NODES OF LARGE DEGREE IN RANDOM TREES AND FORESTS

BERNHARD GITTENBERGER*

ABSTRACT. We study the asymptotic behaviour of the number $N_{k,n}$ of nodes of given degree k in unlabeled random trees, when the tree size n and the node degree k both tend to infinity. It is shown that $N_{k,n}$ is asymptotically normal if $\mathbf{E}N_{k,n} \to \infty$ and asymptotically Poisson distributed if $\mathbf{E}N_{k,n} \to C > 0$. If $\mathbf{E}N_{k,n} \to 0$, then the distribution degenerates. The same holds for rooted, unlabeled trees and forests.

1. INTRODUCTION

1.1. **Statement of the Problem.** Consider the sequence of discrete probability spaces $(\mathcal{T}_n, \mathcal{P}(\mathcal{T}_n), P_n)$, where $\mathcal{T}_n$ denotes the set of unrooted, unlabeled trees with n nodes, $\mathcal{P}(A)$ the power set of A, and P_n the uniform probability measure on $\mathcal{T}_n$. We are interested in the numbers $N_{k,n}$ of nodes of degree k in a random tree in $\mathcal{T}_n$.

It is well known that the number of nodes of fixed degree k (see [25]) is almost proportional to the size of the tree (to be precise: as $n \to \infty$ we have for k fixed $\mathbf{E}N_{k,n} \sim c_k n$ for some $c_k > 0$). Moreover, it is known (see [8]) that $N_{k,n}$ is asymptotically normally distributed.

A natural question is to ask for the behavior of the limiting distribution if we let k grow to infinity as well. When looking at some random mapping parameters (see [2] and [9]), one could expect that there is an order of magnitude for k where the asymptotic normality is no longer valid, but replaced by a Poisson limit law. The limiting behaviour depends on the behaviour of $\mathbf{E}N_{k,n}$. The asymptotic normal limit law still holds as long as $\mathbf{E}N_{k,n} \to \infty$. Note that $\mathbf{E}N_{k,n}$ is decreasing in k. This determines the allowed growth rate for k to preserve the normal limit law. If $\mathbf{E}N_{k,n} \to C > 0$, then the limiting distribution is Poisson. In the case where $\mathbf{E}N_{k,n} \to 0$ the distribution clearly degenerates. A precise formulation of these three cases will be given in the next section (Theorem 2) after presenting the definitions needed.

If we consider the set $\mathcal{T}_n^{(r)}$ of unlabeled, rooted trees instead of $\mathcal{T}_n$, then the behaviour of the analogous random variable satisfies the same limit theorems. The same is true for forests.

1.2. **Historical Remarks.** Earliest references considering the node degree in random trees seem to go back to Otter [22] where among other things trees with certain restrictions on their degree sequence are counted. Later Riordan [24] enumerated trees with given degree sequence. Moon considered nodes of degree one (see [19]) and the maximum degree (see [20]).

The maximum degree has been further studied by Carr, Goh, and Schmutz [4, 11, 10] for various tree classes. The degree distribution for Pólya trees has been investigated by Robinson and Schwenk [25, 26] and Bailey [1]. Simply generated trees have been studied in this respect by Meir and Moon [14, 15, 16, 17, 18] and in a different context by Drmota [6] and Moon and Prodinger [21]. The distribution of the number of nodes of large degree in simply generated trees has been studied in [17]. Multivariate distributions for several tree classes can be found in [8]. For a survey of results for random trees and connections to other random structures (e.g. mappings) as well, see [12].

Date: August 2, 2005.
* Department of Discrete Mathematics and Geometry, Technische Universität Wien, Wiedner Hauptstraße 8-10/104, A-1040 Wien, Austria.
This work has been supported by the Austrian Science Foundation FWF, grant P16053-N05.

1.3. **Plan of the Paper.** In the next section we state some auxiliary results, namely functional equations for the generating functions associated to the various tree classes under consideration. After this we are able to state the main result (Theorem 2).

The final section is devoted to the proof of Theorem 2. We first present an outline of the proof which relies on an appropriate representation of the probability generating function of $\psi_{k,n}(u) = \mathbf{E}u^{N_{k,n}}$. This representation of $\psi_{k,n}(u)$ which is given in Proposition 1 is used to establish Theorem 2 by means of characteristic functions.

In order to be able to do this, this representation has to be analysed in detail. In fact, we need expressions for the local behaviour of the associated generating functions near their singularities which are uniform in k. This task is done in the second part of the last section. The section closes with the proof of Proposition 1

Remark. We would like to mention that our method is applicable to simply generated trees and forests as well, since the generating functions of these classes satisfy simpler functional equations than those for unlabeled trees. In fact, the uniformity of the expansions near the singularity can be easier established by means of Banach's fixed point theorem. The corresponding results have already been obtained by Meir and Moon [17] using different methods.

2. The Generating Functions Related to the Problem

2.1. **Univariate Functions.** Let t_n and $t_n^{(r)}$ denote the cardinalities of the sets $\mathcal{T}_n$ and $\mathcal{T}_n^{(r)}$, respectively. Pólya [23] already discussed the generating function

$$t^{(r)}(x) = \sum_{n\geq 1} t_n^{(r)} x^n$$

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 $|x| = \rho$. Refinements are due to Otter [22] who showed $t^{(r)}(\rho) = 1$ and used the expansion

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

to deduce that

$$t_n^{(r)} \sim \frac{b\sqrt{\rho}}{2\sqrt{\pi}} n^{-3/2} \rho^{-n}. \tag{2}$$

He also calculated $c = b^2/3 \approx 2.3961466$, $\rho \approx 0.3383219$, and $b \approx 2.6811266$. Moreover, he related the generating functions of rooted and unrooted trees:

$$t(x) = \sum_{n\geq 1} t_n x^n = t^{(r)}(x) - \frac{1}{2} t^{(r)}(x)^2 + \frac{1}{2} t^{(r)}(x^2).$$

Hence $t(x)$ has a similar expansion, namely

$$t(x) = \frac{1 + t^{(r)}(\rho^2)}{2} + \frac{b^2 - \rho (t^{(r)})'(\rho^2)}{2} (\rho - x) + bc(\rho - x)^{3/2} + \cdots ,$$

and it follows that

$$t_n \sim \frac{b^3 \rho^{3/2}}{4\sqrt{\pi}} n^{-5/2} \rho^{-n}.$$

Robinson and Schwenk [25] extended Pólya's and Otter's method to obtain $\mathbf{E}N_{k,n} \sim \mu_k n$ ($n \to \infty$, k fixed). The asymptotic behavior of μ_k is given by

$$\mu_k \sim K\rho^k$$

where $K \approx 6.380045$, see [26].

2.2. **Multivariate Functions.** It proves convenient to introduce the class of so-called planted trees, where an additional edge without node is attached to the root of a rooted tree. This increases the node degree of the root by one and simplifies the setting up of the functional equations for the generating functions. Let $t_{nmk}^{(\cdot)}$ denote the number of trees with n vertices m of which have degree k. The superscripts u, r, and f indicate unrooted, rooted trees, and forests (of unrooted trees), respectively. No superscript stands for planted trees. Accordingly, define the generating functions

$$t_k^{(\xi)}(x,u) = \sum_{n,m} t_{nmk}^{(\xi)} x^n u^m, \text{ for } \xi \in \{u,r,f\} \text{ and } t_k(x,u) = \sum_{n,m} t_{nmk} x^n u^m.$$

It is well known from Pólya's enumeration theory (cf. [25] and [8] as well) that these generating functions satisfy the following functional equations:

Let $Z_k(x_1,\ldots,x_k)$ denote cycle index of the symmetric group S_k of k elements. Then we have

$$t_k(x,u) = x\exp\left(\sum_{i\ge1}\frac{t_k(x^i,u^i)}{i}\right) + x(u-1)Z_{k-1}(t_k(x,u),t_k(x^2,u^2),\ldots,t_k(x^{k-1},u^{k-1})), \quad (3)$$

$$t_k^{(r)}(x,u) = x\exp\left(\sum_{i\ge1}\frac{t_k(x^i,u^i)}{i}\right) + x(u-1)Z_k(t_k(x,u),t_k(x^2,u^2),\ldots,t_k(x^k,u^k)), \quad (4)$$

$$t_k^{(u)}(x,u) = t_k^{(r)}(x,u) - \frac{1}{2}t_k(x,u)^2 + \frac{1}{2}t_k(x^2,u^2), \quad (5)$$

$$t_k^{(f)}(x,u) = \exp\left(\sum_{i\ge1}\frac{t_k^{(u)}(x^i,u^i)}{i}\right). \quad (6)$$

The distribution of $N_{k,n}$ has been determined in [8] with the help of these multivariate functions. They showed

Theorem 1. *Let $N_{k,n}$ denote the number of nodes in an unrooted or rooted unlabeled random tree or forest of n nodes that have degree k. Set*

$$F(x,u,y) = xe^y\exp\left(\sum_{i\ge2}\frac{t_k(x^i,u^i)}{i}\right) + x(u-1)Z_{k-1}(y,t_k(x^2,u^2),\ldots,t_k(x^{k-1},u^{k-1})).$$

Furthermore, let

$$f_1 = -\frac{F_u}{F_x}(\rho,1,1),$$

$$f_2 = \left[\frac{1}{F_{yy}F_x}\left(\frac{F_uF_{xy}}{F_x} - F_{xy}\right)^2 - \frac{1}{F_x}\left(\frac{F_u^2F_{xx}}{F_x^2} - \frac{2F_uF_{ux}}{F_x} + F_{uu}\right)\right](\rho,1,1),$$

and

$$\mu_k = -\frac{f_1}{\rho}, \quad \sigma_k^2 = \left(\frac{f_1}{\rho}\right)^2 - \frac{f_2}{\rho}. \quad (7)$$

Then, as $n\to\infty$, $N_{k,n}$ is asymptotically normally distributed with mean value $\sim \mu_k n$ and variance $\sim \sigma_k^2 n$. Moreover for large k:

$$\mu_k \sim \frac{2C}{b^2\rho}\rho^k, \quad \sigma_k^2 \sim \frac{2C}{b^2\rho}\rho^k \text{ where } C = \exp\left(\frac{1}{l}\left(\frac{t_k(\rho^l,1)}{\rho^l} - 1\right)\right) \approx 7.7581604. \quad (8)$$

Remark 1. The theorem in [8] is more general: Actually, choosing a finite number of fixed degrees $k_1,\ldots,k_d$ gives convergence to a multivariate normal distribution. For brevity, we stated only the univariate version.

Remark 2. Similar limit theorems hold for other tree classes as well. For instance, in [8] analogous theorems for plane trees as well as labeled trees are given.

Theorem 2. *Let $N_{k,n}$ denote the number of nodes of degree k in an unrooted or rooted, unlabeled random tree or forest with n nodes, and let $k = k(n)$ be a sequence depending on n. Then the following implications are valid:*

- *If $\lim_{n\to\infty} \mathbf{E}N_{k,n} = \infty$, then the assertion of Theorem 1 is still true, i.e., $N_{k,n}$ is asymptotically normally distributed with mean and variance given by (7).*
- *If there exists a subsequence n_i such that $\lim_{i\to\infty} \mathbf{E}N_{n_i,k(n_i)} = \beta > 0$, then $N_{k(n_i),n_i}$ has a Poisson limiting distribution, precisely, for any $\ell \geq 0$*

$$\lim_{i\to\infty} \mathbf{P}\left\{N_{k(n_i),n_i} = \ell\right\} = e^{-\beta}\frac{\beta^\ell}{\ell!}$$

- *If $\lim_{n\to\infty} \mathbf{E}N_{n,k} = 0$, then $N_{k,n}$ degenerates, i.e., $\lim_{n\to\infty} \mathbf{P}\left\{N_{k,n} = 0\right\} = 1$.*

Remark. The second assertion has to be restricted to a subsequence n_i, because $k = k(n)$ cannot be chosen integer valued such that $\mathbf{E}N_{k,n} \sim Cn\rho^k$ is bounded away from zero and infinity and the limit exists.

3. Proof of Theorem 2

3.1. **Outline of the Proof.** The proof of Theorem 2 relies on the asymptotic representation of the probability generating function of $N_{k,n}$, given by $\mathbf{E}u^{N_{k,n}}$, as a power of a function. Moreover, we need this representation to be uniform in k. In fact, with the help of this proposition Theorem 2 is easily proved.

Proposition 1. *Let ρ be as above and $\rho_k(u)$ be the singularity of the function $x \mapsto t_k(x,u)$ on its circle of convergence, where we assume $|u-1| < \varepsilon$ for some sufficiently small $\varepsilon > 0$. (This implies $\rho_k(1) = \rho$, of course.) Then the probability generating function $\psi_{k,n}(u) = \mathbf{E}u^{N_{k,n}}$ has a uniform asymptotic representation*

$$\psi_{k,n}(u) = \rho^n \rho_k(u)^{-n}\left(1 + O\left(k^{-1/2}\right)\right), \quad \text{as } n \to \infty, \tag{9}$$

uniformly for $|u-1| < \varepsilon$ and sufficiently large k. Furthermore, $\rho_k(u)$ satisfies

$$\rho_k(u) = \rho - f_1(k)(u-1) + f_2(k)(u-1)^2 + O\left(k^2\rho^{3k}|u-1|^3\right) \tag{10}$$

with

$$f_1(k) = \Theta(\rho^k), \quad f_2(k) = o(\rho^k) \text{ as } k \to \infty. \tag{11}$$

From (9) we have, as $n \to \infty$ and uniformly for sufficiently large k

$$\psi_{k,n}(u) \sim \rho^n\rho_k(u)^{-n} = \left(1 + \frac{\rho_k'(1)}{\rho_k(1)}(u-1) + \frac{\rho_k''(1)}{\rho_k(1)}(u-1)^2 + O\left(\rho_k'''(1+\vartheta(u-1))(u-1)^3\right)\right)^{-n}$$

for some $0 < \vartheta < 1$. Thus setting $u = e^{it}$ and using (10) we obtain

$$\psi_{k,n}\left(e^{it}\right) = \left(1 + it\frac{\rho_k'(1)}{\rho_k(1)} - \frac{t^2}{2}\left(\frac{\rho_k'(1)}{\rho_k(1)} - \frac{\rho_k''(1)}{\rho_k(1)}\right) + O\left(\rho^k t^3\right)\right)^{-n}$$

Now using (11), the characteristic function of

$$\frac{N_{k,n} + n\rho_k'(1)/\rho_k(1)}{\sqrt{-n\rho_k'(1)/\rho_k(1)}}$$

is given by

$$\exp\left(-it\sqrt{-n\frac{\rho_k'(1)}{\rho_k(1)}}\right)\psi_{k,n}\left(\exp\left(it \bigg/ \sqrt{-n\frac{\rho_k'(1)}{\rho_k(1)}}\right)\right) = \exp\left(-\frac{t^2}{2} + O\left(t^3(n\rho^k)^{-1/2}\right)\right)$$

which tends to $e^{-t^2/2}$ if $n\rho^k \to \infty$. Thus we have asymptotic normality in this case.

If $\mathbf{E}N_{k,n} \to \beta$ (here $k = k(n_i)$), then

$$\psi_{k,n_i}\left(e^{it}\right) = \left(1 + \frac{\rho_k'(1)}{\rho_k(1)}(e^{it}-1) + O\left(k\rho^{2k}t^2\right)\right)^{-n_i} \to \exp\left(\beta(e^{it}-1)\right)$$

which is the characteristic function of the Poisson distribution with parameter β, if $\beta > 0$. If $\beta = 0$ then we have obviously almost sure convergence to 0.

In order to prove Proposition 1 two main ingredients are needed. First, we need to know the local behaviour of $t_k(x,u)$ near its singularity in terms of an asymptotic expansion which is *uniform* in k. This will be provided in Lemma 3. $t_k(x,u)$ is determined by a functional equation which is closely related to the easier one for Cayley trees. The difference is the perturbation, i.e., the term involving the cycle index. Hence control over this term is needed. Lemmas 1 and 2 provide estimates for the cycle index of the symmetric group which enables us to prove Lemma 3.

Second, we need estimates for the derivates of $\rho_k(u)$ and related functions which can be obtained by implicit differentiation of suitable functional equations (see Lemma 4).

After working out these two tasks, we will present the proof of Proposition 1 to complete the proof of Theorem 2.

3.2. **Actual Proof of Theorem 2.** We start with a short analysis of the cycle index Z_k. The first estimate has been shown by Schwenk [26].

Lemma 1. *If there exists an x with $0 < x < 1$ such that $A_i - 1 \le \lambda x^i$ for $i = 1, \dots, k$ then*

$$Z_k(A_1, \dots, A_k) \sim \exp\left(\sum_{i=1}^{k} \frac{A_i - 1}{i}\right), \text{ as } k \to \infty.$$

Proof. Inspection of the proof of Theorem 3.1 in [26] immediately yields the result. □

Lemma 2. *There exists an $\varepsilon > 0$ such that for $|x-\rho| < \varepsilon$, $|u-1| < \varepsilon$, $a = \rho + 2\varepsilon$, and for all complex y we have*

$$\left|Z_k\left(y, t_k(x^2,u^2), \dots, t_k(x^k,u^k)\right)\right| \le Z_k\left(|y|, t^{(r)}(a^2), \dots, t^{(r)}(a^k)\right). \tag{12}$$

Let $\rho_k(u)$ be as in (14). Then, with x, u, a as above with the restriction $\arg\left(\frac{x}{\rho_k(u)} - 1\right) \neq 0$, *we have*

$$\begin{aligned} Z_k(|t_k(x,u)|, t^{(r)}(a^2), \dots, t^{(r)}(a^k)) &= a^k Z_k\left(\frac{|t_k(x,u)|}{a}, \frac{t^{(r)}(a^2)}{a^2}, \dots, \frac{t^{(r)}(a^k)}{a^k}\right) \\ &\sim a^k \exp\left(\frac{|t_k(x,u)|}{a} - 1 + \sum_{i=2}^{k} \frac{1}{i}\left(\frac{t^{(r)}(a^i)}{a^i} - 1\right)\right), \end{aligned} \tag{13}$$

as $k \to \infty$.

Proof. First observe that the exponents of u in $t_k(x,u)$ are always less than those of x, since there cannot be any node of degree larger or equal the tree size. Hence $|u| \le 1 + \varepsilon$ implies $|t_k(x,u)| \le t^{(r)}(|x|(1+\varepsilon))$. The fact $\rho < 1$ guarantees that $a^2, \dots, a^k$ are inside the circle of convergence of $t^{(r)}(x)$ provided that ε is sufficiently small. Consequently, if $|u-1| < \varepsilon$ and $|x-\rho| < \varepsilon$, then the cycle index Z_k can be estimated by (12).

To show (13) set

$$A_1 = \frac{|y|}{a}, \quad A_i = \frac{t^{(r)}(a^i)}{a^i} \text{ for } i = 1, \dots, k.$$

$t_k(x,u)$ satisfies the functional equation (3). Thus there exists a neighborhood of $(\rho, 1)$ such that $t_k(x,u)$ is uniformly bounded for (x,u) as stated above. Besides, (3) implies $t^{(r)}(0) = t_k(0,u) = 0$. Thus the assumptions of Lemma 1 are fulfilled. Hence (13) follows from the fact that Z_k is homogeneous of degree k and from Lemma 1. □

Next we cite a proposition which yields a first expansion of the function $t_k(x,u)$ (compare with [5] for special cases and [7, 13, 27] for a general formulation). Here and in the sequel the subscripts u, x, and y will denote partial derivatives.

Proposition 2. *Suppose that $F(x,u,y)$ is an analytic function around (x_0,u_0,y_0) such that*

$$\begin{aligned} F(x_0,u_0,y_0) &= y_0, \\ F_y(x_0,u_0,y_0) &= 1, \\ F_{yy}(x_0,u_0,y_0) &\neq 0, \\ F_x(x_0,u_0,y_0) &\neq 0. \end{aligned}$$

Then there exist a neighborhood U of (x_0,u_0), a neighborhood V of y_0 and analytic functions $a(x,u)$, $b(x,u)$, and $c(x,u)=c(u)$ (i.e., $c(x,u)$ is a function which is independent of x) which are defined on U, such that the only solutions $y\in V$ with $y=F(x,u,y)$ ($(x,u)\in U$) are given by

$$y = a(x,u) \pm b(x,u)\sqrt{1-\frac{x}{c(u)}}$$

Furthermore $a(x_0,u_0)=y_0$ and $b(x_0,u_0)=\sqrt{2c(u_0)F_x(x_0,u_0,y_0)/F_{yy}(x_0,u_0,y_0)}$.

Proof. See [7]. □

An immediate consequence (which was used in [8]) is that for fixed k, there exist analytic functions $a_k(x,u)$, $b_k(x,u)$ and $\rho_k(u)$ such that

$$t_k(x,u) = a_k(x,u) - b_k(x,u)\sqrt{1-\frac{x}{\rho_k(u)}} \tag{14}$$

This representation has two (technical) disadvantages. First, k is fixed and it is not obvious what happens if k grows large. Second, a_k and b_k depend on x whereas $\psi_{k,n}$ in Proposition 1 does not (of course). The next lemma gives a representation for t_k without those disadvantages.

Lemma 3. *Let ρ_k as in (14). Then there exist analytic functions g_k and h_k and constants $\varepsilon>0$ and $k_0>0$ such that*

$$t_k(x,u) = g_k(u) - h_k(u)\sqrt{1-\frac{x}{\rho_k(u)}} + O\left(\left|1-\frac{x}{\rho_k(u)}\right|\right) \tag{15}$$

for $|u-1|<\varepsilon$, $|x-\rho_k(u)|<\varepsilon$, such that $\arg\left(\frac{x}{\rho_k(u)}-1\right)\neq 0$, *and uniformly for sufficiently large k.*

Remark. Note that Lemma 3 together with the results in [3] guarantees that the equation $\mathbf{E}N_{n,k} = \mu_k n + O(1)$ holds uniformly in k.

Proof. Taylor expansion of a_k and b_k of (14) w.r.t. x at $x=\rho_k(u)$ yields (15) with $g_k(u) = a_k(\rho_k(u),u)$, and $h_k(u)=b_k(\rho_k(u),u)$.

What remains to be done is showing that the error term in (15) is uniform in k. Therefore let

$$F(x,u,y) = xe^yQ(x,u) + x\cdot(u-1)Z_{k-1}(y,t_k(x^2,u^2),\ldots,t_k(x^{k-1},u^{k-1})) - y. \tag{16}$$

where $Q(x,u)=\exp\left(\sum_{i\geq 2} t_k(x^i,u^i)/i\right)$.

We first show that $y=g_k(u)$ and $x=\rho_k(u)$ are the unique solutions of the system of functional equations

$$F(x,u,y) \equiv 0, \tag{17}$$

$$F_y(x,u,y) \equiv 0. \tag{18}$$

Observe that (17) coincides with (3) and hence by (14) $y=g_k(u)$ and $x=\rho_k(u)$ are solutions of (17).

Note that since $0<\rho<1$ and (x,u) is close to $(\rho,1)$, (x^2,u^2) is inside the domain of convergence of $t_k(x,u)$. Therefore $Q(x,u)$ as well as $t_k(x^i,u^i)$ are analytic functions. Thus the cycle index on the right-hand side of (16) is a polynomial in y with analytic functions as coefficients

t_k is the solution of (17) in a neighborhood U of $(\rho,1)$. Set $(x_0,u_0)=(\rho_k(u),u)$ for some u sufficiently close to 1, in particular such that $(x_0,u_0)\in U$. By analyticity the other assumptions of Proposition 2 are still satisfied. Then (17) has a unique solution, say $\tilde{y}$, satisfying (15) (using the Taylor argument above). Since t_k is the solution of (17) in U, it must coincide with $\tilde{y}$. By the implicit function theorem (18) characterizes the singular points $(\rho_k(u),u)$ of $t_k(x,u)$. Therefore (18) is fulfilled by $(\rho_k(u),u,t_k(\rho_k(u),u))$. But by (15) this is equal to $(\rho_k(u),u,g_k(u))$ and thus the assertion above is proved.

Now we turn to the uniformity of the error term in (15). Expand F in a Taylor series at $(\rho_k(u),u,g_k(u))$. Using $F(\rho_k(u),u,g_k(u))\equiv 0$ and $F_y(\rho_k(u),u,g_k(u))\equiv 0$ we get (omitting the argument u in ρ_k and g_k)

$$\begin{aligned}F(x,u,y) = F_x(\rho_k,u,g_k)(x-\rho_k) &+ \frac{1}{2}F_{yy}(\rho_k,u,g_k)(y-g_k)^2 + F_{xy}(\rho_k,u,g_k)(x-\rho_k)(y-g_k)\\ &+ F_{xx}(\rho_k,u,g_k)(x-\rho_k)^2 + O\left(|x-\rho_k|^3+|y-g_k|^3\right).\end{aligned}$$

Since $F(x,u,t_k(x,u))\equiv 0$, setting $y=t_k$ yields

$$\begin{aligned}(t_k-g_k)^2 = &\frac{2\rho_k F_x(\rho_k,u,g_k)}{F_{yy}(\rho_k,u,g_k)}\left(1-\frac{x}{\rho_k}\right)\\ &+ O\left(F_{xy}(\rho_k,u,g_k)|x-\rho_k|\cdot|t_k-g_k| + F_{xx}(\rho_k,u,g_k)|x-\rho_k|^2\right). \qquad (19)\end{aligned}$$

To proceed we need uniform estimates for ρ_k, F_x, F_{yy}, F_{xy} and F_{xx}.

Next fix $\varepsilon_2>0$ and k. Then by the analyticity of $\rho_k(u)$ there exists an $\varepsilon_1>0$ such that $|\rho_k(u)-\rho|<\varepsilon_2$ for $|u-1|<\varepsilon_1$. Now observe that $t^{(r)}(x)$ satisfies the functional equation $t^{(r)}(x)=xe^{t^{(r)}(x)}Q(x,1)$ and has the only singularity ρ on the circle of convergence. The functional equation for $t_k(x,u)$ is obtained from the one for $t^{(r)}(x)$ by adding the cycle index as perturbation. But by Lemma 2 the cycle index decreases exponentially for $|u-1|<\eta:=\min(\varepsilon_1,\varepsilon_2)$ and $|x-\rho|<\eta$ (provided that ε_2 was chosen sufficiently small). Hence $|\rho_k(u)-\rho|<\eta$ for sufficiently large k. Finally, combining the continuity of $a_k(x,u)$ with equations (17) and (18) shows $|g_k(u)-1|<\varepsilon_3$ for sufficiently large k.

These estimates in conjunction with (12) and (13) can be applied to bound the partial derivatives of $F(x,u,y)$ which on their part can be calculated from (17) by implicit differentiation. By Lemma 2 we get uniform upper bounds for the expressions obtained in this way, in particular for F_x, F_{xy}, and F_{xx}.

Moreover, for $|u-1|\le\varepsilon$ with sufficiently small $\varepsilon>0$ we have uniformly for k large enough $|F_{yy}(\rho_k,u,g_k)|\ge\eta>0$. Thus we finally get from (19)

$$(t_k-g_k)^2 = \Theta(1)\left(1-\frac{x}{\rho_k}\right) + O\left(|x-\rho_k|\cdot|t_k-g_k|+|x-\rho_k|^2\right), \qquad (20)$$

uniformly in k. This implies $t_k-g_k=O\left(\sqrt{1-\frac{x}{\rho_k}}\right)$ and, inserted into (20) again, yields (15) after all. $\square$

Lemma 4. *The functions ρ_k, g_k, and h_k satisfy, as $u\to 1$,*

$$\rho_k(u) = \rho - f_1(k)(u-1) + f_2(k)(u-1)^2 + O\left(k^2\rho^{3k}|u-1|^3\right) \qquad (21)$$

$$g_k(u) = 1 - f_3(k)(u-1) + O\left(k\rho^{2k}|u-1|^2\right) \qquad (22)$$

$$h_k(u) = b\sqrt{\rho} + O\left(\rho^k|u-1|\right) \qquad (23)$$

uniformly in k, where f_1, f_2, and f_3 are functions of k. Moreover, the functions $f_i(k)$ satisfy the asymptotic relations

$$f_1(k)\sim\frac{2C}{b^2}\rho^k, \quad f_2(k)\sim\frac{C^2}{b^4}k\rho^{2k}, \quad f_3(k)\sim\left(C+\frac{2C}{b^2}\right)\rho^k,$$

where b and C are defined in (1) and (8), respectively.

Proof. Let $F(x,u,y)$ as in (16). Note that $F(\rho_k(u),u,g_k(u))\equiv 0$. Hence implicit differentiation yields ρ_k', ρ_k'', and ρ_k''' as well as g_k' and g_k'' in terms of partial derivatives of F up to order 3.

Now from [26, Corollary 4.1] and its generalizations in [8, pp.248] we know that at $(x,u,y) = (\rho,1,1)$ the following relations hold:

$$F_y = F_{yy} = F_{yyy} = 1, \qquad F_x = F_{xy} = F_{xyy} = \frac{b^2}{2}, \qquad F_u \sim F_{uy} \sim F_{uyy} \sim C\rho^k,$$
$$F_{ux} \sim F_{uxy} \sim Ck\rho^k, \qquad F_{uu} = o\left(\rho^{2k}\right), \qquad F_{uuy} = o\left(\rho^{2k}\right)$$

F_{xx} and F_{xxx} are independent of k. Similarly, it can be shown that $F_{uxx} = O\left(k^2\rho^k\right)$, $F_{uux} = o\left(k\rho^{2k}\right)$, and $F_{uuu} = o\left(\rho^{3k}\right)$. This implies

$$\rho_k'(1) \sim -\frac{2C}{b^2}\rho^k, \quad \rho_k''(1) \sim \frac{2C^2}{b^4}k\rho^{2k}, \quad g_k'(1) \sim -\left(C + \frac{2C}{b^2}\right)\rho^k$$

and $\rho_k''' = O\left(k^2\rho^{3k}\right)$, $g_k'' = O\left(k\rho^{2k}\right)$. Thus (21) and (22) are proved.

By Proposition 2 we have $h_k(u) = \sqrt{2\rho_k(u)F_x(\rho_k(u),u,g_k(u))/F_{yy}(\rho_k(u),u,g_k(u))}$. Thus, inserting the expansions for ρ_k and g_k finally proves (23). □

Proof of Proposition 1. To prove this we can follow closely [2]. We use (15) and Cauchy's formula and obtain

$$\psi_{k,n}(u) = \frac{1}{t_n^{(r)}}[x^n]t_k(x,u) = \frac{1}{2\pi i t_n^{(r)}}\int_\Gamma t_k(x,u)\,\frac{dx}{x^{n+1}},$$

with the integration contour $\Gamma = \Gamma_1 \cup \Gamma_2$ with

$$\Gamma_1 = \left\{x = \rho_k(u)\left(1+\frac{s}{n}\right) : s \in \gamma'\right\},$$
$$\Gamma_2 = \left\{x = Re^{i(\vartheta - \arg(u))} : R = |\rho_k(u)|\left|1 + \frac{\log^2 n + i}{n}\right|,\right.$$
$$\left.\arg\left(1+\frac{\log^2 n + i}{n}\right) \le |\vartheta| \le \pi\right\}.$$

where $\gamma' = \{s : |s| = 1, \Re s \le 0\} \cup \{s : 0 < \Re s < \log^2 n, \Im s = \pm 1\}$ is a truncated Hankel contour $\gamma = \{s : |s| = 1, \Re s \le 0\} \cup \{s : 0 < \Re s < \infty, \Im s = \pm 1\}$.

Note that t_k is uniformly bounded on Γ_2 by virtue of $|t_k(x,u)| \le t^{(r)}(|x|)$. Moreover, we have $|x|^{-n} = O\left(\rho_k(u)^{-n}e^{-\log^2 n}\right)$ for $x \in \Gamma_2$. Hence we get

$$\frac{1}{2\pi i t_n^{(r)}}\int_{\Gamma_2} t_k(x,u)\,\frac{dx}{x^{n+1}} = O\left(\rho^n \rho_k(u)^{-n} n^{3/2} e^{-\log^2 n}\right)$$

and therefore this part is of no importance.

On Γ_1 we have (substituting $x = \rho_k(u)(1+s/n)$ and using Lemma 3)

$$\begin{aligned}
\frac{1}{2\pi i}\int_{\Gamma_1} t_k(x,u) &= \frac{1}{2\pi i}\int_{\Gamma_1}\left[-h_k(u)\sqrt{1-\frac{x}{\rho_k(u)}} + O\left(\left|1-\frac{x}{\rho_k(u)}\right|\right)\right]\frac{dx}{x^{n+1}}\\
&= \frac{\rho_k(u)^{-n}}{n}\frac{1}{2\pi i}\int_{\gamma'} -h_k(u)\sqrt{-\frac{s}{n}}\,e^{-s}\left(1 + O\left(\sqrt{\frac{|s|}{n}}\right)\right)ds\\
&= \frac{h_k(u)\rho_k(u)^{-n}}{2\sqrt{\pi}n^{3/2}}\left(1+O\left(\frac{1}{\sqrt{k}}\right)\right) \qquad (24)
\end{aligned}$$

where the last equation follows from $k \le n$ and

$$\frac{1}{2\pi i}\int_{\gamma'} e^{-s}\sqrt{-s}\left(1 + O\left(|s|^{1/2}k^{-1/2}\right)\right)ds = -\frac{1}{2\sqrt{\pi}} + O\left(k^{-1/2}\right).$$

Now divide (24) by $t_n^{(r)}$, then an application of (2) and Lemma 4 completes the proof after all. □

In order to complete the proof of Theorem 2 we have to consider the other tree classes and forests as well. But since the proof totally relies on the structure of the singularity (Proposition 2) and the uniform estimates in Lemma 3, the other cases can easily be treated using the relations (4)–(6).

Acknowledgment. The author thanks two anonymous referees for pointing out some references and several imprecisions in the manuscript.

References

[1] C. K. Bailey. Distribution of points by degree and orbit size in a large random tree. *J. Graph Theory*, 6(3):283–293, 1982.

[2] Gerd Baron, Michael Drmota, and Ljuben Mutafchiev. Predecessors in random mappings. *Combin. Probab. Comput.*, 5(4):317–335, 1996.

[3] Edward A. Bender and L. Bruce Richmond. Central and local limit theorems applied to asymptotic enumeration. II. Multivariate generating functions. *J. Combin. Theory Ser. A*, 34(3):255–265, 1983.

[4] Robin Carr, William M. Y. Goh, and Eric Schmutz. The maximum degree in a random tree and related problems. In *Proceedings of the Fifth International Seminar on Random Graphs and Probabilistic Methods in Combinatorics and Computer Science (Poznań, 1991)*, volume 5, pages 13–24, 1994.

[5] Michael Drmota. Asymptotic distributions and a multivariate Darboux method in enumeration problems. *J. Combin. Theory Ser. A*, 67(2):169–184, 1994.

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

[7] Michael Drmota. Systems of functional equations. *Random Structures Algorithms*, 10(1-2):103–124, 1997. Average-case analysis of algorithms (Dagstuhl, 1995).

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

[9] Michael Drmota and Michèle Soria. Images and preimages in random mappings. *SIAM J. Discrete Math.*, 10(2):246–269, 1997.

[10] William Goh and Eric Schmutz. Limit distribution for the maximum degree of a random recursive tree. *J. Comput. Appl. Math.*, 142(1):61–82, 2002. Probabilistic methods in combinatorics and combinatorial optimization.

[11] William M. Y. Goh and Eric Schmutz. Unlabeled trees: distribution of the maximum degree. *Random Structures Algorithms*, 5(3):411–440, 1994.

[12] Valentin F. Kolchin. *Random mappings.* Translation Series in Mathematics and Engineering. Optimization Software Inc. Publications Division, New York, 1986. Translated from the Russian, With a foreword by S. R. S. Varadhan.

[13] Steven P. Lalley. Finite range random walk on free groups and homogeneous trees. *Ann. Probab.*, 21(4):2087–2130, 1993.

[14] A. Meir and J. W. Moon. On the maximum out-degree in random trees. *Australas. J. Combin.*, 2:147–156, 1990. Combinatorial mathematics and combinatorial computing, Vol. 2 (Brisbane, 1989).

[15] 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.

[16] A. Meir and J. W. Moon. A note on trees with concentrated maximum degrees. *Utilitas Math.*, 42:61–64, 1992.

[17] 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.

[18] A. Meir and J. W. Moon. Corrigendum to: "A note on trees with concentrated maximum degrees". *Utilitas Math.*, 43:253, 1993.

[19] J. W. Moon. Enumerating labelled trees. In *Graph Theory and Theoretical Physics*, pages 261–272. Academic Press, London, 1967.

[20] J. W. Moon. On the maximum degree in a random tree. *Michigan Math. J.*, 15:429–432, 1968.

[21] John W. Moon and Helmut Prodinger. A bijective proof of an identity concerning nodes of fixed degree in planted trees. *Ars Combin.*, 55:91–92, 2000.

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

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

[24] John Riordan. The enumeration of labeled trees by degrees. *Bull. Amer. Math. Soc.*, 72:110–112, 1966.

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

[26] Allen J. Schwenk. An asymptotic evaluation of the cycle index of a symmetric group. *Discrete Math.*, 18(1):71–78, 1977.

[27] A. Woods. Coloring rules for finite trees, and probabilities of monadic second order sentences. *Random Structures and Algorithms*, 10:453–485, 1997.