Asymptotic properties of some minor-closed classes of graphs

Mireille Bousquet-Mélou CNRS, LaBRI, Bordeaux, France

joint work with Kerstin Weller, Oxford

• Simple graphs on the vertex set $\{1, 2, \ldots, n\}$

- Simple graphs on the vertex set $\{1, 2, \ldots, n\}$
- A set of graphs is a class if it is closed under relabelling of the vertices.

- Simple graphs on the vertex set $\{1, 2, \ldots, n\}$
- A set of graphs is a class if it is closed under relabelling of the vertices.
- A minor of an unlabelled graph G is obtained by contracting edges, and deleting edges and isolated vertices.

- Simple graphs on the vertex set $\{1, 2, \ldots, n\}$
- A set of graphs is a class if it is closed under relabelling of the vertices.
- A minor of an unlabelled graph G is obtained by contracting edges, and deleting edges and isolated vertices.
- A class \mathcal{A} is minor-closed if all minors of graphs of \mathcal{A} are in \mathcal{A} .

- Simple graphs on the vertex set $\{1, 2, \ldots, n\}$
- A set of graphs is a class if it is closed under relabelling of the vertices.
- A minor of an unlabelled graph G is obtained by contracting edges, and deleting edges and isolated vertices.
- A class \mathcal{A} is minor-closed if all minors of graphs of \mathcal{A} are in \mathcal{A} .
- Examples: \mathcal{A} can be the set of forests, the set of planar graphs...

- Simple graphs on the vertex set $\{1, 2, \ldots, n\}$
- A set of graphs is a class if it is closed under relabelling of the vertices.
- A minor of an unlabelled graph G is obtained by contracting edges, and deleting edges and isolated vertices.
- A class \mathcal{A} is minor-closed if all minors of graphs of \mathcal{A} are in \mathcal{A} .
- Examples: \mathcal{A} can be the set of forests, the set of planar graphs...
- Theorem [Robertson & Seymour]: every minor-closed class of graphs can be described by excluding a finite number of minors.

- Simple graphs on the vertex set $\{1, 2, \ldots, n\}$
- A set of graphs is a class if it is closed under relabelling of the vertices.
- A minor of an unlabelled graph G is obtained by contracting edges, and deleting edges and isolated vertices.
- A class \mathcal{A} is minor-closed if all minors of graphs of \mathcal{A} are in \mathcal{A} .
- Examples: \mathcal{A} can be the set of forests, the set of planar graphs...
- Theorem [Robertson & Seymour]: every minor-closed class of graphs can be described by excluding a finite number of minors.
- Examples: For forests, exclude the 3-cycle. For planar graphs, exclude K_5 and $K_{3,3}$.

Random graphs

For fixed (and large) n, let A_n be the set of graphs of the class A having size n (i.e., n vertices). Let G_n be a random graph of A_n , taken uniformly at random: if G has size n,

$$\mathbb{P}(\mathcal{G}_n = G) = \frac{1}{a(n)}$$

where $a(n) = |\mathcal{A}_n|$.

A uniform random forest on 75 vertices (2 components)



Some properties of (proper) minor-closed classes

• No excluded minor:

$$a(n) = 2^{\binom{n}{2}}.$$

The expected number of edges is quadratic:

$$\mathbb{E}(e(\mathcal{G}_n)) = \frac{n(n-1)}{4}.$$



Some properties of (proper) minor-closed classes

• No excluded minor:

$$a(n) = 2^{\binom{n}{2}}.$$

The expected number of edges is quadratic:

$$\mathbb{E}(e(\mathcal{G}_n)) = \frac{n(n-1)}{4}.$$



• At least one excluded minor:

 $a(n) \le n! \mu^n$

for some constant μ [Norine, Seymour, Thomas & Wollan 06].

The expected number of edges is at most linear: for any graph of A_n ,

 $e(G) \le \alpha n$

for a constant α that depends on the excluded minors. [Kostochka 82, Thomason 84]

Refined properties: The number and size of components

- What is the probability that \mathcal{G}_n is connected?
- More generally, what is the distribution of N_n , the number of components?

Refined properties: The number and size of components

- What is the probability that \mathcal{G}_n is connected?
- More generally, what is the distribution of N_n , the number of components?
- What is the size S_n of the root component, that is, the component containing the vertex 1?
- What is the size L_n of the largest component?

Refined properties: The number and size of components

- What is the probability that \mathcal{G}_n is connected?
- More generally, what is the distribution of N_n , the number of components?
- What is the size S_n of the root component, that is, the component containing the vertex 1?
- What is the size L_n of the largest component?

Remark: if no minor is excluded, the probability that \mathcal{G}_n is connected tends to 1 as $n \to \infty$, and these questions have simple answers.

Generating functions

• Let a(n) be the number of graphs of size n in \mathcal{A} , and let

$$A(z) = \sum_{n \ge 0} a(n) \frac{z^n}{n!}$$

be the associated exponential generating function.

• Use similar notation (c(n) and C(z)) for connected graphs of \mathcal{A} .

 \bullet If all forbidden minors are connected, graphs of ${\cal A}$ are arbitrary unions of graphs of ${\cal C},$ and

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

When all excluded minors are 2-connected

Theorem [McDiarmid 09]: If all excluded minors are 2-connected, then

- C(z) and A(z) converge at their (common) radius of convergence ρ
- the probability that \mathcal{G}_n is connected tends to $1/A(\rho),$ which belongs to $[1/\sqrt{e},1)$





Forests: $1/A(\rho) = 1/\sqrt{e} = 0.60...$

Planar graphs: $1/A(\rho) = 0.97...$ [Gimenez & Noy 09]

When all excluded minors are 2-connected

Theorem [McDiarmid 09]: If all excluded minors are 2-connected, then

- C(z) and A(z) converge at their (common) radius of convergence ρ
- the probability that \mathcal{G}_n is connected tends to $1/A(\rho)$, which belongs to $[1/\sqrt{e},1)$
- in fact, $N_n 1$ converges to a Poisson distribution of parameter $C(\rho)$:

$$\mathbb{P}(N_n = i+1) \to \frac{C(\rho)^i}{i!A(\rho)}$$

When all excluded minors are 2-connected

Theorem [McDiarmid 09]: If all excluded minors are 2-connected, then

- C(z) and A(z) converge at their (common) radius of convergence ρ
- the probability that \mathcal{G}_n is connected tends to $1/A(\rho)$, which belongs to $[1/\sqrt{e},1)$ (*)
- in fact, $N_n 1$ converges to a Poisson distribution of parameter $C(\rho)$:

$$\mathbb{P}(N_n = i+1) \to \frac{C(\rho)^i}{i!A(\rho)}$$

• the root component contains almost all vertices; more precisely,

$$\mathbb{P}(S_n = n - k) \to \frac{1}{A(\rho)} \frac{a(k)\rho^k}{k!}$$

• the same holds for the largest component.

(*) [Addario Berry, McDiarmid & Reed 11], [Kang & Panagiotou 11]

What if some excluded minors are not 2-connected?

The above properties cannot always hold.

Example: exclude the one-edge graph. Then

- \mathcal{G}_n is never connected (for $n \geq 2$),
- the number of components is maximal: $N_n = n$,
- the largest component has size $L_n = 1$.

More components, of a smaller size

Our results: a collection of examples

For a collection of minor-closed classes avoiding some non-2-connected minors,

- exact determination of C(z) and A(z),
- detailed asymptotics properties of N_n (number of components),
- detailed asymptotics properties of S_n (size of the root component),
- for some classes, detailed asymptotics properties of L_n (size of the largest component)

Key point: behaviour of C(z) near its radius ρ

Note: all forbidden minors are connected, so that $A(z) = \exp(C(z))$

Excluded	2-conn.				all conn. graphs
minors	or	Φ and $igta$		or	of size $k + 1$
				Δ and λ	
C(ho)	$<\infty$	$\log rac{1}{1-z/ ho}$	$rac{1}{\sqrt{1\!-\!z/ ho}}$	$rac{1}{1-z/ ho}$	entire
ℙ(conn.)	ightarrow p > 0	$\rightarrow 0$	$\rightarrow 0$	$\rightarrow 0$	$\rightarrow 0$
N_n	<i>O</i> (1) Poisson	log n gaussian	n ^{1/3} gaussian	n ^{1/2} gaussian	$\frac{n/k}{gaussian}$ (*)
S_n	$n - F_n$ $F_n o ext{discr.}$	$n = \frac{1}{4}(1-x)^{-3/4}$	$n^{2/3}$ $2\sqrt{\frac{x}{\pi}}e^{-x}$	$n^{1/2}$ xe^{-x}	<mark>k</mark> Dirac
L_n	$n - F_n$ $F_n o ext{discr.}$	<i>n</i> Dickman	?	n ^{1/2} log n Gumbel	<mark>k</mark> Dirac

(*) [Canfield 77]

Example: forests of paths (no \triangle nor \checkmark)

• Exact enumeration:

$$C(z) = \frac{z(2-z)}{2(1-z)}, \quad A(z) = \exp(C(z))$$

• Asymptotic enumeration:

$$c_n = \frac{n!}{2}$$
 and $a_n \sim \frac{n!}{2} \frac{(e/2)^{1/4}}{\sqrt{\pi}n^{3/4}} e^{\sqrt{2n}}.$

In particular, $\frac{c(n)}{a(n)} \rightarrow 0$.

Example: forests of paths (no \triangle nor \checkmark)

• Exact enumeration:

$$C(z) = \frac{z(2-z)}{2(1-z)}, \quad A(z) = \exp(C(z))$$

• Asymptotic enumeration:

$$c_n = \frac{n!}{2}$$
 and $a_n \sim \frac{n!}{2} \frac{(e/2)^{1/4}}{\sqrt{\pi}n^{3/4}} e^{\sqrt{2n}}.$

In particular, $\frac{c(n)}{a(n)} \rightarrow 0$.

• The mean and variance of N_n satisfy:

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

and the normalized random variable $\frac{N_n - \sqrt{n/2}}{(n/8)^{1/4}}$ converges in law to a standard normal distribution:

$$\mathbb{P}\left(\frac{N_n - \sqrt{n/2}}{(n/8)^{1/4}} \le y\right) \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-x^2/2} dx.$$

• The mean and variance of S_n satisfy:

$$\mathbb{E}(S_n) \sim 2\sqrt{2n}, \qquad \mathbb{V}(S_n) \sim 4n,$$

and the normalized variable $S_n/\sqrt{2n}$ converges in law to a Gamma(2):

$$\mathbb{P}\left(S_n/\sqrt{2n} \le y\right) \to \int_0^y x e^{-x} dx.$$

A local limit law also holds:

$$\sqrt{2n} \mathbb{P}(S_n = \lfloor x\sqrt{2n} \rfloor) \sim xe^{-x}.$$

• The mean and variance of S_n satisfy:

$$\mathbb{E}(S_n) \sim 2\sqrt{2n}, \qquad \mathbb{V}(S_n) \sim 4n,$$

and the normalized variable $S_n/\sqrt{2n}$ converges in law to a Gamma(2):

$$\mathbb{P}\left(S_n/\sqrt{2n} \le y\right) \to \int_0^y x e^{-x} dx.$$

A local limit law also holds:

$$\sqrt{2n} \mathbb{P}(S_n = \lfloor x\sqrt{2n} \rfloor) \sim xe^{-x}.$$

• The normalized variable $\frac{L_n - \sqrt{n/2} \log n}{\sqrt{n/2}}$ converges in distribution to a Gumbel law:

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

Excluded	2-conn.				all conn. graphs
minors	or	Φ and $igta$		or	of size $k + 1$
				Δ and λ	
C(ho)	$<\infty$	$\log rac{1}{1-z/ ho}$	$rac{1}{\sqrt{1\!-\!z/ ho}}$	$rac{1}{1-z/ ho}$	entire
ℙ(conn.)	ightarrow p > 0	$\rightarrow 0$	$\rightarrow 0$	$\rightarrow 0$	$\rightarrow 0$
N_n	<i>O</i> (1) Poisson	log n gaussian	n ^{1/3} gaussian	n ^{1/2} gaussian	$\frac{n/k}{gaussian}$ (*)
S_n	$n - F_n$ $F_n o ext{discr.}$	$n = \frac{1}{4}(1-x)^{-3/4}$	$n^{2/3}$ $2\sqrt{\frac{x}{\pi}}e^{-x}$	$n^{1/2}$ xe^{-x}	<mark>k</mark> Dirac
L_n	$n - F_n$ $F_n o ext{discr.}$	<i>n</i> Dickman	?	n ^{1/2} log n Gumbel	<mark>k</mark> Dirac

(*) [Canfield 77]

Techniques: Analytic combinatorics





Philippe Flajolet, Robert Sedgewick

A. Exact enumeration

Write directly equations from the recursive description of the objects

Construction		Numbers	Generating function	
Union	$\mathcal{A} = \mathcal{B} \sqcup \mathcal{C}$	a(n) = b(n) + c(n)	A(z) = B(z) + C(z)	
Product	$\mathcal{A} = \mathcal{B} \times \mathcal{C}$	$a(n) = \sum_{i} {n \choose i} b(i) c(n-i)$	$A(z) = B(z) \cdot C(z)$	
$ (\beta,\gamma) = \beta + \gamma $				
Sequence	$\mathcal{A} = \operatorname{Seq}(\mathcal{B})$		$A(z) = \frac{1}{1 - B(z)}$	
Set	$\mathcal{A} = Set(\mathcal{B})$		$A(z) = \exp\left(B(z)\right)$	
Cycle (directed)	$\mathcal{A} = Cyc(\mathcal{B})$		$A(z) = \log \frac{1}{1 - B(z)}$	

Constructible labelled objects

$$A(z) = \sum_{n} a(n) \frac{z^{n}}{n!}$$

A basic example: forests (no \triangle)

• A (vertex-)rooted tree is a root vertex (z) and a set of (sub-)trees (T(z) for each):

 $T(z) = z \exp(T(z)).$

• A tree with n nodes has n - 1 edges, and edge-rooted trees consist of a pair of vertex-rooted trees:

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

• A forest is a set of unrooted trees:

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





The number and size of components

• The generating function of forests having k components is

 $\frac{C(z)^k}{k!}.$

• The number of forests of \mathcal{A} with n vertices and root component of size k:

$${n-1 \choose k-1}c(k)a(n-k)$$

$$\Rightarrow \mathbb{P}(S_n = k) = \frac{c(k)a(n-k)\binom{n-1}{k-1}}{a(n)}.$$

B. Asymptotic enumeration and limit laws

How to extract from a series the asymptotic behaviour of its coefficients?

A general correspondence between

the singular expansion of a series A(z) near its dominant singularities

and

the asymptotic expansion of the *n*th coefficient a(n) of this series

Example: If the dominant singularity has modulus ρ , then the coefficients satisfy

 $\limsup a(n)^{1/n} = 1/\rho.$

• The generating functions of rooted trees, unrooted trees and forests are given by

$$T(z) = z \exp(T(z)), \quad C(z) = T(z) - \frac{1}{2}T(z)^2, \quad A(z) = \exp(C(z))$$

• The generating functions of rooted trees, unrooted trees and forests are given by

$$T(z) = z \exp(T(z)), \quad C(z) = T(z) - \frac{1}{2}T(z)^2, \quad A(z) = \exp(C(z))$$

• All have radius $\rho = 1/e$, with singular expansions

$$T(z) = 1 - \sqrt{2}(1 - ze)^{1/2} + \frac{2}{3}(1 - ze) - \frac{11}{18\sqrt{2}}(1 - ze)^{3/2} + O((1 - ze)^2),$$

• The generating functions of rooted trees, unrooted trees and forests are given by

$$T(z) = z \exp(T(z)), \quad C(z) = T(z) - \frac{1}{2}T(z)^2, \quad A(z) = \exp(C(z))$$

• All have radius $\rho=1/e,$ with singular expansions

$$T(z) = 1 - \sqrt{2}(1 - ze)^{1/2} + \frac{2}{3}(1 - ze) - \frac{11}{18\sqrt{2}}(1 - ze)^{3/2} + O((1 - ze)^2),$$

$$C(z) = 1/2 - (1 - ze) + \frac{2\sqrt{2}}{3}(1 - ze)^{3/2} + O((1 - ze)^2),$$

$$A(z) = e^{1/2} \left(1 - (1 - ze) + \frac{2\sqrt{2}}{3}(1 - ze)^{3/2}\right) + O((1 - ze)^2)$$

• The generating functions of rooted trees, unrooted trees and forests are given by

$$T(z) = z \exp(T(z)), \quad C(z) = T(z) - \frac{1}{2}T(z)^2, \quad A(z) = \exp(C(z))$$

• All have radius $\rho = 1/e$, with singular expansions

$$T(z) = 1 - \sqrt{2}(1 - ze)^{1/2} + \frac{2}{3}(1 - ze) - \frac{11}{18\sqrt{2}}(1 - ze)^{3/2} + O((1 - ze)^2),$$

$$C(z) = 1/2 - (1 - ze) + \frac{2\sqrt{2}}{3}(1 - ze)^{3/2} + O((1 - ze)^2),$$

$$A(z) = e^{1/2} \left(1 - (1 - ze) + \frac{2\sqrt{2}}{3}(1 - ze)^{3/2}\right) + O((1 - ze)^2)$$

• By singularity analysis,

$$c(n) \sim n! rac{e^n}{\sqrt{2\pi}n^{5/2}}$$
 and $a(n) \sim e^{1/2}c(n).$

The probability that \mathcal{G}_n is connected tends to $e^{-1/2}$.

Forests: the number N_n of components

• We have

$$\mathbb{P}(N_n = k) = \frac{1}{k!} \frac{[z^n] C(z)^k}{a(n)}.$$

Forets: the number N_n of components

• We have

$$\mathbb{P}(N_n = k) = \frac{1}{k!} \frac{[z^n] C(z)^k}{a(n)}.$$

With

$$C(z) = 1/2 - (1 - ze) + \frac{2\sqrt{2}}{3}(1 - ze)^{3/2} + O((1 - ze)^2),$$

this gives

$$C(z)^{k} = \frac{1}{2^{k}} + \alpha(1 - ze) + \frac{k}{2^{k-1}} \frac{2\sqrt{2}}{3} (1 - ze)^{3/2} + O((1 - ze)^{2}),$$

Forets: the number N_n of components

• We have

$$\mathbb{P}(N_n = k) = \frac{1}{k!} \frac{[z^n] C(z)^k}{a(n)}.$$

With

$$C(z) = 1/2 - (1 - ze) + \frac{2\sqrt{2}}{3}(1 - ze)^{3/2} + O((1 - ze)^2),$$

this gives

$$C(z)^{k} = \frac{1}{2^{k}} + \alpha(1 - ze) + \frac{k}{2^{k-1}} \frac{2\sqrt{2}}{3} (1 - ze)^{3/2} + O((1 - ze)^{2}),$$

so that

$$\mathbb{P}(N_n = k) \to \frac{1}{2^{k-1}} \frac{1}{(k-1)!} \frac{1}{\sqrt{e}}$$

and $N_n - 1$ follows a Poisson distribution of parameter 1/2.

Forests: the size of the root component

We have found

$$c(n) \sim n! rac{e^n}{\sqrt{2\pi}n^{5/2}}$$
 and $a(n) \sim e^{1/2}c(n).$

Thus for any fixed k,

$$\mathbb{P}(S_n = n - k) = \frac{c(n-k)a(k)\binom{n-1}{k}}{a(n)} \rightarrow \frac{1}{e^{1/2}}\frac{a(k)}{k!e^k}.$$

With the same tools: when \mathcal{A} is dominated by forests

Definition: the class ${\mathcal A}$ is dominated by forests if

$$C(z) = T(z) - T(z)^2/2 + D(z),$$

where D(z) has radius of convergence strictly larger than 1/e (the radius of T).

Proposition: In this case,

$$c(n) \sim n! \frac{e^n}{\sqrt{2\pi}n^{5/2}}$$
 and $a(n) \sim A(1/e)c(n),$

- $N_n 1$ converges in law to a Poisson distribution of parameter C(1/e) = 1/2 + D(1/e),
- the root component contains almost all vertices, and

$$\mathbb{P}(S_n = n - k) = \frac{1}{A(1/e)} \frac{a(k)e^{-k}}{k!},$$

• the same holds for the largest component.

When \mathcal{A} is dominated by forests: examples

Proposition: let $k \ge 1$. Let \mathcal{A} be a minor-closed class of graphs containing all trees, but not the k-spoon. Then \mathcal{A} is dominated by forests:

$$C(z) = T(z) - T(z)^2/2 + D(z),$$

where $\rho(D) > 1/e$.



A more complex example: Excluding the bowtie \bowtie

- Analysis of the structure of such graphs, and partition into disjoint sub-classes
- Generating function of connected graphs:

Asymptotic results and limit laws via the saddle point method



C. Random generation: Boltzmann samplers

A class \mathcal{A} of (labelled) objects that has a "simple" recursive decription and a moderate growth (at most exponential) can often be sampled efficiently according to the associated Boltzmann distribution: an object G of size n has probability

$$\mathbb{P}(G) = \frac{x^n}{n! A(x)}$$

where $A(x) = \sum_{n \in \mathbb{N}} a(n) \frac{x^n}{n!}$ and $x \in (0, \rho]$.

In particular, all objects of size n have the same probability.

[Duchon, Flajolet, Louchard, Schaeffer 04]

A random planar graph of size 727

[Fusy 09] (100,000 vertices in a few seconds)



 $\mathbb{E}(N_n) = \Theta(1) \qquad \mathbb{E}(S_n) \sim n$

A random forest of size 745



A random graph of size 975 avoiding the diamond and the bowtie



A random graph of size 956 avoiding the bowtie



 $\mathbb{E}(N_n) \simeq n^{1/3} \qquad \mathbb{E}(S_n) \simeq n^{2/3}$

A random forest of paths of size 597 (no \triangle nor \checkmark)



(In the test of caterpillars of size 486 (no Δ nor Λ)



A random graph of maximal degree 2 of size 636 (no \checkmark)



 $\mathbb{E}(N_n) \simeq n^{1/2} \qquad \mathbb{E}(S_n) \simeq n^{1/2}$

Final comments

- We do not cover McDiarmid's results (2-connected excluded minors)
- Other classes? (outer-planar graphs with an additional excluded minor...)
- Other parameters? number of edges, distribution of vertex degrees...
- Describe conditions on the excluded minors that determine the nature of the singularities of C(z). For instance,
 - 2-connected minors $\Rightarrow C(\rho)$ finite
 - exclude one spoon but not tree $\Rightarrow C(\rho)$ finite with a singularity in $(1-ze)^{3/2}$

Excluded	2-conn.				all conn. graphs
minors	or	Φ and $igNeta$		or	of size $k + 1$
				Δ and λ	
C(ho)	$<\infty$	$\log rac{1}{1-z/ ho}$	$rac{1}{\sqrt{1\!-\!z/ ho}}$	$rac{1}{1-z/ ho}$	entire
₽(conn.)	ightarrow p > 0	ightarrow 0	$\rightarrow 0$	ightarrow 0	$\rightarrow 0$
N_n	<i>O</i> (1) Poisson	log n gaussian	n ^{1/3} gaussian	n ^{1/2} gaussian	n/kgaussian (*)
S_n	$n - F_n$ $F_n o ext{discr.}$	$n = \frac{1}{4}(1-x)^{-3/4}$	$n^{2/3}$ $2\sqrt{\frac{x}{\pi}}e^{-x}$	$n^{1/2}$ xe^{-x}	<mark>k</mark> Dirac
L_n	$n - F_n$ $F_n o ext{discr.}$	<i>n</i> Dickman	?	n ^{1/2} log n Gumbel	<mark>k</mark> Dirac

Singularity analysis

Saddle point method