Planar Maps and Airy Phenomena

Cyril Banderier[1], Philippe Flajolet[1], Gilles Schaeffer[2], and Michèle Soria[3]

[1] Algorithms Project, INRIA, Rocquencourt, 78150 Le Chesnay (France)
[2] Loria, CNRS, Campus scientifique, B.P. 239, 54506 Vandœuvre-lès-Nancy (France)
[3] Lip6, Université Paris 6, 8 rue du Capitaine Scott, 75005 Paris (France).

Abstract. A considerable number of asymptotic distributions arising in random combinatorics and analysis of algorithms are of the exponential-quadratic type (e^{-x^2}), that is, Gaussian. We exhibit here a new class of "universal" phenomena that are of the exponential-cubic type (e^{ix^3}), corresponding to nonstandard distributions that involve the Airy function. Such Airy phenomena are expected to be found in a number of applications, when confluences of critical points and singularities occur. About a dozen classes of planar maps are treated in this way, leading to the occurrence of a common Airy distribution that describes the sizes of cores and of largest (multi)connected components. Consequences include the analysis and fine optimization of random generation algorithms for multiply connected planar graphs.

Maps are planar graphs presented together with an embedding in the plane, and as such, they model the topology of many geometric arrangements in the plane and in low dimensions (e.g., 3-dimensional convex polyhedra). This paper concerns itself with the statistical properties of random maps, *i.e.*, the question of what such a random map typically looks like. We focus here on connectivity issues, with the specific goal of finely characterizing the size of the highly connected "core" of a random map.

The bases of an enumerative theory of maps have been laid down by Tutte [22] in the 1960's, in an attempt to attack the four-colour conjecture. The present paper builds upon Tutte's results and upon the detailed yet partial analyses of largest components given by Bender, Richmond, Wormald, and Gao [2, 11]. We establish the common occurrence of a new probability distribution, the "map–Airy distribution", that precisely quantifies the sizes of cores in about a dozen varieties of maps, including general maps, triangulations, 2-connected maps, etc. As a corollary, we are able to improve on the complexity of the best known random samplers for multiply connected planar graphs and convex polyhedra [19].

The analysis that we introduce is largely based on a method of "coalescing saddle points" that was perfected in the 1950's by applied mathematicians [3, 24, 1] and has found scattered applications in statistical physics and the study of phase transitions [16]. However, this method does not appear to have been employed so far in the field of random combinatorics. We claim some generality for the approach proposed here on at least two counts. First, a number of enumerative problems are known to be of the "Lagrangean type", being related to the Lagrange inversion theorem and its associated combinatorics. The classical saddle point method is then instrumental in providing asymptotics of simpler problems. However, confluence of saddle points is a stumbling block of the basic method. As we show here, planar maps are precisely instances of this special situation. Next, the method extends to the analysis of a new composition scheme. Indeed, it is known, in the realm of analytic combinatorics, that asymptotic properties of random structures are closely related to singular exponents of counting generating functions. For "most" recursive objects the exponent is $\frac{1}{2}$ and the probabilistic phenomena are described by classical laws, like Gaussian, exponential, or Poisson. Methods of the paper permit us to quantify distributions associated with singular exponents $\frac{3}{2}$ present in maps and unrooted trees and leading to Airy laws.

Very roughly, the classical saddle point method gives rise to probabilistic and asymptotic phenomena that are in the scale of $n^{1/2}$ and the analytic approximations are in the form of an "exponential-quadratic" (e^{-x^2}) corresponding to Gaussian laws. The coalescent saddle-point method presented here gives rise to phenomena in the scale of $n^{1/3}$, with analytic approximations of the "exponential-cubic type" (e^{ix^3}), which, as we shall explain, is conducive to Airy laws. The Airy phenomena that we uncover in random combinatorics should thus be expected to be of a fair degree of universality. To support this claim, here are scattered occurrences of what we recognize as Airy phenomena in the perspective of this paper: the emergence of first cycles and of the giant component in the Erdös-Rényi graph model [8, 13], the enumeration of random forests of unrooted trees [14], clustering formation in the construction of linear probing hash tables [10], the area under excursions and the cumulative storage cost of dynamically varying stacks [15], the area of certain polyominoes [7], path length in combinatorial tree models [21], and (we conjecture) the threshold phenomena involved in the celebrated random 2-SAT problem [4]. We propose to elaborate on these connections in future papers.

Plan of the paper. Basics of maps are introduced in Section 1 where the Airy distribution is presented. The enumerative theory can be developed along two parallel lines, one Lagrangean, the other based on singularity analysis. We first approach the analysis of core size via the Lagrangean framework and variations on the saddle point method: a fine analysis of the geometry of associated complex curves is shown to open access to the size of the core, with the Airy distribution arising from double or "nearby" saddles (Section 2); a refined analysis based on the method of coalescent saddle points then enables us to quantify the distribution of core size over a wide range with precise large deviation estimates (Section 3). The method applies to about a dozen of types of planar maps, it provides a precise quantification of largest components, with consequences on the random generation of highly connected planar graphs (Section 4). Finally, we show that the very same Airy law is bound to occur in any instance of a general composition scheme of analytic combinatorics (Section 5).

1 Basics of maps

A *map* is a planar graph given together with an embedding in the plane considered up to continuous deformations. Following Tutte, we consider *rooted* maps, that is, maps with an oriented edge called the *root*—this simplifies this analysis without essentially affecting statistical properties (see [17] and Section 4). Generically, we take $\mathcal{M}$ and $\mathcal{C}$ to be two classes of maps, with $\mathcal{M}_n$, $\mathcal{C}_n$ the subsets of elements of size n (typically elements with $n+1$ edges). Here, $\mathcal{C}$ is always a subset of $\mathcal{M}$ that satisfies additional properties (*e.g.* higher connectivity). The elements of $\mathcal{M}$ are then called the "basic maps" and the elements of $\mathcal{C}$ are called the "core-maps". We define the *core-size* of a map $m \in \mathcal{M}$ as the size of the largest $\mathcal{C}$–component of m that contains the root of m. As a pilot example, we shall specialize the basic maps $\mathcal{M}_n$ to be the class of nonseparable maps (*i.e.*, 2-connected loopless maps) with $n+1$ edges and $\mathcal{C}_k$ to be the set of 3-connected maps with $k+1$ edges.

Our major objective is to characterize the probabilistic properties of core-size of a random element of $\mathcal{M}_n$, that is, of a random map of size n, when all elements are taken equally likely. Core-size then becomes a random variable X_n defined on $\mathcal{M}_n$. In essence, the pilot example thus deals with 3-connectivity in random 2-connected maps. The paradigm that we illustrate by a particular example is in fact of considerable generality as can be seen from Sections 4, 5 below.

The physics of maps. From earlier works [2, 11, 18], it is known that a random map of $\mathcal{M}_n$ has with high probability a core that is either "small" (roughly of size $k = O(1)$) or "large" (being $\Theta(n)$). The probability distribution $\Pr(X_n = k)$ thus has two distinct modes. The small region (say $k = o(n)$) has been well quantified by previous authors, see [2, 11, 18]: a fraction $p_s = \frac{65}{81}$ of the probability mass is concentrated there. The large region is also known from these authors to have probability mass $p_\ell = 1 - p_s = \frac{16}{81}$ concentrated around $\alpha_0 n$ with $\alpha_0 = \frac{1}{3}$ but this region has been much less explored as it poses specific analytical difficulties. Our results precisely characterize what happens in terms of an Airy distribution.

The Airy function $\mathrm{Ai}(z)$, as introduced by the Royal Astronomer Sir George Bidell Airy, is a solution of the equation $y'' - zy = 0$ that can be defined by a variety of integral or power series representations including [23]:

$$\mathrm{Ai}(z) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{i(zt + t^3/3)}\, dt = \frac{1}{\pi 3^{2/3}} \sum_{n=0}^{\infty} \frac{\Gamma((n+1)/3)}{n!} \sin \frac{2(n+1)\pi}{3} \left(3^{1/3} x\right)^n. \tag{1}$$

Equipped with this definition, we present the main character of the paper.

Definition 1. *The (standard) "map–Airy" distribution is the probability distribution whose density is*

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

The "map–Airy" distribution of parameter c is defined by its density, $c\mathcal{A}(cx)$.

Note the nonobvious fact that the map–Airy distribution is a probability distribution, *i.e.*, $\int_{\mathbb{R}} \mathcal{A}(x)\, dx = 1$, which can be checked by Mellin transform techniques. An unusual feature is the fact that the tails are extremely asymmetric: $\mathcal{A}(x) = O\left(|x|^{-5/2}\right)$, as $x \to -\infty$, and $\mathcal{A}(x) = O\left(x^{1/2} \exp\left(-\frac{4}{3}x^3\right)\right)$, as $x \to +\infty$.

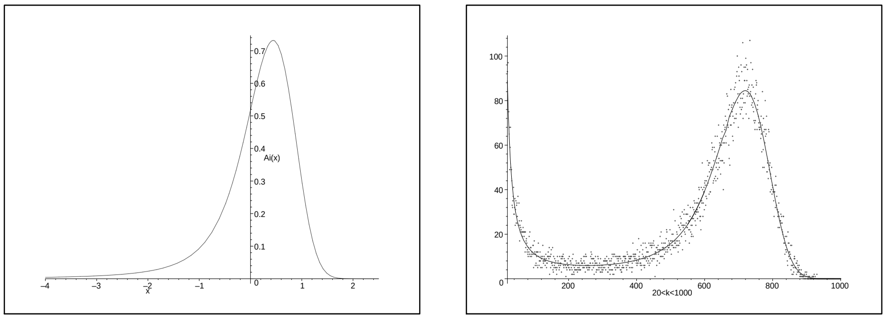

Fig. 1. (i) The map–Airy distribution. (ii) Observed frequencies of core-sizes $k \in [20, 1000]$ in 100,000 random maps of size 2000, against predictions of Thms 3, 4.

We shall find that the size of the core (conditioned upon the large region) is described asymptotically by an Airy law of this type; see Figure 1.

The combinatorics of maps. Let M_n and C_k be the cardinalities of $\mathcal{M}_n$ and $\mathcal{C}_k$. The *generating functions* of $\mathcal{M}$ and $\mathcal{C}$ are respectively defined by

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

(i) *Root-face decomposition.* As shown by Tutte, there results from a root-face decomposition and from the quadratic method [12, Sec. 2.9] that the generating function of $M(z)$ is Lagrangean, which means that it can be parametrized by a system of the form

$$M(z) = \psi(L(z)) \quad \text{where} \quad L(z) = z\phi(L(z)), \tag{2}$$

for two power series ψ, ϕ, with L being determined implicitly by ϕ. For nonseparable maps, we have $\phi(y) = (1+y)^3, \quad \psi(y) = y(1-y)$. There results from the form (2) and from the Lagrange inversion theorem [12] an explicit form for the coefficients of $M(z)$, namely,

$$M_n \equiv [z^n]M(z) = \frac{1}{n}[y^n]\psi'(y)\phi(y)^n, \tag{3}$$

where $[z^n]f(z)$ denotes the coefficient of z^n in the series expansion of $f(z)$. For nonseparable maps, this instantiates to $M_n \equiv [z^n]M(z) = \frac{4(3n)!}{n!(2n+2)!}$.

(ii) *Substitution decomposition.* As shown again by Tutte, maps satisfy additionally relations of the "substitution type": one has: $M(z) = \left(z + \frac{2M(z)^2}{1+M(z)}\right) + C(M(z))$, meaning that each map (left part) either has no core (right part, the first term) or is formed of a nondegenerate core in which maps are substituted (right part, the second term). This equation effectively gives access to the exact enumeration of objects of type $\mathcal{C}$ that are more "complex", *i.e.*, more highly connected than the initial maps of $\mathcal{M}$.

Our interest lies in the probability $\Pr(X_n = k)$ that a map of $\mathcal{M}_n$ has a core with $k+1$ edges. Let $\mathcal{M}_{n,k}$ be the set of maps with this property; we define the *bivariate generating function* $M(z,u) = \sum_{n,k} M_{n,k} u^k z^n$, with $M_{n,k} = \mathrm{card}\,(\mathcal{M}_{n,k})$. Tutte proved the following refinement: $M(z,u) = C(uM(z))$. This determines the *probability distribution of the core-size*:

$$\Pr(X_n = k) = \frac{C_k\,[z^n]M(z)^k}{M_n}, \quad [z^n]M(z)^k = \frac{k}{n}[y^{n-1}]y\psi'(y)\psi(y)^{k-1}\phi(y)^n, \quad (4)$$

where the second equality results from Lagrange inversion.

All the involved generating functions are algebraic functions leading to complicated alternating binomial sums expressing $\Pr(X_n = k)$. The exponential cancellations involved are however not tractable in this elementary way, and complex asymptotic methods must be resorted to.

The asymptotics of maps. There are here two sides to the coin: one evoked now and explored further in Section 5 relies on singularity analysis [9], a method that establishes a general correspondence between the expansion of a generating function at a singularity and the asymptotic form of its coefficients; the other discussed in the next two sections makes use of the power forms provided by the Lagrange inversion theorem that can be exploited asymptotically by the saddle point method.

An implicitly defined function like $L(z)$ in (2) has a singularity of the square-root type $L(z) = \tau - c(1 - z/\rho)^{1/2} + O(1 - z/\rho)$, where the singularity ρ and the singular value τ are determined by the equations $\tau\phi'(\tau) - \phi(\tau) = 0$, $\rho = \frac{\tau}{\phi(\tau)}$. This expansion yields the singular expansion of the generating function of maps,

$$M(z) = \psi(\tau) - a(1 - z/\rho) + b(1 - z/\rho)^{3/2} + O((1 - z/\rho)^2) \quad (5)$$

(in all known map-related cases, one has $\psi'(\tau) = 0$ which induces the singular exponent of 3/2). According to singularity analysis (or the Darboux-Pólya method), this last expansion entails

$$M_n \sim \frac{3b}{4\sqrt{\pi}}\frac{\rho^{-n}}{n^{5/2}}, \ \tau = \frac{1}{2}, \ \rho = \frac{4}{27}, \ M_n^{(\text{nonsep.})} \sim \frac{\sqrt{3}}{2\sqrt{\pi}}\left(\frac{27}{4}\right)^n n^{-5/2} \quad (6)$$

Finally, this approach also yields by direct inversion the asymptotic number of core maps: $C(z)$ has a singularity at $\psi(\tau)$ and singular exponent $\frac{3}{2}$. We have

$$C(z) = c_0 - a'(1 - z/\psi(\tau)) + b'(1 - z/\psi(\tau))^{3/2} + O((1 - z/\psi(\tau))^2), \quad (7)$$

$$C_k = [z^k]C(z) \sim \frac{3b'}{4\sqrt{\pi}}\psi(\tau)^{-k}k^{-5/2}, \ C_k^{(3-\text{conn.})} \sim \frac{8}{243\sqrt{\pi}}4^k k^{-5/2} \quad (8)$$

The foregoing discussion is then conveniently summarized by a statement that constitutes the starting point of our analysis.

Proposition 1. *The distribution of the size of the core in nonseparable maps is characterized by Eq. (4), where the core map counts C_k are determined asymptotically by (8). The basic maps in $\mathcal{M}_n$ are enumerated exactly by (3) and asymptotically by (6).*

2 Two saddles

The probability distribution of core-size in maps is determined by Proposition 1, especially by Equation (4). What is needed is a way to estimate $[z^n]M(z)^k$. The approach starts from the contour integral representation deriving from Cauchy's coefficient formula,

$$[z^n]M^k(z) = \frac{k}{n}\frac{1}{2i\pi}\int_\Gamma z(\psi(z)^k)'\phi(z)^n\frac{dz}{z^{n+1}} = \frac{k}{n}\frac{1}{2i\pi}\int_\Gamma G(z)\,\psi(z)^k\,(\phi(z)/z)^n\;dz \tag{9}$$

where Γ is a contour encircling the origin anticlockwise and $G(z) = \psi'(z)/\psi(z) = (1-2z)/(z(1-z))$.

In simpler cases, integrals over complex contours involving large powers are amenable to the basic saddle point method. The idea consists in deforming the contour Γ in the complex plane, this, in order to have it cross a saddle point of the integrand (*i.e.*, a zero of the derivative) and to take advantage of concentration of the integral near the saddle point. Then local expansions are of the "exponential quadratic" type and the (real-variable) Laplace method permits one to estimate the integral asymptotically [5].

For the problem at hand, there are two saddle points, given by the equation $\frac{\partial}{\partial z}(k\ln\psi + n\ln(\phi/z)) = 0$:

$$z_+(n,k) = \frac{1}{2} \quad\text{and}\quad z_-(n,k) = \frac{n-k}{n+k}.$$

The basic saddle point method applies when these two points are distinct, that is, as long as k/n is "far away" from $\frac{1}{3}$. This corresponds to the situation already well-known from the works of [2, 11, 18]. The "interesting" region is however when $k = n/3$ and when k is close to $n/3$ in the scale of $n^{2/3}$. In that case, the basic version of the saddle point method is no longer applicable. This is precisely where we fit in: we prove that a detailed examination of the analytic geometry of the saddle points in conjunction with suitable integration contours "captures" the major contributions and leads to a precise quantification of core-size in random maps.

Distinct saddles When k is far enough from $n/3$, one of the saddle points is nearer to the origin and predominates. In that case, the basic method applies using a contour that is a circle centered at the origin, passing through the dominant saddle point. This corresponds to the already known results of [2, 11] supplemented by [18].

Theorem 1 (Tails and distinct saddles [11]). *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*

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

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). The right tail ($n > 3k$) has $z_- = (n-k)/(n+k)$ dominating. In each case, the basic saddle point method applies.

A double saddle Here we attack directly the analysis of the "center" of the distribution, that is, the case where $n = 3k$ exactly. Then, the saddle points become equal: $z_- = z_+$. This case serves to introduce with minimal apparatus the enhancements that need to be brought to the basic saddle point method. Observe that the complete confluence of the saddle points precludes the use of "exponential-quadratic" approximations and the problem becomes of an "exponential cubic" type. (See also [2] for a partial discussion of this case based on a method of Van der Corput.)

Theorem 2 (Central part and a double saddle). *The probability distribution of the core of random element of $\mathcal{M}_n$ satisfies, when $n = 3k$,*

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

Proof. When $n = 3k$, equation (9) becomes

$$[z^{3k}]M^k(z) = \frac{1}{6i\pi}\int_\Gamma G(z)P(z)^k dz, \tag{10}$$

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. 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,

$$n\delta^3 \to \infty, \qquad n\delta^4 \to 0, \quad \text{specifically} \quad \delta = (\ln n)n^{-1/3}. \tag{11}$$

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 (10) 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 the constraints on δ (condition $n\delta^3 \to \infty$ in (11)) is exponentially small.

The contribution $\mathcal{E}_{1,2}$ along $\Delta_1 \cup \Delta_2$ to the integral in (10) 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 (11) 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

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

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 $v = u(kd)^{1/3}\exp(2i\pi/3)$ on Δ_1 and $v = u(kd)^{1/3}\exp(-2i\pi/3)$ on Δ_2 then shows that the completed integral equals

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

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.

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 $e = O(1)$.

Nearby saddles When k is close to $n/3$, we choose in the representation (9) an integration contour Γ that catches *simultaneously* the contributions of the two saddle points z_- and z_+. For this purpose, we adopt a contour that goes through the mid-point, $\zeta := (z_- + z_+)/2$, and, like in the previous case, meets the positive real line at an angle of $\pm 2\pi/3$. Local estimates of the integrand, once suitably normalized, lead to a complex integral representation that eventually reduces to Airy functions.

Theorem 3 (Local limit law and nearby saddles). *The probability distribution* $\Pr(X_n = k)$ *admits a local limit law of the map–Airy type: for any real numbers* a, b, *one has*

$$\sup_{a \le \frac{k-n/3}{n^{2/3}} \le b}\left| n^{2/3}\Pr(X_n = k) - \frac{16}{81}\frac{3^{4/3}}{4}\mathcal{A}\left(\frac{3^{4/3}}{4}\frac{k-n/3}{n^{2/3}}\right)\right| \to 0.$$

Proof. We set $k = n/3 + xn^{2/3}$ where x lies in a finite interval of the real line, and define $H := \ln(\psi^{k/n}\phi/z)$ (this replaces $\ln(P)$ in the previous argument). The starting point is again the integral representation (9) taken along a contour Γ that comprises Γ_0, a circle minus a small arc, together with two connecting small segments Δ_1, Δ_2 of length δ, now meeting at ζ, where δ is chosen according to the requirement (11). The arc Γ_0 lies below the level curve of ζ, and the corresponding contribution $\mathcal{E}_0$ is estimated to be exponentially negligible.

We turn next to the contribution $\mathcal{E}_{1,2}$ arising from $\Delta_1 \cup \Delta_2$. The distance between the two saddle points z_-, z_+ is $O(n^{-1/3})$ which represents the "scale" of the problem. One thus sets $z = \zeta + vn^{-1/3}$. Local expansions of H and G are then best carried out with the help (suitably monitored!) of a computer algebra system like `Maple`. The computation relies on the assumption $x = O(1)$, but some care in performing expansions is required because of the relations (11). We find eventually

$$\mathcal{E}_{1,2} = \left(\frac{27}{4}\right)^n 4^{-k} n^{-2/3}\exp\left(-\frac{27}{32}x^3\right)\int_{\Delta_1'\cup\Delta_2'}(9x/2 - 8v)\exp\left(-\frac{64}{27}v^3 - \frac{9}{4}x^2 v\right)(1+\eta)\,dv,$$

where the error term η satisfies $\eta = O(\delta^4 n + n^{-1/12})$ and the segments Δ_1', Δ_2' each have length $\delta n^{1/3}$ tending to infinity according to our assumptions. Perform finally the change of variable $v = \left(\frac{64}{9}\right)^{-1/3} t$ and complete the integration path to $e^{\pm 2i\pi/3}\infty$: the integral then reduces to $\mathrm{Ai}(x), \mathrm{Ai}'(x)$ through contour integrals representations equivalent to (1) (by Cauchy's theorem, with integration path $e^{\pm 2i\pi/3}\infty$ changed to $e^{\pm i\pi/2}\infty$). Thus, for $x = O(1)$ and $k = n/3 + xn^{2/3}$, the main estimate found is

$$[z^n]M^k(z) = \frac{k}{n} 4^{-k} \left(\frac{27}{4}\right)^n n^{-2/3} \frac{3^{4/3}}{4} \mathcal{A}\left(\frac{3^{4/3}}{4} x\right)(1 + o(1)),$$

where $\mathcal{A}(x)$ is the map–Airy density function. This form is equivalent to the statement. The argument also gives a speed of convergence to the limit law of $O(n^{-1/12+o(1)})$.

3 Coalescing saddles

In the present section, we provide a uniform description of the transition regions around $n/3$, allowing k to range anywhere $o(n)$ and $n - o(n)$, precisely, between $\lambda(n)$ and and $n - \lambda(n)$, for any $\lambda(n) = o(n)$ with $\lambda(n) \to \infty$. For the study of this wide region in the scale of n, we set

$$k = \alpha_0 n + \beta n = (1/3 + \beta)n,$$

with estimates valid uniformly for β in any compact subinterval of $]-\frac{1}{3}, \frac{2}{3}[$.

Theorem 4 (Large range and coalescent saddles). *Let $k = n(1/3 + \beta)$, and γ, a_1, a_4 be the functions of β given below. Let $\lambda(n)$ be any function with $\lambda(n) = o(n)$ and $\lambda(n) \to +\infty$. Then, with $\chi = n^{1/3}\gamma$, $\Pr(X_n = n/3 + \beta n)$ equals*

$$\frac{16}{81(1+3\beta)^{3/2} n^{2/3}} \left(\frac{a_1}{2}\mathcal{A}(\chi) + \frac{a_4}{n^{2/3}} \exp\left(-\frac{2}{3}\chi^3\right) \mathrm{Ai}(\chi^2)\right)(1 + O(1/n)), \quad (12)$$

where the error term is uniform for β in any compact subinterval of $]-\frac{1}{3}, \frac{2}{3}[$ and, up to replacing $O(1/n)$ by $O(\lambda(n)^{-1})$, it is also uniform for any $k > \lambda(n)$. With $\mathcal{L}(x) = x \ln x$, the quantities γ, a_1, and a_4 are:

$$\gamma = \left(2\mathcal{L}(1 + 3\beta/4) - \frac{1}{2}\mathcal{L}(1 - 3\beta/2) - \frac{1}{4}\mathcal{L}(1 + 3\beta) - \frac{9}{4}\beta \ln 2\right)^{1/3} \quad (13)$$

$$\frac{a_1}{2} = \frac{9}{8}\left(\frac{\beta/\gamma}{(1 + 3\beta/4)(1 - 3\beta/2)(1 + 3\beta)}\right)^{1/2} \qquad a_4 = \frac{4}{9\beta^2}\sqrt{\frac{\gamma}{\beta}} - \frac{a_1}{4\gamma^2} \quad (14)$$

The estimates involve Airy functions composed with the quantity χ that depends nonlinearly on β. In particular, formula (12) extends the estimates of Section 2 when $k = n/3 + xn^{2/3}$, since in that case $\chi \propto x$ while $\beta \to 0$ and the following approximations apply:

$$\gamma = \frac{3^{4/3}}{4}\beta + O(\beta^2), \quad \frac{a_1}{2} = \frac{3^{4/3}}{4} + O(\beta), \quad a_4 = -\frac{15}{64} 3^{2/3} + O(\beta), \quad \beta \to 0.$$

(The resulting speed of convergence to the Airy law appears to be $O(n^{-2/3})$.) As soon as k leaves the $n/3 \pm O(n^{2/3})$ region, the two Airy terms in (12) start interfering and large deviations are then precisely quantified by (12). When k drifts away to the left of $n/3$ (and $\chi \to -\infty$), basic asymptotics of Airy functions show that the formula simplifies to agree with the results of Section 2.

Proof. The transition phenomenon to be described is the coalescence of two simple saddle points into a double one; see [3, 24]. The simplest occurrence of the phenomenon appears in the integration of $\exp[nf(t,\gamma)]$ with

$$f'(t,\gamma) = t^2 - \gamma^2.$$

Indeed in this case there are two saddle points $\pm\gamma$, coalescing into a double saddle point as $\gamma \to 0$. The strategy consists in performing a change of variable in order to reduce the original problem (9) to this simpler case. Denote the kernel of the integral as $H(z,\beta) = \ln(\psi^{k/n}\phi/z)$ with $k = (1/3 + \beta n)$ and the dependency on β made explicit. The integral in (9) is

$$I(n,\beta) = \int_\Gamma G(z)\exp[nH(z,\beta)]dz,$$

and we seek a change of variable of the form

$$H(z,\beta) = -\left(t^3/3 - \gamma^2 t\right) + r = f(t,\gamma). \tag{15}$$

It turns out that, taking $\gamma = \gamma(\beta)$ to be the real cubic root of $\gamma^3 = \frac{3}{4}[H(z_+,\beta) - H(z_-,\beta)]$, (the relation is expressed by (13)) and $r = r(\beta)$ to be

$$r = \frac{1}{2}[H(z_+,\beta) + H(z_-,\beta)] = H(z_+,\beta) - \frac{2}{3}\gamma^3 = \ln(\psi(\tau)^{k/n}/\rho) - \frac{2}{3}\gamma^3, \tag{16}$$

there exists a conformal map $z \to t$ from the disc D of diameter $[\frac{1}{4}, \frac{3}{4}]$ to a domain D_β satisfying (15) and mapping $z_\pm$ onto $\pm\gamma$. For simplicity, we restrict β to $[-\frac{1}{4}, \frac{1}{4}]$. The domain D_β contains the disc D' of diameter $[-\frac{1}{4}, \frac{1}{4}]$. Let us denote by $z(t)$ the inverse mapping and $G_0(t,\beta) = G(z(t))\dot{z}(t)$ where $\dot{z}(t) = \frac{dz}{dt}$. Remark that $G_0(t,\beta)$ is regular in D'. To guide his intuition, the reader may think of the map $z \to t$ as a slight deformation of the map $z \to 2(z - r)$.

Let us now proceed with the integral. As is usual with saddle point integrals we first localise the integral in D, neglecting the parts of the path down in valleys,

$$I(n,\beta) = \int_\Gamma G(z)\exp[nH(z,\beta)]\,dz = \int_{\Gamma\cap D} G(z)\exp[nH(z,\beta)]\,dz + \mathcal{E}_1(n,\beta),$$

where $\mathcal{E}_1(n,\beta)$ is exponentially negligible when $n \to \infty$, uniformly in β. Inside the disc D we apply the change of variables (15), then restrict attention to the disc D', and deform the contour onto the relevant part of $\Delta_\infty = \{te^{\pm\frac{2i\pi}{3}},\, t \geq 0\}$:

$$\begin{aligned} I(n,\beta) &= \int_{\Gamma_\beta\cap D_\beta} G(z(t))\exp[nf(t,\gamma)]\dot{z}(t)\,dt + \mathcal{E}_1(n,\beta) \\ &= \int_{\Delta_\infty\cap D'} G_0(t,\beta)\exp[nf(t,\gamma)]\,dt + \mathcal{E}_2(n,\beta). \end{aligned}$$

In order to evaluate this integral one needs to dispose of the modulation factor $G_0(t,\beta)$. This can be done via an integration by part: A local expansion near γ yields

$$G_0(t,\beta) = (\gamma - t)a_1 + (t^2 - \gamma^2)H_0(t,\beta),$$

where $H_0(t,\beta)$ is regular in D', and a_1 is given by (14). The integral $I(n,\beta)$ is thus

$$I(n,\beta) = \exp(nr)\int_{\Delta_\infty\cap D'} (\gamma - t)a_1 \exp\left(-n\left(t^3/3-\gamma^2 t\right)\right)\, dt + R_0(n,\beta),$$

where after integration by part, and up to another exponentially negligible term,

$$R_0(n,\beta) = \frac{\exp(nr)}{n}\int_{\Delta_\infty\cap D'}\left(\frac{d}{dt}H_0(t,\beta)\right)\exp\left[-n\left(\frac{t^3}{3}-\gamma^2 t\right)\right] dt + \mathcal{E}_3(n,\beta).$$

The integration by part has reduced the order of magnitude by a factor n, but $R_0(n,\beta)$ is amenable to the same treatment as $I(n,\beta)$. We shall content ourselves with the next terms: let $\frac{d}{dt}H_0(t,\beta) = a_2\gamma + a_3 t + (t^2-\gamma^2)H_1(t,\beta)$, with $H_1(t,\beta)$ regular in D', a_2, a_3 functions of β, so that we have

$$I(n,\beta) = \exp(nr)\int_{\Delta_\infty}\left(\gamma\left(a_1+\frac{a_2}{n}\right) - t\left(a_1-\frac{a_3}{n}\right)\right)\exp\left[-n\left(\frac{t^3}{3}-\gamma^2 t\right)\right] dt + R_1(n,\beta).$$

where the integral has been extended to the whole of Δ_∞ at the expense of yet another exponentially negligible term. The error term is

$$R_1(n,\beta) = \frac{\exp(nr)}{n^2}\int_{\Delta_\infty\cap D'}\left(\frac{d}{dt}H_1(t,\beta)\right)\exp\left[-n\left(\frac{t^3}{3}-\gamma^2 t\right)\right] dt + \mathcal{E}_4(n,\beta).$$

In terms of the Airy function, we thus have

$$I(n,\beta) = 2i\pi\frac{\exp(nr)}{n^{2/3}}\left(\gamma n^{1/3}\left(a_1+\frac{a_2}{n}\right)\mathrm{Ai}(n^{2/3}\gamma^2) - \left(a_1-\frac{a_3}{n}\right)\mathrm{Ai}'(n^{2/3}\gamma^2)\right) + R_1(n,\beta),$$

and the error term $R_1(n,\beta)$ can be estimated: there exist d_0 and d_1 positive such that

$$|R_1(n,\beta)| \le \frac{\exp(nr)}{n^2}\left(\frac{d_0}{n^{1/3}}|\mathrm{Ai}(n^{2/3}\gamma^2)| + \frac{d_1}{n^{2/3}}|\mathrm{Ai}'(n^{2/3}\gamma^2)|\right).$$

The theorem follows from formulae (6), (8), (16) and the definition of the map–Airy law, upon setting $a_4 = \gamma(a_2+a_3)$.

4 Applications to maps and random sampling

The results obtained in the particular case of 3-connected cores of nonseparable maps are instances of a very general pattern in the physics of random maps. Indeed all families in the table below obey the Lagrangean framework and are amenable to the saddle point methods developed in previous sections.

Table 1. A selection of composition schemes ($\mathcal{X}$ an edge, $\mathcal{L},\mathcal{D}$ auxiliary families).

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

Theorem 5. *Consider any scheme of Table 1 with parameters α_0 and c. The probability $\Pr(X_n = k)$ that a map of size n has a core of size k has a local limit law of the map–Airy type with centering constant α_0 and scale parameter c.*

The technique of [11] relates the size of the core to the size of the largest component in random maps. Also, since maps have almost surely no symmetries [17], the analysis extends to unrooted maps. As a consequence:

Theorem 6. *(i) Consider any scheme of Table 1 with parameters α_0 and c. Let X_n^* be the size of the largest component of in a random map of size n with uniform distribution. Then*

$$\Pr\left(X_n^* = \lfloor \alpha_0 n + x n^{2/3} \rfloor\right) = \frac{c\mathcal{A}(cx)}{n^{2/3}}\,(1 + O(n^{-2/3})),$$

uniformly for x in any bounded interval. Furthermore, if x is restricted to the shorter range $|x| < \lambda(n)^{-1}$ for a fixed function $\lambda(n)$ going to infinity with n, then

$$\Pr\left(X_n^* = \lfloor \alpha_0 n + x n^{2/3} \rfloor\right) = \frac{c}{n^{2/3}} \frac{3^{1/6}\Gamma\left(2/3\right)}{\pi}\,(1 + O(\lambda(n)^{-1})).$$

(ii) The same results hold for random unrooted maps.

Theorem 6 extends results of Bender, Gao, Richmond, and Wormald [2, 11] who proved that X_n^* lies in the range $\alpha_0 n \pm \lambda(n) n^{2/3}$ with probability tending to 1, where $\lambda(n)$ is any function going to infinity with n.

Random sampling algorithms for various families of planar maps were described in [19]. For general, nonseparable, bipartite, and cubic nonseparable maps, an algorithm `Map` is given there that takes an integer n and outputs in linear time a map of size n uniformly at random. For the other families of Table 1, a probabilistic algorithm `Core` described below is used.

Probabilistic algorithm `Core(k)` with parameter $f(k)$
1. use `Map(n)` to generate a random map $M \in \mathcal{M}$ of size $n = f(k)$;
2. extract the largest component C of M with respect to the scheme;
3. if C does not have size k, then go back to step 1;
4. output C.

Safe for a set of measure that is exponentially small, this algorithm produces a uniform element of $\mathcal{C}_k$. The expected number of loops made by `Core` is exactly $\ell_n = \Pr(X_n = k)^{-1}$. The results of the paper enable us to precisely analyse this and a number of related algorithms of [18, 19]. We cite just here:

Theorem 7. *In all extraction/rejection algorithms of [19], the choice $f(k) = n/\alpha_0$ yields an algorithm whose average number of iterations satisfies*

$$\ell_n \sim n^{2/3}/(\mathcal{A}(0)c).$$

Let $x_0 \approx 0.44322$ be the position of the peak of the map-Airy density function ($(1 - 4x_0^3)\mathrm{Ai}(x_0^2) + 4x_0^2\mathrm{Ai}'(x_0^2) = 0$). The optimal choice $f(k) = k/\alpha_0 - \frac{x_0}{\alpha_0 c}(k/\alpha_0)^{2/3}$ reduces the expected number of loops by $1 - \mathcal{A}(0)/\mathcal{A}(x_0) \approx 30\%$.

This proves that the extraction/rejection algorithms have overall complexity $O(k^{5/3})$, as do variant algorithms of [18, 19] that are uniform over all $\mathcal{C}_k$. The complexity becomes $O(k)$ if some small tolerance is allowed on the size of the multiply connected map generated. Theorems of the paper enable us to quantify precisely various trade-offs and fine-tune algorithms (details in the full paper). As exemplified by Fig. 1(ii), the predictions fit nicely with experimental results.

5 Composition of Singularities

Map enumeration can be approached through the Lagrangean framework and the saddle point analysis developed so far takes off from there. An alternative approach to the problem relies on singularity analysis [9], as introduced in Section 1. The results of this section contribute to the general classification of combinatorial schemas according to the nature of their singularities [20].

First, a definition. Let M and C be two generating functions with dominant singularities at ρ and σ, such that $M(z) = \sigma - a(1-z/\rho) + b(1-z/\rho)^{3/2} + O((1-z/\rho)^2)$, and $C(z) = c_0 - a'(1-z/\sigma) + b(1-z/\sigma)^{3/2} + O((1-z/\sigma)^2)$, in an indented domain extending beyond the circle of convergence (see [9]). Then the bivariate substitution scheme $C(uM(z))$ is said to be a *critical composition scheme* of type $(3/2, 3/2)$. The functional composition $C(uM(z))$ describes the size of the $\mathcal{C}$ component in a combinatorial substitution $\mathcal{C}[\mathcal{M}]$. The scheme is called critical since the singular value of the inner function (M) equals the singularity of the outer function (C). It will be recognized that Tutte's construction is an instance (with σ replacing the map specific $\psi(\tau)$ of formulae (5) and (7)). Schemes of this broad form have been only scantily analysed, a notable exception being the critical composition scheme of type $(-1, 3/2)$ that shows up in ordered forests and in random mappings (functional graphs): in that case, the density is known to be of the Rayleigh type [6, 20]. The results of this section somehow recycle in a different realm the intuition gathered by the method of coalescing saddles, although the technical developments are a bit different.

Theorem 8. (*i*) *For* $k = \alpha n + \lambda(n)$, *with* $0 \leqslant \alpha < \alpha_0 = \frac{\sigma}{a}$ *and* $\lambda(n) = o(n)$, *the probability distribution of the size* X_n *of the* $\mathcal{C}$*–component of random element of* $\mathcal{C}[\mathcal{M}]$ *of size* n *satisfies*

$$\Pr(X_n = k) \sim \frac{b'}{\sigma\sqrt{\pi}}\, \alpha^{-3/2}(1-\alpha/\alpha_0)^{-5/2}\, n^{-3/2} \qquad \text{for } \alpha > 0;$$

$$\Pr(X_n = k) \sim \frac{b'}{\sigma\sqrt{\pi}}\, \lambda(n)^{-3/2} \qquad \text{for } \alpha = 0 \text{ and } \lambda(n) \to +\infty\,.$$

(*ii*) *For* $k = \alpha_0 n + x n^{2/3}$, $\alpha_0 = \sigma/a$, $x = o(n^{1/3})$, *an Airy-map law holds:*

$$n^{2/3} \Pr(X_n = \alpha_0 n + x n^{2/3}) \;\sim\; \frac{b'}{\alpha_0^{3/2} b} c\mathcal{A}(cx) \qquad \textit{where } c = \left(\frac{b}{3a}\right)^{2/3} / \alpha_0.$$

The proof relies on a modification of the Hankel contour used in classical singularity analysis together with a different scaling. It will be developed in the full paper. The theorem is a companion to Theorems 1, 2, 3, 4 that can also be used to analyse forests of unrooted trees [14] in the critical region, a problem itself relevant to the emergence of the giant component in random graphs [13, 14].

References

1. Bender, C. M., and Orszag, S. A. *Advanced mathematical methods for scientists and engineers. I.* Springer-Verlag, New York, 1999. Asymptotic methods and perturbation theory, Reprint of the 1978 original.
2. Bender, E. A., Richmond, L. B., and Wormald, N. C. Largest 4-connected components of 3-connected planar triangulations. *Random Structures & Algorithms 7*, 4 (1995), 273–285.
3. Bleistein, N., and Handelsman, R. A. *Asymptotic Expansions of Integrals.* Dover, New York, 1986. A reprint of the second Holt, Rinehart and Winston edition, 1975.
4. Bollobás, B., Borgs, C., Chayes, J. T., Kim, J. H., and Wilson, D. B. The scaling window of the 2-sat transition. Preprint, 1999. Available as document math.CO/9909031 at the Los Alamos archive `http://xxx.lanl.gov/form/`.
5. De Bruijn, N. G. *Asymptotic Methods in Analysis.* Dover, 1981. A reprint of the third North Holland edition, 1970 (first edition, 1958).
6. Drmota, M., and Soria, M. Marking in combinatorial constructions: Generating functions and limiting distributions. *Theoretical Comput. Sci. 144*, 1–2 (June 1995), 67–99.
7. Duchon, P. Q-grammars and wall polyominoes. *Annals of Combinatorics 3* (1999), 311–321.
8. Flajolet, P., Knuth, D. E., and Pittel, B. The first cycles in an evolving graph. *Discrete Mathematics 75* (1989), 167–215.
9. Flajolet, P., and Odlyzko, A. M. Singularity analysis of generating functions. *SIAM Journal on Applied Mathematics 3*, 2 (1990), 216–240.
10. Flajolet, P., Poblete, P., and Viola, A. On the analysis of linear probing hashing. *Algorithmica 22*, 4 (Dec. 1998), 490–515.
11. Gao, Z., and Wormald, N. C. The size of the largest components in random planar maps. *SIAM J. Discrete Math. 12*, 2 (1999), 217–228 (electronic).
12. Goulden, I. P., and Jackson, D. M. *Combinatorial Enumeration.* John Wiley, New York, 1983.
13. Janson, S., Knuth, D. E., Luczak, T., and Pittel, B. The birth of the giant component. *Random Structures & Algorithms 4*, 3 (1993), 233–358.
14. Kolchin, V. F. *Random Graphs*, vol. 53 of *Encyclopedia of Mathematics and its Applications.* Cambridge University Press, Cambridge, U.K., 1999.
15. Louchard, G. The Brownian excursion: a numerical analysis. *Computers and Mathematics with Applications 10*, 6 (1984), 413–417.
16. Prellberg, T. Uniform q-series asymptotics for staircase polygons. *Journal of Physics A: Math. Gen. 28* (1995), 1289–1304.
17. Richmond, L. B., and Wormald, N. C. Almost all maps are asymmetric. *J. Combin. Theory Ser. B 63*, 1 (1995), 1–7.
18. Schaeffer, G. *Conjugaison d'arbres et cartes combinatoires aléatoires.* PhD thesis, Université Bordeaux I, 1998.
19. Schaeffer, G. Random sampling of large planar maps and convex polyhedra. In *Proceedings of the thirty-first annual ACM symposium on theory of computing (STOC'99)* (Atlanta, Georgia, may 1999), ACM press, pp. 760–769.
20. Soria-Cousineau, M. *Méthodes d'analyse pour les constructions combinatoires et les algorithmes.* Doctorate in sciences, Université de Paris–Sud, Orsay, July 1990.
21. Takacs, L. A Bernoulli excursion and its various applications. *Advances in Applied Probability 23* (1991), 557–585.
22. Tutte, W. T. Planar enumeration. In *Graph theory and combinatorics (Cambridge, 1983).* Academic Press, London, 1984, pp. 315–319.
23. Watson, G. N. *A Treatise on the Theory of Bessel Functions.* Cambridge University Press, 1980.
24. Wong, R. *Asymptotic Approximations of Integrals.* Academic Press, 1989.