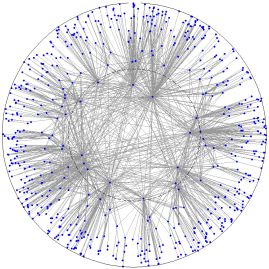
Analytical Approach of Sparse Random Graphs Phase Transitions

VLADY RAVELOMANANA

IRIF – University of Paris 7.

vlad@irif.fr

Saint-Etienne



Outline of the talk

- Introduction & motivations.
- Some tractable graph phase transitions.
- Shifting the critical threshold.
- Open problems.
- Conclusion and perspectives.

Main motivations

In Computer Science, two core problems (cf. [Garey, Johnson'79]).

- Decision problems: write an algorithm such that

  INPUT : an instance $\mathcal{I}$ and a property $\mathcal{P}$
  OUTPUT : YES (resp. NO) if $\mathcal{I}$ satisfies (resp. does not satisfy) $\mathcal{P}$

- Optimization problems: find the **best solution** from all feasible solutions.

Other communities: Probability/Combinatorics/Physics

C:configuration of the system, **E:**energy function, **T:**temperature, **Z:**normalization cf. [Baldassi, Braunstein, Ramezanpour, Zecchina'09]

$$\mathbb{P}(C) = \frac{1}{Z} e^{-E(C)/T}$$

- if $T = \infty$ then all configurations C are equiprobable, the system is "disordered"
- if $T = 0$ $\mathbb{P}(C)$ is "concentrated on the minimum energy function" (the *ground state*). Optimizing $\rightarrow$ finding a zero energy ground state of E.

Translating CS-langage to Physics-langage: the k-SAT example

Computer Science	Statistical Physics
Boolean variable $x \in \{\text{True, False}\}$	Ising spin $s \in \{\text{spin up}(+1), \text{Spin down}(-1)\}$
Clauses	Couplings and fields acting on spins
Number of clauses violated	Energy E of spins configuration
2-SAT : (x or $\bar{y}$) and ($\bar{x}$ or z)	$E = \frac{1}{4}(1 - s_x)(1 + s_y) + \frac{1}{4}(1 + s_x)(1 - s_z)$
3-SAT : (x or $\bar{y}$ or z)	$E = \frac{1}{8}(1 - s_x)(1 + s_y)(1 - s_z)$
The problem is SAT	Ground state energy $= 0$
The problem is UNSAT	Ground state energy > 0

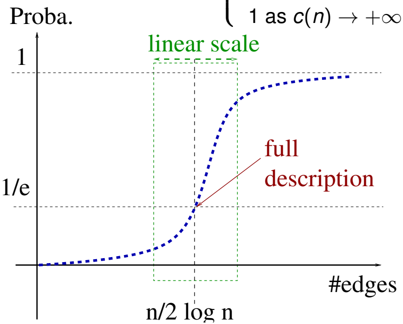
From **[MONASSON'02]** in ALEA lectures (CIRM).

The connectivity phase transition: a typical theorem

From "*On Random Graphs I*" [ERDŐS, RÉNYI 59]

Theorem. As $M = \frac{1}{2} n \Big(\log n + c(n) \Big)$

$$\lim_{n \to +\infty} \texttt{Proba}\,[G(n,M) \text{ is connected}] = \begin{cases} 0 \text{ as } c(n) \to -\infty \\ e^{-e^{-c}} \text{ as } c(n) \to c \text{ (constant)} \\ 1 \text{ as } c(n) \to +\infty. \end{cases}$$



Main research directions

- Random *k*-**SAT** formulas ($k > 2$) are subject to phase transition phenomena [FRIEDGUT,BOURGAIN'99]
- Main research tasks include:
  1. **Localization** of the threshold (ex: 3-**SAT** $4.2\cdots$??? 3-XORSAT $0.91\cdots$ proved in [DUBOIS, MANDLER'03]).
  2. Nature of the transition: **sharp/coarse**. [CREIGNOU, DAUDÉ'09]
  3. **Scaling window** (e.g. 2-SAT [BOLLOBÀS, BORGS, KIM, WILSON'01]) and/or (if possible) details inside the **window of transition**.
  4. **Algorithmic complexity** of the decision problem (e.g. "2-SAT is in P" [TARJAN '79]). **Only tractable problems today!**

Outline of the talk

- Introduction & motivations.
- Some tractable graph phase transitions.
- Shifting the critical threshold.
- Open problems.
- Conclusion and perspectives.

- **2-COL (bipartiteness):** is a graph **2−COLORABLE**?
- **2-XORSAT :** is the formula

$$\begin{cases} x_1 \oplus x_2 = 1 \\ x_2 \oplus x_3 = 0 \\ \dots \\ x_i \oplus x_j = \varepsilon \in \{0, 1\} \\ \dots \end{cases} \qquad \texttt{SAT}?$$

- **2-SAT:** is the formula

$$\begin{cases} \dots \\ r \vee s \quad \texttt{SAT}? \\ \dots \end{cases}$$

  (where r and $s \in \{x_1, \cdots, x_n\} \bigcup \{\bar{x_1}, \cdots, \bar{x_n}\}$)
- **Planarity :** is a graph **planar**?
- $\dots$

Random 2-XORSAT [Daudé, R. '11]

- **General form :**

$$A.X = \varepsilon$$

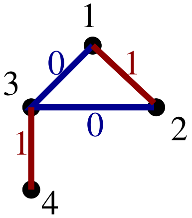
  where A has m rows and ε a m-dimensional $0/1$ vector.
- **Distribution :** uniform. Pick m clauses of the form $x_i \oplus x_j \in \{0, 1\}$ from the set of $n(n-1)$ clauses.
- **Underlying structures :** `graphs with weighted edges`.

SAT characterization

SAT iff no elementary cycle of odd weight.

SAT iff no elementary cycle of odd weight

$$\begin{cases} x_1 \oplus x_2 = 1 \\ x_2 \oplus x_3 = 0 \\ x_1 \oplus x_3 = 0 \\ x_3 \oplus x_4 = 1 \end{cases}$$



Proof.

- UNSAT $\Longleftarrow$ Fix a cycle of odd weight, ...
- SAT $\Longleftarrow$ No cycles of odd weight, a DFS affectation based proof.

Combinatorics

A basic scheme

- Enumeration of "SAT"-graphs (graphs without cycles of odd weight by means of generating functions.

- Use the results to compute

$$\text{PROBA SAT} = \frac{\sharp\text{CONFIG. WITHOUT CYCLES OF ODD WEIGHT}}{\sharp\text{TOTAL CONFIGURATIONS}}.$$

Works on 2-XORSAT

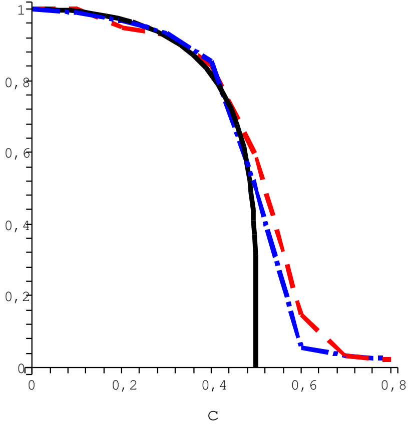
- Using statistical physics methods, [MONASSON '07] inferred that

$$\lim_{n\to+\infty} n^{\text{critical exponent}} \times \mathbb{P}\left[\text{2-XORSAT}\left(n, \frac{n}{2}\right)\right] = \mathbf{O(1)}$$

where "critical exponent" must be $\frac{1}{12}$.

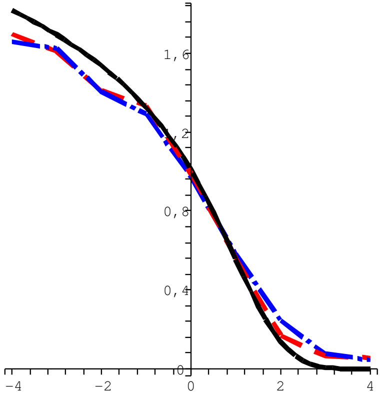
- In [DAUDÉ, R. '11], it is shown that **YES** critical exponent is $1/12$ and the " **O(1)**"are explicitly computed.

Taste of our results : the whole window



p(n, cn) $\stackrel{\text{def}}{=}$ `Proba` **[2** – XOR with **n** variables , **cn** clauses] is **SAT**
for $n = 1000$, $n = 2000$ and the **theoretical** function : $\mathbf{e^{c/2}(1-2c)^{1/4}}$.

Taste of our results: rescaling the critical window



Rescaling at the point “zero”, i.e $c = 1/2$: $n = 1000$, $n = 2000$ and $\lim_{n\to\infty} \; n^{1/12}\times \; p(n, n/2 + \mu n^{2/3})$ as a **function of** μ.

Basic refresher of Symbolic Method on EGFs

The **EXPONENTIAL GENERATING FUNCTION** (EGF) associated to a **labelled class** $\mathcal{A}$ of combinatorial objects is

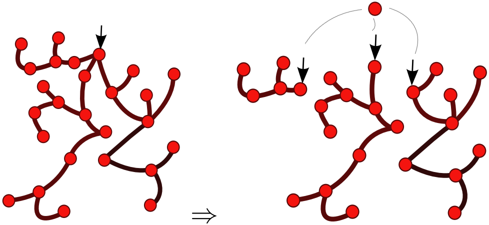
$$A(z) = \sum_{a \in \mathcal{A}} \frac{z^{|a|}}{|a|!} = \sum_{n \geq 0} a_n \frac{z^n}{n!} .$$

cf [FLAJOLET – SEDGEWICK'09]

Construction	Notation	Comments	EGFs
Disjoint union	$\mathcal{A} + \mathcal{B}$	disjoint copies of objects from $\mathcal{A}$ and $\mathcal{B}$	$A(z) + B(z)$
Labeled product	$\mathcal{A} \star \mathcal{B}$	ordered pairs of copies one from $\mathcal{A}$ and one from $\mathcal{B}$	$A(z)B(z)$
Sequence	$\mathrm{Seq}(\mathcal{A})$	sequences of objects from $\mathcal{A}$	$\frac{1}{1-A(z)}$
Set	$\mathrm{Set}(\mathcal{A})$	set of objects from $\mathcal{A}$	$e^{A(z)}$
Cycle	$\mathrm{Cyc}(\mathcal{A})$	cycles of k objects from $\mathcal{A}$	$\log \frac{1}{1-A(z)}$

Trees

We apply the previous grammar to count *rooted* trees



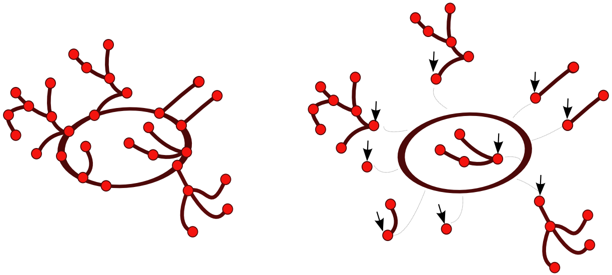
$$\mathcal{T} = \bullet \times \mathrm{Set}(\mathcal{T}) \to T(x) = xe^{T(x)}$$

To forget the root, we just integrate: $(xU'(x) = T(x))$

$$\int_0^x \frac{T(s)}{s} ds = \quad \begin{array}{c} T(s) = u \\ T'(s)\,ds = du \end{array} \quad = \int_{T(0)}^{T(x)} 1 - u\,du = T(x) - \frac{1}{2}T(x)^2$$

and the general version

Unicyclic graphs



$$\mathcal{V} = \bigcirc_{\geq 3}(\mathcal{T}) \to V(x) = \sum_{n=3}^{\infty} \frac{1}{2} \frac{(n-1)!}{n!} (T(x))^n$$

We can write $V(x)$ in a compact way:

$$\frac{1}{2} \quad -\log\left(1 - T(x)\right) - T(x) - \frac{T(x)^2}{2} \quad \to e^{V(x)} = \frac{e^{-T(x)/2 - T(x)^2/4}}{\sqrt{1 - T(x)}}.$$

Connected graphs without cycles of odd weight

We want to **enumerate**

- **labelled connected graphs** (nodes labelled in $[1, n]$) with edges of weight in $\{0, 1\}$ (cf. **[Harary, Palmer '73]**)
- according to two parameters
  1. number of vertices n
  2. number of edges $n + k$ (k is the excess of the component, $k \geq -1$).
- SAT $=$ **those without cycles of odd weight.**

Let

$$C_k(z) = \sum_{k \geq 0} c_{n,n+k} \frac{z^n}{n!} .$$

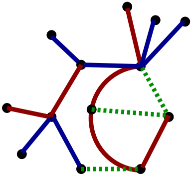
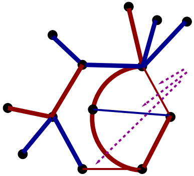
Th. [Daudé, R. 11]

Let Wright(z) be the number of unweighted connected graphs of excess k (**[Wright '77]**). Then

$$C_k(z) = \frac{1}{2} \, \textsc{ConnectedGraphs}_k(2z) .$$

Proof.

- Recall *SAT* iff **no cycles of odd weight**
- Consider a "SAT" connected component g
- Pick a spanning tree of g with n nodes:

- The edges of the spanning tree can be weighted in 2^{n-1} ways. By **[Diestel'00]**, the weights of the other edges are **determined:**

$$C_k(z) = \sum_{k^0} 2^{n-1} \underbrace{\textsc{connected graphs}(n, n+k)}_{\textbf{[Wright '77]}} \frac{z^n}{n!} .$$

Combinatorics of random 2-XORSAT

Main ideas: n variables, m clauses

$$\textsc{Proba}\ \mathtt{SAT} = n! \times \frac{[z^n]\mathtt{SAT\text{-}GF}(z)}{\binom{n(n-1)}{m}}.$$

As the number of clauses m increases, random formulas behave as random graphs with **gazeous**, **liquid** and **hardening** phases.

- **Sub-critical phase (gaz) :** forest of trees and set of unicyclic components.
- **Critical phase (liquid) :** forest of trees, set of unicyclic components and few multicyclic components.
- **Super-critical phase (hardening) :** forest of trees, set of unicyclic components and a single (baby) giant component.

The Random 2-XORSAT Transition

Th. [Daudé, R. '11]

The probability that a random formula with n variables and m clauses is **SAT** satisfies the following :

(i) Sub-critical phase. As $0 < n - 2m \ll n^{2/3}$

$$\mathbb{P}(n,\, m) = e^{m/2n}\left(1 - 2\frac{m}{n}\right)^{1/4} + O\left(\frac{n^2}{(n-2m)^3}\right) .$$

(ii) Critical phase. As $m = \frac{n}{2} + \mu n^{2/3}$, μ fixed real

$$\lim_{n \to +\infty} n^{1/12}\mathbb{P}\left(n,\, \frac{n}{2}(1 + \mu n^{-1/3})\right) = \Psi(\mu),$$

where Ψ can be explicitly expressed (in terms of the Airy function).

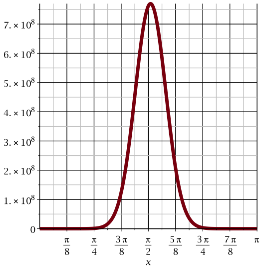
(iii) Super-critical phase. As $m = \frac{n}{2} + \mu n^{2/3}$ with $\mu = o(n^{1/12})$

$$\mathbb{P}\left(n,\, \frac{n}{2}(1 + \mu n^{-1/3})\right) = \mathrm{Poly}(n, \mu) e^{-\frac{\mu^3}{6}} .$$

The classical saddle-point method

The real case looks like (Laplace's method):

$$R(n) = \int_{x_1}^{x_2} g(x) e^{nh(x)} dx \,.$$



$$\int_0^{\pi} x e^{20 \sin x} dx = 4.29 \times 10^8 \,.$$

$$\int_0^{\pi} x e^{20(1-\frac{1}{2}(x-\pi/2)^2)} dx = 4.27 \times 10^8 \,.$$

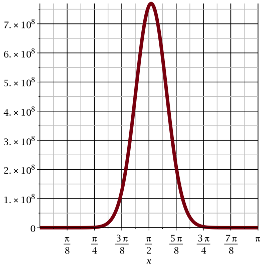
Normal saddle point x_0 s.t $h'(x_0) = 0$ and $h''(x_0) \neq 0$.

What if $h''(x_0) = 0$?

The classical saddle-point method

The real case looks like (Laplace's method):

$$R(n) = \int g(x) e^{nh(x)} dx \; g(x_0) \qquad \sim \qquad \sqrt{\frac{2\pi}{-h''(x_0)n}}\, e^{nh(x_0)}$$



$$\int_0^{\pi} x e^{20 \sin x} dx = 4.29 \times 10^8 .$$

$$\int_0^{\pi} x e^{20(1-\frac{1}{2}(x-\pi/2)^2)} dx = 4.27 \times 10^8 .$$

Normal saddle point x_0 s.t $h'(x_0) = 0$ and $\underline{\underline{h''(x_0) \neq 0}}$.

The case with $h''(z_0) = 0$

Consider

$$\Psi''(z) - z\Psi(z) = 0 \qquad (\textbf{Airy equation}).$$

Look for a solution of the form

$$\Psi(z) = \int_\Gamma e^{tz}\Phi(t)dt \textbf{ with } \Gamma \textbf{ in the complex } t - \textbf{plane}$$

(Φ with no singularities on Γ.)

$$\Psi''(z) = \int_\Gamma t^2 e^{tz}\Phi(t)dt \qquad \textbf{and} \qquad z\Psi(z) = [e^{tz}\Phi(t)]_{\delta\Gamma} - \int_\Gamma e^{tz}\Phi'(t)dt\,.$$

We have then

$$t^2\Phi(t) + \Phi'(t) = 0 \textbf{ leading to solution of the form } \Psi(z) = c\int_\Gamma \exp\left(tz - \frac{t^3}{3}\right)dt$$

(**plus** the condition $e^{tz}\Phi(t) = 0$ on the boundary of the contour Γ).

Proof of the Theorem : the sub-critical phase

1. As $0 < n - 2m \ll n^{2/3}$,the probability that the underlying graph has no multicyclic component is
$$1 - O\left(\frac{n^2}{(n-2m)^3}\right).$$

2. Then, the probability that the random formula is SAT is close to
$$\frac{n!}{\binom{n(n-1)}{m}}[z^n]\frac{(\text{unrooted trees}(z))^{n-m}}{(n-m)!} \times \text{set of even weighted unicycles}(z).$$
That is Stirling$(n, m) \times$ Cauchy(z) with
$$\text{Cauchy}(z) = \frac{1}{2\pi i}\oint\left(\frac{T(2z)}{2} - \frac{T(2z)^2}{4}\right)^{n-m}\frac{e^{-T(2z)/4 - T(2z)^2/8}}{(1-T(2z))^{1/4}}\frac{dz}{z^{n+1}}$$
and $T = -LambertW(-z) = \sum_{n>0} n^{n-1}\frac{z^n}{n!}$.

3. Lagrangian subs. $u = T(2z)$ leads to integral of the form $\oint g(u)\exp(nh(u))du$ with $h(u) = u - \frac{m}{n}\log u + (1 - m/n)\log(2-u)$ so that $h'(2m/n) = 0$ and $h''(2m/n) > 0 \rightarrow$ **classical saddle point method** applies on circular path $|z| = 2m/n$.

Proof of the Theorem : the critical phase

Some **multicyclic components** can appear and the general formula looks like

1.
$$\mathrm{coeff(n, m,\ r)} \times \frac{1}{2\pi i} \oint \left(\frac{T(2z)}{2} - \frac{T(2z)^2}{4} \right)^{n-m+r} \frac{e^{-T(2z)/4 - T(2z)^2/8}}{(1 - T(2z))^{1/4+3r}} \frac{dz}{z^{n+1}}$$

2. Again
$$\mathrm{Stirling(n, m, r)} \times \frac{1}{2\pi i} \oint g_r(u) e^{nh(u)} du$$

3. $h(u) = u - \frac{m}{n} \log u + (1 - m/n) \log (2 - u)$ but this time there are 2 **saddle-points**:
$$u_0 = 2m/n = 1 + 2\,\mu\, n^{-1/3} \text{ and } u_1 = 1\,.$$
(observe $u_0 \sim u_1$ as $n \gg 1$.)

4. Moreover, $h(1) = h'(1) = h''(1) = 0$ so that the **usual “Gaussian” approximation does not hold**.

The critical phase via Airy function

Integral representation on the complex plane

The Airy function is given by

$$\mathrm{Ai}(z) = \frac{1}{2\pi i}\int_{\mathcal{C}} \exp\left(\frac{t^3}{3} - z\,t\right) dt = \frac{\sqrt{3}}{2\pi}\int_0^{\infty} \exp\left(-\frac{t^3}{3} - \frac{z^3}{3t^3}\right) dt$$

where the integral is over a path $\mathcal{C}$ starting at the point at infinity with argument $-\frac{\pi}{3}$ and ending at the point at infinity with argument $+\frac{\pi}{3}$

References

- Books: [WONG'89], [OLVER'97].
- Random graphs (analytic combinatorics approaches): [FLAJOLET, KNUTH, PITTEL'89], [JANSON, KNUTH, ŁUCZAK, PITTEL'93].

Proof of the critical phase of the Th.

Let $p_r(n, m) =$ Proba. to have SAT-graph of excess r.

Thus, $p(n, m) = \sum_{r \geq 0} p_r(n, m)$ is the probability of a random formula being SAT.

The proof of part (ii) of the Theorem (and confirmation of **[Monasson '07]**) is completed as

1. For fixed r, we compute (by means of the Airy stuff)

$$n^{1/12} \times p_r(n, m) \sim c_r A(3r + 1/4, \mu)$$

2. and there exist R, C and ε such that for $r \geq R$ and all n:

$$n^{1/12} p_r(n, m) \leq Ce^{-\varepsilon r}$$

The proof is completed using dominated convergence theorem.

Generalization

The Random 2-XORSAT Problem has been generalized in **[de Panafieu '14]**'s PhD thesis.

Recap. of some tractable phase transitions

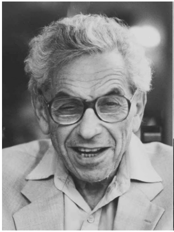
- **2-XORSAT :** is a random formula **SAT**? Done.
- **2-COL (bipartiteness):** is a random graph **2−COLORABLE**? Done by **[Pittel, Yeum '10]**.
- **2-SAT:** is a random formula like

$$\begin{cases} x_1 \vee \bar{x}_{19} \\ \dots \\ \bar{x}_{27} \vee \bar{x}_{36} \\ \dots \end{cases}$$

SAT? Open problem!.

- **Planarity :** is a random graph **PLANAR**?
- ...

Graph planarity phase transition



Paul Erdős (1913-1996)

Alfréd Rényi (1921-1970)

ON THE EVOLUTION OF RANDOM GRAPHS

by

P. ERDŐS and A. RÉNYI

Dedicated to Professor P. Turán at his 50th birthday.

We can show that for $N(n) = \frac{n}{2} + \lambda\sqrt{n}$ with any real λ the probability of $\Gamma_{n,N(n)}$ not being planar has a positive lower limit, but we cannot calculate ts value. It may even be I, though this seems unlikely.

Prior works and contribution

- **[Łuczak, Pittel and Wierman '93]** proved the 1960's conjecture

$$p(\mu) = \lim_{n \to +\infty} \mathbb{P}\left(G\left(n, \frac{n}{2}(1 + \mu n^{-1/3})\right) \texttt{ is planar}\right)$$

exists and $0 < p(\mu) < 1$.

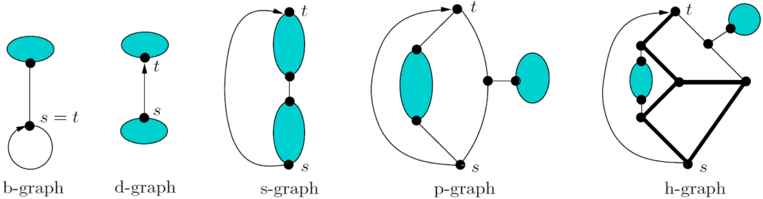
- **[Janson, Łuczak, Knuth, Pittel '93]** gave bounds

$$0.9870 \cdots < p(0) < 0.997 \cdots$$

- **Our contribution :** the whole description of $p(\mu)$ **[Noy, R., Rué '15]**.

Our strategy

- At the critical phase transition, the **kernels**(3-core) that matter are the **cubic ones** (same for random graphs, random 2-XORSAT formulas, $\cdots$).
- Decomposition of [BODIRSKY, KANG, LÖFFLER, MCDIARMID '07]



The decomposition in EGFs

$G^{(k)}(x) = \sum_{n \geq 0} \frac{1}{n!} g_n^{(k)} x^n$ the EGF of k-connected cubic planar graphs with n vertices.
In particular, observe that $g_n^{(k)} = 0$ when n is odd. By convention we write $g_0^{(0)} = 1$.
$G^{(1)}(x)$ is defined in terms of the following equations:

$$
\begin{aligned}
G^{(0)}(z) &= \exp G^{(1)}(z) \\
3x \frac{dG^{(1)}(z)}{dx} &= D(z) + C(z) \\
B(z) &= \frac{z^2}{2}(D(z) + C(z)) + \frac{z^2}{2} \\
C(z) &= S(z) + P(z) + H(z) + B(z) \\
D(z) &= \frac{B(x)^2}{z^2} \\
S(z) &= C(z)^2 - C(z)S(z) \\
P(x) &= z^2 C(z) + \frac{1}{2} z^2 C(z)^2 + \frac{z^2}{2} \\
2(1 + C(x))H(x) &= \left(v(1 - 2v) - v(1 - v)^3 \right) \\
x^2 (C(x) + 1)^3 &= v(1 - v)^3 .
\end{aligned}
$$

A little check!

We find for the cubic planar multigraphs

$$G^{(0)}(z) = \sum_{r\geq 0} G_r \frac{z^{2r}}{(2r)!} = 1 + \frac{5}{24}z^2 + \frac{385}{1\,152}z^4 + \frac{83\,933}{82\,944}z^6 + \cdots$$

$$\text{All cubic}(z) = \sum_{r\geq 0} \frac{(6r)!}{(3r)!2^{3r}6^{2r}} \frac{z^{2r}}{(2r)!} = 1+\frac{5}{24}z^2+\frac{385}{1\,152}z^4+\frac{85\,085}{82\,944}z^6+\cdots$$

First discrepancy is at z^6 with $\frac{85\,085}{82\,944} - \frac{83\,933}{82\,944} = \frac{1}{72}$. But $\frac{1}{72} = \frac{10}{6!}$ and there are 10 possible manners to label $K_{3,3}$!

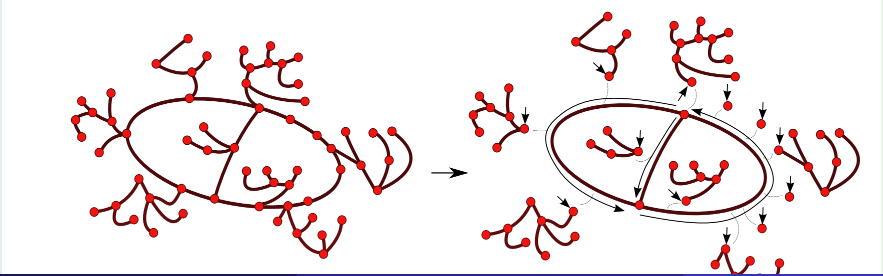
How to get the window?

Main terms at the phase transition

- Only trees (planar), unicyclic (planar) and connected components with cubic planar kernels matter.
- From this observation: use complex integration.

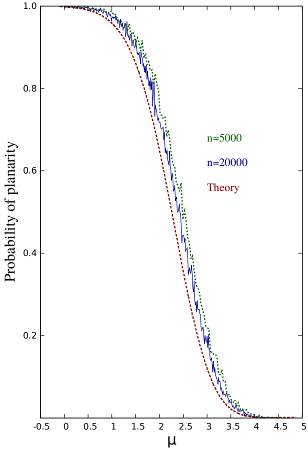
Concretely

$$\frac{n!}{\binom{\binom{n}{2}}{m}} \frac{1}{2\pi i} \oint \frac{\text{Trees}(z)^{n-m+r}}{(n-m+r)!} \times e^{Unicycles(z)} \times \text{Cubic}\left(\text{RootedTrees}(z), \frac{1}{(1-\text{RootedTrees}(z))}\right) \frac{dz}{z^{n+1}}.$$



A picture of the whole window

- Using the decomposition, we get the generating functions (of planar connected components whose kernels are cubic).



-

Outline of the talk

- Introduction & motivations.
- Some tractable graph phase transitions.
- Shifting the critical threshold.
- Open problems.
- Conclusion and perspectives.

A) Structural point of view	B) Analytical point of view
At the beginning, very simple graphs: - set of trees - set of unicycles	Simple saddle point approach
and then at the threshold: - appearance of more complex connected components	Airy-like saddle point

Question:

what about graphs with restricted degrees (degrees in set $\mathcal{D}$)?

Dovgal, R'18

New set of thresholds as long as $1 \in \mathcal{D}$

The idea: until how many edges we have

$$\lim_{n\to+\infty} \frac{\sharp \textbf{Set of Trees}_{\mathcal{D}} \times \textbf{Set of Unicycles}_{\mathcal{D}}}{\sharp \textbf{All configurations}_{D}} \stackrel{?}{=} 1$$

(remark: [DE PANAFIEU, RAMOS'16] obtained the asymptotics of "$\sharp$all configurations$_D$" with n vertices and $O(n)$ edges.)

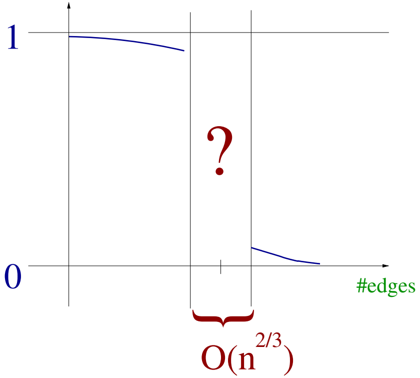
Random 2-SAT: what is inside the window?

- A long and rich problem: the only of the K-SAT family with **proven threshold** [GOERDT '92], [CHVÁTAL, REED '92], [DE LA VEGA '92].
- Best result up to date [BOLLOBÀS, BORGS, CHAYES, KIM, WILSON '01]: the window of transition is of size $O(n^{2/3})$ i.e. let $p(n, m) = \mathbb{P}(\texttt{Formula}(n \texttt{ clauses}, m \texttt{ variables}) \texttt{ is SAT})$

$$p(n, m) = \begin{cases} \bullet\ 1 - O\left(\frac{n^2}{(n-m)^3}\right) \text{ as } n - m \gg n^{2/3} \\ \bullet\ O(1) \text{ as } |n - m| = O(n^{2/3}) \\ \bullet\ \exp\left(-\frac{(m-n)^3}{n^2}\right) \text{ as } m - n \gg n^{2/3} \end{cases}$$

Random 2-SAT: what is known and what is next?

- Concretely **[Bollobàs et al. '01]**



- A question of **[Flajolet '08]**. Note that physicists **[Deroulers, Monasson '06]** computed numerically $\mathbb{P}\left[\text{formula}(10^6\text{variables, } 10^6 \text{ clauses is SAT})\right] \sim 0.913\cdots$

The characterization of 2-SAT

$$\mathcal{F} : \begin{cases} x_1 \vee \bar{x}_2 \\ \dots \\ \bar{x}_3 \vee \bar{x}_2 \\ \dots \end{cases}$$

- $r \vee t$ is equivalent to $\bar{r} \Longrightarrow t$ **AND** $\bar{t} \Longrightarrow r$
- Generate a directed graph D

  with 2 correlated arcs $\bar{r} \mapsto t$ AND $\bar{t} \mapsto r$
  for each clause $r \vee t$.

- Formula $\mathcal{F}$ is **SAT** **iff** there is **no directed path** from s to $\bar{s}$ and vice-versa for all variables s.

An easier problem : random digraphs

- Consider random directed graphs $\overrightarrow{D}(n, n)$ with n nodes and n arcs OR $\overrightarrow{D}(n, p = 1/n)$.
- What is known:
  1. [KARP '90] $p = c/n$ with $c < 1$ or $c > 1$.
     OUTSIDE the scaling window!
  2. [ŁUCZAK '90] $p = c/n$ $c < 1$ or $c > 1$
     small or giant strongly connected components

     See also [PITTEL, POOLE '14] for asymptotic normality).
     OUTSIDE the scaling window!

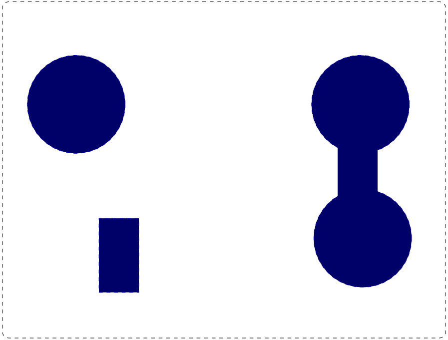
Stacked on to-do list

Go INSIDE that scaling window.

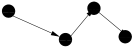
Characterize typical strongly connected components of $\overrightarrow{D}(n,\ n + O(n^{2/3}))$.

Conjecture/intuition.

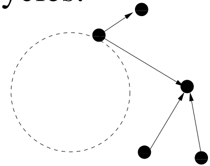
Connected giant component: O(n)



tree−like components:



cycles:



Strongly connected sub−components: $O(n^{1/3})$

Outline of the talk

- Introduction & motivations.
- Some tractable graph phase transitions.
- Open problems.
- Conclusion and perspectives.

Conclusion

Enumerative approaches for SAT-like problems

- From typical anatomy of the graphs, count only the main components (3-regular or cubic planar components)
- If we have some decompositions : enumerative/analytic combinatorics work well.
- [DE PANAFIEU, RAMOS'16] : new approaches on connected graphs with large excess.
- [COLLET, DE PANAFIEU, GARDY, GITTENBERGER, R.'18] : analytic combinatorics of models of graphs with forbidden subgraphs.

Stack of what to do?

- Mixtures of formula:
  1. $(2 - XOR, 2 - SAT)$ is difficult.
     Interpolation between **coarse** and **sharp** phases transitions.
  2. $(2 + p)$-XORSAT is less difficult. (Same kind of interpolation.)
  3. $(2 + p) - SAT$ is **extremely difficult.**
     Interpolation between **tractable/intractable** problems.
- From complexity theory: **QBF** (Quantified Boolean Formulas).