COLLOQUIA MATHEMATICA SOCIETATIS JÁNOS BOLYAI
37. FINITE AND INFINITE SETS, EGER (HUNGARY), 1981.

# ON THE AVERAGE HEIGHT OF MONOTONICALLY LABELLED BINARY TREES

P. KIRSCHENHOFER — H. PRODINGER

## 1. INTRODUCTION

In a recent paper [3], [4] Flajolet and Odlyzko have studied the average height of binary trees (in the sense of Knuth [5]): if the height $h(t)$ of a binary tree is given inductively by

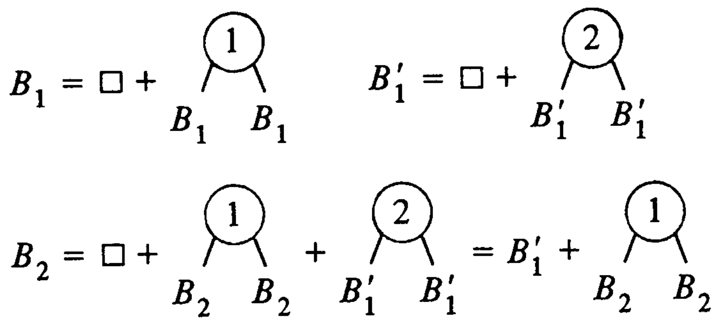
$$h(\square) = 1$$

(1) $$h\left(\begin{smallmatrix} & \circ & \\ / & & \backslash \\ t_1 & & t_2 \end{smallmatrix}\right) = 1 + \max\{h(t_1), h(t_2)\},$$

then the average height of all binary trees with $n$ internal nodes is shown to be asymptotically equivalent to $2\sqrt{\pi n}$.

The present paper deals with an asymptotic evaluation of the average height of binary trees, the nodes of which have been labelled by the numbers 1 and 2 such that any sequence of labels starting from the root of the tree is monotone.

If we define the families $B_1, B_1', B_2$ of labelled binary trees by the formal equations

(2) 

then $B_1$ is the family of all binary trees with all nodes labelled by 1, $B_1'$ with all nodes labelled by 2 and $B_2$ the family of monotonically labelled binary trees as mentioned above. An easy consequences of (2) are the following equations for the generating functions $y_1, y_2$ of $B_1, B_2$, respectively:

(3) $$y_1 = 1 + zy_1^2; \quad y_2 = y_1 + zy_2^2.$$

That means

$$y_1(z) = \sum_{n \geq 0} y_{1,n} z^n = \frac{1 - (1 - 4z)^{\frac{1}{2}}}{2z},$$

and so

$$y_{1,n} = \frac{1}{n+1} \binom{2n}{n} \sim \pi^{-\frac{1}{2}} 4^n n^{-\frac{3}{2}} \qquad (n \to \infty).$$

The generating function

$$y_2(z) = \sum_{n \geq 0} y_{2,n} z^n = \frac{1 - (-1 + 2(1 - 4z)^{\frac{1}{2}})^{\frac{1}{2}}}{2z}$$

has been studied in the paper [7]; the asymptotic behaviour of its coefficients $y_{2,n}$ is established by Darboux's method (see e.g. [6]) to be

(4) $$y_{2,n} \sim 4(6\pi)^{-\frac{1}{2}} \left(\frac{16}{3}\right)^n n^{-\frac{3}{2}}.$$

If we denote by $B_{2,n}$ the set of trees of $B_2$ with exactly $n$ internal nodes and by

(5) $$H_{2,n} = \sum_{t \in B_{2,n}} h(t)$$

(where the height $h(t)$ is given by (1) disregarding the labelling of the tree) we will prove the following asymptotic formula:

$$(6) \qquad H_{2,n} \sim \frac{8}{3} \left(\frac{16}{3}\right)^n n^{-1} \qquad (n \to \infty).$$

The average height of trees of $B_{2,n}$ is then given by the ratio $\frac{H_{2,n}}{y_{2,n}}$ and we get the following main result:

**Theorem.** *The average height of binary trees labelled monotonically by the numbers* 1 *and* 2 *fulfills the asymptotic relation*

$$(7) \qquad \frac{H_{2,n}}{y_{2,n}} \sim \left(\frac{8\pi n}{3}\right)^{\frac{1}{2}} \qquad (n \to \infty).$$

The proof of (6) will be given by studying the following two notions of "height" $h_1(t), h_2(t)$:

$$h_1(\square) = 1$$

$$(8) \qquad h_1\left(\begin{smallmatrix} & \textcircled{1} & \\ & / \quad \backslash & \\ t_1 & & t_2 \end{smallmatrix}\right) = 1 + \max\{h_1(t_1), h_1(t_2)\}$$

$$h_1\left(\begin{smallmatrix} & \textcircled{2} & \\ & / \quad \backslash & \\ t_1 & & t_2 \end{smallmatrix}\right) = \max\{h_1(t_1), h_1(t_2)\} \; (= 1),$$

$$h_2(\square) = 1$$

$$(9) \qquad h_2\left(\begin{smallmatrix} & \textcircled{1} & \\ & / \quad \backslash & \\ t_1 & & t_2 \end{smallmatrix}\right) = \max\{h_2(t_1), h_2(t_2)\}$$

$$h_2\left(\begin{smallmatrix} & \textcircled{2} & \\ & / \quad \backslash & \\ t_1 & & t_2 \end{smallmatrix}\right) = 1 + \max\{h_2(t_1), h_2(t_2)\} = h\left(\begin{smallmatrix} & \textcircled{2} & \\ & / \quad \backslash & \\ t_1 & & t_2 \end{smallmatrix}\right)$$

($h_i(t) - 1$ counts the maximal number of consecuting $i$'s in a sequence starting from the root.) In analogy to (5) we define the quantities

(10) $$H^{(1)}_{2,n} = \sum_{t \in B_{2,n}} h_1(t), \quad H^{(2)}_{2,n} = \sum_{t \in B_{2,n}} h_2(t).$$

It is trivial to see that the following relations hold:

(11) $$y_{2,n} \leqslant H^{(1)}_{2,n}, H^{(2)}_{2,n} \leqslant (n+1) y_{2,n},$$

(12) $$H^{(1)}_{2.n} \leqslant H_{2,n} \leqslant H^{(1)}_{2,n} + H^{(2)}_{2,n}.$$

The idea of proof of relation (6) is now to establish independently the following two results:

(13) $$H^{(1)}_{2,n} \sim \frac{8}{3}\left(\frac{16}{3}\right)^n n^{-1} \qquad (n \to \infty),$$

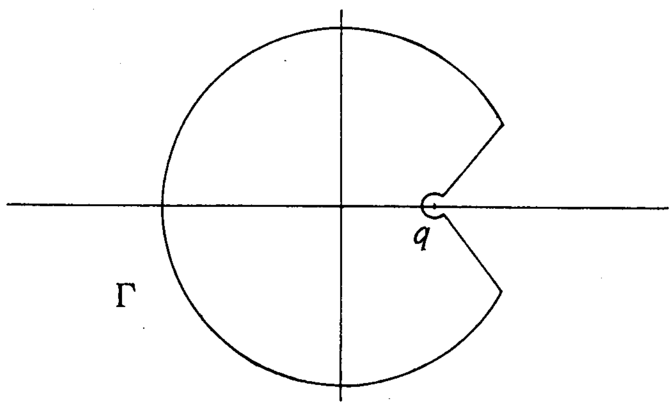
(14) $$H^{(2)}_{2,n} = o(H^{(1)}_{2,n}) \qquad (n \to \infty).$$

In both cases the method is to study the behaviour of an appropriate analytical continuation of the functions

(15) $$H^{(1)}_2(z) = \sum_{n \geqslant 0} H^{(1)}_{2,n} z^n$$

(16) $$H^{(2)}_2(z) = \sum_{n \geqslant 0} H^{(2)}_{2,n} z^n$$

in a sector around their singularities nearest to the origin and to evaluate the coefficients by means of Cauchy's integral formula using a contour of integration $\Gamma$ as depicted in the following diagram




($q$ is the singularity in question.) The advantage of the use of $\Gamma$ is that it gives predominance to the behaviour of the function around its singularity.

Section 2 will be devoted to the proof of (13) by showing that the function $H_2^{(1)}(z)$ has a logarithmic singularity at $q = \frac{3}{16}$ and that the study of its behaviour around the singularity is implicitly given by Flajolet and Odlyzko's method of deriving the average height of nonlabelled binary trees in [3].

The main difficulty which remains is the determination of the behaviour of the function $H_2^{(2)}(z)$ around its singularity at $q = \frac{3}{16}$. In Section 3 we derive the following result:

$$(17) \qquad H_2^{(2)}(z) = H_2^{(2)}\left(\frac{3}{16}\right) + O\left(\left|z - \frac{3}{16}\right|^{\frac{1}{4}}\right)$$

for $z$ in a sector around $\frac{3}{16}$ and thereby

$$(18) \qquad H_{2,n}^{(2)} = O\left(\left(\frac{16}{3}\right)^n n^{-\frac{5}{4}}\right) \qquad (n \to \infty),$$

which is a sharper version of (14) and completes the proof of the theorem.

## 2. ASYMPTOTIC BEHAVIOUR OF THE NUMBERS $H_{2,n}^{(1)}$

By relation (11) the radius of convergence of $H_2^{(1)}(z)$ is the same as of the function $y_2(z)$, that is $\rho = \frac{3}{16}$.

Similar to [3] we define the generating functions $\nu^{[h]}$ of the numbers of trees $t \in B_{2,n}$ with $h_1(t) \leqslant h$. The functions $\nu^{[h]}$ fulfill the recursion (compare (8))

$$(19) \qquad \nu^{[0]} = 0, \qquad \nu^{[h+1]} = y_1 + z(\nu^{[h]})^2.$$

An immediate consequence is the identity

$$(20) \qquad H_2^{(1)}(z) = \sum_{h \geqslant 1} h(\nu^{[h]} - \nu^{[h-1]}) = \sum_{h \geqslant 0} (y_2 - \nu^{[h]}).$$

By setting

$$(21) \qquad f_h(z) = \frac{y_2(z) - \nu^{[h]}(z)}{2y_2(z)}$$

we have

$$(22) \qquad H_2^{(1)}(z) = 2y_2(z) \sum_{h\geq 0} f_h(z)$$

and recursion (19) is transformed into

$$(23) \qquad \begin{aligned} f_0(z) &= \frac{1}{2} \\ f_{h+1}(z) &= (1 - \varphi(z)) f_h(z)(1 - f_h(z)) \end{aligned}$$

with $\varphi(z) = (-1 + 2(1-4z)^{\frac{1}{2}})^{\frac{1}{2}} = 1 - 2zy_2(z)$. After having taken the complex plane cut along the axis $w < 0$, $w^{\frac{1}{2}}$ is to be positive for $w > 0$ throughout the whole paper!

$\varphi(z)$ has its (algebraic) singularity nearest to the origin at $\frac{3}{16}$, where $2(1-4z)^{\frac{1}{2}} = 1$, and no further singularities inside $|z| = \frac{1}{4}$. This situation is quite the same as in [3], where the recursion

$$(24) \qquad \begin{aligned} e_0(z) &= \frac{1}{2} \\ e_{h+1}(z) &= (1 - \epsilon(z)) e_h(z)\,(1 - e_h(z)) \end{aligned}$$

with $\epsilon(z) = (1-4z)^{\frac{1}{2}}$ is studied and the local behaviour of $2y_1(z) \sum_{h\geq 0} e_h(z)$ in a sector around the (logarithmic) singularity $\frac{1}{4}$ turns out to be

$$(25) \qquad 2y_1(z) \sum_{h\geq 0} e_h(z) = -2 \log(1-4z) + K + O(|1-4z|^{\nu})$$

for any $\nu < \frac{1}{4}$. Quite the same methods lead to the following result

$$(26) \qquad 2y_2(z) \sum_{h\geq 0} f_h(z) = -\frac{8}{3} \log\left(1 - \frac{16z}{3}\right) + K_1 + O\left(\left|1 - \frac{16z}{3}\right|^{\nu}\right)$$

for any $\nu < \frac{1}{4}$, and integration using the contour $\Gamma$ mentioned in the introduction with $q = \frac{3}{16}$ yields the result

$$(27)\qquad H^{(1)}_{2,n}=\frac{8}{3}\left(\frac{16}{3}\right)^n n^{-1}+O\left(\left(\frac{16}{3}\right)^n n^{-1-\nu}\right)\qquad (n\to\infty)$$

for any $\nu<\frac{1}{4}$.

### 3. ASYMPTOTIC BEHAVIOUR OF THE NUMBERS $H^{(2)}_{2,n}$

We introduce the generating function $w^{[h]}$ of the number of trees $t\in B_{2,n}$ with $h_2(t)\leqslant h$. If $B^{[h]}$ denotes, as in the paper [3], the generating function of the number of trees $t\in B_{1,n}$ with $h(t)\leqslant h$, then the following relation holds (compare (9)):

$$(28)\qquad w^{[h]}=B^{[h]}+z(w^{[h]})^2.$$

As in (20) we get

$$(29)\qquad \begin{aligned} H^{(2)}_2(z)&=\sum_{h\geqslant 1}h(w^{[h]}-w^{[h-1]})=\sum_{h\geqslant 0}(y_2-w^{[h]})=\\ &=\frac{1}{2z}\sum_{h\geqslant 0}t^{[h]}(z),\end{aligned}$$

where $t^{[h]}(z)=2z(y_2(z)-w^{[h]}(z))$.

By subtracting the equalities $y_2=y_1+zy_2^2$ and (28) the following equation is obtained:

$$(30)\qquad (t^{[h]}(z))^2+2\varphi(z)t^{[h]}(z)-4z(y_1(z)-B^{[h]}(z))=0$$

and $\left\{\dfrac{y_1(z)-B^{[h]}(z)}{2y_1(z)}\right\}$ is just the function system $\{e_h(z)\}$. (Check the recursion (24) above.) Hence the following formal identity

$$(31)\qquad \begin{aligned} H^{(2)}_2(z)&=\frac{1}{2z}\sum_{h\geqslant 0}t^{[h]}(z)=\\ &=\frac{1}{2z}\sum_{h\geqslant 0}\left(-\varphi(z)+(\varphi^2(z)+8zy_1(z)e_h(z))^{\frac{1}{2}}\right)\end{aligned}$$

(the determination of the square root being taken as declared above).

The aim of the following technical lemmas is to show that (31) leads to an analytical continuation of the function $H^{(2)}_2(z)$ (which has again radius of convergence $\frac{3}{16}$ by (11)) in a sector around $\frac{3}{16}$.

The first lemma points out that the function

$$(\varphi^2(z) + 8zy_1(z)e_h(z))^{\frac{1}{2}}$$

has no singularities in an appropriate domain. (Remark that all functions beyond the square root are analytic in $|z| < \frac{1}{4}$.)

**Lemma 3.1.** [*Localization of the zeroes of the functions* $\varphi^2(z) + 8zy_1(z)e_h(z)$]. *There exists an* $\eta > 0$ *such that* $\varphi^2(z) + 8zy_1(z)e_h(z) \neq 0$ *for all* $h \geqslant 0$, $|z| < \frac{3}{16} + \eta$, $z - \frac{3}{16} \notin \mathbf{R}^+$.

**Proof.** The equation $\varphi^2(z) + 8zy_1(z)e_h(z) = 0$ is equivalent with $B^{[h]}(z) = \frac{1}{4z}$.

There exists a sequence $(x_h)$ of real solutions with $x_h \downarrow \frac{3}{16}$. (This is an immediate consequence of the definition of $B^{[h]}(z)$ and their monotonic convergence to $y_1(z)$.)

The functions $1 - 4zB^{[h]}(z)$ are polynomials and so it is possible to find an $\eta > 0$ such that there are no zeroes of all those polynomials on the circle $|z| = \frac{3}{16} + \eta$. The sequence $(1 - 4zB^{[h]}(z))_{h \geqslant 0}$ converges uniformly on all compact subsets of $|z| < \frac{1}{4}$ to $1 - 4zy_1(z)$, which has a simple zero at $\frac{3}{16}$ and no other zeroes in this region.

So by a theorem of Hurwitz (compare [2], p. 148) there exists $N$ such that $1 - 4zB^{[h]}(z)$ has exactly one zero in $|z| < \frac{3}{16} + \eta$ for all $h \geqslant N$. If $N$ is large enough, this zero must be $x_h$ from above, and so there is no further zero inside $|z| < \frac{3}{16} + \eta$.

Any zero $z_h$ of $1 - 4zB^{[h]}(z)$ must fulfill

$$\frac{1}{4|z_h|} = |B^{[h]}(z_h)| \leqslant B^{[h]}(|z_h|)$$

(as $B^{[h]}$ has positive coefficients), and it is immediately seen that this inequality can only be valid for $|z_h| \geqslant x_h$. So by taking $\eta$ small enough there are no zeroes of the functions $1 - 4zB^{[h]}(z)$ in the domain mentioned in the lemma. ▮

**Lemma 3.2.** [*Uniform geometric bound of the functions* $e_h(z)$ *defined by* (24)] $\exists \eta > 0$ *such that* $|e_h(z)| < p^h$ *for* $|z| < \frac{3}{16} + \eta$, $h \in \mathbf{N}$, *with* $p = 2^{-\frac{3}{4}}$.

**Proof.** First we note that for $|z| \leqslant x = 2^{-\frac{31}{16}}(1 - 2^{-\frac{31}{16}}) \approx$ $\approx 0{,}1929\ (> \frac{3}{16})$

$$|1 - \epsilon(z)| \leqslant 2^{-\frac{15}{16}}, \tag{32}$$

because $1 - \epsilon(z) = p_1 e^{it}$ implies $z = \frac{p_1}{2} e^{it} - \frac{p_1^2}{4} e^{2it}$ and therefore

$$|z| \geqslant \frac{p_1}{2} \left(1 - \frac{p_1}{2}\right)$$

which yields the desired bound for $p_1 = 2^{-\frac{15}{16}}$.

Now the assumption can be proved by induction: The estimation can be checked directly in the cases $h = 0, \ldots, 4$; the inequality

$$|e_{h+1}| \leqslant |1 - \epsilon|(1 + |e_h|)|e_h| \leqslant$$
$$\leqslant 2^{-\frac{15}{16}}(1 + p^h)p^h \leqslant 2^{-\frac{15}{16}}\left(1 + 2^{-\frac{15}{16}} \frac{1}{4}\right) p^h \leqslant pp^h$$

for all $h \geqslant 4$ gives the result. ▮

**Lemma 3.3.** [*Local convergence of* $e_h(z)$ *to* $e_h\left(\frac{3}{16}\right)$]. $\exists \eta, C > 0$ *such that*

$$\left| e_h(z) - e_h\left(\frac{3}{16}\right) \right| \leqslant C \left| z - \frac{3}{16} \right| p^h \quad \textit{for} \quad \left| z - \frac{3}{16} \right| < \eta,\ h \in \mathbf{N}_0.$$

**Proof.** Recursion (24) yields

$$e_{h+1}(z) - e_{h+1}\left(\frac{3}{16}\right) =$$

$$= \frac{1}{2}\left(e_h(z) - e_h\left(\frac{3}{16}\right)\right)\left(1 - \left(e_h(z) + e_h\left(\frac{3}{16}\right)\right)\right) +$$

$$+ \frac{\frac{1}{2} - \epsilon(z)}{1 - \epsilon(z)}\, e_{h+1}(z)$$

and so by Lemma 3.2 for $\left|z - \frac{3}{16}\right| < \eta$

$$\left|e_{h+1}(z) - e_{h+1}\left(\frac{3}{16}\right)\right| \leqslant$$

$$\leqslant \frac{1}{2}\left|e_h(z) - e_h\left(\frac{3}{16}\right)\right|(1 + 2p^h) + C_1\left|\frac{1}{2} - \epsilon(z)\right| p^{h+1}.$$

For $\eta$ small enough

$$\left|\frac{1}{2} - \epsilon(z)\right| < C_2\left|z - \frac{3}{16}\right|.$$

By taking $C$ such that $C_1 C_2 \leqslant C \cdot 0{,}03$ and

$$\left|e_h(z) - e_h\left(\frac{3}{16}\right)\right| \leqslant C\left|z - \frac{3}{16}\right| p^h \quad \text{for} \quad 0 \leqslant h \leqslant 5,$$

we get the desired result by induction, as the estimation for $e_h(z)$ yields

$$\left|e_{h+1}(z) - e_{h+1}\left(\frac{3}{16}\right)\right| \leqslant$$

$$\leqslant C\left|z - \frac{3}{16}\right| p^{h+1}\left(\frac{1}{2}p^{-1}(1 + 2p^h) + \frac{C_1 C_2}{C}\right) \leqslant$$

$$\leqslant C\left|z - \frac{3}{16}\right| p^{h+1}(0{,}966 + 0{,}03) \quad \text{for all} \quad h \geqslant 5. \blacksquare$$

**Lemma 3.4.** $\left|\operatorname{Argt}\left(z - \frac{3}{16}\right)\right| \geqslant \frac{\pi}{3} \Rightarrow |\operatorname{Argt} \varphi^2(z)| \leqslant \frac{2\pi}{3}.$

**Proof.** If $z = \frac{3}{16} + \rho e^{\frac{i\pi}{3}}$ $(\rho \geqslant 0)$ then

$$\varphi^2(z) = -1 + 2\sqrt{1 - 4z} = x + iy$$

with

$$(x+1)^2 - y^2 - \frac{2}{\sqrt{3}}(x+1)y = 1,$$

so that $\varphi^2(z)$ maps the half-ray on the part of the right branch of the hyperbola with $y \leqslant 0$. Because of $y'(0) = \sqrt{3} = \operatorname{tg}\frac{\pi}{3}$ we obtain the result. ∎

For brevity's sake let $E_h(z) = 8zy_1(z)e_h(z)$.

**Lemma 3.5.** $\left|\varphi^2(z) + E_h\left(\frac{3}{16}\right)\right|^2 \geqslant \frac{1}{2}\left(E_h\left(\frac{3}{16}\right)^2 + |\varphi^2(z)|^2\right)$ for $\left|\operatorname{Argt}\left(z - \frac{3}{16}\right)\right| \geqslant \frac{\pi}{3}$.

**Proof.** As $E_h\left(\frac{3}{16}\right) \in \mathbf{R}^+$ we get by the cosine theorem

$$\left|\varphi^2(z) + E_h\left(\frac{3}{16}\right)\right|^2 =$$

$$= E_h\left(\frac{3}{16}\right)^2 + |\varphi^2(z)|^2 - 2E_h\left(\frac{3}{16}\right)|\varphi^2(z)| \cos\alpha,$$

with $\alpha \geqslant \frac{\pi}{3}$ by Lemma 3.4. $\cos\alpha < \frac{1}{2}$ yields the result. ∎

**Lemma 3.6.** $\exists \eta,\ C > 0$ *such that*

$$\left|\frac{E_h(z) - E_h\left(\frac{3}{16}\right)}{\sqrt{\varphi^2(z) + E_h\left(\frac{3}{16}\right)}}\right| \leqslant C\left|z - \frac{3}{16}\right|^{\frac{1}{2}} p^h$$

*for* $\left|z - \frac{3}{16}\right| < \eta$, $\left|\operatorname{Argt}\left(z - \frac{3}{16}\right)\right| \geqslant \frac{\pi}{3}$ *and* $h \in \mathbf{N}_0$.

**Proof.** According to Lemma 3.4, $\varphi^2(z) + E_h\left(\frac{3}{16}\right) \neq 0$ and for $\eta$ small enough

$$\left|\frac{E_h(z) - E_h\left(\frac{3}{16}\right)}{\sqrt{\varphi^2(z) + E_h\left(\frac{3}{16}\right)}}\right| \leqslant C_1 \frac{\left|z - \frac{3}{16}\right| p^h}{\left(|\varphi^2(z)|^2 + E_h^2\left(\frac{3}{16}\right)\right)^{\frac{1}{4}}} \leqslant$$

$$\leqslant C_1 \frac{\left|z-\frac{3}{16}\right| p^h}{|\varphi^2(z)|^{\frac{1}{2}}} \leqslant C\left|z-\frac{3}{16}\right|^{\frac{1}{2}} p^h$$

by Lemma 3.4 and 3.5 and

$$|\varphi^2(z)|^{\frac{1}{2}} = 2\sqrt{2}\left|z-\frac{3}{16}\right|^{\frac{1}{2}} + O\left(\left|z-\frac{3}{16}\right|\right) \qquad \left(z \to \frac{3}{16}\right). ∎$$

**Lemma 3.7.** $\exists \eta,\ C > 0,\ N \in \mathbf{N}_0$ *such that*

$$\left|\sqrt{\varphi^2(z)+E_h(z)} - \sqrt{\varphi^2(z)+E_h\left(\frac{3}{16}\right)}\right| \leqslant C\left|z-\frac{3}{16}\right|^{\frac{1}{2}} p^h$$

*for* $\left|z-\frac{3}{16}\right| < \eta,\ \left|\operatorname{Argt}\left(z-\frac{3}{16}\right)\right| \geqslant \frac{\pi}{3},\ h \geqslant N.$

**Proof.** We have

$$\sqrt{\varphi^2(z)+E_h(z)} - \sqrt{\varphi^2(z)+E_h\left(\frac{3}{16}\right)} = \sqrt{\sigma_h(1+\tau_h)} - \sqrt{\sigma_h}$$

with

$$\sigma_h := \varphi^2(z) + E_h\left(\frac{3}{16}\right), \qquad \tau_h := \frac{E_h(z) - E_h\left(\frac{3}{16}\right)}{\varphi^2(z) + E_h\left(\frac{3}{16}\right)}.$$

By Lemma 3.3 and 3.5 with suited $\eta$:

$$|\tau_h| \leqslant C_1 \frac{\left|z-\frac{3}{16}\right| p^h}{|\varphi^2(z)|} \leqslant C_2 p^h \quad \text{for all } h \in \mathbf{N}.$$

By taking $N$ large enough $C_2 p^h < \frac{1}{2}$ for all $h \geqslant N$ and therefore $|\operatorname{Argt}(1+\tau_h)| < \frac{\pi}{4}$. Furthermore

$$|\operatorname{Argt} \sigma_h| < |\operatorname{Argt} \varphi^2(z)| \leqslant \frac{2\pi}{3}.$$

So we have $\sqrt{\sigma_h(1+\tau_h)} = \sqrt{\sigma_h}\sqrt{1+\tau_h}$ and we derive

$$|\sqrt{\sigma_h(1+\tau_h)} - \sqrt{\sigma_h}| =$$

$$= \sqrt{|\sigma_h|}\,|(1+\tau_h)^{\frac{1}{2}} - 1| \leqslant C_3 \sqrt{|\sigma_h|}\,|\tau_h| =$$

$$= C_3 \left| \frac{E_h(z) - E_h\left(\frac{3}{16}\right)}{\sqrt{\varphi^2(z) + E_h\left(\frac{3}{16}\right)}} \right| \leqslant C \left| z - \frac{3}{16} \right|^{\frac{1}{2}} p^h \qquad (h \geqslant N)$$

by Lemma 3.6.∎

**Lemma 3.8.** [*Local behaviour of the functions* $t^{[h]}(z)$ *in a sector around* $\frac{3}{16}$]. $\exists \eta,\ C > 0,\ N \in \mathbf{N}_0$ *such that*

$$\left| -\varphi(z) + \sqrt{\varphi^2(z) + E_h(z)} - \sqrt{E_h\left(\frac{3}{16}\right)} \right| \leqslant C \left| z - \frac{3}{16} \right|^{\frac{1}{4}} p^{\frac{h}{4}}$$

*for all* $\left| z - \frac{3}{16} \right| < \eta,\ \left| \text{Argt}\left(z - \frac{3}{16}\right) \right| \geqslant \frac{\pi}{3},\ h \geqslant N.$

**Proof.**

$$\left| -\varphi(z) + \sqrt{\varphi^2(z) + E_h(z)} - \sqrt{E_h\left(\frac{3}{16}\right)} \right| \leqslant$$

$$\leqslant \left| -\varphi(z) + \sqrt{\varphi^2(z) + E_h\left(\frac{3}{16}\right)} - \sqrt{E_h\left(\frac{3}{16}\right)} \right| +$$

$$+ \left| \sqrt{\varphi^2(z) + E_h(z)} - \sqrt{\varphi^2(z) + E_h\left(\frac{3}{16}\right)} \right| \leqslant$$

$$\leqslant \left| \frac{2\varphi(z)\sqrt{E_h\left(\frac{3}{16}\right)}}{\sqrt{\varphi^2(z) + E_h\left(\frac{3}{16}\right)} + \varphi(z) + \sqrt{E_h\left(\frac{3}{16}\right)}} \right| +$$

$$+ C_1 \left| z - \frac{3}{16} \right|^{\frac{1}{2}} p^h$$

by Lemma 3.7 for $h \geqslant N$. Now we have

$$\left| \sqrt{\varphi^2(z) + E_h\left(\frac{3}{16}\right)} + \varphi(z) + \sqrt{E_h\left(\frac{3}{16}\right)} \right|^2 \geqslant$$

$$\geqslant \left| \sqrt{\varphi^2(z) + E_h\left(\frac{3}{16}\right)} \right|^2 + \left| \varphi(z) + \sqrt{E_h\left(\frac{3}{16}\right)} \right|^2 \geqslant$$

$$\geqslant \left| \varphi(z) + \sqrt{E_h\left(\tfrac{3}{16}\right)} \right|^2 \geqslant |\varphi^2(z)| + E_h\left(\tfrac{3}{16}\right) \geqslant$$

$$\geqslant 2\sqrt{|\varphi^2(z)| E_h\left(\tfrac{3}{16}\right)},$$

again using the cosine theorem and regarding $E_h\left(\frac{3}{16}\right) \in \mathbf{R}^+$.

Putting everything together we get the estimation

$$\sqrt{2}\,|\varphi(z)|^{\frac{1}{2}} \left(E_h\left(\tfrac{3}{16}\right)\right)^{\frac{1}{4}} + C_1 \left| z - \tfrac{3}{16} \right|^{\frac{1}{2}} p^h \leqslant C \left| z - \tfrac{3}{16} \right|^{\frac{1}{4}} p^{\frac{h}{4}}. \blacksquare$$

**Lemma 3.9.** [*Local behaviour of the function* $H_2^{(2)}(z)$ *in a sector around* $\frac{3}{16}$]. $\exists \eta > 0$ *such that the function series*

$$\sum_{h \geqslant 0} t^{[h]}(z) = \sum_{h \geqslant 0} \left(-\varphi(z) + \sqrt{\varphi^2(z) + E_h(z)}\right)$$

*converges absolutely and uniformly in the sector* $\left| z - \frac{3}{16} \right| < \eta$, $\left| \operatorname{Argt}\left(z - \frac{3}{16}\right) \right| > \frac{\pi}{3}$, *and behaves like*

$$\sum_{h \geqslant 0} t^{[h]}(z) = \sum_{h \geqslant 0} \left(E_h\left(\tfrac{3}{16}\right)\right)^{\frac{1}{2}} + O\left(\left| z - \tfrac{3}{16} \right|^{\frac{1}{4}}\right).$$

*In particular the function series is an analytical continuation of the function* $2zH_2^{(2)}(z)$ *for* $z \neq \frac{3}{16}$ *in the sector mentioned above.*

**Proof.** We choose $\eta$, $N$ as in Lemma 3.7 and split the total series in the following way

$$\sum_{h \geqslant 0} t^{[h]}(z) = \sum_{h=0}^{N-1} t^{[h]}(z) + \sum_{h \geqslant N} t^{[h]}(z).$$

For $\eta$ small enough each of the functions $t^{[h]}(z)$ is analytic in the sector for $z \neq \frac{3}{16}$ by Lemma 3.1.

The function $\sum_{h=0}^{N-1} t^{[h]}(z)$ is analytical in a small disc around $\frac{3}{16}$ apart from the algebraic singularity $\frac{3}{16}$ coming from $\varphi(z)$ and therefore

$$\sum_{h=0}^{N-1} t^{[h]}(z) = \sum_{h=0}^{N-1} t^{[h]}\left(\frac{3}{16}\right) + O\left(\left|z - \frac{3}{16}\right|^{\frac{1}{2}}\right) =$$

$$= \sum_{h=0}^{N-1} \left(E_h\left(\frac{3}{16}\right)\right)^{\frac{1}{2}} + O\left(\left|z - \frac{3}{16}\right|^{\frac{1}{2}}\right).$$

Lemma 3.8 establishes the estimation of

$$\sum_{h \geq N} t^{[h]}(z) - \sum_{h \geq N} t^{[h]}\left(\frac{3}{16}\right)$$

as well as the absolute and uniform convergence.∎

**Lemma 3.10.** [*Local behaviour of the function* $H_2^{(2)}(z)$ *apart from* $\frac{3}{16}$]. *With* $\eta$ *from Lemma 3.9 for all sufficiently small* $\eta_1 > 0$ *the function series* $\sum_{h \geq 0} t^{[h]}(z)$ *converges absolutely and uniformly in the domain* $|z| < \frac{3}{16} + \eta_1$, $\left|z - \frac{3}{16}\right| \geq \frac{\eta}{2}$ *and is an analytical continuation of the function* $2zH_2^{(2)}(z)$.

**Proof.** For $\eta_1$ sufficiently small $t^{[h]}(z)$ is an analytical function in the considered domain by Lemma 3.1. Now we choose $N$ such that (in consequence of Lemma 3.2)

$$\left|\frac{E_h(z)}{\varphi^2(z)}\right| < \frac{1}{2} \quad \text{for all} \quad \left|z - \frac{3}{16}\right| \geq \frac{\eta}{2} \qquad (h \geq N).$$

Then we have $\left|\text{Argt}\left(1 + \frac{E_h(z)}{\varphi^2(z)}\right)\right| \geq \frac{\pi}{3}$ and therefore $|\text{Argt}\,\varphi^2(z)| \leq \frac{2\pi}{3}$ by Lemma 3.4. So

$$|t^{[h]}(z)| = \left|-\varphi(z) + \sqrt{\varphi^2(z)\left(1 + \frac{E_h(z)}{\varphi^2(z)}\right)}\right| =$$

$$= |\varphi(z)|\left|-1 + \left(1 + \frac{E_h(z)}{\varphi^2(z)}\right)^{\frac{1}{2}}\right| \leq$$

$$\leq C_1\,|\varphi(z)|\left|\frac{E_h(z)}{\varphi^2(z)}\right| = C_1\left|\frac{E_h(z)}{\varphi^2(z)}\right| \leq Cp^h,$$

where $C_1, C$ can be taken independently from $h$ and $z$.∎

**Theorem 3.11.** $H_{2,n}^{(2)} = O\left(\left(\frac{16}{3}\right)^n n^{-\frac{5}{4}}\right) \quad (n \to \infty).$

**Proof.** By Lemma 3.9 and 3.10 $\sum_{h>0} t^{[h]}(z)$ is an analytical continuation of the function $2zH_2^{(2)}(z)$ in the domain

$$|z| < \frac{3}{16} + \eta_1, \quad \left|\operatorname{Argt}\left(z - \frac{3}{16}\right)\right| > \frac{\pi}{3}, \quad z \neq \frac{3}{16}.$$

As announced in the introduction we evaluate the coefficients by means of Cauchy's integral formula choosing the contour of integration $\Gamma$ in the following way: $\Gamma = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$ with

$$\Gamma_1 = \left\{z \,\middle|\, |z| = \frac{3}{16} + \frac{\eta_1}{2}, \ \left|\operatorname{Argt}\left(z - \frac{3}{16}\right)\right| \geq \frac{5\pi}{6}\right\}$$

$$\Gamma_{2,3} = \left\{z \,\middle|\, z = \frac{3}{16} + \lambda \exp\left(\pm i\frac{5\pi}{6}\right), \ \frac{3}{16} \leq |z| \leq \frac{3}{16} + \frac{\eta_1}{2}\right\}.$$

(To be quite rigorous we should take a part of a small circle around $\frac{3}{16}$ connecting the two line segments and let shrink the radius of this circle to 0, but this causes no troubles.)

The contribution of the integral along $\Gamma_1$ is exponentially small compared with $\left(\frac{16}{3}\right)^n$; the line segments $\Gamma_2, \Gamma_3$ can be trated symmetrically and an easy estimation (compare e.g. [4]) yields

$$\int_{\Gamma_2} \left|1 - \frac{16z}{3}\right|^{\frac{1}{4}} \frac{dz}{|z|^n} = O\left(\left(\frac{16}{3}\right)^n n^{-\frac{5}{4}}\right).$$

So the proof of the theorem is complete.∎

## 4. CONCLUDING REMARKS

We would like to emphasize that the present approach to the evaluation of the average height of binary trees gives more information than announced in the main theorem, as the considerations on $h_1$ and $h_2$ may be regarded of some special interest for themselves.

A natural extension of the problems treated in this paper is the case of labels taken from $\{1, \ldots, k\}$ (the numbers of the corresponding trees have been evaluated in the paper [7]). We feel that the methods of this paper may be well suited for an investigation of this general case, too.

**Acknowledgement**. The authors wish to express their thanks to G. Baron, Ph. Flajolet and R. Schnabl for helpful suggestions and remarks.

## REFERENCES

[1] N.G. de Bruijn, *Asymptotic methods in analysis,* North-Holland Publ. Comp., Amsterdam, 1961.

[2] J.B. Conway, *Functions of one complex variable,* Springer-Verlag, New York – Heidelberg – Wien, 1975.

[3] Ph. Flajolet – A. Odlyzko, Exploring binary trees and other simple trees, In: *Proc. of the 21st Annual Symposium on Foundations of Computer Science,* Syracuse, 1980, 207–216.

[4] Ph. Flajolet – A. Odlyzko, The average height of binary trees and other simple trees, *J. Comput. Syst. Sci.,* 25 (1982), 171–213.

[5] D.E. Knuth, *The art of computer programming,* Addison Wesley, Reading, 1968.

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

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

P. Kirschenhofer – H. Prodinger

Institut f. Algebra und Diskrete Mathematik der Technischen Universität Wien, Gusshausstr. 27–29, A-1040 Vienna, Austria.