Discrete Applied Mathematics 18 (1987) 9–24
North-Holland

MONOTONICALLY LABELLED MOTZKIN TREES

Johann BLIEBERGER
Siemens AG Oesterreich, PSE 2B1, Siemensstr. 88–92, A-1211 Wien, Postfach 83, Austria

Received 3 March 1986
Revised 18 November 1986

Consider a rooted tree structure the nodes of which have been labelled monotonically by elements of $\{1, 2, \ldots, k\}$, which means that any sequence connecting the root of the tree with a leaf is weakly monotone.

For fixed k asymptotic equivalents of the form $C_k q_k^{-n} n^{-3/2}$ ($n \to \infty$) to the numbers of such tree structures with n nodes are obtained for the family of extended unary-binary trees (i.e., Motzkin trees) and for the family of extended unary-t-ary trees. Furthermore the numbers of (not extended) monotonically labelled binary and unary-binary trees are studied.

For each of these families the asymptotic behaviour of q_k as $k \to \infty$ is determined. This is done by investigating a non-linear function sequence. The roots of the functions of this function sequence equal q_k. Thus one finds for instance $q_k \sim (\log 2)/k$ ($k \to \infty$) for the family of extended unary-binary trees, and $q_k \sim \pi/2k$ ($k \to \infty$) for the family of binary trees.

1. Introduction

A large number of recent papers deal with the investigation of generalized classes of tree structures. Compare e.g. [6], [7], [8], [9], [11]. Consider the nodes of a tree labelled by elements of $\{1, 2, \ldots, k\}$ in such a way that any sequence connecting the root of the tree with a leaf is weakly monotone.

These tree structures are of special interest e.g. in connection with some kind of *order preserving maps* (cf. [11]) or in the *enumeration of expression trees.*

Let us study the latter case in some detail:

It is well-known that an arithmetic expression can be transformed to a corresponding expression tree. The connection between formulas and trees is very important in *computer science.* It occurs in a number of contexts in compiling, symbolic manipulation systems, and related areas.

In most cases only the mapping of formulas to expression trees is of interest. Sometimes, however, it is necessary to construct an *algorithm* that given an expression tree produces a corresponding expression. Since in this case usually many different expressions can be constructed from the same tree, it is convenient to reduce the number of parentheses as much as possible. This is done by providing the operators with distinct priorities.

We call expressions that do not involve parentheses or that do involve parentheses only because some operators are not associative, *canonical expressions.* (For instance $a*b+c*d$ or $a''*b''+c*d'''+e$ would be such expressions, while $a*(b+c)$ or $(a+b)''$ would not be.)

Given a fixed set of operators (with priorities) and an algorithm like the one mentioned above, the question arises how often the algorithm will produce canonical expressions.

If we consider the priorities to be labels of an expression tree, canonical expressions can only occur if the labeling is weakly monotone. Hence the question can be answered by counting monotonically labelled trees.

In [11] Prodinger and Urbanek have considered the problem of finding *asymptotic equivalents* to the numbers of monotonically labelled tree structures with n nodes in the case of some special families such as extended binary trees, extended t-ary trees, and ordered trees. For fixed k they obtained asymptotic equivalents of the form $C_k q_k^{-n} n^{-3/2} (n \to \infty)$ to the numbers of these families of trees. Prodinger and Urbanek additionally showed that the sequence q_k obeys a simple nonlinear recurrence relation.

In the present paper we want to investigate the number of *monotonically labelled extended Motzkin trees* (i.e., unary-binary trees) and the number of *monotonically labelled* (*not extended*) *binary trees.* The methods developed in this paper easily generalize to families such as *extended unary-t-ary trees* and (*not extended*) *unary-binary trees* with weights attached to their nodes. In all these cases asymptotic equivalents of the form $C_k q_k^{-n} n^{-3/2}$ $(n \to \infty)$ to the numbers of these families of trees are obtained. The essential difference to the paper of Prodinger and Urbanek is that the sequence q_k does not obey a simple recurrence relation, but appears as the roots of the functions of a certain function sequence which satisfies a nonlinear recurrence relation.

A detailed investigation of this function sequence allows to establish the asymptotic behaviour of q_k as $k \to \infty$. Thus the following results are obtained for

Extended Motzkin Trees. *The number $M_{k,n}$ of Motzkin trees with n internal nodes which are monotonically labelled by elements of $\{1, 2, \ldots, k\}$ fulfills as $n \to \infty$*

$$M_{k,n} \sim C_k q_k^{-n} n^{-3/2}.$$

Here q_k is the only root of $p_k(z) - z$ in $(0, 1)$, where $p_k(z)$ is defined by $(k \geq 0)$

$$p_0(z) = \frac{1-z}{2}, \qquad p_{k+1}(z) = p_k(z)(1 - z - p_k(z)),$$

and C_k is a constant. Moreover, q_k fulfills as $k \to \infty$

$$q_k = \frac{\log 2}{k} + O\left(\frac{\log k}{k^2}\right).$$

Binary Trees. *The number $B_{k,n}$ of binary trees with n nodes which are monotonically labelled by elements of $\{1, 2, \ldots, k\}$ fulfills as $n \to \infty$*

$$\begin{cases} B_{k,n} = 0 & \textit{for } n \equiv 0\ (2), \\ B_{k,n} \sim C_k q_k^{-n} n^{-3/2} & \textit{for } n \equiv 1\ (2). \end{cases}$$

Here q_k is the only root of $p_k(z)$ in $(0, 1)$, where $p_k(z)$ is defined by $(k \geq 0)$

$$p_0(z) = \tfrac{1}{2}, \qquad p_{k+1}(z) = p_k(z)(1 - p_k(z)) - z^2,$$

and C_k is a constant. Moreover, q_k fulfills as $k \to \infty$

$$q_k = \frac{\pi}{2k} + O\left(\frac{\log k}{k^2}\right).$$

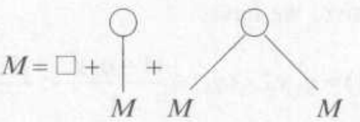
Similar, but more complicated asymptotic formulas can be obtained for the more general tree structures mentioned above.

Properties of monotonically labelled tree structures have been studied extensively in literature. In [8] Kirschenhofer and Prodinger have treated the problem of the average height of monotonically labelled binary trees, in [6] and [7] Kirschenhofer has studied the average shape, and in [9] Kirschenhofer and Prodinger have considered the average oscillation of the contour of monotonically labelled ordered trees.

Remark. We will frequently use the symbol $\langle y(z), z^n \rangle$ for the coefficient of z^n in the power series $y(z)$.

2. Monotonically labelled Motzkin trees

A *Motzkin tree* or *unary-binary* tree is either a single leave or it is build up by an internal node ◯ with either one or two (ordered) subtrees. This can be illustrated by the following symbolic equation:

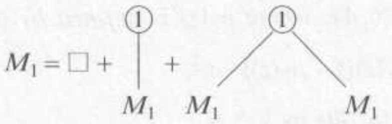


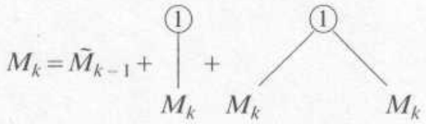
Let M_k denote the family of Motzkin trees the internal nodes of which are labelled monotonically by elements of $\{1, 2, \ldots, k\}$. Let $\langle y_k, z^n \rangle$ be the number of trees in M_k with n internal nodes, and let

$$y_k(z) = \sum_{n \geq 0} \langle y_k(z), z^n \rangle z^n \tag{2.1}$$

be a corresponding generating function.

Furthermore let $\tilde{M}_k$ be the family of Motzkin trees the internal nodes of which are labelled monotonically by elements of $\{2, 3, \ldots, k+1\}$. Then the M_k fulfill the following system of symbolic equations:





(2.2)

This system of symbolic equations can be translated into a system of recurrence relations for the generating functions y_k (cf. [3]):

$$y_k(z) = y_{k-1}(z) + zy_k(z) + zy_k^2(z) \quad (k \geq 1), \qquad y_0(z) = 1. \tag{2.3}$$

To determine the asymptotic behaviour of $\langle y_k, z^n \rangle$ we use a method originally due to Darboux [2] which is described in Harary/Palmer [5, pp. 211f] and in [1].

Let q_k denote the positive singularity of y_k of smallest modulus. We define:

$$F(z, w) = y_{k-1}(z) + (z-1)w + zw^2. \tag{2.4}$$

The singularity q_k must now fulfill (cf. [5], [1])

$$F(q_k, w_k) = F_w(q_k, w_k) = 0 \quad \text{with } w_k = y_k(q_k).$$

The solution of $F_w = 0$ is $w = (1-z)/2z$.

Substituting this into (2.4) we see that q_k is the smallest positive root of

$$zy_{k-1}(z) = (1-z)^2/4. \tag{2.5}$$

Moreover, q_k is the only singularity with smallest modulus:

To see this consider $z = \bar{q}_k$ to be a root of (2.5) with $\bar{q}_k \neq q_k$, but $|\bar{q}_k| = q_k$. Since the coefficients of $y_{k-1}(z)$ are nonnegative, we have

$$|\bar{q}_k y_{k-1}(\bar{q}_k)| \leq |\bar{q}_k| \, y_{k-1}(|\bar{q}_k|) = q_k y_{k-1}(q_k) = \frac{(1-q_k)^2}{4} < \frac{|1-\bar{q}_k|^2}{4},$$

which is obviously a contradiction.

By the method described in Harary/Palmer [5] we get

$$\langle y_k(z), z^n \rangle \sim C_k q_k^{-n} n^{-3/2} \quad (n \to \infty).$$

Since $\langle y_k(z), z^n \rangle \geq \langle y_1(z), z^n \rangle$, it is clear that

$$0 < q_k \leq q_1 = 3 - 2\sqrt{2} < 1. \tag{2.6}$$

We proceed to show that q_k is the only root of $p_k(z)=z$ in (0,1), where $p_k(z)$ is a function sequence that satisfies a nonlinear recurrence relation.

Definition 2.1. Let the function sequence $p_n(z)$ be defined by

$$p_0(z)=\frac{1-z}{2}, \qquad p_{n+1}(z)=p_n(z)(1-z-p_n(z)). \tag{2.7}$$

From this definition it follows that $q_k y_{k-1}(q_k)=p_1(q_k)$.

From (2.3) we have for $0\le i\le k-1$

$$zy_{k-i-1}=zy_{k-i}-z^2y_{k-i}-z^2y_{k-i}^2, \tag{2.8}$$

so that $q_k y_{k-i}(q_k)=p_i(q_k)$ for $1\le i\le k$. In particular we have

$$q_k y_0(q_k)=q_k=p_k(q_k).$$

Since q_k lies in (0,1) by (2.6), q_k is a root of $p_k(z)=z$ in (0,1).

Our next step will be to prove that q_k is the **only** root of $p_k(z)=z$ in (0,1).

Definition 2.2. $h_n(z):=p_n(z)/(1-z)$.

From this definition the following recurrence relation for h_n can easily be derived:

$$h_0(z)=\tfrac{1}{2}, \qquad h_{n+1}(z)=(1-z)h_n(z)(1-h_n(z)).$$

Using this recurrence relation we will show by induction that

$$0\le h_{n+1}(z)<h_n(z)<\tfrac{1}{2} \quad \text{for } 0<z<1 \text{ and } n\ge 1.$$

The case $n=1$ holds true, and if we assume $0<1-z<1, 0\le h_n(z)<\frac{1}{2}$, and $\frac{1}{2}<1-h_n(z)\le 1$, it is an easy consequence that

$$0\le(1-z)h_n(z)(1-h_n(z))=h_{n+1}(z)<h_n(z)<\tfrac{1}{2}.$$

We proceed to prove that $h_n'(z)<0$ for $0<z<1$ and $n\ge 1$: We have $h_1'(z)=-\frac{1}{4}<0$ and for $n\ge 1$

$$h_{n+1}'(z)=-h_n(z)(1-h_n(z))+h_n'(z)(1-z)(1-2h_n(z)).$$

Since $h_n(1-h_n)\ge 0$ and $(1-z)(1-2h_n)>0$ (see above), we have $h_n'(z)<0$ for $n\ge 0$.

From these results and from the definition of h_n it follows that

(1) $\quad 0\le p_{n+1}(z)\le p_n(z)<\frac{1}{2}(1-z)=p_0(z)$ for $0<z<1$ and $n\ge 1$, and

(2) $\quad p_n'(z)<0$ for $0<z<1$ and $n\ge 0$.

Remark. Let $q_k(B)$ be the smallest positive singularity of the generating function of the monotonically labelled extended binary trees. Then the following hold (cf. [11]):

(i) $q_{k+1}(B)=q_k(B)(1-q_k(B)), \qquad q_0(B)=\frac{1}{2}$,

(ii) $q_k(B)=\frac{1}{k}+\mathrm{O}\left(\frac{\log k}{k^2}\right) \quad (k\to\infty)$.

Observing $p_k(0)=q_k(B)$ we conclude from this remark and the considerations above that q_k is the **only** root of $p_k(z)=z$ in (0,1).

Theorem 2.3. *The number $\langle y_k, z^n\rangle$ of Motzkin trees which are monotonically labelled by elements of $\{1,2,\ldots,k\}$ fulfills as $n\to\infty$*

$$\langle y_k, z^n\rangle \sim C_k q_k^{-n} n^{-3/2}.$$

Here q_k is the only root of $p_k(z)-z$ in (0,1), where $p_k(z)$ is defined by ($k\geq 0$)

$$p_0(z)=\frac{1-z}{2}, \qquad p_{k+1}(z)=p_k(z)(1-z-p_k(z)),$$

and

$$C_k=\frac{1}{2\sqrt{\pi q_k}}\left((1-p_k'(q_k))\prod_{l=1}^{k-1}(1-q_k-2p_l(q_k))^{-1}\right)^{1/2}.$$

Proof. It remains to show the formula for C_k.

Using Theorem 5 stated in [1] we see that

$$C_k=\frac{1}{2\sqrt{\pi}}\left(F_z(q_k, y_k(q_k))\right)^{1/2},$$

so that, since

$$F_z(z,w)=y_{k-1}'+w+w^2,$$

we only have to investigate y_{k-1}'.

Differentiating (2.8) we get the recurrence relation

$$y_{k-i}'(1-z-2zy_{k-i})=y_{k-i-1}'+y_{k-i}+y_{k-i}^2.$$

Observing

$$y_{k-i}(q_k)=p_i(q_k)/q_k$$

we get at $z=q_k$

$$y_{k-i}'(q_k)=q_k^{-2}\sum_{j=1}^{k-i}(q_k p_{k-j}(q_k))\prod_{l=j}^{k-i}(1-q_k-2p_{k-l}(q_k))^{-1}.$$

Hence

$$F_z(q_k, y_k(q_k)) = q_k^{-2} \sum_{j=0}^{k-1} (q_k p_j(q_k) + p_j^2(q_k)) \prod_{l=1}^{j} (1 - q_k - 2p_l(q_k))^{-1}.$$

This reduces to the form stated in the proposition if we observe (2.7) and the derivation of (2.7). □

To determine the asymptotic behaviour of q_k as $k \to \infty$, we establish an asymptotic form of $p_k(z)$.

The following lemma is inspired by Lemma 5 in [4], where a similar function sequence plays a central role in determining the average height of binary trees.

Lemma 2.4. *For* $0 < z < 1$ *we have*

$$p_n(z) = \frac{z(1-z)^{n+1}}{1-(1-z)^n} \left(1 + \frac{\varepsilon_n(z) \cdot z}{1-(1-z)^n}\right)^{-1},$$

where

$$\varepsilon_n(z) = 2 + \sum_{k=0}^{n-1} (1-z)^k \frac{h_k(z)}{1-h_k(z)}.$$

Proof. We start from the recurrence

$$h_{j+1}(z) = (1-z) h_j(z)(1-h_j(z))$$

and we take out the $(1-z)^j$ factor present in $h_j(z)$. Let $r_j = h_j(z)/(1-z)^j$, then we have

$$r_{j+1} = r_j(1-h_j).$$

Since we have $r_j > 0$ for $0 < z < 1$, we can define $s_j := r_j^{-1}$ and get

$$s_{j+1} = s_j(1-h_j)^{-1} = s_j\left(1 + h_j + \frac{h_j^2}{1-h_j}\right) = s_j + (1-z)^j + \frac{h_j(1-z)^j}{1-h_j}.$$

If we sum up these identities for $j = 0, \ldots, n-1$ and use $s_0 = r_0^{-1} = h_0^{-1} = 2$ we get

$$s_n = \sum_{j=0}^{n-1} (1-z)^j + 2 + \sum_{j=0}^{n-1} (1-z)^j \frac{h_j}{1-h_j} = \frac{1-(1-z)^n}{z} + \varepsilon_n(z).$$

From this the lemma follows. □

Lemma 2.5. *For* $0 \le z \le q_n(B) = 1/n + O((\log n)/n^2)$ *the following holds uniformly*:

$$p_n(z) = \frac{z(1-z)^{n+1}}{1-(1-z)^n}\left(1 + O\left(\frac{\log n}{n}\right)\right).$$

Proof. Since $x/(1-x)$ is monotone increasing in (0,1), and $h_k(z)$ is monotone decreasing in (0,1), we have

$$h_k \le h_k(0) = q_k(B)$$

and

$$0 \le \varepsilon_n(z) \le 2 + \sum_{k=0}^{n-1} \frac{h_k}{1-h_k} \le 2 + \sum_{k=0}^{n-1} \frac{q_k(B)}{1-q_k(B)} = \mathrm{O}(\log n). \tag{2.9}$$

A simple argument shows that $f(z) = z/(1-(1-z)^n)$ is monotone increasing in (0,1). Hence

$$f(z) \le f(q_n(B)) = \mathrm{O}(1/n).$$

Combining these results we get the desired estimate. □

Lemma 2.5 enables us to determine the asymptotic behaviour of the root of $p_n(z) - z$ in (0,1).

Lemma 2.6. *Let z_n be the root of $p_n(z) = z$ in (0,1). Then*

$$z_n = \frac{\log 2}{n} + \mathrm{O}\left(\frac{\log n}{n^2}\right).$$

Proof. If we observe $z_n > 0$,

$$\frac{z_n(1-z_n)^{n+1}}{1-(1-z_n)^n}\left(1 + \mathrm{O}\left(\frac{\log n}{n}\right)\right) = z_n$$

implies

$$(1-z_n)^n(2-z_n) = 1 + (1-z_n)^{n+1}\,\mathrm{O}\left(\frac{\log n}{n}\right).$$

Since $z_n = \mathrm{O}(1/n)$ in $(0, q_n(B))$,

$$1/(2-z_n) = \tfrac{1}{2} + \mathrm{O}(1/n).$$

Hence

$$(1-z_n)^n = \tfrac{1}{2}\left(1 + \mathrm{O}\left(\frac{\log n}{n}\right)\right).$$

So we get ($z_n \in \mathbb{R}^+$):

$$\begin{aligned} 1 - z_n &= \exp\left(-\frac{1}{n}\log 2\right)\left(1 + \mathrm{O}\left(\frac{\log n}{n^2}\right)\right) \\ &= 1 - \frac{\log 2}{n} + \mathrm{O}\left(\frac{\log n}{n^2}\right), \end{aligned}$$

and the lemma is proved. □

So we have shown the following

Theorem 2.7. *The smallest positive singularities q_k of the generating functions y_k fulfill as $k \to \infty$*

$$q_k = \frac{\log 2}{k} + O\left(\frac{\log k}{k^2}\right).$$

It is possible to derive more precise estimates by iteration of this process, e.g.:

Theorem 2.8. *The smallest positive singularities q_k of the generating functions y_k fulfill as $k \to \infty$*

$$q_k = \log 2\left(\frac{1}{k} - \frac{\log k}{k^2} + O\left(\frac{1}{k^2}\right)\right).$$

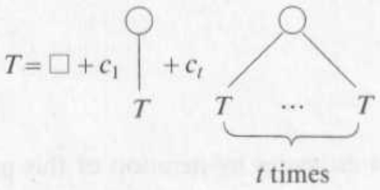
Numerical results corresponding to Theorem 2.7 are displayed in Table 1.

Table 1

k	q_k	$q_k k/\log 2$
1	0.171572875	0.248
2	0.129158910	0.373
3	0.105115939	0.455
4	0.089186654	0.515
5	0.077717283	0.561
6	0.069007052	0.597
7	0.062139224	0.628
8	0.056570040	0.653
9	0.051954174	0.675
10	0.048060707	0.693
50	0.012421752	0.896
100	0.006511413	0.939
200	0.003346614	0.966
300	0.002254027	0.976
400	0.001699733	0.981
500	0.001364415	0.984
600	0.001139674	0.987
700	0.000978537	0.988
800	0.000857343	0.990
900	0.000762874	0.991
1000	0.000687166	0.991

The method developed in this section may be generalized to *unary-t-ary trees* with weights attached to their nodes.

Such a tree consists of leaves and of internal nodes which have either one or t ordered subtrees. The internal nodes with one successor are weighted with $c_1 > 0$, those with t successors are weighted with $c_t > 0$. So the family of *weighted unary-t-ary trees* T may be defined by the formal equation:



Let T_k be the family of weighted unary-t-ary trees which are monotonically labelled by elements of $\{1, \ldots, k\}$. Furthermore let $\langle y_k(z), z^n \rangle$ be the number of trees of T_k with exactly n internal nodes.

Then one can show the following

Theorem 2.9. *The number of the monotonically labelled unary-t-ary trees with weights attached to their nodes in the manner described above fulfills as* $n \to \infty$

$$\langle y_k, z^n \rangle \sim C_k q_k^{-n} n^{-3/2}.$$

Here q_k *is the only root of* $p_k(z) - z$ *in* $(0, 1/c_1)$, *where* $p_k(z)$ *is defined by* $(k \geq 0)$

$$p_0(z) = \left(\frac{1 - c_1 z}{t c_t} \right)^{1/(t-1)}, \qquad p_{k+1}(z) = p_k(z)(1 - c_1 z - c_t p_k^{t-1}(z)),$$

and C_k *is a constant. Moreover,* q_k *fulfills as* $k \to \infty$

$$q_k = \frac{\log(1 + c_1/c_t)}{c_1(t-1)} \frac{1}{k} + \mathrm{O}\left(\frac{\log k}{k^2} \right).$$

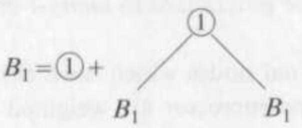
Remark. If we set $c_t = 1$ and let $c_1 \to 0$, we formally obtain the results of Prodinger and Urbanek [11] concerning the family of t-ary trees.

3. Monotonically labelled binary trees

In this section we illustrate how to apply the method developed in the previous section to the case of not extended binary trees.

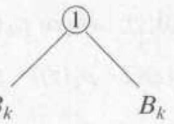
The family of *not extended binary trees* consists of trees the internal nodes of which have two ordered successors.

Let the family of monotonically labelled (not extended) binary trees be defined by the following symbolic equations (in this case the leaves are also considered to be labelled):



(3.1)

$$B_k = \tilde{B}_{k-1} + ① +$$



Here B_k denotes the family of binary trees the nodes of which are labelled by elements of $\{1, \ldots, k\}$, and $\tilde{B}_k$ denotes the same family except that the nodes are labelled by $\{2, 3, \ldots, k+1\}$. Let $\langle y_k(z), z^n \rangle$ be the number of trees in B_k with n nodes (it is clear that $\langle y_k(z), z^{2n} \rangle = 0$) and let $y_k(z) = \sum_{n \geq 0} \langle y_k(z), z^n \rangle z^n$ be a corresponding generating function. Then (3.1) can be translated into the following system of recurrence relations:

$$y_k(z) = y_{k-1}(z) + z + z y_k^2(z), \qquad y_0(z) = 0. \tag{3.2}$$

From this we see that y_k has two singularities, namely $\pm q_k$, $q_k > 0$.

If we define the function sequence $p_n(z)$ by

$$p_0(z) = \tfrac{1}{2}, \qquad p_{n+1}(z) = p_n(z)(1 - p_n(z)) - z^2, \tag{3.3}$$

it is easy to show that the two singularities $+q_k$ and $-q_k$ of the generating function y_k are roots of $p_k(z) = 0$.

In the following we will show that q_k is the only root of $p_k(z) = 0$ in $0 < z < 1$.

Remark. Since $p_n(z)$ is an even function for all n, it suffices to study $p_n(z)$ for $0 < z < 1$.

We will show by induction that $p_{n+1}(z) < p_n(z) < \frac{1}{2}$ for $0 < z \leq 1$, $n \geq 1$. We have

$$p_1(z) = \tfrac{1}{4} - z^2 < \tfrac{1}{2}$$

and

$$p_{n+1}(z) = p_n(1 - p_n) - z^2 < p_n - p_n^2 \leq p_n.$$

Thus the proposition follows.

We proceed to prove that $p_n'(z) < 0$ for $0 < z < 1$, $n \geq 1$. We have $p_1'(z) = -2z < 0$ and using induction we see that

$$p_{n+1}' = p_n'(1 - 2p_n) - 2z < p_n'(1 - 2p_n) < 0.$$

The last estimate holds, because of the assumption $p_n' < 0$ and because $p_n < \frac{1}{2}$, which was shown above.

From these two estimates we see that q_k is the only root of $p_k(z)$ in $(0,1)$. Hence we have shown

Theorem 3.1. *The number* $\langle y_k(z), z^n \rangle$ *fulfills as* $n \to \infty$

$$\begin{cases} \langle y_k(z), z^n \rangle = 0 & \textit{for } n \equiv 0 \ (2), \\ \langle y_k(z), z^n \rangle \sim C_k q_k^{-n} n^{-3/2} & \textit{for } n \equiv 1 \ (2). \end{cases}$$

Here q_k is the only root of $p_k(z)$ in (0,1), where $p_k(z)$ is defined by ($k \geq 0$)

$$p_0(z) = \tfrac{1}{2}, \qquad p_{k+1}(z) = p_k(z)(1 - p_k(z)) - z^2,$$

and C_k is a constant.

We are now going to establish the asymptotic behaviour of q_k as $k \to \infty$.

We define $h_n(z)$ by

$$h_n(z) := \frac{p_n(z) + \mathrm{i}z}{1 + 2\mathrm{i}z},$$

where i denotes the imaginary unit, i.e., $\mathrm{i}^2 = -1$.

Thus $h_n(z)$ fulfills the following recurrence relation for $n \geq 0$:

$$h_0(z) = \tfrac{1}{2}, \qquad h_{n+1}(z) = (1 + 2\mathrm{i}z)h_n(z)(1 - h_n(z)).$$

Lemma 3.2. *For $0 < z < 1$ we have*

$$p_n(z) = \frac{2\mathrm{i}z(1+2\mathrm{i}z)^{n+1}}{(1+2\mathrm{i}z)^n - 1}\left(1 + \frac{\varepsilon_n(z)2\mathrm{i}z}{(1+2\mathrm{i}z)^n - 1}\right)^{-1} - \mathrm{i}z,$$

where

$$\varepsilon_n(z) = 2 + \sum_{j=0}^{n-1}(1+2\mathrm{i}z)^j \frac{h_j}{1 - h_j}.$$

Proof. Very similar to the proof of Lemma 2.4. □

Before we continue, we need a crude estimate for q_k.

Lemma 3.3. *We have $0 < q_k < C/k$ for a $C > 0$ and $q_k \lesssim 2/k$ for $k \to \infty$.*

Proof. Let $B_{2n+1,k}$ denote the family of binary trees where all $(2n+1)$ nodes (even the leaves) are monotonically labelled by $\{1, 2, \ldots, k\}$; let $a_{2n+1,k}$ be the number of different trees in $B_{2n+1,k}$. Then we have (cf. Theorem 3.1)

$$a_{2n+1,k} \sim C_k q_k^{-(2n+1)} n^{-3/2} \quad (n \to \infty).$$

Let $B'_{n,k}$ be the family of binary trees with n internal nodes (and $n+1$ leaves) the internal nodes of which are monotonically labelled by $\{1, 2, \ldots, \lceil k/2 \rceil\}$, and the leaves of which are labelled by $\{\lceil k/2 \rceil + 1, \ldots, k\}$, where $k \geq 2$ and

$$\lceil k/2 \rceil = \begin{cases} k/2 & \text{if } k \equiv 0\,(2), \\ (k+1)/2 & \text{if } k \equiv 1\,(2). \end{cases}$$

Let $b_{n,k}$ be the number of different trees in $B'_{n,k}$. Then we have [11]

$$b_{n,k} \sim C'_k (q_{\lceil k/2 \rceil}(B))^{-n} (k - \lceil k/2 \rceil)^{n+1} n^{-3/2} \quad (n \to \infty),$$

where $q_i(B)$ denotes like in the previous section the singularities of the generating functions of the monotonically labelled extended binary trees.

Since for sufficiently large n we have $a_{2n+1,k} \geq b_{n,k}$, we get

$$q_k^2 \leq (k - \lceil k/2 \rceil)^{-1} q_{\lceil k/2 \rceil}(B) \sim 4/k^2,$$

because of $q_k(B) \sim 1/k$, which has been shown by Prodinger and Urbanek [11]. □

Lemma 3.4. *For $0 < z \leq q_n$ we have*

$$p_n(z) = \frac{2\mathrm{i}z(1+2\mathrm{i}z)^{n+1}}{(1+2\mathrm{i}z)^n - 1}\left(1 + \mathrm{O}\left(\frac{\log n}{n}\right)\right) - \mathrm{i}z.$$

Proof. For $0 < z \leq q_n$ we have $|h_n(z)| \leq p_n(0) + z$, because of

$$|h_n| = \frac{\sqrt{p_n^2(z) + z^2}}{\sqrt{1+4z^2}} \leq p_n(z) + z \leq p_n(0) + z.$$

Since $|1 + 2\mathrm{i}z|^j = \mathrm{O}(1)$ for $0 < z \leq q_n$ and because of the estimate above, we get

$$|\varepsilon_n(z)| \leq 2 + \sum_{j=0}^{n-1} |1 + 2\mathrm{i}z|^j \frac{|h_j|}{|1 - h_j|} \leq 2 + \mathrm{O}\left(\sum_{j=0}^{n-1} (p_j(0) + z)\right).$$

Since

$$p_j(0) = \frac{1}{j} + \mathrm{O}\left(\frac{\log j}{j^2}\right) \quad \text{and} \quad 0 < z < \frac{C}{n},$$

we have $\varepsilon_n(z) = \mathrm{O}(\log n)$ for $0 < z \leq q_n$.

The lemma follows now if we observe

$$\left|\frac{2\mathrm{i}z}{(1+2\mathrm{i}z)^n - 1}\right| = \mathrm{O}\left(\frac{1}{n}\right) \quad \text{for } 0 < z < \frac{C}{n}. \quad \square$$

Lemma 3.5. *If z_n is the root of $p_n(z) = 0$* in $(0, 1)$, then we have as $n \to \infty$

$$z_n = \frac{\pi}{2n} + \mathrm{O}\left(\frac{\log n}{n^2}\right).$$

Proof. Lemma 3.4 implies

$$(1 + 2\mathrm{i}z_n)^n (1 + 4\mathrm{i}z_n) = -1 + \mathrm{O}\left(\frac{\log n}{n}\right).$$

Since

$$\frac{1}{1 + 4\mathrm{i}z_n} = 1 + \mathrm{O}\left(\frac{1}{n}\right) \quad \text{for } 0 < z_n < \frac{C}{n},$$

this equals

$$(1+2iz_n)^n=(-1)\left(1+O\left(\frac{\log n}{n}\right)\right).$$

Thus we get

$$1+2iz_n=\exp\left(\frac{1}{n}\log(-1)\right)\left(1+O\left(\frac{\log n}{n^2}\right)\right).$$

Since $\log(-1)=i\pi(2l+1)$, $l\in\mathbb{Z}$, we find that for $(1/n)i\pi(2l+1)$, $l\geq 1$, we would have $z_n>2/n$, which would contradict Lemma 3.3. Hence

$$1+2iz_n=\exp\left(\frac{1}{n}i\pi\right)\left(1+O\left(\frac{\log n}{n^2}\right)\right)=1+\frac{i\pi}{n}+O\left(\frac{\log n}{n^2}\right). \qquad \square$$

So we can state

Theorem 3.6. *The smallest positive singularities q_k of the generating functions $y_k(z)$ fulfill as $k\to\infty$*

$$q_k=\frac{\pi}{2k}+O\left(\frac{\log k}{k^2}\right).$$

Numerical results corresponding to Theorem 3.6 are displayed in Table 2.

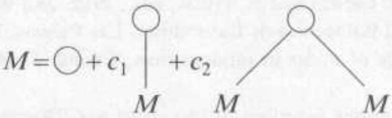
Table 2

k	q_k	$q_k k\,2/\pi$
1	0.500000000	0.318
2	0.340625019	0.434
3	0.265821288	0.508
4	0.220330088	0.561
5	0.189147001	0.602
6	0.166208351	0.635
7	0.148520885	0.662
8	0.134412536	0.685
9	0.122866540	0.704
10	0.113224938	0.721
50	0.028398195	0.904
100	0.014822830	0.944
200	0.007601486	0.968
300	0.005115922	0.977
400	0.003856380	0.982
500	0.003094894	0.985
600	0.002584716	0.987
700	0.002219022	0.989
800	0.001944030	0.990
900	0.001729710	0.991
1000	0.001557971	0.992

Theorem 3.6 can be generalized in the following way. The family M of the *weighted unary-binary trees* may be defined in the following manner:

A unary-binary tree with weights attached to its nodes consists of a node with either one or two ordered subtrees, where the weight $c_1 > 0$ is attached to the nodes with one successor and the weight $c_2 > 0$ is attached to the nodes with two successors. This is a special case of the so-called *simply generated families of trees* introduced by Meir and Moon [10].

The family M fulfills the following symbolic equation:



If the nodes (even the leaves) are labelled monotonically by $\{1, \ldots, k\}$, one gets the family M_k, the family of *weighted monotonically labelled unary-binary trees.*

Let $\langle y_k, z^n \rangle$ be the number of trees in M_k with n nodes. Then one can show the following

Theorem 3.7. *The number* $\langle y_k, z^n \rangle$ *fulfills as* $n \to \infty$

$$\langle y_k, z^n \rangle \sim C_k q_k^{-n} n^{-3/2},$$

where q_k *is the only root of* $p_k(z)$ in $(0, 1/c_1)$. *Here* $p_k(z)$ *is defined by* $(k \geq 0)$

$$p_0(z) = \frac{1 - c_1 z}{2c_2}, \qquad p_{k+1}(z) = p_k(z)(1 - c_1 z - c_2 p_k(z)) - z^2.$$

Moreover, q_k *fulfills as* $k \to \infty$

$$q_k = \frac{\log((c_1 + \beta)/(c_1 - \beta))}{\beta} \frac{1}{k} + \mathrm{O}\left(\frac{\log k}{k^2}\right),$$

where $\beta = \sqrt{c_1^2 - 4c_2}$.

Remark. The results of Theorem 3.7 are formally still valid, if $c_1 \to 0$.

References

[1] E.A. Bender, Asymptotic methods in enumeration, SIAM Review 16 (1974) 485–515.

[2] G. Darboux, Mémoire sur l'approximation des fonctions de très grands nombres, et sur une classe étendu de développements en série, J. Math. Pures Appl. (3) 4 (1878) 5–56.

[3] P. Flajolet, Analyse d'algorithmes de manipulation d'arbres et de fichiers, Thèse Paris-Sud-Orsay, 1979, also in: Cahiers de BURO 34–35 (1981), 1–209.

[4] P. Flajolet and O. Odlyzko, The average height of binary trees and other simple trees, J. Comput. System Sci. 25 (1982) 171–213.

[5] F. Harary and E.M. Palmer, Graphical Enumeration (Academic Press, New York and London, 1973).

[6] P. Kirschenhofer, On the average shape of monotonically labelled tree structures, Discrete Appl. Math. 7 (1984) 161–181.

[7] P. Kirschenhofer, On the height of leaves in binary trees, J. Combin. Inform. System Sci. 8 (1983) 44–60.

[8] P. Kirschenhofer and H. Prodinger, On the average height of monotonically labelled binary trees, in: A. Hajnal et al., eds., Finite and Infinite Sets, Coll. Soc. Janos Bolyai 10 (North-Holland, Amsterdam, 1984) 509–525.

[9] P. Kirschenhofer and H. Prodinger, On the average oscillation of the contour of monotonically labelled ordered trees, in: A. Ballester, D. Cardus and E. Trillas, eds., Proc. 2nd World Conf. on Math. at the Service of Man (Universidad Politecnica de Las Palmas, Las Palmas, 1982) 394–399.

[10] A. Meir and J.W. Moon, On the altitude of nodes in random trees, Canad. J. Math. 30 (1978) 997–1015.

[11] H. Prodinger and F.J. Urbanek, On monotone functions of tree structures, Discrete Appl. Math. 5 (1983) 223–239.