THE DISTRIBUTION OF HEIGHT AND DIAMETER IN RANDOM NON-PLANE BINARY TREES

NICOLAS BROUTIN AND PHILIPPE FLAJOLET

ABSTRACT. This study is dedicated to precise distributional analyses of the height of non-plane unlabelled binary trees ("Otter trees"), when trees of a given size are taken with equal likelihood. The height of a rooted tree of size n is proved to admit a limiting theta distribution, both in a central and local sense, as well as obey moderate as well as large deviations estimates. The approximations obtained for height also yield the limiting distribution of the diameter of unrooted trees. The proofs rely on a precise analysis, in the complex plane and near singularities, of generating functions associated with trees of bounded height.

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*, and 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 article, we study the (random) *height* H_n of a tree sampled uniformly from $\mathcal{Y}_n$.

Most of the results concerning random trees of fixed size are relative to the situation where one can *distinguish* the neighbours of a node, either by their labels (labelled trees), or by the order induced on the progeny through an emebdding in the plane (plane trees); see the reference books [14, 19] and the discussion by Aldous [3] who globally refers to these as *"ordered trees"*. In this range of models, Meir and Moon [33] determined that the depth of nodes is typically $O(\sqrt{n})$ for all *"simple varieties"* of trees, which are determined by restricting in an arbitrary way the collection of allowed node degrees. Regarding height, a few special cases were studied early: Rényi and Szekeres [38] proved in particular that the average height of *labelled* non-plane trees of size n is asymptotic to $2\sqrt{\pi n}$; De Bruijn, Knuth, and Rice [12] dealt with *plane* trees and showed that the average height is equivalent to $\sqrt{\pi n}$ as $n \to \infty$. Eventually, Flajolet and Odlyzko [18] developed an approach for height that encompasses all simple varieties of trees; see also [20] for additional results.

Under such models with distinguishable neighbourhoods, trees of a fixed size n may be seen as Galton–Watson processes (branching processes) conditioned on the size being n, see [1, 26, 28], and there are natural random walks associated to various tree traversals. Accordingly, probabilistic techniques have been successfully applied to quantify tree height and width [8, 9], based on Brownian excursion. An important probabilistic approach consists in establishing the existence of a continuous limit of suitably rescaled random trees of increasing sizes—one can then read off, to first asymptotic order at least, some of the limit parameters directly on the limiting object. The latter point of view has been adopted

by Aldous [2, 3, 4] in his definition of the *continuum random tree* (CRT): see the survey by Le Gall [30] for a recent account of probabilistic developments along these lines.

The case of trees (as are considered here) that have *indistinguishable neighbourhoods* is essentially different. Such trees cannot be generated by a branching process conditioned by size and no direct random walk approach appears to be possible, due to the inherent presence of symmetries. (An analysis of such symmetries otherwise occurs in the recent article [6].) The analysis of unlabelled non-plane trees finds its origins in the works of Pólya [36] and Otter [35]. However, these authors mostly focused on enumeration—the problem of characterizing typical parameters of these random trees remained largely untouched. Recently, in an independent study, Drmota and Gittenberger [15] have examined the profile of *"general" trees* (where all degrees are allowed) and shown that the joint distribution of the number of nodes at a 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.

The foregoing discussion suggests that, although there is no clear *exact* reduction of unlabelled non-plane trees to random walks, such trees largely behave like simply generated families of ordered trees. In particular, it suggests that the rescaled height $H_n/\sqrt{n}$ is likely to admit a limit distribution of the theta-function type [16, 18, 27, 38]. We shall prove that such is indeed the case for *non-plane binary 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 moments (Theorem 5), see §5. Equiped with solid analytic estimates regarding height, we can then proceed to characterize the *diameter of unrooted trees* in §6, this both in a local and central form (Theorems 6 and 7). Some *a posteriori* observations that complete the picture are offered in our Condusion section, §7.

A preliminary investigation of the distribution of height in rooted trees is reported in the extended abstract [7]. Our interest in this range of problems initially arose from questions of Jean-François Marckert and Grégory Miermont [32], in their endeavour to extend the probabilistic methods of Aldous to non-plane trees and develop corresponding continuous models—we are indebted to them for being at the origin of the present study.

1. **Trees and generating functions**

Tree enumeration. Our approach is entirely based on *generating functions*. The class $\mathcal{Y}$ of (non-plane, unlabelled, rooted) binary 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 of Sloane's *On-line Encyclopedia of Integer Sequences*. The trees of $\mathcal{Y}$ with size at most 6 are shown in Figure 1.

A binary tree is either an external node or a root appended to an unordered pair of two (not necessarily distinct) binary trees. In the language of analytic combinatorics [19], this corresponds to the (recursive) specification

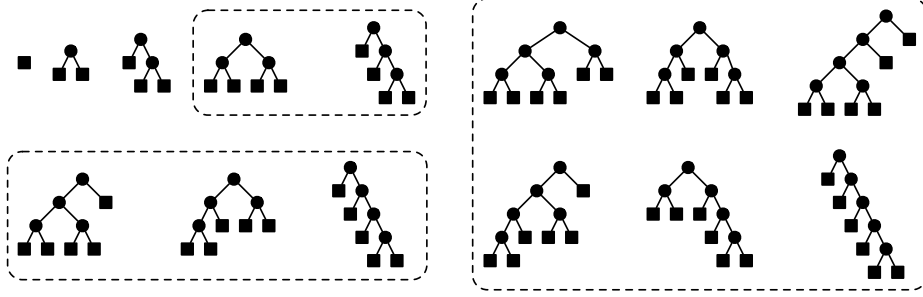$$\mathcal{Y} = \mathcal{Z} + \mathrm{MSet}_2(\mathcal{Y}),$$

Figure 1. The binary unlabelled trees of size less than six.

where $\mathcal{Z}$ represents a generic atom (of size 1) and MSet_2 forms multisets of two elements. The basic functional equation

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

closely related to the early works of Pólya (1937; see [36, 37], and first studied by Otter (1948; see [35]), follows from fundamental principles of combinatorial enumeration [19, 25]. The term $\frac{1}{2}y(z^2)$ accounts for potential symmetries—hereafter, we refer to such terms involving functions of $z^2, z^3, \ldots$, as *Pólya terms*. According to the general theory of analytic combinatorics, we shall operate in an essential manner with properties of generating functions in the *complex plane*. The following lemma is classical but we sketch a proof, as its ingredients are needed throughout our work.

Lemma 1 (Otter [35])**.** *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 $\rho + \frac{1}{2}y(\rho^2) = \frac{1}{2}$. As $z \to \rho^-$, the generating function $y(z)$ satisfies*

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

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

$$(3) \qquad y_n = \frac{\lambda}{2\sqrt{\pi}} \cdot n^{-3/2}\rho^{-n}\left(1 + O\left(\frac{1}{n}\right)\right),$$

Proof. The number of plane binary trees with n external nodes is given by the Catalan number $C_{n-1} = \frac{1}{n}\binom{2n-2}{n-1}$. The number of symmetries in a tree of size n being *a priori* between 1 and 2^{n-1}, one has the bounds

$$C_{n-1}2^{1-n} \leq y_n \leq C_{n-1}.$$

As it is well known, the Catalan numbers satisfy $C_n \sim \pi^{-1/2}4^n n^{-3/2}$, hence the bounds $1/4 \leq \rho \leq 1/2$. It follows that $y(z^2)$ is analytic in a disc of radius $\sqrt{\rho}$, which properly contains $|z| \leq \rho$. Then, from (1), upon solving for $y(z)$, we obtain

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

which can only become singular when the argument of the square root vanishes. By Pringsheim's Theorem [19, p. 240], the value ρ is then the smallest positive solution of $2z + y(z^2) = 1$, corresponding to a simple root, and, at this point, we must have $y(\rho) = 1$,

given (4). This reasoning also justifies the singular expansion (2), which is seen to be valid in a Δ-domain [19, §VI.3], i.e., a domain of the form

(5) $$\{z \ : \ |z| < \rho + \epsilon, z \neq \rho, |\arg(z - \rho)| > \theta\} \qquad \epsilon, \theta > 0$$

that extends beyond the disc of convergence $|z| \leq \rho$.

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 [25, 35, 36] or singularity analysis [19]. □

Numerically, one finds [17, 19, 35]:

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

Height. In a tree, *height* is defined as the maximum number of edges along branches connecting the root to an external node. Let $y_{h,n}$ be the number of trees of size n and height *at most* h and let $y_h(z) = \sum_{n \geq 1} y_{h,n} z^n$ be the corresponding generating function. The arguments leading to (1) yield the fundamental recurrence

(6) $$y_{h+1}(z) = z + \frac{1}{2} y_h(z)^2 + \frac{1}{2} y_h(z^2), \qquad h \geq 0,$$

with initial value $y_0(z) = z$, and

(7) $$\begin{cases} y_1(z) & = & z + z^2, \qquad y_2(z) = z + z^2 + z^3 + z^4, \\ y_3(z) & = & z + z^2 + z^3 + 2z^4 + 2z^5 + 2z^6 + z^7. \end{cases}$$

A central rôle in what follows is played by the generating function of trees with height *exceeding* h:

$$e_h(z) \equiv \sum_{n \geq 1} e_{h,n} z^n := y(z) - y_h(z),$$

Then, a trite calculation shows that the $e_h(z)$ satisfy the main recurrence

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

on which our subsequent treatment of height is entirely based.

Analysis. The distribution of height is accessible by

(9) $$\mathbb{P} \{ H_n > h \} = \frac{y_n - y_{n,h}}{y_n} = \frac{e_{h,n}}{y_n},$$

where $e_{h,n} = [z^n] e_h(z)$. Lemma 1 provides an estimate for y_n, and we shall get a handle on the asymptotic properties of $e_{h,n}$ by means of Cauchy's coefficient formula,

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

upon choosing a suitable integration contour γ in (10), of the form commonly used in singularity analysis theory [19]; see Figure 2 below. This task necessitates first developing 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 $|z| = \rho$, as well as inside a *"sandclock"* anchored at ρ (see Figure 2).

Definition 1. *The* "tube" $\mathcal{T}(\mu, \eta)$ *of width* μ *and angle* η *is defined as*

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

The "sandclock" *of radius* r_0 *and angle* θ_0 *anchored at* ρ *is defined as*

(12) $$\mathcal{S}(r_0, \theta_0) := \{z : \quad |z - \rho| < r_0, \quad \pi/2 - \theta_0 < |\arg(z - \rho)| < \pi/2 + \theta_0\}.$$
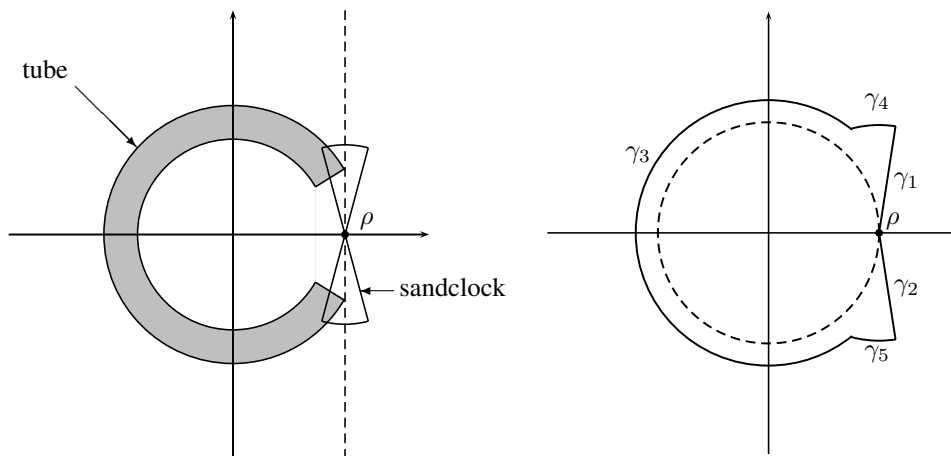
Figure 2. *Left*: the "tube" and "sandclock" regions. *Right*: the Hankel contour used to estimate $e_{h,n}$ (details are given in Figure 3).

Strategy. 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, are comparatively easy: they form the subject of Section 2. In particular, Proposition 1 states that we can always find thinner and thinner tubes that come arbitrarily close to the singularity ρ and where the convergence $y_h \to y, e_h \to 0$ is ensured. The bulk of the technical work is relative to the sandclock, in Section 3, where Proposition 2 grants us the existence of a suitable sandclock for convergence. We can then develop in Section 4 our main approximation, summarized by Proposition 3. Sections 2–4 closely follow the general strategy of the original paper [18]; however, nontrivial adaptations are needed, due to the presence of Pólya terms, so that the problem is no longer of a "pure" iteration type. We then reap the crop in Section 5, where theorems relative to the distribution of height are stated and proved— these appear to parallel the central and local limit laws known in the planar case [20]. A further reduction to the rooted tree case finally allows us to analyse the diameter of un-rooted trees in Section 6.

2. **Convergence away from the singularity in tubes**

Our aim in this section[1] is to extend the domain where e_h is analytic beyond the disc of convergence $|z| \leq \rho$, when z stays in a "tube" $\mathcal{T}(\mu, \eta)$ as defined in (11) 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 the e_h converge to 0, equivalently, $y_h \to y$, in the closed disc of radius ρ (this property is the consequence of the $n^{-3/2}$ subexponential factor in the asymptotic form of y_n, which implies convergence of $y(\rho)$); second, a general criterion for convergence of the e_h to 0,

---

[1]In what follows, we freely omit the arguments of $y(z), e_h(z), y_h(z) \ldots$, whenever they are taken at z. (We reserve h for height and n for size, so that no ambiguity should arise: y_h means $y_h(z)$, whereas y_n invariably represents $[z^n]y(z)$.)

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 $|z| \leq \rho$, and $h \geq 1$, one has*

$$|e_h(z)| \leq \frac{1}{\sqrt{h}} \left(\frac{|z|}{\rho} \right)^h.$$

Proof. To have height at least h, a tree needs at least $h + 1$ nodes, so that $|e_h(z)| \leq \sum_{n>h} y_n |z|^n$. We first note an easy numerical refinement of (3), namely, $y_n \leq \frac{1}{2} \rho^{-n} n^{-3/2}$, obtained by combining the first few exact values of y_n with the asymptotic estimate (3). (See [22] for a detailed proof strategy in the case of a similar but harder problem.) This implies

$$|e_h(z)| \leq \frac{1}{2} \left(\frac{|z|}{\rho} \right)^h \sum_{n>h} \frac{1}{n^{3/2}} \leq \frac{1}{2} \left(\frac{|z|}{\rho} \right)^h \int_h^\infty \frac{dt}{t^{3/2}} = \left(\frac{|z|}{\rho} \right)^h \frac{1}{\sqrt{h}},$$

and the statement results. $\qquad\square$

We now devise a criterion for the convergence of $e_h(z)$ to zero. This criterion, adapted from [18, 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 3).

Lemma 3 (Convergence criterion). *Define the domain[2]*

(13) $$\mathcal{D} := \{ z : |y| < 1 \}.$$

Assume that z satisfies the conditions $z \in \mathcal{D}$ and $|z| < \sqrt{\rho}$. The sequence $\{ |e_h(z)|, h \geq 0 \}$ converges to 0 if and only if there exist an integer $m \geq 1$ and real numbers $\alpha, \beta \in (0, 1)$, such that the following three conditions are simultaneously met:

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

Furthermore, if (14) holds then, for some constant C and $\beta_0 \in (0, 1)$, one has the geometric convergence

(15) $$|e_h| \leq Ch\beta_0^h,$$

for all $h \geq m$.

Proof. (i) *Convergence implies that (14) is satisfied, for some m.* Assume that $z \in \mathcal{D}$, $|z| < \sqrt{\rho}$, and $e_h(z) \to 0$ as $h \to \infty$. Then choose β such that $|y| < \beta < 1$. This gives a possible value for α, say, $\alpha = (\beta - |y|)/2$. Choose m_0 such that, for all $\mu > m_0$, one has $|e_\mu| < \alpha$; then choose m_1 large enough, so that the third condition of (14) is satisfied. The three conditions of (14) are now satisfied by taking $m = \max(m_0, m_1)$.

(ii) *Condition (14) implies convergence and the bound (15).* Conversely, assume the three conditions in (14), for some value m. Then, they also hold for $m + 1$. Indeed, recalling (8), we see that, for any $h \geq 1$,

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

---

[2]This domain will sometimes be referred to as the "cardiod-like" domain, as it contains the unit disc punctured at ρ (Proposition 1) and has a cusp at $z = \rho$, associated to the square root singularity of $y(z)$ at ρ.

where the Pólya term involving $|e_h(z^2)|$ has been bounded using Lemma 2. The hypotheses of (14) together with (16) above taken at $h = m$, yield the inequality $|e_{m+1}| < \alpha$. So, once the conditions (14) hold for some m, they hold for $m + 1$; hence, for all $h \geq m$.

The fact that, under these conditions, there is convergence, $e_h \to 0$, now results from unfolding the recurrence (8): we find, for all $h \geq m$,

$$|e_{h+1}| \leq \beta^{h-m+1}|e_m| + \sum_{i=0}^{h-m} \beta^i \left(\frac{|z|^2}{\rho}\right)^{h-i} \leq \beta^{h-m+1}|e_m| + h \max\left\{\beta, \left(\frac{|z|^2}{\rho}\right)\right\}^h,$$

where Lemma 2 has been used again to bound the Pólya term. The additional assertion that $|e_h| \leq Ch\beta_0^h$ in (15) finally follows from choosing $\beta_0 := \max(\beta, |z|^2/\rho)$. $\square$

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

Proposition 1 (Convergence in "tubes"). *For any, however small, 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. We thus start from a fixed η, assumed to be suitably small. If we exclude a small sector of opening angle 2η around the positive real axis, then the quantity,

$$\lambda_0 := \sup\left\{|y(z)|; \quad |z| = \rho, \ |\arg(z)| \geq \eta\right\},$$

satisfies $\lambda_0 < 1$: this results from the strong triangle inequality (see also the "Daffodil Lemma" of [19]) and the fact that $y(\rho e^{i\theta})$ is a continuous function of θ. (By the argument introduced in the proof of Lemma 1, the function $y(z)$ is analytic at all points of $|z| = \rho$, $z \neq \rho$, hence continuous.) Fix then ϵ by $\lambda_0 = 1 - 2\epsilon$. By continuity of y again, for each z on the circle of radius ρ satisfying $|\arg(z)| \geq \eta$, there exists a small *open* disc $\delta(z)$, centred at z and such that $|y(\zeta)| < 1 - \epsilon$ for all $\zeta \in \delta(z)$. From now on, we assume that the discs $\delta(z)$ are taken small enough, so that they are entirely contained in the larger disc $\{w \in \mathbb{C} : |w| < \sqrt{\rho}\}$.

We can then make use of the convergence criterion of Lemma 3, supplemented once more by a continuity argument. In the notations of (14), choose first $\alpha = \epsilon$, then $\beta = 1 - \epsilon/2$. For all sufficiently large m, say $m \geq \nu$, the last two conditions of (14) are satisfied. Then, since the $e_h(z)$ are analytic (hence continuous) at every point of the unit circle punctured at ρ, there exists, around each z on $|z| = \rho$ with $|\arg(z)| \geq \eta$, a small open disc $\delta_1(z) \subseteq \delta(z)$ and an integer $M(z)$ such that $|e_m| < \alpha$ for all $m \geq M(z)$. We may also freely assume that $M(z) > \nu$.

Finally, by *compactness* of the arc $\{\rho e^{i\theta}\}$ defined by $|\theta| \geq \eta$, there exists a covering of the arc by a *finite* collection of small discs, say $\{\delta_1(z_j)\}_{j=1}^r$. The union of these small discs must then contain a tube of angle η and width $\mu > 0$. By design, in this tube, all three conditions of the convergence criterion of Lemma 3 (Equation (14)) are now satisfied, with $m = \max_{j=1}^r M(z_j)$. $\square$

3. **Convergence near the singularity in a sandclock**

We now focus on the behaviour of $e_h(z)$ in a "sandclock" around the singularity. When z approaches ρ, the quantity $|y|$ is no longer bounded away from 1, so that the criterion for convergence obtained earlier (Lemma 3) cannot be used directly. We then need to proceed in two stages: first, we prove in Subsection 3.1 that, in a suitable sandclock, the initial terms decay "enough"; next, in Subsection 3.2, we establish the existence of a sandclock where convergence of the e_h to y is ensured—this is expressed by the main Proposition 2 below. We shall then be able to build upon these results in the next section and derive suitable singular approximations of the e_h *outside* of the original disc of convergence $|z| \leq \rho$ of $y(z)$, when z is near ρ.

Alternative recurrence. So far, we have operated with the main recurrence (8) relating the e_h, then applied some partial unfolding supplemented by simple continuity arguments. To proceed with our programme, we need to adapt a classical technique in the study of slowly convergent iterations near an indifferent fixed point [11, p. 153], which simply amounts to "taking inverses" and leads to a useful alternative form of the original recurrence.

Lemma 4 (Alternative recurrence). *Assume, for a value z, the conditions*

$$e_i(z) \neq 0 \qquad and \qquad e_i(z)\left[1 - e_i(z^2)/e_i(z)^2\right] \neq 2y, \quad for \ i = 0, \ldots, h-1.$$

Then, the following two recurrence relations hold

$$(17) \quad \frac{y^h}{e_h} = \frac{1}{e_0} + \frac{1}{2y}\sum_{i=0}^{h-1} y^i \left[1 - \frac{e_i(z^2)}{e_i^2}\right]\left(1 - \frac{e_i}{2y}\left[1 - \frac{e_i(z^2)}{e_i^2}\right]\right)^{-1}$$

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

The form (17) is referred to as the *simplified alternative recurrence*; the form (18) is the *extended alternative recurrence*.

Proof. Starting with the recurrence relation (8), 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* (cf also [18]). The identity $(1-u)^{-1} = 1 + u(1-u)^{-1}$ implies

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

Summing the terms of this equality for $i = 0, \ldots, h-1$ then yields the first version. The extended version follows from the expansion $(1-u)^{-1} = 1 + u + u^2(1-u)^{-1}$.   □

Lemma 4 is used to complete the proof of Lemma 5 below (see Equation (29)) and it serves as the starting point of the proof of Proposition 2 (see Equation (31)). It then proves central in establishing the main approximation of Proposition 3 in the next section. The interest of these alternative recurrences is that they relate the *inverse* $1/e_h$ to essentially polynomial forms in the previous e_i. In particular they *serve to convert lower bounds into upper bounds, and vice versa.*

3.1. **Initial behaviour of e_h.** We establish in this subsection (cf Lemma 6) that the quantities $|e_h(z)|$ first exhibit a decreasing behaviour for $h \leq N$, with some appropriate $N = N(z)$. At that point, $|e_N(z)|$ appears to be small enough to guarantee that the criterion of Lemma 3 becomes applicable, whence eventually the convergence $|e_h(z)| \to 0$ as $h \to \infty$ in a sandclock.

The following preparatory lemma serves to control the effect of Pólya terms, when z is close to ρ, so that $z' = z^2$ is close to ρ^2, well inside of the disc of radius ρ. It is evocative of the theory of iteration near an attractive fixed point (see, e.g., [34, Ch. 8]).

Lemma 5 (Smooth iteration for Pólya terms). *Fix $z_0 \in (0, \rho)$. There exists a constant $R_0 > 0$, dependent upon z_0, such that, for all $h \geq 0$, and for all z satisfying $|z - z_0| < R_0$, one has*

$$e_h(z) = C_h(z) \cdot y(z)^h,$$

where, uniformly with respect to z, $C_h(z) = C(z) + o(1)$, as $h \to \infty$, and $C(z)$ is analytic at z_0. Furthermore, for some K_1, K_2, c_0 all positive, one has[3], in the disc $|z - z_0| < R_0$,

$$K_1 < |C_h(z)| < K_2 \qquad \text{and} \qquad |\arg(e_h(z))| \le c_0(h+1)|z - z_0|.$$

Proof. Starting from the main relation (8) and unfolding only the e_h that is factored, we obtain

$$(19) \qquad \frac{e_{h+1}}{y^{h+1}} = e_0 \prod_{i=0}^{h} \left(1 - \frac{e_i}{2y}\right) + \frac{e_h(z^2)}{2y^{h+1}} + \frac{1}{2y} \sum_{i=0}^{h-1} \left[\frac{e_i(z^2)}{y^i} \prod_{j=i+1}^{h} \left(1 - \frac{e_j}{2y}\right)\right].$$

We operate with $C_h(z) := e_h(z)/y(z)^h$ and proceed to prove properties of these quantities.

(i) *Upper bound on C_h and existence of $C(z)$.* When z lies in a small enough neighbourhood of $z_0 \in (0, \rho)$, the convergence of e_i to zero is geometric by Lemma 2, and it remains so, uniformly with respect to z restricted to a small neighbourhood of z_0. Furthermore, the inequality $|y(z)| > |z|$, which holds at $z = z_0$, persists, by continuity, for z in a suitably small neighbourhood of z_0. It follows that both the product and the sum in the right-hand side of (19) converge geometrically and uniformly, so that $C_h(z) \to C(z)$ as $h \to \infty$, where $C(z)$ is analytic at z_0. These arguments also imply that $|C_h(z)|$ remains bounded from above by an absolute constant: $|C_h(z)| < K_2$.

(ii) *Lower bound on C_h.* We next observe that, in a small enough neighbourhood of z_0, the quantity $|C(z)|$ must be bounded from below. Indeed, *a contrario*, if this was not the case, then we would need to have $C(z_0) = 0$. Now, because of the convergence of $C_h(z)$ to $C(z)$, we would have $C_h(z_0) = o(1)$, implying $e_h(z_0) = o(y(z_0)^h)$. This last fact is finally seen to contradict Equation (19), since the left hand side taken at z_0 would tend to 0, while the right hand side remains bounded from below by the positive quantity $e_0 \prod_{i=0}^{\infty}(1 - e_i/(2y))$ taken at z_0. A contradiction has been reached. Thus, we must have $|C(z)| > K_1^\star$ for some $K_1^\star > 0$; hence the claimed inequality $|C_h(z)| > K_1$ for all h large enough, say $h > h_0$. (For $h \le h_0$, we can complete the argument by referring again to Equation (19), which precludes the possibility that $e_h(\zeta) = 0$ for $\zeta \in (0, \rho)$. A continuity argument then provides a small domain around z_0 where $C_j(z)$ is bounded from below, for all $j \in \{1, \ldots, h_0\}$.)

(iii) *Bound on the argument.* Finally, the argument of e_h can be expressed as follows:

$$(20) \qquad \arg e_h = \Im(\log e_h) = \Im(\log C_h) + h\Im(\log y) \qquad (\text{mod } 2\pi).$$

We now consider a disc $|z - z_0| < R$ and momentarily examine the effect of letting $R \to 0$. By analyticity of $y(z)$ at z_0 and since $y(z_0)$ is positive real, we have $\Im(\log y(z)) = O(R)$. Next, since $|C_h(z)|$ is bounded from above and below in a small enough fixed neighbourhood of z_0, $C_h(z_0)$ is positive real, and $C_h(z) \to C(z)$, we have, similarly, $\Im(\log C_h(z)) = O(R)$, where the implied constant in $O(\cdot)$ can be taken independent of h. This means that, there exist constants $d_0, d_1 > 0$ such that, provided R is chosen small enough, one has $|\arg(e_h(z))| < d_0 R + d_1 R h$. This last form implies the stated bound on the argument of e_h. $\qquad \square$

With Lemma 5 in hand, we can obtain a first set of properties of $e_h(z)$, which hold for z in a sandclock $\mathcal{S}(r_0, \theta_0)$ and for h "not too large". These will be used in Proposition 2 to derive an upper bound on $|e_N|$ (for some suitably chosen N depending on z), to the effect that e_N eventually satisfies the criterion of Lemma 3. In the following, we only need to consider $z \in \mathcal{S}(r_0, \theta_0)$, with $\Im(z) \ge 0$, since we clearly have $e_h(\bar{z}) = \overline{e_h(z)}$, where $\bar{z}$ denotes the complex conjugate of z.

---

[3] The *argument* of a complex number $w \ne 0$ taken to be the number $\theta \in (-\pi, +\pi]$ such that $w = |w|e^{i\theta}$.

Lemma 6 (Initial behaviour of e_h)**.** *Suppose $\Im(z) > 0$. Define the integer*

$$(21) \qquad N(z) := \left\lfloor \frac{\arccos(1/4)}{\arg y(z)} \right\rfloor - 2.$$

Fix $\theta_0 \leq \frac{\pi}{8}$, with $\theta_0 > 0$. There exists a constant $r_0 > 0$ such that, if z lies in the sandclock $\mathcal{S}(r_0, \theta_0)$, then, for all h such that $1 \leq h \leq N(z)$, the following inequalities hold:

$$(22) \qquad \frac{|y|^{h+1}}{2(h+1)} < |e_h(z)| < 1/2 \qquad \text{and} \qquad 0 \leq \arg(e_h) \leq (h+2)\arg(y).$$

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

Observe that we can also assume, in a small enough sandclock,

$$(23) \qquad \frac{1}{2} < |e_0(z)| < \frac{2}{3},$$

since $e_0(\rho) = 1 - \rho$ has numerical value $\doteq 0.59730$ and $e_0(z)$ is continuous at $z = \rho$.

Proof. As a preamble, we note that $N(z)$ tends to infinity as $z \to \rho$, since $y(\rho) = 1$ is real, hence has argument 0. Consider next the basic recurrence relation (8) rewritten as

$$(24) \qquad e_{h+1}/y = y \cdot e_h/y \left(1 - \frac{e_h/y}{2}\right) + \frac{e_h(z^2)}{2y}.$$

The behaviour of the first term in (24) is dictated by properties of the mapping

$$(25) \qquad g \; : \; w \mapsto w(1 - w/2).$$

(A very similar function appeared in the analysis of Flajolet and Odlyzko [18, Lemma 3].) By a simple modification of the proof in [18], we can check elementarily the implication

$$(26) \qquad \begin{cases} |w| \leq 1 \\ 0 \leq \arg w \leq \arccos(1/4) \end{cases} \Rightarrow \begin{cases} |g(w)| \leq |w| \\ 0 \leq \arg g(w) \leq \arg w. \end{cases}$$

(i) *Weak upper bounds on modulus and bounds on argument.* We are first going to use (26) and induction on h (with $1 \leq h \leq N(z)$), in order to establish a suitably weakened form of (22); namely,

$$(27) \qquad |e_h/y| < 1 \qquad \text{and} \qquad 0 \leq \arg(e_h/y) \leq (h+1)\arg y.$$

We start with the basis of the induction relative to (27), the case $h = 1$, where $e_1 = y - z - z^2$. Observe that $e_1(\rho) = 1 - \rho - \rho^2 \doteq 0.43$, so that $|e_1(z)| < 1/2$ (and, *a fortiori*, $|e_1/y| < 1$) is granted for z close enough to ρ. Write next $z/\rho = 1 + re^{i\theta}$, with θ close to $\pi/2$ and r a small positive number. Then, by virtue of the singular expansion (2) of Lemma 1, we have,

$$y(z) = 1 + i\lambda\sqrt{r}e^{i\theta/2} + O(r),$$

as $r \to 0$, hence

$$\frac{e_1}{y} = 1 - \rho - \rho^2 + i\lambda(\rho + \rho^2)e^{i\theta/2}\sqrt{r} + O(r).$$

Since $\theta/2$ now lies in $(\pi/4 - \pi/16, \pi/4 + \pi/16)$, there results from the last expansion that the argument of e_1/y is essentially a small positive multiple of $\sqrt{r}$. A precise comparison of the arguments of y and e_1/y, as provided by the last two displayed equations, confirms (routine details omitted) that we can choose a small enough r_0 such that, in the sandclock $\mathcal{S}(r_0, \pi/8)$, we have both $|e_1/y| < 1$ and $0 < \arg(e_1/y) \leq 2\arg(y)$.

Suppose now that (27) holds for all integers up to h. In order to determine whether it also holds for $h + 1$, we have to take into account the Pólya term, that is, the second term in the right-hand side of (24). By possibly further restricting r_0, we can guarantee that, for all $z \in \mathcal{S}(r_0, \pi/8)$, this second term does not contribute any increase in the argument of

e_h/y. Indeed, observe that for $z \in \mathcal{S}(r_0, \pi/8)$, we have $\arg(y) \geq \delta r^{1/2}$, with some $\delta > 0$. In addition, by Lemma 5, Equation (20), we have $\arg e_h(z^2) \leq c_0(h+1)|z^2 - \rho^2|$, so that $|z^2 - \rho^2| = O(r)$ is of a smaller order than $O(\sqrt{r})$. Thus, in (24), the second (Pólya) term on the right hand side of the equality has an argument which is of order hr, and, for r small enough, may be taken to satisfy

$$0 \leq \arg(e_h(z^2)/(2y)) \leq h/2 \cdot \arg(y).$$

Now, the simple geometry of parallelograms implies that two complex numbers ζ and ζ', whose arguments lie in $[0, \frac{\pi}{2}]$, satisfy $\arg(\zeta + \zeta') \leq \max(\arg(\zeta), \arg(\zeta'))$. There results, from the induction hypothesis, the chain of inequalities

$$
\begin{aligned}
0 \leq \arg(e_{h+1}/y) &\leq \max\{\arg(e_h/y) + \arg(y), \arg(e_h(z^2)/y)\} \\
&\leq \max\{(h+1)\arg y + \arg(y)\} \\
&\leq (h+2)\arg(y).
\end{aligned}
$$

At this stage, the induction is complete and (27) is established.

(ii) *Improved upper bound on modulus.* The upper bound on the modulus provided by (27), being (slightly) weaker than the upper bound on $|e_h|$ asserted in (22), needs to be strengthened. The $y(z)$ and $e_h(z)$ are analytic, hence continuous, in the domain $\mathcal{D}$ of (13) and all the sandclocks it contains. Also, we have seen that $e_1(\rho) < 1/2$, while, by definition, $e_1(\rho) > e_2(\rho) > \cdots$. So, after possibly restricting r_0 to a smaller value once more, for all $z \in \mathcal{S}(r_0, \pi/8)$, the inequality $|e_h(z)| \leq 1/2$ is guaranteed to hold with $h = 1, \ldots, 6$, this by virtue of continuity. Next, if $h \geq 6$, the alternative recurrence and the fact that $|e_h/y| < 1$ (asserted in (27) and proved in Part (i) above) imply, via the triangle inequality

$$\left|\frac{e_h}{y}\right| \leq |y|\,|g(e_h/y)| + \left|\frac{e_h(z^2)}{2y}\right|, \qquad \text{where} \quad g(e_h/y) = \frac{e_h}{y} \cdot \left(1 - \frac{e_h}{2}\right)$$

($g(w)$ is as defined in (25)). Now, for $h < N(z)$, the quantity $g(w)$ is taken at $w = e_h/y$, which is such that $|w| < 1$ and $\arg(w) < \arccos(1/4)$, so that, by (26),

(28)
$$
\begin{aligned}
|e_{h+1}/y| &\leq |e_6(z)/y| + \frac{1}{2} \cdot \sum_{i=6}^{h+1} |e_i(z^2)/y| \\
&\leq |e_6(\rho)| + O(\sqrt{r}) + \frac{1}{2} \cdot \sum_{i=6}^{\infty} (\rho + 3r)^i,
\end{aligned}
$$

where the last line makes use of the inequality $|e_i(z^2)| \leq (|z|^2/\rho)^i$ granted by Lemma 2. It follows easily that, for $h \geq 6$,

$$|e_{h+1}/y| \leq |e_6(\rho)| + \frac{1}{2} \cdot \frac{\rho^6}{1 - \rho} + O(\sqrt{r}),$$

for all $r \leq r_0$, with r_0 chosen small enough. In particular, for $h \in [6, N(z)]$ and $r \leq r_0$ small enough, we have $|e_h/y| < 1/5$. Combined with previous observatiosn regarding the initial values of $e_j(z)$, this implies the inequality $|e_h/y| < \frac{1}{2}$ for all $h \leq N(z)$, as asserted.

(iii) *Lower bound on modulus.* It finally remains to establish the lower bound on $|e_h|$ in (22). We start with the recurrence relation (24). For $h \leq N(z)$, the additional Pólya term $e_h(z^2)$ only contributes to making $|e_{h+1}|$ larger. Indeed, for $z \in \mathcal{S}(r_0, \theta_0)$, by Lemmas 5 and the upper bound on arguments proved in Part (i), both terms are such that, for $h < N(z)$,

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

Since $x \mapsto x(1 - x/2)$ is increasing in $[0, 1]$, we have $|e_h/y| \geq f_h,$, for all $h \geq 0$, where the sequence $(f_h)_{h \geq 0}$ is defined by

$$f_0 = \frac{|e_0|}{|y|} = \frac{|y - z|}{|y|} \qquad \text{and} \qquad f_{h+1} = |y| \cdot f_h \cdot \left(1 - \frac{f_h}{2}\right).$$

(The latter recurrence relation is precisely the one analysed by Flajolet and Odlyzko [18] in the case of simply generated trees.) For r_0 small enough, a process analogous to the derivation of the alternative recurrence in Lemma 6 yields

$$(29) \quad \frac{|y|^h}{f_h} = \frac{1}{f_0} + \frac{1}{2} \cdot \sum_{i=0}^{h-1} |y|^i + \frac{1}{2} \cdot \sum_{i=0}^{h-1} \frac{f_i}{1 - f_i/2} \cdot |y|^i \leq \frac{1}{f_0} + \frac{3}{2} \cdot \sum_{i=0}^{h-1} |y|^i \leq 2 + \frac{3h}{2}.$$

This last last bound directly implies the lower bound on $|e_h|$ asserted in (22). $\qquad \square$

3.2. **Existence of a convergence sandclock.** We can now turn to a proof of the main result of this section, Proposition 2 stated below, which establishes the existence of a sandclock in which the e_h converge to 0. This proof follows the lines of the analogous statement [18, Proposition 4], where the iteration is "pure". In the present context, we need once more to control the effect of the Pólya terms, which can be done thanks to an easy auxiliary result, Lemma 7.

Lemma 7. *There exist $r_0, \theta_0 > 0$ small enough, so that, for $z \in \mathcal{S}(r_0, \theta_0)$ and for all h satisfying $0 \leq h \leq N$, with $N \equiv N(z)$ as specified in (21), one has*

$$\frac{|e_h(z^2)|}{|e_h(z)^2|} \leq \frac{1}{2} \cdot \left(\frac{4}{5}\right)^h.$$

Proof. Set $z = \rho + re^{i\theta}$, with $z \in \mathcal{S}(r_0, \theta_0)$, for $r_0, \theta_0 > 0$ taken small enough, which will be successively constrained, as the need arises. The inequality

$$(30) \qquad \frac{|e_h(z^2)|}{|e_h(z)^2|} \leq \frac{4(h+1)^2}{y^2} \left|\frac{z^2}{\rho y^2}\right|^h, \qquad 0 \leq h \leq N,$$

combines the upperbound on $e_h(z^2)$ provided by Lemma 2 (with z in the statement to be replaced by z^2) and the lower bound on $e_h(z)$ guaranteed by Lemma 6.

Now, at $z = \rho$, the upper bound (30) takes the form $4(h+1)^2\rho^h$, where $\rho \doteq 0.40269$, so that its decay is about 0.4^h. By continuity of the exponential rate $|z^2/(\rho y^2)|$, there exists a small sandclock such that the decrease is less than $4(h + 1)^2 0.45^h$ (say). Furthermore, we verify easily that this last quantity is less than $\frac{1}{2} \cdot 0.8^h$ for all $h \geq 13$. Thus, the statement is established for h large enough ($h \geq 13$). On the other hand, examination of initial values shows that the ratio $e_j(\rho^2)/e_j(\rho)$ decreases rapidly from a value of about 0.0543, at $j = 0$, to about $7.8 \, 10^{-10}$, at $j = 12$; furthermore, we observe numerically that $2 \cdot 0.8^{-j} e_j(\rho^2)/e_j(\rho)$ is always less than $1/9$, for $j = 0 \ldots, 12$. Thus, by continuity again, in a small enough sandclock, we must have $|e_j(z^2)/e_j(z)| < \frac{1}{2} \cdot 0.8^j$ for $j = 0, \ldots, 12$. $\qquad \square$

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

Proof. It suffices to verify that, for $h = N \equiv N(z)$ as specified in Equation (21), the quantity e_N satisfies the convergence criterion of Lemma 3, which then grants us convergence of the e_j to 0 for $j > N$. For this purpose, we appeal to the alternative recurrence stated in

Lemma 4

$$(31) \qquad \frac{y^h}{e_h} = \underbrace{\frac{1}{2y}\frac{1-y^h}{1-y}}_{M} + \underbrace{\frac{1}{e_0} - \sum_{i=0}^{h-1}\frac{y^{i-1}e_i(z^2)}{2e_i^2}}_{A} + \underbrace{\frac{1}{4y^2}\sum_{i=1}^{h-1}\frac{y^i e_i\left[1-\frac{e_i(z^2)}{e_i^2}\right]^2}{1-\frac{e_i}{2y}\left[1-\frac{e_i(z^2)}{e_i^2}\right]}}_{B}$$

and devise an asymptotic lower bound for the right-hand side. (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 $1 - y(z) = \epsilon e^{it}$. We need to establish properties of the various quantities, which intervene in (31); this, in a small sandclock, that is, 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 $\epsilon \to 0$:

$$(32) \qquad \begin{array}{ll} 1 - |y| = \epsilon\cos t + O(\epsilon^2), & N(z) = -\varphi/(\epsilon\sin t) + O(1) \\ \arg(y) = -\epsilon\sin t + O(\epsilon^2), \quad \text{and} & 1 - |y|^N = 1 - e^{\varphi\cdot\cot t} + O(\epsilon), \end{array}$$

where $\varphi := \arccos(1/4)$.

The first term M of the right-hand side of (31) will be seen to bring the main contribution. It satisfies, as $\epsilon \to 0$:

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

Next, regarding the term A, we have

$$|A| \leq \frac{1}{|e_0|} + \left|\frac{1}{2y}\sum_{i=0}^{N-1}\frac{y^i e_i(z^2)}{e_i(z)^2}\right| \leq \frac{1}{|e_0|} + \frac{1}{2|y|}\sum_{i=0}^{N-1}\left|\frac{e_i(z^2)}{e_i(z)^2}\right| = O(1),$$

since, by Lemma 7, the summands decrease geometrically.

It now remains to analyse B. We split it further: by Lemma 7, for all i such $18 \leq i \leq N$, we have $|e_i(z^2)|/|e_i(z)^2| \leq 1/100$. Then, by Lemma 6, and for ϵ small enough, we obtain

$$|B| \leq \frac{9}{4}\cdot\frac{\left(\frac{3}{2}\right)^2}{1 - \frac{1/2}{2-2\epsilon}\frac{3}{2}} + \frac{1/5}{4}\cdot\frac{\left(\frac{101}{100}\right)^2}{1 - \frac{1/5}{2-2\epsilon}\frac{101}{100}}\frac{1-|y|^N}{1-|y|} < \frac{K}{2} + \frac{3}{50}\cdot\frac{1-|y|^N}{1-|y|}.$$

It follows that

$$(33) \qquad \frac{|y^N|}{|e_N|} \geq \frac{1-e^{\varphi\cdot\cot t}}{\epsilon}\left(\frac{1}{2} - \frac{3}{50\cos t}\right) + O(1) > \frac{2}{5}\cdot\frac{1-e^{\varphi\cdot\cot t}}{\epsilon},$$

where the last inequality holds 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 return to the criterion for convergence (Lemma 3) and verify that in a small enough sandclock the conditions in (14) are satisfied for $m = N$ and some well chosen parameters α and β. Equation (33) provides the required upper bound on e_N, which fixes our choice for α:

$$|e_N| \leq \alpha := \frac{5}{2}\cdot\frac{\epsilon\cdot e^{\varphi\cot t}}{1-e^{\varphi\cdot\cot t}}.$$

We now focus on the second condition in Lemma 3. From (32) and (33) we have, for $\epsilon > 0$ small enough,

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

Next, one can verify that there exists $\delta_0 > 0$ such that for all $t \in [-\pi/4 - \delta_0, -\pi/4 + \delta_0]$, we have

$$\cos t - \frac{5}{4} \cdot \frac{e^{\varphi \cdot \cot t}}{1 - e^{\varphi \cdot \cot t}} > \frac{1}{4}.$$

Thus, for all $\epsilon > 0$ small enough, we can choose $\beta \in (0,1)$, so that the first two conditions in Lemma 3 are satisfied; namely,

$$(34) \qquad |e_N| \le \alpha = \frac{5}{2} \cdot \frac{\epsilon \cdot e^{\varphi \cot t}}{1 - e^{\varphi \cdot \cot t}} \qquad \text{and} \qquad |y| + \frac{\alpha}{2} < \beta := 1 - \frac{\epsilon}{4}.$$

One then easily verifies that the third condition also holds for small enough $\epsilon > 0$: here, $\alpha(1 - \beta) = \Omega(\epsilon^2)$, and, by (32), we have $(|z|^2/\rho)^N = o(\epsilon^2)$. So for ϵ small enough, $(|z|^2/\rho)^N < \alpha(1 - \beta)$. This shows that the criterion for convergence of Lemma 3 is satisfied with values of α and β specified in (34). As a consequence, $e_h(z) \to 0$ for all z in a small enough sandclock. $\qquad \square$

4. **Main approximation**

In this section, we develop precise quantitative estimates of $e_h(z)$ near the singularity ρ and in a sandclock; these estimates serve as the main ingredient required for developing limit laws for height in the next section.

Proposition 3 (Main estimate for e_h in a sandclock). *There exist $r_1, \theta_1 > 0$ and $K, K' > 0$ such that for all $z \in \mathcal{S}(r_1, \theta_1)$ and all $h \ge 1$,*

$$(35) \quad \frac{y^h}{e_h} = \frac{1}{2} \frac{1 - y^h}{1 - y} + R_h(z), \quad where \quad |R_h(z)| \le K \min\left\{ \log \frac{1}{1 - |y|}, \log(1 + h) \right\}.$$

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

In order to prove this proposition, we need a better control on error terms, which can be achieved by extending the bounds obtained in Section 3 for $h > N$, knowing now that the e_h converge (Proposition 2). The proof requires the bounds to be *uniform* both in the distance to the singularity $|z - \rho|$ *and* in the height h, as expressed by Lemmas 8 and 9 below. The bound (36) below serves as a useful complement to the lower bound in (22), only holds for $h \le N$.

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 one has*

$$(36) \qquad |e_h(z)| \ge \frac{(1 - \delta)^{h+2}}{2(h+1)}, \quad for \ all \ h \ge 0.$$

Proof. Let $\delta \in (0,1)$. We have $|y| > 1 - \delta/4$ provided $r := |z - \rho| < r_0$ small enough. Then, by Lemma 6, the estimate (36) holds for $h \le N$.

We thus only need to consider now the case $h > N$. Assume further that $z \in \mathcal{S}(r_0, \theta_0)$, as in Proposition 2. Then, $|e_h| \le \alpha$, for α as in (34). The recurrence relation (8) implies

$$(37) \quad |e_{h+1}| \ge |y| \, |e_h| \left(1 - \frac{|e_h|}{2|y|}\right) - \frac{|e_h(z^2)|}{2} \ge |y| \left(1 - \frac{\alpha}{2|y|}\right) \cdot |e_h| - \frac{|e_h(z^2)|}{2}.$$

However, by (34), we have $|y| + \alpha/2 < 1$ so that $|y| - \alpha/2 > 1 - \delta/2$. Lemma 2, which serves to bound the Pólya term $|e_h(z^2)|$, then yields $|e_{h+1}| \ge (1 - \delta/2)|e_h| - (\rho + r)^h$. Therefore dividing both side of the recurrence relation by $(\rho + r)^{h+1}$, we obtain for $h \ge N$,

$$(38) \qquad \frac{|e_{h+1}|}{(\rho + r)^{h+1}} \ge \left(\frac{1 - \delta/2}{\rho + r}\right) \frac{|e_h|}{(\rho + r)^h} - \frac{1}{\rho + r}.$$

The remainder of the proof then consists in extracting the desired bound (36) from the latter relation by unfolding the recurrence from h down to N. To this effect, recall that, by Lemma 6, $|e_N| > y^{N+1}/(2(N+1))$ and $N < K/\sqrt{r}$, for some constant K. Hence, we can set r to a value small enough that,

$$(39) \qquad \frac{|e_N|}{(\rho+r)^N} > \frac{2}{\delta} \qquad \text{and} \qquad \frac{1-\delta}{\rho+r} > 1.$$

Then, for such r, using (38) and (39), it is easly verified by induction on h (with $h \geq N$) that $|e_h|/(\rho+r)^h > 2/\delta$. Using his last bound in (38) gives, for $h \geq N$:

$$\frac{|e_{h+1}|}{(\rho+r)^{h+1}} \geq \left(\frac{1-\delta}{\rho+r}\right)\frac{|e_h|}{(\rho+r)^h} \geq \left(\frac{1-\delta}{\rho+r}\right)^{h+1-N}\frac{|e_N|}{(\rho+r)^N}.$$

We can finally recover the information on $|e_h|$ by means of the lower bound for $|e_N|$ in Lemma 6. For all $h > N$, we then have

$$|e_h| \geq |e_N| \cdot (1-\delta)^{h-N+1} \geq \frac{(1-\delta)^{h+2}}{2(N+1)} \geq \frac{(1-\delta)^{h+2}}{2(h+1)},$$

and the proof is complete. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

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

Lemma 9 (Uniform upper bound for $|e_h(z)|$)**.** *There exist constants $r_1, \theta_1 > 0$ and $c_1 > 0$ such that, for any $h \geq 1$, and $z \in \mathcal{S}(r_1, \theta_1)$, we have*

$$|e_h(z)| \leq \frac{c_1}{h}.$$

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 $-\pi/4$. Observe that ϵ is of the order of $\sqrt{z - \rho}$.

Our starting point is again (31), which we now use to obtain an upper bound on $|e_h|$. The first term M is such that

$$(40) \qquad |M| = \left| \frac{1}{2y}\frac{1-y^h}{1-y} \right| = \frac{1}{2|y|} \cdot \frac{|1-y^h|}{|1-y|} \geq \frac{1-|y|^h}{2|1-y|} = \frac{1-|y|^h}{2\epsilon}.$$

On the other hand, for all $h \geq 0$ and $\epsilon > 0$ small enough, by Lemmas 2 and 8, the first error term A in (31) satisfies

$$|A| \leq \frac{1}{|e_0|} + \frac{1}{2|y|}\sum_{i=0}^{h-1}\left|\frac{e_i(z^2)}{e_i(z)^2}\right| \leq \frac{1}{|e_0|} + \frac{1}{2(1-\epsilon)}\sum_{i=0}^{\infty}4(i+1)^2\left(\frac{\rho+\epsilon}{(1-\epsilon)^2}\right)^i.$$

There exists $\epsilon_1 > 0$ such that for all $\epsilon < \epsilon_1$ the geometric term in the series above is at most $2\rho < 1$; together with the fact that $e_0 = y - z = 1 - \rho + O(\epsilon)$, this implies that $|A| \leq 11/(1-2\rho)^3$.

We now bound the second error term B in (31). Note first that, for all ϵ small enough, we have $|e_i| \leq 1/2$ for all $i \geq 0$: for $h \leq N$, this is implied by Lemma 6, while for $h \geq N$ we have $|e_h| < \alpha < 2(1-|y|) = O(\epsilon)$. Furthermore, by Lemmas 2 and 8, for all $\epsilon < \epsilon_1$ small enough $|e_i(z^2)|/|e_i(z)^2| < 1/100$, for $i \geq h_0$ depending only on ϵ_1. It follows that

$$|B| \leq h_0 + \frac{1}{8} \cdot \frac{(101/100)^2}{1 - \frac{1}{4} \cdot \frac{101}{100}} \cdot \sum_{i=h_0+1}^{h} |y|^{i-2} \leq h_0 + \frac{1}{5} \cdot \frac{1-|y|^h}{1-|y|}.$$

As a consequence, using Lemma 4, and combining the bounds just obtained on $|A|$ and $|B|$ with (40), one sees that, for all $h \geq 0$,

$$(41) \quad \frac{|y^h|}{|e_h|} \geq \frac{1 - |y|^h}{\epsilon}\left(\frac{1}{2} - \frac{1}{5\cos t}\right) - h_0 - \frac{11}{(1 - 2\rho)^3} > \frac{1 - |y|^h}{5\epsilon} - h_0 - \frac{11}{(1 - 2\rho)^3},$$

for $|\pi/4 + t|$ small enough.

The relation above provides a decent upper bound on $|e_h|$ provided that $|y^h|$ is small enough. With this in mind, we now prove an upper bound on $|y^h|$ for all $h \geq 0$. First, when h is not too large, $|y|^h$ should decrease at least linearly in h: we show that for some small enough $\delta > 0$, $|y|^h \leq 1 - \delta h\epsilon$ for all $h \leq N$. For some fixed z, the sequence $(|y|^h, h \geq 0)$ is convex; thus if $|y|^{N'} \leq 1 - \delta N'\epsilon$ for some $N' \geq N$, then $|y|^h \leq 1 - \delta h\epsilon$ for all $0 \leq h \leq N$. Recall that $\varphi = \arccos(1/4)$; we now prove that we might take $N' := -2\varphi/(\epsilon \sin t)$. By (32), for ϵ small enough, $|y| \leq 1 - \frac{\epsilon}{2}\cos t$ and $N < N'$ and

$$|y|^{N'} \leq \left(1 + \frac{\epsilon \cos t}{2}\right)^{N'} \leq \exp\left(-\frac{N'\epsilon \cos t}{2}\right) = e^{\varphi \cot t}.$$

However, for $|t + \pi/4| < 1/100$, then $e^{\varphi \cot t} < 1/2$, so that we can pick $\delta > 0$ such that

$$e^{\varphi \cot t} < 1 + \frac{2\delta\varphi}{\sin t} = 1 - \delta N'\epsilon.$$

It follows by (41) that there exists $\delta > 0$ small enough such that $|e_h| \leq 10/(\delta h)$, for $0 \leq h \leq N'$, for all $|\pi/4 + t| < 1/100$ and $\epsilon > 0$ small enough.

On the other hand, if $h \geq N$ and $\epsilon > 0$ is small enough and $|\pi/4 + t| < 1/100$, then $1 - |y|^h \geq 1 - 2e^{\varphi \cot t} > 1/4$ by (32). As a consequence,

$$|e_h| \leq 40\epsilon|y|^h \leq 40\epsilon(1 - \epsilon\cos t + O(\epsilon^2))^h \leq 40\epsilon(1 - \epsilon/2)^h,$$

for all ϵ small enough and t close enough to $-\pi/4$. Now, seen as a function of ϵ, the maximum of the right-hand side above is obtained for $\epsilon = 2/(h + 1)$, which implies that $|e_h| \leq 80/(h + 1)$, for $h \geq N$. Finally, by Lemma 6, and the bounds above, the result follows by choosing $c_1 = \max\{h_0, 10/\delta, 80\}$. $\square$

Proof of Proposition 3. The proof consists in using Lemma 9 above to bound the error terms in (31) 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

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

which proves the main statement of Proposition 3. Finally, since A and B are partial sums, we obtain

$$|R_h - R_{h+1}| = \left|\frac{y^{h-1}}{2}\frac{e_h(z^2)}{e_h(z)^2} + \frac{y^{h-2}}{4}e_h\left[1 - \frac{e_h(z^2)}{e_h(z)^2}\right]^2\left(1 - \frac{e_h}{2y}\left[1 - \frac{e_h(z^2)}{e_h(z)^2}\right]\right)^{-1}\right|,$$

a quantity which is easily seen to be uniformly $O(1/h)$, thanks to Lemmas 8 and 9. $\square$

5. Asymptotic analysis and distribution estimates

The basis of our estimates relative to the distribution of height is the main approximation of e_h in Proposition 3, which is valid in a *fixed sandclock* at ρ. Given its importance, we repeat it under the simplified form:

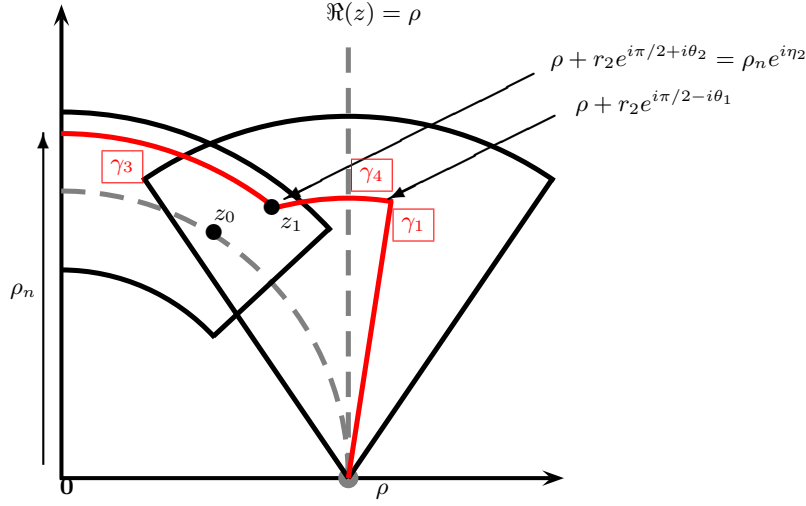$$(42) \qquad\qquad e_h(z) \equiv y(z) - y_h(z) \approx 2\frac{1 - y}{1 - y^h}y^h.$$

Figure 3. Fine details of the Cauchy integration contour γ in the vicinity of ρ.

(Here, the symbol "$\approx$" is to be loosely interpreted in the sense of "approximately equal".) This approximation acquires a precise meaning, when z remains fixed and h tends to infinity, in which case it expresses the geometric convergence of $e_h(z)$ to 0 (since $|y| < 1$); also, when h remains fixed and z tends to ρ, it reduces to the numerical approximation $e_h(\rho) \approx 2/h$, whose accuracy increases with increasing values of h. In other words, the precise version of (42) provided by Proposition 3 consistently covers, in a uniform manner, the case when *both* $z \to \rho$ and $h \to \infty$. (Analogues of the formula (42) surface in the case of general plane trees in [12], plane binary trees [18], and labelled Cayley trees in [38].)

The exploitation of the enhanced versions of (42) relies on Cauchy's coefficient formula (10). The contour γ in Cauchy's integral (10) will be comprised of several arcs and line segments[4] that lie *outside* of the disc $|z| \leq \rho$ and taken in the union of a suitable sandclock (as granted by Proposition 3) and of a tube, overlapping with the sandclock (where properties of Proposition 1 are in effect). The strategy just described belongs to the general orbit of singularity analysis methods expounded in [19, Ch. VI–VII]. We propose to apply it to the height-related generating functions $e_h(z)$ ("central" limit law) and $e_{h-1}(z) - e_h(z)$ (local limit law).

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*

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

where $\qquad \Theta(x) := \sum_{k \geq 1} (k^2 x^2 - 2) e^{-k^2 x^2/4}.$

---

[4] In order to have well-defined determinations of square roots, one may 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.

Proof. We aim at using Cauchy's formula (10) with a well-chosen[5] integration contour γ. The reader should consult Figures 2 and 3. First, we choose *a priori* a sandclock $\mathcal{S}$, whose existence is granted by Proposition 2 and such that the approximation properties of Proposition 3 hold. By design, this sandclock contains *in its interior* a small arc of the circle $\{|z| = \rho\}$. Choose arbitrarily a point z_0 on this small arc, with $z_0 \neq \rho$, $\Im(z_0) > 0$, and set $z_0 = r_2 e^{i\pi/2 + i\theta_0}$. Proposition 1 guarantees the existence of a tube $\mathcal{T}$ that has z_0 *in its interior* and for which the convergence $e_h \to 0$ is ensured. We have now determined a sandclock and a partially overlapping tube, whose union will be seen to contain the contour γ (where $e_h \to 0$) and whose intersection contains $z_0 = r_2 e^{i\theta_0}$.

The contour γ is essentially a Hankel contour escaping from ρ along rectilinear portions γ_1 and γ_2 such that

$$\gamma_1 = \bar{\gamma}_2 = \left\{ \rho + \xi e^{i\pi/2 - i\theta_1} \ : \ \xi \in [0, r_2] \right\},$$

where θ_1 is chosen positive and strictly less than the half-angle of the sandclock $\mathcal{S}$. By design, the segments γ_1 and γ_2 lie entirely inside the sandclock $\mathcal{S}$.

The component γ_3 of the contour is a subarc of the circle

$$C_n := \{ z \ : \ |z| = \rho_n \}, \qquad \text{where} \quad \rho_n := \rho \left(1 + \frac{\log^2 n}{n} \right).$$

Precisely, let $z_1 \equiv z_1(n)$ be the intersection point in the upper half-plane of the circle C_n and the circle $\{|z| = r_2\}$. When n gets large, this point z_1 comes closer and closer to z_0, so that, for all n large enough, it must belong to the intersection $\mathcal{S} \cap \mathcal{T}$. In other words, we can write

$$z_1 = \rho_n e^{i\eta_2} = \rho + r_2 e^{i\pi/2 + i\theta_2},$$

where $\eta_2 = \eta_2(n)$ and $\theta_2 \equiv \theta_2(n)$ both depend on n and tend to finite limits as $n \to +\infty$ (in particular, $\theta_2 \to \theta_0$). Then we take

$$\gamma_3 = \left\{ \rho_n e^{i\theta} : \theta \in [\eta_2, 2\pi - \eta_2] \right\},$$

and for n large enough, the arc γ_3 entirely lies in the tube $\mathcal{T}$.

We can finally complete the contour to make it connected, with joining arcs γ_4 and γ_5, which are arcs of $\{|z - \rho| = r_2\}$ defined by

$$\gamma_4 = \bar{\gamma}_5 = \{ r_2 e^{i\pi/2 + i\theta} : \theta \in [-\theta_1, \theta_2] \},$$

so that both arcs lie inside the sandclock $\mathcal{S}$.

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[6] constant K. On the other hand the Cauchy kernel z^{-n} is small on the contour, so that

$$(43) \qquad \left| \int_{\gamma_3} e_h(z) \frac{dz}{z^{n+1}} \right| < K_1 \rho^{-n} \exp\left(-\log^2 n \right).$$

Join portions γ_4, γ_5. By Proposition 2, one has $e_h \to 0$ uniformly on $\gamma_4 \cup \gamma_5$ as $h \to \infty$. In particular $|e_h(z)| \leq K_2$ for some absolute constant K_2. By definition, for all

---

[5]It might be that none of the tubes corresponding to Proposition 1 includes points to the right of the vertical line $\Re(z) = \rho$, hence the need to insert "joins" γ_4 and γ_5. (The discussion of this case was inadvertently omitted from the earlier version [7].) An alternative would be to make use of a contour that is squeezed in between the circle $|z| = \rho$ and the vertical line $\Re(z) = \rho$ (this is done in [38], where the circle itself is used); but then the near stationarity of the modulus of the Cauchy kernel, $|z|^{-n}$, makes it technically harder, or at least less transparent, to translate approximations of generating functions into coefficient estimates.

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

$z \in \gamma_4 \cup \gamma_5$, $|z| \geq \rho_n$ so that, for the same reasons as in (43),

$$(44) \qquad \left| \int_{\gamma_4 \cup \gamma_5} e_h(z) \frac{dz}{z^{n+1}} \right| \leq K_3 \rho^{-n} \exp(-\log^2 n).$$

Outer rectilinar parts of γ_1 and γ_2. Let $\mathcal{D}_n := \{|z - \rho| \geq \delta_n\}$, with

$$\delta_n = \frac{\log^2 n}{n}.$$

Note that for $z \in \gamma_1 \cap \mathcal{D}_n$, we have $|z| \geq \rho + \delta_n \sin \theta_1$. For the same reason as before,

$$(45) \qquad \left| \int_{(\gamma_1 \cup \gamma_2) \cap \mathcal{D}_n} e_h(z) \frac{dz}{z^{n+1}} \right| \leq K_4 \rho^{-n} \exp(-K_4' \log^2 n).$$

The total contribution of the outer circular arc γ_3, of both join portions γ_4 and γ_5, and of the outer rectilinear parts $\gamma_1 \cap \mathcal{D}_n$, $\gamma_2 \cap \mathcal{D}_n$ are thus *exponentially small* compared to y_n, hence totally negligible in the present context.

Inner rectilinear parts of γ_1 and γ_2. This is where action takes place. From now on, we operate with the normalization

$$h = \lambda^{-1} x \sqrt{n},$$

where x is taken to range over a fixed compact interval of $\mathbb{R}_{>0}$. We now focus on the portions of γ_1 and γ_2 lying outside $\mathcal{D}_n$. We denote them by $\widetilde{\gamma}_1$ and $\widetilde{\gamma}_2$, respectively, and note that all their points are at a distance from ρ that tends to 0, as $n \to +\infty$. Our objective is to replace e_h by the simpler quantity

$$(46) \qquad \widehat{e}_h(z) \equiv \widehat{e}_h := 2 \frac{1 - y}{1 - y^h} y^h,$$

as suggested by Proposition 3. Along $\widetilde{\gamma}_1, \widetilde{\gamma}_2$, the singular expansion of $y(z)$ applies, so that $1 - y = O((\log n)/n^{1/2})$ and the error term $R_h(z)$ from Equation (35) is $O(\log n)$. There results that $(1 - y^h)/(1 - y)$ is always at least as large in modulus as $K_5 \sqrt{n}/\log n$ (this, by a study of the variation of $|1 - e^{-h\tau}|/|1 - e^{-\tau}|$), and we have

$$(47) \qquad \frac{y^h}{e_h} = \frac{y^h}{\widehat{e}_h} \left(1 + O\left(\frac{\log^2 n}{\sqrt{n}} \right) \right).$$

It proves convenient to define

$$(48) \qquad E(h, n) := \frac{1}{2i\pi} \int_{\widetilde{\gamma}_1 \cup \widetilde{\gamma}_2} \widehat{e}_h \frac{dz}{z^{n+1}},$$

and effect the change of variables

$$(49) \qquad z = \rho \left(1 - \frac{t}{n} \right), \qquad dz = -\frac{\rho}{n} dt.$$

With this rescaling, the point t then starts from $-i\rho^{-1} n \delta_n e^{-i\theta_1}$, loops to the right of the origin, then steers away to $i\rho^{-1} n \delta_n e^{i\theta_1}$. Given the singular expansion of $y(z)$ in (2), we have on the small arcs $\widetilde{\gamma}_1, \widetilde{\gamma}_2$,

$$(50) \qquad z^{-n} = \rho^{-n} e^t \left(1 + O\left(\frac{\log^4 n}{n} \right) \right), \qquad y(z) = 1 - \lambda \sqrt{\frac{t}{n}} + O\left(\frac{t}{n} \right),$$

and, with $h = \lambda^{-1} x \sqrt{n}$ and $|t| \leq K_6 \log^2 n$, since $\delta_n = \log^2 n / n$:

$$(51) \qquad y^h = \exp\left(-x\sqrt{t} \right) \left(1 + O\left(\frac{t}{\sqrt{n}} \right) \right) = \exp\left(-x\sqrt{t} \right) \left(1 + O\left(\frac{\log^2 n}{\sqrt{n}} \right) \right).$$

We also find[7], for the range of values of t corresponding to $\widetilde{\gamma}_1, \widetilde{\gamma}_2$:

$$
\begin{aligned}
\frac{1-y^h}{1-y} &= \frac{1-\exp(-x\sqrt{t})(1+t/\sqrt{n})}{\lambda\sqrt{t/n}}\left(1+O\left(\frac{\log^\star n}{\sqrt{n}}\right)\right)\\
&= \left[\sqrt{n}\cdot\frac{1-\exp(-x\sqrt{t})}{\lambda\sqrt{t}}+O(\sqrt{t})\right]\left(1+O\left(\frac{\log^\star n}{\sqrt{n}}\right)\right)\\
&= \left[\sqrt{n}\cdot\frac{1-\exp(-x\sqrt{t})}{\lambda\sqrt{t}}\right]\left(1+O\left(\frac{\log^\star n}{\sqrt{n}}\right)\right).
\end{aligned}
\tag{52}
$$

The approximations (50), (51), and (52) motivate considering, as an approximation of $E(h,n)$ in (48), the contour integral

$$
\text{(53)}\quad J(X) := \frac{1}{2i\pi}\int_{\mathcal{L}}\frac{\exp(-X\sqrt{t})}{1-\exp(-X\sqrt{t})}\sqrt{t}e^t\,dt = \frac{1}{2i\pi}\sum_{k\geq 1}\int_{\mathcal{L}}\exp(-kX\sqrt{t})\sqrt{t}e^t\,dt,
$$

where $\mathcal{L}$ goes from $-\infty+i\infty$ to $-\infty-i\infty$ and winds to the right of the origin. We now make $J(X)$ explicit. Each integral on the right side of (53) can be evaluated by the change of variables $w=i\sqrt{t}$, equivalently, $t=-w^2$. By completing the square and flattening the image contour $\mathcal{L}'$ onto the real line, we obtain:

$$
\text{(54)}\qquad\qquad J(X) = \frac{1}{4\sqrt{\pi}}\sum_{k\geq 1}e^{-k^2X^2/4}(k^2X^2-2).
$$

From the chain of approximations in Equations (46) to (53), we are then led to expect the approximation

$$
E(h,n) \sim 2\lambda\rho^{-n}n^{-3/2}J(x),
$$

which is justified next.

Error management. In order to justify the replacement of e_h by $\widehat{e}_h$, following (47) and (48), we observe the estimate

$$
\text{(55)}\qquad\left|\int_{\widetilde{\gamma}_1\cup\widetilde{\gamma}_2}|y|^h\frac{|1-y|}{|1-y^h|}\frac{|dz|}{|z|^{n+1}}\right| = O\left(\rho^{-n}\frac{\log^4 n}{n^{3/2}}\right).
$$

This results from the discussion of the lower bound on $(1-y^h)/(1-y)$ that follows (46), the inequality $|y^h|\leq 1$, and the fact that the length of the integration interval is $O(\log^2 n/n)$. The error in our approximation has three sources: the two successive replacements

$$
\text{(56)}\qquad\qquad e_h\mapsto\widehat{e}_h,\qquad\frac{1-y^h}{1-y}\mapsto\frac{1-\exp(-\lambda x\sqrt{t})}{\lambda\sqrt{t/n}}
$$

and the integration on a finite contour. We have, for $z\in\widetilde{\gamma}_1\cup\widetilde{\gamma}_2$:

$$
e_h = \widehat{e}_h\left(1+O\left(\frac{\log^2 n}{\sqrt{n}}\right)\right) = 2\lambda\sqrt{\frac{t}{n}}\cdot\frac{\exp(-x\sqrt{t})}{1-\exp(-x\sqrt{t})}\cdot\left(1+O\left(\frac{\log^\star n}{\sqrt{n}}\right)\right).
$$

Finally, the infinite extension of the contour only entails an additive error term of the form $O(\exp(-K\log^4 n))$, since

$$
\int_{\log^2 n}^{\infty}e^{-w^2}\,dw = O(e^{-\log^4 n}).
$$

This implies, for $h=\lambda^{-1}x\sqrt{n}$:

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

---

[7]The expression $\log^\star n$ represents an unspecified positive power of $\log n$.
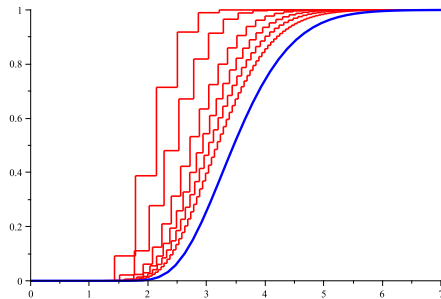
Figure 4. The normalized distribution functions $\mathbb{P}(H_n \leq \lambda^{-1}x\sqrt{n})$, for $n = 10, 20, 50, 100, 200, 500$, as a function of x, and the limit distribution function $1 - \Theta(x)$, where $\Theta(x)$ is specified in Theorem 1.

The explicit form of $J(X)$ in (54) and the asymptotic form of y_n (Lemma 1) jointly yield the statement. □

The main message of the proof of Theorem 1 is twofold: (i) for any "reasonable" expression involving e_h, *the estimation of the Cauchy coefficient formula can be limited to a small neighbourhood of ρ* (parts $\widetilde{\gamma}_1$ and $\widetilde{\gamma}_2$), since the other parts of the contour γ have exponentially negligible contributions; (ii) *the approximation provided by Proposition 3 and Equation* (35) *is normally sufficient to derive first-order asymptotic estimates.*

The convergence in law expressed by Theorem 1 is illustrated by Figure 4. The proof of the Theorem point to an error term, in the convergence to the limit, that is of the form $O((\log^a n)/\sqrt{n})$, with an unspecified exponent a. As a matter of fact, the value $a = 1$ is suggested by the logarithmic character of the error term in (35) of Proposition 3. Convergence is, at any rate, somewhat slow, a fact that is perceptible from Figure 4.

On an other register, the distribution function $1-\Theta(x)$ belongs to the category of elliptic theta functions [42, Ch. XXI], which are of the rough form[8] $\sum q^{k^2}e^{2ikz}$ and are well-known to satisfy transformation formulae [42, p. 475]. Regarding $\Theta(x)$, such formulae provide an alternative form, which we state for the *density function*, $\theta(x) := -\Theta'(x)$:

$$(57) \qquad \theta(x) = \frac{8\sqrt{\pi^3}}{x^3}\theta\left(\frac{4\pi}{x}\right).$$

Theorem 2 (Local limit law of height). *The distribution of the height H_n of a random tree taken uniformly from $\mathcal{Y}_n$ admits a local limit: for x in a compact set of $\mathbb{R}_{>0}$ and $h = \lambda^{-1}x\sqrt{n}$ an integer, there holds uniformly*

$$\mathbb{P}(H_n = h) \sim \frac{\lambda}{\sqrt{n}}\vartheta(x),$$

where $\qquad \vartheta(x) = -\Theta'(x) = (2x)^{-1}\sum_{k\geq 1}(k^4x^4 - 6k^2x^2)e^{-k^2x^2/4}.$

Proof. We abbreviate the discussion, since it is technically very similar to the proof of Theorem 1: only the approximations near $z = \rho$ differ. Proceeding in this way, based on

---

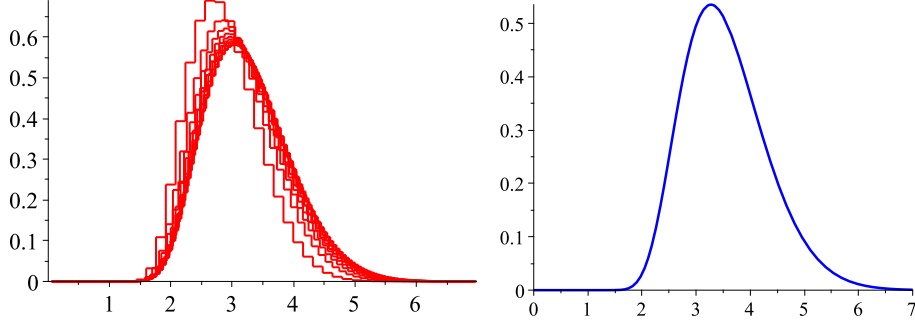[8] Do $q = e^{-x^2/4}$ and $z = 0$ to recover $\theta(x)$.

Figure 5. *Left*: the normalized histograms of the distribution of height $\mathbb{P}(H_n = h)$ (as a function of x, with $h = \lfloor \lambda^{-1} x \sqrt{n} \rfloor$), for $n = 100, 200, \dots, 500$. *Right*: the limit density $\theta(x) = -\Theta'(x)$.

Proposition 3, we can justify approximating the number of trees of height exactly h and size n by the integral

$$\frac{1}{2i\pi} \int_{\widetilde{\gamma}_1 \cup \widetilde{\gamma}_2} (\widehat{e}_{h-1} - \widehat{e}_h) \, \frac{dz}{z^{n+1}},$$

with $\widehat{e}_h$ as defined in (46). We have

(58) $$\widehat{e}_{h-1} - \widehat{e}_h = 2y^{h-1} \frac{(1-y)^2}{(1-y^h)(1-y^{h-1})}.$$

The approximations (51) and (52) then motivate considering the quantity

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

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

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

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

Figure 5 displays the normalized histograms of the distribution of height and a plot of the corresponding limit density.

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 [20, Th. 1.1]):

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

(59) $$\left| \mathbb{P}(H_n \geq \lambda^{-1} x \sqrt{n}) - \Theta(x) \right| \leq \frac{C}{n^B}.$$

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

$$\mathbb{P}(H_n \ge \lambda^{-1} x\sqrt{n}) \sim x^2 e^{-x^2/4}.$$

Similar estimates hold for the local law. These estimates can furthermore be supplemented by (very) large deviation estimates in the style of [20, Th.1.4]:

Theorem 4 (Very large deviations)**.** *There exists a continuous increasing function $I(u)$ satisfying $I(u) > 0$ for $0 < u \le 1$ and such that, given any fixed $\delta > 0$, one has for all $x \in [\delta, 1 - \delta]$ and all n*

$$\mathbb{P}(H_n \ge xn) \le Kn^{3/2}e^{-nI(x)},$$

where K only depends on δ.

Proof. We propose to use saddle point bounds [19, p. 246]: for *any* $r \in (0, \rho)$, one has

(60) $$\mathbb{P}\{H_n \ge h\} = \frac{e_{h,n}}{y_n} \le \frac{1}{y_n}\left(\frac{e_h(r)}{r^n}\right).$$

The first step is to obtain an upper bound on $e_h(r)$, for $r \in (0, \rho)$. For such r, all terms in the recurrence relation (8) are non-negative and expanding the relation with the help of Lemma 2 yields, for all $h \ge 0$, the inequality

$$e_{h+1}(r) \le y(r)e_h(r) + \left(\frac{r^2}{\rho}\right)^h \le y(r)^h\left(\sum_{i=1}^{h}\left(\frac{r^2}{\rho y(r)}\right)^i + e_1(r)\right).$$

However, it is easily verified that for all $r \in (0, \rho)$, we have $y(r) > r + r^2 + r^3 \ge r^2/\rho$. As a consequence, the series above converges and there exists a universal constant K such that

$$e_h(r) \le Ky(r)^h, \quad \text{for } h \ge 0 \text{ and } r \in (0, \rho).$$

The last estimate, the saddle point bound (60), and Lemma 1 yield, in the region $h = xn$,

$$\mathbb{P}\{H_n \ge h\} \le K'n^{3/2}\left(y(r)^x\frac{\rho}{r}\right)^n,$$

for some other universal constant K' and for any $r \in (0, \rho)$.

The goal is now to make an *optimal choice* of the value of r. For x kept fixed and regarded as a parameter, we consider

$$J(r, x) := \frac{y(r)^x}{r}$$

as a function of r, and henceforth abbreviated as $J(r)$. We have $J(0) = +\infty$ and $J(\rho) = \rho^{-1}$. The point, to be justified shortly, is that $J(r)$ decreases from $+\infty$ to some minimal value $J(\xi)$, when $r = \xi$; then it increases again to ρ^{-1} for $r \in (\xi, \rho)$. In particular, we must have $J(\xi) < \rho^{-1}$, which suffices to imply a non-trivial exponential bound on the probabilities.

The unimodality of $J(r) \equiv J(r, x)$ results from the usual convexity properties of generating functions (see [13] or [19, pp. 250 and 580]). Indeed it suffices to observe that the logarithmic derivative (all derivatives being taken with respect to r), namely,

$$\frac{J'(r)}{J(r)} = x\left(\frac{ry'(r)}{y(r)} - \frac{1}{x}\right),$$

varies monotonically from $-\infty$ to $+\infty$, as r varies from 0 to 1. This last fact is a consequence of the positivity of

$$v := \frac{\partial}{\partial r}\left(\frac{rJ'(r)}{J(r)} - \frac{1}{x}\right),$$

itself granted, since $V = rv$ is the variance of a random variable X with probability generating function $\mathbb{E}\left[u^X\right] = J(ru)/J(r)$.

In summary, from the preceding considerations, the system

$$\begin{cases} \xi(x) & : & x\xi y'(\xi) - y(\xi) = 0 \\ I(x) & = & x\log y(\xi) - \log\xi + \log\rho \end{cases}$$

uniquely determines a function $I(x)$, which precisely satisfies the properties asserted in Theorem 4 $\qquad\square$

Finally, the approximation of e_h by $\widehat{e}_h$ in (46) is good enough to grant us access to moments (cf also [18]).

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

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

Proof. The problem reduces to estimating generating functions of the form

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

which are accessible to the Mellin transform technology [21], upon setting $y = e^{-t}$. If we let $F_r(t) = \sum_{h\geq 1} h^r \frac{e^{-ht}}{1-e^{-ht}}$, then the Mellin transform $F_r^\star(s)$ is given by

$$F_r^\star(s) = \zeta(s-r)\zeta(s-1)\Gamma(s),$$

and is valid in the fundamental strip $s > r+1$. The information relative to the distribution is concentrated around the singularity, hence for values of y such that $y \to 1$, or equivalently $t \to 0$. The asymptotics of $F_r(t)$ as $t \to 0$ correspond to the singular expansion of its Mellin transform $F_r^\star(s)$ to the left of the strip.

For $r \geq 2$, the main contribution is due to the simple pole at $s = r + 1$, which has residue $\zeta(r)\Gamma(r+1)$. It follows that

$$F_r(t) \sim \zeta(r)\Gamma(r+1)t^{-r-1} \qquad r \geq 2.$$

Since $1 - y \sim \lambda\sqrt{1-z/\rho}$, and $y = e^{-t}$, we have $t \sim \lambda\sqrt{1-z/\rho}$ and

$$M_r(z) \sim 2\zeta(r)\Gamma(r+1)\lambda^{-r+1}(1-z/\rho)^{-(r-1)/2}.$$

Singularity analysis theorems imply

$$[z^n]M_r(z) \sim 2\zeta(r)\lambda^{-r+1}\Gamma(r+1)\rho^{-n}\frac{n^{-(r+1)/2}}{\Gamma((r-1)/2)}.$$

The duplication formula for the Gamma function, combined with the estimate for y_n, then yields:

$$\mathbb{E}\left[H_n^r\right] \sim \frac{[z^n]M_r(z)}{y_n} \sim \left(\frac{2}{\lambda}\right)^r \zeta(r)r(r-1)\Gamma(r/2)n^{r/2} \qquad r \geq 2.$$

When $r = 1$, the Mellin transform $F_r^\star(s)$ has a double pole at $s = 2$ and the asymptotic form of $F_r(t)$ at zero involves logarithmic terms. We then obtain, as $n \to \infty$,

$$\mathbb{E}\left[H_n\right] \sim 2\lambda^{-1}\sqrt{\pi n}$$
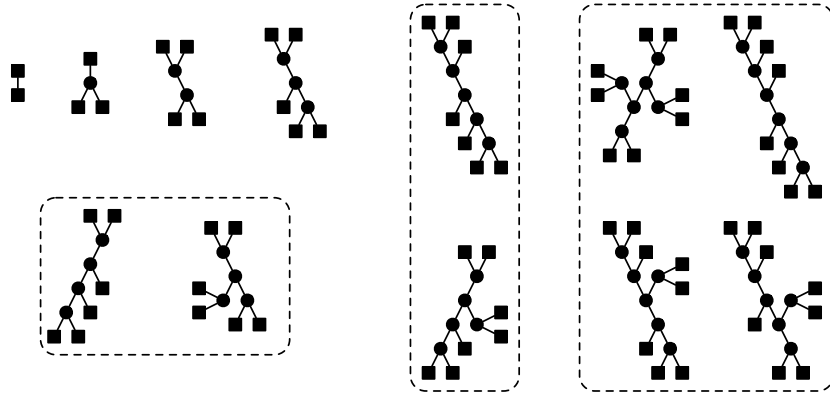
using similar arguments $\qquad\square$

Figure 6. The unlabelled trees of sizes from 2 to 8, with external nodes (leaves) represented by squares.

6. **The diameter of unrooted trees**

In this section, we put to use the approximations of Section 4 in order to quantify extreme distances in random unrooted trees. Developments parallel those of Riordan [39], as regards formal generating functions, and especially Szekeres [41], as regards asymptotic developments.

In the class $\mathcal{Y}$ of rooted binary trees, every node has total degree three or one, except for the root, which has degree two. Consider now the class $\mathcal{U}$ of *unrooted* ternary trees where each node has degree either three or one, without exception (no special root node is now distinguished). Let $\mathcal{U}_n$ be comprised of the elements of $\mathcal{U}$ with n nodes of degree one, the leaves, which determine *size*, hence $(n-2)$ nodes of degree three. Denote by u_n the number of such trees. The trees of $\mathcal{U}$ of size at most 8 are displayed in Figure 6. We write the generating function of $\mathcal{U}$ as $u(z) := \sum_{n \geq 0} u_n z^n$, so that

$$u(z) = z^2 + z^3 + z^4 + z^5 + 2z^6 + 2z^7 + 4z^8 + 6z^9 + 11z^{10} + 18z^{11} + \cdots,$$

and the coefficients constitute sequence A000672 of Sloane's *On-line Encyclopedia of Integer Sequences* [40].

Using considerations about the dissimilarity characteristic of trees found in Otter's work [35] and developed in [5, 25], we obtain

$$(62) \qquad u(z) = z^2 + u^{\bullet}(z) - \frac{1}{2}y(z)^2 + \frac{1}{2}y(z^2),$$

where $u^{\bullet}(z)$ is the generating function of unrooted trees with a distinguished node. (Note that because of the special degree condition in rooted trees $u^{\bullet}(z) \neq y(z)$.) The distinguished node might be a leaf or a node of degree three, which leads to

$$(63) \qquad u^{\bullet}(z) = z(y(z) - z) + \frac{1}{6}y(z)^3 + \frac{1}{2}y(z^2)y(z) + \frac{1}{3}y(z^3).$$

The equations (62) and (63) fully characterize $u(z)$ and, together with Lemma 1, they determine the singular expansion of $u(z)$. The following classical lemma reduces to simple manipulations based on Lemma 1, supplemented by routine singularity analysis of the generating function.

Lemma 10. *The generating function $u(z)$ of unrooted ternary trees expands in a neigh-bourhood of ρ as follows*

$$u(z) = \mu_0 + \mu_1(1 - z/\rho) + \frac{1}{3}\lambda^3(1 - z/\rho)^{3/2} + O\big((1 - z/\rho)^2\big),$$

for some constants $\mu_0, \mu_1 \in \mathbb{R}$ and $\lambda = \sqrt{2\rho + 2\rho^2 y'(\rho^2)}$. Furthermore, the number u_n of unrooted trees of size n satisfies the asymptotic estimate

$$u_n = \frac{\lambda^3}{4\sqrt{\pi}} \cdot n^{-5/2} \rho^{-n}\left(1 + O\left(\frac{1}{n}\right)\right).$$

We now turn to the analysis of the diameter of unrooted trees. A *diameter* in a graph or a tree is any simple path of maximal length and we also refer to the common length of all diameters as the *diameter* of the tree. Let $u_{d,n}$ be the number of unrooted trees on n leaves with diameter exactly equal to d, and let $u_d(z) = \sum_{n \geq 0} u_{d,n} z^n$ denote the associated generating function[9]. To simplify notations, we set

$$g_h(z) := e_{h-1}(z) - e_h(z),$$

which is the generating function of rooted unlabelled binary trees having height *exactly* h.

We have $u_1(z) = z^2$ and $u_2(z) = z^3$. Unrooted trees of size at least 4 may be recur-sively decomposed into sets of rooted trees; the decomposition depends on the parity of the diameter d. If $d = 2h + 1$ is odd, with $d \geq 3$, all diameters share a unique edge (bicentre) that splits the tree into a pair of two rooted trees of height exactly h each, so that

(64) $$u_{2h+1}(z) = \frac{1}{2}g_h(z)^2 + \frac{1}{2}g_h(z^2).$$

On the other hand, trees with even diameter $d = 2h$, with $d \geq 4$, decompose into three rooted trees around a central vertex (center), with two of the trees of height exactly h and a third subtree of height at most h:

$$u_{2h}(z) = \frac{1}{6}g_{h-1}(z)^3 + \frac{1}{2}g_{h-1}(z^2)g_{h-1}(z) + \frac{1}{3}g_{h-1}(z^3)$$

(65) $$+ \frac{1}{2}g_{h-1}(z)^2 y_{h-2}(z) + \frac{1}{2}g_{h-1}(z^2)y_{h-2}(z).$$

In this way, one can enumerate the trees of odd and even diameter (the "bicentred" and "centred" trees), whose generating functions start, respectively, as

$$\begin{array}{rcl}
u^{\mathrm{odd}}(z) &=& z^2 + z^4 + z^6 + +z^7 + 2z^8 + 2z^9 + 6z^{10} + 8z^{11} + \cdots \\
u^{\mathrm{even}}(z) &=& z^3 + z^5 + z^6 + z^7 + 2z^8 + 4z^9 + 5z^{10} + 10z^{11} + \cdots,
\end{array}$$

with coefficients forming sequences A000673 and A000675 of Sloane's *Encyclopedia*.

We now turn to singular asymptotics in a Δ–domain[10] (see [19, §VI.3] and Equation (5), and Figure 3). As usual, the Pólya terms in (64), (65), which are the ones containing func-tional terms involving z^2 or z^3, will turn out to be of negligible effect. Indeed, Lemma 2 guarantees, for $|z| < \sqrt{\rho}$:

(66) $$\left|g_h(z^2)\right| \leq \left|e_{h-1}(z^2)\right| \leq \frac{1}{\sqrt{h-1}}\left(\frac{|z|^2}{\rho}\right)^h.$$

---

[9] We reserve n for the size of trees, so that $u_n \equiv [z^n]u(z)$ is the number of trees of size n; we make use of indices $d, 2h, 2h + 1$ for diameter and occasionally abbreviate $u_d(z), \ldots$, as $u_d, \ldots$, so that no ambiguity should occur.

[10] To be precise, we only need to consider the *part* of a Δ-domain that is interior to a γ-contour of the type introduced in the previous section.

Thus, fixing some R with $\rho < R < \sqrt{\rho}$, we have for some $C > 0$ and $A \in (0, 1)$:

$$(67) \qquad \left| g_h(z^2) \right| < C \cdot A^h,$$

whenever z lies in a suitable Δ–domain anchored at ρ, and the same bound on the right of (67) obviously holds for $g_h(z^3)$. In other words, the Pólya terms involving z^2 and z^3 are exponentially small. This gives us, relative to (64) and (65) and for z in a Δ–domain, the estimate

$$(68) \qquad u_{2h+1}(z) = \frac{1}{2} g_h(z)^2 + O(A^h)$$

and, similarly,

$$(69) \qquad u_{2h}(z) = \frac{1}{2} g_{h-1}(z)^2 y_{h-2}(z) + \frac{1}{6} g_{h-1}(z)^3 + O(A^h).$$

The latter asymptotic form may be further simplified: by Lemmas 1 and 9, for $z \to \rho$ in a sandclock, we have

$$y - e_h = 1 - O(\sqrt{1 - z/\rho}) - e_h, \qquad |g_h| \le |e_{h-1}| = O(1/h),$$

and it follows that, in this sandclock,

$$(70) \qquad u_{2h}(z) = \frac{1}{2} g_{h-1}(z)^2 \Big(1 + O(1/h) + O(\sqrt{1 - z/\rho}) \Big).$$

(The cubic term $\frac{1}{6} g_{h-1}(z)^3$ essentially corresponds to trees having a centre from which there spring three trees of equal height; such configurations are still negligible, but now polynomially, rather than exponentially.) Additionally, in a tube, *all* terms in (68) and (69) are exponentially small, by virtue of Equation (15) of Lemma 3 and Proposition 1; the induced contributions for coefficients are thus going to be exponentially small, and we do not need to discuss these any further.

In a way similar to the asymptotic simplification (58) of $e_h - e_{h+1} \equiv g_{h+1}$, the estimates of (68) and (70) now suggest to introduce the following approximation of u_d,

$$(71) \qquad \widehat{u}_d := 2(1 - y)^4 \frac{y^d}{(1 - y^{d/2})^4},$$

regardless of the the parity of d: we have (in a sandclock)

$$(72) \qquad u_d = \widehat{u}_d \Big(1 + O(1/d) + O(\sqrt{1 - z/\rho}) \Big).$$

Following the line of proof of Theorems 1, 2, and 5, it is now a routine matter to work out the consequences, at the level of coefficients, of the main approximations (71) and (72). Note that, since we have access to generating functions of diameter *exactly* h, we start with a *local limit* law, then proceed to estimate the distribution function by summation. Figure 7 presents supporting numerical data for the local limit law of diameter.

Theorem 6 (Local limit law for diameter). *The diameter D_n of an unrooted tree sampled from $\mathcal{U}_n$ uniformly at random satisfies a local limit law: for x in any compact set of $\mathbb{R}_{>0}$, uniformly, with $(x/\lambda)\sqrt{n}$ an integer, one has:*

$$\lim_{n \to \infty} \mathbb{P} \left\{ D_n = (x/\lambda)\sqrt{n} \right\} = \frac{\lambda}{\sqrt{n}} \widetilde{\vartheta}(x)$$

where $\qquad \widetilde{\vartheta}(x) = \frac{1}{768} \sum_{k \ge 1} k(k^2 - 1) \big(k^5 x^5 - 80 k^3 x^3 + 960 k x \big) e^{-k^2 x^2/16}.$
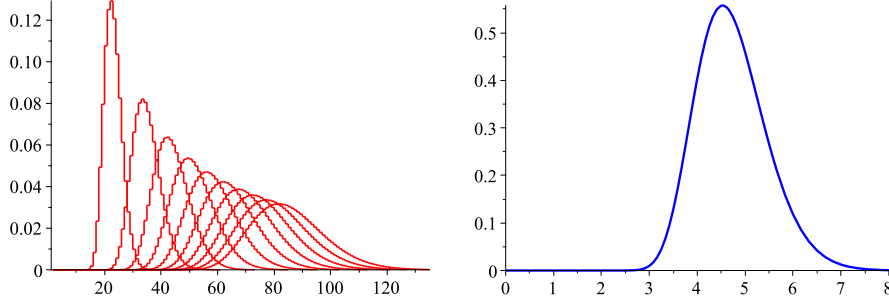
Figure 7. *Left*: the raw histograms of the distribution of diameter in unrooted trees, for $n = 50, 100, 150, \dots, 500$. *Right*: a plot of the limit density function $\widetilde{\theta}(x)$ of Theorem 6.

Proof. We start from the approximations (71) and (72), then make use of Cauchy's coefficient formula together with the contour γ specified in the proof of Theorem 1. As noted already, the contributions of the outer circle γ_3, the joins γ_4 and γ_5 and the further portions of the rectilinear pieces γ_1 and γ_2 are exponentially small, so that we can restrict attention to what happens in a small sandclock, along $\widetilde{\gamma}_1$ and $\widetilde{\gamma}_2$.

The change of variable $z = \rho(1 - t/n)$ and approximations that are justified in Equations (48) to (52) of the proof of Theorem 1 lead to

$$[z^n]\widehat{u}_d(z) = -2\rho^{-n}n^{-3}\lambda^4 J_2(\lambda x/2) + O(\rho^{-n}n^{-7/2}\log^\star n),$$

where we have set

$$J_2(X) := \frac{1}{2i\pi}\int_{\mathcal{L}}\frac{e^{-2X\sqrt{t}}}{(1 - e^{-X\sqrt{t}})^4}t^2 e^t\,dt,$$

with $\mathcal{L}$ that goes from $-\infty + i\infty$ to $-\infty - i\infty$ and winds to the right of the origin. As in Equations (53) and (54), we can make $J_2(X)$ explicit:

$$(73) \qquad J_2(X) = \frac{1}{2i\pi}\sum_{k\geq 3}\frac{k(k - 1)(k - 2)}{6}\int_{\mathcal{L}}e^{-X(k-1)\sqrt{t}}t^2 e^t\,dt$$

$$= -\frac{1}{192\sqrt{\pi}}\sum_{k\geq 2}k(k^2 - 1)\big(k^5 X^5 - 20k^3 X^3 + 60\big)e^{-k^2 X^2/4}.$$

A normalization by u_n, as provided by Lemma 10, then yields the claim. $\qquad\square$

Theorem 7 (Limit distribution of diameter). *The diameter D_n of a unrooted tree sampled from $\mathcal{U}_n$ uniformly at random admits a limit distribution: for x in a compact set of $\mathbb{R}_{>0}$, we have*

$$\lim_{n\to\infty}\mathbb{P}D_n \geq (x/\lambda)\sqrt{n} = \widetilde{\Theta}(x),$$

where $\widetilde{\Theta}(x) \equiv \int_x^\infty \widetilde{\vartheta}(w)\,dw = \frac{1}{96}\sum_{k\geq 1}(k^2 - 1)(k^4 x^4 - 48k^2 x^2 + 192)e^{-k^2 x^2/16}.$

Proof. The convergence of distribution functions results from earlier approximations through integration. Indeed, approximating a Riemann sum by the corresponding integral, we find, for $d = x\sqrt{n}$,

$$[z^n]\sum_{\ell\geq d}u_\ell \sim [z^n]\sum_{\ell\geq d}\widehat{u}_\ell \sim -2\lambda^4\rho^n n^{-3/2}\int_x^\infty J_2(\lambda s/2)ds,$$

as $n \to \infty$. The integral is easily computed from (73): write $X = \lambda x/2$ to obtain

$$\int_x^\infty J_2(\lambda s/2)ds = \frac{1}{3\lambda}\sum_{k\geq 1}(k^2-1)\frac{1}{2i\pi}\int_{\mathcal{L}} e^{kX\sqrt{t}}t^{3/2}e^t dt$$

$$= -\frac{1}{3\cdot 2^4\sqrt{\pi}}\sum_{k\geq 1}(k^2-1)(k^4X^4-12k^2X^2+12)e^{-k^2X^2/4}.$$

A final normalization based on Lemma 10 yields the result. $\qquad\square$

Theorem 8 (Moments of diameter). *The moments of the diameter D_n of a random un-rooted tree with n leaves satisfy*

$$\mathbb{E}\left[D_n\right] \sim \frac{8}{3\lambda}\sqrt{\pi n} \qquad \mathbb{E}\left[D_n^2\right] \sim \frac{16}{3\lambda^2}\left(1+\frac{\pi^2}{3}\right)n \qquad \mathbb{E}\left[D_n^3\right] \sim \frac{64}{\lambda^3}\sqrt{\pi n^3},$$

and, for all $r > 3$,

$$\mathbb{E}\left[D_n^r\right] \sim \frac{2^{2r}}{3}r(r-1)(r-3)\Gamma(r/2)(\zeta(r-2)-\zeta(r))\lambda^{-r}n^{r/2}.$$

Proof. By definition, the moments of D_n are given by

$$(74) \qquad\qquad \mathbb{E}\left[D_n^r\right] = \frac{1}{u_n}[z^n]\sum_{d\geq 1}d^r u_d(z),$$

and, from (71) and (72) once more, we are led to the approximation $\mathbb{E}\left[D_n^r\right] \sim ([z^n]\widehat{M_r})/u_n$, where

$$\widehat{M_r}(z) := 2(1-y)^4\sum_{d\geq 1}d^r \frac{y^d}{(1-y^{d/2})^4}$$

results from replacing u_d by $\widehat{u}_d$ in the generating function of (74).

As for the moments of height, the singular asymptotic form of $\widehat{M_r}(z)$ is conveniently determined by means of the Mellin transform technology. Set $y = e^{-\tau}$, so that $z \to \rho$ corresponds to $\tau \to 0$, with $\tau \sim \lambda\sqrt{1-z/\rho}$. We then need the asymptotic estimation of $\widehat{M_r}(e^{-\tau})$ when $\tau \to 0$. Define

$$F_r(\tau) := \sum_{d\geq 1}d^r \frac{e^{-d\tau}}{(1-e^{-d\tau/2})^4},$$

which is such that $M_r(z) \sim 2\lambda^4\tau^4 F_r(\tau)$. By the "harmonic sum rule" [21], the Mellin transform $F_r^\star(s)$ of $F_r(\tau)$ satisfies, for $\Re(s) > \max\{1+r, 4\}$,

$$F_r^\star(s) = \frac{2^s}{6}\zeta(s-r)(\zeta(s-3)-\zeta(s-1))\Gamma(s).$$

The singularities in a right half-plane are known to dictate the asymptotic expansion of $F_r(\tau)$, as $\tau \to 0$. For $r > 3$, the main contribution comes from a simple pole at $s = r+1$ (due to the factor $\zeta(s-r)$), and we find

$$F_r(\tau) \sim \frac{2^r}{6}(\zeta(r-2)-\zeta(r))\Gamma(r+1)\tau^{-r-1}, \qquad \tau \to 0,$$

which provides in turn the main term in the expansion of $\widehat{M_r}(z)$ as $z \to \rho$:

$$M_r(z) \sim \frac{2^r}{3}\lambda^{-r}\left(\zeta(r-2)-\zeta(r)\right)F(r+1)\left(1-z/\rho\right)^{-r+1}, \qquad z \to \rho.$$

	Cayley trees	Otter trees
mean depth	$\sqrt{\dfrac{\pi n}{2}}$	$\dfrac{1}{\lambda}\sqrt{\pi n}$
mean height	$\sqrt{2\pi n}$	$\dfrac{2}{\lambda}\sqrt{\pi n}$
mean diameter	$\dfrac{4}{3}\sqrt{2\pi n}$	$\dfrac{8}{3\lambda}\sqrt{\pi n}$

Figure 8. A table comparing the asymptotic forms of the expectations of several parameters of trees, for the two models of Cayley trees (non-plane labelled trees) and Otter trees (non-plane unlabelled binary trees), based on [33, 38, 41] and the present paper. *Depth* refers to the depth of a randomly chosen node in the tree; *height* is the maximum distance of any node from the root; *diameter* is relative to the unrooted version of the trees under consideration.

Singularity analysis combined with the estimate of u_n in Lemma 10 and the duplication formula for the Gamma–function then automatically yields the asymptotic form of $\mathbb{E}\left[D_n^r\right]$, in the case $r > 3$.

For $r \le 3$, the approach is similar, but a little more care is required. For $r = 1, 2$ one needs to consider the second terms of the singular expansion of $F_r^{\star}(s)$, at $s = 2$ and $s = 3$, respectively. Also, the cases $r = 1$ and $r = 3$ involve logarithmic terms due to double poles of $F_1^{\star}(s)$ and $F_3^{\star}(s)$ at $s = 2$ and $s = 4$. The claim follows by routine Mellin technology and singularity analysis. $\square$

7. **Conclusion**

We finally conclude with two corollaries and a general comment. First, as a byproduct of (71) and (72), via summation and singularity analysis, we can estimate the proportion of centred and bicentred trees.

Corollary 1. *There are asymptotically as many centred trees (trees of even diameter) as bicentred trees (trees of odd diameter):*

$$[z^n]u^{\mathrm{odd}}(z) \sim [z^n]u^{\mathrm{even}}(z) \sim \frac{1}{2}u_n.$$

This perhaps unsurprising observation parallels one made by Szekeres [41, p. 394] in the case of labelled trees, where all degrees are allowed.

Next, a comparison of expectations of height and diameter in random nonplane trees shows the following.

Corollary 2. *The ratio of the expected diameter of an unrooted tree and the expected height of a rooted tree of the same size satisfies asymptotically*

$$\lim_{n\to\infty} \frac{\mathbb{E}\left[D_n\right]}{\mathbb{E}\left[H_n\right]} = \frac{4}{3}.$$

Again, a similar observation was made by Szekeres [41, p. 396] regarding labelled trees and the same property, with a "universal" $\frac{4}{3}$ factor is expected to hold for any "ordered" tree family (i.e., trees whose nodes have neighbours that are distinguishable; cf our Introduction), as argued semi-heuristically by Aldous in [3].

The fact, established rigorously in the present paper (Theorems 1 to 8 and Corollaries 1, 2), is that, up to scaling, height and diameter behave for some non-plane unlabelled trees

similarly to what is known for ordered trees: see Figure 8 for some striking data. This brings further evidence for the hypothesis that probabilistic models, such as the *Continuum Random Tree*, may be applicable to unordered trees—this has indeed been recently confirmed, in the binary case at least, by Marckert and Miermont [32]. It is piquant to note that the probabilistic approach of [32] relies in part on large deviation estimates for height, which were developed analytically by us in the earlier conference version [7] of the present paper. (Recently, Haas and Miermont [24] have developed an alternative approach that further allows them to prove the convergence of a large class of trees towards continuum limits. This encompasses a self-contained proof of the result in [32] and other more examples with stable tree limits.) An analytic treatment of the height of unordered trees with all degrees allowed has been given recently by Drmota and Gittenberger (see [15] and the account in [14]). Together with the present study, it confirms, among unordered trees, the existence of universal phenomena regarding height and profile, which parallel what has been known for a long time regarding their ordered counterparts. As usual, the analytic approach advocated in the present paper has the advantage of providing precise estimates, with speed of convergence estimates, local limit laws, and convergence of moments.

Finally, the fact that, up to a possible linear change of scale, some of the main characteristics of trees, such as height and diameter, are not sensitive to whether trees are planar (ordered) or not, is also of some relevance to the emerging field of "probabilistic logic" [29, 31]. For instance, there is interest there in determining the probability of satisfiability of random boolean formulae obeying various randomness models (see, e.g., [10, 23]). In this context, our results suggest that the commutativity of logical conjunction and disjunction (reflected by the non-planarity of associated expression trees) should not, in many cases, have a major effect on complexity properties of random Boolean expressions.

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] D. J. Aldous. The continuum random tree II: an overview. In M. T. Barlow and N. H. Bingham, editors, *Proccedings of the 1990 Durham Symposium on Stochastic Analysis*, volume 21 of *London Math. Soc. Lecture Note Ser.*, pages 23–70. Cambridge University Press, 1991.

[4] D. J. Aldous. The continuum random tree III. *The Annals of Probability*, 21:248–289, 1993.

[5] F. Bergeron, G. Labelle, and P. Leroux. *Combinatorial species and tree-like structures*. Cambridge University Press, 1998. ISBN 0-521-57323-8.

[6] M. Bóna and P. Flajolet. Isomorphism and symmetries in random phylogenetic trees. *Journal of Applied Probability*, 46:1005–1019, 2009.

[7] N. Broutin and P. Flajolet. The height of random binary unlabelled trees. In U. Rösler, editor, *Proceedings of Fifth Colloquium on Mathematics and Computer Science: Algorithms, Trees, Combinatorics and Probabilities*, volume AI of *Discrete Mathematics and Theoretical Computer Science Proceedings*, pages 121–134, Blaubeuren, 2008.

[8] 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. ISSN 1077-8926.

[9] 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, 2000.

[10] B. Chauvin, P. Flajolet, D. Gardy, and B. Gittenberger. And/Or Trees Revisited. *Combinatorics, Probability and Computing*, 13(4–5):501–513, 2004. Special issue on Analysis of Algorithms.

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

[12] 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.

[13] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*. Jones and Bartlett Publishers, Boston and London, 1993.

[14] M. Drmota. *Random Trees*. Springer Verlag, 2009.

[15] M. Drmota and B. Gittenberger. The shape of unlabeled rooted random trees. Technical report, Technical University of Vienna, 2008. Revised version available as Arxiv preprint `arXiv:1003.1322`, 2010.

[16] R. T. Durrett and D. L. Iglehart. Functionals of Brownian meander and Brownian excursion. *The Annals of Probability*, 5(1):130–135, 1977.

[17] S. Finch. *Mathematical Constants*. Cambridge University Press, 2003.

[18] 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.

[19] P. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 2009. URL http://algo.inria.fr/flajolet. 824 pages. Also available electronically from the authors' home pages.

[20] 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.

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

[22] 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.

[23] D. Gardy. Random Boolean expressions. In *Computational Logic and Applications (CLA'05)*, volume AF of *Discrete Mathematics and Theoretical Computer Science Proceedings*, pages 1–36, 2005.

[24] B. Haas and G. Miermont. Scaling limits of Markov branching trees, with applications to Galton–Watson and random unordered trees. Technical Report 1003.3632, arXiv, 2010.

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

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

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

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

[29] Z. Kostrzycka and M. Zaionc. Asymptotic densities in logic and type theory. *Studia Logica*, 88(3):385–403, 2008.

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

[31] H. Lefmann and P. Savický. Some typical properties of large AND/OR Boolean formulas. *Random Structures & Algorithms*, 10:337–351, 1997.

[32] J.-F. Marckert and G. Miermont. The CRT is the scaling limit of unordered binary trees. *Random Structures & Algorithms*, 2010. In press. 35 pages.

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

[34] J. Milnor. *Dynamics in one complex variable*. Friedr. Vieweg & Sohn, 1999. ISBN 3-528-03130-1.

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

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

[37] G. Pólya and R. C. Read. *Combinatorial Enumeration of Groups, Graphs and Chemical Compounds*. Springer Verlag, 1987.

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

[39] J. Riordan. Enumeration of trees by height and diameter. *IBM Journal of Research and Development*, 4:473–478, 1960.

[40] N. J. A. Sloane. *The On-Line Encyclopedia of Integer Sequences*. 2008. Published electronically at `www.research.att.com/˜njas/sequences/`.

[41] G. Szekeres. Distribution of labelled trees by diameter. In *Combinatorial Mathematics X*, Lecture Notes in Mathematics, pages 392–397. Springer, 1982.

[42] E. T. Whittaker and G. N. Watson. *A Course of Modern Analysis*. Cambridge University Press, fourth edition, 1927. Reprinted 1973.

P.F., N.B.: ALGORITHMS PROJECT, INRIA-ROCQUENCOURT, F-78153 LE CHESNAY (FRANCE)