Connectivity and related properties for graph classes

Kerstin B. Weller



New College



University of Oxford

Submitted for degree of Doctor of Philosophy

Trinity Term 2013

Connectivity and related properties for graph classes

Kerstin B. Weller, New College

Submitted for degree of Doctor of Philosophy, Trinity Term 2013

There has been much recent interest in random graphs sampled uniformly from the set of (labelled) graphs on n vertices in a suitably structured class $\mathcal{A}$. An important and well-studied example of such a suitable structure is *bridge-addability*, introduced in 2005 by McDiarmid et al. in the course of studying random planar graphs. A class $\mathcal{A}$ is bridge-addable when the following holds: if we take any graph G in $\mathcal{A}$ and any pair u, v of vertices that are in different components in G, then the graph G' obtained by adding the edge uv to G is also in $\mathcal{A}$. It was shown that for a random graph sampled from a bridge-addable class, the probability that it is connected is always bounded away from 0, and the number of components is bounded above by a Poisson law.

What happens if 'bridge-addable' is replaced by something weaker? In this thesis, this question is explored in several different directions.

After an introductory chapter and a chapter on generating function methods presenting standard techniques as well as some new technical results needed later, we look at minor-closed, labelled classes of graphs. The excluded minors are always assumed to be connected, which is equivalent to the class $\mathcal{A}$ being *decomposable* - a graph is in $\mathcal{A}$ if and only if every component of the graph is in $\mathcal{A}$. When $\mathcal{A}$ is minor-closed, decomposable and bridge-addable various properties are known (McDiarmid 2010), generalizing results for planar graphs. A minor-closed class is decomposable and bridge-addable if and only if all excluded minors are 2-connected. Chapter **3** presents a series of examples where the excluded minors are not 2-connected, analysed using generating functions as well as techniques from graph theory. This is a step towards a classification of connectivity behaviour for minor-closed classes of graphs. In contrast to the bridge-addable case, different types of behaviours are observed. Chapter **4** deals with a new, more general variant of bridge-addability related to edge-expander graphs. We will see that as long as we are allowed to introduce 'sufficiently many' edges between components, the number of components of a random graph can still be bounded above by a Poisson law. In this context, random forests in $K_{n,n}$ are studied in detail. Chapter **5** takes a different approach, and studies the class of labelled forests where some vertices belong to a specified stable set. A weighting parameter y for the vertices belonging to the stable set is introduced, and a graph is sampled with probability proportional to y^s where s is the size of its stable set. The behaviour of this class is studied for y tending to ∞. Chapters **6** concerns random graphs sampled from general decomposable classes. We investigate the minimum size of a component, in both the labelled and the unlabelled case.

Contents

1. Introduction **3**

1.1. Overview of thesis 9

2. Generating function methods **12**

2.1. Generatingfunctionology for graphs 12

2.2. Generating functions next to their radius of convergence 16

2.3. Singularity Analysis 20

2.3.1. Labelled forests 23

2.4. Hayman-admissibility and the saddle point method 25

2.5. The method of moments and special probability distributions 35

2.6. A generating function with a simple pole 37

2.7. Random generation using Boltzmann models 40

3. Excluding connected minors **42**

3.1. Introduction: Excluding 2-connected minors 42

3.2. Tree-like structures 44

3.2.1. Excluding a spoon 50

3.3. Excluding the diamond and the bow tie: a logarithmic singularity . 53

3.4. Forests of paths, caterpillars and combs 57

3.5. Forests of graphs of bounded component size 66

3.6. Graphs with maximal degree 2 70

3.7. Excluding the bow tie 73

3.8. Random generation 86

3.9. Summary and conlusion 88

4. Relatively bridge-addable classes of graphs **90**

4.1. Bridge-addable and relatively bridge-addable 90

4.2. Expander host graphs 94

4.3. $K_{n/2,n/2}$ as a host graph 96

4.4. Rooted graphs in expander host graphs 106

4.5. Remarks 112

5. Forests with a stable set **114**
5.1. Forests with a stable set – a partially bridge-addable class 114
5.1.1. The setup 114
5.2. Rooted trees with a stable set 116
5.3. Results on unrooted trees with a stable set 119

6. Smallest component in decomposable structures **128**
6.1. The size of the smallest component in decomposable structures . . . 128
6.2. The size of the smallest component in general smooth decomposable structures 131
6.2.1. Labelled structures 132
6.2.2. Unlabelled structures 137

7. Conclusion **144**

A. Statement of originality **145**

B. Acknowledgements **146**

1. Introduction

Let $\mathcal{A}$ be a *class*, that is a set of labelled graphs fulfilling suitable structural conditions. We are interested in the properties of R_n, a graph taken uniformly at random from the set of all n-vertex graphs in $\mathcal{A}$. We write $R_n \in_u \mathcal{A}$. One of the first graph classes for which $R_n \in_u \mathcal{A}$ was analysed was the class of planar graphs. A graph is *planar* if it can be embedded in the plane, i.e. if it is isomorphic to a plane graph (see [36] for an introduction). Studying planar graphs initially focussed on the famous Four Colour Problem: can the regions of every map on a globe be coloured with four colours so that neighbouring regions which share a boundary have different colours? This problem, first proposed in October 1852 by the student Francis Guthrie, inspired generations of mathematicians until it was finally solved in 1976 by Appel and Haken using a computer-based proof for the first time. An accessible history of the Four Colour Problem can be found in [90]. The theory of random planar graphs $R_n \in \mathcal{A}$ however started with [35] in 1996, followed by several papers around early to mid 2000's, for instance [76] and [52], and by [53] ([54]) in 2005 where Giménez and Noy were able to asymptotically count the number of labelled planar graphs on n vertices. Of the many properties studied, the two most important ones for this thesis are the probability that $R_n \in_u \mathcal{A}$ is connected and more generally the limit law for the number of components. It was shown in [53], [54] that the probability that $R_n \in \mathcal{A}$ is connected converges to a computable constant (≈ 0.96325) and the number of components is asymptotically distributed as $1 + X$ where X is a Poisson law of parameter $\nu \approx 0.037439$. The results are based on generating function methods and analytic tools, and previously obtained results on maps play a key role in their proofs. A *map* is a connected planar graph which is embedded in the plane. The theory of map enumeration was started by Tutte in 1960's [89] and has grown enormously since then, in part due to the fact that it has rich applications in other fields of mathematics and physics (see [68]).

A graph is *c-(vertex)-connected* if we are required to remove at least c vertices before the graph is disconnected. In general, in enumerating c-connected graphs or maps for c a small constant it is natural to proceed from 3-connected to 2-connected

to 1-connected structures where graphical decomposition theorems are represented by functional compositions of the corresponding generating functions. Using techniques from map enumeration as well as the fact that any 3-connected planar graph has a unique embedding on the sphere [36, Theorem 4.3.2], 2-connected labelled planar graphs were counted asymptotically for the first time in [17]. This was a crucial step in the enumeration of planar graphs as well as other graph classes. For a more detailed overview and many more results on various graph classes see the survey [55]. Furthermore, it should be noted that graph enumeration on its own is a well-established branch of combinatorics and one of the most famous and most beautiful examples is Cayley's counting formula for labelled trees, first published in [29]. A tree is a connected graph that does not contain a cycle and a forest is a set of trees. The famous book [59] mainly presents results on various types of unlabelled graphs and Polya's Enumeration Theorem [2, p.253] is applied to count unlabelled structures. Unlabelled structures are generally harder to analyse than labelled structures and they will only be mentioned occasionally in this thesis.

Parallel to the development on labelled random planar graphs, a more probabilistic approach was used to answer similar questions for random graphs on surfaces [75]. In 2011 Chapuy et al. counted graphs embeddable on surfaces of fixed genus [30] applying generating function methods and similar techniques to proofs presented in [54]. A graph is *series-parallel* if its 2-connected components are obtained from a single edge by recursive subdivision of edges and duplication of edges. A graph is *outerplanar* if it is planar and can be embedded in the plane in such a way that all of the vertices belong to the unbounded face of the drawing. Similar results to the ones proved for planar graphs were also obtained for other graph classes such as outerplanar, series-parallel graphs [22] and $K_{3,3}$-free graphs [51].

Various other parameters, such as the maximum degree in a random planar graphs [75], [43], its degree distribution [41] as well as the maximum degree in series-parallel graphs [42] and the distribution of vertices of a given degree in series-parallel graphs [40] have been studied.

Furthermore, looking at the world from a more graph theoretical rather than enumerative angle, it was observed that several of the properties of planar graphs – for instance the fact that asymptotically the probability that $R_n \in_u \mathcal{A}$ is connected is strictly bounded away from 0 – generalize to much larger classes of graphs. A class $\mathcal{A}$ is called *bridge-addable* if for every graph G in $\mathcal{A}$ and vertices u, v in different components of G the graph obtained by adding the edge uv is still in $\mathcal{A}$. In [76] it was proven that for every bridge-addable class of graphs $\mathcal{A}$ and for each n

$$\mathbb{P}(R_n \text{ connected}) \geq e^{-1} \tag{1.1}$$

holds. Clearly, the class of planar graphs is bridge-addable. It was conjectured in [76] that (1.1) is not tight (see [7], [62], [1] for some recent developements). We will generalize the result (1.1) to further graph classes and discuss related conjectures in detail in Chapter 4. Furthermore, 'weighted' bridge-addable random graphs sampled from a suitable, more general distribution (where the sampling probablity depends on the number of bridges in a graph as well as the number of components) than the uniform distribution were considered in [73]. Various bounds known for random graphs sampled uniformly from bridge-addable classes of graphs were amplified and generalised.

One other characteristic feature of planar graphs is that they can be described by forbidden minors. A *minor* of a graph G is a subgraph that is obtained by contracting and deleting edges, deleting isolated vertices and deleting all loops and multiple edges. A class of graphs $\mathcal{A}$ is called *minor-closed* if for every graph G in $\mathcal{A}$ every minor of G is also in $\mathcal{A}$. To each such class $\mathcal{A}$ one can associate a set of excluded minors S: an (unlabelled) graph is excluded if its labelled versions do not belong to $\mathcal{A}$ but the labelled versions of each of its proper minors do. The class of planar graphs $\mathcal{A}$ is minor-closed and by Kuratowski's/Wagner's Theorem [36, Theorem 4.4.6] we know that a graph is planar if and only if it does not contain either K_5 or $K_{3,3}$ as a minor. The famous Robertson and Seymour Theorem, proved in a series of papers [86], states that any infinite sequence $G_1, G_2, \ldots$ of finite graphs includes two graphs G_i and G_j with $i < j$ such that G_i is a minor of G_j. It is a straightforward consequence that if $\mathcal{A}$ is minor-closed the number of forbidden or excluded minors $|S|$, is always finite. Thus every minor-closed class has a characterization in terms of its excluded minors which is of Kuratowski-Wagner type. We write $\mathcal{A} = \mathrm{Ex(S)}$ for the minor-closed class $\mathcal{A}$ defined by excluding the minors in S, for example for planar graphs $\mathcal{A} = \mathrm{Ex}(\{\mathrm{K}_5, \mathrm{K}_{3,3}\})$. Many other graph classes mentioned at the beginning of this chapter, such as series-parallel graphs ($= \mathrm{Ex}(\mathrm{K}_4)$), graphs embeddable on surfaces and outerplanar graphs ($= \mathrm{Ex}(\{\mathrm{K}_4, \mathrm{K}_{2,3}\})$) are minor-closed too.

Are generating functions generally a good means to analyse minor-closed classes of graphs? Speaking in the language of analytic combinatorics we ask the question whether minor-closed classes are *constructible*, i.e. do they admit a finite recursive specification that starts from single atoms and involves only the standard SUM, PRODUCT, SET and CYC constructions? The relevant constructions will be introduced in Chapter 2 and we refer to [49, p.250] for a formal definition of constructible and related results. The Theorem of Computability of Growth [49, Theorem IV.8] states that for every constructible class (labelled or unlabelled) the radius of convergence of the corresponding generating function (either exponential

or ordinary) is either $+\infty$ or a strictly positive computable constant. A class of graphs is *proper* if it does not contain all graphs. Let A_n denote the number of n-vertex graphs in a labelled class $\mathcal{A}$. In 2006 [80] it was shown that for every labelled, proper, minor-closed class of graphs

$$\sup_{n\in\mathbb{N}} \left(\frac{A_n}{n!}\right)^{1/n} < \infty. \tag{1.2}$$

Hence the (exponential) generating function corresponding to the class $\mathcal{A}$ denoted by $A(z)$ (see Chapter 2 for a formal definition) has a strictly positive radius of convergence and minor-closed classes of graphs are certainly not 'too big' to be analysed by generating functions. The result was extended to unlabelled proper minor-closed classes of graphs in 2010 [47] (obviously without the normalizing factor $n!$). We say that $\mathcal{A}$ has a *growth constant* γ if $\lim_{n\to\infty} \left(\frac{A_n}{n!}\right)^{1/n}$ in the labelled and $\lim_{n\to\infty} (A_n)^{1/n}$ in the unlabelled case exists, respectively.

In [76] it was shown that the class of (labelled) planar graphs has a growth constant. Following [65], a class of graphs $\mathcal{A}$ is *decomposable* if a graph G is in $\mathcal{A}$ if and only if all its component are in $\mathcal{A}$. A few years later in [71] it was shown that more generally every minor-closed, decomposable, bridge-addable class of graphs has a growth constant. The same is true for graphs (labelled and unlabelled) embeddable on fixed surfaces [70]. In [20] it was conjectured that in fact every (labelled) minor-closed class of graphs has a growth constant. We refer to the same paper and to [55] for a discussion of possible growth constants of labelled graphs. Some of the results obtained in [20] even extend to the unlabelled case [19]. It is worth noting that in the labelled as well as in the unlabelled case the set of possible growth constants Γ of minor-closed classes of graphs is countable, as the number of excluded minors S is always finite as previously noted.

In [76] it was conjectured that for planar graphs nA_{n-1}/A_n tends to a limit as $n \to \infty$ where A_n is the number of labelled planar graphs on n vertices. Observe that if nA_{n-1}/A_n tends to a limit it has to be γ^{-1}, the inverse of the growth constant. Following [70], a labelled class $\mathcal{A}$ is called *smooth* if nA_{n-1}/A_n tends to a limit as $n \to \infty$ and planar graphs are indeed smooth [54]. Other classes that are smooth include series-parallel [22], outerplanar graphs [22], apex graphs [70], and several other minor-closed classes of graphs [51], [20]. In each case this is because we know the exact asymptotic counting formula or, going back one step further, we know the asymptotic expansion of the generating function next to its radius of convergence, which makes obtaining smoothness as well has having a growth constant trivial. As it turns out many properties of a class of graphs $\mathcal{A}$ are only accessible if $\mathcal{A}$ is smooth (see for example [70, Section 3]). In [16]

it was shown that graphs embeddable on a fixed surface are smooth, settling a conjecture of [70]. In fact a general tool for proving smoothness of functional compositions was proved thus providing a first means of obtaining smoothness in further cases. Graphs embeddable on a fixed surface were counted asymptotically in [30] generalizing results obtained for planar graphs in [53]. In [71] it was shown using the result of [16] that more generally minor-closed, decomposable, bridge-addable classes of graphs are smooth and the question has arisen whether the same is true for any minor-closed class of labelled graphs. As already mentioned above this would imply the conjecture of [20] is valid, that every minor-closed class has a growth constant. Smoothness of classes of combinatorial objects has also been a property of interested in other contexts more closely related to logic and number theory. It crops up as a key property in finding sufficient conditions for labelled and unlabelled $0-1$ limit laws and in particular when trying to determine the limit law of the number of components in general decomposable structures (see [26], [12], [14], [27], [13]). For an overview of related work in number theory, see the book [25]. We observe that no general smoothness result is known for unlabelled, minor-closed classes of graphs.

Many properties of planar graphs – other than smoothness – generalise to minor-closed, decomposable, bridge-addable classes of graphs sometimes referred to as minor-closed, *addable* classes of graphs [76]. One example of such a property is the probability that $R_n \in_u \mathcal{A}$ is connected is asymptotically strictly larger than 0. Observe that a minor-closed class of graphs is decomposable if and only if all its excluded minors are connected. We refer to [67] and [66] for results on non-decomposable minor-closed classes of (labelled) graphs and to [61] for an example of an unlabelled non-decomposable class of graphs. A minor-closed class is bridge-addable if its excluded minors have no cut point or edge component. Therefore minor-closed, decomposable, bridge-addable classes of graphs are precisly the minor-closed classes of graphs that can be characterised by excluding only 2-connected minors. We observe that for all the minor-closed classes of graphs mentioned so far the excluded minors are 2-connected apart from [67], [66], [61]. But even in these papers the analysed minor-closed classes of graphs are still bridge-addable and in the labelled cases [67] and [66] the obtained results are similar to what we know for for minor-closed, decomposable, bridge-addable classes of graphs, i.e. the probability that $R_n \in_u \mathcal{A}$ is connected is still bounded away from 0.

Following [71] let Big(G) of a graph G denote the (lexicographically first) component with the most vertices and Frag(G) the subgraph induced on the vertices not in the largest component (which may be empty if G is connected). Let $\kappa(G, H)$

denote the number of components of G isomorphic to H (H is unlabelled). Suppose we are interested in a labelled class of graphs $\mathcal{A}$, with $\rho > 0$ and $A(\rho) < \infty$. Let $\mathcal{UA}$ denote the corresponding unlabelled class and choose $\rho > 0$ such that $A(\rho)$ is finite. Define $\lambda(H) = \frac{\rho^{|H|}}{\text{aut}(H)}$. We can write $A(\rho) = \sum_{H \in \mathcal{UA}} \lambda(H)$. Then the Boltzmann Poisson random graph BP takes values in $\mathcal{UA}$ with

$$\mathbb{P}(BP = H) = \frac{\lambda(H)}{A(\rho)}. \tag{1.3}$$

In [71] it was shown that if $\mathcal{A}$ is labelled, decomposable and closed under isomorphisms, the random variables $\kappa(BP, H)$ for $H \in \mathcal{UC}$ are independent. Furthermore, $\kappa(BP, H)$ is Poisson distributed with parameter $\lambda(H)$.

Furthermore, it was shown that if $\mathcal{A}$ is decomposable, minor-closed and bridge-addable with radius of convergence $\rho > 0$ then $A(\rho) < \infty$. Also the limiting distribution of the random graph $\text{Frag}(R_n)$ was obtained. Let F_n denote the unlabelled random graph corresponding to $\text{Frag}(R_n)$ where R_n was taken uniformly at random from all graphs in $\mathcal{A}$ on n vertices. In [71] it was shown that if $\mathcal{A}$ is minor-closed, decomposable and bridge-addable then $F_n \to BP$ pointwise in probability where BP is the Boltzmann Poisson random graph for $\mathcal{A}$ (where the parameter ρ is equal to the radius of convergence of $\mathcal{A}$). From this, the limit law for the number of components as well as the limit law of the largest component can be obtained. We will spell these results out further in Chapter 3. Furthermore, for any minor-closed, decomposable, bridge-addable class of graphs $\mathbb{E}(\text{frag}(R_n)) < \infty$ where $\text{frag}(G)$ counts the number of vertices in $\text{Frag}(G)$. In [73] it was proved that in fact for any bridge-addable class of graphs $\mathbb{E}(\text{frag}(R_n)) < \infty$. As already mentioned we know that for any minor-closed class of graphs $\mathcal{A}$ with connected substructures $\mathcal{C}$ where all excluded minors are 2-connected, the corresponding generating function $A(\rho) < \infty$ and even $A'(\rho) < \infty$ and hence $C(\rho)$ and $C'(\rho)$ are always finite too. These results are taken from [71] and their relevance is discussed in Chapter 2. Observe that whether $C(\rho)$ is finite or not plays a crucial role for the limiting probability that R_n is connected and more generally for the distribution of the number of components of $R_n \in_u \mathcal{A}$. In fact, there is a wealth of literature dealing with general decomposable combinatorial structures and what happens if $C(\rho) < \infty$ and $C(\rho) = \infty$ without knowing the exact asymptotic expansion of the generating function next to its radius of convergence. The research started as early as the 1960s by Wright [91], [92] and has received quite a lot of attention since then (see [27], [11], [10],[31]). The most important results in this context are stated in Section 2.2.

One of the fundamental principles of analytic combinatorics states that the *nature of the singularity* of $C(z)$ at ρ, i.e. the precise asymptotic behaviour (the dominant term in a singular expansion) next to its singularity depends on the class. More precisely the dominant singularity, i.e. the dominant singularity in the expansion, for forests behaves like $(1 - z/\rho)^{3/2}$ and more generally so-called subcritical graph classes as defined in [39] have the same type of singularity. Subcritical class are not necessarily minor-closed. A *block* of a graph is a maximal induced 2-connected subgraph. Following [39] a class of graphs $\mathcal{A}$ is *block-stable* if a graph G is in $\mathcal{A}$ if and only if all its blocks are in $\mathcal{A}$. A minor-closed class of graph can only be subcritical if it is block-stable and a minor-closed class is *block-stable* if and only if its excluded minors are 2-connected. Hence a minor-closed class of graphs defined by excluding non-2-connected minors can never be block-stable. Subcritical classes include outerplanar graphs and series-parallel graphs. It should also be observed that the results of [39] extend to the case of unlabelled subcritical classes of graphs. It is worth noting that the nature of the dominant singularity of planar graphs is different to that of forests and behaves like $(1 - z/\rho)^{5/2}$ (see [55]). A more detailed discussion and an approach generalizing results on planar graphs, series-parallel graphs and outerplanar graphs, which applies to any minor-closed class that exclude 3-connected minors, can be found in [56]. We observe again that all mentioned classes of graphs as well as all the tools for analysing them require that the classes are bridge-addable. A first example of an analysis of a minor-closed class of graph that is not bridge-addable can be found in [72]. It was mentioned in [71] that a non-bridge-addable minor-closed class of graphs might behave rather differently. For further recent literature references as well as for results for random graphs sampled from weighted minor-closed classes of graphs see [74].

1.1. Overview of thesis

In this thesis, I will investigate what happens if the assumption that $\mathcal{A}$ is bridge-addable is relaxed. This will be done combining probabilistic, graph theoretical methods as well as tools from analytic combinatorics.

In Chapter 2 the notation used will be provided as well as a brief overview of standard techniques from analytic combinatorics. We will start in Section 2.1 with some generating function methods for analysing graphs in particular, an explanation of the symbolic method and deduce some important relationships between classes of graphs and their generating functions. In Section 2.2, much in the spirit of Wright, we will have a closer look at general tools that allow us to deduce

asymptotic properties of the class from the generating function of a class without knowing anything about the specific analytic properties of the expansion of the function around its radius of convergence. Afterwards in Section 2.3 we give an introduction to singularity analysis. Singularity is a standard method which provides access to a function's coefficients for functions which do not grow too fast next to their singularities. The counterpart for functions that grow very fast next to their singularities – called the saddle point method – is explained in Section 2.4. This section includes a new closure property for Hayman-admissible functions as well as a result on the number of components of $R_n \in_u \mathcal{A}$ in the case when the generating function is Hayman-admissible. Following this, in Section 2.6, the method is applied in one particular case as the results will be useful in Chapter 3.

In Chapter 3 we will examine various examples of minor-closed classes of graphs that are defined by excluding non-2-connected minors. We will focus on the asymptotic number of the graphs, the probability that $R_n \in_u \mathcal{A}$ is connected, the number of components of R_n, the size of the largest component as well as the size of the root component (=the component containing the vertex labelled 1) in $R_n \in_u \mathcal{A}$. As we will see, a whole range of different types of behaviours can be observed and the results are summarized in Table 3.1 at the end of the chapter.

In Chapter 4 we will discuss what happens if we do not have a 'fully' bridge-addable class of graphs but a class of graphs where only certain bridges (which are in a previously fixed host graph) are allowed to be inserted. We will see that as long as the host graph has good expander properties, i.e. we are allowed to introduce quite a few possible bridges without leaving the class, the behaviour is similar to the bridge-addable case. We will obtain a lower bound on the probability that R_n is connected and investigate whether it is tight or not. In this conext, detailed results about random forests in $K_{n,n}$ are obtained.

In Chapter 5 a different approach is taken. We fix $\mathcal{A}$ to be the class of labelled forests. But now some vertices belong to a stable set – we can think of them as being coloured – and the probablity that a specific forest is sampled depends on the size of the stable set as well as on some fixed weighting parameter y. We will see that as long as y is finite the behaviour is similar to random forests without a stable set but it changes in the limiting case.

We finish with Chapter 6, in the spirit for Wright. Taking a labelled or unlabelled decomposable, smooth combinatorial class, we prove a result on the number of components of fixed size. From this we deduce results about the distribution of the smallest component in these structures, which depends on whether the

generating function of the connected structures converges or diverges at its radius of convergence.

2. Generating function methods

2.1. Generatingfunctionology for graphs

Let $\mathcal{A}$ be a set of graphs. If not stated otherwise, we will assume that $\mathcal{A}$ contains labelled graphs. We denote by $\mathcal{A}_n$ the subset of $\mathcal{A}$ formed of graphs with n vertices, and by A_n the cardinality of $\mathcal{A}_n$. The associated exponential generating function is $A(z) = \sum_{n\geq 0} A_n z^n/n!$ and we write $A_n = [z^n]A(z)$ for its n^{th} coefficient. We use similar notation (C_n and $C(z)$) for the subset $\mathcal{C}$ of $\mathcal{A}$ consisting of connected graphs. Let us very briefly review how one can use combinatorial constructions to build complex labelled classes from simpler ones and how they correspond to operations on the corresponding generating functions (see [49, Chapter II] for more details). Suppose we have three classes $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ of labelled combinatorial objects with generating functions $A(z) = \sum_{n\geq 0} A_n z^n/n!$, $B(z) = \sum_{n\geq 0} B_n z^n/n!$ and $C(z) = \sum_{n\geq 0} C_n z^n/n!$, respectively. Suppose $A(z) = B(z) \cdot C(z)$ holds. For the number of objects in $\mathcal{A}_n$ this translates into $A_n = \sum_{k=0}^{n} \binom{n}{k} B_k C_{n-k}$, i.e. an object of size n in $\mathcal{A}$ is an ordered pair of one object from $\mathcal{B}$ and one from $\mathcal{C}$ which sizes sum up to n. For the sets we write $\mathcal{A} = \mathcal{B} \times \mathcal{C}$ and $\times$ is called the *labelled product* (see [49, p. 101] for a formal definition). Hence we find the basic property

$$\mathcal{A} = \mathcal{B} \times \mathcal{C} \Rightarrow A(z) = B(z) \cdot C(z). \tag{2.1}$$

This is the principal idea of the symbolic method: that there exists a one-to-one correspondence of operations on combinatorial classes and their corresponding generating functions. In the same spirit, we define the sequence class of $\mathcal{C}$ to be $\text{SEQ}(\mathcal{C}) = \epsilon + \mathcal{C} + \mathcal{C} \times \mathcal{C} + \ldots$, where $+$ for a class is equivalent to the disjoint union and ϵ is the empty class. Then using (2.1) we obtain

$$\mathcal{A} = \text{SEQ}(\mathcal{C}) \Rightarrow A(z) = \sum_{k=0}^{\infty} C(z)^k = \frac{1}{1 - C(z)}.$$

A 'set' of objects is like a sequence of objects but the components – objects of $\mathcal{C}$ – are not ordered between themselves. Let $\text{SEQ}_k(\mathcal{C})$ denote all sequences of fixed

length k of $\mathcal{C}$. We define $\mathrm{SET}_k(\mathcal{B}) = \mathrm{SEQ}_k(\mathcal{C})/R$ where the equivalence relation R identifies two sequences when the components of one are the permutation of the components of the other. Let $\mathrm{SET}(\mathcal{C}) := \bigcup_{k\geq 0} \mathrm{SET}_k(\mathcal{C})$. Then

$$\mathcal{A} := \mathrm{SET}(\mathcal{C}) \Rightarrow A(z) = e^{C(z)}. \tag{2.2}$$

Let us now go back to our starting point, so let $\mathcal{A}$ be a class of labelled graphs. In general, we will assume that $\mathcal{A}$ is decomposable in the sense of Kolchin [65]: a graph belongs to $\mathcal{A}$ if and only if every component belongs to $\mathcal{A}$. This fact is precisely captured by $\mathcal{A} = \mathrm{SET}(\mathcal{C})$: every graph corresponds to a set of connected components which are also in $\mathcal{A}$ that are not ordered between themselves. By (2.2) we have

$$A(z) = \exp(C(z)).$$

Several variants of this series are of interest, for instance the bivariate generating function that keeps track of the number of (connected) components. Let $A_{n,k}$ be the number of graphs on n vertices in $\mathcal{A}$ with k components. Then

$$\begin{aligned} A(z,u) &= \sum_{n\in\mathbb{N}} \sum_{k=1}^{n} \frac{A_{n,k}}{n!} z^n u^k \\ &= \sum_{G\in\mathcal{A}} u^{c(G)} \frac{z^{|G|}}{|G|!} \end{aligned}$$

where $|G|$ is the size of G and $c(G)$ the number of its components. We say that u *marks* the number of components. In this case by an argument similar to (2.2)

$$A(z,u) = \exp(uC(z)).$$

The following lemma links the structural behaviour of a class to its bivariate generating function. Let $R_n \in_u \mathcal{A}$ denote a graph taken uniformly at random from the set $\mathcal{A}_n$.

Lemma 2.1.1 (Prop.III.3 in [49])**.** *Let χ be a parameter of $R_n \in_u \mathcal{A}$ and let $A(z,u)$ be the corresponding bivariate generating function, where u marks parameter χ. The r^{th} factorial moment of χ is determined by the bivariate generating function $A(z,u)$ by r-fold differentiation followed by evaluation at 1:*

$$\mathbb{E}\left(\chi(\chi-1)\dots(\chi-r+1)\right) = \frac{[z^n]\partial_u^r A(z,u)|_{u=1}}{[z^n A(z,1)]}.$$

Let $\kappa(R_n)$ denote the number of components of $R_n \in_u \mathcal{A}$. Then

$$\mathbb{P}(\kappa(R_n) = i) = \frac{[z^n]C(z)^i}{i![z^n]A(z)} \tag{2.3}$$

and by Lemma 2.1.1 the i^{th} factorial moment of $\kappa(R_n)$ is

$$\mathbb{E}(\kappa(R_n)(\kappa(R_n)-1)\cdots(\kappa(R_n)-i+1)) = \frac{[z^n]\frac{\partial^i A}{\partial u^i}(z,1)}{[z^n]A(z)} = \frac{[z^n]C(z)^i A(z)}{[z^n]A(z)}.$$

Several general results provide a limit law for $\kappa(R_n)$ if $C(z)$ satisfies certain conditions. For some examples we refer to [49, Chapter IX].

Furthermore, in Chapter 3, we will study the size $s(G)$ of the *root component*, which is the component containing the vertex with label 1. This parameter was considered for the first time in our paper [24]. We define

$$\bar{A}(z,v) = \sum_{G\in\mathcal{A},G\neq\emptyset} v^{s(G)-1}\frac{z^{|G|-1}}{(|G|-1)!}$$

where $s(G)$ denotes the size of the component that contains the vertex 1. The choice of $|G|-1$ instead of $|G|$ slightly simplifies some calculations.

Lemma 2.1.2 (Generating function of the root component). *The generating function $\bar{A}(z,v)$ is given by*

$$\bar{A}(z,v) = C'(zv)A(z) \tag{2.4}$$

where $A(z)$ is the generating function of all graphs and $C(z)$ the one of connected ones. In particular,

$$\bar{A}(z,1) = A'(z) = C'(z)A(z).$$

Let S_n denote the size of the root component in $R_n \in_u \mathcal{A}$. We have

$$\mathbb{P}(S_n = k) = \frac{C_k A_{n-k}\binom{n-1}{k-1}}{A_n} = \frac{k}{n}\frac{C_k}{k!}\frac{A_{n-k}}{(n-k)!}\frac{n!}{A_n}. \tag{2.5}$$

Proof. We prove (2.5) first. To construct a graph with $S_n = k$ we have $\binom{n-1}{k-1}$ to choose the $k-1$ vertices of the connected component containing the root, C_k choices to pick a connected graph on these k vertices and A_{n-k} choices to build a graph on the remaining vertices. Hence (2.5) follows.

Let $\mathcal{A}^\bullet$ denote rooted graphs, i..e. in every graph every component has a unique specified vertex, and $\mathcal{C}^\bullet$ rooted connected graphs, respectively. A graph in $\mathcal{A}^\bullet$ is a set of connected components, where one component is rooted. Hence

$$\mathcal{A}^\bullet = \mathcal{C}^\bullet \times \mathcal{A}$$

and rooting connected graphs corresponds to the pointing operation [49, Chapter II, p.136]

$$\mathcal{C}_n^\bullet = [1, \ldots, n] \times \mathcal{C}_n \Rightarrow C^\bullet(z) = z \cdot C'(z).$$

Let v mark the number of vertices in the component containing the root and z the number of components in the whole graph. We get (2.4). □

A consequence of the Lemma is the following observation

Corollary 2.1.3 (**Factorial moments of the root component**). *The i^{th} factorial moment of $S_n - 1$ is*

$$\mathbb{E}((S_n - 1) \cdots (S_n - i)) = \frac{[z^{n-1}]\frac{\partial^i \bar{A}}{\partial v^i}(z, 1)}{[z^{n-1}]\bar{A}(z, 1)} = \frac{[z^{n-i-1}]C^{(i+1)}(z)A(z)}{n[z^n]A(z)}. \tag{2.6}$$

Proof. Apply Lemma 2.1.1 to the formula given in Lemma 2.1.2. □

In fact, the form (2.4) of the generating function shows that this parameter is bound to give rise to interesting (i.e. non-Gaussian, which are rare in analytic combinatorics) limit laws, as both the location and nature of the singularity change as v moves from $1-\varepsilon$ to $1+\varepsilon$. Using the terminology of Flajolet and Sedgewick [49, Sec. IX.11], a *phase transition* occurs. A more systematic study of this parameter than presented in this thesis is currently work in progress with M. Bousquet-Mélou. This project is also reminiscent of the study of the 2-connected component containing the root vertex in a planar map, which also leads to a non-gaussian continuous limit law, namely an Airy distribution [8]. This distribution is also related to the size of the largest 2- and 3-connected components in various classes of graphs [56].

Finally, we denote by $C^{[k]}(z)$ the generating function of connected graphs of $\mathcal{A}$ of size less than k:

$$C^{[k]}(z) = \sum_{n=1}^{k-1} C_n \frac{z^n}{n!}.$$

This gives access to the size L_n of the largest component. We have

$$\mathbb{P}(L_n < k) = \frac{[z^n]\exp(C^{[k]}(z))}{[z^n]A(z)}. \tag{2.7}$$

2.2. Generating functions next to their radius of convergence

There are a few general results that are applicable to a wide range of generating functions and that do not require checking analyticity or similar properties. We will see that in all cases the behaviour of the function at its radius of convergence is crucical. The following is a simple but useful result on products of series, see also [15, Thm. 2]. A class $\mathcal{A}$ is *smooth* if $\lim_{n\to\infty} \frac{nA_{n-1}}{A_n}$ in the labelled case and $\lim_{n\to\infty} \frac{A_{n-1}}{A_n}$ in the unlabelled case exists.

Lemma 2.2.1 (Schur [15])**.** *Let $A(z)$ and $B(z)$ be two power series and let $A(z)$ be smooth with $R_A > 0$. Suppose $B(z)$ has larger radius of convergence than $A(z)$ and $B(R_A) \neq 0$. Then*

$$[z^n]B(z)A(z) \sim B(R_A)[z^n]A(z).$$

In particular $B(z)A(z)$ is smooth.

As seen in the previous lemma, smoothness is a very important property. In 2003 Burris and Bell (see theorem below) proved a useful tool for providing smoothness for any decomposable class of graphs, labelled or unlabelled : it suffices to analyse the connected structures. In the unlabelled case the equivalent of the SET construction explained in Section 2.1 is the $MSET$ construction. We have

$$\mathcal{A} = MSET(\mathcal{C}) \Rightarrow A(z) = e^{\sum_{k=1}^{\infty} \frac{C(z^k)}{k}}.$$

Theorem 2.2.2 (**Burris and Bell — smoothness for decomposable classes**)**.** *Suppose $\mathcal{C}$ is labelled or unlabelled and smooth. Then $\mathcal{A} = \mathrm{SET}(\mathcal{C})$ in the labelled and $\mathcal{A} = MSET(\mathcal{C})$ in the unlabelled case respectively, is smooth. Thus, for decomposable classes it is enough to prove smoothness for the connected substructures.*

Proof. For a proof of the labelled case see [13]. A proof for the unlabelled case can be found in [14]. □

This for instance simplifies the proof for smoothness for labelled proper, minor-closed decomposable, bridge-addable classes of graphs given in [71]. Recall that a graph class is *proper* if it does not contain all graphs. A general tool how to prove smoothness for functional compositions of combinatorial class can be found in [16].

We now focus on decomposable labelled classes and the probability that the random graph $R_n \in_u \mathcal{A}$ is connected and its number of components, $\kappa(R_n)$.

The following theorem tells us that in the *converging case*, i.e. when $C(z)$ converges at $z = \rho$ ([9]), the number of components of $R_n \in_u \mathcal{A}$ is very likely to have Poisson-type behaviour.

Theorem 2.2.3 ([10])**.** *Suppose $\mathcal{A}$ is decomposable with radius of convergence $\rho > 0$, $C(\rho) < \infty$ and $\mathcal{C}$ is smooth. Then the following statements are equivalent:*

1. $\frac{n![z^n](C(z))^2}{C_n} \sim 2 \cdot C(\rho)$
2. *If* $\omega(n) \to \infty$ *then* $\sum_{k=\omega}^{n-\omega} \frac{C_k}{(k!)} \frac{C_{n-k}}{(n-k)!} = o\left(\frac{C_n}{n!}\right)$
3. $\lim_{n\to\infty} \frac{C_n}{A_n}$ *exists.*
4. *Let $\kappa(R_n)$ denote the number of components of $R_n \in_u \mathcal{A}$. Then*

$$\kappa(R_n) \xrightarrow{d} 1 + \mathrm{Po}(C(\rho)) \tag{2.8}$$

for $n \to \infty$ and

$$\mathbb{E}(\kappa(R_n)) \to 1 + C(\rho). \tag{2.9}$$

In particular for $R_n \in_u \mathcal{A}$

$$\lim_{n\to\infty} \mathbb{P}(R_n \in_u \mathcal{A} \text{ is connected}) \to e^{-C(\rho)}. \tag{2.10}$$

For classes that fulfil the assumptions of Theorem 2.2.2 we expect that they only have few components so the largest component probably contains almost all vertices. It is not hard to prove that the same holds for the root component (they are identical). In fact, the tails of the random variables S_n and L_n are related by the following simple result.

Lemma 2.2.4. *Let $\mathcal{A}$ be a decomposable, smooth class with $\rho > 0$, $C(\rho) < \infty$. Assume that $\lim_{n\to\infty} \mathbb{P}(R_n \text{ is connected})$ exists. For $k < n/2$,*

$$\mathbb{P}(S_n = n - k) = \frac{n-k}{n}\,\mathbb{P}(L_n = n - k).$$

Proof. Let us denote by B_n the (lexicographically first) biggest component of R_n. Its size is thus L_n. We have, for $n > 2k$,

$$\begin{aligned}
\mathbb{P}(S_n = n-k) &= \mathbb{P}(S_n = n-k \text{ and } 1 \in B_n) + \mathbb{P}(S_n = n-k \text{ and } 1 \notin B_n)\\
&= \mathbb{P}(L_n = n-k \text{ and } 1 \in B_n) + \mathbb{P}(S_n = n-k \text{ and } 1 \notin B_n)\\
&= \mathbb{P}(1 \in B_n | L_n = n-k)\mathbb{P}(L_n = n-k) + \mathbb{P}(S_n = n-k \text{ and } 1 \notin B_n)\\
&= \frac{n-k}{n}\,\mathbb{P}(L_n = n-k).
\end{aligned}$$

Indeed, there cannot be two components of size $n-k$ or more. This implies that $\mathbb{P}(S_n = n-k \text{ and } 1 \notin B_n) = 0$. □

Proposition 2.2.5 (The root component and the largest component — converging case). *Let $\mathcal{A}$ be a decomposable, smooth class with $\rho > 0$, $C(\rho) < \infty$ and assume that $\lim_{n\to\infty} \mathbb{P}(R_n \text{ is connected})$ exists Let $k \in \mathbb{N}$ be fixed. The random variables $n - S_n$ and $n - L_n$ both converge to a discrete limit distribution X given by*

$$\mathbb{P}(X = k) = \frac{1}{A(\rho)}\frac{A_k \rho^k}{k!}.$$

Proof. By Lemma 2.2.4, the two statements are equivalent. Let $k \geq 0$ be fixed. By (2.5),

$$\mathbb{P}(S_n = n-k) = \frac{C_{n-k} A_k \binom{n-1}{k}}{A_n} = \frac{A_k}{k!}\,\frac{C_{n-k}}{A_{n-k}}\,\frac{(n-1)!A_{n-k}}{(n-k-1)!A_n}.$$

By Theorem 2.2.3, the term C_{n-k}/A_{n-k}, which is the probability that a graph of size $n-k$ is connected, converges to $1/A(\rho)$. Moreover, the class $\mathcal{A}$ is smooth, so that $\frac{(n-1)!A_{n-k}}{(n-k-1)!A_n}$ converges to ρ^k. The result follows. □

The result on L_n has been proved by McDiarmid [71, Cor. 1.6] for minor-closed classes of graphs where the excluded minors are 2-connected.

Let $\mathrm{frag}(R_n)$ be the number of vertices of R_n that are not in the largest component.

Proposition 2.2.6. *Let $\mathcal{A}$ be a decomposable, smooth class with $\rho > 0$, $C(\rho) < \infty$ and $\lim_{n\to\infty} \frac{C_n}{A_n}$ exists. Then as $\mathrm{frag}(R_n) \xrightarrow{d} X$ where X is given in Proposition 2.2.5 we have*

$$\mathbb{E}(X) = \frac{\partial C}{\partial z}(z)\bigg|_{z=\rho} \rho$$

and

$$\liminf_{n\in\mathbb{N}} \mathbb{E}(\mathrm{frag}(R_n)) \geq \mathbb{E}(X) = \rho C'(\rho).$$

Proof. The first part is a simple consequence of the distribution of X (Proposition 2.2.5)

$$\begin{aligned}\mathbb{E}(X) &= \frac{1}{A(\rho)} \cdot \sum_{k=1}^{\infty} \frac{kA_k\rho^k}{k!} \\ &= \frac{1}{e^{C(\rho)}\rho} \frac{\partial A}{\partial z}(z)\bigg|_{z=\rho} = \rho \frac{\partial C}{\partial z}(z)\bigg|_{z=\rho}.\end{aligned}$$

For a fixed k_0 we can write

$$\liminf_{n\in\mathbb{N}} E(\mathrm{frag}(R_n)) \geq 1/A(\rho) \sum_{k=1}^{k_0} \frac{kA_k\rho^k}{k!}$$

and thus deduce that

$$\liminf E(\mathrm{frag}(R_n)) \geq E(X) = \rho C'(\rho).$$

□

Observe, that in particular Proposition 2.2.6 shows that if $C'(\rho)$ diverges, the expectation of the fragment is unbounded. As we will see later, this for example applies to forests of rooted trees.

In the diverging case, i.e. when $C(\rho) = \infty$, the following lemma, which can be found in [10], states that the behaviour is likely to change drastically.

Lemma 2.2.7. *Suppose $\mathcal{A}$ is decomposable, smooth and $\rho > 0$. If $C(\rho)$ diverges, then*

$$\lim_{n\to\infty} \mathbb{P}\left(R_n \in \mathcal{A} \text{ is connected }\right) = 0.$$

Unfortunately, no equivalent to Theorem 2.2.3 is known. From our examples, it would be natural to conjecture that for a decomposable, smooth class $\mathcal{A}$ with $\rho > 0$ and $C(\rho) = \infty$ (and maybe further conditions) the number of components should follow a Gaussian limit distribution.

2.3. Singularity Analysis

In Section 2.1 a general method how to obtain the generating function of a given class was presented. In the following section we will explain why the generating functions tells us a lot about the underlying class. The premises of analytic combinatorics is simple – a function's singularities are reflected in the function's coefficients. Therefore, if we have a generating function and would like to deduce the asymptotics of its coefficients it suffices to look at the behaviour of the function around its singularities. More precisely, we get access to the coefficients via the following theorem known as Cauchy's coefficient formula ([49] Theorem IV.4). The proof is a consequence of the Residue Theorem.

Theorem 2.3.1 (**Cauchy's coefficient formula).** *Let $f(z)$ be analytic in a region Ω containing 0 and let λ be a simple loop around 0 in Ω that is positively oriented. Then the n^{th} coefficient $[z^n]f(z)$ admits the integral representation*

$$[z^n]f(z) = \frac{1}{2\pi i}\int_\lambda f(z)\frac{\partial z}{z^{n+1}}.$$

Hence the problem of how to find the coefficients of a generating function can be reformulated. Suppose one wants to approximate a complex integral of an analytic function, i.e. an integral of the form

$$F(n) = \int_P \phi(z)dz \tag{2.11}$$

where P is an integration path in the complex plane and $\phi(z)$ is analytic in some neighbourhood of P. Both the integration path P and the function $\phi(z)$ depend on a real (large) parameter n and we are interested in the asymptotic behaviour $F(n)$ when $n \to \infty$.

In [48], Flajolet et al. presented an approach to analyse the coefficient of a large class of generating functions by evaluating the corresponding integral. Their method applies to a large class of functions of moderate growth/decay at the unique dominant singularity (=the singularity of minimal modulos). The rough idea can be explained as follows. We want to analyse the coefficients of a function $A(z)$. Suppose we already have a catalogue of functions of which coefficients can be easily analysed asymptotically. Furthermore, suppose that our function $A(z)$ has a unique dominant singularity and we have an asymptotic expansion of our function $A(z)$ near this singularity which is valid in a Δ-domain, i.e. an area of the complex plane extending beyond the disc of convergence of the original series (that looks like a camembert) (see 2.3.3 for a formal definition). The asymptotic expansion should

only contain simpler functions which are covered by our catalogue. The theorem of singularity analysis then states that the coefficients of $A(z)$ asymptotically behave like the coefficients of the asymptotic approximation of $A(z)$. Therefore, using our catalogue the coefficients are easy to deduce.

More precisely, we have the following catalogue of functions.

Theorem 2.3.2 (Thm VI.2 [49]). *Let α be an arbitrary complex number in $\mathbb{C}\backslash\mathbb{Z}_{\leq 0}$. The n^{th} coefficient of*

$$f(z) = \left(\frac{1}{1-z}\right)^{\alpha} \left(\frac{1}{z}\log\frac{1}{1-z}\right)^{\beta} \tag{2.12}$$

admits for large n a full asymptotic expansion in descending powers of $\log n$,

$$f_n \equiv [z^n]f(z) \sim \frac{n^{\alpha-1}}{\Gamma(\alpha)} \cdot (\log n)^{\beta} \cdot \left(1 + \frac{C_1}{\log n} + \frac{C_2}{\log^2 n} + \dots\right) \tag{2.13}$$

where $C_k = \binom{\beta}{k}\Gamma(\alpha)\frac{d^k}{ds^k}\frac{1}{\Gamma(s)}\Big|_{s=\alpha}$.

The proof is based on Cauchy's formula and a systematic use of Hankel-like contours (see [49] Appendix B.3. for details on Hankel contours).

Let S denote the set of such singular functions

$$S = \left\{ \left(\frac{1}{(1-z)}\right)^{\alpha} \lambda(z)^{\beta} : \alpha, \beta \in \mathbb{C},\ \lambda(z) := \frac{1}{z}\log\frac{1}{1-z} \right\}.$$

The following definition is illustrated in Figure 2.1.

Definition 2.3.3 (**Delta domain**). *Given two numbers Φ, R with $R > 1$ and $0 < \Phi < \frac{\pi}{2}$, the open domain $\Delta(\Phi, R)$ is defined as*

$$\Delta(\Phi, R) = \{z : |z| < R, z \neq 1, |\arg(z-1)| > \Phi\}.$$

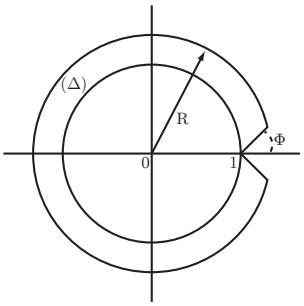


Figure 2.1.: The camembert Δ-domain around 1.

A domain is a Δ-domain at 1 if it is a $\Delta(\Phi, R)$ for some R and Φ. For a complex number $\zeta \neq 0$, a Δ-domain at ζ is the image by the mapping $z : z \longmapsto \zeta \cdot z$ of a Δ-domain at 1. A function is Δ-analytic if it is analytic in some Δ-domain.

The following theorem is the heart of singularity analysis, and justifies the translation of a function's asymptotic expansion into an asymptotic expansion of its coefficients.

Theorem 2.3.4 (Singularity analysis, single singularity, Thm VI.4, [49]). *Let $f(z)$ be function analytic at 0 with a singularity at ζ, such that $f(z)$ can be continued to a domain of the form $\zeta\Delta_0$, for a Δ domain Δ_0, where $\zeta\Delta_0$ is the image of Δ_0 by the mapping $z : \Delta \longmapsto \zeta z$. Assume that there exists two functions δ and τ, where δ is a (finite) linear combination of functions in S and $\tau \in S$, so that*

$$f(z) = \delta\left(\frac{z}{\zeta}\right) + O\left(\tau\left(\frac{z}{\zeta}\right)\right) \tag{2.14}$$

as $z \to \zeta$ in $\zeta \cdot \Delta_0$. Then, the coefficients of $f(z)$ satisfy the asymptotic estimate

$$f_n = \zeta^{-n}\delta_n + O\left(\zeta^{-n}\tau_n^*\right)$$

where $\delta_n = [z^n]\delta(z)$ has its coefficients determined by Theorem 2.3.2 (Theorems VI.1, VI.2 in [49] respectively) and $\tau_n^* = n^{a-1}(\log n)^b$, if $\tau(z) = (1-z)^{-a}\lambda(z)^b$.

As mentioned earlier we are not only interested in the number of graphs on n vertices but also in the number of graphs on n vertices with k components. As it turns out it is reasonable to hope that if the univariate generating function is very-well behaved than so is the bivariate generating function counting the number of components. In fact, using singularity perturbation techniques (studying the second parameter u around 1) singularity analysis applies again and we can study how the radius of the generating function varies depending on u. We refer to [49, p.648 following] for details and only quote the for us relevant result ([49, Prop. IX.14].

Proposition 2.3.5 (Sets of logarithmic structure [49]). *Let $\mathcal{A} = \mathrm{SET}(\mathcal{C})$. Assume that $C(z)$ has radius of convergence ρ and that for $\lambda, \kappa \in \mathbb{R}$ the following expansion is valid in a Δ-domain*

$$C(z) = \kappa \log \frac{1}{1 - z/\rho} + \lambda + O\left(\frac{1}{\log^2(1 - z/\rho)}\right).$$

Then the normalized version of the number of components $\kappa(R_n)$ converges in distribution to a standard normal distribution

$$\frac{\kappa(R_n) - \mathbb{E}(\kappa(R_n))}{\sqrt{\mathbb{V}(\kappa(R_n))}} \xrightarrow{d} \mathcal{N}(0,1)$$

and

$$\mathbb{E}(\kappa(R_n)) \sim Var(\kappa(R_n)) \sim \kappa \log(n).$$

2.3.1. Labelled forests

A very well studied and important class of graphs that can be analysed by singularity analysis is the class of labelled forests. A rooted tree is a tree with a specified vertex. The subtrees stemming from the root are not ordered between themselves. Hence we can think of a rooted tree as a specified vertex (=the root) and a set of rooted trees stemming from the root. Denoting $\mathcal{T}$ the set of all rooted trees we have

$$\mathcal{T} = \mathcal{Z} \times \mathrm{SET}(\mathcal{T}).$$

Using (2.2) we get the well-known functional equation

$$T(z) = z \cdot e^{T(z)}$$

(see [49, Chapter II, p. 127]). As $T(z)$ counts rooted trees, each tree is counted n times in $T(z)$. But we are interested in unrooted trees. $\frac{T(z)^2}{2}$ counts the number of unordered pairs of rooted trees or equivalently edge-rooted trees (introduce an edge between the two roots and call it the root edge). Every tree has $n-1$ edges and hence it is counted $n-1$ times in $\frac{T(z)^2}{2}$. Let $C(z)$ be the generating function of unrooted trees. We have

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

The series $T(z)$, $C(z)$ and $A(z) = e^{C(z)}$ have radius $\rho = 1/e$, with the following singular expansions at this point (see [49, Chap. VI, p. 405] for a general theorem on the asymptotic expansions of these types of functions):

$$\begin{array}{llllll} T(z) = 1 & -\sqrt{2}(1-ze)^{1/2} & +\frac{2}{3}(1-ze) & -\frac{11\sqrt{2}}{36}(1-ze)^{3/2} & +O((1-ze)^2), \\ C(z) = \frac{1}{2} & & -(1-ze) & +\frac{2\sqrt{2}}{3}(1-ze)^{3/2} & +O((1-ze)^2), \\ A(z) = \sqrt{e} & & -\sqrt{e}(1-ze) & +\sqrt{e}\frac{2\sqrt{2}}{3}(1-ze)^{3/2} & +O((1-ze)^2). \end{array} \tag{2.15}$$

The three series are analytic in a Δ-domain and singularity analysis applies. By Theorem 2.3.4 it suffices to deduce the coefficients of the asymptotic expansion of $T(z)$, $C(z)$, $A(z)$. For example

$$\begin{aligned}[z^n]T(z) &\sim [z^n]1 - \sqrt{2}(1-ze)^{1/2} + \frac{2}{3}(1-ze) - \frac{11\sqrt{2}}{36}(1-ze)^{3/2} + O((1-ze)^2) \\ &\sim [z^n] - \sqrt{2}(1-ze)^{1/2} + O\left(\frac{1}{n}\right).\end{aligned}$$

Fortunately these expansions are all built up of well-behaved functions, so by applying Theorem 2.3.2 their coefficients can be deduced.

$$T_n \sim n!\frac{e^n}{\sqrt{2\pi}n^{3/2}}, \quad C_n \sim n!\frac{e^n}{\sqrt{2\pi}n^{5/2}}, \quad \text{and} \quad A_n \sim C(1/e)C_n = \sqrt{e}\,C_n.$$

We can now prove results on random forests, first obtained in [84].

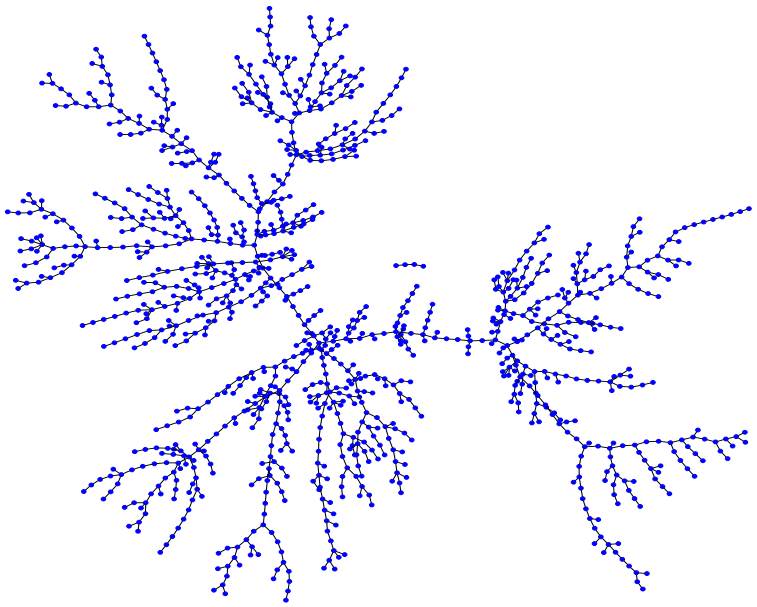


Figure 2.2.: A random forest of size $n = 1165$ with two connected components.

Proposition 2.3.6 (Number of components in a random forest). *Let $R_n \in_u \mathcal{A}$ be a forest chosen uniformly at random from the set of all (labelled) forests. Then*

$$\kappa(R_n) \xrightarrow{d} 1 + \mathrm{Po}\left(\frac{1}{2}\right).$$

Furthermore, as n tends to ∞

$$\mathbb{E}(\mathrm{frag}(R_n)) \to 1. \tag{2.16}$$

Proof. With generating functions the proof of these classic results becomes very simple: from (2.15) we know that $C(z)$ has radius of convergence $\rho = e^{-1}$ and $C\left(\frac{1}{e}\right) = \frac{1}{2}$. The coefficients of $C(z)$ are smooth and $\lim_{n\to\infty} \frac{C_n}{A_n} = \frac{1}{\sqrt{e}}$. Applying Theorem 2.2.3 gives the result. Equation (2.16) is a consequence of Proposition 2.2.6 and Proposition 5.2. [70]. The proof of [70, Proposition 5.2.] shows that for forests $\mathbb{E}(\text{frag}(R_n)) \to \mathbb{E}(X) = \frac{\partial C}{\partial z}(z)\Big|_{z=\rho} \rho$ as n tends to ∞. By (2.15) $\rho = \frac{1}{e}$ and by [49, Theorem VI.8.] we know that it suffices to differentiate the asymptotic expansion of $C(z)$ given in (2.15). Hence $C'(\rho) = e$ and the result follows. □

2.4. Hayman-admissibility and the saddle point method

With singularity analysis we now have a ready tool to analyse generating functions of moderate growth. What about generating functions that are *not* of moderate growth/decay, but that behave rather violently around their singularities? Cauchy's coefficient formula (Theorem 2.3.1) still applies in this case. We are left with the same task as before, i.e. we need to find a way to evaluate

$$F(n) = \int_P \phi(z)dz \tag{2.17}$$

where P is an integration path in the complex plane and $\phi(z)$ is analytic in some neighbourhood of P. Both the integration path P and the function $\phi(z)$ depend on a real (large) parameter n and we are interested in the asymptotic behaviour $F(n)$ when $n \to \infty$. Following [34] we suppose now for a moment that we are not interested in evaluating (2.17) but we only want to have an upper estimate for $F(n)$. For a path P of finite length l_P we have

$$|F(n)| \leq \int_P |\phi(z)|dz \leq l_P \cdot \max_P |\phi(z)|$$

where $\max_P |\phi(z)|$ is the maximum of $|\phi(z)|$ along P. By Cauchy's Theorem (Homotopy Form) we know that P can be replaced by any path C, which has the same endpoints as P, provided that C can be continuously transformed into P without leaving the domain of analyticity of $\phi(z)$. Hence we now wish to determine C such that $\max_C |\phi(z)|$ is minimal. Obviously, such a C does not have to exist, but following [34] "it commonly does". Furthermore, we hope that the path C is at the same time a good path for the evaluation of $F(n)$ itself, i.e. we hope that we can parametrise the path C and write $\int_C \phi(z)dz$ as an integral along a real interval, to which we can try to apply the Laplace method (see Chapter 4 of [34] for a detailed

introduction on Laplace method and Appendix B.6 of [49] for a short overview). It should also be remarked once more that this method only has a chance to succeed if $\phi(z)$ behaves violently. Therefore it is often considered to be somehow a counterpart to singularity analysis (which was presented in Section 2.3) as singularity analysis typically applies to functions of moderate growth.

Typically, for the cases that we will be concerned about, the paths have to go to some higher altitude than the two end points of P. Remember that a saddle point is a point in the domain of a function that is a stationary point but not a local extremum. Then a path C that traverses a saddle point by connecting two points at lower altitude on the surface of $\phi(z)$ and by following two steepest decent lines across the saddle point is a local minimum for $\max_C |\phi(z)|$. This indicates why choosing a path which traverses a saddle point might be a good idea and why the general method is called the saddle point method.

As mentioned previously we would like to analyse the coefficients of rapidly varying analytic generating functions and the saddle point method specialises. By Cauchy's Coefficient Formula we have for $G(z)$ analytic in a region Ω containing 0

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

where λ is a simple, positively oriented loop around 0 that lies within Ω. This still fits our framework as we can see λ as a path with two coinciding end points A and B lying somewhere along the real line. Following [49, p. 549] we can take $A, B = -ae^{i\epsilon}$ for $a > 0$ and $\epsilon \to \infty$. Suppose now that $G(z)$ has a saddle-point on the real axis and therefore by Cauchy's Theorem we can assume that we can choose a path $\mathcal{C}$ which connects the two end points A and B lying in opposite valleys of the saddle-point σ. Changing to $z = re^{i\theta}$ in (2.4) with r close to the saddle point value we obtain

$$[z^n]G(z) = \frac{r^{-n}}{2\pi}\int_{-\pi}^{\pi} G(re^{i\theta})e^{-in\theta}d\theta. \tag{2.18}$$

A natural possibility to choose r is to adapt the saddle point equation $\zeta G'(\zeta)/G(\zeta) = n+1$ to a version relative to polar coordinates. So r is defined by

$$r\frac{G'(r)}{G(r)} = n.$$

The idea is then to decompose the loop $\lambda = \mathcal{C}^{(0)} \cup \mathcal{C}^{(1)}$. $\mathcal{C}^{(0)}$ is called the *central part* and contains the saddle point (or passes very near to it). It is an arc of the circle of radius r determined by $|\theta| \leq \theta_0$ for suitable choice of θ_0. Define $F(z) = G(z)/z^{n+1}$ and $f(z) := \log G(z) - n \log z$. On C_0 $F(z)$ should be well-behaved, i.e. for fixed r (and so fixed n) and $|\theta| \leq \theta_0$ we want a **Gaussian approximation** to hold, i.e.

$$F(re^{i\theta}) \sim F(r)e^{-\theta^2 r^2 f''(r)/2}.$$

This is equivalent to

$$\left| f(re^{i\theta}) - f(r) + \theta^2 \beta(r) f''(r)/2 \right| \to 0$$

for $\theta \to 0$ and some function β. Normally, it suffices to take a Taylor expansion of $g : \theta \mapsto f(re^{i\theta})$ in the neighbourhood of $\theta = 0$ and make sure that the higher order terms in θ disappear (by a suitable choice of θ_0).

The second part of the integral can be neglected (Tails pruning). On the contour $\mathcal{C}^{(1)}$, i.e. for $|\theta| > \theta_0$ the contribution of the integral $\int_{C_1}$ is negligible

$$\frac{\left| \int_{C_1} F(z)dz \right|}{\left| \int_{C=C_0 \cup C_1} F(z)dz \right|} \to 0.$$

Completion of the Gaussian integral. We are left to verify that we can actually evaluate (2.18). By the Gaussian approximation along C_0 we know that

$$\int_{-\theta_0}^{\theta_0} F(re^{i\theta}) \frac{re^{i\theta} d\theta}{2\pi} \sim \frac{rF(r)}{2\pi} \int_{-\theta_0}^{\theta_0} e^{-\theta^2 r^2 f''(r)/2} e^{i\theta} d\theta.$$

We need to verify that this is asymptotically equivalent to

$$\sim \frac{F(r)}{\sqrt{2\pi f''(r)}}$$

using a similar trick used for the evalutation of $\int_0^\infty e^{-x^2} dx = \frac{\sqrt{\pi}}{2}$ (see [49, Chapter B.3, p. 744] for further explanation). An example where the saddle point method is applied step by step can be found in the proof of Proposition 3.4.3. The so-called *saddle point algorithm* can be found in [49, Theorem VIII.3].

Many combinatorial classes whose leading construction is a SET operation are amenable to the saddle-point method. Examples include the generating function of involutions ($A(z) = e^{z+z^2/2}$), set partitions ($A(z) = e^{e^z - 1}$) and fragmented

permutations ($A(z) = e^{z/(1-z)}$) (see [49] p. 558 and following pages for details). However, it is often cumbersome to apply the method step-by-step. The concept of *Hayman-admissible (H-admissible) generating functions*, first mentioned in [60], provides some simplified testable conditions in order to determine if a function is amenable to the saddle-point method or not. Furthermore, once a function is found to be H-admissible useful closure properties give access to a range of new functions which are then automatically H-admissible as well. We follow the definition given in Flajolet and Sedgewick [49] page 564.

Theorem 2.4.1 (Hayman 1956)**.** *Let $A(z)$ be a power series with real coefficients and radius $\rho \in (0, \infty]$. Assume $A(r)$ is positive in some subinterval $r \in (R, \rho) \subset (0, \rho)$. Define*

$$a(r) = r\frac{A'(r)}{A(r)} \quad \textit{and}$$
$$b(r) = r\frac{A'(r)}{A(r)} + r^2\frac{A''(r)}{A(r)} - r^2\left(\frac{A'(r)}{A(r)}\right)^2 = rC'r + r^2C''(r).$$

Assume that the following three properties hold:

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

$\mathbf{H_2}$ [Locality condition] For some function $\theta_0(r)$ defined over (R, ρ), and satisfying $0 < \theta_0 < \pi$, one has
$$\sup_{|\theta|\le\theta_0(r)} \left|\frac{A(re^{i\theta})}{A(r)} e^{-i\theta a(r)+\theta^2 b(r)/2} - 1\right| \to 0 \qquad \text{as } r \to \rho,$$

$\mathbf{H_3}$ [Decay condition]
$$\sup_{|\theta|\in[\theta_0(r),\pi)} \left|\frac{A(re^{i\theta})}{A(r)}\sqrt{b(r)}\right| \to 0 \qquad \text{as } r \to \rho.$$

Then $A(z)$ is Hayman-admissible (short H-admissible) and the n^{th} coefficient of $A(z)$ satisfies, as $n \to \infty$,

$$[z^n]A(z) \sim \frac{A(\zeta)}{\zeta^n\sqrt{2\pi b(\zeta)}}$$

where $\zeta \equiv \zeta_n$ is the unique solution in (R, ρ) of the saddle point equation $\zeta A'(\zeta) = nA(\zeta)$.

Conditions $\mathbf{H_2}$ and $\mathbf{H_3}$ are usually stated in terms of *uniform equivalence* as $r \to \rho$ (see for example [49, page 564]) but we find the above formulation more explicit.

As already mentioned before, the set of H-admissible series has several very useful and well-known closure properties already mentioned in the original paper [60].

Theorem 2.4.2 (Hayman 1956)**.** *Let $G(z)$ and $H(z)$ be H-admissible functions and let $P(z)$ be a polynomial with real non-negative, aperiodic coefficients. Then the following functions are also H-admissible:*

1. *$G(z)H(z)$, $G(z)P(z)$*
2. *$e^{G(z)}$, $e^{P(z)}$, $P(G(z))$.*

Even though the concept has been around for many decades now, there was one closure property which we were not able to find in the literature but which happened to be a very useful tool for later applications.

Theorem 2.4.3. *Let $A(z) = F(z)G(z)$ where $F(z)$ and $G(z)$ are power series with real coefficients and radii of convergence $0 < \rho_F < \rho_G \leq \infty$. Assume that $F(z)$ has non-negative coefficients and is Hayman-admissible, and that $G(\rho_F) > 0$. Then $A(z)$ is Hayman-admissible.*

Proof. Let us first prove that the radius of convergence ρ of $A(z)$ is ρ_F. Clearly, $\rho \geq \rho_F$. Now, suppose $\rho > \rho_F$. Then $A(z)$ is analytic at ρ_F. Together with $G(\rho_F) > 0$ this implies that $F(z) = A(z)/G(z)$ has an analytic continuation at ρ_F, which is impossible by Pringsheim's Theorem (Theorem IV.6 of [49], p. 240) since $F(z)$ has non-negative coefficients. Note also that $A(r)$ is positive on an interval of the form $[R, \rho)$ (by continuity of G). Let us now check the three conditions of Theorem 2.4.1. We have

$$\begin{aligned} a(r) &= a_F(r) + a_G(r) \\ b(r) &= b_F(r) + b_G(r), \end{aligned}$$

with

$$a_G(r) = r\frac{G'(r)}{G(r)} \quad \text{and} \quad b_G(r) = r\frac{G'(r)}{G(r)} + r^2\frac{G''(r)}{G(r)} - r^2\left(\frac{G'(r)}{G(r)}\right)^2,$$

and similarly for a_F and b_F.

$\mathbf{H_1}$. Since $G(\rho_F) = c > 0$, all derivatives of $G(z)$ are finite at $|z| \leq \rho_F$ as they have the same radius of convergence as $G(z)$. The capture condition then clearly holds since $F(z)$ satisfies it.

$\mathbf{H_2}$. We need to show that

$$\sup_{|\theta| \leq \theta_0(r)} \left| \frac{A(re^{i\theta})}{A(r)} \cdot e^{-ia(r)\theta + \theta^2 b(r)/2} - 1 \right| \to 0$$

as $r \to \rho$. Choose $\theta_0(r) = \theta_0^F(r)$ where $\theta_0^F(r)$ is a function for which $F(z)$ satisfies $\mathbf{H_2}$ and $\mathbf{H_3}$. We have

$$\frac{A(re^{i\theta})}{A(r)} \cdot e^{-ia(r)\theta + \theta^2 b(r)/2} = \frac{F(re^{i\theta})}{F(r)} e^{-ia_F(r)\theta + \theta^2 b_F(r)/2} \frac{G(re^{i\theta})}{G(r)} \cdot e^{-ia_G(r)\theta + \theta^2 b_G^2(r)/2}. \tag{2.19}$$

By assumption, F satisfies the locality condition. Hence

$$\frac{F(re^{i\theta})}{F(r)} \cdot e^{-ia_F(r)\theta + \frac{b_F(r)\theta^2}{2}} = 1 + M(r, \theta) \tag{2.20}$$

where

$$\sup_{|\theta| \leq \theta_0(r)} |M(r, \theta)| \to 0 \tag{2.21}$$

as $r \to \rho$. Let $r \in [R, \rho)$ be fixed for a moment. Take θ with $|\theta| \leq \theta_0(r)$, and let us consider a Taylor expansion of $\log G(re^{i\theta})$ in powers of θ around $\theta = 0$:

$$\log G(re^{i\theta}) = \log G(r) + i\theta a_G(r) - \frac{\theta^2}{2} b_G(r) + \theta^3 S(r, \theta)$$

where

$$\begin{aligned} S(r, \theta) = &\frac{1}{3!} \frac{\partial^3}{\partial \alpha^3} \left(\log G(re^{i\alpha}) \right) \\ &- ire^{i\alpha} \left(r^2 e^{2i\alpha} G'''(re^{i\alpha}) + 3re^{i\alpha} G''(re^{i\alpha}) + G'(re^{i\alpha}) \right) \end{aligned}$$

is bounded uniformly in a neighborhood of $(\rho, 0)$. In the original paper, it was shown that $\theta_0(r) \to 0$ as $r \to \rho$ (see [60, Eq. (12.1)]). Thus,

$$\frac{G(re^{i\theta})}{G(r)} \cdot e^{-ia_G(r)\theta + \theta^2 b_G(r)/2} = e^{\theta^3 S(r,\theta)} = 1 + N(r, \theta) \tag{2.22}$$

where

$$\sup_{|\theta| \leq \theta_0(r)} |N(r, \theta)| \to 0 \tag{2.23}$$

as $r \to \rho$. Putting together Equations (2.19 - 2.23), we obtain that $A(z)$ satisfies the locality condition $\mathbf{H_2}$.

$\mathbf{H_3}$. We have:

$$\begin{aligned}\left|\frac{A(re^{i\theta})}{A(r)}\sqrt{b(r)}\right| &= \left|\frac{F(re^{i\theta})G(re^{i\theta})}{F(r)G(r)}\sqrt{b_F(r)+b_G(r)}\right| \\ &\leq \left|\frac{F(re^{i\theta})}{F(r)}\cdot\frac{G(re^{i\theta})}{G(r)}\cdot\sqrt{2b_F(r)}\right|\end{aligned}$$

for r close to ρ, because $b_F(r) \to \infty$ as $r \to \rho$ while $b_G(r)$ is bounded around ρ. Also, since G is analytic at ρ and $G(\rho) > 0$, the term $G(re^{i\theta})/G(r)$ is uniformly bounded in a neighborhood of the circle of radius ρ. Since by assumption, $F(z)$ satisfies $\mathbf{H_3}$, this shows

$$\lim_{r\to\rho_F}\sup_{\theta_0(r)<|\theta|}\frac{|F(re^{i\theta})|}{|F(r)|}\sqrt{2\cdot b_F(r)}\cdot\frac{|G(re^{i\theta}|}{G(r)} = 0$$

so $A(z)$ satisfies $\mathbf{H_3}$ as well. □

In later chapters not only the number of structures of size n will be of interest, but also obtaining limit laws of various parameters of our class $\mathcal{A}$. The number of components of $R_n \in_u \mathcal{A}$, denoted by $\kappa(R_n)$ is of particular interest. For this purpose it will be convenient to have a **uniform** version of H-admissibility for series of the form $e^{uC(z)}$. In 2010 Drmota et al. [44] extended the concept of H-admissible to several variables. We here only need the version for bivariate generating functions.

The bivariate exponential generating function of an array $(F_{n,k})_{n,k\in\mathbb{N}}$ is the formal power series $F(z,u) = \sum_{n,k\geq 1} F_{n,k}\frac{z^n}{n!}u^k$. To simplify the notation define $F_z(z,u) = \frac{\partial}{\partial z}F(z,u)$, $F_u(z,u) = \frac{\partial}{\partial u}F(z,u)$, respectively.

Definition 2.4.4 (**Extended Hayman-admissible, [44]**). *Let*

$$\begin{aligned} a(z,u) &= z\frac{A_z(z,u)}{A(z,u)}, \qquad & b(z,u) &= za_z(z,u), \\ \overline{a}(z,u) &= u\frac{A_u(z,u)}{A(z,u)}, \qquad & \overline{b}(z,u) &= u\overline{a}_u(z,u), \\ c(z,u) &= ua_u(z,u). \end{aligned}$$

With this notation a function $A(z,u) = \sum_{n,k\geq 0} a_{n,k}z^n u^k$ is called extended admissible if there exists $0 < R \leq \infty$ such that the following conditions are satisfied.

1. *$A(z,u)$ is analytic in $\{(z,u) : |z| < R, |u| < 1+\zeta\}$ where $\zeta > 0$ and for some $R_0 < R$ and $R_0 < r < R$ we have $A(r,1) > 0$.*

H_2' *[Locality condition] Set* $\epsilon(r) = K\left(\overline{b}(z,u) - \frac{c(r,1)^2}{b(r,1)}\right)^{-\frac{1}{2}} > 0$ *where* $K > 0$ *is an arbitrary constant. Then there exists a function* $\delta(r) : (R_0, R) \to (0, \pi)$ *such that for* $R_0 < r < R$ *we have*

$$A(re^{i\theta}, u) \sim A(r,u)e^{i\theta a(r,u) - \frac{\theta^2}{2}b(r,u)} \tag{2.24}$$

as r *tends to* R *from below and uniformly for* $|\theta| \leq \delta(r)$ *and* $u \in [1-\epsilon(r), 1+\epsilon(r)]$.

H_3' *[Decay condition] For* $R_0 < r < R$ *we have*

$$A(re^{i\theta}, u) = o\left(\frac{A(r,u)}{\sqrt{b(r,u)}}\right) \tag{2.25}$$

as r *tends to* R *and uniformly for* $\delta(r) \leq |\theta| \leq \pi$ *and* $u \in [1-\epsilon(r), 1+\epsilon(r)]$.

2. $b(r,1) \to \infty$ *as* r *tends to* R *from below.*

3. $b(r,u) \sim b(r,1)$ *for* r *tending to* R *from below and uniformly for* $u \in [1-\epsilon(r), 1+\epsilon(r)]$.

4. $a(r,u) = a(r,1) + c(r,1)(u-1) + O(c(r,1)(u-1)^2)$ *for* $r \in (R_0, R)$ *and* $u \in [1-\epsilon(r), 1+\epsilon(r)]$.

5. $\overline{a}(r,u) = O(\overline{a}(r,1))$ *and* $\overline{b}(r,u) = O(\overline{b}(r,1))$ *for all* u *in an arbitrary but fixed complex neighbourhood of 1 and all* r.

6. $\overline{b}(r,1) - \frac{c(r,1)^2}{b(r,1)} \to \infty$ *as* r *tends to* R *from below.*

7. $\epsilon(r)^3\overline{b}(r,1) \to 0$ *as* r *tends to* R *from below.*

8. $\overline{a}(r,1) = O\left(A(r,1)^\eta\right)$ *and* $\overline{b}(r,1) = O\left(A(r,1)^\eta\right)$ *for every* $\eta > 0$.

Theorem 2.4.5. *Let* $C(z)$ *be a power series with non-negative coefficients and radius of convergence* ρ. *Assume* $A(z) = e^{C(z)}$ *has radius* ρ *and is Hayman-admissible. Define*

$$b(r) = rC'(r) + r^2C''(r) \quad \textit{and} \quad V(r) = C(r) - \frac{(rC'(r))^2}{rC'(r) + r^2C''(r)}.$$

Assume that, as $r \to \rho$,

$$\begin{aligned} V(r) &\to +\infty, && (2.26)\\ \frac{C(r)}{V(r)^{3/2}} &\to 0, && (2.27)\\ b(r)^{1/\sqrt{V(r)}} &= O(1). && (2.28) \end{aligned}$$

Then $A(z,u) := e^{uC(z)}$ is extended Hayman-admissible.

If then $\kappa(R_n)$ is a sequence of random variables such that

$$\mathbb{P}(\kappa(R_n) = k) = \frac{[z^n u^k] e^{uC(z)}}{[z^n] e^{C(z)}},$$

then the mean and variance of $\kappa(R_n)$ satisfy

$$\mathbb{E}(\kappa(R_n)) \sim C(\zeta_n), \qquad \mathbb{V}(\kappa(R_n)) \sim V(\zeta_n), \tag{2.29}$$

where $\zeta = \zeta_n$ is the unique solution in $(0,\rho)$ of the saddle point equation $\zeta C'(\zeta) = n$. Moreover, the normalized version of $\kappa(R_n)$ converges in distribution to a standard normal distribution:

$$\frac{\kappa(R_n) - \mathbb{E}(\kappa(R_n))}{\sqrt{\mathbb{V}(\kappa(R_n))}} \to \mathcal{N}(0,1).$$

Proof. With the previously defined notation, we have

$$\begin{aligned} &a(r,u) = c(r,u) = ruC'(r), \qquad b(r,u) = ruC'(r) + r^2 uC''(r) = ub(r),\\ &\bar{a}(r,u) = \bar{b}(r,u) = uC(r), \qquad \varepsilon(r) = \frac{K}{\sqrt{V(r)}} \end{aligned} \tag{2.30}$$

for a fixed constant K. Condition (1) holds for $R = \rho$, any $\zeta > 0$ and any $R_0 \in [0,\rho)$: indeed, the series $A(z,u)$ is analytic for $|z| < \rho$ and $u \in \mathbb{C}$, and $A(z,1)$ is positive on $[0,\rho)$. Conditions (8) and (9) hold by the above assumptions (2.26) and (2.27). Condition (4) is that $b(r) \to +\infty$ as $r \to \rho$ and holds because A is Hayman admissible. Condition (5) requires that $b(r,u) \sim b(r,1)$ for $r \to \rho$, uniformly for $u \in [1-\epsilon(r), 1+\epsilon(r)]$: this holds because $b(r,u)/b(r,1) = u$ and $\varepsilon(r) \to 0$ as $r \to \rho$. Condition (6) requires that $a(r,u) = a(r,1) + c(r,1)(u-1) + O(c(r,1)(u-1)^2))$ uniformly for $r \in (0,\rho)$ and $u \in [1-\epsilon(r), 1+\epsilon(r)]$. Since $a(r,u) = a(r,1) + c(r,1)(u-1)$, Condition (6) holds.

We are thus left with Conditions (2) and (3), which are uniform versions (in u) of the locality and decay conditions $\mathbf{H_2}$ and $\mathbf{H_3}$ defining H-admissibility. We begin

with $\mathbf{H}_\mathbf{2}'$. Since $A(z)$ is H-admissible, let $\theta_0(r)$ be a function for which $\mathbf{H_2}$ (and $\mathbf{H_3}$) holds:

$$\frac{A(re^{i\theta})}{A(r)}e^{-i\theta a(r)+\theta^2 b(r)/2} = 1 + M(r,\theta)$$

where

$$M(r) := \sup_{|\theta|\le\theta_0(r)} |M(r,\theta)|$$

tends to 0 as $r \to \rho$. Then, for $u \in [1-\varepsilon(r), 1+\varepsilon(r)]$,

$$\begin{aligned}\frac{A(re^{i\theta},u)}{A(r,u)}e^{-i\theta a(r,u)+\theta^2 b(r,u)/2} &= \exp\Big(u\Big(C(re^{i\theta}) - C(r) - i\theta a(r) + \theta^2 b(r)/2\Big)\Big)\\ &= (1+M(r,\theta))^u\end{aligned}$$

where we have taken the principal branch of log to define $(1+M(r,\theta))^u = \exp(u\log(1+M(r,\theta)))$ (because $M(r,\theta)$ is close to 0). Thus

$$\begin{aligned}\sup_{\substack{|\theta|\le\theta_0(r)\\ |u-1|\le\varepsilon(r)}}\left|\frac{A(re^{i\theta},u)}{A(r,u)}e^{-i\theta a(r,u)+\theta^2 b(r,u)/2} - 1\right| &= \sup_{\substack{|\theta|\le\theta_0(r)\\ |u-1|\le\varepsilon(r)}}|(1+M(r,\theta))^u - 1|\\ &\le (1+\varepsilon(r))M(r) + O\Big((1+\varepsilon(r))M(r)^2\Big),\end{aligned}$$

and this upper bound tends to 0 as $r \to \rho$. This proves $\mathbf{H}_\mathbf{2}'$ with $\delta(r) = \theta_0(r)$.

We finally address $\mathbf{H}_\mathbf{3}'$. Since $A(z)$ satisfies the decay condition $\mathbf{H_3}$, the quantity

$$N(r,\theta) := \frac{A(re^{i\theta})}{A(r)}\sqrt{b(r)}$$

satisfies

$$\sup_{|\theta|\in[\theta_0(r),\pi]} |N(r,\theta)| \to 0 \tag{2.31}$$

as $r \to \rho$. We have, for $u \in [1-\varepsilon(r), 1+\varepsilon(r)]$,

$$\begin{aligned}\left|\frac{A(re^{i\theta},u)}{A(r,u)}\sqrt{b(r,u)}\right| &= |N(r,\theta)|^u \sqrt{b(r)}^{1-u}\sqrt{u}\\ &\le |N(r,\theta)|^{1+\varepsilon(r)} \sqrt{b(r)}^{\varepsilon(r)}\sqrt{1+\varepsilon(r)},\end{aligned}$$

and this tends to 0 uniformly for $|\theta| \in [\theta_0(r), \pi]$ thanks to (2.31), (2.30), (2.26) and (2.28).

As explained in [44] just below Theorem 2, these eight conditions give the estimates (2.29) of $\mathbb{E}(N_n)$ and $\mathbb{V}(N_n)$ and imply the existence of a gaussian limit law. □

2.5. The method of moments and special probability distributions

As mentioned at the beginning of this Chapter (Section 2.1) we are not only interested in the number of components of $R_n \in \mathcal{A}$ but also in the size of the component containing the vertex labelled 1. As we will see later, this parameter gives rise to non-gaussian limit laws which we define in this section. To obtain these limit laws we are going to apply the *method of moments*. Let $(X_n)_{n\in\mathbb{N}}, X$ be random variables with distribution functions $(F_n)_{n\in\mathbb{N}}, F$, respectively. We say that the sequence of random variables $(X_n)_{n\in\mathbb{N}}$ converges to X in distribution if for every $x \in \mathbb{R}$ $\lim_{n\to\infty} F_n(x) = F(x)$.

In general, the method of moments is a way of proving convergence in distribution by proving convergence of the corresponding moments. Suppose we are given a probability distribution where all moments are finite and the moment generating function of this probability distribution has positive radius of convergence. Then the probability distribution is uniquely defined by its moments [21, Theorem 30.1.], [49, Thm. C.2].

Theorem 2.5.1 (**Moment Convergence Theorem**). *Let X be a random variable with finite moments which is uniquely determined by its moments. Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of random variables and let $\mathbb{E}(X_n^k)$ denote the k^{th} factorial moment of X_n. If*

$$\lim_{n\to\infty} \mathbb{E}(X_n^k) = \mathbb{E}(X^k)$$

for all values of k, then the sequence $X_{n\in\mathbb{N}}$ converges to X in distribution.

Clearly the factorial moments can be obtained from the moments and vice versa. For our applications, we will show convergence of the factorial moments and deduce convergence in distribution.

In the following we present a rough overview of the non-standard probability distributions that appear in this thesis. The definitions are taken from [21].

The gamma distribution is a two-parameter family of continuous probability distribution.

Definition 2.5.2 (**Gamma distribution**). *Let $r, s > 0$. For $x \in (0, \infty)$ the density of the Gamma distribution Gamma(r, s) is given by*

$$f(x, r, s) = \frac{x^{r-1} e^{-\frac{x}{s}}}{s^r \Gamma(r)}.$$

The k^{th} factorial moment of the Gamma distribution of parameter r and $s = 1$ is

$$\mathbb{E}(X_n^{(k)}) = \frac{\Gamma(k+r)}{\Gamma(r)}.$$

We observe that it is not difficult to see that the sum of r independent exponential random variables, each with mean s has a gamma density with parameters r, s.

The beta distribution is a family of continuous probability distributions defined on the interval $[0, 1]$ parametrized by two positive shape parameters.

Definition 2.5.3. Beta distribution *Let $r, s > 0$. The density of the Beta distribution beta(r, s) is given by*

$$f(x, r, s) = \frac{1}{B(r, s)} x^{r-1}(1-x)^{s-1}$$

where $B(r, s) = \frac{\Gamma(r)\Gamma(s)}{\Gamma(r+s)}$ and $\Gamma(x)$ stands for the standard Gamma function. The k^{th} factorial moment of the Beta distribution is

$$\mathbb{E}(X_n^{(k)}) = \frac{\Gamma(r+s)k!}{\Gamma(k+r+s)}$$

The Gumbel distribution named after Emil Julius Gumbel is used to model the distribution of the maximum (or the minimum) of a number of samples of various distributions. In particular, the maximum (or last order statistic) of a series of independent random variables following an exponential distribution converges in distribution to a Gumbel distribution.

Definition 2.5.4. The Gumbel distribution *Let $x \in (-\infty, \infty)$. The cumulative distribution function of the Gumbel distribution with parameters $\mu, \beta > 0$ is*

$$F(x; \mu, \beta) = \exp^{-e^{-(x-\mu)/\beta}}.$$

The Poisson-Dirichlet distribution was named by Kingman in 1975 [64] and Kingman himself ([63, Section 9.6]) says that 'the distribution $PD(\theta)$ is rather less than user-friendly'. It is related to the Poisson process and was introduced during the study of the length of cycles in random permutations.

Definition 2.5.5. The Poisson-Dirichlet distribution *Let $Y_1 \geq Y_2 \geq \ldots$ be the points of a Poisson point process with intensity $\theta e^{-x}/x, x > 0$ that is $(Y_i)_{i \in \mathbb{N}}$ are independent, $Y_0 = 0$ and $\mathbb{P}(Y_{x+h} - Y_x = 1) = \theta e^{-x}/x \cdot h$. Define $X_{(i)} = \frac{Y_{(i)}}{Y}$ where $Y = \sum_{i=1}^{\infty} Y_i$. The Poisson-Dirichlet distribution is defined as the probability*

distribution of the vector $(X_{(1)}, X_{(2)}, \ldots)$. *Observe that* $Y = \sum_{i=1}^{\infty} Y_i$ *has a Gamma distribution with parameters* $(\theta, 1)$ *and is independent of* $X_{(i)}$.

Let $L_r^{(n)}$ be the length of the r^{th} longest cycle in a random permutation of $[n]$, $L_r^{(n)} = 0$ if the permutation has fewer than r cycles. Kingman [64] showed that for each r fixed

$$n^{-1}(L_1^{(n)}, \ldots, L_r^{(n)}) \xrightarrow{d} (L_1, \ldots, L_r)$$

where the random vector $(L_1, \ldots, L_r)$ has Poisson-Dirichlet distribution with parameter $\theta = 1$, see also [5, Section 5.4, Section 5.7] for further details.

2.6. A generating function with a simple pole

Suppose we are given a generating function $C(z)$ of connected graphs that has a unique dominant singularity ρ, which is a simple pole, so it looks like

$$C(z) = \frac{\alpha}{1 - z/\rho} + D(z) \tag{2.32}$$

where $D(\rho) = \beta$, $\alpha > 0$ and $D(z)$ has radius of convergence larger than ρ. Several of the combinatorial examples we will study in later chapters are going to have generating functions of this type. In the following, we will analyse this type of generating function completely so that later on we will only have to apply the ready-made results.

Proposition 2.6.1 (The number of graphs — when C has a simple pole). *Assume that* $C(z)$ *is of form* (2.32) *and that* $A(z) = e^{C(z)}$. *As* $n \to \infty$,

$$C_n \sim n!\, \alpha \rho^{-n} \quad \textit{and} \quad A_n \sim n!\, \frac{\alpha^{1/4} e^{\alpha/2+\beta}}{2\sqrt{\pi} n^{3/4}} \rho^{-n} e^{2\sqrt{\alpha n}}. \tag{2.33}$$

In particular, the probability that a random graph of $\mathcal{A}_n$ *is connected tends to* 0 *at speed* $n^{3/4} e^{-2\sqrt{\alpha n}}$ *as* $n \to \infty$.

Proof. The asymptotic behaviour of C_n follows from singularity analysis (see for instance [49, Thm. IV.10, p. 258]). To obtain the asymptotic behaviour of A_n, we first write

$$A(z) = F(z)G(z) \quad \text{with} \quad F(z) = \exp\left(\frac{\alpha}{1 - z/\rho}\right) \quad \text{and} \quad G(z) = e^{D(z)}, \tag{2.34}$$

where $G(z)$ has radius larger than ρ. Now we are going to work out the coefficients of $F(z)$. Instead of working out the details of the saddle point method, we apply the ready-to-use results of Macintyre and Wilson [69, Eqs. (10)–(14)], according to which, for $\alpha, \gamma > 0$:

$$[z^n]\frac{1}{(1-z)^\gamma}\exp\left(\frac{\alpha}{1-z}\right) \sim \frac{\alpha^{1/4}e^{\alpha/2}}{2\sqrt{\pi}n^{3/4}}\left(\frac{n}{\alpha}\right)^{\gamma/2} e^{2\sqrt{\alpha n}}. \tag{2.35}$$

This gives

$$f_n := [z^n]F(z) \sim \frac{\alpha^{1/4}e^{\alpha/2}}{2\sqrt{\pi}n^{3/4}}\rho^{-n}e^{2\sqrt{\alpha n}}.$$

This shows in particular that f_{n-1}/f_n tends to ρ as $n \to \infty$. We can apply Schur's Lemma (Lemma 2.2.1) to (2.34) which gives

$$[z^n]A(z) \sim [z^n]F(z)\cdot e^{D(\rho)} \sim [z^n]F(z)\cdot e^{\beta}.$$

□

Proposition 2.6.2 (Number of components — when C has a simple pole). *Assume* (2.32) *holds and* $A(z) = e^{C(z)}$. *The mean and variance of* $\kappa(R_n)$ *satisfy*

$$\mathbb{E}(\kappa(R_n)) \sim \sqrt{\alpha n}, \qquad \mathbb{V}(\kappa(R_n)) \sim \sqrt{\alpha n}/2,$$

and the random variable $\frac{\kappa(R_n)-\sqrt{\alpha n}}{(\alpha n/4)^{1/4}}$ *converges in law to a standard normal distribution.*

Proof. We want to apply Theorem 2.4.5. The fact that $A(z) = \exp\left(\frac{\alpha}{1-z/\rho} + D(z)\right)$ is Hayman-admissible follows from Theorem 2.4.3, using (2.34) and the H-admissibility of $\exp\left(\frac{\alpha}{1-z}\right)$ (see [49, p. 562]). We have

$$\begin{aligned}
C(r) &\sim \frac{\alpha}{1-r/\rho},\\
b(r) &= r\alpha\left(1-\frac{r}{\rho}\right)^{-2}\rho^{-1} + 2\,r^2\alpha\left(1-\frac{r}{\rho}\right)^{-3}\rho^{-2} + rD'(r) + r^2D''(r)\\
&\sim \frac{2\alpha r^2}{(1-r/\rho)^3},\\
V(r) &= \frac{\alpha}{1-z/\rho} - \frac{r\alpha\left(1-\frac{r}{\rho}\right)^{-1}}{b(r)}\\
&\sim \frac{\alpha}{2}\frac{1}{1-r/\rho}
\end{aligned}$$

as $r \to \rho$. So conditions (2.26)–(2.28) are satisfied and Theorem 2.4.5 applies. Thus, we can conclude that the normalized version of $\kappa(R_n)$ converges in law to

a standard normal distribution. For the asymptotic estimates $\mathbb{E}(\kappa(R_n)) \sim C(\zeta_n)$ and $\mathbb{V}(\kappa(R_n)) \sim V(\zeta_n)$ given by (2.29) where ζ_n is defined as the unique root of the saddle-point equation

$$\zeta_n \, \alpha \left(1 - \frac{\zeta_n}{\rho}\right)^{-2} \rho^{-1} = n.$$

This gives

$$\begin{aligned}\zeta_n &= \frac{-1}{2} \frac{\left(-\alpha - 2\,n + \sqrt{\alpha^2 + 4\,\alpha\,n}\right)\rho}{n} \\ &= \rho - \sqrt{\alpha}\rho\,\sqrt{n^{-1}} + 1/2\,\frac{\alpha\,\rho}{n} - 1/8\,\alpha^{3/2}\rho\,\left(n^{-1}\right)^{3/2} + O\left(\left(n^{-1}\right)^{5/2}\right).\end{aligned}$$

Using the saddle point estimate $\zeta_n = \rho - \rho\sqrt{\alpha/n} + O(1/n)$ we obtain the result. □

Since there are approximately $\sqrt{n}$ components, one may expect the size S_n of the root component to be of the order of $\sqrt{n}$. More precisely, we have the following result.

Proposition 2.6.3 (Size of the root component — when C has a simple pole). *Assume that $C(z)$ is of form (2.32) and that $A(z) = e^{C(z)}$. For $i \geq 0$, the i^{th} factorial moment of S_n satisfies, as $n \to \infty$,*

$$\mathbb{E}(S_n^i) \sim (i+1)!(n/\alpha)^{i/2}.$$

Consequently, the normalized variable $S_n/\sqrt{n/\alpha}$ converges in distribution to a Gamma$(2,1)$ *law, of density xe^{-x} on $[0,\infty)$.*

A local limit law also holds: for $x > 0$ and $k = \lfloor x\sqrt{n/\alpha} \rfloor$,

$$\sqrt{n/\alpha}\ \mathbb{P}(S_n = k) \sim xe^{-x}.$$

Proof. The existence of a local limit law implies the existence of a global limit law. It would therefore make sense to start with the local limit law right away. But spotting the right scaling limit for k is rather difficult. Of course, if a local limit law exists k should be of order of the expected value of S_n. We therefore start with the moments and the global limit law. To obtain the global limit law we are going to use the method of moments explained in Chapter 2.5.

We start from (2.6), with

$$C^{(i+1)}(z) = \frac{\alpha(i+1)!}{\rho^{i+1}(1 - z/\rho)^{i+2}} + D^{(i+1)}(z).$$

Let us first observe that (2.33) implies that $A_n/n!$ is smooth. We can thus apply Schur's Lemma 2.2.1 to the product $D^{(i+1)}(z)A(z)$. This gives

$$\frac{[z^{n-i-1}]D^{(i+1)}(z)A(z)}{n[z^n]A(z)} \sim \frac{D^{(i+1)}(\rho)}{n}\frac{A_{n-i-1}}{(n-i-1)!}\frac{n!}{A_n} \sim \frac{D^{(i+1)}(\rho)}{n}\rho^i \to 0.$$

Thus, we have

$$\frac{A_n}{(n-1)!}\,\mathbb{E}(S_n^i) \sim [z^{n-i-1}]\frac{\alpha(i+1)!}{\rho^{i+1}(1-z/\rho)^{i+2}}\exp\left(\frac{\alpha}{1-z/\rho}+D(z)\right). \tag{2.36}$$

Now (2.35) gives

$$[z^{n-i-1}]\frac{\alpha(i+1)!}{\rho^{i+1}(1-z/\rho)^{i+2}}\exp\left(\frac{\alpha}{1-z/\rho}\right) \sim \alpha(i+1)!\,\frac{\alpha^{1/4}e^{\alpha/2}}{2\sqrt{\pi}n^{3/4}}\left(\frac{n}{\alpha}\right)^{i/2+1}\rho^{-n}e^{2\sqrt{\alpha n}}. \tag{2.37}$$

In particular, this sequence of coefficients is smooth. Hence by Schur's Lemma 2.2.1, the asymptotic behaviour of (2.36) only differs from (2.37) by a factor e^β, where $\beta = D(\rho)$. Combined with (2.33), this gives the limiting i^{th} moment of S_n. Since these moments characterize the above Gamma distribution, we can conclude by the method of moments that $S_n/\sqrt{n/\alpha}$ converges in law to Gamma$(2,1)$.
For the local (and hence global) limit law, we combine (2.5) with (2.33)

$$\begin{aligned}\mathbb{P}(S_n=k) &= \frac{k}{n}\frac{C_k}{k!}\frac{A_{n-k}}{(n-k)!}\frac{n!}{A_n}\\ &\sim \frac{k}{n}\alpha\frac{n^{3/4}}{(n-k)^{3/4}}e^{2\sqrt{\alpha}(\sqrt{n-k}-\sqrt{n})}.\end{aligned}$$

We try k of the order of the first moment, i.e. $k = x\sqrt{n/\alpha}$. In this case $\left(\frac{n}{n-k}\right)^{3/4} \to 1$ as $n\to\infty$ and $2\sqrt{\alpha}(\sqrt{n-k}-\sqrt{n}) \to -x$ which implies the result. □

2.7. Random generation using Boltzmann models

Suppose we wish to generate a random combinatorial object of size n from some class $\mathcal{A}$, and all objects of size n should be equally likely. For n very large, we might also be happy with an object of size $n+1$ rather than n given that the uniformity condition still holds. A *Boltzmann model* relative to a labelled combinatorial class $\mathcal{A}$ depends on a real-valued control parameter $x>0$. Any object $H\in\mathcal{A}$ is sampled with probability

$$\mathbb{P}(H) = \frac{x^{|H|}}{|H|!A(x)}, \tag{2.38}$$

provided $A(x)$ is finite. So any two objects of equal size are equally likely to be sampled. The distribution is spread all over $\mathcal{A}$ but on each size the distribution is uniform. A *Boltzmann sampler* $\Gamma A(x)$ for the class $\mathcal{A}$ is a process that produces objects from $\mathcal{A}$ according to this model. The main idea is now to adjust the parameter x so that the Boltzmann draws a structure of size close to n, instead using a procedure that draws samples of exactly size n. In fact the construction of the Boltzmann sampler is rather straightforward if we know the symbolic specification of the class $\mathcal{A}$.

For each of the minor-closed classes $\mathcal{A}$ that we will study in Chapter 3, we have designed an associated *Boltzmann sampler*. I refer to [46, Sec. 4] for general principles on the construction of exponential Boltzmann samplers, and only briefly describe how we have addressed certain specific difficulties at the end of Chapter 3. Most of them are related to the fact that our graphs are unrooted.

3. Excluding connected minors

3.1. Introduction: Excluding 2-connected minors

Let $\mathcal{A}$ be a labelled minor-closed class closed under isomorphisms. We study the asymptotic properties of a random graph R_n taken uniformly in $\mathcal{A}_n$, the set of graphs of $\mathcal{A}$ having n vertices ($R_n \in_u \mathcal{A}$): what is the probability that R_n is connected? More generally, what is the number $\kappa(R_n)$ of connected components? What is the size S_n of the *root component*, that is, the component containing the vertex labelled 1? Or the size L_n of the largest component? These are the questions we would like to explore in this chapter much in the same spirit as [71], [73], [75]. We start with a well studied case, the case of excluding 2-connected minors. All graph classes in the following are labelled and we will not state this explicitly any more. Recall that many interesting classes of graphs such as forests, series-parallel graphs, outer-planar graphs, planar graphs are minor-closed and their minimal forbidden minors (=excluded minors) are 2-connected. Several results are known in this case, see [71]. We state some of them as well as a new result dealing with the size of the root component. The general picture is that the class $\mathcal{A}$ shares many properties with the class of forests analysed in detail in Section 2.3.1. As we have seen in the previous chapter, not only having a growth constant but *smoothness* is one of the crucial fundamental properties.

Theorem 3.1.1 (**Smoothness[71]**). *Let $\mathcal{A}$ be a minor-closed class of graphs and let all excluded minors be 2-connected, that is, $\mathcal{A}$ is minor-closed, decomposable and bridge-addable. Then $\mathcal{A}$ is smooth.*

Recall that every any proper minor-closed class has a strictly positive radius of convergence [80] and in particular by Theorem 3.1.1 every minor-closed class of graphs where the excluded minors are 2-connected has a growth constant [76]. Furthermore, by [20] the growth constant which is the inverse of the radius of convergence in this case is strictly positive. The following Theorem then gives us information about the behaviour of the generating function near its radius of convergence.

Theorem 3.1.2 ([71]). *Let $\mathcal{A}$ be a minor-closed class of graphs and let all excluded minors be 2-connected. The generating functions $C(z)$ and $A(z) = e^{C(z)}$ are finite at their (positive) radius of convergence ρ.*

Furthermore, we have the following results about $\kappa(R_n)$ the number of components of $R_n \in_u \mathcal{A}$.

Proposition 3.1.3 (**Number of components – converging case [71]**). *Let $\mathcal{A}$ be a minor-closed class of graphs and let all excluded minors be 2-connected. Then*

$$\mathbb{P}(R_n \text{ is connected}) \to \frac{1}{A(\rho)}$$

as $n \to \infty$ and moreover

$$\kappa(R_n) \xrightarrow{d} 1 + \mathrm{Po}\left(C(\rho)\right).$$

Observe that every minor-closed closed class where all excluded minors are 2-connected is bridge-addable. By [76, Theorem 2.1] and [1],[62] we have

$$\lim_{n\to\infty} \mathbb{P}(R_n \text{ is connected}) \geq \frac{1}{\sqrt{e}}.$$

The latter value is reached when $\mathcal{A}$ is the class of forests. For further discussion about bridge-addable classes which are not necessarily minor-closed see Chapter 4.

Furthermore, Lemma 2.2.4 and Proposition 2.2.5 apply and we obtain for $k \in \mathbb{N}$ fixed

Proposition 3.1.4. *Let $\mathcal{A}$ be a minor-closed class of graphs and let all excluded minors be 2-connected. The random variables $n - S_n$ and $n - L_n$ both converge to a discrete limit distribution X given by*

$$\mathbb{P}(X = k) = \frac{1}{A(\rho)} \frac{A_k \rho^k}{k!}.$$

A more precise result is actually available. Recall that $\mathrm{Frag}(R_n)$ is the union of the components that are not in the biggest component. Then in [71, Thm. 1.5] the limit law of $\mathrm{Frag}(R_n)$ is described, not only of its size.

3.2. Tree-like structures

What happens if we exclude non-2-connected minors? In [71] it was already indicated that the behaviour of the class might change drastically. But as we will see in the following this is certainly not always the case, i.e. there exist minor-closed classes of graphs defined by excluding a non-2-connected minor that behave very similarly to the 2-connected case described in Theorem 3.1.3.
Suppose $\mathcal{A}$ is a decomposable class of graphs (not necessarily minor-closed). We say $\mathcal{A}$ is *dominated by trees* if it satisfies the following conditions:

1. $\mathcal{A}$ contains all trees,

2. the generating function $D(z)$ that counts the connected graphs of $\mathcal{A}$ that are **not** trees has radius of convergence (strictly) larger than $1/e$ (which is the radius of trees), i.e. there are exponentially fewer graphs than trees.

In this case, the properties that hold for forests still hold, except that the probability C_n/A_n that R_n is connected tends to a limit that is now *at most* $1/\sqrt{e}$ which is the lower bound for *bridge-alterable* classes of graphs proved by [1], [62]. We will see that there exist examples of minor-closed classes for which this limit is still bounded away from 0 but arbitrarily small.

Proposition 3.2.1 (**The number of graphs — when trees dominate).** *Let $\mathcal{A}$ be a class of graphs dominated by trees and let $T(z) = ze^{T(z)}$ be the generating function of rooted trees. Write the generating function of connected graphs in the class $\mathcal{A}$ as*

$$C(z) = T(z) - \frac{T(z)^2}{2} + D(z).$$

The generating function of graphs of $\mathcal{A}$ is $A(z) = e^{C(z)}$. As $n \to \infty$,

$$C_n \sim n! \frac{e^n}{\sqrt{2\pi} n^{5/2}} \quad \text{and} \quad A_n \sim A(1/e) C_n.$$

In particular, the probability that R_n is connected tends to $1/A(1/e) = e^{-1/2 - D(1/e)}$ as $n \to \infty$.

Proof. We use singularity analysis described in Chapter 2.3. By assumption, $D(z)$ has radius of convergence strictly larger than $1/e$. Hence it is analytic at $1/e$ and we get the following Taylor expansion around $z = 1/e$

$$D(z) = D(1/e) + (1/e - z)D'(1/e) + O((1 - ze)^2). \tag{3.1}$$

The singular behaviour of $C(z)$ is therefore that of unrooted trees given by (2.15).

Combining (2.15) and (3.1) we obtain

$$C(z) = \frac{1}{2} + D(1/e) - (1-ze)\left(1 + \frac{D'(1/e)}{e}\right) + \frac{2\sqrt{2}}{3}(1-ze)^{3/2} + O((1-ze)^2), \quad (3.2)$$

as z approaches $1/e$ and this expansion is valid in a Δ-domain. By the Transfer Theorem 2.3.4 it suffices to analyse the coefficients of the expansion of $C(z)$. Applying Theorem 2.3.2 to the expansion of $C(z)$ (the dominant term for the growth of the coefficients is $(1-ze)^{3/2}$ in this case) gives the estimate of C_n. Observe that we can write

$$C(z) = \frac{1}{2} + D(1/e) + F(z) \quad (3.3)$$

and $F(z) \to 0$ as z tends to $\frac{1}{e}$. For the series $A(z) = e^{\frac{1}{2}+D(1/e)+F(z)}$ we therefore find

$$\begin{aligned} A(z) &\sim e^{1/2+D(1/e)} \cdot (1 + F(z)) \\ &= e^{\frac{1}{2}+D(1/e)} \left(1 - (1-ze)\left(1 + \frac{D'(1/e)}{e}\right) + \frac{2\sqrt{2}}{3}(1-ze)^{3/2} + O((1-ze)^2)\right), \end{aligned}$$

and the estimate of A_n follows by singularity analysis (same argument as for C_n). □

Proposition 3.2.2 (**Number of components — when trees dominate).** *We have for $n \to \infty$*

$$\kappa(R_n) \sim 1 + \mathrm{Po}\left(C\left(\frac{1}{e}\right)\right)$$

and in particular $\mathbb{E}(\kappa(R_n)) \sim 1 + C(1/e)$.

Proof. There are several ways to prove the result. The easiest is to apply Theorem 2.2.3. By Proposition 3.2.1 we know that $\mathcal{A}$ is smooth, i.e. nC_{n-1}/C_n converges and $C(z)$ has strictly positive radius of convergence equal to $\frac{1}{e}$. Furthermore, (3.2) implies that $C(\frac{1}{e}) < \infty$. So it remains to observe that C_n/A_n converges. □

Proposition 3.2.3 (**Size of components — when trees dominate).** *The random variable $n - S_n$ converges to a discrete limit distribution X given by*

$$\mathbb{P}(X = k) = \frac{1}{A(1/e)} \frac{A_k e^{-k}}{k!},$$

where A_k and $A(z)$ are given in Proposition 3.2.1*. The same holds for $n - L_n$.*

Proof. The two ingredients used in the proof of Proposition 2.2.5 to establish the limit law of $n - S_n$ (namely, smoothness of $A_n/n!$ and convergence of C_n/A_n), still hold here (see Proposition 3.2.1). Lemma 2.2.4 then gives the law of $n - L_n$. □

We now present a collection of classes dominated by trees. The following Lemma will be useful later.

Lemma 3.2.4. *Fix $k \in \mathbb{N}$. Let $\mathcal{T}_k$ denote the class of rooted (labelled) trees of height $\leq k$. Then*

$$T_k(z) = z \cdot e^{T_{k-1}(z)} \tag{3.4}$$

where $T_1(z) = z$. Denoting ρ_k the radius of convergence of $T_k(z)$ we have $\rho_k > \frac{1}{e}$, $T_k\left(\frac{1}{e}\right) < 1$ and $\lim_{k\to\infty} T_k(\rho_k) = 1$.

Proof. Every rooted tree of height $\leq k+1$ consists of a specified vertex (the root) and trees of height $\leq k$ attached to it. We obtain the recursive specification $\mathcal{T}_k = \mathcal{Z} \times \mathrm{SET}(\mathcal{T}_k)$ which translates into (3.4) by the symbolic method. Note that for fixed k the function $T_k(z)$ is entire. Also, $T_k(z) < T(z)$ for all $z \in [0, 1/e]$ since T_k counts fewer trees than $T(z)$ which is the function of all rooted trees. In particular, $T_k(1/e) < T(1/e) = 1$ and $\lim_{k\to\infty} T_k\left(\frac{1}{e}\right) = T\left(\frac{1}{e}\right) = 1$. □

Proposition 3.2.5. *Let $k \geq 1$. Let $\mathcal{A}$ be a decomposable class of graphs not necessarily minor-closed that contains all trees, and such that all graphs of $\mathcal{A}$ exclude the k-spoon (shown in Figure 3.2). Then $\mathcal{A}$ is dominated by trees, and the results of Propositions* 3.2.1, 3.2.2 *and* 3.2.3 *hold.*

Proof. We have $\mathcal{A} \subset$ Ex(k-spoon), so clearly it suffices to prove the result for $\mathcal{A} =$ Ex(k-spoon) which we henceforth assume. We partition the set $\mathcal{C}$ of connected

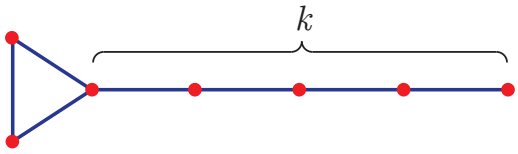


Figure 3.1.: The k-spoon where k counts the number of edges in the handle.

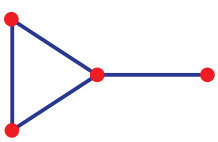


Figure 3.2.: The 1-spoon.

graphs of $\mathcal{A}$ into three subsets: the set $\mathcal{C}_0$ of trees, counted by $C_0 = T - T^2/2$ with $T \equiv T(z)$, the set $\mathcal{C}_1$ of unicyclic graphs (counted by C_1), and finally the set $\mathcal{C}_2$ containing graphs with at least two cycles (counted by C_2). Hence $C = T - T^2/2 + C_1 + C_2$. We will prove that C_1 has radius strictly larger than $1/e$, and that C_2 is entire. A unicyclic graph belongs to $\mathcal{C}$ if and only if all trees attached to its unique cycle have height strictly less than k. We want to obtain the generating function of cycles. There are $(n-1)!$ directed cycles (say anticlockwise) for $n \geq 3$, as a directed cycle is determined by the succession of elements that ' follow' 1 and so a permutation of $n-1$ elements. Hence there are $\frac{1}{2}(n-1)!$ undirected cycles for $n \geq 3$. The generating function of cycles is therefore

$$\mathrm{Cyc}(z) = \frac{1}{2}\sum_{n\geq 3}\frac{(n-1)!}{n!}z^n = \frac{1}{2}\sum_{n\geq 3}\frac{z^n}{n} = \frac{1}{2}\left(\log\frac{1}{1-z} - z - \frac{z^2}{2}\right). \tag{3.5}$$

Graphs in $\mathcal{A}$ that contain exactly one cycle might have trees of height strictly less than k grafted to them. Hence to construct all graphs in $\mathcal{A}$ that contain exactly one cycle we have to substitute rooted trees of height strictly less than k for vertices in the cycles. By the symbolic method (see the beginning of Chapter 2.1 and [49, Chap. II] for more details) this translates into substituting the generating function of rooted trees $T_{k-1}(z)$ into the generating function of cycles

$$C_1(z) = \mathrm{CYC}(T_{k-1}) = \frac{1}{2}\left(\log\frac{1}{1-T_{k-1}(z)} - T_{k-1}(z) - \frac{T_{k-1}(z)^2}{2}\right). \tag{3.6}$$

So by Lemma 3.2.4 $C_1(z)$ has radius of convergence strictly larger than $\frac{1}{e}$.

We now want to prove that $C_2(z)$ is entire. The (2)-*core* of a connected graph G is the (possibly empty) unique maximal subgraph of minimal degree (at least) 2. Subsequently, any connected graph of minimal degree ≥ 2 is therefore called a *core*. Let $\bar{\mathcal{C}}_2$ denote the set of cores in Ex(k-spoon) which have ≥ 2 cycles and $\bar{C}_2$ the associated generating function. Similarly to unicyclic graphs where rooted trees could be substituted for vertices, for any core in $\mathcal{A}$ vertices can be substituted by rooted trees of height strictly less than k. By the symbolic method we have $C_2(z) = \bar{C}_2(T_{k-1}(z))$ and we want to show that the contribution of $C_2(z)$ for $z = 1/e$ is negligible. More precisely, we need to show that $\bar{C}_2(z)$ is entire. It follows from [20, Thm. 3.1] that it suffices to prove that no core in $\bar{\mathcal{C}}_2$ contains a path of length $3k-1$. We are going to show the following: Every graph H with $\delta(H) \geq 2$ that contains at least 2 cycles and a path of length $l \geq 3k-1$ contains a k-spoon as a minor. Let $P = (v_0, v_1, \ldots, v_\ell)$ be a path of maximal length in G, and assume that $\ell \geq 3k-1$. We will prove that G contains the k-spoon as a minor.

Since P is maximal and G is a core, there exist v_i and v_j, with $i \geq 2$ and $j \leq \ell - 2$, such that the edges $\{v_0, v_i\}$ and $\{v_j, v_\ell\}$ belong to G.

If $i = \ell$ or $j = 0$, let $\bar{P}$ be the cycle of G formed of P and the edge $\{v_0, v_\ell\}$. Let $\bar{Q}$ be another cycle of G. If $\bar{Q}$ contains at most one vertex of P (Figure 3.3(a)), we find an ℓ-spoon by deleting one edge of $\bar{P}$, contracting $\bar{Q}$ into a 3-cycle and one of the paths joining P to $\bar{Q}$ into a point. If $\bar{Q}$ contains at least two vertices v_a and v_b of P, with $a < b$ (Figure 3.3(b)), we may assume that $\bar{Q}$ consists of the edges $\{v_a, v_{a+1}\}, \ldots, \{v_{b-1}, v_b\}$ and of a path Q that only meets P at v_a and v_b. Let $\bar{R}$ denote the cycle formed of the path Q and the path $(v_b, v_{b+1}, \ldots, v_\ell, v_0, \ldots, v_a\}$. Then we obtain a p-spoon, with $p \geq \lceil 3k/2 \rceil - 1 \geq k$, by contracting the shortest of the cycles $\bar{Q}$ and $\bar{R}$ into a 3-cycle and deleting an edge ending at v_a from the other.

Assume now that $i < \ell$ and $j > 0$. Suppose first that $i \leq j$ (Figure 3.3(c)). By symmetry, we may assume that the cycle $\bar{P}_1 = (v_0, \ldots, v_i)$ is shorter than (or equal in length to) the cycle $\bar{P}_2 = (v_j, \ldots, v_\ell)$. In particular, $i \leq \ell/2$. Contract $\bar{P}_1$ into a 3-cycle, and remove the edge $\{v_j, v_\ell\}$ from $\bar{P}_2$: this gives a p-spoon with $p = \ell - i \geq \lceil \ell/2 \rceil \geq k$. Assume now that $j < i$ (Figure 3.3(d)). Consider the three following paths joining v_i and v_j: $(v_i, v_{i-1}, \ldots, v_j)$, $(v_i, v_0, v_1, \ldots, v_j)$ and $(v_i, v_{i+1}, \ldots, v_\ell, v_j)$. Since the sum of the lengths of these paths is $\ell + 2 \geq 3k + 1$, one of them, say $(v_i, v_0, v_1, \ldots, v_j)$, has length at least $k+1$. That is, $j \geq k$. Delete from this path the edge $\{v_i, v_0\}$, and contract the cycle formed by the other two paths into a 3-cycle: this gives a j-spoon, with $j \geq k$. □

The simplest non-trivial class of graphs satisfying the conditions of Proposition 3.2.5 is $\mathcal{A} = \mathrm{Ex}(1\text{-spoon})$. By specializing to $k = 1$ the previous proof gives $C_1 = \mathrm{Cyc}(z)$ and $C_2 = 0$ (since no core having several cycles avoids the 2-path). Hence

$$C(z) = T(z) - \frac{T(z)^2}{2} + \frac{1}{2}\left(\log \frac{1}{1-z} - z - \frac{z^2}{2}\right).$$

More generally, consider the class $\mathcal{A}^{(k)} = \mathrm{Ex}(\mathrm{K}_4 - \mathrm{e}, \text{bow tie}, \text{k-spoon})$ where $K_4 - e$ and bow tie are shown in Figure 3.4 and 3.5, respectively: excluding the latter two graphs means that no graph of $\mathcal{C}$ can have several cycles, so that $C_2 = 0$ and the following result follows.

Proposition 3.2.6 (No diamond, bow tie or k-spoon). *Let $k \geq 1$. Define $\mathcal{A}^{(k)} = \mathrm{Ex}(\mathrm{K}_4 - \mathrm{e}, \textit{bow tie}, \mathrm{k}\textit{-spoon})$ the class of graphs excluding the diamond K_4-e, the bow tie and the k-spoon as a minor. The class $\mathcal{A}^{(k)}$ is dominated by trees, and*

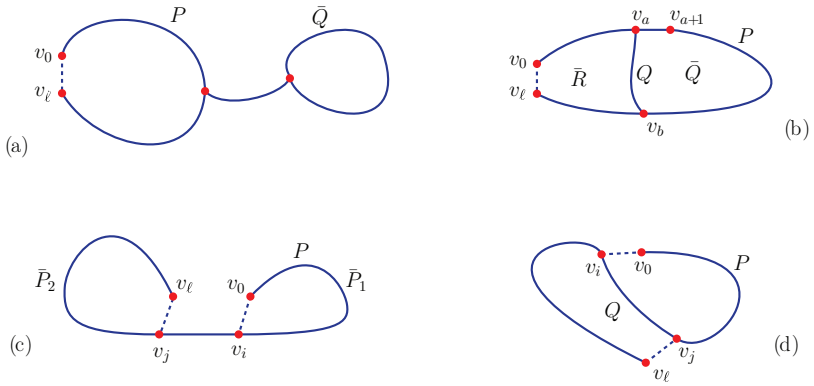


Figure 3.3.: A core having several cycles and excluding the k-spoon cannot contain a path of length $3k-1$.

the results of Propositions 3.2.1, 3.2.2 *and* 3.2.3 *hold. In particular, the probability that* $R_n \in_u \mathcal{A}_n^{(k)}$ *is connected tends to* $p_k = e^{-C^{(k)}(1/e)} > 0$ *as* $n \to \infty$ *where*

$$C^{(k)}(z) = T(z) - \frac{T(z)^2}{2} + D^{(k)}(z)$$

and

$$D^{(k)}(z) = \frac{1}{2}\left(\log\frac{1}{1-T_k(z)} - T_k(z) - \frac{T_k(z)^2}{2}\right)$$

but $\lim_{k\to\infty} p_k = 0$.

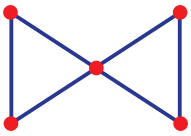


Figure 3.4.: The bow tie graph

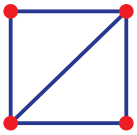


Figure 3.5.: The diamond graph K_4 minus an edge.

Proof. Let $T(z)$ be the generating function of rooted trees, given by $T(z) = ze^{T(z)}$, and let $T_k(z)$ be the generating function of rooted trees of height strictly less than k. The connected structures in $\mathcal{A}^{(k)}$ are precisely trees and cycles where vertices have been substituted for rooted trees of height $\leq k$. So they are clearly a subset of so-called unicyclic graphs. By (3.6) the generating function of these graphs is

$$C^{(k)}(z) = T(z) - \frac{T(z)^2}{2} + D^{(k)}(z)$$

where

$$D^{(k)}(z) = \frac{1}{2}\left(\log\frac{1}{1-T_k(z)} - T_k(z) - \frac{T_k(z)^2}{2}\right).$$

By Lemma 3.2.4 $T_k(z)$ is entire and $T_k\left(\frac{1}{e}\right) < 1$. Hence the radius of convergence of $D^{(k)}(z)$ is strictly larger than $\frac{1}{e}$ and $\mathcal{A}^{(k)}$ is dominated by trees. On the other hand, $\lim_{k\to\infty} T_k\left(\frac{1}{e}\right) = 1$ (Lemma 3.2.4) and therefore $D^{(k)}(\frac{1}{e}) \to \infty$ as $k \to \infty$ which implies $\lim_{k\to\infty} p_k = \lim_{k\to\infty} e^{-1/2-D(1/e)} = 0$. □

3.2.1. Excluding a spoon

For the class of graphs excluding the 2-spoon as a minor we can compute the corresponding generating function and constants explicitly. For spoons with longer handles, i.e. $k > 2$ technically the same method (effectively determine the cores

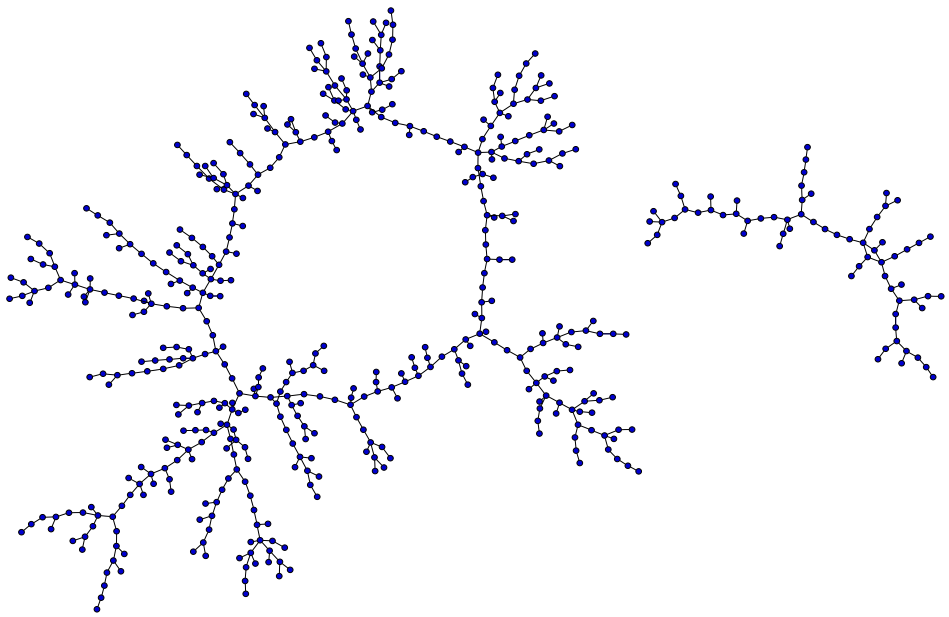


Figure 3.6.: A random graph of size $n = 541$ avoiding the diamond, the bow tie and the 20-spoon.

and substituting trees of small height for some vertices) applies, but the generating functions get more and more involved.

Proposition 3.2.7. *The class of graphs $\mathcal{A} = \mathrm{Ex}(2\text{-spoon})$ is dominated by trees, and the results of Propositions* 3.2.1*,* 3.2.2 *and* 3.2.3 *apply: As $n \to \infty$,*

$$C_n \sim n! \frac{e^n}{\sqrt{2\pi} n^{5/2}} \quad \text{and} \quad A_n \sim A(1/e) C_n$$

where $A(1/e) \approx e^{1/2+0.046}$. The number of components $\kappa(R_n)$ satisfies:

$$\kappa(R_n) \sim 1 + \mathrm{Po}(C(1/e))$$

with $C(1/e) \approx 1/2 + 0.046$ and the random variable $n - S_n$ converges to a discrete limit distribution X given by

$$\mathbb{P}(X = k) = \frac{1}{A(1/e)} \frac{A_k e^{-k}}{k!}.$$

The same holds for $n - L_n$.

Proof. Let $T(z)$ be the generating function of rooted trees, given by $T(z) = ze^{T(z)}$. Then $A(z) = e^{C(z)}$ and we claim that the generating function of connected graphs is

$$C(z) = T(z) - \frac{T(z)^2}{2} + D(z)$$

with

$$D(z) = \frac{1}{2}\left(\log \frac{1}{1 - ze^z} - ze^z - \frac{z^2 e^{2z}}{2}\right) + \frac{z^4}{4!} + z^2 e^{2z}\left(e^z - 1 - z - \frac{z^2}{4}\right). \quad (3.7)$$

We first follow the proof of Proposition 3.2.5: we write $C = T - T^2/2 + C_1 + C_2$, where C_1 is given by (3.6) with $T_k = T_2 = ze^z$, and C_2 counts connected graphs having several cycles and avoiding the 2-spoon. Note that C_1 is the first term in the above expression of $D(z)$. Let us now focus on C_2. Clearly, K_4 does not contain the 2-spoon as a minor. Moreover, none of its vertices can be replaced by a non-trivial tree. This gives the term $z^4/4!$ in $D(z)$. If we have more than 1 cycle in a graph, all cycles have to share at least 1 edge otherwise we would create a bow tie which then itself contains the 2-spoon as a minor. Every cycle of length at least 5 and a chord contains a 2-spoon. So we are left with the enumeration of graphs which contain only cycles of length 3 or 4 and where some vertices have been replaced by stars. The two possible families are shown in Figure 3.7.

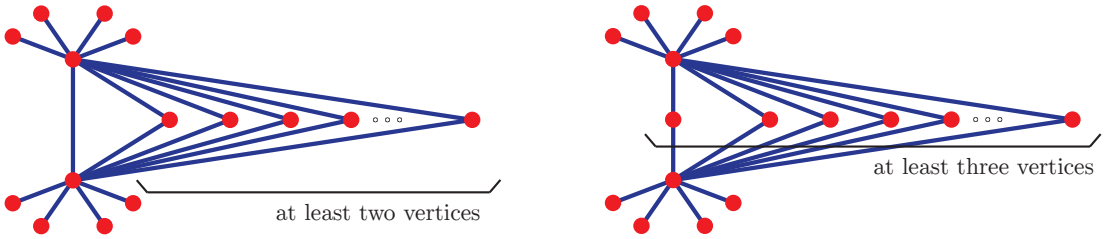


Figure 3.7.: Graphs with several cycles and no 2-spoon minor.

Stars are a set of vertices where one vertex is specified (the center vertex) and all vertices are connected to the center vertex. We have

$$\text{Stars} = \mathcal{Z} \times \text{SET}(\mathcal{Z}). \tag{3.8}$$

Applying the symbolic method see [49, Chap. II]) gives that the two families have the contribution

$$\frac{1}{2}(ze^z)^2\left(e^z - 1 - z\right) + \frac{1}{2}(ze^z)^2\left(e^z - 1 - z - \frac{z^2}{2}\right).$$

□

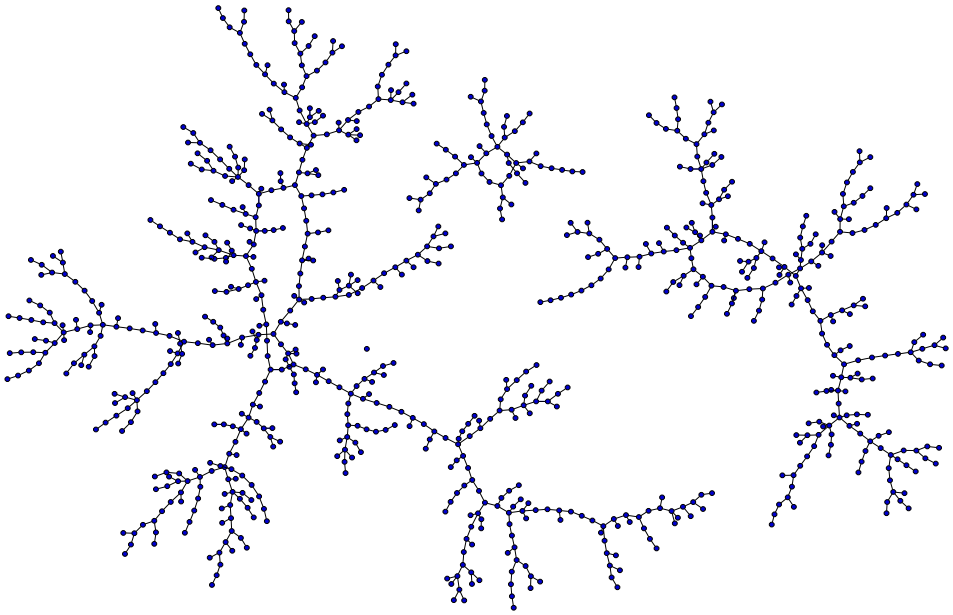


Figure 3.8.: A random graph of size $n = 859$ avoiding the diamond and the bow tie.

3.3. Excluding the diamond and the bow tie: a logarithmic singularity

Let $\mathcal{A} = \mathrm{Ex}(\mathrm{K}_4\text{-e}, \text{bow tie})$ be the class of graphs excluding the diamond and the bow tie shown in Figure 3.4 and 3.5. As we will see, these are the graphs precisely unicyclic graphs or pseudoforests and each component has at most one cycle. They were studied in the 70's by Wright [93] and crop up at various points in Combinatorics, in particular when $G_{n,p}$ reaches the so-called critical point. The enumeration results mentioned in this chapter are originally due to Rényi in 1959 [84], we refer to [23, p. 119] or [49, p.133].

Proposition 3.3.1 (**The number of graphs excluding a diamond and a bow tie as a minor**). *Let $\mathcal{A} = \mathrm{Ex}(\mathrm{K}_4 - \mathrm{e}, \textit{bow tie})$. As $n \to \infty$,*

$$C_n \sim n! \frac{e^n}{4n} \quad \textit{and} \quad A_n \sim n! \, \frac{1}{(2e)^{1/4}\Gamma(1/4)} \frac{e^n}{n^{3/4}}. \tag{3.9}$$

In particular, the probability that R_n is connected tends to 0 at speed $n^{-1/4}$ as $n \to \infty$.

Proof. Let $T(z)$ be the generating function of rooted trees, defined by $T(z) = ze^{T(z)}$. The following expression of $C(z)$ is obtained by taking the limit $k \to \infty$ in Proposition 3.2.6 (equivalently, see [49, equation (52), p.133])

$$C(z) = \frac{T(z)}{2} - \frac{3T(z)^2}{4} + \frac{1}{2} \log \frac{1}{1 - T(z)}.$$

The generating function of graphs of all graphs $\mathcal{A}$ is $A(z) = e^{C(z)}$. Following [49, p.133] we now estimate C_n and A_n via singularity analysis, see Chapter 2.3 for an explanation. Recall from Example 2.3.1 Chapter 2.3 that $T(z)$ has a unique dominant singularity at $z = 1/e$ with the singular expansion given in (2.15) valid in a Δ-domain. Thus $1/e$ is also the unique dominant singularity of $C(z)$ and of $A(z)$, and the following expansion is valid in a Δ-domain

$$C(z) \sim \frac{1}{4} \log \left(\frac{1}{1 - ze} \right) \quad \text{and} \quad A(z) \sim \frac{1}{(2e)^{1/4}(1 - ez)^{1/4}}. \tag{3.10}$$

The asymptotic estimates of C_n and A_n follow. □

Proposition 3.3.2 (**Number of components — no bow tie nor diamond**). *Let $\mathcal{A} = \mathrm{Ex}(\mathrm{K}_4 - \mathrm{e}, \textit{bow tie})$. The mean and variance of $\kappa(R_n) \in_u \mathcal{A}$ satisfy:*

$$\mathbb{E}(\kappa(R_n)) \sim \log n/4, \qquad \mathbb{V}(\kappa(R_n)) \sim \log n/4,$$

and the random variable $\frac{\kappa(R_n)-\log n/4}{\sqrt{\log n/4}}$ *converges in law to a standard normal distribution.*

Proof. Using (2.15), the estimate (3.10) can be refined into

$$C(z) = \frac{1}{4}\log\left(\frac{1}{1-ze}\right) + \lambda + O(\sqrt{1-ze}), \tag{3.11}$$

where λ is a constant. The proposition is then a direct application of Proposition 2.3.5. □

The number of connected components is about $1/4 \log n$. However, the size of the root component is found to be of order n. More precisely, we have the following result.

Proposition 3.3.3 (**Size of the root component — no bow tie nor diamond**). *The normalized variable* S_n/n *converges in distribution to a beta law with parameters* $\alpha = 1, \beta = 1/4$ *and density* $(1-x)^{-3/4}/4$ *on* $[0,1]$. *In fact, a local limit law holds: for* $x \in (0,1)$ *and* $k = \lfloor xn \rfloor$,

$$n\,\mathbb{P}(S_n = k) \to \frac{1}{4}(1-x)^{-3/4}.$$

The convergence of factorial moments holds as well: for $i \geq 0$,

$$\mathbb{E}(S_n^i) \sim \frac{\Gamma(5/4)i!}{\Gamma(i+5/4)} n^i.$$

Proof. Recall that the existence of a local limit law implies the existence of a global one [21, Thm. 3.3]. Thus it suffices to prove the local limit law. But once we have the first moment it will be much easier to guess the correct limit law. We therefore start with the moments. Let us start from (2.6). Our first task is to obtain an estimate of $C^{(i+1)}(z)$ near $1/e$. By [49, Thm. VI.8, p. 419] it suffices to differentiate (3.11). For $i \geq 1$,

$$C^{(i+1)}(z) \sim \frac{i!}{4}\left(\frac{e}{1-ze}\right)^{i+1}.$$

Then by (3.10) and singularity analysis (Theorems 2.3.4 and 2.3.2)

$$\begin{aligned}[z^{n-i-1}]C^{(i+1)}(z)A(z) &\sim \frac{i!e^{i+1}}{4(2e)^{1/4}}[z^{n-i-1}]\left(\frac{1}{1-ze}\right)^{i+1+1/4} \\ &\sim \frac{i!e^{i+1}}{4(2e)^{1/4}} e^{n-i-1}\frac{(n-i-1)^{i+1/4}}{\Gamma(i+5/4)}\end{aligned}$$

Using (3.9) we obtain, that the i^{th} factorial moment of S_n is asymptotically

$$\frac{[z^{n-i-1}]C^{(i+1)}(z)A(z)}{n[z^n]A(z)} \sim \frac{i!\Gamma(1/4)}{4\Gamma(i+5/4)} \cdot \frac{(n-i-1)^{i+1/4}}{n^{1/4}} \sim \frac{i!\Gamma(1/4)}{4\Gamma(i+5/4)} n^i$$

as $\Gamma(5/4) = \Gamma(1/4) \cdot 1/4$. Since these moments characterize the above beta distribution, we conclude by the method of moments (see proof of Proposition 2.6.3 for an explanation) that S_n/n converges in law to this distribution. For the local limit law we start from the rightmost expression in (2.5) and using (3.9):

$$\mathbb{P}(S_n = k) \sim \frac{k}{n}\frac{e^k}{4k}\frac{1}{(2e)^{1/4}\Gamma(1/4)}\frac{e^{n-k}}{(n-k)^{3/4}} \cdot \frac{(2e)^{1/4}\Gamma(1/4)}{1}\frac{n^{3/4}}{e^n} \sim \frac{1}{4}n^{-1/4}(n-k)^{-3/4}.$$

So for $k = \lfloor xn \rfloor$ which is of order of the first moment, $\mathbb{P}(S_n = k) \sim n^{-1}\frac{1}{4}(1-x)^{-3/4}$. □

We conclude with the law of the size of the largest component, which we derive from general results dealing with components of *logarithmic structures* [5]. The definitions used in [5] are very different from the definitions used in the study of minor-closed classes of graphs and we will explain them 'on the go' in the proof. For more information on the Poisson-Dirichlet distribution see Chapter 2.5. The following proposition is illustrated by Figure 3.9.

Proposition 3.3.4 (Size of the largest component — no bow tie nor diamond). *The normalized variable L_n/n converges in law to the first component of a Poisson-Dirichlet distribution of parameter $1/4$: for $x \in (0,1)$,*

$$\mathbb{P}(L_n < xn) \to r_{1/4}(1/x),$$

where $r_{1/4} : \mathbb{R}^+ \to [0,1]$ is the unique continuous function such that $r_{1/4}(x) = 1$ for $x \in [0,1]$ and for $x > 1$,

$$x^{1/4} r'_{1/4}(x) + \frac{1}{4}(x-1)^{-3/4} r_{1/4}(x-1) = 0.$$

A local limit law also holds: for $x \in (0,1)$ and $1/x \notin \mathbb{N}$,

$$n\,\mathbb{P}(L_n = \lfloor xn \rfloor) \to \frac{(1-x)^{-3/4}}{4x}\, r_{1/4}\left(\frac{1-x}{x}\right).$$

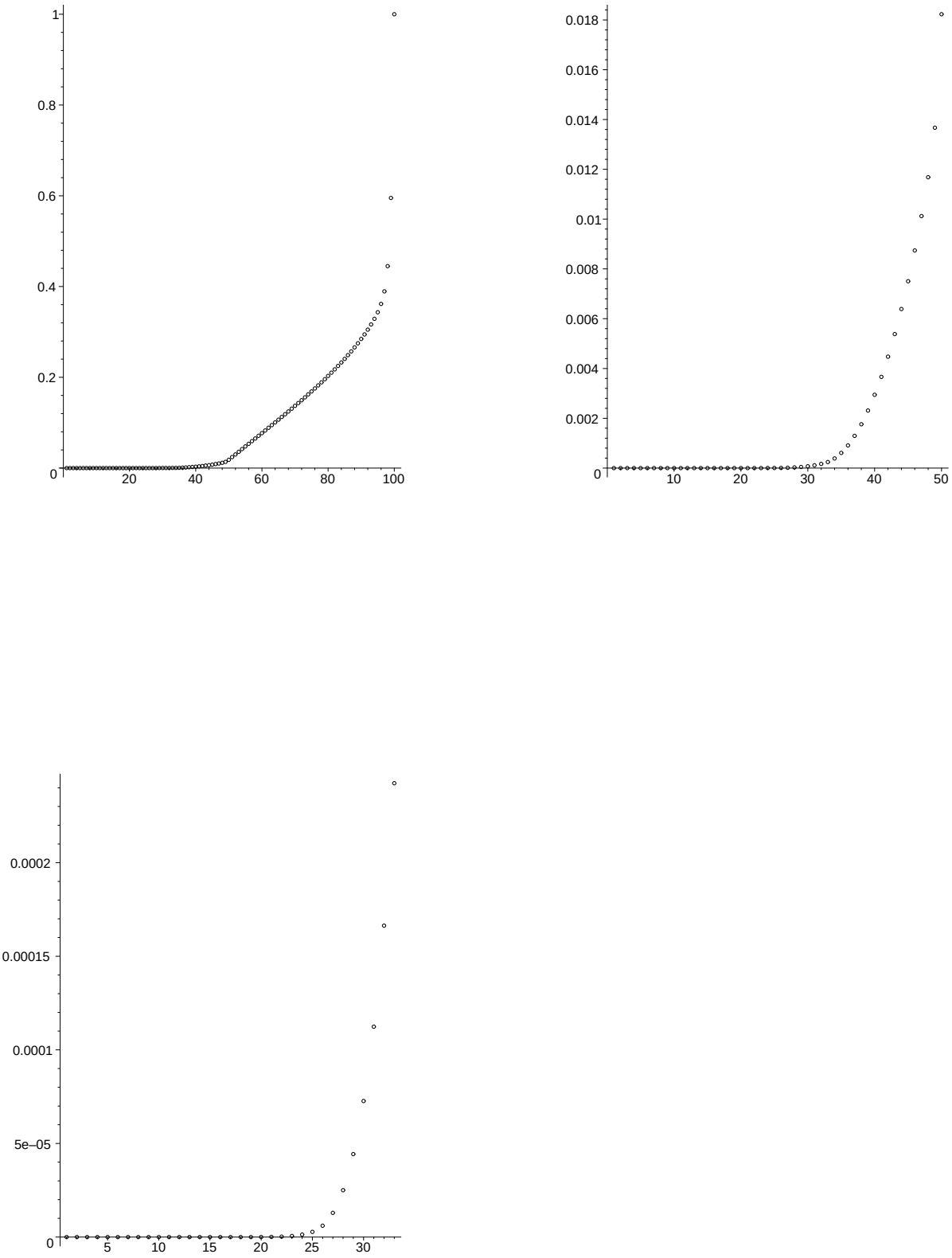


Figure 3.9.: The distribution function of L_n for $n = 100$. On the x-axis the different values of n are shown. On the y-axis $\mathbb{P}(L_n \geq n)$ is plotted. We observe three different phases. The first plot shows the change at $n/2$, the second the change at $n/3$, the third the change at $n/4$.

Proof. A decomposable class of graphs $\mathcal{A}$ is called an *assembly* in [5] (see also [5, Table 2.2] for an illustration). A decomposable class of graphs is a logarithmic assembly [5, p.51] if it satisfies

$$\frac{C_i}{i!} \sim \frac{\theta y^i}{i}$$

as $i \to \infty$ for some $y, \theta > 0$ (see [5, Eq.(2.16)]). Observe that in particular permutations are a logarithmic assembly for if C_i counts the number of permutations of size i $\frac{C_i}{i!} = \frac{1}{n}$. In fact, the study of logarithmic assemblies started with the study of permutations. Clearly, the class of graphs avoiding the bow tie and the diamond is decomposable and hence an assembly. By Proposition 3.3.1 the class avoiding the bow tie and the diamond is also logarithmic as we can take $y = e$ and $\theta = 1/4$. Our random variable L_n coincides with the random variable $L_1^{(n)}$ of [5]. Theorem 6.12 of [5] tells us that for general logarithmic assemblies L_n has a local and a global limit law and Theorem 6.8 tells us that the local limit law is the first component of a Poisson-Dirichlet distribution of parameter θ. The distribution function of the limit law for the largest component (the first component of the Poisson-Dirichlet distribution) is given by [5, Eq. (5.29)], and the differential equation satisfied by $r_{1/4}(x)$ follows from [5, Eq. (4.23)]. □

3.4. Forests of paths, caterpillars and combs

Forests of paths were the first mentioned minor-closed class of graphs where the excluded minors are not 2-connected, see [71] and Figure 3.10 for its excluded minors. In the same paper it was already observed that in this case the expected number of components of $R_n \in_u \mathcal{A}$ is of order $\sqrt{n}$ so the behaviour is very different from the 2-connected case. A similar class of graphs is the class of forests of

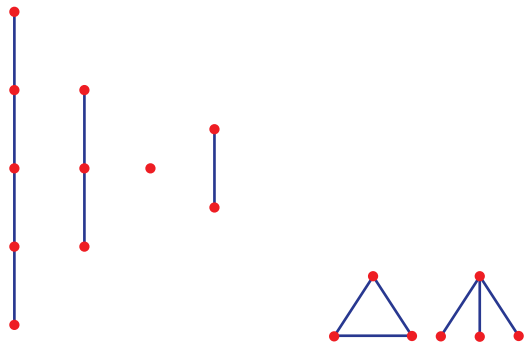


Figure 3.10.: An unlabelled path forest on 11 vertices (left) and excluded minors of forests of paths (right).

caterpillars. A *caterpillar* is a tree made of a simple path called the *spine* together with leaves adjacent to the spine Figure 3.11. Forests of caterpillars are precisely the graphs of pathwidth equal to 1. See [85] for a definition of pathwidth. A *comb* is a tree made of a simple path called the *spine* together with at most one additional leaf adjacent to the spine. Forests of combs are not minor-closed. Nevertheless, they are of (some) general interested and so-called combs inequalities (with combs resembling the ones just defined) have been studied by Chvátal and other for the Travelling Salesman Problem. Expectedly, they behave very similarly to forests of caterpillars and forests of paths. For the growth constants of forests of paths and caterpillars as well as related minor-closed graph classes see [20]. One would expect the two classes to behave quite similarly asymptotically and the following Theorem affirms this intuition.

Theorem 3.4.1 (Number of graphs and components — forests of paths, caterpillars and combs). *Suppose $\mathcal{A}$ is one of the following classes: Forests of paths, forests of caterpillars, forests of combs. Then the following results hold: the number of graphs on n vertices is asymptotically*

$$C_n \sim n!\,\alpha\rho^{-n} \quad \text{and} \quad A_n \sim n!\,\frac{\alpha^{1/4}e^{\alpha/2+\beta}}{2\sqrt{\pi}n^{3/4}}\rho^{-n}e^{2\sqrt{\alpha n}}. \tag{3.12}$$

The mean and variance of $\kappa(R_n)$ satisfy:

$$\mathbb{E}(\kappa(R_n)) \sim \sqrt{\alpha n}, \qquad \mathbb{V}(\kappa(R_n)) \sim \sqrt{\alpha n}/2,$$

and the random variable $\frac{\kappa(R_n)-\sqrt{\alpha n}}{(\alpha n/4)^{1/4}}$ converges in law to a standard normal distribution. The normalized variable $S_n/\sqrt{n/\alpha}$ converges in distribution to a Gamma$(2,1)$ *law, of density xe^{-x} on $[0,\infty)$.*

A local limit law also holds: for $x > 0$ and $k = \lfloor x\sqrt{n/\alpha}\rfloor$,

$$\sqrt{n/\alpha}\ \mathbb{P}(S_n = k) \sim xe^{-x}.$$

The constants are as follows:
Forests of paths: $\rho = 1$, $\alpha = \frac{1}{2}$, $\beta = 0$.

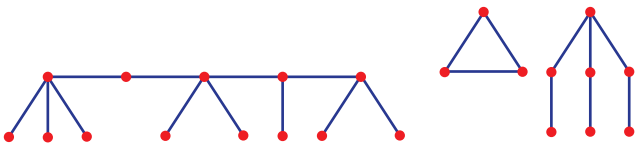


Figure 3.11.: A caterpillar (left) and excluded minors of forests of caterpillars (right).

Forests of caterpillars: $\rho \approx 0.567$ *and* ρ *is the unique solution to* $\rho e^{\rho} = 1$, $\alpha = \frac{(1-\rho)^2}{2(1+\rho)} \approx 0.06$, $\beta = \frac{\rho\left(10+3\,\rho-4\,\rho^2-\rho^3\right)}{4(1+\rho)^2} \approx 0.59$
Forests of combs: $\rho = \frac{\sqrt{5}-1}{2} \approx 0.61$, $\alpha = \frac{-1}{10}\frac{\left(\sqrt{5}-3\right)\sqrt{5}}{\sqrt{5}-1}$, $\beta = \frac{\sqrt{5}}{20}\sqrt{5} - \frac{7}{20}$.

Proof. We are going to show that in all three cases the generating function $C(z)$ counting the connected structures is of type mentioned in (3.13) so it has a simple pole and is of the form

$$C(z) = \frac{\alpha}{1 - z/\rho} + D(z) \tag{3.13}$$

where $D(\rho) = \beta$, $\alpha > 0$ and D has radius of convergence strictly larger than ρ. Hence Propositions 2.6.1, 2.6.2 and 2.6.3 give the results. □

Lemma 3.4.2. *The generating function of forests of (labelled) paths is*

$$A(z) = \exp\left(\frac{1}{2}\frac{z}{1-z} + \frac{1}{2}z\right).$$

Proof. For $n \geq 2$ there are $\frac{n!}{2}$ different ways to label a path of length n. Therefore $A_n = \frac{n!}{2}$ for $n \geq 2$ and $A_1 = 1$. Hence the EGF of all paths is

$$A(z) = \sum_{n=0}^{\infty} A_n \frac{z^n}{n!} = \sum_{n=2}^{\infty} \frac{n!}{2}\frac{z^n}{n!} + z = \frac{1}{2}\frac{z}{1-z} + \frac{1}{2}z = \frac{1}{2}\frac{1}{1-z} - \frac{1}{2} + \frac{1}{2}z \tag{3.14}$$

As we are interested in the forests of paths we have to apply the SET construction to all paths, so we get $A(z) = e^{A(z)}$, or rather $A(z,u) = e^{uA(z)}$, where the second variable u marks the number of components. □

For forests of paths, we have also obtained the limit law of the size L_n of the largest component. It is significantly larger than the root component ($\sqrt{n}\log n$ instead of $\sqrt{n}$).

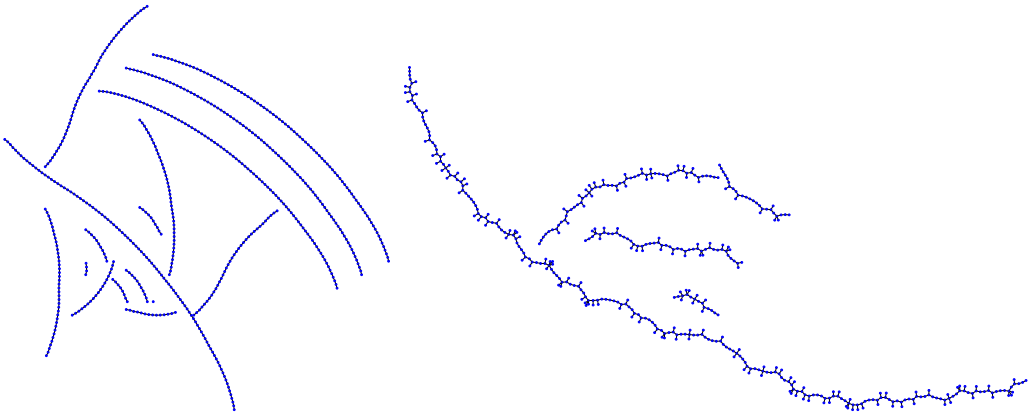


Figure 3.12.: A random forest of paths of size $n = 636$ (left) and a random forest of caterpillars of size $n = 486$ (right).

Proposition 3.4.3 (Size of the largest component — forests of paths). *In forests of paths, the (normalized) size of the largest component converges in law to a Gumbel distribution: for $x \in \mathbb{R}$ and as $n \to \infty$,*

$$\mathbb{P}\left(\frac{L_n - \sqrt{n/2}\log n}{\sqrt{n/2}} < x\right) \to \exp\left(-\frac{e^{-x/2}}{\sqrt{2}}\right).$$

The appearances of a Gumbel distribution might not be too surprising in this context as the following argument illustrates. Coming from the Boltzmann model with parameter x we know that a path of length k is generated according to a geometric distribution with parameter k [46]. This is a simple consequence of the fact that the generating function of paths is basically a sequence (plus minor modifications). In the limit the largest component, i.e. the largest path, in a forests of paths is then expected to be distributed according to a Gumbel distribution as we argued in Chapter 2.5.

Proof. We start from (2.7), where

$$k = \sqrt{n/2}\,(\log n + x) \tag{3.15}$$

and the generating function of paths of size less than k is:

$$C^{[k]}(z) = \sum_{n=1}^{k-1} C_n \frac{z^n}{n!} = \frac{z}{2} + \frac{z - z^k}{2(1-z)}. \tag{3.16}$$

Using a saddle point approach for integrals as described in Chapter 2.4, we will find an estimate of

$$[z^n]\exp(C^{[k]}(z)) = \frac{1}{2i\pi}\int_{\mathcal{C}_r} \exp(C^{[k]}(z))\frac{dz}{z^{n+1}}, \tag{3.17}$$

where the integration contour is any circle $\mathcal{C}_r$ of center 0 and radius $r < 1$.

Let us first introduce some notation: Define $K(z) = C^{[k]}(z)$ and we denote the integrand in (3.17) by F, and its logarithm by f:

$$F(z) = \frac{\exp(K(z))}{z^{n+1}}, \qquad f(z) = K(z) - (n+1)\log z.$$

We choose the radius $r \equiv r_n$ that satisfies the saddle point equation

$$F'(r) = 0, \quad \text{or equivalently} \quad f'(r) = 0 \quad \text{or} \quad rK'(r) = n+1.$$

Note that $rK'(r) = \sum_{n=1}^{k-1} \frac{nr^n}{2} + \frac{r}{2}$ is strictly increasing in $[0, \infty]$ as r grows from 0 to 1 and therefore the solution of the saddle point equation is unique.

By bootstrapping see [49, p.309] and [49, p.562] for a very similar example, we find:

$$r = 1 - \frac{1}{\sqrt{2n}} + \frac{e^{-x/2}}{4\sqrt{2}} \frac{\log n}{n} + O\Big(\frac{1}{n}\Big). \tag{3.18}$$

Gaussian approximation. Let $\theta_0(r) \in (0, \pi)$. We want to find a function θ_0 such that for fixed r and $|\theta| < \theta_0$

$$F(re^{i\theta}) \sim F(r)e^{-\theta^2 r^2 f''(r)/2}$$

holds. Let r be fixed. By taking a taylor expansion with a Lagrange remainder of the function $g : \theta \mapsto f(re^{i\theta})$ in the neighbourhood of $\theta = 0$, we find, for $|\theta| \leq \theta_0$:

$$\Big| f(re^{i\theta}) - f(r) + \theta^2 r^2 f''(r)/2 \Big| \leq \frac{\theta_0^3}{6} \sup_{|\alpha| \leq \theta_0} \Big| g^{(3)}(\alpha) \Big|, \tag{3.19}$$

with

$$\begin{aligned} \Big| g^{(3)}(\alpha) \Big| &= \Big| -ire^{i\alpha} f'(re^{i\alpha}) - 3ir^2 e^{2i\alpha} f''(re^{i\alpha}) - ir^3 e^{3i\alpha} f'''(re^{i\alpha}) \Big| \\ &\leq K'(r) + \frac{n+1}{r} + 3K''(r) + 3\frac{n+1}{r^2} + K'''(r) + 2\frac{n+1}{r^3}. \end{aligned}$$

By combining the expression (3.16) of $K(z) = C^{[k]}(z)$ and the saddle point estimate (3.18), we find that $K'(r) = (n+1)/r \sim n$, that $K''(r) \sim 2\sqrt{2}n^{3/2}$ and finally that $K'''(r) \sim 12n^2$. This last term dominates the above bound on $|g^{(3)}(\alpha)|$. Hence, if

$$\theta_0 \equiv \theta_0(n) = o(n^{-2/3}), \tag{3.20}$$

we find that for some constant $M \in \mathbb{R}$

$$\frac{\theta_0^3}{6} \sup_{|\alpha| \leq \theta_0} \Big| g^{(3)}(\alpha) \Big| \leq \frac{\theta_0^3}{6} \cdot M \cdot 12n^2 \to 0$$

Then by taking the exponential of (3.19),

$$F(re^{i\theta}) \sim F(r)e^{-\theta^2 r^2 f''(r)/2},$$

uniformly in $|\theta| \leq \theta_0$.

Completion of the Gaussian integral. We split the integral (3.17) into two parts, depending on whether $|\theta| \leq \theta_0$ or $|\theta| \geq \theta_0$. The first part is

$$\int_{-\theta_0}^{\theta_0} F(re^{i\theta}) \frac{re^{i\theta} d\theta}{2\pi} \sim \frac{rF(r)}{2\pi} \int_{-\theta_0}^{\theta_0} e^{-\theta^2 r^2 f''(r)/2} e^{i\theta} d\theta.$$

As argued above, $r^2 f''(r) \sim K''(r) \sim 2\sqrt{2} n^{3/2}$. Hence, if we choose $\theta_0 \equiv \theta_0(n)$ such that $\theta_0^2 n^{3/2} \to \infty$ which is compatible with (3.20), for instance if

$$\theta_0 = n^{-5/6}, \tag{3.21}$$

we get $\theta_0 r \sqrt{f''(r)} \to \infty$.

Changing to $\alpha = \theta r \sqrt{f''(r)})$ we obtain

$$\begin{aligned}
\int_{-\theta_0}^{\theta_0} F(re^{i\theta}) \frac{re^{i\theta} d\theta}{2\pi} &\sim \frac{F(r)}{2\pi\sqrt{f''(r)}} \int_{-\theta_0 r\sqrt{f''(r)}}^{\theta_0 r\sqrt{f''(r)}} e^{-\alpha^2/2} e^{i\alpha/(r\sqrt{f''(r)})} d\alpha \\
&\sim \frac{F(r)}{2\pi\sqrt{f''(r)}} \int_{\mathbb{R}} e^{-\alpha^2/2} d\alpha \\
&\sim \frac{F(r)}{\sqrt{2\pi f''(r)}} \qquad (3.22) \\
&\sim \frac{e^{1/4} e^{\sqrt{2n}}}{2^{1/4} 2\sqrt{\pi} n^{3/4}} \exp\left(-\frac{e^{-x/2}}{\sqrt{2}}\right) \qquad (3.23)
\end{aligned}$$

by (3.18) and observing that with $t = \frac{\alpha^2}{2}$

$$\int_{\mathbb{R}} e^{-\alpha^2/2} d\alpha = 2 \int_0^\infty e^{-t} \frac{dt}{\sqrt{2t}} = \sqrt{2}\Gamma(1/2) = \sqrt{2\pi},$$

see for example [49, p.744] for basic properties of the Gamma function.
The second part of the integral can be neglected.

The second part of the integral (3.17) is

$$\int_{\theta_0 < |\theta| < \pi} F(re^{i\theta}) \frac{re^{i\theta} d\theta}{2\pi},$$

and we want to prove that it is dominated by (3.23), i.e. that

$$\int_{\theta_0 < |\theta| < \pi} F(re^{i\theta}) \frac{re^{i\theta} d\theta}{2\pi} \cdot \left(\int_{-\theta_0}^{\theta_0} F(re^{i\theta}) \frac{re^{i\theta} d\theta}{2\pi} \right)^{-1} \to 0$$

as $n \to \infty$. By (3.22) this is equivalent to showing that

$$\int_{\theta_0<|\theta|<\pi} F(re^{i\theta})\frac{re^{i\theta}d\theta}{2\pi} \cdot \frac{\sqrt{f''(r)}}{F(r)} \to 0.$$

Hence it suffices to prove that for $\theta_0 < |\theta| < \pi$, we have

$$|F(re^{i\theta})| = o\left(\frac{F(r)}{\sqrt{f''(r)}}\right). \tag{3.24}$$

Let us denote $z = re^{i\theta}$ and $z_0 = re^{i\theta_0}$ and remember that $|e^{re^{i\theta}}| = e^{\mathrm{Re}(re^{i\theta})}$. We have

$$\begin{aligned}
\frac{|F(re^{i\theta})|}{F(r)} &= |\exp(K(z) - K(r))| \\
&\leq \exp(|K(z)| - K(r)) = \exp\left(\left|\frac{z}{2} + \frac{z - z^k}{2(1-z)}\right| - \frac{r}{2} - \frac{r - r^k}{2(1-r)}\right) \\
&\leq \exp\left(\frac{r + r^k}{2|1-z|} - \frac{r - r^k}{2(1-r)}\right) \\
&\leq \exp\left(\frac{r + r^k}{2|1-z_0|} - \frac{r - r^k}{2(1-r)}\right) = \exp\left(-\frac{n^{1/14}}{\sqrt{2}}(1 + o(1))\right),
\end{aligned}$$

given the values (3.15), (3.18) and (3.21) of k, r and θ_0. Since $f''(r) \sim 2\sqrt{2}n^{3/2}$, we conclude that (3.24) holds.

Conclusion. We have now established that the integral (3.17) is dominated by its first part, and is thus equivalent to (3.23). To obtain the limiting distribution function, it remains to divide this estimate by $A_n/n!$. The asymptotic behaviour of A_n is given by (2.33), with $\alpha = 1/2$, $\rho = 1$ and $\beta = 0$, and this concludes the proof.

□

Lemma 3.4.4. *The exponential generating function of forests of caterpillars is*

$$A(z) = \exp\left(\frac{1}{2}z^2 \cdot (e^z - 1)^2\frac{1}{1 - ze^z} + ze^z - \frac{z^2}{2}\right) \tag{3.25}$$

and

$$\frac{1}{2}z^2 \cdot (e^z - 1)^2\frac{1}{1 - ze^z} + ze^z - \frac{z^2}{2} \sim \frac{\alpha}{1 - z/\rho} + \beta + O\left(1 - z/\rho\right)$$

with $\rho \approx 0.567$ *where* ρ *is the unique solution to* $\rho e^\rho = 1$, $\alpha = \frac{(1-\rho)^2}{2(1+\rho)} \approx 0.06$ *and* $\beta = \frac{\rho\left(10+3\,\rho-4\,\rho^2-\rho^3\right)}{4(1+\rho)^2}$.

Proof. As in the case of forests of paths, we start with the exponential generating function for connected caterpillars. We can think of a caterpillar as a sequence of stars, where the centres of the stars are connected by edges. Then the only problem is to avoid double counting, so to make sure that each caterpillar has a unique spine. We first of all assume that the caterpillar has at least 4 vertices and that it contains a path of length 3. Without loss of generality we assume that the star at the end and the beginning of the spine has at least one leg. Otherwise we can simply think of the last vertex on the spine as a leg of the previous star. A star is simply a set of vertices, with one specified vertex - the centre - adjacent to all the others, therefore

$$\text{Stars} = \mathcal{Z} \times SET_{\geq 2}(\mathcal{Z}) + \mathcal{Z} + \frac{\mathcal{Z} \times \mathcal{Z}}{2}$$

in order to avoid to have 2 stars on 2 vertices.
In order to get stars with at least one leg, we have to exclude the empty set from the SET-construction, so rooted stars have the specification

$$\text{rStars}_{\geq 1 \text{ leg}} = \mathcal{Z} \times SET_{\geq 1}(\mathcal{Z}).$$

Thus the specification of caterpillars with at least 4 vertices containing a path of length 3 is

$$\frac{1}{2}\left(\text{rStars}_{\geq 1 \text{ leg}} \times SEQ(\text{Stars}) \times \text{Stars}_{\geq 1 \text{ leg}}\right).$$

The only caterpillars that don't have 4 vertices or that have at least 4 vertices but don't contain a path of length at least 3 are precisely the stars. So the class of connected caterpillars admits the following specification

$$\frac{1}{2}\left(\text{rStars}_{\geq 1 \text{ leg}} \times SEQ(\text{Stars}) \times \text{Stars}_{\geq 1 \text{ leg}}\right) + \text{Stars}.$$

Applying the symbolic method the specification translates into the following EGF

$$C(z) = \frac{1}{2} z^2 \left(e^z - 1\right)^2 \frac{1}{1 - z e^z} + z e^z - \frac{z^2}{2}.$$

As usual forests of caterpillars $\mathcal{A} = \text{SET}(\mathcal{C})$ where $\mathcal{C}$ denotes the class of connected caterpillars, thus

$$A(z) = e^{\frac{1}{2} z^2 (e^z - 1)^2 \frac{1}{1 - z e^z} + z e^z - \frac{z^2}{2}}.$$

The series $C(z)$ is meromorphic on $\mathbb{C}$, with a unique dominant pole at ρ. We need to show that its behaviour around this point is of type

$$C(z) = \frac{\alpha}{1 - z/\rho} + \beta + O\left(1 - z/\rho\right)$$

for some $\alpha,\beta \in \mathbb{R}$. First we rewrite

$$C(z) = \frac{1}{2}\frac{z\mathrm{e}^z\left(z\mathrm{e}^z + 2\,z - 2 - z^2.\right)}{-1 + z\mathrm{e}^z} \tag{3.26}$$

Then the idea is to desingularize (3.26) by taking a Taylor expansion of the numerator and denominator around ρ. For the numerator this gives

$$-1 - \rho^2 + 2\,\rho + \left(4 - \rho - \rho^2\right)(z-\rho) + O\left((z-\rho)^2\right)$$

and for the denumerator

$$2\,\frac{\rho+1}{\rho}(z-\rho) + O\left((z-\rho)^2\right).$$

Dividing the two expansions and expanding the result in powers of $(1 - z/\rho)$ we obtain

$$\frac{1}{2}\,\frac{1+\rho^2-2\,\rho}{\rho+1}\cdot\left(1-\frac{z}{\rho}\right) + \frac{\rho\,(10+3\,\rho-4\,\rho^2-\rho^3)}{4\,(1+\rho)^2} + O\left(\left(1-\frac{z}{\rho}\right)^{-1}\right).$$

□

Recall that a *comb* is a tree made of a simple path called the *spine* together with at most one additional leaf adjacent to the spine.

Lemma 3.4.5. *The EGF of forests of combs is*

$$A(z) = \exp\left(\frac{z^2}{2(1-z-z^2)}\right) \tag{3.27}$$

and

$$C(z) = \frac{\alpha}{1-z/\rho} + \beta + O\left((1-z/\rho)\right)$$

with $\rho = \frac{\sqrt{5}-1}{2}$, $\alpha = \frac{-1}{2}\,\frac{-1+\rho}{\rho\,(1+2\,\rho)}$, $\beta = \frac{-1}{2}\,\frac{3\,\rho+4\,\rho^2-1}{(1+2\,\rho)^2}$.

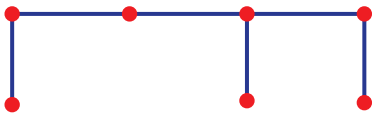


Figure 3.13.: A comb

Proof. A comb is an unoriented sequence of vertices, where some vertices have been replaced by single edges. To avoid double counting, we think of it as rooted on both ends and obtain

$$\frac{1}{2}\,\mathcal{Z} \times \frac{1}{1-(\mathcal{Z}+\mathcal{Z}^2)} \times \mathcal{Z}$$

which translates into

$$C(z) = \frac{z^2}{2(1-z-z^2)}. \tag{3.28}$$

The radius of convergence ρ of $C(z)$ is the positive solution to $z + z^2/2 = 1$ which is $\rho = \sqrt{3} - 1$. Take an expansion of $z^2/2$ and an expansion of $1 - (z + z^2/2)$ around ρ and divide the two. Expanding the result in powers of $(1 - z/\rho)$ gives the result. □

3.5. Forests of graphs of bounded component size

Let $\mathcal{C}$ be a finite class of connected graphs, and let $\mathcal{A}$ be the class of graphs with connected components in $\mathcal{C}$ (or where all connected components are in $\mathcal{C}$).

Note that $\mathcal{A}$ is minor-closed if and only if $\mathcal{C}$ itself is minor-closed. This is the case for instance if $\mathcal{C}$ is the class of graphs of size at most k. In general, we denote by k the size of the largest graphs of $\mathcal{C}$.

Fix $k \geq 1$. Let $\mathcal{A}$ be the class of graphs in which all components have size at most k. This class is minor-closed and the excluded minors are all trees of size $k+1$.

The following proposition is more precise than the standard result on exponentials of polynomials [49, Cor. VIII.2, p. 568], since it makes the behaviour of $b(\zeta_n)$ occurring in the saddle point estimate (3.31) explicit.

Proposition 3.5.1 (The number of graphs with small components). *Write the generating function of graphs of $\mathcal{C}$ as*

$$C(z) = \sum_{i=0}^{k} \frac{C_i}{i!} z^i. \tag{3.29}$$

where C_i denotes the number of connected graphs on i vertices. The generating function of graphs of $\mathcal{A}$ is $A(z) = e^{C(z)}$. As $n \to \infty$,

$$A_n \sim n!\,\frac{1}{\sqrt{2\pi k n}}\,\frac{A(\zeta)}{\zeta^n} \tag{3.30}$$

where $\zeta_n \equiv \zeta_n$ *is defined by* $\zeta C'(\zeta) = n$ *and satisfies*

$$\zeta = \alpha n^{1/k} + \beta + O(n^{-1/k}) \tag{3.31}$$

with

$$\alpha = \left(\frac{(k-1)!}{C_k}\right)^{1/k} \quad \text{and} \quad \beta = -\frac{(k-1)C_{k-1}}{kC_k}. \tag{3.32}$$

The probability that $R_n \in_u \mathcal{A}$ *is connected is of course equal to* 0 *for* $n > k$.

Proof. The series $A(z)$ is H-admissible ([49, Thm. VIII.5, p. 568]) and Theorem 2.4.1 applies. We need to find an asymptotic expansion of ζ_n. The saddle point equation $\zeta C'(\zeta) = n$ is an irreducible (bivariate) polynomial in ζ and n, because it has degree 1 in n. Observe that it has degree k in ζ. Consider $1/n$ as a small parameter x. By [88, Prop. 6.1.6], the saddle point ζ admits an expansion of the form

$$\zeta = \sum_{i \geq i_0} \alpha_i n^{-i/k}, \tag{3.33}$$

for some integer i_0 and complex coefficients α_i. Using Newton's polygon method [49, p. 499], one finds $i_0 = -1$ and the values (3.32) of the first two coefficients.

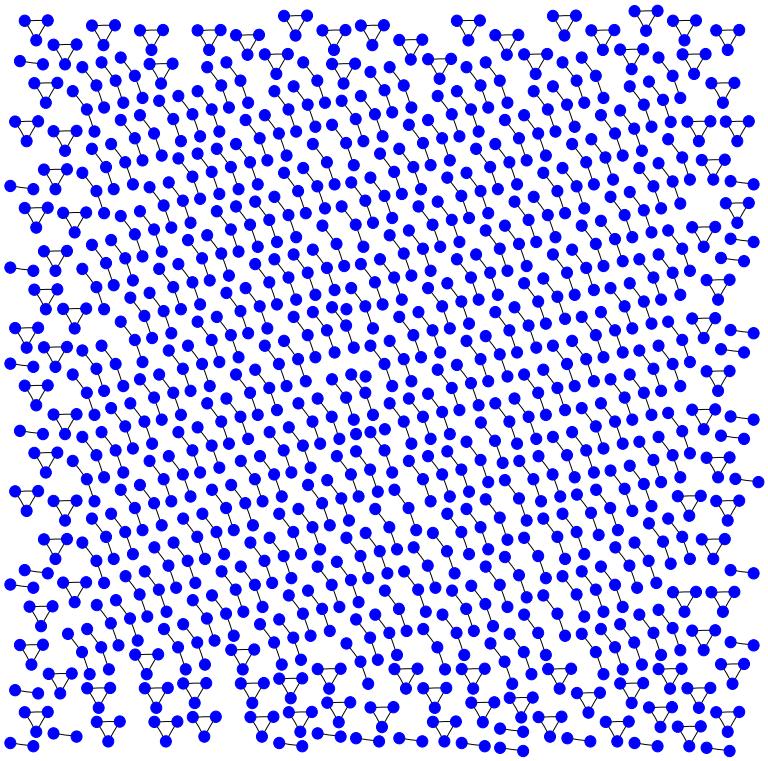


Figure 3.14.: A random graph of size $n = 1171$ with component size at most 3. Observe that most components have size 3, so that the root component is very likely to have size 3.

Since $b(r) = rC'(r) + r^2C''(r)$ has leading term $kC_kr^k/(k-1)!$, the first order expansion of $b(\zeta)$ reads

$$b(\zeta) = kn + O(n^{(k-1)/k}),$$

and the asymptotic behaviour of A_n follows. □

The following statement can be found in [28, Thm. I], but the version presented here is more precise because our estimates of $\mathbb{E}(\kappa(R_n))$ and $\mathbb{V}(\kappa(R_n))$ are explicit. In particular $\mathbb{E}(\kappa(R_n)) \sim n/k$ suggests that most components have maximal size k.

Proposition 3.5.2 (**Number of components — Graphs with small components**). *Assume that the coefficient C_{k-1} in* (3.29) *is non-zero. Let $\kappa(R_n)$ be the number of components of $R_n \in_u \mathcal{A}$. As $n \to \infty$*

$$\kappa(R_n) \sim \mathcal{N}\left(\mathbb{E}(R_n), \mathbb{V}(R_n)\right)$$

and the mean and variance of $\kappa(R_n)$ satisfy:

$$\mathbb{E}(\kappa(R_n)) \sim \frac{n}{k}, \qquad \mathbb{V}(\kappa(R_n)) \sim \frac{C_{k-1}}{k \cdot k!}\alpha^{k-1}n^{(k-1)/k},$$

where α is given by (3.32).

Proof. We apply [28, Thm. I] which gives the convergence in distribution. We can also apply Theorem 2.4.5 if $k > 3$. For $k < 3$ the function $A(z)$ is not extended Hayman-admissible as condition 7 of the definition fails (ϵ gets too large). As already observed in [44] for functions of the form $\exp(P(z,u)$ where $P(z,u)$ is a polynomial the concept of extended Hayman-admissibility is in some sense too strict, i.e. there are cases, for instance the cases $k = 2, 3$ in this example, where a normal limit law occurs although the generating function is not extended Hayman-admissible.

Still denoting the saddle point by $\zeta \equiv \zeta_n$, we have to find (asymptotic) estimates of

$$\mu_n = C(\zeta) \quad \text{and} \quad \sigma_n^2 = C(\zeta) - \frac{(\zeta C'(\zeta))^2}{\zeta C'(\zeta) + \zeta^2 C''(\zeta)}.$$

Given (3.31), we obtain

$$\mu_n = \frac{n}{k} + \frac{C_{k-1}}{k!}\alpha^{k-1}n^{(k-1)/k} + O(n^{(k-2)/k}),$$

$$\zeta^2 C''(\zeta) = (k-1)n - \frac{C_{k-1}}{(k-2)!}\alpha^{k-1}n^{(k-1)/k} + O(n^{(k-2)/k}),$$

and finally

$$\sigma_n^2 = \mu_n - \frac{n^2}{n + \zeta^2 C''(\zeta)} = \frac{C_{k-1}}{k \cdot k!} \alpha^{k-1} n^{(k-1)/k} + O(n^{(k-2)/k}).$$

□

Since there are approximately n/k components, one expects the size S_n of the root component to be k. This is indeed the case as the following proposition demonstrates. See also Figure 3.14 for an illustration.

Proposition 3.5.3 (**Size of the components — Graphs with small components).** *The distribution of S_n converges to a Dirac law at k:*

$$\mathbb{P}(S_n = j) \to \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{otherwise.} \end{cases}$$

The same holds for the size L_n of the largest component.

Proof. We combine the second formulation in (2.5) with the estimate (3.30) of A_n. This gives

$$\mathbb{P}(S_n = j) \sim \frac{1}{n} \frac{C_j}{(j-1)!} \frac{A(\zeta_{n-j})}{A(\zeta_n)} \frac{\zeta_n^n}{\zeta_{n-j}^{n-j}}.$$

It suffices to prove that this probability tends to 0 for every j with $j < k$. So let us assume $j < k$. Since ζ_n is increasing with n we have $\frac{A(r_{n-j})}{A(r_n)} < 1$. Therefore it suffices to prove that

$$\frac{1}{n} \frac{\zeta_n^n}{\zeta_{n-j}^{n-j}} \to 0. \tag{3.34}$$

Recall from (3.31) and (3.33) that ζ_n admits an expansion of the form

$$\zeta_n = \alpha n^{1/k} + \sum_{i=0}^{k-1} n^{-i/k} \beta_i + O\left(1/n\right).$$

This gives, for some constants γ_i,

$$n \log \zeta_n = \frac{n}{k} \log n + n \log \alpha + \sum_{i=1}^{k} \gamma_i n^{1-i/k} + O(n^{-1/k}).$$

Hence

$$(n-j) \log \zeta_{n-j} = \frac{n-j}{k} \log n + (n-j) \log \alpha + \sum_{i=1}^{k} \gamma_i n^{1-i/k} + O(1).$$

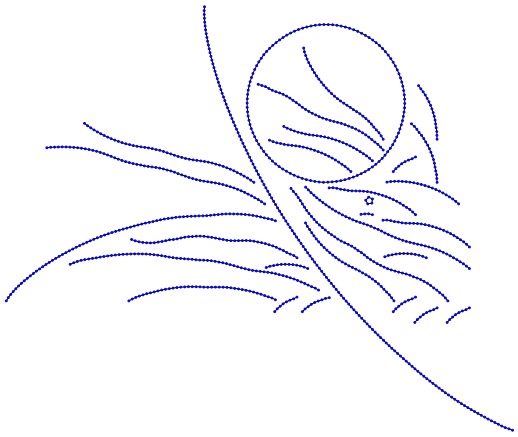


Figure 3.15.: A random graph of size $n = 1034$ with no 3-star minor.

This gives

$$n \log \zeta_n - (n-j) \log \zeta_{n-j} - \log n = \frac{j-k}{k} \log n + O(1),$$

and (3.34) follows, since $j < k$. Since $L_n \geq S_n$, the behaviour of L_n is then clear. □

3.6. Graphs with maximal degree 2

Let $\mathcal{A}$ be the class of graphs of maximal degree 2, or equivalently, $\mathcal{A} = \text{Ex}(3\text{-star})$. That is the minor-closed class of graphs defined by excluding the star on 4 vertices as a minor. The class $\text{Ex}((k+1)3\text{-star})$ for $k \in \mathbb{N}$, i.e. the class excluding $k+1$ disjoint 3-stars, was enumerated in [72]. In $\mathcal{A}$ every vertex has degree at most 2 otherwise it would contain a 3-star as a minor. The connected components of such graphs are therefore paths or cycles. The generating function $C(z)$ of this class has a simple pole corresponding to forests of paths examined in section 3.4 and a logarithmic singularity at its radius of convergence ρ which reflects the cycles that are included. As we shall see, the logarithm changes the asymptotic behaviour of the numbers A_n, but the other results remain unaffected. This reflects the fact that we expect that most of the vertices in $R_n \in \mathcal{A}$ are in forests of paths and not in cycles.

Proposition 3.6.1 (The number of graphs of maximal degree 2). *Let $\mathcal{A} = Ex(3\text{-}star)$. Then*

$$C_n = n!/2 + (n-1)!/2 \quad \textit{for } n \geq 3 \textit{ with } C_1 = C_2 = 1.$$

As $n \to \infty$,

$$A_n \sim n! \, \frac{1}{2\sqrt{e\pi} n^{1/2}} e^{\sqrt{2n}}.$$

In particular, the probability that R_n is connected tends to 0 at speed $n^{1/2} e^{-\sqrt{2n}}$ as $n \to \infty$.

Proof. Recall that by (3.5) the generating function of cycles is

$$\mathrm{Cyc}(z) = \frac{1}{2} \log \frac{1}{1-z} - \frac{z}{2} - \frac{z^2}{4}.$$

The generating function of connected graphs (paths or cycles) of size n in the class $\mathcal{A}$ is

$$C(z) = \frac{z}{2(1-z)} - \frac{1}{2} + \frac{z}{2} + \frac{1}{2} \log \frac{1}{1-z} - \frac{z}{2} - \frac{z^2}{4} = \frac{z(2-z+z^2)}{4(1-z)} + \frac{1}{2} \log \frac{1}{1-z}. \tag{3.35}$$

Hence the generating function of graphs of $\mathcal{A}$ is therefore

$$A(z) = e^{C(z)} = \frac{1}{\sqrt{1-z}} \exp\left(\frac{z(2-z+z^2)}{4(1-z)} \right).$$

To obtain the asymptotic behaviour of A_n, we write

$$A(z) = F(z)G(z) \quad \text{with} \tag{3.36}$$

$$F(z) = \frac{1}{\sqrt{1-z}} \exp\left(\frac{1}{2(1-z)} \right) \quad \text{and} \quad G(z) = \exp\left(-\frac{1}{2} - \frac{z^2}{4} \right). \tag{3.37}$$

By Schur's Lemma 2.2.1 we have

$$[z^n]A(z) \sim [z^n]F(z) \cdot e^{-3/4}.$$

Applying Macintyre's and Wilson's estimate (2.35) with $\gamma = \frac{1}{2}$, $\alpha = \frac{1}{2}$ then gives the result. □

For the number of components, we find the same behaviour as in the case of forests of paths (Theorem 3.4.1 with $\alpha = 1/2$). We have also determined the expected number of cyclic components.

Proposition 3.6.2 (Number and nature of components — Graphs of maximal degree 2). *The mean and variance of N_n satisfy:*

$$\mathbb{E}(\kappa(R_n)) \sim \sqrt{n/2}, \qquad \mathbb{V}(\kappa(R_n)) \sim \sqrt{n/8},$$

and the random variable $\frac{\kappa(R_n)-\sqrt{n/2}}{(n/8)^{1/4}}$ *converges in law to a standard normal distribution. The expected number of cycles in* $R_n \in_u \mathcal{A}$ *is asymptotically equal to* $(\log n)/4$. *The size of the root component is still described by Proposition 2.6.3, with* $\alpha = 1/2$.

Proof. We want to apply Theorem 2.4.5. To prove that $A(z)$ is Hayman-admissible, we apply Theorem 2.4.3 to (3.36). This reduces our task to proving that $F(z)$ is H-admissible which follows from [60, Theorem XII, p. 93] (see also [45, Lemma 1]). Conditions (2.26)–(2.28) are checked with Maple.

The asymptotic estimates of $\mathbb{E}(\kappa(R_n))$ and $\mathbb{V}(\kappa(R_n))$ are obtained through (2.29), using the saddle point estimate $\zeta_n = 1 - 1/\sqrt{2n} + O(1/n)$. Recall that by (3.5) the generating function of cycles is

$$\mathrm{Cyc}(z) = \frac{1}{2}\log\frac{1}{1-z} - \frac{z}{2} - \frac{z^2}{4}.$$

The bivariate generating function of graphs of $\mathcal{A}$, where z counts the number of vertices and v the number of cycles is

$$\tilde{A}(z,v) = \exp\left(z + \frac{z^2}{2(1-z)} + v\mathrm{Cyc}(z)\right).$$

We apply Lemma 2.1.1. By differentiating with respect to v, the expected number of cycles in $\mathcal{G}_n$ is found to be

$$\mathbb{E}(\text{ number of cycles in } R_n) \sim \frac{[z^n]\mathrm{Cyc}(z)A(z)}{[z^n]A(z)}.$$

The asymptotic behaviour of $[z^n]A(z) = A_n/n!$ has been established in Proposition 3.6.1. We are now determining an estimate of $[z^n]\mathrm{Cyc}(z)A(z)$. Using Schur's Lemma 2.2.1

$$[z^n]\mathrm{Cyc}(z)A(z) \sim [z^n]\frac{1}{2}\log\frac{1}{1-z}\cdot A(z) - [z^n]\frac{3}{4}A(z).$$

Then (2.35) gives

$$[z^n]\mathrm{Cyc}(z)A(z) \sim \frac{\log n}{8\sqrt{e\pi}n^{1/2}}e^{\sqrt{2n}}$$

and the result follows on the expected number of cycles follows. The proof for the root component is exactly the same as for the proof of Proposition 2.6.3 with now (for $i \geq 1$)

$$\begin{aligned} C^{(i+1)}(z) &= \left(\frac{d}{dz}\right)^{i+1} \frac{1}{2(1-z)} \\ &\quad + \left(\frac{d}{dz}\right)^{i+1} \log \frac{1}{1-z} - \frac{1}{2}\mathbf{1}_{i=1} \\ &= \frac{(i+1)!}{2(1-z)^{i+2}} + \frac{i!}{2(1-z)} - \frac{1}{2}\mathbf{1}_{i=1} \\ &= \frac{i!(2+i-z)}{(1-z)^{i+2}} - \frac{1}{2}\mathbf{1}_{i=1}. \end{aligned}$$

Then (2.35) gives

$$\frac{[z^{n-i-1}]C^{(i+1)}(z)A(z)}{n[z^n]A(z)} \sim (i+1)!(2n)^{i/2} + i!\frac{1}{2}\frac{(2n)^{i/2+1/2}}{n} \sim (i+1)!(2n)^{i/2}$$

and so the same result as in Proposition 2.6.3 holds. □

3.7. Excluding the bow tie

We now denote by $\mathcal{A} = \text{Ex(bow tie)}$ the class of graphs excluding the bow tie as a minor (Figure 3.4). The question of studying the class of Ex(bow tie) was first raised in [72].

Proposition 3.7.1 (The generating function of graphs – no bow tie). *Let $\mathcal{A} = Ex(bow\ tie)$ and $T \equiv T(z)$ be the generating function of rooted trees, defined by $T(z) = ze^{T(z)}$. The generating function of connected graphs in $\mathcal{A}$ is*

$$\begin{aligned} C(z) = \frac{T^2\,(1-T+T^2)\,\mathrm{e}^T}{1-T} + \frac{1}{2}\log\left(\frac{1}{1-T}\right) \\ + \frac{T\,(12-54\,T+18\,T^2-5\,T^3-T^4)}{24(1-T)}. \end{aligned} \tag{3.38}$$

The generating function of graphs of $\mathcal{A}$ is $A(z) = e^{C(z)}$.

The key point is the following characterization of cores (graphs of minimum degree 2) excluding a bow tie.

Proposition 3.7.2. *The cores that exclude the bow tie as a minor are:*

- *the empty graph,*

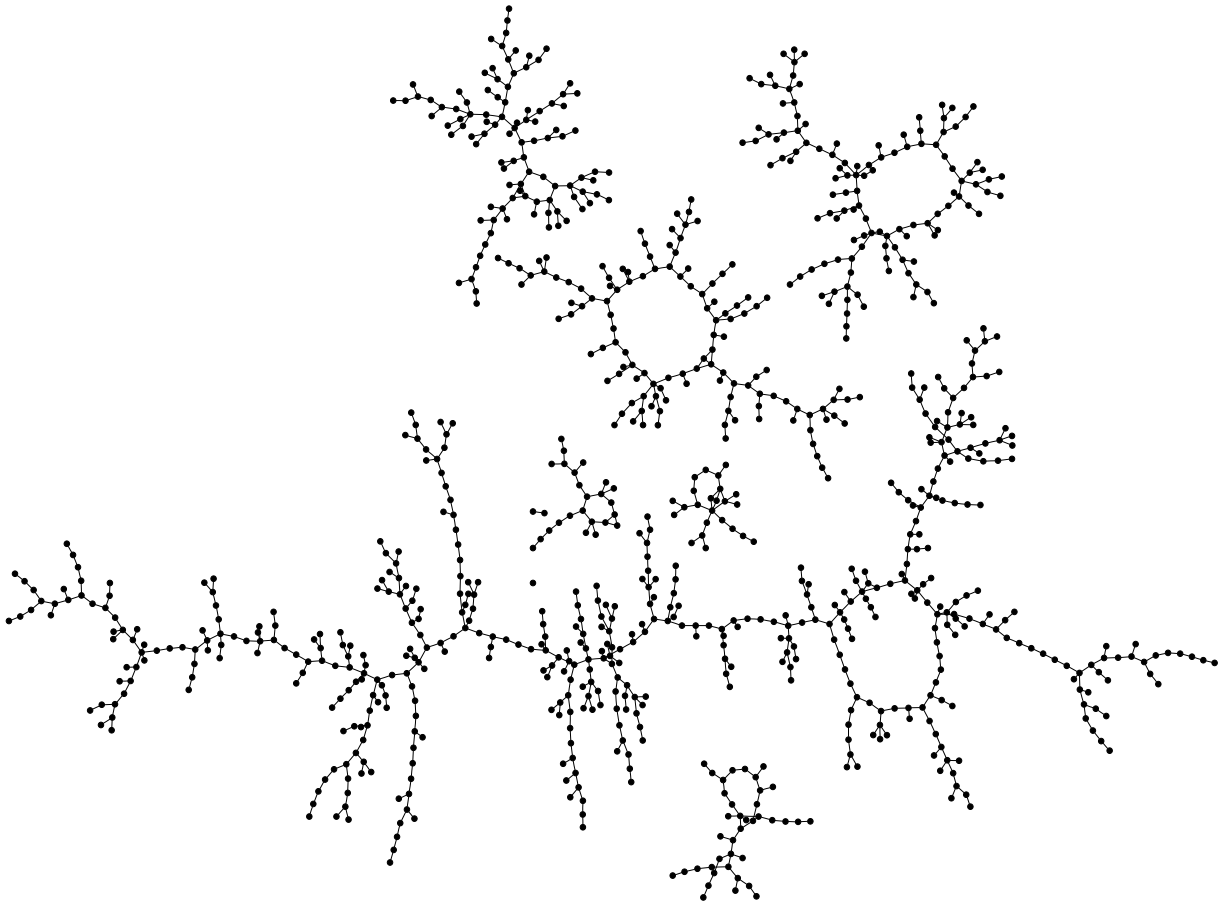


Figure 3.16.: A random graph of size $n = 956$ with no bow tie minor.

- *cycles,*
- K_4*, with one edge possibly subdivided edge,*
- *the graphs of Figures* 3.17 *and* 3.18*.*

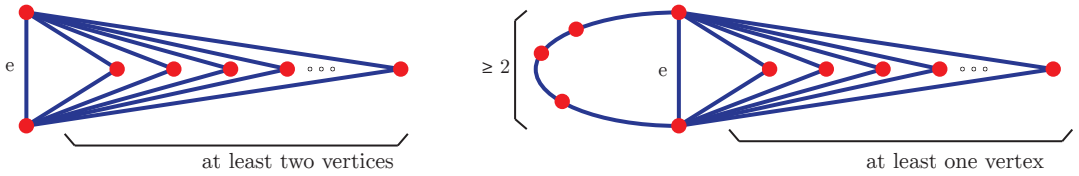


Figure 3.17.: A maximal cycle with one chord e and external vertices.

We will first establish a number of properties of cores excluding the bow tie as a minor. A *chord* of a cycle C is an edge, not in C, joining two vertices of C.

Lemma 3.7.3. *Let $C = (v_0, v_1, \ldots, v_{n-1})$ be a cycle in a core G excluding the bow tie. Then every chord of C joins vertices that are at distance 2 on C (we say that it is a short chord). Moreover, C has at most two chords. If it has two chords, say*

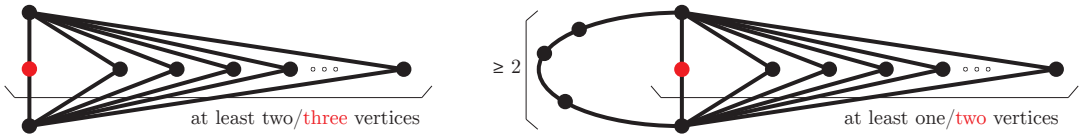


Figure 3.18.: A maximal cycle C with no chord and several external vertices.

$\{v_0, v_2\}$ and $\{v_i, v_{i+2}\}$, with $1 \leq i \leq n-1$ and $v_n = v_0, v_{n+1} = v_1$, then $v_i = v_1$ or $v_{i+2} = v_1$.

Proof. If a chord is not short, contracting it (together with some edges of C) gives a bow tie. Figure 3.19 then illustrates the second statement, which can be loosely restated as follows: the two chords cross and their four endpoints are consecutive on C. □

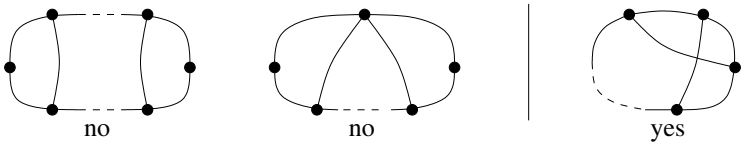


Figure 3.19.: The relative positions of two (short) chords in a cycle.

Lemma 3.7.4. *Let C be a cycle of maximal length in a core G with no bow tie minor. Let v be an external vertex, that is, a vertex not belonging to C. Then v is incident to exactly two edges, both ending on C and the endpoints of these edges are at distance 2 on C.*

Proof. Since G is a core, v belongs to a cycle C'. Since C is connected and excludes the bow tie, C' shares at least two vertices with C. Thus let P_1 and P_2 be two vertex-disjoint paths (taken from C') starting at v and ending in v_1 and v_2 on C, respectively where they hit C for the first time.Then v_1 and v_2 lie at distance at least 2 on C, otherwise C would not have maximal length. The same argument implies that P_1 and P_2 have length 1 each. If the distance between $v1$, $v2$ is ≥ 3 we can obtain a bow tie by contracting vv_1, $vv2$. So the distance of $v1, v2$ on C is equal to two.

Assume now that v has degree at least 3, and let e be a third edge (different from P_1 and P_2) adjacent to v. Again, e must belong to a cycle, sharing at least two vertices with C, and the same argument as before shows that e ends on C. But then Figure 3.20 shows that G contains a bow tie as a minor. □

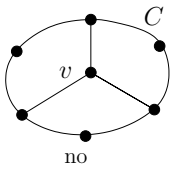


Figure 3.20.: A cycle C with an external vertex v of degree at least 3.

Lemma 3.7.5. *Let C be a cycle of maximal length in a core G excluding the bow tie as a minor. If C has two chords, it contains all vertices of G.*

Proof. Let e_1 and e_2 be the two chords of C. Lemma 3.7.3 describes their relative positions. Let v be a vertex not in C. Lemma 3.7.4 describes how it is connected to C. Contract one of the two edges incident to v to obtain a chord of C. By Lemma 3.7.3, this chord must be a copy of e_1 or e_2. But then Figure 3.21(a) shows that G contains a bow tie (delete the two bold edges). □



Figure 3.21.: A cycle C with (a) two chords and an external vertex v (b) one chord and an external vertex v.

Lemma 3.7.6. *Let C be a cycle of maximal length in a core G which does not contain the bow tie as a minor. If C has a chord e, all external vertices of C are adjacent to the endpoints of e.*

Proof. Let v be external to C. Contract one of the incident edges. This gives a chord e'. If e' is a copy of e, then we are done. Otherwise, the relative positions of e and e' are described by Lemma 3.7.3. But then Figure 3.21(b) shows that C has not maximal length (consider the cycle shown with the dotted line). □

Lemma 3.7.7. *Let C be a cycle of maximal length in a core G excluding the bow tie as a minor. If C has several external vertices, they are adjacent to the same points of C.*

Proof. Consider two external vertices v_1 and v_2. Lemma 3.7.4 describes how each of them is connected to C. Contract an edge incident to v_1 and an edge incident to v_2. This gives two chords of C. Either these two chords are copies of one another, which means that v_1 and v_2 are adjacent to the same points of C. Or the relative

position of these two chords is as described in Lemma 3.7.3. But then Figure 3.22 shows that G contains a bow tie (contract e). □

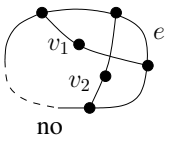


Figure 3.22.: A cycle C with two external vertices.

Proof of Propositions 3.7.1 and 3.7.2. Observe that a graph G does not contain bow tie as a minor if and only if its core (as defined in the proof of Proposition 3.2.5) does not contain it as a minor. Hence, if $\bar{C}(z)$ denotes the generating function of non-empty cores excluding the bow tie as a minor, we have

$$C(z) = T(z) - \frac{T(z)^2}{2} + \bar{C}(T(z)). \tag{3.39}$$

Using the above lemmas, we can now describe and count these non-empty cores. We start with cores which are cycles: their contribution to $\bar{C}(z)$ is given by (3.5). We now consider cores G having several cycles. Let C be a cycle of G of maximal length, chosen so that it has the largest possible number of chords. By Lemma 3.7.3, this number is 2, 1 or 0.

If C has two chords, it contains all vertices of G (Lemma 3.7.5). By Lemma 3.7.3 and Figure 3.19 (right), either all vertices have degree 3 and $G = K_4$, or G consists of K_4 where one edge is subdivided. This subdivision can also be seen as substituting an edge by a (directed) path. (Figure 3.23).

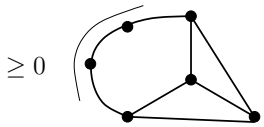


Figure 3.23.: K_4 with a subdivided edge.

This gives the generating function

$$\frac{z^4}{4!} + \frac{z^4}{4!} \cdot 6 \cdot \frac{z}{1-z}. \tag{3.40}$$

The second term comes from the following observation: We pick four vertices, these will be the vertices of degree 3 forming a K_4. Now, we have $\binom{4}{2} = 6$ choices to pick an edge in K_4 (our subdivided edge) which we replace by a (directed) path of length at least 1. See 3.4.2 for the generating function of a path.

Assume now that C has exactly one chord e. By Lemma 3.7.6, all external vertices are adjacent to the endpoints of e. To avoid problems with symmetries, we count separately the cores where C has length 4, or length ≥ 5 (Figure 3.17). This gives the generating function

$$\frac{z^2}{2}(e^z - 1 - z) + \frac{z^2}{2}(e^z - 1)\frac{z^2}{1-z}. \tag{3.41}$$

The first term counts cores where C has length 4. They can be constructed by picking two vertices and a set of size at least 2 of indistinguishable vertices. Hence they have the specification

$$\mathcal{Z} \times \mathcal{Z} \times SET_{\geq 1}(\mathcal{Z}).$$

To count cores where C has length ≥ 5, we have to modify our construction from the first term by taking into account a directed path from the first picked vertex to the second picked vertex and its specification is

$$\mathcal{Z} \times \mathcal{Z} \times SET_{\geq 1}(\mathcal{Z}) \times (\text{directed path of lenght at least 1}).$$

The factor here $z^2/(1-z)$ accounts for this directed path which is (a large) part of the maximal cycle.

Assume finally that C has no chord. By Lemma 3.7.7, all external vertices are adjacent to the same points of C. Again, we treat separately the cases where C has length 4, or length ≥ 5 (Figure 3.18). This gives the generating function

$$\frac{z^2}{2}(e^z - 1 - z - z^2/2) + \frac{z^2}{2}(e^z - 1 - z)\frac{z^2}{1-z}. \tag{3.42}$$

Putting together the contributions (3.5), (3.40), (3.41) and (3.42) gives the value of $\bar{C}(z)$ (the generating function of cores) from which we obtain the series $C(z)$ using (3.39). □

We now derive asymptotic results from Proposition 3.7.1.

Proposition 3.7.8 (**The asymptotic number of graphs avoiding the bow tie**). *As* $n \to \infty$,

$$C_n \sim n!\,\frac{e - 5/4}{\sqrt{2\pi}}\frac{e^n}{\sqrt{n}}$$

and

$$A_n \sim n!\,\frac{(e-5/4)^{1/6}e^{19/8-11e/3}}{\sqrt{6\pi}}\frac{e^n}{n^{2/3}}\,\exp\left(\frac{3}{2}(e-5/4)^{2/3}n^{1/3}\right).$$

Proof. Let us first recall that the series $T(z)$ has radius $1/e$, and can be continued analytically on the domain $\mathcal{D} := \mathbb{C} \setminus [1/e, +\infty)$. In fact, $T(z) = -W(-z)$, where W is the (principle branch of the) *Lambert function* [32]. The singular behaviour of $T(z)$ near $1/e$ is given by (2.15). Moreover, the image of $\mathcal{D}$ by T avoids the half-line $[1, +\infty)$. It thus follows from the expression (3.38) of $C(z)$ that $C(z)$ and $A(z)$ are analytic in the domain $\mathcal{D}$. Recall that by (2.15)

$$T(z) = 1 - \sqrt{2}(1 - ze)^{1/2} + \frac{2}{3}(1 - ze) - \frac{11\sqrt{2}}{36}(1 - ze)^{3/2} + O((1 - ze)^2)$$

as z approaches $1/e$ in a Δ-domain. Indeed it can be shown that $T(z)$ admits a full expansion in powers of $Z = \sqrt{1 - ze}$ around $z = 1/e$, i.e. around $z = 1/e$ $T(z) = \sum_{i=0}^{\infty} c_i Z^{n/2}$ for $c_i \in \mathbb{R}$. Take an expansion of $C(T)$ around $T(z) = 1$ and substitute T by $\sum_{i=0}^{\infty} c_i Z^{n/2}$. Expanding the whole expression in powers of Z and replacing Z by its value $\sqrt{1 - ze}$ we obtain

$$C(z) = \frac{e - 5/4}{\sqrt{2}\sqrt{1 - ze}} + \frac{1}{2} \log \frac{1}{\sqrt{2}\sqrt{1 - ze}} + O\left((1 - ze)^{-1}\right) \tag{3.43}$$

as z approaches $1/e$ in a Δ-domain. The above estimate of C_n then follows from singularity analysis.

Theorem VI.8 [49, p.419] tells us that in this case we can obtain an asymptotic expansion of Moreover, we derive from (2.15) that,

Estimating A_n is more involved. For the moment, let's assume that $A(z)$ is H-admissible. We then apply Theorem 2.4.1. The saddle point equation is $\zeta C'(\zeta) = n$ and we need to find an asymptotic estimate for ζ. First, we need to find an expression for $zC'(z)$. By [49, Thm VI.8] $C'(z)$ is analytic in the same Δ-domain than $C(z)$ is and we get an asymptotic expansion of $C'(z)$ by differentiating the asymptotic expansion (3.43) of $C(z)$ term by term. Hence the resulting function $zC'(z)$ is rational as the logarithm disappears after differentiating it. The saddle point equation then becomes

$$\frac{e - 5/4}{2\sqrt{2}(1 - \zeta e)^{3/2}} + \frac{1}{4(1 - \zeta e)} + O\left(\frac{1}{(1 - \zeta e)^{1/2}}\right) = n. \tag{3.44}$$

This gives the saddle point as

$$\zeta = \frac{1}{e} - \frac{(e - 5/4)^{2/3}}{2en^{2/3}} - \frac{1}{6en} + O(n^{-4/3}). \tag{3.45}$$

The idea is that $\frac{e-5/4}{2\sqrt{2}(1-\zeta e)^{3/2}} = n$ gives the first order of the saddle point and from there a good guess suggests that the next term should be of order n^{-1}. We now want to obtain estimates of the values $A(\zeta)$, ζ^n and $b(\zeta)$ occurring in Theorem 2.4.1.

Refining (3.43) into

$$C(z) = \frac{e-5/4}{\sqrt{2}\sqrt{1-ze}} + \frac{1}{4}\log\frac{1}{2(1-ze)} + \frac{19}{8} - \frac{11e}{3} + O(\sqrt{1-ze}), \tag{3.46}$$

we find

$$C(\zeta) = (e-5/4)^{2/3}n^{1/3} + \frac{1}{6}\log\frac{n}{e-5/4} + \frac{53}{24} - \frac{11e}{3} + O(n^{-1/3}),$$

which gives

$$A(\zeta) \sim \frac{e^{\frac{53}{24}-\frac{11e}{3}}}{(e-5/4)^{1/6}} n^{1/6}\exp((e-5/4)^{2/3}n^{1/3}). \tag{3.47}$$

It then follows from (3.45) that

$$\zeta^n \sim e^{-1/6}\exp\left(-n - (e-5/4)^{2/3}\, n^{1/3}/2\right). \tag{3.48}$$

Finally,

$$b(r) = rC'(r) + r^2C''(r) \sim \frac{3\sqrt{2}(e-5/4)}{8(1-re)^{5/2}}, \tag{3.49}$$

so that

$$b(\zeta) \sim \frac{3}{(e-5/4)^{2/3}}\, n^{5/3}.$$

Putting this estimate together with (3.47) and (3.48), we obtain the estimate of $A_n/n!$ given in the proposition. □

It remains to show that $A(z)$ from Proposition 3.7.1 is actually Hayman-admissible.

Proposition 3.7.9. *Let $C(z)$ and $A(z)$ be the series given in Proposition 3.7.1. Then the series $A(z)$ and $C^{(i+1)}(z)A(z)$ are H-admissible, for any $i \geq 1$.*

Proof. We begin with the series $A(z)$. Recall the analytic properties of $T(z)$, listed at the beginning of the proof of Proposition 3.7.8. For the capture condition $\mathbf{H_1}$ observe that

$$\begin{aligned} a(r) = rC'(r) &\sim \frac{e-5/4}{2\sqrt{2}(1-re)^{3/2}} \quad \text{and} \\ b(r) = rC'(r) + r^2C''(r) &\sim \frac{3\sqrt{2}(e-5/4)}{8(1-re)^{5/2}}. \end{aligned} \tag{3.50}$$

Let us now prove $\mathbf{H_2}$. By Taylor's formula applied to the function $f : \theta \mapsto C(re^{i\theta})$, we have, for $r \in (0, 1/e)$ and $\theta \in [-\theta_0, \theta_0]$:

$$\left|C(re^{i\theta}) - C(r) - i\theta a(r) + \theta^2 b(r)/2\right| \leq |\theta_0^3|/6 \sup_{|\alpha|\leq\theta_0} |f^{(3)}(\alpha)|$$

with

$$\begin{aligned}\left|f^{(3)}(\alpha)\right| &= \left|-ire^{i\alpha}C'(re^{i\alpha}) - 3ir^2e^{2i\alpha}C''(re^{i\alpha}) - ir^3e^{3i\alpha}C'''(re^{i\alpha})\right| \\ &\leq rC'(r) + 3r^2C''(r) + r^3C'''(r) \sim \frac{\kappa}{(1-re)^{7/2}}\end{aligned}$$

as $r \to 1/e$, for some constant κ. Hence for every $\theta_0 \equiv \theta_0(r) = o((1-re)^{7/6})$,

and r sufficiently close to $1/e$ (r tending to $1/e$)

$$\left|C(re^{i\theta}) - C(r) - i\theta a(r) + \theta^2 b(r)/2\right| \leq |\theta_0^3|/6 \sup_{|\alpha|\leq\theta_0} |f^{(3)}(\alpha)| \sim \frac{|\theta_0^3|\kappa}{6(1-re)^{7/2}} \to 0. \tag{3.51}$$

Then by exponentiating as $A(z) = e^{C(z)}$ we get

$$\sup_{|\theta|\leq\theta_0(r)} \left|\frac{A(re^{i\theta})}{A(r)} e^{-i\theta a(r)+\theta^2 b(r)/2} - 1\right| = \left|e^{o(1)} - 1\right| \to 0$$

as $r \to 1/e$. Thus $\mathbf{H_2}$ holds for such values of θ_0. We now take

$$\theta_0(r) = (1-re)^{6/5}, \tag{3.52}$$

and want to prove that $\mathbf{H_3}$ also holds. The condition is chosen precisely so that $\frac{\theta^2}{(1-re)^{5/2}} \to \infty$ and (3.51) both hold.

Recall that $C(z)$ is analytic on $\mathbb{C} \setminus [1/e, \infty)$, and let us isolate in $C(z)$ the part that diverges at $z = 1/e$:

$$C(z) = \frac{c}{\sqrt{1-ze}} + \frac{1}{4}\log\frac{1}{1-ze} + O(1) \tag{3.53}$$

where $c = \frac{e^{-5/4}}{\sqrt{2}} > 0$. It follows that For $\{|z| < 1/e\}$, there exists $M \in \mathbb{R}$ s.t.

$$B(z) := C(z) - \frac{c}{\sqrt{1-ze}} - \frac{1}{4}\log\frac{1}{1-ze} < M,$$

so $B(z)$ is uniformly bounded on . Hence, writing $z = re^{i\theta}$, we have

$$\sup_{|\theta|\in[\theta_0,\pi)} \left|\frac{A(z)}{A(r)}\sqrt{b(r)}\right| = \sup_{|\theta|\in[\theta_0,\pi)} \left|e^{C(z)-C(r)} \cdot \sqrt{b(r)}\right|$$
$$\leq M \sup_{|\theta|\in[\theta_0,\pi)} \left|\frac{1-re}{1-ze}\right|^{1/4} \left|\exp\left(\frac{c}{\sqrt{1-ze}} - \frac{c}{\sqrt{1-re}}\right)\right| \sqrt{b(r)}$$

for some constant M.

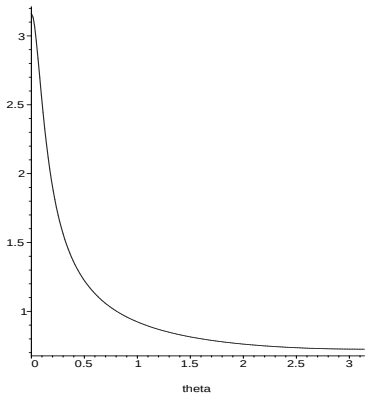


Figure 3.24.: Plot of $\Re\left(1/\sqrt{1-re^{i\theta}}\right)$ for $r = 0.9$.

For any z of modulus $r < 1/e$, we have $|1-re| \leq |1-ze|$, and we can bound the first factor above by 1. Also

$$\Re\left(1/\sqrt{1-ze}\right) = (1 + r^2 - 2r\cos(\theta))^{-1/4} \tag{3.54}$$

which can be obtained with Maple (see also [37, Theorem 4.1 proof] for a similar function). As $2r\cos(\theta)$ is a decreasing function in $\theta \in (0,\pi)$ we obtain that the function in (3.54) is decreasing in $\theta \in (0,\pi)$ (see Figure 3.24 for an illustration). Hence, denoting $z_0 = re^{i\theta_0}$

$$\sup_{|\theta|\in[\theta_0,\pi)} \left|\frac{A(z)}{A(r)}\sqrt{b(r)}\right| \leq M \exp\left(\Re\left(\frac{c}{\sqrt{1-z_0e}}\right) - \frac{c}{\sqrt{1-re}}\right)\sqrt{b(r)}.$$

Using Maple, we see that as $r \to 1/e$, the choice (3.52) of θ_0 implies that

$$\Re\left(\frac{c}{\sqrt{1-z_0e}}\right) - \frac{c}{\sqrt{1-re}} = -\frac{3c}{8(1-re)^{1/10}} + o(1) \to -\infty.$$

Hence (for M, T real constants) using the estimate (3.50) of $b(r)$

$$M \exp\left(\Re\left(\frac{c}{\sqrt{1-z_0 e}}\right) - \frac{c}{\sqrt{1-re}}\right)\sqrt{b(r)}$$
$$\leq T \exp\left(-\frac{3c}{8(1-re)^{1/10}} + o(1)\right) \cdot (1-re)^{-5/4} \to 0$$

and Condition $\mathbf{H_3}$ follows, .

Let us now consider the series $A_i(z) := C^{(i+1)}(z)A(z)$, for $i \geq 1$. By induction on i we have that for $i \geq 1$,

$$C^{(i)}(z) = \frac{(e-5/4)}{\sqrt{2}} \frac{\delta}{\delta i} \frac{1}{(1-ze)^{1/2}} \tag{3.55}$$
$$= \frac{(e-5/4)(2i)!e^i}{\sqrt{2}4^i i!} \cdot \frac{1}{(1-ze)^{i+1/2}} + O\left(\frac{1}{(1-ze)^i}\right). \tag{3.56}$$

This can be proved either from the expression of $C(z)$ given in Proposition 3.7.1, or by starting from the singular expansion (3.53) of $C(z)$ and applying [49, Thm. VI.8, p. 419].

Recall the behaviour (3.50) of the functions $a(r)$ and $b(r)$ associated with $A(z)$. It follows, with obvious notation, that as $r \to 1/e$,

$$a_i(r) = r\frac{\delta}{\delta r}\left(\log(C^{(i+1)}(r) + C(r))\right) \tag{3.57}$$
$$= a(r) + r\frac{C^{(i+2)}(r)}{C^{(i+1)}(r)} = a(r) + O\left(\frac{1}{1-re}\right) \tag{3.58}$$

and

$$b_i(r) = r\frac{\delta}{\delta r}\left(\log(C^{(i+1)}(r) + C(r))\right) + r^2\left(\frac{\delta}{\delta r}\right)^2\left(\log(C^{(i+1)}(r) + C(r))\right) \tag{3.59}$$
$$= b(r) + r\frac{C^{(i+2)}(r)}{C^{(i+1)}(r)} + r^2\frac{C^{(i+3)}(r)}{C^{(i+1)}(r)} - r^2\left(\frac{C^{(i+2)}(r)}{C^{(i+1)}(r)}\right)^2 \tag{3.60}$$
$$= b(r) + O\left(\frac{1}{(1-re)^2}\right) \tag{3.61}$$

both tend to infinity. Thus $\mathbf{H_1}$ holds.

Let us now prove that $A_i(z)$ satisfies $\mathbf{H_2}$ with the same value of θ_0 as for $A(z)$ (that is, $\theta_0 = (1-re)^{6/5}$). Thanks to (3.57–3.59) we have, for $|\theta| \le \theta_0$ and uniformly in θ,

$$\begin{aligned} e^{-i\theta a_i(r)+\theta^2 b_i(r)/2} &= e^{-i\theta a(r)+\theta^2 b(r)/2 - i\theta O((1-re)^{-1})+\theta^2 b(r)/2 O((1-re)^{-2})} \\ &= e^{-i\theta a(r)+\theta^2 b(r)/2}\left(1+O((1-re)^{1/5})\right). \end{aligned}$$

Now using (3.55), we obtain, denoting $z = re^{i\theta}$,

$$\frac{C^{(i+1)}(z)}{C^{(i+1)}(r)} = \left(\frac{1-ze}{1-re}\right)^{-i-3/2}\left(1+O((1-re)^{1/2})\right) = 1+O((1-re)^{1/5}).$$

Hence

$$\frac{A_i(z)}{A_i(r)}e^{-i\theta a_i(r)+\theta^2/2 b_i(r)} = \frac{A(z)}{A(r)}e^{-i\theta a(r)+\theta^2/2 b(r)}\left(1+O((1-re)^{1/5})\right)$$

and Condition $\mathbf{H_2}$ holds for A_i since it holds for A.

Finally, since $C(z)$ has non-negative coefficients, we have $\left|C^{(i+1)}(z)\right| \le C^{(i+1)}(r)$ for $z = re^{i\theta}$. Thus the fact that $A_i(z)$ satisfies $\mathbf{H_3}$ follows from the fact that $A(z)$ satisfies $\mathbf{H_3}$, together with $b_i(r) \sim b(r)$. □

Proposition 3.7.10 (Number of components — no bow tie). *Let $\mathcal{A} = Ex(\text{bow tie})$ and $\kappa(R_n)$ the number of components of $R_n \in_u \mathcal{A}$. Then*

$$\frac{\kappa(R_n) - \mathbb{E}(\kappa(R_n))}{\mathbb{V}(\kappa(R_n))} \xrightarrow{d} \mathcal{N}(0,1)$$

in distribution as $n \to \infty$ and the mean and variance of $\kappa(R_n)$ satisfy:

$$\mathbb{E}(\kappa(R_n)) \sim (e-5/4)^{2/3} n^{1/3}, \qquad \mathbb{V}(\kappa(R_n)) \sim \frac{2}{3}(e-5/4)^{2/3} n^{1/3},$$

Proof. We want to apply Theorem 2.4.5. By Proposition 3.7.9, $A(z)$ is H-admissible. Conditions (2.26)–(2.28) are readily checked, using

$$C(z) \sim \frac{e-5/4}{\sqrt{2}\sqrt{1-ze}}, \quad b(r) \sim \frac{3\sqrt{2}(e-5/4)}{8(1-ze)^{5/2}} \quad \text{and} \quad V(r) \sim \frac{\sqrt{2}(e-5/4)}{3\sqrt{1-ze}}.$$

The asymptotic estimates of $\mathbb{E}(\kappa(R_n))$ and $\mathbb{V}(\kappa(R_n))$ are obtained through (2.29), using the saddle point estimate (3.45). □

Since there are approximately $n^{1/3}$ components, one may expect the size S_n of the root component to be of the order of $n^{2/3}$. More precisely, we have the following result.

Proposition 3.7.11 (Size of the root component — no bow tie). *The normalized variable* $(e-5/4)^{2/3}S_n/(2n^{2/3})$ *converges in distribution to a* $\mathrm{Gamma}(3/2,1)$ *law, of density* $2\sqrt{x}e^{-x}/\sqrt{\pi}$ *on* $[0,\infty)$. *In fact, a local limit law holds: for* $x>0$ *and* $k=\lfloor x\frac{2n^{2/3}}{(e-5/4)^{2/3}}\rfloor$,

$$\frac{2n^{2/3}}{(e-5/4)^{2/3}}\ \mathbb{P}(S_n=k)\to 2\sqrt{\frac{x}{\pi}}e^{-x}.$$

The convergence of the factorial moments holds as well: for $i\geq 0$,

$$\mathbb{E}(S_n^i)\sim\frac{\Gamma(i+3/2)}{\Gamma(3/2)}\left(\frac{2n^{2/3}}{(e-5/4)^{2/3}}\right)^i.$$

Proof. The local (and hence global) limit law follows directly from Proposition 3.7.8, using (2.5). For the convergence of the moments, we start from (2.6). By Proposition 3.7.9 $C^{(i+1)}(z)A(z)$ is H-admissible. We then apply Theorem 2.4.1 to estimate the coefficient of z^n in this series (we will replace n ny $n-i-1$ later). Our calculations mimic those of Proposition 3.7.8, but the saddle point equation now reads

$$\zeta C'(\zeta)+\zeta\,\frac{C^{(i+2)}(\zeta)}{C^{(i+1)}(\zeta)}=n,$$

where $\zeta\equiv\zeta_n^{(i)}$ depends on i and n. Comparing with the original saddle point equation (3.44), and using the estimate (3.55) of $C^{(i)}(z)$, this reads

$$\frac{e-5/4}{2\sqrt{2}(1-\zeta e)^{3/2}}+\frac{7+4i}{4(1-\zeta e)}+O\left(\frac{1}{(1-\zeta e)^{1/2}}\right)=n.$$

This gives the saddle point as

$$\zeta=\frac{1}{e}-\frac{(e-5/4)^{2/3}}{2en^{2/3}}-\frac{7+4i}{6en}+O(n^{-4/3}). \tag{3.62}$$

We now want to obtain estimates of $C^{(i+1)}(\zeta)A(\zeta)$, ζ^n and $b_i(\zeta)$. We first derive from (3.46) that

$$C(\zeta)=(e-5/4)^{2/3}n^{1/3}+\frac{1}{6}\log\frac{n}{e-5/4}+\frac{29}{24}-\frac{11e}{3}-\frac{2i}{3}+O(n^{-1/3}).$$

This gives

$$A(\zeta)\sim\frac{e^{\frac{53}{24}-\frac{11e}{3}-\frac{2i}{3}}}{(e-5/4)^{1/6}}n^{1/6}\exp((e-5/4)^{2/3}n^{1/3}). \tag{3.63}$$

Moreover, we derive from (3.55) that

$$C^{(i+1)}(\zeta)\sim\frac{(2i+1)!}{2^i i!}e^{i+1}\frac{n^{1+2i/3}}{(e-5/4)^{2i/3}}. \tag{3.64}$$

It then follows from (3.62) that

$$\zeta^n \sim e^{-7/6-2i/3} \exp\left(-n - (e-5/4)^{2/3}\, n^{1/3}/2\right). \tag{3.65}$$

Finally, (3.59) and (3.49) give

$$b_i(\zeta) \sim b(\zeta) \sim \frac{3}{(e-5/4)^{2/3}}\, n^{5/3}.$$

Putting this estimate together with (3.63), (3.64) and (3.65), we obtain

$$[z^n]C^{(i+1)}(z)A(z) \sim \frac{(2i+1)!}{2^i i!} \frac{(e-5/4)^{1/6} e^{19/8-11e/3}}{\sqrt{6\pi}} \frac{e^{n+i+1}}{n^{2/3}} \frac{n^{1+2i/3}}{(e-5/4)^{2i/3}} \exp\left(\frac{3}{2}(e-5/4)^{2/3} n^{1/3}\right).$$

We finally replace n by $n-i-1$ (the only effect is to replace e^{n+i+1} by e^n), and divide by the estimate of $nA_n/n!$ given in Proposition 3.7.8: this gives the estimate of the i^{th} factorial moment of S_n as stated in the proposition, and concludes the proof. □

3.8. Random generation

As explained in Section 2.7 we will give the Boltzmann samplers of each of the discussed graph classes and briefly describe how we have addressed certain specific difficulties.

Trees and forests. Designing a Boltzmann sampler for rooted trees is explained in [46]. Note that the calculation of $T(x)$ can be avoided by feeding the sampler directly with the parameter $t = T(x)$, taken in $(0, 1]$. To sample *unrooted* trees, a first solution is to sample a rooted tree G and keep it with probability $1/|G|$. However, this sometimes generates large rooted trees that are rejected with high probability. A much better solution is presented in [33, Sec. 2.2.1]. In order to obtain an unrooted tree distributed according to (2.38), one calls the sampler of rooted trees with a *random parameter* t. The density of t must be chosen to be $(1-t)/C(x)$ on $[0, T(x)]$, with $C(x) = T(x) - T(x)^2/2$. To sample t according to this density, we set $t = 1 - \sqrt{1 - 2uC(x)}$, where u is uniform in $[0, 1]$. Again, we actually avoid computing the series $C(x)$ by feeding directly our sampler with the value $T(x) \in (0, 1]$. We use this trick for all classes that involve the series $T(x)$.

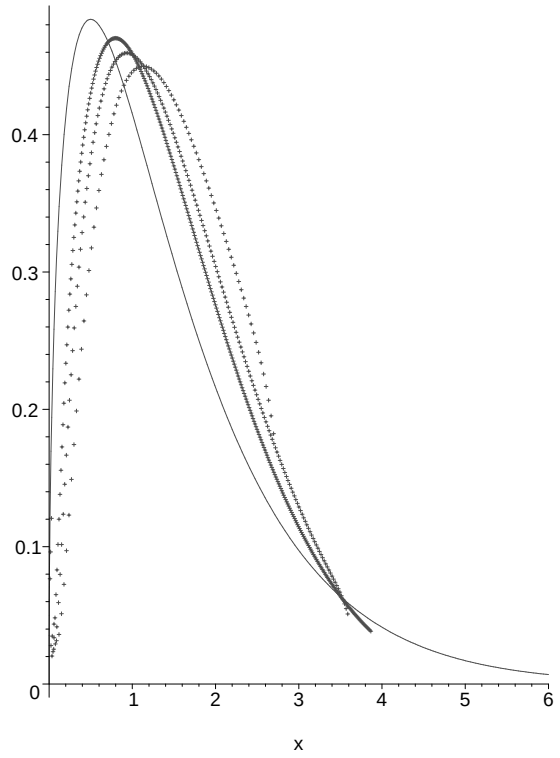


Figure 3.25.: Plots of $\frac{2n^{2/3}}{(e-5/4)^{2/3}}\mathbb{P}\left(S_n = \lfloor x\frac{2n^{2/3}}{(e-5/4)^{2/3}}\rfloor\right)$ against x, for $n = 100$, $n = 200$ and $n = 400$. The continuous curve is the expected limit, $2\sqrt{x/\pi}e^{-x}$.

To obtain large forests (Figure 2.2), we actually sample forests with a distinguished vertex; that is, a rooted tree plus a forest [46, Sec. 6.3].

Paths, cycles and stars. The `sequence` operator of [46, Sec. 4] produces *directed* paths, while we need undirected paths. Let u be uniform on $[0, 1]$. Our generator generates the one-vertex path if $u < x/C_p(x)$, with

$$C_p(x) = \frac{1}{2} \cdot \frac{1}{1-z} + \frac{z-1}{2}$$

and $C_p(x)$ is the generating function of paths, see Lemma 3.4.2. Otherwise it generates a path of length $2 + \mathrm{Geom}(x)$. An alternative is to generate a directed path, and reject it with probability $1/2$ if its size is at least 2.

Although the `cycle` operator of [46, Sec. 4] generates *oriented* cycles, this does not create a similar problem for our non-oriented cycles: indeed, a cycle of length at least 3 has exactly two possible orientations.

Designing a Boltzmann sampler ΓRS for rooted stars is elementary. For unrooted stars, we simply call ΓRS, but reject the star with probability 1/2 if it has size 2 (because the only star with two rootings has size 2).

Graphs excluding the bow tie. This is the most complex of our algorithms, because the generation of connected graphs involves 7 different cases (see the proof of Proposition 3.7.1). There is otherwise no particular difficulty. We specialize this algorithm to the generation of graphs avoiding the 2-spoon (Proposition 3.2.7). However, the probability to obtain a forest is about 0.95, and thus there is no point in drawing a random graph of this class.

The graphs shown in this chapter are taken from [24] and have been drawn using the `graphviz` software.

3.9. Summary and conlusion

In this chapter, we studied various examples of minor-closed classes of graphs with different types of excluded minors. A summary of these results can be found in Table 3.1. We observe that different behaviours can be observed and the behaviour of the generating function of the class of connected structures $C(z)$ next to its radius of convergence ρ seems to be crucial. Therefore in the table are ordered according to the speed of divergence of $C(z)$ next to ρ.

This raises the following question: is it possible to describe an explicit correlation between the properties of the excluded minors and the nature of the singularity of $C(z)$? For instance, it is known that $C(\rho)$ is finite when all excluded minors are 2-connected, but Section 3.2 shows that this happens as well with some non-2-connected excluded minors. Which excluded minors give rise to a simple pole in $C(z)$ (as in caterpillars)? or to a logarithmic singularity (as for graphs with no bow tie nor diamond), or to a singularity in $(1 - z/\rho)^{-1/2}$ (as for graphs with no bow tie)? In terms of the generating function the studied examples lead us to predict the following results when $C(z)$ has a unique dominant singularity and a singular behaviour of the form $(1 - z/\rho)^{-\alpha}$, with $\alpha > 0$:

- the mean and variance of the number $\kappa(R_n)$ of components scale like $n^{\alpha/(1+\alpha)}$, and $\kappa(R_n)$ admits a gaussian limit law after normalization,
- the mean of S_n scales like $n^{1/(1+\alpha)}$, and S_n, normalized by its expectation, converges to a Gamma distribution of parameter $\alpha+1$.

The second point as well as a general study of the behaviour of the size of the root component in decomposable structures is work in progress with M. Bousquet-Mélou.

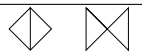
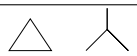
Excluded minors	$C(\rho)$	Sing. of $C(z)$	$\frac{C_n}{A_n}$	number $\kappa(R_n)$ of comp.	root comp. S_n	largest comp. L_n	Refs. and methods
2-connected	$<\infty$	?	$\geq 1/\sqrt{e}$ <1	$O(1)$ Poisson	$n-R_n$ $R_n \to$ discr.	$n-F_n$ $F_n \to$ discr.	[71], [62], [1] et al.
at least a spoon, no tree	$<\infty$	$(1-ze)^{3/2}$	>0 $\leq 1/\sqrt{e}$	id.	id.	id.	Sec. 3.2 sing. an.
	∞	$\log$ $(+\sqrt{\ })$	0	$\log n$ gaussian	n $\frac{1}{4}(1-x)^{-3/4}$	$\mathrm{PD}^{(1)}(1/4)$	Sec. 3.3 sing. an.
	∞	$1/\sqrt{\ }$	0	$n^{1/3}$ gaussian	$n^{2/3}$ $2\sqrt{x/\pi}e^{-x}$	?	Sec. 3.7 saddle
 (path forests)	∞	simple pole	0	$\sqrt{n}$ gaussian	$\sqrt{n}$ xe^{-x}	$\sqrt{n}\log n$ Gumbel	Sec. 3.4 saddle
 (forests of caterpillars)	∞	id.	0	id.	id.	?	Sec. 3.4 saddle
 (max. deg. 2)	∞	id. $(+\log)$	0	id.	id.	?	Sec. 3.6 saddle
all conn. graphs of size $k+1$	∞	entire (polyn.)	0	n/k gaussian	k Dirac	k Dirac	Sec. 3.5 saddle

Table 3.1.: Summary of the results: for each quantity $\kappa(R_n)$, S_n and L_n, we give an equivalent of the expected value (up to a multiplicative constant, except in the last line where constants are exact) and a description (name or density) of the limit law. The examples are ordered according to the speed of divergence of $C(z)$ near its radius ρ. The colour code indicates how violently $C(z)$ diverges. Blue stands for convergence and orange represents very strong divergence. Spoons are defined in Figure 3.3. As we get lower in the table, the graphs have more components, of a smaller size. The symbol $\mathrm{PD}^{(1)}(1/4)$ stands for the first component of a Poisson-Dirichlet distribution of parameter $1/4$, see Section 2.5. Observe that the class Ex(diamond), that is K_4 minus one edge, belongs to the class of graphs in the first line of the table.

4. Relatively bridge-addable classes of graphs

4.1. Bridge-addable and relatively bridge-addable

A class $\mathcal{A}$ of graphs is called *bridge-addable* if for any graph G in $\mathcal{A}$ and any pair of vertices u, v in different components the graph $G \cup \{uv\}$ obtained by adding the edge uv to G is also in $\mathcal{A}$. The concept of bridge-addability was introduced by McDiarmid et al.[76] in 2005 in the course of studying random planar graphs, and examples include forests, series-parallel graphs, planar graphs, and more generally graphs embeddable on any fixed surface, and minor-closed classes of graphs where the forbidden minors are all 2-connected.

If S is a finite non-empty set we write $R \in_u S$ to mean that R is drawn uniformly at random from S. For a collection $\mathcal{A}$ of graphs we write $R_n \in_u \mathcal{A}$ to mean that R_n is drawn uniformly at random from $\mathcal{A}_n$, the set of all graphs in $\mathcal{A}$ on vertex set $[n] := \{1, ..., n\}$. We will always assume that $\mathcal{A}_n$ is non-empty at least for large n.

By Theorem 2.1 in [76], for every finite bridge-addable set of graphs $\mathcal{A}$ and for $R \in_u \mathcal{A}$

$$\mathbb{P}(R \text{ connected}) \geq e^{-1}; \tag{4.1}$$

and indeed the random number $\kappa(R)$ of components is stochastically at most $1 +$ Po(1), that is

$$\kappa(R) \leq_s 1 + \text{Po}(1). \tag{4.2}$$

As mentioned before, here we assume implicitly that there is an n-vertex graph in $\mathcal{A}$; and for random variables X and Y, we say that X is *stochastically at most* Y and write $X \leq_s Y$ if $\mathbb{P}(X \leq t) \geq \mathbb{P}(Y \leq t)$ for each t.

Let $\mathcal{F}$ be the class of forests (acyclic graphs). By a result of Rényi [84], for $F_n \in_u \mathcal{F}$ we have

$$\mathbb{P}(F_n \text{ is connected}) \to e^{-1/2} \quad \text{as } n \to \infty,$$

and indeed

$$\kappa(F_n) \xrightarrow{d} 1 + \mathrm{Po}\left(\frac{1}{2}\right) \quad \text{as } n \to \infty.$$

We also refer to Section 2.3.1 for further details.

Since forests are plausibly the 'least connected' bridge-addable class of graphs, it was natural to think that at least asymptotically (4.1) is not tight, and the following conjecture was made:

Conjecture 4.1.1 ([77]). *If $\mathcal{A}$ is bridge-addable and $R_n \in_u \mathcal{A}$ then*

$$\liminf_{n\to\infty} \mathbb{P}(R_n \textit{ is connected}) \geq e^{-1/2}.$$

It was proved independently in [1] and [62] that under a further assumption on the class $\mathcal{A}$ – the class $\mathcal{A}$ has to be *bridge-alterable* – Conjecture 4.1.1 holds. A class of graphs $\mathcal{A}$ is *bridge-alterable* if it is closed under adding and deleting bridges. Thus $\mathcal{A}$ is bridge-alterable exactly when it satisfies the condition that for every graph H and every bridge e in H the graph H belongs to $\mathcal{A}$ if and only if $H \setminus e$ (the graph obtained by deleting e) belongs to $\mathcal{A}$. For $\mathcal{A}$ bridge-alterable and $R_n \in \mathcal{A}$ we have

$$\liminf_{n\in\mathbb{N}} \mathbb{P}(R_n \text{ is connected}) \geq e^{-1/2}. \tag{4.3}$$

Clearly, the class of forests is bridge-alterable and so the result is best possible.

Also all the examples of graph classes mentioned previously (that is, minor-closed classes of graphs defined by forbidding 2-connected minors and graphs embeddable on a fixed surface) are bridge-alterable. For *bridge-addable* classes of graphs, Conjecture 4.1.1 is still open. The best known bound is due to Balister, Bollobás and Gerke [7] where it was proved that

$$\liminf_{n\to\infty} \mathbb{P}(R_n \text{ is connected}) \geq e^{-0.7983}. \tag{4.4}$$

Given a sequence of non-negative integer-valued random variables $X_1, X_2, \ldots$ and Y we say that X_n is *stochastically at most Y asymptotically* and write $X_n \lesssim_s Y$, if for each fixed $t \geq 0$

$$\limsup_{n\to\infty} \mathbb{P}(X_n \geq t) \leq \mathbb{P}(Y \geq t).$$

In [73, Theorem 2.4] it was shown that the result (4.3) also yields a better bound on $\kappa(R_n)$: for any bridge-alterable class $\mathcal{A}$ of graphs and corresponding random graph $R_n \in_u \mathcal{A}$

$$\kappa(R_n) \lesssim_s 1 + \mathrm{Po}\left(\frac{1}{2}\right).$$

Similarly, the proof of (4.4) together with [73, Lemma 3.3] give that for any bridge-addable class $\kappa(R_n) \lesssim_s 1 + \mathrm{Po}\,(0.7983)$.

Let the *fragment size* $\mathrm{frag}(G)$ of a graph G be the number of vertices outside the largest component. In [73] it was shown that for each bridge-addable class $\mathcal{A}$ and each n, for $R_n \in_u \mathcal{A}$

$$\mathbb{E}[\mathrm{frag}(R_n)] < 2, \tag{4.5}$$

generalising and improving Lemma 2.6 in [71].

For $F_n \in_u \mathcal{F}$, where $\mathcal{F}$ is the class of forests, we know that $\mathbb{E}[\mathrm{frag}(F_n)] \to 1$ (see for example [71]) which leads us to the next conjecture, extending Conjecture 4.1.1:

Conjecture 4.1.2 ([73])**.** *Let $\mathcal{A}$ be bridge-addable and $R_n \in_u \mathcal{A}$. Then*

1. $\kappa(R_n) \lesssim_s 1 + \mathrm{Po}\left(\frac{1}{2}\right)$ *and*
2. $\limsup_{n\to\infty} \mathbb{E}[\mathrm{frag}(R_n)] \leq 1$.

Let us consider *rooted* bridge-addable classes of graphs denoted by $\mathcal{A}^\bullet$, i.e a bridge-addable class $\mathcal{A}$ where in every graph $G \in \mathcal{A}$ every component has a unique distinguished vertex called the *root*. It was proved in [73] that for $R^\bullet \in_u \mathcal{A}^\bullet$

$$\kappa(R^\bullet) \leq_s 1 + \mathrm{Po}(1). \tag{4.6}$$

Observe that this bound is tight: let $\mathcal{F}^\bullet$ be the class of forests of rooted trees. By a result of Rényi [84], for $F_n^\bullet \in_u \mathcal{F}^\bullet$ we have

$$\mathbb{P}(F_n^\bullet \text{ is connected}) \to e^{-1} \quad \text{as } n \to \infty,$$

and indeed

$$\kappa(F_n^\bullet) \xrightarrow{d} 1 + \mathrm{Po}\,(1) \quad \text{as } n \to \infty.$$

We also refer to Section 2.3.1 for further details. We will see that in some cases, having information about the rooted case allows us to deduce results about the unrooted cases.

In the following we will generalise the definition of bridge-addable and show that the results (4.1), (4.5) and (4.6) are in fact special cases of a more general picture.

Fix a graph G and call it the *host graph*. Let $\mathcal{A}$ be a collection of subgraphs of G. We call $\mathcal{A}$ *relatively bridge-addable (with respect to* G*)* if for every $H \in \mathcal{A}$ and every pair of vertices u, v which are adjacent in G and lie in different components of H the graph $H \cup uv$ is also in $\mathcal{A}$. An illustration can be found in Figure 4.1. Observe that for $G = K_n$ *relatively bridge-addable* is equivalent to *bridge-addable* in the usual sense of [76]. An example of a relatively bridge-addable class of graphs for any host graph G is the set of acyclic subgraphs of G.

For disjoint sets B and C of vertices in G we let $e(B, C)$ denote the number of edges between B and C. If $e(B, C) \geq \alpha |B||C|$ for each partition $B \cup C$ of $V(G)$ we say that G is an α*-edge-expander*. Thus the complete graph K_n is a 1-edge-expander. Another natural example of an expander is a d-regular graph on n vertices. In this case we may take $\alpha = \frac{d-\lambda}{n}$, where λ is the second largest eigenvalue of the adjacency matrix of G, see [4, Theorem 9.1.2.]. Hence (or otherwise) $K_{n,n}$ is a $\frac{1}{2}$-edge-expander. Another example is the classical random graph $G_{n,p}$: with high probability $G_{n,p}$ is a $(1-\epsilon)p$-edge-expander for $p \geq C \log n/n$. To see this observe that for any fixed partition with $|B| = i$ and $|C| = n - i$, the probability that $e(B, C) < (1-\epsilon)p\, i(n-i)$ is at most $e^{-\Theta(\epsilon^2 i(n-i)p)}$ by standard Chernoff bounds, see for example [23]. Hence, for $p \geq C \log n/n$, with $C = C(\epsilon)$ large enough, we see that the probability that $G_{n,p}$ is not an $(1-\epsilon)p$-edge-expander is at most $\sum_{1 \leq i \leq n/2} n^i n^{-2i} = o(1)$.

The following Lemma will be a very important tool.

Lemma 4.1.3 (Lemma 3.4. in [73])**.** *Let the random variable X take non-negative integer values. Let $\alpha > 0$ and suppose*

$$\mathbb{P}(X = k+1) \leq \frac{\alpha}{k+1} \mathbb{P}(X = k) \quad \textit{for each } k = 0, 1, 2, \ldots$$

Then $X \leq_s Y$ where $Y \sim \mathrm{Po}(\alpha)$.

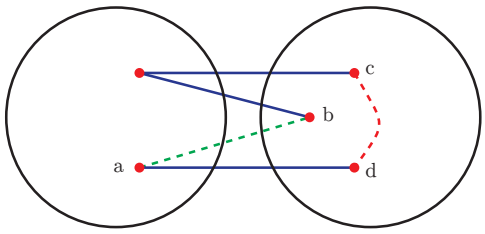


Figure 4.1.: $K_{n,n}$ as a host graph G. Suppose the graph with the blue edges is in our bridge-addable class. Then the graph consisting of the blue edges and the green edge also has to be in the class but not necessarily the graph with the blue edges and the red edge.

4.2. Expander host graphs

Recall that K_n is an α-edge-expander for $\alpha = 1$. The following result then generalises inequality (4.1) which is Theorem 2.2 of [76].

Theorem 4.2.1. *Let $0 < \alpha \leq 1$, let G be an α-edge-expander, let the set $\mathcal{A}$ of subgraphs of G be relatively bridge-addable, and let $R \in_u \mathcal{A}$. Then*

$$\kappa(R) \leq_s 1 + \mathrm{Po}(1/\alpha),$$

and in particular

$$\mathbb{P}(R \text{ is connected}) \geq e^{-1/\alpha}.$$

Proof. For a subgraph H of G, let Bridge(H) denote the set of bridges of H, with |Bridge(H)| = bridge(H); and let Cross(H) denote the set of edges of G between components of H, with cross(H) = |Cross(H)|. Let $\mathcal{A}^k$ denote the set of graphs in $\mathcal{A}$ with k components. Let G have n vertices. For $k = 1, \ldots, n$ define

$$a_k = \min_{H \in \mathcal{A}^k} \mathrm{cross(H)}$$

and

$$b_k = \max_{H \in \mathcal{A}^k} \mathrm{bridge(H)}.$$

Then for each $k = 1, \ldots, n-1$

$$a_{k+1}|\mathcal{A}^{k+1}| \leq \sum_{H \in \mathcal{A}^{k+1}} \mathrm{cross(H)} \leq \sum_{H \in \mathcal{A}^k} \mathrm{bridge(H)} \leq b_k|\mathcal{A}^k|.$$

We shall see that $b_k \leq n - k$ and $a_{k+1} \geq \alpha k(n-k)$. Then

$$|\mathcal{A}^{k+1}| \leq \frac{b_k}{a_{k+1}}|\mathcal{A}^k| \leq \frac{1}{\alpha k}|\mathcal{A}^k|$$

and so

$$\mathbb{P}(\kappa(R) = k+1) \leq \frac{1}{\alpha k}\mathbb{P}(\kappa(R) = k). \tag{4.7}$$

Applying Lemma 4.1.3 together with (4.7) then gives the theorem.

It remains to bound a_{k+1} and b_k. But observe that $b_k \leq n - k$ since the number of edges in a spanning forest of a graph on n vertices with k components is $n - k$; so it remains now to bound a_{k+1}.

We want to find a graph H with $k+1$ components which minimises cross(H). Take an arbitrary graph $H \in \mathcal{A}$ with $k+1$ components, where the components have $n_1, \ldots, n_{k+1}$ vertices. As G is an α-edge-expander, we get

$$\text{cross}(H) \;\geq\; \frac{\alpha}{2} \sum_{i=1}^{k+1} n_i(n-n_i) \;=\; \frac{\alpha}{2}\left(n^2 - \sum_{i=1}^{k+1} n_i^2\right).$$

Thus we want to maximise $\sum_{i=1}^{k+1} n_i^2$ subject to $\sum_{i=1}^{k+1} n_i = n$ and $n_i \geq 1$ for all i. The maximum is attained at $n_1 = n-k$ and $n_2 = \ldots = n_{k+1} = 1$ with value $(n-k)^2 + k = n^2 - 2kn + k^2 + k$. Hence

$$\text{cross}(H) \geq \frac{\alpha}{2}\left(2nk - k^2 - k\right) = \frac{\alpha}{2}\left(2k(n-k) + k^2 - k\right) \geq \alpha k(n-k);$$

and so $a_{k+1} \geq \alpha k(n-k)$, as required. □

Recall that frag(G) is the number of vertices which are not in the largest component. The following result on the expected number of vertices not in the largest component generalises the result (4.5) above (from [73]).

Theorem 4.2.2. *Let G be an α-edge-expander, let the set $\mathcal{A}$ of subgraphs of G be relatively bridge-addable, and let $R \in_u \mathcal{A}$. Then*

$$\mathbb{E}[\text{frag}(R)] < 2/\alpha.$$

Proof. Let G have n vertices. Let us prove the following claim first:

$$\text{cross}(H) \geq \alpha\, \frac{n}{2}\, \text{frag}(H) \tag{4.8}$$

for each graph H in $\mathcal{A}$.

Observe that it is trivially true for $\text{frag}(H) \leq \frac{n}{2}$ as G is an α-edge-expander. For the general case, let $H \in \mathcal{A}$ have k components, with $n_1 \geq \cdots \geq n_k$ vertices. As G is an α-edge-expander, as earlier

$$\text{cross}(H) \;\; \geq \tfrac{\alpha}{2} \textstyle\sum_{i=1}^{k} n_i(n-n_i) \;\; \geq \dfrac{\alpha}{2}\left(n^2 - \displaystyle\sum_{i=1}^{k} n_i^2\right).$$

But

$$\sum_{i=1}^{k} n_i^2 \leq \sum_{i=1}^{k} n_1\, n_i = n_1\, n.$$

Thus since $\mathrm{frag}(H) = n - n_1$ we get

$$\mathrm{cross}(H) \geq \frac{\alpha}{2}\, n(n-n_1) = \frac{\alpha\, n}{2}\, \mathrm{frag}(H),$$

which establishes the claim (4.8).

By (4.8) we have

$$\frac{\alpha\, n}{2} \sum_{H\in\mathcal{A}} \mathrm{frag}(H) \leq \sum_{H\in\mathcal{A}} \mathrm{cross}(H) \leq \sum_{H\in\mathcal{A}} \mathrm{bridge}(H).$$

Hence

$$\mathbb{E}[\mathrm{frag}(R)] \leq \frac{1}{|\mathcal{A}|} \cdot \frac{2}{\alpha\, n} \sum_{H\in\mathcal{A}} \mathrm{bridge}(H) < \frac{2}{\alpha},$$

completing the proof of the theorem. □

Are the bounds of Theorem 4.2.1 and Theorem 4.2.2 tight? In the bridge-addable case they are conjectured (Conjectures 4.1.1 and 4.1.2) not to be tight: for $G = K_n$ the 'correct' bounds are believed to be $1+\mathrm{Po}\left(\frac{1}{2}\right)$ and $\mathbb{E}[\mathrm{frag}(R_n)] < 1$, which would correspond to the class of forests.

4.3. $K_{n/2,n/2}$ as a host graph

What happens for particular other α-edge-expanders? Are the bounds of Theorem 4.2.1 and Theorem 4.2.2 tight? In the following, we will have a closer look at the complete bipartite graph $K_{n/2,n/2}$, which is a $\frac{1}{2}$-edge-expander. Here, $K_{n/2,n/2}$ really means $K_{\lfloor n/2\rfloor,\lceil n/2\rceil}$ but we write $K_{n/2,n/2}$ to simplify notation. To be precise $K_{n/2,n/2}$ is only a $\alpha - \frac{2}{n}$-edge-expander as for $|S| \leq \frac{n}{2}$ we have (deleting one vertex)

$$e(S,\bar{S}) \geq e(S,\bar{S}-v) \geq \frac{1}{2}|S|(n-|S|-1) = |S||\bar{S}|\left(\frac{1}{2} - \frac{2}{n}\right)$$

but this will not be of any importance.

Theorem 4.3.1. *Consider $K_{n/2,n/2}$, say with parts $\{1,\ldots,\lfloor n/2\rfloor\}$ and $\{\lfloor n/2\rfloor + 1,\ldots,n\}$. Let the random graph R_n be sampled uniformly at random from the forests in this graph. Then as $n\to\infty$*

(a)

$$\kappa(R_n) \to_d 1 + \mathrm{Po}(1)$$

and so in particular $\mathbb{P}(R_n$ is connected$) \to e^{-1}$;

(b) the unlabelled trees T appear in $\mathrm{Frag}(R_n)$ *asymptotically independently, with distribution* $\mathrm{Po}(\lambda(T))$ *where* $\lambda(T) = 2/(e^{v(T)}aut(T))$ *and*

(c)

$$\mathbb{E}[\mathrm{frag}(R_n)] \to 2.$$

For comparison, let F_n be sampled uniformly from the n-vertex forests in K_n. Then $\mathbb{P}(F_n \text{ is connected}) \to e^{-\frac{1}{2}}$ as $n \to \infty$, and indeed $\kappa(F_n)$ converges in distribution to $1 + \mathrm{Po}(\frac{1}{2})$, as mentioned earlier. Furthermore, as $n \to \infty$, $\mathbb{E}[\mathrm{frag}(F_n)] \to 1$, and the unlabelled trees T appear in $\mathrm{Frag}(F_n)$ asymptotically independently, with distribution $\mathrm{Po}(\mu(T))$ where $\mu(T) = 1/(e^{v(T)}\mathrm{aut}(T))$ (see [76]).

Thus not only is it true asymptotically that $\mathbb{E}[\kappa(\mathrm{Frag}(R_n))] \sim 2\,\mathbb{E}[\kappa(\mathrm{Frag}(F_n))]$ and $\mathbb{E}[\mathrm{frag}(R_n)] \sim 2\,\mathbb{E}[\mathrm{frag}(F_n)]$ but it is even true for each unlabelled tree T that the expected number of appearances of T in $\mathrm{Frag}(R_n)$ is twice that in $\mathrm{Frag}(F_n)$.

For the proof we need the following exact counting result.

Lemma 4.3.2 ([87], [6])**.** *The number of spanning (unrooted) trees of $K_{r,s}$ (fixed partition) is*

$$r^{s-1} \cdot s^{r-1}.$$

Proof. The proof of [87] uses generating functions and Lagrange inversion. Unfortunately, [6] is not accessible. □

More references and the proof of this result can be found in [79, p. 30-32].

Proof. We introduce a variant S_n of the random forest R_n, and estimate the probability that $\mathrm{Frag}(S_n)$ is a given forest F, in terms of the probability that R_n is connected, see (4.9) below. By summing over F we estimate the probability that $\mathrm{Frag}(R_n)$ consists of j trees each of size at most ℓ, again in terms of the probability that R_n is connected, see (4.10) below. From that result we determine the limiting probability that R_n is connected, and complete the proof of part (a). For part (b) we start from (4.9) but replace $\mathbb{P}(R_n \text{ is connected})$ with its limiting value $1/e$. We see that the probability that $\mathrm{Frag}(R_n)$ contains a large tree is small and complete the proof. Finally, part (c) follows from part (b), once we show that the expected contribution from large components is negligible. To see this we compare the contribution of large trees to the fragments of random forests sampled from $K_{n/2,n/2}$ to the ones sampled from K_n.

Proof of part (a)

Consider the random graph S_n generated by choosing a partition $\mathbf{U} = (U_1, U_2)$ of $[n]$ with $|U_1| = \lceil \frac{n}{2} \rceil$ uniformly at random, letting $\tilde{H}_n$ be the complete bipartite graph with parts U_i, and picking a forest in $\tilde{H}_n$ uniformly at random. Clearly $\kappa(S_n)$ and $\kappa(R_n)$ have the same distribution and indeed the probability that S_n is isomorphic to a fixed, unlabelled forest F is equal to the probability that R_n is isomorphic to F. Fix a forest F on $[n]$ consisting of j trees with $t_1, \ldots, t_j$ vertices, and let $t = \sum_i t_i$. We claim that

$$\mathbb{P}(\mathrm{Frag}(S_n) = F) \sim \mathbb{P}(R_n \text{ is connected}) \cdot 2^j (en)^{-t}. \tag{4.9}$$

Let us prove this. Note first that the probability that $\mathbf{U} = (U_1, U_2)$ gives a proper 2-colouring of F is asymptotically equal to $2^{-(t-j)}$. To see this fix a colouring $\Phi : V(F) \to \{1, 2\}$. Let a be the number of $1'$s and b the number of $2'$s in F. For fixed U we have

$$\mathbb{P}(\ U \text{ restricted to } V(F) \text{ is } \Phi) = \frac{\binom{n-a-b}{\frac{n}{2}-a}}{\binom{n}{\frac{n}{2}}} \sim 2^{-(a+b)} = 2^{-t}$$

and so the probability that U is a proper 2-colouring of F is

$$\mathbb{P}(U \text{ is a proper 2-colouring of } F) \sim 2^j \cdot 2^{-t}.$$

Fix a partition $\mathbf{U}^0 = (U_1^0, U_2^0)$ of $[n]$ that gives a proper 2-colouring of F. Let x_i be the number of vertices of F in U_i^0, so that $\sum_i x_i = t$. We may assume that $t < \lfloor n/2 \rfloor$, and so each $n_i \geq 1$. Then by Lemma 4.3.2 and the fact that $(1 + y/m)^m \to e^y$ as $m \to \infty$,

$$\begin{aligned}
& \mathbb{P}(\mathrm{Frag}(S_n = F) \,|\, \mathbf{U} = \mathbf{U}^0) / \mathbb{P}(R_n \text{ is connected}) \\
= \ & \left(\frac{n}{2} - x\right)^{\frac{n}{2}-y-1} \left(\frac{n}{2} - y\right)^{\frac{n}{2}-x-1} \cdot \left(\frac{n}{2}\right)^{-n+2} \\
\sim \ & \left(\frac{n}{2}\right)^{n-t-2} \left(1 - \frac{2x}{n}\right)^{n/2} \left(1 - \frac{2y}{n}\right)^{n/2} \cdot \left(\frac{n}{2}\right)^{-n+2} \\
\sim \ & \left(e\frac{n}{2}\right)^{-t}.
\end{aligned}$$

This holds uniformly for each relevant partition $\mathbf{U}^0$, and hence we can relax the conditioning, to obtain

$$\frac{\mathbb{P}(\mathrm{Frag}(S_n) = F \,|\, \mathbf{U} \text{ properly colours } F)}{\mathbb{P}(R_n \text{ is connected})} \sim (2\,en)^{-t}.$$

Hence

$$\begin{aligned}
&\frac{\mathbb{P}(\mathrm{Frag}(S_n)=F)}{\mathbb{P}(R_n \text{ is connected})}\\
= \;& \frac{\mathbb{P}(\mathrm{Frag}(S_n)=F)\,|\,\mathbf{U} \text{ properly colours } F)}{\mathbb{P}(R_n \text{ is connected})}\cdot \mathbb{P}(\mathbf{U} \text{ properly colours } F)\\
\sim \;& (2\,en)^{-t}\cdot 2^{t-j} \;=\; 2^{-j}\cdot (en)^{-t},
\end{aligned}$$

which gives (4.9).

Our next task is to use (4.9) to prove (4.10) below. We need some notation. Let $\beta_\ell = \sum_{j=1}^{\ell} \frac{j^{j-2}}{j!\,e^j}$, which is the truncated generating function of (unrooted labelled) trees, evaluated at its radius of convergence $1/e$. Let $\tilde{\beta}_\ell = 2\beta_\ell$. From Section 2.3.1 we know that β_ℓ increases to $\frac{1}{2}$ as $\ell \to \infty$, and so $\tilde{\beta}_\ell$ increases to 1. Thus $\sum_{j=0}^{\ell} \frac{(\tilde{\beta}_\ell)^j}{j!}$ increases to e^1 as $\ell \to \infty$.

Let $p(n, j, \leq \ell)$ be the probability that $\mathrm{Frag}(R_n)$ consists of j component trees, each of order at most ℓ. We claim that, for each fixed j and ℓ, as $n \to \infty$

$$p(n, j, \leq \ell) \sim \mathbb{P}(R_n \text{ is connected}) \cdot \frac{(\tilde{\beta}_\ell)^j}{j!}. \tag{4.10}$$

Let us prove this. Let $\mathcal{F}_n(j, \ell)$ denote the set of forests on $[n]$ consisting of j trees each with at most ℓ vertices. Then

$$\begin{aligned}
& p(n, j, \leq \ell)/\mathbb{P}(R_n \text{ is connected})\\
= \;& \sum_{F \in \mathcal{F}_n(j,\ell)} \mathbb{P}(\mathrm{Frag}(S_n)=F)/\mathbb{P}(R_n \text{ is connected})\\
\sim \;& \sum_{F \in \mathcal{F}_n(j,\ell)} 2^j \cdot (en)^{-v(F)}\\
\sim \;& \frac{1}{j!} \sum_{1 \leq t_1 \leq \ell} \cdots \sum_{1 \leq t_j \leq \ell} \binom{n}{t_1}\binom{n-t_1}{t_2}\cdots\binom{n-(t_1+\cdots t_{j-1})}{t_j} \cdot \prod_{i=1}^{j} {t_i}^{t_i-2} \cdot 2^j (en)^{-\sum_i t_i}.
\end{aligned}$$

Using $\binom{n-\sum_i t_i}{n_i} \sim \frac{n^{t_i}}{t_i!}$

$$\begin{aligned}
\sim \;& \frac{1}{j!} \sum_{1 \leq t_1 \leq \ell} \cdots \sum_{1 \leq t_j \leq \ell} \prod_{i=1}^{j} \frac{n^{t_i}}{t_i!} {t_i}^{t_i-2} \cdot 2^j (en)^{-\sum_i t_i}\\
= \;& \frac{1}{j!} \sum_{1 \leq t_1 \leq \ell} \cdots \sum_{1 \leq t_j \leq \ell} \prod_{i=1}^{j} \left(\frac{{t_i}^{t_i-2}}{t_i! e^{t_i}} \cdot 2 \right)\\
= \;& \frac{1}{j!} \left(\sum_{i=1}^{\ell} \frac{i^{i-2}}{i! e^i} \cdot 2 \right)^j \;=\; \frac{1}{j!} \left(\tilde{\beta}_\ell\right)^j,
\end{aligned}$$

which completes the proof of (4.10).

We may now use (4.10) to complete the proof of part (a). Let

$$p(n, \leq \ell, \leq \ell) = \sum_{j=0}^{\ell} p(n, j, \leq \ell),$$

the probability that $\mathrm{Frag}(R_n)$ consists of at most ℓ components each of size at most ℓ. Recall that $\mathbb{E}[\mathrm{frag}(R_n)] \leq 4$ for each n (Theorem 4.2.2). Hence by Markov's inequality

$$1 - p(n, \leq \ell, \leq \ell) \leq \mathbb{P}(\mathrm{frag}(R_n) > \ell) < \mathbb{E}[\mathrm{frag}(R_n)]/\ell \leq c/\ell.$$

So by taking ℓ large enough we may ensure that $p(n, \leq \ell, \leq \ell)$ is arbitrarily close to 1. But by (4.10), for each (fixed) ℓ

$$p(n, \leq \ell, \leq \ell) \sim \mathbb{P}(R_n \text{ is connected}) \cdot \sum_{j=0}^{\ell} \frac{1}{j!} \left(\tilde{\beta}_\ell\right)^j.$$

Also, as we noted earlier, the sum on the right side tends to e^1 as $\ell \to \infty$. Thus

$$\mathbb{P}(R_n \text{ is connected}) \to e^{-1} \quad \text{as } n \to \infty. \tag{4.11}$$

Furthermore, we may argue as above to see that for each fixed j, by taking ℓ large enough we can ensure that $p(n, j, \leq \ell)$ is arbitrarily close to $\mathbb{P}(\kappa(S_n) = 1 + j)$, and $\frac{(\tilde{\beta}_\ell)^j}{j!}$ is arbitrarily close to $\frac{\lambda^j}{j!}$. But by (4.10) and (4.11), as $n \to \infty$

$$p(n, j, \leq \ell) \sim e^{-1} \cdot \frac{(\tilde{\beta}_\ell)^j}{j!}$$

and so

$$\mathbb{P}(\kappa(S_n) = 1 + j) \to e^{-1} \cdot \frac{\lambda^j}{j!}.$$

Thus $\kappa(R_n)$ converges in distribution to $1+\mathrm{Po}(\lambda)$. But we know also that $\kappa(R_n) \leq_s 1 + \mathrm{Po}(2)$ by Theorem 4.2.1, and we obtain convergence in any mean of $\kappa(R_n)$ to $1 + \mathrm{Po}(\lambda)$.

Proof of part (b)
Let $\mathcal{T}$ be the class of all unlabelled trees. Let $A \subseteq \mathcal{T}$ be finite and non-empty.

Let $n(T)$ be a non-negative integer for each $T \in A$, and let E_n^1 be the event that $\kappa(R_n, T) = n(T)$ for each $T \in A$. We wish to show that with $\lambda(T) = \frac{2}{e^{v(T)}\mathrm{aut}(T)}$

$$\mathbb{P}(E_n^1) \to p := \prod_{T \in A} e^{-\lambda(T)} \frac{\lambda(T)^{n(T)}}{n(T)!} \quad \text{as } n \to \infty. \tag{4.12}$$

Let $0 < \epsilon < 1$. We shall show that $|\mathbb{P}(E_n^1) - p| < 4\epsilon$ for n sufficiently large, to complete the proof.

Enumerate the trees in $\mathcal{T}$ as $T_1, T_2, \ldots$. We shall choose (large) constants k and b. For each positive integer j let $B_j = \{T_1, \ldots, T_j\}$, and let $B_j' = B_j \setminus A$ and $\overline{B_j} = \mathcal{T} \setminus B_j$.

Choose m large enough that $A \subseteq B_k$, and each tree in $\overline{B_m}$ has size at least $4/\epsilon$. Then since $\mathbb{E}[\mathrm{frag}(R_n)] \leq 4$ for each n by Theorem 4.2.2

$$\mathbb{P}\left(\kappa(\mathrm{Frag}(R_n), T) > 0 \text{ for some } T \in \overline{B_m}\right) \leq \epsilon.$$

Let $\mathcal{T}$ denote the class of labelled trees (in K_n). Recall that for the corresponding exponential generating function $U(z)$,

$$U(z) = \sum_{T \in \mathcal{T}} \frac{z^n}{v(T)!} = \sum_{n \geq 1} \frac{n^{n-2} z^n}{n!} \tag{4.13}$$

and $U(e^{-1}) = \frac{1}{2}$ (see Section 2.3.1). But for each $T \in \mathcal{U}$

$$\frac{z^{v(T)}}{\mathrm{aut}(T)} = \sum_{T' \in \mathcal{T}} \frac{z^{v(T)}}{v(T)!}.$$

Hence

$$\sum_{T \in \mathcal{T}} \lambda(T) = 2U\left(\frac{1}{e}\right) = 1. \tag{4.14}$$

Now increase m if necessary so that it is large enough that also

$$\prod_{T \in \overline{B_m}} e^{-\lambda(T)} \geq (1+\epsilon)^{-1}.$$

Now choose b large enough that both $\mathbb{P}(\kappa(R_n) > b) \leq \epsilon$ for all n, and

$$\prod_{T \in B_k} \mathbb{P}(\mathrm{Po}(\lambda(T)) \leq b) \geq (1+\epsilon)^{-1}.$$

Let the (unlabelled) forest F consists of k_T copies of tree T for each T in the finite set $S \subset \mathcal{T}$. For $T \in \mathcal{T}$ the random variable $\kappa(R_n, T)$ counts the number

of components isomorphic to T in the random graph $R_n \in_u \mathcal{A}$. Recall that (??) states that for any fixed labelled forest $\tilde{F}$ (on a finite set of positive integers)

$$\mathbb{P}(\mathrm{Frag}(R_n) = \tilde{F}) \sim \mathbb{P}(R_n \text{ is connected}) \cdot 2^j (en)^{-t} \tag{4.15}$$

and by part (a)

$$\mathbb{P}(R_n \text{ is connected}) \to \frac{1}{e} \quad \text{for } n \to \infty. \tag{4.16}$$

Applying (4.15),(4.16) and (4.14), we have for n large

$$\begin{aligned}
&\mathbb{P}(\kappa(R_n, T) = k_T \,\forall\, T \in S,\ \kappa(\mathrm{Frag}(R_n), T') = 0 \,\forall\, T \in \mathcal{T} \setminus S) \\
&= \mathbb{P}(\mathrm{Frag}(R_n) \cong F) \\
&\sim \prod_{T \in S} \left(\binom{n}{v(T)} \frac{v(T)!}{\mathrm{aut}(T)} \right)^{k_T} \frac{1}{k_T!} \cdot \left(\frac{en}{2} \right)^{-v(T)k_T} \cdot e^{-\sum_{T \in \mathcal{T}} \lambda(T)} \\
&\sim \prod_{T \in S} \left(\left(\frac{n^{v(T)}}{\mathrm{aut}(T)} \left(\frac{2}{en} \right)^{v(T)} \right)^{k_T} \frac{1}{k_T!} e^{-\lambda T} \right) \cdot \prod_{T' \notin S} e^{-\lambda(T')} \\
&\sim \prod_{T \in S} e^{-\lambda(T)} \frac{\lambda(T)^{k_T}}{k_T!} \cdot \prod_{T' \notin S} e^{-\lambda(T)}
\end{aligned}$$

Next, we are going to use this result. Let $m(T)$ be a non-negative integer for each $T \in B'_m$, and let E_n^2 be the event that $\kappa(R_n, T) = m(T)$ for each $T \in B'_m$. Let E_n^3 be the event that $\kappa(\mathrm{Frag}(R_n), T) = 0$ for each $T \in \overline{B_m}$. From the previous result we know that

$$\mathbb{P}\left(E_n^1 \wedge E_n^2 \wedge E_n^3\right) \to p \cdot \prod_{T \in B'_m} e^{-\lambda(T)} \frac{\lambda(T)^{m(T)}}{m(T)!} \prod_{T \in \overline{B_m}} e^{-\lambda(T)} \quad \text{as } n \to \infty.$$

Summing over the finite set of values $m(T) = 0, 1, \ldots, b$ for $T \in B'_m$ we see that, as $n \to \infty$

$$\mathbb{P}\left(E_n^1 \wedge (\kappa(R_n, T) \le b \,\forall T \in B'_m) \wedge E_n^3\right) \to p \cdot \prod_{T \in B'_m} \mathbb{P}(\mathrm{Po}(\lambda(T) \le b) \prod_{T \in \overline{B_m}} e^{-\lambda(T)}.$$

Call the left side x_n and the right side p'. Then by our choice of k and b we have $x_n \le \mathbb{P}(E_n^1) \le x_n + 2\epsilon$, and

$$p' \le p \le p'(1+\epsilon)^2 \le p' + 3\epsilon.$$

Hence

$$|\mathbb{P}(E_n^1) - p| \le |x_n - p'| + 3\epsilon < 4\epsilon$$

for n sufficiently large, as required.

Proof of part (c)
Proof of part (c) using part (b). Let $N_k = N_k(G)$ be the number of components of size k in a graph G. Let S_n be define as in the proof of part (a).

We show that the contribution to $\mathbb{E}[\text{frag}(S_n)]$ from large components is negligibly small. In particular we prove the following claim.

Claim For any $\epsilon > 0$ there is a k_1 such that for all n

$$\sum_{k_1 \le k \le n/2} k\,\mathbb{E}[N_k(S_n)] < \epsilon. \tag{4.17}$$

Then by part (b), for each unlabelled tree T

$$\mathbb{E}[v(T)\kappa(S_n, T)] \to 2\frac{v(T)}{e^{v(T)}\text{aut}(T)} \quad \text{as } n \to \infty.$$

Summing up contributions from the finite set $\mathcal{U}^1$ of unlabelled trees T with size less than k_1,

$$\mathbb{E}[\sum_{T\in\mathcal{U}^1} v(T)\kappa(S_n, T)] \to 2\sum_{T\in\mathcal{U}^1} \frac{v(T)}{e^{v(T)}\text{aut}(T)}.$$

But since the generating function of rooted labelled trees takes value 1 at $1/e$ (see Section 2.3.1), we have

$$\sum_{T\in\mathcal{U}} \frac{v(T)}{e^{v(T)}\text{aut}(T)} = \sum_{T\in\mathcal{T}} \frac{v(T)}{e^{v(T)}v(T)!} = 1.$$

Hence for all ϵ there exists a k_1 such that

$$\mathbb{E}[\text{frag}(S_n) \ge \mathbb{E}[\sum_{T\in\mathcal{U}^1} v(T)\kappa(S_n, T)] \ge 2\sum_{T\in\mathcal{U}_1} \frac{v(T)}{e^{v(T)}\text{aut}(T)} = 2(1-\epsilon)$$

and therefore

$$\liminf_{n\in\mathbb{N}} \mathbb{E}[\text{frag}(S_n) \ge 2.$$

On the other hand by the claim (4.17) for all $\epsilon > 0$ we can choose k_1 such that

$$\mathbb{E}[\text{frag}(S_n) \le \mathbb{E}[\sum_{T\in\mathcal{U}^1} v(T)\kappa(S_n, T)] + \epsilon.$$

Hence

$$\mathbb{E}[\text{frag}(S_n)] \to 2 \quad \text{as } n \to \infty.$$

Thus it remains to establish the claim (4.17).

We need the following observation. Let $f(x) = x^{n-x-1}(n-x)^{x-1}$ for $1 \leq x \leq n-1$. Clearly $f(x) = f(n-x)$. Also $\log f(x) = (n-x-1)\log x + (x-1)\log(n-x)$, with derivative

$$\begin{aligned} & \frac{n-x-1}{x} - \log x - \frac{x-1}{n-x} + \log(n-x) \\ = \quad & \frac{(n-x)(n-x-1) - x(x-1)}{x(n-x)} + \log(n-x) - \log x \end{aligned}$$

which is positive for $1 < x < n/2$. Thus $f(x)$ is maximised at $x = n/2$, with value $(n/2)^{n-2}$. So

$$x^{n-x-1}(n-x)^{x-1} \leq (n/2)^{n-2} \quad \text{for } 1 \leq x \leq n-1. \tag{4.18}$$

We upper bound $\mathbb{E}[N_k(S_n)]$ by comparing it to $\mathbb{E}[N_k(R_n)]$, where $R_n \in_u \mathcal{F}$ (forests in K_n). Note that by Cayley's counting formula and Theorem 4.2.1

$$\mathbb{E}[N_k(R_n)] \geq \binom{n}{k} k^{k-2} \frac{(n-k)^{n-k-2}}{n^{n-2}} \frac{|\mathcal{T}_n|}{|\mathcal{F}_n|} \geq \binom{n}{k} k^{k-2} \frac{(n-k)^{n-k-2}}{n^{n-2}} \cdot \frac{1}{e}.$$

We now want to use this bound.

We may write $\mathbb{E}[N_k(S_n)]$ as a multiple sum, which starts with summing over k-sets $S \subseteq [n]$, then trees T on S, then possible partitions of $n' = n_1 + n_2$ of $[n] \setminus S$ and then counting possible forests on K_{n_1,n_2}. We claim that up to constant it is enough to count trees in K_{n_1,n_2}. Fix a k-set $S \subseteq [n]$ and a tree T on S. T gives an (unordered) partition of S into stable sets S_1 and S_2. Let $\mathcal{C}_k$ be the set of all pairs consisting of a tree T on S and a proper 2 colouring of T. Fix $0 < \epsilon < \frac{1}{2}$ and let $\mathcal{C}'_k$ be the set of all pairs in $\mathcal{C}_k$ where some colour set has size at least $(1+\epsilon)k/2$. Let $\tilde{\mathcal{C}}_k$ be $\mathcal{C}_k \setminus \mathcal{C}'_k$. We claim that $|\mathcal{C}'_k| \leq \tilde{\mathcal{C}}_k$ for k large and hence

$$|\mathcal{C}_k| \leq 2\tilde{\mathcal{C}}_k \text{ for all sufficiently large } k. \tag{4.19}$$

To see this, observe that applying (4.18) together with standard bounds for sums of binomial coefficients gives

$$\sum_{r=1}^{\lfloor (1-\epsilon)k/2 \rfloor} \binom{k}{r} r^{k-r-1}(k-r)^{r-1} \leq 2^{ck} \left(\frac{k}{2}\right)^{k-2}$$

with $c < 0$ which is still exponentially smaller than $\binom{k}{k/2}\left(\frac{k}{2}\right)^{k-2}$. We are interested only in large k (at least some large constant to be chosen below) so by (4.19), rather than considering $\mathcal{C}_k$ we may restrict our attention to $\tilde{\mathcal{C}}_k$. Let now $0 < \epsilon \leq \frac{1}{4}$, $n^{'} = n - k$ and $|S_i| \leq \frac{(1+\epsilon)k}{2}$. Then $n_i = \frac{n}{2} - |S_i|$ and it follows that $n_i \leq (1+\epsilon)\frac{n^{'}}{2}$. To see this observe that $n_1 > \frac{(1+\epsilon)(n-k)}{2}$ together with $\epsilon \geq 2\epsilon k$ would imply that

$$n_2 < \frac{(1-\epsilon)(n-k)}{2} = \frac{n-k-\epsilon k-\epsilon n+2\epsilon k}{2} < \frac{n}{2} - \frac{(1+\epsilon)k}{2}.$$

We claim that the graph K_{n_1,n_2} is still an $\frac{1}{4}$-edge-expander. To see this let A be a set of $a \leq n^{'}/2$ vertices, with a_1, a_2 vertices in the 2 parts. Then

$$\begin{aligned} e(A,\bar{A}) &= \sum_i a_i(n-a-(n_i-a_i)) \\ &= a(n-a) - \sum_i a_i(n_i-a_i) \\ &\geq a(n-a) - a\cdot\max_i n_i + \sum_i a_i^2 \\ &\geq a(n-a) - a(1+\epsilon)\frac{n}{2} + \frac{a^2}{2} \\ &= a\left((n-a) - \epsilon\frac{n}{2}\right) \\ &\geq (1-2\epsilon)\cdot a(n-a) \end{aligned}$$

since $\epsilon\frac{n}{2} \leq 2\epsilon\frac{n-a}{2}$. Hence, Theorem 4.2.1 applies and

$$\mathbb{P}(R_{n'} \text{ is connected}) \geq \rho. \tag{4.20}$$

We are now back to counting the individual sums in $\mathbb{E}[N_k(S_n)]$.

We want to sum over all partitions $R_1 \cup R_2$ of $[n] \setminus S$ such that $|S_1| + |R_1| = \lfloor n/2 \rfloor$ and $|S_2| + |R_2| = \lceil n/2 \rceil$. Note that the number of choices for this partition is at most $\binom{n-k}{\lfloor (n-k)/2 \rfloor}$. Also, given the partition, the number of trees on $R_1 \cup R_2$ is

$$|R_1|^{|R_2|-1}|R_2|^{|R_1|-1} \leq \left(\frac{n-k}{2}\right)^{n-k-2}$$

by (4.18).

By (4.20)

$$\begin{aligned}\mathbb{E}[N_k(S_{\frac{n}{2}})] &\leq O(1)\frac{\binom{n}{k}k^{k-2}\binom{n-k}{\lfloor (n-k)/2\rfloor}\left(\frac{n-k}{2}\right)^{n-k-2}}{\binom{n}{\lfloor n/2\rfloor}\left(\frac{n}{2}\right)^{n-2}} \\ &= O(1)\binom{n}{k}k^{k-2}\frac{\Theta(n^{-\frac{1}{2}})2^{n-k}\left(\frac{n-k}{2}\right)^{n-k-2}}{\Theta(n^{-\frac{1}{2}})2^{n}\left(\frac{n}{2}\right)^{n-2}} \\ &= \Theta(1)\cdot\binom{n}{k}k^{k-2}\frac{(n-k)^{n-k-2}}{n^{n-2}}.\end{aligned}$$

Hence there is a constant c_0 such that

$$\mathbb{E}[N_k(S_n)] \leq c_0 \cdot \mathbb{E}[N_k(R_n)]$$

for all n and all $k = 1, \ldots, n$. But for any $\epsilon > 0$ there is a k_0 such that for all n we have

$$\sum_{k_0 \leq k \leq n/2} k\,\mathbb{E}[N_k(R_n)] < \epsilon$$

(see for example [74, p.9]) and (4.17) follows.

□

Theorem 4.2.1 says that, given any relatively bridge-addable class $\mathcal{A}$ of graphs in $K_{n/2,n/2}$, for $R_n \in_u \mathcal{A}$

$$\mathbb{P}(R_n \in_u \mathcal{A}) \geq e^{-2}.$$

However, Theorem 4.3.1 suggests that this bound may not be tight (if relatively bridge-addable classes of graphs behave at all similarly to bridge-addable classes), and we could replace e^{-2} by $e^{-1+o(1)}$. Also $\mathbb{E}(\text{frag}(R_n)) < 4$ by Theorem 4.2.2, but Theorem 4.3.1 suggests that we could replace 4 by $2 + o(1)$.

We think that indeed, analogously to the fully bridge-addable case, Theorem 4.2.1 is not best possible and forests are the relatively bridge-addable class of graphs that are the least likely to be connected. A formal conjecture is given at the end of this chapter.

4.4. Rooted graphs in expander host graphs

In fact, the exact same statement as in Theorem 4.2.1 holds in the rooted case, i.e. when we consider relatively bridge-addable classes in which every component

of every graph has a unique root. As we will see later, having information about the rooted cases allows us to deduce further information about the unrooted case. We say that a connected graph is *rooted* if it has a specified root vertex; and a graph is *rooted* if each component is.

Theorem 4.4.1. *Let $0 < \alpha \leq 1$, let G be an α-edge-expander and let the set $\mathcal{A}^\bullet$ of rooted subgraphs of G be relatively-bridge-addable. Let $R^\bullet \in_u \mathcal{A}^\bullet$. Then*

$$\kappa(R^\bullet) \leq_s 1 + \mathrm{Po}(1/\alpha),$$

and in particular

$$\mathbb{P}(R^\bullet \text{ is connected}) \geq e^{-1/\alpha}.$$

Proof. Let $\mathcal{P}^k$ be the set of pairs $(G^\bullet, e)$ where $G^\bullet$ is a rooted graph on $[n]$ with k components and e is a bridge in $G^\bullet$. Let Q^{k+1} denote the set of pairs $(H^\bullet, uv)$ with $H^\bullet \in \mathcal{A}^\bullet$ where H has $k+1$ components and uv is an ordered pair of vertices such that u is the root of one of its components and v is in a different component. As described in [73] there is an injection from $\mathcal{Q}^{k+1}$ to $\mathcal{P}^k$.
Fix $H \in \mathcal{A}^{(k+1)}$ with components $C_1, \ldots, C_{k+1}$ of size $n_1, \ldots, n_{k+1}$ respectively and let $e_G(u_i, \bar{C}_i)$ denote the number of possible edges from the root u_i of some component C_i going to all other components. Then (summing over all possible rootings in each C_i)

$$\begin{aligned}
&\sum_{u_1 \in C_1} \cdots \sum_{u_{k+1} \in C_{k+1}} \sum_{i=1}^{k+1} e_G(u_i, \bar{C}_i) \\
&= \sum_{i=1}^{k+1} \sum_{u_i \in C_i} e_G(u_i, \bar{C}_i) \cdot \sum_{\substack{\text{rootings of } C_j \\ j \neq i}} 1 \\
&= \sum_{i=1}^{k+1} \sum_{u_i \in C_i} e_G(u_i, \bar{C}_i) \cdot \prod_{j \neq i} n_j \\
&= \sum_{i=1}^{k+1} e_G(C_i, \bar{C}_i) \cdot \prod_{j \neq i} n_j \\
&\geq \sum_{i=1}^{k+1} \alpha \, n_i(n - n_i) \cdot \prod_{j \neq i} n_j \\
&= \alpha \prod_{j=1}^{k+1} n_j \cdot nk \\
&= \prod_{j=1}^{k+1} n_j \cdot \alpha nk.
\end{aligned}$$

So on average a pair $H^\bullet \in \mathcal{Q}$ corresponds to more than $\alpha\, nk$ pairs $(H^\bullet, uv)$ in $\mathcal{P}$. Then

$$A^{(k)\bullet} \cdot (n-k) \geq \sum_{(G^\bullet,e)\in\mathcal{P}^k} 1 \geq \sum_{(H^\bullet,uv)\in\mathcal{Q}^{k+1}} 1 \geq \alpha k \cdot nA^{(k+1)\bullet}.$$

Hence

$$|\mathcal{A}^{(k+1)\bullet}| \leq \frac{n-k}{\alpha kn}|\mathcal{A}^{(k)\bullet}|$$

and so the result follows by applying Lemma 4.1.3. □

Observe that for $G = K_n$ contrary to the unrooted case the bound given in Theorem 4.4.1 is tight and the lower bound $1 + \mathrm{Po}(1)$ is reached by the class of forests of rooted trees (see Section 2.3.1). Furthermore, it was shown in [73, Theorem 2.6] that for $R_n^\bullet$ taken uniformly at random from the class of rooted forests in K_n, $\mathbb{E}[\mathrm{frag}(R_n^\bullet)] \to \infty$. Hence there is (at least for K_n) no analogue for rooted graphs for Theorem 4.2.2 (which bounds $\mathbb{E}[\mathrm{frag}(R_n^\bullet)]$).

As the following result shows the lower bound proved in Theorem 4.4.1 is also tight for rooted graphs in $K_{n/2,n/2}$ and similarly attained by forests of rooted trees. Also in this case $\mathbb{E}[\mathrm{frag}(R_n^\bullet)] \to \infty$ which implies that $C'(\rho) = \infty$.

Proposition 4.4.2. *Consider the complete bipartite graph $K_{n/2,n/2}$, say with parts $\{1,\ldots,n/2\}$ and $\{n/2+1,\ldots,n\}$. Let the random graph $R_n^\bullet$ be sampled uniformly at random from the forests of rooted trees in this graph. Then $\mathbb{P}(R_n^\bullet$ is connected$) \to 1/e^2$ as $n \to \infty$. and indeed $\kappa(R_n^\bullet)$ converges in distribution to $1 + \mathrm{Po}(2)$. In contrast to the unrooted case, $\mathbb{E}[\mathrm{frag}(R_n^\bullet)] \to \infty$.*

Proof. By Lemma 4.3.2 we know that the number of rooted spanning trees $\tilde{\mathcal{T}}_n^\bullet$ in $K_{n/2,n/2}$ is

$$\tilde{\mathcal{T}}_n^\bullet = n^{n-1}/2^{n-2} \tag{4.21}$$

Assume for the moment that the number of (spanning) forests of rooted trees in $K_{n/2,n/2}$ is

$$\tilde{\mathcal{F}}_n^\bullet = \left(1+\frac{n}{2}\right)^{n-2}(n+1) \tag{4.22}$$

and hence

$$\mathbb{P}(R_n \text{ is connected}) = \frac{\tilde{\mathcal{T}}_n^\bullet}{\tilde{\mathcal{F}}_n^\bullet} = \left(\frac{n}{n+2}\right)^{n-1} \cdot \frac{n+2}{n+1} \sim \frac{1}{e^2}. \tag{4.23}$$

It remains to prove (4.22) which can be done using Lagrange inversion but a bijective proof is given here. Let $f(i,j)$ be the number of rooted spanning forests

in a complete bipartite graph with parts of sizes i and j. Let A and B be disjoint sets with $|A| = i+1$ and $|B| = j+1$. Let $\mathcal{T}$ be the set of edge-rooted spanning trees in the complete bipartite graph with parts A and B. Let $\mathcal{F}$ be the set of triples (a, b, F) where $a \in A$, $b \in B$ and F is a rooted spanning forest in the complete bipartite graph with parts $A\backslash a$ and $B\backslash b$. We define a bijection ϕ from $\mathcal{T}$ to $\mathcal{F}$. Start with a $T \in \mathcal{T}$. Let the root edge be ab, delete this edge, and then delete each edge incident with a or b and make the other end a root. This gives $\phi(T) = (a, b, F) \in \mathcal{F}$. This is a bijection. The size of $\mathcal{T}$ is $(i+j+1)(i+1)^j(j+1)^i$ by the counting result in Lemma 4.3.2 and the fact that there are $n-1$ edges in a tree with n vertices. The size of $\mathcal{F}$ is $(i+1)(j+1)f(i,j)$. Hence $f(i,j) = (i+j+1)(i+1)^{j-1}(j+1)^{i-1}$.

Furthermore, by a result of Szwarc and Wintgen (1965) (see [79, p.17]) we have that the number of forests of rooted trees in $K_{n/2,n/2}$ on n vertices with k components and k specified root vertices is

$$F_{n,k} = k\left(\frac{n}{2}\right)^{n-k-1} - \sum_{s=0}^{k} s(k-s)\left(\frac{n}{2}\right)^{n-k-2}.$$

Hence

$$\mathbb{P}(\kappa(R_n) = k) = \frac{\binom{n}{k} F_{n,k}}{\left(1+\frac{n}{2}\right)^{n-2}(n+1)} \to e^{-2}\frac{2^{k-1}}{(k-1)!}.$$

For the result on the Fragment of $R_n^\bullet$ let $1 \le j \le n^{1/3}$. Let T_0 be a rooted subtree in $K_{n/2,n/2}$ with $v(T_0) = j$. Suppose the split is a, b where $a + b = j$. The number of rooted trees T such that $T_0 \cup T$ is in $\tilde{\mathcal{F}}_n^\bullet$ is

$$\begin{aligned}
& (n-j)\left(\frac{n}{2}-a\right)^{\frac{n}{2}-b-1}\left(\frac{n}{2}-b\right)^{\frac{n}{2}-a-1} \\
= \;& (n-j)\left(\frac{n}{2}\right)^{n-j-2}\left(1-\frac{2a}{n}\right)^{\frac{n}{2}-b-1}\left(1-\frac{2b}{n}\right)^{\frac{n}{2}-a-1} \\
\sim \;& n\left(\frac{n}{2}\right)^{n-j-2} e^{-j} \\
= \;& \frac{n^{n-1}}{2^{n-2}}\left(\frac{2}{en}\right)^j \;=\; |\tilde{\mathcal{T}}_n^\bullet|\left(\frac{2}{en}\right)^j
\end{aligned}$$

uniformly over all a, b, j.

As a contribution to $\mathrm{frag}(R_n^\bullet)$ we shall count only forests consisting of one small balanced tree and one other (large) component. In the sums below let j run through the even integers from say $\log n$ to $\sqrt{n}$.

$$\begin{aligned}\mathbb{E}[\mathrm{frag}(R_n^\bullet)] &\geq (1+o(1))|\tilde{\mathcal{F}}_n^\bullet|^{-1}\sum_j \binom{n/2}{j/2}^2 j\cdot|\tilde{\mathcal{T}}_j^\bullet||\,\tilde{\mathcal{T}}_n^\bullet|\left(\frac{2}{en}\right)^j\\ &= \Theta(1)\left(|\tilde{\mathcal{T}}_n^\bullet|/|\tilde{\mathcal{F}}_n^\bullet|\right)\sum_j \frac{(n/2)^j}{(j/2)^{j+1}e^{-j}}\frac{j^j}{2^{j-2}}\left(\frac{2}{en}\right)^j\\ &= \Theta(1)\sum_j j^{-1}\to\infty.\end{aligned}$$

□

Observe that the class of rooted forests in $K_{n/2,n/2}$ is not decomposable as the class is not closed under isomorphisms. Furthermore, the exponential formula does not apply as we are not allowed to relabel components.

Following [71] a graph H is *freely addable* to a class $\mathcal{A}$ of graphs if, given any graph G disjoint from H, the union of G and H is in $\mathcal{A}$ if and only if G is in $\mathcal{A}$. In particular, a class is decomposable if and only if each graph in $\mathcal{A}$ is freely addable to $\mathcal{A}$.

The following Proposition provides a first link between the rooted and the unrooted case for bridge-addable classes of graphs.

Proposition 4.4.3. *Let $\mathcal{A}$ be bridge-addable and smooth, $\rho>0$, and suppose that there is a graph H freely addable to $\mathcal{A}$. Let $R_n\in_u\mathcal{A}$ and let $R_n^\bullet\in_u\mathcal{A}^\bullet$. Then*

$$\liminf_{n\in\mathbb{N}}\frac{\mathbb{P}(R_n\ \textit{connected})}{\mathbb{P}(R_n^\bullet\ \textit{connected})}>1.$$

Proof. Since $\mathcal{A}$ is bridge-addable we may assume that $|V(H)|=h\geq 2$, since if K_1 is freely addable to $\mathcal{A}$ then so is $2K_1$, and thus so is K_2.

Let $f(n)\sim\log n$ say. Let $\tilde{\mathcal{A}}_n$ be the set of graphs G in $\mathcal{A}_n$ with $\mathrm{frag}(G)\leq f(n)$. Then by Markov's inequality and [76, Theorem 2.1] (or equivantely Theorem 4.2.1)

$$|\tilde{A}_n|/|\mathcal{A}_n| \;=\mathbb{P}(\mathrm{frag}(R_n)\leq f(n))=1-\mathbb{P}(\mathrm{frag}(R_n)\geq f(n))\geq 1-\tfrac{2}{f(n)}=1+o(1).$$

Suppose that H is freely addable to $\mathcal{A}$ and let $h = v(H) \geq 2$. Let ρ be the radius of convergence for $\mathcal{A}$, so $0 < \rho < \infty$. Let $\mathcal{B}_n$ be the set of graphs of the form $H \cup G$ where $G \in \mathcal{C}_{n-h}$. Recall that a class is bridge-addable if the corresponding host graph is K_n and K_n is an α-edge-expander with $\alpha = 1$. Using smoothness and then Theorem 4.2.1 we find

$$\begin{aligned}\frac{|\mathcal{B}_n|}{|\mathcal{A}_n|} &\geq \binom{n}{h}\frac{|\mathcal{C}_{n-h}|}{|\mathcal{A}_n|} \\ &\sim \frac{\rho^h}{h!}\frac{|\mathcal{C}_{n-h}|}{|\mathcal{A}_{n-h}|} \\ &\geq \frac{\rho^h}{e \cdot h!}.\end{aligned}$$

Let $\delta_0 = \frac{1}{2}(h-1)\frac{\rho^h}{e \cdot h!}$. Define the *volume of a graph* H with components of sizes $|H_1|, \ldots, |H_k|$ to be $\mathrm{vol}(H) := \prod_i^k |H_i|$. Then $\delta_0 > 0$, and as for n sufficiently large $\mathrm{Frag}(H \cup G) = H$

$$\sum_{G \in \mathcal{B}_n} (\mathrm{vol}(\mathrm{Frag}(G)) - 1) = (h-1)|\mathcal{B}_n| \geq \delta_0 |\mathcal{A}_n|.$$

Let $\delta = \frac{1}{2}\delta_0$. Then

$$\begin{aligned}|\mathcal{A}_n^\bullet| \geq |\tilde{\mathcal{A}}_n^\bullet| &= \sum_{G \in \tilde{\mathcal{A}}_n} \mathrm{big(G)vol(Frag(G))} \\ &\geq (n - f(n)) \sum_{G \in \tilde{\mathcal{A}}_n} \mathrm{vol}(\mathrm{Frag}(G)) \\ &\geq (n - f(n)) \left(|\tilde{\mathcal{A}}_n| + \sum_{G \in \mathcal{B}_n} (\mathrm{vol}(\mathrm{Frag}(G)) - 1) \right) \\ &\geq (1 + o(1))n \left(|\mathcal{A}_n| + \delta_0 |\mathcal{A}_n| \right) \\ &\geq (1 + \delta) n |\mathcal{A}_n|\end{aligned}$$

for n sufficiently large. But then

$$\frac{|\mathcal{C}_n^\bullet|}{|\mathcal{A}_n^\bullet|} \leq \frac{n|\mathcal{C}_n|}{(1+\delta)n|\mathcal{A}_n|}$$

and so

$$\mathbb{P}(R_n \text{ connected}) \geq (1 + \delta)\mathbb{P}(R_n^\bullet \text{ connected}).$$

□

4.5. Remarks

Looking at forests of rooted and unrooted trees in K_n, it seems as if even a non-asymptotic analogue to Proposition 4.4.3 should hold. Let R_n be sampled uniformly at random from the set of forests of trees in K_n and let $R_n^\bullet$ be sampled uniformly at random from the set of forests of rooted trees in K_n. It seems that

$$\frac{\mathbb{P}(R_n \text{ connected})}{\mathbb{P}(R_n^\bullet \text{ connected})} > 1 \text{ for } n \geq 7.$$

See also Figure 4.2.

Also, the question remains whether there is an analogue of Proposition 4.4.3 for relatively bridge-addable classes of graphs.

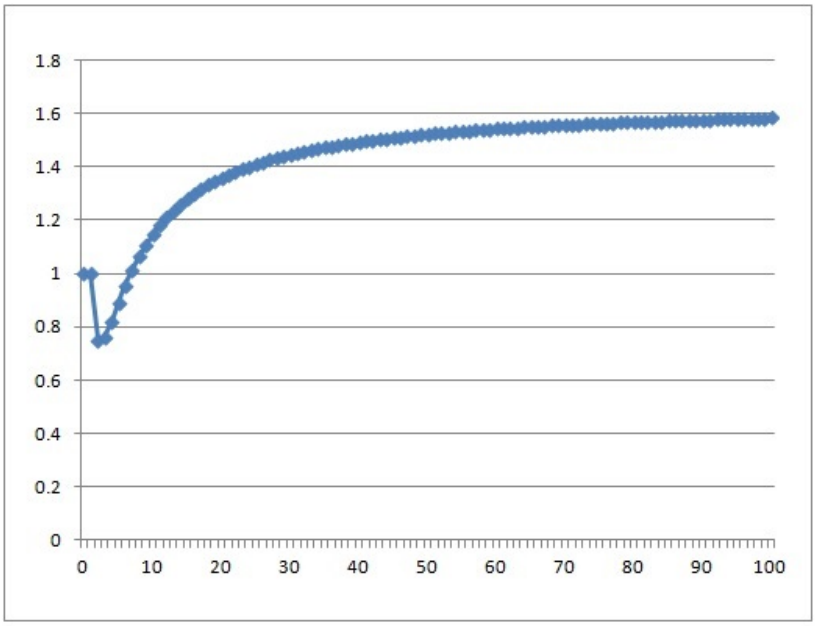


Figure 4.2.: Let R_n be a random forest of unrooted trees and $R_n^\bullet$ a random forest of rooted trees. The graph shows $\mathbb{P}(R_n \text{ connected})/\mathbb{P}(R_n^\bullet \text{ connected})$ for $n = 1 \ldots 100$. As expected the ratio tends to $e^{1/2}$ as n tends to ∞.

We observe that the existence of an equivalent result would imply that the (asymptotic) bound of Theorem 4.2.1

$$\liminf_{n\in\mathbb{N}} \mathbb{P}(R_n \text{ is connected}) \geq e^{-\alpha}$$

cannot be tight for unrooted, relatively bridge-addable classes that fulfil the assumptions as of Proposition 4.4.3 as by Theorem 4.4.1 the exact same bound holds in the rooted case. In fact, as mentioned before we think that indeed, analogously to the fully bridge-addable case, Theorem 4.2.1 is not best possible and forests are the relatively bridge-addable class of graphs that are the least likely to be connected.

Conjecture 4.5.1. *Let G be a finite connected graph, let $\mathcal{A}$ be relatively bridge-addable with respect to G, and let $\mathcal{F}$ be the class of forests in G. For $R \in_u \mathcal{A}$ and $F \in_u \mathcal{F}$*

$$\mathbb{P}(R \textit{ is connected}) \geq \mathbb{P}(F \textit{ is connected})$$

and indeed

$$\kappa(R) \leq_s \kappa(F).$$

5. Forests with a stable set

5.1. Forests with a stable set – a partially bridge-addable class

5.1.1. The setup

Theorem 2.1 in [76] states that the probability for any bridge-addable class $\mathcal{A}$ the probability that $R \in_u \mathcal{A}$ is connected is bounded away from 0. [1] and [62] independently showed that under the slightly stronger assumption of having a bridge-alterable class, i.e. the class is also closed under deletion of bridges, forests are the least likely to be connected. So the probability of a graph drawn uniformly at random from any bridge-addable labelled class to be connected is always at least the probability to pick a tree out of the set of all forests. What happens if we replace 'bridge-addable' by something weaker? Does the behaviour change dramatically? A natural class of graphs to look at seems to be forests where some vertices belong to a specified and fixed stable set, so we can think of them as being coloured. For a random coloured forest there are some pairs of vertices u and v in different components such that the graph obtained by adding the edge uv is still in the class but obviously not all such pairs are allowed. Hence we can think of the class as being 'partially-bridge-addable', and the extent to which this happens is controlled by the parameter y, the weighting parameter of the vertices belonging to the stable set. Let

$$\mathcal{F}_{n,s} = \{(F, S)| \text{ F is a forest on } [n] \text{ vertices and S is a stable set in F with } |S| = s\}$$

and $F_{n,s} = |\mathcal{F}_{n,s}|$.
Let $F(x, y) = \sum_{n\geq 1, s\geq 0} F_{n,s} \frac{x^n}{n!} y^s$ be the corresponding generating function, where x counts the number of vertices and y marks the number of coloured vertices

in a forest of size n. Let (F_n, S) denote the random forests, sampled with with probability proportional to $y^{|S|}$, i.e. with probability

$$\frac{y^{|S|}}{n![x^n]F(x,y)},$$

from all forests on n vertices. Let $U(x,y)$ be the generating function of unrooted trees with a stable set. Then the probability that we pick a connected forest, i.e. that F_n is a tree, is

$$\frac{[x^n]U(x,y)}{[x^n]F(x,y)}.$$

The aim is to analyse the limiting probability

$$\begin{aligned}\lim_{n\to\infty}\frac{U_n}{F_n} &= \lim_{n\to\infty}\frac{U_n(y)}{F_n(y)}\\ &= \lim_{n\to\infty}\frac{n![x^n]U(x,y)}{n![x^n]F(x,y)} = \lim_{n\to\infty}\frac{[x^n]U(x,y)}{[x^n]F(x,y)}\end{aligned}$$

dependent on y. As in the case with no stable set it is convenient to start with rooted trees. Let $T_{n,k}$ denote the number of rooted trees on n vertices with k of them belonging to a specified stable set. Then the generating function of rooted trees with a stable set $\mathcal{T}$ is

$$T(x,y) = \sum_{n\geq 1, s\geq 0} T_{n,k}\frac{x^n}{n!}y^k.$$

Then

$$T(x,y) = T_0(x,y) + T_1(x,y)$$

where $T_0(x,y)$ denotes the generating function of rooted trees where the root is not in the stable set and $T_1(x,y)$ the generating function of rooted trees where the root belongs to the stable set. Similar to rooted trees with no stable set the classes $\mathcal{T}_0$ and $\mathcal{T}_1$ are defined by a symbolic equation.

$$\mathcal{T}_0 = \mathcal{Z} \times SET(\mathcal{T}) \tag{5.1}$$

translates the fact that the class of all rooted trees with a stable set with a root not belonging to the stable set can be constructed by taking one vertex (the root) and any set of forests of rooted trees with stable set and connecting their roots to the new root. The set construction translates the fact that subtrees stemming from the root are not ordered between themselves. Similar for $\mathcal{T}_1$

$$\mathcal{T}_1 = \mathcal{Z} \times \mathcal{Y} \times SET(\mathcal{T}_0) \tag{5.2}$$

holds. Trees in $\mathcal{T}_1$ have a root belonging to the stable set and therefore rooted subtrees stemming from the root must have a root which does not belong to the stable set. Applying the symbolic method see Chapter 2 the function $T(x, y)$ has to satisfy the following implicit equation

$$T(x, y) = xe^{T(x,y)} + xye^{xe^{T(x,y)}}. \tag{5.3}$$

5.2. Rooted trees with a stable set

To begin with, we count rooted trees with a stable set using a bijective proof. The same result can also be obtained by Lagrange inversion.

Lemma 5.2.1. *The number of rooted trees on n vertices with j of them forming a specified stable set is equal to*

$$\binom{n}{j} n^{n-j}(n-j)^{j-1}. \tag{5.4}$$

Hence for the generating function of rooted trees with a stable set, $T(x, y)$, we have

$$[x^n y^j]T(x, y) = \frac{n^{n-j}(n-j)^{j-2}}{j!(n-j-1)!}. \tag{5.5}$$

Proof. We are going to adapt Joyal's bijection (see [3, p.177]). We want to prove (5.5) by constructing a bijection between rooted trees with a specified stable j-set and a marked vertex which does not belong to the stable set and the set of functions $\{f | f : \{1, \dots n\} \to \{1, \dots n\}\}$ equipped with the same specified j-stable set, i.e. some states in $\{1, \dots n\}$ are coloured and a coloured state can only be mapped to an uncoloured state. We first want to map such a function to a tree.

Pick a function f. Think of it as in oriented graph, i.e. the functional graph associated with the function. Then

- Each component of the oriented graph has to have a cycle.
- Each cycle has a smallest non-coloured vertex, called the end-vertex.

Order all components, so that we have an increasing sequence according to the labels of their end-vertex. We "open" each cycle by deleting the edge after the end-vertex. Call this first vertex on the path the start vertex. Concatenate 2 cycles which are consecutive elements in our sequence of components by putting an edge between the end-vertex of the component with the smaller end-vertex and the start-vertex of the other one. Delete all orientations. Call the very first start-vertex the root (either black or white). Call the end-vertex of the last cycle in the sequence the marked vertex (necessarily uncoloured). We constructed a rooted tree with a stable set and an uncoloured marked vertex.

Pick a rooted tree with a stable set and a marked vertex that is not in the stable set.
Oriente the edges "towards" the root. Look at the unique maximal path from the marked vertex to the root. Whenever there is a noncoloured vertex with a smaller label than the previous one, delete the edge before this vertex. Put an edge between the last vertex considered and the noncoloured vertex with the smallest label in that piece. Orient this edge towards the uncoloured vertex.
Then the second mapping is the inverse of the first one. Hence we constructed a bijection.

We can now count the considered graphs.

- Choose j states out of $\{1, \dots, n\}$ and colour them. Gives $\binom{n}{j}$ choices.
- Each noncoloured state can be connected to every other state. Gives n^{n-j} choices.
- Each coloured state can be connected to every non-coloured state. Gives $(n-j)^j$ choices.

Hence, the number of rooted trees with a stable set is

$$\frac{\binom{n}{j} n^{n-j}(n-j)^j}{(n-j)} = \binom{n}{j} n^{n-j}(n-j)^{j-1}.$$

□

Corollary 5.2.2. *Let $\mathcal{T}_{n,j}^{marked}$ denote the set of rooted trees, with j marked vertices. Then*

$$\mathbb{P}(R_n \in \mathcal{T}_{n,j}^{marked},\ j \text{ vertices form a stable set}) = \left(\frac{n-j}{n}\right)^{j-1}. \tag{5.6}$$

Proof. Let the number j of marked vertices be fixed. Then

$$\begin{aligned}
&\mathbb{P}(R_n \in \mathcal{T}_{n,j}^{\text{marked}},\ j \text{ vertices form a stable set}) \\
=&\frac{\#\text{rooted trees, } j \text{ vertices form a stable set}}{\#\text{ rooted trees with } j \text{ marked vertices}} \\
=&\frac{\binom{n}{j} n^{n-j}(n-j)^{j-1}}{\binom{n}{j} n^{n-1}} \\
=&\left(\frac{n-j}{n}\right)^{j-1}.
\end{aligned}$$

□

Lemma 5.2.3. *Let $\mathcal{RF}$ be the class of forests of rooted trees with a stable set. The number of forests in $\mathcal{RF}$ on n vertices with a stable set of size j is*

$$\binom{n+1}{j}(n+1)^{n-j}(n+1-j)^{j}. \tag{5.7}$$

The probability that (RF_n, S) sampled with probability $\frac{y^{|S|}}{n![x^n]RF(x,y)}$ is connected, i.e. that RF_n is a tree, is

$$\frac{[x^n y^j]T(x,y)}{[x^n y^j]RF(x,y)} = \left(\frac{n}{n+1}\right)^{n-j}\left(\frac{n-j}{n+1-j}\right)^{j-1}, \tag{5.8}$$

where $RF(x,y)$ denotes the generating function of forests of rooted trees with a stable set.
In the case with no stable set ($j = 0$) this probability is equal to

$$\left(\frac{n}{n+1}\right)^{n} \longrightarrow \frac{1}{e} \tag{5.9}$$

as $n \to \infty$ as expected.

Proof. Let $\mathcal{RF}$ denote forests of rooted trees with a stable set. Then $\mathcal{RF} = SET(\mathcal{T})$ and hence

$$\begin{aligned}
RF(x,y) &= e^{T(x,y)} \\
T_0(x,y) &= x \cdot e^{T(x,y)} = x \cdot RF(x,y).
\end{aligned}$$

The number of rooted trees on n vertices, j of them belonging to the stable set and the root does not belong to the stable set is equal to

$$n![x^n]T_0(x,y) = \frac{n-j}{n} \cdot \binom{n}{j} n^{n-j}(n-j)^{j-1}.$$

Furthermore,

$$[x^n y^j]RF(x,y) = [x^n y^j]\frac{T_0(x,y)}{x} = [x^{n+1}y^j]T_0(x,y).$$

and hence

$$\begin{aligned} n![x^n y^j]RF(x,y) &= n![x^{n+1}y^j]T_0(x,y) \\ &= \frac{n+1-j}{n+1}\cdot\binom{n+1}{j}(n+1)^{n+1-j}(n+1-j)^{j-1} \\ &= \binom{n+1}{j}(n+1)^{n-j}(n+1-j)^j. \end{aligned}$$

We refer to Section 2.3.1 for results on forests of rooted trees with no stable set. □

5.3. Results on unrooted trees with a stable set

We are interested in forests of unrooted trees with a stable set. Hence we need to find a relation between rooted and unrooted trees with a stable set.

Lemma 5.3.1. *Let $T(x,y)$ denote the generating function of the rooted trees with a stable set and $U(x,y)$ the generating function of unrooted trees with a stable set. Then*

$$U(x,y) = T(x,y) - \left(\frac{(T(x,y))^2}{2} - \frac{(T_1(x,y))^2}{2}\right) \tag{5.10}$$

where $T_1(x,y)$ is the generating function of trees with a stable set where the root is in the stable set as defined in (5.2).

Proof. We proceed similarly to unrooting trees with no stable set, see Section 2.3.1. Each unrooted tree with a stable set is counted n times in $T(x,y)$. On the other hand $\frac{(T(x,y))^2}{2}$ counts the number of unordered pairs of rooted trees with a stable set. Thus

$$E(x,y) = \frac{(T(x,y))^2}{2} - \frac{(T_1(x,y))^2}{2}$$

counts the number of unordered pairs of trees with a stable set where at most one of the roots of the trees in the pair belongs to the stable set. We can now think of those pairs as edge rooted trees with a stable set as we can simply insert an edge between the two roots. This is clearly possible as at most one of the roots is in the stable set. Thus $E(x,y)$ counts the number of edge rooted trees with a stable set, and every such unrooted tree is counted $n-1$ times. Together with the first observation the result follows. □

Applying the SET construction see Chapter 2 we obtain that the generating function of forests of trees where some vertices belong to a stable set is

$$F(x,y) = e^{T(x,y) - \left(\frac{(T(x,y))^2}{2} - \frac{(T_1(x,y))^2}{2}\right)}. \tag{5.11}$$

The next step now is to analyse (5.11).

Take y as fixed for the moment. Then $T(x,y)$ fulfils the *smooth implicit function schema* ([49, page 467]). To see this recall that by (6.7)

$$T(x,y) = xe^{T(x,y)} + xye^{xe^{T(x,y)}}. \tag{5.12}$$

Hence

$$T(x,y) = G(x, T(x,y)) \tag{5.13}$$

with $G(x,w) = xe^w + xye^{xe^w}$ and the following conditions hold. To simplify the notation we write $\frac{\partial}{\partial w}G(x,w) = G_w(x,w)$, $\frac{\partial}{\partial x}G(x,w) = G_x(x,w)$ and $\frac{\partial}{\partial w}G_w(x,w) = G_{ww}(x,w)$.

Analyticity $G(x,w)$ is analytic in a domain $|x| < R$ and $|w| < S$ for some $R, S > 0$.

Non-linear dependency The coefficients satisfy $[x^m w^n]G(x,w) \geq 0$, $[x^0 w^0]G(x,w) = 0$, $[x^0 w^1]G(x,w) > 1$ and they are eventually positive

Characteristic system There exist two real-valued numbers r, s satisfying the systems of equations

$$G(r,s) = s \qquad G_w(r,s) = 1. \tag{5.14}$$

The conditions on the coefficients ensure that the problem is normalized, irreducible (not of the form $T(x,y) = T(x,y) + ...$) and that the dependency on $T(x,y)$ is non-linear. Then a slightely modified Theorem due to Meir and Moon [78] – for our purpose we need to push the expansion further – states that $T(x,y)$ has a square-root singularity at $x = r$. We also refer to the standard book on trees by Drmota [38] for this and further results.

Theorem 5.3.2 (Extension of smooth implicit function schema [49])**.** *Let y be fixed. Let $F(x,y)$ be a bivariate exponential generating function that belongs to the smooth-implicit function schema, defined by $G(x,w)$ with (r,s) the positive*

solution of the characteristic system. Then $F(r, y) < \infty$, where $r(y)$ is the radius of convergence of $F(x, y)$ parametrized by y and

$$F(x,y) = s(y) - \gamma(y) \cdot \sqrt{1 - \frac{x}{r(y)}} + c(y)\left(1 - \frac{x}{r(y)}\right) + d(y)\left(1 - \frac{x}{r(y)}\right)^{\frac{3}{2}} + O\left(\left(1 - \frac{x}{r(y)}\right)^{2}\right)$$

as x tends to r with

$$\gamma = \sqrt{\frac{2rG_x(r,s)}{G_{ww}(r,s)}}$$
$$c = \frac{(-\gamma^2 G_{www} + 6\, rG_{xw})}{6G_{ww}}$$
$$d = \frac{(-5\,\gamma^4 G_{www}^2 + 36\,\gamma^2 rG_{www}G_{xw} - 36\,r^2G_{xw}^2 - 36\,rG_{xww}\gamma^2 G_{ww} + 36\,G_{xx}r^2G_{ww})}{72\gamma\, G_{ww}^2}$$

and the expansion is valid in a Δ-domain.

Proof. We need to adapt the proof of Theorem VII.3 in Flajolet and Sedgewick [49]. More precisely, we need to get need further terms in the expansion of Lemma VII.3. Locally, close to (r, s) the function $G(z, w) = w - F(z, w)$ behaves like

$$\begin{aligned} G(z,w) = {} & G(r,s) + (z-r)G_z(r,s) + (w-s)G_w(r,s) \\ & + \frac{1}{2}\Big((z-r)^2G_{zz}(r,s) + 2(z-r)(w-s)G_{zw}(r,s) + (y-s)^2G_{ww}(r,s)\Big) \\ & + \frac{1}{3!}(w-s)^3G_{www}(r,s). \end{aligned}$$

With $G(r, s) = G_w(r, s) = 0$ this simplifies the expression above to

$$\begin{aligned} G(z,w) = {} & (z-r)G_z(r,s) \\ & + \frac{1}{2}\Big((z-r)^2G_{zz}(r,s) + 2(z-r)(w-s)G_{zw}(r,s) + (y-s)^2G_{ww}(r,s)\Big) \\ & + \frac{1}{3!}(w-s)^3G_{www}(r,s). \end{aligned}$$

According to the proof of Lemma VII.3 in Flajolet and Sedgewick [49] we know that a solution around $z = r$ of

$$(z-r)G_z(r,s) + \frac{1}{2}\Big((z-r)^2 G_{zz}(r,s) + 2(z-r)(w-s)G_{zw}(r,s) + (y-s)^2 G_{ww}(r,s)\Big) + \frac{1}{3!}(w-s)^3 G_{www}(r,s) = 0$$

for $(w-s)$ in powers of $(1-z/r)^{1/2}$ exists. The result is then obtained by using Maple with its commands mtaylor and [gfun]algeqtoseries. □

Solving the characteristic system explicitly for r, s is not possible. Instead, we have the following result.

Lemma 5.3.3. *Let $T(x,y)$ be the generating function of rooted trees where some vertices belong to a fixed stable set – x marks the number of vertices, y marks the number of vertices in the stable set. We have*

$$T(x,y) = s + b\cdot\left(1-\frac{x}{r}\right)^{1/2} + c\cdot\left(1-\frac{x}{r}\right) + d\cdot\left(1-\frac{x}{r}\right)^{3/2} + O\left(\left(1-\frac{x}{r}\right)^2\right) \quad (5.15)$$

as x tends to r with

$$r(y) = \frac{1}{\beta+1} e^{-\frac{1}{\beta+1}-\beta}$$
$$s(y) = \frac{1}{\beta+1} + \beta$$

where

$$y = \beta\,(\beta+1)\,\mathrm{e}^{\left(\beta^2+\beta+1\right)\left(\mathrm{e}^{-\frac{\beta^2+\beta+1}{\beta+1}}+\beta+1\right)(\beta+1)^{-2}}.$$

and b, c, d as in Theorem 5.3.2.

In other words, we found a parametrized form of y for some function $\beta = \beta(y)$ and this parametrization of y in terms of β is unique for $y > 0$.

Proof. Let $y \geq 0$. Let

$$f(\beta) = \beta\,(\beta+1)\,\mathrm{e}^{\left(\beta^2+\beta+1\right)\left(\mathrm{e}^{-\frac{\beta^2+\beta+1}{\beta+1}}+\beta+1\right)(\beta+1)^{-2}}.$$

Then $f(0) = 0$ and one can check with Maple that $f'(x) > 0$ holds as well as $\lim_{x\to\infty} f(x) = \infty$. Hence, for each $y > 0$ there exists a unique $\beta(y)$. We have

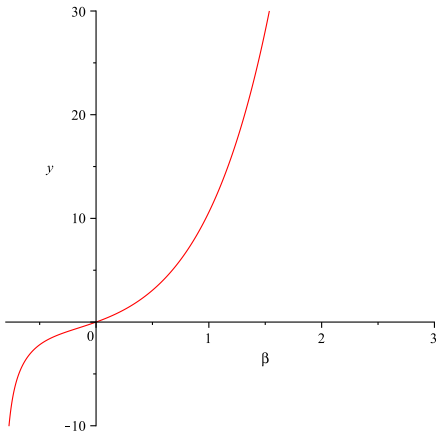


Figure 5.1.: y in dependency of β.

already seen that $T(x, y)$ belongs to the smooth-implicit function schema. In case of $T(x, y)$ the system of characteristic equations is

$$G(r, s) = s, \quad G'_w(r, s) = 1$$

which reads as

$$re^s + rye^{rs} = s \tag{5.16}$$

$$re^s + r^2ye^{s+re^s} = 1. \tag{5.17}$$

Let us eliminate y between these two equations. By taking (5.17)-(5.16)$\cdot re^s$ and writing $re^s := 1/(\beta + 1)$ we obtain

$$\frac{1}{\beta + 1} + \beta = s. \tag{5.18}$$

From which it follows that

$$r = \frac{1}{\beta + 1}e^{-s} = \frac{1}{\beta + 1}e^{-\frac{1}{\beta+1}-\beta}. \tag{5.19}$$

Returning to the very first equation, we obtain

$$y = \beta\,(\beta + 1)\,\mathrm{e}^{\left(\beta^2+\beta+1\right)\left(\mathrm{e}^{-\frac{\beta^2+\beta+1}{\beta+1}}+\beta+1\right)(\beta+1)^{-2}} \tag{5.20}$$

We have thus solved the characteristic system in parametric form: for each $y \geq 0$, there exists a unique $\beta \geq 1$ such that (5.20) holds, since y is a strictly increasing function in β as discussed before. Then one checks that r and s defined by 5.18 and 5.19 satisfy the characteristic system. □

Hence the three interesting cases for us are

1. $y \to \infty$ which corresponds to $\beta \to \infty$. The intuition is that the bigger y the more 'weight' we put on the vertices in the stable set. Hence the less close we are to being bridge-addable.

2. $y = 1$ which corresponds to $\beta \approx 0.3339242037$ and $F(x,1) = \sum_{n\geq 1} \frac{x^n}{n!} \sum_{s\geq 0} F_{n,s}$.

3. $y \to 0$ which corresponds to $\beta \to 0$. In fact $F(x,y) = \sum_{n\geq 1} \frac{x^n}{n!} \left(F_{n,0} + \sum_{s\geq 1} F_{n,s} y^s\right)$, so for $y = 0$ we obtain exactly the EGF of trees with no stable set.

Lemma 5.3.4. *Let $T(x,y)$ be the generating function of rooted trees with a stable set and let $U(x,y)$ be the generating function for unrooted trees with a stable set. Suppose $T(x,y)$ has a complete asymptotic expansion in powers of $\left(1-\frac{x}{r}\right)^{1/2}$ (as in (5.15)). Then $U(x,y)$ admits a complete asymptotic expansion in powers of $\left(1-\frac{x}{r}\right)^{1/2}$. Hence (using the notation of Lemma 5.3.3)*

$$U(x,y) = A + C\cdot\left(1-\frac{x}{r}\right) + D\cdot\left(1-\frac{x}{r}\right)^{3/2} + O\left(\left(1-\frac{x}{r}\right)^2\right)$$

with $A = s - s\cdot re^s + \frac{1}{2}r^2e^{2s}$, $C = -s$ and $D = \frac{-2b}{3}$ as $x \to r$

Proof. The complete expansion of $T(x,y)$ for $x \to r$ in powers of $\left(1-\frac{x}{r}\right)^{1/2}$ follows from Lemma 5.3.3. By Lemma 5.3.1, we only need to prove that $T_1(x,y)$ admits an asymptotic expansion in powers of $\left(1-\frac{x}{r}\right)^{1/2}$ in order to show, the first part of the statement. Observe

$$\begin{aligned} T_1(x,y) &= T(x,y) - T_0(x,y) \\ T_0(x,y) &= x\cdot e^{T(x,y)}. \end{aligned}$$

We know that $T(x,y) \to s$ as $x \to r$. For x close to r, we obtain

$$\begin{aligned} T_0(x,y) =& r\cdot\left(e^s + e^s(T(x,y)-s) + \frac{1}{2}e^s(T(x,y)-s)^2 + \frac{1}{6}e^s(T(x,y)-s)^3\right) \\ &+ O\left((T(x,y)-s)^3\right) \end{aligned}$$

Therefore, by (5.15) $T_0(x,y)$ has a complete asymptotic expansion in powers of $\left(1-\frac{x}{r}\right)^{1/2}$ and so does $U(x,y)$.
As in the uncoloured case, there are n possible ways to root an unrooted tree, hence

$$x\frac{\partial U}{\partial x}(x,y) = T(x,y).$$

We obtain

$$-\left(1-\left(1-\frac{x}{r}\right)\right)\frac{\partial U}{\partial\left(1-\frac{x}{r}\right)}\left(\frac{x}{r},y\right)-T\left(\frac{x}{r},y\right)=0,$$

or equivalently

$$\begin{aligned}&-1/2\,\frac{B}{\sqrt{\left(1-\frac{x}{r}\right)}}-C-s\\&+1/2\,\left(-3\,D+B-2\,b\right)\sqrt{\left(1-\frac{x}{r}\right)}\\&+\left(C-c\right)\left(1-\frac{x}{r}\right)+1/2\,\left(3\,D-2\,d\right)\left(1-\frac{x}{r}\right)^{3/2}\\&+O(\left(1-\frac{x}{r}\right)^{2})=0.\end{aligned}$$

By comparing coefficients, we deduce $B=0$, $C=-s$, $D=\frac{-2b}{3}$. □

We are now able to state and prove our final result.

Theorem 5.3.5 (**Random forests with a stable set).** *Let $\mathcal{F}$ be the class of forests where each tree has a fixed stable set S. Let (F_n,S) with $F\in\mathcal{F}$ and S the stable set of F be picked with probability*

$$\frac{y^{|S|}}{n![x^n]F(x,y)}.$$

Let y be fixed. Then the limiting probability that (F_n,S) is connected, i.e. that F_n is a tree is equal to

$$\rho(y)=\rho_\beta=\lim_{n\to\infty}\frac{U_n}{F_n}=\left(\exp\left(1/2\,\frac{2\,\beta^3+2\,\beta^2+2\,\beta+1}{\left(\beta+1\right)^2}\right)\right)^{-1}>0 \tag{5.21}$$

for fixed β which is implicitly defined by

$$y=\beta\left(\beta+1\right)\mathrm{e}^{\left(\beta^2+\beta+1\right)\left(\exp-\frac{\beta^2+\beta+1}{\beta+1}+\beta+1\right)\left(\beta+1\right)^{-2}}.$$

for $y>0$. Furthermore,

$$\lim_{y\to\infty}\rho(y)=\lim_{\beta\to\infty}\rho_\beta=0. \tag{5.22}$$

Observe that

$$\lim_{y\to0}\rho(y)=\lim_{y\to0}\exp\left(\frac{1}{2}\,\frac{2\,\beta^3+2\,\beta^2+2\,\beta+1}{\left(\beta+1\right)^2}\right)=e^{1/2}.$$

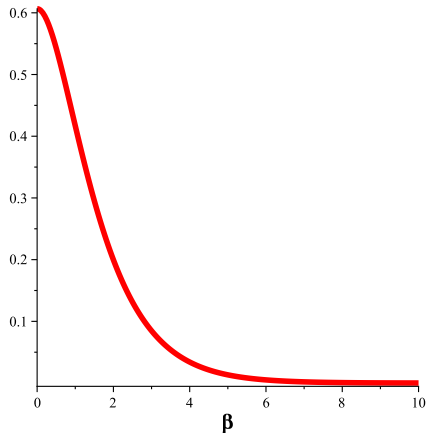


Figure 5.2.: The probability that F is a forests for randomly sampled (F, S) versus β.

Hence we recovered the known result for forests without a stable set (see Section 2.3.1). Furthermore, the result affirms the intuition: the fewer bridges we are allowed to introduce between components (the bigger y gets), the less likely we are to pick a connected graph.

Proof. First of all, by singularity analysis which was described in Chapter 2.3 (Theorem VI.4 (p. 393) and Theorem VI.2 (p. 385) in [49])

$$[x^n]U(x,y) = D \cdot \frac{1}{\pi n^5}\left(\frac{3}{4} + O\left(\frac{1}{n}\right)\right).$$

For $x > 0$ small, we have $e^x \sim 1 + x$. Hence as $U(x,y) \to 0$ for $x \to r$, we get

$$\begin{aligned}
[x^n]e^{U(x,y)} &\sim [x^n]e^{A+C\cdot\left(1-\frac{x}{r}\right)+D\cdot\left(1-\frac{x}{r}\right)^{3/2}+O\left(\left(1-\frac{x}{r}\right)^{\frac{1}{2}}\right)} \\
&\sim [x^n]e^A \cdot \left(1 + C\cdot\left(1-\frac{x}{r}\right) + D\cdot\left(1-\frac{x}{r}\right)^{\frac{3}{2}} + O\left(\left(1-\frac{x}{r}\right)^{\frac{1}{2}}\right)\right) \\
&= e^A \frac{D}{\sqrt{\pi n^5}}\left(\frac{3}{4} + O\left(\frac{1}{n}\right)\right)
\end{aligned}$$

Therefore

$$\lim_{n\to\infty}\frac{U_n}{F_n} = e^{-A} = \left(\exp\left(1/2\,\frac{2\,\beta^3 + 2\,\beta^2 + 2\,\beta + 1}{(\beta+1)^2}\right)\right)^{-1}. \tag{5.23}$$

□

Theorem 5.3.6. *Let $\kappa(F_n)$ denote the number of components of (F_n, S) the random coloured forest that we picked with probability proportional to $y^{|S|}$, i.e. with probability*

$$\frac{y^{|S|}}{n![x^n]F(x,y)}.$$

Then

$$\kappa(F_n) \sim 1 + \mathrm{Po}\left(\frac{1}{2}\,\frac{2\,\beta^3 + 2\,\beta^2 + 2\,\beta + 1}{(\beta+1)^2}\right). \tag{5.24}$$

Proof. Let R_U denote the radius of convergence of $U(x,y)$ for fixed $y > 0$. Then $R_U = r(y)$ and hence $r > 0$ for $y < \infty$. By singularity analysis we know that the coefficients of $U(x,y)$ are smooth and $U(r(y), y) = s(y) < \infty$.
In (5.21) we proved that the limiting probability to be connected exists. Applying Theorem 2.2.3 gives the result. □

6. Smallest component in decomposable structures

6.1. The size of the smallest component in decomposable structures

Suppose we are given a decomposable set of combinatorial objects which are built up of simpler objects called 'components'. Components themselves cannot be further decomposed and might consist of labelled or unlabelled atoms. What is the probability distribution and the expected size of the smallest component? Of course the behaviour might vary depending on the set. As we have seen in previous chapters the behaviour of the generating function next to its radius of convergence might be of interest.

Many combinatorial examples such as trees, permutations, planar graphs have component generating functions which are of *alg-log* type and there is an extensive literature dealing with different aspects of such decomposable structures. The class of *alg-log* component generating functions was introduced in [50] in the course of studying the number of components in combinatorial structures.

Definition 6.1.1 ([50]). *A component generating function $C(z)$ (exponential or ordinary) is of alg-log type at singularity ρ if $C(z)$ is analytic in a Δ-domain at ρ (see Definition 2.3.3) and we have*

$$C(z) = c + d\left(1 - \frac{z}{\rho}\right)^{-\alpha}\left(\log\frac{1}{1 - z/\rho}\right)^{\beta}(1 + o(1)) \tag{6.1}$$

for $z \to \rho$ in this Δ-domain and some $\alpha, \beta, c, d \in \mathbb{R}$. α is called the algebraic exponent, β the logarithmic exponent.

By Theorem 2.3.4 the coefficients of every generating function of alg-log type are smooth. For various values of α, β general results are known about the properties

of $R_n \in_u \mathcal{A}$. In [58] the general framework for the expected size of the largest component in such decomposable structures was developed. In the case $\alpha = 0$, $\beta = 1$ the components are said to be of *logarithmic* type and the set $\mathcal{A}$ is called an *exp-log* class. The number of components in these structures was studied in [50]. In Chapter 5 of [57] the size of the r^{th} largest component and its distribution was obtained. In [5] exp-log classes were studied from a probabilistic point of view. Furthermore, information on exp-log classes with a restricted pattern (see definition below) and the distribution of the size of the smallest component can be found in [82]. The following is a modified version of the definition of a restricted pattern in [37].

Definition 6.1.2. *A* simple restricted pattern *is a mapping S from a finite subset J of positive integers to the non-negative integers. A* restricted pattern *is a sequence of such mapping S_n for each $n = 1, 2, \ldots$. We say that a combinatorial object G has the* simple restricted pattern *S if G has exactly $S(j)$ components of size j for each $j \in J$ (and may have any number of components of other sizes). An n-vertex object has* restricted pattern *S if it has the simple restricted pattern S_n.*

Observe that the corresponding pattern for the smallest component having size at least k is $S(j) = 0$ for $j \leq k - 1$. Several results on the size of the smallest component in decomposable classes were proved (see for example [83] and [18] for related work).

Amongst other results the following theorem was proven (combination of Theorem 3.1 and Theorem 4.1 in [37]). A generating function *algebraic* if it is algebraic-logarithmic in the sense of Definition 6.1.1 and $\beta = 0$, $\alpha < 0$. Similarly, we call a generating function *mixed* if it is algebraic-logarithmic in the sense of Definition 6.1.1 and $0 < \alpha < 1$ and arbitrary logarithmic exponent β.

Theorem 6.1.3 ([37])**.** *Let $\mathcal{A}$ be a labelled decomposable class and let its component generating function $C(z)$, i.e. the generating function of its labelled, connected structures, be either algebraic or mixed. Suppose that in addition the following holds: if $C(z)$ is mixed then $\sum_{j \in J} jS(j) = o(n^{\frac{1}{-\alpha+1}})$, if $C(z)$ is algebraic then $\sum_{j \in J} jS(j) = o(n)$. Define $\lambda_j = \frac{C_j \rho^j}{j!}$. Then as $n \to \infty$*

$$\mathbb{P}(R_n \text{ has restricted pattern } S) \sim \prod_{j \in J} \frac{1}{S(j)!} (\lambda_j)^{S(j)} e^{-\lambda_j}.$$

Let now $\mathcal{A}$ be an unlabelled decomposable combinatorial class. Let $\mathcal{C}$ denote the connected substructures in $\mathcal{A}$ and $C(z)$, $A(z)$ the corresponding ordinary generat-

ing function. Then $\mathcal{A} = MSET(\mathcal{C})$ (see [49, Chapter I, p. 29] for an explanation). For the generating function this translates into

$$A(z) = \sum_{n=0}^{\infty} A_n z^n = e^{\left(C(z)+\frac{C(z^2)}{2}+\frac{C(z^3)}{3}+\ldots\right)}$$

and we recall the proof here for completeness. To simplify we assume that $\mathcal{A}$ is finite. For the limit argument for the infinite case see [49, p.29]. Let $\mathcal{A} = \{\alpha_1, \alpha_2, \ldots\}$ be some listing of the elements in $\mathcal{A}$. A multiset construction can be viewed as a sequence of repeated elements $\alpha_1 \in \mathcal{A}$, a sequence of repeated elements α_2, and so on. By the sequence construction (see Chapter 2.1 for the labelled case, the exact same result holds in the unlabelled case)

$$\begin{aligned} A(z) &= \prod_{\alpha\in\mathcal{A}} (1 - z^{|\alpha|})^{-1} = \prod_{n=1}^{\infty} (1 - z^n)^{A_n} \\ &= \exp\left(\sum_{n=1}^{\infty} A_n \log(1 - z^n)^{-1}\right) \\ &= \exp\left(\frac{A(z)}{1} + \frac{A(z^2)}{2} + \ldots\right). \end{aligned}$$

In [37] the following counter part to Theorem 6.1.3 for unlabelled structures was obtained.

Theorem 6.1.4 ([37])**.** *Let $\mathcal{A}$ be an unlabelled decomposable class and let its ordinary component generating function $C(z)$, i.e. the generating function of its connected structures, be of alg-log type. Suppose that either $C(z)$ is mixed ($\beta = 0$, $\alpha < 0$) or it is algebraic ($0 < \alpha < 1$, $\beta \in \mathbb{R}$).*

Suppose that in addition if $C(z)$ is mixed $\sum_{j\in J} jS(j) = o(n)$ and if $C(z)$ is algebraic $\sum_{j\in J} jS(j) = o(n^{\frac{1}{-\alpha+1}})$. Then as $n \to \infty$

$$\mathbb{P}(R_n \textit{ has restricted pattern } S) \sim \prod_{j\in J} \binom{C_j + S(j) - 1}{S(j)} \rho^{jS(j)} (1 - \rho^j)^{C_j}.$$

The proofs for both Theorem 6.1.3 and Theorem 6.1.4 are based on analysing generating functions. Let $A(z, S)$ be the generating function counting the structures with the restricted pattern S. Then

$$\mathbb{P}(R_n \text{ has restricted pattern } S) = \frac{[z^n]A(z, S)}{[z^n]A(z)}$$

and the coefficients of these generating functions need to be estimated. Depending on whether we are in the mixed or in the algebraic case either singularity analysis

or the saddle point method have to be used. This might be not too surprising as in the mixed case $C(\rho) = c < \infty$ versus in the purely algebraic case $C(\rho) = \infty$.

6.2. The size of the smallest component in general smooth decomposable structures

In the conclusion section of [37] it was remarked that somewhat the authors were surprised to find that among all labelled structures regardless of the nature of $C(z)$, i.e. the ranges of the parameters α, β, they obtained the same expression for the limiting probability $\mathbb{P}(R_n$ has restricted pattern $S)$. The same holds in the unlabelled structures. It therefore seems natural to expect that there should exist a more general version of their results. In the following, decomposable, smooth classes of combinatorial objects will be investigated – regardless of the nature of the singularity. We will see that in the labelled as well as in the unlabelled case, unlabelled finite connected combinatorial objects appear asymptotically independently from each other (at least if they appear only a finite number of times). Using this result will give access to the probability distribution of the size of the smallest component up to a constant. We will need one more definition. For a random variable X and a distribution F we write $X \sim_d F$ to mean that the distribution of X is F.

Definition 6.2.1. *Let $X_1^n, \ldots, X_j^n$ be a set of non-negative integer valued random variables which all depend on some parameter n. We say that $X_1^n, \ldots, X_j^n$ are asymptotically independent with limiting distributions $F_1, \ldots, F_j$ if there exist independent random variables $\tilde{X}_1, \ldots, \tilde{X}_j$ with $\tilde{X}_i \sim_d F_i$ for $1 \leq i \leq j$ such that*

$$(X_1^n, \ldots, X_j^n) \xrightarrow{d} (\tilde{X}_1, \ldots, \tilde{X}_j).$$

The following lemma can be proved by simple but tedious calculation and the proof is omitted here.

Lemma 6.2.2. *Let $X_1^n, \ldots, X_j^n$ be asymptotically independent non-negative integer valued random variables with limiting distributions $F_1, \ldots, F_j$. Let $\tilde{X}_1, \ldots, \tilde{X}_j$ be independent random variables with $\tilde{X}_i \sim_d F_i$ for $1 \leq i \leq j$. Let $S_1, \ldots, S_k$ be disjoint subsets of $[j]$ and define $Y_i^n = \sum_{t \in S_i} X_t^n$. Then $Y_1^n, \ldots, Y_k^n$ are asymptotically independent with limiting distributions $G_1, \ldots, G_k$ where $G_i = \sum_{t \in S_i} F_t$.*

Here $\sum_t F_t$ denotes the distribution of the sum of independent random variables with distributions F_t.

6.2.1. Labelled structures

Let $\mathcal{A}$ be a **labelled** set of combinatorial objects. The following theorem can be found in [76, Lemma 5.3]. In [76] the result was stated for planar graphs, but the proof works for general decomposable and smooth classes of combinatorial objects. Let H be an object in $\mathcal{C}$ and let $\kappa(R_n, H)$ be the random variable that counts the number of components isomorphic to H in $R_n \in_u \mathcal{A}$.

Theorem 6.2.3 ([76])**.** *Let $\mathcal{A}$ be a labelled, decomposable and smooth class so $\mathcal{A} = \mathrm{SET}(\mathcal{C})$ and let $\rho > 0$ denote its radius of convergence. Define $n_i = |V(H_i)|$ for a combinatorial object $H_i \in \mathcal{C}$. For each fixed j let $H_1, \ldots, H_j$ be a collection of pairwise non-isomorphic objects in $\mathcal{C}$. We have for all $i \leq j$ as $n \to \infty$*

$$\kappa(R_n, H_i) \xrightarrow{d} \mathrm{Po}\left(\frac{\rho^{|H_i|}}{aut(H_i)}\right) \tag{6.2}$$

and $\kappa(R_n, H_1), \ldots, \kappa(R_n, H_j)$ are asymptotically independent.

Let $\lambda_{H_i} = \frac{\rho^i}{\mathrm{aut}(H_i}$ and $\mathbb{E}((X)_k)$ the k^{th} factorial moment. The theorem is proved using the method of moments for joint distributions, explained in Chapter 2.5. Recall that the k^{th} factorial moment of a Poisson distribution of parameter λ is equal to λ^k. Similarly, a result which can for example be found in [76, Lemma 5.4.] on the joint distribution of independent Poisson random variables tells us that to show Theorem 6.2.3 it suffices to show the following statement

$$\mathbb{E}[(\kappa(R_n, H_1)_{k_1} \cdots (\kappa(R_n, H_j))_{k_j}] \to \lambda_{H_1}^{k_1} \cdots \lambda_{H_j}^{k_j}.$$

We are interested in the distribution of the number of components of a fixed size.

Proposition 6.2.4. *Let $\mathcal{A}$ be a decomposable, smooth class with radius of convergence $\rho > 0$. Let X_s denote the number of components of size s in $R_n \in_u \mathcal{A}$. Then for each fixed j and all $s \leq j$*

$$X_s \xrightarrow{d} \mathrm{Po}(\lambda_s) \text{ where } \lambda_s = \frac{C_s \rho^s}{s!}$$

and the random variables $X_1, \ldots, X_j$ are asymptotically independent.

Recall that for decomposable, smooth classes with $\rho > 0$, $C(\rho) < \infty$ and where the limit $\lim_{n\to\infty} \mathbb{P}(R \text{ connected})$ exists we know that $\kappa(R_n) \to 1 + \mathrm{Po}(C(\rho))$ and $C(\rho) = \sum_{s=1}^{\infty} \lambda_s$.

Proof. Fix s. Let $\mathcal{U}_s$ denote all unlabelled structures of size s in $\mathcal{C}$ and we can write

$$X_s = \sum_{G \in \mathcal{U}_s} \kappa(R_n, G).$$

By Theorem 6.2.3

$$\kappa(R_n, G) \xrightarrow{d} \mathrm{Po}\left(\frac{\rho^{|G|}}{\mathrm{aut}(G)}\right)$$

for every object G in the finite set $\mathcal{U}_s$, and the random variables are asymptotically independent. Then by Lemma 6.2.2

$$\begin{aligned} X_s = \sum_{G \in \mathcal{U}_s} \kappa(R_n, G) &\xrightarrow{d} \sum_{G \in \mathcal{U}_s} \mathrm{Po}\left(\frac{\rho^{|G|}}{\mathrm{aut}(G)}\right) \\ &= \mathrm{Po}\left(\sum_{G \in \mathcal{U}_s} \frac{\rho^{|G|}}{\mathrm{aut}(G)}\right) = \mathrm{Po}\left(\frac{\rho^s}{s!} \sum_{G \in \mathcal{U}_s} \frac{|G|!}{\mathrm{aut}(G)}\right) \\ &= \mathrm{Po}\left(\frac{C_s \rho^s}{s!}\right) \end{aligned}$$

and $X_1, \ldots, X_j$ are asymptotically independent by Lemma 6.2.2 as they are sums over disjoint subsets of the finite set $\bigcup_{s=1}^{j} \mathcal{U}_s$ of asymptotically independent random variables. □

For any simple finite pattern S we have the following result.

Corollary 6.2.5. *Let* $\lambda_j = \frac{C_j \rho^j}{j!}$. *Then*

$$\mathbb{P}(R_n \textit{ has pattern } S) \sim \prod_{j \in J} \frac{1}{S(j)!} \left(\lambda_j\right)^{S(j)} e^{-\lambda_j}.$$

Observe that, in comparison to Theorem 4.1.4 (of [37]) we consider only a simple restricted pattern but we allow a much more general class of combinatorial objects. Once more, we are interested in the size of the smallest component in decomposable structures.

We are interested in the size of the smallest component in decomposable structures.

Lemma 6.2.6. *Let* $\mathcal{A}$ *be a decomposable and smooth class with radius of convergence* $\rho > 0$. *Let* $s(R_n)$ *be the size of the smallest component in* $R_n \in_u \mathcal{A}$. *Recall* $\lambda_i = \frac{C_i \rho^i}{i!}$. *Then for each* $k \in \mathbb{N}$ *as* $n \to \infty$

$$\mathbb{P}(s(R_n) \geq k+1) \sim e^{-\sum_{i=1}^{k} \lambda_i}. \tag{6.3}$$

Proof. We apply Corollary 6.2.5. Recall that according to Definition 6.1.2 the smallest component has size at least $k+1$ translates into $S(1) = \ldots = S(k) = 0$ and as $n \to \infty$

$$\mathbb{P}(s(R_n) \geq k+1) \to \exp\left(-\sum_{i=1}^{k} \lambda_i\right).$$

□

Observe that the RHS in (6.10) tends to $1/A(\rho)$ if $A(\rho) < \infty$, and tends to 0 if $A(\rho) = \infty$. Thus if $A(\rho) = \infty$ we have a (non-defective) limiting distribution for $s(R_n)$. The following Corollary should be compared to [10, Theorem 3 b] and [76, Theorem 5.5].

Corollary 6.2.7. *Let $\mathcal{A}$ be a decomposable and smooth class. Then*

$$\limsup_{n\in\mathbb{N}} \mathbb{P}(R_n \text{ connected}) \leq e^{-C(\rho)} = \frac{1}{A(\rho)}.$$

Equality clearly holds for $C(\rho) = \infty$ (see also [10, Theorem 3]). Furthermore, equality holds if $C(\rho) < \infty$ and $\lim \mathbb{P}(R_n \text{ connected})$ exists, (see Theorem 2.2.3).

Proof. By Lemma 6.2.6 we have for each fixed k and each $n > k$

$$\mathbb{P}(R_n \text{ connected}) \leq \mathbb{P}(s(R_n) > k) \to \exp\left(-\sum_{i=1}^{k} \lambda_i\right).$$

Hence

$$\limsup_{n\in\mathbb{N}} \mathbb{P}(R_n \text{ connected}) \leq \limsup_{k\in\mathbb{N}} \exp\left(-\sum_{i=1}^{k} \lambda_i\right) = \exp(-C(\rho)).$$

□

In some special diverging cases we can compute the limiting distribution of the size of the smallest component explicitly.

Corollary 6.2.8. *If $\mathcal{A}$ is decomposable, smooth and additionally*

$$\frac{C_i \rho^i}{i!} = \frac{C_j \rho^j}{j!} = \lambda \tag{6.4}$$

holds for all i, j – clearly $C(\rho) = \infty$ in this case – then

$$s(R_n) \xrightarrow{d} X$$

where X is Geom$\left(1-e^{-\lambda}\right)$. In particular for the limiting distribution $\mathbb{E}(X) = \frac{1}{1-e^{-\lambda}} \geq 1$.

An example of a decomposable class that fulfils (6.4) is the class of permutations or equivalently the class of directed paths. Here $\frac{C_i\rho^i}{i!} = 1$ for all $i \in \mathbb{N}$ and the limiting distribution of the smallest component is Geom $(1-e^{-1})$ and so $1 < \mathbb{E}(X) < \infty$.

In fact, it is not difficult to see that in the divergent case, we have natural classes with finite mean.

Corollary 6.2.9. *If $\mathcal{A}$ is decomposable, smooth and additionally*

$$\liminf_{i\in\mathbb{N}} \frac{C_i\rho^i}{i!} > 0 \tag{6.5}$$

holds – clearly $C(\rho) = \infty$ in this case – then for X with the limiting distribution of $s(R_n)$

$$\mathbb{E}(X) < \infty.$$

Proof. Let $\lambda_i = \frac{C_i\rho^i}{i!}$. We have

$$\mathbb{E}(X) = \sum_{k\geq 1} \mathbb{P}(X \geq k) = \sum_{k\geq 1} \exp(-\sum_{i=1}^{k} \lambda_i)$$

and for k large enough

$$\lambda_k \geq \frac{1}{2} \liminf_{i\in\mathbb{N}} \lambda_i > 0.$$

Hence

$$\frac{\exp(-\sum_{i=1}^{k} \lambda_i)}{\exp(-\sum_{i=1}^{k-1} \lambda_i)} = \exp(-\lambda_k) \leq \exp\left(-\frac{1}{2} \liminf_i \lambda_i\right) < 1.$$

By the ratio test for series we know that $\mathbb{E}(X) = \sum_{k\geq 1} \exp(-\sum_{i=1}^{k} \lambda_i) = c < \infty$. □

Observe that

$$\mathbb{E}(s(R_n)) = \sum_{k=1}^{j} \mathbb{P}(s(R_n) \geq k) + \sum_{k=j+1}^{n} \mathbb{P}(s(R_n) \geq k)$$

and for every ϵ there exists a large enough j large

$$\sum_{k=1}^{j} \mathbb{P}(s(R_n) \geq k) \to \mathbb{E}(X) - \epsilon.$$

Then

$$\liminf_{n\in\mathbb{N}} \mathbb{E}(s(R_n)) \geq \liminf_{j\in\mathbb{N}} \sum_{k=1}^{j} \mathbb{P}(s(R_n) \geq k) = \mathbb{E}(X)$$

– even if $\mathbb{E}(R_n) = \infty$.

Corollary 6.2.9 shows that given certain assumptions on the coefficients of the class the expected value of the limiting distribution is finite. Of course it would be nice to show that in fact $\lim_{n\to\infty} \mathbb{E}(s(R_n)) = \mathbb{E}(X)$ holds and therefore the expected value of the smallest component in $R_n \in_u \mathcal{A}$ is finite. But in general, no such result is known. For forests of paths and forests of caterpillars we have $\liminf_{i\in\mathbb{N}} \frac{C_i\rho^i}{i!} = \frac{1}{2} > 0$, $\liminf_{i\in\mathbb{N}} \frac{C_i\rho^i}{i!} = \alpha > 0$ (see Theorem 3.4.1) and their generating functions diverge at the radius of convergence. Similarly for forests of bow ties. Hence Corollary 6.2.9 applies and the expected value of the limiting distribution of the smallest component in these classes tends to a finite limit. Furthermore, we know that forests of paths and forests of caterpillars have a simple pole (Theorem 3.4.1) and the asymptotic expansion of their generating function is of alg-log type around their singularity. In fact, they are both of exactly the type of generating functiont that was analysed in [37]. Hence by [37, Corollary 4.3] in these cases the limiting expected value of $s(R_n)$ is the limiting expected value of the limiting distribution. Hence the expected value of the smallest component for forests of paths is given by

$$\begin{aligned}\mathbb{E}(s(R_n)) &\to e^0 + \exp(-1) \cdot \sum_{k\geq 0} \exp(-1/2 \cdot k) \\ &= 1 + (e - e^{1/2})^{-1} \approx 1.93.\end{aligned}$$

For forests of caterpillars getting the precise asymptotic value of $\mathbb{E}(s(R_n))$ is difficult as we do not have access to C_n for all n. But we observe that $\mathbb{E}(s(R_n))$ is bounded.

In the convergent case, the behaviour changes and we have access to the expected value of the size of the smallest component in the limiting object $R_n \in_u \mathcal{A}$.

Proposition 6.2.10. *Let $\mathcal{A}$ be decomposable, smooth, and additionally $C(\rho) < \infty$, $\lim_n \frac{C_n}{A_n}$ exists and $\mathbb{E}(\mathrm{frag}(R_n)) < c$ then*

$$\mathbb{E}(s(R_n)) \sim n \cdot e^{-C(\rho)} = \frac{n}{A(\rho)}. \tag{6.6}$$

Proof. We have

$$\mathbb{E}(s(R_n)) = n \cdot \frac{C_n}{A_n} + \mathbb{E}[s(R_n) \cdot \mathbb{1}_{R_n \text{ not connected}}].$$

But $s(R_n)\mathbb{1}_{R_n \text{ not connected}} \leq \text{frag}(R_n)$ and so

$$\mathbb{E}[s(R_n) \cdot \mathbb{1}_{R_n \text{ not connected}}] \leq \mathbb{E}(\text{frag}(R_n)) \leq c.$$

By Theorem 2.2.3 we know that if $\mathcal{A}$ is decomposable, smooth, $C(\rho) < \infty$ and $\frac{C_n}{A_n}$ tends to a limit it has to be $e^{-C(\rho)}$. □

Recall that any (proper) minor-closed classes of graphs where the excluded minors are 2-connected satisfies all assumptions of Proposition 6.2.10 (see Chapter 3 and [71] for the fact that $\mathbb{E}(\text{frag}(R_n))$ is bounded). For these examples Proposition 6.2.10 tells us precisely what we expect: We know that asymptotically R_n is connected with probability $\sim e^{-C(\rho)} > 0$ and therefore with positive probability there is only one component (which is obviously also the smallest component) and the smallest component has expected size $n \cdot e^{-C(\rho)}$.

We also observe that Proposition 6.2.10 presents a natural extension of [37, Corollary 3.7] for labelled structures. [37, Corollary 3.7] applies to generating function of alg-log type with $0 < \alpha < 1$ and any β and gives the same asymptotic expression for $\mathbb{E}(s(R_n))$. Lemma 2.2.6 implies that classes, which have connected component generating functions of alg-log type with $0 < \alpha < 1$ and any β, have $\mathbb{E}(\text{frag}(R_n)) \to \infty$ as $n \to \infty$. Hence Proposition 6.2.10 does not apply. An example of such a combinatorial class is the class of forests of rooted trees. In this case, $\rho = e^{-1}$, $\alpha = \frac{1}{2}$, $\beta = 0$, $d = \sqrt{2}$ and $c = 1$ (see Section 2.3). Observe that by [37, Corollary 3.7] the expected value of the smallest component in a random forest of rooted trees is still $\mathbb{E}(s(R_n)) \sim n \cdot e^{-C(\rho)} = \frac{n}{e}$ which suggests that Proposition 6.2.10 is not best possible.

On the other hand our result applies to any generating function of alg-log type with $\alpha > 1$ and any β and therefore in particular to forests of unrooted trees (see Section 2.3). By Proposition 6.2.10 the expected value of the smallest component in a random forest of unrooted trees is $\mathbb{E}(s(R_n)) \sim \frac{n}{\sqrt{e}}$.

6.2.2. Unlabelled structures

We are now investigating unlabelled structures. In the following $\mathcal{A}$ and $\mathcal{C}$ denote **unlabelled** sets of combinatorial objects. Let H be a connected, unlabelled object in $\mathcal{C}$ and let $\kappa(R_n, H)$ denote the number of components of R_n equal to H where $R_n \in_u \mathcal{A}$ is now taken uniformly at random from the set of all unlabelled objects of size n in $\mathcal{A}$.

Theorem 6.2.11. *Let $\mathcal{A}$ be an unlabelled, decomposable and smooth class so $\mathcal{A} = \mathrm{MSET}(\mathcal{C})$ and let $0 < \rho < 1$ denote its radius of convergence. Define $h_i = |V(H_i)|$ for a combinatorial object $H_i \in \mathcal{C}$. For each fixed j let $H_1, \ldots, H_j$ be a collection of pairwise non-identical objects in $\mathcal{C}$. We have for all $i \leq j$ as $n \to \infty$*

$$\kappa(R_n, H_i) \xrightarrow{d} \textit{Geometric}(1 - \rho^{h_i}) \tag{6.7}$$

and $\kappa(R_n, H_1), \ldots, \kappa(R_n, H_j)$ are asymptotically independent.

The condition $0 < \rho < 1$ translates the fact that the class $\mathcal{A}$ needs to grow at least exponentially.

Proof. For each $k \in \mathbb{N}$ fixed we can construct a combinatorial object in $\mathcal{A}$ with at least k copies of $H \in \mathcal{C}$ by putting down the k copies of H and then constructing any object in $\mathcal{A}$ on the remaining $n - kh$ atoms. Let $h = V(H)$. We have [1]

$$\mathbb{P}(\kappa(R_n, H) \geq k) = \frac{A_{n-hk}}{A_n} \sim \rho^{hk} \tag{6.8}$$

by smoothness. This gives

$$\mathbb{P}\kappa(R_n, H) = k) = \frac{A_{n-hk}}{A_n} - \frac{A_{n-h(k+1)}}{A_n} \sim \rho^{hk}(1 - \rho^h).$$

Asymptotically the number of appearances of any fixed object H is

$$\mathrm{Geometric}(1 - \rho^h).$$

Suppose we take any finite collection of distinct combinatorial objects $H_1, \ldots, H_t \in \mathcal{C}$ of sizes $h_1, \ldots, h_t$. Then for n sufficiently large the copies will not interfere with each other, so we can fix the copies of $H_1, \ldots, H_t$ first and then build any object on the remaining vertices

$$\begin{aligned}
\mathbb{P}(\kappa(R_n, H_1) \geq k_1, \ldots, \kappa(R_n, H_t) \geq k_t) &= \frac{A_{n-\sum_{i=1}^t k_i h_i}}{A_n} \\
&= \frac{A_{n-\sum_{i=1}^t k_i h_i}}{A_{n-\sum_{i=1}^t k_i h_i + 1}} \cdot \frac{A_{n-\sum_{i=1}^t k_i h_i + 1}}{A_{n-\sum_{i=1}^t k_i h_i + 2}} \cdots \frac{A_{n-1}}{A_n} \\
&\sim \rho \cdot \rho \cdot \ldots \rho \\
&= \rho^{\sum_{i=1}^t k_i h_i}.
\end{aligned}$$

[1] For further comments about the next line, please see the discussion after Corollary 6.2.16.

and a short calculation implies

$$\begin{aligned}&\mathbb{P}(\kappa(R_n, H_1) = k_1, \ldots, \kappa(R_n, H_t) = k_t)\\ \sim&\rho^{h_1 k_1}(1-\rho_1^h)\cdot\rho^{h_2 k_2}(1-\rho^{h_2})\cdot\ldots\rho^{h_t k_t}(1-\rho^{h_t}).\end{aligned}$$

Hence for any finite collection of distinct combinatorial objects $H_1, \ldots, H_t \in \mathcal{C}$ the random variables $\kappa(R_n, H_1), \ldots, \kappa(R_n, H_t)$ will be asymptotically independent Geom$(1-\rho^{h_i})$. □

Let X be the number of failures before the first success, in a series of n independent trials with probability of success p. Then

$$\mathbb{P}(X = k) = \binom{n+k-1}{k}(1-p)^k p^n \text{ for } k = 0, 1, \ldots.$$

This is (a form of) the *negative binomial distribution* with parameters n and p. Recall that C_i is the number of connected (unlabelled) objects in the unlabelled class $\mathcal{A}$.

Proposition 6.2.12. *Let $\mathcal{A}$ be an unlabelled, decomposable and smooth class $\mathcal{A} = MSET(\mathcal{C})$ with radius of convergence $0 < \rho < 1$. Let X_i denote the number of components of size i in $R_n \in_u \mathcal{A}$. Then for each fixed j and for all $i \leq j$ as $n \to \infty$*

$$\mathbb{P}(X_i = k_i) \sim \binom{C_i + k_i - 1}{k_i}\rho^{ik_i}(1-\rho^i)^{C_i} \tag{6.9}$$

and the random variables $X_1, X_2, \ldots X_j$ are asymptotically independent.

The display (6.9) shows that X_i are asymptotically independent negative binomials with parameters C_i and $(1-\rho^i)$.

Proof. Fix i. We can construct each object with k_i components of size i by picking k_i objects in C_i and constructing any graph on the remaining vertices. For every given H_l in $\mathcal{C}_i$ and every fixed integer t_l by Theorem 6.2.11

$$\mathbb{P}(\kappa(R_n, H_l) = t_l) \sim \rho^{it_l}(1-\rho^i).$$

Hence for each C_i-tuple $t_1, \ldots, t_{C_i}$ with $\sum_{l=1}^{C_i} t_l = k_i$ of non-negative integers by Lemma 6.2.2 as C_i is finite

$$\begin{aligned}\mathbb{P}\left(\bigwedge_{l=1}^{C_i} \kappa(R_n, H_l) = t_l\right) &\sim \prod_{l=1}^{C_i} \rho^{it_l}(1-\rho^i) \\ &= (1-\rho^i)^{C_i}\rho^{ik_i}.\end{aligned}$$

We have $\binom{C_i+k_i-1}{k_i}$ choices to pick the C_i-tuple $t_1, \ldots, t_{C_i}$ with $\sum_{l=1}^{C_i} t_l = k_i$ of non-negative integers and hence

$$\mathbb{P}(X_i = k_i) \sim \binom{C_i + k_i - 1}{k_i}\rho^{ik_i}(1-\rho)^{C_i}.$$

By Lemma 6.2.2 $X_1, \ldots, X_j$ are asymptotically independent as they are sums over disjoint subsets over the finite set $\cup_{i=1}^{j} C_j$ of asymptotically independent random variables. □

For any simple pattern S, we have

Corollary 6.2.13.

$$\mathbb{P}(R_n \textit{ has pattern } S) \sim \prod_{j\in J}\binom{C_j + S(j) - 1}{S(j)}\rho^{jS(j)}(1-\rho^j)^{C_j}.$$

Observe that, in comparison to Theorem 4.1.4 (of [37]) we consider only a simple restricted pattern but we allow a much more general class of combinatorial objects. Once more, we are interested in the size of the smallest component in decomposable structures.

Lemma 6.2.14. *Let $\mathcal{A}$ be a decomposable, unlabelled and smooth class with radius of convergence $0 < \rho < 1$. Let $s(R_n)$ be the size of the smallest component in $R_n \in_u \mathcal{A}$. Then for each $k \in \mathbb{N}$ as $n \to \infty$*

$$\mathbb{P}(s(R_n) \geq k+1) \to \prod_{j=1}^{k}\left(1-\rho^j\right)^{C_j}. \tag{6.10}$$

Proof. We apply Corollary 6.2.13. Recall that according to Definition 6.1.2 the smallest component has size at least $k+1$ translates into $S(1) = \ldots = S(k) = 0$ and as $n \to \infty$

$$\mathbb{P}(s(R_n) \geq k+1) \to \prod_{j=1}^{k}(1-\rho^j)^{C_j}.$$

□

Observe that the RHS in (6.10) tends to $1/A(\rho)$ if $A(\rho) < \infty$, and tends to 0 if $A(\rho) = \infty$. For an explanation of this fact, see the proof for the formula of the multiset construction at the beginning of this chapter. Thus if $A(\rho) = \infty$ we have a (non-defective) limiting distribution for $s(R_n)$.

Analogously to Corollary 6.2.7 and [10, Theorem 3 b] we have the following statement.

Corollary 6.2.15. *Let $\mathcal{A}$ be an unlabelled, decomposable and smooth class with $0 < \rho < 1$. Then*

$$\limsup_{n\in\mathbb{N}} \mathbb{P}(R_n \text{ connected}) \leq e^{-\sum_{j\geq 1} \frac{C(\rho^j)}{j}} = \frac{1}{A(\rho)}.$$

Equality holds if $C(\rho) = \infty$. For $C(\rho) < \infty$ equality holds if $\lim_{n\to n} \mathbb{P}(R_n \text{ connected})$ exists (see [10, Theorem 2, unlabelled version]).

Proof. For each fixed k and for each $n > k$

$$\mathbb{P}(R_n \text{ connected}) \leq \mathbb{P}(s(R_n) > k) \to \prod_{j=1}^{k} (1-\rho^j)^{C_j}.$$

Hence

$$\limsup_{n\in\mathbb{N}} \mathbb{P}(R_n \text{ connected}) \leq \inf_{k\in\mathbb{N}} \prod_{j=1}^{k} (1-\rho^j)^{C_j} = e^{-\sum_{j\geq 1} \frac{C(\rho^j)}{j}} = 1/A(\rho).$$

□

Similarly to Corollary 6.2.9, we have

Corollary 6.2.16. *If $\mathcal{A}$ is decomposable, smooth, $0 < \rho < 1$ and additionally*

$$\liminf_{i\in\mathbb{N}} C_i \rho^i > 0 \tag{6.11}$$

holds – clearly $C(\rho) = \infty$ in this case – then for X with the limiting distribution of $s(R_n)$

$$\mathbb{E}(X) < \infty.$$

Proof. Let $\lambda_i = C_i\rho^i$. We write

$$\mathbb{E}(X) = \sum_{k\geq 1} \mathbb{P}(X \geq k) = \sum_{k\geq 1} \prod_{j=1}^{k} (1-\rho^j)^{C_j}$$

and observe that for k large enough

$$\rho^{k+1} C_{k+1} \geq \frac{1}{2} \liminf_{i\in\mathbb{N}} \lambda_i > 0.$$

Then

$$\frac{\prod_{j=1}^{k+1}(1-\rho^j)^{C_j}}{\prod_{j=1}^{k}(1-\rho^j)^{C_j}} = (1-\rho^{k+1})^{C_{k+1}} \leq e^{-\rho^{k+1}C_{k+1}} \leq e^{-\frac{1}{2}\liminf_{i\in\mathbb{N}}\lambda_i} < 1.$$

□

Unfortunately, all results presented so far and in particular the ones about diverging ordinary generating functions do not apply to ordinary generating functions with radius of convergence equal to 1. The same holds for the results in [37]. Hence they do not apply to many interesting combinatorial sets such as forests of unlabelled paths or integer compositions.

Observe that for a smooth class $\mathcal{A}$ with $\rho = 1$ equation (6.8) still holds and for every fixed $k \in \mathbb{N}$ and object H of size h

$$\mathbb{P}(\kappa(R_n, H) \geq k) = \frac{A_{n-hk}}{A_n} \sim 1.$$

Hence asymptotically R_n contains with probability 1 any finite number of any connected, finite object. This suggests that to obtain any non-trivial result on the limiting distribution of $X_1, \ldots, X_j$ and similarly to Chapter 3, one would first have to normalise $X_1, \ldots, X_j$ (using the appropriate mean and variance, which are not going to be constant anymore) before deducing the limit laws. Observe that a connected caterpillar is equivalent to an integer composition. A forest of paths corresponds to an unordered sequence of integers which sum up to n. Hence forests of path are precisely integer partitions. In [37] some of the results were extended to the special case of integer partitions and they obtained the following result.

Proposition 6.2.17 (**Integer partitions [37]**). *Let $\mathcal{A}$ be the class of forests of paths or equivalently the class of integer partitions. Then as $n \to \infty$*

$$\mathbb{P}(R_n \textit{ has finite pattern } S) \to \prod_{j\in J} \frac{\pi j}{\sqrt{6n}}.$$

The formula above, i.e. the product on j, indicates that the contribution of each part in the restricted pattern is still asymptotically independent of each other.
Given an unlabelled object G with largest component H let $\mathrm{Frag}(G)$ be the object obtained by deleting H. As in the labelled case $\mathrm{Frag}(G)$ is the empty graph if G is connected. Similarly to Proposition 6.2.10 we have the following result.

Proposition 6.2.18. *If $\mathcal{A}$ is decomposable, smooth, $0 < \rho < 1$, $C(\rho) < \infty$ and $\lim_{n\to\infty} \frac{C_n}{A_n}$ exists and $\mathbb{E}(\mathrm{frag}(R_n)) \leq c < \infty$ then*

$$\mathbb{E}(s(R_n)) \sim n \cdot e^{-\sum_{k=1}^{\infty} \frac{C(\rho^k)}{k}} = \frac{n}{A(\rho)}. \tag{6.12}$$

Proof. Observe that for unlabelled structures (and non-finite generating function) $\rho < 1$ and $C(\rho) < \infty$ implies $\sum_{k=1}^{\infty} \frac{C(\rho^k)}{k} < \infty$. To see this observe that $C(\rho^{k+1}) \leq \rho C(\rho^k)$ and hence

$$\frac{C(\rho^{k+1})}{k+1} \cdot \frac{k}{C(\rho^k)} = \frac{k}{k+1} \cdot \rho < \rho < 1$$

and so $A(\rho) < \infty$. We have

$$\mathbb{E}(s(R_n)) = n \cdot \frac{C_n}{A_n} + \mathbb{E}(s(R_n) \cdot [\mathbb{1}_{R_n \text{ not connected}}].$$

But $s(R_n)\mathbb{1}_{R_n \text{ not connected}} \leq \mathrm{frag}(R_n)$ and so

$$\mathbb{E}[s(R_n) \cdot \mathbb{1}_{R_n \text{ not connected}}] \leq \mathbb{E}(\mathrm{frag}(R_n)) \leq c.$$

By [10, Theorem 2 unlabelled version] we know that if $\mathcal{A}$ is decomposable, smooth, $C(\rho) < \infty$ and $\frac{C_n}{A_n}$ tends to a limit it has to be equal to $e^{-\sum_{k=1}^{\infty} \frac{C(\rho^k)}{k}} = \frac{1}{A(\rho)}$. □

Once more our result complements [37, Corollary 3.7 unlabelled part]. An example where our result applies (but not the one of [37]) are forests of unlabelled trees. Here all the assumptions of Proposition 6.2.18 are fulfilled, in particular the coefficients are smooth and $\lim_{n\to\infty} \frac{C_n}{A_n} \approx 0.5228$ by [81] and the expectation of the fragment is bounded (see [61]). Hence we expect the smallest component in a random unlabelled forests to be of size $0.5228 \cdot n$. For $k \leq 10$ Proposition 6.2.18 also applies to Ex((k+1)C) – unlabelled, minor-closed classes of graphs excluding $(k+1)$ disjoint cycles (see [61, Table 3] for values of $\mathbb{E}(\mathrm{frag}(R_n))$ and $\lim_{n\to\infty} C_n/A_n$).

7. Conclusion

In this thesis we investigated random graphs sampled from a suitably structured class. We focused on the probability that the sampled object is connected and on its number of components. In particular, we studied variations of bridge-addability and can conclude that bridge-addability seems to be quite robust. Even if not all but only a few bridges are allowed to be introduced, the typical bridge-addable behaviour, i.e. the probability of the random object to be connected is bounded away from 0, the expected number of components is stochastically at most a shifted Poisson law and the number of vertices in the fragment is bounded, can be observed (Chapter 4 and Chapter 5). Given these results, it seems natural to extend a conjecture of McDiarmid et al. from 2006 and conjecture that for every relatively bridge-addable class, the forests belonging to this class are the class which is the least likely to be connected (Conjecture 4.5.1).

For minor-closed classes of graphs it turns out that bridge-addability is not the only determining factor: all minor-closed, bridge-addable classes exhibit the typical bridge-addable behaviour but there exist classes of graphs (for example graphs excluding a spoon) that are not bridge-addable but still behave very similarly (Chapter 3). A complete classification of minor-closed classes of graphs seems to be very desirable but still a distant prospect. In particular, a classification ofconnectivity behaviour of minor-closed classes of graphs according to the behaviour of the generating function next to its radius of convergence seems to be a crucial but unaccomplished and probably difficult task.

In Chapter 6, we observed that, in many respects, labelled and unlabelled decomposable classes of combinatorial objects behave quite similarly: the unlabelled graph appear asymptotically independently and the limit laws can be determined. From this the limit law of the smallest component can be obtained. To determine the expected size of the smallest component, the behaviour of the generating function next to its radius of convergence is again crucial.

A. Statement of originality

The results in Chapters 2, 3 and first part of Chapter 5 are (mostly) joint work with Mireille Bousquet-Mélou, Université de Bordeaux. An extended abstract of the results in Chapter 2 and 3 was published (Title: "Asymptotic properties of some minor-closed classes of graphs") at the conference Formal Power Series and Algorithms, Paris, June 2013. A full version including proofs of the results of Chapter 2 and 3 was accepted for publication for a special issue of Combinatorics, Probability and Computing dedicated to the memory of Philippe Flajolet. First results on coloured forests, forests of paths, forests of caterpillars and general proof ideas in particular involving Hayman- and extended Hayman-admissibility very similar to the ones used in Chapter 2 appeared in my transfer thesis (Michelmas Term 2011). Chapter 4 and 6 are joint work with my supervisor Colin McDiarmid, University of Oxford. Parts of Chapter 4 were submitted as an extended abstract to Latin American Theoretical INformatics Symposium 2014.

B. Acknowledgements

First and foremost I would like to thank my supervisor Colin McDiarmid. It has been an honour to be his DPhil student. With patience and calmness he has taught me what mathematical research is about and provided both scientific and moral support over the course of my PhD. I appreciate all his contributions of time and ideas and the freedom he gave me to make my DPhil experience productive and worthwhile. The enthusiasm he has for his research was contagious and highly motivational for me, even during tough times.

I would also like to thank my collaborator and friend Mireille Bousquet-Mélou. Without her listening to me and dancing with me on a conference dinner in July 2011 large part of this work would not have been possible. Her fantastic mathematical knowledge as well as her friendship has been truly inspiring and stimulating.

I am also thankful for the support of the Combinatorics Group at Oxford. The group has been a source of friendships as well as good advice. In particular, I would like to thank Kitty Meeks for many nice lunches, coffees and encouraging discussions. Oxford would have been a very different place for me without her. Lutz Warnke for long conversations about mathematics and life in general, as well as scientific advise. Fiona Skerman, Anne Hillebrand, Puck Rombach for many uplifting conversations, lunches, dinner, baking cookies, making birds and many more.

I have very much appreciated the camaraderie of my 'French gang'. Jérémie Lambruso and Basile Morcrette for being wonderful friends and organizing the 'French-British Workshop in Analytic Combinatorics' with me. Basile, for the music we played together at Alea and in Montreal – Mozart's Duo will always be remembered as conference music! Jérémie for numerous hilarious phone calls from all over the world that always cheered me up. Julien Courtiel for his good mood, consistent optimism, his companionship in Bordeaux and elsewhere and many, many fun movie nights at the cinema.

My time at Oxford was made enjoyable in large part due to the many friends that became a part of my life. My biggest thank you goes to my wonderful house mates Nikita Loik, Emily Spiers and Liz Williams for the fantastic and fun almost two years that we spend together in Albert Street 19.

Megan Robb for our walks in University Parks and the many cups of tea we drank together.

Andreas Ditter for the fantastic music we played together.

Maria Wuithschick – thank you very much for the nice breaks we had together during the last three years, for always being here there and listening to me when I needed you.

Alastair for proof-reading my thesis and supporting me during tough times. Ich bin unglaublich froh und dankbar, dass ich dich habe!

Lastly, I would like to thank my family for all their love and encouragement. Thank you, Eva-Maria und Lothar, for supporting me in all my pursuits. Thank you my sisters, Heidrun and Silke, for listening and encouraging me in difficult moments.

Vielen Dank.

Bibliography

[1] Addario-Berry, L., McDiarmid, C., and Reed, B. (2012). Connectivity for bridge-addable monotone graph classes. *Combinatorics, Probability and Computing*, 21:803–815.

[2] Aigner, M. (2007). *A course in enumeration*, volume 238 of *Graduate Texts in Mathematics*. Springer, Berlin.

[3] Aigner, M. and Ziegler, G. (2004). *Proofs from THE BOOK (3. ed.)*. Springer.

[4] Alon, N. and Spencer, J. H. (2008). *The probabilistic method*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons Inc., Hoboken, NJ, third edition.

[5] Arratia, R., Barbour, A. D., and Tavaré, S. (2003). *Logarithmic combinatorial structures: a probabilistic approach*. EMS Monographs in Mathematics. European Mathematical Society (EMS), Zürich.

[6] Austin, T. L. (1960). The enumeration of point labelled chromatic graphs and trees. *Canadian Journal of Mathematics*, 12:535–545.

[7] Balister, P., Bollobás, B., and Gerke, S. (2008). Connectivity of addable graph classes. *Journal of Combinatorial Theory, Series B*, 98:577–584.

[8] Banderier, C., Flajolet, P., Schaeffer, G., and Soria, M. (2001). Random maps, coalescing saddles, singularity analysis, and Airy phenomena. *Random Structures Algorithms*, 19(3-4):194–246.

[9] Barbour, A. and Granovsky, B. L. (2005). Random combinatorial structures: the convergent case. *Journal of Combinatorial Theory, Series A*, 109(2):203 – 220.

[10] Bell, J., Bender, E., Cameron, P., and Richmond, L. (2000). Asymptotics for the probability of connectedness and the distribution of number of components. *Electronic Journal of Combinatorics*, 7:R33.

[11] Bell, J. P. (2000). When structures are almost surely connected. *Electronic Journal of Combinatorics*, 7:3–6.

[12] Bell, J. P. (2002). Sufficient conditions for zero-one laws. *Transactions of the American Mathematical Society*, 354(2):613–630 (electronic).

[13] Bell, J. P. and Burris, S. N. (2003). Asymptotics for logical limit laws: When the growth of the components is in an rt class. *Transactions of the American Mathematical Society*, 355:3777–3794.

[14] Bell, J. P., Burris, S. N., and Yeats, K. A. (2012). Monadic second-order classes of forests with a monadic second-order 0-1 law. *Discrete Mathematics & Theoretical Computer Science*, 14(1):87–107.

[15] Bender, E. A. (1974). Asymptotic methods in enumeration. *SIAM Review*, 16:485–515.

[16] Bender, E. A., Canfield, E. R., and Richmond, L. B. (2008). Coefficients of functional compositions often grow smoothly. *Electronic Journal of Combinatorics*, 15:Research Paper 21, 8.

[17] Bender, E. A., Gao, Z., and Wormald, N. C. (2000). The number of labeled 2-connected planar graphs. *Journal of Combinatorics*, 9:43.

[18] Bender, E. A., Mashatan, A., Panario, D., and Richmond, L. B. (2004). Asymptotics of combinatorial structures with large smallest component. *Journal of Combinatorial Theory, Series A*, 107:117–125.

[19] Bernardi, O. (2013). Personal communication.

[20] Bernardi, O., Noy, M., and Welsh, D. (2010). Growth constants of minor-closed classes of graphs. *Journal of Combinatorial Theory, Series B*, 100(5):468 – 484.

[21] Billingsley, P. (1995). *Probability and Measure*. John Wiley & Sons Inc.

[22] Bodirsky, M., Giménez, O., Kang, M., and Noy, M. (2007). Enumeration and limit laws for series-parallel graphs. *European Journal of Combinatorics*, 28:2091–2105.

[23] Bollobás, B. (2001). *Random Graphs*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, second edition.

[24] Bousquet-Mélou, M. and Weller, K. (2013). Asymptotic properties of some minor-closed classes of graphs. *Combinatorics, Probability and Computing (accepted)*.

[25] Burris, S. (2001). *Number theoretic density and logical limit laws*. Mathematical Surveys and Monographs 86. American Mathematical Society.

[26] Burris, S. and Yeats, K. (2008). Sufficient conditions for labelled 0-1 laws. *Discrete Mathematics & Theoretical Computer Science*, 10:–.

[27] Cameron, P. J. (1997). On the probability of connectedness. *Discrete Mathematics*, 167-168:175 – 187. Selected Papers 15th British Combinatorial Conference.

[28] Canfield, E. R. (1977). Central and local limit theorems for the coefficients of polynomials of binomial type. *Journal of Combinatorial Theory, Series A*, 23(3):275–290.

[29] Cayley, P. (1881). On the Analytical Forms Called Trees. *American Journal of Mathematics*, 4(1-4):266–268.

[30] Chapuy, G., Fusy, E., Giménez, O., Mohar, B., and Noy, M. (2011). Asymptotic enumeration and limit laws for graphs of fixed genus. *Journal of Combinatorial Theory, Series*, 118:748–777.

[31] Compton, K. J. (1987). Some methods for computing component distribution probabilities in relational structures. *Discrete Mathematics*, 66:59–77.

[32] Corless, R. M., Gonnet, G. H., Hare, D. E. G., Jeffrey, D. J., and Knuth, D. E. (1996). On the Lambert W function. *Advances in Computational Mathematics*, 5(4):329–359.

[33] Darrasse, A., Panagiotou, K., Roussel, O., and Soria, M. (2012). Biased boltzmann samplers and generation of extended linear languages with shuffle. In *Proceedings of the 23rd International Meeting on Probabilistic, Combinatorial, and Asymptotic Methods in the Analysis of Algorithms (AofA'12)*, pages 125–140. Discrete Mathematics &Theoretical Computer Science.

[34] de Bruijn, N. ((1958) 1981). *Asymptotic methods in analysis*. Dover Publications, Inc., New York.

[35] Denise, A., Vasconcellos, M., and Welsh, D. J. A. (1996). The random planar graph. *Congressus Numerantium*, 113:61–79.

[36] Diestel, R. (2005). *Graph theory*, volume 173 of *Graduate Texts in Mathematics*. Springer-Verlag, Berlin, third edition.

[37] Dong, L., Gao, Z., Panario, D., and Richmond, L. (2010). Asymptotics of smallest component sizes in decomposable structures of alg-log type. *Discrete Mathematics & Theoretical Computer Science*, 12(2):197–222.

[38] Drmota, M. (2009). *Random trees*. SpringerWienNewYork, Vienna.

[39] Drmota, M., Fusy, É., Kang, M., Kraus, V., and Rué, J. (2011a). Asymptotic study of subcritical graph classes. *SIAM Journal of Discrete Mathematics*, 25(4):1615–1651.

[40] Drmota, M., Giménez, O., and Noy, M. (2010). Vertices of given degree in series-parallel graphs. *Random Structures Algorithms*, 36(3):273–314.

[41] Drmota, M., Giménez, O., and Noy, M. (2011b). Degree distribution in random planar graphs. *Journal of Combinatorial Theory Series A*, 118(7):2102–2130.

[42] Drmota, M., Giménez, O., and Noy, M. (2011c). The maximum degree of series-parallel graphs. *Combinatorics, Probability and Computing*, 20(4):529–570.

[43] Drmota, M., Giménez, O., Noy, M., Panagiotou, K., and Steger, A. (2012). The maximum degree of random planar graphs. In *Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '12, pages 281–287. SIAM.

[44] Drmota, M., Gittenberger, B., and Klausner, T. (2005). Extended admissible functions and gaussian limiting distributions. *Mathematics of Computation*, 74:1953–1966.

[45] Drmota, M. and Soria, M. (1995). Marking in combinatorial constructions: Generating functions and limiting distributions. *Theoretical Computer Science*, 144(1-2):67 – 99.

[46] Duchon, P., Flajolet, P., Louchard, G., and Schaeffer, G. (2004). Boltzmann samplers for the random generation of combinatorial structures. *Combinatorics, Probability and Computing*, 13(4-5):577–625.

[47] Dvorak, Z. and Norine, S. (2010). Small graph classes and bounded expansion. *Journal of Combinatorial Theory, Series B*, 100(2):171 – 175.

[48] Flajolet, P. and Odlyzko, A. (1990). Singularity analysis of generating functions. *SIAM Journal of Discrete Mathematics*, 3(2):216–240.

[49] Flajolet, P. and Sedgewick, R. (2009). *Analytic combinatorics*. Cambridge University Press, New York, NY, USA.

[50] Flajolet, P. and Soria, M. (1990). Gaussian limiting distributions for the number of components in combinatorial structures. *Journal of Combinatorial Theory, Series A*, 53(2):165–182.

[51] Gerke, S., Giménez, O., Noy, M., and Weißl, A. (2008). The number of graphs not containing $K_{3,3}$ as a minor. *Electron. J. Combin.*, 15(1):Research Paper 114, 20.

[52] Gerke, S., McDiarmid, C., Steger, A., and Weißl, A. (2005). Random planar graphs with n nodes and a fixed number of edges. In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 999–1007. ACM.

[53] Giménez, O. and Noy, M. (2005). The number of planar graphs and properties of random planar graphs. In *2005 International Conference on Analysis of Algorithms*, Discrete Math. Theor. Comput. Sci. Proc., AD, pages 147–156 (electronic). Assoc. Discrete Math. Theor. Comput. Sci., Nancy.

[54] Giménez, O. and Noy, M. (2009a). Asymptotic enumeration and limit laws of planar graphs. *Journal of the American Mathematical Society*, 22:309–329.

[55] Giménez, O. and Noy, M. (2009b). Counting planar graphs and related families of graphs. In *Surveys in combinatorics 2009*, volume 365 of *London Mathematical Society Lecture Note Series*, pages 169–210. Cambridge University Press, Cambridge.

[56] Giménez, O., Noy, M., and Rué, J. (2013). Graph classes with given 3-connected components: asymptotic enumeration and random graphs. *Random Structures Algorithms*, 42(4):438–479.

[57] Gourdon, X. (1996). *Combinatoire, algorithmique et géométrie des polynômes*. PhD thesis, École Polytechnique.

[58] Gourdon, X. (1998). Largest component in random combinatorial structures. *Discrete Mathematics*, 180(1-3):185 – 209. Proceedings of the 7th Conference on Formal Power Series and Algebraic Combinatorics.

[59] Harary, F. and Palmer, E. M. (1973). *Graphical enumeration*. Academic Press, New York.

[60] Hayman, W. (1956). A generalisation of stirling's formula. *Journal für die reine und angewandte Mathematik (Crelles Journal)*, 1956:67–95.

[61] Kang, M. and McDiarmid, C. (2011). Random unlabelled graphs containing few disjoint cycles. *Random Structures Algorithms*, 38(1-2):174–204.

[62] Kang, M. and Panagiotou, K. (2013). On the connectivity of random graphs from addable classes. *Journal of Combinatorial Theory, Series B*, 103(2):306 – 312.

[63] Kingman, J. F. C. (1993). *Poisson processes*, volume 3 of *Oxford Studies in Probability*. The Clarendon Press Oxford University Press, New York. Oxford Science Publications.

[64] Kingman, J. F. C., Taylor, S. J., Hawkes, A. G., Walker, A. M., Cox, D. R., Smith, A. F. M., Hill, B. M., Burville, P. J., and Leonard, T. (1975). Random discrete distribution. *J. Roy. Statist. Soc. Ser. B*, 37:1–22.

[65] Kolchin, V. F. (1999). *Random graphs*, volume 53 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge.

[66] Kurauskas, V. and McDiarmid, C. (2011a). Random graphs containing few disjoint excluded minors. submitted.

[67] Kurauskas, V. and McDiarmid, C. (2011b). Random graphs with few disjoint cycles. *Combinatorics, Probability and Computing*, 20:763–775.

[68] Lando, S. K. and Zvonkin, A. K. (2004). *Graphs on surfaces and their applications*, volume 141 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag, Berlin. With an appendix by Don B. Zagier, Low-Dimensional Topology, II.

[69] Macintyre, A. J. and Wilson, R. (1954). Operational methods and the coefficients of certain power series. *Mathematische Annalen*, 127:243–250.

[70] McDiarmid, C. (2008). Random graphs on surfaces. *Journal of Combinatorial Theory, Series B*, 98(4):778 – 797.

[71] McDiarmid, C. (2009). Random graphs from a minor-closed class. *Combinatorics, Probability and Computing*, 18:583–599.

[72] McDiarmid, C. (2011). On graphs with few disjoint t-star minors. *European Journal of Combinatorics*, 32(8):1394–1406.

[73] McDiarmid, C. (2012). Connectivity for random graphs from a weighted bridge-addable class. *Electronic Journal of Combinatorics*, 19(4):Paper 53, 20.

[74] McDiarmid, C. (2013). Random graphs from a weighted minor-closed class. *Electronic Journal of Combinatorics*, 20(2):P52.

[75] McDiarmid, C. and Reed, B. (2008). On the maximum degree of a random planar graph. *Combinatorics, Probability and Computing*, 17(4):591–601.

[76] McDiarmid, C., Steger, A., and Welsh, D. J. (2005). Random planar graphs. *Journal of Combinatorial Theory, Series B*, 93(2):187 – 205.

[77] McDiarmid, C., Steger, A., and Welsh, D. J. A. (2006). Random graphs from planar and other addable classes. In *Topics in discrete mathematics*, volume 26 of *Algorithms Combin.*, pages 231–246. Springer, Berlin.

[78] Meir, A. and Moon, A. W. (1989). On an asymptotic method in enumeration. *Journal of Combinatorial Theory Series A*, 51:77–89.

[79] Moon, J. W. (1970). *Counting labelled trees*, volume 1969 of *From lectures delivered to the Twelfth Biennial Seminar of the Canadian Mathematical Congress (Vancouver)*. Canadian Mathematical Congress, Montreal, Quebec.

[80] Norine, S., Seymour, P., Thomas, R., and Wollan, P. (2006). Proper minor-closed families are small. *Journal of Combinatorial Theory, Series B*, 96(5):754 – 757.

[81] Palmer, E. and Schwenk, A. (1979). On the number of trees in a random forest. *Journal of Combinatorial Theory, Series B*, 27(2):109 – 121.

[82] Panario, D. and Richmond, B. (2001a). Smallest components in decomposable structures: exp-log class. *Algorithmica*, 29(1-2):205–226. Average-case analysis of algorithms (Princeton, NJ, 1998).

[83] Panario, D. and Richmond, B. (2001b). Smallest components in decomposable structures: exp-log class. *Algorithmica*, 29(1-2):205–226. Average-case analysis of algorithms (Princeton, NJ, 1998).

[84] Rényi, A. (1959). Some remarks on the theory of trees. *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, 4:73–85.

[85] Robertson, N. and Seymour, P. (1983). Graph minors. i. excluding a forest. *Journal of Combinatorial Theory, Series B*, 35(1):39 – 61.

[86] Robertson, N. and Seymour, P. (1983–2004). Graph minors I–XX. *Almost all in Journal of Combinatorial Theory Series B.*

[87] Scoins, H. I. (1962). The number of trees with nodes of alternate parity. *Proc. Cambridge Philos. Soc.*, 58:12–16.

[88] Stanley, R. P. (1999). *Enumerative combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge.

[89] Tutte, W. T. (1968). On the enumeration of planar maps. *Bulletin of the American Mathematical Society*, 74:64–74.

[90] Wilson, R. (2002). *Four colors suffice*. Princeton University Press, Princeton, NJ.

[91] Wright, E. M. (1967). A relationship between two sequences. *Journal of the London Mathematical Society*, 17:269–304.

[92] Wright, E. M. (1968). A relationship between two sequences iii. *Journal of the London Mathematical Society*, 43:720–724.

[93] Wright, E. M. (1977). The number of connected sparsely edged graphs. *Journal of Graph Theory*, 1(4):317–330.