Enumerating alternating tree families

Markus Kuba, Alois Panholzer

FPSAC 2008, Valparaiso-Viña del Mar, Chile DMTCS proc. **AJ**, 2008, 105–116

Enumerating alternating tree families

Markus Kuba[1] and Alois Panholzer[1†]

[1] *Institute of Discrete Mathematics and Geometry, Vienna University of Technology, Wiedner Hauptstr. 8-10/104, A-1040 Wien, Austria.* {`Markus.Kuba,Alois.Panholzer`}`@tuwien.ac.at`

Abstract. We study two enumeration problems for *up-down alternating trees*, i.e., rooted labelled trees T, where the labels $v_1, v_2, v_3, \dots$ on every path starting at the root of T satisfy $v_1 < v_2 > v_3 < v_4 > \cdots$. First we consider various tree families of interest in combinatorics (such as unordered, ordered, d-ary and Motzkin trees) and study the number T_n of different up-down alternating labelled trees of size n. We obtain for all tree families considered an implicit characterization of the exponential generating function $T(z)$ leading to asymptotic results of the coefficients T_n for various tree families. Second we consider the particular family of up-down alternating labelled ordered trees and study the influence of such an alternating labelling to the average shape of the trees by analyzing the parameters *label of the root node*, *degree of the root node* and *depth of a random node* in a random tree of size n. This leads to exact enumeration results and limiting distribution results.

Résumé. Nous étudions deux problèmes de dénombrement d'*arbres alternés haut-bas* : par définition, ce sont des arbres munis d'une racine et tels que, pour tout chemin partant de la racine, les valeurs $v_1, v_2, v_3, \dots$ associées aux noeuds du chemin satisfont la chaîne d'inégalités $v_1 < v_2 > v_3 < v_4 > \cdots$. D'une part, nous considérons diverses familles d'arbres intéressantes du point de vue de l'analyse combinatoire (comme les arbres de Motzkin, les arbres non ordonnés, ordonnés et d-aires) et nous étudions pour chaque famille le nombre total T_n d'arbres alternés haut-bas de taille n. Nous obtenons pour toutes les familles d'arbres considérées une caractérisation implicite de la fonction génératrice exponentielle $T(z)$. Cette caractérisation nous renseigne sur le comportement asymptotique des coefficients T_n de plusieurs familles d'arbres. D'autre part, nous examinons le cas particulier de la famille des arbres ordonnés : nous étudions l'influence de l'étiquetage alterné haut-bas sur l'allure générale de ces arbres en analysant trois paramètres dans un arbre aléatoire (valeur de la racine, degré de la racine et profondeur d'un noeud aléatoire). Nous obtenons alors des résultats en terme de distribution limite, mais aussi de dénombrement exact.

Keywords: Alternating trees, Generating functions, Functional equations, Asymptotic enumeration results

1 Introduction

The family $\mathcal{T}$ of unrooted unordered alternating trees (also called intransitive trees) consists of all unrooted unordered labelled trees T, where the nodes of T with $|T| = n$ (the number $|T|$ of nodes of T will be called the size of T) are labelled by distinct integers of $\{1, 2, \dots, n\}$ in such a way that for every path $v_1, v_2, v_3, \dots$ in T it holds $v_1 < v_2 > v_3 < v_4 > \cdots$ or $v_1 > v_2 < v_3 > v_4 > \cdots$ (we always identify a node $v \in T$ with its label).

[†]This work has been supported by the Austrian Science Foundation FWF, grant S9608-N13.

This tree family appears in various contexts in combinatorics as in the enumeration of admissible bases of certain hypergeometric systems [4], in the enumeration of so called local binary search trees [7] and when enumerating the number of regions of certain hyperplane arrangements [8].

The enumeration problem for the number T_n of unrooted unordered alternating trees of size n has been solved by A. Postnikov in [7] by obtaining the following formula for T_n:

$$T_n = \frac{1}{n2^{n-1}} \sum_{k=1}^{n} \binom{n}{k} k^{n-1}.$$

The corresponding problem for rooted ordered alternating trees has been addressed and solved by C. Chauve, S. Dulucq and A. Rechnitzer in [1]. They considered trees T, where one node of T is distinguished as the root and where the subtrees of each node of T are linearly ordered, which are labelled by distinct integers of $\{1, 2, \dots, |T|\}$ in an "alternating way", i.e., in such a way that for every path $v_1, v_2, v_3, \dots$ in T it holds $v_1 < v_2 > v_3 < v_4 > \cdots$ or $v_1 > v_2 < v_3 > v_4 > \cdots$. The authors of [1] found that the number T_n of rooted ordered alternating trees of size $n \geq 2$ is given by the surprisingly simple formula $T_n = 2(n-1)^{n-1}$.

The aim of the present work is to address and give (up to some extent) solutions to the following two problems for alternating trees. First we consider the enumeration problem for other alternating labelled tree families, as, e.g., for binary trees, d-ary trees and Motzkin trees, where the corresponding families of unordered or arbitrary labelled trees appear frequently in combinatorics or computer science. We remark that all trees considered in this paper are rooted trees and we remark further that it is sufficient for the enumeration problem to count "up-down alternating labelled trees", i.e., it holds for every path $v_1, v_2, v_3, \dots$ starting at the root of a tree: $v_1 < v_2 > v_3 < v_4 > \cdots$. Of course, the number of all alternating labelled trees of size $n \geq 2$ of a rooted tree family is twice the number of up-down alternating labelled trees of size n. Our study relies on a description of the combinatorial decomposition of an up-down alternating tree T of a tree family considered with respect to the largest element $n = |T|$ in T. This decomposition leads to a recursive description of the enumeration problem and to quasilinear first order partial differential equations for suitably defined multivariate generating functions. For all tree families considered the differential equation appearing can be solved implicitly, which also leads to an implicit characterization of the exponential generating function $T(z)$ of the number T_n of trees of size n by means of certain functional equations. With few exceptions, amongst them the already known results for rooted ordered and rooted (or unrooted) unordered alternating trees, it does not seem that there are explicit formulæ for T_n available. However, the appearing functional equations for the generating functions of the number T_n of up-down alternating labelled trees are particularly useful to obtain asymptotic results of T_n for various tree families.

Second we are interested in the influence of an alternating labelling to the average structure or shape of the trees in a tree family compared to an arbitrary labelling. We do this by considering one particular tree family, namely the family of up-down alternating labelled ordered trees, and studying several tree parameters for random trees of size n (i.e., each of the T_n different trees of size n is chosen with equal probability). In particular we obtain limiting distribution results for the label of the root node, the degree of the root node, and the depth (i.e., the distance to the root) of a randomly chosen node in a random tree of size n. Interestingly one even obtains exact formulæ for the number of up-down alternating labelled ordered trees of size n, where the root is labelled by j, as well as for the number of up-down alternating labelled ordered trees of size n, where the root has degree m. To show these results we again use the basic decomposition of an up-down alternating tree T with respect to the largest element $n = |T|$, which again leads to certain partial differential equations for suitably introduced bivariate generating functions. A study of these differential equations leads then to exact or asymptotic results for the parameters studied.

2 Results

2.1 Enumeration results for up-down alternating labelled trees

We give here our results for the number T_n of up-down alternating labelled trees of size n for various tree families $\mathcal{T}$ as described in Subsection 3.1.

Theorem 1 *The exponential generating function $T(z) := \sum_{n\geq 1} T_n \frac{z^n}{n!}$ of the numbers T_n is for the tree families considered implicitly given as solution of the following functional equations:*

$$\text{Ordered trees:} \quad z = (1 - T(z)) \log \frac{1}{1 - T(z)}, \quad \text{or explicitly} \quad T(z) = 1 - e^{-W(z)},$$

$$\text{Unordered trees:} \quad z = \frac{2T(z)}{1 + e^{T(z)}}, \quad \text{or explicitly} \quad T(z) = \frac{z}{2} + W\big(\frac{ze^{\frac{z}{2}}}{2}\big),$$

$$d\text{-ary trees:} \quad z = \frac{2}{\big(1 + (1 + T(z))^{d+1}\big)^{\frac{d-1}{d+1}}} \int_0^{T(z)} \frac{dx}{\big(1 + (1 + x)^{d+1}\big)^{\frac{2}{d+1}}},$$

$$d\text{-bundled ordered trees:} \quad z = \frac{2}{\big(1 + \big(\frac{1}{1-T(z)}\big)^{d-1}\big)^{\frac{d+1}{d-1}}} \int_0^{T(z)} \Big(1 + \big(\frac{1}{1-x}\big)^{d-1}\Big)^{\frac{2}{d-1}} dx,$$

$$\text{Motzkin trees:} \quad z = \int_0^{T(z)} \frac{dx}{\frac{\big(4r(x)+4\sqrt{108+r^2(x)}\big)^{\frac{2}{3}}}{16} + \frac{9}{\big(4r(x)+4\sqrt{108+r^2(x)}\big)^{\frac{2}{3}}} - \frac{3}{4}},$$

$$\text{with} \quad r(x) = 24(T(z) - x) + 12(T^2(z) - x^2) + 8(T^3(z) - x^3) + 10,$$

$$\text{Strict binary trees:} \quad z = \int_0^{T(z)} \frac{dx}{\frac{\big(4r(x)+4\sqrt{4+r^2(x)}\big)^{\frac{2}{3}}}{4} + \frac{4}{\big(4r(x)+4\sqrt{4+r^2(x)}\big)^{\frac{2}{3}}} - 1},$$

$$\text{with} \quad r(x) = 3(T(z) - x) + (T^3(z) - x^3),$$

where the function $W(z) := \sum_{n\geq 1} n^{n-1} \frac{z^n}{n!}$ appearing is the so called tree function.

Theorem 2 *The numbers T_n are for each of the families of up-down alternating labelled ordered, unordered, d-ary and d-bundled ordered trees asymptotically given by*

$$T_n \sim C\rho^{-n} n^{-\frac{3}{2}} n!,$$

where ρ is the radius of convergence of the corresponding exponential generating function $T(z)$ and C is some computable constant, which may differ for every tree family considered. For these tree families the radius of convergence ρ is given as follows:

$$\text{Ordered trees:} \quad \rho = \frac{1}{e} \approx 0.367879\ldots,$$

$$\textit{Unordered trees:}\quad \rho = -2W(-e^{-1}) \approx 0.556929\ldots,$$

$$d\textit{-ary trees:}\quad \rho = \frac{2}{(d-1)(1+\tau)^d}, \quad \textit{with } \tau \textit{ the positive real solution of the equation}$$

$$\frac{\big(1+(1+\tau)^{d+1}\big)^{\frac{d-1}{d+1}}}{(d-1)(1+\tau)^d} = \int_0^\tau \frac{dx}{\big(1+(1+x)^{d+1}\big)^{\frac{2}{d+1}}},$$

$$d\textit{-bundled ordered trees:}\quad \rho = \frac{2(1-\tau)^d}{d+1}, \quad \textit{with } \tau \textit{ the positive real solution of the equation}$$

$$\frac{\big(1+\big(\frac{1}{1-\tau}\big)^{d-1}\big)^{\frac{d+1}{d-1}}}{(d+1)\big(\frac{1}{1-\tau}\big)^d} = \int_0^\tau \Big(1+\big(\frac{1}{1-x}\big)^{d-1}\Big)^{\frac{2}{d-1}} dx.$$

Furthermore the numbers T_n are for the families of ordered, unordered and 3-bundled ordered trees given by the following exact formulæ:

$$\textit{Ordered trees:}\quad T_n = (n-1)^{n-1},$$

$$\textit{Unordered trees:}\quad T_n = \frac{1}{2^n}\sum_{k=1}^{n}\binom{n}{k}k^{n-1},$$

$$3\textit{-bundled ordered trees:}\quad T_n = \frac{(n-1)!}{2^{n+1}}\sum_{k=0}^{2n}\binom{2n}{k}\binom{\frac{5n-3}{2}-k}{n-1}.$$

We remark that the exact formulæ of T_n for the families of ordered trees and unordered trees already appear (or are easily deduced from results) in [1] and [7], respectively.

2.2 Results for tree parameters in up-down alternating labelled ordered trees

We give here our exact and asymptotic results of parameters described in Section 4 for the family of up-down alternating labelled ordered trees.

Theorem 3 *Let $T_{n,j}$ denote the number of up-down alternating labelled ordered trees of size n, where the root node has label j, with $1 \le j \le n$, and let L_n be the random variable, which gives the label of the root node of a randomly chosen up-down alternating labelled ordered tree of size-n. Then $T_{n,j}$ is given by the following exact formula:*

$$T_{n,j} = (n-j)(n-1)^{j-2}n^{n-j-1},$$

and the normalized random variable $\frac{L_n}{n}$ converges for $n \to \infty$ in distribution to a random variable L, i.e., $\frac{L_n}{n} \xrightarrow{(d)} L$, with density function $f(x) = (1-x)e^{1-x}$, for $0 \le x \le 1$.

Theorem 4 *Let $T_{n,m}$ denote the number of up-down alternating labelled ordered trees of size n, where the root node has degree m, with $0 \le m \le n$, and let R_n be the random variable, which gives the degree of the root node of a randomly chosen up-down alternating labelled ordered tree of size-n. Then $T_{n,m}$ is given by the following exact formula:*

$$T_{n,m} = H_m(n-1)^{n-1} + \sum_{\ell=1}^{m}\binom{m}{\ell}(-1)^\ell\frac{\ell+1}{\ell}(n-1-\ell)^{n-1},$$

and the random variable R_n converges for $n \to \infty$ in distribution to a discrete random variable R, i.e., $R_n \xrightarrow{(d)} R$, whose distribution is given by

$$\mathbb{P}\{R = m\} = \big(\frac{e-1}{e}\big)^m - 1 + \sum_{\ell=1}^{m} \frac{\big(\frac{e-1}{e}\big)^\ell}{\ell}, \quad \textit{for} \;\; m \in \mathbb{N}.$$

Here $H_m := \sum_{k=1}^{m} \frac{1}{k}$ denotes the m-th harmonic number.

Theorem 5 *Let D_n be the random variable, which counts the depth (i.e., the distance to the root) of a randomly chosen node in a random up-down alternating labelled ordered tree of size n. Then the normalized random variable $\frac{D_n}{\sqrt{n}}$ converges for $n \to \infty$ in distribution to a Rayley-distributed random variable R_α, i.e., $\frac{D_n}{\sqrt{n}} \xrightarrow{(d)} R_\alpha$, with parameter $\alpha = \frac{2}{3}$, where R_α has density function $f(x) = \frac{x}{\alpha^2} e^{-\frac{x^2}{2\alpha^2}}$, for $x \geq 0$.*

3 Enumeration of up-down alternating trees

3.1 Tree families considered

In the following we describe the combinatorial families $\mathcal{T}$ of trees that we consider here. Basically all trees contained in the tree families are up-down alternating labelled rooted trees. This means that we only consider labelled trees: the nodes in a tree T of size $|T| = n$ are labelled by distinct integers of $\{1, 2, \dots, n\}$, where one of the n nodes of T is distinguished as the root node. Furthermore the labelling of any tree T is an "up-down alternating labelling", i.e., it must hold for any sequence of nodes $v_1, v_2, v_3, \dots$ lying on the path from the root to an arbitrary node in T that $v_1 < v_2 > v_3 < v_4 > \cdots$, where we identify a node with its label.

To specify the up-down alternating labelled tree families $\mathcal{T}$ we are dealing with, we describe the corresponding arbitrary labelled tree families $\tilde{\mathcal{T}}$ and define that $\mathcal{T} \subseteq \tilde{\mathcal{T}}$ contains exactly those trees $T \in \tilde{\mathcal{T}}$ with an up-down alternating labelling.

Labelled ordered trees: Every tree $T \in \tilde{\mathcal{T}}$ of size n consists of a root node, where a (possibly empty) sequence of labelled ordered trees is attached and where the whole tree is relabelled with the labels $\{1, 2, \dots, n\}$ in an order preserving way.

Labelled unordered trees: Every tree $T \in \tilde{\mathcal{T}}$ of size n consists of a root node, where a (possibly empty) set of labelled unordered trees is attached and where the whole tree is relabelled with the labels $\{1, 2, \dots, n\}$ in an order preserving way.

Labelled d-ary trees: Every tree $T \in \tilde{\mathcal{T}}$ of size n consists of a root node, which has d positions, where either a labelled d-ary tree is attached or not and where the whole tree is relabelled with the labels $\{1, 2, \dots, n\}$ in an order preserving way.

Labelled d-bundled trees: Every tree $T \in \tilde{\mathcal{T}}$ of size n consists of a root node, which has d positions, where a (possibly empty) sequence of labelled d-bundled trees is attached and where the whole tree is relabelled with the labels $\{1, 2, \dots, n\}$ in an order preserving way. Alternatively one might think of a d-bundled tree as an ordered tree, where the sequence of subtrees attached to any node in the tree is separated by $d - 1$ bars into d bundles.

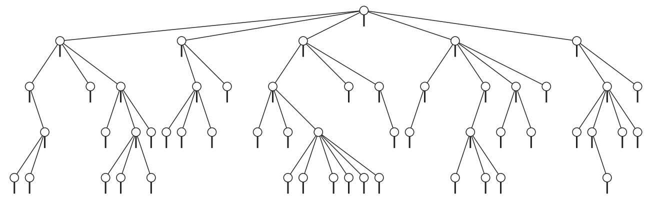

Fig. 1: An example of a 2-bundled tree. A bar separates the subtrees into left and right ones.

Labelled Motzkin trees: Every tree $T \in \tilde{\mathcal{T}}$ of size n consists of a root node, where a sequence of 0, 1 or 2 labelled Motzkin trees is attached and where the whole tree is relabelled with the labels $\{1, 2, \dots, n\}$ in an order preserving way.

Labelled strict binary trees: Every tree $T \in \tilde{\mathcal{T}}$ of size n consists of a root node, where a sequence of 0 or 2 labelled strict binary trees is attached and where the whole tree is relabelled with the labels $\{1, 2, \dots, n\}$ in an order preserving way.

Whereas the families of ordered, unordered, d-ary, strict binary and Motzkin-trees are well-known tree families with a lot of applications (see, e.g., [2, 9]), we remark that d-bundled trees appear, e.g., in the context of certain "preferential attachment" growth models for trees (see [5]) and furthermore, that they are satisfying some randomness preservation properties when studying cutting-down procedures for random trees (see [6]). An example of a 2-bundled tree is given in Fig. 1.

3.2 Combinatorial decompositions

Fundamental to our approach is the description of the decomposition of an up-down alternating tree T of size $|T| = n$ in a tree family $\mathcal{T}$ with respect to node n. If we cut-off all edges incident with node n and relabel the resulting trees with labels from 1 up to their sizes in an order-preserving way we obtain for $n \geq 2$ an alternating tree $\widehat{T}$ of size $k \geq 1$ that contains the original root of the tree T and alternating trees $T_1, T_2, \dots, T_r$ of sizes $k_1, \dots, k_r \geq 1$, which correspond to the subtrees originally attached to node n. Of course, it holds that $k + k_1 + \cdots + k_r = n - 1$.

If $\mathcal{T}$ is one of the families of up-down alternating labelled unordered, ordered, d-ary, d-bundled ordered or Motzkin trees it follows that all the resulting trees $\widehat{T}, T_1, \dots, T_r$ are again alternating trees of the family $\mathcal{T}$. This is not true if $\mathcal{T}$ is the family of up-down alternating labelled strict binary trees: although $T_1, \dots, T_r$ are alternating trees of $\mathcal{T}$ this does not hold for the tree $\widehat{T}$, since there exists now a node in $\widehat{T}$, where only one subtree is attached. We first consider the decomposition for those tree families $\mathcal{T}$, where all the resulting trees are again contained in $\mathcal{T}$ and discuss the decomposition for the family of strict binary trees later.

In order to use the decomposition of an up-down alternating labelled tree of a family $\mathcal{T}$ to get a recursive description of the numbers T_n of different size-n trees of $\mathcal{T}$ we are now interested in an answer to the following question. What is the number of different trees $T \in \mathcal{T}$ of size $|T| = n$ that we can obtain by starting with alternating trees $\widehat{T}, T_1, \dots, T_r \in \mathcal{T}$ of corresponding sizes $k, k_1, \dots, k_r$, with $k + k_1 + \cdots + k_r = n - 1$, distributing the labels $\{1, 2, \dots, n-1\}$ amongst the nodes of $\widehat{T}, T_1, \dots, T_r$ and relabelling all these trees in an order-preserving way and afterwards attaching $T_1, \dots, T_r$ to a new vertex labelled by n and attaching node n to a node in $\widehat{T}$?

Obviously for all tree families we have a contribution of $\binom{n-1}{k, k_1, \dots, k_r}$ stemming from the distribution of the labels to the trees. The more interesting contribution is the factor coming from the number of possible positions, where node n can be attached to a node in $\widehat{T}$ of size $|\widehat{T}| = k$, such that the up-down alternating labelling is preserved for the resulting tree. This contribution, which will be denoted here by w, depends on the specific tree family considered and will be studied now. We require there the notion of the depth $h(v)$ of a node v in a tree $\widehat{T}$, which is given by the distance of v to the root of $\widehat{T}$, i.e., the number of edges lying on the unique path from the root of $\widehat{T}$ to node v. It holds then that node n can only be attached to nodes $v \in \widehat{T}$, which are at an even level, i.e., where it holds $h(v) \equiv 0 \mod 2$. The set of nodes of $\widehat{T}$ at an even level will be denoted by $V := \{v \in \widehat{T} : h(v) \equiv 0 \mod 2\}$ and its cardinality by $\ell := |V|$. Furthermore we denote by $d^+(v)$ the out-degree (the number of children) of a node v in a tree $\widehat{T}$.

Ordered trees: The number of positions w of attaching node n to one of the nodes of V is given by

$$w = \sum_{v \in V} (d^+(v) + 1) = \sum_{v \in V} 1 + \sum_{v \in V} d^+(v) = |V| + |\widehat{T} \setminus V| = |\widehat{T}| = k,$$

since $\sum_{v \in V} d^+(v)$ gives exactly the number of nodes in $\widehat{T}$ at an odd level. Thus we obtain that, independent of the specific tree $\widehat{T}$, there are always $|\widehat{T}| = k$ positions of attaching node n, such that the resulting tree is again up-down alternating labelled.

Unordered trees: The number of positions w of attaching node n to one of the nodes of V is now simply given by $w = |V| = \ell$.

d**-ary trees:** We obtain now for w:

$$w = \sum_{v \in V} (d - d^+(v)) = d \sum_{v \in V} 1 - \sum_{v \in V} d^+(v) = d|V| - |\widehat{T} \setminus V| = d\ell - (k - \ell) = (d+1)\ell - k.$$

d**-bundled ordered trees:** Now w is given as follows:

$$w = \sum_{v \in V} (d^+(v) + d) = d \sum_{v \in V} 1 + \sum_{v \in V} d^+(v) = d|V| + |\widehat{T} \setminus V| = d\ell + (k - \ell) = (d-1)\ell + k.$$

Motzkin trees: To count the number of positions w of attaching node n to one of the nodes of V we define the set of nodes in V with out-degree 0 and 1, respectively, by $V^{[0]} := \{v \in V : d^+(v) = 0\}$, $V^{[1]} := \{v \in V : d^+(v) = 1\}$, and use the notation $\ell^{[0]} := |V^{[0]}|$, $\ell^{[1]} := |V^{[1]}|$ for their cardinalities. We obtain then:

$$w = \sum_{v \in V^{[0]}} 1 + \sum_{v \in V^{[1]}} 2 = |V^{[0]}| + 2|V^{[1]}| = \ell^{[0]} + 2\ell^{[1]}.$$

Thus this combinatorial decomposition only leads for the family of ordered trees directly to a recurrence for the number T_n of alternating labelled trees of size n, whereas we have to store additional information for the other tree families considered: for unordered trees, d-ary trees and d-bundled ordered trees we will introduce the number $T_{n,m}$ of alternating labelled trees of size n, where exactly m nodes are at an even level, and for Motzkin trees we will introduce the number $T_{n,m^{[0]},m^{[1]}}$ of alternating labelled trees, where exactly $m^{[0]}$ nodes at an even level have out-degree 0 and $m^{[1]}$ nodes at an even level have out-degree 1. Of course, it holds $T_n = \sum_{m \geq 0} T_{n,m}$ and $T_n = \sum_{m^{[0]},m^{[1]} \geq 0} T_{n,m^{[0]},m^{[1]}}$, respectively.

Strict binary trees: As mentioned above when applying this decomposition to an up-down alternating tree T of the family $\mathcal{T}$ of strict binary trees the resulting tree $\widehat{T}$ containing the original root of T is no more a strict binary tree. To treat the family of strict binary trees with the same approach as before we consider a larger tree family $\mathcal{S} \supseteq \mathcal{T}$ containing $\mathcal{T}$. The family $\mathcal{S}$ consists now of all up-down alternating labelled rooted trees T, where every node $v \in T$ at an odd level ($h(v) \equiv 1 \mod 2$) has a sequence of 0 or 2 children and where every node $v \in T$ at an even level ($h(v) \equiv 0 \mod 2$) has a sequence of 0, 1 or 2 children. Of course, the family $\mathcal{T}$ of alternating labelled strict binary trees contains exactly those trees of $\mathcal{S}$, which do not contain any nodes of out-degree 1. As immediately seen the basic decomposition of an up-down alternating labelled tree $T \in \mathcal{S}$ of size $|T| = n$ with respect to n leads to up-down alternating labelled trees $\widehat{T}, T_1, \dots, T_r$, which are all elements of $\mathcal{S}$. Thus we can repeat the considerations made above for the family $\mathcal{S}$. It remains to study the number w of possible positions, where node n can be attached to a node in $\widehat{T}$ of size $|\widehat{T}| = k$, such that the up-down alternating labelling is preserved for the resulting tree. We denote as above the set of nodes in $\widehat{T}$ at an even level with V and define the set of nodes in V with out-degree 0 and 1, respectively, by $V^{[0]} := \{v \in V : d^+(v) = 0\}$, $V^{[1]} := \{v \in V : d^+(v) = 1\}$, and use the notation $\ell^{[0]} := |V^{[0]}|$, $\ell^{[1]} := |V^{[1]}|$ for their cardinalities. We obtain then:

$$w = \sum_{v \in V^{[0]}} 2 + \sum_{v \in V^{[1]}} 1 = 2|V^{[0]}| + |V^{[1]}| = 2\ell^{[0]} + \ell^{[1]}.$$

Thus in order to treat the enumeration problem for strict binary trees we will introduce the number $T_{n,m^{[0]},m^{[1]}}$ of alternating labelled trees of the family $\mathcal{S}$ defined above, where exactly $m^{[0]}$ nodes at an even level have out-degree 0 and $m^{[1]}$ nodes at an even level have out-degree 1. The basic combinatorial decomposition leads then to a recursive description of these numbers $T_{n,m^{[0]},m^{[1]}}$. Of course, we are interested in particular in the number T_n of up-down alternating labelled strict binary trees, which are given by $T_n = \sum_{m^{[0]} \ge 0} T_{n,m^{[0]},0}$.

General tree families: We remark that one could use this basic decomposition to obtain a recursive description of the number of up-down alternating labelled trees for any family $\mathcal{T}$ of so called simply generated trees (see, e.g., [3] for a definition and results), where the out-degree of a node $v \in T$ is bounded a priori for all trees $T \in \mathcal{T}$ by some fixed bound d.

As we have seen for the families of up-down alternating labelled unordered, ordered and d-bundled ordered trees the approach also works for some instances of tree families, where the degree of a node $v \in T$ is not bounded by some universal constant for all $T \in \mathcal{T}$; however, in general one would be forced to store then the whole sequence $(m^{[0]}, m^{[1]}, m^{[2]}, \dots)$ of the numbers $m^{[i]}$ of nodes v in a tree T at an even level with out-degree $d^+(v) = i$.

3.3 Generating functions

The basic decomposition of an up-down alternating labelled tree T of size $|T| = n$ with respect to node n described in Subsection 3.2 immediately leads to recurrences for the numbers T_n (ordered trees), $T_{n,m}$ (unordered trees, d-ary trees and d-bundled ordered trees) or $T_{n,m^{[0]},m^{[1]}}$ (Motzkin trees and strict binary trees) introduced there. We will not give these recurrences here, since they are obtained directly from the description given in Subsection 3.2. We treat the recurrences appearing by introducing suitable generating functions: $T(z) := \sum_{n\ge1} T_n \frac{z^n}{n!}$, $F(z,u) := \sum_{n\ge1}\sum_{m\ge0} T_{n,m}\frac{z^n}{n!}u^m$, and $F(z,u_0,u_1) := \sum_{n\ge1}\sum_{m^{[0]}\ge0}\sum_{m^{[1]}\ge0} T_{n,m^{[0]},m^{[1]}}\frac{z^n}{n!}u_0^{m^{[0]}}u_1^{m^{[1]}}$.

This leads, apart from the instance of ordered trees, where a nonlinear ordinary differential equation occurs, to first order quasilinear partial differential equations for the generating functions introduced.

These equations are given below, where we use the abbreviation $F := F(z, u_0, u_1)$; additionally the initial conditions $T(0) = F(0, u) = F(0, u_0, u_1) = 0$ hold:

$$\begin{aligned}
\text{Ordered trees:} \quad & T'(z) - 1 = \frac{zT'(z)}{1 - T(z)}, \\
\text{Unordered trees:} \quad & F_z(z, u) - u = ue^{F(z,u)} F_u(z, u), \\
d\text{-ary trees:} \quad & F_z(z, u) - u = (1 + F(z, u))^d \big((d + 1)uF_u(z, u) - zF_z(z, u)\big), \\
d\text{-bundled ordered trees:} \quad & F_z(z, u) - u = \frac{1}{(1 - F(z, u))^d} \big((d - 1)uF_u(z, u) + zF_z(z, u)\big), \\
\text{Motzkin trees:} \quad & F_z - u_0 = \big(1 + F + F^2\big)\big(u_1 F_{u_0} + 2F_{u_1}\big), \\
\text{Strict binary trees:} \quad & F_z - u_0 = \big(1 + F^2\big)\big(2u_1 F_{u_0} + F_{u_1}\big).
\end{aligned}$$

All the differential equations appearing can be solved by using the method of characteristics for first order quasilinear partial differential equations, see, e.g., [10] for a description of this method. The exponential generating functions $T(z) = \sum_{n\ge1} T_n \frac{z^n}{n!}$ of the number T_n of up-down alternating trees of size n that we are mainly interested in, can be obtained from these solutions via $T(z) = F(z, 1)$ (for unordered trees, d-ary trees and d-bundled ordered trees), $T(z) = F(z, 1, 1)$ (for Motzkin trees) and $T(z) = F(z, 1, 0)$ for strict binary trees. We obtain then that the functions $T(z)$ are given implicitly as solutions of certain functional equations, which are collected in Theorem 1.

For the instance of ordered and unordered trees the functions $T(z)$ can be expressed via the so called tree function and exact formulæ for the numbers T_n can be obtained by extracting coefficients. This leads to results obtained by [1] and [7]. For most of the other tree families considered it does not seem that there are explicit formulæ for the numbers T_n available. We only remark the somewhat curious fact that for the instance of 3-bundled ordered trees the equation for $T(z)$ can be simplified to $z = \frac{2T(z)(1-T(z))^3(2-T(z))}{(1+(1-T(z))^2)^2}$, which also leads to an exact formula for T_n. These exact results are collected in Theorem 2.

3.4 Asymptotic enumeration results

An advantage of the implicit characterization of the generating functions $T(z) = \sum_{n\ge1} T_n \frac{z^n}{n!}$ as the solution of certain functional equations (as obtained in Subsection 3.3) is that this often allows to apply analytic techniques leading to asymptotic results for the coefficients T_n, for $n \to \infty$.

We will not show here details, but refer to the very general treatment discussed in [3]. Basically one has to determine the radius of convergence ρ of the analytic function $T(z)$, which already leads to some growth estimates of T_n. However, for a detailed description of the growth of the coefficients T_n one has to locate all singularities on the radius of convergence (the so called dominant singularities) and describe the behaviour of $T(z)$ locally in a complex neighbourhood of their dominant singularities. By applying transfer lemmata, i.e., singularity analysis, this leads then in many instances to precise asymptotic results for T_n.

For the families $\mathcal{T}$ of up-down alternating labelled ordered, unordered, d-ary and d-bundled ordered trees one can determine the radius ρ of convergence by an application of the implicit function theorem and one can also show that $z = \rho$ is the unique dominant singularity. Furthermore one is able to determine the local behaviour of $T(z)$ around $z = \rho$ and to apply singularity analysis. The asymptotic results for T_n obtained for these tree families are given in Theorem 3.

For the families $\mathcal{T}$ of up-down alternating labelled Motzkin trees and strict binary trees we also determined the radius ρ of convergence, but up to now we could not show that $z = \rho$ is the unique dominant singularity (in the instance of Motzkin trees) or that there are exactly two dominant singularities $z = \rho$ and $z = -\rho$ (in the instance of strict binary trees), respectively. Thus we defer a more detailed analysis of the asymptotic behaviour of T_n for these tree families to future research.

4 Parameters in up-down alternating labelled rooted trees

We study now certain parameters for the family $\mathcal{T}$ of up-down alternating labelled ordered trees. We recall that the number of size-n trees in $\mathcal{T}$ is given by $T_n = (n-1)^{n-1}$ and that its exponential generating function $T(z) = \sum_{n\geq 1} T_n \frac{z^n}{n!}$ is given by $T(z) = 1 - e^{-W(z)}$, where $W(z) = \sum_{n\geq 1} n^{n-1}\frac{z^n}{n!}$ is the tree function.

4.1 Label of the root node

First we want to count the number $T_{n,j}$ of up-down alternating labelled ordered trees of size n, where the root node has label j, with $1 \leq j \leq n$. In a probabilistic setting we introduce the random variable L_n, where $\mathbb{P}\{L_n = j\} = \frac{T_{n,j}}{T_n}$ gives the probability that the root node of a random size-n alternating tree has label j. By using the basic decompostion of an alternating tree of size n with respect to the largest element n and counting the number of ways, where the root in the tree labelled by j will get, after an order preserving relabelling with elements $\{1, 2, \dots, k\}$, the label ℓ in the subtree of size k containing the original root, one obtains a recurrence for the probabilities $\mathbb{P}\{L_n = j\}$, $1 \leq j \leq n$.

This recurrence will be treated by introducing the bivariate generating function $F(z,u) := \sum_{n\geq 1}\sum_{1\leq j\leq n} T_n \mathbb{P}\{L_n = j\}\frac{z^{j-1}}{(j-1)!}\frac{u^{n-j}}{(n-j)!}$, which leads to the following first order linear partial differential equation with initial condition $F(z,0) = 1$:

$$\big(1 - \frac{u}{1-T(z+u)}\big)F_u(z,u) - \frac{z}{1-T(z+u)}F_z(z,u) - \frac{1}{1-T(z+u)}F(z,u) = 0.$$

By an application of the method of characteristics we obtain that the solution of this differential equation is given by

$$F(z,u) = e^{ue^{W(z+u)}}.$$

Extracting coefficients by applying the Lagrange inversion formula leads then to the exact results for $T_{n,j}$ stated in Theorem 3. The limiting distribution result for the probabilities $\mathbb{P}\{L_n = j\}$ as given also in Theorem 3 easily follows from this exact result.

One also easily obtains an exact expression for the expected value $\mathbb{E}(L_n)$: $\mathbb{E}(L_n) = \sum_{j=1}^{n} j\mathbb{P}\{L_n = j\} = 3n-1-\frac{n^n}{(n-1)^{n-1}}$, which gives $\mathbb{E}(L_n) \sim (3-e)n = (0.281718\dots)n$. Thus the result matches with the intuition that smaller labels are preferred to become the label of the root node in alternating ordered trees, but the exact amount of this preference is covered in the findings above.

4.2 Root degree

Next we are interested in the behaviour of the root degree in up-down alternating labelled ordered trees. To do this we introduce the random variable R_n, which counts the root degree of a randomly chosen up-down alternating ordered tree of size n. If we denote now by $T_{n,m}$ the number of up-down alternating labelled ordered trees of size n, where the root node has degree m, with $0 \leq m \leq n$, the probabilities $\mathbb{P}\{R_n = m\}$ are just given by $\mathbb{P}\{R_n = m\} = \frac{T_{n,m}}{T_n}$. The numbers $T_{n,m}$ satisfy a recurrence, which can be obtained from the basic decomposition of an alternating tree of size n with respect to the largest element n and using the fact that by cutting off node n the subtree of size k containing the root has either also degree m (this is obtained for $k-m-1$ positions of n) or has degree $m-1$ (this is obtained for m positions of n).

This recurrence can be treated by introducing the bivariate generating function $F(z,v) := \sum_{n\geq 1}\sum_{m\geq 0} T_{n,m}\frac{z^n}{n!}v^m$. This leads then to the following first order linear partial differential equation with initial condition $F(0,v) = 0$:

$$\big(1 - \frac{z}{1-T(z)}\big)F_z(z,v) + \frac{v(1-v)}{1-T(z)}F_v(z,v) = 1 - \frac{1-v}{1-T(z)}F(z,v).$$

The solution of this differential equation, which can be obtained again by applying the method of characteristics, is given as follows:

$$F(z,v) = \frac{z}{(1-v)W(z)}\log\Big(1 - (1-v)\big(1 - \frac{W(z)}{z}\big)\Big).$$

Extracting coefficients by using the Lagrange inversion formula leads to the exact formula for $T_{n,m}$ given in Theorem 4. Of course, the exact distribution of R_n is then determined by $\mathbb{P}\{R_n = m\} = \frac{T_{n,m}}{(n-1)^{n-1}}$. The discrete limiting distribution result for R_n as given in Theorem 5 can be obtained either from these exact results or easier (due to the alternating sum involved in this exact expression) by applying singularity analysis to the functions $[v^m]F(z,v)$.

We further remark that the expected value of R_n is given by the following exact formula: $\mathbb{E}(R_n) = \frac{1}{2}\Big[\big(\frac{n+1}{n-1}\big)^{n-1} - 1\Big]$, which gives that $\mathbb{E}(R_n) \sim \frac{e^2-1}{2} \approx 3.194528\dots$. If we compare this result with the corresponding result for unlabelled (or equivalently randomly labelled) ordered trees, where it holds $\mathbb{E}(R_n) \sim 3$, we obtain that on average the root of an alternating tree has a slightly higher degree than the root of a randomly labelled tree.

4.3 Depth of nodes

An important parameter when analysing the structure of random trees in rooted tree families is the depth of a randomly chosen node. Thus we are studying the random variable D_n, which counts the depth of a randomly chosen node in a random up-down alternating labelled ordered tree of size n. Due to the nature of the basic decomposition of alternating trees after the node with label n in a size-n tree, we require for a study of D_n an auxiliary parameter, which we call "the depth of a random insertion point": if we choose an alternating ordered tree T of size n there are exactly n possibilities to attach a node with label $n+1$ to one of the nodes in T in such a way that the resulting tree is again an alternating ordered tree, now of size $n+1$. The depth of node $n+1$ when attached to a node in T in an appropriate way is then of interest here. We introduce thus the auxiliary random variable X_n, which counts the depth of node $n+1$ in a tree obtained by choosing a random alternating tree of size n and attaching a node with label $n+1$ at random at one of the n possibilities such that the resulting tree is an alternating tree of size $n+1$.

Using the basic decomposition of an up-down alternating labelled ordered tree with respect to the node with the largest label we obtain a system of recurrences for the probabilities $\mathbb{P}\{X_n = m\}$ and $\mathbb{P}\{D_n = m\}$. When introducing the bivariate generating functions

$F(z,v) := \sum_{n\geq 1}\sum_{m\geq 0} nT_n \mathbb{P}\{X_n = m\}\frac{z^n}{n!}v^m$, and $G(z,v) := \sum_{n\geq 1}\sum_{m\geq 0} nT_n \mathbb{P}\{D_n = m\}\frac{z^n}{n!}v^m$, we get the following system of differential equations with initial conditions $F(0,v) = G(0,v) = 0$:

$$\big(1 - \frac{z}{1-T(z)}\big)F_z(z,v) - \frac{1}{1-T(z)}F(z,v) - \frac{v}{(1-T(z))^2}F(z,v)^2 - v = 0,$$
$$\big(1 - \frac{z}{1-T(z)}\big)G_z(z,v) - \frac{1}{1-T(z)}F(z,v) - \frac{vF(z,v)}{(1-T(z))^2}G(z,v) - 1 = 0.$$

In order to study the asymptotic behaviour of the depth D_n of a random node in a random alternating ordered tree of size n we use the so called method of moments. By studying the evaluation at $v = 1$ of the r-derivatives of $F(z,v)$ and $G(z,v)$ w.r.t. v we are able to show that, for r fixed and $n \to \infty$, the r-th moment of the normalized random variable $\frac{D_n}{\sqrt{n}}$ converges to the r-th moment of a Rayley-distributed random variable R_α, with parameter $\alpha = \frac{2}{3}$. Since the Rayleigh-distribution is fully characterized by its moments an application of the Theorem of Fréchet and Shohat shows the convergence in distribution of $\frac{D_n}{\sqrt{n}}$ to R_α which is stated as Theorem 5.

If one compares this result with the depth D_n of a random node in a randomly labelled ordered tree of size n, where $\frac{D_n}{\sqrt{n}}$ is also asymptotically Rayleigh distributed, but with a larger parameter $\alpha = 1$, one gets that on average the depth of a randomly chosen node in a randomly chosen alternating ordered tree is about $1/3$ smaller than the depth of a randomly chosen node in a random labelled tree of the same size.

References

[1] C. Chauve, S. Dulucq and A. Rechnitzer, Enumerating alternating trees, *Journal of Combinatorial Theory, Series A* 94, 142–151, 2001.

[2] P. Flajolet and R. Sedgewick, *An introduction to the analyis of algorithms*, Addison-Wesley, Reading, 1996.

[3] P. Flajolet and R. Sedgewick, *Analytic combinatorics*, to appear.
Draft available at `http://algo.inria.fr/flajolet/Publications/books.html`

[4] I. M. Gelfand, M. I. Graev and A. Postnikov, Combinatorics of hypergeometric functions associated with positive roots, in: *"Arnold-Gelfand mathematical seminars: Geometry and Singularity Theory"*, 205–221, Birkhäuser, Basel, 1997.

[5] M. Kuba and A. Panholzer, On the degree distribution of the nodes in increasing trees, *Journal of Combinatorial Theory, Series A* 114, 597–618, 2007.

[6] A. Panholzer, Cutting down very simple trees, *Quaestiones Mathematicae* 29, 211–227, 2006.

[7] A. Postnikov, Intransitive trees, *Journal of Combinatorial Theory, Series A* 79, 360–366, 1997.

[8] A. Postnikov and R. Stanley, Deformations of Coxeter hyperplane arrangements, *Journal of Combinatorial Theory, Series A* 91, 544–597, 2000.

[9] R. Stanley, *Enumerative Combinatorics, Vol. 2*, Cambridge University Press, Cambridge, 1999.

[10] M. E. Taylor, *Partial differential equations. Basic theory. Texts in Applied Mathematics, 23*, Springer, New York, 1996.