Analytic Combinatorics of Non-crossing Configurations

Philippe Flajolet
Algorithms Project, INRIA Rocquencourt, F-78153 Le Chesnay (France)
Marc Noy
Dept. Applied Mathematics, Universitat Politècnica de Catalunya,
Pau Gargallo, 5, 08028 Barcelona (Spain)

March 2, 1998

Abstract

This paper describes a systematic approach to the enumeration of "non-crossing" geometric configurations built on vertices of a convex n-gon in the plane. It relies on generating functions, symbolic methods, singularity analysis, and singularity perturbation. Consequences are both exact and asymptotic counting results for trees, forests, graphs, connected graphs, dissections, and partitions. Limit laws of the Gaussian type are also established in this framework; they concern a variety of parameters like number of leaves in trees, number of components or edges in graphs, etc.

Introduction

The enumeration of planar configurations defined on vertices of a convex n-gon has a long and dignified history. In 1753, Euler and Segner counted triangulations —the well-known answer involves the Catalan numbers— and on this occasion Euler invented combinatorial generating functions. Since then, many other configurations have been enumerated: see for instance Comtet's book [7], for an account of known results. The interest for such configurations comes first and foremost from the combinatorics of classical structures [7], but also from computational geometry, and even the interpretation of perturbative expansions in statistical physics [10].

The purpose of this paper is to re-examine these problems in the light of recent general methods of *analytic combinatorics* [17, 33]. First, thanks to symbolic methods developed by various schools [4, 17, 18, 21, 33, 35, 38], there is a systematic and purely formal correspondence between combinatorial constructions and *generating functions*. In this way, specifications of combinatorial structures can be translated automatically into generating function equations. This approach is, as we propose to show, especially effective here, since planarity entails neat decompositions for the configurations to be enumerated. Second, analytic methods based on the analysis of singularities [16] give a transparent access to asymptotic counts that plainly appear as morphic images of the local expansions of generating functions near a singularity.

This programme is carried out here on six of the most basic planar "non-crossing" configurations: trees and forests, graphs and connected graphs, dissections and partitions. The generating functions involved are all *algebraic functions*, a property to be somewhat expected given the context-free character of these objects. However, their forms are sometimes more complicated than what is encountered in the Catalan domain comprehensively reviewed by Gould in [20]. *Singularity analysis* then makes it possible to derive precise estimates; see especially our Theorem 4. In addition, a general approach of "singularity perturbation asymptotics" [15] permits us to refine the counting estimates and derive *limit laws* for many parameters of interest.

Given the vast literature on the subject, we cannot expect to derive only new results; our hope is that the unified treatment presented here could be of methodological interest and that the present paper could also serve as a partial survey of the enumerative, asymptotic, and probabilistic aspects of non-crossing configurations. The analytic approach followed here, when contrasted to more classical combinatorial bijective proofs, proves especially effective when exact formulæ either become too intricate or fade away.

In the first sections of this paper, numbered 1,2,3, we make explicit the basic decompositions of the six fundamental types of planar configurations considered. We characterize in each case the counting generating functions by the minimal polynomial equation they satisfy, which serves two goals: first, this leads to explicit counting results; second, the equations can be fed into the asymptotic machinery of Section 4, leading eventually to the precise asymptotic estimates of Theorem 4. In addition, many parameters of interest are easily taken into account by bivariate generating functions, the corresponding equations serving as input to the bivariate asymptotic process of Section 5. A consequence, stated in Theorem 5, is that all the parameters discussed, e.g., the number of edges or components in non-crossing graphs of a fixed size, have distributions that are Gaussian in the asymptotic limit.

Combinatorial preliminaries. Let $\Pi_n = \{v_1, v_2, \ldots, v_n\}$ be a fixed set of points in the plane, conventionally ordered counter-clockwise, that are vertices of a convex polygon, for instance, the vertices of a regular n-gon. Define a *non-crossing graph* of size n as a graph that has vertex set Π_n and whose edges are straight line segments that do not cross. Several classical combinatorial objects can be viewed as non-crossing graphs (we omit the qualifier non-crossing from now on). For instance, triangulations of a convex polygon are graphs with the maximum number of edges; dissections of a convex polygon are graphs containing the edges $v_1v_2, v_2v_3, \ldots, v_nv_1$; non-crossing partitions are graphs whose components are either points, edges, or cycles. Note that the vertices have a fixed ordering and that two graphs are considered different even if they are equivalent up to a symmetry of the regular polygon. (See for instance [27] for the enumeration of planar configurations considered equivalent under reflection.)

We recall that a graph is connected if any two vertices can be joined by a path. A tree is a connected acyclic graph, and the number of edges in a tree is one less than the number of vertices. A forest is an acyclic graph, or a graph whose components are trees.

Let $\mathcal{A}$ be a class of combinatorial objects and let $|\alpha|$ be the size of an object $\alpha \in \mathcal{A}$. If $\mathcal{A}_n$ denotes the objects in $\mathcal{A}$ of size n and $a_n = |\mathcal{A}_n|$, then the (ordinary) *generating function*, GF for short, of the class $\mathcal{A}$ is

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

Here, the size of a graph is its number of vertices and we consider various classes of non-crossing graphs.

There is a direct correspondence between set-theoretic operations (or "constructions") on combinatorial classes and algebraic operations on GFs. For an exposition of the symbolic enumeration method, see for instance [17, 33]. Table 1 summarizes this correspondence for the

Construction		*Operation on GF*
Union	$\mathcal{A} = \mathcal{B} \cup \mathcal{C}$	$A(z) = B(z) + C(z)$
Product	$\mathcal{A} = \mathcal{B} \times \mathcal{C}$	$A(z) = B(z)C(z)$
Sequence	$\mathcal{A} = \mathsf{Seq}(\mathcal{B})$	$A(z) = 1/(1 - B(z))$
Substitution	$\mathcal{A} = \mathcal{B} \circ \mathcal{C}$	$A(z) = B(C(z))$

Table 1: The basic combinatorial constructions and their translation into generating functions.

operations that are used in the paper. There "union" means union of disjoint copies, "product" is the usual cartesian product, "sequence" forms sequences, and "substitution" $\mathcal{A} = \mathcal{B} \circ \mathcal{C}$ corresponds to grafting objects of $\mathcal{C}$ on nodes of $\mathcal{B}$.

Enumerations according to size and an auxiliary parameter χ are described by bivariate generating functions, or BGFs,

$$A(z, w) = \sum_{\alpha \in \mathcal{A}} z^{|\alpha|} w^{\chi(\alpha)} = \sum_{n,k \geq 0} A_{n,k} z^n w^k,$$

with $A_{n,k}$ the number of objects of size n with χ-parameter equal to k. Throughout the paper the variable z is reserved for marking the number of vertices (the size) of the different kinds of graphs, and the variable w is reserved for marking a secondary parameter, like leaves in trees or edges in graphs. Classes and their GFs are consistently denoted by the same letters.

We will need repeatedly the Lagrange-Bürmann inversion theorem in order to extract coefficients of GFs that satisfy functional equations of the implicit type [7, 21, 33, 38]:

Lagrange inversion. *Let $\phi(u)$ be a formal power series with $\phi(0) \neq 0$, and let $Y(z)$ be the unique formal power series solution of the equation $Y = z\phi(Y)$. Then the coefficient of z in $\psi(Y)$, for an arbitrary series ψ, is given by*

$$[z^n]\psi(Y(z)) = \frac{1}{n}[u^{n-1}]\phi(u)^n \psi'(u).$$

In particular, for every $k > 0$ we have

$$[z^n]Y(z)^k = \frac{k}{n}[u^{n-k}]\phi(u)^n.$$

Lagrange inversion obviously applies to bivariate generating functions upon treating the auxiliary variable as a parameter.

1 Trees and forests

In this section a tree means a non-crossing tree, and a forest means a non-crossing forest. Trees are considered *rooted* at vertex v_1 and the degree of a vertex in a tree is its out-degree;

leaves are vertices of degree zero. As mentioned above, throughout the paper the *size* of a non-crossing graph is its number of vertices.

Basic decompositions reflect the geometric structure of trees and forests (Fig. 1), which leads to algebraic generating functions that prove to be amenable to Lagrange expansion.

Theorem 1 (*i*) *The number of non-crossing trees of size n is equal to*

$$T_n = \frac{1}{2n-1}\binom{3n-3}{n-1},$$

and the number of non-crossing trees of size n and k leaves is equal to

$$T_{n,k} = \frac{1}{n-1}\binom{n-1}{k}\sum_{j=0}^{k-1}\binom{n-1}{j}\binom{n-k-1}{k-1-j}2^{n-2k+j}.$$

(*ii*) *The number of trees with degree partition $(n_0, n_1, \ldots, n_r)$, where $\sum n_i = n$ and $\sum in_i = n-1$, is equal to*

$$\frac{1}{n(n-1)}\binom{n}{n_0, n_1 \ldots, n_r}1^{n_0}2^{n_1}\cdots(r+1)^{n_r}\sum_{i=1}^{r}\frac{i}{i+1}n_i.$$

(*iii*) *The number of forests of size n is*

$$F_n = \sum_{j=1}^{n}\frac{1}{2n-j}\binom{n}{j-1}\binom{3n-2j-1}{n-j}, \tag{1}$$

and the number of forests of size n and k components is

$$F_{n,k} = \frac{1}{2n-k}\binom{n}{k-1}\binom{3n-2k-1}{n-k}. \tag{2}$$

(*iv*) *The GF of forests, the BGF of trees and leaves, and the BGF of forests and components, are algebraic functions given by (10), (6) and (11).*

Trees were first enumerated by Dulucq and Penaud [12], and their result is summarized in part (*i*) of the theorem; the enumeration of forests by GF in (10) below is due to Noy [29]. We recover both results, as well as several new ones in the form of multivariate extensions. In particular, the counting of trees according to the number of leaves as stated in (*i*) solves a problem that was left open in [29]. The explicit forms for the number of forests in part (*iii*), formulæ (1) and (2), provide explicit expansions for the GF computations of [29].

Trees. We use the following basic decomposition for counting trees. Let d be the degree of v_1 in a tree τ. Then τ can be viewed as a sequence attached to v_1 of d ordered pairs of trees sharing a common vertex. This motivates the following definition: a *butterfly* is an ordered pair of trees with a common root. The name aims to convey the idea that the pair of trees looks like the two wings of a butterfly. If v_1 has degree d, then τ can be identified with a sequence of d butterflies hanging from v_1. In the example of Fig. 1a there are 3 butterflies, rooted at x, y and z. Observe that the left wing of the butterfly at y is reduced to a point.

If $T(z)$ is the GF for trees and $B(z)$ is the GF for butterflies, we have the following equations:

$$\begin{aligned} T &= \frac{z}{1-B} \\ B &= T^2/z. \end{aligned} \tag{3}$$

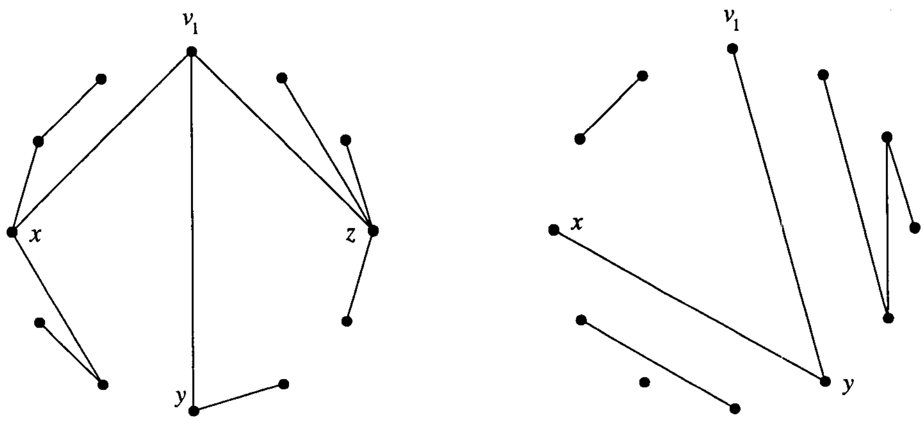

Figure 1: (a) A tree; (b) A forest.

The first equation corresponds to the sequence construction, and the factor z to the root v_1. The division by z in the second equation is needed since we identify two root vertices to form a butterfly. It follows that T satisfies

$$T^3 - zT + z^2 = 0. \tag{4}$$

If we set $z = \zeta^2, T = \zeta U$, the equation becomes $U - U^3 = \zeta$, a direct case of application of the Lagrange inversion theorem. As a consequence, we get the first assertion in part (i) of the theorem. An alternative transformation that is useful for the sequel is as follows. Set $T = z + zy$, and "solve" for z in terms of y; this gives

$$z = \frac{y}{(1+y)^3}, \qquad y = z(1+y)^3, \tag{5}$$

which is amenable to Lagrange inversion. These derivations also show that T_{n+1} is the number of ternary trees drawn in the plane (without reference to a fixed convex polygon) that have n internal nodes. This can be proved by a more direct decomposition argument (as in [12]); the interest of introducing butterflies is that the same decomposition technique proves useful later.

Let now $T(z,w)$ and $B(z,w)$ be the bivariate generating functions of trees and butterflies, respectively, where z marks vertices and w marks leaves. Then we have

$$\begin{aligned} T(z,w) &= \frac{z}{1-B}; \\ B(z,w) &= \frac{T^2}{z} - z + zw. \end{aligned}$$

The first equation is the same as (3), since the number of leaves in τ is just the sum of the number of leaves in the sequence of butterflies defining τ. The second equation reflects the fact that we have a leaf when the two wings of a butterfly are empty. Hence the term z in $B(z,w)$ has to be replaced with zw. Eliminating B we obtain

$$T^3 + (z^2w - z^2 - z)T + z^2 = 0. \tag{6}$$

Expansion of $T(z,w)$ can be carried out by the same process as in (5). Set $T = z + zy$, and solve for z, which gives

$$z = \frac{y}{(y+1)(y^2+2y+w)}, \qquad y = z(y+1)(y^2+2y+w).$$

Then, by Lagrange inversion, one has

$$[z^n]T(z,w) = [z^{n-1}]y = \frac{1}{n-1}[u^{n-2}]\left((u+1)(u^2+2u+w)\right)^{n-1},$$

and upon extracting $[w^k]$,

$$\begin{aligned}[z^n w^k]T(z,w) &= \frac{1}{n-1}[u^{n-2}w^k]\left((u+1)(u^2+2u+w)\right)^{n-1} \\ &= \frac{1}{n-1}\binom{n-1}{k}[u^{n-2}](u+1)^{n-1}(u^2+2u)^{n-1-k}.\end{aligned}$$

This last form yields directly the expression of $T_{n,k}$ stated in part (i) of the theorem.

The *degree partition* of a tree τ of size n and maximum degree r is the sequence $(n_0, n_1, \ldots, n_r)$, where n_i is the number of vertices of degree i in τ, for $i = 0, \ldots, r$. Clearly $\sum n_i = n$ and, since the number of edges is $n-1$, $\sum i n_i = n-1$. Given a sequence of non-negative integers $(n_0, n_1, \ldots, n_r)$ with $\sum n_i = n$ and $\sum i n_i = n-1$, we consider the problem of determining the number of trees of size n having partition $(n_0, n_1, \ldots, n_r)$.

To solve this problem we have to look again at butterflies. A butterfly β has a left and a right tree with a common vertex v. If d is the degree of v, then β can be seen in turn as a sequence of d butterflies attached to v. There are $d+1$ ways of distributing them among the left and right trees, hence we have $B = z(1 + 2B + 3B^2 + \cdots)$. Let now $u_0, u_1, \ldots$ be a sequence of formal variables, where u_i marks a vertex of degree i, either in trees or in butterflies. Then the equation becomes

$$B = z(u_0 + 2u_1 B + 3u_2 B^2 + \cdots + (r+1)u_r B^r + \cdots), \tag{7}$$

where $B = B(z, u_0, u_1, \ldots)$ is a GF in an infinite number of variables. On the other hand, the basic equation (3) becomes

$$T = z(u_0 + u_1 B + u_2 B^2 + \cdots + u_r B^r + \cdots). \tag{8}$$

Using Lagrange inversion in (7) we find that

$$[u_0^{n_0} u_1^{n_1} \cdots u_r^{n_r} z^n](z u_k B^k) = \frac{k}{n-1}\binom{n-1}{n_0, \ldots, n_k - 1, \ldots, n_k} 1^{n_0} 2^{n_1} \cdots (k+1)^{n_k - 1} \cdots (r+1)^{n_r}.$$

Now we use (8) to express the coefficient of $[u_0^{n_0} u_1^{n_1} \cdots u_r^{n_r} z^n]$ in T as the sum of the above expression for $k = 1, \ldots, r$. A straightforward manipulation gives the final compact solution stated in part (ii) of the theorem.

Forests. A forest is an acyclic graph, i.e., a graph whose connected components are trees. Let ϕ be a forest and let r be the number of vertices in the component τ containing v_1. Then ϕ has to be completed with r additional forests (some of them possibly empty), one to the left of every vertex of τ. An example is shown in Fig. 1b. Here $r = 3$ and the forest consists of τ and 3 additional forests of sizes $2, 3$ and 4 to the left of v_1, x and y, respectively.

Thus the class of forests is obtained from the class of trees by substituting a vertex by a pair (vertex, forest). If F is the GF of forests, then the substitution construction yields

$$F = 1 + T(zF), \tag{9}$$

where T is the GF of trees as before, and 1 is the GF of the empty forest of size 0. Since T satisfies (4) one can eliminate T and recover a result from [29] (the equation here is marginally different since we are taking the constant term of F to be 1):

$$F^3 + (z^2 - z - 3)F^2 + (z+3)F - 1 = 0. \tag{10}$$

In order to expand, we set $F = 1 + y$, then "solve" for z, which yields

$$y = z(1+y)\left(\frac{1-\sqrt{1-4y}}{2y}\right),$$

an equation of the Lagrange type that also suggests a Catalan tree decomposition for non-crossing forests. Formula (1) then results from the Lagrange expansion of powers of the Catalan GF.

Let now $F(z, w)$ be the bivariate GF for forests, where w marks components. We only have to include the factor w in (9) to take into account the component of v_1 that was singled out, to obtain $F(z, w) = 1 + wT(zF)$. Eliminating T as before we get

$$F^3 + (w^3z^2 - w^2z - 3)F^2 + (w^2z + 3)F - 1 = 0. \tag{11}$$

This equation also admits a Lagrange form, upon setting $F = 1 + wy$,

$$y = z(1+wy)\left(\frac{1-\sqrt{1-4y}}{2y}\right),$$

hence the explicit formula for $F_{n,k}$ in part (*iii*). Note that (1) follows again from (2) by summation on k. Finally, we remark that counting edges instead of components is an equivalent problem, since the number of edges in a forest is equal to the number of vertices minus the number of components.

2 Connected graphs and general graphs

As before, a graph means a non-crossing graph. Planarity once more entails strong decomposition properties (Fig. 2) reflected by algebraic generating functions and Lagrange expansions.

Theorem 2 (*i*) *The number of connected graphs of size n is given by*

$$C_n = \frac{1}{n-1}\sum_{j=n-1}^{2n-3}\binom{3n-3}{n+j}\binom{j-1}{j-n+1}. \tag{12}$$

The number of connected graphs of size n with k edges is given by

$$C_{n,k} = \frac{1}{n-1}\binom{3n-3}{n+k}\binom{k-1}{k-n+1}. \tag{13}$$

(ii) The number of graphs of size $n \geq 3$ is expressible in terms of Schröder numbers,

$$G_n = 2^n c_{n-1}, \qquad c_n := \sum_{0 \leq \nu \leq (n/2)} (-1)^\nu \frac{1 \cdot 3 \cdots (2n-2\nu-3)}{\nu!\,(n-2\nu)!} 3^{n-2\nu} 2^{-\nu-2}, \tag{14}$$

the number of graphs of size n with k edges is

$$G_{n,k} = \frac{1}{n-1} \sum_{j=0}^{n-2} \binom{n-1}{k-j} \binom{n-1}{j+1} \binom{n-2+j}{n-2}, \tag{15}$$

and the number of graphs of size n with k connected components is

$$\widehat{G}_{n,k} = \frac{1}{n} \binom{n}{k-1} \sum_{j=0}^{n-k} \binom{n+j-1}{j} \binom{2n-2k-j}{n-k} \frac{j 2^{n-k-j}}{2n-2k-j}. \tag{16}$$

(iii) The BGFs of connected graphs and the BGF of graphs counted according to edges are algebraic functions given by (18) and (22). The BGF of graphs and number of connected components is an algebraic function given by (23).

Note. Our original derivation for $C_{n,k}$ in (13) involved a summation instead of the closed form stated in the theorem. We are grateful to an anonymous referee for pointing out the current form that is simpler, together with the derivation that follows Eq. (17) and (18) below.

The univariate generating functions of connected graphs and general graphs were obtained by Domb and Barrett [10] after considerable effort. In both cases, these authors also obtained the bivariate GF according to the number of edges, building upon the work of the Rev. T. P. Kirkman in 1857; see [10] for a thorough historical discussion. We recover all the results of [10] plus two new ones, namely the enumeration of graphs according to the number of components by GF (part *(iii)*) and an explicit formula for the number of connected graphs (part *(i)*). The result concerning $G_{n,k}$ is roughly equivalent to Kirkman's results in view of Eq. (9–10) of [10], while the one concerning $\widehat{G}_{n,k}$ seems to be new. Our approach to this problem is a direct adaptation of the scheme we used for counting trees and forests, and as such it is purely "algebraic"; in contrast, in [10], recourse had to be had to a combination of algebraic and differential arguments. The Schröder numbers[1] c_n count generalized bracketings (equivalently, plane trees with n leaves and internal nodes of degree ≥ 2), and they are defined in [7, p. 57].

Connected graphs. We use a decomposition technique analogous to that for counting trees. Let d be the degree of vertex v_1 in a connected graph Γ, and let v_i and v_j be two consecutive neighbours of v_1 in Γ. Then the subgraph induced on the vertex set $\{v_i, v_{i+1}, \ldots, v_j\}$ is either a connected graph (not reduced to a point), or two disjoint connected graphs containing v_i and v_j, respectively. The two possibilities are exemplified in Fig. 2a, where $d = 3$. The graph induced between x and y has two components, each of size 2, and the graph between y and z consists of a single component of size 3.

If we let C be the GF for connected graphs, the first possibility is counted by $C - z$, and the second one by C^2. If v_i is the first neighbour of v_1 then one has a connected graph on $\{v_2, \ldots, v_i\}$, whereas if v_j is the last neighbour of v_1 one has a connected graph

[1]Stanley observes in a vivid account [36] that the 10th Schröder number 103,049 was already known to Hipparchus in the second century B.C., as recorded by Plutarch in the first century A.D.

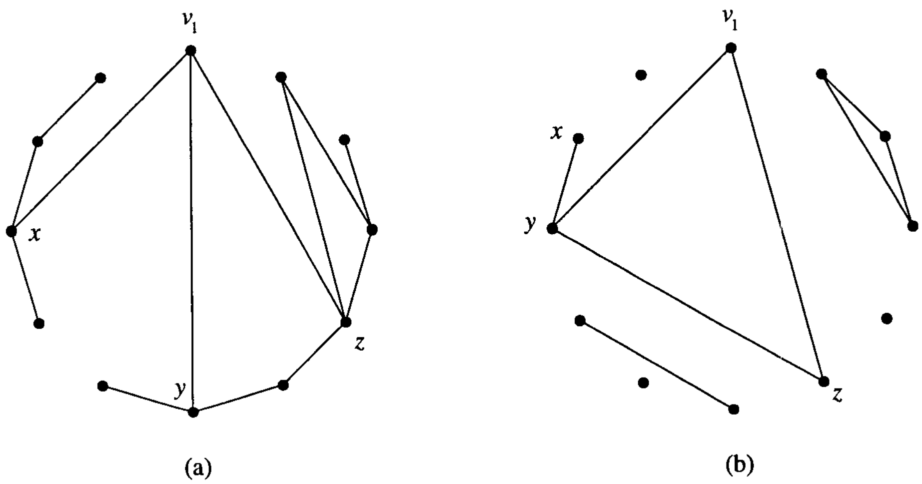

Figure 2: (a) A connected graph; (b) An arbitrary graph.

on $\{v_{j+1}, \ldots, v_n\}$. Taking into account that the d neighbours of v_1 are counted twice, we obtain

$$\begin{aligned} C &= z + z\frac{C^2}{z} + z\frac{C^2(C-z+C^2)}{z^2} + \cdots + z\frac{C^2(C-z+C^2)^{d-1}}{z^d} + \cdots \\ &= z\left(1 + \frac{C^2}{z-(C-z+C^2)}\right). \end{aligned}$$

Simplification gives

$$C^3 + C^2 - 3zC + 2z^2 = 0. \tag{17}$$

Setting $C = zy + z$ the equation becomes

$$z(1+y)^3 = y - y^2 = y(1-y).$$

Expanding the Lagrange form

$$y = z\frac{(1+y)^3}{1-y}$$

we obtain the formula stated in part (i).

Let now $C(z, w)$ be the GF for connected graphs, where w marks edges. If v_1 has degree d we have to introduce a factor w^d in the corresponding summand before (17), and a simple computation gives

$$wC^3 + wC^2 - z(1+2w)C + z^2(1+w) = 0. \tag{18}$$

This is equation (47) of [10]. Setting $C = zy + z$ it becomes

$$wz(1+y)^3 = y(1-wy),$$

and the Lagrange form

$$y = z\frac{w(1+y)^3}{1-wy}$$

gives the formula stated in part (i).

Observe that (12) follows directly from (13), and that the extreme values of k in (12) correspond to trees ($k = n - 1$) and to triangulations ($k = 2n - 3$).

Graphs. Let Γ be a graph and let r be the number of vertices in the component Γ_1 containing v_1. Then Γ has to be completed with r additional graphs (some of them possibly empty), one to the left of every vertex of Γ_1. For instance, $r = 4$ in Fig. 2b, where the graph to the left of x is empty. The graphs to the left of v_1, y and z have sizes $1, 3$ and 4, respectively.

If G is the GF of graphs and C the GF of connected graphs as above then, as in the case of forests, the substitution construction gives

$$G = 1 + C(zG). \tag{19}$$

Taking into account that C satisfies (17), we can eliminate C and obtain (after cancelling a factor G),

$$G^2 + (2z^2 - 3z - 2)G + 3z + 1 = 0. \tag{20}$$

This equation appears in [10], but in a slightly different form since we are taking the constant term of G to be 1. Solving the quadratic yields

$$G(z) = 1 + \frac{3}{2}z - z^2 - \frac{z}{2}\sqrt{1 - 12z + 4z^2}, \tag{21}$$

which is a recognizable variant of the GF of Schröder numbers [7].

Using this scheme we can easily enumerate graphs according to the number of edges. Let w mark edges, and let $C(z,w)$ be the bivariate GF for connected graphs. Then (19) becomes

$$G(z,w) = 1 + C(zG(z,w), w),$$

because the number of edges in a graph is simply the sum of the number of edges in its components. Eliminating C in (18) we arrive at

$$wG^2 + ((1+w)z^2 - (1+2w)z - 2w)G + w + z(1+2w) = 0. \tag{22}$$

This equation becomes amenable to Lagrange inversion upon the change of variables $G = 1 + z + zy$ that transforms it into

$$y = z(1+w)\left(\frac{1+y}{1-wy}\right).$$

Similarly, let now w mark components. Then (19) becomes

$$G(z,w) = 1 + wC(zG(z,w)),$$

where $C(z)$ is the univariate GF for connected graphs and the factor w takes into account the component containing v_1. Eliminating C in (17) we arrive at

$$G^3 + (2w^3z^2 - 3w^2z + w - 3)G^2 + (3w^2z - 2w + 3)G + w - 1 = 0. \tag{23}$$

The explicit expansion obeys principles similar to what has been done before. Set $G = 1 + wy$, solve for z, and obtain the Lagrange form,

$$y = 4z(1+yw)\left(3 + \sqrt{1-8y}\right)^{-1}.$$

What is required now is an expansion of the negative powers of $q(u) = 3 + \sqrt{1-8u}$. A change of variables similar to the one that underlies Lagrange inversion in Cauchy coefficient integrals, namely $q(u) = 4 - 4t$, then shows that

$$[u^a]4^b q(u)^{-b} = [t^a](1-t)^{-b}(1-2t)^{-a-1}(1-4t).$$

The rest of the computation is routine.

3 Dissections and Partitions

A *dissection* of a convex polygon $\Pi_n = \{v_1, v_2, \ldots, v_n\}$ is a partition of Π_n into polygonal regions by means of non-crossing diagonals; that is, a non-crossing graph containing the edges $v_1v_2, v_2v_3, \ldots, v_nv_1$. A *non-crossing partition* of size n is a partition of $[n] = \{1, 2, \ldots, n\}$ such that if $a < b < c < d$ and a block contains a and c, then no block contains b and d. One can draw such a partition on a circle by representing each block as a convex polygon on the points belonging to the block. Then non-crossing partitions are the same as non-crossing graphs whose connected components are points, edges and cycles. (We do not consider here triangulations as they have been investigated so extensively since Euler's time; see [20]).

Theorem 3 *(i) For $n \geq 3$, the number of dissections of size n is the Schröder number c_{n-1} defined in (14). It can also be expressed as*

$$D_n = \frac{1}{n-1}\sum_{i=0}^{n-2}(-1)^i\binom{n-1}{i}\binom{2n-4-i}{n-2-i}2^{n-2-i}.$$

The number of dissections of size $n \geq 3$ with k regions satisfies

$$D_{n,k} = \frac{1}{k}\binom{n-3}{k-1}\binom{n+k-2}{k-1}.$$

(ii) The number of non-crossing partitions of size n is a Catalan number,

$$P_n = \frac{1}{n+1}\binom{2n}{n},$$

and the number of partitions of size n with k blocks is a Narayana number

$$P_{n,k} = \frac{1}{n}\binom{n}{k}\binom{n}{k-1}.$$

The results in the above theorem are all classical and can be found in many sources. We include them in order to show how the general methodology allows easy derivations. Kreweras first discussed non-crossing partitions in [26] while results for dissections are summarized in [7, p. 74]. See also [5] for more references and related work.

Dissections of a convex polygon. Let δ be a dissection of Π_n and let ρ be the region containing the edge v_1v_2. If ρ has $r+1$ sides, then δ is identified with a sequence of $r \geq 2$ dissections (some of them possibly reduced to a single edge) attached to ρ. In the example of Fig. 3a the region ρ is shaded and $r = 4$. The dissections in the sequence have sizes 2 ("empty"), 4, 6, and 3, in anticlockwise order.

If z^2 marks the "empty" dissection consisting of a single edge, then

$$D = z^2 + \frac{D^2}{z} + \cdots + \frac{D^r}{z^{r-1}} + \cdots \qquad (24)$$

where the denominator z^{r-1} indicates that $r-1$ pairs of vertices have been identified. Summation and simplification gives

$$2D^2 - z(1+z)D + z^3 = 0.$$

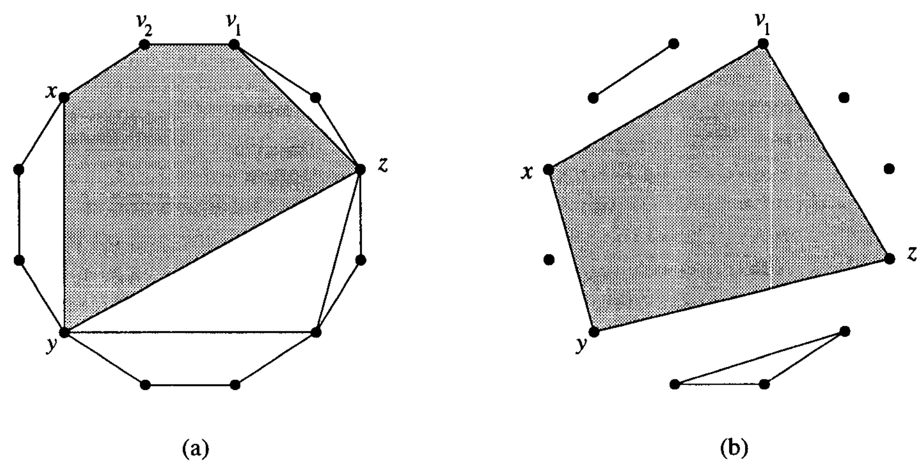

Figure 3: (a) A dissection; (b) A non-crossing partition.

Solving the quadratic equation yields again a GF that is a variant of the GF of Schröder numbers.

There is an alternative way to expand the GF, not to be found in Comtet's book [7]. Set $D = zy$. Then y satisfies an equation similar to (24),

$$y = z + \frac{y^2}{1-y} \qquad \text{or} \qquad z = y\frac{1-2y}{1-y}.$$

This equation is of the Lagrange type and it can be subjected to inversion,

$$[z^n]y(z) = \frac{1}{n}[u^{n-1}]\left(\frac{1-u}{1-2u}\right)^n,$$

which gives the first relation of part (i). This relation also reveals a combinatorial curiosity: the quantity $nc_n = n[z^n]y(z)$ equals the number of n-tuples of integer compositions with grand total sum equal to $n-1$.

Let now z mark vertices and w mark regions. Then (24) becomes

$$D = z^2 + w\left(\frac{D^2}{z} + \frac{D^3}{z^2} + \cdots\right),$$

where the factor w marks the region containing v_1v_2. This is equivalent to

$$(1+w)D^2 - z(1+z)D + z^3 = 0.$$

As before, we set $y(z,w) = D(z,w)/z$ and get

$$y = z + w\frac{y^2}{1-y}, \qquad y = z\left(1 - w\frac{y}{1-y}\right)^{-1}.$$

This is again an equation of the Lagrange type and inversion gives

$$[z^n]y(z,w) = \frac{1}{n}[u^{n-1}]\left(1 - w\frac{u}{1-u}\right)^{-n}.$$

From there, the explicit form stated in part (i) results by extracting the coefficient of w^k. We remark that $D_{n,k}$ is also the number of plane trees of the Schröder type, built on $n-1$ external nodes that have k internal nodes, each of degree ≥ 2.

Let us also remark that once we know how to enumerate dissections we can enumerate general graphs. Indeed, a graph is the set of internal diagonals of a dissection plus any set of boundary edges. As a consequence, the number of graphs of size $n \geq 3$ is $G_n = 2^n D_n$. If the graph has k edges, j of them are internal diagonals and $k-j$ are boundary edges. Hence we obtain $G_{n,k} = \sum_{j=0}^{k} \binom{n}{j} D_{n,j+1}$ as an alternative to the formula stated in Theorem 2.

Non-crossing partitions. Let π be any non-crossing partition and let r be the size of the block β containing vertex v_1. Then π can be encoded as a sequence of r partitions (some of them possibly empty), one to the left of every point in β. In the example of Fig. 3b the block β is shaded and $r=4$. The four partitions have sizes $2, 1, 3$ and 2.

If P is the GF of non-crossing partitions and 1 denotes the empty partition, then

$$P = \frac{1}{1 - zP},$$

and of course we recover the GF for the Catalan numbers (see [20]),

$$zP^2 - P + 1 = 0,$$

with the corresponding Lagrange form for $y = zP$ that reads $y = z(1-y)^{-1}$.

If z marks vertices and w marks blocks, then

$$P = 1 + wzP + wz^2P^2 + \cdots = 1 + \frac{wzP}{1 - zP},$$

and we get

$$zP^2 + (wz - z - 1)P + 1 = 0.$$

With $y = zP$, this can be written as $y = z(1 + wy/(1-y))$, and Lagrange inversion gives the classical Narayana numbers,

$$[z^n w^k]P(z,w) = \frac{1}{n}\binom{n}{k}\binom{n}{k-1},$$

that also enumerate general plane trees of size $n+1$ that have k leaves.

We close this section, and the first part of the paper, with two tables. Table 2 gives some numerical values for the six basic configurations, and Table 3 summarizes the various GF equations that we have encountered in the first three sections. Four out of the six sequences appear in the *Encyclopedia of Integer Sequences* (*EIS*, [34]) and the last column gives the corresponding entries in *EIS*.

n	1	2	3	4	5	6	7	8	9	10	11	*EIS*
T_n	1	1	3	12	55	273	1428	7752	43263	246675	1430715	**M2926**
F_n	1	2	7	33	181	1083	6854	45111	305629	2117283	14929212	—
C_n	1	1	4	23	156	1162	9192	75819	644908	5616182	49826712	**M3594**
G_n	1	2	8	48	352	2880	25216	231168	2190848	21292032	211044352	—
D_n	0	1	1	3	11	45	197	903	4279	20793	103049	**M2898**
P_n	1	2	5	14	42	132	429	1430	4862	16796	58786	**M1459**

Table 2: Numerical values for the six basic configurations.

Configuration	*GF equation*
Trees	$T^3 - zT + z^2 = 0$
—, leaves	$T^3 + (z^2w - z^2 - z)T + z^2 = 0$
Forests	$F^3 + (z^2 - z - 3)F^2 + (z + 3)F - 1 = 0$
—, components	$F^3 + (w^3z^2 - w^2z - 3)F^2 + (w^2z + 3)F - 1 = 0$
Connected graphs	$C^3 + C^2 - 3zC + 2z^2 = 0$
—, edges	$wC^3 + wC^2 - z(1 + 2w)C + z^2(1 + w) = 0$
Graphs	$G^2 + (2z^2 - 3z - 2)G + 3z + 1 = 0$
—, edges	$wG^2 + ((1 + w)z^2 - (1 + 2w)z - 2w)G + w + z(1 + 2w) = 0$
—, components	$G^3 + (2w^3z^2 - 3w^2z + w - 3)G^2 + (3w^2z - 2w + 3)G + w - 1 = 0$
Dissections	$2D^2 - z(1 + z)D + z^3 = 0$
—, regions	$(1 + w)D^2 - z(1 + z)D + z^3 = 0$
Partitions	$zP^2 - P + 1 = 0$
—, blocks	$zP^2 + (wz - z - 1)P + 1 = 0$

Table 3: GF equations: z marks vertices and w marks the secondary parameter.

4 Asymptotic counting

In this section, we prove that each class of non-crossing configurations leads to an asymptotic estimate of the form

$$f_n \sim \gamma \frac{\omega^n}{\sqrt{\pi} n^{3/2}}, \tag{25}$$

where f_n is the number of objects of size n, and γ, ω are context-dependent algebraic numbers. Such estimates are for instance familiar in the theory of tree enumerations [11, 22, 28, 30].

Roughly, each of the six counting generating functions is an algebraic function, as seen in Sections 1,2,3. It is known that the singularities of GFs determine the asymptotics of their coefficients. Here, we *a priori* expect local singular expansions in the form of Puiseux expansions, that is to say expansions involving fractional exponents. Generically, singularities of the square-root type are expected, as in many implicitly defined functions [11, 22]. All our GFs appear to be of this type, with a local expansion near the dominant singularity ρ being

$$f(z) \sim c_0 + c_1\sqrt{1 - z/\rho}. \tag{26}$$

Then singularity analysis [16] is used to achieve the transfer of (26) to coefficients leading to estimates of the form (25).

Rather than examining each case separately, we develop here a common strategy that is adequate for treating all classes discussed in previous sections (in one case, the argument needs to be mildly amended) and is systematic enough to be amenable to treatment by a computer algebra system.

Theorem 4 *Consider the configurations of trees, forests, connected graphs, graphs, dissections, and partitions. The corresponding counts each satisfy an asymptotic estimate of the form*

$$f_n = \gamma \frac{\omega^n}{\sqrt{\pi} n^{3/2}} \left(1 + \mathcal{O}(\frac{1}{n})\right),$$

where γ, ω are algebraic numbers given in Table 4.

The asymptotic counting of graphs was obtained by Domb and Barrett [10] using Darboux's method; the asymptotic form of Schröder numbers is certainly known to many and is close to the framework of simple families of trees introduced by Meir and Moon [28]. The asymptotics of trees and partitions can be directly obtained from explicit formulæ and Stirling's approximation. The present approach is introduced because it has the merit of providing a global approach while lending itself naturally to a perturbation analysis that leads to Gaussian laws, as discussed in the next section.

PROOF. The generating functions considered so far satisfy a system of polynomial equations. They are then *algebraic* functions, since from classical elimination theory any system can be reduced to a single polynomial equation,

$$P(z, y) = 0, \qquad P \in \mathbf{Q}[z, y], \tag{27}$$

and reduction to such a form may be achieved systematically by either resultant or Groebner basis elimination [19]. Here, our combinatorial specifications being simple enough, elimination is immediate, so that the form (27) is directly available from previous sections.

	Class	ω	*Num. value*	γ
(T)	Trees	$\frac{27}{4}$	6.75000	$\frac{\sqrt{3}}{27}$
(F)	Forests	$\frac{1}{\xi}$	8.22469	0.07465
(C)	Connected graphs	$6\sqrt{3}$	10.39230	$\frac{\sqrt{6}}{9} - \frac{\sqrt{2}}{6}$
(G)	Graphs	$6+4\sqrt{2}$	11.65685	$\frac{1}{4}\sqrt{-140+99\sqrt{2}}$
(D)	Dissections	$3+2\sqrt{2}$	5.82842	$\frac{1}{4}\sqrt{-140+99\sqrt{2}}$
(P)	Partitions	4	4.00000	1

Table 4: The constants appearing in the statement of Theorem 4. There, ξ denotes the root of the polynomial $4 - 32x - 8x^2 + 5x^3$ that is near 0.121, and 0.07465 represents the explicit algebraic number of degree 6 equal to $\beta/2$, with β given in the text.

Consider a polynomial equation

$$P(z,y) \equiv \sum_{j=0}^{d} a_j(z)y^j = 0. \tag{28}$$

It has in general (that is, except for a finite set of exceptional values) d distinct solutions that are then analytic branches of a complex algebraic curve; see for instance the discussion of the Weierstrass Preparation Theorem in [1] or [23].

A finite set Ω of candidate singularities can be determined systematically by a general process explained below. The problem is then to determine which of the elements of Ω are dominant singularities (that is, singularities of smallest modulus) of the branch that coincides with the counting generating function under study and is thereby identified by its expansion at 0. In all generality, such a determination implies solving a so-called connection problem between branches [6]. However, the problems under consideration are once more simple enough, so that Ω can be "filtered" and reduced, in each case, to a single element by means of elementary arguments. We find that each generating function $f(z)$ has a unique dominant and positive real singularity at some $\rho > 0$ near which it satisfies an expansion of the square-root type,

$$f(z) = c_0 + c_1(1 - z/\rho)^{1/2} + c_2(1 - z/\rho) + \mathcal{O}((1 - z/\rho)^{3/2}). \tag{29}$$

Then, by Darboux's method [7, 22] or singularity analysis [16], *transfer* from the singular expansion (29) to coefficients is permissible and

$$[z^n]f(z) = \frac{c_1}{\Gamma(-1/2)} \frac{\rho^{-n}}{\sqrt{n^3}} \left(1 + \mathcal{O}(\frac{1}{n})\right), \tag{30}$$

a form that matches (25) with $\omega = \rho^{-1}$ and $\gamma = -c_1/2$.

The last phase of asymptotic transfer is a standard one. We thus concentrate on the problem of singularity localization and singular expansion and refer to the papers by Klarner and Woodworth [25] as well as by Canfield [6] for background.

A partial algorithm. The polynomial equation $P(z,y) = 0$ has in general d roots or branches for a fixed value of z. When the leading coefficient $a_d(z)$ vanishes, some of these branches escape to infinity and are thus potential singularities. Singularities may otherwise only arise at points z such that the two equations

$$P(z,y) = 0, \qquad \frac{\partial}{\partial y}P(z,y) = 0$$

have a common root y. In this case, two branches meet and there is possibly a branch point. Such places where branches meet are thus zeros of the resultant polynomial [2],

$$R(z) := \text{Result}_y\left(P(z,y), \frac{\partial}{\partial y}P(z,y)\right). \tag{31}$$

At all other points, there are d distinct branches that are each analytic by Weierstrass preparation. Then, a superset of the set of singularities is

$$\Omega = \{z \mid R(z)\cdot a_d(z) = 0\}. \tag{32}$$

The generating functions of non-crossing configurations all have a radius of convergence in the interval $[0,1]$ since their coefficients satisfy combinatorial bounds of the form $A^n < f_n < B^n$, for some A, B with $1 < A < B < \infty$. Thus, one need only consider

$$\Omega_1 = \Omega \cap \{z \mid |z| < 1\},$$

which must contain at least one positive element ρ. (Pringsheim's theorem asserts that a function with nonnegative coefficients is singular at its radius of convergence [37].) If Ω_1 has cardinality 1, a unique dominant singularity has been found[3]. We thus assume the uniqueness condition to be satisfied.

In all cases under consideration, the function $f(z)$ remains finite at its singularity since $a_d(\rho) \neq 0$. We set

$$\tau := \lim_{z\to\rho^-} f(z),$$

so that τ also equals the quantity c_0 in (26). The quantity τ is a double root of $P(\rho, y) = 0$ and it has to be positive. It is thus a root of the resultant polynomial

$$S(y) := \text{Result}_z\left(P(z,y), \frac{\partial}{\partial y}P(z,y)\right). \tag{33}$$

(If these conditions are not sufficient, at least τ could be isolated by carefully controlled numerical analysis of $f(z)$ for $z \in (0,\rho)$.)

By the general theory of algebraic functions [23], a Puiseux expansion —an expansion into fractional powers, that is, into powers of $(1 - z/\rho)^{1/r}$— holds locally at $z = \rho$, for some integer $r > 1$. Such an expansion derives explicitly from the bivariate expansion of $P(z,y)$ at (ρ,τ),

$$P(z,y) = p_{00} + p_{10}Z + p_{01}Y + p_{20}Z^2 + p_{11}ZY + p_{02}Y^2 + \cdots, \tag{34}$$

[2]The resultant of a polynomial with its derivative is its "discriminant" and its non-vanishing is a well-known criterion for distinct roots.

[3]This situation covers five out of our six cases. The exception is the case of connected graphs where $\Omega_1 = \{-1/(6\sqrt{3}),\ 1/(6\sqrt{3})\}$, but for which a rational parametrization, $z = x(1-x)(1-2x)$ and $C = z(1-x)^{-1}$, permits us to eliminate the negative value from the set of candidate singularities by simply following the branch at the origin that corresponds to the combinatorial GF. (Domb and Barrett [10] do not address this issue explicitly.) Alternatively, one could appeal to the powerful theorems of Drmota [11].

$$p_{ij} := \frac{1}{i!\,j!}\frac{\partial^{i+j}}{\partial z^i \partial y^j} P(z,y)\bigg|_{(\rho,\tau)}, \qquad Z = z-\rho,\ Y = y-\tau.$$

By assumption, $p_{00} = p_{01} = 0$. Provided the condition,

$$p_{02} \neq 0, \tag{35}$$

holds, then the dependency between Y and Z is locally quadratic, and as $z \to \rho$,

$$f(z) = c_0 + c_1(1-z/\rho)^{1/2} + \mathcal{O}((1-z/\rho)), \qquad c_0 = \tau, \quad c_1 = -\left(\frac{2\rho\, p_{10}}{p_{02}}\right)^{1/2} \tag{36}$$

(The minus sign in c_1 must be adopted here since the generating function increases with its argument.)

In summary, if the condition (35) is satisfied, then the singular expansion (29) holds, and the asymptotic forms of coefficients (25,30) have been established. Condition (35) is itself satisfied generically and is easily checked numerically in each individual case. The coefficients in the expansions are then all explicitly computable algebraic numbers. □

The above programme has been carried out for all non-crossing configurations defined in previous sections. Computations have been performed under the Maple system for symbolic manipulations, together with the `Gfun` extension due to Salvy and Zimmermann [32]. In particular, the `Gfun` package provides automatically Puiseux expansions of algebraic functions, a great help here.

Here is an outline of the computation for the case of forests, where $y(z) = T(z)$ is defined by (10). There, some care is needed in selecting correct algebraic conjugates amongst various possibilities. The basic GF equation is (10). The resultant polynomial $R(z)$ defined in (31) is found mechanically to be

$$R(z) = -z^3(4 - 32z - 8z^2 + 5z^3),$$

whose roots are the four algebraic numbers,

$$\Omega = \{0,\ -1.93028,\ 0.12158,\ 3.40869\}$$

(approximately). Therefore, a unique dominant singularity of $F(z)$ has been isolated,

$$\Omega_1 = \{\xi \doteq 0.12158,\ 4 - 32\xi - 8\xi^2 + 5\xi^3 = 0\}.$$

The three branches of the cubic give rise at $z = \rho$ to one branch that is analytic when $z = \xi$, with value numerically close to 0.67816, and two conjugate branches with value 1.21429 at $z = \rho$. The expansion of the two conjugate branches starts as

$$\alpha \pm \beta\sqrt{1 - z/\xi} + \cdots,$$

where

$$\alpha = \frac{43}{37} + \frac{18}{37}\xi - \frac{35}{74}\xi^2 \doteq 1.21429, \qquad \beta = \frac{1}{37}\sqrt{228 - 981\xi - 5290\xi^2} \doteq 0.14931,$$

and the determination with the minus sign must be taken for the combinatorial GF. The computation can be conveniently based upon `Gfun`'s ability to determine Puiseux expansions. The data for our six families are summarized in Table 4.

5 Limit laws

The six basic combinatorial types of Sections 1–3 give rise to seven basic parameters for which BGFs $f(z,w)$ have been found to satisfy polynomial equations of the form

$$P(z,w,f(z,w)) = 0.$$

These equations, together with a few initial conditions provided by the combinatorics of the problems, fully determine the BGFs. The problem of estimating the coefficients

$$f_{n,k} = [z^n u^k] f(z,w)$$

is then a bivariate asymptotic problem.

The quantities

$$\pi_{n,k} = \frac{f_{n,k}}{f_n},$$

represent discrete probability distributions. Let μ_n and σ_n^2 be the mean and variance of such a distribution $\pi_{n,k}$. Classically, the distribution $\pi_{n,k}$ is said to be *asymptotically normal* (or Gaussian) if, pointwise for each $x \in \mathbf{R}$,

$$\lim_{n\to\infty} \sum_{k \le \mu_n + x\sigma_n} \pi_{n,k} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt. \tag{37}$$

In other words, the distribution of the random variable X_n representing parameter χ taken on non-crossing configurations of size n, has a distribution function that, after normalization, tends to the Gaussian distribution function. We establish now that our seven reference parameters all have laws that are asymptotically normal. For background information on these analytic techniques, we refer globally to [2, 3, 11, 24] and the exposition in [14] or [17, Ch. 9]. (Extremal parameters, like maximum degree or longest chord, require different techniques; see [9].)

Theorem 5 *Consider the following parameters: number of leaves in trees, components in forests, edges in connected graphs, components in graphs, edges in graphs, regions in dissections, blocks in partitions. The corresponding distributions over objects of size n each have mean μ_n and variance σ_n^2 that satisfy*

$$\mu_n \sim \kappa n, \qquad \sigma_n^2 \sim \lambda n,$$

where κ, λ are algebraic numbers given in Table 5. The laws are in each case asymptotically normal.

PROOF. As seen in the proof of Theorem 4, each of the counting GFs $f(z)$ has a unique dominant singularity ρ that is of the square-root type, see (26,29). This in turn entails, by singularity analysis, that the various types of non-crossing configurations all obey an asymptotic formula of the form (30).

Consider a parameter χ like the number of leaves, edges, components, etc, and let $f(z,w)$ be the corresponding bivariate GF. Our goal is to establish a lifted form of the singular expansion (29),

$$f(z,w) = c_0(w) + c_1(w)\sqrt{1 - z/\rho(w)} + \mathcal{O}(1 - z/\rho(w)), \tag{38}$$

Class, Parameter	κ (mean)		λ (variance)	
Trees, leaves	$\frac{4}{9}$	0.444	$\frac{28}{243}$	0.115
Forests, components	$\frac{8}{37}-\frac{13}{37}\xi+\frac{15}{74}\xi^2$	0.176	$\frac{192}{1369}+\frac{5}{2738}\xi-\frac{47}{2738}\xi^2$	0.140
Connected graphs, edges	$\frac{1}{2}+\frac{\sqrt{3}}{2}$	1.366	$\frac{1}{4}$	0.250
Graphs, edges	$\frac{1}{2}+\frac{\sqrt{2}}{2}$	1.207	$\frac{1}{4}+\frac{\sqrt{2}}{8}$	0.426
Graphs. components	$\frac{5}{7}-\frac{3}{7}\sqrt{2}$	0.108	$\frac{50}{2401}+\frac{255}{4802}\sqrt{2}$	0.095
Dissections, regions	$\frac{\sqrt{2}}{2}$	0.707	$\frac{\sqrt{2}}{8}$	0.176
Partitions, blocks	$\frac{1}{2}$	0.500	$\frac{1}{8}$	0.125

Table 5: The constants appearing in the statement of Theorem 5. There, ξ denotes the root near 0.121 of the polynomial $4-32z-8z^2+5z^3$.

uniformly with respect to w for w in a small neighbourhood of 1, and with $\rho(w), c_0(w), c_1(w)$ analytic at $w=1$. There, $\rho(w)$ is the dominant singularity (assumed to be unique) of $f(z,w)$, where w is treated as a parameter. If (38) is granted, then, by singularity analysis,

$$f_n(w) := [z^n]f(z,w) = \gamma(w)\left(\frac{1}{\rho(w)}\right)^n\left(1+O(\frac{1}{n^{1/2}})\right), \tag{39}$$

for some analytic function $\gamma(w)$, with an error term that is uniform with respect to w. Uniformity is crucial and is guaranteed in full generality by the constructive character of the singularity analysis method. (See the discussion in [16].)

The probability generating function of χ satisfies

$$q_n(w) := \frac{f_n(w)}{f_n} = \frac{\gamma(w)}{\gamma(1)}\left(\frac{\rho(1)}{\rho(w)}\right)^n\left(1+O(\frac{1}{\sqrt{n}})\right). \tag{40}$$

This means that $q_n(w)$ is a so-called "quasi-power". In particular, the mean $\mu_n = q_n'(1)$ and the variance $\sigma_n^2 = q_n''(1)+q_n'(1)-q_n'(1)^2$ result by differentiation of (40), so that

$$\kappa = -\frac{\rho'(1)}{\rho(1)}, \qquad \lambda = -\frac{\rho''(1)}{\rho(1)} - \frac{\rho'(1)}{\rho(1)} + \left(\frac{\rho'(1)}{\rho(1)}\right)^2. \tag{41}$$

Then, by extensions due to Bender, Richmond and Hwang of the central limit theorem, a limiting Gaussian law for the distribution of χ results from (40). Basically, from the quasi-powers form, the normalized characteristic functions $\phi_n(t) = e^{-it\mu_n/\sigma_n}q_n(e^{it/\sigma_n})$ converge to the characteristic function of a standard normal distribution, namely $e^{-t^2/2}$. The limit law then follows as a consequence of the continuity theorem for characteristic functions.

At this stage, the proof of the theorem is completed as soon as one can establish the lifted expansion (38) for each of the seven parameters under consideration. The proof relies on the permanence of analytic relations under "perturbation" by an auxiliary parameter, a property that is technically granted by the Weierstrass preparation theorem.

Consider the lifted version of the resultant of (31),

$$R(z,w) = \mathrm{Result}_y\left(P(z,y,w), \frac{\partial}{\partial z}P(z,y,w)\right). \tag{42}$$

This is a polynomial whose restriction $P(z,1)$ has, by the developments of the proof of Theorem 4 and the companion computations, a simple isolated root at $z=\rho$. By the implicit function theorem and the Weierstrass preparation theorem [1, 23], this root lifts to a simple root near ρ that is an analytic branch $\rho(w)$ of an algebraic function, for w in a small neighbourhood of 1:

$$R(\rho(w),w)=0, \qquad \rho(1)=1. \tag{43}$$

Then, by Weierstrass preparation again, the analytic factorization

$$P(z,y) = (y^2+m_1(z)y+m_2(z))\cdot H(z,y),$$

with $H(\rho,\tau)\neq 0$, that corresponds to a square root singularity, lifts to

$$P(z,y,w) = (y^2+m_1(z,w)y+m_2(z,w))\cdot H(z,y,w),$$

with $H(\rho,\tau,1)\neq 0$. Then, the quadratic formula yields

$$f(z,w)=\frac{1}{2}\left(-m_1(z,w)-\sqrt{m_1(z,w)^2-4m_2(z,w)}\right).$$

It then suffices to expand $f(z,w)$ near $(\rho(w),w)$ in order to get the uniform family of singular expansions (38), hence eventually, the Gaussian limit law[4]. □

Globally, the process discussed here is one of "singularity perturbation" where one establishes that the singular expansion of a BGF has a smooth analytic behaviour when the auxiliary parameter w varies in a small neighbourhood of 1. Computationally, the process is simple. The algebraic function $\rho(w)$ is determined by Eq. (43). The regular expansion of the branch that coincides with ρ at $w=1$ provides the first two moments.

For instance, for edges in connected graphs, the algebraic equation is (18). The resultant polynomial is found to be

$$R(z,w)=w^2z^2\left(27w(w+1)^2z^2+2(w-1)(2w+1)(w+2)z-w\right).$$

The expansion of $\rho(w)$ at $w=1$ is determined by the implicit function theorem, and its coefficients are simply rational functions of ρ as $\rho(w)$ is analytic. The computation is again conveniently handled by the `Gfun` package of Maple,

$$\rho(w)=\frac{1}{18}\sqrt{3}-\left(\frac{1}{12}+\frac{1}{36}\sqrt{3}\right)(w-1)+\left(\frac{1}{12}+\frac{5}{144}\sqrt{3}\right)(w-1)^2+\mathcal{O}((w-1)^3).$$

[4]In addition, by the Berry-Esseen inequalities [13], the speed of convergence to the Gaussian limit is $\mathcal{O}(n^{-1/2})$ uniformly.

The result found is then best expressed in logarithmic-exponential form (see [8]), where the mean and variance coefficients of (41) read directly:

$$\log\left(\frac{\rho(1)}{\rho(e^s)}\right) = \kappa s + \frac{1}{2}\lambda s^2 + \mathcal{O}(s^3) = \left(\frac{1}{2} + \frac{\sqrt{3}}{2}\right) s + \frac{1}{8}s^2 + \mathcal{O}(s^3).$$

This gives $\kappa = \frac{1}{2} + \frac{\sqrt{3}}{2}$ and $\lambda = \frac{1}{4}$. The data for the seven parameters under consideration are all obtained in this way and summarized in Table 5.

6 Conclusion

Symbolic methods in combinatorial enumerations lead in many cases to easy derivations of generating function equations. This observation applies specially to non-crossing configurations, since planarity constraints and the distinguishable character of vertices entail strong decomposition properties. As a result, the generating functions are all algebraic. Singularity analysis and singularity perturbation methods then allow for a transparent treatment that is also computationally effective. A graphical illustration of the chain of computations is presented in Fig. 4.

It is clear that a large number of similar problems are amenable to this chain. Instances are leaves in forests and isolated points or vertices in graphs, for which Gaussian laws can be proved to hold by the methods employed here. Trees whose degrees are bounded by some fixed integer b can be enumerated for each fixed b, their generating functions remain algebraic, and similarly for 1-regular and 2-regular graphs. In all these cases, symbolic methods in conjunction with complex asymptotics allow for a concise and unified characterization of properties of random structures, a distinctive feature of analytic combinatorics.

Note added. After this paper was ready for publication, the authors have learnt of an interesting paper [31], where the author obtained, among other results, formulæ (12) and (13).

Acknowledgements. This work was supported in part by the Long Term Research Project *Alcom-IT* (# 20244) of the European Union and by Project SEUID PB96-0005-C02-02 of the Spanish Government. The authors are grateful to Frédéric Cazals for his contribution in establishing the data in Tables 4, 5 and his help with random generation and plots. Special thanks are also due to anonymous referees for a very careful reading of the manuscript and for constructive suggestions; in particular one referee found the simplified form of the enumeration of connected graphs that is stated in Theorem 2.

References

[1] ABHYANKAR, S.-S. *Algebraic geometry for scientists and engineers.* American Mathematical Society, 1990.

[2] BENDER, E. A. Central and local limit theorems applied to asymptotic enumeration. *Journal of Combinatorial Theory 15* (1973), 91–111.

[3] BENDER, E. A., AND RICHMOND, L. B. Central and local limit theorems applied to asymptotic enumeration II: Multivariate generating functions. *Journal of Combinatorial Theory,* Series A *34* (1983), 255–265.

[4] BERGERON, F., LABELLE, G., AND LEROUX, P. *Théorie des espèces et combinatoire des structures arborescentes.* No. 19 in Publications du LACIM. Université du Québec à Montréal, 1994.

[5] BERNHART, F. R. Catalan, Motzkin, and Ring Numbers. (For this volume.)

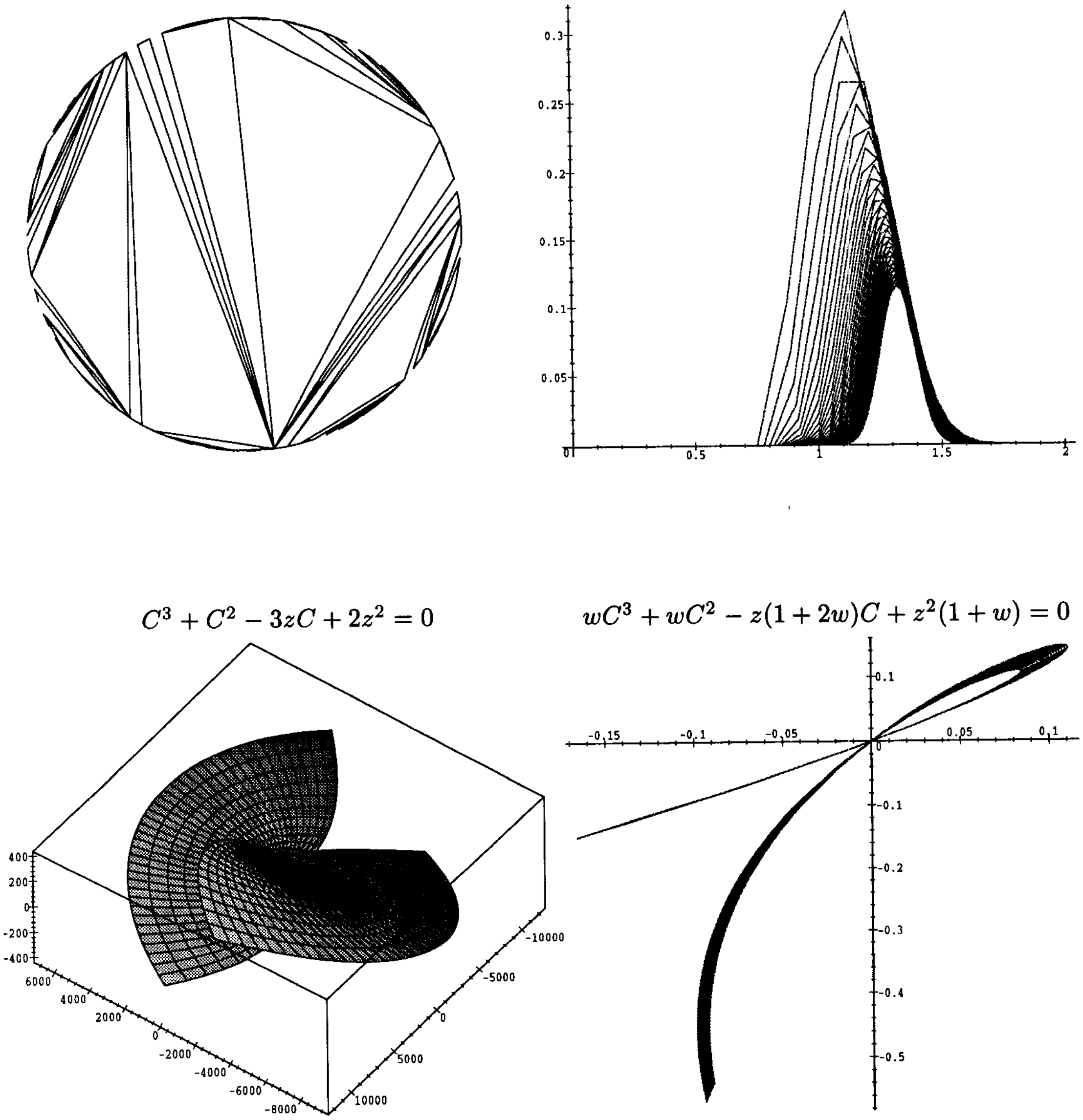

Figure 4: Non-crossing connected graphs (top, left: a random instance of size 100) have a combinatorial decomposition of the "cubic" type, reflected by GFs which satisfy a cubic equation. The counting generating function (bottom left: a 3-dimensional plot of $\Im C(z)$ for complex z) has an algebraic branch point of the square-root type that induces an asymptotic count of type $\omega^n/n^{3/2}$. The family of generating functions $\{C(z,w)\}_w$ where w records the number of edges (bottom right: plot of $C(z,w)$ for real z, when w varies in between 0.9 and 1.1) exhibit a common square-root singularity that moves analytically with w, a fact that induces a limit law of the Gaussian type for the number of edges (top right: histograms of the distribution for $n = 8\,..\,50$, with x-axis scaled to n).

[6] CANFIELD, E. R. Remarks on an asymptotic method in combinatorics. *Journal of Combinatorial Theory,* Series A *37* (1984), 348–352.

[7] COMTET, L. *Advanced Combinatorics.* Reidel, Dordrecht, 1974.

[8] DAVID, F. N., AND BARTON, D. E. *Combinatorial Chance.* Charles Griffin, London, 1962.

[9] DEVROYE, L., FLAJOLET, P., HURTADO, F., NOY, M., AND STEIGER, W. Random triangulations (extended abstract), Apr. 1996. Accepted for publication in *Discrete and Computational Geometry.*

[10] DOMB, C., AND BARRETT, A. Enumeration of ladder graphs. *Discrete Mathematics 9* (1974), 341–358.

[11] DRMOTA, M. Systems of functional equations. *Random Structures and Algorithms 10,* 1–2 (1997), 103–124.

[12] DULUCQ, S., AND PENAUD, J.-G. Cordes, arbres et permutations. *Discrete Mathematics 117* (1993), 89–105.

[13] FELLER, W. *An Introduction to Probability Theory and Its Applications,* vol. 2. John Wiley, 1971.

[14] FLAJOLET, P., HWANG, H.-K., AND SORIA, M. The ubiquitous Gaussian law in analytic combinatorics, 1997. In preparation.

[15] FLAJOLET, P., AND LAFFORGUE, T. Search costs in quadtrees and singularity perturbation asymptotics. *Discrete and Computational Geometry 12,* 4 (1994), 151–175.

[16] FLAJOLET, P., AND ODLYZKO, A. M. Singularity analysis of generating functions. *SIAM Journal on Discrete Mathematics 3,* 2 (1990), 216–240.

[17] FLAJOLET, P., AND SEDGEWICK, R. Analytic combinatorics. Book in preparation, 1998. (Individual chapters are available as INRIA Research Reports 1888, 2026, 2376, 2956, 3162.).

[18] FOATA, D. *La série génératrice exponentielle dans les problèmes d'énumération.* S.M.S. Montreal University Press, 1974.

[19] GEDDES, K. O., CZAPOR, S. R., AND LABAHN, G. *Algorithms for Computer Algebra.* Kluwer Academic Publishers, Boston, 1992.

[20] GOULD, H. W. Research bibliography on two number sequences. In *Mathematica Monongaliae,* 1971. (A comprehensive bibliography on Bell and Catalan numbers).

[21] GOULDEN, I. P., AND JACKSON, D. M. *Combinatorial Enumeration.* John Wiley, New York, 1983.

[22] HARARY, F., AND PALMER, E. M. *Graphical Enumeration.* Academic Press, 1973.

[23] HILLE, E. *Analytic Function Theory.* Blaisdell Publishing Company, Waltham, 1962. 2 Volumes.

[24] HWANG, H.-K. *Théorèmes limites pour les structures combinatoires et les fonctions arithmetiques.* PhD thesis, École Polytechnique, Dec. 1994.

[25] KLARNER, D. A., AND WOODWORTH, P. Asymptotics for coefficients of algebraic functions. *Aequationes Mathematicae 23* (1981), 236–241.

[26] KREWERAS, G. Sur les partitions non croisées d'un cycle. *Discrete Mathematics 1* (1972), 333–350.

[27] LISONEK, P. Closed forms for the number of polygon dissections. *Journal of Symbolic Computation 20* (1995), 595–601.

[28] MEIR, A., AND MOON, J. W. On the altitude of nodes in random trees. *Canadian Journal of Mathematics 30* (1978), 997–1015.

[29] NOY, M. Enumeration of non-crossing trees on a circle. *Discrete Mathematics 180* (1998), 301-313.

[30] PÓLYA, G., AND READ, R. C. *Combinatorial Enumeration of Groups, Graphs and Chemical Compounds.* Springer Verlag, New York, 1987.

[31] ROGERS, D. G. The enumeration of a family of ladder graphs. Part II: Schroeder and superconnective relations. *Quarterly Journal of Mathematics Oxford (2) 31* (1980), 491-506.

[32] SALVY, B., AND ZIMMERMANN, P. GFUN: a Maple package for the manipulation of generating and holonomic functions in one variable. *ACM Trans. Math. Softw. 20,* 2 (1994), 163–167.

[33] SEDGEWICK, R., AND FLAJOLET, P. *An Introduction to the Analysis of Algorithms.* Addison-Wesley Publishing Company, 1996.

[34] SLOANE, N. J. A., AND PLOUFFE, S. *The Encyclopedia of Integer Sequences.* Academic Press, 1995.

[35] STANLEY, R. P. Generating functions. In *Studies in Combinatorics,* M.A.A. Studies in Mathematics, Vol. 17. (1978), G.-C. Rota, Ed., The Mathematical Association of America, pp. 100–141.

[36] STANLEY, R. P. Hipparchus, Plutarch, Schröder and Hough. *American Mathematical Monthly 104* (1997), 344–350.

[37] TITCHMARSH, E. C. *The Theory of Functions*, second ed. Oxford University Press, 1939.

[38] WILF, H. S. *Generatingfunctionology.* Academic Press, 1990.