VARIETIES OF INCREASING TREES

François Bergeron
LACIM
Université du Québec à Montréal
Case Postale 8888, Succursale A
Montréal, Québec H3C3P8
Canada

Philippe Flajolet
Algorithms Project
INRIA Rocquencourt
78153 Le Chesnay
France

Bruno Salvy
Algorithms Project
INRIA Rocquencourt
78153 Le Chesnay
France

Abstract

An increasing tree is a labelled rooted tree in which labels along any branch from the root go in increasing order. Under various guises, such trees have surfaced as tree representations of permutations, as data structures in computer science, and as probabilistic models in diverse applications.

We present a unified generating function approach to the enumeration of parameters on such trees. The counting generating functions for several basic parameters are shown to be related to a simple ordinary differential equation,

$$\frac{d}{dz}Y(z) = \phi(Y(z)),$$

which is non linear and autonomous.

Singularity analysis applied to the intervening generating functions then permits to analyze asymptotically a number of parameters of the trees, like: root degree, number of leaves, path length, and level of nodes. In this way it is found that various models share common features: path length is $O(n \log n)$, the distribution of node levels and number of leaves are asymptotically normal, etc.

Introduction

A *labelled tree* of size n is a rooted tree comprising n nodes that are labelled by distinct integers of the set $\{1, \ldots, n\}$. An *increasing tree* is a labelled tree such that the sequence of labels along any branch starting at the root is increasing.

The enumeration of trees is a major branch of combinatorial analysis. A classical result due to Arthur Cayley in 1889 states that the number of labelled non-plane trees with n nodes is n^{n-1}. For plane trees, the corresponding count is $(n-1)!\binom{2n-2}{n-1}$ since the number of unlabelled plane trees is the Catalan number $\frac{1}{n}\binom{2n-2}{n-1}$ and there are $n!$ possible labellings given a fixed traversal order of trees, e.g., preorder. On this and other standard combinatorial analysis results, we refer the reader to the treatises of Comtet [6], Goulden and Jackson [19], or Bergeron *et al.* [3].

This paper concerns the enumeration of parameters on various families of increasing trees. The families to be considered are of two types: (i) *non-plane* trees, which are taken in the graph theoretic sense so that subtrees stemming from a node are not ordered between themselves; (ii) *plane trees*, where a plane embedding is specified so that subtrees stemming from a node are ordered between themselves. Rather arbitrary conditions can be imposed on the node degrees that are allowed.

Definition 1 Let $\{s_r\}_{r=0}^{\infty}$ be a sequence of non negative integers, such that $s_0 \neq 0$ and $s_r \neq 0$ for some $r \geq 2$. The *variety of trees* associated to $\{s_r\}$ and the specification of an element of

{Plane, Non-plane} is the collection of all increasing trees (plane or non plane depending on the specification) with s_r sorts of nodes of outdegree (arity) r for all r.

The *degree function* of a variety of trees associated with $\{s_r\}$ is defined as follows.

$$\text{In the plane case:} \qquad \phi(w) := \sum_{r\geq 0} s_r w^r.$$

$$\text{In the non-plane case:} \qquad \phi(w) := \sum_{r\geq 0} s_r \frac{w^r}{r!}.$$

The degree function condenses all the information needed for the analysis of tree parameters considered in this paper. We denote the coefficients of the degree function $\phi(w)$ by ϕ_r, so that $\phi(w) = \sum_{r=0}^{\infty} \phi_r w^r$, with $\phi_r = s_r$ in the plane case, while $\phi_r = s_r/r!$ in the non plane case. Whenever the collection of node types allowed is finite, $\phi(w)$ is a polynomial. In that case we call the variety a *polynomial variety* and let d denote the degree of $\phi(w)$; the integer d is then the maximum node degree allowed in the variety, and we call it the *degree* of the variety.

Our treatment is directed towards asymptotic estimates via generating functions (GF's). It aims at global results applicable across varieties of increasing trees. In that sense, it can be viewed as a transposition to increasing trees of a programme carried out by Meir and Moon who extensively studied so-called *simple families of trees*, see for instance [32]. However, due to the constraint of increasing labels, we are typically facing *algebraic differential equations* rather than plain algebraic equations. The key analytic method employed here is that of singularity analysis, developed by Flajolet and Odlyzko [11], though Darboux's method [6, p. 277] could have been employed instead in a few places at perhaps the expense of a little more work.

As a byproduct, the generating function approach sometimes provides explicit form for various tree counting problems, in the case of *exactly solvable models*. In this way we are able to unify several results that have appeared scattered in the literature.

For combinatorial counting purposes, we appeal to *exponential generating functions*: let $\{f_n\}_{n\geq 0}$ be a sequence of numbers; the exponential generating function (EGF) of the sequence is defined as

$$f(z) = \sum_{n\geq 0} f_n \frac{z^n}{n!}. \tag{1}$$

(Note that we use the same letters for a sequence and its EGF. We shall henceforth adhere to this convention, except for a few explicitly indicated situations where we have to resort to different types of generating functions.)

Our main results are as follows. Fix a variety of trees $\mathcal{Y}$, i.e., the degree function $\phi(w)$. Let Y_n be the number of trees of size n in the variety. The EGF of the variety of trees,

$$Y(z) = \sum_{n=1}^{\infty} Y_n \frac{z^n}{n!}, \tag{2}$$

is defined implicitly by

$$\int_0^{Y(z)} \frac{dw}{\phi(w)} = z. \tag{3}$$

The inversion problem is solvable in terms of special functions in a few particular cases of interest and we then call the corresponding models *solvable*.

Under fairly general conditions—most notably whenever $\phi(w)$ is a polynomial, which means a finite set of allowed degrees—the equation (3) can be analyzed near its dominant singularity. For polynomial varieties, this leads to an asymptotic counting result of the form

$$\frac{Y_n}{n!} \sim K \cdot \left(\frac{1}{\rho}\right)^n \cdot n^{-(d-2)/(d-1)} \qquad \text{with} \qquad \rho = \int_0^{\infty} \frac{dw}{\phi(w)}, \tag{4}$$

where $K = K_\phi$ is a constant that depends on ϕ alone, and d is the degree of the variety. The quantity ρ appears to be always a logarithmic form in algebraic numbers.

(Technically, this approach is analogous to the analytic method used by Meir and Moon in [33], where the authors enumerate recursive trees without unary nodes.)

Let $s[.]$ be a tree function that admits an inductive definition. Path length or the number of leaves in the tree are typical examples, and the precise meaning is explained below. Let $S(z)$ be the associated EGF,

$$S(z) = \sum_{t \in \mathcal{Y}} s[t] \frac{z^{|t|}}{|t|!},$$

where as usual $|t|$ represents the size of t. Then it is found that $S(z)$ is expressible as a sort of integral transform

$$S(z) = Y'(z) \int_0^z \left(\frac{d}{dt} F(t) \right) \cdot \frac{dt}{Y'(t)}. \tag{5}$$

The fact that this transform is determined by the variety $\mathcal{Y}$ under consideration is materialized by the occurrences of Y', while F reflects in a direct manner the inductive definition of the parameter $s[.]$ under consideration.

In solvable models where $Y(z)$ admits an explicit expression, the transform (5) is itself explicit. For instance, for binary increasing trees, and an inductive parameter

$$s[t] = f_{|t|} + s[t_{\text{left}}] + s[t_{\text{right}}],$$

it becomes

$$S(z) = \frac{1}{(1-z)^2} \left[f_0 + \int_0^z \left(\frac{d}{dt} F(t) \right) \cdot (1-t)^2 \, dt \right], \tag{6}$$

where $F(t)$ is in this particular case the ordinary generating function of the number sequence $\{f_n\}$, $F(t) = \sum_n f_n t^n$. In this way, for solvable models, we can get exact EGF's for parameters like path length, number of leaves, and so forth, on the trees.

However, in general, the solution is *asymptotic* rather than exact. It consists in viewing Eq. (5) as a "singularity transformer". We find, for finite families and for the other classical families, that path length is on average $\lambda n \log n$, that the expected number of leaves is asymptotic to αn, for some constants λ, α dependent upon ϕ.

A variation of this scheme in line with Bender's work [2] and with [12, 13] leads to limit distributions. For instance, the distribution of nodes in strata of a tree or the number of leaves both asymptotically conform to a Gaussian law.

Most existing works (with the notable exception of Meir and Moon's studies [32, 33]) appeal to special recurrence relation, often based on the insertion of a new node, and to the existence of closed–form solutions. In contrast, in this work, we resort to a combination of algebraic and analytic generating function methods which is versatile and widely applicable. In this way, it becomes possible to cast into a unifying framework a number of existing analyses concerning increasing trees and also to vastly extend the range of problems and models amenable to asymptotic analysis.

1 Classical Tree Models

We review here some of the models arising from diverse areas that constitute varieties of increasing trees.

Binary increasing trees. Any permutation σ of n elements can be written as a word σ whose ith letter is $\sigma(i)$; if $\min(\sigma)$ is the minimal element of σ, then, as a word, σ can be decomposed into $\sigma_{\text{left}} \cdot \min(\sigma) \cdot \sigma_{\text{right}}$, where σ_{left} (σ_{right}) designates the factors that appear to the left (resp. right) of the minimal element. One can construct a binary tree, $\mathbf{T}(\sigma)$, by repeated use of this decomposition:

> The root of the tree $\mathbf{T}(\sigma)$ is $\min(\sigma)$, with the left and right root subtrees being constructed recursively as $\mathbf{T}(\sigma_{\text{left}})$ and $\mathbf{T}(\sigma_{\text{right}})$; the tree associated to the empty permutation is the empty tree.

In this way, we associate bijectively to each permutation of $\{1, 2, \ldots, n\}$ a labelled tree with n labelled internal nodes and $n+1$ unlabelled external nodes.

Equivalently, by eliminating the unlabelled external nodes, we obtain a labelled unary–binary tree with two sorts of unary nodes, the left branching nodes and the right branching nodes. Thus, the binary increasing trees associated to permutations correspond to a plane family defined by $s_0 = 1, s_1 = 2, s_2 = 1$, so that $\phi(w) = 1 + 2w + w^2$. From the correspondence, there results in particular that the number of binary increasing trees of size n is $n!$.

This construction is recalled in Stanley's book [35, pp. 23–41] who attributes it "to the French". Here, we refer to Françon's work (see [15] and references therein) which is based in part on earlier methods developed by Foata and Schützenberger and in part on a pioneering paper [4] written by Burge in 1972.

Now, a number of classical permutation parameters have direct translations into basic tree parameters. For instance, the distribution of the number of nodes on the leftmost branch of the tree (i.e., the number of left–to–right minima in the permutation) is given by the Stirling numbers of the first kind; the distribution of the number of nodes with a left son in the tree (the descents in the permutation) is given by the Eulerian numbers, etc. Some of these classical results appear here as corollaries in Section 5

Binary increasing trees (under the alternative names of heap ordered trees or tournament trees) can also be used as a data structure to represent mergeable priority queues, with algorithms that can be precisely analyzed, see especially Burge's paper and Vuillemin's survey in [39], or [18, 38] for an overview. Finally, this model is of special importance since it is isomorphic to the analytic models of standard *binary search tree* and the *Quicksort algorithm*, see, e.g., [4, 9, 15, 21, 24, 29, 38, 39]. There, the model is equivalent to a splitting process in which n elements are split into a "root" and into two subgroups of cardinalities K and $n-1-K$, with the distribution of the random variable K being uniform over its range, $\Pr\{K = k\} = 1/n$ for all $k \in \{0, \ldots, n-1\}$.

Strict binary increasing trees. They correspond to binary increasing trees in which each node has either 0 or 2 sons, so that $\phi(w) = 1 + w^2$. By the standard correspondence with permutations described above, these trees give rise to *alternating permutations*, also called up-and-down, that are of the form $\sigma = \sigma_1\sigma_2\ldots\sigma_n$ with $\sigma_1 > \sigma_2 < \sigma_3 > \sigma_4 \cdots$. By implicitly using such a construction, Desiré André obtained in 1881 that the number of such trees over $2n+1$ nodes is equal the Taylor coefficient[1] $(2n+1)!\,[z^{2n+1}]\tan(z)$. (This number is known as a tangent number or Euler number.) The use of this principle in enumerating alternating permutations is detailed for instance in [19, p. 169].

Recursive trees. Meir and Moon [32] define recursive trees as the variety of non-plane increasing trees such that all node degrees are allowed. Thus, the degree function is $\phi(w) = \exp(w)$.

[1] As usual, $[z^n]f(z)$ denotes the coefficient of z^n in the expansion of $f(z)$ into powers of z, see [20, §5.4].

This model lies at the basis of Burge's sorting method of [4]. It was proposed as a statistical model in philology for problems related to the identification of terminal copies of manuscripts [34]; it has also been used for the legal assessment of chain letters and pyramids [17]. Its basic property for applications is that it can be generated by successive insertions of nodes, where at each stage each node is taken with equal likelihood as the father of the new node inserted.

All authors have noticed a close connection between recursive trees and permutations, the number of increasing trees of size n being $(n-1)!$. It is however a little less known (though it is already explicit in [4]) that this model is equivalent to that of binary increasing trees by the following remarks.

> *1).* Take a recursive tree of size n. *2).* Make it into a special plane increasing tree by ordering brother subtrees at each place in the tree from left to right according to increasing values of their roots. *3).* Chop off the root of the tree and apply the classical rotation correspondence [23, Sec. 2.3.2] that transforms a forest into a binary tree. The resulting tree is isomorphic to a binary increasing tree labelled on $\{2,\ldots,n\}$, itself isomorphic to a similar tree canonically labelled on $\{1,\ldots,n-1\}$.

From this correspondence and the observations above regarding binary increasing trees, one gets directly that the distribution of root degrees is given by Stirling numbers of the first kind, that the distribution of the number of leaves is Eulerian, and that path length is on average $\sim n\log n$. We rederive these and other properties in Section 5.

Plane recursive trees. This is a model introduced by Szymański [36] and further developed by Mahmoud *et al.* [30, 31]. In our terminologyit corresponds exactly to a variety of plane trees with degree set $\Omega = \{0,1,2,3,\ldots\}$, so that $\phi(w) = 1/(1-w)$. It also admits a construction by successive insertions, where at each stage each insertion slot is taken with equal likelihood. Alternatively, each node is selected as an insertion node with a probability proportional to its degree. As picturesquely described by Mahmoud, this is a propagation model in which "success brings success". (Footnote: Some of our results of Section 5 regarding plane recursive trees have also been independently obtained by Wen–Chin Chen and Wen–Chun Ni [5].)

2 Exact Enumeration of Varieties

We express here the basic counting problem for varieties of trees in terms of generating functions. This leads to a few cases of interest—the solvable models—where the generating functions are expressible in terms of standard functions. The definition of the generating function of a variety $\mathcal{Y}$ has been given in the introduction, see (2). The following result is a folk theorem. Fascinating combinatorial variations around it form the subject of a series of papers by Leroux and Viennot [26, 27] regarding the combinatorics of elementary calculus.

Theorem 1 *The exponential generating function $Y(z)$ of a variety of trees defined by the degree function ϕ is given implicitly by*

$$\int_0^{Y(z)} \frac{dw}{\phi(w)} = z. \tag{7}$$

Proof. This can be obtained from standard counting lemmas [19, 35, 40]. (Alternatively, the reader could return to underlying recurrences.) In terms of EGFs, forming a forest of k trees enumerated by $Y(z)$ corresponds to the EGF $Y^k(z)$ if the forest is ordered (plane case) and to $Y^k(z)/k!$ if it is unordered (non plane case). Appending a root with a minimal label to a forest

Name	Diff. eqn.	Explicit solution	ρ
Plane d–ary	$y' = (1+y)^d$	$y(z) = -1 + [1-(d-1)z]^{-1/(d-1)}$	$\frac{1}{d-1}$
Plane Strict d–ary	$y' = 1 + y^d$	$d=2 \quad y(z) = \tan z$ $d>2 \quad$ —	$\frac{\pi}{d}\frac{1}{\sin\frac{\pi}{d}}$
Non plane Strict d–ary	$y' = 1 + \frac{y^d}{d!}$	$d=2 \quad y(z) = \sqrt{2}\tan\frac{z}{\sqrt{2}}$ $d>2 \quad$ —	$\frac{\pi}{d}\frac{(d!)^{1/d}}{\sin\frac{\pi}{d}}$
Plane unary-binary	$y' = 1 + y + y^2$	$y(z) = \frac{\sqrt{3}}{2}\tan\left(\frac{\sqrt{3}}{2}z + \frac{\pi}{6}\right) - \frac{1}{2}$	$\frac{2\pi\sqrt{3}}{9}$
Non plane unary-binary	$y' = 1 + y + y^2/2$	$y(z) = \tan\left(\frac{z}{2} + \frac{\pi}{4}\right) - 1$	$\frac{\pi}{2}$
Plane "Recursive"	$y' = \frac{1}{1-y}$	$y(z) = 1 - \sqrt{1-2z}$	$\frac{1}{2}$
(Non plane) "Recursive"	$y' = \exp(y)$	$y = \log\frac{1}{1-z}$	1

Table 1: Some varieties of increasing trees and solvable models. For each type, we have listed the differential equation of the EGF, the explicit forms available, and the radius of convergence which dictates the exponential growth, ρ^{-n}, of the family.

enumerated by $W(z)$ corresponds to the EGF $\int_0^z W(t)\,dt$. Thus, with $\phi(w) = \sum_{r=0}^{\infty} \phi_r u^r$, we obtain

$$Y(z) = \int_0^z \left(\sum_{r=0}^{\infty} \phi_r Y^r(t)\right) dt.$$

From there, we derive

$$Y'(z) = \phi(Y(z)), \quad Y(0) = 0 \qquad \text{or} \qquad \frac{Y'(z)}{\phi(Y(z))} = 1, \tag{8}$$

and the result follows by integration. □

From Theorem 1, there results an explicit expression for the EGF of a variety of trees provided the integral $\int dw/\phi(w)$ is expressible in terms of special functions *and* its inverse is reducible to special functions. This is notably the case for the classical tree varieties described in the previous section.

Corollary 1 *(i) For binary trees, $\phi(w) = (1+w)^2$, we have*

$$Y(z) = \frac{z}{1-z}, \qquad Y_n = n\,!.$$

(ii) For strict binary trees, $\phi(w) = 1 + w^2$,

$$Y(z) = \tan(z), \qquad Y_{2n+1} = (2n+1)!\,[z^{2n+1}]\tan(z).$$

(iii) For recursive trees, $\phi(w) = \exp(w)$,

$$Y(z) = \log\frac{1}{1-z}, \qquad Y_n = (n-1)!\,.$$

(iv) For plane recursive trees, $\phi(w) = (1-w)^{-1}$,

$$Y(z) = 1 - \sqrt{1-2z}, \qquad Y_n = 1\cdot 3\cdot 5 \cdots (2n-3) = \frac{(n-1)!}{2^{n-1}}\binom{2n-2}{n-1}.$$

Proof. It only involves elementary integration.

$$\begin{array}{rclrcl} \displaystyle\int_0^y \frac{dw}{(1+w)^2} &=& \displaystyle\frac{y}{1+y}, & \displaystyle\int_0^y \frac{dw}{1+w^2} &=& \arctan(y), \\ \displaystyle\int_0^y \frac{dw}{\exp(w)} &=& 1-e^{-y}, & \displaystyle\int_0^y \frac{dw}{(1-w)^{-1}} &=& y - \displaystyle\frac{y^2}{2}, \end{array}$$

and all the inverse functions are explicitly computable. □

The results of Corollary 1 are each well known under one guise or another. See, *e.g.*, [3, 19, 32, 36, 39].

A collection of explicit results is summarized in Table 1. An interesting solvable model whose asymptotic behaviour is characteristic of polynomial families is given below.

EXAMPLE 1. *d–ary trees.* The class of d–ary increasing trees corresponds to $\phi(w) = (1+w)^d$. (It can be viewed as d-ary trees in which only internal nodes are labelled.) We have

$$\int_0^y \frac{dw}{(1+w)^d} = \frac{1}{d-1}\left[1 - \frac{1}{(1+y)^{d-1}}\right] \qquad \text{and} \qquad Y(z) = -1 + [1-(d-1)z]^{-1/(d-1)}, \quad (9)$$

so that

$$Y_n = \delta^{1-n}(1+\delta)\cdot(2+\delta)\cdots(n-1+\delta) \qquad \text{where} \qquad \delta = \frac{1}{d-1}, \qquad \text{whence}$$

$$\frac{Y_n}{n!} \sim \frac{1}{\Gamma(\delta)}\delta^{-n} n^{-1+\delta}.$$

□

EXAMPLE 2. *Special ternary trees.* This is a rather artificial example just meant to demonstrate that explicit EGF's may arise in unexpected contexts. Take $\phi(w) = w^3 + 6w^2 + 11w + 6$, which corresponds to a variety of plane ternary trees with six sorts of leaves, eleven sorts of unary nodes, six sorts of binary nodes, and one sort of ternary node. We find

$$Y(z) = -2 + \frac{2}{\sqrt{4-3e^{2z}}} = 6\,z + 66\,\frac{z^2}{2!} + 1158\,\frac{z^3}{3!} + 28290\,\frac{z^4}{4!} + 887046\,\frac{z^5}{5!} + \&c.$$

□

3 Asymptotic Enumeration of Polynomial Varieties

In general, no exact form is available and we have to resort to asymptotic analysis. Our discussion at this point focuses on *polynomial varieties.*

A preliminary observation regarding periodicity phenomena is in order here. If $\phi(w)$ is a function of w^p for some $p \geq 2$, so that $\phi(w) = \psi(w^p)$ for some power series ψ;, we say that $\phi(w)$ is *periodic* and the maximum possible p is called its *period.* Otherwise, $\phi(w)$ is said to be aperiodic (and we take $p = 1$). For strict binary trees the period is $p = 2$, but for unary-binary trees $p = 1$. For period $p \geq 2$, we have $Y(z) = zY^*(z^p)$, for some power series Y^*; accordingly, the non–zero coefficients Y_n are those whose index satisfies the congruence condition $n \equiv 1 \pmod{p}$. In subsequent statements, this restriction $n \equiv 1 \pmod{p}$ is implicitly assumed in all periodic cases.

Theorem 2 *Let $\mathcal{Y}$ be a polynomial variety associated with the degree function $\phi(w) = \phi_d w^d + \cdots + \phi_0$. The number Y_n of elements of $\mathcal{Y}$ with size n satisfies[2] for $d > 3$*

$$\frac{Y_n}{n!} = \frac{p}{\eta\Gamma(\delta)}\left(\frac{1}{\rho}\right)^n n^{-1+\delta} + \frac{ph_2\eta}{\Gamma(-\delta)}\rho^{-n}n^{-1-\delta} + O(n^{-1-2\delta}). \tag{10}$$

where

$$\rho = \int_0^\infty \frac{dw}{\phi(w)}, \quad \delta = \frac{1}{d-1}, \quad \eta = \left(\frac{\phi_d \rho}{\delta}\right)^\delta,$$

p is the period of ϕ, and h_2 is as defined in Lemma 2 below. In particular, if ϕ has distinct roots, then

$$\rho = -\sum_{j=1}^{d} \frac{1}{\phi'(\zeta_j)} \log(-\zeta_j),$$

where the ζ_j are the roots of ϕ and the principal determination of the logarithm is taken.

Proof. We first determine the radius of convergence ρ of $Y(z)$; next we analyze $Y(z)$ in the vicinity of $z = \rho$; finally we translate the behaviour of $Y(z)$ into asymptotics of the coefficients Y_n by means of the method of singularity analysis, see [11].

Recall that a *dominant singularity* is one of smallest modulus. By Pringsheim's theorem [37, §7.21] we know that one of the dominant singularities is real positive, and this singularity is equal to the radius of convergence of $Y(z)$.

Lemma 1 *Given a degree function $\phi(w)$ that is polynomial or entire, the dominant real positive singularity of the function $Y(z)$, solution to $Y' = \phi(Y)$ and $Y(0) = 0$, is*

$$\rho = \int_0^\infty \frac{dy}{\phi(y)}.$$

Furthermore, if ϕ is non periodic, then ρ is the only dominant singularity of $Y(z)$. If ϕ has period $p \geq 2$, then $Y(z) = zY^*(z^p)$, where Y^* has a unique dominant singularity at $\rho^{1/p}$.

Proof. The integral is clearly defined since $\phi(w)$ does not vanish on the positive real axis and increases at least like w^2 at infinity. For any y with $0 < y < +\infty$, the integral $\int_0^y dw/\phi(w)$ is an analytic function of y with a non zero derivative; it is therefore invertible. Thus we find that $Y(z)$ is analytic at least for all real z, with $0 < z < \rho$. Clearly $Y(z)$ becomes infinite as $z \to \rho^-$, so that ρ is an actual singular point of $Y(z)$.

Let $z_0 = r_0 \exp(i\theta)$, with $r_0 < \rho$. Since Y has positive Taylor coefficients, we have $|Y(z_0)| \leq Y(r_0)$ by the triangular inequality. By a well known lemma, equality $|Y(z_0)| = Y(r_0)$ is possible for $\theta \neq 0$ only if $Y(z) = z^a Y^*(z^p)$ for some integers a, p with $p \geq 2$, in which case $\theta = 2m\pi/p$. (This in turn implies that $\phi(w)$ is periodic.)

Assume first that we are in the non periodic case. Thus $|Y(z_0)| < Y(r_0)$. Let r_1 be a positive real such that $|Y(z_0)| = Y(r_1)$; we have $r_1 < r_0$ by growth of $Y(z)$ on the positive real axis. Consider the function ψ solution to $\psi'(z) = \phi(\psi(z))$, with $\psi(r_0) = Y(r_1)$. Then ψ and Y are related by $\psi(z) = Y(z - r_0 + r_1)$ since the system is autonomous (i.e., there is no explicit dependency of z). From the positivity of the problem (i.e., ϕ), we then have $|Y(z)| \leq \psi(|z|)$. Thus the modulus of the solution $Y(z)$ is upper bounded by $Y(|z| - r_0 + r_1)$. In plain words a delay in the growth of $Y(z)$ along a non real ray "propagates" along that ray. In particular, $Y(z)$ exists along the ray of angle θ for $|z| < \rho - r_0 + r_1$, and it is analytic there.

[2]For $d = 2$, the term containing $n^{-1-\delta}$ disappears and the error after the first term is exponentially small; for $d = 3$, the form of h_2 given in Eq. (10) has to be modified.

In the periodic case, a suitable amended form of the argument applies, with the exceptional angles being simply the multiples of π/p. □

Lemma 2 *Let $\phi(w) = \phi_0 + \cdots + \phi_d w^d$ be a polynomial degree function with degree $d \geq 2$. Then, in a complex neighbourhood of ρ, the solution $Y(z)$ of (7) is of the form*

$$Y(z) = \frac{1}{\Delta(z)} \cdot H(\Delta(z)) \qquad \textit{where} \qquad \Delta(z) = \eta(1 - z/\rho)^{\delta}, \tag{11}$$

and $H(w) = \sum_{m=0}^{\infty} h_m w^m$ is analytic at $w = 0$,

$$h_0 = 1, \;\; h_1 = -\frac{\phi_{d-1}}{d\phi_d}, \;\; h_2 = -\frac{2d\phi_d\phi_{d-2} - (d-1)\phi_{d-1}^2}{2d(d+1)\phi_d^2}.$$

Proof. We start with the expansion of $1/\phi$ as $w \to +\infty$,

$$\frac{1}{\phi(w)} = \frac{1}{\phi_d w^d} - \frac{\phi_{d-1}}{\phi_d^2 w^{d+1}} + \frac{\phi_{d-1}^2 - \phi_d\phi_{d-2}}{\phi_d^3 w^{d+2}} + \cdots.$$

By integration, we find as $Y \to +\infty$,

$$\int_Y^{\infty} \frac{dw}{\phi(w)} = \frac{\delta}{\phi_d} Y^{-d+1} - \frac{\phi_{d-1}}{d\phi_d^2} Y^{-d} + \cdots. \tag{12}$$

But from

$$\int_0^{Y(z)} \frac{dw}{\phi(w)} = z \qquad \text{and} \qquad \int_0^{\infty} \frac{dw}{\phi(w)} = \rho,$$

we have

$$\rho - z = \int_{Y(z)}^{\infty} \frac{dw}{\phi(w)}. \tag{13}$$

Therefore, by comparing (12) and (13), we get

$$\frac{\delta}{\phi_d} Y^{1-d} - \cdots = (\rho - z).$$

Inverting this relation, leads to the singular expansion of $Y(z)$ as a function of $(\rho - z)^{1/(d-1)}$.

This process is purely formal. The analytic character of the resulting series $H(w)$ is easily established by means of the method of majorizing series or by observing that $1/\phi(w)$ and derived functions are analytic at ∞. □

We return to the proof of Theorem 2. The bare principle of singularity analysis consists in applying the following rules,

$$\begin{aligned} Y(z) = (1 - z/\rho)^{\alpha} \qquad &\Longrightarrow \qquad [z^n]Y(z) \sim \rho^{-n} \frac{n^{-\alpha-1}}{\Gamma(-\alpha)} (1 + O(\frac{1}{n})) \\ Y(z) = O((1 - z/\rho)^{\alpha}) \qquad &\Longrightarrow \qquad [z^n]Y(z) = O(\rho^{-n} n^{-\alpha-1}), \end{aligned}$$

inside the local asymptotic expansion of a function which is singular at $z = \rho$. These two rules are applied to the expansion (11) of $Y(z)$.

Validity of the process in the aperiodic case is ensured because $Y(z)$ exists in a domain larger than the disk of convergence as guaranteed by the two lemmas (see [11]). In the periodic case, contributions from the p dominant singularities must be added, which accounts for the extra factor of p.

This concludes the proof of Theorem 2. □

EXAMPLE 3. *Strict ternary trees.* For the variety of strict ternary trees, we have $\phi(w) = 1 + w^3$. Thus, $Y(z)$ satisfies

$$\frac{1}{3}\log(1+Y) - \frac{1}{6}\log(1-Y+Y^2) + \frac{1}{\sqrt{3}}\arctan(\frac{2Y-1}{\sqrt{3}}) + \frac{\pi}{6\sqrt{3}} = z.$$

This relation is *not* explicitly solvable for $Y(z)$. However, the singularity is easily determined to be $\rho = 2\pi\sqrt{3}/9$, and

$$Y(z) = \frac{1}{\sqrt{2}\sqrt{\frac{2\pi\sqrt{3}}{9} - z}} + h_2(1 - z/\rho) + O((1 - z/\rho)^{5/2})$$

so that

$$\frac{Y_n}{n!} = \frac{1}{\sqrt{2\pi}}\left(\frac{2\pi\sqrt{3}}{9}\right)^{-n-1/2} n^{-1/2}[1 - \frac{1}{8n} + \frac{1}{128n^2} + O(n^{-3})].$$

Full expansions are readily obtained. □

4 Inductive Parameters

We show here how a fairly general theory of cost measures on varieties of trees can be developed, based on the algebraic and analytic methods used for basic tree enumerations. In order not to obscure the line of reasoning too much we have limited the discussion to "inductive maps" defined below, and to a particular "elementary" subclass with interesting analytic properties.

4.1 Exact Generating Functions.

Definition 2 A function from trees to complex numbers is called an *inductive map* if it is definable by a relation

$$s[t] = f_{|t|} + \sum_{\tau \propto t} s[\tau],$$

for some number sequence $\{f_n\}$, where the sum is over all root subtrees τ of t (noted $\tau \propto t$).

Given a tree function $s[.]$ and a variety $\mathcal{Y}$, the generating function of $s[.]$ (over $\mathcal{Y}$) is

$$S(z) = \sum_{t \in \mathcal{Y}} s[t]\frac{z^{|t|}}{|t|!}.$$

Typical examples of inductive maps are tree size ($f_n = 1$), number of leaves ($f_n = \delta_{n,1}$) and path length ($f_n = n$). The GF $S(z)$ is an EGF of cumulated values since $n![z^n]S(z) = \sum_{|t|=n} s[t]$.

Theorem 3 *Let $s[t]$ be an inductive map,*

$$s[t] = f_{|t|} + \sum_{\tau \propto t} s[\tau].$$

The generating function of $s[t]$, on a variety $\mathcal{Y}$, is computable by

$$S(z) = Y'(z)\int_0^z \left(\frac{d}{dt}F(t)\right)\frac{dt}{Y'(t)}, \tag{14}$$

where $F(z)$ is defined from $\{f_n\}$ and $\mathcal{Y}$ by

$$F(z) = \sum_{n \geq 0} f_n Y_n \frac{z^n}{n!}. \tag{15}$$

Proof. The bivariate generating function

$$Y(z,u) = \sum_{t\in\mathcal{Y}} u^{s[t]} z^{|t|}$$

is easily seen to satisfy

$$\begin{aligned} Y(z,u) &= \textstyle\sum_n Y_n u^{f_n} \frac{z^n}{n!} + \int_0^z \sum_{r=1}^{\infty} \phi_r Y^r(t,u)\,dt \\ &= \textstyle\sum_n Y_n u^{f_n} \frac{z^n}{n!} + \int_0^z \phi(Y(t,u))\,dt. \end{aligned} \tag{16}$$

The argument is of the same type as that used for equation (7): the contribution to $Y(z,u)$ comes either from the root (the first sum), or—when the root has degree r—from one of the r root subtrees with multiplicity ϕ_r.

From equation (16) and the fact that $Y(z,1) = Y(z)$ and $\partial Y/\partial u(z,1) = S(z)$, we obtain by differentiation

$$S(z) = F(z) + \int_0^z S(t)\phi'(Y(t))\,dt. \tag{17}$$

Eq. (17) next translates into a linear differential equation

$$S'(z) = F'(z) + S(z)\phi'(Y(z)), \quad S(0) = 0.$$

First solve the associated homogeneous equation, which yields $S(z) = CY'(z)$. Then solve the inhomogeneous equation by the variation-of-parameter method, which gives

$$S(z) = Y'(z)\int_0^z \frac{F'(t)}{Y'(t)}\,dt. \tag{18}$$

□

4.2 Asymptotic Estimates.

Definition 3 An inductive map $s[.]$ is called *elementary* when the associated number sequence $\{f_n\}$ is of the form

$$f_n = Cn^\alpha \log^r n,$$

for some real C, α and non-negative r. The triple (C, α, r) is called the *parameter* of the map.

Theorem 4 *Let s be an elementary map with parameter (C,α,r) defined on a polynomial variety $\mathcal{Y}$. Then the average value $\overline{S_n}$ of s on the elements of $\mathcal{Y}$ with size n satisfies asymptotically:*

1. *if $\alpha < 1$, then*

$$\overline{S_n} \sim \lambda\, n \qquad \textit{with} \qquad \lambda = \frac{1}{\rho}\int_0^\rho \frac{F'(t)}{Y'(t)}\,dt;$$

2. *if $\alpha = 1$, then*

$$\overline{S_n} \sim \lambda n \log^{r+1} n \qquad \textit{with} \qquad \lambda = \frac{C(\delta+1)}{r+1};$$

3. *if $\alpha > 1$, then*

$$\overline{S_n} \sim \lambda n^\alpha \log^r n \qquad \textit{with} \qquad \lambda = \frac{C}{\alpha - 1}(\alpha + \delta).$$

Proof. The result follows from singularity analysis and from the analysis of $Y(z)$ already effected in the proof of Theorem 2. We assume without loss of generality that $\mathcal{Y}$ is non periodic. The major steps are as follows.

A. It is *a priori* clear from growth conditions, namely $S_n = O(Y_n n^{\alpha+2}) = O(\rho^{-n} n^{\alpha+1/2})$, that $S(z)$ has $z = \rho$ as a dominant singularity. We must therefore analyse $S(z)$ locally there.

B. The modified generating function $F(z) = \sum f_n Y_n z^n$ is a Hadamard product (a "termwise" product) of $Y(z)$ with the ordinary generating function $\mathcal{F} = \sum_{n=1}^{\infty} f_n z^n$. With the standard notation of Hadamard products ($\odot$), we have

$$F(z) = \mathcal{F}(z) \odot Y(z).$$

The analysis of $F(z)$ associated with $f_n = Cn^{\alpha} \log^r n$ is then effected by the following steps.

> $\mathcal{F}(z)$ is analytically continuable and has the expected asymptotic expansion—it has an algebraic–logarithmic singularity—when $z \to 1$ in the complex plane. This results from classical complex integral representations, see for instance Ford's book [14] or Evgrafov [8, p. 165].
>
> Hadamard products preserve analytic continuation. (This is Hadamard's celebrated theorem on composition of singularities by Hadamard products.)
>
> Hadamard products preserve algebraic–logarithmic singularities. This theorem is due to Pólya, see [41] for a discussion relevant to our goals.

C. Once the singular behaviour of F is known, the corresponding analysis for S is done mechanically using rules for differentiation, integration, and usual Cauchy products. The necessary expansions are summarized by the following lemma.

Lemma 3 *If s has parameter (C, α, r) then the singular expansion of $S(z)$ near ρ is given by*

$$S(z) \sim \lambda(1 - z/\rho)^{-1-\delta}\psi(z), \quad z \to \rho,$$

where λ is a constant and

$$\psi(z) = \begin{cases} 1 & \text{if } \alpha < 1, \\ \log^{r+1} \frac{1}{1-z/\rho} & \text{if } \alpha = 1, \\ (1 - z/\rho)^{1-\alpha} \log^r \frac{1}{1-z/\rho} & \text{if } \alpha > 1. \end{cases}$$

Proof. First, $\mathcal{F}$ has an isolated singularity at $z = 1$ as follows from integral representations [14, 8]. Next, F has an algebraic–logarithmic singularity at $z = \rho$: by Theorem 2, its coefficients grow as

$$\frac{C}{\eta\Gamma(\delta)} \rho^{-n} n^{\alpha-1+\delta} \log^r n,$$

and by the closure theorems discussed by Wilson [41], $F(z)$ has the corresponding singularity. Then,

$$F'(t) \sim \lambda_0 (1 - t/\rho)^{-\alpha-1-\delta} \log^r \frac{1}{1 - t/\rho} \quad (t \to \rho) \text{ with } \lambda_0 = \frac{C}{\eta\rho} \frac{\Gamma(\alpha + \delta + 1)}{\Gamma(\delta)}.$$

Thus, the integrand in (14) grows like

$$\lambda_0 \frac{\eta\rho}{\delta} (1 - t/\rho)^{-\alpha} \log^r \frac{1}{1 - t/\rho}, \quad t \to \rho,$$

and several cases need to be distinguished.

1. If $\alpha < 1$, then the integral is convergent and we get the statement of the lemma with
$$\lambda = \frac{\delta}{\eta\rho}\int_0^\rho \left(\frac{d}{dt}F(t)\right)\frac{dt}{Y'(t)};$$
2. if $\alpha = 1$, then integrating by parts, we get the statement with $\lambda = \lambda_0\rho/(r+1)$;
3. if $\alpha > 1$, then again integrating by parts, we get the statement with $\lambda = \lambda_0\rho/(\alpha - 1)$.

This completes the evaluation of $S(z)$ near ρ and the lemma is established. □

From the last lemma and singularity analysis applied to $S(z)$, the proof of Theorem 4 is in turn completed. □

Theorem 4 admits a number of extensions. For instance, the same methodology accommodates exponentially decaying f_n like $\theta^n n^\alpha$, with $\theta < 1$, boundary condition terms like $\delta_{n,0}$ or $\delta_{n,1}$, and any finite linear combinations. In all these cases, and more generally whenever the generating function $\mathcal{F}(z)$ is analytic in a disk that properly contains the unit disk, the formula of case 1 of Theorem 4 applies. Quantities f_n that are asymptotically equivalent (rather than equal) to $n^\alpha \log^r n$ are also amenable to these methods.

5 Characteristics of Increasing Trees

The estimates presented here illustrate direct consequences of the theorems and methods introduced in the last section, when specialized to a few classical tree parameters. This concerns statistics on path length, mean number of leaves, and root degree.

In addition, extensions of the method lead to asymptotic probability distributions for levels and number of leaves, which are proved to be Gaussian in the limit.

5.1 Path length

Path length of a tree t is by definition the sum of the distances of all nodes in t to the root of t, distances being measured by number of nodes[3] on the connecting branch. An alternative inductive definition is thus
$$s[t] = |t| + \sum_{\tau \propto t} s[\tau].$$
Path length is directly amenable to techniques of the last section, and in this case, function F introduced in Theorem 3 takes the explicit form $F(z) = zY'(z)$.

Theorem 5 *Let the variety of trees $\mathcal{Y}$ be defined by the degree function ϕ. The generating function of path length $S(z)$ is given by*
$$S(z) = Y'(z)\int_0^z \frac{Y'(t) + tY''(t)}{Y'(t)}\,dt.$$

Asymptotically, for a polynomial variety of degree d, we have
$$\overline{S_n} = \frac{S_n}{Y_n} = (\delta+1)n\log n + Cn + O\left(\frac{\log n}{n^{\min(0,2\delta-1)}}\right),$$
with
$$C = -1 - \psi(\delta) + \frac{1}{\rho}\int_0^\rho \left[1 + t\frac{Y''(t)}{Y'(t)} - \frac{\delta+1}{1-t/\rho}\right]dt,$$
where ψ is the logarithmic derivative of Euler's Γ function.

[3] Path length is sometimes defined by measuring the distance to the root as the number of connecting edges. This variant of path length, $s^*[t]$ satisfies $s^*[t] = s[t] - |t|$, so that $\overline{S_n^*} = \overline{S_n} - n$.

Proof. A direct application of Theorem 4 to the case $f_n = n$, plus some extra care for the subdominant terms. □

Corollary 2 *(i) For binary trees, $\phi(w) = (1+w)^2$, we have*

$$\overline{S_n} = 2n\log n + (2\gamma - 3)n + 2\log n + 2\gamma + 1 + O(1/n).$$

(ii) For strict binary trees, $\phi(w) = 1 + w^2$,

$$\overline{S_n} = 2n\log n + (2\gamma + 3 + \frac{\pi^2}{36} + \frac{\pi^4}{7200})n + 2\log n + (2\gamma + 1 - \frac{\pi^2}{36} - \frac{\pi^4}{7200}) + O(1/n).$$

(iii) For recursive trees, $\phi(w) = \exp(w)$,

$$\overline{S_n} = n\log n + \gamma n + \frac{1}{2} + O(1/n).$$

(iv) For plane recursive trees, $\phi(w) = (1-w)^{-1}$,

$$Y_n = \frac{1}{2}n\log n + (\log 2 + \frac{\gamma+1}{2})n - \frac{1}{4}\log n - \frac{2\log 2 + \gamma}{4} + O(1/n).$$

Proof. The GF $S(z)$ is determined by integration, and we find the four expressions:

$$\frac{2\log(1-z)^{-1} - z}{(1-z)^2}, \quad \frac{1}{\cos^2 z}\int_0^z (1 + 2t\tan t)\,dt, \quad \frac{\log(1-z)^{-1}}{(1-z)}, \quad \frac{2z + \log(1-2z)^{-1}}{4(1-2z)^{1/2}}.$$

This corresponds to more or less complicated explicit forms, e.g. $\overline{S_n} = nH_n$ for recursive trees and the asymptotic forms follow by singularity analysis. □

Variance. From equation (16), we can compute the variance of a general inductive map by

$$\sigma^2 = \frac{[z^n]\frac{\partial^2}{\partial u^2}Y(z,1)}{Y_n} + \overline{S_n} - \overline{S_n}^2.$$

The second derivative is obtained in a way similar to the first one,

$$\frac{\partial^2}{\partial u^2}Y(z,1) = Y'(z)\int_0^z \frac{H'(t) + \phi''(Y(t))S^2(t)}{Y'(t)}dt,$$

where

$$H(z) = \sum_n Y_n f_n(f_n - 1)\frac{z^n}{n!} = Y(z) \odot (\mathcal{F}(z) \odot \mathcal{F}(z) - \mathcal{F}(z)).$$

From this equation, using the same arguments as in the proof of Theorem 4, it is possible to attain a classification of the possible variances in the polynomial case, and a similar but complicated process would yield moments of higher order.

In the particular case of path length, we have $f_n = n$, and

$$H(z) = \sum_{n\geq 0} n(n-1)Y_n\frac{z^n}{n!} = z^2 Y''(z),$$

so that $\frac{\partial^2 Y}{\partial u^2}(z,1)$ reduces to

$$Y'(z)\int_0^z \frac{2tY''(t) + t^2Y'''(t) + \phi''(Y(t))S^2(t)}{Y'(t)}dt.$$

This has a singularity at ρ, and in the polynomial case, the local behaviour yields the following.

Corollary 3 *The variance of path length on a polynomial variety $\mathcal{Y}$ is asymptotic to λn^2, with*

$$\lambda = \psi(\delta+1)(\delta+1)(2\delta+4-\frac{2C}{\rho}) - \psi'(\delta+1)(\delta+1)^2 + 2(\delta+1)(\delta+2) - \frac{4C}{\rho} + \frac{\delta+2}{\rho^2},$$

and C as in Theorem 5.

This result is well known in the case of quicksort and path length in binary search trees for which $\lambda = 7 - 2\frac{\pi^2}{3}$.

5.2 Node Sorts

The next example shows a parameter that does not strictly speaking falls in the general category treated by Theorem 4, though exactly the same method of proof applies.

Theorem 6 *Let $s[t]$ be the number of nodes of outdegree i in a random tree t of size n from a variety defined by $\phi(w)$. With $\phi_i = [w^i]\phi(w)$, the generating function of $s[.]$ is*

$$S(z) = \phi_i Y'(z) \int_0^z \frac{Y^i(t)}{Y'(t)}\,dt. \tag{19}$$

For a polynomial family, the mean number of i–nodes is

$$\overline{S_n} = \lambda_i n[1 + O(n^{-\delta})] \qquad \textit{where} \qquad \lambda_i = \frac{\phi_i}{\rho} \int_0^\rho \frac{Y^i(t)}{Y'(t)}\,dt.$$

Proof. The parameter $s[t]$ which represents the number of nodes of degree i satisfies

$$s[t] = \mathbf{1}_{\text{degree}(\text{root}(t))=i} + \sum_{u \propto t} s[u].$$

Hence, the integral equation for the EGF $S(z)$:

$$S(z) = \phi_i \int_0^z Y^i(t)\,dt + \int_0^z S(t)\phi'(Y(t))\,dt,$$

whose solution is

$$S(z) = \phi_i Y'(z) \int_0^z \frac{Y^i(t)}{Y'(t)}\,dt.$$

The dominant singularities of S are therefore those of Y, and the singular behaviour at ρ is

$$\frac{\phi_i \delta}{\eta\rho}(1 - z/\rho)^{-1-\delta} \int_0^\rho \frac{Y^i(t)}{Y'(t)}dt,$$

the integral being convergent. From there we get the coefficients via singularity analysis. □

Corollary 4 *(i) For binary trees, $\phi(w) = (1+w)^2$, the expected numbers of nodes of degrees $0, 1, 2$ are equal to*

$$\frac{n+1}{3}, \frac{n+1}{3}, \frac{n-2}{3}.$$

(ii) For recursive trees, $\phi(w) = \exp(w)$, the expected number of nodes of outdegree i is asymptotic to

$$\lambda_i n \qquad \textit{where} \qquad \lambda_i = \frac{1}{2^{i+1}}.$$

(iii) For plane recursive trees, $\phi(w) = (1-w)^{-1}$, the expected number of nodes of outdegree i is asymptotic to

$$\lambda_i n \qquad \textit{where} \qquad \lambda_i = \frac{2}{(i+1)(i+2)(i+3)}.$$

Proof. For binary trees, the result follows by elementary integration. (For strict binary trees, the problem is degenerate!) For the other two cases, one estimates simultaneously all the λ_i by means of their ordinary generating function, $\Lambda(x) = \sum_i \lambda_i x^i$.

For recursive trees, we find

$$\begin{aligned}\Lambda(x) &= \sum_{i=0}^{\infty} \frac{x^i}{i!} \int_0^1 (1-t) \log^i \frac{1}{1-t}\, dt \\ &= \int_0^1 \frac{(1-t)}{(1-t)^x}\, dt = \left[\frac{T^{2-x}}{2-x}\right]_0^1 = \frac{1}{2-x}.\end{aligned}$$

For plane recursive trees, we find

$$\begin{aligned}\Lambda(x) &= \int_0^{1/2} \frac{\sqrt{1-2t}}{1-x(1-\sqrt{1-2t})}\, dt \\ &= \frac{(1-x)^2}{x^3} \log \frac{1}{1-x} + \frac{3x-2}{2x^2}.\end{aligned}$$

□

The result for binary trees is of course classical. For recursive trees, the result appears in [1, Thm 3.3] in the combinatorial literature but seems to have been first noted by Gastwirth [16].

5.3 Root degree

Let $r[t]$ denote the root degree of tree t. Then the EGF of trees with root degree j is clearly

$$\phi_j \int_0^z Y^j(t)\, dt. \tag{20}$$

The bivariate generating function for the distribution of root degree in trees in a variety $\mathcal{Y}$ is defined by

$$R(z,u) = \sum_{t \in \mathcal{Y}} u^{r[t]} \frac{z^{|t|}}{|t|\,!}.$$

Theorem 7 *The bivariate generating function of root degrees, $R(z,u)$, is given by*

$$R(z,u) = \int_0^z \phi(uY(t))\, dt.$$

Let $S(z) = R_u(z,1)$ and $T(z) = R_{uu}(z,1)$ be the generating functions of the first and second factorial moments of root degree. Then

$$\begin{aligned}\tfrac{d}{dz} S(z) &= Y(z) \tfrac{d}{dz} \log Y'(z) \\ \tfrac{d}{dz} T(z) &= \tfrac{Y^2(z)}{Y'(z)} \tfrac{d^2}{dz^2} \log Y'(z)\end{aligned} \tag{21}$$

For a polynomial variety of degree d, almost all trees of size n have root degree equal to d. The probability π_{nj} that a tree of size n has root degree j, with $1 \le j \le d$, is

$$\pi_{nj} \sim \alpha_j n^{-1+(j-1)\delta} \qquad \textit{with} \qquad \alpha_j = \frac{\rho \phi_j}{\eta^{j-1}} \frac{\Gamma(\delta)}{\Gamma(j\delta)}.$$

Proof. The bivariate generating function follows from (20). For computational purposes, it is advantageous to deal with $\frac{\partial}{\partial z}R(z,u)$ from which the forms of S and T follow.

The case of polynomial varieties results immediately from singularity analysis. □

Corollary 5 *(i) For recursive trees, $\phi(w) = \exp(w)$, the number of trees with size n and root degree k is $s_{n-1,k}$. The mean root degree is H_{n-1}.*

(ii) For plane recursive trees, $\phi(w) = (1-w)^{-1}$, the number of trees with size n and root degree k is

$$\frac{(2n-3-k)!}{2^{n-1-k}(n-1-k)!}.$$

The mean root degree is $\sqrt{\pi n} + O(1)$.

Proof. The theorem provides the generating functions

$$\frac{(1-z)^{1-u}-1}{u-1} \quad \text{and} \quad \frac{1-\sqrt{1-2z}}{u} + \frac{1-u}{u^2}\log(1-u+u\sqrt{1-2z}),$$

from which the results are deduced. □

The result for recursive trees is classical, at least under equivalent formulations (see the section on models). The mean for plane recursive trees constitutes Thm 1 of Mahmoud *et al.*'s paper [31]. The distribution result appears to be new and the expression is related to ballot numbers.

5.4 Profiles of trees

Consider the quantity L_{nk} representing the expected number of nodes at depth k on all trees of a variety $\mathcal{Y}$ with size n. (By convention the depth of the root is taken to be 0.) For fixed n, the sequence $\{L_{n,k}\}_{k=0}^{n}$ describes the mean "profile" of trees in the variety (see [32]). Let $L(z,u)$ be the bivariate generating function,

$$L(z,u) = \sum_{n,k\geq 0} L_{nk}\frac{z^n}{n!}u^k,$$

Theorem 8 *The bivariate generating function $L(z,u)$ of node levels satisfies*

$$L(z,u) = (Y'(z))^u \int_0^z (Y'(t))^{1-u}\,dt. \tag{22}$$

Let Ω_n be the depth of a random node in a random tree of $\mathcal{Y}$ with size n, i.e.,

$$Pr(\Omega_n = k) = \frac{L_{nk}}{\sum_k L_{nk}}.$$

For a polynomial variety of degree d, the mean μ_n and the variance σ_n^2 of Ω_n satisfy

$$\mu_n = (\delta+1)\log n + O(1) \quad \textit{and} \quad \sigma_n^2 = (\delta+1)\log n + O(1).$$

The distribution of Ω_n is asymptotically normal,

$$(\Omega_n - \mu_n)/\sigma_n \to \mathcal{N}(0;1),$$

in the sense of convergence in distribution.

Proof. Define $s[t]$, the level polynomial of a tree t, to be the sum $\sum_\nu u^{\mathrm{depth}(\nu)}$ taken over all nodes ν of t. Then $s[t]$ is inductively defined by

$$s[t] = \begin{cases} 1 & \text{if } |t| = 1 \\ u \sum_{\tau \propto t} s[\tau] & \text{otherwise.} \end{cases}$$

Algebraically, the generating polynomial $L_n(u) = \sum_k L_{nk} u^k$ behaves like the expectation of an inductive parameter. Using the same reasoning as for counting trees in Section 2 or for inductive maps, we get the equation

$$L(z,u) = Y(z) + u \int_0^z L(t,u)\phi'(Y(t))\,dt,$$

which translates into a linear differential equation

$$\frac{\partial}{\partial z} L(z,u) = Y'(z) + uL(z,u)\frac{\phi'(Y(z))Y'(z)}{\phi(Y(z))}, \quad L(0,u) = 1.$$

Integrating the homogeneous equation first, we get the solution

$$\exp(u \log Y'(z)) = (Y'(z))^u.$$

The integral form of $L(z,u)$ is then obtained by the variation–of–parameter method.

In the case of a polynomial ϕ, using Lemma 2, we determine

$$\log Y'(z) = (\delta + 1)\log\frac{1}{1 - z/\rho} + C + O[(1 - z/\rho)^{2\delta}].$$

A theorem of Flajolet and Soria [12] states that a bivariate scheme of the form $\exp(uL(z))$ for some function $L(z)$ with a dominant logarithmic singularity induces Gaussian distributions in the asymptotic limit. It applies here to $(Y'(z))^u = \exp(u \log Y'(z))$.

In the case of $L(z,u)$ in Eq. (22), the integral is convergent for u in a complex neighbourhood of 1 and $|z| \le \rho$ so that it plays the role of an unessential perturbation. A simply amended form of the main result of [12] then applies. □

The result for μ_n is also consistent with the estimate of expected path length which is precisely $n\mu_n + n$.

Corollary 6 *(i) For binary trees, $\phi(w) = (1 + w)^2$, we have*

$$L(z,u) = \frac{(1-z)^{-2u} - (1-z)^{-1}}{2u - 1},$$

$$\mu_n = 2\log n + 2\gamma - 4 + O(\log n/n), \quad \sigma_n^2 = 2\log n + 4 + 2\gamma - \frac{2\pi^2}{3} + O(\log^2 n/n).$$

(ii) For recursive trees, $\phi(w) = \exp(w)$,

$$L(z,u) = \frac{(1-z)^{-u} - 1}{u},$$

$$\mu_n = \log n + \gamma - 1 + O(1/n), \quad \sigma_n^2 = \log n + \gamma - \frac{\pi^2}{6} + O(1/n).$$

(iii) For plane recursive trees, $\phi(w) = (1-w)^{-1}$,

$$L(z,u) = \frac{(1-2z)^{-u/2} - (1-2z)^{1/2}}{1+u},$$

$$\mu_n = \frac{1}{2}\log n + \frac{\gamma}{2} + \log 2 - \frac{1}{2} + O(\log n/n), \quad \sigma_n^2 = \frac{1}{2}\log n + \frac{\gamma}{2} + \log 2 - \frac{1}{4} - \frac{\pi^2}{8} + O(\log^2 n/n).$$

For binary trees, the distribution has been given by a number of authors independently in the 1980's, see [28]; for recursive trees, it was found by Devroye [7] using the theory of records. Both explicit forms and limit distributions are derivable from these GF's. For recursive trees we find the coefficients $s_{n,k+1}$, while the two other ones give rise to convolution of Stirling numbers of the first kind. (We recall that the Stirling number of the first kind, s_{nk} is defined as $[u^k]\, u(u+1)(u+2)\cdots(u+n-1)$; Stirling numbers are ubiquitous in this category of problems.)

5.5 Distribution of the number of leaves

In order to approach the distribution of the number of leaves and more generally of nodes by sorts, we introduce

$$\phi(\mathbf{u}; y) = \sum_{r=0}^{\infty} \phi_r u_r y^r,$$

a series involving in general infinitely many indeterminates, (u_r). Let $N(n_0, n_1, \ldots)$ be the number of trees of size n (in the variety defined by ϕ) that have n_0 leaves, n_1 nodes of degree 1, etc. The multivariate GF

$$Y(\mathbf{u}; z) = \sum_{n_0, n_1, \ldots} N(n_0, n_1, \ldots) u_0^{n_0} u_1^{n_1} \cdots \frac{z^{n_0+n_1+\cdots}}{(n_0+n_1+\cdots)!}$$

satisfies

$$\int_0^{Y(\mathbf{u};z)} \frac{dt}{\phi(\mathbf{u};t)} = z.$$

This function condenses all the distribution information of all node sorts in a variety. In particular, it specializes to the bivariate GF for the distribution of the number of leaves, $Y(u;z) = Y(u,1,1,\ldots;z)$.

Theorem 9 *The bivariate generating function for leaves $Y(u;z)$ is defined implicitly by*

$$\int_0^{Y} \frac{dt}{(u-1)\phi_0 + \phi(t)} = z.$$

For random trees of size n in a polynomial variety, the number of leaves tends to a Gaussian limit (in the sense of convergence in distribution).

Proof. (Sketch) The argument for the Gaussian limit is based on analyzing the "perturbation" caused by the variable u in the non linear differential equation $Y' = \phi(Y)$, with u near 1. We consider some fixed u close enough to 1.

First, the radius of convergence of $Y(u;z)$ is

$$\rho(u) = \int_0^{\infty} \frac{dt}{(u-1)\phi_0 + \phi(t)}.$$

We have

$$\begin{aligned}\rho(u)-\rho(1) &= \int_0^\infty \frac{dt}{\phi(t)}\left(\frac{1}{1+\phi_0\frac{u-1}{\phi(t)}}-1\right)\\ &= -(u-1)\phi_0\int_0^\infty \frac{dt}{\phi^2(t)}+(u-1)^2\phi_0^2\int_0^\infty \frac{dt}{\phi^3(t)}-\cdots. \qquad (23)\end{aligned}$$

The singular exponent of $Y(u;z)$ is still δ, and we have

$$Y(u;z)=H_0(u)(1-z/\rho(u))^{-\delta}+H_1(u)+etc.$$

By a suitable use of majorizing series arguments, a uniform version of this expansion is established, when u lies in a real neighbourhood of 1. It is known from Bender's work that similar analytic schemas lead to Gaussian laws [2]. Here, we use a slightly stronger form of Bender's theorem discussed in [13]. In this way, the result is established. □

Let $Y_{n,k}$ denote the number of trees of size n having k leaves in variety $\mathcal{Y}$.

Corollary 7 *(i). For binary trees, $\phi(w)=(1+w)^2$, $Y_{n,k}$ is a pseudo-Eulerian number,*

$$Y_{n,k}=n![u^kz^n]Y(u;z) \qquad \textit{where}\ \ Y(u;z)=\xi\frac{\xi\tan(\xi z)+1}{\xi-\tan(\xi z)}-1 \ \ \textit{and}\ \ \xi=(u-1)^{1/2}.$$

(ii). For recursive trees, $Y_{n,k}$ is a shifted Eulerian number,

$$Y_{n,k}=A_{n-1,k}=(n-1)![u^kz^{n-1}]\,(u-1+\frac{1-u}{1-ue^{z(1-u)}}).$$

(iii). For plane recursive trees, $Y_{n,k}$ is a second order Eulerian number,

$$Y_{n,k}=n![u^kz^n]\frac{C(ue^{-u}e^{z(u-1)^2})-C(ue^{-u})}{u-1}, \qquad \textit{where} \qquad C(z)=\sum_{n=1}^{\infty}n^{n-1}\frac{z^n}{n!}$$

is Cayley's function.

In all three cases, the distribution is Gaussian in the limit.

Proof. The generating functions follow from straight integration. In the case of recursive trees, we get a modified GF for the Eulerian numbers in the form

$$Y(u;z)=\log\frac{1-u}{1-ue^{z(1-u)}}.$$

For plane recursive trees, we arrive at second order Eulerian numbers (for a definition, see [20, p. 256]). Cayley's function arises through inversion of $Ce^C=z$. □

The result for binary trees is classical. The one for recursive trees was found independently by several authors, an early reference being [34]. Mahmoud *et al.* [31] discovered the connection with second order Eulerian numbers. The EGF given above does not seem to have appeared in the literature however. The Gaussian law in this case is in [31] where it is derived from limit theorems on Pólya urn models.

Similar limit distribution results hold for other sorts of nodes, like i–nodes.

6 Extensions

Results in this paper can be extended in various directions. On the algebraic side, variants of our basic algebraic schemes may be considered. On the analytic side, we may treat varieties defined by degree functions that are either entire or have singularities at a finite distance.

6.1 Algebraic Schemes

The basic scheme of Thm. 1,

$$Y(z) = \int_0^z \phi(Y(t))\, dt,$$

arises with ϕ being rather generally the GF of a "species" of structures in the sense of Joyal [3, 22]. For instance, we can fit into our algebraic framework, the cases of

$$\phi(w) = 1 + \log \frac{1}{1-w}, \quad \phi(w) = \tan(w).$$

The first case corresponds to "mobiles" where subtrees dangling from a node are arranged in cyclic order and thus constitute a freely rotating cycle. The second case corresponds to "festoon" trees in which labels of the sons of a node go up an down, forming an alternating permutation. (Of course the increasing tree property is still assumed in these constructions.)

The algebra of such series is indistinguishable from that of increasing trees varieties considered earlier. The analysis, as we discuss below, can be treated along similar lines.

6.2 Analytic Schemes

The principles of analysis employed generalize to either entire functions or functions with singularities at a finite distance. The argument runs as follows.

The singular behaviour of $Y(z)$ is obtained by inverting the singular expansion of $\int dw/\phi(w)$ near its smallest positive singularity. The asymptotic form of the coefficients of $Y(z)$ derives from that singular behaviour by means of singularity analysis.

Assume that $\phi(w)$ which has positive coefficients becomes singular at $\sigma > 0$. We have either $\sigma = +\infty$ or $\sigma < \infty$ depending on whether ϕ is entire or not. In both cases, however, the argument employed in the proof of Thm. 2 generalizes, and the radius of convergence of $Y(z)$ is

$$\rho = \int_0^\sigma \frac{dw}{\phi(w)}. \tag{24}$$

Eq. (24) is consistent with what we have found (see Cor. 1) for recursive trees, where $\phi(w) = \exp(w)$ so that $\sigma = +\infty$, and for plane recursive trees where $\phi(w) = (1-w)^{-1}$ so that $\sigma = 1$. In these two cases, we have

$$\rho = \int_0^{+\infty} e^{-w}\, dw = 1 \quad \text{and} \quad \rho = \int_0^1 (1-w)\, dw = \frac{1}{2}.$$

Singularity at a finite distance. Consider the typical case of

$$\phi(Y) = \frac{1}{P(Y)},$$

with P a polynomial. In this case, the equation $P(Y)Y' = 1$ has the integral $Q(Y) = z$, with $Q' = P$ and $Q(0) = 0$. Thus Y is an algebraic function and its coefficients are *a priori* of the asymptotic form $C\rho^{-n}n^{p/q}$ for some rational p/q. The singularity of $Y(z)$ is at

$$\rho = \int_0^\sigma P(y)\,dy = Q(\sigma),$$

where σ is the smallest positive root of P. The singular expansion of Y is then found by inverting

$$\rho - z = \int_{Y(z)}^\sigma P(y)\,dy = Q(\sigma) - Q(Y(z)).$$

In the generic case, $P'(\sigma) \neq 0$, and we obtain the singular expansion

$$Y(z) = \sigma - \sqrt{\frac{2(\rho - z)}{|P'(\sigma)|}} + O(\rho - z), \quad z \to \rho,$$

from which we deduce

$$\frac{Y_n}{n!} = \sqrt{\frac{\rho}{\pi|P'(\sigma)|}}\rho^{-n}n^{-3/2}[1 + O(1/n)].$$

Thus the particular explicit form obtained for plane recursive trees is indeed attached to a fairly general scheme.

EXAMPLE 4. *Special plane trees and festoon trees.* As an illustration, consider plane recursive trees in which nodes degrees are all multiples of a fixed integer t, so that $\phi(w) = (1 - w^t)^{-1}$. In this case, the equation for Y, which is $Y'(1 - Y^t) = 1$ leads to

$$Y - \frac{Y^{t+1}}{t+1} = z.$$

The trinomial equation is solvable by Lagrange inversion, from which we find

$$Y_{nt+1} = \frac{(nt)!}{(t+1)^n}\binom{n(t+1)}{n},$$

a formula that generalizes that of plane recursive trees and is also in agreement with the asymptotic estimates above.

Festoon trees already discussed correspond to $\phi(w) = \tan(w)$, so that $\sigma = \frac{\pi}{2}$. We find $\rho = \frac{\pi}{4}$, and asymptotically, by a direct extension of the situation of rational functions, $Y_n/n! \sim C(4/\pi)^n n^{-3/2}$, for some $C > 0$. □

EXAMPLE 5. *Mobile trees.* They correspond to $\phi(w) = 1 + \log(1 - w)^{-1}$. The asymptotic analysis takes us a little out of beaten tracks. The radius of convergence ρ is

$$\rho = \int_0^\infty \frac{e^{-t}}{1+t}\,dt = 1! - 2! + 3! - 4! + \cdots = -e\gamma + e\sum_{n=1}^\infty \frac{(-1)^{n-1}}{n \cdot n!} \approx 0.59634\,73623.$$

The local expansion of $Y(z)$ is provided by inverting the famous divergent series of Euler,

$$Y(z) = 1 - (\rho - z)\log\frac{1}{1 - z/\rho} + (\rho - z)\log\log\frac{1}{1 - z/\rho} + O\Big(\frac{(\rho - z)(\log\log(1 - z/\rho)^{-1})^2}{\log(1 - z/\rho)}\Big).$$

This type of singular expansion itself necessitates the full power of singularity analysis, and we find

$$Y_n = \rho^{1-n}\left[\frac{1}{n^2} - \frac{1}{n^2\log n} + O\Big(\frac{1}{n^2\log^2 n}\Big)\right].$$

□

Entire functions. The same principles apply. If $\phi(w) \sim K w^m e^w$, for some constant K and integer m, when $w \to +\infty$, then

$$\rho - z = \int_Y^\infty \frac{dw}{\phi(w)} \sim \frac{1}{K e^Y Y^m},$$

so that $Y(z)$ is dominated by a logarithmic term at $z = \rho$, and

$$\frac{Y_n}{n!} \sim \rho^{-n} n^{-1}. \tag{25}$$

EXAMPLE 6. *Even and Odd Trees.* Non plane even trees have all their nodes of even degree, so that $\phi(w) = \cosh(w)$. Odd trees have all their non leaf nodes of odd degree, so that $\phi(w) = 1 + \sinh(w)$. The EGF's are explicitly computable in both cases, and we find:

$$Y^{\text{even}}(z) = \log \tan\left(\frac{z}{2} + \frac{\pi}{4}\right), \quad Y^{\text{odd}} = \log\left[\frac{(\sqrt{2}+1)(e^{\sqrt{2}z}+1)}{3 + 2\sqrt{2} - e^{\sqrt{2}z}}\right].$$

The corresponding singularities are at

$$\rho^{\text{even}} = \frac{\pi}{2}, \quad \rho^{\text{odd}} = \frac{\sqrt{2}}{2}\log(3 + 2\sqrt{2}).$$

The coefficients obey Eq. (25). □

The example of $\phi(w) = e^w - w$ is treated in full detail by Meir and Moon [33]. Their main result,

$$\frac{Y_n}{n!} \sim \rho^{-n} n^{-1}, \quad \rho = \int_0^\infty \frac{dw}{e^w - w},$$

which is Thm. 3 of [33], also matches with our Eq. (25).

7 Conclusion

We have demonstrated here a fairly general approach to the analysis of tree parameters where a basic equation for trees is studied from the point of view of its singularities. Here, we have been dealing with a non linear autonomous differential equation or order 1, $Y' = \phi(Y)$.

Major characteristic parameters of trees have GF's that are expressible in terms of the basic GF $Y(z)$ by means of transformations (here integrals). Once we view these expressions as "singularity transformers", it becomes possible to study a large number of statistical problems in a unified manner. In this context, composition theorems for singularities of analytic functions prove especially valuable.

The techniques developed here are not restricted to varieties of increasing trees. Notably, Knuth and Pittel's results regarding union–find trees [25] are amenable to analytic techniques based on singularity analysis instead of recurrences, this being done along the very same steps as in this paper.

The perturbation techniques used here in order to derive limit distributions by means of bivariate analytic schemes of a general nature certainly deserve further attention. For instance, some counterparts in linear cases have already proved useful in analyzing distributions of quadtrees [10]. Quite clearly general bivariate analytic schemes on differential equations are conducive to Gaussian laws under quite a wide range of conditions.

Acknowledgements. This work was supported in part by the Basic Research Action of the E. C. under contract No. 3075 (Project ALCOM).

The authors are grateful to Hosam Mahmoud for communicating many bibliographical pointers and preprints.

References

[1] BALIŃSKA, K. T., QUINTAS, L. V., AND SZYMANSKY, J. Random recursive forests. Preprint, August 1991.

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

[3] BERGERON, F., LABELLE, G., AND LEROUX, P. Computation of the expected number of leaves in a tree having a given automorphism, and related topics. preliminary version, 1991.

[4] BURGE, W. H. An analysis of a tree sorting method and some properties of a set of trees. In *First USA-JAPAN Computer Conference Proceedings* (October 1972), AFIPS and IPSJ, pp. 372–378.

[5] CHEN, W., AND NI, W. Heap ordered trees. Personal communication, October 1991. Based on Wen–Chun Ni's Master Thesis, National Taiwan University.

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

[7] DEVROYE, L. Applications of the theory of records in the study of random trees. *Acta Informatica 26* (1988), 123–130.

[8] EVGRAFOV, M. A. *Analytic Functions.* Dover, New York, 1966.

[9] FLAJOLET, P. Random tree models in the analysis of algorithms. In *PERFORMANCE'87* (1988), P.-J. Courtois and G. Latouche, Eds., Elsevier Science Publishers (North Holland), pp. 171–187. (Invited lecture).

[10] FLAJOLET, P., GONNET, G., PUECH, C., AND ROBSON, J. M. Analytic variations on quadtrees. *Algorithmica* (1992). 24 pages, to appear.

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

[12] FLAJOLET, P., AND SORIA, M. Gaussian limiting distributions for the number of components in combinatorial structures. *Journal of Combinatorial Theory,* Series A *53* (1990), 165–182.

[13] FLAJOLET, P., AND SORIA, M. General combinatorial schemas: Gaussian limit distributions and exponential tails. *Discrete Mathematics* (1991). To appear in the Special Issue on *Combinatorics and Algorithms*, A. S. Fraenkel Editor. Available as Research Report 632, LRI, Université Paris-Sud, January 1991. 21 pages.

[14] FORD, W. B. *Studies on divergent series and summability and the asymptotic developments of functions defined by Maclaurin series*, 3rd ed. Chelsea Publishing Company, New York, 1960.

[15] FRANÇON, J. Arbres binaires de recherche : Propriétés combinatoires et applications. *RAIRO Informatique Théorique 10*, 12 (Dec. 1976), 35–50.

[16] GASTWIRTH, J. L. A probability model of a pyramid scheme. *American Statistician 31* (1977), 79–82.

[17] GASTWIRTH, J. L., AND BHATTACHARYA, P. K. Two probability models of pyramid or chain letter schemes demonstrating that their promotional claims are unreliable. *Operations Research 32*, 3 (May-June 1984), 527–536.

[18] GONNET, G. H., AND BAEZA-YATES, R. *Handbook of Algorithms and Data Structures: in Pascal and C*, second ed. Addison–Wesley, 1991.

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

[20] GRAHAM, R., KNUTH, D., AND PATASHNIK, O. *Concrete Mathematics.* Addison Wesley, 1989.

[21] HENNEQUIN, P. Combinatorial analysis of quicksort algorithm. *RAIRO Theoretical Informatics and Applications 23*, 3 (1989), 317–333.

[22] JOYAL, A. Une théorie combinatoire des séries formelles. *Advances in Mathematics 42*, 1 (1981), 1–82.

[23] KNUTH, D. E. *The Art of Computer Programming*, vol. 1: Fundamental Algorithms. Addison-Wesley, 1968. Second edition, 1973.

[24] KNUTH, D. E. *The Art of Computer Programming*, vol. 3: Sorting and Searching. Addison-Wesley, 1973.

[25] KNUTH, D. E., AND PITTEL, B. A recurrence related to trees. *Proceedings of the American Mathematical Society 105*, 2 (Feb. 1989), 335–349.

[26] LEROUX, P., AND VIENNOT, G. X. Combinatorial resolution of systems of differential equations I: Ordinary differential equations. In *Combinatoire Énumérative* (1986), G. Labelle and P. Leroux, Eds., no. 1234 in Lecture Notes in Mathematics, Springer-Verlag, pp. 210–245.

[27] LEROUX, P., AND VIENNOT, X. G. Combinatorial resolution of systems of differential equations. IV. separation of variables. *Discrete Mathematics 72* (1988), 237–250.

[28] LOUCHARD, G. Exact and asymptotic distributions in digital and binary search trees. *RAIRO Theoretical Informatics and Applications 21*, 4 (1987), 479–495.

[29] MAHMOUD, H. *Evolution of Random Search Trees.* John Wiley, New York, 1991. In press.

[30] MAHMOUD, H. M. Distances in random plane-oriented recursive trees, 1991. Preprint. To appear in the *Journal of Applied and Computational Mathematics*, special issue on asymptotic analysis in discrete mathematics.

[31] MAHMOUD, H. M., SMYTHE, R. T., AND SZYMAŃSKI, J. On the structure of random plane-oriented recursive trees and their branches, 1991. Preprint.

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

[33] MEIR, A., AND MOON, J. W. Recursive trees with no nodes of out-degree one. *Congressus Numerantium 66* (1988), 49–62.

[34] NAJOCK, D., AND HEYDE, C. C. On the number of terminal vertices in certain random trees with an application to stemma constructions in philology. *Journal of Applied Probability 19* (1982), 675–680.

[35] STANLEY, R. P. *Enumerative Combinatorics*, vol. I. Wadsworth & Brooks/Cole, 1986.

[36] SZYMAŃSKI, J. On a non-uniform random recursive tree. In *Random Graphs '85*, M. Karoński and Z. Palka, Eds., vol. 33 of *Annals of Discrete Mathematics.* North-Holland, 1987, pp. 297–306.

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

[38] VITTER, J. S., AND FLAJOLET, P. Analysis of algorithms and data structures. In *Handbook of Theoretical Computer Science*, J. van Leeuwen, Ed., vol. A: Algorithms and Complexity. North Holland, 1990, ch. 9, pp. 431–524.

[39] VUILLEMIN, J. A unifying look at data structures. *Communications of the ACM 23*, 4 (Apr. 1980), 229–239.

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

[41] WILSON, R. Functions with dominant singularities of the generalized algebraic-logarithmic type (II): On the order of the Hadamard product. *Proceedings of the London Mathematical Society*, Series 2 *43*, 2190 (1937), 417–438.