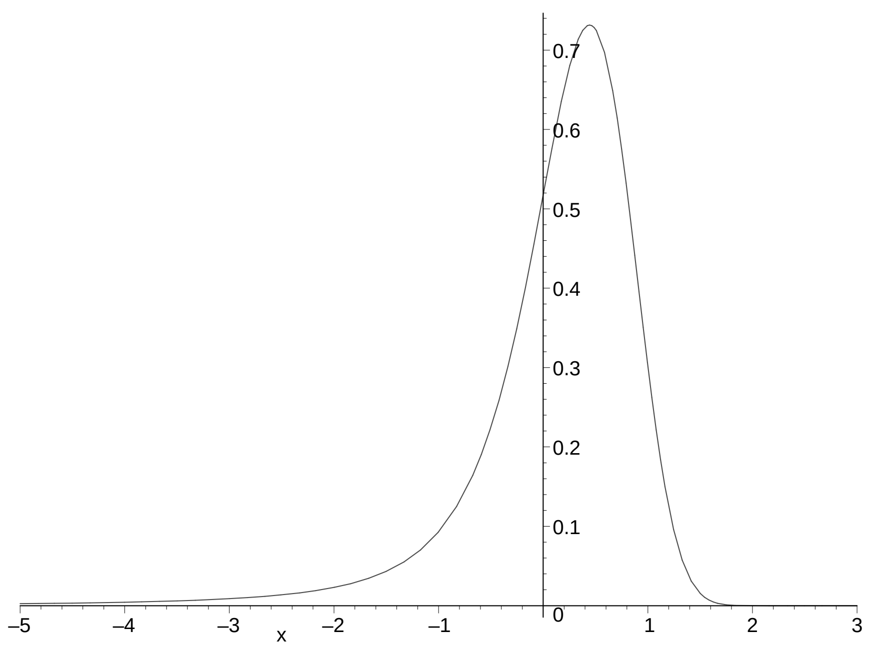
Projet Algorithmes
INRIA-Rocquencourt

La distribution d'Airy et les graphes planaires



Cyril Banderier (INRIA)
Philippe Flajolet (INRIA)
Gilles Schaeffer (LORIA)
Michèle Soria (LIP6)

Contrat

- Graphes aléatoires
- Transition de phase graphes aléatoires
- Zoom sur une transition de phase
- Quid des cartes planaires ?

- Méthode de Laplace
- Col simple $\Rightarrow$ gaussien
- Col double $\Rightarrow$ Airy
- Cols coalescents

- La distribution d'Airy pour les cartes

- Tutte, multiconnexité et cartes planaires

- Application : algorithme avec rejet

- Schémas de composition

- Prospective (Airy everywhere)

Graphes aléatoires

Paul Erdős (1913-1996) et Alfréd Rényi (1921-1970)

articles fondateurs (1959-1960) 2 modèles :

- probabliste : binomial random graph $G(n,p)$
- combinatoire : uniform random graph $G(n,m)$

$\Pr(G) = \binom{\binom{n}{2}}{m}^{-1}$

Théorème, Łuczak (1990) $G(n,p) \Leftrightarrow G(n,m)$
(i.e. même proba d'avoir une propriété avec
$m = \binom{n}{2} p$)

Zoom sur une transition de phase

Théorème (Bollobás & Thomason,1986) : Loi 0-1 pour propriétés “stables par ajout d’arêtes”

$\Rightarrow$ phénomène de seuil

$\Rightarrow$ zoom sur intervalle critique

ex: contenir un certain motif (sous-graphe), connexité

Sinon, aussi seuil pour : hamiltonicité, planarité

Convergence loi normale, Poisson, OK.

Combinatoire analytique

structures combinatoires

$\Downarrow$

équations fonctionnelles

(dictionnaire sur les séries génératrices)

$\Downarrow$

asymptotique

(analyse de singularités, col)

$\Downarrow$

lois limites

La loi d'Airy des cartes

George Biddell Airy (1801-1892)

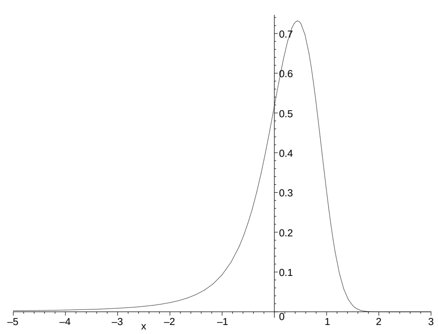
Fonction d'Airy $y'' - zy = 0$

$$\mathrm{Ai}(z) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{i(zt+t^3/3)}\,dt =$$

$$\frac{1}{\pi 3^{2/3}}\sum_{n=0}^{\infty}\frac{\Gamma((n+1)/3)}{n!}\sin\left(\frac{2(n+1)\pi}{3}\right)\left(3^{1/3}x\right)^n$$

La distribution d'Airy pour les cartes

$$\mathcal{A}(x) = 2\exp\left(-\frac{2}{3}x^3\right)\left(x\mathrm{Ai}(x^2) - \mathrm{Ai}'(x^2)\right).$$



The Map Airy Distribution, $\int_{-\infty}^{+\infty}\mathcal{A}(x)dx = 1$

Grandes puissances et méthodes de cols

Théorème des résidus :

$$F(z) = a_{-2}z^{-2} + a_{-1}z^{-1} + a_0 + a_1 z + a_2 z^2 + \dots$$

$$\int \Downarrow$$

$$\int_\gamma F(z)dz = [a_{-2}\frac{z^{-1}}{-1}]_\gamma + [a_{-1}\ln(z)]_\gamma + [a_0 z]_\gamma + \dots$$

$$\int_\gamma F(z)dz = 0 + 2i\pi a_{-1} + 0 + \dots$$

Formule de Cauchy (coeff $\leftrightarrow$ intégrale) :

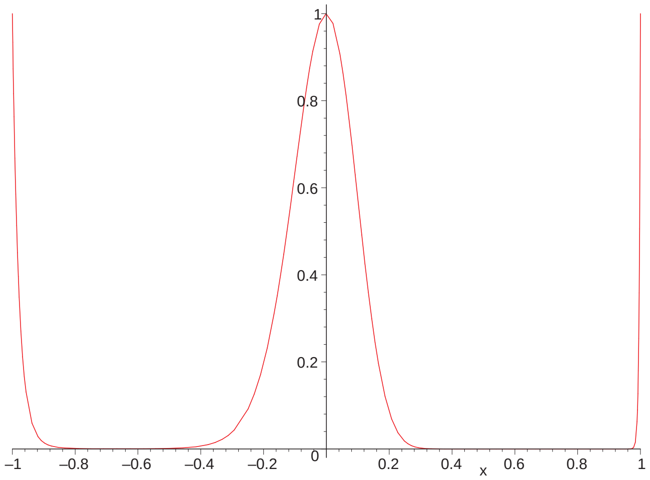
$$[z^n]F(z) = \int_\gamma F(z)/z^{n+1}dz$$

Grande puissance

$$[z^n]F(z) = \int e^{H_n(z)}dz$$

La méthode de Laplace

Pierre-Simon Laplace (1749-1827)



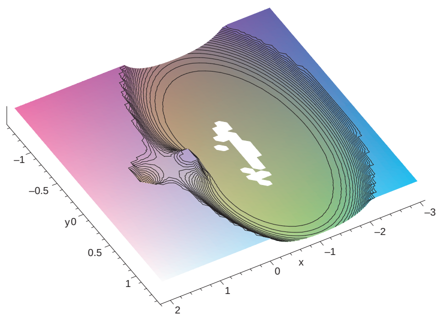
$(1 - x^2 - x^3)^n$ sur $[-1, 1]$ $(n = 50)$

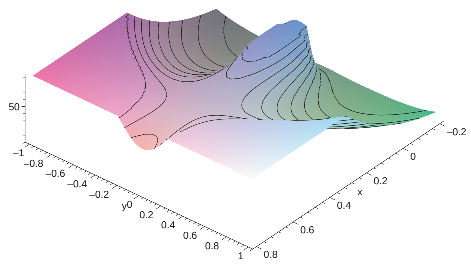
Asymptotiquement, le maximum dicte sa loi

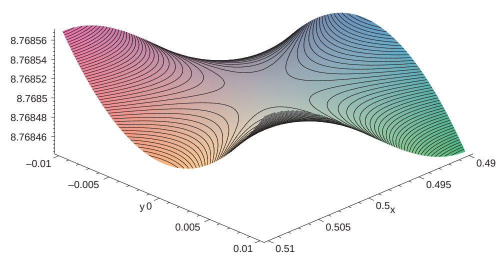
Col simple $\rightarrow$ gaussien

Saddle point

Col double → Airy







Monkey saddle point

Cols Coalescents

Cartes

Définition : Carte = graphe planaire enracinée (= on oriente arbitrairement une arête + coulisse sur la sphère)

Deux familles :

la famille $\mathcal{M}$ des cartes “générales” ;

la famille $\mathcal{C}$ des cartes vérifiant une certaine propriété (ex: multiconnexité, séparabilité...).

Définition : noyau de M = plus grande composante (de type C) de M et contenant sa racine.

But : Donner proba qu'une carte de taille n ait un noyau de taille k

$$\Pr(X_n = k) = ???$$

Schéma lagrangien

$$M(z) := \sum_{n\geqslant 1} M_n z^n \quad \text{et} \quad C(z) := \sum_{k\geqslant 1} C_k z^k.$$

(i) Root-face decomposition.

Tutte (quadratic method)

$$M(z) = \psi(L(z)) \quad \text{avec} \quad L(z) = z\phi(L(z))$$

(ex: cartes non séparables
$\phi(y) = (1+y)^3, \quad \psi(y) = y(1-y))$.

Théorème d'inversion de Lagrange $\Rightarrow$
$M_n \equiv [z^n]M(z) = \frac{1}{n}[y^n]\psi'(y)\phi(y)^n = \frac{4(3n)!}{n!(2n+2)!}$

(ii) Substitution decomposition. Tutte,

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

carte = “pas de noyau” ou “noyau dans lequel on subsitue des cartes”

maps ($\mathcal{M}$)	$\mathcal{M}_n$	cores ($\mathcal{C}$)	scheme	α_0	c
general	n edges	nonseparable	$\mathcal{M} \simeq \mathcal{C} \circ (\mathcal{X}\mathcal{M}^2)$	1/3	$3/4^{2/3}$
general	n edges	bridgeless	$\mathcal{M} \simeq \mathcal{C} \circ (\mathcal{X}(\mathcal{X}\mathcal{M})^*)$	4/5	$(5/3)^{2/3}/4$
general	n edges	loopless	$\mathcal{M} \simeq \mathcal{L} + \mathcal{C} \circ (\mathcal{X}((\mathcal{X}\mathcal{M})^*)^2)$	2/3	3/2
loopless	n edges	simple	$\mathcal{M} \simeq \mathcal{C} \circ (\mathcal{X}\mathcal{M})$	2/3	$3^{4/3}/4$
bipartite	n edges	bip. simples	$\mathcal{M} \simeq \mathcal{C} \circ (\mathcal{X}\mathcal{M})$	5/9	$3^{8/3}/20$
bipartite	n edges	bip. nonsep.	$\mathcal{M} \simeq \mathcal{C} \circ (\mathcal{X}\mathcal{M}^2)$	5/13	$(13/6)^{5/3} \cdot 3/10$
bipartite	n edges	bip. bridgeless	$\mathcal{M} \simeq \mathcal{C} \circ (\mathcal{X}(\mathcal{X}\mathcal{M})^*)$	3/5	$(15/2)^{5/3}/18$
nonseparable	n edges	simple nonsep.	$\mathcal{M} \simeq \mathcal{C} \circ (\mathcal{X}\mathcal{M})$	4/5	$15^{5/3}/36$
nonseparable	$n+1$ edges	3-connected	$\mathcal{M} \simeq \mathcal{D} + \mathcal{C} \circ \mathcal{M}$	1/3	$3^{4/3}/4$
cubic nonsep.	$n+2$ faces	cubic 3-conn.	$\mathcal{M} \simeq \mathcal{C} \circ (\mathcal{X}(1+\mathcal{M})^3)$	1/2	$(3/2)^{1/3}$
cubic 3-conn.	$n+2$ faces	cubic 4-conn.	$\mathcal{M} \simeq \mathcal{M} \cdot \mathcal{C} \circ (\mathcal{X}\mathcal{M}^2)$	1/2	$6^{2/3}/3$

Quelques schémas de composition ($\mathcal{X}$ arête, $\mathcal{L}, \mathcal{D}$ familles auxiliaires).

Théorème (Tails and distinct saddles [?]). *Let $\lambda(n)$ be an arbitrary function with $\lambda(n) \to +\infty$ and $\lambda(n) = o(n^{1/3})$. Then, the probability distribution of the core of random element of $\mathcal{M}_n$ satisfies*

$$\begin{aligned} \Pr(X_n = k) &\sim \frac{32}{243\sqrt{\pi}} \cdot \frac{n^{5/2}}{k^{3/2}(n-3k)^{5/2}}, &\quad \textit{uniformly for } \lambda(n) < k < n/3 - n^{2/3}\lambda(n) \\ \Pr(X_n = k) &= O\left(\exp(-n(k/n - 1/3)^3)\right), &\quad \textit{uniformly for } k > n/3 + n^{2/3}\lambda(n). \end{aligned}$$

Proof (Sketch). The left tail ($n < 3k$) corresponds to the saddle point $z_+ = \frac{1}{2}$ that is dominant (*i.e.*, nearer to the origin and providing the major asymptotic contribution; Fig. **??**, top). The right tail ($n > 3k$) has $z_- = (n-k)/(n+k)$ dominating (Fig. **??**, bottom). In each case, the basic saddle point method applies. □

Théorème **(Central part and a double saddle).** *The probability distribution of the core of random element of $\mathcal{M}_n$ satisfies, when $n = 3k$,*

$$\Pr(X_{3k} = k) = \frac{4}{27}\frac{\Gamma(2/3)}{3^{1/6}\pi}k^{-2/3}\left(1 + O((\ln(k))^4 k^{-1/3})\right) \qquad \textit{where} \quad \frac{4}{27}\frac{\Gamma(2/3)}{3^{1/6}\pi} \approx 0.0531.$$

Proof. When $n = 3k$, equation (**??**) becomes

$$[z^{3k}]M^k(z) = \frac{1}{6i\pi}\int_\Gamma G(z)P(z)^k dz, \qquad (1)$$

where $P(z) := \psi\phi^3/z^3 = (1-z)(1+z)^9/z^2$ and the "kernel" $\ln(P)$ (together with P, P^k) now has a double saddle point at $\tau = z_- = z_+ = \frac{1}{2}$, sometimes called a "monkey saddle", *viz.*, a saddle with places for two legs and a tail. The idea consists in choosing a contour that is no longer a circle centered at the origin, but, rather, approaches the real axis at an angle (see Fig. **??**, middle). Specifically, the integration path Γ consists of the following: the part Γ_0 of a circle centered at 0 from which a small arc is taken out, joining with two (small) segments Δ_1, Δ_2 of length δ that intersect at $\frac{1}{2}$ at an angle of $\pm 2\pi/3$.

We shall adopt a value of δ satisfying two conflicting requirements,

$$n\delta^3 \to \infty, \qquad n\delta^4 \to 0, \quad \text{specifically} \quad \delta = (\log n) n^{-1/3}. \tag{2}$$

The kernel $\ln(P)$ has a double saddle point in τ, meaning that its local expansion is of the cubic type:

$$\ln(P(z)) = \ln(P(\tau)) - d(z-\tau)^3 + O((z-\tau)^4), \qquad d = \frac{64}{9}.$$

The geometry of the level curves of the kernel shows that the contribution $\mathcal{E}_0$ along Γ_0 to the integral in (1) is bounded by a constant times the value of $P(z)^k$ at the endpoints of Γ_0. This contribution then satisfies

$$\mathcal{E}_0 \equiv \int_{\Gamma_0} G(z) \exp[k \ln(P(z))] \, dz = O(P(\tau)^k \exp(-kd\delta^3)),$$

which, given one of our constraints on δ (the condition $n\delta^3 \to \infty$ in (2)) is exponentially small.

The contribution $\mathcal{E}_{1,2}$ along $\Delta_1 \cup \Delta_2$ to the integral in (1) provides the dominant contribution and is estimated next by a local analysis of P^k for values of z near τ. Set $u = z - \tau$. The condition $n\delta^4 \to 0$ in (2) implies that terms of order 4 and higher do not matter asymptotically, and a simple calculation, using the fact that $G(\tau+u) = -8u + O(u^2)$, yields

$$\mathcal{E}_{1,2} \equiv \int_{\Delta_1\cup\Delta_2} G(z)\exp[k\ln(P(z))]\,dz = -8P(\tau)^k \int_{\Delta_1\cup\Delta_2} u\exp\left(-kdu^3\right)\left(1+O(k\delta^4)\right)du\,.$$

The integral along $\Delta_1 \cup \Delta_2$ can be extended to two full half lines of angle $\pm 2\pi/3$ emanating from the origin, this at the expense of introducing only exponentially small error terms (since $n\delta^3 \to \infty$). The rescaling $v = t(kd)^{1/3}\exp(2i\pi/3)$ on Δ_1 and $v = t(kd)^{1/3}\exp(-2i\pi/3)$ on Δ_2 then shows that the completed integral equals

$$(kd)^{-2/3}(e^{4i\pi/3} - e^{-4i\pi/3})\int_0^{+\infty} t\exp(-t^3)dt = -(kd)^{-2/3}\frac{i}{\sqrt{3}}\Gamma(2/3),$$

where the evaluation results from a cubic change of variable. In summary, we have found

$$[z^n]M^k(z) = \frac{1}{6i\pi}\left(\mathcal{E}_0 + \mathcal{E}_{1,2}\right) = \frac{3^{5/6}}{12\pi}\frac{P(\tau)^k}{k^{2/3}}\Gamma(2/3)\left(1 + O(k\delta^4)\right),$$

which, given our choice of δ, is equivalent to the statement. $\square$

A similar reasoning proves that the estimate remains valid for $n = 3k + e$ with $e = 1$ or $e = 2$, and more generally with any e satisfying $e = O(1)$.

Algorithme avec rejet

Génération aléatoire

	Algorithme probabiliste `Core(k)` avec paramètre $f(k)$
1.	`Map(n)` engendre une carte $M \in \mathcal{M}$ de taille $n = f(k)$;
2.	extraire la plus grande composante C de M en fonction du schéma ;
3.	si taille de $C \neq k$, alors goto 1 ;
4.	renvoyer C.

Théorème. *Pour tous les algorithmes avec extraction-rejet (cf G. Schaeffer), choisir $f(k) = n/\alpha_0$ donne $\ell_n \sim n^{2/3}/(\mathcal{A}(0)c)$ itérations en moyenne. Le choix optimum $f(k) = k/\alpha_0 - \frac{x_0}{\alpha_0 c}(k/\alpha_0)^{2/3}$ apporte un gain de 30%. (x_0 pic de la distribution)*

Schémas de composition

Analyse de singularités

$$\begin{aligned} M(z) &= \sigma - a(1-z/\rho) + b(1-z/\rho)^{3/2} + O((1-z/\rho)^2) \\ C(z) &= c_0 - a'(1-z/\sigma) + b(1-z/\sigma)^{3/2} + O((1-z/\sigma)^2), \end{aligned}$$

$C(uM(z))$ composition critique de type (3/2,3/2)

Théorème. *(i) Pour $k = \alpha n + \lambda(n)$, avec $0 \le \alpha < \alpha_0 = \frac{\sigma}{a}$ et $\lambda(n) = o(n)$, la proba de distribution de la taille X_n de la $\mathcal{C}$–composante de l'élément aléatoire $\mathcal{C}[\mathcal{M}]$ de taille n est*

$$\begin{aligned} \Pr(X_n = k) &\sim \frac{b'}{\sigma\sqrt{\pi}}\, \alpha^{-3/2}(1-\alpha/\alpha_0)^{-5/2}\, n^{-3/2} \quad (\alpha > 0) \\ &\sim \frac{b'}{\sigma\sqrt{\pi}}\, \lambda(n)^{-3/2} \quad (\alpha = 0, \lambda(n) \to +\infty). \end{aligned}$$

(ii) Pour $k = \alpha_0 n + x n^{2/3}$ avec $\alpha_0 = \sigma/a$ et $x = o(n^{1/3})$, on a une loi d'Airy de type carte de paramètre $c = \left(\frac{b}{3a}\right)^{2/3}/\alpha_0$:

$$n^{2/3}\Pr(X_n = \alpha_0 n + x n^{2/3}) \sim \frac{b'}{\alpha_0^{3/2} b} c\mathcal{A}(cx)$$

Une loi universelle

Détectée :

• amas dans les tables de hachage linéaire (Flajolet, Poblete, Viola 98)
• aire sous diverses excursions (Louchard 84)
• aire de quelques familles de polyominos (Duchon 99)
• émergence du premier cycle dans les modèles de graphes aléatoires (Knuth, Flajolet, Pittel 89)
• émergence de la première composante géante (Gao, Wormald 99)

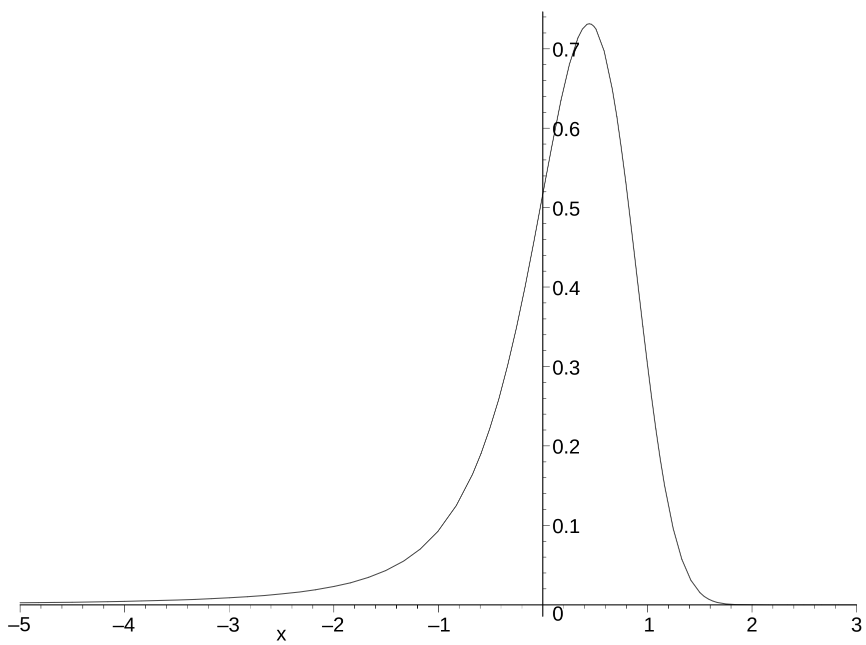
Prospective :

• forêts aléatoires d'arbres non-enracinés
• phénomène de seuil pour 2-SAT

À regarder

• coalescence de cols multiples
• compositions de type (α, β)

Enchantée d'avoir fait votre connaissance,
au revoir et à bientôt !



Signé : the map–Airy law.