrangles follows a gaussian law with linear expectation, but we compute the exact value of the constants, i.e., we prove that the expected number of triangles in a connected series-parallel graph with n vertices is asymptotically μn, with $\mu \approx 0.39418$. Our combinatorial and analytic tools also give enumerative results on the number of series-parallel graphs without triangles almost “for free”, so we compute these values. Finally, we prove a Gaussian law with linear expectation on the number of occurrences of triangles and quadrangles in 2-connected series parallel graphs, which cannot be proven by means of the general theorem of the previous section.

In Chapter 4 we enumerate graphs where the degree of every vertex belongs to a fixed subset of the natural numbers. More formally, given $\mathcal{D} \subset \mathbb{N}$, we give an asymptotic estimate, that depends on $\mathcal{D}$, on the number of graphs with n vertices where the degree of every vertex belongs to $\mathcal{D}$. Many particular cases have been already studied. When $\mathcal{D}$ consists of a single element k, we deal with the family of k-regular graphs, which were counted in [63]. More complex subsets have been also considered, like graph with minimum degree $\delta \geq k$ [10], or classes of Eulerian graphs, defined as graphs where all vertices have even degree [25] [73]. As expected, our result matches all the particular cases when we specialize our formula. In the case of Eurlerian graphs, our formula provides enumeration for a given number of vertices and edges, strengthening the previously existing results, which only take vertices into account.

The techniques used in this chapter are slightly different to the ones from the previous chapters. Some of the tools we use do not behave well when the generating function of the class of graphs has radius of convergence 0, as is the case for general graphs, even if we constrain the degree of the vertices. Instead, we work with the so-called configuration model. In this model, we consider that each vertex has a number of half-edges equivalent to its degree, and we match all the possible half-edges between them. This allows us to set the degree of every vertex in advance, but the main drawback is that we allow all the possible matchings between half-edges, which leads to the possibility of loops and multiple edges, i.e., when we match two half-edges belonging to the same vertex, or when we match more than one pair of half-edges between the same pair of vertices. In other words, the resulting graph is not necessarily a simple graph, but a multigraph.

Since all the previous results were related to simple graphs, and enumeration does not behave well with multigraphs, due to the existence of an infinite number of multigraphs with a fixed number of vertices, we have to find a way to get rid of the loops and multiple edges. Actually, since we work with asymptotic enumeration, we can leave some multigraphs if their density is low enough so that the asymptotic estimate is not affected. The good news are that the probability of having a triple edge or a double loop tends to zero when the size of the graph is large enough, so we only have to consider double edges and single loops. We deal with this situation by means of the inclusion-exclusion principle, inspired in a particular approach given in [5].

All the chapters of this thesis where written as papers for different journals and conferences. However, not all of them have been published as for today. At the time of finishing writing this thesis, the status of the different chapter are as follows:

- Chapter 2 was written to be published as a paper. An extended abstract was accepted in Eurocomb 2013 [46]. We are currently preparing an extended version to submit to a journal.

- Chapter 3 has been accepted in Random Structure and Algorithms [59]. We are currently waiting for its publication.

- Chapter 4 has been published as a preprint in the proceedings of the Workshop on Analytic Algorithmics and Combinatorics 2016 [30]. We are currently working in a longer version to be submitted to a journal.

Contents

Chapter 1

Background and definitions

This chapter is devoted to introduce all the contents needed to follow the rest of the thesis. It includes both the main definitions that appear in other chapters, and the main theorems already existing in the literature. For the sake of clarity, the concepts are divided into three sections. In Section 1.1, the combinatorial structures we will deal with are defined, and the notation for them is presented. In Section 1.2, we show how to use analytic tools to extract properties from combinatorial classes, and we state the main theorems that we use. In Section 1.3 we show well-known relations between combinatorial classes that can be translated into analytic relations so that we have the equations needed to apply the theorems of Section 1.2.

1.1 Structures

In this section we define and fix notation for the structures that we use, mostly graphs and maps. In Section 1.1.1 we define what a graph is, and we precisely state which kind of graphs we are working with. In Section 1.1.2 we define the other main structure we work with: maps. Finally, in Section 1.1.3 we define properties in graphs that define graph classes that we will count later.

1.1.1 Graphs

The main combinatorial objects in this thesis are graphs, in particular labelled graphs. A *labelled graph* is a pair $(V(G), E(G))$, where $V = V(G)$ is the set of *vertices* and $E(G)$ is a set of unordered pairs of different vertices in V, called edges. The two vertices of an edge are called *endpoints*, and two vertices v and w are *adjacent* if and only if there exists an edge in the graph whose endpoints are v and w. All the vertices are distinct, and we will use the set $[n] = \{1, 2, \ldots, n\}$ to label the n vertices of a labelled graph. Since all the graphs in this thesis are labelled, from now on we just say *graph* instead of labelled graph. Sometimes, a graph is called *simple* graph to distinguish it from multigraphs. A *labelled multigraph*, or, in this thesis, a *multigraph* is a pair $(V(G), E(G))$, where $V = V(G)$ is a set of vertices, as in the case of simple graphs, and $E(G)$ is a finite multiset of unordered pairs of (non necessarily different) vertices in V. In other words, a multigraph allows both *multiple edges* and *loops*. A multiple edge is a set of edges with the same endpoints, and a loop is an edge where the endpoints are equal. Note that this definition allows multiple loops. Also, unless stated differently, the edges of a multigraph are not distinguishable. In this thesis we only study properties related to simple graphs, but some proofs require multigraphs as intermediate steps, which in some cases are easier to work with.

A *subgraph* of a graph G is a graph H such that the set of vertices of H is a subset of the set of vertices of G, and the set of edges of H is a subset of the set of edges of G, whose endpoints are vertices in H. A *path* in a graph is a finite sequence of different vertices of the graph $v_1, v_2, \ldots, v_n$ such that for every $i \in \{1, \ldots, n-1\}$ there is an edge between v_i and v_{i+1}. A *cycle* of a graph is a path such that there is an edge between v_1 and v_n and $n \geq 3$. A graph G is *connected* if it is not empty and there is a path between each pair of vertices of G. A *connected component* of a graph G is a maximal connected subset of vertices of G. A graph G is *k-connected* if it has at least k vertices and there is no subset of $k-1$ vertices such that after removing them the graph becomes not connected. Note that the definition of connected is equivalent to the definition of 1-connected. A 2-*connected component* or a *block* of a graph G is a maximal 2-connected subgraph of G. A *tree* is a connected graph without cycles. A *rooted tree* is a tree with a distinguished vertex, which is called the *root* of the tree. A *minor* of a graph G is a subgraph H whose vertices matches disjoint subsets of the vertices of G, and two vertices v and w of H are adjacent if an only if there is a vertex in the subset corresponding to v adjacent to a vertex in the subset corresponding to w in G. A graph is *complete* if it contains every possible edge; in other words, if every pair of different vertices is adjacent. We write K_n as the complete graph with n vertices. A *bipartite* graph is a graph where the set of vertices is divided into two subsets, and any two vertices belonging to the same subset are not adjacent. A complete bipartite graph is a bipartite graph where all the pair of vertices belonging to different sets are adjacent. We write $K_{n,m}$ as the complete bipartite graph whose vertices are divided into two subsets of sizes n and m respectively.

1.1.2 Maps

Another structure that we consider in this thesis are maps. Roughly speaking, a map is a connected planar multigraph embedded in the sphere. We use the definitions and notation given by Landon and Zvonkin in [29]. More formally, a *rooted map* is a subdivision of the sphere into sets homeomorphic to dots, segments and discs. They are respectively called *vertices*, *edges* and *faces*. Moreover, there is a distinguished oriented edge called the *root* edge. The two extremes of an edge are called its *endpoints*. The *degree* of a vertex is the number of different endpoints of edges incident to it. Note that a vertex can be incident to the two endpoints of a loop, and in this case such an edge contributes twice to the degree of the vertex. The *degree* of a face is the number of edges appearing when traversing the boundary of the face. Note that a single edge can appear twice in the boundary, and in this case it is counted twice. Two maps are considered distinct if there is no homeomorphism between them preserving the internal structures. In Figure 1.1 we can see two equivalent rooted maps with three vertices and four edges each. Note that the lower triangle in the map at the left maps to the outer face of the map at the right. On the other hand, in Figure 1.2 we can see two different rooted maps. Although they are isomorphic as mutigraphs, the embedding is different, since in the first map the face at the left of the root edge has degree five, whereas in the second map it has degree three.

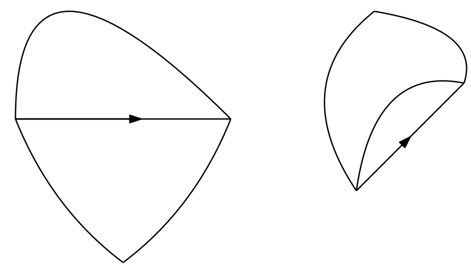

Figure 1.1: Two equivalent maps.

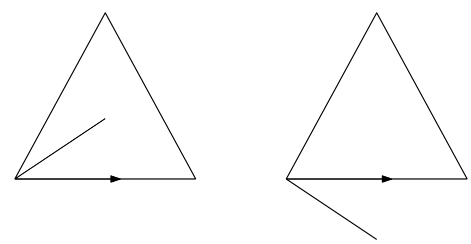

Figure 1.2: Two different maps.

In this thesis the vertices and edges of maps are not labelled, but one can prove that, after rooting and orienting an edge, all the other vertices and edges are distinguishable. In Figure 1.3 we can see a map with four unlabelled vertices, but vertices 1 and 2 are distinct, since vertex 1 is adjacent to one endpoint of the root edge, whereas vertex 2 is adjacent to the other endpoint. For the same reason, the edge between the origin of the root and vertex 1 is different from the edge between the end of the root and vertex 2. Note that, as a planar multigraph, all the vertices and edges are equivalent.

1.1.3 Classes of graphs

In this subsection we present the classes of graphs we work with. A class of graphs is given by the property that the graphs must fulfil in order to be in the class. For example, the class of connected graphs is the set of graphs that are connected. The total number of graphs with n vertices, using our definition of graph, is $2^{\binom{n}{2}}$. This kind of quadratic exponential does not behave well with our tools, since generating functions

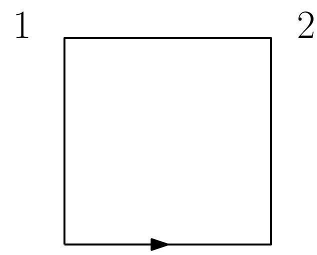

Figure 1.3: Two distinguishable vertices of a map.

typically encode additive parameters, which usually have linear growth, so for Chapter 2 and Chapter 3 we consider classes of graphs that only contain asymptotically $\theta(n)k^n n!$ graphs with n vertices, for some constant k and subexponential function $\theta(n)$. These classes are the class of planar graphs and the class of series parallel graphs.

We say that a graph G is *planar* if there is an embedding of each connected component of G in the sphere such that no two edges intersect. An embedding of a connected graph is a subdivision of the sphere into sets homeomorphic to dots, segments and discs such that there is a bijection between sets homeomorphic to dots and vertices in G, sets homeomorphic to segments and edges in G, such that an edge is incident to a vertex if an only if the corresponding sets are incident as well. For some applications, an alternative definition is the characterization given by Wagner's Theorem, that states that a graph is planar if and only if it does not contain K_5 or $K_{3,3}$ as a minor. However, this definition is not used for enumerative purposes. Instead, we use Whitney's Theorem, that states that every 3-connected graph admits a unique embedding in the sphere, which implies that we can enumerate 3-connected graphs in terms of 3-connected maps.

The other class that we consider is that of series-parallel graphs. There are several equivalent definitions for them. In our context, a *series-parallel* graph is a graph whose 2-connected components can be obtained from a graph consisting of two vertices and an edge, by means of a concatenation of series and parallel operations. Given a multigraph G, we say that a multigraph H is obtained from G by means of a *series* operation if H is the result of replacing an edge e of G with a vertex adjacent to the endpoints of e. Analogously, we say the H is obtained from G by means of a *parallel* operation if H is the result of duplicating an edge e of G with another edge between the same endpoints. Note that after a parallel operation we always have a non-simple multigraph. Since we are working with simple series-parallel graphs, we only consider those that end up without loops or multiple edges. In particular, the last operation must be a series operation. Also, note that according to our definition it is enough if the two connected components are obtained by means of this operation.

There are alternative equivalent definitions for series-parallel graphs. One of the best known states that a graph is series-parallel if it does not contain K_4 as a minor. Alternatively, series-parallel graphs are those whose connected components are subgraphs of a 2-tree, which is defined as the result of starting with an edge, and repeatedly adding vertices adjacent to the endpoints of an existing edge. These definitions are useful for other applications, but our tools behave well with the constructive definition of series and parallel operations, so we will use this one.

1.2 Analysis

In this section we present the analytic tools that we use to study combinatorial classes. Most of the content is based on the book Analytic Combinatorics by P. Flajolet and R. Sedgewick [5]. In Section 1.2.1 we define generating functions, which are used to translate combinatorial relations into analytic functions, so that we can apply the existing theorems. Section 1.2.2 is devoted to the study of the asymptotic enumeration of classes counted by generating functions, whereas Section 1.2.3 shows how to deal with objects chosen uniformly at random of such combinatorial classes. Finally, Section 1.2.4 discusses the particular case of random objects when they are defined by a system of an infinite number of equations.

1.2.1 Generating functions

A *combinatorial class* is a set of objects together with a size function. More formally, it is a pair $(\mathcal{A}, |\ .\ |)$, where $|\ .\ | : \mathcal{A} \to \mathbb{N}$ is called the size function. The size function must satisfy that the number of objects with any given size is finite. More formally, for every $n \in \mathbb{N}$, $|\{a \in A||a| = n\}| < \infty$. Also there might be objects of size zero, so we consider zero to be a natural number. Let $\mathcal{A}(n)$ be the set of elements of size n. We define the *generating function* of $\mathcal{A}$ as the formal power series $A(z) = \sum_{a \in A} z^{|a|} = \sum_{n=0}^{\infty} |\mathcal{A}(n)| z^n$. Sometimes, we call the generating function $A(z)$ the *ordinary* generating function, to distinguish it from the *exponential generating function*. The exponential generating function of $\mathcal{A}$ is the formal power series $A(x) = \sum_{a \in A} x^{|a|}/|a|! = \sum_{n=0}^{\infty} |\mathcal{A}(n)| x^n/n!$. Conversely, we write $[z^n]A(z) = |\mathcal{A}(n)| = a_n$.

We are often interested in keeping track of the value of a parameter of the combinatorial objects. A parameter is a function $f : \mathcal{A} \to \mathbb{N}$, but in contrast to the size, we do not impose that the number of elements with a given parameter is finite. Instead, we will distinguish objects of the same size according to the value of its parameter. Let $\mathcal{A}(n, m)$ be the set of elements of size n and parameter m. We define the *bivariate (ordinary) generating function* of a combinatorial class $\mathcal{A}$ with parameter f as $A(z, y) = \sum_{a \in A} z^{|a|} y^{f(a)} = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} |\mathcal{A}(n, m)| z^n y^m$. The *exponential bivariate generating function* of $\mathcal{A}$ is the formal power series $A(x, y) = \sum_{a \in A} x^{|a|} y^{f(a)}/|a|! = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} |\mathcal{A}(n, m)| x^n y^m/n!$. Note that in the bivariate exponential generating function we divide by the factorial of the size of an object, and not by the factorial of its parameter.

In order to translate combinatorial relations into generating functions we use the so-called *Symbolic Method*. This method gives a dictionary that allows us to obtain equations defining generating functions from the combinatorial relations that define the combinatorial class. The two basic classes are *neutral* class $\mathcal{E}$, consisting of a single element of size 0, and the *atomic* class $\mathcal{Z}$, consisting of a single element of size 1. Given two combinatorial classes $\mathcal{A}$ and $\mathcal{B}$, we define the *union* of $\mathcal{A}$ and $\mathcal{B}$, denoted

$\mathcal{A} \cup \mathcal{B}$, as the class consisting of the disjoint union of both classes. More formally, $\mathcal{A} \cup \mathcal{B} = (\mathcal{A}, 0) \cup (\mathcal{B}, 1)$, and the size of each element of the class is inherited from the initial classes. We define the *cartesian product* of $\mathcal{A}$ and $\mathcal{B}$, denoted $\mathcal{A} \times \mathcal{B}$, as the class of ordered pairs of elements in $\mathcal{A}$ and $\mathcal{B}$, where the size of each pair is the sum of the sizes of the elements in the pair. The *sequence* of a combinatorial class $\mathcal{A}$ such that $|\mathcal{A}(0)| = 0$ consists of the set of ordered finite sequences of any size of elements in $\mathcal{A}$, and the size is again the sum of the sizes of the elements in the sequence. More formally, $\mathrm{Seq}(\mathcal{A}) = \mathcal{E} \cup \mathcal{A} \cup (\mathcal{A} \times \mathcal{A}) \cup (\mathcal{A} \times \mathcal{A} \times \mathcal{A}) \ldots$. The *pointing* operator, $\mathcal{A}^\bullet$, consists in replacing each object of size n with n distinct objects of the same size. On the other hand, the *deriving* operator, $\mathcal{A}^\circ$, replaces each object of size n with n distinct objects of size $n-1$. The rationale with respect to these two operators is the following: we assume that each object of size n consists of n *atoms*, like for example vertices in the case of graphs. Then, the pointing operator consists in distinguishing one of the atoms of each object, whereas the deriving operator does the same but the atom is no longer considered as such, so the size of the object becomes $n-1$. The *composition* operator, $\mathcal{A} \circ \mathcal{B}$, replaces each element of $\mathcal{A}$ of size n with a sequence of n objects in $\mathcal{B}$, and the size is the sum of the sizes of the objects in $\mathcal{B}$. Again, this operator conceptually deals with atoms; in this case each atom of each object of $\mathcal{A}$ is replaced with an object in $\mathcal{B}$.

The cartesian product, as defined above, does not behave well when dealing with classes of labelled objects, i.e., classes where the objects of size n consist of n atoms labelled from 1 to n, like in the case of labelled graphs. Indeed, a pair of elements of sizes n and m has size $n+m$ and $n+m$ atoms, but the labels are chosen from two disjoint sets $\{1, \ldots, n\} \times \{0\}$ and $\{1, \ldots, m\} \times \{1\}$, whereas one might expect the labels of the atoms to be chosen from the set $\{1, \ldots, n+m\}$. This is why we define the *labelled product* between two combinatorial classes $\mathcal{A}$ and $\mathcal{B}$ as follows: each object of $\mathcal{A} \times \mathcal{B}$ consists in an ordered pair of objects $a \in \mathcal{A}$ and $b \in \mathcal{B}$, of size respectively n and m. Such an object has size $n+m$ and consists of $n+m$ atoms, n of those in a and m in b, inherited from the original objects. Each atom is labelled with a distinct label from 1 to $n+m$, and the relative order is preserved, in the sense that for each pair of atoms p and q in a, the label of p is larger than the label of q in a if an only if the label of p is larger than the label of q in the object $\langle a, b \rangle$, and the same with the atoms of b. Note that this means that every pair of objects of size n and m corresponds to $\binom{n+m}{n}$ objects of the labelled product. In Figure 1.4 one can see a possible relabelling of the product of two labelled objects that preserves the relative order of the labelling of each object of the pair.

In labelled contexts, the sequence operator is defined using this notion of labelled product. In other words, an object of the sequence class is the labelled product of a number of object of the original class. Also, the composition operator takes the labelling into account. We can also define the *set* operation of a labelled combinatorial class $\mathcal{A}$, as the class of unordered sets of objects in $\mathcal{A}$, where the size of an object is the sum of the sizes of the objects in the set.

Once we have defined combinatorial classes, we need a dictionary to translate them into algebraic relations between generating functions. The labelled operations work with exponential generating functions, whereas the standard product, sequence and

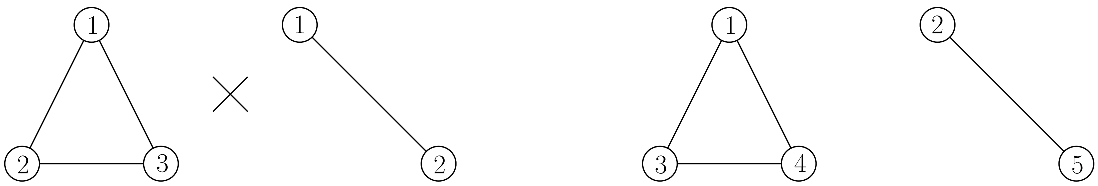

Figure 1.4: A relabelling of the product of labelled objects that preserves the relative order.

composition operations are easier to translate into ordinary generating functions. Table 1.1 includes the main conversions. Note that because of the convenient definition of exponential generating functions, the relations between exponential generating functions representing labelled operations are the same as the relations between ordinary generating functions representing ordinary operations, so we only include one column for both. The exception is the set operation, which only makes sense in the case of labelled classes, so the corresponding equation only works for exponential generating functions. In Chapter 3 we use other non-standard relations, like cycles and restricted sets, that are defined there. The operations of Table 1.1 also apply when there is an additive parameter in the combinatorial class and the generating function counts it. The composition, pointing and deriving operators can also be applied to the parameter in the way one might expect: if we replace the variable that counts the parameter with another generating function, the resulting generating function conceptually represents the result of replacing each object counted by the parameter with an object of the other combinatorial class. One can also differentiate with respect to the parameter, and the result will consist conceptually in distinguishing one of the objects counted by the parameter, and either keeping it or not, depending on the operator.

Construction	Class	Equations
Sum	$\mathcal{C} = \mathcal{A} \cup \mathcal{B}$	$C(x) = A(x) + B(x)$
Product	$\mathcal{C} = \mathcal{A} \times \mathcal{B}$	$C(x) = A(x) \cdot B(x)$
Sequence	$\mathcal{C} = \text{Seq}(\mathcal{A})$	$C(x) = 1/(1 - A(x))$
Set	$\mathcal{C} = \text{Set}(\mathcal{A})$	$C(x) = \exp(A(x))$
Composition	$\mathcal{C} = \mathcal{A} \circ \mathcal{B}$	$C(x) = A(B(x))$
Pointing	$\mathcal{C} = \mathcal{A}^\bullet$	$C(x) = A^\bullet(x) = x\frac{d}{dx}A(x)$
Deriving	$\mathcal{C} = \mathcal{A}^\circ$	$C(x) = A^\circ(x) = \frac{d}{dx}A(x)$

Table 1.1: The Symbolic Method translating combinatorial constructions into operations on counting series.

1.2.2 Enumeration

The whole point of computing the generating function of combinatorial classes is to analyse it in order to obtain properties about the class that it counts. The most natural property of a combinatorial class is the enumeration, defined as the number of objects of any given size, which must be finite by definition. In our applications, the exact number is often hard to compute, so instead we try to obtain an *asymptotic estimate*. More formally, given a combinatorial class $\mathcal{A}$, we are interested in finding a function $f : \mathbb{N} \to \mathbb{N}$ such that $\lim_{n\to\infty} f(n)/|\mathcal{A}(n)| = 1$. We use the notation $\mathcal{A}(n) \sim f(n)$ to state that $f(n)$ is the asymptotic estimate of $\mathcal{A}(n)$. There exist analytic tools that provide precise estimates from the knowledge of the generating function of a combinatorial class. This is explained in [5][Chapter IV], where the authors state the following principles:

First Principle of Coefficient Asymptotics. The location of the singularities of a function dictates the exponential growth of its coefficients.

Second Principle of Coefficient Asymptotics. The nature of the singularities of a function determines the associate subexponential factor.

Here, the exponential growth of the coefficients is the limit $\lim_{n\to\infty} \sqrt[n]{\mathcal{A}(n)}$, whereas the subexponential growth is a factor $\theta(n)$ such that $\lim_{n\to\infty} \sqrt[n]{\theta(n)} = 1$. In our applications, $\mathcal{A}(n) \sim \theta(n)\alpha^n$ typically holds, and we compute $\theta(n)$ and α by means of tools based in the previous principles. Note that when dealing with exponential generating functions, an extra $n!$ factor must be added. The exponential growth of a generating function can be obtained by means of Pringsheim's Theorem, as given in [5], which states that, under some technical conditions, the exponential growth of a generating function is given by the inverse of its dominant singularity. The exact statement of the theorem is the following:

Theorem 1.2.1 (Location of Singularities)**.** *Let $A(z)$ be a generating function. analytic at 0. Assume that $\rho = |\rho|$ is a singularity of $A(z)$, and every other singularity has absolute value larger than ρ. Then*

$$\sqrt[n]{[z_n]A(z)} \sim \rho^{-1}$$

The techniques for computing the subexponential growth are usually more involved than the ones used for the exponential term. See [5] for a complete description. In our case, we deal with square-root singularities, i.e., our generating functions behave like a square root near the singularity. We use the so-called Transfer Theorem. This theorem gives the subexponential growth assuming that the generating function can be analytically extended to a Δ-domain, defined as the region $\Delta(\alpha, \rho, R) = \{x \in \mathbb{C} : |x| < R,\ x \neq \rho,\ |\arg(x-\rho)| > \alpha\}$, where ρ is the dominant singularity, for some real $R > \rho$ and $\alpha > 0$, as in Figure 1.5. After we are able to prove that the generating function behaves like a square root near the main singularity, and that it can be analytically extended to a Δ-domain, we apply the following version of the Transfer Theorem, as given in [20]:

Theorem 1.2.2 (Transfer Theorem)**.** *Let a be an arbitrary complex number in* $\mathbb{C} \setminus \mathbb{Z}_{\leq 0}$ *and let* $R > \rho > 0$ *assume that* $f(x)$ *is analytic in a domain* $\Delta = \Delta(\alpha, \rho, R)$*.*

If, as $x \to \rho$ *in* Δ*,*

$$f(z) \sim (1 - z)^{-a}$$

then

$$[z^n]f(x) \sim \frac{n^{a-1}}{\Gamma(a)}.$$

Actually, the complete version given in [5][Theorem VI.1] gives a complete asymptotic expansion in descending powers of n, but for our purposes the first term is enough.

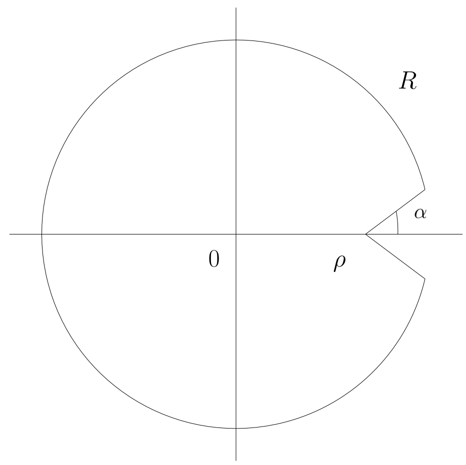

Figure 1.5: A Δ-domain near a singularity ρ with outer radius R and angle α.

1.2.3 Limit Laws

Given a combinatorial class $\mathcal{A}$ with a parameter $f : \mathcal{A} \to \mathbb{N}$, we are interested in the behaviour of the parameter of an object of a given size chosen uniformly at random. As in the case of univariate enumeration, computing the exact number of objects of size n and parameter m is typically too difficult for the kind of combinatorial classes we work with. Instead we will use probability analysis to find the asymptotic behaviour when the size n tends to infinity. First we need some definitions regarding probability analysis. Let $(\Omega, \mathcal{P}(\Omega), \mathbf{p})$ be a probability space, and let $\mathbf{X}$ be a random variable over this probability space. We denote $F_{\mathbf{X}}(x)$ the *probability distribution function* of $\mathbf{X}$, defined as $F_{\mathbf{X}}(x) = \mathbf{p}(\{\mathbf{X} \leq x\})$. If the probability distribution function is differentiable, we denote its derivative as $f_{\mathbf{X}}$, and we define the *expectation* $\mathbb{E}[\mathbf{X}]$ as follows:

$$\mathbb{E}[\mathbf{X}] = \int_{-\infty}^{+\infty} t f_{\mathbf{X}}(t)dt.$$

In general, given a function $g : \mathbb{R} \to \mathbb{R}$, we define $\mathbb{E}[g(\mathbf{X})]$ analogously as:

$$\mathbb{E}[g(\mathbf{X})] = \int_{-\infty}^{+\infty} g(t) f_{\mathbf{X}}(t) dt.$$

The *variance* is defined as $\mathbb{V}[\mathbf{X}] = \mathbb{E}[(\mathbf{X} - \mathbb{E}[\mathbf{X}])^2]$, or, equivalently, $\mathbb{V}[\mathbf{X}] = \mathbb{E}[\mathbf{X}^2] - \mathbb{E}[\mathbf{X}]^2$. Sometimes, when there is no ambiguity, we just write $\mathbb{E}\mathbf{X}$ and $\mathbb{V}\mathbf{X}$.

Analogously to the definition of generating functions of combinatorial classes, we can define the generating function of a discrete random variable, taking values in the natural numbers $\mathbb{Z}_{\geq 0}$. If $\mathbf{X}$ is such a random variable, the associated probability generating function $p_{\mathbf{X}}(z)$ is defined as $p_{\mathbf{X}}(z) = \sum_{n=0}^{\infty} \mathbf{p}(\mathbf{X} = n) z^n$. Note that $p_{\mathbf{X}}(1) = 1$, $p'_{\mathbf{X}}(1) = \mathbb{E}\mathbf{X}$, and $p''_{\mathbf{X}}(1) + p'_{\mathbf{X}}(1) - p'_{\mathbf{X}}(1)^2 = \mathbb{V}\mathbf{X}$, provided that these values exist.

Given a parameter f of a combinatorial class $\mathcal{A}$, and given $n \in \mathbb{N}$, we can define a random variable $\mathbf{X}_n$ as follows: we consider all the objects of $\mathcal{A}(n)$, and we pick one of them, a, uniformly at random. The value of the random variable is $f(a)$. Again, computing the exact value of $\mathbf{X}_n$ is typically difficult, so instead we compute the *limit* distribution of $\mathbf{X}_n$ when n tends to infinity. There are many definitions of convergence of random variables. In this thesis, the convergence is in *distribution*, even if we do not mention it. Given a sequence $(\mathbf{X}_n)_n$ of random variables over a probability space, we say that $(\mathbf{X}_n)_n$ *converges in distribution* to $\mathbf{X}$, denoted by $\mathbf{X}_n \xrightarrow{d} \mathbf{X}$, if for every $x \in \mathbb{R}$, $\lim_{n\to\infty} F_{\mathbf{X}_n}(x) = F_{\mathbf{X}}(x)$. Since this is the only kind of convergence that appears in this thesis, we just write $\mathbf{X}_n \to \mathbf{X}$ to express $\mathbf{X}_n \xrightarrow{d} \mathbf{X}$.

The main tool that we use to compute the limit law of a sequence of random variables that count a parameter is the *Quasi-powers Theorem*, which proves convergence to a Gaussian distribution, and gives easy tools to compute its expectation and variance. The Quasi-powers Theorem was obtained by Hwang in [34]. Before stating it, we need to define some notation, based on the version appearing in [5]. Let $f(u)$ be a complex function, analytic in $u = 1$, and assume that $f(1) \neq 0$. Then define

$$\mathfrak{m}(f) = \frac{f'(1)}{f(1)}, \qquad \mathfrak{v}(f) = \frac{f''(1)}{f(1)} + \frac{f'(1)}{f(1)} - \left(\frac{f'(1)}{f(1)}\right)^2.$$

With this notation, the statement in [5][Theorem IX.8], is as follows:

Theorem 1.2.3 (Quasi-powers Theorem)**.** *Let $\mathbb{X}_n$ be non-negative discrete random variables (supported by $\mathbb{Z}_{\geq 0}$), with probability generating functions $p_n(u)$. Assume that, uniformly in a fixed complex neighbourhood of $u = 1$, for sequences β_n, $\kappa_n \to +\infty$, there holds*

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

where $A(u)$, $B(u)$ are analytic at $u = 1$ and $A(1) = B(1) = 1$. Assume finally that $B(u)$ satisfies the so-called "variability condition",

$$\mathfrak{v}(B(u)) \equiv B''(1) + B'(1) - B'(1)^2 \neq 0.$$

Under these conditions, the mean and variance of $\mathbf{X}_n$ *satisfy*

$$\mu_n \equiv \mathbb{E}\mathbf{X}_n = \beta_n \mathfrak{m}(B(u)) + \mathfrak{m}(A(u)) + O(\kappa_n^1)$$
$$\sigma_n^2 = \mathbb{V}\mathbf{X}_n = \beta_n \mathfrak{v}(B(u)) + \mathfrak{v}(A(u)) + O(\kappa_n^1).$$

The distribution of $\mathbf{X}_n$ *is, after standardization, asymptotically Gaussian, and the speed of convergence to the Gaussian limit is* $O(\kappa_n^1 + \beta_n^{-1/2})$*:*

$$\mathbb{P}\left\{\frac{\mathbf{X}_n - \mathbb{E}\mathbf{X}_n}{\sqrt{\mathbb{V}\mathbf{X}_n}} \le x\right\} = \mathbf{\Phi}(x) + O\left(\frac{1}{\kappa_n} + \frac{1}{\sqrt{\beta_n}}\right),$$

where $\mathbf{\Phi}(x)$ *is the distribution function of a standard normal,*

$$\mathbf{\Phi}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-w^2/2} dw.$$

In Chapter 3 we want to compute a parameter whose probability generating function is not analytic for $u > 1$, and therefore it is not analytic near $u = 1$. Instead, we show that if the generating function is the solution of a system of equations satisfying a number of additional properties, then we can relax the constraint, so we only need that the probability generating function is analytic near $u = 1$ in an arc of circle of center 0 and radius 1. See Theorem 3.3.1 for further details.

1.2.4 Systems of equations

In some of our applications, our generating function are obtained as the solution of a system of equations. Under some conditions of positivity, the solutions of the system have the same dominant singularity, and behave like a square-root near it, which is sufficient to apply Transfer Theorem or Quasi-powers Theorem, as stated in the previous subsections. One of the main references in this area is [47][Chapter 2], which shows how to prove that the solution of a system of equations has the square-root behavior, distinguishing whether we have one or many equations, or whether we consider or not additional parameters. We reproduce here [47][Theorem 2.33], as it is general enough for our purposes.

Theorem 1.2.4. *Let* $\mathbf{F}(x, \mathbf{y}, \mathbf{u}) = (F_1(x, \mathbf{y}, \mathbf{u}), \ldots, F_N(x, \mathbf{y}, \mathbf{u}))$ *be a non-linear system of functions analytic around* $x = 0$*,* $\mathbf{y} = (y_1, \ldots, y_N) = \mathbf{0}$*,* $\mathbf{u} = (u_1, \ldots, u_k) = \mathbf{0}$*, whose Taylor coefficients are all non-negative, such that* $\mathbf{F}(0, \mathbf{y}, \mathbf{u}) = \mathbf{0}$*,* $\mathbf{F}(x, \mathbf{0}, \mathbf{u}) \neq 0$*,* $\mathbf{F}_x(x, \mathbf{y}, \mathbf{u}) \neq 0$*. Furthermore assume that the dependency graph of* $\mathbf{F}$ *is strongly connected and that the region of convergence of* $\mathbf{F}$ *is large enough that there exists a complex neighborhood* U *of* $\mathbf{u} = \mathbf{1} = (1, \ldots, 1)$*, where the system*

$$y = \mathbf{F}(x, \mathbf{y}, \mathbf{u}),$$
$$0 = \det(\mathbf{I} - \mathbf{F}_{\mathbf{y}}(x, \mathbf{y}, \mathbf{u})),$$

has solutions $x = x_0(\mathbf{u})$ *and* $\mathbf{y} = \mathbf{y}_0(\mathbf{u})$ *that are real, positive and minimal for positive real* $\mathbf{u} \in U$*.*

Let

$$\mathbf{y} = \mathbf{y}(x, \mathbf{u}) = (y_1(x, \mathbf{u}), \ldots, y_N(x, \mathbf{u}))$$

denote the analytic solutions of the system

$$\mathbf{y} = \mathbf{F}(x, \mathbf{y}, \mathbf{u})$$

with $\mathbf{y}(0, \mathbf{u}) = \mathbf{0}$*.*

Then there exists $\epsilon > 0$ *such that* $y_j(x, \mathbf{u})$ *admit a representation of the form*

$$y_j(x, \mathbf{u}) = g_j(x, \mathbf{u}) - h_j(x, \mathbf{u})\sqrt{1 - \frac{x}{x_0(\mathbf{u})}}$$

for $\mathbf{u} \in U$*,* $|x - x_0(\mathbf{u})| < \epsilon$ *and* $|\arg(x - x_0(\mathbf{u}))| \neq 0$*, where* $g_j(x, \mathbf{u}) \neq 0$ *and* $h_j(x, \mathbf{u}) \neq 0$ *are analytic functions with* $(g_j(x_0(\mathbf{u}), \mathbf{u}))_j = (y_j(x_0(\mathbf{u}), \mathbf{u}))_j = \mathbf{y}_0(\mathbf{u})$*.*

Furthermore, if $[x^n]y_j(x, \mathbf{1}) > 0$ *for* $1 \leq j \leq N$ *and for sufficiently large* $n \geq n_1$*, then there exists* $0 < \delta < \epsilon$ *such that* $y_j(x, \mathbf{u})$ *is analytic in* $(x, \mathbf{u})$ *for* $\mathbf{u} \in U$ *and* $|x - x_0(\mathbf{u})| \geq \epsilon$ *but* $|x| \leq |x_0(\mathbf{u})| + \delta$ *(this condition guarantees that* $\mathbf{y}(x, \mathbf{u})$ *has a unique smallest singularity with* $|x| = |x_0(\mathbf{u})|$*).*

In some cases, we cannot express a combinatorial class as the solution of a system of a finite number of equations. Instead, we express it as the solution of an infinite system of equations with an infinite number of variables. Again, we are interested in evaluating a parameter from an object of the class chosen uniformly at random. The main result in this area is given in [61], where the authors show that, if some constraints of compactness are fulfilled, then the solution of such a system admits a representation of square-root type in a Δ-domain, which proves that the parameter follows asymptotically a Gaussian law. For completeness the statement of the theorem is given here [61][Theorem 1]. Here ℓ^p, for a given real p with $1 \leq p < \infty$, denotes the Banach space of all complex valued sequences $(t_n)_{n \in \mathbb{N}}$ satisfying $\sum_{n=1}^{\infty} |t_n|^p < \infty$.

Theorem 1.2.5. *Let* $1 \leq p < \infty$*,* $1 \leq r \leq \infty$ *and* $\mathbf{F} : \mathbb{C} \times \ell^p \times \ell^r \to \ell^p$*,* $(x, \mathbf{y}, \mathbf{v}) \mapsto \mathbf{F}(x, \mathbf{y}, \mathbf{v})$ *be a function satisfying:*

(1) there exist open balls $B \in \mathbb{C}$*,* $U \in \ell^p$ *and* $V \in \ell^p$ *such that* $(0, \mathbf{0}, \mathbf{0}) \in B \times U \times V$ *and* $\mathbf{F}$ *is analytic in* $B \times U \times V$*,*

(2) The function $(x, \mathbf{y}) \mapsto \mathbf{F}(x, \mathbf{y}, \mathbf{0})$ *is a positive function,*

(3) $\mathbf{F}(0, \mathbf{y}, \mathbf{v}) = \mathbf{0}$ *for all* $\mathbf{y} \in U$ *and* $\mathbf{v} \in V$*,*

(4) $\mathbf{F}(x, \mathbf{0}, \mathbf{v}) \not\equiv \mathbf{0}$ *in* B *for all* $\mathbf{v} \in V$*,*

(5) $\frac{\partial \mathbf{F}}{\partial \mathbf{y}}(x, \mathbf{y}, \mathbf{0}) = A(x, \mathbf{y} + \alpha(x, \mathbf{y})\mathbf{I}_p$ *for all* $(x, \mathbf{y}) \in B \times U$*, where* α *is an analytic function and there exists an integer* n *such that* A^n *is compact,*

(6) $A(x, \mathbf{y})$ *is irreducible for strictly positive* $(x, \mathbf{y})$ *and* $\alpha(x, \mathbf{y})$ *has nonnegative Taylor coefficients.*

Let $\mathbf{y} = \mathbf{y}(x, \mathbf{v})$ *be the unique solution of the functional equation*

$$\mathbf{y} = \mathbf{F}(x, \mathbf{y}, \mathbf{v})$$

with $\mathbf{y}(0, \mathbf{v}) = \mathbf{0}$*. Assume that for* $\mathbf{v} = \mathbf{0}$ *the solution has a finite radius of convergence* $x_0 > 0$ *such that* $\mathbf{y}_0 := \mathbf{y}(x_0, \mathbf{0})$ *exists and* $(x_0, \mathbf{y}_0) \in B \times U$*.*

Then there exists $\epsilon > 0$ *such that* $\mathbf{y}(x, \mathbf{v})$ *admits a representation of the form*

$$\mathbf{y}(x, \mathbf{v}) = \mathbf{g}(x, \mathbf{v}) - \mathbf{h}(x, \mathbf{v})\sqrt{1 - \frac{x}{x_0(\mathbf{v})}}$$

for $\mathbf{v}$ *in a neighbourhood of* $\mathbf{0}$*,* $|x - x_0(\mathbf{v})| < \epsilon$*, and* $\arg(x - x_0(\mathbf{v})) \neq 0$*, where* $\mathbf{g}(x, \mathbf{v})$*,* $\mathbf{h}(x, \mathbf{v})$ *and* $x_0(\mathbf{v})$ *are analytic functions with* $\mathbf{h}_i(x_0(\mathbf{0}, \mathbf{0}) > 0$ *for all* $i \geq 1$*.*

Moreover, if there exist two integers n_1 *and* n_2 *that are relatively prime such that* $[x^{n_1}]\mathbf{y}_1(x, \mathbf{0}) > 0$ *and* $[x^{n_2}]\mathbf{y}_1(x, \mathbf{0}) > 0$ *then* $x_0(\mathbf{v}$ *is the only singularity of* $\mathbf{y}(x, \mathbf{v})$ *on the circle* $|x| = x_0(\mathbf{v})$ *and there exist constants* $0 < \delta < \pi/2$ *and* $\eta > 0$ *such that* $\mathbf{y}(x, \mathbf{v})$ *is analytic in a region of the form*

$$\Delta := \{x : |x| < x_0(\mathbf{0}) + \eta, |\arg(x/x_0(\mathbf{v}) - 1)| > \delta\}.$$

In our applications, conditions (1) $\sim$ (4), (6) are usually fulfilled. We state a set of conditions that implies (5), so that if the system of equations satisfy them, then Theorem 1.2.5 applies. We call this set of conditions the *partition condition*, which roughly speaking is defined as follows. Assume that the vector $\mathbf{F}$ is indexed by $j \in \mathbb{N}$. Then, for every $j \in \mathbb{N}$ there must exist a function $\widetilde{F}(x, y)$ such that $F_j(x, \mathbf{y}, \mathbf{1}) = \widetilde{F}_j(x, y_1 + y_2 + \ldots)$. In other words, if the parameters are not taken into account, then all the solutions of the equation $\mathbf{y} = \mathbf{F}(x, \mathbf{y}, \mathbf{v})$ have the same contribution to the system of equations. In Chapter 3 we define precisely this property, and we show how can we use it to prove that all the assumptions of Theorem 1.2.5 are fulfilled, and therefore that some complex parameters converge asymptotically to a Gaussian law.

1.3 Combinatorial techniques

In this section we show combinatorial properties that allow one to define generating functions as the solution of some equations, or as functions of other generating functions that we already know. Section 1.3.1 deals with the decomposition of a graph into connected and 2-connected components. The given combinatorial relations show how to obtain the generating function of 2-connected graphs in terms of the generating function of all the graphs, and vice versa. Section 1.3.2 shows how to *unroot* trees, and classes of graphs that have a tree structure, that is, how to obtain the generating function of trees from the generating function of rooted trees.

1.3.1 Block decomposition

Let us consider a class of graphs $\mathcal{G}$ such that a graph is in $\mathcal{G}$ if and only if every connected component is in $\mathcal{G}$. Let $\mathcal{C}$ be the subclass of connected graphs in $\mathcal{G}$. Then the following equation holds:

$$\mathcal{G} = \mathrm{Set}(\mathcal{C}).$$

Let $G(x)$ be the exponential generating function of $\mathcal{G}$, and let $C(x)$ be the exponential generating function of $\mathcal{C}$. Then

$$G(x) = \exp(C(x)).$$

In order for this equation to be true, we need that the empty graph is included in the class of general graphs, but not in the class of the connected graphs. This is why we impose that a connected graph has to be a non-empty graph.

In the literature, a 2-connected component is called a block. A class of graphs is *block stable* if a graph is in the class if and only if all its connected components and 2-connected blocks belong to the class as well. For technical reasons, the root of a connected graph will be part of it, whereas the root of a 2-connected graph will not. If a class is block-stable, we can relate the classes of *pointed* connected graphs and *derived* 2-connected graphs. The combinatorial argument is the following. The root of a rooted connected graph belongs to zero or more 2-connected blocks, and every vertex of every block apart from the root is incident as well to a rooted connected graph, which may consist of a single vertex. Let $\mathcal{C}^\bullet$ be the class of pointed connected graphs of the class, and $\mathcal{B}^\circ$ the class of derived 2-connected graphs of the class. Then

$$\mathcal{C}^\bullet = \mathcal{Z} \times \mathrm{set}(\mathcal{B}^\circ \circ \mathcal{C}^\bullet),$$

Recall that $\mathcal{Z}$ is the atomic class, consisting of a single element of size one. This combinatorial equation can be easily translated into generating functions. If $B'(x)$ and $xC'(x)$ are the generating functions of $\mathcal{B}^\circ$ and $\mathcal{C}^\bullet$ respectively, then the following equation holds:

$$xC'(x) = x\exp(B'(xC'(x)).$$

Note that, if needed, this equation also relates the equations for (unrooted) connected and 2-connected graphs, which can be obtained from $B'(x)$ and $xC'(x)$ by means of integration.

1.3.2 Dissymmetry theorem

The relation between 2-connected and 3-connected graphs of a class is not as straightforward as those shown in Section 1.3.1. The standard approach uses an intermediate class called networks. A *network* is a 2-connected graph where we root, direct and, possibly, remove an edge. The endpoints of the rooted edge are called *poles*, and we denote them as 0 and ∞, such that the direction of the edge goes from 0 to ∞. In this way, we keep track of the orientation of the root edge even if we remove it. Note that the poles are not labelled. Indeed, we use 0 and ∞ instead of the old labels of the vertices, so a network corresponding to a 2-connected graph of size n will have size $n-2$. Of course, a relabelling that preserves the relative order must be performed so that the other vertices have labels in $\{1, \ldots, n-2\}$. A *series network* is a network with at least one cut vertex, whereas a *parallel network* is a 2-connected network. The following relation between the combinatorial classes of 2-connected graphs $\mathcal{B}$ and networks $\mathcal{N}$ is straightforward:

$$(\mathcal{Y} \times \mathcal{B}^{\circ-\circ}) \cup (\mathcal{Y} \times \mathcal{B}^{\circ-\circ}) = \mathcal{X}^2 \times \mathcal{N},$$

where $\mathcal{Y}$ represents the class consisting of a single element of size 0 and one edge with generating function y, and $\mathcal{B}^{\circ-\circ}$ consists in rooting and removing one edge of the graph. This is translated to generating functions as follows:

$$2y^2 \frac{\partial}{\partial y} B(x, y) = x^2 N(x, y).$$

From this, it is clear that one could theoretically obtain the generating function of 2-connected graphs of a class by integrating the generating function of networks with respect to y. However, in some cases, we do not have access to the enriched generating function that counts edges, and in other cases integrating this function is not easy. This is why we use the so-called combinatorial integration approach defined in [2]. This approach uses the Dissymmetry Theorem to relate the generating functions of 2-connected graphs and networks of a given class of graphs.

The Dissymmetry Theorem states that there exists a bijection between classes of labelled trees with different rootings. This theorem was proven in [2], and the statement is the following:

Theorem 1.3.1 (Dissymmetry theorem)**.** *Let $\mathcal{T}$ be the class of labelled unrooted trees. Let $\mathcal{T}_\bullet$ be the class of labelled rooted trees. Let $\mathcal{T}_{\bullet-\bullet}$ be the class of labelled trees rooted in an edge. Let $\mathcal{T}_{\bullet-\circ}$ be the class of labelled trees rooted in an oriented edge. Then the following bijection holds:*

$$\mathcal{T} \cup \mathcal{T}_{\bullet-\circ} \approx \mathcal{T}_\bullet \cup \mathcal{T}_{\bullet-\bullet}$$

Note that this equation allows to express the generating function of unrooted trees in terms of the generating function of rooted trees, which is typically simpler. According to the dictionary of Section 1.2.1, one could also obtain the generating function of unrooted trees by integrating that of rooted trees. However, integration is typically more difficult than just applying the dissimetry theorem, since the generating function is usually not explicit. In any case, this gives an alternative name for the technique consisting in applying the dissymetry theorem: *combinatorial integration.*

In our case, we do not apply the theorem to trees, but to *tree-decomposable* classes. A tree-decomposable class is a class where each object has a tree associated to it. The tree represents different parts of the object, and we take advantage of the rooting in those cases where it is easier to define the object if a given part is distinguished. In particular, we know that the class of two connected graphs is tree-decomposable, by using a mapping from graphs to trees inspired in [17], where the authors give a canonical way to decompose a 2-connected graph into 3-connected components. In particular, this decomposition assigns a tree to every 2-connected graph where the nodes represent either cycles, or multiedges, or 3-connected graphs. Then the Dissymmetry Theorem applies, and it gives a system of equations that relates 3-connected graphs and 2-connected graphs, as the one shown in [2][Section 5.4]. This approach is particularly useful with series-parallel graphs, since there are no 3-connected series-parallel graphs, and planar graphs, since we know that there is a unique embedding of a 3-connected planar graph, and we know how to count 3-connected embedded planar graphs.

Chapter 2

Planar graphs of minimum degree two and three

This chapter is based in a joint work with M. Noy. We determine the asymptotic growth of planar maps and graphs with a condition on the minimum degree, and properties of random graphs from these classes. In particular we show that the expected size of the largest tree attached to the core of a planar graph of size n chosen uniformly at random is asymptotically $c\log(n)$ for an explicit constant c. These results provide new information on the structure of random planar graphs.

2.1 Introduction

The main goal of this chapter is to enumerate planar graphs subject to a condition on the minimum degree δ, and to analyze the corresponding planar random graphs. Asking for $\delta \geq 1$ is not very interesting, since a random planar graph contains in expectation a constant number of isolated vertices. The condition $\delta \geq 2$ is directly related to the concept of the core of a graph. Given a connected graph G, its *core* (also called 2-core in the literature) is the maximum subgraph C with minimum degree at least two. The core C is obtained from G by repeatedly removing vertices of degree one. Conversely, G is obtained by attaching rooted trees at the vertices of C. The *kernel* of G is obtained by replacing each maximal path of vertices of degree two in the core C with a single edge. The kernel has minimum degree at least three, and C can be recovered from K by replacing edges with paths. Notice that G is planar if and only C is planar, if and only if K is planar.

As shown in Figure 2.1, the kernel may have loops and multiple edges, which must be taken into account since our goal is to analyze simple graphs. Another issue is that when replacing loops and multiple edges with paths the same graph can be produced several times. To this end we weight multigraphs according to the number of loops and edges of each multiplicity. We remark that the concepts of core and kernel of a graph are instrumental in the theory of random graphs [64, 37].

For the sake of brevity, it is convenient to introduce the following definitions: a

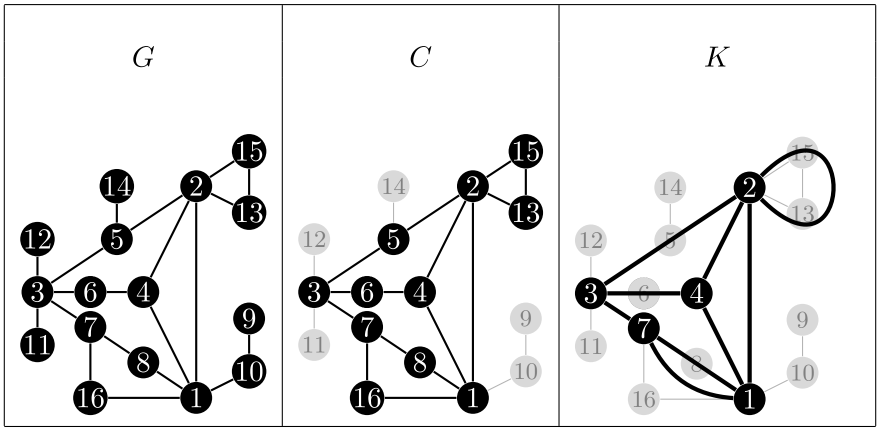

Figure 2.1: Core and kernel of a graph.

2-graph is a connected graph with minimum degree at least two, and a *3-graph* is a connected graph with minimum degree at least three. From now on all graphs are labelled and generating functions are of the exponential type. Let c_n, h_n and k_n be, respectively, the number of planar connected graphs, 2-graphs and 3-graphs with n vertices, and let

$$C(x) = \sum c_n \frac{x^n}{n!}, \qquad H(x) = \sum h_n \frac{x^n}{n!}, \qquad K(x) = \sum k_n \frac{x^n}{n!}$$

be the associated generating functions. Also, let $t_n = n^{n-1}$ be the number of (labelled) rooted trees with n vertices and let $T(x) = \sum t_n x^n/n!$. The decomposition of a connected graph into its core and the attached trees implies the following equation

$$C(x) = H(T(x)) + U(x), \tag{2.1}$$

where $U(x) = T(x) - T(x)^2/2$ is the generating functions of *unrooted* trees. Since $T(x) = xe^{T(x)}$, we can invert the above relation and obtain

$$H(x) = C(xe^{-x}) - x + \frac{x^2}{2}.$$

The equation defining $K(x)$ is more involved and requires the bivariate generating function

$$C(x, y) = \sum c_{n,k}\, y^k \frac{x^n}{n!},$$

where $c_{n,k}$ is the number of connected planar graphs with n vertices and k edges. We can express $K(x)$ in terms of $C(x, y)$ as

$$K(x) = C(A(x), B(x)) + E(x), \tag{2.2}$$

where $A(x), B(x), E(x)$ are explicit elementary functions (see Section 2.3).

From the expression of $C(x)$ as the solution of a system of functional-differential equations [8], it was shown that

$$c_n \sim \kappa n^{-7/2} \gamma^n n!,$$

where $\kappa \approx 0.4104 \cdot 10^{-5}$ and $\gamma \approx 27.2269$ are computable constants. In addition, analyzing the bivariate generating function $C(x, y)$ it is possible to obtain results on the number of edges and other basic parameters in random planar graphs. Our main goal is to extend these results to planar 2-graphs and 3-graphs.

Using Equations (2.1) and (2.2) we obtain precise asymptotic estimates for the number of planar 2- and 3-graphs:

$$\begin{aligned} h_n \sim \quad & \kappa_2 n^{-7/2} \gamma_2^n n!, \qquad & \gamma_2 \approx 26.2076, \quad \kappa_2 \approx 0.3724 \cdot 10^{-5}, \\ k_n \sim \quad & \kappa_3 n^{-7/2} \gamma_3^n n!, \qquad & \gamma_2 \approx 21.3102, \quad \kappa_3 \approx 0.3107 \cdot 10^{-5}. \end{aligned}$$

As is natural to expect, h_n and k_n are exponentially smaller than c_n. Also, the number of 2-connected planar graphs is known to be asymptotically $\kappa_c n^{-7/2} 26.1841^n n!$ (see [71]), smaller than the number of 2-graphs. This is consistent, since a 2-connected has minimum degree at least two.

By enriching Equations (2.1) and (2.2) taking into account the number of edges, we prove that the number of edges in random planar 2-graphs and 3-graphs are both asymptotically normal with linear expectation and variance. The expected number of edges in connected planar graphs was shown to be [8] asymptotically μn, where $\mu \approx 2.2133$. We show that the corresponding constants for planar 2-graphs and 3-graphs are

$$\mu_2 \approx 2.2614, \qquad \mu_3 \approx 2.4065.$$

This conforms to our intuition that increasing the minimum degree should increase the expected number of edges.

We also analyze the size X_n of the core in a random connected planar graph, and the size Y_n of the kernel in a random planar 2-graph. We show that both variables are asymptotically normal with linear expectation and variance and that

$$\begin{aligned} \mathbf{E}\, X_n \sim \quad & \lambda_2 n, \qquad & \lambda_2 \approx 0.9618, \\ \mathbf{E}\, Y_n \sim \quad & \lambda_3 n, \qquad & \lambda_3 \approx 0.8259. \end{aligned}$$

We remark that the value of λ_2 has been recently found by McDiarmid [16] using alternative methods. Also, we remark that the expected size of the largest block (2-connected component) in random connected planar graphs is asymptotically $0.9598n$ [28]. Again this is consistent since the largest block is contained in the core.

The picture is completed by analyzing the size of the trees attached to the core. We show that the number of trees with k vertices attached to the core is asymptotically normal with linear expectation and variance. The expected value is asymptotically

$$C \frac{k^{k-1}}{k!} \rho^k n,$$

where $C > 0$ is a constant and $\rho \approx 0.03673$ is the radius of convergence of $C(x)$. For k large, the previous quantity grows like

$$\frac{C}{\sqrt{2\pi}} \cdot k^{-3/2} (\rho e)^k n.$$

This quantity is negligible when $k \gg \log(n)/(\log(1/\rho e))$. Using the method of moments, we show that the size L_n of the largest tree attached to the core is in fact asymptotically

$$\frac{\log(n)}{\log(1/\rho e)}.$$

Our last result concerns the distribution of the vertex degrees in random planar 2-graphs and 3-graphs. We show that for each fixed $k \geq 2$ the probability that a random vertex has degree k in a random planar 2-graph tends to a positive constant $d_H(k)$, and for each fixed $k \geq 3$ the probability that a random vertex has degree k in a random planar 3-graph tends to a positive constant $d_K(k)$. Moreover $\sum_{k\geq 2} p_H(k) = \sum_{k\geq 3} p_K(k) = 1$, and the probability generating functions

$$p_H(w) = \sum_{k\geq 2} d_H(k) w^k, \qquad p_K(w) = \sum_{k\geq 3} d_K(k) w^k$$

are computable in terms of the probability generating function $p_C(w)$ of connected planar graphs, which was fully determined in [21].

The previous results show that almost all planar 2-graphs have a vertex of degree two, and almost all planar 3-graphs have a vertex of degree three. Hence asymptotically all our results hold also for planar graphs with minimum degree exactly two and three, respectively. In addition, all the results for connected planar graphs easily extend to arbitrary planar graphs. This is because the expected size of the largest component in a random planar graph is $n - O(1)$ (see [28]). We will not repeat for each of our results the corresponding statement for graphs of minimum degree exactly two or three.

It is natural to ask why we stop at minimum degree three. The reason is that there seems to be no combinatorial decomposition allowing to deal with planar graphs of minimum degree four or five (a planar graph has always a vertex of degree at most five). It is already an open problem to enumerate 4-regular planar graphs. In contrast, the enumeration of cubic planar graphs was completely solved in [38].

The contents of the chapter are as follows. In Section 2.2 we find similar results for planar maps, that is, connected planar graphs with a fixed embedding. They are simpler to derive and serve as a preparation for the results on planar graphs, while at the same time they are new and interesting by themselves. In Section 2.3 we find equations linking the generating functions of connected graphs, 2-graphs and 3-graphs; to this end we must consider multigraphs as well as simple graphs. In Section 2.4 we use singularity analysis in order to prove our main results on asymptotic enumeration and properties of random planar 2-graphs and 3-graphs. The analysis of the distribution of the degree of the root, which is technically more involved, is deferred to Section 2.5. We conclude with some remarks and open problems.

We assume familiarity with the basic results of analytic combinatorics as described in [5]. Given a complex number $\zeta \neq 0$, a Δ-domain at ζ is an open set of the form

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

We wiil need the following result [5, Corollary VI.1].

> *Transfer Theorem.* If $f(z)$ is analytic in a Δ-domain and satisfies, locally around its dominant singularity ρ, the estimate
>
> $$f(z) \sim (1 - z/\rho)^{-\alpha}, \qquad z \to \rho,$$
>
> with $\alpha \notin \{0, -1, -2, \dots\}$, then the coefficients of $f(z)$ satisfy
>
> $$[z^n]f(z) \sim \frac{n^{\alpha-1}}{\Gamma(\alpha)}\,\rho^{-n}.$$

We also need and a simplified version of [5, Theorem IX.8].

> *Quasi-powers Theorem.* Let the X_n be non-negative discrete random variables with probability generating functions $p_n(u)$. Assume that, uniformly in a fixed complex neighbourhood of $u = 1$
>
> $$p_n(u) = A(u)B(u)^n \left(1 + O\left(\frac{1}{n}\right)\right)$$
>
> where $A(u), B(u)$ are analytic at $u = 1$ and $A(1) = B(1) = 1$. Assume finally that $B(u)$ satisfies $B''(1) + B'(1) - B'(1)^2 \neq 0$.
>
> Then the distribution of X_n is, after standardization, asymptotically Gaussian, and the mean and variance satisfy
>
> $$\mathbf{E}\,X_n \sim \left(\frac{B'(1)}{B(1)}\right) n, \qquad \mathbf{Var}\,X_n \sim \left(\frac{B''(1)}{B(1)} + \frac{B'(1)}{B(1)} - \left(\frac{B'(1)}{B(1)}\right)^2\right) n.$$
>
> In our applications we will have $B(u) = \rho(1)/\rho(u)$, where $\rho(u)$ will be the dominant singularity (as a function of z) of a bivariate generating function $f(z, u)$. The former expressions become then
>
> $$\mathbf{E}\,X_n \sim \left(\frac{-\rho'(1)}{\rho(1)}\right) n, \qquad \mathbf{Var}\,X_n \sim \left(-\frac{\rho''(1)}{\rho(1)} - \frac{\rho'(1)}{\rho(1)} + \left(\frac{\rho'(1)}{\rho(1)}\right)^2\right) n.$$

In order to apply the former results we need to show that the corresponding generating functions can be defined in a Δ-domain. This relatively straightforward for planar maps, since we have explicit algebraic expressions for the generating functions. It is rather more involved for planar graphs. The expressions obtained in Section 2.3 are not enough for this purpose and we have to use alternative equations related to the decomposition of connected graphs into 2-connected components (see Section 2.4).

2.2 Planar maps

We recall that a planar map is a connected planar multigraph embedded in the plane up to homeomorphism. A map is rooted if one of the edges is distinguished and oriented. In this way a rooted map has a root edge and a root vertex (the tail of the root edge). We define the root face as the face on the right of the directed root edge. A rooted map has no automorphisms, in the sense that every vertex, edge and face is distinguishable. From now on all maps are planar and rooted. We stress the fact that maps may have loops and multiple edges.

The enumeration of rooted planar maps was started by Tutte in his seminal paper [1]. Let m_n be the number of rooted maps with n edges, with the convention that $m_0 = 0$. Then

$$m_n = \frac{2 \cdot 3^n}{(n+2)(n+1)} \binom{2n}{n}, \quad n \geq 1$$

The generating function $M(z) = \sum_{n \geq 0} m_n z^n$ is equal to

$$M(z) = \frac{18z - 1 + (1-12z)^{3/2}}{54z^2} - 1. \tag{2.3}$$

Either from the explicit formula or from the expression for $M(z)$ and the transfer theorem, it follows that

$$m_n \sim \frac{2}{\sqrt{\pi}} n^{-5/2} 12^n. \tag{2.4}$$

If $m_{n,k}$ is the number of maps with n edges and degree of the root face equal to k, then $M(z,u) = \sum m_{n,k} z^n u^k$ satisfies the equation

$$M(z,u) = zu^2(M(z,u)+1)^2 + uz\left(\frac{uM(z,u) - M(z,1)}{u-1} + 1\right). \tag{2.5}$$

By duality, $M(z,u)$ is also the generating function of maps in which u marks the degree of the root vertex. This is a convenient modifications of the usual equations for maps, where the empty map is also counted.

The core C of a map M is obtained, as for graphs, by removing repeatedly vertices of degree one, so that C has minimum degree at leat two. Then M is obtained from C by placing a planar tree at each corner (pair of consecutive half-edges) of C. This is equivalent to replacing each edge with a non-empty planar tree rooted at an edge. The number t_n of planar trees with $n \geq 1$ edges is equal to the n-th Catalan number and the generating function $T(z) = \sum t_n z^n$ satisfies

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

We define a 2-map as a map with minimum degree at least two, and a 3-map as a map with minimum degree at least three. Let h_n and k_n be, respectively, the number of 2-maps and 3-maps with n edges.

Theorem 2.2.1. *The generating functions $H(z)$ and $K(z)$ of 2-maps and 3-maps, respectively, are given by*

$$H(x) = \frac{1-x}{1+x}\left(M\left(\frac{x}{(1+x)^2}\right) - x\right), \qquad K(x) = \frac{H\left(\frac{x}{1+x}\right) - x}{1+x}.$$

The following estimates hold:

$$h_n \sim \kappa_2 n^{-5/2}(5+2\sqrt{6})^n, \qquad k_n \sim \kappa_3 n^{-5/2}(4+2\sqrt{6})^n, \tag{2.6}$$

where

$$\kappa_2 = \frac{2}{\sqrt{\pi}}\left(\frac{2}{3}\right)^{5/4} \approx 0.6797, \quad \kappa_3 = \frac{2}{\sqrt{\pi}}\left(4 - 4\sqrt{\frac{2}{3}}\right)^{5/2} \approx 0.5209.$$

Proof. The decomposition of a map into its core and the collection of trees attached to the corners implies the following equation:

$$M(z) = T(z) + H\left(T(z)\right)\frac{1+T(z)}{1-T(z)}. \tag{2.7}$$

The first summand corresponds to the case where the map is a tree, and the second one where the core is not empty: each edge is replaced with a non-empty tree whose root corresponds to the original edge. The factor

$$\frac{1+T(z)}{1-T(z)} = 1 + \frac{2\,T(z)}{1-T(z)}$$

is interpreted as follows. The first summand corresponds to the case where the root of the map is in the core, and the second one to the case where it is in a pendant rooted tree τ, which we place at the left-back corner of the root edge of the core. In this case there is a non-empty sequence of non-empty trees from the root edge e of τ to the root edge of the core, and the factor 2 distinguishes the two possible directions of e.

In order to invert the former relation let $x = T(z)$, so that

$$z = \frac{x}{(1+x)^2}.$$

We obtain

$$H(x) = \frac{1-x}{1+x}\left(M\left(\frac{x}{(1+x)^2}\right) - x\right) = x + 3x^2 + 16x^3 + 96x^4 + 624x^5 + \cdots \tag{2.8}$$

Let now C be a 2-map. The kernel K of C is defined as follows: replace every maximal path of vertices of degree two in C with a single edge (see Figure 2.2). Clearly K is a 3-map and C can be obtained by replacing edges in K with paths. It follows that

$$H(z) = K\left(\frac{z}{1-z}\right)\frac{1}{1-z} + \frac{z}{1-z}. \tag{2.9}$$

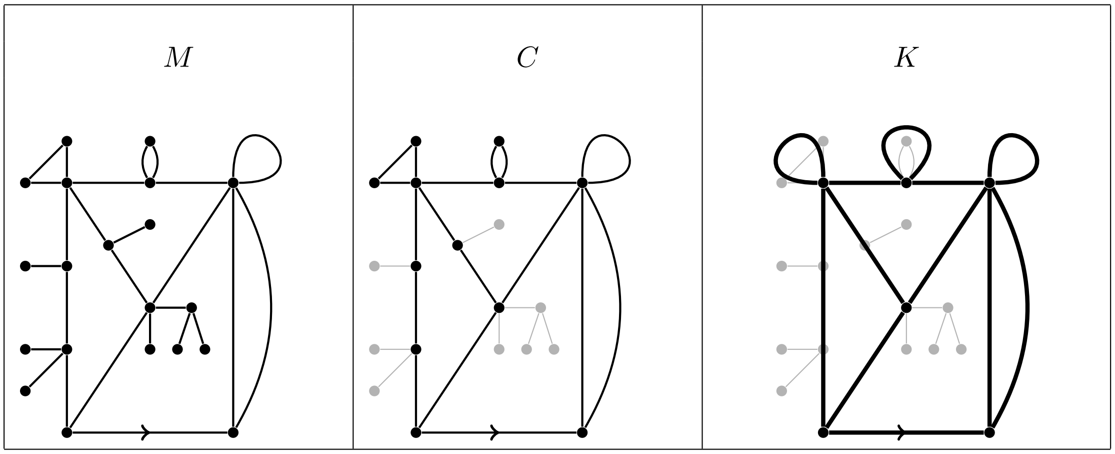

Figure 2.2: Core and kernel of a map.

The first term corresponds to the substitution of paths for edges, and the extra factor $1/(1-z)$ indicates where to locate the new root edge in the path replacing the original root edge. The last term corresponds to cycles, whose kernel is empty. Inverting the relation $x = z/(1-z)$ we obtain

$$K(x) = \frac{H\left(\frac{x}{1+x}\right) - x}{1+x} = 2z^2 + 9z^3 + 47z^4 + 278z^5 + \cdots \tag{2.10}$$

In order to obtain asymptotic estimates for h_n and k_n we need to locate the dominant singularities of $H(z)$ and $K(z)$ and show that these functions are defined on suitable Δ-domains. $M(z)$ has a unique singularity at $\rho = 1/12$ and is analytic in $\mathbb{C}$ minus the ray $[1/12, +\infty)$, and $T(z)$ is singular only at $1/4$. Hence $H(z)$ has a singularity at $\sigma = \rho T(\rho)^2 = 5 - 2\sqrt{6}$. We show next that $H(z)$ is analytic in $|x| < \sigma$ and has no other singularities in $|x| = \sigma$. By continuity it is analytic in a Δ-domain at σ.

From Equation (2.8), the singularities of H are at -1 and at the points x where $t = x/(1+x)^2 \in [1/12, \infty)$. We show that these points either satisfy $|x| \geq 1$ or belong to the real segment $[\sigma, 1)$. If we solve the equation for x we get

$$x = \frac{1 - 2t \pm \sqrt{1-4t}}{2t}, \qquad t \in [1/12, \infty).$$

We analyze two cases. For $t > 1/4$ we can rewrite $x = (1 - 2t \pm i\sqrt{4t-1})/2t$ and obtain $|x| = 1$. When $1/12 \leq t \leq 1/4$, x must be real. Consider the solution $x(t) = (1 - 2t + \sqrt{1-4t})/2t$. It is non-increasing since the derivative

$$x'(t) = -\frac{1 - 2t + \sqrt{1-4t}}{2\sqrt{1-4t} \cdot t^2}.$$

is negative . Since $x(1/4) = 1$ it follows that $x \geq 1$. For the solution $x(t) = (1 - 2t - \sqrt{1-4t})/2t$ the derivative is positive and $x(1/12) = \sigma$. Hence $x \geq \sigma$.

From Equation (2.10) it follows that $K(z)$ has singularity at $\tau = \sigma/(1-\sigma) = (\sqrt{6}-2)/4$.

For future reference we display these basic constants, that is, the dominant singularities for 2- and 3-maps:

$$\sigma = 5 - 2\sqrt{6}, \qquad \tau = \frac{\sqrt{6}-2}{4}.$$

The singular expansion of $M(z)$ at the singularity $z = 1/12$ can be obtained directly from the explicit formula (2.3), and is equal to

$$M(z) = \frac{1}{3} - \frac{4}{3}Z^2 + \frac{8}{3}Z^3 + O(Z^4),$$

where $Z = \sqrt{1-12z}$. Plugging this expression into (2.8) and expanding gives

$$H(x) = H_0 + H_2X^2 + \frac{8}{3}\left(\frac{2}{3}\right)^{5/4} X^3 + O(X^4),$$

where now $X = \sqrt{1-x/\sigma}$. A similar computation using (2.10) gives

$$K(x) = K_0 + K_2X^2 + \frac{8}{3}\left(4 - 4\sqrt{\frac{2}{3}}\right)^{5/2} X^3 + O(X^4),$$

where $X = \sqrt{1-x/\tau}$.

The estimates for h_n and k_n follow by the transfer theorem and the equality $\Gamma(-3/2) = 4\sqrt{\pi}/3$. □

Our next result is a limit law for the size of the core and the kernel in random maps.

Theorem 2.2.2. *The size X_n of the core of a random map with n edges, and the size Y_n of the kernel of a random 2-map with n edges are asymptotically Gaussian with*

$$\mathbf{E}\,X_n \sim \frac{\sqrt{6}}{3}n \approx 0.8165n, \qquad \mathbf{Var}\,X_n \sim \frac{n}{6} \approx 0.1667n,$$

$$\mathbf{E}\,Y_n \sim (2\sqrt{6}-4)n \approx 0.8990n, \qquad \mathbf{Var}\,Y_n \sim (18\sqrt{6}-44)n \approx 0.0908n.$$

The size Z_n of the kernel of a random map with n edges is also asymptotically Gaussian with

$$\mathbf{E}\,Z_n \sim \left(4 - \frac{4\sqrt{6}}{3}\right)n \approx 0.7340n, \qquad \mathbf{Var}\,Z_n \sim \left(\frac{128}{3} - \frac{52}{3}\sqrt{6}\right)n \approx 0.2088n.$$

Proof. If u marks the size of the core in maps then and immediate extension of (2.7) yields

$$M(z,u) = H\left(uT(z)\right)\frac{1+T(z)}{1-T(z)} + T(z). \tag{2.11}$$

It follows that the singularity $\xi(u)$ of the univariate function $z \mapsto M(z,u)$ is given by

$$\xi(u) = \frac{\sigma u}{(\sigma+u)^2}.$$

An easy calculation gives

$$-\frac{\xi'(1)}{\xi(1)} = \frac{\sqrt{6}}{3}, \qquad -\frac{\xi''(1)}{\xi(1)} - \frac{\xi'(1)}{\xi(1)} + \left(\frac{\xi'(1)}{\xi(1)}\right)^2 = \frac{1}{6}.$$

If now u marks the size of the kernel in 2-maps then an extension of (2.9) gives

$$H(z,u) = K\left(\frac{uz}{1-z}\right)\frac{1}{1-z} + \frac{z}{1-z}. \tag{2.12}$$

The singularity $\chi(u)$ of $z \mapsto K(z,u)$ is now given by

$$\chi(u) = \frac{\tau}{\tau+|u|}.$$

Applying the quasi-powers theorem and we obtain

$$-\frac{\chi'(1)}{\chi(1)} = 2\sqrt{6}-4, \qquad -\frac{\chi''(1)}{\chi(1)} - \frac{\chi'(1)}{\chi(1)} + \left(\frac{\chi'(1)}{\chi(1)}\right)^2 = 18\sqrt{6}-44.$$

The last statement concerning Z_n follows by combining equations (2.11) and (2.12), obtaining an expression of $M(z,u)$ in terms of $K(z)$, and repeating the same computations as before for the corresponding singularity function. In order to apply the quasi-powers theorem we need to show that, for u_0 close enough to 1 we can extend both $H(z,u_0)$ and $K(z,u_0)$ to respectives Δ-domains. Both proofs are analogous to the one of Theorem 2.4.3. $\square$

It is interesting to compare the previous result with the known results on the largest block (2-connected components) in random maps [43]. The expected size of the largest block in random maps is asymptotically $n/3$, rather smaller than the size of the core. In other words, the core C consists of the largest block B together with smaller blocks attached to B comprising in total $\frac{\sqrt{6}-1}{3}n \approx 0.4832n$ edges. An explanation for this is the presence of a linear number of loops, which belong to the core, but do not belong to the largest block.

Our next goal is to analyze the size of the trees attached to the core of a random map.

Theorem 2.2.3. *Let $X_{n,k}$ count trees with k edges attached to the core of a random map with n edges. Then $X_{n,k}$ is asymptotically normal and*

$$\mathbf{E}\, X_{n,k} \sim \alpha_k n$$

where

$$\alpha_k = (4+\frac{5}{3}\sqrt{6})\frac{1}{k+1}\binom{2k}{k}\left(\frac{1}{12}\right)^k, \quad k \geq 1.$$

Moreover, $\sum_{k\geq 1}\alpha_k = \sqrt{6}/3$.

Proof. The generating function for trees, where variable w_k marks trees with k edges, is equal to

$$T(z, w_k) = T(z) + (w_k - 1)t_k z^k,$$

where $t_k = C_k = \frac{1}{k+1}\binom{2k}{k}$ is the k-th coefficient of $T(x)$. The scheme for the core decomposition is then

$$M(z, w_k) = H(T(z, w_k))\frac{1 + T(z)}{1 - T(z)} + T(z).$$

It follows that the singularity $\rho_k(w_k)$ of the univariate function $z \mapsto M(z, w_k)$ is given by the equation

$$T(\rho_k(w_k)) + (w_k - 1)t_k\rho_k(w_k)^k = \sigma.$$

An easy calculation gives

$$-\frac{\rho_k'(1)}{\rho_k(1)} = \frac{1 - \sigma}{\sigma(1 + \sigma)}C_k\left(\frac{1}{12}\right)^k.$$

The first part of the proof is concluded by noticing that $(1 - \sigma)/(\sigma(1 + \sigma)) = 4 + \frac{5}{3}\sqrt{6}$.

Finally, $\sum_{k\geq 1} \alpha_k = \sqrt{6}/3$ follows from the closed form of the generating function for the Catalan numbers. □

Recall that the size of the core is asymptotically $\frac{1-\sigma}{1+\sigma}n = \frac{\sqrt{6}}{3}n$. Hence the asymptotic probability that a random tree attached to the core has size k is

$$\beta_k = \frac{\alpha_k}{\sqrt{6}/3} \sim \frac{1}{\sigma\sqrt{\pi}}k^{-3/2}3^{-k}, \qquad k \to \infty, \tag{2.13}$$

It follows that if $k \gg \log(n)/\log(3)$, the expected number $\alpha_k n$ of trees of size k tends to zero. This indicates that the size L_n of the largest tree attached to the core is at most $\log(n)/\log(3)$ with high probability. We show that in fact $L_n/\log(n)$ tends to $1/\log(3)$.

Theorem 2.2.4. *Let L_n be the size the largest tree attached to the core of a random map with n edges. Then*

$$\frac{L_n}{\log n} \to \frac{1}{\log(3)} \approx 0.9102 \qquad \textit{in probability}$$

and

$$\mathbb{E}L_n \sim \frac{1}{\log(3)}\log n \qquad (n \to \infty).$$

The proof is completely analogous to the proof of Theorem 2.4.5 and is not presented here. After proving Theorem 2.4.5 we indicate how to adapt the proof to establish Theorem 2.2.4, which is technically easier.

Degree distribution. Our last result in this section deals with the distribution of the degree of the root vertex in 2-maps and 3-maps. We let $M(z,u)$ be the GF of maps, where z marks edges and u marks the degree of the root vertex. Similarly, $H(z,u)$ is the GF for 2-maps, and $T(z,u) = 1/(1-uz(T(z)+1)) - 1$ for trees, where again u marks the degree of the root. Then we have

$$M(z,u) = H\left(T(z), \frac{u(T(z,u)+1)}{T(z)+1}\right)(T(z,u)+1) + H(T(z))\frac{T(z,u)}{1-T(z)} + T(z,u).$$

The first term corresponds to the case where the root belongs to the core: we replace each edge with a tree, and each edge incident to the root vertex is replaced with a possibly empty tree, where u marks the degree of the root. The term $T(z)+1$ in the denominator ensures that an edge is not replaced twice with a tree. The factor $T(z,u)+1$ allows to place a possibly empty tree in the root corner. The second term corresponds to the case where the root belongs to a tree attached to the core: the denominator $1-T(z)$ encodes a sequence of trees going from the core to the root edge. The last term corresponds to the case where the core is empty, and therefore the map is a tree.

If we change variables $x = T(z)$ and $w = u(T(u,z)+1)/(T(z)+1)$, the inverse is

$$z = \frac{x}{(1+x)^2}, \qquad u = \frac{w(1+x)}{1+wx}.$$

The former equation becomes

$$H(x,w) = \frac{M\left(\frac{x}{(1+x)^2}, \frac{w(1+x)}{1+wx}\right)}{1+wx} - \frac{wx}{1+x}M\left(\frac{x}{(1+x)^2}\right) + \frac{1}{1+wx} + \frac{wx^2}{1-x} - 1. \tag{2.14}$$

The first terms are

$$H(x,u) = w^2x + \left(w^2+2w^4\right)x^2 + \left(3w^2+4w^3+4w^4+5w^6\right)x^3 + \cdots$$

The relationship between $H(z,u)$ and $K(z,u)$ is simpler:

$$H(z,u) = K\left(\frac{z}{1-z}, u\right) + K\left(\frac{z}{1-z}\right)\frac{zu^2}{1-z} + \frac{zu^2}{1-z}.$$

Inverting gives

$$K(x,u) = H\left(\frac{x}{1+x}, u\right) - \frac{xu^2}{1+x}H\left(\frac{x}{1+x}\right) - \frac{xu^2}{1+x}, \tag{2.15}$$

and the first terms are

$$K(z,u) = 2u^4z^2 + (4u^3+5u^6)z^3 + (9u^3+9u^4+15u^5+14u^8)z^4 + \cdots$$

In order to analyze $H(z,u)$ and $K(z,u)$ we need the expansion of $M(z,u)$ near the singularity $\rho = 1/12$. As we have seen, the expansion of $M(z)$ near $z = 1/12$ is

$$M(z) = \frac{1}{3} - \frac{4}{3}Z^2 + \frac{8}{3}Z^3 + O(Z^4),$$

where $Z = \sqrt{1-12z}$. Since $M(z,u)$ satisfies (2.5) we obtain

$$M(z,u) = M_0(u) + M_2(u)Z^2 + M_3(u)Z^3 + O(Z^4). \tag{2.16}$$

A simple computation by indeterminate coefficients gives

$$M_3(u) = \frac{8u}{\sqrt{3(2+u)(6-5u)^3}}.$$

The limiting probability that a random map has a root vertex (or face) of degree k is equal to

$$p_M(k) = \frac{[u^k][z^n]M(z,u)}{[z^n]M(z)}.$$

Both coefficients can be estimated using transfer theorems and we get that the probability generating function of the distribution is given by

$$p_M(u) = \sum p_M(k)u^k = \frac{M_3(u)}{M_3(1)} = \frac{u\sqrt{3}}{\sqrt{(2+u)\,(6-5\,u)^3}} \tag{2.17}$$

Our goal is to obtain analogous results for 2-maps and 3-maps.

Theorem 2.2.5. *Let $p_M(u)$ be as before, and let $p_H(u)$ and $p_K(u)$ be the probability generating functions for the distribution of the root degree in 2-maps and 3-maps, respectively. Then we have*

$$p_H(u) = \frac{p_M\left(\dfrac{u(1+\sigma)}{1+u\sigma}\right)\dfrac{1+\sigma}{1+u\sigma} - u\sigma}{1-\sigma},$$

$$p_K(u) = \frac{p_H(u) - u^2\sigma}{1-\sigma},$$

where $\sigma = 5 - 2\sqrt{6}$, as in Theorem 2.2.1. Furthermore, the limiting probabilities that the degree of the root vertex is equal to k exist, both for 2-maps and 3-maps, and are asymptotically

$$p_H(k) \sim \nu_2 k^{1/2} w_H^k,$$

$$p_M(k) \sim \nu_3 k^{1/2} w_K^k,$$

where $w_H = w_K = \sqrt{2/3} \approx 0.8165$, $\nu_2 = \sqrt{3(1-\sigma)/(64\pi)} \approx 0.1158$, $\nu_3 = \sqrt{3/(64\pi(1-\sigma))} \approx$ 0.1288.

The correction terms $u\sigma$ in $p_H(u)$ and $u^2\sigma$ in $p_K(u)$ are due to the fact, respectively, that 2-maps have no vertices of degree one and 3-maps no vertices of degree two.

Proof. Since $M(z,u)$ satisfies (2.16) and $H(x,w)$ satisfies (2.14), we obtain

$$H(z,u) = H_0(u) + H_2(u)Z^2 + H_3(u)Z^3 + O(Z^4),$$

where $Z = \sqrt{1 - z/\sigma}$, and $H_3(u)$ can be computed as

$$H_3(u) = \left(\frac{1-\sigma}{1+\sigma}\right)^{3/2} \left(\frac{M_3\left(u(1+\sigma)/(1+u\sigma)\right)}{1+u\sigma} - \frac{M_3(1)u\sigma}{1+\sigma}\right).$$

The probability generating function of the distribution is given by

$$p_H(u) = \frac{H_3(u)}{H_3(1)} = \frac{p_M\left(\dfrac{u(1+\sigma)}{1+u\sigma}\right)\dfrac{1+\sigma}{1+u\sigma} - u\sigma}{1-\sigma}, \tag{2.18}$$

as claimed in the statement.

Now by (2.15), $K(u,z)$ satisfies

$$K(z,u) = K_0(u) + K_2(u)Z^2 + K_3(u)Z^3 + O(Z^4),$$

where now $Z = \sqrt{1 - z/\tau}$ and $K_3(u)$ is

$$K_3(u) = \left(\frac{1}{1+\tau}\right)^{3/2} \left(H_3(u) - H_3(1)\sigma u^2\right).$$

The probability generating function of the distribution is given by

$$p_K(u) = \frac{K_3(u)}{K_3(1)} = \frac{p_H(u) - u^2\sigma}{1-\sigma}. \tag{2.19}$$

The asymptotics of the distributions can be obtained from that of $p_M(u)$. The singularity of $p_M(u)$ is at $u_M = 6/5$, and its expansion is computed from the explicit formula in (2.17) as

$$p_M(u) = P_{-3}U^{-3} + O(U^{-2}), \tag{2.20}$$

where $U = \sqrt{1 - 5u/6}$ and $P_{-3} = 1/(4\sqrt{10})$. The singularity of p_H and p_K is obtained by solving the equation

$$\frac{u(1+\sigma)}{1+u\sigma} = u_M = \frac{6}{5},$$

giving $u_H = u_K = \sqrt{3/2}$. Hence, the exponential growth constants are $w_H = w_K = \sqrt{2/3}$. The singular expansion of $p_H(u)$ is obtained by composing (2.18) and (2.20), giving as a result

$$p_H(u) = Q_{-3}U^{-3} + O(U^{-2}), \tag{2.21}$$

where now $U = \sqrt{1 - u\sqrt{2/3}}$, and $Q_{-3} = P_{-3}\sqrt{15(1-\sigma)/8} = \sqrt{3(1-\sigma)}/16$. The singular expansion of $p_K(u)$ is obtained by composing (2.19) and (2.21) giving as a result

$$p_K(u) = R_{-3}U^{-3} + O(U^{-2}), \tag{2.22}$$

where U is as before and $R_{-3} = Q_{-3}/(1-\sigma) = \sqrt{3/(1-\sigma)}/16$.

The estimates for $p_H(k)$ and $p_M(k)$ follow by the transfer theorem, provided that the probability generating functions can be extended to a Δ-domain. Since we know explicitly $p_M(u)$, we also know that it is analytic at $D = \mathbb{C} \setminus (-\infty, -2] \cup [6/5, \infty)$. By Equation (2.18) we know that if $u(1+\sigma)/(1+u\sigma) \in D$ then p_H and p_K are analytic at u. By inverting the expression we can check that if $u(1+\sigma)/(1+u\sigma) \notin D$ then $u \in (-\infty, -1/(8-3\sqrt{6})] \cup [\sqrt{3/2}, \infty)$, and therefore p_H and p_K are analytic in a Δ-domain. $\square$

2.3 Equations for 2-graphs and 3-graphs

In this section we find expressions for the generating functions of 2- and 3-graphs in terms of the generating function of connected graphs. The results are completely general and specialize to the generating functions of planar graphs, since a graph is planar if and only if its core is planar, and in turn the core is planar if and only if its kernel is planar.

Let $C(x,y)$ be the generating function of connected graphs, where x marks vertices and y marks edges. Denote by $H(x,y)$ and $K(x,y)$ the generating functions, respectively, of 2-graphs and 3-graphs. We will find equations of the form

$$\begin{aligned} H(x,y) &= C(A_1(x,y), B_1(x,y)) + E_1(x,y) \\ K(x,y) &= C(A_2(x,y), B_2(x,y)) + E_2(x,y), \end{aligned}$$

where A_i, B_i and E_i are explicit functions.

From now on all graphs are labelled, and all generating functions are of the exponential type.

2-graphs. Let $\mathcal{G}$ be a connected graph. The core $\mathcal{C}$ of $\mathcal{G}$ is obtained by removing repeatedly vertices of degree one, so that $\mathcal{G}$ is obtained from $\mathcal{C}$ by replacing each vertex of $\mathcal{G}$ with a rooted tree. The number T_n of rooted trees with n edges is known to be n^{n-1}, and the generating function $T(x) = \sum T_n x^n/n!$ satisfies

$$T(x) = xe^{T(x)}.$$

The core of $\mathcal{G}$ can be empty, in which case $\mathcal{G}$ must be an (unrooted) tree. The number U_n of unrooted trees is known to be n^{n-2}, and the generating function $U(x) = \sum u_n x^n/n!$ is equal to

$$U(x) = T(x) - \frac{T(x)^2}{2}.$$

Theorem 2.3.1. *Let h_n be the number of 2-graphs with n vertices. Then $H(x) = \sum h_n x^n/n!$ is given by*

$$H(x) = C(xe^{-x}) - x + \frac{x^2}{2}. \tag{2.23}$$

Proof. The decomposition of a graph into its core and the attached rooted trees implies the following equation:

$$C(z) = H(T(z)) + U(z). \tag{2.24}$$

The first summand corresponds to the case where the core is non-empty, and the second summand corresponds to the case where the graph is a tree. In order to invert the former relation let $x = T(z)$, so that

$$z = xe^{-x}, \qquad U(z) = x - \frac{x^2}{2}.$$

We obtain

$$H(x) = C(xe^{-x}) - x + \frac{x^2}{2} = \frac{x^3}{3!} + 10\frac{x^4}{4!} + 252\frac{x^5}{5!} + \dots$$

□

Equation (2.23) can be enriched by taking edges into account. The generating functions $T(x, y)$ and $U(x, y)$ are easily obtained as $T(x, y) = T(xy)/y$ and $U(x, y) = U(xy)/y$, and a quick computation gives

$$H(x, y) = C(xe^{-xy}, y) - x + \frac{x^2 y}{2} = y^3\frac{x^3}{3!} + (3y^4 + 6y^5 + y^6)\frac{x^4}{4!} + \dots \tag{2.25}$$

3-graphs. A multigraph is a graph where loops and multiple edges are allowed. As in the case of simple graphs, we define a k-multigraph as a connected multigraph in which the degree of each vertex is at least k. Let $\widetilde{\mathcal{C}}$ be a 2-multigraph. The kernel $\widetilde{\mathcal{K}}$ of $\widetilde{\mathcal{C}}$ is defined as follows: replace every maximal path of vertices of degree two in $\widetilde{\mathcal{C}}$ with a single edge. Clearly $\widetilde{\mathcal{K}}$ is a 3-multigraph, and $\widetilde{\mathcal{C}}$ can be obtained by replacing edges in $\widetilde{\mathcal{K}}$ with paths.

Let $\widetilde{\mathcal{G}}$ be a multigraph. For each $i \geq 1$, let α_i be the number of vertices in $\widetilde{\mathcal{G}}$ that are incident to exactly i loops, and let β_i be the number of i-edges, that is, edges of multiplicity i. The weight of $\widetilde{\mathcal{G}}$ is defined as

$$w(\widetilde{\mathcal{G}}) = \prod_{i\geq 1}\left(\frac{1}{2^i i!}\right)^{\alpha_i} \cdot \prod_{i\geq 1}\left(\frac{1}{i!}\right)^{\beta_i}.$$

This definition is justified by the fact that when replacing an i-edge with i different paths, the order of the paths is irrelevant. Similarly, when replacing a loop with a path, the orientation is irrelevant. Note that the weight satisfies $0 < w(\widetilde{\mathcal{G}}) \leq 1$, and moreover $w(\widetilde{\mathcal{G}}) = 1$ if and only if $\widetilde{\mathcal{G}}$ is simple. With this definition, the sum $\widetilde{K}_n$ of the weights of all 3-multigraphs with n vertices is finite.

As a preliminary step to computing the generating function of 3-graphs, we establish a relation between 3-multigraphs and connected multigraphs. In order to distinguish between edges of different multiplicity, we introduce infinitely many variables as follows. Let $\widetilde{C}_{n,m,l_1,l_2,\ldots}$ be the sum of the weights of connected multigraphs with n vertices, m loops and l_i i-edges for each $i \geq 1$. Define similarly $\widetilde{K}_{n,m,l_1,l_2,\ldots}$ for 3-multigraphs, and let

$$\widetilde{C}(x,z,y_1,y_2,\ldots) = \sum \widetilde{C}_{n,m,l_1,l_2,\ldots} x^n z^m y_1^{l_1} y_2^{l_2} \ldots / n!$$

and

$$\widetilde{K}(x,z,y_1,y_2,\ldots) = \sum \widetilde{K}_{n,m,l_1,l_2,\ldots} x^n z^m y_1^{l_1} y_2^{l_2} \ldots / n!.$$

Theorem 2.3.2. *Let $\widetilde{C}(x,z,y_1,y_2,\ldots)$ and $\widetilde{K}(x,z,y_1,y_2,\ldots)$ be as before. Then*

$$\widetilde{K}(x,z,y_1,y_2,\ldots) = \widetilde{C}\left(xe^{-x(y_1+s)}, -sxy_1 - xy_2 + z, s + y_0, s^2 + 2y_1 s + y_2, \ldots, \sum_{j=0}^{k} \binom{k}{j} y_j s^{k-j}, \ldots\right) + E(x,y_1), \tag{2.26}$$

where

$$y_0 = 1, \quad s = -\frac{xy_1^2}{1+xy_1}, \quad E(x,y) = -x + \frac{x^2 y}{2+2xy} - \ln\sqrt{1+xy} + \frac{xy}{2} - \frac{(xy)^2}{4}.$$

The proof of Theorem 2.3.2 is quite technical and is given below. As a corollary we obtain the generating function of 3-graphs. Recall that $C(x,y)$ is the generating function of connected graphs.

Corollary 2.3.3. *Let $K_{n,m}$ be the number of 3-graphs with n vertices and m edges. The generating function $K(x,y) = \sum K_{n,m} x^n y^m / n!$ is given by*

$$K(x,y) = C\left(A(x,y), B(x,y)\right) + E(x,y), \tag{2.27}$$

where

$$A(x,y) = xe^{(x^2y^3 - 2xy)/(2+2xy)}, \qquad B(x,y) = (y+1)e^{-xy^2/(1+xy)} - 1,$$

and $E(x,y)$ is as in Theorem 2.3.2.

Proof. Since the weight of a simple graph is one, the number of simple 3-graphs is equivalent to the number of weighted 3-multigraphs without loops or multiple edges. This observation leads to

$$K(x,y) = \widetilde{K}(x,0,y,0,\ldots,0,\ldots). \tag{2.28}$$

Moreover, for each connected multigraph $\widetilde{\mathcal{G}}$, a connected simple graph $\mathcal{G}$ can be obtained by removing loops and replacing each multiple edge with a single edge. Then $\widetilde{\mathcal{G}}$ is obtained from $\mathcal{G}$ by replacing each edge with a multiple edge, and attaching zero or more loops at each vertex. This can be encoded as

$$\widetilde{C}(x, z, y_1, y_2, \ldots, y_k, \ldots) = C\left(xe^{z/2}, \sum_{i\geq 1} \frac{y_i}{i!}\right), \tag{2.29}$$

where the exponential and the $1/i!$ terms take care of the weights. Finally, Equation (2.27) can be obtained by combining (2.28), (2.26) and (2.29). □

We remark that a formula equivalent to (2.27) was obtained by Jackson and Reilly [67], using the principle of inclusion and exclusion. Our approach emphasizes the assignment of weights to multigraphs, which are needed in the various combinatorial decompositions.

Note that taking $y = 1$ in Equation (2.27) we obtain the univariate generating function $K(x)$ of 3-graphs as

$$K(x) = K(x, 1) = C(A(x, 1), B(x, 1)) + E(x, 1) \tag{2.30}$$

The proof of Theorem 2.3.2 requires the generating function of 2-multigraphs. Let $\widetilde{H}_{n,m,l_1,l_2,\ldots}$ be the sum of the weights of 2-multigraphs with n vertices, m loops and l_i i-edges ($i \geq 1$), and let

$$\widetilde{H}(x, z, y_1, y_2, \ldots) = \sum \widetilde{H}_{n,m,l_1,l_2,\ldots} x^n z^m y_1^{l_1} y_2^{l_2} \ldots /n!.$$

Lemma 2.3.4. *Let $\widetilde{H}(x, z, y_1, y_2, \ldots)$ and $\widetilde{K}(x, z, y_1, y_2, \ldots)$ be as before. The following equation holds:*

$$\begin{aligned}\widetilde{K}(x, z, y_1, y_2, \ldots, y_k, \ldots) =& \widetilde{H}\left(x, -sxy_1 - xy_2 + z, y_1 + s, y_2 + 2y_1 s + s^2, \ldots, \sum_{j=0}^{k} \binom{k}{j} y_j s^{k-j}, \ldots\right) \\ & - \ln\sqrt{1 + xy_1} - \frac{xz}{2} + \frac{x^2 y_2}{4} + \frac{xy_1}{2} - \frac{(xy_1)^2}{4},\end{aligned} \tag{2.31}$$

where

$$s = -\frac{xy_1^2}{1 + xy_1}.$$

Proof. The kernel of a 2-multigraph is obtained by replacing each edge with a path. This implies the following equation:

$$\begin{aligned}\widetilde{H}(x, z, y_1, y_2, \ldots, y_k, \ldots) =& \widetilde{K}\left(x, sxy_1 + xy_2 + z, y_1 + s, y_2 + 2y_1 s + s^2, \ldots, \sum_{j=0}^{k} \binom{k}{j} y_j s^{k-j}, \ldots\right) \\ & - \ln\sqrt{1 - xy_1} + \frac{xz}{2} + \frac{x^2 y_2}{4} - \frac{xy_1}{2} - \frac{(xy_1)^2}{4},\end{aligned} \tag{2.32}$$

where

$$s = \frac{xy_1^2}{1 - xy_1}.$$

The first summand corresponds to the case where there is at least one vertex of degree ≥ 3, and thus the kernel is not empty. The other summands correspond to cycles (each vertex is of degree exactly two): from the logarithm encoding cycles we must take care of cycles of length one or two.

If the kernel is not empty, we replace every edge and every loop with a path. The expression s encodes a nontrivial path, consisting of at least one vertex. Each loop can be replaced with either another loop, or a vertex and a double edge, or a path consisting of at least two vertices; these operations are encoded, respectively, by z, xy_2 and s. Note that if the kernel has an i-loop, then we can replace any of the loops with a path, in both orientations. Therefore there are $2i$ ways to obtain the same graph, which compensates the fact that the weight of the new graph will be $2i$ times the weight of the old graph. Each k-edge can be replaced with a j-edge and $k - j$ nontrivial paths, where $0 \leq j \leq k$. There are $(k - j)!$ ways to obtain the same graph, and the weight becomes $k!/j!$ times the previous weight. Therefore y_k is replaced with $\binom{k}{j} y_j s^{k-j}$, for $j = 0, \ldots, k$.

A simple computation shows that inverting (2.32) gives (2.31), as claimed. □

Proof of Theorem 2.3.2. Given a multigraph it is clear that every vertex incident to a loop or to a multiple edge belongs to the core. Therefore, Equation (2.25) can be easily extended to multigraphs, giving the equation

$$\widetilde{H}(x, z, y_1, y_2, \ldots, y_k, \ldots) = \widetilde{C}\left(xe^{-xy_1}, z, y_1, y_2, \ldots, y_k, \ldots\right) - x + \frac{x^2 y_1}{2}. \qquad (2.33)$$

Finally, Equation (2.26) is obtained by composing (2.31) and (2.33). □

As mentioned before, Theorem 2.3.1 and Corollary 2.3.3 hold for planar graphs as well. In the next section we use them to enumerate and analyze planar 2- and 3-graphs.

2.4 Planar graphs

In this section we follow the ideas of Section 2.2 on planar maps in order to obtain related results for planar 2-graphs and 3-graphs. The asymptotic enumeration of planar graphs was solved in [8]. From now on we assume that we know the generating function $C(x, y)$ of connected planar graphs, where x marks vertices and y marks edges, as well as its main properties, such as the dominant singularities and the singular expansions around them.

In this section we use the equations obtained in Section 2.3 to compute some parameters in planar graphs. Most of the computations will be analogous to the ones of maps, but technically more involved. In order to compare the following results, we recall [8] that the number of connected planar graphs is $c_n \sim \kappa n^{-7/2} \gamma^n$, where $\kappa \approx 0.4104 \cdot 10^{-5}$ and $\gamma \approx 27.2269$. As expected, there will be slightly fewer connected 2-graphs and

3-graphs than connected planar graphs. Besides, the expected degree of 2-graphs and 3-graphs will be slightly higher.

2.4.1 Planar 2-graphs

We start our analysis with planar 2-graphs. The analysis for 3-graphs in the next subsection is a bit more involved.

Theorem 2.4.1. *Let h_n be the number of planar 2-graphs. The following estimate holds:*

$$h_n \sim \kappa_2 n^{-7/2} \gamma_2^n n!,$$

where $\gamma_2 \approx 26.2076$ and $\kappa_2 \approx 0.3724 \cdot 10^{-5}$.

Proof. Recall Equation (2.23) from Section 2.3:

$$H(x) = C(xe^{-x}) - x + \frac{x^2}{2}.$$

In order to obtain an asymptotic estimate for h_n we need to locate the dominant singularity of $H(x)$. The singularity of $C(x)$ is $\rho = \gamma^{-1} \approx 0.0367$ [8]. Hence the singularity of $H(x)$ is at $\sigma = T(\rho) \approx 0.0382$. Therefore, the exponential growth constant of h_n is $\gamma_2 = \sigma^{-1} \approx 26.2076$. Note that we use the same symbol σ as in Section 2.2 for maps, but they correspond to different constants. No confusion should arise and it helps emphasizing the parallelism between planar maps and graphs.

The singular expansion of $C(x)$ at the singularity $x = \rho$ is

$$C(x) = C_0 + C_2X^2 + C_4X^4 + C_5X^5 + O(X^6),$$

where $X = \sqrt{1 - x/\rho}$, and $C_5 \approx -0.3880 \cdot 10^{-5}$ is computed in [8]. Plugging this expression into (2.23) and expanding gives

$$H(x) = H_0 + H_2X^2 + H_4X^4 + H_5X^5 + O(X^6),$$

where now $X = \sqrt{1 - x/\sigma}$ and $H_5 = C_5(1-\sigma)^{5/2} \approx -0.3520 \cdot 10^{-5}$. The estimate for h_n follows directly by the transfer theorem, provided that H can be extended to a Δ-domain. As opposed to the case of maps, we do not have an exact expression for C, and because of the relation of Equation (2.23), it is not enough to assume that C can be extended to a Δ-domain, since $|(-\sigma)\exp(-(-\sigma))| > \rho$. Instead, we use an alternative expression for H.

Define $A(x)$ as the generating function of connected planar graphs with an unlabelled root where all the vertices except, perhaps, the root, have degree at least 2. If the root has degree 2 then graphs in A are encoded by $H'(x)$. Otherwise either the graph is reduced to a single vertex or the root is connected to a rooted 2-graph through a path of arbitrary length and they are encoded by $\frac{x}{1-x}H'(x)$. Hence we have

$$A(x) = \frac{H'(x)}{1-x} + 1. \tag{2.34}$$

Let $B(x)$ be the generating of planar 2-connected graphs. The unique decomposition of a rooted connected graph into blocks is reflected (see [8]) into the basic equation $C'(x) = \exp(B'(xC'(x)))$. The radius of convergence R of B is given by $R = \rho C'(\rho)$, and R is the only singularity in the circle of convergence of $B(x)$.

A straightforward modification including paths as building blocks in the decomposition gives

$$A(x) = \exp(B'(xA(x)) - x). \tag{2.35}$$

By subtracting x we remove the source of vertices of degree 1: leaves of the block tree decomposition consisting of a single edge. Let $F(x) = \exp(B'(xA(x)) - x)$ be the right-hand side of (2.35). Equation (2.34) shows that A has the same singularities as H in the open ball of radius 1. We now use (2.35) to prove that A, and therefore H, can be extended to a Δ-domain.

The proof has two parts. First we have to prove that A behaves like a square root near its singularity $x = \sigma$. This follows from [47, Theorem 2.31], using $r(x) = R/x$ (in the notation of [47]). Then we need to prove that there is no branch point when solving $A = F(A, x)$ for x in the circle of convergence $|x| = \sigma$. Since $F_A(A, x) = xAB''(xA)$ is a positive function, and $F_A(A(\sigma), \sigma) = RB''(R) < 1$, we have that $|F_A(A(x), x)| < 1$, so it is analytic in a neighbourhood of x. Using the neighbourhood of the singularity, and a finite covering of its compact complement we obtain that A is analytic in a Δ-domain. □

Our next result is a limit law for the number of edges in a random planar 2-graph. We recall [8] that the expected number of edges in random connected planar graphs is asymptotically μn, where $\mu \approx 2.2133$, and the variance is λn with $\lambda \approx 0.4303$.

Theorem 2.4.2. *The number X_n of edges in a random planar 2-graph with n vertices is asymptotically Gaussian with*

$$\mathbf{E}\, X_n \sim \mu_2 n \approx 2.2614n,$$

$$\mathbf{Var}\, X_n \sim \lambda_2 n \approx 0.3843n.$$

Proof. Equation (2.25) from Section 2.3

$$H(x, y) = C(xe^{-xy}, y) - x + \frac{x^2 y}{2}$$

implies that the singularity $\sigma(y)$ of the univariate function $x \mapsto H(x, y)$ is given by

$$\sigma(y)e^{-\sigma(y)y} = \rho(y),$$

where $\rho(y)$ is the singularity of the univariate function $x \mapsto C(x, y)$. An easy calculation gives

$$\mu_2 = -\frac{\sigma'(1)}{\sigma(1)} = \frac{-\rho'(1)/\rho - \sigma}{1 - \sigma} = \frac{\mu - \sigma}{1 - \sigma} \approx 2.2614,$$

which provides the constant for the expectation. Similarly

$$\lambda_2 = -\frac{\sigma''(1)}{\sigma(1)} - \frac{\sigma'(1)}{\sigma(1)} + \left(\frac{\sigma'(1)}{\sigma(1)}\right)^2 =$$
$$\frac{\frac{-\rho''(1)}{\rho(1)} - 3\sigma'(1) - \frac{3\sigma'(1)^2}{\sigma} + \sigma'(1)^2 + 2\sigma'(1)\sigma + \sigma^2 - \frac{\sigma'(1)}{\sigma} + \left(\frac{\sigma'(1)}{\sigma}\right)^2}{1-\sigma}.$$

This value can be computed from the known values of μ, λ and σ. In order to apply the quasi-powers theorem we need to prove that $H(x, y)$ is analytic in a Δ-domain for y close enough to 1. Define $A(x, y)$ as the generating function of connected planar graphs with an unlabelled root where all the vertices except the root have degree at least 2. The following equations are a direct extension of (2.34) and (2.35):

$$A(x, y) = \frac{H_x(x, y)}{1 - xy} + 1,$$
$$A(x, y) = \exp(B_x(xA(x, y), y) - xy) = F(A, x, y).$$

From the first equation we know that A and H have the same singularities for x, y such that $xy < 1$, so we just need to prove that for values y_0 near 1 the function $A(x, y_0)$ is analytic in a Δ-domain. The proof is analogous to that of Theorem 2.4.1. First, $A(x, y)$ behaves like a square root near the singularity $\sigma(y_0)$, again by [47, Theorem 2.31] taking $r(x, u) = R(u)/x$. Then we need that, when $|x| = R(y)$, $F_A(A(x, y), x, y) \neq 1$ holds. Since F_A is positive, $F_A(A(x, 1), x, 1) < 1$, and since both F and A are continuous in y, for values of y close enough to 1 the inequality holds, so again we can extend the generating function to a Δ-domain. □

Next we determine a limit law for the size of the core and the kernel in random connected planar graphs.

Theorem 2.4.3. *The size X_n of the core of a random connected planar graph with n edges is asymptotically Gaussian with*

$$\mathbf{E}\, X_n \sim (1 - \sigma)n \approx 0.9618n, \qquad \mathbf{Var}\, X_n \sim \sigma n \approx 0.0382n.$$

Proof. The generating function $\widehat{C}(x, u)$ of connected planar graphs, where u marks the size of the core, is given by

$$\widehat{C}(x, u) = H(uT(x)) + U(x). \tag{2.36}$$

It follows that the singularity $\xi(u)$ of the univariate function $x \mapsto \widehat{C}(x, u)$ is given by the equation

$$uT(\xi(u)) = \sigma.$$

We can isolate $\xi(u)$ obtaining the explicit formula

$$\xi(u) = \frac{\sigma e^{-\sigma/u}}{u}.$$

An easy calculation gives

$$-\frac{\xi'(1)}{\xi(1)} = 1 - \sigma, \qquad -\frac{\xi''(1)}{\xi(1)} - \frac{\xi'(1)}{\xi(1)} + \left(\frac{\xi'(1)}{\xi(1)}\right)^2 = \sigma.$$

In order to apply the quasi-powers theorem we need to show that, for u_0 close enough to 1 we can extend the generating function $C(x, u_0)$ to a Δ-domain. As in the proof of Theorem 2.4.1, two steps are needed. First, we have to prove that $C(x, u)$ is analytic near $x = \rho(u)$ if $\arg(x/\rho(u) - 1) > \alpha$. We know that this is the case for $H(x)$ near σ, for some angle β. Since $uT(x)$ is analytic, it is conformal and preserves angles locally, hence for u close enough to 1 and x close enough to $\rho(u)$, if $\arg(x/\xi(u) - 1) > \alpha$ for some $\alpha > \beta$, then $uT(x)$ is close to σ and $\arg(T(x)u/\sigma - 1) > \beta$. Then $T(x)u$ is in the region of convergence of H and $C(x, u)$ is analytic in x. On the other hand, if $u = 1$ then $uT(x)$ is a positive function, hence if $|x| = \xi(1)$ but $x \neq \xi(1)$ then $|T(x)| < \sigma$. This implies that if u is close enough to 1 and $|x| = |\xi(u)|$ but far enough from $\xi(u)$, then $|uT(x)u| < \sigma$ by the continuity of $uT(x)u$, so $C(x, u)$ is analytic in a neighbourhood of x. By compactness, a finite number of neighbourhoods is enough, and their union gives a Δ-domain in which $C(x, u)$ is analytic. □

Our next goal is to analyze the size of the trees attached to the core of a random connected planar graph.

Theorem 2.4.4. *Let $X_{n,k}$ count trees with k vertices attached to the core of a random connected planar graph with n vertices. Then $X_{n,k}$ is asymptotically normal and*

$$\mathbf{E}\, X_{n,k} \sim \alpha_k n, \qquad \mathbf{Var}\, X_n \sim \beta_k n,$$

where

$$\alpha_k = \frac{1 - \sigma}{\sigma} \frac{k^{k-1}}{k!} \rho^k,$$

and β_k is described in the proof.

Proof. The generating function of trees where variable w_k marks trees with k vertices is equal to

$$T(x, w_k) = T(x) + (w_k - 1) T_k x^k,$$

where $T_k = k^{k-1}/k!$ is the k-th coefficient of $T(x)$. The composition scheme for the core decomposition is then

$$C(x, w_k) = H(T(x, w_k)) + U(x).$$

It follows that the singularity $\rho_k(w_k)$ of the univariate function $x \mapsto C(x, w_k)$ is given by the equation

$$T(\rho_k(w_k)) + (w_k - 1) T_k (\rho_k(w_k))^k = \sigma.$$

An easy calculation gives

$$\alpha_k = -\frac{\rho_k'(1)}{\rho_k(1)} = \frac{1 - \sigma}{\sigma} \frac{k^{k-1}}{k!} \rho^k$$

$$\beta_k = -\frac{\rho_k''(1)}{\rho_k(1)} - \frac{\rho_k'(1)}{\rho_k(1)} + \left(\frac{\rho'(1)}{\rho(1)}\right)^2 = \frac{1}{\sigma^2}\left(T_k\rho^k(T_k\rho^k(1-2k+4\sigma-2k\sigma^2)+\sigma-\sigma^2\right)$$

The proof that $C(x, w_k)$ can be extended analytically to a Δ-domain is analogous to the proof of Theorem 2.4.3. □

As expected, $\sum_{k\geq 0}\alpha_k = 1-\sigma$, since there are σn vertices not in the core, and therefore there are $(1-\sigma)n$ trees attached to the core. Moreover, $\sum_{k\geq 0} k\alpha_k = 1$, since a connected graph is the union of the trees attached to its core.

To conclude this section, we consider the parameter L_n equal to the size of largest tree attached to the core of a random planar connected graph.

Theorem 2.4.5. *Let L_n be the size of largest tree attached to the core of a random planar connected graph. Then*

$$\frac{L_n}{\log n} \to \frac{1}{\log(1/(e\rho))} \approx 0.4340 \qquad \textit{in probability}$$

and

$$\mathbb{E}L_n \sim \frac{1}{\log(1/(e\rho))}\log n \qquad (n\to\infty).$$

For the proof we need Theorem 1.1 from [70], a kind of Master theorem for proving results on the maximum degree of certain classes random graphs, which can be adapted to other extremal parameters. For completeness we reproduce the statement in full.

Theorem 2.4.6 (Master theorem). *Let $d_{n,k}$ denote the probability that a randomly selected vertex of a certain class of random graphs of size n has degree k, and let $d_{n,k,\ell}$ denote the probability that two different randomly selected (ordered) vertices have degrees k and ℓ. Suppose that we have the following properties.*

1. *There exists a limiting degree distribution $\overline{d}_k$ $(k > 1)$ with an asymptotic behaviour of the form*
$$\log \overline{d}_k \sim k\log q \quad (k\to\infty),$$
where q is a real constant with $0 < q < 1$.

2. *We have, as $n\to\infty, k\to\infty, \ell\to\infty$, and uniformly for $k,\ell \leq C\log n$ (for an arbitrary constant $C > 0$)*
$$d_{n,k} \sim \overline{d}_k \quad \textit{and} d_{n,k,\ell} \sim \overline{d}_k\overline{d}_\ell.$$

3. *There exists $\overline{q} < 1$ such that, uniformly for all $n,k,\ell \geq 1$,*
$$d_{n,k} = O(\overline{q}^k) \quad \textit{and} \quad d_{n,k} = O(\overline{q}^{k+\ell}).$$

Let Δ_n denote the maximum degree of a random graph of size n in this class. Then

$$\frac{\Delta_n}{\log n} \to \frac{1}{\log(1/q)} \quad \textit{in probability}$$

and

$$\mathbf{E}\,\Delta_n \sim \frac{1}{\log(1/q)} \log n \quad (n \to \infty).$$

Proof. The main idea in the proof is to generalize the previous theorem, assigning a numerical "label" ν to each vertex instead of its vertex degree. Given the same hypothesis of the Master theorem in the behaviour of this parameter, the conclusion still holds and we obtain an estimate on the maximum label.

In our case the label is the size of the tree attached to the core that contains the a given vertex. If the graph is itself a tree then all labels are equal to 0 by convention. Therefore, in the rewording of the theorem, $d_{n,k}$ denotes the probability that a randomly selected vertex of a random planar graph of size n has label k, and $d_{n,k,l}$ denotes the probability that two different randomly selected (ordered) vertices have labels k and l. In order to compute such probabilities we define the generating functions $\widehat{C}(x,z)$ and $\widehat{C}(x,z,w)$ as follows: $\widehat{C}(x,z)$ is for connected planar graphs with a root vertex, where x marks vertices and z marks the label of the root. Analogously, $\widehat{C}(x,z,w)$ is for connected planar graphs with two different ordered root vertices, where x marks vertices, z marks the label of the first root, and w the label of the second root. The following relations hold:

$$\widehat{C}(x,z) = H'(T(x))T^{\bullet}(zx) + T(x)$$

$$\widehat{C}(x,z,w) = H''(T(x))T^{\bullet}(zx)T^{\bullet}(wx) + H'(T(x))T^{\bullet\bullet}(zwx) + T^{\bullet}(x)$$

We then have

$$d_{n,k} = \frac{[x^n z^k]\widehat{C}(x,z)}{[x^n]\widehat{C}(x,1)}, \qquad d_{n,k,l} = \frac{[x^n z^k w^l]\widehat{C}(x,z,w)}{[x^n]\widehat{C}(x,1,1)}.$$

Also, note that $\widehat{C}(x,1) = C^{\bullet}(x)$, and $\widehat{C}(x,1,1) = C^{\bullet\bullet}(x)$, which are well-known functions. Next we verify that all the conditions in the Master theorem hold.

Condition 1. Define α_k as in Theorem 2.4.4. Then

$$\overline{d}_k = k \cdot \alpha_k = \frac{1-\sigma}{\sigma}\frac{k^k}{k!}\rho^k.$$

And one easily check that $\log \overline{d}_k \sim k \log(e\rho)$ as $k \to \infty$, as required.

Condition 2. To check this condition we cannot use the quasi-powers theorem, since it only proves the desired result for fixed k. Since we only need the result for k tending to infinity, we can dismiss the graphs whose core is empty. Therefore, for $k \to \infty$,

$$[x^n z^k]\widehat{C}(x,z) \sim [x^n]H'(T(x))[z^k]T^{\bullet}(xz) = [x^{n-k}]H'(T(x)) \cdot [z^k]T^{\bullet}(z).$$

From this we obtain

$$d_{n,k} \sim \frac{[x^{n-k}]H'(T(x))}{[x^n]C^\bullet(x)} \cdot [z^k]T^\bullet(z) \sim \frac{1-\sigma}{\sigma}\left(\frac{n-k}{n}\right)^{-5/2} \rho^k \cdot \frac{1}{\sqrt{2\pi k}} e^k.$$

Finally, when $k \le C \log n$ we have $\left(\frac{n-k}{n}\right)^{-5/2} \to 1$ and thus $d_{n,k} \sim \overline{d}_k$.

Now we have to prove the same for $d_{n,k,l}$. Let $\widehat{C}(x,z,w) = S_1 + S_2 + S_3$, where

$$S_1 = H''(T(x))T^\bullet(zx)T^\bullet(wx), \quad S_2 = H'(T(x))T^{\bullet\bullet}(zwx), \quad S_3 = T^\bullet(x).$$

We know that the coefficients of S_3 are 0 when k and l tend to infinity. Since we differentiate H once instead of twice, it follows that $[x^n z^k w^k]S_2 = O((k/n)[x^n z^k w^k]S_1)$. Since $k = O(\log n)$, the coefficients of S_2 are asymptotically smaller than those of S_1. Therefore, the main asymptotic part comes from S_1. We have

$$[x^n z^k w^l]S_1(x,z,w) = [x^{n-k-l}]H''(T(x)) \cdot [z^k]T^\bullet(z) \cdot [w^l]T^\bullet(w).$$

Then

$$d_{n,k,l} = \frac{[x^{n-k-l}]H''(T(x))}{[x^n]C^{\bullet\bullet}(x)} \cdot [z^k]T^\bullet(z) \cdot [w^l]T^\bullet(w) \cdot (1+o(1)) \sim$$

$$\sim \left(\frac{1-\sigma}{\sigma}\right)^2 \left(\frac{n-k-l}{n}\right)^{-3/2} \rho^{k+l} \cdot \left(\frac{1}{\sqrt{2\pi k}}\right)^2 e^{k+l}.$$

When $k, \ell = O(\log n)$ we have $\left(\frac{n-k-l}{n}\right)^{-3/2} \to 1$, and thus $d_{n,k,l} \sim \overline{d}_k \overline{d}_l$, as required.

Condition 3. We already proved that, for k, $l \ge 1$, and uniformly for any k, l, n, we have

$$[x^n z^k]\widehat{C}(x,z) = f(n,k)(e\rho)^k, \quad [x^n z^k w^l]\widehat{C}(x,z,w) = g(n,k,l)(e\rho)^{k+l},$$

where f and g are subexponential functions, so for any $\overline{q} > e\rho$ we have that $d_{n,k} = O(\overline{q}^k)$ and $d_{n,k,l} = O(\overline{q}^{k+l})$.

Thus the Master theorem applies and we conclude the proof. □

2.4.2 Planar 3-graphs

We recall again that the generating function of connected planar graphs $C(x,y)$, where x marks vertices and y marks edges, was computed in [8].

Theorem 2.4.7. *Let k_n be the number of planar 3-graphs. The following estimate holds:*

$$k_n \sim \kappa_3 n^{-7/2} \gamma_3^n n!,$$

where

$$\gamma_3 \approx 21.3102, \qquad \kappa_3 \approx 0.3107 \cdot 10^{-5}.$$

Proof. Recall Equation (2.30) from Section 2.3:

$$K(x) = C\,(A(x), B(x)) + E(x), \tag{2.37}$$

where $A(x), B(x)$ and $E(x)$ are explicit functions. In order to obtain an estimate for k_n we need to locate the dominant singularity of $K(x)$. The singularities of $C(x, y)$ is given by $(X(t), Y(t))$, where $t \in (0, 1)$ and X, Y are explicit functions defined in [8]. Hence the singularity τ of $K(x)$ is obtained by solving

$$X(t) = A(\tau), \qquad Y(t) = B(\tau).$$

The smallest positive solution τ of the system can be computed numerically and is $\tau \approx 0.0469$. The exponential growth constant is then $\gamma_3 = \tau^{-1} \approx 21.3102$.

The singular expansion of $C(x, y)$ at the singularity $x = \rho(y)$ is of the form

$$C(x, y) = C_0(y) + C_2(y)X^2 + C_4(y)X^4 + C_5(y)X^5 + O(X^6),$$

where $X = \sqrt{1 - x/\rho(y)}$, and $C_5(y)$ is an explicit expression computed in [8]. Plugging this expression into (2.37) and expanding gives

$$K(z) = K_0 + K_2 Z^2 + K_4 Z^4 + K_5 Z^5 + O(Z^6), \tag{2.38}$$

where $Z = \sqrt{1 - z/\tau}$. In order to compute the dominant coefficient K_5, we need to expand $C_5(B(z))\,(1 - D(z))^{5/2}$, where $D(z) = A(z)/\rho(B(z))$, at $z = \tau$. Consider the first-order Taylor expansion of $D(z)$:

$$D(z) = D(\tau) + D'(\tau)(z - \tau) + O((z - \tau)^2).$$

Since $(A(\tau), B(\tau))$ is a singular point of $C(x, y)$, we have

$$A(\tau) = \rho(B(\tau)), \qquad D(\tau) = \frac{A(\tau)}{\rho(B(\tau))} = 1.$$

Therefore, $\sqrt{1 - D(z)}$ is computed as

$$\sqrt{\tau D'(\tau)(1 - z/\tau) + O((x - \tau)^2)} = \sqrt{\tau D'(\tau)}Z + O(Z^2),$$

hence $(1 - D(z))^{5/2} = (\tau D'(\tau))^{5/2} Z^5 + O(Z^6)$. Since $C_5(y)$ is analytic at $y = B(\tau)$, we conclude that $K_5 = C_5(B(\tau))(\tau D'(\tau))^{5/2} \approx -0.2937 \cdot 10^{-5}$. The estimate for k_n follows directly by the transfer theorem, with $\kappa_3 = K_5/\Gamma(-5/2) \approx 0.3107 \cdot 10^{-5}$, provided that K can be analytically extended to a Δ-domain at τ. This is shown in the next lemma, completing the proof of the theorem. □

Lemma 2.4.8. *The generating function $K(x)$ is analytic in a Δ-domain at its dominant singularity τ.*

Proof. Before starting the proof we introduce some definitions:

- $K^\bullet(x)$ is the generating function of rooted planar graphs with minimum degree at least 3. Note that $K^\bullet(x)$ has the same radius of convergence τ as $K(x)$.
- For $i = 1, 2$, $K_i^\bullet(x)$ is the generating function of rooted planar graphs where all the vertices have degree at least 3 except for the root, which has degree exactly i.
- $\widehat{B}(x, u)$ is the generating function of 2-connected planar graphs where x counts vertices of degree at least 3, u counts vertices of degree exactly two, and both kind of vertices are labelled with the same set of labels. In particular $\widehat{B}(x, u) = \sum_{n,m\geq 0} b_{n,m}x^n u^m/(n + m)!$, where $b_{n,m}$ counts the number of 2-connected planar graphs with n vertices of degree at least 3 and m vertices of degree exactly 2. Note that we do not consider a single edge in $\widehat{B}(x, u)$ since it has no vertices of degree 2 or more.

We can relate these generating function using the following positive system of equations:

$$\begin{aligned} K_1^\bullet &= F_1(x, K^\bullet, K_2^\bullet), \\ K_2^\bullet &= F_2(x, K^\bullet, K_1^\bullet, K_2^\bullet), \\ K^\bullet &= F_3(x, K^\bullet, K_1^\bullet, K_2^\bullet), \end{aligned} \tag{2.39}$$

where

$$\begin{aligned} &F_1(x, z, z_2) = x(z + z_2), \\ &F_2(x, z, z_1, z_2) = x\left(\frac{(z + z_2)^2}{2} + B_u\right), \\ &F_3(x, z, z_1, z_2) = x\left(B_x + (z + z_2)(B_u + B_x) + \frac{(B_u + B_x)^2}{2} + \exp_{\geq 3}(z + z_2 + B_u + B_x)\right). \end{aligned}$$

The terms B_x and B_u in the previous system are defined as

$$\begin{aligned} B_x &= \widehat{B}_x(x + z + z_1 + z_2, z + z_1 + z_2), \\ B_u &= \widehat{B}_u(x + z + z_1 + z_2, z + z_1 + z_2). \end{aligned}$$

First, we need to check that F_1, F_2 and F_3 are analytic in a neighbourhood of 0. This is equivalent to checking that B_x and B_u are analytical at 0. We derive this from the following properties of $\widehat{B}$:

- $\widehat{B}$ is a positive series in x and u.
- B_x and B_u are analytic at (x_0, u_0) if and only if $\widehat{B}$ is analytic at (x_0, u_0).
- If $\widehat{B}$ is analytic near (x_0, u_0), then it is analytic at (x_1, u_1), for $|x_1| \leq x_0$ and $|u_1| \leq u_0$.
- $\widehat{B}(x, x) = B(x) - x^2/2$, hence $\widehat{B}(x, u)$ is for $(x_0, u_0) < (R, R)$, although it might be analytic for (x_0, u_0) where $u_0 < R \leq x_0$ or $x_0 < R \leq u_0$ as well.

This implies that $\widehat{B}(x,u)$ is analytic at 0, and the same holds for F_1, F_2 and F_3. Since $K^\bullet(0) = K_1^\bullet(0) = K_2^\bullet(0) = 0$ we have that the system (2.39) holds in a neighbourhood of $x = 0$, and it ceases to hold at the singularity of $K^\bullet$. First note that $K_1^\bullet$, $K_2^\bullet$ and $K^\bullet$ have all the same radius of convergence, τ, because all of them are the sum of the others plus some positive terms, so it is not possible that a singularity appears in one of them and not in the others. Note that, in these cases, there are three sources of singularities:

- Poles at F_1, F_2 and F_3. This is not possible, since all the involved functions are analytic in $\mathbb{C}$ except for B_x and B_u.
- Branching point in solving F_1, F_2 and F_3. This is not possible either, since in this case the singular analysis of $K^\bullet$ would be of the form $K^\bullet = K_0^\bullet + K_1^\bullet Z + O(Z^2)$, where $K_1^\bullet \neq 0$, and we have seen in Equation (2.38) that this is not the case.
- A singularity in $B_x(x)$ and $B_u(x)$ (note that both functions share singularities). This must be the source of singularity, and in fact the singularity of $B_x(x)$ must be exactly at $x = \tau$. If the singularity was at a given $x_0 < \tau$, then there would be an unbounded derivative of B_x at x_0, and since $K^\bullet$ is $x \cdot B_x$ plus some positive terms, then $K^\bullet$ would have an unbounded derivative at $x_0 < \tau$, and that is impossible since $K^\bullet$ is analytic for x with $|x| < \tau$. The singularity cannot be at an $x_0 > \tau$ either, because we discarded the other sources of singularities and this would imply that $K^\bullet$ is analytic for $x > \tau$, which is impossible.

Therefore the equations hold for x such that B_x is analytic at $(x + K_1^\bullet(x) + K_2^\bullet(x) + K^\bullet(x), K_1^\bullet(x) + K_2^\bullet(x) + K^\bullet(x))$. Now, consider x such that $|x| = \tau$ but $x \neq \tau$. Then, by positivity of $\widehat{B}$ and $K_i^\bullet$, we have: $(|x + K_1^\bullet(x) + K_2^\bullet(x) + K^\bullet(x)|, |K_1^\bullet(x) + K_2^\bullet(x) + K^\bullet(x)|) < (\tau + K_1^\bullet(\tau) + K_2^\bullet(\tau) + K^\bullet(\tau), K_1^\bullet(\tau) + K_2^\bullet(\tau) + K^\bullet(\tau))$, so $\widehat{B}$ is analytic and the equations hold. Therefore $K^\bullet(x)$ is analytic as well.

We just have to check that, if $|x| = \tau$ and $x \neq \tau$ then there is not branching point when solving the system of equations. Let A be the Jacobian matrix of (F_1, F_2, F_3). According to [47, Section 2.2.5], the maximum positive eigenvalue of A is a positive function in x, $K_i^\bullet$. We know that such an eigenvalue must be smaller than 1 when evaluated at $(\tau, K_1^\bullet(\tau), K_2^\bullet(\tau), K^\bullet(\tau))$, since otherwise there would exist a real x with $|x| < \tau$ such that the system evaluated at x has a branching point, and we know this is not possible. Hence, by positivity of the maximum eigenvalue, if $|x| = \tau$ but $x \neq \tau$ then the maximum eigenvalue of A evaluated at $(x, K_1^\bullet(x), K_2^\bullet(x), K^\bullet(x))$ cannot be 1, so we can apply the Implicit Function Theorem and there is an analytic continuation of $K^\bullet$ in a neighbourhood of x, and by compactness it can be extended to a Δ-domain, as we wanted to prove. □

Our next result is a limit law for the number of edges in a random planar 3-graph.

Theorem 2.4.9. *The number X_n of edges in a random planar 3-graph with n vertices is asymptotically Gaussian with*

$$\mathbf{E}\, X_n \sim \mu_3 n \approx 2.4065n, \qquad \mathbf{Var}\, X_n \sim \lambda_3 n \approx 0.3126n$$

Proof. Recall Equation (2.27) from Section 2.3:

$$K(x,y) = C\left(A(x,y), B(x,y)\right) + E(x,y), \tag{2.40}$$

where

$$A(x,y) = xe^{(x^2y^3-2xy)/(2+2xy)}, \qquad B(x,y) = (y+1)e^{-xy^2/(1+xy)} - 1,$$

$$E(x,y) = -x + \frac{x^2y}{2+2xy} - \ln\sqrt{1+xy} + \frac{xy}{2} - \frac{(xy)^2}{4}.$$

It follows that the singularity $\tau(y)$ of the univariate function $x \mapsto K(x,y)$ is given by the equation

$$A(\tau(y), y) = \rho(B(\tau(y), y)),$$

where $\rho(y)$ is as before the singularity of $x \mapsto C(x,y)$. The value of $\tau(1) = \tau$ is already known. In order to compute $\tau'(1)$ we differentiate and obtain

$$A_x(\tau,1)\tau'(1) + A_y(\tau,1) = \rho'(B(\tau,1))\left[B_x(\tau,1)\tau'(1) + B_y(\tau,1)\right].$$

Solving for $\tau'(1)$ we obtain

$$\tau'(1) = -\frac{\rho'(B(\tau,1))B_y(\tau,1) - A_y(\tau,1)}{\rho'(B(\tau,1))B_x(\tau,1) - A_x(\tau,1)}.$$

Since $\rho = X \circ Y^{-1}$, where X and Y are explicit functions defined in [8], $\rho'(y)$ can be computed as $X'(Y^{-1}(y))/Y'(Y^{-1}(y))$. After some calculations we finally get a value of $\tau'(1) \approx -0.1129$ and

$$\mu_3 = -\frac{\tau'(1)}{\tau(1)} \approx 2.4065.$$

Using the same procedure we can isolate $\tau''(1) \approx 0.3700$ and obtain the variance as

$$\lambda_3 = -\frac{\tau''(1)}{\tau(1)} - \frac{\tau'(1)}{\tau(1)} + \left(\frac{\tau'(1)}{\tau(1)}\right)^2 \approx 0.3126.$$

In order to apply quasi-powers theorem we just have to prove that $K(x,y)$ is analytic in a Δ-domain for y close enough to 1. The proof is analogous to the proof of Theorem 2.4.7. We just have to adapt the equations so that they count edges as well:

$$K_1^\bullet = F_1(x, K^\bullet, K_2^\bullet, y),$$

$$K_2^\bullet = F_2(x, K^\bullet, K_1^\bullet, K_2^\bullet, y),$$

$$K^\bullet = F_3(x, K^\bullet, K_1^\bullet, K_2^\bullet, y),$$

where

$$F_1(x, K^\bullet, K_2^\bullet, y) = xy(K_2^\bullet + K^\bullet),$$

$$F_2(x, K^\bullet, K_1^\bullet, K_2^\bullet, y) = x\left(\frac{y^2(K_2^\bullet + K^\bullet)^2}{2} + B_u\right),$$

$$F_3(x, K^\bullet, K_1^\bullet, K_2^\bullet, y) =$$

$$= x\left(B_x + y(K_2^\bullet + K^\bullet)(B_u + B_x) + \frac{(B_u + B_x)^2}{2} + \exp_{\geq 3}(yK_2^\bullet + yK^\bullet + B_u + B_x)\right),$$

and where we use, for short, the notation:

$$B_x = \widehat{B}_x(x + K_1^\bullet + K_2^\bullet + K^\bullet, K_1^\bullet + K_2^\bullet + K^\bullet, y)$$

$$B_u = \widehat{B}_u(x + K_1^\bullet + K_2^\bullet + K^\bullet, K_1^\bullet + K_2^\bullet + K^\bullet, y),$$

where $\widehat{B}(x, u, y)$ is the enriched version of $\widehat{B}(x, u)$, defined in Theorem 2.4.7, where y counts edges. The rest of the proof is identical: since F_i are continuous in y and positive, the largest eigenvalue will have absolute value smaller than 1 if y is close enough to 1, so there will be no branching point when solving the system of equations. □

Next we determine the limit law for the size of the kernel in random planar 2-graphs.

Theorem 2.4.10. *The size Y_n of the kernel of a random planar 2-graph with n edges is asymptotically Gaussian with*

$$\mathbf{E}\, Y_n \sim \mu_K n \approx 0.8259n, \qquad \mathbf{Var}\, Y_n \sim \lambda_K n \approx 0.1205n \tag{2.41}$$

Proof. Recall that the decomposition of a simple 2-graph into its kernel gives

$$H(x) = \widetilde{H}(x, 0, 1, 0, \ldots) = \widetilde{K}\left(x, \frac{x^2}{1-x}, \frac{1}{1-x}, \ldots, k\left(\frac{x}{1-x}\right)^{k-1} + \left(\frac{x}{1-x}\right)^k, \ldots\right) + E(x, 1).$$

If u marks the size of the kernel then

$$H(x, u) = \widetilde{K}\left(ux, \frac{x^2}{1-x}, \frac{1}{1-x}, \ldots, k\left(\frac{x}{1-x}\right)^{k-1} + \left(\frac{x}{1-x}\right)^k, \ldots\right) + E(x, 1).$$

Composing with Equations (2.26) and (2.29) we get

$$H(x, u) = C\left(A(x, u), B(x, u)\right) + F(x, u)$$

where

$$A(x, u) = ux \exp\left(\frac{-x\left(2u + x + u^2x - 2\,ux\right)}{2(1 - x + ux)}\right),$$

$$B(x, u) = -1 + 2\exp\left(\frac{x\,(1-u)}{1 - x + ux}\right),$$

and $F(x, u)$ is a correction term which does not affect the singular analysis. It follows that the singularity $\chi(u)$ of the univariate function $x \mapsto H(x, u)$ is given by the equation

$$A(\chi(u), u) = \rho(B(\chi(u), u)),$$

If we differentiate the former expression and replace u with 1 we get

$$A_x(\sigma,1)\chi'(1) + A_y(\sigma,1) = \rho'(1)(B_x(\sigma,1)\chi'(1) + B_y(\sigma,1)).$$

Note that $\chi(1) = \sigma$, where σ is, as before, the singularity of the generating function $H(x)$ of planar 2-graphs. Moreover, $B(x,1) = 1$. After some calculations we finally get $\chi'(1) \approx -0.03135$ and

$$\mu_K = -\frac{\chi'(1)}{\chi(1)} = \frac{2\rho'(1)e^{\sigma} + \sigma^2 - \sigma + 1}{1-\sigma}.$$

This is computed using the known values of σ and $\rho'(1) = -\rho\mu$. Using the same procedure we can isolate $\chi''(1) \approx 0.05295$ and compute λ_K as

$$\lambda_K = -\frac{\chi''(1)}{\chi(1)} - \frac{\chi'(1)}{\chi(1)} + \left(\frac{\chi'(1)}{\chi(1)}\right)^2 \approx 0.1205.$$

We need to show that $H(x,u)$ is analytic in a Δ-domain. If $u = 1$ we already know it for $H(x,1)$. Since $A(x,u)$ and $B(x,u)$ are both analytic, and $A(\sigma,1) = \rho$ and $B(x,1) = 1$, then for u close enough to 1 and x close enough to $\chi(u)$, by continuity, if $\arg(x/\chi(u) - 1) > \alpha$ then $\arg(A(x,u)/\rho(B(x,u)) - 1) > \beta$ for some $\beta > 0$, as in the proof of Theorem 2.4.3. Also, if $|x| = \sigma$ but $x \neq \sigma$, then we know that $H(x,1)$ is analytic near x. Again by continuity, if u is close enough to 1 then $H(x,u)$ is analytic at (x,u), and by compactness this is sufficient to prove analyticity in a Δ-domain. □

Note that, since the expected size of the core of a random connected planar graph is $1-\sigma$, the expected size of the kernel of a random connected planar graph with n vertices is asymptotically $(1-\sigma)\mu_K n = (2\rho'(1)e^{\sigma} + \sigma^2 - \sigma + 1)n \approx 0.7944n$.

2.5 Degree distribution

In this section we compute the limit probability that a vertex of a planar 2-graph or 3-graph has a given degree. In order to do that, we compute the probability distribution of the root of a rooted planar 2-graph and 3-graph. Since every vertex is equally likely to be the root, we conclude that the average distribution is the same. Note that this is not true for maps, so in this section we only compute the distribution for graphs. This section is rather technical, especially the part of 3-graphs, so that is why we separate its content from that of Section 2.4.

Let $c_n^\bullet$ be the number of rooted connected planar graphs with n vertices, i.e., $c_n^\bullet = n \cdot c_n$. Let $C^\bullet(x) = \sum c_n^\bullet x^n = xC'(x)$ be its associated generating function. Let $c_{n,k}^\bullet$ be the number of rooted connected planar graphs with n vertices and such that the root degree is exactly k. Let $C^\bullet(x,w) = \sum c_{n,m}^\bullet x^n u^m$ be its associated generating function. The limit probability d_k that the root vertex has degree k can be obtained as

$$d_k = \lim_{n\to\infty} \frac{c_{n,k}^\bullet}{c_n^\bullet} = \lim_{n\to\infty} \frac{[x^n][w^k]C^\bullet(x,w)}{[x^n]C^\bullet(x)}.$$

Therefore, the probability distribution $p(w) = \sum d_k w^k$ can be obtained from the knowledge of $C^\bullet(w,u)$. In [21] this function is computed, and d_k is proven to be asimptotically

$$d_k \sim c \cdot k^{-1/2} q^k,$$

where $c \approx 3.0175$ and $q \approx 0.6735$ are computable constants. Our goal is to obtain similar results for 2-graphs and 3-graphs, by respectively computing generating function $H^\bullet(x,w)$ and $K^\bullet(x,w)$ in terms of $C^\bullet(x,w)$.

2.5.1 2-graphs

Theorem 2.5.1. *Let $h^\bullet_{n,k}$ be the number of rooted 2-graphs with n vertices and with root degree k. Let $H^\bullet(x,w) = \sum h^\bullet_{n,k} x^n w^k$ be its associated generating function. The following equation holds*

$$H^\bullet(x,w) = e^{x(1-w)} C^\bullet(xe^{-x}, w) - xwC^\bullet(xe^{-x}) - x + x^2 w \tag{2.42}$$

Proof. The decomposition of a graph into ins core and the attached rooted trees implies the following equation:

$$C^\bullet(z,w) = H^\bullet(T(z),w)\frac{T(z,w)}{T(z)} + H^\bullet(T(z))\frac{wT(z,w)}{1-T(z)} + T(z,w),$$

where $T(z,w) = z \cdot e^{wT(z)}$ is the generating function of rooted trees where w marks the degree of the root. The first addend corresponds to the case where the root is in the core. In this case, the degree of the graph root is the degree of the core root plus the degree of the root of its appended tree. The second addend corresponds to the case where the root is in an attached tree. In this case there is a sequence of trees between the core and the root, and finally a rooted tree. The degree of the graph root is the degree of the root of the rooted tree plus one. The last addend corresponds to the case where the graph is a tree, and therefore its core is empty.

In order to invert the former relation let $x = T(z)$ so that

$$z = xe^{-x}, \quad T(z,w) = xe^{-x(1-w)}, \quad H^\bullet(T(z)) = (1-x)C^\bullet(xe^{-x}) + x^2 - x.$$

After some calculations we obtain

$$H^\bullet(x,w) = e^{x(1-w)} C^\bullet(xe^{-x}, w) - xwC^\bullet(xe^{-x}) - x + x^2 w =$$

$$= \frac{1}{2}w^2x^3 + \left(w^2 + \frac{2}{3}w^3\right)x^4 + \left(\frac{9}{2}w^2 + \frac{13}{3}w^3 + \frac{41}{24}w^4\right)x^5 + \dots$$

□

The probability distribution $p(w)$ can be computed using transfer theorems. The expansion of $C^\bullet(x,w)$ near the singularity $x = \rho$ gives the following equation

$$C^\bullet(x,w) = C_0(w) + C_2(w)X^2 + C_3(w)X^3 + O(X^4), \tag{2.43}$$

where $X = \sqrt{1 - x/\rho}$. The probability distribution can be computed as

$$p(w) = \frac{C_3(w)}{C_3(1)}.$$

Our goal is to obtain the same result by applying the relation obtained in (2.42).

Theorem 2.5.2. *Let e_k be the limit probability that a random vertex has degree k in a 2-graph. Let $p_H(w) = \sum e_k w^k$ be its probability distribution. Let $p(x)$ be as before. The following equation holds:*

$$p_H(w) = \frac{e^{\sigma(1-w)}p(w) - \sigma w}{1 - \sigma}, \tag{2.44}$$

where $\sigma = T(\rho)$, as in Theorem 2.4.1. Furthermore, the limiting probability that the degree of a random vertex is equal to k exists, and is asymptotically

$$p_H(k) \sim \nu_2 k^{-1/2} q^k,$$

where $q \approx 0.6735$ and $\nu_2 \approx 3.0797$.

Proof. Since $C^\bullet(x, w)$ satisfies (2.43), and $H^\bullet(x, w)$ satisfies (2.42), we obtain

$$H^\bullet(x, w) = H_0(w) + H_2(w)X^2 + H_3(w)X^3 + O(X^4),$$

where $X = \sqrt{1 - x/\sigma}$, and $H_3(w)$ is computed as

$$H_3(w) = e^{\sigma(1-w)}C_3(w)(1 - \sigma)^{3/2} - w\sigma C_3(1)(1 - \sigma)^{3/2}$$

The probability generating function of the distribution is given by

$$p_H(w) = \frac{H_3(w)}{H_3(1)} = \frac{(1 - \sigma)^{3/2}\left(e^{\sigma(1-w)}C_3(w) - w\sigma C_3(1)\right)}{(1 - \sigma)^{3/2}C_3(1)(1 - \sigma)} = \frac{e^{\sigma(1-w)}p(w) - \sigma w}{1 - \sigma}.$$

The asymptotics of the distribution can be obtained from $p(w)$. The singularity of $p(w)$ is obtained in [21] as $r \approx 1.4849$. The expansion of $p(w)$ near the singularity is computed as

$$p(w) = P_{-1}W^{-1} + O(1),$$

where $P_{-1} \approx 5.3484$ is a computable constant, and $W = \sqrt{1 - w/r}$. Plugging this expression into (2.44) we get

$$p_H(w) = Q_{-1}W^{-1} + O(1),$$

where $Q_{-1} = P_{-1}e^{\sigma(1-r)}/(1 - \sigma) \approx 5.4586$. The estimate for $p_H(k)$ follows directly by singularity analysis. ☐

2.5.2 3-graphs

In order to prove a similar result for 3-graphs, we need to extend the generating function $C^\bullet(x.w)$ so that it takes edges into account. This function $C^\bullet(x,y,w)$ was computed in [21], and our goal is to obtain the analogous generating function for 3-graphs, $K^\bullet(x,w)$, in terms of We remark that the expression given in [21] for $C^\bullet(x,y,w)$ is extremely involved and needs several pages to write it down.

Theorem 2.5.3. *Let $k^\bullet_{n,k}$ be the number of rooted 3-graphs with n vertices and with root degree k. Let $K^\bullet(x,w) = \sum k^\bullet_{n,k} x^n w^k$ be its associated generating function. The following equation holds*

$$K^\bullet(x,w) = B_0(x,w) \cdot C^\bullet\left(B_1(x), B_2(x), B_3(x,w)\right) + A(x,w) \tag{2.45}$$

where

$$B_0(x,w) = e^{(w^2-1)x^2/(2+2x)+x(1-w)/(1+x)}, \qquad B_1(x) = xe^{(x^2-2x)/(2+2x)},$$

$$B_2(x) = 2e^{-x/(1+x)} - 1, \qquad B_3(x,w) = \frac{(1+w)e^{-wx/(1+x)} - 1}{2e^{-x/(1+x)} - 1},$$

$$A(x,w) = A_0(x) + A_1(x)w + A_2(x)w^2,$$

and $A_0(x)$, $A_1(x)$, $A_2(x)$ are analytic functions.

In order to prove this theorem we need some technical lemmas that relate different classes of graphs.

Lemma 2.5.4. *Let $\widetilde{C}^\bullet(x,w,z,y_1,\dots,y_k,\dots)$ be the generating function of rooted connected planar weighted multigraphs where x marks vertices, w marks the root degree, z marks loops, and y_k marks k-edges. The following equation holds*

$$\widetilde{C}^\bullet(x,w,z,y_1,\dots,y_k\dots) = e^{z\cdot(w^2-1)/2} C^\bullet\left(xe^{z/2}, \sum_{i\geq 1}\frac{y_i}{i!}, \frac{\sum_{i\geq 1} w^i \cdot y_i/i!}{\sum_{i\geq 1} y_i/i!}\right). \tag{2.46}$$

Proof. Given a simple connected planar graph $\mathcal{G}$, a connected planar multigraph can be obtained from $\mathcal{G}$ by replacing each edge with a multiple edge, and placing 0 or more loops in each vertex (see proof of Corollary 2.3.3 for details). In the case of rooted graphs, if we replace an edge incident to the root with a i edge, its root degree is increased in $i-1$. Therefore, instead of replacing such an edge with a multiple edge with generating function $y_i/i!$, we replace it with a multiple edge with generating function $w^i y_i/i!$. Similarly, when we add a loop incident to the root vertex, the root degree is increased by 2. Therefore, its associated generating function is not z, but zw^2. □

Lemma 2.5.5. *Let $\widetilde{H}^\bullet(x,w,z,y_1,\dots,y_k,\dots)$ be the generating function of rooted planar weighted 2-multigraphs where x marks vertices, w marks the root degree, z marks loops, and y_k marks k-edges. The following equation holds*

$$\widetilde{H}^\bullet(x,w,z,y_1,\dots,y_k\dots) = e^{y_1 x(1-w)}\widetilde{C}^\bullet(xe^{-y_1 x}, w, z, y_1,\dots,y_k\dots) - w\cdot A(x,z,y_1,\dots y_k,\dots) - x, \tag{2.47}$$

for a given function $A(x,z,y_1,\dots y_k,\dots)$ that does not depend on w.

Proof. The decomposition of a planar connected weighted multigraph into its core and the attached rooted trees implies the following equation:

$$\widetilde{C}^{\bullet}(x,w,z,y_1,\ldots,y_k,\ldots) = \widetilde{H}^{\bullet}(T(x,y_1),w,z,y_1,\ldots,y_k,\ldots)\frac{T(x,y_1,w)}{T(x,y_1)}+$$

$$+\widetilde{H}^{\bullet}(T(x,y_1),z,y_1,\ldots,y_k,\ldots)\frac{wT(x,y_1,w)}{1-T(x,y_1)}+T(x,y_1,w),$$

where $T(x,y) = T(xy)/y$ is the generating function of rooted trees where x marks vertices and y marks edges, and $T(x,y,w) = T(xy,w)/y$ is the generating function of rooted trees where x marks vertices, y marks edges, and w marks the root degree. The justification of this relation is analogous to the proof of Theorem 2.5.1, as well as the inverse. □

Lemma 2.5.6. *Let $K^{\bullet}(x,w)$ be the generating function of rooted simple planar 3-graphs where x marks vertices and w marks the root degree. The following equation holds*

$$K^{\bullet}(x,w) = \widetilde{H}^{\bullet}(x,w,-sx,1+s,2s+s^2,\ldots,ks^{k-1}+s^k,\ldots)+w^2A(x), \qquad (2.48)$$

for a given function $A(x)$, and where $s=-x/(1+x)$.

Proof. Recall from Section 2.3 the decomposition (2.32) of a planar 2-multigraph into its kernel and paths of vertices

$$\widetilde{H}(x,z,y_1,y_2,\ldots,y_k,\ldots) = \widetilde{K}\left(x, sxy_1+xy_2+z, y_1+s, y_2+2y_1s+s^2,\ldots,\sum_{j=0}^{k}\binom{k}{j}y_js^{k-j},\ldots\right)$$

$$-\ln\sqrt{1-xy_1}+\frac{xz}{2}+\frac{x^2y_2}{4}-\frac{xy_1}{2}-\frac{(xy_1)^2}{4},$$

where $s = xy_1^2/(1-xy_1)$ is a nonempty path of edges and vertices. If we root a vertex of a planar 2-multigraph there are two options: either it belongs to the kernel or it belongs to an edge of the kernel. In the former case, its degree corresponds to the degree of the corresponding vertex in the kernel. In the latter case its degree must be 2. With this observation we can extend this equation so that it considers rooted graphs and it takes the root degree into account, as

$$\widetilde{H}^{\bullet}(x,w,z,y_1,y_2,\ldots,y_k,\ldots) =$$
$$\widetilde{K}^{\bullet}\left(x,w,sxy_1+xy_2+z,y_1+s,y_2+2y_1s+s^2,\ldots,\textstyle\sum_{j=0}^{k}\binom{k}{j}y_js^{k-j},\ldots\right)+w^2A(x,z,y_1,\ldots,y_k,\ldots),$$

where $A(x,z,y_1,\ldots,y_k,\ldots)$ is a function that does not depend on w. This relation can be inverted as in Section 2.3, and finally we can conclude (2.48) from the following equation

$$K^{\bullet}(x,w) = \widetilde{K}^{\bullet}(x,w,0,1,0,\ldots,0,\ldots).$$

□

Using these lemmas we finally prove Theorem 2.5.3.

Proof. The equation (2.45) is obtained by combining equations (2.48), (2.47) and (2.46). □

The expression obtained in Theorem 2.5.3 allows us to prove the following result.

Theorem 2.5.7. *Let f_k be the limit probability that a random vertex has degree k in a planar 3-graph. The limit probability distribution $p_K(w) = \sum f_k w^k$ exists and is computable.*

Proof. The generating function $C^\bullet(x, y, w)$ is expressed in [21] as

$$C^\bullet(x, y, w) = C_0(y, w) + C_2(y, w)X^2 + C_3(y, w)X^3 + O(X^4),$$

where $X = \sqrt{1 - x/\rho(y)}$. If we compose this expression with (2.45) we obtain

$$K^\bullet(x, w) = B_0(x, w)\times \\ \big[C_0(B_2(x)), B_3(x, w)) + C_2(B_2(x), B_3(x, w))X^2 + C_3(B_2(x), B_3(x, w))X^3 + O(X^4)\big] + A(x, w), \tag{2.49}$$

where $X = \sqrt{1 - B_1(x)/\rho(B_2(x)}$. If we define $D(x) = B_1(x)/\rho(B_2(x))$ then we can proceed as in the proof of Theorem 2.4.7, obtaining that $X = \sqrt{\tau D'(\tau)}Z + O(Z^2)$, where $Z = \sqrt{1 - x/\tau}$. Plugging this expression into (2.49) we obtain

$$K^\bullet(z, w) = K_0(w) + K_2(w)Z^2 + K_3(w)Z^3 + O(Z^4),$$

where $Z = \sqrt{1 - z/\tau}$ and

$$K_3(w) = B_0(\tau, w)C_3(B_2(\tau), B_3(\tau, w))(\tau D'(\tau))^{3/2} + a_0 + a_1 w + a_2 w^2,$$

for some constants a_0, a_1 and a_2. The limit probability distribution of the root vertex being of degree k is computed as

$$p_K(w) = \frac{K_3(w)}{K_3(1)} = \frac{B_0(\tau, w)C_3(B_2(\tau), B_3(\tau, w))(\tau D'(\tau))^{3/2} + a_0 + a_1 w + a_2 w^2}{B_0(\tau, 1)C_3(B_2(\tau), 1)(\tau D'(\tau))^{3/2} + a_0 + a_1 + a_2}.$$

Since we know that a 3-graph has no vertices of degree 0, 1 or 2, we can choose suitable values of a_0, a_1 and a_2 such that the probability distribution $p_K(w) = \sum f_k w^k$ satisfies $f_0 = f_1 = f_2 = 0$. The function $C_3(y, w)$ is described in [21], and every other function that appears in the previous expression is explicit. Therefore, p_K is computable, as we wanted to prove. □

We remark that $p_K(w)$ is expressed in terms of $C_3(x, w)$, which is a very involved (although elementary) function, given in the appendix in [21].

2.6 Concluding remarks

Most of the results we have obtained can be extended to other classes of graphs. Let G be a class of graphs closed under taking minors such that the excluded minors of G are 2-connected. Interesting examples are the classes of series-parallel and outerplanar graphs. Given such a class G, a connected graph is in G if and only if its core is in G. Hence Equation (2.23) also holds for graphs in G. Using the results from [24], we have performed the corresponding computations for the classes of series-parallel and outerplanar graphs (there are no results for kernels since outerplanar and series-parallel have always minimum degree at most two). The results are displayed in the next table, together with the data for planar graphs. The expected number of edges is μn, and the expected size of the core is κn. It is worth remarking that the size of the core is always linear, whereas the size of the largest block in series-parallel and outerplanar graphs is only $O(\log n)$ [28, 32].

Graphs	Growth constant	μ (edges)	κ (core)
Outerplanar	7.32	1.56	0.84
Outerplanar 2-graphs	6.24	1.67	
Series-parallel	9.07	1.62	0.875
Series-parallel 2-graphs	8.01	1.70	
Planar	27.23	2.21	0.962
Planar 2-graphs	26.21	2.26	

The k-core of a graph G is the maximum subgraph of G in which all vertices have degree at least k. Equivalently, it is the subgraph of G formed by deleting repeatedly (in any order) all vertices of degree less than k. In this terminology, what we have called the core of a graph is the 2-core. Using the results from [8] it is not difficult to show that the 3-core, 4-core and 5-core of a random planar graph have all linear size with high probability (there is no 6-core since a planar graph has always a vertex of degree at most five). The interesting question is however whether the k-core has a connected component of linear size (as is the case for $k = 2$). We have performed computational experiments on random planar graphs, using the algorithm described in [74], and based on the results we formulate the following conjecture.

Conjecture. With high probability the 3-core of a random planar graph has one component of linear size. With high probability the components of the 4-core of a random planar graph are all sublinear.

We have not been able to prove neither of the conjectures. As opposed to the kernel, the 3-core is obtained by repeatedly *removing* vertices of degree two. These deletions may have long-range effects that appear difficult to analyze. Even more challenging appears the analysis of the 4-core.

Chapter 3

Subgraph statistics in series-parallel graphs

This chapter is based in a joint work with J. Rué and M. Drmota, accepted in the journal Random Structures and Algorithms. Let H be a fixed graph and $\mathcal{G}$ a subcritical graph class. We show that the number of occurrences of H (as a subgraph) in a graph in $\mathcal{G}$ of order n, chosen uniformly at random, follows a normal limiting distribution with linear expectation and variance. The main ingredient in our proof is the analytic framework developed by Drmota, Gittenberger and Morgenbesser to deal with infinite systems of functional equations [61]. As a case study, we obtain explicit expressions for the number of triangles and cycles of length 4 in the family of series-parallel graphs.

3.1 Introduction

The study of subgraphs in random discrete structures is a central area in graph theory, which dates back to the seminal works of Erdős and Rényi in the sixties [35]. Since then, a lot of effort has been devoted to locate the threshold function for the number of copies of a given subgraph in the $G(n,p)$ random graph model, as well as the limiting distribution of the corresponding counting random variable (see for instance [36, 4, 52], and the monograph [41, Chapter 3]). The number of copies of a fixed graph and its statistics had been also addressed in other restricted graph classes, including regular graphs, random graphs with specified degree sequences (see for instance, [56, 22, 58, 53, 55], see also [57]) and random planar maps [50, 11].

In this chapter we study subgraphs of a graph chosen at random from a so-called subcritical class. Recall that a block is a maximal 2-connected subgraph. Roughly speaking, a graph class is called *subcritical* if the largest block of a random graph in the class with n vertices has $O(\log(n))$ vertices (see the precise analytic definition in Section 3.3). Indeed, graphs in these classes have typically a tree-like structure and share several properties with trees. Just to mention some families, prominent subcritical graph classes are forests, cacti trees, outerplanar graphs and series-parallel graphs, and more generally graph families defined by a finite set of 3-connected components (see the work of Giménez, Noy and Rué [28]). Let us mention that the analysis of subcritical graph classes is intimately related to the study of the random planar graph model: it is conjectured that a graph class defined by a set of excluded minors is subcritical if and only if at least one of the excluded graphs is planar (see [45]).

The systematic study of subcritical graph classes was started by Bernasconi, Panagiotou and Steger in [65] when studying the expected number of vertices of given degree. Later, Drmota, Fusy, Kang, Kraus and Rué in [12] extended the analysis to unlabelled graph classes, and obtained normal limiting probability distributions for different parameters, including the number of cut-vertices, blocks, edges and the vertex degree distribution. Drmota and Noy [27] investigated several extremal parameters in these graph classes. They showed, for instance, that the expected diameter D_n of a random connected graph from a subcritical graph class on n vertices satisfies $c_1\sqrt{n} \leq \mathbb{E}[D_n] \leq c_2\sqrt{n\log n}$ for some constants c_1 and c_2. More recently, the precise asymptotic estimate has been deduced to be of order $\Theta(\sqrt{n})$ (see [49]). Furthermore, the normalized metric space $(V(G), d_G/\sqrt{n})$ (where $d_G(u,v)$ denotes the distance between u and v in the graph G) is shown to converge (with respect to the Gromov-Hausdorff metric) to the so-called *Brownian Continuum Random Tree*, multiplied by a scaling factor that depends only on the class under study (see [49] for details, and also [40] for extensions to the unlabelled setting). Let us also mention that even more recently, the Benjamini-Schramm convergence had been addressed as well in [40, 31] for these graph families. Finally, the maximum degree and the degree sequence of a random series-parallel graph have been studied in [70] and [65, 75], respectively.

Our results: this chapter is a contribution to the understanding of the structure of a random graph from one of these graph classes. More precisely, we present a very general framework to deal with subgraph statistics in subcritical graph classes. Our main result is the following theorem:

Theorem 3.1.1. *Let $\mathcal{C}_n$ be the set of connected graphs in a subcritical graph class $\mathcal{G}$ of order n, and let H be a fixed connected graph. Let X_n^H be the number of copies of H in an object in $\mathcal{C}_n$, chosen uniformly at random. Then,*

$$\mathbb{E}[X_n^H] = \mu_H n + O(1) \text{ and } \mathbb{V}\mathrm{ar}[X_n^H] = \sigma_H^2 n + O(1)$$

for some constants $\mu_H > 0$, $\sigma_H^2 \geq 0$ that only depend on H (and on the subcritical graph class under study). Moreover, if $\sigma_H^2 > 0$, then the random variable

$$\frac{X_n^H - \mathbb{E}[X_n^H]}{\sqrt{\mathbb{V}\mathrm{ar}[X_n^H]}}$$

converges in distribution to $\mathrm{N}(0,1)$.

We want to stress the fact that the constant μ_H can be expressed in terms of the Benjamini-Schramm limit (see [40]) but this does not apply for σ_H^2 since the Benjamini-Schramm limit implies only a law of large numbers.

The strategy we use on the proof is based on analytic combinatorics. More precisely, given a subgraph H we are able to get expressions for the counting formulas encoding the number of copies of H. As we will show, even if H has a very simple structure, we will need infinitely many equations and infinitely many variables to encode all the possible copies. We will be able to fully analyze the infinite system of equations that we obtain using an adapted version of the main theorem of Drmota, Gittenberger and Morgenbesser [61], which provides the necessary analytic ingredient in order to study infinite functional systems of equations. This result extends the classical Drmota-Lalley-Woods Theorem for (finite) systems of functional equations (see [60]).

Let us also discuss some similar results from the literature. The study of induced subgraphs (also called *patterns*) in random trees was done by Chyzak, Drmota, Klausner and Kok in [66], showing normal limiting distributions with linear expectation and variance. This covers in particular the distribution of the number of vertices of given degree in random trees. In the more general setting of subcritical graph classes, the number of vertices of degree k was studied in [12].

In another direction, the study of *pendant copies* of a given graph in random planar graphs was started by McDiarmid and Steger and Welsh in [44]. Let H being fixed rooted graph, with vertex set $\{1,\ldots,h\}$ and root r. We say that H *appears* in the graph G at $W \subset V(G)$ (or H is a *pendant copy* in G) if (i) there is an increasing bijection from $\{1,\ldots,h\}$ to W giving a graph isomorphism between H and the induced subgraph $G[W]$ of G, and (ii) there is exactly one edge in G between W and the rest of G, and this edge is incident with the root r. The number of pendant copies of a fixed subgraph H in a subcritical graph class was studied in [28]. In particular, it was shown that this number follows a Central Limit Theorem with linear expectation. As every pendant copy defines a subgraph, this result implies that the expected number of subgraphs in a subcritical graph of order n, chosen uniformly at random is at least linear. Our result strongly strengths this fact by showing the precise limiting probability distribution and the order of magnitude of the expectation.

As mentioned, Theorem 3.1.1 provides Central Limit Theorems with linear expectation and variance. However, our method does not give an effective method for computing the corresponding constants. In particular if H is not 2-connected then we need to operate with an infinite system of equations. Nevertheless if H is 2-connected and if we have a very precise characterization of the corresponding counting generating function of 2-connected graphs (for example given by a finite system of functional equations) then it is possible to compute the constants μ_H and σ_H^2 to arbitrary precision. As a case study, we provide explicit computations for specific (2-connected) subgraphs of series-parallel graphs. Recall that a graph is series-parallel if it excludes K_4 as a minor. Another equivalent definition states that a series-parallel graph is a graph whose 2-connected components are obtained from an edge by means of series and parallel operations. In this setting, we are able to show the following explicit result for triangles:

Theorem 3.1.2. *The number of copies of K_3, $X_n^{\blacktriangle}$, on series-parallel graph with n vertices chosen uniformly at random is asymptotically normal, with*

$$\mathbb{E}[X_n^{\blacktriangle}] = \mu_{\blacktriangle} n + O(1), \qquad \mathbb{V}\mathrm{ar}[X_n^{\blacktriangle}] = \sigma_{\blacktriangle}^2 n + O(1),$$

where $\mu_{\blacktriangle} \approx 0.39481$ *and* $\sigma^2_{\blacktriangle} \approx 0.41450$.

Additionally, our approach based on generating functions and analytic combinatorics let us also analyze the asymptotic number of triangle-free series-parallel graphs on n vertices. Finally, the more involved case of studying the number of copies of C_4 as well as the number of series-parallel graphs with a given girth is discussed as well.

Structure of the chapter: Section 3.2 is devoted to the introduction of the notation concerning generating functions and results needed later in the chapter. Section 3.3 covers the analytic preliminaries of the chapter. This section includes a modified version of the main theorem of Drmota, Gittenberger and Morgenbesser in [61], which is our main analytic ingredient on the proof of Theorem 3.1.1. Section 3.4 deals with the easier situation where the subgraph under study is 2-connected. The arguments to deal with the general connected case are developed in Section 3.5. In order to prepare the reader for the involved notation used to deal with general subgraphs, some simpler cases are fully developed. Section 3.6 is devoted to explicit computations in the family of series-parallel graphs. Section 3.7 discusses the results obtained so far and possible future investigations.

3.2 Preliminaries

All graphs we study are assumed to be simple (no loops nor multiple edges) and labelled. A graph on n vertices will be always labelled with different elements in $\{1, \ldots, n\}$. Finally, all convergences of random variables considered in this chapter are in distribution, as in the rest of this thesis.

3.2.1 Combinatorial classes. Exponential generating functions.

We follow the notation and definitions in [5]. A *labelled combinatorial class* is a set $\mathcal{A}$ together with a size function, such that for each $n \geq 0$ the set of elements of size n, denoted by $\mathcal{A}_n$, is finite. Each object a of $\mathcal{A}_n$ is made of n atoms (typically, vertices in graph classes) assembled in a specific way, the atoms bearing distinct labels in the set $\{1, \ldots, n\}$. We always assume that a combinatorial class is stable under graph isomorphism, namely, $a \in \mathcal{A}$ if and only if all graphs a' isomorphic to a are also elements of $\mathcal{A}$.

In counting problems it is convenient to use the exponential generating function (shortly the EGF) associated to the labelled class $\mathcal{A}$:

$$A(x) := \sum_{n \geq 0} |\mathcal{A}_n| \frac{x^n}{n!}.$$

In our setting, we use the (exponential) indeterminate x to encode vertices. In the opposite direction, we also write $[x^n]A(x) = |\mathcal{A}_n|/n!$.

The basic constructions we consider in this chapter are described in Table 3.1. Let us briefly explain each construction (see [5] for all the details). The *disjoint union* $\mathcal{A} \cup \mathcal{B}$ of two classes $\mathcal{A}$ and $\mathcal{B}$ refers to the disjoint union of the classes (and the corresponding induced size). The *labelled Cartesian product* $\mathcal{A} \times \mathcal{B}$ of two classes $\mathcal{A}$ and $\mathcal{B}$ is the set of pairs (a, b) where $a \in \mathcal{A}$ and $b \in \mathcal{B}$, joint with a redistribution of the labels of both a and b. The size of (a, b) is the sum of the sizes of a and b. The *sequence* of a set $\mathcal{A}$ (denoted by $\mathrm{Seq}(\mathcal{A})$) is $\{\varepsilon\} \cup \mathcal{A} \cup (\mathcal{A} \times \mathcal{A}) \cup \dots$ (ε denotes an element in the class of size 0). The *set* construction $\mathrm{Set}(\mathcal{A})$ is $\mathrm{Seq}(\mathcal{A})/\sim$, where $(a_1, a_2, \dots, a_r) \sim (\widehat{a}_1, \widehat{a}_2, \dots, \widehat{a}_r)$ when there exists a permutation of indices τ in $\{1, \dots, r\}$ such that equality $a_i = \widehat{a}_{\tau(i)}$ holds for all i. The *cycle* construction $\mathrm{Cyc}(\mathcal{A})$ is similar to the set construction, but now two elements are equivalent if and only if one can be obtained from the second by a cyclic permutation. The *restricted set* construction is equivalent to the set construction but here the Cartesian product has only a fixed number of terms. Finally, the *composition* of two combinatorial classes $\mathcal{A}$ and $\mathcal{B}$ (or or *substitution* of $\mathcal{B}$ in $\mathcal{A}$) is obtained by substituting each atom of each element of $\mathcal{A}$ by an element of $\mathcal{B}$.

We additionally consider classes of graphs of various types depending on whether one marks vertices or not. A (vertex-)*pointed graph* is a graph with a distinguished (labelled) vertex. A *derived graph* is a graph where one vertex is distinguished but not labelled (the other $n-1$ vertices have distinct labels in $\{1, \dots, n-1\}$). In particular, isomorphisms between two pointed graphs (or between two derived graphs) have to respect the distinguished vertex.

Given a graph class $\mathcal{A}$, the *pointed* class $\mathcal{A}^\bullet$ is the class of pointed graphs arising from $\mathcal{A}$. Similarly, the *derived* graph class $\mathcal{A}^\circ$ is obtained by taking all derived graphs built from $\mathcal{A}$. Hence, $|\mathcal{A}^\bullet_n| = n|\mathcal{A}_n|$ and $|\mathcal{A}^\circ_{n-1}| = |\mathcal{A}_n|$, and we have $A^\bullet(x) = x\frac{d}{dx}A(x)$ and $A^\circ(x) = \frac{d}{dx}A(x)$.

Construction	Class	Equations
Sum	$\mathcal{C} = \mathcal{A} \cup \mathcal{B}$	$C(x) = A(x) + B(x)$
Product	$\mathcal{C} = \mathcal{A} \times \mathcal{B}$	$C(x) = A(x) \cdot B(x)$
Sequence	$\mathcal{C} = \mathrm{Seq}(\mathcal{A})$	$C(x) = 1/(1 - A(x))$
Set	$\mathcal{C} = \mathrm{Set}(\mathcal{A})$	$C(x) = \exp(A(x))$
Restricted Set	$\mathcal{C} = \mathrm{Set}_{\geq k}(\mathcal{A})$	$C(x) = \exp_{\geq k}(A(x)) = \exp(A(x)) - \sum_{i=0}^{k-1} \frac{A(x)^i}{i!}$
Cycle	$\mathcal{C} = \mathrm{Cyc}(\mathcal{A})$	$C(x) = -\frac{1}{2}\log(1 - A(x)) - \frac{A(x)}{2} - \frac{A(x)^2}{4}$
Composition	$\mathcal{C} = \mathcal{A} \circ \mathcal{B}$	$C(x) = A(B(x))$
Pointing	$\mathcal{C} = \mathcal{A}^\bullet$	$C(x) = A^\bullet(x) = x\frac{d}{dx}A(x)$
Deriving	$\mathcal{C} = \mathcal{A}^\circ$	$C(x) = A^\circ(x) = \frac{d}{dx}A(x)$

Table 3.1: The Symbolic Method translating combinatorial constructions into operations on counting series.

Pointing and deriving operators will be only used over vertices. When dealing with *ordinary* parameters over combinatorial classes (for instance, edges or copies of a fixed subgraph) we use extra variables in the corresponding counting formulas. The partial derivatives of counting series with respect to parameters are denoted by subindices of the corresponding indeterminate. For instance, a generating function of the form $A_y(x, y)$ means $\frac{\partial}{\partial y} A(x, y)$.

3.2.2 Graph decompositions

Recall from Chapter 1 that a *block* of a graph g is a maximal 2-connected subgraph of g. A graph class $\mathcal{G}$ is *block-stable* if it contains the edge-graph e (the unique connected graph with two labelled vertices), and satisfies the property that a graph g belongs to $\mathcal{G}$ if and only if all the blocks of g belong to $\mathcal{G}$. Block-stable classes cover a wide variety of natural graph families, including graph classes specified by a finite list of forbidden minors that are all 2-connected. Planar graphs ($\mathrm{Ex}(K_5, K_{3,3})$) or series-parallel graphs ($\mathrm{Ex}(K_4)$) are block-stable.

For a graph class $\mathcal{G}$, we write $\mathcal{C}$ and $\mathcal{B}$ for the subfamilies of connected and 2-connected graphs in $\mathcal{G}$, respectively. In particular, the following combinatorial specifications hold:

$$\mathcal{G} = \mathrm{Set}(\mathcal{C}), \;\; \mathcal{C}^\bullet = \bullet \times \mathrm{Set}(\mathcal{B}^\circ \circ \mathcal{C}^\bullet).$$

The previous formulas can be interpreted as follows: first, each general graph in $\mathcal{G}$ is a set of elements in $\mathcal{C}$. Secondly, a pointed connected graph in $\mathcal{C}^\bullet$ can be decomposed as the root vertex, and a set of derived blocks (the ones incident with the root vertex) where we substitute on each vertex a pointed connected graph. See, for instance, [15, 47, 14] for full details. By means of Table 3.1 these expressions translate into equations of EGF in the following way:

$$G(x) = \exp(C(x)), \;\; C^\bullet(x) = x\exp(B^\circ(C^\bullet(x))).$$

See [17] for further results on graph decompositions and connectivity on graphs.

3.3 Analytic preliminaries

In this part we include the analytic results necessary in the forthcoming sections of the chapter.

3.3.1 Subcritical graphs

We start with the notion of subcritical graph class. Further details concerning these graph classes can be found in [12]. The next definition is based on the notation introduced in Section 3.2.2:

Definition 1. *A block-stable class of (vertex labelled) graphs is called* subcritical *if*

$$C^\bullet(\rho_C) < \rho_B,$$

where ρ_B denotes the radius of convergence of $B^\circ(x)$ and ρ_C the radius of convergence of $C^\bullet(x)$.

Roughly speaking, the subcritical condition means that the singular behaviour of $B^\circ(x)$ does not interfere with the singular behaviour of $C^\bullet(x)$. Only the behaviour of $B^\circ(x)$ for $|x| \le (1+\varepsilon)C^\bullet(\rho_C)$ matters (where $\varepsilon > 0$ is sufficiently small). This analytic behaviour translates into qualitative properties: as it is shown in [32], this condition assures that the largest block of a random graph in the class with n vertices has $O(\log(n))$ vertices. This behaviour differs from the one in random planar graphs, where the largest block follows an Airy-map law with linear expectation (see [28]).

Let us analyze the equations. From the subcritical condition and the arguments in [12, Section 4.1] (see also [5, Chapter VI.9.]) it follows that $y = C^\bullet(x)$ becomes singular for $x = x_0 = \rho_C$ if x_0 (and $y_0 = C^\bullet(x_0)$) satisfies the system of equations

$$y_0 = x_0 e^{B^\circ(y_0)},$$
$$1 = x_0 e^{B^\circ(y_0)} B^{\circ\prime}(y_0),$$

or equivalently if

$$1 = y_0 B''(y_0),$$
$$x_0 = y_0 e^{-B'(y_0)}.$$

In particular, we just have to assure that the equation $1 = yB''(y)$ has a solution $y < \rho_B$. Equivalently this is granted if

$$\rho_B B''(\rho_B) > 1.$$

It also follows from general theory (see [5, Chapter VI.9.]) that the solution function $C^\bullet(x)$ has a square-root type singularity at $x = \rho_C$ and can be (locally) written in the form

$$C^\bullet(x) = h_1(x) - h_2(x)\sqrt{1 - \frac{x}{\rho_C}},$$

where $h_1(x)$ and $h_2(x)$ are analytic functions at $x = \rho_C$ and satisfy the condition $h_1(\rho_C) = C^\bullet(\rho_C)$ and $h_2(\rho_C) > 0$.

It is convenient to assume that our graph class is an *aperiodic class*. That is, there is a positive integer n_0 such that we have $[x^n]C^\bullet(x) > 0$ for $n \ge n_0$. Then it follows that $x = \rho_C$ is the only singularity on the circle of convergence $|x| = \rho_C$. Additionally, there is an analytic continuation of $C^\bullet(x)$ to a domain of the form $\{x \in \mathbb{C} : |x| < \rho' \text{ and } \arg(x - \rho_C) \notin [-\theta, \theta]\}$ for some real number $\rho' > \rho_C$ and some positive angle $0 < \theta < \pi/2$. We call such a domain Δ*-region* or *domain dented* at ρ_C.

More precisely, if $|x| = \rho_C$ but $x \neq \rho_C$ then

$$|C^\bullet(x) B^{\circ\prime}(C^\bullet(x))| < 1.$$

Thus by the Implicit Function Theorem, $C^\bullet(x)$ has no singularity there and can be analytically continued. Consequently, we get by singularity analysis over $C^\bullet(x)$ that

$$[x^n]C^\bullet(x) = \frac{h_2(\rho_C)}{2\sqrt{\pi}} n^{-3/2} \rho_C^{-n} \left(1 + O\left(n^{-1}\right)\right).$$

Since $C^\bullet(x) = xC'(x)$ we also obtain the local singular behavior of $C(x)$ which is of the form

$$C(x) = h_3(x) + h_4(x)\left(1 - \frac{x}{\rho_C}\right)^{3/2},$$

for some functions $h_3(x)$ and $h_4(x)$ which are analytic at $x = \rho_C$. Since $G(x) = \exp(C(x))$ this also provides the local singular behavior of $G(x)$:

$$G(x) = h_5(x) + h_6(x)\left(1 - \frac{x}{\rho_C}\right)^{3/2},$$

where again $h_5(x)$ and $h_6(x)$ are analytic at $x = \rho_C$. This implies (applying again singularity analysis) that

$$[x^n]G(x) = \frac{3h_6(\rho_C)}{4\sqrt{\pi}} n^{-5/2}\rho_C^{-n}\left(1 + O\left(n^{-1}\right)\right).$$

In what follows we will heavily make use of these properties of subcritical graph classes.

3.3.2 A single equation

We first state a Central Limit Theorem that is a slight modification of [47, Theorem 2.23]. Let $F(x,y,u) = \sum_{n,m} F_{n,m}(u)x^n y^m$ be an analytic function in x, y around 0, and u is a complex parameter with $|u| = 1$. Suppose that the following conditions hold:

($F1$) $F(0,y,u) \equiv 0$.

($F2$) $F(x,0,u) \not\equiv 0$.

($F3$) All coefficients $F_{n,m}(1)$ of $F(x,y,1)$ are real and non-negative.

($F4$) For $|u| = 1$, then $|F_{n,m}(u)| \leq F_{n,m}(1)$.

($F5$) The function $t \mapsto F(x,y,e^{it})$ is at least three times continuously differentiable and all derivatives are analytic, too, in x and y.

($F6$) The region of convergence of $F(x,y,u)$ is large enough such that there exist non-negative solutions $x = x_0$ and $y = y_0$ of the system of equations

$$\begin{aligned} y &= F(x,y,1), \\ 1 &= F_y(x,y,1), \end{aligned} \tag{3.1}$$

with $F_x(x_0,y_0,1) \neq 0$ and $F_{yy}(x_0,y_0,1) \neq 0$.

Then, by the implicit function Theorem it is clear that the functional equation

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

has a unique analytic solution $y = y(x,u) = \sum_n y_n(u)x^n$ with $y(0,u) = 0$ that is three times continuously differentiable with respect to t if $u = e^{it}$. Furthermore the coefficients $y_n(1)$ are non-negative.

It is easy to show that there exists an integer $d \geq 1$ and a residue class r modulo d such that $y_n(1) > 0$ if $n \not\equiv r \pmod{d}$. In order to simplify the following presentation we assume that $d = 1$ (namely, we discuss the aperiodic case). The general case can be reduced to this case by a proper substitution in the original equation.

We have then the following theorem:

Theorem 3.3.1. *Assume that $F(x,y,u)$ satisfies assumptions $(F1)-(F6)$ and $y(x,u)$ is a power series in x that is the (analytic) solution of the functional equation $y = F(x,y,u)$. Suppose that X_n is a sequence of random variables such that*

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

where $|u| = 1$. Set

$$\begin{aligned}\mu &= \frac{F_u}{x_0 F_x},\\ \sigma^2 &= \frac{1}{x_0 F_x^3 F_{yy}}\Big(F_x^2(F_{yy}F_{uu} - F_{yu}^2) - 2F_xF_u(F_{yy}F_{xu} - F_{yx}F_{yu}) + F_u^2(F_{yy}F_{xx} - F_{yx}^2)\Big) +\\ &\quad \mu + \mu^2,\end{aligned}$$

where all partial derivatives are evaluated at the point $(x_0, y_0, 1)$ which is a solution to the system of equations (3.1). *Then we have that*

$$\mathbb{E}[X_n] = \mu n + O(1),\ \mathbb{V}\mathrm{ar}[X_n] = \sigma^2 n + O(1)$$

and if $\sigma^2 > 0$ then

$$\frac{X_n - \mathbb{E}[X_n]}{\sqrt{\mathbb{V}\mathrm{ar}[X_n]}} \to N(0,1).$$

Proof. The proof runs along the same lines as that of [47, Theorem 2.23]. We just indicate the differences.

By the Implicit Function Theorem it follows that there exist functions $f(u)$ and $g(u)$ (for $|u-1| < \varepsilon$ and $|u| = 1$ for some $\varepsilon > 0$) which are three times differentiable with respect to t if $u = e^{it}$ that satisfy

$$\begin{aligned}g(u) &= F(f(u), g(u), u),\\ 1 &= F_y(f(u), g(u), u)\end{aligned}$$

with $f(1) = x_0$ and $g(1) = y_0$. Furthermore, by applying a proper variant of the Weierstrass Representation Theorem it follows (as in the proof of [47, Theorem 2.23]) that we have a presentation of the form

$$y(x,u) = h_1(x,u) - h_2(x,u)\sqrt{1 - \frac{x}{f(u)}} \tag{3.2}$$

locally around $x = x_0$, $u = 1$, where $h_1(x,u)$, and $h_2(x,u)$ are analytic in x and three times continuously differentiable with respect to t if $u = e^{it}$, where $h_1(f(u),u) = g(u)$ and

$$h_2(f(u),u) = \sqrt{\frac{2f(u)F_x(f(u),g(u),u)}{F_{yy}(f(u),g(u),u)}} \neq 0.$$

Since $d = 1$ we also get

$$y_n(u) = [x^n]y(x,u) = \sqrt{\frac{f(u)F_x(f(u),g(u),u)}{2\pi F_{yy}(f(u),g(u),u)}} f(u)^{-n} n^{-3/2} \left(1 + O(n^{-1})\right) \tag{3.3}$$

uniformly for $|u-1| < \varepsilon$ and $|u| = 1$. Hence,

$$\mathbb{E}\left[u^{X_n}\right] = \frac{[x^n]y(x,u)}{[x^n]y(x,1)} = \frac{h_2(f(u),u)}{h_2(f(1),1)} \left(\frac{f(1)}{f(u)}\right)^n \left(1 + O\left(n^{-1}\right)\right). \tag{3.4}$$

By using the local expansion of $f(u)$ we get for $u = e^{it}$

$$\frac{f(1)}{f(u)} = e^{it\mu - \sigma^2 t^2/2 + O(t^3)},$$

which directly implies

$$\lim_{n\to\infty} \mathbb{E}\left[e^{it(X_n - \mu n)/(\sigma\sqrt{n})}\right] = e^{-t^2/2}.$$

By Levi's Theorem this proves the Central Limit Theorem. □

Remark 3.3.2. *In our applications, the function $y(x,u)$ will be the generating function $C^\bullet(x,u) = x\frac{\partial}{\partial x}C(x,u)$ of connected graphs. Since $[x^n]C^\bullet(x,u) = n[x^n]C(x,u)$ it follows that*

$$\frac{[x^n]\,C(x,u)}{[x^n]\,C(x,1)} = \frac{[x^n]\,C^\bullet(x,u)}{[x^n]\,C^\bullet(x,1)}$$

and, thus, it is sufficient to work with $C^\bullet(x,u)$ instead of $C(x,u)$. However, if we are interested in all graphs (not necessarily connected) we need to study the behaviour of $G(x,y)$. By means of the set construction $G(x,u) = \exp(C(x,u))$ we have to replace $y(x,u) = C^\bullet(x,u)$ by the function

$$G(x,u) = \tilde{y}(x,u) = \exp\left(\int_0^x \frac{y(\xi,u)}{\xi}\,d\xi\right)$$

and the new random variable $\tilde{X}_n$ that is defined by $\mathbb{E}\left[u^{\tilde{X}_n}\right] = \frac{[x^n]\tilde{y}(x,u)}{[x^n]\tilde{y}(x,1)}$. Indeed, $\tilde{y}(x,u)$ has a slightly different singular behaviour: from (3.2) we obtain

$$\int_0^x \frac{y(\xi,u)}{\xi}\,d\xi = h_3(x,u) + h_4(x,u)\left(1 - \frac{x}{f(u)}\right)^{3/2}$$

and consequently

$$\tilde{y}(x,u) = h_5(x,u) + h_6(x,u)\left(1 - \frac{x}{f(u)}\right)^{3/2}$$

for proper function $h_3(x,u)$, $h_4(x,u)$, $h_5(x,u)$, $h_6(x,u)$. However, from that expression we obtain the same kind of asymptotic behavior as in (3.4) *and a Central Limit Theorem for $\tilde{X}_n$ with the same asymptotic behaviour for mean and variance as for X_n.*

Remark 3.3.3. *In most of the applications, the condition $\sigma^2 > 0$ is satisfied. As it is shown in [12, Lemma 4], if $y = F(x,y,u) = \sum_{n,m,k} a_{n,m,k} x^n y^m u^k$ satisfies some natural analytic conditions (see [12]), and assuming that there are three integer vectors (n_j, m_j, k_j), $j = 1,2,3$ with $m_j > 0$, $j = 1,2,3$ with*

$$\begin{vmatrix} n_1 & m_1 - 1 & k_1 \\ n_2 & m_2 - 1 & k_2 \\ n_3 & m_3 - 1 & k_3 \end{vmatrix} \neq 0$$

and $a_{n_j,m_j,k_j} \neq 0$ for $j = 1,2,3$, then $\sigma^2 > 0$.

Remark 3.3.4. *Finally we remark that Theorem 3.3.1 extends to a finite system of equations $y_j = F_j(x, y_1, \ldots, y_K, u)$, $1 \le j \le K$, provided that the system is strongly connected (compare with [47, Theorem 2.35]). We will use this extension in Section 3.6.2.*

3.3.3 An infinite system of equations

The main reference for this subsection is the work of Drmota, Gittenberger and Morgenbesser [61]. We start again with an equation of the form $y = F(x,y)$, where F satisfies (almost) the same assumptions as that of Theorem 3.3.1.

($I1$) Let $F(x,y) = \sum_{n,m} F_{n,m} x^n y^m$ be an analytic function in x, y around 0 such that the conditions ($F1$), ($F2$), ($F3$), and ($F6$) are satisfied (we just omit the properties concerning u).

In particular this means that the solution $y = y(x)$ has a square-root type singularity at x_0 and the coefficients $y_n = [x^n]y(x)$ have an asymptotic expansion of the form given by (3.3), where $u = 1$. Furthermore we suppose that the following conditions hold:

($I2$) There exist functions $y_j(x,u)$, $j = 1,2,\ldots$, that are power series in x and are defined for $|u| = 1$ such that

$$y(x) = \sum_{j\ge 1} y_j(x,1).$$

($I3$) The infinite system of functions $\mathbf{y} = (y_j(x,u))_{j\ge1} = (y_j)_{j\ge1}$ satisfy an (infinite) system of equations of the form

$$y_j = F_j(x, \mathbf{y}, u), \qquad j \ge 1,$$

where F_j has a power series expansion

$$F_j(x, \mathbf{y}, u) = \sum_{n, m_1, m_2, \ldots} F_{j;n,m_1,m_2,\ldots}(u) x^n y_1^{m_1} y_2^{m_2} \cdots$$

with coefficients that satisfy $|F_{j;n,m_1,m_2,\ldots}(u)| \le F_{j;n,m_1,m_2,\ldots}(1)$. In particular, these coefficients are non-negative for $u = 1$.

$(I4)$ For every $j \ge 1$ there exists a function $\tilde{F}_j(x, y)$ with

$$F_j(x, \mathbf{y}, 1) = \tilde{F}_j(x, y_1 + y_2 + \cdots) \tag{3.5}$$

and

$$\sum_{j \ge 1} \tilde{F}_j(x, y) = F(x, y). \tag{3.6}$$

$(I5)$ For each $j \ge 1$, F_j is three times continuously differentiable with respect to t with $u = e^{it}$ such that the series

$$\sum_{j \ge 1} j^m \frac{\partial^r}{\partial t^r} F_j(x, \mathbf{y}, e^{it}), \qquad r \in \{0, 1, 2, 3\}, \tag{3.7}$$

converges absolutely for $|x| < x_0 + \varepsilon$ and $\|\mathbf{y}\|_1 < y_0 + \varepsilon$ (for some $\varepsilon > 0$).

From properties $(I2)$–$(I4)$ it immediately follows that F_j is well defined (and also analytic) for x and $\mathbf{y} = (y_j)_{j \ge 1}$ for which $F(|x|, \|\mathbf{y}\|_1)$ is analytic (recall that $\|\mathbf{y}\|_1 = \sum_{j \ge 1} |y_j|$). Consequently, under the same conditions, $\tilde{F}_j(|x|, \|\mathbf{y}\|_1)$ is convergent. Actually we only need convergence for $|x| < x_0 + \varepsilon$ and $\|\mathbf{y}\|_1 < y_0 + \varepsilon$ for some $\varepsilon > 0$. This suggests to work in the space $\ell^1(\mathbb{C})$ for $\mathbf{y} = (y_j)_{j \ge 1}$. However, in the present situation we have to be slightly more careful since we have to take also into account derivatives with respect to t (with $u = e^{it}$). For this purpose we use weighted ℓ^1 spaces of the form

$$\ell^1(m, \mathbb{C}) = \Big\{ \mathbf{y} = (y_j)_{j \ge 1} \in \mathbb{C}^{\mathbb{N}} : \|\mathbf{y}\|_{m,1} := \sum_{j \ge 1} j^m |y_j| < \infty \Big\},$$

for some non-negative real number m (see also Remark 3.3.6). Since $\|\mathbf{y}\|_1 \le \|\mathbf{y}\|_{m,1}$ the functions F_j are also well defined (and analytic) if $|x| < x_0 + \varepsilon$ and $\|\mathbf{y}\|_{m,1} < y_0 + \varepsilon$ for some $\varepsilon > 0$.

Note that the case $r = 0$ in condition $(I5)$ just says that for each j the mapping $(x, \mathbf{y}) \mapsto F_j(x, \mathbf{y}, u)$ is well defined in the space $\mathbb{C} \times \ell^1(m, \mathbb{C})$ with $|x| < x_0 + \varepsilon$ and $\|\mathbf{y}\|_{m,1} < y_0 + \varepsilon$ (for some $\varepsilon > 0$).

Finally we want to mention that $(I4)$ informally means that the infinite system can be interpreted as a partition of the main equation $y = F(x, y)$. Hence, we refer to Equation (3.6) as the *Partition Property*).

The main theorem in this context is the following:

Theorem 3.3.5. *Assume that $F(x, y)$ and $F_j(x, y_1, y_2, \ldots, u)$, $j \geq 1$, satisfy the conditions (I1)–(I5) so that the functions $y_j(x, u)$, $h \geq 1$, satisfy the infinite system of equations $y_j(x, u) = F_j(x, y_1(x, u), y_2(x, u), \ldots, u)$, $j \geq 1$. Furthermore set $y(x, u) = \sum_{j \geq 1} y_j(x, u)$ and suppose that X_n is a sequence of random variables with*

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

for $|u| = 1$. Then we have

$$\mathbb{E}[X_n] = \mu n + O(1) \quad \text{and} \quad \mathbb{V}\mathrm{ar}[X_n] = \sigma^2 n + O(1)$$

for some real constants $\mu > 0$ and $\sigma^2 \geq 0$. Furthermore if $\sigma^2 > 0$ then

$$\frac{X_n - \mathbb{E}[X_n]}{\sqrt{\mathbb{V}\mathrm{ar}[X_n]}} \to N(0, 1).$$

Remark 3.3.6. *We note that a corresponding theorem for a finite system is also true ([60, 47]) but in our context we just need the infinite version.*

Furthermore, Theorem 3.3.5 even holds in slightly more general situations. For example, if the functions $y_j(x, u)$ are not indexed by an integer $j \geq 1$ but by a multi-index $\mathbf{j} = (j_1, \ldots, j_d)$ of integers $j_i \geq 1$ then we can also adapt the space $\ell^1(m, \mathbb{C})$ to the space

$$\Big\{ \mathbf{y} = (y_{\mathbf{j}})_{\mathbf{j} \geq \mathbf{1}} \in \mathbb{C}^{\mathbb{N}^d} : \|\mathbf{y}\|_{m,1} := \sum_{\mathbf{j} \geq \mathbf{1}} \|\mathbf{j}\|_1^m |y_{\mathbf{j}}| < \infty \Big\}.$$

Actually we will need this generalization if we consider subgraphs H with more than one cut-vertex.

Finally, as for Theorem 3.3.1 the Central Limit Theorem transfers to $\tilde{X}_n$ that is defined with the help of $\tilde{y}(x, u) = \exp\left(\int_0^x y(\xi, u)/\xi \, d\xi\right)$. Compare this fact with Remark 3.3.2.

Proof. We first note that Theorem 3.3.5 will be deduced from [61, Theorem 1] with a slight adaption corresponding to u – here we just require differentiability with respect to t if $u = e^{it}$ and not analyticity – and corresponding to the underlying space – we replace $\ell^1(\mathbb{C})$ by $\ell^1(m, \mathbb{C})$. Actually the modification corresponding to u can be treated as in the proof of Theorem 3.3.1 and the change of the underlying space does not change the proof at all, so we will not discuss these issues.

Next we note that Equation (3.5) implies

$$y_j(x, 1) = \tilde{F}_j(x, y(x)),$$

where $y = y(x)$ is the solution of the equation $y = F(x, y)$. Thus, we study two cases. First, if $\tilde{F}_j$ does not depend on y then $y_j(x, 1)$ is analytic at $x = x_0$. This also implies that $y_j(x, u)$ is analytic for $|u| = 1$ and for $|x| < 1 + \varepsilon$ for some $\varepsilon > 0$. Let I_1 denote the set of indices j with this property. Furthermore, since $F(x, y)$ is also analytic at $x = x_0$ it also follows that $\sum_{j \in I_1} y_j(x, u)$ is analytic in x for $|x| < x_0 + \varepsilon$ and for $|u| = 1$.

In the second case $y_j(x,1)$ has a square-root singularity of the form

$$y_j(x) = h_{1,j}(x) - h_{2,j}(x)\sqrt{1 - x/x_0},$$

which is inherited from that of $y(x)$. Furthermore it follows that $F_j(x, y_1, y_2, \ldots, u)$ depends on all variables y_i, $i \ge 1$. Let I_2 denote the set of indices j of the second case.

If we reduce now the infinite system to those equations with $j \in I_2$, where we consider $y_j(x,u)$ with $j \in I_1$ as already known functions, then we get a strongly connected system of equations

$$y_j = F_j(x, (y_i(x,u))_{i\in I_1}, (y_\ell)_{\ell \in I_2}, u), \qquad j \in I_2$$

that satisfies all regularity assumptions of [61, Theorem 1]. In particular, since

$$|F_j(x, y_1, y_2, \ldots, u)| \le F_j(|x|, |y_1|, |y_2|, \ldots, 1) = \tilde{F}_j(|x|, |y_1| + |y_2| + \cdots)$$

and $\tilde{F}_j(x,y)$ is analytic (at least) in the region where $F(x,y)$ is analytic, it follows that the function $F_j(x, y_1, y_2, \ldots, u)$ is well defined (and analytic in x and $y_1, y_2, \ldots$) for x in a proper neighborhood of 0, $\mathbf{y} = (y_j)_{j\ge 1}$ in a proper neighborhood of 0 in $\ell^1(m, \mathbb{C})$ and u with $|u| = 1$.

The only remaining assumption that has to be checked is that the operator

$$J = \left(\frac{\partial F_j}{\partial y_i}(x, \mathbf{y}, 1) \right)_{i,j \in I_2}$$

is compact. Since the property

$$F_j(x, y_1, y_2, \ldots, 1) = \tilde{F}_j(x, y_1 + y_2 + \cdots)$$

is satisfied, it follows that

$$\frac{\partial F_j}{\partial y_i}(x, \mathbf{y}, 1) = \frac{\partial \tilde{F}_j}{\partial y}(x, y_1 + y_2 + \cdots)$$

is independent of the choice of i. Hence the rank of J equals 1 which implies that J is a compact operator.

Thus we can apply [61, Theorem 1] and obtain that all functions $y_j(x,u)$, $j \in I_2$, have a common square-root type singularity, and an expression of the form

$$y_j(x,u) = h_{1,j}(x,u) - h_{2,j}(x,u)\sqrt{1 - \frac{x}{f(u)}}.$$

with functions $f(u), h_{1,j}(x,u), h_{2,j}(x,u)$ that are three times differentiable in t, where $u = e^{it}$ and analytic in x around x_0.

Summing up we, thus, obtain a square-root singularity for $y(x,u)$. So we are precisely in the same situation as in the proof of Theorem 3.3.1. And so the result follows. □

3.4 2-Connected Subgraphs

The purpose of this section is to consider 2-connected subgraphs H. This case is much easier than the general case since a 2-connected subgraph can only appear in a block. Due to its shortness, we include the proof for this specific subgraph case.

Theorem 3.4.1. *Suppose that H is a 2-connected graph that appears as a subgraph in an aperiodic subcritical graph class $\mathcal{G}$. Let $X_n = X_n^{(H)}$ denote the number of copies of H in a connected or general graph in $\mathcal{G}$ of order n, chosen uniformly at random.*

Then, X_n satisfies a Central Limit Theorem with $\mathbb{E}[X_n] \sim \mu n$ and $\mathbb{V}\mathrm{ar}[X_n] \sim \sigma^2 n$ as $n \to \infty$.

Proof. Let $b^\circ_{n,k}$ be the number of rooted 2-connected graphs in $\mathcal{G}$ with $n-1$ non-root vertices such that H appears precisely k times as a subgraph. Furthermore let

$$B^\circ(x,u) = \sum_{n,k} b^\circ_{n,k} \frac{x^n}{n!} u^k$$

be the corresponding generating function.

Let $C^\bullet(x,u)$ be the corresponding generating function of connected graphs in $\mathcal{G}$ (where the root is non discounted). Since H is assumed to be 2-connected the number of copies of H in a connected graph is just the sum of its copies in the blocks. Hence we have

$$C^\bullet(x,u) = xe^{B^\circ(C^\bullet(x,u),u)}.$$

If $u = 1$ then $B^\circ(x,1)$ and $C^\bullet(x,1)$ are the usual counting functions that satisfy the equation $C^\bullet(x,1) = xe^{B^\circ(C^\bullet(x,1),1)}$.

In order to prove Theorem 3.4.1 we just have to check the conditions of Theorem 3.3.1. It is straightforward to check that conditions (F1)-(F4) are satisfied. By the subcritical condition we certainly have $x_0 = \rho_C$ and $y_0 = C^\bullet(\rho_C, 1)$ that satisfy

$$\begin{aligned} y_0 &= x_0 e^{B^\circ(y_0,1)}, \\ 1 &= x_0 e^{B^\circ(y_0,1)} B^{\circ\prime}(y_0,1). \end{aligned}$$

Furthermore, since $C^\bullet(\rho_C,1) < \rho_B$ the region of convergence of $F(x,y,u) = xe^{B^\circ(y,u)}$ is large enough and condition $(F6)$ in Theorem 3.3.1 is also satisfied.

The only missing assumption that has to be (finally) checked is that the mapping $t \mapsto xe^{B^\circ(y,e^{it})}$ is three times continuously differentiable in t. Of course it is sufficient to study the mapping $t \mapsto B^\circ(y,e^{it})$. First we note that $|B^\circ(y,u)| \le B^\circ(|y|,1)$. From this it follows that $B^\circ(y,u)$ exists (and is also analytic in y) for all $|y| < \rho_B$ and for $|u| = 1$. Next we note that the number of occurrences of a graph H of order L in a graph with n vertices is bounded by $O(n^L)$. Write b°_n for the number of rooted 2-connected graphs in $\mathcal{G}$ with $n-1$ non-root vertices. Thus it follows that

$$\left| \frac{\partial^r}{\partial u^r} \sum_k b^\circ_{n,k} u^k \right| \le \sum_k k^r b^\circ_{n,k} = O(n^{rL} b^\circ_n)$$

for u with $|u| = 1$; for notational convenience we have taken the derivatives formally with respect to u. However, since all all derivatives $\frac{\partial^m}{\partial y^m} B^\circ(y,1)$ are finite it follows that all derivatives $\frac{\partial^r}{\partial u^r} B^\circ(y,u)$ exist for $|u| = 1$. (Alternatively we can use the bound $n^{rL} = O((1+\varepsilon)^n)$ for every $\varepsilon > 0$ which implies that

$$\left| \frac{\partial^r}{\partial u^r} B^\circ(y,u) \right| = O(B^\circ(|y|(1+\varepsilon),1)).$$

Consequently all assumptions of Theorem 3.3.1 are satisfied and the result follows for the connected case. In the general case, where we have to work with $G(x,u) = \exp\left(\int_0^x C^\bullet(\xi,u)/\xi \, d\xi\right)$, we get the same result, see Remark 3.3.2. $\square$

3.5 Connected Subgraphs

The purpose of this section is to extend Theorem 3.4.1 to subgraphs H that are not 2-connected, and hence prove Theorem 3.1.1. The main difference between the 2-connected case and the (general) connected case is that occurrences of H are not necessarily separated by cut-vertices. This means that we have to cut H also into pieces (more precisely, into its blocks) and to count all combinations of these pieces when two (or several) blocks are joined by a cut-vertex, or several cut-vertices.

We start this section by illustrating the arguments with the base case $H = P_2$, which is the simplest case of a graph H that is not 2-connected. Later, as a warm-up for the general case (where notation could be especially involved), we show the combinatorics behind two particular cases: copies of subgraphs with one cut-vertex and exactly three blocks (Section 3.5.2) and the number of copies of P_3 (Section 3.5.3). In both cases we show again the type of functional equations we obtain in this setting and the main difficulties that arise when encoding the counting formulas. At the end of the section we indicate how the method can be modified to cover the general case, both combinatorially and analytically.

3.5.1 Counting copies of P_2

We start the study of connected subgraphs dealing with the simplest case study. This subsection will show some of the ideas that will be used later in the proof of Theorem 3.1.1, and can be considered as a warm-up for the forthcoming subsections.

Let H be a path of length 2, namely P_2. In this situation, H separates by a cut-vertex into two edges. In particular, observe that if we join two blocks at a cut-vertex and the two corresponding degrees of these two blocks at the cut-vertex are k_1 and k_2 then there we create $k_1 k_2$ occurrences of H. Hence, copies of P_2 may arise from the following sources:

(A) Copies where the vertex in P_2 of degree 2 is the root vertex, but not contained in one of the blocks of the graph. In other words, each edge of the copy of P_2 under study is contained in a different block incident with the root vertex.

(B) Copies that are not incident with the root vertex, and built by taking and edge in one of the blocks incident with the root vertex and completing it to a copy of P_2 by means of one of the attached rooted connected graphs.

(C) Copies already existing in the blocks incident with the root vertex.

The previous observation suggests that we distinguish between infinitely many situations (depending on the root degrees of the blocks) which leads to an infinite system of equations. Let us start by introducing corresponding generating functions for 2-connected graphs. We denote by

$$B_j^\circ(w_1, w_2, w_3, \ldots; u), \qquad j \geq 1,$$

the generating function of blocks in $\mathcal{G}$, where the root has degree j, where w_i counts the number of non-root vertices of degree i, and where u counts the number of occurrences of $H = P_2$. Formally this is a generating function in infinitely many variables. Of course we have

$$B_j^\circ(x, x, \ldots; u) = B_j^\circ(x, u),$$

where x counts the number of non-root vertices. Consequently if $B_j^\circ(x, u)$ is convergent for some positive x and for u with $|u| = 1$ then $B_j^\circ(w_2, w_3, \ldots; u)$ converges for all w_i with $|w_i| < x$ and for all u with $|u| = 1$. Next let

$$C_j^\bullet(x, u), \qquad j \geq 0,$$

denote the generating function of connected rooted graphs in $\mathcal{G}$, where the root vertex has degree j, where x counts the number of (all) vertices and u the number of occurrences of $H = P_2$. Then by the same principle as above we have $C_0^\bullet(x) = x$ and for $j \geq 1$

$$C_j^\bullet(x, u) = x \sum_{s \geq 0} \frac{1}{s!} \sum_{j_1 + \cdots + j_s = j} u^{\sum_{i_1 < i_2} j_{i_1} j_{i_2}} \prod_{i=1}^{s} B_{j_i}^\circ \left(\sum_{\ell_1 \geq 0} u^{\ell_1} C_{\ell_1}^\bullet(x, u), \ \sum_{\ell_2 \geq 0} u^{2\ell_2} C_{\ell_2}^\bullet(x, u), \ \ldots \ ; \ u \right).$$

Note that s is the number of blocks containing the root vertex, and the index j_r (for $r = 1, \ldots, s$) is the degree of the root of the block r. In particular, the term $u^{\sum_{i_1 < i_2} j_{i_1} j_{i_2}}$, the terms $\sum_{\ell_k \geq 0} u^{\ell_k} C_{\ell_k}^\bullet(x, u)$ and the last argument in $B_{j_i}^\circ(w_1, w_2, \ldots; u)$ are the contributions to Case (A), (B) and (C), respectively.

It is convenient to replace all occurrences of $C_0^\bullet(x)$ by x. Thus we can view the infinite dimensional vector $\mathbf{y} = \mathbf{y}(x, u) = (C_j^\bullet(x, u))_{j \geq 1}$ as the solution of an infinite dimensional system of the form $y_j = F_j(x, \mathbf{y}, u)$, where F_j is defined by

$$F_j(x, \mathbf{y}, u) = x \sum_{s \geq 0} \frac{1}{s!} \sum_{j_1 + \cdots + j_s = j} u^{\sum_{i_i < i_2} j_{i_1} j_{i_2}} \prod_{i=1}^{s} B_{j_i}^\circ \left(x + \sum_{\ell_1 \geq 1} u^{\ell_1} y_{\ell_1}, \ x + \sum_{\ell_2 \geq 1} u^{2\ell_2} y_{\ell_2}, \ \ldots \ ; \ u \right). \tag{3.8}$$

We now show that this system of equations satisfies all assumptions of Theorem 3.3.5. First of all, let us check the Partition Property (3.6). If $u = 1$ the function $F_j(x, \mathbf{y}, 1)$

can be written as a function $\tilde{F}_j(x, y_1 + y_2 + \cdots)$, where

$$\tilde{F}_j(x,y) = x \sum_{s\geq 0} \frac{1}{s!} \sum_{j_1+\cdots+j_s=j} \prod_{i=1}^{s} B^{\circ}_{j_i}(x+y,1)\,.$$

In particular, $F(x,y) = \sum_j \tilde{F}_j(x,y)$ is equal to $x\exp(B^{\circ}(x+y))$, which is analytic in x. Since $|\tilde{F}_j(x,y)| \leq \tilde{F}_j(|x|,|y|)$ it is sufficient to study $\tilde{F}_j$ for positive x and y. By Equation (3.8) it follows that for all $n < j$ we have $[x^n]\tilde{F}_j(x,y) = 0$. Consequently we have (for positive x and y)

$$j^m \tilde{F}_j(x,y) \leq \left(x\frac{\partial}{\partial x}\right)^m \tilde{F}_j(x,y).$$

By analyticity of $\sum_j \tilde{F}_j(x,y)$, it directly follows then that

$$\sum_{j\geq 1} j^m \tilde{F}_j(x,y) \leq \left(x\frac{\partial}{\partial x}\right)^m F(x,y).$$

Thus, the infinite system is well defined (and analytic) on $\ell^1(m,\mathbb{C})$ for every $m \geq 0$.

It remains to check Condition (I5) (Equation (3.7)) for $r \in \{1,2,3\}$. For the sake of brevity we only work out the details of the case $r = 1$. The remaining cases are more involved but can be handled similarly. We first note that u appears in F_j at three different places:

(i) As the power $u^{\sum_{i_i<i_2} j_{i_1}j_{i_2}}$ (contribution of Case (A)),

(ii) In sums of the form $\sum_{\ell\geq 1} u^{m\ell}y_\ell$ as an argument of one of the terms $B^{\circ}_{j_i}$ (contribution of Case (B)) and

(iii) As the last argument in one of the terms $B^{\circ}_{j_i}(w_1, w_2, \ldots; u)$ (contribution of Case (C)).

As above it is sufficient to consider positive x and $y = y_1 + y_2 + \cdots$ in order to assure absolute convergence.

If we substitute $u = e^{it}$ and take the derivative with respect to t it follows that in Case (i) the derivative gives a factor of the form

$$i\sum_{i_i<i_2} j_{i_1}j_{i_2} e^{it\sum_{i_i<i_2} j_{i_1}j_{i_2}}$$

which can be absolutely bounded by

$$\sum_{i_i<i_2} j_{i_1}j_{i_2} \leq s(s-1)j^2.$$

Thus we are led to consider the sum (which is an upper bound)

$$\sum_{j\geq 1} j^m x \sum_{s\geq 0} \frac{1}{s!} s(s-1) j^2 \sum_{j_1+\cdots+j_s=j} \prod_{i=1}^{s} B_{j_i}^{\circ}(x+y,1)$$
$$\leq \sum_{j\geq 1} j^{m+2} x \sum_{s\geq 2} \frac{1}{(s-2)!} \sum_{j_1+\cdots+j_s=j} \prod_{i=1}^{s} B_{j_i}^{\circ}(x+y,1)$$
$$\leq \left(x\frac{\partial}{\partial x}\right)^{m+2} \sum_{j\geq 1} x \sum_{s\geq 2} \frac{1}{(s-2)!} \sum_{j_1+\cdots+j_s=j} \prod_{i=1}^{s} B_{j_i}^{\circ}(x+y,1)$$
$$= \left(x\frac{\partial}{\partial x}\right)^{m+2} xB^{\circ}(x+y)^2 \exp\left(B^{\circ}(x+y)\right)$$

which is certainly bounded (for positive x and y). Now we study Case (ii). If we take derivatives we get

$$\frac{\partial}{\partial t} B_j^{\circ}\left(x + \sum_{\ell_1\geq 1} e^{it\ell_1} y_{\ell_1},\ x + \sum_{\ell_2\geq 1} e^{2it\ell_2} y_{\ell_2},\ \ldots\ ;\ u\right)$$
$$= \sum_{m\geq 1} \frac{\partial}{\partial w_m} B_j^{\circ}\left(x + \sum_{\ell_1\geq 1} e^{it\ell_1} y_{\ell_1},\ x + \sum_{\ell_2\geq 1} e^{2it\ell_2} y_{\ell_2},\ \ldots\ ;\ u\right) \sum_{\ell_m\geq 1} im\ell_m e^{mit\ell_m} y_{\ell_m}$$

which can be bounded from the above by

$$\left(\sum_{\ell\geq 1} \ell |y_\ell|\right) \sum_{m\geq 1} m \frac{\partial B_j^{\circ}}{\partial w_m}(x+y,1)$$

Note that the sum $\sum_{m\geq 1} m \frac{\partial B_j^{\circ}}{\partial w_m}$ corresponds to the sum of the degrees of the non-root vertices. Since this sum is bounded by twice the number of edges it is bounded by $n(n-1)$, where n denotes the number of vertices. This leads us to the upper bound

$$\left(\sum_{\ell\geq 1} \ell |y_\ell|\right) (x+y)^2 \frac{\partial^2 B_j^{\circ}}{\partial x^2}(x+y,1).$$

This upper bound also implies the upper bound (recall that the derivative here is only restricted to Case (ii)):

$$\left|\frac{\partial}{\partial t} \prod_{i=1}^{s} B_{j_i}^{\circ}\right| \leq \left(\sum_{\ell\geq 1} \ell |y_\ell|\right) (x+y)^2 \frac{\partial^2}{\partial x^2} \prod_{i=1}^{s} B_{j_i}^{\circ}(x+y,1).$$

Finally, summing up over j (with the weight j^m) we obtain the upper bound (for positive

x and y)

$$\left(\sum_{\ell\geq 1}\ell|y_\ell|\right)\sum_{j\geq 1}j^m x\sum_{s\geq 0}\frac{1}{s!}\sum_{j_1+\cdots+j_s=j}(x+y)\frac{\partial^2}{\partial x^2}\prod_{i=1}^{s}B^\circ_{j_i}(x+y,1)$$
$$\leq\left(\sum_{\ell\geq 1}\ell|y_\ell|\right)\left(x\frac{\partial}{\partial x}\right)^{m+2}(x+y)^3\exp\left(B^\circ(x+y)\right).$$

Recall that we are working in the space $\ell_1(m,\mathbb{C})$ with $m\geq 1$. Thus the series $\sum_\ell \ell|y_\ell|$ is finite by definition Hence, the whole term is bounded. Finally in Case (iii) we can argue in the same way as in the proof of Theorem 3.4.1 and obtain

$$\left|\frac{\partial}{\partial t}B^\circ_j(x+y,e^{it})\right|\leq j^3 B^\circ_j(x+y,1)$$

(in the case of $H=P_2$ we have $L=3$). This leads us to consider the sum

$$\sum_{j\geq 1}j^m x\sum_{s\geq 0}\frac{1}{s!}\sum_{j_1+\cdots+j_s=j}\left(\sum_{i=1}^{s}j_i^3\right)\prod_{i=1}^{s}B^\circ_{j_i}(x+y,1)$$

which can be bounded (similarly to Case (i)) by

$$\left(x\frac{\partial}{\partial x}\right)^{m+3}xB^\circ(x+y)^3\exp\left(B^\circ(x+y)\right).$$

By putting the Case (i)–(iii) together it follows that Condition (3.7) is satisfied for $r=1$. As mentioned above the cases $r=2$ and $r=3$ can be similarly handled. This completes the proof of the Central Limit Theorem in the case of $H=P_2$ for connected graphs.

3.5.2 Main Example 1: Connected graphs with one cut-vertex

In the following two subsections we will illustrate the methods used in the general situation for a simpler subgraph. We start discussing the number of copies of a connected graph H with exactly 1 cut-vertex and three different blocks attached to it. Let H_1, H_2 and H_3 denote these blocks, and v the cut-vertex of H. Furthermore we denote by $H_{1,2}$ the graph spanned by the vertices of H_1 and H_2, and similarly $H_{1,3}$ and $H_{2,3}$. The unique cut-vertex in H induces a vertex in each H_i, $H_{i,j}$ that we denote by $c(H_i)$ and $c(H_{i,j})$, respectively.

Before analyzing the corresponding equations, let us show how a copy of H may be created in the graph class under study. When creating a rooted connected graph by joining blocks to the root, we have the following possibilities for a copy of H:

(A) Copies that are incident with the root vertex but not contained in one of the blocks of the graph.

($\underline{\text{B}}$) Copies that are not incident with the root vertex and not contained in any block incident with the root vertex, built by taking subgraphs of H already existing in the blocks incident with the root vertex and completing them by pasting elements in $\mathcal{C}^\bullet$.

($\underline{\text{C}}$) Copies already existing in the blocks incident with the root vertex.

See the analogy with the study of the number of copies of P_2 (Case (A) and Case ($\underline{\text{A}}$), etc).

Let encode now the problem by means of generating functions. All indices in this subsection are vectors with six components, of the form $L = (l_1, l_2, l_3; l_{1,2}, l_{1,3}, l_{2,3})$. As we will show, such an index will encode the number of copies of H_i and $H_{i,j}$ incident with a certain vertex. Let $\mathbf{w}$ be the infinite vector with components w_K, with $K = (k_1, k_2, k_3; k_{1,2}, k_{1,3}, k_{2,3})$ being an index with 6 entries. We denote by $B^\circ_L(\mathbf{w}; u)$, $L = (l_1, l_2, l_3; l_{1,2}, l_{1,3}, l_{2,3})$ the generating function of derived blocks in $\mathcal{G}$, where the root vertex is incident with l_i copies of H_i ($i \in \{1, 2, 3\}$) and $l_{i,j}$ copies of $H_{i,j}$ ($i \neq j$) at $c(H_i)$ and $c(H_{i,j})$, respectively. We use the variable w_K to encode the number of vertices which are incident with k_i copies of H_i, and $k_{i,j}$ copies of the subgraphs $H_{i,j}$ at $c(H_i)$. We also use the variable u to count the number of copies of H. We note that different copies of the same subgraph H_i or $H_{i,j}$ could be overlapping. From the previous definition, it is obvious that writing $w_K = x$ in $B^\circ_L(\mathbf{w}; u)$ for all K we obtain the generating function $B^\circ_L(x, u)$ where now x counts the total number of vertices. As in the analysis of P_2, if this generating function is convergent for some positive x and for u with $|u| = 1$ then $B^\circ_L(\mathbf{w}; u)$ converges for all w_K with $|w_K| < x$ and for all u with $|u| = 1$.

For a vector index $R = (r_1, r_2, r_3; r_{1,2}, r_{1,3}, r_{2,3})$, let $C^\bullet_R(x, u)$ be the generating function of vertex-rooted connected graphs in $\mathcal{G}$, where the root vertex is incident with r_i copies of H_i at $c(H_i)$ and similarly for the numbers r_{ij} and the subgraphs H_{ij}, and where u counts the number of occurrences of H. Each of these functions satisfies the following equation

$$C^\bullet_R(x, u) = x \sum_{s \geq 0} \frac{1}{s!} \sum^{*}_{\{L_1, \ldots, L_s\}} u^{\sum_{i\neq j\neq k}(l_1^i l_2^j l_3^k) + \sum_{i \neq j}\left(l_{1,2}^i l_3^j + l_{1,3}^i l_2^j + l_{1,2}^i l_3^j + l_{2,3}^i l_1^j\right)} \prod_{i=1}^{s} B^\circ_{L_i}(\mathbf{W}, u) \tag{3.9}$$

where the sum $\sum_{i \neq j \neq k}$ is taken over triplets with pairwise different indices, and the sum $\sum^*_{\{L_1, \ldots, L_s\}}$ is taken over all sets of s indices $L_i = (l_1^i, l_2^i, l_3^i; l_{1,2}^i, l_{1,3}^i, l_{2,3}^i)$, $i = 1, \ldots, s$ satisfying

$$\sum_{i=1}^{s} l_1^i = r_1, \ \sum_{i=1}^{s} l_2^i = r_2, \ \sum_{i=1}^{s} l_3^i = r_3,$$
$$\sum_{i=1}^{s} l_{1,2}^i + \sum_{i \neq j} l_1^i l_2^j = r_{1,2}, \ \sum_{i=1}^{s} l_{1,3}^i + \sum_{i \neq j} l_1^i l_3^j = r_{1,3}, \ \sum_{i=1}^{s} l_{2,3}^i + \sum_{i \neq j} l_2^i l_3^j = r_{2,3},$$

and the infinite vector $\mathbf{W}$ has components

$$W_K = \sum_P u^{k_1 p_{2,3} + k_2 p_{1,3} + k_3 p_{1,2} + k_{1,2} p_3 + k_{1,3} p_2 + k_{2,3} p_1} C_P^\bullet(x,u). \tag{3.10}$$

Observe that the term $u^{\sum_{i\neq j\neq k}(l_1^i l_2^j l_3^k)+\sum_{i\neq j}\left(l_{1,2}^i l_3^j + l_{1,3}^i l_2^j + l_{1,2}^i l_3^j + l_{2,3}^i l_1^j\right)}$, the terms in u in $\mathbf{W}$ and the last argument in the term $B^\circ_{L_i}(\mathbf{W},u)$ are the contributions to Case (A), (B) and (C), respectively.

Equation (3.9) reads in the following way: a pointed connected graph in the family where the root vertex is incident with r_i copies of H_i at $c(H_i)$ (and similarly for the numbers r_{ij} and the subgraphs H_{ij}) is obtained by pasting a set of blocks at the root vertex, and adding the extra copies of H created, both those arising from the root vertex (Case (A)), and those arising from the decomposition of the blocks with the copies of connected rooted graphs (Case (B)). This last term is encoded by means of the term in u after the sum $\sum^*_{L_i}$.

The analysis of this system of equations is very similar to the study of the number of copies of P_2 and can be mimic without any difficulty. The only technical point in the analysis is that we have to check several properties in the functional space introduced in Remark 3.3.6. Let us also mention that very similar arguments (with more indices) apply for subgraphs H with exactly one cut-vertex (even with more than three blocks and possible block repetitions).

3.5.3 Main Example 2: Counting copies of P_3

We present an additional warm-up example, where we show a new difficulty that arises for subgraphs with more than one cut-vertex. As we will see, it is not enough to express the infinite system of equations in terms of 'indexed' block families counting formulas. Indeed, for each block in the class (and for each set of blocks) we will need a very precise information of its internal structure. It will turn out that Theorem 3.3.5 does not directly apply. However, we will show how this problem can be overcome.

For illustrative reasons of this phenomenon, we just study the number of copies of P_3 on the subcritical class graph where all 2-connected blocks are isomorphic to K_4 minus an edge. This family is indeed subcritical due to the fact that the generating function for blocks is analytic (see [28]). We denote by $C^\bullet_{k,l}(x,u)$ the generating function of (vertex) rooted connected graphs in the family where the root vertex has degree k and is the starting point of l paths of type P_2 (possibly intersecting). As usual, u marks occurrences of P_3.

In our setting, we have $B^\circ(x) = x^3$. Observe that (up to the labellings of the vertices) K_4 minus an edge has two different ways to be rooted: either over a vertex of degree 2 or degree 3. We call the resulting derived objects b_1° and b_2° with generating functions $b_1^\circ(x) = b_2^\circ(x) = \frac{1}{2}x^3$.

Let us now describe the system of equations satisfied by $C^\bullet_{k,l}(x,u)$, or at least the form of the first equations for small indices. It is obvious that $C^\bullet_{0,0}(x,u) = x$, that for every choice of $l \neq 0$, $C^\bullet_{0,l}(x,u) = 0$. Also, for every choice of $l \geq 0$, $C^\bullet_{1,l}(x,u) = 0$.

Expressions for $C^{\bullet}_{2,l}(x,u)$ and $C^{\bullet}_{3,l}(x,u)$ become more involved: in both cases we may have a block of type b_1° (and b_2°, respectively) incident with the root of the connected object. See Figure 3.1 for a general structure of both cases.

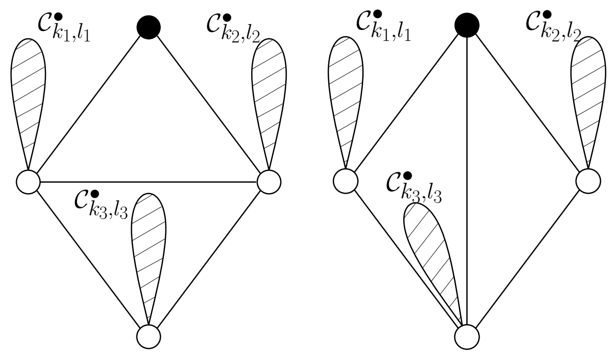

Figure 3.1: Block structure when dealing with $C^{\bullet}_{2,l}$ (left) and $C^{\bullet}_{3,l}$ (right). The shaded regions represent different rooted connected copies.

Following the notation in Figure 3.1, by writing $I = (k_1, l_1, k_2, l_2, k_3, l_3)$ we have the following relations:

$$C^{\bullet}_{2,l}(x,u) = \frac{1}{2}xu^6 \sum_{I}^{*} u^{k_1k_2+k_1k_3+k_2k_3+4(k_1+k_2+k_3)+3(l_1+l_2)+2l_3} C^{\bullet}_{k_1,l_1}(x,u)C^{\bullet}_{k_2,l_2}(x,u)C^{\bullet}_{k_3,l_3}(x,u),$$

$$C^{\bullet}_{3,l}(x,u) = \frac{1}{2}xu^6 \sum_{I}^{**} u^{k_1k_3+k_2k_3+4(k_1+k_2+k_3)+2(l_1+l_2)+3l_3} C^{\bullet}_{k_1,l_1}(x,u)C^{\bullet}_{k_2,l_2}(x,u)C^{\bullet}_{k_3,l_3}(x,u),$$

where the first sum $\sum_I^*$ is taken over indices I satisfying $k_1 + k_2 + 4 = l$ and the second sum $\sum_I^{**}$ is taken over indices I satisfying $k_1 + k_2 + k_3 + 4 = l$. Both formulas are easily explained by checking the structure depicted in Figure 3.1. In particular, the initial term u^6 encodes the 6 existing different paths of length 3 in the graph K_4 minus an edge.

Let us describe more precisely the leftmost term defining $C^{\bullet}_{2,l}(x,u)$. An object counted in $C^{\bullet}_{2,l}(x,u)$ is obtained by pasting three rooted connected objects over vertices of a block of type b_1°. Then, the final number of paths of length 3 arise from the following contributions:

(a) The existing paths of length 3 in each of the pasted rooted connected components.

(b) The existing paths of length 3 in b_1° (6 in total).

(c) Paths that are created by concatenating paths of length 1 in b_1° with paths of length 2 in each pasted rooted connected component.

(d) Paths that are created by concatenating paths of length 2 in b_1° with paths of length 1 in each pasted rooted connected component.

(e) Paths created by using 2 paths of length 1 in a pair rooted connected components which are linked in b_1° by a path of length 1.

As mentioned above, the most difficult term to be encoded is the one in Case (e) and it is given by the correlation term $k_1k_2 + k_2k_3 + k_2k_3$, which is built explicitly using the internal structure of b_1° and the set of indices I.

The situation is even more involved if several blocks are attached to the root. For example, the equations for $C^\bullet_{4,l}(x,u)$ and $C^\bullet_{5,l}(x,u)$ require the whole information of the two attached blocks. Nevertheless, it is clear how to set up an infinite system of equations for the functions $C^\bullet_{k,l}(x,u)$.

Unfortunately this system does not satisfy all assumptions of Theorem 3.3.5. Namely if we set $u = 1$ we obtain for example

$$C^\bullet_{2,l}(x,1) = \frac{1}{2}x \sum_I^* C^\bullet_{k_1,l_1}(x,1)C^\bullet_{k_2,l_2}(x,1)C^\bullet_{k_3,l_3}(x,1)$$

where this sum is taken over indices I satisfying $k_1 + k_2 + 4 = l$. This means that the right hand side cannot be written in terms of $C^\bullet(x) = \sum_{k,l} C^\bullet_{k,l}(x,1)$, and hence the Partition Property (3.6) cannot be satisfied.

However, it is possible to modify our setting slightly. Instead of analyzing the block decomposition related to the equation $C^\bullet(x) = x\exp(B^\circ(C^\bullet(x)))$ we iterate this equation and replace it by

$$C^\bullet(x) = x\exp(B^\circ(x\exp(B^\circ(C^\bullet(x))))),$$

which means that we specify first a tree of height two of (rooted) blocks before we substitute each vertex by $C^\bullet$ in order to obtain a recursive description for $C^\bullet$. Observe that this equation is well-defined for $x < \rho_C$, because of the subcritical condition $C^\bullet(\rho_C) < \rho_B$.

We show this procedure by considering one special instance that is part of the equation for $C^\bullet_{2,6}(x,u)$, compare with Figure 3.2. Here the root block is of type b_1°. One non-root vertex of this block is attached to another block of type b_1°, a second non-root vertex is attached to a block of type b_2°, whereas the third non-root vertex has no block attached. It is clear that such a block structure will lead to a connected graph of type $(k,l) = (2,6)$ - and there is another instance similar to that which cover then all situations of this form (switch the bottommost K_4 minus an edge and paste it using a vertex of degree 3 instead a vertex of degree 2, as drawn in Figure 3.2).

The corresponding generating function is then of the form

$$x^4u^{64}\sum_K u^{H(K)}\prod_{i=1}^{6} C^\bullet_{k_i,l_i}(x,u),$$

where the sum is taken over all indices $K = (k_1,l_1,k_2,l_2,k_3,l_3,k_4,l_4,k_5,l_5,k_6,l_6)$ and

$$\begin{aligned} H(K) &= k_1k_2 + k_1k_3 + k_2k_3 + 3(l_1+l_2) + 2l_3 + 7(k_1+k_2) + 4k_3 + \\ &\quad k_4k_6 + k_5k_6 + 2(l_4+l_5) + 3l_6 + 6(k_4+k_5+k_6). \end{aligned}$$

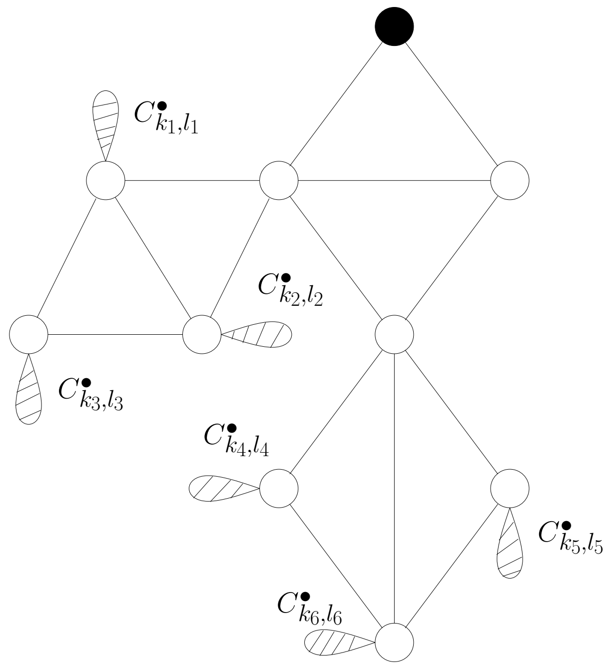

Figure 3.2: Construction of rooted connected objects of type $(k, l) = (2, 6)$.

Observe that the term u^{64} encodes the number of copies of P_3 in the block structure drawn in Figure 3.2: 6 copies on each of the three blocks (18 in total), 40 copies using edges of two blocks (2 edges in one block, another edge in an adjacent block) and finally 6 copies with an edge on each block.

In the same way we can deal with the other case which leads to a functional equation for $C^\bullet_{2,6}(x, u)$ of the form

$$C^\bullet_{2,6}(x, u) = F_{2,6}\left(x, u, (C^\bullet_{k,l}(x, u))_{k,l\geq 0}\right).$$

3.5.4 The general case. Proof of Theorem 3.1.1

We finally deal with the study of the number of copies of a general fixed connected subgraph H. Recall that the new difficulty emerging when considering copies of P_3 was the existence of a correlation between the root type and the root types of the attached connected graphs. In this section we show how we can overcome this problem. We start with the observation that the equation characterizing (rooted) connected graphs in terms of blocks can be iteratively written as follows:

$$C^\bullet(x) = x\exp(B^\circ(C^\bullet(x))) = x\exp(B^\circ(x\exp(B^\circ(C^\bullet(x))))) = \dots. \quad (3.11)$$

When stopping after h iterations, Equation (3.11) says that a rooted connected graph is obtained by repeating h times the operation of pasting a set of rooted blocks on vertices, and finally substituting recursively rooted connected graphs on each vertex – in the previous section we did just one interation.

We introduce now some terminology and notation. Recall that the *block graph* of a graph G (namely, the graph whose vertices are the blocks and cut-vertices of G, and edges are defined between each block and its incident cut-vertices) is a tree. Let $c^\bullet \in \mathcal{C}^\bullet$ be a rooted graph in our graph class. We define a distance for blocks with respect to the root vertex as follows: blocks which are incident with the root vertex are at distance 1 from the root vertex. Blocks that are not at distance 1 from the root vertex but are incident with some block of this type are at distance 2 from the root vertex, and so on. We say then that the set of blocks of $c^\bullet$ which are at distance at most h from the root in its block graph is the *h-root block* of $c^\bullet$. We define $\mathcal{B}^{(h),\circ}$ to be the family of all possible h-root blocks. We write $B^{(h),\circ}(x,w)$ for the EGF associated to $\mathcal{B}^{(h),\circ}$, where x encodes vertices on the h-root block until level $h-1$, while the extra parameter w encodes vertices belonging to the blocks pasted in the last step of the iteration (namely, at level h). Then, the following holds:

$$B^{(h),\circ}(x,w) = \exp\left(B^\circ(xB^{(h-1),\circ}(x,w))\right)$$

with initial condition $B^{(1),\circ}(x,w) = \exp(B^\circ(w))$. From Equation (3.11) we get that for each $h \geq 1$, $C^\bullet(x) = xB^{(h),\circ}(x, C^\bullet(x))$. In particular $C^\bullet(x) = xB^{(1),\circ}(x, C^\bullet(x)) = x\exp(B^\circ(C^\bullet(x)))$. Due to the fact that each iteration in Equation 3.11 has well-defined unique solution $C^\bullet(x)$ arising from a subcritical composition scheme, we may assume that all the analysis will be done for points x, w where the function $B^{(h),\circ}(x,w)$ is analytic. We also write $B^{(h),\circ}(x,w,u)$ for the counting formula of h-root blocks, where u marks copies of the subgraph H.

Let us now study substructures of H that will be necessary for the encoding. Assume that the block graph of H has diameter h. The main observation we exploit is that all copies of H which are incident to the root vertex of $c^\bullet$ are contained in the h-root block of $c^\bullet$. Let $\mathcal{H}_0 = \{H_1, \dots, H_s\}$ be all the connected subgraphs spanned by subsets of blocks of H. For a given $H_i \in \mathcal{H}_0$ we denote by $\overline{H_i}$ the set of blocks in H not contained in H_i. Given $H_i \in \mathcal{H}_0$ we say that a vertex v in H_i is a *virtual cut-vertex* if it is either a cut-vertex in H_i, or when we embed H_i in H, the resulting vertex becomes a cut-vertex in the ambient graph H. See Figure 3.3 for an example of a subgraph with 3 blocks and 4 virtual cut-vertices (2 of them cut-vertices).

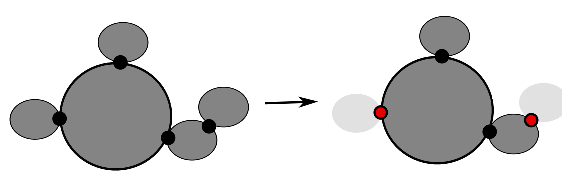

Figure 3.3: The graph H and a subgraph of H with four virtual cut-vertices (two of them, in red, are not cut-vertices in the subgraph). Each dark circle denotes a block.

We denote then by $\mathcal{H}$ the family of graphs constructed from the graphs in $\mathcal{H}_0$ by rooting one of its virtual cut-vertices. Let d denote the cardinality of $\mathcal{H}$ and let $\mathcal{I} = \mathbb{N}^d$ be the set of d-dimensional indices I. For every $I = (i_1, \dots, i_d) \in \mathcal{I}$ (which we also call *profile*) we consider the combinatorial family $\mathcal{C}_I^\bullet$ (with exponential generating function $C_I^\bullet(x, u)$) of rooted connected graphs in $\mathcal{C}^\bullet$ with i_j copies of the j-th subgraph of $\mathcal{H}$, $1 \leq j \leq d$, where the virtual cut-vertex coincides with the root vertex of the connected graph in $\mathcal{C}_I^\bullet$. Similarly, we define $\mathcal{B}_I^{(h),\circ}$ the family of h-rooted blocks whose profile is equal to I. Hence, $\mathcal{B}^{(h),\circ} = \bigcup_{I \in \mathcal{I}} \mathcal{B}_I^{(h),\circ}$.

Let $c^\bullet \in \mathcal{C}_I^\bullet$. There are three different types of copies of H in $c^\bullet$:

Case (c) Copies of H already existing in the rooted connected graphs that we attach at the h-root block of $c^\bullet$.

Case (bc) Copies created by using some subgraph of H from the h-root block of $c^\bullet$ and completing it to H by using attached rooted connected graphs with appropriate profiles.

Case (b) Copies of H already existing in the h-root block of $c^\bullet$.

See Figure 3.4 for an example of a subgraph H with $h = 2$, and three different copies of H arising from these 3 different sources.

We can now write an expression for $C_I^\bullet(x, u)$. Let b_h be a h-root block. Denote by $|b_h|_2$ the number of vertices of b_h on the h-level of the root block, and $|b_h|_1 = |b_h| - |b_h|_2$. We write $\mathbf{I}(|b_h|_2) = (I_1, \dots, I_{|b_h|_2})$. This set of profiles will be the ones of the rooted connected graphs that we will attach to each vertex of the h-level of the h-root block. Also, given fixed b_h and a set of profiles $\mathbf{I}(|b_h|_2)$, we write

- $G_1(b_h)$ for the number of copies of H in Case (b).
- $G_2(b_h, \mathbf{I}(|b_h|_2))$ for the number of copies of H in Case (bc).

Observe that both $G_1(b_h)$ and $G_2(b_h, \mathbf{I}(|b_h|_2))$ depend on the specific structure of b_h (and also on the set of the profiles $\mathbf{I}(|b_h|_2)$ in G_2). With this terminology in mind, now it is easy to write an equation for $C_I^\bullet(x, u)$:

$$C_I^\bullet(x, u) = x \sum_{b_h \in \mathcal{B}_I^{(h),\circ}} \sum_{\mathbf{I}(|b_h|_2)} u^{G_1(b_h) + G_2(b_h, \mathbf{I}(|b_h|_2))} \frac{x^{|b_h|_1}}{|b_h|_1!} \prod_{i=1}^{|b_h|_2} C_{I_i}^\bullet(x, u), \tag{3.12}$$

where the second sum is extended to all possible sets of $|b_h|$ profiles. We are now ready to prove Theorem 3.1.1 by analyzing Equation (3.12). We write $\mathbf{y} = (C_I^\bullet(x, u))_I$, which is a solution to the infinite system of equations $y_I = F_I(x, \mathbf{y}, u)$ with

$$y_I = F_I(x, \mathbf{y}, u) = x \sum_{b_h \in \mathcal{B}_I^{(h),\circ}} \sum_{\mathbf{I}(|b_h|_2)} u^{G_1(b_h) + G_2(b_h, \mathbf{I}(|b_h|_2))} \frac{x^{|b_h|_1}}{|b_h|_1!} \prod_{j=1}^{|b_h|_2} y_{I_j}.$$

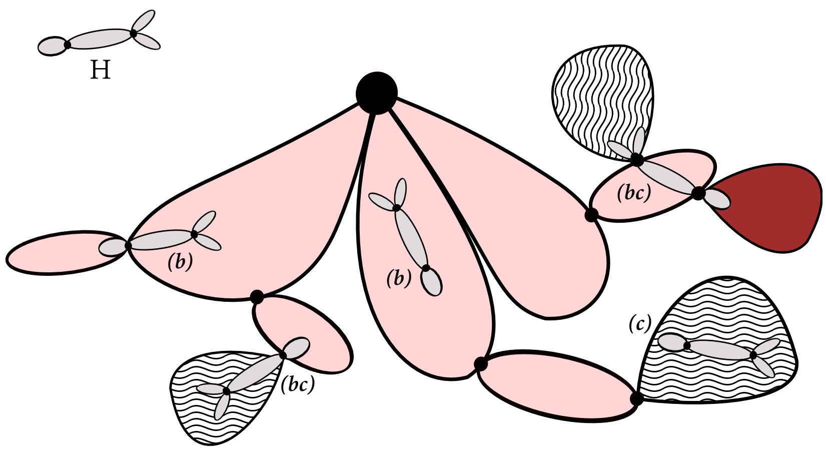

Figure 3.4: A generic copy of H in the construction may arise from three different sources. Examples of different contributions are shown with the corresponding case (either (b), (c) or (bc)).

We can now check that this system of equations satisfies the conditions of Theorem 3.3.5. We may assume in all the analysis that all variables y_I, x are positive. Let us start with the Partition Property (3.6). By writing $u = 1$, we get that $F_I(x, \mathbf{y}, 1)$ is equal to

$$F_I(x, \mathbf{y}, 1) = x \sum_{b_h \in \mathcal{B}_I^{(h),\circ}} \frac{x^{|b_h|_1}}{|b_h|_1!} \prod_{j=1}^{|b_h|_2} y_{I_j} = x B_I^{(h),\circ}\left(x, \sum_J y_J\right).$$

Hence, $F_I(x, \mathbf{y}, 1)$ is equal to $xB_I^{(h),\circ}(x, \sum_J y_J) = \widetilde{F}_I(x, \sum_J y_J)$, and Condition (3.5) is satisfied. Let us check now Condition (3.6). Observe that

$$\sum_I \widetilde{F}_I(x, y) = \sum_I x B_I^{(h),\circ}(x, y) = x B^{(h),\circ}(x, y) = F(x, y),$$

which is analytic in x and y due to the subcritical condition (recall that with this notation, $C^\bullet(x) = F(x, C^\bullet(x))$). Also, the condition assuring that this system of equations is well defined and analytic in the functional space considered in Remark 3.3.6 is satisfied by taking a sufficiently large (but bounded) number of derivatives of $\widetilde{F}_I(x,y)$ with respect to y. Again, by the subcritical condition all these derivatives are bounded and consequently, for each choice of $m \geq 1$

$$\sum_I \|I\|_1^m \widetilde{F}_I(x,y) \leq \left(xy\frac{\partial^2}{\partial x \partial y}\right)^{mf_1(|H|)} F(x,y),$$

for a certain function f_1 that only depends on the size of $|H|$ (and hence, it is bounded). This fact finally proves the first part of the conditions.

Let us show now Condition (3.7). We only argue the case $r = 1$, as the arguments for the second and the third derivatives are very similar (but much longer). We will show that the terms can be bounded by a constant number of derivatives (depending on $|H|$) of an analytic function, hence the resulting value will be bounded as well. We need first to bound the following derivative at $t = 0$:

$$\begin{aligned}\left|\frac{\partial}{\partial t}F_I(x,\mathbf{y},e^{it})\right| &= \left|x\frac{\partial}{\partial t}\sum_{b_h\in\mathcal{B}_I^{(h),\circ}}\sum_{\mathbf{I}(|b_h|_2)} e^{it(G_1(b_h)+G_2(b_h,\mathbf{I}(|b_h|_2)))}\frac{x^{|b_h|_1}}{|b_h|_1!}\prod_{j=1}^{|b_h|_2} y_{I_j}\right| \\ &\leq x\sum_{b_h\in\mathcal{B}_I^{(h),\circ}}\sum_{\mathbf{I}(|b_h|_2)} (G_1(b_h)+G_2(b_h,\mathbf{I}(|b_h|_2)))\frac{x^{|b_h|_1}}{|b_h|_1!}\prod_{j=1}^{|b_h|_2} y_{I_j}\end{aligned}$$

Hence we have two different contributions, namely expressions $G_1(b_h)$ and $G_2(b_h, \mathbf{I}(|b_h|_2))$. Observe first that $G_1(b_h)$ counts the number of copies of H in b_h, hence it is bounded by $O(|b_h|^{|H|})$. Consequently, we have the bound

$$x\sum_{b_h\in\mathcal{B}_I^{(h),\circ}}\sum_{\mathbf{I}(|b_h|_2)} G_1(b_h)\frac{x^{|b_h|_1}}{|b_h|_1!}\prod_{j=1}^{|b_h|_2} y_{I_j} \leq \left(xy\frac{\partial^2}{\partial x\partial y}\right)^{|H|}\widetilde{F}_I\left(x,\sum_J y_J\right)$$

which is bounded, and hence

$$\begin{aligned}&x\sum_I \|I\|_1^m \sum_{b_h\in\mathcal{B}_I^{(h),\circ}}\sum_{\mathbf{I}(|b_h|_2)} G_1(b_h)\frac{x^{|b_h|_1}}{|b_h|_1!}\prod_{j=1}^{|b_h|_2} y_{I_j} \\ &\leq x\sum_I \|I\|_1^m\left(xy\frac{\partial^2}{\partial x\partial y}\right)^{|H|}\widetilde{F}_I\left(x,\sum_J y_J\right) \leq x\left(y\frac{\partial^2}{\partial x\partial y}\right)^{mf_1(|H|)+|H|} F\left(x,\sum_J y_J\right).\end{aligned}$$

It finally remains to study the contribution $G_2(b_h, \mathbf{I}(|b_h|_2))$, which is the number of copies of H created in Case (bc). We can obtain a bound for $G_2(b_h, \mathbf{I}(|b_h|_2))$ by using that the size of H is bounded by $|H|$. Observe that any copy created in Case (bc)

(and hence counted by G_2) is obtained by taking a subgraph of H in the h-root block, and completing it to H by attaching at most $|H|$ substructures arising from pending connected graphs. The number of subgraphs of $|H|$ in b_h is bounded by $|b_h|^{f_2(|H|)}$, for a certain function f_2. This means that

$$G_2(b_h, \mathbf{I}(|b_h|_2)) \leq |b_h|^{f_2(|H|)} \sum_{*} \|I_{j_1}\|_1 \dots \|I_{j_{|H|}}\|_1,$$

where the sum in the previous expression is extended to all subsets of size $|H|$ of $\{1, \dots, |b_h|\}$. Observe that the total number of terms of the sum is bounded then by $|b_h|^{|H|}$. Putting it all together we get the following:

$$\begin{aligned}
& x \sum_{b_h \in \mathcal{B}_I^{(h),\circ}} \sum_{\mathbf{I}(|b_h|)} G_2(b_h; \mathbf{I}(|b_h|_2)) \frac{x^{|b_h|_1}}{|b_h|_1!} \prod_{j=1}^{|b_h|_2} y_{I_j} \\
& \leq x \left(\sum_I \|I\|_1 |y_I| \right)^{|H|} \sum_{b_h \in \mathcal{B}_I^{(h),\circ}} |b_h|^{f_2(|H|)} |b_h|^{|H|} \frac{x^{|b_h|_1}}{|b_h|_1!} \prod_{j=1}^{|b_h|_2} y_{I_j} \\
& \leq \left(\sum_I \|I\|_1 |y_I| \right)^{|H|} \left(xy \frac{\partial^2}{\partial x \partial y} \right)^{f_2(|H|)+|H|} \widetilde{F}_I \left(x, \sum_J y_J \right).
\end{aligned}$$

By assumption, the sum $\sum_I \|I\|_1 |y_I|$ is bounded, hence the previous term is bounded as well. Finally we can get bounded expressions for the weighted sum with coefficients $\|I\|_1^m$, as we did when analyzing the function $G_1(b_h)$. This concludes the study for the first derivative. As mentioned, case $r = 2$ and $r = 3$ can be similarly handled and obtain similarly bounded expressions. This concludes the proof of Theorem 3.1.1.

3.6 Computations

The method used to prove our Theorem 3.1.1 provides normal limiting distributions with linear expectation and variance. In general, the constants appearing in both the expectation and the variance are complicated to be estimated, as they depend on the solution of a non-linear system of equations with infinitely many equations and variables. However, in some few cases we can get explicit computational results. In this section we address this problem by analyzing the statistics of some small subgraphs in series-parallel graphs. Recall that a series-parallel graph is a graph such that its 2-connected components are obtained from an edge by means of series and parallel operations. In particular, we compute the subgraph statistics for triangles in a 2-connected and connected series-parallel graph of order n (see further comments on this graph class in Section 3.1). In this prominent case, it is straightforward to apply Remark 3.3.3 in order to justify that the corresponding constant σ_H^2 is positive. Hence, for all subgraphs in the connected level the second statement in Theorem 3.1.1 will hold. Additionally, our methodology gives easily the asymptotic enumeration of series-parallel graphs avoiding a triangle. The advantage of series-parallel graphs is that there is a constructive definition that allows us to encode and count triangles.

In order to analyze series-parallel graphs we use a variant of Tutte's decomposition into 3-connected components, as depicted in [17]. Recall that this strategy is used when a class of graphs satisfies that a graph belongs to the family if and only if its connected, 2-connected and 3-connected components also belong to, as it is the case in series-parallel graphs.

Note that in Section 3.4 we already proved that the number of copies of a 2-connected subgraph in a connected graph is normally distributed. In this section we prove the same for the 2-connected level in the particular class of series-parallel graphs. In this situation Theorem 3.1.1 does not apply. However, we can again get a Central Limit Theorem by means of the Quasi-Powers Theorem (see [34]).

3.6.1 Triangles in series-parallel graphs

As already mentioned, a connected graph is obtained from its tree decomposition into 2-connected blocks. We can also decompose 2-connected graphs into 3-connected graphs by means of networks. Recall that a *network* is a graph with two distinguished vertices, called *poles*, such that the graph obtained by adding an edge between the two poles (if they are not adjacent) is 2-connected (see Tutte's monograph [17]). In the case of series-parallel graphs there are no 3-connected graphs, so we start with networks as the basic building blocks. The key point here is that networks are easy enough to be built, so we can control the copies of simple structures, like cycles. If these structures are 2-connected then they can only appear inside 2-connected blocks, so if we count them at the 2-connected level then Tutte's decomposition gives the total number for general graphs.

In one of the steps of the decomposition we have to obtain a 2-connected graph from a network. In general, a network is obtained by picking an edge of the 2-connected graph and performing some minor corrections. Therefore, in order to obtain a 2-connected graph from a network we have to 'forget' a root edge. Since we can translate the action of rooting an edge in terms of generating functions as differentiating with respect to the variable that counts edges, we can translate the opposite action (forgetting the root) as the integration with respect to the same variable. This was done in [8] to obtain the generating function of 2-connected planar graphs. However, we will use a more recent approach, purely combinatorial, defined following the ideas of the grammar developed in [2]. This approach uses the so-called Dissymmetry Theorem for trees [15]. This technique gives a bijection that relates unrooted trees and trees rooted in both a vertex and an edge, which is used to express the generating function of unrooted trees in terms of the generating function of rooted trees. In [2] the authors consider the decomposition of a 2-connected graph into networks. Since the class is tree-decomposable, they show that the Dissymmetry Theorem can be used to obtain the generating function of 2-connected graphs in terms of the generating function of the networks, with no integration involved.

Equations

Since there are no 3-connected series-parallel graphs, we start by computing the generating functions $D^{\blacktriangle}(x, y, u)$ of networks, where x mark vertices and y marks edges. We add the additional parameter u which is used to encode triangles.

Recall that a network is obtained from a 2-connected series-parallel graph by choosing and orienting an edge. It might not occur in the graph, and the vertices incident to it, the poles, are not labelled, but instead one of them is consider to be 0, and the other one is ∞. For convenience we split both series and parallel generating functions as follows. We define $P_0^{\blacktriangle} := P_0^{\blacktriangle}(x, y, u)$ as the generating function of parallel networks that do not contain an edge between the poles, whereas $P_1^{\blacktriangle} := P_1^{\blacktriangle}(x, y, u)$ is the generating function of parallel networks where there is an edge connecting the poles. For convenience, we include the network consisting of a single edge in $P_1^{\blacktriangle}$.

We define $S_2^{\blacktriangle} := S_2^{\blacktriangle}(x, y, u)$ as the generating function of series networks where there is a path of length exactly 2 between the poles, or equivalently where there exists a single cut-vertex. We also define $S_3^{\blacktriangle} := S_3^{\blacktriangle}(x, y, u)$ as the remaining series networks. Namely, the ones where the graph distance between the poles is at least 3. Figure 3.5 shows the structures for $P_0^{\blacktriangle}$, $P_1^{\blacktriangle}$ and $S_2^{\blacktriangle}$.

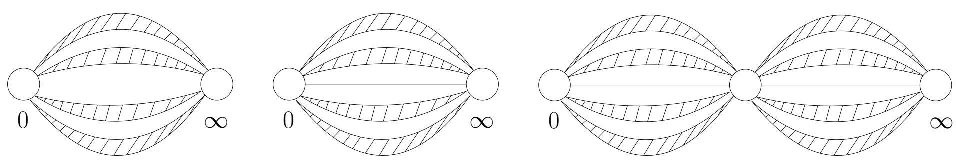

Figure 3.5: Structure of classes $P_0^{\blacktriangle}$, $P_1^{\blacktriangle}$ and $S_2^{\blacktriangle}$ respectively.

The generating function $D^{\blacktriangle} := D^{\blacktriangle}(x, y, u)$ can be expressed then as the solution of the following system of equations:

$$\begin{aligned} D^{\blacktriangle} &= P_0^{\blacktriangle} + P_1^{\blacktriangle} + S_2^{\blacktriangle} + S_3^{\blacktriangle} \\ P_0^{\blacktriangle} &= \exp_{\geq 2}(S_2^{\blacktriangle} + S_3^{\blacktriangle}) \\ P_1^{\blacktriangle} &= y \exp(uS_2^{\blacktriangle} + S_3^{\blacktriangle}) \\ S_2^{\blacktriangle} &= x(P_1^{\blacktriangle})^2 \\ S_3^{\blacktriangle} &= xD^{\blacktriangle}P_0^{\blacktriangle} + xP_1^{\blacktriangle}(P_0^{\blacktriangle} + S_2^{\blacktriangle} + S_3^{\blacktriangle}). \end{aligned} \tag{3.13}$$

The previous formulas read as follows: a graph in $P_0^{\blacktriangle}$ is obtained as a set of at least two series networks in parallel, since no series network has an edge between the poles. A graph in $P_1^{\blacktriangle}$ is obtained by putting a set of series networks in parallel with an edge. Note that each series network that contains a path of length 2 between the poles will produce a triangle. A graph in $S_2^{\blacktriangle}$ has a single cut-vertex, and an edge joining it to both poles, which might be in parallel with other series networks, so we need two copies of $P_1^{\blacktriangle}$. A graph in $S_3^{\blacktriangle}$ has at least one cut-vertex. Let x be the cut-vertex closest to pole 0. There are two options: either x is joined to pole 0 by an edge, and therefore by a network in $P_1^\blacktriangle$, or it is joined to pole 0 by a network in $P_0^\blacktriangle$. In the former case, there cannot be an edge between x and pole ∞, so there must be a network in $P_0^\blacktriangle$, $S_2^\blacktriangle$ or $S_3^\blacktriangle$ that joins x and pole ∞. In the latter case any network is possible, since the distance between the poles will be at least 3.

From these equations we deduce that triangles can only come up from parallel constructions where the poles are connected by an edge and at least one path of length 2. Observe that we cannot get a single equation for $D^\blacktriangle(x, y, u)$ from the System (3.13), due to the exponential operator. However we can obtain the first terms of its expansion near 0:

$$D^\blacktriangle(x,y,u) = y + \left(y^2 + uy^3\right)x + \left(2y^3 + 3y^4 + 4uy^4 + 5u^2y^5\right)\frac{x^2}{2!} + \left(6y^4 + 30y^5 + 7y^6 + 18uy^5 + 48uy^6 + 36u^2y^6 + 49u^3y^7\right)\frac{x^3}{3!} + O(x^4).$$

Now that we know $D^\blacktriangle$ and the auxiliary functions $P_0^\blacktriangle$, $P_1^\blacktriangle$, $S_2^\blacktriangle$ and $S_3^\blacktriangle$, we can use the Dissymmetry Theorem for trees in order to obtain the generating function $B^\blacktriangle(x, y, u)$ of 2-connected series-parallel graphs, where x, y, u marks vertices, edges and triangles, respectively. We will use the same approach as in [2]: since the class of 2-connected series-parallel graphs is tree-decomposable, the Dissymmetry Theorem gives the following bijection

$$B^\blacktriangle \; + \; B^\blacktriangle_{\circ\to\circ} \simeq B^\blacktriangle_\circ \; + \; B^\blacktriangle_{\circ-\circ},$$

where $B^\blacktriangle_\circ$ represents the class of 2-connected series-parallel graphs with a distinguished vertex in the tree decomposition, i.e., either a ring or a multiedge, $B^\blacktriangle_{\circ-\circ}$ represents the class of 2-connected series-parallel graphs with a distinguished edge in the tree decomposition, which must be incident to both a ring and a multiedge, and $B^\blacktriangle_{\circ\to\circ}$ represents the class of 2-connected series-parallel graphs with a distinguished oriented edge in the tree decomposition. A class of graphs is tree-decomposable if each graph has an associated tree whose nodes are distinguishable in some way. See [2] for further details. This leads to the following expressions:

$$\begin{aligned}
B_R^\blacktriangle(x,y,u) &= \mathrm{Cyc}(x(P_0^\blacktriangle + P_1^\blacktriangle)) + (u-1)\frac{(xP_1^\blacktriangle)^3}{6}, \\
B_M^\blacktriangle(x,y,u) &= \frac{x^2}{2}\left(y\exp_{\geq 2}(uS_2^\blacktriangle + S_3^\blacktriangle) + \exp_{\geq 3}(S_2^\blacktriangle + S_3^\blacktriangle)\right), \\
B_{MR}^\blacktriangle(x,y,u) &= \frac{x^2}{2}\left((S_2^\blacktriangle + S_3^\blacktriangle)(P_0^\blacktriangle + P_1^\blacktriangle - y) + (u-1)(P_1^\blacktriangle - y)S_2^\blacktriangle\right).
\end{aligned}$$

Let us explain each term. $B_R^\blacktriangle$ represents the class of 2-connected series-parallel graphs with a distinguished ring in the tree decomposition, where a ring, as in [2], is a cycle of length at least 3. $B_M^\blacktriangle$ represents the class of 2-connected series-parallel graphs with a distinguished multiedge in the tree decomposition, and $B_{MR}^\blacktriangle$ represents the class of 2-connected series-parallel graphs with a distinguished pair of incident ring and multiedge. In the case of $B_R^\blacktriangle$ we have to consider the special case where the ring is of length 3, and the parallel networks that replace the edges of the ring are of the kind $P_1^{\blacktriangle}$, since this generates a new triangle, as it is shown in Figure 3.6. In the case of $B_M^{\blacktriangle}$ we distinguish 2 cases, depending on whether one of the edges of the multiedge is not replaced with a series network, but with an edge, since this generates a new triangle for every other edge replaced with a series network in $S_2^{\blacktriangle}$. Note that, according to the definition of the tree decomposition, a multiedge has to be composed by at least three series networks. In the case of $B_{MR}^{\blacktriangle}$ we have to take into account the special situation where both conditions happen at the same time.

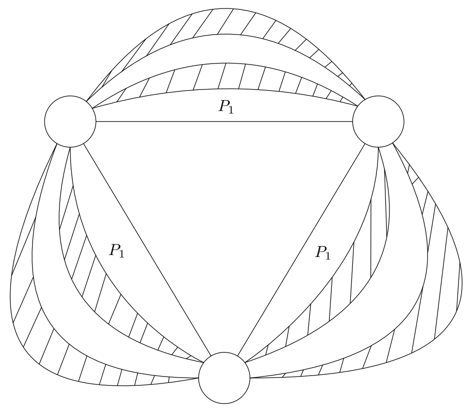

Figure 3.6: 2-connected series-parallel graph rooted in a ring of length 3. The shaded regions plus the edges represent parallel networks of the kind $P_1^{\blacktriangle}$.

Observe that we have obtained an explicit expression of $B_R^{\blacktriangle}$, $B_M^{\blacktriangle}$ and $B_{MR}^{\blacktriangle}$ in terms of $P_0^{\blacktriangle}$, $P_1^{\blacktriangle}$, $S_2^{\blacktriangle}$ and $S_3^{\blacktriangle}$. Finally, following [2, Section 5.3], the generating function $B^{\blacktriangle}(x,y,u)$ is obtained as

$$B^{\blacktriangle}(x,y,u) = \frac{1}{2}x^2y + B_R^{\blacktriangle}(x,y,u) + B_M^{\blacktriangle}(x,y,u) - B_{MR}^{\blacktriangle}(x,y,u). \tag{3.14}$$

In the last step we just compute the generating function $G^{\blacktriangle}(x,y,u)$ as the set of its connected components, encoded as the exponential of $C^{\blacktriangle}(x,y,u)$, which at the same time can be obtained from the decomposition into 2-connected components, encoded as $B^{\blacktriangle}(x,y,u)$, by a standard integration. This determines the generating function $G^{\blacktriangle}(x,y,u)$ of series-parallel graphs where x counts vertices, y counts edges and u counts triangles.

Number of triangles

Theorem 3.1.1 assures that the number of triangles in a series-parallel graph on n vertices choosing uniformly at random follows a Central Limit Theorem. We will proceed with the explicit computation of the constants. We will not consider the number of edges any more, so we can assume that $y = 1$. In all this section all generating functions are considered evaluated at $y = 1$, and hence, we only use variable x and u.

The first lemma gives the singularity type of the networks:

Lemma 3.6.1. *The generating function $S_3^\blacktriangle(x,u)$ of networks where the distance between the poles is greater than 2 satisfies*

$$S_3^\blacktriangle(x,u) = g_3(x,u) - h_3(x,u)\sqrt{1-\frac{x}{R^\blacktriangle(u)}},$$

for functions $g_3(x,u)$ and $h_3(x,u)$ analytic in a neighbourhood of the point $(x,u) = (R,1)$, $R = R^\blacktriangle(1) \approx 0.12800$, and $R^\blacktriangle(u)$ is the singularity curve of $S_3^\blacktriangle(x,u)$. A similar result is also true for $S_2^\blacktriangle(x,u)$ with certain functions $g_2(x,u)$ and $h_2(x,u)$ instead of $g_3(x,u)$ and $h_3(x,u)$.

Proof. We will use the techniques shown in [47]. In particular, we will use [12, Theorem 2], which is a consequence of [47, Theorem 2.33]. First, we need to adapt the equations so that they satisfy the hypothesis of [12, Theorem 2]. The new equations are:

$$\begin{aligned} S_2^\blacktriangle =& x\exp(2uS_2^\blacktriangle + 2S_3^\blacktriangle) = F_1(x, S_2^\blacktriangle, S_3^\blacktriangle, u) \\ S_3^\blacktriangle =& x\left(\exp_{\geq 2}(S_2^\blacktriangle + S_3^\blacktriangle)(\exp_{\geq 1}(S_2^\blacktriangle + S_3^\blacktriangle) + \exp(uS_2^\blacktriangle + S_3^\blacktriangle)) + \right. \\ & \left. \exp(uS_2^\blacktriangle + S_3^\blacktriangle)\exp_{\geq 1}(S_2^\blacktriangle + S_3^\blacktriangle)\right) = F_2(x, S_2^\blacktriangle, S_3^\blacktriangle, u) \end{aligned}$$

Since the functions F_1, F_2 are analytic in the complex plane, they satisfy the hypothesis of [12, Theorem 2]. Moreover, in [24] the authors show that for $u = 1$ the system has a unique solution, for which $x = R \approx 0.12800$. Since the system is aperiodic, there is a unique singularity, which implies the existence of a square-root expansion around $(x = R, u = 1)$, which in particular implies the statement. □

By means of Equation (3.13), all different network classes can be expressed in terms of both $S_3^\blacktriangle(x,u)$ and $S_2^\blacktriangle(x,u)$. Hence, all network classes have a similar expression. This observation makes the following lemma an straightforward result:

Lemma 3.6.2. *The generating function $B^\blacktriangle(x,u)$ of 2-connected series-parallel graphs where x marks vertices and u marks triangles satisfies*

$$B^\blacktriangle(x,u) = g_B(x,u) - h_B(x,u)\sqrt{1-\frac{x}{R^\blacktriangle(u)}}, \tag{3.15}$$

where g_B and h_B are analytic in a neighbourhood of the point $(x,u) = (R,1)$, $R = R^\blacktriangle(1) \approx 0.12800$, and where $R^\blacktriangle(u)$ is the function described in Lemma 3.6.1.

As described in [28], the dominant singularity of both $C(x,u)$ and $G(x,u)$ arises from a branch point of the equation defining $C^\bullet(x,u)$ in terms of $B^\circ(x,u)$. We write $\tau^\blacktriangle(u)$ the solution to the equation $\tau^\blacktriangle(u)B^{\circ'}(\tau^\blacktriangle(u),u) = 1$. The singularity of $C(x,u)$ (and also $G(x,u)$) is located at $\rho^\blacktriangle(u) = \tau^\blacktriangle(u)\exp(-B^\circ(\tau^\blacktriangle(u),u))$. Note that, since both $C(x,u)$ and $G(x,u)$ are aperiodic, the singularity is unique.

The next step is to deduce from the previous lemmas the limiting distribution for the number of triangles. We already know, due to Theorem 3.1.1, that in the connected level this random variable follows a normal limit distribution. However, we cannot claim from Theorem 3.1.1 a normal limit distribution for a subclass of connected series-parallel graphs. In the next theorem we obtain the result in the case of 2-connected graphs in the family, without using Theorem 3.1.1 :

Theorem 3.6.3. *The number of triangles $W_n^{\blacktriangle}$ on a 2-connected series-parallel graph with n vertices chosen uniformly at random is asymptotically normal distributed, with*

$$\mathbb{E}[W_n^{\blacktriangle}] = \mu_{\blacktriangle,2} n + O(1), \qquad \mathbb{V}\mathrm{ar}[W_n^{\blacktriangle}] = \sigma^2_{\blacktriangle,2} n + O(1),$$

where $\mu_{\blacktriangle,2} \approx 0.45242$ and $\sigma^2_{\blacktriangle,2} \approx 0.45997$.

Proof. To get the constant in the expectation and variance, we compute both $R^{\blacktriangle\prime}(u)$ and $R^{\blacktriangle\prime\prime}(u)$ by means of the equations for networks. As both parameters

$$\mu_{\blacktriangle,2} = -\frac{R^{\blacktriangle\prime}(1)}{R^{\blacktriangle}(1)}, \ \sigma^2_{\blacktriangle,2} = -\frac{R^{\blacktriangle\prime\prime}(1)}{R^{\blacktriangle}(1)} - \frac{R^{\blacktriangle\prime}(1)}{R^{\blacktriangle}(1)} + \left(\frac{R^{\blacktriangle\prime}(1)}{R^{\blacktriangle}(1)}\right)^2.$$

are strictly greater than 0, we can apply the Quasi-Powers Theorem over the expression in Equation (3.15), and the result holds straightforward. □

Finally, we are able to compute the number of triangles in a (connected or general) series-parallel graph of order n, chosen uniformly at random.

Theorem 3.6.4. *The number of triangles $X_n^{\blacktriangle}$ of a connected series-parallel graph with n vertices, chosen uniformly at random is asymptotically normal, with*

$$\mathbb{E}[X_n^{\blacktriangle}] = \mu_{\blacktriangle} n + O(1), \qquad \mathbb{V}\mathrm{ar}[X_n^{\blacktriangle}] = \sigma^2_{\blacktriangle} n + O(1),$$

where $\mu_{\blacktriangle} \approx 0.39481$ and $\sigma^2_{\blacktriangle} \approx 0.41450$. The same result holds for a general series-parallel graph with n vertices, chosen uniformly at random.

Proof. The normality of the random variable in the connected level is assured by the fact that series-parallel graphs are subcritical, and hence Theorem 3.1.1 applies in this setting. Normality in the general level is a consequence of Remark 3.3.2. Finally, we obtain the explicit value of the multiplicative constants in $\mathbb{E}[X_n^{\blacktriangle}]$ and $\mathbb{V}\mathrm{ar}[X_n^{\blacktriangle}]$ from the explicit expression of F in terms of B°. □

One may compare these values with the expected number of pendant triangles in a uniformly at random series-parallel graph on n vertices, computed in [28] as approximately $2.2313 \cdot 10^{-3} n$. As expected, the number of copies of a triangle in a random series-parallel graph is much smaller than the number of occurrences of the triangle in a random series-parallel graph.

Enumeration of triangle-free series-parallel graphs

If we write $u = 0$ in the equations of the previous section we get the generating function of triangle-free series-parallel graphs. In this subsection we provide the asymptotic analysis of such family, which is interesting by itself.

In all this section we use the equations in the introduction of Section 3.6.1 with the value $u = 0$. In order to emphasize that we are considering triangle-free families, we use the superindex $\triangle$ instead of $\blacktriangle$. We start studying the singular behaviour for networks.

Lemma 3.6.5. *Fix y in a small neighbourhood of 1. The generating function $S_3^{\triangle}(x,y)$ of triangle-free series networks where the poles are at a distance greater than 2 has a positive singularity $R^{\triangle}(y)$, and the following singular expansion in a dented domain Δ at a certain $x = R_{\triangle}(y)$:*

$$S_3^{\triangle}(x,y) = a_0^{\triangle}(y) + a_1^{\triangle}(y)X + a_2^{\triangle}(y)X^2 + a_3^{\triangle}(y)X^3 + O(X^4),$$

where $X = \sqrt{1 - x/R_{\triangle}(y)}$. In particular, $R_{\triangle}(1) \approx 0.19635$.

Proof. First, note that if we assign $u = 0$ in the equations defining the networks in Equation (3.13), then we can express $S_3^{\triangle}$ as the solution of the following single implicit equation:

$$S_3^{\triangle} = x\left(\left(\exp_{\geq 2}(xy^2 e^{2S_3^{\triangle}} + S_3^{\triangle})\right)\left(\exp_{\geq 1}(xy^2 e^{2\,S_3^{\triangle}} + S_3^{\triangle}) + ye^{S_3^{\triangle}}\right) + ye^{S_3^{\triangle}}\exp_{\geq 1}(xy^2 e^{2\,S_3^{\triangle}} + S_3^{\triangle})\right). \tag{3.16}$$

We write the right hand side of Equation (3.16) as $G(S_3^{\triangle}, x, y)$. Then, for every choice of y in a neighbourhood of 1, we need to check that $S_3^{\triangle}$ satisfies a so-called *Smooth Implicit-function Scheme* (see the work of Meir and Moon [33], se also [5, Section VII. 4.1.]) of the form $S_3^{\triangle} = G(S_3^{\triangle}, x, y)$. In this context, if G verifies some analytic conditions, then the solution $S_3^{\triangle}(x, y_0)$ of the equation admits an square root expansion in a domain dented at its singularity. We now check the conditions:

(a) $G(U, x, y)$ must be analytic in a given complex region. In our case it is an entire function.

(b) The coefficients $g_{m,n}(y)$ of the Taylor expansion of $G(U, x, y)$ with respect to U and x must be non-negative, as it is the case. Moreover $g_{0,0}(y) = 0$, and $g_{0,1}(y) = 0 \neq 1$.

(c) $g_{m,n}(y)$ must be positive for some m and some $n \geq 2$. Since $g_{1,2}(y) = 2y$, this holds for any y in a small neighbourhood of 1.

(d) The singularity must be unique, which is true since the generating function is aperiodic.

(e) Finally, for each choice of y in a small neighbourhood of 1, we need the existence of a solution $R_{\triangle}(y)$ and $a_0^{\triangle}(y)$ satisfying the characteristic system

$$a^{\triangle}(y) = G(a_0^{\triangle}(y), R_{\triangle}(y), y),\ 1 = G_U(a_0^{\triangle}(y), R_{\triangle}(y), y). \tag{3.17}$$

Direct computations for $y = 1$ gives that such system of equations has a valid solution at $a^{\triangle}(1) \approx 0.15545$ and $R_{\triangle} \approx 0.19635$. Finally, this statement is also true in a small neighbourhood of $y = 1$ by the fact that both equations in System (3.17).

In conclusion, the Implicit-function Scheme $S_3^{\triangle} = G(x, S_3^{\triangle}, y)$ is smooth for $y = 1$, by continuity of the equations it is also smooth for y in a small neighbourhood of y. Hence for each choice of y in a small neighbourhood of 1, $S_3^{\triangle}$ admits a square-root expansion in a domain dented at $R_{\triangle}(y)$, as we wanted to show. □

Once we have the singularity behaviour of $S_3^{\triangle}(x, y)$ for y close enough to 1, we can compute the coefficients of its singular expansion at a given value of $y = 1$. This computation is enclosed in the following lemma.

Lemma 3.6.6. *We have that the coefficients on the singular expansion of $S_3^{\triangle}(x, 1)$ in a domain dented at $x = R_{\triangle}(1) \approx 0.19635$ are equal to:*

$$\begin{aligned} a_0^{\triangle}(1) = a_0^{\triangle} \approx 0.15545, \quad a_1^{\triangle}(1) = a_1^{\triangle} \approx -0.34792, \\ a_2^{\triangle}(1) = a_2^{\triangle} \approx 0.27799, \quad a_3^{\triangle}(1) = a_3^{\triangle} \approx -0.16276. \end{aligned} \tag{3.18}$$

In particular, $a_i^{\triangle}(y) \neq 0$ for $i \in \{0, \ldots, 3\}$ and y in a small neighbourhood of 1.

Proof. We just apply undeterminate coefficients over Equation (3.16). By continuity of the functions $a_i^{\triangle}(y)$, the final statement holds as well. □

By using this singular expansion for $S_3^{\triangle}(x, 1)$ we can obtain the corresponding coefficients of the singular expansion of the rest of the networks counting formulas:

Lemma 3.6.7. *The generating functions $P_0^{\triangle}$, $P_1^{\triangle}$, $S_2^{\triangle}$ and $D^{\triangle}$ have the following singular expansions in a domain dented at $x = R_{\triangle}(y)$:*

$$\begin{aligned} P_0^{\triangle} &= p_0^{\triangle}(y_0) + p_1^{\triangle}(y_0)X + p_2^{\triangle}(y_0)X^2 + p_3^{\triangle}(y_0)X^3 + O(X^4) \\ P_1^{\triangle} &= q_0^{\triangle}(y_0) + q_1^{\triangle}(y_0)X + q_2^{\triangle}(y_0)X^2 + q_3^{\triangle}(y_0)X^3 + O(X^4) \\ S_2^{\triangle} &= s_0^{\triangle}(y_0) + s_1^{\triangle}(y_0)X + s_2^{\triangle}(y_0)X^2 + s_3^{\triangle}(y_0)X^3 + O(X^4) \\ D^{\triangle} &= d_0^{\triangle}(y_0) + d_1^{\triangle}(y_0)X + d_2^{\triangle}(y_0)X^2 + d_3^{\triangle}(y_0)X^3 + O(X^4), \end{aligned}$$

where $X = \sqrt{1 - x/R_{\triangle}(y)}$. In particular, when $y = 1$ and $R_{\triangle}(1) \approx 0.19635$ we have that

$$p_0^{\triangle}(1) \approx 0.10374,\ p_1^{\triangle}(1) \approx -0.28169,\ p_2^{\triangle}(1) \approx 0.33606,\ p_3^{\triangle}(1) \approx -0.31761,$$

$$q_0^{\triangle}(1) \approx 1.16818,\ q_1^{\triangle}(1) \approx -0.40643,\ q_2^{\triangle}(1) \approx 0.39544,\ q_3^{\triangle}(1) \approx -0.31132,$$

$$d_0^{\triangle}(1) \approx 1.69532,\ d_1^{\triangle}(1) \approx -1.22249,\ d_2^{\triangle}(1) \approx 0.95538,\ d_3^{\triangle}(1) \approx -0.81117,$$

$$s_0^{\triangle}(1) \approx 0.26795,\ s_1^{\triangle}(1) \approx -0.18645,\ s_2^{\triangle}(1) \approx -0.05411,\ s_3^{\triangle}(1) \approx -0.01948,$$

Additionally, all the terms are different to 0 when y belongs to a small neighbourhood of 1.

Proof. Observe that writing $u = 0$ in the system of Equations (3.13), $P_0^{\triangle}$, $P_1^{\triangle}$, $S_2^{\triangle}$ and $D^{\triangle}$ can be expressed explicitly in terms of $S_3^{\triangle}$. Hence we can use the coefficients of the expansion of $S_3^{\triangle}$ obtained in (3.18) to compute the expansion for the all other functions. The last statement follows from continuity and the fact that the computations give coefficients different from 0. □

We can go now directly to get the singular expansions for $B^{\triangle}(x,y)$:

Theorem 3.6.8. *The generating function $B^{\triangle}(x,y)$ of triangle-free series-parallel networks has the following singular expansion in a domain dented at $R_{\triangle}$ of the form*

$$B^{\triangle} = b_0^{\triangle}(y) + b_2^{\triangle}(y)X^2 + b_3^{\triangle}(y)X^3 + O(X^4),$$

where $X = \sqrt{1 - x/R_{\triangle}(y)}$. In particular, when $y = 1$ and $R_{\triangle}(1) \approx 0.19635$ we have that

$$b_0^{\triangle}(1) \approx 0.01964, \qquad b_2^{\triangle}(1) \approx -0.04123, \qquad b_3^{\triangle}(1) \approx 0.00359.$$

Moreover, $b_0^{\triangle}(y)$, $b_2^{\triangle}(y)$ and $b_3^{\triangle}(y)$ are different to 0 for y close enough to 1, and the term $b_1^{\triangle}(y)$ of X is 0 for any y close enough to 1.

Proof. Replacing $P_0^{\triangle}$, $P_1^{\triangle}$, $S_2^{\triangle}$, $S_3^{\triangle}$ with their singular expansion in the Equation (3.14) gives directly square-root expansions for $B^{\triangle}$. Note that $B^{\triangle}$ can also be obtained from $P_1^{\triangle}$ by the equation

$$2y\frac{\partial B^{\triangle}}{\partial y} = x^2 P_1^{\triangle}.$$

This is true because by our encoding the only networks which contains the root edge are the ones considered in $P_1^{\triangle}$. This implies that the singular expansion of $B^{\triangle}$ must start at X^3, so that after differentiating it we get the singular expansion of $P_1^{\triangle}$. □

We can now apply the Transfer Theorem for singularity analysis [51] in order to get the first asymptotic counting formulas:

Theorem 3.6.9. *The number $b_n^{\triangle}$ of 2-connected triangle-free series-parallel graphs with n vertices is asymptotically equal to*

$$b^{\triangle} \cdot n^{-5/2} \cdot R_{\triangle}^{-n} \cdot n!\,(1 + o(1)),$$

where $b^{\triangle} \approx 0.00152$ and $R_{\triangle}^{-1} \approx 5.09289$.

Proof. Apply the Transfer Theorem to the singular expansion. □

Now we can move to the connected level. In this case, the solution of the equation $\tau B^{\bullet\prime}(\tau) = 1$ is located at $\tau = 0.19631$, and hence the singularity of $C^{\bullet}(x)$ is located at $\rho = 0.19403$. We can then state the final enumerative theorem in this subsection:

Theorem 3.6.10. *The number of connected and general triangle-free series-parallel graphs with n vertices, $c_n^{\triangle}$ and $g_n^{\triangle}$, is asymptotically equal to*

$$c^{\triangle} \cdot n^{-5/2} \cdot \rho_{\triangle}^{-n} \cdot n!\,(1+o(1)), \quad g^{\triangle} \cdot n^{-5/2} \cdot \rho_{\triangle}^{-n} \cdot n!\,(1+o(1)),$$

where $c^{\triangle} \approx 0.00473$, $g^{\triangle} \approx 0.00563$ *and* $\rho_{\triangle}^{-1} \approx 6.28155$.

Proof. This is an straightforward computation. Due to the subcritical scheme, the singularity of both $C^{\triangle}(x,1)$ and $G^{\triangle}(x,1)$ arise from a branch point. The solution to the equation $xB^{\circ\prime}(x) = 1$ is given by $\tau_{\triangle} = 0.19629$. Such value gives that $C(x)$ ceases to be analytic at $x = \rho_{\triangle} = 0.15920$ .We apply then the Transfer Theorem to the resulting singular expansion, joint with the expressions of the coefficients of the singular expansions that were obtained in [28, Proposition 3.10.]. □

As a direct consequence of these computations, the probability that triangle-free series-parallel graph of order n, chosen uniformly at random, is connected is equal to $\exp(-C_0) \approx 0.83962$ (see [28, Theorem 4.6.]).

These enumerative results complement previous ones concerning series-parallel graphs with certain obstructions. In Table 3.2, the constant growth for (connected) series-parallel graphs, triangle-free series-parallel graphs and bipartite series-parallel graphs is shown. The constant for the full family was obtained in [24], while the asymptotic enumeration for bipartite series-parallel graphs can be found in [69].

Family	Constant growth
Series-Parallel	9.07359
Triangle-free Series-Parallel	6.28155
Bipartite Series-Parallel	5.30386

Table 3.2: Constants growth for series-parallel graphs, and for subfamilies (triangle-free and bipartite).

Observe that the asymptotics for triangle-free graphs and bipartite graphs are different. This fact contrasts with the picture that emerges in the general graph setting: as it was proven by Erdős, Kleitman and Rothschild in [9], the number of triangle-free graphs with n vertices is asymptotically equal to the number of bipartite graphs with n vertices. The main difference in our setting is that series-parallel graphs are very sparse, and hence an Erdős-Kleitman-Rothschild type result does not hold in this setting.

3.6.2 4-cycles

For the sake of conciseness and in order to show a new set of equations we analyze the statistics of 4-cycles C_4 in 2-connected series-parallel graphs. We proceed as we did in the previous section: we get first the equations defining networks, and then we build the counting formulas of 2-connected series-parallel graphs. We do not deal with the connected and general setting, because we are in the subcritical case and the procedure will be very similar to the case of triangles.

The combinatorial ideas to get the generating functions for networks (encoding now the number of cycles of length 4) are similar to the ones used before. We denote by $S_2^{\blacksquare}$, $S_3^{\blacksquare}$ and $S_\infty^{\blacksquare}$ series networks where the poles are at distance 2, 3 and more than 3, respectively. Observe that the first two networks could contribute to the creation of 4-cycles (by means of parallel operations), while the term $S_\infty^{\blacksquare}$ cannot. Similarly, we define $P_1^{\blacksquare}$, $P_2^{\blacksquare}$ and $P_\infty^{\blacksquare}$ parallel networks where the distance of the poles is equal to 1, 2 or more than 2. In particular the single edge is encoded in P_1. The total counting formula for networks is encoded by $D^{\blacksquare}$.

Note that there is a difference with respect to triangles. A series network where the poles are at distance 2 has a unique path of length 2 between the poles. This is not true for the case of paths of length 3: both $S_2^{\blacksquare}$ and $S_3^{\blacksquare}$ may have an arbitrary number of paths of length 3, and each of those will form a 4-cycle if we put it in parallel with an edge. This is why we need two additional functions, $\overline{S_2^{\blacksquare}}$ and $\overline{S_3^{\blacksquare}}$, that count series networks that will be put in parallel with an edge, and therefore each path of length 3 will contribute with a new 4-cycle, in addition to other 4-cycles already contained in the network either touching or not touching the poles. In other words, in $\overline{S_2^{\blacksquare}}$ and $\overline{S_3^{\blacksquare}}$ the variable u counts both paths of length 3 between the poles and 4-cycles.

In order to count paths of length 3 we need to note that some of them come from paths of length 2 in the parallel networks that we put in series. Therefore, we will denote as $\overline{P_1^{\blacksquare}}$ and $\overline{P_2^{\blacksquare}}$ the parallel networks where u counts both 4-cycles and paths of length 2 between the poles. These paths in turn come from series networks, so we use again the property that a series network whose poles are at distance 2 has a unique path of length 2 between the poles. This gives the following main equation

$$D^{\blacksquare} = S_2^{\blacksquare} + S_3^{\blacksquare} + S_\infty^{\blacksquare} + P_1^{\blacksquare} + P_2^{\blacksquare} + P_\infty^{\blacksquare}$$

with

$$\begin{aligned}
S_2^{\blacksquare} &= x(P_1^{\blacksquare})^2,\\
\overline{S_2^{\blacksquare}} &= x\left(\overline{P_1^{\blacksquare}}\right)^2,\\
S_3^{\blacksquare} &= x(P_1^{\blacksquare}(P_2^{\blacksquare} + S_2^{\blacksquare}) + P_2^{\blacksquare}P_1^{\blacksquare}),\\
\overline{S_3^{\blacksquare}} &= x(P_1^{\blacksquare}(\overline{P_2^{\blacksquare}} + uS_2^{\blacksquare}) + \overline{P_2^{\blacksquare}}P_1^{\blacksquare}),\\
S_\infty^{\blacksquare} &= x(P_1^{\blacksquare}(P_\infty^{\blacksquare} + S_3^{\blacksquare} + S_\infty^{\blacksquare}) + P_2^{\blacksquare}(S_2^{\blacksquare} + S_3^{\blacksquare} + S_\infty^{\blacksquare} + P_2^{\blacksquare} + P_\infty^{\blacksquare}) + P_\infty^{\blacksquare}D^{\blacksquare}),
\end{aligned}$$

concerning series networks and

$$
\begin{aligned}
P_1^{\blacksquare} &= y\left(\exp(\overline{S_3^{\blacksquare}}+S_\infty^{\blacksquare})\sum_{k\geq 0}u^{\binom{k}{2}}\frac{\left(\overline{S_2^{\blacksquare}}\right)^k}{k!}\right),\\
\overline{P_1^{\blacksquare}} &= y\left(\exp(\overline{S_3^{\blacksquare}}+S_\infty^{\blacksquare})\sum_{k\geq 0}u^{\binom{k+1}{2}}\frac{\left(\overline{S_2^{\blacksquare}}\right)^k}{k!}\right),\\
P_2^{\blacksquare} &= S_2^{\blacksquare}\exp_{\geq 1}(S_3^{\blacksquare}+S_\infty^{\blacksquare})+\exp(S_3^{\blacksquare}+S_\infty^{\blacksquare})\sum_{k\geq 2}u^{\binom{k}{2}}\frac{(S_2^{\blacksquare})^k}{k!},\\
\overline{P_2^{\blacksquare}} &= uS_2^{\blacksquare}\exp_{\geq 1}(S_3^{\blacksquare}+S_\infty^{\blacksquare})+\exp(S_3^{\blacksquare}+S_\infty^{\blacksquare})\sum_{k\geq 2}u^{\binom{k+1}{2}}\frac{(S_2^{\blacksquare})^k}{k!},\\
P_\infty^{\blacksquare} &= \exp_{\geq 2}(S_3^{\blacksquare}+S_\infty^{\blacksquare}).
\end{aligned}
$$

concerning parallel networks. The equations for series networks are obtained by fixing the network type which is incident with the 0-pole (which must be of parallel type), and the network incident with the ∞-pole. In particular, the indices must sum the corresponding index in S. The equations for parallel networks are more involved: in this case sets of networks of type $S_2^{\blacksquare}$ can create a quadratic number of copies of C_4, hence the infinite sums with quadratic exponents in u.

Starting from these equations, we can go to deduce counting formulas for 2-connected objects. In order to apply the dissymmetry theorem we need to obtain the corresponding $B_R^{\blacksquare}$, $B_M^{\blacksquare}$ and $B_{RM}^{\blacksquare}$ as follows:

$$
\begin{aligned}
B_R^{\blacksquare} =& \mathrm{Cyc}(x(P_1^{\blacksquare}+P_2^{\blacksquare}+P_\infty^{\blacksquare}))+\frac{x^3}{6}\left(\left(\overline{P_1^{\blacksquare}}\right)^3-(P_1^{\blacksquare})^3+3(P_1^{\blacksquare})^2\overline{P_2^{\blacksquare}}-3\left(P_1^{\blacksquare}\right)^2P_2^{\blacksquare}\right)+\\
&(u-1)\frac{(xP_1^{\blacksquare})^4}{8}\\
B_M^{\blacksquare} =& \frac{x^2}{2}\left(P_1^{\blacksquare}+P_2^{\blacksquare}+P_\infty^{\blacksquare}\right)-\\
&\frac{x^2}{2}\left(y+y(\overline{S_2^{\blacksquare}}+\overline{S_3^{\blacksquare}}+S_\infty^{\blacksquare})+\frac{u(S_2^{\blacksquare})^2}{2}+S_2^{\blacksquare}(S_3^{\blacksquare}+S_\infty^{\blacksquare})+\frac{(S_3^{\blacksquare}+S_\infty^{\blacksquare})^2}{2}\right)\\
B_{RM}^{\blacksquare} =& \frac{x^2}{2}(\overline{S_2^{\blacksquare}}(\overline{P_1^{\blacksquare}}-y)+S_2(\overline{P_2^{\blacksquare}}+P_\infty^{\blacksquare})+(\overline{S_3^{\blacksquare}}+S_\infty^{\blacksquare})(P_1^{\blacksquare}-y)+(S_3^{\blacksquare}+S_\infty^{\blacksquare})(P_2^{\blacksquare}+P_\infty^{\blacksquare}))\\
B^{\blacksquare} =& \frac{1}{2}x^2y+B_R^{\blacksquare}+B_M^{\blacksquare}-B_{RM}^{\blacksquare},
\end{aligned}
$$

where $B_R^{\blacksquare}$ represents 2-connected series-parallel graphs rooted at a ring, $B_M^{\blacksquare}$ represents 2-connected series-parallel graphs rooted at a multiedge, and $B_{RM}^{\blacksquare}$ represents 2-connected series-parallel graphs rooted at a ring and a multiedge that are adjacent at the decomposition tree. In the case of $B_R^{\blacksquare}$ we have to deal with several special cases, since if the length of the ring is 3 or 4, then 4-cycles might appear. If the length is 4, a single cycle appears. If the length is 3, many cycles might appear if we replace at least two edges of the ring with a parallel network of the kind $P_1^{\blacksquare}$: in particular, any path of length 2 in the other parallel network will produce a 4-cycle. In the case of $B_M^{\blacksquare}$, the parallel networks already count all the 4-cycles, but since the number of edges must be at least three, we have to remove the cases with one and two series networks or edges in parallel. In the case of $B_{RM}^{\blacksquare}$ we have to consider the special cases depending on the length of the ring and whether there is an edge between the poles of the multiedge. If the length of the ring is three, then any path of length two in the parallel network will produce a 4-cycle, whereas if there is an edge between the poles of the multiedge, then any path of length 3 in the series network will produce a 4-cycle, as it is shown in Figure 3.7.

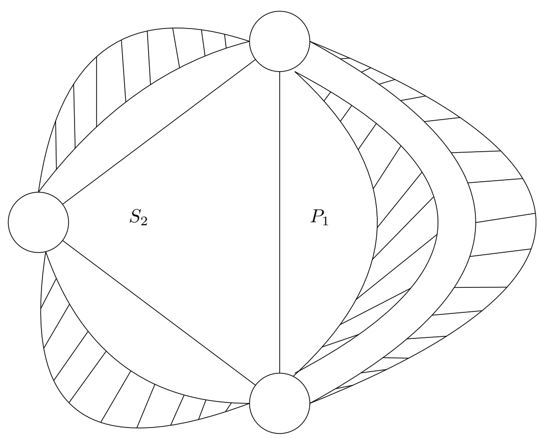

Figure 3.7: 2-connected series-parallel graph rooted in a ring of length 3 and a multiedge with an edge between the poles. Each path of length 3 in $S_2^{\blacksquare}$ and each path of length 2 in $P_1^{\blacksquare}$ will produce a new 4-cycle.

Using these expression and setting $y = 1$, $u = 0$ we obtain the radius of convergence and the singularity analysis of $B^{\blacksquare}(x, 1, 0)$, which, by means of the Transfer Theorem, gives the asymptotic enumeration of 2-connected series-parallel-graphs without 4-cycles.

Theorem 3.6.11. *The number of 2-connected quadrangle-free series-parallel graphs with n vertices ($b_n^{\square}$) is asymptotically equal to*

$$b^{\square} \cdot n^{-5/2} \cdot R_{\square}^{-n} \cdot n!\,(1 + o(1)),$$

where $b^{\square} \approx 0.00145$ *and* $R_{\square}^{-1} \approx 5.13738$.

The proof of the next result is analogous to the proof of triangle-free series-parallel graphs.

Theorem 3.6.12. *The number of connected and general quadrangle-free series-parallel graphs with n vertices, $c_n^{\square}$ and $g_n^{\square}$, is asymptotically equal to*

$$c^{\square} \cdot n^{-5/2} \cdot \rho_{\square}^{-n} \cdot n!\,(1+o(1)), \quad g^{\square} \cdot n^{-5/2} \cdot \rho_{\square}^{-n} \cdot n!\,(1+o(1)),$$

where $c^{\square} \approx 0.00233$, $g^{\square} \approx 0.00276$ and $\rho_{\square}^{-1} \approx 6.41498$.

We use Remark 3.3.4 to obtain the following result about 4-cycles. It is a modification of [47, Theorem 2.35], which provides a way to compute the expectation and variance of generating functions that satisfy the following system of equations:

$$\boldsymbol{y} = \boldsymbol{F}(x, \boldsymbol{y}, \boldsymbol{u}), \tag{3.19}$$

$$0 = \det(\boldsymbol{I} - \boldsymbol{F}_{\boldsymbol{y}}(x, \boldsymbol{y}, \boldsymbol{u})). \tag{3.20}$$

According to that theorem, the expectation $\boldsymbol{\mu}$ and the variance $\boldsymbol{\Sigma}$ of the parameters $\boldsymbol{u}$ can be computed as

$$\boldsymbol{\mu} = -\frac{x_{0,\boldsymbol{u}}(1)}{x_0(1)},$$

$$\boldsymbol{\Sigma} = -\frac{x_{0,\boldsymbol{uu}}(1)}{x_0(1)} + \boldsymbol{\mu}\boldsymbol{\mu}^T + \mathrm{diag}(\boldsymbol{\mu}),$$

where $x = x_0(\boldsymbol{u})$ and $\boldsymbol{y} = \boldsymbol{y}_0(\boldsymbol{u})$ are the solutions of the system (3.19) and (3.20). After half an hour of execution time in Maple we get the following theorem:

Theorem 3.6.13. *The number of quadrangles $W_n^{\blacksquare}$ on a uniformly at random 2-connected series-parallel graph on n vertices is asymptotically Gaussian, with*

$$\mathbb{E}[W_n^{\blacksquare}] = \mu_{\blacksquare,2} n + O(1), \qquad \mathbb{V}\mathrm{ar}[W_n^{\blacksquare}] = \sigma_{\blacksquare,2}^2 n + O(1),$$

where $\mu_{\blacksquare,2} \approx 0.51235$ and $\sigma_{\blacksquare,2}^2 \approx 0.25418$.

3.6.3 Girth

Now we can generalize the previous results to obtain the generating function of series-parallel graphs with girth at least k, for $k \geq 4$. We need new notation for the series and parallel networks. In particular, in order to express the generating function of networks with girth k we define S_i, for $2 \leq i \leq k-2$ as the generating function of series networks with girth $\geq k$ and where the distance between the poles is exactly i. The generating function of series networks with girth $\geq k$ and distance to the poles $\geq k-1$ is expressed as S_∞. Analogously, we define P_i, for $1 \leq i \leq k-3$ and P_∞ as the generating function of parallel networks with the same condition on the distance between the poles. In this case the generating functions satisfy the following system of equations:

$$P_i = \begin{cases} y\exp(S_\infty) & \text{if} \quad i = 1 \\ S_i \exp_{\geq 1}(\sum_{j \geq k-i} S_j) & \text{if} \quad 1 < i < k/2 \\ S_i \exp_{\geq 1}(\sum_{j \geq i+1} S_j) + \exp_{\geq 2}(S_i)\exp(\sum_{j \geq i+1} S_j) & \text{if} \quad k/2 \leq i \leq k-3 \\ \exp_{\geq 2}(S_{k-2} + S_\infty) & \text{if} \quad i = \infty \end{cases}$$

$$S_i = \begin{cases} x\left[\sum_{j=1}^{i-1} P_j(S_{i-j}+P_{i-j})\right] & \text{if} \quad 1 \le i \le k-2 \\ x\left[\sum_{j\ge 1} P_j \sum_{t\ge k-j-1}(S_t+P_t)\right] & \text{if} \qquad i=\infty \end{cases}$$

Note that for convenience we are considering that S_1 exists, with a value of $S_1 = 0$. This equations generalize the corresponding ones for girth 3. For the case of parallel networks we impose that the the distance between the poles of the two shortest series networks is at most k. This can be done by distinguishing two cases: if the distance between the poles is $i < k/2$ then there must be one single series network with distance i between the poles. This implies that all the other series networks must have distance at least $k-i$ between the poles, because otherwise there would be a cycle of length less than k. If the distance between the poles is $i \ge k/2$, then no cycle of length less than k can be produced, so we just have to be sure that the shortest series network that we put in parallel is of length i. For the case of series networks no further constraint is needed, since no new cycle can be produced.

This gives a way to compute the exponential growth for any possible girth. Since the computations are involved and analogous to the ones of girth 4 we do not include the results.

3.7 Concluding remarks

In this work we have shown normal limiting distributions for the number of copies of a given graph for subcritical graph classes. From our study several challenging questions might be investigated in the future. First, the proof of our main theorem does not give a systematic way to compute both the expectation and the variance of the corresponding random variable (we only get that both of them are linear in n). In Section 3.6 we have exploited extra information concerning the structure of series-parallel graphs in order to get precise constants, but getting a full numerical analysis seems to be very difficult in a more general context. Nevertheless, we can use the Benjamni-Schramm limit given in [40, 31] to get the constant for the mean value (for details see [40]). However, it seems to be very difficult to obtain a general procedure for computing the constant for the variance.

Second, we cannot immediately obtain local limit laws for the number of copies of a given graph. In our analysis we only provided asymptotic information of our generating functions a neighborhood of $u = 1$ (for $|u| = 1$). In order to obtain a Local Limit Theorem we need asymptotic information for all u with $|u| = 1$. This is certainly not out of reach but needs a lot of extra work.

Finally, our techniques do not apply to subgraphs in planar-like families (see [28]). Technically speaking, when analyzing subcritical graph classes we have continuously exploited the assumption that the counting formula for the blocks can be considered to be analytic. Unfortunately, the picture changes dramatically when dealing with planar graphs, as a critical composition scheme arises (see [8, 28]). In this context, very little is known concerning the number of subgraphs in the random planar graph model, even concerning the number of triangles. The only result we know so far is [72, 39], where the authors exploit the fact that triangles in cubic planar graph do not intersect. Using this combinatorial fact, they are able to show normality for the number of triangles in cubic planar graphs. This method does not apply in the general planar setting, as an edge can be incident with many triangles. So new ideas from different sources are needed to attack this problem.

Chapter 4

Counting graphs with given degree sequence

This chapter is based in a paper reported and presented in the Meeting on Analytic Algorithms and Combinatorics 2016. Given a set $\mathcal{D}$ of nonnegative integers, we derive the asymptotic number of graphs with a given number of vertices, edges, and such that the degree of every vertex is in $\mathcal{D}$. This generalizes existing results, such as the enumeration of graphs with a given minimum degree, and establishes new ones, such as the enumeration of Euler graphs, *i.e.* where all vertices have an even degree. Those results are derived using analytic combinatorics.

4.1 Introduction

4.1.1 Related works

The asymptotics of several families of simple graphs with degree constraints have been derived. Regular graphs, where all vertices have the same degree, have been enumerated by [63], graphs with minimum degree at least δ by [10]. An *Euler graph*, or *even graph*, is a graph where all vertices have an even degree. An exact formula for the number of such graphs, for a given number of vertices and without consideration of the number of edges, has been derived by [25] and [73]. In the present work, we generalize those results and derive the asymptotic number of graphs with degrees in any given set.

A similar problem has been addressed with probabilistic tools by the *configuration model*, introduced independently by [3] and [54]. This model inputs a distribution F on the degrees, and outputs a random multigraph where the degree of each vertex follows F. The main difference with the model analyzed in this article is that the number of edges in the configuration model is a random variable. The link between both models is discussed in Section 4.5.1. For more information on the configuration model, we recommend the book of [42].

Other related problems include the enumeration of graphs with a given degree sequence, given in [63], and the enumeration of graphs with degree parities, investigated by [26]. The case of multigraphs can be seen as symmetric matrices with nonnegative coefficients. The enumeration where the sum of the rows is constant is given in [23], and the regularity of the case where the possible entries of the matrix are fixed is given in [48].

4.1.2 Model and notations

A *multiset* is an unordered collection of objects, where repetitions are allowed. Sets are then multisets without repetitions. A *sequence* is an ordered multiset. We use the parenthesis notation $(u_1, \dots, u_n)$ for sequences, and the brace notation $\{u_1, \dots, u_n\}$ for sets and multisets. Open real intervals are denoted by open square brackets $]a, b[$.

A *simple graph* G is a set $V(G)$ of labelled vertices and a set $E(G)$ of edges, where each edge is an unordered pair of distinct vertices. In a *multigraph*, the edges form a multiset and the vertices in an edge need not be distinct. An edge $\{v, w\}$ is a *loop* if $v = w$, a *multiple edge* if it has at least two occurrences in the multiset of edges, and a *simple edge* otherwise. Thus, the simple graphs are the multigraphs that contain neither loops nor multiple edges, *i.e.* that contain only simple edges. The set of multigraphs with n vertices and m edges is denoted by $\mathrm{MG}_{n,m}$, and the subset of simple graphs by $\mathrm{SG}_{n,m}$.

The *degree* of a vertex is defined as its number of occurrences in $E(G)$. In particular, a loop increases its degree by 2. The set of multigraphs from $\mathrm{MG}_{n,m}$ where each vertex has its degree in a set $\mathcal{D}$ is denoted by $\mathrm{MG}_{n,m}^{(\mathcal{D})}$. The subset of simple graphs is $\mathrm{SG}_{n,m}^{(\mathcal{D})}$. The set $\mathcal{D}$ may be finite or infinite. We denote its generating function by

$$\mathrm{Set}_{\mathcal{D}}(x) = \sum_{d \in \mathcal{D}} \frac{x^d}{d!}.$$

For any natural number i, $\mathcal{D} - i$ denotes the set $\{d - i \in \mathbb{Z}_{\geq 0} \mid d \in \mathcal{D}\}$. In particular, observe that $\mathrm{Set}'_{\mathcal{D}}(x) = \mathrm{Set}_{\mathcal{D}-1}(x)$. If the minimum of $\mathcal{D}$ is smaller than i, then $\mathcal{D} - i$ contains less elements than $\mathcal{D}$. We also define the *valuation* $r = \min(\mathcal{D})$ and *periodicity* $p = \gcd\{d_1 - d_2 \mid d_1, d_2 \in \mathcal{D}\}$ of the set $\mathcal{D}$ (by convention, the periodicity is infinite when $|\mathcal{D}| = 1$).

4.1.3 Main Theorem and applications

Our main result is an asymptotic expression for the number of graphs in $\mathrm{SG}_{n,m}^{(\mathcal{D})}$, when m grows linearly with n.

Theorem 4.1.1. *Assume $\mathcal{D}$ contains at least two integers, has valuation $r = \min\{d \in \mathcal{D}\}$ and periodicity $p = \gcd\{d_1 - d_2 \mid d_1, d_2 \in \mathcal{D}\}$. Let m, n denote two integers tending to infinity, such that $2m/n$ stays in a fixed compact interval of $]r, \max(\mathcal{D})[$ and p divides $2m - rn$, then the number of simple graphs in $\mathrm{SG}_{n,m}^{(\mathcal{D})}$ is*

$$\frac{(2m)!}{2^m m!} \frac{\mathrm{Set}_{\mathcal{D}}(\zeta)^n}{\zeta^{2m}} \frac{p}{\sqrt{2\pi n \zeta \phi'(\zeta)}} e^{-W_{\frac{n}{m}}(\zeta)^2 - W_{\frac{n}{m}}(\zeta)} \times (1 + O(n^{-1})),$$

where $\phi(x) = \frac{x \operatorname{Set}_{\mathcal{D}-1}(x)}{\operatorname{Set}_{\mathcal{D}}(x)}$*,* ζ *is the unique positive solution of* $\phi(\zeta) = \frac{2m}{n}$*, and* $W_{\frac{n}{m}}(x) = \frac{n}{4m} \frac{x^2 \operatorname{Set}_{\mathcal{D}-2}(x)}{\operatorname{Set}_{\mathcal{D}}(x)}$*. If* p *does not divide* $2m - rn$*, if* $2m/n < r$ *or if* $2m/n > \max(\mathcal{D})$*, then* $\mathrm{SG}_{n,m}^{(\mathcal{D})}$ *is empty.*

When $\mathcal{D} = \mathbb{Z}_{\geq 0}$, the degrees are not constrained, so $\mathrm{SG}_{n,m}^{(\mathcal{D})} = \mathrm{SG}_{n,m}$. Using Stirling formula, it can indeed be checked that $\binom{\binom{n}{2}}{m}$, the total number of simple graphs with n vertices and m edges, is asymptotically equal to the result of Theorem 4.1.1

$$\frac{n^{2m}}{2^m m!} \frac{(2m)!}{(2m)^{2m} e^{-2m} \sqrt{2\pi 2m}} e^{-\left(\frac{m}{n}\right)^2 - \frac{m}{n}} \left(1 + O(n^{-1})\right).$$

[10] have derived the asymptotics of simple graphs with minimum degree at least δ. They used probabilitic and analytic elementary tools, in a sophisticated way. In the present chapter, we have addressed the enumeration of a broader family of graphs with degree constraints, using more powerful tools (analytic combinatorics). For graphs with minimum degree at least δ, the asymptotics derived in Theorem 4.1.1, for $\mathcal{D} = \mathbb{Z}_{\geq \delta}$, matches their result.

Euler graphs are simple graphs where each vertex has an even degree. An exact, but complicated, formula for the number of such graphs, for given number of vertices and without consideration of the number of edges, has been derived by [25] and [73]. Applying Theorem 4.1.1, we are now able to derive the asymptotic number of Euler graphs with n vertices and m edges, when $2m/n$ stays in a fixed compact interval of $\mathbb{R}_{>0}$

$$\begin{aligned} |\mathrm{SG}_{n,m}^{(\text{even})}| = & \frac{(2m)!}{2^m m!} \frac{\cosh(\zeta)^n}{\zeta^{2m}} \frac{2}{\sqrt{2\pi n \zeta \phi'(\zeta)}} \\ & \times e^{-\left(\frac{n}{4m}\zeta^2\right)^2 - \frac{n}{4m}\zeta^2} (1 + O(n^{-1})), \end{aligned}$$

where $\phi(x) = x \tanh(x)$ and $\tanh(\zeta) = 2m/n$.

In the rest of the chapter we provide a proof for Theorem 4.1.1.

4.2 Analytic preliminaries

Our tool for the analysis of graphs with degree constraints is *analytic combinatorics*, as presented by [5]. Its principle is to associate to the combinatorial family studied its *generating function*. The asymptotics of the family is then linked to the analytic behavior of this function.

In the analysis of a graphs family $\mathcal{F}$ with analytic combinatorics, the main difficulty is the fast growth of its cardinality, which often implies a zero radius of convergence for the corresponding generating function

$$\sum_{G \in \mathcal{F}} w^{|E(G)|} \frac{z^{|V(G)|}}{|V(G)|!}.$$

This feature drastically reduces the number of tools from complex analysis that can be applied. Graphs with degree constraints are no exception, but our approach completely avoid this classic issue. In fact, the only analytic tool we use is the following lemma, a variant of [5, Theorem VIII.8].

Lemma 4.2.1. *Consider a non-monomial series $B(z)$ with nonnegative coefficients, analytic on $\mathbb{C}$, with valuation $r = \min\{n \mid [z^n]B(z) \neq 0\}$ and periodicity $p = \gcd\{n \mid [z^{n-r}]B(z) \neq 0\}$. Let $\phi(z)$ denote the function $\frac{zB'(z)}{B(z)}$, and K a compact interval of the open interval $]r, \lim_{x\to\infty} \phi(x)[$. Let N, n denote two integers tending to infinity while N/n stays in K, and let ζ denote the unique positive solution of $\phi(\zeta) = N/n$. Finally, consider a compact Y and a function $A(y,z)$, $\mathcal{C}^2$ on $Y \times \mathbb{C}$, such that for all y in Y, the function $z \mapsto A(y,z)$ is analytic on $\mathbb{C}$ and $A(y,\zeta^p)$ is nonzero. Then we have, uniformly for N/n in K and y in Y,*

$$[z^N]A(y,z^p)B(z)^n = \begin{cases} \frac{pA(y,\zeta^p)}{\sqrt{2\pi n\zeta\phi'(\zeta)}}\frac{B(\zeta)^n}{\zeta^N}\left(1+O(n^{-1})\right) \\ \quad \text{if } p \text{ divides } N - nr, \\ 0 \text{ otherwise.} \end{cases}$$

4.3 Multigraphs with degree constraints

The work of [68] and [64] demonstrates that multigraphs are more suitable to the analytic combinatorics approach than simple graphs. Moreover, the results on multigraphs can usually be extended to simple graphs. Following this observation, multigraphs are analyzed in this section, before turning so simple graphs in Section 4.4.

4.3.1 Preliminaries

The main model of random multigraphs with n vertices and m edges is the *multigraph process*, analyzed by [68] and [64]. It samples uniformly and independently $2m$ vertices $(v_1, v_2, \dots, v_{2m})$ in $\{1,\dots,n\}$, and outputs a multigraph with set of vertices $\{1,\dots,n\}$ and set of edges

$$\{\{v_{2i-1}, v_{2i}\} \mid 1 \leq i \leq m\}.$$

Given a simple or multi graph, one can order the set of edges and the vertices in each edge. The result is a sequence of ordered pairs of vertices, that we call an *ordering* of G. Let $\mathsf{orderings}(G)$ denote the number of such orderings. For example, the multigraph on 2 vertices with edges $\{\{1,1\},\{1,2\},\{1,2\}\}$ has 12 orderings, amongst them $((1,2),(1,1),(2,1))$. For simple graphs, the number of orderings is equal to $2^m m!$, because each edge has two possible orientations and all edges can be permuted. For non-simple multigraphs, $\mathsf{orderings}$ is smaller. [68] and [64] introduced the *compensation factor* $\kappa(G)$ of a multigraph G with m edges, defined as

$$\kappa(G) = \frac{\mathsf{orderings}(G)}{2^m m!}.$$

The compensation factor of a multigraph is 1 if and only if it is simple.

Observe that in the random distribution induced by the multigraph process, each multigraph receives a probability proportional to its compensation factor. Therefore, when the output of the multigraph process is constrained to be a simple graph, the sampling becomes uniform on $\mathrm{SG}_{n,m}$. The *total weight* of a family $\mathcal{F}$ of multigraphs is the sum of their compensation factors. For example, the total weight of $\mathrm{MG}_{n,m}$ is equal to $\frac{n^{2m}}{2^m m!}$. When $\mathcal{F}$ contains only simple graphs, its total weight is equal to its cardinality.

4.3.2 Exact and asymptotic enumeration

We derive an exact expression for the number of multigraphs with degree constraints in Theorem 4.3.1, then translates it into an asymptotics in Theorem 4.3.2.

Theorem 4.3.1. *The total weight of all multigraphs in* $\mathrm{MG}_{n,m}^{(\mathcal{D})}$ *is*

$$\sum_{G \in \mathrm{MG}_{n,m}^{(\mathcal{D})}} \kappa(G) = \frac{(2m)!}{2^m m!} [x^{2m}] \operatorname{Set}_{\mathcal{D}}(x)^n.$$

Proof. By definition of the compensation factor, the number of multigraphs of the theorem is equal to

$$\frac{1}{2^m m!} \sum_{G \in \mathrm{MG}_{n,m}^{(\mathcal{D})}} \mathsf{orderings}(G).$$

Let us consider an ordering

$$((v_1, v_2), (v_2, v_3), \dots, (v_{2m-1}, v_{2m})).$$

of a multigraph G from $\mathrm{MG}_{n,m}^{(\mathcal{D})}$. For all $1 \leq i \leq n$, let $P_i = \{j \mid v_j = i\}$ denote the set of positions of the vertex i in this ordering. Since the vertices have their degrees in $\mathcal{D}$, each P_i has size in $\mathcal{D}$. This implies a bijection between

- the orderings of multigraphs in $\mathrm{MG}_{n,m}^{(\mathcal{D})}$,
- the sequences of sets $(P_1, \dots, P_n)$, where the size of each set is in $\mathcal{D}$, and $(P_1, \dots, P_n)$ is a partition of $\{1, \dots, 2m\}$ (*i.e.* the sets are disjoints and $\bigcup_{i=1}^n P_i = \{1, \dots, 2m\}$).

We now interpret $(P_1, \dots, P_n)$ as a sequence of sets that contain labelled objects and apply the *Symbolic Method* (see [5]). The exponential generating function of sets of size in $\mathcal{D}$ is $\operatorname{Set}_{\mathcal{D}}(x)$. The bijection then implies

$$\sum_{G \in \mathrm{MG}_{n,m}^{(\mathcal{D})}} \mathsf{orderings}(G) = (2m)! [x^{2m}] \operatorname{Set}_{\mathcal{D}}(x)^n,$$

and the theorem follows, after division by $2^m m!$. □

Now applying Lemma 4.2.1 to the exact expression, we derive the asymptotics of multigraphs with degree constraints. Let us first eliminate three simple cases.

- When $\mathcal{D}$ contains only one integer $\mathcal{D} = \{d\}$, $\mathrm{MG}_{n,m}^{(\mathcal{D})}$ is the set of d-regular multigraphs. The total weight of $\mathrm{MG}_{n,m}^{(\mathcal{D})}$ is then 0 if $2m \neq nd$, and $\frac{(2m)!}{2^m m! d!^n}$ otherwise.
- The sum of the degrees of the vertices is equal to $2m$, so $\mathrm{MG}_{n,m}^{(\mathcal{D})}$ is empty when $2m/n < \min(\mathcal{D})$ or $2m/n > \max(\mathcal{D})$.
- The periodicity p of $\mathcal{D}$ is equal to $\gcd\{d - r \mid d \in \mathcal{D}\}$. For each vertex v of a multigraph from $\mathrm{MG}_{n,m}^{(\mathcal{D})}$, it follows that p divides $\deg(v) - r$. By summation over all vertices, we conclude that if p does not divide $2m - nr$, then the set $\mathrm{MG}_{n,m}^{(\mathcal{D})}$ is empty.

The two last points obviously hold for $\mathrm{SG}_{n,m}^{(\mathcal{D})}$.

Theorem 4.3.2. *Consider a set $\mathcal{D} \subset \mathbb{Z}_{\geq 0}$ of size at least 2. Let $r = \min(\mathcal{D})$ denote its valuation and $p = \gcd\{d_1 - d_2 \mid d_1, d_2 \in \mathcal{D}\}$ its periodicity. Let m, n denote two integers tending to infinity, such that $2m/n$ stays in a fixed compact interval of the open interval $]r, \max(\mathcal{D})[$, and p divides $2m - rn$, then the total weight of $\mathrm{MG}_{n,m}^{(\mathcal{D})}$ is equal to*

$$\sum_{G \in \mathrm{MG}_{n,m}^{(\mathcal{D})}} \kappa(G) = \frac{(2m)!}{2^m m!} \frac{p}{\sqrt{2\pi n \zeta \phi'(\zeta)}} \frac{\mathrm{Set}_{\mathcal{D}}(\zeta)^n}{\zeta^{2m}} \times \left(1 + O(n^{-1})\right)$$

where $\phi(x) = \frac{x\,\mathrm{Set}_{\mathcal{D}-1}(x)}{\mathrm{Set}_{\mathcal{D}}(x)}$ and ζ is the unique positive solution of $\phi(\zeta) = \frac{2m}{n}$.

4.3.3 Typical multigraphs with degree constraints

Let us recall that an edge is *simple* if it is neither a loop nor a multiple edge. Before turning to the enumeration of simple graphs with degree constraints, we first describe the behavior of non-simple edges in a typical multigraph from $\mathrm{MG}_{n,m}^{(\mathcal{D})}$. No proofs are given here, as stronger results will be derived later.

Using random sampling, we observe that in most of the multigraphs from $\mathrm{MG}_{n,m}^{(\mathcal{D})}$, all non-simple edges have low multiplicity and are well separated. This motivates the following definition. A multigraph from $\mathrm{MG}_{n,m}^{(\mathcal{D})}$ is in $\mathrm{MG}_{n,m}^{(\mathcal{D},\star)}$ if all its non-simple edges are loops or double edges, and each vertex belongs to at most one loop or (exclusive) one double edge. Let $|E|_e$ denote the number of occurrences of the element e in the multiset E. Formally, $\mathrm{MG}_{n,m}^{(\mathcal{D},\star)}$ is characterized as the set of multigraphs G from $\mathrm{MG}_{n,m}^{(\mathcal{D})}$ such that for all vertices u, v, w, we have

$$\begin{aligned}
&|E(G)|_{\{v,v\}} \leq 1,\\
&|E(G)|_{\{v,w\}} \leq 2,\\
&|E(G)|_{\{u,v\}} = |E(G)|_{\{v,w\}} = 2 \implies u = w,\\
&\{v,v\} \in E(G) \implies \forall w,\ |E(G)|_{\{v,w\}} \leq 1.
\end{aligned}$$

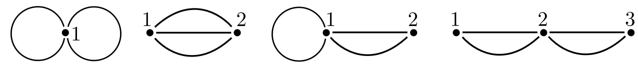

Figure 4.1: Four examples of multigraphs from $\mathrm{MG}_{n,m}^{(\mathcal{D},0)}$.

The complementary set, $\mathrm{MG}_{n,m}^{(\mathcal{D})} \setminus \mathrm{MG}_{n,m}^{(\mathcal{D},*)}$, is denoted by $\mathrm{MG}_{n,m}^{(\mathcal{D},0)}$, and illustrated in Figure 4.1.

4.4 Simple graphs with degree constraints

We introduce the notation $\mathrm{SG}_{n,m}^{(\mathcal{D})}$ for the set of simple graphs with n vertices, m edges and all degrees in $\mathcal{D}$, *i.e.* multigraphs from $\mathrm{MG}_{n,m}^{(\mathcal{D})}$ that contain neither loops nor multiple edges. The enumeration of simple graphs with degree constraints is derived in Theorem 4.1.1. First, in Section 4.4.2, we describe an inclusion-exclusion process that outputs $|\,\mathrm{SG}_{n,m}^{(\mathcal{D})}\,|$ when applied to $\mathrm{MG}_{n,m}^{(\mathcal{D},*)}$. In Section 4.4.3, this process is then applied to $\mathrm{MG}_{n,m}^{(\mathcal{D})}$, and the error introduced is proven to be negligible in Section 4.4.4.

In order to forbid loops and multiple edges in multigraphs from $\mathrm{MG}_{n,m}^{(\mathcal{D})}$, we introduce the notion of *marked multigraphs*.

4.4.1 Marked multigraphs

A *marked multigraph* G is a triplet $(V(G), E(G), \bar{E}(G))$, where $V(G)$ denotes the set of vertices, $E(G)$ the multiset of *normal edges*, and $\bar{E}(G)$ the multiset of *marked edges*, where both normal and marked edges are unordered pairs of vertices. We say that a marked multigraph G belongs to a family $\mathcal{F}$ of (unmarked) multigraphs if the unmarked multigraph $(V(G), E(G) \cup \bar{E}(G))$ is in $\mathcal{F}$.

We now extend to marked multigraphs the definitions of degree, orderings and compensation factors, introduced for multigraphs in Section 4.2. The *degree* of a vertex from a marked multigraph G is equal to its number of occurrences in the multiset $E(G) \cup \bar{E}(G)$. An *ordering* of a marked multigraph G with $m = |E(G)| + |\bar{E}(G)|$ edges is a sequence

$$S = ((v_1, w_1, t_1), \dots, (v_m, w_m, t_m))$$

from $(V(G) \times V(G) \times \{0,1\})^m$ such that the multiset $\{\{v_i, w_i\} \mid (v_i, w_i, 0) \in S\}$ is equal to $E(G)$, and the multiset $\{\{v_i, w_i\} \mid (v_i, w_i, 1) \in S\}$ is equal to $\bar{E}(G)$. The number of orderings of a given marked multigraph G is denoted by $\mathsf{orderings}(G)$, and its *compensation factor* is

$$\kappa(G) = \frac{\mathsf{orderings}(G)}{2^m m!}.$$

For example, consider the marked multigraph G with

$$V(G) = \{1, 2\},$$
$$E(G) = \{\{1, 2\}\},$$
$$\bar{E}(G) = \{\{1, 2\}, \{1, 2\}\}.$$

Its number of orderings is 24, and therefore its compensation factor is $\kappa(G) = 1/2$, whereas it is $1/6$ for G without the marks,

$$V(G) = \{1,2\}, \quad E(G) = \{\{1,2\},\{1,2\},\{1,2\}\}.$$

In the following, we will consider families of marked multigraphs where the marked edges are loops or multiple edges. Given a marked multigraph G, then $\ell(G)$ denotes the number of loops in $\bar{E}(G)$, and $k(G)$ the number of distinct edges from $\bar{E}(G)$ that are not loops. The generating function of a family $\mathcal{F}$ or marked multigraphs is

$$F(u,v) = \sum_{G \in \mathcal{F}} \kappa(G) u^{k(G)} v^{\ell(G)}.$$

4.4.2 Inclusion-exclusion process

In this section, we build an operator Marked that inputs a family of multigraphs and outputs a family of marked multigraphs. It is designed so that the asymptotics of its generating function $\mathrm{Marked}_{\mathrm{MG}_{n,m}^{(\mathcal{D})}}(u,v)$ is linked to the asymptotics of $|\,\mathrm{SG}_{n,m}^{(\mathcal{D})}\,|$. In order to justify the construction, we first introduce the operators $\mathrm{Marked}^{(1)}$ and $\mathrm{Marked}^{(2)}$.

First marking. If we could mark all loops and multiple edges from $\mathrm{MG}_{n,m}^{(\mathcal{D})}$, the enumeration of simple graphs with degree constraints would be easy. Indeed, given a family $\mathcal{F}$ of multigraphs, let $\mathrm{Marked}_{\mathcal{F}}^{(1)}$ denote the marked multigraphs from $\mathcal{F}$ with all loops and multiple edges marked. Since the simple graphs are the multigraphs that have neither loops nor multiple edges, we have

$$\begin{aligned}\mathrm{Marked}^{(1)}_{\mathrm{MG}_{n,m}^{(\mathcal{D})}}(0,0) &= \sum_{G \in \mathrm{MG}_{n,m}^{(\mathcal{D})}} \kappa(G) 0^{k(G)} 0^{\ell(G)} \\ &= \sum_{G \in \mathrm{SG}_{n,m}^{(\mathcal{D})}} \kappa(G),\end{aligned}$$

which is equal to $|\,\mathrm{SG}_{n,m}^{(\mathcal{D})}\,|$, because simple graphs have a compensation factor equal to 1. Unfortunately, we do not have a description of this family in the symbolic method formalism.

Second marking. The inclusion-exclusion principle advises us to mark *some* of the non-simple edges. Let $\mathrm{Marked}_{\mathcal{F}}^{(2)}$ denote the set of marked multigraphs G from $\mathcal{F}$ such that each edge from $\bar{E}(G)$ is either a loop, or has multiplicity at least 2 in $\bar{E}(G)$ and does not belong to $E(G)$. This construction implies the relation

$$\mathrm{Marked}_{\mathcal{F}}^{(2)}(u,v) = \mathrm{Marked}_{\mathcal{F}}^{(1)}(u+1, v+1),$$

and therefore

$$|\,\mathrm{SG}_{n,m}^{(\mathcal{D})}\,| = \mathrm{Marked}^{(2)}_{\mathrm{MG}_{n,m}^{(\mathcal{D})}}(-1,-1).$$

The natural idea to build a marked multigraph G from $\mathrm{Marked}^{(2)}_{\mathrm{MG}^{(\mathcal{D})}_{n,m}}$ is to first choose some loops and multiple edges to put in $\bar{E}(G)$, then complete $E(G)$ with unmarked edges, which may well form other loops and multiple edges, in a way that ensures $G \in \mathrm{Marked}^{(2)}_{\mathrm{MG}^{(\mathcal{D})}_{n,m}}$. However, the description of the set of marked edges is complicated, because of the numerous possible intersection patterns.

Third marking. We have seen in Section 4.3.3 that in most of the multigraphs from $\mathrm{MG}^{(\mathcal{D})}_{n,m}$, non-simple edges do not intersect. This motivates the following definition. Given a set $\mathcal{F}$ of multigraphs, let $\mathrm{Marked}(\mathcal{F})$ denote the set of marked multigraphs from $\mathcal{F}$ such that each vertex is in exactly one of the following cases:

- the vertex belongs to no marked edge,
- the vertex belongs to one marked loop and no other marked edge,
- the vertex belongs to two identical marked edges and no other marked edge.

Therefore, each marked edge is a loop of multiplicity 1 or a double edge. This marking process links the multigraphs from $\mathrm{MG}^{(\mathcal{D},*)}_{n,m}$, defined in Section 4.3.3, to the simple graphs with degree constraints.

Lemma 4.4.1. *The value* $\mathrm{Marked}_{\mathrm{MG}^{(\mathcal{D},*)}_{n,m}}(-1,-1)$ *is equal to the number of simple graphs in* $\mathrm{SG}^{(\mathcal{D})}_{n,m}$.

Proof. As explained in the paragraphs **First markink** and **Second marking** of Section 4.4.2, the following relations hold

$$\mathrm{Marked}^{(1)}_{\mathrm{MG}^{(\mathcal{D},*)}_{n,m}}(0,0) = |\,\mathrm{SG}^{(\mathcal{D})}_{n,m}\,|,$$
$$\mathrm{Marked}^{(2)}_{\mathrm{MG}^{(\mathcal{D},*)}_{n,m}}(u,v) = \mathrm{Marked}^{(1)}_{\mathrm{MG}^{(\mathcal{D},*)}_{n,m}}(u+1,v+1).$$

Furthermore, by construction of $\mathrm{MG}^{(\mathcal{D},*)}_{n,m}$, we have

$$\mathrm{Marked}_{\mathrm{MG}^{(\mathcal{D},*)}_{n,m}}(u,v) = \mathrm{Marked}^{(2)}_{\mathrm{MG}^{(\mathcal{D},*)}_{n,m}}(u,v),$$

so $\mathrm{Marked}_{\mathrm{MG}^{(\mathcal{D},*)}_{n,m}}(-1,-1) = |\,\mathrm{SG}^{(\mathcal{D})}_{n,m}\,|$. □

Applying the operator Marked to the decomposition

$$\mathrm{MG}^{(\mathcal{D})}_{n,m} = \mathrm{MG}^{(\mathcal{D},*)}_{n,m} \uplus \mathrm{MG}^{(\mathcal{D},0)}_{n,m},$$

we find

$$\mathrm{Marked}_{\mathrm{MG}^{(\mathcal{D})}_{n,m}}(u,v) = \mathrm{Marked}_{\mathrm{MG}^{(\mathcal{D},*)}_{n,m}}(u,v) + \mathrm{Marked}_{\mathrm{MG}^{(\mathcal{D},0)}_{n,m}}(u,v)$$

which implies, after evaluation at $(u,v) = (-1,-1)$ and reordering of the terms,

$$\begin{aligned} |\operatorname{SG}^{(\mathcal{D})}_{n,m}| = &\operatorname{Marked}_{\operatorname{MG}^{(\mathcal{D})}_{n,m}}(-1,-1) \\ &- \operatorname{Marked}_{\operatorname{MG}^{(\mathcal{D},0)}_{n,m}}(-1,-1). \end{aligned}$$

We compute the asymptotics of $\operatorname{Marked}_{\operatorname{MG}^{(\mathcal{D})}_{n,m}}(-1,-1)$ in Section 4.4.3, and prove that $\operatorname{Marked}_{\operatorname{MG}^{(\mathcal{D},0)}_{n,m}}(-1,-1)$ is negligible in Section 4.4.4.

4.4.3 Application of the inclusion-exclusion process to all multigraphs with degree constraints

We start with an exact expression of $\operatorname{Marked}_{\operatorname{MG}^{(\mathcal{D})}_{n,m}}(u,v)$ in Lemma 4.4.2, then derive its asymptotics in Lemma 4.4.4.

Lemma 4.4.2. $\operatorname{Marked}_{\operatorname{MG}^{(\mathcal{D})}_{n,m}}(u,v)$ *is equal to*

$$\frac{(2m)!}{2^m m!}[x^{2m}]\left(\sum_{k,\ell\geq 0} a_{n,m,2k+\ell}\frac{(uW_{\frac{n}{m}}(x)^2)^k}{k!}\frac{(vW_{\frac{n}{m}}(x))^\ell}{\ell!}\right) \times \operatorname{Set}_{\mathcal{D}}(x)^n,$$

where $a_{n,m,j} = 0$ *when* j *is greater than* $\min(n,m)$*, otherwise*

$$\begin{aligned} a_{n,m,j} &= \frac{n!}{(n-j)!n^j}\frac{m!}{(m-j)!m^j}\frac{(2m-2j)!(2m)^{2j}}{(2m)!}, \\ W_{\frac{n}{m}}(x) &= \frac{n}{4m}\frac{x^2 \operatorname{Set}_{\mathcal{D}-2}(x)}{\operatorname{Set}_{\mathcal{D}}(x)}. \end{aligned}$$

Proof. To build an ordering of a multigraph from $\operatorname{Marked}_{\operatorname{MG}^{(\mathcal{D})}_{n,m}}$ with $2k$ vertices in marked double edges and ℓ vertices in marked loops, we perform the following steps:

1. choose the labels of the $2k$ vertices that appear in the marked double edges, and the ℓ vertices that appear in the marked loops. There are $\binom{n}{2k,\ell,n-2k-\ell}$ such choices.
2. choose the distinct k edges of distinct vertices, among the chosen $2k$ vertices, that will become the marked double edges. There are $\frac{(2k)!}{2^k k!}$ such choices.
3. order the $2k$ marked double edges and the vertices in each of them. There are $\frac{(2k)!4^k}{2^k}$ ways to order them.
4. order the ℓ loops. There are $\ell!$ ways to do so.
5. choose among the m edges of the final ordering which ones receive marked loops and which ones receive marked double edges. There are $\binom{m}{2k,\ell,m-2k-\ell}$ choices.

6. to fill the rest of the final ordering, build an ordering of length $2m-4k-2\ell$ where the $2k$ vertices that belong to marked double edges and the ℓ vertices that appear in marked loops have degree in $\mathcal{D}-2$, while the other $n-2k-\ell$ vertices have degree in $\mathcal{D}$. The number of such orderings is $(2m-4k-2\ell)![x^{2m-4k-2\ell}]\operatorname{Set}_{\mathcal{D}-2}(x)^{2k+\ell}\operatorname{Set}_{\mathcal{D}}(x)^{n-2k-\ell}$.

This bijective construction implies the following enumerative result

$$\begin{aligned}&\sum_{G\in\operatorname{Marked}(\mathrm{MG}^{(\mathcal{D})}_{n,m})}\kappa(G)u^{k(G)}v^{\ell(G)}\\ =&\frac{1}{2^m m!}\sum_{k,\ell\geq 0}\binom{n}{2k,\ell,n-2k-\ell}\frac{(2k)!}{2^k k!}\frac{(2k)!4^k}{2^k}\ell!\\ &\times\binom{m}{2k,\ell,m-2k-\ell}(2m-4k-2\ell)!\\ &\times[x^{2m-4k-2\ell}]\operatorname{Set}_{\mathcal{D}-2}(x)^{2k+\ell}\operatorname{Set}_{\mathcal{D}}(x)^{n-2k-\ell}u^{k(G)}v^{\ell(G)}.\end{aligned}$$

After simplification, this last expression can be rewritten

$$\begin{aligned}\operatorname{Marked}_{\mathrm{MG}^{(\mathcal{D})}_{n,m}}(u,v)&=\frac{(2m)!}{2^m m!}\\ \times[x^{2m}]&\left(\sum_{k,\ell\geq 0}a_{n,m,2k+\ell}\frac{(uW_{\frac{n}{m}}(x)^2)^k}{k!}\frac{(vW_{\frac{n}{m}}(x))^\ell}{\ell!}\right)\\ \times\operatorname{Set}&_{\mathcal{D}}(x)^n.\end{aligned}$$

□

We observe that when $2k+\ell$ is fixed while n,m tends to infinity, then $a_{n,m,2k+\ell}$ tends to 1. The double sum can then be approximated by an exponential, and it is tempting to conclude

$$\operatorname{Marked}_{\mathrm{MG}^{(\mathcal{D})}_{n,m}}(u,v)\sim[x^{2m}]e^{uW_{\frac{n}{m}}(x)^2+vW_{\frac{n}{m}}(x)}\operatorname{Set}_{\mathcal{D}}(x)^n.$$

The next lemma formalizes this intuition. A multivariate generating function $f(x_1,\dots,x_n)$ is said to *dominate coefficient-wise* another series $g(x_1,\dots,x_n)$ if for all $k_1,\dots,k_n\geq 0$,

$$\left|[x_1^{k_1}\cdots x_n^{k_n}]g(x_1,\dots,x_n)\right|\leq[x_1^{k_1}\cdots x_n^{k_n}]f(x_1,\dots,x_n).$$

Lemma 4.4.3. *When m/n stays in a fixed compact interval of $\mathbb{R}_{>0}$, there is an entire bivariate analytic function $C(u,v)$ such that, for n large enough, $\frac{1}{n}C(u,v)$ dominates coefficient-wise*

$$e^{u+v}-\sum_{k,\ell\geq 0}a_{n,m,2k+\ell}\frac{u^k}{k!}\frac{v^\ell}{\ell!}. \tag{4.1}$$

Proof. Developing the exponential as a double sum

$$e^{u+v} = \sum_{k,\ell \geq 0} \frac{u^k}{k!} \frac{v^\ell}{\ell!},$$

the result can be rewritten

$$n \frac{|1 - a_{n,m,2k+\ell}|}{k!\ell!} \leq [u^k v^\ell] C(u,v)$$

for all k, ℓ. We prove that when n is large enough, we have

$$n \frac{|1 - a_{n,m,2k+\ell}|}{k!\ell!} \leq \left(1 + \frac{n}{m}\right) \frac{(2k+\ell)^2 e^{4k+2\ell}}{\sqrt{k!\ell!}} \tag{4.2}$$

for all $k, \ell \geq 1$. Since the right-hand side are the coefficients of a function analytic on $\mathbb{C}^2$, this will conclude the proof.

Let $b_{n,j}$ denote the value $\prod_{i=0}^{j-1} \left(1 - \frac{i}{n}\right)$, then observe that $a_{n,m,j}$ is equal to $b_{n,j} b_{m,j} / b_{2m,2j}$. Since $b_{n,j} \leq 1$, if $(c_{n,j})$ denotes a sequence such that $c_{n,j} \leq b_{n,j}$ for all (n,j), then $c_{n,j} c_{m,j} \leq a_{n,m,j} \leq c_{2m,2j}^{-1}$, which implies

$$\begin{aligned} & n \frac{|1 - a_{n,m,2k+\ell}|}{k!\ell!} \\ & \leq n \frac{\max(c_{2m,4k+2\ell}^{-1} - 1, 1 - c_{n,2k+\ell} c_{m,2k+\ell})}{k!\ell!}. \end{aligned} \tag{4.3}$$

We now prove that Equation (4.2) holds both for $2k+\ell \leq \sqrt{m}/2$ and for $2k+\ell > \sqrt{m}/2$.

Case $2k+\ell \leq \sqrt{m}/2$. We prove by recurrence on j that $b_{n,j} \geq 1 - \frac{j^2}{n}$. The recurrence is initialized with $b_{n,0} = 1$. Assuming it is satisfied at rank j, then

$$\begin{aligned} b_{n,j+1} = \left(1 - \frac{j}{n}\right) b_{n,j} &\geq \left(1 - \frac{j}{n}\right)\left(1 - \frac{j^2}{n}\right) \\ &\geq 1 - \frac{(j+1)^2}{n}, \end{aligned}$$

which concludes the proof of the recurrence. This implies, using Inequality (4.3),

$$\begin{aligned} n \frac{|1 - a_{n,m,2k+\ell}|}{k!\ell!} \leq \frac{n}{k!\ell!} \max\Bigg(& \frac{1}{1 - \frac{(4k+2\ell)^2}{2m}} - 1, \\ & 1 - \left(1 - \frac{(2k+\ell)^2}{n}\right)\left(1 - \frac{(2k+\ell)^2}{m}\right)\Bigg). \end{aligned}$$

Since $2k + \ell \leq \sqrt{m}/2$, the first argument of the maximum function is at most 1. The second argument is smaller than $(n^{-1} + m^{-1})(2k+\ell)^2$. Therefore, we have

$$n \frac{|1 - a_{n,m,2k+\ell}|}{k!\ell!} \leq \left(1 + \frac{n}{m}\right) \frac{(2k+\ell)^2}{k!\ell!},$$

and Inequality (4.2) is satisfied.

Case $2k+\ell > \sqrt{m}/2$**.** We first prove $b_{n,j} \geq e^{-j}$. To do so, we apply a sum-integral comparison in the expression

$$\begin{aligned}\log(b_{n,j}) &= \sum_{i=0}^{j-1} \log\left(1-\frac{i}{n}\right)\\ &\geq \int_0^j \log\left(1-\frac{x}{n}\right) dx\\ &= -(n-j)\log\left(1-\frac{j}{n}\right) - j \geq -j.\end{aligned}$$

Inequality (4.3) then implies

$$\begin{aligned}n\frac{|1-a_{n,m,2k+\ell}|}{k!\ell!} &\leq \frac{n}{k!\ell!}\max\left(e^{4k+2\ell}-1, 1-e^{-(4k+2\ell)}\right)\\ &\leq \frac{n}{\sqrt{k!\ell!}}\frac{e^{4k+2\ell}}{\sqrt{k!\ell!}}.\end{aligned}$$

We now prove that $n/\sqrt{k!\ell!}$ is smaller than 1 for n large enough. Indeed, $2k+\ell > \sqrt{m}/2$ implies $\max(k,\ell) \geq \sqrt{m}/8$, so

$$\frac{n}{\sqrt{k!\ell!}} \leq \frac{n}{\sqrt{\max(k,\ell)!}} \leq \frac{n}{(\sqrt{m}/8)!},$$

and since m/n stays in a compact interval of $\mathbb{R}_{>0}$, this last term tends to 0 with n. We then conclude

$$n\frac{|1-a_{n,m,2k+\ell}|}{k!\ell!} \leq \frac{e^{4k+2\ell}}{\sqrt{k!\ell!}}$$

for n large enough, so Inequality (4.2) is satisfied. □

We can now derive the asymptotics of $\mathrm{Marked}_{\mathrm{MG}_{n,m}^{(\mathcal{D})}}(u,v)$. As observed in the discussion preceding Theorem 4.3.2, the result is trivial when $\mathcal{D}$ contains only one integer, when $2m/n$ is outside $[\min(\mathcal{D}), \max(\mathcal{D})]$ and when p does not divide $2m - \min(\mathcal{D})n$.

Lemma 4.4.4. *Assume $\mathcal{D}$ has size at least 2, valuation r and periodicity p. Let m, n denote two integers tending to infinity, such that $2m/n$ stays in a fixed compact interval of $]r, \max(\mathcal{D})[$ and p divides $2m - rn$. When u, v stay in a fixed compact, then*

$$\begin{aligned}\mathrm{Marked}_{\mathrm{MG}_{n,m}^{(\mathcal{D})}}(u,v) = &\frac{(2m)!}{2^m m!}\frac{\mathrm{Set}_{\mathcal{D}}(\zeta)^n}{\zeta^{2m}}\frac{p}{\sqrt{2\pi n\zeta\phi'(\zeta)}}\\ &\times e^{uW_{\frac{n}{m}}(\zeta)^2+vW_{\frac{n}{m}}(\zeta)}(1+O(n^{-1})),\end{aligned} \tag{4.4}$$

where $W_{\frac{n}{m}}(x) = \frac{n}{4m}\frac{x^2\,\mathrm{Set}_{\mathcal{D}-2}(x)}{\mathrm{Set}_{\mathcal{D}}(x)}$, $\phi(x) = \frac{x\,\mathrm{Set}_{\mathcal{D}-1}(x)}{\mathrm{Set}_{\mathcal{D}}(x)}$ *and* $\phi(\zeta) = \frac{2m}{n}$.

Proof. We start with the expression of $\mathrm{Marked}_{\mathrm{MG}^{(\mathcal{D})}_{n,m}}(u,v)$ derived in Lemma 4.4.2

$$\begin{aligned}&\mathrm{Marked}_{\mathrm{MG}^{(\mathcal{D})}_{n,m}}(u,v) = \\ &\frac{(2m)!}{2^m m!}[x^{2m}]\left(\sum_{k,\ell\geq 0} a_{n,m,2k+\ell}\frac{(uW_{\frac{n}{m}}(x)^2)^k}{k!}\frac{(vW_{\frac{n}{m}}(x))^\ell}{\ell!}\right) \\ &\quad\times \mathrm{Set}_{\mathcal{D}}(x)^n.\end{aligned}$$

Using the notation

$$\begin{aligned}A(x) =& e^{uW_{\frac{n}{m}}(x)^2+vW_{\frac{n}{m}}(x)} \\ &- \sum_{k,\ell\geq 0} a_{n,m,2k+\ell}\frac{(uW_{\frac{n}{m}}(x)^2)^k}{k!}\frac{(vW_{\frac{n}{m}}(x))^\ell}{\ell!},\end{aligned}$$

this implies

$$\begin{aligned}&\frac{(2m)!}{2^m m!}[x^{2m}]e^{uW_{\frac{n}{m}}(x)^2+vW_{\frac{n}{m}}(x)}\,\mathrm{Set}_{\mathcal{D}}(x)^n \\ &\quad - \mathrm{Marked}_{\mathrm{MG}^{(\mathcal{D})}_{n,m}}(u,v) \\ &= \frac{(2m)!}{2^m m!}[x^{2m}]A(x)\,\mathrm{Set}_{\mathcal{D}}(x)^n.\end{aligned}$$

Observe that $W_{\frac{n}{m}}(x)$ has valuation 0 and period p. According to Lemma 4.2.1, we have

$$\begin{aligned}&\frac{(2m)!}{2^m m!}[x^{2m}]e^{uW_{\frac{n}{m}}(x)^2+vW_{\frac{n}{m}}(x)}\,\mathrm{Set}_{\mathcal{D}}(x)^n \\ &= \frac{(2m)!}{2^m m!}\frac{\mathrm{Set}_{\mathcal{D}}(\zeta)^n}{\zeta^{2m}}\frac{p}{\sqrt{2\pi n\zeta\phi'(\zeta)}} \\ &\times e^{uW_{\frac{n}{m}}(\zeta)^2+vW_{\frac{n}{m}}(\zeta)}(1+O(n^{-1})),\end{aligned}$$

so the demonstration is complete if we prove

$$\frac{(2m)!}{2^m m!}[x^{2m}]A(x)\,\mathrm{Set}_{\mathcal{D}}(x)^n = \frac{(2m)!}{2^m m!}O\left(n^{-1}\frac{\mathrm{Set}_{\mathcal{D}}(\zeta)^n}{\zeta^{2m}\sqrt{n}}\right).$$

The Taylor coefficients of $W_{\frac{n}{m}}(x)$ need not be positive, so we introduce the entire function

$$\tilde{W}_{\frac{n}{m}}(x) = \sum_{n\geq 0}|[z^n]W_{\frac{n}{m}}(z)|x^n,$$

which dominate $W_{\frac{n}{m}}(x)$ coefficient-wise. By application of Lemma 4.4.3, $\frac{1}{n}C(u\tilde{W}_{\frac{n}{m}}(x)^2, v\tilde{W}_{\frac{n}{m}}(x))$ dominates coefficient-wise $A(x)$, and therefore

$$\begin{aligned}&\left|\frac{(2m)!}{2^m m!}[x^{2m}]A(x)\,\mathrm{Set}_{\mathcal{D}}(x)^n\right| \\ &\leq \frac{(2m)!}{2^m m!}[x^{2m}]\frac{1}{n}C(u\tilde{W}_{\frac{n}{m}}(x)^2, v\tilde{W}_{\frac{n}{m}}(x))\,\mathrm{Set}_{\mathcal{D}}(x)^n.\end{aligned}$$

Finally, according to Lemma 4.2.1, we have

$$\begin{aligned}&\frac{(2m)!}{2^m m!}[x^{2m}]\frac{1}{n}C(u\tilde{W}_{\frac{n}{m}}(x)^2, v\tilde{W}_{\frac{n}{m}}(x))\operatorname{Set}_{\mathcal{D}}(x)^n\\ &= \frac{(2m)!}{2^m m!}O\left(n^{-1}\frac{\operatorname{Set}_{\mathcal{D}}(\zeta)^n}{\zeta^{2m}\sqrt{n}}\right).\end{aligned}$$

□

4.4.4 Negligible marked multigraphs

Recall that $\mathrm{MG}_{n,m}^{(\mathcal{D},0)}$ denotes the set $\mathrm{MG}_{n,m}^{(\mathcal{D})}\setminus\mathrm{MG}_{n,m}^{(\mathcal{D},*)}$. In Lemma 4.4.6, we prove that $\mathrm{Marked}_{\mathrm{MG}_{n,m}^{(\mathcal{D},0)}}(-1,-1)$ is negligible. To do so, we first bound $\mathrm{Marked}_R(1,1)$ for a family R of marked multigraphs from $\mathrm{MG}_{n,m}^{(\mathcal{D})}$ with mandatory edges.

Lemma 4.4.5. *Let $e_1,\dots,e_j$ denote j edges on the set of vertices $\{1,\dots,n\}$, and R the set of multigraphs from $\mathrm{MG}_{n,m}^{(\mathcal{D})}$ that contain those edges, with multiplicities (i.e. an edge with k occurrences in the list has at least k occurrences in the multiset of edges of the multigraph)*

$$R = \left\{G\in\mathrm{MG}_{n,m}^{(\mathcal{D})}\;\middle|\;\forall i,\ e_i\in E(G)\ \textit{with multiplicities}\right\}.$$

Assume $\mathcal{D}$ contains at least two integers and has valuation r. Let m, n denote two integers tending to infinity, such that $2m/n$ stays in a fixed compact interval of $]r,\max(\mathcal{D})[$, then

$$\mathrm{Marked}_R(1,1) = O\big(n^{-j}\,\mathrm{Marked}_{\mathrm{MG}_{n,m}^{(\mathcal{D})}}(1,1)\big).$$

Proof. Let $\tilde{R}$ denote the set of multigraphs from $\mathrm{MG}_{n,m}^{(\mathcal{D})}$ with j distinguished mandatory edges

$$e_1=\{v_1,v_2\},\dots,e_j=\{v_{2j-1},v_{2j}\}.$$

Given an ordering of a multigraph from R, we can distinguish the first occurrences of the mandatory edges, in order to obtain the ordering of a multigraph from $\tilde{R}$. Therefore, the number of orderings of multigraphs from R is at most equal to the number of orderings of multigraphs from $\tilde{R}$. Dividing by $2^m m!$, this implies

$$\sum_{G\in R}\kappa(G)\le\sum_{G\in\tilde{R}}\kappa(G),$$

so $\mathrm{Marked}_R(1,1)\le\mathrm{Marked}_{\tilde{R}}(1,1)$.

Let W denote the fixed set of vertices that appear in the mandatory edges, and for all $w\in W$, let d_w denote the number of occurrences of the vertex w in the mandatory edges

$$d_w = \big|\{i\mid v_i=w\}\big|.$$

Let also $G_{n,m}^{(d)}$ denote the set of multigraphs with n vertices and m edges, where each vertex w from the mandatory edges has degree in $\mathcal{D}-d_w$ and the other vertices have degrees in $\mathcal{D}$. To construct an ordering from a multigraph in $\text{Marked}_{\tilde{R}}$, we choose the j positions of the mandatory edges among the m positions available, the order of the vertices in those edges, and mark or not each of them. Then the rest of the ordering is filled with an ordering from $\text{Marked}_{G^{(d)}_{n,m-j}}$. Therefore, the number of orderings from $\text{Marked}_{\tilde{R}}$ is at most

$$m^j 2^j 2^j 2^{m-j}(m-j)! \, \text{Marked}_{G^{(d)}_{n,m-j}}(1,1).$$

Dividing by $2^m m!$ and using the fact that j is fixed, we obtain

$$\text{Marked}_{\tilde{R}}(1,1) = O\left(\text{Marked}_{G^{(d)}_{n,m-j}}(1,1)\right). \tag{4.5}$$

Following the steps of Lemma 4.4.2, $\text{Marked}_{G^{(d)}_{n,m-j}}(1,1)$ is not greater than

$$\begin{aligned}&\frac{(2m-2j)!}{2^{m-j}(m-j)!}[x^{2m-2j}]\left(\sum_{k,\ell\geq 0} a_{n,m-j,2k+\ell}\frac{W_{\frac{n}{m}}(x)^{2k+\ell}}{k!\ell!}\right)\\ &\times\left(\prod_{v\in W}\text{Set}_{\mathcal{D}-d_v}(x)\right)\text{Set}_{\mathcal{D}}(x)^{n-|W|}.\end{aligned}$$

An application of the same argument as in the proof of Lemma 4.4.4 leads to

$$\begin{aligned}\text{Marked}_{G^{(d)}_{\mathcal{D}}(n,m-j)}(1,1) =&\frac{(2m-2j)!}{2^{m-j}(m-j)!}\\ &\times O\left(\frac{\text{Set}_{\mathcal{D}}(\zeta)^{n-|W|}}{\zeta^{2(m-j)}\sqrt{n-|W|}}\right).\end{aligned}$$

Since $|W|$ and j are fixed, this implies, using Lemma 4.4.4,

$$\begin{aligned}\text{Marked}_{G^{(d)}_{n,m-j}}(1,1) =&\frac{(2m-2j)!}{2^{m-j}(m-j)!}\frac{2^m m!}{(2m)!}\\ &\times OO\left(\text{Marked}_{\text{MG}^{(\mathcal{D})}_{n,m}}(1,1)\right).\end{aligned}$$

Simplifying and injecting this relation from Equation (4.5), we obtain

$$\text{Marked}_{\tilde{R}}(1,1) = O\big(n^{-j}\,\text{Marked}_{\text{MG}^{(\mathcal{D})}_{n,m}}(1,1)\big).$$

□

Figure 4.1 displays four multigraphs from $\text{MG}^{(\mathcal{D},0)}_{n,m}$. Actually, any multigraph from $\text{MG}^{(\mathcal{D},0)}_{n,m}$ contains one of those four graphs as a subgraph, and this property can be described in terms of mandatory edges. In the following lemma, we use this fact to bound $\text{Marked}_{\text{MG}^{(\mathcal{D},0)}_{n,m}}(-1,-1)$.

Lemma 4.4.6. *Assume $\mathcal{D}$ contains at least two integers, has valuation r and periodicity p. Let m, n denote two integers tending to infinity, such that $2m/n$ stays in a fixed compact interval of $]r, \max(\mathcal{D})[$, and p divides $2m - nr$, then*

$$\begin{aligned}&\mathrm{Marked}_{\mathrm{MG}_{n,m}^{(\mathcal{D},0)}}(-1,-1)\\&= O\left(n^{-1}\,\mathrm{Marked}_{\mathrm{MG}_{n,m}^{(\mathcal{D})}}(-1,-1)\right).\end{aligned}$$

Proof. By definition, a multigraph G belongs to $\mathrm{MG}_{n,m}^{(\mathcal{D},0)}$ if and only if it contains a vertex v that is in one of the following configurations:

1. the loop {v,v} appears at least twice in $E(G)$,
2. there is a vertex u such that the edge {u,v} appears at least three times,
3. there is a vertex u such that {v,v} is in $E(G)$ and {u,v} appears at least twice,
4. there are vertices u and w such that {u,v} and {v,w} both appear at least twice.

Let $\tilde{R}_1$ (resp. $\tilde{R}_2$, $\tilde{R}_3$, $\tilde{R}_4$) denote the set of multigraphs from $\mathrm{MG}_{n,m}^{(\mathcal{D})}$ that contain a vertex in configuration 1 (resp. 2, 3, 4). We then have

$$\mathrm{MG}_{n,m}^{(\mathcal{D},0)} = \tilde{R}_1 \cup \tilde{R}_2 \cup \tilde{R}_3 \cup \tilde{R}_4.$$

Let also R_1, R_2, R_3 and R_4 denote four subsets of $\mathrm{MG}_{n,m}^{(\mathcal{D})}$, such that

1. the multigraphs from R_1 contain two occurrences of the loop $\{1,1\}$,
2. the multigraphs from R_2 contain three occurrences of the edge $\{1,2\}$,
3. the multigraphs from R_3 contain an occurrence of $\{1,1\}$ and two occurrences of $\{1,2\}$,
4. the multigraphs from R_4 contain two occurrences of $\{1,2\}$ and two occurrences of $\{1,3\}$

(see Figure 4.1). Given the symmetric roles of the vertices, the number of orderings from multigraphs in $\tilde{R}_1$ (resp. $\tilde{R}_2$, $\tilde{R}_3$, $\tilde{R}_4$) is lesser than or equal to n times (resp. n^2, n^2, n^3) the number of orderings from multigraphs in R_1 (resp. R_2, R_3, R_4). This implies

$$\begin{aligned}&\mathrm{Marked}_{\tilde{R}_1}(1,1) \le n\,\mathrm{Marked}_{R_1}(1,1),\\&\mathrm{Marked}_{\tilde{R}_2}(1,1) \le n^2\,\mathrm{Marked}_{R_2}(1,1),\\&\mathrm{Marked}_{\tilde{R}_3}(1,1) \le n^2\,\mathrm{Marked}_{R_3}(1,1),\\&\mathrm{Marked}_{\tilde{R}_4}(1,1) \le n^3\,\mathrm{Marked}_{R_4}(1,1),\end{aligned}$$

so

$$\begin{aligned}\mathrm{Marked}_{\mathrm{MG}_{n,m}^{(\mathcal{D},0)}}(1,1) \leq & n \,\mathrm{Marked}_{R_1}(1,1)\\ &+ n^2 \,\mathrm{Marked}_{R_2}(1,1)\\ &+ n^2 \,\mathrm{Marked}_{R_3}(1,1)\\ &+ n^3 \,\mathrm{Marked}_{R_4}(1,1).\end{aligned}$$

The multigraphs from R_1 (resp. R_2, R_3, R_4) have 2 mandatory edges (resp. 3, 3, 4). Four applications of Lemma 4.4.5 lead to

$$\mathrm{Marked}_{\mathrm{MG}_{n,m}^{(\mathcal{D},0)}}(1,1) = O(n^{-1})\,\mathrm{Marked}_{\mathrm{MG}_{n,m}^{(\mathcal{D})}}(1,1).$$

Finally, according to Lemma 4.4.4,

$$\mathrm{Marked}_{\mathrm{MG}_{n,m}^{(\mathcal{D})}}(1,1) = O\left(\mathrm{Marked}_{\mathrm{MG}_{n,m}^{(\mathcal{D})}}(-1,-1)\right).$$

□

Now we have all the ingredients to prove Theorem 4.1.1.

of Theorem 4.1.1. In Lemma 4.4.1, we have proven that the number of simple graphs in $\mathrm{SG}_{n,m}^{(\mathcal{D})}$ is equal to $\mathrm{Marked}_{\mathrm{MG}_{n,m}^{(\mathcal{D},*)}}(-1,-1)$. By a set manipulation, this quantity can be rewritten

$$\mathrm{Marked}_{\mathrm{MG}_{n,m}^{(\mathcal{D})}}(-1,-1) - \mathrm{Marked}_{\mathrm{MG}_{n,m}^{(\mathcal{D},0)}}(-1,-1),$$

where $\mathrm{MG}_{n,m}^{(\mathcal{D},0)} = \mathrm{MG}_{n,m}^{(\mathcal{D})} \setminus \mathrm{MG}_{n,m}^{(\mathcal{D},*)}$. Replacing the second term with the result of Lemma 4.4.6, we obtain

$$|\,\mathrm{SG}_{n,m}^{(\mathcal{D})}\,| = \mathrm{Marked}_{\mathrm{MG}_{n,m}^{(\mathcal{D})}}(-1,-1)\big(1 + O(n^{-1})\big).$$

Finally, the asymptotics of $\mathrm{Marked}_{\mathrm{MG}_{n,m}^{(\mathcal{D})}}(-1,-1)$ has been derived in Lemma 4.4.4. □

4.5 Random generation

In order to keep a combinatorial interpretation, we focused on generating functions $\mathrm{Set}_{\mathcal{D}}(x)$ with coefficients in $\{0,1\}$. Our results hold more generally for any generating function $D(x)$ with nonnegative coefficients and large enough radius of convergence (so that the saddle-point from Lemma 4.2.1 is well defined). Multigraphs are then counted with a weight that depends of the degrees of their vertices

$$\mathrm{weight}(G) = \kappa(G) \prod_{v \in V(G)} \deg(v)! [x^{\deg(v)}] D(x).$$

The present work has been guided by experiments on large random graphs with degree constraints. We used exact and Boltzmann sampling [13]. Observe that to build a random simple graph from $\mathrm{SG}_{n,m}^{(\mathcal{D})}$, one can sample multigraphs from $\mathrm{MG}_{n,m}^{(\mathcal{D})}$ and reject until the multigraph is simple. As a consequence of Theorem 4.1.1, the expected number of rejections is $e^{-W_{\frac{n}{m}}(\zeta)^2 - W_{\frac{n}{m}}(\zeta)}$ (using the notations of the theorem).

4.5.1 Boltzmann sampling

The construction of the Boltzmann algorithm is straightforward from Theorem 4.3.1. To build a random multigraph with degrees in $\mathcal{D}$, n vertices and approximately m edges, the algorithm first computes a positive value x, according to the number of edges targeted. It then draws independently n integers $(d_1, \dots, d_n)$, following the law

$$\mathbb{P}(d) = \frac{\left([z^d]D(z)\right) x^d}{D(x)} \tag{4.6}$$

with $D(x) = \mathrm{Set}_{\mathcal{D}}(x)$. If their sum is odd, a new sequence is drawn. Otherwise, the algorithm outputs a random multigraph with sequence of degrees $(d_1, \dots, d_n)$. To do so, as in the configuration model ([3], [54]), each vertex v_i receives d_i half-edges, and a random pairing on the half-edges is drawn uniformly.

Therefore, the random distribution induced on multigraphs by the Boltzmann sampling algorithm is identical to the configuration model. Conversely, given a probability distribution on $\mathbb{Z}_{\geq 0}$, one can choose $D(x)$ so that the distribution is equal to the one described by Equation (4.6). Thus, we expect random multigraphs from the configuration model and multigraphs with degree constraints to share many statistical properties.

4.5.2 Recursive method

For the sampling of a multigraph in $\mathrm{MG}_{n,m}^{(\mathcal{D})}$, the generator first draws a sequence of degrees, and then performs a random pairing of half-edges, as in configuration model and the Boltzmann sampler. Each sequence $(d_1, \dots, d_n)$ from $\mathcal{D}^n$ is drawn with weight $\prod_{v=1}^{n} 1/(d_v)!$. In the first step, we use dynamic programming to precompute the values $(S_{i,j})_{0 \leq i \leq n, 0 \leq j \leq 2m}$, sums of the weights of all the sequences of i degrees that sum to j

$$S_{i,j} = \sum_{\substack{d_1,\dots,d_i \in \mathcal{D} \\ d_1 + \dots + d_i = j}} \prod_{v=1}^{i} \frac{1}{d_v!},$$

using the initial conditions and the recursive expression

$$S_{i,j} = \begin{cases} 1 & \text{if } (i,j) = (0,0), \\ 0 & \text{if } i = 0 \text{ and } j \neq 0, \text{ or if } j < 0, \\ \sum_{d \in \mathcal{D}} \frac{S_{i-1,j-d}}{d!} & \text{otherwise.} \end{cases}$$

After this precomputation, we generate the sequence of degrees as follows: first we sample the last degree d_n of the sequence according to the distribution

$$\mathbb{P}(d_n = d) = \frac{S_{n-1,2m-d}}{d! S_{n,2m}},$$

then we recursively generate the remaining sequence $(d_1, \dots, d_{n-1})$, which must sum to $2m - d_n$. Once the sequence of degrees is computed, we generate a random pairing on the corresponding half-edges.

4.6 Forthcoming research

The results presented can be extended in several ways. The case where $2m/n$ tends to $\min(\mathcal{D})$ or $\max(\mathcal{D})$ could be considered. For example, [10] have derived, using elementary tools, the asymptotics of graphs with a lower bound on the minimum degree when $m = O(n\log(n))$. This extension would only require to adjust the saddle-point method from Lemma 4.2.1.

We have also derived results on the enumeration of graphs where the degree sets vary with the vertices. The model inputs an infinite sequence of sets $(\mathcal{D}_1, \mathcal{D}_2, \dots)$ and output graphs where each vertex v has its degree in $\mathcal{D}_v$. The techniques presented in this chapter can be extended to this case, if some technical conditions are satisfied, such as the convergence of the series $n^{-1}\sum_{v\geq 1}\log(\mathrm{Set}_{\mathcal{D}_v}(x))$. This extension will be part of a longer version of the paper in which this chapter is based. Two examples of such families are graphs with degree parities [26], and graphs with a given degree sequence [63].

We believe that complete asymptotic expansion can be derived for graphs with degree constraints. This would require to apply a more general version of Lemma 4.2.1, such as presented in Chapter 4 by [6], and we would have to consider more complex families than $\mathrm{MG}_{n,m}^{(\mathcal{D},*)}$.

The asymptotics of connected graphs from $\mathrm{SG}_{n,m}$ when $m-n$ tends to infinity has first been derived by [62]. Since then, two new proofs were given, one by [19], the other by [18]. The proof of Pittel and Wormald relies on a link between connected graphs and graphs from a particular family of graphs with degree constraints (graphs with degrees at least 2). In [7], following the same approach, but using analytic combinatorics, we obtained a short proof for the asymptotics of connected multigraphs from $\mathrm{MG}_{n,m}$ when $m-n$ tends to infinity. We now plan to extend this result to simple graphs, and to derive a complete asymptotic expansion.

In this chapter, we have focused on the enumeration of graphs with degree constraints. We can now start the investigation on the typical structure of random instances of such graphs. An application would be the enumeration of Eulerian graphs, *i.e.* connected Euler graphs.

Finally, the inclusion-exclusion technique we used to remove loops and double edges can be extended to forbid any family of subgraphs.

Bibliography

[1] W. T. Tutte. A census of planar maps. *Canad. J. Math.* 15:249–271, 1963.

[2] G. Chapuy, É. Fusy, M. Kang, B. Shoilekova. A complete grammar for decomposing a family of graphs into 3-connected components. *Electronic Journal of Combinatorics*, 15(1):R148, 2008.

[3] B. Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European Journal of Combinatorics*, 1:311–316, 1980.

[4] S. Janson, T. Łuczak, A. Ruciński. An exponential bound for the probability of nonexistence of a specified subgraph in a random graph. *Random graphs '87 (Poznań, 1987)*, 73–87. Wiley, Chichester, 1990.

[5] P. Flajolet and R. Sedgewick. *Analytic combinatorics*. Cambridge University Press, 2009.

[6] R. Pemantle, M. C. Wilson. *Analytic combinatorics in several variables*. Cambridge University Press, New York, NY, USA, 2013.

[7] E. de Panafieu. Analytic combinatorics of graphs, hypergraphs and inhomogeneous graphs. PhD thesis, Université Paris-Diderot, Sorbonne Paris-Cité, 2014.

[8] O. Giménez, M. Noy. Asymptotic enumeration and limit laws of planar graphs. *Journal of the American Mathematical Society*, 22(2):309–329, 2009.

[9] P. Erdős, D.J. Kleitman, B.L. Rothschild. Asymptotic enumeration of kn-free graphs. *International Colloquium on Combinatorial Theory, Atti dei Convegni Lincei, Roma*, 2:19–27, 1976.

[10] B. Pittel, N. C. Wormald. Asymptotic enumeration of sparse graphs with a minimum degree constraint. *J. Comb. Theory, Ser. A*, 101(2):249–263, 2003.

[11] Z. Gao, N. C. Wormald. Asymptotic normality determined by high moments, and submap counts of random maps. *Probab. Theory Related Fields*, 130(3):368–376, 2004.

[12] M. Drmota, É. Fusy, M. Kang, V. Kraus, J. Rué. Asymptotic study of subcritical graph classes. *SIAM J. Discrete Math.*, 25(4):1615–1651, 2011.

[13] P. Duchon, P. Flajolet, G. Louchard, and G. Schaeffer. Boltzmann samplers for the random generation of combinatorial structures. *Combinatorics, Probability and Computing*, 13:, 2004.

[14] I. P. Goulden, D. M. Jackson. *Combinatorial enumeration.* A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1983.

[15] F. Bergeron, G. Labelle, P. Leroux. Combinatorial species and tree-like structures *Cambridge University Press*, 67, 1998.

[16] C. McDiarmid. Connectivity for random graphs from a weighted bridge-addable class. *Electron. J. Combin.*, 19(4), Paper 53, 20 pp, 2012.

[17] W. T. Tutte. *Connectivity in graphs*, volume 285. University of Toronto Press, 1966.

[18] R. van der Hofstad, J. Spencer. Counting connected graphs asymptotically. *European Journal on Combinatorics*, 26(8):1294–1320, 2006.

[19] B. Pittel, N. C. Wormald. Counting connected graphs inside-out. *Journal of Combinatorial Theory, Series B*, 93(2):127–172, 2005.

[20] O. Giménez, M. Noy. Counting planar graphs and related families of graphs. *Surveys in combinatorics* 169–210, 2009.

[21] M. Drmota, O. Giménez, M. Noy. Degree distribution in random planar graphs. *J. Combin. Theory Ser. A* 118:2102–2130, 2011.

[22] Z. Gao, N.C. Wormald. Distribution of subgraphs of random regular graphs. *Random Structures & Algorithms*, 32(1):38–48, 2008.

[23] F. Chyzak, M. Mishna, B. Salvy. Effective scalar products of d-finite symmetric functions. *Journal of Combinatorial Theory, Series A*, 112(1):1–43, 2005.

[24] M. Bodirsky, O. Giménez, M. Kang, M. Noy. Enumeration and limit laws for series–parallel graphs. *European Journal of Combinatorics*, 28(8):2091–2105, 2007.

[25] R. W. Robinson. Enumeration of euler graphs. *Proof Techniques in Graph Theory*, 147–153, 1969.

[26] R. Read, R. Robinson. Enumeration of labelled multigraphs by degree parities. *Discrete Mathematics*, 42(1):99–105, 1982.

[27] M. Drmota, M. Noy. Extremal parameters in sub-critical graph classes. *ANALCO'13*, 1–7, 2013.

[28] O. Giménez, M. Noy, J. Rué. Graph classes with given 3-connected components: asymptotic enumeration and random graphs. *Random Structures and Algorithms*, 42(4):438–479, 2013.

[29] S. K. Lando, A. K. Zvonkin. *Graphs on surfaces and their applications.* Low-dimensional topology. Springer, Berlin, New York, Paris, 2004.

[30] É. de Panafieu, L. Ramos. Graphs with degree constraints *Proceedings of the Thirteenth Workshop on Analytic Algorithmics and Combinatorics*, 34–45, 2016.

[31] A. Georgakopoulos, S. Wagner. Limits of subcritical random graphs and random graphs with excluded minors. Available on-line on `arXiv:1512.03572`.

[32] K. Panagiotou, A. Steger. Maximal biconnected subgraphs of random planar graphs. *ACM Trans. Algorithms*, 6(2), 2010.

[33] A. Meir, J. W. Moon. On an asymptotic method in enumeration. *Journal of Combinatorial Theory, Series A*, 51(1):77–89, 1989.

[34] H. Hwang. On convergence rates in the central limit theorems for combinatorial structures. *European Journal of Combinatorics*, 19(3):329–343, 1998.

[35] P. Erdős, A. Rényi. On the evolution of random graphs. 5:17–61, 1960.

[36] M. Karoński, A. Ruciński. On the number of strictly balanced subgraphs of a random graph. *Graph theory (Łagów, 1981)*, volume 1018 of *Lecture Notes in Math.*, 79–83. Springer, Berlin, 1983.

[37] M. Noy, V. Ravelomanana, J. Rué. On the probability of planarity of a random graph near the critical point. *Proc. Amer. Math. Soc.* (to appear).

[38] M. Bodirsky, M. Kang, M. Löffler, C. McDiarmid. Random cubic planar graphs. *Random Structures Algorithms* 30:78–94, 2007.

[39] M. Noy, C. Requilé, J. Rué. Random cubic planar graphs revisited. In preparation.

[40] B. Stufler. Random enriched trees with applications to random graphs. Available on-line on `arXiv:1504.02006`.

[41] S. Janson, T. Łuczak, A. Ruciński. *Random graphs.* Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.

[42] R. van der Hofstad. *Random graphs and complex networks*, Vol. I, 2014.

[43] C. Banderier, P. Flajolet, G. Schaeffer, M. Soria. Random Maps, Coalescing Saddles, Singularity Analysis, and Airy Phenomena. *Random Strucures Algorithms*, 19:194–246, 2001.

[44] C. McDiarmid, A. Steger, D. J. A. Welsh. Random planar graphs. *Journal of Combinatorial Theory, Series B*, 93(2):187–205, 2005.

[45] M. Noy. Random planar graphs and beyond. *Proceedings of the International Congress of Mathematicians 2014*, Vol. IV:407–430, 2014.

[46] M. Noy, L. Ramos. Random planar graphs with minimum degree two and three *The Seventh European Conference on Combinatorics, Graph Theory and Applications*, 251–258, 2013.

[47] M. Drmota. Random trees: an interplay between combinatorics and probability. *SpringerWienNewYork*, 2009.

[48] M. Mishna. Regularity in weighted graphs: A symmetric function approach. 2015.

[49] K. Panagiotou, B. Stufler, K. Weller. Scaling limits of random graphs from subcritical classes. To appear in *Annals of Probability*. Available on-line at `arXiv:1411.1865`.

[50] Z. Gao, N. C. Wormald. Sharp concentration of the number of submaps in random planar triangulations. *Combinatorica*, 23(3):467–486, 2003.

[51] P. Flajolet, A. Odlyzko. Singularity analysis of generating functions. *SIAM Journal on discrete mathematics*, 3(2):216–240, 1990.

[52] A. Ruciński. Small subgraphs of random graphs—a survey. *Random graphs '87 (Poznań, 1987)*, 283–303. Wiley, Chichester, 1990.

[53] J. H. Kim, B. Sudakov, V. Vu. Small subgraphs of random regular graphs. *Discrete Math.*, 307(15):1961–1967, 2007.

[54] N. Wormald. *Some problems in the enumeration of labelled graphs.* Newcastle University, 1978.

[55] B. D. McKay. Subgraphs of dense random graphs with specified degrees. *Combin. Probab. Comput.*, 20(3):413–433, 2011.

[56] B. D. McKay. Subgraphs of random graphs with specified degrees. *Proceedings of the Twelfth Southeastern Conference on Combinatorics, Graph Theory and Computing, Vol. II (Baton Rouge, La., 1981)*, 33:213–223, 1981.

[57] B. D. McKay. Subgraphs of random graphs with specified degrees. *Proceedings of the International Congress of Mathematicians*, IV:2489–2501. Hindustan Book Agency, New Delhi, 2010.

[58] P. Lieby, B.n McKay, J. McLeod, I. Wanless. Subgraphs of random k-edge-coloured k-regular graphs. *Combinatorics, Probability and Computing*, 18(4):533–549, 2009.

[59] M. Drmota, L. Ramos, J. Rué. Subgraph statistics in subcritical graph classes. *Random Structures & Algorithms* (to appear).

[60] M. Drmota. Systems of functional equations. *Random Structures Algorithms*, 10(1-2):103–124, 1997.

[61] M. Drmota, B. Gittenberger, J. F. Morgenbesser. Systems of functional equations and infinite dimensional gaussian limits distributions in combinatorial enumeration. Submitted, available on-line at `http://www.dmg.tuwien.ac.at/drmota/`.

[62] E. R. Canfield, E. A. Bender, B. D. McKay. The asymptotic number of labeled connected graphs with a given number of vertices and edges. *Random Structures and Algorithm*, 1:129–169, 1990.

[63] E. A. Bender, E. R. Canfield. The asymptotic number of labeled graphs with given degree sequences. *Journal of Combinatorial Theory, Series A*, 24(3):296–307, 1978.

[64] S. Janson, T. Łuczak, D. E. Knuth, B. Pittel. The birth of the giant component. *Random Structures and Algorithms*, 4(3):233–358, 1993.

[65] N. Bernasconi, K. Panagiotou, A. Steger. The degree sequence of random graphs from subcritical classes. *Combinatorics, Probability and Computing*, 18(5):647–681, 2009.

[66] F. Chyzak, M. Drmota, T. Klausner, G. Kok. The distribution of patterns in random trees. *Comb. Probab. Comput.*, 17(1):21–59, 2008.

[67] D. M. Jackson, J. W. Reilly. The enumeration of homeomorphically irreducible labelled graphs. *J. Combin. Theory Ser. B*, 19:272–286, 1975.

[68] P. Flajolet, D. E. Knuth, B. Pittel. The first cycles in an evolving graph. *Discrete Mathematics*, 75():167–215, 1989.

[69] C. Requilé, J. Rué, K. Weller. The ising model on planar graphs: grammar and applications to bipartite enumeration. In preparation.

[70] M. Drmota, O. Giménez, M. Noy. The maximum degree of series-parallel graphs. *Combinatorics, Probability and Computing*, 20(4):529–570, 2011.

[71] E. A. Bender, Z. Gao, N. C. Wormald. The number of labeled 2-connected planar graphs. *Electron. J. Combin.*, 9, Research Paper 43, 13 pp, 2002.

[72] C. Requilé, J. Rué. Triangles in random cubic planar graphs. *Electronic Notes in Discrete Mathematics, Proceedings of Eurocomb 2015*, 49:383–391, 2015.

[73] C. L. Mallows, N. J. A. Sloane. Two-graphs, switching classes and euler graphs are equal in number. *Journal of applied mathematics*, 28(4), 1975.

[74] E. Fusy. Uniform random sampling of planar graphs in linear time. *Random Structures Algorithms* 35:464–522, 2009.

[75] M. Drmota, O. Giménez, M. Noy. Vertices of given degree in series-parallel graphs. *Random Structures Algorithms*, 36(3):273–314, 2010.