Fifth Colloquium on Mathematics and Computer Science DMTCS proc. **AI**, 2008, 121–134

The height of random binary unlabelled trees

Nicolas Broutin and Philippe Flajolet

Algorithms Project, INRIA-Rocquencourt, F-78153 Le Chesnay (France)

This extended abstract is dedicated to the analysis of the height of non-plane unlabelled rooted binary trees. The height of such a tree chosen uniformly among those of size n is proved to have a limiting theta distribution, both in a central and local sense. Moderate as well as large deviations estimates are also derived. The proofs rely on the analysis (in the complex plane) of generating functions associated with trees of bounded height.

Keywords: Average case analysis, height, limit distribution, local limit theorem, generating functions.

1 Introduction

We consider trees that are *binary, non-plane, unlabelled, and rooted*; that is, a tree is taken in the graph-theoretic sense and it has nodes of (out)degree two or zero only; a special node is distinguished, the root, which has degree two. In this model, the nodes are indistinguishable, while no order is assumed between the neighbours of a node. Let $\mathcal{Y}$ denote the class of such trees, and let $\mathcal{Y}_n$ be the subset consisting of trees with n external nodes (i.e., nodes of degree zero).

In this extended abstract, we study the (random) *height* H_n of a tree sampled uniformly from $\mathcal{Y}_n$. The depth of nodes has been analysed for many "*simple varieties*" of trees by Meir and Moon [22]. Regarding height, a few special cases were studied early: Rényi and Szekeres [25] proved that the average height of *labelled* non-plane trees of size n is asymptotic to $2\sqrt{\pi n}$; De Bruijn, Knuth, and Rice [7] dealt with *plane* trees and showed that the average height is equivalent to $\sqrt{\pi n}$ as $n \to \infty$. Finally, Flajolet and Odlyzko [10] developed an approach that encompasses all simple varieties of trees (see [12] for more results). These results are relative to trees where one can distinguish the neighbours of a node, either by their labels (labelled trees), or by the order induced on the progeny (plane trees).

In such models with distinguishable progeny, there are natural random walks associated to the random trees. Also, trees of a fixed size n may be seen as Galton–Watson processes conditioned on the size of the total progeny being n; cf [1, 17, 19]. Probabilistic techniques have been applied to the random walk associated to a tree traversal in order to derive asymptotic results about the trees, in particular regarding height and width [4, 5]. Yet an other approach is to find the continuous limit of suitably rescaled random trees of increasing sizes. One can then read off some of the limit parameters directly on the limiting object. This point of view has been adopted by Aldous [2] in his definition of the continuum random tree. For a recent account of the probabilistic developments, see the survey by Le Gall [20].

The case of trees with indistinguishable progeny is essentially different, and no direct random walk approach would seem to be possible, due to the symmetries inherent in unlabelled structures. The analysis of unlabelled non-plane trees finds its origins in the works of Pólya [24] and Otter [23]. However, these authors mostly focused on enumeration—the problem of characterizing typical parameters of these random trees remains largely untouched. Gittenberger [15] and recently, in an independent study that predates ours by a couple of weeks, Drmota–Gittenberger [8] have examined the profile of non-plane unlabelled *"general" trees* (where all degrees are allowed) and shown that the joint distribution of the number of nodes at any finite number of levels converges weakly to the finite-dimensional distribution of Brownian excursion local times. They further extended the result to a convergence of the entire profile to the Brownian excursion local time. This gives in particular the limit law for the height of these trees. This suggests that, although there is no clear reduction of unlabelled trees to random walks, such trees largely behave like simply generated families. In particular, this suggests that the rescaled height $H_n/\sqrt{n}$ should admit a limit distribution of the theta type [10, 18, 25]. We shall prove that such is indeed the case for *binary non-plane trees* in Theorems 1 and 2 below. We also provide moderate and large deviations estimates (Theorems 3 and 4) as well as asymptotic estimates for the moments (Theorem 5); see §6. Some of the more technical proofs are omitted and we limit ourselves to the global structure of the arguments; the details may be found in the long version [3].

This note arose from questions of Jean-François Marckert and Grégory Miermont [21]. Their motivation comes from an attempt at extending the probabilistic methods of Aldous to non-plane trees and developing corresponding continuous models—we are indebted to them for being at the origin of the present study. We also express our gratitude to Alexis Darrasse and Carine Pivoteau for designing and programming for us efficient Boltzmann samplers of binary trees and providing detailed statistical data that guided our first analyses of this problem.

2 Trees and generating functions

Tree enumeration. Our approach is entirely based on *generating functions*. The class $\mathcal{Y}$ of binary (non-plane unlabelled rooted) trees is defined to include the tree with a single external node. A tree has size n if it has n external nodes (hence $n-1$ internal nodes). The cardinality of the subclass $\mathcal{Y}_n$ of trees of size n is denoted by y_n and the generating function (GF) of $\mathcal{Y}$ is

$$y(z) := \sum_{n\geq 1} y_z z^n = z + z^2 + z^3 + 2z^4 + 3z^5 + 6z^6 + 11z^7 + 23z^8 + \cdots,$$

the coefficients corresponding to the entry A001190 (Wedderburn–Etherington numbers) of Sloane's *On-line Encyclopedia of Integer Sequences*.

Since a binary tree is either an external node or a root appended to an unordered pair of two (not necessarily distinct) binary trees, one has the basic functional equation

$$y(z) = z + \frac{1}{2}y(z)^2 + \frac{1}{2}y(z^2), \tag{1}$$

as follows from fundamental principles of combinatorial enumeration [11, 16, 23, 24]. According to the general principles of analytic combinatorics, we shall operate in an essential manner with properties of generating functions in the *complex plane*. The following lemma is classical:

Lemma 1 *Let ρ be the radius of convergence of $y(z)$. Then, one has $1/4 \leq \rho \leq 1/2$, and ρ is determined implicitly by $2\rho + y(\rho^2) = 1$. As $z \to \rho^-$, the generating function $y(z)$ satisfies*

$$y(z) = 1 - \lambda\sqrt{1 - z/\rho} + O\left(1 - z/\rho\right), \qquad \lambda = \sqrt{2\rho + 2\rho^2 y'(\rho^2)}. \tag{2}$$

Furthermore, the number y_n of trees of size n satisfies asymptotically

$$y_n = \frac{\lambda}{2\sqrt{\pi}} \cdot n^{-3/2} \rho^{-n} (1 + O(1/n)). \tag{3}$$

Proof: Let $C_{n-1} = \frac{1}{n}\binom{2n-2}{n-1}$ be the number of plane binary trees of size n. One has combinatorially $C_{n-1}2^{-n} \leq y_n \leq C_n$. The C_n, which are the Catalan numbers and they are well-known to grow as $4^n n^{-3/2}$. The bounds $1/4 \leq \rho \leq 1/2$ result. It follows that $y(z^2)$ is analytic in a disc properly containing $|z| \leq \rho$. Then, from (1), upon solving for $y(z)$, we obtain

$$y(z) = 1 - \sqrt{1 - 2z - y(z^2)}, \tag{4}$$

which becomes singular when the argument of the square root vanishes. The value ρ is then the positive solution of $2z + y(z^2) = 1$ and, at this point, we must have $y(\rho) = 1$. This reasoning also justifies the singular expansion (2), seen to be valid in a Δ-domain extending beyond the disc of convergence $|z| < \rho$.

Equation (3) constitutes Otter's celebrated estimate: it results from translating the square-root singularity of $y(z)$ by means of either Darboux's method or singularity analysis [11, 16, 23, 24]. □

Numerically, one finds [9, 11, 23]:

$$\rho \doteq 0.40269\,750367, \qquad \lambda \doteq 1.13003\,37163, \qquad \frac{\lambda}{2\sqrt{\pi}} \doteq 0.31877\,66259.$$

Height. Let $y_{h,n}$ be the number of trees of size n and height *at most* h. (Height is measured as the maximum number of edges along branches from the root to external nodes.) Let $y_h(z) = \sum_{n\geq 1} y_{h,n} z^n$. The arguments leading to (1) yield the fundamental recurrence

$$y_{h+1}(z) = z + \frac{1}{2} y_h(z)^2 + \frac{1}{2} y_h(z^2), \qquad \text{and} \qquad y_0(z) = z. \tag{5}$$

We also set $e_h(z) \equiv \sum_{n\geq 1} e_{h,n} z^n := y(z) - y_h(z)$, which is the generating function of trees with height *exceeding* h. Then, a trite calculation shows that the $e_h(z)$ satisfy the main recurrence

$$e_{h+1}(z) = y(z) e_h(z) \left(1 - \frac{e_h(z)}{2y(z)}\right) + \frac{1}{2} e_h(z^2), \qquad \text{and} \qquad e_0(z) = y(z) - z, \tag{6}$$

on which our subsequent treatment of height is entirely based.

Analysis. The distribution of height is accessible by

$$\mathbb{P}\{H_n > h\} = \frac{y_n - y_{h,n}}{y_n} = \frac{e_{h,n}}{y_n}, \tag{7}$$

where $e_{h,n} = [z^n] e_h(z)$. We shall get a handle on its asymptotic properties by means of Cauchy's coefficient formula,

$$e_{h,n} = \frac{1}{2i\pi} \int_\gamma e_h(z) \frac{dz}{z^{n+1}}, \tag{8}$$

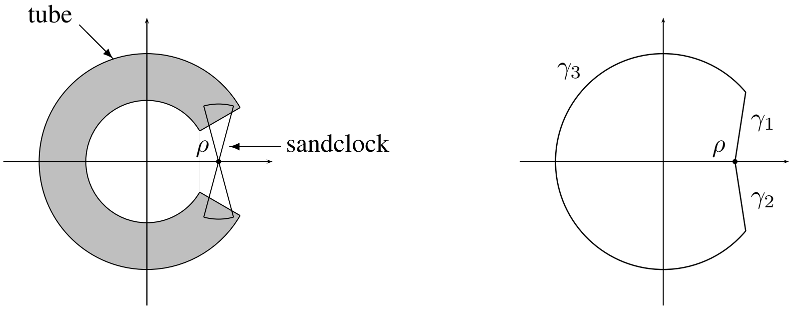

Fig. 1. The "tube" and "sandclock" regions on the left, and the Hankel contour used to estimate $e_{h,n}$ on the right.

upon choosing a suitable integration contour γ in (8). This task necessitates first developping suitable estimates of $e_h(z)$, for values of z both *inside* and *outside* of the disc of convergence $|z| < \rho$. Precisely, we shall need estimates valid in a *"tube"* around an arc of the circle $|z| = \rho$,

$$\mathcal{T}(\mu, \eta) := \{z : \quad -\mu < |z| - \rho < \mu, \ |\arg(z)| > \eta\}. \tag{9}$$

as well as inside a *"sandclock"* anchored at ρ

$$\mathcal{S}(r_0, \theta_0) := \{z : \quad |z - \rho| < r_0, \quad \pi/2 - \theta_0 < |\arg(z - \rho)| < \pi/2 + \theta_0\}. \tag{10}$$

(See Figure 1 for a rendering).

Plan. Estimates of the sequence of generating functions $(e_h(z))$ within the disc of convergence and a tube, where z stays away from the singularity ρ form the subject of Section 3. The bulk of the technical work is relative to the sandclock, in Section 4. We then develop our main approximation in Section 5. Sections 3–5 closely follow the general strategy of the original paper [10]; however, nontrivial adaptations are needed, due to the presence of Pólya terms, so that the problem is no longer of a "pure" iteration type, as in [10]. We finally reap the crop in Section 6, where our main theorems relative to the distribution of the height are stated and proved (these somewhat parallel the local limit law of [12] in the planar case).

3 Convergence away from the singularity

Our aim in this section[(i)] is to extend the domain where e_h is analytic beyond the disc of convergence $|z| \le \rho$, when z stays in a "tube" $\mathcal{T}(\mu, \eta)$ as defined in (9) and is thus away from ρ. The main result is summarized by Proposition 1, at the end of this section. Its proof relies on the combination of two ingredients: first, the fact, expressed by Lemma 2, that $e_h \to 0$ (equivalently: $y_h \to y$) in the closed disc of radius ρ; second, a general criterion for convergence of the e_h to 0, which is expressed by Lemma 3. The criterion implies in essence that the convergence domain is an open set, and this fact provides the basic analytic continuation of the generating functions of interest.

Lemma 2 *For all z such that $|z| \le \rho$, and $h \ge 1$, one has $|e_h(z)| \le (|z|/\rho)^h/\sqrt{h}$.*

[(i)] In what follows, we freely omit the arguments of y, $e_h, \ldots$, whenever they are taken at z.

Proof: To have height at least h, a tree needs at least $h+1$ nodes, so that $|e_h(z)| \le \sum_{n>h} y_n |z|^n$. We first note an easy numerical refinement of (3), namely, $y_n \le \frac{1}{2}\rho^{-n} n^{-3/2}$. (See [14] for a detailed proof strategy in the case of a similar but harder problem.) The claim follows by bounding the sum using an integral. $\square$

We now devise a criterion for the convergence of the e_h to 0. This criterion, adapted from [10, Lemma 1], is crucial in obtaining extended convergence regions, both near the circle $|z| = \rho$ (in this section) and near the singularity ρ (in Section 4).

Lemma 3 (Convergence criterion) *Let $z \in \mathcal{D} := \{z : |y| < 1\}$ and assume that $|z| < \sqrt{\rho}$. The sequence $\{|e_h(z)|, h \ge 0\}$ converges to 0 if and only if there exist an integer m and real numbers $\alpha, \beta \in (0,1)$, such that the following three conditions are simultaneously met:*

$$|e_m| < \alpha, \qquad |y| + \alpha/2 < \beta, \qquad \alpha\beta + \left(|z|^2/\rho\right)^m < \alpha. \tag{11}$$

Furthermore, if (11) holds then, for some C and $\beta_0 \in (0,1)$, one has $|e_h| \le Ch\beta_0^h$, for all $h \ge m$.

Proof: (i) Since $|y| < 1$ in $\mathcal{D}$ and $z < \sqrt{\rho}$, the convergence $e_h \to 0$ clearly implies that (11) holds.

(ii) Conversely, assume the three conditions in (11), for some value m. Then, they also hold for $m+1$. Indeed, recalling (6), we see that, for any $h \ge 0$,

$$|e_{h+1}| \le |e_h|\left(|y| + \frac{|e_h|}{2}\right) + \frac{|e_h(z^2)|}{2} \le |e_h|\left(|y| + \frac{|e_h|}{2}\right) + \left(\frac{|z|^2}{\rho}\right)^h, \tag{12}$$

where the Pólya term involving $|e_h(z^2)|$ has been bounded using Lemma 2. Using the hypotheses of (11) together with (12) above taken at $h = m$, yields $|e_{m+1}| < \alpha$. By induction, (11) then holds for all $h \ge m$.

(iii) The assertion that $|e_h| \le Ch\beta_0^h$ follows by expanding (6) and using $\beta_0 = \max\{\beta, |z|^2/\rho\}$. $\square$

We can now state the main convergence result of this section:

Proposition 1 (Convergence in "tubes") *For any angle $\eta > 0$, there exists a tube $\mathcal{T}(\mu, \eta)$ with width $\mu > 0$, such that $|e_h(z)| \to 0$, as $h \to \infty$, whenever $|z|$ lies in $\mathcal{T}(\mu, \eta)$.*

Proof sketch: If we exclude a small sector of opening angle 2η around the positive real axis, then the quantity, $\lambda_0 := \sup\{\, |y(z)|; \quad |z| = \rho, \, |\arg(z)| \ge \eta \,\}$, satisfies $\lambda_0 < 1$. The continuity of y and Lemma 3 then imply the convergence of $e_h(z)$ to 0 in a small disc around each $z \in \{\rho e^{i\theta} : |\theta| \ge \eta\}$. The latter set being compact, we can extract a finite covering, which must then contain the desired tube where the convergence holds. $\square$

4 Convergence near the singularity

We now focus on the behaviour of $e_h(z)$ in a "sandclock" around the singularity. When z approaches ρ, $|y|$ is no longer bounded away from 1 and the criterion for convergence given by Lemma 3 cannot be used directly. However, as we prove below, the quantities $|e_h(z)|$ first exhibit a decreasing behaviour for $h \le N$, for some appropriate $N = N(z)$. At that point, $|e_N(z)|$ is small enough for the criterion of Lemma 3 to be satisfied, whence enventually the convergence $|e_h(z)| \to 0$ as $h \to \infty$ in a sandclock.

The upper bound we use for $|e_N|$ is based on the following alternative recurrence.

Lemma 4 (Alternative recurrence) *If* $e_i \neq 0$ *and* $e_i\left[1-e_i(z^2)/e_i^2\right] \neq 2y$ *for* $i=0,\ldots,h-1$ *then the following recurrence relations hold*

$$\frac{y^h}{e_h} = \frac{1}{e_0} + \frac{1}{2y}\frac{1-y^h}{1-y} - \sum_{i=0}^{h-1}\frac{y^{i-1}e_i(z^2)}{2e_i^2} + \sum_{i=1}^{h-1}\frac{y^{i-2}e_i}{4}\left[1-\frac{e_i(z^2)}{e_i^2}\right]^2\left(1-\frac{e_i}{2y}\left[1-\frac{e_i(z^2)}{e_i^2}\right]\right)^{-1}. \quad (13)$$

Proof: The proof relies on a classical idea in the study of slowly convergent iterations [6, 10]. Starting with the recurrence relation (6), rewritten as

$$\frac{e_{i+1}}{y^{i+1}} = \frac{e_i}{y^i}\left(1-\frac{e_i}{2y}\left[1-\frac{e_i(z^2)}{e_i^2}\right]\right),$$

the trick is to *take inverses*. This is a classical technique in the study of slowly convergent iterations (near an indifferent fixed point): see for instance [6, page 152]. Using the identity $(1-u)^{-1} = 1+u+u^2(1-u)^{-1}$ shows that

$$\frac{y^{i+1}}{e_{i+1}} - \frac{y^i}{e_i} = \frac{y^{i-1}}{2}\left[1-\frac{e_i(z^2)}{e_i^2}\right]\left(1-\frac{e_i}{2y}+\frac{y^{i-2}}{4}\left[1-\frac{e_i(z^2)}{e_i^2}\right]^2\left[1-\frac{e_i(z^2)}{e_i^2}\right]\right)^{-1}.$$

Summing the terms of this equality for $i=0,\ldots,h-1$ yields the claim. □

An important step of the proof of Proposition 2 consists in controlling the behavior the terms in the alternative recurrence for all $h \leq N$ (we will fix N later). In particular, we need good enough upper and lower bounds for $|e_h(z)|$, $h \leq N$, and for z around ρ and ρ^2. Obtaining such estimates requires to study the recurrence relation (6) more carefully and to quantify the effect of the analytic Pólya term $e_h(z^2)$ for z in the "sandclock" $\mathcal{S}(r_0,\theta_0)$. The following lemma is evocative of the theory of iteration near an attractive fixed point and gives some estimates on the behaviour of e_h in the interior of the disc of convergence.

Lemma 5 *There exist constants* $K_1, K_2 > 0$ *and* $r > 0$*, such that, for all* $h \geq 0$*, and for* $|z-\rho^2| < r$ *one has* $e_h(z) = C_h(z)\cdot y(z)^h$*, where* $C_h(z) = (C(z)+o(1))y(z)^h$ *and* $C(z)$ *is analytic at* ρ^2*. Furthermore,* $K_1 < |C_h(z)| < K_2$ *and* $\arg(e_h(z)) \leq c_0(h+1)r$*.*

With Lemma 5 in hand, we can obtain a first set of properties of $e_h(z)$ for $z \in \mathcal{S}(r_0,\theta_0)$ and h not too large (depending on z). These will be used to derive an upper bound on $|e_N|$ and prove that e_N satisfies the criterion of Lemma 3. In the following, we only need to consider $z \in \mathcal{S}(r_0,\theta_0)$, with $\Im(z) \geq 0$, since we clearly have $e_h(\bar{z}) = \overline{e_h(z)}$, where $\bar{z}$ denotes the complex conjugate of z.

Lemma 6 (The initial behavior of $|e_h|$**)** *Suppose* $\Im(z) \geq 0$*. Let* $N(z) = \lfloor \arccos(1/4)/\arg y(z) \rfloor$*. There exist constants* $r_0 > 0$ *and* $\theta_0 > 0$ *such that if* z *lies in the sandclock* $\mathcal{S}(r_0,\theta_0)$ *then, for* $1 \leq h \leq N(z)$*:*

$$\frac{|y|^{h+1}}{2(h+1)} < |e_h(z)| < 1/2 \qquad \textit{and} \qquad 0 \leq \arg(e_{h+1}) \leq (h+1)\arg(y). \quad (14)$$

Furthermore, one has $|e_h(z)| < 1/5$*, for* $6 \leq h \leq N(z)$*.*

Proof: (i) We first focus on the proof of the upper bound in (14). Consider the recurrence relation (6) rewritten as

$$e_{h+1}/y = y \cdot e_h/y\left(1 - \frac{e_h/y}{2}\right) + \frac{e_h(z^2)}{2y}. \tag{15}$$

The behavior of the first term in (15) is dictated by properties of $g : w \mapsto w(1 - w/2)$. A very similar function appeared in the analysis of Flajolet and Odlyzko [10], Lemma 3. By a simple modification of the proof in [10], we have

$$\left\{\begin{array}{c} |w| \le 1 \\ 0 \le \arg w \le \arccos(1/4) \end{array}\right. \quad \Rightarrow \quad \left\{\begin{array}{c} |g(w)| \le |w| \\ 0 \le \arg g(w) \le \arg w. \end{array}\right. \tag{16}$$

We use (16) and an induction argument to prove that, for $0 \le h \le N(z)$,

$$|e_h| \le 1/2 \qquad \text{and} \qquad \arg e_h \le (h+1)\arg y. \tag{17}$$

Write $z = \rho + re^{it}$. Then, by Lemma 1, we have $e_0/y = 1 - z/y$ so that $|e_0/y| \le 1 - \rho + O(\sqrt{r})$ and $0 \le \arg(e_0/y) = O(\sqrt{r})$. In particular, for r small enough, (17) holds for $h = 0$.

Suppose now that (17) holds for all integers up to h. To determine whether it also holds for $h+1$, we have to take the second term in the right-hand side of (15) into account. For r_0 small enough and $z \in \mathcal{S}(r_0, \pi/8)$, Lemma 5 ensures that this second term cannot contribute any increase in the argument of e_h/y. Therefore,

$$\arg(e_{h+1}/y) = \arg(e_h/y) + \arg(y) \ \le \ (h+2)\arg(y).$$

Furthermore, since y is analytic in $\mathcal{D}$, we have $|e_i| = |e_i(\rho)| + O(\sqrt{r})$ as $r \to 0$, for all fixed $i \ge 0$. So for r_0 small enough and $z \in \mathcal{S}(r_0, \pi/8)$, $|e_{h+1}(z)| \le 1/2$ if $h \le 5$. Now, if $h \ge 5$,

$$|e_{h+1}/y| \ \le \ |e_6(z)/y| + \frac{1}{2}\cdot\sum_{i=6}^{h+1}|e_i(z^2)/y| \ \le \ |e_6(\rho)| + O(\sqrt{r}) + \frac{1}{2}\cdot\sum_{i=6}^{\infty}(\rho+3r)^i,$$

since $e_i(z^2) \le (|z|^2/\rho)^i$ by Lemma 2. One can then verify that, for $h \ge 6$, $|e_{h+1}/y| < 1/5$, for r_0 small enough. So, a fortiori, among the conditions in (16), the one on the modulus holds as long as that on the argument does. The latter one holds for all $h \le N$.

(ii) It remains to prove the lower bound in (14); we start with (15). For $h \le N(z)$, the additional Pólya term $e_h(z^2)$ only contributes to making $|e_{h+1}|$ larger: for $z \in \mathcal{S}(r_0, \theta_0)$, by Lemma 5 and the upper bound we have just proved, the arguments of both terms are such that, for $h < N(z)$,

$$|e_{h+1}/y| \ge |y| \cdot |e_h/y| \cdot \left(1 - \frac{|e_h/y|}{2}\right).$$

Since $x \mapsto x(1 - x/2)$ is increasing in $[0,1]$, we have, for all $h \ge 0$, $|e_h/y| \ge f_h$, where the sequence $(f_h)_{h\ge 0}$ is defined by $f_{h+1} = |y| \cdot f_h \cdot (1 - f_h/2)$ and $f_0 = |e_0|/|y|$. The latter recurrence relation has been analysed by Flajolet and Odlyzko [10] in the case of simply generated trees; the bound follows. □

It is then easy to show that, as expected, the Pólya terms are essentially negligible:

Lemma 7 *There exist $r_0 > 0$ and $\theta_0 > 0$ small enough such that, for $z \in \mathcal{S}(r_0, \theta_0)$ and for all $h \le N$, one has $|e_h(z^2)|/|e_h(z)^2| \le \min\left\{12 \cdot (2\rho)^{h+1}, \frac{1}{2}\right\}$.*

Finally, we prove the main result of this section. The proof follows the lines of the analogous statement [10, Proposition 4], where the iteration is "pure", but now needs to control the effect of the Pólya terms, using Lemma 7.

Proposition 2 (Convergence in a "sandclock" around ρ**)** *There exist constants* r_0 *and* θ_0 *such that the sequence* $\{e_h(z), h \geq 0\}$ *converges to zero for* $z \in \mathcal{S}(r_0, \theta_0)$*, where the "sandclock"* $\mathcal{S}(r_0, \theta_0)$ *is defined by* (10).

Proof: We prove that for $h = N \equiv N(z)$ defined in Lemma 6, e_N satisfies the convergence criterion of Lemma 3. For this purpose, we use the alternative recurrence stated in Lemma 4

$$\frac{y^h}{e_h} = \frac{1}{2y}\frac{1-y^h}{1-y} + \underbrace{\frac{1}{e_0} - \sum_{i=0}^{h-1} \frac{y^{i-1} e_i(z^2)}{2e_i^2}}_{A_h} + \underbrace{\sum_{i=1}^{h-1} \frac{y^{i-2} e_i}{4} \left[1 - \frac{e_i(z^2)}{e_i^2}\right]^2 \left(1 - \frac{e_i}{2y}\left[1 - \frac{e_i(z^2)}{e_i^2}\right]\right)^{-1}}_{B_h} \tag{18}$$

and devise an asymptotic lower bound for the right-hand side. First observe that we can indeed use the relation, since by Lemmas 6 and 7, for all $i = 0, \ldots, N$, the denominators do not vanish.

Write $1 - y(z) = \epsilon e^{it}$. Obtaining the claim reduces to proving properties for small $\epsilon > 0$ and t close to $-\pi/4$. The following expansions are valid uniformly for $t \in [-\pi/4 - \delta, -\pi/4 + \delta]$ with $0 < \delta < \pi/4$ when $\epsilon \to 0$:

$$\left\{ \begin{array}{rcl} 1 - |y| & = & \epsilon \cos t + O(\epsilon^2), \\ \arg(y) & = & -\epsilon \sin t + O(\epsilon^2), \end{array} \right. \quad \text{and} \quad \left\{ \begin{array}{rcl} N(z) & = & -\varphi/(\epsilon \sin t) + O(1) \\ 1 - |y|^N & = & 1 - e^{\varphi \cdot \cot t} + O(\epsilon), \end{array} \right. \tag{19}$$

where $\varphi := \arccos(1/4)$. The first term of the right-hand side of (18) brings the main contribution:

$$\left| \frac{1}{2y} \frac{1 - y^N}{1 - y} \right| = \frac{1}{2|y|} \cdot \frac{|1 - y^N|}{|1 - y|} \geq \frac{1}{2} \frac{1 - |y|^N}{|1 - y|} = \frac{1 - e^{\varphi \cdot \cot t}}{2\epsilon} + O(1),$$

as $\epsilon \to 0$. On the other hand, we have $|A_N| = O(1)$ for, by Lemma 7, the summands decrease geometrically. The second error term B_N appearing in (18) can be bounded by splitting the sum into two at $h = K$. Let $\nu > 0$. By Lemma 7, there is $K \geq 6$ large enough that for all i satisfying $K \leq i \leq N$, $|e_i(z^2)|/|e_i(z)^2| \leq \nu$. Then, by Lemma 6, and for ϵ small enough, we have

$$|B_N| \leq \frac{K/2}{4} \cdot \frac{(3/2)^2}{1 - \frac{1/2}{2-2\epsilon} \cdot (3/2)} + \frac{1/5}{4} \cdot \frac{(1+\nu)^2}{1 - \frac{1/5}{2-2\epsilon}(1+\nu)} \frac{1 - |y|^N}{1 - |y|} < \frac{K}{2} + \frac{3}{50} \cdot \frac{1 - |y|^N}{1 - |y|}, \tag{20}$$

upon choosing $\nu = 1/100$. It follows that

$$\frac{|y^N|}{|e_N|} \geq \frac{1 - e^{\varphi \cdot \cot t}}{\epsilon} \left(\frac{1}{2} - \frac{3}{50 \cos t} \right) + O(1) > \frac{2}{5} \cdot \frac{1 - e^{\varphi \cdot \cot t}}{\epsilon}, \tag{21}$$

for all $z \in \mathcal{D}$ such that $\epsilon < \epsilon_0$ and $|t - \pi/4| < \delta_0$, as soon as both ϵ_0 and δ_0 are small enough.

We can now focus on the criterion for convergence (Lemma 3) and the conditions in (11) for $m = N$. From (19) and (21) we have for $\epsilon > 0$ small enough,

$$|y| + \frac{|e_N|}{2} \leq 1 - \epsilon \cdot \left(\cos t - \frac{5}{4} \cdot \frac{e^{\varphi \cdot \cot t}}{1 - e^{\varphi \cdot \cot t}} \right) + O(\epsilon^2).$$

A simple verification shows that the coefficient of ϵ above is at most $-1/4$ for all t close enough to $-\pi/4$. Thus, for all $\epsilon > 0$ small enough,

$$|e_N| \leq \frac{5}{2} \cdot \frac{\epsilon \cdot e^{\varphi \cot t}}{1 - e^{\varphi \cdot \cot t}} =: \alpha \qquad \text{and} \qquad |y| + \frac{|e_N|}{2} < 1 - \frac{\epsilon}{4} =: \beta. \tag{22}$$

Also, by (19), we have $(|z|^2/\rho)^N = o(\alpha(1-\beta))$. Hence the criterion for convergence of Lemma 3 is satisfied with the values of α and β specified in (22), as soon as ϵ is small enough. □

5 Main approximation

In this section, we estimate $e_h(z)$ around the singularity.

Proposition 3 (Main estimate for e_h in a sandclock around ρ) *There exist r_1, θ_1 and K, K' such that for all $z \in \mathcal{S}(r_1, \theta_1)$ and all $h \geq 1$,*

$$\frac{y^h}{e_h} = \frac{1}{2} \cdot \frac{1 - y^h}{1 - y} + R_h(z) \qquad \textit{where} \qquad |R_h(z)| \leq K \min\left\{\log\frac{1}{1 - |y|}, \log(1 + h)\right\}.$$

Furthermore, $|R_h - R_{h+1}| \leq K'/h$.

The main idea is to obtain a better control on the error term using bounds extending those obtained in Section 4 for $h > N$, knowing the *a priori* information that e_h converges. The proof of Proposition 3 also requires the bounds to be *uniform* in the distance to the singularity $|z - \rho|$ and in h.

Lemma 8 (Uniform lower bound for $|e_h|$) *For any $\delta \in (0,1)$, there exist constants $r_1, \theta_1 > 0$ such that if $z \in \mathcal{S}(r_1, \theta_1)$, then for all $h \geq 0$, one has $|e_h(z)| \geq (1-\delta)^{h+2}/(2(h+1))$.*

Proof: Let $\delta \in (0,1)$. For r small enough, we have $|y| > 1 - \delta/2$ provided $r := |z - \rho| < r_0$. Then, by Lemma 6, the desired lower bound is clear for $h \leq N$. So we now assume that $h > N$. The recurrence relation (6) implies that

$$|e_{h+1}| \geq |y||e_h|\left(1 - \frac{|e_h|}{2|y|}\right) - \frac{|e_h(z^2)|}{2}.$$

By Lemma 3, $|e_h|$ is decreasing for $h \geq N$, when the criterion is satisfied. So, for $z \in \mathcal{S}(r_0, \theta_0)$ as in Proposition 2, we have

$$|e_{h+1}| \geq |y|\left(1 - \frac{|e_N|}{2|y|}\right) \cdot |e_h| - \frac{|e_h(z^2)|}{2}.$$

However, we have set N in such a way that $|y| + |e_N|/2 < 1$, and by Lemma 2, we have $|e_{h+1}| \geq (1-\delta)|e_h| - (\rho + r)^h$. Routine verifications then show, for r small enough, that $1 - \delta > \rho + r$, the negative contribution decreases fast enough so that $|e_h|$ remains bounded from below as desired. □

We can now proceed with the proof of a uniform upper bound for $|e_h|$ when $z \in \mathcal{S}(r_1, \theta_1)$.

Lemma 9 (Uniform upper bound for $|e_h(z)|$ around ρ) *There exist constants c_1, r_1 and θ_1 such that, for any $h \geq 1$, and $z \in \mathcal{S}(r_1, \theta_1)$, we have $|e_h(z)| \leq c_1/h$.*

Proof: Write $1 - y = \epsilon e^{it}$ for some $\epsilon > 0$ and t. It suffices to prove that the result holds for all such z provided ϵ is small enough and t close enough to $-\pi/4$. We use again (18). Then, we have,

$$\left| \frac{1}{2y} \frac{1 - y^h}{1 - y} \right| = \frac{1}{2|y|} \cdot \frac{|1 - y^h|}{|1 - y|} \geq \frac{1 - |y|^h}{2|1 - y|} = \frac{1 - |y|^h}{2\epsilon}. \tag{23}$$

The error terms A_h and B_h are bounded as in the proof of Proposition 2 and we obtain, for all $h \geq 0$ and $\epsilon > 0$ small enough, $|A_h| \leq K_1$ and $|B_h| \leq K_2 + \frac{1}{5}(1 - |y|^h)/(1 - |y|)$[(ii)]. One then sees that, for all $h \geq 0$, and all t close enough to $-\pi/4$,

$$\frac{|y^h|}{|e_h|} \geq \frac{1 - |y|^h}{\epsilon} \left(\frac{1}{2} - \frac{1}{5 \cos t} \right) - K_3 > \frac{1 - |y|^h}{5\epsilon} - K_3 \qquad \text{and} \qquad |e_h| \leq 10\epsilon \frac{|y|^h}{1 - |y|^h}. \tag{24}$$

For h not too large, $|y|^h$ decreases at least linearly in h, and one can show that $|y|^h \leq 1 - \delta h \epsilon$, for some small $\delta > 0$ and all $0 \leq h \leq N$ as long as $|t + \pi/4|$ is small enough. Equation (24) then implies that $|e_h| \leq K_4/h$, for all $h \leq N$.

On the other hand, by (19), if $h \geq N$ then the factor $|y|^h$, ensures the desired decreasing behaviour. Indeed, for ϵ small enough and t close enough to $-\pi/4$, one has $|y| < (1-\epsilon/2)$ and $|e_h| \leq K_5\epsilon(1-\epsilon/2)^h$. The maximum of the right-hand side above is obtained for $\epsilon = 2/(h + 1)$, which implies that $|e_h| \leq 2K_5/(h + 1)$ for $h \geq N$. This completes the proof. □

Proof of Proposition 3: The proof consists in using Lemma 9 above to bound the error terms in (18) for $z \in \mathcal{S}(r_2, \theta_2)$, with $r_2 = \min\{r_0, r_1\}$ and $\theta_2 = \min\{\theta_0, \theta_1\}$. For some constants c_2 and c_3, we have

$$|A_h| + |B_h| \leq \frac{11}{(1 - 2\rho)^3} + c_2 \left(1 + \sum_{i=1}^{h} \frac{|y^i|}{i} \right) \leq c_3 \min \left\{ \log \left(\frac{1}{1 - |y|} \right), 1 + \log h \right\}.$$

Also, since A_h and B_h are partial sums, $R_h - R_{h+1}$ only contains one summand, which is easily seen to be $O(1/h)$ uniformly by Lemmas 8 and 9. □

6 Asymptotic analysis and distribution estimates

The basis of our estimates relative to the distribution of height is Proposition 3 in conjunction with Cauchy's coefficient formula (8) where γ is a contour comprised of the arc of an outer circle of radius larger than ρ (and interior to the cardioid-shaped region, where $|y| < 1$) and a set of two connecting segments passing through the singularity ρ (Figure 1). In addition, it proves useful, in order to garantee well-defined determinations of square roots, to think of the two segments as in fact joined by an infinitesimal arc of a circle that passes to the *left* of the singularity ρ. The strategy just described belongs to the general orbit of singularity analysis methods.

Theorem 1 (Limit law of height) *The height H_n of a random tree taken uniformly from $\mathcal{Y}_n$ admits a limiting theta distribution: for any fixed $x > 0$, there holds*

$$\lim_{n\to\infty} \mathbb{P}(H_n \geq x\sqrt{n}) = \sum_{k\geq 1} (k^2\lambda^2x^2 - 2)e^{-k^2\lambda^2x^2/4}, \qquad \lambda := \sqrt{2\rho + 2\rho^2 y'(\rho^2)}.$$

[(ii)] In what follows, we use generically $K, K_1, \ldots$ to denote positive absolute constants, not necessarily of the same value at different occurrences.

Proof: The integration contour in Cauchy's formula (8) is $\gamma = \gamma_1 \cup \gamma_2 \cup \gamma_3$. There, γ_1 is the segment lying in the half-plane $\Im(z) \geq 0$, γ_2 is the complex-conjugate image of γ_1, and γ_3 is the outer circular arc. For the radius of the latter circle, we adopt $r_n = \rho(1 + \log^2 n/n)$. We assume that γ_3 lies in a legal tube (granted by Proposition 1) and that γ_1 and γ_2 are in an overlapping sandclock such that Proposition 3 applies. We set, for some $\theta_1 > 0 : \gamma_1 = \bar{\gamma}_2 = \left\{\rho\big(1 + xe^{i\pi/2+i\theta_1}\big) : x \in [0, \delta_n]\right\}$, with δ_n such that the rectilinear portions γ_1 and γ_2 connect with the outer circle γ_3. So, we have $\delta_n \sim \log^2 n/(n \sin\theta_1)$.

Outer circular arc (γ_3). By Proposition 1, we have $e_h(z) \to 0$ uniformly on γ_3 as $h \to \infty$. In particular, all moduli $|e_h(z)|$ are bounded by an absolute constant K. On the other hand the Cauchy kernel z^{-n} is small on the outer part of the contour, so that

$$\left|\int_{\gamma_3} e_h(z)\,\frac{dz}{z^{n+1}}\right| < K_1\rho^{-n}\exp\left(-\log^2 n\right). \tag{25}$$

This contribution is thus exponentially small compared to y_n.

Rectilinear parts (γ_1, γ_2). Our objective is to replace e_h by the simpler quantity

$$\widehat{e}_h(z) \equiv \widehat{e}_h := 2\frac{1-y}{1-y^h}y^h, \tag{26}$$

as suggested by Proposition 3. Along γ_1, γ_2, the singular expansion of $y(z)$ applies, so that $1 - y = O((\log n)/n^{1/2})$ and the error term $R_h(z)$ is $O(\log n)$. There results that $(1 - y^h)/(1 - y)$ is always at least as large in modulus as $K_2\sqrt{n}/\log n$ (by a study of the variation of $|1 - e^{-h\tau}|/|1 - e^{-\tau}|$), and we have

$$\frac{y^h}{e_h} = \frac{y^h}{\widehat{e}_h}\left(1 + O\left(\frac{\log^2 n}{\sqrt{n}}\right)\right). \tag{27}$$

It proves convenient to define the following approximation for $e_{h,n}$:

$$E(h,n) := \frac{1}{2i\pi}\int_{\gamma_1\cup\gamma_2} \widehat{e}_h\,\frac{dz}{z^{n+1}}, \tag{28}$$

and effect the change of variables $z = \rho\left(1 - \frac{t}{n}\right)$. The quantity t then varies from $-in\delta_n e^{-i\theta_1}$, loops to the right of the origin, then moves away to $in\delta_n e^{i\theta_1}$. With the singular expansion of $y(z)$ as in (2), we have on γ_1, γ_2,

$$z^{-n} = \rho^{-n}e^t\left(1 + O\left(\log^4 n/n\right)\right), \qquad y(z) = 1 - \lambda\sqrt{t/n} + O\left(t/n\right). \tag{29}$$

and, with $h = x\sqrt{n}$ and $|t| \leq K_2\log^2 n$:

$$y^h = \exp(-\lambda x\sqrt{t})\cdot\left(1 + O\left(\log^2 n/\sqrt{n}\right)\right). \tag{30}$$

We also find[(iii)], for the range of values of t corresponding to γ_1, γ_2:

$$\frac{1-y^h}{1-y} = \left[\sqrt{n}\cdot\frac{1-\exp(-\lambda x\sqrt{t})}{\lambda\sqrt{t}}\right]\left(1 + O\left(\frac{\log^\star n}{\sqrt{n}}\right)\right). \tag{31}$$

[(iii)] The expression $\log^\star n$ represents an unspecified positive power of $\log n$.

The approximations (29), (30), and (31) motivate considering, as an approximation of $E(h,n)$ in (28), the contour integral

$$J(X) := \frac{1}{2i\pi}\int_{\mathcal{L}} \frac{\exp(-X\sqrt{t})}{1-\exp(-X\sqrt{t})}\sqrt{t}e^t\,dt = \frac{1}{2i\pi}\sum_{k\geq 1}\int_{\mathcal{L}} \exp(-kX\sqrt{t})\sqrt{t}e^t\,dt, \tag{32}$$

where $\mathcal{L}$ is the image of $\gamma_1 \cup \gamma_2$ in the change of variable. We can now make $J(X)$ explicit. Each integral on the right side can be evaluated by the change of variables $w = i\sqrt{t}$, $t = -w^2$. By completing the square and flattening the image contour $\mathcal{L}$ onto the real line, we obtain:

$$J(X) = \frac{1}{4\sqrt{\pi}}\sum_{k\geq 1} e^{-k^2X^2/4}(k^2X^2-2). \tag{33}$$

Error management. It can be checked that the replacements: $e_h \mapsto \widehat{e}_h$, $y \mapsto 1-\lambda\sqrt{t/n}$, and $y^h \mapsto \exp(-\lambda x\sqrt{t})$ only entail error terms of order $\rho^{-n}n^{-2}\log^\star n$, which implies, for $h = x\sqrt{n}$, for $h = x\sqrt{n}$:

$$e_{h,n} = 2\lambda\rho^{-n}n^{-3/2}J(\lambda x) + O\left(\rho^{-n}\log^\star n/n^2\right).$$

The explicit form of $J(X)$ in (33) and the asymptotic form of y_n (Lemma 1), yield the statement. □

Theorem 2 (Local limit law of height) *The height H_n of a random tree taken uniformly from $\mathcal{Y}_n$ admits a local limiting distribution: for x in a compact set of $\mathbb{R}_{>0}$ and $h = x\sqrt{n}$ an integer, there holds*

$$\mathbb{P}(H_n = h) \sim \frac{1}{2x\sqrt{n}}\sum_{k\geq 1}(k^4\lambda^4x^4 - 6k^2\lambda^2x^2)e^{-k^2\lambda^2x^2/4}.$$

Proof: Proceeding like in the proof of Theorem 1, we can justify approximating the number of trees of height exactly h by the integral

$$\frac{1}{2i\pi}\int_\gamma (\widehat{e}_h - \widehat{e}_{h+1})\,\frac{dz}{z^{n+1}}, \qquad \text{where} \qquad \widehat{e}_h - \widehat{e}_{h+1} = 2y^h\frac{(1-y)^2}{(1-y^h)(1-y^{h+1})}.$$

The approximations (30) and (31) then motivate consideration of the integral

$$J_1(X) := \frac{1}{2i\pi}\int_{\mathcal{L}} \frac{\exp(-X\sqrt{t})}{(1-\exp(-X\sqrt{t}))^2}te^t\,dt,$$

and one finds (with the auxiliary estimate $R_h - R_{h+1} = O((\log^\star n)/\sqrt{n})$ provided by Proposition 3):

$$y_{n,h} - y_{n,h+1} = 2\lambda^2\rho^{-n}n^{-2}J_1(\lambda x) + O\left(\rho^{-n}\frac{\log^\star n}{n^{5/2}}\right).$$

On the other hand, differentiation under the integral sign yields $J_1(X) = -J'(X)$, which proves the statement. □

Revisiting the proof of Theorems 1 and 2 shows that one can allow x to become either small or large, albeit to a limited extent. Indeed, it can be checked, for instance, that allowing x to get as large as $O(\sqrt{\log n})$ only introduces extra powers of $\log n$ in error estimates. However, such extensions are limited by the fact that the main theta term eventually becomes smaller than the error term. We state (compare with [12, Th. 1.1]):

Theorem 3 (Moderate deviations) *There exist constants $A, B, C > 0$ such that for $h = x\sqrt{n}$ with $A/\sqrt{\log n} \le x \le A\sqrt{\log n}$ and n large enough, there holds*

$$\left| \mathbb{P}(H_n \ge x\sqrt{n}) - \sum_{k\ge 1}(k^2\lambda^2x^2 - 2)e^{-k^2\lambda^2x^2/4} \right| \le \frac{C}{n^B}. \tag{34}$$

In particular, if $x \to \infty$ in such a way that $x \le A\sqrt{\log n}$, then, uniformly,

$$\mathbb{P}(H_n \ge x\sqrt{n}) \sim \lambda^2 x^2 e^{-\lambda^2x^2/4}.$$

Similar estimates hold for the local law.

These estimates can be supplemented by (very) large deviation estimates in the style of [12, Th.1.4]: it suffices to make use of the fact that e_h is bounded from above by a large power and optimize on $r \in (0, \rho)$ the saddle-point bound

$$e_{h,n} \le \frac{e_h(r)}{r^n}, \qquad 0 < r \le \rho.$$

The probability of a linearly height is then exponentially small:

Theorem 4 (Very large deviations) *There exists a continuous increasing function $I(\xi)$ satisfying $I(\xi) > 0$ for $0 < \xi \le 1$, such that, given any fixed $\delta > 0$, one has for $h = \xi n$, and $\delta < \xi < 1 - \delta$,*

$$\mathbb{P}(H_n \ge \xi n) \le K n^{3/2} e^{-nI(\xi)},$$

where K depends on δ.

Finally, the approximation of e_h by $\widehat{e}_h$ in (26) is good enough to grant us access to moments (cf also [10]). The problem reduces to estimating generating functions of the form

$$M_r(z) = 2(1-y)^2 \sum_{h\ge 1} h^r \frac{y^h}{(1-y^h)^2},$$

which are accessible to the Mellin transform technology [13], upon setting $y = e^{-\tau}$.

Theorem 5 (Moments of height) *Let $r \ge 1$. The* rth *moment of the height H_n satisfies*

$$\mathbb{E}[H_n] \sim \frac{2}{\lambda}\sqrt{\pi n} \qquad \text{and} \qquad \mathbb{E}[H_n^r] \sim r(r-1)\zeta(r)\Gamma(r/2)\left(\frac{2}{\lambda}\right)^r n^{r/2}, \quad r \ge 2.$$

References

[1] D. J. Aldous. The random walk construction of uniform spanning trees and uniform labelled trees. *SIAM Journal on Discrete Mathematics*, 3(4):450–465, 1990.

[2] D. J. Aldous. The continuum random tree I. *The Annals of Probability*, 19:1–28, 1991.

[3] N. Broutin and P. Flajolet. The distribution of the height in random binary unlabelled trees. In preparation, 2008.

[4] P. Chassaing and J.-F. Marckert. Parking functions, empirical processes, and the width of rooted labeled trees. *Electronic Journal of Combinatorics*, 8(1):Research Paper 14, 19 pp. (electronic), 2001.

[5] P. Chassaing, J.-F. Marckert, and M. Yor. The height and width of simple trees. In *Mathematics and computer science (Versailles, 2000)*, Trends Math., pages 17–30. Birkhäuser Verlag, Basel, 2000.

[6] N. G. de Bruijn. *Asymptotic Methods in Analysis*. Dover, 1981. A reprint of the third North Holland edition, 1970 (first edition, 1958).

[7] N. G. de Bruijn, D. E. Knuth, and S. O. Rice. The average height of planted plane trees. In R. C. Read, editor, *Graph Theory and Computing*, pages 15–22. Academic Press, 1972.

[8] M. Drmota and B. Gittenberger. The shape of unlabeled rooted random trees. Technical report, Technical University of Vienna, 2008.

[9] S. Finch. *Mathematical Constants*. Cambridge University Press, New-York, 2003.

[10] P. Flajolet and A. M. Odlyzko. The average height of binary trees and other simple trees. *Journal of Computer and System Sciences*, 25:171–213, 1982.

[11] P. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 2008. URL `http://algo.inria.fr/flajolet`. Free electronic version of a book to be published in 2008; 808p.+xii, available electronically from the authors' home pages.

[12] P. Flajolet, Z. Gao, A. Odlyzko, and B. Richmond. The distribution of heights of binary trees and other simple trees. *Combinatorics, Probability and Computing*, 2:145–156, 1993.

[13] P. Flajolet, X. Gourdon, and P. Dumas. Mellin transforms and asymptotics: Harmonic sums. *Theoretical Computer Science*, 144(1–2):3–58, June 1995.

[14] P. Flajolet, P. Grabner, P. Kirschenhofer, and H. Prodinger. On Ramanujan's Q–function. *Journal of Computational and Applied Mathematics*, 58(1):103–116, Mar. 1995.

[15] B. Gittenberger. The profile of unlabeled trees. *Discrete Mathematics & Theoretical Computer Science Proceedings*, AD:125–138, 2005.

[16] F. Harary and E. M. Palmer. *Graphical Enumeration*. Academic Press, 1973.

[17] D. P. Kennedy. The Galton–Watson process conditioned on the total progeny. *Journal of Applied Probability*, 12(4):800–806, dec 1975.

[18] D. P. Kennedy. The distribution of the maximum brownian excursion. *Journal of Applied Probability*, 13(2): 371–376, jun 1976.

[19] V. F. Kolchin. *Random Mappings*. Optimization Software Inc., New York, 1986. Translated from *Slučajnye Otobraženija*, Nauka, Moscow, 1984.

[20] J. F. Le Gall. Random trees and applications. *Probability Surveys*, 2:245–311, 2005.

[21] J.-F. Marckert and G. Miermont. A limit theorem for rescaled binary rooted non plane trees. In preparation, 2008.

[22] A. Meir and J. W. Moon. On the altitude of nodes in random trees. *Canadian Journal of Mathematics*, 30: 997–1015, 1978.

[23] R. Otter. The number of trees. *Annals of Mathematics*, 49(3):583–599, 1948.

[24] G. Pólya. Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und chemische Verbindungen. *Acta Mathematica*, 68:145–254, 1937.

[25] A. Rényi and G. Szekeres. On the height of trees. *Australian Journal of Mathematics*, 7:497–507, 1967.