THE SHAPE OF UNLABELED ROOTED RANDOM TREES

MICHAEL DRMOTA AND BERNHARD GITTENBERGER

Abstract. We consider the number of nodes in the levels of unlabelled rooted random trees and show that the stochastic process given by the properly scaled level sizes weakly converges to the local time of a standard Brownian excursion. Furthermore we compute the average and the distribution of the height of such trees. These results extend existing results for conditioned Galton-Watson trees and forests to the case of unlabelled rooted trees and show that they behave in this respect essentially like a conditioned Galton-Watson process.

1. Introduction

We consider the profile and height of unlabelled rooted random trees. This kind of trees is also called Pólya trees, because the enumeration theory developed by Pólya allows an analytical treatment of this class of trees by means of generating functions (see [43]). The profile of a rooted tree T is defined as follows. First we define the k-th level of T to be the set of all nodes having distance k from the root (where we use the usual shortest path graph metric). Let $L_k(T)$ denote the number of nodes of the k-th level. The *profile* of T is the sequence $(L_k(T))_{k\geq 0}$. For a random tree this sequence becomes a stochastic process.

The first investigations of the profile of random trees seem to go back to Stepanov [46] who derived explicit formulas for the distribution of the size of one level. Further papers deal mainly with simply generated trees as defined by Meir and Moon [36]. Note that simply generated trees are defined by the functional equation

$$y(x) = x\phi(y(x)) \tag{1}$$

for their generating function but can also be viewed as family trees of a Galton-Watson process conditioned on the total progeny. Kolchin (see [32, 33]) related the level size distributions to distributions occurring in particle allocation schemes. Later Takács [47] derived another expression for the level sizes by means of generating functions. Aldous [1] conjectured two functional limit theorems for the profile in two different ranges which were proved in [13, 23]. The first author [10] studied restrictions of the profile to nodes of fixed degree. An extension to random forests of simply generated trees is given by the second author [24].

Later other tree classes have been considered as well. The profile of random binary search trees has been first studied by Chauvin et al. [5] and later by Drmota and Hwang (see [11] and [16]). Random recursive trees have been investigated recently by Drmota and Hwang [17] and van der Hofstad et al. [49]. Related research was done by Chauvin et al. [5], Fuchs et al. [22], Hwang [29, 30], Louchard et al. [34], and Nicodème [39]. Extremal studies of the profile (called the width of trees) of simply generated trees have been started by Odlyzko and Wilf [40]. The distribution including moment convergence has been presented independently in Chassaing et al. [4] and the authors of this paper [15]. For other tree classes we refer to the work of Devroye and Hwang [9] and Drmota and Hwang [17]. A general overview on random trees which also strongly highlights the profile of trees can be found in the first author's book on random trees [12].

Date: March 8, 2010.

Key words and phrases. unlabelled trees, profile, height, Brownian excursion, local time.

Address: Institut für Diskrete Mathematik und Geometrie, Technische Universität Wien, Wiedner Hauptstr. 8-10/104, A-1040 Wien, Austria.

Email: michael.drmota@tuwien.ac.at, gittenberger@dmg.tuwien.ac.at.

This work has been supported by the Austrian Science Foundation FWF, NFN-Project S9604, as well as by the EU FP6-NEST-Adventure Programme, contract no. 028875 (NEST).

Whereas simply generated trees have an average height of order $\sqrt{n}$, the other tree classes mentioned above have height of order $\log n$. Pólya trees do not belong to the class of simply generated trees. Since we are not aware of any rigorous proof of this assertion in the literature, we will present a (rather simple) proof of this in the next section. To our knowledge, so far the fact was only underpinned by the following argument concerning the generating functions of both tree classes. The argument works for many tree classes considered in the literature, e.g. Cayley trees, plane trees, Motzkin trees, binary trees, and many more. In all these cases the function $\phi(y)$ in (1) is entire function or meromorphic. But it is not at all clear that this argument is still true if $\phi(y)$ has a more complicated singularity structure. In these cases (i.e., entire or meromorphic $\phi(y)$), the generating functions enumerating the number of simply generated trees and Pólya trees, respectively, have a fundamentally different singularity structure. Whereas the first one has one or a finite number of singularities (the latter occurs in the periodic case) on the circle of convergence and allows analytic continuation to a slit plane (with the possible exception of finitely many isolated singularities which are of algebraic type even if the function itself is not algebraic), the generating function associated to Pólya trees is much more complicated. In fact, for the latter function the unit circle is a natural boundary (i.e., no analytic continuation beyond it is possible). There is exactly one singularity on the circle of convergence of the power series expansion at 0, but the analytic continuation has an infinite number of singularities inside the unit circle. Each point on the unit circle is an accumulation point of the set of singularities. These facts follow from the functional equation defining this generating function and the fact that the power series expansion around zero has radius of convergence strictly smaller than one (see next section). It also involves an analytically complicated structure like the cycle index of the symmetric group. Due to this difference with respect to the analytic behaviour of the generating function Pólya trees are not simply generated and therefore they cannot be represented as branching processes.

Note that the rather complicated singularity structure does not affect the asymptotics of statistical parameters like the number of trees, the profile or the height. The behaviour of these parameters is determined by the dominant singularity which is, as all other singularities as well, an isolated singularity, exactly as in the case of simply generated trees. Indeed, Pólya trees behave in many respects similar to simply generated trees (compare with [44, 27, 37, 38, 14, 25]) or the recent work of Marckert and Miermont [35] who showed that binary unlabelled trees converge in some sense to the continuum random tree, i.e., the same limit as that of simply generated trees. Moreover, Broutin and Flajolet [3] showed $\sqrt{n}$-behaviour for the height of binary unlabelled trees. Hence it is expected that the order of the height is $\sqrt{n}$ as well. In this paper we will give an affirmative answer to this question. This justifies the choice of $\sqrt{n}$ for the scaling of the level sizes in the subsequent theorems.

The plan of the paper is as follows. In the next section we present our main results. Then we will set up the generating functions for our counting problem of trees with nodes in certain levels marked. This function is given as solution of a recurrence relation which has to be analyzed in detail. Knowing the singular behaviour of the considered generating functions allows us to show that the finite dimensional distributions (fdd's) of the profile, i.e., the distributions of the sizes of several levels considered simultaneously, converge to the fdd's of Brownian excursion local time. The singularity analysis is carried out in Section 4 and the computation of the fdd's in Section 5. In order to complete the functional limit theorem we need to prove tightness. This means, roughly speaking, that the sample paths of the process do not have too strong fluctuations (see [2] for the general theory).

In the final section we turn to the height. The pioneering work on this topic was done by Flajolet and Odlyzko [20] and Flajolet et al. [19] in their studies of simply generated trees where they completed the program started in [8]. What we have to do is to show that the generating function appearing in the analysis of the height has a local structure which is amenable to the steps carried out in [20] and [19]. This is done in the last section and leads to average and distribution of the height.

2. Preliminaries and Results

First we collect some results for unlabelled unrooted trees. Let $\mathcal{Y}_n$ denote the set of unlabelled rooted trees consisting of n vertices and y_n be the cardinality of this set. Pólya [43] already discussed the generating function

$$y(x) = \sum_{n \geq 1} y_n x^n$$

and showed that the radius of convergence ρ satisfies $0 < \rho < 1$ and that $x = \rho$ is the only singularity on the circle of convergence $|z| = \rho$. He also showed that $y(x)$ satisfies the functional equation

$$y(x) = x \exp\left(\sum_{i \geq 1} \frac{y(x^i)}{i} \right). \tag{2}$$

Nowadays, this functional equation is easily derived by using the theory of combinatorial constructions which is presented in the comprehensive book of Flajolet and Sedgewick [21]. Indeed, a Pólya tree can be viewed as a root with a multiset of Pólya trees attached to it. Then the functional equation (2) pops out immediately from the multiset construction and its generating function. This functional equation can be used to compute the coefficients:

$$y(x) = x + x^2 + 2x^3 + 4x^4 + 9x^5 + 20x^6 + 48x^7 + 115x^8 + 286x^9 + 719x^{10} + \dots \tag{3}$$

Later Otter [41] showed that $y(\rho) = 1$ as well as the asymptotic expansion

$$y(x) = 1 - b(\rho - x)^{1/2} + c(\rho - x) + d(\rho - x)^{3/2} + \cdots \tag{4}$$

which he used to deduce that

$$y_n \sim \frac{b\sqrt{\rho}}{2\sqrt{\pi}} n^{-3/2} \rho^{-n}. \tag{5}$$

Furthermore he calculated the first constants appearing in this expansion: $\rho \approx 0.3383219$, $b \approx 2.6811266$, and $c = b^2/3 \approx 2.3961466$.

We will return to the function $y(x)$ in Section 4 and list a couple of useful properties in Lemma 1 after introducing some notations.

Theorem 1. *Pólya trees are not simply generated.*

Proof: Let us assume that Pólya tree are simply generated. Then the generating function $y(x)$ given by (2) must have a representation in the form (1) where $\phi(y)$ is a power series with non-negative coefficients. By (3) the functional inverse $y^{-1}(x)$ exists and we have $y(x) \sim x$ and $y^{-1}(x) \sim x$, as $x \to 0$. This implies $\phi(0) = 1$. Plugging $y^{-1}(x)$ into (1) we obtain

$$x = y(y^{-1}(x)) = y^{-1}(x)\phi(y(y^{-1}(x))) = y^{-1}(x)\phi(x)$$

and consequently

$$\begin{aligned} \phi(x) = \frac{x}{y^{-1}(x)} &= \frac{x}{x - x^2 + x^4 - x^5 + x^6 - 4x^7 + 11x^8 - 18x^9 + 18x^{10} + \dots} \\ &= 1 + x + x^2 - x^5 + 3x^6 - 5x^7 + 7x^8 - 8x^9 + x^{10} + \dots \end{aligned}$$

which violates the requirement of non-negative coefficients for $\phi(x)$. □

Remark. The sequence of the coefficients of $y(x)$ is A000081, that of $y^{-1}(x)$ is A050395 in Sloane's On-line Encyclopedia of Integer Sequences [45]

The height of a tree is the maximal number of edges on a path from the root to another vertex of the tree. It turns out that the average height is of order $\sqrt{n}$.

Theorem 2. *Let H_n denote the height of an unlabelled rooted random tree with n vertices. Then we have*

$$\mathbf{E}H_n \sim \frac{2\sqrt{\pi}}{b\sqrt{\rho}}\sqrt{n}. \tag{6}$$

and

$$\mathbf{E}H_n^r \sim \left(\frac{2}{b\sqrt{\rho}}\right)^r r(r-1)\Gamma(r/2)\zeta(r)\, n^{r/2} \tag{7}$$

for every integer $r \geq 2$.

The proof of this theorem is deferred to the last section, since the proofs of the auxiliary lemmas which will eventually establish the assertion will utilize similar techniques as needed to prove the next three theorems.

Remark. Note that more information on the limiting distribution is available. Indeed, a local limit theorem holds as well. Let $y_n^{(h)}$ denote the number of unlabelled rooted trees with n vertices and height equal to h and let $\delta > 0$ arbitrary but fixed. If we set $\beta = 2\sqrt{n}/hb\sqrt{\rho}$, then, as $n \to \infty$, we have

$$\mathbf{P}\left\{H_n = h\right\} = \frac{y_n^{(h)}}{y_n} \sim 4b\sqrt{\frac{\rho\pi^5}{n}}\beta^4 \sum_{m\geq 1} m^2(2m^2\pi^2\beta^2 - 3)e^{-m^2\pi^2\beta^2}$$

uniformly for $\frac{1}{\delta\sqrt{\log n}} \leq \frac{h}{\sqrt{n}} \leq \delta\sqrt{\log n}$. A rigorous proof of this theorem was given by Broutin and Flajolet [3] for binary unlabelled trees. They also provide a moment convergence theorem, a weak limit theorem as well as large deviation results.

Let $L_n(t)$ denote the number of nodes at distance t from the root of a randomly chosen unlabelled rooted tree of size n. If t is not an integer, then define $L_n(t)$ by linear interpolation:

$$L_n(t) = (\lfloor t\rfloor + 1 - t)L_n(\lfloor t\rfloor) + (t - \lfloor t\rfloor)L_n(\lfloor t\rfloor + 1), \quad t \geq 0. \tag{8}$$

We will show the following theorem.

Theorem 3. *Let*

$$l_n(t) = \frac{1}{\sqrt{n}}L_n\left(t\sqrt{n}\right).$$

Then $l_n(t)$ satisfies the following functional limit theorem:

$$(l_n(t))_{t\geq 0} \xrightarrow{w} \left(\frac{b\sqrt{\rho}}{2\sqrt{2}} \cdot l\left(\frac{b\sqrt{\rho}}{2\sqrt{2}} \cdot t\right)\right)_{t\geq 0}$$

in $C[0,\infty)$, as $n \to \infty$. Here b and ρ are the constants of Equation (4) and $l(t)$ denote the local time of a standard scaled Brownian excursion.

In order to prove this result we have to show the following two theorems

Theorem 4. *Let b, ρ, and $l_n(t)$ be as in Theorem 3, then for any d and any choice of fixed numbers $t_1, \ldots, t_d$ the following limit theorem holds:*

$$(l_n(t_1), \ldots, l_n(t_d)) \xrightarrow{w} \frac{b\sqrt{\rho}}{2\sqrt{2}}\left(l\left(\frac{b\sqrt{\rho}}{2\sqrt{2}} \cdot t_1\right), \ldots, l\left(\frac{b\sqrt{\rho}}{2\sqrt{2}} \cdot t_d\right)\right),$$

as $n \to \infty$.

Theorem 5. *For all non-negative integers n, r, h we have*

$$\mathbf{E}\left(L_n(r) - L_n(r+h)\right)^4 \leq C\, h^2 n \tag{9}$$

where C denotes some fixed positive constant. Consequently, the process $l_n(t)$ is tight.

3. Combinatorial Setup

In order to compute the distribution of the number of nodes in some given levels in a tree of size n we have to calculate the number $y_{k_1 m_1 k_2 m_2 \cdots k_d m_d n}$ of trees of size n with m_i nodes in level k_i, $i = 1, \ldots, d$ and normalize by y_n.

Therefore we introduce the generating functions $y_k(x,u)$ defined by the recurrence relation

$$\begin{aligned} y_0(x,u) &= u y(x) \\ y_{k+1}(x,u) &= x \exp\left(\sum_{i \geq 1} \frac{y_k(x^i, u^i)}{i} \right), \quad k \geq 0. \end{aligned} \tag{10}$$

The function $y_k(x,u)$ represents trees where the nodes in level k are marked (and counted by u). If we want to look at two levels at once, say k and ℓ, then we have to take trees with height at most k and substitute the leaves in level k by trees with all nodes at level $\ell - k$ marked (counted by v) and marking their roots as well (counted by u). This leads to the generating function $y_{k,\ell}(x,u,v) = \tilde{y}_{k,\ell-k}(x,u,v)$ satisfying the recurrence relation

$$\begin{aligned} \tilde{y}_{0,\ell}(x,u,v) &= u y_\ell(x,v) \\ \tilde{y}_{k+1,\ell}(x,u,v) &= x \exp\left(\sum_{i \geq 1} \frac{\tilde{y}_{k,\ell}(x^i, u^i, v^i)}{i} \right), \quad k \geq 0. \end{aligned} \tag{11}$$

In general we get therefore

$$\begin{aligned} y_{k_1,\ldots,k_d}(x,u_1,\ldots,u_d) &= \sum_{m_1,\ldots,m_d,n \geq 0} y_{k_1 m_1 k_2 m_2 \cdots k_d m_d n} u_1^{m_1} \cdots u_d^{m_d} x^n \\ &= \tilde{y}_{k_1, k_2 - k_1, k_3 - k_2, \ldots, k_d - k_{d-1}}(x, u_1, \ldots, u_d) \end{aligned}$$

where

$$\begin{aligned} \tilde{y}_{0,\ell_2,\ldots,\ell_d}(x,u_1,\ldots,u_d) &= u_1 \tilde{y}_{\ell_2,\ldots,\ell_d}(x,u_2,\ldots,u_d) \\ \tilde{y}_{k+1,\ell_2,\ldots,\ell_d}(x,u_1,\ldots,u_d) &= x \exp\left(\sum_{i \geq 1} \frac{\tilde{y}_{k,\ell_2,\ldots,\ell_d}(x^i, u_2^i, \ldots, u_d^i)}{i} \right), \quad k \geq 0. \end{aligned}$$

The coefficients of these function are related to the process $L_n(t)$ (see (8) and the lines before) by

$$\mathbf{P}\left\{ L_n(k_1) = m_1, L_n(k_1 + k_2) = m_2, \ldots, L_n(k_1 + k_2 + \cdots + k_d) = m_d \right\} = \frac{y_{k_1 m_1 k_2 m_2 \cdots k_d m_d n}}{y_n}$$

where the k_i are integers and the probability space of the measure $\mathbf{P}$ is the set of Pólya trees with n vertices equipped with the uniform distribution.

As claimed in Theorem 3, the process $l_n(t) = \frac{1}{\sqrt{n}} L_n(t)$ converges weakly to Brownian excursion local time. From [28] (cf. [6, 13] as well) we know that the characteristic function $\phi(t)$ of the total local time of a standard Brownian excursion at level κ is

$$\phi(t) = 1 + \frac{\sqrt{2}}{\sqrt{\pi}} \int_\gamma \frac{t\sqrt{-s}\exp(-\kappa\sqrt{-2s}\,)}{\sqrt{-s}\exp(\kappa\sqrt{-2s}\,) - it\sqrt{2}\sinh\left(\kappa\sqrt{-2s}\,\right)} e^{-s}\, ds \tag{12}$$

where $\gamma = (c - i\infty, c + i\infty)$ with some arbitrary $c < 0$. The characteristic function of the joint distribution of the local time at several levels $\kappa_1, \ldots, \kappa_d$ was computed in [13] (for $d = 2$ already in [6] albeit written down in a form which does not exhibit the recursive structure) and is given by

$$\phi_{\kappa_1 \ldots \kappa_d}(t_1, \ldots, t_d) = 1 + \frac{\sqrt{2}}{i\sqrt{\pi}} \int_\gamma f_{\kappa_1,\ldots,\kappa_d}(x, t_1, \ldots, t_d) e^{-x}\, dx, \tag{13}$$

where

$$f_{\kappa_1,\ldots,\kappa_p}(x, t_1, \ldots, t_d) = \Psi_{\kappa_1}(x, t_1 + \Psi_{\kappa_2 - \kappa_1}(\ldots \Psi_{\kappa_{d-1} - \kappa_{d-2}}(x, t_{d-1} + \Psi_{\kappa_d - \kappa_{d-1}}(x, t_d)) \cdots))$$

with

$$\Psi_\kappa(x,t) = \frac{it\sqrt{-x}\exp(-\kappa\sqrt{-2x}\,)}{\sqrt{-x}\exp(\kappa\sqrt{-2x}\,) - it\sqrt{2}\sinh\left(\kappa\sqrt{-2x}\right)}.$$

For further studies of this distribution see [26, 42, 48].

In order to show the weak limit theorem we have to show pointwise convergence of the characteristic function $\phi_{k_1,\cdots,k_d,n}(t_1,\dots,t_d)$ of the joint distribution of $\frac{1}{\sqrt{n}}L_n(k_1),\dots,\frac{1}{\sqrt{n}}L_n(k_d)$ to the corresponding characteristic function of the local time in some interval containing zero. We have

$$\phi_{k_1,\cdots,k_d,n}(t_1,\dots,t_d) = \frac{1}{y_n}[x^n]y_{k_1,\dots,k_d}\left(x, e^{it_1/\sqrt{n}},\dots e^{it_d/\sqrt{n}}\right).$$

This coefficient will be calculated asymptotically by singularity analysis (see [18]) for $k_j = \lfloor \kappa_j\sqrt{n}\rfloor$. Thus knowing the local behaviour of $y_k(x,u)$ near its dominant singularity is the crucial step in proving Theorem 3. This is provided by the following theorem will be the crucial step of the proof.

Theorem 6. *Set $w_k(x,u) = y_k(x,u) - y(x)$. Let $x = \rho\left(1+\frac{s}{n}\right)$, $u = e^{it/\sqrt{n}}$, and $k = \lfloor\kappa\sqrt{n}\rfloor$. Moreover, assume that $|\arg s| \ge \theta > 0$ and, as $n\to\infty$, we have $s = O\left(\log^2 n\right)$ whereas t and κ are fixed. Then $w_k(x,u)$ admits the local representation*

$$w_k(x,u) \sim \frac{b^2\rho}{2\sqrt{n}}\cdot\frac{it\sqrt{-s}\exp\left(-\kappa b\sqrt{-\rho s}\,\right)}{\sqrt{-s} - \frac{itb\sqrt{\rho}}{4}\left(1-\exp\left(-\kappa b\sqrt{-\rho s}\,\right)\right)} \tag{14}$$

$$= \frac{b\sqrt{2\rho}}{\sqrt{n}}\Psi_{\kappa b\sqrt{\rho}/(2\sqrt{2})}\left(s, \frac{itb\sqrt{\rho}}{2\sqrt{2}}\right) \tag{15}$$

uniformly for $k = O\left(\sqrt{n}\right)$.

The proof is deferred to the next section.

Note that Theorem 6 implies Theorem 4 for the case $d=1$. If we set $d=1$ in Theorem 4, then we have

$$\begin{aligned}\phi_{k;n}(t) &= \frac{1}{y_n}[x^n]y_k\left(x, e^{it/\sqrt{n}}\right)\\ &= \frac{1}{2\pi i y_n}\int_\Gamma y_k\left(x, e^{it/\sqrt{n}}\right)\frac{dx}{x^{n+1}}\\ &= 1 + \frac{1}{2\pi i y_n}\int_\Gamma w_k\left(x, e^{it/\sqrt{n}}\right)\frac{dx}{x^{n+1}}\end{aligned}$$

where Γ is a suitable closed contour encircling the origin. Using Theorem 6 it is easy to show that $\phi_{\kappa\sqrt{n};n}(t)$ converges to $\phi_{\kappa b\sqrt{\rho}/(2\sqrt{2})}(tb\sqrt{\rho}/2\sqrt{2})$ as desired.

The higher dimensional case is more involved, but relies on the same principles. A complete proof is given in Section 5.

4. The Local Behaviour of $y_k(x,u)$ – Proof of Theorem 6

4.1. **Notation.** We will provide some frequently used notations now.

We will study the local behaviour of y_k by analyzing the quantity

$$w_k(x,u) = y_k(x,u) - y(x)$$

which frequently involves the term

$$\Sigma_k(x,u) := \sum_{i\ge 2}\frac{w_k(x^i,u^i)}{i}.$$

Furthermore, estimates of the partial derivatives

$$\gamma_k(x,u) = \frac{\partial}{\partial u}y_k(x,u) \text{ and } \gamma_k^{[i]}(x,u) = \frac{\partial^i}{\partial u^i}y_k(x,u), \qquad i\ge 2,$$

will be needed.

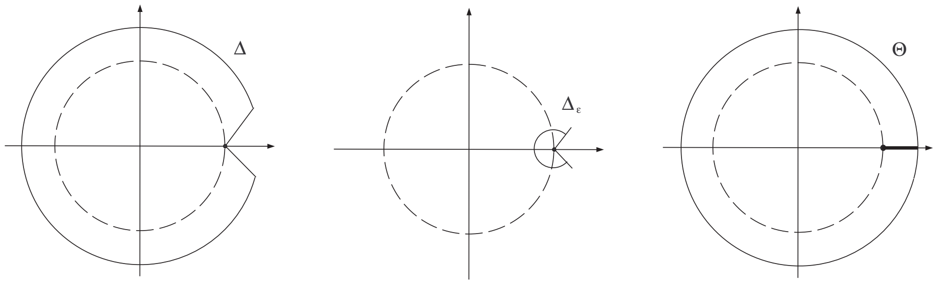

Figure 1: The domains Δ, Δ_ε, and Θ. For analyticity arguments we will need the domain Δ. Sometimes, the proof goes without change for the larger domain Θ as well. In that case we will use Θ for the sake of more generality. For arguments where the asymptotic behaviour near the singularity is important the domain Δ_ε is used. The domains Ξ_k (not depicted here) are needed whenever the asymptotic behaviour for $u \sim 1$ is considered. Here uniformity is often important such that a universal ε-neighbourhood of $u = 1$ which is independent of k does not serve our needs.

The asymptotic analysis of w_k (resp. y_k) enables us to apply Cauchy's integral formula and get the coefficients of $y_k(x,u)$ asymptotically (see the proof of Theorem 4 in the next section) which eventually leads to an integral of the form (12). Therefore estimates for $y(x)$, provided in the next lemma, and the other functions appearing in our analysis are needed. The estimates will be valid in various domains. Therefore let us introduce

$$\Delta = \{x \in \mathbb{C} : |x| < \rho + \eta,\ |\arg(x-\rho)| > \theta\}, \tag{16}$$

$$\Delta_\varepsilon = \{x \in \mathbb{C} : |x-\rho| < \varepsilon,\ |\arg(x-\rho)| > \theta\} \tag{17}$$

$$\Theta = \{x \in \mathbb{C} : |x| < \rho + \eta,\ |\arg(x-\rho)| \neq 0\}$$

$$\Xi_k = \{u \in \mathbb{C} : |u| \le 1, k|u-1| \le \tilde{\eta}\} \tag{18}$$

with $\varepsilon, \eta, \tilde{\eta} > 0$ and $0 < \theta < \frac{\pi}{2}$.

Remark. In all the arguments which we will use in the following proofs it is always assumed that ε and η are sufficiently small even if it is not explicitly mentioned.

4.2. **Analysis of the local behaviour of $y_k(x,u)$.** Obviously, $w_k(x,1) \equiv 0$. Since $y_k(x,u)$ represents the set of trees where the vertices of level k are marked, we expect that $\lim_{k\to\infty} y_k(x,u) = y(x)$ inside the domain of convergence. This is not obvious, but follows from what we derive in the sequel. We start with a useful property of $y(x)$.

Lemma 1. *Provided that η in (16) is sufficiently small, the generating function $y(x)$ has the following properties:*

a) *For $x \in \Delta$ we have $|y(x)| \le 1$. Equality holds only for $x = \rho$.*
b) *Let $x = \rho\left(1 - \frac{1+it}{n}\right)$ and $|t| \le C \log^2 n$ for some fixed $C > 0$. Then there is a $c > 0$ such that*
$$|y(x)| \le 1 - c\sqrt{\frac{\max(1,|t|)}{n}}.$$
c) *For $|x| \le \rho$ we have $|y(x)| \le y(|x|) \le 1$. Moreover, near $x = 0$ the asymptotic relation $y(x) \sim x$ holds.*
d) *There exists an $\varepsilon > 0$ such that*
$$|y(x)| \ge \min\left(\frac{\varepsilon}{2}, \frac{|x|}{2}\right) \tag{19}$$
for all $x \in \Theta$.

Proof: The first statement, when restricted to $|x| \le \rho$, follows from the facts that $y(x)$ has only positive coefficients (except $y_0 = 0$), $y(\rho) = 1$ and there are no periodicities. Extension to Δ is easily established by using (4) and continuity arguments.

The second statement is an immediate consequence of the singular expansion (4) of $y(x)$.

The first inequality of the third statement follows from the positivity of the coefficients y_n. The same fact also implies that $y(x)$ is strictly increasing in the interval $[0, \rho]$ and therefore bounded by $y(\rho) = 1$. The asymptotic relation near 0 follows from $y_0 = 0$ and $y_1 = 1$ or from the functional equation (2)

Finally, for proving the last statement we split the circle $|x| \le \rho$ into the smaller circle $|x| < \varepsilon$ (let us call it C) and the annulus $A := \{x : \varepsilon \le |x| \le \rho\}$. The function $y(x)$ is analytic in $C \cup A$ except at $x = \rho$, but still continuous there. Since the annulus A is compact, $|y(x)|$ attains a minimum there. By the functional equation (2) this minimum must be positive, since both factor on the right-hand side of (2) are nonzero. Since $y(x) \sim x$, as $x \to 0$, a sufficiently small ε guarantees that $\min_{|x|=\varepsilon} |y(x)| > \varepsilon/2$. Since this inequality holds for every smaller ε as well, we can choose an ε such that

$$\min_{\varepsilon \le |x| \le \rho} |y(x)| > \frac{\varepsilon}{2} \quad \text{and} \quad |y(x)| \ge \frac{|x|}{2} \text{ for } |x| < \varepsilon$$

hold. By continuity this can be extended to (19). □

Next we derive an *a priori* estimate of w_k in a small domain.

Lemma 2. *Let $|x| \le \rho^2 + \varepsilon$ for sufficiently small ε and $|u| \le 1$. Then there exist a constant L with $0 < L < 1$ and a positive constant C such that*

$$|w_k(x, u)| \le C|u - 1| \cdot |x| \cdot L^k$$

for all non-negative integers k.

Proof: We first note that by using the recurrence relation (10) we obtain

$$\begin{aligned} w_{k+1}(x, u) &= y_{k+1}(x, u) - y(x) \\ &= x \exp\left(\sum_{i \ge 1} \frac{1}{i} y_k(x^i, u^i)\right) - y(x) \\ &= y(x) \left(\exp\left(w_k(x, u) + \sum_{i \ge 2} \frac{w_k(x^i, u^i)}{i}\right) - 1\right) \end{aligned} \tag{20}$$

For $k = 0$ we have $|w_0(x, u)| = |u - 1| \cdot |y(x)| \le C|u - 1||x|$ since $y(x) = O(x)$ as $x \to 0$. We will then use the trivial inequality

$$|e^x - 1| \le \frac{|x|}{1 - \frac{|x|}{2}} \tag{21}$$

for the induction steps. However, in order to apply this tool we need some a-priori estimates.

Obviously we have for $|x| \le \rho$ and $|u| \le 1$

$$|w_k(x, u)| \le 2y(|x|)$$

and consequently

$$\left| \sum_{i \ge 1} \frac{w_k(x^i, u^i)}{i} \right| \le 2 \sum_{i \ge 1} \frac{y(|x|^i)}{i} = 2 \log \frac{y(|x|)}{|x|}.$$

Since the function $\frac{y(x)}{x}$ is convex for $0 \le x \le \rho$ and $y(\rho) = 1$ we get $\frac{y(|x|)}{|x|} \le 1 + \frac{|x|}{\rho}$ because of the value of ρ. Consequently

$$\log \frac{y(|x|)}{|x|} \le \log\left(1 + \frac{|x|}{\rho}\right) \le \frac{|x|}{\rho}.$$

Thus, if $|x| \le \rho^2 + \varepsilon$ (for a sufficiently small $\varepsilon > 0$ we have

$$\left| \sum_{i\ge 1} \frac{w_k(x^i,u^i)}{i} \right| \le 2\rho + 2\frac{\varepsilon}{\rho}.$$

By using (21) we thus obtain

$$\left| \exp\left(\sum_{i\ge 1} \frac{w_k(x^i,u^i)}{i} \right) - 1 \right| \le \frac{1}{1-\rho-\frac{\varepsilon}{\rho}} \left| \sum_{i\ge 1} \frac{w_k(x^i,u^i)}{i} \right|.$$

Therefore, if we assume that we already know $|w_k(x,u)| \le C|u-1||x|L^k$ (for $|x| \le \rho^2+\varepsilon$, $|u|\le 1$ and some L with $0<L<1$) then we also get

$$\begin{aligned} |w_{k+1}(x,u)| &\le |y(x)| \cdot \left| \exp\left(\sum_{i\ge 1} \frac{w_k(x^i,u^i)}{i} \right) - 1 \right| \\ &\le \frac{|y(x)|}{1-\rho-\frac{\varepsilon}{\rho}} \left| \sum_{i\ge 1} \frac{w_k(x^i,u^i)}{i} \right| \\ &\le \frac{|y(x)|}{1-\rho-\frac{\varepsilon}{\rho}} CL^k \sum_{i\ge 1} \frac{|u^i-1|}{i} |x|^i \\ &\le \frac{|y(x)|}{1-\rho-\frac{\varepsilon}{\rho}} CL^k |u-1| \frac{|x|}{1-|x|}. \end{aligned}$$

By convexity we have $y(x) < x/\rho$ for $0<x<\rho$ and, thus, there exists $\varepsilon>0$ with $y(\rho^2+\varepsilon)\le\rho$. Consequently we get for $|x|\le\rho^2+\varepsilon$ the estimate

$$|w_{k+1}(x,u)| \le CL'L^k|x||u-1|$$

holds where

$$L' = \frac{\rho}{\left(1-\rho-\frac{\varepsilon}{\rho}\right)(1-\rho^2-\varepsilon)}.$$

The value of L' is smaller than 1 if $\varepsilon>0$ is sufficiently small. Thus, an induction proof works for $L=L'$. □

Corollary 1. *For $|u|\le 1$ and $|x|\le\rho+\varepsilon$ ($\varepsilon>0$ small enough) there is a positive constant $\tilde{C}$ such that (for all $k\ge 0$)*

$$|\Sigma_k(x,u)| \le \tilde{C}|u-1|L^k$$

with the constant L from the previous lemma.

Proof: We have

$$\begin{aligned} |\Sigma_k(x,u)| &\le \sum_{i\ge 2} \frac{1}{i}|w_k(x^i,u^i)| \le C \sum_{i\ge 2} \frac{1}{i}|u^i-1|\cdot|x|^i L^k \\ &\le C|u-1|L^k \frac{|x|^2}{1-|x|} \le C|u-1|L^k \frac{1}{1-(\rho+\varepsilon)} = \tilde{C}|u-1|L^k \end{aligned}$$

□

Corollary 2. *Let $u\in\Xi_k$ and $x\in\Delta_\varepsilon$. Then*

$$\sum_{i\ge 2} \gamma_k(x^i,u^i) = O\left(L^k\right).$$

Proof: The assertion follows immediately from the previous corollary and Taylor's theorem. □

We will eventually need precise upper and lower bounds of $w_k(x,u)$ near its singularity. By Taylor's theorem $w_k(x,u)$ can be expressed in terms of $\gamma_k(x,u)$. Hence, let us consider $\gamma_k(x,u)$ now.

Lemma 3. *For $x \in \Theta$ (where $\eta > 0$ is sufficiently small) the functions $\gamma_k(x)$ can be represented as*

$$\gamma_k(x) := \gamma_k(x,1) = C_k(x) y(x)^k, \tag{22}$$

where the $C_k(x)$ form a sequence of analytic functions which converges uniformly to an analytic limit function $C(x)$ (for $x \in \Theta$) with convergence rate

$$C_k(x) = C(x) + O(L^k), \tag{23}$$

for some L with $0 < L < 1$. Furthermore we have $C(\rho) = \frac{1}{2} b^2 \rho$.

There exist constants $\varepsilon, \theta, \tilde{\eta} > 0$ and $\theta < \frac{\pi}{2}$ such that

$$|\gamma_k(x,u)| = O(|y(x)|^{k+1}) \tag{24}$$

uniformly for $x \in \Delta_\varepsilon$ and $u \in \Xi_k$.

Remark. Note that the constants $\varepsilon, \theta, \tilde{\eta} > 0$ determine the domains Δ_ε and Ξ_k. Thus the theorem guarantees that there are suitable domains Δ_ε and Ξ_k of the shape (17) and (18) where the estimates are valid.

Proof: The first statement we have to show is that the functions $\gamma_k(x)$ are analytic functions in Θ. We prove this by induction. Obviously, the assertion holds for $k = 0$ since $\gamma_0(x) = y(x)$. Assume it is true for k. Then $\Gamma_k(x) := \sum_{i \ge 2} \gamma_k(x^i)$ is analytic for $|x| < \sqrt{\rho}$ and hence also in Θ. Using the recurrence relation of $y_k(x,u)$, Equation (10), we get

$$\begin{aligned}
\gamma_{k+1}(x,u) &= \frac{\partial}{\partial u} x e^{y_k(x,u) + \Sigma_k(x,u)} \\
&= x e^{y_k(x,u) + \Sigma_k(x,u)} \sum_{i \ge 1} \frac{\partial}{\partial u} y_k(x^i, u^i) u^{i-1} \\
&= y_{k+1}(x,u) \sum_{i \ge 1} \gamma_k(x^i, u^i) u^{i-1}. \qquad (25)
\end{aligned}$$

This implies

$$\gamma_{k+1}(x) = y(x)\gamma_k(x) + y(x)\Gamma_k(x)$$

which finally implies that $\gamma_{k+1}(x)$ is analytic in Θ as well.

By solving this recurrence we obtain also the analyticity of $C_k(x) = \frac{\gamma_k(x)}{y(x)^k}$ in Θ. Furthermore, we will show that the sequence $(C_k(x))_{k \ge 0}$ has a uniform limit $C(x)$ which has the desired properties.

Setting $u = 1$ we can rewrite (25) to

$$C_{k+1}(x) y(x)^{k+2} = C_k(x) y(x)^{k+2} + y(x) \left(C_k(x^2) y(x^2)^{k+1} + C_k(x^3) y(x^3)^{k+1} + \ldots \right). \tag{26}$$

resp. to

$$C_{k+1}(x) = \sum_{i \ge 1} C_k(x^i) \frac{y(x^i)^{k+1}}{y(x)^{k+1}}. \tag{27}$$

Set

$$L_k := \sup_{x \in \Theta} \sum_{i \ge 2} \frac{|y(x^i)|^{k+1}}{|y(x)|^{k+1}}.$$

If $\eta > 0$ is sufficiently small then due to Lemma 1 we have

$$\sup_{x \in \Theta} \frac{|y(x^i)|}{|y(x)|} < 1 \quad \text{for all } i \ge 2 \text{ and } \quad \sup_{x \in \Theta} \frac{|y(x^i)|}{|y(x)|} = O(\overline{L}^i)$$

for some $\overline{L}$ with $0<\overline{L}<1$. Consequently we also get

$$L_k = O(L^k)$$

for some L with $0 < L < 1$ (actually we can choose $L = \overline{L}^2$). Thus, if we use the notation $\|f\| := \sup_{x\in\Theta} |f(x)|$ then (27) yields

$$\|C_{k+1}\| \le \|C_k\|(1+L_k) \tag{28}$$

and also

$$\|C_{k+1} - C_k\| \le \|C_k\| L_k. \tag{29}$$

But (28) implies that the functions $C_k(x)$ are uniformly bounded in the given domain by

$$\|C_k\| \le c_0 := \prod_{\ell\ge 1}(1+L_\ell).$$

Furthermore, (29) guarantees the existence of a limit $\lim_{k\to\infty} C_k(x) = C(x)$ which is analytic in Θ; and we have uniform exponential convergence rate

$$\|C_k - C\| \le c_0 \sum_{\ell \ge k} L_\ell = O(L^k).$$

Hence, we get (22) as desired.

Finally, note that (for $|x| \le \rho$)

$$\sum_{k\ge 0} \gamma_k(x,1) = \sum_{n\ge 1} n y_n x^n = xy'(x).$$

On the other hand,

$$\sum_{k\ge 0} \gamma_k(x,1) = \sum_{k\ge 0} (C(x) + O\left(L^k\right)) y(x)^{k+1} n = \frac{C(x)y(x)}{1-y(x)} + O\left(1\right).$$

Since $C(x)$ is continuous in Θ this implies

$$C(\rho) = \lim_{x\to\rho} \frac{xy'(x)(1-y(x))}{y(x)} = \frac{b^2\rho}{2}$$

where we used (4).

Let us turn to the second assertion. In order to obtain the upper bound (24) we set for $\ell \le k$

$$\overline{C}_\ell = \sup_{x\in\Delta_\varepsilon,\ u\in\Xi_k} |\gamma_\ell(x,u) y(x)^{-\ell}|. \tag{30}$$

Observe that by Lemma 1 $|\gamma_\ell(x,u)| \le \overline{C}_\ell \sup_{x\in\Delta_\varepsilon} |y(x)|^k \le \overline{C}_\ell$. Therefore, by Taylor's theorem and Corollary 1 we obtain

$$\begin{aligned} |y_{\ell+1}(x,u)| &\le |y(x)| \exp\left(\sum_{i\ge 1} \frac{|w_\ell(x^i,u^i)|}{i} \right) \\ &\le |y(x)| e^{\overline{C}_\ell |u-1| + O(L^\ell)}. \end{aligned} \tag{31}$$

By (25) we have

$$\overline{C}_{\ell+1} = \sup_{x\in\Delta_\varepsilon,\ u\in\Xi_k} \left| \frac{y_{\ell+1}(x,u)}{y(x)^{\ell+1}} \right| \cdot \left| \sum_{i\ge 1} \gamma_\ell(x^i,u^i) u^{i-1} \right|.$$

Applying (31) and the estimate in Corollary 2 we get

$$\begin{aligned} \overline{C}_{\ell+1} &\le \sup_{x\in\Delta_\varepsilon,\ u\in\Xi_k} \left| \frac{e^{\overline{C}_\ell |u-1| + O\left(L^\ell\right)}}{y(x)^\ell} \right| \cdot \left| \gamma_\ell(x,u) + \sum_{i\ge 2} \gamma_\ell(x^i,u^i) u^{i-1} \right| \\ &\le e^{\overline{C}_\ell \tilde{\eta}/k} \overline{C}_\ell \left(1 + O\left(L^\ell\right)\right) \end{aligned} \tag{32}$$

Set

$$c_0 = \prod_{j\ge 0}(1+O(L^j))$$

and choose $\tilde\eta > 0$ such that $e^{2c_0\tilde\eta} \le 2$. We also choose $\varepsilon \le \tilde\eta$ and $0<\theta<\frac{\pi}{2}$ such that $|y(x)|\le 1$ for $|x-\rho|<\varepsilon$ and $|\arg(x-\rho)|\ge\theta$. If $k>0$ is fixed, then it follows by induction that, provided that $|u-1|\le \tilde\eta/k$,

$$\overline{C}_\ell \le \prod_{j<\ell}(1+O(L^j))\cdot e^{2c_0\tilde\eta\ell/k} \le 2c_0 \qquad (\ell\le k).$$

This completes the proof of the lemma, since $\bar C_0 = 1$. □

The representation (22) from Lemma 3 gives us a first indication of the behaviour of $w_k(x,u)$ for u close to 1. We expect that

$$w_k(x,u) \approx (u-1)\gamma_k(x,1) \sim (u-1)C(x)y(x)^k. \tag{33}$$

This actually holds (up to constants) in a proper range for u and x, although it is only partially true in the range of interest (see Theorem 6).

In order to make this more precise we derive estimates for the second derivatives $\gamma_k^{[2]}(x,u)$.

Lemma 4. *Suppose that $|x|\le\rho-\eta$ for some $\eta>0$ and $|u|\le 1$. Then uniformly*

$$\gamma_k^{[2]}(x,u) = O(y(|x|)^{k+1}). \tag{34}$$

There also exist constants $\varepsilon,\theta,\tilde\eta>0$ such that

$$\gamma_k^{[2]}(x,u) = O(k\,|y(x)|^{k+1}) \tag{35}$$

uniformly for $u\in\Xi_k$ and $x\in\Delta_\varepsilon$.

Proof: By definition we have the recurrence (compare with (25))

$$\begin{aligned}\gamma_{k+1}^{[2]}(x,u) &= y_{k+1}(x,u)\sum_{i\ge1} i\gamma_k^{[2]}(x^i,u^i)u^{2i-2}\\ &\quad + y_{k+1}(x,u)\left(\sum_{i\ge1}\gamma_k(x^i,u^i)u^{i-1}\right)^2\\ &\quad + y_{k+1}(x,u)\sum_{i\ge2}(i-1)\gamma_k(x^i,u^i)u^{i-2}\end{aligned} \tag{36}$$

with initial condition $\gamma_0^{[2]}(x)=0$.

First suppose that $|x|\le\rho-\eta$ for some $\eta>0$ and $|u|\le1$. Then we have $|\gamma_k^{[2]}(x,u)|\le\gamma_k^{[2]}(|x|,1)$. Thus, in this case it is sufficient to consider non-negative real $x\le\rho-\eta$. We proceed by induction. Suppose that we already know that $\gamma_k^{[2]}(x) := \gamma_k^{[2]}(x,1)\le D_k y(x)^{k+1}$ (where $D_0=0$). Then we get from (36) and the already known bound $\gamma_k(x,1)\le Cy(x)^k$ from Lemma 3 the upper bound

$$\begin{aligned}\gamma_{k+1}^{[2]}(x) &\le D_k y(x)^{k+2} + D_k y(x)^{k+2}\sum_{i\ge2} i\frac{y(x^i)^{k+1}}{y(x)^{k+1}}\\ &\quad + y(x)\left(C^2y(x)^{2k} + \sum_{i\ge2}\gamma_k(x^i,1)\right)^2\\ &\quad + Cy(x)^{k+2}\sum_{i\ge2}(i-1)\frac{y(x^i)^{k+1}}{y(x)^{k+1}}\\ &\le y(x)^{k+2}\left(D_k(1+O(L^k)) + C^2y(\rho-\eta)^k + O(L^k)\right)\end{aligned}$$

where we used Corollary 2 in the last step. Consequently we can set

$$D_{k+1} = D_k(1+O(L^k)) + C^2y(\rho-\eta)^k + O(L^k)$$

and obtain that $D_k = O(1)$ as $k \to \infty$ which proves (34).

Next choose $\varepsilon, \theta, \tilde\eta$ as in the proof of Lemma 3 and set (for $\ell \le k$)

$$\overline{D}_\ell = \sup_{x\in\Delta_\varepsilon,\ u\in\Xi_k} \left|\gamma_\ell^{[2]}(x,u) y(x)^{-\ell-1}\right|.$$

By the same reasoning as in the proof of Lemma 3, where we use the already proved bound $|\gamma_\ell(x,u) \le \overline{C}|y(x)|^{\ell+1}$, we obtain

$$\overline{D}_{\ell+1} \le \overline{D}_\ell\, e^{\tilde\eta C/k}(1+O(L^\ell)) + C^2\, e^{\tilde\eta C/k} + O(L^\ell),$$

that is, we have

$$\overline{D}_{\ell+1} \le \alpha_\ell \overline{D}_\ell + \beta_\ell$$

with $\alpha_\ell = e^{\tilde\eta C/k}(1+O(L^\ell))$ and $\beta_\ell = C^2\, e^{\tilde\eta C/k} + O(L^\ell)$. Hence we get

$$\begin{aligned}\overline{D}_k &\le \sum_{j=0}^{k-1} \beta_j \prod_{i=j+1}^{k-1} \alpha_i \\ &\le k \max_j \beta_j\, e^{\tilde\eta C} \prod_{\ell\ge 0}(1+O(L^\ell)) \\ &= O(k).\end{aligned}$$

This completes the proof of (35). □

Using the estimates for $\gamma_k(x,u)$ and $\gamma_k^{[2]}(x,u)$ we derive the following representations for $w_k(x,u)$ and $\Sigma_k(x,u)$.

Lemma 5. *Let $\varepsilon, \theta, \tilde\eta$ and $C_k(x) = \gamma_k(x,1)/y(x)^{k+1}$ as in Lemma 3. Then we have*

$$w_k(x,u) = C_k(x)(u-1)y(x)^{k+1}\left(1 + O(k|u-1|)\right), \tag{37}$$

uniformly for $u \in \Xi_k$ and $x \in \Delta_\varepsilon$.

Furthermore we have for $|x| \le \rho + \eta$ (for some $\eta > 0$) and $|u| \le 1$

$$\Sigma_k(x,u) = \tilde C_k(x)(u-1)y(x^2)^{k+1} + O\left(|u-1|^2 y(|x|^2)^k\right), \tag{38}$$

where the analytic functions $\tilde C_k(x)$ are given by

$$\tilde C_k(x) = \sum_{i\ge 2} C_k(x^i)\left(\frac{y(x^i)}{y(x^2)}\right)^{k+1}. \tag{39}$$

They have a uniform limit $\tilde C(x)$ with convergence rate

$$\tilde C_k(x) = \tilde C(x) + O(L^k)$$

for some constant L with $0 < L < 1$.

Proof: The first relation (37) follows from Lemma 3, Lemma 4 and Taylor's theorem.

In order to prove (38) we first note that $|x^i| \le \rho - \eta$ for $i \ge 2$ and $|x| \le \rho + \eta$ (if $\eta > 0$ is sufficiently small). Hence, by a second use of Taylor's theorem we get uniformly

$$w_k(x^i,u^i) = C_k(x^i)(u^i-1)y(x^i)^{k+1} + O\left(|u^i-1|^2 y(|x^i|)^{k+1}\right)$$

and consequently

$$\begin{aligned}\Sigma_k(x,u) &= \sum_{i\ge 2}\frac{1}{i}\, C_k(x^i)(u^i-1)y(x^i)^{k+1} + O\left(|u-1|^2 y(|x|^2)^k\right) \\ &= (u-1)\tilde C_k(x) y(x^2)^{k+1} + O\left(|u-1|^2 y(|x|^2)^k\right).\end{aligned}$$

Here we have used the property that the sum

$$\sum_{i\ge 2} C_k(x^i)\frac{u^i-1}{i(u-1)}\frac{y(x^i)^{k+1}}{y(x^2)^{k+1}}$$

represents an analytic function in x and u, since due to $|x| < \rho + \varepsilon < 1$ we have $x^i \to 0$ and therefore (3) implies $y(x^i) \sim x^i$. Finally, since $C_k(x) = C(x) + O(L^k)$, it also follows that $\tilde{C}_k(x)$ has a limit $\tilde{C}(x)$ and the same order of convergence. □

With these auxiliary results we are able to get a precise result for $w_k(x,u)$.

Lemma 6. *For $u \in \Xi_k$ and $x \in \Delta_\varepsilon$ ($\varepsilon, \theta, \tilde{\eta}$ as in Lemma 3) we have*

$$w_k(x,u) = \frac{C_k(x) w_0(x,u)\, y(x)^k}{1 - \frac{1}{2} C_k(x) w_0(x,u) \frac{1-y(x)^k}{1-y(x)} + O\left(|u-1|\right)}.$$

Proof: Since by Lemma 5 $\Sigma_k(x,u) = O(w_k(x,u))$, we observe that $w_k(x,u)$ satisfies the recurrence relation (we omit the arguments now)

$$\begin{aligned} w_{k+1} &= y\left(e^{w_k+\Sigma_k} - 1\right) \\ &= y\left(w_k + \frac{w_k^2}{2} + \Sigma_k + O(w_k^3) + O(\Sigma_k^2)\right) \\ &= y w_k \left(1 + \frac{w_k}{2} + O\left(w_k^2\right) + O\left(\Sigma_k\right)\right)\left(1 + \frac{\Sigma_k}{w_k}\right). \end{aligned}$$

Equivalently, we have

$$\begin{aligned} \frac{y}{w_{k+1}} + \frac{y\,\Sigma_k}{w_k w_{k+1}} &= \frac{1}{w_k}\left(1 - \frac{w_k}{2} + O\left(w_k^2\right) + O\left(\Sigma_k\right)\right) \\ &= \frac{1}{w_k} - \frac{1}{2} + O\left(w_k\right) + O\left(\frac{\Sigma_k}{w_k}\right), \end{aligned}$$

and consequently

$$\frac{y^{k+1}}{w_{k+1}} = \frac{y^k}{w_k} - \frac{\Sigma_k y^{k+1}}{w_k w_{k+1}} - \frac{1}{2} y^k + O\left(w_k y^k\right) + O\left(\frac{\Sigma_k y^k}{w_k}\right).$$

Thus we get by recurrence

$$\frac{y^k}{w_k} = \frac{1}{w_0} - \sum_{\ell=0}^{k-1} \frac{\Sigma_\ell}{w_\ell w_{\ell+1}} y^{\ell+1} - \frac{1}{2}\frac{1-y^k}{1-y} + O\left(\frac{1-L^k}{1-L}\right) + O\left(w_0 \frac{1-y^{2k}}{1-y^2}\right).$$

Now we use again Lemma 5 to obtain

$$\begin{aligned} w_0 \sum_{\ell=0}^{k-1} \frac{\Sigma_\ell}{w_\ell w_{\ell+1}} y^{\ell+1} &= \sum_{\ell=0}^{k-1} \frac{\tilde{C}_\ell(x) y(x^2)^{\ell+1} + O\left(|u-1| y(|x|^2)^\ell\right)}{C_\ell(x) C_{\ell+1}(x) y(x)^{\ell+1}(1 + O(\ell|u-1|))} \\ &= \sum_{\ell=0}^{k-1} \frac{\tilde{C}_\ell(x)}{C_\ell(x) C_{\ell+1}(x)} \frac{y(x^2)^{\ell+1}}{y(x)^{\ell+1}} + O(|u-1|) \\ &= c_k(x) + O(u-1) \end{aligned}$$

where $c_k(x)$ denotes the sum in the penultimate line above. Observe, too, that $w_0 \frac{1-y^{2k}}{1-y^2} = O(1)$, if $k|u-1| \le \tilde{\eta}$. Hence we obtain the representation

$$w_k = \frac{w_0 y^k}{1 - c_k(x) - \frac{w_0}{2}\frac{1-y^k}{1-y} + O\left(|u-1|\right)}. \tag{40}$$

Thus, it remains to verify that $1 - c_k(x) = 1/C_k(x)$. By using (27) and (39) it follows that

$$\tilde{C}_k(x) = \sum_{i\ge 2} C_k(x^i) \left(\frac{y(x^i)}{y(x^2)}\right)^{k+1} = \left(C_{k+1}(x) - C_k(x)\right)\left(\frac{y(x)}{y(x^2)}\right)^{k+1} \tag{41}$$

and consequently by telescoping

$$c_k(x) = \sum_{\ell=0}^{k-1} \frac{C_{\ell+1}(x) - C_\ell(x)}{C_\ell(x) C_{\ell+1}(x)} = \frac{1}{C_0(x)} - \frac{1}{C_k(x)} = 1 - \frac{1}{C_k(x)}.$$

Alternatively we can compare (40) with (37) for $u \to 1$ which also shows $1 - c_k(x) = 1/C_k(x)$. This completes the proof of the lemma. $\square$

The proof of Theorem 6 is now immediate. We substitute $x = \rho\left(1 + \frac{s}{n}\right)$ (where $s/n \to 0$), $u = e^{it/\sqrt{n}}$ and set $k = \kappa\sqrt{n}$. We also use the local expansion $y(x) = 1 - b\sqrt{\rho}\sqrt{1 - x/\rho} + O(|x - \rho|)$. That leads to

$$y(x)^k = \exp\left(-\kappa b\sqrt{-\rho s}\right)\left(1 + O\left(\frac{\kappa}{\sqrt{n}}\right)\right).$$

Finally, since the functions $C_k(x)$ are continuous and uniformly convergent to $C(x)$, they are also uniformly continuous and, thus,

$$C_k(x) \sim C(\rho) = \frac{1}{2}b^2\rho. \tag{42}$$

Altogether this leads to

$$\begin{aligned}
&\frac{w_0(x,u)\, y(x)^k}{\frac{1}{C_k(x)} - \frac{w_0(x,u)}{2}\frac{1-y(x)^k}{1-y(x)} + O\left(|u-1|\right)} \\
&\quad = \frac{C_k(x)(u-1)(1-y(x))\, y(x)^{k+1}}{1 - y(x) - C_k(x)\frac{w_0(x,u)}{2}(1-y(x)^k) + O\left(|u-1| \cdot |1-y(x)|\right)} \\
&\quad \sim \frac{b^2\rho}{2\sqrt{n}} \cdot \frac{it\sqrt{-s}\exp\left(-\frac{1}{2}\kappa b\sqrt{-\rho s}\,\right)}{\sqrt{-s}\exp\left(\frac{1}{2}\kappa b\sqrt{-\rho s}\,\right) - \frac{itb\sqrt{\rho}}{2}\sinh\left(\frac{1}{2}\kappa b\sqrt{-\rho s}\,\right)}
\end{aligned}$$

as proposed.

5. The Finite Dimensional Limiting Distributions – Proof of Theorem 4

For $d = 1$ (in Theorem 4) we have

$$\begin{aligned}
\phi_{k,n}(t) &= \frac{1}{y_n}[x^n]y_k\left(x, e^{it/\sqrt{n}}\right) \\
&= \frac{1}{2\pi i y_n}\int_\Gamma y_k\left(x, e^{it/\sqrt{n}}\right)\frac{dx}{x^{n+1}}
\end{aligned} \tag{43}$$

where the contour $\Gamma = \gamma \cup \Gamma'$ consists of a line

$$\gamma = \{x = \rho\left(1 - \frac{\sigma + i\tau}{n}\right) \mid -C\log^2 n \le \tau \le C\log^2 n\}$$

with an arbitrarily chosen fixed constants $C > 0$ and $\sigma > 0$, and Γ' is a circular arc centered at the origin and making Γ a closed curve.

The contribution of Γ' is exponentially small since for $x \in \Gamma'$ we have $\frac{1}{y_n}|x^{-n-1}| = O\left(n^{3/2}e^{-\log^2 n}\right)$ whereas $\left|y_k\left(x, e^{it/\sqrt{n}}\right)\right|$ is bounded.

If $x \in \gamma$, then the local expansion (14) is valid. Insertion into (43), using (5), and taking the limit for $n \to \infty$ yields the characteristic function of the distribution of $\frac{b\sqrt{\rho}}{2\sqrt{2}}l\left(\frac{b\sqrt{\rho}}{2\sqrt{2}}\kappa\right)$ as desired.

Now we proceed with $d = 2$. The computation of the two dimensional limiting distributions shows the general lines of the proof. An iterative use of the techniques will eventually prove Theorem 4. We confine ourselves with the presentation of the case $d = 2$.

We have to show

$$\left(\frac{1}{\sqrt{n}}L_n(\kappa_1\sqrt{n}),\, \frac{1}{\sqrt{n}}L_n(\kappa_2\sqrt{n})\right) \xrightarrow{w} \left(\frac{b\sqrt{\rho}}{2\sqrt{2}} \cdot l\left(\frac{b\sqrt{\rho}}{2\sqrt{2}}\kappa_1\right),\, \frac{b\sqrt{\rho}}{2\sqrt{2}} \cdot l\left(\frac{b\sqrt{\rho}}{2\sqrt{2}}\kappa_2\right)\right). \tag{44}$$

Since the characteristic function of the two dimensional distribution satisfies

$$
\begin{aligned}
\phi_{k,k+h,n}(t_1,t_2) &= \frac{1}{y_n}[x^n]\tilde{y}_{k,h}\left(x, e^{it_1/\sqrt{n}}, e^{it_2/\sqrt{n}}\right) \\
&= \frac{1}{2\pi i y_n}\int_\Gamma \tilde{y}_{k,h}\left(x, e^{it_1/\sqrt{n}}, e^{it_2/\sqrt{n}}\right)\frac{dx}{x^{n+1}} \\
&= 1 + \frac{1}{2\pi i y_n}\int_\Gamma w_{k,k+h}\left(x, e^{it_1/\sqrt{n}}, e^{it_2/\sqrt{n}}\right)\frac{dx}{x^{n+1}}, \qquad (45)
\end{aligned}
$$

where

$$
w_{k,k+h}(x,u_1,u_2) = y_{k,k+h}(x,u_1,u_2) - y(x) = \tilde{y}_{k,h}(x,u_1,u_2) - y(x),
$$

we need to analyze the asymptotic behaviour of $w_{k,k+h}(x,u_1,u_2)$ for k and h proportional to $\sqrt{n}$. Furthermore, note that $y(x)$ and $\tilde{y}_{k,h}(x,u_1,u_1)$ are analytic functions for $x \in \Delta$ and thus $w_{k,k+h}$ is bounded for $x \in \Delta$. Hence the contribution of Γ' to the Cauchy integral (45) is $O\left(1/y_n \cdot e^{-c\log^2 n}\right)$ with some suitable constant $c > 0$ and therefore negligibly small. Extending γ to infinity, as it is required for the two-dimensional version of (12), again introduces an exponentially small and therefore negligible error.

Thus it is sufficient to know the behaviour of $w_{k,k+h}\left(x, e^{it_1/\sqrt{n}}, e^{it_2/\sqrt{n}}\right)$ for $x \in \gamma$.

Then, equation (44) follows from the following proposition and (5).

Proposition 1. *Let $\kappa_2 > \kappa_1 > 0$ and t_1, t_2 be given with $|\kappa_2 t_1| \le c$ and $|\kappa_2 t_2| \le c$. Furthermore, set $u_i = e^{it_i/\sqrt{n}}$ and $k_i = \lfloor \kappa_i\sqrt{n}\rfloor$ for $i = 1, 2$. Define s by $x = \rho\left(1 + \frac{s}{n}\right)$. Then we have*

$$
w_{k_1,k_2}(x,u_1,u_2) \sim \frac{b\sqrt{2\rho}}{\sqrt{n}}\,\Psi_{\frac{\kappa_1 b\sqrt{\rho}}{2\sqrt{2}}}\left(s, \frac{it_1 b\sqrt{\rho}}{2\sqrt{2}} + \Psi_{\frac{(\kappa_2-\kappa_1)b\sqrt{\rho}}{2\sqrt{2}}}\left(s, \frac{it_2 b\sqrt{\rho}}{2\sqrt{2}}\right)\right), \qquad (46)
$$

as $n \to \infty$, uniformly for x, u_1 and u_2 such that $k_1|u_1 - 1| \le c$, $k_2|u_2 - 1| \le c$, and $s = O\left(\log^2 n\right)$ such that $\Re(s) = -\sigma$ where $\sigma > 0$ and sufficiently large but fixed.

Proof: Note that $w_{k_1,k_2}(x,1,1) = 0$, $w_{k_1,k_2}(x,u_1,1) = w_{k_1}(x,u_1)$, and $w_{k_1,k_2}(x,1,u_2) = w_{k_2}(x,u_2)$. Thus, the first derivatives are given by

$$
\begin{aligned}
\left[\frac{\partial}{\partial u_1} w_{k_1,k_2}(x,u_1,u_2)\right]_{u_1=u_2=1} &= \gamma_{k_1}(x,1), \\
\left[\frac{\partial}{\partial u_2} w_{k_1,k_2}(x,u_1,u_2)\right]_{u_1=u_2=1} &= \gamma_{k_2}(x,1).
\end{aligned}
$$

It is also possible to get bounds for the second derivatives of the form $O(k_2 y(x)^{k_1})$, if x, u_1, and u_2 are in the domain given in the assertion above. Hence, we can approximate $w_{k_1,k_2}(x,u_1,u_2)$ by

$$
\begin{aligned}
w_{k_1,k_2}(x,u_1,u_2) &= C_{k_1}(x)(u_1-1)y(x)^{k_1+1} + C_{k_2}(x)(u_2-1)y(x)^{k_2+1} \qquad (47)\\
&\quad + O\left(k_2 y(x)^{k_1}(|u_1-1|^2 + |u_2-1|^2)\right).
\end{aligned}
$$

Similarly (and even more easily, compare with the proof of Lemma 5) we obtain a representation for

$$
\begin{aligned}
\Sigma_{k_1,k_2}(x,u_1,u_2) &= \sum_{i\ge 2}\frac{w_{k_1,k_2}(x^i,u_1^i,u_2^i)}{i} \\
&= \tilde{C}_{k_1}(x)(u_1-1)y(x^2)^{k_1+1} + \tilde{C}_{k_2}(x)(u_2-1)y(x^2)^{k_2+1} \\
&\quad + O\left(y(|x|^2)^{k_1}|u_1-1|^2 + y(|x|^2)^{k_2}|u_2-1|^2\right). \qquad (48)
\end{aligned}
$$

In order to identify the asymptotic main term of $\Sigma_{k_1,k_2}(x,u_1,u_2)$ note that $k_2 - k_1 \sim (\kappa_2 - \kappa_1)\sqrt{n}$ and $|y(x^2)| < y(|x|^2) < 1$ for $|x| \le \rho + \eta < 1$. Moreover, observe that the terms $u_1 - 1$ and $u_2 - 1$ are proportional and hence

$$
\frac{C_{k_2}(x)(u_2-1)y(x)^{k_2}}{C_{k_1}(x)(u_1-1)y(x)^{k_1}} \sim \frac{t_2}{t_1}\exp\left(-(\kappa_2-\kappa_1)b\sqrt{-\rho s}\right). \qquad (49)
$$

The κ_i and the t_i are fixed, so we can choose σ such that the right-hand side in (49) is different from 1. This guarantees that the asymptotic main terms of (47), $C_{k_1}(x)(u_1-1)y(x)^{k_1}$ and $C_{k_2}(x)(u_2-1)y(x)^{k_2}$, do not cancel each other.

By using the same reasoning as in the proof of Lemma 6 we get the representation

$$w_{k_1,k_2} = \frac{w_{0,k_2-k_1}y^{k_1}}{1-f_{k_1}-\frac{w_{0,k_2-k_1}}{2}\frac{1-y^{k_1}}{1-y}+O(|u_1-1|+|u_2-1|)},$$

where

$$\begin{aligned} f_{k_1} &= f_{k_1}(x,u_1,u_2) \\ &= w_{0,k_2-k_1}(x,u_1,u_2)\sum_{\ell=0}^{k_1-1}\frac{\Sigma_{\ell,k_2-k_1+\ell}(x,u_1,u_2)\,y(x)^{\ell+1}}{w_{\ell,k_2-k_1+\ell}(x,u_1,u_2)w_{\ell+1,k_2-k_1+\ell+1}(x,u_1,u_2)}. \end{aligned}$$

Note that

$$\begin{aligned} w_{0,k_2-k_1} &= u_1t_{k_2-k_1}(x,u_2)-y(x) \\ &= (u_1-1)y(x)+u_1w_{k_2-k_1}(x,u_2) \\ &= U+W, \end{aligned}$$

where U and W abbreviate $U=(u_1-1)y(x)$ and $W=u_1w_{k_2-k_1}(x,u_2)$. The assumptions on u_i and k_i given in the statement of Proposition 1 imply that by Lemma 5 we have ($A\asymp B$ means that A and B have same order of magnitude)

$$W \asymp u_1C_{k_2-k_1}(x)(u_2-1)y(x)^{k_2-k_1+1}=o(u_2-1)$$

whereas $U\asymp u_1-1\asymp u_2-1$. Thus we may safely assume that $|W|<\frac{1}{2}|U|$, so that there is no cancellation.

Next by (47) we have

$$\begin{aligned} w_{\ell,k_2-k_1+\ell} =& C_\ell(x)(u_1-1)y(x)^{\ell+1}+C_{k_2-k_1+\ell}(x)(u_2-1)y(x)^{k_2-k_1\ell+1} \\ &+O\left((k_2-k_1\ell)y(x)^\ell(|u_1-1|^2+|u_2-1|^2)\right) \\ =& C_\ell(x)y(x)^\ell U+C_{k_2-k_1+\ell}(x)(u_2-1)y(x)^{k_2-k_1\ell+1} \\ &+O\left((k_2-k_1\ell)y(x)^\ell(|u_1-1|^2+|u_2-1|^2)\right) \end{aligned}$$

But since $C_{k_2-k_1+\ell}(x)\sim C_{k_2-k_1}(x)$ by formula (23) and Lemma 5 relates the second term to W, we obtain

$$w_{\ell,k_2-k_1+\ell}=\left(C_\ell(x)U+W\right)y(x)^\ell\left(1(k_2-k_1)|u_2-1|\right)\right).$$

Hence, f_k can be approximated by (for simplicity we omit the error terms)

$$\begin{aligned} f_{k_1} &\sim U(U+W)\sum_{\ell=0}^{k_1-1}\frac{\tilde{C}_\ell(x)\,(y(x^2)/y(x))^{\ell+1}}{(C_\ell(x)U+W)(C_{\ell+1}(x)U+W)} \\ &= (U+W)\sum_{\ell=0}^{k_1-1}\frac{(C_{\ell+1}(x)U+W)-(C_\ell(x)U+W)}{(C_\ell(x)U+W)(C_{\ell+1}(x)U+W)} \\ &= (U+W)\left(\frac{1}{U+W}-\frac{1}{C_kU+W}\right) \\ &= 1-\frac{U+W}{C_kU+W}, \end{aligned}$$

where we have used the formula (41) and telescoping. Consequently, it follows that

$$\begin{aligned} w_{k_1,k_2} &\sim \frac{(U+W)y^{k_1}}{\frac{U+W}{C_kU+W}-\frac{U+W}{2}\frac{1-y^{k_1}}{1-y}} \\ &= \frac{(C_kU+W)y^{k_1}}{1-\frac{C_kU+W}{2}\frac{1-y^{k_1}}{1-y}}. \end{aligned} \tag{50}$$

Now we will approximate all the terms in (50). First recall $C_k(x) \sim \frac{b^2\rho}{2}$ (formula (42)). The assertion on $x = \rho\left(1 + \frac{s}{n}\right)$ implies $x \to \rho$ and hence $y(x) \to 1$. Thus we obtain

$$U = y(u_1 - 1) \sim \frac{it_1}{\sqrt{n}}.$$

The asymptotic expansion (4) and Theorem 6 imply

$$\begin{aligned} y^{k_1} &\sim \exp\left(-\frac{1}{2}\kappa_1 b\sqrt{-\rho s}\right), \\ 1 - t &\sim b\sqrt{\rho}\sqrt{\frac{s}{n}}, \\ W = u_1 w_{k_2-k_1}(x, u_2) &\sim \frac{b\sqrt{2\rho}}{\sqrt{n}}\,\Psi_{\frac{(\kappa_2-\kappa_1)b\sqrt{\rho}}{2\sqrt{2}}}\left(s, \frac{it_2 b\sqrt{\rho}}{2\sqrt{2}}\right), \end{aligned}$$

Applying all these approximations we finally get (46). □

6. Tightness – Proof of Theorem 5

In this section we will show that the sequence of random variables $l_n(t) = n^{-1/2}L_n(t\sqrt{n})$, $t \ge 0$, is tight in $C[0,\infty)$. By [31, p. 63] it suffices to prove tightness for $C[0, A]$. Hence we consider $L_n(t)$ for $0 \le t \le A\sqrt{n}$, where $A > 0$ is an arbitrary real constant.

By [2, Theorem 12.3] tightness of $l_n(t) = n^{-1/2}L_n(t\sqrt{n})$, $0 \le t \le A$, follows from tightness of $L_n(0)$ (which is trivial) and from the existence of a constant $C > 0$ such that (9) holds for all non-negative integers n, r, h.

The fourth moment in Equation (9) can be expressed as the coefficient of a suitable generating function. Indeed the generating function counting tree according to size as well as the quantity $L_n(r) - L_n(r + h)$ is $\tilde{y}_{r,h}\left(x, u, \frac{1}{u}\right)$ ($\tilde{y}_{r,h}(x, u, v)$ is defined by (11)) since assigning a weight u to the vertices in level r and weight $1/u$ to those in level $r + h$ means that any tree having n vertices and with $L_n(r) - L_n(r + h) = m$ contributes to the coefficient of $x^n u^m$ where m can be negative. The fourth moment can then be obtained by applying the operator $\left(u\frac{\partial}{\partial u}\right)^4$ and setting $u = 1$ afterwards. Therefore we have

$$\mathbf{E}\left(L_n(r) - L_n(r + h)\right)^4 = \frac{1}{y_n}[x^n]\left[\left(\frac{\partial}{\partial u} + 7\frac{\partial^2}{\partial u^2} + 6\frac{\partial^3}{\partial u^3} + \frac{\partial^4}{\partial u^4}\right)\tilde{y}_{r,h}\left(x, u, \frac{1}{u}\right)\right]_{u=1}$$

Thus, (9) is equivalent to

$$[x^n]\left[\left(\frac{\partial}{\partial u} + 7\frac{\partial^2}{\partial u^2} + 6\frac{\partial^3}{\partial u^3} + \frac{\partial^4}{\partial u^4}\right)\tilde{y}_{r,h}\left(x, u, \frac{1}{u}\right)\right]_{u=1} \le C\,\frac{h^2}{\sqrt{n}}\rho^{-n} \tag{51}$$

In order to prove (51) we use a result from [18] saying that

$$F(x) = O\left((1 - x/\rho)^{-\beta}\right) \qquad (x \in \Delta)$$

implies

$$[x^n]F(x) = O\left(\rho^{-n}n^{\beta-1}\right),$$

where Δ is the region of (16)

Hence, it is sufficient to show that

$$\left[\left(\frac{\partial}{\partial u} + 7\frac{\partial^2}{\partial u^2} + 6\frac{\partial^3}{\partial u^3} + \frac{\partial^4}{\partial u^4}\right)\tilde{y}_{r,h}\left(x, u, \frac{1}{u}\right)\right]_{u=1} = O\left(\frac{h^2}{1 - |y(x)|}\right) = O\left(\frac{h^2}{\sqrt{|1 - x/\rho|}}\right) \tag{52}$$

for $x \in \Delta$ and $h \ge 1$. (Note that $\theta < \frac{\pi}{2}$ implies that $1 - |y(x)| \ge c\sqrt{|1 - x/\rho|}$ for some constant $c > 0$.)

Now we define

$$\gamma_k^{[j]}(x) = \left[\frac{\partial^j y_k(x, u)}{\partial u^j}\right]_{u=1} \qquad \text{and} \qquad \gamma_{r,h}^{[j]}(x) = \left[\frac{\partial^j \tilde{y}_{r,h}\left(x, u, \frac{1}{u}\right)}{\partial u^j}\right]_{u=1}$$

and derive the following upper bounds.

Lemma 7. *We have*

$$\gamma_k^{[1]}(x) = \begin{cases} O(1) & \text{uniformly for } x \in \Delta, \\ O\left(|x/\rho|^k\right) & \text{uniformly for } |x| \le \rho \end{cases} \tag{53}$$

and

$$\gamma_{r,h}^{[1]}(x) = \begin{cases} O\left(\frac{h}{r+h}\right) & \text{uniformly for } x \in \Delta, \\ O\left(|x/\rho|^r\right) & \text{uniformly for } |x| \le \rho, \end{cases} \tag{54}$$

where L is constant with $0 < L < 1$.

Proof: We already know that $\gamma_k^{[1]}(x) = C_k(x)y(x)^k$, where $C_k(x) = O(1)$ and $|y(x)| \le 1$ for $x \in \Delta$. Furthermore, by convexity we also have $|y(x)| \le |x/\rho|$ for $|x| \le \rho$. Hence, we obtain $\gamma_k^{[1]}(x) = O\left(|x/\rho|^k\right)$ for $|x| \le \rho$.

The functions $\gamma_{r,h}^{[1]}(x)$ are given by the recurrence

$$\gamma_{r+1,h}^{[1]}(x) = y(x) \sum_{i \ge 1} \gamma_{r,h}^{[1]}(x^i)$$

with initial value $\gamma_{0,h}^{[1]}(x) = y(x) - \gamma_h(x)$. To show this, just differentiate (11) w.r.t. u and then plug in $u = 1$. Hence, the representation $\gamma_{r,h}^{[1]}(x) = \gamma_r^{[1]}(x) - \gamma_{h+r}^{[1]}(x)$ follows by induction. Since, $\gamma_r^{[1]}(x) = (C(x) + O(L^r))y(x)^r$ we thus get that

$$\gamma_{r,h}^{[1]}(x) = O\left(\sup_{x\in\Delta} |y(x)^r(1 - y(x)^h)| + L^r\right)$$

However, it is an easy exercise to show that

$$\sup_{x\in\Delta} |y(x)^r(1 - y(x)^h)| = O\left(\frac{h}{r+h}\right). \tag{55}$$

For this purpose observe that if $x \in \Delta$ then we either have $|y(x) - 1| \le 1$ and $|y(x)| \le 1$, or $|y(x)| \le 1 - \eta$ for some $\eta > 0$. In the second case we surely have

$$|y(x)^r(1 - y(x)^h)| \le 2(1 - \eta)^r = O(L^r).$$

For the first case we set $y = 1 - \rho e^{i\varphi}$ and observe that

$$\left|1 - (1 - \rho e^{i\varphi})^h\right| \le (1 + \rho)^h - 1.$$

Hence, if $r \ge 3h$ we thus obtain that

$$|y(x)^r(1 - y(x)^h)| \le \max_{0\le\rho\le1} (1 - \rho)^r \left((1 + \rho)^h - 1\right) \le \frac{h}{r-h} \le \frac{2h}{r+h}.$$

If $r < 3h$ we obviously have

$$|y(x)^r(1 - y(x)^h)| \le 2 \le \frac{4h}{r+h}$$

which completes the proof of (55). Of course, we also have $L^r = O\left(\frac{h}{h+r}\right)$. This completes the proof of the upper bound of $\gamma_{r,h}^{[1]}(x)$ for $x \in \Delta$.

Finally, the upper bound $\gamma_{r,h}^{[1]}(x) = O(|x/\rho|^r)$ follows from (53). □

Lemma 8. *We have*

$$\gamma_k^{[2]}(x) = \begin{cases} O\left(\min\left\{k, \frac{1}{1-|y(x)|}\right\}\right) & \text{uniformly for } x \in \Delta, \\ O\left(|x/\rho|^k\right) & \text{uniformly for } |x| \le \rho - \eta \end{cases} \tag{56}$$

and

$$\gamma_{r,h}^{[2]}(x) = \begin{cases} O\left(\min\left\{h, \frac{1}{1-|y(x)|}\right\}\right) & \text{uniformly for } x \in \Delta, \\ O\left(|x/\rho|^r\right) & \text{uniformly for } |x| \le \rho - \eta \end{cases} \tag{57}$$

for every $\eta > 0$.

Remark. By doing a more precise analysis similarly to Lemma 3 we can, for example, show that $\gamma_k^{[2]}(x)$ can be represented as

$$\gamma_k^{[2]}(x) = y(x)^k \sum_{\ell=1}^{k} D_{k,\ell}(x) y(x)^{\ell-1}, \tag{58}$$

where the functions $D_{k,\ell}(x)$ are analytic in Δ. For every ℓ there is a limit $D_\ell(x) = \lim_{k\to\infty} D_{k,\ell}(x)$ with

$$D_{k,\ell}(x) = D_\ell(x) + O(\tilde{L}^{k+\ell}),$$

where $0 < \tilde{L} < 1$. Furthermore these limit functions $D_\ell(x)$ satisfy

$$D_\ell(x) = C(x)^2 + O(\tilde{L}^\ell).$$

Since we will not make use of this precise representation we leave the details to the reader.

Proof: The bound $\gamma_k^{[2]}(x) = O(|x/\rho|^k)$ (for $|x| \le \rho - \eta$) and the bound $\gamma_k^{[2]}(x) = O(k)$ follow from Lemma 4. In order to complete the analysis for $\gamma_k^{[2]}(x)$ we recall the recurrence derived from (10) (compare also with (36))

$$\gamma_{k+1}^{[2]}(x) = y(x) \sum_{i\ge 1} i\gamma_k^{[2]}(x^i) + y(x) \left(\sum_{i\ge 1} \gamma_k^{[1]}(x^i) \right)^2 + y(x) \sum_{i\ge 2} (i-1)\gamma_k^{[1]}(x^i) \tag{59}$$

that we rewrite to

$$\gamma_{k+1}^{[2]}(x) = y(x)\gamma_k^{[2]}(x) + b_k(x), \tag{60}$$

where

$$b_k(x) = y(x) \sum_{i\ge 2} i\gamma_k^{[2]}(x^i) + y(x) \left(\sum_{i\ge 1} \gamma_k^{[1]}(x^i) \right)^2 + y(x) \sum_{i\ge 2} (i-1)\gamma_k^{[1]}(x^i). \tag{61}$$

Note that for $i \ge 2$ we have $|x^i| \le \rho - \eta$. Therefore we can apply the second estimate of (56) (first and third sum of (61)) and the estimate (53) (the first for $i = 1$ in the second sum of (61), the second for the other summands) and obtain then

$$b_k(x) = O(1) \quad \text{uniformly for } x \in \Delta. \tag{62}$$

Since $\gamma_0^{[2]}(x) = 0$, the solution of the recurrence (60) can be written as

$$\gamma_k^{[2]}(x) = b_{k-1}(x) + y(x)b_{k-2}(x) + \cdots + y(x)^{k-1} b_0(x). \tag{63}$$

So (62) implies finally

$$\gamma_k^{[2]}(x) = O\left(\sum_{j=0}^{k-1} |y(x)|^j \right) = O\left(\frac{1}{1 - |y(x)|} \right).$$

which completes the proof of (56).

The recurrence for $\gamma_{r,h}^{[2]}(x)$ is similar to that of $\gamma_k^{[2]}(x)$:

$$\gamma_{r+1,h}^{[2]}(x) = y(x) \sum_{i\ge 1} i\gamma_{r,h}^{[2]}(x^i) + y(x) \left(\sum_{i\ge 1} \gamma_{r,h}^{[1]}(x^i) \right)^2 + y(x) \sum_{i\ge 2} (i-1)\gamma_{r,h}^{[1]}(x^i) \tag{64}$$

with initial value $\gamma_{0,h}^{[2]}(x) = \gamma_h^{[2]}(x)$. We again use induction. Assume that we already know that $|\gamma_{r,h}^{[2]}(x)| \le D_{r,h}|x/\rho|^k$ for $|x| \le \rho - \eta$ and for some constant $D_{r,h}$. By (57) we can set $D_{0,h} = D_h$

which is bounded as $h \to \infty$. We also assume that $|\gamma_{r,h}^{[1]}(x)| \le C|x/\rho|^k$ for $|x| \le \rho - \eta$. Then by (64) we get

$$|\gamma_{r+1,h}^{[2]}(x)| \le D_{r,h}|x/\rho|^{k+1} + D_{r,h}|x/\rho| \frac{2|x|^{2k}/\rho^k}{(1-|x|^k)^2} + C^2|x/\rho| \left(\frac{|x/\rho|^k}{1-|x|^k}\right)^2 + C|x/\rho| \frac{2|x|^{2k}/\rho^k}{(1-|x|^k)^2}.$$

Thus, we can set

$$D_{r+1,h} = D_{r,h}\left(1 + \frac{2(\rho-\eta)^k}{(1-\rho^k)^2}\right) + C^2 \frac{(\rho-\eta)^k}{(1-\rho^k)^2} + C\frac{2(\rho-\eta)^k}{(1-\rho^k)^2}$$

which shows that the constants $D_{r,h}$ are uniformly bounded. Consequently $\gamma_{r,h}^{[2]}(x) = O\left(|x/\rho|^k\right)$ for $|x| \le \rho - \eta$.

Next we start from (64) and assume that $|\gamma_{r,h}^{[2]}(x)| \le \bar{D}_{r,h}$ for $x \in \Delta$. We already know that $|\gamma_{r,h}^{[1]}(x)| \le C\frac{h}{h+k}$ for $x \in \Delta$. Hence,

$$\begin{aligned}|\gamma_{r+1,h}^{[2]}(x)| &\le \bar{D}_{r,h} + D_{r,h}\sum_{i\ge 2} i|x^i/\rho|^k \\ &\quad + C^2\left(\frac{h}{k+h} + \sum_{i\ge 2}|x^i/\rho|^k\right)^2 + C\sum_{i\ge 2}(i-1)|x^i/\rho|^k \\ &\le \bar{D}_{r,h} + 8D_{r,h}(\rho+\eta)^{2k}/\rho^k \\ &\quad + C^2\left(\frac{h}{k+h} + 2(\rho+\eta)^{2k}/\rho^k\right)^2 + 4C(\rho+\eta)^{2k}/\rho^k.\end{aligned}$$

Thus, we can set

$$\bar{D}_{r+1,h} = \bar{D}_{r,h} + 8D_{r,h}(\rho+\eta)^{2k}/\rho^k + C^2\left(\frac{h}{k+h} + 2(\rho+\eta)^{2k}/\rho^k\right)^2 + 4C(\rho+\eta)^{2k}/\rho^k$$

with initial value $\bar{D}_{0,h} = \bar{D}_h = O(h)$ and obtain a uniform upper bound of the form

$$\bar{D}_{r,h} = O(h).$$

Consequently $\gamma_{r,h}^{[2]}(x) = O(h)$ for $x \in \Delta$.

Thus, in order to complete the proof of (57) it remains to prove $\gamma_{r,h}^{[2]}(x) = O(1/(1-|y(x)|))$ for $x \in \Delta$. Analogously to the way we obtained (63) from (59) we obtain from (64) the representation of $\gamma_{r,h}^{[2]}(x)$ as

$$\gamma_{r,h}^{[2]}(x) = \gamma_{0,h}^{[2]}(x) + c_{k-1,h}(x) + y(x)c_{k-2,h}(x) + \cdots + y(x)^{k-1}c_{0,h}(x), \tag{65}$$

where

$$c_{j,h}(x) = y(x)\sum_{i\ge 2} i\gamma_{j,h}^{[2]}(x^i) + y(x)\left(\sum_{i\ge 1}\gamma_{j,h}^{[1]}(x^i)\right)^2 + y(x)\sum_{i\ge 2}(i-1)\gamma_{j,h}^{[1]}(x^i).$$

Observe that there exists $\eta > 0$ such that $|x^i| \le \rho - \eta$ for $i \ge 2$ and $x \in \Delta$. Hence it follows in a similar fashion as we showed (62) that $c_{j,h}(x) = O(1)$ for $x \in \Delta$. Since $\gamma_{0,h}^{[2]}(x) = \gamma_h^{[2]}(x) = O(1/(1-|y(x)|))$, we consequently get

$$\gamma_{r,h}^{[2]}(x) = \gamma_h^{[2]}(x) + O\left(\frac{1}{1-|y(x)|}\right) = O\left(\frac{1}{1-|y(x)|}\right).$$

□

Remark. Note that the theorem our proof relies on, namely [2, Theorem 12.3], actually requires the existence of $\alpha > 0$ and $\beta > 1$ such that

$$\mathbf{E}\,|L_n(r) - L_n(r+h)|^\alpha = O\left((h\sqrt{n})^\beta\right).$$

The estimates of Lemma 8 already prove that

$$\mathbf{E}\,(L_n(r) - L_n(r+h))^2 = O\left(h\sqrt{n}\right).$$

Unfortunately this estimate is slightly too weak to prove tightness. Third moments are technically unpleasant they attain positive and negative signs. So we actually have to deal with 4-th moments.

Before we start with bounds for $\gamma_k^{[3]}(x)$ and $\gamma_k^{[4]}(x)$ we need an auxiliary bound.

Lemma 9. *We have uniformly for* $x \in \Delta$

$$\sum_{r\ge 0} |\gamma_{r,h}^{[1]}(x)\gamma_{r,h}^{[2]}(x)| = O\left(h^2\right). \tag{66}$$

Proof: We use the representation (65), where we can approximate $c_{j,h}(x)$ by

$$c_{j,h}(x) = y(x)\gamma_{j,h}^{[1]}(x)^2 + O\left(L^j\right) = O\left(\frac{h^2}{(r+h)^2}\right)$$

uniformly for $x \in \Delta$ with some constant L that satisfies $0 < L < 1$. Furthermore, we use the approximation

$$\gamma_{r,h}^{[1]}(x) = C(x)y(x)^r(1 - y(x)^h) + O\left(L^r\right)$$

that is uniform for $x \in \Delta$. For example, this shows

$$\sum_{r\ge 0} |\gamma_{r,h}^{[1]}(x)| = |C(x)|\frac{|1 - y(x)^h|}{1 - |y(x)|} + O\left(1\right).$$

Now observe that for $x \in \Delta$ there exists a constant $c > 0$ with $|1 - y(x)| \ge c(1 - |y(x)|)$. Hence it follows that

$$\frac{|1 - y(x)^h|}{1 - |y(x)|} = O\left(\left|\frac{1 - y(x)^h}{1 - y(x)}\right|\right) = O\left(h\right)$$

and consequently

$$\sum_{r\ge 0} |\gamma_{r,h}^{[1]}(x)| = O\left(h\right).$$

Similarly we get

$$\begin{aligned}
\sum_{r\ge 1}|\gamma_{r,h}^{[1]}(x)| \left|\sum_{j<r} y(x)^{r-j-1}c_{j,h}(x)\right| &\le \sum_{j\ge 0} |c_{j,h}(x)|\,|y(x)|^{-j-1} \sum_{r>j} |y(x)|^r|\gamma_{r,h}^{[1]}(x)| \\
&= \sum_{j\ge 0} |c_{j,h}(x)|\,|y(x)|^{-j-1}\left(|C(x)|\,|y(x)|^{2j+2}\frac{|1 - y(x)^h|}{1 - |y(x)|^2} + O\left(|y(x)|^{j+1}L^j\right)\right) \\
&= O\left(\sum_{j\ge 0}\frac{h^3}{(j+h)^2}\right) \\
&= O\left(h^2\right).
\end{aligned}$$

Hence, we finally obtain

$$\begin{aligned}
\sum_{r\ge 0} |\gamma_{r,h}^{[1]}(x)\gamma_{r,h}^{[2]}(x)| &\le \sum_{r\ge 0} |\gamma_{r,h}^{[1]}(x)|\,|\gamma_{0,h}^{[2]}(x)| + \sum_{r\ge 1} |\gamma_{r,h}^{[1]}(x)| \left|\sum_{j<r} y(x)^{r-j-1}c_{j,h}(x)\right| \\
&= |\gamma_h^{[2]}(x)| \cdot \sum_{r\ge 0} |\gamma_{r,h}^{[1]}(x)| + \sum_{r\ge 1} |\gamma_{r,h}^{[1]}(x)| \left|\sum_{j<r} y(x)^{r-j-1}c_{j,h}(x)\right| = O\left(h^2\right).
\end{aligned}$$

where we used the fact $|\gamma_h^{[2]}(x)| = O\left(h\right)$ from Lemma 8. □

Lemma 10. *We have*

$$\gamma_k^{[3]}(x) = \begin{cases} O\left(\min\{k^2, \frac{k}{1-|y(x)|}\}\right) & \textit{uniformly for } x \in \Delta, \\ O\left(|x/\rho|^k\right) & \textit{uniformly for } |x| \le \rho - \eta \end{cases} \tag{67}$$

and

$$\gamma_{r,h}^{[3]}(x) = \begin{cases} O\left(\min\{h^2, \frac{h}{1-|y(x)|}\}\right) & \textit{uniformly for } x \in \Delta, \\ O\left(|x/\rho|^r\right) & \textit{uniformly for } |x| \le \rho - \eta \end{cases} \tag{68}$$

for every $\eta > 0$.

Proof: The recurrence for $\gamma_k^{[3]}(x)$ is again obtain by differentiation of (10) and given by

$$\begin{aligned} \gamma_{k+1}^{[3]}(x) &= y(x)\sum_{i\ge1} i^3\gamma_k^{[3]}(x^i) + y(x)\left(\sum_{i\ge i}\gamma_k^{[1]}(x^i)\right)^3 + 3y(x)\left(\sum_{i\ge1}\gamma_k^{[1]}(x^i)\right)\left(\sum_{i\ge1} i\gamma_k^{[2]}(x^i)\right) \\ &\quad + 3y(x)\left(\sum_{i\ge1}\gamma_k^{[1]}(x^i)\right)\left(\sum_{i\ge1}(i-1)\gamma_k^{[i]}(x^i)\right) + 3y(x)\sum_{i\ge1} i(i-1)\gamma_k^{[2]}(x^i) \\ &\quad + y(x)\sum_{i\ge1}(i-1)(i-2)\gamma_k^{[1]}(x^i) \end{aligned} \tag{69}$$

By inspecting the proof of Lemmas 7 and 8 one expects that the only *important* part of this recurrence if given by

$$\gamma_{k+1}^{[3]}(x) = y(x)\gamma_k^{[3]}(x) + y(x)\gamma_k^{[1]}(x)^3 + 3y(x)\gamma_k^{[1]}(x)\gamma_k^{[2]}(x) + R_k \tag{70}$$

and R_k collects the *less important* remainder terms that only contributes exponentially small terms. Thus, in order to shorten our presentation we will only focus on these terms. In particular it is easy to show the bound $\gamma_k^{[3]}(x) = O\left(|x/\rho|^k\right)$ for $|x| \le \rho - \eta$. (We omit the details.)

Next, since $y(x)\gamma_k^{[1]}(x)^3 + 3y(x)\gamma_k^{[1]}(x)\gamma_k^{[2]}(x) + R_k = O\left(k\right)$ for $x \in \Delta$, it directly follows that $\gamma_k^{[3]}(x) = O\left(k^2\right)$.

Now we proceed by induction and observe that a bound of the form $|\gamma_k^{[3]}(x)| \le E_k/(1-|y(x)|)$ leads to

$$|\gamma_{k+1}^{[3]}(x)| \le \frac{E_k}{1-|y(x)|} + O\left(\frac{1}{1-|y(x)|}\right) + |R_k|$$

and consequently to $E_{k+1} \le E_k + O\left(1\right)$. Hence, $E_k = O\left(k\right)$ and $\gamma_k^{[3]}(x) = O\left(k/(1-|y(x)|)\right)$.

Similarly, the leading part of the recurrence for $\gamma_{r,h}^{[3]}(x)$ is given by

$$\begin{aligned} \gamma_{r+1,h}^{[3]}(x) &= y(x)\gamma_{r,h}^{[3]}(x) + y(x)\gamma_{r,h}^{[1]}(x)^3 + 3y(x)\gamma_{r,h}^{[1]}(x)\gamma_{r,h}^{[2]}(x) + \bar{R}_{r,h} \\ &= y(x)\gamma_{r,h}^{[3]}(x) + d_{r,h}(x), \end{aligned} \tag{71}$$

where

$$d_{r,h}(x) = y(x)\gamma_{r,h}^{[1]}(x)^3 + 3y(x)\gamma_{r,h}^{[1]}(x)\gamma_{r,h}^{[2]}(x) + \bar{R}_{r,h} = O\left(h\right)$$

and the initial value is given by

$$\gamma_{0,h}^{[3]}(x) = -\gamma_h^{[3]}(x) - 3\gamma_h^{[2]}(x) = O\left(\min\left\{h^2, \frac{h}{1-|y(x)|}\right\}\right).$$

Note that we also assume that $\gamma_{r,h}^{[3]}(x) = O\left(|x/\rho|^r\right)$ for $|x| \le \rho - \eta$ (which can be easily proved). Consequently it directly follows that

$$\begin{aligned}\gamma_{r,h}^{[3]}(x) &= \gamma_{0,h}^{[3]}(x) + d_{r-1,h}(x) + y(x)d_{r-1,h}(x) + \cdots + y(x)^{r-1}d_{0,h}(x)\\ &= O\left(\frac{h}{1-|y(x)|}\right).\end{aligned}$$

Next observe that Lemmas 7–9 ensure that

$$\sum_{j\ge 0} |d_{j,h}(x)| = O\left(h^2\right)$$

uniformly for $x \in \Delta$. Hence, we finally get

$$\gamma_{r,h}^{[3]}(x) = O\left(h^2\right)$$

which completes the proof of Lemma 10. □

Lemma 11. *We have*

$$\gamma_k^{[4]}(x) = \begin{cases} O\left(\frac{k^2}{1-|y(x)|}\right) & \text{uniformly for } x \in \Delta,\\ O\left(|x/\rho|^k\right) & \text{uniformly for } |x| \le \rho-\eta \end{cases} \tag{72}$$

and

$$\gamma_{r,h}^{[4]}(x) = \begin{cases} O\left(\frac{h^2}{1-|y(x)|}\right) & \text{uniformly for } x \in \Delta,\\ O\left(|x/\rho|^r\right) & \text{uniformly for } |x| \le \rho-\eta \end{cases} \tag{73}$$

for every $\eta > 0$.

Proof: The proof is very similar to that of Lemma 10. First, the recurrence for $\gamma_k^{[4]}(x)$ is essentially of the form

$$\begin{aligned}\gamma_{k+1}^{[4]}(x) &= y(x)\gamma_k^{[4]}(x) + y(x)\gamma_k^{[1]}(x)^4 + 4y(x)\gamma_k^{[1]}(x)\gamma_k^{[3]}(x)\\ &\quad + 6y(x)\gamma_k^{[1]}(x)^2\gamma_k^{[2]}(x) + 3y(x)\gamma_k^{[2]}(x)^2 + R_k,\end{aligned} \tag{74}$$

where R_k collects all exponentially small summands. We assume that we have already proved the upper bound $\gamma_k^{[4]}(x) = O\left(|x/\rho|^k\right)$ for $|x| \le \rho-\eta$. Now, by induction and the assumption $|\gamma_k^{[4]}(x)| \le F_k/(1-|y(x)|)$ and the known estimates $\gamma_k^{[1]}(x) = O(1)$, $\gamma_k^{[2]}(x) = O\left(\min\{k, 1/(1-|y(x)|)\}\right)$, and $\gamma_k^{[3]}(x) = O\left(k/(1-|y(x)|)\right)$ we get

$$|\gamma_{k+1}^{[4]}(x)| \le \frac{F_k}{1-|y(x)|} + O\left(|y(x)|^k\right) + O\left(\frac{k}{1-|y(x)|}\right) + O\left(\frac{1}{1-|y(x)|}\right) + |R_k|$$

and consequently $F_k = O\left(k^2\right)$.

Finally, the essential part of the recurrence for $\gamma_{r,h}^{[4]}(x)$ is given by

$$\begin{aligned}\gamma_{r+1,h}^{[4]}(x) &= y(x)\gamma_{r,h}^{[4]}(x) + y(x)\gamma_{r,h}^{[1]}(x)^4 + 4y(x)\gamma_{r,h}^{[1]}(x)\gamma_{r,h}^{[3]}(x)\\ &\quad + 6y(x)\gamma_{r,h}^{[1]}(x)^2\gamma_{r,h}^{[2]}(x) + 3y(x)\gamma_{r,h}^{[2]}(x)^2 + \bar{R}_{r,h}\\ &= y(x)\gamma_{r,h}^{[4]}(x) + e_{r,h}(x),\end{aligned} \tag{75}$$

where

$$\begin{aligned}e_{r,h}(x) &= y(x)\gamma_{r,h}^{[1]}(x)^4 + 4y(x)\gamma_{r,h}^{[1]}(x)\gamma_{r,h}^{[3]}(x)\\ &\quad + 6y(x)\gamma_{r,h}^{[1]}(x)^2\gamma_{r,h}^{[2]}(x) + 3y(x)\gamma_{r,h}^{[2]}(x)^2 + \bar{R}_{r,h}.\end{aligned}$$

As above, $\bar{R}_{r,h}$ collects all exponentially small terms. Thus,

$$\gamma_{r,h}^{[4]}(x) = \gamma_{0,h}^{[4]}(x) + e_{r-1,h}(x) + y(x)e_{r-1,h}(x) + \cdots + y(x)^{r-1}e_{0,h}(x).$$

If we use the known estimates $\gamma_{r,h}^{[1]}(x) = O(1)$, $\gamma_{r,h}^{[2]}(x) = O(h)$, and $\gamma_{r,h}^{[3]}(x) = O\left(h^2\right)$ which gives $d_{r,h} = O\left(h^2\right)$ and the initial condition

$$\gamma_{0,h}^{[4]}(x) = 12\gamma_h^{[2]}(x) + 8\gamma_h^{[3]}(x) + \gamma_h^{[4]}(x) = O\left(\frac{h^2}{1-|y(x)|}\right)$$

we obtain

$$\gamma_{r,h}^{[4]}(x) = O\left(\frac{h^2}{1-|y(x)|}\right).$$

This completes the proof of Lemma 11. □

The proof of (52) is now immediate. As already noted this implies (51) and proves Theorem 5.

7. The Height – Proof of Theorem 2

Let $y_n^{(k)}$ denote the number of trees with n nodes and height less than k. Then the generating function $y_k(x) = \sum_{n\geq 1} y_n^{(k)} x^n$ satisfies the recurrence relation

$$y_0(x) = 0$$
$$y_{k+1}(x) = x\exp\left(\sum_{i\geq 1}\frac{y_k(x^i)}{i}\right), \quad k \geq 0.$$

Obviously $y_k(x) = y_k(x,0)$ where the function on the right-hand side is the generating function of (10) which we used to analyze the profile in the previous sections. So w_k and Σ_k could be defined accordingly. However, the proof given here relies heavily on the seminal work of Flajolet and Odlyzko [20] on the height of binary trees. Therefore, to be in accordance with the notation used there, we work with the opposite sign and set

$$e_k(x) = y(x) - y_k(x),$$

that is $e_k(x) = -w_k(x,0)$. Then e_k satisfies the recurrence

$$e_{k+1}(x) = y(x)\left(1 - e^{-e_k(x)-E_k(x)}\right), \tag{76}$$

where

$$E_k(x) = \sum_{i\geq 2}\frac{e_k(x^i)}{i} = -\Sigma_k(x,0). \tag{77}$$

The function $e_k(x)$ is the generating function for the number of trees with height at least k.

The proof of Theorem 2 follows the same principles as the proof of the corresponding properties of the height of Galton-Watson trees (see [20, 19]). However, the term $E_k(x)$ needs some additional considerations.

Proposition 2. *If $e_k(x)$ satisfies*

$$e_k(x) = \frac{y(x)^k}{\frac{1}{2}\frac{1-y(x)^k}{1-y(x)} + O\left(\min\left\{\log k, \log\frac{1}{1-|y(x)|}\right\}\right)}. \tag{78}$$

for $x \in \Delta_\varepsilon$, then Theorem 2 follows.

Proof: Since $e_k(x)$ is bounded in $\Delta \setminus \Delta_\varepsilon$, only the local behaviour near $x = \rho$ determines the asymptotic height of Pólya trees. The shape (78) of $e_k(x)$ precisely matches that of the corresponding quantity for simply generated trees. Flajolet and Odlyzko showed that (78) implies (6), see [20, p. 204] where this argument was used to derive the average height as well as the other moments of the height of simply generated trees. □

Proposition 3. *Suppose that for $x \in \Delta_\varepsilon$ the estimate*

$$\frac{|E_k(x)|}{|e_k(x)|^2} = O(L^k) \tag{79}$$

holds for some $L < 1$ and that

$$|e_k(x)y(x)^k| = O\left(1/k\right). \tag{80}$$

Then we have (78) for $x \in \Delta_\varepsilon$.

Proof: Equation (76) can be rewritten to (omitting the argument x)

$$e_{k+1} = ye_k \left(1 - \frac{e_k}{2} + O\left(e_k^2 + \frac{E_k}{e_k}\right)\right),$$

resp. to

$$\frac{y}{e_{k+1}} = \frac{1}{e_k} + \frac{1}{2} + O\left(e_k + \frac{E_k}{e_k^2}\right).$$

This leads to the representation

$$\frac{y^k}{e_k} = \frac{1}{e_0} + \frac{1}{2}\frac{1-y^k}{1-y} + O\left(\sum_{\ell<k} |e_\ell y^\ell|\right) + O\left(\sum_{\ell<k} \frac{|E_\ell|}{|e_\ell^2|}|y^\ell|\right). \tag{81}$$

Recall that $e_0 = y(x)$. By (79) and $\frac{1}{|y(x)|} = O\left(1\right)$ (for $x \in \Delta_\varepsilon$), a consequence of Lemma 1, this implies

$$e_k = \frac{y^k}{\frac{1}{2}\frac{1-y^k}{1-y} + O\left(\sum_{\ell<k} |e_\ell y^\ell|\right) + O(1)} \tag{82}$$

and (80) yields (78) after all. □

Remark. Note that the proof above does not make explicit use of the domain of x. Thus the implication of Proposition 3 is still true if we write, for instance, $0 \le x \le \rho$ instead of $x \in \Delta_\varepsilon$ in (78), (79), and (80). We remark that we will use such modifications of Proposition 3 in the sequel, though we do not state several almost identical propositions differing only in the domain of x.

Note that (81) and (78) can be made more precise. Set

$$S_k = \frac{e_k^2\left(e^{-E_k} - 1\right)}{\left(e^{e_k} - 1\right)\left(1 - e^{-e_k - E_k}\right)},$$

and define a function $h(v)$ by

$$\frac{v}{1 - e^{-v}} = 1 + \frac{v}{2} + v^2 h(v).$$

Then the recurrence $e_{k+1} = y(1 - e^{-e_k - E_k})$ rewrites to

$$\frac{y}{e_{k+1}} = \frac{1}{e_k} + \frac{1}{2} + e_k h(e_k) + \frac{S_k}{e_k^2}$$

and leads to the explicit representations

$$\frac{y^k}{e_k} = \frac{1}{e_0} + \frac{1}{2}\frac{1-y^k}{1-y} + \sum_{\ell<k} e_\ell h(e_\ell) y^\ell + \sum_{\ell<k} \frac{S_\ell}{e_\ell^2} y^\ell \tag{83}$$

and

$$e_k = \frac{y^k}{\frac{1}{e_0} + \frac{1}{2}\frac{1-y^k}{1-y} + \sum_{\ell<k} e_\ell h(e_\ell) y^\ell + \sum_{\ell<k} \frac{S_\ell}{e_\ell^2} y^\ell}. \tag{84}$$

This formula is a refinement of (78) since it makes the error term explicit. We will use it in the sequel.

Furthermore, note that if we just assume $e_k \to 0$ and $E_k = o(e_k)$ as $k \to \infty$, then

$$S_k \sim -E_k \text{ and } h(e_k) = O\left(1\right). \tag{85}$$

We start our precise analysis with an a priori bound for $e_k(x)$. The next step is proving (78) for $0 \le x \le \rho$. Then we will, little by little, enlarge the allowed domain for x and arrive finally at Proposition 3 as stated above.

Lemma 12. *Let $|x| \le \rho$. Then there is a $C > 0$ such that*

$$|e_k(x)| \le \frac{C}{\sqrt{k}} \left|\frac{x}{\rho}\right|^k.$$

Proof: Obviously, we have

$$|e_k(x)| = \sum_{n>k} (y_n - y_n^{(k)})|x|^n \le \sum_{n>k} y_n |x|^n.$$

The assertion follows now from $y_n \sim c\rho^{-n} n^{-3/2}$ for some constant $c > 0$. □

Lemma 12 applies to $E_k(x)$.

Corollary 3. *Suppose that $|x| < \sqrt{\rho}$. Then there exists a constant $C_0 > 0$ with*

$$|E_k(x)| \le \frac{C_0}{\sqrt{k}} \left|\frac{x^2}{\rho}\right|^k.$$

Remark. Observe that in the definition of $E_k(x)$, Equation (77), the arguments in the sum are raised to a power of at least 2. Therefore $E_k(x)$ is an analytic function in the domain $|x| < \sqrt{\rho}$ and not only in the smaller domain $|x| \le \rho$.

The next lemma shows that $e_k(x)$ behaves as expected if x is on the positive real axis.

Lemma 13. *Suppose that $0 \le x \le \rho$ is real. Then (78) holds in this domain for x.*

Remark. As remarked before, we will show a weaker version of (78). During the proof we will make use of the analogous weaker versions of (79) and (80). Therefore, in the following proof a reference to one of these formulas means the referred formula, but for $0 \le x \le \rho$ and not for $x \in \Delta_\varepsilon$.

Proof: Let $\tilde{e}_k(x)$ be defined by $\tilde{e}_0(x) = y(x)$ and by $\tilde{e}_{k+1}(x) = y(x)(1 - e^{-\tilde{e}_k(x)})$ (for $k \ge 0$). Then $\tilde{e}_k(x)$ is precisely the analogue of e_k for Cayley trees, a class of simply generated trees (preceisely: the class of labelled rooted trees). So $\tilde{e}_k(x)$ behaves like (78) in Δ.

However, if $0 \le x \le \rho$ then we obtain by induction that $e_k(x) \ge \tilde{e}_k(x)$. Hence, by combining (78) with the upper bound from Lemma 12 we have

$$\frac{E_k(x)}{e_k(x)^2} \le \frac{E_k(x)}{\tilde{e}_k(x)^2} = O(L^k)$$

for some L with $0 < L < 1$. Thus (79) is satisfied.

In order to show the second assumption (80) of Proposition 3 note that by Lemma 12 $e_k(x)$ is even exponentially small for $x < \rho$. For the case $x = \rho$ observe that (79) in conjunction with Lemma 12 guarantee (85). Applying this to (84) implies

$$e_k(x) = \frac{y(x)^k}{\frac{1}{2}\frac{1-y(x)^k}{1-y(x)} + O\left(\sqrt{k}\right)}.$$

This equation yields

$$e_k(\rho) \sim \frac{2}{k} \tag{86}$$

and completes the proof. □

The analysis of $e_k(x)$ for complex x with $|x| \le \rho$ is not too difficult. The next two lemmas consider the case $|x| \le \rho$ and $|x - \rho| \le \varepsilon$ and the case $|x| \le \rho - \varepsilon$.

Lemma 14. *There exists $\varepsilon > 0$ such that (78) holds for all x with $|x| \le \rho$ and $|x - \rho| \le \varepsilon$.*

Proof: First recall that $|e_k(x)| \le C/\sqrt{k}$ and $|y(x)| < 1$ for $x \ne \rho$. Moreover, in the proof of the previous lemma we showed $e_k(\rho) \sim 2/k$. Hence (80) is satisfied.

Suppose that we can show that $|E_k/e_k^2| \le 1$ or, equivalently, $|E_k/e_k| \le |e_k| \le C/\sqrt{k}$. Then it follows that

$$\begin{aligned}
|e_{k+1}| &= |y|\,|e_k| \left|\frac{1-e^{-e_k-E_k}}{e_k}\right| \\
&= |y|\,|e_k| \left(1+\frac{E_k}{e_k}+O\left(\frac{(e_k+E_k)^2}{e_k}\right)\right) \\
&\ge |y|\,|e_k|\,1-C_1|e_k| \\
&\ge |y|\,|e_k| e^{-C_2k^{-1/2}}. \qquad (87)
\end{aligned}$$

where C_1, C_2 are suitable constants.

Now we choose k_0 sufficiently large such that

$$e^{-2C_2\sqrt{k}} \le \frac{1}{k} \quad \text{and} \quad C_0\rho^{k/2}e^{4C_2\sqrt{k}} \le 1$$

hold for all $k \ge k_0$. By continuity, (86) implies the existence of an $\varepsilon > 0$ with $|e_{k_0}(x)| \ge \frac{1}{k_0}$ and $|y(x)| \ge \rho^{1/4}$ for $|x| \le \rho$ and $|x-\rho| \le \varepsilon$. These assumptions imply

$$|e_{k_0}(x)| \ge \frac{1}{k_0} \ge e^{-2C_2\sqrt{k_0}} \ge |y(x)|^{k_0}e^{-2C_2\sqrt{k_0}}$$

and by Corollary 3 (since $|x| \le \rho < \sqrt{\rho}$ this is applicable)

$$\frac{|E_{k_0}|}{|e_{k_0}^2|} \le C_0\rho^{k_0}|y|^{-2k_0}e^{4C_2\sqrt{k_0}} \le C_0\rho^{k_0/2}e^{4C_2\sqrt{k_0}} \le 1.$$

The goal is to show by induction that for $k \ge k_0$ and for $|x| \le \rho$ and $|x-\rho| \le \varepsilon$

$$|e_k| \ge |y|^k e^{-3C_2\sqrt{k}} \quad \text{and} \quad \left|\frac{E_k}{e_k^2}\right| \le 1. \qquad (88)$$

Assume that (88) is satisfied for $k = k_0$. Now suppose that (88) holds for some $k \ge k_0$. Then (87) implies

$$\begin{aligned}
|e_{k+1}| &\ge |y|\,|e_k|e^{-C_2k^{-1/2}} \\
&\ge |y|^{k+1}e^{-3C_2\sqrt{k}}e^{-C_2k^{-1/2}} \\
&\ge |y|^{k+1}e^{-3C_2\sqrt{k+1}}.
\end{aligned}$$

Furthermore

$$\frac{|E_{k+1}|}{|e_{k+1}^2|} \le C_0\rho^{k+1}|y|^{-2k-2}e^{4C_2\sqrt{k+1}} \le C_0\rho^{(k+1)/2}e^{4C_2\sqrt{k+1}} \le 1.$$

Hence, we have proved (88) for all $k \ge k_0$.

In the last step of the induction proof we also obtained the upper bound

$$\frac{|E_k|}{|e_k^2|} \le C_0\rho^{k/2}e^{4C_2\sqrt{k}}$$

which is sufficient to obtain the asymptotic representation (78). □

Lemma 15. *Suppose that $|x| \le \rho - \varepsilon$ for some $\varepsilon > 0$. Then we have uniformly*

$$e_k(x) = C_k(x)y(x)^k = (C(x)+o(1))y(x)^k \qquad (89)$$

for some analytic function $C(x)$. Consequently we have uniformly for $|x| \le \sqrt{\rho-\varepsilon}$

$$E_k(x) = \tilde{C}_k(x)y(x^2)^k = (\tilde{C}(x)+o(1))y(x^2)^k \qquad (90)$$

with an analytic function $\tilde{C}(x)$.

Proof: If $|x| \le \rho - \varepsilon$ then by Lemma 12 we have $|e_k(x)| \le e_k(\rho - \varepsilon) = O\left(\left(1 - \frac{\varepsilon}{\rho}\right)^k\right)$. Thus, we can replace the upper bound $|e_k(x)| \le C/\sqrt{k}$ in the proof of Lemma 14 by an exponential bound which leads to a lower bound for $e_k(x)$ of the form

$$|e_k(x)| \ge c_0 |y(x)|^k.$$

Hence, by using (84) the result follows with straightforward calculations.

In order to show the second assertion we start from (77) and insert (89). Then we obtain

$$\begin{aligned} E_k(x) &= \frac{C(x^2) + o(1)}{2} y(x^2)^k + \sum_{i \ge 3} \frac{(C(x^i) + o(1)) y(x^i)^k}{i} \\ &= \left(\frac{C(x^2)}{2} + \sum_{i \ge 3} \frac{(C(x^i) + o(1)) y(x^i)^k}{i y(x^2)} + o(1) \right) y(x^2). \end{aligned} \tag{91}$$

Since $C(x)$ is analytic, it is bounded in the compact interval $[0, \rho - \varepsilon]$. Furthermore, observe that using the bound from Lemma 1 it is easy to see that the series in (91) is uniformly convergent. Hence the representation (90) follows. □

The disadvantage of the previous two lemmas is that they only work for $|x| \le \rho$. In order to obtain some progress for $|x| > \rho$ fix a constant $C > 0$ such that

$$\left| e^{-E_k(x)} - 1 \right| \le \frac{C}{\sqrt{k}} \left(\frac{|x|^2}{\rho} \right)^k$$

for all $k \ge 1$ and for all $|x| \le \sqrt{\rho}$.

Lemma 16. *Let $x \in \Delta$ and suppose that there exist real numbers D_1 and D_2 with $0 < D_1, D_2 < 1$ and some integer $K \ge 1$ with*

$$|e_K(x)| < D_1, \quad |y(x)| \frac{e^{D_1} - 1}{D_1} < D_2, \quad D_1 D_2 + e^{D_1} \frac{C}{\sqrt{K}} \left(\frac{|x|^2}{\rho} \right)^K < D_1. \tag{92}$$

Then we have $|e_k(x)| < D_1$ for all $k \ge K$ and

$$e_k(x) = O(y(x)^k)$$

as $k \to \infty$, where the implicit constant might depend on x.

Remark. Note that the second inequality in (92) implies $|y(x)| < 1$ and thus an exponential decay of $e_k(x)$. Hence, in particular, the assumptions of this lemma imply that (80) holds.

Proof: By definition we have $e_{K+1} = y(1 - e^{-e_K - E_K})$. Hence, if we write $e^{-E_k} = 1 + R_k$ we obtain

$$|e_{K+1}| \le |e_K| \, |y| \frac{e^{|e_K|} - 1}{|e_K|} + e^{|e_K|} R_K.$$

If (92) is satisfied then it follows that

$$|e_{K+1}| \le D_1 D_2 + e^{D_1} \frac{C}{\sqrt{K}} \left(\frac{|x|^2}{\rho} \right)^K < D_1.$$

Now we can proceed by induction and obtain $|e_k| < D_1$ for all $k \ge K$. Note that $D_2 < 1$ and since $x \in \Delta$ we have $|x| \le \rho + \varepsilon$, Corollary 3 implies

$$|E_k(x)| \le \frac{C_0}{\sqrt{k}} \left| \frac{(\rho + \varepsilon)^2}{\rho} \right|^k < (\rho + 3\varepsilon)^k.$$

Moreover, $R_k = e^{-E_k} - 1 = O(|E_k|)$ and hence $\sum_k R_k = O(1)$. Thus we have

$$e_k = O(D_2^k).$$

If we set $a_k = y(1 - e^{-e_k})/e_k$ and $b_k = -y e^{-e_k} R_k$ we obtain the recurrence

$$e_{k+1} = e_k a_k + b_k$$

with an explicit solution of the form

$$e_k = e_K \prod_{K\le i<k} a_i + \sum_{K\le j<k} b_j \prod_{j<i<k} a_i.$$

Since $e_k = O(D_2^k)$, we have

$$\prod_{j<i<k} a_i = O\left(y(x)^{k-j}\right)$$

and hence by Corollary 3

$$b_j \prod_{j<i<k} a_i = O\left(y(x)^{k-j}|x^2/\rho|^j\right) = O\left(y(x)^k L^j\right)$$

for some L with $0 < L < 1$. Hence,

$$e_k = O\left(e_K y(x)^{k-K}\right) + O\left(y(x)^k L^K\right) = O\left(y(x)^k\right). \tag{93}$$

□

Recall that $e_k(x) \to 0$ for $|x| \le \rho$. Using Lemma 16 we deduce that $e_k \to 0$ in a certain region that extends the circle $|x| \le \rho$.

Lemma 17. *For every $\varepsilon > 0$ there exists $\delta > 0$ such that $e_k(x) \to 0$ at an exponential rate, if $|x| \le \rho + \delta$ and $|x - \rho| \ge \varepsilon$.*

Proof: If we show that Lemma 16 is applicable, then we are done. First observe that since the assumption on x implies $|y(x)| < 1$, for sufficiently small D_1 there exist $D_2 < 1$ and K (sufficiently large) such that the second and third inequality if (92) hold. Next note that $|e_K(x)| \le e_K(\rho)$ for $|x| \le \rho$ and thus $|e_K(x)| \le 2e_K(\rho)$ by continuity. By (86) we have $e_K(\rho) \sim \frac{2}{K}$ we can make $|e_K(x)|$ arbitrarily small and the proof is complete. □

Now we turn to the most important range, namely for $x \in \Delta$ and $|x - \rho| \le \varepsilon$.

Lemma 18. *There exists $\varepsilon > 0$ and a constant $c_1 > 0$ such that for all $x \in \Delta_\varepsilon$ the conditions (92) are satisfied for*

$$k = K(x) = \left\lfloor \frac{c_1}{|\arg(y(x))|} \right\rfloor$$

and properly chosen real numbers D_1, D_2. Consequently, $e_k(x) \to 0$ at exponential rate.

Proof: Suppose that $\arg(x)$ and $\arg(y(x))$ are positive and that $K(x) = \lfloor c_1/|\arg(y(x))|\rfloor$, where $c_1 = \arccos(1/4) - \varepsilon_1$ and $\varepsilon_1 > 0$ is arbitrarily small. Note that $K(x)$ can be made as large as we desire, since $y(x) \sim 1$ and therefore $\arg(y(x))$ is small for small ε.

Now fix an integer k_0 and ε small enough to guarantee $k_0 < K(x)$. Moreover, fix two small positive real numbers ε_2 and ε_3. First we will prove by induction that for $k_0 \le k \le K(x)$ we have (for sufficiently small ε)

$$|e_k(x)| \le \varepsilon_2, \tag{94}$$

$$\arg(e_k(x)) \le k \arg(y(x)) + \varepsilon_3, \tag{95}$$

$$|e_k(x)| \le c\frac{|y(x)|^k}{k} \text{ for some } c > 0. \tag{96}$$

The first step of the induction proof is to show (94). Observe that due to the choice of c_1 and formula (95) of the induction hypothesis we have

$$0 < \arg(e_k(x)) < \arccos(1/4). \tag{97}$$

This implies $|e_{k+1}| \le |e_k| + |E_k|$ and consequently, by the second statement of Lemma 15 and the property $e_{k_0}(\rho) \sim c/k_0$ (compare with (86)) it follows that

$$|e_k(x)| \le |e_{k_0}| + \sum_{k_0\le \ell<k} |E_\ell| < \varepsilon_2 \tag{98}$$

provided that k_0 is chosen sufficiently large.

Next we show (95). We start with (76) and obtain

$$\begin{aligned} e_{k+1} &= y(x)e_k\left(\frac{1-e^{-e_k}}{e_k} + O\left(\frac{E_k}{e_k}\right)\right) \\ &= y(x)e_k\left(1-\frac{e_k}{2} + O\left(e_k^2\right) + O\left(\frac{E_k}{e_k}\right)\right). \end{aligned} \tag{99}$$

Note that by (94) the first of the two error terms is much smaller than e_k. In order to estimate the second error term, note that by the second statement of Lemma 15 we know $E_k(x) = (\tilde{C}(x) + o(1))y(x^2)^k$. Combining this with (94) we obtain

$$\left|\frac{E_k}{e_k}\right| = O\left(\left|\frac{ky(x^2)^k}{y(x)^k}\right|\right) = O\left(L^k\right)$$

with some $0 < L < 1$. Since $|y(x^2)/y(x)| < \rho$ (a consequence of the convexity of $y(x)$ on the positive real line) whereas $|y(x)| \sim \rho$, we can have $L < |y(x)|$ provided that ε is small enough. But this together with (96) implies that also the second error term in (99) is small in comparison to e_k. Hence (97) implies that the argument of the last factor in (99) is negative. Thus we conclude

$$\begin{aligned} \arg(e_{k+1}(x)) &\le \arg(y(x)) + \arg(e_k) \\ &\le (k-k_0+1)\arg(y(x)) + \arg(e_{k_0}(x)) \\ &\le k\arg(y(x)) + \varepsilon_3 \end{aligned}$$

where the last inequality follows the fact that by $e_{k_0}(\rho) \sim c/k_0$ (compare with (86)) and continuity we can always achieve $|\arg(e_{k_0}(x))| < \varepsilon_3$.

The third step is to prove the lower bound (96) for $e_k(x)$ for $k_0 \le k \le K(x)$. of the form $|e_k(x)| \ge c|y(x)|^k/k$ for some $c > 0$. By Lemma 15 $E_k(x) = (\tilde{C}(x) + o(1))y(x^2)^k$ behaves nicely, if $|x-\rho| \le \varepsilon$. Suppose that $|x| \ge \rho$ and $x \in \Delta_\varepsilon$. Since $\arg(y(x^2))$ is of order $\arg(y(x))^2$ we deduce that $\arg(E_k(x)) = O(\arg(y(x)))$ for $k_0 \le k \le K(x)$. In particular, it follows that (for $k_0 \le k \le K(x)$)

$$|e_{k+1}(x)| = \left|y(x)(1-e^{-e_k(x)-E_k(x)})\right| \ge |y(x)|(1-e^{-|e_k(x)|}).$$

Treating the nonlinear recurrence $a_{k+1} = 1-e^{-a_k}$ with the methods of de Bruijn [7, p. 156], it is possible to show inductively that $a_k \sim \frac{c}{k}$ and thus $|e_k(x)| \ge c|y(x)|^k/k$ for some $c > 0$.

The last task is to find D_1 and D_2 such that the conditions (92) are satisfied for $k = K(x)$. In order to do this, we first show that in the formula

$$e_k = \frac{y^{k-k_0}}{\frac{1}{e_{k_0}} + \frac{1}{2}\frac{1-y^{k-k_0}}{1-y} + \sum_{k_0\le\ell<k} e_\ell h(e_\ell)y^\ell + \sum_{k_0\le\ell<k}\frac{S_\ell}{e_\ell^2}y^\ell}, \tag{100}$$

where k_0 is fixed, the second term in the denominator dominates. Since k_0 is fixed, the first term is bounded. For estimating the third term note that by (98) the terms in the sum can be made arbitrarily small if k_0 is chosen sufficiently large. Finally, due to the already obtained bounds $|e_k(x)| \ge c|y(x)|^k/k$, $E_k(x) = O(y(x^2)^k)$, and the property $S_\ell \sim -E_\ell$ the last term satisfies

$$\sum_{k_0\le\ell<k}\frac{S_\ell}{e_\ell^2}y^\ell = O(1)$$

and therefore it does not contribute to the main term, either. Summing up we have

$$|e_k| \le \frac{|y|^{k-k_0}}{\left|\frac{1}{2}\frac{1-y^{k-k_0}}{1-y}\right|}(1+\varepsilon_4)$$

for an arbitrarily small $\varepsilon_4 > 0$.

We set $\arg(\rho - x) = \theta$, where we assume that $\theta \in \left[-\frac{\pi}{2}-\varepsilon_5, \frac{\pi}{2}+\varepsilon_5\right]$ (for some $\varepsilon_5 > 0$ that has to be sufficiently small), and $r = b|\rho - x|^{1/2}$, where b is the constant appearing in (4). Then we have

$$\begin{aligned}|y| &= 1 - r\cos\frac{\theta}{2} + O(r^2),\\ \log|y| &= -r\cos\frac{\theta}{2} + O(r^2),\\ \arg(y) &= -r\sin\frac{\theta}{2} + O(r^2).\end{aligned}$$

Hence with $k = K(x) = \lfloor c_1/|\arg(y)|\rfloor$ we have

$$|y^{k-k_0}| \sim e^{-c_1\cot(\theta/2)+O(r^2)} \le e^{-c_1\cot(\frac{\pi}{4}+\frac{\varepsilon_5}{2})+O(r^2)} \le e^{-c_1}(1-\varepsilon_6)$$

for some arbitrarily small $\varepsilon_6 > 0$ (depending on ε_5). Consequently

$$|e_k| < D_1 := 2\frac{e^{-\arccos(1/4)}}{1-e^{-\arccos(1/4)}}r(1+\varepsilon_7) = c'r,$$

where $\varepsilon_7 > 0$ can be chosen arbitrarily small. Moreover

$$|y| = 1 - r\cos\frac{\theta}{2} + O(r^2) \le 1 - \frac{r}{\sqrt{2}}(1-\varepsilon_8)$$

for some (small) $\varepsilon_8 > 0$ and consequently

$$|y|\frac{e^{D_1}-1}{D_1} = 1 - \left(\frac{1}{\sqrt{2}} - \frac{e^{-\arccos(1/4)}}{1-e^{-\arccos(1/4)}}\right)r(1-\varepsilon_9) + O(r^2).$$

Thus, we are led to set

$$D_2 := 1 - \frac{1}{2}\left(\frac{1}{\sqrt{2}} - \frac{e^{-\arccos(1/4)}}{1-e^{-\arccos(1/4)}}\right)r = 1 - c''r$$

and the first two conditions of (92) are satisfied if r is sufficiently small. Since $D_1 - D_1D_2 = c'c''r^2$ we just have to check whether

$$e^{D_1}\frac{C}{\sqrt{k}}\left|\frac{x^2}{\rho}\right|^k < c'c''r^2.$$

However, since $k = K(x) \ge c_1\sqrt{2}\,r^{-1}$ the left hand side of this inequality is definitely smaller than $c'c''r^2$ if r is sufficiently small. Hence all conditions of (92) are satisfied for $k = K(x)$. □

Lemma 19. *There exists $\varepsilon > 0$ such that (78) holds for all x with $x \in \Delta_\varepsilon$ with $|x| \ge \rho$.*

Proof: We recall that the properties $e_k = O(1/k)$ and (79) imply (78). By Lemma 18 we already know that the first condition holds. Furthermore, we have upper bounds for E_k (see Lemma 15). Hence, it remains to provide proper lower bounds for e_k.

Since we already know that $e_k \to 0$ and $|E_k| = O(L^k)$ (for some $L < 1$), the recurrence (76) implies

$$|e_{k+1}| \ge (1-\delta)\left(|e_k| - |E_k|\right)$$

for some $\delta > 0$ provided that $x \in \Delta$ and $|x-\rho| < \varepsilon$. Without loss of generality we can assume that $L < (1-\delta)^2$. Hence

$$|e_k| \ge (1-\delta)^k - \sum_{\ell<k}|E_\ell|(1-\delta)^{k-\ell} \ge c_0(1-\delta)^k$$

for some constant $c_0 > 0$. Consequently

$$\left|\frac{E_k}{e_k^2}\right| = O\left(\left(\frac{L}{(1-\delta)^2}\right)^k\right).$$

As noted above, this upper bound is sufficient to deduce (78). □

Acknowledgements. We express our gratitude to Nicolas Broutin for his careful reading of an earlier version of our manuscript as well as to him and Philippe Flajolet for inspiring discussions on the topic and for providing us with a preprint of their paper [3]. Furthermore, we thank two anonymous referees whose detailed and very informative reports led to the correction of numerous imprecisions and errors in the manuscript as well as to a considerable improvement of the presentation.

References

[1] David Aldous. The continuum random tree. II. An overview. In *Stochastic analysis (Durham, 1990)*, volume 167 of *London Math. Soc. Lecture Note Ser.*, pages 23–70. Cambridge Univ. Press, Cambridge, 1991.

[2] Patrick Billingsley. *Convergence of probability measures.* John Wiley & Sons Inc., New York, 1968.

[3] N. Broutin and P. Flajolet. The height of random binary unlabelled trees. *Discrete Math. Theor. Comput. Sci. Proc.*, AI:121–134, 2008.

[4] 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, Basel, 2000.

[5] Brigitte Chauvin, Michael Drmota, and Jean Jabbour-Hattab. The profile of binary search trees. *Ann. Appl. Probab.*, 11(4):1042–1062, 2001.

[6] J. W. Cohen and G. Hooghiemstra. Brownian excursion, the $M/M/1$ queue and their occupation times. *Math. Oper. Res.*, 6(4):608–629, 1981.

[7] N. G. de Bruijn. *Asymptotic methods in analysis.* Dover Publications Inc., New York, third edition, 1981.

[8] N. G. de Bruijn, D. E. Knuth, and S. O. Rice. The average height of planted plane trees. In *Graph theory and computing*, pages 15–22. Academic Press, New York, 1972.

[9] Luc Devroye and Hsien-Kuei Hwang. Width and mode of the profile for some random trees of logarithmic height. *Ann. Appl. Probab.*, 16(2):886–918, 2006.

[10] M. Drmota. On nodes of given degree in random trees. In *Probabilistic methods in discrete mathematics (Petrozavodsk, 1996)*, pages 31–44. VSP, Utrecht, 1997.

[11] Michael Drmota. On Robson's convergence and boundedness conjectures concerning the height of binary search trees. *Theoret. Comput. Sci.*, 329(1-3):47–70, 2004.

[12] Michael Drmota. *Random trees.* SpringerWienNewYork, Vienna, 2009. An interplay between combinatorics and probability.

[13] Michael Drmota and Bernhard Gittenberger. On the profile of random trees. *Random Structures Algorithms*, 10(4):421–451, 1997.

[14] Michael Drmota and Bernhard Gittenberger. The distribution of nodes of given degree in random trees. *J. Graph Theory*, 31(3):227–253, 1999.

[15] Michael Drmota and Bernhard Gittenberger. The width of Galton-Watson trees conditioned by the size. *Discrete Math. Theor. Comput. Sci.*, 6(2):387–400 (electronic), 2004.

[16] Michael Drmota and Hsien-Kuei Hwang. Bimodality and phase transitions in the profile variance of random binary search trees. *SIAM J. Discrete Math.*, 19(1):19–45 (electronic), 2005.

[17] Michael Drmota and Hsien-Kuei Hwang. Profiles of random trees: correlation and width of random recursive trees and binary search trees. *Adv. in Appl. Probab.*, 37(2):321–341, 2005.

[18] Ph. Flajolet and A. M. Odlyzko. Singularity analysis of generating functions. *SIAM Journal on Discrete Mathematics*, 3:216–240, 1990.

[19] Philippe Flajolet, Zhicheng Gao, Andrew Odlyzko, and Bruce Richmond. The distribution of heights of binary trees and other simple trees. *Combin. Probab. Comput.*, 2(2):145–156, 1993.

[20] Philippe Flajolet and Andrew Odlyzko. The average height of binary trees and other simple trees. *J. Comput. System Sci.*, 25(2):171–213, 1982.

[21] Philippe Flajolet and Robert Sedgewick. *Analytic combinatorics.* Cambridge University Press, Cambridge, 2009.

[22] Michael Fuchs, Hsien-Kuei Hwang, and Ralph Neininger. Profiles of random trees: limit theorems for random recursive trees and binary search trees. *Algorithmica*, 46(3-4):367–407, 2006.

[23] Bernhard Gittenberger. Convergence of branching processes to the local time of a Bessel process. In *Proceedings of the Eighth International Conference "Random Structures and Algorithms" (Poznan, 1997)*, volume 13, pages 423–438, 1998.

[24] Bernhard Gittenberger. On the profile of random forests. In *Mathematics and computer science, II (Versailles, 2002)*, Trends Math., pages 279–293. Birkhäuser, Basel, 2002.

[25] Bernhard Gittenberger. Nodes of large degree in random trees and forests. *Random Structures Algorithms*, 28(3):374–385, 2006.

[26] Bernhard Gittenberger and Guy Louchard. The Brownian excursion multi-dimensional local time density. *J. Appl. Probab.*, 36(2):350–373, 1999.

[27] Frank Harary, Robert W. Robinson, and Allen J. Schwenk. Twenty-step algorithm for determining the asymptotic number of trees of various species. *J. Austral. Math. Soc. Ser. A*, 20(4):483–503, 1975.

[28] G. Hooghiemstra. On the explicit form of the density of Brownian excursion local time. *Proc. Amer. Math. Soc.*, 84(1):127–130, 1982.

[29] Hsien-Kuei Hwang. Profiles of random trees: plane-oriented recursive trees (extended abstract). In *2005 International Conference on Analysis of Algorithms*, Discrete Math. Theor. Comput. Sci. Proc., AD, pages 193–200 (electronic). Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2005.

[30] Hsien-Kuei Hwang. Profiles of random trees: plane-oriented recursive trees. *Random Structures Algorithms*, 30(3):380–413, 2007.

[31] Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1988.

[32] V. F. Kolchin. Branching processes, random trees and a generalized particle distribution scheme. *Mat. Zametki*, 21(5):691–705, 1977.

[33] Valentin F. Kolchin. *Random mappings*. Translation Series in Mathematics and Engineering. Optimization Software Inc. Publications Division, New York, 1986.

[34] Guy Louchard, Wojciech Szpankowski, and Jing Tang. Average profile of the generalized digital search tree and the generalized Lempel-Ziv algorithm. *SIAM J. Comput.*, 28(3):904–934 (electronic), 1999.

[35] J.-F. Marckert and G. Miermont. The crt is the scaling limit of unordered binary trees.

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

[37] A. Meir and J. W. Moon. On nodes of large out-degree in random trees. In *Proceedings of the Twenty-second Southeastern Conference on Combinatorics, Graph Theory, and Computing (Baton Rouge, LA, 1991)*, volume 82, pages 3–13, 1991.

[38] A. Meir and J. W. Moon. On nodes of given out-degree in random trees. In *Fourth Czechoslovakian Symposium on Combinatorics, Graphs and Complexity (Prachatice, 1990)*, volume 51 of *Ann. Discrete Math.*, pages 213–222. North-Holland, Amsterdam, 1992.

[39] Pierre Nicodème. Average profiles, from tries to suffix-trees. In *2005 International Conference on Analysis of Algorithms*, Discrete Math. Theor. Comput. Sci. Proc., AD, pages 257–266 (electronic). Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2005.

[40] Andrew M. Odlyzko and Herbert S. Wilf. Bandwidths and profiles of trees. *J. Combin. Theory Ser. B*, 42(3):348–370, 1987.

[41] Richard Otter. The number of trees. *Ann. Math.*, 49(2):583–599, 1948.

[42] Jim Pitman. The SDE solved by local times of a Brownian excursion or bridge derived from the height profile of a random tree or forest. *Ann. Probab.*, 27(1):261–283, 1999.

[43] George Pólya. Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und chemische Verbindungen. *Acta Math.*, 68:145–254, 1937.

[44] Robert W. Robinson and Allen J. Schwenk. The distribution of degrees in a large random tree. *Discrete Math.*, 12(4):359–372, 1975.

[45] N. J. A. Sloane. *The on-line Encyclopedia of Integer Sequences*. Published electronically at `http://www.research.att.com/~njas/sequences/`., 2006.

[46] V. E. Stepanov. The distribution of the number of vertices in the layers of a random tree. *Teor. Verojatnost. i Primenen.*, 14:64–77, 1969.

[47] Lajos Takács. Conditional limit theorems for branching processes. *J. Appl. Math. Stochastic Anal.*, 4(4):263–292, 1991.

[48] Lajos Takács. On the distribution of the number of vertices in layers of random trees. *J. Appl. Math. Stochastic Anal.*, 4(3):175–186, 1991.

[49] Remco van der Hofstad, Gerard Hooghiemstra, and Piet Van Mieghem. On the covariance of the level sizes in random recursive trees. *Random Structures Algorithms*, 20(4):519–539, 2002.